1

Using CRISPR/Cas9 to study SPOCK1 in pancreatic cancer

Kingston University London

By Maria GEORGIADI

First supervisor: Dr Terry Gaymes

Other supervisors: Dr Natasha Hill and Dr Karen Whiting

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

Faculty of Science, Engineering and Computing

June 2020

Table of Contents

Declaration

This thesis has been submitted for a Masters by Research degree at Kingston University. I declare that this thesis contains my original research and any contribution by other individuals has been fully acknowledged.

Maria Georgiadi

Acknowledgements

Abstract

Pancreatic ductal adenocarcinoma (PDAC) is among the most aggressive tumours, with a devastating 5-year survival rate of less than 9%. This is in part the result of the development of an abundant desmoplastic stroma which surrounds, protects, and actively promotes a tumour conducive environment. The stroma is largely composed of extracellular matrix (ECM) proteins, cells such as fibroblasts and stellate cells that are activated in response to the tumour and soluble proteins. Among these is a group of non-structural proteins that play a central role in the mediating interactions between cells and the ECM. SPOCK1 is a member of the Secreted Protein Acidic and Rich in Cysteine (SPARC) family of matricellular proteins. In various tumours, several oncogenic roles have been described for SPOCK1 such as promoting invasiveness and metastasis. Clinical samples correlate SPOCK1 expression with advanced PDAC tumours and poor prognosis. However, very little is currently known on the mechanisms of action in PDAC but interactions with matrix metalloproteinases (MMP) and growth factors, and activation of the Phosphatidylinositol 3-kinase (PI3K)/Akt pathway are suspected.

This research project aimed to understand the role of *SPOCK1* in stromal and pancreatic cancer cell growth and adhesion. Clustered regularly interspaced short palindromic repeats/CRISPR associated protein 9 (CRISPR/Cas9) gene editing technique was used to attempt to knockdown (KD) *SPOCK1*. However, while a preliminary T7 endonuclease 1 (T7E1) assay indicated the presence of a mutation, Clustal omega analysis of sequencing of the CRISPR/Cas9 transfected cell lines failed to show a mutation in the *SPOCK1* region. Despite the lack of mutation in the *SPOCK1* target region, functional assays showed effects on both cell growth and adhesion suggesting off-target binding of Cas9 to the single guide RNA (sgRNA). Several off-target gene were identified with sgRNA sequence similarity to *SPOCK1.* Further experiments searching for interactions between SPOCK1 protein and ECM components revealed fibronectin, fibroblast growth factor, collagen, and membrane type 1 matrix metalloproteinase as direct binding partners of SPOCK1, suggesting that the SPOCK1 protein has diverse roles in the PDAC ECM.

Key words: PDAC, ECM, matricellular proteins, CRISPR/Cas9, SPOCK1, MMP

List of Figures

List of Tables

List of Abbreviations

AF488	Alexafluor488
ANOVA	Analysis of variance
α-SMA	Alpha-smooth muscle actin
BLESS	Breaks Labelling, Enrichments on Streptavidin and next generation Sequencing
BMPB	Bone morphogenetic protein 4
BPAE	Bovine pulmonary artery endothelial cells
BRCA1/-2	Breast cancer type 1/-2
BSA	Bovine serum albumin
CDKN2A	Cyclin dependent kinase inhibitor 2A
CA-19	Cancer antigen 19
Cas9	CRISPR associated protein 9
CCL2	Chemokines C-C motif chemokine ligand 2
CCN	Connective tissue growth factor and Nov family
CD8	Cluster of differentiation 8
COLL	Collagen
COL1A1	Collagen type 1 alpha 1
CRISPR	Clustered regularly interspaced short palindromic repeats
crRNA	CRISPR RNA
CSF1	Colony stimulating factor 1
CSF1R	Colony stimulating factor 1 receptor
dH_20	Distilled water
DAPI	4',6-diamidino-2-phenylindole
DNA	Deoxyribonucleic acid
DMD	Duchenne muscular dystrophy
DMSO	Dimethyl sulfoxide
DSB	Double strand break
EC	Extracellular
ECM	Extracellular matrix
EDTA	Ethylenediaminetetraacetic acid
EMEM	Eagle's minimal essential medium
EMT	Epithelial-mesenchymal transition
FAK	Focal adhesion kinase
FAMMM	Familial atypical multiple mole melanoma
FAP	Fibroblast associated protein
FBS	Foetal bovine serum
FDA	Food and drug administration
FGF	Fibroblast growth factor

FN	Fibronectin
FOLFIRINOX	Fluorouracil, leucovorin, irinotecan and oxaliplatin
FSTL-1	Follistatin like 1
gDNA	Genomic DNA
H_2SO_4	Sulphuric acid
HDR	Homology directed repair
HER2	Human epidermal growth factor receptor 2
HIF	Hypoxia-inducible factor
HIV	Human immunodeficiency virus
HRP	Horse radish peroxidase
ICAM-1	Intercellular adhesion molecule 1
IPMN	Intraductal papillary mucinous neoplasm
KRAS	Kirsten rat sarcoma viral oncogene
MEF	Mouse embryonic fibroblasts
MMP	Matrix metalloproteinase
MT1-MMP	Membrane type 1 matrix metalloproteinase
MT3-MMP	Membrane type 3 matrix metalloproteinase
MCN	Mucinous cystic neoplasm
MCP1	Monocyte chemotactic protein 1
NET	Neuroendocrine tumour
NF-κB	Nuclear factor-kappa beta
NGF-β	Nerve growth factor -beta
NHEJ	Non-homologous end joining
NO	Nitric oxide
nt	Nucleotide
OS	Overall survival
KD	Knockdown
P-value	Probability value
PAI-1	Plasminogen activator inhibitor type 1
PALB2	Partner and localizer of BRCA2
PAM	Protospacer adjacent motif
PanIN	Pancreatic intraepithelial neoplasia
PARP	Poly adenosine diphosphate-ribose polymerase
PBS	Phosphate buffered saline
PCNA1	Proliferating cell nuclear antigen
PCR	Polymerase chain reaction
PDAC	Pancreatic ductal adenocarcinoma
PEDF	Pigment epithelium derived factor
PFS	Progression-free survival
PI3K	Phosphatidylinositol 3-kinase
POSTN	Periostin

PS-1	Pancreatic stellate cell line
RIPA	Radioimmunoprecipitation assay buffer
RNA	Ribonucleic acid
RNP	Ribonucleoprotein
RPMI-1640	Roswell Park Memorial Institute-1640 medium
RT	Room temperature
SC	Single colony
sgRNA	Single guide RNA
SLRPS	Small leucine-rich proteoglycans
SMOC	Secreted modular calcium-binding protein
SNPs	Single nucleotide polymorphisms
SPARC	Secreted protein acidic and rich in cysteine
STAT3	Signal transducer and activator of transcription 3
T7E1	T7 endonuclease I
TAMs	Tumour associated macrophages
TBS	Tris buffered saline
TGF-β	Transforming growth factor- beta
TH-302	Evofosfamide
TIC	Tumour initiating cell
TIMP	Tissue inhibitors of metalloproteinases
TMB	3,3',5,5'-tetramethylbenzidine
TME	Tumour microenvironment
TN-C	Tenascin-C
TNF-α	Tumour necrosis factor - alpha
TP53	Tumour protein 53
TPZ	Tirapazamine
tracrRNA	Trans activating RNA
TSP-1	Thrombospondin-1
UK	United Kingdom
UPR	Unfolded protein response
VCAM-1	Vascular cell adhesion molecule 1
VEGF	Vascular endothelial growth factor
VEGFR	Vascular endothelial growth factor receptor

1 Introduction

1.1 Pancreatic cancer

1.1.1 Pancreatic ductal adenocarcinoma: Epidemiology and Aetiology

The predominant form of pancreatic cancer, pancreatic adenocarcinomas, arises in the exocrine glands of the pancreas (85%) while the rarer pancreatic neuroendocrine tumours occur in the endocrine tissue (Rawala, Sunkara and Guduputi, 2019). Clinically, pancreatic ductal adenocarcinoma (PDAC) present as a firm off-white mass encapsulated by non-malignant atrophic fibrotic tissue with dilated ducts. Most PDACs arise from three common precursor lesions; pancreatic intraepithelial neoplasia [PanIN], intraductal papillary mucinous neoplasm [IPMN] and mucinous cystic neoplasm [MCN] (Ying *et al.*, 2016). Unfortunately, patients with PDAC are often asymptomatic or present non-specific symptoms until advanced stages of the disease therefore only 15-20% of patients are operable at time of diagnosis (Bekkali and Oppong, 2017). Due to its aggressive nature and late diagnosis, pancreatic cancer is the seventh leading cause of cancer related deaths globally with 5-year survival rate of 9% (Sung *et al.*, 2021; Rawala, Sunkara and Guduputi, 2019).

There are several well-established risk factors for pancreatic cancer. These include age, tobacco smoke, high alcohol consumption, chronic pancreatitis, high BMI, poor oral health (association with periodontal disease, tooth loss and microbiota), physical inactivity and certain inherited predispositions and genetic mutations (*BRCA2*, *BRCA1*, *PALB2*, *FAMMM* and *CDKN2A*) (Simoes *et al.*, 2017). The most frequent genetic alterations/mutations in PDAC cells occur in the proto-oncogene *KRAS* (95%) and tumour suppressor genes *CDKN2A* (95%), *TP53* (75%), and *SMAD4* (55%) (Gore and Korc, 2014).

Current first line treatment such as cytidine analogues, FOLFIRINOX or gemcitabine/nab-paclitaxel show limited success against PDAC due to chemoresistance and poor understanding of underlying disease mechanism (Orth *et al*., 2019).

1.1.2 The tumour stroma

The treatment of PDAC is complicated by the development of a dense fibrotic stroma (microenvironment) which obstructs the delivery of chemotherapy to the tumour at the core. Despite initial assumptions that the surrounding stroma is a physical barrier, much research since also shown the active role the stroma plays in the development and progression of PDAC (Gore and Korc, 2014; Provenzano *et al.,* 2012).

The stroma comprises of a complex network of structural extracellular matrix (ECM) proteins, soluble factors, cells and matricellular proteins (Fig. 1.1). The ECM is a cross-linked 3D scaffolding of fibrillar proteins (e.g., collagens and laminins), proteoglycans and glycoproteins that bind cell surface receptors to mediate biochemical and biophysical signalling to direct cell fate (Walma and Yamada, 2020).
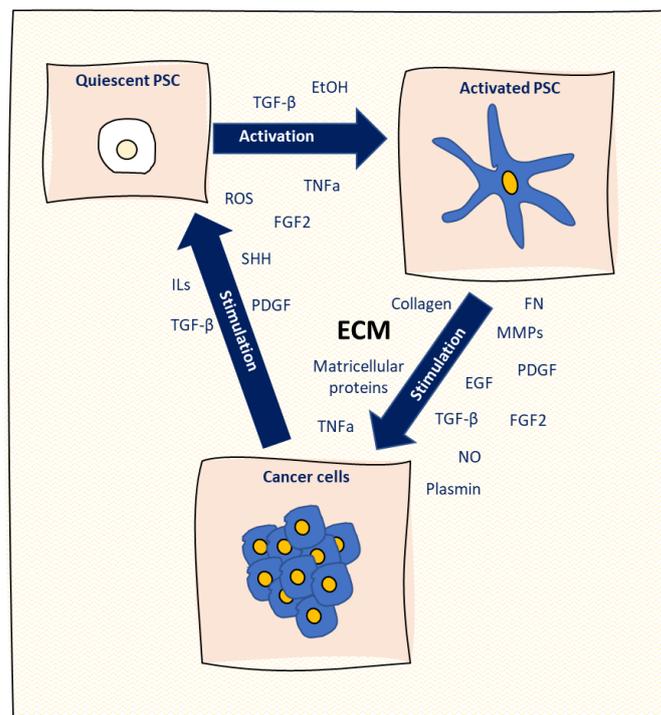
Figure 1.1: The tumour and its stroma. Pancreatic stellate cells (PSC) are activated in response to the rise of cancerous cells. This new myofibroblast-like phenotype is accompanied by a significant increase in secretion of soluble factors such as growth factors, precursor ECM proteins and matricellular proteins which in turn promote cancer proliferation, invasion, metastasis, and chemoresistance. Image drawn from information acquired from Erkan *et al*. 2012. EGF: Epidermal growth factor, EtOH: ethanol, FGF2: fibroblast growth factor-2, ILs: Interleukins, MMP: matrix metalloproteases, NO: nitric oxide, ROS: reactive oxygen species, SHH: sonic hedgehog, TGF-β: transforming growth factor beta, TNF: tumour necrosis factor.

Due to the restricted perfusion in PDAC tumours, PDAC cells (90% present with KRAS mutation) use membrane shuffling to aid micropinocytosis to acquire essential nutrients from their environment. PDAC cells can metabolise collagen fragments under glucose deprived conditions as well as proliferate in the absence of essential amino acids in albumin supplemented medium (Weniger *et al*., 2018).

Matrix metalloproteinases (MMPs), tissue inhibitors of metalloproteinases (TIMPs) and other enzymes work together to ensure matrix turnover and remodelling. In PDAC elevated amounts of MMPs can compromise basement membrane integrity to aid invasion and metastasis (Venkatasubramanian, 2012). These proteins are secreted by a variety of cells that are activated in response to the presence of malignant cells. Cancer associated fibroblasts, pancreatic stellate cells and some activated tumour cells all contribute to a tumour conducive stroma by secretion of growth factors (e.g., FGF, TGF-β, VEGF) that bind to the ECM and help promote cell proliferation, growth, invasiveness, chemoresistance and metastasis. Unsurprisingly, this has led to research into understanding and targeting the tumour stroma for enhanced cancer therapy (Lai *et al*., 2020; Valkenburg, de Groot and Pienta, 2018; Provenzano *et al.,* 2012).

There are four primary therapeutic strategies employed in targeting the tumour microenvironment (TME) for cancer therapy; 1) target the tumour vasculature, 2) target cancer associated inflammation, 3) target the hypoxia in the TME and 4) target communication between the tumour and TME (Fang and DeClerck, 2013).

Targeting the tumour vasculature

Vasoendothelial growth factor *(VEGF)* is the predominant soluble driver of angiogenesis. This makes *VEGF* an attractive target for therapy. Sunitinib, sorafenib, and pazopanib are some of the small molecule inhibitors of *VEGF/VEGFR* signalling molecules that have been approved by the Food and Drug Administration (FDA) for the treatment of advanced renal cell carcinoma (Eichelberg *et al*., 2015; Sternberg *et al*., 2013). Bevacizumab, a monoclonal antibody against VEGF, has been successfully employed for the treatment of several solid tumours such as colon cancer and hepatocellular carcinoma (Garcia *et al*., 2020). However, in 2011, bevacizumab was revoked by the FDA as a treatment option for patients with metastatic HER2-negative breast cancer due to significant adverse effects

such as haemorrhaging and heart failure, without the benefit of improving overall survival (Kümler *et al.*, 2014; Choueiri *et al.*, 2011; Hapani *et al.*, 2010). This highlights the need to study organ specific microenvironments. PDAC tumours are hypoxic and hypovascularized with high microvascular density and low microvascular integrity and can mimic vasculogenesis. Thus, unsurprisingly, chemoradiation therapy and chemotherapy with bevacizumab did not improve survival outcome in patients with locally advanced pancreatic cancer in a phase II clinical trial (Crane *et al.*, 2009). Anti-angiogenic therapies may still have potential in the less frequent pancreatic neuroendocrine tumours (NET) as these are highly vascularised (Pozas *et al.,* 2019; Wang *et al.*, 2018).

Targeting the stromal immune cells

Inhibiting immune stromal cells is another method to target the TME. Polarization of macrophages to a pro-tumourigenic phenotype, also known as M2-type tumour associated macrophages [TAMs]), by chemokines C-C motif chemokine ligand 2 (*CCL2*), monocyte chemotactic protein 1 (*MCP1*) and colony stimulating factor 1 (*CSF1*) is key in oncogenesis (Fang and DeClerk, 2013). Tumour educated macrophages and inflammatory monocytes have been demonstrated to induce tumour initiating cell (TIC) properties in PDAC cells by activating *STAT3* – this crosstalk initiates TAM-mediated TIC suppression of cytotoxic T cell activity. Inhibiting *CCL2* and *CSF1R* signalling improved chemotherapeutic response to gemcitabine, inhibited metastasis, improved the antitumour response of CD8$^+$ T lymphocytes and decreased TICs in PDAC tumours (Mitchem *et al.*, 2013).

Targeting the TME hypoxia

Like many TMEs, the PDAC stroma is severely hypoxic. Tumour cells adapt and utilize this environment to promote genomic instability and alter their metabolism which ultimately drives proliferation, metastasis, invasion, vasculogenesis, chemoresistance, radioresistance and resistance to cell death (Riffle and Hegde, 2017). A way to use hypoxia for therapy is to use bioreductive prodrugs which are cytotoxins activated by enzymatic reduction under hypoxic conditions. Tirapazamine (TPZ) is the most well-known bioreductive prodrug; its hypoxic tumour cells TPZ is reduced to free radical which causes single and double strand breaks in DNA. Despite its promising results in early studies, several phase III clinical trials have shown limited survival benefit of combination therapy

with TPZ in non-small cell lung cancer, head, and neck cancer. A phase II clinical trial with bioreductive prodrug TH-302 and gemcitabine in patients previously treated, locally advanced or metastatic pancreatic cancer showed significant improvement in progression-free survival (PFS), tumour response and CA-19 response compared with gemcitabine alone (Borad *et al.,* 2015). A randomized, double-blind, placebo-controlled phase III study (ClinicalTrials.gov Identifier: NCT01746979) in patients with locally advanced unresectable or metastatic PDAC was completed in 2017. Median overall survival (OS) and PFS showed modest improvement with TH-302 when combined with gemcitabine compared to placebo/gemcitabine treatment (OS: 8.7 vs. 7.6 months and PFS: 3.7 vs. 5.5 months). Unfortunately, grade 3 hematologic adverse effects were also more frequent in the TH-302/gemcitabine arm compared to the placebo/gemcitabine group.

A second strategy to tackle hypoxia in the TME is to inhibit molecules and pathways key to the survival of hypoxic cell such as hypoxia-inducible factor (HIF), the unfolded protein response (UPR) and mTOR pathway (Wilson and Hay, 2011). Jung *et al.,* (2020) reported elevated expression of *HIF-1α/-2α* in *PanIN, IPMN*, neuroendocrine tumour (NET) and PDAC tissue when compared with normal pancreatic tissue. Currently, the most studied HIF targets are ENZ-2968, a HIF-1α antisense mRNA, and PX478, a small molecule inhibitor of HIF (Fang and DeClerck, 2013). Pilot studies and a phase I clinical trial using EZN-2968 in patients with advanced solid tumours or lymphoma have been setup, with one of the formers indicating potential for further study for clinical use (Jeong *et al.*, 2014; ClinicalTrials.gov Identifiers: NCT00466583 and NCT01120288). An *in vitro* and mouse model study showed that combined treatment of gemcitabine with PX-478 inhibited PDAC tumour growth, potentially due to initiation of gemcitabine induced immunogenic cell death post HIF-1α inactivation (Zhao *et al.,* 2019). Importantly, HIF-1α has also been reported to promote glucose homeostasis β by improving β-cell function in db/db and streptozotocin-induced mouse models of diabetes (Ilegems *et al.,* 2022). Clinical trials using combination therapy with PX-478 in patients diagnosed with PDAC may prove to be promising.

Targeting the tumour stroma communication

Communication between the tumour cell and its microenvironment can occur with direct contact or indirectly via soluble factors.

Integrins are a family of cell surface receptors that interact with the ECM and cell-cell communication and play important role in a variety of biological processes. Aggregation of integrins with the ECM and focal adhesion kinase (FAK) can mediate cell growth, differentiation, adhesion, motility, apoptosis, and angiogenesis (Van der Flier and Sonnenberg, 2001). In PDAC, integrins β5, αvβ6 and α6β4 have been reported to promote growth and β1 and β8 to promote gemcitabine resistance and radiochemoresistance respectively (Humphries *et al.*, 2022; Hurtado de Mendoza *et al.*, 2021; Jin *et al.,* 2019; Reader *et al.*, 2019; Yan *et al.*, 2018;). A randomised phase II clinical trial using gemcitabine and cilengitide, a potent inhibitor of avβ3 and avβ5, in patients with advanced unresectable pancreatic cancer showed no positive impact on OS (6.7 vs 7.7 months), efficacy or safety compared to gemcitabine treatment alone (Friess *et al.*, 2006). Several phase I and phase II clinical trials using FAK inhibitors for patients with pancreatic cancer are currently in progress (ClinicalTrials.gov Identifiers: NCT03727880, NCT02546531, NCT02428270).

A major group of soluble factors responsible for communication in PDAC tumours are matricellular proteins. These will be discussed in more detail in Section 1.2.

1.2 Matricellular proteins

Among the many components that compromise the tumour stroma, there is rather unique group of non-structural extracellular proteins that are collectively referred to as matricellular proteins. The primary role of matricellular proteins lie in the regulation of cell-cell and cell-matrix communications by interaction with cell surface receptors, hormones, proteases, the structural matrix proteins, and other bio-effector molecules (Murphy-Ullrich and Sage, 2014; Bornstein, 2009). Consequently, the extensive binding/interaction capabilities of these proteins allow for regulation of a multitude of biological processes such as cell adhesion, proliferation, invasion, migration, survival, and angiogenesis (Wong and Rustgi, 2013). Classical members of this group of proteins are SPARC, thrombospondin-1 (*TSP-1*) and tenascin-C (*TN-C*). However, research now shows

the connective tissue growth factor and Nov (*CCN*) family, periostin (*POSTN*), small leucine-rich proteoglycans (*SLRPS*), R-spondins, fibulins, galectins, pigment epithelium derived factor (*PEDF*), plasminogen activator inhibitor type 1 (*PAI-1*) and autotaxin to also fit this description (Murphy-Ullrich and Sage, 2014; Bornstein, 2009).

1.2.1 The SPARC family

The 8 members of the SPARC family of matricellular proteins consists of SPARC, hevin, *SPOCK-1/-2/-3, SMOC-1/-2* and follistatin-like 1 (*FSTL-1*). These proteins all share structural similarity in the 3 main domains: 1) a highly acidic domain with low calcium binding affinity 2) a follistatin-like domain and 3) a calcium binding extracellular domain (EC) (Fig. 1.2). The shared structure suggests similarities/overlap in their function (Viloria *et al.,* 2016). While the function of *SPARC* has been studied extensively, the mechanisms underlying the role of the wider SPARC family, particularly in reference to pancreatic tumorigenesis, remains unclear. Table 1.1 summarises some of the known role of the SPARC family.
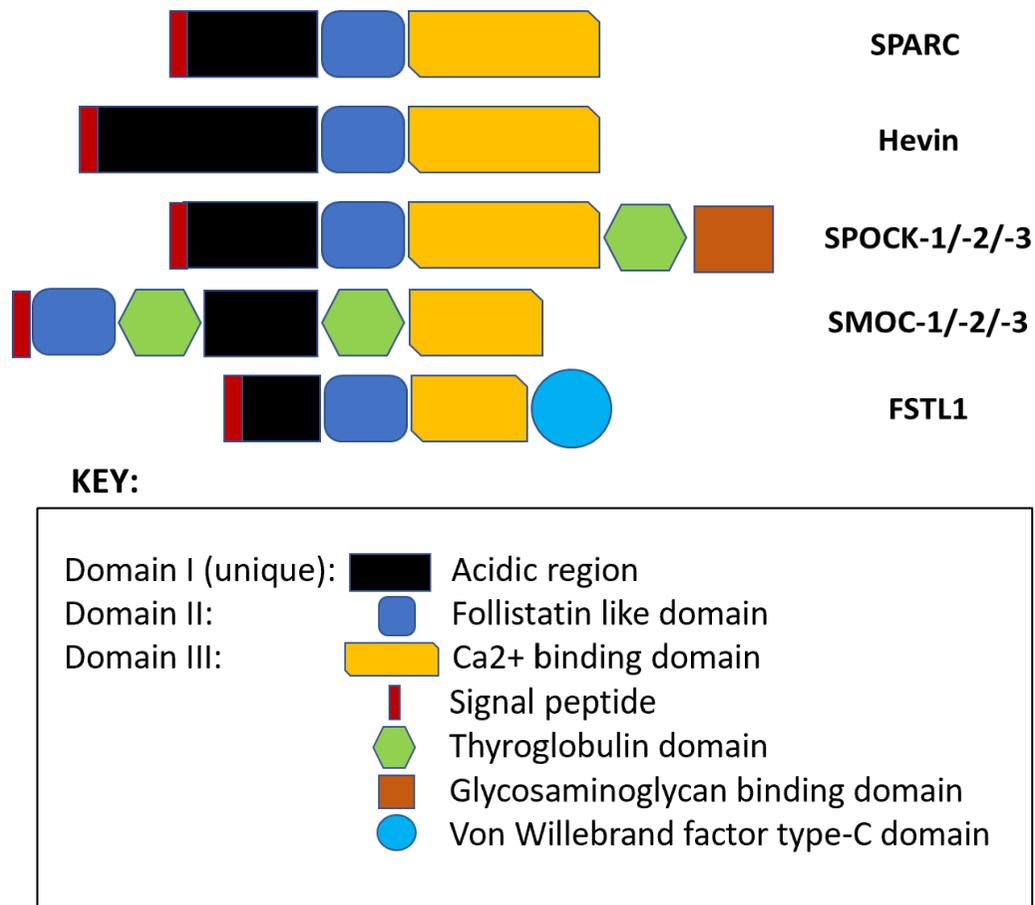
Figure 1.2: Diagram of the structural domains of the SPARC family members. All members of the SPARC family of proteins share domains I (black), II (blue) and III (yellow). With the exception of the SMOC proteins, the signal peptide (red) is located in domain I. In addition to the three main domain, the SPOCK proteins contain a single thyroglobulin domain (green) while the SMOCs contain two. The glycosaminoglycan binding domain is distinctive to the SPOCK proteins (orange) as is the von Willebrand factor type-c domain (cyan) to FSTL1. Information acquired from Viloria *et al*., 2020.

