Periostin Splice Variants in Pancreatic Cancer

By

Eleni Georgiadi

First supervisor: Dr Terry Gaymes

Other supervisors: Dr Karen Whiting and Dr Natasha Hill

A thesis submitted in partial fulfilment of the requirements of Kingston University for the award of Masters by Research

Faculty of Science, Engineering and Computing

June 2022

DECLARATION

This thesis is titled “Periostin Splice Variants in Pancreatic Cancer” has been submitted for the degree of Masters by Research and has not been submitted for another degree at any institutions. This thesis contains the candidate’s original work and the contributions of work made by others have been acknowledged in text and permission was granted from authors for certain figures used in this thesis and the work of others has been referenced.

ACKNOWLEDGEMENTS

I would like to thank Dr Terry Gaymes, Dr Karen Whiting and Dr Natasha Hill for their extensive guidance, support, and expertise throughout my research. I also like to express my gratitude to Dr Francesca Mackenzie and Mahduri Sawant for their assistance with the cloning work and guiding me during the trouble-shooting process. I would like to thank Musamma Jamila Rahim whose work helped form part of the foundation for this project. I am grateful for the training and support provided by Amanda Dandagama Mudiyanselage who was always there for me.

I want to extend my gratitude to my family, my mother, Aggeliki and my sisters, Maria and Lazaria for always being supportive and encouraging to pursue this project.

ABSTRACT

Introduction: Pancreatic ductal adenocarcinoma (PDAC) is one of the deadliest solid tumours, with 5-year survival of less than 5%. The aggressive nature of the tumour, the tumour-supportive environment (stroma), delayed diagnosis and poor chemotherapy response lead to poor patient prognosis. The tumour-stroma is a dynamic network of immune, mesenchymal, and endothelial stromal cells surrounded by an extracellular matrix rich in matricellular and architectural proteins. Periostin (POSTN), a matricellular protein primarily secreted by stromal cells, is often associated with cell proliferation and migration. POSTN overexpression has been observed in many tumour types, however downregulation in bladder carcinoma rises some questions. *POSTN* is an alternatively sliced gene, that encodes multiple different isoforms, whose function in PDAC is still unknown. The aim of this study is to isolate and clone these isoforms in order to produce recombinant proteins for further investigation.

Methods: The transcript variants were extracted from PS1 cells (a pancreatic stromal cell line), converted to cDNA and amplified by RT-PCR. The synthesised DNA was ligated with plasmid vectors and cloned by traditional cloning into *E.coli* cells. Single colonies were sequenced and analyses, determining the transcript variant being cloned. The plasmids ligated with the variants of interest were isolated and transfected into PS1 cells. The recombinant proteins synthesised by the PS1 cells were then purified.

Results: Multiple isoforms of *POSTN* were cloned in *E.coli*. Sequencing analysis revealed the exons present in each sample of the cloned colonies and the isoforms were identified. The plasmids of interest were successfully extracted from the colonies by using purification technique appropriate for subsequent mammalian cell transfection. Substantial progress was made in producing recombinant POSTN proteins.

Conclusion: In this study the isolation of three cloned *POSTN* transcript variants were achieved. These variants are POSTN-202, 209 and 201-b.

Key words: PDAC, POSTN/periostin, tumour stroma, PS1 cells, cloning

TABLE OF CONTENTS

LIST OF FIGURES

7 Appendix **Page**

LIST OF TABLES

LIST OF ABBREVIATIONS

APS	Ammonium persulfate
AS	Alternative splicing
ATM	A-T mutated
bp	Base pairs
BPS	Branch point sequence
BRCA-1/-2	Breast cancer gene-1/-2
CA	Carcinoembryonic antigen
CAFS	Carcinoma-associated fibroblasts
CD8	Cluster of differentiation 8
CDH1	Cadherin1
CDKN2A	Cyclin-dependent kinase inhibitor 2A
cDNA	Complementary DNA
CFTR	Cystic fibrosis transmembrane conductance regulator
CHO	Chinese hamster ovary
CO_2	Carbon dioxide
CT	Computed tomography
CTGF	Connective tissue growth factor
dH_2O	Distilled water
DMSO	Dimethylsulfoxide
DNA	Deoxyribonucleic acid
dNTPs	Deoxynucleotide triphosphates
E. coli	Escherichia coli
ECM	Extracellular matrix
EDTA	Ethylenediamine tetraacetic acid
EGFR	Epidermal growth factor receptor
ERK	Extracellular signal-regulated kinase
ESE	Exonic splicing enhancer
FAK	Focal adhesion kinase
FAS1	Fasciclin 1
FBS	Foetal bovine serum
FBXW7	F-box and WD repeat domain containing 7
FGF	Fibroblast growth factor
FOX-1/-2	Forkhead box
HCL	Hydrochloric Acid
His-tag	Polyhistidine tag
hnRNP	Heterogeneous nuclear ribonucleoprotein
hPPARgamma	Human peroxisome proliferator-activated receptor gamma
ICI	Immune checkpoint inhibitor
IDH1	Isocitrate Dehydrogenase (NADP(+)) 1
IDR	intrinsic disorder region
IL-1/-6	Interleukin-1/-6
INK4A	Inhibitors of CDK4
IPMN	intraductal papillary mucinous neoplasms
IPTG	Isopropyl-β-D-thiogalactopyranoside
ISE	Intronic splicing enhancers
ISS	Intronic splicing silencer
JNK	Jun N-terminal Kinase

kDa	Kilodalton
KDR/flk-1	Kinase insert domain receptor-containing receptor/fetal liver kinase-**1**
KRAS	Kirsten rat sarcoma virus
LB	Lysogeny broth
LCR	Low complexity region
MAP	Mitogen-activated protein
MCN	Mucinous cystic neoplasms
MCS	Multiple cloning site
MLH1	MutL homolog-1/-2
MMP	Matrix metalloproteinases
MMR	Mismatch repair
MRI	Magnetic resonance imaging
mRNA	Messenger RNA
MSH1	MutS homolog 2
NaCl	Sodium chloride
NF1	Neurofibromatosis type 1
NRT	No reverse transcriptase
nt	Nucelotide
NTC	No template control
OD	Optical density
PALB2	Partner and localizer of BRCA2
PanIN	Pancreatic intraepithelial neoplasm
PBS	Phosphate buffered saline
PC	Pancreatic cancer
PCR	Polymerase chain reaction
PCR	Polymerase chain reaction
PD1	Programmed death 1
PDAC	Pancreatic ductal adenocarcinoma
PDGF	Platelet derived growth factor
PDL1	Programmed death 1 ligand 1
PI3K	Phosphoinositide 3-kinase
PKB	Protein kinase B
PMS2	Postmeiotic Segregation Increased, S. Cerevisiae, 2
POSTN	Periostin
PS1	Pancreatic stellate cells
PSCs	Pancreatic stellate cells
QARS	glutaminyl-tRNA synthetase
RIPA	Radio immune precipitation assay
RNA	Ribonucleic acid
RPM	Rotations per minute
RPMI-1640	Roswell park memorial institute medium
RT	Reverse transcriptase
RT	Room temperature
SABR	Stereotactic ablative radiotherapy
SDGF	Smooth muscle cell derived growth factor
SDS	Sodium dodecyl sulfate
SDS-PAGE	Sodium dodecyl sulfate–polyacrylamide gel electrophoresis
SF1	Splicing factor 1
SMAD4/DPC4	Mothers against decapentaplegic homolog 4/Deleted in pancreatic cancer 4
SMARCA4	SWI/SNF Related, Matrix Associated, Actin Dependent Regulator Of Chromatin, Subfamily A, Member 4

snRNPs	Small nuclear ribonucleoprotein particle
SPARC	Secreted protein acidic and rich in cysteine
SR	Serine/arginine-rich proteins
SS	Splice site
STK11	Serine/threonine kinase 11
TAE	Tris-acetate-EDTA buffer
TAM	Tumour associated macrophages
TEMED	Tetramethylethylenediamine
TGF-β	Transforming growth factor beta
TNF-α	Tumour necrosis factor alpha
TP53	Tumor protein 53
TTBS	Tween20 – Tris buffered saline
U2AF	U2 auxiliary factor
UV	Ultraviolet light
VEGF	Vascular endothelial growth factor
WHO	World health organisation
Wnt	Wingless/Integrated
WT	Wild type

1 INTRODUCTION

1.1 PANCREATIC CANCER

1.1.1 Epidemiology

Cancer has afflicted various multicellular organisms for millions of years, however only in the last couple of decades cancer has been progressed to one of the most dreaded diseases. Changes in lifestyle, modern habits and increased life expectancy are only a few of the factors that have alarmingly increased the incidence of the disease (Roy and Saikia, 2016; Hausman, 2019).

Being the second deadliest disease, after heart disease, cancer increases human mortality rate worldwide. According to Word Health Organisation (WHO), in 2020 the deaths caused by cancer were as high as 9,958,133 with new incidences estimated at approximately 20 million per year.

Developed continents including Europe and North America have incidence rates close to 257 per 100,000 individuals. Developing countries with lower net income such as South America, Asia and South Africa are not far behind, with incidence rates of 188.8-257.1 per 100,000 people (Fig.1.1).

Despite being ranked as the 15th most common cancer, pancreatic cancer remains one of the deadliest cancers with estimated 5-years survival rate of less than 10% (Ilic and Ilic, 2016). The morbidities and mortalities from PDAC in 2020 were so close in number that only 29,770 out of almost half a million patients survived. Pancreatic cancer is initially asymptomatic despite being an aggressive form of cancer, and its aetiology remains unclear. Numerous studies have attempted to determine the risk factors that contribute to the development of the disease (Loveday, Lipton and Thomson, 2019; McGuigan *et al.,* 2018; Goral, 2015).

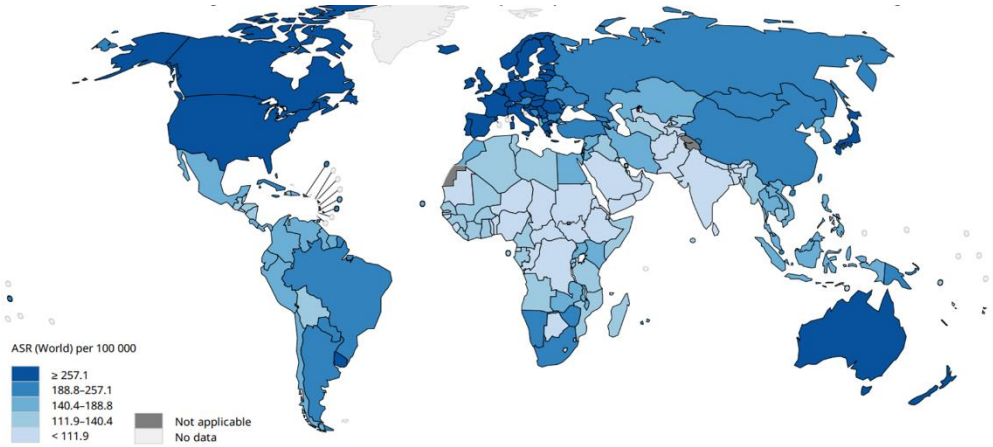


Figure 1.1: Map indicating the incidences of cancer worldwide. Countries with high incidence rates are indicated in dark blue and countries with lower incidence are indicated in shades of light blue colours. By continent, North America, Europe and Australia have the highest incidence rates, whereas South America, Africa and Asia and have lower rates. Image acquired: (*Word Health Organization.* 2022)

1.1.2 Risk factors

Risk factors can be divided into two main groups: non-modifiable and modifiable.

Non-modifiable risk factors include age, sex, ethnicity, genetic susceptibility/family history and other aspects that cannot be influenced by human habits. Pancreatic cancer predominately affects the elderly, with 90% of all patients being over 55 years of age. There is a slightly higher incidence in males over females possibly due to hormonal differences and habits. Genetics and lifestyle choices are the two factors that are also considered to explain the higher levels of incidences within distinct ethnicity groups such as in African Americans compared to Caucasian. Familial pancreatic cancer accounts for 5-10% of new cases and the familial risk factors elevate the risk of occurrence at least nine times. (Loveday, Lipton and Thomson, 2019; McGuigan *et al.,* 2018). Pancreatic cancer (PC) is more prevalent in individuals with certain gene mutations and associated hereditary syndromes. Hereditary pancreatitis with Serine Protease 1 (*PRSS1*) mutation was found to increase the risk of developing PC more than 10-fold. A similar increase in risk is found in patients with mutation in the Breast cancer genes 1 and 2 (*BRCA1* and *BRCA2*). Peutz-Jeghers syndrome is an autosomal dominant hereditary disease caused by a mutation in of the Serine/threonine kinase 11 (*STK11*) gene and the risk of developing PC in patients with this syndrome is 132-fold higher compared to general population. Other mutations that elevate the risk of PC are in the Cyclin-dependent kinase inhibitor 2A (*CDKN2A/p16)* gene associated with familial atypical mole and multiple melanomas, the Cystic fibrosis transmembrane conductance regulator (*CFTR*) gene and the mismatch repair genes, MutL homolog 1, 2, 6 (*MLH1*, *MSH2*, *MSH6)* and Postmeiotic Segregation Increased, S. Cerevisiae 2 (*PMS2*). This list was recently extended to include mutations in the ataxia telangiectasia mutated (*ATM*) and Partner and localizer of BRCA2 (*PALB2*) genes, with the last one also been associated with breast cancer (Capurso *et al.,* 2015; Del Chiaro *et al.,* 2014; Chen, F., Roberts and Klein, 2017).

Modifiable risk factors refer to lifestyle choices that can negatively impact the genome. Cigarette smoking has been assessed as the most important modifiable risk factor. A meta-analysis of 82 studies showed that a smoker has 74% increased risk of been diagnosed with pancreatic cancer over a non-smoker (Loveday, Lipton and Thomson, 2019; McGuigan *et al.,* 2018). Other risk factors include alcohol consumption, obesity and dietary choices (Ilic and Ilic, 2016; Loveday, Lipton and Thomson, 2019).

1.1.3 Symptoms

Early diagnosis of pancreatic cancer is hindered by many obstacles. The symptoms are insidious at onset and even at examination they are non-specific and overlap with other diseases (McGuigan

et al., 2018). Common symptoms of PC include diarrhoea (specifically steatorrhea) or constipation, malaise, nausea and vomiting (Loveday, Lipton and Thomson, 2019). Upper abdominal pain that radiates to the sides or back and is alleviated when sitting still or bending forward is a typical finding in 79% of the cases (Gupta, Amanam and Chung, 2017; Goral, 2015). Unexplained weight loss (in 85% of the cases) in combination with the symptoms above may suggest the presence of the disease (Gupta, Amanam and Chung, 2017). New-onset and persistent jaundice is characteristic in patients with cancer at the head of the pancreas due to the physical obstruction of the bile duct. This clinical presentation alerts the need of further investigation by specialist (Fig.1.2). However, at the time of diagnosis, in more than 80% of cases the cancer has progress to an advanced stage where the prognosis is very poor (Hawes *et al.,* 2000; McGuigan *et al.,* 2018).

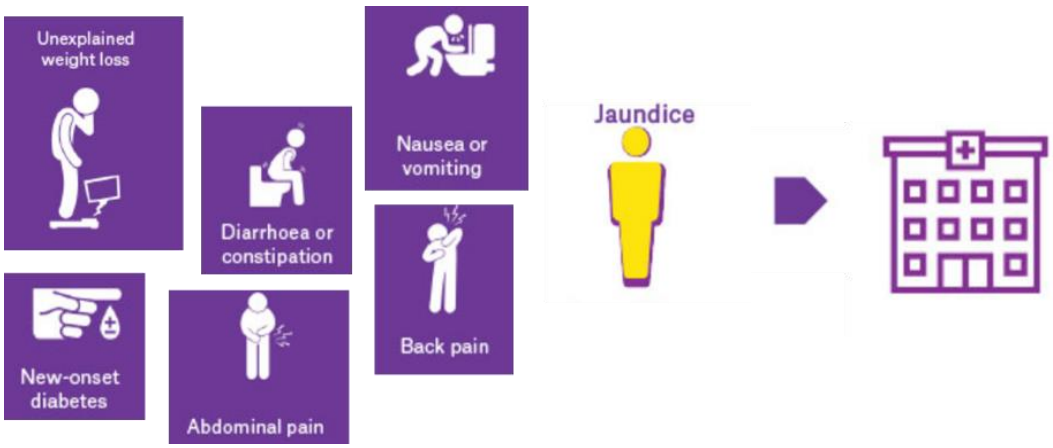


Figure 1.2: Most profound symptoms of patients with pancreatic cancer. Unexplained weight loss, diarrhoea or constipation, nausea or vomiting, new-onset diabetes, abdominal and back pain, and jaundice are the symptoms that usually lead a patient to the doctor for further investigation. Image adapted from: (Loveday, Lipton and Thomson, 2019).

1.1.4 Diagnosis and treatment

There are many different approaches utilised for the diagnosis of pancreatic cancer. Screening techniques such as computed tomography (CT) and magnetic resonance imaging (MRI) scans are routinely used for the detection, location, and metastases of pancreatic tumours (Canto *et al.,* 2013). Biopsy and endoscopic approaches are also performed for determining invasion and staging of the tumour (Goral, 2015). Tumour markers including carcinoembryonic antigen (CA) 19-9, CA 72-4, CA 50, CA 242, fibrinogen gamma and sialylated plasma protease C1 are promising tools for diagnosis. CA 19-9 is still the only biomarker approved by the United States Food and Drug Administration for routine monitoring of pancreatic cancer, but it can only be detected in blood in the later stages of the disease (McGuigan *et al.,* 2018). Early diagnosis of pancreatic

cancer is challenging but finding novel markers remains vital as this will pave the way for better clinical outcomes (Zaimy *et al.,* 2017; Canto *et al.,* 2013).

Since the first anatomical resection of a pancreatic tumour by Friedrich Trendelenberg in 1882, surgery remains the only treatment offering a potential cure for pancreatic cancer. There are three main pancreatic resections, all determined by the anatomical location of the tumour. For tumours found on the head or body of pancreas, pancreaticoduodenectomy is usually performed, whereas distal pancreatectomy is performed for tumours in the tail of the pancreas (Torphy, Fujiwara and Schulick, 2020). Surgery is routinely followed by chemotherapy with drugs including FOLFIRINOX and Gemcitabine (Loveday, Lipton and Thomson, 2019). New approaches administer chemotherapy prior to operation (Neoadjuvant chemotherapy) to down stage the tumour and offer a better surgical outcome (Heinrich and Lang, 2017). Radiotherapy and ablative therapies involving the infusion of effector immune cells are not routinely used however they have shown encouraging results for future use (Loveday, Lipton and Thomson, 2019).

Targeting of the programmed immune checkpoint inhibitor (ICI) pathway molecule, programmed death 1 or programmed death - ligand 1 (PD-1/PD-L1), is a novel therapy for many cancer types including PDAC (Saka *et al.,* 2020; Henriksen *et al.,* 2019). PD-1 are co-inhibitory receptors expressed on the surface of tumour-infiltrating T-cells. Tumour cells selectively express its ligand PD-L1, aiming to create a negative feedback mechanism that favours immune suppression against the tumour. Binding of PD-1 with its ligand leads to T-cell anergy or death (known as exhaustion) and reduces the number of activated immune cells. By blocking this interaction, T-cell exhaustion can be reversed, and anti-tumour activities reinforced. Despite its great potential this therapy has limited success in clinical trials with PDAC mainly due to the naturally weak anti-tumour T cell immune response (Saka *et al.,* 2020; Yang *et al.,* 2020; Henriksen *et al.,* 2019). Stereotactic ablative radiotherapy (SABR) is a novel highly focused radiation treatment that delivers an intense dose of radiation to the target tumour with high accuracy, while limiting damage to the surrounding tissues. The intense radiation releases tumour antigens and enhances anti-tumour immunity. SABR combined with ICI immunotherapy has shown promising results in the treatment of various cancer types, including breast and renal carcinomas (Stewart *et al.,* 2021; All, Garant and Hannan, 2021). SABR could be used in the future for strengthening the anti-tumour immune response of T-cells in patients with PDAC and enhance the performance of ICI therapies (Li and Shen, 2020; Palma *et al.,* 2020).

1.1.5 From healthy pancreas to PDAC

The pancreas is primarily an exocrine gland, producing various digestive enzymes, but also has essential endocrine function. The endocrine pancreas constitutes of pancreatic islets which are clusters of cells that secrete hormones to regulate glucose metabolism. Cancer can arise from different cell types in any part of the pancreas, however in most of the cases (85-95%) cancer affects the epithelial lining of the pancreatic ducts, giving rise to exocrine pancreatic ductal adenocarcinoma (PDAC) (Loveday, Lipton and Thomson, 2019). Pancreatic endocrine tumours compose the second most common form, seen in less than 5% of the total cases (Gupta, Amanam and Chung, 2017; Vincent *et al.,* 2011).

PDAC evolves through neoplastic precursor lesions (Gupta, Amanam and Chung, 2017; Vincent *et al.,* 2011). Pancreatic intraepithelial neoplasm (PanIN) is the commonest form of lesions, followed by intraductal papillary mucinous neoplasms (IPMN) and mucinous cystic neoplasms (MCN). PanIN are non-invasive microscopic lesions (<5cm), usually located at the small pancreatic duct in the head of the pancreas (Goral, 2015). PanINs are divided into three subtypes (PanIN-1, PanIN-2 and-PanIN 3) which indicate the genetic, epigenetic and morphological changes leading to carcinoma *in situ* and malignant tumours (Fig.1.3). PanINs are not detectable by pancreatic imaging and are associated with invasive carcinomas (Goral, 2015; McGuigan *et al.,* 2018). Intraductal papillary mucinous neoplasms (IPMN) have low malignant potential but if left untreated can progress to uncurable invasive malignancies. They are commonly found among smokers and affect the ducts of the pancreas (Goral, 2015; Vincent *et al.,* 2011). Mucinous cystic neoplasms (MCN) are less frequent precursors usually found in the head or the tail of the pancreas but do not arise in the pancreatic ducts. MCN predominately effects females with a ratio of 20:1 over males (Goral, 2015; Vincent *et al.,* 2011).

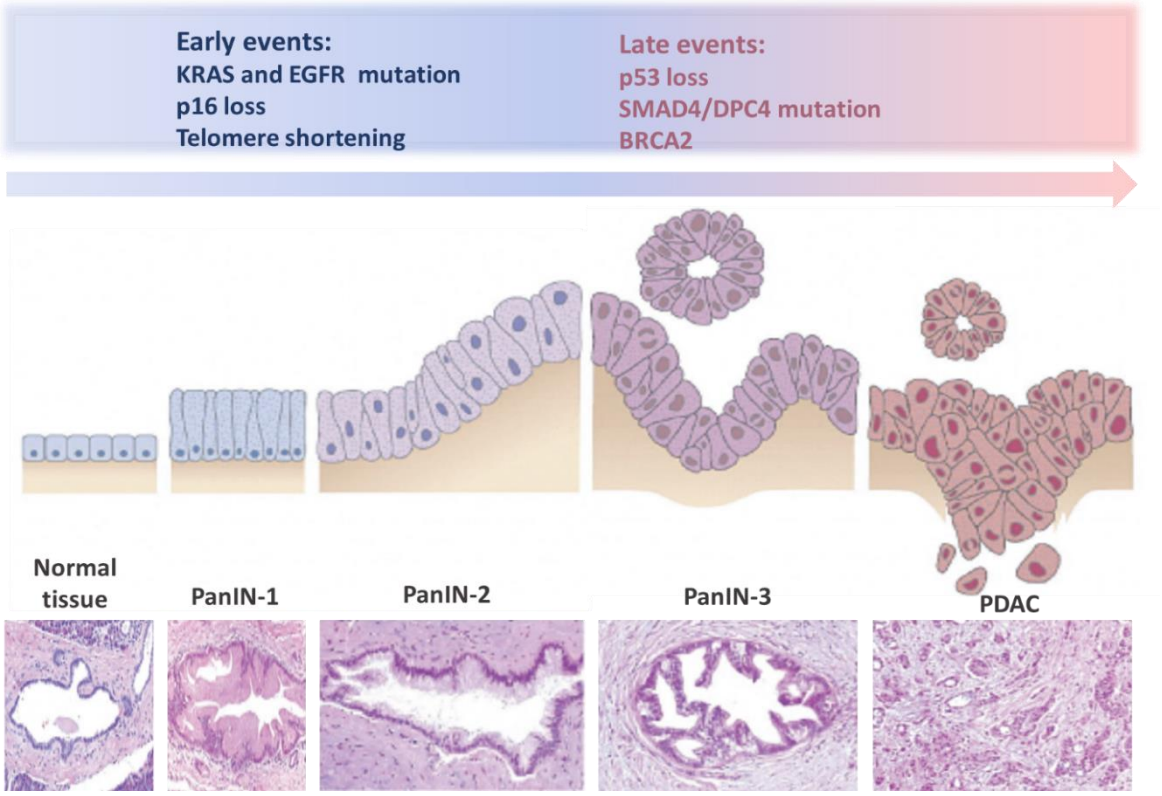


Figure 1.3: Step by step the transformation from a health pancreas to PDAC. Normal tissue image is shown on the left-hand side. When mutations in oncogenes (such as KRAS) occur, PanIN-1 develops. The accumulation of various mutations and loss of p16 (in many cases) lead to PanIN-2. When DNA damage checkpoints are lost, and the number of oncogenes increases the lesions progress to PanIN-3 and eventually PDAC. Image modified from: (Goral, 2015).

Oncogenes are mutated genes that encode proteins such as transcription factors, apoptotic inhibitors, growth factors and signal transducers. These proteins have a fundamental role in cell proliferation, differentiation, migration and apoptosis, and when mutated, they favour cancer growth. On the other hand, tumour-suppressor genes encode proteins that antagonise oncogenes and prevent cancer formation and progression. Such genes encode members of the DNA repair systems, transcription factors and activators of apoptosis. Successive accumulation of both oncogenes and mutated tumour-suppressor genes are usually required for tumour development (Hidalgo, 2012; Maddalena *et al.,* 2021). As observed in ninety percent of tumours, most PDAC tumours have mutations activating the Kirsten rat sarcoma virus (*KRAS*) oncogene. *KRAS* mutations can alter the differentiation and proliferation profile of a cell, but this alone is insufficient to develop PC (Grant, Hua and Singh, 2016). Inactivating mutations of tumour-suppressor genes are also required. Commonly silenced tumour-suppressor genes in PDAC are Tumor protein 53 (*TP53*), *CDKN2A/p16* and Mothers against decapentaplegic homolog 4/Deleted in pancreatic cancer 4 (*SMAD4/DPC4*). *CDKN2A* encodes the p16/INK4A protein, a kinase inhibitor responsible for arresting the cell cycle at the G1/S checkpoint and this gene is mutated in 60% of patients with PDAC (Grant, Hua and Singh, 2016). *TP53* missense mutations are also associated

with enhanced fibrosis and an immunosuppressive microenvironment in PDAC (Maddalena *et al.,* 2021). Less prevalent mutations involve in PDAC are on Epidermal growth factor receptor (*EGFR*), SWI/SNF Related, Matrix Associated, Actin Dependent Regulator Of Chromatin, Subfamily A, Member 4 (*SMARC4A*), Neurofibromatosis type 1 (*NF1*), Cadherin1 (*CDH1*), F-box and WD repeat domain containing 7 (*FBXW7*) and Isocitrate Dehydrogenase (NADP(+)) 1 (*IDH1*) genes (Hidalgo, 2012; Wood and Hruban, 2015).

1.1.6 The tumour microenvironment

Desmoplasia, a dense extracellular matrix surrounding the tumour, is the most prominent histological feature of PDAC. As shown in Figure 1.4, the cancer cells are surrounded by a complex and heterogeneous stroma that can constitute up to 90% of tumour mass. Its heterogenicity arises from a combination of cellular and acellular components (Yang *et al.,* 2020). The cellular components include immune cells, carcinoma-associated fibroblasts (CAFs) and pancreatic stellate cells (PSCs). These cells are surrounded by an extracellular matrix (ECM) rich in proteins, including non-structural matricellular proteins (e.g, periostin, the SPARC family, tenascins, osteopontin), structural ECM proteins (e.g., glycoproteins, proteoglycans and collagens) and secreted soluble proteins (e.g., VEGF and CTGF) (Hidalgo, 2012; Farrow, Albo and Berger, 2008; Fujita *et al.,* 2009).

