PhD Degree in Systems Medicine
Curriculum in Molecular Oncology - XXXV cycle
European School of Molecular Medicine (SEMM)
University of Milan

Integrated molecular investigations of MYRF and SFRP2: unveiling regulatory mechanisms of ER stress and ECM remodelling in pancreatic cancer

MED/04

Fabiana Arco

Tutor: Dr. Gioacchino Natoli
European Institute of Oncology (IEO), Milan

PhD Coordinator: Prof. Saverio Minucci

Academic Year 2022-2023

Table of Contents

List of abbreviations

(d)FBS	(dialyzed) Fetal Bovine Serum
BM	Basement Membrane

CAFs	Cancer-Associated Fibroblasts
CNX	CalNeXin
CRD	Cystein-Rich Domain
CRISPR	Cluster Regularly Interspaced Short Palindromic Repeats
CRT	CalReTiculin
DBD	DNA-Binding Domain
ECM	ExtraCellular Matrix
ELISA	Enzyme-Linked ImmunoSorbent Assay
EMT	Epithelial-to-Mesenchymal Transition
ER	Endoplasmic Reticulum
ERAD	ER-Associated Degradation
ERQC	ER Quality Control
FC	Fold Change
FDR	False Discovery Rate
FzD	FrizzleD
GO	Gene Ontology
GSEA	Gene Set Enrichment Analysis
ICA	Intramolecular Chaperon Auto-processing
IP	ImmunoPrecipitation
IPMN	Intraductal Papillary Mucinous Neoplasm
KO	Knock Out
LC-MS	Liquid Chromatography-Mass Spectrometry
LMD	Laser Micro-Dissection
MCN	Mucinous Cystic Neoplasm
MMPs	Matrix MetalloProteinases
MoA	Mode of Action
NLS	Nuclear Localization Signal
NTR	NeTRin-like
PanIN	Pancreatic Intraepithelial Neoplasia
PDAC	Pancreatic Ductal AdenoCarcinoma
PSCs	Pancreatic Stellate Cells
RNA-seq	RNA-sequencing
RT-qPCR	Real Rime-quantitative Polymerase Chain Reaction
SDS-PAGE	Sodium Dodecyl Sulphate-PolyAcrylamide Gel Electrophoresis
SHG	Second Harmonic Generation
TF	Transcription Factor

TGF-β	Transforming Growth Factor beta
TMD	TransMembrane Domain
TME	Tumor MicroEnvironment
UPR	Unfolded Protein Response
WT	Wild Type

Figures index

Tables index

Abstract

Pancreatic ductal adenocarcinoma (PDAC) stands out as the most prevalent type of pancreatic tumors and the deadliest among common solid malignancies. One of the causes of this high lethality is the significant intratumor cell diversity, highlighting the need for enhanced molecular understanding. Indeed, PDACs are ensembles of three different subtypes with distinct morphology and gene expression programs that coexist in variable proportions in every tumor.

Following the identification of transcriptional regulatory networks in the various biotypes, we decided to better characterize two pivotal proteins controlling specific hallmarks of pancreatic cancer biology, like evading growth suppression and activating invasion.

First, we focused on the regulation of the myelin regulatory factor (MYRF), a poorly characterized transcription factor. Previous studies conducted in our laboratory demonstrated that MYRF, an endoplasmic reticulum (ER) transmembrane protein, maintains the homeostasis in the ER of pancreatic secretory glandular cells. After trimerization and self-cleavage, MYRF releases the N-terminal fragment which translocates into the nucleus to exert its transcriptional role by regulating the expression of genes encoding secretory proteins. However, the role of C-terminal fragment which remains inserted in the ER is unclear. Since the regulation of MYRF release from the ER membrane is still unknown, our hypothesis is that this fragment may interact with proteins in the ER lumen and regulate MYRF trimerization and cleavage. Using enzyme-catalyzed proximity labeling followed by biotin pull-down and mass spectrometry, we identified MYRF proximal interactors in the ER lumen. Currently, we were able to obtain a first validation of MYRF interaction with UGGT1, which could provide quality control for MYRF folding.

Furthermore, in order to investigate the role of MYRF C-terminal fragment in the maintenance of ER functionality, we generated clonal cell lines lacking one of the two critical MYRF functional domains – DNA binding or transmembrane domains - by CRISPR/Cas9-mediated in-frame deletion. The analysis of the gene expression profile of MYRF domain-specific mutants revealed that MYRF has a structural functionality in the ER that is independent of its transcriptional activity.

Additionally, we focused on the role of SFRP2 as a potential regulator influencing the characteristics of PDAC transitional cells. One of the key aspects of transitional cancer cells is the ability to contribute to the extracellular matrix (ECM) remodeling: we observed that SFRP2 led to a noteworthy augmentation in this process, along with increased migratory and invasive capabilities of PDAC cells. We also demonstrated that the abundance of ECM components contributes to different tissue density and rigidity, leading cells to mount a mechanoresponsive transcription program.

This work could provide i) a mechanistic understanding of MYRF in the regulation of ER activity. Considering the impact of ER stress on cancer cell fitness and drug-induced toxicity, the possibility to influence ER regulation by modulating MYRF trimerization and cleavage by targeting UGGT1 may provide therapeutic opportunities; ii) a better characterization of SFRP2 and its functional effects on PDAC transitional cells, improving anti-cancer therapies.

Introduction

1. Pancreatic ductal adenocarcinoma

Pancreatic ductal adenocarcinoma (PDAC) stands out as the most prevalent type of pancreatic tumors. It ranks as the third most common cause of cancer-related deaths when considering both men and women, with a consistent annual increase of approximately 1%[1]. Surgically resectable disease is present in less than 20% of patients[2]. All the others are ineligible for surgery due to locally advanced or metastatic disease at the time of diagnosis[3], and standard chemotherapy offers a modest improvement in survival[4,5]. The 5-year survival rate after diagnosis is observed in only approximately 13% of patients[1].

The high lethality of PDAC can be ascribed to different concomitant factors, including the often latent clinical manifestation, wherein symptoms typically arise only after the disease has already advanced significantly. Additionally, the absence of early detection markers and effective treatments contributes to the high mortality

rate of this tumor. Finally, the complex biology of the tumor further compounds the challenges associated with managing and treating PDAC. The biological complexity of the tumor, particularly, is evidenced by the significant intratumor cell diversity, highlighting the need for enhanced molecular understanding and improved classification schemes.

1.1 PDAC development

Numerous risk factors contribute to the development of pancreatic ductal adenocarcinoma. For instance, smoking more than doubles the risk for an individual[6], although unlike other diseases associated with smoking, clear carcinogen-related DNA mutational signatures have not been observed[7]. Family history is a well-established factor influencing the likelihood of PDAC, with 10% of patients having a family history of the disease[8–10]. Moreover, the occurrence of germline mutations in known cancer predisposition genes is higher than expected in PDAC patients without a family history[11,12], posing challenges in interpreting genetic variants and estimating an individual's risk in the absence of relevant family background[13,14].

The majority of PDAC cases has been conceptualized as arising from three types of microscopic preneoplastic lesions: *pancreatic intraepithelial neoplasia* (PanIN), *intraductal papillary mucinous neoplasm* (IPMN) and *mucinous cystic neoplasm* (MCN). PanINs are the most common and extensively studied, described as the lesions of small calibre pancreatic ducts[15,16]. They were initially categorized into four stages, delineated by a progression from moderate dysplasia to high-grade dysplasia with a continuous increase in the number of genetic alterations: PanIN-1 (1A and 1B) and PanIN-2, as hyperplastic lesions characterized by the low-to-intermediate epithelial dysplasia, and PanIN-3, the most dysplastic lesion till the invasive carcinoma[17]. Nowadays, the definition of PanIN is constantly changing. Basturk and colleagues[15] proposed a two-tiered system classifying PanIN into low-grade (including PanIN-1 and PanIN-2) and high-grade (including PanIN-3). Consequently, the high-grade dysplasia is reserved exclusively for the most advanced dysplasia at the upper end of the spectrum, known as “carcinoma in situ” type lesions.

A smaller proportion of PDAC cases develop in the presence of cystic lesions in the pancreas, such as IPMNs and MCNs[15,16], which are also classified into two tiers of dysplasia based on the highest grade of dysplasia detected. PanINs and IPMNs exhibit significant similarities in terms of histological features, including columnar,

mucin-containing cells and varying degrees of papilla formation, but IPMNs later evolve into large cystic structures. MCNs are large cystic lesions with a distinctive tendency to produce mucin[18].

1.2 Transcription factors and epigenetic regulators involved in PDAC development

The primary development and progression of pancreatic cancer are linked to inactivating mutations in several tumor suppressor genes, including KRAS, TP53, DPC4 and CDKN2A, involved in various essential developmental and signalling pathways, such as DNA repair, cell cycle regulation, chromatin remodelling and proliferation[19].

KRAS gene belongs to the RAS family of GTP-binding proteins, participating in signal transduction across the membrane in response to growth factor signals[20]. It plays a crucial role in various cellular functions, such as proliferation, differentiation, cell shape, migration, endocytosis, cell survival and senescence[21]. Large-scale exome sequencing, copy number analysis and the exploration of the genetic landscape of pancreatic cancer have indicated that KRAS mutation is present in more than 95% of pancreatic cancer cases[22]. The most common mutation entails the replacement of glycine with aspartate at codon 12. Less frequent point mutations may also occur at codons 11, 13, 61 and 146[23]. A mutated KRAS results in sustained activation of the KRAS protein, triggering the mitogen-activated protein (MAP) and/or phosphoinositide 3-kinase (PI3K) pathway. This activation, in turn, promotes cell proliferation, migration, transformation and survival[24]. KRAS mutations can be identified in low-grade PanIN, thus representing an early event in the progression of pancreatic cancer[25].

The TP53 gene is a transcription factor (TF) that regulates the G1/S checkpoint, DNA repair and apoptosis in response to various stimuli, such as DNA damage, oncogene activation, hypoxia and stress[26]. It suppresses cell division and survival, thus serving as a crucial mechanism in the defense against cancer[27]. TP53 is found to be mutated in 70% of pancreatic cancer, with the majority of cases predominantly resulting from the loss of a single allele followed by an intragenic mutation in the second allele[23,28]. Mutations in TP53 are linked to the onset of genomic instability, leading to substantial copy number alterations, chromosomal rearrangements (polyploidy) and chromothripsis[23]. Recent studies have indicated a strong association between TP53 mutations and metastasis in PDAC: TP53 plays a role in promoting tumor invasion and the loss of its activity contributes to metastasis by

downregulating genes that inhibit cell migration and epithelial–mesenchymal transition (EMT)[29]. The inactivation of TP53 is a late event in the progression of PDAC, typically detected in advanced PanIN-3 lesions[30,31]. However, TP53 mutations appear to be more common in IPMN[32].

DPC4 is a tumor suppressor gene encoding for SMAD4. This protein is a component of transforming growth factor beta (TGF-β) pathway. The TGF-β pathway orchestrates diverse signalling networks that have dual effects, both promoting and suppressing cancer. On one hand, TGF-β can elevate the expression of p21 (cyclin-dependent kinase inhibitor 1), inhibiting tumor growth, fostering cell differentiation and inducing apoptosis in a manner dependent on SMAD4[33]. On the other hand, as the tumor advances, TGF-β progressively triggers the epithelial–mesenchymal transition of tumor cells[34,35]. The inactivation of DPC4 occurs in 55% of pancreatic cancer, resulting from the loss of function of both alleles[23,36]. Mutations in SMAD4 result in the generation of dysfunctional or truncated SMAD4 proteins, which are subsequently degraded through the ubiquitin–proteasome pathway[23]. As with p53, the loss of SMAD4 also is a late event in tumorigenesis, typically occurring during the advanced stages of PanIN-3 lesions[37].

The CDKN2A is a tumor suppressor gene responsible for encoding proteins that play a role in regulating the G1/S checkpoint, thereby inducing cell cycle arrest[38]. Specifically, it gives rise to two proteins with distinct biological functions: p16 and p14. p16 serves as a cyclin-dependent kinase inhibitor, impeding cell cycle progression from the G1 to S phase by suppressing the interaction between cyclin D1 and CDK4/CDK6[39]. Meanwhile, p14 induces cell growth arrest or apoptosis by hindering MDM2-dependent degradation of the p53 protein[39]. CDKN2A is commonly inactivated in pancreatic cancers through various mechanisms, including homozygous deletions, intragenic mutations, loss of the second wild-type allele and epigenetic silencing through promoter hypermethylation[23]. The loss of CDKN2A function typically takes place during the PanIN-2 stage of cancer development[40].

A comprehensive genome-wide association study, conducted on a cohort comprising thousands of pancreatic cancer patients and matched controls, successfully identified other top-ranking genes involved in the origin and progression of PDAC precursor lesions, including PDX1, SOX9, HES1, NR5A2 and MIST[41]. PDX1 serves as a master transcription factor in pancreas development[42]. In mouse models the knockout of PDX1 during the early stages of tumor progression hinders the transition from PanIN lesions to PDAC. Interestingly, its deletion in established PDAC inhibits tumor cell growth, underscoring the regulatory role of this

transcription factor in controlling differentiation at different stages of tumor progression[43]. Moreover, the simultaneous expression of SOX9 with oncogenic KRAS was found to accelerate the formation of precursor lesions[44].

Mutational profiles regard not only alterations in transcription factors, but also in genes encoding chromatin regulators[45]. In the context of pancreatic cancer, the significance of epigenetic regulators is highlighted by the elevated incidence of somatic mutations affecting this class of proteins. Notably, members of the SWItch/Sucrose NonFermentable (SWI/SNF) chromatin remodeling complex and the histone-lysine N-methyltransferase 2 (KMT2) family exhibit a high somatic mutation rate[46]. Individual members of the SWI/SNF complex display mutation rates below 5%, tending to occur in a mutually exclusive manner. However, when the mutations are collectively considered as contributions by the complex as a whole, they may account for up to 30% of tumor cases in pancreatic cancer[47].

1.3 PDAC intratumor heterogeneity

Morphological heterogeneity is a prevalent observation in pancreatic ductal adenocarcinoma that is not limited to differences between distinct patients, but it is almost invariably observed within the same tumor. Furthermore, it encompasses not only the microscopic characteristics of the tumor cell population, but it also extends to other crucial aspects of the cancer, such as the degree of differentiation and the growth pattern[48].

More specifically, the extensive cellular and architectural intratumor heterogeneity of PDAC is represented by varying proportions of tumor cells with diverse sizes, shapes and subcellular features that coexist in a rich and dense fibrotic stroma. In the stroma the cells are organized into growth patterns and cytological appearances that span from pseudo-glandular structures to compact nests and that can be observed in different regions of the same tumor[48]. RNA-seq data established a morphological threshold, defined as “≥40% non-gland forming”, to identify two distinct morphological patterns: *gland-forming* and *non-gland-forming* [49] (**Figure 1**).

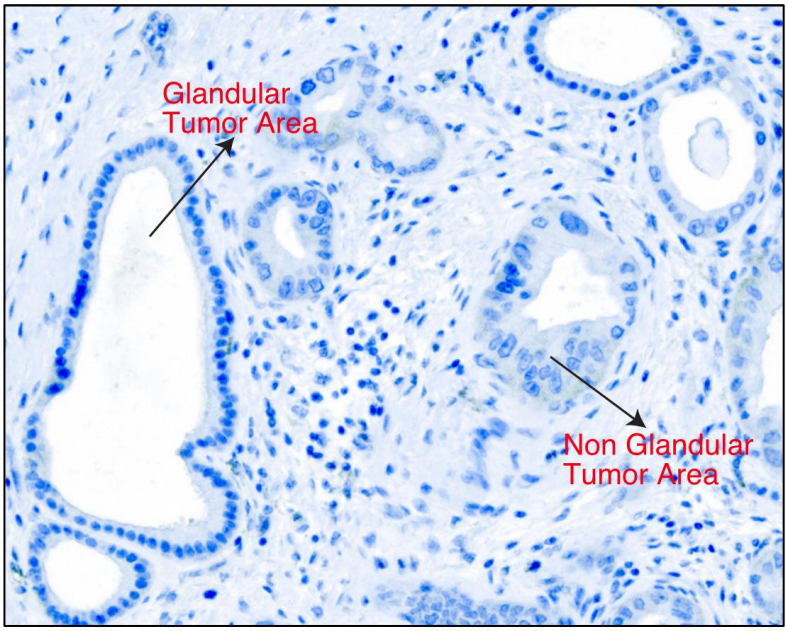

Figure 1: PDAC morphological heterogeneity. Hematoxylin staining of a human PDAC tumor highlights the typical features: tumor cells are organized into well-formed glands (indicated as glandular tumor area) or more compact nests (indicated as non glandular tumor area) and they are surrounded by an important stromal reaction.

Clusters of tumor cells exhibiting diverse levels of differentiation are frequently mixed, with poor differentiation group often marked by the presence of individual infiltrating cancer cells or small solid cell clusters[48]. Microscopically, the infiltrative nature of pancreatic ductal adenocarcinoma is distinguished by an atypically dispersed growth pattern. Indeed, unlike the compact masses of densely packed tumor cells commonly observed in cancers like colorectal cancer, PDAC cell clusters tend to grow more loosely, often separated by significant distances[48].

The high heterogeneity of PDAC cells correlates with poorer overall survival: the occurrence of over 40% of the non-gland-forming component was linked to a poorer prognosis, while cancers with less than 40% of the non-gland-forming component were associated with a more favourable prognosis[50]. This suggests that the coexistence of diverse tumor cells may facilitate rapid adaptation to therapy and/or the selection of aggressive, treatment-resistant tumor cells. Hence, although PDAC intratumor heterogeneity is not typically included in diagnostic pathology reporting, it plays a prominent role in drug resistance and therapeutic failures, holding significant implications for both patient management and research[51,52].

1.4 Molecular characterization of PDAC subtypes

At the molecular level, PDAC intratumor heterogeneity is reflected in the variability of gene expression programs. Bulk transcriptomic data deriving from several studies resulted in the molecular characterization of PDAC subtypes[50,53–56]. According to

these classifications, pancreatic cancer is described by the presence of a *classical* (or pancreatic progenitor) signature and *quasi-mesenchymal, basal* or *squamous* signature. The largest subgroup (*ca.* 80%) of PDAC is represented by the classical signature with main expression of typical endodermal or epithelial differentiation genes, such as GATA6, HNF1B and FOXAs[56–58], and with a better prognosis. Quasi-mesenchymal, basal or squamous signatures lack endodermal lineage fidelity and differentiation, due to the methylation of genes like HNF4A and GATA6[50,55,58], and are associated with high tumor grade and worse clinical outcome. These signatures are partially overlapping yet distinct and their different names are indicative of the most prevalent gene expression program[55,56,59]. Specifically, the quasi-mesenchymal signature encompasses genes associated with epithelial-to-mesenchymal transition (EMT)[60]. In contrast, both the basal and squamous signatures involve regulators of basal cells and genes expressed by their differentiating progeny, featuring keratins that are characteristic of stratified epithelia. A notable characteristic of these subtypes is their association with mutations in genes related to chromatin modification, including DNA methylation and acetylation, such as MLL2, MLL3, and KDM6A[58].

Although several studies classifying PDAC subtypes have predominantly identified shared expression features, bulk transcriptomic analyses involve averaging the transcriptomes from various coexisting tumor cells, obscuring heterogeneity at the single-cell level. Indeed, in this approach each tumor is classified based on the most abundant population, while the contribution of less represented cancer cells is overlooked. The effects of such averaging in a heterogeneous tumor might explain why these classification systems have limited overall clinical impact. A recent study based on the use of single-cell resolution recognized the stratification of malignant states, where individual cells can be grouped based on their similarity in expression subtypes (e.g., basal), while exhibiting variation in their expression of tumor microenvironment-influenced programs, such as inflammatory and drug response[61].

1.5 Microenvironment in pancreatic cancer

In recent years, the heterogeneity of PDAC cells seems to be primarily influenced by local signals originating from the tumor microenvironment (TME)[62,63]. Pancreatic ductal adenocarcinoma is renowned for featuring an extensively studied TME, the complexity of which involves a multitude of elements, ranging from hypoxia and pH gradients to the modulation of cell signalling and interactions between stromal and tumor cells[51,52].

PDAC stroma has been described to be highly heterogeneous, encompassing both structural and functional variations[48]. In comparison to adenocarcinomas in other organs, a histopathological characteristic of PDAC is the presence of a fibrotic desmoplastic stroma, which is evident in both primary and metastatic tumors[64] (**Figure 2**). The heterogeneity of stromal compartment consists of diverse cellular and non-cellular components, including cancer-associated fibroblasts, myofibroblasts, pancreatic stellate cells, immune cells and blood vessels (**Figure 2**). Additionally, the stroma is rich in extracellular matrix (ECM) components and various soluble proteins with signalling activity, such as cytokines and growth factors[65,66].

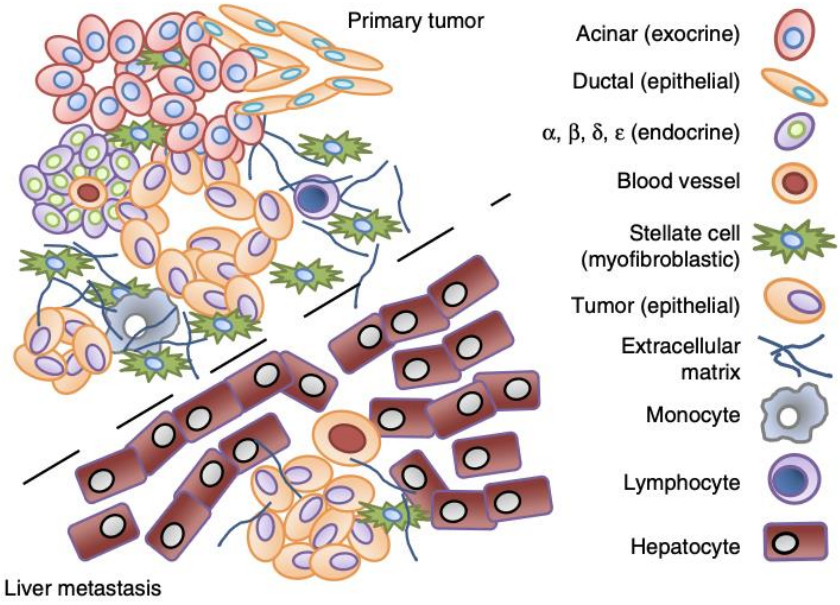

Figure 2: PDAC stroma components. Schematic representation of the major cell types in primary and metastatic PDAC. Moffitt R.A. et al., *Nature Genetics*, 2015.

Cancer-associated stromal cells or cancer-associated fibroblasts (CAFs) represent a diverse group of cells recognized as significant producers of ECM proteins, including collagens and hyaluronans. These typically spindle-shaped cells express one or more activated fibroblast markers, like ACTA2 encoding alpha-smooth muscle actin (α-SMA) (described below), and they have been associated with various tumor-promoting functions including tumorigenesis, angiogenesis, immunosuppression and metastasis[67]. Specifically, fibroblasts in close proximity to cancer cells undergo induction by TGF-β from cancer cells into myofibroblastic CAFs[68]. These myofibroblastic CAFs contribute to the creation of a mechanical barrier with both tumor-promoting and antitumor effects[69,70].

Pancreatic stellate cells (PSCs) get activated by cancer cells when exposed to cytokines, such as platelet-derived growth factor (PDGF) and TGF-β that induce the transformation of PSCs to myofibroblast[71]. Indeed, when become activated, PSCs

exhibit increased proliferation, synthesis of collagen and other extracellular matrix proteins and expression of α-SMA[71]. This is an actin isoform that, upon incorporation in stress fibers, enhances cell contractile activity and resistance to tensile stress[72]. These findings imply a crucial role for stellate cells in the process of pancreatic fibrogenesis, a frequent response to chronic pancreatic injury[71,73,74]. This desmoplastic reaction is recognized for creating a dense mechanical barrier surrounding the tumor, providing specialized niches for cancer cells and influencing tumor growth and invasive behavior[67].

PDAC is commonly considered an immunologically “cold tumor” due to several characteristics that lead to the evasion of an effective cytotoxic immune response[75]. For instance, the oncogenes in PDAC contribute to the creation of an immunosuppressive tumor microenvironment. Mutated KRAS is linked to downstream effects, such as the production of granulocyte-macrophage colony-stimulating factor (GM-CSF), which results in the recruitment of immunosuppressive myeloid cells[76]. Additionally, it downregulates major histocompatibility complex-1 and promotes the increase of regulatory T cells[77]. Immunosuppression stands out as a pervasive and nearly universal hallmark of PDAC, forming a significant barrier to the success of immunotherapy in the treatment of this tumor[78].

The extracellular matrix surrounding the tumor has long been implicated in the regulation of cancer progression, including processes such as migration and invasion. ECM is an external meshwork composed of structural proteins, proteoglycans, adaptor proteins and tissue remodelling enzymes. A diverse array of ECM proteins and macromolecules, originating from either stromal cells or tumor cells, are deposited in the desmoplastic stroma of PDAC[79]. Studies revealed differential expression of proteolytic matrix metalloproteinases (MMPs) and tissue inhibitors of MMPs between non-transformed pancreas and pancreatic ductal adenocarcinoma tissues. In particular, elevated expression of specific MMPs has been associated with metastatic disease and/or poorer prognosis[80–82].

