Periostin Splice Variants in Pancreatic Cancer

By

Musamma Jamila RAHIM

First supervisor: Dr Natasha J Hill

Other supervisors: Dr Lucy Jones and Dr Jessica Buxton

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

August 2019

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 Natasha Hill and my supervisory team for their extensive support, guidance and expertise throughout this project. To Dr Jessica Buxton for proof-reading my work and providing me with constructive feedbacks.

To Dr Ali Ryan for allowing me to work in his lab for a year and for training me for the cloning work and guiding me during the whole trouble-shooting process.

And to technical staff and Jayne Reeves for training me in all the equipment such as the Biodrop spectrophotometer.

I would like to thank my collaborators who gave me the cells for this project, Prof Helmout Modjtahedi and Dr Athina Myrto-Chioni from Kingston University, and Prof. Hemant Kocher from Barts Cancer Institute at Queen Mary University of London.

I am grateful for the initial training provided by colleagues (Sharan Asher, Amanda Dandagama Munasinghe, and Amtul Bhunnoo) in RT PCR, RNA extraction and cDNA synthesis, and basic cell culture. Also, to Sissy Spiritosanto for training me in Western Blot.

I want to thank colleagues and friends in the Integrated Research and Teaching Laboratory (IRTL) for their moral support and thanks to Dr Lauren Mulcahy-Ryan for managing the laboratory and for being there when I needed the most help. To Rosalind Percival for helping behind the scenes with all the admin work required for research to happen.

I want to extend my gratitude to my family for always being supportive and giving me endless encouragement to pursue on with this project even when I was so close to giving up. I want to also thank my close friends, Amtul Bhunnoo for inspiring me with quotes and always being there for me in the difficult times; Mahreen, Frankie, Salma, Ambreen and Aanisha for instilling hope in me and keeping my sanity throughout this long stressful period.

Last but not the least, I want to thank Allah for granting me with this wonderful opportunity to embark on this project and for blessing me in so many ways.

Abstract

Introduction: Pancreatic ductal adenocarcinoma (PDAC) is typically characterised with fibrotic stroma. It is a lethal disease due to its incidence rate being almost equal to its mortality rate. The poor prognosis of PDAC has not changed for the past forty years is due to lack of early detection, late diagnosis of the cancer and ineffectiveness of conventional treatments due to chemoresistance. Recent studies show a critical role of the tumour microenvironment in cancer progression. Periostin (POSTN) is a matricellular protein involved in development and tissue injury. POSTN is overexpressed in multiple types of cancer including PDAC. Alternative splicing has been shown to generate multiple POSTN transcript variants in other tissues but has not been examined in PDAC. The aim of this project was to investigate the POSTN isoforms in PDAC and carry out functional studies to characterise these isoforms.

Methods: Bioinformatics analysis of POSTN was carried out to investigate the potential POSTN isoforms and an RT-PCR approach was used to detect the expression of POSTN isoforms in multiple cell lines. Western Blot analysis was carried out on PS-1 cell lysates to detect POSTN protein expression. Traditional cloning was utilized to clone all POSTN isoforms to produce recombinant proteins to carry out functional studies.

Results: Bioinformatics analysis reveal ten POSTN transcript variants across three databases (Ensembl, Genbank and AceView). Predictions made from domain databases show two potential intracellular POSTN proteins. RT PCR analysis showed expression of seven POSTN transcript variants namely POSTN-001, 002,003,004, 201a, 201 band novel POSTN isoform in pancreatic stellate (PS-1) cells. POSTN was not expressed in any cancer cell lines tested. Similarly, Western blot analysis showed the presence of potential POSTN protein isoforms. Substantial progress was made with cloning POSTN isoforms.

Conclusion: This study revealed the presence of several POSTN isoforms in pancreatic stellate cells which has the potential to be used as clinical biomarkers or therapeutic targets of PDAC. Further studies are warranted to fully characterise these isoforms.

Table of Contents

List of figures

List of Tables

List of abbreviations

Akt	Serine–threonine kinase
ATP	Adenosine triphosphate
cDNA	Complementary Deoxyribonucleic acid
CDS	Coding sequence
CT	Computed Tomography
CNV	Copy Number Variation
ECM	Extracellular Matrix
ERK	Extracellular Signal-regulated kinase
ESE	Exonic Splicing Enhancer
ESS	Exonic Splicing Silencers
FAK	Focal Adhesion Kinase
FBS	Fetal Bovine Serum
HEK	Human embryonic kidney
Herg	Human ether-a-gogo related gene
HnRNP	Heterogeneous nuclear ribonucleoprotein
HPSC	Human Pancreatic Stromal Cells
IARC	International Agency for Research on Cancer
ISE	Intronic Splicing Enhancer
ISS	Intronic Splicing Silencer
JNK	c-Jun N-terminal Kinases
MCS	Multiple Cloning Site
MRI	Magnetic Resonance Imaging
NRT	No Reverse Transcriptase
NTC	No Template Control
ORF	Open Reading Frame
OS	Overall Survival
OSF-2	Osteoblast Specific Factor-2
PDAC	Pancreatic Ductal Adenocarcinoma
PKB	Protein Kinase B
POSTN	Periostin

PS-1	Pancreatic stellate cells
RIPA	Radio immune Precipitation Assay
RT	Reverse Transcriptase
RT-PCR	Reverse Transcription-Polymerase Chain Reaction
SEER	Surveillance, Epidemiology and End Results Program
SF1	Splicing Factor 1
snRNP U1	small nuclear Ribonucleoproteins U1
SR proteins	Serine/arginine-rich family of nuclear phosphoproteins
TNM	Tumour, Node, Metastasis
TSL	Transcript Support Level
U2AF	U2 Auxiliary Factor

Chapter 1- Introduction

1.1 Pancreatic cancer

1.1.1 Incidence and mortality

Cancer is a global issue which ranks the second most common cause of mortality in the US (Siegel *et al.,* 2018). Pancreatic ductal adenocarcinoma (PDAC) affects the pancreatic ductal cells and accounts for 95% of cases in pancreatic cancer (Tingle *et al.,* 2015). Figure 1.1 show PDAC incidence is relatively rare compared with other cancers such as breast, prostate and lung cancers. However, according to the International Agency for Research on Cancer (IARC), PDAC is the seventh major cause of all cancer mortality in developed countries. However, it is the third major cause of cancer mortality in the US according to Surveillance, Epidemiology and End Results Program (SEER) (National Cancer Institute, 2019). The 5-year survival rate in the US is estimated to be 8.5 % according to the data from SEER, which has not changed in the past forty years. There is a small difference between the incidence and mortality rates of PDAC in both males and females across different populations which indicates the lethality of this disease (Figures 1.2).

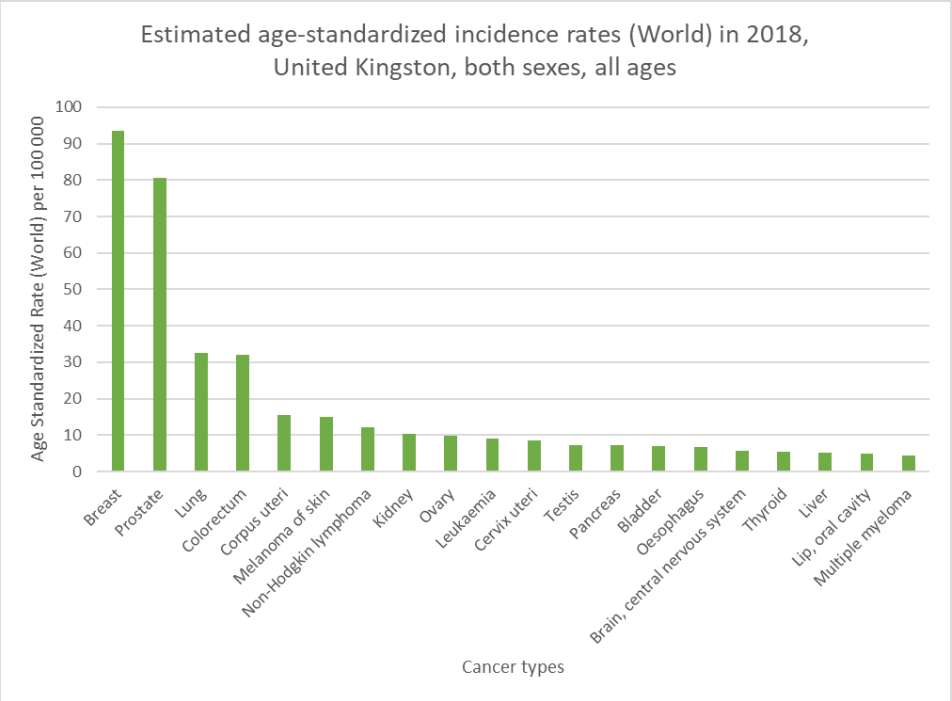
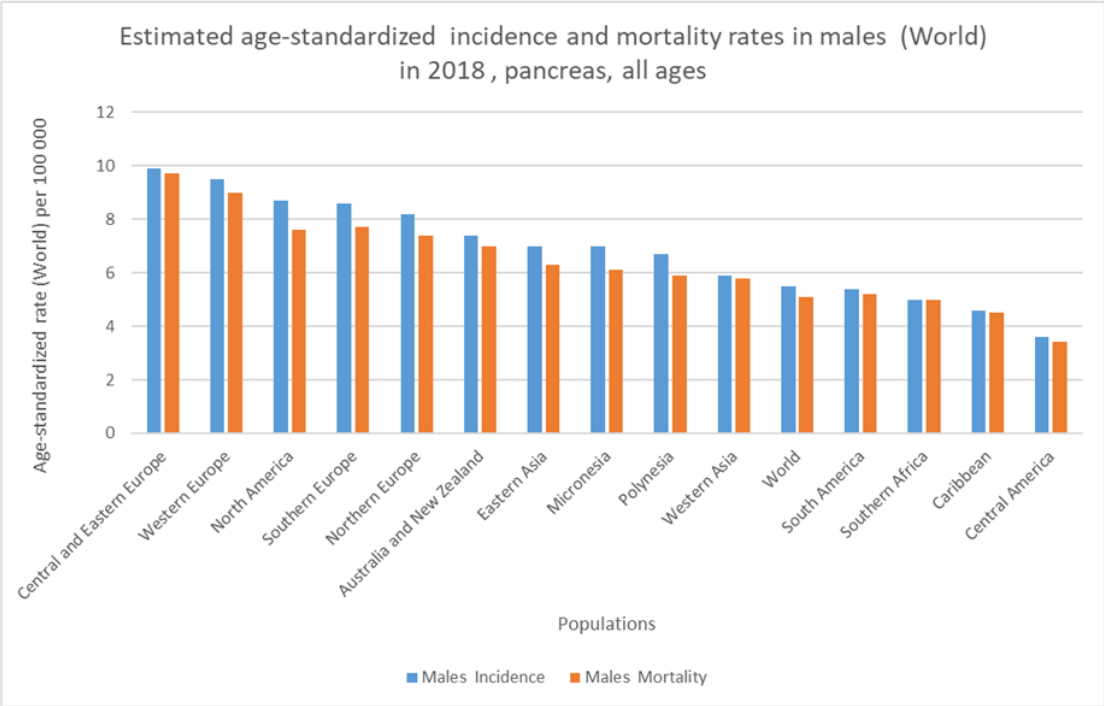


Figure 1. 1 Pancreatic cancer incidence in men and women of all ages in the United Kingdom, adapted from GLOBOCAN 2018 estimates.

(a)



(b)

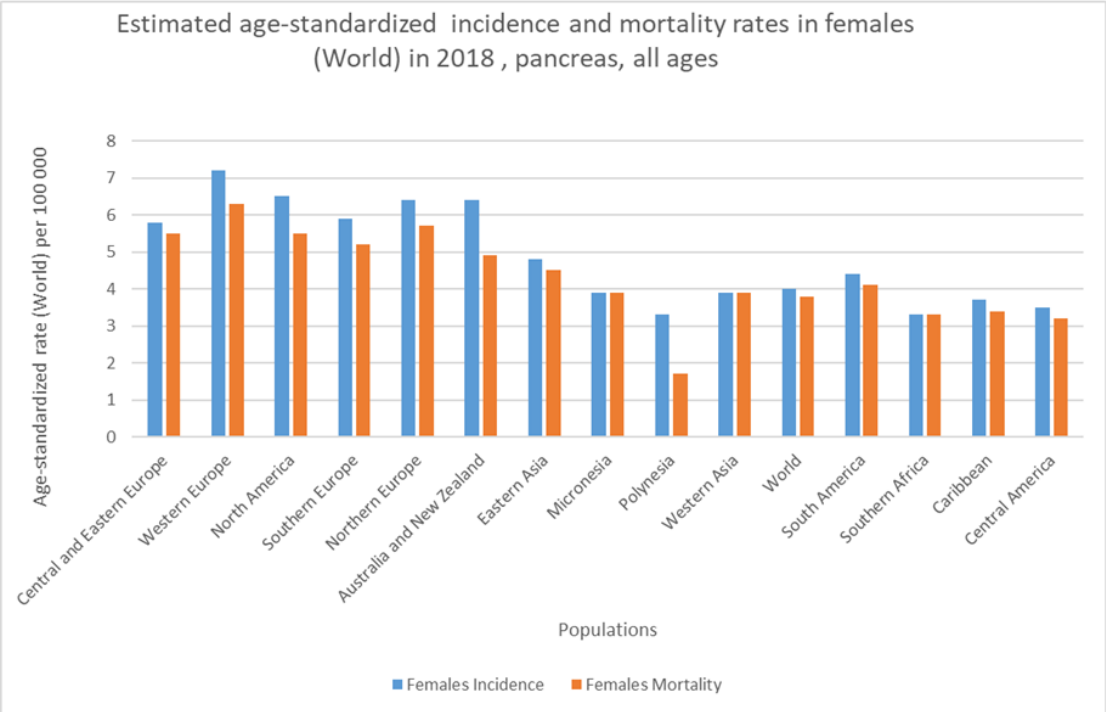


Figure 1. 2 Pancreatic cancer incidence and mortality in (a) males and (b) females in different populations, adapted from GLOBOCAN 2018 estimates.

1.1.2 Diagnosis and Treatment

Pancreatic cancer is diagnosed according to Tumour, Node, Metastasis (TNM) staging and imaged using Computed Tomography (CT) and Magnetic resonance imaging (MRI) (Mohammad, 2018). Carbohydrate antigen, CA 19-9 is associated with tumour progression, is a useful biomarker for diagnosis in PDAC patients (Park *et al.*, 2013). However, it was reported that CA 19-9 cannot be used as a screening tool for an asymptomatic population suffering from PDAC (Kim *et al.,* 2004). CA19-9 can be used during PDAC management to assess the response to treatment and to predict the prognosis of the disease (Park *et al.*, 2013).

The poor survival rate of PDAC, depicted in Figure 1.2, is partly due to the late diagnosis of cancer when it has already metastasised to neighbouring lymph nodes and organs such as the liver (Dong *et al.,* 2018). A systematic review on survival times of PDAC patients showed that the general median survival time after diagnosis was 6.1 months for patients in median ages between 62-67 years but was much shorter for the elderly population (1 to 3.2 months) (Carrato *et al.,* 2015). PDAC is asymptomatic in the early stages but the symptoms only appear in the advanced stage of the disease, so the diagnosis is often delayed (Dong *et al.,* 2018). Hence, PDAC is known to be a silent killer due to its asymptomatic nature in the early stages of the disease (Rahman & Washington, 2019). The presentation of non-specific symptoms that appear in the later stage of PDAC include weight loss, jaundice, dark urine, light stool colour, abdominal pain and nausea (Rahman & Washington, 2019).

Typically, patients diagnosed with PDAC are treated with conventional methods such as chemotherapy, radiation, combination therapy and surgery (Rossi *et al.,* 2014). A phase III study showed a combination of chemotherapeutic drugs, nab-paclitaxel plus gemcitabine was effective as it increased the median survival time (8.5 months) than gemcitabine alone (6.7 months) (Von Hoff *et al.,* 2013). However, numerous studies have shown PDAC is resistant to current therapies. The only potential cure for PDAC and extended survival rate is surgery but it is ineffective when cancer has progressed to the advanced stage. Patients who underwent pancreatectomy had a higher median survival rate by up to 10 months compared with non-resected patients (Bilimoria *et al.,* 2007). The poor prognosis

of PDAC is also due to the reoccurrence of this disease even after surgery. There had been no shift in the poor prognosis of PDAC over the past 40 years so there is a need to investigate novel therapeutic drugs to improve survival rate and to reduce reoccurrence rates (Tingle *et al.,* 2015).

1.2 The stroma and its importance in tumour development

There are multiple reasons why chemoresistance develops in PDAC which includes the tumour microenvironment, cancer stem cells and non-coding RNAs (Figure 1.3). A typical feature that surrounds PDAC is the dense stroma known as desmoplasia (Farrow *et al.*, 2008). The stroma accounts for 90% of the tumour volume of PDAC (Xie & Xie, 2015). It consists of the non-cellular component which is the extracellular matrix (ECM) which contains proteoglycans, hyaluronic acid, and structural proteins such as collagen, fibronectin and laminin; it includes soluble factors such as growth factors, chemokines, cytokines, antibodies, metabolites and cellular compartment: stromal cells, neural cells, vascular cells and inflammatory cells (Werb & Lu, 2015). In the tumour microenvironment, there is a bi-directional interaction of stromal and cancer cells that contribute to chemoresistance, thus reducing the efficacy of the current chemotherapeutic drugs in PDAC (Neumann *et al.,* 2018).

Cancer stem cells has been identified in PDAC and has the capacity to renew itself and initiate tumour growth during metastasis and result in chemoresistance (Li *et al*., 2007; Gnanamony & Gondi, 2017). Cancer stem cells have high expression of c-Jun N-terminal kinases (JNK) which is important in developing chemoresistance to two important chemotherapeutic drugs, fluorouracil and gemcitabine (Suzuki *et al.,* 2015). Suzuki *et al.* (2015) showed that knockdown of JNK allowed the sensitisation of cancer stem cells to chemotherapeutic drugs and induce reactive oxygen species production.

Non-coding RNA has a role in translation, and different processes including DNA replication, splicing, and epigenetic regulation (Taucher *et al.,* 2016). Micro-RNA is a type of non-coding RNA involved in regulating gene expression and has a role in chemoresistance to current therapeutic drugs for PDAC (Gnanamony & Gondi, 2017). A study has shown that micro-RNA-17-92 cluster consists of six members of micro-RNA which is downregulated in pancreatic cancer stem cells compared with the differentiated

counterpart and overexpression of this micro-RNA in cancer stem cells led to reduction in self-renewal capacity, hence, sensitivity to chemotherapeutic drugs (Cioffi *et al.,* 2015).

Chemoresistance of PDAC is also due to the heterogeneity of the disease. PDAC has a complex mutational landscape from studies that carried out exome and copy number variation (CNV) analysis as well as whole genome studies (Biankin *et al.,* 2012; Waddell *et al.*, 2015; Murphy *et al.,* 2016). The mutational landscape of PDAC is due to multiple somatic mutations of different oncogenes and tumour suppressor genes (Hanahan & Weinberg, 2011; Felsenstein *et al.*, 2018). However, the five major genes that are well characterised in PDAC include KRAS which is an oncogene, and tumour suppressor genes such as TP53, SMAD4, CDKN2A and ARID1A according to the COSMIC database (Heestand & Kurzrock, 2015).

Many key targeted therapies have been tried out in clinical trials (Lai *et al.,* 2019). Molecular targets for the main mutated genes (TP53, SMAD4 and CDKN2A) of PDAC was shown to have poor prognosis (Oshima *et al.,* 2013). MEK is a protein kinase which is a key downstream target of KRAS signalling (Pasca Di Magliano *et al.,* 2013). Trametinib is a MEK inhibitor was used in a randomised phase II trial in combination with gemcitabine but it did not prolong the median overall survival (OS) of PDAC patients with metastatic disease (Infante *et al.,* 2014). Another potential therapeutic target for PDAC is gamma secretase inhibitor targeting the driver gene, GATA6, involved in Wnt/Notch signalling (Zhong *et al.,* 2011; Mizuma *et al.*, 2012).

The stromal components enable the formation of a complex microenvironment that facilitates tumour development (Yuan *et al.,* 2019). A study by Hwang *et al.* (2008) showed that there is a crosstalk between human pancreatic stromal cells (HPSC) and pancreatic cancer cells to promote progression of the tumour in PDAC and the importance of stroma in inhibiting the response of chemotherapy and radiation of tumour cells. This shows that the tumour microenvironment is dynamic which aids in tumorigenesis and chemoresistance (Fujita *et al.*, 2009).

As there is a growing interest in the tumour microenvironment, several studies of anti-stromal therapeutics have been tested preclinically and some in early human clinical trials. A study by Özdemir *et al.* (2014) depleted the stroma *in vivo* which lead to tumour

invasion and poor survival. This shows the stroma does have tumour suppressive properties and shows the complexity of the stromal components. It is important to note that many proteins within the ECM are affected by splicing that has been linked to cancer progression. The next section illustrates how splicing takes place in the cell.

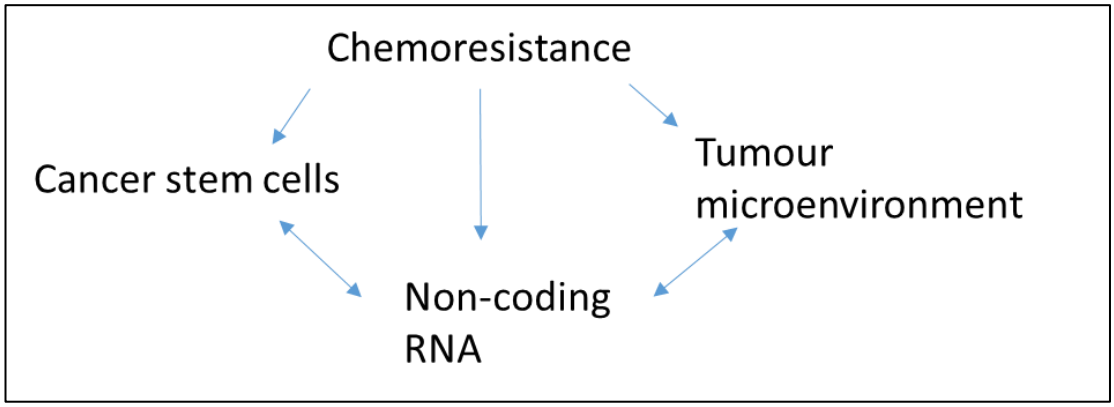


Figure 1. 3 Different factors linked with chemoresistance of pancreatic cancer.

An interplay between cancer cells, the tumour microenvironment, cancer stem cells which are controlled by small non-coding RNAs. Image adapted from (Gnanamony & Gondi, 2017).

1.3 RNA processing

Eukaryotic nascent RNA is synthesised by RNA polymerase II. It undergoes four main processing steps to become mature mRNA including 5' capping, 3' end cleavage and polyadenylation and splicing.

1.3.1 5' capping

Eukaryotic mRNA possess a 5' cap structure (Topisirovic *et al.*, 2011). The 5' cap is covalently bonded with the mRNA by 5'-5' triphosphate bridge. The 5' capping occurs after the synthesis of ~20 bp of the nascent RNA by RNA polymerase II (Jurado *et al.*, 2014). The 5' cap is necessary for RNA processing, prevent degradation, help with nuclear export of mature transcript, stabilize eukaryotic mRNA and translation (Furuichi *et al.,* 1977; Shimotohnot *et al*, 1977).

1.3.2 3' end cleavage and polyadenylation

Most eukaryotic pre-mRNA undergoes endonucleolytic cleavage where the sequences at the 3' end are lost and the addition of polyadenylated tail (Poly (A) tail) at the 3' end of pre-mRNA (Wahle & Rüegsegger, 1999). The mammalian signal for polyadenylation is the consensus sequence AAUAAA which is positioned between 11 and 25 base pairs before

the cleavage and polyadenylation site (Cooke *et al.*, 1999). Downstream of the cleavage side resides a less conserved GU-rich element and poly (A) tail is added at the cleavage site (Zhao *et al.,* 1999). Polyadenylation is a highly coordinated process which happens alongside splicing (Cooke *et al.*, 1999).

1.3.3 Splicing process

Pre-mRNA splicing in eukaryotes occurs with the exclusion of introns from pre-mRNA and ligation of exons to give mature mRNA (Smathers & Robart, 2019) (Figure 1.4). As shown in figure 1.4, the basic mechanism is initiated by 2' hydroxyl of the intron branch point (adenine) which carry out a nucleophilic attack in the 5' splice site. This is the first transesterification reaction that leads to free 5' exon and an intermediate lariat-3' exon. The free 5' exon hydroxyl group subsequently carry out nucleophilic attack on the 3' splice site which is the second transesterification. This results in ligated exons with the lariat as a by-product. This is a regulated process which is mediated by the spliceosome, a ribonucleoprotein complex.

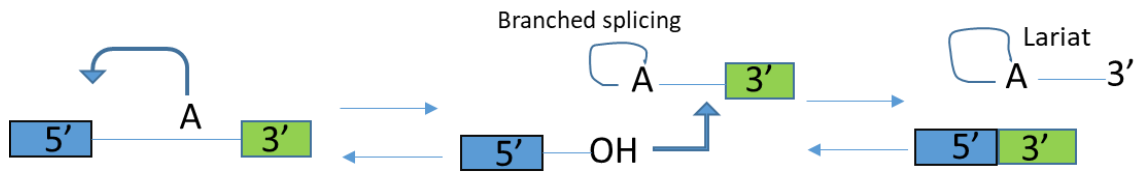


Figure 1. 4 Basic splicing mechanism of pre-mRNA consists of two transesterification reactions to produce ligated exons and intron lariat, adapted from (Smathers & Robart, 2019).

1.3.4 Splicing mechanism

The mechanism of splicing is well characterised, and it is a regulated process which is mediated by the spliceosome. Figure 1.5 depicts the detailed stages of the splicing process (Chen & Manley, 2009). The spliceosome assembly is initiated with the binding of small nuclear ribonucleoproteins U1 (snRNP U1) to the 5' splice site and the branch point is bound by splicing factor 1 (SF1). The U2 auxiliary factor (U2AF) heterodimer binds to both polypyrimidine tract and 3' terminal AG. This results in the E complex formation which is adenosine triphosphate (ATP) independent. The U2 snRNP replaces SF1 at the branch point resulting in A complex formation. This enables the addition of U4, U6-U5 tri-snRNP complex which forms the B complex. At this point, all the spliceosome components

conduct the splicing of pre-mRNA. The spliceosome undergoes conformational changes and remodelling where U1 and U4 are lost which subsequently results in catalytically active C complex which completes the splicing process.

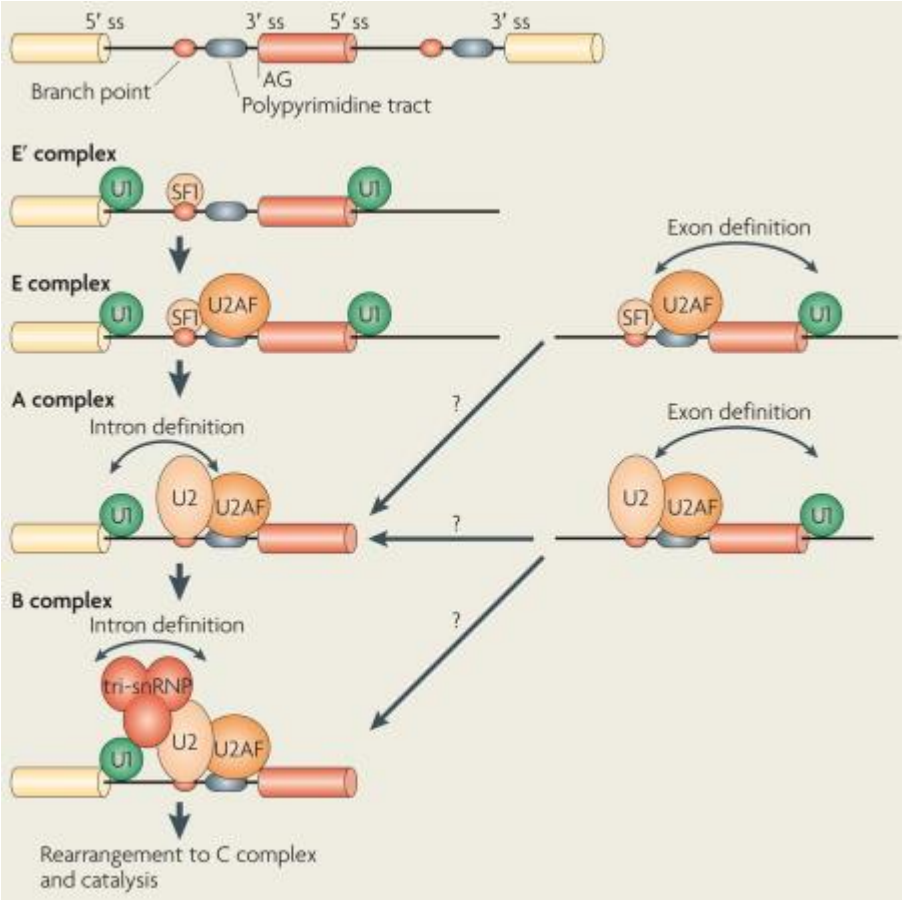


Figure 1. 5 Splicing mechanism using spliceosome.

Image acquired from (Chen & Manley, 2009) with permission granted. Intron definition is when spliceosome assembly leads to the removal of an intron. Likewise, exon definition is when the spliceosome assembly leads to exon skipping.

1.3.5 Alternative splicing

Section 1.4 describes constitutive splicing whereby all introns are spliced out resulting in the ligation of all exons of this gene. However, alternative splicing is a regulated process which gives rise to multiple mature mRNAs, some of which results in proteins isoforms with potentially different functions, from the same gene. Thus, increasing protein diversity. Alternative splicing results in different combination of exons and sometimes intron retention due to the selection of different splice sites (Blencowe, 2006)(Figure 1.6).

Alternative splicing has an important role in development and tissue differentiation (Yeo *et al.,* 2004).

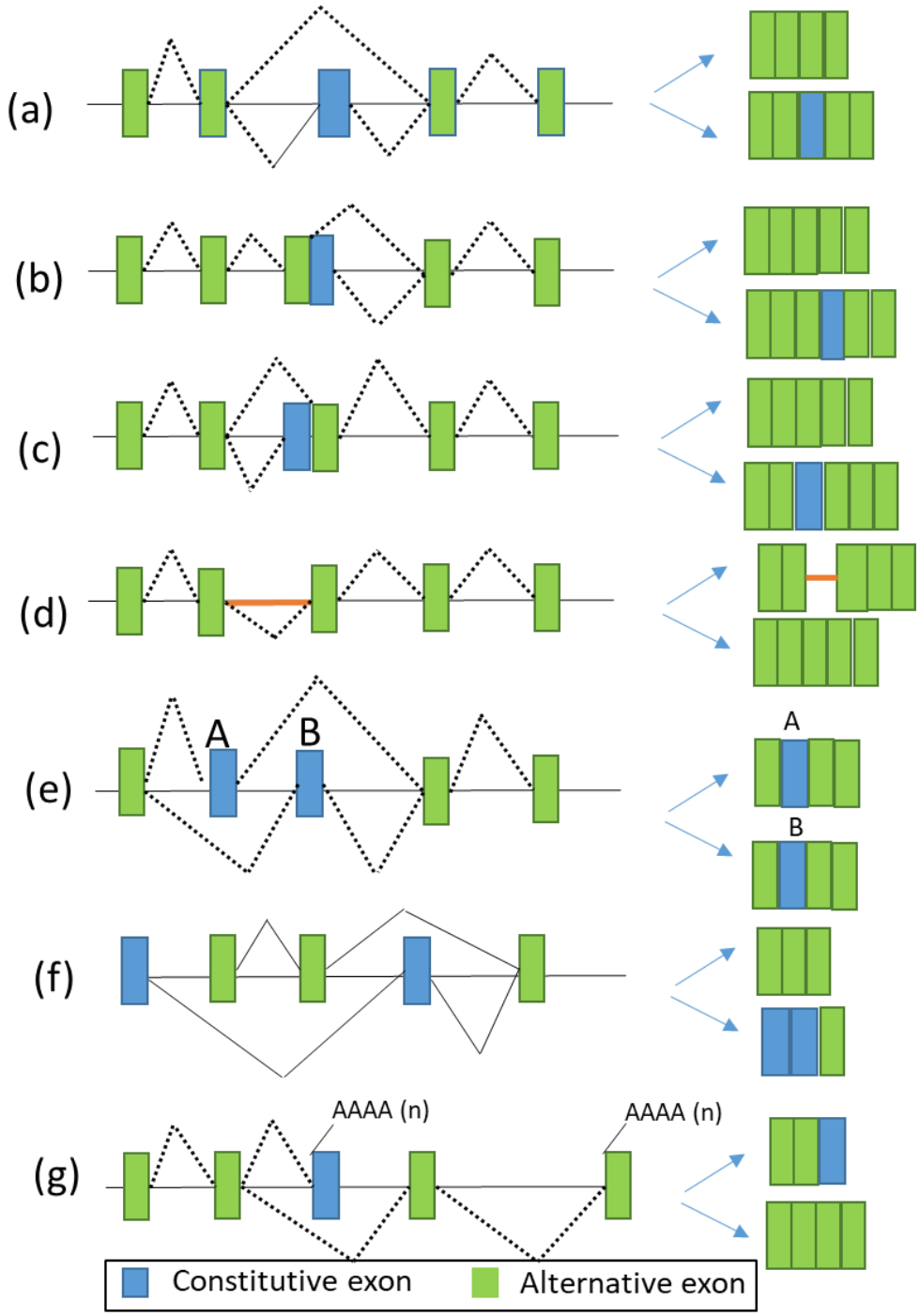


Figure 1. 6 Seven types of alternative splicing based on systemic analyses of expressed sequence tag (EST) and cDNA and microarray data.

