Dynamics of the microbiome in pancreatic cancer from inception to invasion

DISSERTATION

for the award of the degree

Doctor rerum naturalium (Dr. rer. nat.)

of the Georg-August-Universität Göttingen

within the doctoral program Molecular Medicine

of the Georg-August-University School of Science (GAUSS)

submitted by

Nina Pfisterer

born in Stuttgart, Germany

Göttingen, 2023

Thesis Advisory Committee

Prof. Dr. Dr. Albrecht Neeße
Department of Gastroenterology, Gastrointestinal Oncology and Endocrinology, University Medical Center Göttingen

Prof. Dr. Matthias Dobbelstein
Department of Molecular Oncology, University Medical Center Göttingen

Prof. Dr. Argyris Papantonis
Institute of Pathology, University Medical Center Göttingen

Members of the Examination Board

1. Referee Prof. Dr. Dr. Albrecht Neeße
Department of Gastroenterology, Gastrointestinal Oncology and Endocrinology
University Medical Center Göttingen

2. Referee Prof. Dr. Matthias Dobbelstein
Department of Molecular Oncology
University Medical Center Göttingen

Further Members of the Examination Board

Prof. Dr. Argyris Papantonis
Institute of Pathology, University Medical Center Göttingen

Prof. Dr. Michael Zeisberg
Department of Nephrology and Rheumatology, University Medical Center Göttingen

Prof. Dr. mult. Thomas Meyer
Department of Psychosomatic Medicine and Psychotherapy, University Medical Center Göttingen

Prof. Dr. Ralf Dressel
Institute of Cellular and Molecular Immunology, University Medical Center Göttingen

Date of Oral Examination

September 11, 2023

THE PHD MOVIE 2

by Jorge Cham

SCENE

Staring Cecelia in her room at her desk on the phone with her mom.

CECELIA [optimistic]

I have some good news.

MOM [enthusiastic]

You’re pregnant!

CECELIA [defensive]

NO! No.

I’m finally writing my thesis.

MOM [disappointed]

Ouh.

Contents

ABBREVIATIONS

	× g	times gravity
	°C	degree Celsius
	μL	microliter
	μm	micrometer
	12w	twelve weeks
	2D	two-dimensional
	3D	three-dimensional
	5-FU	5-fluorouracil
	6w	six weeks
A, a |	AB	avidin-biotin
	abv	archaea, bacteria, viruses
	ADEX	aberrantly differentiated endocrine exocrine
	ADM	acinar-to-ductal metaplasia
	AS	alignment score
	AUC	area under the curve
	AUPR	area under precision-recall curve
B, b |	BC	Bray-Curtis dissimilarity
	BET	bromodomain and extra-terminal motif
	bEV	bacterial extracellular vesicle
	BL	basal-like
	BRCA2	breast cancer gene 2
	BSA	bovine serum albumin
C, c |	C57BL/6	mouse number 57 black substrain 6
	CAF	cancer-associated fibroblast

	CDKN2A	cyclin-dependent kinase inhibitor 2A gene
	cJUN	cellular transcription factor Jun
	CK5	cytokeratin 5
	CLA	classical
	Cov	alignment coverage
	CRC	colorectal carcinoma
	Cre	'causes recombination' recombinase
	CTRL	control
D, d |	DAB	diaminobenzidine
	DAPI	4′,6-diamidino-2-phenyl-indol
	DMEM	Dulbecco's Modified Eagle Medium
	DNA	desoxyribonucleic acid
	dsDNA	double-stranded desoxyribonucleic acid
E, e |	EB	elution buffer
	ECM	extracellular matrix
	EDTA	ethylenediaminetetraacetic acid
	EP	end point
	EtOH	ethanol
F, f |	FC	flow cell
	FCS	fetal calf serum
	FFPE	formalin-fixed paraffin-embedded
	FISH	fluorescence-*in situ*-hybridization
	fmol	femtomole
	FMT	fecal microbiome transplantation
	FOLFIRINOX	fluorouracil, irinotecan, leucovorin, oxaliplatin
	Foxn1	forkhead box protein N1 gene

G, g |	GATA6	GATA binding protein 6
	GEM	gemcitabine
	GEMM	genetically engineered mouse model
H, h |	h	hour
	H&E	hematoxylin and eosin
	H_2O	water
	H_2O_2	hydrogen peroxide
	HCl	hydrochloride
	HRP	horse radish peroxidase
	HS	high sensitivity
	HYB	hybrid
I, i |	IARC	International Agency for Research on Cancer
	ICI	immune checkpoint inhibitor
	IHC	immunohistochemistry
	IHMS	International Human Microbiome Standards
	InvS	inverse Simpson index
K, k |	kb	kilobases
	KC	*LSL-Kras$^{G12D/+}$;P48-Cre*
	KPC	*LSL-Kras$^{G12D/+}$;LSL-Trp53$^{R172H/+}$;Pdx-1-Cre*
	KRAS/*Kras*	Kirsten rat sarcoma viral oncogene homolog
L, l |	LAVES	Landesamt für Verbraucherschutz und Lebensmittelsicherheit
	LDA	linear discriminant analysis
	LEfSe	linear discriminant analysis effect size
	LPS	lipopolysaccharide
	LSL	Lox-Stop-Lox
	LTS	long-term survivor

M, m	M	molar
	Mb	megabases
	MCP1	monocyte chemoattractant protein 1
	MeTaPONT	Metagenomic Taxonomy Pipeline for ONT
	mg	milligram
	min	minute
	mL	milliliter
	mM	millimolar
N, n	nab-PAC	nano-formulated albumin-bound paclitaxel
	NaCl	sodium chloride
	NCBI	National Center for Biotechnology Information
	NET	neuroendocrine tumor
	ng	nanogram
	NGS	next generation sequencing
	nm	nanometer
	NMDS	non-metric multidimensional scaling
	NMRI	Naval Medical Research Institute
	NTC	negative control
	nu	nude
O, o	ONT	Oxford Nanopore Technologies
	OS	observed species
	OSCC	oral squamous cell carcinoma
P, p	P/S	penicillin/streptomycin
	P48/Ptf1	p48 DNA-binding subunit of pancreas transcription factor 1
	PanIN	pancreatic intraepithelial neoplasia
	PBS	phosphate-buffered saline
	PC	pancreatic cancer

	PCoA	principal coordinates analysis
	PCR	polymerase chain reaction
	PDAC	pancreatic ductal adenocarcinoma
	Pdx-1/Ipf1	pancreatic and duodenal homeobox 1/insulin promoter factor 1
	PERMANOVA	permutational multivariate analysis of variance
	pH	pondus hydrogenii
	P-MAPS	Pancreatitis – Microbiome As Predictor of Severity
	PP	pancreatic polypeptide
Q, q |	QM	quasimesenchymal
	QS	quality score
R, r |	RAP	rapid adapter
	RefSeq	Reference Sequence
	RNA	ribonucleic acid
	ROC	receiver operating characteristic
	rpm	rounds per minute
	RPMI	Roswell Park Memorial Institute
	rRNA	ribosomal ribonucleic acid
	RT	room temperature
S, s |	s	second
	S	Svedberg unit
	SCFA	short-chain fatty acid
	SFB	short fragment buffer
	SI	Shannon index
	SIBO	small intestine bacterial overgrowth
	SOP	standard operating procedure
	spp.	species
	STS	short-term survivor

T, t |	taxID	taxonomic identification number
	TBS	Tris-buffered saline
	TBS-T	Tris-buffered saline with Tween 20
	TE	Tris-EDTA
	TME	tumor microenvironment
	TNFα	tumor necrosis factor α
	TP53/*Trp53*	tumor protein 53 gene
	Tris	tris(hydroxymethyl)aminomethane
U, u |	UMG	University Medical Center Göttingen
	USA	United States of America
	uwUF	unweighted UniFrac distance
V, v |	V	variable region
	VC	vehicle control
W, w |	wUF	weighted UniFrac distance
Z, z |	ZTE	Zentrale Tierexperimentelle Einrichtung

LIST OF FIGURES

LIST OF TABLES

ABSTRACT

Recently, the microbiome was associated with hallmark features of pancreatic ductal adenocarcinoma (PDAC). This dismal human disease with one of the lowest survival rates of below 10% is characterized by clinical challenges such as unspecific symptoms, lack of biomarkers, late diagnosis, refractoriness against systemic therapies, and a high recurrence rate after surgery. The biology behind these difficulties has been approached in multiple ways, for instance by tackling the pronounced desmoplastic stroma or investigating chemotherapy-inactivating enzymes. To date, no significant progress with clinical usage was made. After being long known for its impact on a plethora of physiological and pathophysiological processes, studies of considerable impact have lately demonstrated the microbiome to be involved in all kinds of aspects of various cancer types. Moreover, the discovery of tumors and healthy organs to harbor differential microbiota was groundbreaking. These findings were also and most notably shown in PDAC. Being one of the deadliest cancers, PDAC would benefit immensely from systematic, reliable microbiome research. Being convinced that third generation sequencing is the most effective technology for precise and reproducible microbiome studies, a sequencing platform by Oxford Nanopore Technologies (ONT) together with a comprehensive workflow for DNA extraction, library preparation and bioinformatic analysis was established as a major project within the scope of this thesis. This comprised testing seven different DNA isolation protocols and benchmarking several other parameters. Further, the pipeline was optimized for murine high (feces) and low microbial biomass (pancreatic) samples. Ensuing successful establishment of ONT sequencing, the presence of bacterial material in murine KPC tumor and human PDAC tissue was shown via IHC targeting lipopolysaccharide and FISH targeting the 16s rRNA gene. The occurrence of both components was visualized and, for KPC tumors, also quantified demonstrating a significantly higher abundance in diseased tissue as compared to healthy controls. The characterization of the fecal microbiome at various time points during disease progression as well as the end stage tumor microbiome in the KPC mouse model via ONT sequencing impressively showed significant differences in beta diversity metrics compared to their healthy counterparts. Furthermore, the similarity with recently published human stool and PDAC microbiomes could be ascertained via regression models. Finally, first experiments involving FFPE tissue material as most

commonly available sample source in the clinical setting were performed aiming at evaluating the impact of PDAC subtypes on their intratumoral microbiomes. Therefore, samples from tumors derived from orthotopically transplanted human PDAC cell lines of the classical and basal-like subtype, as well as xenografted classical tumors treated with TNFα causing lineage reprogramming, and resected human PDAC which was subtyped via IHC were employed. In the course of these experiments, contamination issues were encountered and are explicitly discussed in the following. The results were not as conclusive as the KPC data, however, first evidence for the PDAC subtype to impinge on the microbiome is given. Taken together, new sequencing technology by ONT was installed in a customized way, distinct local microbiomes of the KPC model were characterized and found to be similar to the situation in humans, and PDAC subtypes seem to influence the microbiome.

1 Introduction

1.1 Pancreas physiology

1.1.1 Macro- and microscopic structures of the pancreas

Gastroenterology, the study of the digestive tract and associated diseases, well-defines the pancreas as a solitary organ, specifically one of the accessory organs of digestion which also include gallbladder and liver. It is a heterocrine gland characterized by both endocrine and exocrine function (**1.1.2**). Anatomically, it is situated in the abdomen beneath and partially behind the stomach and stretches horizontally in the medial plane between duodenum and spleen (Dolenšek et al., 2015). In humans, the pancreas may roughly be subdivided into head, body, and tail, however, these parts are not precisely demarcated (Bockman, 1993) (**figure 1**).

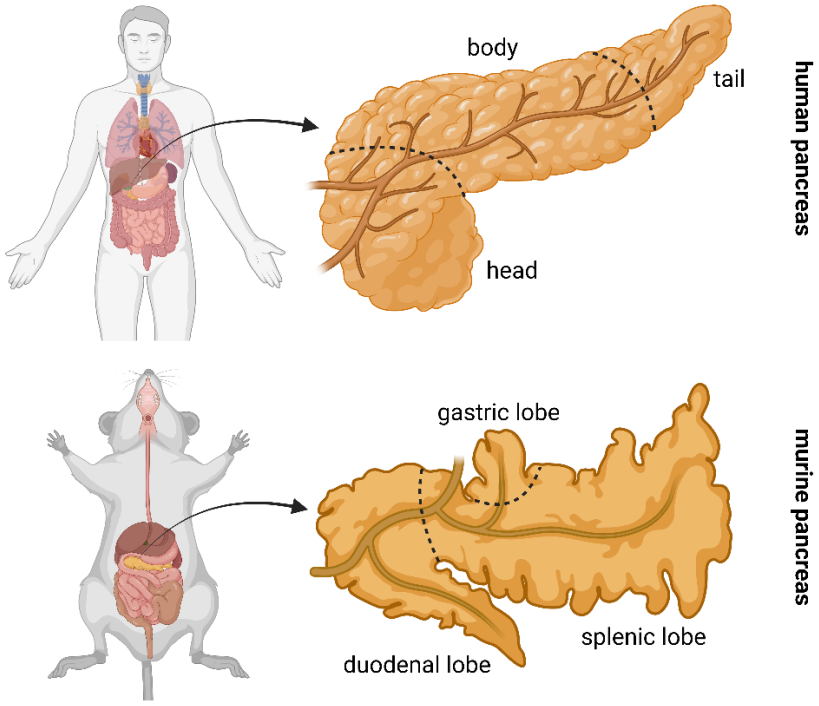

Figure 1: Illustration of the human and murine pancreas. While they are both situated partially behind the stomach between duodenum and spleen, they present with partially different anatomy. Although the murine pancreas is rather dispersed, both organs are loosely divided in three parts. The human pancreas is subdivided into head, body, and tail; the murine pancreas in duodenal lobe, gastric lobe, and splenic lobe. Both are heavily lobulated and intercalated with a plethora of ducts. Own illustration created with (biorender.com).

It weighs about 90 g on average in adults (Caglar et al., 2014; Innes and Carey, 1994). The pancreas is irregularly lobulated with each lobe consisting of smaller lobules. These glandular lobules are each provided with a duct uniting in the main duct that runs through the entire pancreas opening into the duodenum. Additionally, the common bile duct joins the pancreatic main duct (Standring, 2016). These ductal structures are relevant access points to the pancreatic tissue which are of great importance regarding the matter of this thesis.

Both endocrine and exocrine parts of the pancreas are of great clinical and scientific interest. Their structures have been studied and characterized intensely over a long period of time. In his thesis of 1869, Paul Langerhans described the Islets of Langerhans which form the endocrine tissue of the pancreas (Langerhans, 1869). They vary in size from a few to a couple of thousands of endocrine cells that form irregular roundish clusters in the sea of exocrine cells from where also the term islets arises. Endocrine cells comprise beta cells, alpha cells, delta cells, pancreatic polypeptide (PP) cells, and epsilon cells in order of decreasing abundance. The endocrine part makes up for only 1% of the pancreatic tissue. In terms of severely malignant diseases of the pancreas and, by association, in the context of this thesis, the more relevant exocrine tissue accounts for the vast majority of the pancreas accordingly, about 99%. As illustrated in **figure 2**, the lobules are structured in so-called acini, clusters of pyramidal acinar cells arranged in domes to secrete zymogens into the intercalating ducts which in turn drain into the intra-, then interlobular ducts. Acinar cells are characterized by a large endoplasmic reticulum and their nuclei are oriented towards the basement membrane (Young et al., 2014).

As for most other diseases, pancreas research still strongly relies on animal experiments. Observations and findings are then translated to humans what makes the understanding of similarities, but in particular, the characterization of differences absolutely crucial. In pancreas research, mice are the most often and best studied animal model and, to a great extent, today's knowledge about the human pancreas is based on murine studies (Dolenšek et al., 2015). Paradoxically, the murine pancreas is less clearly defined, although it has been studied even in living mice since decades ago (Covell, 1928; Flory and Thal, 1947). This is mainly due to its rather diffuse appearance. However, analogously to the human organ, attempts have been undertaken to define three major parts, namely the duodenal, splenic, and gastric lobe (Liu et al., 2010) (**figure 1**). These lobes are usually interspaced by adipose

and connective tissue. Regarding endocrine cells, there are more beta cells and less alpha cells in murine Islets of Langerhans as compared to humans (Dolenšek et al., 2015). The ductal organization also differs between humans and mice. Still, despite these structural differences, mice are genetically very similar to humans with many gene homologues as well as many pathophysiological and functional aspects being highly similar in both organisms rendering the mouse a suitable model regarding many applications, particularly within the scope of this thesis.

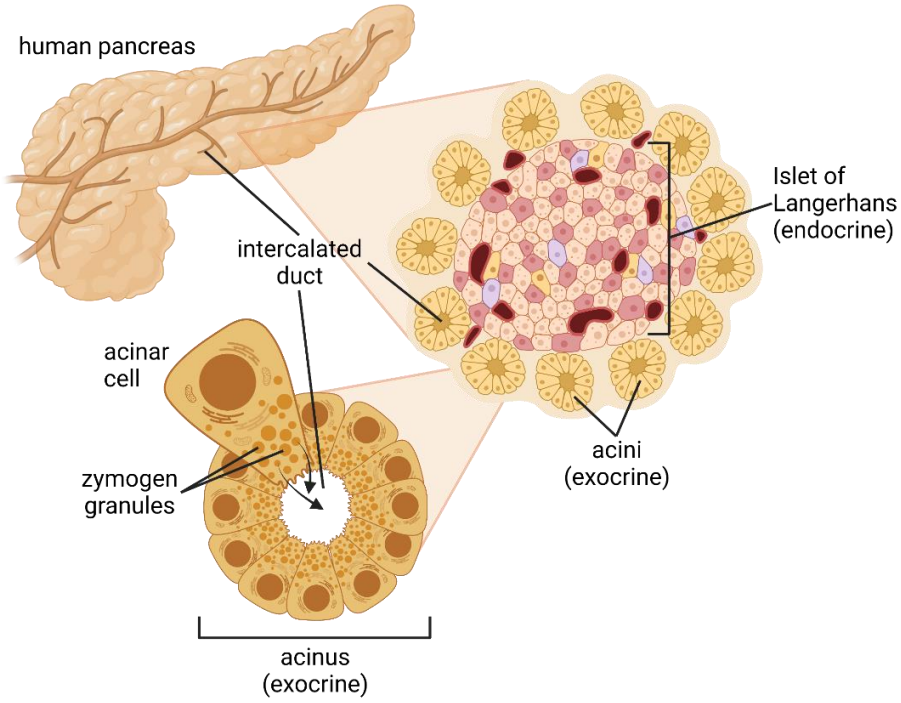

Figure 2: Illustration of human exocrine pancreas structures. Both endocrine (Islets of Langerhans) and exocrine (acini) pancreas compartments coexist. Although single acini are usually smaller than Islets of Langerhans, in total they make up the largest part of the pancreatic tissue. The exocrine pancreas also contains ductal structures which intercalate the acini. Acinar cells produce zymogens, precursors of enzymes, and secrete them in digestive fluid into the ducts. In human adults, the exocrine pancreas is the organ with the highest protein synthesis levels wherefore acinar cells are packed with corresponding cell organelles such as endoplasmic reticula. Own illustration created with (biorender.com).

1.1.2 Dual function of the pancreas

Besides the gonads, the pancreas is the only organ that executes simultaneous endocrine and exocrine functions. Hence, it is also referred to as the organ of the endocrine system and the digestive system with the endocrine Islets of Langerhans producing essential hormones

controlling the blood sugar level and the exocrine pancreas serving as a source of digestive enzymes (Zhou and Melton, 2018). The blood glucose regulation is achieved by beta cells secreting insulin if the blood glucose level is too high, and by their counterparts, the alpha cells, which secrete glucagon to increase the level (Baum et al., 1962; Lacy, 1959). Somatostatin inhibits the release of both insulin and glucagon and is synthesized by delta cells (Young et al., 2014). PP cells secrete pancreatic polypeptide which generally regulates pancreatic secretion in response to food intake (Batterham et al., 2003). Finally, epsilon cells release ghrelin, a growth hormone-releasing peptide increasing with the sensation of hunger (Müller et al., 2015). Taken together, pancreatic islet cells are responsible for hormone production and secretion contributing to homeostasis of glucose levels in the wake of food uptake. The function of the exocrine cells is also vital. Acinar cells contain granules that hold the pre-form of digestive enzymes, so-called zymogens. Proenzymes prevent autodigestion of the pancreas. Only when the digestive fluids secreted by the acini ultimately reaches the duodenum which is the first part of the small intestine receiving food from the stomach, the proenzymes are activated via cleavage. Enzymes such as trypsin, lipase and amylase then help to break down proteins, lipids and carbohydrates contained in the food. Although structurally very different, Henderson and colleagues have already suggested the functional relationship between endocrine and exocrine tissue as well as the hormonal effects of the endocrine over the exocrine pancreas in the early 80s (Henderson et al., 1981). For instance, insulin had been found to increase the synthesis on amylase in fed animals (Söling and Unger, 1972).

1.2 Pancreatic ductal adenocarcinoma (PDAC)

1.2.1 Description and dissociation of PDAC from other pancreatic cancer forms

Of all pancreatic diseases, the most malignant is pancreatic cancer (PC) which is even among cancer entities one of the most aggressive malignancies with one of the worst prognoses and highest death rates. In this, pancreatic cancer is an umbrella term comprising endocrine tumors as well as neoplasms of the exocrine pancreas. Only about 1-2% of the PC cases arise from the endocrine compartment of the pancreas, so-called neuroendocrine tumors (NET) (Sonbol et al., 2022). The vast majority derives from the exocrine part. The most common

type of PC also belongs to this latter type, namely pancreatic ductal adenocarcinoma (PDAC) accounting for 95% of all PC cases (Milan et al., 2021). Thus, the general term pancreatic cancer is often used synonymously for PDAC. Yet, it is important to discriminate between the different entities as NETs are generally less aggressive than adenocarcinomas of the pancreas.

PDAC is not the most common, yet one of the deadliest cancers worldwide. Both incidence and mortality of PDAC are either stable or even increasing in many countries according to the Global Cancer Statistics 2020 (Sung et al., 2021). It is projected to become the third-leading cause of cancer-related deaths by 2025 after lung and colorectal cancers, then even the second-leading by 2030 with only lung cancer to stay at the top of predicted cancer deaths per year (Ferlay et al., 2016; Rahib et al., 2014). These numbers are impressively accentuated by the alarming 5-year survival residing below 10% worldwide. Only for the United States of America (USA), the American Cancer Society stated an overall 5-year survival rate of 10% for the first time in 2020 (Scholten, 2020). An interplay of various factors contributes to this dreadful prognosis. To begin with, patients usually present with unspecific clinical symptoms, such as abdominal or back pain, unexplained weight loss and loss of appetite, or even no symptoms at early disease stages (Siegel et al., 2021). This is why PDAC is often diagnosed late, and at the time of diagnosis, the disease is already locally advanced or even metastasized. Options of systemic therapy are very limited and mostly ineffective. The only curative treatment option is surgery, however, only 10-15% patients presenting with early local disease are qualified for surgery at diagnosis, but even if operated, the post-surgical recurrence rate is high (Park et al., 2021). All of this results in a median survival of 6-9 months for locally advanced PDAC, respectively 3 months for metastatic PDAC (Adamska et al., 2017; Kleeff et al., 2016).

PDAC is very heterogenous, and the associated lack of biomarkers poses a major problem (Neesse et al., 2019). Compared to homogenous cancers such as breast cancer that is also characterized by a much higher survival rate due to personalized medicine, for instance, PDAC features an extremely high degree of heterogeneity among tumors but especially also within one tumor alone (Peng et al., 2019; Waks and Winer, 2019). Depending on where the samples are taken from in the 3D tumor mass, they present with very different looking histology. This can even hold true for smaller regions on the same sectional plane. This

intratumoral and regional heterogeneity is a hallmark of PDAC (Grünwald et al., 2021). Generally, PDAC is characterized by a highly desmoplastic stroma. Up to 90% of the tumor mass are stromal cells, thus, PDAC has a relatively low content of cancerous cells (Neesse et al., 2011). Stromal cells comprise the non-malignant cells in the tumor microenvironment (TME) such as endothelial cells, neurons, immune cells, and cancer-associated fibroblasts (CAF). Moreover, acellular components of the TME include collagen, hyaluronic acid, and proteoglycans forming a robust extracellular matrix (ECM) (Neesse et al., 2019, 2011). Small signaling molecules, such as cytokines and growth factors, and blood vessels form part of the TME as well, however, PDAC is remarkably hypovascularized as a result of high intratumoral pressure by the ECM destroying the vessels (Neesse et al., 2019; Pereira et al., 2019).

Only in recent years, a new contributor was found and described to play a considerable role in the complex TME of cancers, in particular pancreatic cancer. Microbiota, meaning a community of microorganisms including bacteria, archaea, viruses, and fungi, are not only characteristic but also functional players in PDAC. They are also referred to as the microbiome, the genetic collectivity of all microbiota, which will be further introduced in the second half of this chapter being of central interest in this thesis.

1.2.2 Risk factors and development of PDAC

While other rare forms of PC may also occur in the younger population, PDAC is rarely diagnosed in patients under the age of 40. In fact, the median age at diagnosis is 71 in the USA and almost all deaths occur after the age of 55 (Rawla et al., 2019; Ryan et al., 2014). For this reason, PDAC is also referred to as a disease of the elderly. Moreover, men are somewhat more often affected than women (Rawla et al., 2019). The age-standardized incidence rate of PC is the highest in Europe and North America and remarkably lower in Asia and especially Africa (Sung et al., 2021). While this might also, at least partially, be attributed to the incomplete reporting of case numbers from these regions, some risk factors for PDAC can be rather unambiguously ascribed to Western countries. Risk factors may be grouped into modifiable and non-modifiable variables. Modifiable risk factors include tobacco smoking, heavy alcohol consumption, obesity, certain diets, and the exposure to

toxic substances (Aune et al., 2012; Bosetti et al., 2012; Larsson and Wolk, 2012; Rawla et al., 2019; Wang et al., 2016). Besides gender, age and ethnicity as non-modifiable risk factors, also medical conditions, such as long-term diabetes mellitus, infections, and chronic pancreatitis, may be accounted to this group since their onset can only be influenced to a very limited extent (Ben et al., 2011; Duell et al., 2012; Maisonneuve and Lowenfels, 2015). Some conditions may also be hereditary and other genetic factors, such as genetic variation or (germ-line) mutations, are also not amenable to influence (Ghiorzo, 2014). Some mutations have been identified to play a role in increasing the risk to hereditary forms of PC.

The exact mechanisms by which most of these risk factors transform cells and induce PDAC are not entirely understood. However, the progression from neoplastic cells to PDAC is comparatively well described. The majority of adenocarcinomas, about 60 to 70%, manifest in the pancreas head (Ryan et al., 2014). Although PDAC belongs to the group of exocrine tumors, it does not rise from acinar cells but starts in the epithelial cells of the pancreatic ducts as the name suggests. However, a process called acinar-to-ductal metaplasia (ADM) in which acinar cells undergo differentiation to duct-like cells due to genetic and environmental pressure can potentially lead to precancerous lesions in the event of oncogenic genetic alterations (Wang et al., 2019). In any case, PDAC is preceded by precursor lesions, mostly so-called pancreatic intraepithelial neoplasia (PanIN). PanINs develop in several stages through the accumulation of somatic mutations in a well-studied order. The four stages are PanIN-1A, -1B, -2, and -3 (**figure 3**). PanIN-1A, -1B, and -2 are classified as low-grade, PanIN-3 as high-grade (Hruban et al., 2001). Normal ducts have a cuboidal epithelium with the nuclei being tightly packed and no atypia. The first mutation to occur in over 90% of PDAC cases is a *KRAS* mutation in PanIN-1A and -1B (Fischer and Wood, 2018). These two are highly similar. The cells become tall and columnar with the nuclei arranged at the basement membrane. Solely, in PanIN-1B, the lesions are more papillary. With the progression to PanIN-2, an inactivating *CDKN2A* mutation and nuclear abnormalities accumulate with, for the most part, loss of nuclear polarity (Kanda et al., 2012). PanIN-3 lesions, also formerly referred to as carcinoma *in situ*, are characterized by dystrophic goblet cells and small epithelial cell clusters in the lumen ready for dissemination (Hruban et al., 2001). Typical sites of metastasis are liver, lymph nodes, and lung (Park et al., 2021). This last stage of PanIN lesions is associated with *TP53*, *SMAD4* and *BRCA2* mutations (Guo et al., 2016; Yachida and Iacobuzio-Donahue, 2009). There are many more genes known to be

possibly mutated in PDAC contributing to the extreme heterogeneity of this cancer entity on the molecular level which will be of further interest in the following.

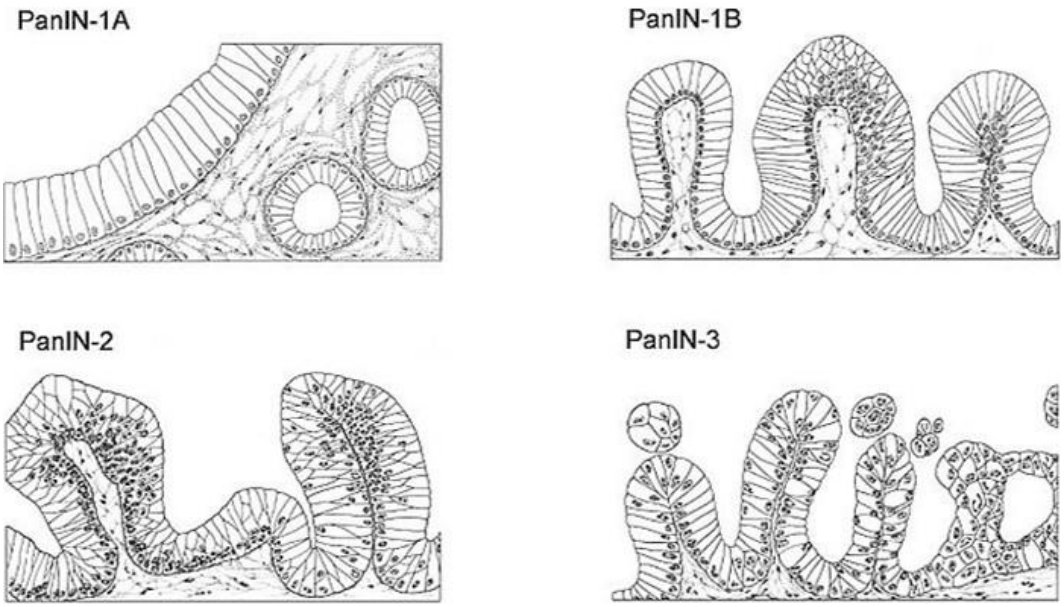

Figure 3: Stages of pancreatic intraepithelial neoplasia PanIN-1A, -1B, -2, and -3. The malignant transformation of the epithelial cells in the pancreatic ducts marks the start in the development of PDAC. PanINs are precursor lesions that develop in several stages through the accumulation of somatic mutations. PanIN-1A (tall columnar cells; *KRAS* mutation), -1B (papillary architecture; *KRAS* mutation), and -2 (nuclear abnormalities; *CDKN2A* mutation) are low-grade, PanIN-3 (cell clusters budding off, carcinoma *in situ*; *TP53*, *SMAD4* and *BRCA2* mutations) is classified as high-grade. Illustration taken from (pathology.jhu.edu/pancreas/medical-professionals/duct-lesions, accessed on 2023-07-24 12:26:00).