The exploration of intratumoral heterogeneity represents a crucial frontier in comprehending how the tumor microenvironment propels malignant progression. Indeed, signals from the TME play a crucial role in regulating the state, plasticity and therapeutic response of cancer cells.

Recently, it has been shown that multiple subpopulations of CAFs and immune cell communities play a central role in the formation of subtypes within the tumor microenvironment (subTMEs). Grünwald and colleagues[83] discovered the presence of two fundamental microenvironmental states, originating from fibroblast plasticity,

that give rise to functional intratumoral heterogeneity within the pancreatic TME. They defined a *reactive subTME* – characterized by complex yet functionally coordinated fibroblast communities, being immune-rich and associated with aggressive tumor cell phenotypes – and a *deserted subTME* – that contained fewer activated fibroblasts, exhibited tumor-suppressive features and that is matrix-rich. As a result, the heterogeneity within the abundant pancreatic tumor microenvironment is not random, but it signifies fundamental tissue organizational units. Tumors with such heterogeneous TMEs have simultaneous access to both tumor-promoting and chemo-protective elements. Consequently, the poor survival observed is likely a result of the complementary effects of subtypes within the TME on tumor pathobiology.

1.6 Coexisting morpho-biotypes in PDAC

As described before, the concept of diverse signatures coexisting within the same PDAC is firmly established in the field. Nevertheless, there is currently no rational framework that links this idea with morphological heterogeneity. Pathological observations indicating morphological heterogeneity in PDAC can be traced back at least two decades. However, in no instance has a clear and high-resolution intersection of morphological and gene expression data been reported.

In order to address the disparity between the widely recognized intratumor morphological diversity and the molecular knowledge of PDAC heterogeneity, we have recently unveiled the regulatory mechanisms governing distinct morphological patterns within individual PDAC cases[84]. We employed morphological variation as an indicator of the underlying regulatory and gene expression diversity in PDAC. Specifically, we used laser micro-dissection (LMD) to isolate single or multiple morphologically distinct tumor areas (gland-forming or non-gland-forming[49]) from primary PDACs of treatment-naïve patients and then carry out RNA-seq on the microdissected areas. Following differential gene expression analysis, the differentially expressed genes were subjected to unsupervised clustering, resulting in the partitioning of genes into eight clusters (**Figure 3A**). This enabled us to identify transcriptional differences and similarities among the morpho-biotypes. More specifically, morpho-biotypes A and B exhibited a shared high-level expression of genes in clusters 1-3 (**Figure 3B**). These clusters were enriched, among other functions, with terms related to ribosome biogenesis and protein translation, indicating high protein synthesis activity. Additionally, terms associated with cell-cell adhesion and epithelial polarization were identified. Nevertheless, morpho-biotype

A exhibited a distinction from B due to the specific expression of a set of genes (cluster 4) linked to endodermal differentiation and functions, and to a tubular architecture, therefore defined as *glandular* morpho-biotype. In contrast, biotype B exhibited selective expression of genes (cluster 5) encoding proteins involved in extracellular matrix synthesis, cell migration and epithelial-to-mesenchymal transition (**Figure 3B**). In order to highlight the induction of mesenchymal gene expression within cell groups that, however, retained a significant expression of ductal and endodermal genes, we defined this morpho-biotype as *transitional*. Morpho-biotype D was distinguished by the significant downregulation of a majority of genes associated with morpho-biotypes A and B and the upregulation of proteins associated with neuronal activities and functions (**Figure 3B**). Neuronal gene expression represents a common, yet abortive lineage priming event in cancer cells incapable of differentiating towards endodermal lineages. Indeed, these areas tended to coincide with the highly undifferentiated regions identified by pathologists[85]. Hence, we designated this morpho-biotype as *undifferentiated*. In addition, there were tumor areas exhibiting a gene expression profile that fell between the glandular and transitional biotypes on one hand, and the undifferentiated biotype on the other (**Figure 3A**). This hybrid or intermediate profile (group C) may not signify a distinct entity but rather reflect the outcome of the considerable cellular plasticity characteristic of PDAC[86].

Therefore, PDACs are ensembles of three different biotypes with distinct morphology and gene expression programs that coexist in variable proportions in every tumor.

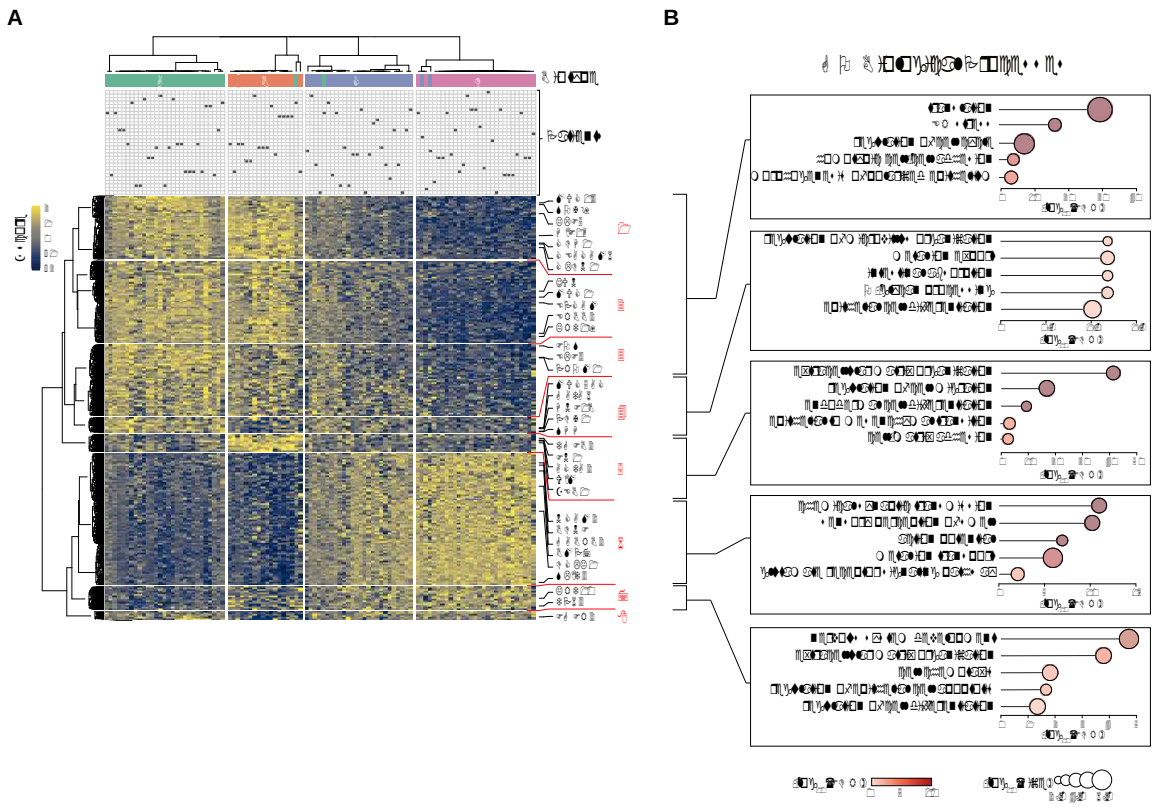

Figure 3: Transcriptional differences and similarities among the PDAC morpho-biotypes. A- Heatmap showing the four sample groups derived from differentially expressed genes clustered by graph-based clustering approach. Samples were re-clustered using ConsensusClusterPlus[87] with Pearson's correlation distance, k-means clustering algorithm and 1000 iterations. Differentially expressed genes identified based on a fold change ≥ ± 2 in either direction and a false discovery rate (FDR) ≤ 0.01 were clustered using KNN and the Louvain graph clustering method. Biotype annotations are shown for each sample. **B-** Representative Gene Ontology classes (GO, biological processes) enriched in the indicated gene cluster. The dot color scale indicates the FDR while the dot size shows the frequency of the GO term in the GO Annotation (GOA) database. Adapted from Di Chiaro P. et al. (preprint in *Biorxiv*).

This approach provides much higher resolution than bulk approaches and significantly differs from single-cell approaches, where relationships between signatures and morphologies are ignored and the distribution of cells with distinct gene expression programs in the tumor tissue is not analyzed.

1.7 Transcriptional regulatory networks controlling morpho-biotype programs

Following the demonstration of the coexistence of morpho-biotype signatures in a clinical dataset, our recent study also involved the identification of transcriptional regulatory networks linking transcription factors and their target genes in the various biotypes (**Figure 4**). To this aim, we used ARACNe-VIPER approach and the gene signatures from each biotype, based on LMD data. Endodermal lineage TFs (e.g., HNF1B, GATA6 and FOXA2) are differentially expressed in the glandular biotype

(green), EMT-associated TFs (e.g., ZEB1 and GLIS2) in the transitional biotype (orange), as well as TFs associated with neural development (e.g., POU5F2) but also stemness of basal epithelia (TP63) in the undifferentiated one (pink).

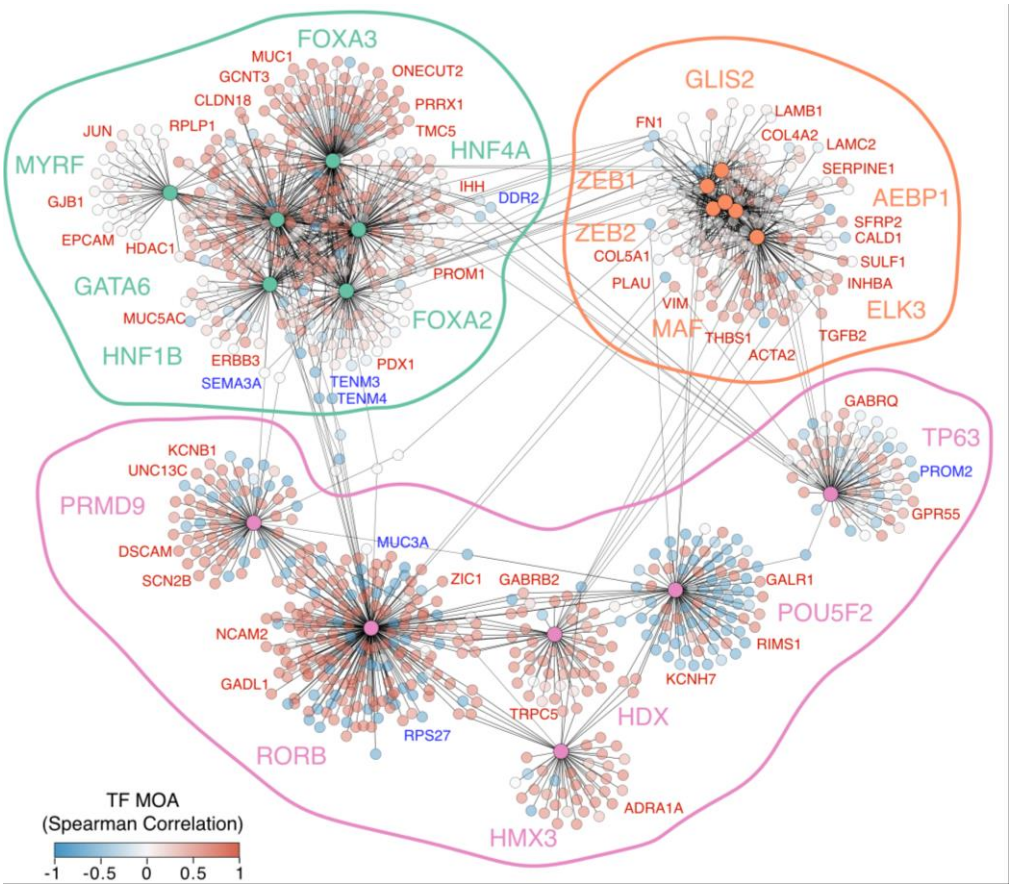

Figure 4: Transcriptional networks controlling PDAC biotype-specific programs. Gene network representation of selected differentially activated transcription factors in the three biotypes as determined by ARACNe-VIPER. Each node represents a target gene that is predicted to be positively (red) or negatively (blue) regulated by its corresponding TF (TF MoA, mode of action of a TF based on Spearman correlation with its target genes).

Currently, we focused our attention on glandular and transitional morpho-biotypes, whose gene expression programs are controlling specific hallmarks of pancreatic cancer biology, like evading growth suppression and activating invasion[88].

A key master regulator in the transcriptional network of glandular biotype is MYRF (**Figure 4**), an endoplasmic reticulum (ER) transmembrane protein that has recently assumed a prominent role in the prevention of ER stress in pancreatic cancer[89].

There are also transcription factors involved in the regulation of pancreas development and cellular differentiation in PDAC. For instance, FOXA2 (Forkhead Box A2) is a canonical pioneer TF with the ability to bind DNA embedded in nucleosomes[90]. It is broadly expressed by the endoderm from the earliest phases of gastrulation and is essential for its development[91,92]. In the context of human PDAC, FOXA2 is broadly expressed in tumor cells regardless of their differentiation

grade, reflecting its role at early stages of endodermal development[93]. FOXA2, along with its close paralog FOXA1, is necessary for the induction of the PDX1 gene, a target in the transcriptional network of glandular biotype (**Figure 4**) that encodes the major driver of pancreatic specification. FOXA1/FOXA2 combined deletion results in nearly complete failure in pancreas development[94].

GATA6 also has a role in the endoderm specification. It is expressed in the foregut endoderm and its haploinsufficiency in humans is the most common cause of pancreas agenesis[95]. Genetic manipulation of GATA6 levels in human cell lines has also demonstrated that GATA6 inhibits epithelial-to-mesenchymal transition, invasion and cell dissemination *in vivo*[96].

A transcription factor involved in pancreas ductal development is HNF1B, initially expressed in pancreatic progenitors and later confined to ductal cells. In humans, HNF1B haploinsufficiency is linked to pancreatic hypoplasia[97]. Specifically, selective inactivation of HNF1B in ductal cells generates alterations in tubular structures, increased inflammation and a fibrotic reaction, collectively creating a conducive environment for KRAS-driven tumor initiation in acinar cells[98].

Regarding the master regulators identified in the transcriptional network of transitional biotype, there is a marked upregulation of genes associated with extracellular matrix synthesis and epithelial-mesenchymal transition (**Figure 4**). Within the set of genes of the ECM that play crucial roles in cell adhesion, migration and signalling there is SFRP2, a signalling molecule currently under exploration as a therapeutic target for fibrotic disorders and cancer[99,100]. LAMC2, a fundamental component of epithelial basement membranes that regulates cell motility and adhesion[101], may contribute to tumor proliferation, migration and invasion through the activation of PI3K/Akt signalling in pancreatic cancer[102].

Mesenchyme□related TFs include ZEB1, which stands out as a potent inhibitor of epithelial identity and its increased expression is associated with the acquisition of mesenchymal features. Specifically, ZEB1 suppresses the expression of epithelial genes, such as CDH1 (encoding E-cadherin), and microRNAs of the miR-200 family. These microRNAs play a crucial role in maintaining epithelial identity by repressing a network of targets that control cytoskeleton dynamics, invasion and migration[103]. ZEB1 and miR-200 family are considered central players in the regulation of EMT, invasion and metastasis in pancreatic and other cancers[104].

Following the identification of transcriptional regulatory networks in the various biotypes, this study delved into a more focused exploration of two pivotal proteins. More specifically:

- in the glandular biotype we investigated MYRF. Since we have already demonstrated its role in maintaining homeostasis in the endoplasmic reticulum of pancreatic secretory glandular cells, now we focused on the molecular understanding of its regulation;
- in the transitional biotype we investigated SFRP2. Since it is already used as therapeutic target, we focused on its role in remodelling the extracellular matrix to unveil any clinical implication also in PDAC.

2. MYRF as master regulator of the glandular biotype

As described before, we established the transcriptional regulatory networks that sustain the distinct morpho-biotypes observed in pancreatic ductal adenocarcinoma. A substantial number of master regulators were identified within each biotype. Among them, MYRF exhibits elevated expression levels in the context of the glandular morpho-biotype, with a prominent role in pancreatic cancer[89].

MYRF is an endoplasmic reticulum transmembrane protein highly conserved from nematodes to humans, predicted to be essential for viability[105,106]. It derives its name from its first characterization as a transcription factor involved in the maturation of oligodendrocytes in the central nervous system. Knocking out MYRF in these cells in mice results in the absence of myelin production, subsequent apoptosis of differentiating oligodendrocytes and ultimately death of the mice [107].

2.1 MYRF protein structure

From a structure point of view, the MYRF human protein is mainly characterized by three domains (**Figure 5A**):

- the cytoplasmic N-terminus contains the *DNA-binding domain* (DBD) (Q349-R547). It's homolog to the DNA binding domain of the yeast protein Ndt80, a transcription factor responsible for meiosis completion in yeast[108]. It has been observed that only Ndt80 orthologs in higher eukaryotes, like MYRF, carry the Ndt80-like DBD fused with the intramolecular chaperone auto-processing domain (described below), probably due to a lateral gene transfer from bacteriophages[109]. MYRF most N-terminal fragment is also characterized by a transactivation domain to induce transcription and its sumoylation adds a further layer of control of MYRF transcriptional activity[110];

- the DBD is followed by an *intramolecular chaperone auto-processing domain* (ICA domain) (G558-R701)[109]. It's the ortholog of the same domain found in the tailspike proteins of bacteriophages. During bacteria infection, these proteins are able to trimerize and bind to the surface of the host cell. Then, the self-cleavage of the ICA domain activates the endosialidase that consequently degrades the bacteria wall[111]. The ICA domain of the MYRF protein maintains its function and induces trimerization and self-cleavage of MYRF (**Figure 5B**). Therefore, the N-terminal trimeric fragment is released to translocate into the nucleus and act as a *bona fide* TF to induce the transcription of the target genes[109,112] (**Figure 5B**). The remaining part of the protein, including most of the ICA domain and the transmembrane domain, remains inserted in the ER. Whether such autocatalytic release of MYRF from the ER membrane is regulated is currently unknown;
- the C-terminus includes the *transmembrane domain* (TMD) (G767-Y787) which allows MYRF to be inserted into the endoplasmic reticulum membrane.

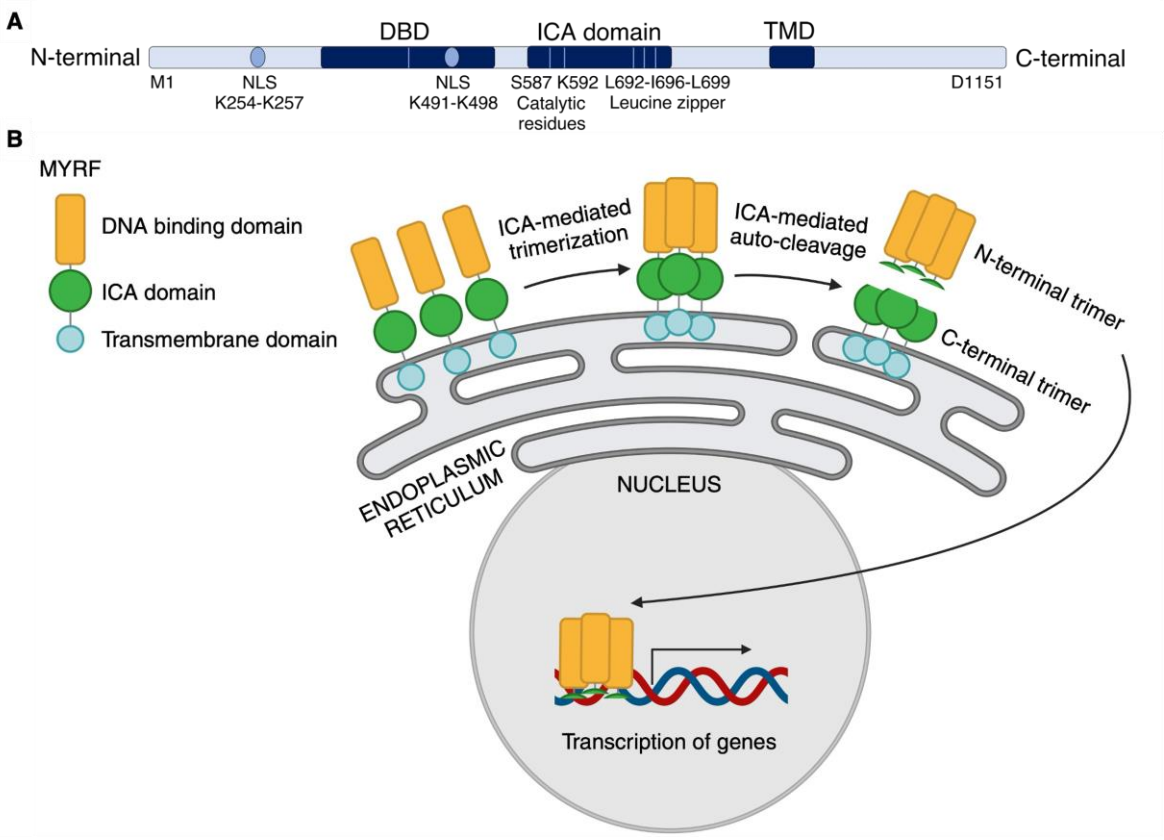

Figure 5: Structure of MYRF protein. A- Schematic representation of MYRF protein structure with principal functional domains. NLS – Nuclear Localization Signal, DBD – DNA binding domain, ICA – Intramolecular Chaperone Auto-processing, TMD – transmembrane domain. **B-** Diagram of MYRF trimerization and cleavage in the ER and nuclear translocation and transcriptional activation in the nucleus. Image created with Biorender. ICA – Intramolecular Chaperone Auto-processing. Adapted from Li Z. et al., *PloS Biology*, 2013.

2.2 MYRF functions

Several studies have highlighted the involvement of MYRF in physiological processes associated with extensive protein secretion. As mentioned before, it was first identified as a transcription factor controlling myelin expression in the transition from pre-oligodendrocytes to mature oligodendrocytes in the central nervous system[107]. The N-terminal product resulting from MYRF cleavage directly interacts with a diverse set of genes associated with myelination, promoting their transcription. Numerous binding sites for MYRF have been identified, revealing previously uncharacterized enhancers for these myelin-related genes[109].

Additionally, MYRF ortholog in *C.elegans* is involved in molting, the process through which the external collagenous cuticle is lost and substituted by a new slightly bigger cuticle in the maturation toward the adulthood[113].

Interestingly, the correlation between MYRF and processes depending on secretion was also observed in highly secretory PDAC cells. In a previous study from our laboratory, it has been shown that MYRF limits the expression of genes encoding secretory proteins with complex folding, thus preventing ER overload. Indeed, RNA-seq analysis of the differentially expressed transcripts in MYRF KO pancreatic cells revealed that numerous genes promoting cell proliferation and DNA synthesis – particularly, the direct targets of MYRF such as FOS and FOSB – experienced significant downregulation (**Figure 6A**). On the contrary, genes linked to protein maturation, ER stress and unfolded protein response (UPR) are upregulated (**Figure 6A**). The upregulated genes related to ER stress included CHOP/DDIT3 and CHAC1, components of the PERK-ATF4 arm of the UPR, and HERPUD1/HERP, a crucial component of the ER-associated degradation (ERAD) machinery. Moreover, the deletion of MYRF *in vivo* gives rise to hypocellular tumors with enlarged extracellular spaces filled with secretory products. To investigate the function of the ER in MYRF KO cells, we employed transmission electron microscopy to analyze the morphological features of KO cells (**Figure 6B**): it has been demonstrated that ER alterations occur, with a complete loss of the typical perinuclear lamellar stacks of ER cisternae which are replaced by enlarged sacs consistent with ER overload. These ER alterations are associated with an increase in the number of lysosomes and autophagic bodies, a common response to ER stress.

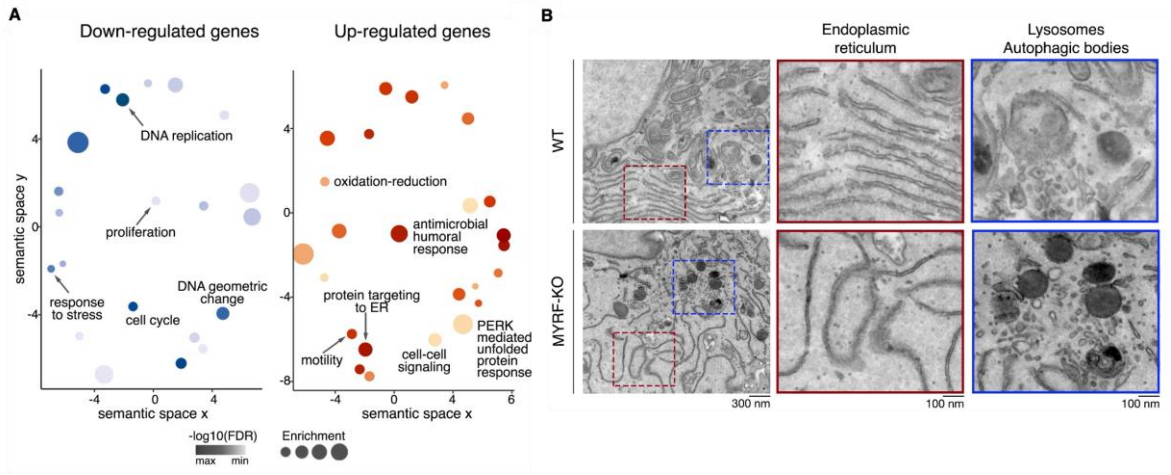

Figure 6: MYRF deletion effects. A- Functional categories associated with genes differentially expressed in MYRF KO cells. **B-** Transmission electron microscopy images of WT and MYRF KO PDAC cells. From left to the right, a lower magnification, magnification focusing on ER morphology, magnification showing lysosomes and autophagic bodies. Milan M. et al., *Developmental Cell*, 2020.

