Fall 12-14-2018

Functional and Mechanical Role of Splice Variant of Mucin4 (MUC4/X) and Trefoil Factors in Pancreatic Cancer Pathogenesis

Rahat Jahan
University of Nebraska Medical Center

Functional and Mechanical Role of Splice Variant of Mucin4 (MUC4/X) and Trefoil Factors in Pancreatic Cancer Pathogenesis

By

Rahat Jahan

A DISSERTATION

Presented to the Faculty of

The Graduate College in the University of Nebraska

In Partial Fulfillment of Requirements

For the Degree of Doctor of Philosophy

Under the Supervision of

Dr. Surinder K. Batra

Department of Biochemistry and Molecular Biology

University of Nebraska Medical Center

Omaha, Nebraska

November 2018

Supervisory Committee:

Kaustubh Datta, Ph.D.

Apar K. Ganti, MD.

Michael A. Hollingsworth, Ph.D.

Sukhwinder Kaur, Ph.D.

Amar B. Singh, Ph.D.

Paul Sorgen, Ph.D.

Functional and mechanical role of splice variant of Mucin4 (MUC4/X) and Trefoil Factors (TFFs) in Pancreatic Cancer Pathogenesis

University of Nebraska Medical Center, 2018

Advisor: Surinder K Batra, Ph.D.

Pancreatic Cancer (PC) is one of the vicious cancers as it ranks third in the race of leading cause of cancer-related death. Lack of early diagnostic marker, poor understanding of molecular mechanism of the disease and failure to conventional chemotherapy makes this disease dreadful.

Mucin 4 (MUC4), a high molecular weight glycoprotein is one of the top differentially expressed molecules in PC while not expressed in normal pancreas. Accumulating evidence from our lab suggested its tumorigenic role in PC by increasing cell proliferation, invasion, chemotherapy resistance, tumor growth, and metastasis. Previously, our lab and other has identified 24 different splice variant of MUC4 among them MUC4/X is devoid of exon 2 and 3 and MUC4/Y is devoid of exon 2. Exon 2 encodes for the largest domain of MUC4 suggesting that MUC4/X is devoid of the largest domain of MUC4 which variable tandem repeat. Though lots of effort has been made to identify its role in PC, there is still a gap on understanding its splice variant in PC as splice variant has an invaluable role in tumor pathogenesis. Recently splice variant has emerged as one of the key players for tumorigenesis and MUC4 is one of the key players for PC pathogenesis, we aim to identify the functional and mechanical role of MUC4/X, a

splice variant which is devoid of the largest domain of MUC4 yet contains all other functional domain, in PC pathogenesis.

Thus, in this part of dissertation, we sought to identify the role of splice variant MUC4/X, a unique splice variant of wild-type MUC4 which contain all functional domain except largest tandem repeat. First, we identified that, MUC4/X in aberrantly expressed in poorly differentiated PC clinical sample. Then our invitro experimental evidence suggested overexpression of MUC4/X in PC cells is involved in increased cell proliferation, invasion and metastasis. Moreover, our orthotopic transplantation system also corroborated our in-vitro findings which showed increased volume of tumor and metastasis to distant organ. Using inducible tet-on system to overexpress MUC4/X in the presence of WT-MUC4 in CAPAN-1 cells, we identified that MUC4/X has increased cell proliferation and invasion suggesting their role as tumorigenic alone as well as in the presence of WT-MUC4. Our mechanical investigation indicate that overexpression of MUC4/X led to upregulation of integrin β1-FAK-ERK pathway which might be potential mechanism for MUC4/X mediated PC tumorigenesis.

Lack of early effective diagnostic marker and resistance to chemotherapy are the major reasons for poor PC patient outcome. There is a pressing need to identify highly specific and sensitive biomarker as well as precise understanding of chemoresistance of PC. Trefoil factors (TFFs) are small secretory molecules mostly associated with mucin. Their primary role is to protect gastrointestinal tract partnering with mucin. Report on aberrant expression, potential as biomarker and role in tumorigenicity has conveyed for many cancers, however, their role in PC is

still elusive. Recently they have emerged as a part of gene signature of classical subtype of PC, a subtype which showed gemcitabine resistance towards PC. As it is high time to identify effective biomarker and understanding the role of chemoresistance in PC, in this part of my thesis, we focused to evaluate TFFs diagnostic potential using a training and validation cohort of PC clinical sample. Here, we comprehensively investigated the diagnostic potential of all the member of trefoil family, i.e., TFF1, TFF2, and TFF3 (TFFs) in combination with CA19.9 for detection of PC. In silico analysis of publicly available datasets and expression analysis from human and spontaneous PC mouse model revealed a significantly increased expression of TFFs in precursor lesions and PC cases. Additionally, we performed a comprehensive analysis in the sample set (n= 377) comprising of independent training and validation set using ELISA consisted of benign controls (BC), chronic pancreatitis (CP), and various stages of PC. Our analysis revealed that TFF1 and TFF2 were significantly elevated in early stages of PC in comparison to BC ($P<0.005$) and CP group ($P<0.05$) while significant elevation in TFF3 levels were in CP group. In receiver operating curve (ROC) analyses, combination of TFFs with CA19.9 emerged as promising panel for discriminating early stage of PC from BC (AUCTFF1+TFF2+TFF3+CA19.9=0.928) as well as CP (AUCTFF1+TFF2+TFF3+CA19.9 =0.943). Notably, at 90% specificity, TFFs combination improved CA19.9 sensitivity by 10% and 25% to differentiate early stage of PC from BC and CP respectively. Similar findings were observed in an independent validation set proving unique biomarker capabilities of TFFs. Overall

our study demonstrated that the combination of TFFs enhanced sensitivity and specificity of CA19.9 to discriminate early stage of PC from BC and CP.

Additionally, we also aim to identify the molecular landscape of TFFs role in gemcitabine resistance of PC which integrates analyzing publicly available cancer genome dataset, dissecting transcriptomic and signaling pathways and identification of biochemical interaction. From TCGA database analysis revealed a significant positive correlation between TFF1 and GR predictor of PC (P=0.0001). Our in vitro studies showed that SW1990-TFF1-KD cells induced apoptosis, reduced colony formation capacity and modulated many apoptotic regulators such increase of cleaved caspases and decrease of CIAP in the presence of gemcitabine. Furthermore, TFF1 was observed to be colocalized with MUC5AC, in human and mouse PC tissues suggesting their partnering are critical for PC pathogenesis. Interestingly, our chromatin immunoprecipitation indicates that 16 fold enrichment of GATA-6, an overexpressed transcription factor in classical subtype of PC, was observed on two distinct TFF1 promoter sites and GATA-6-siRNA repressed expression of TFF1. Moreover, protein-protein docking studies revealed the interaction of TFF1 with CXCR4 at Phe-172, Ser-122 and Glu-1 and TFF1 recombinant protein treatment in SW1990 cells increased CXCR4 mediated downstream signaling critical for GR. In this part, our overall data demonstrate that TFF1 may play a crucial role in gemcitabine resistance which is regulated by GATA6 and by interacting with MUC5AC.

Table of Content

List of Tables
List of Figures
Abbreviations
Acknowledgment

Table of Content

List of Table:

Chapter 1a

Chapter 1b

Chapter 2:

Chapter 4:

List of Figures

Chapter 1a

Chapter 1b

Chapter 3:

Chapter 4:

Figure 3: Higher levels of TFF1-3 are present in circulation during the early stages of PC.

Figure 4: (A) Evaluation of the diagnostic significance of TFFs in combination with CA19.9

Figure 4: (B). Evaluation of the diagnostic performance of TFF1-3 in CA19.9 low (<37U/ml) and high (≥37 U/ml) groups.

Figure S1: Expression of TFF1-3 in the publicly available cBioPortal database http://www.cbioportal.org/index.do) consisting of 169 studies from 30 different tumor types.

Figure S2: (A) Expression of TFF1-3 in PC tissues

Figure S2: (B-C) Expression of TFF1-3 in metastatic tissues and islet of Langerhans.

Figure S3: Diagnostic performance of individual TFF and ratio of TFF1/TFF3 in the training cohort.

Figure S4: Correlation between TFF1, TFF2, TFF3 and CA19.9 in all patients and within each group.

Figure S5: Overall diagnostic performance of individual TFF, combination of TFF1-3 and combination of TFFs and CA19.9 biomarkers in the training cohort

Figure S6: Diagnostic performance of individual TFF and the combination of TFF1-3 biomarkers in the validation cohort.

Chapter 5:

Figure 1: Correlation of TFF1 and GSR

Figure 2 (A): Effect of long-term treatment in Side Population (SP) cells.

Figure 2 (B-E): Effect of long-term treatment in stem cell marker, GSR and TFF1

Figure 3 (A-B). Role of TFF1 knockdown in gemcitabine resistance.

Figure 3 (C): TFF1 knockdown decreased colony formation in the presence of gemcitabine (1 μM)

Figure 3 (D-E): Effect of TFF1 knockdown in apoptotic protein molecules

Figure 4 (A-B).: Coexpression of GATA-6 in same PC duct

Figure 4 (C-D): GATA6 binding to the promoters of TFF1 detected by ChIP-qPCR in SW1990 cells

Figure 4 (E): GATA6 regulates TFF1 in SW1990 cells.

Figure 5: Colocalization of MUC5AC and TFF1 in PC human and spontaneous mouse tissues

Figure 5: Increase level of MUC5AC in long-term gemcitabine-treated SW1990 PC cells

Figure 6 (A): In Silico analysis of CXCR4 -TFF1 interaction

Figure 6 (B-D): CXCR4 antagonist partially abolished TFF1 downstream signaling.

List of Abbreviations

JAK/STAT: Janus kinase-signal transducers and activation of transcription;

GKN2: gastrokine 2

Lnc RNA: Long noncoding RNA

CDS: Coding DNA sequence

TSS: Transcription Start Site

LYVE-1: Lymphatic vessel endothelial hyaluronan receptor 1

REG1A: Regenerating Family Member 1 Alpha)

AP-1: activator protein 1

HNF3: Hepatocyte nuclear family 3

SV40: Simian vacuolating virus 40

ER: Estrogen Receptor

PRINS: Psoriasis-susceptibility-Related RNA Gene Induced by Stress

2-DG: 2-Deoxy-d-glucose

TPA: Tissue Plasminogen Activator

MMP-9: Matrix Metalloprotease

PAR: Protease Activated Receptor

LacdiNAc-N,N-di-N-acetyllactose diamine

TFF1: Trefoil Factor 1,

TFF2: Trefoil Factor 2,

TFF3: Trefoil Factor 3,

CA19.9: Cancer Antigen 19.9,

AUC: Area under the ROC curve,

SN: Sensitivity, SP: Specificity,

ROC: Receiver Operating Curve,

PC: Pancreatic Cancer.

BC: Benign Cancer,

CP: Chronic Pancreatitis,

EPC: Early Pancreatic Cancer,

LPC: Late Pancreatic Cancer.

Acknowledgments

First and foremost, I am grateful to my God for the good health and wellbeing that were necessary to complete this long journey. I am especially indebted to my supervisor, Dr. Surinder Kumar Batra, Professor and Chairman, Department of Biochemistry and Molecular Biology, for all the selfless time, wisdom, care, and unparallel support that kept me going. His unwavering guidance helped me in all the time of research and writing of this thesis.

I wholeheartedly express my sincere thanks to all the members of my supervisory committee, Michael A. Hollingsworth, Ph.D., Kaustubh Datta, Ph.D., Paul Sorgen Ph.D., Apar Ganti MD, Amar Singh, PhD and Sukhwinder Kaur, PhD not only for their insightful comments and encouragement, but also for the critical question which incented me to widen my research from various perspectives.

My thanks also go to all the member of my comprehensive exam committee, Justin Mott MD, Ph.D., Kelly Trujillo Ph.D., Kate Hyde Ph.D. and Punita Dhawan Ph.D. for their enthusiastic remarks and suggestions over the comprehensive exam period.

I am incredibly thankful and indebted to Dr. Sukhwinder Kaur for her constructive criticism and insightful suggestions towards my project, discussion and helping me in writing paper. I also want to extend my gratefulness to Dr. Muzafar Macha for valuable guidance and suggestions towards my project.

Dozens of people have helped and taught me immensely in my lab. I take this opportunity to express gratitude to all of the lab members for their help and support. I especially thank Dr. Maneesh Jain, Dr. Moorthy P. Ponnusamy, Dr. Sushil Kumar, Dr. Rachagani, Dr. Nasser, Dr. Partha Sarathy for their insightful feedback and advice that has been instrumental in shaping my research, I want to separately thank Dr. Imay for giving insightful comments and opportunity to work with him. My sincere thanks also go to Dr. Wade, Dr. Prakash, Dr. Abhijit, Dr. Shailendra, Dr. Sakthivel, Dr. Samy, Dr. Rakesh, Dr. Javed, Dr. Rohitesh, Dr. Ranjana, Dr. Naveen for giving advice and suggestion over the period. A very special and warm thanks goes to Dr. Raghupathy for being a good friend. I want to express my special appreciation and thanks to Dr. Sanjib, Dr. Ramesh and Dr.

Ashu for providing me suggestions any time when I needed it and also for being supportive labmate.

I want to express my sincere gratitude for the support and help from Kavita Mallya, our lab manager for being so kind and patient with me and helping me out when I was expecting a baby in the middle of my Ph.D. journey.

I am also very grateful to all the teaching and research staff that have taken some time to discuss and enrich my work. My warm thanks and gratitude go for Karen, Amy, Jeanette, Coleen, and April for all the support they have provided during my graduate training.

A special group from the lab is not mentioned yet, because they deserve their own part. Koelina Ganguly, a very good friend, a science-discussion partner over the last two years, and always managed to make me feel special. Some special words of gratitude go to my friends- Pranita, Sanchita, Sunandini, Saswati, Kasturi, Garima who have always been a major source of support. Their presence was very important in a process that is often felt as tremendously solitaire. I also thanks previous graduate students Dr. Priyanka, Dr. Suhasini, Dr. Priya, Dr. Suprit, Dr. Shiv, Dr. Seema for giving advice many times when I needed, even after they left the institution. I also acknowledge Joey Carmichael for being there any time needed, Chris and Andrew for their help any time necessary. My special gratitude also goes to Rama Krishn Nimmakayala for discussing experiments with great patience.

I also thank my parents, Shamsun Nahar and Abdul Wohab Chowdhury as well as my mother in law Kohinoor Begum for the unceasing encouragement, support and endless prayer from 8000 miles away from the USA. I also thank my brother-Sadman Abtahee, sister-Nusrat Jahan, brother in law- Ahmad Tawsif for blindly and selflessly believing in me. In this occasion, I remember my late father in law, Dr. Ruhul Amin, whom I definitely know would be one of the proudest people after I receive my degree.

I am grateful to my birth country Bangladesh for providing me an education. I feel blessed to have my training in two great alma mater Holy Cross Girls High School and University of Dhaka (Department of Pharmacy). In this moment I

especially all teachers from Holy Cross Girls High School for their unparallel guidance in my school life.

Last, but not the least, I am grateful and dedicate this thesis to my life-partner, Ahmad Tanwir who supported me each and every moment and being patient more than me throughout this entire journey. I can clearly say that, without his tremendous devotion and sacrifices, the journey would not be possible at all. Lastly, I thank my 2.5 years old son, Umair Ahmad for being the most lovable and understanding boy ever in the universe.

Chapter 1: Introduction

1(A). Transmembrane Mucins (MUCs): Structure and Molecular Signaling

Parts of this chapter are driven from:

Jahan R, Kaur S, Macha MA, Batra SK. Encyclopedia of Signaling Molecules. Mucins (MUCs). December 2016.DOI:10.1007/978-1-4614-6438-9_101739-1

Synonym

MUCIN 1: MUC1, transmembrane mucin 1, mucin-1, Muc1, Polymorphic epithelial mucin (PEM), Peanut lectin binding urinary mucin (PUM), Episialin, MAM-6, DF3 antigen, H23 Antigen, Epithelial Membrane Antigen (EMA), H23AG, Krebs von den Lungen-6 (KL-6), episialin, polymorphic epithelial mucin (PEM), cell membrane-associated polymorphic mucin, tumor-associated epithelial membrane antigen, tumor-associated mucin, MUC1/TR, MUC1 apo-mucin, mammary serum antigen (MSA), human milk fat globule antigen (HMFG), CAM 123-6, polymorphic urinary mucin (PUM), peanut-reactive urinary mucin, and the commonly used serum markers for breast cancer CA15.3 and CA27.29.

MUCIN 4: MUC4, Muc4, Ascites sialoglycoprotein-1 (ASGP-1).

MUCIN 16: MUC16, Mucin 16, Muc16, CA-125.

MUCIN 17: Mucin-17, membrane mucin 17, Muc3, intestinal membrane mucin 17, secreted mucin 17.

Historical background

The word “mucin” originated from the Greek word “*slimy,”* and denotes the glycoproteins of the mucosal lining of epithelial cells. By the 1980s, mucins were found to be an important molecule present in the mucosal epithelial lining of the respiratory, gastrointestinal, and genitourinary tract. Due to the large size and complex structure with extensive glycosylation, sequencing and functional studies were restricted until 1990. At that time, a revolution happened in understanding mucin structure, when four independent groups found an identical protein core

named MUC1, which was encoded from isolated cDNA clones from mammary and pancreatic mucins (Apostolopoulos et al. 2015). Cloning and nomenclature of other mucins followed MUC1, with full-length MUC1 cDNA clones obtained from the mammary gland or milk mucin. Interestingly, MUC1 lacks sequence homology with other transmembrane mucins, with MUC1 N- and C-ter domains bearing more resemblance. These were hypothesized to evolve from repeated region of secretory mucin MUC5B (Kufe, 2009). In 1991, MUC4 was cloned from the human tracheobronchial cDNA library and a human pancreatic tumor cell line (Chaturvedi et al. 2008). In 1999, the complete genomic sequence of MUC4 was established **(**Moniaux et al. 1999**)**. MUC16, CA125, a well-known serum biomarker for ovarian cancer, was first identified by an antibody (OC-125) developed by Bast and colleagues. This antibody was reactive against ovarian cancer cell OVCA 433, derived from a patient with serous ovarian carcinoma (Bast, Jr. et al. 1981). Later on, after almost two decades, from an independent study of two researchers, Yin and O'Brien showed that CA125 corresponds to one of the epitopes of one of the largest mucins, named MUC16 (Das and Batra, 2015). Van Klinken first identified MUC17 in the form of chimera in five tandem repeats, each composed of 59 amino acids (aa) located upstream of 17 aa tandem repeats of MUC3. In 2002, Gum et al. identified the origin of 59 amino acids as new membrane-bound mucin MUC17 from partial cDNA fragment. This TM mucin has five tandem repeats (TR), two epidermal growth factor (EGF)-like domains, one sea urchin sperm protein, the enterokinase, and agrin (SEA) domain, transmembrane domain (TM), and 80 amino acid (aa) long cytoplasmic tails (CT). In 2006, the full coding sequence of

MUC17 was characterized, and two splice variants of MUC17 (membrane bound and secreted, MUC17/SEC) were reported (Moniaux et al. 2006).

Structure and isoforms

To date, 21 different mucins have been identified (MUC1-MUC21), grouped into transmembrane, gel forming, and secreted family member. Due to the differential expression in multiple organs and various benign and malignant pathologies, among all transmembrane mucins, MUC1, MUC4, MUC16, and MUC17 have been the focus for research, with extensive efforts made to understand the structure, evolution, cell signaling, and role in disease progression. Thus, in this chapter, we will primarily discuss these four widely but not exclusively studied TM mucin molecules. Transmembrane mucins are characterized by the presence of a variable number of TR (VNTR) domains, along with a membrane-spanning trans-membrane, and a relatively small cytoplasmic tail (CT). The presence of TM domain tethers mucin to cell membranes and the domain is absent in secretory and gel-forming mucins. Tandem repeat (TR) domain, a hallmark of the mucin family, is rich in serine, threonine, and proline residues and is heavily O-glycosylated. TR sequence, number, and length are variable among members of the mucin family in orthologues. Their CT often contains a phosphorylation site to interact with the cytoplasmic cell signaling mediator (Hanson and Hollingsworth, 2016). Furthermore, transmembrane mucins MUC1, MUC4, MUC16, and MUC17 are characterized by the presence of unique domains and structural features discussed below and summarized in **Table 1.**

MUC1: MUC1 is a transmembrane glycoprotein (molecular weight 300 to 600 kDA)

located on human chromosome 1q21. MUC1 undergoes auto-cleavage at the N-terminal to the serine residue in the GSVVV motif found in the SEA module of the MUC1 extracellular domain. Its self-cleavage results in two subunits: MUC1-N terminal (MUC1-N or MUC1α) and MUC1-Carboxy terminal (MUC1-C or MUC1β). The MUC1-N subunit possesses a membrane localization signal sequence and a variable number of TR (VNTR) (20 aa) that are extensively O-glycosylated on serine and threonine residues. MUC1-C contain 58 aa extracellular domain (MUC1-ED), 28 aa transmembrane domain (MUC1-TM), and 72 aa cytoplasmic domain (MUC1-CT) (Pillai et al. 2015). The MUC1 gene covers seven exons on chromosome 1q21, where MUC1α is transcribed from exon1-4 and MUC1β is transcribed from exons 4–7. The largest exon, exon 2, consists of 20-200 units of TR (Zhang et al. 2013) **(Table 1).** To date, seventy-eight MUC1 isoforms have been identified that are generated from various alternative splicing mechanisms. Among them, the well-known MUC1 isoforms are the full-length form encoded by all seven exons within the intact VNTR region (MUC1/TR or MUC1/REP); variant of MUC1/TR i.e. MUC1/A, MUC1/B, MUC1/C, MUC1/D, MUC1/SEC (MUC1 secreted isoform) and the isoforms lacking VNTR, such as MUC1/X, MUC1/Y, MUC1/Z, and MUC1/ZD (Obermair et al. 2002). Studies have shown that variants A, D, X, Y, and Z were more frequently expressed in malignant than in benign tumors (Apostolopoulos et al. 2015). The schematic for MUC1 structure is shown in **Fig 1.**

MUC4: MUC4 has a transmembrane mucin domain, with some additional unique domains such as the nidogen-like (NIDO), the adhesion-associated domain in mucin MUC4 and other proteins (AMOP), the von Willebrand factor (vWD), TR domain with the VNTR, three EGF domains, a TM, and a CT domain. The MUC4 gene comprises 26 exons (E1–E26) and resides at chromosome 3q29. E2, the largest exon, codes for the central domain **(Table 1).** Because rat MUC4 (rMUC4) has shown cleavage at the GDPH site and rMuc4 and human MUC4 contain the conserved putative GDPH proteolytic cleavage site, human MUC4 is also hypothesized to be proteolytically cleaved at GDPH, generating two subunits: MUC4α (3000-7300 aa) and MUC4β (1156aa). MUC4α is a heavily glycosylated, mucin-like subunit containing TR, NIDO, and AMOP domains, whereas MUC4β is a growth factor-like subunit containing vWD, EGF-like domains, TM and a CT domain. With several alternative splicing mechanisms, 23 splice forms have been identified and grouped into secretory, membrane-bound with TR, and membrane-bound lacking TR (*MUC4*/X and *MUC4*/Y) (Chaturvedi et al. 2008). The schematic of the MUC4 structure with its unique domain is illustrated in **Figure 2.**

MUC16: MUC16, the largest cell surface mucin with a molecular weight of approximately 2.5 MDa (22152 aa), was mapped to chromosome 19p13.2. MUC16 is composed of three different domains: the N-terminal domain, a central TR region interspersed with SEA domains, and a carboxy-terminal domain. The N-terminal domain (12,000 aa) consists of the heavily O-glycosylated region, followed by a central region having 60 TRs of 156 aa, each of which is to date the largest repeat sequence among various transmembrane mucin (**Table 1**). The carboxy-terminal

domain is the smallest (284 aa), and consists of the extracellular region, TM, and a CT. 56 SEA (sea urchin, enterokinase, and agrin) domains are present in MUC16, some of which are interspersed near tandem repeats of MUC16. The presence of multiple SEA domains make MUC16 unique in structure from other transmembrane mucins, including MUC1, MUC12, MUC13, and MUC17, given that they possess only a single SEA domain. The CT of MUC16 (32 aa) contains tyrosine, threonine, and also has an RRRKK motif that is hypothesized to bind to the ezrin/radixin/moesin family, suggesting there may be an association with MUC16 and actin cytoskeleton (Haridas et al. 2014). It was speculated earlier that MUC16 undergoes proteolytic cleavage at another site, probably 50 aa upstream of the transmembrane domain, but our group studies using a recombinant construct indicate that cleavage might take place in the juxtamembrane region of MUC16 in the acidic pH of Golgi/post-Golgi compartments.This cleavage is primarily structure-dependent and not primarily amino acid sequence-dependent (Das and Batra, 2015). To date, isoform/splice variants of MUC16 have not been identified and this can be a potential area of study to understand the biology of MUC16**.** The unique features of MUC16 structure are summarized in **Table 1.**

MUC17: MUC17 resides on chromosome 7q22 in proximity with three other transmembrane mucin genes: MUC3A, MUC3B, and MUC12 (Gum, Jr. et al. 2002). Full-length MUC17 contains a 25 aa membrane targeting signal, a central domain of 63 TR consisting of 59-amino acid, two EGF-like domains, a SEA domain, a TM domain, and a CT (80 aa). Alternative splicing excludes exon 7 to generate a secreted form of the protein, MUC17/SEC, missing the second EGF

domain, TM, and carboxy terminal CT. MUC17 comprises 13 exons, where the largest exon, E3, encodes the tandem repeat domain (Moniaux et al. 2006). The unique features of MUC17 structure are summarized in **Table 1.**

Table1: Characteristic features of well studied transmembrane mucins.

Gene	Chromosomal Location	Predicted M.W. with glycosyaltion	No of Exons	Unique Domain	No of AA. in Tandem repeat	Ectodomain Tandem repeat sequence	Number of Tandem Repeat
MUC1	1q21	~250–500 kDa	8	Variable number of Tandem Repeat region, SEA, TM, cytoplasmic domain	20	PDPRPATGSTAPPAH GVTSA	25-125
MUC4	3q29	~550 to 930 kDa	26	Variable number of Tandem Repeat region, NIDO, AMOP, VWD, EGF, TM, and cytoplasmic domain	16	TSSASTGHATPLPVT D	145-500
MUC16	19p13.2	~3–5 million Da	84	Tandem Repeat, SEA, TM, cytoplasmic domain	156	FNPWSSVPTTSTPGT STPGTSTVHLATSGT PSSLPGHTAPVPLLIP FTLNFTITNLHYEENM QHPGSRKFNTTERVL QGLLKPLFKSTSVGP LYSGCRLTLLRPEKH GAATGVDAICTLRLD PTGPGLDRERLYWEL SQLTNSVTELGPYTL DRDSLYVNG	>60
MUC17	7q22	~450 kDa	13	Tandem repeat domain, two EGF-like domains, a SEA domain, transmembrane domain, and cytoplasmic tail domain.	59	LSTTPVASSEASTLST SPVDTSTPVTNSSPT NSSPTTAEVTSMPTS TAGEGSTPLTNMP	63

Abbreviations for Table 1:SEA: sea-urchin sperm protein, the enterokinase, and agrin, NIDO: nidogen like AMOP: adhesion-associated domain (AMOP), VWD: von Willebrand factor, TM: Transmembrane domain EGF: Epidermal Growth Factor.

Evolution of Transmembrane Mucin

Assessment of the transmembrane mucin sequence across species suggests the development of domains of transmembrane mucin from various sequences. The aa sequence of the MUC-N terminus upstream of the SEA domain and MUC1-CT appears to be evolved from MUC5B, whereas the SEA domain of MUC1 evolved from heparin sulfate proteoglycan 2 (HSGP2) (Kufe, 2009). On the other hand, the MUC16 SEA domain emerged from agrin, but MUC4 does not have any SEA domain. A homology analysis of MUC4 domains revealed positional conservation for crucial amino acids in various domains of MUC4, with a high degree of sequence similarity among multiple orthologues. This suggests the development of individual domains from common ancestral domains. The NIDO- and EGF-like domains appear to have evolved from a common ancestor to the nidogen protein, while AMOP and vWD domains evolved from Susd2 protein (Chaturvedi et al. 2008).

Physiological Function and Tissue Expression

Mucins are multifaceted molecules that function differently under normal and pathological conditions. Their function under pathological conditions varies owing to structural alteration, cell, and tissue-specific differential expression levels, as

well as differential post-translational modification during pathological development. Typical functions of various members of the mucin family include lubrication, protection from proteases, and defense against pathogens. MUC1 and MUC4 are strongly expressed in the lungs, where they serve as a structural barrier to foreign antigens and constitute first-line innate immunity, helping to expel mucus from airways in conjunction with secretory mucins MUC5AC and MUC5B. In the human uterine epithelium, MUC16 undergoes shedding to facilitate blastocyst adherence. The physiological function of the relevantly newly discovered mucin MUC17 is not well-studied. However, it is hypothesized to serve as a physical barrier against microorganisms, like other transmembrane mucins, and a high expression of MUC17 is observed in intestinal villi. It is interesting that in the administration of a recombinant protein containing both the EGF-like domains of Muc3 and the murine homolog of MUC17, significantly reduced mucosal ulceration by promoting cell migration and inhibition of apoptosis (Ho et al. 2006). Further, treatment with MUC17-CRD1-L-CRD2 following induction of colitis with DSS exhibited significantly improved crypt damage scores and severity, grade III ulceration compared to control animals (Luu et al. 2010). Tissue expression of mucin in normal and disease conditions has been summarized in Table 2.

Table 2. Expression of transmembrane mucin in healthy tissue as well as in disease state

Name	Normal Tissue Expression	Dysregulated Expression in Cancer and Other Diseases
MUC1	Glandular or luminal epithelial cells of the mammary gland, esophagus, stomach, duodenum, pancreas, uterus, spermatocytes, ocular surface epithelium, kidney, testis, prostate, and lungs, and to a lesser extent, in hematopoietic cells (non-epithelial cells). It is absent in the skin epithelium and mesenchymal cells.	Cancer (breast, pancreas, colon, lung and endometrial), hairy cell leukemia, follicular lymphoma, acute myelogenous leukemia, acute lymphoblastic leukemia, Sjogren's syndrome, chronic otitis, etc.
MUC4	Epithelial surfaces of the eye, oral cavity, middle ear, lachrymal glands, salivary glands, mammary gland, prostate gland, stomach, colon, lung, trachea, and female reproductive tract.	Cancer (pancreas, breast, colon, lung, ovarian, prostate), and Crohn's disease, etc.
MUC16	Epithelial lining upper respiratory tract, cornea, conjunctiva, female reproductive organ, pleura, peritoneum, pericardium, abdominal cavity and cervical mucus.	Cancer (ovarian, lung, pancreas, uterine cervix) inflammatory bowel disease, Crohn's disease, liver cirrhosis, pulmonary tuberculosis, and Sjogren's dry eye.
MUC17	Conjunctival epithelium, gastrointestinal tract with the highest expression in the duodenum, transverse colon, and terminal ileum.	Pancreatic cancer, inflammatory bowel disease, and colon cancer.

Transmembrane Mucin Mediated Cellular Signaling in Cancer

MUC1: Molecular signaling by transmembrane mucin MUC1 has been extensively studied in both normal and pathological conditions. The extracellular MUC1-N-terminal sub-unit remains attached with the transmembrane MUC1-C terminal subunit by non-covalent interactions to form a heterodimeric complex. The release of MUC1-N sub-unit from the cell surface leaves MUC1-C subunit to function as a putative receptor to interact with other cell surface receptors, specially Receptor Tyrosine kinases (RTKs). Galectin has been shown to serve as a ligand for MUC1, and it was interesting to note that Galectin-3 binds to MUC1-N subunit, which in turn triggers intracellular signaling leading to the recruitment of β-catenin to MUC1-

CT (Tanida et al. 2014) (**Fig. 1**). Similar to other mucins, underglycosylation of MUC1 is observed during malignant conditions. The interaction between underglycosylated, tumor-specific MUC1 and E-selectin promotes binding of MUC1 with the intercellular adhesion molecule 1(ICAM 1) on endothelial cells, and this allows the cancer cell to metastasize by migrating faster through the tumor microenvironment. This heterotypic interaction also triggers intracellular signaling through MUC1-Src-Crk-Rac1/Cdc42 to remodel the actin cytoskeleton of the cell and induce higher transendothelial migration of MUC1 overexpressing cells (Rahn et al. 2004;Shen et al. 2008). MUC1-TM serves as a connection between MUC1-ED and MUC1-CT to communicate signals from the extracellular milieu to the intracellular cytoplasm. Another group has shown that glycosylation of Asn 36 in MUC1-ED serves as a docking site for galectin-3 , which later serves as a bridge between MUC1-C with EGFR on the cell membrane (Ramasamy et al. 2007). MUC1-CT is the most studied cytoplasmic domain among the members of transmembrane mucin families. Many studies have showed that MUC1-CT induces cell transformation by interacting with various kinases, including PKCδ, glycogen synthase kinase 3β (GSK-3β), c-Src, ErbB, FGFR3, transcription factors, and chaperones associated with multiple cancers (**Fig. 1**). Phosphorylation of Threonine in the CT domain stimulates interaction between MUC1 and β-catenin and HSP 90, whereas phosphorylation of Ser inhibits the interaction of β-catenin with MUC1. Additionally, FGFR3 and platelet-derived growth factor receptor β (PDGFR β) interact with MUC1, leading to phosphorylation on the tyrosine residue in the cytoplasmic tail and resulting in MUC1-mediated invasion. The interaction

between MUC1-CT with EGFR triggers PI3K-AKT signaling cascade stimulating glycolysis, increases VEGF production, promotes angiogenesis and cell survival (**Fig. 1**). MUC1-CT contain motifs responsible for its nuclear localization, initiation of the signaling cascade, and promotion of cell proliferation and tumor progression. The absence of a classical nuclear localization signal (NLS) sequence in MUC1 did not restrict its transportation through the nuclear pore due to its association with importin B and a nucleoporin, NUP62, and more importantly, the CQC motif present in the cytoplasmic tail for MUC1 oligomerization. Moreover, mutation of CQC to a AQA motif has a dominant negative function on cancer cell growth, indicating its role in tumorigenesis (Kufe, 2009). β-catenin is another accomplice of MUC1-CT that facilitates MUC1-mediated tumorigenesis. The serine-rich motif (SRM, **S**AGNGG**SSLS**) is a crucial but not limiting factor in mediation of the interaction between MUC1-CT and β-catenin for activating β-catenin-Tcf–mediated transcription. Moreover, MUC1-C abrogates GSK3 beta-mediated phosphorylation and degradation of β-catenin, which can be attenuated by a mutation in the SRM motif (Huang et al. 2005). 17 beta-estradiol (E2) stimulates the binding of MUC1-CT with ER alpha (ERα) DNA-binding domain to block ubiquitination and degradation of ERα and to stimulate ERα-mediated growth and survival of breast cancer cells (Wei et al. 2006) (**Fig. 1**).

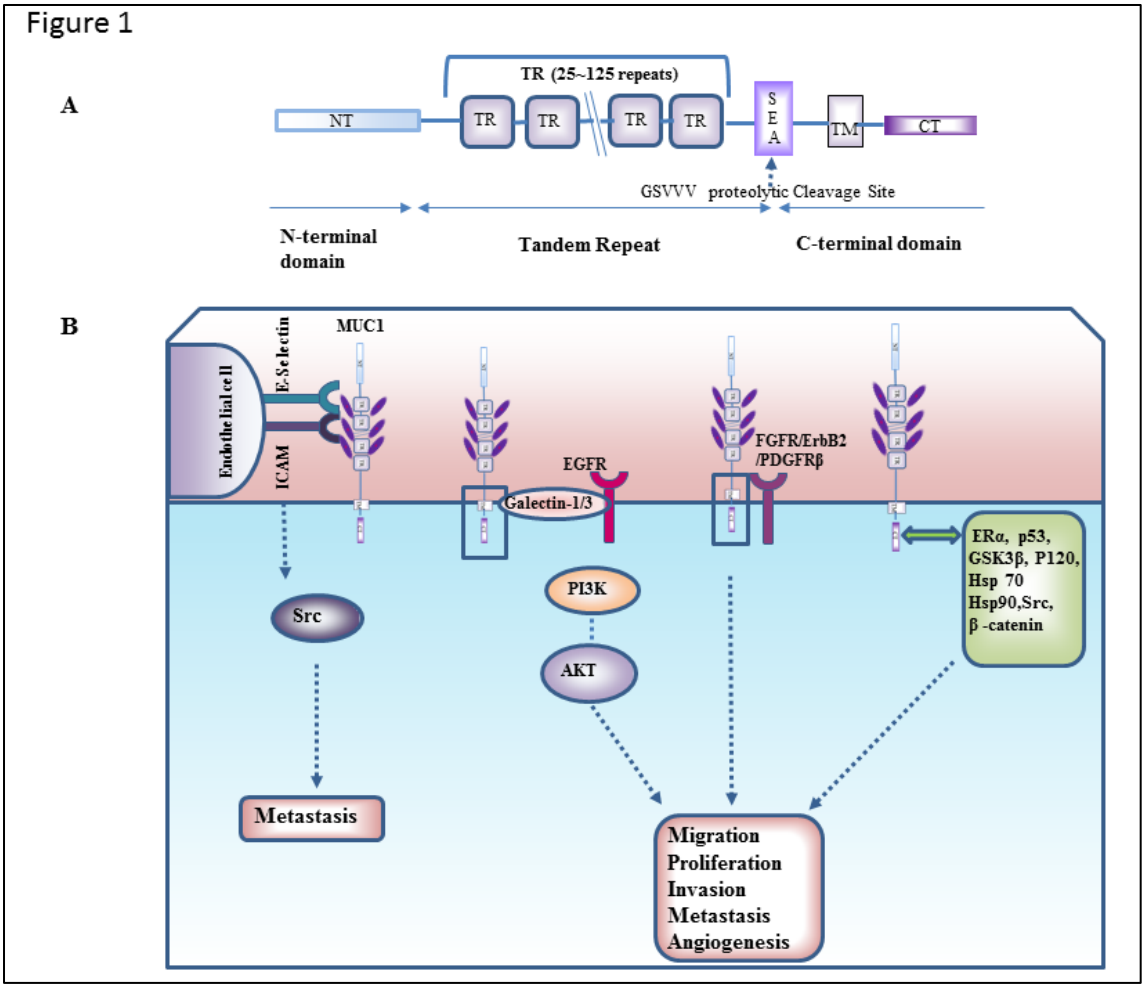


Figure 1: A: Schematic representation of MUC1 domain and structure. Membrane-tethered MUC1α (MUC1 N-terminus) and non-covalently attached MUC1β (MUC1-Cterminus) heterodimer are generated by auto-proteolytic cleavage in the SEA domain. MUC1α contains N-terminal region (NT), central tandem repeat domain (TR), extracellular region while MUC1β contains TM and CT. **B.** MUC1α interacts with ICAM and E-Selectin on endothelial cells that in-turn facilitate metastasis. MUC1α also interacts with EGFR, FGFR, Galectin 1 and 3, PDGFRβ to initiate cell signaling like PI3K-AKT to increase cell proliferation, metastasis, and migration. The cytoplasmic tail of MUC1 interacts with various molecules including ERα, p53, GSK3β, P120, HSP 70 Hsp90, Src, and β -catenin to initiate cell signaling cascade.

Abbreviations used: SEA: Sea urchin sperm protein-enterokinase-agrin; TM, Transmembrane; TR: Tandem Repeat Domain; vWD: von Willebrand D, Erα-estrogen receptor, GSK3β- Glycogen synthase kinase 3β, Hsp 70-Heat Shock Protein -70, Hsp90- Heat Shock Protein-90.

MUC4: Overexpression of MUC4 has been associated with reduced patient survival and enhanced tumorigenesis properties both *in-vitro* and *in vivo* in multiple malignancies, including pancreatic, breast, and ovarian cancers (Singh et al. 2007). MUC4 interacts with HER2/HER3 to initiate cancer cell proliferation, migration, and invasion signaling cascades, such as PI3K-Akt, FAK, and ras-Erk (Singh et al. 2007) (**Fig. 2**). It is hypothesized that the EGF domain present on the extracellular domain of MUC4 interacts with and activates HER family members. Recently, this interaction has demonstrated a novel role of MUC4 in neural invasion, by upregulating axonal guidance factor netrin-1 *via* HER2/AKT/NF-κB pathway in pancreatic cancer (PC) (Wang et al. 2015). Splice variants of MUC 4, i.e. (MUC4/Y), as well as the function of individual domains of MUC4, i.e. AMOP and NIDO, have received attention in recent studies. MUC4/Y overexpression demonstrated a higher invasion and tumorigenesis property of MUC4 in PC (Tang et al. 2016). MUC4 NIDO domains increased metastasis by interacting with ECM protein fibulin -2, whereas the AMOP domain facilitated angiogenesis by increasing VEGF-A. (Senapati et al. 2012;Tang et al. 2016) (**Fig. 2**).

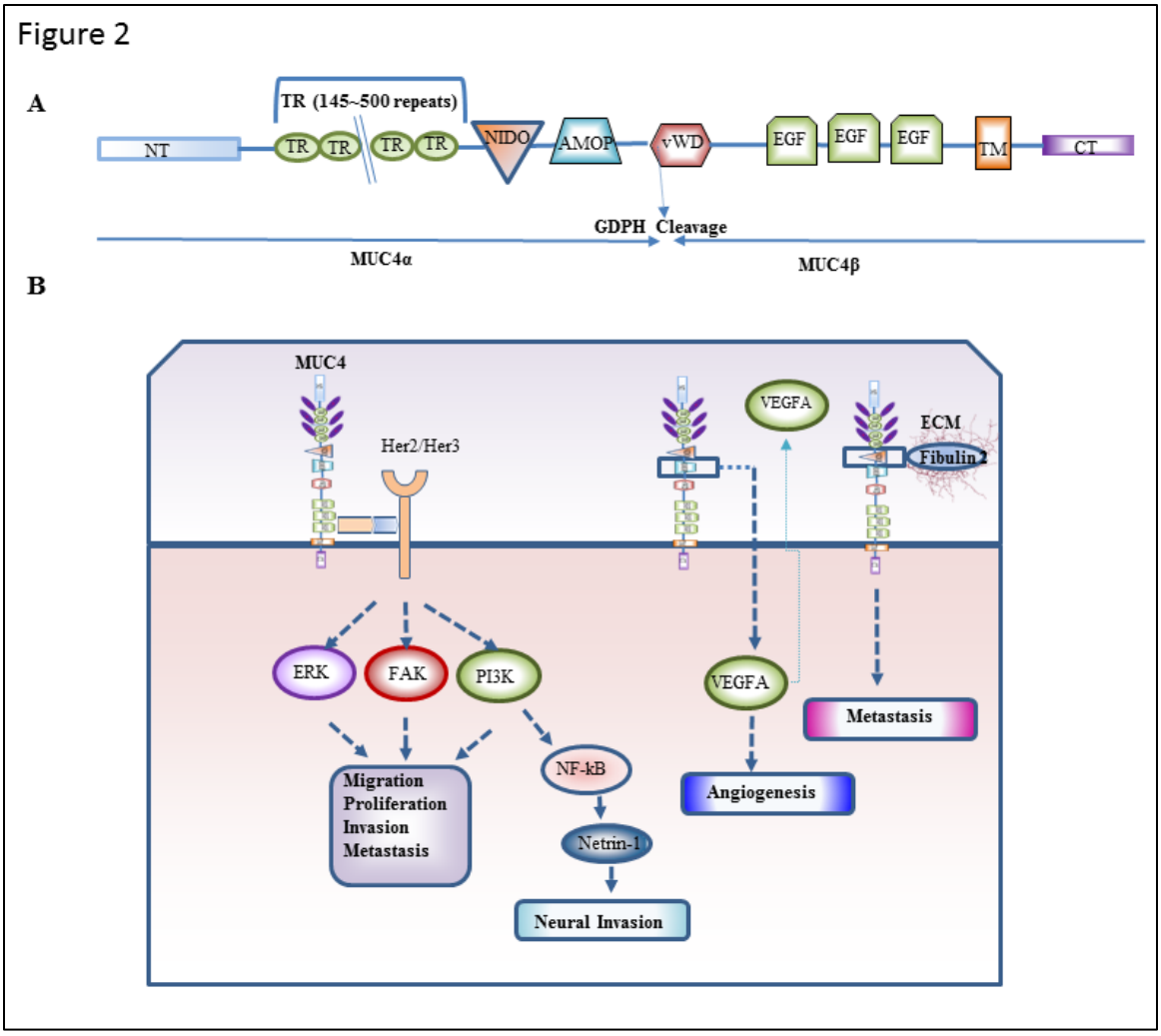


Figure 2: A: Schematic representation of the MUC4 structure. MUC4 is hypothesized to be cleaved at GDPH to generate two subunit- MUC4α and MUC4β. Extracellular MUC4α contains N-terminal region (NT), central tandem repeat domain (TR), NIDO, AMOP while MUC4β is comprised of vWD, EGF, TM, and CT. **B.** EGF domain in MUC4 is hypothesized to interacts with EGFR family member HER2 and initiate MUC4 facilitated downstream signaling including cell proliferation, migration, and invasion. AMOP domain of MUC4 has shown a major role in angiogenesis by increasing production of VEGF-A. Further, the nidogen (NIDO) domain have high homology with the nidogen-EGF domain of ancestral

nidogen proteins and interacts with fibulin-2 of extra cellular matrix (ECM) and facilitate metastasis of tumor cells. Abbreviations used: NIDO: Nidogen AMOP-Adhesion-associated domain in mucin MUC4 and other protein, vWD- von Willebrand D, VEGF-A-Vascular Endothelial Growth Factor –A, FAK-Focal Adhesion Kinase, PI3K-phosphoinositide-3-kinase, ERK-extracellular signal-regulated kinases, NF-kB -nuclear factor-kB.

MUC16: Elevated expression of transmembrane mucin MUC16 has been observed in benign and in malignant disease. MUC16 interaction has been observed with Siglec-9, Galectin1/3, Mesothelin, Selectin, and JAK-2 (Das and Batra, 2015). Binding of MUC16 with Siglec-9 present on Natural Killer (NK) cells exerts inhibitory signal which protects cancer cell from the cytotoxic impact of immune cells (Das and Batra, 2015). The interaction of N- and O- glycans of MUC16 with E-selectin on endothelial cells and L-Selectin on lymphocytes is hypothesized to facilitate tumor cell migration (Das and Batra, 2015) **(Fig. 3**). Mesothelin-to-MUC16 interaction is one of the most well-studied, shown to be critical for invasion and metastasis in various cancers, including pancreatic and ovarian. Mesothelin binding to MUC16 expressed by metastatic pancreatic cancer cells increases their motility and invasive potential by selectively upregulating MMP-7 through the upregulation of the phosphorylated p38 mitogen-activated protein kinase (MAPK)-dependent pathway in pancreatic cancer (Haridas et al. 2014). Moreover, MUC16 interaction has been observed with mesothelin. Galectin and showed tumorigenesis through FAK-mediated Akt and MAPK activation in PC

cells (Muniyan et al. 2016).The interaction between MUC16 and cytoplasmic JAK2 further leads to STAT-dependent upregulation of cyclin-D1 and results in cell proliferation (Das and Batra, 2015) (**Fig.3**).

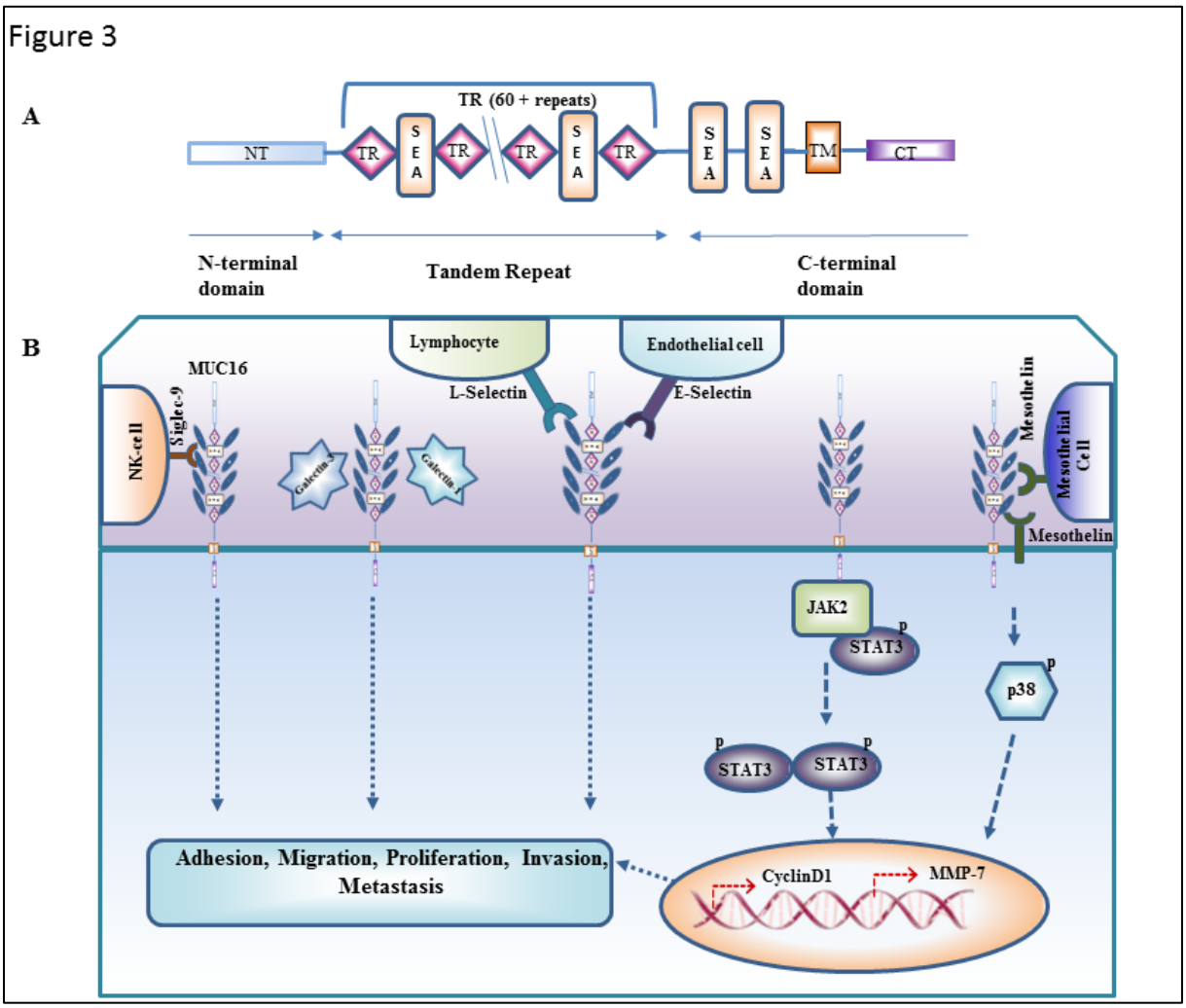


Figure 3: A: Schematic representation of transmembrane mucin MUC16 domains and structure. The N-terminus of MUC 16 is characterized by the presence of N-terminal region (NT), central tandem repeat domain (TR). The c-terminal has SEA domain, a single membrane-spanning transmembrane domain, and a cytoplasmic tail. The TR is interspersed by the SEA domain. **B** Extracellular portion of MUC16 interacts with Galectin-1, Galectin-3, Mesothelin, Siglec-9 (Natural Killer cells), E-

Selectin (endothelial cell), L-Selectin (lymphocyte) to initiate signaling cascade for cell migration and metastasis. The cytoplasmic tail of MUC 16 interacts with JAK2 to initiate STAT-3 dependent upregulation of cyclin D1 resulting in cell proliferation and metastasis. Abbreviations used: JAK2- Janus Kinase-2, STAT- signal transducers and activators of transcription, SEA: Sea urchin sperm protein-enterokinase-agrin; TM: Transmembrane; TR: Tandem Repeat Domain, MMP-7: matrix metalloproteinase-7.

MUC17: Among other transmembrane mucins, MUC17 has been the least studied. The carboxy terminal tail of MUC17 was shown to interact with the PDZ motif present on protein PDZK1 (PDZ domain containing 1), and it is believed to play a significant role in the proper localization MUC17. It is hypothesized that this interaction is essential for post-endocytic sorting of MUC17 to the membrane (Malmberg et al. 2008) (**Fig. 4**). MUC17 recombinant protein does not induce cell proliferation, however, but enhances cell migration and apoptosis by stimulating ERK phosphorylation, and this is hypothesized to play a therapeutic role in the restitution of epithelial cells after injury or inflammation (Luu et al. 2010). Sho Kitamoto and group demonstrated that MUC17 expression is enhanced by the HIF1α-mediated hypoxic response in pancreatic cancer (Kitamoto et al. 2012). A schematic representation of structure and cell signaling mediated by MUC17 is shown in **Figure 4.**

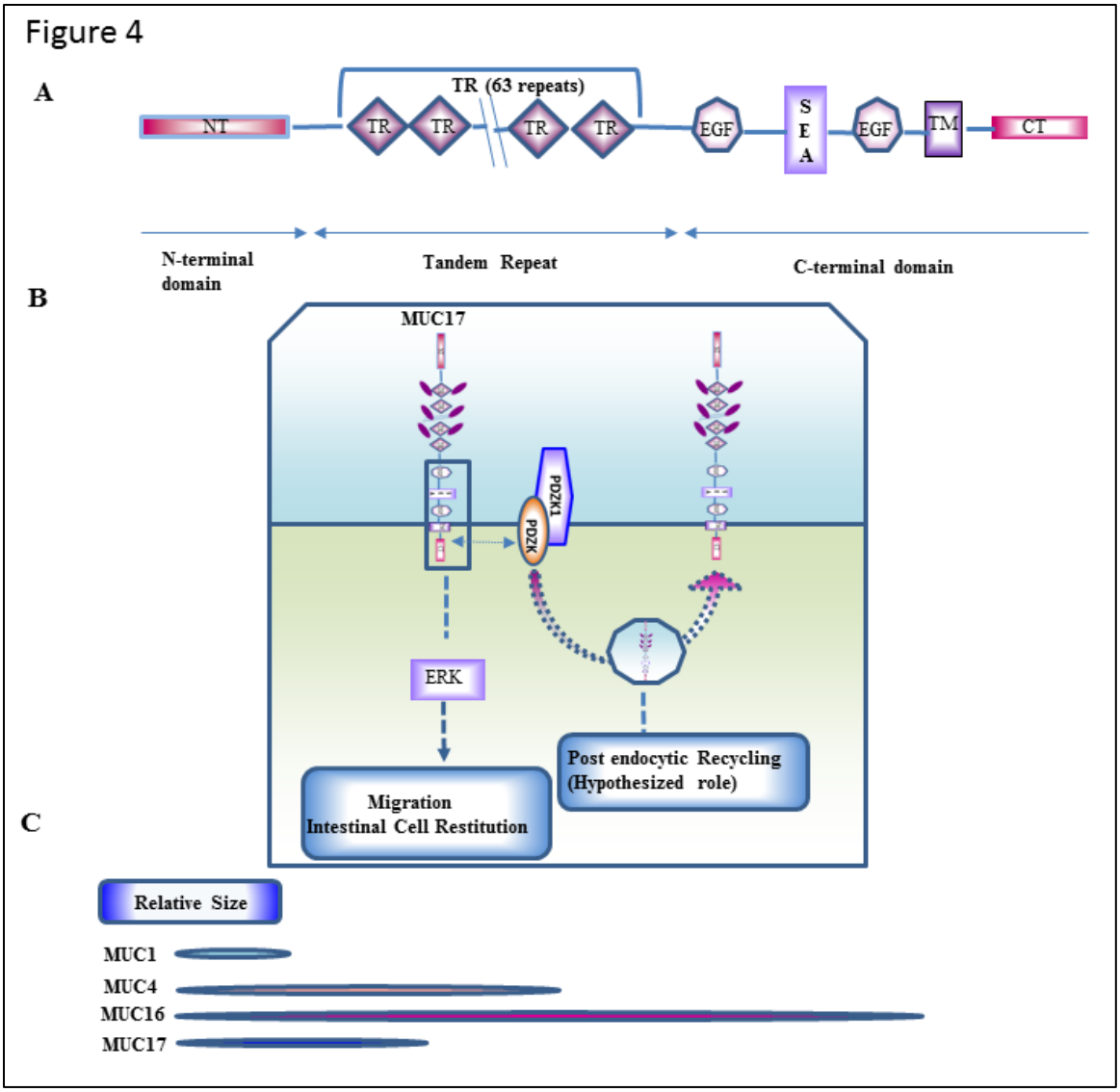


Figure 4: A: Schematic representation of the MUC17 structure. MUC17 contains N-terminal region (NT), central tandem repeat domain (TR) and C-terminal region. **B**. Carboxy terminal tail of MUC17 interact with PDZK1 (PDZ domain containing 1) which is hypothesized to play a role in post-endocytic recycling of MUC17 to the cell membrane. **C**. Representation of relative size of transmembrane mucins MUC1, MUC4, MUC16, and MUC17.

Table 3: Identified Transmembrane Mucin Interacting Protein Partner.

Human Mucin Gene	Mucin Interacting Partner
MUC1	EGFR, ICAM-1, Met, PDGFRβ, c-Src, β-catenin, APC,GSK3β, Erα, HSP70, HSP90, p53, NUP62, FGFR3, IKKβ, IKKγ, GRB2, ErbB2/HER2, ErbB3 and ErbB4
MUC4	ErbB2/HER2, HER3
MUC16	Galectin 1, Galectin 3, Siglec-9, Selectin, JAK2, Mesothelin, moesin
MUC17	PDZK1

Transmembrane Mucin Mouse Models

Joshi et al. have elegantly reviewed the Muc1 and Muc16 knockout (KO) mouse model **(**Joshi et al. 2015**)**. The phenotype observed for both mucin knockouts are detailed in **Table 4.** Of interest, human **MUC1** and murine *Muc1* have similarities in tissue expression patterns as well as sequence, highlighting the high significance of KO animal models for delineating the molecular mechanism and role of transmembrane mucins in disease and in normal physiology. Although an 87% sequence homology lies in non-tandem repeat domains, including the TM and cytoplasmic domains, deletion studies on the exon 2 of MUC1 (VNTR region) in mice have provided valuable insights into understanding the role of MUC1 in disease progression. Both human and mouse **MUC4** have 25 exons with conserved critical aa residues in the MUC4-specific domains, such as NIDO, AMOP, vWD, and TM. Not surprisingly, the highest variability found in tandem repeats both in a number of aa and sequences of aa (146-500 times of repeat of 16 AA in human TR while mouse Muc4 1 has TR of 24–126 aa with 20–21 repeat). Muc4 $^{-/-}$ did not show any abnormality in development, but assessment in a

colorectal cancer model delineated its impact on delayed cancer progression (Das et al. 2016). Exons 1 and 3 of human **MUC16** have sequence resemblance with two large exons of mouse Muc16 at the N-terminal region. The number of SEA domains varies in both murine and human Muc16 SEA domain in their extracellular (EC) region. Similar to MUC4 KO animals, Muc16 KO mice did not have any defect in normal development and reproduction of mice. **MUC17** is considered an ortholog of Muc3 because there is more similarity of human MUC17 with murine Muc3. 59.6 %, 52 %, and 63.5 % similarity has been found in the carboxyl terminal, first EGF, and second EGF domains of MUC17 with Muc3, respectively. To date, no KO mice have been developed for MUC17 to study its role and molecular signaling in different disease.

Mucin(s) as Therapeutic Target

Many groups have made MUC1 a strong focus as a potential therapeutic target because of its aberrant overexpression in cancer and its role in cancer progression, metastasis, poor response to therapy, and survival. Several therapeutic interventions have been developed to evaluate the role of MUC1 in therapeutics, mainly targeting MUC1-C, MUC1-N, and the junction of MUC1-C and MUC1-N subunits. C595, a murine MAb that recognizes a tetrapeptide motif (RPAP) of the MUC1 core protein, has shown the inhibitory effect on tumor growth and ascites production in the OVCAR-3 mouse xenograft model. Moreover, a combination of MAb C595 and docetaxel has been shown to inhibit intraperitoneal tumor growth while prolonging survival *in vivo,* suggesting the effectiveness of this combination as a potent therapeutic (Wang et al. 2011). Another anti-MUC1

antibody GP1.4 that recognizes four aa.(DTRP) motifs present in variable numbers of tandem repeats (VNTR) of the MUC1 extracellular region, activated the internalization of EGFR and suppressed proliferation and migration of pancreatic cancer cells in a MUC1-dependent manner (Hisatsune et al. 2011). Another cell surface-specific MUC1 antibody, DMC209, which exclusively recognizes the cleaved junction between MUC1α and MUC1β and is present on tumor cells but not on circulatory MUC1α, showed a tumor suppressive role in mice (Rubinstein et al. 2009). Considering the role of MUC1-C in metastatic progression, several peptide-based therapeutics have been developed to disrupt MUC1-C dimerization and its interaction with other protein-like β-catenin, Src, EGFR. The PTD-4 MUC1 inhibitory peptide (PIMP) decreased metastatic progression in a mouse model by blocking interaction between MUC1-CTand β-catenin. On the other hand, treatment with GO-201 that targets CQC of MUC1-C motif to inhibit oligomerization resulted in tumor regression in cancer, suggesting that targeting MUC1 CD could represent a potential cancer therapy (Pillai et al. 2015). Nano-capsulated GO-203, which also inhibits MUC1-C sub-unit homodimerization, has completed the phase 1 trial (Hasegawa et al. 2015). Antibodies that interrupt MUC16 and interaction with its binding partner can be a likely therapeutic target to control tumor progression. A human immunoadhesin, HN125, has been developed by Xiang and colleagues that competitively inhibits the interaction of MUC16 and mesothelin by binding with MUC16, and it has been shown to mediate antibody-dependent cancer cell death. Oregovomab and Abgovomab are two anti-MUC16 antibodies developed for ovarian cancer

immunotherapy, but in clinical trials these have shown disappointing results in cancer immunotherapy (Felder et al. 2014).Though the potential of MUC4 and MUC17 as a combined targeted therapy has been shown *in vitro*, neither any cancer treatments nor any clinical trials have been reported.

Table 4: Summary of phenotypic observation of Mucin knockout mouse models.

Human Mucin Gene	Mouse Homologue	Chromosomal location in mice	Deleted region to construct knock out mouse model	Phenotypic observation
MUC1*	Muc1	3F1	exon 2	**1**. Healthy appearance, no alterations in survival rate and fertility. **2**. Decreased cholesterol absorption in small intestine. **3**. Reduced intestinal mucus accumulation suggesting reduced intestinal obstruction. **4**. Demonstrated higher survival, reduced high-grade PanIN lesions, and reduced metastasis as compared to control mice when $Muc1^{-/-}$ mice were crossed with mice expressing constitutively active Kras in the pancreas. **5.** Elevated susceptibility to *Helicobacter pylori* infection.
MUC4**	Muc4	16B3	exons 2–6	**1**. Increased resistance to DSS-induced colitis. **2.** No histologic abnormality on colon. **3.** Delayed colorectal tumorigenesis progression after AOM/DSS treatment. **4.** Reduced infiltration of histocytes and T-lymphocytes and expression of inflammatory cytokines in the colon following DSS treatment.
MUC16*	Muc16	9A3	Major part of exon 3 and a portion of intron 3	**1**. No significant change in normal development and reproduction. **2**. Exhibited high reproductive ability in male Muc16−/− mice.
MUC17	Muc3a	5G2	Not established yet	Not Applicable

*(Joshi et al. 2015) ** (Das et al. 2016)

Perspective

This chapter reviews the structure, molecular evolution, and mechanism of action for transmembrane mucins under healthy and pathological conditions. MUCs are high molecular weight, heavily O-glycosylated glycoproteins, which are expressed in all epithelial tissues and protect against infection by forming a mucus barrier. While functioning as allies under normal physiological conditions, they act as adversaries in different disease states, especially in tumor progression. Although MUC1 is extensively studied regarding its structure and function, other TM mucins, such as MUC4, MUC16, and MUC17 are still in their infancy as to our understanding of their various characteristics such as posttranslational modification, differential glycosylation, and interacting protein partners etc. A systemic effort is required to identify the promoter region and splice variants of TM mucins to properly understand the mechanisms behind disease progression. The role of splice variants identified in MUC4 and MUC17 still requires exploration. Conditional Muc16, Muc4, and Muc3(MUC17) KO mice could be helpful in validating and defining the role of these mucins in oncogenesis in the background of $p53^{-/-}$, as could constitutively active Kras mice in future. MUCs have gained immense attention over the past decades regarding their role in cancer progression, and many promising *in vitro* and preclinical works are ongoing. However, more research and attention is needed regarding the biology and therapeutic aspects of MUCs, and these aspects must be addressed to fully understand the proper function of these complex, giant molecules.

Reference List

Apostolopoulos V, Stojanovska L, Gargosky SE. MUC1 (CD227): a multi-tasked molecule. Cell Mol Life Sci. 2015; 72: 4475-500. doi:10.1007/s00018-015-2014-z.

Bast RC, Jr., Feeney M, Lazarus H, Nadler LM, Colvin RB, Knapp RC. Reactivity of a monoclonal antibody with human ovarian carcinoma. J Clin Invest. 1981; 68: 1331-37.

Chaturvedi P, Singh AP, Batra SK. Structure, evolution, and biology of the MUC4 mucin. FASEB J. 2008; 22: 966-81. doi:10.1096/fj.07-9673rev.

Das S, Batra SK. Understanding the Unique Attributes of MUC16 (CA125): Potential Implications in Targeted Therapy. Cancer Res. 2015; 75: 4669-74. doi: 10.1158/0008-5472.CAN-15-1050.

Das S, Rachagani S, Sheinin Y, Smith LM, Gurumurthy CB, Roy HK, Batra SK. Mice deficient in Muc4 are resistant to experimental colitis and colitis-associated colorectal cancer. Oncogene. 2016; 35: 2645-54. doi: 10.1038/onc.2015.327.

Felder M, Kapur A, Gonzalez-Bosquet J, Horibata S, Heintz J, Albrecht R, Fass L, Kaur J, Hu K, Shojaei H, Whelan RJ, Patankar MS. MUC16 (CA125): tumor biomarker to cancer therapy, a work in progress. Mol Cancer. 2014; 13: 129. doi:10.1186/1476-4598-13-129.

Gum JR, Jr., Crawley SC, Hicks JW, Szymkowski DE, Kim YS. MUC17, a novel membrane-tethered mucin. Biochem Biophys Res Commun. 2002; 291: 466-75. doi:10.1006/bbrc.2002.6475.

Hanson RL, Hollingsworth MA. Functional Consequences of Differential O-glycosylation of MUC1, MUC4, and MUC16 (Downstream Effects on Signaling). Biomolecules. 2016; 6. doi:10.3390/biom6030034.

Haridas D, Ponnusamy MP, Chugh S, Lakshmanan I, Seshacharyulu P, Batra SK. MUC16: molecular analysis and its functional implications in benign and malignant conditions. FASEB J. 2014; 28: 4183-99. doi: 10.1096/fj.14-257352.

Hasegawa M, Sinha RK, Kumar M, Alam M, Yin L, Raina D, Kharbanda A, Panchamoorthy G, Gupta D, Singh H, Kharbanda S, Kufe D. Intracellular Targeting of the Oncogenic MUC1-C Protein with a Novel GO-203 Nanoparticle Formulation. Clin Cancer Res. 2015; 21: 2338-47. doi: 10.1158/1078-0432.CCR-14-3000.

Hisatsune A, Nakayama H, Kawasaki M, Horie I, Miyata T, Isohama Y, Kim KC, Katsuki H. Anti-MUC1 antibody inhibits EGF receptor signaling in cancer cells. Biochem Biophys Res Commun. 2011; 405: 377-81. doi: 10.1016/j.bbrc.2011.01.029.

Ho SB, Dvorak LA, Moor RE, Jacobson AC, Frey MR, Corredor J, Polk DB, Shekels LL. Cysteine-rich domains of muc3 intestinal mucin promote cell migration, inhibit apoptosis, and accelerate wound healing. Gastroenterology. 2006; 131: 1501-17.

Huang L, Chen D, Liu D, Yin L, Kharbanda S, Kufe D. MUC1 oncoprotein blocks glycogen synthase kinase 3beta-mediated phosphorylation and degradation of beta-catenin. Cancer Res. 2005; 65: 10413-22. doi: 10.1158/0008-5472.CAN-05-2474.

Joshi S, Kumar S, Bafna S, Rachagani S, Wagner KU, Jain M, Batra SK. Genetically engineered mucin mouse models for inflammation and cancer. Cancer Metastasis Rev. 2015; 34: 593-609. doi:10.1007/s10555-015-9549-1.

Kitamoto S, Yokoyama S, Higashi M, Yamada N, Matsubara S, Takao S, Batra SK, Yonezawa S. Expression of MUC17 is regulated by HIF1alpha-mediated hypoxic responses and requires a methylation-free hypoxia responsible element in pancreatic cancer. PLoS One. 2012; 7: e44108. doi:10.1371/journal.pone.0044108.

Kufe DW. Mucins in cancer: function, prognosis and therapy. Nat Rev Cancer. 2009; 9: 874-85. doi: 10.1038/nrc2761.

Luu Y, Junker W, Rachagani S, Das S, Batra SK, Heinrikson RL, Shekels LL, Ho SB. Human intestinal MUC17 mucin augments intestinal cell restitution and enhances healing of experimental colitis. Int J Biochem Cell Biol. 2010; 42: 996-1006. doi: 10.1016/j.biocel.2010.03.001.

Malmberg EK, Pelaseyed T, Petersson AC, Seidler UE, De JH, Riordan JR, Hansson GC. The C-terminus of the transmembrane mucin MUC17 binds to the scaffold protein PDZK1 that stably localizes it to the enterocyte apical membrane in the small intestine. Biochem J. 2008; 410: 283-89. doi: 10.1042/BJ20071068.

Moniaux N, Junker WM, Singh AP, Jones AM, Batra SK. Characterization of human mucin MUC17. Complete coding sequence and organization. J Biol Chem. 2006; 281: 23676-85. doi:10.1074/jbc.M600302200.

Moniaux N, Nollet S, Porchet N, Degand P, Laine A, Aubert JP. Complete sequence of the human mucin MUC4: a putative cell membrane-associated mucin. Biochem J. 1999; 338 (Pt 2): 325-33.

Muniyan S, Haridas D, Chugh S, Rachagani S, Lakshmanan I, Gupta S, Seshacharyulu P, Smith LM, Ponnusamy MP, Batra SK. MUC16 contributes to the metastasis of pancreatic ductal adenocarcinoma through focal adhesion mediated signaling mechanism. Genes Cancer. 2016; 7: 110-124. doi:10.18632/genesandcancer.104.

Obermair A, Schmid BC, Packer LM, Leodolter S, Birner P, Ward BG, Crandon AJ, McGuckin MA, Zeillinger R. Expression of MUC1 splice variants in benign and malignant ovarian tumours. Int J Cancer. 2002; 100: 166-71. doi:10.1002/ijc.10456.

Pillai K, Pourgholami MH, Chua TC, Morris DL. MUC1 as a potential target in anticancer therapies. Am J Clin Oncol. 2015; 38: 108-18. doi:10.1097/COC.0b013e31828f5a07.

Rahn JJ, Shen Q, Mah BK, Hugh JC. MUC1 initiates a calcium signal after ligation by intercellular adhesion molecule-1. J Biol Chem. 2004; 279: 29386-90. doi:10.1074/jbc.C400010200.

Ramasamy S, Duraisamy S, Barbashov S, Kawano T, Kharbanda S, Kufe D. The MUC1 and galectin-3 oncoproteins function in a microRNA-dependent regulatory loop. Mol Cell. 2007; 27: 992-1004. doi: 10.1016/j.molcel.2007.07.031.

Rubinstein DB, Karmely M, Pichinuk E, Ziv R, Benhar I, Feng N, Smorodinsky NI, Wreschner DH. The MUC1 oncoprotein as a functional target: immunotoxin binding to alpha/beta junction mediates cell killing. Int J Cancer. 2009; 124: 46-54. doi:10.1002/ijc.23910.

Senapati S, Gnanapragassam VS, Moniaux N, Momi N, Batra SK. Role of MUC4-NIDO domain in the MUC4-mediated metastasis of pancreatic cancer cells. Oncogene. 2012; 31: 3346-56. doi: 10.1038/onc.2011.505.

Shen Q, Rahn JJ, Zhang J, Gunasekera N, Sun X, Shaw AR, Hendzel MJ, Hoffman P, Bernier A, Hugh JC. MUC1 initiates Src-CrkL-Rac1/Cdc42-mediated actin cytoskeletal protrusive motility after ligating intercellular adhesion molecule-1. Mol Cancer Res. 2008; 6: 555-67. doi: 10.1158/1541-7786.MCR-07-2033.

Singh AP, Chaturvedi P, Batra SK. Emerging roles of MUC4 in cancer: a novel target for diagnosis and therapy. Cancer Res. 2007; 67: 433-36. doi: 10.1158/0008-5472.CAN-06-3114.

Tang J, Zhu Y, Xie K, Zhang X, Zhi X, Wang W, Li Z, Zhang Q, Wang L, Wang J, Xu Z. The role of the AMOP domain in MUC4/Y-promoted tumour angiogenesis and metastasis in pancreatic cancer. J Exp Clin Cancer Res. 2016; 35: 91. doi:10.1186/s13046-016-0369-0.

Tanida S, Mori Y, Ishida A, Akita K, Nakada H. Galectin-3 binds to MUC1-N-terminal domain and triggers recruitment of beta-catenin in MUC1-expressing mouse 3T3 cells. Biochim Biophys Acta. 2014; 1840: 1790-1797. doi:10.1016/j.bbagen.2014.02.008.

Wang L, Chen H, Liu F, Madigan MC, Power CA, Hao J, Patterson KI, Pourgholami MH, O'Brien PM, Perkins AC, Li Y. Monoclonal antibody targeting MUC1 and increasing sensitivity to docetaxel as a novel strategy in treating human epithelial ovarian cancer. Cancer Lett. 2011; 300: 122-33. doi: 10.1016/j.canlet.2010.09.013.

Wang L, Zhi X, Zhu Y, Zhang Q, Wang W, Li Z, Tang J, Wang J, Wei S, Li B, Zhou J, Jiang J, Yang L, Xu H, Xu Z. MUC4-promoted neural invasion is mediated by the axon guidance factor Netrin-1 in PDAC. Oncotarget. 2015; 6: 33805-22. doi:10.18632/oncotarget.5668.

Wei X, Xu H, Kufe D. MUC1 oncoprotein stabilizes and activates estrogen receptor alpha. Mol Cell. 2006; 21: 295-305. doi:10.1016/j.molcel.2005.11.030.

Zhang L, Vlad A, Milcarek C, Finn OJ. Human mucin MUC1 RNA undergoes different types of alternative splicing resulting in multiple isoforms. Cancer Immunol Immunother. 2013; 62: 423-35. doi:10.1007/s00262-012-1325-2.