Table 1.1: Known functions of the SPARC family

SPARC family member	ECM activity	Growth factor activity	Cell adhesion	Role in PDAC
SPARC	• Mediates procollagen I fibrillation and collagen I deposition into the ECM (Rentz *et al.,* 2007) • Decreases collagen IV in basement membrane (Morrisey *et al.,* 2016) • Modulates cell secretion of fibronectin and laminin (Kamihagi *et al.,* 1994)	• Modulates VEGF (Chandrasekaran *et al.* 2007), TGF-β (Frankie *et al.,* 2004), NGF-β, PDGF-BB and FGF-2 activity (Okura *et al.,* 2019)	• Mediates focal adhesion disassembly (Murphy-Ullrich *et al.,* 1995)	• Overexpression in tumour stroma correlated with decreased survival (Murakawa *et al.* 2019) • Inhibits and enhances cell proliferation in vitro (Munasinghe *et al.,* 2020)
SPARC-like1/ Hevin	• Increases collagen fibrillogenesis (Sullivan *et al.* 2006)	• Undetermined	• Inhibits focal adhesion formation (Girard and Springer, 1996)	• Overexpressed in PDAC. Correlated with increased vascularity and reduced invasiveness (Esposito *et al.,* 2007)
SMOC-1	• Binds tenacin-c (Brellier et al, 2011), fubulin-1 and vitronectin (Novinec *et al.,* 2008) • Localisation in the basement membrane (Vannahme *et al.,* 2002)	• Inhibits TGF-β/BMP signalling (Rainger *et al.,* 2011; Dreieicher *et al.,* 2009) • Binds pro-EGF (Thomas *et al.,* 2016)	• Undetermined	• Undetermined
SMOC-2	• Undetermined	• Enhances VEGF and bFGF-dependent angiogenesis (Rocnik *et al.,* 2006)	• Promotes epidermal cell attachment and focal adhesion formation (Maier,	• Undetermined

		• SMOC-2 expression is upregulated by TGF-β (Gerarduzzi *et al.,* 2017)	Paulsson and Hartmann, 2008)	
SPOCKS/ Testicans	• Effects cathepsin-k • SPOCK-1 & -3 inhibits pro-MMP2 via MT1-MMP or MT3-MMP (Nakada *et al*. 2001) • SPOCK-2 prevents pro-MMP2 inhibition by SPOCK-1 & -3 (Nakada *et al.,* 2003) • SPOCK-1 alters collagen deposition (Veenstra *et al.,* 2017)	• SPOCK 1 expression is upregulated by TGF-β and PDGF-BB (Du *et al.,* 2020)	• SPOCK-1/-2 prevented neurite extensions (Schnepp *et al.,* 2005; Marr and Edgell, 2003)	• SPOCK1 overexpression in tumour stroma correlated with increased invasiveness and poor prognosis (Veenstra *et al.,* 2017) • SPOCK1 induces EMT and metastasis in PDAC by activation of the NF-κB pathway (Cui *et al.,* 2022)
FSTL-1	• Promotes expression of MMP-1 and MMP-13 (Hu *et al.,* 2019) • Promotes ECM protein expression (Chaly *et al.,* 2015)	• Inhibits activin A and BMP-4 signalling but not BMP-2 (Geng *et al.,* 2011; Eijken *et al.,* 2007) • VEGF may inhibit FLST-1 expression (Niu *et al.,* 2021)	• FSTL1 reduced TNF-α-induced VCAM-1 and ICAM-1 (Ghim *et al.,* 2021)	• Inhibits PDAC cell proliferation (Viloria *et al.,* 2016)

1.2.2 *SPOCK1*/testican1

At the point of discovery, the *SPOCK*/testican proteins were referred to simply as glycosaminoglycan-bearing polypeptide in human seminal plasma with further characterisation eventually leading to giving the protein a name (Fig. 1.3) (Alliel *et al.*, 1993; Bonnet *et al*., 1992). While SPOCK2 (N-glycosylated) and SPOCK3 (Multiple O-glycans) are pure heparan sulfate proteoglycans, SPOCK1 carries an additional chondroitin sulphate chain- suggesting each has different functions (Sun *et al*., 2020).
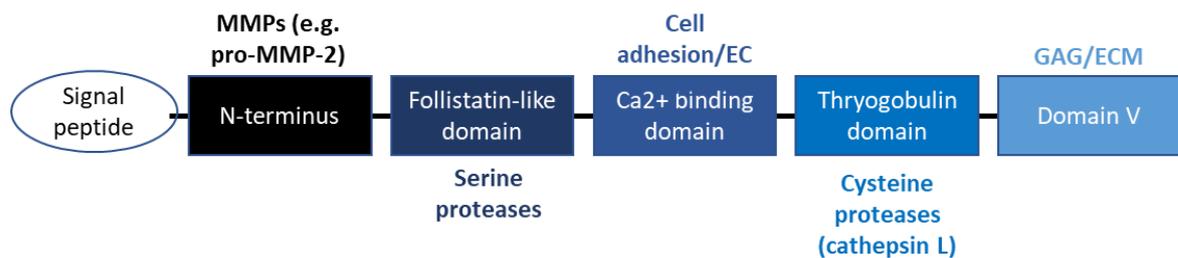
Figure 1.3: The structure of SPOCK-1 protein and potential interaction sites for the ECM and soluble factors. Adapted from Sun *et al*. (2020)

The biological functions of *SPOCK1* have been an area of interest, as most matricellular proteins seem to be key drivers of tumorigenesis. *SPOCK1* is highly expressed in the central nervous system but it can be detected in other tissues and organs (Sun *et al*., 2020). *SPOCK1* has been implicated in the growth, proliferation, colony formation, invasion, metastasis and EMT of gallbladder cancer by activating the PI3K/AKT pathway (Shu *et al.,* 2015). Similarly, *SPOCK1* expression correlated with high histological graded breast cancer samples, high tumour size and triple negative phenotype (Fan *et al.,* 2016). The authors suggest a link between TGF-β and *SPOCK1*, the interaction likely drives EMT in breast tissue. Since the first link between *SPOCK1* and oncogenesis was made in 2012, *SPOCK-1* had been negatively linked to 36 tumour types by 2020 (Váncza *et al.,* 2022).

Even in recent years, the understanding of *SPOCK1'*s role and importantly, its mechanistic function in PDAC, are severely lacking. In mouse neuroblasts cultures, *SPOCK1* blocked attachment sites on culture ware and prevented the ability of the neuroblasts to form neurite extensions although this did not prevent attachment of *SPOCK1* pre-treated cells (Marr and Edgell, 2003). A later study by Schnepp *et al.* (2005) demonstrated similar results. This may suggest a role of SPOCK-1 in cell adhesion.

Currently, there is no experimental evidence showing direct binding between *SPOCK1* and the ECM. Visualisation of PDAC and *SPOCK1* knockdown (KD) mouse embryonic fibroblast (MEF) co-cultures using picrosirius red suggests that SPOCK1 mediates collagen deposition via stromal cell manipulation (Veenstra *et al.,* 2017). In addition, they showed that SPOCK1 in PDAC show its overexpression in the stromal compartment strongly correlated with markers of an activated stroma such as SPARC, αSMA/ACTA2 and FAP and poor patient prognosis. Importantly, Panc-1 proliferation index was significantly reduced in *SPOCK1* knockdown MEF co-cultures suggesting tumour promoting functions of *SPOCK1* in PDAC.

SPOCK-1 has been shown to inhibit pro-MMP2 (Nakada *et al.* 2001), with SPOCK2 having the ability to negate this effect (Nakada *et al.,* 2003). It may be possible that SPOCK1 plays a role in the matrix remodelling via MMPs in PDAC where ECM proteins are highly upregulated.

A recent study of *SPOCK1* function in PDAC was published by Li *et al.* in 2019. They reported that *SPOCK1* knockdown in PCNA-1 and MIA PaCa-2 cells curtailed their proliferative and metastatic potential. Cell cycle arrest at G0/G1 was observed in SPOCK-1 knockdown Panc-1 and MIA PaCa-2 cells and increased cleavage of caspase-3/-9 and PARP. Collectively, they suggest that these observations may be potentially linked to enhanced EMT and PI3K/AKT pathway activity and downregulation of the apoptotic pathway.

1.3 CRISPR/Cas9

CRISPR/Cas9 (clustered regularly interspaced short palindromic repeats) gene editing technique originates from the endogenous bacterial and archaea immune response against viruses. The organisms, after the first encounter with the virus, were incorporating small regions of the viruses' genome (CRISPR sequences/spacers) into their DNA as clustered interrupted repeats. When these organisms encounter the same virous, the Cas endonuclease and the CRISPR sequences will be used to detect and cleave the viral DNA (Han, Pang and Soh, 2020).

CHAPTER 1: Introduction

The CRISPR/Cas9 system includes a specificity determining CRISPR RNA (crRNA), an auxiliary trans activating RNA (tracrRNA) and a Cas9 (CRISPR associated protein 9) endonuclease which will cause the double stranded break on the DNA, allowing gene modification and a protospacer adjacent motif (PAM) locating at downstream of the target site (Zhang, Wen and Guo, 2014).

The crRNA and tracrRNA can be found as a chimeric single-guide RNA (sgRNA) (Redman *et al.,* 2016, Zhang *et al.,* 2015). More than 40 different types of Cas protein families have been discovered (Zhang, Wen and Guo, 2014). Due to the variety of Cas, further classification was in order. CRISPR/Cas9 got arranged into three major categories: type I, II and III. The Cas9 nuclease was later observed in *Streptococcus pyogenes* bacteria, and it is a type II category (Han, Pang and Soh, 2020).

The CRISPR/Cas9 system has been utilised to modify eukaryotic genome via silencing, knocking down or replacing sequences on a gene (Fig. 1.4). When this system acts on the gene it induces a break on both strands of the DNA (double stranded break or DSB). The DSB (double stranded break) will be fixed by the DNA repair mechanisms. Non-homologous end joining (NHEJ) is an error prone DNA repair mechanism with reduced accuracy and often generate insertions or/and deletions which could disrupt the function of the gene. Homology directed repair (HDR) is a more efficient error free method of DNA repair which requires a homology containing DNA donor sequence as a repair template (Zhang, Wen and Guo, 2014). CRISPR/Cas9 utilises both mechanisms depending on the desired result. NHEJ is used for gene Knockdown whereas HDR is used for knockins (addition of a predesigned short sequence) (TAKARA).

CRISPR/Cas9 is a gene editing technique which allows genetic modification and has promising application in biomedical research. Other knockdown methods have been available to us, but CRISPR provides a more quick, cheap and relatively easy way of switching on or off genes (Redman *et al*., 2016).

Numerous studies have applied the CRISPR/Cas9 genome editing technique for the treatment of diseases. Bella *et al.* (2018) attempted to eliminate the human immunodeficiency virus (HIV) virus by using CRISPR/Cas9. In more detail, mice having peripheral blood mononuclear cells derived from HIV-I positive patients were treated with the CRISPR/Cas9 complex. This reduced HIV-I viral DNA fragments in the mice. Other uses

involve correction of mutations like Duchenne muscular dystrophy (DMD) and tyrosinemia where the compromised sequence causing the disease is being replaced with a functioning sequence. gene knockout etc (Mohanty, Dash and Pradhan, 2019).

Although CRISPR/Cas9 is a very promising tool, there are still some milestones that need to be overcome. A major concern has been the off-target effects that may occur even though it is considered an accurate technique, it is not rare to have off target effects which can be unpredictable. A PAM site 3-5nt downstream the target sequence is compulsory for the correct assembly and action of CRISPR/Cas9 (specificity). The implications of long-term consequences of germline modifications are unpredictable. Lastly, the administration of CRISPR/Cas9 Type II derived from *Streptococcus pyogenes,* which causes an immune response destroying the Cas9 endonuclease.
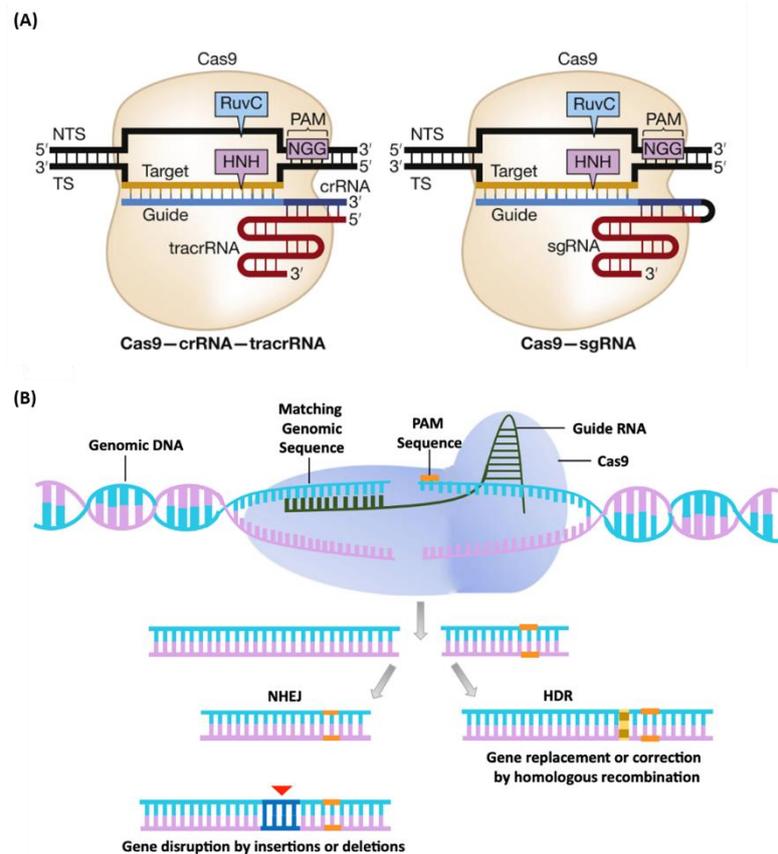
Figure 1.4: CRISPR/Cas9 structure and function. (A) Diagram of Cas9/CRISPR RNA (crRNA) and Cas9-single guide RNA (sgRNA). Higashijima, Y., Hirano, S., Nangaku, M. and Nureki, O., 2017. Applications of the CRISPR-Cas9 system in kidney research. *Kidney International*, *92*(2), pp.324-335. **(B)** Diagram showing RNA guided DNA cleavage by CRISPR-SpCas9. CRISPR/Cas9 binds to the sgRNA targeted region and cleaves the DNA. The Cas9 endonuclease introduces double stranded breaks (DSB) in the DNA. DNA repair mechanisms will attempt to repair the breakage by either NHEJ or HDR. NHEJ is quick and error prone mechanism which may introduce mutations during its attempt to correct the sequence. The mutation will cause the gene to be non-functional (knockdown). HDR is a more accurate error free mechanism which can be used to introduce a new desired sequence in the sgRNA target region (knockin). Ghosh, D., Venkataramani, P., Nandi, S. and Bhattacharjee, S., 2019. CRISPR–Cas9 a boon or bane: the bumpy road ahead to cancer therapeutics. *Cancer cell international*, *19*(1), pp.1-10. TS: target DNA strand, NTS: non-target DNA strand, RuvC and NHN: Cas9 endonuclease domains, PAM: protospacer adjacent motif, NGG: 5'-NGG-'3 PAM sequence of *S. pyogenes* Cas9 (SpCas9), NHEJ: non-homologous end joining, HDR: homology-directed repair

1.4 Aims

The aim of this study is to determine the effect of CRISPR/Cas9-mediated knockout of *SPOCK1* on pancreatic cancer cell and stromal cell adhesion and growth, and to determine whether SPOCK1 protein interacts directly with key extracellular drivers of PDAC tumorigenesis such as MT1-MMP, collagen and FGF.

2 Materials and Methods

2.1 Cell culture

All cell culture reagents were purchased from Fisher Scientific UK and relevant catalogue numbers are referred to throughout this section. Items purchased from other manufacturers/suppliers are specified. A summary table of all reagents and consumables used in this thesis is provided in Materials and Methods Section 2.8.

2.1.1 Cell line acquisition

Pancreatic stellate cells (PS-1) were kindly donated by Professor Hemant Kocher of Queen Mary University of London. The cells were isolated from a donated human pancreas, verified to be of stellate cell origin and immortalised as described by Froeling *et al.* (2009).

AsPC-1, Panc-1 and Capan-1 PDAC cell lines were kindly donated by Dr Charlotte Edling, previously at Queen Mary University of London. Table 2.1 shows the characteristics of the human PDAC and stromal cell lines.

Bovine pulmonary artery endothelial (BPAE) cells were purchased from ATCC.

Table 2.1 **Characteristics of PDAC and stellate cell lines used in this study**

	Cell Line	Derivation	Metastasis	Differentiation
PDAC cell type	**AsPC-1**	Ascites	Yes	Poor
	Capan1	Liver	Yes	Well
	Panc-1	Pancreas	Yes	Poor
Stromal cell type	**PS-1**	Pancreas	N/A	N/A
Data acquired from Kaleağasıoğlu and Berger (2014) and Munasinghe *et al.* (2020)				

2.1.2 Passage and maintenance

Routine PS-1 cell maintenance and passaging was performed in RPMI-1640 basal medium supplemented with 2mM L-glutamine, 100µg/ml penicillin-streptomycin and 10% foetal bovine serum. BPAE cells were maintained in EMEM supplemented with 2mM L-glutamine, 100µg/ml penicillin-streptomycin and 20% foetal bovine serum.

Cells were cultured in filtered T75 flasks in HERAcell 150i incubators set at 37°C, 5% CO_2 in a humidified environment. PS-1 and BPAE cells were passaged at a 1:10 ratio every 3 days using 0.05% trypsin-EDTA. All cell lines were utilised for experiments within 10 passages post thawing.

2.1.3 Cell counting

PS-1 cells were washed with PBS, trypsinized for 3-4min at 37°C, centrifuged at 1250rpm, supernatant discarded, and cell pellet resuspended in 1 ml complete medium. Cell counting was performed prior to cell seeding using disposable haemocytometers. Corning trypan blue cell viability stain was used for cell dilution and the cells were counted under a microscope. The following formula was used to calculate the cell concentration. The cell concentration was adjusted with medium as required for each experiment.

Average number of cells x dilution factor x 10^4 = cell concentration (cells/ml)

2.1.4 Cryopreservation and storage

PS-1 cells were cryopreserved in RPMI-1640 supplemented with 10% DMSO, 20% FBS and penicillin-streptomycin. Each cryogenic vial contained $1x10^6$ cells in 1ml cryopreservation medium. The vials were placed in a Mr Frosty freezing container filled with isopropanol to ensure controlled temperature decrease at an approximate rate of -1°C/minute in an ultra-low temperature freezer (-80°C) overnight. The vials were either stored here for short-term storage or liquid nitrogen (-196°C) for long-term.

2.2 SPOCK1 knockdown in PS-1 cells

2.2.1 SPOCK1 transfection using CRISPR/Cas9

PS-1, ASPC-1 and Panc-1 cells were cultured in a 6-well plate at a density of 0.5-1x10^{6}cells/well in complete medium. 1μM of Cas9 protein was incubated with 1.62μg hSPOCK1 sgRNA in 10mM Tris buffer (pH 7.4) for 5 minutes at room temperature (RT) to form the ribonucleoprotein (RNP). A non-targeting control single guide RNA (sgRNA) was used as a control. The pre-cultured cells were washed in PBS, trypsinised, pelleted and resuspended in 70μl Ingenio electroporation solution. The cell/Ingenio mixture was combined with RNP (100μl total volume) resulting in a final concentration of 3μM guide RNA (gRNA) and 1μM Cas9 and transferred to a 0.2cm cuvette. The cells were electroporated in the cuvette using a 4D-Nucleofector X Unit (Lonza) set with a CA163 pulse. The CA163 pulse was chosen based on previous work in our lab by Fatemia Mohamedi *et al.* (unpublished data). The cells were subsequently transferred to a 6-well plate in 2ml complete medium and cultured in an incubator for 72hrs before performing a (T7 endonuclease 1) T7E1 assay.

2.2.2 T7 endonuclease 1 assay

SPOCK1 primer design

Primers were designed using NCBI Primer Blast to have at least 2 mismatches to any off-target sequences and at least 2 mismatches in the last 5bp at the 3' end (Fig. 2.1). No off-target hits were identified. The primer pair is henceforth referred to as hSPOCK1_T7_for/rev.

```
>hSPOCK1 genomic gene ID 6695
GGGCACCCGTGATATGCCACATACGGTATTGGGGCTGATTAAGCCATATTCCCTCACCGT
TTTTTTTCTTTTTTTTTCTGTGTGTTGTAAAAATCTTCCCATGTCAGTAGTCATGTTTTC
AATGGCTACATTGTTTTCCTCAGTCCATCTGGTTTGTTGAATCTATGCTTCTGGGACAAA
ATCCTCTGCTTCCTATCAATGAGTTGGAGCAAATGGAAGTGCTGGCTAAGGCCCTCAGGA
TCCCTTGGCCTGGTTGGATAGTCCTTTATTTATTTGGCCTTTTATTCTGGGAACTTAGAG
GAAAGACTCATAGGGTCTGGAGCTTCAACTAATTTATCCCCTTTTTTCTCATCTTTAGGC
AAAAGAAGGGGAACGTGGCCCAGAAACACTGGGTTGGACCTTCGAATTTGGTCAAGTGCA
AGCCCTGTCCCGTGGCACAGTCAGCCATGGTCTGCGGCTCAGATGGCCACTCCTACACAT
CCAAGGTTGGTTTTCTTGTCTTTTTATGTTATCAAAAATACTTCATGTCGGCTTGTTCTA
ACACTTCTCTATGCCAAAACTGGGGCTTGGGAAGCGTGGTGGGCTCCAGGCCTTGCCCTT
```

Figure 2.1: SPOCK-1 primer designed with NCBI Primer Blast. Image shows a region of the SPOCK1 gene. The region the forward and reverse primers bind is indicated by the black boxes. The PAM site is highlighted in red. The single guide RNA (sgRNA) binds to the region of interest which is highlighted in bright yellow. The expected cleaved PCR product size is 170 and 380bps.

T7E1 assay

CRISPR/Cas9 induces mutations at the genomic target (PAM) site which consequently results in insertions/deletions in the DNA. These mutations lead to mismatches in the double helix. The T7E1 enzyme recognizes the resulting heteroduplex DNA and therefore cleaves DNA at these mismatch sites. PCR amplification around the PAM site generates PCR products that will be cleaved by T7E1 if Cas9-induced mutations are present. The rate of PCR product cleavage by T7E1 therefore indicates the efficiency of Cas9 cleavage at the target site.

A genomic DNA extraction kit was used to harvest DNA from lysed CRISPR/Cas9 transfected PS-1 cells. PCR was used to amplify endogenous loci using the following conditions: in an Eppendorf, 10µl of 5x Phusion HF buffer was combined with 0.5µl of 500nM forward and reverse primer, 1µl of 10mM dNTPs, 1µl of Phusion Hot Start II High Fidelity DNA Polymerase, 32µl of RNAs free water and 5µl of cell lysate. The thermal conditions applied are shown in Table 2.1.

Table 2.2: **PCR conditions used for the T7E1 assay**

Cycle step	Temperature ℃	Time (min/sec)	Cycles
Initial denaturation	98	3 min	1
Touchdown Annealing	55	10 sec 15 sec	10
Extension	72	30 sec	
Denature	98	10 sec	25
Annealing	55	15 sec	25
Extension	72	30 sec	25
Final extension	72	10 min	1

A 15µl reaction was made by combining 10U/µl of T7E1 in 10x NEBuffer 2 with 300ng of hybridized PCR product for 25min at 37°C. The resulting products were separated on a 2% agarose gel and the cleaved band quantified using GelAnalyser software. The % gene edit was calculated using the Horizon's Beta Tool (Available at: https://horizondiscovery.com/en/ordering-and-calculation-tools/t7ei-calculator).

2.2.3 Agarose gel electrophoresis

50x TAE buffer (242g Tris, 57.1ml glacial acetic acid, 19.66g or 100ml of 0.5M EDTA) was diluted to 1X in dH$_2$0. 1g of agarose powder was dissolved in 50ml 1x TEA buffer by heating in a microwave for 90 seconds. 5µl GelRed Nucleic Acid Stain was added to the gel and poured into a pre-taped casting tray. A gel comb was added to the molten gel and allowed to solidify at RT. Post solidification, the tape was removed, and the gel was placed in a BIO-RAD gel electrophoresis machine. The comb was removed and 5µl of 1kb DNA ladder and 15µl of PCR product was loaded into the wells. The gel was run for 60-90 minute at 75V to separate the bands and viewed using a Syngene G-Box.

2.2.4 Isolation of single cells from CRISPR/Cas9 transfected cells using an array serial dilution

Cells transfected with the RNP complex were cultured in a 6-well plate until 70% confluent. A cell concentration of 20,000 cells/ml was made, and 4,000 cells were seeded in well A1 (200µl) in a 96-well cell culture plate. 100µl/well of complete medium was added to the remaining 95 wells. A two-fold serial dilution was performed in this plate to isolate single cells from the heterogenous gene pool: 1) 100µl of cell suspension was

transferred vertically from well A1 through to H1 2) 100µl of cell suspension was transferred horizontally in each row from wells 1 through to wells 12 (Figure 2.2). After performing the serial dilution, well A1 is empty and a 100µl of the volume in well H1 was discarded to ensure that all the remaining wells in the plate will have a total volume of 100µl/well. The cells were cultured at 37°C, checked for single cell colonies, and expanded for 2.5 weeks to create monoclonal colonies. The single cell colonies were subsequently transferred into a 24-well plate and into a 6-well plate after another 2.5 weeks of culture for further expansion.
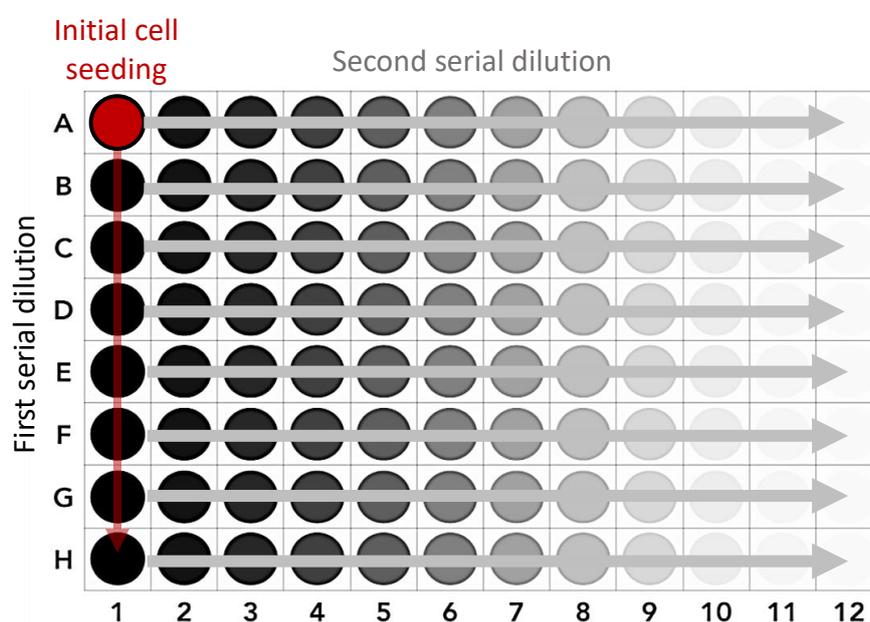
Figure 2.2: Performing an array serial dilution for the isolation of single cells. 4,000 cells were seeded in well A1 and a 2-fold serial dilution was first performed in the column from A1 to H1. A second 2-fold serial dilution was subsequently performed horizontally in each row from wells 1 to 12.

2.2.5 DNA extraction

A Purelink Genomic DNA Mini Kit was used to extract DNA from the CRISPR/Cas9 transfected PS-1 cells. The manufacturers' protocol was followed. Briefly, the cell pellet ($\leq 5 \times 10^6$ cells) was resuspended in 200µl PBS and treated with 20µl proteinase K and 20µl RNase A for 2min at RT. Lysis/binding buffer (200µl) was added to the mixture and incubated at 55°C for 10min. 200µl of absolute ethanol was added to the lysate. The mixture (total volume: 640µl) was placed into a spin column and centrifuged at 10,000g for 1min. Discard flow through and place column in a new collection tube. Wash Buffer 1

(500µl) was added to the column and centrifuged at 10,000g for 1min. Discard flow through and place column in a new collection tube. The spin column was transferred to a new collection tube and washed with washed with Wash Buffer 2 (500µl) at 13,000g for 3min. The collection tube was discarded. The collection tube was placed in a sterile 1.5ml microcentrifuge tube. 25µl genomic elution buffer was added to the column and incubated for 1min at RT. The tube was centrifuged at 13,000g for 1min and the extracted DNA was stored at -20˚C until required for PCR.

2.2.6 Preparation of transfected cell lines for sequencing

The concentration of DNA extracted was quantified using a Biodrop and amplified using PCR. A PCR was prepared as shown in Table 2.3.

Table 2.3: Components of a 50µl PCR reaction

DreamTaq Green PCR Mastermix (2X)	25µl
Forward primer	5µl (of 10µM stock primer)
Reverse primer	5µl (of 10µM stock primer)
Template DNA	1µg (volume dependent on concentration of extracted DNA)
Nuclease free water	Adjust as required
Total volume	50µl

PCR was performed using the thermal cycling conditions recommended by ThermoScientific for their DreamTaq Green PCR MasterMix [K1081] (Table 2.4).

Table 2.4: PCR thermal cycling conditions before sequencing

Step	Temperature (˚C)	Time	Number of Cycles
Initial denaturation	95	1-3 min	1
Denaturation	95	30 sec	
Annealing	55	30 sec	35
Extension	72	1 min	
Final extension	72	5-15 min	1

5ul volume of each PCR product was transferred to an Eppendorfs. 5µl of hSPOCK1_T7_for/rev primer was then added to each sample. In an initial experiment, in order to identify which primer is more effective, the transfected lines were sent with both primers in different Eppendorfs. Samples were sent to GENEWIZ Inc. for sequencing.

2.2.7 Sequence analysis

The results were uploaded on the official GENEWIZ website. The sequence traces were viewed using BIOEDIT software. Subsequently, the sequences were extracted from BIOEDIT as FASTA files for further analysis.

To identify any mutations on the transfected lines, the sequences were aligned against the same area of the *hSPOCK1* gene using CLUSTAL OMEGA software available at EMBL-EBI (https://www.ebi.ac.uk/).