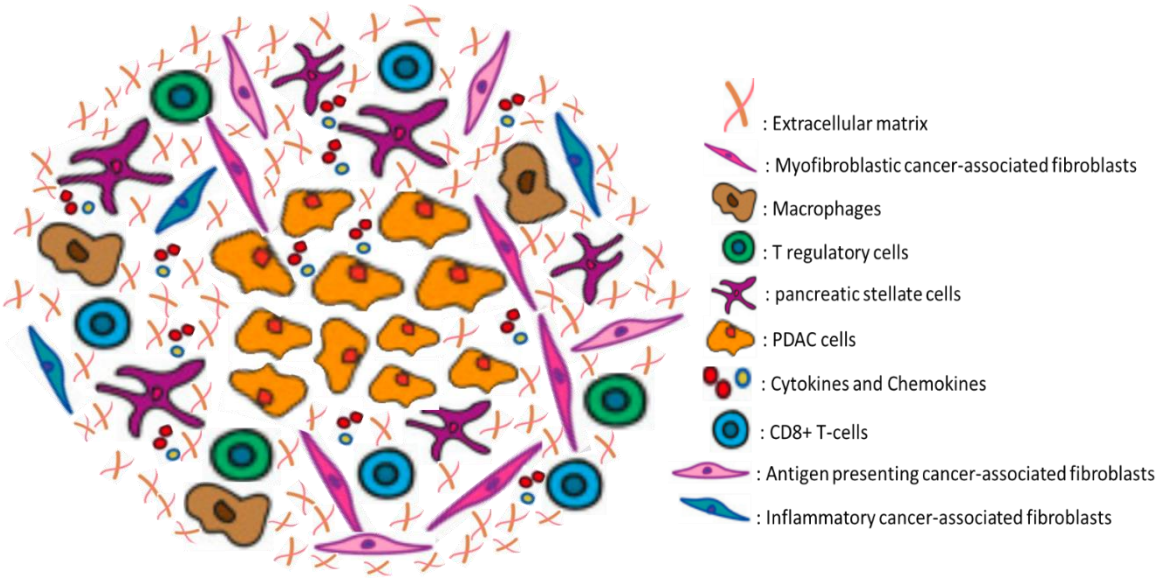


Figure 1.4: The PDAC microenvironment. Diagram shows the dense stroma, a characteristic feature of PDAC, surrounding the cancer cells. It is composed of numerus different cell types, including immune cells, activated pancreatic stellate cells and cancer-associated fibroblasts. All these cells are surrounded by an extracellular matrix rich in cytokines and chemokines. Image drawn from cell diagrams acquired from (Narayanan, Vicent and Ponz-Sarvisé, 2021)

Even though initially the desmoplastic reaction was considered as a host barrier against the tumour invasion, it has now been proven to create a sophisticated microenvironment that facilitates the growth, invasion and metastases of the tumour (Yuan *et al.,* 2019; Fujita *et al.,* 2009). Pancreatic stellate cell (PSCs), also known as cancer-associated myofibroblasts, play a key role in the stroma formation and turnover (Fujita *et al.,* 2009). Upon activation by soluble factors released by cancer cells (e.g., Tumour necrosis factor alpha (TNF-α), Interleukin-1/-6 (IL-1/-6), Platelet derived growth factor (PDGF) and Fibroblast growth factor (FGF)), PSCs secrete large amounts of ECM proteins (e.g., collagen), growth factors (e.g., Smooth muscle cell derived growth factor (SDGF), Transforming growth factor beta (TGF-β), Connective tissue growth factor (CTGF), Vascular endothelial growth factor (VEGF) and Epidermal growth factor (EGF)), integrins, cytokines and chemokines. Such components favour the activation of cancer cells, that facilitates the continued activation of PSCs creating a loop of cell activation essential for tumour growth, survival and metastasis (Hidalgo, 2012; Farrow, Albo and Berger, 2008; Hwang *et al.,* 2008). Hwang *et al.* (2008) and Yuan *et al*. (2019) found that conditioned medium from activated PSCs stimulated pancreatic tumour cell proliferation, migration and invasion.

The stroma plays a pivotal role in tumour progression and chemoresistance by promoting the formation of an immunosuppressive tumour microenvironment with low vascularisation (Yuan *et al.,* 2019; Xie and Xie, 2015). Due to its significance in PDAC, the stroma has been a target of many chemotherapy trials but has demonstrated limited success (Haqq *et al.,* 2014; Bhaw-Luximon and Jhurry, 2015; Feig *et al.,* 2012). Matrix metalloproteinases (MMPs) are cell secreted enzymes that promote ECM degradation and turnover and are associated with tumour progression. Clinical trials using synthetic MMPs inhibitor Marimastat (either alone or in combination with gemcitabine) showed no benefit to patient survival (Bramhall *et al.,* 2002; Bramhall *et al.,* 2001). Strikingly, in another phase 3 clinical trial, treatment with MMP inhibitor BAY-12-9566 had a negative effect on patient survival (Moore *et al.,* 2003).

Connective tissue growth factor (CTGF) is another overexpressed protein in PDAC that plays an important role in cell biology and the ECM. Pamrevlumab is a human monoclonal antibody that targets CTGF. Administration of pamrevlumab with gemcitabine in clinical trials resulted in enhanced tumour sensitivity to gemcitabine. In stage 2 and 3 trials, combination therapy of pamrevlumab, gemcitabine and erlotinib enhanced tumour response to the chemotherapy and improved resection rates in patients with advanced PC. The clinical trial is still in progress, and more data is required for commercial and widespread use of pamrevlumab (Picozzi *et al.,* 2020).

Focal adhesion kinase (FAK) is a coordinator of cell migration and proliferation and signals downstream integrins and growth factor receptors. Combination treatment of gemcitabine with

tamoxifen, a drug which induces FAK deletion in endothelial cells, was shown to reduce liver cancer metastasis and increase overall survival in mouse models and patients (Roy-Luzarraga *et al.,* 2022). The authors further clarified that this effect was not a result of alterations on primary tumour angiogenesis, blood vessel leakage, or early metastatic events. Specifically, tumour cells were less able to colonize the liver even though there was no difference in the homing and seeding of the cells. Liver metastases are less hypoxic compared to the primary tumour, and different signalling pathways (including endothelial cell angiocrine signals) can be triggered in different oxygen conditions. This suggests that different hypoxia levels in the environment could explain a contrasting angiocrine regulation by FAK, resulting in decreased metastasis but no change in the size of the primary tumour (Roy-Luzarraga *et al.,* 2022).

The blood vessels of pancreatic tumours have been a target of chemotherapy trials. Blood vessels play an important role in drug delivery, hypoxia, metastasis and the desmoplastic reaction. Antiangiogenic therapy aimed at reducing the tumour size, controversially resulted in increased hypoxia, impaired drug delivery and potentially increased metastasis in mouse models (Ebos *et al.,* 2009; Pàez-Ribes *et al.,* 2009). Vascular normalisation, resulting from the reduction of blood vessel density, leading to temporarily increase of bloody flow gave promising results. However, it is time and dose dependant and considered challenging to be clinically implemented (Webb, 2005). A different strategy with proangiogenic agents was used. Vascular promotion aims to increase vessel density, dilation, blood flow and leakiness, resulting in increased drug delivery and intracellular uptake of chemotherapy and decreased hypoxia. Indeed, treatment with low dose of Cilengitide and Verapamil, in combination with Gemcitabine led to reduction of tumour size and metastasis. According to the authors, this affect was potentially due to a combination of the increased intratumoural delivery of these chemotherapeutic drugs, increased oxygenation which improves chemosensitivity, together with the direct effect of Cilengitide on Gemcitabine uptake into cancer cells (Wong *et al.,* 2015).

1.2 ALTERATIVE SPLICING

1.2.1 RNA splicing

In the late 1970s, researchers comparing the mRNA to the genomic sequence in adenovirus, observed that regions of the sequence were removed from the pre-mRNA and the remaining regions were re-joined together prior to the release of the mRNA into the cytosol (Chow *et al.,* 1977; Berget, Moore and Sharp, 1977). This discovery led to the fundamental principal of RNA

splicing (Kelemen *et al.,* 2013a; Smathers and Robart, 2019). RNA splicing occurs in two forms; 1) constitutive splicing, where introns are removed and the exons are ligated in the same order as they appear in the gene and 2) alternative splicing (AS), where exons from the pre-mRNA are spliced in different arrangements resulting in various isoforms of the mRNA (Blencowe, 2006; Chen, M. and Manley, 2009). During AS the exons can be skipped or enclosed in different patterns. Figure 1.5 shows the most frequently occurring alternative splicing events. The vast majority of human genes are subject to alternative splicing, resulting in an exceptional proteomic diversity essential for higher eukaryotes (Wang and Burge, 2008).

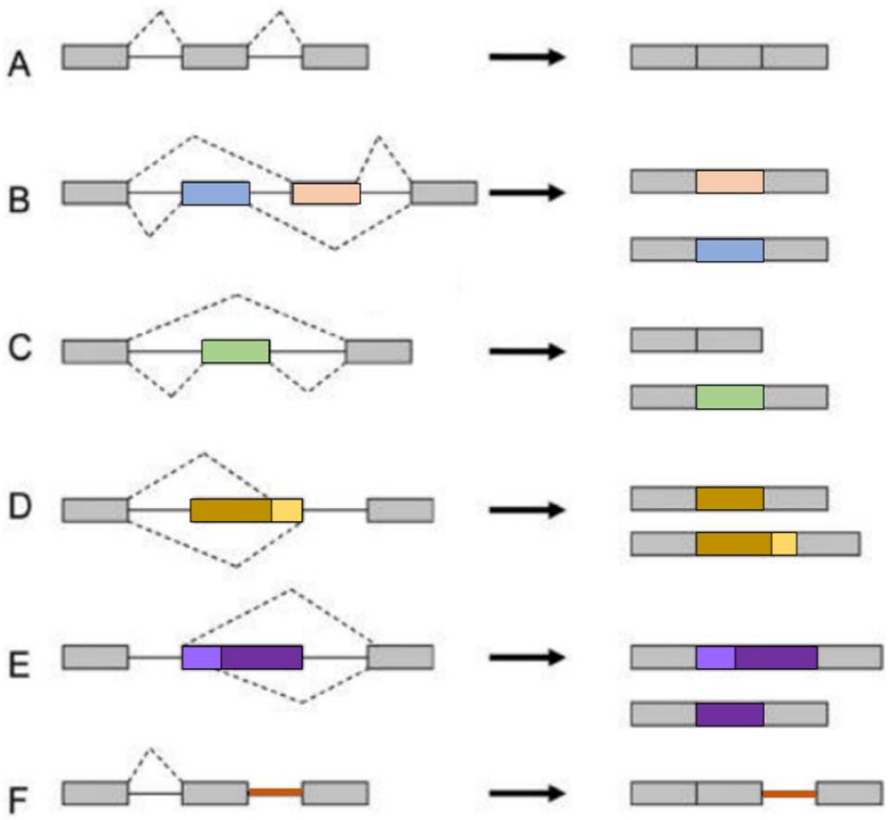


Figure 1.5: The five main types of alternative splicing events. (A) Constitutive splicing: where the exon is always present **(B)** Mutually exclusive exons: where one or the other exon is included **(C)** Cassette alternative exon: where exon is either included or excluded, **(D)** Alternative 3' splice site: where the end of the exon can be included or excluded **(E)** Alternative 5' splice site: where the beginning of the exon can be included or excluded **(F)** Intron retention: where an intron between two exons is included. Adapted from: (Chen and Manley, 2009).

1.2.2 Mechanism of splicing

Exons represent only a small sequence (50-300nt) separated by large introns (~3,500nt) within the pre-mRNA. A very precise recognition of the intron and exon boundaries is necessary for accurate splicing (Smathers and Robart, 2019; Chen and Manley, 2009). There are 3 sites found within every intron sequence that are characterised as the core splicing signals: the 5' splice site (5'ss), the 3' splice site (3'ss), and the branch point sequence (BPS) (Fig. 1.6) (Hertel and Graveley, 2005).

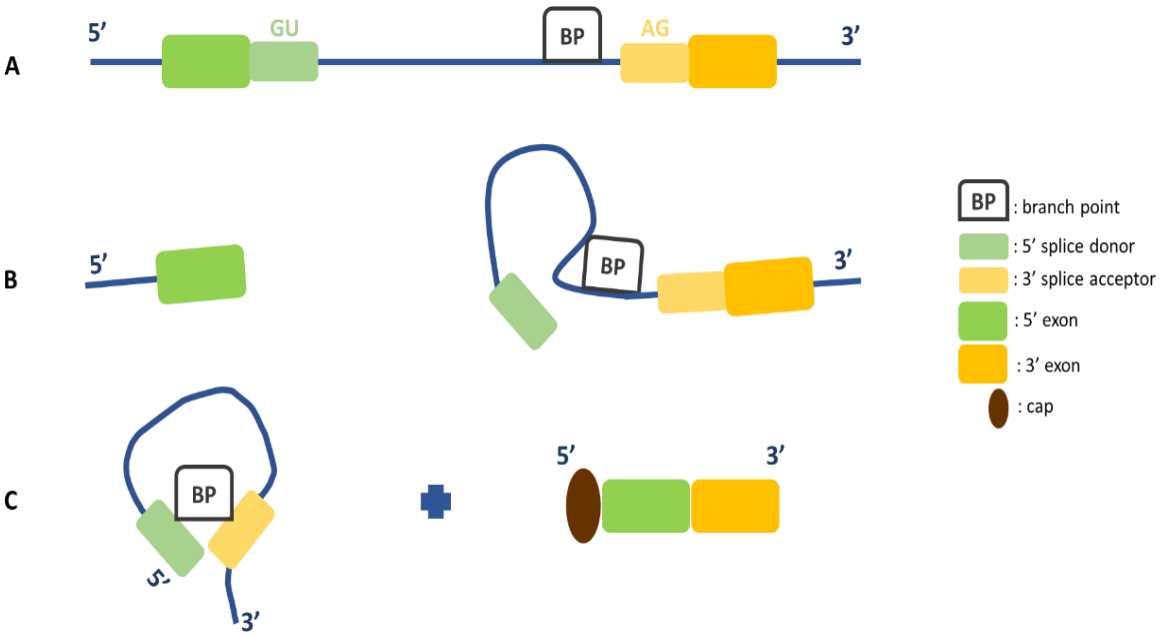


Figure 1.6: mRNA processing in cells. The three steps of mRNA processing with the interactions of the core splicing signals; the 5' splice site (ss) (5' splice donor), the 3' splice site (ss) (3' splice acceptor), and the branch point (BP): (A) Exon removal requires the presence of specific dinucleotide sequences found on its termini; conserved GU at the 5' splicing site and AG at the 3' slicing site. (B) Cleavage of the pre-mRNA at the 5'ss, is followed by the joining of the 5' end to an adenosine unit found within the BP. (C) Lariat is formed by the ligation of the 5' and 3' ss and the intron is realised. The mRNA in now formed (on the right-hand site).

The Spliceosome is a large macromolecular complex composed of five small nuclear ribonucleoprotein particle (snRNPs) and at least 170 auxiliary proteins (Chen and Manley, 2009; Kelemen *et al.,* 2013). The assembly is initiated by the binding of U1 snRNP to 5' splicing site, following by the binding of splicing factor 1 (SF1) which facilitates the association of U2 auxiliary factor (U2AF) heterodimer to the 3' splice site. Through the exchange and recruitment of more factors the spliceosome complex undergoes changes leading to intron removal and the joining of the exons in a two steps transesterification reaction (Fig. 1.7) (Hertel and Graveley, 2005; Kelemen *et al.,* 2013; Chen and Manley, 2009). The core splicing sequences are recognized multiple times during spliceosome assembly.

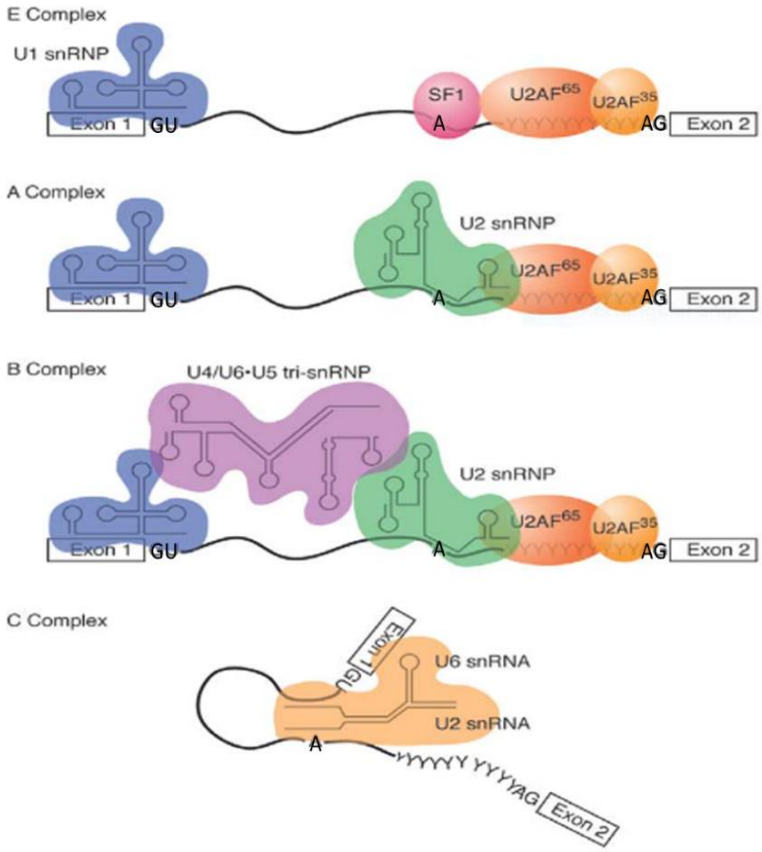


Figure 1.7. Spliceosome assembly onto pre-mRNA; E complex consist of U1 snRNP bound to the 5' splicing site, SF1 bound to the branch point and both U2AFs bound, one to the pyrimidine tract and the other to 3'ss AG. In complex A, U2 snRNP replaces SF1. Then, the recruitment of U4/C6·U5 tri-snRNP leads to the formation of complex B. The catalytically active complex C if formed through rearrangements. In this complex U2 and U6 interact and U6 replaces U1 at the 5'ss. Adapted from: (Hertel and Graveley, 2005).

1.2.3 Alternative splicing regulation

Despite its high potential for errors, pre-mRNA splicing is a very orchestrated and accurate process. This fidelity implies the involvement of various transcript features. The inclusion or exclusion of exons during the pre-mRNA processing is entirely defined by the interaction between trans-acting factors and cis-acting elements. Cis-acting elements are divided into four categories depending on their position and function. Elements that enhance or inhibit the expression of an exon that they associate with are classified as exonic splicing enhancers (ESEs) or silencers (ESSs) respectively. The elements that promote or inhibit the use of an adjacent splice site or exon from an intronic location are categorised as intronic splicing enhancers (ISEs) or silencers (ISSs) respectively (Yeo *et al.,* 2004; Chen and Manley, 2009). ESEs and ISE are bound by positive trans-acting factors, such as serine/arginine-rich family of nuclear phosphoproteins (SR proteins) while ESSs and ISSs are bound by negative acting factors, such as heterogeneous nuclear ribonucleoproteins (hnRNPs) as well as neuro-oncological ventral antigen-1/-2 and FOX-1/-2 (Shenasa and Hertel, 2019; Rooke *et al.,* 2003). Alternative splicing is thought to be a very complex

process within a cell and is influenced by numerus other factors such RNA secondary structure, RNA polymerase II transcription speed (faster speed leads to exon skipping) and intron-exon architecture (longer introns inhibit exon expression) (Shenasa and Hertel, 2019).

1.2.4 Role of alternative splicing

Alternative splicing has a pivotal role in the normal function of the cell. For instance, a protein can be both pro-apoptotic or anti-apoptotic but alternative splicing can allow the selection of either function. Such example is caspase 3, where the deletion of exon 6 generates an isoform with anti-apoptotic function (Vegran *et al.,* 2005). In contrast, the inclusion of exons 13 and 16 generates a pro-apoptotic isoform of MAP-kinase activating death domain that inhibit cell proliferation (Mulherkar, Prasad and Prabhakar, 2007). Insertion of exon 3 in human peroxisome proliferator-activated receptor gamma (hPPARgamma) generates a truncated isoform that was found to enhance proliferation of Chinese hamster ovary (CHO) cells and induce anchorage-independent colony growth on agar (Kim, H. J. *et al.,* 2006). The role of alternative splicing in normal and pathological cell processes is an emerging area of research (Kelemen *et al.,* 2013b).

1.3 POSTN AND CANCER

1.3.1 Periostin

Periostin (POSTN), previously termed osteoblast-specific factor 2, is a 93.3 kDa extracellular matrix protein encoded by POSTN gene on chromosome 13. The protein sequence consists of a cysteine-rich amino-terminal EMI domain, repeats of four fasciclin domains (FAS1) that interact with integrins, and a carboxy-terminal domain (hydrophilic C-terminal), including a heparin-binding site and the alternatively spliced region. These domains function as binding sites for ECM proteins such as collagen, fibronectin (EMI domain), tenascin-C (FAS1 domains) and heparin (heparin-binding site) (Fig.1.8) (Park *et al.,* 2016).

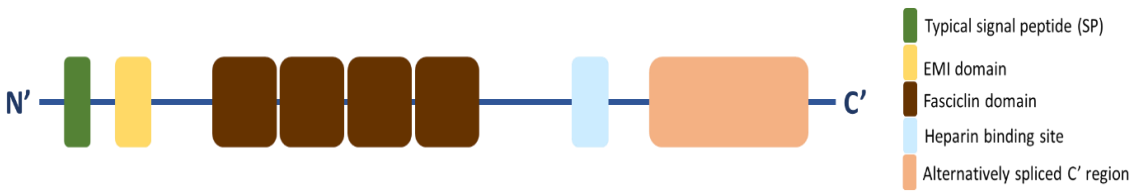


Figure 1.8: POSTN protein domains. Periostin consists of a typical signal peptide and an EMI domain followed by four fasciclin domains. Just before the alternatively spliced region on the C' termini, a heparin-binding site is found.

1.3.2 Exonic analysis and alternative splicing of Periostin

The periostin gene consists of 23 exons, with a genomic footprint covering approximately 36 kilobases. Exons 1 to 16 are constitutive exons which encode known domains of the protein including fasciclin and EMI domains. Exons 15-23 encode the C-terminal region. This region is devoid of known conserved functional domains and seems to be intrinsically disordered. Within this region, exons 17-21 appear to be exons that can be included or excluded from the mature mRNA (known as cassette exons), giving rise to multiple different transcript variants (Fig. 1.9) (Viloria and Hill, 2016; Kim, C. J., Isono, Tambe, Chano, Okabe, Okada and Inoue, 2008a). This alternatively spliced region appears to be important as it contains known sequence motifs, including regions of low compositional complexity (LCRs) and intrinsic disorder (IDRs). LCRs are regions with fewer amino acid types compare to an average compositional region, and even though their role is still unclear, it is thought to be structural (Mier *et al.,* 2020). IDRs are polypeptide segments that are not likely to form a defined three-dimensional structure due to absence of hydrophobic amino acids and appear to be related to the binding properties of the synthesised protein (Uversky, 2013). Lastly, a C-terminal nuclear localization signal is encoded in the AS region. These signal fragments mediate the transport of proteins from the cytoplasm into the nucleus, however this effect is in contrast with the known function of periostin as secreted protein (Hoersch and Andrade-Navarro, 2010a).

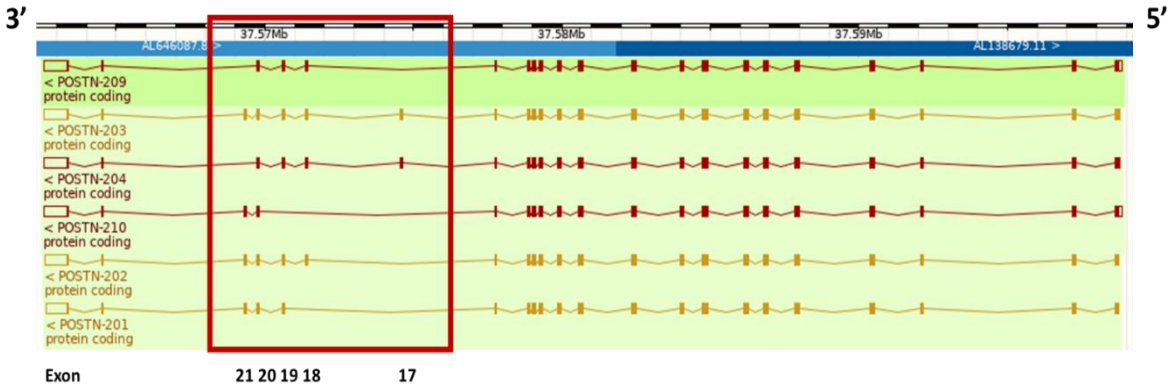


Figure 1.9: c Six transcript variants are synthesised through the alternative splicing of POSTN. The region where the alternative exons are found is in the red box and the exons are numbered on the bottom of the image. Figure was obtained from the ENSEMBL database, accessed on 02/05/2022.

1.3.3 The function of POSTN in normal tissues

Periostin is defined by its ability to interact with ECM proteins, through its EMI domain and C-terminal region, and cell-surface integrin receptors through the fasciclin domains. It is preferentially expressed in the membrane covering the bones known as the periosteum, which

contributes to the vertical growth of bones, its cortical thickness, and the bone diameter (Kormann *et al.,* 2020). Periostin is also expressed in other connective tissue rich in collagen which undergo mechanical stress. Examples of such tissue include the heart valves, periodontal ligament, tendons, and respiratory tissue. Periostin also possesses an important role in the organisation and function of ECM structures, such as by acting as a scaffold in the collagen fibrillogenic and remodelling process. It has been shown to bind to various integrins, including αvβ1, αvβ3 and αvβ5, eliciting activation of focal adhesion kinase (FAK), protein kinase B (AKT) or phosphoinositide 3-kinases (PI3K) signal pathways. In skin, periostin is significantly involved in excisional wound repair by facilitating myofibroblast differentiation through a β1 integrin/FAK dependent mechanism. In gingival healing, it regulates ECM synthesis by increasing fibronectin and collagen synthesis via integrin β1, FAK and Jun N-terminal Kinase (JNK) dependant mechanism (Nikoloudaki *et al.,* 2020).The ability of POSTN to activate these pathways signifies its contribution to cell proliferation, survival, adhesion and migration (O'Dwyer and Moore, 2017; Nikoloudaki *et al.,* 2020).

1.3.4 Periostin in tumorigenesis

Numerous studies have revealed the upregulation of periostin in cancer tissues compared to their respective normal tissue. Periostin is frequently overexpressed in breast, colon and ovarian cancer and is reported promote tumour progression through enhanced cancer cell survival and metastasis (Hoersch and Andrade-Navarro, 2010).

Periostin has FAS1 domains that allow the interaction with proteins of the integrin family, including aVb5 and aVb3. Integrins are transmembrane heterodimeric receptors involved in cell-ECM and cell-cell interactions. Their wider function involves the regulation of not only cell adhesion and migration but also cell proliferation and survival via activation of cytosolic signalling cascade (Mishra *et al.,* 2020; Kudo *et al.,* 2007).

Changes in cell morphology and motility are essential for progression of a carcinoma *in situ* to an invasive/metastatic tumour (Kudo *et al.,* 2007). A study comparing the expression of an array of proteins in oral cancer cell line clones showed that periostin expression had the highest fold-change in expression in the more invasive clones compared their less aggressive counterparts (Kudo *et al.,* 2004). Periostin overexpression in 293T cells induced cell invasiveness and increased migration (Yan and Shao, 2006). These results suggest the involvement of periostin in invasiveness of various cancer cells (Kudo *et al.,* 2007).

Cells need to overcome programmed cell death and stimulate proliferation in order to become established tumours. They need to overcome cellular stresses such as hypoxia and nutrient

deprivation and periostin has been reported to aid cells in this (Baril *et al.,* 2007a). Additionally, treatment of colon cancer cells with anti-periostin antibody activated apoptosis (Tai, Dai and Chen, 2005). Together these studies suggests that periostin plays a critical role in cancer cell survival (Kudo *et al.,* 2007).

Angiogenesis is vital for the later stages of tumorigenesis and is required for tumour growth. It requires endothelial cell proliferation, migration, and tube formation. Vascular endothelial growth factor (VEGF) and its receptor (VEGFR2) have been demonstrated to be involved in the induction of angiogenesis. Integrin αvß3 engagement with activated VEGFR2 is essential for endothelial cells migration and adhesion. Periostin promotes neovascularization and enhance capillary formation whilst enhancing VEGF-C receptor VEGFR2 expression in endothelial cells (Siriwardena *et al.,* 2006; Shao *et al.,* 2004) . These findings suggest the involvement of periostin in epithelial cell survival and capillary formation and consequently to the progression of the disease (Ruan, Bao and Ouyang, 2009; Kudo *et al.,* 2007).

1.3.5 Alternative splicing of periostin in cancer

Periostin expression is upregulated in many cancer types including breast, lung and colon cancer. In contrast, periostin appears to be downregulated in bladder cancer (Kim, H. R. *et al.,* 2016; Oh *et al.,* 2017; Kim *et al.,* 2008; Fujikawa *et al.,* 2021). These controversial findings suggest the possibility of different functions of periostin transcript variants. Such evidence is further supported by Hoersch and Andrade-Navarro (2010) who suggest the possibility that the C-terminal domain of the protein, which is subject to alternative splicing, has a critical role in POSTN function (Hoersch and Andrade-Navarro, 2010b).