1.2.3 PDAC subtypes

Although PDAC is characterized by complex heterogeneity, attempts have been undertaken to classify it into different subtypes. This subtyping is of special interest regarding personalized medicine. However, the extreme genetic heterogeneity, including over 100 to 150 described somatic mutations, in turn explains why precision medicine is particularly difficult to practice in PDAC (Martincorena and Campbell, 2015). In addition, the fact that most somatic mutations are currently not druggable is part of the problem (Stott et al., 2022).

During the last couple of years, several groups developed classification systems of PDAC that are similar or even overlapping. In 2011, Collisson and colleagues pointed out the necessity to define PDAC populations since multiple drug trials failed in unselected ones

(Collisson et al., 2011). They analyzed transcriptional profiles of primary PDAC and defined three subtypes, classical (CLA), quasimesenchymal (QM), and exocrine-like. Moreover, they associated these subtypes with significantly different overall survival between individuals of all three subtypes. Hereby, patients with CLA PDAC had the highest median survival, whereas QM PDAC patients performed extremely worse and were characterized by a significantly shorter overall survival (Collisson et al., 2011). A couple of years later, Moffitt et al. made use of a PDAC gene expression microarray to identify tumor subtypes (Moffitt et al., 2015). Two tumor-specific subtypes, CLA and basal-like (BL), were identified with the BL subtype resembling the QM subtype of Collisson et al. and also having a significantly worse overall survival (Collisson et al., 2011; Moffitt et al., 2015). Since PDAC tumors are mainly composed of stroma, they additionally defined two stroma-specific subtypes termed "normal" and "activated". In this, the activated stromal subtype presented with a worse median survival than the normal subtype. In combination, CLA tumors with normal stroma performed the best, while BL tumors performed the worst almost independently of their stromal subtype (Moffitt et al., 2015). Just one year later, Bailey and colleagues ran an integrated genomic analysis on more than 450 PDAC samples by which they identified 32 recurrently mutated genes (Bailey et al., 2016). From this, the famous "Bailey circle" arose summarizing these genes under ten pathways. As in the other studies, expression analysis led to the definition of four subtypes, namely squamous, pancreatic progenitor, immunogenic, and aberrantly differentiated endocrine exocrine (ADEX). The squamous subtype is comparable to Collisson's QM and Moffitt's BL subtype and it is also presented as the subtype of poor prognosis (Bailey et al., 2016; Collisson et al., 2011; Moffitt et al., 2015). Eventually, Puleo and colleagues presented another gene expression analysis-based classification system comprising five PDAC subtypes: pure BL, stroma activated, desmoplastic, pure CLA, and immune CLA (Puleo et al., 2018). The overlap with the previous publications is evident, yet, they questioned the existence of an ADEX subtype as did Moffitt et al. regarding the predictive power of Collisson's subtypes when applied to their data. Puleo et al. also associated their subtypes with patient outcomes and found the pure BL to have the worst outcome, while the pure and immune CLA had a good prognosis (Puleo et al., 2018). Taken together, general recurrent subtypes of PDAC are the QM/BL/squamous and the CLA subtypes characterized by a worse prognosis and bad

therapy response, and a better prognosis as well as more effective therapy response, respectively. Therefore, they will be referred to as BL and CLA in this thesis.

Considering the pronounced heterogeneity in PDAC, however, it is not trivial to classify these tumors strictly into two categories. Usually, PDAC samples show both CLA and BL features in adjacent areas. Depending on the degree of this subtype heterogeneity, it is advisable to take a mixed form into account. Most recently, Chan-Seng-Yu and colleagues introduced the so-called hybrid (HYB) PDAC subtype as they found BL and CLA programs to co-exist intratumorally (Chan-Seng-Yue et al., 2020). The hybrids presented multiple expression signatures wherefore they were classified inconsistently in previous studies. Because of that, the HYB subtype was additionally included in the analysis of this thesis.

Apart from these patient stratification approaches based on genomic and transcriptomic signatures, metabolic subtypes have been defined since metabolic reprogramming toward high glucose consumption and increased lipid uptake characterizes PDAC tumors (Espiau-Romera et al., 2020; Liang et al., 2016). The subtypes termed “glycolytic” and “lipogenic” are characterized by distinct metabolite levels corresponding to glycolysis and lipogenesis, respectively (Daemen et al., 2015). Most importantly, the lipogenic subtype is associated with the CLA signature, while the glycolytic correlates with the BL profile thereby also providing evidence for differential prognosis and the potential to predict PDAC patient survival (Espiau-Romera et al., 2020).

1.2.4 Therapeutic strategies

One of the main challenges with PDAC are the late diagnosis at an advanced or even metastasized disease state which demands for potent treatment options. However, available therapeutics are limited and mostly ineffective. PDAC is very refractory to systemic therapies due to rapid resistance mechanisms and its unique tumor architecture posing physical barriers, such as the pronounced stroma and hypovascularization, to the agents, although stromal depletion did not result in enhanced therapeutic response either (Neesse et al., 2015).

In line with this clinical situation, the only curative treatment option is resection, however, only few patients are diagnosed early enough for surgery (Heinemann et al., 2013;

Konstantinidis et al., 2013). Resectability depends on the disease progress, so in order to stratify patients, tumor staging is performed (**figure 4**). At stages I and II, tumors may be surgically resectable if they are not involved in the local vessels. These stages are further defined by no invasion to lymph nodes and no spread to distant sites (stage I) or only local spread to close organs or lymph nodes (stage II). In stage III, the tumor has reached the major blood vessels, and in stage IV, it has metastasized to distant organs, such as typically liver and lung.

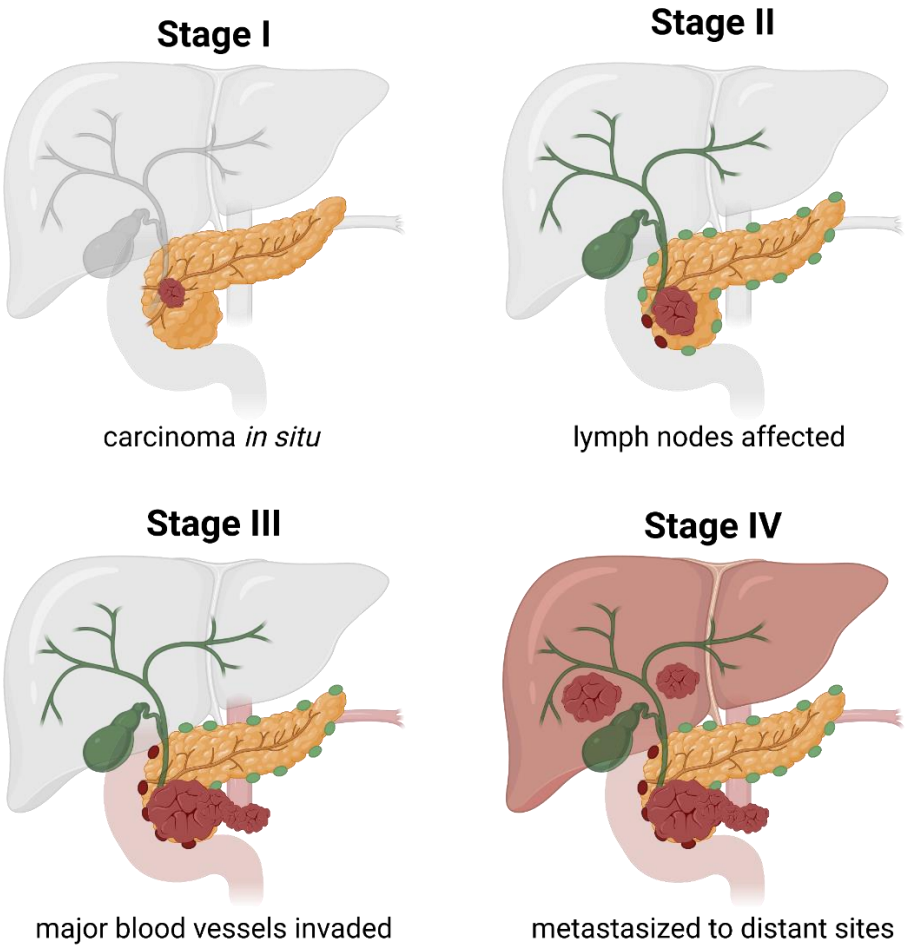

Figure 4: Tumor stages of pancreatic ductal adenocarcinoma. The development of PDAC may be divided in several stages (I-IV) depending on the disease progress and invasiveness of the tumor. In this illustration, surrounding organs that are not (yet) affected in the regarded stage are grayed out. Stage I is also referred to as carcinoma *in situ* which usually developments in the pancreas head. In stage II, nearby tissues which may include lymph nodes are affected by the growing tumor. When PDAC reaches stage III, it invades the major surrounding blood vessels facilitating spreading to distant sites what happens in stage IV, the metastasis to typical organs such as liver and lung. Own illustration created with (biorender.com).

Neoadjuvant (before surgery) chemotherapy is administered in order to increase the percentage of surgical candidates from patients that are considered "borderline resectable" although this may also unfavorably delay surgery (Heinemann et al., 2013; Park et al., 2021). Adjuvant (after surgery) chemotherapy is given due to the poor prognosis despite surgery due to the high recurrence rate (Ryan et al., 2014). The two most frequently deployed chemotherapeutic agents are in fact usually combination therapies. Antimetabolites, such as gemcitabine (GEM) and 5-fluorouracil (5-FU), or agents with DNA damage effects, such as oxaliplatin and irinotecan, are the main chemotherapeutics used in the clinics (Park et al., 2021). GEM is usually administered to patients with worse performance status, i.e., who are older and show overall stronger symptoms and weakness. A common combinatory drug is nano-formulated albumin-bound paclitaxel (nab-PAC). GEM alone does not significantly prolong survival and patients who received GEM show a median survival of 6.7 months in a study by Von Hoff and colleagues (Von Hoff et al., 2013). In the same study, patients treated with GEM/nab-PAC combination therapy had a significantly prolonged survival of 8.5 months as compared to GEM monotherapy. Another multiagent chemotherapy commonly made use of in the clinics is the combination of fluorouracil, irinotecan, leucovorin, and oxaliplatin (FOLFIRINOX). A study from 2011 found a significant difference in the median survival of patients receiving FOLFIRINOX (11.1 months) and GEM (6.8 months) (Conroy et al., 2011). However, FOLFIRINOX is only recommended for patients with high functional status after resection, usually meaning younger patients with overall good health (Park et al., 2021).

Radiotherapy is controversially discussed as an adjuvant therapy for resected PDAC, however, there are novel therapies considered to treat stratified subgroups of PDAC patients. For example, for individuals with *BRCA* mutations targeted (immune) therapy are tested (O'Reilly et al., 2020). However, although immunotherapies revolutionized cancer treatment across various tumor entities, immune checkpoint inhibitors (ICI) have proven to be ineffective in PDAC (Hosein et al., 2022; Ott et al., 2019; Royal et al., 2010). Still, targeting immune cells or other components of the TME such as CAFs in a subset-specific manner are again on the rise (Biffi et al., 2019; Buechler et al., 2021; Elyada et al., 2019). Gene alterations in *KRAS*, the most frequently mutated gene in PDAC, are subject to targeted agent development and may be of great value. Despite all of these approaches, PDAC remains one

of the deadliest cancers and the need for novel therapies and biomarkers persists. In this thesis, the latter will be of special interest with regard to the aforementioned microbiome.

1.3 Mouse models of PDAC

1.3.1 The KPC model

The traits and challenges of PDAC have been elucidated in **1.2**. Many observations were made in humans first, however, in order to understand the origins of this malignancy, to investigate tumor biology, and to develop new therapies and biomarkers, model systems are required. As for most other diseases, animal models are indispensable to recapitulate this complex human disease and to conduct comprehensive research on PDAC. Mice are the most often used animals to study PDAC and a plethora of models are established for different aspects of the disease, such as different mutations or tumor progression rates. Conclusions drawn from these mouse models may then later be translated to the clinics and tested on humans. Hereby, it is primarily crucial to understand and characterize the model itself in order for it to serve as reliable research foundation which is also part of this thesis regarding the microbiome.

The most commonly used genetically engineered mouse model (GEMM) in PDAC research is the Lox-Stop-Lox (LSL) system-based *LSL-Kras*$^{G12D/+}$*;LSL-Trp53*$^{R172H/+}$*;Pdx-1-Cre* (KPC) model. It has first been developed and described about 20 years ago in order to define the genetic basis of PDAC, and comprises oncogenic *Kras* and a heterozygous *Trp53* mutation (Hingorani et al., 2005). Both transgenic gene alterations are activated exclusively in the pancreas by the tissue-specific Cre recombinase under the Pdx-1/Ipf1 promotor (Gopinathan et al., 2015). However, Pdx-1 promotor leakage still exists to a low percentage causing benign papilloma growth in the facial and anal area of these mice. Therefore, the Cre in this GEMM may also be expressed under the P48/Ptf1 promotor whose expression is restricted to acinar cells (Hingorani et al., 2003). Pdx-1 is expressed in a stochastic manner in the early developmental phase of the pancreas and confined to islet cells (Hingorani et al., 2003; Offield et al., 1996). In this thesis, exclusively KPC mice with Pdx-1 Cre were used.

As mentioned before, the Kirsten rat sarcoma viral oncogene homolog (*KRAS*) is the most frequently mutated proto-oncogene in humans and postulated a critical event in the oncogenesis of exocrine PCs (Almoguera et al., 1988). The most common codon mutation is the G12D variant with a frequency of 39.2% (Luo, 2021). It causes uncontrolled cell grow and proliferation via a constitutively active *KRAS* (Gopinathan et al., 2015). Therefore, it is an essential part of the KPC model.

Tumor protein p53 (*TP53*) is a tumor suppressor and the most frequently mutated gene in all cancers. Also in PDAC, it has a mutation rate of 50 to 75% (Morton et al., 2010). Mutated *TP53* prevents apoptosis and simultaneously upregulates the cell division rate (Gopinathan et al., 2015). Several hotspot mutations in the DNA binding domain of TP53 are described. They are mostly missense mutations with oncogenic gain-of-function. This is also the case for the R175H which has the highest occurrence in cancer (Chiang et al., 2021). In PDAC, it has a mutational frequency of 4.5% (Cicenas et al., 2017). The respective homolog in mice R172H is a crucial part of the KPC model as a requirement for aggressive disease progression (Hingorani et al., 2005).

Driven by the *Kras* and *Trp53* transgenes, the KPC model has almost 100% tumor penetrance. After approximately 8 weeks, PanIN lesions develop, and between 12 to 50 weeks, fibrotic tumors grow out. This stage is accompanied by the typical features of human PDAC including liver and lung metastases, cachexia, jaundice, and ascites (Gopinathan et al., 2015; Lee et al., 2016). Further, also on the molecular, histological, and pathological level, murine KPC tumors resemble human PDAC to a great extent, including an immunosuppressive TME, for instance (Lee et al., 2016; Mallya et al., 2021). Generally, the KPC model is well-described and has been studied in many regards including different therapeutic options and biomarkers for PDAC (Eresen et al., 2020; Ligat et al., 2015; Orr et al., 2023; Ray et al., 2012; Zhang et al., 2022). KPC mice feature a median survival of 5 to 6 months (Westphalen and Olive, 2012).

1.3.2 Orthotopic transplantation of tumor cells

Besides *in vivo* experiments, *in vitro* studies with established and primary PDAC cell lines are put into practice in PDAC research. The combination of both via the orthotopic

transplantation of tumor cells into the pancreas of mice offers some advantages over transgenic mouse models. For instance, experiments may be timed according to the researcher's needs and tumors grow much faster to the desired size, wherefore the costs for orthotopic transplantation experiments are comparably low. Orthotopic models are more standardized with less variance in tumor latency and progression (Qiu and Su, 2013a). Still, orthotopically transplanted mice show metastasis and ascites formation like GEMMs (Mallya et al., 2021).

When referring to orthotopic transplantation, the direct injection of tumor cells to the site of interest is meant, in this case the pancreas. Various publications have elucidated the procedure and provide guidance for the cautionary steps during the surgery (Erstad et al., 2018; Qiu and Su, 2013a). Two types of orthotopic transplantation may be discriminated. In syngenic mouse models, cells that are genetically (sufficiently) identical and immunologically compatible to the receiving animal are used for transplantation. For instance, KPC-based tumor cells are orthotopically transplanted to the pancreas of wildtype C57BL/6 mice. These are immunocompetent allograft tumor models that can replicate essential features of PDAC such as the complex TME since all involved cells and components are murine. Predictable growth kinetics make these models very suitable for therapeutic studies (Erstad et al., 2018). In the other scenario of orthotopic transplantation, the goal is to graft human PDAC cells to mice (xenograft) in order to grow more humanized tumor material. However, the disadvantage of this model is that immunodeficient mice, such as athymic nude mice, are needed which do not fully recapitulate the course of natural human PDAC oncogenesis (Qiu and Su, 2013a). Immunodeficiency is of course necessary to prevent graft rejection in this xenograft tumor model. Further, it is important to consider that the tumor stroma derives from the mouse. Still, this mouse model is widely used for its advantages, such as a high metastasis rate and the option to label the tumor cells in culture prior to transplantation in order to track tumor growth *in vivo* (Loukopoulos et al., 2004; Qiu and Su, 2013b).

Within the scope of this thesis, tumor samples of orthotopic pancreatic cancer xenograft models were used. Tumors from nude mice orthotopically transplanted with the established human PDAC cells lines MiaPaCa2, Capan-2, and Capan-1 were subject to microbiome signature investigation in a PDAC subtype-dependent manner since these cell lines are

representatives of more BL or CLA tumors. MiaPaCa2 derived from the primary tumor of a 65-year-old male patient which had infiltrated surrounding tissues (Yunis et al., 1977). It has a mesenchymal phenotype and is poorly differentiated. Therefore, it is classified as BL. Capan-2 also originates from a primary tumor of a 56-year-old male, however, it is well differentiated and therefore assigned to the CLA subtype (Kyriazis et al., 1986). Finally, the Capan-1 cells arose from the liver metastasis of a 40-year-old man with pancreatic adenocarcinoma, and are also members of the CLA subtype as they are well differentiated (Kyriazis et al., 1982).

MiaPaCa2- and Capan-2-derived tumors are used in direct comparison in this thesis with histologically adjacent normal tissue as additional differential tissue compartment groups. Capan-1-derived tumors were treated with tumor necrosis factor α (TNFα) and its vehicle control (VC) water (H_2O), respectively (Tu et al., 2021). TNFα was reported to cause lineage reprogramming of CLA tumor cells towards a BL phenotype, and this effect was even reversible via bromodomain and extra-terminal motif (BET) protein inhibitors (Tu et al., 2021). Due to these highly interesting new insights into PDAC subtype plasticity, these tumors samples were of great interest within the scope of this thesis.

1.4 The microbiome in health and disease

1.4.1 The gut microbiome

The entirety of bacteria, archaea, viruses, and fungi are summarized as microbiota, or microbiome if referring to their genetic total more precisely, but these terms are also used synonymously. The microbiome may also refer to a microbial community with their accompanying structural elements in a habitat (Berg et al., 2020). Since all ecosystems and organisms, living and dead, are colonized by microbiota, specific sub-microbiomes are usually discussed and compared. For instance, site-specific microbiomes in soil, fresh or salt water are examined in environmental sciences, and organ-specific microbiomes of skin, oral cavity, or intestines in humans are regarded in biomedical sciences. In addition, institution-specific microbiomes from different hospitals or laboratories can be of interest.

Before the age of modern sequencing techniques, microbiota, in particular bacteria, have been studied traditionally on agar plates, in liquid cultures or under the microscope. The advent of next generation sequencing (NGS) then offered a method to investigate microbial populations on a large scale and to characterize them more precisely based on genetic relationship rather than on similar visual appearances. This led to the fundamental observation in biomedical microbiome research which is a dependency of the microbiome on a plethora of extrinsic factors such as ethnicity, food, lifestyle, and wealth, as well as intrinsic factors such as age and diseases leading to a massive inter- and even intra-group variability of the microbiome. While significant inter-group differences are widely published, the intra-group differences can also be profound meaning that although two samples derive from healthy subjects, they might look highly different in their microbial compositions (del Castillo et al., 2019).

The gut microbiome is by far the best-studied and in many regards the largest human sub-microbiome. Up to 1,000 different inhabiting microorganisms have been described, with the vast majority being represented by bacteria and only up to 40 different species of such (Beaugerie and Petit, 2004; Guarner and Malagelada, 2003; Sears, 2005). This is the largest number compared to other body sites (Quigley, 2013). Feces consist of up to 60% of bacteria in their dry mass which makes them an excellent sample material to study the microbiome (Stephen and Cummings, 1980). Gut bacteria are mostly anaerobic with few site-specific exceptions, however, in feces, there is a gradient of aerobic to anaerobic bacteria towards the inside. It is also known that the microbial composition varies along the intestinal system with the colon harboring the largest microbial density (Shapira, 2016). Typical human gut bacterial genera are *Bacteroides*, *Clostridium*, *Faecalibacterium*, *Eubacterium*, various *Cocci*, and *Bifidobacterium* (Beaugerie and Petit, 2004; Guarner and Malagelada, 2003).

The healthy gut microbiome is much more than a community of commensals which refers to a mere non-harmful coexistence of microorganisms and human cells. Instead, an essential symbiotic relationship between the human body and the gut microbiome establishes throughout lifetime where the latter exceeds a plethora of physiological functions. As reflected by the enormous number of publications since the start of gut flora research in the mid-90s, particularly the mere existence of several review articles all simply termed “The gut microbiome”, it is a topic of great extent and importance (Brody, 2020; Gibson and

Roberfroid, 1995; Kuziel and Rakoff-Nahoum, 2022; Sidhu and van der Poorten, 2017). The intestinal microbiome is involved in basically all aspects of the organism, such as the circadian rhythm, brain function, psychiatric concerns, immune system, metabolism, pH regulation, and many more (Kiss et al., 2020; Mohajeri et al., 2018; Shi et al., 2017; Van Ameringen et al., 2019; Voigt et al., 2016).

In diseased states of the organism, the gut flora may be the cause, a contributor or an indicator, so-called biomarker. However, it is not trivial to determine whether the microbiome is in fact causal for the respective disease. When investigating the microbiome in a disease context, it is first a shift in the microbial composition that is observed, called dysbiosis. The central question in disease-related microbiome research is the "chicken or egg" paradigm, so whether changes in the microbiome are the cause or an epiphenomenon of the disorder (Gorkiewicz and Moschen, 2018). The gut microbiome is often involved in gastrointestinal diseases such as bowel perforation, inflammatory bowel diseases, and intestinal cancers, in the leaky gut syndrome, and subsequently in maladies such as sepsis, but also connections to other cancer types have been found (Adelman et al., 2020; Liu et al., 2019). In the context of the latter, the gut microbiome is often discussed as a possible biomarker and many studies reported associations between the microbiome and the regarded disease. The challenge is then to transfer these associations to the clinics. To date, there are several options established how to modulate the microbiome. Besides unintentional gut flora disturbances via drugs such as antibiotics, there are pre-, pro-, and postbiotics available. Moreover, procedures such as dietary modulation, phage therapy, and fecal microbiome transplantations (FMT) are available (Cullin et al., 2021). The relation between the gut microbiome and pharmaceuticals is in fact bidirectional. It is manipulated by the drugs (pharmacodynamics) but microbes can just as well metabolize drugs and influence their structure, bioavailability, activity, and toxicity (pharmacokinetics) (Weersma et al., 2020).

Investigating the microbiome requires appropriate measures to quantify and display the composition of samples and differences between groups. The most commonly used readouts are alpha and beta diversity. Other than that, there is gamma diversity, though rarely applied in this context. These terms derive from the field of ecology and describe species diversity at different scales. As illustrated in **figure 5**, alpha diversity means the mean

species diversity at a local site, while beta diversity describes the differentiation between those sites, and gamma diversity comprises the total species diversity in a certain area (Whittaker, 1972, 1960).

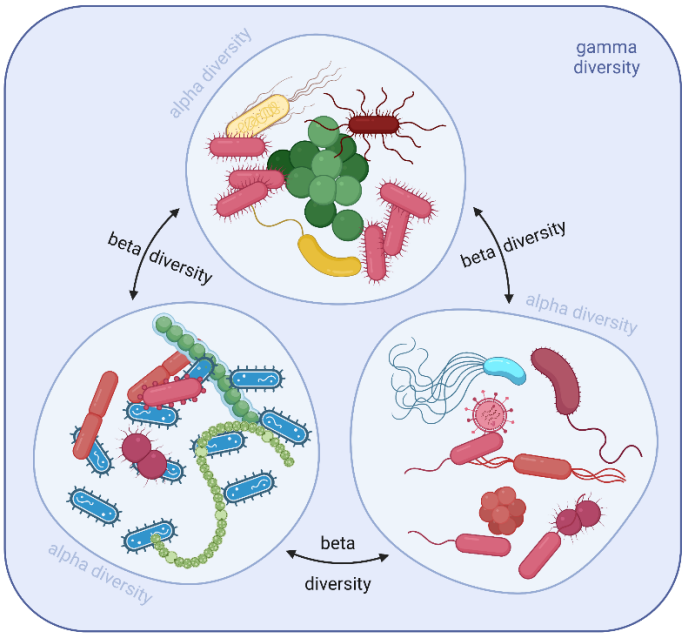

Figure 5: Schematic representation of alpha, beta, and gamma diversity. These three diversity levels originating from the field of ecology describe heterogeneity at different scales. Applied to microbiology, alpha diversity refers to the diversity in single samples and considers their richness and evenness, for instance. Beta diversity compares two sample (sites) or more among each other and works with distances and dissimilarities. Finally, gamma diversity is about an even broader spectrum and means the total species diversity in a defined area, e.g., the laboratory. Own illustration created with (biorender.com).

In the context of molecular biology, samples can be conceived as sites. Regarding alpha diversity, the most common measure is observed species (OS), also known as richness. It counts the number of different species in one sample. Moreover, the Shannon index (SI) and inverse Simpson index (InvS) are applied in this thesis. SI gives an idea about the species entropy within one sample, i.e., how evenly the different species are distributed (evenness) and it is the higher the more equal the abundances of the present species are. InvS derives from the Simpson index which gives a probability that two species randomly picked from a population are the same (Simpson, 1949). Thus, the lower the Simpson index value is, the higher is the diversity in the regarded system. Since this is rather non-intuitive, InvS is often reported. It is defined by one divided by the Simpson index. Regarding beta diversity,

measures based on species abundance and their phylogenetic relation are commonly used. UniFrac distance metrics reflect the relative relation between the contained species via the phylogenetic distance. Unweighted UniFrac distance (uwUF) only respects the presence or absence of species, while weighted UniFrac distance (wUF) also takes their abundances into account (Lozupone and Knight, 2005; Lozupone et al., 2007). In contrast, Bray-Curtis dissimilarity (BC) focuses on the compositional dissimilarities between the groups by considering species counts (Bray and Curtis, 1957).

Making use of the gut microbiome as a biomarker is a strongly pursued research area as there are many diseases for which no reliable markers are established. In particular, the possibility to predict cancer cases from the microbiome would be a milestone, especially in PDAC research. The interdependence between gut and pancreas microbiome has been described and some studies have already suggested the fecal microbiome as a non-invasive biomarker (Half et al., 2019; Kartal et al., 2022; Li et al., 2020; Nagata et al., 2022; Ren et al., 2017). This is also of special interest within the scope of this thesis.

1.4.2 The pancreatic microbiome

As shortly introduced in **1.2.1**, the microbiome has recently become a highly intriguing new player in the predominantly TME- and subtype-driven PDAC research area. This is due to today's understanding of the microbiome as an integrated part of the human organism rather than a pure bystander. This perception started to change historically with the introduction of high-throughput sequencing methods allowing to discover microbiomes in internal organs that had long thought to be sterile. It is now broadly accepted that they harbor at least a low biomass of microbiota. Even more so, the human microbiome has been termed a complex organ itself (Baquero and Nombela, 2012).

The pancreas as one of the organs that was discovered to be colonized by a low amount of microbes in samples of organ donors is now offered a new vantage point for PDAC research (Geller et al., 2017). However, in contrast to the gut microbiome which can easily be sampled via stool or swabs, for instance, the pancreatic microbiome is harder to study in living organisms as it requires invasive biopsies. Although it is very important to learn about the pancreatic microbiome, the connection between gut and pancreas may also be exploited in

this regard. The pancreas influences the gut via the secretion of antimicrobial peptides physiologically, and pancreatic insufficiency in turn may lead to small intestine bacterial overgrowth (SIBO) and dysbiosis in the intestinal microbiome (Ammer-Herrmenau et al., 2020) (**figure 6**). Therefore, the correlation of the intestinal microbiome with diseases of the pancreas is of great interest (Ammer-Herrmenau et al., 2021a).