2.2.8 Predicting off-target mutations

In an effort to investigate possible off target effects of CRISPR/Cas9, BLAST software was used to identify sequences in the human genome with high similarity to the sgRNA used, where the sgRNA could have potentially bound. This was done by searching against 'Genomic + transcript databases' and 'Human genomic plus transcript' for the region of the *SPOCK1* that the sgRNA was meant to bind. The top ten genes with the highest percentage of similarities were assessed for the presence of downstream PAM site.

2.3 Western blot

Western blot was used to test SPOCK1 protein in PS-1 cells. PS-1 cells were cultured until approximately 70% confluent. PS-1 cell lysate was produced by incubation with RIPA lysis buffer containing protease inhibitor cocktail for 20 minutes at 4°C. The cell lysate was cleared by centrifugation at 12,000rpm for 15min, diluted in 5x loading buffer (312mM Tris-HCL of pH6.8, 10% Sodium dodecyl sulphate, 10% β-mercaptoethanol, 25% glycerol and 0.015% bromophenol blue) and placed in a heat block set at 100°C for 6 minutes. The sample was loaded onto a 12% SDS-PAGE (Resolving gel: 3.75ml 1M Tris pH8.8, 3.6ml 30% v/v acrylamide, 50μl 20% w/v SDS, 2.3ml dH$_2$O, 12.5μl TEMED and 300μl 10% w/v APS. Stacking gel: 625μl 1M Tris pH6.8, 900μl 30% v/v acrylamide, 25μl 20% w/v SDS, 3.29ml dH$_2$O, 12.5 μl TEMED and 150μl 10% w/v APS) for gel electrophoresis at 120V for 1.5 hours. Subsequently the samples were transferred onto a 0.2μm pore nitrocellulose membrane using a Turbo-blot semi-transfer system and kit. Post transfer, the membrane was blocked with 5% milk solution and incubated overnight at 4°C with SPOCK1 antibody diluted 1:1000 in TTBS and mouse beta actin at 1:5000.

The following day, the membrane was washed with TTBS, incubated with IRDye 680RD goat anti-mouse IgG (to detect beta actin) and IRDye 800CW goat anti-rabbit IgG (to detect SPOCK1) secondary antibodies, washed with TTBS and scanned with an Odyssey CLx Infrared Imaging System (Li-cor).

2.4 Solid phase binding assay to determine SPOCK1 binding partners

SPOCK1 (positive control for assay), FN, MT1-MMP, FGF2 and COL1A1 proteins purchased from Bio-techne were diluted in a 100mM bicarbonate/carbonate coating buffer (pH 9.6) to a final concentration of 0.1µg/ml. These protein solutions (100µl/well) were incubated overnight at 4°C in a non-treated 96-well microplate to allow the proteins to bind to the plate. The solution in the wells were discarded and washed in TBS to remove excess unbound protein. The wells were blocked with 2% casein for 1hr at RT to reduce non-specific binding of the SPOCK1 protein to the well in subsequent steps. The wells were washed in TBS and incubated with 50µl/well binding solution (SPOCK1 protein solution) at 37°C for 1hr to allow for any potential binding between SPOCK1 and ECM proteins to occur. The wells were washed to remove any unbound SPOCK1 protein and incubated with SPOCK1 antibody solution (1:500 in TBS) for 2hrs at RT. Subsequently, the wells were washed and incubated with an anti-rabbit HRP-linked secondary antibody solution (1:350 in TBS) for 1hr at RT. After a final washing step, the wells were incubated with 50µl/well TMB ELISA substrate solution for 3-15 minutes at RT for colour to form but not saturate. 50µl of a 0.5M H_2SO_4 was added to each well to stop the reaction and the plate was read immediately at 450nm using a spectrophotometer (Bio-Tek Instruments, Epoch 2).

2.5 Alamar blue viability assay

AsPC-1 and Panc-1 (± *SPOCK1* CRISPR/Cas9 treatment) cells cultured at density of 4,000 cells/well in a 96-well plate for 3 days. On the third day of culture, 10µl of alamarBlue reagent was added to the culture medium for the final 3hrs of the culture period. An Epoch 2 microplate spectrophotometer was used to read the optical density at 570nm and 600nm.

2.6 Live cell imaging and image analysis

Cells were seeded in 6, 12 or 24 Corning cell culture plates for experiments were placed in an IncuCyte Zoom live cell imaging system (Sartorius) for imaging and image analysis for cell confluence and eccentricity. At least 3 technical replicates were used per treatment.

Using the IncuCyte Zoom software, the position of the plates inserted into the machine was selected from the 'Drawer Setup', the Corning plate catalogue number was inserted into 'Add Vessel' and a 'Scan Pattern' of four images per well was selected. Plate maps were created to label samples. The software was set to capture images at 24hr intervals up to 96 hrs for PS-1 cells. For PDAC cell lines, images were captured every hour for 96 hrs.

From the 'Analysis Job Utilities' tab an image collection was created for each cell line imaged (PS-1, AsPC-1 and Panc-1) by adding 20 images from all time points into individual collections. A 'New Processing Definition' was launched by using these images to adjust parameters for the software to accurately recognise cells from the background. Cell debris or artefacts were eliminated from detection/analysis by excluding any area below 300μm^2. All images captured were analysed overnight using the parameters set from the image collections for each cell line by selecting 'Launch New Analysis'. Once complete, the analysis was opened and 'Export Metrics' was selected to include all the wells and time points of interest (0-24hrs for attachment assay; 0-24, 24, 48, 72 and 96hrs for growth assay). Cell confluence percentage or eccentricity was chosen under 'Phase Metrics' and the 'Data Export' tab was used to export the raw data. Data was exported using the average of the four images/well to improve accuracy.

2.7 Focal adhesion assay

An Actin Cytoskeleton / Focal Adhesion Staining Kit was purchased from Merck Life Science. 5,600 BPAE cells/well were cultured overnight in Lab-Tek II chamber slides. Post culture, the cells were washed with PBS, fixed with 4% PFA for 10 minutes, washed again and permeabilised with 0.1% triton-x100 for 4 minutes at RT. The cells were washed and blocked with 1% BSA for 30minutes at RT before incubation with anti-vinculin antibody

(1:200) for 1 hour at RT. Following incubation, the cells were washed and incubated for 1 hour with a secondary anti-mouse Alexaflor488 antibody (1:200) and TRITC-conjugated phalloidin (1:200). The cells were washed 3 times in PBS and incubated with DAPI (1:1000) for 5 minutes at RT. A drop of ProLong Gold anti-fade mounting medium was used before applying a coverslip (No. 1.5) and sealing with clear nail varnish. A 63x objective was used to image the cells under oil immersion on a confocal microscope (Carl Zeiss LSM800).

2.8 Statistics

Raw data were analysed using two-tailed T-test or two-way ANOVA using Microsoft Excel or GraphPad Prism 9 software. P-values less than or equal to 0.05 was considered statistically significant.

2.9 Reagents and consumables

Table 2.5. lists the primary reagents and consumables used throughout this study.

Table 2.5: Reagents and consumables

REAGENT	CATALOGUE NUMBER	PRODUCER/DISTRIBUTOR
Cell Culture		
6-well cell culture plates	10578911	Fisher Scientific
24-well cell culture plates	10732552	Fisher Scientific
96-well cell culture plates	10695951	Fisher Scientific
Cryogenic tubes	11750573	Fisher Scientific
DMSO	D2650-100ML	Merck Life Science
DPBS (no CaCl$_2$, no MgCL$_2$)	12037539	Fisher Scientific
EMEM	30-2003	ATCC
FastRead counting slides	BVS100	Immune Systems
Fetal bovine serum	11573397	Fisher Scientific
L-glutamine (200 mM)	11500626	Fisher Scientific
Mr. Frosty freezing container	10110051	Fisher Scientific
Penicillin-streptomycin (10,000 U/mL)	11548876	Fisher Scientific
RPMI 1640 medium (1x)	12004997	Fisher Scientific
T75 cell Culture Flasks	10364131	Fisher Scientific
Trypan blue	15393661	Fisher Scientific

Trypsin-EDTA (0.05%)	11590626	Fisher Scientific
PCR		
DreamTaq Green PCR Master Mix (2X)	K1081	Thermo Scientific
DNA extraction		
Purlink Genomic DNA Mini kit	10053293	FisherScientific
CRISPR/CAS9		
10 mM Tris-HCl buffer, pH 7.4	B-006000-100	Horizon Discovery
Cas9 nuclease protein NLS	CAS12205	Horizon Discovery
Edit-R predesigned synthetic sgRNA	SG-013724-01-0002	Horizon Discovery
Edit-R synthetic sgRNA non-targeting controls	U-009501-01-02	Horizon Discovery
Electroporation solution	MIR 50111	Cambridge Bioscience
hSPOCK1_T7_for primer	Custom primer design: CACATACGGTATTGGGGCTGAT	Merck Life Science
hSPOCK1_T7_rev primer	Custom primer design: CAAGCCCCAGTTTTGGCATAG	Merck Life Science
T7E1 Assay		
Genomic DNA mini kit	10053293	Fisher Scientific
NEBuffer 2	B7003S	New England Biolabs
Phusion HF buffer pack	10492088	Fisher Scientific
Phusion hot start II high-fidelity DNA polymerase (2 U/μL)	10441338	Fisher Scientific
T7 endonuclease I (10U/μl)	M0302S	New England Biolabs
Agarose Gel Electrophoresis		
Agarose	R0491	Thermo Fisher Scientific
EDTA	10306983	Fisher Scientific
GelRed	41003	VWR
GeneRuler DNA ladder	SM0311	Thermo Fisher Scientific
Tris(hydroxymethyl)aminomethane	252859-500G	Merck Life Science
UltraPure agarose	16500500	Thermo Fisher Scientific

CHAPTER 2: Materials and Methods

Solid Phase Binding Assay		
0.5M H_2SO_4	ARK2197	Merck Life Science
3,3',5,5'-tetramethylbenzidine (TMB) liquid substrate system for ELISA	T0440-100ML	Merck Life Science
96-well microplates (non-treated)	10216341	Fisher Scientific
Actin Cytoskeleton / Focal Adhesion Staining Kit	FAK100	Merck Life Science
Alexa Fluor 488 goat anti-rabbit IgG	10236882	Fisher Scientific
Bovine Serum Albumin	A2153	Merck Life Science
Casein	10545691	Fisher Scientific
Coverslips No. 1.5, 22 × 22 mm	Z692271	Merck Life Science
Alamar Blue Assay		
96-well cell culture plate	10695951	Fisher Scientific
AlamarBlue dye	15549604	Fisher Scientific
Focal Adhesion Kinase		
HRP-linked anti-rabbit goat IgG antibody	7074S	Cell Signaling
Human fibronectin protein	1918-FN-02M	Bio-techne
Lab-Tek II Chamber slides	177399	ThermoFisher
ProLong Gold anti-fade mounting medium	11539306	Fisher Scientific
Recombinant human FGF basic/FGF2/bFGF (146 aa) protein	233-FB-025/CF	Bio-techne
Recombinant human MMP-14/MT1-MMP Protein	918-MP-010	Bio-techne
Recombinant human pro-collagen I alpha 1/COL1A1 Protein	6220-CL-020	Bio-techne
Recombinant human testican 1/SPOCK1 protein	2327-PI-050	Bio-techne
Sodium bicarbonate	10020510	Fisher Scientific
Sodium carbonate	10538070	Fisher Scientific
SPOCK1 anti-human rabbit polyclonal antibody	16541322	Fisher Scientific
Triton-x 100	T8787	Merck Life Science
Western Blot		
2-mercaptoethanol	M3148	Merck Life Science

CHAPTER 2: Materials and Methods

30% ProtoGel	EC-890	National DIagnostics
6M HCL	15606880	Fisher Scientific
Ammonium persulphate	A3678-25G	Merck Life Science
Anti-beta actin mouse monoclonal antibody	ab8224	Abcam
Bromophenol Blue	114391	Merck Life Science
Coomassie brilliant blue	11876744	Fisher Scientific
Glycerol	G7757	Merck Life Science
Glycine	G8898	Merck Life Science
IRDye 680RD goat anti-mouse IgG	926-68070	LI-COR
IRDye 800CW goat anti-rabbit IgG	926-32211	LI-COR
PageRuler plus protein ladder	11832124	Fisher Scientific
Paraformaldehyde	158127-500G	Merck Life Science
Protease Inhibitor Cocktail (100X)	10720825	Fisher Scientific
RIPA lysis and extraction buffer	10017003	Fisher Scientific
SDS	L3771-100G	Merck Life Science
SPOCK1 anti-human rabbit polyclonal antibody	16541322	Fisher Scientific
TEMED	T22500	Merck Life Science
Tris(hydroxymethyl)aminomethane	252859-500G	Merck Life Science
Triton -X 100	T8787-50ml	Merck Life Science
Turbo-blot turbo transfer system	17001915	Bio-Rad
Tween-20	P2287-500ml	Merck Life Science

3 Results

3.1 Creating SPOCK1 KD cell lines for functional study

To study the role of SPOCK-1 in PDAC, two PDAC cell lines and a stellate cell line were transfected with CRISPR/Cas9 to knockout *SPOCK1* expression as described in Section 2.2.1. It was therefore necessary to determine the CRISPR/Cas9 transfection efficiency and this was tested using the PS-1 cell line. gDNA was first extracted from the CRISPR/Cas9 transfected cells and a PCR was performed. Prior to performing the PCR, *SPOCK1* primers [hSPOCK1_T7_for/rev] were designed (Section 2.2.2) and these were tested for specificity on the untransfected PS-1 cell line using agarose gel electrophoresis (Fig. 3.1A). Bands were detected at ~170bp, within the expected range of 170-380bp for the cleaved PCR product. After the primers were tested and proven efficient, the PCR was performed using the transfected line. The resulting products were digested with T7E1, an enzyme that recognizes deformities in DNA heteroduplexes (Section 2.2.2), and these were run on agarose gels (Fig 3.2B). A faint second cut band was detected in the *SPOCK1* KD PS-1 cells digested with T7E1 (Fig. 3.2B, red arrow) indicating that the CRISPR/CAS9 transfection was successful.

(A)

(B)
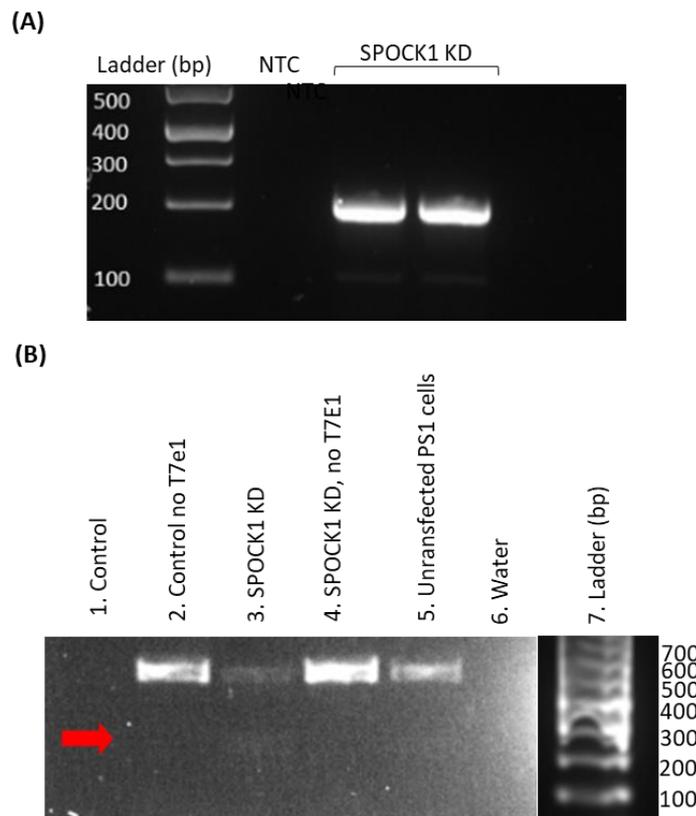
Figure 3.1: Determining CRISPR mutation efficiency using the T7 endonuclease assay. The newly designed hSPOCK1_T7_for and hSPOCK1_T7_rev PCR primers were tested for specificity on 175ng of AsPC1 cell lysate. The expected cleaved products should be within 170-380bp **(A).** PS1 cells were treated with SPOCK1 sgRNA, or a control non-targeting sgRNA. The gDNA was then extracted and PCR amplification performed with SPOCK1 primers flanking the target site. The PCR product was then either digested with T7E1 or left uncut. Untransfected PS1 cells were used as an additional control to verify template integrity. The PCR products were then run on a 2% agarose gel at 75w for an hour. A low exposure image of the gel is shown on the left to display the ladder and a high exposure image on the right shows the samples. The expected size of the uncut band is 550bp, and the cut bands at 170bp and 380bp [red arrow] **(B)**. A low exposure image of the gel was captured to visualize the ladder unsaturated and this image was cropped and attached to the high exposure image for band size reference in image B. NTC: no template control

The CRISPR/Cas9 transfection results in a heterogenous cell population. A 2-fold serial dilution of the CRISPR transfected cells was performed to isolate single cells from the heterogenous population (Section 2.2.3). Wells containing single cells were cultured to allow cell division and repopulate the culture vessel with a clonal population of cells. The DNA was extracted from these cell populations in order to amplify the *SPOCK1* gene and this product was sequenced to detect any mutations present. A gel electrophoresis was performed on the PCR product prior to the sequencing to confirm the amplification (Fig. 3.2).
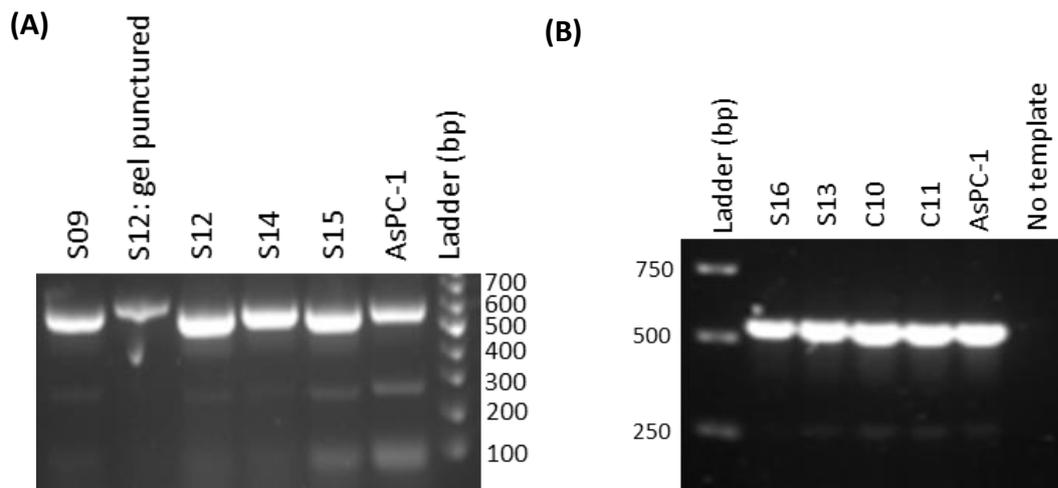
Figure 3.2: SPOCK1 detection in SPOCK1/Control CRISPR/Cas9 transfected PS1 single cell colonies. CRISPR/CAS9 transfected PS1 single cell colony DNA was extracted. A PCR was performed using hSPOCK1_T7_for and hSPOCK1_T7_rev primers. The expected PCR products ~552bp. AsPC-1 DNA was used as a positive control. Gel shows SPOCK1 CRISPR/Cas9 transfected single cell colonies (S09, S12, S14, S15), AsPC-1 cells, a no template control and a 100bp ladder **(A).** Gel shows a ladder, SPOCK1 CRISPR/Cas9 transfected single cell colonies (S13 & S16), control PS1 colonies (C10 &C11), AsPC-1 cells and a no template control **(B)**

While the samples were being sequenced, functional studies were conducted to test the effect of the CRISPR/Cas9 transfection on cell growth and adhesion.

3.2 SPOCK1 knockdown inhibits stromal cell growth and may affect stromal cell adhesion

Stromal cell growth and adhesion were tested using the CRISPR transfected PS-1 cells. Cell count data using a haemocytometer of the heterogenous PS-1 cell population at day 4 of culture showed a 34% decrease in the SPOCK1 knockdown (KD) cell number compared to control (Fig. 3.3A, p=0.0005, n=9 from 3 independent experiments). To confirm this data, a more extensive study was conducted by measuring cell confluence and eccentricity throughout the 4-day culture period via an IncuCyte Zoom live cell imaging system (Sartorius). Image analysis of the SPOCK1 KD heterogenous cell population showed the consistent and significant inhibition of PS-1 cell growth (Fig 3.3B&D). The extent of the attachment of these cells to the culture ware was performed by image analysis measuring cell eccentricity (0=perfect circle, 1=furthest cell spread) at 24hrs when the process of cell

adherence should be complete (Fig. 3.3C&D). The transfected cells were more spread than the non-targeted control cells which were rounder in shape (p=0.05, day1) suggesting the SPOCK1 KD promotes cell adhesion in stromal cells.
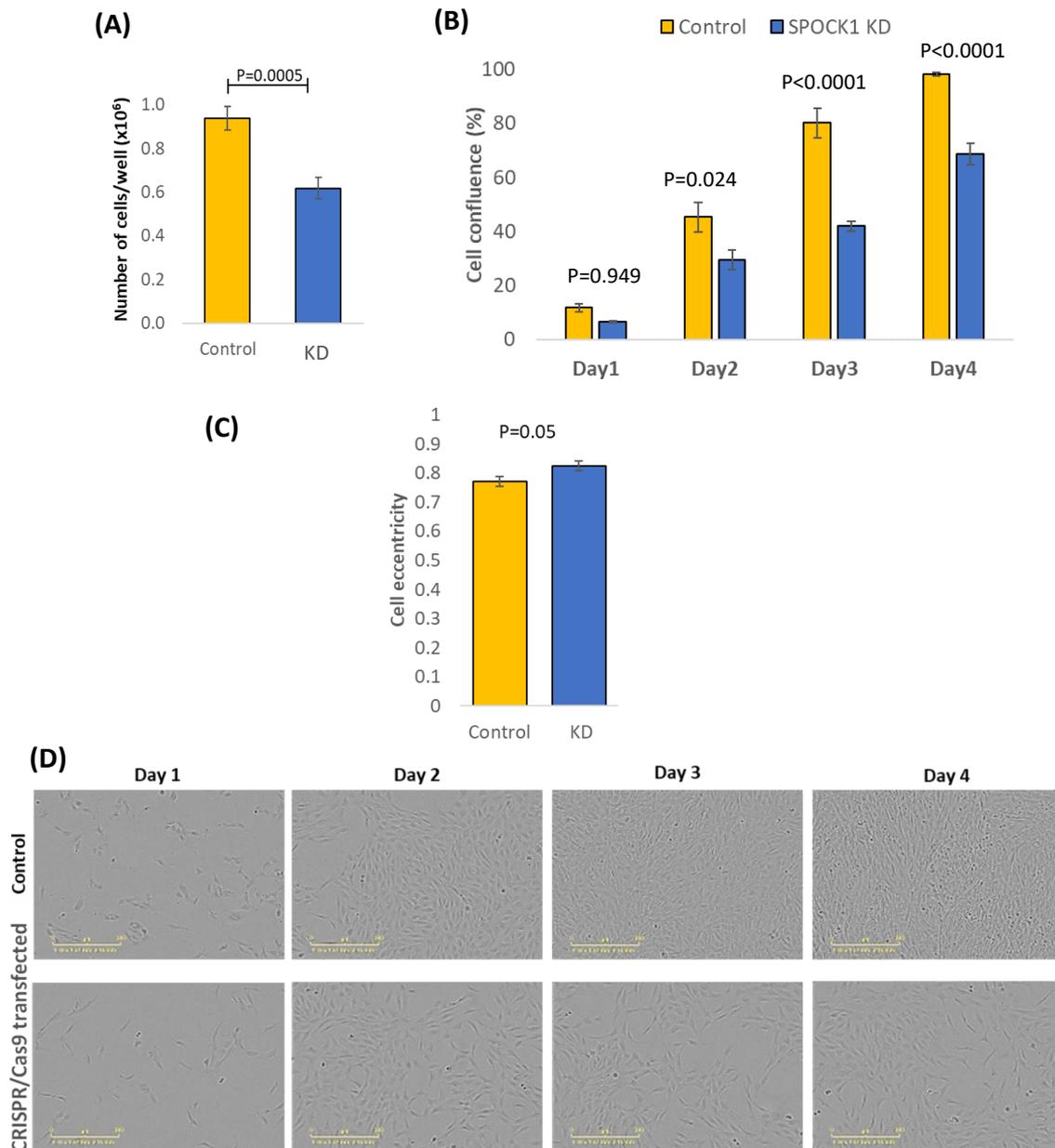
Figure 3.3: SPOCK1 KD inhibits cell growth and promotes adhesion in the heterogenous PS1 cell population. Heterogenous PS1 populations (±SPOCK1 KD) were plated at density of 1.5×10^5 cells/well in a 6-well culture plate and cultured for 4 days. Cells were counted at day 4 of culture. Graph shows number of cells ±SEM, n=9 from 3 independent experiments, T-test. **(A)** A subsequent experiment was conducted to measure cell confluence and eccentricity throughout a 4-day culture period using an IncuCyte Zoom live cell imaging system. Graphs show average confluence **(B)** and cell eccentricity **(C)** N=8-9 from 3 independent experiment, two-way ANOVA. Representative images for (A) and (C) **(D)**

An array dilution of the CRISPR/Cas9 transfected cells was performed to isolate single colonies from the heterogenous population (Section 2.2.3).

Growth was then measured using a haemocytometer in one of the homogenous PS-1 cell populations selected at random which showed even stronger growth inhibition in the SPOCK1 KD population (63% less cells than control, p=0.002, 2-tailed T-test) (Fig. 3.4A). The previous 4-day experiment on the mixed heterogenous PS-1 cell population was then repeated on four randomly selected PS-1 single cell colonies (SC1, SC2, SC3 and SC4) (Fig. 3.4B) compared against a control PS-1 colony (isolated from heterogenous population created using a non-targeting CRISPR/Cas9 transfection). Growth was inhibited significantly in SC1 (Day3: p=0.027, Day 4: p=0.049) and in SC4 (Day3: p=0.035) when compared to control cells. While displaying the same pattern, growth was not inhibited significantly in SC2 and SC3 when (Fig. 3.4B). SPOCK1 KD had no significant effect on cell eccentricity in any of the SCs.