These data indicate that MYRF maintains ER homeostasis and proliferative capacity in highly secretory PDAC cells[114].

2.3 ER stress in pancreatic cancer

In the context of cancer, the heightened cell proliferation and subsequent increased rates of protein synthesis can instigate ER stress. Additionally, localized depletion of oxygen, nutrients and glucose in the tumor microenvironment can contribute to ER stress[115]. It is then clear that tumor cells survive, grow and metastasize through a critical adaptive response to ER stress[116].

PDAC tumors and the cells constituting their surrounding microenvironments are documented to display basal endoplasmic reticulum stress, characterized by the accumulation of unfolded proteins in the ER lumen[117]. Indeed, ER stress condition results from an imbalance between the demand for protein folding and the folding capacity of the ER[118]. Consequently, secretory cancer cells must adapt ER function to the increased demand imposed by proteins with a complex processing[119]. Such complex folding is mainly due to the extensive glycosylation of the protein or to the presence of multiple disulfide bridges, the formation of which relies on a controlled oxidative environment in the ER[120,121].

The cellular response to ER stress involves the activation of ERAD and UPR[122] (**Figure 7**). These responses have the following objectives to restore ER homeostasis: reduce overall protein synthesis to alleviate the protein burden in the

ER, enhance the ER capacity to manage misfolded proteins and, if normal function restoration fails, initiate cell death through the induction of apoptosis[123].

The ER lumen is characterized by the presence of a high concentration of molecular chaperones, including calnexin, calreticulin and BiP. These chaperones play a primary role in folding and preventing the aggregation of newly synthesized proteins[124]. When misfolded proteins accumulate in the ER, chaperone proteins are recruited to the misfolded proteins releasing the three UPR sensors – ER transmembrane proteins, namely IRE1a, PERK and ATF6 – that stimulate their signaling cascade[123], and ERAD is activated. The primary function of ERAD is to alleviate ER stress by retrotranslocating misfolded proteins from the ER lumen to the cytosol[122]. This retrotranslocation process is often coupled with ubiquitination by an ER-associated E3 ubiquitin ligase[122].

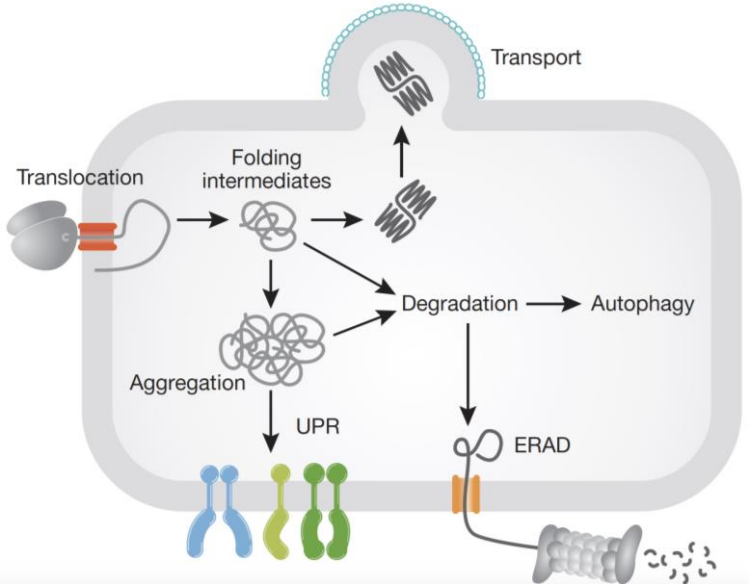

Figure 7: Cellular responses to ER stress. Upon ER stress, UPR and ERAD are activated to release ER stress or, if necessary, induce apoptosis. UPR – Unfolded Protein Response, ERAD – ER-Associated Degradation. Douglas M. et al., *EMBO Reports*, 2009.

Increasing evidence suggests that the UPR is consistently active in pancreatic ductal adenocarcinoma and may play a role in tumorigenesis and disease progression[125]. This underscores the importance of understanding the intricate dynamics of the UPR in the context of PDAC, as it may offer crucial insights into potential therapeutic targets and strategies for managing this challenging cancer type[125].

3. SFRP2 in ECM remodelling of transitional cells

Another important aspect of this study lies in uncovering novel regulators influencing PDAC properties. One of the key aspects of transitional cancer cells is the ability to contribute to the extracellular matrix remodeling and SFRP2 stands out as a potential regulator of such process. Indeed, molecules targeting SFRP2 are currently under exploration for fibrotic disorders and cancer[99,100].

SFRP2 is a soluble glycoprotein belonging to the secreted frizzled-related protein family. These proteins share structural similarities with frizzled proteins, the serpentine receptors playing a pivotal role in the widely utilized cell-cell communication pathway involving Wnt signalling. Due to this homology, SFRPs were promptly identified as antagonists to the Wnt pathway: they hindered the activation of the signalling pathway by binding to Wnt proteins. Nevertheless, some Wnt-unrelated effects have also been reported[126,127].

SFRP2 is overexpressed by CAFs in tumors characterized by stromal signatures and elevated mesenchymal gene expression levels[128,129]. Within this context, SFRP2 plays a role in the CAF-driven enhancement of invasive properties in cancer cells, ultimately contributing to an unfavourable prognosis, as shown in the case of melanoma[128].

3.1 Secreted frizzled-related protein family

The secreted frizzled-related protein (SFRP) family comprises five secreted glycoproteins, each approximately 300 amino acids in length: SFRP1, SFRP2, SFRP3, SFRP4 and SFRP5 (**Figure 8**). According to sequence comparison and phylogenetic analysis, SFRP1, SFRP2 and SFRP5 are closely related forming a subgroup that diverges from the subgroup constituted by SFRP3 and SFRP4[130]. Post-translational modifications seem to introduce additional distinctions that could contribute to the functional diversity among various SFRP family members. For instance, SFRP1 undergoes N-glycosylation, resulting in a shift of approximately 2.8 kDa[131]. It is also sulphated at two tyrosine residues, a modification highly conserved in SFRP5 but not present in SFRP2, SFRP3 and SFRP4.

All members of SFRP family are modular proteins sharing a distinctive structure that comprises two independent domains[131] (**Figure 8**):

- the N-terminus includes a secretion signal peptide, absent in the mature secreted protein, followed by a *cysteine-rich domain* (CRD). The CRD is characterized by the presence of ten cysteine residues at conserved positions, forming a pattern of disulfide bridges identical to that observed in the extracellular Wnt binding domain of frizzled (FzD) receptors[131,132]. This

domain facilitates the interaction between SFRP proteins and Wnt ligands, ultimately leading to the antagonization of Wnt signalling[129];

- the C-terminus is characterized by segments containing positively charged residues, suggesting heparin-binding properties[133]. This region is also marked by six cysteine residues forming three disulfide bridges, defining the *netrin-like domain* (NTR). The NTR domain exhibits homology to the netrin domain found in the complement proteins C3, C4, C5, type I procollagen C-proteinase enhancer proteins and tissue inhibitors of metalloproteinases[134].

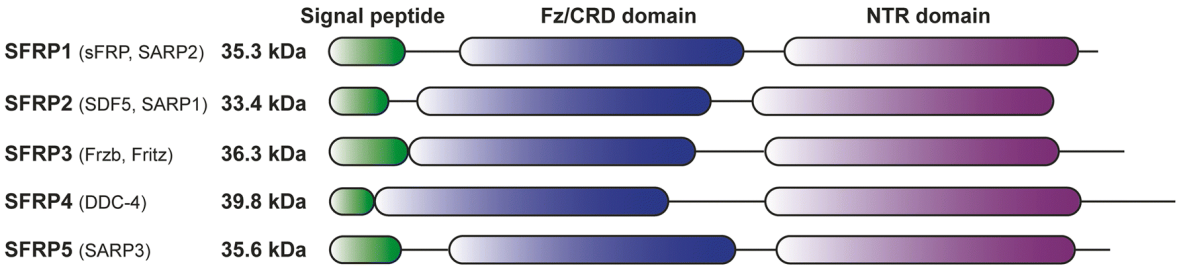

Figure 8: SFRP proteins structure. The members of the SFRP family share a signal peptide (green), a frizzled-like cysteine-rich domain (Fz/CRD, blue) and a netrin domain (NTR, purple). Size and synonyms for each protein are indicated. van Loon K. et al., *Cancer Metastasis Reviews*, 2021.

3.3 SFRP2 in cancer

As mentioned before, the SFRP family is recognized for its involvement in the regulation of Wnt signalling pathway. Activation of the Wnt signalling pathway occurs through the binding of soluble Wnt ligands to FzD receptors on the cell surface, ultimately resulting in the transcription of Wnt target genes.

Overactivation of the Wnt signalling pathway plays a crucial role not only in embryonic development, tissue regeneration and cell proliferation, but also in the initiation, progression and formation of metastases in cancer[135,136]. In colorectal cancer a significant proportion of cases exhibits overactivation of the canonical Wnt signalling pathway, often due to mutations in genes such as β-catenin[137]. Mutations in β-catenin are also frequently observed in various other cancer types, including hepatocellular carcinoma, gastric carcinoma, ovarian carcinoma and melanoma [137]. Given its significance as a key player in the Wnt signalling pathway, SFRP2 has the capacity to affect various aspects of tumorigenesis, functioning as both a tumor suppressor and a tumor promoter. Indeed, SFRP2 is downregulated through its promoter hypermethylation in various types of cancer, like bladder, breast, gastric, lung, ovarian and pancreatic cancer[136]. On the contrary, SFRP2 expression was significantly higher in osteosarcoma tumors compared to mesenchymal stem cells and in the serum of breast cancer patients compared to controls, thus promising as

a biomarker and prognostic prediction tool[136]. Collectively, these findings provide evidence that SFRP2 can exhibit dual roles, depending on the specific context and conditions within the tumor microenvironment.

While the overactivation of the Wnt signalling pathway is closely linked to cancer initiation and progression, it also plays a significant role in the development of the tumor vasculature. In this context, SFRP2 is interestingly demonstrated to be upregulated, being a crucial player in the process of tumor angiogenesis[136].

3.4 SFRP2 in tumor microenvironment

While many of the discussed functions of SFRPs are related to their effects on Wnt signalling, it's worth noting that SFRPs exhibit promiscuity by interacting with molecules unrelated to components of the Wnt signalling cascades. Their ability to engage with diverse molecules suggests a broader functional role beyond Wnt modulation, indicating that SFRPs may participate in various cellular processes and pathways with no apparent connection to Wnt signaling[126].

The relationship between SFRP2 and the tumor stroma involves the interaction of SFRP2 with the extracellular matrix, playing a role in shaping the tumor microenvironment. Specifically, SFRP2 is known to accumulate in the extracellular matrix and to physically interact with components, such as fibronectin-integrin complexes. This interaction can influence cell adhesion and viability, as it occurs in mammary tumor cells by providing protection against apoptotic stimuli[138]. Through its interactions in the extracellular matrix, SFRP2 could potentially modulate signalling pathways involved in tumor-stroma crosstalk, impacting the behaviour of both cancer cells and stromal cells. Indeed, the expression of SFRP2 is frequently elevated in primary tumors compared to their normal counterparts, potentially due to heightened production by the tumor stroma. In this case SFRP2, commonly linked to unfavourable patient outcomes, is a component of a shared gene program associated with epithelial-to-mesenchymal transition observed across a diverse array of cancers[139].

As mentioned before, stromal cells within the tumor microenvironment, such as cancer-associated fibroblasts and immune cells, play a crucial role in tumor progression. Kaur et al. discovered that SFRP2 expressed by aged fibroblasts promoted melanoma angiogenesis and metastasis by suppressing β-catenin in melanoma cells[140].

In summary, SFRP2 interaction with the extracellular matrix and potential influence on the tumor stroma suggest a multifaceted role in the TME. Further research is

needed to elucidate the specific mechanisms and consequences of SFRP2 involvement in the tumor-stroma interplay in different cancer types.

3.5 The extracellular matrix

The extracellular matrix is a complex assembly of multidomain macromolecules, including glycoproteins, proteoglycans and fibrous proteins, like collagen and elastin. Collagen fibers, a major ECM component, provide tensile strength, while elastin imparts elasticity[141]. The composition and the arrangement of ECM components are tissue-specific, providing a unique microenvironment tailored to precise physiological requirements and contributing to the mechanical properties of tissues[142,143]. The interaction among ECM components contributes to the formation of a structurally stable composite, influencing tissue mechanics[144]. This mechanical stability is vital for tissue integrity and function. Additionally, the ECM acts as a reservoir for various growth factors and bioactive molecules, influencing cellular processes such as proliferation and differentiation[145]. Growth factors are stored in the ECM and released in a regulated manner, playing a pivotal role in tissue development, repair, and homeostasis.

The ECM exists in two primary forms, varying in function, composition and location. One form is the *interstitial matrix*, which creates porous three-dimensional networks enveloping cells, linking them within the stroma and establishing connections to the *basement membrane* (BM), the other ECM structure.

The interstitial matrix not only ensures the structural integrity of tissues and organs, but also influences processes like cell differentiation and migration. Its protein composition predominantly consists of collagens I, III, V, fibronectin and elastin[146]. Cancer involves the restructuring of the interstitial ECM, leading to a diverse array of biophysical and biochemical alterations that impact cell signalling, ECM rigidity, cell migration and the progression of tumors[147].

The basement membrane is a slim layer of extracellular matrix situated beneath epithelial and endothelial tissues. It is constructed from an intricate interwoven network of polymerized laminins – which offers signalling cues to cells – and type IV collagen – which is considered the primary structural backbone of the BM – with additional crosslinking provided by proteins like nidogen and perlecan (**Figure 9**). The remodelling of the basement membrane is essential for cancer cells to invade the stromal tissue and develop into a malignant tumor[148].

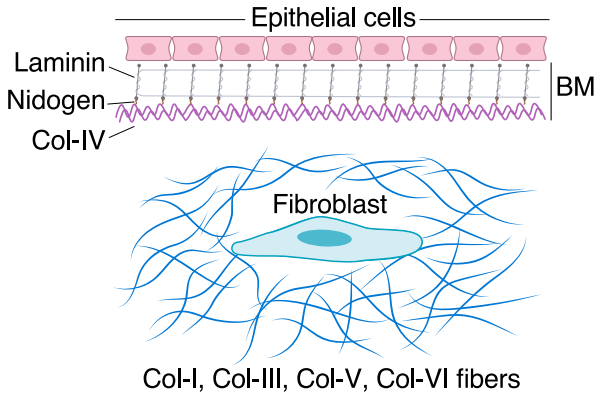

Figure 9: The basement membrane. Schematic representation of the basement membrane (BM) and its main components (top), and the fibrous extracellular matrix (ECM) containing typical collagen fibers produced by fibroblasts (bottom). Di Chiaro P. et al. (preprint in *Biorxiv*).

The intricate processes of ECM remodelling result in alterations to the overall abundance, concentration, structure and organization of individual ECM components. This, in turn, influences the three-dimensional spatial topology of the matrix surrounding cells, as well as its biochemical and biophysical properties, ultimately impacting cell fate. In pathological conditions, such as inflammatory diseases, tissue fibrosis and cancer, cells exhibit dysregulation of this process.

Recent research emphasizes the significance of tumor-mediated systemic aberrations in the ECM, particularly in the context of establishing metastasis[149]. The mechanics of the basement membrane are crucial to consider due to cell mechanotransduction, which is the process by which cells sense and respond to mechanical cues. Given the thinness of the BM underlying epithelial tissues, a critical question is whether cells can sense the stiffness of only the immediate BM or if they can also perceive the stiffness of the stromal matrix beyond the BM. Research indicates that cells can detect stiffness on the micron scale, suggesting that the effective stiffness sensed by cells in direct contact with the BM could be influenced by the mechanical properties of the surrounding stromal matrix[150,151]. A common mechanoresponse to heightened stiffness and mechanical stresses is the enhanced deposition of fibrous extracellular matrix, a modification directly associated with tumor promotion[152]. This response serves as a protective mechanism for parenchymal cells experiencing prolonged mechanical stress. However, it can also initiate a detrimental feedforward loop, amplifying tissue mechanics and leading to an extended fibrotic response[153] like desmoplasia, a significant characteristic observed in various cancers, including PDAC[148,154].

In our LMD-seq data, the mRNA levels of laminins and type IV collagens were well expressed in the glandular biotype, significantly increased in the transitional biotype and nearly undetectable in the undifferentiated biotype (**Figure 10**).

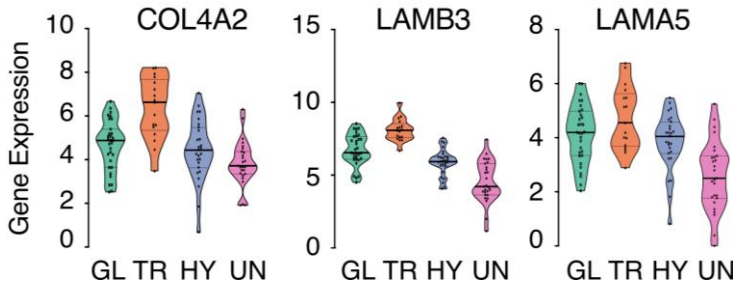

Figure 10: Expression of basement membrane genes in PDAC morpho-biotypes. Violin plots showing the expression levels stratified by biotype of representative differentially expressed genes. Gene expression represents the $\log_2$ (normalized read counts scaled by cDNA length + 1). GL – Glandular, TR – Transitional, HY – Hybrid, UN – Undifferentiated. Di Chiaro P. et al. (preprint in *Biorxiv*).

These data suggest that the transitional and undifferentiated biotypes exhibit a severely impaired organization of the basement membrane, ranging from its complete absence to significant alterations in its structure. This may be linked to the heightened propensity of these malignant cells to invade the surrounding tissue.

Aim of the project

The aim of this project is to characterize two key proteins – MYRF and SFRP2 – involved in the regulation of two specific hallmarks of pancreatic cancer, the ER stress and the ECM remodelling respectively.

Firstly, the focus is on identifying and studying the interactors of MYRF in the luminal side of the ER to unveil cleavage regulatory factors. The study of this protein also aims to dissect the role of its two critical domains – the DNA binding domain and the transmembrane domain – in MYRF-mediated maintenance of ER functionality. In this way, we will obtain a mechanistic understanding of MYRF in the regulation of ER activity in pancreatic cancer cells.

Secondly, our objectives were to observe the impact of SFRP2 on extracellular matrix remodelling and assess its influence on the migratory and invasive capabilities of PDAC cells. Understanding the association between SFRP2 and tumor-promoting functions could offer substantial potential for the development of anti-cancer therapies targeting this protein.

Materials and methods

Cell culture

Cell lines and media

CFPAC-1 cells were obtained from the American Type Culture Collection (ATCC, CRL-1918). Stable CFPAC1 expressing MYRF TurboID fusion proteins were generated using lentiviral transduction: HEK-293 were transfected with each construct using calcium phosphate. The transfected cells were incubated at 37°C for 8h, after which they were replenished with fresh medium and further incubated at 32°C for 48h. The culture media was filtered through a 0.45um filter and added to CFPAC-1 cells along with Polybrene (8mg/mL). At 24h after transduction, puromycin (1.5ug/mL) was added to the target cells.

For the MYRF TurboID experimental set up, WT CFPAC-1 and CFPAC-1 expressing MYRF TurboID fusion proteins were maintained in 5.0% CO2 at 37°C in IMDM (Sigma) or DMEM (EuroClone) supplemented with 10% fetal bovine serum (HyClone) or dialyzed serum (dFBS) (HyClone) for 5 days, according to the experimental conditions described in the results. Media were all supplemented with 2 mM L-glutamine and 1% pen/strep. For the streptavidin pull-down experiment, the same cells were maintained in DMEM with 10% dialyzed fetal bovine serum for 5 days. At day 5, 50 mM biotin (Sigma cat. B4639) was added to the cells for 10 minutes.

HCT116 were obtained from the American Type Culture Collection (ATCC, CCL-247). They were maintained in 5.0% CO2 at 37°C in McCoy's 5A medium supplemented with 10% fetal bovine serum (FBS) (HyClone) and 2 mM L-glutamine.

The human primary MDA-PATC69 cells (kindly gifted by Dr. Michael Kim MD Anderson Cancer Center) were generated from a xenotransplanted human tumor fragment[155]. Cells were maintained in DMEM (EuroClone) supplemented with 10% fetal bovine serum (HyClone), 2 mM L-glutamine and 1% pen/strep.

For recombinant protein experiments, MDA-PATC69 cells were treated with 200ng/mL of recombinant human sFRP-2 protein (R&D Systems, #6838-FR-025/CF) 24 hours after seeding. For gene overexpression experiments, 1x10e6 MDA-PATC69 cells were transfected with 2µg of the SFRP2-Flag or the empty vector (pcDNA3.1, ThermoFisher Scientific) using the P1 Primary Cell 4D-Nucleofector X kit (Lonza Bioscience #V4XP-1012) and program EA-104 of the 4D-Nucleofector system (Lonza Bioscience) following manufacturer's instructions. After nucleofection, cells were grown for 48 h in culture medium before proceeding with further assays.

For mechano-response experiments, 5x10e4 MDA-PATC69 cells were plated onto collagen-coated hydrogels with defined stiffness (1 and 12 kPa) bound to 12-well plates (Matrigen Softwell, CELL guidance systems) and harvested after 48 h.

Random migration and Boyden chamber assays

For random migration assays, 2x10e4 untreated or nucleofected MDA-PATC69 cells were left adhering for 24h in a 12-well plate. Before starting time-lapse imaging, media was replaced and supplemented with Nuclight Rapid Red (Incucyte, 1:2000) to live-stain nuclei and cells stimulated with recombinant sFRP2 protein. Bright field and 647 nm fluorescence images were acquired every 15 min for 72h with Leica DFC9000 GTC camera using the objective HC PL Fluotar 10X/0,30, mounted on a Leica Thunder imager system microscope equipped with an incubator chamber (OKOlab) maintained at 37C in a 5% CO2 atmosphere. Tracking of migrating cells was performed using TrackMate plug-in (Fiji software) with LoG detector and Lack Tracker algorithms.

For Boyden chamber assays, 1x10e5 serum-starved nucleofected cells or cells treated with recombinant sFRP2 for 48h were seeded on top of 8mm-pore polycarbonate membranes (Costar) precoated with 20 µg Collagen (for migration assay, non-gelling type I collagen, C8919, Sigma-Aldrich) or 35 µg Matrigel (for invasion assay, growth factor reduced 356231, BD biosciences). Cells were allowed to migrate for 24 hours towards a serum gradient and fixed in 4%PFA. Nuclei were stained with DAPI and counted using the local maxima algorithm by Fiji software.

In-Cell ELISA

2x10e3 untreated or nucleofected MDA-PATC69 cells were left adhering for 24h in a 96 well plate. Culture media was replaced and untreated cells stimulated with recombinant sFRP2 protein. Cells were cultured for 96h to allow ECM deposition and lysed with 0.075% ammonium hydroxide to decellularize the plate[156]. The wells were rinsed in PBS, blocked with 5% BSA and incubated overnight at 4˚C with the following antibodies: collagen IV (Millipore AB769, 0.4 µg/ml), COL3A1 (Sigma-Aldrich #HPA007583, 0.5 µg/ml), COL5A1 (Sigma-Aldrich #HPA030769, 0.5 µg/ml), LAMB3 (Sigma-Aldrich #AMAb91160, 1 µg/ml), fibronectin (Sigma-Aldrich F3648, 1:10000). After extensive washing, wells were incubated with the respective HRP-conjugated secondary antibodies and signal revealed with tetramethylbenzidine (TMB) substrate. The colorimetric reaction was allowed to develop for 1 hour at room temperature. Reaction was stopped with 0.6M sulfuric acid and 450 nm absorbance

measured at Glomax Explorer instrument (Promega). Signals were normalized to the DNA content assessed by QuantiFluor assay (Promega) from the cell lysate and empty wells coated with 5% BSA were used as background controls for each antibody.

Cloning and plasmids construction

For the generation of MYRF TurboID constructs, TurboID-V5 sequence was amplified from C1(1-29)-TurboID-V5_pCDNA3 (Addgene cat. #107173)[157] to be cloned into a pSCALP vector – carrying FLAG tag at the N-terminal of the cloned protein and cut with XhoI and XbaI – with the following primers:

FWTURBO – 5'GCCCCCTCGAGAAAGACAATACTGTGCCTCT3'

RVTURBO – 5'CCGTCTAGATTAGGTGCTGTCCAGGCCCA3'

The resulting construct was pSCALP_TurboID-V5 vector.