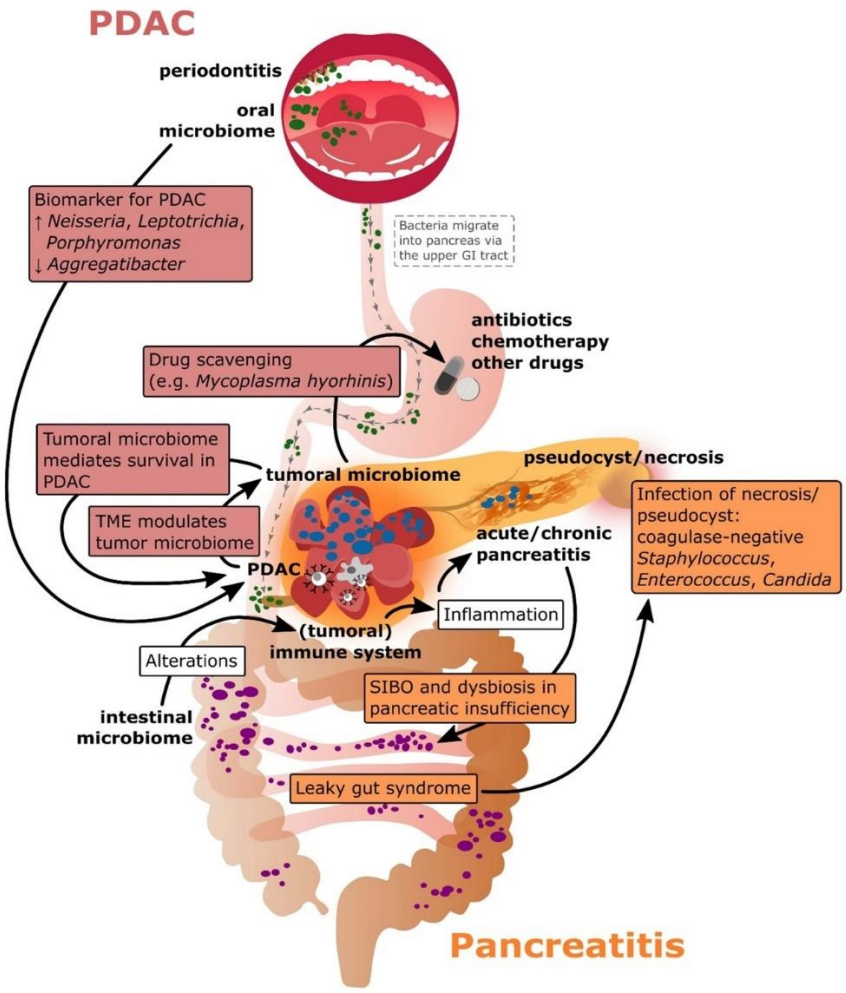

Figure 6: Summary of current findings and knowledge about the orointestinal and tumoral microbiome in PDAC and pancreatitis. Red boxes refer to PDAC-related, orange boxes to pancreatitis-related key findings. Illustration taken from (Ammer-Herrmenau et al., 2020) (CC BY 4.0).

The pancreatic microbiome of healthy individuals has been studied only sparsely compared to the PDAC tumor microbiome. This is mostly due to technical difficulties in obtaining healthy pancreas sample material. Most studies only applied histologically normal tissue from adjacent to tumor tissue or even benign pancreatic disease samples as controls for their tumor samples. Yet, in order to elucidate microbiome alterations associated with PDAC, a

profound characterization of the normal pancreatic microbiome is necessary. A few studies examined normal human pancreatic tissue and found *Chlamydiales*, *Brevibacterium*, *Acinetobacter*, *Enterobacter*, *Pseudomonas*, and *Lactobacillus* in higher relative abundances as compared to PDAC samples, however, not all of these differences were reported to be significant (del Castillo et al., 2019; Pushalkar et al., 2018; Thomas et al., 2018).

Although the microbiomes of all pancreatic diseases may offer platforms for new research, the focus within the frame of this thesis is set on PDAC. To date, the tumoral pancreatic microbiome has been investigated in multiple studies. Generally, microbes were found to be more abundant in tumors as compared to healthy pancreatic tissue. The most prominent phylum across published microbial compositions is *Proteobacteria*, comprising mainly *Gamma-* and *Alphaproteobacteria* (Chakladar et al., 2020; Pushalkar et al., 2018; Riquelme et al., 2019). The constantly increasing number of publications has shown the impact of the microbiome on cancer, and PDAC in particular, including aspects such as subtypes, development, and therapy response. It can be broken down to three most significant findings. First, tumors harbor their own specific microbiome which differs amongst tumor entities and is distinct to their respective healthy control tissue (del Castillo et al., 2019; Nejman et al., 2020; Pushalkar et al., 2018). Second, the BL subtype of PDAC which shows a higher inflammatory response has a more diverse microbiome as compared to the less aggressive subtypes, CLA and HYB (Guo et al., 2021). And last, there is evidence for the tumoral microbial signature and diversity of PDAC to predict patient long-term survival (Riquelme et al., 2019). These findings underpin the importance and the potential of the microbiome in being a highly relevant factor in PDAC. In the following, its role in PDAC pathogenesis, diagnostics, therapy, and prognosis will be described in more detail.

1.5 The impact of the microbiome on different aspects of PDAC

1.5.1 The role of the microbiome in PDAC pathogenesis

Regarding the microbiome as causal for PDAC pathogenesis in the “chicken or egg” paradigm is not too far-fetched considering the well-established knowledge about infectious microorganisms leading to cancer formation via chronic inflammation. Furthermore, the fact

that healthy pancreas does not harbor that kind of microbes to that extend seems to at least make this consideration plausible. Over 16% of all cancer cases can directly be traced back to pathogen infections and of most the mechanisms are well-described (de Martel et al., 2012; Schwabe and Jobin, 2013). So far, eleven distinct "oncomicrobes" are known to induce cancer (IARC Working Group on the Evaluation of Carcinogenic Risks to Humans, 2012). This is not yet the case for PDAC. However, some microbes leading to the formation of other cancer entities, for instance, via dysregulation of the Wnt/β-catenin signaling pathway, were also found in PDAC, such as *Clostridium*, *Bacteroides*, and *Escherichia* (Silva-García et al., 2019). The Wnt/β-catenin signaling pathway is also one of the significantly altered pathways in PDAC, wherefore associations like this offer vantage points for functional studies of the microbiome in PDAC carcinogenesis (Bailey et al., 2016).

Besides the activation of oncogenic signaling, there are some additional feasible ways of how microbes may contribute to PDAC onset and maintenance. On the one hand, direct carcinogenic effects can be achieved via mutagenesis. Here, a similar analogue as for oncogenic signaling can be drawn which once again underpins the likelihood of microbes to at least contribute to PDAC. *Fusobacterium nucleatum* is speculated to contribute to carcinogenesis and oncogenesis of oral squamous cell carcinoma cells by causing DNA double strand breaks via the Ku70/p53 pathway (Geng et al., 2020). This bacterial species was found in human PDAC samples as well (Nejman et al., 2020). On the other hand, indirect carcinogenic effects may be conferred via chronic inflammation. Association studies have shown the oral microbiome and periodontitis to connect with PC development in some way (Farrell et al., 2012; Michaud et al., 2007). Another study stated that bacterial structures such as lipopolysaccharide (LPS), also called microbe-associated molecular patterns, promote PC (Ochi et al., 2012). Indirect oncogenic effects can also be transmitted via microbial metabolites which has been shown in colonic cancer for short-chain fatty acids (SCFA), for example (P. Liu et al., 2021). Eventually, the interplay between the microbiome and the immune system is a highly relevant point to revisit with regard to immunogenic TME reprogramming. Since PDAC is characterized by a pronounced inflammatory reaction, this consideration seems natural. Based on various study results, microbial invasion is perceived as a potential driver of carcinogenesis via immune system remodeling (Alam et al., 2022; Chakladar et al., 2020; Pushalkar et al., 2018).

Although the exact mechanisms of how microbes, as a cause or contributor, are involved in PDAC are not clear to date, there is evidence suggesting that there indeed is a link between the microbiome and the pathogenesis of PDAC. Detailed functional studies are not part of this thesis; however, their meaning and impact are noted in **chapter 6**.

1.5.2 Application of the microbiome in PDAC diagnosis

As a consequence of knowing that the microbiome is an important, possibly even central part of the PDAC TME, its exploitation for urgent open needs in the clinical context of PDAC is the next logical step. The high demand for reliable biomarkers for PDAC has been addressed earlier in this chapter, and both the tumoral and intestinal microbiomes are considered relevant sampling sources. At this point, the imperative of thorough microbiome characterizations becomes clear. By defining a common base despite inter- and intra-group differences, the microbiome may ideally be used as a biomarker to distinguish cancer from non-cancerous cases and as a target for new therapeutic strategies which will be further covered in **1.5.3**.

The missing of a reliable, non-invasive screening method for PC even today poses a major problem in the clinics. The tumoral microbiome can obviously not be considered for this purpose, yet, needs to be studied to understand tumor biology to tackle the other fundamental problems of PDAC such as the alarmingly limited therapy options. Other local microbiomes of PDAC patients are studied to find patterns and differences compared to healthy individuals. Some of the most convenient microbial sites for clinical practice are the oral cavity and gut. Indeed, the orointestinal microbiome was investigated in a great number of studies as potential PDAC biomarker (Fan et al., 2016; Nagata et al., 2022; Olson et al., 2017; Pushalkar et al., 2018). The sampling is non-invasive and easily realizable via oral washes, salvia, different oral and rectal swabs, or stool. However, this is also what bears the risk of heterogenous results and makes it more difficult to standardize the outcome. Moreover, the sequencing technique applied in these studies introduces great variability which will be discussed in detail as part of the results in this thesis.

What is also routinely sampled from patients at hospital admission, both outpatient and in-patient, is blood. Blood draws are minimally invasive, easy to access and a new parameter

to be determined could potentially be implemented in laboratory routine analyses. In 2013, Michaud et al. conducted a prospective study with blood samples from PDAC patients pre diagnosis and found high levels of the oral *Porphyromonas gingivalis*, the periodontitis-causing pathogen correlated with 2-fold increased PDAC risk (Michaud et al., 2013). A more recent but less PDAC-specific study analyzed microbial nucleic acid plasma profiles of more than 10,000 patients and found that these were capable of predicting different cancer types (Poore et al., 2020). A third microbial structural element that are currently considered to be exploited for diagnostic assessments are bacterial extracellular vesicles (bEV). A metagenomic study analyzing bEVs in serum samples demonstrated their diagnostic potential to differentiate ovarian cancer from benign ovarian tumors (Kim et al., 2020). This differential diagnosis is of strong importance in the context of PDAC as well since other benign pancreatic conditions are often confused with it and needless surgeries could potentially be avoided. In a retrospective propensity score matching study, the composition of bEVs was indeed found to be distinguishable in blood samples of PC patients (Kim et al., 2021). However, these studies are still quite preliminary and require refinement, in particular regarding the biggest issue in microbiome research, namely contamination.

1.5.3 Influence of the microbiome on PDAC therapy and prognosis

The fact that the very poor prognosis of PDAC did not improve significantly during the last couple of decades is a wake-up call for researchers to reach out for alternative ways how to tackle this problematic challenge. Learning about the microbiome, a new option to perform prognostic assessment via microbial signature profiling could indeed be feasible. Moreover, exploiting the microbiome for therapy is no longer hypothetical but in fact possible in multiple ways.

Anticancer treatment involving the microbiome is not performed routinely yet, but there are several conceivable strategies. Considering the existence of a tumoral microbiome which differs from the healthy organ's microbiome, the most straightforward concept is of course to modulate the microbiome. Intervening with the microbiome has the potential to support and improve other standard anticancer therapy effects as it was reported to act against chemotherapy, for instance (Geller et al., 2017). Possible intervention techniques include

change of diet, pre- (compounds to enhance beneficial microorganisms), pro- (living microorganisms), post- (metabolic products of microorganisms), and antibiotics (antimicrobial substance), phage therapy, and FMT (Han et al., 2021; Rodriguez-Arrastia et al., 2021; Routy et al., 2018; Tanoue et al., 2019; Zheng et al., 2019). In phage therapy, bacteriophages are equipped with nano-sized chemotherapy molecules that are directly delivered to the requiring site. This could potentially limit side effects that are known from systemic therapy. Moreover, it offers the possibility to target specific bacteria depending on the bacteriophage that is chosen. Compared to broad spectrum antibiotic treatment, this would be a real game changer also for *in vivo* studies that aim at elucidating the effect of certain bacteria within the microbial community of a particular microbiome.

Microbes may have a direct impact on anticancer therapies, including chemotherapy, immunotherapy, and even radiotherapy (Geller et al., 2017; J. Liu et al., 2021; Sivan et al., 2015; Vétizou et al., 2015). For example, dysbiosis in the gut microbiome was found to elicit ICI resistance in cancer patients (Routy et al., 2018). The evidence on microbes influencing chemotherapy efficiency, positively and negatively, has remarkably increased recently (Guenther et al., 2022; Kesh et al., 2022; Panebianco et al., 2022). Also, it is not limited to PDAC, a fact that further confirms the validity of the hypothesis.

Besides the sole association of the microbiome with subtypes of PDAC mentioned in **1.4.2** as one of the major findings in PDAC microbiome research of recent years, this reveals the predictive value of the microbiome as subtypes have varying prognosis as a consequence of their differing response to chemotherapy (Guo et al., 2021). When thinking about prognosis in PDAC, the rare cases of long-term survival of this dismal disease, i.e., more than 5 years after surgery, come to mind. To date, there is no uniform explanation why this is the case for a few patients to live for 10 or 20 years after PDAC diagnosis. Riquelme and colleagues found a specific microbiome signature in long-term survivors (LTS) that was not present in short-term survivors (STS) (Riquelme et al., 2019). With this microbial indicator of patient outcome, a new era of PDAC perspectives was heralded.

2 Thesis objectives

Conceiving the enormous shortcomings in basically all aspects of PDAC which is about to become the deadliest human malignancy, this is a wake-up call for researchers hard to beat for its urgency. Having approached PDAC research in many other regards before, it was now the time to explore a new field with great potential to be exploited for all aspects of the disease including diagnostics, therapy, and prognosis. There is emerging evidence and promising data from the field for a tumor- and tumor entity-specific microbiome to exist as well as for its involvement in regards such as patient long-term survival. Thus, the project of this thesis aimed at this new player in the highly dynamic network system of PDAC: the microbiome. The objective within the scope of this thesis towards this new research focus comprises four precise aims:

(i) Provision of evidence for the presence of bacteria in murine KPC tumors and human PDAC samples.

(ii) Establishment of third generation sequencing by Oxford Nanopore Technologies and its optimization for murine fecal and pancreatic samples.

(iii) Analysis of the microbial composition differences in KPC mice in
   a. fecal samples during tumor progression,
   b. tumor vs. healthy pancreas samples, and
   c. tumor center vs. periphery samples,
   with the comparison to published human stool and PDAC microbiomes.

(iv) Investigation of impact of transcriptional PDAC subtypes on the microbiome in
   a. orthotopically transplanted human PDAC cell lines of the CLA (Capan-2) vs. BL (MiaPaCa2) subtype,
   b. orthotopically transplanted human PDAC cell lines of the CLA subtype (Capan-1) with TNFα and control treatment, and
   c. human PDAC samples with their subtyping via IHC targeting marker proteins GATA6 and CK5.

With intent to answer open questions asking for the dynamics of the microbiome during tumor evolution and the host genetic/transcriptomic profile influences on the microbiome, the overall hypothesis of this thesis is that PDAC subtypes impinge on their intratumoral

microbiome potentially offering new therapeutic options. Using the microbiome as a key to change BL to CLA subtypes and exploiting it to improve therapy response is highly theoretical but a somewhat ideal conception.

3 Material and methods

3.1 Material

3.1.1 Laboratory equipment

Device, Type	Company, Headquarters
Analog rotator, RS-RD 5	Pheonix Instrument GmbH, Garbsen, Germany
Autoclave Fedegari FVA2/A1	ibs/tecnomara GmbH, Fernwald, Germany
Automatic Benchtop Tissue Processor, Leica TP1020, semi-enclosed	Leica Biosystems Nussloch GmbH, Nussloch, Germany
Beaker, 500 mL, 1 L, Duran	Schott AG, Mainz/DWK Life Sciences GmbH, Wertheim, Germany
Biological safety cabinet, class II, Thermo Scientific Safe 2020	Thermo Electron LED GmbH, Langenselbold, Germany
Cellometer Cell Counter Auto 1000	Nexcelom Bioscience LLC, Lawrence, MA, USA
Centrifuge, Heraeus Megafuge 1	Thermo Fisher Scientific Waltham, MA, USA
CO_2 incubator, HERAcell 240i	Thermo Fisher Scientific, Waltham, MA, USA
Cryo boxes with grid inserts, Labsolute	Th. Geyer GmbH & Co. KG, Renningen, Germany
Dewar flask for liquid nitrogen	KGW-Isotherm/Karlsruher Gastechnisches Werk – Schieder GmbH, Karlsruhe, Germany
Dry bath incubator, BSH 5002-E	Benchmark Scientific Inc., Sayreville, NJ, USA
Flattening table for histopathology, HI 1220	Leica Biosystems Nussloch GmbH, Nussloch, Germany
Fluorescence microscope, Leica DMi8	Leica Microsystems CMS GmbH, Wetzlar, Germany
Fluorometer, Qubit 3.0	Life Technologies/Thermo Fisher Scientific, Waltham, MA, USA
Fridge and freezer, Mediline, Profi line, Glass line	Liebherr-International Deutschland GmbH, Biberach an der Riß, Germany
Glass bottles, 100 mL, 250 mL, 500 mL, 1 L, Duran	Schott AG, Mainz/DWK Life Sciences GmbH, Wertheim, Germany
Graduated cylinders, 500 mL, 1 L, Silberbrand Eterna	Brand GmbH + Co. KG, Wertheim, Germany
GridION Mk1/X5	Oxford Nanopore Technologies Ltd., Oxford, UK

Hybridizer, ThermoBrite StatSpin	Abbott Molecular, Des Plaines, IL, USA
Ice machine, CoolNat, S-Nr. 202086	Ziegra Eismaschinen GmbH, Isernhagen, Germany
Immunostaining slide rack and slides, Thermo Scientific Shandon Sequenza	Thermo Shandon LTD./Thermo Fisher Scientific, Waltham, MA, USA
Incubator, UF260	Memmert GmbH & Co. KG, Schwabach, Germany
Incubator, UN55pa	Memmert GmbH & Co. KG, Schwabach, Germany
Magnet, DynaMag-2	Invitrogen/Thermo Fisher Scientific, Waltham, MA, USA
Microscope, Olympus BX43F	Olympus Corp., Tokyo, Japan
Microscope, Olympus CKX53	Olympus Corp., Tokyo, Japan
Microtome, Leica RM2265	Leica Biosystems Nussloch GmbH, Nussloch, Germany
Microwave, NN-E201WM	Panasonic Marketing Europe GmbH, Hamburg, Germany
Micro scales, AC1215, max. 120 g	Sartorius AG, Goettingen, Germany
Minicentrifuge, Sprout	Heathrow Scientific, Vernon Hills, IL, USA
MinION Mk1B (MIN-101B)	Oxford Nanopore Technologies Ltd., Oxford, UK
Multi dispenser, Multipette plus	Eppendorf AG, Hamburg, Germany
Nanophotometer, P330	Intas Science Imaging, Göttingen, Germany
Paraffin embedding module, EG1150 H, HistoCore Arcadia C	Leica Biosystems Nussloch GmbH, Nussloch, Germany
pH meter, FiveEasy Plus	Mettler Toledo AG, Schwerzenbach, Switzerland
Pipets, ErgoOne, 10, 20, 200, 1000 µL	Starlab International GmbH, Hamburg, Germany
Pipets, Research plus, 2.5, 100 µL	Eppendorf AG, Hamburg, Germany
Pipette filler, pipetus akku	Hirschmann Laborgeräte GmbH & Co. KG, Eberstadt, Germany
Precision scales, EW, max. 2200 g	Kern & Sohn GmbH, Balingen, Germany
Reaction tube racks, 80 wells for 1.5, 2 mL tubes, polypropylene	Th. Geyer GmbH & Co. KG, Renningen, Germany
Refridgerated microcentrifuge, Perfect Spin 24R, Peqlab	VWR International, LLC/Avantor, Inc., Radnor, PA, USA
Staining troughs, type Schiefferdecker, chalk biocarbonate glass	Th. Geyer GmbH & Co. KG, Renningen, Germany
ThermoBrite Humidity Control Cards	Abbott Molecular, Des Plaines, IL, USA
ThermoMixer compact	Eppendorf AG, Hamburg, Germany
Timer, WB-388	Oregon Scientific, Gennevilliers, France
Tissue homogenizer, Precellys 24	Bertin Technologies, Montigny-le-bretonneux, Ile-de-France, France
Tweezers and dissecting scissors	Karl Hammacher GmbH, Solingen, Germany

Ultrapure water system, Arium pro	Sartorius AG, Goettingen, Germany
Virtual Slide Microscope, VS120	Olympus Corp., Tokyo, Japan
Vortexer, REAX2000	Heidolph Instruments GmbH & Co. KG, Schwabach, Germany
Vortexer, RS-VA 10	Pheonix Instrument GmbH, Garbsen, Germany
Water bath, WNB 14	Memmert GmbH & Co. KG, Schwabach, Germany

3.1.2 Consumables

Item, Type	Company, Headquarters
Adhesion slides for IHC, SuperFrost Plus	Gerhard Menzel B.V. & Co. KG, Braunschweig, Germany
Bead Tubes	Invitrogen/Thermo Fisher Scientific, Waltham, MA, USA
Bed Mat, 60 x 90 cm, MoliCare Premium	Paul Hartmann AG, Heidenheim, Germany
Cell culture dish, 100 x 22 mm, Labsolute	Th. Geyer GmbH & Co. KG, Renningen, Germany
Cellometer Cell Counting Chambers, SD1000	Nexcelom Bioscience LLC, Lawrence, MA, USA
CombiTips advanced, 5 mL, sterile	Eppendorf AG, Hamburg, Germany
CryoPure tube, 1.6 mL, red	Sarstedt AG & Co. KG, Nuembrecht, Germany
Embedding cassettes	Th. Geyer GmbH & Co. KG, Renningen, Germany
Feather disposable scalpel, No. 15, No. 21, sterile	Feather Safety Razor Co. LTD., Osaka, Japan
Filter tips, TipOne, 10, 20, 100, 1000 µL	Starlab International GmbH, Hamburg, Germany
Flexible mounting adhesive, Fixogum	Marabu GmbH & Co. KG, Tamm, Germany
Flongle Adapter, ADP-FLG001	Oxford Nanopore Technologies Ltd., Oxford, UK
Flongle Flow Cell, Pk. 1, FLO-FLG001	Oxford Nanopore Technologies Ltd., Oxford, UK
Folded filters, 185 mm	Schleicher & Schuell BioScience GmbH, Dassel, Germany
Gloves, latex/nitrile, Starguard	Starlab International GmbH, Hamburg, Germany
Kimwipes, Delicate Task Wipes 7558, white, 20.5 x 20 cm	Kimberly-Clark Professional, Dallas, TX, USA
Microscope cover slips, 24 x 32 mm	Th. Geyer GmbH & Co. KG, Renningen, Germany
Microtome blades, Feather S35	pfm medical AG, Cologne, Germany

Microtube 1.5, 2 mL DNA LowBind	Sarstedt AG & Co. KG, Nuembrecht, Germany
Microtube, 0.5, 1.5, 2 mL	Sarstedt AG & Co. KG, Nuembrecht, Germany
Needles, Microlance 3, 0.9 x 40 mm, 20 G x 1 ½ ‘‘, Nr. 1	Becton Dickinson S.A., Fraga, Spain
Needles, Sterican, 0.45 x 12 mm, 26 G x ½ ‘‘	B. Braun Melsungen AG, Melsungen, Germany
PCR tubes, 200 µL, Multiply µStrip Pro	Sarstedt AG & Co. KG, Nuembrecht, Germany
Pipette tips, TipOne, 10, 200, 1000 µL	Starlab International GmbH, Hamburg, Germany
Qubit Assay Tubes	Invitrogen/Thermo Fisher Scientific, Waltham, MA, USA
Rolled edge glass vials with polyethylene lids, 5 mL	Th. Geyer GmbH & Co. KG, Renningen, Germany
Serological pipette, 2, 25, 50 mL	Sarstedt AG & Co. Nuembrecht, Germany
Serological pipette, 5, 10 mL	Greiner Bio-One GmbH, Frickenhausen, Germany
Single use filter unit, 0.2 µm, sterile, Labsolute	Th. Geyer GmbH & Co. KG, Renningen, Germany
Spot-ON Flow Cell, R9 version (R9.4.1)	Oxford Nanopore Technologies Ltd., Oxford, UK
TC Dish 100, Standard, sterile	Sarstedt AG & Co. KG, Nuembrecht, Germany
TC flask T75	Sarstedt AG & Co. KG, Nuembrecht, Germany
Tube, 5, 15, 50 mL	Sarstedt AG & Co. KG, Nuembrecht, Germany
Weighing boats	Lab Logistics Group GmbH, Meckenheim, Germany

3.1.3 Chemicals, reagents, and enzymes

Chemical/Reagent/Enzyme, Type	**Company, Headquarters**
2-Propanol, z. A., 2.5 L, Chemsolute	Th. Geyer GmbH & Co. KG, Renningen, Germany
5X NEBNext Quick Ligation Reaction Buffer	New England Biolabs, Ipswich, MA, USA
Acetic acid, 100% p.A., 2.5 L	PanReac AppliChem GmbH/ITW Reagents, Darmstadt, Germany
Actinomycin D-Mannitol	Sigma-Aldrich/Merck KGaA, Darmstadt, Germany

Albumin Bovine Fraction V, pH 7.0	Serva Electrophoresis GmbH, Heidelberg, Germany
Ammonium Acetate (Emsure, ACS, Reag. Ph Eur)	Sigma-Aldrich/Merck KGaA, Darmstadt, Germany
AMPure XP beads	Beckman Coulter Inc., Brea, CA, USA
Ampuwa, water for injection purposes, 1000 mL	Fresenius Kabi Deutschland GmbH, Bad Homburg, Germany
Blunt/TA Ligase Master Mix	New England Biolabs, Ipswich, MA, USA
Citric acid monohydrate, p.a.	Carl Roth GmbH + Co. KG, Karlsruhe, Germany
Cytoseal 60	Thermo Fisher Scientific, Waltham, MA, USA
DAPI, 4',6-Diamidino-2-phenyl-indol-dihydrochlorid, 5 mg	Sigma-Aldrich/Merck KGaA, Darmstadt, Germany
Desinfectant, Desomed Rapid AF	Desomed Dr. Trippen GmbH, Freiburg, Germany
Dulbecco's Modified Eagle Medium (DMEM), gibco	Life Technologies/Thermo Fisher Scientific, Waltham, MA, USA
Dulbecco's Phosphate Buffered Saline (PBS), gibco	Life Technologies/Thermo Fisher Scientific, Waltham, MA, USA
Embedding wax (Paraffin)	Engelbrecht GmbH, Edermünde/Besse, Germany
Eosin Y solution, aqueous, 1 L	Sigma-Aldrich/Merck KGaA, Darmstadt, Germany
Ethanol, 99%, denatured/absolute p. a., Chemsolute	Th. Geyer GmbH & Co. KG, Renningen, Germany
Ethylenediaminetetraacetic acid tetra-sodium salt dihydrate	Sigma-Aldrich/Merck KGaA, Darmstadt, Germany
Fetal bovine serum (FBS), gibco	Life Technologies/Thermo Fisher Scientific, Waltham, MA, USA
Formaldehyde solution, 4.5%, phosphate buffered, Chemsolute	Th. Geyer GmbH & Co. KG, Renningen, Germany
Hematoxylin solution according to Mayer, 1 L	Merck & Co., Inc., Kenilworth, NJ, USA
Hydrochloric acid, 2 N	Carl Roth GmbH + Co. KG, Karlsruhe, Germany
Hydrogen Peroxide, 30% p.a., Rotipuran	Carl Roth GmbH & Co. KG, Karlsruhe, Germany
Isoflurane CP 1 mL/mL	CP-Pharma Handelsgesellschaft mbH, Burgdorf, Germany
LongAmp Hot Start Taq 2X Master Mix	New England Biolabs, Ipswich, MA, USA
NEBNext FFPE DNA Repair Buffer	New England Biolabs, Ipswich, MA, USA
NEBNext FFPE DNA Repair Mix	New England Biolabs, Ipswich, MA, USA
Nuclease free water, 2 mL	Integrated DNA Technologies, Inc. (IDT), Coralville, IA, USA
Nuclease-free water, 25 mL	New England Biolabs, Ipswich, MA, USA

PBS Dulbecco, powder	Biochrom GmbH, Berlin, Germany
Penicillin-Streptomycin solution 100X, 100 mL	Biowest LLC, Riverside, MO, USA
Proteinase K solution, 20 mg/mL, 10 mL	PanReac AppliChem GmbH/ITW Reagents, Darmstadt, Germany
Quick T4 DNA Ligase	New England Biolabs, Ipswich, MA, USA
Ribonuclease, >70 U/mg, salt-free, protease-free, 100 mg	Carl Roth GmbH & Co. KG, Karlsruhe, Germany
RPMI 1640 Medium, gibco	Life Technologies/Thermo Fisher Scientific, Waltham, MA, USA
Sodium chloride, Fluka	Honeywell International Inc., Charlotte, NC, USA
Tris, Tris-HCl, p. a., Pufferan	Carl Roth GmbH + Co. KG, Karlsruhe, Germany
Tris-EDTA buffer solution, BioUltra for molecular biology, pH 7.4	Sigma-Aldrich/Merck KGaA, Darmstadt, Germany
Trypsin-EDTA 0.5% 10X, gibco	Life Technologies/Thermo Fisher Scientific, Waltham, MA, USA
Tween 20	PanReac AppliChem GmbH/ITW Reagents, Darmstadt, Germany
Ultra II End-prep enzyme mix	New England Biolabs, Ipswich, MA, USA
Ultra II End-prep reaction buffer	New England Biolabs, Ipswich, MA, USA
Vectashield Antifade Mounting Medium	Biozol Diagnostica Vertrieb GmbH, Eching, Germany
Xylene, 2.5 L, J. T. Baker	Avantor, Inc., Radnor, PA, USA

3.1.4 Laboratory-made buffers

Buffer	**Composition**
Citrate buffer, pH 6.0	10 mM Citric acid monohydrate in aqua bi-dest
Elution buffer (16S rRNA sequencing)	10 mM Tris-HCl, pH 8.0 50 mM NaCl in aqua bi-dest
TBS buffer, pH 7.4	50 mM Tris-HCl 150 mM NaCl in aqua bi-dest
TBS-T buffer	0.1% Tween 20 in TBS buffer
TE buffer, pH 9.0	1 mM EDTA 1 mM Tris-HCl in aqua bi-dest

3.1.5 Antibodies and probes

Antibody/Probe, Specification	Company, Headquarters
Cytokeratin 5, clone XM26, monoclonal, mouse, IgG1 κ, NCL-L-CK5	Leica Biosystems, Nussloch, Germany
EUB338, 5′- GCT GCC TCC CGT AGG AGT -3′, 5′ and 3′ modification: Cyanine 3, 100 pmol/µL	biomers.net GmbH, Ulm, Germany
GATA6, polyclonal, goat, IgG, AF1700	R&D Systems, Minneapolis, MN, USA
Lipopolysaccharide (LPS), polyclonal, rabbit, IgG, PAB526Ge01	Cloud-Clone Corp., Katy, TX, USA
Non-specific complement probe, 5′- CGA CGG AGG GCA TCC TCA -3′, 5′ and 3′ modification: Cyanine 3, 100 pmol/µL	biomers.net GmbH, Ulm, Germany

3.1.6 Kits

Kit, Specification	Company, Headquarters
16S Barcoding Kit 1-24 (SQK-16S024)	Oxford Nanopore Technologies Ltd., Oxford, UK
Flow Cell Priming Kit (EXP-FLP002)	Oxford Nanopore Technologies Ltd., Oxford, UK
ImmPACT DAB Substrate Kit, Peroxidase (HRP)	Vector Laboratories, Inc., Burlingame, CA, USA
Ligation Sequencing Kit (SQK-LSK109)	Oxford Nanopore Technologies Ltd., Oxford, UK
Native Barcoding Expansion 1-12 (EXP-NBD104)	Oxford Nanopore Technologies Ltd., Oxford, UK
Native Barcoding Expansion 13-24 (EXP-NBD114)	Oxford Nanopore Technologies Ltd., Oxford, UK
OneStep PCR Inhibitor Removal Kit	ZymoResearch, Irvine, CA, USA
PureLink Microbiome DNA Purification Kit	Invitrogen/Thermo Fisher Scientific, Waltham, MA, USA
QIAamp DNA Investigator Kit	Qiagen, Venlo, the Netherlands
Qubit dsDNA HS Assay Kit	Invitrogen/Thermo Fisher Scientific, Waltham, MA, USA
RecoverAll Total Nucleic Acid Isolation Kit	Invitrogen/Thermo Fisher Scientific, Waltham, MA, USA
Vectastain ABC Kit, peroxidase (Rabbit IgG) (PK-4001) (with normal goat serum)	Vector Laboratories, Inc., Burlingame, CA, USA
Vectastain ABC Kit, peroxidase, Goat IgG (PK-4005) (with normal rabbit serum)	Vector Laboratories, Inc., Burlingame, CA, USA

Vectastain ABC Kit, peroxidase, Mouse IgG (PK-4002) (with normal horse serum)	Vector Laboratories, Inc., Burlingame, CA, USA
Zyto*Light* FISH-Tissue Implementation Kit	ZytoVision GmbH, Bremerhaven, Germany

3.1.7 Software and data bases

Software	**Developer/Company, Headquarters**
Google Scholar	Google LLC, Mountain View, CA, USA
GraphPad Prism 9.4.0.673	GraphPad Software Inc., La Jolla, CA, USA
ImageJ 1.50b open-source software	Wayne Rasband, Research Services Branch, National Institutes of Health, Bethesda, MD, USA
Leica LAS X – Application Suite X, Imaging and Analysis Software	Leica Microsystems CMC GmbH, Wetzlar, Germany
Microsoft Word/Excel/PowerPoint for Microsoft 365 MSO (Version 2208, Build 16.0.15601.20660), 32 Bit	Microsoft Corp., Redmond, WA, USA
MinKNOW software v. 22.08.09 with Guppy toolkit 6.2.11	Oxford Nanopore Technologies Ltd., Oxford, UK
Olympus cellSens Entry v2.3 (Build 18987)	Olympus Corp., Tokyo, Japan
PubMed	National Institute of Health (NIH)/National Library of Medicine/National Center for Biotechnology Information (NCBI), Bethesda, MD, USA
R, RStudio v2.4 and later	Posit PBC, Boston, MA, USA
Web of Science	Clarivate Analytics, Philadelphia, PA, USA

3.2 Methods

3.2.1 Animal studies

All animal experiments were performed in accordance with national and institutional laws and regulations. They were conducted in the "Zentrale Tierexperimentelle Einrichtung" (ZTE) at the University Medical Center Göttingen (UMG), Göttingen, Germany, after approval through the "Landesamt für Verbraucherschutz und Lebensmittelsicherheit" (LAVES) of Lower Saxony, Germany. The experiments involving KPC mice run under the application number 19/3085 and were entirely conducted within the scope of this thesis. Studies on the pancreatic microbiome of NMRI-*Foxn1*$^{nu/nu}$ (nude) mice orthotopically transplanted with human pancreatic tumor cells were performed with archived formalin-fixed paraffin-embedded (FFPE) tissue samples. The transplantation experiments run under the application number 15/2057 and were conducted by Dr. Mengyu Tu and Lukas Klein, working group of Dr. Shiv Singh, Department of Gastroenterology, Gastrointestinal Oncology and Endocrinology, UMG, Göttingen, Germany.