(a) Cassette alternative exon (exon skipping). (b) Alternative 5' splice site. (c) Alternative 3' splice site. (d) Intron retention. (e) Mutually exclusive alternative exons. (f) Alternative promoter and first exon. (g) Alternative poly (A) site and terminal exon. Figure adapted from (Blencowe, 2006)

1.3.6 Regulation of alternative splicing

Studies have shown the complexity of the regulation of alternative splicing. Alternative splicing involves the production of splice variants which is a result of the interaction of trans-acting splice factors with the cis-regulatory elements in either exons or introns (Yeo *et al.,* 2004). Splicing regulatory elements in exons include exonic splicing enhancers (ESEs) or silencers (ESS) which dictates the fate of the exon either promoting or inhibiting it and likewise, splicing regulatory elements for introns include intronic splicing enhancer (ISE) and silencers (ISS)(Wang & Burge, 2008; Wang *et al.*, 2015). Two of the most characterised splicing regulatory factors include RNA-binding proteins which plays an important role in the control of alternative splicing include the serine/arginine-rich family of nuclear phosphoproteins (SR proteins) and heterogeneous nuclear ribonucleoproteins (HnRNPs) (Wang *et al.,* 2015). It is thought generally SR proteins bind to ESE and ISE to promote the splicing assembly whereas HnRNPs bind to ESS and ISS to oppose splicing at alternative splice sites (Chen & Manley, 2009; Kelemen *et al.,* 2013). Although the roles of SR protein and HnRNP are more complex than that. A study found HnRNP H can act both as a positively or negatively in the regulation of splicing (Rooke *et al.*, 2003). SR proteins are required to be phosphorylated otherwise they act as an inhibitor of splicing (Furuyama & Bruzik, 2002).

1.4 POSTN

Splicing affects many matricellular proteins (Viloria & Hill, 2016). Matricellular proteins are part of the non-structural components of the ECM and is involved in tissue development and repair and up-regulation of matricellular proteins have a role in many diseases including cancer (Sage, 2014). Periostin (POSTN), previously termed osteoblast-specific factor-2 (OSF-2), is a 93.3 kDa matricellular protein which was first found in periosteum in bone tissues (Takeshita *et al.,* 1993; Horiuchi *et al.*, 1999). POSTN protein structure consists of a conserved N- terminal region containing the signal peptide enabling the secretion into the ECM, an EMI domain followed by four homologous FAS I domains and terminating at the C-terminal region which is affected by alternative splicing (Horiuchi et al., 1999; Takeshita *et al.,* 1993). POSTN binds to several integrin receptors such as αVβ1, αVβ3, αVβ5 and α6β4 which activates the serine–threonine kinase/ protein kinase B (Akt/PKB) and focal adhesion kinase (FAK) signal pathways which demonstrates

the role in cell adhesion (Murakami *et al.,* 2017). POSTN has been shown to promote cell motility (Gillan *et al.,* 2002; Chuanyu *et al.,* 2017). Integrin binding occurs via FAS I domains 2 and 4 (Hoersch & Andrade-Navarro, 2010). It is thought that the C' terminal region of POSTN regulates the binding to ECM proteins such as collagen I, collagen V, tenascin-C and fibronectin (Takayama *et al.,* 2006; Norris *et al.,* 2007; Kii *et al.,* 2010). Hoersch *et al.* (2010) hypothesised that the C-terminal region of POSTN may mediate the binding of ECM proteins, although this has not been tested experimentally.

1.4.1 POSTN in different diseases

In the literature, elevated levels of POSTN have been observed in many inflammatory diseases and tissue fibrosis including in several respiratory diseases, liver disease and kidney disease (Wallace *et al.,* 2014; Huang *et al.,* 2015; Izuhara *et al.,* 2016; Bian *et al.,* 2019). In Asthma, POSTN is shown to be induced by interleukin- 4 and interleukin-13 from epithelial cells in the airways and bind with other ECM proteins such as collagen V and fibronectin, which results in subepithelial fibrosis, a feature that is present in bronchial asthma (Takayama *et al.,* 2006; Sehra *et al.,* 2011). Hence, studies have looked at POSTN serum levels as a useful biomarker for certain phenotype of asthma (Matsusaka *et al.,* 2015).

1.4.2 POSTN in cancer

Several studies have shown an overexpression of POSTN is linked to many types of cancers compared to normal tissue counterparts and has been associated with invasiveness, metastasis promoting angiogenesis (Table 1.1). This shows POSTN is a marker of tumour progression. However, findings from bladder cancer showed the opposite pattern as POSTN appeared to be downregulated in bladder cancer compared with normal bladder tissue (Kim *et al.,* 2005). This suggests that POSTN functions as a tumour suppressor in bladder cancers. This contradiction may be explained due to tissue-specific expression of POSTN isoforms due to alternative splicing in the C-terminal region which could explain the tumour suppressive effects of POSTN.

1.4.3 POSTN isoforms in cancer

Limited studies have looked at different POSTN splice variants in cancer. Table 1.2 shows the different POSTN isoforms found in the literature. Of these, four POSTN isoforms have been fully characterised: POSTN isoforms 1, 2, 3 and 4 (Table 1.2). The expression of four

POSTN isoforms namely isoforms 1, 2, 3 and 4 expressed in normal bladder tissues (Kim *et al.,* 2008). Morra *et al.* (2011) found renal fetal tissues revealed expression of POSTN isoforms 1, 2, 5, 6 and 7 whereas renal cell carcinoma and normal tissues expressed POSTN isoform 3, 4 and 8. This clearly shows the unique POSTN profile in embryogenesis. It demonstrates tissue specificity as normal bladder tissue had a different POSTN profile to normal renal tissue. In renal cell carcinoma, POSTN isoform 8 was seen more frequently than normal tissue suggesting a role in tumorigenesis. Similarly, POSTN isoform 4 was predominately observed in bladder cancer tissues and functional studies showed that POSTN isoform 4 had encouraged *in vitro* invasiveness and *in vivo* metastasis. It has been speculated that these isoforms linked with tumour progression may lack crucial domains in the C-terminal region (Viloria & Hill, 2016).

Table 1. 1 POSTN expression in cancer

Cancer type	POSTN expression	POSTN function	References
Gastric cancer	Upregulated POSTN expression in advanced gastric cancers	Promote metastasis and invasion. Promote xenogeneic tumour growth.	(Zhong, Li *et al.,* 2019)
Colon cancer	Upregulated POSTN expression	Promote metastasis in vivo and induce angiogenesis in metastatic growth. Activation of Akt and PKB survival pathway.	(Bao *et al.,* 2004)
Breast cancer	Upregulated POSTN expression	Promote angiogenesis.	(Shao *et al.,* 2004)
Ovarian cancer	Upregulated POSTN expression	Promote tumour angiogenesis and metastasis, decrease apoptosis *in vivo.*	(Gillan *et al.*, 2002)
Non-small cell lung cancer	Upregulated POSTN expression	POSTN highly associated with poor prognosis.	(Murakami *et al.*, 2017)
Cholangiocarcinoma	Upregulated POSTN expression	Promote proliferation and invasion.	(Utispan et al., 2010)
Esophageal cancer	Upregulated POSTN expression	Promote migration and invasion.	(Michaylira et al., 2010)
Head and Neck cancer	Upregulated POSTN expression	Promote invasion.	(Kudo et al., 2006)
Neuroblastoma	Upregulated POSTN expression	Higher POSTN mRNA expression levels linked to advanced stage of Neuroblastoma.	(Sasaki et al., 2002)
Renal Cell carcinoma	Upregulated POSTN expression	Higher POSTN correlated with advanced stage.	(Morra et al., 2011a)
Prostate cancer	Upregulated POSTN expression	Higher POSTN is associated with tumour aggressiveness.	(Tian et al., 2015)
PDAC	Upregulated POSTN expression	Promote invasion and motility.	(Baril et al., 2007)
Bladder cancer	Downregulation of POSTN expression	POSTN suppress cell invasiveness without altering cell proliferation and tumour growth *in vivo.* Suppress invasion and metastasis in human bladder cancer.	(Kim et al., 2005)

Table 1. 2 POSTN transcript variants encoded by *POSTN* gene. Acquired from (Viloria & Hill, 2016).

Common name (Morra and Moch, 2011)	Common name (Kim et al., 2008)	Ensembl variant	AA	TSL	Genbank ID	Exons present in alternatively spliced region
"Isoform 1"	"WT"	POSTN-001	836	1	NM_006475	17-21
"Isoform 3"	"Variant II"	POSTN-201	781	1	NM_001135935	18,19,20
		POSTN-003	809	1	NM_001286665	18-21
"Isoform 7"		POSTN-202	749	1	NM_001286666	20,21
"Isoform 2"	"Variant III"	POSTN-002	779	1	NM_001135934	19,20,21
"Isoform 5"		POSTN-004	808	5	NM_001330517	17-20
"Isoform 4"	"Variant I"	-			NM_001135936	19,20
"Isoform 8"		-			NM_001286667	20
"Isoform 6"		-			XM_005266232	17,19,20,21

Additionally, other studies have found POSTN isoforms in non-small lung cancer (Morra *et al.,* 2012). Fetal lung expressed eight POSTN isoforms including isoforms from 1-8 (Table 1.2), but also POSTN isoform 9 (not shown in Table 1.2 contain exons 17, 19 and 20). In this study POSTN isoform 1, 5 and 9 were unique to fetal lung tissue again suggesting POSTN isoforms are differentially expressed in embryogenesis. Five POSTN isoforms 2, 3, 4, 7 and 8 were detected in Non-small lung cancer tissue and corresponding normal tissue. This shows there was no difference in POSTN isoform expression profile but it does highlight the tissue-specific expression of POSTN. Similarly, in thyroid carcinoma, there was no difference in the POSTN isoform expression pattern between thyroid carcinoma and its corresponding normal tissue (Bai *et al.,* 2010). It has been speculated that despite the same isoforms being present in normal thyroid tissue and its cancerous tissue, altered expression levels of POSTN isoforms may play an important role in tumour invasiveness and metastasis (Bai *et al.,* 2010).

1.4.4 POSTN in PDAC

As shown in Table 1.1, POSTN is upregulated in PDAC. Numerous studies have shown POSTN is exclusively expressed by activated pancreatic stellate cells (Erkan *et al.,* 2007; Fukushima *et al*., 2008; Kanno *et al.*, 2008; Ben *et al.*, 2011; Liu *et al.*, 2016; Liu *et al.*, 2017). Hence, the pancreatic stellate cell line was selected in this study. There is evidence of interaction between pancreatic cancer cells, pancreatic stellate cells and POSTN (Figure

1.7). Erkan *et al,* (2007) showed that pancreatic cancer cells interact with pancreatic stellate cells to allow the secretion of POSTN and other ECM proteins. POSTN can activate pancreatic stellate cells in an autocrine manner to release more ECM proteins, TGFβ1 as well as POSTN itself thus creating a supportive tumour environment (Erkan et al., 2007). POSTN knockdown in pancreatic stellate cells *in* vivo showed a reduction in tumour growth which demonstrates POSTN acts as a tumour promoter in the progression of PDAC (Liu et al., 2017). Interestingly, a study by Fukushima *et al.* (2008) showed POSTN expression in the non-invasive stage of PDAC (intraductal papillary mucinous neoplasm) using immunohistochemistry. It has been speculated the interaction between pancreatic cancer cells and pancreatic stellate cells may happen in the non-invasive stage of PDAC.

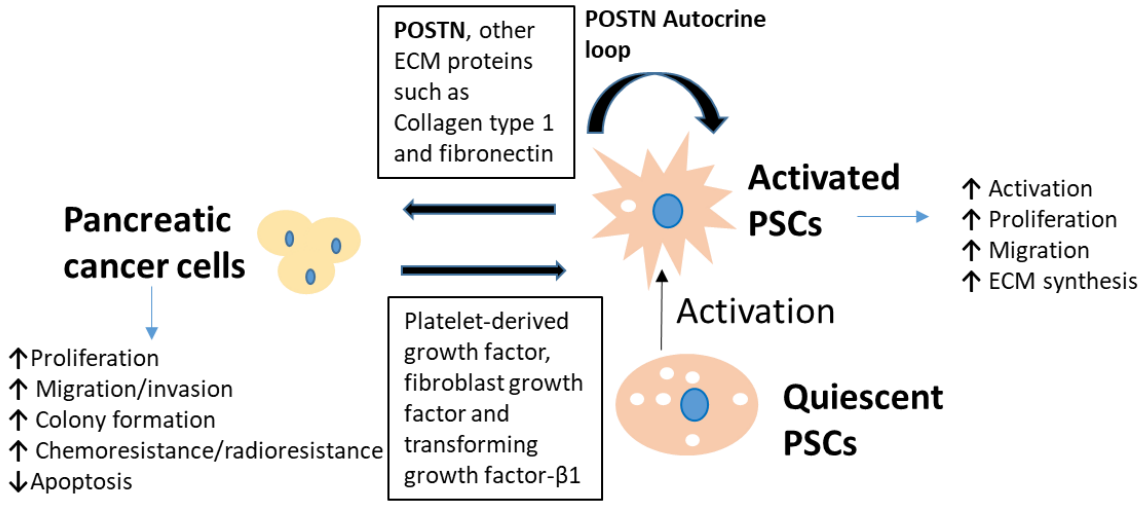


Figure 1. 7 Tumour supportive microenvironment of PDAC progression showing cross-talk between pancreatic cancer cell, pancreatic stellate cells (PSC) and POSTN.

Pancreatic cancer cells produce stimulants which activate quiescent PSCs into activated PSCs. Activated stellate cells secrete POSTN and other ECM proteins into the matrix creating a stroma-rich environment. POSTN creates an autocrine positive feedback loop on activated PSCs which perpetuates PSCs activation.

POSTN is shown to be involved in tumour progression of PDAC. POSTN has been found to promote angiogenesis in PDAC by activating the extracellular-signal-regulated kinase (ERK) pathway which promotes the proliferation, migration and invasion of endothelial cells (Liu *et al.*, 2016). POSTN promotes PDAC progression by regulating several signal transduction pathways. Erkan *et al.* (2007) reported that POSTN can stimulate the proliferation of pancreatic cancer cells under serum deprivation and hypoxia. At the molecular level, Baril *et al.* (2007) demonstrated that POSTN binds to specific integrin

receptor α6β4 in PDAC cells and starts signalling pathways such as P13 kinase pathway which promotes the activation of FAK and AKT. POSTN has been shown to facilitate the adhesion, migration and survival of PDAC cells (Ben *et al.,* 2011; Liu *et al.,* 2017).

Conversely, a study by Kanno *et al.* (2008) showed POSTN have tumour suppressive effects on PDAC which is inconsistent with the previous findings. This may be due to biphasic effects of POSTN in PDAC where Kanno *et al.* (2008) demonstrated low POSTN concentration (100 ng/ml) inhibited migration whereas high POSTN concentration (1 µg/ml) increased the migration of pancreatic cancer cells. Interestingly, phosphorylation of AKT occurred when higher concentration (1µg/ml) of POSTN was used. Similarly, this biphasic effect of POSTN was observed on signal pathway AKT where low concentration (100 ng/mL) decreased AKT phosphorylation but the opposite happened when POSTN concentration was increased (1 µg/mL). Interestingly, Kanno *et al.* (2008) found Panc-1 cells expressing POSTN did not metastasise to the liver compared with control which suggests that POSTN may inhibit spreading of PDAC.

1.5 Aims and Objectives

To date, no studies have looked at POSTN isoform expression in PDAC. Thus, this study aims to investigate POSTN transcript variants expression in pancreatic cells using semi-quantitative reverse transcription-polymerase chain reaction (RT-PCR). One hypothesis is that POSTN isoforms likely to have different matrix binding properties due to alternative splicing in the C-terminal region and to differentially affect cell adhesion and migration (Viloria & Hill, 2016). To test this novel hypothesis, the second aim was to carry out mechanistic analysis of these POSTN isoforms by producing recombinant proteins via traditional cloning method.

Chapter 2- Materials and methods

2.1.1 Bioinformatics Analysis of POSTN Isoforms

Human periostin (POSTN) transcript variants with complete coding sequence (CDS) were selected for bioinformatics analysis. The six POSTN transcripts were: *POSTN* 203 (ENST00000379747), *POSTN* 201 (ENST00000379742), *POSTN* 202 (ENST00000379743), *POSTN* 204 (ENST00000379749), *POSTN* 209 (ENST00000541179) and *POSTN* 210 (ENST00000541481). The corresponding old names for these variants can be found in figure 3.1. Ensembl database (https://www.ensembl.org/index.html) was used to download both exon and complementary DNA (cDNA) sequences of human POSTN transcripts in the RTF and FASTA file formats, respectively (Accessed: October 2015). These exon sequences of the individual transcript were used to help annotate exons in the cDNA FASTA sequence by highlighting each exon with a distinct colour. This allowed to map out all exons of the transcripts. These transcripts sequences were used to design RT-PCR primers to detect a specific region of the transcript. Expected PCR product sizes were calculated by mapping both forward and reverse primers on these transcripts.

Additional POSTN transcripts were identified in Genbank (https://www.ncbi.nlm.nih.gov/genbank/) were included in the analysis: *POSTN* 201b (NM_001135936) and *POSTN* 202b (NM_001286667). The names (*POSTN* 201b and 202b) were given as their sequences were identical to *POSTN* 201a and *POSTN* 202a, respectively but they had missing one exon (see Figure 3.2). Further analysis was carried out using AceView database (www.ncbi.nlm.nih.gov/IEB/Research/Acembly) to make comparison with POSTN transcript variants from Ensembl and Genbank. Additional potential intracellular forms of POSTN were identified namely *POSTN* g and *POSTN* i.

The following domain databases were used to predict the domains of POSTN protein variants (Table 2.1). Individual predicted POSTN protein sequences were put into the query box for the prediction of potential domains.

Table 2. 1 Domain database used to predict potential domains of POSTN protein

Database	URL
Expasy Prosite	https://prosite.expasy.org/
Uniprot KB	https://www.uniprot.org/uniprot/Q15063
SignalP-5.0 Server	http://www.cbs.dtu.dk/services/SignalP/

2.1.2 Primer Design

Primers were designed to detect POSTN transcript variants using 203 POSTN cDNA sequence from Genbank (accession number: NM_006475.2) in Primer-BLAST (https://www.ncbi.nlm.nih.gov/tools/primer-blast/) which were subsequently ordered (Sigma oligos). Specific parameters were set when designing the primers and the primer specificity stringency were adjusted to ensure the primers were specific to the desired template. Specific primers were designed to span exon-exon junction on at least one primer to eliminate any genomic DNA amplification. In certain experiments, exon spanning primers were not required.

2.2 Cell lines and cell culture

2.2.1 Cell lines

Several cell lines were used for this project were given by several collaborators. The human pancreatic stellate cells (PS-1) were given by Prof. Hemant Kocher, Barts Cancer Institute at Queen Mary University of London. PS-1 cells were originally isolated from the donated human pancreas and were immortalised (Froeling et al., 2009). The breast cancer cell lines used were MCF-7 (Soule *et al.*, 1973) and MDA-MB-468 cells (Brinkley et al., 1980) which were generously given by Prof Helmout Modjtahedi and Dr Athina Myrto-Chioni at Kingston University, respectively.

2.2.2 Cell culture media

PS-1 cells were grown in complete media using 500 mL RPMI -1640 (Gibco) supplemented with 2mM L-glutamine, 100 U/mL penicillin, 100 µg/mL streptomycin and 10% Fetal Bovine Serum (FBS). Breast cancer cell lines: MCF-7 and MDA-MB-468 cells were grown in DMEM (Gibco) with the same supplements as the RPMI-1640. These cells were cultured

in T-75 or T-25 flasks and grown in an incubator at 37 °C and 5% CO_2 under humidified conditions.

2.2.3 Cell cryopreservation and thawing cells out of liquid nitrogen

Cells with low passage number were put in the cryovials with the correct media and DMSO. Samples were stored in a Mr Frosty container with isopropanol at -20 °C for 2 hours. They were moved to -80 °C for 24 hours and for long-term storage, the individual cryovials were kept at -196 °C in liquid nitrogen. Cells needed for an experiment are taken out of the liquid nitrogen to be thawed. 1 mL of complete medium was added to the vial and was transferred in 25 mL universal tube. Either 10 mL (T25 flask) or 30 mL (T75 flask) of complete media was added gradually in the universal tube. This cell suspension was added in a T25 or T75 flask. The flask was incubated at 37 °C and 5% CO_2 under humid conditions. After overnight incubation, the old medium was removed and replaced with fresh complete medium without passaging.

2.2.4 Cell passaging

Cells were passaged every 2-3 days when the cell confluency reached to 70-100%. When the cells were passaged from T-25 flask, the old media was removed and 5 mL PBS was added to remove residual media which may hinder the activity of trypsin-EDTA. 3 mL trypsin-EDTA (0.05%) was added to allow the adherent cells to detach from the flask and was incubated at 37 °C with 5% CO_2 for 3-4 minutes (time dependent on cell type). The cells were viewed using a bright-field inverted microscope to see floating cells to show the complete detachment of the cell. 7 mL of fresh media was added to deactivate trypsin-EDTA. Cells were split 1 in 10 for PS-1, MCF-7 and MDA-MB-468 cells in a new flask with fresh complete media. Cells up to 10 passages were used in experiments.

2.2.5 Cell counting

Cells were first detached from flasks described in section 2.2.4. The cell suspension was aspirated from the flask into a 25 mL universal tube and was centrifuged at 1000 RPM for 3-4 minutes. The supernatant was aspirated without disturbing the pellet. The pellet was disrupted by flicking at the bottom of the tube. The cells were resuspended in 1 mL fresh complete media. The solution was mixed well so it was homogenous. 10 µL of cell suspension was transferred into 2 mL Eppendorf tube for cell counting. 40 µL of trypan blue (Fisher Scientific) was added to the Eppendorf tube which was used to detect viable

and dead cells. Dead cells are stained blue. 10 µL of the resulting mixture was added into the cell counter slide (#BSV100H, Immune Systems). The 10X objective lens of the haemocytometer was used to view the 2 x 5 grid. Three out of ten squares on the haemocytometer were counted. The cell numbers were averaged and the concentration of the cells per mL was calculated using the formula below:

$$cells/ml = Average\ no.of\ cells\ x\ Dilution\ factor\ x\ 10^4$$

2.2.6 RNA isolation and cDNA synthesis

Cells were cultured and total RNA was extracted from PS-1, MCF-7 and MDAMB-468 cells according to the protocol provided from RNeasy mini kit (Qiagen). Approximately, 1 x10 ^{7} cells were used which was determined by cell counting as described in section 2.2.5 or when the cells in the T25 flask were 70% confluent. These cells were trypsinized and the addition of media enabled the deactivation of the trypsin. The cells were transferred in a universal tube and were centrifuged at 1200 RPM for 3-4 minutes. The supernatant was aspirated. The pellet was disrupted and resuspended in 1 mL of media before it was transferred in a 1.5 ml Eppendorf tube. The cells were centrifuged at 6000 RPM for 2 minutes and the supernatant was aspirated. The cell pellet was resuspended in a mixture of 600 µl RLT lysis buffer and 10 µl β-mercaptoethanol as mentioned in the manufacturer's protocol. An on-column DNase digestion was included in the protocol to eliminate any genomic contamination. This included the addition of 70 µl of RDD buffer and 5 µl of DNase I master mix. The RNA purity (A260/230 and A260/280 ratios) and concentration (ng/µl) was measured using the Biodrop spectrophotometer (Fisher Scientific). The unused extracted RNA was stored at -80 °C.

The RNA volume needed in the reverse transcription reaction was calculated using the formula below to synthesise 500 ng of cDNA. The first-strand cDNA of each cell lines was synthesised using reverse transcriptase for reverse transcription reaction according to the manual of RevertAid First Strand cDNA Synthesis kit (Thermo Fisher Scientific). Each cell line had reverse transcriptase (RT) which had all the components of the reaction and two negative controls were no reverse transcriptase (NRT) and no template control (NTC). NRT control had all the components of the reaction apart from the reverse transcriptase enzyme, which was replaced with dH_2O. Similarly, the NTC had everything apart from RNA, again replaced with water. These controls were used in the semi-quantitative RT

PCR reaction. All cDNA templates were diluted 1 in 10 from the stock concentration 25 ng/µL before using it the RT PCR reaction.

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

2.2.7 Semi-quantitative Reverse transcriptase-polymerase chain reaction

POSTN PCR products were amplified by semi-quantitative reverse transcriptase-polymerase chain reaction.

QARS primers: 5'-CAGGGGTCTGGCTTATGTGT-3' (forward primer) and 5'-CTCTGAAAACTTGCCCTTGC-3'(reverse primer) (Sigma oligos) was used as a housekeeping gene which acts as a positive control. Many factors inhibit PCR. The positive control shows the PCR product can be amplified. PCR was confirmed when an expected product size of 150 bp was detected in various cell lines on a gel. All PCR reactions had negative controls include NRT and NTC as described in section 2.2.6. Also, a dH_2O control is used to rule out any contamination from the reagents by substituting cDNA template with water in the PCR reaction.

The RT PCR reaction was set up as shown in table 2.2. The protocol was adapted from the manual from DreamTaq Green PCR Master Mix (Thermo Fisher, K1081). As shown in table 2.2, the 1X concentration of DreamTaq Green PCR Master Mix was used. It consists of DreamTaq DNA polymerase, 2X DreamTaq Green buffer, free nucleotides (dATP, dCTP, dGTP and dTTP), 0.4 mM each, and 4 mM $MgCl_2$. A 100 µM stock concentration of primers (Sigma Aldrich) were made by resuspending it in nuclease-free dH_2O according to the datasheet provided. 1 in 10 dilutions was made of these primers to make 10 µM aliquots. RT PCR was performed under the following conditions shown in table 2.3 using Thermal Cycler (Applied Biosystems Veriti). The annealing temperature was adjusted according to the primer pair used.

Table 2. 2 RT PCR reaction mixture using DreamTaq Green PCR Master mix

Components	Final concentration
DreamTaq Green PCR Master Mix (2X)	1X
Forward primer (10 µM)	0.5-1 µM
Reverse primer (10 µM)	0.5-1 µM
Template cDNA	6.25 ng - 25 ng
Nuclease-free dH_2O	Up to the reaction volume

Table 2. 3 Thermocycling conditions for PCR using Dreamtaq PCR master mix (2X)

Step	**Temperature (°C)**	**Number of Cycles**	**Time**
Initial Denature	95	1	2 min
Denature	95	35	30 s
Annealing	**(56°C for QARS primer)**		30 s
Extension	72		30 s
Final Extension	72	1	5 min
Storage	4		∞

2.2.8 Agarose gel electrophoresis and analysis

Table 2.4 shows the buffer stocks composition. From the 50 X stock, a litre of 1 X TAE buffer was made by taking out 20 mL of 50 X TAE buffer and diluting it in 980 ml of dH_2O. Agarose gels were made by weighing out different amounts of agarose powder in either 50 mL TAE buffer or 100 ml TAE buffer, which is dependent on the percentage of the gel and size of the gel casting tray. The agarose and TAE mixture was heated in the microwave to dissolve the agarose powder in a beaker. GelRed Nucleic Acid stain (10,000 X) was added to the molten agarose and was diluted 1 in 10,000. Hence, 5 µl of Gel Red was added in 50 ml of molten agarose. The casting tray was prepared by double taping both ends using autoclave tape and adding the gel comb (BIO-RAD) to form wells. The molten gel was poured into the casting tray, ready to be solidified. Between 10 µL to 50

µL of RT PCR products were loaded into the wells of the agarose gel. The well size was determined by the thickness of the gel and the size of the comb. Between 3 µL to 5 µL of GeneRuler 100 bp DNA ladder (SM024, Thermo Scientific) was loaded into the well as a molecular weight marker for product sizes ranging from 100 bp to 1500 bp.

Table 2. 4 Material used for gel electrophoresis

1 X TAE buffer was used as a running buffer and to make different % w/v agarose gel.

Buffer	Components	Volume or Weight
50 X TAE	Tris Base Glacial acetic acid EDTA	242 g 57.1 ml 100 mL of 0.5 M Or 19.66 g
1 X TAE	Tris Base Glacial acetic acid EDTA	2.4 g 0.57 g 1 ml of 0.5 M EDTA

The gel was run between 60-100 V for 80 minutes (BIO-RAD powerpack). The running time was dependent on the size of the expected product size for the separation of the bands. Larger product sizes were run longer than smaller PCR products. Once the bands were separated, the gel image was visualised for semi-quantitative analysis of the PCR product using the Syngene G-Box via Genesys software.

2.2.9 Gel purification

After gel electrophoresis, PCR products that were sent for Sanger sequencing were excised from 1% UltraPure Low Melting Point agarose gel under UV light in a dark room. It was collected in a sterile Eppendorf tub. Bands were extracted and purified using QIAquick Gel Extraction Kit (Qiagen). In the final step of the protocol, 30 µl of Elution buffer was added rather than of 50 µl for increased DNA concentration.

2.3.0 Sanger sequencing and analysis

After gel extraction procedure, the Biodrop was used to measure DNA purity and yield. The purified PCR products were sent off to Genewiz along with the corresponding forward and reverse primers for Sanger sequencing. The results were accessed online on

the Genewiz website. Each forward and reverse sequence had a trace file and a sequence file. The individual trace files were viewed in Finch TV software (https://finchtv.software.informer.com/1.4/) to detect the quality of the sequencing. If the sequence read "N", it means the base pair could not be detected. However, the trace file can be edited where there is "N" if the signal is strong for a particular base pair. Once the sequence trace was edited, the reverse sequence was reverse complemented to allow the forward and reverse sequences to be aligned together using Clustal Omega alignment tool (https://www.ebi.ac.uk/Tools/msa/clustalo/). The start of the reverse sequence contains the forward primer. Conversely, the end of the forward sequence has the reverse primer. This gave a composite sequence which was pasted on BLASTn (https://blast.ncbi.nlm.nih.gov/Blast.cgi?PAGE_TYPE=BlastSearch) search box to check for the best-matched transcript. Although optional, it is possible to restrict the search to pick up specific organism such as *Homo sapiens.* Under the program selection, highly similar sequences (megablast) was selected and blasted. The composite sequence was then aligned against the expected transcript to check for similarity.

2.3.1 Western blot

2.3.2 Sample preparation

PS-1 cells were cultured in a T-75 flask. The medium was removed from the flask and PBS was added to the flask to wash the cells and then removed from the flask (# 14190, Life Technologies). 500 µl Radio immune Precipitation Assay (RIPA) buffer was added to the flask. Using a cold plastic scraper, the adherent cells were scraped and the cell suspension was transferred into an Eppendorf tube. It was incubated on ice for 20 minutes and was centrifuged at 1500 RPM for 5 minutes. The supernatant was transferred into a new Eppendorf tube. 10 µl of the sample was used for protein quantification using Bradford assay to calculate how much sample to load for 30 µg of protein. The cell lysate was added to 5X sample buffer (312mM Tris-HCL of pH 6.8, 10% Sodium dodecyl sulphate, 10% β-mercaptoethanol, 25% glycerol and 0.015% bromophenol blue) and boiled at 100 °C for 6 minutes and placed back on ice.

2.3.3 Loading and Gel electrophoresis

30 µg of cell lysate and protein molecular ladder were loaded in polyacrylamide gel (**12% Resolving gel**: 3.75 mL of 1M Tris pH 8.8, 3.6 mL 30 % (v/v) acrylamide, 50 µL 20 % (w/v)

SDS, 2.3 mL dH_2O, 12.5 µl TEMED and 300 µL 10% (w/v) APS. **Stacking gel**: 625 µL 1M Tris pH 6.8, 900 µL 30 % (V/V) Acrylamide, 25 µL 20% (W/V) SDS, 3.29 mL dH_2O, 12.5 µL TEMED and 150 µL 10 % (w/v) APS) which were placed on a Mini-Protean Tetra Cell (#165-800, BIO-RAD). Tris-glycine SDS running buffer was used to run the samples (30 g Tris, 145 g glycine, 10 g SDS) at 120 Volts for 90 minutes to ensure good separation of protein.