Chapter 1(B): An overview on functional and mechanical role of Trefoil Factors in cancer

Abstract

Trefoil Factors 1, 2, and 3 (TFFs) are a family of small secretory molecules which function to protect gastrointestinal tract injuries by conferring epithelial mucosal repair. Besides weaving epithelial structural integrity, they are also involved in the signal transduction pathways for presiding cell migration, proliferation and invasion. Early studies using biochemical and genetic mouse model deemed TFFs as tumor suppressors; however, recent intriguing evidences from experimental and clinical studies have supported their role in the development and progression of many solid tumors. In addition, their clinical utility as potential biomarkers are emerging in various diseases including cancer. Although much research is focused to explore their role in gastrointestinal defense, their molecular and cellular mechanism of action as a disease facilitator and their binding partners remain largely unknown. Since trefoil factors are looming in cancer pathogenesis, so it is time to evaluate and appreciate the major studies to date and identify future prospective. We hereby review the recent advances of their involvement in precancerous lesions, tumorigenesis, and their emerging potential for subtyping various cancers. We will also provide a brief assemblage of their well-studied structure and expression status along with its mucin partner in different organs and lessons from mouse models.

Keywords: Trefoil Factor, TFF1, TFF2, TFF3, cancer, cancer subtype, cell signaling, biomarker, mouse model, precancerous lesion

Introduction

Trefoil factors ,comprised of three small secretory cysteine-rich peptides - TFF1, TFF2 & TFF3 are involved in mucosal repair and protect mucus epithelia from a variety of environmental insults. The gastrointestinal (GI) tract adopts an ultimate defense mechanism to combat unceasingly pathogenic and microbial attacks by integrating the epithelial production of abundant mucus and trefoil factors (TFFs) [1]. During different inflammatory conditions, TFFs are upregulated to impact cell-cell contacts and migration of cells to wounded area thereby helping wound repair [2]. Unfortunately, dysregulated expression of TFFs can result many pathological conditions including cancer. Although initially recognized as tumor suppressors, recent evidence across many solid tumors have demonstrated TFFs oncogenic role by increasing cell proliferation, invasion, apoptosis, migration and chemoresistance. However, even after four decades of their identification, the molecular landscape of each member of trefoil family is not fully understood regarding their involvement in tumor initiation and progression. Moreover, understanding of mechanical role of TFFs in precursor inflammatory stages are also in infancy. In this review, we have accentuated the recent findings on the molecular and clinical significance of TFFs in precancerous lesions along with tumor progression with detailed description of their well-studied structure, regulation and expression status in normal and diseases state.

Overview of Trefoil Factors structure and expression status

All three TFFs -TFF1, TFF2, TFF3 are contiguously encoded on human chromosome 21q22.3 clustered in a 55 kb region (**Figure 1a**) [3]. They are small

secretory molecules and share common features including N-terminal signal domain, at least one trefoil domain which also previously known as P-domain characterized by three pairs of conserved disulfide bond and C-terminal Cys-X-X motif (**Figure 1a**) [4]. Among all TFFs, TFF2 is the only member which contain two TFF domain. The clove-like structure of the trefoil domain, a domain which is defines as a conserved sequence of 42-43 amino acid, is formed by disulfide bridges at Cys1 - Cys5, Cys2 - Cys4, Cys3 – Cys6 configuration [13]. Trefoil domain of TFF1, TFF3 and second trefoil domain of TFF2 is consist of 42 amino acid whereas first trefoil domain of TFF2 is made of 43 aa. They all possess a signal secretion sequence which is removed by proteolysis during passage through the endoplasmic reticulum [5] .Secondary structure of TFFs are represented in **Figure 1b**.

TFF1 was identified as an estrogen-regulated gene previously known as Breast Cancer Estrogen-Inducible Protein (BCEI/pS2), pNR-2, HP1.A, D21S21. It was cloned from MCF7 breast cancer (BC) cell line cDNA in 1982 [6]. TFF1 is an 84 amino acid (aa) long peptide with approximate molecular weight of 9 kDa. TFF2, previously named spasmolytic polypeptide /SP/ SML1, was cloned in 1982 from porcine pancreas cDNA and contains four exons with peptide length of 129 aa and molecular weight of 14.28 kDa in size. Interestingly, porcine pancreatic spasmolytic polypeptide (TFF2)/(pSP) was the first of the trefoil factors to be discovered during an investigation of an insulin-poor side-fraction from the purification of porcine insulin [7] . It was named as “spasmolytic polypeptide “ because of its dose dependent inhibitory potential to contract intestinal pig ileum

[8].Though its higher expression and discovery in porcine pancreas anticipated much of its role in pancreas physiology, later on studies revealed that porcine is the only mammalian system with high pancretic TFF2 expression [9]. One of the unique characteristics of human TFF2 is that it contains N-linked fucosylated N,N'-diacetyl lactosediamine (LacdiNAc) oligosaccharide whereas porcine, murine and rat gastric TFF2 lacks an N–glycosylation site [10]. Notably, glycosylated recombinant TFF2 was also demonstrated to be more effective in reducing gastric damage *in vivo* than the non-glycosylated form of TFF2 [11]. The last member identified in this family is TFF3 (previously called intestinal trefoil factor, ITF or hP1.B) which was discovered from the rat cDNA sequence in 1991 [12] and is primary expressed by goblet cells of the large and small intestines [13]. In 1996, their previous nomenclature was changed to the present nomenclature at Philippe Laudat Conference on trefoil/P-domain peptides [9].

While the specific intramolecular disulfide bonds make TFFs very resistant to high heat and trypsin, chymotrypsin and carboxypeptidases, the unpaired seventh cysteine of TFF1 and TFF3 mediates their homo or heterodimerization [5]. Specifically, Cys-58 residue of TFF1 mediates homodimerization and also heterodimerization between Cys38 of brichos domain of TFZI1 [14]. TFF1 forms either homo or hetero dimers, but TFF3 is mostly observed as a monomer with few exceptions as a dimer in colonic tissue and gastric mucus [15]. It is reported that gastric TFF1 heterodimerizes with gastrokine-2/TFIZ1, while colonic TFF3 mostly hetero-dimerizes with mucus-associated IgGFc-binding (FCGBP) protein [15]. Gastrokine-2/TFIZ1 expression also decreases during gastric cancer progression

similar to TFF1 [16, 17], suggesting their cooperative role in gastric cancer progression. Recent studies reported the presence of both TFF3 monomers and TFF1-TFF3 heterodimers in small intestinal specimens from patients with Crohn's disease. [18]. Interestingly, parenteral administration of TFF3 monomer aggravated the colitis but intracolonic administration of TFF2/TFF3 dimers improved experimentally induced colitis in rats [19]. Structurally, both monomer and homodimer of TFF1 is more asymmetric and exhibit extended conformation than TFF3 due to possessing very positively charged carboxy terminus of TFF1 as compared to TFF3 [20]. In case of TFF2, unlike TFF1 and TFF3 homo dimer, the two trefoil domains are connected by a disulphide bond, between Cys6 to Cys104, with an additional peptide chain from Leu50 to Asp56 which make TFF2 more compact than the dimer forms of TFF1 and TFF3 (ref).

TFF1 is mainly expressed in the superficial and foveolar epithelium of the stomach and is also expressed in the brain, liver, kidney, respiratory tract and the salivary glands. In contrast to diffused expression of TFF1 in the stomach, TFF2 is mainly expressed by neck of the stomach cells and Brunner's gland of proximal duodenum in the distal stomach [8]. TFF3 is majorly expressed by the mature goblet cells in the small and large intestine TFF3 whereas no expression was detected in liver, pancreas, or stomach in rodents [21]. Expression of TFF2 and TFF3 was also reported in lymphoid tissues, such as spleen, thymus, lymph nodes, and bone marrow, and they stimulate migration of monocytes [22]. All three trefoil peptides are expressed in fetal organs such as fetal stomach and duodenum but not in placental or fetal membranes, though increased levels of serum level of

TFF2 and TFF3 during pregnancy suggest their potential role in fetal development [23]. Moreover, orthologues of these TFFs are found in other animals such as rats, dogs, cats, cows, wolves, rhesus monkeys, short-tailed opossum, sheep, chimpanzees, and pigs. Organ specific expression of individual TFF are listed in Table 1 [23].

Table 1: Expression status of Trefoil Factors in normal physiology

Trefoil Factor	Expression in normal Physiology	Ref.
TFF1	• Stomach-Gastric foveolar cell, surface epithelial cell, gastric mucosa cells of the pyloric part • Brunners gland-Upper duct • Efferent tear ducts • Salivary gland (submandibular and parotid gland), • Small and large intestine and rectum- Goblet cells • Mucous acini of sublingual gland • Weak expression in bronchial goblet cells of the pseudostratified epithelia in respiratory tract • Vagina- Submucosal gland • Trachea-Submucosal (seromucous) glands • Absence in acinar cell of pancreas, normal ductal epithelium • Conjunctival goblet cell • Breast: Focal expression in duct luminal cell • Gall Bladder: Patchy epithelial expression	[24-32]

TFF2	• Stomach-Gastric mucosa cells of the pyloric part • Esophagus-Submucosal glands in the • Sublingual gland-Weak expression of mucous acini • Duodenum-Brunner's gland acini and distal ducts • Gall Bladder- Focal expression in duct epithelium	[24, 29, 30]
TFF3	• Brain and spinal cord- oxytocinergic neurons • Pituitary Gland-Hypothalamus • Eye-Human efferent tear ducts, conjunctivital goblet cell, cornea, retina • Epithelium of nasolacrimal duct • Vagina-Submucosal gland • Stomach-Faint staining expression in the upper part of the gastric mucosa • Small and large intestine and rectum- goblet cells for the • Esophagus-submucosal glands • mucous acini of sublingual gland • Bronchial goblet cells of the pseudostratified epithelia in the respiratory system • Conjunctival goblet cells • Surface epithelium of the endocervix and in gland-like structures of the cervical epithelium • Uterus • Gall bladder-patchy epithelial expression	[24, 29, 33-35]

TFFs in gastrointestinal defense

TFF peptides has been termed as luminal surveillance peptides by Playford in 1995 as they aid to epithelial repair mechanism in collaboration with mucin by firming/stabilizing the mucus layer [36]. Numerous studies using cell line, human tissues and mouse model demonstrated the protective implication of TFF peptides in the GI tract.

Evidence from mouse model demonstrated that lacking any TFF in mouse model featured an increased susceptibility to damage in injury models which potentially indicates that TFFs role in protecting the gastrointestinal tract. Moreover, TFF1 is an essential component for normal differentiation of the antral and pyloric gastric mucosa as well as for retaining gastric and stomach regular structural architecture. Lefebvr O. et al. demonstrated that homozygous mutation of TFF1 in mice displayed noticeable gastric mucosa abnormalities with severe hyperplasia and dysplasia and developed antropyloric adenomas [23]. Moreover, Tff1-/- mice displayed enlarged villi in the small intestine and serrated, thickened stomach wall with protruding nodules [23, 24]. In case of Tff2-/- knockout mice, they were fertile with no apparent gastrointestinal abnormalities however, they increased acid secretion and were prone to non-steroidal anti-inflammatory drug injury [25]. Likewise, Tff3-null mice also did not manifest any unusual characteristics but are prone to dextran sulfate sodium-induced colonic damage [37]. Taken together these studies indicate that, TFFs are integral component for stomach and protect GI tract from environmental or chemical insult. Recently, Thiem S et al. developed an unique tri-transgenic Tg(Tff1-CreERT2;Tff2-rtTA;Tff3-

Luc) mouse model has been developed employing both tamoxifen-inducible Cre recombinase (CreERT2) and the doxycycline-inducible reverse tetracycline transactivator (rtTA) [38]. In this tri-transgenic mouse model, doxycycline-induced Tff2 positive gastric cells demonstrated higher expression of gastric stem cell marker such as Lgr5. Lrig1, Troy, Mist1 than tamoxifen induced Tff1 positive isolated stomach cells suggesting role of Tff2 in the gastric stem cell. This mouse model will be beneficial to study TFFs role in organ development and disease progression as it permits consecutive genetic perturbation of all the TFF in the same subset of epithelial cells.

To enhance mucosal defense, TFF-mucin is hypothesized to work together. Gouyer et al. intially reported co-expression of TFF1 and TFF3 in a mucin layer of colon carcinoma cell line HT-29 MTX which secretes mucins and expresses a goblet cell-like phenotype [44]. TFF1 was demonstrated to physically interact with the carboxy terminus of MUC2 and MUC5AC through the von Willebrand factor C cysteine-rich domains in murine stomach and duodenum [43]. Moreover, copper dependent dimerization of TFF1 leads binding with MUC5AC in the mucus of colorectal cancer cell [39]. Additionally, TFF1 is found to be indispensable for H. Pylori binding with MUC5AC cell explain the tropism of this bacteria for gastric tissue as a defense mechanism [15]. Interestingly, TFF3 dimer is more compact and less asymmetric structure than TFF1 monomer and it is speculated that as more The variable distance and orientation of the TFF1 binding sites offer more interaction with mucin [5]. Likewise, TFF2 interaction with α1-4-linked GlcNAc capped MUC6 (possibly through non-covalent lectin and covalent disulfide

bridges) was demonstrated to inhibit the growth of H. Pylori [45]. Moreover, TFF2 also modulates the visco-elasticity of gastric mucins present in the solution into a gel-like state [46, 47]. Notably, change in viscoelastic nature of mucins was found to depend on the state of TFFs as monomers or dimers. TFF3 monomer was able to modulate the mucin viscosity while TFF1 and 2 monomers failed to change it [47]. Systemically administered radiolabeled TFF2 secreted in gastric lumen and increased viscosity of gastric mucus however radiolabeled TFF3 is secreted by mucous neck cells without alterations in the viscosity [22]. Because of their colocalization with mucins in different organs and differential binding to gastric mucins, TFFs have been hypothesized as a "link peptide" for stabilization of the gastric mucus and thought to impact rheological properties [44]. These observations cumulatively highlighted the importance of TFFs in mucus formation possibly through physical interaction with mucins for protection of gastrointestinal tract [44]. Though it is a long-standing hypothesis that TFF-mucin interaction protects gastrointestinal mucosa from various environmental insults, the importance of this interaction and concomitant expression has not been extensively explored in many diseases including cancers.

For rapid epithelial restitution after acute injury, migration of healthy cells to wounded area as well as removal of damages cells is required not only for preventing bacterial access but also to prevent fluid and electrolyte loss. Numerical studies indicated that TFF peptides are “mitogen” and promote GI epithelial restitution is through accelerating cell migration and inhibiting apoptosis of the migrating cells. Though exact mechanism is not clearly understood, however,

TFFs has crucial role in disrupting cell-cell adhesion, modifying cytoskeleton, and signal transduction to promote cell migration and prevent apoptosis which are needed during epithelial restitution. TFF3 has shown diverse role in regulating cell -cell junction. TFF3 is involved in degradation of E-cadherin as well as disrupting ecadehrin/β-catenin complex to promote cell migration by disturbing cell-cell junction, whereas, it inhibits platelet activating factor mediated induction of claudin -1 and ZO-1 to decrease trans epithelial leakage suggesting TFF is not only responsible for epithelial restitution but also promote barrier function via regulating cell-cell junction [40-42].

Transcriptional and epigenetic regulation of Trefoil Factors

The expression of TFFs is tightly controlled in normal gastro physiology, but evasion of this regulation in disease conditions results in their aberrant expression. Promoter analysis has revealed that 5' flanking sequence of the TFF1 gene contains many enhancer elements responsive to tissue plasminogen activator TPA, EGF, H-RAS, c-jun and estrogen responsive element (ERE) [43]. In addition, the TFF1 promoter also contain AP-1 [(a complex of JUN, FOS, ATF (activating transcription factor) and MAF (musculoaponeurotic fibrosarcoma) protein families)] binding motif and regulates TFF1 by 2,3,7,8-Tetrachlorodibenzo-p-dioxin-mediated inhibition of 17 beta-estradiol-induced TFF1 expression [44, 45]. Similar to TFF1, 5’ flanking region of TFF2 contains several putative binding sites for Myc, Ets-like factor and PEA3 [3].In addition to that, putative NFκB binding sequence GGGA/G(C/A/T)TT/CT/CCC has been found in all TFF1-3 promoters [46]. Interestingly, all TFFs has the ability to upregulate their expression.

In situ analysis suggested binding of Sp1 and Sp3 to the TFF1 promoter which mediates recruitment of HDACs and histone acetyltransferases (HATs) to promote acetylation of histones in TFF1 promoter [47]. In addition, C/EBPβ also binds to the 5'-flanking region and negatively regulates TFF1 expression in gastric cancer [48, 49]. In contrast to TFF1, the TFF2 promoter contains a CCAAT/enhance binding proteins (C/EBPs) site as a cell line-specific cis-acting regulatory element and positively regulates TFF2 expression in MCF7 cells [50].Goblet cell specific transcription is of TFF3 is possibly through “goblet -cell response element” (GCRE) that is present in proximal TFF3 promoter [51].

Table 2: Regulation of Trefoil Factors

<table>
<tr><th></th><th colspan="2">Regulator</th><th>Remarks</th></tr>
<tr><td rowspan="2">TFF1</td><td>UP</td><td>HNF-3α and β, GATA6, STAT6, PPARγ, Gastrin, TFF1</td><td rowspan="2">I. HNF-3α and β increases TFF1 expression in pancreatic and gastric cancer cells with with no effect on intestinal cell lines [52]
II. Methylation of GATA6 promoter in gastric cancers suppressed TFF1 and 2 through the activation of STAT3 [53].
III. Gastrin Responsive Element is present in 300 base pairs upstream of the transcriptional start site of TFF1 which is critical for gastrin-dependent activation of TFF1. [54].
IV.TFF1 can autoactivate its own expression TFF1 responsive element was present between −583 and −212 bp of its promoter [55].</td></tr>
<tr><td>Down</td><td>Foxo3a, nucleophosmin/B23, C/EBPβ, NFκB, IL-6, IL-1β</td></tr>
</table>

			V.Foxo3a binds to the Forkhead responsive sequence of the TFF1 promoter and negatively regulates its expression in BC cell [31]. VI. In endometrial cancers, a nucleolar phosphoprotein (nucleophosmin/B23) and a transcriptional repressor of ERα have been shown to negatively regulate TFF1 expression through a complex formation with activator protein-2γ (TFAP2C/AP2γ) [36].
TFF2	Up	CCAAT/enhance binding proteins (C/EBPs), GATA-6	I.TFF2 promoter contains a CCAAT/enhance binding proteins (C/EBPs) site as a cell line-specific cis-acting regulatory element and positively regulates TFF2 expression in MCF7 cells [39]
	Down	IL-1β, IL-6, IFN-γ	
TFF3	Up	STAT3, IL-4 and IL-13	I. In silico analysis has revealed the presence of a STAT3 binding site in the TFF3 promoter region and positive regulation of TFF3 [40]. II. IL-4 and IL-13 upregulates TFF3 transcription which is STAT6-dependent [56].
	Down	IL-1β, IL-6, C/EBPβ	

Altered epigenetic programming which includes aberrant DNA methylation and histone modifications is one of the leading factors for regulating tumorigenesis [57]. Transcriptional regulation of TFFs via epigenetic modifications in cancer has been studied by multiple groups. In gastric cancer, epigenetic silencing of TFF1 expression by promoter methylation in the CpGs at −20 bp and −457 bp upstream of transcription start site is considered as one of the mechanisms for gastric tumorigenesis [58]. Likewise, TFF2 silencing due to aberrant promoter hypermethylation is also reported in gastric cancers [59]. While the promoter

region in TFF1 and TFF3 was hypomethylated in prostate cancer (PCa) patient samples and PCa cell lines, compared to benign prostatic hyperplasia, TFF1 and TFF3 were overexpressed in PCa and the TFF2 promoter was hypermethylated in PCa cells with no protein expression [60]. Similarly, TFF3 promoter hypomethylation was observed in another study using a large cohort of PCa patient samples, suggesting epigenetics as one of the mechanisms for TFF3 regulation [61]. However, the study also observed that decreased levels of TFF3 mRNA correlated with a higher Gleason score, advanced stage and early PSA recurrence, suggesting a negative correlation with advanced disease of prostate cancer in contradiction with the findings where TFF3 was shown to be tumorigenic in prostate cancer [61, 62]. A hypomethylated TFF2 promoter was also observed in a panel of PC cell lines as well as in micro dissected ductal cells from pancreatic tumors with no expression in normal pancreatic epithelium. Further, treatment of non-expressing PC MIAPaCa cells with 5Aza-dC and TSA restored TFF2, demonstrating its regulation by promoter methylation [63]. Similarly, the hypomethylated TFF1 promoter is reported in estrogen receptor-positive BC tumors which have an estrogen responsive element binding site [64]. A recent *in silico* analysis from 45 BC cell lines revealed TFF1 as one of the top epigenetically controlled genes [65].

Trefoil Factors and Precancerous Lesions

According to Basil Morson, “a precancerous lesion is a histological abnormality in which cancer is more likely to occur” [66]. Epithelial malignancies

are progressive (multi-factorial) accumulation of genotypic and phenotypic changes which start from metaplastic lesion resulting from chronic inflammation to precursor intraepithelial neoplasia to invasive cancer. Both epithelial and inflamed cells has deregulated expression of secreted soluble factors and cytokines which alter key development genes, genomic and epigenomic profile of stem cells leading to the development of metaplasia [67]. Many appealing pieces of evidence suggest role of TFFs in precancerous lesion which is often originated from inflammation. For example, gastric cancer which starts from a long-lasting inflammatory conditions triggered by Helicobacter pylori infection leading to gastritis with subsequent metaplasia and finally steps into invasive stage. In the stomach, TFF1 typically is expressed in partially committed pre-pit cell progenitor in the oxyntic gland of the stomach which has both morphological and functional feature of stem cell [68].

Interestingly, deletion of TFF1 demonstrated increased gastric pit length and reduced acid-secreting parietal cells suggesting its direct contribution in proliferation and pre-pit to pre-parietal commitment programming which leads to more metaplastic state. Thus, TFF1 knockout (KO) mouse models represents precancerous hyperplastic lesions in early progenitor/stem cell in the oxyntic region of stomach and their critical role in gastric stem/progenitor cells leading to gastric cancer (**Figure 2a**) [69]. Though all TFFs are differentially expressed in a unique manner in mouse oxyntic mucosa of the stomach, deficiency of TFF1 also modulates the expression of TFF2 and TFF3, and a probable reason is that the genes are contiguous in short DNA segment [69].

Loss of parietal cells is also an early event which leads to the development of spasmolytic /TFF2 expressing metaplasia (SPEM), an antral phenotype lineage that is responsible for gastric repair in response to acute toxic injury or *Helicobacter* infection and considered as an essential step of pre-neoplastic gastric lesion as well [70, 71]. One of the molecular characteristics of SPEM is co-expression of TFF2 in mRNA and protein level, intrinsic factor (IF) and GSII lectin. Unique co-expression of TFF2 with *denovo* expression CD44 was noticed in SPEM of young mice but was deficient in aged mice, indicating a possible relationship of TFF2 and CD44 exclusively in SPEM development early aged mice [70]. On the other hand, Quante M. et al. demonstrated TFF2 mRNA transcripts but not the protein is present in progenitor cells within the stem cell region in the gastric isthmus and potentially indicates origin of glandular progenitors [72]. However, these progenitor cells are not responsible for giving rise TFF2- expressing SPEM lineage, instead trans-differentiating chief cells which not express TFF2 typically are partially accountable for giving rise SPEM suggesting how a change in microenvironment which leads to gastric atrophy induce TFF2 expression [72]. As SPEM is positively correlated with the presence of both dysplasia and gastric adenocarcinoma, attention is required to understand the role of TFF2 in *H. pylori*-associated SPEM development and subsequent progression to gastric cancer.

Very recently, Strobel et al. identified a novel and unique pancreatic ductal progenitor niche coined as pancreatic duct gland (PDG) which has demonstrated a substantial impact on pancreatic ductal healing after acute inflammatory injury [73]. Restitution after injury in pancreatic ducts is imparted by TFF1 and TFF2

which promote cells to migrate from PDG to pancreatic ductal epithelium [74]. This PDG gland morphologically distinct with gland-like outpouches or coiled glands mainly found within the mesenchymal cuff of surrounding large ducts and expresses gastric mucins (MUC6), and TFF2 and stem cell progenitor markers such as Shh, Pdx-1 and Hes-1. While PDG can form both panIN and IPMN, loss of TFF2 from the PDG compartment has a pronounced effect on the progression and development of intraductal papillary mucinous neoplasia (IPMN) which has characteristic ductal papillary structures and cystic metaplasia in this novel progenitor glands [75] (**Figure 1b)**.

In case of oesophageal cancer, Barrett's esophagus (BE) is a known intestinal metaplasia which progresses into oesophageal cancer [76]. Morphologically, in BE, normal squamous epithelium is replaced by a metaplastic columnar epithelium as a consequence of chronic gastroesophageal reflux disease [77]. BE shows all the hallmarks of stem cells such as self-renewal, multipotent differentiation, and commitment to the respective lineages from which they were derived. Moreover, their stem cell molecular signature is distinct from those of the esophagus or the stomach [77]. Barret's stem cells express MUC2 which is a marker of goblet cells and expressed by intestinal stem cells, and also expresses MUC5AC and TFF2 which are not expressed by intestinal but gastric stem cells [77]. *In vivo* xenograft studies showed that the mutational profile of the esophageal tumor generated from transformed Barret's stem cells are distinct from transformed esophageal stem cells, supporting the notion that the origin of the two distinct forms of esophageal cancer (squamous cell carcinoma and intestinal

adenocarcinoma) are indeed from two different stem cell population. It would be interesting to study the role of TFF2 in BE to understand the origin of esophageal carcinoma as it can arise from Barret's as well as esophageal stem cell.

In colon cancer, serrated polyps are known heterogeneous precursor lesions characterized histologically by infolding of colonic crypt epithelium which attributes "saw-tooth" appearances. They are classified into three groups: hyperplastic polyps (HP), sessile serrated adenoma/polyp (SSA/P), and traditional serrated adenoma (TSA). 30-35% of colorectal cancer developed from an unorthodox serrated pathway which carries similar molecular features to classical colorectal carcinoma progression model such as BRAFV600E and DNA hypermethylation [78, 79]. TFF1 and TFF2, along with MUC2 and MUC5AC, have been shown to be differentially upregulated in sessile serrated adenoma which is in line with earlier published reports [80, 81]. In future, TFF1 and TFF2 knockout mouse model would be logical step for identifying the mechanistic role of TFFs in the serrated pathway and possible relationship with BRAFV600E.

An autocrine role for TFF1 has been shown to be involved in differentiating Clara cells to goblet cell and earlier studies demonstrated that extensive goblet metaplasia harbors precancerous molecular signature for lung cancer. However, role of TFF1 in goblet cell metaplasia and their direct involvement in lung cancer is not established yet [82, 83]. Much surprisingly, there is not much exploration of TFF3 in precursor lesion of cancer specifically gastric cancer as it is highly expressed in gastric intestinal metaplasia, a precancerous lesion of intestinal type

of gastric cancer as well as in pancreatic cancer [84].TFF3 knock out mouse model would be a great model to start to identify its role in precancerous lesion of cancer.

Functional and mechanical role of Trefoil Factors in cancer

TFF1: TFF1 is observed to serve both tumor suppressive as well as oncogenic role by modulating different cancer-associated trademarks. Aberrant overexpression of TFF1 has been observed in prostate (PCa) and pancreatic cancers (PC) but is downregulated in gastric cancer (GC). Recent studies suggest that peroxisome proliferator-activated receptors (PPAR) δ inhibits TFF1 expression resulting in decreased PCa tumor growth [85]. Using the inducible TFF1 expression system, it was further demonstrated that induction of TFF1 resulted in increased colony formation and tumor growth of PCa. These results were in agreement with earlier observations by Radiloff et al. who demonstrated that loss of TFF1 expression in PC and PCa cells decreased cell viability and reduced tumor xenograft growth in mice by inducing senescence [86, 87]. Overexpression of TFF1 also enhanced PCa invasion and metastasis by repressing E cadherin expression and strengthening the mesenchymal phenotype [88]. In mucinous ovarian carcinoma, TFF1 promoted cell invasion by regulating cyclin D1 and c-Myc overexpression through Wnt-β catenin signaling [89]. TFF1 also increased angiogenesis through COX2 and thromboxane A2 receptor (TXA2-R) [90, 91]. TFF1 can also auto induce its own expression both in normoxia and hypoxia which is responsible for invasion and EMT [55].

While these studies showed the tumorigenic activity of TFF1, employing *in vitro*, *in vivo* and mouse models, other studies reported its tumor suppressive activities in gastric and retinoblastoma cancers. As mentioned earlier, TFF1 KO mice accelerated GC tumorigenesis [92]. Soutto et al. demonstrated that TFF1 negatively regulated c-Myc and cyclin D1 expression by inhibiting nuclear localization of β-catenin [93]. While TFF1 overexpression was shown to inhibit TNF-α mediated binding of Tumor necrosis factor receptor 1 (TNFR1) to TNF Receptor Associated Factor 2 (TRAF2) and subsequent NF-kB activation, Cobler et al. indicated that TNF-α mediated NF-κB activation also results in TFF1 downregulation to promote gastric tumorigenesis [92, 94]. These findings prompted investigators to suggest that loss of TFF1 mediated NFkB activation upregulated proinflammatory and antiapoptotic genes and resulted in neoplastic transformation in gastric cancer [92]. Also, loss of TFF1 in GC also increased tumor invasion and was associated with worse survival outcome, particularly in patients who have undergone curative surgery without adjuvant therapy. A recent study further illustrated that TFF1 KO mice have increased infiltration of T lymphocytes and monocytes in gastric tissues, increased inflammation and suppressed cytotoxic response by CD4+Th17 and CD8+Tc17 cells respectively [95]. In case of breast cancer, the contradictory role of TFF1 has been proposed. Buache et al. demonstrated TFF1 knockdown favored BC tumorigenicity *in vitro* and *in vivo* suggesting its tumor suppressive role in BC [96]. Further studies suggested that 7,12-Dimethylbenz[a]anthracene (DMBA) treatment to TFF1 KO mice resulted in the development of aggressive mammary gland tumors compared

to wild-type mice [96]. While all these studies indicate a tumor suppressor role of TFF1 in BC, with a stark contrast, its oncogenic role in BC has also been demonstrated by increased the expression levels of cell cycle-regulatory molecules cyclin D1 cyclin E1 along with CDK2 and the transcription factors like c-MYC, c-JUN and c-FOS, the components of the activating protein-1 (AP-1). [97]. The overall, schematic of TFF1 mediated cell signaling is shown in **Figure 2**.

TFF2: Like TFF1, TFF2 also demonstrated divergent role on tumorigenicity with significant implications on cellular proliferation, apoptosis and cell migration. In breast cancer, TFF2 exerted anti-apoptosis and migratory potential towards BC cells [98, 99] . In addition, TFF2 promotes cell proliferation of cholangiocarcinoma cells by activating EGFR/MAPK pathway [100]. While the TFF2 expression (wild type) is associated with overall poor survival, the ratio of splice variant of TFF2 (Exon 2 deleted) and wild-type TFF2 is associated with better prognosis of patients with cholangiocarcinoma [101]. Microarray analysis of tumors in APC revealed significant upregulation of TFF2 and TFF1 in the enlarged intestinal tumor and TFF2 overexpression promoted intestinal adenocarcinoma cell proliferation *in vitro* and tumor growth *in vivo*. Likewise, overexpressing TFF2 in colon cancer increased cell proliferation and tumor size *in vitro* and *in vivo* [102].

In contrast with the tumorigenic role, several lines of evidence reported a tumor suppressor role of TFF2 in gastric cancer and its negative correlation with advanced tumor stage, metastasis and reduced disease-free survival (DFS) [103]. Genomic analysis revealed promoter hyper-methylation as a possible mechanism for TFF2 silencing in both GC patients and cell lines models [59]. Moreover, Cai Y.

et al. demonstrated the interaction of TFF2 with specificity protein/Krüppel-like factor (SP/KLF) transcription factor family member Sp3 in the cytoplasm which is responsible for the anti-tumor activity of gastric cancer cells [104]. Subsequent studies also suggested that TFF2 as the downstream target of p53 regulation and its overexpression by mutant $p53^{R175H}$ induced an anti-apoptotic response with increased migratory potential of gastric cancer cells [105]. In addition, recent studies from Pdx1-Cre; LSL-$KRAS^{G12D}$-TFF2-knockout mouse model provided evidence of formation of high grade pancreatic intraepithelial neoplasias and intraductal papillary mucinous neoplasm compared to control mice. Furthermore, when TFF2 was overexpressed in pancreatic cancer cell, it corroborated with in vivo findings by abrogating pancreatic cancer cell proliferation by activating SMAD4 [75].

Though cell surface receptors for TFF2 are not well documented, transmembrane G protein-coupled receptors such as CXCR4 and Protease-Activated Receptor (PAR) are being implicated as signaling receptors for TFF2. Recent studies have implicated low-affinity TFF2-CXCR4 interaction in Ca^{2+} signaling activation in lymphoblastic Jurkat cells that are accompanied by increased cell proliferation [106]. This signaling activation was abrogated by receptor desensitization pre-treatment with the specific antagonist AMD3100 or an anti-CXCR4 antibody [106]. In colorectal cancer (CRC) patients BRAFV600E mutation is correlated with higher TFF2 and CXCR4 levels in sessile serrated polyps, a point mutation found in ~10% CRC patients [107], suggesting the TFF2-CXCR4 axis as a novel therapeutic avenue for CRC cases originating from

serrated polyps [108]. The TFF2-CXCR4 axis was also shown to activate MAPK signaling in GC and PC cells [106]. In addition, overexpression of TFF2 also enhanced pancreatic β-cell proliferation by activating ERK1/2 activation and upregulation of cell cycle regulatory proteins [109].

Likewise, another G-protein coupled receptor, Protease-Activated Receptor (PAR), came in the limelight for signaling receptor of TFFs. PAR is implicated in GI tract barrier function and is found in pancreas, colon, small intestine and stomach [110]. TFF2 isolated from frog skin secretions have been shown to induce human platelet aggregate through direct interaction with PAR1 [111]. Moreover, TFF2 and PAR4 overexpression is observed in CRCs and their interaction increases motility and invasive behavior of CRC cells *via* MAPK/ERK pathway activation [111, 112]. Schematic representation of TFF2 mediated cell signaling has been portrayed in **Figure 3**.

TFF3: TFF3 has been demonstrated to be tumorigenic in breast, brain, cervical, hepatocellular, pituitary, papillary thyroid carcinoma**.** In breast cancer, TFF3 expression was positively correlated with aggressive clinicopathological parameters and worse survival. Moreover, *in vitro* and *in vivo* studies suggested that overexpression of TFF3 resulted in increased cell proliferation and invasion by enhancing phosphorylation of c-SRC along with STAT3 activation [113]. Elevated expression of TFF3 has also been correlated with higher grade glioma and poor clinical outcome [114]. Knockdown of TFF3 in glioma cell lines reduced cell proliferation, migration and invasion and increased apoptosis and slowed tumor progression by downregulating AKT phosphorylation, MMP-9, cleaved

caspases and dependent of HIF 1α [114]. Analysis from four publicly available genomic datasets revealed TFF3 as a distinguished upregulated molecule in prostate cancer [139]. On this note, overexpression of TFF3 in prostate cancer resulted in increased tumorigenicity and resistance to ionizing radiation [62]. In a similar vein, repression of TFF3 induced mitochondrial apoptosis by decreasing BCL2 along with elevating cytochrome C ,Smac/DIABLO and mitochondrial translocation of BAX in prostate cancer suggesting antiapoptotic role of TFF3 in prostate cancer [115]. Recently, Rickman DS et al . demonstrated that TFF3 overexpression expression is dependent on the status of ERG rearrangement in castration-resistant prostate cancer which is a frequent event that is characterized by a rearrangement between the 5' regulatory elements such as TMPRSS2, SLC45A3, and NDRG1 of androgen-regulated gene and the coding region of ETS gene family of transcription factor.Thus, overexpression of TFF3 is responsible for ERG rearrangement mediated castration resistant prostate cancer cell invasion [37]. In cervical cancer, Yuan et al. demonstrated forced expression of TFF3 increased cell proliferation and invasion likely through STAT3-mediated suppression of the epithelial E cadherin coding gene CDH1. TFF3 silencing also increased the sensitivity of cervical cancer cell lines to etoposide chemotherapy and induced apoptosis by modulating apoptotic molecules such as BCL-2 and Bax [116]. Additionally, chemoresistance in cervical cancer cells by TFF3 is partly mediated by overexpression of P-glycoprotein, an efflux transporter commonly known as multiple drug resistance (MDR-1) gene [116]. In hepatocellular carcinoma, elevated expression of TFF3 was associated with shorter relapse-free

survival and poor survival outcome. Furthermore, overexpression of TFF3 enhanced hepatocellular carcinoma cell proliferation by modulating cell cycle, apoptotic and EMT regulating genes [117]. Furthermore, TFF3 is also involved in hepatocellular chemoresistance by aggravating the cancer stem cell phenotype and stimulating stem cell survival through phosphorylation of AKT at Ser473 and elevated BCL-2 expression [117]. In pituitary adenoma, enhanced expression of TFF3 protein was documented and siRNA-mediated knockdown of TFF3 protein expression induced tumor cell apoptosis via modulation of the mitochondrial pathway, suggesting its tumorigenic role in pituitary cancer [118]. Wu J. et al. indicated that knockdown of TFF3 in the papillary thyroid carcinoma cell line TPC-1 lowered cell proliferation, invasion and migration and slowed tumor growth and volume by downregulating AKT phosphorylation, MMP-9 and BCL-2 [119].

While all these above mentioned studies reflect oncogenic role of TFF3, in oppose to that, tumor suppressive role of TFF3 has been revealed in colorectal cancer and retinoblastoma. Uchino H. et al. demonstrated that overexpression of TFF3 reduced cellular growth *in vitro* and *in vivo* in colon cancer cells [120]. Similarly, In retinoblastoma, forced expression of TFF3 decreased cell viability, proliferation and growth while significantly increasing cell death in retinoblastoma cells [121].

Like TFF2, CXCR4 has also been suggested as a potential receptor for TFF3. X-ray structure-based modeling between TFF3 and CXCR4 demonstrated high-affinity interaction mostly with the extracellular loops and partly in the pore region of the CXCR4 structure [122]. Interestingly, cell proliferation through TFF3

was independent of CXCR4 or CXCR7, however, migration was dependent on these two chemokine receptors in human conjunctival epithelial (HCjE) cells [122]. It will be fascinating to dissect molecular mechanism of TFF and G-Protein couple receptors to identify TFF mediated signaling which will be useful for therapeutic targeting. Schematic representation of TFF3 mediated cell signaling has been portrayed in **Figure 4**.

Surprisingly, while TFF1 and TFF2 is well studied in gastric cancer, till date no mechanical and functional studies has been performed to dissect role of TFF3 in gastric cancer though there is report on elevated expression of TFF3 has correlation with poor prognosis of gastric cancer patients [135]. Similarly to date, role of TFF3 in PC has been elusive though there has been report of their upregulation in mRNA level in these lethal form of cancer [123]. Similarly, differential TFF3 expression in follicular adenomas and follicular thyroid carcinomas using fine-needle aspirates has also been reported but still there is a gap in information regarding their mechanistic and functional role in this cancer [141].

Trefoil Factors and non-coding RNAs in cancer

Emerging evidence support that non-coding RNAs, specifically microRNA (miRNA) and long non-coding RNAs (lncRNA) played as a frontier in regulating diverse functions of oncogenes or tumor suppressors [124]. A recent study has reported the presence of many murine miRNAs (mmu-miRNA) interspersed in the promoter regions CDS, 3’ UTR and 5’UTR of the murine TFF gene cluster (chromosome 17) [A.A. Shah, N. Blin]. These include mmu-mir-770-5p, mmu-miR-

107, mmu-mR-103, mmu-miR-324-3p, mmu-miR-693-5p and mmu-miR-383 suggesting positive correlation for regulation of TFF(s) by miRNA under normal and pathological conditions [A.A. Shah, N. Blin]. Further crosstalk between miRNA, epigenetic control and TFFs are implicated as a mechanism of gastric tumorigenesis [119]. For example, the loss of TFF1 resulted in demethylation of HOXA10 promoter causing reduction in its expression that subsequently reduces miR-196b-5p expression and in turn promoting gastric tumorigenesis [125, 126]. However, molecular mechanism of TFF1 mediated alteration in HOXA10methylation requires further investigations. Soutto et al. observed that TFF1 induces p53 expression and suppress gastric tumor growth by down-regulating **miR-504** [127, 128]. Likewise, overexpression of TFF1 in retinoblastoma cell lines induced p53 expression by downregulating **miRNA-18a** (encoded by the miR-17–92 clusters) and promoted caspase-dependent apoptotic cell death [129, 130]. Overexpression of TFF3 downregulated tumor suppressor miR-491-5p (the mature form of miR-491) and resulted in nuclear accumulation of Psoriasis-susceptibility-Related RNA Gene Induced by Stress (PRINS), a lncRNA involved in mediating inflammation and cancer in colon cancer cases. Nuclear translocated PRINS and pro-apoptotic Phorbol-12-myristate-13-acetate-induced protein 1 (PMAIP1) in nucleus results in evasion of apoptosis [114, 115].

In addition to regulation of miRNA expression by TFFs, miRNA mediated alteration in expression is also observed by multiple groups. Indeed, overexpression of **miRNA 423-5p** downregulated TFF1 in gastric cancer cells which resulted into increased cell proliferation and metastasis by increasing cyclin

D1, cyclin D3 and β-catenin expression particularly in absence of TFF1 suggesting miRNA-423-5P mediates downstream signaling via negatively regulating TFF1[131]. In gastric cancer cells, tumor suppressor miR-30 was observed to negatively regulates HNF4γ expression in conjunction with reduced expression of VIL1 and TFF3 (intestinal metaplasia markers) as well as SPEM marker (TFF2) expression in GC cells. Therefore, reduced miR-30 expression mediated upregulation of HNF4γ, TFF2 and TFF3 and was responsible for inducing early metaplasia to invasive GC progression. Recently, Guo et al. found TFF3 as a novel target for **miR-7-5p** and via repressing TFF3 expression along with decreasing PI-3K/AKT signalling, **miR-7-5p** mediated decreased proliferation and migration of colorectal adenocarconima LS174T cell [132]. On the same track, suppression of miR-7-5p expression by glycolytic inhibitor 2-Deoxy-d-Glucose upregulated TFF3 through activating PI3K/AKT signaling in GBM cell line suggesting TFF3 indeed a target for miR-7-5p [133]. Though small steps has been made to explore the relation between TFF1 /TFF3 and miRNA regarding tumorigenesis, TFF2 -miRNA axis in cancer pathogenesis has not been yet explored. As interplay between miRNA-transcription factor is well established and TFF is emerging as novel targets for miRNA,it would be interesting to identify the crosstalk between transcription factor and miRNA which epigenetically regulate TFFs in cancer using genome wide computational high throughput analysis, identifying TFF centric mirNA-transcription factor hub and subsequent validation [134].

Trefoil Factors and cancer subtyping: Clue for therapeutic stratification

Exploration of genomic and transcriptomic technologies has reached a new pinnacle not only for unwinding the molecular complexity of disease progression, but also their role in precision care by stratifying patients based on the molecular subtype of cancers. Major cancers are now being defined as subtype specific because of tumor heterogeneity, though identification of patient and treatment interventions based on subtype are still in early stages. TFFs have been recognized as a part of a distinct gene signature in molecular subtyping in several cancers.

Most recently, Moffitt et al. have identified classical and basal-type subgroups of PC based on transcriptomic analysis from human pancreatic tumor tissues, mouse tissues and cell line. Interestingly, TFFs were overexpressed in classical subtype of PC [135]. In addition, Collisson et al. identified TFF1 and TFF3 overexpression in classical subtype of PC [135-138]. They further reported gemcitabine resistance as an important characteristic of the classical subtype of PC [136].

In BC, four major subtypes have been identified with distinct molecular signatures-Luminal A, Luminal B, Her2+ and Triple negative which further classified into basal, claudin low [139, 140]. The ER-negative basal-like subtype which has poor survival rate is represented by high expression of keratins 5 and 17, laminin, metallothionein 1X and fatty acid binding protein 7 whereas the ERBB2+ subtype is characterized by high expression of several genes in the ERBB2 amplicon at 17q22.24 including ERBB2 and GRB [141]. ER-positive

(predominant) luminal A subtype demonstrates better survival outcome with a distinct characteristic expression of the ER α gene, GATA binding protein 3, X-box binding protein 1, TFF1, TFF3, HNF3, and estrogen-regulated LIV-1. Luminal subtype C also shares molecular feature from both luminal A and luminal B which is a feature they share with the basal-like and ERBB2+subtypes [139, 142]. More interestingly, Heiser et al. have demonstrated that chemotherapeutic agents have displayed preferential sensitivity in different subgroups. For instance, luminal cells demonstrated favorable sensitivity to HDAC inhibitor vorinostat while lapatinib (EGFR inhibitor) has displayed a stronger association with HER2 positive than the luminal subtype. On the other hand, the basal subtype showed better response to etoposide, docetaxel and cisplatin treatment [143]. When superpathway network analysis was performed to identify the mechanism of subtype-specific chemotherapeutic response, upregulation of FOXA2 which controls TFF1 in luminal A subtype was observed [143].

Anguraj et al. defined six subtypes of CRC including stem-like, inflammatory, cetuximab-resistant transit amplifying, cetuximab sensitive transit amplifying, goblet-like and enterocyte type with a proposed candidate biomarker as well as a probable guide to therapeutic agents for these subtypes. The gene signature for goblet-like subtype with better disease-free survival was demonstrated to express MUC2 and TFF3 expression [144]. Mucinous colorectal carcinoma, a histological subtype of CRC, has characteristic upregulated KRAS and BRAF mutations and promoter hypermethylation which correlate with poor survival and worsened therapy response [145, 146]. Intriguingly, microarray and

pathway analysis showed augmented expression of TFF2 and MUC1 in this mucinous subtype of CRC, implying the importance of TFFs in this lethal subtype [147]. While all the accumulating studies suggest TFFs' participation in subtyping of cancer, there is no further study in any of the cancer how they modulate subtyping characteristics along with drug tolerance.

Prospect of TFF as biomarker in cancer

Several computational and meta-analysis studies have suggested the role of TFFs as potential biomarkers in many diseases including cancers [148].Serum analysis of TFF1, TFF2 and TFF3 from 94 breast cancer patients and 84 healthy individuals suggested that TFF1 and TFF3, but not TFF2, are higher in women with BC, and interestingly combination of all TFFs has demonstrated an area under curve (AUC) value of 0.96 suggesting their potential as breast cancer diagnosis [149]. A urinary biomarker panel including LYVE-1(Lymphatic vessel endothelial hyaluronan receptor 1), REG1A (Regenerating Family Member 1 Alpha) and TFF1 demonstrated promising diagnostic accuracy to detect early stage PC from healthy urine with an AUC value of 0.90 and 0.93 in both training and validation datasets, respectively. Moreover, the panel resulted into a higher AUC value of 0.991 with a specificity/sensitivity (SN/SP) of 94.4/100 after addition of CA19.9 to the panel as compared to CA19.9 only (AUC value of 0.88 and SN/SP 83.1/92.9) to differentiate PDAC stage I &II from healthy control [150]. In addition, higher levels of TFF3 correlated with poor overall survival of CRC patients and performed with better sensitivity (76.29%) than CEA (72.16%) and CA19.9 (46.39%) while similar

specificity (approximately 97%) to detect distant metastasis [151]. Elevated TFF3 serum levels were observed in lung cancer (n=130) patients who didn't go for treatment including surgery, chemotherapy and radiotherapy compared to normal controls (n=60), however, the study didn't further explore the diagnostic potential of TFF3 alone or in combination with other two TFF-TFF1 and TFF2 to diagnose lung cancer [152]. In gastric cancer, serum level of TFF2 and TFF3 was significantly elevated in cancer patient as compared to with or without H. pylori infection control group. However, further analysis from these 183 patients serum before treatment and 45 healthy controls demonstrated that TFF3 performed better as a diagnostic marker with SN/SP (80.9%/81.0%) using a cutoff of 3.6 ng/mL, while the pepsinogen test, major gastric cancer screening modality in japan, had SN/SP (44.8%/87.4%). Furthermore, combination of TFF3 measurement with pepsinogen I/II testing may provide a screening modality with increased sensitivity [84]. With a very similar finding from serum analysis of 72 gastric cancer patients and 37 healthy controls, Huang Z et al. concluded that TFF3 and pepsinogen has better diagnostic potential than either TFF3 or pepsinogen [153].TFF3 is also elevated in patients who responded to endocrine therapy and exhibited more specificity and sensitivity with a SN/SP 91%/ 79% as a predictive biomarker than progesterone receptor (SN/SP-91%/69%), TFF1(SN/SP), and estrogen receptor (SN/SP 69%/72%) in unstratified metastatic breast cancer patients In addition, elevated expression of TFF1 and TFF3 in serum from untreated and hormone-refractory PCa patients with advanced disease compared to the localized tumors were also reported [154].

Conclusion and Future Directions

TFFs mediate signaling through various pathways, so it is not surprising that they will work through multiple receptors and to therapeutically target TFF, identification of TFF receptors is of high importance. More efforts should be focused on identifying signaling receptor for trefoil factors using both binding and functional assay approach to confirm binding partner of TFF peptide and whether the binding partner is required for TFF mediated downstream signaling. Though efforts have been made to identify TFF putative receptors, most have been inconclusive [155-157]. Similarly, kinetics studies confirmed a TFF3 specific receptor on intestinal epithelial cells but was unable to identify the receptor [158]. As mucins are highly correlated with TFFs and several tyrosine kinase receptors, especially EGFR, TFF may be an interlinking agent between mucins and tyrosine kinase receptors. Recently, GPCR coupled receptors has been reported as a receptor for TFF2 and TFF3, however, CXCR-TFF1 axis is yet to be uncovered. Particular interest should be focused on identifying the active domain(s) of CXCR4 which interact with TFF1.

From a biomarker point of view, TFFs are contextual in cancers. Nonetheless, their prominence as biomarkers is undeniable. Their expression in precursor cancer lesions is one of the attractive characteristics for why they are being investigated as potential biomarkers. Trefoil factors are being proposed as intervention treatment for gut inflammatory diseases due to their high potential. However, we believe like May et al. that without proper understanding of their role

in cancer progression, their binding receptor and their function as autocrine or paracrine mechanisms, their potential as a biomarker will be imperceptible [159].

In our review we have mentioned the role of trefoil factors in precancerous lesion, however, tailoring the treatment of precancerous lesions according to each individual case is a multidisciplinary challenge, which involves expertise from radiology, pathology, surgery and oncology. Moreover, treating precancerous lesion is still under debate because of the higher risk of recurrence or progression or unnecessary overtreatment. Though TFF has been emerged as potential marker for many precancerous lesion especially gastric and pancreatic cancer, their usefulness in translational science is not explored yet.

The divergence of evidence on the role of trefoil factors as tumor suppressors or tumor promoters has questioned how they will be useful as a therapeutic intervention. TFF1 as a pharmacological treatment on gut inflammatory disorders, notably the following chemo- or radio-therapies, is being investigated [160, 161]. Based on the tumor suppressive role of TFF1, oral intake or local administration of TFF1 peptide has been proposed to combat tumor invasion of gastric cancer [162]. Contrarily, targeting TFF1 has been proposed as a therapeutic target for mammary carcinoma [97]. In a phase II randomized, double-blind, placebo-controlled study, it was demonstrated that both low (10 mg/mL) and high dose (80 mg/mL) of recombinant human TFF3 oral spray was safe and effective to reduce chemotherapy-associated oral mucositis in colorectal cancer [163]. While it is proposed as a therapeutic alternative in colorectal cancer, studies also suggest that TFF3 is responsible for increased migration, invasion,

anti-apoptosis, and tumorigenic potential in colon cancer [164, 165]. Therefore, as TFFs are critical molecules in preserving epithelial integrity in the gastrointestinal tract, more precise and cautious study is warranted to elucidate at which level they become tumorigenic. As they are also involved in precursor lesions, their molecular involvement in the initiation of precursor lesions needs to be explored and exploited in strategies for preventing cancers arising from metaplastic precursors.

Another interesting future direction would be to elucidate TFFs role in glycosylation. TFF2 is mostly N-glycosylated at N-15 and has lectin binding property critical for barrier function in *H. pylori* infection, the exploration of the role of TFF1 and TFF3 in glycan modification is appealing as the single trefoil domain of TFF2 was demonstrated to be responsible for inducing LacdiNAc, an indicator of protein specific GalNAc transfer by the glycosyltransferases4GalNAc-T3/T4 [166, 167].

Lastly, the implicit question remains what is the role of TFFs in stratifying patients for precision therapeutic intervention? Does their importance lie only in identification of “subtyping” the patients or further “pharmacotyping” patients based on subtype specific gene signatures along with high throughput drug assay? What will be the consequence if we inhibit subtype-specific genes like TFFs which are upregulated genes in classical subtype PC which has shown high gemcitabine resistance and displayed erlotinib sensitivity? It is currently unknown whether their inhibition will increase gemcitabine sensitivity or erlotinib resistance. Undoubtedly, a deeper understanding of TFF in subtype-specific cancer will be a great addition into stratified medicine. Sophisticated genomics-based assays and patient-derived

subtype specific organoids will be useful starting models to answer these pertinent questions which would provide a unique opportunity to employ personalized therapeutic interventions. For instance, in CRC, TFFs are implicated in serrated subtype which harbored BRAFV600E-mutations. Thus, BRAFV600E-mutant organoids can be used as a valuable *in vitro* model for identifying TFFs role in serrated adenomas. Although, organoids are great future models for identifying patient specific medicine, establishing organoids from patient biopsies, performing gene sequencing, analyzing drug sensitivity with high throughput systems and ultimately making a clinical decision requires at minimum an 8 week time frame, a long window for starting therapy for a swiftly lethal disease like PC [168]. In this case, we propose to first perform next-generation sequencing on patient biopsies to identify cancer subtypes and then to treatment based on pre-identified potential drugs for that gene signature. In this respect, understanding the role of subtype-specific gene in drug sensitivity is indispensable.

Overall, in this review, we have provided a comprehensive overview of the trefoil factor family with anticipation to facilitate new thoughts and connecting ideas that could assist in elucidating TFF function.

Figure 1a

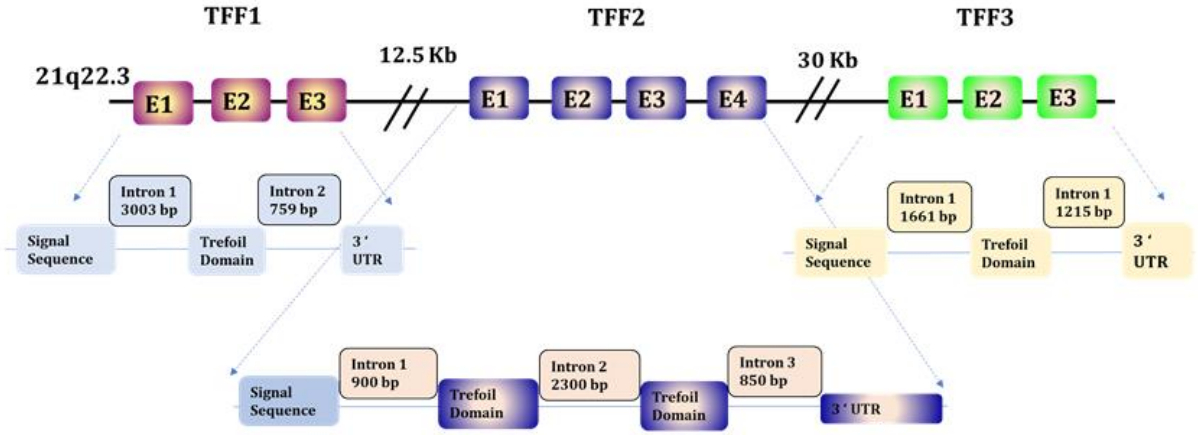


Figure 1b

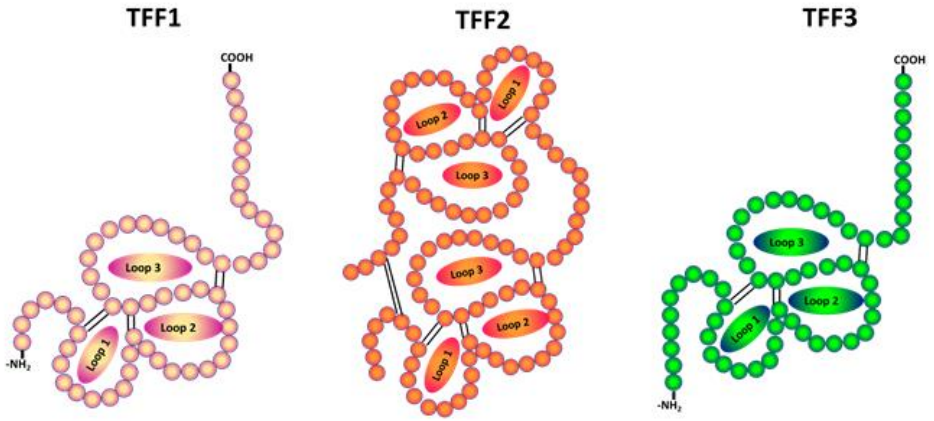


Figure Legends

Figure 1: (a) Schematic diagram of genomic organization of TFF1, TFF2 and TFF3. All TFF is clustered in 21q22.3. TFF1 and TFF3 contain one trefoil domain whereas TFF2 contain two trefoil domains. **(b) Schematic diagram of secondary structure of TFF1, TFF2 and TFF3.** Loops are indicated as shown in the picture.

Figure 2a

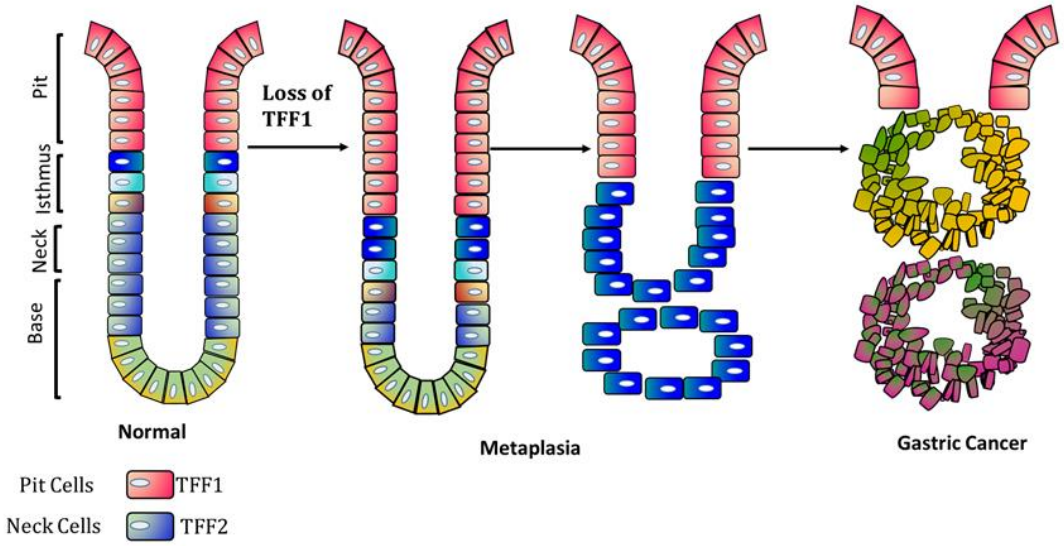


Figure 2b

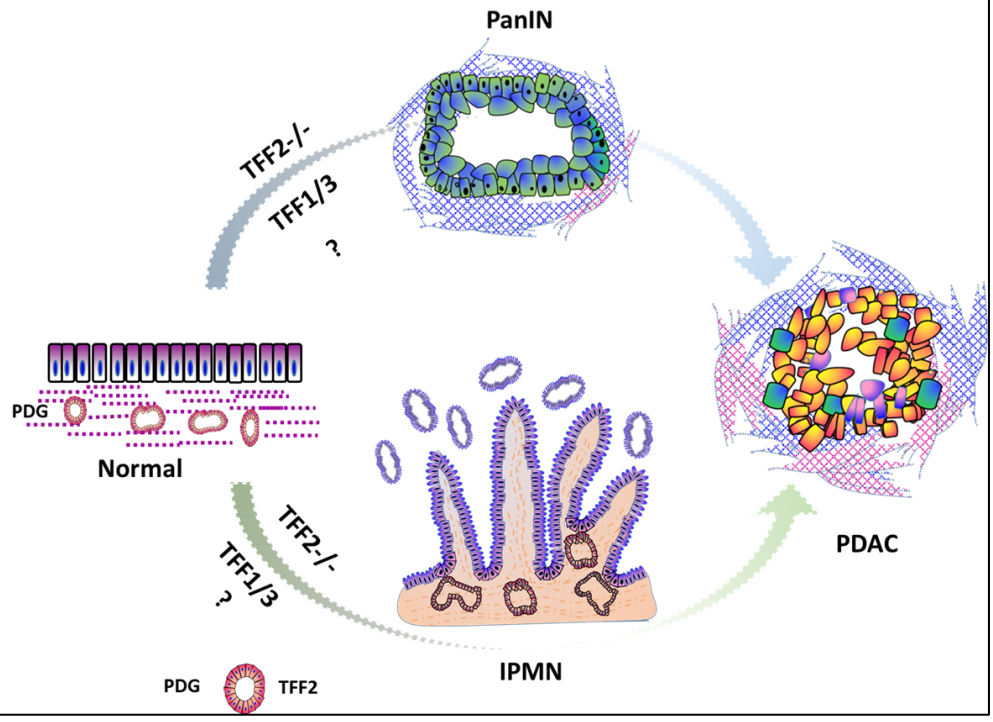


Figure 2 : Schematic diagram depicting influence of TFF in precancerous stage which ultimately progress in cancer. (**a**) Loss of TFF1 accelerates gastric cancer progression by increasing the length of pit of gastric which leads to metaplasia and subsequent gastric carcinogenesis. **(b)** Loss of TFF2 from the pancreatic duct compartment (PDG) can lead to development of IPMN and PanIN type precancerous lesion which ultimately results into pancreatic cancer.

Figure 3

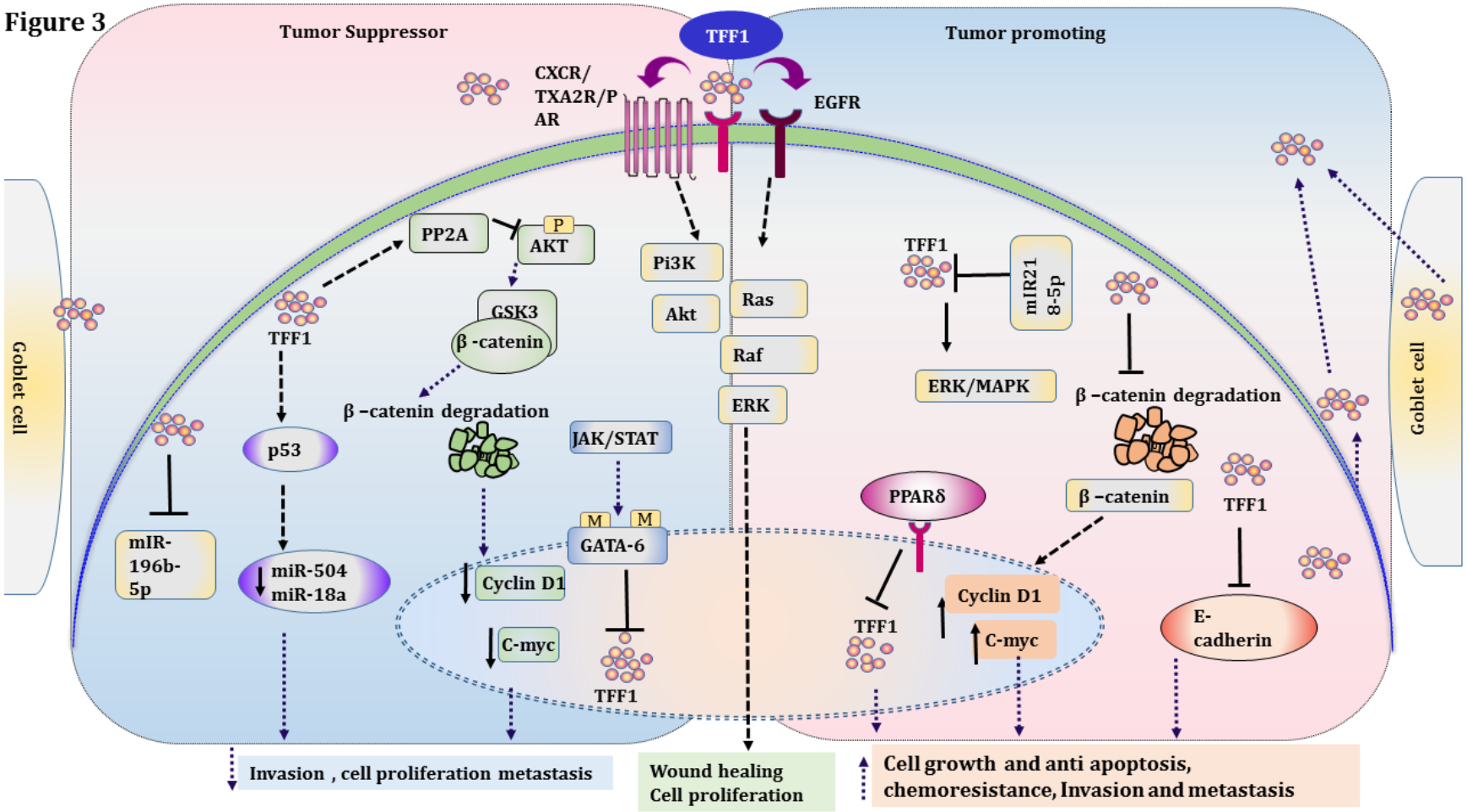

Tumor Suppressor
Tumor promoting
TFF1
CXCR/
TXA2R/P
AR
EGFR
PP2A
P
AKT
Pi3K
Ras
Raf
ERK
Akt
GSK3
β -catenin
β –catenin degradation
TFF1
p53
JAK/STAT
M
M
GATA-6
Cyclin D1
C-myc
TFF1
mIR-
196b-
5p
miR-504
miR-18a
TFF1
mIR21
8-5p
ERK/MAPK
β –catenin degradation
PPARδ
β –catenin
TFF1
Cyclin D1
C-myc
TFF1
E-
cadherin
Goblet cell
Goblet cell
Invasion , cell proliferation metastasis
Wound healing
Cell proliferation
Cell growth and anti apoptosis,
chemoresistance, Invasion and metastasis

Figure 3: Schematic representation how TFF1 mediated traveling of cell signal lead to tumor suppressor as well as tumor promoting effect. CXCR, PAR, EGFR are postulated signaling receptor for TFF1 which incites downstream signaling by activating AKT, PI3K, ERK. **Left box** represent role of TFF1 as tumor suppressor. Activation of p53 and subsequent downregulation of miR-504 and 18a mediated by TFF1 increased cell death. Decreased c-myc and cyclin D1 gene expression by fostering β-catenin degradation also lead to decreased cell proliferation, invasion and metastasis. Decreased expression of TFF1 because of aberrant methylation of GATA-6 lead to aggressive gastric cancer suggest its tumor suppressive role. **Right box** represent role of TFF1 as tumorigenic phenotype facilitator. TFF1 can mediate phosphorylation of ERK for increased cell proliferation which can be inhibited by miR218-5p. It is also involved in downregulating e-cadherin and miR-196-5bp which lead to increased cell proliferation, invasion and metastasis. In mucinous ovarian carcinoma it increased cyclin D1 and c-myc expression through activating the β-catenin signaling which in turn increases cell proliferation, invasion and chemo resistance.

Figure 4

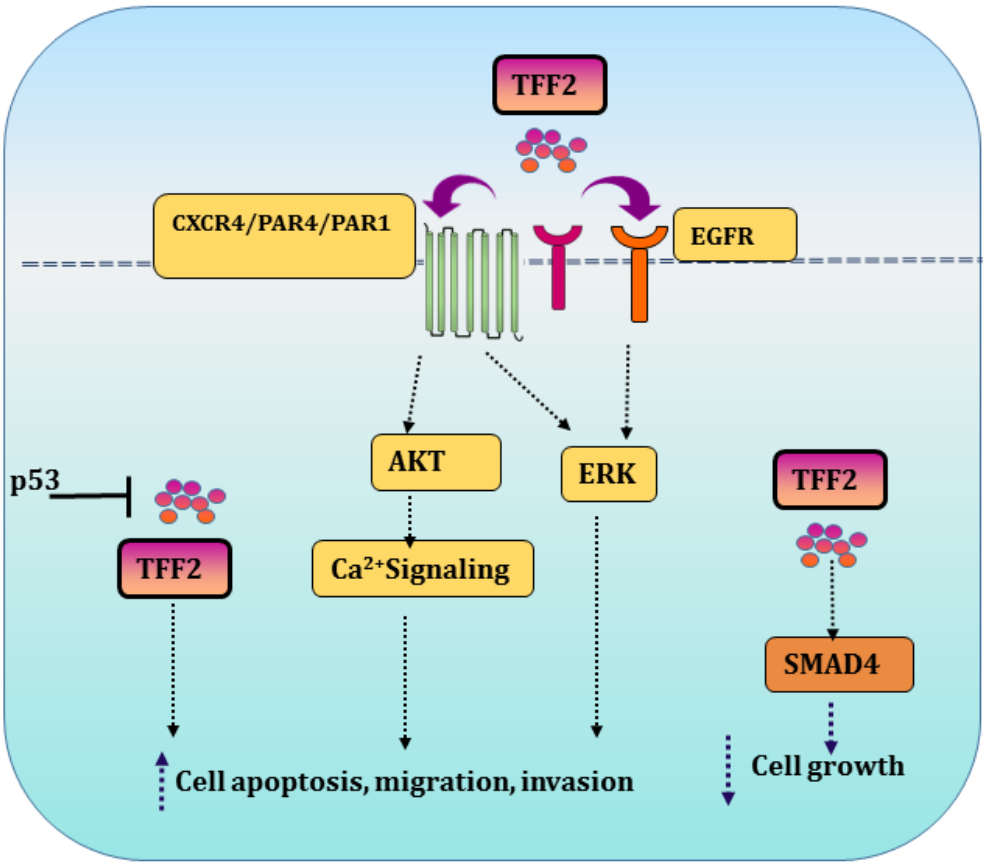

TFF2
CXCR4/PAR4/PAR1
EGFR
AKT
ERK
p53
TFF2
Ca2+Signaling
TFF2
SMAD4
Cell apoptosis, migration, invasion
Cell growth

Figure 4: Schematic representation how TFF2 mediated traveling of cell signal lead to tumor suppressor as well as tumor promoting role. Tumor suppressor p53 induces cell apoptosis and inhibits cell migration by downregulating TFF2. TFF2 has shown to be low affinity ligand for CXCR4 and thus initiates downstream signaling cascade for inciting Ca^{2+} signaling and cue for cell proliferation, migration which are hallmarks for malignancy phenotype. EGFR has been identified as a putative receptor for all TFF. On contrary, overexpression of TFF2 activates SMAD4 which leads to decrease cell proliferation suggesting tumor suppressor role of TFF2.

Figure 5

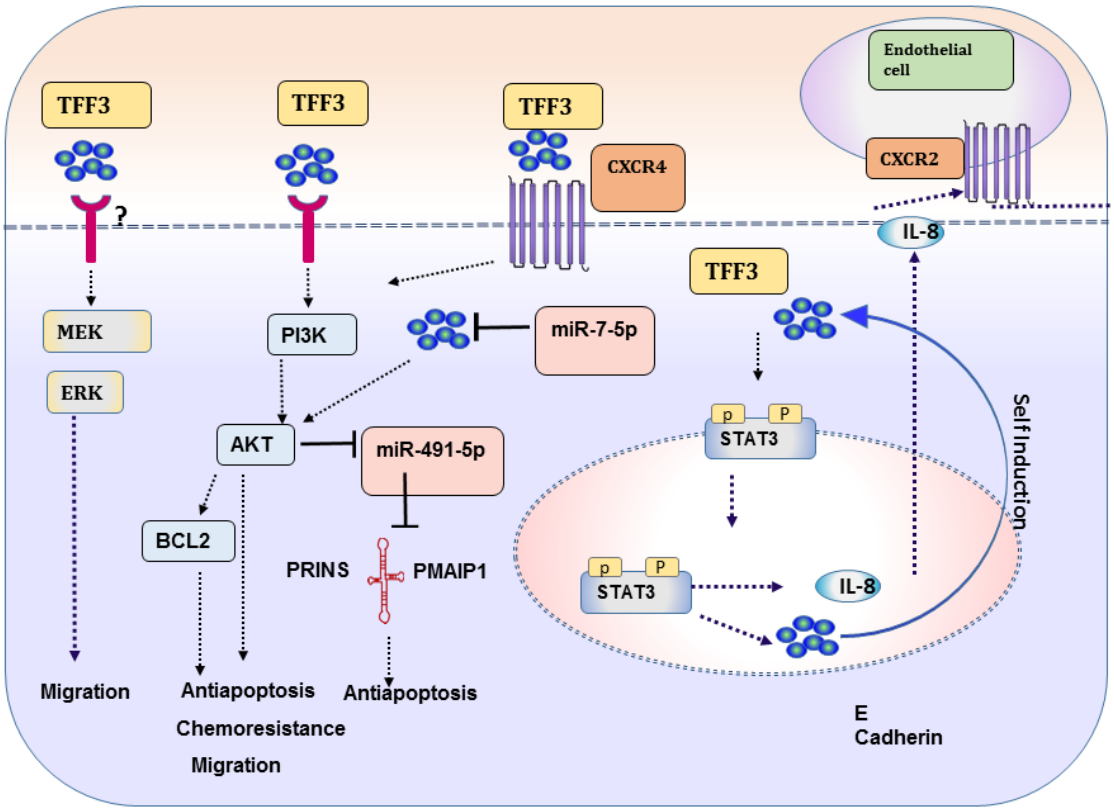

TFF3
TFF3
TFF3
CXCR4
Endothelial cell
CXCR2
IL-8
?
MEK
ERK
PI3K
miR-7-5p
TFF3
AKT
miR-491-5p
BCL2
PRINS
PMAIP1
STAT3
STAT3
IL-8
Self Induction
Migration
Antiapoptosis
Chemoresistance
Migration
Antiapoptosis
E
Cadherin

Figure 5: Schematic representation of traveling of TFF3 mediated cell signaling. TFF3 induces activation of ERK and AKT signaling for migration, anti-apoptosis. TFF3 also regulates miRNA and lncRNA for inducing anti-apoptosis. TFF3 mediate downstream signaling through CXCR4 and EGFR receptors for migration, cell proliferation and invasion. It can phosphorylate SMAD4 which led to induction of IL-8 which consequently result into angiogenesis, cell proliferation, invasion and migration. Moreover, TFF3 can autoinduce itself by activating SMAD4.