Mice were housed under constant conditions in the ZTE, including a light-dark-cycle of 12 h/12 h, 23 ± 1 °C room temperature (RT), and 40-60% humidity. They were kept in individually ventilated cages with up to five animals of the same sex per cage and were fed regular vegetarian mouse chow together with normal tap water. Animal care takers and scientists handling the mice were both male and female.

KPC mice were bred in the Animal Facility of the European Neuroscience Institute Göttingen, Göttingen, Germany, and transferred to the ZTE at the age of approximately 5 weeks after weening and genotyping. Control (CTRL) mice of either *LSL-Kras*$^{G12D/+}$;*LSL-Trp53*$^{R172H/+}$ or *Pdx-1-Cre* genotype were taken from the same litters. General termination criteria at which the experiments had to be discontinued by sacrificing the animals comprise overall morbidity, lethargy, signs of pain, rough fur, loss of self-care and social behavior, ascites, and cachexia as manifested by a body weight loss of 20%. Respecting these criteria, the end point (EP) was defined as 6 months or when tumor-bearing for CTRL and KPC mice, respectively. Subsequently, the EP varied for KPC mice. For each batch, the age range is indicated separately in the respective results section.

The orthotopic transplantation of immortalized human PDAC cell lines Capan-2, MiaPaCa2, and Capan-1, the treatment of the Capan-1-transplanted animals with TNFα and H_2O, respectively, and the sacrifice of the animals including tumor isolation was performed by Dr. Mengyu Tu and Lukas Klein as previously published (Tu et al., 2021). The FFPE tumor samples were received and further processed as described in the following.

3.2.2 Sample acquisition, processing, and storage

Four different types of samples were used in this thesis: snap-frozen fecal samples of KPC and CTRL mice, snap-frozen pancreatic samples (bulk tumor and healthy control pancreas, tumor center and tumor periphery) of KPC and CTRL mice, FFPE pancreatic samples (tumor with adjacent normal pancreas, TNFα- and H_2O-treated tumor) from NMRI-*Foxn1*$^{nu/nu}$ mice orthotopically transplanted with different human PDAC cell lines, and FFPE pancreatic samples (tumor and healthy control pancreas) from human PDAC patients.

Feces were collected from KPC and CTRL mice at the age of 6 weeks (6w), 12 weeks (12w), and the EP. Therefore, the animals were placed separately in clean cages without embedding under the hood until defecation. EP feces were directly taken from the colon at sacrifice. Importantly, only feces that were not in contact with urine were collected in cryovials and subsequently snap-frozen in liquid nitrogen. The samples were stored at -80 °C until further use.

Tumor and healthy pancreas from KPC and CTRL mice were sampled at EP sacrifice. After isoflurane anesthesia and cervical dislocation, organs were harvested from the animals and either snap-frozen or formalin-fixed. Regarding the pancreas, one part was transferred to cryovials and snap-frozen in liquid nitrogen for subsequent DNA extraction. These vials were then stored at -80 °C until further use. The other part was fixed in 4.5% formaldehyde (formalin) for at least 24 h and subsequently dehydrated in an ethanol (EtOH) and xylene series of increasing percentage. This dehydration process was performed automatically overnight in a benchtop tissue processor. The program is shown in **table 1**. Eventually, these samples were embedded in paraffin and stored in blocks at RT protected from light until further use.

Table 1: Automated overnight program of the semi-enclosed Leica TP1020 benchtop rotary tissue processor for tissue dehydration.

Solution	Time [hh:mm]
Formalin	01:15
55% EtOH	00:30
85% EtOH	00:45
96% EtOH	01:00
99% EtOH 1	01:15
99% EtOH 2	01:10
99% EtOH 3	01:30
Xylene 1	00:20
Xylene 2	00:30
Xylene 3	01:10
Liquid paraffin 1	00:30
Liquid paraffin 2	00:45
Liquid paraffin 3	01:30

Large tumors of KPC mice were cut in half and used specifically to study the microbiome of tumor center and periphery if the tumor showed a fibrotic and necrotic whitish center macroscopically clearly distinguishable from the softer pinkish periphery. Of one tumor half, tumor center and periphery were carefully separated with a sterile scalpel in a sterile petri dish and transferred into two separate cryovials, snap-frozen in liquid nitrogen and stored at -80 °C until further use. The other tumor half was formalin-fixed and paraffin-embedded as described above with the sectional plane horizontally embedded in order to obtain full cross sections of the entire tumor. The paraffin blocks were stored at RT protected from light until further use.

Tumor samples from NMRI-*Foxn*$^{nu/nu}$ mice orthotopically transplanted with human PDAC cell lines Capan-2 and MiaPaCa2, as well as Capan-1 with additional TNFα or H_2O treatment, respectively, were directly obtained as FFPE samples in paraffin blocks. All preceding work was done and the samples kindly provided by Dr. Mengyu Tu and Lukas Klein (Tu et al., 2021).

Human PDAC and corresponding healthy pancreas from the same patients were kindly provided as FFPE samples in separate paraffin blocks by Prof. Dr. Philip Ströbel and Dr. Hanibal Bohnenberger from the Institute of Pathology, UMG, Göttingen, Germany. The samples were taken from patients at surgery by Prof. Dr. Michael Ghadimi, Prof. Dr. Jochen

Gaedcke and Prof. Jens Jakob, Department of General, Visceral and Pediatric Surgery, UMG, Göttingen, Germany. Sampling was approved by the ethics committee (application number 5/11/19) within the scope of the P-MAPS study (NCT04777812).

Required FFPE samples were cut at 4 µm thickness using a microtome. The sections were dried at 37 °C overnight and subsequently subject to various staining methods described in the following.

3.2.3 Hematoxylin and eosin staining

To perform hematoxylin and eosin (H&E) staining, the tissue sections were first subject to paraffin removal with xylene (**table 2**). Secondly, rehydration of the tissue was performed in an EtOH series of decreasing percentages (**table 2**). Then, the slides were quickly washed in aqua bidest followed by 6 min incubation in double-filtered hematoxylin. From there, they were transferred to 0.2% acetic acid for 15 s. Subsequently, the sections were placed under running tap water for about 6 min until the hematoxylin turned blue. Next, they were incubated in eosin for 3 min, then washed again in tap water for 30 s. Afterwards, the tissue was again dehydrated by passing an EtOH series of increasing percentages followed by incubation in xylene (**table 2**).

Table 2: Paraffin removal, tissue rehydration and dehydration of tissue sections for hematoxylin and eosin staining.

Step	Solution	Time [mm:ss]
Paraffin removal	Xylene 1	10:00
	Xylene 2	10:00
Tissue rehydration	99% EtOH	01:00
	96% EtOH	01:00
	70% EtOH	01:00
Tissue dehydration	70% EtOH	00:30
	96% EtOH	01:30
	99% EtOH	02:00
EtOH removal	Xylene 1	10:00
	Xylene 2	10:00

Finally, the tissue sections were mounted with Cytoseal 60 mounting medium and a cover slip of adequate size. The slides were left to set under the hood overnight before prolonged microscoping and image acquisition using an Olympus light microscope with camera controlled by the Olympus cellSens Entry software v2.3.

3.2.4 Immunohistochemistry

Immunohistochemistry (IHC) was performed on murine and human pancreatic tissue sections targeting LPS, GATA binding protein 6 (GATA6), and cytokeratin 5 (CK5). For all three targets, the basic protocol was the same. Differences are listed in **table 3**. Initially, the tissue sections were subject to paraffin removal and tissue rehydration for which solutions and times are indicated in **table 4**. Then, the slides were quickly washed in tap water. Next, epitope unmasking was performed by heat in the respective buffer (**table 3**). Therefore, the slides were cooked for approximately 15 min in the buffer in a microwave with a subsequent cool down for 10 min inside the microwave. Eventually, the cuvette containing buffer and slides was transferred to ice for further cool down. Thereafter, the slides were rinsed three times in tap water before immersing them in freshly prepared 3% H_2O_2 for 10 min at RT to quench endogenous peroxidases. Again, they were briefly rinsed three times in tap before aligning them in a Shandon Sequenza immunostaining slide rack with matching plastic cover plate assemblies. The formation of air bubbles between slide and cover plate was explicitly precluded. Next, the slides were washed three times with Tris-buffered saline (TBS) with 0.1% Tween 20 (TBS-T) wash buffer. Subsequently, unspecific antigen blockage was carried out by adding 200 µL TBS-T with 10% normal serum (**table 3**) and 1% bovine serum albumin (BSA) to each slide and incubating it for 1 h at RT. Slides were again rinsed once with TBS-T wash buffer before 200 µL of the respective antibody in the respective dilution (**table 3**) in TBS-T with 1% BSA was added and incubated at 4 °C overnight. The next day, the slides were washed three times with TBS-T before 200 µL of the respective secondary antibody (**table 3**) from the respective Vectastain ABC Kit diluted 1:200 in TBS-T with 1% BSA were added to each slide. Subsequently, the tissue sections were incubated for 1 h at 37 °C. Meanwhile, the AB complex to enhance the antibody signal was made from 10 µL of each reagent A and B from the same Vectastain kit and 1 mL TBS-T. The

components were merged in a reaction tube which was then carefully inverted. The mix was incubated for 30 min at RT. Finally, the tissue sections were washed again three times with TBS-T before 200 µL of the AB complex were added per slide. After another 1 h incubation at RT, the slides were once more subject to three TBS-T washes before being disassembled from the cover plate assemblies in the slide rack and quickly washed in tap water three times. Diaminobenzidine (DAB) working solution was prepared from 30 µL DAB chromogen per 1 mL DAB diluent from the ImmPACT DAB Substrate Kit, Peroxidase (HRP) and mixed well by pipetting. Approximately 100 µL DAB working solution or a sufficient volume to cover the entire tissue section was directly added on top of each slide retrieved from the tap water rinse. The DAB working solution was incubated on the tissue to develop for the respective time indicated in **table 3** before the reaction was stopped by placing the slides in tap water.

Table 3: Specification of parameters for antibodies against the respective IHC targets.

Target	Buffer	Serum	Primary antibody dilution	Secondary antibody	Development time [mm:ss]
LPS	citrate	goat	1:50	anti-rabbit	04:00
GATA6	citrate	rabbit	1:350	anti-goat	03:30
CK5	TE	horse	1:750	anti-mouse	02:00

Eventually, counterstaining was performed by 3 min incubation in double-filtered hematoxylin. The sections were placed under running tap water for about 6 min until the hematoxylin turned blue followed by dehydration of the tissue sections in an EtOH series with increasing concentrations and xylene incubation (**table 4**). Lastly, the tissue sections were mounted with Cytoseal 60 mounting medium and a cover slip of adequate size. The slides were left to set under the hood overnight before prolonged microscoping and image acquisition using an Olympus light microscope with camera controlled by the Olympus cellSens Entry software v2.3. Quantification was carried out using a color deconvolution macro in ImageJ 1.50b open-source software. Plotting and statistical evaluation were performed in GraphPad Prism 9.4.0.673, the latter by checking for normal distribution using the Shapiro-Wilk test and subsequently calculating the appropriate significance test. GATA6 and CK5 stainings were imaged with an Olympus whole slide scanner for overview scans. After excluding staining artifacts, positive-stained cells were quantified using the QuPath

v0.4.3 software for subtyping. Z-scores were calculated from the percentages. The CLA, HYB and BL subtypes were discriminated by dividing the z-scores in three quantiles and comparing those of GATA6 and CK5. The subtyping results were entered in the metadata of the human FFPE tumor samples sequencing data.

Table 4: Paraffin removal, tissue rehydration and dehydration of tissue sections for immunohistochemistry.

Step	Solution	Time [mm:ss]
Paraffin removal	Xylene 1	10:00
	Xylene 2	10:00
Tissue rehydration	99% EtOH 1	04:00
	99% EtOH 2	04:00
	96% EtOH	04:00
	80% EtOH	04:00
	70% EtOH	04:00
Tissue dehydration	70% EtOH	03:00
	80% EtOH	03:00
	96% EtOH	03:00
	99% EtOH	03:00
EtOH removal	Xylene 1	05:00
	Xylene 2	05:00
	Xylene 3	05:00
	Xylene 4	05:00

3.2.5 Fluorescence-*in situ*-hybridization

To visualize the bacterial 16S rRNA gene fluorescence-*in situ*-hybridization (FISH) was conducted. Therefore, the Zyto*Light* FISH-Tissue Implementation Kit was used according to the manufacturer's instructions with minor modifications. Two slides of each sample were processed in order to stain one with the target probe and the other with the control probe. To dewax, the tissue slides were incubated in a hybridizer for 10 min at 70 °C in addition to subsequent xylene incubation (**table 5**). Next, the tissue was rehydrated in an EtOH series of decreasing concentrations (**table 5**). It was then washed twice in aqua bidest for 2 min and subsequently incubated for 15 min in the kit's Heat Pretreatment Solution Citric preheated to 98 °C. Again, the slides were transferred to aqua bidest and washed twice for 2 min. Using a clean Kimwipe, the slides were carefully dried and transferred to the hybridizer. Proteolysis

was performed by applying three drops of the kit's Pepsin Solution directly to each tissue section and incubating them for 10 min at 37 °C. Afterwards, the slides were washed in the kit's Wash Buffer SSC for 5 min and in aqua bidest for 1 min. Next, the tissue was subject to dehydration via an EtOH series (**table 5**). Eventually, the slides were air-dried before 3 µL of the respective probe, EUB338 or non-specific complement probe (nonEUB338), were applied directly onto the tissue sections. The probe sequences had been published previously (Amann et al., 1990; Prudent and Raoult, 2019; Wallner et al., 1993). To spread out the probe evenly, the tissue sections were covered with cover slips of adequate size without introducing air bubbles, and the edges were sealed with Fixogum. Then, the slides were transferred to the hybridizer again and incubated for 10 min at 75 °C to denature the DNA. Thereafter, wet stripes were placed in the hybridizer to keep the interior moist and prevent the slides from drying out. Finally, the slides were incubated at 37 °C overnight to hybridize the probes to the DNA. The next day, the slides were degummed and transferred to the kit's 1X Wash Buffer A pre-warmed to 37 °C. The wash buffer-containing cuvette was carefully shaken after approximately 2 min to facilitate the cover slip to come off. Then, the slides were again washed twice for 5 min in 1X Wash Buffer A pre-warmed to 37 °C. Subsequently, the tissue was dehydrated according to the EtOH series in **table 5** and air-dried protected from light.

Table 5: Paraffin removal, tissue rehydration and dehydration of tissue sections for fluorescence-*in situ*-hybridization.

Step	Solution	Time [mm:ss]
Paraffin removal	Xylene 1	05:00
	Xylene 2	05:00
Tissue rehydration	99% EtOH 1	05:00
	99% EtOH 2	05:00
	90% EtOH	02:00
	70% EtOH	02:00
Tissue dehydration	70% EtOH	01:00
	90% EtOH	01:00
	99% EtOH	01:00

Finally, the tissue sections were mounted with one drop of Vectashield Antifade Mounting Medium supplemented with Actinomycin D-Mannitol (1:20) and DAPI (1:1000) and a cover slip of adequate size. The slides were kept at 4 °C protected from light and images were

acquired using a Leica fluorescence microscope with camera controlled by the LAS X software (n = 4 conditions, n = 10 slides per condition, n = 5 images per slide). Quantification was carried out manually by counting light signals in a total of 200 image fields (200x magnification). Plotting and statistical evaluation were performed in GraphPad Prism 9.4.0.673, the latter by checking for normal distribution using the Shapiro-Wilk test and subsequently calculating the appropriate significance test.

3.2.6 DNA extraction from snap-frozen pancreatic bulk tissue (Qiagen)

To establish the most efficient and reliable microbial DNA extraction method for the different sample types, particularly the challenging pancreatic tumor tissue with its high fibrotic content and low microbial biomass, two different DNA extraction kits and protocols were tested. First, the Qiagen QIAamp DNA Investigator Kit with its protocol "Isolation of Total DNA from Tissues" was assessed. In addition to the establishment sample batch, a negative control (NTC) was introduced to control for environmental contamination using only consumables, buffers, and reagents but no true sample material (buffer control). Other than that, it was treated the same way. The deep-frozen tissue samples were placed on a sterile petri dish on dry ice to avoid thawing of the entire sample. Using sterile scalpels, small pieces were cut off, transferred to a 1.5 mL reaction tube and 180 µL of the kit's Buffer ATL were added immediately. Additionally, 20 µL proteinase K from the kit were added and the tube mixed in pulses with the help of a vortex for approximately 15 s. Then, the tubes were put on a thermomixer set to 56 °C and 650 rpm overnight for most efficient tissue lysis. The next day, the tubes were briefly centrifuged to collect the liquid drops in the lid before 200 µL of the kit's Buffer AL were added to each tube. These were then pulse-vortexed for approximately 15 s. Next, 200 µL 99% EtOH were added, the tubes again quickly pulse-vortexed and incubated for 5 min at RT. Subsequently, they were briefly spun down to collect the lysate before it was transferred to a QIAamp MinElute column in a collection tube. In case there was a viscous part of the lysate remaining in the tube, it was discarded in order to prevent obstruction of the column. The column was centrifuged at 6,000 × g for 1 min. After this step, the collection tube was exchanged and the flow-through discarded. Next, the membrane-bound DNA was washed several times. First, 500 µL of the kit's Buffer AW1 were

added to the column which was again centrifuged at 6,000 × g for 1 min and assembled with a new collection tube. This step was repeated with 700 µL Buffer AW2 from the kit and with 700 µL 99% EtOH. Eventually, the column was once more centrifuged at 15,000 × g for 3 min to free the membrane from residual EtOH. Then, the column was transferred to a new 1.5 mL reaction tube and further dried for 3 min at 56 °C with the lid open to allow residual EtOH to evaporate. Finally, 50 µL of the kit's Buffer ATE were directly applied to the membrane, incubated for 5 min at RT and the column centrifuged at 14,000 × g for 1 min. The eluted DNA was stored at -20 °C until further use.

3.2.7 DNA extraction from snap-frozen pancreatic bulk tissue and feces (Invitrogen)

The second kit tested for consistent DNA extraction was the Invitrogen PureLink Microbiome DNA Purification Kit. The user manual was adjusted according to recommendations by the International Human Microbiome Standard (IHMS). The resulting protocol was published and referred to as "Invitrogen IHMS" (Ammer-Herrmenau et al., 2021b). Aside from samples used to establish the method (Results Chapter 1), all snap-frozen tissue and feces samples in this thesis were isolated according to this protocol. With each sample batch to be isolated, at least 1-2 NTCs were introduced to control for environmental contamination using only consumables, buffers, and reagents but no true sample material (buffer controls). Apart from that, they were not treated differently than real samples. To prepare for the protocol, small pieces were cut off the snap-frozen bulk tissue samples on a sterile petri dish on dry ice using sterile scalpels and transferred to 2 mL bead tubes. Feces were directly transferred to bead tubes (2 feces from each sample). To each tube, 600 µL of the kit's Lysis Buffer S1 and 100 µL Lysis Enhancer S2 were added. Then, the tubes were vortexed vigorously before they were incubated on a thermomixer shaking at 850 rpm for 15 min at 95 °C. Thereafter, the sample-containing bead tubes were transferred to a tissue homogenizer and a program at 6500 rpm was run twice with a 5 min break for cool down to RT between both cycles (**table 6**). After this mechanical disruption of the samples, they were centrifuged at 16,000 × g for 5 min at 4 °C. Subsequently, the supernatant was transferred to 2 mL low binding reaction tubes. Then again, 300 µL Lysis Buffer S1 and 50 µL Lysis Enhancer S2 were added to the bead tubes in which residual solid parts of the samples

remained and the same process was repeated. After centrifuging, the supernatant was added to the previous supernatant in the 2 mL low binding tubes from the first lysis round. Next, 260 µL 10 M ammonium acetate were added to these tubes which were then inverted to mix. Subsequently, they were incubated for 10 min on ice followed by centrifugation at 16,000 × g for 10 min at 4 °C. The resulting supernatant was transferred to new 2 mL low binding tubes without disturbing the pellet of debris. To each tube, 250 µL Cleanup Buffer S3 from the kit were added and the tubes were immediately vortexed to ensure a thorough mixing of the contained components. Again, incubation for 10 min on ice and centrifugation at 16,000 × g for 10 min at 4 °C followed before the supernatant was transferred to new 2 mL low binding tubes without touching the pellet. Subsequently, DNA precipitation was induced by adding 1 mL isopropanol, vortexing the tubes thoroughly, and incubating them for at least 30 min on ice. Afterwards, the precipitated DNA was pelleted by centrifugation at 16,000 × g for 15 min at 4 °C. The supernatant was discarded, and the remaining DNA pellet was washed with 300 µL 70% EtOH. To prevent discarding loosened DNA after this washing step, the tubes were again centrifuged at 16,000 × g for 5 min at 4 °C. The EtOH was then discarded. With the lid open, the tubes were turned upside down and placed on clean Kimwipes to dry the DNA pellet for approximately 15 min. Thereafter, the pellet was resuspended in 100 µL sterile filtered Tris-EDTA (TE) buffer. For fecal samples which returned a very firm large pellet, the TE buffer was incubated for 10 min at RT on the DNA pellet to solvate it before resuspending it by pipetting. Next, 2 µL DNase-free RNase (10 mg/mL) were directly added to the resolved DNA and incubated for 30 min at 37 °C. Then, both 15 µL proteinase K and 200 µL of the kit's Binding Buffer S4 were added to the samples, the tubes were vortexed and incubated for 10 min at 65 °C. Then, 200 µL 99% EtOH were added and the tubes again vortexed. Eventually, the sample mix was transferred to a spin column with collection tube from the kit and centrifuged at 14,000 × g for 1 min at 4 °C. Then, the sample was washed again with 200 µL 99% EtOH which was incubated for 3 min before the column was again centrifuged at 14,000 × g for 1 min at 4 °C. Thereafter, the flow-through-containing tube was discarded, and the column placed into a new collection tube. Subsequently, 500 µL of the kit's Wash Buffer S5 were added to the column which was again centrifuged at 14,000 × g for 1 min at 4 °C. The flow-through-containing tube was once more discarded, the column placed into a new collection tube and again centrifuged at 14,000 × g for 1 min at 4 °C to remove residual wash buffer. Finally, the columns were placed in new 1.5 mL low binding

reaction tubes and 55 µL of the kit's Elution Buffer S6 were directly added to the center of each column membrane. After an incubation for 5 min, the DNA was eluted from the membrane via centrifugation at 14,000 × g for 1 min at 4 °C. The samples were stored at -20 °C until further use.

Table 6: Precellys 24 Tissue Homogenizer program for DNA extraction from snap-frozen samples.

Step	Revolutions [rpm]	Time [mm:ss]
Shake	6500	00:30
Pause	0	00:30
Shake	6500	00:30
Pause	0	00:30
Shake	6500	00:30
Pause	0	00:30
Cool down	0	05:00
Shake	6500	00:30
Pause	0	00:30
Shake	6500	00:30
Pause	0	00:30
Shake	6500	00:30
Pause	0	00:30

3.2.8 DNA extraction from FFPE pancreatic tissue

To extract DNA from FFPE orthotopic and human patient pancreatic tissue, the Invitrogen RecoverAll Total Nucleic Acid Isolation Kit was applied with the protocol "RecoverAll Multi-Sample RNA/DNA Workflow" which allows sequential extraction of RNA and DNA from FFPE samples. The manual was followed with some adaptations. However, only the DNA was made use of within the scope of this thesis wherefore only the corresponding parts of the protocol are elucidated in the following. Using H&E stainings of each sample and a light microscope, the area of interest was located and circled with a marker pen on the cover slips. Regarding FFPE orthotopic samples, tumor and adjacent normal tissue were located on the same slide, so both were marked. Concerning FFPE patient samples, tumor and normal pancreatic samples were embedded separately wherefore only on the tumor slides the area of interest had to be marked. This was done by pathologists. From the normal pancreatic samples, the entire section was used. The paraffin blocks were cut at 10 µm and dried overnight at 37 °C. Depending on the size of the tissue area of interest, DNA was isolated

from 2-8 sections per sample. The initial step of the protocol involving paraffin removal was skipped. Regarding NTCs, not only buffer controls but also paraffin controls were introduced since paraffin as supporting material is a considerable source of contamination. The kit's Digestion Buffer was diluted as suggested by mixing 25 µL Digestion Buffer and 75 µL nuclease-free H_2O per sample. 100 µL of this dilution were aliquoted to the corresponding number of 1.5 mL low binding reaction tubes. Then, the 10 µm sections were aligned in registry with their corresponding marked H&E stainings if applicable. With the help of a sterile scalpel which was pre-wetted by dipping it into the diluted Digestion Buffer, the tissue was scratched of the slide. The tissue was collected in a cohesive mass and transferred to the reaction tube containing the Digestion Buffer with the scalpel. For each sample, a new sterile scalpel was used. Paraffin controls were prepared by scratching off paraffin from around the tissue which was exactly treated like the paraffin-embedded true sample material. For each FFPE orthotopic experiment, 6 paraffin controls and 2 buffer controls to control for environmental contamination using only consumables, buffers, and reagents but no paraffin or tissue were introduced. For the FFPE human patient sample experiment, one paraffin control from each sample (n = 24 tumor paraffin controls, n = 24 normal pancreas paraffin controls) and two buffer controls per DNA isolation batch (n = 10 buffer controls) were introduced. After collecting the respective material in the Digestion Buffer, 4 µL of the kit's Protease were added. The samples were then mixed by gently flicking the tubes. They were quickly spun down in a tabletop centrifuge to collect droplets in the lid. Sample material that sticked to the walls of the tube was pushed back down into the Digestion Buffer using a sterile pipet tip. Subsequently, the samples were incubated for 1 h at 55 °C. Thereafter, the tubes were again quickly spun down to collect condensation before they were subject to incubation at 90 °C for 1 h followed by another brief centrifugation. Next, 120 µL of Isolation Additive from the kit were added and mixed with the sample by pipetting. The mixture was then transferred to a column placed in a collection tube. Here, Filter Cartridges with Collection Tubes from the RecoverAll Total Nucleic Acid Isolation Kit were used instead of PureLink RNA Mini Columns and PureLink Viral Collection Tubes as originally suggested in the "RecoverAll Multi-Sample RNA/DNA Workflow". Then, the columns were centrifuged at 10,000 × g for 30 s. The flow-through was subsequently used for RNA isolation while thereafter DNA was recovered from this column. Therefore, 600 µL Wash 1 Buffer from the kit were applied to the column. Again,

centrifugation at 10,000 × g for 30 s was carried out. The flow-through was discarded and the column put back into the same collection tube. Next, the column was washed with 500 µL Wash 2/3 Buffer from the kit followed by centrifugation at 10,000 × g for 30 s. The flow-through was again discarded and the column reinserted into the collection tube. This wash step involving Wash 2/3 Buffer was repeated. Then, the column was once more centrifuged to remove residual fluid (14,000 × g, 2 min) and transferred to a 1.5 mL low binding reaction tube. A respective amount of the kit's elution buffer sufficient for the respective number of samples was aliquoted to a 2 mL reaction tube and pre-heated to 95 °C. The pipet tip was pre-wetted three times with the hot elution buffer before 50 µL were directly added to the center of the column. Incubation for 1 min at RT was followed by centrifugation at 14,000 × g for 1 min. The eluted DNA was stored at -20 °C until further use.