Then, full length MYRF sequence and MYRF sequence lacking the fragment facing the ER lumen (MYRF ΔER) were amplified from our pSCALP_MYRF vector[89] to be cloned into pSCALP_TurboID-V5 vector with the following primers:

FMYRFTI – 5'GCCGGATCCATGGAGGTGGTGGACGAGAC3'

RMYRFTI – 5'CCGGCGGCCGCATGTCACACAGGCGGTAGAAGT3'

RMYERTI – 5'CCGGCGGCCGCATGTACAGTGTGGACATGGACA3'

The resulting constructs (MYRF TurboID and MYRF ΔER TurboID) carried TurboID-V5 at the C-terminal of the cloned proteins.

PCR fragments were amplified using KAPA Taq DNA Polymerase. The vectors were double-digested using standard enzymatic restriction digest and ligated to gel purified PCR products by T4 DNA ligation. Ligated plasmid products were transformed in bacteria and colonies were screened. The constructs were verified by Sanger sequencing.

For the generation of SFRP2-FLAG construct, human SFRP2 sequence was cloned into the EcoRI sites of the pcDNA3.1 vector (Invitrogen), adding FLAG tag at the C-terminal of the protein using NEB Gibson Assembly Cloning Kit and the following primers:

SFRP2_1– 5'CCGCCAGTGTGCTGGAATTCGCCACCATGCTGCAGGGCCCTG3'

SFRP2_3 – 5'GTGTGATGGATATCTGCAGAATTCTCACTTGTCATCGTCGTCCTTGTAGTCG CACTGCAGCTTGCGGA3'

The construct was verified by Sanger sequencing.

Protein manipulation techniques

Mass spectrometry

WT CFPAC-1 cells and CFPAC-1 expressing MYRF TurboID and MYRF ΔER TurboID were harvested, washed twice with ice-cold PBS and centrifuged at 1500 rpm for 5 minutes. Cell pellets were lysed with RIPA buffer containing 50mM Tris-HCl pH 8, 150mM NaCl, 1% NP40, 0.1% sodium deoxycholate, 0.1% SDS, a cocktail of protease inhibitors (c-complete EDTA-free 3x20 tablets, Roche) and 1mM PMSF and incubated by rotation for 30 minutes at 4°C. After sonication and clarification through centrifugation (15000 rpm at 4°C for 15 minutes), the supernatant was collected, quantified and 1mg was used for streptavidin pull-down. 50 ml of Dynabeads MyOne Streptavidin C1 beads (Invitrogen cat. 65002) were then added to each sample and incubated overnight with rotation at 4°C. Beads were then washed using RIPA buffer (7x5 minutes washes) and then PBS 1X (2x5 minutes washes). The beads from the streptavidin pull-downs were then prepared for MS using the iST Sample Preparation Kit (Preomics cat. 00027) following manufacturer's specifications. Sample preparation was performed according to the in-StageTip protocol[158]. Briefly, samples were incubated in PreOmics lysis buffer (catalogue number P.O. 00001, PreOmics) for cysteine reduction and alkylation, followed by a protein denaturation step at 95°C for 10 minutes. Then, proteins from each sample were processed by adding the Trypsin/LysC mix (Digestion buffer, PreOmics) at 1:50 enzyme-to-protein ratio, at 37°C overnight. Tryptic peptides were eluted from the solid-phase extraction material and dried out completely with a SpeedVac centrifuge (Eppendorf). Finally, tryptic peptides were re-suspended in 5 µL of Load buffer (PreOmics) and analysed by LC-MS/MS using an EASY-nLC 1200 (Thermo Fisher Scientific, cat. no LC140) connected to a Q-Exactive HF instrument (Thermo Fisher Scientific) through a nano-electrospray ion source. In all cases, the nano-LC system was operated in one column set-up with an EasySpray PEPMAP RSLC C18 column (Thermo Fisher Scientific) kept at 45°C (constant temperature). Solvent A was 0.1% formic acid (FA) and solvent B was 0.1% FA in 80% Acetonitrile (CAN). Peptides were separated with a gradient of 3–30% solvent B over 64 minutes, followed by a gradient of 30–60% for 10 minutes and 60–95% over 5 minutes, at a flow rate of 250 nL/min. The Q-Exactive was operated in the data-dependent acquisition (DDA) to automatically switch between MS and MSMS mode. Spray voltage was set to 1.7 kV, s-lens RF level at 50, and heated capillary temperature at 275°C. Selected target ions were dynamically excluded for 20

seconds and all experiments were acquired using positive polarity mode. MS spectra (from m/z 375-1550) were analysed in the Orbitrap detector with resolution R=60,000 at m/z 200. The 15 most intense peptide ions with charge states ≥2 were sequentially isolated to a target value of 3e6 (Top15). MS2 data was acquired at R=15,000 resolution and an ion target value of 1e5. Higher-energy collisional dissociation (HCD) fragment scans was acquired with optimal setting for parallel acquisition using 1.4 m/z isolation width and normalized collision energy of 28. The maximum allowed ion accumulation times were 20ms for full scans and 80ms for MSMS.

Western blots and immunoprecipitation

CFPAC1 cells were lysed in RIPA buffer containing 50mM Tris-HCl pH 8, 150mM NaCl, 1% NP40, 0.1% sodium deoxycholate, 0.1% SDS and incubated for 30 minutes on ice. HCT116 cells were lysed in another RIPA buffer containing 50mM Tris-HCl pH 8, 150mM NaCl, 1% NP40, 0.5% sodium deoxycholate, 1% SDS and incubated for 30 minutes on ice. Both buffers were prepared with fresh addition of a cocktail of protease inhibitors (c-complete EDTA-free 3x20 tablets, Roche) and 1mM PMSF. After sonication and clarification through centrifugation (15,000 rpm at 4°C for 15 minutes), protein extracts were resolved on SDS–polyacrylamide gel, blotted onto nitrocellulose membranes, and probed with the following antibodies: MYRF (Home-made, 2μg/mL), V5 (Invitrogen R960-25, 1:1000), UGGT1 (*Bethyl* Laboratories A305-530A, 0.1μg/mL), streptavidin-HRP (Abcam ab7403, 1:5000), vinculin (Sigma-Aldrich, V9131, 1:10000), tubulin (Santa Cruz Biotechnologies sc-32293, 1:5000). The anti-MYRF antibody was generated in house by immunizing rabbits with a GST-fused peptide mapping on MYRF aminoacids 29-125.

For the validation of MS results, streptavidin pull-down was performed as described above and beads at the final step were eluted in Laemmli buffer (Bio-Rad cat. 1610747). Protein inputs and pulled-down extracts were resolved on SDS–polyacrylamide gel, blotted onto nitrocellulose membranes and probed with the following antibodies: streptavidin-HRP (Abcam ab7403, 1:5000), V5 (Invitrogen R960-25, 1:1000), UGGT1 (Bethyl Laboratories A305-530A, 0.1μg/mL).

For SFRP2 experiments, cells were lysed in RIPA buffer (50mM Tris-HCl pH 8, 150mM NaCl, 1% NP-40, 0.1 Na-Deoxycholate, 0.1% SDS) containing a cocktail of protease inhibitors (c-complete EDTA-free 3x20 tablets, Roche) and 1mM PMSF. 10 μg of clarified cell extracts was resolved on SDS–polyacrylamide gel and blotted

onto nitrocellulose membranes. After blocking, membranes were probed with antibodies SFRP2 (OriGene, TA807387, 1 µg/ml) or vinculin (Sigma-Aldrich, V9131, 1:10000).

2 ml of 0.22µm-filtered conditioned medium from PATC69 cells nucleofected with the empty vector or with the SFRP2-FLAG vector were incubated with 5 µg of anti-FLAG antibody (Sigma-Aldrich F4799) pre-bound to 50 µl of G protein-coupled paramagnetic beads (Dynabeads) overnight at 4°C (with 1mM PMSF, 1µg/ml aprotinin and 1µg/ml leupeptin). Beads were washed extensively with RIPA buffer and once with 1x TE 50mM NaCl. Immunoprecipitates were eluted in 2x Laemmli buffer, resolved by SDS–PAGE and immunoblotted with the SFRP2 antibody.

After incubation with primary antibodies, all membranes were washed and probed with the respective secondary HRP antibodies. Chemiluminescent HRP signal was revealed using Clarity ECL (Bio-Rad). Images were acquired using a Chemidoc imaging system (Bio-Rad).

Confocal microscopy

Immunofluorescence and confocal analysis were performed on WT CFPAC-1 cells and CFPAC-1 expressing MYRF TurboID and MYRF ΔER TurboID, grown onto glass coverslips. Briefly, methanol fixed cells where permeabilized with 0.5% Triton X-100 for 10 minutes, blocked with a blocking buffer (PBS +3% BSA + 0.2% Triton X-100) for 1 hour and stained with primary antibodies. Specifically:

- for the localization of MYRF TurboID and MYRF ΔER TurboID, V5 (Bio-Rad SV5-PK1 MCA1360, 10ug/ml) was used overnight;
- for the colocalization between UGGT1 and biotinylation signal, UGGT1 (Bethyl Laboratories A305-530A, 5ug/mL) was used overnight.

All cells were then stained with specific secondary antibodies linked to fluorophores (to detect the primary antibodies), streptavidin AF568 (Invitrogen S11226, 1:200) (only for UGGT1 colocalization study) and with calnexin (Invitrogen MA3-027-A488, 1:200). Nuclei were counterstained with DAPI and samples mounted with glycerol.

The confocal images were acquired using the Eclipse Ti2 microscope (Nikon Europe B.V.) paired with the X-Light V3 spinning disk (CrestOptics S.p.A), solid-state lasers from Lumencor Celesta light engine, a multiband dichroic mirror, single-band emission filters, and an sCMOS camera (Kinetix, Teledyne Photometrics). Three-dimensional multi-channel stacks, each 3 microns in depth, were obtained using a 100x/1.49 oil immersion objective lens, resulting in a voxel size of 70x70x200 nm.

Two-photon second harmonic generation

Hematoxylin and Eosin (H&E)-stained slides cut from formalin□fixed paraffin□ embedded (FFPE) sections were used for a two-photon excitation (2PE) technique using a pulsed infrared laser (Chameleon Ultra II; Coherent) at 880 nm. Collagen SHG images were acquired by a confocal microscope (Leica; TCS SP5) on an upright microscope (DM6000 CFS) equipped with blue (argon, 488 nm), yellow (561 nm solid state laser) and red (633 nm solid state laser) excitation laser lines with a 20x water immersion objective lens (HCX Apochromat L 20x NA 1.00 W; Leica) and controlled by Leica LAS AF software (Leica).

Deletion techniques

CRISPR/Cas9-mediated genome editing

To generate MYRF deleted CFPAC-1 clones, single guide RNA sequences targeting either the second or the third exon of MYRF were designed using the CRISPR design tool (http://tools.genome-engineering.org) and cloned into lentiCRISPv2 plasmid (Addgene #52961). The following guide sequence were selected and cloned following manufacturer's instructions:

sgRNA 1 – 5'-GCCACGACATCAACGGTGCCC-3'

sgRNA 2 – 5'-GCCAGGGCACCGTTGATGTCG-3'

Following infection and puromycin selection (1.5 µg/ml), single cells were seeded in 96-well plates by dilution and expanded. Clones were screened using western blot.

For the generation of MYRF domain-specific mutants in HCT116, single guide sequence specific to MYRF DBD (exon 8, exon 10) and TMD (exon 16, exon 18) were designed using the CRISPR design tool (http://tools.genome-engineering.org) and cloned into pX330 (Addgene cat. # 42230)[159,160]. The following combinations of single guide RNAs were selected and cloned following manufacturer's instructions:

DBD deletion – 5'ATCGGTTCCTCCTACCTGCT3'-5'GAGCCCCTCCATTTCTGAGGC3'

TMD deletion – 5'CCCAGGGACACCCCTGATCC3'-5'TACCTGGAGGGGCACTAAAC3'

After transfection of cells using FuGENE HD reagent and puromycin selection (300ng/mL for DBD deletion, 500ng/mL for TMD deletion), single cells were seeded in 96-well plates by limiting dilution and expanded. Clones were screened using western blot.

RNA manipulation techniques

Total RNA extraction

Total RNA was extracted using Zymo Quick-RNA kit (Zymo Research, R1055) and used for either retrotranscription and RT-qPCR analysis or RNA-seq.

RT-qPCR

For RT-qPCR experiments, cDNA was prepared from 500 ng of total RNA with ImProm-II Reverse Transcription System (Promega cat. no. A3800) following manual instructions. RT-qPCR was assembled with Fast SYBR Green Master Mix (Applied Biosystems cat. 4385614) and run on a QuantStudio 6 realtime PCR machine (Applied Biosystems). Analysis was done on the Thermo Fisher Cloud platform. Primers were designed using Primer3:

C1ORF43 – 5'GGATGAAAGCTCTGGATGCC3'-5'GCTTTGCGTACACCCTTGAA3'

MYRF – 5'AACATGCGTAAGAAGGGCAAG3'-5'CCAGCGTGTAGTTCTGGTTC3'

TM4SF4 – 5'CTGTGGTTGGATTCTTGGGAG3'-5'GGGTAGCCCCATGTACTATTGG3'

CHAC1 – 5'GAACCCTGGTTACCTGGGC3'-5'CGCAGCAAGTATTCAAGGTTGT3'

BiP – 5'CATCACGCCGTCCTATGTCG3'-5'CGTCAAAGACCGTGTTCTCG3'

CHOP – 5'GGAAACAGAGTGGTCATTCCC3'-5'CTGCTTGAGCCGTTCATTCTC3'

Next-generation sequencing

RNA-seq library preparation and sequencing

RNA-seq in HCT116 was carried out using the TruSeq Stranded Total RNA Sample Preparation kit. Libraries from 1 µg of total RNA were quantified with the Quantifluor (Promega) reagent and the quality of the size was controlled with the Tapestation instrument (Agilent) using the high-sensitivity assay HD5000 (Agilent). cDNA libraries were sequenced on an Illumina NovaSeq 6000 platform with 51-bp paired-end reads.

RNA-seq in MDA-PATC69 cells was carried out using the SMART-seq2 protocol[161] with minor modifications. Briefly, 10 ng of total RNA were reverse transcribed with template-switching using oligo(dT) primers and an LNA-containing template-switching oligo (TSO). The resulting cDNA was pre-amplified, purified and tagmented with Tn5 transposase produced in-house using a described protocol[162]. cDNA fragments generated after tagmentation were gap-repaired, enriched by PCR and purified to create the final cDNA library. Libraries were quantified by QuantiFluor assay run on a Glomax Explorer instrument (Promega) and sequenced on an Illumina NovaSeq 6000 platform to obtain 50 bp paired-end reads.

Computational methods

Mass-spectrometry data processing and statistics

Acquired raw data were analysed using the integrated MaxQuant version 1.6.2.3[163] using the Andromeda search engine[164] and Human Fasta Database downloaded from UniprotKB (74470 Entries) was used. Carbamidomethylation of Cysteine was set as a fixed modification. Enzyme specificity was set as carboxy-terminal to arginine and lysine as expected, using Trypsin and LysC as proteases. A maximum of two missed cleavages was allowed. Peptide identification was performed in Andromeda with an initial precursor mass deviation of up to 7 ppm and a fragment mass deviation of 20 ppm. False discovery rate (FDR) of all peptide was set to a maximum of 1% for the identifications. All proteins and peptides matching the reversed database were filtered out. The LFQ intensity calculation was enabled, as well as the match between runs (MBRs) feature[165]. The “protein groups” output file from MaxQuant was analysed using Perseus software[164]. Briefly, no imputation was used, and the data were filtered to have 3 valid values in at least one group. To identify significant regulated proteins an FDR=0.05 was imposed in the t-test analysis.

Immunofluorescence colocalization analysis

Before conducting image analysis, the multichannel Z-stacks underwent deconvolution using the Lucy-Richardson algorithm within the NIS software. To evaluate colocalization coefficients within the cytoplasmic region, a single optical section was chosen from the middle plane of the nucleus and a custom Python-based image segmentation pipeline that leverages the CellPose algorithm for cellular segmentation[166] was used. This pipeline allowed the identification of nuclear and cellular regions of interest (ROIs) exploiting the DAPI and the calnexin signal. Furthermore, a custom script in Fiji[167] was employed for obtaining cytoplasmic ROIs, executing background subtraction, and conducting colocalization analysis in the cytoplasm through the JACoP plug-in. The degree of colocalization between MYRF and both UGGT1 and calnexin signals was quantified specifically in V5-positive cells, as determined by visual examination of the captured fields of view. All data analysis was carried out using RStudio software (version 2023.6.0.421, RStudio Team, 2020), a platform specifically designed for R programming (URL: http://www.rstudio.com/).

Two-photon second harmonic generation analysis

Image analysis was performed using a custom-made macro in Fiji[167]. Briefly, collagen fibers in SHG images were first pre-processed with a Median and a Top Hat filter and then identified using the Ridge Detection plugin in Fiji. For the quantification of fibers thickness, images were processed with an Enhance Local Contrast (CLAHE) function and an auto local threshold (method=Bernsen). Then the "Local Thickness (masked, calibrated, silent)" plugin (https://imagej.net/imagej-wiki-static/Local_Thickness) was executed.

RNA-seq data analysis

For RNA-seq data the following analysis pipeline was used. Paired-end reads with a minimum of 30 bp (after quality trimming and adapter clipping) were aligned to the human genome build GRCh38/hg38 using TopHat2[168] with parameter "–b2□very□ sensitive". The RefSeq gene annotations was provided as the reference transcriptome. Only uniquely mapped reads were retained. At the gene level, the expression counts were estimated using summarizeOverlaps function of GenomicRanges R package (https://bioconductor.org/packages/release/bioc/html/GenomicRanges.html) with "IntersectionStrict" mode. Normalisation was against a set of invariant genes across samples, defined as those with differences in read counts within 1σ from the mean difference (μdiff), assuming a *quasi*-normal distribution of gene read counts[169]. Read counts were Log_2-transformed and used as proxy of expression for downstream analyses.

Only for the experiment of MDA-PATC69 cells the Read counts were normalized using DEseq2's median of ratios[170].

Gene set enrichment analysis

Gene Set Enrichment Analysis (GSEA) was performed using the fgsea R package (https://bioconductor.org/packages/release/bioc/html/fgsea.html) with the following parameters eps=0.0, minSize=15 and maxSize = 2000. It was run using genes ranked based on the log_2 fold change[170] between stiff (12kPa) and soft (1kPa) groups in MDA-PATC69 cells. Biotype-specific signatures, mechanoresponse[171] and ECM[79] programs were used as gene sets to compute the enrichments. Statistical significance was estimated by an FDR value ≤ 0.05.

For HCT116 transcriptional gene signatures reported previously (ER stress response, GO:0034976) were used to calculate a score for all samples using

Seurat's AddModuleScore function. Default parameters were used except for the number of control features that was set equal to the number of gene signature features.

Results

1. Identification of MYRF interactors

If so far we have observed a mainly transcriptional regulation role of MYRF played by the N-terminal fragment (M1-A591), we now want to elucidate the function and regulatory role of the MYRF C-fragment (K592-D1120), which remains inserted in the luminal side of the ER after self-cleavage and which includes most of the ICA domain, the TMD and the C-terminus.

Since the regulation of MYRF release from the ER membrane is still unknown, our hypothesis is that this fragment may sense the ER luminal content or properties of the ER membrane and eventually regulate MYRF trimerization and cleavage. Usually, cleavage and subsequent release of other ER membrane-associated TFs, such as SREBP2 and ATF6, are highly regulated processes activated in response to specific signals coming from either the ER membrane or the ER lumen[172,173]. However, in contrast to such TFs, MYRF is capable of self-cleavage, implying that its processing and nuclear entry does not require additional players. Indeed, what we have observed so far is that in physiological conditions of a specific PDAC cell line (*CFPAC-1*) MYRF is partially full length (FL, 150 kDa) (**Figure 11A**). On the contrary, when MYRF is overexpressed, it is almost completely self-cleaved (N-terminal, 80 kDa) (**Figure 11B**). Hence, it is tempting to hypothesize that the level of MYRF expression regulates its cleavage, but this requires further validation.

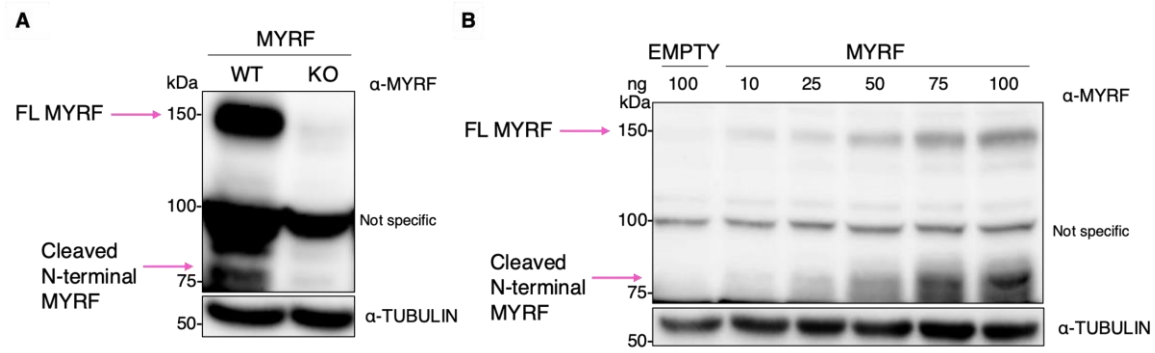

Figure 11: Evaluation of MYRF cleavage. A- Immunoblot of a pool of 3 clones of CFPAC-1 expressing (WT) or knocking-out (KO) MYRF. **B-** Immunoblot of HEK293T transfected with an empty vector or overexpressing MYRF at increasing concentrations. In both figures A and B, the upper left arrow indicates the full length (FL) MYRF protein (150 kDa), while the lower left arrow indicates the cleaved N-terminal MYRF fragment (80 kDa). Tubulin is shown as loading control. Molecular weight markers are indicated.

Alternatively, an additional layer of regulation could be represented by the factors interacting with MYRF C-fragment in the ER lumen. To test this hypothesis, we decided to start by identifying MYRF interactors to reveal potential cleavage regulatory factors.

1.1 TurboID-catalyzed proximity labeling

MYRF is a transmembrane protein and establishing interactions at the membrane level is a challenging task. Membrane proteins require native chaperones and lipid environments to fold properly and they usually undergo significant posttranslational modifications[174]. Consequently, the recombinant expression of transmembrane proteins often results in the production of incorrectly folded forms, limiting the biochemical study of native function[175]. Furthermore, purifying membrane proteins poses greater challenges for several reasons. Firstly, the majority of membrane proteins are found in low abundance. Secondly, being embedded in the lipid bilayer, they necessitate detergents to achieve solubility in aqueous systems. The careful selection of detergents tailored to the solubilization and purification needs of the particular membrane protein is crucial in the purification process[175].

Hence, we decided to set up and optimize an enzyme-catalyzed proximity labeling approach, mainly used for proteomic analysis of macromolecular complexes, organelles and protein interaction networks[176]. Compared with classical immunoprecipitation, this biochemical approach aims at identifying proteins proximal to (i.e., within 10 nm radius from) a protein of interest fused in frame to a biotin ligase – such as TurboID – even if the proteins do not form a stable complex[177]. After the addition of biotin, the enzymatic moiety catalyses its activation

and reaction with nucleophilic lysine residues of proximal proteins, eventually biotinylating them and thus creating a covalent biotin tag for subsequent streptavidin-mediated purification and mass spectrometry for protein identification (**Figure 12**).

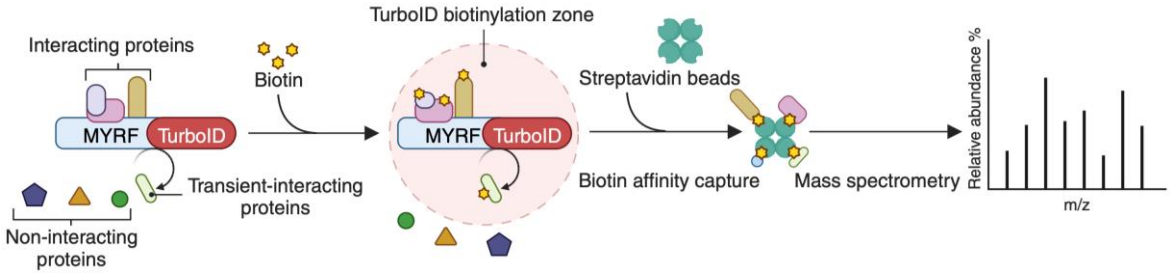

Figure 12: TurboID-catalyzed proximity labeling. Schematic workflow of proximity labeling in MYRF TurboID cells. Image created with Biorender.

In the last few years, proximity labeling approaches are considered as high-throughput methods for the investigation of protein-protein interactions. They are commonly based on BirA, a mutant of *E. coli* biotin ligase with slow kinetics requiring 18-24 hours of biotin labeling to produce sufficient biotinylated material[178,179]. In order to overcome the slow labeling kinetics of existing biotin ligase, new hyper-active variants of BirA have recently been generated. Among these, TurboID gives clearly detectable biotinylated product in 10 minutes of biotin labeling[157] and it is the biotin ligase that we chose for our study.