2.3.4 Semi-dry transfer, antibody probing and detection

After the running of the gel, the gel was sandwiched with filter papers soaked in transfer buffer (3 g Tris, 15 g Glycine, 200 mL methanol and dH_2O up to 1L) with nitrocellulose membrane (pre-soaked with transfer buffer) which was positioned on the gel. The sandwiched gel was laid in a Trans-Blot SD Semi-Dry Transfer Cell (#170-3940, BIO-RAD) at 15 Volts for 90 minutes. To visualise the protein in the membrane, it was stained with Ponceau S (P7170, Sigma-Aldrich). The Ponceau solution was washed by TBST (100 ml 10X Tris-buffered saline, 1 mL Tween 20, 900 mL dH2O). To avoid non-specific binding, the membrane was blocked using 5% skimmed milk in 10 mL of TBST for 45 minutes. The blocked solution was removed and primary antibodies were incubated at 4°C overnight: anti-POSTN rabbit polyclonal antibody (#ab92460, Abcam, 1:1000 dilution) and β-actin mouse monoclonal antibody (#ab8224, Abcam, 1:1000) as loading control with 5% BSA in TBST.

The next day, the solution was discarded from the membrane and washed thrice with TBST and agitated for 10 minutes per wash to ensure removal of residual primary antibodies. The membrane was incubated for 60 minutes with the secondary antibodies: IRDye 680RD goat anti-mouse IgG (#926-68070, Licor, 1 in 15 000) and IRDye 800CW goat anti-rabbit IgG (#926-32211, Licor, 1 in 15,000). The membrane was washed again with TBST before the membrane was ready for imaging. The membrane was viewed using the Odyssey CLx Infrared Imaging System (Licor). Image Studio was used to export the image and was used just for viewing the bands while no quantification was done.

2.4.1 Cloning of POSTN transcript variants

2.4.2 Ampicillin stock

To make ampicillin stock, 100 mg of ampicillin powder (#A0104, Melford) was dissolved in 1 mL of dH_2O to get an antibiotic concentration of 100 mg/mL. The antibiotic was stored in a -20 °C freezer.

2.4.3 Agar plate preparation

For agar plate preparation, 35 g of Lysogeny broth agar (LB agar) (35g/L, Sigma) was transferred to a 1 L glass bottle. dH_2O was topped up to 1 L and stirred before it was autoclaved. The molten LB agar was cooled down to 55 °C before adding 1 mL of ampicillin (100 µg/mL) under the fume hood. Approximately 30 mL of molten LB agar containing the antibiotic was poured into sterile petri dishes. Once the agar was solidified, these were placed upside down to prevent condensation and stored at 4 °C in the refrigerator wrapped in parafilm for up to two weeks before use.

2.4.4 Heat shock transformation

The vector pcDNA™4/TO (Invitrogen) was kindly provided by Dr Athina Myrto-Chioni at Kingston University. This plasmid has a gene that confers resistance to ampicillin. To extract sufficient plasmid, initial purified plasmid pcDNA4™/TO was transformed into chemically competent *E.coli* cells using the heat shock method as described in the protocol from Addgene (https://www.addgene.org/protocols/bacterial-transformation/). The competent XL-1 strain of *E.coli* bacteria was used for initial transformation.

The competent *E.coli* cells were taken out of the -80 °C freezer. 50 µl of chemically competent *E.coli* cells were thawed on ice for 20 minutes. It was added to 1 µl plasmid in 1.5 mL Eppendorf tube. This mixture was left on the ice to re-equilibrate for 5 minutes. The tube was placed in the water bath so that the bottom half of the tube was submerged in water at exactly 42°C for 45 seconds. After heat shock, it was placed back on the ice for an additional 5 minutes. The transformation mixture was suspended in 1 mL of fresh Lysogeny broth (LB broth) (25 g/L Miller's modification, Alfa Aesar) under the fume hood and incubated for 1 hour at 37 ^{o}C. LB agar plates were made using 30 mL LB agar (35g/L, Sigma) and 30 µl of ampicillin (100 µg/mL). Transformation mixture were centrifuged at 13 000 RPM for 10 seconds to form a pellet. Most of the supernatant was

discarded. The remaining 200 µL of LB broth was resuspended so the pellet was disrupted before steaking using a spreader in ampicillin-resistant plates under the fume hood. LB agar plates were placed in the incubator overnight at 37 ^{o}C.

2.4.5 Overnight culture of bacteria

Under the fume hood, an overnight culture was inoculated using 5 mL of LB broth (25 g/L Miller's modification, Alfa Aesar) containing 5 µL of ampicillin (100 µg/mL). *E.coli* XL-1 strains containing the plasmids were scraped using a sterile tip from the plasmid glycerol stock and dropped in the sterile falcon tube. It was placed in a shaking incubator overnight at 37 ^{o}C.

2.4.6 Glycerol stocks for long term storage of plasmid

Glycerol stocks were prepared for long-term storage of bacteria containing the plasmids. An overnight culture was inoculated. Under the fume hood, 1 mL of LB broth (25g/l Miller's modification, Alfa Aesar) was added to sterile falcon tubes. 1 µL of ampicillin was added. A small tip was used to scrape off the surface of colonies of bacteria which contains the plasmid from the LB agar plate and dropped into the sterile falcon tube. The bacterial culture was incubated overnight in a shaking incubator at 37 ^{o}C. After 16-20 hours, the liquid turned turbid indicating bacterial growth. The liquid from the falcon tube was transferred into an Eppendorf tube and 5% v/v of glycerol (#A16205, Alfa Aesar) was added. These glycerol stock tubes were flash-frozen in liquid nitrogen for a few seconds and immediately stored at a -80 ^{o}C freezer. The glycerol added enabled the bacteria to remain stable in freezing conditions and kept the cells viable.

2.4.7 Plasmid extraction

When the bacterial culture was left overnight, the liquid turned turbid indicating bacterial growth. The tip was discarded from the falcon tube and 5 mL of the bacterial overnight culture was centrifuged at 4500 RPM for 5 minutes before plasmid DNA was extracted from the overnight cultures using Monarch® Plasmid Miniprep Kit (New England Biolabs). Purified plasmid DNA was quantified using Biodrop measuring the purity (A260/230 ratio and A260/280 ratio) and concentration (ng/µL) ready for restriction digestion.

2.4.8 *POSTN* gene amplification for cloning into pcDNA™4/TO

POSTN gene sequences were obtained from the Ensembl database (Accessed: October 2015). *POSTN* 203 (*POSTN-001*) sequence was used for designing cloning specific forward and reverse primers. Since *POSTN* 203 sequence is the largest transcript, the primers should amplify all the other splice variants. Integrated DNA technologies Oligoanalyzer tool was used to analyse primers for predicted melting temperatures and hairpin analysis of the cloning primers. In the initial experiments, the forward and reverse cloning primers had restriction sites of NotI-HF and XhoI, respectively. These cloning primers were redesigned to include Hind III-HF and ApaI sites in the forward and reverse primer respectively. Using these primers, semi-quantitative RT PCR was carried out to amplify the transcripts of *POSTN* splice variants using Dream Taq (2X) master mix. The components of the PCR reaction are listed in Table 2.1 and the thermal cycling conditions are reported in Table 2.5. The PCR products were analysed on 0.8%-1.5% w/v agarose gels stained with Gel Red (10,000X) as described in section 2.2.8 to visualise the RT-PCR products.

Table 2. 5 Thermocycling conditions for PCR using Dreamtaq PCR master mix (2X). It was adapted from HotStar HiFidelity PCR Handbook (Qiagen) to amplify *POSTN* splice variants using cloning primers.

Step	Temperature (°C)	Number of Cycles	Time
Initial Denature	95	1	5 min
Denature	94	**32**	15 s
Annealing	**68**		60 s
Extension	68		5 min
Final Extension	72	1	10 min
Storage	4		∞

2.4.9 Quick Cip phosphatase digestion

Digested plasmids were treated with Quick Cip (Quick Dephosphorylation Kit M0508) which catalyses the dephosphorylation of 5’ ends containing phosphate groups. This reaction helps to prevent re-ligation of the digested plasmid. The samples were prepared according to table 2.6 and incubated at 37 °C for 10 minutes followed by endonuclease denaturation at 80 °C for 2 minutes. The required pmol of DNA ends was calculated. The

linear plasmid was converted from µg to pmol using Biomath Calculator (Promega) based on two variables: µg of DNA and DNA length (kb) (https://www.promega.co.uk/resources/tools/biomath/). The phosphatase was added directly into the digestion reaction during or after DNA digestion.

Table 2. 6 Quick Cip phosphatase digestion components and volumes

Solution	Volume
DNA	1 pmol of DNA ends*
CutSmart Buffer (10X)	2 µL
Quick Cip	1 µL
dH_2O	To 20 µL

*1 pmol of DNA ends is about 1 µg of 3 kb plasmid

2.5.0 DNA purification

Before the restriction digestion, QIAquick PCR purification kit was used to remove the enzymes from the PCR reaction (Qiagen) which enabled the purification of the amplified PCR products. It was quantified using the Biodrop (A260/230 and A260/280 ratios and concentration ng/µL). The purified plasmid was quantified after plasmid extraction. The digested plasmid and PCR products were purified again with Monarch® PCR & DNA Cleanup Kit to remove the enzymes and buffer.

2.5.1 Double endonuclease restriction digest of PCR products and plasmid

Restriction digests were conducted using the components from Table 2.7. Plasmid DNA and PCR product were digested using the following restriction enzymes pairs, NotI-HF and XhoI (NEB) or Hind III-HF and ApaI (NEB). Double Digest Finder tool on the NEB website was used to select the above restriction enzymes which picks the buffer that the enzymes are most compatible with. In this case, the 10X CutSmart Buffer (NEB) was used. Nuclease-free water was added to make up to 50 µL total of the reaction mixture. The reaction was incubated at 37 ^{o}C for 2 hours. The incubation temperature and time are enzyme-dependent. The restriction enzymes in the reaction mixture were heat-inactivated depending on the enzyme used see Table 2.8. Digested plasmids were visualised on 1% w/v agarose gel stained with Gel Red to validate the plasmid digests to see if the product size were at the expected band size.

Table 2. 7 Restriction digest of PCR products and plasmids

Solution	Concentration/Volume
DNA (Plasmid or PCR product)	1-3 µg
Restriction enzyme I	1 µL Or 10 units
Restriction enzyme II	1 µL Or 10 units
10 X CutSmart Buffer	5 µL (1X)
dH_2O	Up to the final volume of 50 µL

The table shows all the reaction components and volume for double digested insert and plasmids. CutSmart buffer (10X) was used for restriction enzymes NotI-HF and XhoI. CutSmart buffer (10X) also used for restriction enzymes Hind III-HF and ApaI.

Table 2. 8 Temperature used for restriction endonuclease heat inactivation

Restriction enzyme	Heat inactivation temperature (ºC)	Heat inactivation time (minutes)
NotI-HF	65	20
XhoI	65	20
Hind III-HF	80	20
ApaI	65	20

2.5.2 Ligation of plasmids and inserts

Ligation of pcDNA4™/TO plasmid with the insert was performed using T4 DNA ligase protocol (Promega), the formula below was used to calculate the molar ratios of plasmid and insert for the ligation reaction.

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

The following plasmid-insert molar ratios were used for ligation reactions 1:1, 1:3 and 3:1 (Table 2.9). The amount of plasmid was kept constant at 100 ng, only varying the amount of insert. The positive control was undigested plasmid and a negative control included the ligation reaction without the insert. The ligation reactions were incubated at room temperature for 3 hours (incubation temperature was optimised in this study), followed by transformation by the heat shock method.

Table 2. 9 Ligation reaction set up of plasmid vector and insert

Components	1:1 (Insert:plasmid)	1:3 (insert:plasmid)	3:1 (Insert:plasmid)
Insert DNA	X	X/3	3X
Vector DNA (ng)	100	100	100
10X T4 DNA ligase buffer (µl)	1	1	1
T4 DNA ligase enzyme (µl)	1	1	1
Nuclease-free water (µl)	Up to 10	Up to 10	Up to 10

2.5.3 Transformation of ligation mixture in *E.coli* competent cells

Heat shock transformation was used as described in section 2.4.4. However, a different strain of bacteria was used. *E.coli* DH5α competent cells (C2987H, New England Biolabs) were thawed on the ice for 10 minutes. Then 5 µl of ligation mixture was added to the competent cells and were incubated on ice for 5 minutes to re-equilibrate. The reaction was heat shocked in the water bath at 42 °C for 45 seconds and was placed back on ice for 5 minutes. 950 µl of fresh LB broth was added to the competent cells and were incubated at 37 °C for an hour shaking at 250 RPM to allow the bacteria to grow. Two-hundred µl of transformed bacteria were transferred on LB agar plates containing 100 µg/mL of ampicillin. The plates were placed upside down and incubated overnight at 37 °C.

2.5.4 Colony PCR to select successful transformants

Individual transformants were picked to be used as a template which was added directly to the PCR reaction with the master mix. Each PCR reaction was set up to 25 µl reaction consisting of 2x DreamTaq, CMV forward and BGH reverse primers (Table 2.10), topped up with dH_2O. Primers used in colony PCR are complementary to the sequences on either side of the multiple cloning site (MCS) of the vector. These primers can also amplify the gene insert if it was successfully ligated in the plasmid in the MCS where the restriction

sites of the plasmid are located. The colony PCR was performed using Applied Biosystems Veriti under the conditions shown in table 2.11. The initial ten minutes of the heating step allows the release of the plasmid DNA from the bacterial cell.

Table 2.10 show primers used to screen for successful transformants which contain plasmid with *POSTN* splice variants. Table 2.10 shows the annealing temperatures of both are starkly different. An annealing temperature of 55 °C was selected. The primers were obtained from pcDNA™4/TO manual (Invitrogen) which mark the multiple cloning site.

Table 2. 10 Primers anneal at the MCS of the plasmid used to screen successful transformants of *POSTN* splice variants. Primers were obtained from pcDNA™4/TO user guide (Invitrogen).

Primer	Sequence (5'-3')	Annealing temperature (ºC)
CMV	CGCAAATGGGCGGTAGGCGTG	70
BGH	TAGAAGGCACAGTCGAGG	56

Table 2. 11 Thermocycling conditions for colony PCR using DreamTaq PCR Master Mix (2X).

Step	**Temperature (°C)**	**Number of Cycles**	**Time**
Initial Denature	95	1	10 min
Denature	95	**34**	45 s
Annealing	**55**		45 s
Extension	72		60 s
Final Extension	72	1	5 min
Storage	4		∞

Chapter 3-Results: investigating the expression of POSTN transcript variants in pancreatic cells

3.1 Bioinformatics analysis of POSTN transcript variants

3.1.1 POSTN transcripts present in Ensembl, Genbank and AceView databases

Bioinformatics analysis was done on human *POSTN* gene located on chromosome 13. Table 3.1 gives a summary of the eight predicted POSTN transcript variants that have protein-coding potential across the three databases: Ensembl, Genbank and AceView. According to the Ensembl database, there are ten POSTN transcript variants present of which only six were identified as protein-coding for the *POSTN* gene (POSTN-001,002,003,004,201a and 202a). These variants were cross-referenced with Genbank which revealed two more variants: POSTN-201b and POSTN-202b.

Table 3. 1 Bioinformatics analysis of eight coding POSTN transcript variants of human POSTN gene encoding POSTN in Ensembl, Genbank and AceView databases.

Updated name (Ensembl, Oct 2017)	Old name (Ensembl Oct 2016)	AceView transcript designation	Genbank ID (RefSeq)	Transcript length (bp)	Predicted protein size (aa)	Molecular weight (kDa)	Flags (TSL)
POSTN-203	POSTN-001	A	NM_006475	3373	836	93.3	1
POSTN-201	POSTN-002	D	NM_001135934	3071	779	87	1
POSTN-202	POSTN-003	B?	NM_001286665	3196	809	90.4	1
POSTN-204	POSTN-004	N/A	NM_001330517.1	3210	808	90.1	5
POSTN-209	POSTN-201a	C	NM_001135935	3214	781	87.3	1
N/A	POSTN-201b*	E	NM_001135936	3135*	751*	83.9*	-
POSTN-210	POSTN-202a	F	NM_001286666	3113	749	83.5	1
N/A	POSTN-202b*	H	NM_001286667	3045*	721*	80.4*	-

All protein-coding POSTN transcript variants were selected from Ensembl, Genbank and AceView were cross-referenced and matched accordingly. The transcript length (bp), the predicted amino acids (aa), the predicted molecular weights and transcript support level (TSL) were obtained from the Ensembl database. * is marked on variants found in Genbank. All cDNA sequence data was originally exported from Ensembl in October 2015 using version 82. Abbreviation: TSL, transcript support level.

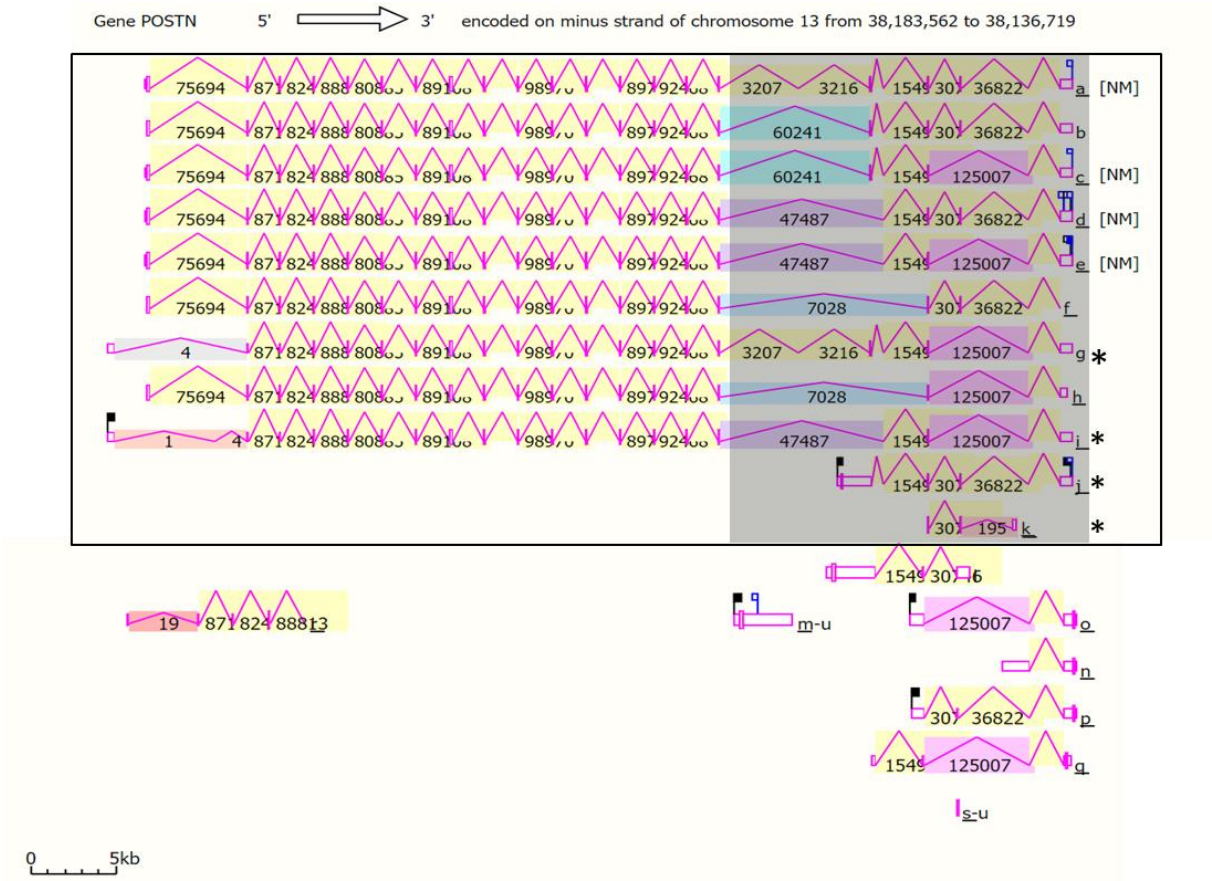


Figure 3. 1 Alignment of mRNAs of POSTN transcripts adapted from AceView (Accessed, May 2019).

The naming of transcripts on the right (a,b,c…) are according to size order. The grey section highlighted is the region affected by alternative splicing. The transcripts boxed are protein-coding. Additional protein-coding transcripts identified only in AceView marked with * (transcripts g, i, j and k). This figure was obtained from the AceView database.

Table 3.1 shows brief details from the Ensembl database the length in base pairs of each transcript variant, as well as the predicted amino acid sequences and molecular weights based on the amino acid sequence. Finally, flags were assigned to each transcript. The transcript support level (TSL) is a method used to highlight transcript models that are well-supported and poorly supported in Ensembl database. The transcript is annotated depends on the type and quality of alignments. The TSL is good for POSTN transcripts 001, 002, 003, 201a and 202a as it was designated 1 which indicates all splice junctions are supported by at least one non-suspect mRNA. “Non-suspect” mRNA comes from a higher quality cDNA library whereas any cDNA from a library that contains artefacts and false positives are labelled “suspect” mRNA. In contrast, POSTN transcript 004 was flagged with

TSL-5 where no single transcript supports the model structure by either an expressed sequence tag (EST) or a suspect mRNA.

Further comparison of transcripts was made using AceView (Figure 3.1). AceView database revealed seventeen POSTN transcript variants and two unspliced forms. However, only eleven are predicted to be protein-coding (transcripts a to k). Of these 11 transcripts, seven transcripts (a,b,c,d,e,f and h) were matched with those found in Genbank and Ensembl by aligning the amino acid sequences using Clustal Omega. Transcripts g, i, j and k are uniquely found in AceView and are not included in the analysis in Table 3.1, since they do not match those transcripts found in Ensembl or Genbank. Both transcripts g and i have missing exon 1 with the start site in the middle of exon 3. Transcripts j and k have truncated 5' coding sequence (CDS) and appears to be within the alternative spliced region.

POSTN gene is expressed in different tissues (Table 3.2). Transcripts from Ensembl came from gene expression studies using RNA sequence data from a variety of tissues. In AceView, cDNA was obtained from tissues not limited to the ones listed in table 3.2. AceView uses experimental cDNA sequences from the species itself. The mRNA models from AceView are from the cDNA sequence data from Genbank (Thierry-Mieg and Thierry-Mieg, 2006).

Table 3. 2 Tissue specific expression patterns of POSTN transcripts from Ensembl and AceView.

Old transcript name Ensembl	Updated transcript name Ensembl	Human tissue from RNA seq data (Ensembl, May 2019)	AceView expression tissue	AceView mRNA reconstructed from cDNA clones
POSTN-001	POSTN-203	Breast, heart, ovary and testes.	Whole embryo (mainly body) and Osteosarcoma.	21
POSTN-201a	POSTN-209	Adipose, colon, kidney, lymph and prostate.	Thyroid gland	33
POSTN-003	POSTN-202	Adrenal	Thyroid gland	7
POSTN-004	POSTN-204	Thyroid	N/A	
POSTN-002	POSTN-201	N/A	Placenta	38
POSTN-201b*	N/A	N/A	Testis, uterus, tongue (tumour tissue), coronary artery, heart, lung, thalamus, trachea, breast, embryonic stem cells (embryoid bodiesderived from H1, H7 & H9 cells, periodontal ligament and hepatoblastoma.	231
POSTN-202a	POSTN-210	N/A	Thyroid gland	2
POSTN-202b*	N/A	N/A	Thyroid gland	2

The cDNA sequences of six individual transcripts (001,002,003,004,201a and 201b) were obtained from Ensembl version 82 (accessed, October 2015). The FASTA sequence of 201b and 202b were obtained from Genbank. These sequences were used to highlight the exon structure of POSTN transcripts to aid with the subsequent primer strategy (Figure 3.2). As shown in Figure 3.2, POSTN 001 is the largest transcript with 23 exons and the shortest transcript is POSTN-202b missing exons 17, 18, 19 and 21. All the other transcripts excluding POSTN-001 are affected by alternative splicing in the 3' end of CDS. Exons affected by splicing mainly include 17, 18, 19 and 21. All POSTN transcript variants are missing exon 17 except POSTN-001 and POSTN-004. POSTN transcripts 002, 201b, 202a and 202b lack exon 18. POSTN transcripts 202a and 202b do not have exon 19. Exon 21 is absent in POSTN transcripts 004, 201a, 201b and 202b.

3.1.2 Prediction of POSTN protein domains and variant protein function

To elucidate the domain structure of POSTN transcripts, domain databases including Expasy Prosite and Uniprot KB were used. Predicted amino acids of individual transcripts were used in the query box which revealed the predicted domain features of POSTN transcript variants. All eight POSTN isoforms (001,002,003,004,201a, 201b, 202a and 202b) contain at the N-terminal the signal peptide (position 1-21 aa) which allows for the secretion of protein in the extracellular matrix (Figure 3.3). Conversely, POSTN protein g and i from AceView had missing signal peptide sequence which indicates these proteins may be intracellular proteins which are interesting since POSTN is known to be a secretory protein.

The EMI domain (position 40-94 aa) which has six conserved cysteine residues and has been predicted to have 3 disulfide linkages between the cysteine residues. All ten POSTN proteins (001,002,003,004,201a, 201b, 202a, 202b, g and i) have identical predicted EMI domain except POSTN protein g and i. All ten POSTN proteins possess four homologous FAS1 domains at positions 97-230 aa, 234-365 aa, 368-492aa and 496-628aa. The position of FAS1 domains in the POSTN proteins g and i are 1-147aa, 151-282aa, 285-409 aa and 413-545aa. There is an amino acid modification within the fourth FAS1 domain, it is predicted to have N-linked glycosylation site at 599 Asn residue.

Lastly, there is a variable length of the C-terminal region of POSTN variants where alternative splicing occurs in the POSTN protein. Hence, there is no functional domain in this region. This region has been proposed to dictate the functional differences between POSTN isoforms.

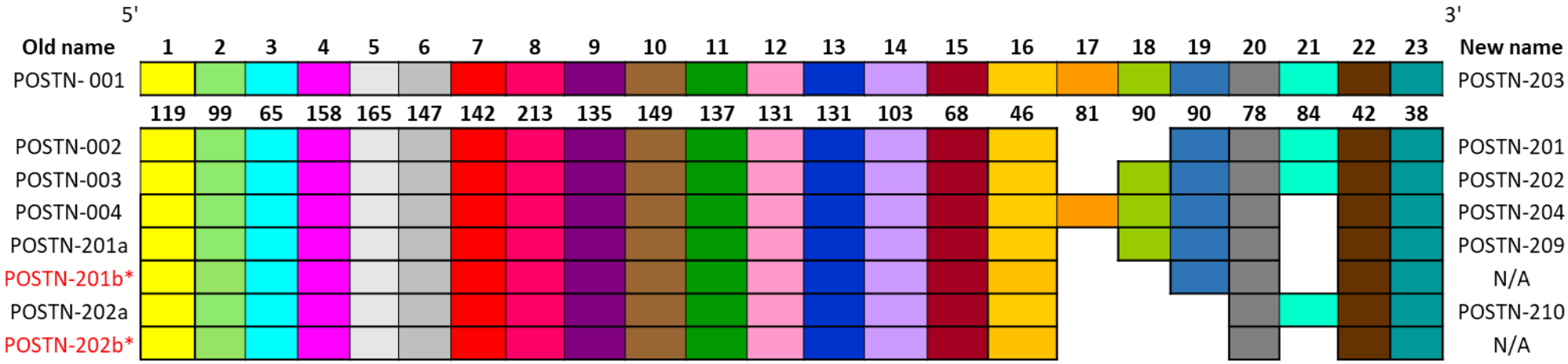


Figure 3. 2 Alignment of cDNA sequences of POSTN alternative transcript variants to create the exon structure.

Individual exon is specified with a distinct colour. The numbers at the top and bottom of POSTN-001 transcript indicates the exon number and length (bp), respectively. These transcripts have a complete coding sequence (CDS) obtained from Ensembl (accessed: October 2015). The transcripts marked with * were obtained from Genbank (POSTN-201b and POSTN 202b). Old naming designations are given on the left while the updated names are on the right.

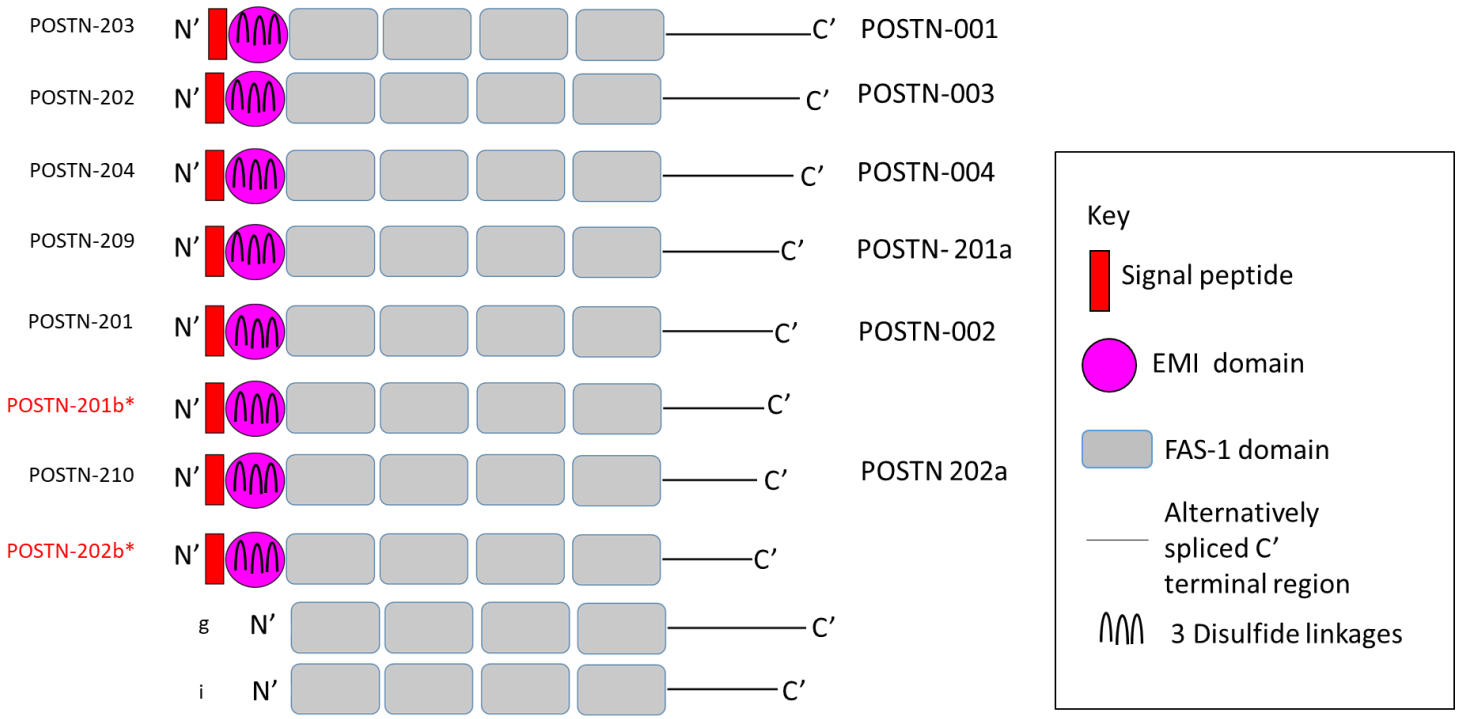


Figure 3. 3 Predicted functional multi-domain structure of POSTN protein variants.

Predicted POSTN protein domains structure were arranged in descending size order. New Ensembl name designation was used for the six Ensembl variants (203, 202, 204, 209, 201 and 210), followed by two new variants from Genbank (201b* and 202b*) and the last two unique variants from AceView that had protein-coding potential. The corresponding old Ensembl name designation given on the right. Predicted amino acid sequences were obtained from the three databases (Ensembl, Genbank and AceView) to determine the relative domains using Expasy prosite. SignalP-5.0 Server was used to predict the signal peptide.

3.2 Analysis of splice variant expression in stromal and cancer cells

3.2.1 Design of RT-PCR primer assay to determine POSTN splice variant expression

The exon structure of POSTN transcript variants depicted in Figure 3.2, helped to design primers to specifically amplify different POSTN transcript variants. The primers designed in the first strategy are listed in table 3.3. Primers were designed using Primer-BLAST programme as described in Chapter 2. Specific parameters were set to ensure specific primer pair was designed to bind only to the desired target of interest while avoiding non-specific amplification by the binding of unintended targets. The template FASTA sequence of POSTN-001 transcript was input in the query box. The position range was set for the forward and reverse primers to ensure the primers are located within specific regions. The amplicon size was chosen to be from 150 bp to 1000 bp. The minimum size was selected to distinguish amplicons from primer dimers which are typically around 100 bp. The melting temperatures (T_m) of both primers were set between 57 °C to 63 °C with the optimum T_m difference of two degrees between two primers. This ensures that primers can anneal to the target template simultaneously to amplify the PCR products.

Most primer pairs were set to span exon/exon junction to amplify the target cDNA sequence and avoid the amplification of genomic DNA. Under “Exon junction match” the number of bases of the primer can be adjusted on each side of the junction. Exon spanning primers were left at the default setting. However, some primers were designed not to be exon spanning. For instance, PN_17_for is a forward primer which binds within exon 17 region is not exon spanning (Table 3.3).