3.2.9 DNA extraction from human pancreatic cancer cell lines

As a control experiment to the analysis of tumors derived from orthotopically transplanted Capan-2 and MiaPaCa2 cells, DNA was extracted directly from these cultured human pancreatic cancer cell lines (n = 5 from each cell line) wherefore the Invitrogen IHMS DNA extraction protocol described in **3.2.7** was applied. The cells were cultured under sterile conditions in T75 cell culture flasks at 37 °C with 5% carbon dioxide. Capan-2 cells were supplied with Roswell Park Memorial Institute (RPMI) 1640 medium with 10% fetal calf serum (FCS) and 1% penicillin/streptomycin (P/S). MiaPaCa2 cells were grown in Dulbecco's Modified Eagle Medium (DMEM) with 10% FCS and 1% P/S. DNA was isolated from 5 different passages of the cells when they reached about 90% confluency in the T75 flask. Additionally, several NTCs were introduced, namely supernatant (Capan-2 n = 6, MiaPaCa2 n = 5), medium only (Capan-2 n = 6, MiaPaCa2 n = 4), and buffer controls (n = 3). 550 µL of both supernatant and pure medium were aliquoted from each passage in 1.5 mL reaction tubes before cells were harvested for DNA extraction. These NTCs were frozen at -20 °C until DNA extraction. After collecting the NTCs, the cells were washed with phosphate-buffered saline (PBS) and detached by incubation with 2 mL trypsin at 37 °C for 5 min. The enzymatic digestion was stopped by adding 8 mL medium to the flask. The cell suspension was transferred to a 15 mL Falcon tube and counted with the help of an automatic cell counter. 1 mL of the suspension was aliquoted to 1.5mL reaction tubes and stored at -20 °C until

further use. At the time of DNA extraction, supernatant and medium NTCs were either directly subject to DNA extraction or first centrifuged at 1,200 rpm for 3 min to extract DNA from a putative contamination pellet after disposal of the liquid. If directly processed, 500 µL of the NTC aliquot were transferred to a bead tube and 600 µL Lysis Buffer S1 from the Invitrogen kit were added. If centrifuged first, the putative contamination pellet was resuspended in 600 µL Lysis Buffer S1 which was then transferred to a bead tube. The buffer controls were prepared from consumables, buffers, and reagents only but treated exactly like sample material. The cell samples were centrifuged at 1,200 rpm for 3 min, the supernatant was discarded, and the cell pellet was resuspended in 600 µL Lysis Buffer S1 which was then transferred to a bead tube. Subsequently, the Invitrogen IHMS protocol (**3.2.7**) was conducted. The eluted DNA was stored at -20 °C until further use.

3.2.10 DNA purification and quantification

All DNA samples from pancreatic bulk tissue, feces, FFPE tissue, and cells were purified prior to sequencing wherefore the ZymoResearch OneStep PCR Inhibitor Removal Kit was used. To prepare, the column containing a powder matrix was inserted into provided collection tubes and 600 µL of the kit's Prep Solution was added to the column. Then, it was centrifuged at 8,000 × g for 3 min. The collection tube containing the flow-through was discarded and the column was inserted to a 1.5 mL low binding reaction tube. The entire DNA sample was added to the prepped column before it was centrifuged at 16,000 × g for 3 min. The recovered DNA was then ready for downstream analysis. Primarily, purity and concentration of the samples were determined by means of a Nanophotometer and Qubit fluorometer, respectively. To assess purity, the A260/A280 and A260/A230 ratios referring to the quotients of absorbance at 260 nm and 280 nm or 230 nm, respectively, were looked at with the photometer by applying 1 µL shortly vortexed DNA sample to it. Regarding DNA, ratios of 1.8 and 2.0-2.2, respectively, indicate purity. For reasons of accuracy, the fluorometer was employed with the Qubit dsDNA HS Assay to assess DNA contents of the samples. For orientation, the concentration calculated by the photometer was noted down as well. At >120 ng/µL, the sample was diluted 1:10 prior to measuring it in the fluorometer. Therefore, 5 µL DNA and 45 µL of the respective elution solution were mixed in a new 1.5 mL low binding reacting tube. A working solution was prepared by diluting the

concentrated assay reagent (fluorescent dye) 1:200 with the provided dilution buffer corresponding to 1 µL and 199 µL per sample to be measured. 199 µL of this working solution were aliquoted to the respective number of Qubit Assay tubes and 1 µL of the (diluted) DNA sample was added. The assay tubes were quickly vortexed, then incubated for 2 min in the dark. Eventually, they were briefly vortexed again just before inserting them into the fluorometer for measurement. The assessed DNA concentration is required for subsequent library preparation.

3.2.11 Oxford Nanopore 16S rRNA gene sequencing

Due to their low microbial biomass, all DNA samples except for feces were subject to 16S rRNA gene sequencing via third generation sequencing by Oxford Nanopore Technologies. The ONT 16S Barcoding Kit 1-24 (SQK-16S024) was employed according to the protocol "16S Barcoding Kit 1-24 (SQK-16S024), Version: 16S_9086_v1_revJ_14Aug2019/Last update: 03/12/2020". In the establishment experiments, different amounts of input DNA (10 ng, 50 ng, 100 ng) and different numbers of polymerase chain reaction (PCR) cycles (25, 35) were tested. In all following experiments, 50 ng input DNA and 35 PCR cycles were applied for library preparation. For each sample to be sequenced, 50 ng DNA were transferred to a PCR tube and the volume was adjusted to 10 µL with nuclease-free H_2O. In case the concentration was too low, the maximum volume of 10 µL DNA was used. This also applied to the NTCs. In the experiment involving FFPE human patient samples, additional PCR NTCs were introduced with every sample batch that underwent library preparation. To this end, 10 µL nuclease-free H_2O were used instead of DNA eluate (no template control). The PCR tube holding the 10 µL input DNA was gently flicked, then shortly centrifuged to collect all liquids at the bottom. Subsequently, a master mix from 5 µL nuclease-free H_2O and 25 µL LongAmp Hot Start Taq 2X Master Mix per sample (includes NTCs) to be sequenced was prepared. 30 µL of this master mix were distributed to each input DNA and again the tube was gently flicked followed by a quick spin down. Next, for each sample, an individual barcode was selected from the kit. The barcodes were mixed by pipetting up and down approximately ten times. 10 µL of each barcode were transferred to its sample and the reaction was mixed thoroughly by pipetting up and down several times. Each barcode is attached to a specific 16S rRNA gene primer (27F and 1492R) and holds a 5' tag that later enables ligase-free attachment of

Rapid Sequencing Adapters. Subsequently, the PCR tubes were transferred to a thermocycler and the program displayed in **table 7** was run. After the PCR, the samples were transferred to 1.5 mL low binding reaction tubes. AMPure XP beads were resuspended vigorously by vortexing. 30 µL beads were added to and mixed with the sample by pipetting until the beads were evenly submerged. Subsequently, the tubes were placed on a spinning rotator mixer and the samples incubated for 5 min to ensure proper annealing of the DNA fragments to the beads. Meanwhile, 500 µL 70% EtOH per sample were freshly prepared with nuclease-free H_2O. After the incubation on the rotator mixer, the samples were briefly spun down to collect all liquids at the bottom of the tubes which were then transferred to a magnet and kept there. With the beads pelleting to the wall of the tube facing the magnet, the supernatant was carefully pipetted off and discarded. The pellet was washed with 200 µL 70% EtOH which was again removed and discarded. This washing step was repeated once. Then, the tubes were quickly spin down und placed back on the magnet. Residual EtOH was pipetted off and the bead pellets holding the DNA samples were dried for approximately 30 s with the tube lids open. These were then removed from the magnet and the bead pellet resuspended in 10 µL 10 mM Tris-HCl (pH 8.0) with 50 mM NaCl buffer. The samples were incubated for 2 min before they were put back on the magnet. When the eluates were cleared from beads, 10 µL containing DNA freed from other PCR components were retained and transferred to new 1.5 mL low binding reaction tubes. The DNA concentration was measured as elucidated in **3.2.10** using a Qubit fluorometer and 1 µL eluate. Based on the concentrations, the samples were pooled in a new 1.5 mL low binding tube in a total volume of 10 µL. As far as possible, equal amounts of a maximum of 12 samples were multiplexed. True samples and NTCs were multiplexed in separate pools but conditioned samples and control-treated samples were run together in one pool on one flow cell (FC) whenever possible in order to avoid batch effects. The pool of barcoded DNA samples was provided with 1 µL rapid adapters (RAP) from the kit. The tube was flicked to mix and subsequently shortly spun down. Then, the reaction was incubated for 5 min at RT before transferred to ice until further use. Meanwhile, the FC was prepared using the ONT Flow Cell Priming Kit (EXP-FLP002). It is essential to avoid the introduction of air into the FC during the entire process. True samples were run on R9.4.1 FCs while NTCs were run on FLO-FLG001 Flongle FCs. For R9.4.1 FCs, a priming mix was prepared by adding 30 µL Flush Tether from the kit directly to one vial of Flush Buffer from the kit which was then mixed by vortexing. These

FCs require the removal of remaining air from the priming channel prior to introducing 800 µL of the priming mix via the priming port. During a 5 min incubation, the library was prepared by combining the components listed in **table 8** in a new 1.5 mL low binding reaction tube. The two FC types have different loading capacities wherefore the libraries had to be prepared from different volumes (**table 8**). The library was mixed by pipetting just prior to loading it onto the respective FC in order to ensure an even distribution of the kit's fast settling Loading Beads. Additionally, just prior to loading, R9.4.1 FCs were again primed with another 200 µL priming mix from the Flush Buffer vial with the SpotON sample port open to ensure a smooth flow through all channels and ports of the FC. Regarding Flongle FCs, 3 µL Flush Tether and 117 µL Flush Buffer were mixed in a separate reaction tube. This priming mix was directly and entirely pipetted into the Flongle FC's port without incubation time prior to library loading. Loading the FCs with the respective library volume was conducted in a dropwise fashion into the sample port. The FCs were run on a ONT GridION or MinION device controlled by the MinKNOW software v. 22.08.09.

Table 7: Adjusted PCR program of the ONT "16S Barcoding Kit 1-24 (SQK-16S024)" protocol used for 16S rRNA gene sequencing.

Number of cycles	Temperature [°C]	Time [mm:ss]
1	95	01:00
35	95	00:20
	55	00:30
	65	02:00
1	65	05:00
Hold	4	∞

Table 8: Volumes of components from the ONT "16S Barcoding Kit 1-24 (SQK-16S024)" protocol for DNA library preparation according to flow cell type.

Component	Volume [µL] for R9.4.1	Volume [µL] for FLO-FLG001
Sequencing Buffer	34	13.5
Loading Beads	25.5	11
Nuclease-free H_2O	4.5	0
DNA library	11	5.5
Total	75	30

3.2.12 Oxford Nanopore metagenomic sequencing

Due to their high quality high molecular weight DNA contents, fecal samples were subject to metagenomic sequencing via third generation sequencing by ONT. The ONT Ligation Sequencing Kit (SQK-LSK109) in combination with the ONT Native Barcoding Expansion 1-12/13-24 (EXP-NBD104/114), respectively, was employed according to the protocol "Native barcoding genomic DNA (with EXP-NBD104, EXP-NBD114, and SQK-LSK109), Version: NBE_9065_v109_revV_14Aug2019/Last update: 21/02/2020" with minor adaptations. The protocol consists of 3 major parts: DNA repair and end-prep, native barcode ligation, and adapter ligation and clean-up. For each sample to be sequenced, 1 µg DNA was transferred to a new 1.5 mL low binding reaction tube in a maximum volume of 48 µL. If the volume was lower, it was adjusted to 48 µL with nuclease-free H_2O. Regarding the NTCs, the maximum volume was applied. The tubes were flicked to mix DNA and nuclease-free H_2O followed by a quick spin down. Next, a master mix was prepared from the components listed in **table 9** per sample (includes NTCs) in a new 1.5 mL reaction tube. 12 µL of the mix were distributed to each sample adding up to a total of 60 µL per reaction. Then, master mix and DNA were mixed thoroughly by pipetting, and the tubes were incubated on a thermomixer at 20 °C for 5 min, then at 65 °C for 5 min. Meanwhile, the AMPure XP beads were resuspended vigorously by vortexing. After the incubation, 60 µL beads were added to each end-prep reaction and both were mixed by flicking the tubes. Subsequently, the tubes were additionally inserted on a spinning rotator mixer for 5 min. During this DNA annealing time, 500 µL 70% EtOH were freshly prepared per sample. The following steps involving washing and elution are equivalent to the procedure after PCR described in **3.2.11** except the bead pellets were resuspended in 25 µL nuclease-free H_2O. After quantifying the end-prepped DNA as elucidated in **3.2.10** using a Qubit fluorometer and 1 µL eluate, a unique barcode was selected for each sample from the kit. 500 ng end-prepped DNA were transferred to a new 1.5 mL low binding reaction tube in a maximum volume of 22.5 µL. If the required volume was lower, it was adjusted to 22.5 µL with nuclease-free H_2O. If the concentration was too low, the maximum volume was applied. Subsequently, 2.5 µL of the respective Native Barcode and 25 µL Blunt/TA Ligase Master Mix were added to each DNA sample and the reaction was mixed well by pipetting. Then, the AMPure XP beads were resuspended vigorously by vortexing before 50 µL were added to each reaction and it was mixed well by

pipetting. Again, the tubes were placed on a spinning rotator mixer followed by a 5 min incubation. The following washing and elution steps involving 500 µL freshly prepared 70% EtOH are equivalent to the procedure after PCR described in **3.2.11** except the bead pellets were resuspended in 26 µL nuclease-free H_2O. After quantifying the barcoded DNA as elucidated in **3.2.10** using a Qubit fluorometer and 1 µL eluate, equal amounts of each sample were pooled in a new 1.5 mL low binding reaction tube adding up to a total of 700 ng in a maximum volume of 66 µL. As for 16S rRNA gene sequencing, a maximum of 12 samples were merged per sequencing run. Also, true samples and NTCs were separately pooled since they were run on different FC types. Depending on the number of samples to be pooled, the individual amounts of barcoded DNA varied. For instance, if 12 samples were to be pooled, 58.33 ng of each sample were merged in the pool. After quantifying the pooled DNA as elucidated in **3.2.10** using a Qubit fluorometer and 1 µL pool, the volume was adjusted to 65 µL with nuclease-free H_2O. Next, adapter ligation was performed by adding 5 µL Adapter Mix II, 20 µL 5X NEBNext Quick Ligation Reaction Buffer, and 10 µL Quick T4 DNA Ligase to the DNA pool with mixing the tube contents by flicking and spinning them down between each sequential addition. After the final spin, the reaction was incubated for 10 min. Then, the AMPure XP beads were resuspended by vortexing them vigorously. 50 µL were added to the reaction and mixed in by pipetting. Again, a 5 min incubation on the spinning rotator mixer followed. The subsequent two washing steps were essentially carried out as the procedure after PCR described in **3.2.11**, however, for both washes 250 µL of the kit's Short Fragment Buffer (SFB) were used and the bead pellet was resuspended in the SFB by flicking the tube, then returning it to the magnet to re-pellet the beads and pipet off the supernatant. After drying the pellet, it was resuspended in 15 µL Elution Buffer (EB) and incubated for 10 min. Subsequently, the eluate was retained after placing the tube back on the magnet into a new 1.5 mL low binding reaction tube and the adapter ligated DNA was quantified as elucidated in **3.2.10** using a Qubit fluorometer and 1 µL eluate. From the DNA concentration, the weight was calculated in the remaining 14 µL eluate. Assuming an average DNA fragment size of 4 kb, the amount of substance in fmol was calculated. The protocol recommends loading a maximum of 50 fmol per FC which correspond to 129.8 ng at 4 kb nucleic acid size. Hence, the volume corresponding to 129.8 ng of the adapter ligated DNA pool was calculated, transferred to a new 1.5 mL low binding reaction tube, and adjusted to 12 µL with EB. By adding the remaining components in the respective volumes

listed in **table 10**, the DNA library was prepared for true samples (R9.4.1 FC) and NTCs (FLO-FLP001 Flongle FC), respectively. The two different FC types were primed and loaded as elaborated in **3.2.11**. Eventually, both were run on a ONT GridION or MinION device controlled by the MinKNOW software v. 22.08.09.

Table 9: Components and corresponding volumes for 1X master mix for DNA repair and end-prep in ONT metagenomic sequencing with the ONT "Native barcoding genomic DNA (with EXP-NBD104, EXP-NBD114, and SQK-LSK109)" protocol.

Component	Volume [µL] per reaction
NEBNext FFPE DNA Repair Buffer	3.5
NEBNext FFPE DNA Repair Mix	2
Ultra II End-prep reaction buffer	3.5
Ultra II End-prep enzyme mix	3
Total	12

Table 10: Volumes of components from the ONT "Native barcoding genomic DNA (with EXP-NBD104, EXP-NBD114, and SQK-LSK109)" protocol for DNA library preparation according to flow cell type.

Component	Volume [µL] for R9.4.1	Volume [µL] for FLO-FLG001
Sequencing Buffer	37.5	13.5
Loading Beads	25.5	11
DNA library	12	5.5
Total	75	30

3.2.13 Bioinformatic analysis pipeline for sequencing data

Sequencing was conducted under Ubuntu 18.04 and the output was analyzed essentially as previously published in our comprehensive method paper for complex microbiota analysis using ONT (Ammer-Herrmenau et al., 2021b). In brief, the workflow consists of pre-processing, taxonomic classification, and post-processing. Specifically, pre-processing comprises high accuracy base calling of the fast5 files (raw data format) performed by the Guppy toolkit 6.2.11 integrated in the MinKNOW software v. 22.08.9 by ONT. Per default, MinKNOW applies a quality score (QS) of 9 which was not changed. Base calling returns

fastq files containing the nucleotide sequences. Next, taxonomic classification of those was conducted by per-read manner using the Metagenomic Taxonomy Pipeline for ONT sequencing (MeTaPONT). This pipeline comprises classification via the classifier Centrifuge. To this end, two different indices were used. The nt library contains all complete and incomplete genomes from the NCBI BLAST's nt database and was built on October 15, 2020. This index was used to classify all sequencing data derived from snap-frozen samples. For all FFPE sample data, the abv library was applied for classification. This index was built manually on September 10, 2020 containing all complete NCBI bacterial, fungal, archaea, human, and mouse RefSeq genomes. MeTaPONT further implies alignment control via Minimap2 including an alignment score (AS) of 1000 and an alignment coverage (Cov) of 50. The workflow was implemented as a wrapper program and can be retrieved as a docker from github (github.com/microbiome-gastro-UMG/MeTaPONT). Eventually, regarding post-processing of the classified sequencing data, subsequent analysis was performed in R version 2.4 or later and R studio. MeTaPONT output and meta data were combined in a phyloseq object. Then, all reads were filtered to bacteria on species level. Next, prevalence filtering was applied using the PERfect package with the simultaneous approach (Smirnova et al., 2019). Subsequently, decontamination was conducted using the decontam package with the prevalence method (Davis et al., 2018). To normalize the microbial data, the samples in the phyloseq object were rarified to a defined read number which was chosen as a compromise between reaching the rarefaction curve asymptote and losing samples due to insufficient sequencing depth. To assess alpha diversity metrics OS, SI, and InvS, non-rarified data was used, while the calculation of beta diversity metrics uwUF, wUF, and BC was conducted from rarified data. Statistics were performed on alpha diversity metrics by testing for normal distribution and homogeneity of variance via Shapiro-Wilk and Levene's test, respectively, followed by *t*-test or Wilcoxon signed-rank test according to the result of the normality test. Beta diversity metrics were ordinated via principal coordinates analysis (PCoA) and tested for significance by permutational multivariate analysis of variance (PERMANOVA) which was conducted using adonis2 of the vegan package. Differential abundances were determined via the linear discriminant analysis effect size (LEfSe) approach. Eventually, regression models using published human microbiome data to be compared with the murine microbial compositions from this thesis were built by logistic regression based L2 regularization (Ridge regression) which was calculated using the cv.glmnet function.

4 Results

4.1 Chapter 1 | Establishment of Oxford Nanopore sequencing and bioinformatic analysis workflow for murine fecal and pancreatic samples

4.1.1 Impact of the DNA extraction protocol on DNA concentration, sequenced species and read count

Modern microbiome research requires a high throughput sequencing technique with accompanying steps such as DNA isolation and bioinformatic analysis. Therefore, the establishment of an experimental pipeline was a prerequisite for this thesis and constitutes a major part of it. Most of the results presented here in **4.1** are published under the Creative Commons Attribution 4.0 International license (CC BY 4.0) in the method paper "Comprehensive Wet-Bench and Bioinformatics Workflow for Complex Microbiota Using Oxford Nanopore Technologies" which I co-authored (Ammer-Herrmenau et al., 2021b).

The experimental pipeline consists of three major steps: DNA preparation, sequencing, and data processing. Regarding sequencing, third generation sequencing by ONT was chosen with the MinION and GridION sequencing platforms. Although Illumina NGS platforms established in other departments have undisputed competence, the nanopore-based technology has some advantages which are of great relevance for microbiome research and will be discussed in **chapter 5**. Establishing the most suitable DNA extraction and preparation method compatible with ONT sequencing required testing different DNA isolation kits in combination with different protocols. The initial experiments were performed with buccal and rectal swabs as elucidated in the method paper because they are easy-to-access samples, relevant for future clinical studies, and prevented wasting valuable murine samples (Ammer-Herrmenau et al., 2021b). Hence, most results in **4.1** were generated using swabs, however, the developed pipeline was eventually validated and accommodated to murine fecal and pancreatic samples in **4.1.4**.

The DNA extraction method is known to have relevant impact on the results of microbial analyses (Bjerre et al., 2019; Teng et al., 2018). Therefore, four different DNA isolation kits from three different vendors were tested. They are referred to as "Applied Biosystems"

(MagMAX Microbiome Ultra Nucleic Acid Isolation Kit), "Invitrogen" (PureLink Microbiome DNA Purification Kit), "Qiagen investigator" (QIAamp DNA Investigator Kit), and "Qiagen microbiome" (QIAamp DNA Microbiome Kit) in the following. For each of these kits, two different protocols were tested, the original manufacturer's protocol (referred to as "original") and an adapted version thereof modified according to the IHMS recommendations (referred to as "IHMS"). For the QIAamp DNA Microbiome Kit, only the original protocol was tested because a meaningful implementation of the IHMS recommendations was not possible. Thus, a total of seven different protocols were assessed. The Qiagen investigator original and Invitrogen IHMS protocols are described in detail in **3.2.6** and **3.2.7** since only these two are relevant for the further experiments presented in this thesis. The pre-experiments narrowed the protocols down to these two as best candidates for murine tissue samples which were then tested in **4.1.4**. The other protocols and specifications can be retrieved from the method paper (Ammer-Herrmenau et al., 2021b).

DNA concentration and purity, number of observed species and read counts were adduced as readouts to determine the best-fitting DNA extraction protocol for ONT 16S rRNA and metagenomic sequencing. For both buccal and rectal swabs, both the Qiagen investigator original and Invitrogen IHMS protocols achieved significantly higher DNA concentrations compared to the other DNA extraction protocols (**figure 7a,b**). They were also the only two protocols that achieved median concentrations above the threshold of 10.42 ng/μL indicated by the dashed line in **figure 7a,b**. This DNA concentration is necessary to provide at least 500 ng input DNA in 48 μL reaction mixture for metagenomic sequencing. The median DNA concentration extracted from buccal swabs was the highest when using the Qiagen investigator original protocol. The median DNA concentration extracted from rectal swabs was the highest when using the Invitrogen IHMS protocol.

Besides the concentration, a high level of DNA purity is essential for sensitive downstream applications such as sequencing. Impurities may potentially interfere with the chemical properties and physical modules of the sequencing platform. A260/A280 and A260/A230 ratios estimated by Nanophotometer measurements are indicators of purity. Regarding the A260/A280 value, little derivation from the desired ratio of 1.8 (dashed line) was registered across all seven protocols (**figure 7c**). Applied Biosystems IHMS and Invitrogen IHMS were the closest in their median A260/A280 ratios to 1.8. Concerning the A260/A230 value, major

derivation from the desired ratio of 2.0-2.2 (dashed line) was noted across the tested DNA extraction protocols (**figure 7d**). The Invitrogen IHMS protocol showed by far the median values closest to 2.0 as compared to the other protocols, however, the variance was still high ranging from 1.0 to 2.5.

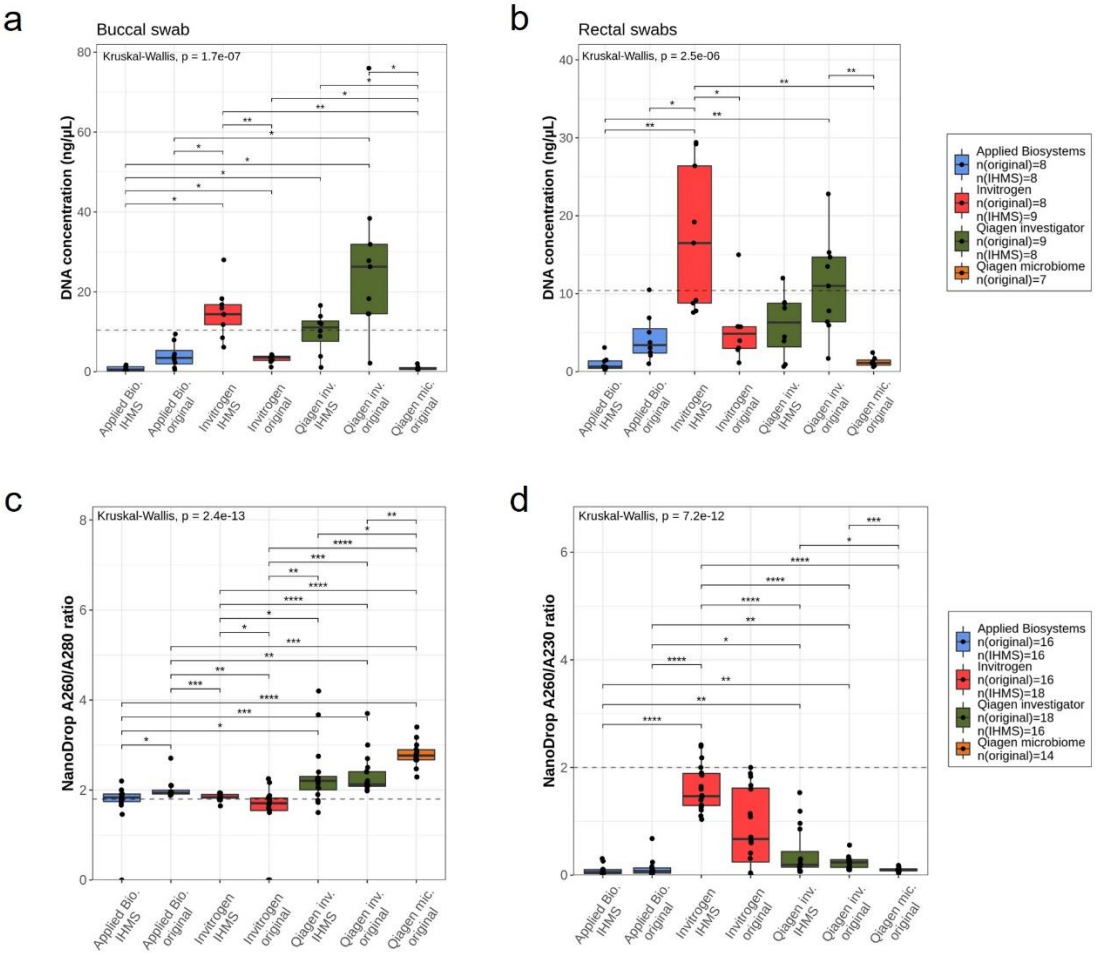

Figure 7: DNA yield and purity as readouts for DNA extraction kit performance on buccal and rectal swab samples. (a) DNA concentrations of buccal swab samples isolated with different original and IHMS-modified DNA extraction kit protocols. N-numbers indicated in graph legend (upper panel row). Dashed line indicates DNA concentration threshold for metagenomic sequencing (10.42 ng/µL). (b) DNA concentrations of rectal swab samples isolated with different original and IHMS-modified DNA extraction kit protocols. N-numbers indicated in graph legend (upper panel row). Dashed line indicates DNA concentration threshold for metagenomic sequencing (10.42 ng/µL). (c) Photometric A260/A280 ratios of buccal and rectal swab samples isolated with different original and IHMS-modified DNA extraction kit protocols. N-numbers indicated in graph legend (lower panel row). Dashed line indicates ideal A260/A280 ratio for pure DNA (1.8). (d) Photometric A260/A230 ratios of buccal and rectal swab samples isolated with different original and IHMS-modified DNA extraction kit protocols. N-numbers indicated in graph legend (lower panel row). Dashed line indicates ideal A260/A230 ratio for pure DNA (2.0-2.2). All statistical testing performed with Kruskal-Wallis and pairwise Wilcoxon signed-rank tests. *, $p \leq 0.05$; **, $p \leq 0.01$; ***, $p \leq 0.001$; ****, $p \leq 0.0001$.

Next, the performance in 16S rRNA and metagenomic sequencing of the differently extracted DNA samples was tested to assess their quality for the target application. For that purpose, the rarefaction curves, in which the read count is plotted against the number of observed species, were evaluated to determine the sequencing depth (read count) threshold at which most samples reached their curve's asymptote (**figure 8a,b**).