Reference

[1] H. Kindon, C. Pothoulakis, L. Thim, K. Lynch-Devaney, D.K. Podolsky, Trefoil peptide protection of intestinal epithelial barrier function: cooperative interaction with mucin glycoprotein, Gastroenterology.109 (1995) 516-523.

[2] W. Hoffmann, Trefoil factors TFF (trefoil factor family) peptide-triggered signals promoting mucosal restitution, Cell Mol Life Sci.62 (2005) 2932-2938.

[3] S. Beck, H. Schmitt, H. Shizuya, N. Blin, P. Gott, Cloning of contiguous genomic fragments from human chromosome 21 harbouring three trefoil peptide genes, Hum Genet.98 (1996) 233-235.

[4] O. Lefebvre, C. Wolf, M. Kedinger, M.P. Chenard, C. Tomasetto, P. Chambon, M.C. Rio, The mouse one P-domain (pS2) and two P-domain (mSP) genes exhibit distinct patterns of expression, J Cell Biol.122 (1993) 191-198.

[5] L. Thim, F.E. May, Structure of mammalian trefoil factors and functional insights, Cell Mol Life Sci.62 (2005) 2956-2973.

[6] P. Masiakowski, R. Breathnach, J. Bloch, F. Gannon, A. Krust, P. Chambon, Cloning of cDNA sequences of hormone-regulated genes from the MCF-7 human breast cancer cell line, Nucleic Acids Res.10 (1982) 7895-7903.

[7] K.H. Jorgensen, L. Thim, H.E. Jacobsen, Pancreatic spasmolytic polypeptide (PSP): I. Preparation and initial chemical characterization of a new polypeptide from porcine pancreas, Regul Pept.3 (1982) 207-219.

[8] K.D. Jorgensen, B. Diamant, K.H. Jorgensen, L. Thim, Pancreatic spasmolytic polypeptide (PSP): III. Pharmacology of a new porcine pancreatic polypeptide with spasmolytic and gastric acid secretion inhibitory effects, Regul Pept.3 (1982) 231-243.

[9] L. Thim, Trefoil peptides: from structure to function, Cell Mol Life Sci.53 (1997) 888-903.

[10] F.G. Hanisch, H. Ragge, T. Kalinski, F. Meyer, H. Kalbacher, W. Hoffmann, Human gastric TFF2 peptide contains an N-linked fucosylated N,N'-

diacetyllactosediamine (LacdiNAc) oligosaccharide, Glycobiology.23 (2013) 2-11.

[11] R.J. Playford, T. Marchbank, R. Chinery, R. Evison, M. Pignatelli, R.A. Boulton, L. Thim, A.M. Hanby, Human spasmolytic polypeptide is a cytoprotective agent that stimulates cell migration, Gastroenterology.108 (1995) 108-116.

[12] R. Chinery, J. Williamson, R. Poulsom, The gene encoding human intestinal trefoil factor (TFF3) is located on chromosome 21q22.3 clustered with other members of the trefoil peptide family, Genomics.32 (1996) 281-284.

[13] D. Taupin, D.K. Podolsky, Trefoil factors: initiators of mucosal healing, Nat Rev Mol Cell Biol.4 (2003) 721-732.

[14] B.R. Westley, S.M. Griffin, F.E. May, Interaction between TFF1, a gastric tumor suppressor trefoil protein, and TFIZ1, a brichos domain-containing protein with homology to SP-C, Biochemistry.44 (2005) 7967-7975.

[15] T.K. Albert, W. Laubinger, S. Muller, F.G. Hanisch, T. Kalinski, F. Meyer, W. Hoffmann, Human intestinal TFF3 forms disulfide-linked heteromers with the mucus-associated FCGBP protein and is released by hydrogen sulfide, J Proteome Res.9 (2010) 3108-3117.

[16] K.A. Oien, F. McGregor, S. Butler, R.K. Ferrier, I. Downie, S. Bryce, S. Burns, W.N. Keith, Gastrokine 1 is abundantly and specifically expressed in superficial gastric epithelium, down-regulated in gastric carcinoma, and shows high evolutionary conservation, J Pathol.203 (2004) 789-797.

[17] F.E. May, S.M. Griffin, B.R. Westley, The trefoil factor interacting protein TFIZ1 binds the trefoil protein TFF1 preferentially in normal gastric mucosal cells but the co-expression of these proteins is deregulated in gastric cancer, Int J Biochem Cell Biol.41 (2009) 632-640.

[18] R. Chinery, P.A. Bates, A. De, P.S. Freemont, Characterisation of the single copy trefoil peptides intestinal trefoil factor and pS2 and their ability to form covalent dimers, FEBS Lett.357 (1995) 50-54.

[19] S.S. Poulsen, H. Kissow, K. Hare, B. Hartmann, L. Thim, Luminal and parenteral TFF2 and TFF3 dimer and monomer in two models of experimental colitis in the rat, Regul Pept.126 (2005) 163-171.

[20] F.E. May, S.T. Church, S. Major, B.R. Westley, The closely related estrogen-regulated trefoil proteins TFF1 and TFF3 have markedly different hydrodynamic properties, overall charge, and distribution of surface charge, Biochemistry.42 (2003) 8250-8259.

[21] S. Suemori, K. Lynch-Devaney, D.K. Podolsky, Identification and characterization of rat intestinal trefoil factor: tissue- and cell-specific member of the trefoil protein family, Proc Natl Acad Sci U S A.88 (1991) 11017-11021.

[22] G.A. Cook, M. Familari, L. Thim, A.S. Giraud, The trefoil peptides TFF2 and TFF3 are expressed in rat lymphoid tissues and participate in the immune response, FEBS Lett.456 (1999) 155-159.

[23] M.H. Samson, S.S. Poulsen, R. Obeid, W. Herrmann, E. Nexo, Trefoil factor family peptides in the human foetus and at birth, Eur J Clin Invest.41 (2011) 785-792.

[24] J. Madsen, O. Nielsen, I. Tornoe, L. Thim, U. Holmskov, Tissue localization of human trefoil factors 1, 2, and 3, J Histochem Cytochem.55 (2007) 505-513.

[25] R. Poulsom, A.M. Hanby, E.N. Lalani, F. Hauser, W. Hoffmann, G.W. Stamp, Intestinal trefoil factor (TFF 3) and pS2 (TFF 1), but not spasmolytic polypeptide (TFF 2) mRNAs are co-expressed in normal, hyperplastic, and neoplastic human breast epithelium, J Pathol.183 (1997) 30-38.

[26] D.A. Devine, A.S. High, P.J. Owen, R. Poulsom, W.A. Bonass, Trefoil factor expression in normal and diseased human salivary glands, Hum Pathol.31 (2000) 509-515.

[27] F.P. Paulsen, M. Hinz, U. Schaudig, A.B. Thale, W. Hoffmann, TFF peptides in the human efferent tear ducts, Invest Ophthalmol Vis Sci.43 (2002) 3359-3364.

[28] E. dos Santos Silva, M. Ulrich, G. Doring, K. Botzenhart, P. Gott, Trefoil factor family domain peptides in the human respiratory tract, J Pathol.190 (2000) 133-142.

[29] W.M. Wong, R. Poulsom, N.A. Wright, Trefoil peptides, Gut.44 (1999) 890-895.

[30] A.M. Hanby, R. Poulsom, S. Singh, G. Elia, R.E. Jeffery, N.A. Wright, Spasmolytic polypeptide is a major antral peptide: distribution of the trefoil peptides human spasmolytic polypeptide and pS2 in the stomach, Gastroenterology.105 (1993) 1110-1116.

[31] M.C. Rio, M.P. Chenard, C. Wolf, L. Marcellin, C. Tomasetto, R. Lathe, J.P. Bellocq, P. Chambon, Induction of pS2 and hSP genes as markers of mucosal ulceration of the digestive tract, Gastroenterology.100 (1991) 375-379.

[32] A.M. Hanby, R. Poulsom, G. Elia, S. Singh, J.M. Longcroft, N.A. Wright, The expression of the trefoil peptides pS2 and human spasmolytic polypeptide (hSP) in 'gastric metaplasia' of the proximal duodenum: implications for the nature of 'gastric metaplasia', J Pathol.169 (1993) 355-360.

[33] W. Hoffmann, W. Jagla, Cell type specific expression of secretory TFF peptides: colocalization with mucins and synthesis in the brain, Int Rev Cytol.213 (2002) 147-181.

[34] J.C. Probst, T. Zetzsche, M. Weber, P. Theilemann, T. Skutella, R. Landgraf, G.F. Jirikowski, Human intestinal trefoil factor is expressed in human hypothalamus and pituitary: evidence for a novel neuropeptide, FASEB J.10 (1996) 1518-1523.

[35] S. Pelden, T. Insawang, C. Thuwajit, P. Thuwajit, The trefoil factor 1 (TFF1) protein involved in doxorubicininduced apoptosis resistance is upregulated by estrogen in breast cancer cells, Oncol Rep.30 (2013) 1518-1526.

[36] T. Marchbank, R.J. Playford, Trefoil factor family peptides enhance cell migration by increasing cellular osmotic permeability and aquaporin 3 levels, FASEB J.32 (2018) 1017-1024.

[37] H. Mashimo, D.C. Wu, D.K. Podolsky, M.C. Fishman, Impaired defense of intestinal mucosa in mice lacking intestinal trefoil factor, Science.274 (1996) 262-265.

[38] S. Thiem, M.F. Eissmann, E. Stuart, J. Elzer, A. Jonas, M. Buchert, M. Ernst, Inducible gene modification in the gastric epithelium of Tff1-CreERT2, Tff2-rtTA, Tff3-luc mice, Genesis.54 (2016) 626-635.

[39] S. Montefusco, R. Esposito, L. D'Andrea, M.C. Monti, C. Dunne, B. Dolan, A. Tosco, L. Marzullo, M. Clyne, Copper promotes TFF1-mediated Helicobacter pylori colonization, PLoS One.8 (2013) e79455.

[40] D. Meyer zum Buschenfelde, R. Tauber, O. Huber, TFF3-peptide increases transepithelial resistance in epithelial cells by modulating claudin-1 and -2 expression, Peptides.27 (2006) 3383-3390.

[41] D. Meyer zum Buschenfelde, H. Hoschutzky, R. Tauber, O. Huber, Molecular mechanisms involved in TFF3 peptide-mediated modulation of the E-cadherin/catenin cell adhesion complex, Peptides.25 (2004) 873-883.

[42] J.A. Efstathiou, M. Noda, A. Rowan, C. Dixon, R. Chinery, A. Jawhari, T. Hattori, N.A. Wright, W.F. Bodmer, M. Pignatelli, Intestinal trefoil factor controls the expression of the adenomatous polyposis coli-catenin and the E-cadherin-catenin complexes in human colon carcinoma cells, Proc Natl Acad Sci U S A.95 (1998) 3122-3127.

[43] A.M. Nunez, M. Berry, J.L. Imler, P. Chambon, The 5' flanking region of the pS2 gene contains a complex enhancer region responsive to oestrogens, epidermal growth factor, a tumour promoter (TPA), the c-Ha-ras oncoprotein and the c-jun protein, EMBO J.8 (1989) 823-829.

[44] B.E. Gillesby, M. Stanostefano, W. Porter, S. Safe, Z.F. Wu, T.R. Zacharewski, Identification of a motif within the 5' regulatory region of pS2 which is responsible for AP-1 binding and TCDD-mediated suppression, Biochemistry.36 (1997) 6080-6089.

[45] R. Eferl, E.F. Wagner, AP-1: a double-edged sword in tumorigenesis, Nat Rev Cancer.3 (2003) 859-868.

[46] M.B. Loncar, E.D. Al-azzeh, P.S. Sommer, M. Marinovic, K. Schmehl, M. Kruschewski, N. Blin, R. Stohwasser, P. Gott, T. Kayademir, Tumour necrosis factor alpha and nuclear factor kappaB inhibit transcription of human TFF3 encoding a gastrointestinal healing peptide, Gut.52 (2003) 1297-1303.

[47] J.M. Sun, V.A. Spencer, L. Li, H. Yu Chen, J. Yu, J.R. Davie, Estrogen regulation of trefoil factor 1 expression by estrogen receptor alpha and Sp proteins, Exp Cell Res.302 (2005) 96-107.

[48] G. Regalo, S. Forster, C. Resende, B. Bauer, B. Fleige, W. Kemmner, P.M. Schlag, T.F. Meyer, J.C. Machado, A. Leutz, C/EBPbeta regulates homeostatic and oncogenic gastric cell proliferation, J Mol Med (Berl).94 (2016) 1385-1395.

[49] N.V. Sankpal, M.W. Mayo, S.M. Powell, Transcriptional repression of TFF1 in gastric epithelial cells by CCAAT/enhancer binding protein-beta, Biochim Biophys Acta.1728 (2005) 1-10.

[50] A.L. Chi, S. Lim, T.C. Wang, Characterization of a CCAAT-enhancer element of trefoil factor family 2 (TFF2) promoter in MCF-7 cells, Peptides.25 (2004) 839-847.

[51] H. Itoh, N. Inoue, D.K. Podolsky, Goblet-cell-specific transcription of mouse intestinal trefoil factor gene results from collaboration of complex series of positive and negative regulatory elements, Biochem J.341 (Pt 2) (1999) 461-472.

[52] S. Beck, P. Sommer, E. dos Santos Silva, N. Blin, P. Gott, Hepatocyte nuclear factor 3 (winged helix domain) activates trefoil factor gene TFF1 through a binding motif adjacent to the TATAA box, DNA Cell Biol.18 (1999) 157-164.

[53] C.S. Wu, K.L. Wei, J.L. Chou, C.K. Lu, C.C. Hsieh, J.M. Lin, Y.F. Deng, W.T. Hsu, H.M. Wang, C.H. Leung, D.L. Ma, C. Li, M.W. Chan, Aberrant JAK/STAT Signaling Suppresses TFF1 and TFF2 through Epigenetic Silencing of GATA6 in Gastric Cancer, Int J Mol Sci.17 (2016).

[54] Z.E. Khan, T.C. Wang, G. Cui, A.L. Chi, R. Dimaline, Transcriptional regulation of the human trefoil factor, TFF1, by gastrin, Gastroenterology.125 (2003) 510-521.

[55] E. Romano, M. Vllahu, V. Bizzarro, R. Belvedere, R. Esposito, A. Petrella, A. Tosco, TFF1 Promotes EMT-Like Changes through an Auto-Induction Mechanism, Int J Mol Sci.19 (2018).

[56] C. Blanchard, S. Durual, M. Estienne, K. Bouzakri, M.H. Heim, N. Blin, J.C. Cuber, IL-4 and IL-13 up-regulate intestinal trefoil factor expression: requirement for STAT6 and de novo protein synthesis, J Immunol.172 (2004) 3775-3783.

[57] S. Sharma, T.K. Kelly, P.A. Jones, Epigenetics in cancer, Carcinogenesis.31 (2010) 27-36.

[58] H. Tomita, S. Takaishi, T.R. Menheniott, X. Yang, W. Shibata, G. Jin, K.S. Betz, K. Kawakami, T. Minamoto, C. Tomasetto, M.C. Rio, N. Lerkowit, A. Varro, A.S. Giraud, T.C. Wang, Inhibition of gastric carcinogenesis by the hormone gastrin is mediated by suppression of TFF1 epigenetic silencing, Gastroenterology.140 (2011) 879-891.

[59] P. Jiang, G. Yu, Y. Zhang, Y. Xiang, Z. Zhu, W. Feng, W. Lee, Y. Zhang, Promoter hypermethylation and downregulation of trefoil factor 2 in human gastric cancer, Oncol Lett.7 (2014) 1525-1531.

[60] E.M. Vestergaard, E. Nexo, N. Torring, M. Borre, T.F. Orntoft, K.D. Sorensen, Promoter hypomethylation and upregulation of trefoil factors in prostate cancer, Int J Cancer.127 (2010) 1857-1865.

[61] M. Norgaard, C. Haldrup, T.M. Storebjerg, E.M. Vestergaard, P.J. Wild, S. Hoyer, M. Borre, T.F. Orntoft, K.D. Sorensen, Comprehensive Evaluation of TFF3 Promoter Hypomethylation and Molecular Biomarker Potential for Prostate Cancer Diagnosis and Prognosis, Int J Mol Sci.18 (2017).

[62] O. Perera, A. Evans, M. Pertziger, C. MacDonald, H. Chen, D.X. Liu, P.E. Lobie, J.K. Perry, Trefoil factor 3 (TFF3) enhances the oncogenic characteristics of prostate carcinoma cells and reduces sensitivity to ionising radiation, Cancer Lett.361 (2015) 104-111.

[63] N. Sato, A. Maitra, N. Fukushima, N.T. van Heek, H. Matsubayashi, C.A. Iacobuzio-Donahue, C. Rosty, M. Goggins, Frequent hypomethylation of multiple genes overexpressed in pancreatic ductal adenocarcinoma, Cancer Res.63 (2003) 4158-4166.

[64] V. Martin, S. Ribieras, X.G. Song-Wang, Y. Lasne, L. Frappart, M.C. Rio, R. Dante, Involvement of DNA methylation in the control of the expression of an estrogen-induced breast-cancer-associated protein (pS2) in human breast cancers, J Cell Biochem.65 (1997) 95-106.

[65] L.A. Loss, A. Sadanandam, S. Durinck, S. Nautiyal, D. Flaucher, V.E. Carlton, M. Moorhead, Y. Lu, J.W. Gray, M. Faham, P. Spellman, B. Parvin, Prediction of epigenetically regulated genes in breast cancer cell lines, BMC Bioinformatics.11 (2010) 305.

[66] B.C. Morson, L.H. Sobin, E. Grundmann, A. Johansen, T. Nagayo, A. Serck-Hanssen, Precancerous conditions and epithelial dysplasia in the stomach, J Clin Pathol.33 (1980) 711-721.

[67] M. Herfs, P. Hubert, P. Delvenne, Epithelial metaplasia: adult stem cell reprogramming and (pre)neoplastic transformation mediated by inflammation?, Trends Mol Med.15 (2009) 245-253.

[68] S.M. Karam, Mouse models demonstrating the role of stem/progenitor cells in gastric carcinogenesis, Front Biosci (Landmark Ed).15 (2010) 595-603.

[69] S.M. Karam, C. Tomasetto, M.C. Rio, Trefoil factor 1 is required for the commitment programme of mouse oxyntic epithelial progenitors, Gut.53 (2004) 1408-1415.

[70] A.C. Engevik, R. Feng, E. Choi, S. White, N. Bertaux-Skeirik, J. Li, M.M. Mahe, E. Aihara, L. Yang, B. DiPasquale, S. Oh, K.A. Engevik, A.S. Giraud, M.H. Montrose, M. Medvedovic, M.A. Helmrath, J.R. Goldenring, Y. Zavros, The Development of Spasmolytic Polypeptide/TFF2-Expressing Metaplasia (SPEM) During Gastric Repair Is Absent in the Aged Stomach, Cell Mol Gastroenterol Hepatol.2 (2016) 605-624.

[71] A.M. Halldorsdottir, M. Sigurdardottrir, J.G. Jonasson, M. Oddsdottir, J. Magnusson, J.R. Lee, J.R. Goldenring, Spasmolytic polypeptide-

expressing metaplasia (SPEM) associated with gastric cancer in Iceland, Dig Dis Sci.48 (2003) 431-441.

[72] M. Quante, F. Marrache, J.R. Goldenring, T.C. Wang, TFF2 mRNA transcript expression marks a gland progenitor cell of the gastric oxyntic mucosa, Gastroenterology.139 (2010) 2018-2027 e2012.

[73] O. Strobel, D.E. Rosow, E.Y. Rakhlin, G.Y. Lauwers, A.G. Trainor, J. Alsina, C. Fernandez-Del Castillo, A.L. Warshaw, S.P. Thayer, Pancreatic duct glands are distinct ductal compartments that react to chronic injury and mediate Shh-induced metaplasia, Gastroenterology.138 (2010) 1166-1177.

[74] J. Yamaguchi, A.S. Liss, A. Sontheimer, M. Mino-Kenudson, C.F. Castillo, A.L. Warshaw, S.P. Thayer, Pancreatic duct glands (PDGs) are a progenitor compartment responsible for pancreatic ductal epithelial repair, Stem Cell Res.15 (2015) 190-202.

[75] J. Yamaguchi, M. Mino-Kenudson, A.S. Liss, S. Chowdhury, T.C. Wang, C. Fernandez-Del Castillo, K.D. Lillemoe, A.L. Warshaw, S.P. Thayer, Loss of Trefoil Factor 2 From Pancreatic Duct Glands Promotes Formation of Intraductal Papillary Mucinous Neoplasms in Mice, Gastroenterology.151 (2016) 1232-1244 e1210.

[76] B.J. Reid, X. Li, P.C. Galipeau, T.L. Vaughan, Barrett's oesophagus and oesophageal adenocarcinoma: time for a new synthesis, Nat Rev Cancer.10 (2010) 87-101.

[77] Y. Yamamoto, X. Wang, D. Bertrand, F. Kern, T. Zhang, M. Duleba, S. Srivastava, C.C. Khor, Y. Hu, L.H. Wilson, H. Blaszyk, D. Rolshud, M. Teh, J. Liu, B.E. Howitt, M. Vincent, C.P. Crum, N. Nagarajan, K.Y. Ho, F. McKeon, W. Xian, Mutational spectrum of Barrett's stem cells suggests paths to initiation of a precancerous lesion, Nat Commun.7 (2016) 10380.

[78] E. Fessler, J.P. Medema, Colorectal Cancer Subtypes: Developmental Origin and Microenvironmental Regulation, Trends Cancer.2 (2016) 505-518.

[79] R. Singh, L. Zorron Cheng Tao Pu, D. Koay, A. Burt, Sessile serrated adenoma/polyps: Where are we at in 2016?, World J Gastroenterol.22 (2016) 7754-7759.

[80] D.A. Delker, B.M. McGettigan, P. Kanth, S. Pop, D.W. Neklason, M.P. Bronner, R.W. Burt, C.H. Hagedorn, RNA sequencing of sessile serrated colon polyps identifies differentially expressed genes and immunohistochemical markers, PLoS One.9 (2014) e88367.

[81] M. Caruso, J. Moore, G.J. Goodall, M. Thomas, S. Phillis, A. Tyskin, G. Cheetham, N. Lerda, H. Takahashi, A. Ruszkiewicz, Over-expression of cathepsin E and trefoil factor 1 in sessile serrated adenomas of the colorectum identified by gene expression analysis, Virchows Arch.454 (2009) 291-302.

[82] S. Tajima, N. Kurabe, K. Okudela, K. Yajima, T. Takahashi, H. Neyatani, H. Sugimura, K. Koda, Extensive goblet cell metaplasia of the peripheral lung may harbor precancerous molecular changes: comparison of two cases, Pathol Int.64 (2014) 533-538.

[83] W. Hoffmann, TFF (trefoil factor family) peptides and their potential roles for differentiation processes during airway remodeling, Curr Med Chem.14 (2007) 2716-2719.

[84] S. Aikou, Y. Ohmoto, T. Gunji, N. Matsuhashi, H. Ohtsu, H. Miura, K. Kubota, Y. Yamagata, Y. Seto, A. Nakajima, J.R. Goldenring, M. Kaminishi, S. Nomura, Tests for serum levels of trefoil factor family proteins can improve gastric cancer screening, Gastroenterology.141 (2011) 837-845 e831-837.

[85] Y. Shi, M. Hon, R.M. Evans, The peroxisome proliferator-activated receptor delta, an integrator of transcriptional repression and nuclear receptor signaling, Proc Natl Acad Sci U S A.99 (2002) 2613-2618.

[86] N. Martin-Martin, A. Zabala-Letona, S. Fernandez-Ruiz, L. Arreal, L. Camacho, M. Castillo-Martin, A.R. Cortazar, V. Torrano, I. Astobiza, P. Zuniga-Garcia, A. Ugalde-Olano, A. Loizaga-Iriarte, M. Unda, L. Valcarcel-Jimenez, A. Arruabarrena-Aristorena, M. Piva, P. Sanchez-Mosquera, A.M.

Aransay, A. Gomez-Munoz, R. Barrio, J.D. Sutherland, A. Carracedo, PPARdelta Elicits Ligand-Independent Repression of Trefoil Factor Family to Limit Prostate Cancer Growth, Cancer Res.78 (2018) 399-409.

[87] D.R. Radiloff, T.P. Wakeman, J. Feng, S. Schilling, E. Seto, X.F. Wang, Trefoil factor 1 acts to suppress senescence induced by oncogene activation during the cellular transformation process, Proc Natl Acad Sci U S A.108 (2011) 6591-6596.

[88] N.M. Bougen, N. Amiry, Y. Yuan, X.J. Kong, V. Pandey, L.J. Vidal, J.K. Perry, T. Zhu, P.E. Lobie, Trefoil factor 1 suppression of E-CADHERIN enhances prostate carcinoma cell invasiveness and metastasis, Cancer Lett.332 (2013) 19-29.

[89] S. Zhao, Y. Ma, X. Huang, Trefoil factor 1 elevates the malignant phenotype of mucinous ovarian cancer cell through Wnt/beta-catenin signaling, Int J Clin Exp Pathol.8 (2015) 10412-10419.

[90] S. Rodrigues, E. Van Aken, S. Van Bocxlaer, S. Attoub, Q.D. Nguyen, E. Bruyneel, B.R. Westley, F.E. May, L. Thim, M. Mareel, C. Gespach, S. Emami, Trefoil peptides as proangiogenic factors in vivo and in vitro: implication of cyclooxygenase-2 and EGF receptor signaling, FASEB J.17 (2003) 7-16.

[91] S. Rodrigues, S. Attoub, Q.D. Nguyen, E. Bruyneel, C.M. Rodrigue, B.R. Westley, F.E. May, L. Thim, M. Mareel, S. Emami, C. Gespach, Selective abrogation of the proinvasive activity of the trefoil peptides pS2 and spasmolytic polypeptide by disruption of the EGF receptor signaling pathways in kidney and colonic cancer cells, Oncogene.22 (2003) 4488-4497.

[92] M. Soutto, A. Belkhiri, M.B. Piazuelo, B.G. Schneider, D. Peng, A. Jiang, M.K. Washington, Y. Kokoye, S.E. Crowe, A. Zaika, P. Correa, R.M. Peek, Jr., W. El-Rifai, Loss of TFF1 is associated with activation of NF-kappaB-mediated inflammation and gastric neoplasia in mice and humans, J Clin Invest.121 (2011) 1753-1767.

[93] M. Soutto, D. Peng, A. Katsha, Z. Chen, M.B. Piazuelo, M.K. Washington, A. Belkhiri, P. Correa, W. El-Rifai, Activation of beta-catenin signalling by TFF1 loss promotes cell proliferation and gastric tumorigenesis, Gut.64 (2015) 1028-1039.

[94] L. Cobler, R. Mejias-Luque, M. Garrido, M. Pera, E. Badia-Garrido, C. de Bolos, Activation of the NF-kB pathway downregulates TFF-1 in gastric carcinogenesis, Virchows Arch.463 (2013) 497-507.

[95] M. Soutto, M. Saleh, M.S. Arredouani, B. Piazuelo, A. Belkhiri, W. El-Rifai, Loss of Tff1 Promotes Pro-Inflammatory Phenotype with Increase in the Levels of RORgammat+ T Lymphocytes and Il-17 in Mouse Gastric Neoplasia, J Cancer.8 (2017) 2424-2435.

[96] E. Buache, N. Etique, F. Alpy, I. Stoll, M. Muckensturm, B. Reina-San-Martin, M.P. Chenard, C. Tomasetto, M.C. Rio, Deficiency in trefoil factor 1 (TFF1) increases tumorigenicity of human breast cancer cells and mammary tumor development in TFF1-knockout mice, Oncogene.30 (2011) 3261-3273.

[97] N. Amiry, X. Kong, N. Muniraj, N. Kannan, P.M. Grandison, J. Lin, Y. Yang, C.M. Vouyovitch, S. Borges, J.K. Perry, H.C. Mertani, T. Zhu, D. Liu, P.E. Lobie, Trefoil factor-1 (TFF1) enhances oncogenicity of mammary carcinoma cells, Endocrinology.150 (2009) 4473-4483.

[98] W. Brugger, J. Frisch, G. Schulz, K. Pressler, R. Mertelsmann, L. Kanz, Sequential administration of interleukin-3 and granulocyte-macrophage colony-stimulating factor following standard-dose combination chemotherapy with etoposide, ifosfamide, and cisplatin, J Clin Oncol.10 (1992) 1452-1459.

[99] L.S. Siu, H. Romanska, P.D. Abel, M. Baus-Loncar, T. Kayademir, G.W. Stamp, N. Lalani el, TFF2 (trefoil family factor2) inhibits apoptosis in breast and colorectal cancer cell lines, Peptides.25 (2004) 855-863.

[100] K. Kosriwong, T.R. Menheniott, A.S. Giraud, P. Jearanaikoon, B. Sripa, T. Limpaiboon, Trefoil factors: tumor progression markers and mitogens via

EGFR/MAPK activation in cholangiocarcinoma, World J Gastroenterol.17 (2011) 1631-1641.

[101] S. Kamlua, S. Patrakitkomjorn, P. Jearanaikoon, T.R. Menheniott, A.S. Giraud, T. Limpaiboon, A novel TFF2 splice variant (EX2TFF2) correlates with longer overall survival time in cholangiocarcinoma, Oncol Rep.27 (2012) 1207-1212.

[102] K. Fujimoto, G. Fujii, K. Taguchi, K. Yasuda, Y. Matsuo, A. Hashiyama, M. Mutoh, H. Tanaka, M. Wada, Involvement of trefoil factor family 2 in the enlargement of intestinal tumors in Apc(Min/+) mice, Biochem Biophys Res Commun.463 (2015) 859-863.

[103] D.K. Dhar, T.C. Wang, R. Maruyama, J. Udagawa, H. Kubota, T. Fuji, M. Tachibana, T. Ono, H. Otani, N. Nagasue, Expression of cytoplasmic TFF2 is a marker of tumor metastasis and negative prognostic factor in gastric cancer, Lab Invest.83 (2003) 1343-1352.

[104] Y. Cai, M. Yi, D. Chen, J. Liu, B. Guleng, J. Ren, H. Shi, Trefoil factor family 2 expression inhibits gastric cancer cell growth and invasion in vitro via interactions with the transcription factor Sp3, Int J Mol Med.38 (2016) 1474-1480.

[105] S.P. Tu, A.L. Chi, W. Ai, S. Takaishi, Z. Dubeykovskaya, M. Quante, J.G. Fox, T.C. Wang, p53 inhibition of AP1-dependent TFF2 expression induces apoptosis and inhibits cell migration in gastric cancer cells, Am J Physiol Gastrointest Liver Physiol.297 (2009) G385-396.

[106] Z. Dubeykovskaya, A. Dubeykovskiy, J. Solal-Cohen, T.C. Wang, Secreted trefoil factor 2 activates the CXCR4 receptor in epithelial and lymphocytic cancer cell lines, J Biol Chem.284 (2009) 3650-3662.

[107] D. Barras, BRAF Mutation in Colorectal Cancer: An Update, Biomark Cancer.7 (2015) 9-12.

[108] M.K. Gala, T. Austin, S. Ogino, A.T. Chan, TFF2-CXCR4 Axis Is Associated with BRAF V600E Colon Cancer, Cancer Prev Res (Phila).8 (2015) 614-619.

[109] K. Orime, J. Shirakawa, Y. Togashi, K. Tajima, H. Inoue, Y. Ito, K. Sato, A. Nakamura, K. Aoki, Y. Goshima, Y. Terauchi, Trefoil factor 2 promotes cell proliferation in pancreatic beta-cells through CXCR-4-mediated ERK1/2 phosphorylation, Endocrinology.154 (2013) 54-64.

[110] W.F. Xu, H. Andersen, T.E. Whitmore, S.R. Presnell, D.P. Yee, A. Ching, T. Gilbert, E.W. Davie, D.C. Foster, Cloning and characterization of human protease-activated receptor 4, Proc Natl Acad Sci U S A.95 (1998) 6642-6646.

[111] Y. Zhang, G. Yu, Y. Wang, Y. Xiang, Q. Gao, P. Jiang, J. Zhang, W. Lee, Y. Zhang, Activation of protease-activated receptor (PAR) 1 by frog trefoil factor (TFF) 2 and PAR4 by human TFF2, Cell Mol Life Sci.68 (2011) 3771-3780.

[112] G. Yu, P. Jiang, Y. Xiang, Y. Zhang, Z. Zhu, C. Zhang, S. Lee, W. Lee, Y. Zhang, Increased expression of protease-activated receptor 4 and Trefoil factor 2 in human colorectal cancer, PLoS One.10 (2015) e0122678.

[113] V. Pandey, Z.S. Wu, M. Zhang, R. Li, J. Zhang, T. Zhu, P.E. Lobie, Trefoil factor 3 promotes metastatic seeding and predicts poor survival outcome of patients with mammary carcinoma, Breast Cancer Res.16 (2014) 429.

[114] S. Diao, Q. Zheng, J. Gao, Y. Yao, S. Ren, Y. Liu, Y. Xu, Trefoil factor 3 contributes to the malignancy of glioma via regulating HIF-1alpha, Oncotarget.8 (2017) 76770-76782.

[115] J. Liu, S.Y. Kim, S. Shin, S.H. Jung, S.H. Yim, J.Y. Lee, S.H. Lee, Y.J. Chung, Overexpression of TFF3 is involved in prostate carcinogenesis via blocking mitochondria-mediated apoptosis, Exp Mol Med.50 (2018) 110.

[116] Z. Yuan, D. Chen, X. Chen, H. Yang, Y. Wei, Overexpression of trefoil factor 3 (TFF3) contributes to the malignant progression in cervical cancer cells, Cancer Cell Int.17 (2017) 7.

[117] M.L. You, Y.J. Chen, Q.Y. Chong, M.M. Wu, V. Pandey, R.M. Chen, L. Liu, L. Ma, Z.S. Wu, T. Zhu, P.E. Lobie, Trefoil factor 3 mediation of oncogenicity and chemoresistance in hepatocellular carcinoma is AKT-BCL-2 dependent, Oncotarget.8 (2017) 39323-39344.

[118] F. Gao, S. Pan, B. Liu, H. Zhang, TFF3 knockout in human pituitary adenoma cell HP75 facilitates cell apoptosis via mitochondrial pathway, Int J Clin Exp Pathol.8 (2015) 14568-14573.

[119] J. Wu, H. Zhang, W. Zhang, J. Zhang, X. Lin, G. Xue, F. Gao, Lentivirus-mediated shRNA interference of trefoil factor 3 blocks cell viability, migration and invasion in the papillary thyroid carcinoma cells, Neoplasma.65 (2018) 169-177.

[120] H. Uchino, H. Kataoka, H. Itoh, R. Hamasuna, M. Koono, Overexpression of intestinal trefoil factor in human colon carcinoma cells reduces cellular growth in vitro and in vivo, Gastroenterology.118 (2000) 60-69.

[121] J. Grosse-Kreul, M. Busch, C. Winter, S. Pikos, H. Stephan, N. Dunker, Forced Trefoil Factor Family Peptide 3 (TFF3) Expression Reduces Growth, Viability, and Tumorigenicity of Human Retinoblastoma Cell Lines, PLoS One.11 (2016) e0163025.

[122] J. Dieckow, W. Brandt, K. Hattermann, S. Schob, U. Schulze, R. Mentlein, P. Ackermann, S. Sel, F.P. Paulsen, CXCR4 and CXCR7 Mediate TFF3-Induced Cell Migration Independently From the ERK1/2 Signaling Pathway, Invest Ophthalmol Vis Sci.57 (2016) 56-65.

[123] H.C. Harsha, K. Kandasamy, P. Ranganathan, S. Rani, S. Ramabadran, S. Gollapudi, L. Balakrishnan, S.B. Dwivedi, D. Telikicherla, L.D. Selvan, R. Goel, S. Mathivanan, A. Marimuthu, M. Kashyap, R.F. Vizza, R.J. Mayer, J.A. Decaprio, S. Srivastava, S.M. Hanash, R.H. Hruban, A. Pandey, A compendium of potential biomarkers of pancreatic cancer, PLoS Med.6 (2009) e1000046.

[124] R. Fatima, V.S. Akhade, D. Pal, S.M. Rao, Long noncoding RNAs in development and cancer: potential biomarkers and therapeutic targets, Mol Cell Ther.3 (2015) 5.

[125] L. Shao, Z. Chen, D. Peng, M. Soutto, S. Zhu, A. Bates, S. Zhang, W. El-Rifai, Methylation of the HOXA10 Promoter Directs miR-196b-5p-Dependent Cell Proliferation and Invasion of Gastric Cancer Cells, Mol Cancer Res.16 (2018) 696-706.

[126] J.F. Sousa, K.T. Nam, C.P. Petersen, H.J. Lee, H.K. Yang, W.H. Kim, J.R. Goldenring, miR-30-HNF4gamma and miR-194-NR2F2 regulatory networks contribute to the upregulation of metaplasia markers in the stomach, Gut.65 (2016) 914-924.

[127] M. Soutto, Z. Chen, M.A. Saleh, A. Katsha, S. Zhu, A. Zaika, A. Belkhiri, W. El-Rifai, TFF1 activates p53 through down-regulation of miR-504 in gastric cancer, Oncotarget.5 (2014) 5663-5673.

[128] W. Hu, C.S. Chan, R. Wu, C. Zhang, Y. Sun, J.S. Song, L.H. Tang, A.J. Levine, Z. Feng, Negative regulation of tumor suppressor p53 by microRNA miR-504, Mol Cell.38 (2010) 689-699.

[129] Y. Yang, Q. Mei, miRNA signature identification of retinoblastoma and the correlations between differentially expressed miRNAs during retinoblastoma progression, Mol Vis.21 (2015) 1307-1317.

[130] M. Busch, J. Grosse-Kreul, J.J. Wirtz, M. Beier, H. Stephan, B. Royer-Pokora, K. Metz, N. Dunker, Reduction of the tumorigenic potential of human retinoblastoma cell lines by TFF1 overexpression involves p53/caspase signaling and miR-18a regulation, Int J Cancer.141 (2017) 549-560.

[131] J. Liu, X. Wang, X. Yang, Y. Liu, Y. Shi, J. Ren, B. Guleng, miRNA423-5p regulates cell proliferation and invasion by targeting trefoil factor 1 in gastric cancer cells, Cancer Lett.347 (2014) 98-104.

[132] J. Guo, L. Xu, X. Teng, M. Sun, MicroRNA-7-5p regulates the proliferation and migration of intestinal epithelial cells by targeting trefoil factor 3 via inhibiting the phosphoinositide 3-kinase/Akt signalling pathway, Int J Mol Med.40 (2017) 1435-1443.

[133] A. Shukla, P. Gupta, R. Singh, D.P. Mishra, Glycolytic inhibitor 2-Deoxy-d-Glucose activates migration and invasion in glioblastoma cells through modulation of the miR-7-5p/TFF3 signaling pathway, Biochem Biophys Res Commun.499 (2018) 829-835.

[134] N.J. Martinez, A.J. Walhout, The interplay between transcription factors and microRNAs in genome-scale regulatory networks, Bioessays.31 (2009) 435-445.

[135] R.A. Moffitt, R. Marayati, E.L. Flate, K.E. Volmar, S.G. Loeza, K.A. Hoadley, N.U. Rashid, L.A. Williams, S.C. Eaton, A.H. Chung, J.K. Smyla, J.M. Anderson, H.J. Kim, D.J. Bentrem, M.S. Talamonti, C.A. Iacobuzio-Donahue, M.A. Hollingsworth, J.J. Yeh, Virtual microdissection identifies distinct tumor- and stroma-specific subtypes of pancreatic ductal adenocarcinoma, Nat Genet.47 (2015) 1168-1178.

[136] E.A. Collisson, A. Sadanandam, P. Olson, W.J. Gibb, M. Truitt, S. Gu, J. Cooc, J. Weinkle, G.E. Kim, L. Jakkula, H.S. Feiler, A.H. Ko, A.B. Olshen, K.L. Danenberg, M.A. Tempero, P.T. Spellman, D. Hanahan, J.W. Gray, Subtypes of pancreatic ductal adenocarcinoma and their differing responses to therapy, Nat Med.17 (2011) 500-503.

[137] P. Bailey, D.K. Chang, K. Nones, A.L. Johns, A.M. Patch, M.C. Gingras, D.K. Miller, A.N. Christ, T.J. Bruxner, M.C. Quinn, C. Nourse, L.C. Murtaugh, I. Harliwong, S. Idrisoglu, S. Manning, E. Nourbakhsh, S. Wani, L. Fink, O. Holmes, V. Chin, M.J. Anderson, S. Kazakoff, C. Leonard, F. Newell, N. Waddell, S. Wood, Q. Xu, P.J. Wilson, N. Cloonan, K.S. Kassahn, D. Taylor, K. Quek, A. Robertson, L. Pantano, L. Mincarelli, L.N. Sanchez, L. Evers, J. Wu, M. Pinese, M.J. Cowley, M.D. Jones, E.K. Colvin, A.M. Nagrial, E.S. Humphrey, L.A. Chantrill, A. Mawson, J. Humphris, A. Chou, M. Pajic, C.J. Scarlett, A.V. Pinho, M. Giry-Laterriere, I. Rooman, J.S. Samra, J.G. Kench, J.A. Lovell, N.D. Merrett, C.W. Toon, K. Epari, N.Q. Nguyen, A. Barbour, N. Zeps, K. Moran-Jones, N.B. Jamieson, J.S. Graham, F. Duthie, K. Oien, J. Hair, R. Grutzmann, A. Maitra, C.A. Iacobuzio-Donahue, C.L. Wolfgang, R.A. Morgan, R.T. Lawlor, V. Corbo, C. Bassi, B. Rusev, P. Capelli, R. Salvia, G. Tortora, D. Mukhopadhyay, G.M. Petersen, I. Australian Pancreatic Cancer Genome, D.M. Munzy, W.E. Fisher, S.A. Karim, J.R. Eshleman, R.H. Hruban, C. Pilarsky, J.P. Morton, O.J. Sansom, A. Scarpa, E.A. Musgrove, U.M. Bailey, O. Hofmann, R.L.

Sutherland, D.A. Wheeler, A.J. Gill, R.A. Gibbs, J.V. Pearson, N. Waddell, A.V. Biankin, S.M. Grimmond, Genomic analyses identify molecular subtypes of pancreatic cancer, Nature.531 (2016) 47-52.

[138] a.a.d.h.e. Cancer Genome Atlas Research Network. Electronic address, N. Cancer Genome Atlas Research, Integrated Genomic Characterization of Pancreatic Ductal Adenocarcinoma, Cancer Cell.32 (2017) 185-203 e113.

[139] T. Sorlie, C.M. Perou, R. Tibshirani, T. Aas, S. Geisler, H. Johnsen, T. Hastie, M.B. Eisen, M. van de Rijn, S.S. Jeffrey, T. Thorsen, H. Quist, J.C. Matese, P.O. Brown, D. Botstein, P.E. Lonning, A.L. Borresen-Dale, Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications, Proc Natl Acad Sci U S A.98 (2001) 10869-10874.

[140] C.M. Perou, T. Sorlie, M.B. Eisen, M. van de Rijn, S.S. Jeffrey, C.A. Rees, J.R. Pollack, D.T. Ross, H. Johnsen, L.A. Akslen, O. Fluge, A. Pergamenschikov, C. Williams, S.X. Zhu, P.E. Lonning, A.L. Borresen-Dale, P.O. Brown, D. Botstein, Molecular portraits of human breast tumours, Nature.406 (2000) 747-752.

[141] C. Sotiriou, S.Y. Neo, L.M. McShane, E.L. Korn, P.M. Long, A. Jazaeri, P. Martiat, S.B. Fox, A.L. Harris, E.T. Liu, Breast cancer classification and prognosis based on gene expression profiles from a population-based study, Proc Natl Acad Sci U S A.100 (2003) 10393-10398.

[142] X. Dai, H. Cheng, Z. Bai, J. Li, Breast Cancer Cell Line Classification and Its Relevance with Breast Tumor Subtyping, J Cancer.8 (2017) 3131-3141.

[143] L.M. Heiser, A. Sadanandam, W.L. Kuo, S.C. Benz, T.C. Goldstein, S. Ng, W.J. Gibb, N.J. Wang, S. Ziyad, F. Tong, N. Bayani, Z. Hu, J.I. Billig, A. Dueregger, S. Lewis, L. Jakkula, J.E. Korkola, S. Durinck, F. Pepin, Y. Guan, E. Purdom, P. Neuvial, H. Bengtsson, K.W. Wood, P.G. Smith, L.T. Vassilev, B.T. Hennessy, J. Greshock, K.E. Bachman, M.A. Hardwicke, J.W. Park, L.J. Marton, D.M. Wolf, E.A. Collisson, R.M. Neve, G.B. Mills, T.P. Speed, H.S. Feiler, R.F. Wooster, D. Haussler, J.M. Stuart, J.W. Gray, P.T. Spellman, Subtype and pathway specific responses to anticancer

compounds in breast cancer, Proc Natl Acad Sci U S A.109 (2012) 2724-2729.

[144] A. Sadanandam, C.A. Lyssiotis, K. Homicsko, E.A. Collisson, W.J. Gibb, S. Wullschleger, L.C. Ostos, W.A. Lannon, C. Grotzinger, M. Del Rio, B. Lhermitte, A.B. Olshen, B. Wiedenmann, L.C. Cantley, J.W. Gray, D. Hanahan, A colorectal cancer classification system that associates cellular phenotype and responses to therapy, Nat Med.19 (2013) 619-625.

[145] M. Fleming, S. Ravula, S.F. Tatishchev, H.L. Wang, Colorectal carcinoma: Pathologic aspects, J Gastrointest Oncol.3 (2012) 153-173.

[146] L.J. Mekenkamp, K.J. Heesterbeek, M. Koopman, J. Tol, S. Teerenstra, S. Venderbosch, C.J. Punt, I.D. Nagtegaal, Mucinous adenocarcinomas: poor prognosis in metastatic colorectal cancer, Eur J Cancer.48 (2012) 501-509.

[147] B. Perez-Villamil, A. Romera-Lopez, S. Hernandez-Prieto, G. Lopez-Campos, A. Calles, J.A. Lopez-Asenjo, J. Sanz-Ortega, C. Fernandez-Perez, J. Sastre, R. Alfonso, T. Caldes, F. Martin-Sanchez, E. Diaz-Rubio, Colon cancer molecular subtypes identified by expression profiling and associated to stroma, mucinous type and different clinical behavior, BMC Cancer.12 (2012) 260.

[148] S. Srivastava, S. Kedia, S. Kumar, V. Pratap Mouli, R. Dhingra, V. Sachdev, V. Tiwari, L. Kurrey, R. Pradhan, V. Ahuja, Serum human trefoil factor 3 is a biomarker for mucosal healing in ulcerative colitis patients with minimal disease activity, J Crohns Colitis.9 (2015) 575-579.

[149] Y. Ishibashi, H. Ohtsu, M. Ikemura, Y. Kikuchi, T. Niwa, K. Nishioka, Y. Uchida, H. Miura, S. Aikou, T. Gunji, N. Matsuhashi, Y. Ohmoto, T. Sasaki, Y. Seto, T. Ogawa, K. Tada, S. Nomura, Serum TFF1 and TFF3 but not TFF2 are higher in women with breast cancer than in women without breast cancer, Sci Rep.7 (2017) 4846.

[150] T.P. Radon, N.J. Massat, R. Jones, W. Alrawashdeh, L. Dumartin, D. Ennis, S.W. Duffy, H.M. Kocher, S.P. Pereira, L. Guarner posthumous, C. Murta-Nascimento, F.X. Real, N. Malats, J. Neoptolemos, E. Costello, W. Greenhalf, N.R. Lemoine, T. Crnogorac-Jurcevic, Identification of a Three-

Biomarker Panel in Urine for Early Detection of Pancreatic Adenocarcinoma, Clin Cancer Res.21 (2015) 3512-3521.

[151] M. Vocka, D. Langer, J. Petrtyl, P. Vockova, T. Hanus, M. Kalousova, T. Zima, L. Petruzelka, Trefoil factor family (TFF) proteins as potential serum biomarkers in patients with metastatic colorectal cancer, Neoplasma.62 (2015) 470-477.

[152] Y. Qu, Y. Yang, D. Ma, W. Xiao, Increased trefoil factor 3 levels in the serum of patients with three major histological subtypes of lung cancer, Oncol Rep.27 (2012) 1277-1283.

[153] Z. Huang, X. Zhang, H. Lu, L. Wu, D. Wang, Q. Zhang, H. Ding, Serum trefoil factor 3 is a promising non-invasive biomarker for gastric cancer screening: a monocentric cohort study in China, BMC Gastroenterol.14 (2014) 74.

[154] E.M. Vestergaard, M. Borre, S.S. Poulsen, E. Nexo, N. Torring, Plasma levels of trefoil factors are increased in patients with advanced prostate cancer, Clin Cancer Res.12 (2006) 807-812.

[155] L. Thim, E. Mortz, Isolation and characterization of putative trefoil peptide receptors, Regul Pept.90 (2000) 61-68.

[156] X.D. Tan, W. Hsueh, H. Chang, K.R. Wei, F. Gonzalez-Crussi, Characterization of a putative receptor for intestinal trefoil factor in rat small intestine: identification by in situ binding and ligand blotting, Biochem Biophys Res Commun.237 (1997) 673-677.

[157] J. Kalabis, I. Rosenberg, D.K. Podolsky, Vangl1 protein acts as a downstream effector of intestinal trefoil factor (ITF)/TFF3 signaling and regulates wound healing of intestinal epithelium, J Biol Chem.281 (2006) 6434-6441.

[158] Z. Yong, W. Lin, S. Yong, L. Guang-Ping, W. Dan, L. Shang-Jun, W. Wei, P. Xi, Kinetic characterization of an intestinal trefoil factor receptor, PLoS One.8 (2013) e74669.

[159] F.E. May, The potential of trefoil proteins as biomarkers in human cancer, Biomark Med.6 (2012) 301-304.

[160] M. Katoh, Trefoil factors and human gastric cancer (review), Int J Mol Med.12 (2003) 3-9.

[161] S. Kjellev, The trefoil factor family - small peptides with multiple functionalities, Cell Mol Life Sci.66 (2009) 1350-1369.

[162] T. Tanaka, J. Nakamura, Y. Kitajima, K. Kai, S. Miyake, M. Hiraki, T. Ide, Y. Koga, H. Noshiro, Loss of trefoil factor 1 is regulated by DNA methylation and is an independent predictive factor for poor survival in advanced gastric cancer, Int J Oncol.42 (2013) 894-902.

[163] D.E. Peterson, N.P. Barker, L.I. Akhmadullina, I. Rodionova, N.Z. Sherman, I.S. Davidenko, G.N. Rakovskaya, E.A. Gotovkin, S.A. Shinkarev, M.V. Kopp, E.P. Kulikov, V.M. Moiseyenko, J.M. Gertner, I. Firsov, T. Tuleneva, A. Yarosh, C.W. Woon, Phase II, randomized, double-blind, placebo-controlled study of recombinant human intestinal trefoil factor oral spray for prevention of oral mucositis in patients with colorectal cancer who are receiving fluorouracil-based chemotherapy, J Clin Oncol.27 (2009) 4333-4338.

[164] C. Hanisch, J. Sharbati, B. Kutz-Lohroff, O. Huber, R. Einspanier, S. Sharbati, Trefoil factor 3 mediates resistance to apoptosis in colon carcinoma cells by a regulatory RNA axis, Cell Death Dis.8 (2017) e2660.

[165] X. Yio, J.Y. Zhang, M. Babyatsky, A. Chen, J. Lin, Q.X. Fan, J.L. Werther, S. Itzkowitz, Trefoil factor family-3 is associated with aggressive behavior of colon cancer cells, Clin Exp Metastasis.22 (2005) 157-165.

[166] F.E. May, J.I. Semple, J.L. Newton, B.R. Westley, The human two domain trefoil protein, TFF2, is glycosylated in vivo in the stomach, Gut.46 (2000) 454-459.

[167] D. Bonar, F.G. Hanisch, Trefoil factor family domains represent highly efficient conformational determinants for N-linked N,N'-di-N-acetyllactosediamine (LacdiNAc) synthesis, J Biol Chem.289 (2014) 29677-29690.

[168] M.R. Aberle, R.A. Burkhart, H. Tiriac, S.W.M. Olde Damink, C.H.C. Dejong, D.A. Tuveson, R.M. van Dam, Patient-derived organoid models help define

personalized management of gastrointestinal cancer, Br J Surg.105 (2018) e48-e60.

General Hypothesis and objective

Pancreatic Cancer (PC) is one of the vicious cancers as it ranks third in the race of leading cause of cancer-related death. Lack of early diagnostic marker, poor understanding of molecular mechanism of the disease and failure to conventional chemotherapy makes this disease dreadful. Mucin 4 (MUC4), a highly upregulated molecule in PC and has role increasing cell proliferation, invasion, chemotherapy resistance, tumor growth and metastasis of PC. MUC4/X is devoid of exon 2, and 3 and MUC4/Y is devoid of exon 2. Exon 2 encodes for the largest domain of MUC4 suggesting that MUC4/X is devoid of the largest domain of MUC4 which variable tandem repeat. Though lots of effort has been made to identify its role in PC, there is still a gap in understanding its splice variant in PC as splice variant has an invaluable role in tumor pathogenesis. Recently splice variant has emerged as one of the key players for tumorigenesis and MUC4 is one of the key player for PC pathogenesis, we aim to identify the functional and mechanical role of MUC4/X, a splice variant which is devoid of the largest domain of MUC4 yet contains all other functional domain, in PC pathogenesis. We hypothesize that because MUC4/X in devoid of tandem repeat which is a major site for O-glycosylation, it is involved in PC tumorigenesis by more adhesion to extracellular matrix

Lack of early effective diagnostic marker and resistance to chemotherapy are the major reasons for poor PC patient outcome. There is a pressing need to identify highly specific and sensitive biomarker as well precise understanding of chemoresistance of PC. Trefoil factors (TFFs) are small secretory molecules mostly associated with mucin. Their primary role is to protect gastrointestinal tract

partnering with mucin. Report on aberrant expression, potential as biomarker and role in tumorigenicity has conveyed for many cancers, however, their role in PC is still elusive. Recently they have emerged as a part of gene signature of classical subtype of PC, a subtype which showed gemcitabine resistance towards PC. As it is high time to identify effective biomarker and understanding the role of chemoresistance in PC, in this part of my thesis, I focused to evaluate TFFs diagnostic potential using a training and validation cohort of PC clinical sample to identify their clinical significance as biomarker to detect early stage of PC. Additionally, I also aim to identify the molecular landscape of TFFs role in gemcitabine resistance of PC which integrates analyzing publicly available cancer genome dataset, dissecting transcriptomic and signaling pathways and identification of biochemical interaction.

Broadly, the aims for my dissertation research were as follows:

1. To delineate the functional and molecular mechanism of MUC4/X in PC pathogenesis.
2. To evaluate the diagnostic potential of Trefoil Factors in PC
3. To dissect functional and mechanical aspects of TFF1 in gemcitabine resistance

Chapter 2: Methods and Materials

(A) Functional and mechanical role of MUC4/X in PC pathogenesis

Clinical sample: Pancreatic tumor tissues and adjacent normal tissues were obtained from the University of Nebraska Medical Center (UNMC) rapid autopsy program (RAP). The study was approved by the Institutional Review Board (IRB) at UNMC, and all participants were consented before tissue collection (IRB-091-01). Tumors were flash frozen in liquid nitrogen and stored at −80°C until analysis.

RNA isolation from cell and frozen tissue, reverse transcription and real-time PCR

Total RNA from cells and frozen tissues were isolated using a mirVana miRNA kit (Ambion, Austin, TX, USA). RNA was reverse transcribed by using 1 µg of total RNA with random hexamer oligos (500 µg/ml), 1µl of 10mM dNTPs, 5x first-strand reverse transcriptase buffer, 1µl of 0.1M dithiothreitol and 1µl of (50 unit) SuperScript RT as described previously [169]. Briefly, 10ng of complementary DNA was amplified using LightCycler® 480 SYBR Green I master mix (Roche Diagnostics, IN, USA) in the Light Cycler 480II (Roche Diagnostics, IN, USA). The amplification was performed in a two-step cyclic process (95°C for 5 min, followed by 45 cycles of 95°C for 10s, 60°C for 10s and 72°C for 10s). The relative expression of mRNA (ΔCt) was normalized with β-actin, and the relative fold change (ΔΔCt) was measured in reference to a normal human pancreatic ductal epithelial (HPDE) cell line. The WT-MUC4 and MUC4/X expression in clinical samples were analyzed and expressed as fold change (log10 transformed) relative

to control group (HPDE). The qPCR primers used are listed in **Supplementary Table S1**.

Cell lines

MIAPaCa, Capan-1, AsPC-1 and CD18/HPAF PC cell lines were obtained from ATCC, and grown in Dulbecco's Modified Eagle's medium (DMEM) containing high glucose (Hyclone,Thermo USA), supplemented with 10% (v/v) fetal bovine serum and 1% penicillin-streptomycin (HyClone, Thermo, USA) at 37°C in a humidified atmosphere containing 5% CO_2. Human mesothelial LP9/TERT-1 cells, an hTERT-immortalized cell line phenotypically and functionally resembling normal human peritoneal mesothelial cells, were obtained from Dr. James Rheinwald (Brigham and Women's Hospital, Harvard Institute of Medicine, Boston, MA) and cultured as detailed previously [170].

Generation and expression of MUC4/X construct

Standard PCR and molecular cloning techniques were utilized to generate MUC4/X overexpression constructs as detailed previously [171]. Briefly, the MUC4/X was amplified by RT-PCR from our miniMUC4 construct, cloned into the mammalian expression vector p3X-FLAG-CMV9 (Sigma-Aldrich, St. Louis, MO, USA) [172] and verified by sequencing. MIAPaCa and AsPC-1 PC cells were stably transfected using lipofectamine (Invitrogen, CA, USA) and stable clones were selected using G418 (400 µg/ml). Cells were analyzed for MUC4/X protein expression by immunoblotting and immunofluorescence using anti-FLAG antibody (1:3000, M2 clone, Sigma-Aldrich).

Generation of Tet-on inducible system

The DNA fragment encoding MUC4/X was amplified by RT-PCR from the p3X-FLAG-CMV9-MUC4/X construct and subcloned into a Topo 2.1 cloning vector using TOPO TA Cloning® kit (Invitrogen, CA, USA). Using SalI and SpeI restriction digestion enzymes, the desired MUC4/X DNA fragment was digested from the Topo-TA, cloned into pTet-Splice vector (Invitrogen, CA, USA) and sequenced. Tetracycline-inducible MUC4/X expressing PC cell line Capan-1 was generated as detailed previously [173]. Briefly, Capan-1 cells were transduced with rtTA lentiviral particles according to the manufacturer's recommendations (Gentarget, Inc, CA, USA) and stably transfected clones were selected using puromycin (2 µg/ml). Subsequently, these rtTA transduced stable Capan-1 cells were transfected with pTet-Splice-MUC4/X plasmid, and after 48 hrs of transfection, MUC4/X expression was induced using doxycycline (2 µg/ml). MUC4/X expression was confirmed by performing immunoblotting and immunofluorescence using anti-FLAG antibody.

Cell proliferation assay (MTT assay)

The effect of MUC4/X overexpression on the viability/ proliferation of PC cells was determined using MTT (3- [4, 5-dimethylthiazol-2-yl]-2, 5 diphenyl tetrazolium bromide) assay as described previously [174]. PC cells (MIAPaCa and AsPC-1) with either an empty vector (MIAPaCa-EV and AsPC-1-EV) as control and MUC4/X over expressed MIAPaCa-MUC4/X and AsPC-1-MUC4/X cells (5×10^3/well) were seeded in triplicates onto a 96-well plate for 24-96 hrs in the presence of 2% serum containing media. Cell viability was assessed by adding 10µl of 5 mg/ml of MTT in each well containing 90µl of media (final working concentration of MTT is 0.5 mg/ml) followed by incubation for 3-4 hrs at 37°C.

100µl of dimethyl sulphoxide (DMSO) was added to dissolve formed formazan crystals. Optical density (OD) was measured at a wavelength of 450 nm and data collected was analyzed using the SOFTMAX PRO software (Molecular Devices Corp., Sunnyvale, CA, USA).

Ethynyl-2-deoxyuridine (EdU) incorporation assay

To assess the cell proliferation, control and MUC4/X overexpressing (MUC4/X-OE) cells ($1X10^6$ /ml) were seeded in six-well plates onto the sterilized coverslip and analyzed for incorporation of 10µM EdU (5-ethynyl-2′-deoxyuridine) using Click-iT EdU Cell Proliferation Assay Kit (Click-iT EDU kit, Thermo Fisher Scientific, MA, USA). Briefly, after incubating the cells with 10µM of EdU for 24 hrs under dark conditions, cells were washed thrice with PBS, fixed with 4% paraformaldehyde and permeabilized with 0.1% Triton X-100 in PBS. The cell nuclei were stained with DAPI (4′, 6-Diamidino-2-Phenylindole, Dihydrochloride, Sigma-Aldrich, St. Louis, MO, USA) and visualized under a Zeiss confocal laser-scanning microscope (Carl Zeiss Microimaging, Thornwood, NY, USA). Percentage of proliferative cells were calculated by the number of EdU positive cells per field/total number of DAPI positive cells per field (*100).

Colony formation assay

500 cells/well were seeded in triplicate in a 6 well plate in 2% FBS containing medium [174]. After 2 weeks, cells were fixed with 100% methanol, stained with 0.4% crystal violet, and colonies were counted [174].

Western blotting

Western blotting was performed as described previously [174]. PC cells in log phase were lysed with ice-cold RIPA (50mM Tris-HCl, 150mM NaCl, 1% NP-40, 0.5% sodium deoxycholate and 0.1% SDS) containing protease and phosphatase inhibitors. Cell lysates was quantified, and equal amount of proteins were resolved on 8-12% SDS-PAGE (for <250 kDa molecular weight proteins) and 2% agarose gel electrophoresis (for high molecular weight MUC4) and transferred onto polyvinylidene difluoride (PVDF) membranes. Membranes were blocked with 5% nonfat milk in phosphate-buffered saline (PBS) for 2 hrs, followed by incubation with primary antibodies at 4°C overnight. Antibodies used includes: anti-FLAG (1:3000; Clone-M2 ,Sigma-Aldrich), MUC4 (In-house generated, 1:1000), β-Actin (1:5000, Sigma-Aldrich), integrin-β1 (1:1000, Cell Signaling), phospho-ERK (1:1000, Cell Signaling), FAK (1:500, Santa Cruz Biotechnology), p-FAK (1:1000, Cell Signaling), HA (1:2000, Cell Signaling), total ERK (1:1000, Cell Signaling). After overnight incubation, membranes were washed (3X10 min in each time) and probed with appropriate secondary antibodies (1:10000) for 1hr at room temperature. The protein of interest was detected by enhanced chemiluminescence (Thermo Fisher Scientific, MA, USA). Protein levels were normalized with β-Actin to determine fold changes using ImageJ software (version 1.50i, NIH).

Immunofluorescence

Immunofluorescence analysis was performed as detailed previously [174]. PC cells were grown to 60 - 70% confluency on sterilized coverslips, washed with Hanks buffer, and fixed in ice-cold methanol for 2 min. Fixed cells were then

blocked with 10% goat serum for 30 min at room temperature followed by incubation with primary anti-FLAG antibody (1 : 200) for 60 min at room temperature. After three consecutive washings with PBST, cells were incubated with Alexa Fluor-conjugated secondary antibody (1:200) for 60 min at room temperature. Cells were then washed with PBST (three times and 10 min in each time) and mounted using anti-fade vectashield mounting medium (Vector Laboratories Inc., Burlingame, CA, USA). Immunofluorescence was observed under Zeiss confocal laser-scanning microscope (Carl Zeiss Microimaging, Thornwood, NY, USA).

Invasion assay

Invasion assay was carried as described previously [169]. Briefly, six-well cell culture inserts were coated with Matrigel (BD Biosciences, Bedford, MA). Next, 5×10^5 cells/well were seeded on top of the chamber in triplicate, in serum-free media. Serum-containing medium was added to the lower chamber of the well. After 24 hrs, non-invaded cells were removed using cotton swabs, and cells that invaded into the lower side of insert were fixed, stained, and photographed (Dade-Behring Inc, Newark, DE 19714, USA). Photographs were taken from 5 random fields from each insert to count numbers of invaded cells per area. The experiment was repeated at least three times.

Wound healing assay

Wound healing assay was performed as per the standard protocol described previously [175]. Control and MUC4/X–OE cells were seeded at a density of 2×10^6

cells/well in six-well plate in growth medium and allowed to reach ~90% confluency. After 24 hrs, an artificial wound was created on the monolayer of the cells using a 200µl sterile pipette tip, washed with PBS to remove damaged and detached cells and allowed the cells to migrate in 2% serum containing media. Images of control and MUC4/X-OE cells were taken at 0, 12 and 24 hrs and percentage of the wound closure was determined by the percentage of the ratio of wound width at 24 hrs to the wound width at 0 hr. The experiment was repeated thrice.

Adhesion assay

Adhesion assays were performed using Millicoat extracellular matrix (ECM) screening kit (#ECM 205 Millipore, Massachusetts, USA) which includes 96 well plate precoated with ECM proteins such as collagen I, collagen IV, fibronectin, vitronectin, and laminin. Control and MUC4/X-OE cells (2×10^5 cells/well) were plated in each well and allowed to attach for 5 hrs at 37°C. Shorter incubation time (5 hrs) was chosen to avoid variations that arise due to differential growth kinetics. Floating cells were carefully aspirated, and plates were washed with PBS. Attached cells were stained using crystal violet (0.1%, w/v in acetone) and solubilized using 200 µl of DMSO. Absorbance (at 570 nm) was measured using a microplate reader and data was analyzed using the SOFTMAX PRO software (Molecular Devices Corp., Sunnyvale, CA). Adhesion assays was performed twice in triplicate.

Mesothelial peritoneal cell adhesion assay

Adhesion of control and MUC4/X-OE cells to a monolayer of peritoneal mesothelial cells (LP9/TERT-1) was done as previously described [176]. Briefly, 2×10^4 LP9/TERT-1 cells per well were plated in flat-bottom 96-well plates and allowed to grow to a confluent monolayer for 48 hrs. Control and MUC4/X-OE cells were trypsinized, washed with PBS, and probed with 5 µmol/L cell tracker green fluorescent CMFDA dye (Thermo Fischer Scientific, MA, USA) for 30 minutes at 37°C. CMFDA labeled cells were washed with M199 medium with 0.1% fetal bovine serum to remove the free dye. Control and MUC4/X-OE cells were added (3×10^4 cells/well) onto the top of the monolayer of mesothelial cells. After 4 hrs incubation at 37°C, total fluorescence in each well was recorded in a spectrofluorimeter (Perkin-Elmer, Turku, Finland) using 485-nm and 535-nm wavelengths for excitation and emission, respectively. Then, nonadherent cells were removed by gentle washing, followed by measurement of fluorescence. To calculate percentage of bound cells, values recorded for bound cell were compared with total fluorescence from cells before washing [177]. This assay was repeated twice in triplicates each time.

In vivo tumorigenesis and metastasis assay

Mouse orthotopic implantation experiments were performed under a protocol approved by the Institutional Animal Care and Use Committee (IACUC). Four to six-week-old athymic nude mice (equal number of males and females, housed separately) were purchased and maintained as described previously [169, 171]. Briefly, after anesthetizing the mice by intraperitoneal administration of 120 mg/kg

ketamine and 16 mg/kg xylazine, exponentially growing MIAPaCa-EV and MIAPaCa-MUC4/X cells (2.5×10^5 cells /50 µl of PBS) were orthotopically implanted into the head of the pancreas (vector control, n=7, and MUC4/X, n=8) [169]. To evaluate tumor growth and metastasis, mice were sacrificed by CO_2 asphyxiation and autopsied on the 50^{th} day after implantation of tumor cells. After inspection of macroscopic tumor growth, the pancreas was resected and weighed. Both primary tumors and organs with metastases were kept in 10% formalin for 48 h followed by embedding in paraffin blocks. Blocks were sectioned into 0.5µM-thick sections and processed for histochemical analysis.

H&E and immunohistochemical (IHC) staining

Tissues sections were stained with hematoxylin and eosin (H&E), and IHC was done using anti-FLAG (1:200, Sigma-Aldrich), anti-Ki67 (1:200, Cell Signaling Technology), anti-integrin-β1 (1:300, Cell Signaling Technology) antibodies as described previously [178].

Microarray Gene Expression Analysis

Total RNA was isolated using the mirVana miRNA kit (Ambion, Austin, TX, USA). RNA yield and purity were measured using a Nanodrop (NanoDrop 1000 spectrophotometer, Thermo Scientific, Delaware USA). Affymetrix microarray was used to identify differentially expressed genes between MIAPaCa-EV and MIAPaCa-MUC4/X cells. Whole genome gene expression profiles were determined by the Microarray Core Facility at the University of Kansas Medical Center (UKMC, USA) using Human Gene 2.0 ST arrays (Affymetrix, Santa Clara,

CA, USA). Target preparation, library labeling, hybridization, post-washing, and signal scanning were performed according to the protocol by UKMC. Microarray data were analyzed using Affymetrix Power software. Ingenuity pathway analysis (IPA) (QIAGEN, Valencia, CA, US) was performed to define canonical pathway differences between control and MUC4/X-OE PC cells and generate networks.

Statistical analysis

Statistical analysis was performed using Medcalc for Windows version 9.6.4.0 software *(MedCalc Software,* Mariakerke, Belgium) for analyzing patient data. Statistical differences between two groups was analyzed using unpaired, two-tailed t-test. The quantification is shown as the mean ± SD for *in vitro* studies where statistics refer to technical replicas denoted by "n".

(B) Clinical significance of Trefoil Factors in diagnosing early stage of PC

Sandwich Enzyme-Linked Immunosorbent Assay (ELISA)

Individual TFF levels in serum were measured quantitatively by sandwich ELISA in accordance with manufacturer instructions, using the DuoSet ELISA kit with few modifications (R&D Systems, Catalog for TFF1: DY5237, for TFF2: DY4077, and for TFF3: DY4407). Before performing the analysis, assay optimization was performed to select appropriate positive controls and the sample dilution factor to be used for ELISA. Standard curves were produced from standard provided with the kit. TFF1 and TFF2 standards were serially (log_2) diluted from 250 pg/ml to 0.4 pg/ml, and TFF3 standard from 750 pg/ml to 1.46 pg/ml. A detailed ELISA protocol is provided in the supplementary section. For the detection step, instead of using the manufacturer provided streptavidin-HRP, we used Pierce Streptavidin Poly-

HRP (Thermo Scientific, USA), (diluted to 0.4 µl/ml in 1% BSA) and incubated for 20 minutes in the dark at room temperature. ELISA plates were read at 450 nm with an absorbance correction at 540 nm. Appropriately diluted sample lysates from TFF expressing cell lines were used as a positive control (MCF7 breast cancer cells for TFF1, LS174T colon cancer cell line for TFF2 and AsPC-1 PC cell line for TFF3). CA19.9 serum levels were measured using DRG® CA 19-9 ELISA (EIA-3940) kit (DRG International, Inc., NJ, USA). Serum samples were diluted appropriately to obtain absorbance in the linear range of the assay. Area under curve (AUC) for CA19.9 analysis was performed with the standard clinical cutoff value (37 U/ml) [179]. All samples were tested in duplicate. Data was analyzed using SOFTMAX PRO software (Molecular Devices Corp., Sunnyvale, CA).

Study cohorts

Two independent sets of samples (exploratory and validation sets) were used to assess the diagnostic performance of TFFs in PC patient serum. Both sample sets were obtained from University of Pittsburgh Medical Center (UPMC, IRB number PRO07030072), with written consent from all patients prior to enrollment in the study. Training sets (n=312) included benign control (BC, n=107), chronic pancreatitis (CP, n=47), and PC (n=158). Within the PC group, samples from early stage (EPC, Stage 1 and 2, n=80) and late stage PC (LPC, Stage 3 and 4, n=73) were segregated for further analysis [180]. Serum samples were shipped from UPMC by overnight mail to the University of Nebraska Medical Center (UNMC) for all experimental analyses. PC staging was determined surgically, based on operative pathology, biopsy of metastatic disease, or radiographic imaging studies.

The grade location of the tumor, and stage, were based on reviewed hospital records. Diagnostic significance of TFFs were further assessed in an independent blinded validation set (n=58) containing BC (n=8), CP (n=27) and PC (n=23). Patient demographic information for both training and validation datasets is included in the Supplementary table S1A-S1B.

Trefoil factor expression in publicly available PC datasets

GEO datasets containing PC-specific gene expression and patient clinical information were used to assess expression of TFFs. Further, to make the comparisons statistically significant and reduce any chance of method-induced statistical bias, two independent datasets (GSE16515 and GSE43288) were selected containing normal, PanIN, and PC samples. For dataset GSE43288, samples were profiled using Affymetrix Human Genome U133A Array (Affymetrix, Inc., Santa Clara, CA), which contains normal pancreas (n=3), PanIN (n=13) and PC (n=4) tumor tissues, while dataset GSE16515 samples were profiled using Affymetrix Human Genome U133 Plus 2.0 Array (Affymetrix, Inc., Santa Clara, CA) and contain normal pancreas (n= 16) and PC (n=36) tumor tissues.[181] Briefly, the raw CEL files were first downloaded and background-corrected. Expression was calculated using the quantile normalization method robust multi-array average (RMA) within the same package for each dataset. The normalized gene expression values (RMA) for TFFs in normal, PanIN, and PC were plotted using MedCalc software. We also analyze all TFFFs expression from cBioPortal database, a freely

available dataset which is comprised of published and provisional TCGA datasets consisting of 169 studies from 30 different tumor types.