1.2 MYRF TurboID: experimental set up

For the identification of MYRF proximal proteins in the ER, we set up the TurboID-catalyzed proximity labeling in CFPAC-1 cells for which the role of MYRF in the homeostasis of the ER has been demonstrated[89].

In order to identify proteins proximal to the MYRF C-fragment in the ER lumen, we first fused TurboID with a V5 tag (37 kDa) to the MYRF C-terminus (M1-D1151), so that the TurboID is exposed to the lumen of the ER (*MYRF TurboID*) (**Figure 13A**). We also fused TurboID-V5 to the MYRF sequence lacking the fragment facing the ER lumen (M1-Y787) (*MYRF ΔER TurboID*) (**Figure 13B**). This represents our negative control to exclude proteins that non-specifically interact with that portion of MYRF possibly due to their abundance in the ER lumen.

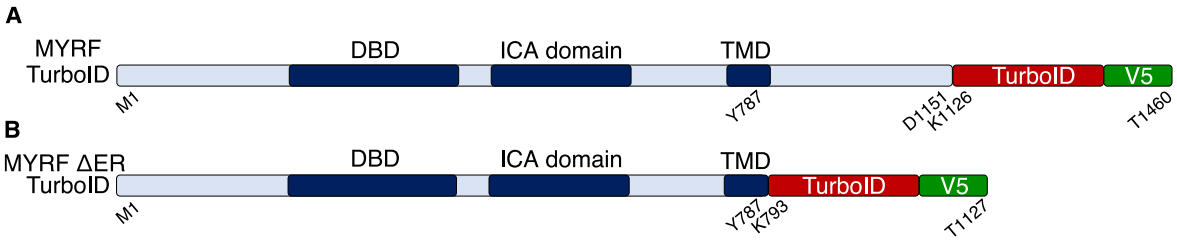

Figure 13: Structure of MYRF TurboID fusion proteins. Schematic representation of MYRF TurboID (**A**) and MYRF ΔER TurboID (**B**) with MYRF principal functional domains and aminoacidic positions indicated. DBD – DNA binding domain, ICA – Intramolecular Chaperone Auto-processing, TMD – transmembrane domain.

Each MYRF TurboID construct was generated by cloning into a lentiviral vector, which was then used to transduce CFPAC-1 cells. Specifically, cells were transduced with different concentrations of lentiviral vectors. Then, we selected the polyclonal population of CFPAC-1 expressing MYRF TurboID or MYRF ΔER TurboID protein at endogenous level (**Figure 14A**) to maintain conditions as physiological as possible.

We were also able to verify the localization of MYRF TurboID fusion proteins on the ER by immunofluorescence (**Figure 14B-C**), taking advantage of the V5 tag fused in frame to the biotin ligase.

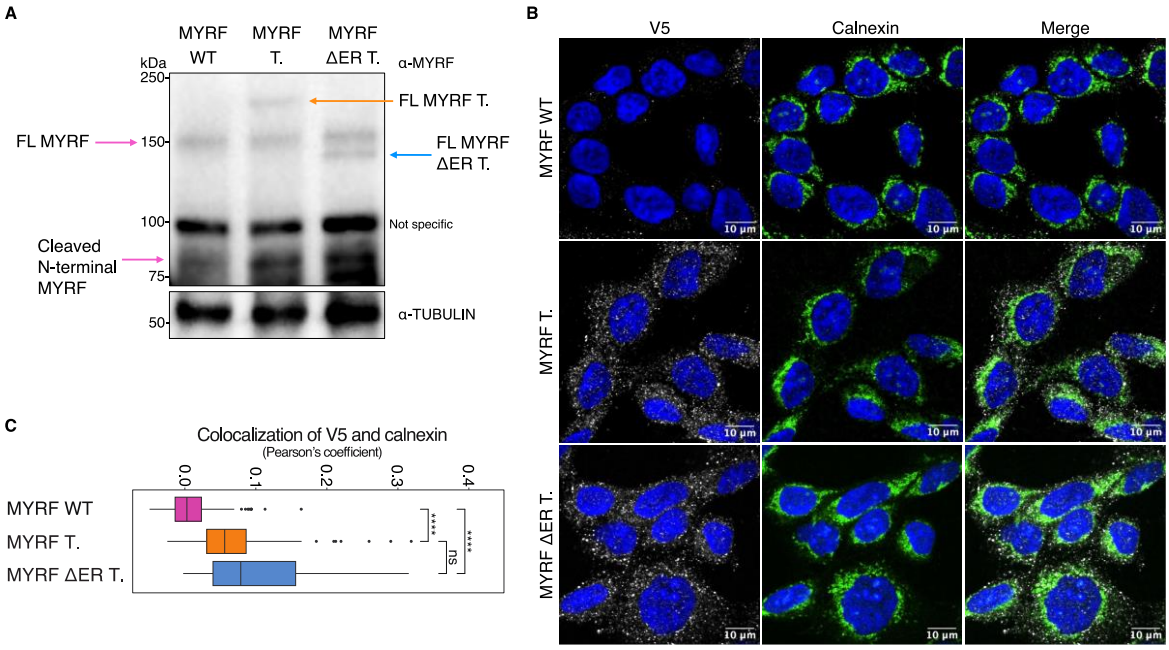

Figure 14: Expression of MYRF TurboID fusion proteins in CFPAC-1. A- Immunoblot analysis to evaluate the expression levels of MYRF TurboID (orange arrow, FL MYRF T., 180 kDa), and MYRF ΔER TurboID (light blue arrow, FL MYRF ΔER T., 125 kDa) compared to MYRF endogenous level (upper pink arrow, FL MYRF, 150 kDa). The molecular weight shift is due to the insertion of the TurboID. The pink arrow at the bottom left indicates the cleaved N-terminal MYRF fragment (80 kDa). Tubulin is shown as loading control. Molecular weight markers are indicated. **B-** Immunofluorescence on CFPAC-1 cells not infected (MYRF WT), infected with MYRF TurboID and MYRF ΔER TurboID vectors. Gray – V5, Green – Calnexin (ER marker), Blue – DAPI (nuclear marker). Merged images

are shown on the right. **C-** Box plot representing the colocalization analysis between V5 and calnexin in CFPAC-1 cells in spinning disk images. Every point represents the colocalization strength for a single cell for a total of 100 segmented cells on average. Two-sided Wilcoxon rank-sum tests (MYRF T. vs. WT or MYRF ΔER T., MYRF ΔER T. vs. WT) are shown.

TurboID leverages endogenous biotin present in cell culture media and serum to exhibit continuous biotinylation under standard cell culture conditions, comparable to the biotinylation levels observed in a typical induced enzyme-proximity labeling experiment[180]. In order to evaluate the proper inducibility of biotinylation by TurboID in our cell line, we cultured WT CFPAC-1 cells and CFPAC-1 expressing MYRF TurboID fusion proteins in different conditions of medium and serum. More specifically, we used IMDM or DMEM, a high- and low-biotin medium, respectively, and regular fetal bovine serum (FBS) or FBS from which biotin was removed by dialysis (dialyzed FBS, dFBS). We observed high basal level of biotinylation in cells expressing MYRF TurboID fusion proteins cultured in IMDM and regular FBS (*condition 1*) (**Figure 15**), due to the endogenous biotin present in the culture medium. On the contrary, culturing the same cells in DMEM and dFBS (*condition 2*) for 5 days (as suggested by literature[180]), the overall biotinylation decreased (**Figure 15**). Despite 5 days of culture without endogenous biotin, cells continued to show low levels of biotinylation, likely attributed to the recycling of biotin subsequent to the degradation of previously biotinylated proteins. Indeed, we still observed the biotinylation of known endogenous proteins that are constitutively biotinylated in mammalian cells: pyruvate carboxylase (130 kDa), 3-methylcrotonyl coA carboxylase (75 kDa) and propionyl coA carboxylase (72 kDa). Supplementing the cells deprived of endogenous biotin with exogenous biotin for 10 minutes (*condition 3*) slightly increased the biotinylation intensity (**Figure 15**). In particular, the biotinylation signal was clearly detectable in CFPAC-1 expressing MYRF TurboID compared to WT cells – not expressing biotin ligase – and cells expressing MYRF ΔER TurboID. We also tried to increase the biotin labelling time, but we did not observe an increase in the biotinylation signal with longer biotin incubation times (data not shown).

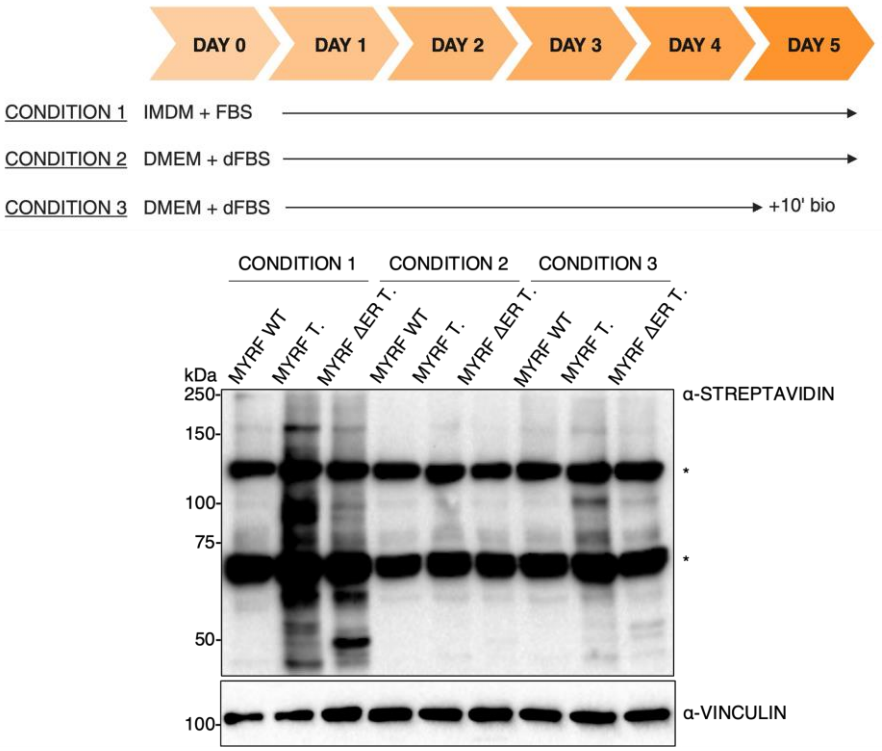

Figure 15: Biotinylation inducibility by TurboID in CFPAC-1. Top panel- Schematic representation of the cell culture conditions used for the experimental set up of TurboID-catalyzed proximity labeling in CFPAC-1. Image created with Biorender. Bottom panel- Streptavidin immunoblot analysis to test for the enrichment of biotinylated proteins in WT CFPAC-1 (MYRF WT) and CFPAC-1 expressing MYRF TurboID (MYRF T.) and MYRF ΔER TurboID (MYRF ΔER T.), all cultured in the three different conditions. Asterisks show known endogenous biotinylated proteins in mammalian cells. Vinculin is shown as loading control. Molecular weight markers are indicated.

In order to reduce the amount of steady-state biotinylation in cells constitutively expressing TurboID, we decided to set up the experimental approach by resorting to the use of low-biotin medium (DMEM) combined with dialyzed serum for 5 days to nearly completely deplete endogenous biotin and supplementation of exogenous biotin for 10 minutes to limit the extent of spurious biotinylation events.

Since the proximity labelling is based on the pull-down of biotinylated proteins, we validated the TurboID system performing the biotin pull-down in CFPAC-1 expressing MYRF TurboID. We cultured the cells in DMEM and dFBS for 1, 3 and 5 days and then we exposed cells to biotin for 10 minutes. By purifying biotinylated proteins with streptavidin-coated paramagnetic beads, the streptavidin immunoblot analysis deriving by the resulting pull-down showed a decrease of overall biotinylation starting at 3 days and higher after 5 days without endogenous biotin (**Figure 16**). Following a biotin incubation of 10 minutes, TurboID starts working again inducing a significant enrichment in the biotinylation signal (**Figure 16**). Due to the essential nature of biotin for cell survival, the duration of these depletion studies was not extended beyond 5 days.

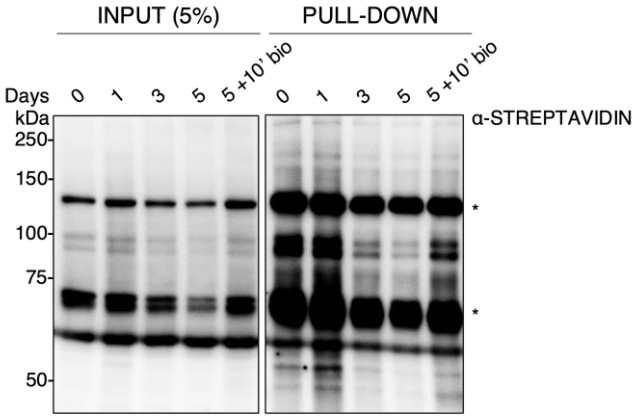

Figure 16: Biotin pull-down test. Streptavidin immunoblot showing a basal biotinylation signal in the pull-down of CFPAC-1 expressing MYRF TurboID and cultured in DMEM and dFBS for 0-1 days. After 3-5 days in dFBS, the biotinylation signal decreases and then increases after the addition of exogenous biotin for 10 minutes. Asterisks show known endogenous biotinylated proteins in mammalian cells. Molecular weight markers are indicated.

These data validated the feasibility of TurboID-catalyzed proximity labeling in this specific setting.

1.3 Identification of MYRF proximal proteins in the ER

In order to identify which MYRF C-terminal associated proteins were biotinylated by TurboID, we cultured WT CFPAC-1 and CFPAC-1 expressing MYRF TurboID fusion proteins as described above. The biotin-affinity pull-down performed under these cell culture conditions confirmed the proper functioning of the biotin ligase in cells expressing MYRF TurboID: indeed, the biotinylation signal was clearly detectable in the streptavidin immunoblot of these cells compared to WT cells and cells expressing MYRF ΔER TurboID (**Figure 17A**).

We further validated the system through immunofluorescence analysis where streptavidin signal was stronger in cells expressing MYRF TurboID and colocalized with calnexin, the ER marker protein (**Figure 17B-C**). These data confirmed that TurboID fused in frame to the MYRC C-terminus, then exposed to the ER lumen, was able to generate a biotinylation cloud in this compartment.

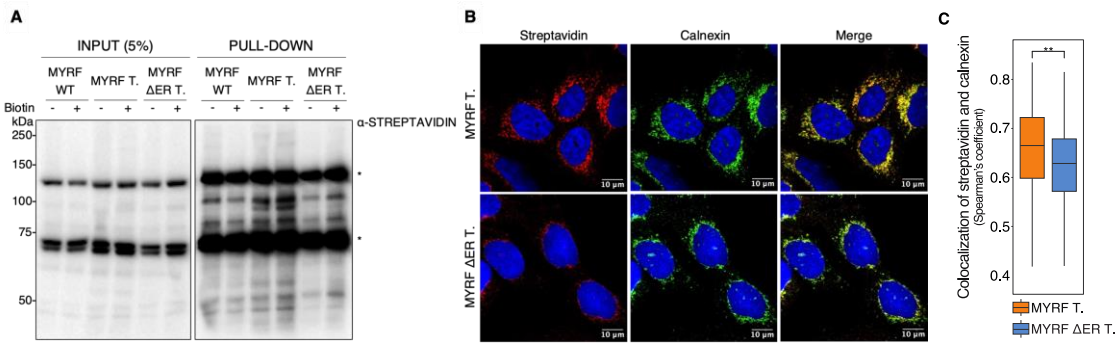

Figure 17: MYRF TurboID biotinylation in CFPAC-1. A- Streptavidin immunoblot showing biotinylated proteins enriched in the pull-down of CFPAC-1 expressing MYRF TurboID and cultured in condition 3. Asterisks show known endogenous biotinylated proteins in mammalian cells. Molecular weight markers are indicated. **B-** Immunofluorescence on CFPAC-1 cells expressing MYRF TurboID or MYRF ΔER TurboID cultured in condition 3. Red – Streptavidin, Green – Calnexin (ER marker), Blue – DAPI (nuclear marker). Merged images are shown on the right. **C-** Box plot representing the colocalization analysis between streptavidin and calnexin in CFPAC-1 cells in spinning disk images. Every point represents the colocalization strength for a single cell for a total of 100 segmented cells on average. Two-sided Wilcoxon rank-sum tests (MYRF T. vs. MYRF ΔER T.) are shown.

Captured proteins through the biotin-affinity pull-down were removed from the streptavidin beads using on-bead digestion, followed by mass spectrometry (LC-MS/MS) on four biological replicates to identify and quantify the relative amount of each biotinylated protein. We focused our attention only on the proteins that were: i) not detected in the common biotinylated background (WT CFPAC-1) and/or ii) enriched in CFPAC-1 expressing MYRF TurboID compared to the control sample (CFPAC-1 expressing MYRF ΔER TurboID) and iii) detected in all four biological replicates (fold change ≥ 2.5, q-value ≤ 0.05). We obtained a high-quality data-set with only 16 proteins significantly enriched over the control (**Table 1**), including MYRF since it autobiotinylates itself. As proteins with a biological function that may have greater relevance in the interaction with MYRF, we found *IFI16* (Gamma-interferon-inducible protein 16), *LRRC59* (Leucine-rich repeat-containing protein 59) and *UGGT1* (UDP-glucose:glycoprotein glucosyltransferase 1) (**Figure 18**).

Gene name	Protein name	Fold change	q-value
CDC5L	Cell division cycle 5-like protein	3,87	2,84E-02
MYRF	Myelin regulatory factor	3,86	2,48E-02
BCAS2	Pre-mRNA-splicing factor SPF27	3,82	1,44E-02
ZNF22	Zinc finger protein 22	3,79	3,30E-02
SSFA2	Sperm-specific antigen 2	3,60	0,00E+00
IFI16	Gamma-interferon-inducible protein 16	3,54	2,91E-02
SRSF4	Serine/arginine-rich splicing factor 4	3,49	3,23E-02
SERPINB12	Serpin B12	3,38	3,23E-02
LRRC59	Leucine-rich repeat-containing protein 59	3,22	3,26E-02
THOC6	THO complex subunit 6 homolog	3,14	2,80E-02
UGGT1	UDP-glucose:glycoprotein glucosyltransferase 1	3,06	3,13E-02
THOC1	THO complex subunit 1	2,96	2,57E-02
GAPDH	Glyceraldehyde-3-phosphate dehydrogenase	2,93	3,35E-02
RUNX1	Runt-related transcription factor 1	2,81	2,62E-02
TSPO	Translocator protein	2,62	3,00E-02
S100A9	Protein S100-A9	2,56	2,18E-02

Table 1: Biotinylated proteins enriched in CFPAC-1 expressing MYRF TurboID. List of proteins enriched in CFPAC-1 expressing MYRF TurboID compared to CFPAC-1 expressing MYRF ΔER TurboID (fold change ≥ 2.5, q-value ≤ 0.05). Biotin-depleted cells were treated with biotin for 10 minutes, followed by the preparation of total lysates, streptavidin pull-down and mass spectrometry. Highlighted proteins include MYRF, as top hit, and IFI16, LRRC59 and UGGT1, as potentially proximal proteins in the interaction with MYRF.

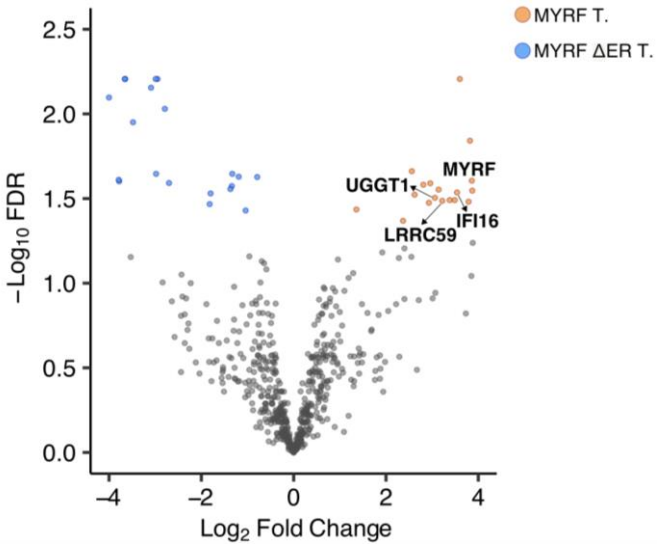

Figure 18: Identification of MYRF proximal proteins. Volcano plot showing the proteins identified by proximity labelling in CFPAC-1 expressing MYRF TurboID (orange) or MYRF ΔER TurboID (light blue). Biotin-depleted cells were treated with biotin for 10 minutes, followed by the preparation of total lysates, streptavidin pull-down and mass spectrometry. The identified proteins are shown according to their relative abundance (log_2 fold change) and statistical significance ($-log_{10}$ FDR) in CFPAC-1 expressing MYRF TurboID vs. MYRF ΔER TurboID. n=4 biological replicates. Selected statistically significant proteins are indicated.

2. UGGT1 is a potential regulator of MYRF activity

We currently focused our study on UGGT1. UGGT1 is a protein of the ER lumen that re-glucosylates unfolded glycoproteins, thus providing quality control for protein transport out of the ER[181,182]. One possibility is that MYRF could be a target of

UGGT1 as a protein to be properly folded. Alternatively, MYRF and UGGT1 may work together in the same pathway controlling the fate of misfolded proteins in the ER.

2.1 Validation of MYRF-UGGT1 interaction

We first validated the system through immunofluorescence analysis. A colocalization signal between UGGT1 and streptavidin was observed in CFPAC-1 expressing MYRF TurboID fusion protein and cultured in condition 3 (**Figure 19A-B**), confirming the biotinylation of UGGT1 by MYRF TurboID.

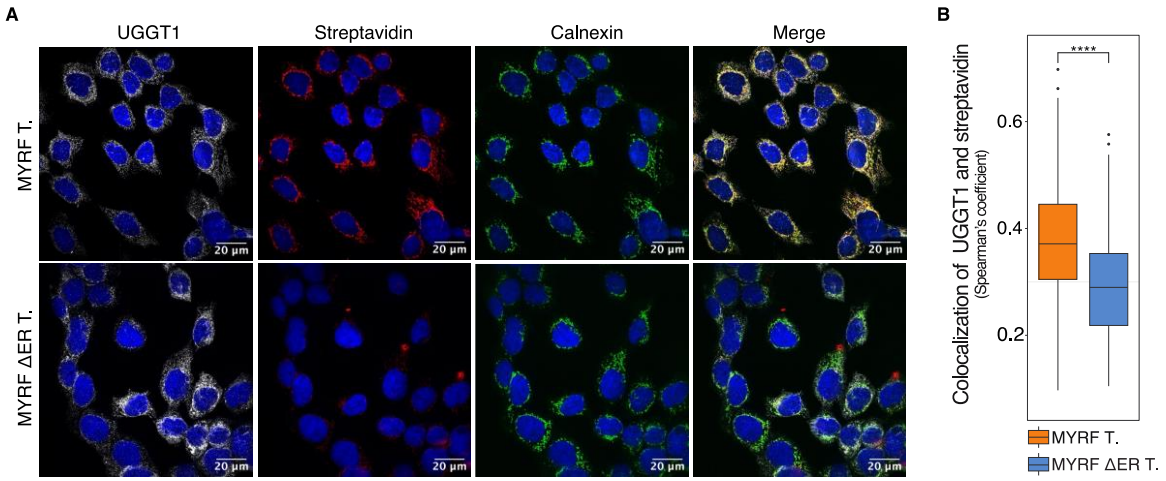

Figure 19: Biotinylation of UGGT1 by MYRF TurboID. A- Immunofluorescence on CFPAC-1 cells expressing MYRF TurboID or MYRF ΔER TurboID cultured in condition 3. Gray – UGGT1, Red – Streptavidin, Green – Calnexin (ER marker), Blue – DAPI (nuclear marker). Merged images are shown on the right. **B-** Box plot representing the colocalization analysis between UGGT1 and streptavidin in CFPAC-1 cells in spinning disk images. Every point represents the colocalization strength for a single cell for a total of 100 segmented cells on average. Two-sided Wilcoxon rank-sum tests (MYRF T. vs. MYRF ΔER T.) are shown.

Then, we validated the interaction between MYRF and UGGT1 by performing the biotin-affinity pull-down in the same conditions used for the identification of MYRF C-terminal proximal proteins in the ER. As expected, the immunoblot on lysates purified by streptavidin beads showed that UGGT1 was enriched in the pull-down of CFPAC-1 expressing MYRF TurboID, while missing in that of WT cells – not expressing biotin ligase – and cells expressing MYRF ΔER TurboID (**Figure 20**). Notably, UGGT1 was still detected despite 5 days of cell culture without endogenous biotin, probably due to a partial biotin depletion from the medium and serum. Moreover, as previously reported by MS analysis, we also observed MYRF TurboID fusion proteins in the pull-down of CFPAC-1 cells expressing them (**Figure 20**), confirming that they autobiotinylate. Specifically, by taking advantage of the V5 fused in the MYRF TurboID frame, we were able to observe the biotinylation of both

the full length and the cleaved fragment, the latter corresponding to the portion of MYRF which remains inserted in the ER after self-cleavage.