Kim *et al*. (2008) made a more in-depth analysis on the transcript variants, in normal and cancer bladder cells and their potential functional differences. The authors identify four transcript variants, corresponding to isoforms 1, 2, 3 and 4 (Table 1.1). Indeed, the expression of the full-length mRNA of POSTN was fully suppressed in both bladder cancer cell lines and tissues, where 48% of bladder cancer tissues and 33% of bladder cancer cell were still expressing alternatively spliced mRNAs, suggesting the importance of ‘’WT’’ variant mRNA in bladder carcinogenesis. Additionally, isoform 4 was detected in high concentrations in bladder cancer tissues compared to other isoforms. Further analysis showed that isoforms 1 and 3 suppress cell invasion and metastasis, a functionality that isoform 4 was lacking. Isoform 4 encodes exons 19 and 20 and lack exon 18, which the authors speculate induce structural changes that modulate the invasive and metastatic functions of periostin (Kim, C. J., Isono, Tambe, Chano, Okabe, Okada and Inoue, 2008b).

Table 1.1: POSTN transcript variants encoded by *POSTN* gene

ENSEMBL variant	Common name (Kim *et al.,* 2008)	Genbank ID	Exons present in alternatively spliced region	AA	TSL
POSTN-209	''Variant II''	NM_001135935.2	18-20	781	1
POSTN-203	''WT''	NM_006475.3	17-21	836	1
POSTN-204		NM_001330517.2	17-20	808	5
POSTN-210		NM_001286666.2	20,21	749	1
POSTN-202		NM_001286665.2	18-21	809	1
POSTN-201	''Variant III''	NM_001135934.2	19-21	779	1
	''Variant I''	NM_001135936.2	19,20	751	
Splice variants as found on Ensembl, literature and Genbank. Additional information on the encoded exons, the amino acids number and the Transcript Support Level (TSL) of the variants are displayed.					

Fujikawa *et al*. (2021) demonstrated that transcript variants of periostin that contain exon 21 enhanced the expression of M2 tumour associated macrophages (TAM) markers and reduced the expression of M1 TAM markers *in vitro*. TAM cells are the most abundant inflammatory cell type in breast cancer and higher levels are associated with poor prognosis. In particular, M1 TAMs enhance tumour activity whereas M2 TAMs antagonize anti-tumor processes by producing high levels of angiogenic and anti-inflammatory factors (Jayasingam *et al*., 2020). Based on these findings, Fujikawa *et al*. suggest that antibodies raised against exon 21 of periostin could reduce tumour growth in Triple-Negative Breast Cancer (Fujikawa *et al*., 2021).

A similar approach has been made towards exon 17 in other studies (Ikeda-Iwabu *et al*., 2021; Kyutoku *et al*., 2011). Kyutoku *et al*. (2011) induced antibodies raised against exon 17 of periostin in breast cancer models. The neutralisation of the variants including exon 17 inhibited cell detachment and cell proliferation. These changes blocked cell migration and invasion and induced tumour necrosis *in vitro*. Furthermore, a study demonstrated the binding of periostin to wnt3a activated wnt signalling, a pathway related to cell differentiation, mobility and polarization in breast cancer cell (Malanchi *et al*., 2011). In addition, Ikeda-Iwabu *et al*. (2021) revealed that a transcript variant with exon 17 strongly bound to wnt3a, an ability that was lacking in the variant that exclude exon 17. Based on these results, they concluded that the presence of exon 17 can alter the function and the properties of periostin isoforms and suggested the therapeutic potential of exon 17 antibodies in the treatment of breast cancer.

1.3.6 Periostin in pancreatic ductal adenocarcinoma

Pancreatic stellate cells (PSCs) are one of the cell types found within the stroma of PDAC. Upon activation by cancer cells, these cells were found to express large amounts of periostin. In

contrast, periostin expression in pancreatic cancer cell was low or absent (Kanno *et al.,* 2008; Liu and Du, 2015)

The influence of periostin in the progression of PDAC has been well established by numerous studies. Kanno *et al*. (2008) demonstrated the key role of periostin in tumour-stromal interactions and the ability of periostin to promote migration and metastasis when overexpressed. Liu *et al*. (2016) demonstrated the pivotal role of periostin in the desmoplastic reaction and the formation of the characteristic microenvironment of PDAC, shortening the overall survival of the patients. Their findings demonstrated that periostin promotes metastasis and angiogenesis in pancreatic cells via Erk/VEGF signalling pathways and silencing of *POSTN* in nude mice suppressed capillary-like vessels formation. Another study showed that periostin could sustain the activation of pancreatic stellate cells via an autocrine loop, creating a tumor-supportive microenvironment through the overexpression of extracellular matrix proteins (Erkan *et al.,* 2007). Lastly, the interaction of periostin with the α6ß4 integrin complex promoted cell migration through focal adhesion kinase (FAK) phosphorylation. Periostin influenced survival of the cancer cells by binding to ß4 integrin, activating Phosphoinositide 3 (PI3) kinase pathway and inducing Protein kinase B (Akt) phosphorylation (Baril *et al.,* 2007b).

Despite the notion that alternative splicing of periostin holds an important role in various cancer types and their progression, there are no studies on the distinct functions of periostin isoforms in PDAC (Fujikawa *et al.,* 2021; Ikeda-Iwabu *et al.,* 2021; Jayasingam *et al.,* 2020; Malanchi *et al.,* 2011).

1.4 AIMS

Currently there is no published data on the periostin isoform expression profile in PDAC. It has been hypothesised that periostin transcript variants may have different binding properties due to alternative splicing in the C-terminal, a region that is thought to affect cell migration and adhesion (Viloria and Hill, 2016). To test this hypothesis a previous study in our laboratory aimed to detect *POSTN* transcript variants in PS1 cells; 6 different isoforms were discovered. However, the isolation of these isoforms was not successful using gel extraction techniques as the transcript variants were too similar in size (Rahim *et al.* 2019, unpublished data).

The aim of this study is to isolate *POSTN* transcript variants by using cloning techniques to express specific isoforms of POSTN. These will then be used to produce recombinant periostin proteins to utilise for functional study.

2 MATERIALS AND METHODS

2.1 CELL CULTURES

All the cell culture reagents and disposables' catalogue numbers are listed on table 2.1. Most of the products were purchased from Fisher Scientific UK. Products supplied by other manufacturers are specified.

2.1.1 Cell line acquisition

Pancreatic stellate cells (PS1) were donated by Professor Hemant Kocher of Queen Mary University of London. The cells were isolated from a donated human pancreas and immortalised as described by (Froeling *et al.,* 2009).

2.1.2 Cell passaging and maintenance

RPMI-1640 basal medium was used for the passaging and maintenance of PS1 cells. The medium was supplemented with 100µg/ml penicillin-streptomycin, 2mM L-glutamine and 10% foetal bovine serum.

PS1 cells were cultured in filtered T75 flasks, placed in HERAcell 150i (Thermo Fisher Scientific) incubators set at 37°C, 5% CO_2 in a humidified environment. Cells were passaged every 3 days at a 1:10 ratio using 0.05% trypsin-EDTA for up to 10 passages post thawing.

2.1.3 Cell counting

PS1 cells were washed with PBS, trypsinized for 4 minutes at 37°C and centrifuged at 1200rpm. The supernatant was discarded, and the cell pellet was resuspended in 1 ml complete medium. Corning trypan blue cell viability stain was used for cell dilution and disposable haemocytometers were used for cell counting under a microscope prior to cell seeding. The formula used for defining cell concentration is displayed bellow. The cell concentration was modified by adding complete medium according on the requirements of each experiment.

$$\textit{Average number of cells x dilution factor x } 10^4 \textit{ = cell concentration (cells/ml)}$$

2.1.4 Cryopreservation for storage

PS1 cells were cryopreserved in RPMI-1640 medium supplemented with 10% FBS and penicillin-streptomycin. Each cryogenic vial contained $1x10^6$ cells suspended in 1ml cryopreservation medium. The vials were transferred in a Mr Frosty freezing container containing isopropanol for

controlled temperature reduction at a rate of -1°C/minute in an ultra-low temperature freezer (-80°C) overnight. The vials were transferred to liquid nitrogen (-196°C) for long-term storage.

Table 2.1 Reagents and disposables used in PC1 cells culture

Reagent	Manufacturer	Catalogue number
RPMI-1640 basal medium	Fisher scientific	12004997
Penicillin-streptomycin	Fisher scientific	11548876
L-glutamine	Fisher scientific	11500626
Foetal bovine serum	Fisher scientific	11573397
T75 flasks	Fisher scientific	10364131
Trypsin-EDTA	Fisher scientific	11590626
PBS	Fisher scientific	12037539
Corning trypan blue cell viability stain	Fisher scientific	15393661
Disposable haemocytometers	Immune Systems	BVS100
DMSO	Merck Life Science	D2650-100ML
Cryogenic vials	Fisher scientific	11750573
Mr Frosty freezing container	Fisher scientific	10110051

2.2 CLONING OF PERIOSTIN TRANSCRIPT VARIANTS

A number of different primers were used during the cloning prosses of *Periostin*. The choice of the primers was based on the area of *POSTN* that needed to be amplified for the experiment. The table bellow illustrates the different primers that were used (table 2.2).

Table 2.2 Summary of the primers used in PCR reactions during *POSTN* cloning

Primer name	Sequence	Amplified area
QARS forward primer	CAGGGGTCTGGCTTATGTGT	QARS gene
QARS reverse primer	CTCTGAAAACTTGCCCTTGC	QARS gene
POSTN-Hind-III forward primer	TTTCGCAAGCTTCGGAGAGACT CAAGATGATTCCC	Full length of POSTN
POSTN-ApaI reverse primer	CAAGTCGGGCCCTCAATGGTG ATGGTGATGATGCT GAGAACGACCTTCCCTTAATCG	Full length of POSTN
CMV-forward	CGCAAATGGGCGGTAGGCGTG	Cloning site of Plasmid
BGH-reverse	TAGAAGGCACAGTCGAGG	Cloning site of Plasmid
POSTN_ex14/15_forward	TCCAGCAGACACACCTGTTG	Exons 15-22
POSTN_ex22_reverse	GTTGGCTTGCAACTTCCTCAC	Exons 15-22

2.2.1 cDNA synthesis

PS1 cells were chosen for cDNA synthesis as they are known to express periostin (Rahim *et al*., 2019, unpublished data). The cells were cultured until they reached 70% confluency and approximately $1x10^7$ cells were used for the cDNA synthesis. The RNA of the cells was extracted using a RNeasy mini kit (#74104, Qiagen). After all the medium from the cells was aspirated, the cells were resuspended in 600µl RLT lysis buffer and 10µl β-mercaptoethanol (#21985023, Thermo Fisher Scientific). Additionally, 70µl of RDD buffer and 5µl of DNase I master mix (#79254, Qiagen)

was added for the elimination of contaminating genomic DNA in the column. The concentration (ng/µl) and purity (A260/230 and A260/280 ratios) of the extracted RNA was measured using a Biodrop µLite spectrophotometer (Thermo Fisher Scientific).

Aiming to synthesise 500ng of cDNA, the exact amount of RNA needed was calculated using the formula below. A RevertAid First Strand cDNA Synthesis kit (#K1621, Thermo Fisher Scientific) was used. Following the manufacturer's instructions, reverse transcriptase was used for the synthesis of the first-strand cDNA. Three different reactions were setup. One that included all the required components for the cDNA synthesis (RT) and two negative controls, one that was lacking reverse transcriptase (NRT) and one with no RNA template control (NTC). Equal reaction volumes were ensured by adding dH_2O. Oligo (dT) primer (#SO132, Thermo Fisher) was used and the thermal conditions were set as shown in Table 2.3. The cDNA template was diluted 1 in 10 from the stock concentration to 25ng/µl prior to use for RT-PCR.

$$\frac{Yield\ of\ cDNA\ \ (ng)}{RNA\ concentration\ (\frac{ng}{\mu l})} = RNA\ volume\ required\ in\ cDNA\ synthesis\ reaction\ (\mu L)$$

Table 2.3 Thermal cycle for cDNA synthesis

Step	Temperature (°C)	Number of Cycles	Time
Initial Denature	94	1	3 min
Denature	94	35	30 sec
Annealing	58		30 sec
Extension	72		45 sec

The concentration and the purity of the synthesised cDNA was checked using a BioDrop. A second RT-PCR reaction was setup to check the performance of the cDNA. Qars primers were used and the thermal cycling conditioned used are shown on table 2.4. The PCR product was run on a 1% agarose gel as described in Section 2.2.3 and an image was taken to confirm the presence of a single band and absence of primer dimers.

2.2.2 Isolation and amplification of *POSTN* gene

POSTN DNA was amplified by RT-PCR. A housekeeping gene was chosen to be amplified as a control, such genes are constitutive genes, required for the basic function of the cells. The chosen gene was the QARS gene and the primers for its amplification were used. QARS primers were used with RT cDNA as positive control. Negative controls were included: NRT and NTC as templates with QARS and the cloning primers (POSTN-Hind-lll and POSTN-ApaI primers) and replacement of primer volume with dH_2O. The master mix preparation, the sequence of the primers and the

thermal conditions used is shown in Table 2.4. The primers were designed by a prior student (Rahim *et al*. 2019, unpublished data). Two restriction enzymes were incorporated to the primers. A Hind-III-site was added to the 5' end of the forward rimer and the ApaI-site was added to the 3'end of the reverse primer. Lastly, a 6x histidine tag was added before the stop codon, to allow for the purification of the recombinant protein (Fig.2.1). Post thermal cycling, 5µl of each reaction was run on a 1% agarose gel as described in Section 2.2.3 and an image was taken to view the bands.

Table 2.4 Master mixes, primer' sequence and thermal cycles used for *POSTN* DNA amplification

Master mix with POSTN and Hind-III and ApaI primers	
Reagent	Volume
HF buffer	10µl
10nM dNTPs	1µl
POSTN-Hind-III forward primer	5µl
POSTN-ApaI reverse primer	5µl
DNA polymerase	1µl
cDNA	10µl (25ng)
dH_2O	18µl

Master mix with QARS primers	
Reagent	Volume
HF buffer	10µl
10nM dNTPs	1µl
QARS forward primer	5µl
QARS reverse primer	5µl
DNA polymerase	1µl
cDNA	10µl (25ng)
dH2O	18µl

Primer	Sequence
QARS forward primer	CAGGGGTCTGGCTTATGTGT
QARS reverse primer	CTCTGAAAACTTGCCCTTGC
POSTN-Hind-III forward primer	TTTCGCAAGCTTCGGAGAGACTCAAGATGATTCCC
POSTN-ApaI reverse primer	CAAGTCGGGCCCTCAATGGTGATGGTGATGATGCT GAGAACGACCTTCCCTTAATCG

Thermal cycling			
Step	Temperature (°C)	Number of Cycles	Time
Initial Denature	98	1	5 min
Denature	98	32	30 sec
Annealing	68		2min 30 sec
Extension	72		2min 30 sec
Final Extension	72	1	5 min

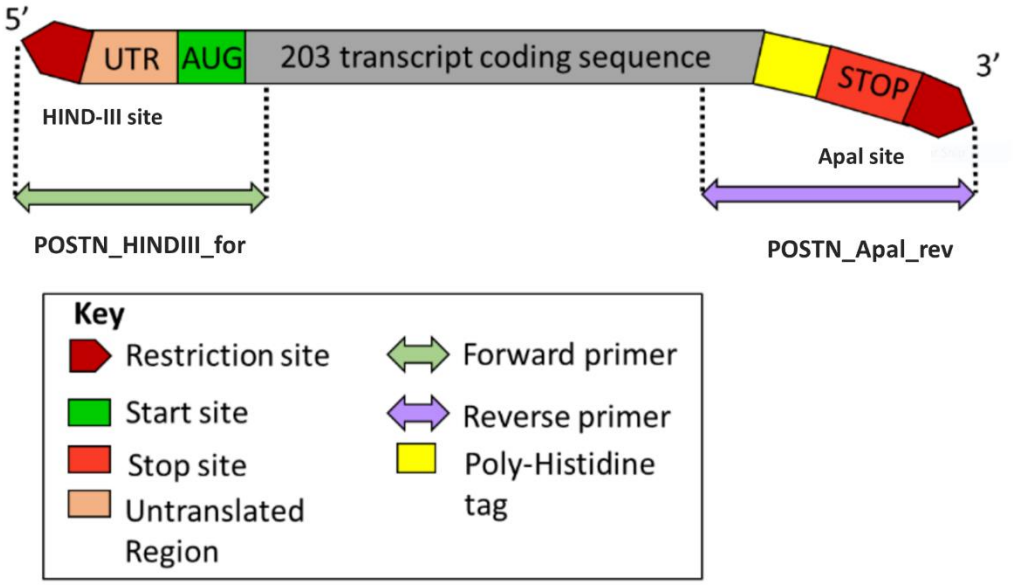


Figure 2.1: Diagram mapping cloning primers on *POSTN* sequence. The forward primer (POSTN-Hind-III) is located at the 5' of the sequence while the reverse primer (POSTN-ApaI) binds in the 3' end of sequence, to amplify the open reading frame (ORF) for cloning. Image adapted from Rahim *et al.,* (2019) unpublished data.

2.2.3 Agarose gel electrophoresis

To prepare the agarose gel, 1x TAE buffer was made by diluting 20ml of 50X TAE buffer (Table 2.5) in 980ml dH_2O. In a beaker, 50ml of 1x TAE and 0.5g of agarose powder (#R0491, TopVisio agarose/ Thermos Fisher scientific for imaging or #16500500, UltraPure agarose/Invitrogen for purification) was combined. The mixture was heated in a microwave for approximately 90 seconds. When the agarose gel was cooled but still molten, 5µl of GelRed Nucleic Acid stain (1x) was added. The gel was poured into a casting tray with both sides taped to prevent leakage and the gel comb (BIO-RAD) was inserted to form wells.

Table 2.5 Components of TAE buffer

Buffer	Components	Weight / Volume
50x TAE	Tris Base	242g
	Glacial acetic acid	57.1ml
	EDTA	19.66g or 100ml of 0.5M
1x TAE	Dilute 50x TAE buffer 1:50 in dH_2O	

When solidified, the tape was removed and the casting tray with the gel was placed into the BIO-RAD gel electrophoresis machine. 15µl of PCR product were loaded into each well of the gel and 5µl of GeneRuler 1kb DNA ladder (#SM0311, Thermo Fisher Scientific) was loaded into a well as a

molecular weight marker. The gel was run for 60-90 minutes at 75V until the bands were separated. The semi-quantitative analysis of the PCR products was visualised using the Syngene G-Box via the GeneSys image acquisition software.

2.2.4 Ampicillin stock

To make an ampicillin stock, 400mg of ampicillin sodium salt (#BP1760-25, Fisher) was dissolved in 4ml of dH_2O to achieve an antibiotic concentration of 100mg/ml. The stock was stored at -20ºC.

2.2.5 LB broth preparation

LB broth was made by dissolving 12.5g of Lysogeny broth (LB broth) powder (#H26760, Alfa Aesar) in 500ml of dH_2O. The solution was sterilised by autoclaving at 121°C for ~1hour and was cooled to 55°C at RT prior to adding 1ml of the ampicillin stock into the broth.

2.2.6 LB agar preparation

In a 500ml glass bottle, 17.5g of Lysogeny broth agar (LB agar) (#L2897, Sigma) were transferred and dH_2O was added up to 500 ml. The mixture was autoclaved at 121°C for ~1 hour. When the temperature of the molten agar had dropped to approximately 55°C, 500µl of ampicillin stock was added. While still molten, 30ml of the agar were transferred into sterile petri dishes. The plates were left to solidify and then wrapped in parafilm and stored upside down at 4°C for up to two weeks.

2.2.7 Overnight culture of *E.coli*

A small amount of the Glycerol stock of *E.coli* cells containing pcDNA4-TO-Puromycin-mVenus-MAP plasmids (#44118, Addgene)(Fig. 2.2) was picked up with a pipet tip and incubated overnight in 50ml falcon tube containing 5ml LB broth. The incubator was set to 37°C at 150 rotations per minute (RPM).

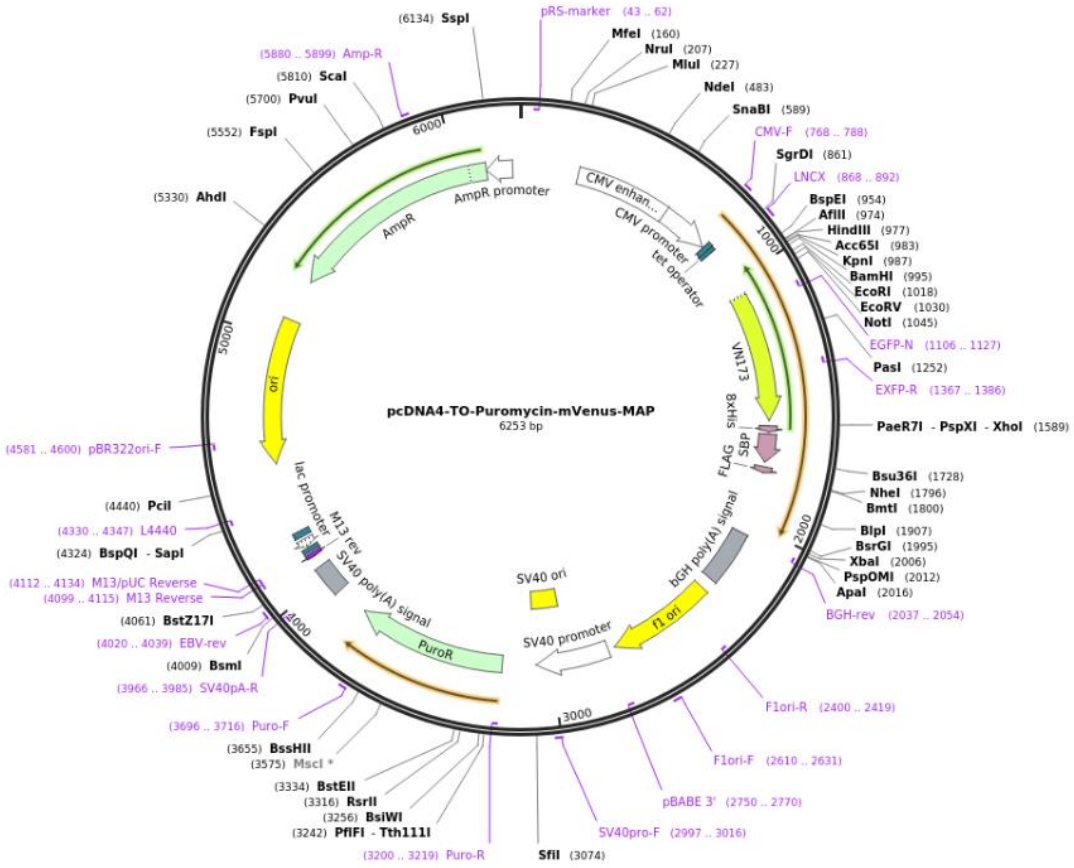


Figure 2.2: Plasmid map. The plasmid was modified from the original pcDNA4-TO backbone. Modified fluorescent protein mVenus with a His8-SBP-FLAG tandem tag was inserted as a NotI/XbaI fragment into pcDNA4-TO-Puromycin vector, which had the Zeocin selection marker in pcDNA4/TO (Invitrogen) replaced by Puromycin. This figure was acquired from the pcDNA™4/TO Invitrogen manual.

2.2.8 Plasmid extraction

Bacterial growth in the overnight *E.coli* culture (Section 2.2.7) was confirmed visually by the cloudiness of the LB broth. The bacteria were collected by centrifugation at 12,000g for 10 minutes. The supernatant was removed, and the plasmid DNA was extracted from the cell pallet by following the manufacturer's instruction for the PureLink HiPure Plasmid Miniprep Kit (#k210002, Thermo Fisher Scientific). The purified plasmid concentration and the purity of the sample were measured using a BioDrop spectrophotometer.

2.2.9 Enzyme digestion

Before the plasmid and the PCR product ligation, both the plasmid and PCR product needed to be double digested with relevant restriction enzymes. Hind III-HF (#R0104, NEB) and ApaI-HF (#R0114S, NEB) restriction enzymes were used for the reaction together with the components in Table 2.6. The 50µl reaction was incubated at 37°C for 15 minutes. For heat-inactivation of the

enzymes, the reaction was then incubated at 80°C for 20 minutes. The samples were then run on 1% agarose gel to confirm the digestion.

Table 2.6 Components of enzyme digestion

Solution	Volume
Plasmid DNA or PCR product	1.5µg
Hind III-HF	1µl
ApaI-HF	1µl
10x CutSmart Buffer	5µl
dH_2O	Up to final reaction volume

2.2.10 Agarose gel purification

Post enzymatic digestion, the plasmid samples were gel purified in order to selectively collect the digested plasmid without the unwanted undigested plasmids and residues from the reaction. The samples were run on a 1% ultra-pure low melting point agarose gel prepared as described in Section 2.2.3. The bands were visualised, cut, and collected in a sterile Eppendorf tub under UV light in a dark room. The plasmids were extracted per manufacturer's instruction by using a QIAquick Gel Extraction Kit (#28706, Qiagen).

2.2.11 DNA purification

The PCR products were purified to achieve better ligation with the plasmids. Monarch PCR & DNA Cleanup Kit (#T1030S, New England Biolab) was used and the purification of the PCR product was performed per manufacturer's instructions.

2.2.12 Plasmid ligation

The gel purified digested plasmids were ligated with the PCR product using a 1:3 vector to insert ratio. 100µg of plasmids used per ligation reaction and the amount of insert required was calculated from the following formula:

$$ng\ of\ insert = \frac{ng\ of\ vector\ \times kb\ size\ of\ insert}{kb\ size\ of\ vector} \times molar\ ratio\ of \frac{insert}{vector}$$

The ligation reaction was setup as shown in Table 2.7. The reaction was prepared on ice and the T4 DNA ligase enzyme was added last. Two different reactions were setup: one containing the insert and vector and the second without an insert to serve as a negative control. The reaction tubes were incubated as per manufacturer's instructions at RT for 10 minutes and then heat-inactivated at 65°C for 10 minutes. The tubes were placed on ice until transformation.

Table 2.7 Components used for ligation of plasmid with insert-PCR product

Components	Volume
T4 Ligase Buffer (10x)	2µl
Vector DNA	100µg
Insert DNA	105.5µg
T4 DNA Ligase enzyme	1µl
Nuclease-free water	Up to 20µl

2.2.13 Transformation

After ligation, the plasmids were ready for transformation into *E.coli* DH5α competent cells (C2987H, New England Biolabs). The cells were thawed on the ice. Four different reactions containing 50µl of *E.coli* were prepared as shown in Table 2.8. and incubated on ice for 30 minutes. Heat shock was performed as per manufacturer instructions at precisely 42°C for 30 seconds. The mixture was then cooled on ice for 5 minutes. 950µl of SOC medium, pre-warmed to RT, was added into each tube and incubated at 37°C while shaking vigorously (250 rpm). After one hour, 100µl of each reaction tube were plated on agar plates that were at 37°C. The plates were incubated overnight at 37°C.

Table 2.8 Tubes containing different components for transformation into *E.coli*

Reaction tube	Plasmid used	Quantity
1	Undigested plasmids (positive control)	12.5ng
2	Digested and ligated plasmid, with no insert (negative control)	50µl
3	No plasmids (negative control)	N/A
4	Digested and ligated plasmid with PCR insert	50µl

2.2.14 Colony PCR

To confirm the success of the ligation and transformation, the product insert was confirmed in the colonies using PCR with CMV-forward and BGH-reverse primers (Table 2.9).

Table 2.9 Primers and their annealing temperature obtained from pcDNA™4/TO manual (Invitrogen)

Primer	Sequence	Annealing temperature
CMV-forward	CGCAAATGGGCGGTAGGCGTG	70°C
BGH-reverse	TAGAAGGCACAGTCGAGG	56°C

These primers were chosen as they are complementary to both sides of the multiple cloning site (MCS) of the plasmid. This would amplify the large insert if the ligation was successful or produce a smaller band if no insert was present (Fig. 2.3).

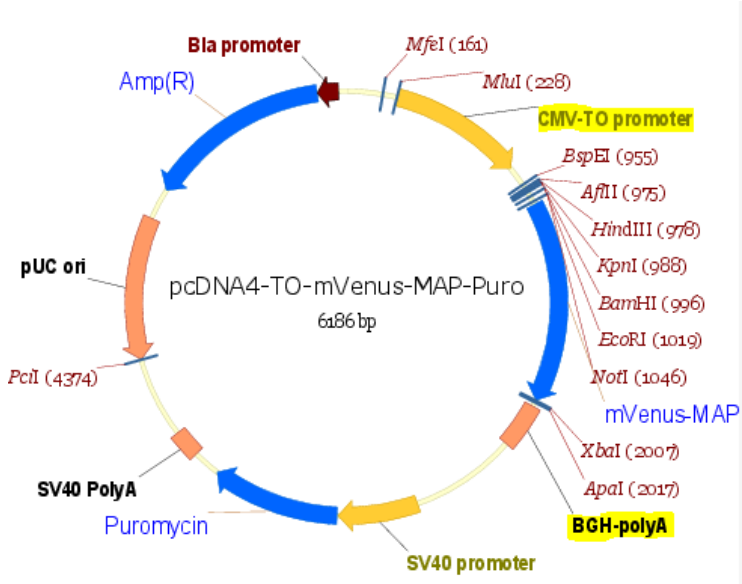


Figure 2.3: pcDNA4-TO-Puromycin-mVenus-MAP plasmid map. Where the primers bind in colony PCR is highlighted in yellow colour. This figure was acquired from the pcDNA™4/TO Invitrogen manual.