Another feature of Primer-BLAST is that it enables the primer to be specific by adjusting the specificity stringency parameters. The default setting of the stringency was changed so that the 3’ end of the primers have at least six or more mismatches to the unintended products in the last ten base pairs (Ye et al., 2012). Any products with six or more mismatches to at least one primer is ignored (Ye et al., 2012). The primer pair become more specific to its target template if there is a higher mismatch values especially towards the 3’ end between primers and unintended product (Ye et al., 2012). This will make it difficult for primers to anneal to unintended targets. Primers were set to amplify splice

variants from the same gene. "Non-redundant " database was used to check for any non-specific products.

The ideal primer pair was selected if it was met by the following criteria. Similar T_m values with 50-60 % GC content, low self-complementarity and low 3' self-complementarity. The latter two criteria were given to ensure there are no formation of secondary structures such as hairpins and primer dimers.

Primers used in the first strategy are listed in Table 3.3. The primers were named according to which exon/ exon junction it binds to and whether it is a forward or reverse primer. Different primer combinations were used to amplify POSTN transcript variants. PN_14/15_for was paired with two different reverse primers (PN_20/21_rev and PN_20/22_revA) to detect all eight variants. POSTN-001 and POSTN-004 transcripts both have exon 17 but differ by presence/ absence of exon 21. Thus, PN_17_for was paired with the same reverse primers (PN_20/21_rev and PN_20/22_revA) to amplify POSTN-001 and POSTN-004 transcripts selectively.

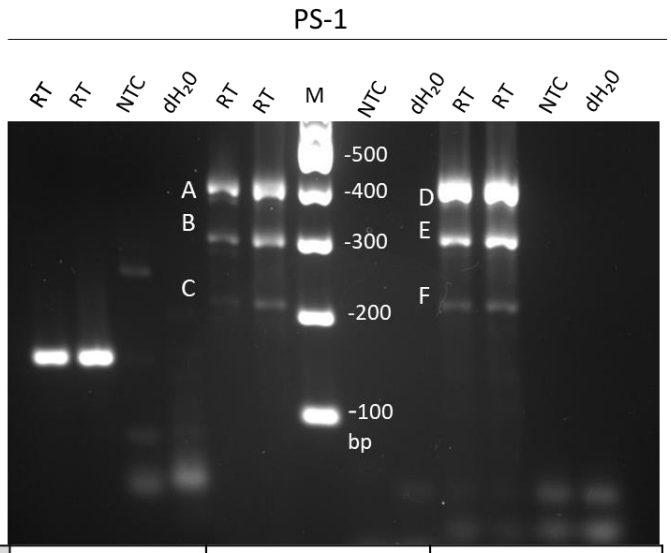
Table 3. 3 List of forward and reverse primers designed in the first strategy.

Primer name	Primer sequence (5'-3')	Tm	POSTN transcripts targeted
PN_14/15_for	TCCAGCAGACACACCTGTTG	60.18	001,002,003,004,201a,201b, 202a & 202b
PN_17_for	GATTGAAGGCAGTCTTCAGCC	59.26	001 &004
PN_21/22_rev	GTGTGTCTCCCTGAAGCAGT	59.61	001,002,003 & 202a
PN_20/22_revA	GGGTGTGTCTTCTTGTAACAATTTC	58.79	004, 201a, 201b & 202b
PN_20/22_revB	CACGGGTGTGTCTTCTTGTAAC	59.46	004,201a,201b & 202b
PN_22_rev	GTTGGCTTGCAACTTCCTCAC	60.27	001,002,003,004,201a,201b, 202a & 202b

3.2.2 Validation of POSTN reverse primers at 20/22 exon junction and its challenges

To validate two new POSTN reverse primers at exon 20/22 junction (PN_20/22_revA and PN_20/22_revB), PS-1 cDNA sample was used as a positive control sample for detection of POSTN expression. QARS primers were used alongside the new primers to validate semi-quantitative RT PCR. As shown in Figure 3.4, amplification with the QARS primers produced a PCR product of the expected size (150 bp) demonstrating successful

amplification. The forward primer for the POSTN gene, PN_14/15_for, was paired with two new reverse primers (PN_20/22_revA and PN_20/22_revB). Three distinct bands were detected in pancreatic stellate cells at varying intensities for both primers. Bands A, B and C were seen using PN_20/22_revA while bands D, E and F were seen using PN_20/22_revB. Note, no-RT control was unintentionally omitted from this experiment. Bands below 100 bp are primer dimers. Figure 3.4 shows the possible POSTN transcripts matching with the expected product size. Due to poor sequencing, the variants corresponding to the bands could not be confirmed.



Primer pair	QARS_for QARS_rev	PN_14/15_for PN_20/22_revA	PN_14/15_for PN_20/22_revB
Transcripts targeted	QARS	**A)** 004 & 001	**D)** 003, 004, 201a
		B) 002 & 201b	**E)** 201b & 202a
		C) 202a & 202b	**F)** 202b
Expected product size (bp)	150	**A)** 469	**D)** 475, 472,391
		B) 298	**E)** 301 & 295
		C) 208	**F)** 211

Figure 3. 4 Validation of two new reverse primers (PN_20/22_revA and PN_20/22_revB).

POSTN transcript variant expression in PS-1 cells on 2.5% w/v agarose gel. Details of the primers used, transcripts targeted, and expected product sizes are included. QARS used as housekeeping gene used as a positive control. Negative controls: no RT- no reverse transcriptase, NTC- no template control and water as reagent control. M is a 100 bp DNA ladder. Bands shown have not been sequenced.

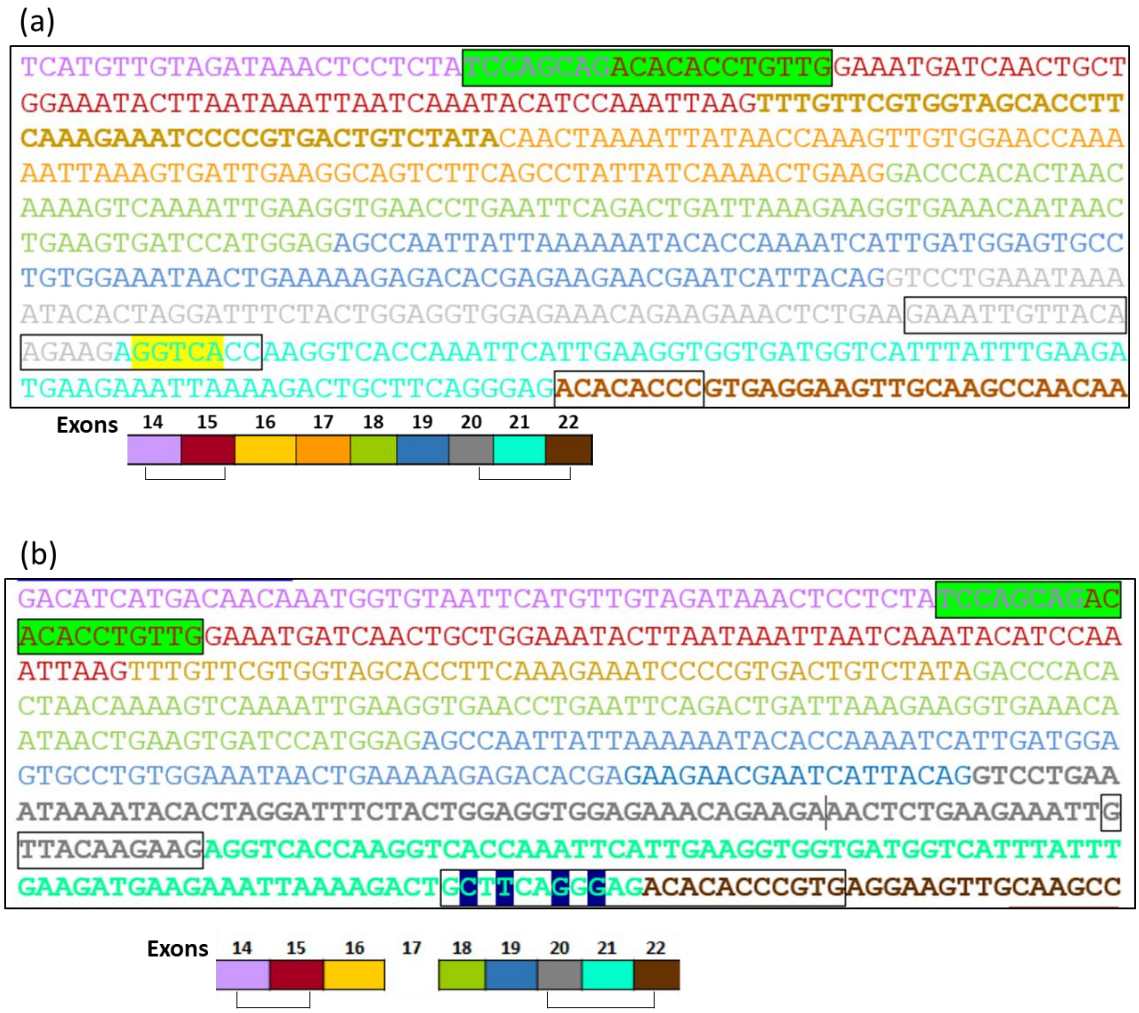


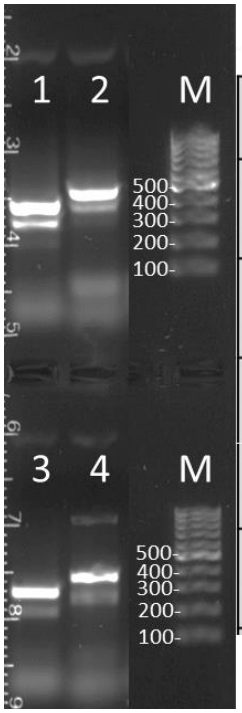
Figure 3. 5 Challenge of exon 20/22 junction of the reverse primers.

(a) PN_14/15 _for and PN_20/22_revA primers mapped on POSTN-001 transcript. Five mismatches at the 5' end of exon 21 of PN_20/22_revA. The mismatches occur in exon 21 as indicated in yellow. Reverse primer binds 20/21 junction instead. **(b)** PN_14/15_for and PN_20/22_revB primers mapped on POSTN-003. Four mismatches indicated in dark blue found in the 3' end of exon 21 of PN_20/22_revB. Reverse primer binds 21/22 junction instead. The sequences shown are from the database used to illustrate the challenge of exon/exon junctions.

Ideally, POSTN reverse primers with exon junction 20/22 should only amplify POSTN transcript variants that have adjacent exons 20 and 22. These include POSTN-004, POSTN-201a, POSTN-201b and POSTN-202b, however, both reverse primers (PN_20/22_revA and PN_20/22_revB) also amplify variants that do not have exon 20/22 junction. As displayed in Figure 3.5 a, analysis of sequences in the database was used to illustrate the exon-exon

junction challenge. POSTN-001 transcript is amplified using PN_20/22_revA when it has no 20/22 junction. This is because the start of exon 21 and the start of exon 22 are similar as it only has five base pair mismatches. Hence, the reverse primer binds to exon 20/21 boundary. Similarly, in Figure 3.5 b, POSTN-003 is amplified using PN_20/22_revB. Again, POSTN-003 does not contain exon 20/22 boundary. The end of exon 20 and the end of exon 21 are similar with only four mismatches in base pairs. Hence, the reverse primer binds to the ex21/22 boundary.

3.2.3 Multiple POSTN transcript variants are detected in stromal cells



Lane no.	Primer pair	Target product from sequencing	Actual band length (bp)
1	PN_14/15_for & PN_20/22_revA	201a	388
		201b	298
2	PN_14/15_for & PN_21/22_rev	003	466
		002	? (380 expected size)
3	PN_17_for & PN_20/22_revA	004	304
		Novel variant (missing 18 and 21)	214
4	PN_17_for & PN_21/22_rev	001	386
		N/A	N/A

Figure 3. 6 Expression of POSTN transcript variant in pancreatic stellate cells on 1% w/v ultrapure gel.

Lanes numbered 1-4 are the RT. M is the DNA ladder. Table adjacent to the gel shows the primers used in each lane with the corresponding PCR products starting at the top with the higher molecular weight bands. RT PCR and sequencing data kindly obtained from Meryem Ozgur. Sequencing data were reanalysed.

After the validation of POSTN reverse primers (PN_20/22_revA and PN_20/22_revB), PN_20/22_revA was selected for the analysis of POSTN in pancreatic stellate, since fewer POSTN transcript variants correspond with a particular band compared with bands from PN_20/22_revB (Figure 3.4).

To find out which POSTN transcripts isoforms are expressed in pancreatic stellate cells (PS-1) total RNA was extracted from PS-1 and were reverse transcribed to make cDNA to be used in the RT PCR analysis. As shown in Figure 3.6, lanes 1-4 show eight bands were seen corresponding to POSTN transcript variants in pancreatic stellate cells. Lanes 1-4 uses four different primer combinations to detect a wide range of POSTN transcript variants in pancreatic stellate cells. Pancreatic stellate cells expressed six POSTN transcript variants 201a, 201b, 003,002, 004 and 001 which were verified by Sanger sequencing (See appendix). Interestingly, a novel variant was expressed in lane 3, containing exon 17, 19, 20 and 22 missing 18 and 21 (Figure 3.7). This POSTN variant was not found in the above three databases Ensembl, Genbank and AceView.

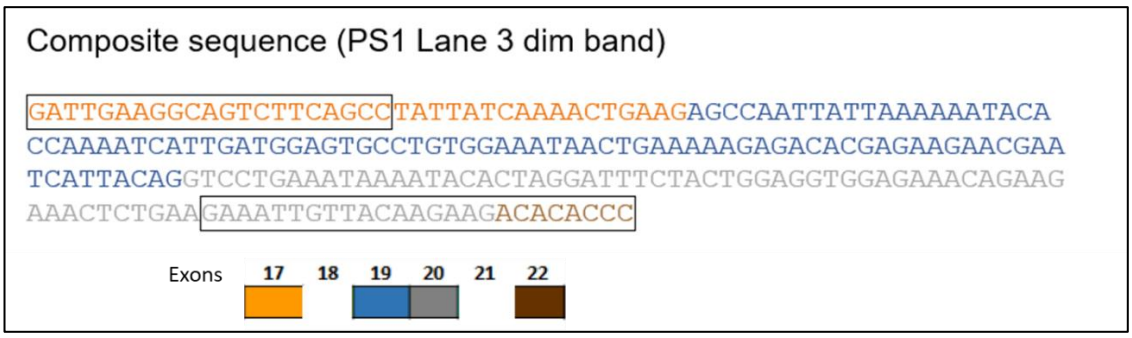
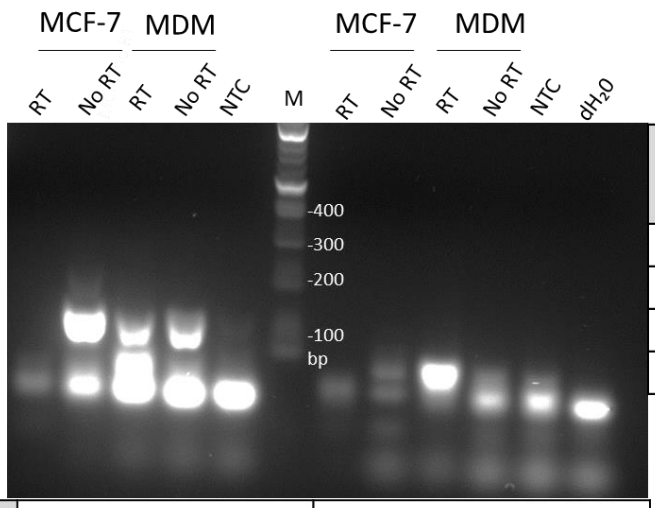


Figure 3. 7 Composite sequence of POSTN novel variant found in Lane 3 of Figure 3.6.

The boxed sequences show the forward and reverse primers.

3.2.4 POSTN expression is not detectable pancreatic or breast cancer cells

The RNA seq data from Ensembl showed that POSTN-001 is expressed in breast tissue (Table 3.2). To confirm if breast cancer cell lines express any POSTN transcript variants, nested RT PCR experiment was conducted to ensure the specific target is amplified. QARS was used as an internal control. Figure 3.8 demonstrates MCF-7 and MDA-MB-468 cell lines did not express any variants of POSTN. Previously, preliminary RT PCR data showed no POSTN expression in pancreatic cell lines: Capan-1, BxPC-3 and PANC-1 (data not shown).



Transcripts targeted	Expected product size (bp)
001	551
002	380
003	470
202	290

Primer pair(s)	QARS_for QARS_rev	PN_14/15_for PN_22_rev
		PN_14/15_for PN_21/22_rev

Figure 3. 8 POSTN transcript variant expression in breast cancer cell lines on 2% w/v agarose gel.

Two breast cancer cell lines MCF-7 and MDA-MB-468 (MDM) were tested for POSTN variant expression using nested RT PCR. M is the DNA ladder.

3.2.5 Design of the second strategy to identify further POSTN transcript variants

The second strategy of primers was designed with the same parameters as described in Section 3.2.1 with the exception that the primers designed were not exon spanning (see Table 3.4). The primers used in the first strategy had its challenges with the exon spanning primers. In particular, the reverse primer with exon 20/22 junction had amplified POSTN transcript variants that did not have this junction (Figure 3.5). For this reason, primers were designed within exon 20 and exon 23 to detect whether POSTN transcripts variants had exon 21 or not. Other reverse primers: PN_20_rev and PN_22_rev were designed to enable the detection of the alternatively-spliced region. The forward primer PN_14/15_for was paired with PN_20_rev to detect the splicing region affecting exons 17, 18 and 19. To detect the amplification of the entire alternatively-spliced region affecting exons 17, 18, 19 and 21, the forward primer PN_14/15_for was paired with PN_22_rev.

Table 3. 4 List of forward and reverse primers designed in the second strategy

Primer name	Primer sequence (5'-3')	Tm	POSTN transcripts targeted
PN_14/15_for	TCCAGCAGACACACCTGTTG	60.18	001,002,003,004,201a,201b, 202a & 202b
PN_20_for	ACTGGAGGTGGAGAAACAGAAG	59.63	001,002,003,004,201a,201b, 202a & 202b
PN_20_rev	GTTTCTTCTGTTTCTCCACCTCC	59.18	001,002,003,004,201a,201b, 202a & 202b
PN_23_rev	CACTGAGAACGACCTTCCCT	59.63	001,002,003,004,201a,201b, 202a & 202b
PN_22_rev	GTTGGCTTGCAACTTCCTCAC	60.27	001,002,003,004,201a,201b, 202a & 202b

Figure 3.9a shows the results of RT PCR which primers were used to amplify the alternatively-spliced region affecting exons 17, 18, 19 and 21. Multiple bands were obtained at various intensities following amplification of cDNA from pancreatic stellate cells. However, only two bands were successfully sequenced (marked in arrows). The top bands are likely to be POSTN-003 at 493 bp and POSTN-201a at 409 bp. It is worth noting, the composite sequence of forward and reverse sequences could not be formed (See appendix). The bottom band had a lower intensity, indicating lower expression levels. It is likely to be either POSTN-002 or POSTN-201b based on the forward sequence trace. The reverse sequence would have revealed if the variant had exon 21 or not, to deduce if it is POSTN-002 or POSTN 201b.

Similarly, in Figure 3.9b primers were used to detect amplification in the alternatively-spliced region affecting exons 17, 18 and 19. Three bands were obtained in pancreatic stellate cells. The first two bands are likely to be POSTN-003 and POSTN-201a both at 357 bp. Although the size differentiation in the agarose gel is different, with the top band at approximately 500 bp and the second band at around 400 bp. The bottom band is likely to be POSTN-002 and POSTN-201b at 267 bp. All the variants identified were confirmed by the forward sequence only as the reverse traces had poor sequencing. Figure 3.9c demonstrates that there are variants with exon 21 (top band) and variants without exon 21 (bottom band).

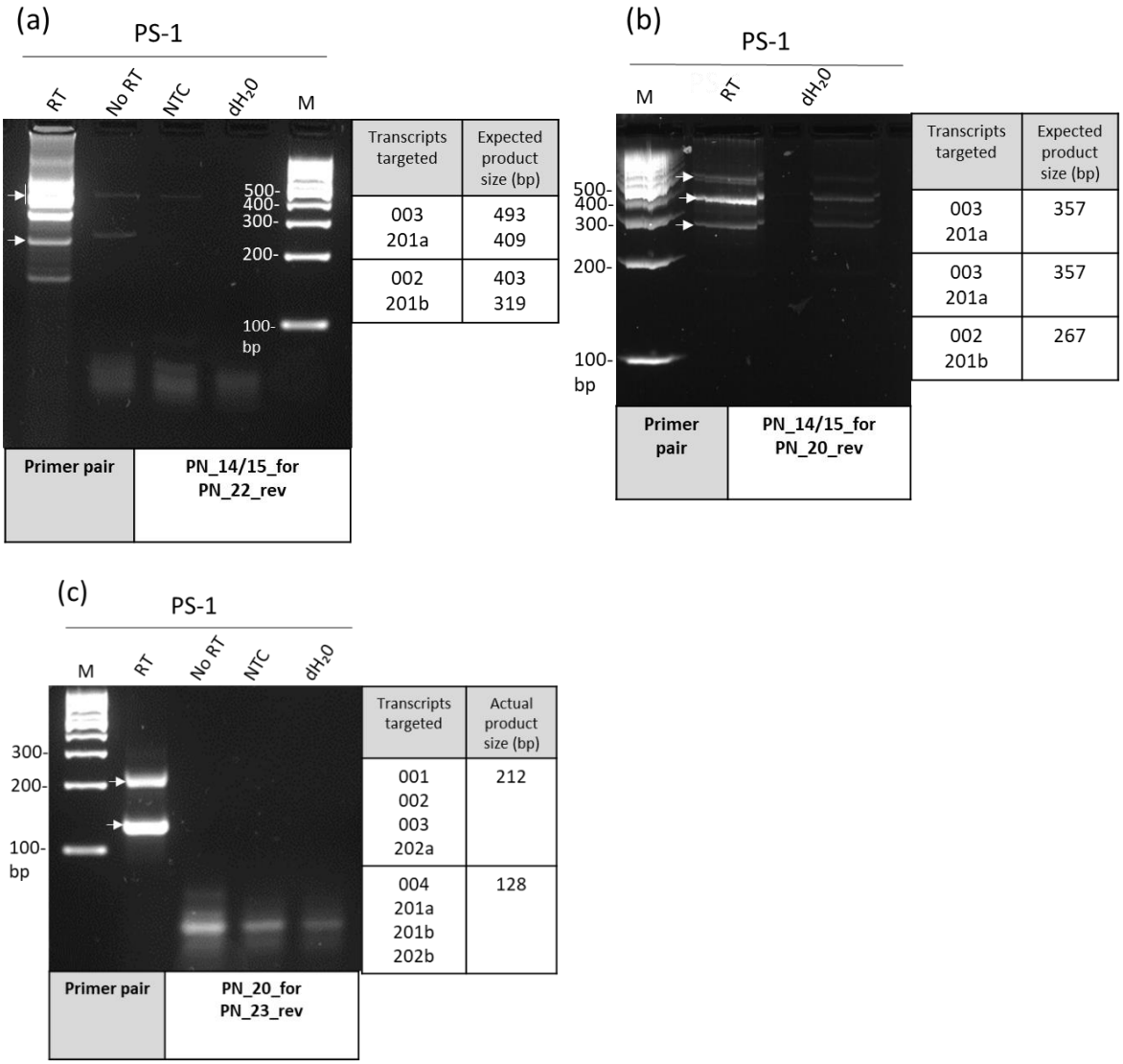


Figure 3. 9 POSTN transcript variant expression in pancreatic stellate cells on 4% w/v agarose gel.

The same forward primer used for (a) and (b) only differing in the reverse primer to amplify the entire alternative spliced region. (c) Primers designed to specifically amplify for the presence or absence of exon 21. Bands marked in arrows were sequenced successfully. M=DNA ladder.

3.2.6 At least seven POSTN splice variants were identified in stromal cells

Table 3.5 shows a summary of seven POSTN transcripts variants expressed in pancreatic stellate cells which were verified by Sanger sequencing. The expression levels could only be judged semi-quantitatively. POSTN transcripts 001,003,004 and 201a had high expression levels in pancreatic stellate cells. Whereas, POSTN-002, POSTN-201b and a novel POSTN variant had lower levels of expression in pancreatic stellate cells. Interestingly, POSTN-202a and POSTN-202b were not detected in pancreatic stellate cells.

Table 3. 5 Splice variant expression of POSTN transcripts in PS-1 cell line.

Figures 3.6 and 3.9 were taken together to get an overview of the range of splice variants expressed in pancreatic stellate cells. Expression level can be judged semi-quantitatively.

Cell line	Splice variants expressed	Exons present in the splicing region	Expression level (semi-quantitative measure)
PS-1	001	17,18,19,20 & 21	High
	002	19,20 & 21	Low
	003	18,19,20 & 21	High
	004	17,18,19 & 20	High
	201a	18,19 &20	High
	201b	19 &20	Low
	Novel variant	17,19 & 20	Low

3.3 Preliminary analysis of the expression of multiple POSTN protein variants in pancreatic stellate cells

Following POSTN transcript variant expression in pancreatic stellate cells (Table 3.5), to investigate whether similar expression of POSTN isoforms is seen at the protein level by western blot. A preliminary experiment was conducted whereby PS-1 cell lysate was prepared and was used to test for POSTN protein expression using anti-POSTN antibody (ab92460, Abcam). The antibody had an epitope at the conserved region in the first Fasciclin (FAS I) domain which enables for the detection of all isoforms (see Figure 3.3). Figure 3.10 shows that at least three bands of varying molecular weight were observed in the 80-93 kDa range expected based on the alternative transcripts observed by RT PCR, at the expected molecular weight sizes as predicted in Table 3.1. Although the bands observed were not sharp, this could be due to multiple POSTN isoforms migrating close together through the gel. The loading control, β-actin was expressed at the expected size at 42 kDa.

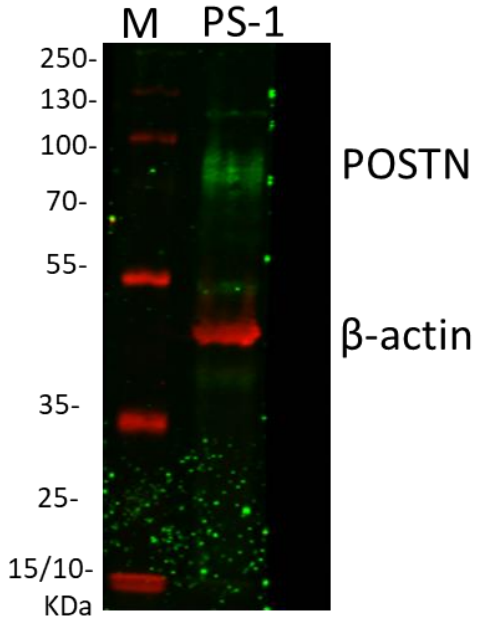


Figure 3. 10 Detection of the expression of POSTN protein variants in pancreatic stellate cells.

30 µg of PS-1 cell lysates were analysed by western blot to detect POSTN using anti-POSTN antibody (ab92460, Abcam). The epitope of POSTN lies in the conserved region of fasciclin domain 1 (FAS-1) enabling the detection of all isoforms. β-actin at 42 kDa used as a loading control. M is the molecular weight marker.

Chapter 4-Results: Cloning of POSTN transcript variants to determine the functional analysis of cellular processes

4.1.1 Initial cloning strategy and design of PCR primers for the preparation of insert

As described in section 3.2.6, multiple POSTN transcript variants expression was observed in pancreatic stellate cells which was confirmed by sequencing. This section describes attempts to clone all these POSTN variants and overexpress them in human stromal cells (PS-1) to investigate the function of POSTN protein isoforms in cellular processes such as pancreatic cancer cell growth, proliferation, adhesion and migration. The first step was to prepare the insert ready to be ligated with the vector pcDNA™4/TO (Figure 4.8).

To amplify the open reading frames (ORF) of all POSTN transcript variants, cloning primers were designed just before the start codon and ending with the stop codon of the longest POSTN transcript, POSTN-203 (POSTN-001), as illustrated in Figure 4.1. The CDS of POSTN-203 was copied and opened in Bioedit software to identify the suitable restriction enzymes that do not cleave within the ORF. A list of restriction enzymes that do not cleave the ORF was obtained. To select the two appropriate restriction sites, the full sequence of pcDNA™4/TO plasmid was obtained from Addgene vector database and imported in Bioedit software to look for restriction enzyme that cut once within the multiple cloning site of the plasmid.

NotI-HF and XhoI were the two chosen restriction enzymes to be incorporated in the cloning primers. As shown in Table 4.1 the forward primer (PN_203cdna_for) consist of a thirteen base pairs upstream and six base pairs downstream of the start codon. The NotI-site was added 5’ end of the forward primer. Similarly, the reverse primer (PN_203cdna_rev) comprises of 24 base pairs of the ORF before the stop codon. Adjacent to the ORF, 6 X histidine tag was added before the stop codon, to allow for the purification of the recombinant protein. The plasmid pcDNA™4/TO does not contain the 6 X histidine tag, therefore, it is important that one of the cloning primers had the 6 X histidine tag. XhoI was added after the stop codon. The reverse complement was generated in the reverse primer to allow for the amplification of the ORF. The cloning

primers were subsequently analysed for secondary structures (hairpin) and primer dimers using the Oligoanalyser tool in Integrated DNA technologies.

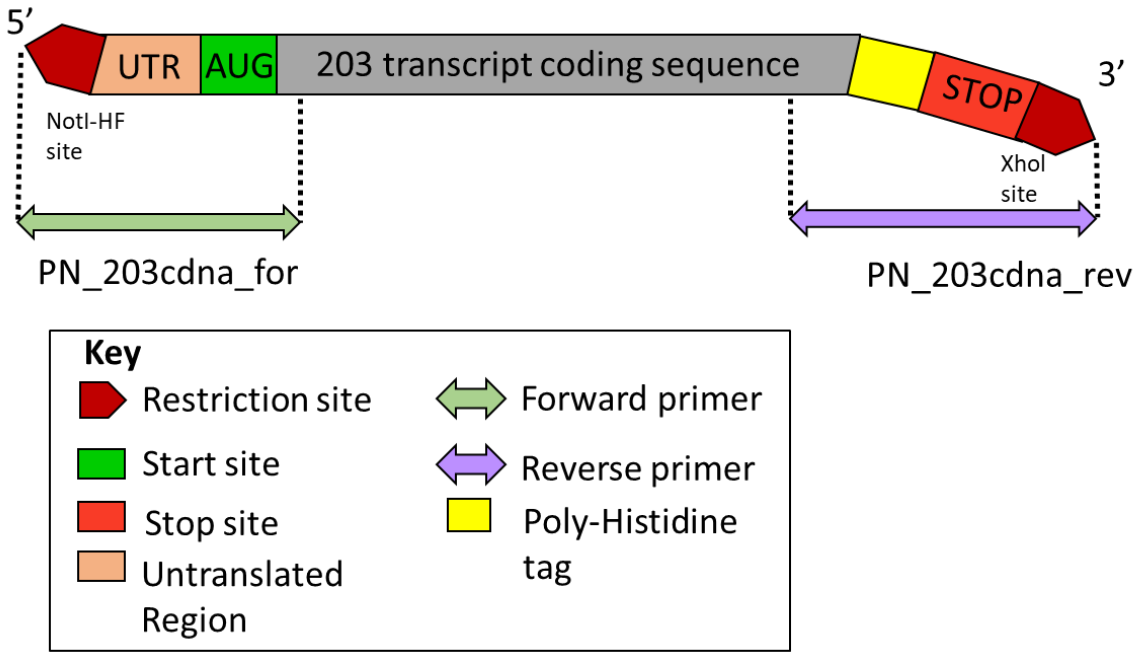


Figure 4. 1 Schematic diagram mapping cloning primers on POSTN-203 transcript coding sequence.

The forward primer (PN_203cdna_for) binds in the 5' of the POSTN-203 sequence while the reverse primer is located at the 3' end of POSTN-203 sequence to amplify the open reading frame (ORF) for cloning. Each cloning primer has different elements added to them. POSTN-203 transcript is the same as POSTN-001 transcript which is the updated nomenclature used throughout the cloning section.

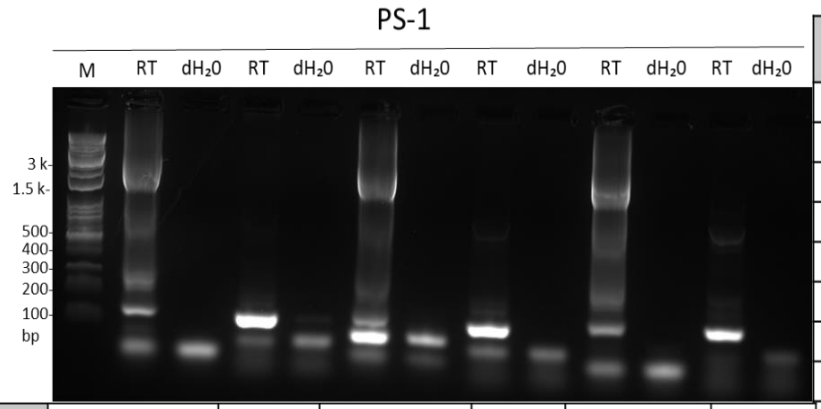
Table 4. 1 Cloning primer sequences used in the first strategy to amplify all POSTN isoforms.