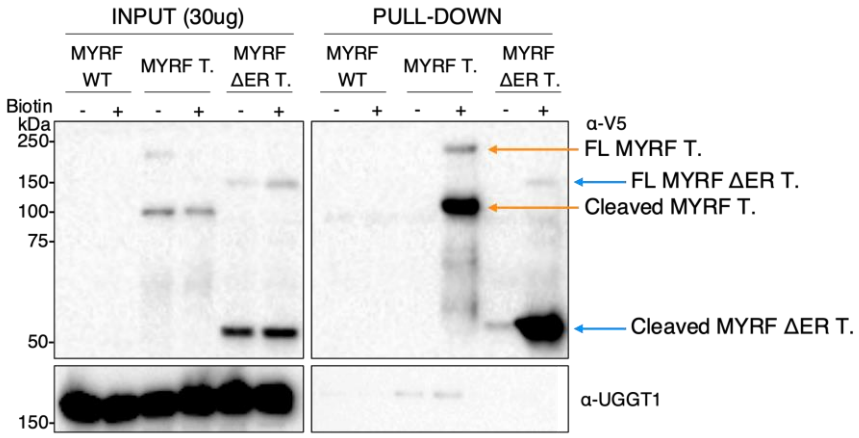

Figure 20: Validation of MYRF-UGGT1 interaction in CFPAC-1. Immunoblot analysis to confirm the enrichment of UGGT1 in the pull-down of CFPAC-1 expressing MYRF TurboID and the autobiotinylation of MYRF TurboID fusion proteins in the pull-downs of cells expressing them. Specifically, in the α-V5 immunoblot the orange arrows indicate the full length (FL MYRF T., 180 kDa) and cleaved C-terminal (cleaved MYRF T., 100 kDa) MYRF TurboID fragment, respectively; the light blue arrows indicate the full length (FL MYRF ΔER T., 125 kDa) and cleaved C-terminal (cleaved MYRF ΔER T., 45 kDa) MYRF ΔER TurboID fragment, respectively. All the cells were cultured without endogenous biotin ± exogenous biotin for 10 minutes, followed by the preparation of total lysates and streptavidin pull-down. Molecular weight markers are indicated.

2.2 Impact of UGGT1 depletion on MYRF transcriptional activity

To discern between the two hypothesis – MYRF as a target of UGGT1 or MYRF as a co-regulator of misfolded proteins together with UGGT1 – we started to generate CFPAC-1 cells knocking-out MYRF or UGGT1 in order to evaluate whether the deletion of one protein influences the expression of the other and vice versa.

We first observed that removing MYRF did not induce any changes in UGGT1 protein expression (**Figure 21**).

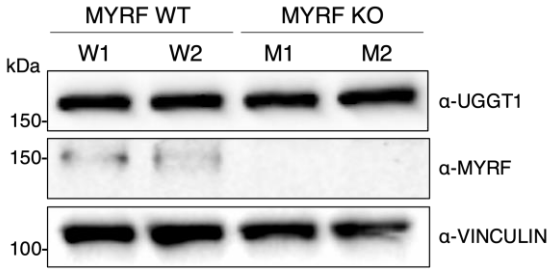

Figure 21: Effect of MYRF KO on UGGT1 protein expression. Immunoblot of 4 CFPAC-1 clones expressing (WT – W1, W2) and knocking-out (KO – M1, M2) MYRF. UGGT1 expression is not affected by the deletion of MYRF. Vinculin is shown as loading control. Molecular weight markers are indicated.

On the contrary, the KO of UGGT1 (**Figure 22A**) strongly reduced the expression of MYRF and TM4SF4 – one of the known MYRF targets[89] – and increased the

expression of ER-stress related genes (**Figure 22B**). All the genes tested were upregulated at comparable levels upon MYRF deletion, as described in our previous work[89].

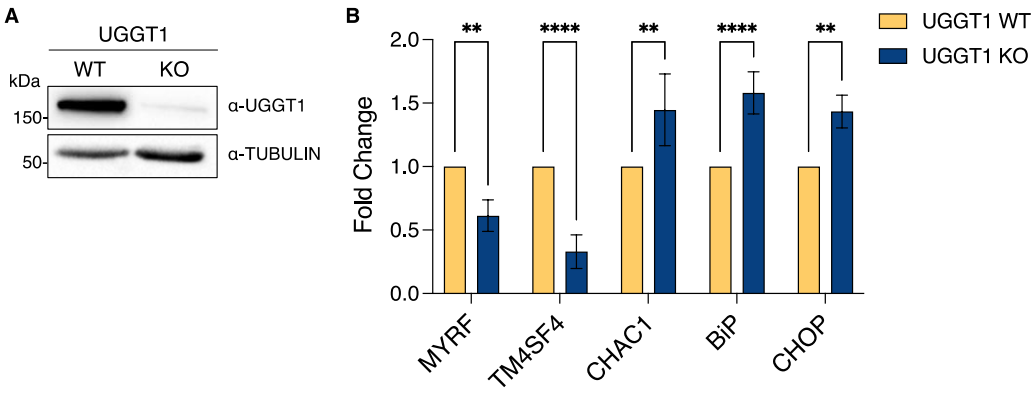

Figure 22: Effect of UGGT1 KO on MYRF and its targets expression. A- Immunoblot showing the deletion of UGGT1. Tubulin is shown as loading control. Molecular weight markers are indicated. **B-** Histograms showing qPCR expression levels of MYRF and its different targets upon UGGT1 deletion. Results deriving from the mean of three experiments are expressed as fold change relative to the respective controls. Standard deviations are shown. p values are derived from a two-way ANOVA test.

These preliminary data suggested that UGGT1 is a potential regulator of MYRF folding and it could affect the transcriptional activity of MYRF.

3. Dissection of the role of MYRF functional domains

The specific reason why MYRF is inserted in the ER membrane and the possible role of MYRF in controlling ER function independently of transcriptional regulation remain unaddressed. To obtain more detailed insight into the relative role of the transcription-activating function of MYRF vs. the role of the protein domain that remains ER-associated after intramolecular cleavage, we dissected the function of two critical domains of MYRF: the DBD and the TMD. In particular:

- an in-frame genomic deletion of the exons encoding the DBD resulted in the production of a protein able to trimerize and self-cleave – as it maintains the ICA domain – but unable to bind the DNA. In this way, we analyzed the role of transcriptional activity vs. non-transcriptional activities of MYRF;
- an in-frame genomic deletion of the exons encoding the TMD resulted in the production of a protein still capable to trimerize, self-cleave and bind the DNA but unable to be inserted into the ER membrane. We thus unveiled if any function is specifically associated with the ER-associated portion of MYRF.

3.1 Generation of MYRF domain-specific mutants

In this conceptual framework, we generated MYRF domain-specific mutants by CRISPR/Cas9-mediated genome editing. Achieving an in-frame deletion of a specific region within the endogenous gene poses a significant challenge, as it requires the removal of extensive genomic regions. Given that the majority of PDAC cell lines exhibit hyperdiploidy, we opted for an alternative cellular model to generate MYRF domain-specific mutants: *HCT116*, a diploid cell line derived from colorectal cancer, a tumor with endodermal origin and secretory activity like PDAC.
DBD and TMD include exons 7 to 12 and 17 to 18, in order (**Figure 23**). To ensure that the protein remains in frame after deletion, we targeted MYRF from exon 8 to 10 or from exon 16 to 18 to remove DBD or TMD, respectively.

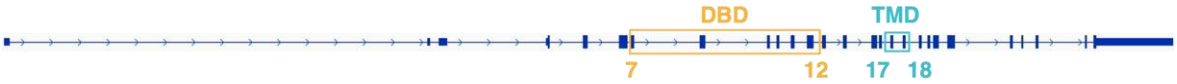

Figure 23: MYRF gene and its domain-specific exons. Representative snapshot of MYRF gene. The blue squares indicate the exons. The DBD and TMD exons are highlighted in yellow and light blue squares, respectively.

To obtain a homozygous deletion, cells lacking MYRF DBD or TMD were grown as single clones and screened by western blot analysis (**Figure 24**), based on the shift of the protein band after domain-specific deletion.

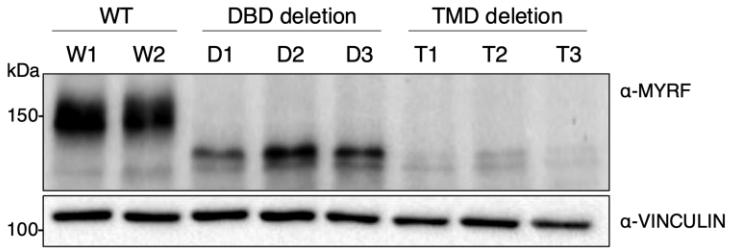

Figure 24: MYRF domain-specific mutants in HCT116. Immunoblot showing the shift of MYRF protein band in DBD (136 kDa) and TMD (141 kDa) deprived clones compared to WT clones (150 kDa). Vinculin is shown as loading control.

3.2 MYRF structural functionality in the ER

To study the impact of these domains on gene expression programs, we analysed 3 mutant clones for each domain by RNA-seq, comparing them with WT clones. Since MYRF has already been associated to the control of ER homeostasis, we checked the enrichment of a gene signature related to ER stress response extracted from the molecular signature database (MSigDB) (GO:0034976) in the respective mutants (**Figure 25A**). Only the TMD clones showed a strong enrichment of this signature and a significant upregulation of the most ER stress genes, previously observed induced in MYRF-deleted CFPAC-1 cells[89] (**Figure 25B**).

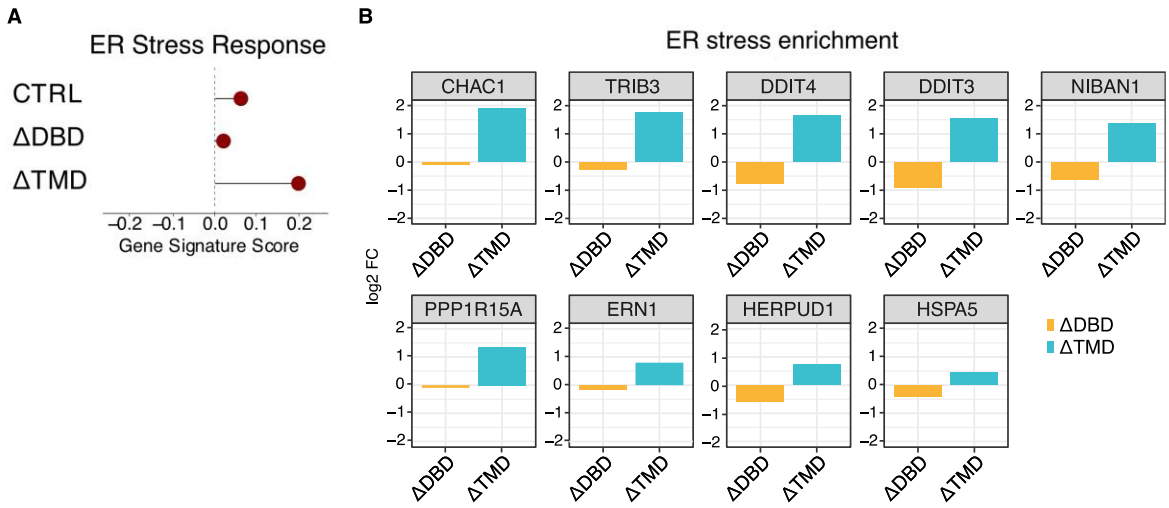

Figure 25: ER stress genes enrichment in HCT116 MYRF domain-specific mutants. A- Dot plot showing gene signature (GO: 0034976) score enriched in control (CTRL) and mutant clones. **B-** Circular bar plots showing the log_2 fold change value of MYRF ΔDBD or MYRF ΔTMD vs. MYRF WT (n=3 per group). Asterisks indicate statistically significant differentially expressed genes compared to WT cells (Log_2FC >0.58, FDR <0.05).

Therefore, after the deletion of TMD, upregulation of ER stress-related genes similar to those observed when completely removing MYRF protein became apparent. On the contrary, the removal of the DBD did not result in such evident alterations. Consequently, these data suggested that MYRF has a structural functionality in the ER that is independent of its transcriptional activity.

Having elucidated the regulatory mechanism and pivotal role of MYRF in glandular and secretory cells, we now shift our focus to explore the implications of SFRP2 in the context of PDAC transitional cells.

4. Induction of ECM synthesis, migration and invasion by SFRP2

Another crucial element of this study involves the discovery of novel regulators that influence PDAC properties. Notably, the contribution of transitional cancer cells to the remodelling of the extracellular matrix is a key feature, and SFRP2 may represent a potential regulator in this process.

The expression of SFRP2 appears to be associated with tumor aggressiveness and invasiveness[136]. More specifically, tumors characterized by stromal signatures and elevated mesenchymal gene expression show CAFs with increased expression of sFRP2[183]. This upregulation contributes to CAF-mediated enhancement of invasive properties of cancer cells, leading to an unfavourable prognosis, as evidenced in melanoma[140]. Given this observation, we explored the possibility that autonomous

production of SFRP2 by PDAC cells might enhance their migratory and invasive capabilities.

We first over-expressed the human SFRP2 protein into a primary human PDAC cell line in which the corresponding gene was not transcriptionally active. Specifically, *MDA-PATC69* cells were transfected with empty or SFRP2-FLAG vector and protein expression was analyzed by immunoblot (**Figure 26**). As expected, the endogenous expression of the protein was undetectable, while the overexpression of SFRP2 in cell lysates and the proper secretion in the culture supernatant (detected by FLAG immunoprecipitation) were confirmed.

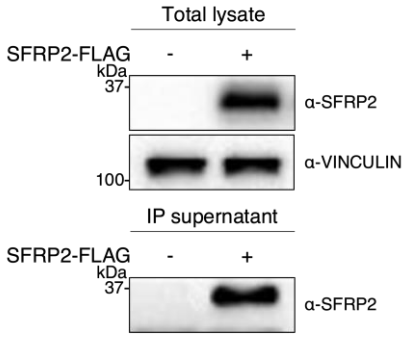

Figure 26: Expression of SFRP2 in MDA-PATC69. Immunoblot of total lysate (top) and of anti-FLAG immunoprecipitation (IP) from the culture supernatant (bottom) of the MDA-PATC69 primary human PDAC cell line transfected with either control or the SFRP2-FLAG vector. Vinculin is shown as loading control.

We then investigated the impact of the overexpressed and recombinant protein on extracellular matrix production by in-cell enzyme-linked immunosorbent assay (ELISA). This assay revealed a significant increase in the deposition of all analyzed proteins of the ECM (**Figure 27**), confirming the contribution of cancer cells in matrix remodelling.

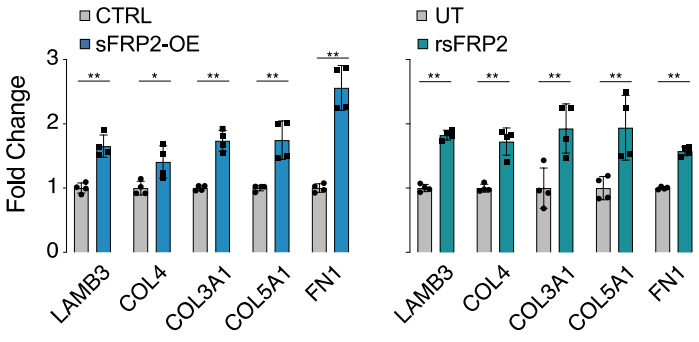

Figure 27: Effect of SFRP2 on ECM proteins synthesis. The abundance of the indicated ECM components was measured by ELISA in the MDA-PATC69 PDAC cell line transfected with control or SFRP2-FLAG vector (left bars) or treated with recombinant SFRP2 (right bars). Data (n = two biological replicates in duplicate) were normalized by DNA content and expressed as fold change relative to the respective controls. Means and standard deviations are shown. Two-tailed t tests are shown.

Cell-ECM interactions are known to generate active pulling forces via actin filament contraction, so we evaluated changes in cell motility. Random migration, chemotactic migration and matrigel invasion assays indicated markedly enhanced migratory and invasive capabilities in MDA-PATC69 PDAC cells overexpressing SFRP2 (**Figure 28**).

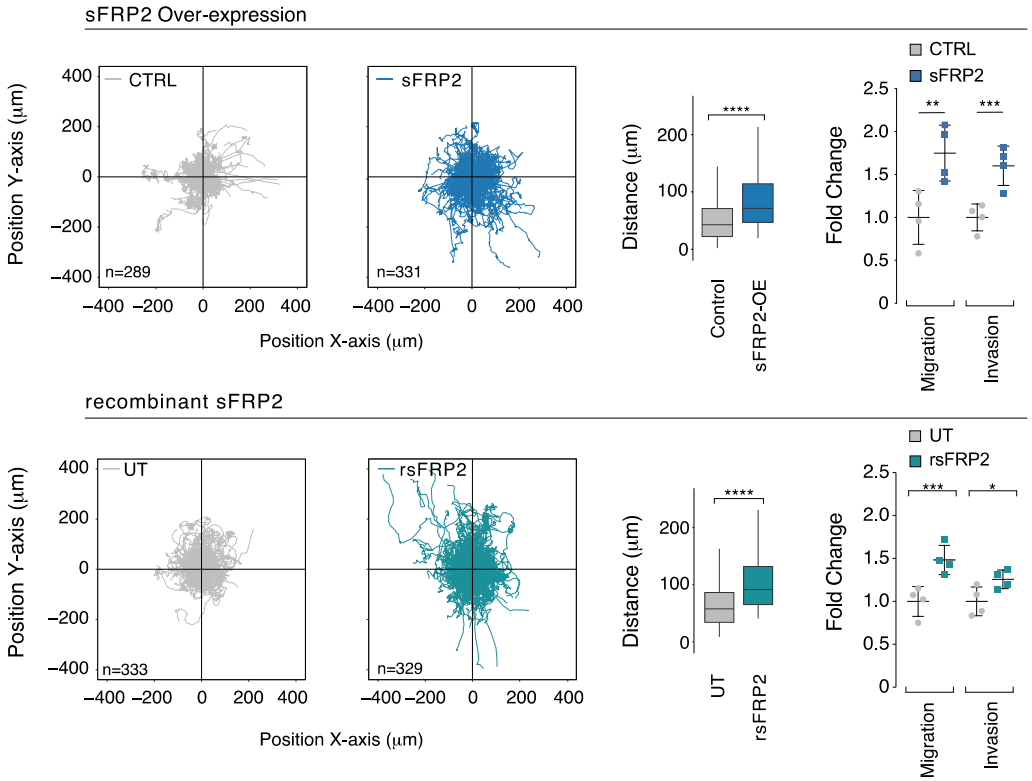

Figure 28: Effect of SFRP2 on invasion and migration of MDA-PATC69. Random migration, chemotaxis migration and matrigel invasion assays of control (CTRL) and overexpressed SFRP2 (sFRP2-oe) (top), and untreated (UT) and recombinant SFRP2 (rsFRP2) (bottom) in the MDA-PATC69 primary human PDAC cell line. Trajectory plots with maxima distance quantification were obtained by time-lapse microscopy in two independent experimental sets. Two-sided Wilcoxon rank-sum tests are shown. Chemotaxis migration and matrigel invasion assays are expressed as fold change compared to CTRL or UT (two independent experimental sets performed in duplicate). Means and standard deviations are shown. Two-tailed t tests are shown.

Taken all together, these findings suggest that the production of SFRP2 by transitional PDAC cells is implicated in both local extracellular matrix remodeling and development of migratory and invasive capabilities in PDAC cells.

5. Effects of ECM stiffness on transitional cells

5.1 ECM stiffness induces mechanoresponse and transitional programs

Since changes in the abundance of ECM components contributes to different tissue density and rigidity, leading cells to mount a mechanoresponsive transcription program[171], we investigated whether this phenomenon occurs in PDAC.

We cultured the MDA-PATC69 on collagen-coated hydrogels with defined rigidity of 1 and 12 kPa, mimicking the stiffness range recorded in PDAC[184]. Our findings revealed that the culture of PDAC cells in the 12kPa stiff matrix induced not only the mechanoresponse gene signature but also the transitional program, marked by the increased expression of components of the interstitial matrix and the myofibroblast marker gene ACTA2 (**Figure 29**).

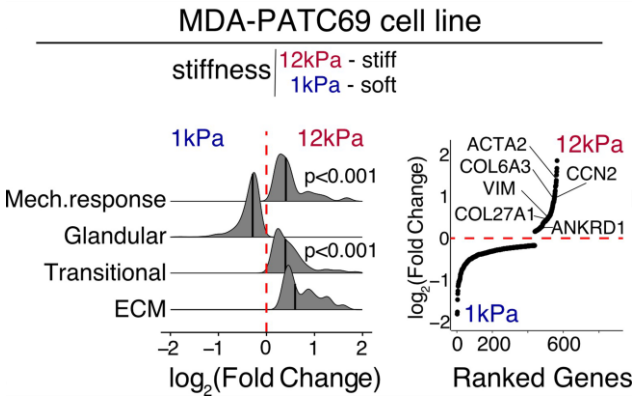

Figure 29: Induction of mechanoresponse and transitional signature in MDA-PATC69. Left: ridge plot showing log_2 fold change distributions of core enriched genes of the mechanoresponse program[171], the glandular biotype, the transitional biotype and a PDAC ECM signature[79] in MDA-PATC69 primary human PDAC cell line cultured in vitro on either a stiff (12kPa) or a soft (1kPa) matrix. Statistically significant enrichments are indicated. Right: log_2 fold change values of enriched genes ranked in ascending order. Selected enriched genes are highlighted.

These data suggest that the induction of the transitional biotype signature may occur in response to the local sensing of a stiff extracellular matrix.

4.1 PDAC heterogeneity in the thickness of collagen fibers

As described before, transitional biotype cells revealed elevated expression of genes associated with extracellular matrix. Therefore, we investigated whether the heightened production of ECM components by PDAC cells of the transitional biotype could directly impact the architectural organization of the surrounding stroma.

To address this point, we utilized second harmonic generation (SHG) signal generated by two-photon illumination[185], a technique that visualizes and quantifies the local supramolecular fibrillar structure of the collagen-rich matrix in PDAC regions of the three different biotypes. Collagen fibers are formed through a mass action-driven process, assembling individual molecules into fibrils (with a diameter ranging from 0.02 to 0.1 µm) and then into fibers (with a diameter from 0.5 to 20 µm). When collagen fibers are exposed to two low-energy incident photons, they generate a scattering photon with exactly twice the frequency of the incident light[186]. This non-linear optical phenomenon is directly proportional to the extent of collagen

assembly into fibers, providing a measurable indication of the architectural organization of the collagen network in specific tissue areas. Examining tissue regions around PDAC areas assigned to the three biotypes, fibers with a diverse range of thickness reproduced the heterogeneity of the PDAC tumor microenvironment (**Figure 30**). However, transitional areas exhibited a significant and substantial increase in thick collagen fibers in close proximity to tumor cells compared to glandular areas and undifferentiated areas (**Figure 30**).

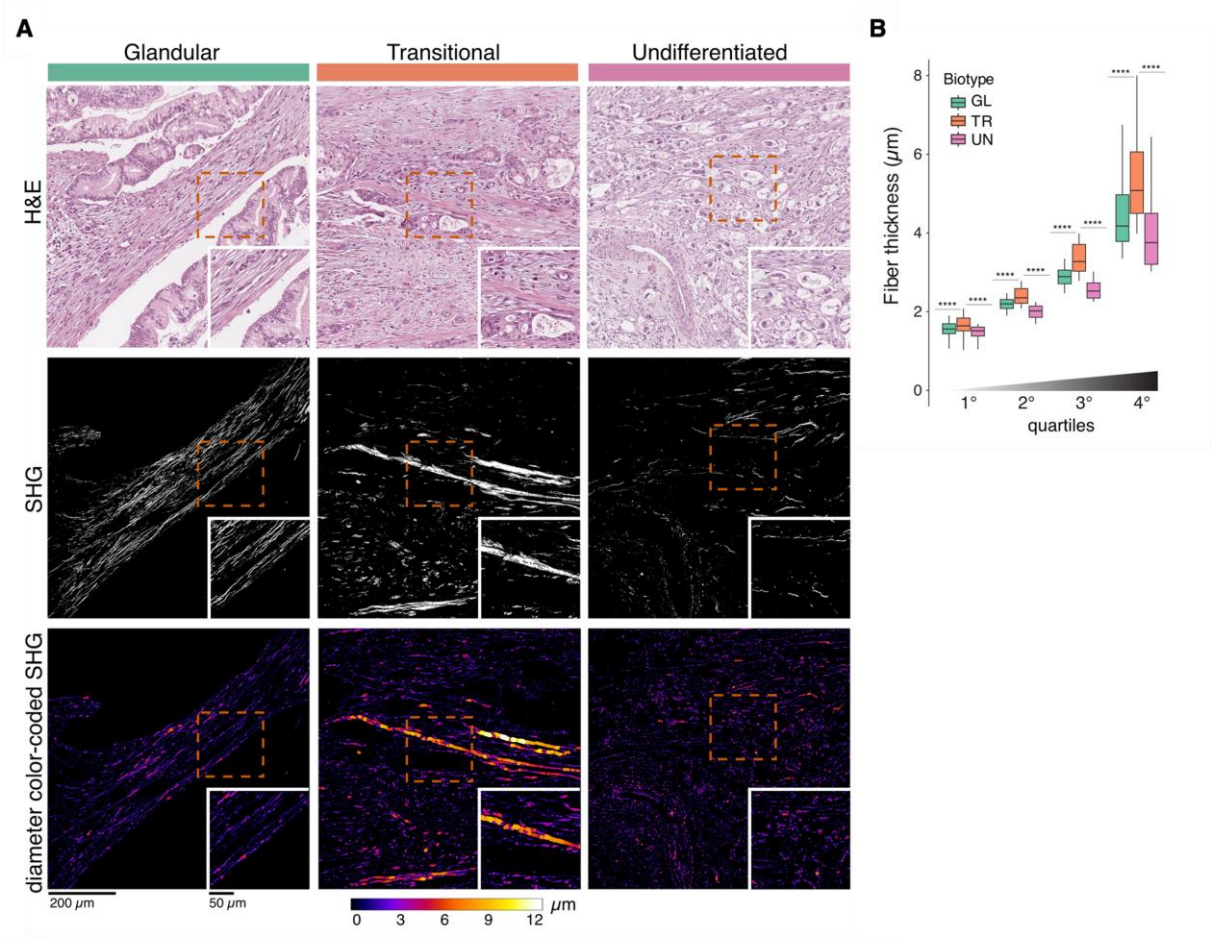

Figure 30: Collagen fibers thickness in PDAC morpho-biotypes. A- Hematoxylin-eosin (H&E) sections (top), second harmonic generation (SHG) images (middle) and diameter color-coded SHG images (bottom) of the PDAC tumor areas of the different biotypes. Orange boxes indicate the area shown at higher magnification on the bottom-right. Scale bars: H&E, SHG and diameter color-coded SHG images 200 μm, inserts 50 μm. **B-** Box plot showing the quantification of fiber thickness derived by SHG images for each tumor area and stratified by biotype (n=6 patients; 5 to 13 tumor areas per patient). The fiber thickness values were divided into 4 quartiles based on increasing signal. Two-sided Wilcoxon rank-sum tests (TR vs. GL; TR vs. UN) are shown. GL – Glandular, TR – Transitional, HY – Hybrid, UN – Undifferentiated.