Several 25µl PCR reactions were prepared on ice as shown in Table 2.10. The 100µM stock primers were diluted 1:10 in sterile water.

Table 2.10 PCR master mix used for colony PCR

Solution	Volume
2x DreamTaq	12µl
CMV-forward	1µl
BGH-reverse	1µl
dH_2O	10µl
Template	~1µl

Individual colonies were picked with a pipette tip, placed directly into a reaction tube and mixed. Undigested plasmids (1µl) replaced the colony template as reference for empty vessel amplification. Two negative controls were setup for the colony PCR. The first contained undigested and unligated *POSTN* PCR product (Synthesised in Section 2.2.2) to ensure that the primers do not amplify *POSTN* that is not ligated to plasmids. The second negative control did not contain a template to detect any genomic contamination and the volume of the template was replaced by dH_2O. The thermal cycle conditions used is shown in Table 2.11. The initial 10 minutes was added prior to the cycle to allow the release of the plasmids from the *E.coli*.

Table 2.11 Thermal cycle conditions used for colony PCR

Step	Temperature (°C)	Number of Cycles	Time
Initial Denature	95	1	10 min
Denature	95	34	45 sec
Annealing	59		45 sec
Extension	72		60 sec
Final Extension	72	1	5 min

Agarose gel electrophoresis was then performed as described in Section 2.2.3 by loading 15µl of the PCR products and 3µl 1kb ladder into the wells. The remaining products were stored at -20°C.

2.2.15 Sample preparation for Sanger Sequencing

Colonies that expressed the ligated PCR product was prepared for sequencing. The primers were chosen in order to amplify the alternatively spliced region of the gene and help to find the different periostin transcript variants (Table 2.12). The region where the primers are expected to bind the *POSTN* sequence is shown in Figure 2.4.

Table 2.12 Primers used for preparation of samples for sequencing and their annealing temperature

Primer	Sequencing	Annealing temperature
POSTN_ex14/15_forward	TCCAGCAGACACACCTGTTG	65.2°C
POSTN_ex22_reverse	GTTGGCTTGCAACTTCCTCAC	65.9°C

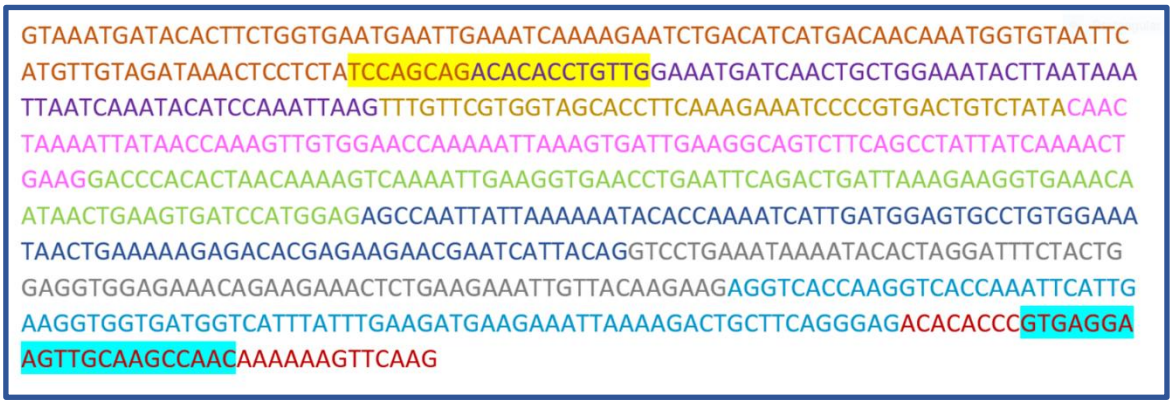


Figure 2.4: The binding sites of POSTN_ex14/15_forward and POSTN_ex22_reverse primers on *POSTN*. The alternatively spliced region of POSTN is shown. Each exon in the sequence is written in different colour: exon14, exon 15, exon 16, exon 17, exon 18, exon 19, exon 20, exon 21, and exon 22. The POSTN_ex14/15_forward primer spans exon 14 and exon 15 and is highlighted in yellow. POSTN_ex22_reverse primer binds to exon 22 and is highlighted in blue.

PCR master mix was prepared on ice as shown in Table 2.13. 100µM primer stock solution was diluted 1:10 in sterile water. The thermal cycling conditions used for the PCR is shown in Table 2.14. Undigested plasmids-PCR product (Synthesised in Section 2.2.14) was used as a template for the negative control.

Table 2.13 PCR master mix for sequencing PCR

Solution	Volume
2x DreamTaq	12µl
POSTN_ex14/15_for	1µl
POSTN_ex22_rev	1µl
dH2O	8µl
Template	3µl

Table 2.14 Thermocycle conditions for sequencing PCR

Step	Temperature (°C)	Number of Cycles	Time
Initial Denature	95	1	3 min
Denature	95	35	30 sec
Annealing	62		30 sec
Extension	72		60 sec
Final Extension	72	1	5 min

The products from the PCR were run on an agarose gel (as described in Section 2.2.3) to confirm the success of the reaction. The remaining products were stored at -20 °C.

2.2.16 Sanger sequencing

The samples were prepared for sequencing by mixing 5µl of the sequencing PCR products containing up to 10ng of DNA with 5µl of 10µM primer. POSTN_ex14/15_forward or POSTN_ex22_reverse primers were used in order to determine which results in a better-quality sequence and to confirm the results.

The samples were sent to Genewiz LTD for Sanger sequencing. The results were accessed online on the Genewiz website. The quality of the sequences was detected using BioEdit (https://bioedit.software.informer.com/7.2/). When the sequence shows "N", it means the nucleotide could not be determined. The reverse sequences were reverse complemented to allow the sequences to be aligned with the forward sequence.

The sequences were copied into the BLAST search box in NCBI (https://blast.ncbi.nlm.nih.gov/Blast.cgi?PROGRAM=blastn&PAGE_TYPE=BlastSearch&LINK_LOC=blasthome) to check the presence of the *POSTN* gene. The settings Homo sapiens (for specific organism) and highly similar sequences(megablast) were selected in the search.

Clustal Omega alignment tool (https://www.ebi.ac.uk/Tools/msa/clustalo/) was used for aligning the sequences with the exons of POSN alternative splicing region. Every sequence was aligned with one exon at a time including exons 15-22, in order to determine the exons included and consequently, the transcript variant.

2.2.17 Plasmid extraction of colonies of interest

Since the colonies expressing the isoforms of interest have been chosen, these colonies were picked and incubated overnight in LB broth at 37°C in a shaking incubator as described in Section 2.2.7. The plasmids were purified for transfection into mammalian cells; therefore, it was important to purchase an endotoxins free product. ZymoPURE™ Plasmid Miniprep Kit (#D4210, Zymo research) was used according to the manufacturer's instructions. The plasmid samples were then quantified using a BioDrop and stored at -20°C.

2.2.18 Glycerol stocks

To make glycerol stock of the colonies of interest, 500µl of the bacterial overnight culture (from Section 2.2.17) was transferred in an Eppendorf tube with 10% v/v of glycerol (#A16205, Alfa Aesar) and gently mixed by inverting the tube. The tubes were then flash-frozen in liquid nitrogen and stored in a ultra-low freezer (-80 ^{o}C).

2.2.19 PS1 cells transfection

PS1 cells were cultured in a 6-well plate at a density of 1x10^{6}cells/well overnight in complete medium. The PS1 cells were washed in PBS, trypsinised, pelleted and resuspended in 70µl Ingenio electroporation solution (MIR 50111, Cambridge Bioscience). The cell/Ingenio mixture and 100 or 200ng of plasmids (from Section 2.2.17) were combined and placed in a 0.2cm cuvette. The mixture was electroporated using a 4D-Nucleofector X Unit (Lonza) set to a CA-137 pulse. The cells were immediately transferred to a 6-well plate in 2ml medium and cultured in an incubator for 48 hours. Post culture, the conditioned medium was collected from the six wells plate. 60µl of the medium was aliquoted into Eppendorf tubes for western blot and the remaining was flash-frozen in liquid nitrogen and stored at -80 ^{o}C for future use. The cells were then washed with PBS. Cell lysate was prepared by adding 2ml of Radio immune Precipitation Assay (RIPA) buffer to the wells for 20minutes on ice. The cells were scraped to facilitate the process and the lysates were transferred into Eppendorf tubes. 60µl of each sample was used immediately for western blot and the remaining was stored at -20°C.

2.2.20 Western Blot

All the reagents that were used during the western blot procedure are listed in Table 2.15. Cell lysates and conditioned medium from Section 2.2.19 was diluted in 5x loading buffer. The mixture was boiled for 6 minutes at 100°C and then cooled on ice.

A SDS-PAGE was prepared by polymerising a stacking gel solution on top of a 12% resolving gel solution in a glass cassette. The gel was transferred to a Mini-Protean Tetra Cell (#165-800, BIO-RAD) with 1x running buffer. 30µl of each sample was loaded into each well and 5µl of protein

ladder (#26619, Thermo Fisher Scientific) as added to a well for reference. The samples were run at 120V for ~90 minutes for the proteins to separate. A Trans-blot Turbo semi-dry transfer machine and kit (17001915, Bio-rad) was used to transfer the proteins from the gel onto a nitrocellulose membrane (high MW setting). The membrane was blocked in 5% skimmed milk for 60 minutes at RT. The membrane was probed with an anti-His tag antibody (#ab18184, Abcam, 1:1000) and a β-actin mouse monoclonal antibody (#ab8224, Abcam, 1:1000) diluted in TTBS and incubated overnight at 4°C on a rocker. The following day, the membrane was washed tree times in TTBS for 20 minutes per wash and probed with a LI-COR secondary antibody (800CW Donkey anti-Mouse IgG Secondary Antibody, IRDye, 1:10,000 in TTBS). The membrane was incubated for 1 hour at RT on a rocker then washed with TTBS trice for 20 minutes per wash and scanned using a Li-cor Odyssey CLx. The membrane was stored in TBS at 4 °C for up to one week.

Table 2.15 Reagents used in western blot

5x Loading buffer	
Reagent	**Quantity**
312mM Tris HCL (pH6.8)	11ml
25% (v/v) glycerol	5ml
0.015% (w/v) bromophenol blue	0.3g
10% (w/v) SDS	2g
10% (v/v) β-mercaptoethanol	2ml

10x TBS	
Reagent	**Quantity**
Tris	18.2g
NaCl	87g
pH 8.0	
dH_2O up to 1L	

1x TTBS	
Reagent	**Quantity**
10x TBS	100ml
Tween20	1ml
dH_2O up to 1L	

10x Running buffer	
Reagent	**Quantity**
SDS	10g
Glycine	145g
Tris	30g
dH_2O up to 1L	
pH 8.3	

12% Resolving gel	
Reagent	**Volume**
1M Tris pH 8.8	3.75ml
30% (V/V) acrylamide	3.6ml
20% (W/V) SDS	50µl
TEMED	12.5µl
dH_2O	2.3ml
10% (W/V) APS	300µl

Stacking gel	
Reagent	**Volume**
1M Tris pH 6.8	625µl
30% (V/V) acrylamide	900µl
20% (W/V) SDS	25µl
TEMED	12.5µl
dH_2O	3.29ml
10% (W/V) APS	150µl

2.2.21 SDS-page

To confirm the presence of different proteins in the cell lysate and the conditioned medium after the transfection, an SDS-page was performed. A 12% SDS-PAGE was prepared, and gel electrophoresis was performed as described in Section 2.2.20. The SDS-PAGE was incubated with a Coomassie brilliant blue solution (Table 2.16) for 30 minutes at RT on a rocker. The dye was then removed, and the gel was washed overnight on a rocker with a de-stain solution (Table 2.17) at RT.

Table 2.16 Coomassie blue solution

Reagent	Quantity
Coomassie blue (National Diagnostics, #R-250)	1g
Glacial acetic acid (Fisher Scientific, #A35-500)	50ml
Methanol (VWR)	225ml
dH_2O	225ml

Table 2.17 Components of de-stain solution

Reagent	Volume
Methanol (VWR)	50ml
Glacial acetic acid (Fisher Scientific)	50ml
dH_2O	400ml

2.2.22 Protein column purification

Gravity column purification kit (#635654, Takara Bio company) was used for isolating the His-tagged periostin derived from the plasmid DNA. Since Periostin is an extracellular matrix protein

that is secreted by cells, the conditioned medium from the transfected cells was used in this experiment. Following the manufacturer's protocol, the conditioned medium was loaded onto the plastic 5ml gravity columns with affinity to bind His-tag, washed with Equilibration buffer, Wash buffer and then eluted in Elution buffer. The elute was collected in 1ml fractions and each ml of solution was kept in seperate Eppendorf tubes and their concentration was monitored using a BioDrop.

2.2.23 Protein extraction from cloned *E.coli*

Cloned *E.coli* colonies were cultured overnight as described in Section 2.2.7. The next day, 1ml of this starting culture was transferred to a falcon tube with 10ml of LB broth and incubated until it reached an OD600 of ~0.6 (~4 hours). To induce protein expression, 10µl of 0.1 mM isopropyl-β-D-thiogalactopyranoside (IPTG) (#15529019, Fisher scientific) was added and the culture was incubated overnight at 18°C on a shaking incubator. After 20 hours of incubation, the culture was centrifugation at 4000*g* for 10 minutes. The LB broth was flash-frozen in liquid nitrogen and stored at -80 ^{o}C.

Lysates were also prepared from these *E.coli* cultures. The *E.coli* cells were treated with Tractor Buffer ((#635654, Takara Bio company) in order to release the proteins. After centrifugation, the supernatant was passed through the gravity columns for purification as described in Section 2.2.22.

2.2.24 POSTN 3D structure

To determine the areas of the periostin proteins that the alternative exons encode, the 3D structure of POSTN protein needed to be predicted. First, the protein sequence of the periostin protein containing all the exons was obtained from Uniprot (POSTN-203, Uniprot identifier: Q15063-1). Then, the Alphafold database (https://alphafold.ebi.ac.uk/entry/Q15063) was used to visualise the 3D tertiary structure of POSTN-203. The UniProt data base (https://www.uniprot.org/uniprot/Q15063#Q15063-5) was then used to determine the protein sequence that is encoded by each alternative exon. Finally, these sequences were compared against the sequence from the POSTN-203. When a part of the sequence is selected on UniProt, the corresponding area on the 3D POSTN-203 structure is highlighted. In this way, it was possible to detect the regions of the periostin protein that are encoded by the alternatively spliced exons.

3 RESULTS

3.1 AMPLIFICATION OF *POSTN* GENE FOR CLONING

To amplify *POSTN*, cDNA was synthesised from PS1 cells. A previous study in our lab found these stromal pancreatic cells to express periostin (Rahim *et al*. 2019, unpublished data). To validate the cDNA synthesis, an initial RT-PCR reaction was performed using routinely used primers to amplify the *QARS* gene. The products were run on a 1% agarose gel with 100bp ladder (Fig. 3.1). Bands were detected at the expected product size for these primers at 150-200bp. The negative controls, with no reverse transcriptase (NRT) and no template control (NTC) showed no bands, confirming that there was no contamination during preparation. Lastly the concentration and the purity of the synthesised cDNA was measured using a BioDrop µLite Spectrophotometer at 454.4 mg/ml with the A260/230 and A260/280 reading at 2.193 and 2.112 respectively, validating the integrity of the cDNA synthesised.

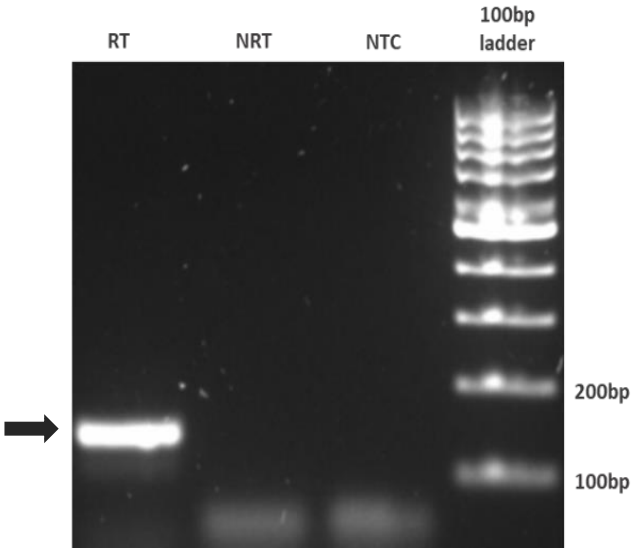


Figure 3.1: Validation of cDNA synthesis using QARS primers. The samples with reverse transcriptase (RT), no reverse transcriptase (NRT) and no template control (NTC) were loaded on 2% agarose gel with 100bp ladder. QARS primers give a roughly 150-200 bp product. The arrow shows the band detected at around 150 bp.

The RT-PCR reaction for amplification of *POSTN* was subsequently tested. cDNA, with reverse transcriptase (RT) and QARS primers was used as a positive control for the reaction. NRT, NTC and dH_2O with both QARS and POSTN primers (Fig. 3.2) were used as negative controls for the reaction. The agarose gel image showed no bands, indicating that there was an issue with the PCR process and amplification of *POSTN*.

The cDNA quality had been confirmed with the QARS primers in previous PCR reaction. However, the DNA polymerase kit that was used had been thawed several times which can have compromise performance. The experiment was therefore repeated with a new kit. The amplification of *POSTN* was confirmed by observation of a strong band at 2000-2500bp when the sample was run on a 1% agarose gel (Fig. 3.3). The thickness of the band and the multiple bands within it can be explained by the presence of multiple transcript variants of similar size in the product. Due to the close proximity of the bands, gel extraction of individual transcript variants would not be feasible. Thus, cloning techniques will be used to overcome this.

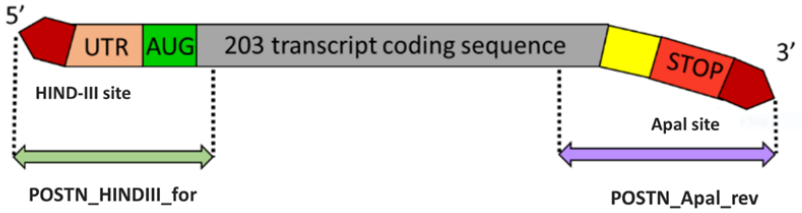


Figure 3.2: Diagram of cloning primers used to amplify *POSTN* gene. The forward primer (POSTN-Hind-III) is located at the 5' of the sequence while the reverse primer (POSTN-ApaI) binds in the 3' end of sequence. These primers amplify the open reading frame (ORF) of POSTN, required for cloning. Image modified from Rahim *et al.,* (2019), unpublished data.

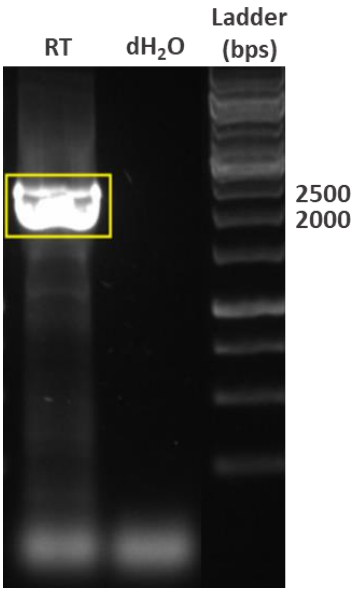


Figure 3.3: Repeating the RT-PCR technique for amplifying POSTN DNA using a new kit. A new DNA polymerase kit was used to amplify *POSTN*. The cDNA was replaced with dH_2O to serve as a no template control in the second well. A thick band was detected between 2000bp and 2500bp in the well containing PS1 cDNA (yellow box). Empty wells were removed from the image.

3.2 PLASMID EXTRACTION

Plasmids (pcDNA4-TO-Puromycin-mVenus-MAP) were extracted from *E.coli* and used as vectors for the cloning process. *E.coli* from a glycerol stock was cultured overnight in LB broth and the cloudiness of the broth indicated bacterial growth. The plasmids were extracted as described in Chapter 2 sub-sections 2.2.2-2.2.8. The products were run on a 1% agarose gel to confirm plasmid extraction (Fig. 3.4). The expected plasmid size was approximately 6300bp. However, a band size of 3500bp was observed consistently in two experimental replicates. These bands were nevertheless still considered to be the pcDNA4-TO-Puromycin-mVenus-MAP plasmids, as when supercoiled, plasmids run faster on an agarose gel. This has been previously observed in our laboratory (Rahim *et al*. 2019, unpublished data). Faint bands of ~8000bp were also detected. These could have been formed due to nicked plasmids that migrate slower in the gel than normal plasmids as one of the strands is relaxed.

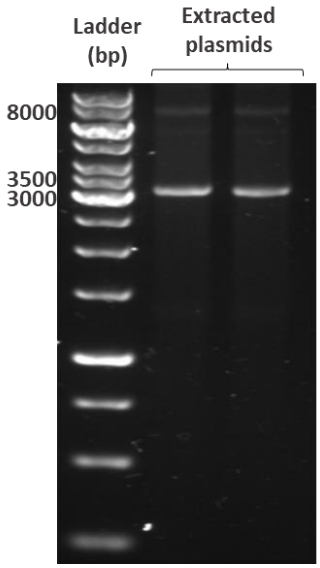


Figure 3.4: Observing purified plasmids extracted from *E.coli*. Plasmids were extracted from *E.coli*, purified and run on a 1% agarose gel. Two strong bands were observed at ~3500bp and two faint bands at ~8000bp. N=2 from 2 independent plasmid extractions.

3.3 ENZYME DIGESTION OF PLASMID AND PCR PRODUCT

Both the vector (plasmid) and insert (PCR product with restriction sites) were digested with the same enzymes to be prepared for ligation. Double digestion with Hind-III and Apal enzymes was performed to minimize self-ligation and ensure correct orientation of the plasmids. To obtain high quality products, digested plasmids and PCR product were purified to remove any reagents and impurities left from digestion.

3.3.1 Digested plasmid agarose gel extraction Plasmid

Plasmids were purified by agarose gel extraction. The bands containing the plasmids were cut out from the gel for purification (Fig. 3.5).

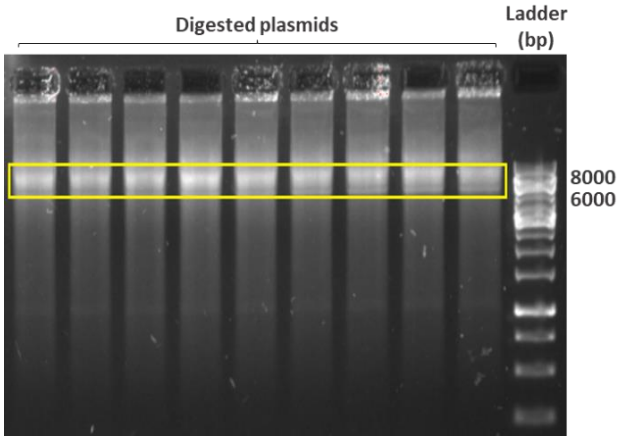


Figure 3.5: Image of the agarose gel prior to plasmid extraction during the plasmid purification process. All 9 wells contained digested plasmids samples to increase the quantity of plasmids extracted. The area of the gel cut out for purification is highlighted in the yellow box.

The purified plasmids were subsequently run on a 1% agarose gel alongside the undigested plasmids to confirm digestion. The success of the digestion was shown by the slower migration of the non-supercoiled plasmid. Figure 3.6 shows the digested plasmids at ~7000bp and the undigested plasmids at 3500bp.

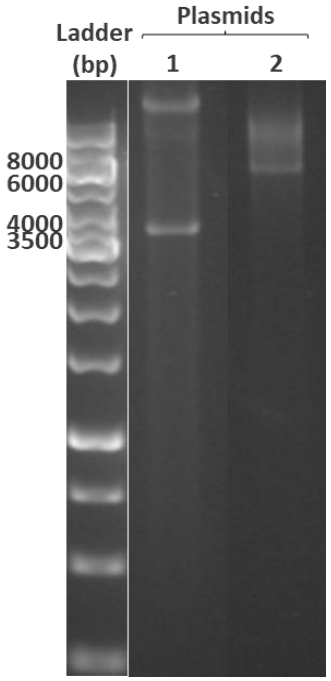


Figure 3.6: Extracted/undigested plasmids and digested plasmids. Undigested plasmids (lane 1) and digested and gel purified plasmids (lane 2) were run on a 1% agarose gel. Bands were observed at ~3500bp in the undigested sample and at ~7000bp for the digested sample. Image has been cropped to exclude empty wells.

3.3.2 Digested PCR product purification

A PCR and DNA clean-up kit was used to remove any residues from the enzyme digests to increase the purity of the PCR product. The PCR product containing the full *POSTN*, that was amplified using the primers shown on Fig. 3.2, was used. The protocol from the manufacturer was followed and the purified product was run on an agarose gel. An undigested sample was not run since the difference in size would be too small to be detected by agarose gel electrophoresis. Figure 3.7 shows image of the agarose gel with the purified digested POSTN suggesting successful purification.

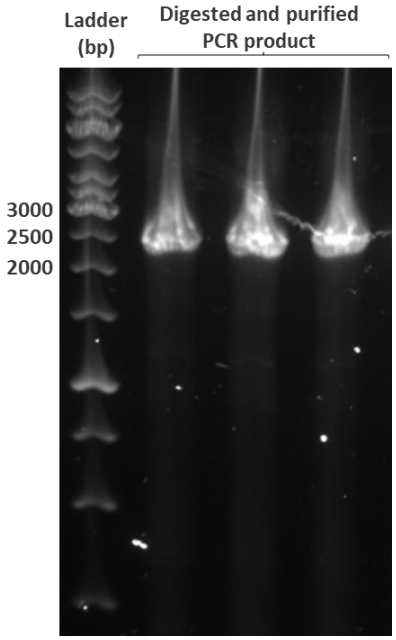


Figure 3.7: Confirming the purification of the digested PCR product using a 1% agarose gel. Purified digested PCR product was run in triplicate on an agarose gel. Periostin bands were detected at ~2500bp. Smearing was observed in the image due to puncturing of the agarose gel during loading.

3.4 LIGATION AND TRANSFORMATION OF PLASMID IN *E.COLI* CELLS

The vector and the insert were ligated and kept on ice before the transformation process. To ensure the transformation of ligated plasmids, one positive and two negative controls were used. The first negative control contained *E.coli* without any plasmids to detect any potential contamination during the experiment and/or inability of the agar plates to inhibit growth of non-ampicillin resistant bacteria. In the second negative control, *E.coli* were transformed with ligated sample that was lacking the insert. If positive, this would be an indicator of plasmid self-ligation. The positive control contained undigested plasmids and this would confirm the transformation of the plasmids in *E.coli*.

Following the overnight culture of transformed *E.coli* on agar plates, images of the plates were captured and analysed (Fig. 3.8). The plates of the transformed *E.coli* with ligated plasmids and the undigested plasmids controls had bacterial growth. However, both negative controls also showed excessive bacterial growth. The negative control that contained ligated plasmid with no insert showed growth and transformation indicating self-ligation of the plasmid and/or the insufficient enzyme digestion.

The *E.coli* transformed without plasmids also showed bacterial growth. This could indicate that an insufficient quantity of ampicillin was added to plate or the ampicillin has been inactivation if the agar had not cooled down sufficiently when the ampicillin was mixed into the solution.

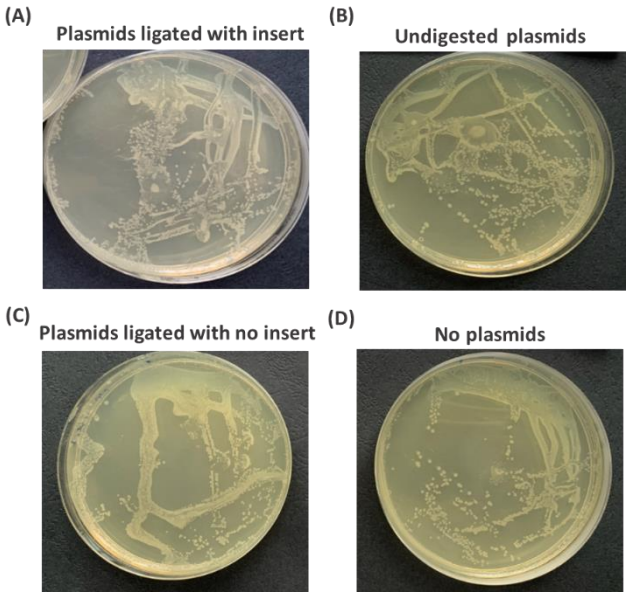


Figure 3.8: Determining transformation of the *E.coli* with ligated plasmid. Transformed *E.coli* were cultured overnight on agar plates to test the transformation of ligated plasmids. *E.coli* were transformed with **(A)** ligated plasmids with insert, **(B)** undigested plasmids [positive control], **(C)** plasmids with no insert [negative control] or **(D)** no plasmid [negative control].