(A)

(B)

46

(C)
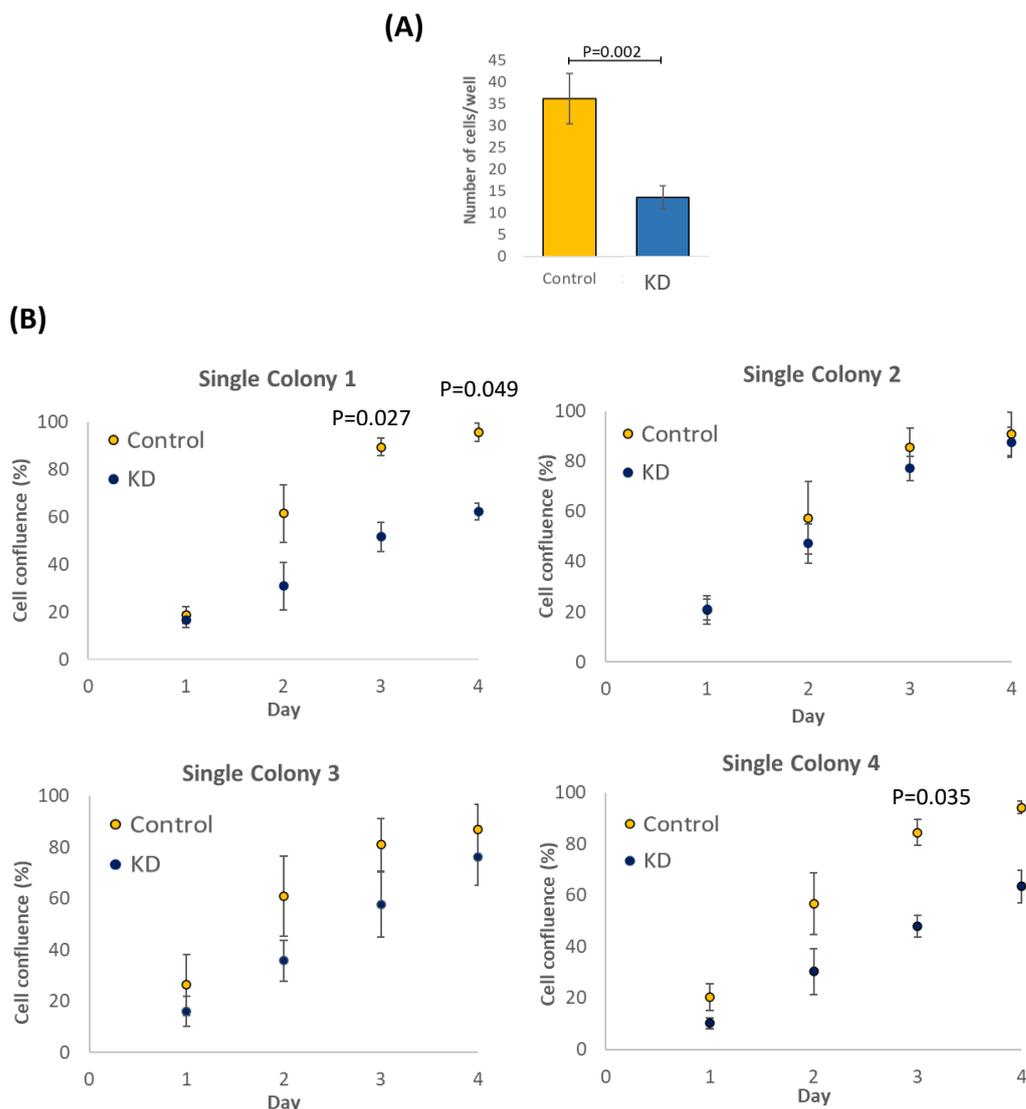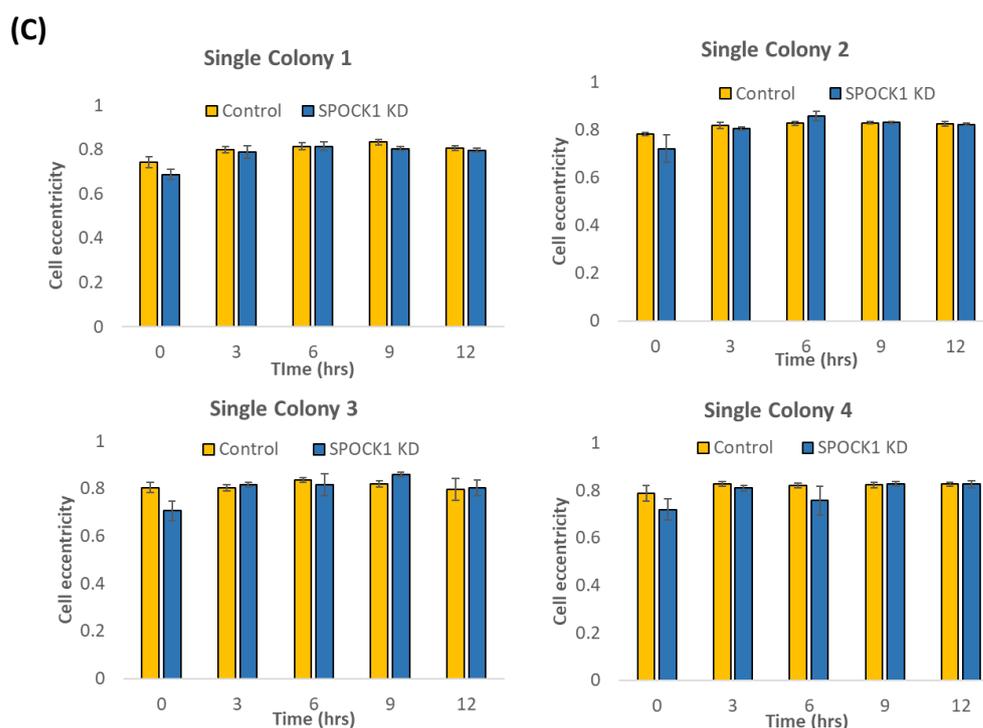
Figure 3.4: SPOCK1 KD promotes PS1 cell growth in select homogeneous PS1 cell populations. Homogenous PS1 cell populations were isolated by a two-fold serial dilution of the heterogenous PS1 cell population. The homogenous PS1 populations (±SPOCK1 KD) were plated at density of 1.5×10^5 cells/well in a 6-well culture plate and cultured for 4 days. Graph show average cell number ±SEM at day 4 of culture. N=12 from 4 independent experiments. Image analysis was performed on images captured with an IncuCyte live cell imaging system. Graph show cell confluence ±SEM **(A)** and eccentricity ±SEM **(C)** in four homogenous single cell colonies (SC1, SC2, SC2 and SC4). N=8-9 from 3 independent experiment. Two-way ANOVA.

3.3 SPOCK1 KD in pancreatic cancer cells affects cell viability

To assess the effect of the SPOCK1 KD in pancreatic cancer cells, Panc-1 and AsPC-1 cell lines were CRISPR/CAS9 transfected in two separate experiments using the same protocol used for PS-1 cells and using electroporation pulse CA163 on a Nucleofector 4. Alamar blue was used to measure PDAC cell viability (Fig. 3.5). Absorbance measurements of the wells after a 4hr incubation with cell viability dye, Alamar blue, showed greater viability in the SPOCK1 KD Panc-1 cells. However, only the cells from the second transfection were slightly above borderline significant (Fig. 3.5A, p=0.06). In stark contrast, cell viability was

reduced in the SPOCK1 KD transfected AsPC-1 cells, with the cells from the second transfection being significantly reduced (p=0.03). Due to the high variability in results between samples in the assay, the experiment was repeated with the Panc-1 cells using live cell imaging (Fig. 3.5B-D). Image analysis of a 4-day culture showed no difference in Panc-1 cell confluency in the SPOCK1 KD cells compared to controls with a slight decrease in confluence at 96hrs however this was not statistically significant (Fig3.5B&C, P=0.204 [CRISPR/Cas9 transfected vs control]). Together this data could suggest that SPOCK1 KD in pancreatic cancer cells can increase or decrease cell viability but the role it assumes seem to be cell type dependent. However, further experiments would be required to validate this preliminary data, including analysis of CRISPR/Cas9 mutation in each colony as it was not possible to perform this analysis within the current study. Capan-1 cells that underwent CRISPR/Cas9 treatment did not survive the procedure and has been excluded in this study as the cells begin rounding and detach from the culture ware (data not shown).

(A)

(B)

(C)
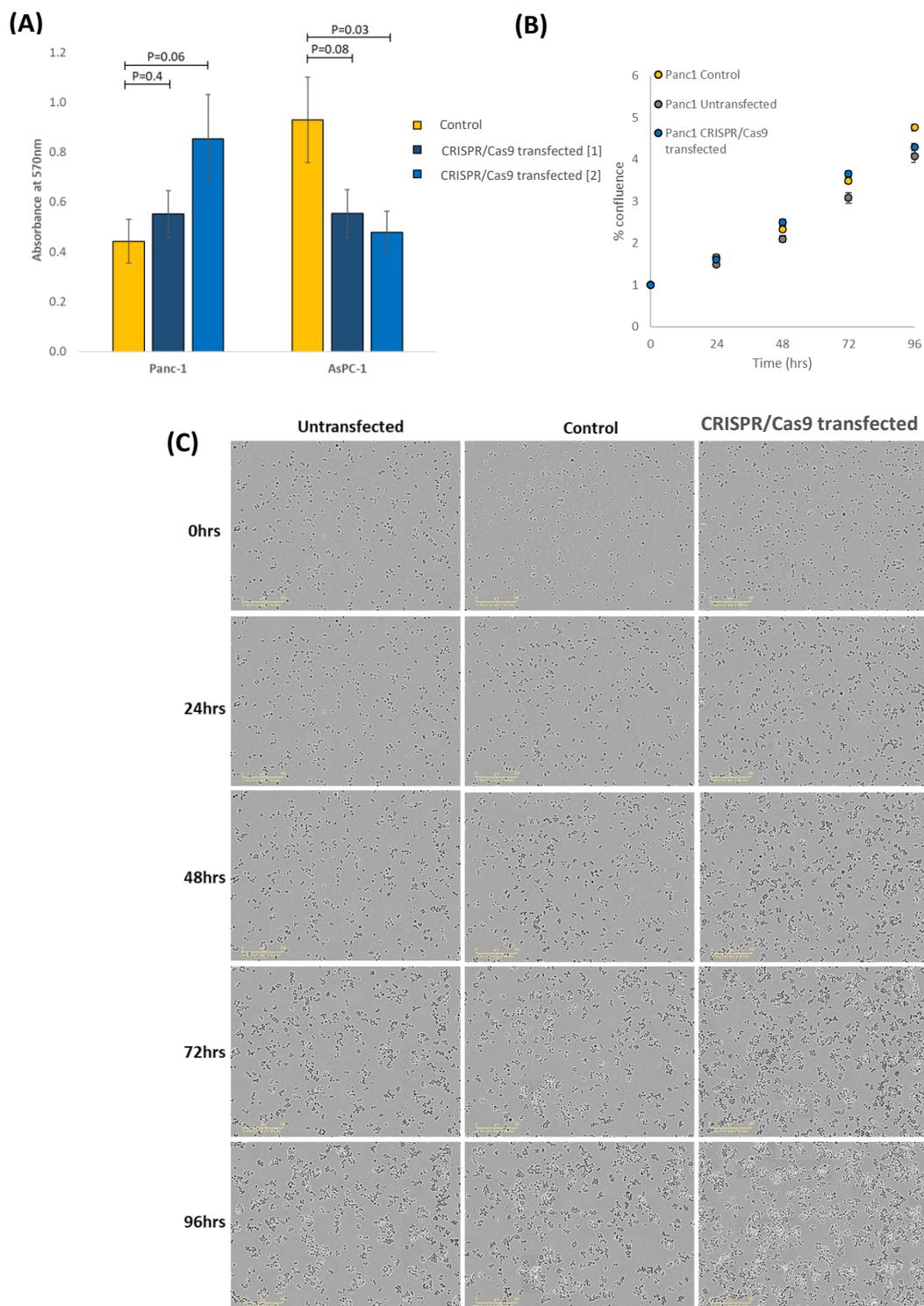
Figure 3.5: CRISPR/Cas9 transfection promotes and inhibits pancreatic cancer cell viability. Alamar blue assay was performed on AsPC-1 and Panc-1 (±CRISPR/Cas9 targeted/non-targeted [control] transfected or untransfected cells) cells cultured at density of 4,000 cells/well in a 96-well plate for 3 days. Graph shows absorbance at 570nm ±SEM, n=12 from 1 independent experiment **(A)**. Growth of Panc-1 cells (±CRISPR/Cas9 targeted/non-targeted [control] transfected or untransfected cells) cultured for 4 days **(B)** with representative images **(C)**. Graph shows % cell confluence standardized to Time 0 ±SEM. N=9 from 3 independent experiments. Two-way ANOVA.

3.4 SPOCK1 KD in pancreatic cancer cells promotes de-adhesion

The effect of SPOCK1 KD on PDAC cell adhesion was subsequently tested. Similar to the PS-1 cells, the extent of the attachment of these cells to the culture ware was performed by image analysis measuring cell eccentricity (0=perfect circle, 1=furthest cell spread) (Fig. 3.6). SPOCK1 KD Panc-1 cells were more spread than the control cells which were rounder in shape at day1 suggesting they are further along in the process of adherence (Fig. 3.6A) This effect is not observed at 48, 72 and 96hrs, likely due to the fact that the cells have fully adhered after this time point with cell death likely to be beginning at day 4. Since cell adherence occurs soon after cell plating, the first 24hrs of culture was analysed in more detail and statistics was carried out at 4hrs intervals (Fig. 3.6B). Interestingly, the SPOCK1 KD cells were more adherent in the first 3 hrs of culture but is similar to the control cells at around 12hrs and continues to adhere strongly thereafter. Together this data suggest that the protein is de-adhesive in PDAC cells.

(A)

(B)

4hrs: P=0.05
8hrs: P=0.03
12hrs: P<0.0001
16hrs: P<0.002
20hrs: P<0.002
24hrs: P<0.001

(D)
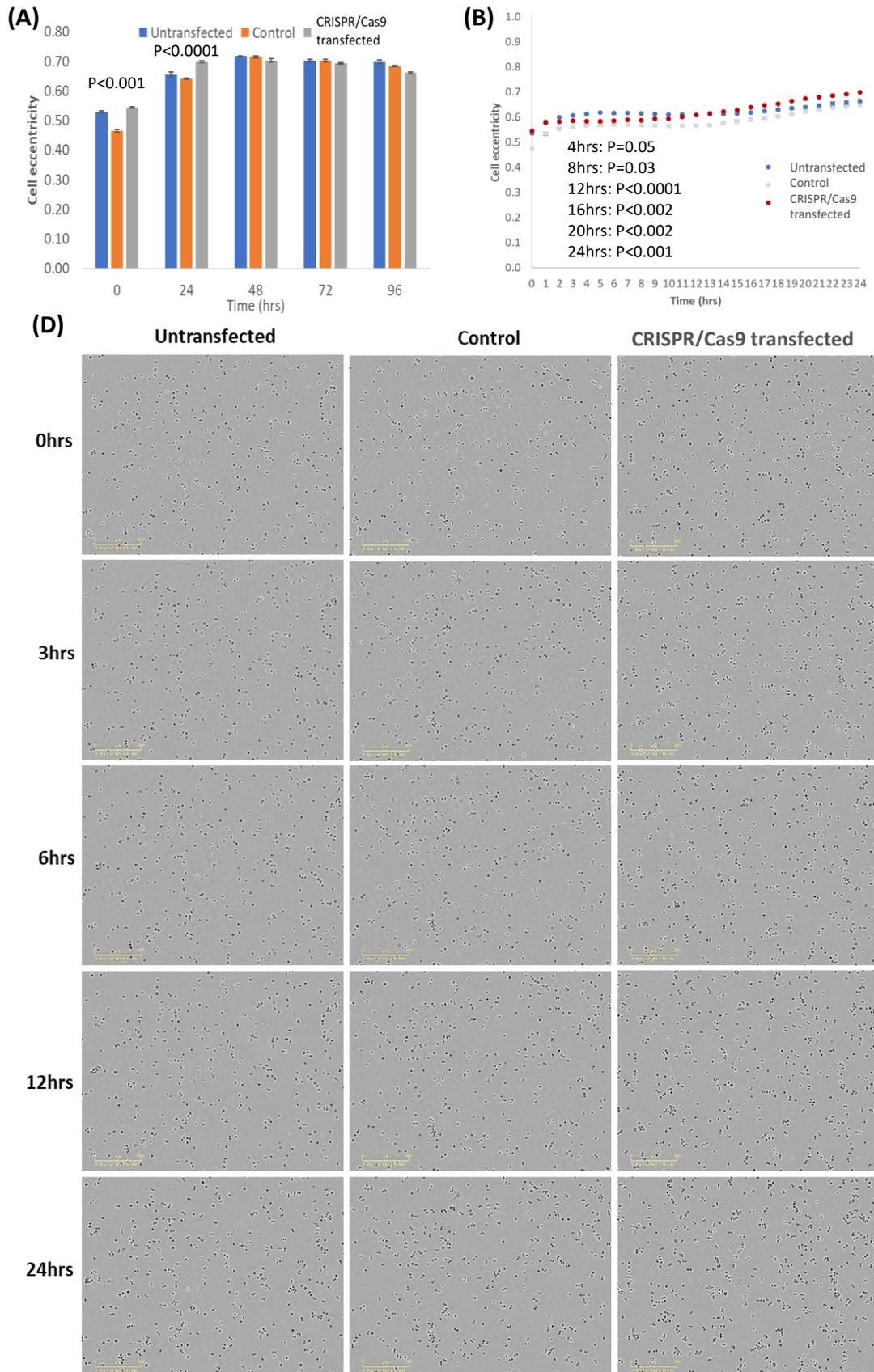
Figure 3.6: SPOCK1 KD inhibits PDAC cell attachment. Cell eccentricity was measured in Panc-1 cells (±SPOCK1 KD or untransfected cells) cultured in 12-well plates for 4 days **(A)**. A detailed analysis of the first 24hrs of the culture is shown **(B)** with representative images shown at 0, 3, 6, 12, and 24hrs of culture **(C)**. Graphs show cell eccentricity ±SEM. N=9 from 3 independent experiments. 2-way ANOVA.

To determine if the de-adhesive effect is achieved by disassembly of focal adhesions, imaging of focal adhesion complexes using an anti-vinculin antibody to visualise focal adhesions, combined with phalloidin to visualise the actin cytoskeleton was performed, initially using (anchorage dependent) BPAE cells to optimise the staining protocol (Fig. 3.7). Unfortunately, due to time constraints the staining was not performed in pancreatic cancer cells, but this remains a potential area for further study.
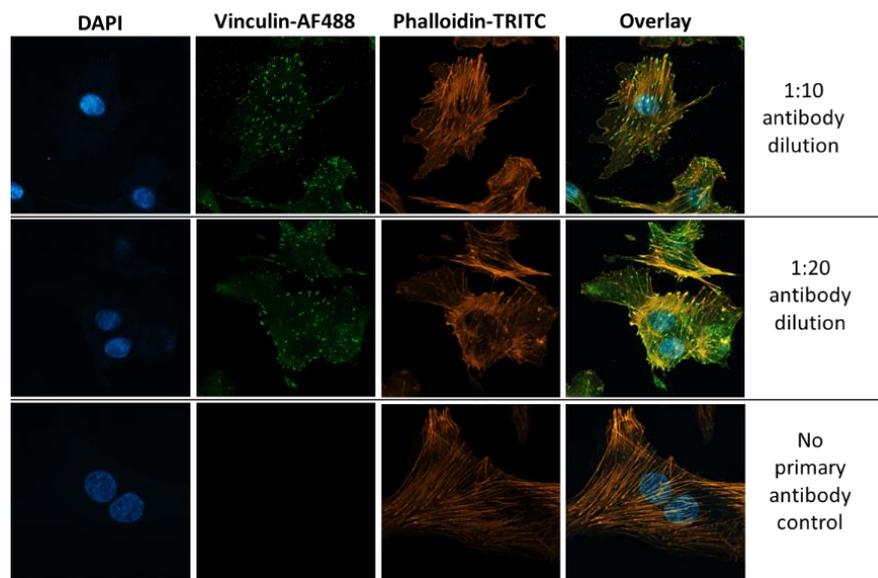
Figure 3.7: Optimization of focal adhesion staining. A FAK kit was purchased from Merck Life Science. BPAE cells cultured on chamber slides (Lab-Tek) were stained according to manufacturer's instruction. Figure shows Vinculin-AF488 antibody staining (green), phalloidin-actin cytoskeleton staining (orange) and DAPI (blue)nuclear staining.

3.5 Sequencing of SPOCK1 KD single cell colonies did not show *SPOCK1* gene mutations

As mentioned previously in Section 3.1, PCR product was sequenced using forward or reverse primers by Genewiz (Azenta Life Sciences) to identify specific *SPOCK1* mutations in the CRISPR/Cas9 transfected cells. After extracting the sequences from BioEdit software 7.1.3.0 (11/04/2011), the quality of each sequence was examined. The samples sent for sequencing with Forward primer presented low quality sequence with numerous N throughout (see Appendix Section 7.2). The samples sent for sequencing with the Reverse primer had good quality sequences and they were used for further analysis.

Using CLUSTAL OMEGA the sequences were examined for mutations. The targeted location of the control transfected line and the targeted transfected line sequences were identical to the original *SPOCK1* gene sequence which confirms the absence of mutation in the area of focus (see Appendix Section 7.1).

A western blot was subsequently performed on the cell lysate to detect SPOCK1 at the protein level however, no bands were detected for any sample or the human recombinant SPOCK1 protein positive control (Appendix Section 7.3).

3.6 Investigating CRISPR/Cas9 off-target mutations

The possibility of an off-target mutation was considered, and possible candidates were predicted using BLAST software provided by NCBI. Ten genes were selected based on the highest similarities to the binding site of the sgRNA (Table 3.1).

Table 3.1: Off-target genes that may have been mutated by CRISPR/Cas9 transfection.

Sequences producing significant alignments									
Select for downloading or viewing reports	Description	Scientific Name	Max Score	Total Score	Query Cover	E value	Per. Ident	Acc. Len	Accession
Transcripts									
Select seq NM_001370.2	Homo sapiens dynein axonemal heavy chain 6 (DNAH6), mRNA	HUMAN	28.2	28.2	73%	24	100.00%	12752	NM_001370.2
Select seq NM_001364293.3	Homo sapiens transportin 1 (TNPO1), transcript variant 4, mRNA	HUMAN	26.3	26.3	68%	96	100.00%	8506	NM_001364293.3
Select seq NM_001363670.2	Homo sapiens calmodulin 1 (CALM1), transcript variant 1, mRNA	HUMAN	26.3	26.3	68%	96	100.00%	4919	NM_001363670.2
Select seq NM_015239.3	Homo sapiens ATP/GTP binding carboxypeptidase 1 (AGTPBP1), transcript variant 2, mRNA	HUMAN	26.3	26.3	68%	96	100.00%	4344	NM_015239.3
Select seq NM_032947.5	Homo sapiens small integral membrane protein 3 (SMIM3), mRNA	HUMAN	26.3	26.3	68%	96	100.00%	1521	NM_032947.5
Select seq NM_015286.6	Homo sapiens synemin (SYNM), transcript variant B, mRNA	HUMAN	26.3	26.3	68%	96	100.00%	6417	NM_015286.6
Select seq NM_001289987.3	Homo sapiens filamin A interacting protein 1 (FILIP1), transcript variant 1, mRNA	HUMAN	26.3	26.3	68%	96	100.00%	4802	NM_001289987.3
Select seq NM_001287059.2	Homo sapiens HECT, C2 and WW domain containing E3 ubiquitin protein ligase 1 (HECW1), transcript variant 2, mRNA	HUMAN	26.3	26.3	68%	96	100.00%	9116	NM_001287059.2
Select seq NM_006289.4	Homo sapiens talin 1 (TLN1), mRNA	HUMAN	26.3	26.3	68%	96	100.00%	8623	NM_006289.4
Select seq NM_018227.6	Homo sapiens ubiquitin like modifier activating enzyme 6 (UBA6), mRNA	HUMAN	26.3	26.3	68%	96	100.00%	9540	NM_018227.6

The sequences of these genes were examined further for the presence of a PAM site downstream, which could lead to off target binding of the CRISPR/Cas9. Several PAM sites were present around the target area (Fig 3.7).

```
NM_001370.2    TCCACAGGCCGTGCAGCCACACTTAAGGAAATGCTTCGACTCCATTTCAAAGCTCGAATT    3780
SPOCK1         ------------------------------------------------TCGAATT       7
                                                               *******

NM_001370.2    TGCTCTCATGCCTCCTGCCGAAGGAAAGATTCCTGGTATTGATGGAGAACCAGAAAAGGT    3840
SPOCK1         TGGTCAAGTGCA------------------------------------------------   19
               ** **    ***

gRNA           ------------------------------------------------------------  0
TNPO1          CAGTGCCTTTGCTACCCTAGAAGAGGAGGCTTGTACAGAACTTGTTCCTTACCTTGCTTA  1680

gRNA           --------------------TCGAATTTGGTC----------------------------  12
TNPO1          TATACTTGATACCCTGGTCTTTGCATTTAGTAAATACCAGCATAAGAACCTGCTCATTCT  1740
                               * * **** **

(CALM1)V3      ACCACGAACCCCTCAGCATACTGGGAATCTCTTCCTGAACAACGAATGTAAATTTGGTCA  3660
gRNA           ------------------------------------------TCGAATTTGGTCA        13
                                                         **********

(CALM1)V3      AGTCTACTCTTCCGTTCATTCAATTATTTTAAGCATTTGAATTATTTATTGTATATCCTA  3720
gRNA           AGTGCA------------------------------------------------------  19
               ***   *

gRNA           ------------------------------------------------------------  0
AGTPBP1)V2     GACTTAATGGTACAGATTCATTCTATTCTTGCAAAGATTGGACCAAAAGATAAAAAATTT  660

gRNA           ---------------------------------------TCGAATTTGGTCAAGTGCA--  19
AGTPBP1)V2     GGAGTAAAGGCTAGAATTAATGGGGCTCTGAATATAACCCTGAATTTGGTCAAGCAGAAT  720
                                                      *************    *

(SMIM3),       AAAGCAGCAGGAGGGACTTTGGGGCATGGACCTGAGTTCTGGTTTTGATTCTGCCACGAG  540
gRNA           ------------------------------------------------------------  0

(SMIM3),       CCAGCTGTGTGAATTTGGTCAAGGGACCTAACTCTCTGAGTTCCAGGTTCCTTATCTTTC  600
gRNA           --------TCGAATTTGGTCAAGTGCA----------------------------------  19
                       * ************* *

(SYNM)VB,      GATTTTGTTTTAGCTGTAACAGGTAATGGTTTTTGGATAGATGATTGACTGGTGAGAATT  5820
gRNA           ----------------------------------------------------TCGAATT    7
                                                                   *****

(SYNM)VB,      TGGTCAAGGTGACAGCCTCCTGTCTGATGACAGGACAGACTGGTGGTGAGGAGTCTAAGT  5880
gRNA           TGGTCAAGTGCA------------------------------------------------  19
               ********    *

(FILIP1)V1,    GCTGAGGGAAGAAGAAGAGAAGCTCAAAGCCATTACTTCCAAATCCAAAGAAGACAGACA  1440
gRNA           ----------------------------------------------------------TC   2

(FILIP1)V1,    GAAATTGCTCAAGTTAGAAGTGGACTTTGAACACAAGGCTTCGAGGTTTTCTCAAGAGCA  1500
gRNA           GAATTTGGTCAAG----------------------------------------TGCA     19
               *** *** *****                                           ***

(HECW1)V2      GAACTCCCCAGGTTTACAGAGAGCCAGTGCAAGAGCCCCTTCCCCCTACCGAAGAGACTT  3900
gRNA           ------------------------------------------------------------  0

(HECW1)V2      TGAGGCCAAGCTCCGCAATTTCTACAGAAAACTGGAAGCCAAAGGATTTGGTCAGGGTCC  3960
gRNA           -------------------------------------TCGAATTTGGTCAAGT---      16
                                                    * ********* *

(TLN1),        CAAGGAGGTGGCCAACAGCACAGCTAATCTTGTCAAGACCATCAAGGCGCTAGATGGGGC  4800
gRNA           ----------------------TCGAATTTGGTCAAGTGCA-------------------  19
                                     * *** * ****** **

(UBA6),        CATGGAAATTTGGTCAAGGTTATTTTGTGATTTCGGTGATGAATTTGAAGTTTTAGATACA  660
gRNA           --TCGAATTTGGTCAAGTGCA---------------------------------------  19
                  * ************* *
                    ____
```

	Homo sapiens dynein axonemal heavy chain 6 (DNAH6), mRNA
	Homo sapiens transportin 1 (TNPO1), transcript variant 4, mRNA
	Homo sapiens calmodulin 1 (CALM1), transcript variant 1, mRNA
	Homo sapiens ATP/GTP binding carboxypeptidase 1 (AGTPBP1), transcript variant 2, mRNA
	Homo sapiens small integral membrane protein 3 (SMIM3), mRNA
	Homo sapiens synemin (SYNM), transcript variant B, mRNA
	Homo sapiens filamin A interacting protein 1 (FILIP1), transcript variant 1, mRNA
	Homo sapiens HECT, C2 and WW domain containing E3 ubiquitin protein ligase 1 (HECW1), transcript variant 2, mRNA
	Homo sapiens talin 1 (TLN1), mRNA
	Homo sapiens ubiquitin like modifier activating enzyme 6 (UBA6), mRNA

Figure 3.8: Ten genes with sequence similarity to *SPOCK1* have PAM sites near the target region

3.7 SPOCK1 protein interacts with multiple ECM components

Currently, little to no experimental evidence exists to show direct binding partners of the SPOCK1 protein. A solid phase binding assay was designed to test binding of the SPOCK1 protein to fibronectin (FN), fibroblast growth factor (FGF) and collagen (COLL) (Section 2.4). A single experiment was also tested with MT1-MMP for binding. SPOCK1 recombinant protein was used in the ELISA as a positive control to confirm binding of the SPOCK1 antibody. SPOCK1 bound significantly to all proteins tested (Fig. 3.9, $P<0.003$ for all, 2-way ANOVA).
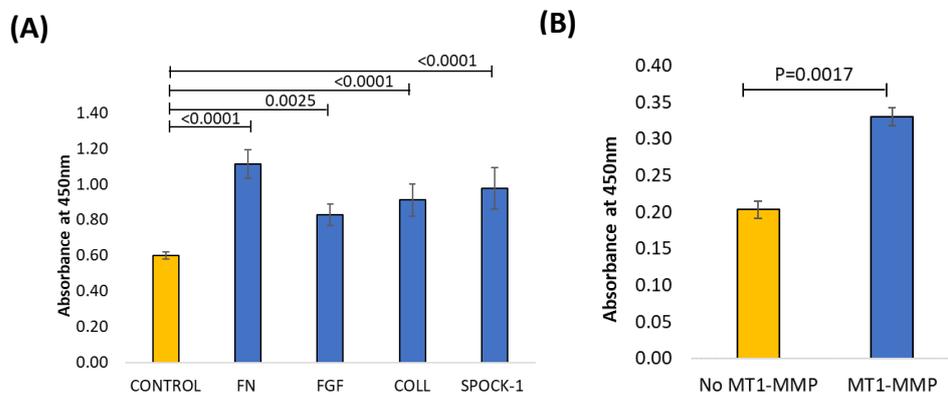
Figure 3.9: SPOCK1 binds directly to ECM proteins, FGF and MT1-MMP. An untreated 96 well plate was coated with FN, FGF, COLL, SPOCK1 and MT1-MMP, washed, blocked and incubated with a SPOCK1 binding solution. After washing, the wells were incubated with a SPOCK1 antibody, washed and an HRP-linked secondary antibody. TMB solution was used for detection. Graphs show absorbance at 450nm± SEM. N=9, from 3 independent experiments **(A)** and n=3 from 1 independent experiment **(B)**. Statistics are from a two-way ANOVA.