Discussion

TurboID proximity labeling for MYRF interactome in the ER lumen

The foundation of cellular regulation and signal transduction lies in protein–protein, protein–RNA and protein–DNA interaction networks. Exploring these networks is crucial for understanding complex biological processes. Traditional methods, like affinity purification and yeast two-hybrid assays, have limitations as they isolate high-affinity molecular interactions mostly under nonphysiological conditions or *in vitro*. Additionally, these methods face challenges in organelle isolation and protein subcellular localization[187]. To overcome these issues, proximity labeling techniques have been introduced in proteomics[188]. Specifically, TurboID is the most active biotin ligase used for proximity labeling, even to identify the proteome in the endoplasmic reticulum lumen of cultured mammalian cells[189].

MYRF is an ER transmembrane protein and establishing interactions at the membrane level is challenging. Indeed, we tried to identify MYRF interactors trough the classical immunoprecipitation without obtaining convincing results (data not shown). We successfully optimized the TurboID-catalyzed proximity labeling in a secretory pancreatic cancer cell line to identify MYRF binding partners in the ER lumen.

MYRF interplay with UGGT1

MYRF is a key factor involved in ER stress adaptive responses of pancreatic secretory glandular cells. Indeed, its N-terminal fragment plays a fundamental role in transcriptional regulation of ER homeostasis[89]. Despite this observation, MYRF is still poorly characterized from a mechanistic and regulatory point of view.

Firstly, it is not well-established whether the cleavage of ER membrane-inserted MYRF is a regulated process, and if so, the mechanisms governing this regulation are not fully understood. To investigate this aspect, the study identified proteins that interact with MYRF C-fragment in the ER lumen revealing potential cleavage regulatory factors. Currently, we were able to obtain a first validation of MYRF interaction with UGGT1. Specifically, UGGT1 was enriched in the pull-down of CFPAC-1 cells expressing MYRF TurboID and treated with biotin, and its depletion influenced the transcription of MYRF and its target genes. Since UGGT1 provides

quality control for protein transport out of the ER[181,182], MYRF could be one of its targets as a protein to be properly folded.

Protein folding in the endoplasmic reticulum is error prone and ER quality control (ERQC) processes play a crucial role in ensuring that only the correctly folded proteins are allowed to exit the ER. Glycoproteins, in particular, can be retained in the ER as part of ERQC. The folding, trafficking and degradation of glycoproteins are intricately linked to the modification of N-linked oligosaccharides[190,191]. Upon translocation into the ER, a preassembled oligosaccharide core ($Glc_3Man_9GlcNAc_2$) is transferred to asparagine residues within the acceptor sequence Asn-Xxx-Ser/Thr. Subsequently, glucosidases I and II sequentially trim this core to a monoglucosylated glycan, facilitating binding to lectin-like ER chaperones calreticulin (CRT) and calnexin (CNX). Interaction with CRT and CNX serves to retain improperly folded substrates in the ER, constituting a crucial step in glycoprotein quality control. Upon release from CRT/CNX, glucosidase II can remove the remaining glucose, producing a product that, if folded correctly, traffics to the Golgi. If folded improperly, the protein may be directed to ERAD or undergo reglucosylation for renewed interaction with CRT/CNX[192]. UGGT1 serves as a central component of glycoprotein ERQC. It recognizes the incompletely folded glycoproteins to reglucosylates the deglucosylated N-glycan, facilitating their reassociation with lectin-like chaperones[191,192]. UGGT1 modified a broad range of proteins, with a particular preference for plasma membrane and transmembrane domain-containing proteins[193], such as MYRF. We suppose a model in which UGGT1 keeps MYRF in a monomeric state by binding it to the lectin-like chaperons. Then, when the monomeric MYRF is correctly folded, it is released, trimerizes and cleaved. It will therefore be interesting to study the effect of UGGT1 depletion on MYRF cleavage and understand whether UGGT1 represents a good regulator of MYRF folding and cleavage.

The role of MYRF in maintaining ER homeostasis in cancer cells carries significant implications for pancreatic cancer biology. Effective adaptation to ER stress provides cancer cells with selective advantages, including potential vulnerabilities that could be exploited for therapeutic purposes. In this context, disrupting MYRF function, such as by inhibiting or reducing trimerization, may intensify ER stress and enhance the susceptibility of cancer cells to drugs that counteract adaptation. Consequently, the function of UGGT1 as sensor for the folding of MYRF in the ER quality control machinery highlights its potential as a pharmacological target in PDAC.

MYRF C-terminal function

A second important issue that requires further investigation pertains to the function of the MYRF C-fragment residing in the luminal side of the ER, persisting in the ER membrane after self-cleavage and including most of the ICA domain, the TMD and the C-terminus. The induction of genes related to ER stress response after the deletion of MYRF transmembrane domain suggests that the C-terminal portion of MYRF could possibly have a structural functionality in the ER, independent of its transcriptional regulation. The deletion of MYRF TMD generated a low level of protein expression, comparable to that obtained after a full KO of MYRF. This may be due to proteins lacking TMD that are misfolded and then degraded by the proteasome. Therefore, removing MYRF from the ER membrane is sufficient to induce ER stress.

SFRP2 influencing transitional cells characteristics

PDAC cells with diverse gene expression programs exhibit distinct morphological characteristics, forming identifiable and spatially discrete tumor areas. This study involves the discovery of novel regulators influencing PDAC characteristics, potentially carrying practical importance and deserving further exploration. Significantly, the involvement of transitional cancer cells in the remodelling of the extracellular matrix is a crucial aspect, and SFRP2 may serve as a potential regulator in this process.

Recent findings suggest that SFRP2 functions as an anti-oncogene. For instance, its downregulation in gastric cancer, colorectal cancer, melanoma and oral squamous cell carcinoma due to epigenetic promoter hypermethylation indicates a potential role as a tumor suppressor[194]. However, a consistent body of research has also indicated that SFRP2 may exert tumor-promoting effects in certain types of cancer[195]. Indeed, it stimulates cell proliferation and decreases apoptosis in renal carcinoma cells, suggesting its oncogenic properties[196]. Furthermore, recent investigations have unveiled that the overexpression of SFRP2 can enhance the invasive, metastatic and therapeutic resistance capabilities in specific cancer types. Specifically, the increased expression of SFRP2 can drive melanoma metastasis and confer resistance to therapy by specifically influencing the tumor microenvironment through the activation of a multistep signalling cascade[140]. Moreover, SFRP2 levels are predictive of survival in pancreatic cancer[197].

In this study, we demonstrated the tumor promoting role of SFRP2 in human primary PDAC cells. The production of SFRP2 led to a noteworthy augmentation in the remodelling of the extracellular matrix, along with increased migratory and invasive capabilities of PDAC cells. Therefore, SFRP2 can be associated with metastasis of pancreatic cancer, holding significant promise for anti-cancer therapies.

A feed-forward loop for transitional biotype signature

PDAC transitional cells showed a remarkable phenomenon of progressive transdifferentiation into myofibroblasts with the production of structural and regulatory components of the extracellular matrix. Significantly, the production of matrisomal proteins by PDAC cells, rather than the stroma, is correlated with a negative prognosis[79,198] and may be associated with prometastatic effects[199], underscoring the central relevance of PDAC cell transdifferentiation towards a myofibroblast-like phenotype in disease progression and metastasis.

We demonstrated that the transitional reprogramming was also characterized by the generation of dense, thick collagen fibers in the immediate surrounding tissue and that the initiation of the transitional signature may result from the local detection of a stiff extracellular matrix. A typical mechanoresponse to elevated stiffness and mechanical stresses is the increased accumulation of fibrous extracellular matrix. This reaction potentially triggers a feed-forward loop wherein the heightened mechanical stress from a rigid matrix induces the transitional program. Simultaneously, increased matrisomal protein synthesis exacerbates tissue desmoplasia and fibrosis[200]. The desmoplastic matrix, in turn, supports tumor invasive characteristics and dissemination through various mechanisms, including facilitating the active migration of tumor cells along immobilized gradients of fibronectin and collagen fibers (haptotaxis)[200,201].

In conclusion, we have provided evidence for the mechanism of two regulators involved in two important hallmarks of pancreatic cancer – the ER stress and the ECM remodelling – that can be exploited for therapeutic purpose.

References

1. Siegel Mph, R. L., Giaquinto, A. N., Ahmedin, |, Dvm, J. & Siegel, R. L. Cancer statistics, 2024. *CA Cancer J Clin* 74, 12–49 (2024).

2. Cancer Facts & Figures 2024 | American Cancer Society. https://www.cancer.org/research/cancer-facts-statistics/all-cancer-facts-figures/2024-cancer-facts-figures.html.
3. Bengtsson, A., Andersson, R. & Ansari, D. The actual 5-year survivors of pancreatic ductal adenocarcinoma based on real-world data. *Sci Rep* 10, (2020).
4. Conroy, T. *et al.* FOLFIRINOX versus gemcitabine for metastatic pancreatic cancer. *N Engl J Med* 364, 1817–1825 (2011).
5. Von Hoff, D. D. *et al.* Increased survival in pancreatic cancer with nab-paclitaxel plus gemcitabine. *N Engl J Med* 369, 1691–1703 (2013).
6. Pancreatic Adenocarcinoma. *https://doi.org/10.1056/NEJMc1412266* 371, 2139–2141 (2014).
7. Alexandrov, L. B. *et al.* Signatures of mutational processes in human cancer. *Nature 2013 500:7463* 500, 415–421 (2013).
8. Petersen, G. M. Familial pancreatic cancer. *Semin Oncol* 43, 548–553 (2016).
9. Klein, A. P. *et al.* Prospective risk of pancreatic cancer in familial pancreatic cancer kindreds. *Cancer Res* 64, 2634–2638 (2004).
10. Brune, K. A. *et al.* Importance of Age of Onset in Pancreatic Cancer Kindreds. *JNCI Journal of the National Cancer Institute* 102, 119 (2010).
11. Grant, R. C. *et al.* Prevalence of germline mutations in cancer predisposition genes in patients with pancreatic cancer. *Gastroenterology* 148, 556–564 (2015).
12. Shindo, K. *et al.* Deleterious Germline Mutations in Patients With Apparently Sporadic Pancreatic Adenocarcinoma. *J Clin Oncol* 35, 3382–3390 (2017).
13. Humphris, J., Chang, D. K. & Biankin, A. V. Inherited susceptibility to pancreatic cancer in the era of next-generation sequencing. *Gastroenterology* 148, 496–498 (2015).
14. Yurgelun, M. B. *et al.* Germline cancer susceptibility gene variants, somatic second hits, and survival outcomes in patients with resected pancreatic cancer. *Genet Med* 21, 213–223 (2019).
15. Basturk, O. *et al.* A Revised Classification System and Recommendations From the Baltimore Consensus Meeting for Neoplastic Precursor Lesions in the Pancreas. *Am J Surg Pathol* 39, 1730–1741 (2015).
16. Hruban, R. H. *et al.* An illustrated consensus on the classification of pancreatic intraepithelial neoplasia and intraductal papillary mucinous neoplasms. *Am J Surg Pathol* 28, 977–987 (2004).
17. Hruban, R. H. *et al.* Pancreatic intraepithelial neoplasia: a new nomenclature and classification system for pancreatic duct lesions. *Am J Surg Pathol* 25, 579–586 (2001).
18. Naqvi, A. A. T., Hasan, G. M. & Hassan, M. I. Investigating the role of transcription factors of pancreas development in pancreatic cancer. *Pancreatology* 18, 184–190 (2018).
19. Naqvi, A. A. T., Hasan, G. M. & Hassan, M. I. Investigating the role of transcription factors of pancreas development in pancreatic cancer. *Pancreatology* 18, 184–190 (2018).
20. Maitra, A. & Hruban, R. H. Pancreatic Cancer. *Annu Rev Pathol* 3, 157 (2008).
21. Rajalingam, K., Schreck, R., Rapp, U. R. & Albert, Š. Ras oncogenes and their downstream targets. *Biochim Biophys Acta* 1773, 1177–1195 (2007).
22. Waddell, N. *et al.* Whole genomes redefine the mutational landscape of pancreatic cancer. *Nature 2015 518:7540* 518, 495–501 (2015).
23. Li, O. *et al.* Biological characteristics of pancreatic ductal adenocarcinoma: Initiation to malignancy, intracellular to extracellular. *Cancer Lett* 574, 216391 (2023).

24. Buscail, L., Bournet, B. & Cordelier, P. Role of oncogenic KRAS in the diagnosis, prognosis and treatment of pancreatic cancer. *Nature Reviews Gastroenterology & Hepatology 2020 17:3* 17, 153–168 (2020).
25. Hayashi, A., Hong, J. & Iacobuzio-Donahue, C. A. The pancreatic cancer genome revisited. *Nature Reviews Gastroenterology & Hepatology 2021 18:7* 18, 469–481 (2021).
26. Lane, D. & Levine, A. P53 Research: The Past Thirty Years and the Next Thirty Years. *Cold Spring Harb Perspect Biol* 2, (2010).
27. Kastenhuber, E. R. & Lowe, S. W. Putting p53 in Context. *Cell* 170, 1062–1078 (2017).
28. Raphael, B. J. *et al.* Integrated Genomic Characterization of Pancreatic Ductal Adenocarcinoma. *Cancer Cell* 32, 185 (2017).
29. Grant, T. J., Hua, K. & Singh, A. Molecular Pathogenesis of Pancreatic Cancer. *Prog Mol Biol Transl Sci* 144, 241–275 (2016).
30. Guo, J., Xie, K. & Zheng, S. Molecular Biomarkers of Pancreatic Intraepithelial Neoplasia and Their Implications in Early Diagnosis and Therapeutic Intervention of Pancreatic Cancer. *Int J Biol Sci* 12, 292 (2016).
31. Gil, Y. R., Sánchez, P. J., Velasco, R. M., García, A. G. & Lobo, V. J. S. A. Molecular Alterations in Pancreatic Cancer: Transfer to the Clinic. *Int J Mol Sci* 22, 1–16 (2021).
32. Kanda, M. *et al.* Mutant TP53 in Duodenal Samples of Pancreatic Juice From Patients With Pancreatic Cancer or High-Grade Dysplasia. *Clinical Gastroenterology and Hepatology* 11, 719-730.e5 (2013).
33. Bauer, J., Sporn, J. C., Cabral, J., Gomez, J. & Jung, B. Effects of Activin and TGFβ on p21 in Colon Cancer. *PLoS One* 7, e39381 (2012).
34. Alvarez, M. A., Freitas, J. P., Mazher Hussain, S. & Glazer, E. S. TGF-β Inhibitors in Metastatic Pancreatic Ductal Adenocarcinoma. *J Gastrointest Cancer* 50, 207–213 (2019).
35. Aiello, N. M. *et al.* EMT Subtype Influences Epithelial Plasticity and Mode of Cell Migration. *Dev Cell* 45, 681-695.e4 (2018).
36. Hahn, S. A. *et al.* DPC4, A Candidate Tumor Suppressor Gene at Human Chromosome 18q21.1. *Science (1979)* 271, 350–353 (1996).
37. Loss of Expression of Dpc4 in Pancreatic Intraepithelial Neoplasia: Evidence That DPC4 Inactivation Occurs Late in Neoplastic Progression1 | Cancer Research | American Association for Cancer Research. https://aacrjournals.org/cancerres/article/60/7/2002/507164/Loss-of-Expression-of-Dpc4-in-Pancreatic.
38. Serrano, M., Hannon, G. J. & Beach, D. A new regulatory motif in cell-cycle control causing specific inhibition of cyclin D/CDK4. *Nature 1993 366:6456* 366, 704–707 (1993).
39. Seo, J., Seong, D., Lee, S. R., Oh, D. B. & Song, J. Post-Translational Regulation of ARF: Perspective in Cancer. *Biomolecules 2020, Vol. 10, Page 1143* 10, 1143 (2020).
40. Maitra, A., Kern, S. E. & Hruban, R. H. Molecular pathogenesis of pancreatic cancer. *Best Pract Res Clin Gastroenterol* 20, 211–226 (2006).
41. Li, D. *et al.* Pathway analysis of genome-wide association study data highlights pancreatic development genes as susceptibility factors for pancreatic cancer. *Carcinogenesis* 33, 1384–1390 (2012).
42. Kaneto, H. *et al.* Expression of Heparin-binding Epidermal Growth Factor-like Growth Factor during Pancreas Development: A POTENTIAL ROLE OF PDX-1 IN TRANSCRIPTIONAL ACTIVATION. *Journal of Biological Chemistry* 272, 29137–29143 (1997).

43. Roy, N. *et al.* PDX1 dynamically regulates pancreatic ductal adenocarcinoma initiation and maintenance. *Genes Dev* 30, 2669–2683 (2016).
44. Kopp, J. L. *et al.* Identification of Sox9-Dependent Acinar-to-Ductal Reprogramming as the Principal Mechanism for Initiation of Pancreatic Ductal Adenocarcinoma. *Cancer Cell* 22, 737–750 (2012).
45. Hayashi, A., Hong, J. & Iacobuzio-Donahue, C. A. The pancreatic cancer genome revisited. *Nat Rev Gastroenterol Hepatol* 18, 469–481 (2021).
46. Waddell, N. *et al.* Whole genomes redefine the mutational landscape of pancreatic cancer. *Nature* 518, 495–501 (2015).
47. Shain, A. H. *et al.* Convergent structural alterations define SWItch/Sucrose NonFermentable (SWI/SNF) chromatin remodeler as a central tumor suppressive complex in pancreatic cancer. *Proc Natl Acad Sci U S A* 109, (2012).
48. Verbeke, C. Morphological heterogeneity in ductal adenocarcinoma of the pancreas - Does it matter? *Pancreatology* 16, 295–301 (2016).
49. Kalimuthu, S. N. *et al.* Morphological classification of pancreatic ductal adenocarcinoma that predicts molecular subtypes and correlates with clinical outcome. *Gut* 69, 317–328 (2020).
50. Moffitt, R. A. *et al.* Virtual microdissection identifies distinct tumor- and stroma-specific subtypes of pancreatic ductal adenocarcinoma. *Nature Genetics 2015 47:10* 47, 1168–1178 (2015).
51. Burrell, R. A. & Swanton, C. Tumour heterogeneity and the evolution of polyclonal drug resistance. (2014) doi:10.1016/j.molonc.2014.06.005.
52. Lou, E., Subramanian, S. & Steer, C. J. Pancreatic cancer: modulation of KRAS, MicroRNAs, and intercellular communication in the setting of tumor heterogeneity. *Pancreas* 42, 1218–1226 (2013).
53. Chan-Seng-Yue, M. *et al.* Author Correction: Transcription phenotypes of pancreatic cancer are driven by genomic events during tumor evolution (Nature Genetics, (2020), 52, 2, (231-240), 10.1038/s41588-019-0566-9). *Nat Genet* 52, 463 (2020).
54. Basal-like and classical cells coexist in pancreatic cancer revealed by single-cell analysis on biopsy-derived pancreatic cancer organoids from the classical subtype | Enhanced Reader.
55. Collisson, E. A. *et al.* Subtypes of pancreatic ductal adenocarcinoma and their differing responses to therapy. (2011) doi:10.1038/nm.2344.
56. Collisson, E. A., Bailey, P., Chang, D. K. & Biankin, A. V. Molecular subtypes of pancreatic cancer. *Nat Rev Gastroenterol Hepatol* doi:10.1038/s41575-019-0109-y.
57. Decker, K., Goldman, D. C., L. Grasch, C. & Sussel, L. Gata6 is an important regulator of mouse pancreas development. *Dev Biol* 298, 415 (2006).
58. Bailey, P. *et al.* Genomic analyses identify molecular subtypes of pancreatic cancer. *Nature 2016 531:7592* 531, 47–52 (2016).
59. Bailey, P. *et al.* Tissue Pathology and Diagnostic Oncology, Royal Prince Alfred Hospital, Camperdown New South Wales 2050, Australia. 21 School of Medicine. *Nature* 5000, 26 (2016).
60. Nieto, M. A., Huang, R. Y. Y. J., Jackson, R. A. A. & Thiery, J. P. P. EMT: 2016. *Cell* 166, 21–45 (2016).
61. Raghavan, S. *et al.* Microenvironment drives cell state, plasticity, and drug response in pancreatic cancer. *Cell* 184, 6119-6137.e26 (2021).
62. Grünwald, B. T. *et al.* Spatially confined sub-tumor microenvironments in pancreatic cancer. *Cell* 184, 5577-5592.e18 (2021).
63. Raghavan, S. *et al.* Microenvironment drives cell state, plasticity, and drug response in pancreatic cancer. *Cell* 184, 6119-6137.e26 (2021).

64. Whatcott, C. J. *et al.* Desmoplasia in Primary Tumors and Metastatic Lesions of Pancreatic Cancer. *Clin Cancer Res* 21, 3561–3568 (2015).
65. Masugi, Y. The Desmoplastic Stroma of Pancreatic Cancer: Multilayered Levels of Heterogeneity, Clinical Significance, and Therapeutic Opportunities. *Cancers (Basel)* 14, (2022).
66. Chu, G. C., Kimmelman, A. C., Hezel, A. F. & DePinho, R. A. Stromal biology of pancreatic cancer. *J Cell Biochem* 101, 887–907 (2007).
67. Ho, W. J., Jaffee, E. M. & Zheng, L. The tumour microenvironment in pancreatic cancer — clinical challenges and opportunities. *Nat Rev Clin Oncol* 17, 527 (2020).
68. Biffi, G. *et al.* IL1-Induced JAK/STAT Signaling Is Antagonized by TGFβ to Shape CAF Heterogeneity in Pancreatic Ductal Adenocarcinoma. *Cancer Discov* 9, 282–301 (2019).
69. Zhao, Z., Li, T., Sun, L., Yuan, Y. & Zhu, Y. Potential mechanisms of cancer-associated fibroblasts in therapeutic resistance. *Biomedicine & Pharmacotherapy* 166, 115425 (2023).
70. Manoukian, P., Bijlsma, M. & van Laarhoven, H. The Cellular Origins of Cancer-Associated Fibroblasts and Their Opposing Contributions to Pancreatic Cancer Growth. *Front Cell Dev Biol* 9, (2021).
71. Apte, M. V *et al.* Pancreatic stellate cells are activated by proinflammatory cytokines: implications for pancreatic fibrogenesis.
72. Gabbiani, G. The myofibroblast in wound healing and fibrocontractive diseases. *J Pathol* 200, 500–503 (2003).
73. Vonlaufen, A. *et al.* Pancreatic stellate cells: partners in crime with pancreatic cancer cells. *Cancer Res* 68, 2085–2093 (2008).
74. Apte, M. V. *et al.* Pancreatic stellate cells are activated by proinflammatory cytokines: implications for pancreatic fibrogenesis. *Gut* 44, 534–541 (1999).
75. Ullman, N. A., Burchard, P. R., Dunne, R. F. & Linehan, D. C. Immunologic Strategies in Pancreatic Cancer: Making Cold Tumors Hot. *Journal of Clinical Oncology* 40, 2789 (2022).
76. Bayne, L. J. *et al.* Tumor-derived granulocyte-macrophage colony-stimulating factor regulates myeloid inflammation and T cell immunity in pancreatic cancer. *Cancer Cell* 21, 822–835 (2012).
77. Carvalho, P. D. *et al.* KRAS Oncogenic Signaling Extends beyond Cancer Cells to Orchestrate the Microenvironment. *Cancer Res* 78, 7–14 (2018).
78. Falcomatà, C. *et al.* Context-Specific Determinants of the Immunosuppressive Tumor Microenvironment in Pancreatic Cancer. *Cancer Discov* 13, 278 (2023).
79. Tian, C. *et al.* Proteomic analyses of ECM during pancreatic ductal adenocarcinoma progression reveal different contributions by tumor and stromal cells. *Proc Natl Acad Sci U S A* 116, 19609–19618 (2019).
80. Jones, L. E., Humphreys, M. J., Campbell, F., Neoptolemos, J. P. & Boyd, M. T. Comprehensive Analysis of Matrix Metalloproteinase and Tissue Inhibitor Expression in Pancreatic CancerIncreased Expression of Matrix Metalloproteinase-7 Predicts Poor Survival. *Clinical Cancer Research* 10, 2832–2845 (2004).
81. Matsuyama, Y., Takao, S. & Aikou, T. Comparison of Matrix Metalloproteinase Expression Between Primary Tumors With or Without Liver Metastasis in Pancreatic and Colorectal Carcinomas. *J Surg Oncol* 80, 105–110 (2002).
82. Imbalance of expression of matrix metalloproteinases (MMPs) and tissue inhibitors of the matrix metalloproteinases (TIMPs) in human pancreatic carcinoma - Bramhall - 1997 - The Journal of Pathology - Wiley Online Library. https://pathsocjournals.onlinelibrary.wiley.com/doi/abs/10.1002/%28SICI%291096-9896%28199707%29182%3A3%3C347%3A%3AAID-PATH848%3E3.0.CO%3B2-J.