This experiment was therefore repeated adding new ampicillin to the agar pre-cooled down to 45°C to ensure that no deactivation of the antibiotic would occur. A test run was first conducted, culturing untreated *E.coli* and the same 3 controls as before (Fig. 3.9). Again, all the samples including the negative controls had bacterial growth. The presence of bacterial growth in the untreated *E.coli* culture helped to rule out the possibility of any cross-contamination of the samples with ligated plasmids during transformation. Collectively, these results indicated that insufficient ampicillin performance was not the cause of the bacterial growth since a new stock of the antibiotic was used and it was not deactivated. Thus, further investigation was required.

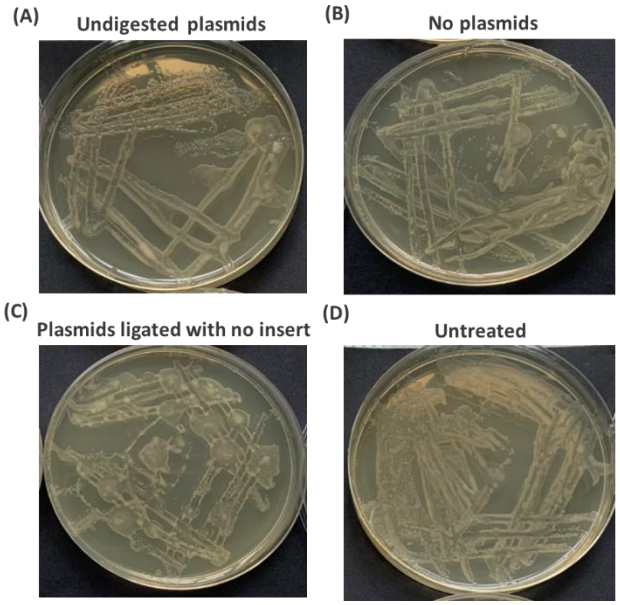


Figure 3.9: Optimising transformation by testing ampicillin performance. Transformed *E.coli* were cultured overnight on agar plates to test re-test the transformation of ligated plasmids. *E.coli* were transformed with **(A)** undigested plasmids [positive control], **(B)** no plasmids [negative control], **(C)** plasmids with no insert [negative control] or **(D)** left untreated [negative control].

Since the issue was not with plate preparation, a new *E.coli* stock was thawed and transformed. The experiment was repeated as described previously but only the controls were plated (Fig 3.10). The cultures of *E.coli* with no plasmids showed no growth and the *E.coli* transformed with plasmids with on insert had a minor growth. The positive control with the undigested plasmids had several colonies. Based on these findings, the issue was with the previous *E.coli* cell stock.

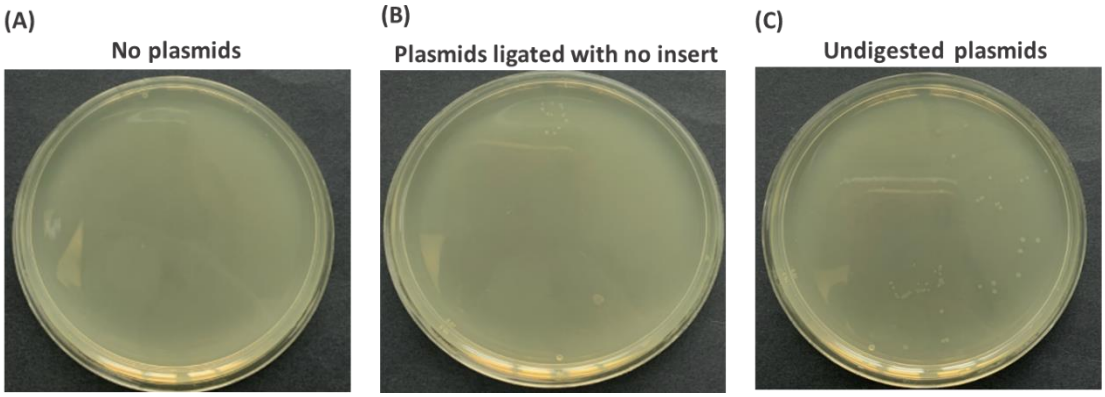


Figure 3.10: Testing a new *E.coli* cell stock. A new cell stock *E.coli* were transformed and cultured overnight on agar plates to test the transformation of ligated plasmids. *E.coli* cultured with **(A)** no plasmids [negative control] **(B)** no insert [negative control] and **(C)** undigested plasmids [positive control].

As the issue was now resolved, a full experiment was repeated with the new *E.coli* stock as described in Section 2.2.13. As shown in Figure 3.11, there was growth of several single colonies of *E.coli* transfected with ligated plasmids. However, there was an intermediate growth of bacteria on the culture containing the ligated plasmids with no insert. It is acceptable to have some minor growth as a small number of plasmids are expected to self-ligate. However, the number of the colonies present were higher than expected indicating insufficient enzyme digestion.

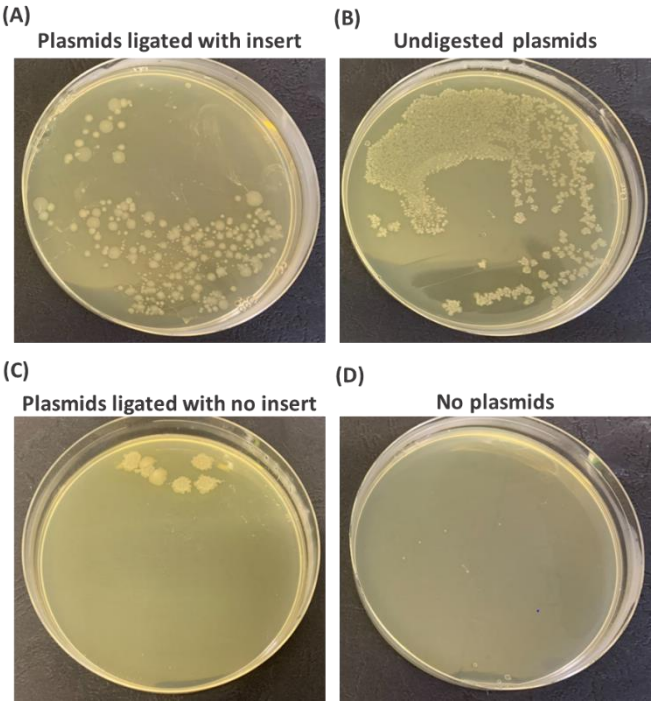


Figure 3.11: Determining transformation of the *E.coli* with ligated plasmid with new *E.coli* stock. Newly transformed *E.coli* were cultured overnight on agar plates to test the transformation of ligated plasmids. *E.coli* were transformed with **(A)** ligated plasmids with insert, **(B)** undigested plasmids [positive control], **(C)** plasmids with no insert [negative control] or **(D)** no plasmid [negative control].

The issue seemed to have arisen during the agarose gel purification of the digested plasmids (Section 3.3.1, Fig. 3.5). The aim of the gel purification was to extract solely the digested plasmid, separating it from undigested plasmids and any reagents of the reaction. As seen on Figure 3.5, the bands extracted were blurry. That could indicate that the extracted bands contained both the digested and undigested plasmids. To correct this, a new agarose gel purification was performed where the samples were electroporated longer to ensure further separation of the bands in the gel (Fig. 3.12). The ligation and transformation steps were repeated with the new plasmids. Following overnight culture, images of the plates were taken (Fig. 3.13). Both the negative controls had no growth, and several colonies were present in the plates with the ligated plasmids with insert. The transformation was considered successful, and the colonies were kept for colony PCR.

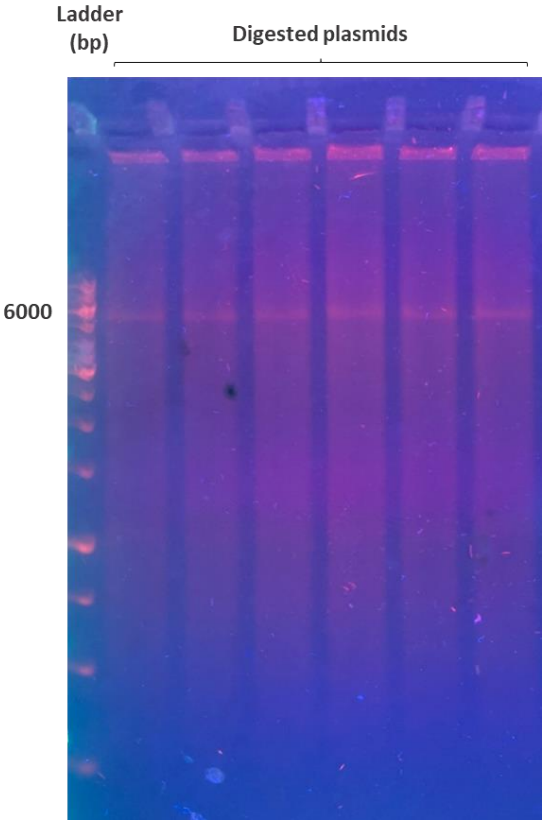


Figure 3.12: Digested plasmids run on an agarose gel for purification. Six wells were loaded with digested plasmids for agarose gel purification. The agarose gel was imaged under UV light and bands were detected at ~6000bp.

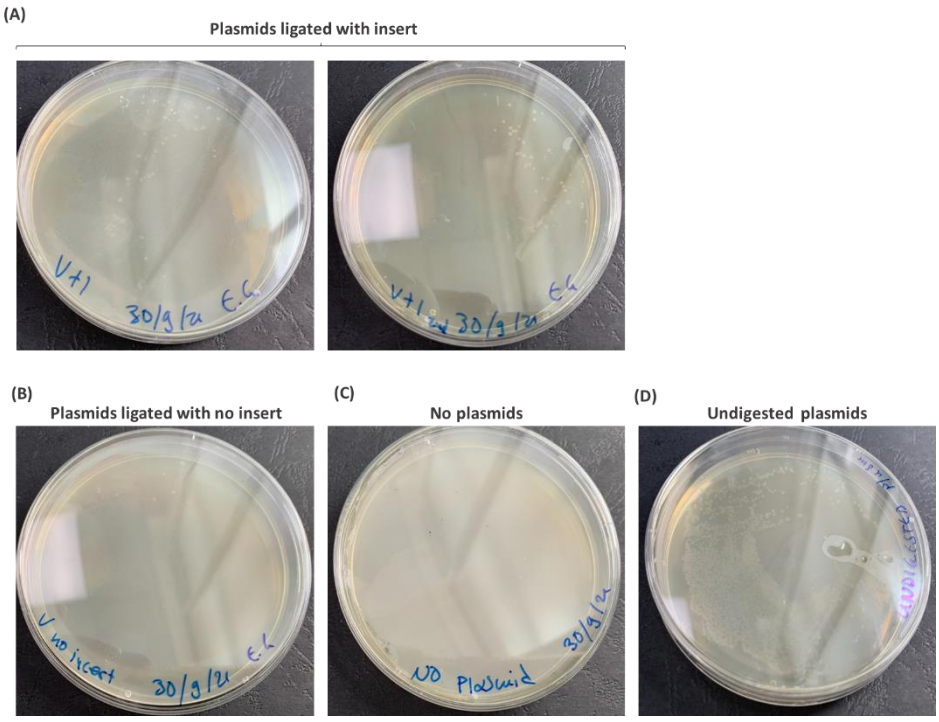


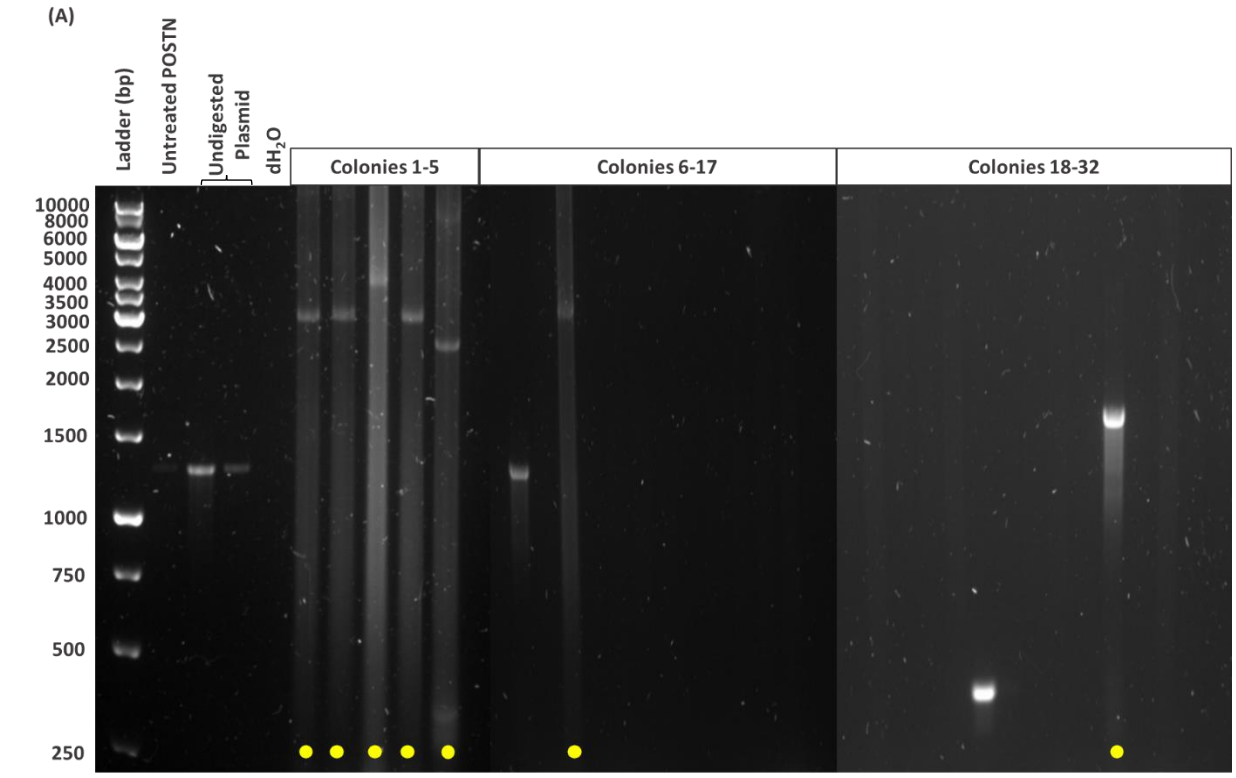
Figure 3.13: Determining transformation of the *E.coli*. *E.coli* were cultured overnight on agar plates to test the transformation of ligated plasmids. *E.coli* were transformed with **(A)** plasmids ligated with insert in duplicate, **(B)** plasmids with no insert [negative control] **(C)** no plasmids [negative control], or **(D)** undigested plasmid [positive control].

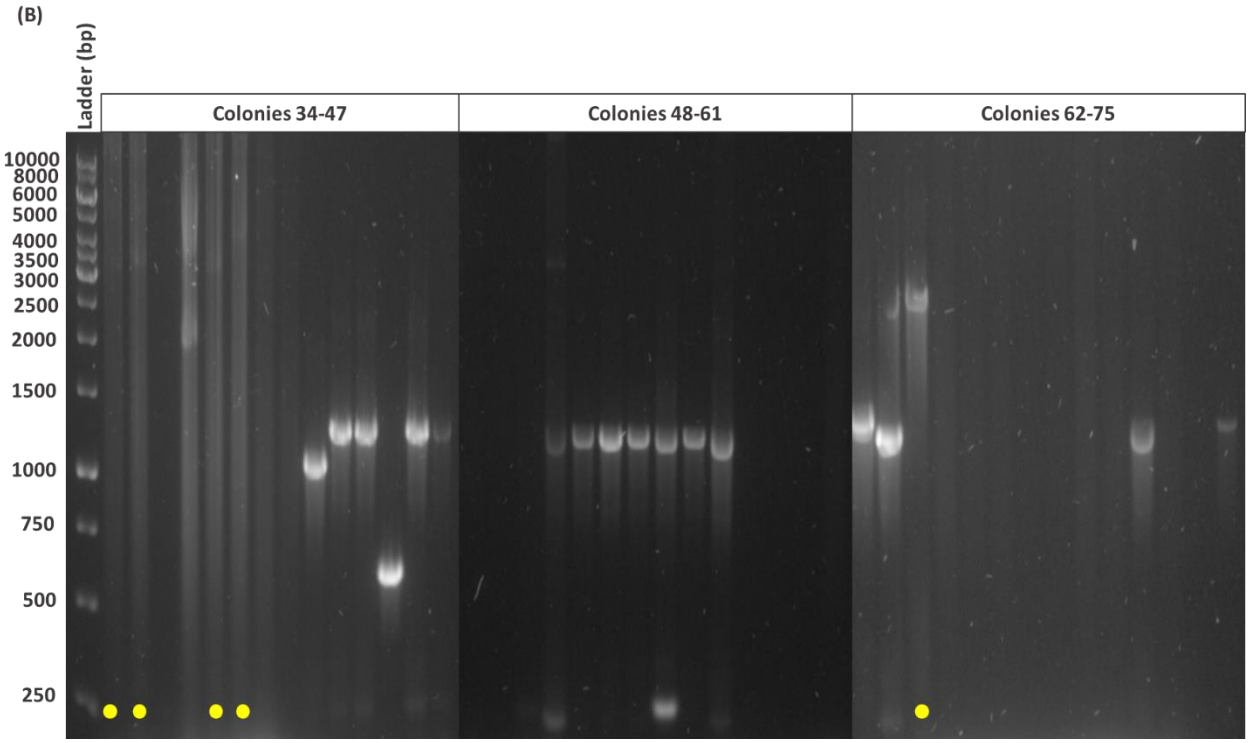
3.5 EXTRACTION OF *POSTN* TRANSCRIPT VARIANTS / COLONY PCR

Single colonies were selected from the agar plates with ligated plasmids. The colonies were added directly into PCR master-mix and PCR with the cloning primers was performed. This reaction would indicate the presence of plasmids containing the insert. If the colony only contained self-ligated plasmids, then there would be amplification of a small PCR product (Fig. 3.14). The PCR products were run on a 1% agarose gel with a 1kb ladder, and an image was taken to assess the presence of the insert (Fig. 3.15).

The expected size of the product is around 3000-3500bp, with exact size depending on the transcript variant amplified. Products with size of 1000-1500bp were considered as not successfully cloned as that is the expected size of an empty vector cloning site. Some PCR products had multiple bands when run on the gel which could indicate the presence of multiple plasmids within a colony. To investigate further, some samples with multiple bands were sent for sequencing. The samples with size between 2500-3500bp were considered most likely to express the insert and chosen for sequencing PCR along with some others as indicated on Table 3.1. The samples that were sent for sequencing were not gel purified. Consequently, if more than one variant was amplified, they would be picked up by the subsequent Sanger sequencing. A summary of the results of the colony PCRs are show in Table 3.1.

Figure 3.14. Cloning primers amplification pattern. The cloning primers initiate amplification from the ends of the plasmid where the insert should be ligated. If the ligation was successful, the primers would amplify a small region of the plasmid and the full insert/ POSTN isoform. If the plasmid was self-ligated, then the primers would amplify a small region of the plasmid.


(A)
Ladder (bp)
Untreated POSTN
Undigested Plasmid
dH2O
Colonies 1-5
Colonies 6-17
Colonies 18-32
10000
8000
6000
5000
4000
3500
3000
2500
2000
1500
1000
750
500
250


(B)
Ladder (bp)
Colonies 34-47
Colonies 48-61
Colonies 62-75
10000
8000
6000
5000
4000
3500
3000
2500
2000
1500
1000
750
500
250

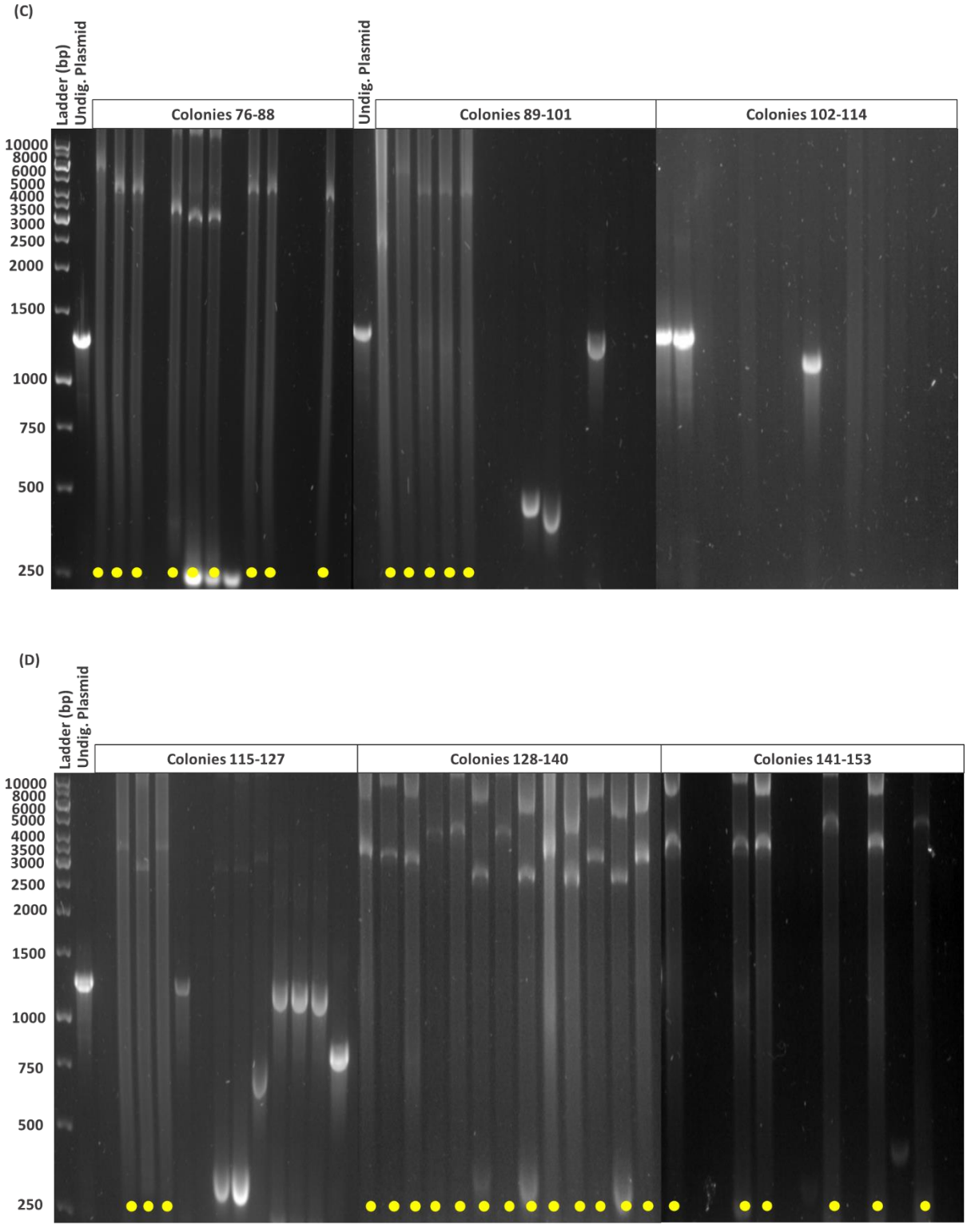


Figure 3.15: PCR of colonies extracted from the overnight culture of transformed *E.coli*. The 153 colony PCR samples were run on 1% agarose gels and images were taken. Each image shown are 3 gels combined with a 1kba ladder. Image shows **(A)** samples 1-32 with controls (undigested periostin and dH_2O) and two wells of undigested plasmids as size reference **(B)** samples 34-75. PCR product from colony 33 was not run on the gel **(C)** samples 76-114 with two undigested plasmid controls **(D)** samples 115-153 with an undigested plasmid control. The colonies that were chosen for sequencing PCR are indicated with a yellow dot on the bottom of the images.

Table 3.1: Table summarising *E.coli* colony PCR results				
Agarose gels (Fig. 3.13)	**Multiple bands present**	**No bands**	**Small sized bands**	**Colonies chosen for sequencing PCR**
A	-	6, 8, 10-22, 24-27 and 29-32	7 and 23	1-5, 9 and 28
B	-	36, 37, 40, 41, 48, 49, 58-61, 65-71, 73 and 74	42-47, 50-57, 62, 63, 72 and 75	34, 35, 38, 39 and 64
C	80-82	79, 83, 86, 87, 94, 95, 98, 100, 101, 104-108 and 110-114	96, 97, 99, 102, 103 and 109	76-78, 80-82, 84, 85 and 88-93
D	128-135, 137-141, 144, 145, 148 and 150	115, 120, 142, 143, 146, 147, 149, 151 and 153	119 and 121-127	116-118, 128-141, 144, 145, 148, 150 and 152
Total number	20	80	34	38

3.6 SEQUENCING RT-PCR

The PCR products from the colonies were used as templates for sequencing PCR. This step was used to amplify only the alternative splicing region of *POSTN* to reduce the product size optimum for Sanger sequencing (Fig. 3.16). The *POSTN* PCR products synthesised in Section 2.2.2 was used as positive control for the reaction and indicated where the sequencing PCR products should be present on the agarose gel. All samples had multiple bands when run on an agarose gel (Fig. 3.17). This could indicate either a nonspecific amplification or the presence of multiple transcript variants within the same sample. Some samples had additional bands of round 1500bp. These bands were considered as products from the colony PCR (that was used as template for sequencing PCR) that had some amplification during the second PCR.

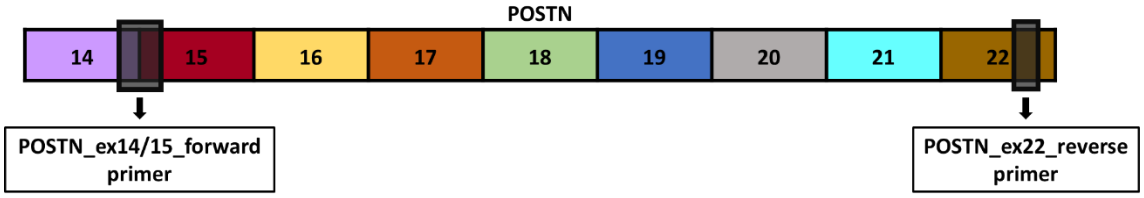


Figure 3.16: POSTN target region of sequencing RT-PCR. The primers used in this PCR reaction were POSTN_ex14/15_forward and POSTN_ex22_reverse primers. The exons of POSTN are numbered and shown in different colour. The binding sites of these primers on the POSTN gene are illustrated by black coloured boxes. By using these primers only, the alternatively spliced region was amplified, resulting in a small PCR product, suitable for sequencing.

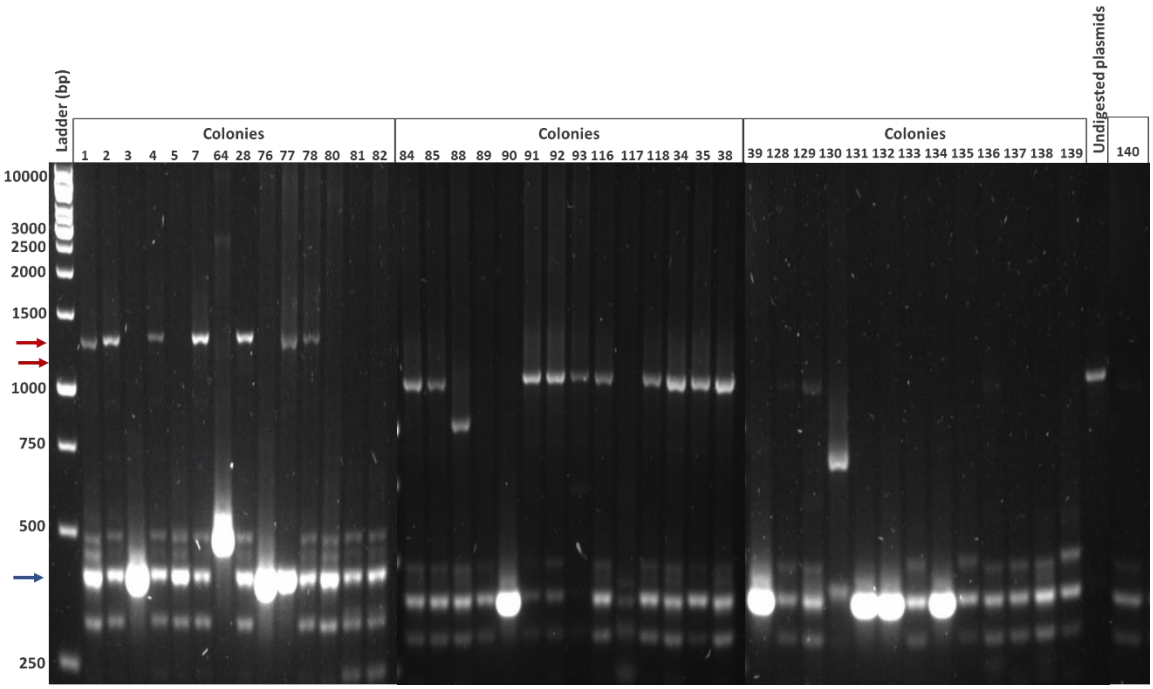


Figure 3.17: Sequencing PCR of colonies chosen to be sent for sequencing. Some transformed *E.coli* that underwent colony PCR was prepared for subsequent sequencing PCR (details in Table 3.1). These samples were run on an 1% agarose gels to confirm amplification. Samples 141-152 were not run on the gel and are not shown on the image. Red arrows indicate nonspecific amplification in some samples and the blue arrow indicates the amplified alternatively sliced region of periostin. PCR product with undigested plasmids was run as negative control. Images from 3 gels were combined to make the above image.