Primer name	Restriction sites	Primer sequence (5'-3')	Tm
PN_203cdna_for	NotI-HF	GCGGCCGCGGAGAGACTCAAGATGATTCCC	72
PN_203cdna_rev	XhoI	CTCGAGTCAATGGTGATGGTGATGATGCTGAGAACGACCTTCCCTTAATCG	71

Different restriction sites are incorporated in the primers are underlined. A green highlight is the start site in the forward primer and the red highlight is the stop site in the reverse primer. The grey highlight is the 5' UTR upstream of the start codon in the forward primer. 6 X Histidine-tags (in yellow) is added in the reverse primer before the stop site. Tm values of primers were calculated using NEB Tm calculator.

4.1.2 Optimising PCR amplification of POSTN variants

The cloning primers designed in section 4.1.1 were used to amplify the insert containing POSTN transcript variants in PS-1 cells. The melting temperatures were optimised (Tm) by gradually increasing the Tm values from 60 °C, 62 °C and 68 °C to judge which Tm values worked the best. Figure 4.2 shows the amplification of a PCR product around 2 kb which indicates POSTN variants. There is a presence of primer dimer at 100 bp and non-specific bands around 300 bp and 500 bp. The Tm value of 68 °C was chosen for subsequent PCR experiments.

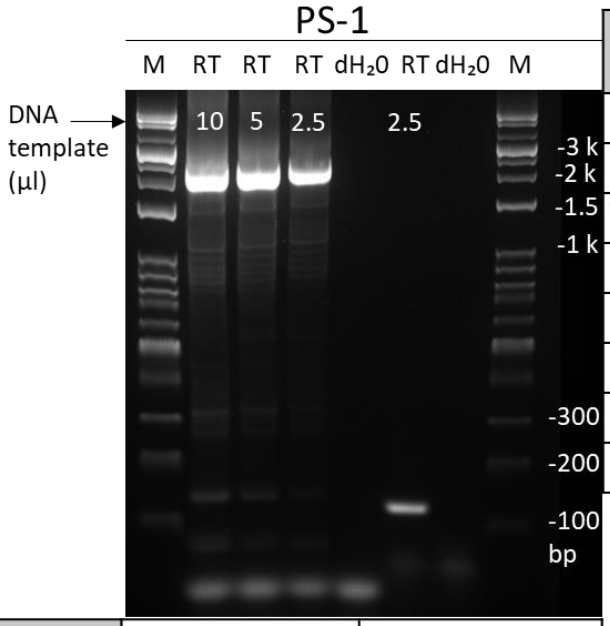


Transcripts targeted	Expected product size (bp)
001	2522
002	2351
003	2441
004	2438
201a	2357
201b	2267
202a	2261
202b	2177

Primer pair	PN_203cdna_for PN_203cdna_rev	QARS_for QARS_rev	PN_203cdna_for PN_203cdna_rev	QARS_for QARS_rev	PN_203cdna_for PN_203cdna_rev	QARS_for QARS_rev
Tm Values (ºC)	60		62		68	

Figure 4. 2 PCR amplification of POSTN transcripts variants in PS-1 cell line on 1.5% w/v agarose gel. Optimisation of RT PCR by increasing the Tm values. QARS used as an internal control. M represents a 1 kb DNA ladder. DreamTaq (2X) polymerase was used to amplify POSTN transcript variants.

Further PCR optimisation was carried out by testing out different starting templates of 10 µl, 5 µl and 2.5 µl (Figure 4.3). The PCR cycle number was reduced to 32. The amplification of insert was around 2.5 kb. The band represents all POSTN transcript variants which have not been resolved well in the agarose gel since the product size differentiation is small. 2.5 µl of cDNA was selected for subsequent PCR experiments for the amplification of POSTN transcript variants.



Transcripts targeted	Expected product size (bp)
001	2522
002	2351
003	2441
004	2438
201a	2357
201b	2267
202a	2261
202b	2177

Primer pair	PN_203cdna_for PN_203cdna_rev	QARS_for QARS_rev

Figure 4. 3 POSTN transcript variants expressed in PS-1 cell line in 1.5% w/v agarose gel.

Further optimisation of PCR conditions by decreasing the starting template (marked in black arrow) from 10 µl, 5µl and 2.5 µl and reducing the cycle number to 32 in the PCR cycling reaction.

4.1.3 Verification of insert by Sanger sequencing suggests the presence of POSTN transcript isoforms

To confirm whether the band expressed at 2.5 kb contain all POSTN transcript variants, sequencing was conducted. Since the band around 2.5 kb did not resolve well, the band was dissected from the top (labelled P2) and bottom (labelled P3) and was sequenced using the reverse primer (PN_203cdna_rev) to deduce the sequencing results. Additionally, the whole 2.5 kb band was excised separately and was sequenced with the same reverse primer (PN_203cdna_rev). As shown in Figure 4.4, all three reverse sequence traces were identical missing exon 21. Although we cannot formally rule out

that exon 21 does not exist since there were mixed traces present in the sequencing data. However, it does suggest that the majority of POSTN isoforms do not contain exon 21.

```
203                 AAAAGTCAAAATTGAAGGTGAACCTGAATTCAGACTGATTAAAGAAGGTGAAACAATAAC
PN2                 ------------------------------------TGATTAAAGAAGGGGAAACAATAAC
POSTN_203_rev_edit  ------------------------------------cTGATTAAAGAAGGTGAAACAATAAC
PN3                 ------------------------------------TGATTAAAGAAGGTGAAACANTAAC
                                                        ************* ****** ****

203                 TGAAGTGATCCATGGAGAGCCAATTATTAAAAAATACACCAAAATCATTGATGGAGTGCC
PN2                 NNAAGTGATCCATGGAGAGCCAATTATTAAAAAATACACCAAAATCATTGATGGAGTNCC
POSTN_203_rev_edit  tGAAGTGATCCATGGAGAGCCAATTATTAAAAAATACACCAAAATCATTGATGGAGTGCC
PN3                 NGAAGTGATCCATGGAGAGCCAATTATTAAAAAATACACCAAAATCATTGATGGAGTGCC
                      ******************************************************** **

203                 TGTGGAAATAACTGAAAAAGAGACACGAGAAGAACGAATCATTACAGGTCCTGAAATAAA
PN2                 TGTGGAAATANNTGAAAAAGAGACACGAGAAGAACGAATCATTACAGGTCNNGAAATAAA
POSTN_203_rev_edit  TGTGGAAATAAcTGAAAAAGAGACACGAGAAGAACGAATCATTACAGGTCcTGAAATAAA
PN3                 TGTGGAAATAACTGAAAAAGAGACACGAGAAGAACGAATCATTACAGGTCCTGAAATAAA
                    **********  ************************************  ********

203                 ATACACTAGGATTTCTACTGGAGGTGGAGAAACAGAAGAAACTCTGAAGAAATTGTTACA
PN2                 ATACACTAGGATTTCTACTGGAGGTGGAGAAACAGAAGAAACTCTGAAGAAATTGTTACA
POSTN_203_rev_edit  ATACACTAGGATTTCTACTGGAGGTGGAGAAACAGAAGAAACTcTGAAGAAATTGTTACA
PN3                 ATACACTAGGATTTCTACTGGAGGTGGAGAAACAGAAGAAACTCTGAAGAAATTGTTACA
                    ************************************************************

203                 AGAAGAGGTCACCAAGGTCACCAAATTCATTGAAGGTGGTGATGGTCATTTATTTGAAGA
PN2                 AGAAG-------------------------------------------------------
POSTN_203_rev_edit  AGAAG-------------------------------------------------------
PN3                 AGAAG-------------------------------------------------------
                    *****

203                 TGAAGAAATTAAAAGACTGCTTCAGGGAGACACACCCGTGAGGAAGTTGCAAGCCAACAA
PN2                 ------------------------------ACACACCCGTGAGGAAGTTGCAAGCCAACAA
POSTN_203_rev_edit  ------------------------------ACACACCCGTGAGGAAGTTGcAAGCCAacAA
PN3                 ------------------------------ACACACCCGTGAGGAAGTTGCAAGCCAACAA
                                                   ******************************


203                 AAAAGTTCAAGGATCTAGAAGACGATTAAGGGAAGGTCGTTCTCAGTGAAAATCCAAAAA
PN2                 AAAAGTtCAAGGATCTAGAAGA ---------------------------------------
POSTN_203_rev_edit  AAAAGTtCAAGgatCtagaag ---------------------------------------
PN3                 AAAAGTtCAAGGATCTAGAAGA ---------------------------------------
                    **********************
```

Exons present: 18 19 20 21 22 23

Figure 4. 4 Multiple alignments of reverse nucleotide sequence traces of 2.5 kb band against POSTN 203 isoform (top row highlighted with relative exons). 203 is the sequence of POSTN 203 transcript obtained from Ensembl used to compare sequence traces. The 2.5 kb band was dissected so that the top and bottom slices are labelled PN2 and PN3 respectively. A whole 2.5 kb band was also separately excised and sent for sequencing is labelled POSTN_203_rev_edit. PN_203_rev reverse primer was used to obtain the reverse traces above.

```
NM_006475.2          CTCTCCAAAGCCCACTGCCAGTTCTCTTCGGGGACTAACTGCAACGGAGAGACTCAAGAT
postn_203_for_edited ------------------------------------------------------------

NM_006475.2          GATTCCCTTTTTACCCATGTTTTCTCTACTATTGCTGCTTATTGTTAACCCTATAAACGC
postn_203_for_edited ---------------------------------GNTGNTTATTGTTAACCCTATAAACGC
                                                     * ** **********************

NM_006475.2          CAACAATCATTATGACAAGATCTTGGCTCATAGTCGTATCAGGGGTCGGGACCAAGGCCC
postn_203_for_edited CAACAATCATTATGANAANATCTTGGCTCATAGTCGTATCAGGGGTCGGGACCAAGGCCC
                     *************** ** ****************************************

NM_006475.2          AAATGTCTGTGCCCTTCAACAGATTTTGGGCACCAAAAAGAAATACTTCAGCACTTGTAA
postn_203_for_edited AAATGTCTGTGCCCTTCAACAGATTTTGGGCACCAAAAAGAAATACTTCAGCACTTGTAA
                     ************************************************************

NM_006475.2          GAACTGGTATAAAAAGTCCATCTGTGGACAGAAAACGACTGTGTTATATGAATGTTGCCC
postn_203_for_edited GAACTGGTATAAAAAGTCCATCTGTGGACAGAAAACGACTGTGTTATATGAATGTTGCCC
                     ************************************************************

NM_006475.2          TGGTTATATGAGAATGGAAGGAATGAAAGGCTGCCCAGCAGTTTTGCCCATTGACCATGT
postn_203_for_edited TGGTTATATGAGAATGGAAGGAATGAAAGGCTGCCCAGCAGTTTTGCCCATTGACCATGT
                     ************************************************************

NM_006475.2          TTATGGCACTCTGGGCATCGTGGGAGCCACCACAACGCAGCGCTATTCTGACGCCTCAAA
postn_203_for_edited TTATGGCACTCTGGGCATCGTGGGAGCCACCACAACGCAGCGCTATTCTGACGCCTCAAA
                     ************************************************************

NM_006475.2          ACTGAGGGAGGAGATCGAGGGAAAGGGATCCTTCACTTACTTTGCACCGAGTAATGAGGC
postn_203_for_edited ACTGAGGGAGGAGATCGAGGGAAAGGGATCCTTCACTTACTTTGCACCGAGTAATGAGGC
                     ************************************************************

NM_006475.2          TTGGGACAACTTGGATTCTGATATCCGTAGAGGTTTGGAGAGCAACGTGAATGTTGAATT
postn_203_for_edited TTGGGACAACTTGGATTCTGATATCCGTAGAGGTTTGGAGAGCAACGTGAATGTTGAATT
                     ************************************************************

NM_006475.2          ACTGAATGCTTTACATAGTCACATGATTAATAAGAGAATGTTGACCAAGGACTTAAAAAA
postn_203_for_edited ACTGAATGCTTTACATAGTCACATGATTAATAAGAGAATGTTGACCAAGGACTTAAAAAA
                     ************************************************************

NM_006475.2          TGGCATGATTATTCCTTCAATGTATAACAATTTGGGGCTTTTCATTAACCATTATCCTAA
postn_203_for_edited TGGCATGATTATTCCTTCAATGTATAACAATTTGGGGCTTTTCATTAACCATTATCCTAA
                     ************************************************************

NM_006475.2          TGGGGTTGTCACTGTTAATTGTGCTCGAATCATCCATGGGAACCAGATTGCAACAAATGG
postn_203_for_edited TGGGGTTGTCACTGTTAATTGTGCTCGAATCATCCATGGGAACCAGATTGCAACAAATGG
                     ************************************************************

NM_006475.2          TGTTGTCCATGTCATTGACCGTGTGCTTACACAAATTGGTACCTCAATTCAAGACTTCAT
postn_203_for_edited TGTTGTCCATGTCATTGACCGTGTGCTTACACAAATTGGTACCTCAATTCAAGACTTCAT
                     ************************************************************

NM_006475.2          TGAAGCAGAAGATGACCTTTCATCTTTTAGAGCAGCTGCCATCACATCGGACATATTGGA
postn_203_for_edited TGAAGCAGAAGATGACCTTTCATCTTTTAGAGCAGCTGCCATCACATCGGACATATTGGA
                     ************************************************************

NM_006475.2          GGCCCTTGGAAGAGACGGTCACTTCACACTCTTTGCTCCCACCAATGAGGCTTTTGAGAA
postn_203_for_edited GGCCCTTGGAAGAGACGGTCACTTCACACTCTTTGCTCCCACCAATGAGGCTTTTGAGAA
                     ************************************************************

NM_006475.2          ACTTCCACGAGGTGTCCTAGAAAGGATCATGGGAGACAAAGTGGCTTCCGAAGCTCTTAT
postn_203_for_edited ACTTCCACGAGGTGTCCTAGAAAGGATCATGGGAGACAAAGTGGCTTCCGAAGCTCTTAT
                     ************************************************************

NM_006475.2          GAAGTACCACATCTTAAATACTCTCCAGTGTTCTGAGTCTATTATGGGAGGAGCAGTCTT
postn_203_for_edited GAAGTACCACATCTTAAATACTCTCCAGTGTTCTGAGTCTATTATGGGAGGAGCAGTCTT
                     ************************************************************

NM_006475.2          TGAGACGCTGGAAGGAAATACAATTGAGATAGGATGTGACGGTGACAGTATAACAGTAAA
postn_203_for_edited TGAGACGCTGGAAGGAAATACAATTGAGATAGGATGTGACGGTGACAGTATAACAGTAAA
                     ************************************************************

NM_006475.2          TGGAATCAAAATGGTGAACAAAAAGGATATTGTGACAAATAATGGTGTGATCCATTTGAT
postn_203_for_edited TGGAATCAAAATGTTGAACA-AAAGGATATTGTGACAAATAATGGTGTGATCCATTTGAT
                     ************* ****** ****************************************

NM_006475.2          TGATCAGGTCCTAATTCCTGATTCTGCCAAACAAGTTATTGAGCTGGCTGGAAAACAGCA
```

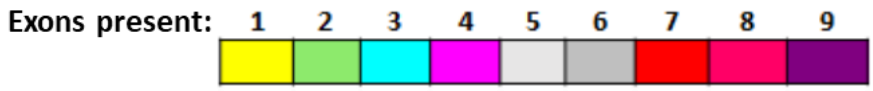


Figure 4. 5 Multiple alignments of forward nucleotide sequence trace of 2.5 kb band aligned against POSTN 203 isoform obtained from Ensembl (top row highlighted with exons). PN_203_for forward primer was used to detect forward trace from exon 1 onwards.

```
NM_006475.2          ACTTCCACGAGGTGTCCTAGAAAGGATCATGGGAGACAAAGTGGCTTCCGAAGCTCTTAT
postn_ex_7_for_edited ------------------------ATcATGGGAGACAAAGTGGCTTCCGAAGCTCTTAT
                                               **********************************

NM_006475.2          GAAGTACCACATCTTAAATACTCTCCAGTGTTCTGAGTCTATTATGGGAGGAGCAGTCTT
postn_ex_7_for_edited GAAGTACCACATCTTAAATACTCTCCAGTGTTCTGAGTCTATTATGGGAGGAGCAGTCTT
                      ************************************************************

NM_006475.2          TGAGACGCTGGAAGGAAATACAATTGAGATAGGATGTGACGGTGACAGTATAACAGTAAA
postn_ex_7_for_edited TGAGACGCTGGAAGGAAATACAATTGAGATAGGATGTGACGGTGACAGTATAACAGTAAA
                      ************************************************************

NM_006475.2          TGGAATCAAAATGGTGAACAAAAAGGATATTGTGACAAATAATGGTGTGATCCATTTGAT
postn_ex_7_for_edited TGGAATCAAAATGGTGAACAAAAAGGATATTGTGACAAATAATGGTGTGATCCATTTGAT
                      ************************************************************

NM_006475.2          TGATCAGGTCCTAATTCCTGATTCTGCCAAACAAGTTATTGAGCTGGCTGGAAAACAGCA
postn_ex_7_for_edited TGATCAGGTCCTAATTCCTGATTCTGCCAAACAAGTTATTGAGCTGGCTGGAAAACAGCA
                      ************************************************************

NM_006475.2          AACCACCTTCACGGATCTTGTGGCCCAATTAGGCTTGGCATCTGCTCTGAGGCCAGATGG
postn_ex_7_for_edited AACCACCTTCACGGATCTTGTGGCCCAATTAGGCTTGGCATCTGCTCTGAGGCCAGATGG
                      ************************************************************

NM_006475.2          AGAATACACTTTGCTGGCACCTGTGAATAATGCATTTTCTGATGATACTCTCAGCATGGA
postn_ex_7_for_edited AGAATACACTTTGCTGGCACCTGTGAATAATGCATTTTCTGATGATACTCTCAGCATGGA
                      ************************************************************

NM_006475.2          TCAGCGCCTCCTTAAATTAATTCTGCAGAATCACATATTGAAAGTAAAAGTTGGCCTTAA
postn_ex_7_for_edited TCAGCGCCTCCTTAAATTAATTCTGCAGAATCACATATTGAAAGTAAAAGTTGGCCTTAA
                      ************************************************************

NM_006475.2          TGAGCTTTACAACGGGCAAATACTGGAAACCATCGGAGGCAAACAGCTCAGAGTCTTCGT
postn_ex_7_for_edited TGAGCTTTACAACGGGCAAATACTGGAAACCATCGGAGGCAAACAGCTCAGAGTCTTCGT
                      ************************************************************

NM_006475.2          ATATCGTACAGCTGTCTGCATTGAAAATTCATGCATGGAGAAAGGGAGTAAGCAAGGGAG
postn_ex_7_for_edited ATATCGTACAGCTGTCTGCATTGAAAATTCATGCATGGAGAAAGGGAGTAAGCAAGGGAG
                      ************************************************************

NM_006475.2          AAACGGTGCGATTCACATATTCCGCGAGATCATCAAGCCAGCAGAGAAATCCCTCCATGA
postn_ex_7_for_edited AAACGGTGCGATTCACATATTCCGCGAGATCATCAAGCCAGCAGAGAAATCCCTCCATGA
                      ************************************************************

NM_006475.2          AAAGTTAAAACAAGATAAGCGCTTTAGCACCTTCCTCAGCCTACTTGAAGCTGCAGACTT
postn_ex_7_for_edited AAAGTTAAAACAAGATAAGCGCTTTAGCACCTTCCTCAGCCTACTTGAAGCTGCAGACTT
                      ************************************************************

NM_006475.2          GAAAGAGCTCCTGACACAACCTGGAGACTGGACATTATTTGTGCCAACCAATGATGCTTT
postn_ex_7_for_edited GAAAGAGCTCCTGACACAACCTGGAGACTGGACATTATTTGTGCCAACCAATGATGCTTT
                      ************************************************************

NM_006475.2          TAAGGGAATGACTAGTGAAGAAAAAGAAATTCTGATACGGGACAAAAATGCTCTTCAAAA
postn_ex_7_for_edited TAAGGGAATGACTAGTGAAGAAAAAGAAATTCTGATACGGGACAAAAATGCTCTTCAAAA
                      ************************************************************

NM_006475.2          CATCATTCTTTATCACCTGACACCAGGAGTTTTCATTGGAAAAGGATTTGAACCTGGTGT
postn_ex_7_for_edited CATCATTCTTTATCACCTGACACCAGGAGTTTTCATTGGAAAAGGATTTGAACCTGGTGT
                      ************************************************************

NM_006475.2          TACTAACATTTTAAAGACCACACAAGGAAGCAAAATCTTTCTGAAAGAAGTAAATGATAC
postn_ex_7_for_edited TACTAACATTTTAAAGACCACACAAGGAAGCAAAATCTTTCTGAAAGAAGTAAATGATAC
                      ************************************************************

NM_006475.2          ACTTCTGGTGAATGAATTGAAATCAAAAGAATCTGACATCATGACAACAAATGGTGTAAT
postn_ex_7_for_edited ACTTCTGGTGAATGAATTGAAATCAAAAGAATCTGACATCATGACAACAAATGGTGT---
                      *********************************************************

NM_006475.2          TCATGTTGTAGATAAACTCCTCTATCCAGCAGACACACCTGTTGGAAATGATCAACTGCT
postn_ex_7_for_edited ------------------------------------------------------------
```

Exons present: 7 8 9 10 11 12 13 14

Figure 4. 6 Multiple alignments of forward nucleotide sequence trace of 2.5 kb band aligned against POSTN 203 isoform obtained from Ensembl (top row highlighted with exons). PN_ex7_for forward primer used to detect sequence from exon 7 onwards.

```
NM_006475.2          GAAAGAGCTCCTGACACAACCTGGAGACTGGACATTATTTGTGCCAACCAATGATGCTTT
postn_ex12_for_edited ------------------------------------------------------------GCT
                                                                                    *

NM_006475.2          TAAGGGAATGACTAGTGAAGAAAAAGAAATTCTGATACGGGACAAAAATGCTCTTCAAAA
postn_ex12_for_edited TTTAGGGATGACTAGTGAAGAAAAAGAAATTCTGATACGGGACAAAAATGCTCTTCAAAA
                     *   ** ****************************************************

NM_006475.2          CATCATTCTTTATCACCTGACACCAGGAGTTTTCATTGGAAAAGGATTTGAACCTGGTGT
postn_ex12_for_edited CATCATTCTTTATCACCTGACACCAGGAGTTTTCATTGGAAAAGGATTTGAACCTGGTGT
                     ************************************************************

NM_006475.2          TACTAACATTTTAAAGACCACACAAGGAAGCAAAATCTTTCTGAAAGAAGTAAATGATAC
postn_ex12_for_edited TACTAACATTTTAAAGACCACACAAGGAAGCAAAATCTTTCTGAAAGAAGTAAATGATAC
                     ************************************************************

NM_006475.2          ACTTCTGGTGAATGAATTGAAATCAAAAGAATCTGACATCATGACAACAAATGGTGTAAT
postn_ex12_for_edited ACTTCTGGTGAATGAATTGAAATCAAAAGAATCTGACATCATGACAACAAATGGTGTAAT
                     ************************************************************

NM_006475.2          TCATGTTGTAGATAAACTCCTCTATCCAGCAGACACACCTGTTGGAAATGATCAACTGCT
postn_ex12_for_edited TCATGTTGTAGATAAACTCCTCTATCCAGCAGACACACCTGTTGGAAATGATCAACTGCT
                     ************************************************************

NM_006475.2          GGAAATACTTAATAAATTAATCAAATACATCCAAATTAAGTTTGTTCGTGGTAGCACCTT
postn_ex12_for_edited GGAAATACTTAATAAATTAATCAAATACATCCAAATTAAGTTTGTTCGTGGTAGCACCTT
                     ************************************************************

NM_006475.2          CAAAGAAATCCCCGTGACTGTCTATACAACTAAAATTATAACCAAAGTTGTGGAACCAAA
postn_ex12_for_edited CAAAGAAATCCCCGTGACTGTCTATA---------------------------------
                     ***************************

NM_006475.2          AATTAAAGTGATTGAAGGCAGTCTTCAGCCTATTATCAAAACTGAAGGACCCACACTAAC
postn_ex12_for_edited ----------------------------------------------NACCCACNCTAAC
                                                                    ****** *****

NM_006475.2          AAAAGTCAAAATTGAAGGTGAACCTGAATTCAGACTGATTAAAGAAGGTGAAACAATAAC
postn_ex12_for_edited AAAAGTCAAAATTGAANGTGAACCTGANTTCGNACTGATTAAAGAAGGTGAAACAATAAC
                     **************** ********** ***  ****************************

NM_006475.2          TGAAGTGATCCATGGAGAGCCAATTATTAAAAAATACACCAAAATCATTGATGGAGTGCC
postn_ex12_for_edited TGAANTGATCCATGGNNANNCNNNTATTAAAAAATACACCANAATCATTGATGGAGTGCC
                     **** **********  *  *   ***************** ******************

NM_006475.2          TGTGGAAATAACTGAAAAAGAGACACGAGAAGAACGAATCATTACAGGTCCTGAAATAAA
postn_ex12_for_edited TGTGGAA-----------------------------------------------------
                     *******
```

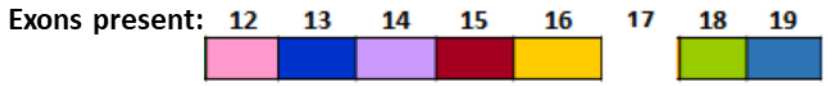


Figure 4. 7 Multiple alignments of forward nucleotide sequence trace of 2.5 kb band aligned against POSTN 203 isoform obtained from Ensembl (top row highlighted with exons). PN_ex12_for forward primer used to detect sequence from exon 12 onwards.

To sequence through the entire POSTN transcript variants, three primers were used including POSTN_203cdna_for, POSTN_7_for and POSTN_12_for. Figures 4.5 and 4.6 show sequencing results revealed that exons 1 to 14 were present. The corresponding traces were clean with no background trace. Figure 4.7 show exons 12 to 19 were present with the absence of exon 17. However, the sequencing trace reveals that there were

overlapping peaks after 400 bp which suggest mixed traces. Hence, we cannot rule out the missing exon 17. It suggests that the majority of POSTN transcript variants do not contain exon 17 which correlates well with the bioinformatics data (Figure 3.2). Merging the forward and reverse sequences together suggest that POSTN 201a, which have both missing exons 17 and 21, seems to be the most common POSTN transcript variant in PS1 cells.

4.1.4 Choice and preparation of the cloning vector for transformation

Plasmid pcDNA™4/TO is a mammalian expression vector used in the cloning experiment. Figure 4.8 shows all the features of the plasmid. The plasmid size is 5078 bp long with a resistance gene against ampicillin. This allows transformants to be grown in LB agar plate supplemented with ampicillin. Any competent cells that did not take up plasmid pcDNA™4/TO fail to grow in ampicillin treated LB plates. pcDNA™4/TO has the multiple cloning site (MCS) which contains all the restriction sites. The two restriction enzymes: NotI-HF and XhoI, were selected and incorporated in the cloning primers to amplify insert in section 4.1.1, are located only in the MCS and do not cut anywhere else in the plasmid. A detailed map of the MCS is shown in figure 4.9.

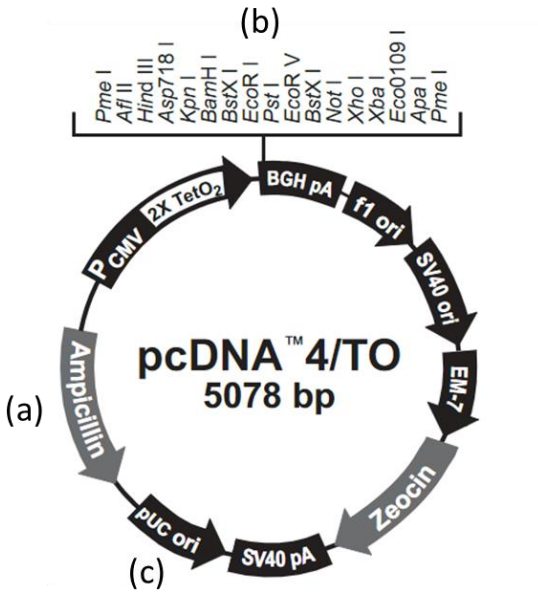


Figure 4. 8 Plasmid vector pcDNA™4/TO mammalian expression vector map used in the cloning experiment. (a) Ampicillin resistance gene. (b) Multiple cloning site with all the restriction sites. (c) Origin of replication. This figure was acquired from the pcDNA™4/TO Invitrogen manual.

```
                                                                                CMV Forward priming site
 721  AAAATCAACG GGACTTTCCA AAATGTCGTA ACAACTCCGC CCCATTGACG CAAATGGGCG

                                   TATA box                 Tetracycline operator (TetO2)
 781  GTAGGCGTGT ACGGTGGGAG GTCTATATAA GCAGAGCTCT CCCTATCAGT GATAGAGATC

      Tetracycline operator (TetO2)
 841  TCCCTATCAG TGATAGAGAT CGTCGACGAG CTCGTTTAGT GAACCGTCAG ATCGCCTGGA

 901  GACGCCATCC ACGCTGTTTT GACCTCCATA GAAGACACCG GGACCGATCC AGCCTCCGGA

                Pme I*  Afl II Hind III  Asp718 I  Kpn I      BamH I                BstX I* EcoR I
 961  CTCTAGCGTT TAAACTTAAG CTTGGTACCG AGCTCGGATC CACTAGTCCA GTGTGGTGGA

            Pst I  EcoR V        BstX I*   Not I     Xho I    Xba I   Eco0109 I  Apa I   Pme I*
1021  ATTCTGCAGA TATCCAGCAC AGTGGCGGCC GCTCGAGTCT AGAGGGCCCG TTTAAACCCG

                   BGH Reverse priming site
1081  CTGATCAGCC TCGACTGTGC CTTCTAGTTG CCAGCCATCT
```

*Please note that there are two *Pme* I sites and two *Bst*X I sites in the polylinker.

Figure 4. 9 Multiple cloning site (MCS) of the pcDNA™4/TO mammalian expression vector. Restriction sites are mapped in the MCS to show the cleavage sites. Potential stop codons are underlined.

This figure was acquired from the pcDNA™4/TO Invitrogen manual. NotI-HF and XhoI restriction enzymes chosen in the cloning experiment.

Preparation of plasmid pcDNA™4/TO was described previously in Chapter 2. The initial plasmid was extracted from competent *E.coli* XL1 strain and quantified using a Biodrop. The extracted plasmid was double digested using the restriction enzymes NotI-HF and XhoI. This creates incompatible cohesive ends which prevent self-ligation and enables the correct orientation of insert into the MCS of the plasmid. After the purification and quantification, the digested plasmid was ligated using T4 DNA ligase with the digested insert in the following molar ratios of the insert to vector: 1:1, 1:3 and 3:1. Negative control was included with digested plasmid alone which was treated with T4 DNA ligase. The ends of the digested plasmid are incompatible due to double digestion with two different restriction enzymes, the plasmid should remain linearized. Hence, no colonies should be seen in the negative plate. The ligated products were transformed using heat shock method into competent *E.coli* DH5α cells. The bacterial cells were grown on ampicillin treated LB agar plates to select cells that contain circular pcDNA™4/TO which has the ampicillin resistance gene. Colonies were seen on all plates the following day

(image not shown). The negative plate with digested plasmid and no insert had colonies present. This may suggest there was incomplete digestion of the plasmid.

As the negative control plate was positive, to test whether two restriction enzymes NotI-HF and XhoI were functioning efficiently, a single digest of each enzyme was set up. As shown in Figure 4.10 colonies were observed on all plates. Figures 4.10a and b show single digests of the plasmid with XhoI and NotI-HF restriction enzyme which was not ligated with T4 DNA ligase. Figure 4.10c shows single digest with NotI-HF which was ligated act as a positive control as colonies are expected to be seen. The digestion incubation was increased from 6 hours to overnight incubation to ensure complete digestion of the plasmid. Colonies were observed in the unligated plates (Figures 4.10 a and b) due to the plasmid being cut once producing compatible ends, hence the plasmid may have self-ligated. Figures 4.10b and c were compared against each other and showed no difference in the number of colonies. Ideally, Figure 4.10 b should have fewer colonies than 4.10c.