83. Grünwald, B. T. *et al.* Spatially confined sub-tumor microenvironments in pancreatic cancer. *Cell* 184, 5577-5592.e18 (2021).
84. Chiaro, P. Di *et al.* Coexisting morpho-biotypes unveil the regulatory bases of phenotypic plasticity in pancreatic ductal adenocarcinoma. *bioRxiv* 2022.12.12.520054 (2022) doi:10.1101/2022.12.12.520054.
85. Verbeke, C. Morphological heterogeneity in ductal adenocarcinoma of the pancreas - Does it matter? *Pancreatology* 16, 295–301 (2016).
86. Halbrook, C. J., Lyssiotis, C. A., Pasca di Magliano, M. & Maitra, A. Pancreatic cancer: Advances and challenges. *Cell* 186, 1729–1754 (2023).
87. Wilkerson, M. D. & Hayes, D. N. ConsensusClusterPlus: a class discovery tool with confidence assessments and item tracking. *Bioinformatics* 26, 1572 (2010).
88. Damanakis, A. I. *et al.* The Hallmarks of Pancreatic Cancer. *Textbook of Pancreatic Cancer* 189–201 (2021) doi:10.1007/978-3-030-53786-9_14.
89. Milan, M. *et al.* Pancreatic Cancer Cells Require the Transcription Factor MYRF to Maintain ER Homeostasis. *Dev Cell* 55, 398-412.e7 (2020).
90. Zaret, K. S. & Grompe, M. Generation and Regeneration of Cells of the Liver and Pancreas. *Science* 322, 1490 (2008).
91. Dufort, D., Schwartz, L., Harpal, K. & Rossant, J. The transcription factor HNF3beta is required in visceral endoderm for normal primitive streak morphogenesis. *Development* 125, 3015–3025 (1998).
92. Ang, S. L. *et al.* The formation and maintenance of the definitive endoderm lineage in the mouse: involvement of HNF3/forkhead proteins. *Development* 119, 1301–1315 (1993).
93. Milan, M. *et al.* FOXA2 controls the cis-regulatory networks of pancreatic cancer cells in a differentiation grade-specific manner. *EMBO J* 38, (2019).
94. Gao, N. *et al.* Dynamic regulation of Pdx1 enhancers by Foxa1 and Foxa2 is essential for pancreas development. *Genes Dev* 22, 3435 (2008).
95. Allen, H. L. *et al.* GATA6 haploinsufficiency causes pancreatic agenesis in humans. *Nat Genet* 44, 20 (2012).
96. Martinelli, P. *et al.* Original article: GATA6 regulates EMT and tumour dissemination, and is a marker of response to adjuvant chemotherapy in pancreatic cancer. *Gut* 66, 1665 (2017).
97. Haumaitre, C. *et al.* Severe pancreas hypoplasia and multicystic renal dysplasia in two human fetuses carrying novel HNF1beta/MODY5 mutations. *Hum Mol Genet* 15, 2363–2375 (2006).
98. Quilichini, E. *et al.* Pancreatic Ductal Deletion of Hnf1b Disrupts Exocrine Homeostasis, Leads to Pancreatitis, and Facilitates Tumorigenesis. *Cell Mol Gastroenterol Hepatol* 8, 487 (2019).
99. Mastri, M. *et al.* Secreted Frizzled-related protein 2 as a target in antifibrotic therapeutic intervention. *Am J Physiol Cell Physiol* 306, (2014).
100. van Loon, K., Huijbers, E. J. M. & Griffioen, A. W. Secreted frizzled-related protein 2: a key player in noncanonical Wnt signaling and tumor angiogenesis. *Cancer Metastasis Rev* 40, 191 (2021).
101. Domogatskaya, A., Rodin, S. & Tryggvason, K. Functional diversity of laminins. *Annu Rev Cell Dev Biol* 28, 523–553 (2012).
102. Wang, Q. *et al.* Laminin γ-2 (LAMC2) promotes the proliferation and invasion of pancreatic cancer cells by regulating PI3K/Akt signaling. *J Pancreatol* (2024) doi:10.1097/JP9.0000000000000166.
103. Bracken, C. P. *et al.* Genome-wide identification of miR-200 targets reveals a regulatory network controlling cell invasion. *EMBO J* 33, 2040–2056 (2014).
104. Wellner, U., Brabletz, T. & Keck, T. ZEB1 in Pancreatic Cancer. *Cancers (Basel)* 2, 1617 (2010).

105. Russel, S., Frand, A. R. & Ruvkun, G. Regulation of the C. elegans molt by pqn-47. *Dev Biol* 360, 297 (2011).
106. Qi, H. *et al.* De novo variants in congenital diaphragmatic hernia identify MYRF as a new syndrome and reveal genetic overlaps with other developmental disorders. *PLoS Genet* 14, (2018).
107. Emery, B. *et al.* Myelin gene regulatory factor is a critical transcriptional regulator required for CNS myelination. *Cell* 138, 172–185 (2009).
108. Pierce, M. *et al.* Sum1 and Ndt80 Proteins Compete for Binding to Middle Sporulation Element Sequences That Control Meiotic Gene Expression. *Mol Cell Biol* 23, 4814 (2003).
109. Bujalka, H. *et al.* MYRF Is a Membrane-Associated Transcription Factor That Autoproteolytically Cleaves to Directly Activate Myelin Genes. *PLoS Biol* 11, e1001625 (2013).
110. Choi, J. ok *et al.* Elucidating the transactivation domain of the pleiotropic transcription factor Myrf. *Scientific Reports 2018 8:1* 8, 1–9 (2018).
111. Schwarzer, D., Stummeyer, K., Gerardy-Schahn, R. & Mühlenhoff, M. Characterization of a novel intramolecular chaperone domain conserved in endosialidases and other bacteriophage tail spike and fiber proteins. *J Biol Chem* 282, 2821–2831 (2007).
112. Li, Z., Park, Y. & Marcotte, E. M. A Bacteriophage Tailspike Domain Promotes Self-Cleavage of a Human Membrane-Bound Transcription Factor, the Myelin Regulatory Factor MYRF. *PLoS Biol* 11, (2013).
113. Russel, S., Frand, A. R. & Ruvkun, G. Regulation of the C. elegans molt by pqn-47. *Dev Biol* 360, 297–309 (2011).
114. Milan, M. *et al.* Pancreatic Cancer Cells Require the Transcription Factor MYRF to Maintain ER Homeostasis. *Dev Cell* 55, 398-412.e7 (2020).
115. Kong, B. *et al.* Hypoxia-induced endoplasmic reticulum stress characterizes a necrotic phenotype of pancreatic cancer. *Oncotarget* 6, 32154–32160 (2015).
116. Pommier, A. *et al.* Unresolved endoplasmic reticulum stress engenders immune-resistant, latent pancreatic cancer metastases. *Science (1979)* 360, (2018).
117. Porter, K. R., Claude, A. & Fullam, E. F. A STUDY OF TISSUE CULTURE CELLS BY ELECTRON MICROSCOPY : METHODS AND PRELIMINARY OBSERVATIONS. *J Exp Med* 81, 233–246 (1945).
118. Almanza, A. *et al.* Endoplasmic reticulum stress signalling - from basic mechanisms to clinical applications. *FEBS J* 286, 241–278 (2019).
119. Wang, M. & Kaufman, R. J. The impact of the endoplasmic reticulum protein-folding environment on cancer development. *Nature Reviews Cancer* vol. 14 581–597 Preprint at https://doi.org/10.1038/nrc3800 (2014).
120. Walter, P. & Ron, D. The unfolded protein response: From stress pathway to homeostatic regulation. *Science* vol. 334 1081–1086 Preprint at https://doi.org/10.1126/science.1209038 (2011).
121. Hetz, C. The unfolded protein response: Controlling cell fate decisions under ER stress and beyond. *Nature Reviews Molecular Cell Biology* vol. 13 89–102 Preprint at https://doi.org/10.1038/nrm3270 (2012).
122. Kim, H., Bhattacharya, A. & Qi, L. Endoplasmic reticulum quality control in cancer: Friend or foe. *Semin Cancer Biol* 33, 25–33 (2015).
123. Ron, D. & Walter, P. Signal integration in the endoplasmic reticulum unfolded protein response. *Nat Rev Mol Cell Biol* 8, 519–529 (2007).
124. Clarke, H. J., Chambers, J. E., Liniker, E. & Marciniak, S. J. Endoplasmic reticulum stress in malignancy. *Cancer Cell* 25, 563–573 (2014).
125. Robinson, C. M. *et al.* An Emerging Role for the Unfolded Protein Response in Pancreatic Cancer. *Cancers (Basel)* 13, 1–16 (2021).

126. Bovolenta, P., Esteve, P., Ruiz, J. M., Cisneros, E. & Lopez-Rios, J. Beyond Wnt inhibition: new functions of secreted Frizzled-related proteins in development and disease. *J Cell Sci* 121, 737–746 (2008).
127. Cruciat, C. M. & Niehrs, C. Secreted and Transmembrane Wnt Inhibitors and Activators. *Cold Spring Harb Perspect Biol* 5, (2013).
128. Kaur, A. *et al.* sFRP2 in the aged microenvironment drives melanoma metastasis and therapy resistance HHS Public Access. *Nature* 532, 250–254 (2016).
129. Vincent, K. M. & Postovit, L. M. A pan-cancer analysis of secreted Frizzled-related proteins: re-examining their proposed tumour suppressive function. *Scientific Reports 2017 7:1* 7, 1–9 (2017).
130. Satoh, W., Gotoh, T., Tsunematsu, Y., Aizawa, S. & Shimon, A. Sfrp1 and Sfrp2 regulate anteroposterior axis elongation and somite segmentation during mouse embryogenesis. *Development* 133, 989–999 (2006).
131. Chong, J. M., Üren, A., Rubin, J. S. & Speicher, D. W. Disulfide bond assignments of secreted Frizzled-related protein-1 provide insights about Frizzled homology and netrin modules. *J Biol Chem* 277, 5134–5144 (2002).
132. Finch, P. W. *et al.* Purification and molecular cloning of a secreted, Frizzled-related antagonist of Wnt action. *Proc Natl Acad Sci U S A* 94, 6770–6775 (1997).
133. Üren, A. *et al.* Secreted frizzled-related protein-1 binds directly to Wingless and is a biphasic modulator of Wnt signaling. *J Biol Chem* 275, 4374–4382 (2000).
134. Bányai, L. & Patthy, L. The NTR module: domains of netrins, secreted frizzled related proteins, and type I procollagen C-proteinase enhancer protein are homologous with tissue inhibitors of metalloproteases. *Protein Sci* 8, 1636–1642 (1999).
135. Jung, Y. S. & Park, J. Il. Wnt signaling in cancer: therapeutic targeting of Wnt signaling beyond β-catenin and the destruction complex. *Exp Mol Med* 52, 183–191 (2020).
136. van Loon, K., Huijbers, E. J. M. & Griffioen, A. W. Secreted frizzled-related protein 2: a key player in noncanonical Wnt signaling and tumor angiogenesis. *Cancer Metastasis Rev* 40, 191 (2021).
137. Giles, R. H., Van Es, J. H. & Clevers, H. Caught up in a Wnt storm: Wnt signaling in cancer. *Biochim Biophys Acta Rev Cancer* 1653, 1–24 (2003).
138. Lee, A. Y. *et al.* Expression of the secreted frizzled-related protein gene family is downregulated in human mesothelioma. *Oncogene* 23, 6672–6676 (2004).
139. Vincent, K. M. & Postovit, L. M. A pan-cancer analysis of secreted Frizzled-related proteins: re-examining their proposed tumour suppressive function. *Scientific Reports 2017 7:1* 7, 1–9 (2017).
140. Kaur, A. *et al.* sFRP2 in the aged microenvironment drives melanoma metastasis and therapy resistance HHS Public Access. *Nature* 532, 250–254 (2016).
141. Karamanos, N. K. *et al.* A guide to the composition and functions of the extracellular matrix. *FEBS J* 288, 6850–6912 (2021).
142. CHAFFEY, N. Alberts, B., Johnson, A., Lewis, J., Raff, M., Roberts, K. and Walter, P. Molecular biology of the cell. 4th edn. *Ann Bot* 91, 401 (2003).
143. Kular, J. K., Basu, S. & Sharma, R. I. The extracellular matrix: Structure, composition, age-related differences, tools for analysis and applications for tissue engineering. *J Tissue Eng* 5, (2014).
144. Kadler, K. E., Hill, A. & Canty-Laird, E. G. Collagen fibrillogenesis: fibronectin, integrins, and minor collagens as organizers and nucleators. *Curr Opin Cell Biol* 20, 495 (2008).
145. Hynes, R. O. Extracellular matrix: not just pretty fibrils. *Science* 326, 1216 (2009).

146. Winkler, J., Abisoye-Ogunniyan, A., Metcalf, K. J. & Werb, Z. Concepts of extracellular matrix remodelling in tumour progression and metastasis. *Nature Communications 2020 11:1* 11, 1–19 (2020).
147. Egeblad, M., Rasch, M. G. & Weaver, V. M. Dynamic interplay between the collagen scaffold and tumor evolution. *Curr Opin Cell Biol* 22, 697–706 (2010).
148. Winkler, J., Abisoye-Ogunniyan, A., Metcalf, K. J. & Werb, Z. Concepts of extracellular matrix remodelling in tumour progression and metastasis. *Nature Communications 2020 11:1* 11, 1–19 (2020).
149. Cox, T. R. & Erler, J. T. Remodeling and homeostasis of the extracellular matrix: implications for fibrotic diseases and cancer. *Dis Model Mech* 4, 165–178 (2011).
150. Chang, J. & Chaudhuri, O. Beyond proteases: Basement membrane mechanics and cancer invasion. (2019) doi:10.1083/jcb.201903066.
151. Lin, Y. C. *et al.* Mechanosensing of substrate thickness. *Phys Rev E Stat Nonlin Soft Matter Phys* 82, (2010).
152. Fang, M., Yuan, J., Peng, C. & Li, Y. Collagen as a double-edged sword in tumor progression. *Tumour Biol* 35, 2871–2882 (2014).
153. Tschumperlin, D. J., Ligresti, G., Hilscher, M. B. & Shah, V. H. Mechanosensing and fibrosis. *J Clin Invest* 128, 74 (2018).
154. Whatcott, C. J. *et al.* Desmoplasia in Primary Tumors and Metastatic Lesions of Pancreatic Cancer. *Clin Cancer Res* 21, 3561–3568 (2015).
155. Kim, M. P. *et al.* Generation of orthotopic and heterotopic human pancreatic cancer xenografts in immunodeficient mice. *Nat Protoc* 4, 1670–1680 (2009).
156. Diaferia, G. R., Cirulli, V. & Biunno, I. SEL1L Regulates Adhesion, Proliferation and Secretion of Insulin by Affecting Integrin Signaling. *PLoS One* 8, e79458 (2013).
157. Branon, T. C. *et al.* Efficient proximity labeling in living cells and organisms with TurboID. *Nature Biotechnology 2018 36:9* 36, 880–887 (2018).
158. Kulak, N. A., Pichler, G., Paron, I., Nagaraj, N. & Mann, M. Minimal, encapsulated proteomic-sample processing applied to copy-number estimation in eukaryotic cells. *Nature Methods 2014 11:3* 11, 319–324 (2014).
159. Das, J., Prateek Bhatia, • & Singh, A. CRISP Points on Establishing CRISPR-Cas9 In Vitro Culture Experiments in a Resource Constraint Haematology Oncology Research Lab. doi:10.1007/s12288-018-1008-z.
160. Cong, L. *et al.* Multiplex Genome Engineering Using CRISPR/Cas Systems HHS Public Access. *Science (1979)* 339, 819–823 (2013).
161. Picelli, S. *et al.* Full-length RNA-seq from single cells using Smart-seq2. *Nature Protocols 2013 9:1* 9, 171–181 (2014).
162. Picelli, S. *et al.* Smart-seq2 for sensitive full-length transcriptome profiling in single cells. *Nat Methods* 10, 1096–1100 (2013).
163. Tyanova, S., Temu, T. & Cox, J. The MaxQuant computational platform for mass spectrometry-based shotgun proteomics. *Nature Protocols 2016 11:12* 11, 2301–2319 (2016).
164. Tyanova, S. *et al.* The Perseus computational platform for comprehensive analysis of (prote)omics data. *Nature Methods 2016 13:9* 13, 731–740 (2016).
165. Cox, J. *et al.* Accurate Proteome-wide Label-free Quantification by Delayed Normalization and Maximal Peptide Ratio Extraction, Termed MaxLFQ. *Mol Cell Proteomics* 13, 2513 (2014).
166. Stringer, C., Wang, T., Michaelos, M. & Pachitariu, M. Cellpose: a generalist algorithm for cellular segmentation. *Nature Methods 2020 18:1* 18, 100–106 (2020).
167. Schindelin, J. *et al.* Fiji: an open-source platform for biological-image analysis. *Nat Methods* 9, 676–682 (2012).
168. Kim, D. *et al.* TopHat2: Accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome Biol* 14, 1–13 (2013).

169. Gualdrini, F. *et al.* SRF Co-factors Control the Balance between Cell Proliferation and Contractility. *Mol Cell* 64, 1048–1061 (2016).
170. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* 15, (2014).
171. Cordenonsi, M. *et al.* The Hippo transducer TAZ confers cancer stem cell-related traits on breast cancer cells. *Cell* 147, 759–772 (2011).
172. Haze, K., Yoshida, H., Yanagi, H., Yura, T. & Mori, K. Mammalian Transcription Factor ATF6 Is Synthesized as a Transmembrane Protein and Activated by Proteolysis in Response to Endoplasmic Reticulum Stress. *Mol Biol Cell* 10, 3787 (1999).
173. Sakai, J. *et al.* Sterol-regulated release of SREBP-2 from cell membranes requires two sequential cleavages, one within a transmembrane segment. *Cell* 85, 1037–1046 (1996).
174. Avila, J. R., Lee, J. S. & Toriia, K. U. Co-Immunoprecipitation of Membrane-Bound Receptors. *The Arabidopsis Book / American Society of Plant Biologists* 13, e0180 (2015).
175. Avila, J. R., Lee, J. S. & Toriia, K. U. Co-Immunoprecipitation of Membrane-Bound Receptors. *The Arabidopsis Book / American Society of Plant Biologists* 13, e0180 (2015).
176. Kim, D. I. & Roux, K. J. Filling the void: Proximity-based labeling of proteins in living cells. *Trends Cell Biol* 26, 804 (2016).
177. Kim, D. I. & Roux, K. J. Filling the Void: Proximity-Based Labeling of Proteins in Living Cells. *Trends in Cell Biology* vol. 26 804–817 Preprint at https://doi.org/10.1016/j.tcb.2016.09.004 (2016).
178. Choi-Rhee, E., Schulman, H. & Cronan, J. E. Promiscuous protein biotinylation by Escherichia coli biotin protein ligase. *Protein Science* 13, 3043–3050 (2004).
179. Roux, K. J., Kim, D. I., Raida, M. & Burke, B. A promiscuous biotin ligase fusion protein identifies proximal and interacting proteins in mammalian cells. *J Cell Biol* 196, 801 (2012).
180. May, D. G., Scott, K. L., Campos, A. R. & Roux, K. J. Comparative Application of BioID and TurboID for Protein-Proximity Biotinylation. *Cells* 9, (2020).
181. Helenius, A. How N-linked oligosaccharides affect glycoprotein folding in the endoplasmic reticulum. *Mol Biol Cell* 5, 253–265 (1994).
182. Sousa, M. & Armando, P. J. The molecular basis for the recognition of misfolded glycoproteins by the UDP-Glc:glycoprotein glucosyltransferase. *EMBO Journal* 14, 4196–4203 (1995).
183. Puram, S. V *et al.* Single-Cell Transcriptomic Analysis of Primary and Metastatic Tumor Ecosystems in Head and Neck Cancer. *Cell* 171, 1611.e1-1611.e24 (2017).
184. Laklai, H. *et al.* Genotype tunes pancreatic ductal adenocarcinoma tissue tension to induce matricellular fibrosis and tumor progression. *Nat Med* 22, 497–505 (2016).
185. Campagnola, P. J. *et al.* Three-Dimensional High-Resolution Second-Harmonic Generation Imaging of Endogenous Structural Proteins in Biological Tissues. (2002).
186. Campagnola, P. Second Harmonic Generation Imaging Microscopy: Applications to Diseases Diagnostics. *Anal Chem* 83, 3224 (2011).
187. Guo, J. *et al.* The development of proximity labeling technology and its applications in mammals, plants, and microorganisms. *Cell Commun Signal* 21, 269 (2023).
188. Qin, W., Cho, K. F., Cavanagh, P. E. & Ting, A. Y. Deciphering molecular interactions by proximity labeling. *Nat Methods* 18, 133–143 (2021).
189. Branon, T. C. *et al.* Efficient proximity labeling in living cells and organisms with TurboID. *Nature Biotechnology 2018 36:9* 36, 880–887 (2018).

190. Helenius, A. & Aebi, M. Roles of N-Linked Glycans in the Endoplasmic Reticulum. *https://doi.org/10.1146/annurev.biochem.73.011303.073752* 73, 1019–1049 (2004).
191. Hebert, D. N. & Molinari, M. Flagging and docking: dual roles for N-glycans in protein quality control and cellular proteostasis. *Trends Biochem Sci* 37, 404 (2012).
192. Ferris, S. P., Jaber, N. S., Molinari, M., Arvan, P. & Kaufman, R. J. UDP-glucose:glycoprotein glucosyltransferase (UGGT1) promotes substrate solubility in the endoplasmic reticulum. *Mol Biol Cell* 24, 2597 (2013).
193. Adams, B. M., Canniff, N. P., Guay, K. P., Larsen, I. S. B. & Hebert, D. N. Quantitative glycoproteomics reveals cellular substrate selectivity of the er protein quality control sensors uggt1 and uggt2. *Elife* 9, 1–27 (2020).
194. Huang, C. *et al.* Secreted Frizzled-Related Protein 2 Is Associated with Disease Progression and Poor Prognosis in Breast Cancer. *Dis Markers* 2019, (2019).
195. Liu, Y. *et al.* Secreted frizzled-related protein 2-mediated cancer events: Friend or foe? *Pharmacol Rep* 69, 403–408 (2017).
196. Yamamura, S. *et al.* Oncogenic functions of secreted Frizzled-related protein 2 in human renal cancer. *Mol Cancer Ther* 9, 1680–1687 (2010).
197. Expression of SFRP2 in pancreatic cancer - The Human Protein Atlas. https://www.proteinatlas.org/ENSG00000145423-SFRP2/pathology/pancreatic+cancer.
198. Tian, C. *et al.* Cancer Cell-Derived Matrisome Proteins Promote Metastasis in Pancreatic Ductal Adenocarcinoma. *Cancer Res* 80, 1461–1474 (2020).
199. Swietlik, J. J. *et al.* Cell-selective proteomics segregates pancreatic cancer subtypes by extracellular proteins in tumors and circulation. *Nature Communications 2023 14:1* 14, 1–17 (2023).
200. Han, W. *et al.* Oriented collagen fibers direct tumor cell intravasation. *Proc Natl Acad Sci U S A* 113, 11208–11213 (2016).
201. Oudin, M. J. *et al.* Tumor Cell-Driven Extracellular Matrix Remodeling Drives Haptotaxis during Metastatic Progression. *Cancer Discov* 6, 516–531 (2016).