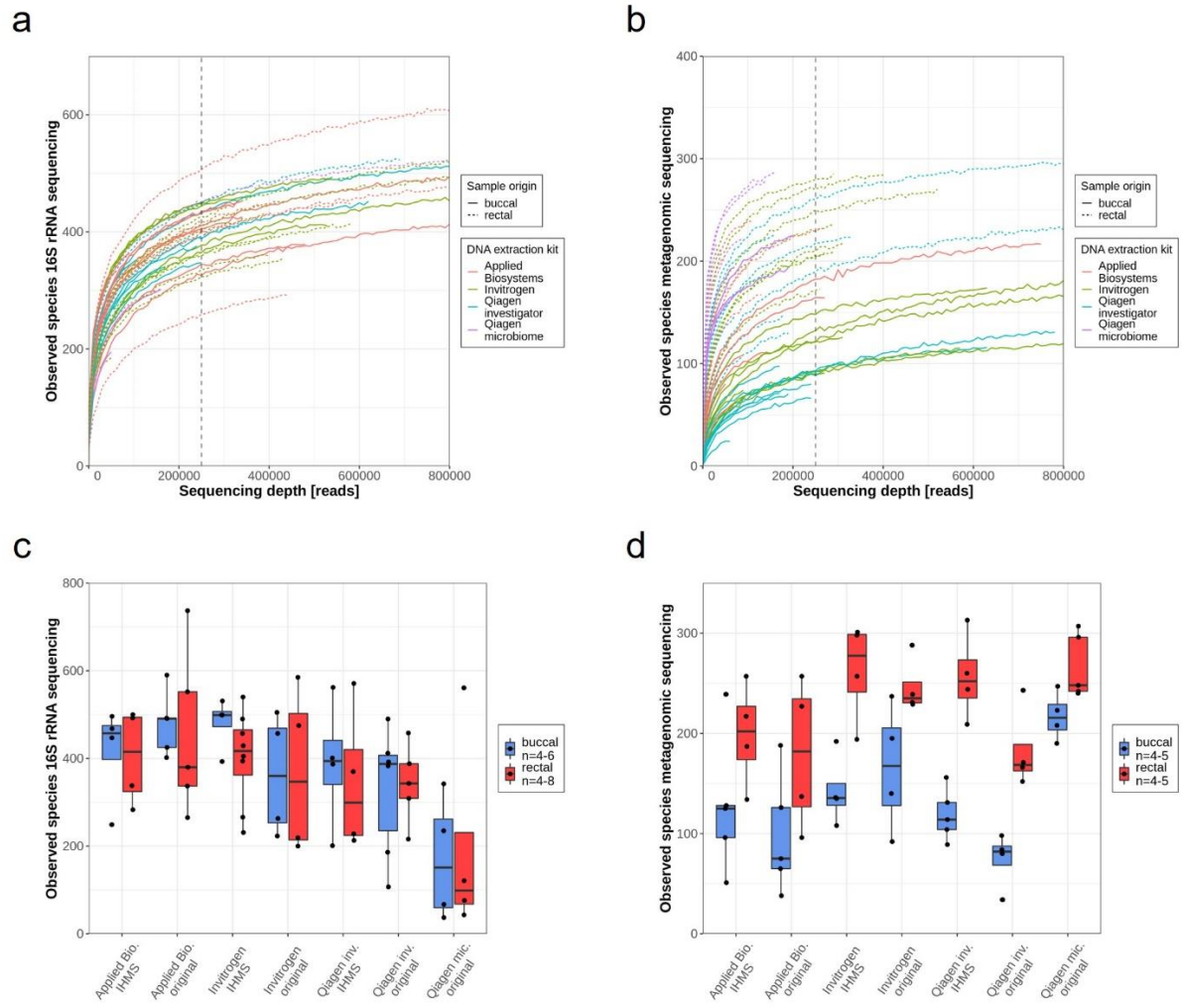

Figure 8: Evaluation of 16S rRNA and metagenomic sequencing of buccal and rectal swab samples via rarefaction curves and alpha diversity metric observed species. (a) Rarefaction curves of buccal and rectal samples isolated with different DNA extraction kits as determined via 16S rRNA sequencing. The vertical dashed line indicates the determined sequencing depth threshold of 250,000 reads. (b) Rarefaction curves of buccal and rectal samples isolated with different DNA extraction kits as determined via metagenomic sequencing. The vertical dashed line indicates the determined sequencing depth threshold of 250,000 reads. (c) Alpha diversity metric OS of buccal and rectal swab samples isolated with different original and IHMS-modified DNA extraction kit protocols and analyzed via 16S rRNA gene sequencing. (d) Alpha diversity metric OS of buccal and rectal swab samples isolated with different original and IHMS-modified DNA extraction kit protocols and analyzed via metagenomic sequencing. All statistical testing performed with Kruskal-Wallis and pairwise Wilcoxon signed-rank tests.

This threshold was set to 250,000 for both 16S rRNA and metagenomic sequencing (dashed line). Except for samples extracted with the Qiagen microbiome protocol, most samples were able to cross this threshold. Especially in metagenomic sequencing, exclusively all samples extracted with the Qiagen microbiome kit failed to reach the threshold. The alpha diversity metric OS showed the highest median absolute numbers for both buccal and rectal swab samples when isolated with Invitrogen IHMS and undergoing 16S rRNA sequencing, although these results were not statistically significant (**figure 8c**). Regarding metagenomic sequencing, the Invitrogen IHMS-isolated rectal swab samples also achieved the highest number of OS as compared to the rectal swab samples isolated via the other protocols, however, the results here did not reach statistical significance either (**figure 8d**).

Eventually, read counts achieved with metagenomic sequencing were compared between the DNA extraction protocols (**figure 9**). The samples were pooled equimolarly in the libraries to ensure highly similar DNA input amounts and the read counts normalized to the Qiagen investigator IHMS protocol to compensate varying FC pore counts. Still, the Invitrogen protocols achieved the highest count of sequenced DNA stretches (**figure 9**).

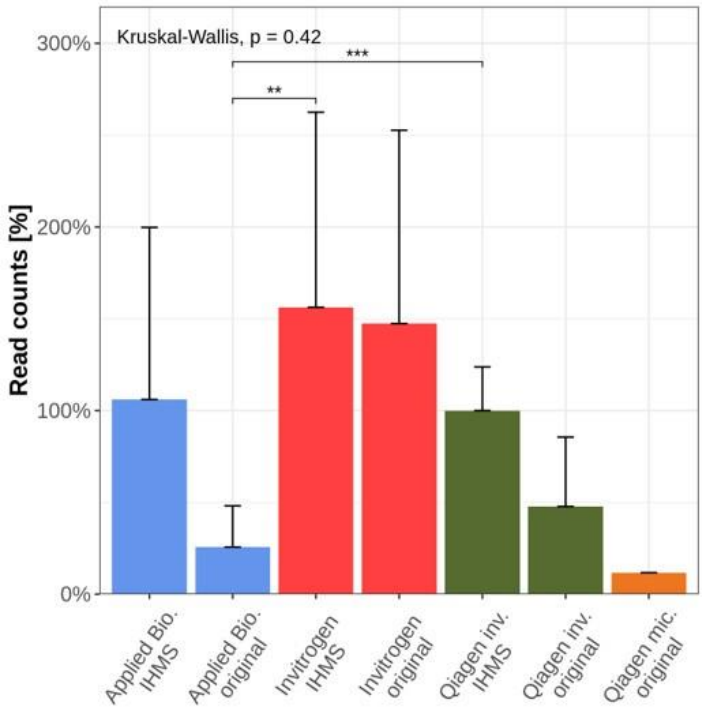

Figure 9: Comparison of read counts achieved via metagenomic sequencing of swab samples isolated with different original and IHMS-modified DNA extraction kit protocols. N = 5 metagenomic sequencing experiments are pooled. Read counts are plotted in percent relatively to the Qiagen investigator IHMS results. Statistical testing performed with Kruskal-Wallis and pairwise Wilcoxon signed-rank tests. **, $p \leq 0.01$; ***, $p \leq 0.001$.

4.1.2 Read classification using Centrifuge and evaluation of different classifier indexes

Sequenced DNA fragments need to be base called meaning the translation of the primary raw data format to bases. In the case of nanopore sequencing, raw data are voltage changes along the membrane when bases with different properties pass the membrane through the pore. For ONT sequencing, base calling is performed by Guppy, the official base caller integrated in the control software MinKNOW. Thereafter, base called DNA fragments need to be classified meaning their assignment to a specific species. Finally, this classification undergoes alignment control. Here, the classifier Centrifuge was chosen for its sensitive and fast annotation with rather low computational capacity. Minimap2 was selected as a sequence alignment tool for its fast mapping of long noisy reads against large databases. Comparing read to taxonomic identification number (taxID) annotations by Centrifuge only and Centrifuge combined with Minimap2 revealed over a quarter of sequence-to-taxID classifications mismatched when using Centrifuge only as visualized by the Venn diagram in **figure 10a**. To dispose of low-quality reads, a Centrifuge QS and a maximum number of annotations per single read were defined. As QS, 150 was chosen as proposed in the literature because, although an increase of this value removed considerably more reads, Minimap2 already removes a great part of these sequences (Sanderson et al., 2018). This is emphasized by the slopes of Minimap2-controlled read counts decreasing with increasing Centrifuge QS (**figure 10b**). However, the Minimap2-controlled read counts are not as much influenced than the Centrifuge only-classified reads. To show that reads with low information content can safely be omitted to save computational time, to minimize false positive hits, and thus to limit the number of matches per single read, a hit length-to-query length ratio was calculated. Centrifuge usually assigned below 50 taxIDs and the required annotation length (hit length) for classification only took up 40% of the sequence length (query length) (**figure 10c**). Eventually, the indexes used for classification had to be judged by their excellence and performance. To this end, four libraries were assessed: two preformed indexes from Centrifuge, pvh (2016) and p (2018), the NCBI complete genomes index (2020), also referred to as abv library in **3.2.13**, as well as the nt database with all complete and incomplete NCBI genomes (2020). Precision, recall and area under precision-recall curve (AUPR) were adduced as readouts to test the four indexes on a commercially

available metagenomically sequenced gut mock community consisting of 17 species (14 common gut bacteria, one archaeon, and two fungi) (**figure 10d**). The indexes pvh and p identified only 11 and 13 species, respectively. The NCBI complete genomes index identified only 14 species. A recall of 100% was only achieved by the complete and incomplete genomes nt database. This index also had the highest AUPR. The NCBI genomes reached the second highest recall and AUPR percentages. All four indexes had about equally low precision percentages.

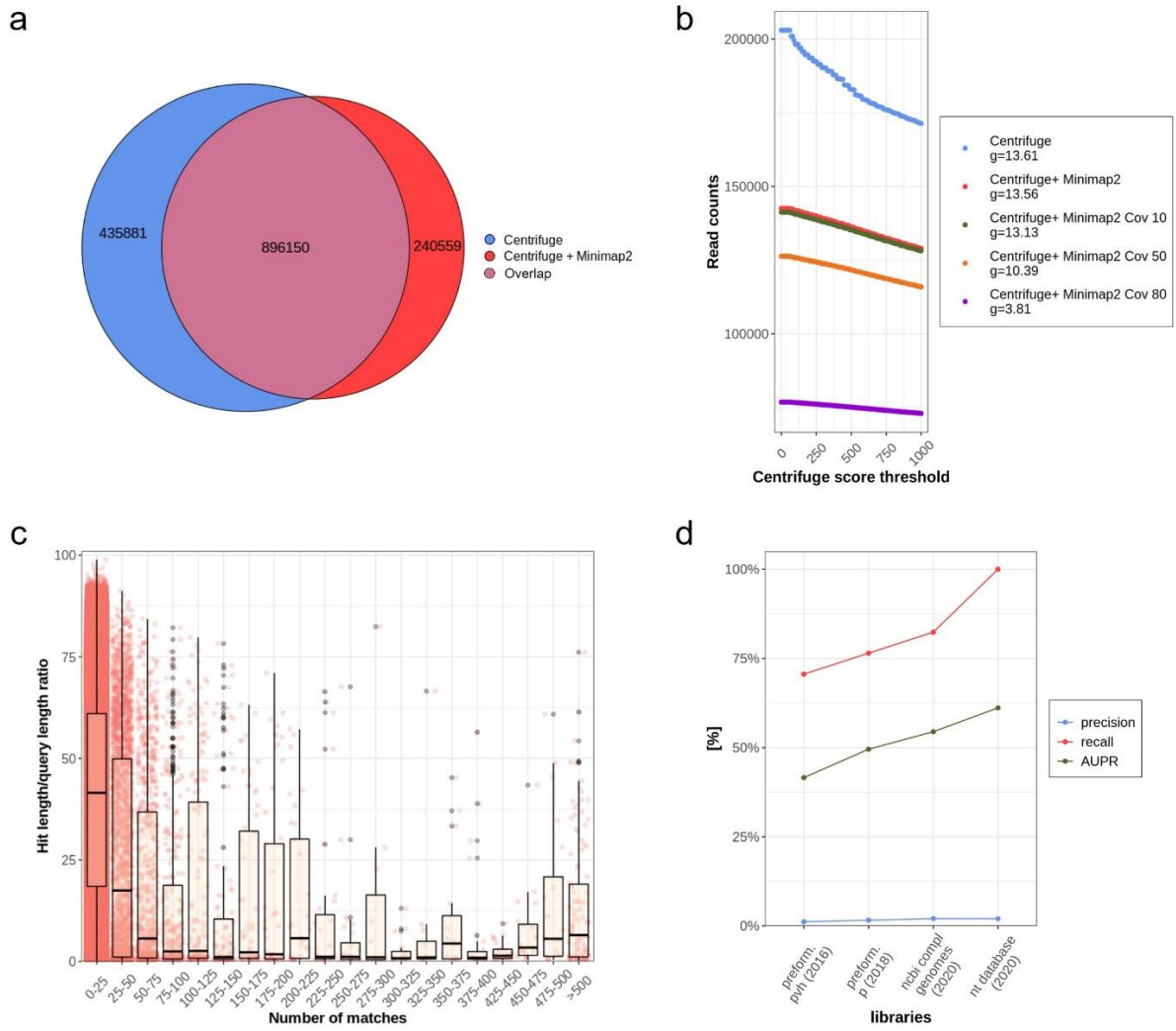

Figure 10: Evaluation of read classifier Centrifuge and different classifier indexes. (a) Venn diagram of sequence-to-taxID classification by Centrifuge only as well as Centrifuge and Minimap2 (no further filters). (b) Influence of Centrifuge quality score on read counts of n = 4 metagenomically sequenced samples classified by Centrifuge and controlled by Minimap2 with different coverages. Difference quotient g is indicated for each classification setting in the graph legend. (c) N = 12 rectal samples classified with Centrifuge and sorted by their number of taxID matches are plotted against the hit length-to-query length (classified length-to-total sequence length) ratio. Each red dot corresponds to one taxID. Sequences with more than 50 taxIDs are dropped. (d) Readout parameters precision, recall, and AUPR of four different Centrifuge classification indexes.

4.1.3 Benchmarking experiment testing different Minimap2 parameters

As elucidated, alignment control by Minimap2 is indispensable since Centrifuge has in fact a high sensitivity, however, also an intolerably high false positive rate. Benchmarking was conducted to determine the best AS and Cov values that remove minimum true positives and maximum false positives by excluding low quality reads. Precision, recall and AUPR percentages were calculated for each AS/Cov combination tested (**figure 11a**). AS1500/Cov50 achieved almost the highest precision and AUPR with a simultaneous recall of 100%. However, although AS2000/Cov50 reached a slightly higher precision and AUPR values, this combination also excluded true positives. Precision and AUPR were comparable between Cov10- and Cov50-classified samples. The read counts declined with increasing AS and Cov as expected (**figure 11b**). In this, a higher AS was more decisive than a higher Cov, in particular for 16S rRNA sequencing. The same trend applied to the alpha diversity metric OS (**figure 11c**). Here, the number of OS was generally lower in metagenomic sequenced samples compared to 16S rRNA sequenced samples except for the Minimap2 settings AS1000/Cov50. Samples classified with these parameters presented with the lowest but an about equal median OS for both 16S rRNA and metagenomic sequencing. Picking up on the impact of Cov changes, a benchmarking experiment with rectal swabs showed that a Cov of 50 (and AS1500) was able to exclude incorrectly annotated environmental bacteria compared to Minimap2 without any thresholds and with Cov10 (and AS1500) as visualized by the microbial composition of the samples in **figure 11d**. All of these establishment experiments resulted in a methodological and bioinformatic workflow (Invitrogen IHMS, nt database, Minimap2 AS1000 or 1500/Cov50) with highly accurate and reproducible sequencing data output and, although tested on swab material, also transferable to other sample material. In the following **4.1.4**, ONT sequencing parameters are optimized for murine tumor material.

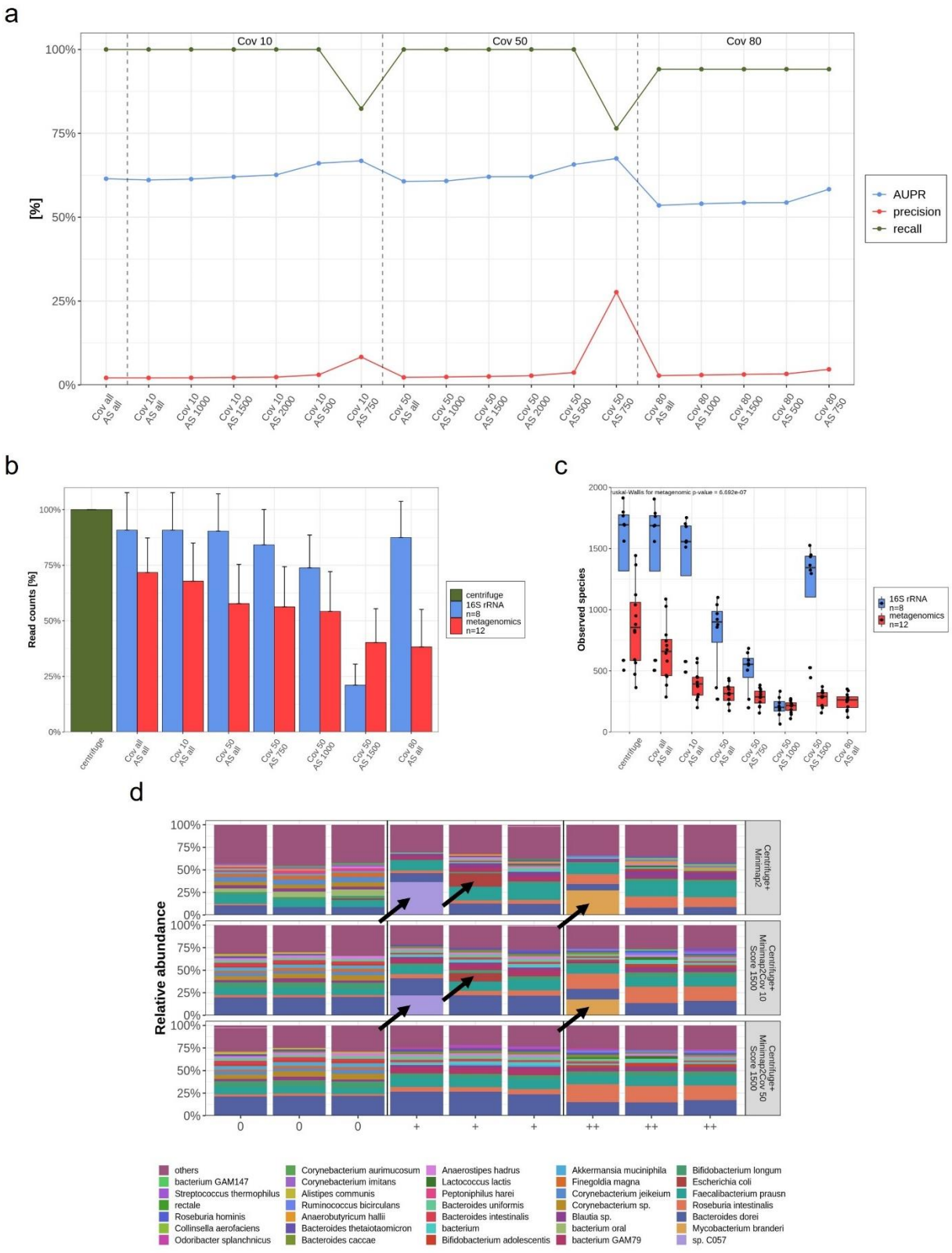

Figure 11: Benchmark experiment testing different Minimap2 parameters. (a) Readout parameters precision, recall, and AUPR of various combinations of AS and Cov tested on metagenomic gut mock community sequencing data. (b) Read counts in percent of 16S rRNA gene and metagenomically sequenced rectal swab samples determined by different AS/Cov combinations. Statistical testing performed with Kruskal-Wallis test. (c) Number of observed species in 16S rRNA gene and metagenomically sequenced rectal swab

samples determined by different AS/Cov combinations. Statistical testing performed with Kruskal-Wallis test. (d) Microbial compositions (species level) of n = 9 rectal swab samples with three different grades of stool contamination (0, +, ++) that were subject to Centrifuge classification and different Minimap2 filtering (no threshold, AS1500/Cov10, AS1500/Cov50). Black arrows point out contaminating species that are dropped by AS1500/Cov50.

4.1.4 Evaluation of different DNA extraction protocols and PCR settings for murine tumor samples

After establishing the experimental pipeline with swab samples, the technical details in the ONT 16S rRNA sequencing protocol were optimized for low biomass murine bulk tumor samples with special requirements. Therefore, the two most promising DNA extraction protocols, Invitrogen IHMS and Qiagen investigator original (a specific version for tissues), were tested as well as three different amounts of DNA input into the PCR, namely 10 ng, 50 ng, and 100 ng, and two different numbers of PCR cycles, 25 and 35. These parameters were evaluated on four different bulk tumor tissue samples in all possible combinations resulting in 48 differently prepared samples.

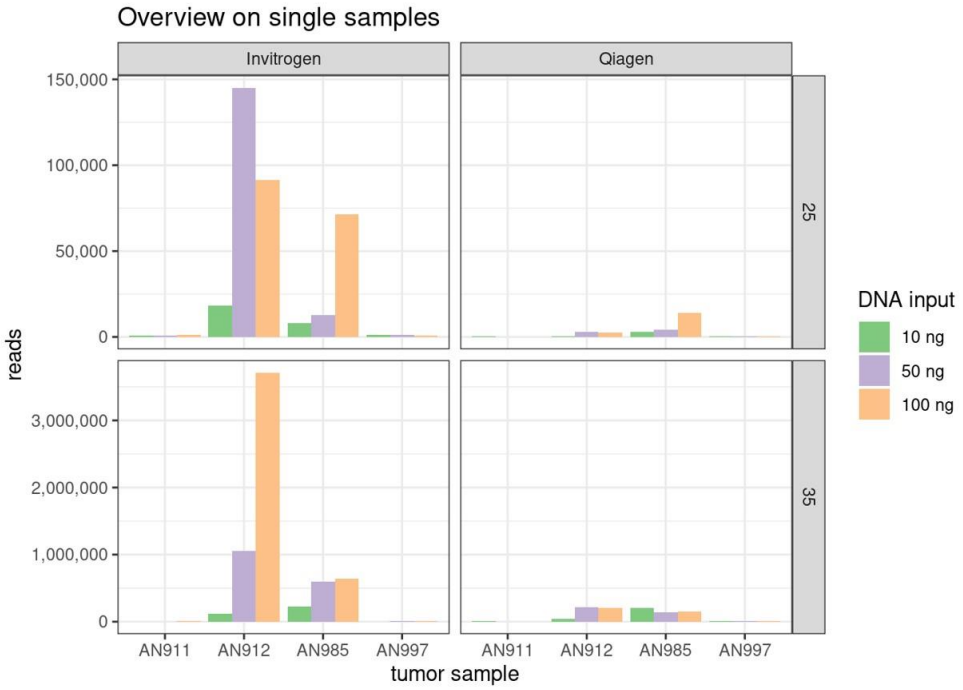

Figure 12: Overview on read counts of four KPC tumor samples (AN-numbers) analyzed in different ways. DNA was isolated according to the Invitrogen IHMS and the Qiagen investigator original DNA extraction protocols. 16S rRNA sequencing was performed with varying parameters (10 ng, 50 ng, or 100 ng DNA input and 25 or 35 PCR cycles).

The number of reads was employed as readout to evaluate the best sequencing parameters knowing high read count samples performed better in the downstream normalization process. In **figure 12**, the read counts of all four tumor DNA samples library-prepared in 12 different ways each are displayed. In this linear notation, the heterogeneity among the samples is directly visible. Even among equally prepared samples, the strongly varying read numbers are remarkable. To decide on the best DNA extraction and library preparation, the samples were clustered according to the different parameters.

To begin with, Invitrogen and Qiagen extraction methods were compared with regard to 25 (**figure 13a**) and 35 PCR cycles (**figure 13b**), respectively. Samples extracted with Invitrogen IHMS had significantly more reads than Qiagen investigator original-extracted samples when undergoing PCR with 25 cycles. Contrary, there was no significant difference in the read count between Invitrogen and Qiagen samples sequenced after PCR with 35 cycles. However, there was a trend towards more reads in Invitrogen samples. Also, the median read count was higher than in Qiagen samples. Generally, 35 PCR cycles resulted in considerably higher read numbers than 25 PCR cycles. For a direct comparison of the number of PCR cycles, the samples were also plotted for the Invitrogen IHMS (**figure 13c**) and Qiagen investigator original DNA extraction protocol (**figure 13d**) separately. On the one hand, DNA samples isolated with Invitrogen IHMS did not reach read counts of significant difference when amplified with 25 and 35 PCR cycles, respectively. Still, there was a considerable number of samples resulting in clearly higher read count if they underwent PCR with 35 cycles which also shows up in the slightly higher median read number. On the other hand, samples extracted with Qiagen had significantly higher read counts after PCR with 35 cycles compared to 25 cycles. However, the determining factor was the overall higher read counts achieved by Invitrogen-isolated DNA samples. Also, although this did not reach statistical significance, 35 PCR cycles tended to result in more reads than 25 cycles. Thus, the DNA extraction protocol of choice was Invitrogen IHMS in combination with 35 PCR cycles for library preparation. Eventually, the three different DNA input amounts were compared among samples extracted by Invitrogen and amplified in 35 PCR cycles (**figure 13e**). 50 ng and 100 ng DNA input for PCR amplification and barcoding resulted in clearly higher median read counts compared to 10 ng DNA input suggested by the original ONT 16S rRNA sequencing protocol, however, these differences were not significant. The median numbers of reads of 50 ng and 100 ng DNA input were the same, wherefore 50 ng DNA as

PCR input were chosen for all further experiments involving 16S rRNA sequencing of murine pancreatic samples.

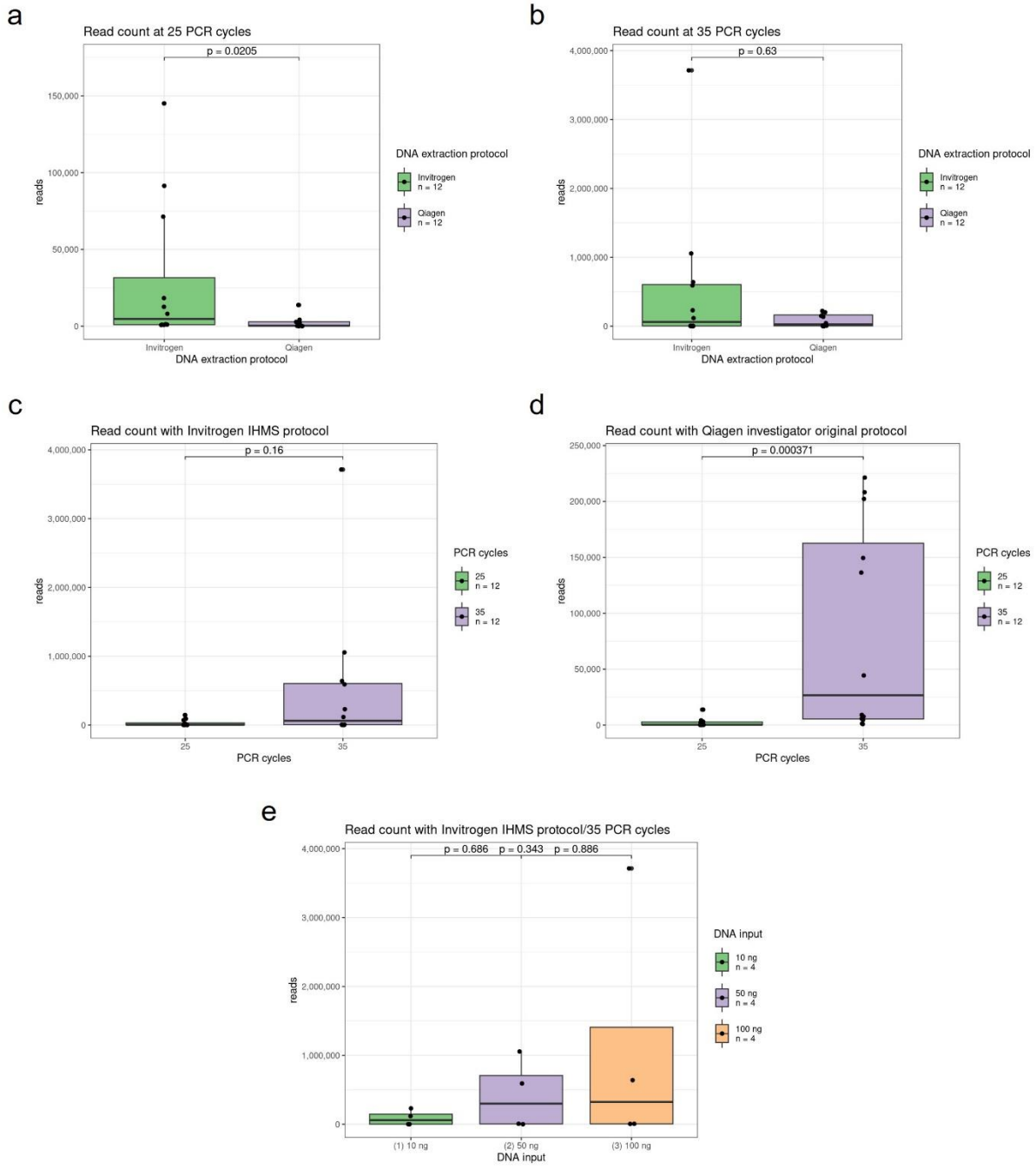

Figure 13: Evaluation of DNA extraction protocols and PCR settings for KPC tumor samples (low microbial biomass). (a) Significant difference between read counts of KPC tumor samples isolated with the Invitrogen IHMS and the Qiagen investigator original protocol, respectively, and 16S rRNA gene sequenced with varying DNA input amounts (10 ng: n = 4, 50 ng: n = 4, 100 ng: n = 4) and 25 PCR cycles. (b) Read counts of KPC tumor samples isolated with the Invitrogen IHMS and the Qiagen investigator original protocol, respectively, and 16S rRNA gene sequenced with varying DNA input amounts (10 ng: n = 4, 50 ng: n = 4, 100 ng: n = 4) and 35 PCR cycles. (c) Read counts of KPC tumor samples isolated with the Invitrogen IHMS protocol, and 16S rRNA gene sequenced with varying

DNA input amounts (10 ng: n = 4, 50 ng: n = 4, 100 ng: n = 4) as well as 25 and 35 PCR cycles, respectively. (d) Significant difference between read counts of KPC tumor samples isolated with the Qiagen investigator original protocol, and 16S rRNA gene sequenced with varying DNA input amounts (10 ng: n = 4, 50 ng: n = 4, 100 ng: n = 4) as well as 25 and 35 PCR cycles, respectively. (e) Read counts of KPC tumor samples isolated with the Invitrogen IHMS protocol, and 16S rRNA gene sequenced with varying DNA input amounts (10 ng: n = 4, 50 ng: n = 4, 100 ng: n = 4) and 35 PCR cycles. All statistical testing performed with Wilcoxon rank test.