3.7 SEQUENCING ANALYSIS

To determine the transcript variant clones in each colony, the samples from the Section 3.6 were sent for sequencing. The results were compared against one exon at a time in order to determine the exons expressed in each sample. Some samples were flagged as they were considered mixed colonies and were not analysed further. The remaining samples were analysed, however the sample quality was low. In some cases, when the same sample was sequenced with both the forward and reverse primers the results were different. Furthermore, the majority of the sequences had low quality sequence on the area where exon 17 or 18 would be present, making it unclear which exon was present. This information suggests that mixed were present indicating that the *E.coli* could have been transformed with two or more plasmids simultaneously giving the mixed results. Therefore, these results were considered unreliable (data not shown).

The *E.coli* to plasmid ratio could be the source of these errors. The number of plasmids was reduced to half of what was used the initial experiment and the *E.coli* cell number was kept the same. This would dilute the plasmid concentration to prevent multiple plasmids from entering a single cell. No other parameters were changed, and the experiment was performed as described in Chapter 2.2.13.

Twenty-eight new colonies were selected. They were all used for colony PCR and then the alternatively spliced region of the cloned *POSTN* was amplified by sequencing PCR. The samples were run on agarose gel to confirm the amplification before sequencing (Fig. 3.18).

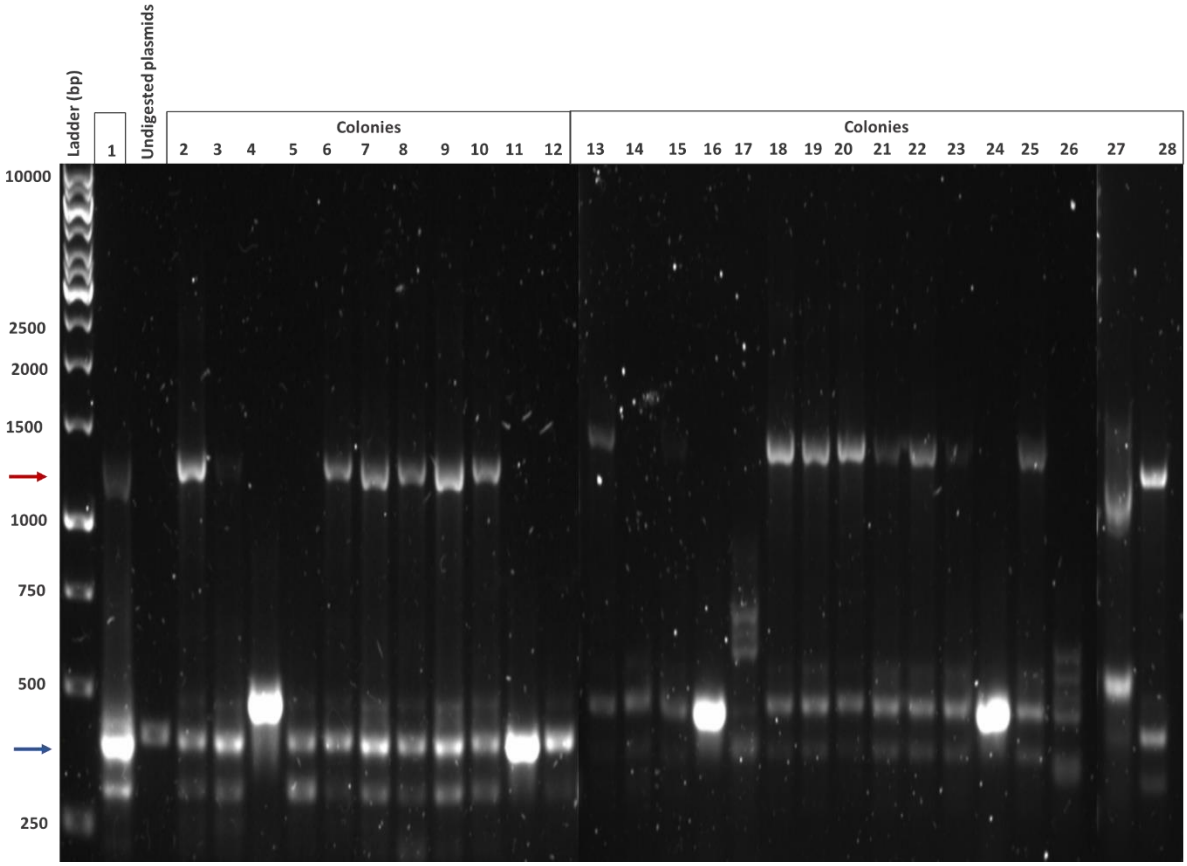


Figure 3.18: Preparing colonies expressing the cloned periostin to repeat sequencing PCR. Samples expressing cloned periostin were run on a 1% agarose gels to confirm amplification. The red arrow indicates nonspecific amplification in some samples and the blue arrow indicates the amplified alternatively sliced region of periostin. PCR product with undigested plasmids was run as a negative control.

The results of the sequencing were subsequently analysed. Some of the samples may have contained mixed colonies and these were excluded from the analysis. The other samples had better quality sequence compared to the previous experiment. The individual exons were checked against the sequence to determine the exons expressed in each sample. Then the *POSTN* variants expressed in each sample were determined. The samples of interest were sent again for sequencing with a different primer to ensure that the sequence was clear with both the forward and reverse primers.

After detecting the exons present in each sequence, this data was compared with the known transcript variants of *POSTN* to determine the isoform cloned in each colony. Table 3.2 shows the information used to determine the variants expressed in the colonies and Figure 3.19 shows a

diagram of the exon assembly of the transcript variants. From all the samples analysed, there was no variant expressing exon 17. On the other hand, exon 19 was strongly expressed and present in every sample. Only three different variants were found to be expressed. In particular, the variants named 202 and 209 on Ensembl were detected. The third variant detected was not found on Ensembl at the time of this study, however it had previously been recorded in the database under the name 201b (Rahim *et al*. 2019, unpublished data). This variant can still be found in NCBI with the ID number 4 (Table 3.2) however it will be referred to as 201b throughout this study. Two colonies were expressing variant POSTN-202, one was expressing variant POSTN-201b and nine were expressing variant POSTN-209. Only colony 3 (expressing variant 202), colony 12 (expressing variant 201b) and colony 13 (expressing variant 209) were chosen to be used for the extraction of the plasmids as they had clearer sequences. The sequencing of these three colonies is shown in the Appendix (Fig. 7.1). The sequences of the remaining samples can be found in the Appendix (Fig. 7.2.).

Table 3.2 Information related to known transcript variant of *POSTN*

Transcript variant name on Ensembl	Transcript variant name on NCBI	Isoform ID	Exons expressed	Expected size of the region from exon 15 to 22 (bps)	Variant size Ensembl and Genebank (bps)
201	2	NM_001135934.2	15, 16, 19, 20, 21, 22	408	3071
-	4	NM_001135936.2	15, 16, 19, 20, 22	324	3046
202	5	NM_001286665.2	15, 16, 18, 19, 20, 21, 22	498	3196
203	1	NM_006475.3	15, 16, 17, 18, 19, 20, 21, 22	579	3301
204	8	NM_001330517.2	15, 16, 17, 18, 19, 20, 22	495	3210
209	3	NM_001135935.2	15, 16, 18, 19, 20, 22	414	3214
210	6	NM_001286666.2	15, 16, 20, 21, 22	318	3113

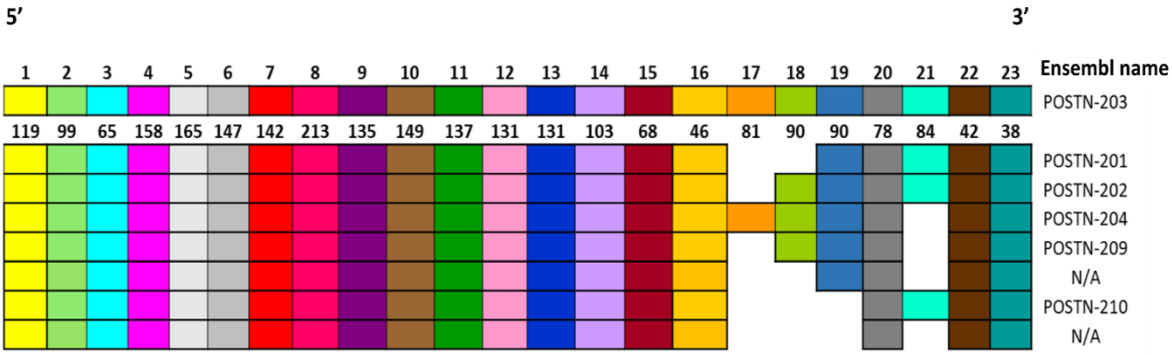


Figure 3.19: Exon structure of POSTN alternative transcript variants. The exons can be distinguished by different colours. The numbers at the top of the POSTN-203 variant indicate the number of the exon and the numbers under indicate the of length (nucleotides) of each exon. These variants have a complete coding sequence obtained from Ensembl (accessed: 28/25/2022). The two variants with the 'N/A' indication are variants that are not found in Ensembl but they used to, according to a previous student and are still found in NCBI (accessed: 28/05/2022). Image adapted from: *Rahim* et al*., (2019)* unpublished data.

3.8 PLASMID EXTRACTION FROM COLONIES OF INTEREST

The three colonies (3, 12 and 13) expressing the periostin transcript variants 202, 201b and 209 were cultured overnight. The plasmids were then extracted with plasmid prep kit compatible for subsequent mammalian cell transfection. The plasmids were extracted as per manufacturer's instructions and the concentration of the plasmids were determined by a BioDrop µLite Spectrophotometer; colony 3: 149 ng/µl, colony 12: 1520 ng/µl and colony 13: 145 ng/µl.

3.9 PLASMID TRANSFECTION INTO HUMAN PANCREATIC STELLATE CELLS (PS1)

PS1 cells were transfected with the plasmids encoding the periostin transcript variants 202, 201b and 209. The cells were cultured for 48 hours to allow them to express the cloned proteins. Cell lysates were prepared and stored at -20°C. As periostin is an extracellular matrix protein it is expected to be secreted by the cells into the culture medium. The conditioned medium was also stored at -80°C for periostin protein purification.

Western blot was performed using an anti-his tag antibody to determine the presence of the periostin protein in the cell lysate. Non-transfected PS1 cells lysate was used as control. An image was captured of the nitrocellulose membrane showed multiple bands (Fig. 3.20). Faint bands were detected at ~60kDa. As the expected size of *POSTN* protein is between 83-93kDa, it was unclear whether this band was *POSTN*. Before repeating the western blot, the transfection was repeated with an increased number of plasmids (200ng) without altering the PS1 cell number. Again, a ~60kDa band was detected by western blot (Fig. 3.21). Thus, further investigation was required.

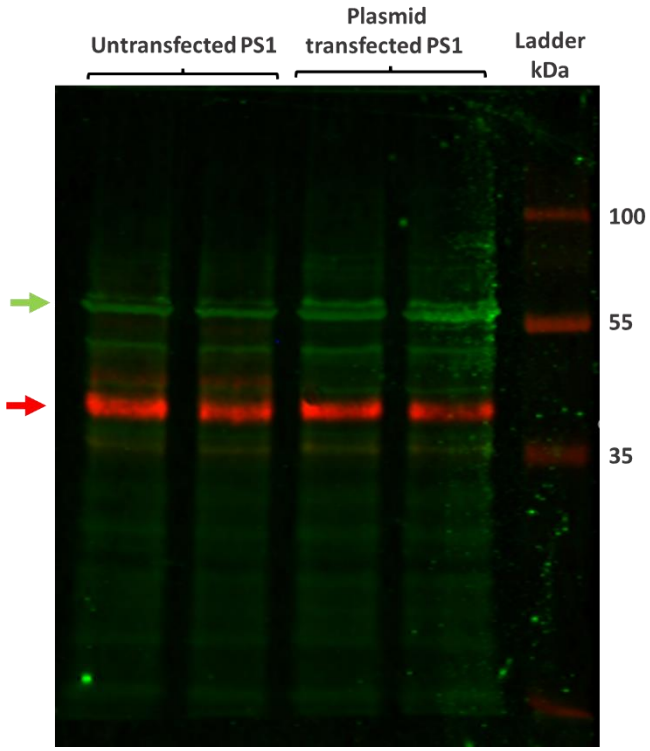


Figure 3.20: Using western blot to determine transfection of the PS1 cells. PS1 cells were transfected with plasmid with the periostin variant 209 and cell lysate was used for western blot. Nitrocellulose membrane was probed with an anti-His tag antibody (green) and β-tubulin (red) loading control antibody.

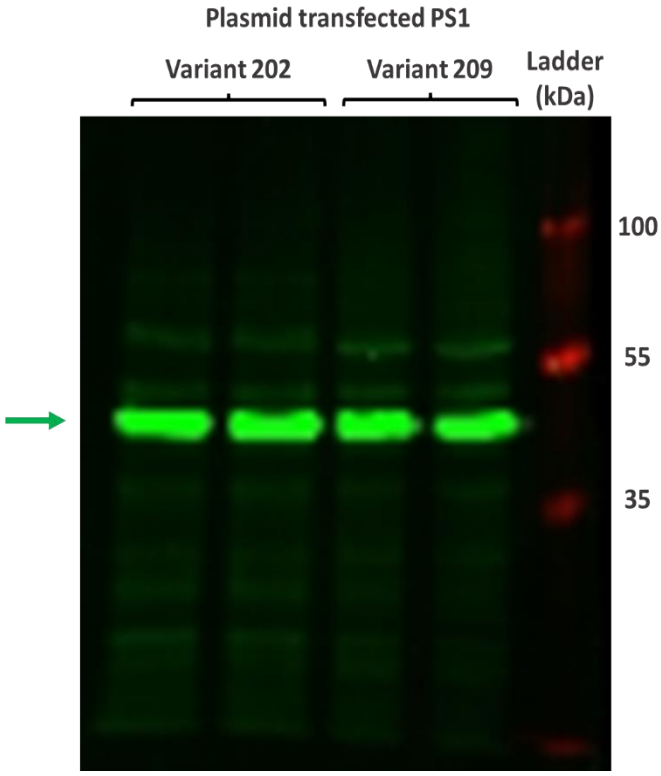


Figure 3.21: Repeating western blot to determine transfection of the PS1 cells. PS1 cells were transfected with plasmid with the periostin variant 202 or 209. Cell lysate was used for western blot. Nitrocellulose membrane was probed with an anti-His tag antibody (green) and β-actin (green) loading control antibody. The green arrow indicates the control bands.

Another western blot was performed, this time with the addition of anti-POSTN antibody as well as the anti-His-tag antibody. This antibody would help determine if the band observed with the anti-His-tag antibody would co-localise with the band detecting POSTN. Multiple bands were detected on the nitrocellulose membrane probed with the anti-POSTN antibody. These bands were found close the expected size of POSTN (80-90KDa). This was expected due to the expression of multiple endogenous transcript variants in the samples. The bands with the strongest signal were ~90kDa however bands were also detected at ~60kDa where the anti-His tag antibody band appeared (Fig. 3.22).

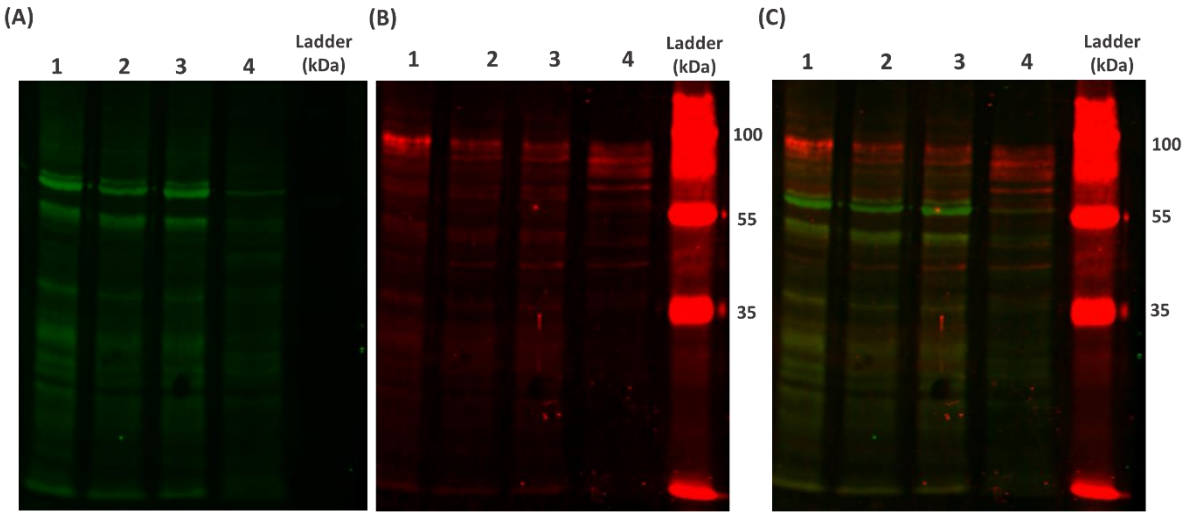


Figure 3.22: Western blot to examine position of bands formed by anti-His tag and anti-POSTN antibody. PS1 cells were transfected with three cloned periostin isoforms. Wells on a SDS PAGE was loaded with cell lysate from clones with isoform [1] 202, [2] 201b, and [3] 209. Well [4] was loaded with conditioned medium from PS1 cells with cloned periostin 202 isoform. A western blot was performed, and the nitrocellulose membrane was probed with an anti-His-tag (green) and periostin (red) antibodies. Images show **(A)** His-tag **(B)** periostin and **(C)** merge.

His-tag protein purification was performed to purify the protein of interest (POSTN) from a mixed sample. After the purification, the samples were used for western blot. No bands were detected from the purified sample by western blot (Fig. 3.23). This suggested that there was no His-tagged protein in the sample or that the protein concentration was too low to be detected by this technique.

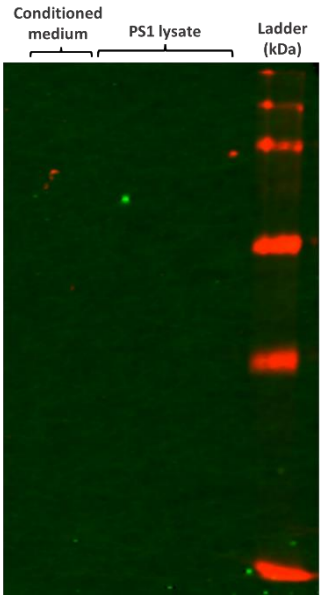


Figure 3.23: Western blot to determine the presence of His-tag POSTN protein in the His-tag purified sample. Both the cell lysate and the condition medium were purified and used for western blot. No positive control was taken. These was no bands detected in either sample.

To determine if the error originated from the PS1 cell transfection, another experiment was performed. The protein at this stage was purified from the cloned *E.coli* that was confirmed to express POSTN. The samples underwent His-tag protein purification. A western blot (Fig. 3.24A) and a Coomassie staining of an SDS-PAGE (Fig. 3.24B) was performed on the purified protein. Bands were not detected using either method, suggesting again that there was no/low protein or an error in the process of detection.

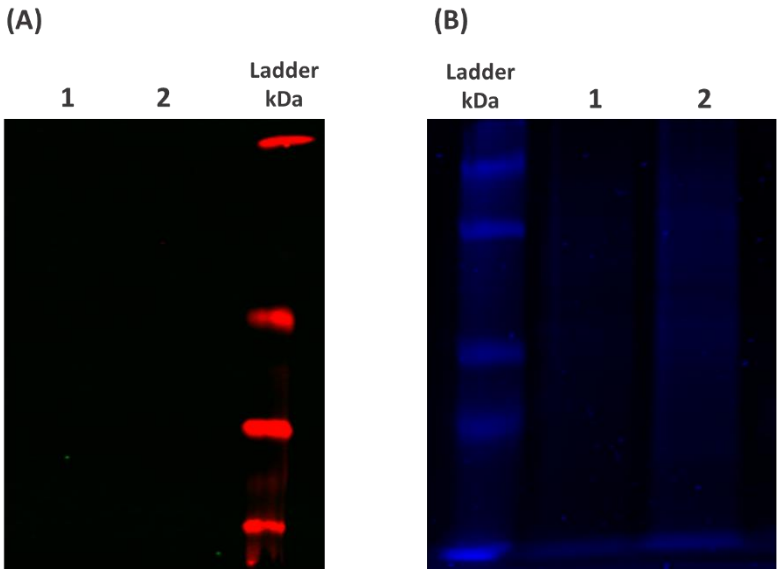


Figure 3.24: Testing for presence of His-tagged periostin in *E.coli* cell lysate and conditioned LB broth. *E.coli* encoded with cloned POSTN-209 isoform and its conditioned LB broth was purified to extract his-tagged periostin. Images show **(A)** western blot probed with anti-His Tag antibody and **(B)** coomsassie stained SDS-PAGE to determine the presence or absence of periostin in the samples. Well [1] contained conditioned LB broth and well [2] the *E.coli* cell lysate.

3.10 PREDICTION OF PROTEIN STRUCTURE

In order to understand how the alternative inclusion or expulsion of exons would affect the structure and potentially the binding abilities of periostin, the exact area of the protein encoded by the alternative exons had to be determined. A protein structure prediction software was used to reveal the position of this area within the 3D structure of the protein and to give an idea of the structural changes that could result from the absence of any of these alternative exon.

Alphafold software was used to predict the structure of periostin. Figure 3.25 (A) shows the periostin protein synthesised from the transcript variant that encodes all the exons (POSTN-203, Uniprot identifier: Q15063-1). Uniprot was used to obtain the FASTA protein sequence of the periostin isoforms. This was used to determine the FASTA sequence of the exons that are alternatively spliced. The parts of the periostin protein that are encoded by the alternatively spliced exon were detected using the Alphafold software (Fig. 3.25B).

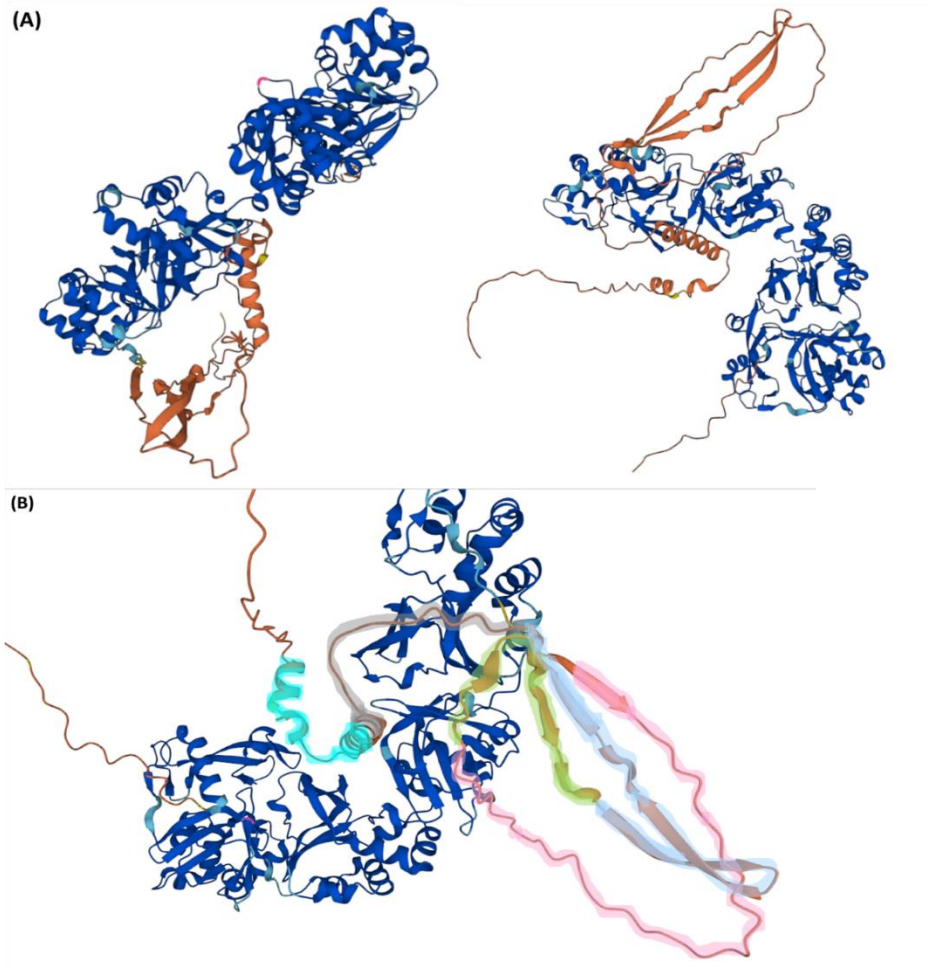


Figure 3.25: Determining the alternatively spliced regions of the periostin protein. (A) Alphafold software was used to draw the 3D protein structure of periostin using the FASTA sequence from Uniprot **(**Uniprot identifier: Q15063-1). Image **(B)** shows structure of the periostin protein with the alternatively spliced regions highlighted. The areas of the protein that are encoded by each alternatively spliced exon are highlighted in the following colours: exon 17-pink, exon 18-green, exon 19-blue, exon 20-grey and exon21- turquoise.

4 DISCUSSION

In recent years, considerable evidence has supported the observation that the transcript variants of periostin hold crucial roles in the progression of various cancer types, including renal, bladder, non-small cell lung and thyroid cancer (Bai *et al.,* 2010; Morra *et al.,* 2012; Morra *et al.,* 2011; Kim *et al.,* 2008). In this study, three periostin isoforms, POSTN-202, POSTN-209 and POSTN-201b as named by Ensembl, were detected and isolated from PS1 cells. These cells were chosen rather than the cancer cells, as upon activation they have been found to produce large amounts of periostin withing the tumour stroma. Periostin isoforms 201b and 209 have been also identified and investigated by Kim *et al*. (2008) in bladder carcinoma, termed variant I and II, and by Morra *et al*. (2011) in renal carcinoma under the name isoform 4 and 3 respectively. In both studies, no POSTN-202 transcript was isolated. In contrast, Bai *et al*. (2010) detected all three variants and named the variants as thy7 (POSTN-201b), thy6 (POSTN-209) and thy2 (POSTN-202). POSTN-202 was further characterised by the author as a novel specific isoform.

Kim *et al*. (2008) demonstrated that the tumour-suppressive effect that variant II had towards cancer cell invasiveness *in vitro* and metastasis *in vivo* was absent for variant I. The author highlighted that even though the difference between the two variants was only the presence or absence of exon 18, the reason of this distinct function are structural changes in the C-terminal region of the protein rather than the loss of the small specific motif of exon 18. Indeed, our analysis of the 3D structure of the protein showed that the alternative exons encode a region that, if a short sequence is missing, it could change the structure of the protein. Of the 3 main regions that make up periostin, one is composed entirely of the protein sequence encoded by the alternatively spliced exons.

Even though POSTN-201b was found to promote invasiveness in bladder cancer, it is uncertain whether this variant has the same affect in PDAC. Periostin transcript variant expression is tissue dependent, and properties of these variants may differ from tissue to tissue. An analysis of binding capacities of these isoforms with factors such as collagen and fibrinogen, and the functional effects of their overexpression in PS1 cells survival, migration and adhesion would provide valuable information regarding their function. However, due to time constraints this analysis could not be performed in this study.

In this study, only three variants were cloned, however a previous study in our laboratory managed to isolate 6 isoforms (Rahim *et al*. 2019, unpublished data). It may be possible that the

reason for detecting fewer variants was because only small number of colonies were sequenced and analysed. Many of the sequences that were analysed showed the samples contained mixed colonies, hence this data could not be used.

POSTN-209 was the predominant isoform detected the in samples, especially in the first set of samples sequenced. This could indicate high expression levels of this particular in the PS1 cells. This may also suggest a potential important role of this variant in the physiological function of PS1 cells. It has been demonstrated that POSTN-209 has tumour suppressor function in bladder cancer (Kim, C. J., Isono, Tambe, Chano, Okabe, Okada and Inoue, 2008c). This may be the case for PDAC as well. On the other hand, POSTN-201b was only detected in one colony, and this isoform promoted tumour invasiveness in the same study, therefore if it has the same function in pancreatic cancer then it is not unexpected to have limited expression in normal cells. Unfortunately, the data from this study was not conclusive enough to determine the expression profile of POSTN in PS1 cells.