Tissue immunohistochemistry (IHC) and immunofluorescence for TFF1, TFF2 and TFF3

Immunohistochemistry analyses were performed on commercial tissue microarray (TMA) (Biomax,USA) as well as TMA from the rapid autopsy program (RAP) at UNMC). Tissue spots on the Biomax TMA (BIC14011, OD-CT-DgPan03-001, OD-CT-DgPan01-006) were examined by their own pathologists to determine the pathological grading and staging. This array contained spots from healthy normal or normal adjacent to pancreatic tumor, chronic pancreatitis (CP), and PC of different stages and grades. Immunohistochemistry was performed as described previously. For staining TFF1, TFF2 and TFF3 in the paraffin embedded murine tissues, 10, 20, 30 and 40 weekold floxed $Kras^{G12D}$ mouse tissues (positive for both Kras and Pdx1-Cre), and their age matched littermate controls, were processed using methods describing previously.[182] Briefly, after deparaffinization with xylene and subsequent rehydration with ethanol, epitope retrieval was achieved by boiling the slides in citrate buffer (pH 6.0) for 15 min. Endogenous peroxidase activity was quenched by immersing the TMAs in hydrogen peroxidase solution [0.3% hydrogen peroxidase in 1:1 solution of methanol: water] for 1 hr. at room temperature, in the dark. Tissues were next blocked with horse serum (ImmPRESS Universal antibody kit; Vector Laboratories, Burlingame, CA) for 2 hrs. at room temperature. Subsequently, the TMAs were incubated with individual primary antibodies for TFF1 (1:200, ab92377, Abcam, Cambridge, MA),

TFF2 (Protein tech, 1:500) and TFF3 (ITF Antibody (FL-80): sc-28927, Santa Cruz for human tissue, for mouse tissue TFF3 antibody was kindly provided by Daniel K. Podolsky, UT Southwestern Medical Center, Dallas, TX)). After overnight incubation with primary antibodies, the slides were washed four times with PBS, followed by incubation with horseradish peroxidase (HRP)–conjugated secondary antibody (ImmPRESS Universal antibody kit; Vector Laboratories, Burlingame, CA) for 1 hr. Following secondary antibody incubation, the TMAs were washed with PBS and color was developed using DAB solution (3,3′-diaminobenzidine solution) (DAB substrate kit; Vector Laboratories, Burlingame, CA). Once the reddishbrown precipitate was developed, the peroxidase reaction was quenched using distilled water, and sections were counterstained with hematoxylin for 1–2 min. After that, dehydration in an increasing percentage of ethanol followed by three consecutive washings with xylene was performed. Finally, the sections were mounted using Vecta-mount mounting medium (Vector Laboratories). Each tissue spot was evaluated by Dr. Yuri Sheinin (pathologist) for the *H*-score, which is the product of the percentage of cells positive in the area for each TFF and intensity of staining on a 0-3 scale (0=no staining, 1=weak staining, 2=moderate staining and 3= strong staining). Colorectal cancer tissues were used as positive controls and no primary antibody control was used as negative control (image not shown). The slides were scanned with Ventana iScan HT from (Ventana Medical Systems, Inc.; Roche Group, Tucson, AZ).

For tissue immunofluorescence studies on TFFs, the initial steps up to blocking, were similar to those of IHC. Following blocking, the slides were incubated with an

appropriate dilution of primary antibodies (1:25, 1:50 and 1:25 for TFF1, TFF2 and TFF3 respectively, the same antibody used for immunohistochemistry) overnight at 4°C. Primary antibody detection, and visualization of TFF1-3 was achieved by Alexa Fluor 488 labelled goat anti-rabbit IgG2b (Invitrogen Molecular Probes, Eugene, OR, diluted 1:200), incubated for one hour and then washed with PBS three times. Finally, the sections were mounted using an anti-fading mounting medium containing DAPI (Vectashield H-1500; Vector Laboratories). The stained sections were analyzed by confocal scanning laser microscopy, using an LSM 510 Meta laser scanning microscope (Carl Zeiss, Jena, Germany) .

Statistical analysis

Serum concentration of each protein was calculated using GraphPad software (GraphPad Software, Inc., San Diego, CA). Logarithm transformation was applied to all serum analyses. If the biomarker had a zero level, before log transformation, the zero was changed to half the next lowest value for that marker. Biomarker levels were compared among assay groups with ANOVA. If the overall test was significant, then pairwise comparisons were conducted, adjusting for multiple comparisons with Tukey's method. Patient characteristics were compared by diagnosis using chi-square or Fisher's exact test for categorical variables and ANOVA for continuous variables. TFF1, TFF2, and TFF3 levels were compared with patient clinicopathological characteristics using t-tests or ANOVA. We examined the correlation between markers using Pearson correlation. Univariate and multivariate logistic regression and ROC curves were used to test individual markers and combinations as predictors of disease status. ROC curves were used

to determine the optimal marker cut points for discriminating the potential of an individual protein. SAS software version 9.3 was used for data analysis (SAS Institute Inc., Cary, NC).

(C) Functional and mechanical role of TFF1 in gemcitabine resistance of pancreatic cancer

Cells

SW1990 and COLO357 PC cell lines were obtained from ATCC, and grown in Dulbecco's Modified Eagle's medium (DMEM) containing high glucose (Hyclone,Thermo USA), supplemented with 10% (v/v) fetal bovine serum and 1% penicillin-streptomycin (HyClone, Thermo, USA) at 37 °C in a humidified atmosphere containing 5% CO2.

RNA isolation, reverse transcription, and real-time PCR

RNA isolation from cells and cDNA preparation was followed as described previously [183]. The relative expression of mRNA (ΔCt) was normalized with β-actin.

Measurement of Gemcitabine Sensitivity Ratio (GSR)

Individual gene expression was downloaded from cBIoportal database. GraphPad Prism 6 (La Jolla, CA, USA) was used to perform Pearson/Spearman correlation test (two-tailed) (La Jolla, CA, USA).

For the calculation of GSR, four gemcitabine metabolic genes quantified were taken into account- hENT1, RRM1 and -2 (ribonucleotide reductase gene), and DCK. GSR were first introduced by Nakano et al. [184]. The gene expression

values for each gene were determined and used in the following formula to obtain the ratio.

GSR= [hENT1]*[DCK]/[RRM1][RRM2]

Because the genes required for metabolism (hENT1 and DCK) are in the numerator and the genes responsible for reducing effectiveness of the drug (RRM1 and RRM2) are in the denominator, lower values of the ratio denote enhanced gemcitabine resistance [185]

Long-term gemcitabine treatment

SW1990 and COLO357 PC cells were treated with increasing concentration of gemcitabine continuously for 8 weeks. Briefly, concentration of gemcitabine for SW1990 range from 0.5 um-2.5 μM and for COLO357 0.5- 0.8 μM. The PC cells were treatment and recovery were in alternative two days/week for eight weeks. After eight weeks of gemcitabine treatment, the PC cells were used for further experiments.

SP analysis

To determine cancer stem cell population, we perform SP analysis. Briefly, $1X10^6$ / ml cells were used for each experimental group including control and test. Control cells in each group (Parental and gemcitabine-treated) were incubated with 50 μM Verapamil (Sigma, St Louis, MO, USA) for 15 min at 37°C. Then cells were incubated for an additional 60 min at 37°C in water bath with 5 μg /ml Hoechst 33342 dye. Cells were placed immediately on ice. SP and non-SP fractions were gated separately and analyzed using an LSRII flow cytometer (BD Biosciences, San Jose, CA, USA) [186, 187] .

TFF1 knockdown in PC cells

SW1990 PC cells were transfected with piLenti-siRNA-GFP (Applied Biological Materials Inc., Richmond, Canada) and empty vector was used as control. The transfection process was performed using Turbofect (Thermo Fischer) following the manufacturer's protocol. The knockdown of TFF1 was analyzed by immunoblotting analysis.

Colony Forming Assay

Colony forming assay was performed as previously described [183]. Briefly, 500 cells/ well were plated in 2% FBS containing medium. SW1990-Scr and SW1990-TFF1-KD cells were treated with gemcitabine 1µM for 14 days. Then cells were fixed with 100% methanol, stained with 0.4% crystal violet, and colonies were counted

Migration Assay

Migration assay was performed as previously described [188]. Motility assays, (0.5-0.8) $x10^6$ cells seeded in serum-free medium on the top chamber of polyethylene teraphthalate membranes (six-well insert, pore size 8µm). Two ml of 10% serum-containing medium was added to the lower chamber of the well and the cells were allowed to migrate for 24 h-48 hr depending upon cell under chemotactic drive. Cells that did not migrate through the pores were removed by scraping the membrane with a cotton swab. The migrated cells on the lower side of the membrane were stained with a Diff-Quick cell stain kit (Dade-Behring Inc., Newark, USA). So first Fixation (5 min), then cytoplasmic staining (5 min-orange color), and nuclear staining (five minute-dark blue color) and then quick dip 3 times

in water for removing excess color. Cut the membrane with a scalpel and put on the microscopic slide. Take photograph in ten random fields of view at 10X magnification. Cell numbers were counted and expressed as the average number of cells/field of view.

Apoptosis Assay

SW1990 Scr and SW1990-TFF1-KD was treated with staurosporine 2μM for 12 hrs (ab120056, Abcam, USA) and gemcitabine 1μM for 24 hrs (G6423, Sigma-Aldrich) and apoptosis assay was performed as previously described [189].

Protein Array

For the assay using apoptosis antibody array, protein lysates (200 μg) prepared from gemcitabine-treated SW1990-Scr, and SW1990-TFF1-Sh analysed by a Human Apoptosis Array Kit (Cat. # ARY009, R&D Systems, Minneapolis, MN, USA) according to the manufacturer's protocol. Appropriate dilutions of protein in the lysates were prepared as per the maximum allowable volume per array recommended by the manufacturer. The recommended quantity of lysates was diluted and pipetted onto the membranes and incubated overnight at 2–8°C on a rocking platform shaker. Biotinylated secondary antibody cocktail provided by the manufacturer was pipetted onto membranes and incubated for 1 h. After the washing process, the membranes were incubated with streptavidin-HRP provided by the manufacturer for 30 min. The signals were developed using chemiluminescent reagents and then exposed to X-ray films. The positive signals were analyzed using ImageJ software (National Institutes of Health, Bethesda, MD, USA).

Immunohistochemistry and immunofluorescence

Immunohistochemistry (IHC) and immunofluorescence (IF) procedure was performed as described previously [183] [190]. Primary antibody dilution used for PC human tissues GATA-6 (1:50, Cell Signaling), TFF1 (1:100, Abcam). For IF, primary antibody dilutions were MUC5AC (CLH2, 1:100), TFF1 (ab190942 1:50 for mouse tissue, ab92377 1:50 for human tissues, Abcam)

Immunoprecipitation

For immunoprecipitation study, SW1990 cells were grown to confluency, and then lysed in non-denaturing lysis buffer [20mM Tris-HCl pH 8.0, 137mM NaCl, 2mM81ethylenediaminetetraacetic acid (EDTA), 1% Nonidet P-40 (NP-40), 1mM NaF, 1mM sodium orthovanadate (Na3VO4), 1mM PMSF, aprotinin 5mg/ml, leupeptin 5mg/ml and containing 1% Triton X-100] for 25–35 min at 4°C. For pre-cleaning, the processed lysates were incubated with protein A+G Sepharose beads (Sigma-Aldrich Corp., St Louis, MO, USA) for 4 hours at 4°C on a rotator. Total protein concentration in pre-cleared lysates was quantified and equal amounts of total protein (1000 µg) in 500µl volumes of non-denaturing lysis buffer were then incubated overnight with anti-MUC5AC antibodies and with respective IgG at 4°C on a rotator. The protein-antibody complexes were incubated with protein A+G Sepharose beads on a rotator for 4 hours at 4°C. The pulled-down Immunocomplexes were washed with the lysis buffer (3X) followed by two wash with PBS. The immunoprecipitates and input were electrophoretically resolved and immunoblotted with anti-TFF1 antibodies.

Chromatin immunoprecipitation

Chromatin immunoprecipitation was performed as previously described [191]. The antibodies used for ChIP assay includes normal IgG rabbit, GATA-6 (D61E4), Cell Signaling). The immunoprecipitated DNA samples were PCR amplified using following primers TFF1 promoter primer. Promoter-1- FP AGCCAAGATGACCTCACCAC, TFF1 Promoter-1- RP AGCCCCGGATTTTATAGGG, TFF1 Promoter-2 FP TTCCGGCCATCTCTCACTAT, TFF1 Promoter-2 RP TCATCTTGGCTGAGGGATCT.

Protein-Protein Docking Study

The three-dimensional structures of CXCR4 (PDB code: 3ODU) and TFF1(PDB code: 1HI7) proteins were downloaded from Protein Data Bank (PDB; http://www.rcsb.org/) and prepared using Protein Preparation Wizard of Schrödinger suite as described in detail [192]. The protein-protein docking was performed using BioLuminate module (Schrödinger Release 2018-2: BioLuminate, Schrödinger, LLC, New York, NY, 2018) present in Schrödinger suite. BioLuminate searched for 50 best complexes from 70,0000 possible protein-protein configurations. The methodology has been described in detail [193]

siRNA knockdown

Knockdown of GATA6 was achieved by specific duplex siRNAs (50nmol/L) purchased from Origene Technologies (Cat. #SR319903). Transfection of siRNAs in SW1990 cells was performed with Turbofect (Thermo Fischer) according to the manufacturer’s instructions.

Statistical analysis

Correlation analyses were performed using GraphPad Prism 6.03 (GraphPad Software Inc.,). In all studies, data represent biological replicates (n) and are depicted as mean values ± s.d. as indicated in the figure legends. Comparison of mean values was conducted with unpaired, two-tailed Student's *t*-test as indicated in the figure legends.

Table 1: List of Primers used in this dissertation:

qPCR primer list (used in chapter 3)

Gene Name	Forward Primer	Reverse Primer
Sortilin 1	GGTGTCTGGAGGAAGTCGTG	ACACCCACAGGCCATTTTCA
Coronin 2B	GGAACCTGGATGTGGGTGAG	ATGGAGAGGTCCTCCTGGTC
Integrin B3	CATCCATAGCACCTCCACATAC	CCAGCCAACTCATGGGAATAA
INFA4	ATGAGGACTCCATCCTGGCT	CTGCTCTGACAACCTCCCAG
MUC4/X	TGGTCCCAGGAGTTTCCCTCTT	GGGAATCACGGAGAGAGGAGC
β-actin	TGGACATCCGCAAAGACCTG	CCGATCCACACGGAGTACTT
h-MUC4	GACTTGGAGCTCTTTGAGAATGG	TGCAATGGCAGACCACAGTCC
Exon 2-MUC4	ACAACCTCCACAGACTCCAC	GAGACTGCTGAGGTCACTGG
Exon 3-MUC4	GAAGACAGACGGTGGGAGAC	AGTGCTGGGAATGGTGGAAA

qPCR primer list (used in chapter 5)

Gene Name	Forward Primer	Reverse Primer
Promoter-1	AGCCAAGATGACCTCACCAC	AGCCCCGGATTTTATAGGG
Promoter-2	TTCCGGCCATCTCTCACTAT	TCATCTTGGCTGAGGGATCT
hENT1	CTCTCAGCCCACCAATGAAAG	CTCAACAGTCACGGCTGGAA
dCK	GAGAAACCTGAACGATGGTCTT	TCTCTGCATCTTTGAGCTTGC
RRM1	CTGCAACCTTGACTACTAAGCA	CTTCCATCACATCACTGAACACT
RRM2	CCACGGAGCCGAAAACTAAAG	CTCTGCCTTCTTATACATCTGCC
GATA-6	GACTTGCTCTGGTAATAG	CTGTAGGTTGTGTTGTGG
TFF1	CAGACAGAGACGTGTACAGTGGCCC	AGCGTGTCTGAGGTGTCCGGT

Table 2: List of antibodies used in this dissertation

Name of the antibodies	Company	Dilution
anti-FLAG	Clone-M2 ,Sigma-Aldrich)s	1:3000
MUC4	In-house generated	1:1000
β-Actin	Sigma-Aldrich	1:5000
integrin-β1	Cell Signaling	1:1000
phospho-ERK	Cell Signaling	1:1000
FAK	Santa Cruz Biotechnology	1:500
p-FAK	Cell Signaling	1:1000
HA	Cell Signaling	1:2000
total ERK	Cell Signaling	1:1000
TFF1	ab92377 (human tissue)	1:200
TFF1	ab190942 (mouse tissue)	1:50
TFF1	Cell Signaling	1:1000
TFF2	Protein tech	1:700
TFF3	Santa Cruz Biotechnology	1:50
GATA-6	Cell Signaling	1:50
MUC5AC	CLH2	1:100

Reference

[1] S. Rachagani, M.A. Macha, M.P. Ponnusamy, D. Haridas, S. Kaur, M. Jain, S.K. Batra, MUC4 potentiates invasion and metastasis of pancreatic cancer cells through stabilization of fibroblast growth factor receptor 1, Carcinogenesis.33 (2012) 1953-1964.

[2] M.A. Dickson, W.C. Hahn, Y. Ino, V. Ronfard, J.Y. Wu, R.A. Weinberg, D.N. Louis, F.P. Li, J.G. Rheinwald, Human keratinocytes that express hTERT and also bypass a p16(INK4a)-enforced mechanism that limits life span become immortal yet retain normal growth and differentiation characteristics, Mol. Cell Biol.20 (2000) 1436-1447.

[3] S. Senapati, V.S. Gnanapragassam, N. Moniaux, N. Momi, S.K. Batra, Role of MUC4-NIDO domain in the MUC4-mediated metastasis of pancreatic cancer cells, Oncogene.31 (2012) 3346-3356.

[4] N. Moniaux, P. Chaturvedi, G.C. Varshney, J.L. Meza, J.F. Rodriguez-Sierra, J.P. Aubert, S.K. Batra, Human MUC4 mucin induces ultra-structural changes and tumorigenicity in pancreatic cancer cells, Br. J. Cancer.97 (2007) 345-357.

[5] M. Gomez-Martinez, D. Schmitz, A. Hergovich, Generation of stable human cell lines with Tetracycline-inducible (Tet-on) shRNA or cDNA expression, J. Vis. Exp.(2013) e50171.

[6] M.A. Macha, S. Rachagani, A.K. Qazi, R. Jahan, S. Gupta, A. Patel, P. Seshacharyulu, C. Lin, S. Li, S. Wang, V. Verma, S. Kishida, M. Kishida, N. Nakamura, T. Kibe, W.M. Lydiatt, R.B. Smith, A.K. Ganti, D.T. Jones, S.K.

Batra, M. Jain, Afatinib radiosensitizes head and neck squamous cell carcinoma cells by targeting cancer stem cells, Oncotarget.8 (2017) 20961-20973.

[7] I. Lakshmanan, S. Rachagani, R. Hauke, S.R. Krishn, S. Paknikar, P. Seshacharyulu, S. Karmakar, R.K. Nimmakayala, G. Kaushik, S.L. Johansson, G.B. Carey, M.P. Ponnusamy, S. Kaur, S.K. Batra, A.K. Ganti, MUC5AC interactions with integrin beta4 enhances the migration of lung cancer cells through FAK signaling, Oncogene.35 (2016) 4112-4121.

[8] J.G. Lee, J.H. Ahn, T. Jin Kim, J. Ho Lee, J.H. Choi, Mutant p53 promotes ovarian cancer cell adhesion to mesothelial cells via integrin beta4 and Akt signals, Sci Rep.5 (2015) 12642.

[9] K. Lessan, D.J. Aguiar, T. Oegema, L. Siebenson, A.P. Skubitz, CD44 and beta1 integrin mediate ovarian carcinoma cell adhesion to peritoneal mesothelial cells, Am. J Pathol.154 (1999) 1525-1537.

[10] S. Das, S. Rachagani, M.P. Torres-Gonzalez, I. Lakshmanan, P.D. Majhi, L.M. Smith, K.U. Wagner, S.K. Batra, Carboxyl-terminal domain of MUC16 imparts tumorigenic and metastatic functions through nuclear translocation of JAK2 to pancreatic cancer cells, Oncotarget.6 (2015) 5772-5787.

[11] S. Scara, P. Bottoni, R. Scatena, CA 19-9: Biochemical and Clinical Aspects, Adv. Exp. Med Biol.867 (2015) 247-260.

[12] S. Kaur, L.M. Smith, A. Patel, M. Menning, D.C. Watley, S.S. Malik, S.R. Krishn, K. Mallya, A. Aithal, A.R. Sasson, S.L. Johansson, M. Jain, S. Singh, S. Guha, C. Are, M. Raimondo, M.A. Hollingsworth, R.E. Brand, S.K. Batra,

A Combination of MUC5AC and CA19-9 Improves the Diagnosis of Pancreatic Cancer: A Multicenter Study, Am J Gastroenterol.112 (2017) 172-183.

[13] H. Pei, L. Li, B.L. Fridley, G.D. Jenkins, K.R. Kalari, W. Lingle, G. Petersen, Z. Lou, L. Wang, FKBP51 affects cancer cell response to chemotherapy by negatively regulating Akt, Cancer cell.16 (2009) 259-266.

[14] S. Rachagani, M.P. Torres, S. Kumar, D. Haridas, M. Baine, M.A. Macha, S. Kaur, M.P. Ponnusamy, P. Dey, P. Seshacharyulu, S.L. Johansson, M. Jain, K.U. Wagner, S.K. Batra, Mucin (Muc) expression during pancreatic cancer progression in spontaneous mouse model: potential implications for diagnosis and therapy, J Hematol Oncol.5 (2012) 68.

[15] R. Jahan, M.A. Macha, S. Rachagani, S. Das, L.M. Smith, S. Kaur, S.K. Batra, Axed MUC4 (MUC4/X) aggravates pancreatic malignant phenotype by activating integrin-beta1/FAK/ERK pathway, Biochim Biophys Acta Mol Basis Dis.1864 (2018) 2538-2549.

[16] Y. Nakano, S. Tanno, K. Koizumi, T. Nishikawa, K. Nakamura, M. Minoguchi, T. Izawa, Y. Mizukami, T. Okumura, Y. Kohgo, Gemcitabine chemoresistance and molecular markers associated with gemcitabine transport and metabolism in human pancreatic cancer cells, Br J Cancer.96 (2007) 457-463.

[17] A. Chakraborty, K.A. Dorsett, H.Q. Trummell, E.S. Yang, P.G. Oliver, J.A. Bonner, D.J. Buchsbaum, S.L. Bellis, ST6Gal-I sialyltransferase promotes

chemoresistance in pancreatic ductal adenocarcinoma by abrogating gemcitabine-mediated DNA damage, J Biol Chem.293 (2018) 984-994.

[18] A.P. Vaz, S. Deb, S. Rachagani, P. Dey, S. Muniyan, I. Lakshmanan, S. Karmakar, L. Smith, S. Johansson, S. Lele, M. Ouellette, M.P. Ponnusamy, S.K. Batra, Overexpression of PD2 leads to increased tumorigenicity and metastasis in pancreatic ductal adenocarcinoma, Oncotarget.7 (2016) 3317-3331.

[19] C.D. Salcido, A. Larochelle, B.J. Taylor, C.E. Dunbar, L. Varticovski, Molecular characterisation of side population cells with cancer stem cell-like characteristics in small-cell lung cancer, Br J Cancer.102 (2010) 1636-1644.

[20] I. Lakshmanan, S. Salfity, P. Seshacharyulu, S. Rachagani, A. Thomas, S. Das, P.D. Majhi, R.K. Nimmakayala, R. Vengoji, S.M. Lele, M.P. Ponnusamy, S.K. Batra, A.K. Ganti, MUC16 Regulates TSPYL5 for Lung Cancer Cell Growth and Chemoresistance by Suppressing p53, Clin Cancer Res.23 (2017) 3906-3917.

[21] I. Lakshmanan, S.K. Batra, Protocol for Apoptosis Assay by Flow Cytometry Using Annexin V Staining Method, Bio Protoc.3 (2013).

[22] R.K. Nimmakayala, P. Seshacharyulu, I. Lakshmanan, S. Rachagani, S. Chugh, S. Karmakar, S. Rauth, R. Vengoji, P. Atri, G.A. Talmon, S.M. Lele, L.M. Smith, I. Thapa, D. Bastola, M.M. Ouellette, S.K. Batra, M.P. Ponnusamy, Cigarette Smoke Induces Stem Cell Features of Pancreatic Cancer Cells via PAF1, Gastroenterology.155 (2018) 892-908 e896.

[23] S. Chaudhary, B. Madhukrishna, A.K. Adhya, S. Keshari, S.K. Mishra, Overexpression of caspase 7 is ERalpha dependent to affect proliferation and cell growth in breast cancer cells by targeting p21(Cip), Oncogenesis.5 (2016) e219.

[24] I.A. Sheikh, Stereoselectivity and the potential endocrine disrupting activity of di-(2-ethylhexyl)phthalate (DEHP) against human progesterone receptor: a computational perspective, J Appl Toxicol.36 (2016) 741-747.

[25] K. Zhu, T. Day, D. Warshaviak, C. Murrett, R. Friesner, D. Pearlman, Antibody structure determination using a combination of homology modeling, energy-based refinement, and loop prediction, Proteins.82 (2014) 1646-1655.

Chapter 3

Functional and Mechanical role of MUC4/X in Pancreatic Cancer Pathogenesis

Parts of this chapter are driven from:

Jahan R, Macha MA, Rachagani S, Smith LM, Kaur S, Batra SK. Axed MUC4 (MUC4/X) aggravates pancreatic malignant phenotype by activating integrin-β1/FAK/ERK pathway. BBA-Molecular Basis of Disease, 2018;1864(8):2538-2549.

Synopsis

Alternative splicing is evolving as an eminent player of oncogenic signaling for tumor development and progression. Mucin 4 (MUC4), a type I membrane-bound mucin, is differentially expressed in pancreatic cancer (PC) and plays a critical role in its progression and metastasis. However, the molecular implications of MUC4 splice variants during disease pathogenesis remain obscure. The present study delineates the pathological and molecular significance of a unique splice variant of MUC4, MUC4/X, which lacks the largest and polymorphic exon 2, along with exon 3. Exon 2 encodes for the highly glycosylated tandem repeat (TR) domain of MUC4 and its absence creates MUC4/X, which is devoid of TR. Expression analysis from PC clinical samples revealed significant upregulation of MUC4/X in PC tissues with most differential expression in poorly differentiated tumors. In vitro studies suggest that overexpression of MUC4/X in wild-type-MUC4 (WT-MUC4) null PC cell lines markedly enhanced PC cell proliferation, invasion, and adhesion to extracellular matrix (ECM) proteins. Furthermore, MUC4/X overexpression leads to an increase in the tumorigenic potential of PC cells in orthotopic transplantation studies. In line with these findings, doxycycline-induced expression of MUC4/X in an endogenous WT-MUC4 expressing PC cell line (Capan-1) also displayed enhanced cell proliferation, invasion, and adhesion to ECM, compared to WT-MUC4 alone, emphasizing its direct involvement in the aggressive behavior of cancer cells. Investigation into the molecular mechanism suggested that MUC4/X facilitated PC tumorigenesis via integrin-β1/FAK/ERK signaling pathway. Overall, these findings

revealed the novel role of MUC4/X in promoting and sustaining the oncogenic features of PC.

Introduction

Pancreatic cancer (PC) is the third leading cause of cancer-related deaths in the USA and is predicted to become second by 2030 [1, 2]. Due to high complexity and degree of heterogeneity, the molecular mechanisms for progression and early metastasis remain obscure. Alternative splicing is one of the mechanisms that contributes to the complexity and effectiveness of disease progression; therefore, it stands to reason that cancer cells adopt alternative splicing to sustain themselves within a hostile environment. Recently, alternative splicing has gained immense attention in cancer research due to its strong association with tumor development as well as an attractive anticancer therapeutic target(s) [3]. For example, splice variant of the CD44 adhesion molecule has been implicated in the metastatic spread of various human tumor cells and has been correlated with poor prognosis [4]. Another example of a well studied oncogenic splice variant in PC, is the paired related homeodomain transcription factor (PrrX1). Two alternative isoforms of PrrX1 (PrrX1a, and PrrX1b) have shown distinct but complementary roles in PC oncogenesis [5]. Prrx1b regulates epithelial to mesenchymal transition (EMT) which promotes invasion, while Prrx1a regulates mesenchymal to epithelial transition (MET) by tumor redifferentiation which enhances metastatic colonization [5]. All these previous studies suggest that splice variants have distinct and pronounced functions at different stages of tumor progression, and therefore

further exploration is merited for delineating their mechanistic and therapeutic significance in a highly lethal malignancy like PC.

The type 1 transmembrane mucin MUC4, is one of the most differentially overexpressed genes in PC, with undetectable expression in normal pancreas and *de novo* expression in early precursor lesions [6]. With this differential expression in PC, MUC4 has been implicated as a primary oncogenic player with prominent roles in neoplastic transformation, tumor progression, metastasis, and chemoresistance [7-11]. It is comprised of 26 exons organized into unique domains including a variable tandem-repeat (TR) domain, nidogen-like (NIDO) domain, adhesion-associated domain in MUC4 and other proteins (AMOP), three EGF-like domains (EGF), transmembrane (TM) domain and a short cytoplasmic tail (CT) domain (**Figure 1a-b**) [12, 13]. We and others have identified 24 distinct variants of MUC4, however the functional implications of these splice variants in PC pathogenesis is not fully elucidated [14]. Specifically, deletion of exons 2 and 3 results in the formation of MUC4/X, and deletion of exon 2 alone results in MUC4/Y [14]. Exon 2 codes for the largest domain of MUC4 and characteristic mucin structural signature defined by a TR region made of 145 - 500 repeats of 16 amino acids that are heavily *O*-glycosylated on serine and threonine residues [13].

Few mucin splice variants have been studied to assess their various pathological implications. Among them, the role of MUC1 splice variants have been studied in cancer and inflammatory diseases [15]. Previous research has shown that MUC1/Y is overexpressed in breast cancer patient tissues compared to normal adjacent tissues [16]. Additionally, overexpression of MUC4/Y has been

associated with PC tumorigenesis by activating the JNK and AKT signaling pathways [17]. Also in PC patients, MUC4/Y overexpression is positively correlated with tumor invasion and metastasis [18].

Various transmembrane mucins, including MUC1 and MUC4, have been shown to impart steric hindrance to cell-ECM interaction due to their large glycosylated TR domain [19-21]. Further, it has been demonstrated that a higher number of tandem repeats in SMC/Muc4, the rat homolog of human MUC4, contributes to decreased adhesion of cancerous cells to the ECM protein fibronectin, suggesting that not only the presence, but extent of the TR region, is also responsible for adhesive properties [22]. Earlier studies from our group have also demonstrated the involvement of WT-MUC4 in impeding PC cells interaction with ECM by interfering with integrin-mediated cell adhesion [19]. These studies suggest a critical role of TR domain of WT-MUC4 for cell - ECM adhesion in PC pathogenesis.

The present study explores the molecular and pathological significance of MUC4/X in the pathogenesis of PC. Overall, our studies revealed significant overexpression of MUC4/X in PC tissues. Our studies also showed that overexpression of MUC4/X in PC cells augmented cell proliferation, invasion and metastasis in both *in vitro* and *in vivo* models. These effects were mediated by boosting the integrin-β1/FAK/ERK signaling pathway.

Results

Aberrant expression of MUC4/X in PC tumor tissues

The expression of WT- MUC4 and MUC4/X was investigated in normal pancreatic tissue adjacent to pancreatic tumor (NAT) (n=8) and PC patient tissues [well-differentiated (WD, n=7), moderately-differentiated (MD, n=9), and poorly-differentiated (PD, n=10)]. For MUC4/X expression assessment, the qPCR forward primer was designed to bind the junctional sequence of exons 1 and 4, and the reverse primer to bind the *3′* end of exon 4 (**Supplementary Table 1**). We observed significant upregulation of MUC4/X expression in PC samples compared to NAT samples ($p<0.005$) (**Figure 1c**). Interestingly, we also observed significantly high MUC4/X expression in PD tissues, compared to the NAT and WD tissues (NAT vs. PD, $p<0.005$; WD vs. PD, $p<0.001$) (**Figure 1d**). We observed that all NAT tissues showed lower level of MUC4/X with a consistent elevated level WT-MUC4 expression as seen in **Figure 1e**. While all the PC tissues displayed WT-MUC4 expression, 20% of WD, 40% of MD, and 80% of PD tumors showed higher expression of MUC4/X suggesting the WT-MUC4 and splice variant may complement each other in PC progression. Overall, deregulated expression of MUC4/X was observed in PC patient tissues (**Figure 1e and Supplementary Figure S1a**). Considering the significant overexpression of MUC4/X in PC, we next focused on delineating its pathobiological significance in PC tumorigenesis.

Overexpression of MUC4/X enhanced the cell proliferation and clonogenicity of PC cells in vitro

We examined the expression of MUC4/X in immortalized pancreatic cells in addition to the panel of PC cell lines with differential differentiation status. While no expression was observed in immortalized normal pancreatic cell line HPDE, varied expression of MUC4/X were observed in PC cell lines with the lowest expression in MIAPaCa and SW-1990 and highest expression in BxPC3 cells (**Supplementary Figure S1b**). Among these, MIAPaCa, PANC-1 and AsPC-1 cell lines do not express WT-MUC4 endogenously while Capan-1, BxPC-3, HPAC, and HPAF-II PC cell lines express moderate to high levels of WT-MUC4 (**Supplementary Figure S1b**) [23]. To investigate the functional attributes of MUC4/X, stable overexpression of MUC4/X was carried out in MIAPaCa and AsPC-1 cells that do not express WT-MUC4 at the protein level, and have very low level of expression at the mRNA level (**Supplementary Figure S1b**). It is worth mentioning, due to lack of antibodies for detecting MUC4/X protein, we generated a dual epitope-tagged MUC4/X mammalian expression construct, with N-terminal FLAG-tag and a C-terminal HA-tag (**Supplementary Figure S2a**). The empty vector transfected cell lines (MIAPaCa-EV/ AsPC-1-EV) were used as controls. Expression of MUC4/X was verified by immunoblot and immunofluorescence analyses, using anti-FLAG and anti-HA antibodies (**Figures 2a and Supplementary Figure S2b**). Additionally, our qPCR analysis also revealed almost no expression of WT-MUC4 in MIAPaCa-MUC4/X and AsPC-1-MUC4/X cells when compared to CD18/HPAF PC cell line (CD18/HPAF is a well-

differentiated, highly aggressive and metastatic PC cell line) which expresses high levels of endogenous WT-MUC4 (**Supplementary Figure S2c**) [24]. MTT assay revealed significant enhancement of cellular proliferation in MIAPaCa-MUC4/X and AsPC-1-MUC4/X cells compared to control cells at 3rd and 4th day time points (**Figure 2b**). Consistent with MTT assay results, EdU incorporation assay also showed a significantly higher percentage of EdU positive MUC4/X-OE cells compared to control cells ($p<0.05$) (**Figure 2c**). As shown in **Figure 2d**, we also observed that overexpression of MUC4/X significantly increased colony forming ability of MIAPaCa-MUC4/X and AsPC-1-MUC4/X cells ($p<0.005$) as compared to control cells. Together, these findings support the notion that MUC4/X promotes proliferation of PC cells.

Overexpression of MUC4/X promotes cell invasion and migration

WT-MUC4 has been associated with increased invasion and motility in various malignancies including PC [7, 25]. To analyze the effect of MUC4/X overexpression on cell invasion and migration, we performed a Boyden chamber assay as well as a wound healing assay. As seen in **Figure 3a**, the number of invasive cells were significantly higher upon MUC4/X overexpression in MIAPaCa and AsPC-1 cells than with the control cells ($p<0.005$, $p<0.05$). The wound healing assay also revealed significantly enhanced wound closure by MUC4/X-OE PC cells compared to control cells ($p<0.005$) (**Supplementary Figure S3a-b**). Overall, overexpression of MUC4/X notably the enhanced invasive and migratory capability of PC cells.

Given that the TR domain is a heavily O-glycosylated domain of MUC4, its absence may influence the ability of MUC4 to interact with ECM [19]. Thus, we next examined how MUC4/X potentiates PC cell adherence to ECM proteins such as fibronectin, vitronectin, laminin, collagen IV, and collagen I. Both MUC4/X-OE and control PC cells were plated onto various ECM proteins coated 96-well plate. A significant increase in cell adhesion ($p<0.001$) was observed on laminin, vitronectin, and fibronectin-coated wells by MUC4/X-OE cells (MIAPaCa and AsPC-1) when compared to control cells **(Figure 3b-c)**.

To assess the molecular mechanisms for differential adhesion to ECM proteins and enhanced migration/invasion by MUC4/X-OE cells, we examined the impact of MUC4/X overexpression on integrin-β1. Integrin-β1 was chosen as it is a known ligand for fibronectin, vitronectin, collagen IV and collagen I [26]. Our results revealed increased expression of integrin-β1 in MUC4/X-OE cells when compared to controls in conjunction with FAK, a non-receptor protein-tyrosine kinase which is involved in cell proliferation, invasion, adhesion, and metastasis [27]. Moreover, integrin-β1 plays a role in cell survival primarily by phosphorylation of FAK [28]. To delineate the molecular pathways associated with increased migration and invasion of PC cell upon MUC4/X overexpression, we evaluated its effect on the FAK/ERK signaling pathway. We observed increased expression of integrin-β1, phosphorylated FAK and ERK in MIAPaCa-MUC4/X and AsPC-1-MUC4/X cells as compared to control cells (**Figure** 3d), signifying that MUC4/X might facilitate the invasive and metastatic potential of PC cells through the integrin-β1/FAK/ERK pathway. Additionally, since our EdU cell proliferation assay showed higher EdU

incorporation which is generally incorporated in DNA synthesis S phase, we also analyzed the expression of cyclin A2, a well-established cyclin that promotes entry into S-phase and also considered as cell proliferation marker [29, 30]. Corroborating with our *in vitro* data from the EdU cell proliferation assay, overexpression of MUC4/X in both PC cell lines elevated the expression of cyclin A2 (**Figure 3d**). Overall, our results demonstrate that MUC4/X overexpression modulates oncogenic molecules and its downstream signaling to exacerbate the PC phenotype.

To better understand the underlying mechanism of MUC4/X tumorogenic potential, microarray analysis was performed in MIAPaCa-MUC4/X and control MIAPaCA-EV cells. Our results revealed ≥2-fold upregulation of CD44, Sortilin 2, Coronin 2B, integrin-α6, integrin-α3, and Interferon α4 (IFN-α4) in MIAPaCa-MUC4/X cells compared to vector control cells (**Supplementary Figure S4a**). A full list of up- and down-regulated (≥ 2-fold) genes is provided in **Supplementary Table 2**. Indeed, validation of up-regulated genes from the microarray data using qPCR verified more than 2-fold increase of CD44, Sortilin 2, Coronin 2B, integrin-α6, integrin-α3, IFN-α4, and MUC4 **(Supplementary Figure S4a-S4b).** MUC4 showed a 44-fold up-regulation in the microarray analysis, resulting from the overexpression of MUC4/X as Human Gene 2.0 ST does not contain splicing probe, and MIAPaCa has a very low basal expression of WT-MUC4 **(Supplementary Figure S4b)**. Among the upregulated genes, CD44 and various integrins have been shown to be involved in promoting the tumorigenic potential of PC [31, 32]. Moreover, CD44 has been associated with increased motility,

adhesion, and chemoresistance in other malignancies [32]. Ingenuity Pathway Analysis (IPA) showed enrichment of ephrin receptor signaling and the axonal guidance pathway in the MIAPaCA-MUC4/X cells **(Supplementary Figure S4c).** These have been implicated as critical pathways in PC tumorigenesis in previous studies [33, 34].

MUC4/X overexpression fosters tumorigenicity in vivo

In vitro studies of MUC4/X-OE PC cells showed higher proliferative and invasive potential when compared to control cells. To ascertain whether the overexpression of MUC4/X can promote PC tumor growth and metastasis, we orthotopically implanted (2.5×10^5 cells in 50µl of PBS) MIAPaCa-MUC4/X cells or control MIAPaCa-EV cells into the pancreas of athymic nude mice (MIAPaCa-EV, n=7, and MIAPaCa-MUC4/X, n=8). At 50 days post-implantation, we euthanized the mice and observed that tumors formed by MIAPaCa-MUC4/X PC cells were significantly larger (average weight of 931±291.69 mg) in comparison to the control cells (566±178.38 mg) indicating that MUC4/X overexpression is involved in exacerbating PC tumorigenicity ($p<0.05$) **(Figure 4a-b)**. Our IHC analysis of tumor tissues using anti-FLAG antibody confirmed maintenance of MUC4/X overexpression during tumor growth and metastasis **(Figure 4c)**. We also observed higher Ki-67, and integrin-β1 expression in MUC4/X-OE xenograft tumors **(Figure 4c)**. Furthermore, all the mice (8/8) implanted with MUC4/X-OE PC cells developed metastasis in distant organs, including diaphragm, spleen, and intestine **(Supplementary Figure S5)**. Interestingly, significant metastasis to

peritoneal cavity (6/8), liver (5/8), kidney (5/8), mesenteric cavity (5/8) and stomach (2/8) was observed compared to control mice (peritoneal 1/7, liver 1/8, kidney 1/7, mesenteric cavity 0/7, and stomach 1/8) **(Figure 4d and Supplementary Figure S5)**. Metastasis to the ovary (2/8), cecum (1/8), colon (1/8), and bladder (1/8) was exclusively observed in MUC4/X-OE cell-bearing mice **(Figure 4d and Supplementary Figure S5)**. Our results suggest the involvement of MUC4/X in tumor growth and distant metastasis.

MUC4/X increases the adhesive capability to peritoneal *LP9/TERT*-1 cells.

Peritoneal metastasis in PC reflects a devastating form of cancer progression that is intensely complex [35]. Probable mechanisms behind this fatal metastasis is augmented expression of adhesion molecules and integrin's, specifically integrin-β1, which play a significant role to help the cancer cells to attach to distant mesothelium [35]. As orthotopic transplantation of MIAPaCa-MUC4/X cells resulted in significant metastasis to the peritoneal cavity and increased expression of integrin-β1, we investigated whether MUC4/X-OE increases the adhesion capability of PC cells to immortalized peritoneal *LP9/TERT*-1 cells. To mimic the *in vivo* conditions where tumor cells would face the surface of the peritoneum, we coated 96 well plate with immortalized *LP9/TERT*-1 peritoneal cells follow by subsequent incubation with tumor cells (control and MUC4/X-OE) on top of the mesothelial monolayer for 5 hrs. Interestingly, MIAPaCa-MUC4/X (30%, $p<0.05$) and AsPC-1-MUC4/X (60%, $p<0.005$) cells significantly adhered to the monolayer

of *LP9/TERT*-1 as compared to control cells **(Figure 4e).** These results suggested that MUC4/X plays a significant role in metastasizing PC cells to the peritoneum.

Overexpression of MUC4/X in the presence of endogenous MUC4, promotes cell proliferation, adhesion, and invasion.

Both PC patient tissue samples, as well as PC cell lines, showed concomitant expression of MUC4 and its splice variant MUC4/X (**Figure 1e, Supplementary Figure S1a, and S1b**). Earlier, MUC4 was shown to enhance proliferation, migration, and invasion of PC cells [8, 19, 25, 36]. We were interested if the concomitant expression of WT-MUC4 and MUC4/X can synergistically enhance the tumorigenic properties of cancer cells, or if their expression is redundant for cells.

To explore the explicit function of MUC4/X in PC, a tet-on inducible system was developed (**Figure 5a**). Conditional activation of MUC4/X expression by doxycycline allowed us to decipher a gain of function of MUC4/X, in the presence of WT-MUC4**.** Immunofluorescence and immunoblotting showed robust MUC4/X protein expression after 48 hrs of doxycycline treatment (**Figure 5b**). Induction of MUC4/X didn't change the expression of WT-MUC4 (**Figure 5b**). Thymidine analog EdU incorporation assay revealed that induction of MUC4/X in Capan-1 cells resulted in higher cell proliferation compared to doxycycline-negative Capan-1 cells (**Figure 5c,** $p<0.05$). Induction of MUC4/X also resulted in a significant increase in cell adhesion to ECM proteins such as laminin, collagen, and fibronectin (**Figure 5d**, $p<0.005$), as well as increased cell invasion ($p<0.05$) as observed via Boyden chamber assay (**Figure 5e**). Overexpression of MUC4/X did

not change the expression of WT-MUC4, suggesting that MUC4/X overexpression mediates the phenomena of increased cell proliferation, invasion, and adhesion to ECM. Given that we observed MUC4/X overexpression in MIAPaCa, and AsPC-1 cell lines upregulated integrin-β1 and its downstream signaling molecules, we next sought to analyze their differential expression in tet-on inducible Capan-1 cell lines, in the presence of abundant expression of WT-MUC4. In concordance with the above-mentioned findings, doxycycline-induced MUC4/X resulted in higher integrin-β1 expression and increased pFAK, pERK and cyclin A2 expression in the presence of WT-MUC4 (**Figure 5f**). This result suggested that conditional induction of MUC4/X in a WT-MUC4 expressing cell line, phenocopied the characteristics of MUC4/X in a WT-MUC4 negative cell line, emphasizing its tumorigenic role independent of WT-MUC4 expression.

Discussion

Transmembrane mucin MUC4 has previously been implicated in oncogenesis by initiating signaling *via* interaction and stabilization of HER2 [25]. Importantly, in the context of adhesion, the highly glycosylated TR domain of MUC4 has been hypothesized to interfere with tumor cell interaction with extracellular matrix (ECM) proteins, in part by blocking accessibility of integrins to ECM ligands sterically [19]. It has also been shown that a nonglycosylated TR of MUC1 facilitates attachment of tumor cells to the ECM to establish metastatic foci [37]. In addition, MUC1 without a TR domain displayed an aggressive PC phenotype *in vitro* and *in vivo* [38].-

Apart from the TR domain, other functional domains of MUC4 has also been associated with invasion and metastasis of PC. Earlier studies from our group have suggested that the NIDO domain of MUC4, a domain similar to the G1 domain of a basement membrane nidogen (entactin), plays a critical role in invasion and metastasis. Deletion of NIDO domain from MUC4 decreased PC cells invasion and metastasis [39]. Moreover, Tang *et al*. showed deletion of the adhesion-associated domain of MUC4 (AMOP domain) reduced invasion and metastasis of PC cells [40].

Overall, transmembrane mucins, with their various unique domains, are implicated in invasion and metastasis of tumor cells. Tumor cell invasion and metastasis is a complex process involving a delicate balance between adhesion and detachment of cancer cells with the ECM. This is initiated by dissemination of cancer cells from primary location followed by intravasation and subsequent extravasation through endothelial cells or lymphatics, to colonize distant organs [41]. Considering the significant role of MUC4 and its domains in intensifying PC malignant phenotype, the present study explored the biological significance of MUC4/X, a splice variant that sans the heavily O-glycosylated TR domain of MUC4, on PC cells capability to adhere, invade, and metastasize.

MUC4/X is considered as a unique isoform of WT-MUC4, not only for lacking the TR domain, but also because of its adhesion-related functional attributes which are supposedly more accessible to the ECM due to the absence of steric hindrance by tandem repeat O-glycosylated domain. To comprehend the contribution of the TR domain deficient MUC4/X, we explored the interaction of MUC4/X

overexpressing cells with various ECM components. ECM is a dense visco-elastic latticework composed of collagen, laminins, fibronectin, vitronectin, and different linker proteins (i.e. nidogen and entactin) which connect collagens with other proteins [42]. Cells can perceive direct or indirect signaling from the ECM by interacting with different ECM proteins [43, 44]. Moreover, integrins can link the ECM to the intracellular actin cytoskeleton to commence intracellular signaling events by elevating gene expression involved in cell proliferation, survival, and migration [44]. In this regard, it has previously been shown that overexpression of MUC1 and MUC4 inhibit integrin-ECM interaction [19]. Of note, earlier studies utilizing a microarray of PC cell lines with a MUC1-TR-deleted construct, revealed upregulation of integrin-β5 [38]. In corroboration with this, our microarray data showed the upregulation of integrins (subunit α3 and α6) and CD44 in MUC4/X-OE PC cells.

Interestingly, knockdown of the integrin-β1 subunit reduced cell adhesion to ECM proteins and decreased PC tumor growth and metastasis [45]. Hence, we investigated the expression of integrin-β1, a binding partner of integrin α3 and α6, in MUC4/X overexpressing cells as well as in an inducible system. Surprisingly both showed upregulation of the β1 subunit of integrin. Integrin-β1 has been observed to play an essential role in cell adhesion, proliferation, and metastasis [26, 45-47]. Notably, collagen (types I, III, and IV), fibronectin and laminin are known ligands for integrin-β1 [26]. The interactions between integrin-β1 and laminin are crucial for cell survival through FAK signaling [28]. Moreover, adhesion of cancer cells to ECM incites the intracellular signaling cascade for cell

proliferation and invasion via integrin-β1/FAK/ERK signaling [48]. In the present study, we observed a similar phenomenon for MUC4/X, which promoted pancreatic malignancy by activating the integrin-β1/FAK/ERK pathway.

Peritoneal metastasis is one of the most frequent and deadly forms of PC metastasis [49]. Pancreatic tumors with peritoneal metastasis are highly resistant to chemotherapies with very poor patient survival (≤12 months). Moreover, the presence of peritoneal carcinomatosis has been associated with the development of intestinal obstruction, massive ascites, and malnutrition and is unfortunately associated with ~22% of patients morbidities [50]. In our orthotopic transplantation studies for MUC4/X, we observed significant metastasis to the liver and peritoneum. Though the mechanism for peritoneal carcinomatosis is still obscure, integrin-ECM interaction between tumor and mesothelial cells is considered as an early event in this process [51]. Evidenced by integrin-β1 in ovarian carcinoma cells being associated with adhesion of ovarian carcinoma cells to mesothelial cells [52]. Our cell adhesion study revealed that MUC4/X overexpression resulted in significant adhesion to fibronectin, vitronectin, and laminin. It was shown earlier that *LP9/TERT*-1 mesothelial cell line expresses collagen types I, III, and IV, fibronectin and laminin [52]. Henceforth, we evaluated adhesion ability of MUC4/X-OE cell to *LP9/TERT*-1 mesothelial cells. Results from peritoneal cell adhesion assay suggested that MUC4/X-OE cells have higher adherence to immortalized *LP9/TERT*-1 peritoneal cells as compared to control. We hypothesize that upregulation of integrin-β1 by MUC4/X, and the higher adhesion capacity of MUC4/X-OE cells to the ECM proteins overexpressed in peritoneal cells, are the

causative factors for enhanced adhesion and increased metastasis to peritoneum by MUC4/X-OE cells [52].

Considering the differential affinity of MUC4/X toward ECM proteins and WT-MUC4 as mentioned in earlier published reports, we speculate that they may either synergistically or reciprocally function at various stages of tumor development by potentiating invasion, adhesion and metastasis phenomena to drive PC malignancy. MUC4/X overexpression also resulted in increased cell proliferation in part by upregulating cyclin A2, a mammalian A-type cyclin which is responsible for initiation and progression of DNA replication and G1-S phase transition and is considered a proliferation marker [29, 30, 53]. Similar to our study, cyclin A2 overexpression has also been observed in multiple malignancies including breast, prostate, colorectal and PC. It has been shown to contribute to the invasion, metastasis and tumorigenesis of these various cancers [54, 55].

We observed concomitant expression of WT-MUC4 and MUC4/X in PC clinical samples. Interestingly, differential expression of MUC4/X was observed in poorly differentiated PC tumors. However, our studies were limited by lack of MUC4/X specific antibodies. In the future, development of such antibodies could lead to the evolution of improved biomarkers for PC detection. Further, to fully resolve the functional significance of MUC4/X, future orthotopic studies with an inducible system of MUC4/X in the background of WT-MUC4 would be carried out. A transgenic mouse model expressing human MUC4/X in the context of a pancreatic tumor will provide further insight into the functional relevance of this unique splice form.

Together, the results of this study propose the novel role of MUC4/X in PC cell proliferation, invasion, adhesion, and metastasis. The role of MUC4/X was investigated using *in vitro* and *in vivo* models and found to be mediated *via* upregulation of the integrin-β1/FAK/ERK pathway (**Figure 6**). We have investigated the expression of MUC4/X in a cohort of patient tissues as well as cell lines, which suggest its aberrant upregulation in PC. The induced expression of MUC4/X in the presence of WT-MUC4 resulted in a high propensity of tumor cells to proliferate in addition to an increased capability of invasion, suggesting a non-redundant role of this splice variant. To our knowledge, this is the first report on the functional and mechanical role of MUC4/X, alone and with WT-MUC4, suggesting its role for conferring tumorigenic potential to PC. We have thus provided experimental evidence that WT-MUC4 and MUC4/X may be mutually beneficial for PC cells to develop and sustain an oncogenic phenotype at different stages of this lethal disease. Studies to determine the oncogenic role of a splice variant (in this case, MUC4/X) will help us to understand the complex molecular mechanism of PC and design much needed personalized therapeutic interventions

Figure 1

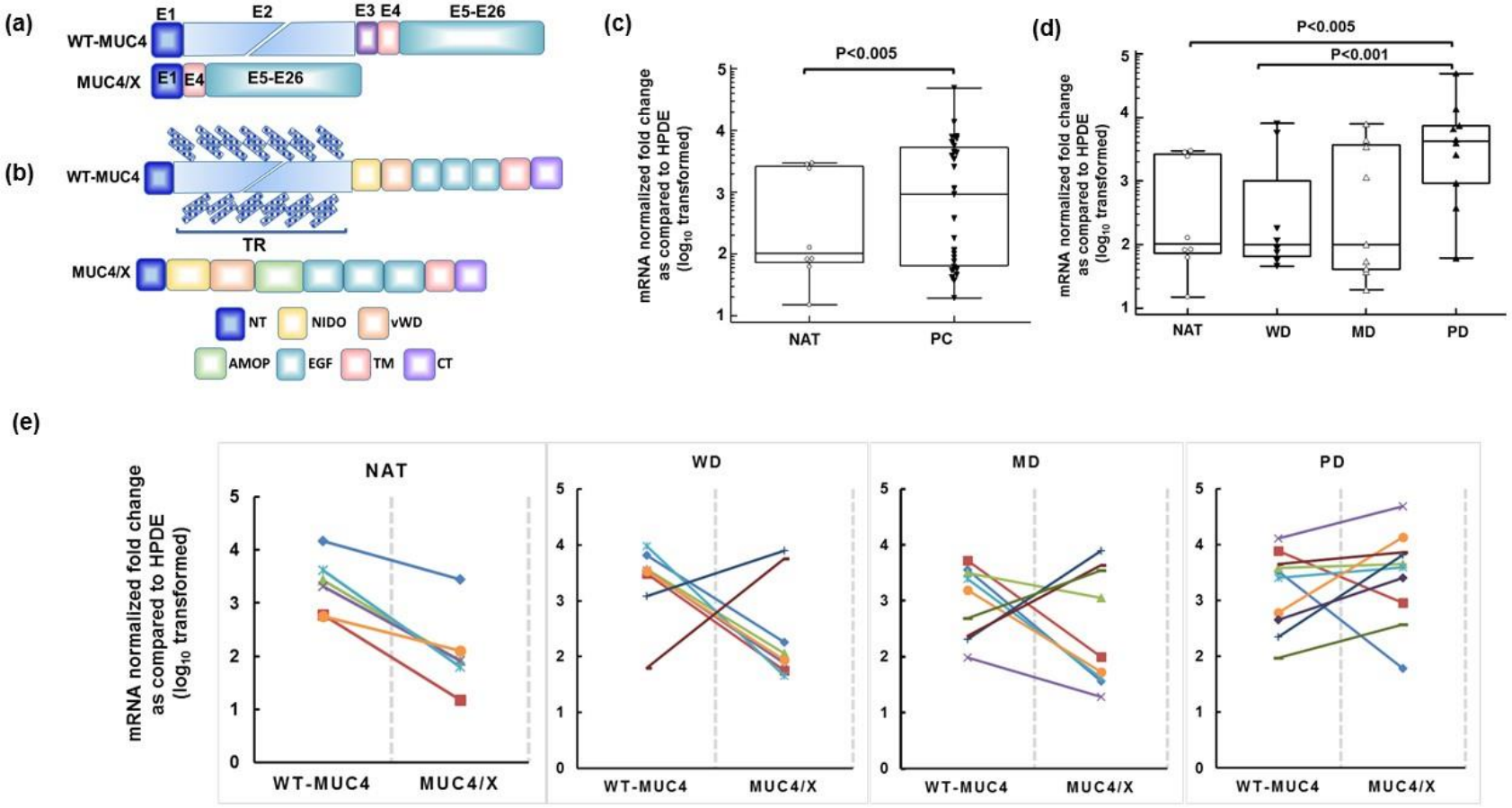

(a)
E1
E2
E3 E4
E5-E26
WT-MUC4
MUC4/X
E1
E4
E5-E26
(b)
WT-MUC4
TR
MUC4/X
NT
NIDO
vWD
AMOP
EGF
TM
CT
(c)
P<0.005
mRNA normalized fold change as compared to HPDE (log10 transformed)
NAT
PC
(d)
P<0.005
P<0.001
mRNA normalized fold change as compared to HPDE (log10 transformed)
NAT
WD
MD
PD
(e)
mRNA normalized fold change as compared to HPDE (log10 transformed)
NAT
WD
MD
PD
WT-MUC4
MUC4/X

Figure Legends

Figure 1: Expression of MUC4/X and wild-type (WT)-MUC4 mRNA in pancreatic cancer (PC) clinical samples and normal adjacent pancreatic tissues (NAT). (a) Schematic diagram of WT-MUC4 and MUC4/X mRNA structure. The WT-MUC4 is comprised of 26 exons while splice variant MUC4/X is devoid of exons 2 and 3. Exons are represented by box. **(b)** Schematic diagram of the domain structure of WT-MUC4 and MUC4/X. WT-MUC4 is a transmembrane mucin characterized by the presence of multiple domains including N-terminal (NT), Tandem Repeat (TR), Nidogen like (NIDO), Adhesion-associated domain in mucin MUC4 and other protein (AMOP), von Willebrand D (vWD), Epidermal Growth Factor (EGF), transmembrane (TM), and cytoplasmic (CT) domain. The unique splice variant, MUC4/X, is characterized by the presence of all domains of WT-MUC4 except the heavily glycosylated and polymorphic TR domain. Diagonally placed blue checker boxes represent O-glycosylation within the TR domain present in WT-MUC4. **(c-d)** Box and whisker plots are representing MUC4/X mRNA fold change (log_{10} transformed) in PC tissues and normal tissues adjacent to tumor (NAT) by qPCR and ΔΔCt method. **(c)** Significantly, elevated expression of MUC4/X was observed in pancreatic tumor tissues in comparison to NAT tissues ($p<0.005$). (**d**) The qPCR analysis of MUC4/X mRNA expression levels in pancreatic NAT, well-differentiated (WD), moderately-differentiated (MD), and poorly-differentiated (PD) cases. Elevated expression of MUC4/X was observed in PD pancreatic tumors in comparison to NAT ($p<0.005$) as well as WD tumors ($p<0.001$). Two-tailed Student's t-test was used to determine the statistical

difference between two groups. The circle represents individual mRNA expression in NAT tissues whereas triangle represents PC cases. The interquartile range (IQR) for MUC4/X expression is presented by box and whisker plot (horizontal line represents the 25th percentile, median and 75th percentile and whisker represents 5th and 95th percentile). **(e)** Line diagram is representing mRNA fold change (log10 transformed) of both WT-MUC4-and MUC4/X in PC tissues and normal tissues adjacent to tumor by qPCR and ΔΔCt methods in NAT, WD, MD, and PD groups. Each line represents individual patient data for WT-MUC4 and MUC4/X where the left and right endpoint represent fold change of WT-MUC4 and MUC4/X expression respectively. β-actin was used as an internal control to normalize the expression of respective gene, and normal human pancreatic ductal epithelial (HPDE) cell line mRNA expression levels were used as calibrant control for determining fold change across patient groups. Y-Axis represent value as fold change (log 10 relative quantification) relative to HPDE.

Figure 2

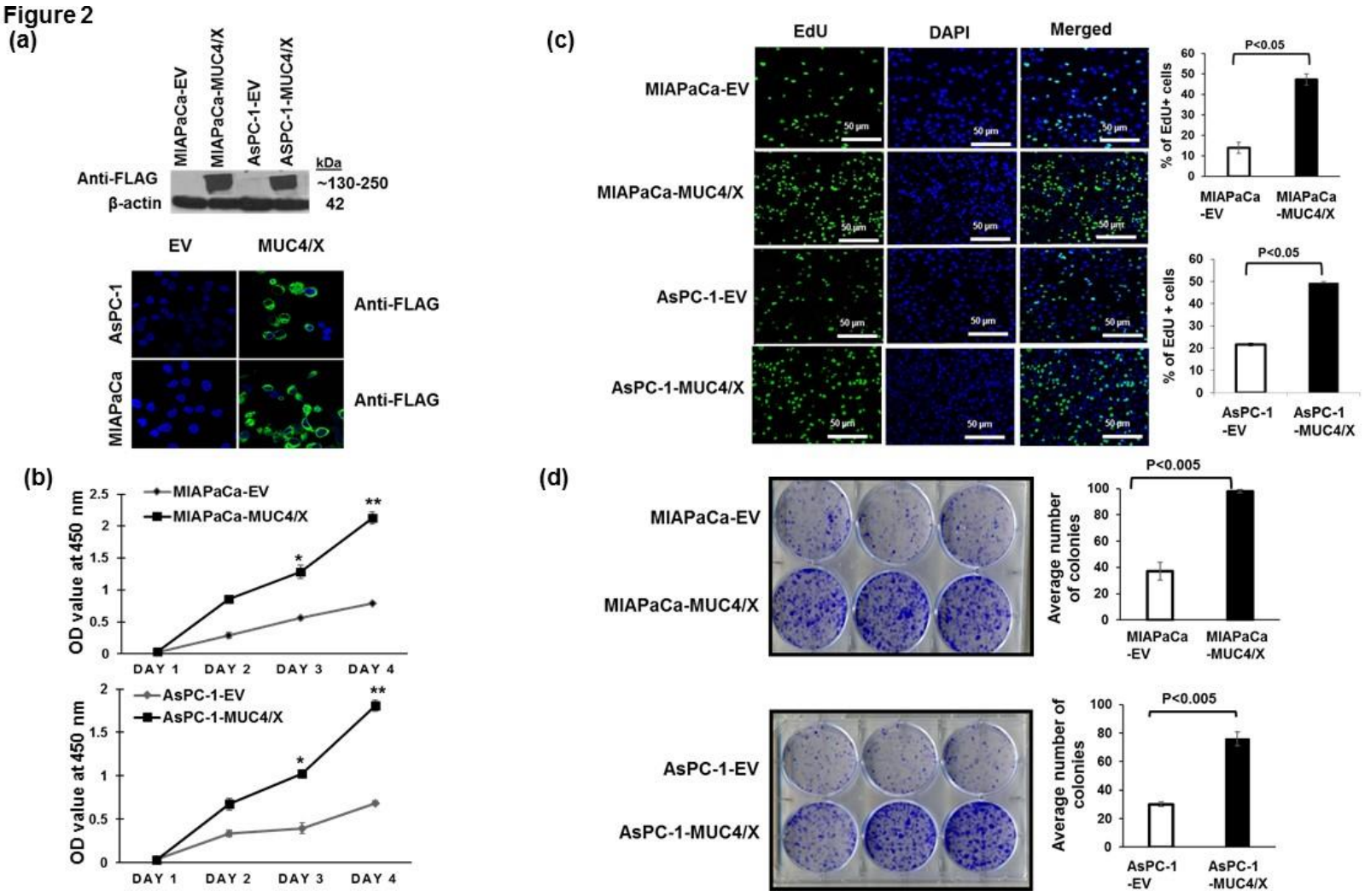

(a)
MIAPaCa-EV
MIAPaCa-MUC4/X
AsPC-1-EV
ASPC-1-MUC4/X
kDa
Anti-FLAG
~130-250
β-actin
42
EV
MUC4/X
AsPC-1
MIAPaCa
Anti-FLAG
Anti-FLAG
(b)
MIAPaCa-EV
MIAPaCa-MUC4/X
OD value at 450 nm
DAY 1
DAY 2
DAY 3
DAY 4
AsPC-1-EV
AsPC-1-MUC4/X
(c)
EdU
DAPI
Merged
MIAPaCa-EV
MIAPaCa-MUC4/X
AsPC-1-EV
AsPC-1-MUC4/X
50 μm
P<0.05
% of EdU+ cells
MIAPaCa -EV
MIAPaCa -MUC4/X
% of EdU + cells
AsPC-1 -EV
AsPC-1 -MUC4/X
(d)
MIAPaCa-EV
MIAPaCa-MUC4/X
P<0.005
Average number of colonies
AsPC-1-EV
AsPC-1-MUC4/X

Figure 2: Functional implications of overexpression of the MUC4/X on PC cell proliferation and colony formation. **(a)** MUC4/X was cloned into the N-terminus FLAG-tagged, and the C-terminus HA-tagged p3XFLAG-CMV™-9 vector followed by stable transfection into WT-MUC4 non-expression PC cell lines MIAPaCa and AsPC-1. (**Upper panel**) Immunoblot analyses using anti-FLAG antibody was performed to assess the expression of MUC4/X. MUC4/X overexpression was observed in stably transfected MIAPaCa-MUC4/X and AsPC-1-MUC4/X cells whereas no expression was detected in empty vector (EV) transfected cells. (**Lower Panel**) Immunofluorescence analyses for MUC4/X using anti-FLAG antibody showed expression of MUC4/X in MIAPaCa-MUC4/X, and AsPC-1-MUC4/X PC cells while no expression was observed in vector alone transfected cells. **(b)** MTT assay was performed to assess the impact of MUC4/X overexpression on cellular proliferation. Significant higher cell proliferation in MIAPaCa-MUC4/X and AsPC-1-MUC4/X cell lines were observed as compared to vector transfected control cells by 3^{rd} and 4^{th} days (MIAPaCa-EV/AsPC-1-EV) (**$p<0.01$, *$p<0.05$). Line diagram represents the OD value (mean± SD, n=3). **(c)** EdU cell proliferation assay was performed to assess the impact of MUC4/X overexpression on PC cell proliferation. MUC4/X-OE cells had a higher number of EdU positive proliferating (green fluorescent) cells as compared to control vector transfected cells ($p< 0.05$) after 24 hrs of incubation with EdU. Nuclei staining with DAPI represents total number of cells. Accompanying bar diagram demonstrates the quantitative measurement for percentages of EdU positive green fluorescent

cells (mean ± SD, n=3). **(d)** Colony forming assay suggested MUC4/X-transfected MIAPaCa and AsPC-1 cells formed significantly greater number of colonies than the control cells ($p<0.005$). The corresponding bar diagram represents quantitative analysis of an average number of colonies (mean ± SD, n=3) observed per well of six-well plates. The two-tailed, unpaired t-test analyses was used to determine significance between two groups. These results suggest MUC4/X overexpression promotes cell proliferation and colony formation.

Figure 3

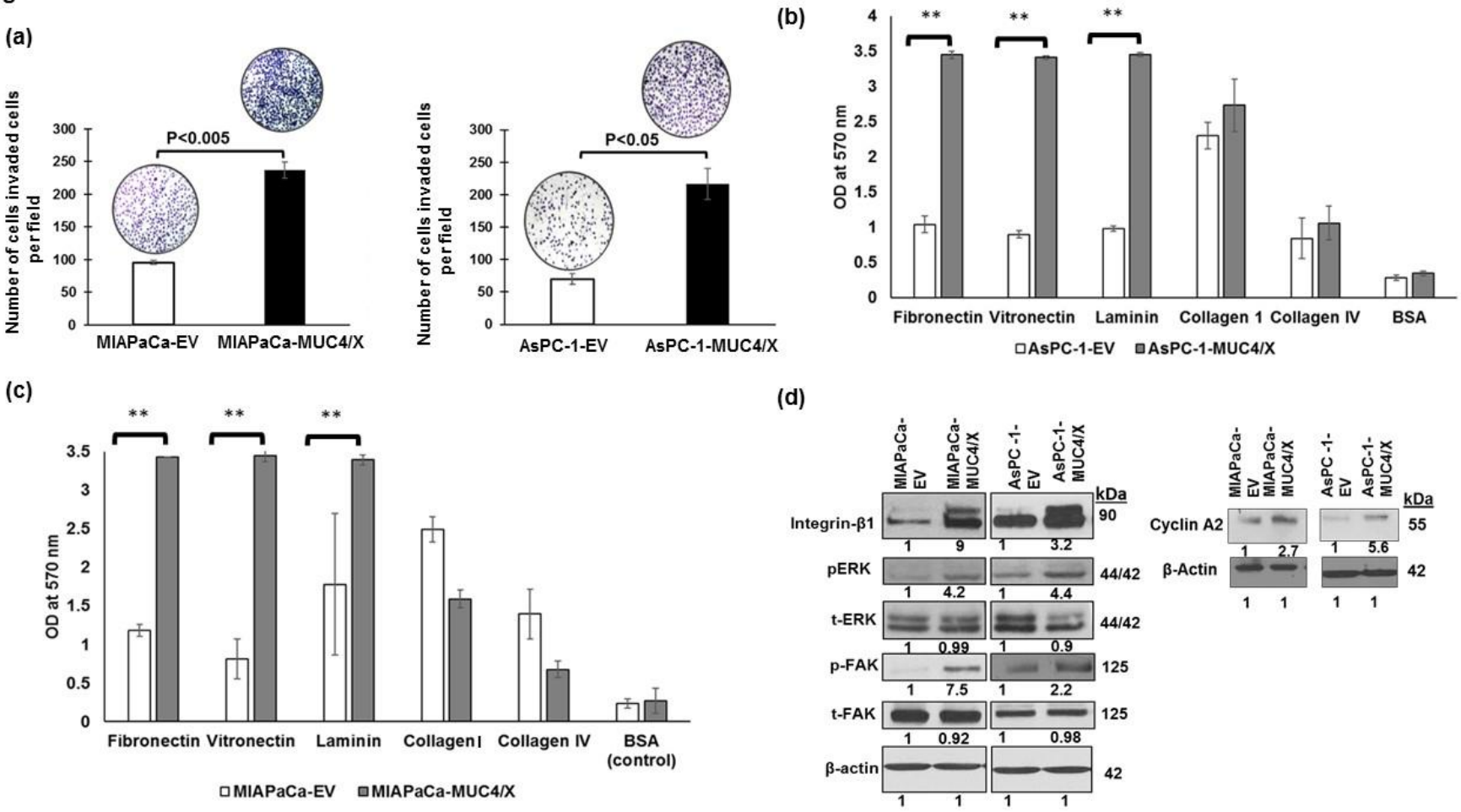

(a)
Number of cells invaded cells per field
P<0.005
MIAPaCa-EV
MIAPaCa-MUC4/X
P<0.05
AsPC-1-EV
AsPC-1-MUC4/X
(b)
OD at 570 nm
Fibronectin
Vitronectin
Laminin
Collagen 1
Collagen IV
BSA
AsPC-1-EV
AsPC-1-MUC4/X
(c)
OD at 570 nm
Fibronectin
Vitronectin
Laminin
CollagenI
Collagen IV
BSA (control)
MIAPaCa-EV
MIAPaCa-MUC4/X
(d)
MIAPaCa-EV
MIAPaCa-MUC4/X
AsPC -1-EV
AsPC-1-MUC4/X
kDa
Integrin-β1
90
pERK
44/42
t-ERK
44/42
p-FAK
125
t-FAK
125
β-actin
42
Cyclin A2
55
β-Actin
42

Figure 3: Functional and molecular implication of MUC4/X overexpression on pancreatic cancer cell. (a) Boyden chamber assay was performed to evaluate the impact of MUC4/X overexpression on invasive property of tumor cells. Higher number of invaded cells were observed in MUC4/X overexpressing MIAPaCa ($p<0.005$) and AsPC-1($p<0.05$) cells in comparison to control cells. Bar diagram represents the average number of invaded cells per field (mean ± SD, n=5). **(b-c)** Cell attachment assay on ECM coated plate was performed to assess adhesion property of MUC4/X-OE and control cells. MUC4/X overexpression increased cell adhesion to extracellular matrix proteins including fibronectin, vitronectin, and laminin (**$p<0.001$). For statistical significance comparison of two groups, a two-tailed, unpaired t-test was performed. Bar diagram represents mean OD value at 570 nm (mean ± SD, n=3). BSA served as negative control for adhesion assay studies. **(d)** To analyze oncogenic signaling pathway activated by MUC4/X, we performed immunoblotting for integrin-β1, pFAK, tFAK (total FAK), pERK, tERK (total ERK) and cyclin A2. β-actin was used as a loading control. Fold change was assessed by quantitative analyses of immunoblotting results upon overexpression of MUC4/X as compared to control cells (represented by numerical value underneath immunoblot results). MUC4/X overexpression led to upregulation of integrin-β1, pFAK, and pERK in both MIAPaCa and AsPC-1 cell line. These results demonstrate that MUC4/X overexpression increased cell invasion and cell adhesion to ECM by modulating integrin-β1/FAK/ERK pathway.