4 Discussion

Phenotypical analysis of the CRISPR/Cas9 transfected cell lines showed differences between the *SPOCK1* sgRNA targeted and the control non-targeted cell lines. An initial T7E1 mismatch detection assay used to determine CRISPR/Cas9 efficiency indicated potential success of the CRISPR/Cas9 transfection due to the presence of cleaved band of the expected product size when run on an agarose gel. To confirm this result, a western blot was performed on the cell lysate to detect SPOCK1 protein however, no bands were detected for any sample or the human recombinant SPOCK1 protein positive control. This failure is likely due to a fault of the antibody as currently, no one has managed to discover a functional SPOCK1 antibody for western blot. Due to time constraints, the functional experiments in this study were conducted prior to receiving the results from the samples sent for sequencing.

Unfortunately, despite the relative ease of conducting a T7E1 assay, the assay can only detect small indels and cannot identify specific mutations. Furthermore, it is not sensitive enough to detect single nucleotide polymorphisms (SNPs) (Zhang *et al.*, 2015). Once the sequencing results were obtained, Clustal Omega analysis revealed the absence of mutations in the *SPOCK1* targeted region in CRISPR/Cas9 transfected cells. It has been reported that, off-target effects of CRISPR/Cas9 are frequent (Naeem *et al.*, 2020; Zhang *et al.*, 2015). This could explain the phenotypic differences that was observed between the CRISPR/Cas9 transfected cells and control cells in the growth and adhesion assays despite the lack of evidence of mutations in the *SPOCK1* target region. Important factors for CRISPR/Cas9 efficiency, according to these studies, include the PAM site nucleotides to be NGG or NRG (R=A or G), the sgRNA to be maximum of 17 nucleotides and the GC content to be more than 50% due to its stabilizing effect on RNA. These criteria were not all met for the SPOCK1 sgRNA used in this project. In particular, the predesigned sgRNA was 19 nucleotides long, that has been previously linked to lower efficiency whilst the GC percentage was also shown to be reduced. The use of CRISPR/Cas9 technology on PS-1 cells could be repeated with these criteria in mind.

The high occurrence of off-target effects could lead scientists not only to further optimise protocols to improve CRISPR/Cas9 transfection efficiency, but also to attempt to locate the genes causing the off-target effects. The detection of CRIPSR/Cas9 efficiency or percentage of mutations present is determined by software usually by the quantification of indels. Indels are insertions and/or deletions of nucleotides in genomic DNA. The off-target localisation includes assays and software based on different premises. Some techniques are proven overall more accurate than others. For example, web-based algorithms and CHIP-seq are being employed despite not being highly sensitive and deep sequencing methods are a widely known and have better sensitivity (Zhang *et al*., 2015). A recent method of locating off-target effects is genome-wide mapping of double stranded breaks (DSB) caused by Cas9 and other nucleases known as Breaks Labelling, Enrichments on Streptavidin and next generation Sequencing (BLESS). This technique is preferred as its more versatile, quantitative, and more sensitive than the other techniques (Naeem *et al*., 2020). Each method carries its advantages and disadvantages and unfortunately there is no conclusive method to locate off-target effects efficiently and accurately.

In this study, the CRISPR/Cas9 induced mutation(s) in unknown genes inhibited stromal cell growth and proliferation as observed in a heterogenous PS-1 cell line population and in two of the four single cell PS-1 colonies tested, (SC1 and SC4). Single colony 2 (SC2) and SC3, however only displayed a similar growth inhibitive pattern upon CRISPR/Cas9 transfection but did not achieve a statistically significant difference in growth between the control and test cells at day 3 and 4 of culture as observed in SC1 and SC4. This could be due to the off-target mutations only being present in some of the clones.

In the heterogenous CRISPR/Cas9 transfected PS-1 cell population, SPOCK1 KD cells promoted cell adhesion (Fig. 3.4 C&D). This effect however was not observed with the four single cell colonies (Fig. 3.5C). Perhaps the unknown proteins' ablation in PS-1 cells only leads to very subtle de-adhesiveness that is apparent in the mixed population but not in the single cell colonies. It may also be possible that protein does not affect stromal cell adhesion. However, it would be important to study the effect on adhesion in more detail within the first 24hrs of cell plating to reach a conclusive outcome.

Interestingly, we found that the effect of the SPOCK1 KD PDAC cell viability was contradictory. A preliminary viability assay using Alamar blue indicated that CRISPR/Cas9 off-target mutated proteins might decrease viability of Panc-1 cells but improved viability of AsPC-1 cells (n=12 from 1 independent experiment). Upon further investigation, these CRISPR/Cas9 off-target mutations had little effect on Panc-1 cell confluence with possible growth inhibitive effects beginning to show at day 4 compared to non-targeted control cells however this was not significant (p=0.204). As AsPC-1 cells were not tested further than the initial experiment, it would be necessary to repeat the growth curve with this cell type as well other pancreatic cancer cell lines to confirm cell specific effects of these proteins. Importantly, it would also be necessary to test to for *SPOCK1* mutations in these cell lines.

In Panc-1 cells, the CRISPR/Cas9 off target unknown proteins had an overall de-adhesive effect (Fig. 3.7). However, upon closer inspection of the first 24 hours of culture, differences in cell eccentricity in the SPOCK1 KD and control populations are lost for a period of few hours around the 12hr culture mark. This may be due to the potential effects of the induction of proliferation which begins shortly after the process of cell adhesion. The population of cells with the gene ablation do however progress to show increased cell spread compared to the controls suggesting an overall de-adhesive role of the unknown protein in PDAC cells. This data however is preliminary and further testing could be done using more single cell colonies in different pancreatic cancer cell lines.

Our collective proliferation data in AsPC-1, Panc-1 and PS-1 cells, while preliminary, could suggest a cell specific role of the unknown protein in cell proliferation as well as potential roles in metastasis in PDAC as de-adhesion is an essential prior step of this process.

Although the exact gene that was mutated by CRISPR/Cas9 in this study has not been determined, a search for sequences with high sgRNA similarity to SPOCK1 revealed 10 genes with a high chance off off-target binding. Some among these such as *TLN-1*, *SYNM* and *HECW1* are known to modulate cell processes such as adhesion and migration which we have observed in this study.

Table 4.1: The function of potential off-target genes/proteins with sequence similarity to SPOCK1

Gene (Homo sapiens)	Function of encoded protein
Dynein axonemal heavy chain 6 (DNAH6)	▪ Encodes part of the microtubule-associated motor protein complex ▪ Important for ciliary function ▪ Mutations may cause primary ciliary dyskinesia (PCD) and heterotaxy [provided by RefSeq, Jun 2016]
Transportin 1 (TNPO1), TV-4	▪ Encodes β-subunit of the karyopherin receptor complex ▪ Important for nuclear localization of proteins [provided by RefSeq, Jun 2018]
Homo sapiens calmodulin 1 (CALM1), TV-1	▪ Encodes calmodulin protein (EF-hand calcium-binding protein) ▪ Calmodulin activation regulates ion channels in heart ▪ Mutation may cause ventricular tachycardia [provided by RefSeq, May 2020]
ATP/GTP binding carboxypeptidase 1 (AGTPBP1), TV-2	▪ Zinc carboxypeptidase, NNA1, contains an ATP/GTP binding motif ▪ Important for nuclear localization of proteins [provided by OMIM, Jul 2002].
Small integral membrane protein 3 (SMIM3)	▪ Prediction: component of cell membrane [provided by Alliance of Genome Resources, Apr 2022]
Synemin (SYNM), TV-B	▪ Encodes cytoskeletal proteins (intermediate filament family member) ▪ Important for cells to resist mechanical stress and muscle structural support ▪ Interacts with desmin and ECM [provided by RefSeq, Jul 2008].
Filamin A interacting protein 1 (FILIP1), TV-1	▪ Encodes protein which bind and degrades filamin A ▪ May have function in cortical neuron migration and dendritic spine morphology [provided by RefSeq, Oct 2016].
HECT, C2 and WW domain containing E3 ubiquitin protein ligase 1 (HECW1), TV-2	▪ Encodes protein which causes ubiquitination of Smad4 ▪ Promotes non-small cell lung cancer cells metastasis (Lu *et al.*, 2021)
Talin 1 (TLN1)	▪ Encodes cell cytoskeletal protein ▪ Important for actin filament assembly and cell adhesion, spread and migration [provided by RefSeq, Feb 2009].
Ubiquitin like modifier activating enzyme 6 (UBA6)	▪ Important for embryonic development and may have role in spermatogenesis [provided by Uniprot]
Information acquired from NCBI database. TV: transcript variant	

BLESS could this be in full again could be carried out to identify the off-target mutation and further research could be conducted into those mutated gene with its potential effects on stromal and PDAC cell function however, the main purpose of this research was to study SPOCK1.

As mentioned in the introduction, SPOCK1 is a matricellular protein with various functions. *SPOCK1* overexpression in the tumour stroma correlated with increased tumour invasiveness and poor prognosis (Veenstra *et al*., 2017). Activated stromal cells produce proteins such as PDGF and TGF-β which in turn upregulate SPOCK1 expression perpetuating a vicious cycle (Du *et al.,* 2020). reported that SPOCK-1 directly promoted proliferation and metastasis in PDAC cell lines, PCNA-1 and MIA PaCa-2 (Li *et al*., 2019).

While there are few functional studies on the role of SPOCK1 PDAC, there is limited information on its mechanisms of action. For the first time, we have shown a direct interaction between SPOCK1 and collagen, fibronectin, MT1-MMP and FGF. The diverse binding capacity of SPOCK1 to key ECM proteins such as collagen and fibronectin, the MT1-MMP enzyme, and the FGF growth factor suggest the protein plays a central in the ECM and cell signalling. MMPs are critical players in ECM degradation and turnover and can promote cancer cell migration. It has been shown that SPOCK1 can inhibit pro-MMP2 via MT1-MMP with SPOCK2 negating this effect (Nakada *et al.,* 2003; Nakada *et al*. 2001). Together, not only does SPOCK1 interact with a vast number of proteins, it also seem to co-ordinate its function in relation to other SPARC family members such as SPOCK2. The SPARC family member, hevin, whose cleavage by MMP3 produces a SPARC-like fragment that could potentially act in the place of SPARC (Weaver *et al*., 2011). In turn, the requirement to study such proteins in more complex systems where all the interacting components are present is imperative (Viloria *et al*., 2016).

As mentioned previously, the major limitation of this study was lack of mutation in the CRISPR/Cas9 transfected cell lines, which limited the research on the study into its function. Other than improving the efficacy of the CRISPR/Cas9 transfection, other ways to study the protein function could be to silence SPOCK1 using siRNA or shRNA or treating cells with recombinant or purified SPOCK1 protein. Unpublished work form our laboratory has reported that SPOCK1 expression in PDAC cells could be absent or present. Therefore, it would be important to first test expression level of SPOCK1 as SPOCK1 may have cell

specific effects. Since the studies of SPOCK1 in PDAC are very limited, there are many possibilities for future areas of research that could be pursued. Since the SPARC family is expected to modulate cell adhesion, cell eccentricity analysis could be could be combined with FAK staining to determine mechanisms of action for cell adhesion. Since SPOCK KD may have affected PS1 cell growth and adhesion, the effects of SPOCK1 on other stromal cell characteristics could be investigated such as activation as quiescent stromal cells can be activated and recruited by PDAC. Since SPOCK1 overexpression in PDAC tissue is correlated with increased invasiveness and poor prognosis (Veenstra *et al.,* 2017), a Matrigel trans well assay could be used to measure invasion and migration in a controlled and isolated environment. Importantly, the direct interaction of SPOCK1 with multiple ECM components as observed in this study suggest it has diverse roles in the complex tumour-stroma and more complex 2D and 3D co-cultures may be required to fully understand its function. After completing functional studies, inhibition of these interactions using small molecule inhibitor or monoclonal antibodies to reverse any negative effects on PDAC cell function could form the foundation for novel PDAC therapies.

5 Conclusions

Analysis of sequencing of the CRISPR/Cas9 transfected cell lines did not show any mutation in the SPOCK1 region. However, functional assays showed effects on cell growth and adhesion suggesting off-target binding of Cas9 to the sgRNA. Ten possible off-target gene were identified with sgRNA sequence similarity to SPOCK1. For the first time, direct binding between SPOCK1 and multiple ECM components have been shown in this study, suggesting SPOCK1 protein has diverse roles in the PDAC ECM. Further research could focus on evaluating specific interactions of SPOCK1 with ECM components and their combined effect on PDAC cell function.

6 References

Alliel, P.M., Perin, J.P., Jollès, P. and Bonnet, F.J., 1993. Testican, a multidomain testicular proteoglycan resembling modulators of cell social behaviour. *European journal of biochemistry*, *214*(1), pp.347-350.

Bekkali, N.L. and Oppong, K.W., 2017. Pancreatic ductal adenocarcinoma epidemiology and risk assessment: Could we prevent? Possibility for an early diagnosis. *Endoscopic ultrasound*, *6*(Suppl 3), p.S58.

Bonnet, F., Perin, J.P., Maillet, P., Jolles, P. and Alliel, P.M., 1992. Characterization of a human seminal plasma glycosaminoglycan-bearing polypeptide. *Biochemical Journal*, *288*(2), pp.565-569.

Borad, M.J., Reddy, S.G., Bahary, N., Uronis, H.E., Sigal, D., Cohn, A.L., Schelman, W.R., Stephenson Jr, J., Chiorean, E.G., Rosen, P.J. and Ulrich, B., 2015. Randomized phase II trial of gemcitabine plus TH-302 versus gemcitabine in patients with advanced pancreatic cancer. *Journal of Clinical Oncology*, *33*(13), p.1475.

Bornstein, P., 2009. Matricellular proteins: an overview. *Journal of cell communication and signaling*, *3*(3), pp.163-165.

Bella, R., Kaminski, R., Mancuso, P., Young, W.B., Chen, C., Sariyer, R., Fischer, T., Amini, S., Ferrante, P., Jacobson, J.M. and Kashanchi, F., 2018. Removal of HIV DNA by CRISPR from patient blood engrafts in humanized mice. *Molecular Therapy-Nucleic Acids*, *12*, pp.275-282.

Brellier, F., Ruggiero, S., Zwolanek, D., Martina, E., Hess, D., Brown-Luedi, M., Hartmann, U., Koch, M., Merlo, A., Lino, M. and Chiquet-Ehrismann, R., 2011. SMOC1 is a tenascin-C interacting protein over-expressed in brain tumors. *Matrix Biology*, *30*(3), pp.225-233.

Chaly Y, Blair HC, Smith SM, Bushnell DS, Marinov AD, Campfield BT, Hirsch R. Follistatin-like protein 1 regulates chondrocyte proliferation and chondrogenic differentiation of mesenchymal stem cells. Annals of the rheumatic diseases. 2015 Jul 1;74(7):1467-73.

Chandrasekaran, V., Ambati, J., Ambati, B.K. and Taylor, E.W., 2007. Molecular docking and analysis of interactions between vascular endothelial growth factor (VEGF) and SPARC protein. *Journal of Molecular Graphics and Modelling*, *26*(4), pp.775-782.

Choueiri, T.K., Mayer, E.L., Je, Y., Rosenberg, J.E., Nguyen, P.L., Azzi, G.R., Bellmunt, J., Burstein, H.J. and Schutz, F.A., 2011. Congestive heart failure risk in patients with breast cancer treated with bevacizumab. In *Database of Abstracts of Reviews of Effects (DARE): Quality-assessed Reviews [Internet]*. Centre for Reviews and Dissemination (UK).

Crane, C.H., Winter, K., Regine, W.F., Safran, H., Rich, T.A., Curran, W., Wolff, R.A. and Willett, C.G., 2009. Phase II study of bevacizumab with concurrent capecitabine and radiation followed by

maintenance gemcitabine and bevacizumab for locally advanced pancreatic cancer: Radiation Therapy Oncology Group RTOG 0411. *Journal of Clinical Oncology*, *27*(25), p.4096.

Cui, X., Wang, Y., Lan, W., Wang, S., Cui, Y., Zhang, X., Lin, Z. and Piao, J., 2022. SPOCK1 promotes metastasis in pancreatic cancer via NF-κB-dependent epithelial-mesenchymal transition by interacting with IκB-α. *Cellular Oncology*, *45*(1), pp.69-84.

Dreieicher, E., Beck, K.F., Lazaroski, S., Boosen, M., Tsalastra-Greul, W., Beck, M., Fleming, I., Schaefer, L. and Pfeilschifter, J., 2009. Nitric oxide inhibits glomerular TGF-β signaling via SMOC-1. *Journal of the American Society of Nephrology*, *20*(9), pp.1963-1974.

Du, Z., Lin, Z., Wang, Z., Liu, D., Tian, D. and Xia, L., 2020. SPOCK1 overexpression induced by platelet-derived growth factor-BB promotes hepatic stellate cell activation and liver fibrosis through the integrin α5β1/PI3K/Akt signaling pathway. *Laboratory Investigation*, *100*(8), pp.1042-1056.

Eichelberg, C., Vervenne, W.L., De Santis, M., von Weikersthal, L.F., Goebell, P.J., Lerchenmüller, C., Zimmermann, U., Bos, M.M., Freier, W., Schirrmacher-Memmel, S. and Staehler, M., 2015. SWITCH: a randomised, sequential, open-label study to evaluate the efficacy and safety of sorafenib-sunitinib versus sunitinib-sorafenib in the treatment of metastatic renal cell cancer. *European urology*, *68*(5), pp.837-847.

Eijken, M., Swagemakers, S., Koedam, M., Steenbergen, C., Derkx, P., Uitterlinden, A.G., van der Spek, P.J., Visser, J.A., de Jong, F.H., Pols, H.A. and van Leeuwen, J.P., 2007. The activin A-follistatin system: potent regulator of human extracellular matrix mineralization. *The FASEB Journal*, *21*(11), pp.2949-2960.

Erkan, M., Adler, G., Apte, M.V., Bachem, M.G., Buchholz, M., Detlefsen, S., Esposito, I., Friess, H., Gress, T.M., Habisch, H.J. and Hwang, R.F., 2012. StellaTUM: current consensus and discussion on pancreatic stellate cell research. *Gut*, *61*(2), pp.172-178.

Esposito, I., Kayed, H., Keleg, S., Giese, T., Sage, E.H., Schirmacher, P., Friess, H. and Kleeff, J., 2007. Tumor-suppressor function of SPARC-like protein 1/Hevin in pancreatic cancer. *Neoplasia*, *9*(1), pp.8-17.

Fan, L.C., Jeng, Y.M., Lu, Y.T. and Lien, H.C., 2016. SPOCK1 is a novel transforming growth factor-β–induced myoepithelial marker that enhances invasion and correlates with poor prognosis in breast cancer. *PLoS One*, *11*(9), p.e0162933.

Fang, H. and DeClerck, Y.A., 2013. Targeting the tumor microenvironment: from understanding pathways to effective clinical trials. *Cancer research*, *73*(16), pp.4965-4977.

Francki, A., McClure, T.D., Brekken, R.A., Motamed, K., Murri, C., Wang, T. and Sage, E.H., 2004. SPARC regulates TGF-beta1-dependent signaling in primary glomerular mesangial cells. *Journal of cellular biochemistry*, *91*(5), pp.915-925.

Friess, H., Langrehr, J.M., Oettle, H., Raedle, J., Niedergethmann, M., Dittrich, C., Hossfeld, D.K., Stöger, H., Neyns, B., Herzog, P. and Piedbois, P., 2006. A randomized multi-center phase II trial

of the angiogenesis inhibitor Cilengitide (EMD 121974) and gemcitabine compared with gemcitabine alone in advanced unresectable pancreatic cancer. *BMC cancer*, *6*(1), pp.1-12.

Froeling, F.E., Mirza, T.A., Feakins, R.M., Seedhar, A., Elia, G., Hart, I.R. and 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. *The American journal of pathology*, *175*(2), pp.636-648.

Garcia, J., Hurwitz, H.I., Sandler, A.B., Miles, D., Coleman, R.L., Deurloo, R. and Chinot, O.L., 2020. Bevacizumab (Avastin®) in cancer treatment: A review of 15 years of clinical experience and future outlook. *Cancer treatment reviews*, *86*, p.102017.

Geng, Y., Dong, Y., Yu, M., Zhang, L., Yan, X., Sun, J., Qiao, L., Geng, H., Nakajima, M., Furuichi, T. and Ikegawa, S., 2011. Follistatin-like 1 (Fstl1) is a bone morphogenetic protein (BMP) 4 signaling antagonist in controlling mouse lung development. *Proceedings of the National Academy of Sciences*, *108*(17), pp.7058-7063.

Gerarduzzi, C., Kumar, R.K., Trivedi, P., Ajay, A.K., Iyer, A., Boswell, S., Hutchinson, J.N., Waikar, S.S. and Vaidya, V.S., 2017. Silencing SMOC2 ameliorates kidney fibrosis by inhibiting fibroblast to myofibroblast transformation. *JCI insight*, *2*(8).

Ghim, M., Pang, K.T., Burnap, S.A., Baig, F., Yin, X., Arshad, M., Mayr, M. and Weinberg, P.D., 2021. Endothelial cells exposed to atheroprotective flow secrete follistatin-like 1 protein which reduces transcytosis and inflammation. *Atherosclerosis*, *333*, pp.56-66.

Girard, J.P. and Springer, T.A., 1996. Modulation of Endothelial Cell Adhesion by Hevin, an Acidic Protein Associated with High Endothelial Venules (∗). *Journal of Biological Chemistry*, *271*(8), pp.4511-4517.

Ghosh, D., Venkataramani, P., Nandi, S. and Bhattacharjee, S., 2019. CRISPR–Cas9 a boon or bane: the bumpy road ahead to cancer therapeutics. *Cancer cell international*, *19*(1), pp.1-10.

Gore, J. and Korc, M., 2014. Pancreatic cancer stroma: friend or foe? *Cancer cell*, *25*(6), pp.711-712.

Han, H.A., Pang, J.K.S. and Soh, B.S., 2020. Mitigating off-target effects in CRISPR/Cas9-mediated in vivo gene editing. *Journal of Molecular Medicine*, *98*(5), pp.615-632.

Hapani, S., Sher, A., Chu, D. and Wu, S., 2010. Increased risk of serious hemorrhage with bevacizumab in cancer patients: a meta-analysis. *Oncology*, *79*(1-2), pp.27-38.

Higashijima, Y., Hirano, S., Nangaku, M. and Nureki, O., 2017. Applications of the CRISPR-Cas9 system in kidney research. *Kidney International*, *92*(2), pp.324-335.

Hu, P.F., Ma, C.Y., Sun, F.F., Chen, W.P. and Wu, L.D., 2019. Follistatin-like protein 1 (FSTL1) promotes chondrocyte expression of matrix metalloproteinase and inflammatory factors via the NF-κB pathway. *Journal of cellular and molecular medicine*, *23*(3), pp.2230-2237.

CHAPTER 6: References

Humphries, J.D., Zha, J., Burns, J., Askari, J.A., Below, C.R., Chastney, M.R., Jones, M.C., Mironov, A., Knight, D., O'Reilly, D.A. and Dunne, M.J., 2022. Pancreatic ductal adenocarcinoma cells employ integrin α6β4 to form hemidesmosomes and regulate cell proliferation. *Matrix Biology*, *110*, pp.16-39.

Hurtado de Mendoza, T., Mose, E.S., Botta, G.P., Braun, G.B., Kotamraju, V.R., French, R.P., Suzuki, K., Miyamura, N., Teesalu, T., Ruoslahti, E. and Lowy, A.M., 2021. Tumor-penetrating therapy for β5 integrin-rich pancreas cancer. *Nature communications*, *12*(1), pp.1-13.

Ilegems, E., Bryzgalova, G., Correia, J., Yesildag, B., Berra, E., Ruas, J.L., Pereira, T.S. and Berggren, P.O., 2022. HIF-1α inhibitor PX-478 preserves pancreatic β cell function in diabetes. *Science Translational Medicine*, *14*(638), p.eaba9112.

Jeong, W., Rapisarda, A., Park, S.R., Kinders, R.J., Chen, A., Melillo, G., Turkbey, B., Steinberg, S.M., Choyke, P., Doroshow, J.H. and Kummar, S., 2014. Pilot trial of EZN-2968, an antisense oligonucleotide inhibitor of hypoxia-inducible factor-1 alpha (HIF-1α), in patients with refractory solid tumors. *Cancer chemotherapy and pharmacology*, *73*(2), pp.343-348.

Jin, S., Lee, W.C., Aust, D., Pilarsky, C. and Cordes, N., 2019. β8 Integrin Mediates Pancreatic Cancer Cell Radiochemoresistanceβ8 Integrin in Pancreatic Cancer Cells. *Molecular Cancer Research*, *17*(10), pp.2126-2138.

Jung, J.H., Sosnowska, D., Weaver, J., Parson, H.K., Casellini, C.M. and Vinik, A.I., 2020. Expression of Hypoxia-Inducible Factors in Different Stages of Pancreatic Tumor Progression. *Reports*, *3*(4), p.30.

Kaleağasıoğlu, F. and Berger, M.R., 2014. SIBLINGs and SPARC families: their emerging roles in pancreatic cancer. *World journal of gastroenterology: WJG*, *20*(40), p.14747.

Kamihagi, K., Katayama, M., Ouchi, R. and Kato, I., 1994. Osteonectin/SPARC regulates cellular secretion rates of fibronectin and laminin extracellular matrix proteins. *Biochemical and biophysical research communications*, *200*(1), pp.423-428.

Kümler, I., Christiansen, O.G. and Nielsen, D.L., 2014. A systematic review of bevacizumab efficacy in breast cancer. *Cancer treatment reviews*, *40*(8), pp.960-973.

Lai, B.F.L., Lu, R.X.Z., Hu, Y., Davenport Huyer, L., Dou, W., Wang, E.Y., Radulovich, N., Tsao, M.S., Sun, Y. and Radisic, M., 2020. Recapitulating Pancreatic Tumor Microenvironment through Synergistic Use of Patient Organoids and Organ-on-a-Chip Vasculature. *Advanced functional materials*, *30*(48), p.2000545.

Li, J., Ke, J., Fang, J. and Chen, J.P., 2020. A potential prognostic marker and therapeutic target: SPOCK1 promotes the proliferation, metastasis, and apoptosis of pancreatic ductal adenocarcinoma cells. *Journal of cellular biochemistry*, *121*(1), pp.743-754.