In order to determine the transcript variants expressed in each colony, sequencing analysis was performed on the area of the gene where the exons 17-21 are found. In contrast, AceView provides data for variants (variant g and i) that not only differ in this region but also in the first three exons of the gene. These two isoforms do not possess an EMI domain and signal peptide. Since the single peptide allows the protein to be exported from the cell, the isoforms lacking the signal peptide may have intracellular function. This raises the question whether the isoforms found in this study have alternative splicing in their exons 1-3. This would not be visible as the analysis was only conducted on exons 17-21 in this study. Consequently, two isoforms differing only in exons 1-3 expression would have been characterised as the same variant. To clarify this, different colonies that express POSTN-209 and POSTN-202 should be sequenced again for the region where exons 1-3 are found.

To determine the full expression profile of periostin transcript variants and their level of expression in the PS1 cells, an analysis should be conducted using a larger number of colonies. The process of sequencing was very time consuming and labour intensive as two subsequent RT-PCR reactions was performed for each sample. This could be alleviated by performing direct colony sequencing provided by GeneWiz. Furthermore, agarose gel purification could be performed on the samples prior to Sanger sequencing to improve the quality and number of viable sequences.

Another limitation is that these variants were isolated only from PS1 cells that were cultured alone. PS1 cells are naturally activated when cultured in tissue culture ware due to the high

stiffness of the plastic, however a true activated myofibroblasts-like phenotype is assumed by these cells when exposed to cancer cells e.g., as with cancer associated fibroblasts. Therefore, the PS1 cells used in this study may not have the same expression profile of POSTN isoforms secreted by activated PS1 cells in cancer. In future studies, the transcript variants of POSTN could be cloned from PS1 cells in monoculture and in co-cultures with pancreatic cancer cells. This would help compare the physiological and pathological expression pattern of POSTN transcripts by PS1 cells. Moreover, an investigation on the expression levels of periostin isoforms expressed by other cell types found within the cancer stroma would give a more comprehensive picture of the expression levels of POSTN transcripts in PDAC. An analysis of periostin expression in primary PDAC biopsies could be performed in future studies.

It is essential to isolate the cloned periostin proteins from PS1 cells as it is important to maintain any posttranslational modification that occur during the protein synthesis. This would help to isolate products that are as identical as possible to the periostin proteins produced *in vivo*. To achieve this, the plasmids encoding the three cloned variants of POSTN were transfected into PS1. Following transfection, the cloned proteins could be isolated. Unfortunately, the transfection in this study seems to have failed and consequently, the expression of the transcript variants at a protein level was not achieved. There could be a number for reasons for this. A common cause of transfection failure is cell death that can occur due to insufficient removal of endotoxins during plasmid extraction from the *E.coli*. In this study, a plasmid prep kit that removes endotoxins was used, and since there was no excessive cell death post transfection, endotoxins were not the source of the transfection failure. Another reason could be that the plasmids did not sufficiently penetrate into the PS1 cells due to sub-optimal transfection conditions. Importantly, the size of the plasmid was large which increases the difficulty to transfect the cells via the technique used. A different approach could be utilised, such as lentiviral transduction. This technique involves the use of lentivirus to induce mammalian cell gene expression and has been used in the past for various cell types (Milone and O'Doherty, 2018; Maru, Orihashi and Hippo, 2016; Petrillo *et al.,* 2018). This type of transfection could not be performed in our laboratory.

Following the failure of PS1 cell transfection, the cloned variant proteins were extracted from the original *E.coli* colonies These were the colonies that were selected as they expressed the cloned plasmids which was confirmed by sequencing. However, the periostin proteins were not detected by SDS-PAGE and western blot. Similar protocols have been followed for the isolation of His-tagged proteins (Takayama, Kii and Kudo, 2009). In this case, the purification was not successful. To enhance the protein yield, the number of *E.coli* used can be increased. Only 10ml of an *E.coli* culture was used to induce protein expression but an amount of up to 1L can be used in column-

based protein purification to collect a sufficient protein yield. Additionally, the amount of IPTG used could titrated to optimise the concentration to increase protein synthesis in the *E.coli*.

The notion that the stroma is a critical factor for the progression of PDAC has been supported by multiple studies (Majumder *et al.,* 2016; Moffitt *et al.,* 2015; Tingle, Moir and White, 2015). POSTN expression was found to be high and played an active role in the desmoplastic reaction (Kanno *et al.,* 2008; Neesse *et al.,* 2015). Liu *et al*. (2015) found periostin to be upregulated in the stroma of pancreatic cancer tissue compared with matched adjacent tissues in a relevant microarray analysis. They also showed that periostin expression correlated with the advanced clinical stage of the disease and poor overall survival. Another study presented similar fundings, adding that periostin promoted tumour invasiveness by increasing cell mobility and enhanced cancer cells survival when exposed to hypoxic conditions (Baril *et al.,* 2007). Despite the evidence of the active involvement of periostin in progression of PDAC, no further investigation of the transcript variants and their functional roles have been made.

5 CONCLUSIONS

In this study, three isoforms of periostin were cloned and isolated, POSTN-202, 209 and 201b. These isoforms have been detected by previous studies to have distinct function and either promote or inhibit cancer invasiveness. Given the important role of periostin in the progression of pancreatic cancer and the functional differences that the transcript variants have been found to have in other tissues and cancer types, highlights the importance further investigation into its expression profile and function in PDAC. It would be interesting to determine the binding properties of these variants with key components of the cancer microenvironment, including collagen and fibrinogen. Additionally, a quantitative analysis on the variants expressed in normal and cancer pancreatic tissue would give more information regarding the tumorigenic effects of certain isoforms. The present treatment of pancreatic cancer adenocarcinoma has been proven to be inadequate and new treatment approaches are essential to improving therapy. Periostin could be a future target of chemotherapy considering the promising evidence however further research is still required.

6 REFERENCES

Word Health Organization. (2022) Available at: https://gco.iarc.fr/today/online-analysis-map?v=2020&mode=population&mode_population=continents&population=900&populations=900&key=asr&sex=0&cancer=39&type=0&statistic=5&prevalence=0&population_group=0&ages_group%5B%5D=0&ages_group%5B%5D=17&nb_items=10&group_cancer=1&include_nmsc=0&include_nmsc_other=0&projection=natural-earth&color_palette=default&map_scale=quantile&map_nb_colors=5&continent=0&show_ranking=0&rotate=%255B10%252C0%255D (Accessed: 25/05/2022).

All, S., Garant, A. and Hannan, R. (2021) 'Stereotactic Ablative Radiation (SAbR) for Oligometastatic RCC', *Seminars in Radiation Oncology,* 31(3), pp.227-234.

Bai, Y. *et al.* (2010) 'Novel isoforms of periostin expressed in the human thyroid', *Japanese Clinical Medicine,* 1 pp.13-20.

Baril, P. *et al.* (2007a) 'Periostin promotes invasiveness and resistance of pancreatic cancer cells to hypoxia-induced cell death: role of the beta4 integrin and the PI3k pathway', *Oncogene,* 26(14), pp.2082-2094.

Baril, P. *et al.* (2007b) 'Periostin promotes invasiveness and resistance of pancreatic cancer cells to hypoxia-induced cell death: role of the beta4 integrin and the PI3k pathway', *Oncogene,* 26(14), pp.2082-2094.

Berget, S.M., Moore, C. and Sharp, P.A. (1977) 'Spliced segments at the 5' terminus of adenovirus 2 late mRNA', *Proceedings of the National Academy of Sciences of the United States of America,* 74(8), pp.3171-3175.

Bhaw-Luximon, A. and Jhurry, D. (2015) 'New avenues for improving pancreatic ductal adenocarcinoma (PDAC) treatment: Selective stroma depletion combined with nano drug delivery', *Cancer Letters,* 369(2), pp.266-273.

Blencowe, B.J. (2006) 'Alternative splicing: new insights from global analyses', *Cell,* 126(1), pp.37-47.

Bramhall, S.R. *et al.* (2001) 'Marimastat as first-line therapy for patients with unresectable pancreatic cancer: a randomized trial', *Journal of Clinical Oncology : Official Journal of the American Society of Clinical Oncology,* 19(15), pp.3447-3455.

Bramhall, S.R. *et al.* (2002) 'A double-blind placebo-controlled, randomised study comparing gemcitabine and marimastat with gemcitabine and placebo as first line therapy in patients with advanced pancreatic cancer', *British Journal of Cancer,* 87(2), pp.161-167.

Canto, M.I. *et al.* (2013) 'International Cancer of the Pancreas Screening (CAPS) Consortium summit on the management of patients with increased risk for familial pancreatic cancer', *Gut,* 62(3), pp.339-347.

Capurso, G. *et al.* (2015) 'Methods and outcomes of screening for pancreatic adenocarcinoma in high-risk individuals', *World Journal of Gastrointestinal Endoscopy,* 7(9), pp.833-842.

Chen, F., Roberts, N.J. and Klein, A.P. (2017) 'Inherited pancreatic cancer', *Chinese Clinical Oncology,* 6(6), pp.58.

Chen, M. and Manley, J.L. (2009) 'Mechanisms of alternative splicing regulation: insights from molecular and genomics approaches', *Nature Reviews.Molecular Cell Biology,* 10(11), pp.741-754.

Chow, L.T. *et al.* (1977) 'An amazing sequence arrangement at the 5' ends of adenovirus 2 messenger RNA', *Cell,* 12(1), pp.1-8.

Del Chiaro, M. *et al.* (2014) 'Early detection and prevention of pancreatic cancer: is it really possible today?', *World Journal of Gastroenterology,* 20(34), pp.12118-12131.

Ebos, J.M. *et al.* (2009) 'Accelerated metastasis after short-term treatment with a potent inhibitor of tumor angiogenesis', *Cancer Cell,* 15(3), pp.232-239.

Erkan, M. *et al.* (2007) 'Periostin creates a tumor-supportive microenvironment in the pancreas by sustaining fibrogenic stellate cell activity', *Gastroenterology,* 132(4), pp.1447-1464.

Farrow, B., Albo, D. and Berger, D.H. (2008) 'The role of the tumor microenvironment in the progression of pancreatic cancer', *The Journal of Surgical Research,* 149(2), pp.319-328.

Feig, C. *et al.* (2012) 'The pancreas cancer microenvironment', *Clinical Cancer Research : An Official Journal of the American Association for Cancer Research,* 18(16), pp.4266-4276.

Froeling, F.E. *et al.* (2009) 'Organotypic culture model of pancreatic cancer demonstrates that stromal cells modulate E-cadherin, beta-catenin, and Ezrin expression in tumor cells', *The American Journal of Pathology,* 175(2), pp.636-648.

Fujikawa, T. *et al.* (2021) 'Periostin Exon-21 Antibody Neutralization of Triple-Negative Breast Cancer Cell-Derived Periostin Regulates Tumor-Associated Macrophage Polarization and Angiogenesis', *Cancers,* 13(20), pp.5072. doi: 10.3390/cancers13205072.

Fujita, H. *et al.* (2009) 'Tumor-stromal interactions with direct cell contacts enhance proliferation of human pancreatic carcinoma cells', *Cancer Science,* 100(12), pp.2309-2317.

Goral, V. (2015) 'Pancreatic Cancer: Pathogenesis and Diagnosis', *Asian Pacific Journal of Cancer Prevention : APJCP,* 16(14), pp.5619-5624.

Grant, T.J., Hua, K. and Singh, A. (2016) 'Molecular Pathogenesis of Pancreatic Cancer', *Progress in Molecular Biology and Translational Science,* 144 pp.241-275.

Gupta, R., Amanam, I. and Chung, V. (2017) 'Current and future therapies for advanced pancreatic cancer', *Journal of Surgical Oncology,* 116(1), pp.25-34.

Haqq, J. *et al.* (2014) 'Pancreatic stellate cells and pancreas cancer: current perspectives and future strategies', *European Journal of Cancer (Oxford, England : 1990),* 50(15), pp.2570-2582.

Hausman, D.M. (2019) 'What Is Cancer?', *Perspectives in Biology and Medicine,* 62(4), pp.778-784.

Hawes, R.H. *et al.* (2000) 'A multispecialty approach to the diagnosis and management of pancreatic cancer', *The American Journal of Gastroenterology,* 95(1), pp.17-31.

Heinrich, S. and Lang, H. (2017) 'Neoadjuvant Therapy of Pancreatic Cancer: Definitions and Benefits', *International Journal of Molecular Sciences,* 18(8), pp.1622. doi: 10.3390/ijms18081622.

Henriksen, A. *et al.* (2019) 'Checkpoint inhibitors in pancreatic cancer', *Cancer Treatment Reviews,* 78 pp.17-30.

Hertel, K.J. and Graveley, B.R. (2005) 'RS domains contact the pre-mRNA throughout spliceosome assembly', *Trends in Biochemical Sciences,* 30(3), pp.115-118.

Hidalgo, M. (2012) 'New insights into pancreatic cancer biology', *Annals of Oncology : Official Journal of the European Society for Medical Oncology,* 23 Suppl 10 pp.135.

Hoersch, S. and Andrade-Navarro, M.A. (2010a) 'Periostin shows increased evolutionary plasticity in its alternatively spliced region', *BMC Evolutionary Biology,* 10 pp.30-30.

Hoersch, S. and Andrade-Navarro, M.A. (2010b) 'Periostin shows increased evolutionary plasticity in its alternatively spliced region', *BMC Evolutionary Biology,* 10 pp.30-30.

Hwang, R.F. *et al.* (2008) 'Cancer-associated stromal fibroblasts promote pancreatic tumor progression', *Cancer Research,* 68(3), pp.918-926.

Ikeda-Iwabu, Y. *et al.* (2021) 'Periostin Short Fragment with Exon 17 via Aberrant Alternative Splicing Is Required for Breast Cancer Growth and Metastasis', *Cells,* 10(4), pp.892. doi: 10.3390/cells10040892.

Ilic, M. and Ilic, I. (2016) 'Epidemiology of pancreatic cancer', *World Journal of Gastroenterology,* 22(44), pp.9694-9705.

Jayasingam, S.D. *et al.* (2020) 'Evaluating the Polarization of Tumor-Associated Macrophages Into M1 and M2 Phenotypes in Human Cancer Tissue: Technicalities and Challenges in Routine Clinical Practice', *Frontiers in Oncology,* 9 pp.1512.

Kanno, A. *et al.* (2008) 'Periostin, secreted from stromal cells, has biphasic effect on cell migration and correlates with the epithelial to mesenchymal transition of human pancreatic cancer cells', *International Journal of Cancer,* 122(12), pp.2707-2718.

Kelemen, O. *et al.* (2013a) 'Function of alternative splicing', *Gene,* 514(1), pp.1-30.

Kelemen, O. *et al.* (2013b) 'Function of alternative splicing', *Gene,* 514(1), pp.1-30.

Kim, C.J. *et al.* (2008a) 'Role of alternative splicing of periostin in human bladder carcinogenesis', *International Journal of Oncology,* 32(1), pp.161-169.

Kim, C.J. *et al.* (2008b) 'Role of alternative splicing of periostin in human bladder carcinogenesis', *International Journal of Oncology,* 32(1), pp.161-169.

Kim, C.J. *et al.* (2008c) 'Role of alternative splicing of periostin in human bladder carcinogenesis', *International Journal of Oncology,* 32(1), pp.161-169.

Kim, H.J. *et al.* (2006) 'Identification of a truncated alternative splicing variant of human PPARgamma1 that exhibits dominant negative activity', *Biochemical and Biophysical Research Communications,* 347(3), pp.698-706.

Kim, H.R. *et al.* (2016) 'SRSF5: a novel marker for small-cell lung cancer and pleural metastatic cancer', *Lung Cancer (Amsterdam, Netherlands),* 99 pp.57-65.

Kormann, R. *et al.* (2020) 'Periostin Promotes Cell Proliferation and Macrophage Polarization to Drive Repair after AKI', *Journal of the American Society of Nephrology : JASN,* 31(1), pp.85-100.

Kudo, Y. *et al.* (2004) 'Invasion and metastasis of oral cancer cells require methylation of E-cadherin and/or degradation of membranous beta-catenin', *Clinical Cancer Research : An Official Journal of the American Association for Cancer Research,* 10(16), pp.5455-5463.

Kudo, Y. *et al.* (2007) 'Periostin: novel diagnostic and therapeutic target for cancer', *Histology and Histopathology,* 22(10), pp.1167-1174.

Li, S. and Shen, L. (2020) 'Radiobiology of stereotactic ablative radiotherapy (SABR): perspectives of clinical oncologists', *Journal of Cancer,* 11(17), pp.5056-5068.

Liu, Y. and Du, L. (2015) 'Role of pancreatic stellate cells and periostin in pancreatic cancer progression', *Tumour Biology : The Journal of the International Society for Oncodevelopmental Biology and Medicine,* 36(5), pp.3171-3177.

Loveday, B.P.T., Lipton, L. and Thomson, B.N. (2019) 'Pancreatic cancer: An update on diagnosis and management', *Australian Journal of General Practice,* 48(12), pp.826-831.

Maddalena, M. *et al.* (2021) 'TP53 missense mutations in PDAC are associated with enhanced fibrosis and an immunosuppressive microenvironment', *Proceedings of the National Academy of Sciences of the United States of America,* 118(23), pp.e2025631118. doi: 10.1073/pnas.2025631118.

Majumder, K. *et al.* (2016) 'A Novel Immunocompetent Mouse Model of Pancreatic Cancer with Robust Stroma: a Valuable Tool for Preclinical Evaluation of New Therapies', *Journal of Gastrointestinal Surgery : Official Journal of the Society for Surgery of the Alimentary Tract,* 20(1), pp.53-65; discussion 65.

Malanchi, I. *et al.* (2011) 'Interactions between cancer stem cells and their niche govern metastatic colonization', *Nature,* 481(7379), pp.85-89.

Maru, Y., Orihashi, K. and Hippo, Y. (2016) 'Lentivirus-Based Stable Gene Delivery into Intestinal Organoids', *Methods in Molecular Biology (Clifton, N.J.),* 1422 pp.13-21.

McGuigan, A. *et al.* (2018) 'Pancreatic cancer: A review of clinical diagnosis, epidemiology, treatment and outcomes', *World Journal of Gastroenterology,* 24(43), pp.4846-4861.

Mier, P. *et al.* (2020) 'Disentangling the complexity of low complexity proteins', *Briefings in Bioinformatics,* 21(2), pp.458-472.

Milone, M.C. and O'Doherty, U. (2018) 'Clinical use of lentiviral vectors', *Leukemia,* 32(7), pp.1529-1541.

Mishra, S.K. *et al.* (2020) 'Periostin Activation of Integrin Receptors on Sensory Neurons Induces Allergic Itch', *Cell Reports,* 31(1), pp.107472.

Moffitt, R.A. *et al.* (2015) 'Virtual microdissection identifies distinct tumor- and stroma-specific subtypes of pancreatic ductal adenocarcinoma', *Nature Genetics,* 47(10), pp.1168-1178.

Moore, M.J. *et al.* (2003) 'Comparison of gemcitabine versus the matrix metalloproteinase inhibitor BAY 12-9566 in patients with advanced or metastatic adenocarcinoma of the pancreas: a phase III trial of the National Cancer Institute of Canada Clinical Trials Group', *Journal of Clinical Oncology : Official Journal of the American Society of Clinical Oncology,* 21(17), pp.3296-3302.

Morra, L. *et al.* (2011) 'Relevance of periostin splice variants in renal cell carcinoma', *The American Journal of Pathology,* 179(3), pp.1513-1521.

Morra, L. *et al.* (2012) 'Characterization of periostin isoform pattern in non-small cell lung cancer', *Lung Cancer (Amsterdam, Netherlands),* 76(2), pp.183-190.

Mulherkar, N., Prasad, K.V. and Prabhakar, B.S. (2007) 'MADD/DENN splice variant of the IG20 gene is a negative regulator of caspase-8 activation. Knockdown enhances TRAIL-induced apoptosis of cancer cells', *The Journal of Biological Chemistry,* 282(16), pp.11715-11721.

Narayanan, S., Vicent, S. and Ponz-Sarvisé, M. (2021) 'PDAC as an Immune Evasive Disease: Can 3D Model Systems Aid to Tackle This Clinical Problem?', *Frontiers in Cell and Developmental Biology,* 9 pp.787249.

Neesse, A. *et al.* (2015) 'Stromal biology and therapy in pancreatic cancer: a changing paradigm', *Gut,* 64(9), pp.1476-1484.

Nikoloudaki, G. *et al.* (2020) 'Periostin and matrix stiffness combine to regulate myofibroblast differentiation and fibronectin synthesis during palatal healing', *Matrix Biology,* 94 pp.31-56.

O'Dwyer, D.N. and Moore, B.B. (2017) 'The role of periostin in lung fibrosis and airway remodeling', *Cellular and Molecular Life Sciences : CMLS,* 74(23), pp.4305-4314.

Oh, H.J. *et al.* (2017) 'Overexpression of POSTN in Tumor Stroma Is a Poor Prognostic Indicator of Colorectal Cancer', *Journal of Pathology and Translational Medicine,* 51(3), pp.306-313.

Pàez-Ribes, M. *et al.* (2009) 'Antiangiogenic therapy elicits malignant progression of tumors to increased local invasion and distant metastasis', *Cancer Cell,* 15(3), pp.220-231.

Palma, D.A. *et al.* (2020) 'Stereotactic Ablative Radiotherapy for the Comprehensive Treatment of Oligometastatic Cancers: Long-Term Results of the SABR-COMET Phase II Randomized Trial', *Journal of Clinical Oncology : Official Journal of the American Society of Clinical Oncology,* 38(25), pp.2830-2838.

Park, S.Y. *et al.* (2016) 'Periostin (POSTN) Regulates Tumor Resistance to Antiangiogenic Therapy in Glioma Models', *Molecular Cancer Therapeutics,* 15(9), pp.2187-2197.

Petrillo, C. *et al.* (2018) 'Cyclosporine H Overcomes Innate Immune Restrictions to Improve Lentiviral Transduction and Gene Editing In Human Hematopoietic Stem Cells', *Cell Stem Cell,* 23(6), pp.820-832.e9.

Picozzi, V. *et al.* (2020) 'Gemcitabine/nab-paclitaxel with pamrevlumab: a novel drug combination and trial design for the treatment of locally advanced pancreatic cancer', *ESMO Open,* 5(4), pp.e000668. doi: 10.1136/esmoopen-000668.

Rooke, N. *et al.* (2003) 'Roles for SR proteins and hnRNP A1 in the regulation of c-src exon N1', *Molecular and Cellular Biology,* 23(6), pp.1874-1884.

Roy, P.S. and Saikia, B.J. (2016) 'Cancer and cure: A critical analysis', *Indian Journal of Cancer,* 53(3), pp.441-442.

Roy-Luzarraga, M. *et al.* (2022) 'Suppression of Endothelial Cell FAK Expression Reduces Pancreatic Ductal Adenocarcinoma Metastasis after Gemcitabine Treatment', *Cancer Research,* 82(10), pp.1909-1925.

Ruan, K., Bao, S. and Ouyang, G. (2009) 'The multifaceted role of periostin in tumorigenesis', *Cellular and Molecular Life Sciences : CMLS,* 66(14), pp.2219-2230.

Saka, D. *et al.* (2020) 'Mechanisms of T-Cell Exhaustion in Pancreatic Cancer', *Cancers,* 12(8), pp.2274. doi: 10.3390/cancers12082274.

Shao, R. *et al.* (2004) 'Acquired expression of periostin by human breast cancers promotes tumor angiogenesis through up-regulation of vascular endothelial growth factor receptor 2 expression', *Molecular and Cellular Biology,* 24(9), pp.3992-4003.

Shenasa, H. and Hertel, K.J. (2019) 'Combinatorial regulation of alternative splicing', *Biochimica Et Biophysica Acta.Gene Regulatory Mechanisms,* 1862(11-12), pp.194392.

Siriwardena, B.S. *et al.* (2006) 'Periostin is frequently overexpressed and enhances invasion and angiogenesis in oral cancer', *British Journal of Cancer,* 95(10), pp.1396-1403.

Smathers, C.M. and Robart, A.R. (2019) 'The mechanism of splicing as told by group II introns: Ancestors of the spliceosome', *Biochimica Et Biophysica Acta.Gene Regulatory Mechanisms,* 1862(11-12), pp.194390.

Stewart, R. *et al.* (2021) 'SABR in oligometastatic breast cancer: Current status and future directions', *Breast (Edinburgh, Scotland),* 60 pp.223-229.

Tai, I.T., Dai, M. and Chen, L.B. (2005) 'Periostin induction in tumor cell line explants and inhibition of in vitro cell growth by anti-periostin antibodies', *Carcinogenesis,* 26(5), pp.908-915.

Takayama, I., Kii, I. and Kudo, A. (2009) 'Expression, purification and characterization of soluble recombinant periostin protein produced by Escherichia coli', *Journal of Biochemistry,* 146(5), pp.713-723.

Tingle, S.J., Moir, J.A. and White, S.A. (2015) 'Role of anti-stromal polypharmacy in increasing survival after pancreaticoduodenectomy for pancreatic ductal adenocarcinoma', *World Journal of Gastrointestinal Pathophysiology,* 6(4), pp.235-242.

Torphy, R.J., Fujiwara, Y. and Schulick, R.D. (2020) 'Pancreatic cancer treatment: better, but a long way to go', *Surgery Today,* 50(10), pp.1117-1125.

Uversky, V.N. (2013) 'The alphabet of intrinsic disorder: II. Various roles of glutamic acid in ordered and intrinsically disordered proteins', *Intrinsically Disordered Proteins,* 1(1), pp.e24684.

Vegran, F. *et al.* (2005) 'Implication of alternative splice transcripts of caspase-3 and survivin in chemoresistance', *Bulletin Du Cancer,* 92(3), pp.219-226.

Viloria, K. and Hill, N.J. (2016) 'Embracing the complexity of matricellular proteins: the functional and clinical significance of splice variation', *Biomolecular Concepts,* 7(2), pp.117-132.

Vincent, A. *et al.* (2011) 'Pancreatic cancer', *Lancet (London, England),* 378(9791), pp.607-620.

Wang, Z. and Burge, C.B. (2008) 'Splicing regulation: from a parts list of regulatory elements to an integrated splicing code', *RNA (New York, N.Y.),* 14(5), pp.802-813.

Webb, T. (2005) 'Vascular normalization: study examines how antiangiogenesis therapies work', *Journal of the National Cancer Institute,* 97(5), pp.336-337.

Wong, P.P. *et al.* (2015) 'Dual-action combination therapy enhances angiogenesis while reducing tumor growth and spread', *Cancer Cell,* 27(1), pp.123-137.

Wood, L.D. and Hruban, R.H. (2015) 'Genomic landscapes of pancreatic neoplasia', *Journal of Pathology and Translational Medicine,* 49(1), pp.13-22.

Xie, D. and Xie, K. (2015) 'Pancreatic cancer stromal biology and therapy', *Genes & Diseases,* 2(2), pp.133-143.

Yan, W. and Shao, R. (2006) 'Transduction of a mesenchyme-specific gene periostin into 293T cells induces cell invasive activity through epithelial-mesenchymal transformation', *The Journal of Biological Chemistry,* 281(28), pp.19700-19708.

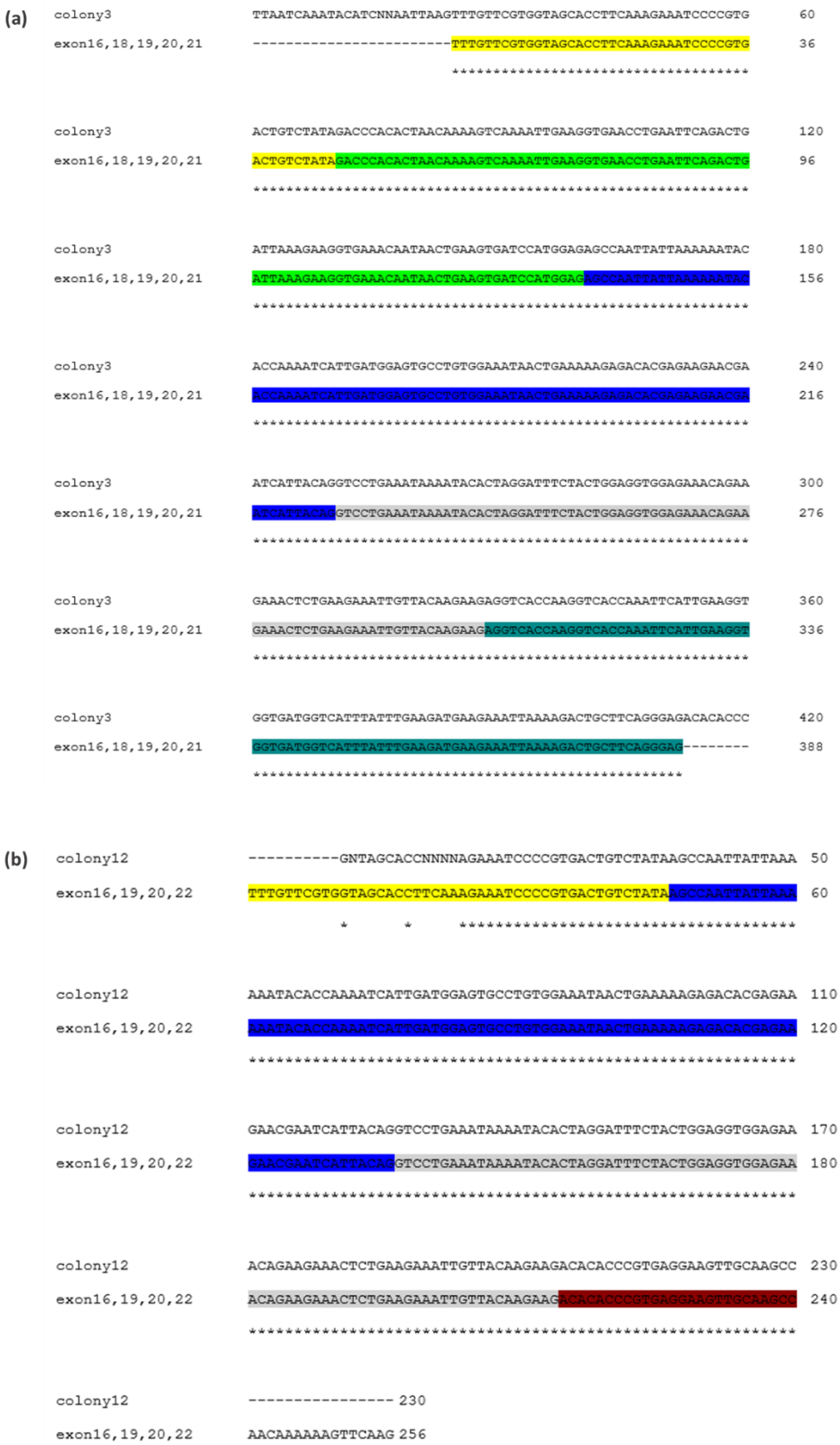
Yang, C.Y. *et al.* (2020) 'Engineering Chimeric Antigen Receptor T Cells against Immune Checkpoint Inhibitors PD-1/PD-L1 for Treating Pancreatic Cancer', *Molecular Therapy Oncolytics,* 17 pp.571-585.