Figure 4

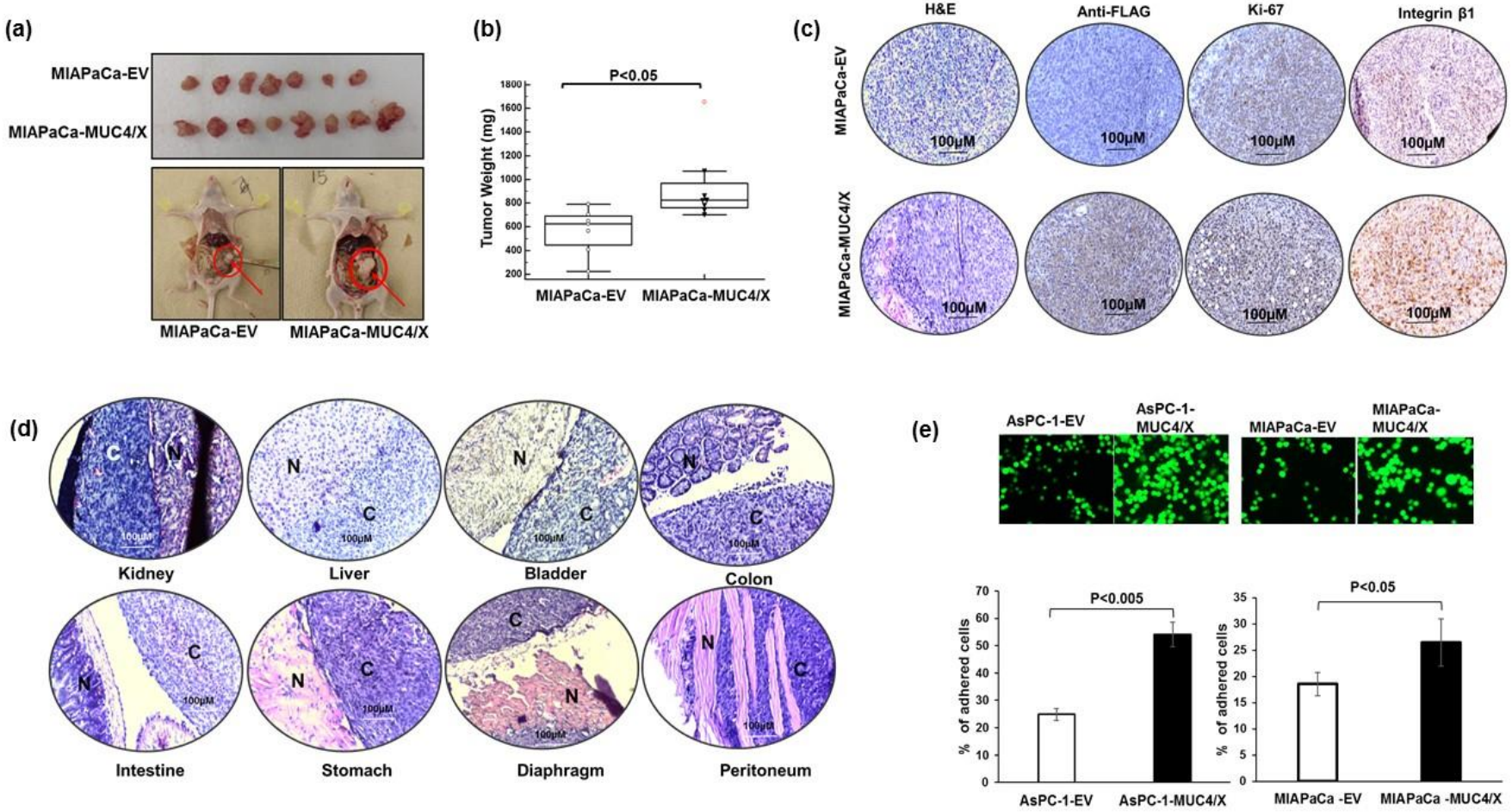

(a)
MIAPaCa-EV
MIAPaCa-MUC4/X
MIAPaCa-EV
MIAPaCa-MUC4/X
(b)
P<0.05
Tumor Weight (mg)
1800
1600
1400
1200
1000
800
600
400
200
MIAPaCa-EV
MIAPaCa-MUC4/X
(c)
H&E
Anti-FLAG
Ki-67
Integrin β1
MIAPaCa-EV
MIAPaCa-MUC4/X
100μM
(d)
Kidney
Liver
Bladder
Colon
Intestine
Stomach
Diaphragm
Peritoneum
C
N
100μM
(e)
AsPC-1-EV
AsPC-1-MUC4/X
MIAPaCa-EV
MIAPaCa-MUC4/X
P<0.005
% of adhered cells
AsPC-1-EV
AsPC-1-MUC4/X
P<0.05
% of adhered cells
MIAPaCa -EV
MIAPaCa -MUC4/X

Figure 4: Effect of MUC4/X overexpression on tumor growth and metastasis *in vivo*. Orthotopic transplantation in the pancreas of nude mice was performed to evaluate the tumorigenic potential of MUC4/X *in vivo*. **(a)** MIAPaCa-EV and MIAPaCa-MUC4/X cells (2.5×10^5) were orthotopically implanted into the pancreas of nude mice (MIAPaCa-EV (n = 7), MIAPaCa-MUC4/X (n=8). Orthotropic implantation of MIAPaCa-MUC4/X cells shows increased tumor-forming ability as compared to the vector-transfected cells. Upper panel image represents extracted tumors from both control and MUC4/X overexpressed groups. The lower panel shows the anatomic image of tumor resection from control and MUC4/X-OE groups. **(b)** The tumors were excised from mice weighed and averaged; the box and whisker plot displays the tumor weights of mice from control and experimental groups. Mice that received MIAPaCa-MUC4/X cells manifest greater tumor weight compared to the mice which received MIAPaCa-EV cells ($p<0.05$). The box and whisker plot represent the interquartile range (IQR) (horizontal line represents the 25^{th} percentile, median and 75th percentile and whisker represents 5th and 95th percentile). The red dot represents outliers. Data were analyzed using two-tailed T-test with unequal variance to determine statistical significance. **(c)** H&E, as well as immunohistochemistry of FLAG (for assessing MUC4/X expression), Ki67, and integrin-β1 was performed on extracted tumor section. IHC revealed elevated level of MUC4/X, Ki67 and integrin-β1 in MIAPaCA-MUC4/X mice group in comparison to MIAPaCa-EV group. **(d)** Micrographs of H&E stained sections showed metastasis to kidney, intestine, colon, bladder, liver, stomach, diaphragm and peritoneum in the MIAPaCa-MUC4/X injected cells in

nude animals. Scale bar is displayed in each image (100 µm). N represents the normal tissues of the organ whereas, C represents cancer cells in the same organ. **(e)** Representative immunofluorescent images from peritoneal adhesion assay are suggesting that AsPC-1-MUC4/X ($p<0.005$) and MIAPaCa-MUC4/X ($p<0.05$) cells have higher cell adhesion to LP-9 peritoneal mesothelial monolayer as compared to control cells. Accompanying bar diagram represents the mean value of the percentage of cells (mean ± SD, n=3) adhered to peritoneal cells. Data was analyzed using two-tailed T-test with unequal variance to determine significance. These results demonstrate that overexpression of MUC4/X promotes tumor growth and metastasis *in vivo*.

Figure 5

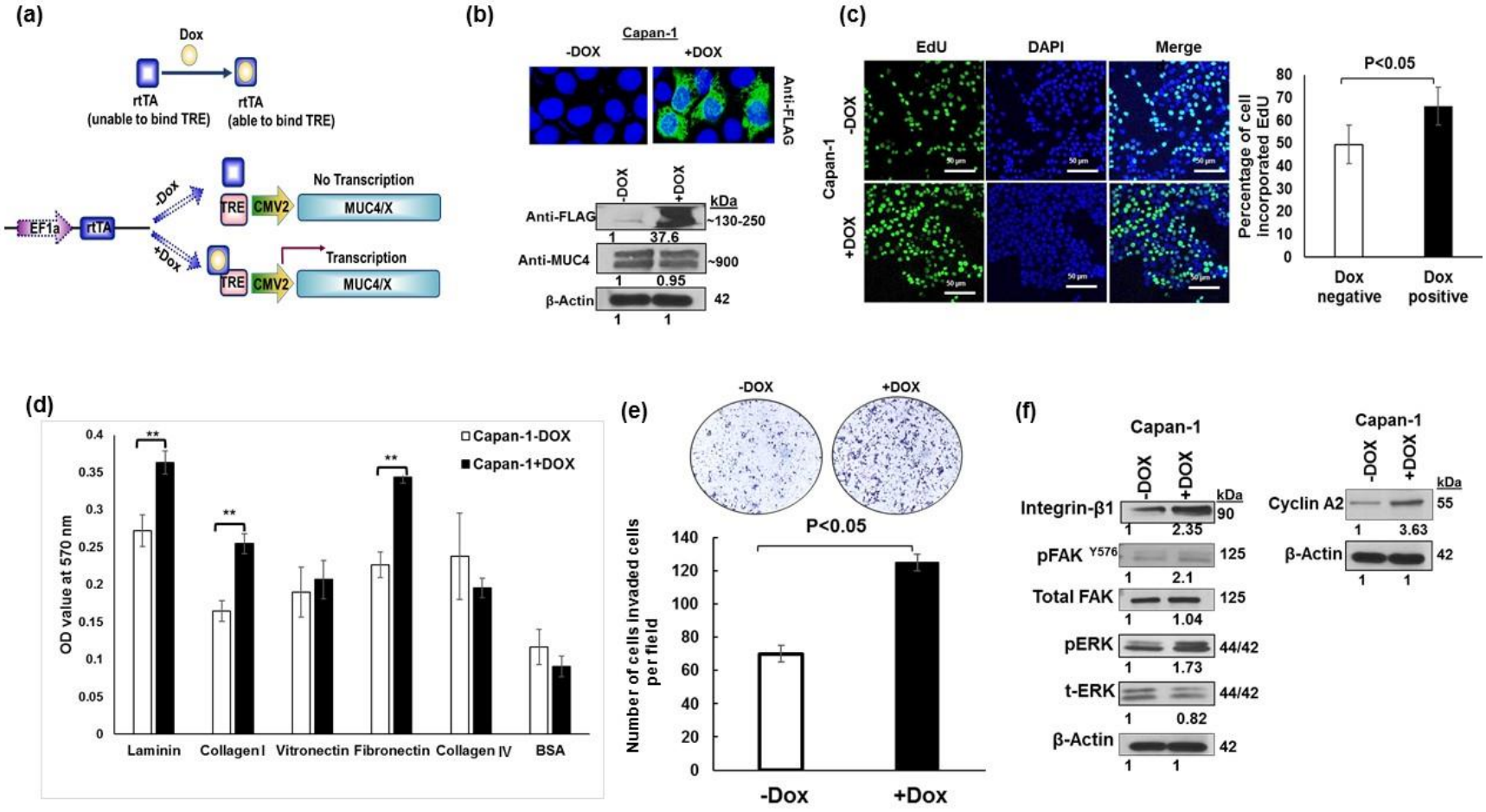

(a)
Dox
rtTA
(unable to bind TRE)
rtTA
(able to bind TRE)
EF1a
rtTA
-Dox
+Dox
TRE
CMV2
No Transcription
MUC4/X
Transcription
(b)
Capan-1
-DOX
+DOX
Anti-FLAG
kDa
Anti-FLAG
~130-250
1
37.6
Anti-MUC4
~900
1
0.95
β-Actin
42
1
1
(c)
EdU
DAPI
Merge
Capan-1
-DOX
+DOX
Percentage of cell incorporated EdU
P<0.05
Dox negative
Dox positive
(d)
Capan-1-DOX
Capan-1+DOX
OD value at 570 nm
Laminin
Collagen I
Vitronectin
Fibronectin
Collagen IV
BSA
(e)
-DOX
+DOX
Number of cells invaded cells per field
P<0.05
-Dox
+Dox
(f)
Capan-1
-DOX
+DOX
kDa
Integrin-β1
90
1
2.35
pFAK Y576
125
1
2.1
Total FAK
125
1
1.04
pERK
44/42
1
1.73
t-ERK
44/42
1
0.82
β-Actin
42
1
1
Capan-1
-DOX
+DOX
kDa
Cyclin A2
55
1
3.63
β-Actin
42
1
1

Figure 5: Mechanistic implication of MUC4/X in the background of WT-MUC4. **(a)** Schematic layout for MUC4/X induction in WT-MUC4 expressing PC cells. Upon doxycycline treatment, the conformational change of rtTA (reverse tetracycline-controlled transactivator) results in binding with TRE (Tet Response Element) which in-turn leads to induction of MUC4/X. **(b)** Immunoblotting (lower panel) and confocal images (upper panel) showing induction of MUC4/X expression upon doxycycline treatment in WT-MUC4 expressing Capan-1 cells. No alteration in the expression of WT-MUC4 was observed on induction of MUC4/X. **(c)** Doxycycline induction of MUC4/X resulted in higher cell proliferation within 24 hrs in the background of WT-MUC4 as indicated by EdU cell proliferation assay. Accompanying bar diagram shows an increase in the percentage of the proliferative cell upon MUC4/X induction ($p<0.05$). **(d)** Induction of MUC4/X resulted in significantly higher cell adhesion to the ECM proteins (**$p<0.005$). The corresponding bar diagram represents OD value (mean ± SD, n=3). BSA served as a negative control. **(e)** Boyden Chamber assay is demonstrating that induction of MUC4/X in Capan-1 cell line resulted in higher number of invaded cells. Accompanying bar diagram represents the number of invaded cells per field (mean ± SD). **(f)** Western blots indicate that increased expression of integrin-β1, pFAK, pERK and cyclin A2 upon induction of MUC4/X in Capan-1 cells. β-actin was used as a loading control. Each blot represents the numerical fold change value upon overexpression of MUC4/X.

Figure 6

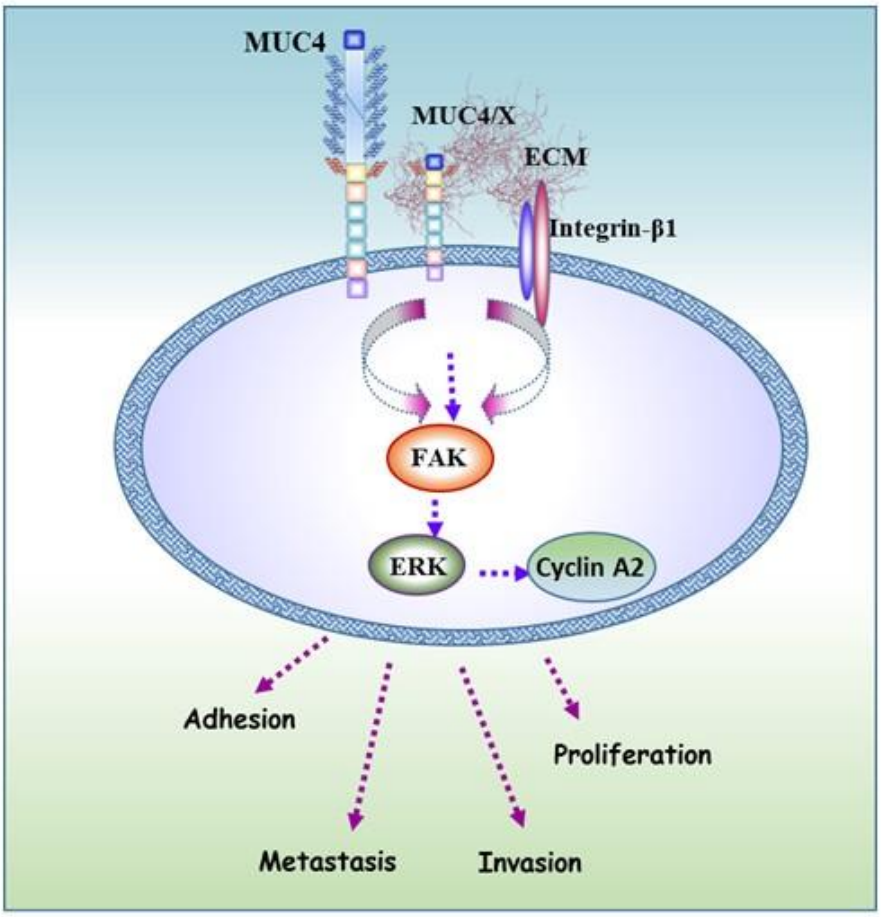
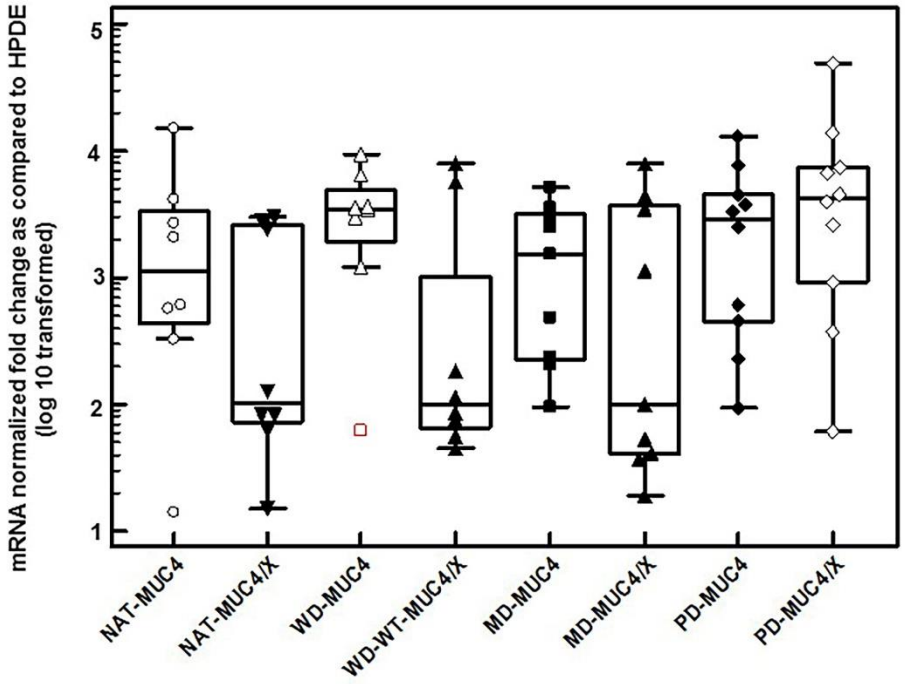

MUC4
MUC4/X
ECM
Integrin-β1
FAK
ERK
Cyclin A2
Adhesion
Proliferation
Metastasis
Invasion

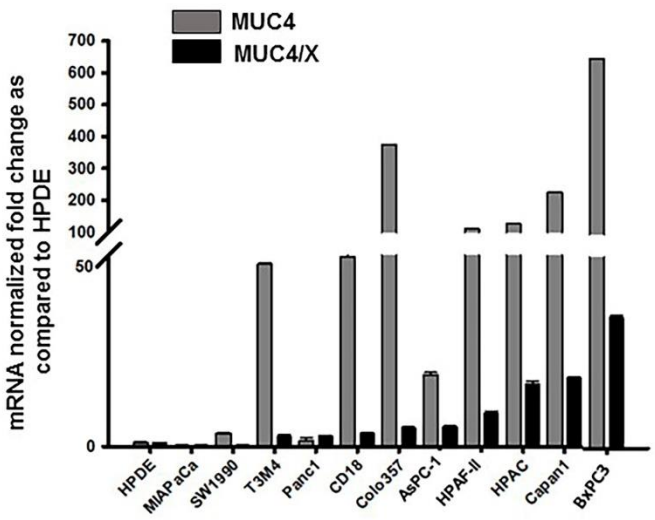
Figure 6: Molecular mechanism for MUC4/X mediated oncogenic signaling in PC tumorigenesis. Schematic diagram showing the oncogenic role of MUC4/X in PC. Ectopic expression of MUC4/X in PC cells resulted in elevated expression of integrin-β1 and increased adhesion to ECM thereby activating downstream FAK/ERK signaling along with cyclin A2 overexpression. Activation of this signaling cascade results in cell proliferation, adhesion, invasion, and metastasis. When the expression of MUC4/X is induced ectopically along with WT-MUC4, it also exhibited a similar aggressive oncogenic phenotype. Overall, overexpression of MUC4/X leads to enhanced aggressiveness of PC.

Supplementary Figure S1

(a)



(b)

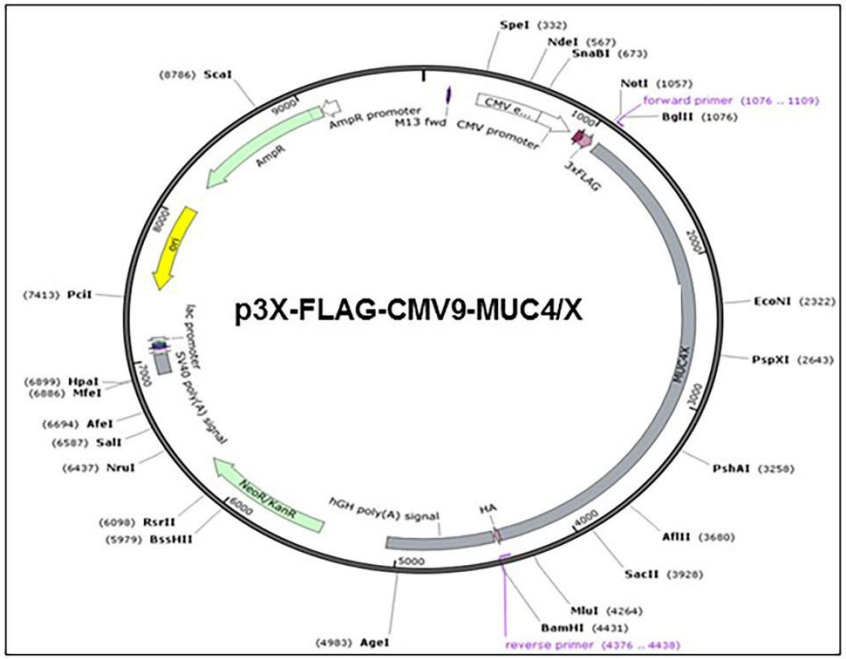


Supplementary Figures

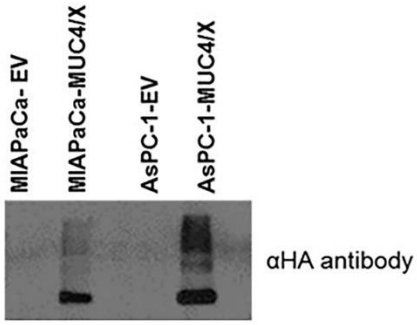
Figure S1: mRNA level of WT-MUC4 and MUC4/X in the human pancreatic tumors and PC cell lines. (a) mRNA expression of WT-MUC4 and MUC4/X in pancreatic tumor tissues as well as NAT tissues. A box and whisker plot represent the interquartile range (IQR) (horizontal line represents the 25th, median and 75th percentile and whisker represents 5^{th} and 95^{th} percentile). Red square represents outlier value. Two-tailed Student's t-test was used to determine the statistically significant difference between the groups. β-actin was used as an internal control to normalize the respective gene, and human pancreatic ductal epithelial cells (HPDE) used as test control for determining fold change. Y-axis represents value as fold change (log 10 relative quantification) relative to the control group. **(b)** Relative fold change in expression level of MUC4/X in a panel of PC cell lines. HPDE served as a test control. β-actin was used as an internal control to normalize the respective gene, and HPDE was used as test control for determining fold change. Y-axis represents normalized fold change in PC cell lines as compared to HPDE.

Supplementary Figure S2

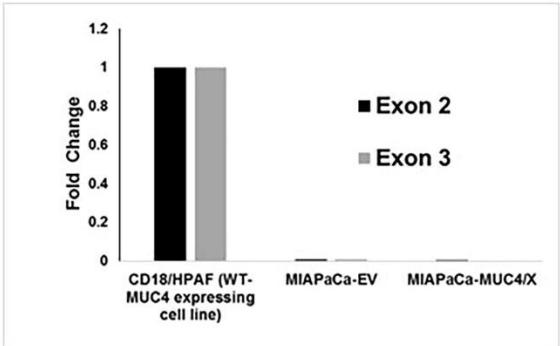
(a)



(b)



(c)



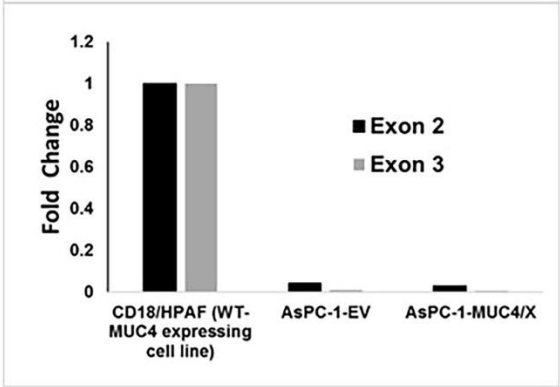


Figure S2: MUC4/X expression plasmid map and confirmation of exons 2 and 3 deletions in MUC4/X-OE PC cell lines by qPCR. (a) Vector map showing that cloned sequence of MUC4/X is flanked by N terminus FLAG tag, and C-terminus HA tag in p3XFLAG-CMV™-9 vectors. The MUC4/X sequence was amplified from prior designed mini-MUC4 construct and cloned into the p3XFLAG-CMV™-9. **(b)** Western blot showing MUC4/X expression using HA antibody suggesting expression of MUC4/X in MIAPaCa and AsPC-1 cell lines while no expression in control cells. **(c)** The qPCR analysis using primer specific to exons 2 as well as exon 3 indicate very minimally, or no expression of exon 2 and exon 3 in MIAPaCa-MUC4/X and AsPC-1-MUC4/X cells suggesting the absence of WT-MUC4. CD18/HPAF cells was used as positive control for WT-MUC4 expression.

Supplementary Figure S3

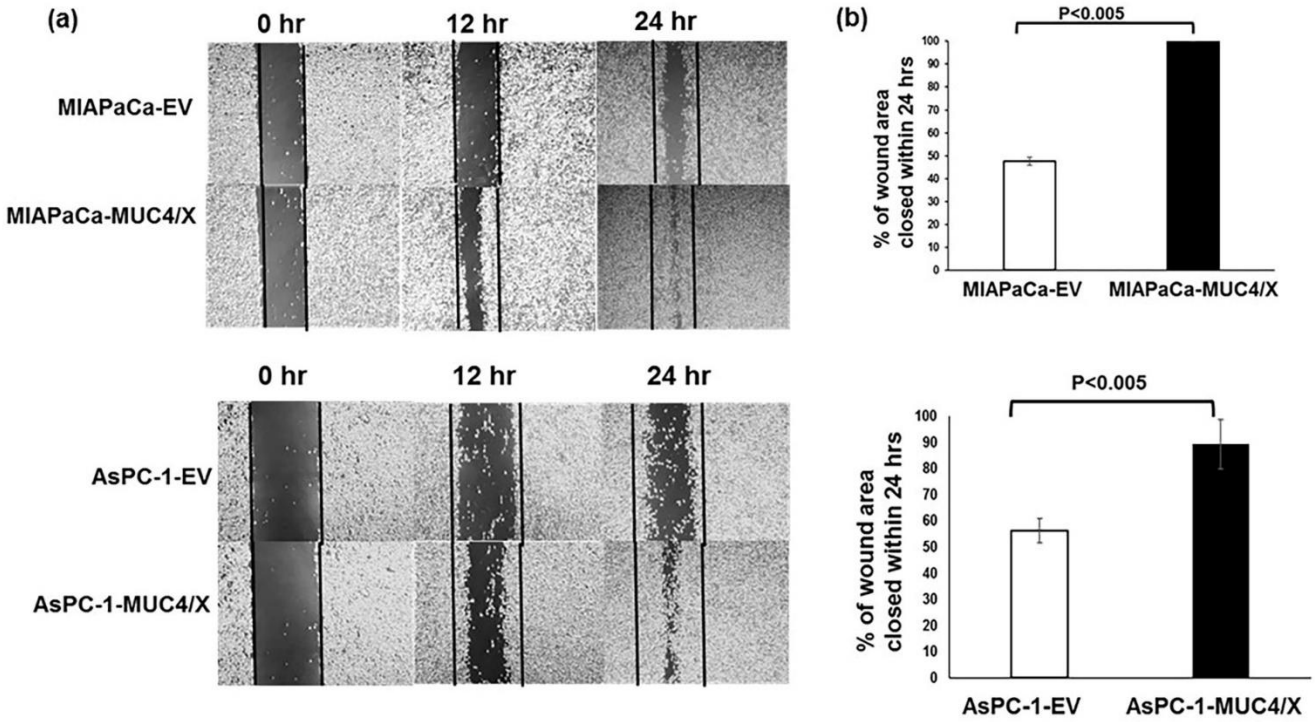
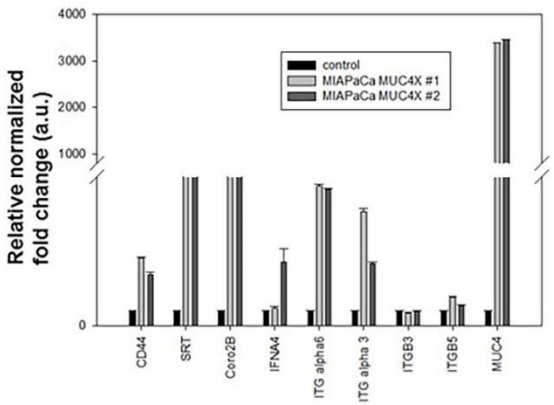


Figure S3: MUC4/X overexpression accelerates migration *in vitro.* (a) Wound healing assay is showing more wound closure by MIAPaCa-MUC4/X and AsPC-1-MUC4/X as compared to control cells after 24 hrs (p <0.005) under minimal medium conditions. Images are showing wound width at 0hr, 12hr, and 24 hrs, after the creation of the wound. The black lines represent the wound boundary. **(b)** Bar graph indicates quantitative results of the percentage of wound area healed/closed (mean ± SD, n=3) within 24 hrs in control and MUC4/X-OE cell

Supplementary Figure S4

a



b

	Microarray Fold Change	QPCR -Fold Change
CD44	2.1	3.4
SRT	4.6	11.8
Coro2B	6.3	51.0
IFNA4	2.5	4.3
ITG alpha6	2.4	9.2
ITG alpha 3	3.1	4.2
ITGB3	1.2	1.0
ITGB5	1.2	1.3
hMUC4	45.9	3460.3

c

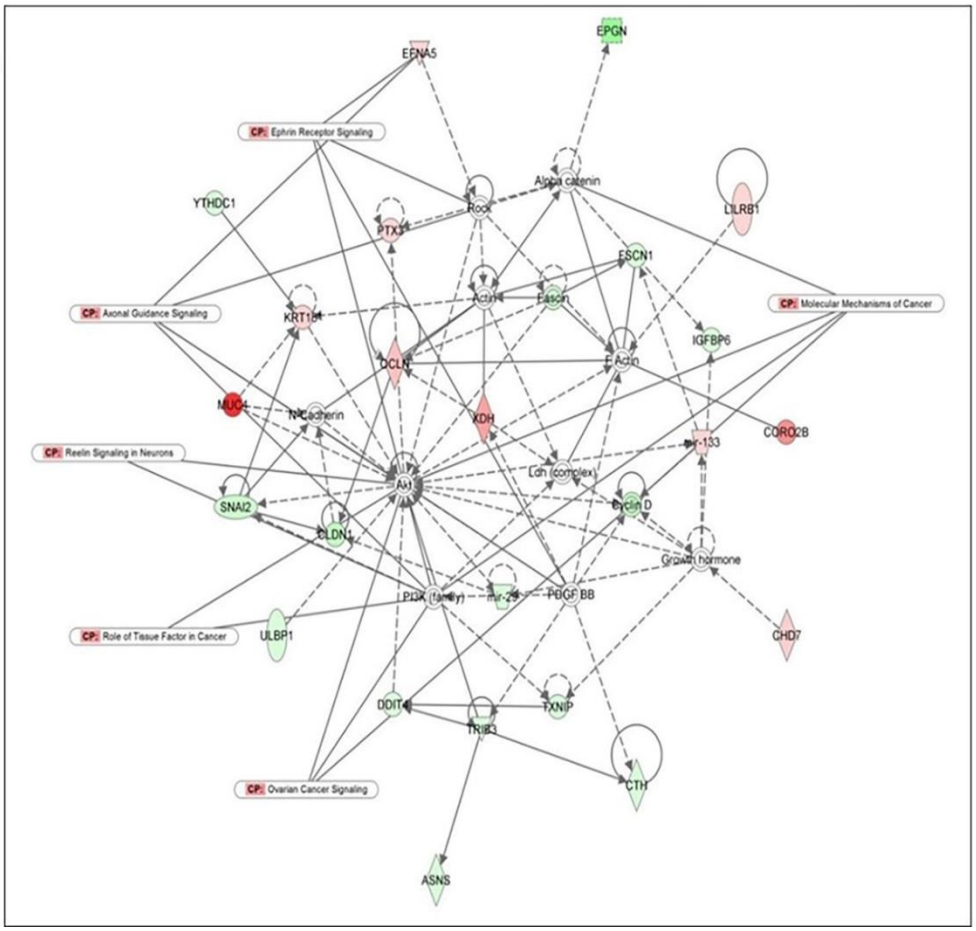


Figure S4: Identification and validation of upregulated molecule modulated by MUC4/X overexpression using microarray analysis along with molecular networks by Ingenuity Pathway Analysis (IPA). (a) The qPCR validation of upregulated molecules in MIAPaCa-MUC4/X as compared to MIAPaCa-EV obtained from microarray analysis. The test genes are listed on the x-axis, and changes in mRNA expression in MIAPaCa-MUC4/X are shown as fold change relative to the mRNA expression in control cells on the y-axis. Fold changes was calculated using the ΔΔCT method. β-actin was used as an internal control to normalize the genes. **(b)** Comparison of fold change of the tested genes between microarray and qPCR data. The intensity of color in the table (red color indicated upregulation whereas green color represents downregulation) represents the degree of upregulation and downregulation. **(c)** IPA pathway analysis of the microarray data obtained from MIAPaCa-EV and MIAPaCa-MUC4/X. Differentially-expressed genes showing ≥2-fold change were selected and analyzed using IPA to illustrate potential interactions and top canonical pathways which are affected by MUC4/X overexpression. Solid arrows represent known direct interactions whereas dotted arrows represent indirect interactions. The intensity of nod color (red color indicated upregulated molecules whereas green color represents downregulated molecules) represents the degree of upregulation and downregulation.

Supplementary Figure S5

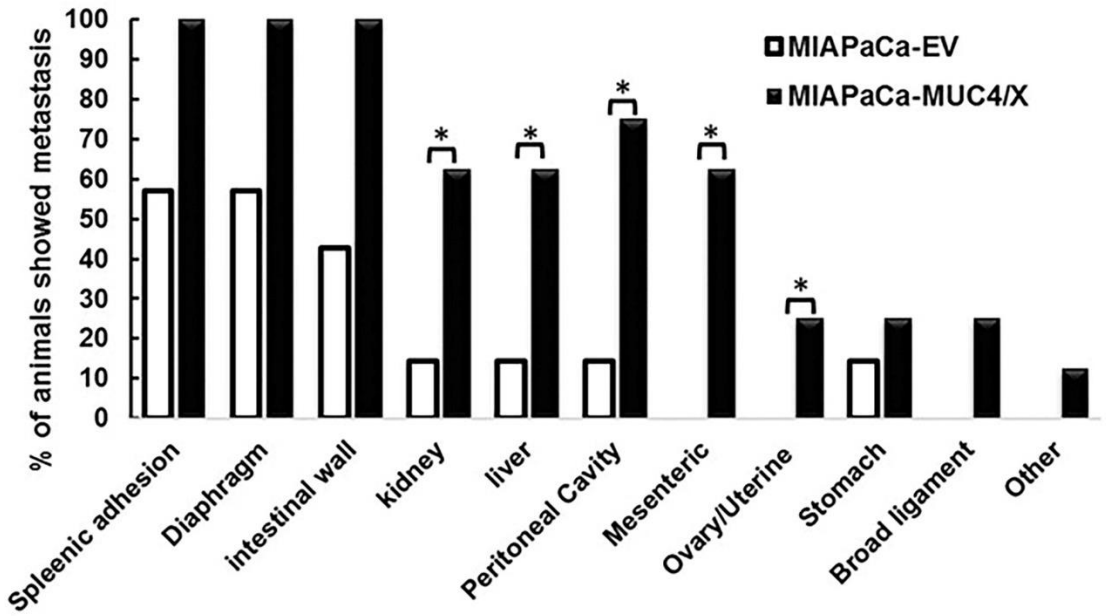


Figure S5: MUC4/X overexpression increases metastasis to different organs. **(a)** Nude mice developed widespread macro-metastases when MIAPaCa-MUC4/X cells were orthotopically injected into the pancreas as compared to MIAPaCa-EV cells. All the mice injected with MIAPaCa-MUC4/X cells developed metastases in the spleen, diaphragm and intestinal wall. In addition, significantly increased metastatic fociwere also observed in kidney, liver, mesentery, peritoneal cavity, and ovary (*****p<0.05) by MUC4/X-OE cells compared to MUC4/X-EV. No metastasis was observed in mesentery, ovary, broad ligament and other organs including cecum, colon, and bladder in control groups. Bar diagram represents the percentage of animals showed metastasis in both groups. Data were analyzed using two-tailed *T*-test with unequal variance between two groups.

References

[1] R.L. Siegel, K.D. Miller, A. Jemal, Cancer statistics, 2018, CA Cancer J Clin.68 (2018) 7-30.

[2] L. Rahib, B.D. Smith, R. Aizenberg, A.B. Rosenzweig, J.M. Fleshman, L.M. Matrisian, Projecting cancer incidence and deaths to 2030: the unexpected burden of thyroid, liver, and pancreas cancers in the United States, Cancer Res.74 (2014) 2913-2921.

[3] M. Ladomery, Aberrant alternative splicing is another hallmark of cancer, Int J Cell Biol.2013 (2013) 463786.

[4] S.C. Lee, H.J. Harn, T.S. Lin, K.T. Yeh, Y.C. Liu, C.S. Tsai, Y.L. Cheng, Prognostic significance of CD44v5 expression in human thymic epithelial neoplasms, Ann. Thorac. Surg.76 (2003) 213-218.

[5] S. Takano, M. Reichert, B. Bakir, K.K. Das, T. Nishida, M. Miyazaki, S. Heeg, M.A. Collins, B. Marchand, P.D. Hicks, A. Maitra, A.K. Rustgi, Prrx1 isoform switching regulates pancreatic cancer invasion and metastatic colonization, Genes Dev.30 (2016) 233-247.

[6] M.J. Swartz, S.K. Batra, G.C. Varshney, M.A. Hollingsworth, C.J. Yeo, J.L. Cameron, R.E. Wilentz, R.H. Hruban, P. Argani, MUC4 expression increases progressively in pancreatic intraepithelial neoplasia, Am. J. Clin. Pathol.117 (2002) 791-796.

[7] P. Mukhopadhyay, I. Lakshmanan, M.P. Ponnusamy, S. Chakraborty, M. Jain, P. Pai, L.M. Smith, S.M. Lele, S.K. Batra, MUC4 overexpression augments cell migration and metastasis through EGFR family proteins in triple negative breast cancer cells, PLoS. One.8 (2013) e54455.

[8] S. Rachagani, M.A. Macha, M.P. Ponnusamy, D. Haridas, S. Kaur, M. Jain, S.K. Batra, MUC4 potentiates invasion and metastasis of pancreatic cancer cells through stabilization of fibroblast growth factor receptor 1, Carcinogenesis.33 (2012) 1953-1964.

[9] N. Jonckheere, N. Skrypek, S. Van, I, Mucins and tumor resistance to chemotherapeutic drugs, Biochim. Biophys. Acta.1846 (2014) 142-151.

[10] P.D. Majhi, I. Lakshmanan, M.P. Ponnusamy, M. Jain, S. Das, S. Kaur, S.T. Shimizu, W.W. West, S.L. Johansson, L.M. Smith, F. Yu, C.E. Rolle, P. Sharma, G.B. Carey, S.K. Batra, A.K. Ganti, Pathobiological implications of MUC4 in non-small-cell lung cancer, J Thorac. Oncol.8 (2013) 398-407.

[11] A.P. Singh, P. Chaturvedi, S.K. Batra, Emerging roles of MUC4 in cancer: a novel target for diagnosis and therapy, Cancer Res.67 (2007) 433-436.

[12] F. Escande, L. Lemaitre, N. Moniaux, S.K. Batra, J.P. Aubert, M.P. Buisine, Genomic organization of MUC4 mucin gene. Towards the characterization of splice variants, Eur. J. Biochem.269 (2002) 3637-3644.

[13] P. Chaturvedi, A.P. Singh, S.K. Batra, Structure, evolution, and biology of the MUC4 mucin, FASEB J.22 (2008) 966-981.

[14] N. Moniaux, F. Escande, S.K. Batra, N. Porchet, A. Laine, J.P. Aubert, Alternative splicing generates a family of putative secreted and membrane-associated MUC4 mucins, Eur J Biochem.267 (2000) 4536-4544.

[15] S. Kumar, E. Cruz, S. Joshi, A. Patel, R. Jahan, S.K. Batra, M. Jain, Genetic Variants of Mucins: Unexplored Conundrum, Carcinogenesis.38 (2016) 671-679

[16] S. Zrihan-Licht, H.L. Vos, A. Baruch, O. Elroy-Stein, D. Sagiv, I. Keydar, J. Hilkens, D.H. Wreschner, Characterization and molecular cloning of a novel MUC1 protein, devoid of tandem repeats, expressed in human breast cancer tissue, Eur. J Biochem.224 (1994) 787-795.

[17] K. Xie, X. Zhi, J. Tang, Y. Zhu, J. Zhang, Z. Li, J. Tao, Z. Xu, Upregulation of the splice variant MUC4/Y in the pancreatic cancer cell line MIA PaCa-2 potentiates proliferation and suppresses apoptosis: new insight into the presence of the transcript variant of MUC4, Oncol Rep.31 (2014) 2187-2194.

[18] Y. Zhu, J.J. Zhang, W.B. Liang, R. Zhu, B. Wang, Y. Miao, Z.K. Xu, Pancreatic cancer counterattack: MUC4 mediates Fas-independent apoptosis of antigen-specific cytotoxic T lymphocyte, Oncol Rep.31 (2014) 1768-1776.

[19] P. Chaturvedi, A.P. Singh, N. Moniaux, S. Senapati, S. Chakraborty, J.L. Meza, S.K. Batra, MUC4 mucin potentiates pancreatic tumor cell proliferation, survival, and invasive properties and interferes with its interaction to extracellular matrix proteins, Mol. Cancer Res.5 (2007) 309-320.

[20] M.A. Hollingsworth, B.J. Swanson, Mucins in cancer: protection and control of the cell surface, Nat. Rev. Cancer.4 (2004) 45-60.

[21] Y. Makiguchi, Y. Hinoda, K. Imai, Effect of MUC1 mucin, an anti-adhesion molecule, on tumor cell growth, Jpn. J Cancer Res.87 (1996) 505-511.

[22] M. Komatsu, C.A. Carraway, N.L. Fregien, K.L. Carraway, Reversible disruption of cell-matrix and cell-cell interactions by overexpression of sialomucin complex, J Biol Chem.272 (1997) 33245-33254.

[23] M.A. Dickson, W.C. Hahn, Y. Ino, V. Ronfard, J.Y. Wu, R.A. Weinberg, D.N. Louis, F.P. Li, J.G. Rheinwald, Human keratinocytes that express hTERT and also bypass a p16(INK4a)-enforced mechanism that limits life span become immortal yet retain normal growth and differentiation characteristics, Mol. Cell Biol.20 (2000) 1436-1447.

[24] S. Senapati, V.S. Gnanapragassam, N. Moniaux, N. Momi, S.K. Batra, Role of MUC4-NIDO domain in the MUC4-mediated metastasis of pancreatic cancer cells, Oncogene.31 (2012) 3346-3356.

[25] N. Moniaux, P. Chaturvedi, G.C. Varshney, J.L. Meza, J.F. Rodriguez-Sierra, J.P. Aubert, S.K. Batra, Human MUC4 mucin induces ultra-structural changes and tumorigenicity in pancreatic cancer cells, Br. J. Cancer.97 (2007) 345-357.

[26] M. Gomez-Martinez, D. Schmitz, A. Hergovich, Generation of stable human cell lines with Tetracycline-inducible (Tet-on) shRNA or cDNA expression, J. Vis. Exp.(2013) e50171.

[27] M.A. Macha, S. Rachagani, A.K. Qazi, R. Jahan, S. Gupta, A. Patel, P. Seshacharyulu, C. Lin, S. Li, S. Wang, V. Verma, S. Kishida, M. Kishida, N. Nakamura, T. Kibe, W.M. Lydiatt, R.B. Smith, A.K. Ganti, D.T. Jones, S.K. Batra, M. Jain, Afatinib radiosensitizes head and neck squamous cell carcinoma cells by targeting cancer stem cells, Oncotarget.8 (2017) 20961-20973.

[28] I. Lakshmanan, S. Rachagani, R. Hauke, S.R. Krishn, S. Paknikar, P. Seshacharyulu, S. Karmakar, R.K. Nimmakayala, G. Kaushik, S.L. Johansson, G.B. Carey, M.P. Ponnusamy, S. Kaur, S.K. Batra, A.K. Ganti, MUC5AC interactions with integrin beta4 enhances the migration of lung cancer cells through FAK signaling, Oncogene.35 (2016) 4112-4121.

[29] J.G. Lee, J.H. Ahn, T. Jin Kim, J. Ho Lee, J.H. Choi, Mutant p53 promotes ovarian cancer cell adhesion to mesothelial cells via integrin beta4 and Akt signals, Sci Rep.5 (2015) 12642.

[30] K. Lessan, D.J. Aguiar, T. Oegema, L. Siebenson, A.P. Skubitz, CD44 and beta1 integrin mediate ovarian carcinoma cell adhesion to peritoneal mesothelial cells, Am. J Pathol.154 (1999) 1525-1537.

[31] S. Das, S. Rachagani, M.P. Torres-Gonzalez, I. Lakshmanan, P.D. Majhi, L.M. Smith, K.U. Wagner, S.K. Batra, Carboxyl-terminal domain of MUC16 imparts

tumorigenic and metastatic functions through nuclear translocation of JAK2 to pancreatic cancer cells, Oncotarget.6 (2015) 5772-5787.

[32] N. Moniaux, G.C. Varshney, S.C. Chauhan, M.C. Copin, M. Jain, U.A. Wittel, M. Andrianifahanana, J.P. Aubert, S.K. Batra, Generation and characterization of anti-MUC4 monoclonal antibodies reactive with normal and cancer cells in humans, J. Histochem. Cytochem.52 (2004) 253-261.

[33] S. Rachagani, S. Senapati, S. Chakraborty, M.P. Ponnusamy, S. Kumar, L.M. Smith, M. Jain, S.K. Batra, Activated KrasG(1)(2)D is associated with invasion and metastasis of pancreatic cancer cells through inhibition of E-cadherin, Br. J. Cancer.104 (2011) 1038-1048.

[34] P. Chaturvedi, A.P. Singh, S. Chakraborty, S.C. Chauhan, S. Bafna, J.L. Meza, P.K. Singh, M.A. Hollingsworth, P.P. Mehta, S.K. Batra, MUC4 mucin interacts with and stabilizes the HER2 oncoprotein in human pancreatic cancer cells, Cancer Res.68 (2008) 2065-2070.

[35] S.C. Bodary, J.W. McLean, The integrin beta 1 subunit associates with the vitronectin receptor alpha v subunit to form a novel vitronectin receptor in a human embryonic kidney cell line, J Biol. Chem.265 (1990) 5938-5941.

[36] V. Bouchard, M.J. Demers, S. Thibodeau, V. Laquerre, N. Fujita, T. Tsuruo, J.F. Beaulieu, R. Gauthier, A. Vezina, L. Villeneuve, P.H. Vachon, Fak/Src signaling in human intestinal epithelial cell survival and anoikis: differentiation

state-specific uncoupling with the PI3-K/Akt-1 and MEK/Erk pathways, J Cell Physiol.212 (2007) 717-728.

[37] A.R. Santos, R.G. Corredor, B.A. Obeso, E.F. Trakhtenberg, Y. Wang, J. Ponmattam, G. Dvoriantchikova, D. Ivanov, V.I. Shestopalov, J.L. Goldberg, M.E. Fini, M.L. Bajenaru, beta1 integrin-focal adhesion kinase (FAK) signaling modulates retinal ganglion cell (RGC) survival, PLoS. One.7 (2012) e48332.

[38] W.W. Gu, J. Lin, X.Y. Hong, Cyclin A2 regulates homologous recombination DNA repair and sensitivity to DNA damaging agents and poly(ADP-ribose) polymerase (PARP) inhibitors in human breast cancer cells, Oncotarget.8 (2017) 90842-90851.

[39] S. Meloche, J. Pouyssegur, The ERK1/2 mitogen-activated protein kinase pathway as a master regulator of the G1- to S-phase transition, Oncogene.26 (2007) 3227-3239.

[40] J.J. Grzesiak, J.C. Ho, A.R. Moossa, M. Bouvet, The integrin-extracellular matrix axis in pancreatic cancer, Pancreas.35 (2007) 293-301.

[41] X.P. Li, X.W. Zhang, L.Z. Zheng, W.J. Guo, Expression of CD44 in pancreatic cancer and its significance, Int J Clin Exp. Pathol.8 (2015) 6724-6731.

[42] E.M. Lisabeth, G. Falivelli, E.B. Pasquale, Eph receptor signaling and ephrins, Cold Spring Harb. Perspect. Biol.5 (2013).

[43] A.V. Biankin, N. Waddell, K.S. Kassahn, M.C. Gingras, L.B. Muthuswamy, A.L. Johns, D.K. Miller, P.J. Wilson, A.M. Patch, J. Wu, D.K. Chang, M.J. Cowley, B.B. Gardiner, S. Song, I. Harliwong, S. Idrisoglu, C. Nourse, E. Nourbakhsh, S. Manning, S. Wani, M. Gongora, M. Pajic, C.J. Scarlett, A.J. Gill, A.V. Pinho, I. Rooman, M. Anderson, O. Holmes, C. Leonard, D. Taylor, S. Wood, Q. Xu, K. Nones, J.L. Fink, A. Christ, T. Bruxner, N. Cloonan, G. Kolle, F. Newell, M. Pinese, R.S. Mead, J.L. Humphris, W. Kaplan, M.D. Jones, E.K. Colvin, A.M. Nagrial, E.S. Humphrey, A. Chou, V.T. Chin, L.A. Chantrill, A. Mawson, J.S. Samra, J.G. Kench, J.A. Lovell, R.J. Daly, N.D. Merrett, C. Toon, K. Epari, N.Q. Nguyen, A. Barbour, N. Zeps, I. Australian Pancreatic Cancer Genome, N. Kakkar, F. Zhao, Y.Q. Wu, M. Wang, D.M. Muzny, W.E. Fisher, F.C. Brunicardi, S.E. Hodges, J.G. Reid, J. Drummond, K. Chang, Y. Han, L.R. Lewis, H. Dinh, C.J. Buhay, T. Beck, L. Timms, M. Sam, K. Begley, A. Brown, D. Pai, A. Panchal, N. Buchner, R. De Borja, R.E. Denroche, C.K. Yung, S. Serra, N. Onetto, D. Mukhopadhyay, M.S. Tsao, P.A. Shaw, G.M. Petersen, S. Gallinger, R.H. Hruban, A. Maitra, C.A. Iacobuzio-Donahue, R.D. Schulick, C.L. Wolfgang, R.A. Morgan, R.T. Lawlor, P. Capelli, V. Corbo, M. Scardoni, G. Tortora, M.A. Tempero, K.M. Mann, N.A. Jenkins, P.A. Perez-Mancera, D.J. Adams, D.A. Largaespada, L.F. Wessels, A.G. Rust, L.D. Stein, D.A. Tuveson, N.G. Copeland, E.A. Musgrove, A. Scarpa, J.R. Eshleman, T.J. Hudson, R.L. Sutherland, D.A. Wheeler, J.V. Pearson, J.D. McPherson, R.A. Gibbs, S.M. Grimmond, Pancreatic cancer genomes reveal aberrations in axon guidance pathway genes, Nature.491 (2012) 399-405.

[44] S. Kusamura, D. Baratti, N. Zaffaroni, R. Villa, B. Laterza, M.R. Balestra, M. Deraco, Pathophysiology and biology of peritoneal carcinomatosis, World J Gastrointest. Oncol.2 (2010) 12-18.

[45] R. Jahan, S. Kaur, M.A. Macha, S.K. Batra, Mucins (MUCs), Encyclopedia of Signaling Molecules, Springer New York. 2016, pp. 1-14.

[46] P. Ciborowski, O.J. Finn, Non-glycosylated tandem repeats of MUC1 facilitate attachment of breast tumor cells to normal human lung tissue and immobilized extracellular matrix proteins (ECM) in vitro: potential role in metastasis, Clin. Exp. Metastasis.19 (2002) 339-345.

[47] K.G. Kohlgraf, A.J. Gawron, M. Higashi, J.L. Meza, M.D. Burdick, S. Kitajima, D.L. Kelly, T.C. Caffrey, M.A. Hollingsworth, Contribution of the MUC1 tandem repeat and cytoplasmic tail to invasive and metastatic properties of a pancreatic cancer cell line, Cancer Res.63 (2003) 5011-5020.

[48] J. Tang, Y. Zhu, K. Xie, X. Zhang, X. Zhi, W. Wang, Z. Li, Q. Zhang, L. Wang, J. Wang, Z. Xu, The role of the AMOP domain in MUC4/Y-promoted tumour angiogenesis and metastasis in pancreatic cancer, J. Exp. Clin. Cancer Res.35 (2016) 91.

[49] G. Bendas, L. Borsig, Cancer cell adhesion and metastasis: selectins, integrins, and the inhibitory potential of heparins, Int. J Cell Biol.2012 (2012) 676731.

[50] M. Barczyk, S. Carracedo, D. Gullberg, Integrins, Cell Tissue Res.339 (2010) 269-280.

[51] P. Lu, V.M. Weaver, Z. Werb, The extracellular matrix: a dynamic niche in cancer progression, J. Cell Biol.196 (2012) 395-406.

[52] D.G. Stupack, D.A. Cheresh, Get a ligand, get a life: integrins, signaling and cell survival, J. Cell Sci.115 (2002) 3729-3738.

[53] J.J. Grzesiak, H.S. Tran Cao, D.W. Burton, S. Kaushal, F. Vargas, P. Clopton, C.S. Snyder, L.J. Deftos, R.M. Hoffman, M. Bouvet, Knockdown of the beta(1) integrin subunit reduces primary tumor growth and inhibits pancreatic cancer metastasis, Int J Cancer.129 (2011) 2905-2915.

[54] V. Bouchard, C. Harnois, M.J. Demers, S. Thibodeau, V. Laquerre, R. Gauthier, A. Vezina, D. Noel, N. Fujita, T. Tsuruo, M. Arguin, P.H. Vachon, B1 integrin/Fak/Src signaling in intestinal epithelial crypt cell survival: integration of complex regulatory mechanisms, Apoptosis.13 (2008) 531-542.

[55] L.A. Hazlehurst, J.S. Damiano, I. Buyuksal, W.J. Pledger, W.S. Dalton, Adhesion to fibronectin via beta1 integrins regulates p27kip1 levels and contributes to cell adhesion mediated drug resistance (CAM-DR), Oncogene.19 (2000) 4319-4327.

[56] P.J. Reddig, R.L. Juliano, Clinging to life: cell to matrix adhesion and cell survival, Cancer Metastasis Rev.24 (2005) 425-439.

[57] A. Deeb, S.U. Haque, O. Olowokure, Pulmonary metastases in pancreatic cancer, is there a survival influence?, J Gastrointest. Oncol.6 (2015) E48-E51.

[58] I. Thomassen, V.E. Lemmens, S.W. Nienhuijs, M.D. Luyer, Y.L. Klaver, I.H. de Hingh, Incidence, prognosis, and possible treatment strategies of peritoneal carcinomatosis of pancreatic origin: a population-based study, Pancreas.42 (2013) 72-75.

[59] S. Kaur, H.A. Kenny, S. Jagadeeswaran, M.R. Zillhardt, A.G. Montag, E. Kistner, S.D. Yamada, A.K. Mitra, E. Lengyel, [218]3-integrin expression on tumor cells inhibits tumor progression, reduces metastasis, and is associated with a favorable prognosis in patients with ovarian cancer, Am. J Pathol.175 (2009) 2184-2196.

[60] N. Arsic, N. Bendris, M. Peter, C. Begon-Pescia, C. Rebouissou, G. Gadea, N. Bouquier, F. Bibeau, B. Lemmers, J.M. Blanchard, A novel function for Cyclin A2: control of cell invasion via RhoA signaling, J. Cell Biol.196 (2012) 147-162.

[61] M.C. Casimiro, M. Crosariol, E. Loro, Z. Li, R.G. Pestell, Cyclins and cell cycle control in cancer and disease, Genes Cancer.3 (2012) 649-657.

[62] Y. Ito, T. Takeda, K. Wakasa, M. Tsujimoto, M. Okada, N. Matsuura, Expression of the G2-M modulators in pancreatic adenocarcinoma, Pancreatology.2 (2002) 138-145

Chapter 4:

Diagnostic implications of Trefoil Factors in pancreatic cancer

Parts of this chapter are driven from:

Jahan R, Smith LM, Sheinin YM, Atri P, Ganguly K, Carmicheal J, Rachagani S, Brand RE, Macha MA, Kaur S, Batra SK.Trefoil Factors and CA19.9: A promising panel for early detection of pancreatic cancer. (Under revision, Ebiomedicine).

Synopsis:

Trefoil factors (TFF1, TFF2, and TFF3) are small secretory molecules that recently have gained significant attention in multiple studies as an integral component of pancreatic cancer (PC) gene signature. Here, we comprehensively investigated the diagnostic potential of all the member of trefoil family, i.e., TFF1, TFF2, and TFF3 (TFFs) in combination with CA19.9 for detection of PC. In silico analysis of publicly available datasets and expression analysis from human and spontaneous PC mouse model revealed a significant increased expression of TFFs in precursor lesions and PC cases. Additionally, we performed a comprehensive analysis in the sample set (n= 377) comprising of independent training and validation set using ELISA consisted of benign controls (BC), chronic pancreatitis (CP), and various stages of PC. Our analysis revealed that TFF1 and TFF2 were significantly elevated in early stages of PC in comparison to BC ($P<0.005$) and CP group ($P<0.05$) while significant elevation in TFF3 levels were in CP group. In receiver operating curve (ROC) analyses, combination of TFFs with CA19.9 emerged as promising panel for discriminating early stage of PC from BC (AUCTFF1+TFF2+TFF3+CA19.9=0.928) as well as CP (AUCTFF1+TFF2+TFF3+CA19.9 =0.943). Notably, at 90% specificity, TFFs combination improved CA19.9 sensitivity by 10% and 25% to differentiate early stage of PC from BC and CP respectively. Similar findings were observed in an independent validation set proving unique biomarker capabilities of TFFs. Overall

our study demonstrated that the combination of TFFs enhanced sensitivity and specificity of CA19.9 to discriminate early stage of PC from BC and CP.

Introduction

Pancreatic cancer (PC) is an aggressive disease with a five-year overall survival rate of <8%. It is the third-leading cause of cancer-related deaths worldwide, and by 2030, it is projected to escalate to second rank of cancer-related death (1, 2). While the five-year survival rate of patients with localized PC is 34.3%, unfortunately, only 10% of total PC patients are diagnosed at an early stage. Approximately 52% of cases are diagnosed at late/metastasized stage, with a worsened five-survival rate of only 2.7% (2). Considering these dire statistics, early detection is key to improved PC patient survival. Therefore, identification of early diagnostic biomarkers may result in a timely therapeutic intervention and lead to improve patient prognosis.

To characterize a prospective diagnostic signature for PC, a compendium of several secretory and membranous proteins was enlisted as potential biomarker candidates that demand methodical validation for clinical effectivity (3). Among the identified 160 secretory molecules, trefoil factors (TFF1, 2 and 3 (TFFs)) were recognized as potential markers for PC (3). TFFs are small, secretory mucin-associated proteins known to protect epithelial cells from various environmental insults (4). Although under physiological conditions they protect the gastric mucosa from inflammation, the oncogenic role of TFFs has been observed in multiple malignancies, including breast, prostate, ovarian, and colon cancers (5). The

secretory nature of TFFs, and their high resistance to proteolytic digestion, acid, and heat degradation qualify them as advantageous from a biomarker perspective. We aimed to explore the individual and combined diagnostic potential of TFFs alone and in combination with CA19.9 in PC. Although they were previously recognized as promising biomarkers, there has been no comprehensive study *assessing the diagnostic capability of TFFs for early detection of PC (6). To evaluate this potential, we explored* publicly available datasets of PC, followed by validation of expression via immunohistochemistry (IHC) in genetically engineered spontaneous mouse model of PC progression and human tissues comprising of normal pancreas adjacent to tumor (NAT), PC precursor lesions (PanIN), and PC tissues. We further evaluated circulatory TFF levels in the sera obtained from *exploratory and validation* clinical *cohorts of PC* patients and control samples and analyzed the biomarker potential of individual TFFs in combination with CA19.9. This study reports a potential diagnostic biomarker panel to identify early-stage PC with improved sensitivity (SN) and specificity (SP).

Results

Expression of TFFs in PanIN lesions and PC from publicly available cancer genome dataset.

We began our exploration with the analysis of TFF1, 2, and 3 expressions in PC using publicly available data sets (GSE43288, GSE16515). Analyzing both data sets, we observed differential expression of all TFFs in PanINs and PC compared to normal controls (**Figure 1 A&B**). Significant upregulation of TFFs was observed

in PanIN (TFF1, P<0.005; TFF2, P<0.005 and TFF3 P<0.05) as compared to normal control (**Figure 1A)**. Similar upregulation of TFFs was observed in PC samples compared to NATs (**Figure 1B**). In line with these results, our analysis of the TCGA genome database from cBioPortal (http://www.cbioportal.org) also showed that TFFs are widely expressed in a variety of cancers, predominantly pancreas, colorectal, breast, and prostate **(Supplementary figure S1A-C)** (9, 10). Interestingly, TFF1 was found to be highest expressed in PC followed by breast and other malignancies **(Supplementary figure S1A).** Similarly, the highest expression of TFF2 was observed in PC followed by colorectal cancer **(Supplementary figure S1B).** Slightly deviating from TFF1 & TFF2, the highest expression of TFF3 was observed in colorectal cancer, followed by PC **(Supplementary figure S1C).** All members of the TFF family were highly expressed in pancreatic tumors, in comparison to other malignancies. Based on the differential upregulation of TFFs from the genomic data, we next sought to comprehensively analyze the expression of TFFs in a panel of PanIN and PC tissues as well as in serum samples.

Expression of TFFs in PC spontaneous mouse model and human clinical samples.

After observing upregulation of TFFs in PC using available data sets, we analyzed the expression of TFF1, 2, and 3 in a well-characterized spontaneous PC mouse model ($Kras^{G12D}$; Pdx1-Cre) (KC) model that recapitulates genetic and histopathological features of human PC (11). Using IHC analysis, we observed elevated expression of all TFF1, 2, and 3 at 10, 20, and 30 weeks of age,

representing early PanIN, late PanIN lesions, and early stage PC, respectively (**Figure 2A**). However, decreased expression of TFF1, 2, and 3 were observed at 40 weeks of age (late stage PC) (**Figure 2A**).

We then analyzed the expression of these TFFs in TMAs representing human NAT, PanIN, and PC tissues (**Figure 2B**). In corroboration with our genomic and KC model data, we observed significant upregulation of all TFFs in PanIN lesions and different grades of PC. No expression of any TFF was detected in NAT. Strong expression of TFFs was observed in PanINs, well differentiated (WD), and moderately differentiated (MD) PC tissues, with moderate expression in poorly differentiated (PD) tissues (**Figure 2B**). The H-score of TFFs was found to be significantly higher for PanIN I-III ($p<0.0005$ for TFF1, $p<0.005$ for TFF2 and $p<0.05$ for TFF3), WD ($p<0.0005$ for TFF1, $p<0.005$ for TFF2 and TFF3) and MD ($p<0.0005$ for TFF1, $p<0.005$ for TFF2 and TFF3), as compared to NAT (**Figure 2C**). Similar results of differential TFFs expression in tumor-associated ducts were also observed using immunofluorescence (**Supplementary figure S2A)**. We also observed strong positive staining of all TFFs in metastatic liver tissues (5/5) (**Supplementary figure S2B).** Strong expression of TFF3 was observed in the Islet of Langerhans (**Supplementary figure S2C)** while no expression of other two TFFs were observed. Overall, our results demonstrated elevated expression of all the TFFs in PanINs and early stages of PCs. Given the significant overexpression in early stages of PC development, we reasoned that TFFs can have the potential for early diagnosis of PC, even before the onset of symptoms.

Circulating levels of TFFs in clinical samples

To investigate the diagnostic potential of TFFs in PC, we analyzed their level in PC patient serum samples, using ELISA. The demographic and clinical characteristics of patients were detailed earlier (exploratory training set from the study) and comprise BC, CP, EPC, and LPC patient serum samples (8). We observed that the median serum levels of TFF1 in patients with BC, CP, EPC, and LPC were 257 pg/ml (Inter Quartile Range, IQR: 156-616 pg/ml), 270.26 pg/ml (IQR: 185-574 pg/ml), 370.41 pg/ml (IQR: 214-1002 pg/ml), and 303 pg/ml (IQR: 186-589 pg/ml) respectively. The median serum levels of TFF2 in patients with BC, CP, EPC, and LPC were 3768 pg/ml (IQR: 2510-5322 pg/ml), 3683 pg/ml (IQR: 2679-7450 pg/ml), 5792 pg/ml (IQR: 3518-8932 pg/ml), and 4807 pg/ml (IQR: 2757-7556 pg/ml), respectively. The median serum levels of TFF3 in patients with BC, CP, EPC, and LPC were 9348 pg/ml (IQR: 6728-13223 pg/ml), 11945 pg/ml (IQR: 7452-19149 pg/ml), 11168 pg/ml (IQR: 7756-19865 pg/ml), and 9183 pg/ml (IQR: 6329-14851 pg/ml), respectively. TFFs levels are plotted on a logarithm scale (**Figure 3 A-C**). The median serum level of TFF1 was significantly higher in EPC as compared to BC ($P< 0.005$) and CP ($P<0.05$) (**Figure 3A**). Serum levels of TFF2 were also significantly higher in the EPC group compared to BC and CP (**Figure 3B**). In contrast, a significant elevation in TFF3 was observed in the CP group compared to the BC group ($P<0.01$), with no further change observed during PC development (**Figure 3C**). Of interest, low circulating levels of all TFFs were observed in late-stage PCs compared to early-stage PCs (**Figure 3A-C**).

We also investigated the possible correlations between TFFs and clinicopathological variables such as gender, age, race, bilirubin level, and alcohol history. No apparent difference was observed in the mean value of TFF levels across gender, alcohol history, race (African American, Asian & Caucasian) and bilirubin level. We noted that levels of TFFs have a strong correlation with age (P<0.0001, P=0.0004 and P=0.0014 for TFF1, TFF2, and TFF3, respectively) (**Supplementary table S2).** Circulatory level of all TFFs level were significantly high in patients aged more than 64 (**Supplementary table S2**)**.**

Diagnostic performance of TFF1-3 individually and in combination in a training cohort

Considering upregulated expression of TFFs across disease groups, we next sought to explore their diagnostic potential either alone or in combination to differentiate various stages of PC from benign controls. The diagnostic performance of TFFs alone as well as in combination was assessed by using ROC curve analysis **(Table 1 and Supplementary table S3)**. In discriminating BCs from PC, individual TFFs showed moderate discriminatory potential with SN/SP values for TFF1, TFF2, and TFF3 being 0.717/0.462, 0.518/0.772 and 0.527/0.643, respectively, and AUCs 0.610, 0.639, and 0.575, respectively **(Table 1, Figure S3A).** For differentiating CP from PC, SN/SP values for TFF1, TFF2, and TFF3 were 0.842/0.362, 0.675/0.525, and 0.459/0.652, respectively, and AUCs of 0.622, 0.576, and 0.551, respectively **(Table 1, Supplementary figure S3A)**. We also analyzed whether the ratio of circulating individual TFF levels could differentiate between disease groups. Among all possible ratios, TFF1/TFF3 showed the most

promising potential to differentiate between PC *vs* CP (AUC= 0.701) and EPC *vs* CP, (AUC= 0.731) (**Supplementary figure S3B**).

We next analyzed the diagnostic performance of TFFs in combination of two and three in each group **(Supplementary table S3)**. For this, we first assessed correlation across TFFs and disease group. Our results suggested that the correlation coefficient was significantly higher between TFF1 and TFF3 (R= 0.514, P=0.0003) in the CP group. (**Supplementary figure S4**). Interestingly, this dual combination of TFF1 and TFF3 could distinguish PC and EPC from CP, AUCs of 0.711 and 0.724, respectively (**Supplementary table S3**).

A combination of all TFFs demonstrated an AUC of 0.664 (95%CI, 0.586-0.742) with SN/SP 0.728/0.542 to segregate PC from BC. A similar prediction trend was observed for this panel to distinguish EPC from BC. Compared to any single TFF, the combination panel of TFFs showed significant improvement in differentiating PC from CP, with an AUC 0.759 (95%CI, 0.671-0.846) and SN/SP of 0.466/0.923 (**Table 2**)**.** Similarly**,** to distinguish early-stage PC from CP, the panel achieved an AUC value of 0.760 (95%CI, 0.664-0.857) with SN/SP 0.509/0.897. Values of AUC and sensitivity/specificity at optimal cutoffs are presented in **Table 1.** Values of AUC and SN/SP at optimal cutoffs for dual combinations of TFFs are presented in **Supplementary table S3**.

The combination of TFFs with CA19.9 improves diagnostic performance

The correlation between individual TFF and CA19.9 was evaluated using Pearson's correlation analysis. The Spearman correlation coefficient (r) was used to delineate the correlation between TFF and CA19.9 serum levels

(Supplementary figure S4). A positive correlation was observed for TFF1 and CA19.9 in BC (r=0.217, P=0.033) and stages 1 and 2 of PCs (r=0.226, P=0.05). Significant positive correlation was also observed for TFF2 and CA19.9 (r=0.389, P=0.013) in the CP group **(Supplementary figure S4)**. Of note, a negative correlation was observed between TFF2 and CA19.9 in EPC (r=-0.186, p=0.17), whereas a significant positive correlation was observed in LPC (r=0.386, p=0.0089). Additionally, a positive but not significant correlation was observed for TFF3 and CA19.9 in all the groups **(Supplementary figure S4)**. Our correlation analysis suggested that these markers can be complementary to each other in various groups, which can improve overall efficacy of diagnosis.

As individual TFF showed moderate discriminatory potential and positive correlation with CA19.9, we next investigated the diagnostic performance of all TFF in combination with CA19.9. Our purpose was to analyze whether various combinations of TFFs can improve the diagnostic ability of CA19.9. CA19.9 differentiated PC from BC with SN/SP 0.864/0.810 (**Figure 4A**). Combining TFFs with CA19.9 showed improved efficiency to distinguish PC from BC with AUC 0.936 (95%CI, 0.874-0.985) compared to CA19.9 alone, AUC 0.910 (95%CI, 0.874-0.946) (**Figure 4A**). Sensitivity also increased from 0.755 to 0.848 (at 90% specificity), to discriminate between PC and BC (**Table 2**). To discriminate between EPC and BC, the panel showed an interesting 10% increase in sensitivity (at 90% specificity) compared to CA19.9 alone **(Table 2)**. In the case of discerning PC from CP, the AUC value escalated from 0.906, (95%CI 0.862-0.950) to 0.943, (95%CI, 0.902-0.984). The combination of all four markers dramatically improved

SN/SP of CA19.9 from 0.867/0.809 to 0.899/0.923 (**Table 2**). Moreover, a 15.7% rise in sensitivity (at 90% specificity) was observed to distinguish PC from CP. Analysis of EPC and CP also demonstrated a sharp increase of sensitivity from 0.667 to 0.925 (at 90% specificity) after the addition of all TFFs to CA19.9 (**Table 2**)**.** In the case of differentiating LPC from BC or CP, the panel reasonably improved sensitivity and specificity as well as the overall AUC value (**Table 2**). AUC values and SN/SP of this panel for the different groups (at optimal cutoffs as well as sensitivity at 90% specificity) for the panel are presented in **Table 2.** The overall performance of all possible combination of TFF with CA19.9 are depicted in **supplementary figure S5A-S5B**.

One of the drawbacks of standard CA 19.9 is that 15-20% people do not express CA19.9 and therefore possess a risk of false negative results (7). Our finding demonstrates differential correlation between CA19.9 and TFF1-3, which suggests that their addition should complement CA19.9 to identify PC. In light of this, we next sought to identify the diagnostic role of TFFs in low expressing CA19.9 (<37U/ml) and high expressing CA19.9 (>37U/ml) PC patient samples, assuming a likely possibility that low CA19.9 PC patients are Lewis negative. We grouped the patients based on the well-established and recommended cut-off value for CA19.9, 37U/ml (12). We found that a combination of TFF1-3 can better discriminate PC from CP in low CA19.9 expressing group, AUC- 0.815, than in high CA19.9 expressing group, AUC 0.728. **(Figure 4B)**. The ability to discriminate between EPC and CP was also improved in low vs. high expressing CA19.9 groups, AUC 0.865 vs. 0.712, and SN/SP 1/0.615 vs. 0.915/0.462 **(Figure 4B**,

Supplementary table S5, and S6). As demonstrated with these correlation and ROC curve results, a combination of TFF1-3 can complement CA19.9 to determine PC status.

Diagnostic performance of TFF1-3 in the validation cohort

Next, we performed an independent study for TFF1-3 to validate their discriminatory potential in a blinded serum sample cohort. In compliance with our earlier results, we observed individual diagnostic performance of TFF1 to be SN/SP 0.348/1.00, TFF2 SN/SP 0.409/0.375, and TFF3 SN/SP 0.348/0.875 (**Supplementary figure S6A, Supplementary table S7**). Moreover, the combination of TFF1-3 held an AUC value of 0.75, to differentiate between PC and BC, SN/SP 0.455/0.680. This combination was also able to differentiate between PC and CP, (AUC 0.84, SN/SP 0.636/0.440) **(Supplementary figure S6B**, **Supplementary table S7**). In the validation cohort, a combination of TFFs also showed potential to discriminate PC from control: this needs further validation in the large multicenter clinical cohort.

Discussion

Trefoil Factors have recently emerged as a prominent player in PC pathogenesis. They have been identified by multiple individual studies as top differentially expressed genes in the classical subtype of PC (13, 14). From the biomarker point of view, TFF1, along with Lymphatic vessel endothelial hyaluronan receptor 1 (LYVE1) and Regenerating Family Member 1 Alpha (REG1A), has shown promising results as urinary markers for PC (15). Moreover, findings have

suggested that TFF1 originates from PC, since its level sharply decreases after surgical removal of the tumor (15). The potential of TFFs to determine disease status is also well evident in other cancers. For instance, TFF3 has been demonstrated as a promising biomarker in colorectal cancer and gastric cancers compared to conventional markers (16-18). Though TFFs have been proposed as potential diagnostic markers for PC in many studies, to our knowledge, there has been no comprehensive study of all trefoil family members for diagnosing this disease. Here, we analyzed the expression of TFFs along PC progression and evaluated their potential to improve diagnosis of PC at early stage with better accuracy.

Preclinical exploration of differentially expressed genes from microarray and GEO datasets, in conjunction with previously published reports, suggest that TFFs are differentially upregulated in PC. Our findings from cBioportal also showed very high expression of TFF1, 2, and 3 in PC compared to 30 other malignancies listed in the databases. Genes and proteins overexpressed in PanIN lesions hold the potential to detect PC at early stages. While investigating early genetic aberrations during PC pathogenesis, Guo et al. identified TFF1 overexpression in the PanIN lesions. Earlier, transcriptomic analysis also revealed that TFF1 was one of the top upregulated molecules in sporadic and familial PanINs (19). This lends credence to the use of TFF1 as a biomarker to identify cystic precursor lesions as well as early stages of PC. Moreover, elevated expression of all TFF1, 2, and 3 were reported in intraductal papillary-mucinous neoplasms (IPMNs) of the pancreas (20). Using IHC analysis, a significant proportion of pancreatic adenocarcinoma

cells 23/45 (55%) and ampullary tumor cells, 8/10 (80%) were shown to overexpress TFF1 (21). In corroboration with earlier reports, our study also revealed higher expression of TFFs in well-differentiated PC tumors compared to undifferentiated tumors (22). In addition, in our study strong expression of TFFs in metastatic liver tissues was in agreement with an earlier study by Moffit et al.(14). While our present work and other previous studies observed overexpression of TFF3 in the Islets of Langerhans, the pathophysiological relevance of this overexpression is still unknown. Serum TFF3 is known to have a proliferative effect on pancreatic islet β-cells and therefore can therapeutically benefit type 1 and 2 diabetic patients (23). Also, overexpression of TFF3 in the liver of diabetic and obese mice was shown to improve glucose tolerance by decreasing blood glucose levels and inhibiting genes involved in gluconeogenesis (24). While insulin resistance and glucose intolerance are associated with PC pathogenesis, it becomes imperative to investigate the role of TFF3 in diabetes, a well-known risk factor for PC (25).