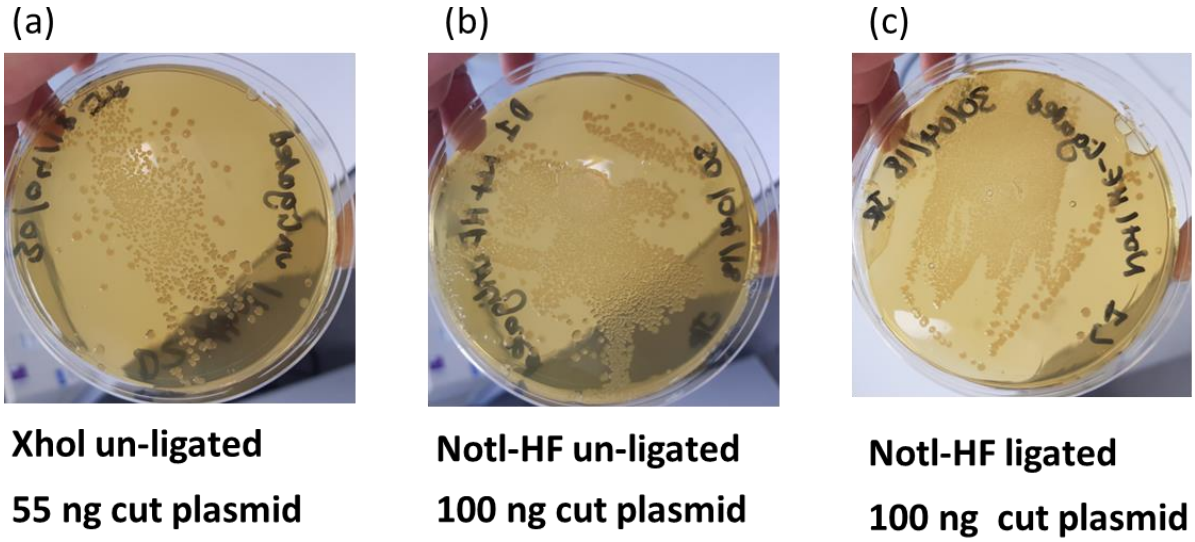


Figure 4. 10 Colonies seen on negative control ampicillin LB agar plates (a) and (b) after ligation and transformation. Two single digestions of plasmid, one with NotI-HF and one with XhoI which were left overnight instead of 6 hours previously used to ensure complete digestion. (a) Transformation of single digest of the plasmid with XhoI without ligase. (b) Transformation of single digest of the plasmid with NotI-HF without ligase. (c) Transformation of single digest of the plasmid with NotI-HF with ligase. The ligation reaction was performed on NotI-HF digested plasmid but not XhoI digested plasmid due to low plasmid yields.

4.1.5 Validating plasmid digest

It was not clear from the LB plates in Figure 4.10 whether the digestion has worked after increasing the plasmid digestion time from 6 hours to overnight. To check the extent of digestion by both restriction enzymes, plasmid pcDNA™4/TO was digested separately by each restriction enzyme and run on 1 % w/v agarose gel to visualise the digested products. The undigested plasmid was used as a positive control.

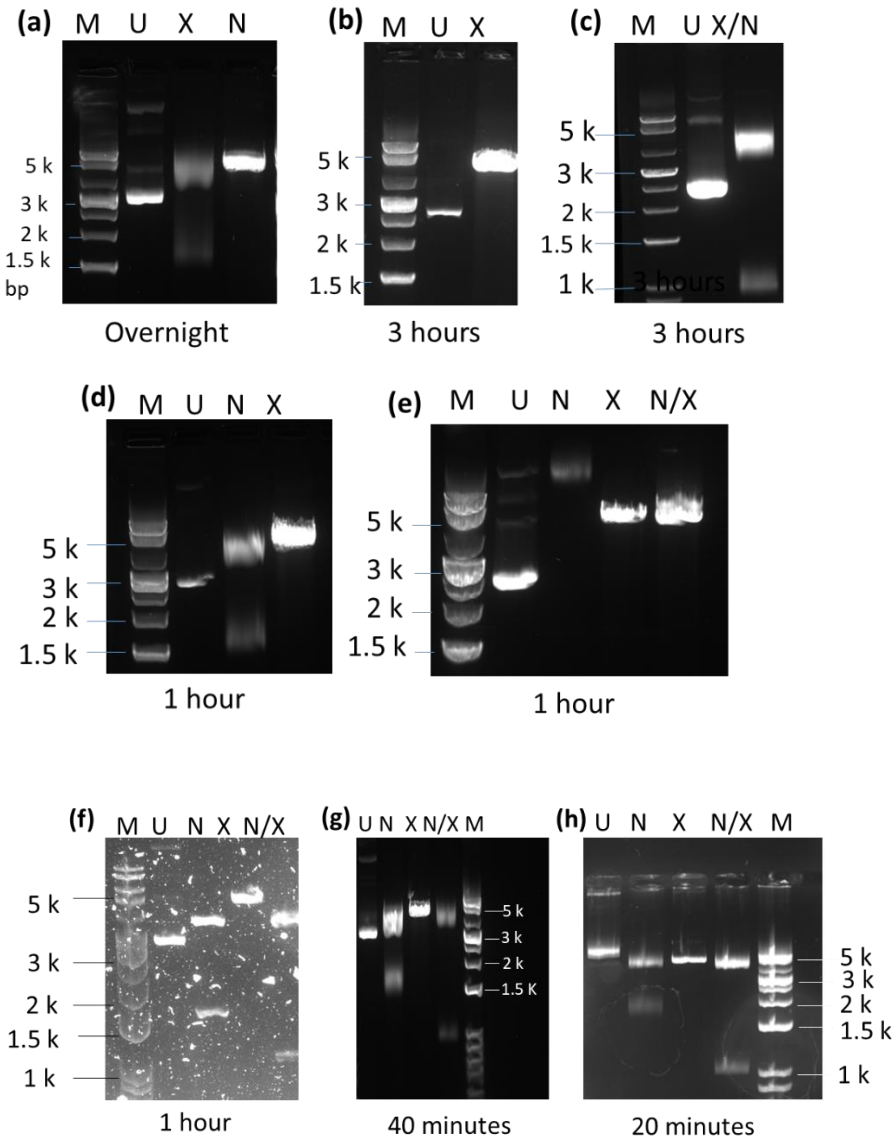


Figure 4. 11 Optimisation of plasmid endonuclease restriction digest by decreasing the digestion incubation time (labelled at the bottom of each gel) at 37 ^{o}C. Plasmid digests were viewed on 1% w/v agarose gel. (a) to (e) had 1 µg of the plasmid in each lane while (f) to (g) had 500 ng of the plasmid in each lane. M= 1kb ladder, U=Uncut plasmid, N=NotI-HF, X=XhoI and N/X= NotI-HF/XhoI.

Figure 4.11 shows the optimisation of plasmid digestion time. There were three bands present in the uncut plasmid. Most uncut plasmids appeared around 3 kb rather than 5kb due to its supercoiled nature, hence it runs faster on a gel. The other forms that it can take is nicked plasmid DNA which appears higher than 5 kb. Also, a band between supercoiled and nicked plasmid is linear plasmid which can occur due to harsh treatment during the plasmid purification procedure.

Initially, the plasmid was digested separately by two different restriction enzymes, NotI-HF and XhoI and was incubated at 37 °C overnight. As shown in Figure 4.11a, two bands are observed at 5 kb and 1.8 kb when plasmid was incubated with XhoI. One explanation is that XhoI may be carrying out star activity due to the restriction enzyme being left for an extended period i.e. overnight. Hence, there was a second cleavage. Figure 4.11b show an expected single band of plasmid digested with XhoI after the reduction to 3 hours. However, when the plasmid was double digested with NotI-HF and XhoI at 3 hours, there was a presence of two bands (Figure 4.11c). Hence, the incubation time was further reduced to an hour (Figures 4.11d, e and f). It showed that plasmid digested with NotI-HF had the presence of two bands. To eliminate the presence of multiple bands from a digested plasmid with NotI-HF, incubation time was further reduced to 40 and 20 minutes. However, the presence of a second band was still observed.

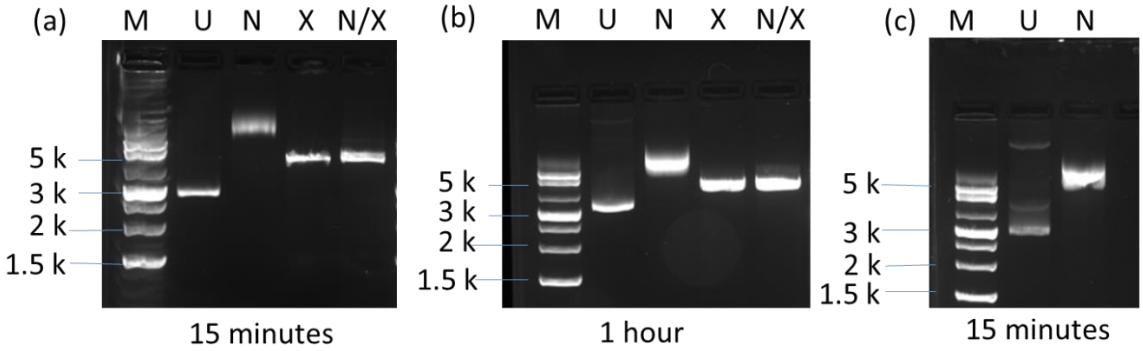


Figure 4. 12 Optimising plasmid digest by decreasing the digestion incubation time further (labelled bottom of each gel) and using new restriction enzyme, NotI-HF due to star activity seen in the previous gels. All gels are made from 1% w/v agarose gel (a) to (c) had 500 ng of the plasmid in each lane. (c) 5 X concentration of NotI-HF was used. M=1 kb ladder, U=uncut plasmid, N=NotI-HF, X=XhoI and N/X= Not I-HF/XhoI.

A new NotI-HF was purchased from NEB since the previous NotI-HF had problems as it cleaved the plasmid more than one place and decreasing the incubation time did not affect. The plasmid was incubated with new restriction enzyme NotI-HF for 15 minutes but the band expressed appeared to be higher than the expected size of 5kb (Figure 4.12). It may possibly due to the digestion being incomplete. So, the digestion incubation time was increased again to an hour (Figure 4.12 b). The band remained to be higher than 5 kb. To ensure complete digestion, five times more NotI-HF was added to the circular plasmid and incubated for 15 minutes. It still did not make a difference to the plasmid size.

4.1.6 Verifying plasmid pcDNA™4/TO to confirm its identity

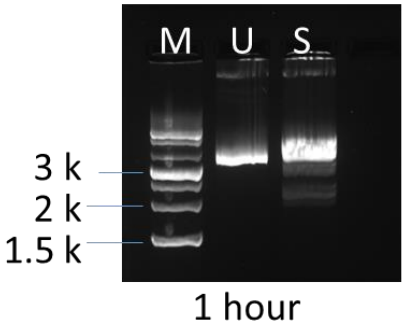


Figure 4. 13 Verifying plasmid Plasmid pcDNA™4/TO using SalI restriction enzyme on 1% w/v agarose gel. 500 ng of plasmid was used in each lane. SalI was incubated with the plasmid at 37 ^{o}C for an hour. M=1kb ladder, U=Uncut plasmid and S=SalI.

Previously in Figure 4.12, a single cut plasmid with NotI-HF gave a product size higher than the expected plasmid length. To confirm the size of the plasmid, a restriction enzyme SalI was used since it cuts at positions 863 bp and 2892 bp of plasmid pcDNA™4/TO. As shown in figure 4.13, two product sizes around 2 kb and 3 kb were expressed confirming the plasmid is around 5 kb. Additionally, the plasmid identity was confirmed by sequencing the entire MCS region using CMV_for and BGH_rev primers. Figure 4.14 shows a band observed around 340 bp of the MCS of plasmid pcDNA™4/TO was extracted and sequenced. Sequencing confirmed the band is the MCS of plasmid pcDNA™4/TO and the actual size is 338 bp (Figure 4.15) and the NotI site can be seen (Figure 4.15).

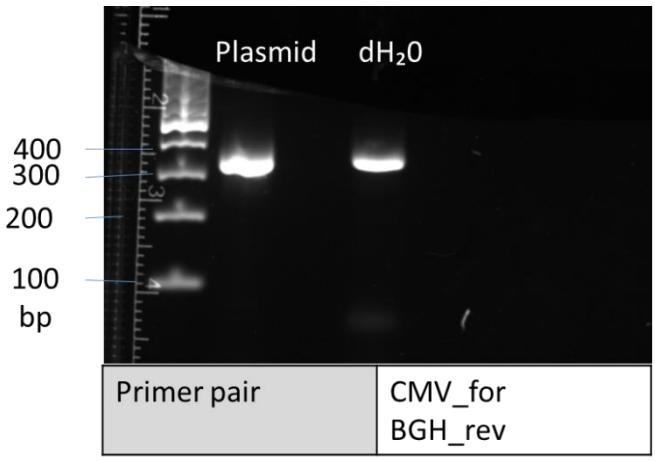


Figure 4. 14 RT PCR amplifying the multiple cloning site (MCS) of the plasmid to verify via Sanger sequencing. 40 ng of plasmid was used in the PCR reaction which was run on 1% w/v ultrapure gel. M is the 100 bp DNA ladder**.**

```
Composite        --------GGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCTCCCTATCA      52
PCDNA4/TO        CGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCTCCCTATCA      60
                         ***************************************************

Composite        GTGATAGAGATCTCCCTATCAGTGATAGAGATCGTCGACGAGCTCGTTTAGTGAACCGTC      112
PCDNA4/TO        GTGATAGAGATCTCCCTATCAGTGATAGAGATCGTCGACGAGCTCGTTTAGTGAACCGTC      120
                 ************************************************************

Composite        AGATCGCCTGGAGACGCCATCCACGCTGTTTTGACCTCCATAGAAGACACCGGGACCGAT      172
PCDNA4/TO        AGATCGCCTGGAGACGCCATCCACGCTGTTTTGACCTCCATAGAAGACACCGGGACCGAT      180
                 ************************************************************

Composite        CCAGCCTCCGGACTCTAGCGTTTAAACTTAAGCTTGGTACCGAGCTCGGATCCACTAGTC      232
PCDNA4/TO        CCAGCCTCCGGACTCTAGCGTTTAAACTTAAGCTTGGTACCGAGCTCGGATCCACTAGTC      240
                 ************************************************************
                                                NotI
Composite        CAGTGTGGTGGAATTCTGCAGATATCCAGCACAGTGGCGGCCGCTCGAGTCTAGAGGGCC      292
PCDNA4/TO        CAGTGTGGTGGAATTCTGCAGATATCCAGCACAGTGGCGGCCGCTCGAGTCTAGAGGGCC      300
                 ************************************************************

Composite        CGTTTAAACCCGCTGATCAGCCT---------------    315
PCDNA4/TO        CGTTTAAACCCGCTGATCAGCCTCGACTGTGCCTTCTA    338
                 **********************
```

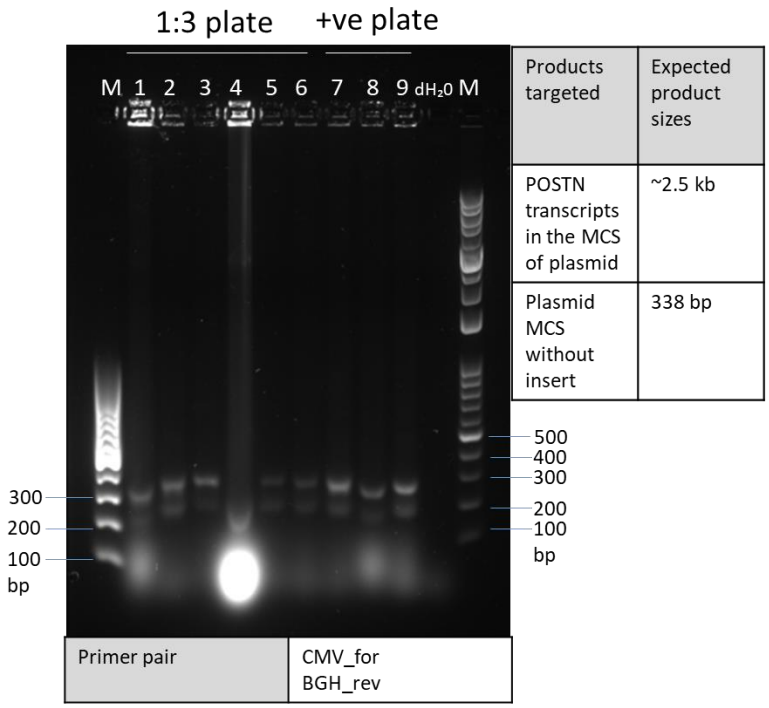
Figure 4. 15 Sequence alignment of composite sequence from CMV and BGH primers against the MCS sequence of plasmid pcDNA™4/TO. Forward and reverse primers are highlighted in yellow and green respectively. NotI site is marked in purple in the MCS.

4.1.7 Screening for potential clones by colony PCR

Plasmids and inserts were prepared by double digesting it with NotI-HF and XhoI creating cohesive ends. To help reduce the background colonies observed from re-ligated vector in the negative control plate (double digested plasmid with no insert that was ligated), the digested plasmid was treated with phosphatase enzyme (Quick CIP, NEB). This helps remove 5’ phosphate groups which in theory prevents re-ligation of the plasmid. Digested

plasmid treated with phosphatase was purified and quantified. The digested plasmid and insert were ligated together in the following insert: vector molar ratios, 1:1, 3:1 and 1:3. The ligated products were transformed into competent *E.coli* DH5α cells via the heat shock method. The cells were grown on ampicillin treated LB agar plates to select cells containing intact plasmid pcDNA™4/TO which has the ampicillin resistance gene. A positive control using undigested plasmid were grown in the same medium. Colonies were observed in all plates including negative control (images not shown).

Colonies of potential clones from (insert: vector molar ratio) 1:3 plate and empty plasmid from the positive control plate were picked to be screened by colony PCR. Primers from the MCS (CMV_for and BGH_rev) of plasmid pcDNA™4/TO were chosen to amplify the vector and insert using the cycling conditions mentioned in Chapter 2. Amplification of the empty vector with these primers results in a PCR product size of around 338 bp but with insert, the expected product size is around 2.5 kb which can be easily detected on an agarose gel. Unfortunately, the cloning of POSTN inserts into plasmid pcDNA™4/TO was unsuccessful (Figure 4.16). Two bands were observed in all lanes. The top band around 340 bp is the MCS of plasmid pcDNA™4/TO. The bottom band remains a mystery and needs to be sequenced to confirm its identity.



Products targeted	Expected product sizes
POSTN transcripts in the MCS of plasmid	~2.5 kb
Plasmid MCS without insert	338 bp

Figure 4. 16 Screening for potential clones by colony PCR on 1% w/v agarose gel. The plasmid was digested using NotI-HF and XhoI. Double digested plasmids were ligated with the insert with the following insert to plasmid ratios: 1:1, 1:3 and 3:1. Controls include undigested plasmid as positive control and double digested linear plasmid without the insert as a negative control. Ligation and transformation were carried out and colonies were viewed the following day. Six (lanes 1-6) and three (lanes 6-9) transformants were picked from 1:3 and +ve control ampicillin LB agar plates respectively to screen for potential clones. M =DNA ladder. Two DNA ladders were used. 100 bp ladder on the left and 1 kb ladder on the right.

4.1.8 Redesign of cloning strategy and PCR primers

In the initial primer strategy, restriction sites NotI and XhoI were close together as displayed in the MCS of plasmid pcDNA™4/TO (Figure 4.9). This means during double digestion of plasmid; the two enzymes bind close together may sterically inhibit each other. For this reason, a new cloning strategy was employed to choose alternative enzymes that have a greater distance between their recognition sites. NotI restriction enzyme require an addition of 10 bp upstream of the recognition site for efficient cleavage (Costa *et al.,* 1994). This may be the reason why there was inefficient cleavage of NotI restriction site. The new cloning primers were designed in a similar approach as described in section 4.1.1 and is depicted in figure 4.17.

The distance between NotI and XhoI sites was only 6 bp. Thus, two different restriction sites HindIII and ApaI were chosen as they were 90 bp apart and did not cleave within the ORF of POSTN-203 as well as with different POSTN splice variants. As shown in figure 4.17, the forward cloning primer consists of the same gene sequence as the initial cloning primer but with HindIII restriction sequence added upstream of the primer. Addition of six random bases was added upstream of the restriction site to enable the restriction enzyme to cleave efficiently as recommended by NEB. The six random base pairs were chosen carefully to avoid palindromes and primer dimers. Similarly, the new cloning reverse primer was composed of the same gene sequence and the 6X histidine tag as with the previous cloning reverse primer apart from a different restriction enzyme site, ApaI, after the stop codon. Again, the addition of six random base pairs was added after the restriction site. The reverse primer was reverse complemented to enable amplification of POSTN insert.

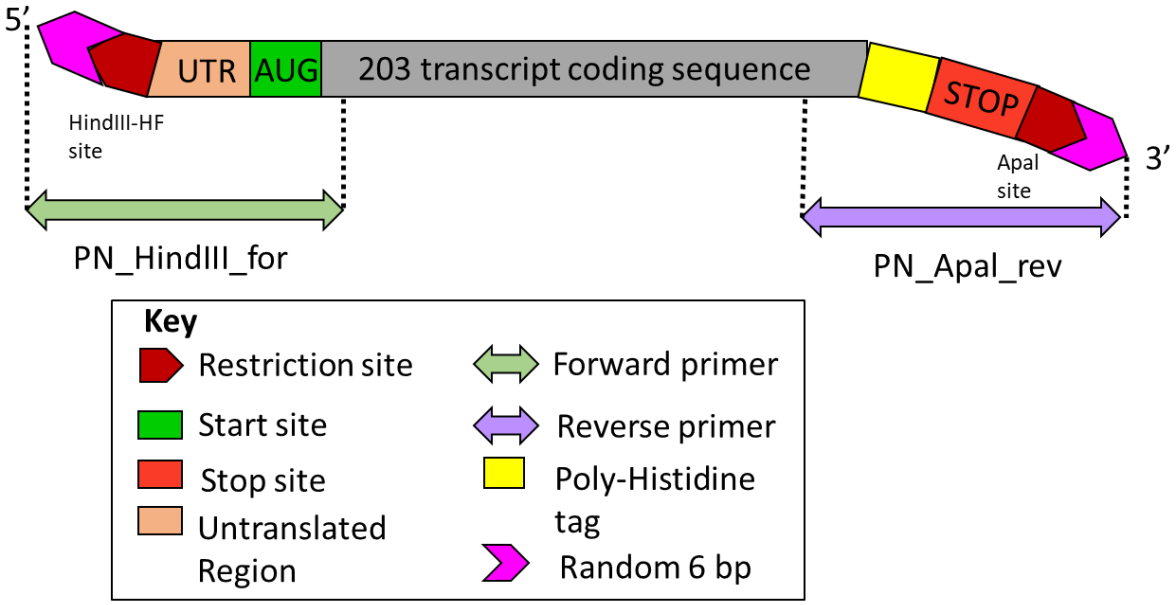


Figure 4. 17 Schematic diagram mapping cloning primers on POSTN-203 transcript coding sequence.

The forward primer (PN_203cdna_for) binds in the 5' end of the POSTN-203 sequence while the reverse primer is located at the 3' end of POSTN-203 sequence to amplify the open reading frame (ORF) for cloning. Each cloning primer has different elements added to them.

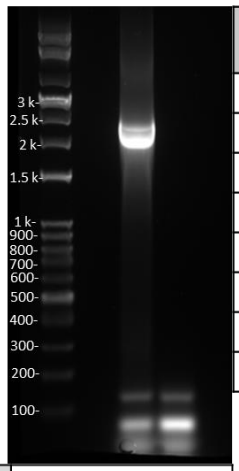
Table 4. 2 Cloning primers sequences used in the second strategy to amplify all POSTN isoforms.

Primer name	Restriction sites	Primer sequence (5'-3')	Tm
PN_HindIII_for	HindIII-HF	TTTCGCAAGCTTCGGAGAGACTCAAGATGATTCCC	67
PN_Apal_rev	Apal	CAAGTCGGGCCCTCAATGGTGATGGTGATGATGCTGAGAACGACCTTCCCT TAATCG	75

Different restriction sites are incorporated in the primers are underlined. A green highlight is the start site in the forward primer and the red highlight is the stop site in the reverse primer. The grey highlight is the 5' UTR in the forward primer. 6 X Histidine-tags (in yellow) is added in the reverse primer before the stop site. Addition of random base pairs (in purple) is added after restriction sites to enable cleavage efficiency. Tm values of primers were calculated using NEB Tm calculator.

4.1.9 PCR amplification of POSTN variants

The new cloning primers designed in section 4.1.8 were used to amplify the insert to be used for cloning. The PCR conditions have been previously optimized with the initial cloning primers: Tm was set at 68 °C which was run at 32 cycles. Figure 4.18 shows a successful amplification of POSTN transcript variants at between 2 kb to 2.5 kb. Again, the large PCR fragments of POSTN transcript variants could not be resolved well. Primer dimers are around 150 bp appear faint. After PCR amplification, the PCR products were purified to remove the enzymes that may hinder in the restriction digestion reaction.



Transcripts targeted	Expected product size (bp)
001	2522
002	2351
003	2441
004	2438
201a	2357
201b	2267
202a	2261
202b	2177

Primer pair	PN_HindIII_for PN_Apal_rev

Figure 4. 18 PCR amplification of POSTN transcript variants using new cloning primers on 1% w/v agarose gel. Expected product sizes are adjacent to the gel image. The primer used is detailed at the bottom of the gel. M is the 1 kb ladder. Water used as reagent control.

4.2.0 Validating plasmid digest and ligation

The plasmid pcDNA™4/TO was extracted from *E.coli* XL1 strain and quantified. Two single digests were carried out using ApaI and HindIII-HF to digest the plasmid. To check for successful digestion by restriction endonucleases, the digested products along with uncut circular plasmid were run on 1% w/v agarose gel (Figure 4.19). The linearized plasmid from the single digest were around 5 kb showing complete digestion.

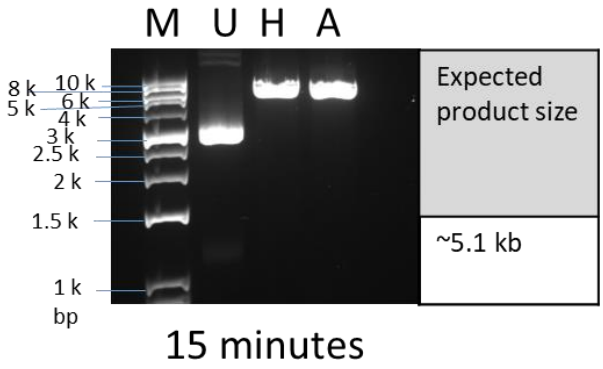


Figure 4. 19 Validating single plasmid digests using two new restriction enzymes (HindIII-HF and ApaI) on 1% w/v agarose gel. 500 ng of the plasmid used in each lane. The restriction endonuclease reaction was incubated at 37 ^{o}C for 15 minutes. M= 1 kb DNA ladder, U= uncut plasmid, H=HindIII-HF and A=ApaI

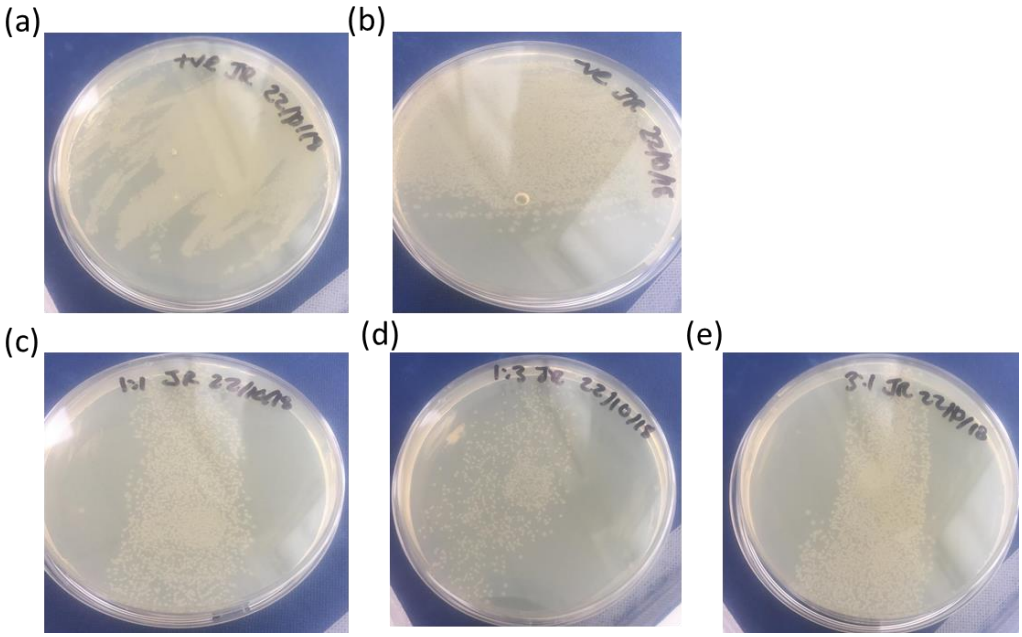


Figure 4. 20 Viewing colonies after ligation and transformation. Before ligation plasmid and insert were double digested using HindIII-HF and ApaI restriction enzymes. (a) The positive control is an undigested plasmid. (b) The negative control is double digested plasmid with no insert. (c) to (e) double digested plasmids ligated with insert in different molar ratios (1:1, 1:3 and 3:1).

The plasmid was double digested with HindIII-HF and ApaI and ligated with a double digested insert at different insert: vector molar ratios (1:1, 1:3 and 3:1). The uncut plasmid was used as a positive control. Negative control was double digested plasmid treated with ligase. Ligated products were transformed by heat shock method and were plated on LB-treated ampicillin agar plates. Figure 4.20 show colonies observed on all plates. A lawn was detected in the positive control plate (Figure 4.20a). Interestingly, the negative control had far more colonies than plasmid ligated with inserts. The negative control was positive indicating there was incomplete digestion although Figure 4.19 shows complete digestion.

Colonies were picked to screen for potential clones from 3:1 insert: vector molar ratio plate as well as empty plasmid from positive control plate. Colonies were screened as mentioned in section 4.1.7. Figure 4.21 show band at 338 bp indicating just the MCS of plasmid pcDNA™4/TO with the absence of insert.

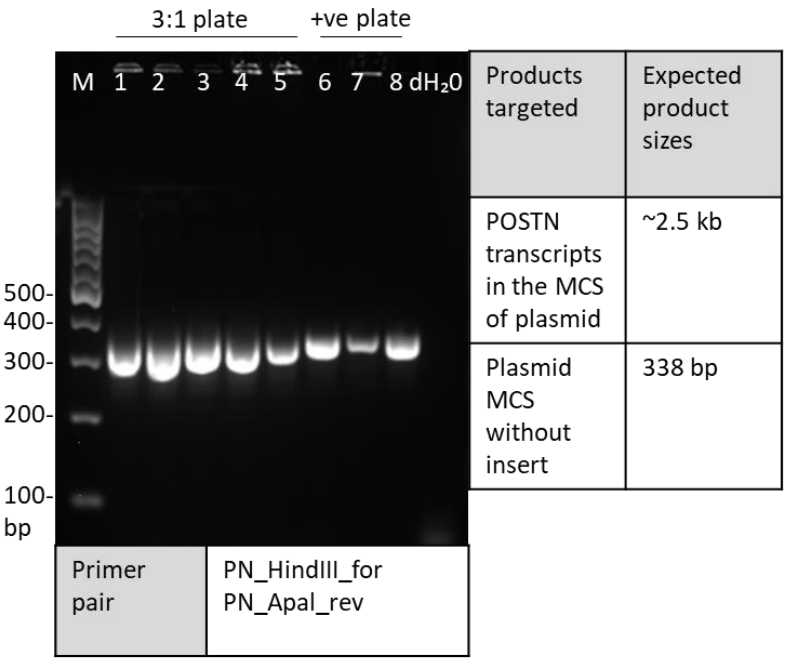


Figure 4. 21 Gel analysis of potential clones using colony PCR on 2% w/v agarose gel. The plasmid was digested using HindIII-HF and ApaI. Double digested plasmids were ligated with the insert with the following insert to plasmid ratios: 1:1, 1:3 and 3:1. Controls include undigested plasmid as positive control and double digested linear plasmid without insert (see figure 4.20 images of colonies). The transformation was carried out and colonies were viewed the following day. Lanes 1-5 are transformants picked from 3:1 plate and lanes 6-8 are transformants picked from +ve control plate. dH_20 used as a negative control on the gel. M =100 bp DNA ladder.

Since colonies were seen in the negative control plate and the ligation had not worked (Figure 4.20), extra controls were put in place. Single plasmid digest was carried out using HindIII-HF restriction enzyme. It was purified using a PCR purification kit and quantitated using Biodrop. Half of the digested plasmid was dephosphorylated but not ligated which was the “no ligation control”. The second half of digested plasmid was not treated with phosphatase but ligated. The uncut plasmid was used as a positive control to check for the viability of competent cells and checks the antibiotic resistance of the plasmid. Negative control of untransfected competent *E.coli* DH5α cells was used, no colonies should be seen since the competent cells did not uptake the plasmid that confers ampicillin resistance. As shown in figure 4.22, a lawn was seen in a positive control as expected (Figure 4.22a). Figure 4.22b and c were compared and showed a similar amount of colonies. Ideally, figure 4.22b should have fewer colonies since the plasmid has been dephosphorylated and not ligated.