Yeo, G. *et al.* (2004) 'Variation in alternative splicing across human tissues', *Genome Biology,* 5(10), pp.R74-r74. Epub 2004 Sep 13.

Yuan, Y. *et al.* (2019) 'BAG3-positive pancreatic stellate cells promote migration and invasion of pancreatic ductal adenocarcinoma', *Journal of Cellular and Molecular Medicine,* 23(8), pp.5006-5016.

Zaimy, M.A. *et al.* (2017) 'New methods in the diagnosis of cancer and gene therapy of cancer based on nanoparticles', *Cancer Gene Therapy,* 24(6), pp.233-243.

7 APPENDIX


(a)
colony3 TTAATCAAATACATCNNAATTAAGTTTGTTCGTGGTAGCACCTTCAAAGAAATCCCCGTG 60
exon16,18,19,20,21 ----------TTTGTTCGTGGTAGCACCTTCAAAGAAATCCCCGTG 36
colony3 ACTGTCTATAGACCCACACTAACAAAAGTCAAAATTGAAGGTGAACCTGAATTCAGACTG 120
exon16,18,19,20,21 ACTGTCTATAGACCCACACTAACAAAAGTCAAAATTGAAGGTGAACCTGAATTCAGACTG 96
colony3 ATTAAAGAAGGTGAAACAATAACTGAAGTGATCCATGGAGAGCCAATTATTAAAAAATAC 180
exon16,18,19,20,21 ATTAAAGAAGGTGAAACAATAACTGAAGTGATCCATGGAGAGCCAATTATTAAAAAATAC 156
colony3 ACCAAAATCATTGATGGAGTGCCTGTGGAAATAACTGAAAAAGAGACACGAGAAGAACGA 240
exon16,18,19,20,21 ACCAAAATCATTGATGGAGTGCCTGTGGAAATAACTGAAAAAGAGACACGAGAAGAACGA 216
colony3 ATCATTACAGGTCCTGAAATAAAATACACTAGGATTTCTACTGGAGGTGGAGAAACAGAA 300
exon16,18,19,20,21 ATCATTACAGGTCCTGAAATAAAATACACTAGGATTTCTACTGGAGGTGGAGAAACAGAA 276
colony3 GAAACTCTGAAGAAATTGTTACAAGAAGAGGTCACCAAGGTCACCAAATTCATTGAAGGT 360
exon16,18,19,20,21 GAAACTCTGAAGAAATTGTTACAAGAAGAGGTCACCAAGGTCACCAAATTCATTGAAGGT 336
colony3 GGTGATGGTCATTTATTTGAAGATGAAGAAATTAAAAGACTGCTTCAGGGAGACACACCC 420
exon16,18,19,20,21 GGTGATGGTCATTTATTTGAAGATGAAGAAATTAAAAGACTGCTTCAGGGAG-------- 388
(b)
colony12 ----------GNTAGCACCNNNNAGAAATCCCCGTGACTGTCTATAAGCCAATTATTAAA 50
exon16,19,20,22 TTTGTTCGTGGTAGCACCTTCAAAGAAATCCCCGTGACTGTCTATAAGCCAATTATTAAA 60
colony12 AAATACACCAAAATCATTGATGGAGTGCCTGTGGAAATAACTGAAAAAGAGACACGAGAA 110
exon16,19,20,22 AAATACACCAAAATCATTGATGGAGTGCCTGTGGAAATAACTGAAAAAGAGACACGAGAA 120
colony12 GAACGAATCATTACAGGTCCTGAAATAAAATACACTAGGATTTCTACTGGAGGTGGAGAA 170
exon16,19,20,22 GAACGAATCATTACAGGTCCTGAAATAAAATACACTAGGATTTCTACTGGAGGTGGAGAA 180
colony12 ACAGAAGAAACTCTGAAGAAATTGTTACAAGAAGACACACCCGTGAGGAAGTTGCAAGCC 230
exon16,19,20,22 ACAGAAGAAACTCTGAAGAAATTGTTACAAGAAGACACACCCGTGAGGAAGTTGCAAGCC 240
colony12 ---------- 230
exon16,19,20,22 AACAAAAAAGTTCAAG 256

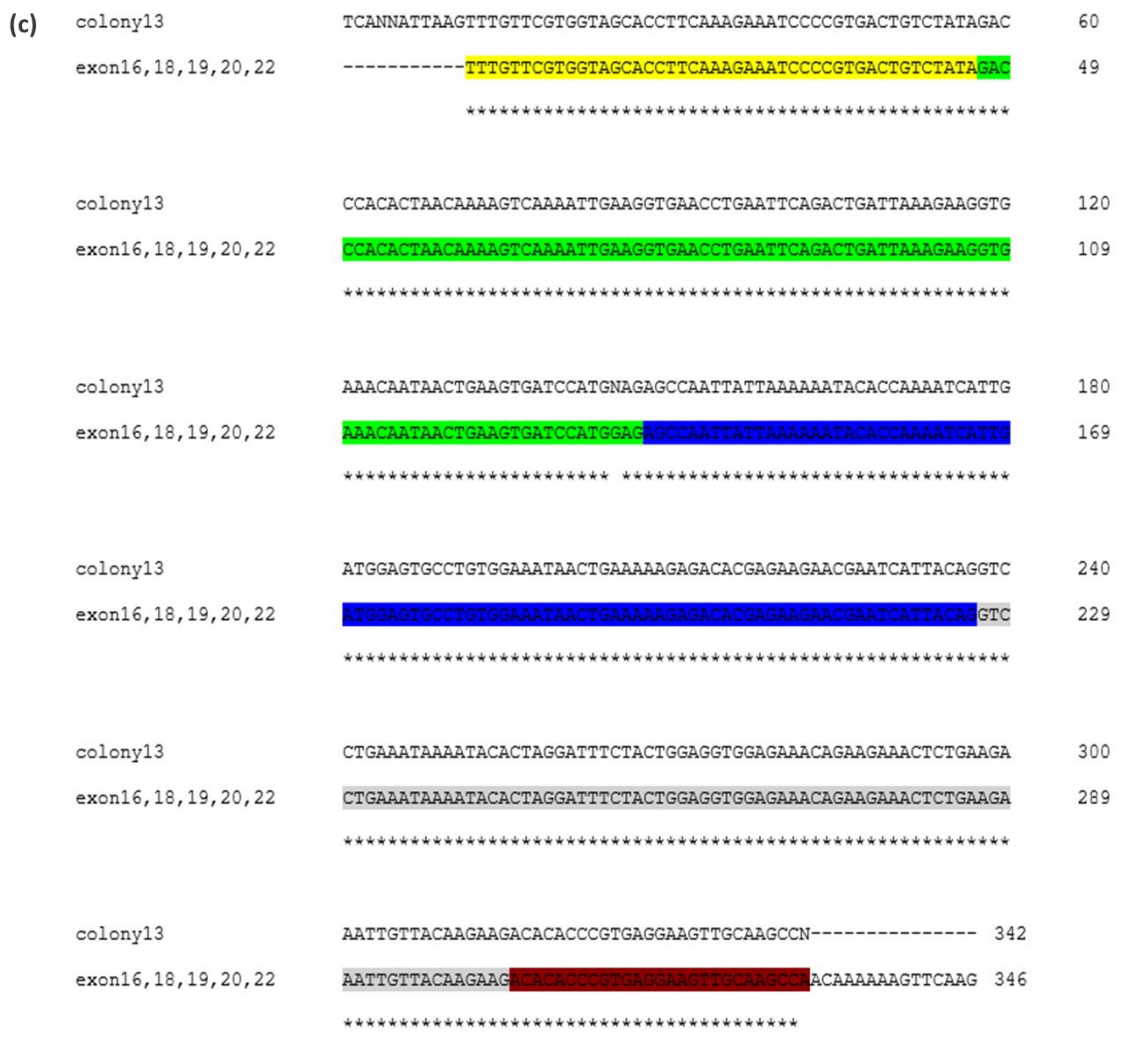
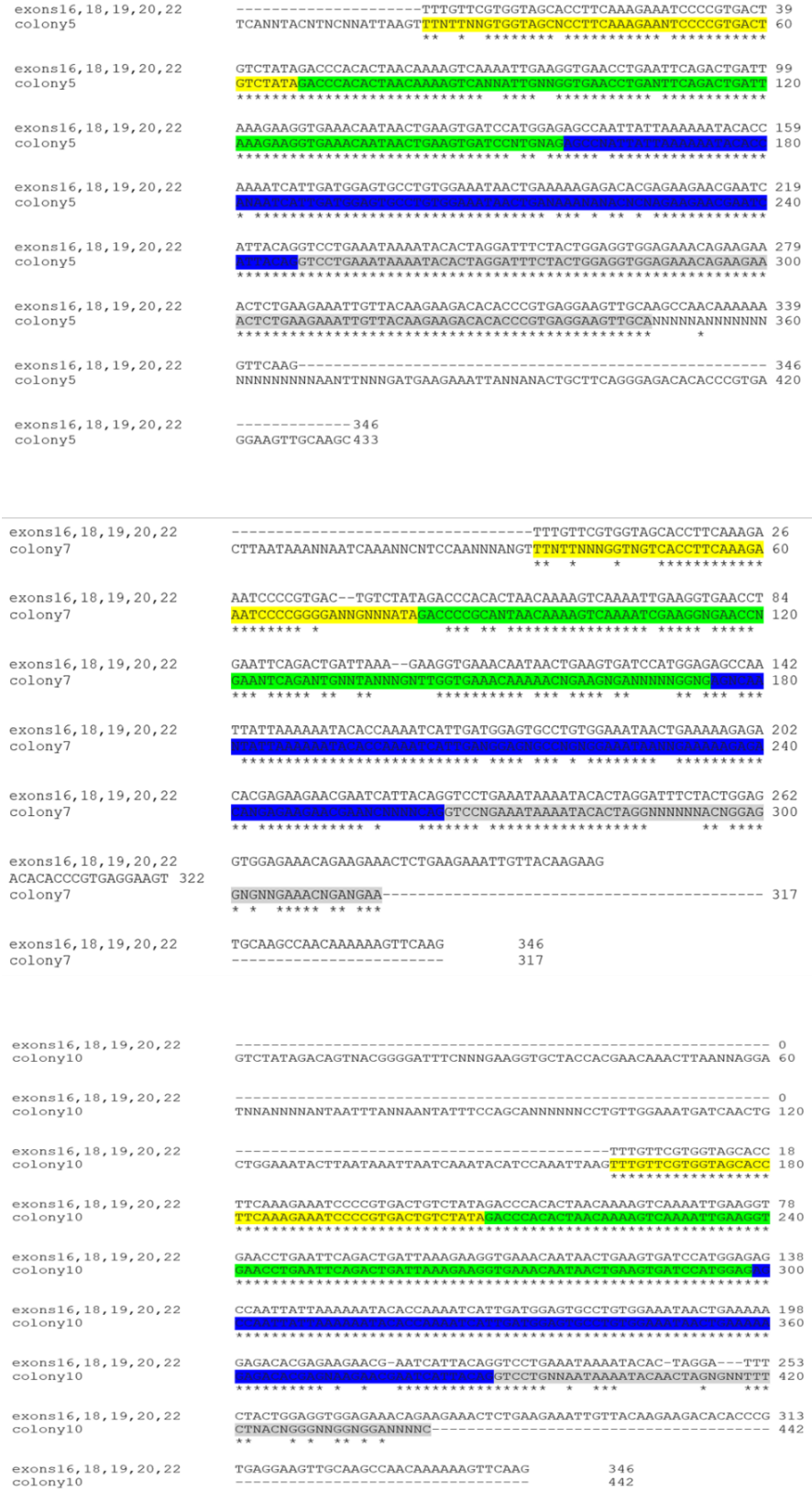
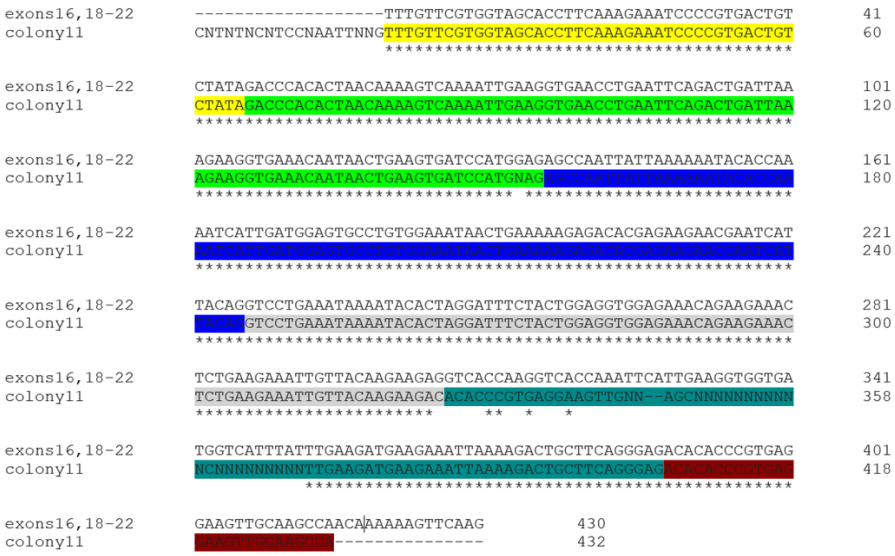
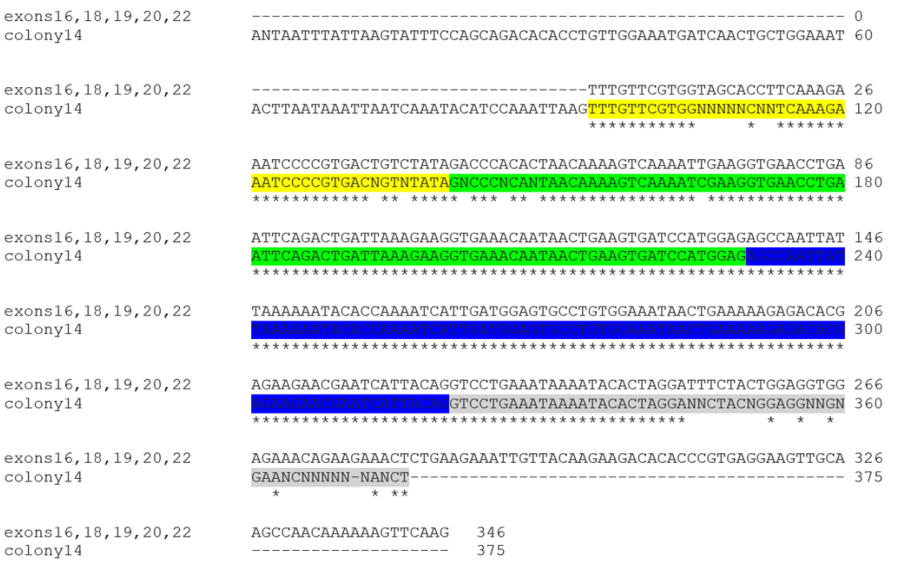


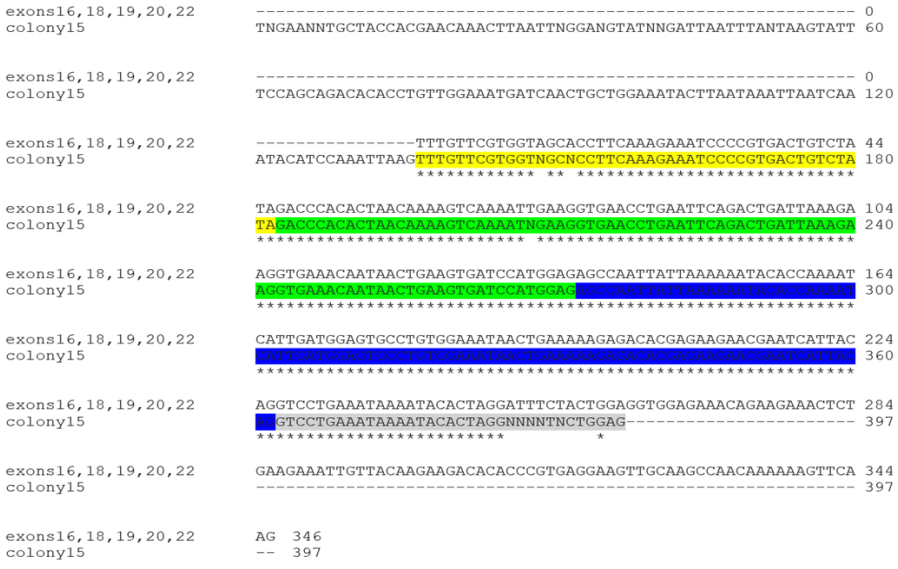
Figure 7.1: Sequencing of the three colonies of interest. The image was created to summarise the exons present in the colonies and does not represent the exact way the analysis was performed. The sequences were compared against one exon at a time, however this data was too long to be shown. When the exons of each sample were determined all the exons were then aligned with the sequence to give the results shown on this figure. Exons are coloured to visualise where are found in the sequence: exon 16-yellow, exon 18-green, exon19-blue, exon 20-grey, exon21-turquoise and exon22-bordeaux.

(a) colony 3 that includes exons 16,18,19,20 and 21. The sequence was too short to show the presence of exon 22 but this exon is a cascade exon and it does not affect the final determination of the isoform **(b)** colony 12 including exons 16,19,20 and 22 **(c)** colony 13 that includes exons 16,18,19,20 and 22.


exons16,18,19,20,22
colony5
exons16,18,19,20,22
colony7
exons16,18,19,20,22
colony10


exons16,18-22 ----------TTTGTTCGTGGTAGCACCTTCAAAGAAATCCCCGTGACTGT 41
colony11 CNTNTNCNTCCNAATTNNGTTTGTTCGTGGTAGCACCTTCAAAGAAATCCCCGTGACTGT 60
exons16,18-22 CTATAGACCCACACTAACAAAAGTCAAAATTGAAGGTGAACCTGAATTCAGACTGATTAA 101
colony11 CTATAGACCCACACTAACAAAAGTCAAAATTGAAGGTGAACCTGAATTCAGACTGATTAA 120
exons16,18-22 AGAAGGTGAAACAATAACTGAAGTGATCCATGGAGAGCCAATTATTAAAAAATACACCAA 161
colony11 AGAAGGTGAAACAATAACTGAAGTGATCCATGNAG 180
exons16,18-22 AATCATTGATGGAGTGCCTGTGGAAATAACTGAAAAAGAGACACGAGAAGAACGAATCAT 221
colony11 240
exons16,18-22 TACAGGTCCTGAAATAAAATACACTAGGATTTCTACTGGAGGTGGAGAAACAGAAGAAAC 281
colony11 GTCCTGAAATAAAATACACTAGGATTTCTACTGGAGGTGGAGAAACAGAAGAAAC 300
exons16,18-22 TCTGAAGAAATTGTTACAAGAAGAGGTCACCAAGGTCACCAAATTCATTGAAGGTGGTGA 341
colony11 TCTGAAGAAATTGTTACAAGAAGACACACCCGTGAGGAAGTTGNN--AGCNNNNNNNNNN 358
exons16,18-22 TGGTCATTTATTTGAAGATGAAGAAATTAAAAGACTGCTTCAGGGAGACACACCCGTGAG 401
colony11 NCNNNNNNNNTTGAAGATGAAGAAATTAAAAGACTGCTTCAGGGAG 418
exons16,18-22 GAAGTTGCAAGCCAACAAAAAAGTTCAAG 430
colony11 ----------- 432


exons16,18,19,20,22 ---------- 0
colony14 ANTAATTTATTAAGTATTTCCAGCAGACACACCTGTTGGAAATGATCAACTGCTGGAAAT 60
exons16,18,19,20,22 ----------TTTGTTCGTGGTAGCACCTTCAAAGA 26
colony14 ACTTAATAAATTAATCAAATACATCCAAATTAAGTTTGTTCGTGGNNNNNCNNTCAAAGA 120
exons16,18,19,20,22 AATCCCCGTGACTGTCTATAGACCCACACTAACAAAAGTCAAAATTGAAGGTGAACCTGA 86
colony14 AATCCCCGTGACNGTNTATAGNCCCNCANTAACAAAAGTCAAAATCGAAGGTGAACCTGA 180
exons16,18,19,20,22 ATTCAGACTGATTAAAGAAGGTGAAACAATAACTGAAGTGATCCATGGAG 146
colony14 ATTCAGACTGATTAAAGAAGGTGAAACAATAACTGAAGTGATCCATGGAG 240
exons16,18,19,20,22 TAAAAAATACACCAAAATCATTGATGGAGTGCCTGTGGAAATAACTGAAAAAGAGACACG 206
colony14 300
exons16,18,19,20,22 AGAAGAACGAATCATTACAGGTCCTGAAATAAAATACACTAGGATTTCTACTGGAGGTGG 266
colony14 GTCCTGAAATAAAATACACTAGGANNCTACNGGAGGNNGN 360
exons16,18,19,20,22 AGAAACAGAAGAAACTCTGAAGAAATTGTTACAAGAAGACACACCCGTGAGGAAGTTGCA 326
colony14 GAANCNNNNN-NANCT---------- 375
exons16,18,19,20,22 AGCCAACAAAAAAGTTCAAG 346
colony14 ---------- 375


exons16,18,19,20,22 ---------- 0
colony15 TNGAANNTGCTACCACGAACAAACTTAATTNGGANGTATNNGATTAATTTANTAAGTATT 60
exons16,18,19,20,22 ---------- 0
colony15 TCCAGCAGACACACCTGTTGGAAATGATCAACTGCTGGAAATACTTAATAAATTAATCAA 120
exons16,18,19,20,22 ----------TTTGTTCGTGGTAGCACCTTCAAAGAAATCCCCGTGACTGTCTA 44
colony15 ATACATCCAAATTAAGTTTGTTCGTGGTNGCNCCTTCAAAGAAATCCCCGTGACTGTCTA 180
exons16,18,19,20,22 TAGACCCACACTAACAAAAGTCAAAATTGAAGGTGAACCTGAATTCAGACTGATTAAAGA 104
colony15 TAGACCCACACTAACAAAAGTCAAAATNGAAGGTGAACCTGAATTCAGACTGATTAAAGA 240
exons16,18,19,20,22 AGGTGAAACAATAACTGAAGTGATCCATGGAGAGCCAATTATTAAAAAATACACCAAAAT 164
colony15 AGGTGAAACAATAACTGAAGTGATCCATGGAG 300
exons16,18,19,20,22 CATTGATGGAGTGCCTGTGGAAATAACTGAAAAAGAGACACGAGAAGAACGAATCATTAC 224
colony15 360
exons16,18,19,20,22 AGGTCCTGAAATAAAATACACTAGGATTTCTACTGGAGGTGGAGAAACAGAAGAAACTCT 284
colony15 GTCCTGAAATAAAATACACTAGGNNNNTNCTGGAG---------- 397
exons16,18,19,20,22 GAAGAAATTGTTACAAGAAGACACACCCGTGAGGAAGTTGCAAGCCAACAAAAAAGTTCA 344
colony15 ---------- 397
exons16,18,19,20,22 AG 346
colony15 -- 397

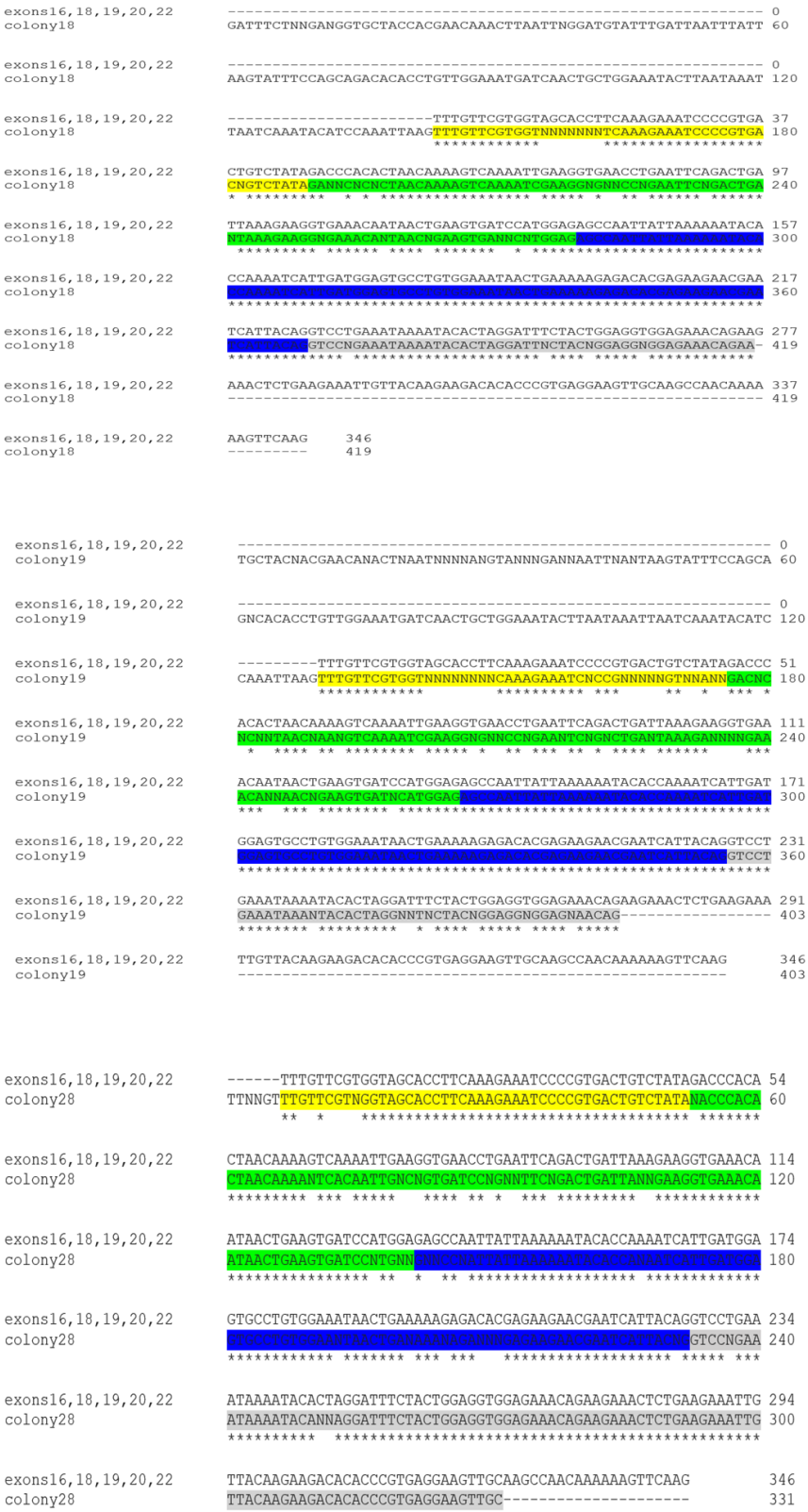


Figure 7.2: Sequencing analysis of cloned single colonies of E.coli. The sequences of colonies: 5, 7, 10, 11, 14, 15, 18, 19 and 28 were aligned with the exons they were expressing. Exons are coloured in exon 16-yellow, exon 18-green, exon19-blue, exon 20-grey, exon21-turquoise and exon22-bordeaux. The colony number can be found on the left-hand side.