DNA isolation from murine fecal samples was tested as well using the Invitrogen IHMS protocol and resulted in excellent DNA concentrations of several hundred ng/µL. Additionally, the purity was very close to the reference ratios for A260/A280 (1.8) and A260/A230 (2.0-2.2). Only very few exceptions deviated from these values which was usually the case when the feces were rather mucoid. In summary, DNA from murine samples was also extracted with Invitrogen IHMS, followed by 50 ng DNA input and 35 PCR cycles in 16S rRNA sequencing for pancreatic samples, and maximum suggested DNA input (1 µg) in metagenomic sequencing for fecal samples. The bioinformatic workflow established with swabs was followed for murine samples as well.

4.2 Chapter 2 | The fecal and pancreatic microbiome in KPC mice

4.2.1 The fecal microbiome of KPC mice is significantly different from healthy controls

The results presented here in **4.2** are submitted for publication and the manuscript entitled "Dynamics of intestinal and intratumoral microbiome signatures in genetically engineered mice and human pancreatic ductal adenocarcinoma" which I wrote as a first author is currently under revision. After establishing the experimental and bioinformatic pipeline for ONT sequencing of microbiota, in particular bacteria, the primary focus was set to the characterization of the fecal and tumoral microbiome of KPC mice. The inevitable necessity of mouse models, in particular GEMMs, as a fundamental basis in translational science was introduced in **1.3**. However, they need to be clearly understood in their properties in order to draw meaningful translatable conclusions. To date, KPC mice are insufficiently studied regarding their microbiome.

Knowing about the microbial presence in the gut and the comparably profound characterization thereof, first, the fecal microbiome of KPC mice was evaluated. Against the background of the microbiome undergoing continuous shifts in response to many influences, fecal samples of KPC and CTRL mice were collected at multiple time points (6w, 12w, and the EP as defined in Methods **3.2.1**) to learn about the dynamics and possible changes in the microbial composition over time. KPC mice at the EP were between 4.5 and 6.5 months old with a median age of 5.75 months. Hence, they were approximately age-matched with the CTRLs sacrificed at an age of 6 months. The fecal DNA of all three age groups (n = 31 KPC and n = 30 CTRL samples in total) was sequenced metagenomically.

Alpha diversity metrics OS, SI, and InvS did not show any significant differences between KPC and CTRL feces at any age (**figure 14**). The normalization process to determine beta diversity metrics reduced the sample number to n = 28 KPC and n = 27 CTRL feces. Calculating uwUF, wUF, and BC returned statistically significant differences between KPC and CTRL feces in all three age groups across all three metrics with decreasing p-values in increasing ages (**figure 15**).

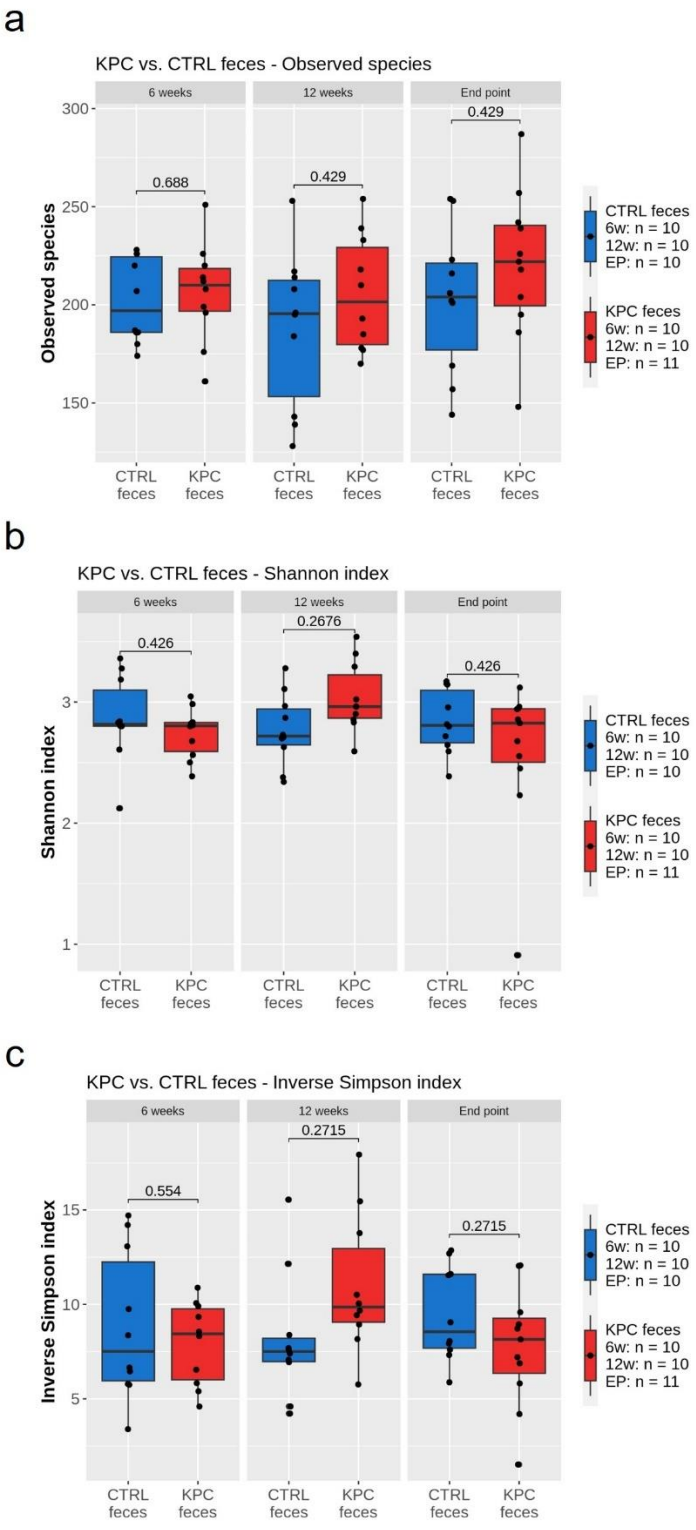

Figure 14: Alpha diversity in feces of 6w, 12w, and EP KPC and CTRL mice. (a) Numbers of observed species are not significantly different between KPC and CTRL fecal samples in any of the regarded age groups. (b) Shannon indexes are not significantly different between KPC and CTRL fecal samples in any of the regarded age groups. (c) Inverse Simpson indexes are not significantly different between KPC and CTRL fecal samples in any of the regarded age groups. All metrics calculated with non-rarified library sizes. All statistical testing performed with t-test or Wilcoxon rank test according to normal distribution test results.

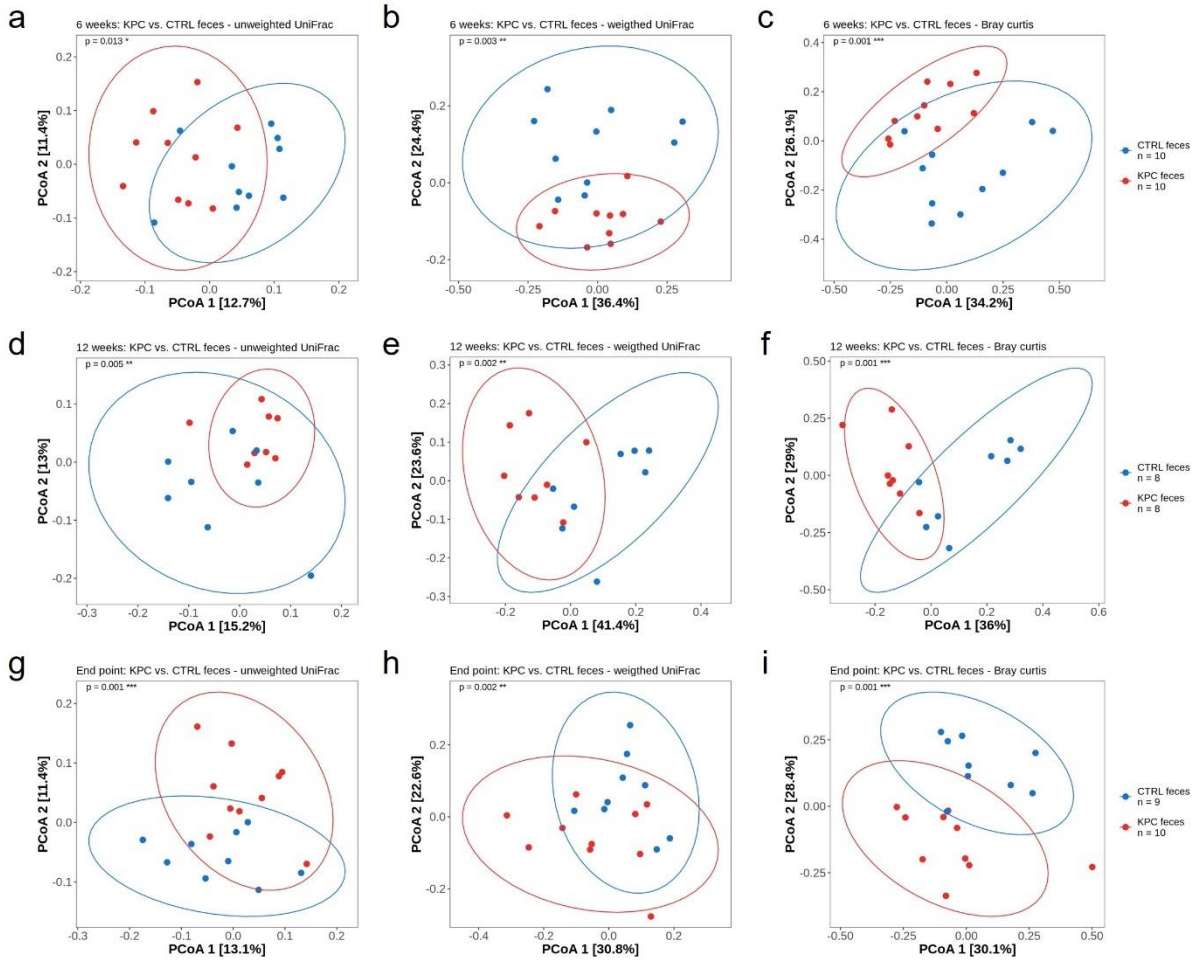

Figure 15: Beta diversity between feces of 6w, 12w, and EP KPC and CTRL mice. (a-c) Beta diversity metrics uwUF, wUF, and BC show significant differences between feces of 6-weeks-old KPC and CTRL mice. Metrics calculated with rarified library sizes. (d-f) Beta diversity metrics uwUF, wUF, and BC show significant differences between feces of 12-weeks-old KPC and CTRL mice. Metrics calculated with rarified library sizes. Two samples dropped in each group by rarefaction. (g-i) Beta diversity metrics uwUF, wUF, and BC show significant differences between feces of EP KPC and CTRL mice. Metrics calculated with rarified library sizes. One sample dropped in each group by rarefaction. All statistical testing performed with PERMANOVA. *, $p \leq 0.05$; **, $p \leq 0.01$; ***, $p \leq 0.001$.

Performing linear discriminant analysis (LDA) with the LEfSe algorithm on all age groups separately to determine the taxa responsible for these differences found *Lactobacillales* significantly enriched in CTRL feces as opposed to KPC feces regardless of age (**figure 16a-c**). In feces of 6w- and 12w-old CTRL mice, *Lactobacillus murinus*, and in feces of EP CTRL mice, *Lactobacillus reuteri* were more pronounced according to the LDA score compared to the corresponding KPC feces. In turn, KPC feces were significantly enriched with different *Bacteroides spp.* independently of age. More precisely, in feces of 6w-old KPC mice,

Duncaniella, in feces of 12w-old KPC mice, *Helicobacter*, and in feces of EP KPC mice, *Escherichia coli* and *Parabacteroides distasonis* were more abundant as calculated by LEfSe. Looking at the more detailed microbial composition of all samples separately, these species were found as well and were clearly responsible for the prominent difference between KPC and CTRL feces (**figure 16d**).

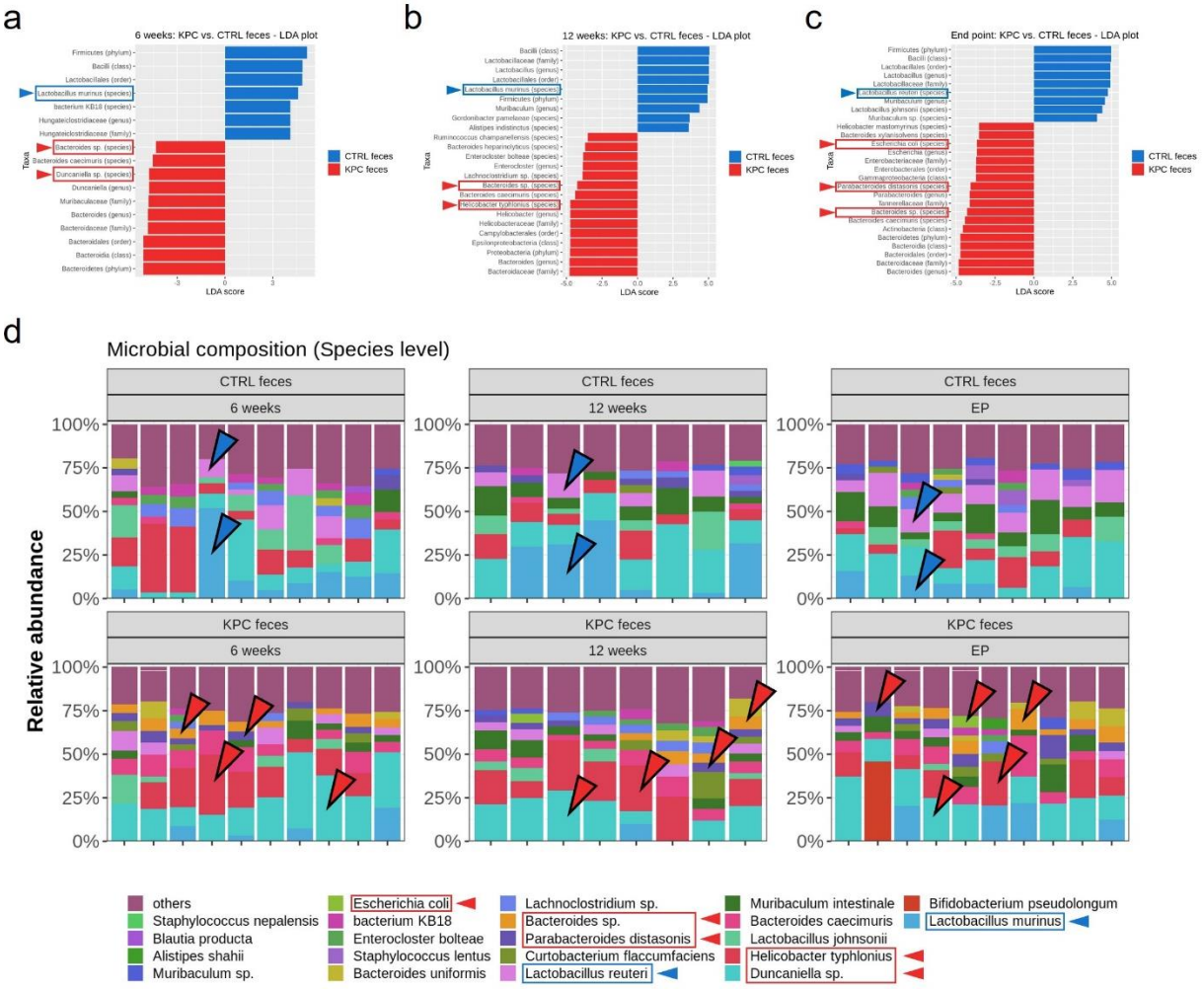

Figure 16: Investigation of the differences between KPC and CTRL fecal microbiomes over time. (a-c) Linear discriminant analysis (LDA) plots of KPC vs. CTRL fecal microbiomes of 6w, 12w, and EP mice. Taxa that reached an LDA score of >4 (6w) or >3.5 (12w, EP) are plotted. Species that also appear in the microbial composition plots in (d) are highlighted with red (KPC) and blue (CTRL) boxes and arrow heads. (d) Microbial compositions of KPC and CTRL fecal samples over time. Relative abundances of species in percent are plotted. Species that also appear in the LDA plots (a-c) are highlighted with red (KPC) and blue (CTRL) boxes and arrow heads.

The objection may be raised that the differences between KPC and CTRL feces are not due to the genotype but other confounding variables. Murine samples are more homogeneous than human samples and less confounded by varying clinical features or lifestyle habits. KPC and CTRL animals were housed together under the same conditions, yet, the variables sex and age (at least in the EP cohort) had to be considered as potential confounders.

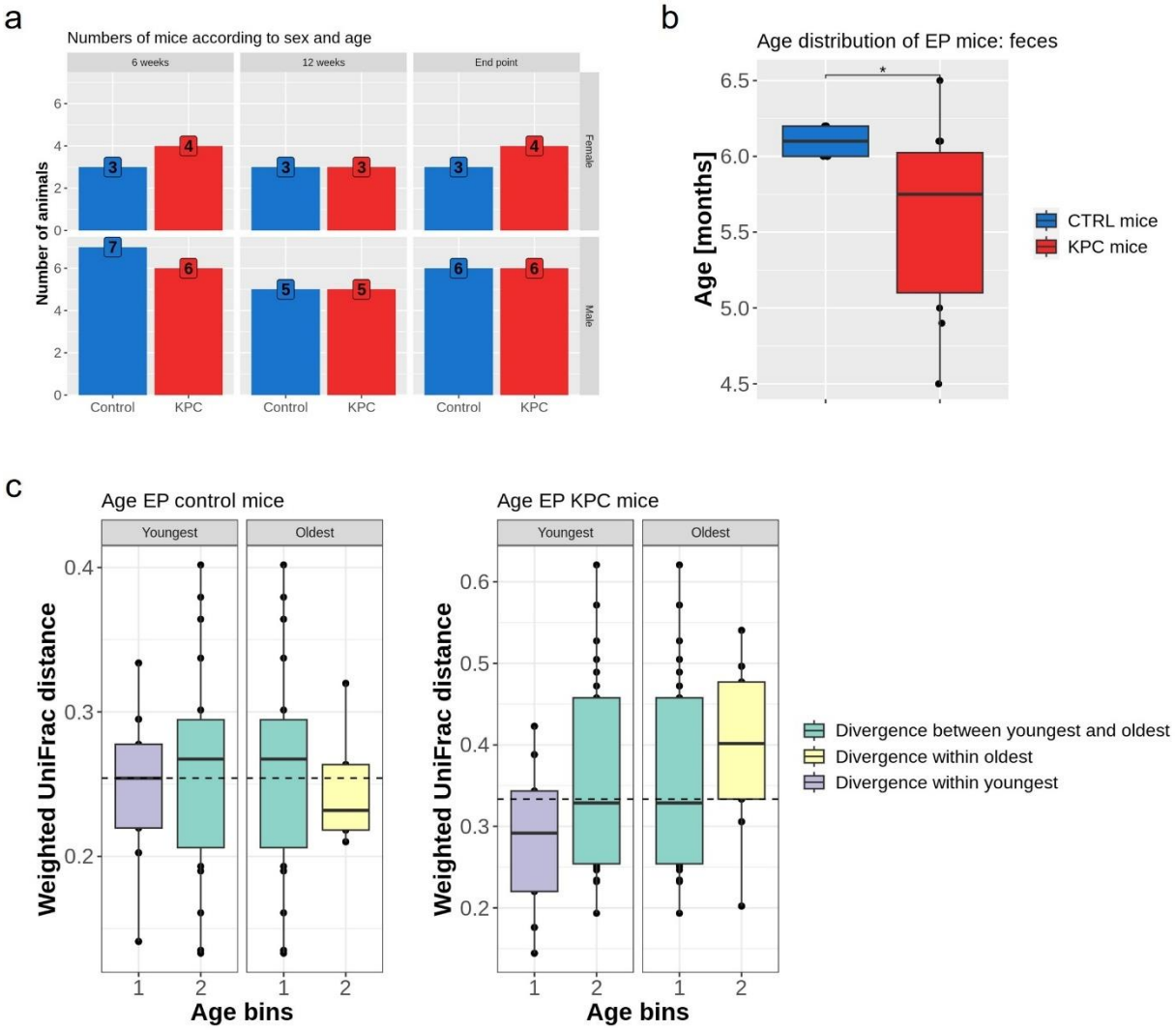

Figure 17: Confounder analysis in KPC vs. CTRL fecal microbiome data. (a) Age- and sex-separated numbers of KPC and CTRL mice from which feces are collected. (b) Ages in months of EP KPC and CTRL mice from which feces are collected. Mann-Whitney *U* test returns a significant difference between the age distributions of feces from KPC and CTRL EP mice. $p \leq 0.05$ *. (c) Weighted UniFrac distance in two separate age bins of EP KPC and CTRL mice. CTRL mice: youngest: 6.0-6.1 months, oldest: 6.2 months. KPC mice: youngest: 4.5-5.7 months, oldest: 5.8-6.5 months. The dashed line marks the overall median of all EP KPC and CTRL mice, respectively.

Sex could be excluded as possible confounder since Fisher's exact test was not significant between male and female mice in any age cohort. **Figure 17a** shows the equal distribution of males and females in both KPC and CTRL groups separated by age. However, testing ages in KPC and CTRL mice of the EP cohort against each other using the Mann-Whitney U test resulted in a significant difference (**figure 17b**). Thus, the influence of age on the beta diversity needed to be evaluated. Two equally distributed age subgroups (bins) were formed from both KPC and CTRL mice separately and pairwise wUF was determined. In **figure 17c,** the dashed line marks the general median across all KPC and CTRL mice, respectively. While the age did not have an impact on the microbiome of KPC mice, the divergence between the youngest and oldest CTRL mice was found to diverge from the general median across all groups. However, stratifying the data for both age groups and applying PERMANOVA still returned significant results for all distance metrics. It is therefore highly likely that beta diversity differences in KPC and CTRL feces do not result from factors sex or age.

4.2.2 KPC tumors harbor significantly more bacterial components than healthy control pancreas

The awareness of tumoral tissue being colonized by microbiota only arose in recent years and although it is increasingly accepted that these diseased organs harbor a microbiome, at least to a certain extent, many studies are questionable due to their sequencing methodology and way of dealing with contamination (Eisenhofer et al., 2019). Moving on to the tumoral microbiome of KPC mice for a comprehensive characterization of the KPC microbiome, first, evidence for bacterial presence in tumor tissue was provided. Staining KPC tumor and healthy murine pancreatic tissue via FISH targeting the 16S rRNA gene and via IHC targeting LPS proved the presence of bacterial components in these samples (**figure 18**). 16S rRNA probes identified intracellular, cytoplasmic localization of this bacterial marker gene in both tumor and healthy pancreas material (**figure 18a**), however, quantifying the stainings confirmed significantly more 16S rRNA in KPC tumor tissue (**figure 18b**). To control for unspecific binding, non-complementary probes were employed which did not show any light signal. Antibodies against LPS revealed distinct dot-shaped staining of this gram-negative bacterial surface marker molecule in the cytoplasm of particularly KPC tumor tissue

cells (**figure 18c**). Quantification demonstrated a significantly higher percentage of LPS-stained areas in these tissues (**figure 18d**).

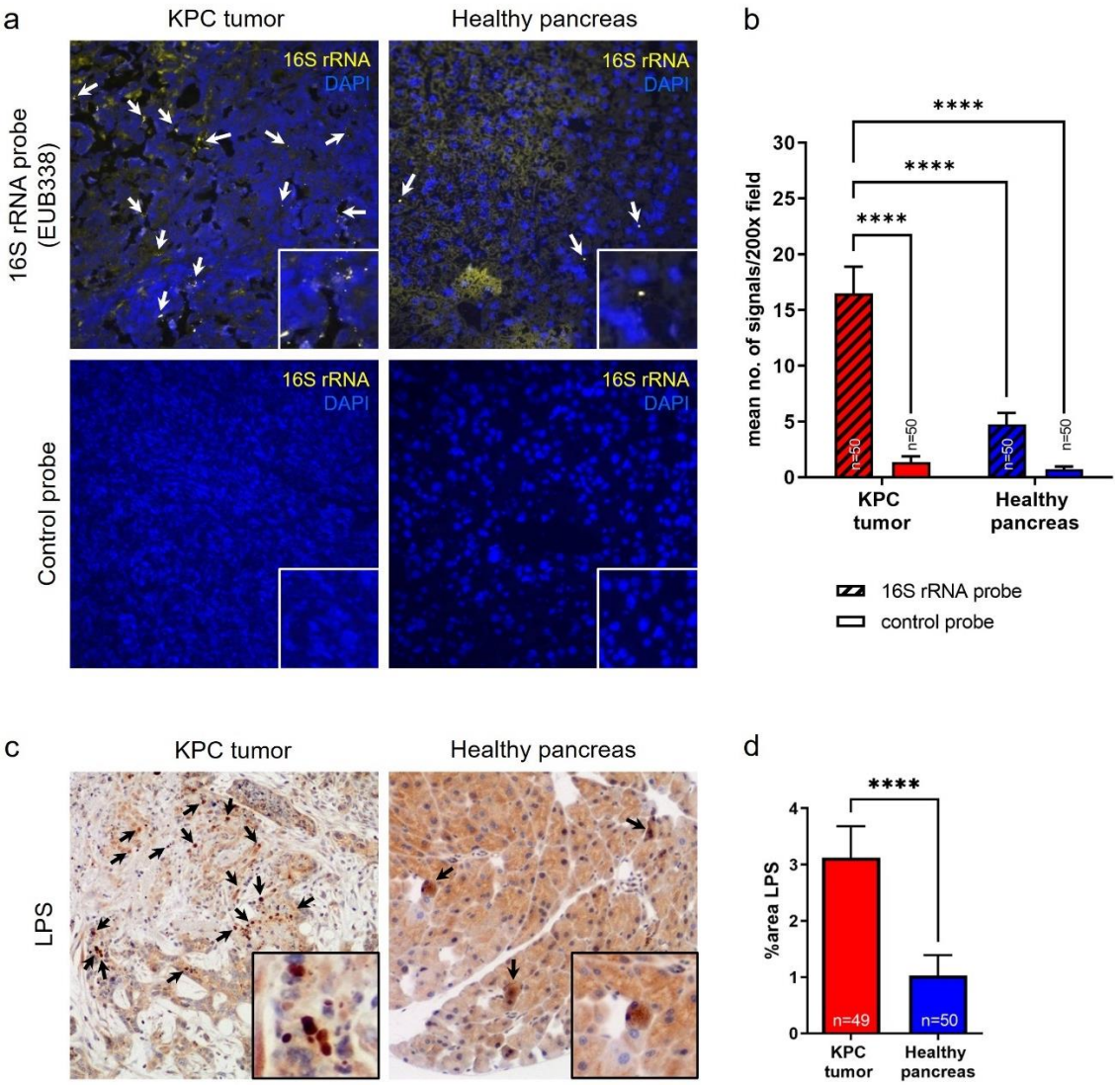

Figure 18: Visualization and quantification of bacterial components in KPC tumor and healthy pancreatic control tissue. (a) FISH of the 16S rRNA probe EUB338 and a non-complementary CTRL probe to DNA of KPC tumor and healthy pancreas FFPE tissue. Nucleic counterstaining with DAPI. White arrows indicate the yellow, dot-shaped 16S rRNA gene light signals against yellow background autofluorescence. (b) Mean number of 16S rRNA light signals of n 200x magnified image frames of KPC tumor and healthy pancreas. One slide of each of 10 mice per group was processed and 5 images of each slide were manually quantified resulting in the given n. (c) IHC staining of LPS on KPC tumor and healthy pancreas FFPE tissue. Black arrows indicate the reddish-brown, dot-shaped LPS antibody signals against the light brown background staining. Nucleic counterstaining with hematoxylin. (d) Mean LPS-stained area in percent of n 200x magnified image frames of KPC tumor and healthy pancreas. One slide of each of 10 mice per group was processed and 5 images of each slide were quantified by a color deconvolution macro in ImageJ resulting in the given n. Regarding the KPC group, from one slide only 4 frames could be acquired. All statistical testing performed with Mann-Whitney *U* test. $p \leq 0.0001$ ****.