Although CA19.9 is by far the most commonly used and standard biomarker for PC, several drawbacks of CA19.9 persist and limit its use. These include false negative results in the 15-20% of patients with a Lewis negative genotype, elevated level in other cancers, and GI diseases (7, 26, 27). Multi-marker diagnostic panels have previously shown promise for many cancers such as breast and CRC (28-31). Our group and others investigated the potential of combining CA19.9 with other biomarkers including intercellular adhesion molecule 1 (ICAM-1), osteoprotegerin (OPG), osteopontin (OPN), human epididymis secretory

protein 4 (HE4), and neutrophil gelatinase-associated lipocalin (NGAL) to improve diagnostic performance (32-35). While all these studies demonstrated the ability to differentiate PC cases from a healthy control, they did not prove to be beneficial for prediagnostic risk assessment for PC (36). Therefore, ongoing efforts to validate the circulating levels of additional biomarkers which are differentially expressed in pancreatic tumors and preneoplastic lesions will be beneficial to increase the detection of PC at early stage. Thus, TFFs hold a promise as a potential biomarker because of their elevated expression in PanIN and PC. While our results revealed the failure of individual TFFs to discriminate PC from BC, the combination of TFFs with CA19.9 demonstrated a surprisingly better diagnostic performance than CA19.9 alone. In addition, this combination of TFFs with CA19.9 significantly improved the diagnostic potential of distinguishing EPC from BC and CP as compared to CA19.9 alone. Another unique finding of our study demonstrated that the combination of TFF1-3 can differentiate between PC and control groups in patients with low CA19.9 expression (<37 U/ml). Taken together, our study suggests the use of TFF1-3 and CA19.9 combination as a potential diagnostic marker for PC diagnosis.

The strength of our study is the comprehensive evaluation of TFFs alone or in combination with CA19.9 as potential PC biomarkers by employing human tissues, large cohort of serum samples, and the use of a mouse progression model of PC. While our study strongly suggests that a combination of TFF1-3 and CA19.9 discriminates early stage PCs from BCs with improved sensitivity and specificity, we also observed decreased TFFs expression in late stages of PC. Though the

underlying mechanisms for this downregulation is still unknown, alteration in methylation patterns between well and poorly differentiated PCs might be one of the reasons (37). Earlier studies have shown hypomethylated TFF2 promoter in 84% of PC tissues and treatment with DNA methyltransferase inhibitor and histone deacetylase inhibitor in PC cell line, where TFF2 is silenced by methylation, led to the activation of this gene (38). Recent studies using spontaneous mouse model demonstrated that loss of TFF1 and TFF2 enhanced PanIN progression along with PC, which reveals their tumor-suppressing role (39, 40). Nevertheless, TFF1 and TFF2 have been shown to increase PC cell proliferation and migration (22, 41). However, based on our findings and earlier reports, we believe that TFFs are very critical factors in initiating PC and that they warrant precise study with mouse model.

The limitations of our study include the small sample size in the validation cohort. In addition, CA19.9 and TFFs are shown to be elevated in other cancers, but our study has investigated the diagnostic potential of this biomarker panel in PC only, and therefore further research in other cancers as also warranted (5, 42). Furthermore, multi-institutional validation and cross-validation for this panel are needed to make it a reliable multimarker panel. Moreover, TFF1 has shown to be a promising urine biomarker for PC (15). Our analysis from the training and validation sets suggest that a combination of TFF1 and TFF3 should be given more emphasis for future biomarker studies, as they performed better in both datasets to differentiate between PC and CP. Therefore, it will be interesting to investigate their diagnostic performance in pancreatic juice and urine as well.

To translate this diagnostic panel from bench to bedside, more effort should be made to uncover the molecular landscape of TFFs in PC. Specifically, it would be interesting to explore whether increased levels of TFFs, both in serum and tissue, are the drivers or the consequence of disease progression. TFF1 was previously correlated with increased PC cell proliferation and metastasis (41), and recombinant treatment with TFF2 has induced PC cell migration (43). By contrast, loss of TFF2 from a newly defined progenitor compartment in PC, coined the pancreatic duct gland, has shown to accelerate IPMN formation (39).

Recently, Collisson et al. identified three PC subtypes: classical, quasi-mesenchymal, and exocrine-like, based on gene signatures from human and mouse PC samples. They observed that classical subtype is more gemcitabine-resistant compared to other subtypes (13). In another study, the Moffitt group identified two subgroups, 'classical' and 'basal-like', where basal-like tumors showed a strong trend toward a better response to adjuvant therapy. Surprisingly, both groups have discovered family members of TFFs to be critical contributors in the classical subtype of PC. Identification of PC subtyping has created a new avenue for PC precision medicine. This will pave the way to improved clinical outcomes and therapeutic response based on intrinsic molecular variabilities among patient groups that clinically progress at different rates and may respond differently to administered therapies (13, 14). We believe that untangling the complex mechanism of PC progression, as well as understanding the genomic landscape of PC subtypes, is thus urgently required for the development of novel

screening strategies and chemopreventive approaches for PC. Uncovering the role of subtype-specific molecules like TFFs is much needed.

The identification of an early diagnostic marker is gaining unprecedented attention not only because it provides insight into disease occurrence but also provides the impetus for developing novel strategies for therapeutic intervention. While still in its infancy, validation of TFFs in combination with CA19.9 in serum will not only predict the presence of PC, but may also have utility in stratifying patients for appropriate therapeutic regimen selection, given that TFFs have been shown to be highly upregulated in classical subtype of PC.

Based on our knowledge from the published literature, this study is the first of its kind to demonstrate the potential role of TFF1, TFF2, and TFF3 as serum-based markers for diagnosing early stage PC. However, the establishment of a clinically valuable biomarker panel requires exhaustive validation and cross-sectional multicenter studies, and this is our ongoing research focus for this panel. Our results suggest compelling evidence from publicly available datasets, tissue, and serum analysis that TFF1, TFF2, and TFF3, along with CA19.9, can be a useful biomarker for identifying PC.

Figure 1

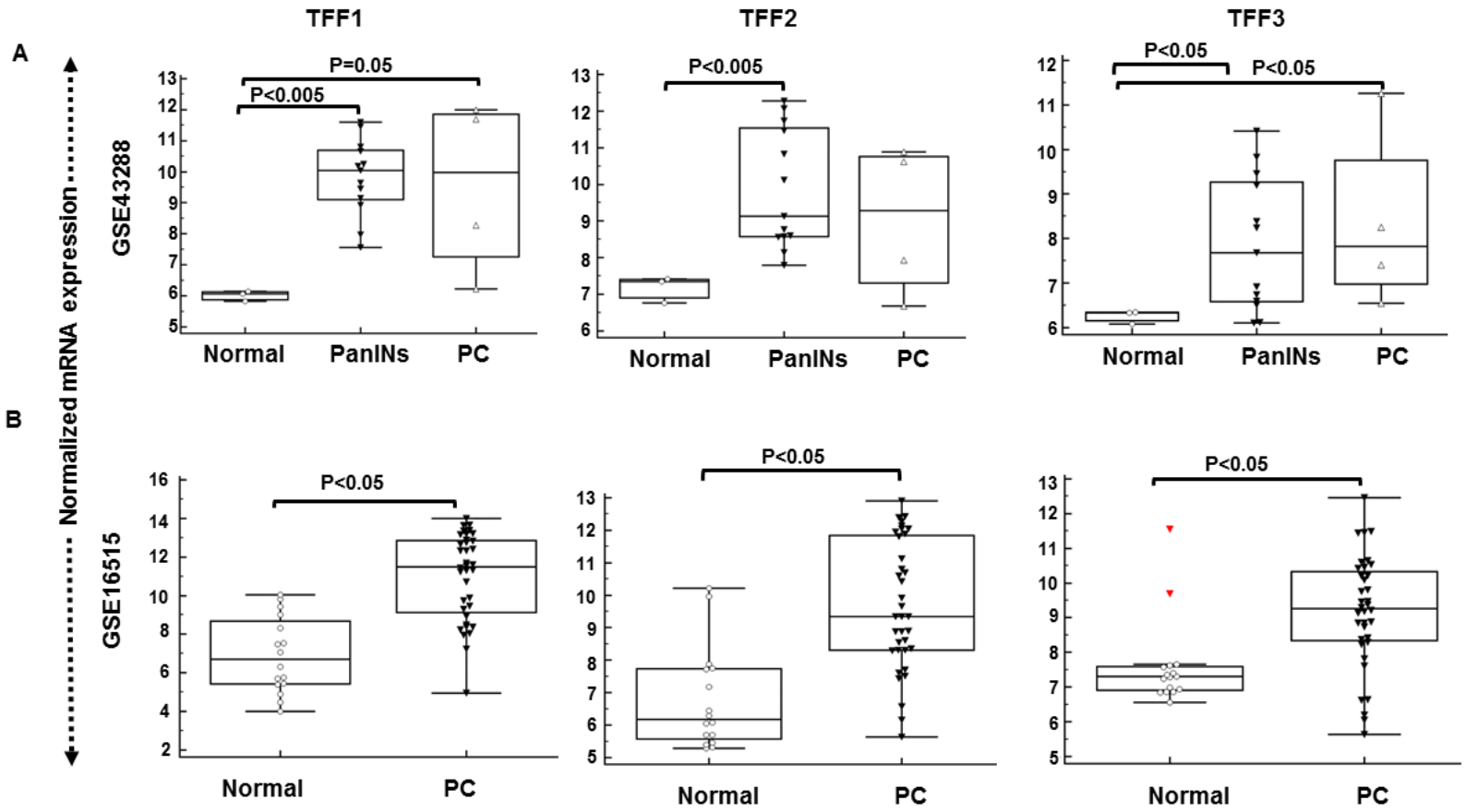

A
B
Normalized mRNA expression
GSE43288
GSE16515
TFF1
TFF2
TFF3
P=0.05
P<0.005
P<0.005
P<0.05
P<0.05
P<0.05
P<0.05
Normal
PanINs
PC

Figure 1: Differential expression of TFF1, TFF2, and TFF3 in PC genomic datasets. (A) Representative box and whisker plots depict the comparison of normalized expression of TFF1-3 mRNA in GSE43288 dataset across normal pancreas (n=3), PC precursor lesions i.e. PanIN (n=13) and pancreatic tumor tissues (n=3). (B) Representative box and whisker plots comparing normalized expression of TFF1-3 mRNA in GSE16515 dataset across normal pancreas (n=16), and pancreatic tumor tissues (n=36). The interquartile range (IQR) for TFF1-3 expression is presented by box and whisker plot (horizontal line represents the 25th percentile, median and 75th percentile and whisker represents 5th and 95th percentile). Publicly available datasets were obtained from Gene Expression Omnibus (GEO, http://www.ncbi.nlm.nih.gov/geo/). Significantly elevated expression of all the TFFs were observed in pancreatic tumors. Further, the elevation in TFFs was found to be significantly higher across precursor lesions of PC. P values were determined using the Mann-Whitney test (two-tailed).

Figure 2

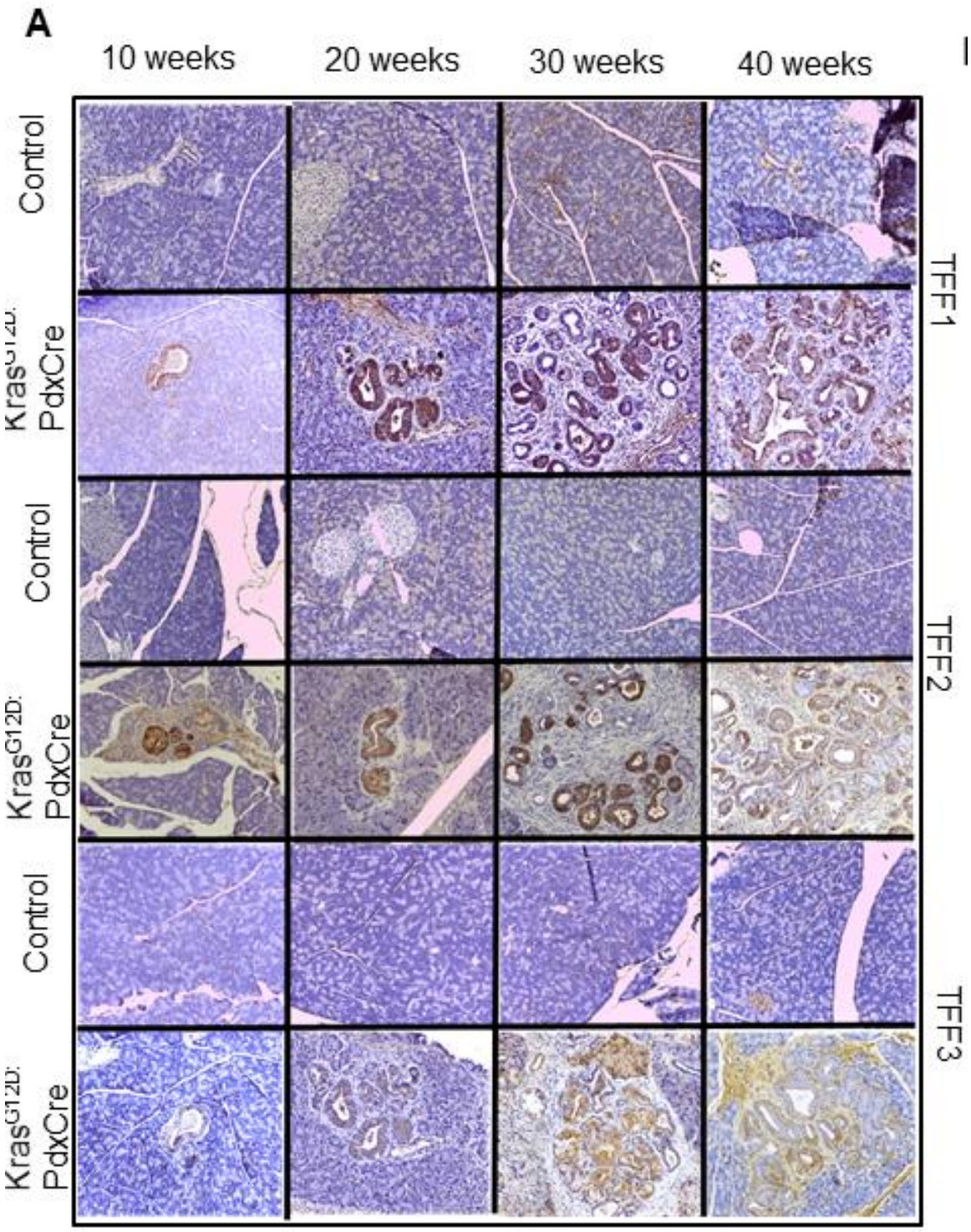

A
10 weeks
20 weeks
30 weeks
40 weeks
Control
KrasG12D; PdxCre
Control
KrasG12D; PdxCre
Control
KrasG12D; PdxCre
TFF1
TFF2
TFF3

Figure 2: Differential expression of TFF1-3 in tissues from precursor lesions, PC tissues from spontaneous PC mouse models. (A) Immunohistochemical analysis of all TFF (TFF1, TFF2, & TFF3) protein expression levels throughout the progression of PC in the spontaneous KrasG12D mouse model, from 10 weeks to 40 weeks. Expression of TFF1, TFF2, and TFF3 was progressively increased from 10 - 30 weeks Pictures are at 10X magnification.

Figure 2B

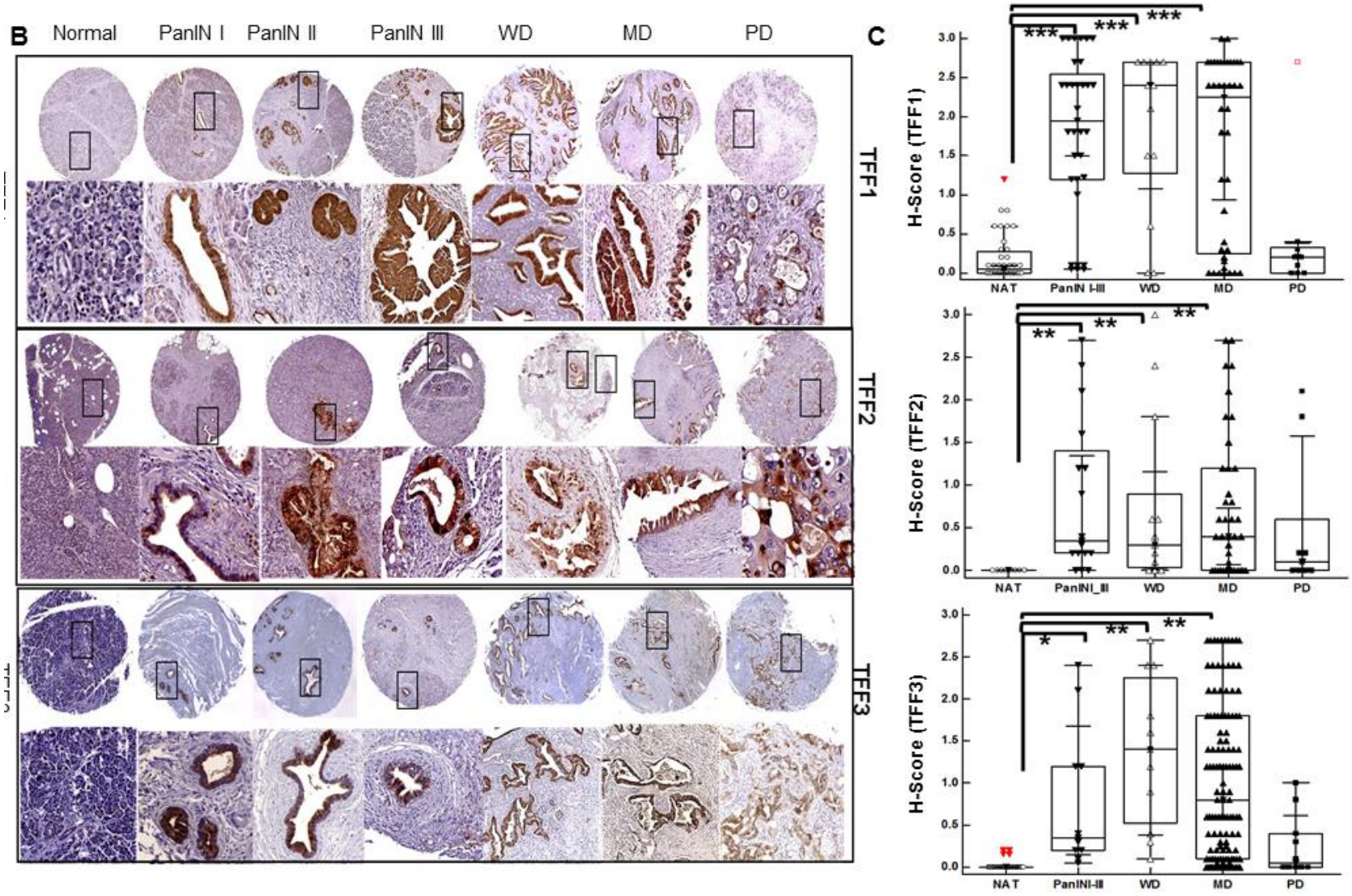

B
Normal
PanIN I
PanIN II
PanIN III
WD
MD
PD
TFF1
TFF2
TFF3
C
H-Score (TFF1)
H-Score (TFF2)
H-Score (TFF3)
NAT
PanIN I-III
PanINI_III
PanINI-III
WD
MD
PD

Figure 2: **Differential expression of TFF1-3 in tissues from precursor lesions, PC tissues from human PC.** (**B)** Immunohistochemical analysis for individual TFF were performed on pancreatic tissue microarrays (TMAs) containing normal pancreatic tissue adjacent to tumor (NAT), pancreatic cancer precursor lesions (PanIN I, II and III), well-differentiated (WD), moderately-differentiated (MD) and poorly-differentiated (PD) pancreatic tumor tissues. No expression of TFF1, TFF2, and TFF3 were observed in normal pancreatic tissues while elevated expression was observed in the ductal compartment across a spectrum of precursor lesions as well as various stages of pancreatic tumor differentiation. Upper panel original magnification (2X) and lower panel magnification (20X). **(C)** Box and whisker plot representing quantitative H-score for TFFs expression across NAT, PanINs and pancreatic tumor tissues. Significant overexpression of TFFs was observed in pancreatic tumor tissues in comparison to normal pancreas. The interquartile range (IQR) for TFFs expression is presented by box and whisker plot (horizontal line represents the 25th percentile, median and 75th percentile and whisker represents 5th and 95th percentile). ***P<0.0005, **P<0.005, *P<0.05.

Figure 3

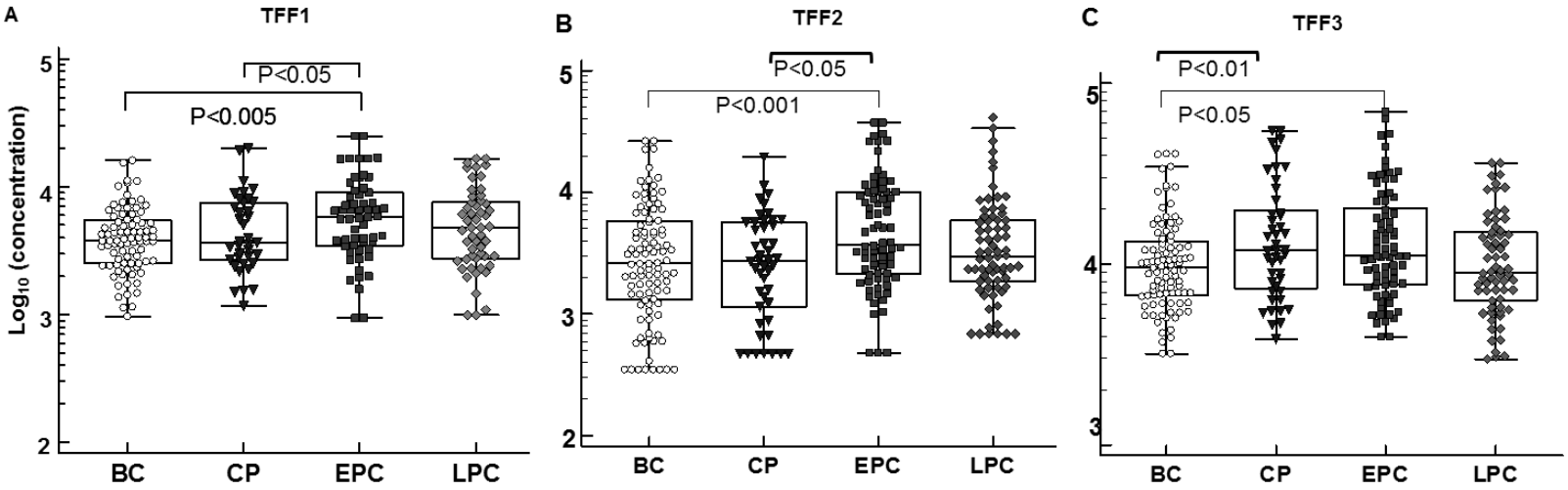

A
TFF1
P<0.05
P<0.005
Log10 (concentration)
5
4
3
2
BC
CP
EPC
LPC
B
TFF2
P<0.05
P<0.001
5
4
3
2
BC
CP
EPC
LPC
C
TFF3
P<0.01
P<0.05
5
4
3
BC
CP
EPC
LPC

Figure 3: Higher levels of TFF1-3 are present in circulation during the early stages of PC. To evaluate diagnostic significance of TFFs, their levels were quantified in serum from various control group (benign controls (BC, N=107) and chronic pancreatitis (CP, N=47) along with early (EPC (Stage 1 and 2, N= 80) and late stage PC cases (LPC, stage 3 and 4, N=73) using duoset sandwich ELISA assay following manufacturer instructions (R&D). (**A-C)** Box and whisker plots showing serum levels of TFF1, TFF2 and TFF3 for benign control (BC), chronic pancreatitis (CP), early-stage PC (EPC, stage I and II) and late stage PC (LPC). The plot shows a significant increase in serum level of TFF1, TFF2, and TFF3 in EPC group as compared to BC. Box and whisker limits represent the fifth and 95th percentiles; the box limits represent IQR where the horizontal lines represent 25^{th}, median and 75^{th} percentile the median concentration of each group. P values are shown above the plots. The P-values were determined by the ANOVA t-test. BC, benign control group; CP, chronic pancreatitis; EPC, early-stage pancreatic cancer (stage 1 and 2); LPC, late-stage pancreatic cancer (stage 3 and 4).

Figure 4

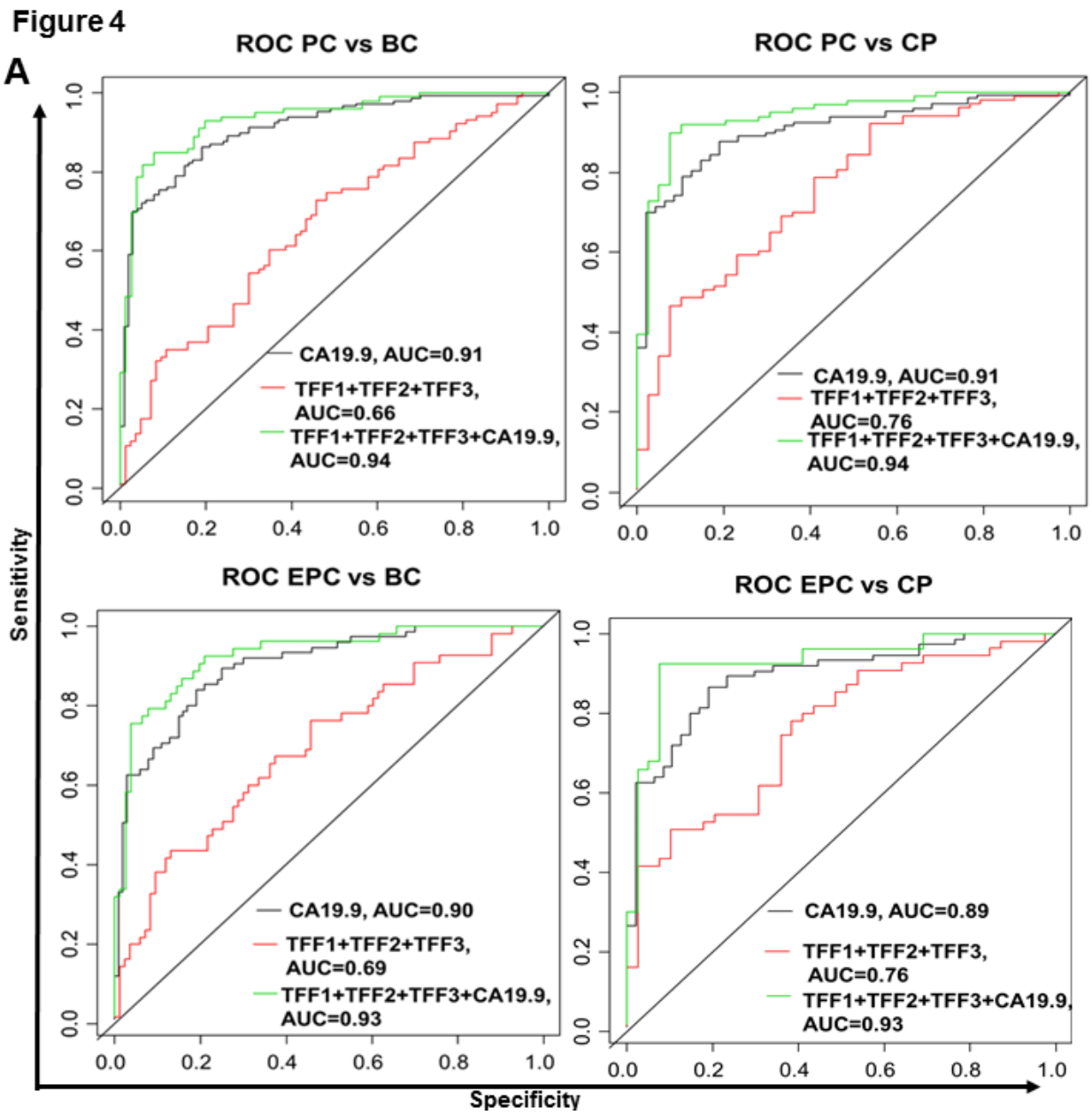

A
ROC PC vs BC
CA19.9, AUC=0.91
TFF1+TFF2+TFF3, AUC=0.66
TFF1+TFF2+TFF3+CA19.9, AUC=0.94
ROC PC vs CP
CA19.9, AUC=0.91
TFF1+TFF2+TFF3, AUC=0.76
TFF1+TFF2+TFF3+CA19.9, AUC=0.94
ROC EPC vs BC
CA19.9, AUC=0.90
TFF1+TFF2+TFF3, AUC=0.69
TFF1+TFF2+TFF3+CA19.9, AUC=0.93
ROC EPC vs CP
CA19.9, AUC=0.89
TFF1+TFF2+TFF3, AUC=0.76
TFF1+TFF2+TFF3+CA19.9, AUC=0.93
Sensitivity
Specificity
0.0
0.2
0.4
0.6
0.8
1.0

Figure 4: Evaluation of the diagnostic significance of TFFs in combination with CA19.9 (A) To evaluate diagnostic significance of the TFFs in combination with CA19.9 for the training set, (OC curves and AUC analyses was carried out for TFF1, TFF2, TFF3 and is presented to distinguish different control groups from PC. The value of the area under curve (AUC) is represented in each box. BC, benign control group; CP, chronic pancreatitis; EPC, early-stage pancreatic cancer (stage 1 and 2); PC, pancreatic cancer. The ROC curve shows comparable performance among all individual TFFs to distinguish different control groups from PC. ROC curves and AUC values for the combination of TFF1-3 and CA19-9 to distinguish different control groups from PC.

Figure 4

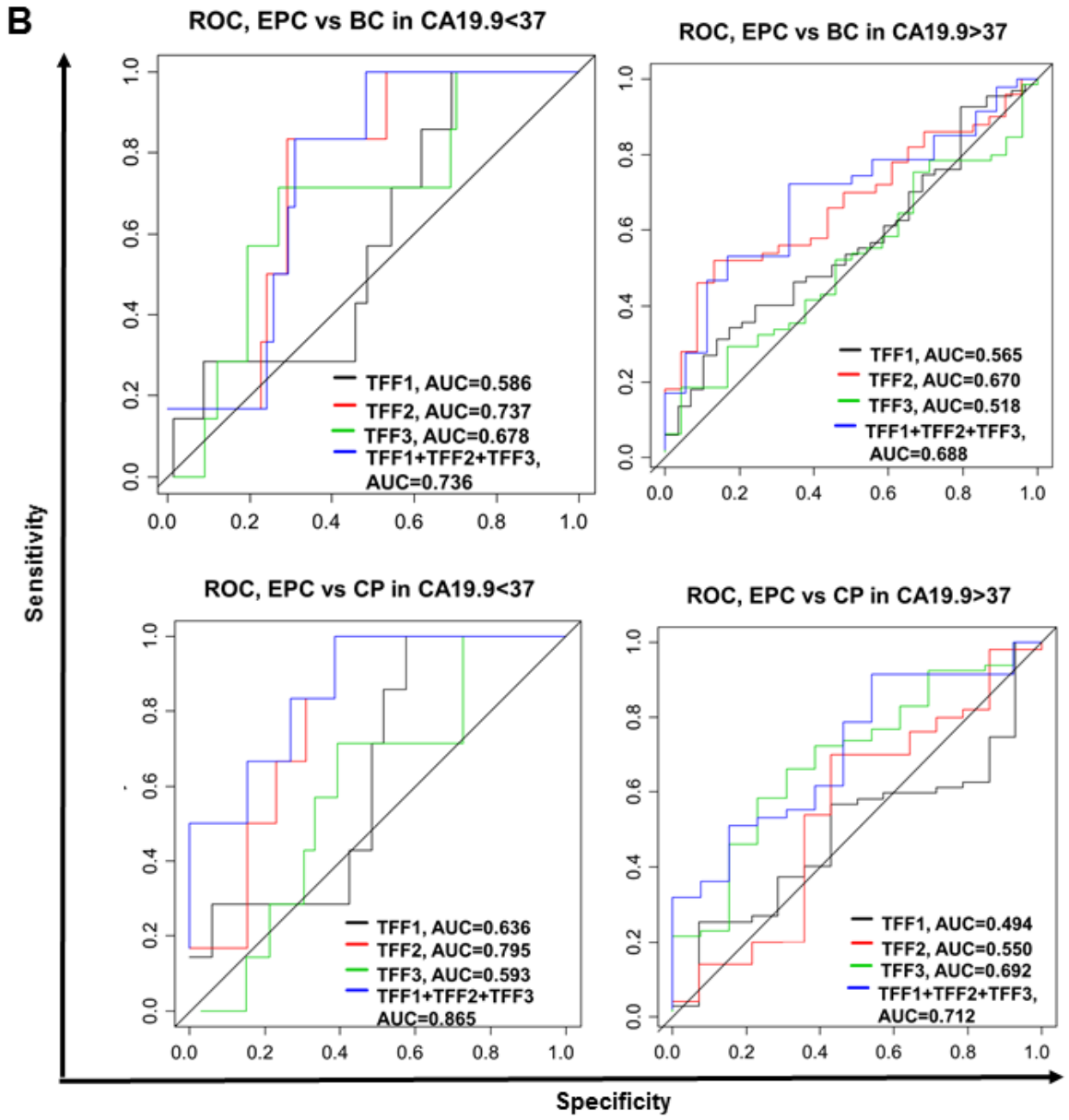

B
ROC, EPC vs BC in CA19.9<37
TFF1, AUC=0.586
TFF2, AUC=0.737
TFF3, AUC=0.678
TFF1+TFF2+TFF3, AUC=0.736
ROC, EPC vs BC in CA19.9>37
TFF1, AUC=0.565
TFF2, AUC=0.670
TFF3, AUC=0.518
TFF1+TFF2+TFF3, AUC=0.688
ROC, EPC vs CP in CA19.9<37
TFF1, AUC=0.636
TFF2, AUC=0.795
TFF3, AUC=0.593
TFF1+TFF2+TFF3 AUC=0.865
ROC, EPC vs CP in CA19.9>37
TFF1, AUC=0.494
TFF2, AUC=0.550
TFF3, AUC=0.692
TFF1+TFF2+TFF3, AUC=0.712
Sensitivity
Specificity

Figure 4: (B). Evaluation of the diagnostic performance of TFF1-3 in CA19.9 low (<37U/ml) and high (≥37 U/ml) groups. ROC curves and AUC values for the combination of TFF1-3 in low CA19.9 (<37 U/ml) and high CA19.9 (>37 U/ml) groups, to distinguish different control groups from PC. The value of the area under curve (AUC) is represented in each box. The ROC curve showed that combination of TFF1-3 retain their diagnostic potential in the setting of low CA19.9, to differentiate between PC and control groups. The AUC is represented in each box. BC, benign control group; CP, chronic pancreatitis; EPC, early-stage pancreatic cancer (stage 1 and 2); PC, pancreatic cancer. The ROC curve showed that combination of TFFs and CA19.9 has better diagnostic accuracy over CA19.9 to distinguish different control groups from PC as well as EPC.

Table 1: **Biomarker performance of individual TFF in PC sample set**

Marker	Comparison Groups	Optimal cutpoint (ln log scale)	Optimal Sensitivity	Optimal Specificity	p value	Total cases	Total controls	AUC	PPV	NPV	Accuracy
TFF1	PC vs. BC	>5.4455	0.717	0.462	0.0078	152	104	0.610	0.659	0.522	0.609
	PC vs. CP	>5.0880	0.842	0.362	0.0035	152	47	0.622	0.81	0.415	0.729
	EPC vs BC	>6.1546	0.474	0.731	0.0078	78	104	0.640	0.569	0.65	0.621
	EPC vs CP	>6.5402	0.385	0.872	0.0081	78	47	0.654	0.833	0.461	0.568
	LPC vs BC	>5.2183	0.768	0.375	0.1751	69	104	0.558	0.449	0.709	0.532
	LPC vs CP	>4.9519	0.855	0.319	0.0614	69	47	0.569	0.648	0.6	0.638
TFF2	PC vs. BC	>8.6181	0.518	0.772	0.0012	114	92	0.639	0.738	0.564	0.631
	PC vs. CP	>8.2287	0.675	0.525	0.1840	114	40	0.576	0.802	0.362	0.636
	EPC vs BC	>8.6181	0.569	0.772	0.0013	58	92	0.665	0.611	0.74	0.693
	EPC vs CP	>8.2387	0.707	0.525	0.0931	58	40	0.607	0.683	0.553	0.633
	LPC vs BC	>8.6181	0.569	0.772	0.0617	52	92	0.594	0.523	0.71	0.653
	LPC vs CP	>8.2287	0.615	0.525	0.7040	52	40	0.527	0.628	0.512	0.576
TFF3	PC vs. BC	>9.2837	0.527	0.643	0.0642	146	98	0.575	0.688	0.477	0.574
	PC vs. CP	<9.1920	0.459	0.652	0.1692	146	46	0.551	0.805	0.273	0.5
	EPC vs BC	>9.2984	0.558	0.643	0.0931	77	98	0.609	0.546	0.643	0.6
	EPC vs CP	<9.3286	0.519	0.565	0.5177	77	46	0.520	0.661	0.406	0.528
	LPC vs BC	>9.4633	0.375	0.745	0.7995	64	98	0.517	0.49	0.646	0.599
	LPC vs CP	<9.1685	0.547	0.652	0.0325	64	46	0.605	0.68	0.5	0.582

BC= Benign Control, CP= Chronic Pancreatitis, PC= Pancreatic Cancer, EPC= Early Stage of Pancreatic Cancer, LPC= Late Stage of Pancreatic Cancer

Table 2: Biomarker Performance of combination of TFFs and CA19.9 in exploratory cohort				
Assay Group	**Prediction Models**	**Optimal Cutpoint**		**90 % SP**
		SN	**SP**	**SN**
PC vs BC	CA19.9	0.864	0.810	0.755
	TFF1+TFF2+TFF3	0.728	0.542	0.330
	TFF1+TFF2+TFF3+CA19.9	0.848	0.921	0.848
PC vs CP	CA19.9	0.878	0.809	0.742
	TFF1+TFF2+TFF3	0.466	0.923	0.466
	TFF1+TFF2+TFF3+CA19.9	0.899	0.923	0.899
EPC vs BC	CA19.9	0.840	0.810	0.693
	TFF1+TFF2+TFF3	0.764	0.542	0.382
	TFF1+TFF2+TFF3+CA19.9	0.755	0.961	0.792
EPC vs CP	CA19.9	0.867	0.809	0.667
	TFF1+TFF2+TFF3	0.509	0.897	0.436
	TFF1+TFF2+TFF3+CA19.9	0.925	0.923	0.925
LPC vs BC	CA19.9	0.821	0.950	0.821
	TFF1+TFF2+TFF3	0.773	0.434	0.205
	TFF1+TFF2+TFF3+CA19.9	0.857	0.974	0.857
LPC vs CP	CA19.9	0.806	0.957	0.821
	TFF1+TFF2+TFF3	0.773	0.641	0.386
	TFF1+TFF2+TFF3+CA19.9	0.881	0.949	0.881

BC= Benign Control, CP= Chronic Pancreatitis, PC= Pancreatic Cancer,
EPC= Early Stage of Pancreatic Cancer, LPC= Late Stage of Pancreatic Cancer,
SN= Sensitivity, SP=Specificity

Supplementary Figure 1

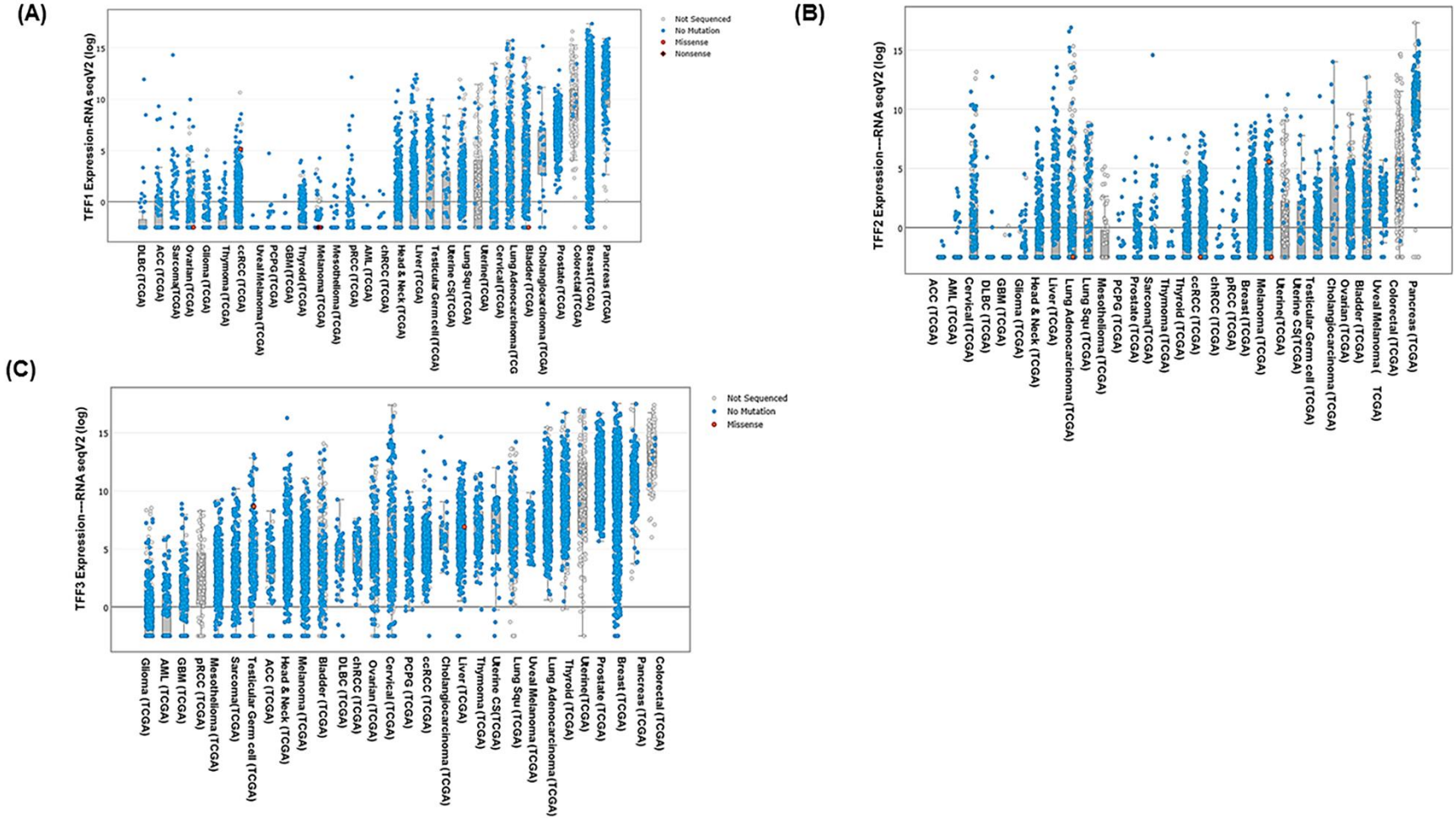


Figure S1: Expression of TFF1-3 in the publicly available cBioPortal database http://www.cbioportal.org/index.do) consisting of 169 studies from 30 different tumor types. (A) Analysis of TFF1 expression level from the cBioPortal database is indicating the highest expression in PC followed by breast and colorectal cancer among all 30 different tumors data found in TCGA. **(B)** Similarly, TFF2 is indicated as top expressed in PC followed by colorectal and uveal melanoma among all 30 different tumors found in TCGA. **(C)** TFF3 showed the highest expression in colorectal cancer followed by PC and breast cancer among all 30 different tumors found in TCGA. Hierarchical organization of different tumors are based on their higher median value (from right to left) where each spot represents a single study. White, blue and red dots denote genes which are not sequenced, gene which are sequenced and missense mutations respectively.

Supplementary Figure S2

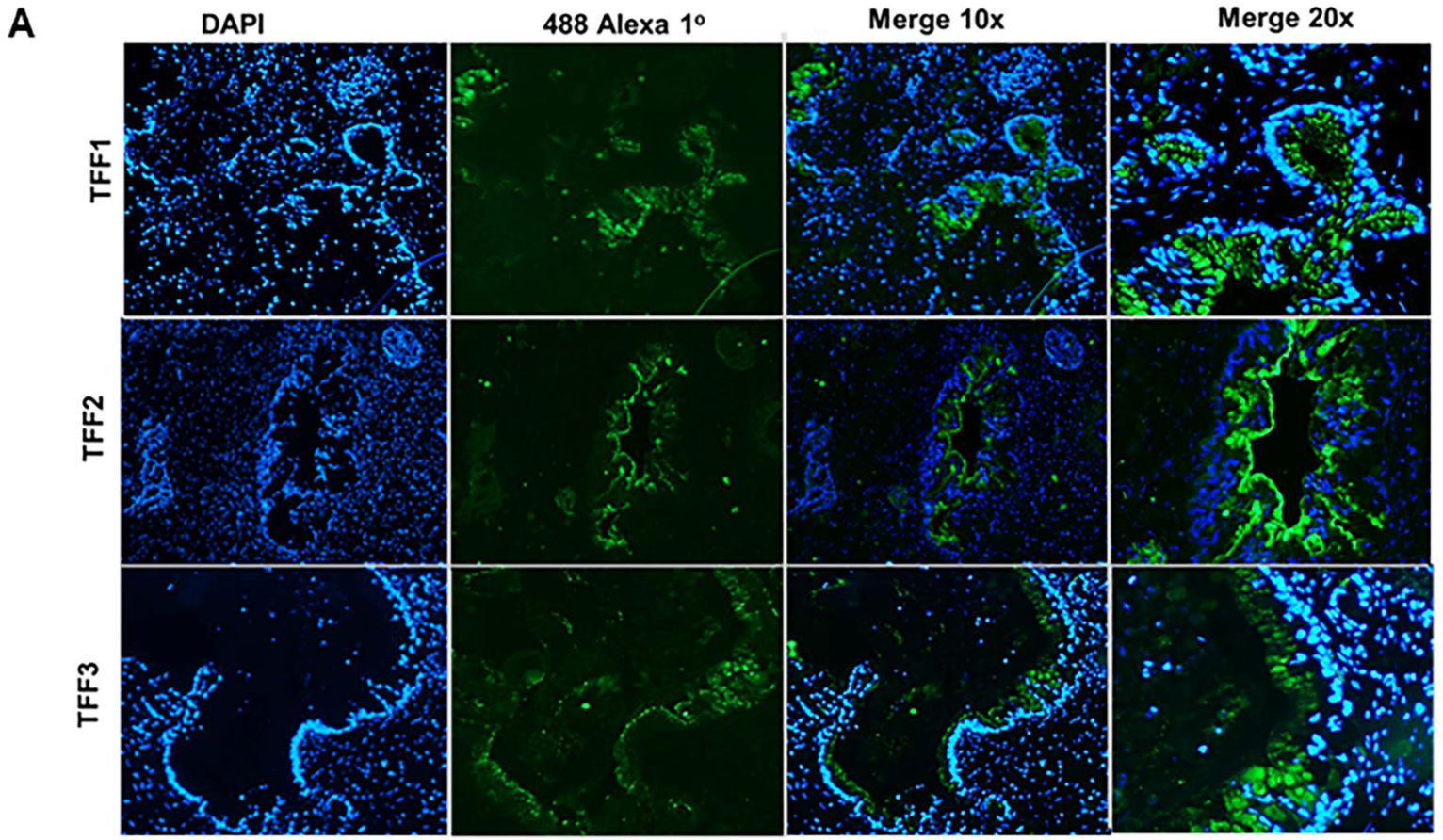


Figure S2: Expression of TFF1-3 in PC tissues (A) Representative confocal image of TFF1-3 expression in PC ducts. Green fluorescence represents Alexa-488 conjugated primary antibodies for individual TFFs. DAPI (blue) is a nuclear marker. The merged picture has been shown in 10X and 20X magnification.

Figure S2

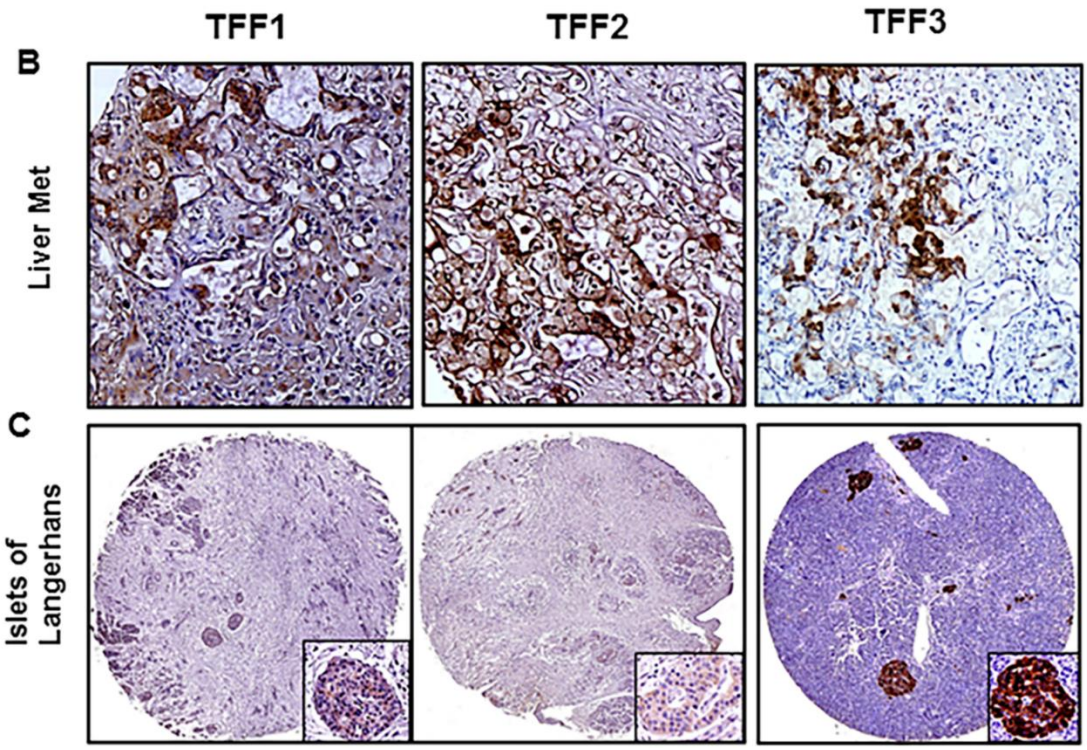
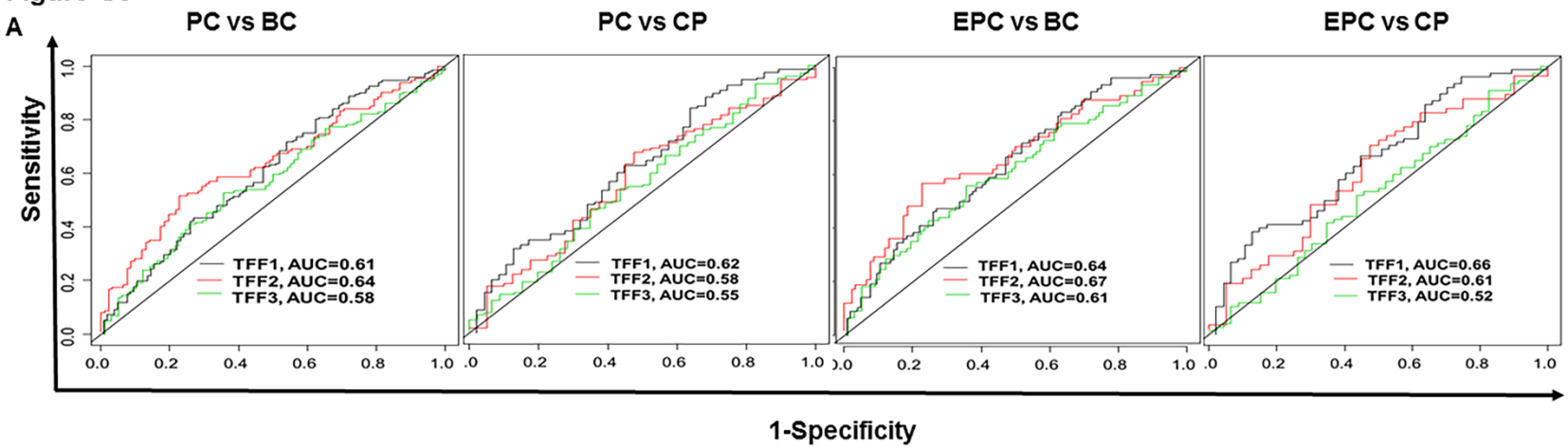


Figure S2: Expression of TFF1-3 in metastatic tissues and islet of Langerhans. (**B**) IHC analysis showed strong positive staining of TFF 1-3 in PC metastatic tissues obtained from liver (n=5). (**C**) IHC analysis showed almost negative staining of TFF1 and TFF2 and strong positive staining for TFF3 in islet of Langerhans.

Figure S3


A
PC vs BC
PC vs CP
EPC vs BC
EPC vs CP
Sensitivity
1-Specificity
TFF1, AUC=0.61
TFF2, AUC=0.64
TFF3, AUC=0.58
TFF1, AUC=0.62
TFF2, AUC=0.58
TFF3, AUC=0.55
TFF1, AUC=0.64
TFF2, AUC=0.67
TFF3, AUC=0.61
TFF1, AUC=0.66
TFF2, AUC=0.61
TFF3, AUC=0.52

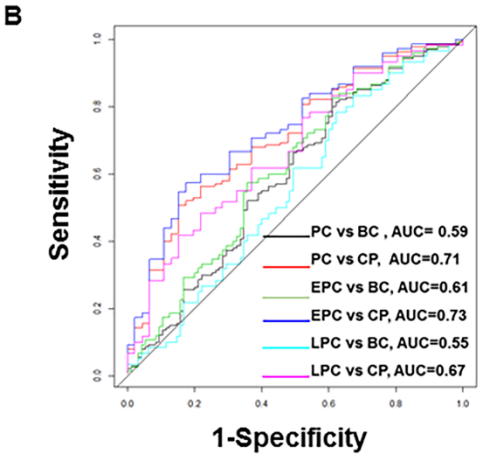

B
Sensitivity
1-Specificity
PC vs BC , AUC= 0.59
PC vs CP, AUC=0.71
EPC vs BC, AUC=0.61
EPC vs CP, AUC=0.73
LPC vs BC, AUC=0.55
LPC vs CP, AUC=0.67

Figure S3: Diagnostic performance of individual TFF and ratio of TFF1/TFF3 in the training cohort. (**A**) ROC curves and AUC values for individual TFFs to distinguish patients in different groups between BC, CP, PC and EPC. (**B**) ROC curves and AUC values for TFF1/TFF3 ratio to distinguish patients in different groups between BC, CP, PC and EPC. The value of AUC is represented in each box. BC, benign control group; CP, chronic pancreatitis; EPC, early-stage pancreatic cancer (stage 1 and 2); PC, pancreatic cancer.

Supplementary Figure S4

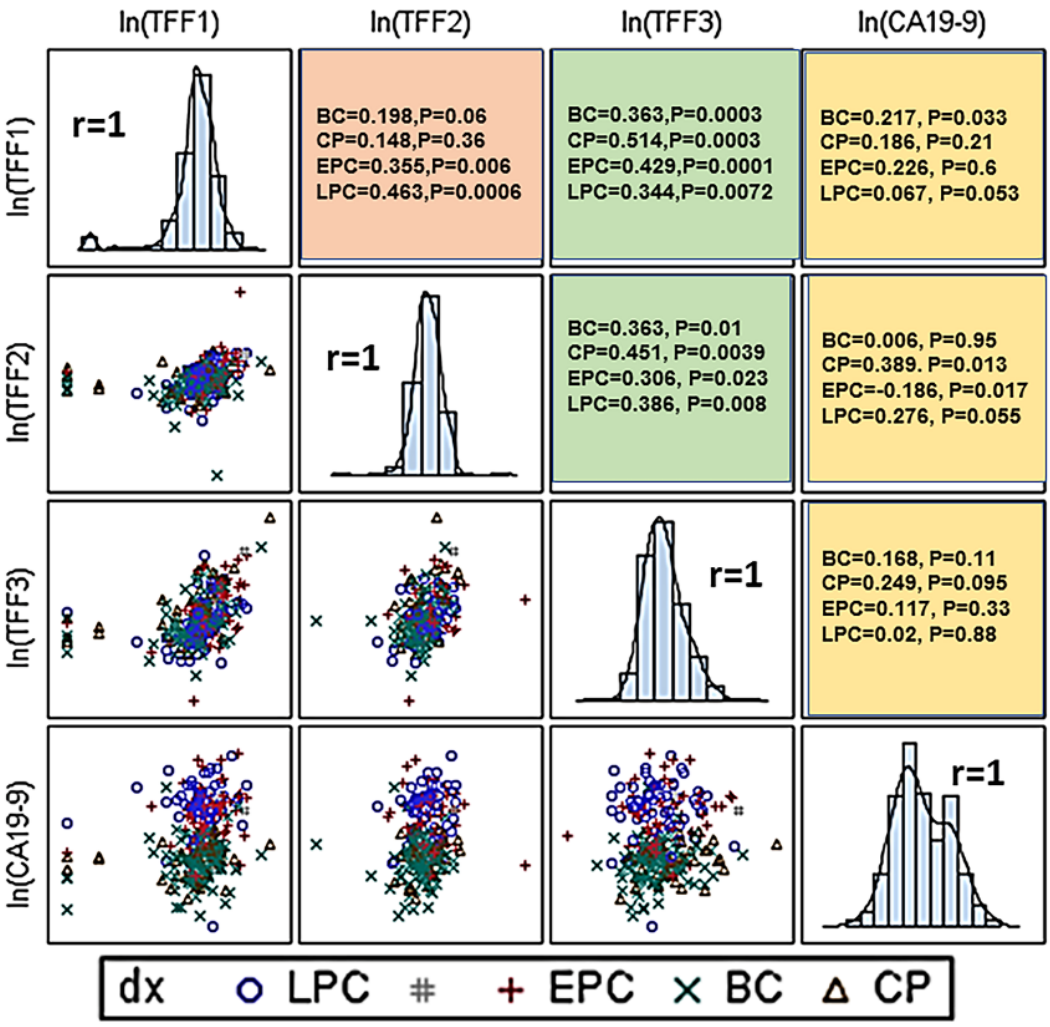


Figure S4: Correlation between TFF1, TFF2, TFF3 and CA19.9 in all patients and within each group. Pearson correlation analysis revealed fair to moderate positive correlation among any two individual molecular group in BC, CP, EPC and LPC pathological groups. Scatter plot is showing overall distribution with corresponding coefficient (r) and P value. In case of CA19.9 and TFF1/2/3, a positive correlation was observed from the analysis of BC and CP group. However, moderate correlation between CA19.9 and TFF1/2/3 in EPC and LPC groups was observed. Each point represents an individual sample. Significance was determined by Spearman's correlation test.

Supplementary Figure S5

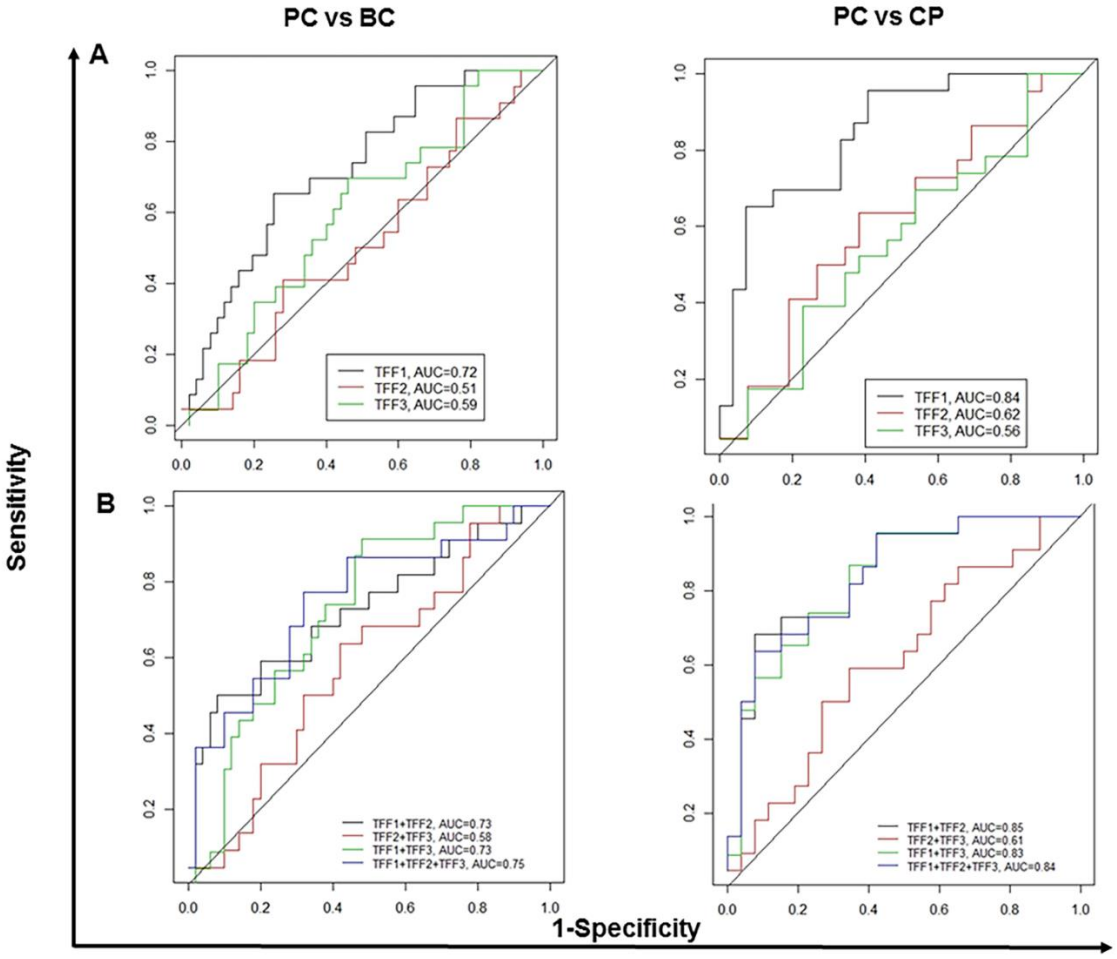


Figure S5: Overall diagnostic performance of individual TFF, combination of TFF1-3 and combination of TFFs and CA19.9 biomarkers in the training cohort. (A-B) Forest plots are depicting AUC values with interquartile range for individual TFFs as well as combination of TFFs with CA19.9 in different groups between PC vs BC, PC vs CP, EPC vs BC and EPC vs CP.

Supplementary Figure S6

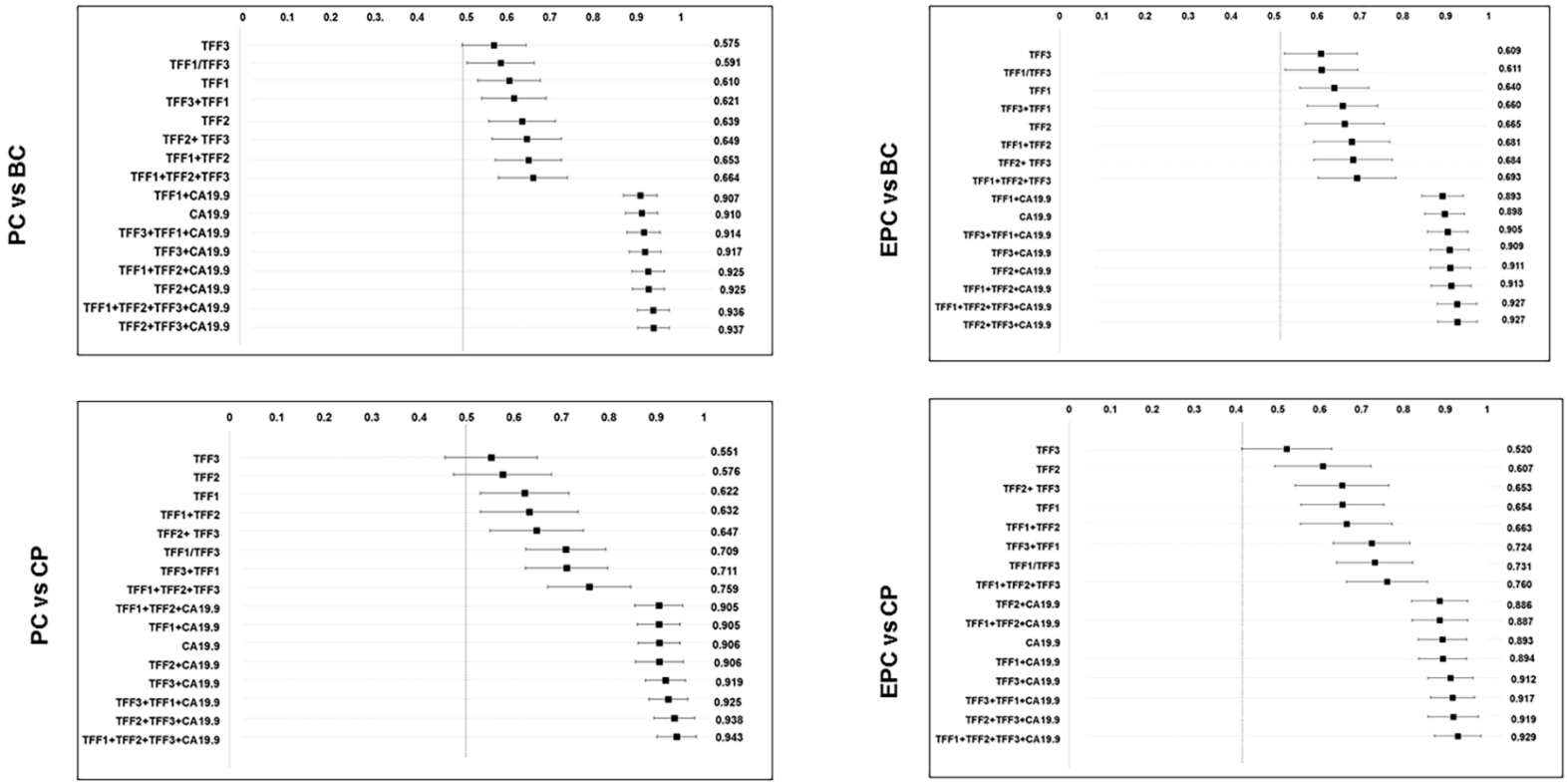


Figure S6: Diagnostic performance of individual TFF and the combination of TFF1-3 biomarkers in the validation cohort. (**A**) ROC curves and AUC values for individual TFFs to distinguish patients in different groups between PC (n=23) vs BC (n=27) and PC vs CP (n=27) (**B**) ROC curves and AUC values for combination of TFF1, TFF2, and TFF3 to distinguish patients in different groups between PC and BC, PC and CP. The ROC curve showed that combination of TFFs has demonstrated improved diagnostic potential as compared to individual TFF to segregate PC group vs. control groups (BC and CP).

Supporting Documents (Supplementary Tables)

Table S1A: Patient characteristics by group- training set					
Variables	**Groups**	**BC (n=107)**	**CP (n=47)**	**PC (n=158)**	**p-value**
Age	<=64	71 (66%)	34 (72%)	61 (39%)	<0.0001
	> 64	36 (34%)	13 (28%)	97 (61%)	
Sex	F	62 (57%)	22 (47%)	77 (49%)	0.26
	M	45 (43%)	25 (53%)	81 (51%)	
Bilirubin	<=1.2	9 (64%)	3 (100%)	40 (45%)	*
	> 1.2	5 (36%)	0	49 (55%)	
Race	AA/Asian	9 (8%)	6 (13%)	9 (6%)	0.26
	Caucasian	98 (92%)	41 (78%)	149 (94%)	
Alcohol hx	Current	4 (4%)	3 (6%)	16 (10%)	0.021
	Ex-drinker	7 (7%)	7 (15%)	20 (13%)	
	Never	3 (3%)	4 (9%)	16 (10%)	
	Unknown	93 (87%)	33 (70%)	106 (67%)	

Table S1B: Patient characteristics by group - validation set					
Variables	**Groups**	**BC (n=9)**	**CP (n=29)**	**PC (n=25)**	**p-value**
Age	<=64	7 (88%)	19 (66%)	10 (40%)	0.082
	> 64	1 (12%)	10 (34%)	15 (60%)	
Sex	F	5 (62%)	13 (45%)	9 (36%)	0.59
	M	3 (38%)	16 (55%)	16 (64%)	
Race	AA/Asian	0	2 (7%)	1 (4%)	0.85
	Caucasian	8 (100%)	27 (93%)	24 (96%)	
Alcohol Status	Current	3 (38%)	11 (37%)	8 (32%)	0.22
	Ex-drinker	0	9 (31%)	7 (28%)	
	Never	5 (62%)	9 (31%)	9 (36%)	
	Unknown	1	0	1 (4%)	

Table S2: TFFs by patient characteristics – log scale													
Characteristics	**Groups**	**TFF1**				**TFF2**				**TFF3**			
		N	**Mean**	**SD**	**p-value**	**N**	**Mean**	**SD**	**p-value**	**N**	**Mean**	**SD**	**p-value**
Age	<=64	161	5.14	2.01	<0.0001	134	8.23	0.9	0.0004	154	9.2	0.75	0.0014
	> 64	142	5.97	1.52		112	8.58	0.67		136	9.46	0.63	
Sex	F	157	5.43	1.93	0.31	125	8.4	0.81	0.83	148	9.32	0.67	0.81
	M	146	5.64	1.74		121	8.38	0.84		142	9.34	0.75	
Bilirubin	<=1.2	52	6.04	1.15	0.58	34	8.57	1.07	0.56	46	9.37	0.72	0.42
	> 1.2	52	5.89	1.56		39	8.69	0.63		47	9.49	0.73	
Race	AA/Asian	24	5.71	1.52	0.62	21	8.25	0.8	0.41	22	9.46	1	0.52
	Caucasian	279	5.51	1.87		225	8.4	0.82		268	9.31	0.68	
Alcohol status	Current	23	5.62	1.16	0.2	11	8.28	0.63	0.068	19	9.12	0.78	0.59
	Ex-drinker	33	5.96	1.19		25	8.6	1.16		30	9.28	0.66	
	Never	23	6.05	0.94		14	8.85	0.77		21	9.39	0.83	
	Unknown	224	5.41	2.02		196	8.34	0.77		220	9.34	0.7	
Stage	1 & 2	78	6.02	1.41	0.13	58	8.67	0.91	0.15	77	9.44	0.74	0.063
	3 & 4	69	5.66	1.46		52	8.43	0.77		64	9.21	0.67	

Table S3:Sensitivity and Specificity of combination of TFF in PC sample set			
Comparison Group	**Combination of TFF**	**Optimal Cutpoint**	
		Sensitivity	**Specificity**
	ln(TFF1)+ln(TFF2)	0.752	0.484
	ln(TFF2)+ ln(TFF3)	0.471	0.786
PC vs. BC	ln(TFF3)+ln(TFF1)	0.821	0.389
	ln(TFF1)+ln(TFF2)	0.310	0.900
	ln(TFF2)+ ln(TFF3)	0.337	0.923
PC vs CP	ln(TFF3)+ln(TFF1)	0.779	0.565
	ln(TFF1)+ln(TFF2)	0.483	0.835
	ln(TFF2)+ ln(TFF3)	0.527	0.786
EPC vs BC	ln(TFF3)+ln(TFF1)	0.827	0.453
	ln(TFF1)+ln(TFF2)	0.379	0.900
	ln(TFF2)+ ln(TFF3)	0.782	0.487
EPC vs CP	ln(TFF3)+ln(TFF1)	0.613	0.739
	ln(TFF1)+ln(TFF2)	0.824	0.396
	ln(TFF2)+ ln(TFF3)	0.444	0.774
LPC vs BC	ln(TFF3)+ln(TFF1)	0.800	0.358
	ln(TFF1)+ln(TFF2)	0.941	0.250
	ln(TFF2)+ ln(TFF3)	0.667	0.667
LPC vs CP	ln(TFF3)+ln(TFF1)	0.617	0.761

Table S4: Sensitivity and Specificity of combination of TFFs and CA19.9 in PC sample set

Assay Group	Combination of TFF	Based on Optimal Cutpoint		Based on specificity of 0.9	Assay Group	Combination of TFF	Based on Optimal Cutpoint		Based on specificity of 0.9
		SN	SP	SN			SN	SP	SN
PC vs. BC	ln(TFF1)+ln(CA19.9)	0.725	0.948	0.754	EPC vs CP	ln(TFF1)+ln(CA19.9)	0.892	0.787	0.676
	ln(TFF2)+ln(CA19.9)	0.807	0.918	0.807		ln(TFF2)+ln(CA19.9)	0.893	0.750	0.661
	ln(TFF3)+ln(CA19.9)	0.752	0.945	0.788		ln(TFF3)+ln(CA19.9)	0.847	0.913	0.847
	ln(TFF1)+ln(TFF2)+ln(CA19.9)	0.852	0.881	0.815		ln(TFF1)+ln(TFF2)+ln(CA19.9)	0.893	0.750	0.661
	ln(TFF2)+ln(TFF3)+ln(CA19.9)	0.850	0.922	0.850		ln(TFF2)+ln(TFF3)+ln(CA19.9)	0.943	0.846	0.736
	ln(TFF3)+ln(TFF1)+ln(CA19.9)	0.742	0.966	0.780		ln(TFF3)+ln(TFF1)+ln(CA19.9)	0.887	0.891	0.859
PC vs CP	ln(TFF1)+ln(CA19.9)	0.838	0.851	0.746	LPC vs BC	ln(TFF1)+ln(CA19.9)	0.825	0.948	0.825
	ln(TFF2)+ln(CA19.9)	0.817	0.850	0.761		ln(TFF2)+ln(CA19.9)	0.816	0.965	0.837
	ln(TFF3)+ln(CA19.9)	0.847	0.913	0.847		ln(TFF3)+ln(CA19.9)	0.817	0.978	0.833
	ln(TFF1)+ln(TFF2)+ln(CA19.9)	0.833	0.850	0.759		ln(TFF1)+ln(TFF2)+ln(CA19.9)	0.813	0.964	0.833
	ln(TFF2)+ln(TFF3)+ln(CA19.9)	0.930	0.872	0.810		ln(TFF2)+ln(TFF3)+ln(CA19.9)	0.860	0.948	0.860
	ln(TFF3)+ln(TFF1)+ln(CA19.9)	0.879	0.891	0.841		ln(TFF3)+ln(TFF1)+ln(CA19.9)	0.821	0.977	0.839
EPC vs BC	ln(TFF1)+ln(CA19.9)	0.892	0.753	0.689	LPC vs CP	ln(TFF1)+ln(CA19.9)	0.810	0.957	0.825
	ln(TFF2)+ln(CA19.9)	0.929	0.741	0.732		ln(TFF2)+ln(CA19.9)	0.918	0.850	0.857
	ln(TFF3)+ln(CA19.9)	0.847	0.824	0.722		ln(TFF3)+ln(CA19.9)	0.833	0.957	0.833
	ln(TFF1)+ln(TFF2)+ln(CA19.9)	0.875	0.810	0.750		ln(TFF1)+ln(TFF2)+ln(CA19.9)	0.833	0.925	0.833
	ln(TFF2)+ln(TFF3)+ln(CA19.9)	0.849	0.870	0.792		ln(TFF2)+ln(TFF3)+ln(CA19.9)	0.860	0.949	0.884
	ln(TFF3)+ln(TFF1)+ln(CA19.9)	0.887	0.773	0.732		ln(TFF3)+ln(TFF1)+ln(CA19.9)	0.893	0.891	0.839

PC= Pancreatic Cancer, EPC= Early Stage of Pancreatic Cancer, LPC= Late Stage of Pancreatic Cancer

Table S5:Biomarker Performance of combination of TFFs in PC sample set, CA19.9 <37							
Comparison	**Combination of TFF**	**AUC**	**SE**	**95% CI**		**Based on optimal Cutpoint**	
				Lower	**Upper**	**Sensitivity**	**Specificity**
PC vs. BC	ln(TFF1)	0.566	0.081	0.408	0.724	0.929	0.309
	ln(TFF2)	0.605	0.111	0.388	0.822	0.700	0.661
	ln(TFF3)	0.554	0.086	0.386	0.722	0.400	0.806
	ln(TFF1)+ln(TFF2)+ln(TFF3)	0.666	0.085	0.500	0.832	0.400	0.828
PC vs CP	ln(TFF1)	0.626	0.087	0.456	0.795	0.929	0.424
	ln(TFF2)	0.639	0.123	0.398	0.879	0.800	0.615
	ln(TFF3)	0.529	0.091	0.352	0.707	1.000	0.152
	ln(TFF1)+ln(TFF2)+ln(TFF3)	0.815	0.074	0.671	0.959	1.000	0.577
EPC vs BC	ln(TFF1)	0.586	0.107	0.377	0.795	1.000	0.309
	ln(TFF2)	0.737	0.082	0.576	0.897	0.833	0.710
	ln(TFF3)	0.678	0.106	0.471	0.885	0.714	0.731
	ln(TFF1)+ln(TFF2)+ln(TFF3)	0.736	0.078	0.582	0.889	0.833	0.690
EPC vs CP	ln(TFF1)	0.636	0.107	0.426	0.847	1.000	0.424
	ln(TFF2)	0.795	0.083	0.633	0.957	1.000	0.615
	ln(TFF3)	0.593	0.108	0.382	0.805	0.714	0.606
	ln(TFF1)+ln(TFF2)+ln(TFF3)	0.865	0.077	0.714	1.000	1.000	0.615
LPC vs BC	ln(TFF1)	0.495	0.120	0.260	0.730	0.833	0.309
	ln(TFF2)	0.593	0.224	0.155	1.000	0.500	0.968
	ln(TFF3)	0.616	0.110	0.400	0.832	0.857	0.448
	ln(TFF1)+ln(TFF2)+ln(TFF3)	0.716	0.151	0.419	1.000	0.500	0.948
LPC vs CP	ln(TFF1)	0.571	0.123	0.330	0.811	0.833	0.424
	ln(TFF2)	0.596	0.243	0.120	1.000	0.500	1.000
	ln(TFF3)	0.706	0.100	0.510	0.902	0.857	0.576
	ln(TFF1)+ln(TFF2)+ln(TFF3)	0.904	0.060	0.787	1.000	1.000	0.808

Table S6:Biomarker Performance of combination of TFFs in PC sample set, CA19.9 >37							
Assay Group	**Combination of TFF**	**AUC**	**SE**	**95% CI**		**Based on optimal cutpoint**	
				Lower	**Upper**	**Sensitivity**	**Specificity**
PC vs. BC	ln(TFF1)	0.523	0.058	0.409	0.636	0.211	0.897
	ln(TFF2)	0.668	0.056	0.559	0.778	0.505	0.870
	ln(TFF3)	0.502	0.060	0.384	0.621	0.279	0.833
	ln(TFF1)+ln(TFF2)+ln(TFF3)	0.667	0.065	0.540	0.795	0.685	0.667
PC vs CP	ln(TFF1)	0.548	0.073	0.405	0.691	0.297	0.929
	ln(TFF2)	0.563	0.094	0.379	0.747	0.707	0.571
	ln(TFF3)	0.706	0.076	0.558	0.854	0.672	0.692
	ln(TFF1)+ln(TFF2)+ln(TFF3)	0.728	0.078	0.575	0.880	0.551	0.846
EPC vs BC	ln(TFF1)	0.565	0.063	0.442	0.687	0.313	0.862
	ln(TFF2)	0.670	0.064	0.545	0.796	0.520	0.870
	ln(TFF3)	0.518	0.067	0.388	0.648	0.185	0.958
	ln(TFF1)+ln(TFF2)+ln(TFF3)	0.688	0.070	0.550	0.825	0.723	0.667
EPC vs CP	ln(TFF1)	0.494	0.079	0.340	0.647	0.254	0.929
	ln(TFF2)	0.550	0.097	0.360	0.740	0.700	0.571
	ln(TFF3)	0.692	0.079	0.537	0.848	0.585	0.769
	ln(TFF1)+ln(TFF2)+ln(TFF3)	0.712	0.078	0.559	0.865	0.915	0.462
LPC vs BC	ln(TFF1)	0.548	0.067	0.416	0.679	0.526	0.655
	ln(TFF2)	0.646	0.068	0.513	0.780	0.467	0.870
	ln(TFF3)	0.532	0.069	0.397	0.667	0.226	0.958
	ln(TFF1)+ln(TFF2)+ln(TFF3)	0.630	0.079	0.475	0.785	0.342	0.944
LPC vs CP	ln(TFF1)	0.635	0.080	0.478	0.792	0.737	0.571
	ln(TFF2)	0.587	0.098	0.395	0.779	0.733	0.571
	ln(TFF3)	0.736	0.078	0.582	0.890	0.698	0.692
	ln(TFF1)+ln(TFF2)+ln(TFF3)	0.745	0.087	0.574	0.916	0.868	0.538

Table S7: Validation of Biomarker Performance of individual TFF in Phase II cohort (cutpoints from training set)							
Gene Name	**Comparison Groups**	**Optimal cutpoint (log scale)**	**Sensitivity**	**Specificity**	**PPV**	**NPV**	**Accuracy**
TFF1	PC vs. BC	>5.4455	0.348	1.000	1	0.348	0.516
	PC vs. CP	>5.0880	0.478	0.926	0.846	0.676	0.72
	EPC vs BC	>6.1546	0.333	1.000	1	0.4	0.538
	EPC vs CP	>6.5402	0.444	0.926	0.8	0.714	0.733
TFF2	PC vs. BC	>8.6181	0.409	0.375	0.643	0.188	0.4
	PC vs. CP	>8.2287	0.546	0.615	0.546	0.615	0.583
	EPC vs BC	>8.6181	0.389	0.375	0.583	0.214	0.385
	EPC vs CP	>8.2387	0.500	0.615	0.474	0.64	0.568
TFF3	PC vs. BC	>9.2837	0.348	0.875	0.889	0.318	0.484
	PC vs. CP	<9.1920	0.652	0.231	0.429	0.429	0.429
	EPC vs BC	>9.2984	0.333	0.875	0.857	0.368	0.5
	EPC vs CP	<9.3286	0.667	0.231	0.375	0.5	0.409
ln(TFF1)+ ln(TFF2)+ ln(TFF3)	PC vs. BC	P(PCvBC)>0.53496	0.455	0.625	0.769	0.294	0.5
	PC vs. CP	P(PCvCP)>0.81703	0.636	0.577	0.56	0.652	0.604
	EPC vs BC	P(EPCvBC)>0.35364	0.389	0.250	0.539	0.154	0.346
	EPC vs CP	P(EPCvCP)>0.70126	0.611	0.692	0.579	0.692	0.659

PC= Pancreatic Cancer, EPC= Early Stage of Pancreatic Cancer, LPC= Late Stage of Pancreatic Cancer

Reference

1. Rahib L, Smith BD, Aizenberg R, Rosenzweig AB, Fleshman JM, Matrisian LM. Projecting cancer incidence and deaths to 2030: the unexpected burden of thyroid, liver, and pancreas cancers in the United States. Cancer Res 2014;74:2913-21.

2. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2018. CA Cancer J Clin 2018;68:7-30.

3. Harsha HC, Kandasamy K, Ranganathan P, Rani S, Ramabadran S, Gollapudi S, et al. A compendium of potential biomarkers of pancreatic cancer. PLoS Med 2009;6:e1000046.

4. Taupin D, Podolsky DK. Trefoil factors: initiators of mucosal healing. Nat. Rev. Mol. Cell Biol 2003;4:721-732.

5. Perry JK, Kannan N, Grandison PM, Mitchell MD, Lobie PE. Are trefoil factors oncogenic? Trends Endocrinol. Metab 2008;19:74-81.

6. Stastna M, Van Eyk JE. Secreted proteins as a fundamental source for biomarker discovery. Proteomics 2012;12:722-35.

7. Scara S, Bottoni P, Scatena R. CA 19-9: Biochemical and Clinical Aspects. Adv. Exp. Med Biol 2015;867:247-260.

8. Kaur S, Smith LM, Patel A, Menning M, Watley DC, Malik SS, et al. A Combination of MUC5AC and CA19-9 Improves the Diagnosis of Pancreatic Cancer: A Multicenter Study. Am J Gastroenterol 2017;112:172-183.

9. Gao J, Aksoy BA, Dogrusoz U, Dresdner G, Gross B, Sumer SO, et al. Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal. Sci Signal 2013;6:pl1.

10. Cerami E, Gao J, Dogrusoz U, Gross BE, Sumer SO, Aksoy BA, et al. The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data. Cancer Discov 2012;2:401-4.

11. Hingorani SR, Petricoin EF, Maitra A, Rajapakse V, King C, Jacobetz MA, et al. Preinvasive and invasive ductal pancreatic cancer and its early detection in the mouse. Cancer Cell 2003;4:437-50.

12. Ballehaninna UK, Chamberlain RS. The clinical utility of serum CA 19-9 in the diagnosis, prognosis and management of pancreatic adenocarcinoma: An evidence based appraisal. J Gastrointest Oncol 2012;3:105-19.

13. Collisson EA, Sadanandam A, Olson P, Gibb WJ, Truitt M, Gu S, et al. Subtypes of pancreatic ductal adenocarcinoma and their differing responses to therapy. Nat. Med 2011;17:500-503.

14. Moffitt RA, Marayati R, Flate EL, Volmar KE, Loeza SG, Hoadley KA, et al. Virtual microdissection identifies distinct tumor- and stroma-specific subtypes of pancreatic ductal adenocarcinoma. Nat. Genet 2015;47:1168-1178.

15. Radon TP, Massat NJ, Jones R, Alrawashdeh W, Dumartin L, Ennis D, et al. Identification of a Three-Biomarker Panel in Urine for Early Detection of Pancreatic Adenocarcinoma. Clin Cancer Res 2015;21:3512-3521.

16. Vocka M, Langer D, Petrtyl J, Vockova P, Hanus T, Kalousova M, et al. Trefoil factor family (TFF) proteins as potential serum biomarkers in patients with metastatic colorectal cancer. Neoplasma 2015;62:470-7.

17. Li Q, Wang K, Su C, Fang J. Serum Trefoil Factor 3 as a Protein Biomarker for the Diagnosis of Colorectal Cancer. Technol Cancer Res Treat 2017;16:440-445.

18. Kaise M, Miwa J, Tashiro J, Ohmoto Y, Morimoto S, Kato M, et al. The combination of serum trefoil factor 3 and pepsinogen testing is a valid non-endoscopic biomarker for predicting the presence of gastric cancer: a new marker for gastric cancer risk. J Gastroenterol 2011;46:736-45.

19. Crnogorac-Jurcevic T, Chelala C, Barry S, Harada T, Bhakta V, Lattimore S, et al. Molecular analysis of precursor lesions in familial pancreatic cancer. PLoS One 2013;8:e54830.

20. Terris B, Blaveri E, Crnogorac-Jurcevic T, Jones M, Missiaglia E, Ruszniewski P, et al. Characterization of gene expression profiles in intraductal papillary-mucinous tumors of the pancreas. Am. J Pathol 2002;160:1745-1754.

21. Collier JD, Bennett MK, Bassendine MF, Lendrum R. Immunolocalization of pS2, a putative growth factor, in pancreatic carcinoma. J Gastroenterol. Hepatol 1995;10:396-400.

22. Guppy NJ, El-Bahrawy ME, Kocher HM, Fritsch K, Qureshi YA, Poulsom R, et al. Trefoil factor family peptides in normal and diseased human pancreas. Pancreas 2012;41:888-896.

23. Fueger PT, Schisler JC, Lu D, Babu DA, Mirmira RG, Newgard CB, et al. Trefoil factor 3 stimulates human and rodent pancreatic islet beta-cell replication with retention of function. Mol Endocrinol 2008;22:1251-9.

24. Xue Y, Shen L, Cui Y, Zhang H, Chen Q, Cui A, et al. Tff3, as a novel peptide, regulates hepatic glucose metabolism. PLoS One 2013;8:e75240.

25. Wang F, Herrington M, Larsson J, Permert J. The relationship between diabetes and pancreatic cancer. Mol Cancer 2003;2:4.

26. Bekaii-Saab T, Goldberg R. Therapeutic advances in pancreatic cancer: miles to go before we sleep. J Natl. Cancer Inst 2015;107.

27. Wang QH, Ji ZG, Chen ZG, Li HZ, Fan H, Fan XR, et al. Serum CA 19-9 as a good prognostic biomarker in patients with bladder cancer. Int J Surg 2015;15:113-116.

28. Zhang F, Deng Y, Drabier R. Multiple biomarker panels for early detection of breast cancer in peripheral blood. Biomed Res Int 2013;2013:781618.

29. Gorelik E, Landsittel DP, Marrangoni AM, Modugno F, Velikokhatnaya L, Winans MT, et al. Multiplexed immunobead-based cytokine profiling for early detection of ovarian cancer. Cancer Epidemiol Biomarkers Prev 2005;14:981-7.

30. Simmons AR, Clarke CH, Badgwell DB, Lu Z, Sokoll LJ, Lu KH, et al. Validation of a Biomarker Panel and Longitudinal Biomarker Performance for Early Detection of Ovarian Cancer. Int J Gynecol Cancer 2016;26:1070-7.

31. Chen H, Qian J, Werner S, Cuk K, Knebel P, Brenner H. Development and validation of a panel of five proteins as blood biomarkers for early detection of colorectal cancer. Clin Epidemiol 2017;9:517-526.

32. Chang JC, Kundranda M. Novel Diagnostic and Predictive Biomarkers in Pancreatic Adenocarcinoma. Int J Mol. Sci 2017;18.

33. Brand RE, Nolen BM, Zeh HJ, Allen PJ, Eloubeidi MA, Goldberg M, et al. Serum biomarker panels for the detection of pancreatic cancer. Clin Cancer Res 2011;17:805-816.

34. Kaur S, Chakraborty S, Baine MJ, Mallya K, Smith LM, Sasson A, et al. Potentials of plasma NGAL and MIC-1 as biomarker(s) in the diagnosis of lethal pancreatic cancer. PLoS. One 2013;8:e55171.

35. Nolen BM, Brand RE, Prosser D, Velikokhatnaya L, Allen PJ, Zeh HJ, et al. Prediagnostic serum biomarkers as early detection tools for pancreatic cancer in a large prospective cohort study. PLoS One 2014;9:e94928.

36. Brand RE, Nolen BM, Zeh HJ, Allen PJ, Eloubeidi MA, Goldberg M, et al. Serum biomarker panels for the detection of pancreatic cancer. Clin Cancer Res 2011;17:805-16.

37. Gayet O, Loncle C, Duconseil P, Gilabert M, Lopez MB, Moutardier V, et al. A subgroup of pancreatic adenocarcinoma is sensitive to the 5-aza-dC DNA methyltransferase inhibitor. Oncotarget 2015;6:746-54.

38. Sato N, Maitra A, Fukushima N, van Heek NT, Matsubayashi H, Iacobuzio-Donahue CA, et al. Frequent hypomethylation of multiple genes overexpressed in pancreatic ductal adenocarcinoma. Cancer Res 2003;63:4158-66.

39. Yamaguchi J, Mino-Kenudson M, Liss AS, Chowdhury S, Wang TC, Fernandez-Del Castillo C, et al. Loss of Trefoil Factor 2 From Pancreatic Duct

Glands Promotes Formation of Intraductal Papillary Mucinous Neoplasms in Mice. Gastroenterology 2016;151:1232-1244 e10.

40. Yamaguchi J, Yokoyama Y, Kokuryo T, Ebata T, Enomoto A, Nagino M. Trefoil factor 1 inhibits epithelial-mesenchymal transition of pancreatic intraepithelial neoplasm. J Clin Invest 2018.

41. Arumugam T, Brandt W, Ramachandran V, Moore TT, Wang H, May FE, et al. Trefoil factor 1 stimulates both pancreatic cancer and stellate cells and increases metastasis. Pancreas 2011;40:815-22.

42. Wang QH, Ji ZG, Chen ZG, Li HZ, Fan H, Fan XR, et al. Serum CA 19-9 as a good prognostic biomarker in patients with bladder cancer. Int J Surg 2015;15:113-6.

43. Guppy NJ, El-Bahrawy ME, Kocher HM, Fritsch K, Qureshi YA, Poulsom R, et al. Trefoil factor family peptides in normal and diseased human pancreas. Pancreas 2012;41:888-96.

Chapter 5

Functional and mechanical role of TFF1 in gemcitabine resistance of pancreatic cancer

Synopsis

Identifying factors facilitating chemoresistance, a major cause of Pancreatic Cancer (PC) mortality, could directly impact the overall patient survival. TFF1, a member of mucin-associated small secretory molecules, is one of the top overexpressed genes in the classical subtype (gemcitabine-resistant subtype) of PC. Multipronged approach suggests significant upregulation of TFF1 in preneoplastic lesion and PC. Interestingly, continuous treatment with gemcitabine significantly upregulated TFF1 in PC cells. Based on this evidence, we hypothesize that TFF1 plays a critical role in gemcitabine resistance (GR) in PC. TCGA database was used to correlate between TFF1 and GR predictor. To understand role of TFF1 in GR, in vitro studies in SW1990-TFF1-knocked-down (SW1990-TFF1-KD) cells were performed. Confocal microscopy, immunoprecipitation, chromatin immunoprecipitation, and siRNA were utilized to identify interacting partner and regulator of TFF1. Protein-protein docking studies using BioLuminate module were performed to predict possible signaling receptor of TFF1.TCGA database analysis revealed a significant positive correlation between TFF1 and GR predictor of PC (P=0.0001). Our in vitro studies showed that SW1990-TFF1-KD cells induced apoptosis, reduced colony formation capacity and modulated many apoptotic regulators such as Bax, Bcl-2, cleaved caspases in the presence of gemcitabine.

Furthermore, TFF1 was observed to be colocalized with MUC5AC, in human and mouse PC tissues suggesting their partnering are critical for PC pathogenesis.

Interestingly, 16 fold enrichment of GATA-6, an overexpressed transcription factor in classical subtype of PC, was observed on two distinct TFF1 promoter sites and GATA-6-siRNA repressed expression of TFF1. Moreover, protein-protein docking studies revealed the interaction of TFF1 with CXCR4 at Phe-172, Ser-122 and Glu-1 and TFF1 recombinant protein treatment in SW1990 cells increased CXCR4 mediated downstream signaling critical for GR. Our studies revealed that TFF1 plays an essential role in GR of PC cells by modulating apoptotic molecules possibly through CXCR4 which needs further validation.

Introduction

Pancreatic cancer (PC) is a major leading cause of cancer-related death in the USA and is expected to take second position surpassing breast cancer, prostate cancer, and colorectal cancer by 2030 [1]. Rapid metastasis, unrestrained chemoresistance, and lack of early prognostic markers are the major causes of poor survival of PC. Especially in case of gemcitabine which is still the first line of therapy, PC patient rapidly develop resistance toward this drug, and thus it is a major cause for poor outcome. Given that resistance to gemcitabine is one of the primary reasons for PC high mortality, an understanding of the molecular mechanisms behind resistance development is of paramount importance to aid in developing personalized therapy.

Very recently, PC has been subtyped into basal, classical and quasimesenchymal based on their gene signature identified from two independent groups of the researcher. While classical subtype demonstrated more resistance toward gemcitabine as compared to quasimesenchymal subtype [2, 3]. Surprisingly, from both independent investigators study, TFF1 emerged as a part of gene signature for classical subtype. TFF1, a small secretory molecule, highly upregulated in PC. While TFF1 has tumor suppressor role in gastric cancer, however, they have a tumorigenic role in prostate and pancreatic cancer [4, 5]. TFF1 is involved inhibiting apoptosis, increased migration, cell proliferation and tumorigenicity in cancer [6].

Although TFF1 is highly upregulated in classical subtype, contributed towards PC tumorigenicity and has played in the antiapoptotic mechanism, however, their role in PC chemoresistance is elusive. Thus, we aim to identify role of TFF1 in

gemcitabine resistance in PC in this study. Herein, we have demonstrated that TFF1 has a negative correlation with gemcitabine sensitivity ratio (GSR) and knocked down of TFF1 in SW1990 PC cell were sensitized to gemcitabine by modulating different apoptotic molecules. Our long-term treatment of PC cancer cells recapitulates the phenotype of gemcitabine-resistant cells, and chronic gemcitabine treatment have upregulated endogenous TFF1 in mRNA and protein level. We have also confirmed that GATA-6, a transcription factor upregulated in classical subtype of PC, binds on TFF1 promoter and positively regulate TFF1 in PC. Additionally, we have demonstrated that TFF1 colocalizes with MUC5AC in human and mouse PC tissues and interestingly interact with MUC5AC particularly in the gemcitabine-treated PC cell line. Taken together our studies have implicated a novel role of TFF1 as a contributing factor towards gemcitabine resistance in PC.

Results

Analysis of the cancer genome atlas (TCGA) database identifies an inverse relationship between GSR and TFF1

First, we analyzed the correlation between TFF1 and Gemcitabine Sensitivity Ratio (GSR), an indicator of gemcitabine resistance [7]. GSR, an indicator which was first developed by Nakao et al. first developed which consider a balance of cellular enzymes of gemcitabine transport and metabolism, that is, hENT1, dCK, RRM1, and RRM2. We have utilized the publicly available TCGA database (TCGA provisional, Pancreatic Cancer) as our primary source of clinical information. We downloaded expression data of TFF1, hENT1, dCK, RRM1, and RRM2 and

performed coexpression analysis between TFF1 and GSR from cBIOportal (http://www.cbioportal.org/). We further subdivided TCGA patient sample into classical, basal, exocrine and quasimesenchymal based on earlier published data and performed the correlation analysis in those subgroups [8]. For our grouped analysis we have excluded that patient samples which are termed as classical by one group and basal by another group and vice versa. Our analysis suggests that TFF1 is significantly (P=0.0001) negatively correlated (Spearman r=-0.2813) with GSR ratio in all PC samples (n=179) from TCGA suggesting TFF1 is positively associated with resistance (**Figure 1A**). Moreover, when we analyzed the correlation of TFF1 and GSR in different subtype, we found that GSR is lowest in classical subtype whereas highest in quasimesenchymal (QM) subtype **(Figure 1B).** Additionally, we also found that TFF1 is negatively correlated with classical subtype (Spearman r=-0.3043) whereas positively correlated with basal (Spearman r=0.00602) and quasimesenchymal (Spearman r=0.04151) subtype (**Supplementary Figure S1A**). We also found negative correlation value in exocrine (Spearman r=-0.112) however the value was not as high as classical. When we analyzed correlation between each component of GSR such as hENT1, dCK, RRM1 and RRM2, we found negative correlation between TFF1 and hENT1 (Spearman r=-0.076, p=0.31), dCK (Spearman r=-0.18, p=0.019) and positive correlation between RRM1 (Spearman r=0.062, p=0.41) and RRM2 (Spearman r=0.019, p=0.013) **(Supplementary Figure S1B).** The negative correlation between GSR values and TFF1 indicate that TFF1 might be associated with resistance to gemcitabine-mediated PC cell death. Our analysis suggested a

negative correlation between TFF1 and GSR in a panel of PC cell line which also corroborate with the analysis from the publicly available dataset by demonstrating a negative correlation.

Long-term gemcitabine-treated resulted in increased cancer stem cell population, migratory potential while decreased GSR value

As TFF1 was correlated with GSR ratio in PC dataset, we hypothesized that if TFF1 activity is important for gemcitabine resistance, then long-term gemcitabine treatment cells should elevate levels of endogenous TFF1. To investigate this hypothesis, we treated SW1990, and COLO357 PC cell lines, cell line which has moderate GSR value (**Supplementary Figure 2A**) with increasing concentration of gemcitabine were exposed to continuous gemcitabine treatment for 10 weeks which lead to the outgrowth of a resistant population. Our long-term gemcitabine-resistant cell demonstrated significant higher number of SP population (Figure 2A) lower GSR value (**Figure 2C**) [9, 10]. We also analyzed the mRNA expression of stem cell populations marker and observed upregulated expression of SOX2, SOX9, NANOG, OCT3/4 in SW1990 GT PC as compared to control parental cell (**Figure 2B**). As we hypothesized that TFF1 is playing role in chemoresistance, then resistant cells should express elevated levels of endogenous TFF1. Corroborating with our expectation, we have found that high level of TFF1 in mRNA and protein level in SW1990-GT and Colo357 cells as compared to respective parental cells **(Figure D-E)**

TFF1 knockdown sensitizes PC cell to gemcitabine

To identify the role of TFF1 in gemcitabine tolerance in PC cells, we have knocked down TFF1 in SW1990 cells PC cells and confirmed by immunoblotting (**Figure 3A**). As TFF1 has played a role as antiapoptotic, we applied staurosporine (2 µM for 12 hrs) to validate its antiapoptotic role as it is one of the mechanisms for pertaining gemcitabine resistance. Our result suggested significant upregulation of apoptotic cells in SW1990-TFF1-KD cells as compared to SW1990-Scr PC cells ($P<0.05$) (**Figure 3B**) upon staurosporine treatment.

Similarly, we have also seen a significant increase of apoptotic cells upon gemcitabine treatment (1µM for 24 hrs) in SW1990-TFF1-KD cells as compared to SW1990-Scr ($P<0.05$) (**Figure 3B**). We also observed significant decrease number of colonies in SW1990-TFF1-Sh cells as compared to SW1990 -Scr cells in the presence of gemcitabine (**Figure 3C**). Next, we investigated the potential mechanisms by which TFF1 modulates apoptosis in SW1990 PC cells by screening targets using a human apoptosis antibody array. As shown in Figure 5C, the density of the signals for the positive control were similar between samples, suggesting that the cell lysates from gemcitabine-treated SW1990-Scr and SW1990-TFF1-Sh were loaded equally onto the arrays. We found that TFF1 knocked down has increased cleaved caspase while decreased Bcl-X, HSP-70, cIAP molecules suggesting TFF1-KD sensitizes PC cells towards gemcitabine by modulating apoptotic signaling molecules (**Figure 3D-E**). Cumulatively, these data suggest TFF1 has a direct impact on gemcitabine tolerance in PC cells.

GATA6 regulates TFF1 in PC

To determine the correlation between GATA-6 and GSR ratio, we performed similar analysis as discussed earlier. Similar to TFF1, GATA-6 also negatively correlated with GSR ratio in classical subtype of PC while positively correlated in exocrine, basal and quasimesenchymal subtype which is analyzed from TCGA database downloaded from cBIOPORTAL (**Supplementary Figure S3A**). Moreover, expression of GATA-6 was highest in classical subtype as compared to another abovementioned subtype of PC (**Supplementary Figure S3B**). Long-term gemcitabine increased GATA-6 expression in SW1900 cells (**Figure 4A**). To validate these findings, we have examined GATA-6 and TFF1 coexpression on human PC tissues using immunohistochemistry on PC tissue microarray (**Figure 4B**). To determine the direct association of GATA6 with TFF1, we first analyzed in silico the promoter sequence for GATA-6 occupancy on the TFF1 promoter. Our analysis revealed that two putative binding site for GATA6 is present 181 and 384 base pair upstream of the transcription start site of TFF1. Having identified their binding site, we performed chromatin immunoprecipitation assay to validate the binding of GATA6 on TFF1 promoter which indeed suggested that 16 fold and 32 fold enrichment of GATA-6 binding on two distinct sites TFF1 promoter (**Figure 4 C-D).** We then utilized web-based tool GEPIA (Gene Expression Profiling Interactive Analysis) to identify the correlation which demonstrates that GATA-6 is significantly positively correlated with TFF1 (Supplementary Figure Interestingly, TFF1-GATA-6 was colocalization was observed in the majority of PC tissues. Next, we transiently knocked down GATA-6 demonstrated that *siRNA-mediated* repression of GATA6 in SW1990 PC cells also repressed TFF1 (**Figure** 4**E**)

Trefoil Factor interacts with MUC5AC in long-term treated gemcitabine cell line

TFF1-mucin interaction is a long-standing partner which play a role in gastrointestinal protection. However, their role in disease pathogenesis is not known yet. To identify potential mucin interacting partner in PC, we analyze correlation between several and transmembrane mucin with TFF1 in PC from web-based tool GEPIA which has the accumulation of TCGA and GTEX database which is consisted of RNA sequencing expression data of 9,736 tumors and 8,587 normal samples across various tumors (http://gepia.cancer-pku.cn/index.html) [11]. Spearman correlation analysis revealed that TFF1 were significantly correlated with MUC5AC (r= 0.71) in human PC using TCGA database followed by MUC1 (r= 0.62) and MUC 13 (r= 0.48) ((**Supplementary Figure S4A**). Moreover, as we earlier revealed that GATA-6 can regulate TFF1, we also analyze the correlation between GATA-6 and MUC5AC. Interestingly the correlation value between GATA-6 and MUC5AC (r= 0.44) was positive and highest followed by MUC1 (r= 0.4) and MUC 13 (r= 0.39) (**Supplementary Figure S4B**). Based on this, we next analyze their colocalization in PC patient tissues. Our immunofluorescence studies suggest strong colocalization of TFF1 and MUC5AC in PC patient sample (**Figure 5A**). We also analyzed their colocalization in 20 and 30 weeks of from spontaneous Kras (G12D/+); Pdx-1-Cre pancreatic tumors which also corroborate with findings from human patient tissue samples (**Figure 5B**). As TFF1 is highly correlated with MUC5AC, next we asked how they correlate with the metabolic genes which are responsible for gemcitabine resistance. Our

analysis has revealed that negative correlation of MUC5AC with hENT1 (Spearman r=-0.05, p=0.51) and dCK (Spearman r=-0.15, p=0.051) whereas positive correlation between RRM1 (Spearman r=0.21, p=0.0059) and RRM2 (Spearman r=0.35, p<0.001) (**Supplementary Figure S5**). This data has demonstrated that MUC5AC showed similar correlation pattern with the component of GSR like TFF1 suggesting MUC5AC may play a role in gemcitabine resistance.

Moreover, MUC5AC was significantly upregulated in SW1990-GT cells as compared to SW1990 cells in both mRNA and protein level (**Figure 5C**). Interestingly, our immunoprecipitation analysis suggested that TFF1 interacts with MUC5AC in SW1990-GT cells where no interaction was observed in SW1990 parental cell (**Figure 5D**). Collectively, these results indicate that TFF1-MUC5AC plays a crucial role in PC pathogenesis and gemcitabine resistance.

TFF1 mediates downstream signaling probably through CXCR4

To understand the mechanism through which TFF1 mediates signaling, we explored the possibility of CXCR4 as a signaling receptor for TFF1 as very recently chemokine receptor CXCR4 has been described as a low-affinity receptor for TFF2**,[12].** We began with identifying possible binding interactions between TFF1 and the CXCR4 in an x-ray structure-based modeling system. To define possible interactions between the chemokine receptors and TFF1 we utilized published x-ray structure data and performed 3D modeling structural analysis. The most probable protein-protein docking arrangement between TFF1 and CXCR4 is displayed in **Figure 6A.** This TFF1-CXCR4 complex is formed by pi-stacking

between amino acid side chains of CXCR4 and TFF1 (Phe172-Chain A: Phe60 chain D), hydrogen bonds (Ser-122 -Chain A: Glu-55-chain D), and (Glu-1-chain C: ile-89-chain A). These interactions of TFF1 mainly occur CXCR4 trans-membrane helices IV and III while one was at the beginning of N terminal. These software modeling results indicate that CXCR4 and may have the possibility to interact with the TFF1 molecule. Functional studies of TFF1-CXCR4 interaction were accomplished by cell culture-based migration assays and evaluation of activation of the ERK/AKT signaling cascade. To identify that TFF1 signals through CXCR4, we blocked that receptor with widely used CXCR4 inhibitor AMD3100. As shown in **Figure 6B**, treatment with TFF1 (10nM) upregulated while AMD3100 (1µM) downregulated pAKT and pERK as compared to control in SW1990 cells confirmed that both TFF1 and AMD3100 are active. However, downstream signaling of TFF1 was partially inhibited by AMD3100 **(Figure 6B).**

Moreover, our functional studies suggest that while migration ($p<0.002$) and colony formation ability ($p<0.05$) of SW1990 cells was significantly increased by TFF1 treatment, AMD 3100 was able to abrogate those functional activities. However, combination treatment of AMD 3100 and TFF1 significantly decreased colony formation and migration index as compared to TFF1 alone which suggest AMD3100 partially abrogated the migration and colony formation ability of TFF1 (**Figure 6C-D**). This result collectively suggest that cell migration and proliferation mediated by TFF1 can partly block by AMD3100 which potentially indicates that TFF1 mediates downstream signaling partially through CXCR4.

Discussion

Chemoresistance is a major cause for high mortality of PC despite numerous studies having been carried out; no significant progress has occurred in the previous two decades regarding overall survival in five years [13]. So delineating molecular mechanism behind current treatment failure demands an immediate action to formulate effective personalized therapy for PC. Studies herein are aimed to identify the role of TFF1 in gemcitabine resistance of PC, a molecule which is significantly upregulated molecule in classical subtype and demonstrated the tumorigenic potential to PC [4].

First, we analyzed the correlation of TFF1 with GSR ratio, an indicator of gemcitabine sensitivity /resistance first coined by Nakano y. et al. in 2007 [14]. Earlier studies suggest that lower abundance of hENT1, a major gemcitabine transporter in human pancreatic cancer is responsible for gemcitabine resistance. Similarly, as well as mutation of dCK, a rate-limiting enzyme for phosphorylation of gemcitabine are one of the primary mechanisms responsible for the development of resistance to gemcitabine. It has been shown that modulation of cellular enzymes of gemcitabine transport and metabolism influences drug activity in vitro.[14]. Nakano et al. first reported that no single gemcitabine metabolic gene is correlated with gemcitabine resistance rather than gemcitabine resistance come from a balance between dCK, RRM1, RRM2, and hENT1 and the balance is represented as gemcitabine ratio using -hENT1 × dCK/RRM1 × RRM2 [14]. In our analysis from TCGA data for identifying correlation between GSR and TFF1, negative correlation between them potentially indicates that TFF1 is positively

correlated with gemcitabine resistance. Interestingly, when we subtyped the patient category based on earlier published report and analyzed the correlation, we found negative correlation especially in classical but not any other subtype suggesting that TFF1 may play crucial role in gemcitabine resistance which might subtype specific. It is not surprising that tolerance of drug is subtyped specific as Heiser et al. also demonstrated that when 77 Food and Drug Administration-approved and investigational compounds were applied on ~50 different breast cancer cell which are well defined based on transcriptional and genomic subtype, at least one third of the drug showed subtype specific response [15] .

Furthermore, when Heiser et al. employed superpathway network to identify potential pathway, they found FOXA1, a transcription factor which regulates TFF1 in luminal subtype of breast cancer, is upregulated in luminal subtype of breast cancer where Lapatinib, Trichostatin A, Triciribine, along with other compounds showed luminal subtype specific sensitivity [15]. So, bearing in mind that a transcription factor regulates a set of the molecule to provide subtype specific response towards drug, we searched whether any transcription factor is related with TFF1 in classical subtype. Collison et al. identified GATA-6 is highly expressed in most classical subtype tumors and cell lines which was also supported by Moffitt et al. and The Cancer Genome Atlas Research Network [2, 3]. Moreover, loss of GATA-6 was found in basal-like tumors with poor outcome [16]. Furthermore, repression of GATA-6 in SW1990 PC cells resulted in decreased cell growth and increased apoptosis in PC [17]. In addition, Al-azzeh et al. demonstrated that GATA-6 activates TFF1 and TFF2 in gastric and intestinal

cell line [18]. Our in-vitro result also corroborates that GATA-6 is negatively correlated with GSR while positively associated with classical subtype gene like TFF1.

Moreover, it is also upregulated in long term treated PC cells which further regulate TFF1 in PC. Though this part of the discussion is indicating that GATA-6 is indeed a master regulator for TFF1 in gemcitabine resistance, further studies are needed to be performed to understand their subtype-specific regulation as they both shown negative correlation with GSR and highly upregulated in classical subtype of PC. In that case, development of PC subtype specific patient-derived cell line and organoid would be next logical step.

GATA-6 is a well-known transcriptional regulator for mucin specially MUC4 and MUC5AC [19, 20]. Moreover, our correlation studies suggested that among all other mucin, MUC5AC was the top positively correlated with GATA-6 as well as TFF1 among all other mucin in PC. MUC5AC, a gel-forming mucin, is de novo expressed and highly upregulated in PC and hypothesized to play a critical role in this disease progression [21]. Furthermore, MUC5AC showed similar correlation to TFF1 with the component of GSR ratio in PC. Earlier studies suggested that TFF1 and MUC5AC interact with each other to provide gastrointestinal protection [21]. Interestingly our invitro studies also indicate that their interaction sustains in terms of pathological condition which is PC chemoresistance in our case. Though role of MUC5AC in lung cancer chemoresistance has been identified, however, its significance in PC chemoresistance is still needed to be investigated as our in vitro

studies also identified its upregulation in long-term gemcitabine-treated PC cell line [22].

Higher invasion potential, EMT and increased number of stem cell population are characteristics features of gemcitabine-resistant cell line [9, 23-25]. TFF1 was also associated with for induction EMT, increase migration and invasion [26]. They are also involved in differentiation of stomach however role of TFF1 in stem cell modulation is not yet explored in cancer [27]. Though the phenotype of gemcitabine resistant PC cell and effect of increased TFF1 is similar, it is the first report to direct exploration between TFF1 and gemcitabine resistance in PC. TFF1 may also modulate gemcitabine resistance by reprogramming tumor cells into CSCs which is further needed to be investigated.

Furthermore, activation of antiapoptotic pathway is one of the major mechanisms of gemcitabine resistance [24]. Earlier report demonstrated TFF1 inhibited doxorubicin mediated apoptosis in breast cancer cell line [28]. Moreover, TFF has been involved in antiapoptosis in other cancer as well (ref). From our studies, it is also evident that TFF1 is involved in upregulating antiapoptotic pathway in gemcitabine resistance. While our studies provided evidence of the involvement of TFF1 in gemcitabine resistance, further exploration is needed to analyze its role in erlotinib tolerance as classical subtype also demonstrated to have higher erlotinib sensitivity.

To further investigate, by which receptor TFF1 mediate antiapoptotic signaling for gemcitabine resistance, we specifically focused on CXCR4 as it has emerged as low-affinity receptor for TFF2 [12]. Moreover, a recent x-ray-based computer

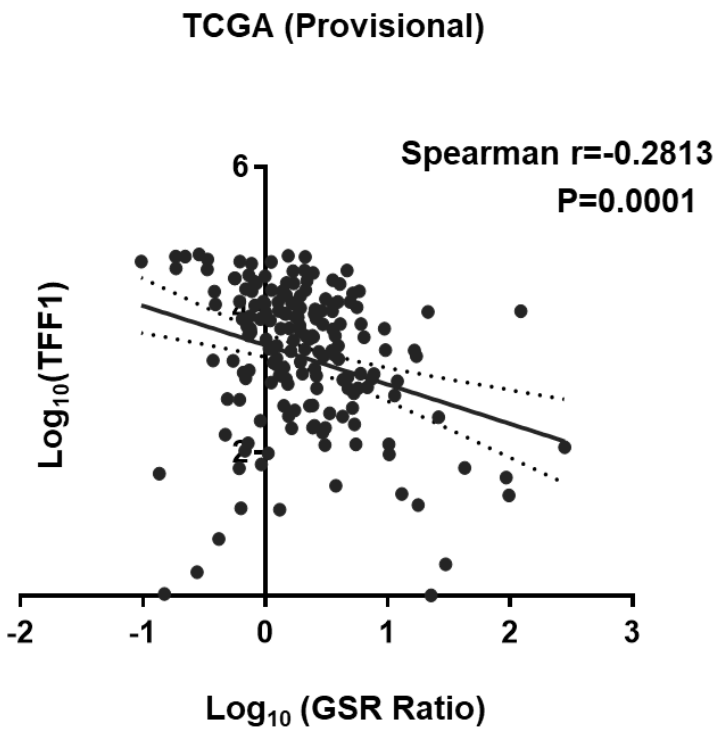
modeling system identified possible interaction with TFF3 and CXCR4 by several salt bridge, hydrogen bond and hydrophobic interaction [29]. Furthermore, role of CXCR4 in stem cell, invasion, migration, EMT, chemoresistance, cell proliferation, antiapoptotic, tumorigenesis and metastasis in PC is well documented [30]. Furthermore, CXCR4 expression increases from PanIN progression to invasive PC in a similar fashion to TFF1 [30]. Notably, treatment with gemcitabine also elevated CXCR4 level in PC similar to TFF1 from our study [31]. Keeping in mind the role of CXCR4 in gemcitabine resistance in PC and their recent correlation with TFFs, we went ahead to identify their interaction using in silico protein-protein docking study which indeed suggested TFF1s' interaction with transmembrane helices of CXCR4. Regarding cell signaling, CXCR4 transmembrane helices contain many critical residues for signaling and undergo conformational changes upon activation CXCR4 which allow transmission of signaling upon binding of extracellular ligand to TM domains [32]. In the future, it would be very interesting to study particular interaction of TFF1 with CXCR4 transmembrane helices using mutagenesis studies. Moreover, our invitro studies suggest that TFF1 probably mediate downstream signaling which is partially blocked by AMD3100. One of the potential mechanisms of this partial blocking is that a critical residue Phe 172 (Transmembrane-IV) is essential for AMD3100 which is a binding site for TFF1 as well suggesting AMD 3100 may block TFF1 interaction with CXCR4 by binding with this residue not with Ser-122 and Glu-1 of CXCR4. It was earlier speculated that this phenylalanine residue is crucial for interactions with the aromatic linker of AMD3100 which may change overall configuration of CXCR4 protein. It will be

interesting to study whether TFF1 has any self-induction/feedback loop mechanism through CXCR4

Overall, our studies highlight novel role of TFF1 in PC gemcitabine resistance. Additionally, we have demonstrated that TFF1 is regulated by GATA-6 and may propagate downstream signaling through CXCR4 which need further investigation. As recently PC has been transcriptionally and genetically subtyped for effective therapy, it is high time to understand the molecular mechanism of gene which are highly upregulated in the different subtype.

Figure 1

A


TCGA (Provisional)
Spearman r=-0.2813
P=0.0001
Log10(TFF1)
6
2
-2
-1
0
1
2
3
Log10 (GSR Ratio)

B

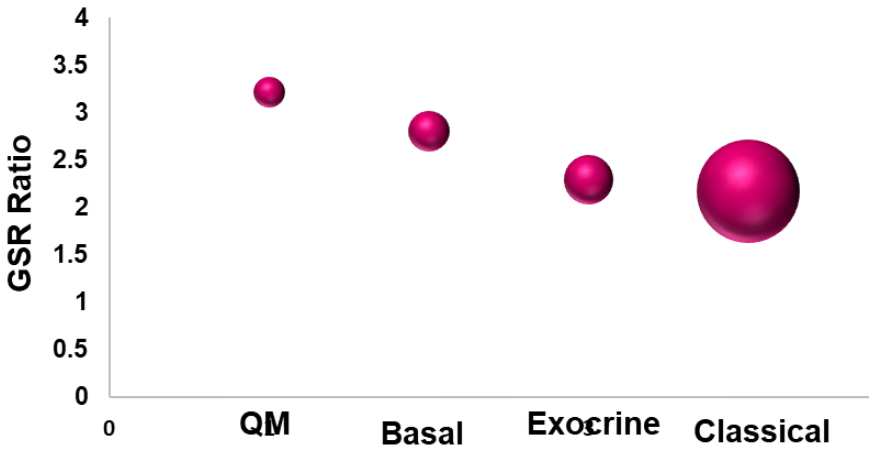

GSR Ratio
4
3.5
3
2.5
2
1.5
1
0.5
0
0
QM
Basal
Exocrine
Classical

Figure 1: Correlation of TFF1 and gemcitabine sensitivity ratio (GSR) (**A**) Correlation analysis between TFF1 mRNA expression and Gemcitabine Sensitivity Ratio (GSR) in the TCGA (Provisional) PC clinical dataset (N=179) showed a significant negative correlation (r=-0.2813). (**B**) Dot plot showing GSR is highest in quasimesenchymal (QM) whereas highest in classical subtype. The radius of the sphere is indicating expression of TFF1 which suggests that TFF1 is highly expressed in classical subtype of PC.

Figure 2 A

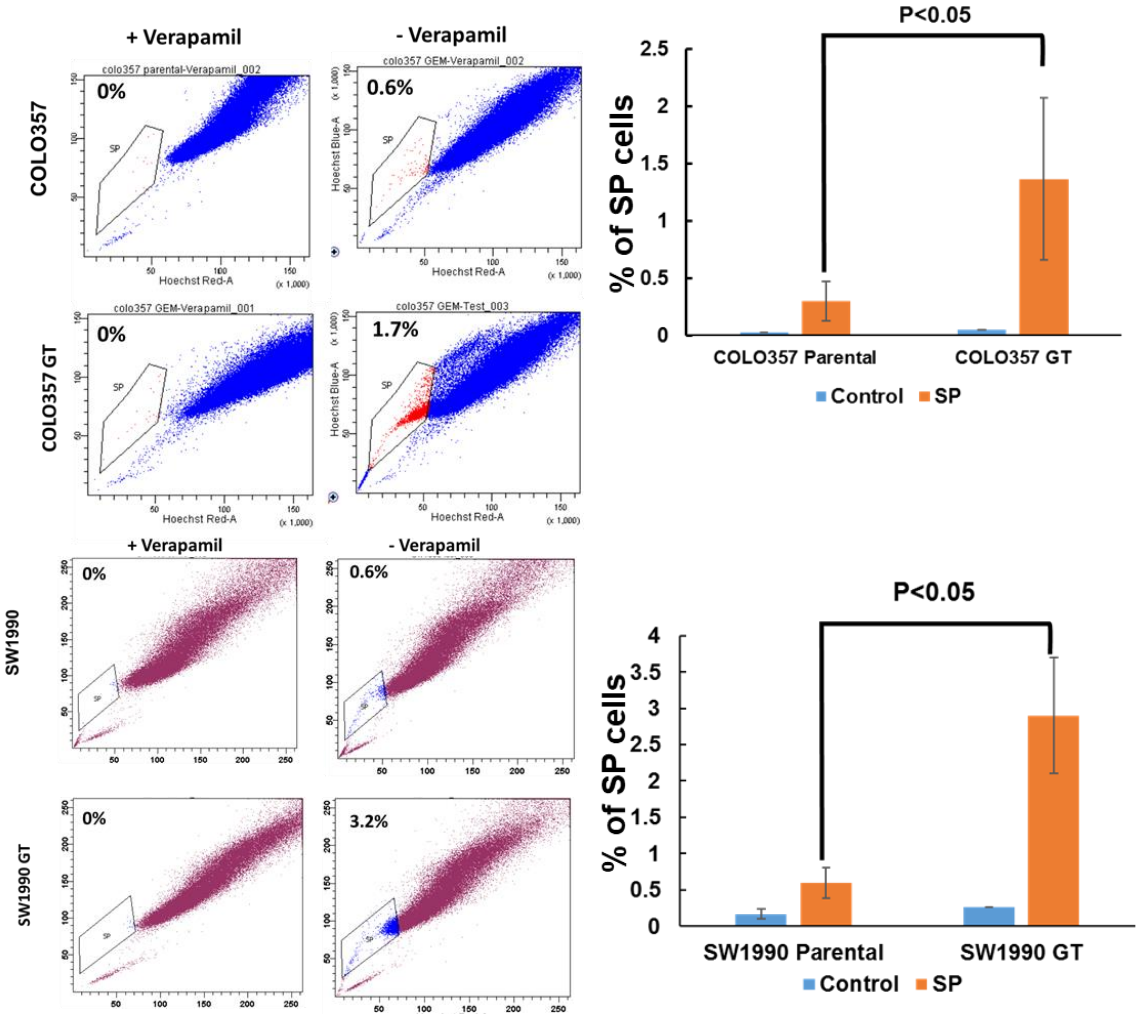
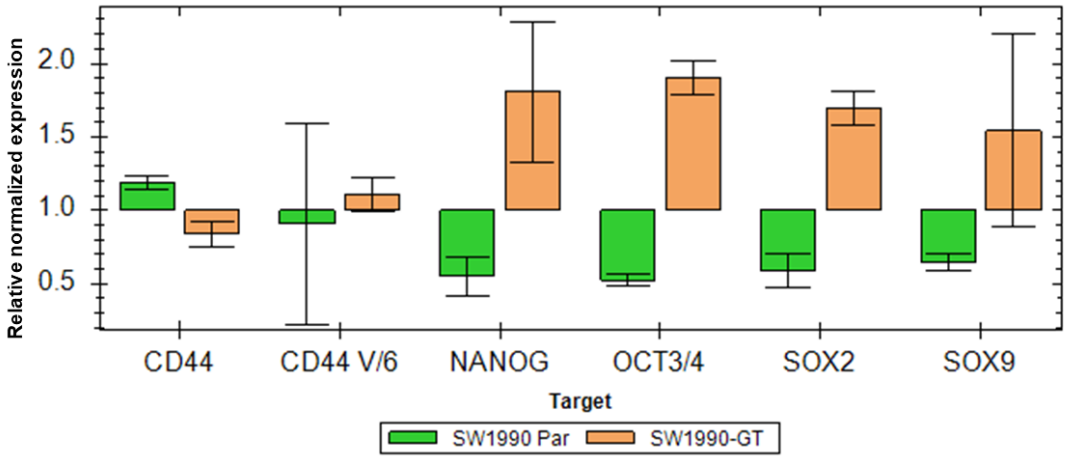

+ Verapamil
- Verapamil
COLO357
COLO357 GT
SW1990
SW1990 GT
0%
0.6%
1.7%
3.2%
SP
Hoechst Red-A
Hoechst Blue-A
P<0.05
% of SP cells
COLO357 Parental
COLO357 GT
SW1990 Parental
SW1990 GT
Control
SP

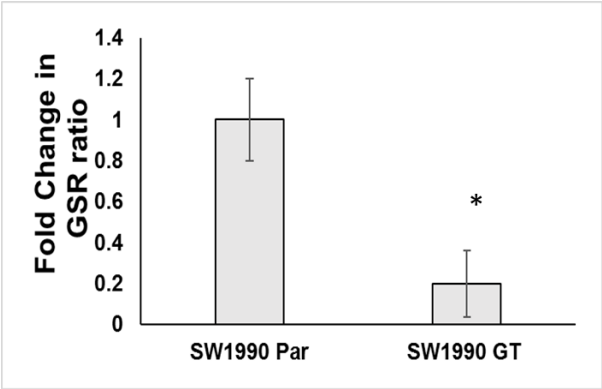
Figure 2 (A): Effect of long-term treatment in Side Population (SP) cells. SW1990 and COLO357 cells were stained with Hoechst 33342 dye in the presence (left) or absence (right) of verapamil and analyzed by flow cytometry. The SP, which disappears in the presence of verapamil, was gated and shown as a percentage of the whole viable cell population. Long-term gemcitabine-treated (GT) PC cells demonstrated an increase in Side Population (SP) in both SW1990-GT and COLO357-GT cells as compared to respective parental cells. Corresponding bar diagram representing the percentage of SP populations ($P<0.05$).

Figure 2 (continued)

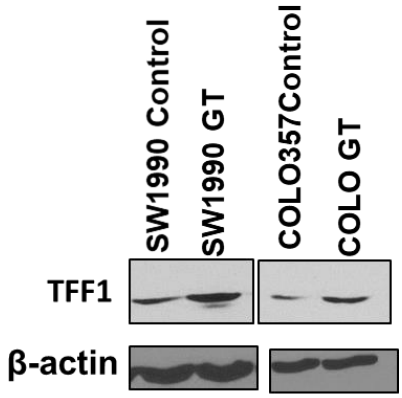
B



C



D



E

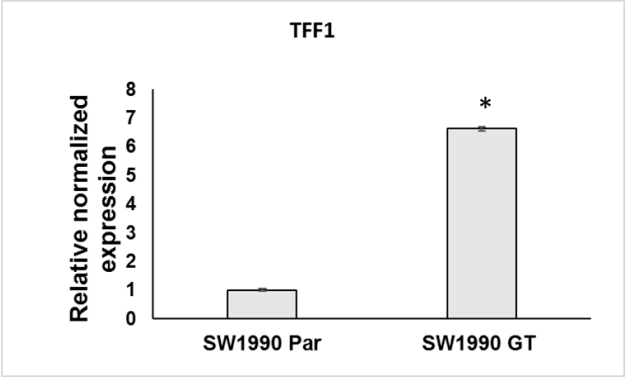
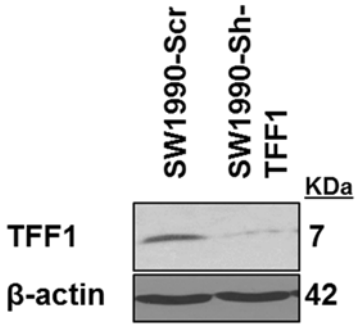


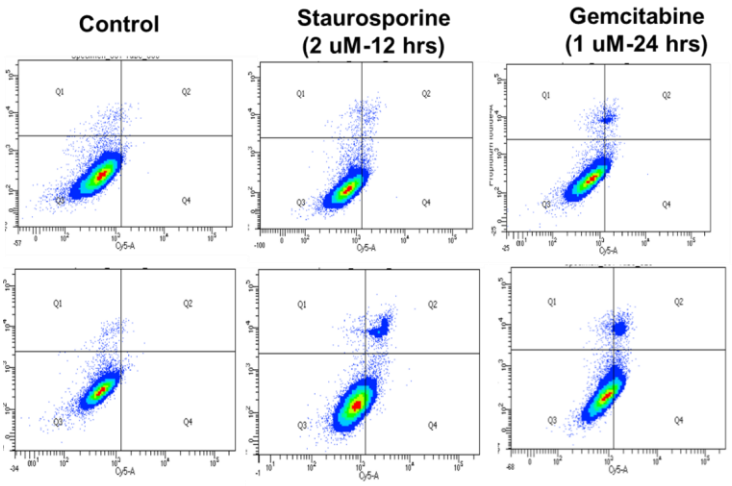
Figure 2 (B-E): Effect of long-term treatment in stem cell marker, GSR and TFF1. (B) qPCR analysis of long-term gemcitabine-treated (GT) PC cells demonstrated increase stem cell marker in mRNA level in both SW1990-GT and COLO357-GT cells as compared to respective parental cells. **(C)** Significant decrease of GSR is observed in long-term gemcitabine-treated SW1990 PC cells. (**D-E**) Western blot and qPCR analysis suggested that long-term gemcitabine treatment increased expression of TFF1 in both protein and mRNA level. * indicate significance $P<0.05$.

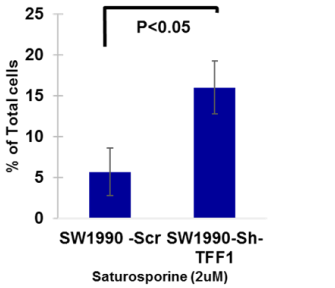
Figure 3

A.


SW1990-Scr
SW1990-Sh-TFF1
KDa
TFF1
7
β-actin
42

B


Control
Staurosporine (2 uM-12 hrs)
Gemcitabine (1 uM-24 hrs)


P<0.05
% of Total cells
25
20
15
10
5
0
SW1990 -Scr
SW1990-Sh-TFF1
Saturosporine (2uM)

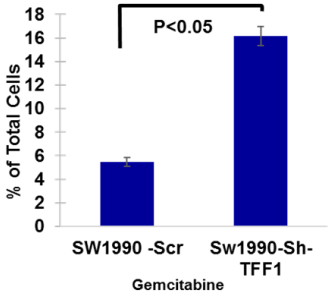
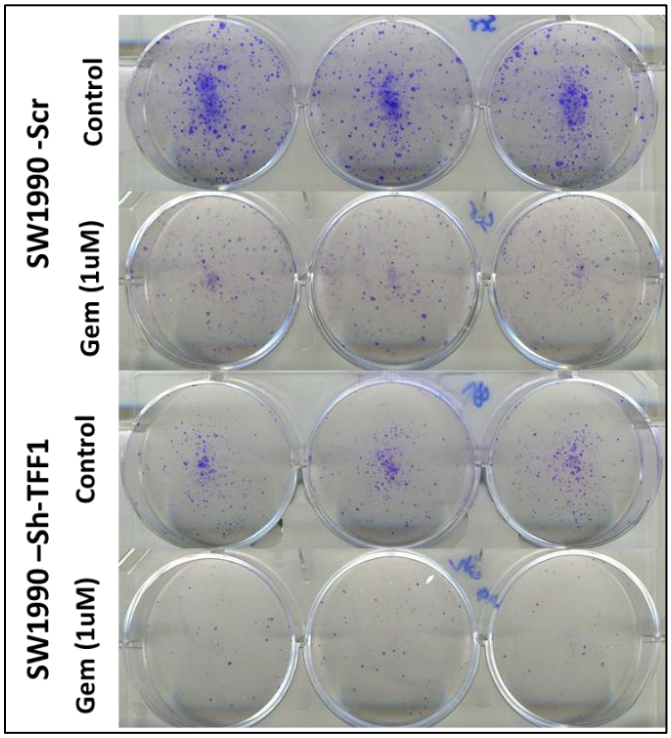

P<0.05
% of Total Cells
18
16
14
12
10
8
6
4
2
0
SW1990 -Scr
Sw1990-Sh-TFF1
Gemcitabine

Figure 3 (A-B). Role of TFF1 knockdown in gemcitabine resistance. (A) Confirmation TFF1 knockdown (KD) in SW1990 PC cells of using western blot image. **(B)** Effect of gemcitabine and staurosporine on apoptosis of SW1990-Scr and SW1990-TFF1-sh cells in vitro. Apoptosis was detected by dual dye staining using Annexin V-FITC/PI. The percentage of apoptotic cells is presented as the mean ± SD in the corresponding bar diagram. In the presence of apoptotic inducer staurosporine (2μM, 8 hrs) and gemcitabine (1μM, 24 hrs) TFF1-KD significantly increased apoptosis in PC cells.

Figure 3 (Continued)

(C)



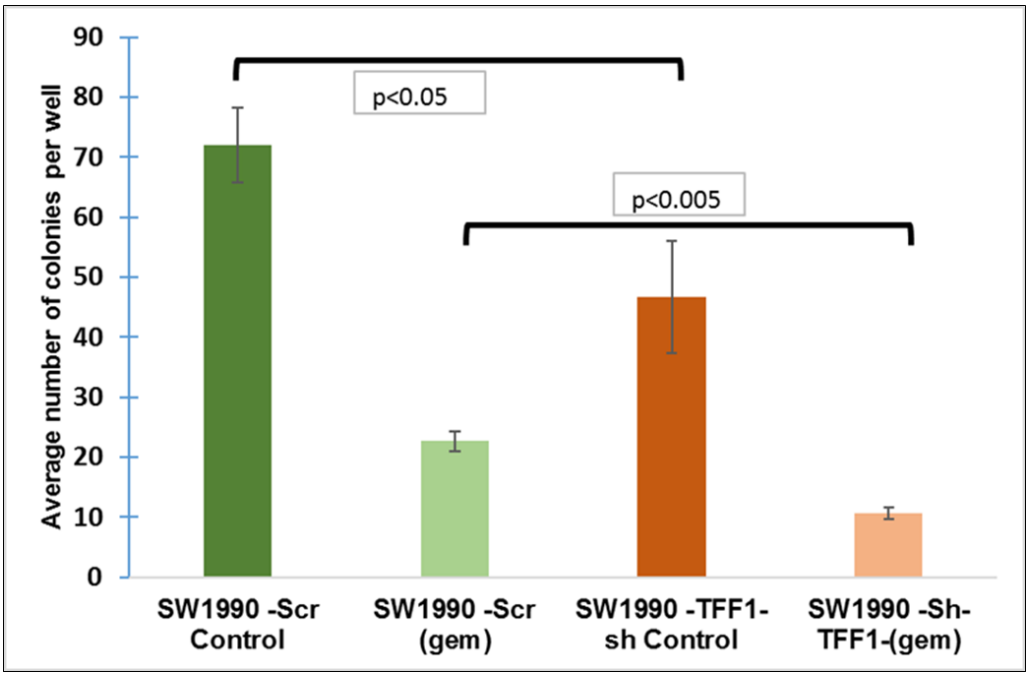
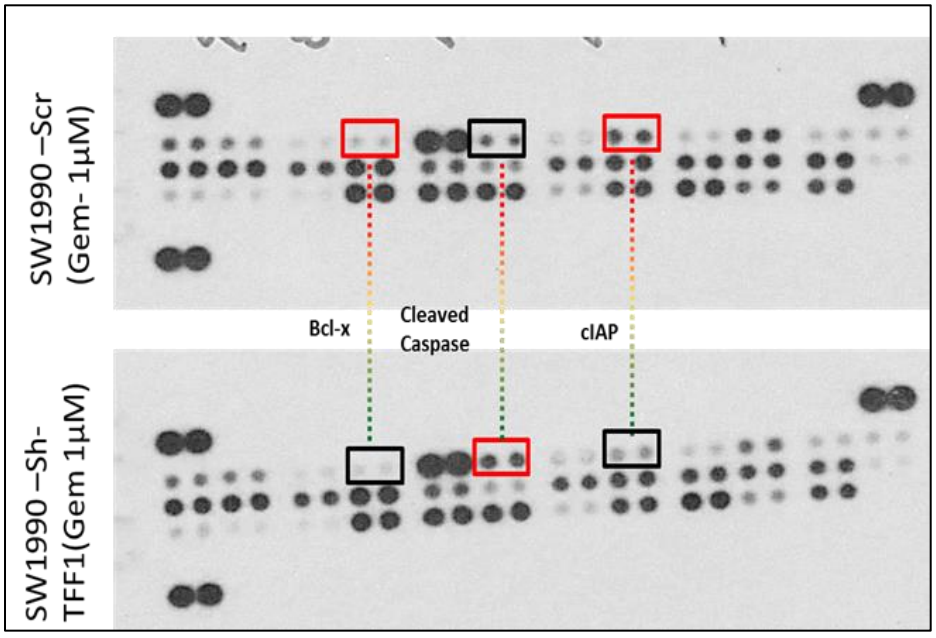


Figure 3 (C): **TFF1 knockdown decreased colony formation in the presence of gemcitabine (1 μM) (C)** Colony forming assay in SW1990-Scr and SW1990-TFF1-Sh cells. Bar diagram showing a decrease in the average number of colonies in SW1990-TFF1-sh PC as compared to SW1990-Scr cells after treating with gemcitabine (1μM).

Figure 3 (Continued)

(D)



(E)

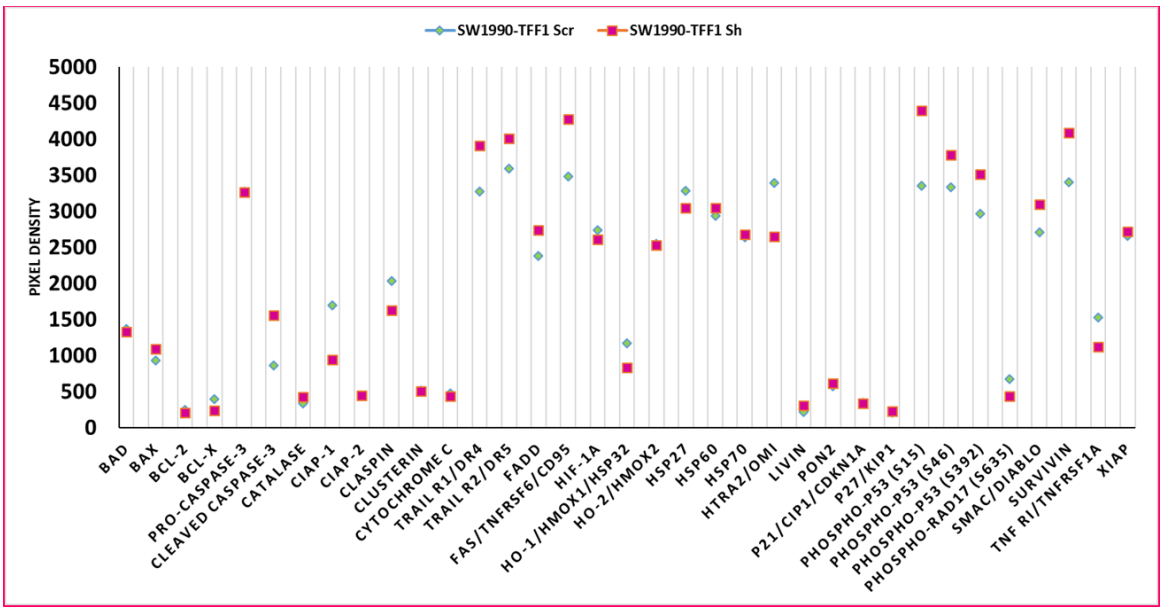
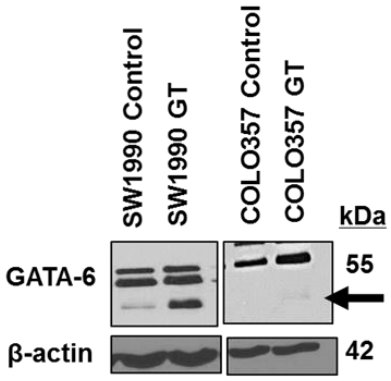


Figure 3: Effect of TFF1 knockdown in apoptotic protein molecules (D) Apoptosis protein arrays for the SW1990-Scr and SW1990-TFF1-Sh cells treated with gemcitabine (1μM, 24 hrs) show the modulation of apoptotic regulators. Right and left upper corner dots represents reference proteins. **(E)** The density of each dot was quantified with Image J and average pixel density of each spot (n=2) were presented as dot plot. The diagram represents the modulation of apoptotic-related molecule upon TFF1 knockdown.

Figure 4:

A


SW1990 Control
SW1990 GT
COLO357 Control
COLO357 GT
kDa
GATA-6
55
β-actin
42

B

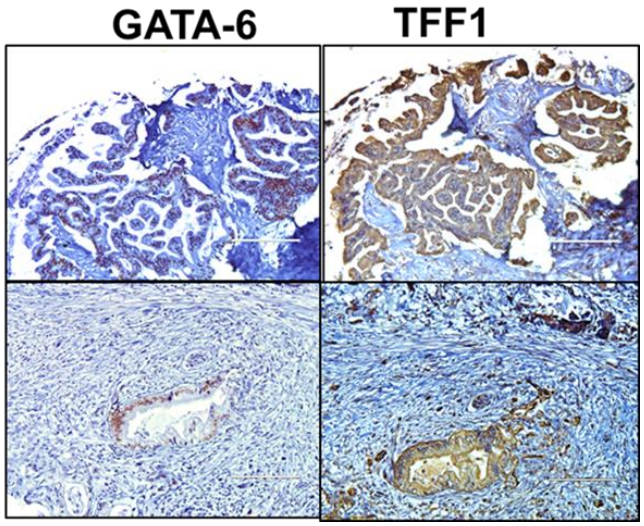
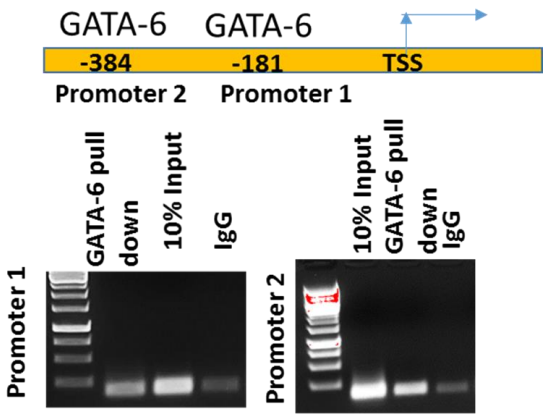

GATA-6
TFF1

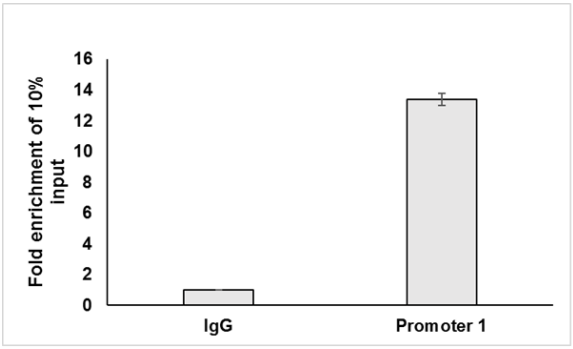
Figure 4: Coexpression of GATA-6 in same PC duct. (A) Western blot image of GATA-6 in long-term treated SW1990, and COLO357 cell demonstrated that there is increased level of expression in SW1990-GT and COLO357 GT cell line**. (B)** Immunohistochemistry in TMA of PC tissues for GATA-6 and TFF1 suggested coexpression in the same duct of PC tissues.

Figure 4 (Continued)

C



D



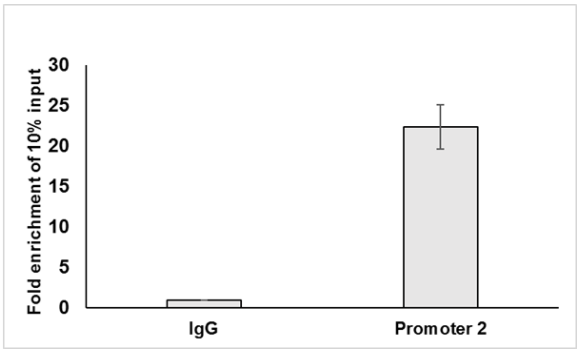
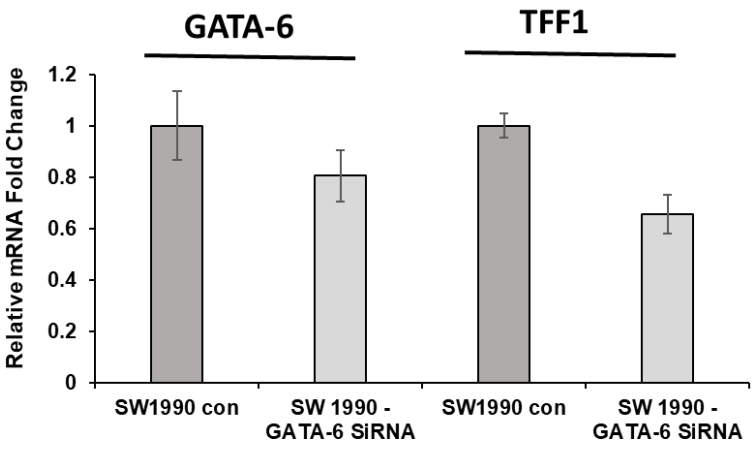


Figure 4: GATA6 binding to the promoters of TFF1 detected by ChIP-qPCR in SW1990 cells (C) In silico analysis suggests that there is two different GATA6 binding site in TFF1 promoter region. The semiquantitative polymerase chain reaction (PCR) result showed binding of GATA-6 on two different sites of TFF1 promoter. **(D)** Quantitate determination of binding of GAT6 by real-time PCR for two TFF1 promoters suggesting 14- and 20-fold enrichment as compared to negative control. ChIP-qPCR data are represented as %of input as compared with binding of non-specific IgG.

Figure 4 (Continued)

E



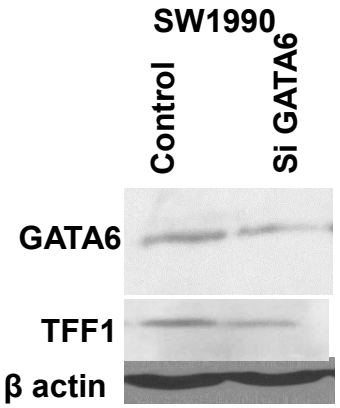
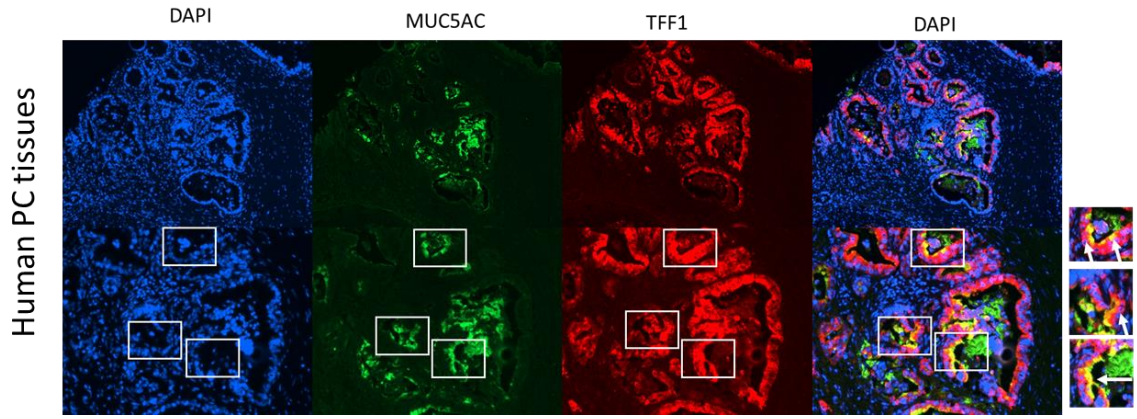


Figure 4 (E): GATA6 regulates TFF1 in SW1990 cells. Expression of TFF1 proteins in the GATA6-silenced SW1990 cell (upper panel) in mRNA and (lower) protein level suggesting repression of GATA-6 also repressed TFF1. β- actin was used as control in both qPCR and western blot experiment**.**

Figure 5

(A)


DAPI
MUC5AC
TFF1
DAPI
Human PC tissues

(B)

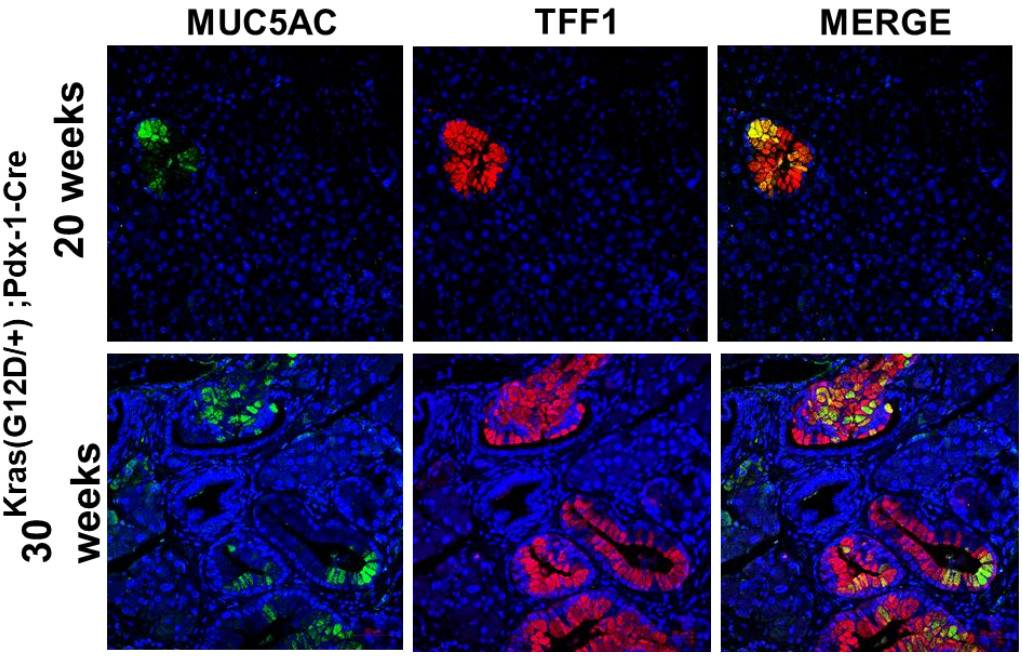
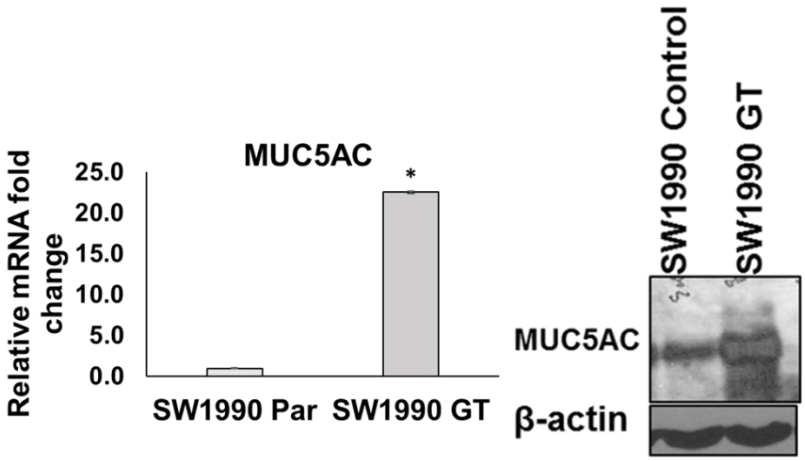

MUC5AC
TFF1
MERGE
Kras(G12D/+) ;Pdx-1-Cre
20 weeks
30 weeks

Figure 5: Colocalization of MUC5AC and TFF1 in PC human and spontaneous mouse tissues. (A) Fluorescence immunohistochemistry staining of PC-TMA with MUC5AC (green), TFF1 (red) antibodies and DAPI (blue) showed coexpression of TFF1 and MUC5AC in the pancreatic duct. **(B)** Fluorescence immunohistochemistry staining of spontaneous PC mouse model tissues (Kras(G12D/+) ; Pdx-1-Cre) from 20 weeks and 30 weeks with MUC5AC (green) , TFF1 (red) antibodies and DAPI (blue) showed coexpression of TFF1 and MUC5AC in mouse PC progression model.