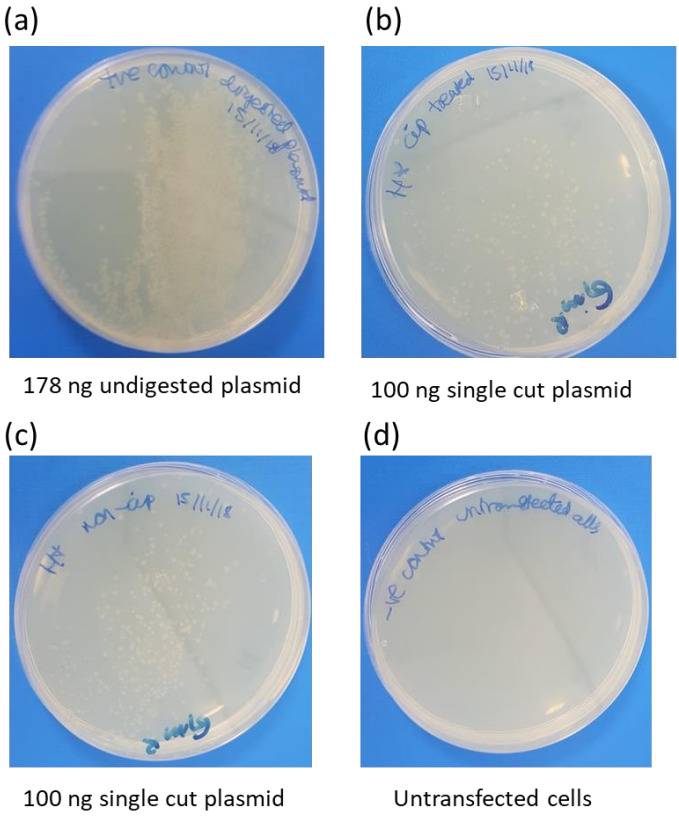


Figure 4. 22 Colonies seen after ligation and transformation of the single digested plasmid. 1 µg of plasmid was digested using Hind III-HF before ligation. (a) A positive control is an undigested plasmid. (b) Single plasmid digest using HindIII-HF treated with cip but unligated. (c) Single plasmid digest using HindIII-HF, no cip treatment but ligated using ligase. (d) Negative control untransfected *E.coli* DH5α cells. All transformations were performed using 12.5 µl competent *E.coli* DH5α cells rather than the standard 50 µl competent *E.coli* DH5α cells.

As colonies were seen in the unligated single digested plasmid treated with phosphatase, it shows that the background is due to uncut vector (Figure 4.22b). Therefore, gel purification of digested plasmid was used instead of on-column purification. Figure 4.23 shows the gel image of the single digested plasmid along with the uncut plasmid. These bands were excised and extracted for the ligation reaction. The ligation and transformation of single digested plasmid were conducted the same way as shown in Figure 4.22. A new T4 DNA ligase was purchased from NEB was used instead of the previous T4 DNA ligase (Promega). As shown in figure 4.24, there were fewer colonies observed with unligated single digested plasmid treated with phosphatase compared with a ligated single digested plasmid with no treatment of phosphatase (Figure 4.24 c and d).

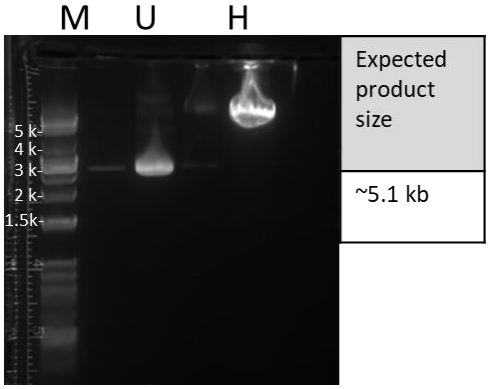


Figure 4. 23 Validating single plasmid digest using HindIII-HF restriction enzyme on 1% w/v ultrapure gel. 1 µg of the plasmid used in each lane. The plasmid was incubated with HindIII-HF at 37 ^{o}C for 15 minutes. Bands were excised for gel purification and then quantified, ready for ligation reaction. M= 1kb DNA ladder, U=uncut plasmid and H=plasmid digested with HindIII-HF

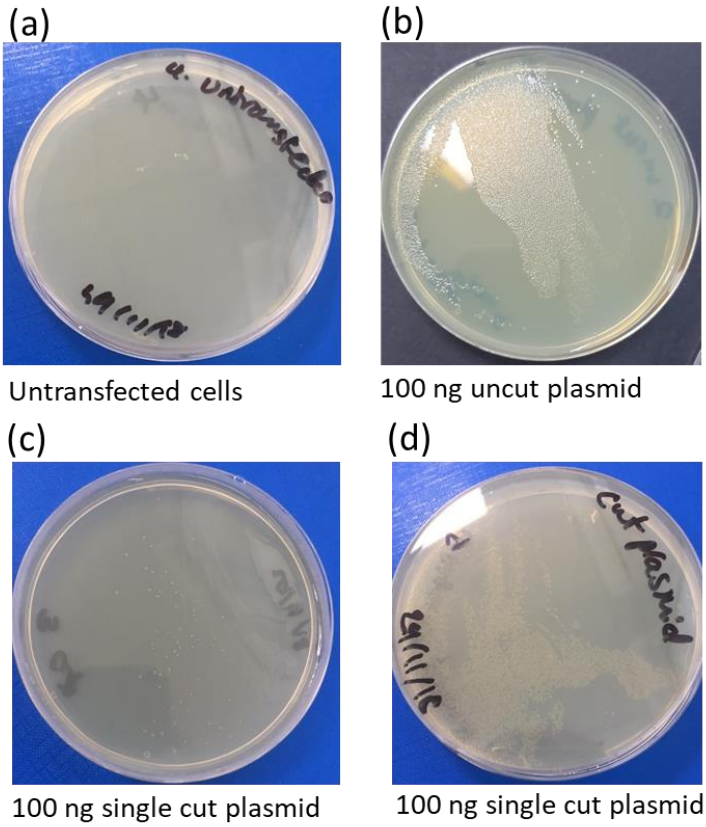


Figure 4. 24 Colonies seen after ligation and transformation of the single digested plasmid. 1 µg of plasmid was digested using Hind III-HF and gel purified before ligation. (a) Negative control of untransfected cells. (b) Positive control uncut plasmid. (c) Single plasmid digest using HindIII-HF treated with quick cip and no ligase. (d) Single plasmid digest using HindIII-HF with no quick cip treatment but with T4 DNA ligase (Monarch). All transformations were performed using 12.5 µl competent *E.coli* DH5α cells rather than the standard 50 µl competent *E.coli* DH5α cells.

4.2.0 Optimising the heat shock transformation step as a positive control was negative

Figure 4.24 shows the ligation reaction has worked. The following experiment was set up the same manner as in Figure 4.24 but digesting plasmid with Apal restriction enzyme instead of HindIII-HF. Two more ligation reactions were added; double digested plasmid treated with phosphatase using HindIII-HF and Apal and ligating with/without insert using the 1:1 insert: vector molar ratio. All digested plasmids were gel purified. Unfortunately, no colonies were observed after ligation and transformation including the positive plate with an undigested plasmid (image not shown). This indicates that transformation has failed.

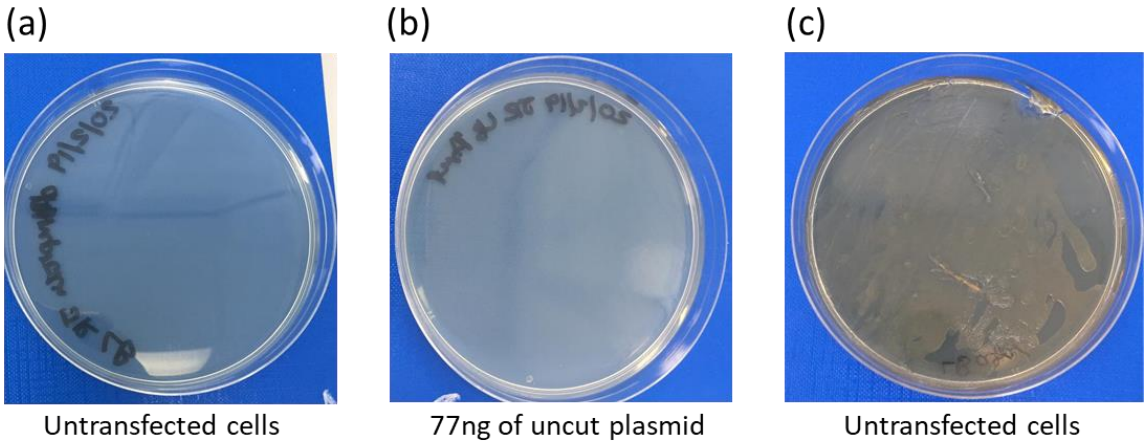


Figure 4. 25 Viewing colonies after transformation of intact circular plasmid. (a) Untransfected competent *E.coli* DH5α cells on LB agar plate treated with ampicillin. (b) Positive control uncut plasmid. (c) Untransfected DH5α cells added on LB agar plate with no ampicillin. All transformations were performed using 25 µl of competent *E.coli* DH5α cells.

As there was an issue with the heat shock transformation step since the positive control was negative, positive control with uncut plasmid was used to see if the transformation has worked. The negative control with untransfected competent *E.coli* DH5α cells was used. Another positive control with untransfected competent *E.coli* DH5α cells was grown

on LB agar plates without ampicillin treatment to test if the cells are viable. Figure 4.25c show that competent *E.coli* DH5α cells are viable since there was growth observed in the non-ampicillin LB agar plate. The positive control was negative showing transformation has not worked (Figure 4.25b).

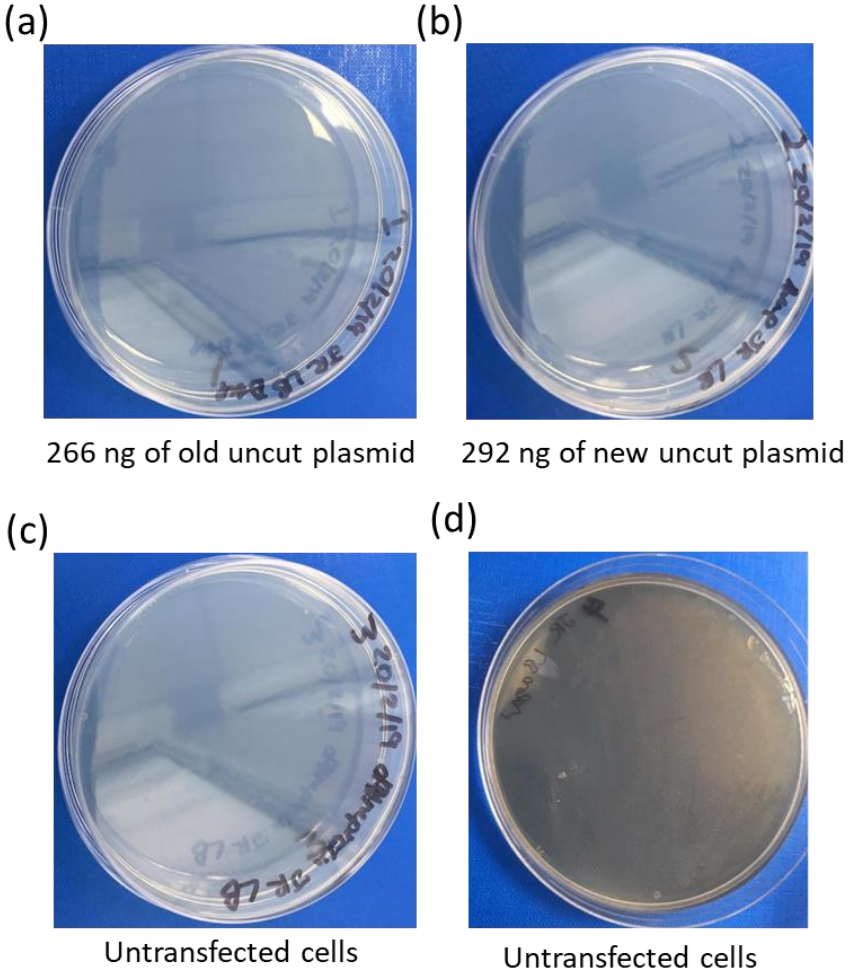


Figure 4. 26 Viewing colonies after transformation of uncut plasmid using competent E.coli JM109 competent cells to check for successful transformation. (a) Transformed old uncut plasmid on selective ampicillin plate. (b) Transformed new uncut plasmid on selective ampicillin plate. (c) Untransfected *E.coli* JM109 competent cells on selective ampicillin plate. (d) Untransfected *E.coli* JM109 cells on LB agar plate with untreated ampicillin. Transformations were performed using 50 µl of JM109 competent cells.

To check for successful transformation via heat shock method, alternative competent cells were used: *E.coli* JM109 cells made in the laboratory. Two positive controls consist of old and newly extracted plasmids to check for transformation. Untransfected competent

E.coli JM109 cells were used on LB plates with/without ampicillin. The untransfected competent *E.coli* JM109 cells did not grow on LB agar plates without ampicillin treatment. This showed that the competent cells were killed off during heat shock transformation step. The lab-grown competent *E.coli* JM109 cells came in 200 µl aliquots in 0.5 ml Eppendorf tubes. One possibility is that the 0.5 ml tube had thin walls which killed the cells during heat shock transformation.

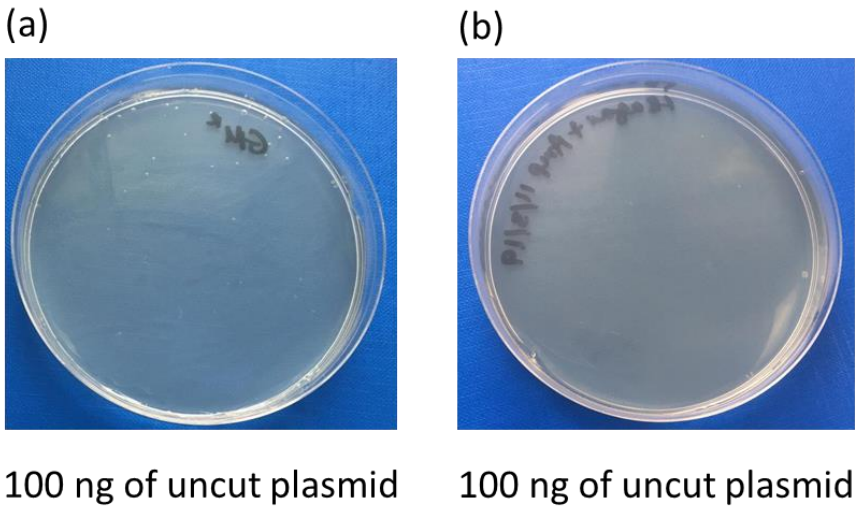


Figure 4. 27

Transformation of intact plasmids using two different plasmids in E.coli JM109 competent cells. (a) Gentamicin treated LB agar plate using 100 ng of plasmid PUC4975 was transformed in 200 µl of competent *E.coli* JM109 cells. (b) Ampicillin treated LB agar plate using 100 ng of plasmid pcDNA™4/TO was transformed in 200 µl of competent *E.coli* JM109 cells. Transformation procedure was conducted by a researcher experienced in this technique.

A control plasmid PUC4975 was transformed using heat shock method in *E.coli* JM109 lab grown competent cells along with plasmid pcDNA™4/TO. The transformed competent *E.coli* JM109 cells were grown on LB agar plates treated with gentamicin or ampicillin. Figure 4.27 shows no growth in either of the plates showing transformation failure. Subsequent experiments in the lab have shown that this was due to lack of competency in the cells used, and future experiments by other students are planned to complete this work.

Chapter 5- Discussion

This study has shown that at least six POSTN transcript variants (POSTN-001, 002, 003, 004, 201a and 201b), including one not previously described, are expressed in pancreatic stellate cells. Stromal cells are known to be important in the development and progression in pancreatic cancer. Additional variants identified in the Ensembl and Genbank databases (POSTN-202a and 202b) were not found to be expressed. Although due to time constraints we were not able to determine the function of these isoforms, it is interesting to speculate on the potential function of these variants and their role in pancreatic cancer. The primary POSTN transcript contains 23 exons while the rest of the POSTN transcript variants have missing one or more cassette exons. This is due to alternative splicing in the C-terminal region of POSTN protein affecting exons 17-21.

Interestingly, a novel POSTN isoform was expressed in PS-1 cells which contain exons 17, 19 and 20 which was not identified in bioinformatics analysis. This novel isoform was found to be expressed in fetal lung known as isoform 9 (Morra *et al.*, 2012). Erkan *et al.* (2007) observed the autocrine loop of POSTN on pancreatic stellate cells (Figure 1.7). It can be speculated that these POSTN transcript variants may perpetuate the activation of pancreatic stellate cells to increase the fibrotic stroma observed in PDAC.

The expression of different POSTN transcript variants have been studied in few cancers such as in bladder, thyroid, renal and small lung (Kim et al., 2008; Bai et al., 2010; Morra et al., 2011;Morra *et al.*, 2012). Functional analysis of POSTN transcript variants in bladder cancer has been studied. In bladder cancer, isoform 4 which corresponds to POSTN-201b was shown to promote cell invasion and metastasis whereas isoforms 1 and 3 corresponds with POSTN-001 and POSTN-201a, inhibited these cellular processes. In context with pancreatic cancer, POSTN-201b, POSTN-001 and POSTN-201a which are expressed in pancreatic stellate cells in this study may have varying roles in inhibiting or promoting pancreatic cancer development. It can be speculated that POSTN-201b may promote pancreatic cancer development. However, these conclusions may not be true as POSTN transcript variants expression are tissue dependent meaning different tissues have different POSTN transcript variant profiles.

Additional POSTN transcript variants g and i from AceView differ at the 5' end. These variants have missing exons 1, 2 and 3 and thus, would not have been detected in this study. POSTN is known to be extracellular and secreted protein as it contains a signal peptide. These POSTN isoforms do not possess a signal peptide and EMI domain hence appear to be intracellular. Further experiments will be required to test these variants are expressed in PS-1 cells, what their location and function maybe.

There are limitations to this study. RT PCR utilized in this study only gave a semi-quantitative measure. To obtain a quantitative measure of POSTN transcript variants, QRT PCR can be employed. This will enable to get an exact measurement of the mRNA expression levels of POSTN transcript variants as this method has both high specificity and sensitivity (Smith & Osborn, 2008; Kuang *et al.*, 2018).

It was difficult to resolve multiple transcript variants by agarose gel electrophoresis- we cannot exclude the presence of additional isoforms. An alternative method may be using higher % agarose gel although the disadvantage of this is that it will be difficult to extract the PCR product for sequencing or using in downstream experiments. Polyacrylamide gels can be used to help resolve the bands although it is toxic compared with agarose. However, cloning may be the best answer of which variants are expressed, as well as determining the function.

Unfortunately, the cloning experiments were unsuccessful within the timeframe of this project. But significant progress was made with refining the cloning strategy and troubleshooting which will be continued by another student.

Different bacterial strains enable the production of mammalian proteins (Donnelly et al., 2001). Protein toxicity has been described as one of the reasons for unsuccessful production of recombinant DNA/ protein (Rosano & Ceccarelli, 2014). There is a possibility that human POSTN transcript may be toxic in *E.coli* cells since no clones had any cDNA insert but just with empty plasmid vector (Figures 4.16 and 4.21). A recent study showed *E.coli* failed to overexpress a human gene called Human ether-a-gogo related gene (hERG) (Vasseur et al., 2019) but worked well using stable human embryonic kidney (HEK) cell line. Different strategies for producing toxic recombinant proteins in

E.coli has been extensively reviewed (Saïda *et al.,* 2006). In this study is it likely to be due to technical challenges as the main reason for difficulties.

Since a wide range of POSTN isoforms were detected in PS-1 cells, the next step will be to successfully clone these POSTN transcript variants. With regards to the cloning procedure, it is important to have both negative and positive controls, as displayed in figure 4.24 along with double digested plasmid ligated with an insert at different ratios (1:1, 1:3 3:1). Gel purification of the digested products appears to be more effective than on-column purification. This is because when the digested plasmid is run on the gel, the linearized plasmid band can be excised while eliminating any undigested plasmid, thus reducing background colonies after transformation. Once POSTN clones of different POSTN transcript variants are successfully obtained, plasmids each expressing a specific POSTN transcript variant can be transfected into human (PS-1) cells to produce recombinant POSTN isoforms. These isoforms attached with His-tag can be purified using Ni columns. The purified proteins added experimentally to pancreatic cancer cells to determine the effect on cell proliferation, survival, adhesion, migration and invasion using standard assays.

POSTN upregulation has been found in the development of PDAC particularly around the neoplastic stroma (Ben *et al.,* 2011). This shows the importance of the tumour microenvironment in mediating tumour progression. Survival studies of patients with PDAC showed POSTN-positive group had lower median survival time compared with POSTN-negative group (Ben *et al.,* 2011). Similarly, Liu *et al.* (2017) showed there is a positive correlation between POSTN expression and clinical stage of PDAC and an increased POSTN expression was linked to poor patient survival. This makes POSTN a good therapeutic target for PDAC. The evidence that POSTN is a good therapeutic target comes from a study by Liu *et* al. (2016) that showed that POSTN contributes to the chemoresistance of PDAC by protecting pancreatic cancer cells against DNA damage induced by gemcitabine. Recently, POSTN is a potentially useful biomarker for early detection of PDAC complementing with other serum biomarkers such as CA242 and CA19.9 (Dong et al., 2018). It would be interesting to see which POSTN transcript variant is linked to PDAC. In our study, POSTN-001. POSTN-003, POSTN-004 and POSTN-201a had the

highest expression in PS-1 cells, although it was only semi-quantitative measure (Table 3.5). It can be speculated that these POSTN isoforms maybe linked with PDAC progression.

In the literature, POSTN is shown to not only be a potential therapeutic target for cancer but is a promising clinical marker. A novel ELISA kit (Biomedica) has been developed which recognises all known POSTN transcript variants by measuring human serum and plasma and this method has been successfully utilized in bone (Walsh *et al.,* 2017). It would be interesting to measure POSTN serum of patients in a different pathological stage of PDAC and see which isoforms correlate with poor prognosis of PDAC. It will be important to determine the expression of POSTN isoforms in PDAC clinical biopsy samples.

5.9 Conclusion and future work

New therapeutic target is required to treat PDAC since conventional methods are ineffective. POSTN is a useful clinical biomarker for PDAC and could potentially be used as a novel therapeutic target. This is the first study to investigate a total of seven POSTN transcript variants: POSTN-001, 002, 003, 004, 201a, 201b and a novel POSTN transcript variant (contain exons 17, 19 and 20) are expressed in PS-1 cells (Table 3.5). The highly expressed variants include POSTN-001, 003, 004 and 201a with the rest of the variants having a lower expression level in PS-1 cells. As primers were designed in the C-terminal region to detect POSTN isoforms, further studies are warranted to study these POSTN transcript variants by cloning these variants to see if functional proteins are produced. This study attempted to produce POSTN clones but was unsuccessful due to technical difficulties. Troubleshooting work for cloning has been laid out in Chapter 4.

It has been hypothesised that POSTN transcript variants may have distinct matrix binding properties and may differentially affect cell adhesion and migration (Viloria & Hill, 2016). Hence, this novel hypothesis remains to be tested. Functional studies need to be conducted to observe the effects in cellular processes such as adhesion, migration and proliferation of pancreatic cancer cells, in order to establish the function of POSTN isoforms in pancreatic cancer.

References

Bai, Y., Nakamura, M., Zhou, G., Li, Y., Liu, Z., Ozaki, T., … Kakudo, K. (2010). Novel isoforms of periostin expressed in the human thyroid. *Japanese Clinical Medicine*, *1*, 13–20. https://doi.org/10.4137/JCM.S5899

Bao, S., Ouyang, G., Bai, X., Huang, Z., Ma, C., Liu, M., … Wang, X.-F. (2004). Periostin potently promotes metastatic growth of colon cancer by augmenting cell survival via the Akt/PKB pathway. *Cancer Cell*, *5*(4), 329–339. https://doi.org/10.1016/S1535-6108(04)00081-9

Baril, P., Gangeswaran, R., Mahon, P. C., Caulee, K., Kocher, H. M., Harada, T., … Lemoine, N. R. (2007). Periostin promotes invasiveness and resistance of pancreatic cancer cells to hypoxia-induced cell death: role of the β4 integrin and the PI3k pathway. *Oncogene*, *26*(14), 2082–2094. https://doi.org/10.1038/sj.onc.1210009

Ben, Q.-W., Jin, X.-L., Liu, J., Cai, X., Yuan, F., & Yuan, Y.-Z. (2011). Periostin, a matrix specific protein, is associated with proliferation and invasion of pancreatic cancer. *Oncology Reports*, *25*(3), 709–716. https://doi.org/10.3892/or.2011.1140

Bian, X., Su, X., Wang, Y., Zhao, G., Zhang, B., & Li, D. (2019). Periostin contributes to renal and cardiac dysfunction in rats with chronic kidney disease: Reduction of PPARα. *Biochimie*, *160*, 172–182. https://doi.org/10.1016/j.biochi.2019.03.003

Biankin, A. V., Waddell, N., Kassahn, K. S., Gingras, M. C., Muthuswamy, L. B., Johns, A. L., … Grimmond, S. M. (2012). Pancreatic cancer genomes reveal aberrations in axon guidance pathway genes. *Nature*, *491*(7424), 399–405. https://doi.org/10.1038/nature11547

Bilimoria, K. Y., Bentrem, D. J., Ko, C. Y., Ritchey, J., Stewart, A. K., Winchester, D. P., & Talamonti, M. S. (2007). Validation of the 6th edition AJCC pancreatic cancer staging system: Report from the National Cancer Database. *Cancer*, *110*(4), 738–744. https://doi.org/10.1002/cncr.22852

Blencowe, B. J. (2006). Alternative Splicing: New Insights from Global Analyses. *Cell*, *126*(1), 37–47. https://doi.org/10.1016/J.CELL.2006.06.023

Brinkley, B. R., Beall, P. T., Wible, L. J., Mace, M. L., Turner, D. S., & Cailleau, R. M. (1980). Variations in cell form and cytoskeleton in human breast carcinoma cells in vitro. *Cancer Research*, *40*(9), 3118–3129. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/7000337

Carrato, A., Falcone, A., Ducreux, M., Valle, J. W., Parnaby, A., Djazouli, K., … Parthenaki, I. (2015, October 1). A Systematic Review of the Burden of Pancreatic Cancer in Europe: Real-World Impact on Survival, Quality of Life and Costs. *Journal of Gastrointestinal Cancer*. Humana Press Inc. https://doi.org/10.1007/s12029-015-9724-1

Chen, M., & Manley, J. L. (2009). Mechanisms of alternative splicing regulation: insights from molecular and

genomics approaches. *Nature Reviews. Molecular Cell Biology*, *10*(11), 741–754. https://doi.org/10.1038/nrm2777

Chuanyu, S., Yuqing, Z., Chong, X., Guowei, X., & Xiaojun, Z. (2017). Periostin promotes migration and invasion of renal cell carcinoma through the integrin/focal adhesion kinase/c-Jun N-terminal kinase pathway. *Tumor Biology*, *39*(4), 101042831769454. https://doi.org/10.1177/1010428317694549

Cioffi, M., Trabulo, S. M., Sanchez-Ripoll, Y., Miranda-Lorenzo, I., Lonardo, E., Dorado, J., … Heeschen, C. (2015). The miR-17-92 cluster counteracts quiescence and chemoresistance in a distinct subpopulation of pancreatic cancer stem cells. *Gut*, *64*(12), 1936–1948. https://doi.org/10.1136/gutjnl-2014-308470

Cooke, C., Hans, H., & Alwine, J. C. (1999). Utilization of splicing elements and polyadenylation signal elements in the coupling of polyadenylation and last-intron removal. *Molecular and Cellular Biology*, *19*(7), 4971–4979. https://doi.org/10.1128/mcb.19.7.4971

Dong, D., Jia, L., Zhang, L., Ma, N., Zhang, A., Zhou, Y., & Ren, L. (2018). Periostin and CA242 as potential diagnostic serum biomarkers complementing CA19.9 in detecting pancreatic cancer. *Cancer Science*, *109*(9), 2841–2851. https://doi.org/10.1111/cas.13712

Donnelly, M. I., Stevens, P. W., Stols, L., Xiaoyin Su, S., Tollaksen, S., Giometti, C., & Joachimiak, A. (2001). Expression of a highly toxic protein, bax, in Escherichia coli by attachment of a leader peptide derived from the GroES cochaperone. *Protein Expression and Purification*. https://doi.org/10.1006/prep.2001.1442

Erkan, M., Kleeff, J., Gorbachevski, A., Reiser, C., Mitkus, T., Esposito, I., … Friess, H. (2007). Periostin Creates a Tumor-Supportive Microenvironment in the Pancreas by Sustaining Fibrogenic Stellate Cell Activity. *Gastroenterology*, *132*(4), 1447–1464. https://doi.org/10.1053/j.gastro.2007.01.031

Farrow, B., Albo, D., & Berger, D. H. (2008). The Role of the Tumor Microenvironment in the Progression of Pancreatic Cancer. *Journal of Surgical Research*, *149*(2), 319–328. https://doi.org/10.1016/j.jss.2007.12.757

Felsenstein, M., Hruban, R. H., & Wood, L. D. (2018). New Developments in the Molecular Mechanisms of Pancreatic Tumorigenesis. *Advances in Anatomic Pathology*. Lippincott Williams and Wilkins. https://doi.org/10.1097/PAP.0000000000000172

Ferlay J, Ervik M, Lam F, Colombet M, Mery L, Piñeros M, Znaor A, Soerjomataram I, Bray F (2018). Global Cancer Observatory: Cancer Today. Lyon, France: International Agency for Research on Cancer. Available at: https://gco.iarc.fr/today (Accessed: 11 June 2019).