CHAPTER 6: References

Lu, C., Ning, G., Si, P., Zhang, C., Liu, W., Ge, W., Cui, K., Zhang, R. and Ge, S., 2021. E3 ubiquitin ligase HECW1 promotes the metastasis of non-small cell lung cancer cells through mediating the ubiquitination of Smad4. *Biochemistry and Cell Biology*, *99*(5), pp.675-681.

Maier, S., Paulsson, M. and Hartmann, U., 2008. The widely expressed extracellular matrix protein SMOC-2 promotes keratinocyte attachment and migration. *Experimental cell research*, *314*(13), pp.2477-2487.

Marr, H.S. and Edgell, C.J.S., 2003. Testican-1 inhibits attachment of Neuro-2a cells. *Matrix biology*, *22*(3), pp.259-266.

Mitchem, J.B., Brennan, D.J., Knolhoff, B.L., Belt, B.A., Zhu, Y., Sanford, D.E., Belaygorod, L., Carpenter, D., Collins, L., Piwnica-Worms, D. and Hewitt, S., 2013. Targeting Tumor-Infiltrating Macrophages Decreases Tumor-Initiating Cells, Relieves Immunosuppression, and Improves Chemotherapeutic ResponsesMacrophages Regulate Tumor-Initiating Cells. *Cancer research*, *73*(3), pp.1128-1141.

Mohanty, S., Dash, A. and Pradhan, C.K., 2019. CRISPR-Cas9 Technology: A magical tool for DNA editing. *Cellular Physiology*, *1*(1), pp.1-11.

Morrissey, M.A., Jayadev, R., Miley, G.R., Blebea, C.A., Chi, Q., Ihara, S. and Sherwood, D.R., 2016. SPARC promotes cell invasion in vivo by decreasing type IV collagen levels in the basement membrane. *PLoS Genetics*, *12*(2), p.e1005905.

Munasinghe, A., Malik, K., Mohamedi, F., Moaraf, S., Kocher, H., Jones, L. and Hill, N.J., 2020. Fibronectin acts as a molecular switch to determine SPARC function in pancreatic cancer. *Cancer Letters*, *477*, pp.88-96.

Murakawa, M., Aoyama, T., Miyagi, Y., Kobayashi, S., Ueno, M., Morimoto, M., Numata, M., Yamamoto, N., Tamagawa, H., Yukawa, N. and Rino, Y., 2019. The impact of SPARC expression on the survival of pancreatic ductal adenocarcinoma patients after curative resection. *Journal of Cancer*, *10*(3), p.627.

Murphy-Ullrich, J.E., Lane, T.F., Pallero, M.A. and Sage, E.H., 1995. SPARC mediates focal adhesion disassembly in endothelial cells through a follistatin-like region and the Ca2+-binding EF-hand. *Journal of cellular biochemistry*, *57*(2), pp.341-350.

Murphy-Ullrich, J.E. and Sage, E.H., 2014. Revisiting the matricellular concept. *Matrix Biology*, *37*, pp.1-14.

Naeem, M., Majeed, S., Hoque, M.Z. and Ahmad, I., 2020. Latest developed strategies to minimize the off-target effects in CRISPR-Cas-mediated genome editing. *Cells,* 9(7), p.1608.

Nakada, M., Miyamori, H., Yamashita, J. and Sato, H., 2003. Testican 2 abrogates inhibition of membrane-type matrix metalloproteinases by other testican family proteins. *Cancer Research*, *63*(12), pp.3364-3369.

CHAPTER 6: References

Nakada, M., Yamada, A., Takino, T., Miyamori, H., Takahashi, T., Yamashita, J. and Sato, H., 2001. Suppression of membrane-type 1 matrix metalloproteinase (MMP)-mediated MMP-2 activation and tumor invasion by testican 3 and its splicing variant gene product, N-Tes. *Cancer research*, *61*(24), pp.8896-8902.

Novinec, M., Kovačič, L., Škrlj, N., Turk, V. and Lenarčič, B., 2008. Recombinant human SMOCs produced by in vitro refolding: calcium-binding properties and interactions with serum proteins. *Protein expression and purification*, *62*(1), pp.75-82.

Niu, R., Nie, Z.T., Liu, L., Chang, Y.W., Shen, J.Q., Chen, Q., Dong, L.J. and Hu, B.J., 2021. Follistatin-like protein 1 functions as a potential target of gene therapy in proliferative diabetic retinopathy. *Aging (Albany NY)*, *13*(6), p.8643.

Okura, Y., Imao, T., Murashima, S., Shibata, H., Kamikavwa, A., Okamatsu-Ogura, Y., Saito, M. and Kimura, K., 2019. Interaction of nerve growth factor β with adiponectin and SPARC oppositely modulates its biological activity. *International journal of molecular sciences*, *20*(7), p.1541.

Orth, M., Metzger, P., Gerum, S., Mayerle, J., Schneider, G., Belka, C., Schnurr, M. and Lauber, K., 2019. Pancreatic ductal adenocarcinoma: Biological hallmarks, current status, and future perspectives of combined modality treatment approaches. *Radiation Oncology*, *14*(1), pp.1-20.

Pozas, J., San Román, M., Alonso-Gordoa, T., Pozas, M., Caracuel, L., Carrato, A. and Molina-Cerrillo, J., 2019. Targeting angiogenesis in pancreatic neuroendocrine tumors: resistance mechanisms. *International journal of molecular sciences*, *20*(19), p.4949.

Provenzano, P.P., Cuevas, C., Chang, A.E., Goel, V.K., Von Hoff, D.D. and Hingorani, S.R., 2012. Enzymatic targeting of the stroma ablates physical barriers to treatment of pancreatic ductal adenocarcinoma. *Cancer cell*, *21*(3), pp.418-429.

Rainger, J., van Beusekom, E., Ramsay, J.K., McKie, L., Al-Gazali, L., Pallotta, R., Saponari, A., Branney, P., Fisher, M., Morrison, H. and Bicknell, L., 2011. Loss of the BMP antagonist, SMOC-1, causes Ophthalmo-acromelic (Waardenburg Anophthalmia) syndrome in humans and mice. *PLoS genetics*, *7*(7), p.e1002114.

Rawla, P., Sunkara, T. and Gaduputi, V., 2019. Epidemiology of pancreatic cancer: global trends, etiology and risk factors. *World journal of oncology*, *10*(1), p.10.

Reader, C.S., Vallath, S., Steele, C.W., Haider, S., Brentnall, A., Desai, A., Moore, K.M., Jamieson, N.B., Chang, D., Bailey, P. and Scarpa, A., 2019. The integrin αvβ6 drives pancreatic cancer through diverse mechanisms and represents an effective target for therapy. *The Journal of pathology*, *249*(3), pp.332-342.

Redman, M., King, A., Watson, C. and King, D., 2016. What is CRISPR/Cas9? *Archives of Disease in Childhood-Education and Practice*, *101*(4), pp.213-215.

CHAPTER 6: References

Rentz, T.J., Poobalarahi, F., Bornstein, P., Sage, E.H. and Bradshaw, A.D., 2007. SPARC regulates processing of procollagen I and collagen fibrillogenesis in dermal fibroblasts. *Journal of Biological Chemistry*, *282*(30), pp.22062-22071.

Riffle, S. and Hegde, R.S., 2017. Modeling tumor cell adaptations to hypoxia in multicellular tumor spheroids. *Journal of Experimental & Clinical Cancer Research*, *36*(1), pp.1-10.

Rocnik, E.F., Liu, P., Sato, K., Walsh, K. and Vaziri, C., 2006. The novel SPARC family member SMOC-2 potentiates angiogenic growth factor activity. *Journal of Biological Chemistry*, *281*(32), pp.22855-22864.

Schnepp, A., Lindgren, P.K., Hülsmann, H., Kröger, S., Paulsson, M. and Hartmann, U., 2005. Mouse testican-2: expression, glycosylation, and effects on neurite outgrowth. *Journal of Biological Chemistry*, *280*(12), pp.11274-11280.

Shu, Y.J., Weng, H., Ye, Y.Y., Hu, Y.P., Bao, R.F., Cao, Y., Wang, X.A., Zhang, F., Xiang, S.S., Li, H.F. and Wu, X.S., 2015. SPOCK1 as a potential cancer prognostic marker promotes the proliferation and metastasis of gallbladder cancer cells by activating the PI3K/AKT pathway. *Molecular cancer*, *14*(1), pp.1-14.

Simoes, P.K., Olson, S.H., Saldia, A. and Kurtz, R.C., 2017. Epidemiology of pancreatic adenocarcinoma. *Chinese Clinical Oncology*, *6*(3), pp.24-24.

Sternberg, C.N., Hawkins, R.E., Wagstaff, J., Salman, P., Mardiak, J., Barrios, C.H., Zarba, J.J., Gladkov, O.A., Lee, E., Szczylik, C. and McCann, L., 2013. A randomised, double-blind phase III study of pazopanib in patients with advanced and/or metastatic renal cell carcinoma: final overall survival results and safety update. *European journal of cancer*, *49*(6), pp.1287-1296.

Sullivan, M.M., Barker, T.H., Funk, S.E., Karchin, A., Seo, N.S., Hook, M., Sanders, J., Starcher, B., Wight, T.N., Puolakkainen, P. and Sage, E.H., 2006. Matricellular hevin regulates decorin production and collagen assembly. *Journal of biological chemistry*, *281*(37), pp.27621-27632.

Sun, L.R., Li, S.Y., Guo, Q.S., Zhou, W. and Zhang, H.M., 2020. SPOCK1 involvement in epithelial-to-mesenchymal transition: a new target in cancer therapy?. *Cancer management and research*, *12*, p.3561.

Sung, H., Ferlay, J., Siegel, R.L., Laversanne, M., Soerjomataram, I., Jemal, A. and Bray, F., 2021. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA: a cancer journal for clinicians*, *71*(3), pp.209-249.

Thomas, J.T., Chhuy-Hy, L., Andrykovich, K.R. and Moos Jr, M., 2016. SMOC binds to pro-EGF, but does not induce Erk phosphorylation via the EGFR. *Plos one*, *11*(4), p.e0154294.

Valkenburg, K.C., De Groot, A.E. and Pienta, K.J., 2018. Targeting the tumour stroma to improve cancer therapy. *Nature reviews Clinical oncology*, *15*(6), pp.366-381.

Van der Flier, A. and Sonnenberg, A., 2001. Function and interactions of integrins. *Cell and tissue research*, *305*(3), pp.285-298.

CHAPTER 6: References

Váncza, L., Tátrai, P., Reszegi, A., Baghy, K. and Kovalszky, I., 2022. SPOCK1 with unexpected function. The start of a new career. *American Journal of Physiology-Cell Physiology*, *322*(4), pp.C688-C693.

Vannahme, C., Smyth, N., Miosge, N., Gosling, S., Frie, C., Paulsson, M., Maurer, P. and Hartmann, U., 2002. Characterization of SMOC-1, a novel modular calcium-binding protein in basement membranes. *Journal of Biological Chemistry*, *277*(41), pp.37977-37986.

Veenstra, V.L., Damhofer, H., Waasdorp, C., Steins, A., Kocher, H.M., Medema, J.P., van Laarhoven, H.W. and Bijlsma, M.F., 2017. Stromal SPOCK 1 supports invasive pancreatic cancer growth. *Molecular oncology*, *11*(8), pp.1050-1064.

Venkatasubramanian, P.N., 2012. Imaging the pancreatic ECM. *Pancreatic Cancer and Tumor Microenvironment*.

Viloria, K., Munasinghe, A., Asher, S., Bogyere, R., Jones, L. and Hill, N.J., 2016. A holistic approach to dissecting SPARC family protein complexity reveals FSTL-1 as an inhibitor of pancreatic cancer cell growth. *Scientific reports*, *6*(1), pp.1-19.

Walma, D.A.C. and Yamada, K.M., 2020. The extracellular matrix in development. *Development*, *147*(10), p.dev175596.

Wang, Z., Peng, L., Song, Y.L., Xu, S., Hua, Z., Fang, N., Zhai, M., Liu, H., Fang, Q., Deng, T. and Zhang, W., 2018. Pseudo-hemorrhagic region formation in pancreatic neuroendocrine tumors is a result of blood vessel dilation followed by endothelial cell detachment. *Oncology Letters*, *15*(4), pp.4255-4261.

Weaver, M., Workman, G., Schultz, C.R., Lemke, N., Rempel, S.A. and Sage, E.H., 2011. Proteolysis of the matricellular protein hevin by matrix metalloproteinase-3 produces a SPARC-like fragment (SLF) associated with neovasculature in a murine glioma model. *Journal of cellular biochemistry*, *112*(11), pp.3093-3102.

Weniger, M., Honselmann, K.C. and Liss, A.S., 2018. The extracellular matrix and pancreatic cancer: a complex relationship. *Cancers*, *10*(9), p.316.

Wilson, W.R. and Hay, M.P., 2011. Targeting hypoxia in cancer therapy. *Nature Reviews Cancer*, *11*(6), pp.393-410.

Wong, G.S. and Rustgi, A.K., 2013. Matricellular proteins: priming the tumour microenvironment for cancer development and metastasis. *British journal of cancer*, *108*(4), pp.755-761.

Ying, H., Dey, P., Yao, W., Kimmelman, A.C., Draetta, G.F., Maitra, A. and DePinho, R.A., 2016. Genetics and biology of pancreatic ductal adenocarcinoma. *Genes & development*, *30*(4), pp.355-385.

Yang, D., Tang, Y., Fu, H., Xu, J., Hu, Z., Zhang, Y. and Cai, Q., 2018. Integrin β1 promotes gemcitabine resistance in pancreatic cancer through Cdc42 activation of PI3K p110β signaling. *Biochemical and biophysical research communications*, *505*(1), pp.215-221.

Zhang, X.H., Tee, L.Y., Wang, X.G., Huang, Q.S. and Yang, S.H., 2015. Off-target effects in CRISPR/Cas9-mediated genome engineering. *Molecular Therapy-Nucleic Acids*, 4, p.e264.

Zhang, F., Wen, Y. and Guo, X., 2014. CRISPR/Cas9 for genome editing: progress, implications and challenges. *Human molecular genetics*, *23*(R1), pp.R40-R46.

Zhao, T., Ren, H., Jia, L., Chen, J., Xin, W., Yan, F., Li, J., Wang, X., Gao, S., Qian, D. and Huang, C., 2015. Inhibition of HIF-1α by PX-478 enhances the anti-tumor effect of gemcitabine by inducing immunogenic cell death in pancreatic ductal adenocarcinoma. *Oncotarget,* 6(4), p.2250.

Websites

CLUSTAL OMEGA/EMBL-EBI:

https://www.ebi.ac.uk/Tools/msa/clustalo/ [last accessed: 27/06/2022]

GENEWIZ:

https://www.genewiz.com/en-GB [last accessed: 19/06/2022]

NCBI/BLAST/Refseq:

https://www.ncbi.nlm.nih.gov/ [last accessed: 26/06/2022]

https://blast.ncbi.nlm.nih.gov/Blast.cgi [last accessed: 27/06/2022]

https://www.ncbi.nlm.nih.gov/refseq/ [last accessed: 28/06/2022]

UNIPROT:

https://www.uniprot.org/ [last accessed: 28/06/2022]

7 Appendix

7.1 Clustal omega analysis on all single colony samples compared to the original hSPOCK1 sequence

>04570586 (S16 Forward)

```
hSPOCK1     CACATACGGTATTGGGGCTGATTAAGCCATATTCCCTCACCGTTTTTTTTCTTTTTTTTT 60
04570586    ------------------------------------------------------------ 0


hSPOCK1     CTGTGTGTTGTAAAAATCTTCCCATGTCAGTAGTCATGTTTTCAATGGCTACATTGTTTT 120
04570586    ----------NAAAANNNNNNNNNTGTNNGTNNTCNTGTTTTCAATGGCNACATTGTTTT 50
                      ****        ***   **   ** ************* **********

hSPOCK1     CCTCAGTCCATCTGGTTTGTTGAATCTATGCTTCTGGGACAAAATCCTCTGCTTCCTATC 180
04570586    CCNCNNTCCATCTGGTTTGTTGAATCTATGCTTCNGGGACAAAATCCNCNGCTTCCNATC 110
            ** *   ***************************** ************ * ****** ***

hSPOCK1     AATGAGTTGGAGCAAATGGAAGTGCTGGCTAAGGCCCTCAGGATCCCTTGGCCTGGTTGG 240
04570586    AATGAGTTGGAGCAAATGGAAGTGCNGGCTAAGGCCCNCNGGATCCCTTGGCCNGGNNGG 170
            ************************* ********** * ************* **   **

hSPOCK1     ATAGTCCTTTATTTATTTGGCCTTTTATTCTGGGAACTTAGAGGAAAGACTCATAGGGTC 300
04570586    ATANTCCTTTATTTN--------------------------------------------- 185
            *** **********

hSPOCK1     TGGAGCTTCAACTAATTTATCCCCTTTTTTCTCATCTTTAGGCAAAAGAAGGGGAACGTG 360
04570586    ------------------------------------------------------------ 185


hSPOCK1     GCCCAGAAACACTGGGTTGGACCTTCGAATTTGGTCAAGTGCAAGCCCTGTCCCGTGGCA 420
04570586    ------------------------------------------------------------ 185


hSPOCK1     CAGTCAGCCATGGTCTGCGGCTCAGATGGCCACTCCTACACATCCAAGGTTGGTTTTCTT 480
04570586    ------------------------------------------------------------ 185


hSPOCK1     GTCTTTTTATGTTATCAAAAATACTTCATGTCGGCTTGTTCTAACACTTCTCTATGCCAA 540
04570586    ------------------------------------------------------------ 185


hSPOCK1     AACTGGGGCTTG     552
04570586    ------------     185
```

73

>04570587 (S16 Reverse)

```
hSPOCK1              CACATACGGTATTGGGGCTGATTAAGCCATATTCCCTCACCGTTTTTTTTCTTTTTTTTT          60

04570587reversecom   ------------------------------------------------------------           0


hSPOCK1              CTGTGTGTTGTAAAAATCTTCCCATGTCAGTAGTCATGTTTTCAATGGCTACATTGTTTT         120

04570587reversecom   ------------------------------------------------------------           0


hSPOCK1              CCTCAGTCCATCTGGTTTGTTGAATCTATGCTTCTGGGACAAAATCCTCTGCTTCCTATC         180

04570587reversecom   ------------------------------------------------------------           0


hSPOCK1              AATGAGTTGGAGCAAATGGAAGTGCTGGCTAAGGCCCTCAGGATCCCTTGGCCTGGTTGG         240

04570587reversecom   ----------------------------------------------------------GG           2
                                                                               **

hSPOCK1              ATAGTCCTTTATTTATTTGGCCTTTTATTCTGGGAACTTAGAGGAAAGACTCATAGGGTC         300

04570587reversecom   ATNGNNNNTNNTTTATTTGGCCTTTTATTCTGGGAACTTAGAGGAAAGACTCATAGGGTC          62
                     ** *     * *************************************************

hSPOCK1              TGGAGCTTCAACTAATTTATCCCCTTTTTTCTCATCTTTAGGCAAAAGAAGGGGAACGTG         360

04570587reversecom   TGGAGCTTCAACTAATTTATCCCCTTTTTTCTCATCTTTAGGCAAAAGAAGGGGAACGTG         122
                     ************************************************************

hSPOCK1              GCCCAGAAACACTGGGTTGGACCTTCGAATTTGGTCAAGTGCAAGCCCTGTCCCGTGGCA         420

04570587reversecom   GCCCAGAAACACTGGGTTGGACCTTCGAATTTGGTCAAGTGCAAGCCCTGTCCCGTGGCA         182
                     ************************************************************

hSPOCK1              CAGTCAGCCATGGTCTGCGGCTCAGATGGCCACTCCTACACATCCAAGGTTGGTTTTCTT         480

04570587reversecom   CAGTCAGCCATGGTCTGCGGCTCAGATGGCCACTCCTACACATCCAAGGTTGGTTTTCTT         242
                     ************************************************************

hSPOCK1              GTCTTTTTATGTTATCAAAAATACTTCATGTCGGCTTGTTCTAACACTTCTCTATGCCAA         540

04570587reversecom   GTCTTTTTANGTTANCAAAANTA--------------------------------------         265
                     ********* **** ***** **

hSPOCK1              AACTGGGGCTTG  552

04570587reversecom   ------------  265
```

>04570590 (S13 Forward)

```
04570590 (S013     ------------------------------------------------------------          0
hSPOCK1            GGGCACCCGTGATATGCCACATACGGTATTGGGGCTGATTAAGCCATATTCCCTCACCGT         60


04570590 (S013     --------TTTTTTTTTTCNNNGNGNNNNNAAAANNNTTNCCNTGTNAGTAGTCNNGTTTTC         52
hSPOCK1            TTTTTTTCTTTTTTTTTTCTGTGTGTTGTAAAAATCTTCCCATGTCAGTAGTCATGTTTTC        120
                           **********    * *      ****      ** ** *** ******* ******


04570590 (S013     AATGGCTACNTTGTTTTCCNCNNNCCATCTGGTTTGTTGAATCTATGCTTCTGGGACAAA        112
hSPOCK1            AATGGCTACATTGTTTTCCTCAGTCCATCTGGTTTGTTGAATCTATGCTTCTGGGACAAA        180
                   ********* ********* *    ***********************************


04570590 (S013     ATCCTCTGCTTCCTATCAATGAGTTGGAGCAAATGGAAGTGCTGGCTAAGGCCCTCNGGA        172
hSPOCK1            ATCCTCTGCTTCCTATCAATGAGTTGGAGCAAATGGAAGTGCTGGCTAAGGCCCTCAGGA        240
                   ******************************************************** ***


04570590 (S013     TCCCTTGGCCTGGTTGGATANTCCTTTATTTNTTTNNNNTTTTNTTNNGNNAACNNAAAA        232
hSPOCK1            TCCCTTGGCCTGGTTGGATAGTCCTTTATTTATTTGGCCTTTTATTCTGGGAACTTAGAG        300
                   ******************** ********** ***    **** **  *  ***   * *


04570590 (S013     GAAANACTCNNAGGGGCGGGANCNTCAANTNATTNNTCTTTTTTTTTCNCTTCTTNAAAA        292
hSPOCK1            GAAAGACTCATAGGGTCTGGAGCTTCAACTAATTTATCCCCTTTTTTCTCATCTTTAGGC        360
                   **** ****  **** * *** * **** * ***   **    ******* * **** *


04570590 (S013     AAAN--------------------------------------------------------        296
hSPOCK1            AAAAGAAGGGGAACGTGGCCCAGAAACACTGGGTTGGACCTTCGAATTTGGTCAAGTGCA        420
                   ***


04570590 (S013     ------------------------------------------------------------        296
hSPOCK1            AGCCCTGTCCCGTGGCACAGTCAGCCATGGTCTGCGGCTCAGATGGCCACTCCTACACAT        480


04570590 (S013     ------------------------------------------------------------        296
hSPOCK1            CCAAGGTTGGTTTTCTTGTCTTTTTATGTTATCAAAAATACTTCATGTCGGCTTGTTCTA        540


04570590 (S013     ------------------------------------------------------------        296
hSPOCK1            ACACTTCTCTATGCCAAAACTGGGGCTTGGGAAGCGTGGTGGGCTCCAGGCCTTGCCCTT        600
```

>04570591 (S13 Reverse)

```
04570591reversecom   ------------------------------------------------------------   0
hSPOCK1              GGGCACCCGTGATATGCCACATACGGTATTGGGGCTGATTAAGCCATATTCCCTCACCGT   60


04570591reversecom   ------------------------------------------------------------   0
hSPOCK1              TTTTTTTCTTTTTTTTTTCTGTGTGTTGTAAAAATCTTCCCATGTCAGTAGTCATGTTTTC   120


04570591reversecom   ------------------------------------------------------------   0
hSPOCK1              AATGGCTACATTGTTTTCCTCAGTCCATCTGGTTTGTTGAATCTATGCTTCTGGGACAAA   180


04570591reversecom   ------------------------------------------------------------   0
hSPOCK1              ATCCTCTGCTTCCTATCAATGAGTTGGAGCAAATGGAAGTGCTGGCTAAGGCCCTCAGGA   240


04570591reversecom   -----------------------------TTATTTGGCCTTTTATTCTGGGAACTTAGAG   31
hSPOCK1              TCCCTTGGCCTGGTTGGATAGTCCTTTATTTATTTGGCCTTTTATTCTGGGAACTTAGAG   300
                                                  ******************************
04570591reversecom   GAAAGACTCATAGGGTCTGGAGCTTCAACTAATTTATCCCCTTTTTTCTCATCTTTAGGC   91
hSPOCK1              GAAAGACTCATAGGGTCTGGAGCTTCAACTAATTTATCCCCTTTTTTCTCATCTTTAGGC   360
                     ************************************************************
04570591reversecom   AAAAGAAGGGGAACGTGGCCCAGAAACACTGGGTTGGACCTTCGAATTTGGTCAAGTGCA   151
hSPOCK1              AAAAGAAGGGGAACGTGGCCCAGAAACACTGGGTTGGACCTTCGAATTTGGTCAAGTGCA   420
                     ************************************************************
04570591reversecom   AGCCCTGTCCCGTGGCACAGTCAGCCATGGTCTGCGGCTCAGATGGCCACTCCTACACAT   211
hSPOCK1              AGCCCTGTCCCGTGGCACAGTCAGCCATGGTCTGCGGCTCAGATGGCCACTCCTACACAT   480
                     ************************************************************
04570591reversecom   CCAAGGTTGGTTTTCTTGTCTT---------------------------------------   233
hSPOCK1              CCAAGGTTGGTTTTCTTGTCTTTTTATGTTATCAAAAATACTTCATGTCGGCTTGTTCTA   540
                     *********************
04570591reversecom   ------------------------------------------------------------   233
hSPOCK1              ACACTTCTCTATGCCAAAACTGGGGCTTGGGAAGCGTGGTGGGCTCCAGGCCTTGCCCTT   600
```

>04570588 (C10 Forward)

```
04570588    ------------------------------------------------------------ 0
hSPOCK1     GGGCACCCGTGATATGCCACATACGGTATTGGGGCTGATTAAGCCATATTCCCTCACCGT 60


04570588    ----------------------------AAAANNNNNNNCNTGNNNGTANTCNTGTTTTC 32
hSPOCK1     TTTTTTTCTTTTTTTTTCTGTGTGTTGTAAAAATCTTCCCATGTCAGTAGTCATGTTTTC 120
                                        ****      *  **    *** ** *******
04570588    AATGGNNNCNTTNTTTTCCNCNNTCCATCTGGTTTGTTGAATCNNTGCNTCNGGGANAAA 92
hSPOCK1     AATGGCTACATTGTTTTCCTCAGTCCATCTGGTTTGTTGAATCTATGCTTCTGGGACAAA 180
            *****    * ** ****** *  ******************  *** ** **** ***
04570588    ATCCNCNNCNTCCNNNNNNNTGANTTGGANCAAATGNAANTGCNNNCTAANGCCCNCAGGA 152
hSPOCK1     ATCCTCTGCTTCCTATCAATGAGTTGGAGCAAATGGAAGTGCTGGCTAAGGCCCTCAGGA 240
            **** *  * ***       *** ***** ****** ** ***    **** **** *****
04570588    TCCCTTGGCCTGGNTGGANAGTCCTTTATTTNGNNG----------------------- 188
hSPOCK1     TCCCTTGGCCTGGTTGGATAGTCCTTTATTTATTTGGCCTTTTATTCTGGGAACTTAGAG 300
            ************* **** ************   *
04570588    ------------------------------------------------------------ 188
hSPOCK1     GAAAGACTCATAGGGTCTGGAGCTTCAACTAATTTATCCCCTTTTTTCTCATCTTTAGGC 360


04570588    ------------------------------------------------------------ 188
hSPOCK1     AAAAGAAGGGGAACGTGGCCCAGAAACACTGGGTTGGACCTTCGAATTTGGTCAAGTGCA 420


04570588    ------------------------------------------------------------ 188
hSPOCK1     AGCCCTGTCCCGTGGCACAGTCAGCCATGGTCTGCGGCTCAGATGGCCACTCCTACACAT 480


04570588    ------------------------------------------------------------ 188
hSPOCK1     CCAAGGTTGGTTTTCTTGTCTTTTTATGTTATCAAAAATACTTCATGTCGGCTTGTTCTA 540


04570588    ------------------------------------------------------------ 188
hSPOCK1     ACACTTCTCTATGCCAAAACTGGGGCTTGGGAAGCGTGGTGGGCTCCAGGCCTTGCCCTT 600
```