Figure 5 (Continued)

C



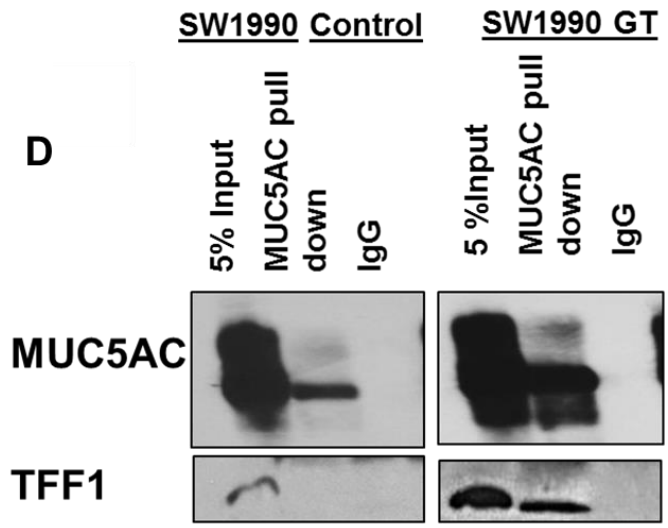
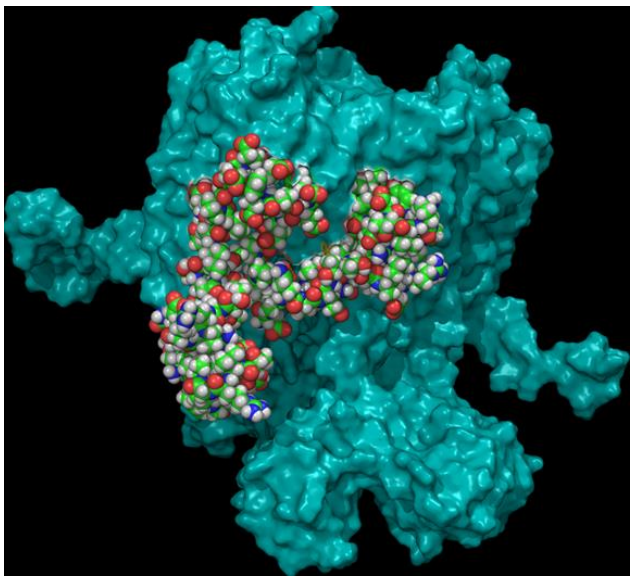


Figure 5: Increase level of MUC5AC in long-term gemcitabine-treated SW1990 PC cells (C) 10 weeks gemcitabine treatment increased MUC5AC expression in mRNA and protein level in PC cells. **(D)** Immunoprecipitation with MUC5AC and immunoblotting with TFF1 in SW1990 parental and SW1990 GT cells show TFF1-MUC5AC interaction in SW1990 GT cells only.

Figure 6

(A)



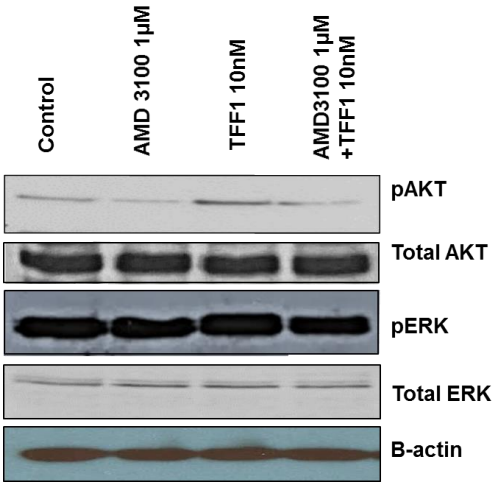
Amino Acid Residue	Specific Interaction	Distance (Å)
Phe-172	pi stacking with Phe-60	3.6
Ser-122	Hydrogen bonding with Glu-55	2.8
Glu-2	Hydrogen bonding with Ile-89	3.0

Figure 6: In Silico analysis of CXCR4 -TFF1 interaction

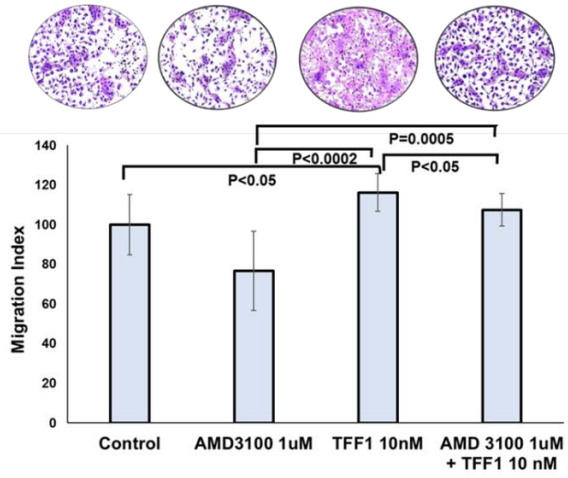
(A) Protein-protein docking analysis reveals the interaction of TFF1 with CXCR4 at Phe-172, Ser-122, and Glu-1 residues through pi-stacking, and hydrogen bonding.

Figure 6

(B)


Control
AMD 3100 1μM
TFF1 10nM
AMD3100 1μM +TFF1 10nM
pAKT
Total AKT
pERK
Total ERK
B-actin

(C)


P=0.0005
P<0.0002
P<0.05
P<0.05
Migration Index
140
120
100
80
60
40
20
0
Control
AMD3100 1uM
TFF1 10nM
AMD 3100 1uM + TFF1 10 nM

(D)

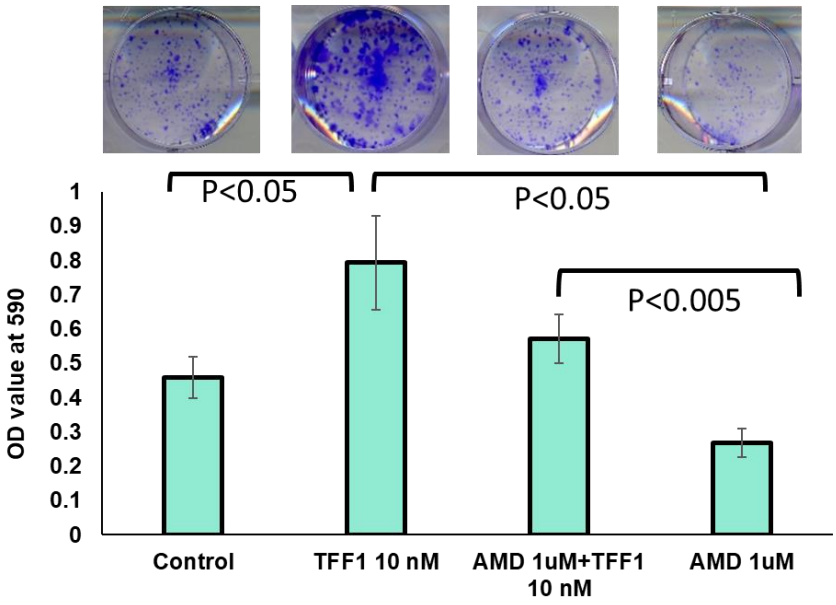
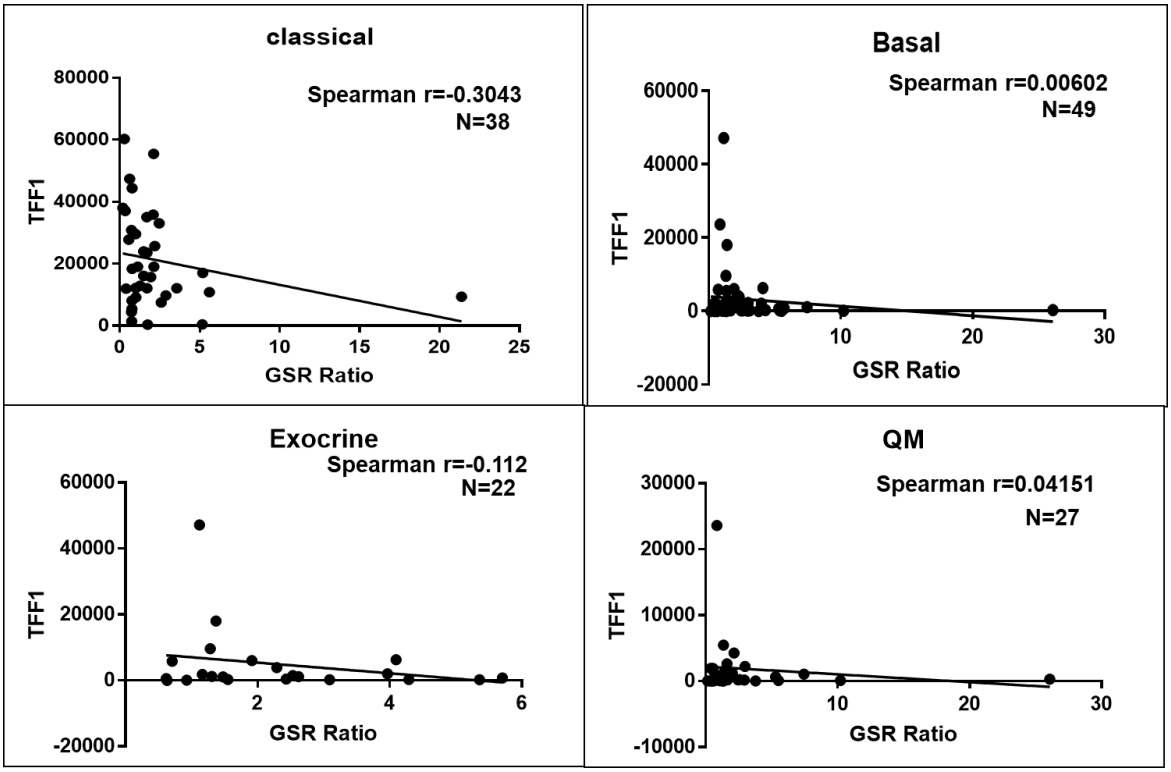

P<0.05
P<0.05
P<0.005
OD value at 590
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
Control
TFF1 10 nM
AMD 1uM+TFF1 10 nM
AMD 1uM

Figure 6: CXCR4 antagonist partially abolished TFF1 mediated downstream signaling. (B) After preincubation with AMD3100 (1 μM) for 1 hour and culture in medium with the indicated drug concentrations (1 μM AMD3100, 10 ng/ml TFF1) for 24 hours, cells were subjected to Western blots. Treatment with TFF1 recombinant protein alone and in combination with CXCR4 inhibitor (AMD3100) reveal that TFF1 modulates downstream signaling partly through CXCR4. (C) After preincubation with AMD3100 (1 μM) for 1 hour, SW1990 cells were used in Transwell assays with indicated drug concentration allowed to migrate for 24 hours. Error bars mean ± SD (n=5). (D) AMD3100 partially abolish TFF1 mediated colony formation. Crystal violet was dissolved with DMSO, and optical density was measured. Error bars mean ± SD (n=3).

Supplementary Figure S1

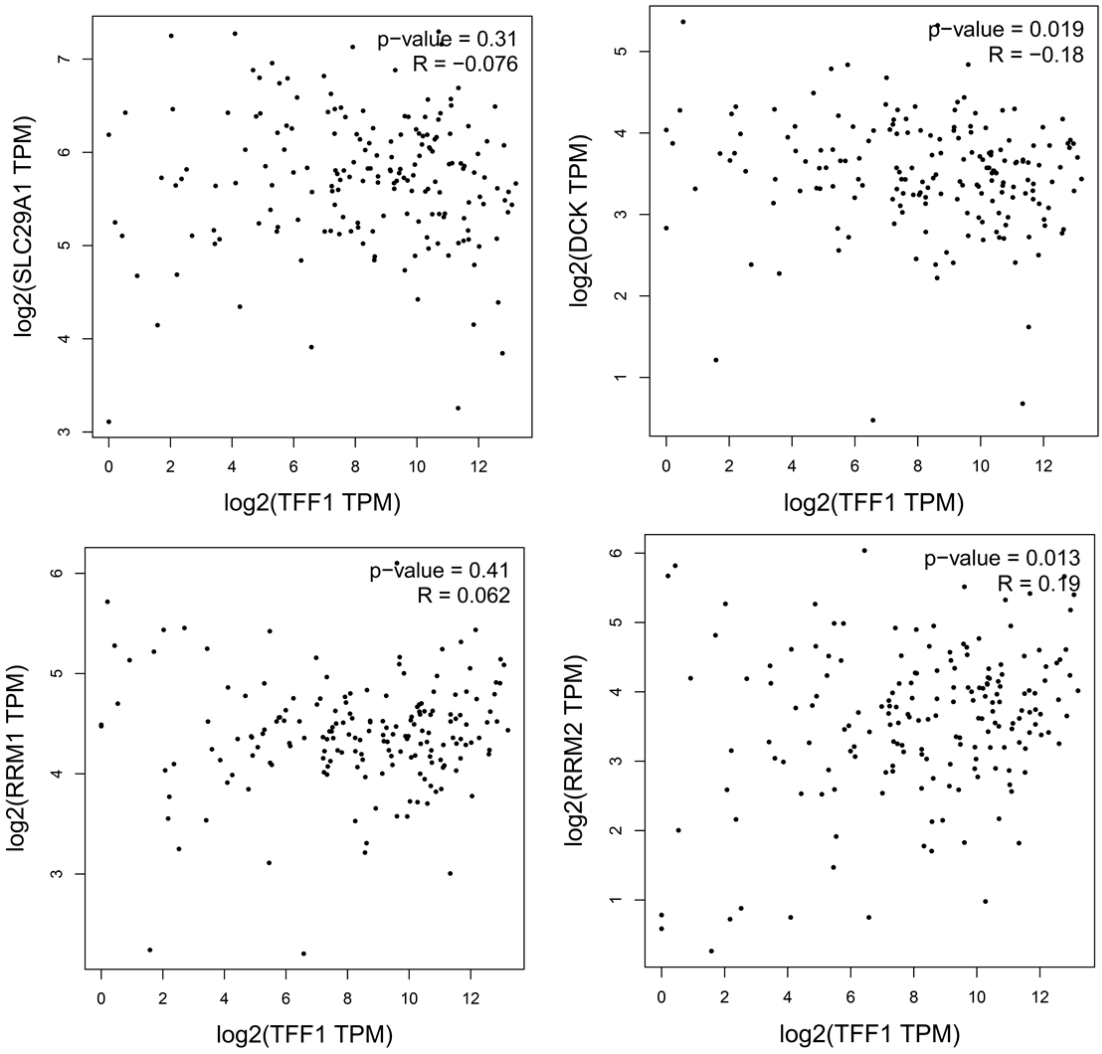
(A)



Supplementary Figure S1 : (A) Spearman correlation analysis of TFF1 and GSR in different PC subtype suggesting the highest negative correlation in classical subtype.

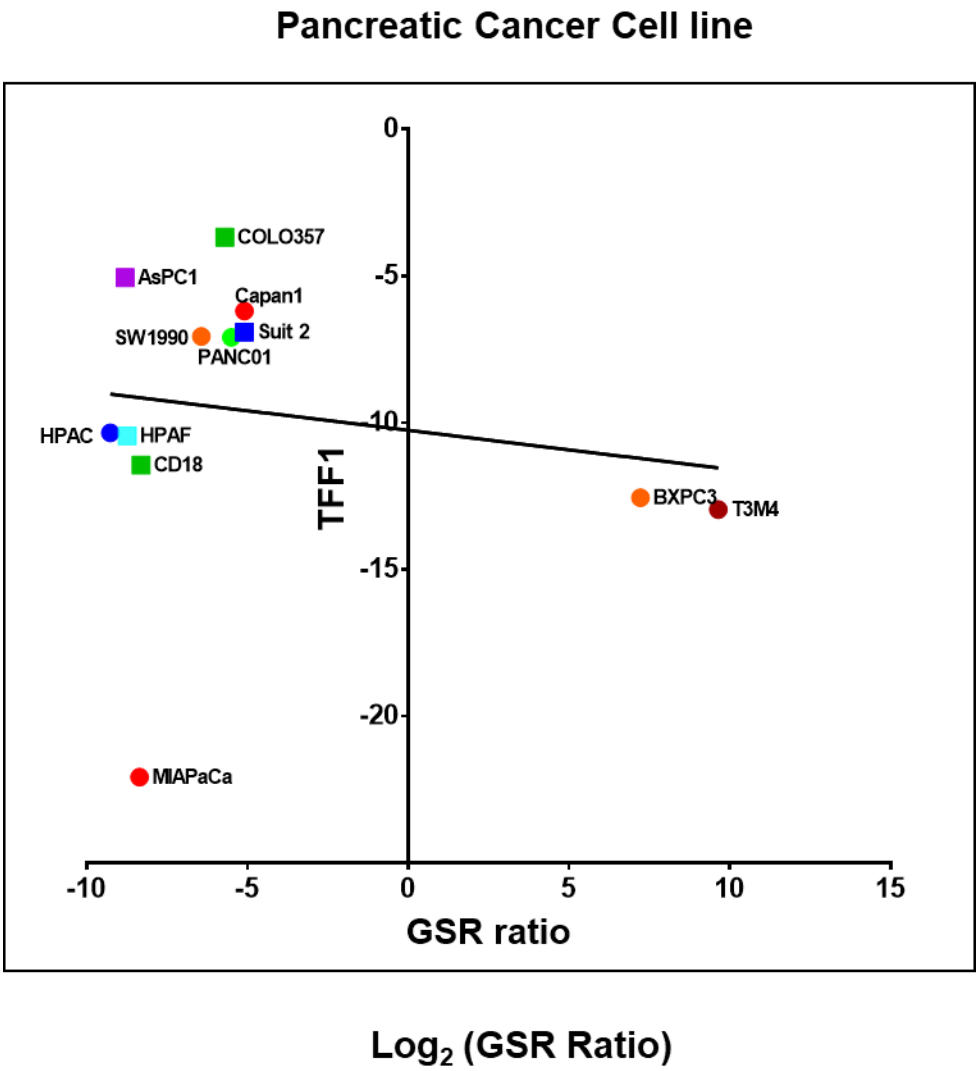
Supplementary Figure S1

(B)



Supplementary Figure S1: (b) Spearman correlation analysis of TFF1 with individual GSR component hENT1, dCK, RRM1 and RRM2 in PC from TCGA database.

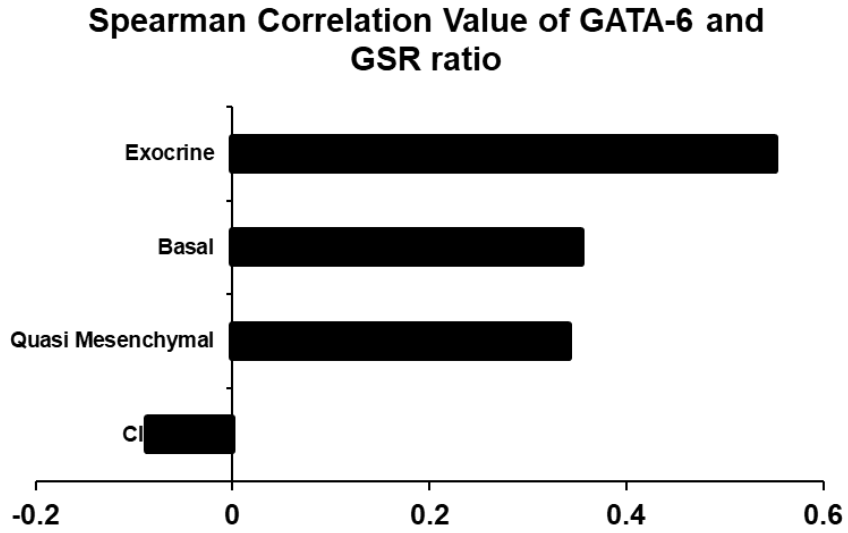
Supplementary Figure S2



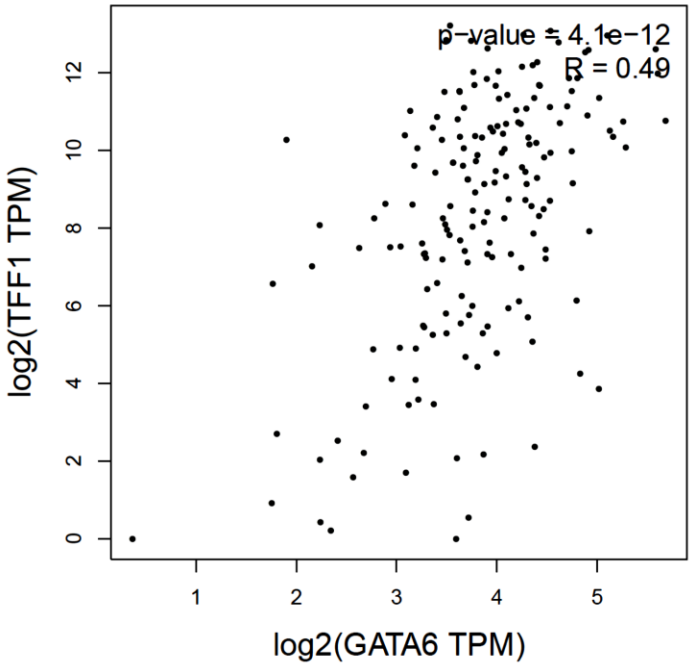
Supplementary Figure S2: Spearman correlation analysis of TFF1 with GSR in PC cell line.

Supplementary Figure S3

A



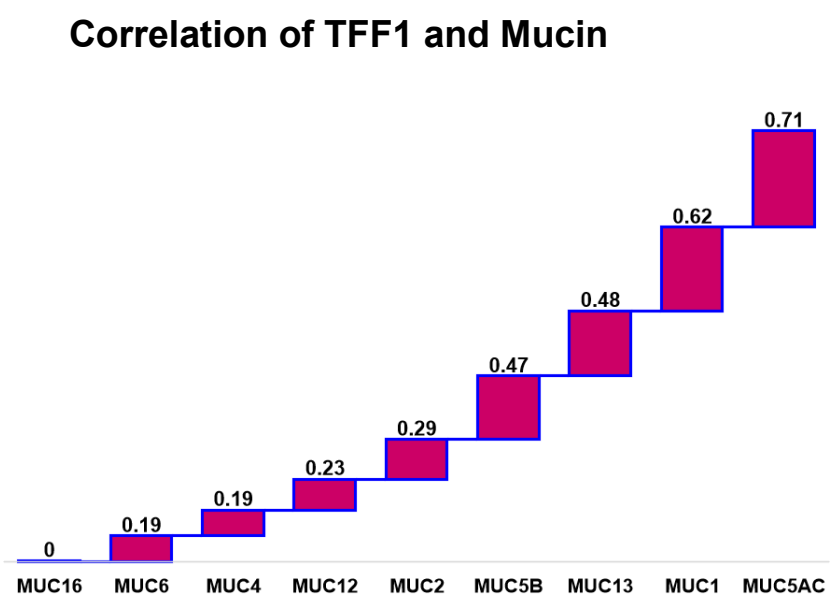
B



Supplementary Figure S3: (A) Spearman correlation between GSR and GATA-6 in different subtype of PC (B) Spearman correlation between GATA-6 and TFF1

Supplementary Figure S4

(A)



(B)

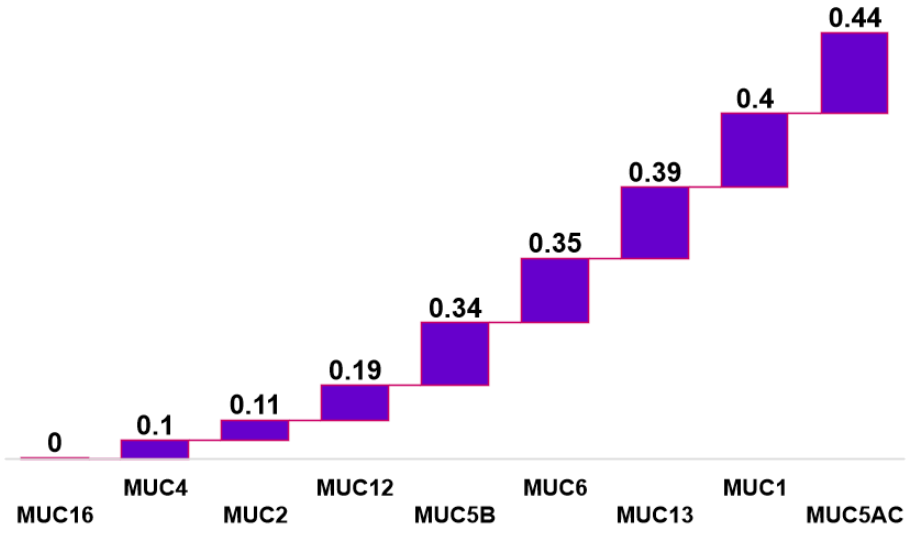


Figure Supplementary S4 : (A) Spearman correlation analysis of MUC5AC with TFF1 and GATA-6

Reference

[1] L. Rahib, B.D. Smith, R. Aizenberg, A.B. Rosenzweig, J.M. Fleshman, L.M. Matrisian, Projecting cancer incidence and deaths to 2030: the unexpected burden of thyroid, liver, and pancreas cancers in the United States, Cancer research, 74 (2014) 2913-2921.

[2] E.A. Collisson, A. Sadanandam, P. Olson, W.J. Gibb, M. Truitt, S. Gu, J. Cooc, J. Weinkle, G.E. Kim, L. Jakkula, H.S. Feiler, A.H. Ko, A.B. Olshen, K.L. Danenberg, M.A. Tempero, P.T. Spellman, D. Hanahan, J.W. Gray, Subtypes of pancreatic ductal adenocarcinoma and their differing responses to therapy, Nat Med, 17 (2011) 500-503.

[3] R.A. Moffitt, R. Marayati, E.L. Flate, K.E. Volmar, S.G. Loeza, K.A. Hoadley, N.U. Rashid, L.A. Williams, S.C. Eaton, A.H. Chung, J.K. Smyla, J.M. Anderson, H.J. Kim, D.J. Bentrem, M.S. Talamonti, C.A. Iacobuzio-Donahue, M.A. Hollingsworth, J.J. Yeh, Virtual microdissection identifies distinct tumor- and stroma-specific subtypes of pancreatic ductal adenocarcinoma, Nat Genet, 47 (2015) 1168-1178.

[4] D.R. Radiloff, T.P. Wakeman, J. Feng, S. Schilling, E. Seto, X.F. Wang, Trefoil factor 1 acts to suppress senescence induced by oncogene activation during the cellular transformation process, Proc Natl Acad Sci U S A, 108 (2011) 6591-6596.

[5] M. Soutto, D. Peng, A. Katsha, Z. Chen, M.B. Piazuelo, M.K. Washington, A. Belkhiri, P. Correa, W. El-Rifai, Activation of beta-catenin signalling by TFF1 loss promotes cell proliferation and gastric tumorigenesis, Gut, 64 (2015) 1028-1039.

[6] J.K. Perry, N. Kannan, P.M. Grandison, M.D. Mitchell, P.E. Lobie, Are trefoil factors oncogenic?, Trends Endocrinol Metab, 19 (2008) 74-81.

[7] A. Chakraborty, K.A. Dorsett, H.Q. Trummell, E.S. Yang, P.G. Oliver, J.A. Bonner, D.J. Buchsbaum, S.L. Bellis, ST6Gal-I sialyltransferase promotes chemoresistance in pancreatic ductal adenocarcinoma by abrogating gemcitabine-mediated DNA damage, J Biol Chem, 293 (2018) 984-994.

[8] a.a.d.h.e. Cancer Genome Atlas Research Network. Electronic address, N. Cancer Genome Atlas Research, Integrated Genomic Characterization of Pancreatic Ductal Adenocarcinoma, Cancer cell, 32 (2017) 185-203 e113.

[9] B.K. Samulitis, K.W. Pond, E. Pond, A.E. Cress, H. Patel, L. Wisner, C. Patel, R.T. Dorr, T.H. Landowski, Gemcitabine resistant pancreatic cancer cell lines acquire an invasive phenotype with collateral hypersensitivity to histone deacetylase inhibitors, Cancer Biol Ther, 16 (2015) 43-51.

[10] Y. An, J. Yao, J.S. Wei, Z.P. Lu, H.H. Cai, C.C. Dai, Z.Y. Qian, Z.K. Xu, Y. Miao, [Establish a gemcitabine-resistant pancreatic cancer cell line SW1990/GZ and research the relationship between SW1990/GZ and pancreatic cancer stem cell], Zhonghua Wai Ke Za Zhi, 48 (2010) 999-1003.

[11] Z. Tang, C. Li, B. Kang, G. Gao, C. Li, Z. Zhang, GEPIA: a web server for cancer and normal gene expression profiling and interactive analyses, Nucleic Acids Res, 45 (2017) W98-W102.

[12] Z. Dubeykovskaya, A. Dubeykovskiy, J. Solal-Cohen, T.C. Wang, Secreted trefoil factor 2 activates the CXCR4 receptor in epithelial and lymphocytic cancer cell lines, J Biol Chem, 284 (2009) 3650-3662.

[13] B.R. Hall, A. Cannon, P. Atri, C.S. Wichman, L.M. Smith, A.K. Ganti, C. Are, A.R. Sasson, S. Kumar, S.K. Batra, Advanced pancreatic cancer: a meta-analysis of clinical trials over thirty years, Oncotarget, 9 (2018) 19396-19405.

[14] Y. Nakano, S. Tanno, K. Koizumi, T. Nishikawa, K. Nakamura, M. Minoguchi, T. Izawa, Y. Mizukami, T. Okumura, Y. Kohgo, Gemcitabine chemoresistance and molecular markers associated with gemcitabine transport and metabolism in human pancreatic cancer cells, Br J Cancer, 96 (2007) 457-463.

[15] L.M. Heiser, A. Sadanandam, W.L. Kuo, S.C. Benz, T.C. Goldstein, S. Ng, W.J. Gibb, N.J. Wang, S. Ziyad, F. Tong, N. Bayani, Z. Hu, J.I. Billig, A. Dueregger, S. Lewis, L. Jakkula, J.E. Korkola, S. Durinck, F. Pepin, Y. Guan, E. Purdom, P. Neuvial, H. Bengtsson, K.W. Wood, P.G. Smith, L.T. Vassilev, B.T. Hennessy, J. Greshock, K.E. Bachman, M.A. Hardwicke, J.W. Park, L.J. Marton, D.M. Wolf, E.A. Collisson, R.M. Neve, G.B. Mills, T.P. Speed, H.S. Feiler, R.F. Wooster, D. Haussler, J.M. Stuart, J.W. Gray, P.T. Spellman, Subtype and pathway specific responses to anticancer compounds in breast cancer, Proc Natl Acad Sci U S A, 109 (2012) 2724-2729.

[16] P. Martinelli, E. Carrillo-de Santa Pau, T. Cox, B. Sainz, Jr., N. Dusetti, W. Greenhalf, L. Rinaldi, E. Costello, P. Ghaneh, N. Malats, M. Buchler, M. Pajic, A.V. Biankin, J. Iovanna, J. Neoptolemos, F.X. Real, GATA6 regulates EMT and tumour dissemination, and is a marker of response to adjuvant chemotherapy in pancreatic cancer, Gut, 66 (2017) 1665-1676.

[17] W.B. Chen, F.T. Huang, Y.Y. Zhuang, J. Tang, X.H. Zhuang, W.J. Cheng, Z.Q. Gu, S.N. Zhang, Silencing of GATA6 suppresses SW1990 pancreatic cancer

cell growth in vitro and up-regulates reactive oxygen species, Dig Dis Sci, 58 (2013) 2518-2527.

[18] E.D. Al-azzeh, P. Fegert, N. Blin, P. Gott, Transcription factor GATA-6 activates expression of gastroprotective trefoil genes TFF1 and TFF2, Biochim Biophys Acta, 1490 (2000) 324-332.

[19] N. Jonckheere, A. Vincent, M. Perrais, M.P. Ducourouble, A.K. Male, J.P. Aubert, P. Pigny, K.L. Carraway, J.N. Freund, I.B. Renes, I. Van Seuningen, The human mucin MUC4 is transcriptionally regulated by caudal-related homeobox, hepatocyte nuclear factors, forkhead box A, and GATA endodermal transcription factors in epithelial cancer cells, J Biol Chem, 282 (2007) 22638-22650.

[20] N. Jonckheere, A. Vincent, H. Franquet-Ansart, J. Witte-Bouma, A. Korteland-van Male, E. Leteurtre, I.B. Renes, I. Van Seuningen, GATA-4/-6 and HNF-1/-4 families of transcription factors control the transcriptional regulation of the murine Muc5ac mucin during stomach development and in epithelial cancer cells, Biochim Biophys Acta, 1819 (2012) 869-876.

[21] S. Kaur, L.M. Smith, A. Patel, M. Menning, D.C. Watley, S.S. Malik, S.R. Krishn, K. Mallya, A. Aithal, A.R. Sasson, S.L. Johansson, M. Jain, S. Singh, S. Guha, C. Are, M. Raimondo, M.A. Hollingsworth, R.E. Brand, S.K. Batra, A Combination of MUC5AC and CA19-9 Improves the Diagnosis of Pancreatic Cancer: A Multicenter Study, Am J Gastroenterol, 112 (2017) 172-183.

[22] I. Lakshmanan, S. Rachagani, R. Hauke, S.R. Krishn, S. Paknikar, P. Seshacharyulu, S. Karmakar, R.K. Nimmakayala, G. Kaushik, S.L. Johansson, G.B. Carey, M.P. Ponnusamy, S. Kaur, S.K. Batra, A.K. Ganti, MUC5AC

interactions with integrin beta4 enhances the migration of lung cancer cells through FAK signaling, Oncogene, 35 (2016) 4112-4121.

[23] W. Wattanawongdon, C. Hahnvajanawong, N. Namwat, S. Kanchanawat, T. Boonmars, P. Jearanaikoon, C. Leelayuwat, A. Techasen, W. Seubwai, Establishment and characterization of gemcitabine-resistant human cholangiocarcinoma cell lines with multidrug resistance and enhanced invasiveness, Int J Oncol, 47 (2015) 398-410.

[24] Y. Jia, J. Xie, Promising molecular mechanisms responsible for gemcitabine resistance in cancer, Genes Dis, 2 (2015) 299-306.

[25] A.N. Shah, J.M. Summy, J. Zhang, S.I. Park, N.U. Parikh, G.E. Gallick, Development and characterization of gemcitabine-resistant pancreatic tumor cells, Ann Surg Oncol, 14 (2007) 3629-3637.

[26] E. Romano, M. Vllahu, V. Bizzarro, R. Belvedere, R. Esposito, A. Petrella, A. Tosco, TFF1 Promotes EMT-Like Changes through an Auto-Induction Mechanism, Int J Mol Sci, 19 (2018).

[27] S.M. Karam, C. Tomasetto, M.C. Rio, Trefoil factor 1 is required for the commitment programme of mouse oxyntic epithelial progenitors, Gut, 53 (2004) 1408-1415.

[28] S. Pelden, T. Insawang, C. Thuwajit, P. Thuwajit, The trefoil factor 1 (TFF1) protein involved in doxorubicininduced apoptosis resistance is upregulated by estrogen in breast cancer cells, Oncol Rep, 30 (2013) 1518-1526.

[29] J. Dieckow, W. Brandt, K. Hattermann, S. Schob, U. Schulze, R. Mentlein, P. Ackermann, S. Sel, F.P. Paulsen, CXCR4 and CXCR7 Mediate TFF3-Induced Cell

Migration Independently From the ERK1/2 Signaling Pathway, Invest Ophthalmol Vis Sci, 57 (2016) 56-65.

[30] R.L. Sleightholm, B.K. Neilsen, J. Li, M.M. Steele, R.K. Singh, M.A. Hollingsworth, D. Oupicky, Emerging roles of the CXCL12/CXCR4 axis in pancreatic cancer progression and therapy, Pharmacol Ther, 179 (2017) 158-170.

[31] S. Singh, S.K. Srivastava, A. Bhardwaj, L.B. Owen, A.P. Singh, CXCL12-CXCR4 signalling axis confers gemcitabine resistance to pancreatic cancer cells: a novel target for therapy, Br J Cancer, 103 (2010) 1671-1679.

[32] M.P. Wescott, I. Kufareva, C. Paes, J.R. Goodman, Y. Thaker, B.A. Puffer, E. Berdougo, J.B. Rucker, T.M. Handel, B.J. Doranz, Signal transmission through the CXC chemokine receptor 4 (CXCR4) transmembrane helices, Proc Natl Acad Sci U S A, 113 (2016) 9928-9933.

Chapter 6

Summary, Conclusion and Future Direction

Summary

Pancreatic cancer (PC) is the third leading cause of cancer-related death in the USA. It will be a second leading cause of cancer-related death shortly after 2020 [1]. Moreover, in the coming future between the years 2010 and 2030, the incidence and death toll associated with PC are expected to increase by 105% and 71%, respectively [1]. The major reasons behind the disastrous rising statistics of PC lack highly sensitive and specific markers for early detection, failure of chemotherapy for rapid development of chemoresistance and gap in knowledge regarding the role of genetic, epigenetic changes and alternative splicing in PC progression. Over the last five years, the overall goal of my projects has been evolved into the following three aspects of PC so that it can help in developing personalized therapeutic intervention and early diagnosis of PC in future.

Goal-1: Contributing in the understanding of complex molecular mechanism of PC by determining of the functional significance of mucin splice variant, MUC4/X in PC pathogenesis. In this part of my thesis, I have delineated the role of a splice variant of mucin4 coined as MUC4/X for the first in time in PC pathogenesis.

Goal-2: Contributing in the identification of novel and effective biomarker panel for early detection of PC by evaluating the diagnostic potential of Trefoil Factors (TFFs) in PC. In this segment of my thesis, I have identified a novel biomarker panel consists of TFFs (TFF1, TFF2, and TFF3) and CA19.9 in diagnosing early stage of PC.

Goal-3. Contributing towards unwinding the mechanisms of PC chemoresistance by delineating the functional and mechanical role of Trefoil

Factor 1(TFF1) in gemcitabine resistance of PC. In this part of my thesis, I have determined the role of TFF1 by analyzing publicly available cancer genome dataset, dissecting transcriptomic and signaling pathways and identification of biochemical interaction of TFF1 with mucin.

The summary of these three projects is discussed below individually.

Part 1. To delineate the functional and molecular mechanism of MUC4/X in PC pathogenesis.

Alternative splicing is the way of gene regulation to provide molecular diversity and controlled regulation in the cell. However, it is emerging as a central hallmark of oncogenic signaling for tumor development, progression, and metastasis. An accumulating body of work demonstrated the critical role of splice variants in cancer pathogenesis.

MUCIN 4 (MUC4), a type I, membrane-bound mucin, is differentially expressed in pancreatic cancer (PC) and plays a crucial role in PC progression and metastasis [2]. Our lab and others have identified 24 splice variants of MUC4. However, information on their expression and molecular implications in the PC Pathogenesis is still unexplored [3, 4]. Among all of them, MUC4/X is devoid of exons 2 encoding for highly glycosylated tandem repeat (TR) domain and exon 3. Thus, it makes a unique structure as it is devoid of largest TR domain yet retains all other functional domain like NIDO, AMOP, vWd, EGF, transmembrane and cytoplasmic domains. As splice variants have shown to have the oncogenic potential, we aim to identify the pathological and molecular significance of splice variant MUC4/X alone or in presence with WT-MUC4 in PC.

First, to identify its clinical significance in PC, we have analyzed the MUC4/X expression with specific primer in clinical tissue samples which comprised of normal adjacent tissues and pancreatic cancer tissues samples. We have isolated RNA from tissues, prepared cDNA and performed qPCR with MUC4/X specific primer. Our results revealed significant upregulation of MUC4/X in PC clinical samples with most differential expression in poorly differentiated tumors tissues while no expression in normal adjacent pancreatic tissues.

Next, to determine the functional significance of MUC4/X without any influence of wild-type (Wild-type) WT-MUC4, we cloned MUC4/X from our previously generated mini-MUC4 construct (ref) and overexpressed it in non-WT MUC4 expressing MIAPaCa and AsPC-1 PC cell line. As we don't have MUC4/X specific antibody, we confirmed the expression with Flag and HA-specific antibody which are flanked on C and N terminus of MUC4/X sequence in the vector. As WT-MUC4 is involved in cell proliferation, invasion, tumorigenesis, and metastasis, we were also curious to know if MUC4/X has role in PC tumorigenicity. To delineate its role in PC we performed cell proliferation assay using MTT, EdU incorporation assay, wound healing assay, invasion assay and adhesion assay. Interestingly in the absence of WT-MUC4, overexpression of MUC4/X in endogenous, wild-type-MUC4 (WT-MUC4) negative PC cell line markedly increased cell proliferation, invasion, and adhesion to extracellular matrix (ECM) proteins. To recapitulate in vitro studies, we performed in vivo orthotopic translation studies where we injected control and MIAPaCa-MUC4/X PC cells into the head of the pancreas. After 50 days of implantation, we sacrificed the mice and observed significant higher tumor

growth and metastasis in MIAPaCa-MUC4/X bearing mice as compared to control mice. We observed significant metastasis in the peritoneal cavity, stomach, kidney and mesenteric cavity in MIAPaCa-MUC4/X bearing mice. Moreover, overexpression of MUC4/X significantly increased the attachment of PC cells to the peritoneal wall which is one of the primary organs for PC metastasis.

As we observed, in PC clinical sample MUC4/X is expressed with WT-MUC4, so we next aim to analyze its functional role in the presence of WT-MUC4. To achieve this, we have generated inducible tet-on system where we can induce MUC4/X with doxycycline in PC cell. Interestingly, doxycycline-induced expression of MUC4/X in an endogenous WT-MUC4 expressing PC cell line (Capan-1) also resulted in enhanced cell proliferation, invasion, and adhesion to ECM emphasizing its direct involvement in enhancing the aggressive behavior of tumor cells.

To understand the molecular mechanism of MUC4/X mediated increased PC tumorigenicity, we analyzed expression of integrin β1 and its downstream signaling as our invitro studies suggested that MUC4/X enhanced increased invasion through extracellular matrix and adhered more with fibronectin, vitronectin which are known ligand for integrin -β1. Integrin-β1 facilitates both intracellular signaling as well as the establishment of a physical link to the ECM which is a well-established role for integrins (ref). Interestingly our analysis also suggested that overexpression of MUC4/X has resulted in increased integrin-β1 expression as well as its downstream signaling pFAK and pERK. Moreover, upregulation of integrin-β1 is one of the primary mechanisms for peritoneal metastasis [5]. Thus,

we concluded that MUC4/X facilitated PC tumorigenesis via triggering integrin-β1/FAK/ERK signaling pathway. Additionally, as our EdU cell proliferation assay showed higher EdU incorporation which is generally incorporated in DNA synthesis S phase, we also analyzed the expression of cyclin A2, a well-established cyclin to promote S-phase entry in cell cycle and also considered as cell proliferation marker [6]. Our analysis revealed increased expression of cyclin A2 in MUC4/X overexpressing PC cell line as compared to control cells.

In a nutshell of this part of my thesis, findings mentioned above indicated the novel role of splice variant MUC4/X in enhancing the oncogenic features of PC.

Part 2. To evaluate the diagnostic potential of Trefoil Factors in PC

Pancreatic cancer (PC) is an aggressive disease with a five-year overall survival rate of <8%. While the five-year survival rate of patients with localized PC is 34.3%, unfortunately, only 10% of total PC patients are diagnosed at an early stage. Approximately 52% of cases are diagnosed at late/metastasized stage, with a worsened five-survival rate of only 2.7% (2). Considering these dire statistics, when developing early detection is the key to improved PC patient survival, it becomes the most attention-grabbing research area to focus on which may lead to enhance patient prognosis.

All member of Trefoil Family (TFF1, TFF2, and TFF3) has emerged as a potential biomarker in a compendium which accumulated potential secretory and membranous proteins as biomarker candidates for PC that was needed experimental validation. Surprisingly, even after a decade of report which indicated potential of TFFs as biomarker, the potential of clinical significance of TFFs as

diagnostic marker was never explored. They are small, secretory mucin-associated proteins known to protect epithelial cells from various environmental insults. Although under physiological conditions they protect the gastric mucosa from inflammation, the oncogenic role of TFFs has been observed in multiple malignancies, including breast, prostate, ovarian, and colon cancers [7]. Because of their secretory nature, their high resistance to proteolytic digestion, acid, and heat degradation which qualify them as advantageous from a biomarker perspective, in this part of thesis, we aimed to comprehensively explore the diagnostic potential of all the member of trefoil family, i.e. TFF1, TFF2, and TFF3 in combination with CA19.9 for detection of PC.

We first analyzed gene expression of Trefoil factors (TFFs) in publicly available cancer genome datasets, and our analysis revealed significantly increased expression of TFF1, TFF2, and TFF3 in pancreatic intraepithelial neoplasia (PanINs) and human PC tissues. Next, we assessed their expression in genetically engineered spontaneous mouse model (GEM) of PC (KrasG12D; Pdx1-Cre (KC)) and in human tissue microarray consisting of normal pancreas adjacent to tumor (NAT), precursor lesions (PanIN), and various pathological grades of PC by immunohistochemistry (IHC). Assessment of KC mouse model suggested upregulated expression of TFFs in PanIN lesions and early stage of PC.

As our in-silico analysis, as well as tissue analysis, suggested their upregulation in precursor lesion of PC as well early stages of PanIN. Next, we evaluated, serum TFFs and CA19.9 levels in comprehensive sample set (n= 377) comprising of independent training and validation set using ELISA comprised of benign controls

(BC), chronic pancreatitis (CP), and various stages of PC. In serum analyses studies, TFF1 and TFF2 were significantly elevated in early stages of PC in comparison to benign ($P<0.005$) and CP control group ($P<0.05$) while significant elevation in TFF3 levels were observed in CP group with no further elevation in its level in early stage PC group.

Next, we analyzed their diagnostic potential by utilizing univariate and multivariate logistic regression and receiver operating characteristic curves (ROC) to examine their diagnostic potential both alone and in combination with CA19.9. Our analysis suggested that combination of TFFs with CA19.9 emerged as promising panel for discriminating early stage of PC from BC (AUCTFF1+TFF2+TFF3+CA19.9=0.928) as well as CP (AUCTFF1+TFF2+TFF3+CA19.9 =0.943). Moreover, at 90% specificity (desired for blood-based biomarker panel), TFFs combination improved CA19.9 sensitivity by 10% and 25% to differentiate early stage of PC from BC and CP respectively. Similar findings were observed in an independent validation set proving unique biomarker capabilities of TFFs.

One of the drawbacks of standard CA 19.9 is that 15-20% people do not express CA19.9 and as our finding demonstrates differential correlation between CA19.9 and TFF1-3, we next sought to identify the diagnostic role of TFFs in low expressing CA19.9 (<37U/ml) and high expressing CA19.9 (>37U/ml) PC patient samples, assuming a likely possibility that low CA19.9 PC patients are Lewis negative. We grouped the patients based on the well-established and recommended cut-off value for CA19.9, 37U/ml. We found that a combination of TFF1-3 can better discriminate PC from CP in low CA19.9 expressing group, AUC-

0.815, than in high CA19.9 expressing group, AUC 0.728. (Figure 4B). The ability to discriminate between EPC and CP was also improved in low vs. high expressing CA19.9 groups, AUC 0.865 vs. 0.712, and SN/SP 1/0.615 vs. 0.915/0.462 (Figure 4B, Supplementary table S5, and S6). As demonstrated with these correlation and ROC curve results, a combination of TFF1-3 can complement CA19.9 to determine PC status

In a nutshell, in silico as well as tissue and serum analyses validated significantly increased level of all TFFs in precursor lesions as well as early stages of PC. The combination of TFFs enhanced sensitivity and specificity of CA19.9 to discriminate early stage of PC from benign controls and CP.

Part 3: To dissect the functional and mechanical aspects of TFF1 in gemcitabine resistance

Chemotherapeutic resistance is one of the major causes of Pancreatic Cancer (PC) mortality and identifying factors facilitating chemoresistance could directly impact the overall patient survival. Trefoil factors (TFF1, TFF2, and TFF3) are small secretory molecules that recently have gained significant attention in multiple studies as an integral component of pancreatic cancer (PC) subtype-specific gene signature. TFF1, a member of mucin-associated small secretory molecules, is one of the top overexpressed genes in the classical subtype (gemcitabine-resistant subtype) of PC. Analysis from The Cancer Genome Atlas (TCGA), immunohistochemistry in both human and spontaneous PC mouse tissues, and serum analysis from a large set of clinical samples suggest upregulation of TFF1

in pre-neoplastic lesion and PC. Based on these evidences, we hypothesize that TFF1 plays a critical role in gemcitabine resistance in PC.

We first analyzed TCGA database to determine the correlation of TFF1 and gemcitabine sensitivity ration (GSR), an indicator for gemcitabine sensitivity which is based on the ratio of gemcitabine metabolic gene (hENT1) *(dCK)/(RRM1) *(RRM2). Our analysis from TCGA database analysis has revealed significant, negative correlation between TFF1 and GSR suggesting TFF1 is positively correlated with gemcitabine resistance.

Next, we treated SW1990 and Colo357, two moderate gemcitabine sensitive PC cell with increasing concentration of for approximately 8 weeks. Interestingly, long-term treatment with gemcitabine (gem) significantly upregulated TFF1 expression in mRNA and protein level in PC cells suggesting TFF1 might play a role in gemcitabine resistance. Next, TFF1 was knocked down in SW1990 PC cells to understand it role in gem resistance. Our in vitro studies showed that TFF1 KD in SW1990 PC cells induced apoptosis, reduced colony formation capacity and modulated many apoptotic regulators such as BAX, Bcl-2, cleaved caspases in the presence of gemcitabine.

Next, we performed, Chromatin immunoprecipitation (ChIP) to identify binding of transcription factor GATA-6 on TFF1 promoter to reveal what controls TFF1 expression in gemcitabine resistance. We specifically analyze GATA- 6 as it also overexpressed in classical subtype of the gene (ref). Interestingly, 16-fold enrichment of GATA-6 was observed on two distinct TFF1 promoter sites. Moreover, GATA-6 KD repressed expression of TFF1 in mRNA and protein level.

To identify its interacting partner in chemoresistance, we first analyze colocalization of TFF1 with MUC5AC in human and mouse tissues. We specifically analyze interaction between MUC5AC and TFF1 as our correlation analysis from TCGA database has demonstrated the highest correlation of TFF1 and MUC5AC in PC clinical samples. Moreover, earlier studies suggest that TFF1 interacts with vWD domain of MUC5AC in normal physiology, however there is no report one their interaction in case of disease pathogenesis [8]. To our surprise, we have found interaction of TFF1-MUC5AC in our long term treated gemcitabine resistant cell line which suggests that TFF1-MUC5AC interaction is needed for gemcitabine resistance

Recently CXCR4 has emerged as one of the signaling receptor for TFF2 and TFF3, however, there is no report on CXCR4 as signaling receptor for TFF1 [9]. To identify that TFF1 mediates downstream signaling through CXCR4 we used protein-protein docking studies using BioLuminate module to analyze TFF1 interaction with potential receptor CXCR4. Our protein-protein docking studies revealed interaction of TFF1 with CXCR4 at Phe-172, Ser-122 and Glu-1 and treatment of PC with TFF1 recombinant protein increased CXCR4 downstream signaling pathway critical for gemcitabine resistance.

In summary, this part of my thesis suggested that TFF1 plays an essential role in gemcitabine resistance of PC cells which is regulated GATA-6, interacting with MUC5AC and by modulating apoptotic molecules possibly through CXCR4 signaling which needs further validation.

Future Directions

Part 1. To delineate the functional and molecular mechanism of MUC4/X in PC pathogenesis.

While we have taken a small step towards elucidating the role of MUC4/X by demonstrating its role in the tumorigenic potential of PC cells by using both in vitro and in vivo overexpression cell-based model system, there is still lot of gap to elucidate its role in PC pathogenesis fully.

(a) Investigating MUC4/X expression in a large cohort of PC tissues comprised of tumor and metastatic tissues.:

From the expression analysis for MUC4/X in PC sample which was comprised of normal tissues adjacent to PC and different graded PC tissues, I was able to demonstrate their progressive increase in expression level from well differentiated to poorly differentiated tissues. However, our patient cohort was lacking metastatic PC patient sample. As my experimental approaches suggested their role in metastasis, analyzing expression level in metastatic tissues will be the next logical step to identify its biological role in PC.

(b) Dissecting the molecular mechanism of MUC4/X in peritoneal metastasis

In my orthotopic transplantation studies using MUC4/X overexpression cell system, I have observed significant metastasis in peritoneum which is the major organ for PC metastasis. The mortality rate of patient with peritoneum metastasis is also alarmingly high. Moreover, in my peritoneal adhesion assay, I have found a significant higher number attachment of MUC4/X overexpressed PC cells with immortalized peritoneal cells. Though, I have proposed a possible mechanism as

upregulation of integrin because of overexpression of MUC4/X. However, the pathway is not fully understood. It would be worthy to investigate, whether knockdown of integrin β1 in MUC4/X overexpressed cell system modify peritoneal metastasis after orthotopically transplantation. Further investigation should also be conducted with other MUC4 splice variant specially MUC4/Y, which is also demonstrated to be tumorigenic for PC. It will answer us whether peritoneal metastasis is only MUC4/X specific or it is mediated by any MUC4 splice variant which is devoid of TR like MUC4/Y. Dissecting the MUC4/X mediated signaling axis in peritoneal metastasis will give us valuable information in future for therapeutic target for PC specific peritoneum metastasis.

My studies also suggested that poorly differentiated PC tumor has higher expression of MUC4/X. However, their expression in metastasis patient tissues is not known yet. Further investigation of MUC4/X expression in peritoneal metastasis from PC tissues would be next step to support the role of MUC4/X in peritoneum metastasis. Moreover, identifying their expression in peritoneum metastatic tissues may serve as a diagnostic and prognostic marker for identifying this lethal metastasis of PC.

(d) Identifying potential interacting partner for MUC4/X:

Our lab has previously demonstrated that tumorigenic potential of MUC4 is imparted by interacting with tyrosine kinase receptor HER2 [10]. As my studies indicated that MUC4/X is tumorigenic for PC and also contain all domain of MUC4 except TR, there is a higher probability that it will interact with HER2, however

validation with experimental approaches is needed to be done. Along with identifying its interaction with HER2 by immunoprecipitating Flag-tagged-MUC4/X, mass spectrometry can be utilized to determine potential novel interacting factor for MUC4/X.

(d) Identifying transcription factor which regulates MUC4/X expression in different stage of tumor progression

Transcription and chromatin regulators impact alternative splicing by recruiting splicing components that subsequently influence splicing in nascent transcripts. Further exploration of splicing regulators which are critical in generating alternative splice variants of MUC4 holds the promise for the discovery of mechanisms and networks of MUC4 specific splice variant and its regulator which may have critical role in cancer development [11].

(e.) Role of MUC4/X in the tumor microenvironment

Tumor microenvironment is comprised of an insoluble extracellular matrix (ECM), a stroma composed of fibroblasts, adipocytes, endothelial and resident immune cells, and a multitude of growth factors and cytokines [12]. While accumulating evidence attests that tumor microenvironment plays a key role in cancer progression; however, little is known regarding how alternative splice affect composition of main key player of tumor microenvironment is little known. As my result suggest that MUC4/X has more adhesive property to extracellular matrix, delineating role of MUC4/X in modulating tumor microenvironment would probably tell us how a splice variant plays a critical role in tumor microenvironment with PC progression. To identify role of MUC4/X in tumor microenvironment regulation,

generating MUC4/X transgenic mouse model would be rational experimental approach.

(f) Identifying role of MUC4/X in PC using transgenic mouse model:

As alternative splicing has been demonstrated to be involved in cancer progression, and their function are often unique as compared to their wild-type protein, gene targeting in mice has been used to create in vivo models to study the regulation and consequences of splice variant in cancer. As MUC4/X has been shown to demonstrate to be tumorigenic from my invitro and in vivo studies, I propose that future directions to delineate its role in pancreatic cancer progression would be generation of transgenic mouse model [13]. I have already generated transgene expression cassette (pTet-MUC4/X) under Tet-regulated expression system. It will enable us to induce overexpression of MUC4/X gene in the pancreas under the control of doxycycline. Based on transactivators (tTA/rtTA), the expression of MUC4/X dynamically controlled in pancreas using doxycycline. Briefly, the MUC4/X transgene expression cassette (ptet-miniMUC4-intron-polyA transgene) will be microinjected and generated MUC4/X transgenic mice. The MUC4/X transgenic mice will be crossed with Pdx1-Cre;LSL-tTA mice and further crossing with the floxed KrasG12D animals for elucidating its role in PC progression.

Part 2. To evaluate the diagnostic potential of Trefoil Factors in PC

(a) Validation of TFFs and CA19.9 in large cohort of sample set and multi-institutional cross-validation

I have analyzed the diagnostic performance in training set along with validation in a small cohort of blinded validation set. To utilize the biomarker panel, consist of TFFs and CA19.9 in a clinical setting, validation in a large cohort of patient sample consisting other benign disease control such as colitis, colorectal polyps, intestinal bowel disease where TFFs are shown to be elevated. Subsequent cross validation in validation set of serum of PC from multi-institutional sample would be next future steps.

(b) Determining the clinical utility of TFFs in jaundice patient:

One caveat of gold standard CA19.9 as a biomarker is that it has false positive results in patients with obstructive jaundice [14]. It would be worthy to investigate TFFs expression in with the patient of jaundice as a clinical need for a marker that is not affected by jaundice. A good control group for evaluating the added value of TFFs should be non-carcinoma conditions leading to jaundice.

(c) Determining the clinical utility of TFFs in the aged group

As my studies suggested that TFFs are highly upregulated in aged (>60 years) patients, it would be interesting to study whether TFFs can act as biomarker for aged patients in determining the disease or stages of the disease.

(d) Understanding the role of TFFs in the initiation of PC

TFF expression analysis from immunohistochemistry of normal, PanIN and different stage of PC tissues suggested that TFFs had very high upregulation in PanIN while decreases in the later stages of PC. Though recently, using invivo TFF1 and TFF2 knock out mouse model suggested their tumor suppressive role, however, there is no exploration of TFF3 in the progression of PC using

spontaneous mouse model. Moreover, the mechanistic role of upregulation of TFFs in PanIN and early stage while decrease in late stage of pancreatic cancer is not fully understood. Moreover, how all three TFF affect while the progression of PC development is not fully understood. So, tri transgenic mouse model developed by Thiem S. and further crossing them with PDX-1-Cre; LSL-KrasG12D would be future direction to elucidate their role in early stage of PC which might be challenging but worthy to investigate [15].

(e) Determining the role of TFFs in metastasis of PC

From my experiment and others, TFFs are expressed in metastatic tissues. Contradictory role of TFF1 is conveyed in case of PC. While Radiloff et al. demonstrated its tumorigenic role in PC, Yamaguchi J et al. revealed that TFF1 suppresses epithelial to mesenchymal transition and decrease invasive properties of PC [16]. Thus, defining TFFs role in PC invasion and metastasis is imperative for proper therapeutic intervention.

(e) Role of TFFs in immune modulation and metabolism

While Koike et al. demonstrated activation of TNFα/NF-κB is reported to increase TFF1 transcription in AGS and MKN48 gastric carcinoma cell lines, contradictory findings was observed by Cobler et al. which demonstrated that TFF1 repression through TNF-α activated NF-κB pathway [17, 18]. Moreover, IL-1β or IL-6 demonstrated to negatively modulate TFF1 and TFF2 promoter activity in the HT-29 (colorectal adenocarcinoma) and KATO-III (gastric carcinoma) which may contribute to ulceration and decreased wound healing during inflammatory bowel disease [19]. So, investigating the crosstalk between immune modulation and

TFFs in PC progression from PanIN to invasive stage with TFFs-knock out mouse model will certainly tell us why TFFs are upregulated in precursor lesion and early stage of PC.

Very recent findings suggest that depletion of Tff3 deficient mice utilize glucose from the bloodstream more effectively and affects the metabolism of fatty acids by increased formation of small lipid vesicles [20]. However, our serum analysis suggests upregulation of TFF3 in early stage of PC. As diabetes and PC are interrelated, it would be worthy to investigate TFFs-glucose metabolism/TFFs-lipid metabolism crosstalk in PC. In this case. TFFs knock out mouse model crossed with Pdx-1 Cre; LSL-KRASG12D (KC) and subsequent glucose and insulin tolerance test, analysis of fatty acids would be further experimental approaches.

Part 3: Dissecting functional and mechanical role of TFF1 in gemcitabine resistance

(a) Delineating the mechanistic role of TFFs in erlotinib tolerance

While classical subtype has more resistance towards gemcitabine, it has demonstrated sensitivity to erlotinib, so identifying role of TFF1 along with TFF2 and TFF3 in gemcitabine resistance would be incomplete without proper understanding its role in PC response to tyrosine kinase receptor inhibitor like erlotinib**.** Generation of erlotinib-resistant cell, knockdown of TFFs in the resistant cell line should be the line of experiments to determine TFFs role in overall drug response.

(b) Identifying predictive potential of TFF as biomarker for chemotherapy

Earlier studies suggested that TFF3 as a therapeutic response marker as it is elevated in patients who responded to endocrine therapy and exhibited more specificity and sensitivity as a predictive biomarker than progesterone receptor and estrogen receptor in unstratified metastatic breast cancer patients [21]. Similarly, TFFs are highly upregulated in classical subtype and disease progression and correlated with GSR ratio; we hypothesize that TFFs can be a guide for gemcitabine response in PC patient. We can analyze their expression in pre-treatment and after clinical treatment samples and identify their potential as biomarker.

(c) Elucidating TFF1/2/3-CXCR4 axis in chemoresistance of PC

CXCR4-CXCL12 axis has been demonstrated to play a key role in gemcitabine resistance [22]. Out initial in silico protein-protein interaction data suggested that there is a possibility of TFF1-CXCR4 interaction, however it is necessary to understand how blocking of TFF1-CXCR4 interaction can overcome gemcitabine resistance in PC.

(d) Identify correlation of TFF1 with cancer stemness in chemoresistance of PC

In my initial data, long-term treatment of gemcitabine has demonstrated increased cancer stem cells property as well as increased level of TFF1. Enhancement of cancer stem cell properties is one the characteristics for gemcitabine resistance in PC [23]. Based on these previous data, we hypothesize TFF1 is correlated with cancer stemness property in cancer gemcitabine resistance. Analysis of cancer stem cells/side population isolated from gemcitabine-resistant cell line and

knocking down TFF1 in stem cell/side population potentially unveil the involvement of TFF1-mediated gemcitabine resistance in PC.

Reference

[1] L. Rahib, B.D. Smith, R. Aizenberg, A.B. Rosenzweig, J.M. Fleshman, L.M. Matrisian, Projecting cancer incidence and deaths to 2030: the unexpected burden of thyroid, liver, and pancreas cancers in the United States, Cancer Res.74 (2014) 2913-2921.

[2] A.P. Singh, P. Chaturvedi, S.K. Batra, Emerging roles of MUC4 in cancer: a novel target for diagnosis and therapy, Cancer Res.67 (2007) 433-436.

[3] P. Chaturvedi, A.P. Singh, S.K. Batra, Structure, evolution, and biology of the MUC4 mucin, FASEB J.22 (2008) 966-981.

[4] F. Escande, L. Lemaitre, N. Moniaux, S.K. Batra, J.P. Aubert, M.P. Buisine, Genomic organization of MUC4 mucin gene. Towards the characterization of splice variants, Eur. J. Biochem.269 (2002) 3637-3644.

[5] K. Lessan, D.J. Aguiar, T. Oegema, L. Siebenson, A.P. Skubitz, CD44 and beta1 integrin mediate ovarian carcinoma cell adhesion to peritoneal mesothelial cells, Am. J Pathol.154 (1999) 1525-1537.

[6] W.W. Gu, J. Lin, X.Y. Hong, Cyclin A2 regulates homologous recombination DNA repair and sensitivity to DNA damaging agents and poly(ADP-ribose) polymerase (PARP) inhibitors in human breast cancer cells, Oncotarget.8 (2017) 90842-90851.

[7] J.K. Perry, N. Kannan, P.M. Grandison, M.D. Mitchell, P.E. Lobie, Are trefoil factors oncogenic?, Trends Endocrinol Metab.19 (2008) 74-81.

[8] C. Tomasetto, R. Masson, J.L. Linares, C. Wendling, O. Lefebvre, M.P. Chenard, M.C. Rio, pS2/TFF1 interacts directly with the VWFC cysteine-rich domains of mucins, Gastroenterology.118 (2000) 70-80.

[9] Z. Dubeykovskaya, A. Dubeykovskiy, J. Solal-Cohen, T.C. Wang, Secreted trefoil factor 2 activates the CXCR4 receptor in epithelial and lymphocytic cancer cell lines, J Biol Chem.284 (2009) 3650-3662.

[10] S. Senapati, S. Das, S.K. Batra, Mucin-interacting proteins: from function to therapeutics, Trends Biochem Sci.35 (2010) 236-245.

[11] H. Han, U. Braunschweig, T. Gonatopoulos-Pournatzis, R.J. Weatheritt, C.L. Hirsch, K.C.H. Ha, E. Radovani, S. Nabeel-Shah, T. Sterne-Weiler, J. Wang, D. O'Hanlon, Q. Pan, D. Ray, H. Zheng, F. Vizeacoumar, A. Datti, L. Magomedova, C.L. Cummins, T.R. Hughes, J.F. Greenblatt, J.L. Wrana, J. Moffat, B.J. Blencowe, Multilayered Control of Alternative Splicing Regulatory Networks by Transcription Factors, Mol Cell.65 (2017) 539-553 e537.

[12] M. Giussani, G. Merlino, V. Cappelletti, E. Tagliabue, M.G. Daidone, Tumor-extracellular matrix interactions: Identification of tools associated with breast cancer progression, Semin Cancer Biol.35 (2015) 3-10.

[13] H. Climente-Gonzalez, E. Porta-Pardo, A. Godzik, E. Eyras, The Functional Impact of Alternative Splicing in Cancer, Cell Rep.20 (2017) 2215-2226.

[14] D. Marrelli, S. Caruso, C. Pedrazzani, A. Neri, E. Fernandes, M. Marini, E. Pinto, F. Roviello, CA19-9 serum levels in obstructive jaundice: clinical value in benign and malignant conditions, Am J Surg.198 (2009) 333-339.
[15] S. Thiem, M.F. Eissmann, E. Stuart, J. Elzer, A. Jonas, M. Buchert, M. Ernst, Inducible gene modification in the gastric epithelium of Tff1-CreERT2, Tff2-rtTA, Tff3-luc mice, Genesis.54 (2016) 626-635.
[16] J. Yamaguchi, Y. Yokoyama, T. Kokuryo, T. Ebata, A. Enomoto, M. Nagino, Trefoil factor 1 inhibits epithelial-mesenchymal transition of pancreatic intraepithelial neoplasm, J Clin Invest.128 (2018) 3619-3629.
[17] T. Koike, T. Shimada, Y. Fujii, G. Chen, K. Tabei, T. Namatame, M. Yamagata, A. Tajima, M. Yoneda, A. Terano, H. Hiraishi, Up-regulation of TFF1 (pS2) expression by TNF-alpha in gastric epithelial cells, J Gastroenterol Hepatol.22 (2007) 936-942.
[18] E. Aihara, K.A. Engevik, M.H. Montrose, Trefoil Factor Peptides and Gastrointestinal Function, Annu Rev Physiol.79 (2017) 357-380.
[19] L. Cobler, R. Mejias-Luque, M. Garrido, M. Pera, E. Badia-Garrido, C. de Bolos, Activation of the NF-kB pathway downregulates TFF-1 in gastric carcinogenesis, Virchows Arch.463 (2013) 497-507.
[20] M. Bujak, I.T. Bujak, S. Sobocanec, M. Mihalj, S. Novak, A. Cosic, M.T. Levak, V. Kopacin, B. Mihaljevic, T. Balog, I. Drenjancevic, M.B. Loncar, Trefoil Factor 3 Deficiency Affects Liver Lipid Metabolism, Cell Physiol Biochem.47 (2018) 827-841.
[21] F.E. May, B.R. Westley, TFF3 is a valuable predictive biomarker of endocrine response in metastatic breast cancer, Endocr Relat Cancer.22 (2015) 465-479.
[22] S. Singh, S.K. Srivastava, A. Bhardwaj, L.B. Owen, A.P. Singh, CXCL12-CXCR4 signalling axis confers gemcitabine resistance to pancreatic cancer cells: a novel target for therapy, Br J Cancer.103 (2010) 1671-1679.
[23] F.H. Sarkar, Y. Li, Z. Wang, D. Kong, Pancreatic cancer stem cells and EMT in drug resistance and metastasis, Minerva Chir.64 (2009) 489-500.

Bibliography of Rahat Jahan

- Jahan R, Macha MA, Rachagani S, Das S, Smith LM, Kaur S, and Batra SK. Axed MUC4 (MUC4/X) aggravates pancreatic malignant phenotype by activating integrin-β1/FAK/ERK pathway. BBA-Molecular Basis of Disease, 2018;1864(8): 2538-2549
- Jahan R, Kaur S, Macha M.A, and Batra S.K. Mucins (MUCs). Encyclopedia of Signaling Molecules. Springer, December 2016. (Book Chapter)
- Qazi AK*, Siddiqui JA*, Jahan R*, Chaudhary S, Walker LA, Sayed Z, Jones DT, Batra SK, and Macha MA. Emerging therapeutic potential of graviola and its constituents in cancers. Carcinogenesis. 2018;39(4), 522-533. (*equal contribution).
- Aithal A, Junker W, Kshirsagar P, Das S, Kaur S, Orzechowski C, Gautam SK, Jahan R, Ponnusamy MP, Batra SK, and Jain M. Development and characterization of carboxy-terminus specific monoclonal antibodies for understanding MUC16 cleavage in human ovarian cancer. PLoS One. 2018;13(4): e0193907
- Macha MA, Rachagani S, Qazi A, Jahan R, Gupta S, Patel A, Lin C, Li S, Wang S, Verma V, Kishida S, Kishida M, Nakamura N, Kibe T, Ganti AK, Jones DT, Batra SK, and Jain M. Afatinib radiosensitizes head and neck squamous cell carcinoma cells by effecting cancer stem cells. Oncotarget. 2017;8(13), 20961-20973.
- Kumar S, Cruz E, Joshi S, Patel A, Jahan R, Batra SK, and Jain M. Genetic variants of mucins: unexplored conundrum. Carcinogenesis. 2017;38(7), 671-679

- Macha MA, Krishn SR, Jahan R, Banerjee K, Batra SK, and Jain M. Emerging potential of natural products for targeting mucins for therapy against inflammation and cancer. Cancer Treatment Reviews, 2015. 41(3),277-88
- Jahan R, Smith LM, Sheinin YM, Atri P, Ganguly K, Carmicheal J, Rachagani S, Brand RE, Macha MA, Kaur S, and Batra SK. Trefoil Factors and CA19.9: A promising panel for early detection of pancreatic cancer. (Under revision, EBiomedicine).
- Ernani V, Jahan R, Smith LM, Marr AS, Kimbrough SE, Kos ME, Tijerina J, Pivovar S, Lakshmanan I, Ketcham M, Rauth S, Mallya K, Kessinger A, Nasser MW, Jain M, Batra SK, and Ganti AK. A Phase I study of weekly doxorubicin and oral topotecan for patients with relapsed or refractory small cell lung cancer (SCLC): A Fred and Pamela Buffet Cancer Center Clinical Trials Network study (Submitted to Clinical Lung Cancer)
- Jahan R, Kaur S, Macha MA, and Batra SK. Implications of Trefoil Factors in cancer. (Manuscript near submission).
- Ganguly K, Krishn SR, Rachagani S, Patel A, Jahan R, Shah A, Kaur S and Batra SK. Role of MUC5AC in pancreatic cancer pathogenesis (Manuscript Near Submission).