Froeling, F. E. M., Mirza, T. A., Feakins, R. M., Seedhar, A., Elia, G., Hart, I. R., & Kocher, H. M. (2009). Organotypic culture model of pancreatic cancer demonstrates that stromal cells modulate E-cadherin, β-catenin, and ezrin expression in tumor cells. *American Journal of Pathology*.

https://doi.org/10.2353/ajpath.2009.090131

Fujita, H., Ohuchida, K., Mizumoto, K., Egami, T., Miyoshi, K., Moriyama, T., … Tanaka, M. (2009). Tumor-stromal interactions with direct cell contacts enhance proliferation of human pancreatic carcinoma cells. *Cancer Science*, *100*(12), 2309–2317. https://doi.org/10.1111/j.1349-7006.2009.01317.x

Fukushima, N., Kikuchi, Y., Nishiyama, T., Kudo, A., & Fukayama, M. (2008). Periostin deposition in the stroma of invasive and intraductal neoplasms of the pancreas. *Modern Pathology*, *21*(8), 1044–1053. https://doi.org/10.1038/modpathol.2008.77

Furuichi, Y., LaFiandra, A., & Shatkin, A. J. (1977). 5′-Terminal structure and mRNA stability. *Nature*, *266*(5599), 235–239. https://doi.org/10.1038/266235a0

Furuyama, S., & Bruzik, J. P. (2002). Multiple roles for SR proteins in trans splicing. *Molecular and Cellular Biology*, *22*(15), 5337–5346. https://doi.org/10.1128/mcb.22.15.5337-5346.2002

Gillan, L., Matei, D., Fishman, D. A., Gerbin, C. S., Karlan, B. Y., & Chang, D. D. (2002). Periostin secreted by epithelial ovarian carcinoma is a ligand for alpha(V)beta(3) and alpha(V)beta(5) integrins and promotes cell motility. *Cancer Research*, *62*(18), 5358–5364. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/12235007

Gnanamony, M., & Gondi, C. S. (2017). Chemoresistance in pancreatic cancer: Emerging concepts (Review). *Oncology Letters*, *13*(4), 2507–2513. https://doi.org/10.3892/ol.2017.5777

Hanahan, D., & Weinberg, R. A. (2011, March 4). Hallmarks of cancer: The next generation. *Cell*. https://doi.org/10.1016/j.cell.2011.02.013

Heestand, G. M., & Kurzrock, R. (2015). Molecular landscape of pancreatic cancer: Implications for current clinical trials. *Oncotarget*, *6*(7), 4553–4561. https://doi.org/10.18632/oncotarget.2972

Hoersch, S., & Andrade-Navarro, M. A. (2010). Periostin shows increased evolutionary plasticity in its alternatively spliced region. *BMC Evolutionary Biology*, *10*(1), 30. https://doi.org/10.1186/1471-2148-10-30

Horiuchi, K., Amizuka, N., Takeshita, S., Takamatsu, H., Katsuura, M., Ozawa, H., … Kudo, A. (1999). Identification and Characterization of a Novel Protein, Periostin, with Restricted Expression to Periosteum and Periodontal Ligament and Increased Expression by Transforming Growth Factor β. *Journal of Bone and Mineral Research*, *14*(7), 1239–1249. https://doi.org/10.1359/jbmr.1999.14.7.1239

Hwang, R. F., Moore, T., Arumugam, T., Ramachandran, V., Amos, K. D., Rivera, A., … Logsdon, C. D. (2008). Cancer-associated stromal fibroblasts promote pancreatic tumor progression. *Cancer Research*, *68*(3), 918–926. https://doi.org/10.1158/0008-5472.CAN-07-5714

Infante, J. R., Somer, B. G., Park, J. O., Li, C. P., Scheulen, M. E., Kasubhai, S. M., … Le, N. (2014). A

randomised, double-blind, placebo-controlled trial of trametinib, an oral MEK inhibitor, in combination with gemcitabine for patients with untreated metastatic adenocarcinoma of the pancreas. *European Journal of Cancer*, *50*(12), 2072–2081. https://doi.org/10.1016/j.ejca.2014.04.024

Izuhara, K., Conway, S. J., Moore, B. B., Matsumoto, H., Holweg, C. T. J., Matthews, J. G., & Arron, J. R. (2016, May 1). Roles of periostin in respiratory disorders. *American Journal of Respiratory and Critical Care Medicine*. American Thoracic Society. https://doi.org/10.1164/rccm.201510-2032PP

Jurado, A. R., Tan, D., Jiao, X., Kiledjian, M., & Tong, L. (2014). Structure and function of pre-mRNA 5'-end capping quality control and 3'-end processing. *Biochemistry*, *53*(12), 1882–1898. https://doi.org/10.1021/bi401715v

Kanno, A., Satoh, K., Masamune, A., Hirota, M., Kimura, K., Umino, J., … Shimosegawa, T. (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), 2707–2718. https://doi.org/10.1002/ijc.23332

Kelemen, O., Convertini, P., Zhang, Z., Wen, Y., Shen, M., Falaleeva, M., & Stamm, S. (2013). Function of alternative splicing. *Gene*, *514*(1), 1–30. https://doi.org/10.1016/j.gene.2012.07.083

Kii, I., Nishiyama, T., Li, M., Matsumoto, K.-I., Saito, M., Amizuka, N., & Kudo, A. (2010). Incorporation of tenascin-C into the extracellular matrix by periostin underlies an extracellular meshwork architecture. *The Journal of Biological Chemistry*, *285*(3), 2028–2039. https://doi.org/10.1074/jbc.M109.051961

Kim, C., Isono, T., Tambe, Y., Chano, T., Okabe, H., Okada, Y., & Inoue, H. (2008). Role of alternative splicing of periostin in human bladder carcinogenesis. *International Journal of Oncology*, *32*(1), 161–169. https://doi.org/10.3892/ijo.32.1.161

Kim, C. J., Yoshioka, N., Tambe, Y., Kushima, R., Okada, Y., & Inoue, H. (2005). Periostin is down-regulated in high grade human bladder cancers and suppressesin vitro cell invasiveness andin vivo metastasis of cancer cells. *International Journal of Cancer*, *117*(1), 51–58. https://doi.org/10.1002/ijc.21120

Kim, J. E., Lee, K. T., Lee, J. K., Paik, S. W., Rhee, J. C., & Choi, K. W. (2004). Clinical usefulness of carbohydrate antigen 19-9 as a screening test for pancreatic cancer in an asymptomatic population. *Journal of Gastroenterology and Hepatology (Australia)*, *19*(2), 182–186. https://doi.org/10.1111/j.1440-1746.2004.03219.x

Kuang, J., Yan, X., Genders, A. J., Granata, C., & Bishop, D. J. (2018). An overview of technical considerations when using quantitative real-time PCR analysis of gene expression in human exercise research. *PLoS ONE*. https://doi.org/10.1371/journal.pone.0196438

Kudo, Y., Ogawa, I., Kitajima, S., Kitagawa, M., Kawai, H., Gaffney, P. M., … Takata, T. (2006). Periostin Promotes Invasion and Anchorage-Independent Growth in the Metastatic Process of Head and Neck Cancer. *Cancer Research*, *66*(14), 6928–6935. https://doi.org/10.1158/0008-5472.CAN-05-4540

Lai, E., Puzzoni, M., Ziranu, P., Pretta, A., Impera, V., Mariani, S., … Scartozzi, M. (2019). New therapeutic targets in pancreatic cancer. *Cancer Treatment Reviews*, *81*, 101926. https://doi.org/10.1016/j.ctrv.2019.101926

Li, C., Heidt, D. G., Dalerba, P., Burant, C. F., Zhang, L., Adsay, V., … Simeone, D. M. (2007). Identification of pancreatic cancer stem cells. *Cancer Research*, *67*(3), 1030–1037. https://doi.org/10.1158/0008-5472.CAN-06-2030

Liu, Y., Li, F., Gao, F., Xing, L., Qin, P., Liang, X., … Du, L. (2016). Periostin promotes the chemotherapy resistance to gemcitabine in pancreatic cancer. *Tumor Biology*, *37*(11), 15283–15291. https://doi.org/10.1007/s13277-016-5321-6

Liu, Y., Li, F., Gao, F., Xing, L., Qin, P., Liang, X., … Du, L. (2016). Periostin promotes tumor angiogenesis in pancreatic cancer via Erk/VEGF signaling. *Oncotarget*, *7*(26), 40148–40159. https://doi.org/10.18632/oncotarget.9512

Liu, Y., Li, F., Gao, F., Xing, L., Qin, P., Liang, X., … Du, L. (2017). Role of microenvironmental periostin in pancreatic cancer progression. *Oncotarget*, *8*(52), 89552–89565. https://doi.org/10.18632/oncotarget.11533

Matsusaka, M., Kabata, H., Fukunaga, K., Suzuki, Y., Masaki, K., Mochimaru, T., … Betsuyaku, T. (2015). Phenotype of asthma related with high serum periostin levels. *Allergology International*, *64*(2), 175–180. https://doi.org/10.1016/j.alit.2014.07.003

Michaylira, C. Z., Wong, G. S., Miller, C. G., Gutierrez, C. M., Nakagawa, H., Hammond, R., … Rustgi, A. K. (2010). Periostin, a cell adhesion molecule, facilitates invasion in the tumor microenvironment and annotates a novel tumor-invasive signature in esophageal cancer. *Cancer Research*, *70*(13), 5281–5292. https://doi.org/10.1158/0008-5472.CAN-10-0704

Mizuma, M., Rasheed, Z. A., Yabuuchi, S., Omura, N., Campbell, N. R., De Wilde, R. F., … Rajeshkumar, N. V. (2012). The gamma secretase inhibitor MRK-003 attenuates pancreatic cancer growth in preclinical models. *Molecular Cancer Therapeutics*, *11*(9), 1999–2009. https://doi.org/10.1158/1535-7163.MCT-12-0017

Mohammad, A. A. (2018). Advanced pancreatic cancer: the standard of care and new opportunities. *Oncology Reviews*, *12*, 98–104. https://doi.org/10.4081/oncol.2018.370

Morra, L., Rechsteiner, M., Casagrande, S., Duc Luu, V., Santimaria, R., Diener, P. A., … Soltermann, A. (2011). Relevance of periostin splice variants in renal cell carcinoma. *The American Journal of Pathology*, *179*(3), 1513–1521. https://doi.org/10.1016/j.ajpath.2011.05.035

Morra, L., Rechsteiner, M., Casagrande, S., von Teichman, A., Schraml, P., Moch, H., & Soltermann, A. (2012). Characterization of periostin isoform pattern in non-small cell lung cancer. *Lung Cancer*, *76*(2), 183–190. https://doi.org/10.1016/J.LUNGCAN.2011.10.013

MURAKAMI, D., TAKAMORI, S., KAWAHARA, A., MITSUOKA, M., KASHIHARA, M., YOSHIYAMA, K., … AKAGI, Y. (2017). Periostin Expression in Non-Small Cell Lung Cancer: Clinical Significance. *The Kurume Medical Journal*, *64*(1.2), 13–20. https://doi.org/10.2739/kurumemedj.MS640012

Murphy, S. J., Hart, S. N., Halling, G. C., Johnson, S. H., Smadbeck, J. B., Drucker, T., … Vasmatzis, G. (2016). Integrated genomic analysis of pancreatic ductal adenocarcinomas reveals genomic rearrangement events as significant drivers of disease. *Cancer Research*, *76*(3), 749–761. https://doi.org/10.1158/0008-5472.CAN-15-2198

National Cancer Institute (2018) SEER Stat Fact Sheet Pancreatic Cancer. Available at: https://seer.cancer.gov/statfacts/html/pancreas.html (Accessed 27 January 2019).

Neumann, C. C. M., von Hörschelmann, E., Reutzel-Selke, A., Seidel, E., Sauer, I. M., Pratschke, J., … Schmuck, R. B. (2018). Tumor–stromal cross-talk modulating the therapeutic response in pancreatic cancer. *Hepatobiliary and Pancreatic Diseases International*, *17*(5), 461–472. https://doi.org/10.1016/j.hbpd.2018.09.004

Norris, R. A., Damon, B., Mironov, V., Kasyanov, V., Ramamurthi, A., Moreno-Rodriguez, R., … Markwald, R. R. (2007). Periostin regulates collagen fibrillogenesis and the biomechanical properties of connective tissues. *Journal of Cellular Biochemistry*, *101*(3), 695–711. https://doi.org/10.1002/jcb.21224

Oshima, M., Okano, K., Muraki, S., Haba, R., Maeba, T., Suzuki, Y., & Yachida, S. (2013). Immunohistochemically detected expression of 3 major genes (CDKN2A/p16, TP53, and SMAD4/DPC4) strongly predicts survival in patients with resectable pancreatic cancer. *Annals of Surgery*, *258*(2), 336–346. https://doi.org/10.1097/SLA.0b013e3182827a65

Özdemir, B. C., Pentcheva-Hoang, T., Carstens, J. L., Zheng, X., Wu, C.-C., Simpson, T. R., … Kalluri, R. (2014). Depletion of Carcinoma-Associated Fibroblasts and Fibrosis Induces Immunosuppression and Accelerates Pancreas Cancer with Reduced Survival. *Cancer Cell*, *25*(6), 719–734. https://doi.org/10.1016/j.ccr.2014.04.005

Park, J. K., Paik, W. H., Ryu, J. K., Kim, Y. T., Kim, Y. J., Kim, J., … Yoon, Y. B. (2013). Clinical significance and revisiting the meaning of CA 19-9 blood level before and after the treatment of pancreatic ductal adenocarcinoma: Analysis of 1,446 patients from the pancreatic cancer cohort in a single institution. *PLoS ONE*, *8*(11), 1–7. https://doi.org/10.1371/journal.pone.0078977

Pasca Di Magliano, M., Logsdon, C. D., & Lonsdon, C. D. (2013). Roles for KRAS in Pancreatic Tumor Development and Progression. *Gastroenterology*, *144*(6), 1220–1229. https://doi.org/10.1053/j.gastro.2013.01.071

Rahman, M., & Washington, L. (2019). The seemingly innocuous presentation of metastatic pancreatic tail cancer: a case report. *Journal of Medical Case Reports*, *13*(1), 178. https://doi.org/10.1186/s13256-019-2125-5

Rooke, N., Markovtsov, V., Cagavi, E., & Black, D. L. (2003). Roles for SR proteins and hnRNP A1 in the regulation of c-src exon N1. *Molecular and Cellular Biology*, *23*(6), 1874–1884. https://doi.org/10.1128/mcb.23.6.1874-1884.2003

Rosano, G. L., & Ceccarelli, E. A. (2014). Recombinant protein expression in Escherichia coli: Advances and challenges. *Frontiers in Microbiology*. https://doi.org/10.3389/fmicb.2014.00172

Rossi, M. L., Rehman, A. A., & Gondi, C. S. (2014). Therapeutic options for the management of pancreatic cancer. *World Journal of Gastroenterology*, *20*(32), 11142–11159. https://doi.org/10.3748/wjg.v20.i32.11142

Sage, E. H. (2014). Revisiting the matricellular concept. *Matrix Biology*, *37*, 1–14. https://doi.org/10.1016/J.MATBIO.2014.07.005

Saïda, F., Uzan, M., Odaert, B., & Bontems, F. (2006). *Expression of Highly Toxic Genes in E. coli: Special Strategies and Genetic Tools*. *Current Protein and Peptide Science* (Vol. 7).

Sasaki, H., Sato, Y., Kondo, S., Fukai, I., Kiriyama, M., Yamakawa, Y., & Fuji, Y. (2002). Expression of the periostin mRNA level in neuroblastoma. *Journal of Pediatric Surgery*, *37*(9), 1293–1297. https://doi.org/10.1053/jpsu.2002.34985

Sehra, S., Yao, W., Nguyen, E. T., Ahyi, A.-N. N., Barbé Tuana, F. M., Ahlfeld, S. K., … Kaplan, M. H. (2011). Periostin Regulates Goblet Cell Metaplasia in a Model of Allergic Airway Inflammation. *The Journal of Immunology*, *186*(8), 4959–4966. https://doi.org/10.4049/jimmunol.1002359

Shao, R., Bao, S., Bai, X., Blanchette, C., Anderson, R. M., Dang, T., … Wang, X.-F. (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), 3992–4003. https://doi.org/10.1128/mcb.24.9.3992-4003.2004

Shimotohnot, K., Kodamat, Y., Hashimotot, J., & Miurat, K.-I. (1977). *Importance of 5'-terminal blocking structure to stabilize mRNA in eukaryotic protein synthesis* (eukaryotic mRNA/confronting nucleotide structure/tobacco pyrophosphatase)*. *Biochemistry* (Vol. 74). Retrieved from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC431268/pdf/pnas00029-0146.pdf

Siegel, R. L., Miller, K. D., & Jemal, A. (2018). Cancer Statistics in USA , 2018. *CA Cancer J Clin*. https://doi.org/10.3322/caac.21442

Smathers, C. M., & Robart, A. R. (2019). The mechanism of splicing as told by group II introns: Ancestors of the spliceosome. *Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms*. https://doi.org/10.1016/J.BBAGRM.2019.06.001

Smith, C. J., & Osborn, A. M. (2008). Advantages and limitations of quantitative PCR (Q-PCR)-based approaches in microbial ecology. https://doi.org/10.1111/j.1574-6941.2008.00629.x

Soule, H. D., Vazquez, J., Long, A., Albert, S., & Brennan, M. (1973). *A Human Cell Line From a Pleural Effusion Derived From a Breast Carcinoma 1,2*. Retrieved from https://academic.oup.com/jnci/article-abstract/51/5/1409/962551

Suzuki, S., Okada, M., Shibuya, K., Seino, M., Sato, A., Takeda, H., … Kitanaka, C. (2015). JNK suppression of chemotherapeutic agents-induced ROS confers chemoresistance on pancreatic cancer stem cells. *Oncotarget*, *6*(1), 458–470. https://doi.org/10.18632/oncotarget.2693

Takayama, G., Arima, K., Kanaji, T., Toda, S., Tanaka, H., Shoji, S., … Izuhara, K. (2006). Periostin: A novel component of subepithelial fibrosis of bronchial asthma downstream of IL-4 and IL-13 signals. *Journal of Allergy and Clinical Immunology*, *118*(1), 98–104. https://doi.org/10.1016/J.JACI.2006.02.046

Takeshita, S., Kikuno, R., Tezuka, K., & Amann, E. (1993). Osteoblast-specific factor 2: cloning of a putative bone adhesion protein with homology with the insect protein fasciclin I. *The Biochemical Journal*, *294 (Pt 1)*(1), 271–278. https://doi.org/10.1042/BJ2940271

Taucher, V., Mangge, H., & Haybaeck, J. (2016, August 1). Non-coding RNAs in pancreatic cancer: challenges and opportunities for clinical application. *Cellular Oncology*. Springer Netherlands. https://doi.org/10.1007/s13402-016-0275-7

Thierry-Mieg, D., & Thierry-Mieg, J. (2006). AceView: a comprehensive cDNA-supported gene and transcripts annotation. *Genome Biology*, *7*(Suppl 1), S12. https://doi.org/10.1186/gb-2006-7-s1-s12

Tian, Y., Choi, C. H., Li, Q. K., Rahmatpanah, F. B., Chen, X., Kim, S. R., … Zhang, H. (2015). Overexpression of periostin in stroma positively associated with aggressive prostate cancer. *PloS One*, *10*(3), e0121502. https://doi.org/10.1371/journal.pone.0121502

Tingle, S. J., Moir, J. A., & 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), 235–242. https://doi.org/10.4291/wjgp.v6.i4.235

Topisirovic, I., Svitkin, Y. V., Sonenberg, N., & Shatkin, A. J. (2011). Cap and cap-binding proteins in the control of gene expression. *Wiley Interdisciplinary Reviews: RNA*, *2*(2), 277–298. https://doi.org/10.1002/wrna.52

Utispan, K., Thuwajit, P., Abiko, Y., Charngkaew, K., Paupairoj, A., Chau-in, S., & Thuwajit, C. (2010). Gene expression profiling of cholangiocarcinoma-derived fibroblast reveals alterations related to tumor progression and indicates periostin as a poor prognostic marker. *Molecular Cancer*, *9*, 13. https://doi.org/10.1186/1476-4598-9-13

Vasseur, Cens, Wagner, Saint, Kugler, Chavanieu, … Boutin. (2019). Importance of the Choice of a Recombinant System to Produce Large Amounts of Functional Membrane Protein hERG. *International Journal of Molecular Sciences*. https://doi.org/10.3390/ijms20133181

Viloria, K., & Hill, N. J. (2016). Embracing the complexity of matricellular proteins: The functional and clinical significance of splice variation. *Biomolecular Concepts*, *7*(2), 117–132. https://doi.org/10.1515/bmc-2016-0004

Von Hoff, D. D., Ervin, T., Arena, F. P., Chiorean, E. G., Infante, J., Moore, M., … Renschler, M. F. (2013). Increased survival in pancreatic cancer with nab-paclitaxel plus gemcitabine. *New England Journal of Medicine*, *369*(18), 1691–1703. https://doi.org/10.1056/NEJMoa1304369

Waddell, N., Pajic, M., Patch, A. M., Chang, D. K., Kassahn, K. S., Bailey, P., … Grimmond, S. M. (2015). Whole genomes redefine the mutational landscape of pancreatic cancer. *Nature*, *518*(7540), 495–501. https://doi.org/10.1038/nature14169

Wahle, E., & Rüegsegger, U. (1999). 3'-End processing of pre-mRNA in eukaryotes. *FEMS Microbiology Reviews*, *23*(3), 277–295. https://doi.org/10.1111/j.1574-6976.1999.tb00400.x

Wallace, D. P., White, C., Savinkova, L., Nivens, E., Reif, G. A., Pinto, C. S., … Fields, T. A. (2014). Periostin promotes renal cyst growth and interstitial fibrosis in polycystic kidney disease. *Kidney Int*, *85*(4), 845–854. https://doi.org/10.1038/ki.2013.488

Walsh, J. S., Gossiel, F., Scott, J. R., Paggiosi, M. A., & Eastell, R. (2017). Effect of age and gender on serum periostin: Relationship to cortical measures, bone turnover and hormones. *Bone*, *99*, 8–13. https://doi.org/10.1016/j.bone.2017.03.041

Wang, Y., Liu, J., Huang, B. O., Xu, Y.-M., Li, J., Huang, L.-F., … Wang, X.-Z. (2015). Mechanism of alternative splicing and its regulation. *Biomedical Reports*, *3*(2), 152–158. https://doi.org/10.3892/br.2014.407

Wang, Z., & 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), 802–813. https://doi.org/10.1261/rna.876308

Xie, D., & Xie, K. (2015). Pancreatic cancer stromal biology and therapy. *Genes and Diseases*. https://doi.org/10.1016/j.gendis.2015.01.002

Ye, J., Coulouris, G., Zaretskaya, I., Cutcutache, I., Rozen, S., & Madden, T. L. (2012). Primer-BLAST: a tool to design target-specific primers for polymerase chain reaction. *BMC Bioinformatics*, *13*, 134. https://doi.org/10.1186/1471-2105-13-134

Yeo, G., Holste, D., Kreiman, G., & Burge, C. B. (2004). Variation in alternative splicing across human tissues. *Genome Biology*, *5*(10), R74. https://doi.org/10.1186/gb-2004-5-10-r74

Yuan, Y., Jiang, J., Wang, J., Sun, J., Li, C., Liu, B., … Wang, H. (2019). BAG3-positive pancreatic stellate cells promote migration and invasion of pancreatic ductal adenocarcinoma. *Journal of Cellular and Molecular Medicine*, jcmm.14352. https://doi.org/10.1111/jcmm.14352

Zhao, J., Hyman, L., & Moore, C. (1999). Formation of mRNA 3' ends in eukaryotes: mechanism, regulation, and interrelationships with other steps in mRNA synthesis. *Microbiology and Molecular Biology*

Reviews : MMBR, *63*(2), 405–445. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/10357856

Zhong, H., Li, X., Zhang, J., & Wu, X. (2019). Overexpression of periostin is positively associated with gastric cancer metastasis through promoting tumor metastasis and invasion. *Journal of Cellular Biochemistry*, *120*(6), 9927–9935. https://doi.org/10.1002/jcb.28275

Zhong, Y., Wang, Z., Fu, B., Pan, F., Yachida, S., Dhara, M., … Iacobuzio-Donahue, C. A. (2011). GATA6 Activates Wnt Signaling in Pancreatic Cancer by Negatively Regulating the Wnt Antagonist Dickkopf-1. *PLoS ONE*, *6*(7), e22129. https://doi.org/10.1371/journal.pone.0022129

Appendix

Composite sequence (PS-1 Lane 1 bright band)

TCCAGCAGACACACCTGTTGGAAATGATCAACTGCTGGAAATACTTAATAAATTAATCAAATACATCCAAATTA
AGTTTGTTCGTGGTAGCACCTTCAAAGAAATCCCCGTGACTGTCTATAGACCCACACTAACAAAAGTCAAAATT
GAAGGTGAACCTGAATTCAGACTGATTAAAGAAGGTGAAACAATAACTGAAGTGATCCATGGAGAGCCAATT
ATTAAAAAATACACCAAAATCATTGATGGAGTGCCTGTGGAAATAACTGAAAAAGAGACACGAGAAGAACGA
ATCATTACAGGTCCTGAAATAAAATACACTAGGATTTCTACTGGAGGTGGAGAAACAGAAGAAACTCTGAAR
AAATTGTTACRAGAAGACACACCC

Exons: 14 15 16 18 19 20 22

Figure A1: Composite sequence of POSTN-201a variant found in Lane 1 of Figure 3.6. The boxed sequences show the forward and reverse primers.

Composite sequence (PS-1 Lane 1 dim band)

TCCAGCAGACACACCTGTTGGAAATGATCAACTGCTGGAAATACTTAATAAATTAATC
AAATACATCCAAATTAAGTTTGTTCGTGGTAGCACCTTCAAAGAAATCCCCGTGACTGTC
TATAAGCCAATTATTAAAAAATACACCAAAATCATTGATGGAGTGCCTGTGGAAATAACT
GAAAAAGAGACACGAGAAGAACGAATCATTACAGGTCCTGAAATAAAATACACTAGGATT
TCTACTGGAGGTGGAGAAACAGAAGAAACTCTGAAGAAATTGTTACAAGAAGACACACCC

Exons: 14 15 16 19 20 22

Figure A2 : Composite sequence of POSTN-201b variant found in Lane 1 of Figure 3.6. The boxed sequences show the forward and reverse primers.

Composite sequence (PS-1 Lane 2 bright band)

TCCAGCAGACACACCTGTTGGAAATGATCAACTGCTGGAAATACTTAATAAATTAATC
AAATACATCCAAATTAAGTTTGTTCGTGGTAGCACCTTCAAAGAAATCCCCGTGACTGT
TATAGACCCACACTAACAAAAGTCAAAATTGAAGGTGAACGAATTCAGACTGATTAAA
GAAGGTGAAACAATAATGAAGTGATCCATGGAGAGCCAATTATTAAAAAATACACCAAA
ATCATTGATGGAGTGCCTGTGGAAATAACTGAAAAAGAGACACGAGAAGAACGAATCATT
ACAGGTCCTGAAATAAAATACACTAGGATTTCTACTGGAGGTGGAGAAACAGAAGAAACT
CTGAAGAAATTGTTACAAGAAGAGGTCACCAAGGTCACCAAATTCATTGAAGGTGGTGAT
GGTCATTTATTTGAAGATGAAGAAATTAAAAGACTGCTTCAGGGAGACACAC

Exons: 14 15 16 18 19 20 21 22

Figure A3: Composite sequence of POSTN-003 variant found in Lane 2 of Figure 3.6. The boxed sequences show the forward and reverse primers.

Composite sequence (PS-1 Lane 2 dim band)

TCCAGCAGACACACCTGTTGGAAATGATCAACTGCTGGAAATACTTAATAAATTAATC
AAATACATCCAAATTAAGTTTGTTCGTGGTAGCACCTTCAAAGAAATCCCCGTGACTGTT
ATAAGCCAATTATTAAAAAATACACCAAAATCATTGATGGAGTGCCTGTGGAAATAACTG
AAAAAGAGACACGAGAAGAACGAATCATTACAGGTCCTGAAATAAAATACACT
AGGATTTCTACTGGAGGTGGAGAAACAGAAGAAACTCTGAAGAAATTGTTACAAGAAGAG
GTCACCAAGGTCACCAAATTCATTGAAGGTGGTGATG

Exons: 14 15 16 19 20 21

Figure A4: Composite sequence of POSTN-002 variant found in Lane 2 of Figure 3.6. The boxed sequence show the forward primer.

Composite sequence (PS1 Lane 3 bright band)

GATTGAAGGCAGTCTTCAGCCTATTATCAAAACTGAAGGACCCACACTAACAAAAGTCA
AAATTGAAGGTGAACCTGAATTCAGACTGATTAAAGAAGGTGAAACAATAACTGAAGTGA
TCCATGGAGAGCCAATTATTAAAAAATACACCAAAATCATTGATGGAGTGCCTGTGGAAA
TAACTGAAAAAGAGACACGAGAAGAACGAATCATTACAGGTCCTGAAATAAAATACACTA
GGATTTCTACTGGAGGTGGAGAAACAGAAGAAACTCTGAAGAAATTGTTACAAGAAGACA
CACCC

Exons: 17 18 19 20 22

Figure A5: Composite sequence of POSTN-004 variant found in Lane 3 of Figure 3.6. The boxed sequences show the forward and reverse primers.

Composite sequence (PS1 bright band Lane 4)

GATTGAAGGCAGTCTTCAGCCTATTATCAAAACTGAAGGACCCACACTAACAAAAGTCA
AAATTGAAGGTGAACCTGAATTCAGACTGATTAAAGAAGGTGAAACAATAACTGAAGTGA
TCCATGGAGAGCCAATTATTAAAAAATACACCAAAATCATTGATGGAGTGCCTGTGGAAA
TAACTGAAAAAGAGACACGAGAAGAACGAATCATTACAGGTCCTGAAATAAAATACACTA
GGATTTCTACTGGAGGTGGAGAAACAGAAGAAACTCTGAAGAAATTGTTACAAGAAGAGG
TCACCAAGGTCACCAAATTCATTGAAGGTGGTGATGGTCATTTATTTGAAGATGAAGAAA
TTAAAAGACTGCTTCAGGGAGACACAC

Exons: 17 18 19 20 21 22

Figure A6: Composite sequence of POSTN-001 variant found in Lane 4 of Figure 3.6. The boxed sequences show the forward and reverse primers.

Forward sequence

taaatTaatcaAATACATCCaAATTAAgTTTGTTCGTGGTAGCACCTTCaAAGAAATCCCCGTGACTGTCTATAgAC
CCACNCTNACAAAA

Exons: 15 16 18

Reverse sequence

CCAATTATTAAAAAATACACCAAAATCATTGATGGAGTGCNTGTGGAAATAACTGAAAAAGAGACACGAGAA
GAACGAATCATtaCAGGTCNNGAAATAAAATaCaCTAGGATTTNTaCTGGAGGTGGAGAAACAGAAGAAa

Exons: 19 20

Figure B1: Forward and reverse sequences of the top band marked in a white arrow in Figure 3.9a. Composite sequences could not be formed. The forward and reverse sequences were blasted (Nucleotide BLAST) and matched with transcripts POSTN-003 (NM_001286665) and POSTN-201a (NM_001135935).

TTAATcAaATACaTCCaAATTAAgTTTGtTCGTGGTAGCACCTTCAAAGAAAtCcccgtGACTGTCTATAAGCCAATTATTAAAAAATACACCA
AAATCNTTGATGGAGtGCCTGtGGAaNTNACTGAAAAAGAGACACNANAAGAACGAATCATTACNGGTCCTGAAATAAaATACACTAN
GATTTCTACT

Exons: 15 16 19 20

Figure B2: Forward sequence only of the bottom band marked in a white arrow in Figure 3.9a. Composite sequences could not be formed as there was poor sequencing in the reverse sequence. The forward sequence was blasted (Nucleotide BLAST) and matched with transcripts POSTN-201b (NM_001135936) and POSTN-002 (NM_001135934).

TaAATTAATcaAATACaTCCaAATTAAGTTTGTTCGTGGTAGCACCTTCa
aAGAAATCCCCGTGACTGTCTATAgACCCACNCTAACAAAAgtCAaAA
TTGANNGTGAACCtGANTTCNNACTGATTAANGAAgGTGAAACANT
AaCTGAAgtGATCCNTGgagAgCCNNTTATTAAAAAaTACACCANNNt
CATTGATGGAGTGcCTGtGGAAaNAACNGaAAAAGAGACACGAGAA
GAACGAATCATTACAGGTCCTGAAATAAAATACACTAGGATTTCTACT
GGAGGTGGAGAAACAGAAGAAAC

Exons: 15 16 18 19 20

Figure C1: Forward sequence of the first band in Figure 3.9b. Composite sequences could not be formed as there was poor sequencing in the reverse sequence. The forward sequence was blasted (Nucleotide BLAST) and matched with transcripts POSTN-003 (NM_001286665) and POSTN-201a (NM_001135935).

aTTAATCaAATACATCCaAATTAAGTTTGTTCGTGGTAGCACCTTCAAAGAAATCCCCGTGACTGTCTATAGACC
CACACTAACAAAAGTCAAAATTGAAGGTGAACCTGAATTCAGACTGATTAAAGAAGGTGAAACAATAACTGA
AGTGATCCATGGAGAGCCAATTATTAAAAAATACACCAAAATCATTGATGGAGTGCCTGTGGAAATAACTGAA
AAAGAGACACGAGAAGAACGAATCATTACAGGTCCTGAAATAAAATACACTAGGATTTCTACTGGAGGTGGA
G

Exons: 15 16 18 19 20

Figure C2: Forward sequence of the second band in Figure 3.9b. Composite sequences could not be formed as there was poor sequencing in the reverse sequence. The forward sequence was blasted (Nucleotide BLAST) and matched with transcripts POSTN-003 (NM_001286665) and POSTN-201a (NM_001135935).

aaTTAATCAaATACATCCaAATTAAGTTTGTTCGTGGTAGCACCTTCAAAGAAATCCC
CGTGACTGTCTATAAGCCAATTATTAAAAAATACACCAAAATCATTGATGGAGTGC
CTGTGGAAATAACTGAAAAAGAGACACGAGAAGAACGAATCATTACAGGTCCTGA
AATAAAATACACTAGGATTTCTACTGGAGGTGGAGAAACAGAAGAAAC

Exons: 15 16 19 20

Figure C3: Forward sequence of the third band in Figure 3.9b. Composite sequences could not be formed as there was poor sequencing in the reverse sequence. The forward sequence was blasted (Nucleotide BLAST) and matched with POSTN-201b (NM_001135936) and POSTN-002 (NM_001135934).

acTGGAGGTGGAGAAACAGAAGAAACTCTGAAGAAATTGTTACAAGAAGAGGTCACCA

AGGTCACCAAATTCATTGAAGGNGGTGATGGTCATTTATTTGAAGATGAAGAAATTAAAA

GACTGCTTCAGGGAGACACACCCGTGAGGAAGTTGCAAGCCAACAAAAAAGTTCAAGGAT

CTAGAAGACGATTAAGGGAAGGTCGTTCTCAGTG

Exons: 20 21 22 23

Figure D1: Composite sequence of the top band from Figure 3.9c. The boxed sequences are forward and reverse primers. The sequence was blasted (Nucleotide BLAST) and matched with POSTN-001 transcript (NM_006475), POSTN-003 transcript (NM_001286665), POSTN-202a transcript (NM_001286666) and POSTN-002 transcript (NM_001135934).

ACTGGAGGTGGAGAAACAGAAGAAACTCTGAAGAAATTGTTACAAGAAGacaCA
CCCGTGAGGAAGTTGCAAGCCAACAAAAAAGTTCAAGGATCTAGAAGACGATT
AAGGGAAGGTCGTTCTCagtg

Exons: 20 22 23

Figure D2: Composite sequence of the bottom band from figure 3.9c. The boxed sequences are forward and reverse primers. The sequence was blasted (Nucleotide BLAST) and matched with 004 transcript (NM_001330517), 202b transcript (NM_001286667), 201b transcript (NM_001135936) & 201a transcript (NM_001135935).