>04570589 (C10 Reverse)

```
hSPOCK1            CACATACGGTATTGGGGCTGATTAAGCCATATTCCCTCACCGTTTTTTTTCTTTTTTTTT      60
04570589reversecom ------------------------------------------------------------       0


hSPOCK1            CTGTGTGTTGTAAAAATCTTCCCATGTCAGTAGTCATGTTTTCAATGGCTACATTGTTTT     120
04570589reversecom ------------------------------------------------------------       0


hSPOCK1            CCTCAGTCCATCTGGTTTGTTGAATCTATGCTTCTGGGACAAAATCCTCTGCTTCCTATC     180
04570589reversecom ------------------------------------------------------------       0


hSPOCK1            AATGAGTTGGAGCAAATGGAAGTGCTGGCTAAGGCCCTCAGGATCCCTTGGCCTGGTTGG     240
04570589reversecom ------------------------------------------------------------       0


hSPOCK1            ATAGTCCTTTATTTATTTGGCCTTTTATTCTGGGAACTTAGAGGAAAGACTCATAGGGTC     300
04570589reversecom -------------TATTTGGCCTTTTATTCTGGGAACTTAGAGGAAAGACTCATAGGGTC      47
                                ***********************************************
hSPOCK1            TGGAGCTTCAACTAATTTATCCCCTTTTTTCTCATCTTTAGGCAAAAGAAGGGGAACGTG     360
04570589reversecom TGGAGCTTCAACTAATTTATCCCCTTTTTTCTCATCTTTAGGCAAAAGAAGGGGAACGTG     107
                   ************************************************************
hSPOCK1            GCCCAGAAACACTGGGTTGGACCTTCGAATTTGGTCAAGTGCAAGCCCTGTCCCGTGGCA     420
04570589reversecom GCCCAGAAACACTGGGTTGGACCTTCGAATTTGGTCAAGTGCAAGCCCTGTCCCGTGGCA     167
                   ************************************************************
hSPOCK1            CAGTCAGCCATGGTCTGCGGCTCAGATGGCCACTCCTACACATCCAAGGTTGGTTTTCTT     480
04570589reversecom CAGTCAGCCATGGTCTGCGGCTCAGATGGCCACTCCTACACNTNCAAGGTTGGTTTTCTT     227
                   ***************************************** * ****************
hSPOCK1            GTCTTTTTATGTTATCAAAAATACTTCATGTCGGCTTGTTCTAACACTTCTCTATGCCAA     540
04570589reversecom GTCTTTNTANG-------------------------------------------------     238
                   ****** ** *
hSPOCK1            AACTGGGGCTTG  552
04570589reversecom ------------  238
```

>04570592 (C11 Forward)

```
04570592     ------------------------------------------------------------ 0
hSPOCK1      GGGCACCCGTGATATGCCACATACGGTATTGGGGCTGATTAAGCCATATTCCCTCACCGT 60


04570592     ---------------------------AANNNNNNNNNNCNTGTCNGTNGTCNNGTTTTC 32
hSPOCK1      TTTTTTTCTTTTTTTTTTCTGTGTGTTGTAAAAATCTTCCCATGTCAGTAGTCATGTTTTC 120
                                      **            * **** ** ***   ******
04570592     AATGNCNNCNTTGTTTTCCNCNNTCCATCTGGTTTGTTGAATCNATGCTTCNGGGACAAA 92
hSPOCK1      AATGGCTACATTGTTTTCCTCAGTCCATCTGGTTTGTTGAATCTATGCTTCTGGGACAAA 180
             **** *   * ********* *   ******************** ****** ********
04570592     ATCCNCNGCTTCCNATCNATGAGTTGGAGCAAATGGAANTGCNGNCTAAGGCCCNCAGGA 152
hSPOCK1      ATCCTCTGCTTCCTATCAATGAGTTGGAGCAAATGGAAGTGCTGGCTAAGGCCCTCAGGA 240
             **** * ****** *** ******************** *** * ********* *****
04570592     TCCCTTGGCCTGGTTGGANANTCCTTTATT------------------------------ 182
hSPOCK1      TCCCTTGGCCTGGTTGGATAGTCCTTTATTTATTTGGCCTTTTATTCTGGGAACTTAGAG 300
             ****************** * *********
04570592     ------------------------------------------------------------ 182
hSPOCK1      GAAAGACTCATAGGGTCTGGAGCTTCAACTAATTTATCCCCTTTTTTCTCATCTTTAGGC 360


04570592     ------------------------------------------------------------ 182
hSPOCK1      AAAAGAAGGGGAACGTGGCCCAGAAACACTGGGTTGGACCTTCGAATTTGGTCAAGTGCA 420


04570592     ------------------------------------------------------------ 182
hSPOCK1      AGCCCTGTCCCGTGGCACAGTCAGCCATGGTCTGCGGCTCAGATGGCCACTCCTACACAT 480


04570592     ------------------------------------------------------------ 182
hSPOCK1      CCAAGGTTGGTTTTCTTGTCTTTTTATGTTATCAAAAATACTTCATGTCGGCTTGTTCTA 540


04570592     ------------------------------------------------------------ 182
hSPOCK1      ACACTTCTCTATGCCAAAACTGGGGCTTGGGAAGCGTGGTGGGCTCCAGGCCTTGCCCTT 600
```

>04570593 (C11 Reverse)

```
04570593reversecom   ------------------------------------------------------------   0
hSPOCK1              GGGCACCCGTGATATGCCACATACGGTATTGGGGCTGATTAAGCCATATTCCCTCACCGT   60


04570593reversecom   ------------------------------------------------------------   0
hSPOCK1              TTTTTTTCTTTTTTTTTTCTGTGTGTTGTAAAAATCTTCCCATGTCAGTAGTCATGTTTTC   120


04570593reversecom   ---------------TNNNTCNNTCCNTNTNGTNGNATGN----ATNNNTGNNTGNNNNA   41
hSPOCK1              AATGGCTACATTGTTTTCCTCAGTCCATCTGGTTTGTTGAATCTATGCTTCTGGGACAAA   180
                                    *   **  *** * * **     **      **   *    *     *

04570593reversecom   NNNNTNNNTNTNCTNNCNATNANTNNGNNCNANNNNANGNNNNNGCTNNNNNNNNCNNNC   101
hSPOCK1              ATCCTCTGCTTCCTATCAATGAGTTGGAGCAAATGGAAGTGCTGGCTAAGGCCCTCAGGA   240
                              *     * ** * ** *   *  * *    * *      ***          *

04570593reversecom   NNNNTNNCCNNGCNNGGNTNGNNNNTNNTNTATTTGGCCTTTTATTCTGGGAACTTAGAG   161
hSPOCK1              TCCCTTGGCCTGGTTGGATAGTCCTTTATTTATTTGGCCTTTTATTCTGGGAACTTAGAG   300
                          *   *   *   ** * *     *  * *****************************

04570593reversecom   GAAAGACTCATAGGGTCTGGAGCTTCAACTAATTTATCCCCTTTTTTCTCATCTTTAGGC   221
hSPOCK1              GAAAGACTCATAGGGTCTGGAGCTTCAACTAATTTATCCCCTTTTTTCTCATCTTTAGGC   360
                     ************************************************************

04570593reversecom   AAAAGAAGGGGAACGTGGCCCAGAAACACTGGGTTGGACCTTCGAATTTGGTCAAGTGCA   281
hSPOCK1              AAAAGAAGGGGAACGTGGCCCAGAAACACTGGGTTGGACCTTCGAATTTGGTCAAGTGCA   420
                     ************************************************************

04570593reversecom   AGCCCTGTCCCGTGGCACAGTCAGCCATGGTCTGCGGCTCAGATGGCCACTCCTACACNT   341
hSPOCK1              AGCCCTGTCCCGTGGCACAGTCAGCCATGGTCTGCGGCTCAGATGGCCACTCCTACACAT   480
                     ********************************************************** *

04570593reversecom   CCAAGGTTGGTTTTCTTGTCTTTTTATGTTNTNA--------------------------   375
hSPOCK1              CCAAGGTTGGTTTTCTTGTCTTTTTATGTTATCAAAAATACTTCATGTCGGCTTGTTCTA   540
                     *****************************  * *

04570593reversecom   ------------------------------------------------------------   375
hSPOCK1              ACACTTCTCTATGCCAAAACTGGGGCTTGGGAAGCGTGGTGGGCTCCAGGCCTTGCCCTT   600
```

>03249341 (S09 Reverse)

```
hSPOCK1          CACATACGGTATTGGGGCTGATTAAGCCATATTCCCTCACCGTTTTTTTTCTTTTTTTTT       60
03249341reverse  ------------------------------------------------------------        0


hSPOCK1          CTGTGTGTTGTAAAAATCTTCCCATGTCAGTAGTCATGTTTTCAATGGCTACATTGTTTT      120
03249341reverse  ------------------------------------------------------------        0


hSPOCK1          CCTCAGTCCATCTGGTTTGTTGAATCTATGCTTCTGGGACAAAATCCTCTGCTTCCTATC      180
03249341reverse  ------------------------------------------------------------        0


hSPOCK1          AATGAGTTGGAGCAAATGGAAGTGCTGGCTAAGGCCCTCAGGATCCCTTGGCCTGGTTGG      240
03249341reverse  ------------------------------------------------------------        0


hSPOCK1          ATAGTCCTTTATTTATTTGGCCTTTTATTCTGGGAACTTAGAGGAAAGACTCATAGGGTC      300
03249341reverse  -----------TTTATTTGGCCTTTTATTCTGGGAACTTAGAGGAAAGACTCATAGGGTC       49
                            *************************************************


hSPOCK1          TGGAGCTTCAACTAATTTATCCCCTTTTTTCTCATCTTTAGGCAAAAGAAGGGGAACGTG      360
03249341reverse  TGGAGCTTCAACTAATTTATCCCCTTTTTTCTCATCTTTAGGCAAAAGAAGGGGAACGTG      109
                 ************************************************************


hSPOCK1          GCCCAGAAACACTGGGTTGGACCTTCGAATTTGGTCAAGTGCAAGCCCTGTCCCGTGGCA      420
03249341reverse  GCCCAGAAACACTGGGTTGGACCTTCGAATTTGGTCAAGTGCAAGCCCTGTCCCGTGGCA      169
                 ************************************************************


hSPOCK1          CAGTCAGCCATGGTCTGCGGCTCAGATGGCCACTCCTACACATCCAAGGTTGGTTTTCTT      480
03249341reverse  CAGTCAGCCATGGTCTGCGGCTCAGATGGCCACTCCTACNCNNNCAAGGTTGGTTTTC--      227
                 *************************************** *      *************


hSPOCK1          GTCTTTTTATGTTATCAAAAATACTTCATGTCGGCTTGTTCTAACACTTCTCTATGCCAA      540
03249341reverse  ------------------------------------------------------------      227


hSPOCK1          AACTGGGGCTTG          552
03249341reverse  ------------          227
```

>03249342 (S012 Reverse)

```
hSPOCK1              CACATACGGTATTGGGGCTGATTAAGCCATATTCCCTCACCGTTTTTTTTCTTTTTTTTT        60
03249342reverse      ------------------------------------------------------------        0


hSPOCK1              CTGTGTGTTGTAAAAATCTTCCCATGTCAGTAGTCATGTTTTCAATGGCTACATTGTTTT        120
03249342reverse      ------------------------------------------------------------        0


hSPOCK1              CCTCAGTCCATCTGGTTTGTTGAATCTATGCTTCTGGGACAAAATCCTCTGCTTCCTATC        180
03249342reverse      ------------------------------------------------------------        0


hSPOCK1              AATGAGTTGGAGCAAATGGAAGTGCTGGCTAAGGCCCTCAGGATCCCTTGGCCTGGTTGG        240
03249342reverse      ------------------------------------------------------------        0


hSPOCK1              ATAGTCCTTTATTTATTTGGCCTTTTATTCTGGGAACTTAGAGGAAAGACTCATAGGGTC        300
03249342reverse      -----------TTTATTTGGCCTTTTATTCTGGGAACTTAGAGGAAAGACTCATAGGGTC        49
                                ************************************************


hSPOCK1              TGGAGCTTCAACTAATTTATCCCCTTTTTTCTCATCTTTAGGCAAAAGAAGGGGAACGTG        360
03249342reverse      TGGAGCTTCAACTAATTTATCCCCTTTTTTCTCATCTTTAGGCAAAAGAAGGGGAACGTG        109
                     ************************************************************


hSPOCK1              GCCCAGAAACACTGGGTTGGACCTTCGAATTTGGTCAAGTGCAAGCCCTGTCCCGTGGCA        420
03249342reverse      GCCCAGAAACACTGGGTTGGACCTTCGAATTTGGTCAAGTGCAAGCCCTGTCCCGTGGCA        169
                     ************************************************************


hSPOCK1              CAGTCAGCCATGGTCTGCGGCTCAGATGGCCACTCCTACACATCCAAGGTTGGTTTTCTT        480
03249342reverse      CAGTCAGCCATGGTCTGCGGCTCAGATGGCCACTCCTACACNTCCAAGGTTGGTTTTCTT        229
                     ******************************************* *****************


hSPOCK1              GTCTTTTTATGTTATCAAAAATACTTCATGTCGGCTTGTTCTAACACTTCTCTATGCCAA        540
03249342reverse      GTCTTTTTANGTTNTNAAAA----------------------------------------        249
                     ********* *** * ****


hSPOCK1              AACTGGGGCTTG        552
03249342reverse      ------------        249
```

>03249343 (S014 Reverse)

```
hSPOCK1          CACATACGGTATTGGGGCTGATTAAGCCATATTCCCTCACCGTTTTTTTTCTTTTTTTTT      60
03249343reverse  ------------------------------------------------------------      0


hSPOCK1          CTGTGTGTTGTAAAAATCTTCCCATGTCAGTAGTCATGTTTTCAATGGCTACATTGTTTT      120
03249343reverse  ------------------------------------------------------------      0


hSPOCK1          CCTCAGTCCATCTGGTTTGTTGAATCTATGCTTCTGGGACAAAATCCTCTGCTTCCTATC      180
03249343reverse  ------------------------------------------------------------      0


hSPOCK1          AATGAGTTGGAGCAAATGGAAGTGCTGGCTAAGGCCCTCAGGATCCCTTGGCCTGGTTGG      240
03249343reverse  ------------------------------------------------------------      0


hSPOCK1          ATAGTCCTTTATTTATTTGGCCTTTTATTCTGGGAACTTAGAGGAAAGACTCATAGGGTC      300
03249343reverse  -----------TTTATTTGGCCTTTTATTCTGGGAACTTAGAGGAAAGACTCATAGGGTC      49
                             *************************************************


hSPOCK1          TGGAGCTTCAACTAATTTATCCCCTTTTTTCTCATCTTTAGGCAAAAGAAGGGGAACGTG      360
03249343reverse  TGGAGCTTCAACTAATTTATCCCCTTTTTTCTCATCTTTAGGCAAAAGAAGGGGAACGTG      109
                 ************************************************************


hSPOCK1          GCCCAGAAACACTGGGTTGGACCTTCGAATTTGGTCAAGTGCAAGCCCTGTCCCGTGGCA      420
03249343reverse  GCCCAGAAACACTGGGTTGGACCTTCGAATTTGGTCAAGTGCAAGCCCTGTCCCGTGGCA      169
                 ************************************************************


hSPOCK1          CAGTCAGCCATGGTCTGCGGCTCAGATGGCCACTCCTACACATCCAAGGTTGGTTTTCTT      480
03249343reverse  CAGTCAGCCATGGTCTGCGGCTCAGATGGCCACTCCTACACNTCCAAGGTTGGTTTTCTT      229
                 ***************************************** *****************


hSPOCK1          GTCTTTTTATGTTATCAAAAATACTTCATGTCGGCTTGTTCTAACACTTCTCTATGCCAA      540
03249343reverse  GTCTTTNTANGNNATCAAAA----------------------------------------      249
                 ****** ** *  *******


hSPOCK1          AACTGGGGCTTG      552
03249343reverse  ------------      249
```

>03249344 (S015 Reverse)

```
hSPOCK1         CACATACGGTATTGGGGCTGATTAAGCCATATTCCCTCACCGTTTTTTTTCTTTTTTTTT       60
0324934reverse  ------------------------------------------------------------        0


hSPOCK1         CTGTGTGTTGTAAAAATCTTCCCATGTCAGTAGTCATGTTTTCAATGGCTACATTGTTTT      120
0324934reverse  ------------------------------------------------------------        0


hSPOCK1         CCTCAGTCCATCTGGTTTGTTGAATCTATGCTTCTGGGACAAAATCCTCTGCTTCCTATC      180
0324934reverse  ------------------------------------------------------------        0


hSPOCK1         AATGAGTTGGAGCAAATGGAAGTGCTGGCTAAGGCCCTCAGGATCCCTTGGCCTGGTTGG      240
0324934reverse  ------------------------------------------------------------        0


hSPOCK1         ATAGTCCTTTATTTATTTGGCCTTTTATTCTGGGAACTTAGAGGAAAGACTCATAGGGTC      300
0324934reverse  -----------TTTATTTGGCCTTTTATTCTGGGAACTTAGAGGAAAGACTCATAGGGTC       49
                           *************************************************


hSPOCK1         TGGAGCTTCAACTAATTTATCCCCTTTTTTCTCATCTTTAGGCAAAAGAAGGGGAACGTG      360
0324934reverse  TGGAGCTTCAACTAATTTATCCCCTTTTTTCTCATCTTTAGGCAAAAGAAGGGGAACGTG      109
                ************************************************************


hSPOCK1         GCCCAGAAACACTGGGTTGGACCTTCGAATTTGGTCAAGTGCAAGCCCTGTCCCGTGGCA      420
0324934reverse  GCCCAGAAACACTGGGTTGGACCTTCGAATTTGGTCAAGTGCAAGCCCTGTCCCGTGGCA      169
                ************************************************************


hSPOCK1         CAGTCAGCCATGGTCTGCGGCTCAGATGGCCACTCCTACACATCCAAGGTTGGTTTTCTT      480
0324934reverse  CAGTCAGCCATGGTCTGCGGCTCAGATGGCCACTCCTACNNNNNCAAGNNNGGTTNTC--      227
                *************************************** ****   **** **


hSPOCK1         GTCTTTTTATGTTATCAAAAATACTTCATGTCGGCTTGTTCTAACACTTCTCTATGCCAA      540
0324934reverse  ------------------------------------------------------------      227


hSPOCK1         AACTGGGGCTTG       552
0324934reverse  ------------       227
```

7.2 BioEdit sequence alignment of all single colony samples

>04570586 (S16 Forward)

NAAAANNNNNNNNNTGTNNGTNNTCNTGTTTTCAATGGCNACATTGTTTTCCNCNNTCCATCTGGTTTGTTGAATCTATGCTTCNGGGACAAAATCC
NCNGCTTCCNATCAATGAGTTGGAGCAAATGGAAGTGCNGGCTAAGGCCCNCNGGATCCCTTGGCCNGGNNGGATANTCCTTTATTTN

>04570586 reverse complement

NAAATAAAGGANTATCCNNCCNGGCCAAGGGATCCNGNGGGCCTTAGCCNGCACTTCCATTTGCTCCAACTCATTGATNGGAAGCNGNGGATTTTG
TCCCNGAAGCATAGATTCAACAAACCAGATGGANNGNGGAAAACAATGTNGCCATTGAAAACANGANNACNNACANNNNNNNNNTTTTN

>04570587 (S16 Reverse)

TANTTTTGNTAACNTAAAAAGACAAGAAAACCAACCTTGGATGTGTAGGAGTGGCCATCTGAGCCGCAGACCATGGCTGACTGTGCCACGGGACAG
GGCTTGCACTTGACCAAATTCGAAGGTCCAACCCAGTGTTTCTGGGCCACGTTCCCCTTCTTTTGCCTAAAGATGAGAAAAAAGGGGATAAATTAGTT
GAAGCTCCAGACCCTATGAGTCTTTCCTCTAAGTTCCCAGAATAAAAGGCCAAATAAANNANNNNCNATCC

>04570587 reverse complement

GGATNGNNNNTNNTTTATTTGGCCTTTTATTCTGGGAACTTAGAGGAAAGACTCATAGGGTCTGGAGCTTCAACTAATTTATCCCCTTTTTTCTCATCT
TTAGGCAAAAGAAGGGGAACGTGGCCCAGAAACACTGGGTTGGACCTTCGAATTTGGTCAAGTGCAAGCCCTGTCCCGTGGCACAGTCAGCCATGG
TCTGCGGCTCAGATGGCCACTCCTACACATCCAAGGTTGGTTTTCTTGTCTTTTTANGTTANCAAAANTA
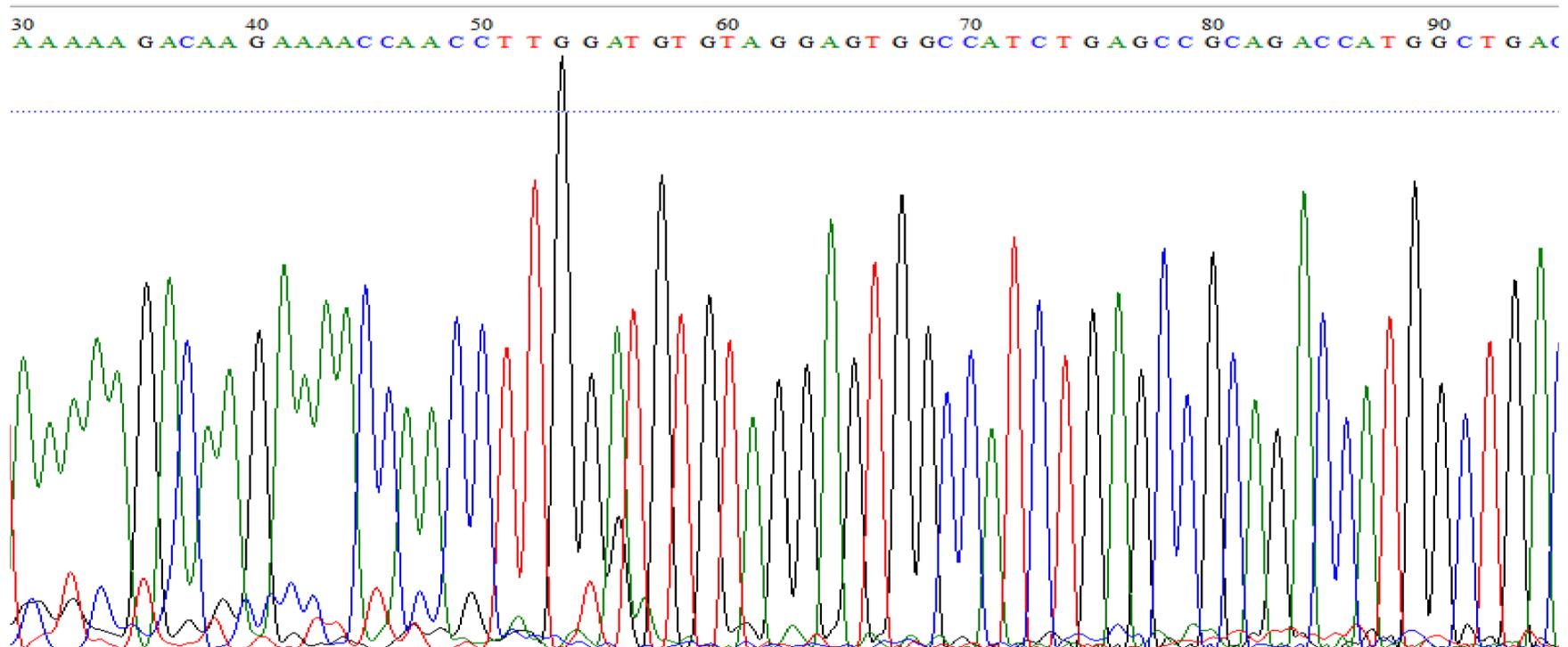
>04570588 (C10 Forward)

AAAANNNNNNNCNTGNNNGTANTCNTGTTTTCAATGGNNNCNTTNTTTTCCNCNNTCCATCTGGTTTGTTGAATCNNTGCNTCNGGGANAAAATC
CNCNNCNTCCNNNNNNNTGANTTGGANCAAATGNAANTGCNNNCTAANGCCCNCAGGATCCCTTGGCCTGGNTGGANAGTCCTTTATTTNGNNG

 >04570588 reverse complement

CNNCNAAATAAAGGACTNTCCANCCAGGCCAAGGGATCCTGNGGGCNTTAGNNNGCANTTNCATTTGNTCCAANTCANNNNNNGGANGNNGNG
GATTTTNTCCCNGANGCANNGATTCAACAAACCAGATGGANNGNGGAAAANAANGNNNCCATTGAAAACANGANTACNNNCANGNNNNNNNT
TTT
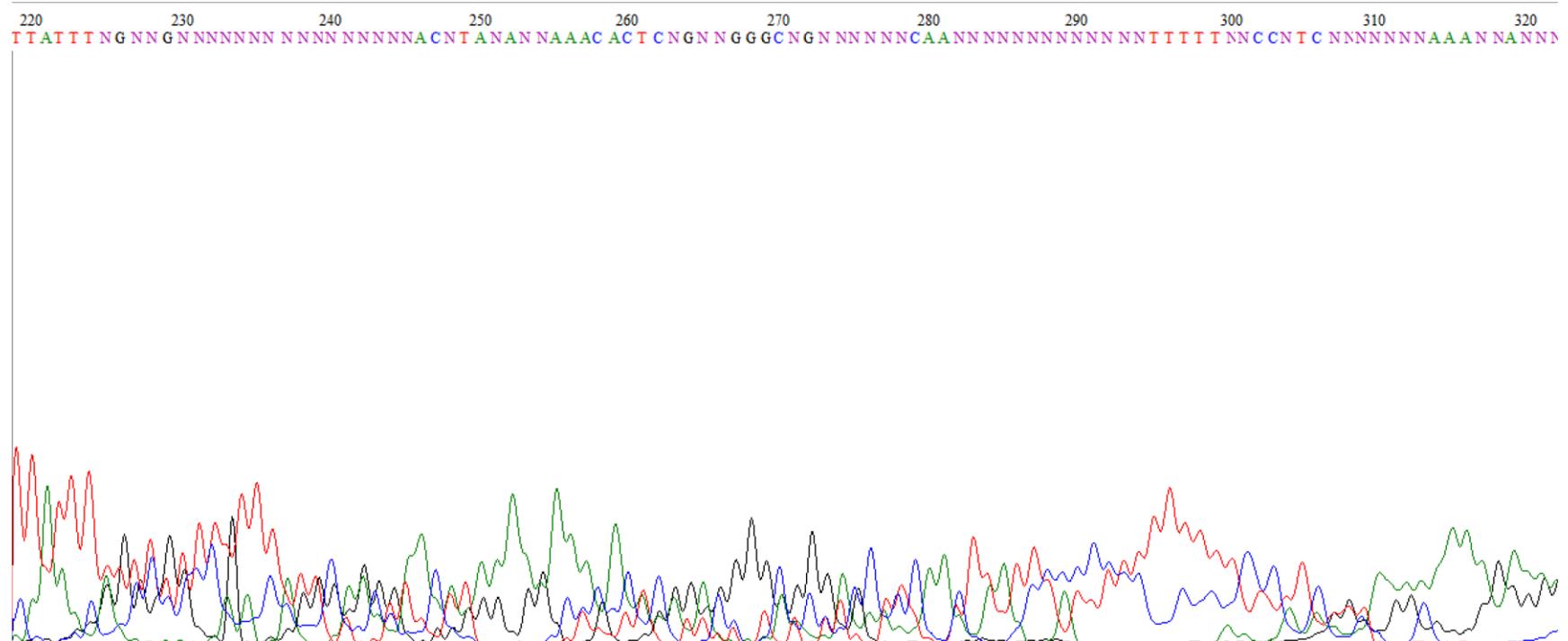
>04570589 (C10 Reverse)

CNTANAAAGACAAGAAAACCAACCTTGNANGTGTAGGAGTGGCCATCTGAGCCGCAGACCATGGCTGACTGTGCCACGGGACAGGGCTTGCACTTGACC
AAATTCGAAGGTCCAACCCAGTGTTTCTGGGCCACGTTCCCCTTCTTTTGCCTAAAGATGAGAAAAAAGGGGATAAATTAGTTGAAGCTCCAGACCCTATGA
GTCTTTCCTCTAAGTTCCCAGAATAAAAGGCCAAATA

>04570589 reverse complement

TATTTGGCCTTTTATTCTGGGAACTTAGAGGAAAGACTCATAGGGTCTGGAGCTTCAACTAATTTATCCCCTTTTTTCTCATCTTTAGGCAAAAGAAGGGGAA
CGTGGCCCAGAAACACTGGGTTGGACCTTCGAATTTGGTCAAGTGCAAGCCCTGTCCCGTGGCACAGTCAGCCATGGTCTGCGGCTCAGATGGCCACTCCT
ACACNTNCAAGGTTGGTTTTCTTGTCTTTNTANG
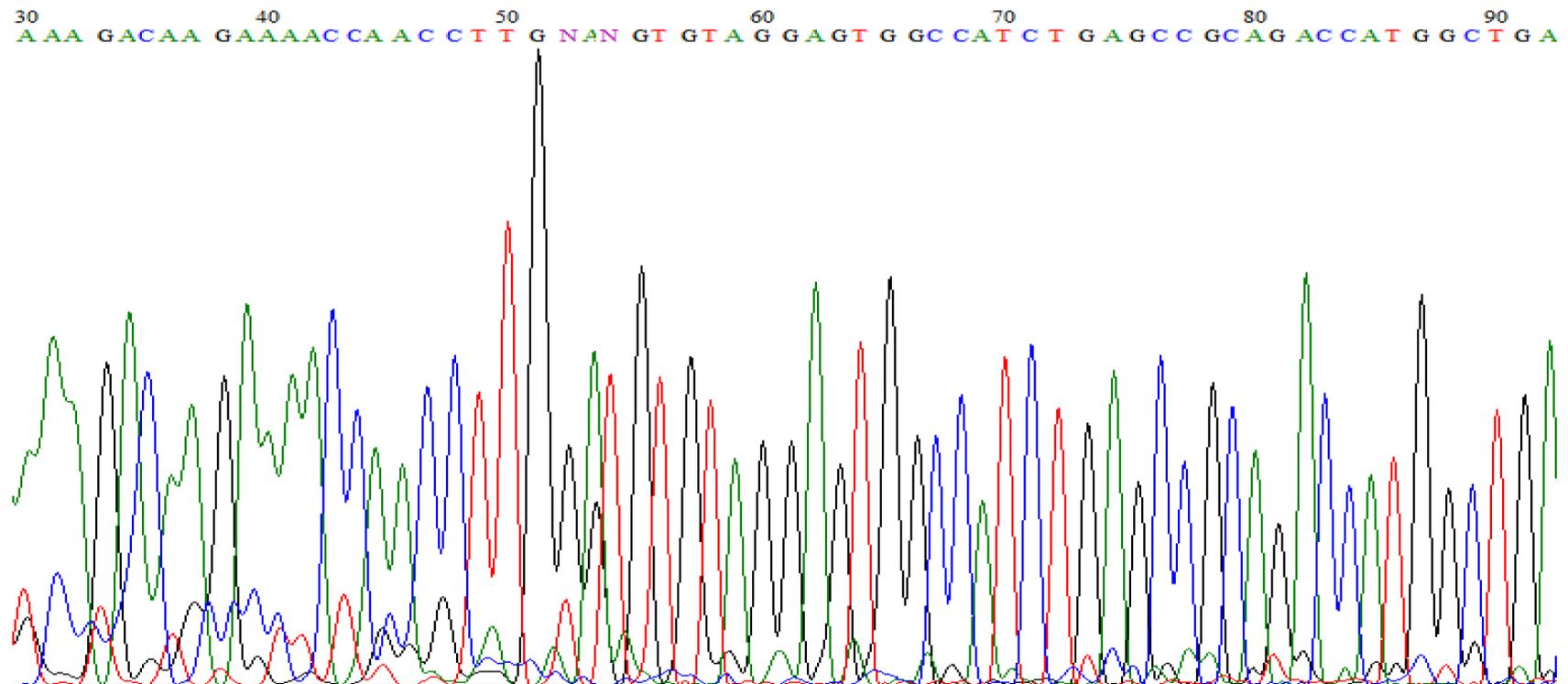
>04570592 (C11 Forward)

AANNNNNNNNNCNTGTCNGTNGTCNNGTTTTCAATGNCNNCNTTGTTTTCCNCNNTCCATCTGGTTTGTTGAATCNATGCTTCNGGGACAAAATCC
NCNGCTTCCNATCNATGAGTTGGAGCAAATGGAANTGCNGNCTAAGGCCCNCAGGATCCCTTGGCCTGGTTGGANANTCCTTTATT

>04570592reverscom

AATAAAGGANTNTCCAACCAGGCCAAGGGATCCTGNGGGCCTTAGNCNGCANTTCCATTTGCTCCAACTCATNGATNGGAAGCNGNGGATTTTGTC
CCNGAAGCATNGATTCAACAAACCAGATGGANNGNGGAAAACAANGNNGNCATTGAAAACNNGACNACNGACANGNNNNNNNNNTT
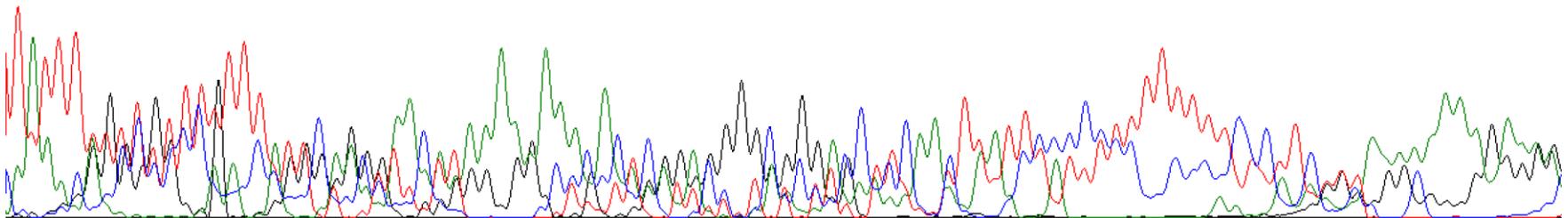
>04570593 (C11 Reverse)
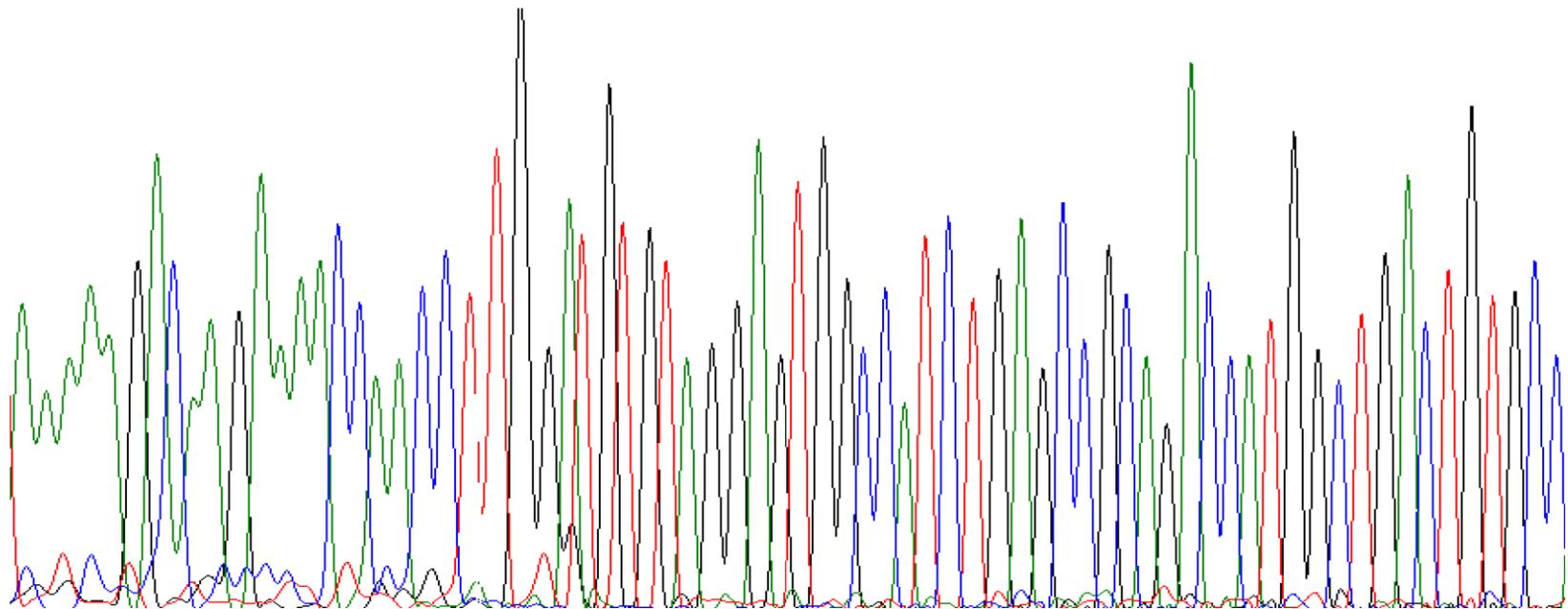
TNANAACATAAAAAGACAAGAAAACCAACCTTGGANGTGTAGGAGTGGCCATCTGAGCCGCAGACCATGGCTGACTGTGCCACGGGACAGGGCTTG
CACTTGACCAAATTCGAAGGTCCAACCCAGTGTTTCTGGGCCACGTTCCCCTTCTTTTGCCTAAAGATGAGAAAAAAGGGGATAAATTAGTTGAAGCTC
CAGACCCTATGAGTCTTTCCTCTAAGTTCCCAGAATAAAAGGCCAAATANANNANNNNCNANCCNNGCNNGGNNANNNNGNNNGNNNNNNNNA
GCNNNNNCNTNNNNTNGNNCNNANTNATNGNNAGNANANNNANNNNTNNNNCANNCANNNATNCATNCACNANANGGANNGANNNA

>04570593 reverse complement

TNNNTCNNTCCNTNTNGTNGNATGNATNNNTGNNTGNNNNANNNNTNNNTNTNCTNNCNATNANTNNGNNCNANNNNANGNNNNNGCTNNN
NNNNNCNNNCNNNNTNNCCNNGCNNGGNTNGNNNNTNNTNTATTTGGCCTTTTATTCTGGGAACTTAGAGGAAAGACTCATAGGGTCTGGAGCT
TCAACTAATTTATCCCCTTTTTTCTCATCTTTAGGCAAAAGAAGGGGAACGTGGCCCAGAAACACTGGGTTGGACCTTCGAATTTGGTCAAGTGCAAGC
CCTGTCCCGTGGCACAGTCAGCCATGGTCTGCGGCTCAGATGGCCACTCCTACACNTCCAAGGTTGGTTTTCTTGTCTTTTTATGTTNTNA
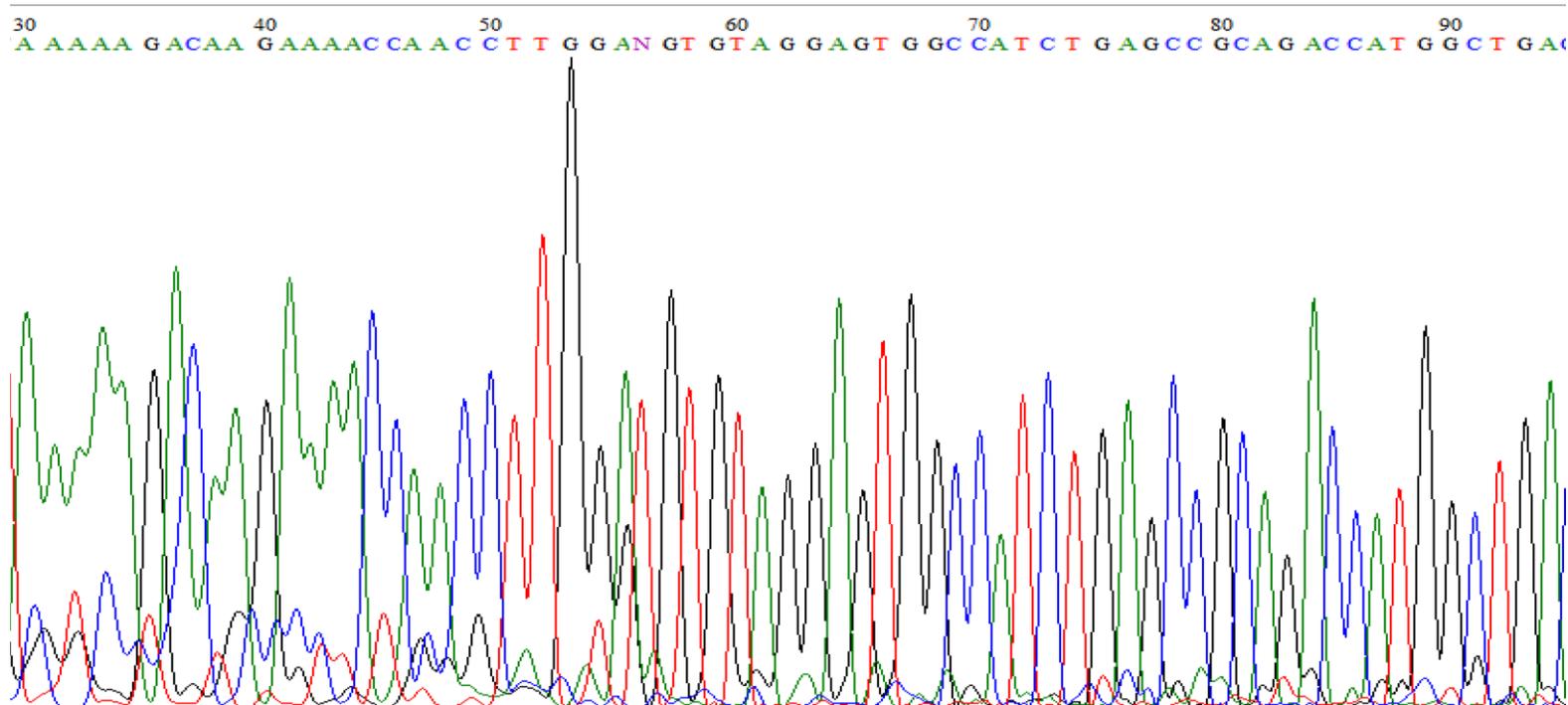
>03249341 (S09)

GAAAACCAACCTTGNNNGNGTAGGAGTGGCCATCTGAGCCGCAGACCATGGCTGACTGTGCCACGGGACAGGGCTTGCACTTGACCAAATTCGAAGGTCCAAC
CCAGTGTTTCTGGGCCACGTTCCCCTTCTTTTGCCTAAAGATGAGAAAAAAGGGGATAAATTAGTTGAAGCTCCAGACCCTATGAGTCTTTCCTCTAAGTTCCCAGA
ATAAAAGGCCAAATAAA

>03249341 reverse

TTTATTTGGCCTTTTATTCTGGGAACTTAGAGGAAAGACTCATAGGGTCTGGAGCTTCAACTAATTTATCCCCTTTTTTCTCATCTTTAGGCAAAAGAAGGGGAACGT
GGCCCAGAAACACTGGGTTGGACCTTCGAATTTGGTCAAGTGCAAGCCCTGTCCCGTGGCACAGTCAGCCATGGTCTGCGGCTCAGATGGCCACTCCTACNCNNN
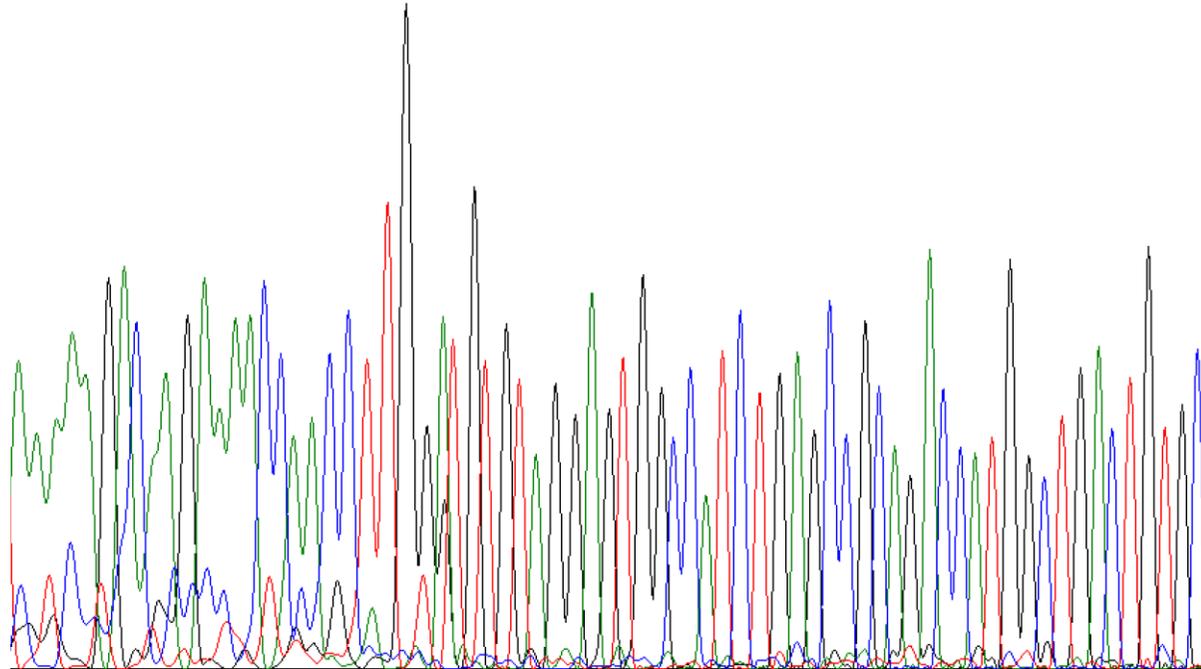CAAGGTTGGTTTTC
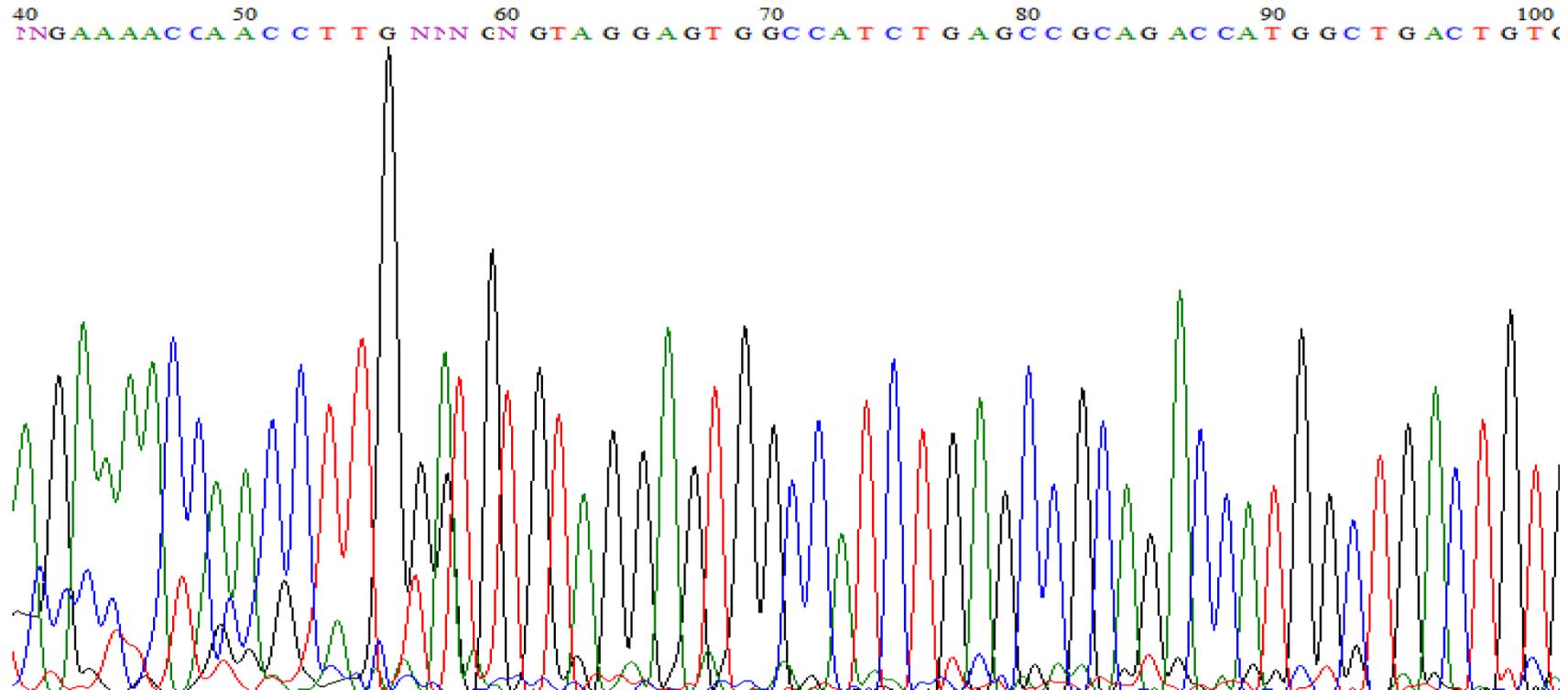
>03249342 (S12)

TTTTNANAACNTAAAAAGACAAGAAAACCAACCTTGGANGTGTAGGAGTGGCCATCTGAGCCGCAGACCATGGCTGACTGTGCCACGGGACAGG
GCTTGCACTTGACCAAATTCGAAGGTCCAACCCAGTGTTTCTGGGCCACGTTCCCCTTCTTTTGCCTAAAGATGAGAAAAAAGGGGATAAATTAGTT
GAAGCTCCAGACCCTATGAGTCTTTCCTCTAAGTTCCCAGAATAAAAGGCCAAATAAA

>03249342 reverse

TTTATTTGGCCTTTTATTCTGGGAACTTAGAGGAAAGACTCATAGGGTCTGGAGCTTCAACTAATTTATCCCCTTTTTTCTCATCTTTAGGCAAAAGAA
GGGGAACGTGGCCCAGAAACACTGGGTTGGACCTTCGAATTTGGTCAAGTGCAAGCCCTGTCCCGTGGCACAGTCAGCCATGGTCTGCGGCTCAG
ATGGCCACTCCTACACNTCCAAGGTTGGTTTTCTTGTCTTTTTANGTTNTNAAAA

>03249343 (S14)

TTTTGATNNCNTANAAAGACAAGAAAACCAACCTTGGANGTGTAGGAGTGGCCATCTGAGCCGCAGACCATGGCTGACTGTGCCACGGGACAGGG
CTTGCACTTGACCAAATTCGAAGGTCCAACCCAGTGTTTCTGGGCCACGTTCCCCTTCTTTTGCCTAAAGATGAGAAAAAAGGGGATAAATTAGTTGA
AGCTCCAGACCCTATGAGTCTTTCCTCTAAGTTCCCAGAATAAAAGGCCAAATAAA

>03249343 reverse

TTTATTTGGCCTTTTATTCTGGGAACTTAGAGGAAAGACTCATAGGGTCTGGAGCTTCAACTAATTTATCCCCTTTTTTCTCATCTTTAGGCAAAAGAAG
GGGAACGTGGCCCAGAAACACTGGGTTGGACCTTCGAATTTGGTCAAGTGCAAGCCCTGTCCCGTGGCACAGTCAGCCATGGTCTGCGGCTCAGAT
GGCCACTCCTACACNTCCAAGGTTGGTTTTCTTGTCTTTNTANGNNATCAAAA

>0324944 (S15)

GANAACCNNNCTTGNNNNNNGTAGGAGTGGCCATCTGAGCCGCAGACCATGGCTGACTGTGCCACGGGACAGGGCTTGCACTTGACCAAATTCGAAGG
TCCAACCCAGTGTTTCTGGGCCACGTTCCCCTTCTTTTGCCTAAAGATGAGAAAAAAGGGGATAAATTAGTTGAAGCTCCAGACCCTATGAGTCTTTCCTCT
AAGTTCCCAGAATAAAAGGCCAAATAAA

>0324944 reverse

TTTATTTGGCCTTTTATTCTGGGAACTTAGAGGAAAGACTCATAGGGTCTGGAGCTTCAACTAATTTATCCCCTTTTTTCTCATCTTTAGGCAAAAGAAGGG
GAACGTGGCCCAGAAACACTGGGTTGGACCTTCGAATTTGGTCAAGTGCAAGCCCTGTCCCGTGGCACAGTCAGCCATGGTCTGCGGCTCAGATGGCCA
CTCCTACNNNNNCAAGNNNGGTTNTC
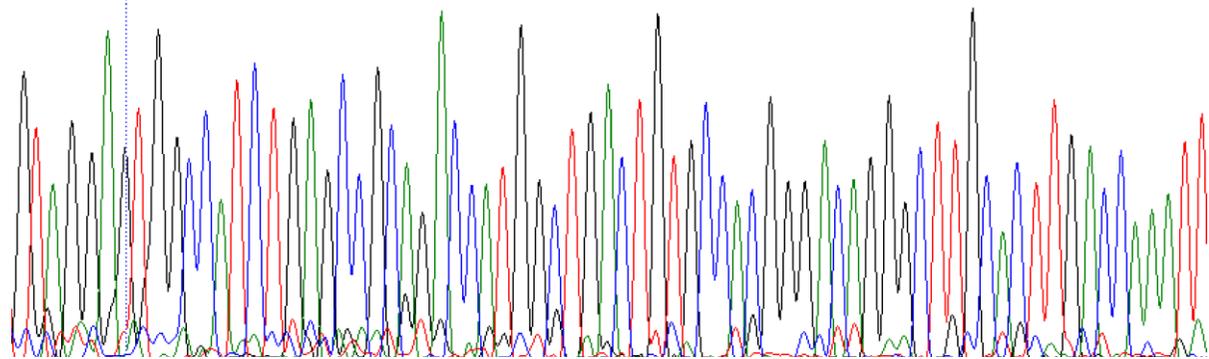
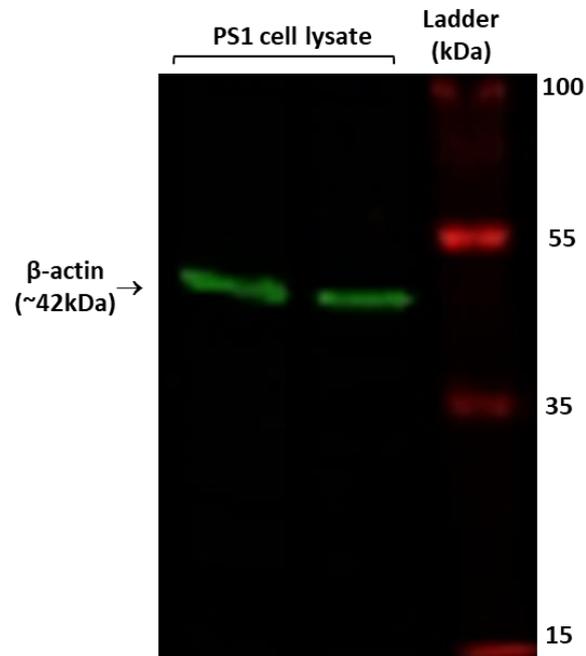## 7.3 Detecting SPOCK1 protein using western blot

Figure 7.1: SPOCK1 protein could not be detected using western blot. PS1 cells are a stellate cell line (stromal cell type) that express and secrete SPOCK1 into the extracellular environment. PS1 cell lysate was used for western blot to test if a rabbit anti-SPOCK1 antibody (Fisher Scientific, 16541322) would be able to detect the protein. Mouse β-actin was used as a loading control. Anti-mouse and anti-rabbit secondary antibodies (LI-COR, IRDye 680RD and IRDye 800CW) and an Odyssey Clx (LI-COR) were used for detection. SPOCK1 should be detected at ~49kDa and β-actin at ~42kDa.