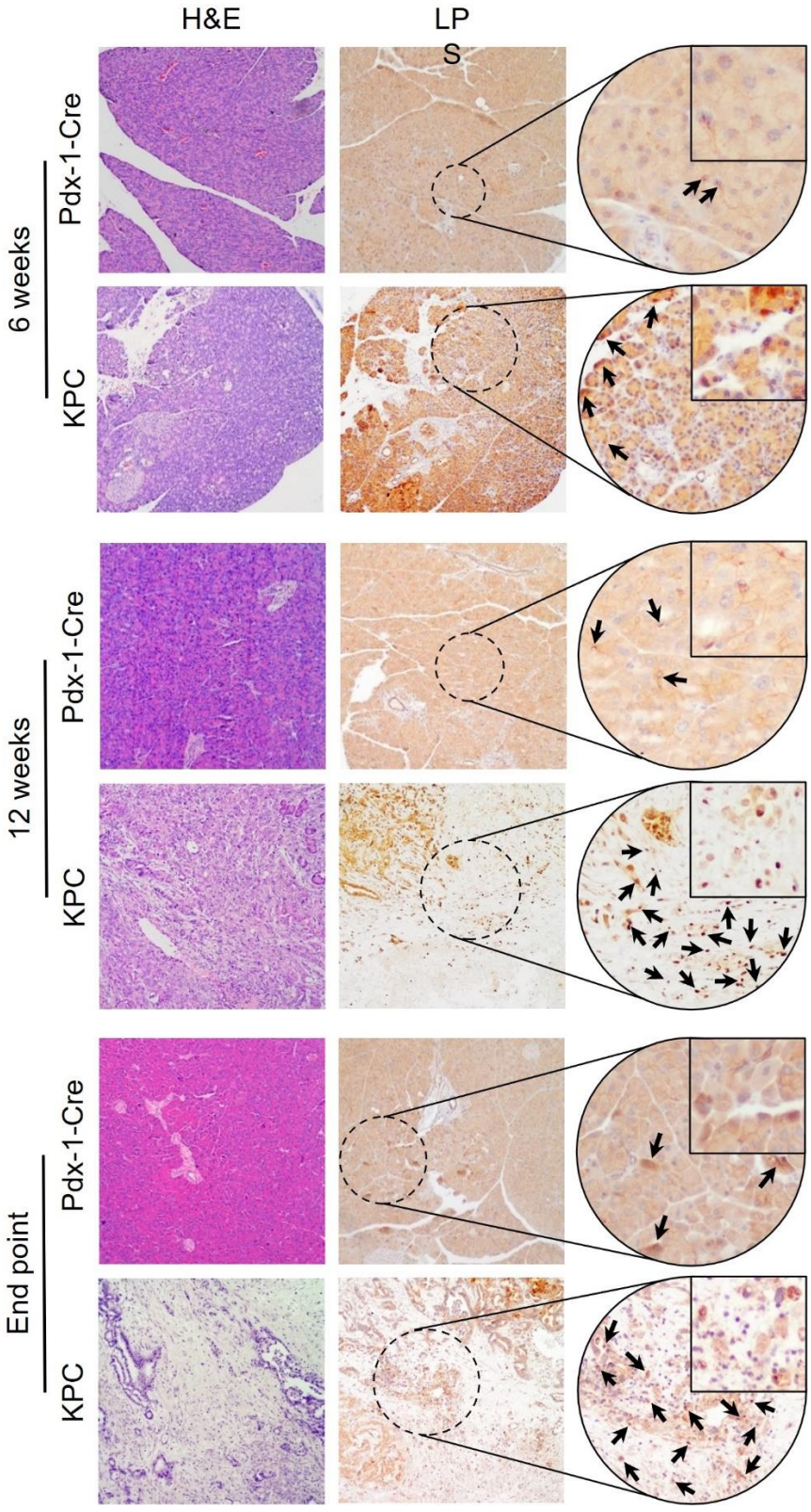

Figure 19: Pancreatic histology of KPC and CTRL (Pdx-1-Cre) mice. H&E staining and IHC targeting LPS on pancreatic tissue samples of 6w, 12w, and EP KPC and Pdx-1-Cre mice qualitatively demonstrate the development of pancreatic lesions and ultimately PDAC in KPC mice over time with simultaneous increasing invasion of gram-negative bacteria. In contrast, healthy Pdx-1-Cre mice are not affected to that extent. Black arrows indicate the reddish-brown, dot-shaped LPS antibody signals against the light brown background staining.

Retrospectively, it was also interesting to evaluate the presence of LPS in pancreatic tissue of 6w- and 12w-old KPC mice in comparison to CTRL mice. However, not enough tissue samples of animals sacrificed at 6w and 12w of age were available to quantify the stainings. Still, the observation made about KPC tumors harboring more LPS than respective CTRLs could be confirmed in 6w- and 12w-old KPC mice presenting with earlier stages of pancreatic tumors which can be seen in the corresponding H&E stainings (**figure 19**). In conclusion, quantitative differences between KPC tumor and healthy CTRL tissue were assessed regarding bacterial components 16S rRNA gene and LPS.

4.2.3 The KPC tumoral microbiome differs significantly from its normal pancreatic counterpart

Providing evidence for bacterial presence in both healthy and diseased murine pancreatic tissue was taken as basis to further extend the characterization of the KPC microbiome. The tumoral and healthy pancreatic microbiomes as two distinct microbial niches were analyzed in detail. KPC tumors of EP mice with a median age of 4 months and healthy pancreas samples from EP CTRL mice sacrificed at 6 months, each n = 15, were subject to 16S rRNA sequencing.

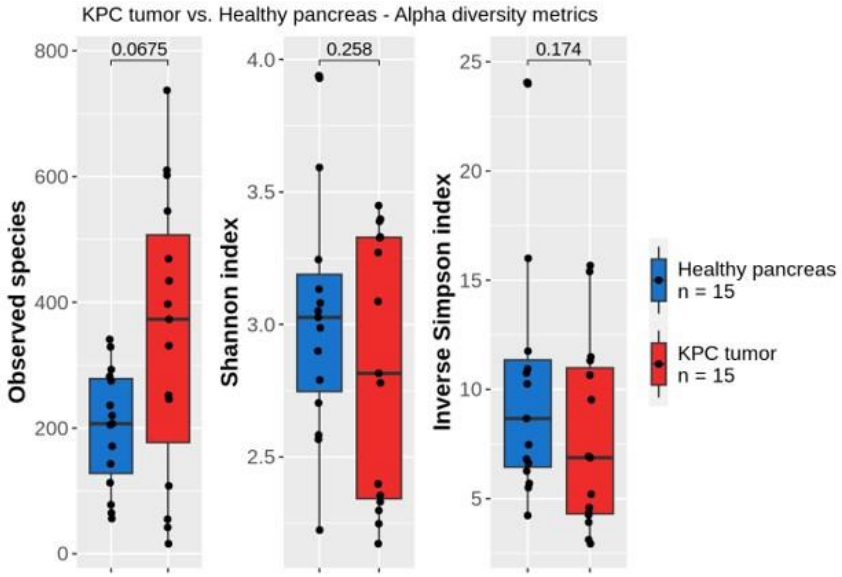

Figure 20: Alpha diversity in KPC tumor and healthy pancreas samples of age-matched CTRL mice. Numbers of observed species, Shannon indexes, and inverse Simpson indexes are not significantly different between KPC tumor tissue and healthy pancreas samples. All metrics calculated with non-rarified library sizes. All statistical testing performed with t-test or Wilcoxon rank test according to normal distribution test results.

As for KPC feces, no significant differences between KPC tumor and healthy CTRL samples were found regarding alpha diversity metrics OS, SI and InvS (**figure 20**). Regarding OS, a tendency towards statistical significance ($p = 0.0675$) may be identified. The normalization process to determine beta diversity metrics reduced the sample number to $n = 12$ KPC tumor and $n = 13$ healthy pancreas samples. Although the separation between KPC and CTRL pancreatic samples was not as clear as for feces, significant differences between these two groups in all three beta diversity metrics uwUF, wUF, and BC were found (**figure 21a-c**).

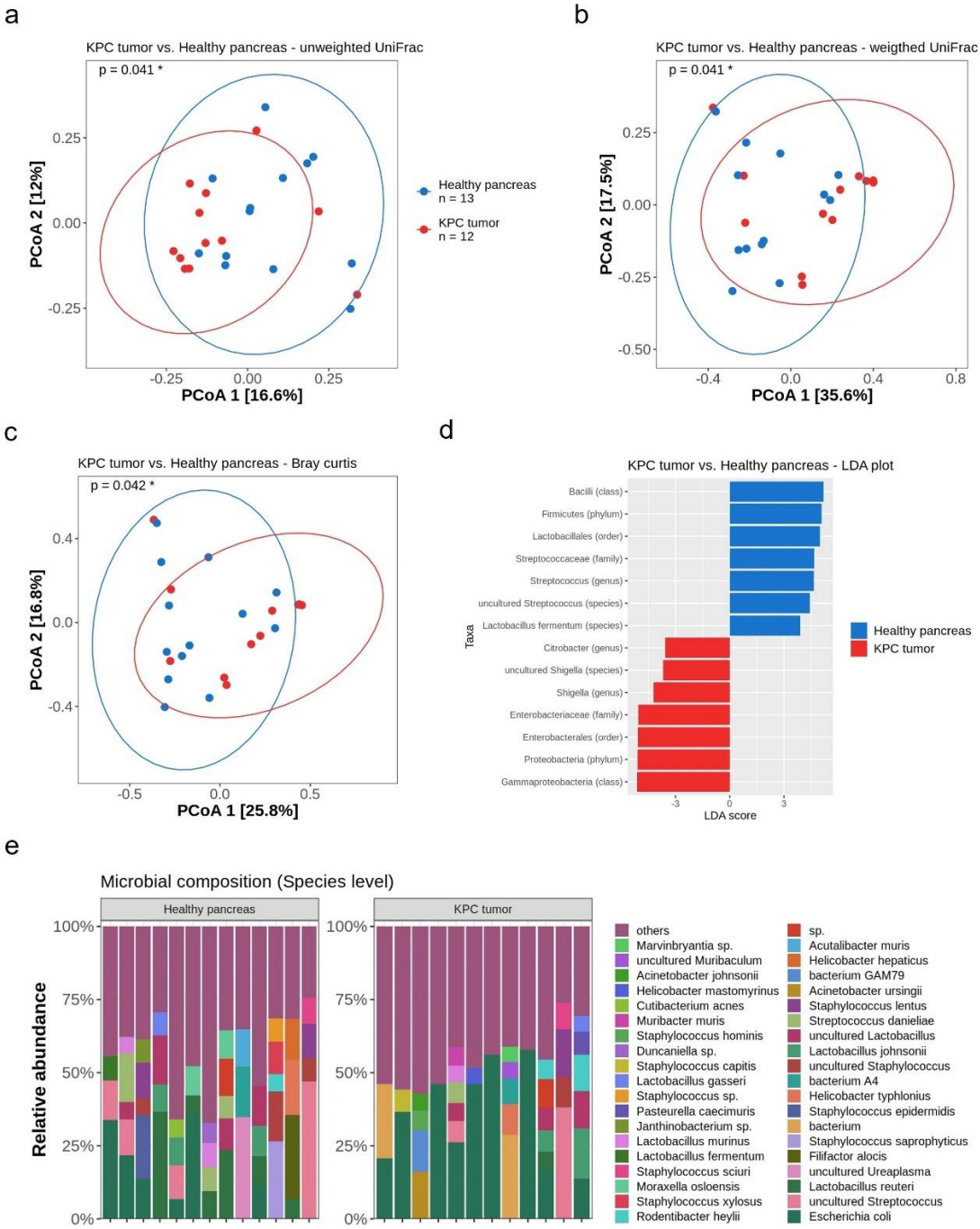

Figure 21: Differences of the microbiomes in KPC tumors and healthy control pancreas samples. (a-c) Beta diversity metrics uwUF, wUF, and BC show significant differences

between KPC tumor and healthy pancreas samples. Metrics calculated with rarified library sizes. Two healthy pancreas and three KPC tumor samples dropped by rarefaction. N-numbers indicated in graph legend in (a) apply to all three panels. Statistical testing performed with PERMANOVA. $p \leq 0.05$ *. (d) LDA plot of KPC tumoral vs. healthy pancreatic microbiomes. Taxa that reached an LDA score of >3.5 are plotted. (e) Microbial compositions of healthy pancreas and KPC tumor samples. Relative abundances of species in percent are plotted.

Looking at the microbial compositions of all samples, *Escherichia coli* appeared to a relatively large extent in both tumor and healthy samples (**figure 21e**). Thus, employing LEfSe, the differentially abundant taxa were again determined (**figure 21d**). Comparably to the fecal samples, *Lactobacillales* were the dominant order in murine healthy pancreas, in particular *Lactobacillus fermentum* and *Streptococcus*. In KPC tumor tissue, *Gammaproteobacteria*, comprising *Enterobacteriaceae* with *Shigella* and *Citrobacter*, were found to be more abundant. Eventually, as for the fecal data, potentially confounding factors had to be re-assessed. Notably, both Fisher's exact test examining sex and Mann-Whitney *U* test checking age returned no significant results. In **figure 22a**, the equally distributed numbers of male and female KPC and CTRL animals are depicted, **figure 22b** shows the age distribution of healthy and tumor mice with no significant difference between them. Therefore, sex and age could be excluded as confounding variables.

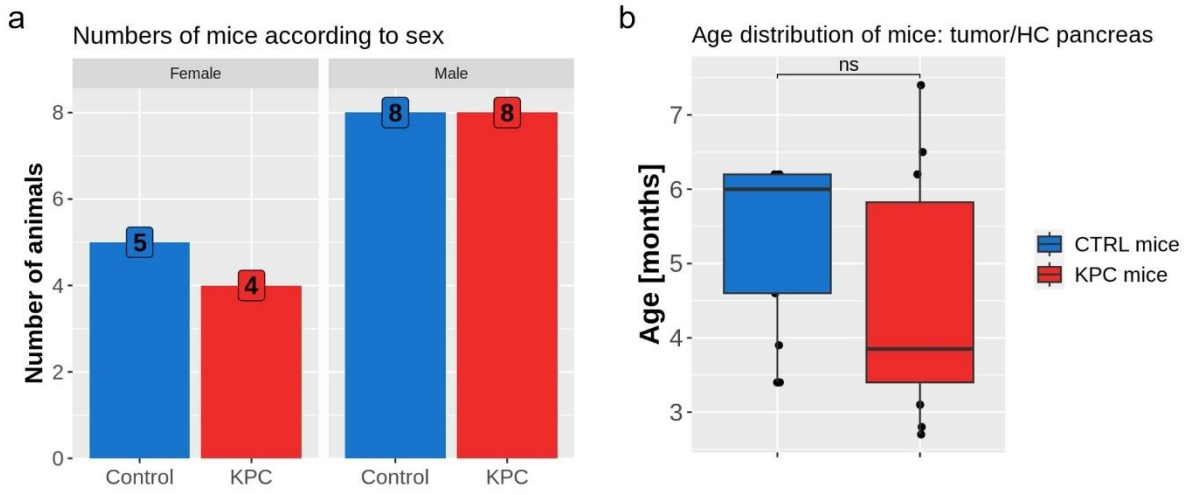

Figure 22: Confounder analysis in KPC tumoral vs. healthy control pancreatic microbiome data. (a) Sex-separated numbers of KPC and CTRL mice from which pancreatic samples were collected. (b) Ages in months of KPC and CTRL mice from which pancreatic samples were collected. Mann-Whitney *U* test did not return a significant difference between the age distributions of pancreatic samples from KPC and CTRL mice. ns, not significant.

4.2.4 The microbiomes of KPC tumor center and periphery are not significantly different from each other

The microstructure of a pancreatic tumor with its pronounced stroma and TME components is of exceptional interest for various reasons such as chemoresistance as elucidated in the introduction. Against this background, the logical consequence was to examine the KPC tumor itself in more detail regarding its microbiome. Macroscopically, two different tissue zones can be discriminated when cutting the solid tumor in half (**figure 23a**). The tumor periphery is well perfused, has a rather reddish color and is comparably translucent with a softer consistency, while the tumor center is firmer and fairly white with hardly any blood vessels. Both regions are clearly demarcated, so it was possible to separate them with a sterile scalpel and to store them in separate tubes. H&E stainings of entire cross sections impressively recapitulate the macroscopic appearance with a distinct border between the fibrotic and necrotic tumor center containing barely any stain and the active tumor periphery with many vital cells taking up the dyes (**figure 23b,c**).

After separating the tumor sub-compartments, they were analyzed by 16S rRNA sequencing. Alpha diversity metric OS reached significantly ($p < 0.01$) different numbers between tumor center and periphery. Notably, KPC tumor centers seem to harbor generally more species than the periphery (**figure 24a**). No significant differences were found regarding SI and InvS (**figure 24b,c**). Surprisingly, there were also no significant differences found between tumor center and tumor periphery in any of the assessed beta diversity metrics, uwUF, wUF, and BC (**figure 24d-f**). The microbial composition of all tumor center and periphery samples on species level reveals that *Escherichia coli* is contained in almost any sample indicating contamination issues (**figure 24g**). Besides *E. coli*, *Helicobacter spp.*, *Lactobacillus spp.* and *Staphylococcus spp.* were recurrently appearing species.

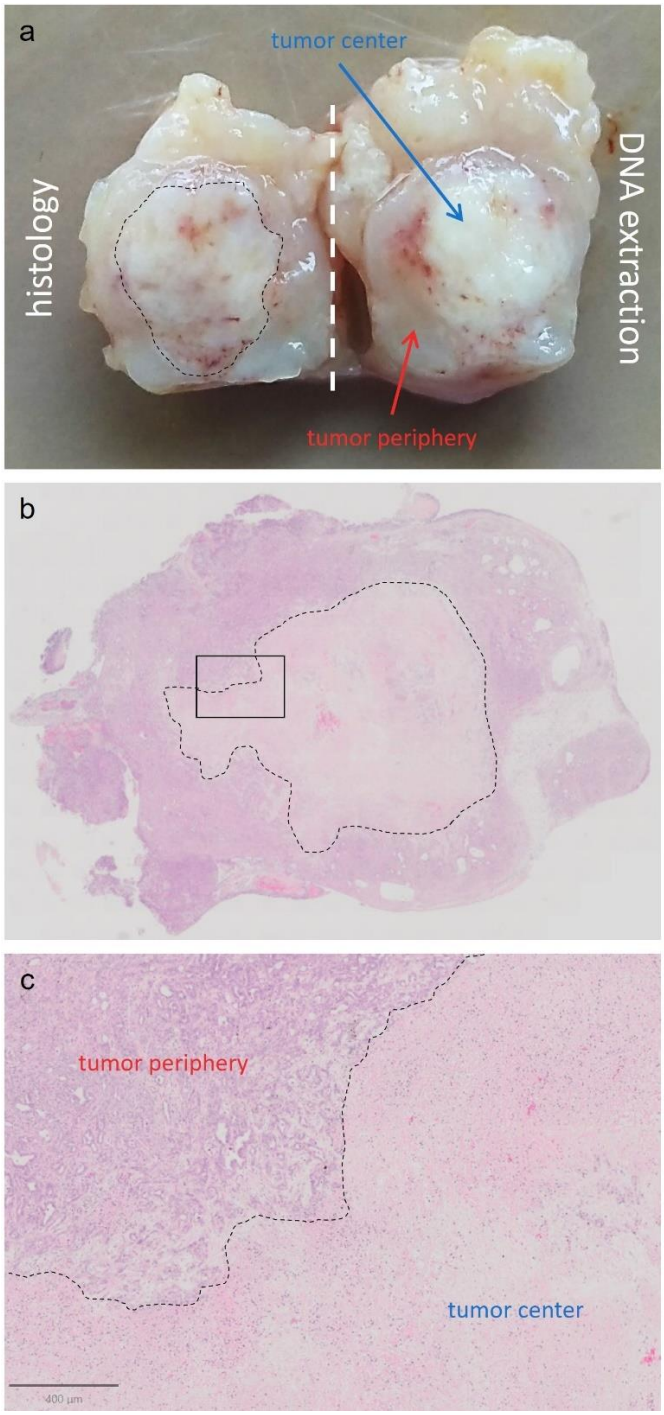

Figure 23: Macro- and microscopic KPC tumor anatomy. (a) Entire KPC tumor resected and cut in half as indicated by the vertical white dashed line. Tumor center (blue arrow) and tumor periphery (red arrow) are macroscopically distinguishable via their different colors and textures. The black dashed line outlines the solid white tumor center which is surrounded by the softer flesh-colored periphery. One half of the tumor was FFPE and used for histology, the other half was subject to compartment-separated DNA extraction. (b) H&E staining of a full cross section of an entire KPC tumor. The necrotic and fibrotic tumor center appearing light pink (eosin) is demarcated by a black dashed line. The black box marks the image detail magnified in (c). (c) Magnified image detail from H&E stained KPC tumor cross section. Histological tumor heterogeneity is clearly visible microscopically. The two zones are delineated by a black dashed line and labelled in red (tumor periphery) and blue (tumor center). The scale bar corresponds to 400 µm.

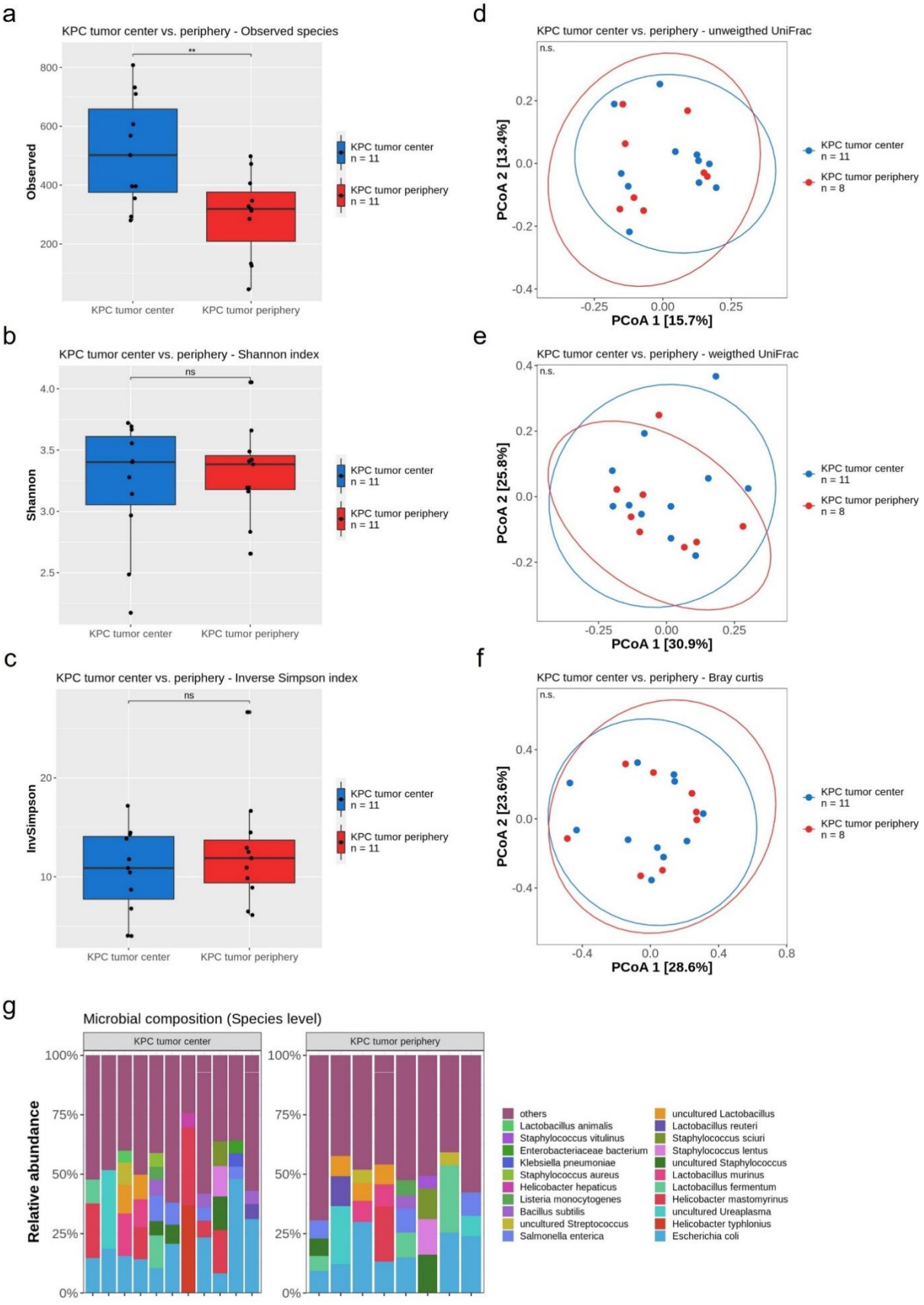

Figure 24: Analysis of the microbiomes in corresponding KPC tumor center and periphery samples. (a-c) Alpha diversity metrics OS, SI and InvS calculated with non-rarified library sizes. Observed species is significantly different between KPC tumor center and periphery samples. Shannon and inverse Simpson indexes are not

significantly different between the two groups. Statistical testing performed with t-test or Wilcoxon rank test according to normal distribution test results. **, $p \leq 0.01$. ns, not significant. (d-f) Beta diversity metrics uwUF, wUF, and BC do not show significant differences between KPC tumor center and periphery samples. Metrics calculated with rarified library sizes. Three KPC tumor periphery samples dropped by rarefaction. Statistical testing performed with PERMANOVA. n.s., not significant. (g) Microbial compositions of KPC tumor center and periphery samples. Relative abundances of species in percent are plotted.

4.2.5 Murine KPC fecal and tumoral microbiomes are similar to corresponding microbiomes of human PDAC patients

After studying locally distinct microbiomes in KPC and CTRL mice and collecting a large set of characterizing data as a profound basis for the future application of this mouse model in (functional) microbiome studies, the logical consequence is to evaluate the comparability of murine and human microbiome with regard to translational and clinical studies. The fecal and tumoral microbiomes were considered separately.

First, the murine feces microbiome data gathered in **4.1.1** was compared to recently published human PDAC patients' stool microbiome data. Kartal et al. and Nagata et al. provide metagenomic classifiers for stool samples from which species and genera were taken to perform L2 regularization-based logistic regression (Kartal et al., 2022; Nagata et al., 2022). Kartal and colleagues published 10 species/genera as their stool microbiota set in PDAC patients' stool which were narrowed down to 8 genera (**figure 25a**). Nagata and colleagues came up with a list of 16 microbes, whereby one article was termed "unknown cellular organism" which was omitted leaving 15 species and 9 genera to be compared for the murine fecal microbiome data (**figure 25a**). In these data, 4 genera and 1 species from the Kartal data were found as well as 4 genera and 3 species from the Nagata data. These taxa are printed in bold in **figure 25a**. The others were not prevalent in the here gathered murine fecal microbiome data. To assess the applicability of these adjusted human PDAC classifiers in order to predict disease status from murine feces in the KPC context, the receiver operating characteristic (ROC) was determined. The model was built for fecal samples of EP mice, as the human stool samples also stemmed from patients with manifested disease. This was done for both Kartal and Nagata data with the exception of the Kartal data on species level since only one of their listed species was confirmed in the murine feces data. The ROC curves

revealed area under the curve (AUC) values of 80% for Nagata et al. data on species level, 70% for Nagata et al. data on genus level, and 66.7% for Kartal et al. data on genus level (**figure 25b**).

a

Published microbial compositions in PDAC patients' stool					
Kartal et al., 2022	**Species extracted from Kartal et al., 2022**	**Genera extracted from Kartal et al., 2022**	**Nagata et al., 2022**	**Species extracted from Nagata et al., 2022**	**Genera extracted from Nagata et al., 2022**
Methanobrevibacter smithii	*Methanobrevibacter smithii*	*Methanobrevibacter*	*Streptococcus anginosus*	***Streptococcus anginosus***	***Streptococcus***
Veillonella atypica	*Veillonella atypica*	*Veillonella*	*Clostridium symbiosum*	*Clostridium symbiosum*	***Clostridium***
Firmicutes sp.	*Firmicutes sp.*	*Firmicutes*	unknown *Mogibacterium*	*Mogibacterium*	*Mogibacterium*
Clostridium sp.	***Clostridium sp.***	***Clostridium***	*Streptococcus oralis*	***Streptococcus oralis***	*Veillonella*
Bacteroides finegoldii	*Bacteroides finegoldii*	***Bacteroides***	*Clostridium clostridioforme*	*Clostridium clostridioforme*	*Sutterella*
bacterium LF-3	bacterium LF-3	*Alloscardovia*	*Veillonella atypica*	*Veillonella atypica*	*Actinomyces*
Alloscardovia omnicolens	*Alloscardovia omnicolens*	***Prevotella***	*Streptococcus vestibularis*	*Streptococcus vestibularis*	***Erysipelotrichaceae***
Prevotella species	*Prevotella species*	***Butyrivibrio***	*Sutterella wadsworthensis*	*Sutterella wadsworthensis*	***Hungatella***
Veillonella species	*Veillonella species*		*Actinomyces sp.* ICM39	*Actinomyces sp.* ICM39	*Anaerotruncus*
Butyrivibrio crossotus	*Butyrivibrio crossotus*		*Veillonella parvula*	*Veillonella parvula*	
			Clostridium boltae/clostridioforme	*Clostridium boltae*	
			Erysipelotrichaceae sp.	*Erysipelotrichaceae sp.*	
			Hungatella hathewayi	***Hungatella hathewayi***	
			unknown cellular organisms	*Anaerotruncus colihominis*	
			Anaerotruncus colihominis	*Streptococcus sp.* HSISM	
			Streptococcus sp. HSISM		

b

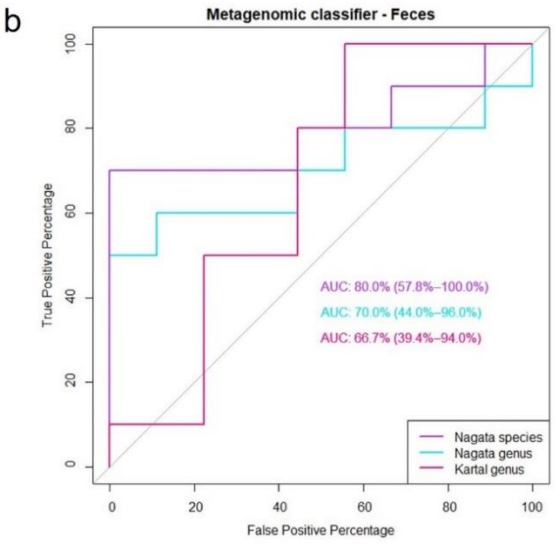

Figure 25: Comparison of murine KPC fecal microbiome data to published stool microbiome data of human PDAC patients. (a) Table containing bacterial species found in PDAC patients' stool samples by (Kartal et al., 2022) and (Nagata et al., 2022). Further, the species and genera extracted from these publications and used for the Ridge regression model in (b) are listed. Species and genera printed in bold are also found in the KPC fecal microbiome data. (b) Applying Ridge linear regression, the adjusted human metagenomic classifiers published by (Kartal et al., 2022) and (Nagata et al., 2022) were tested for their applicability on the murine KPC data in terms of disease prediction. ROC curves of the Ridge regression model of the KPC vs. CTRL fecal microbiome data are plotted. AUC values with confidence intervals are indicated in the plot.

Next, the murine tumor microbiome data gathered in **4.1.3** was compared to its corresponding CTRL data by means of data from Nejman et al. on human PDAC patients' tumor material (Nejman et al., 2020). Relative abundances of Nejman's most abundant genera and species in the murine data were determined. In n = 12 KPC tumors and n = 13 healthy pancreas samples, n = 3 genera (*Citrobacter*, *Enterobacter*, *Klebsiella*) and n = 4 species (*Citrobacter freundii*, *Enterobacter asburiae*, *Enterobacter cloacae*, *Klebsiella pneumoniae*) from Nejman et al. human PDAC tumors were found. Both genera and species from human PDAC were found to be significantly more abundant in KPC tumors than in healthy murine pancreas (**figure 26a,b**).

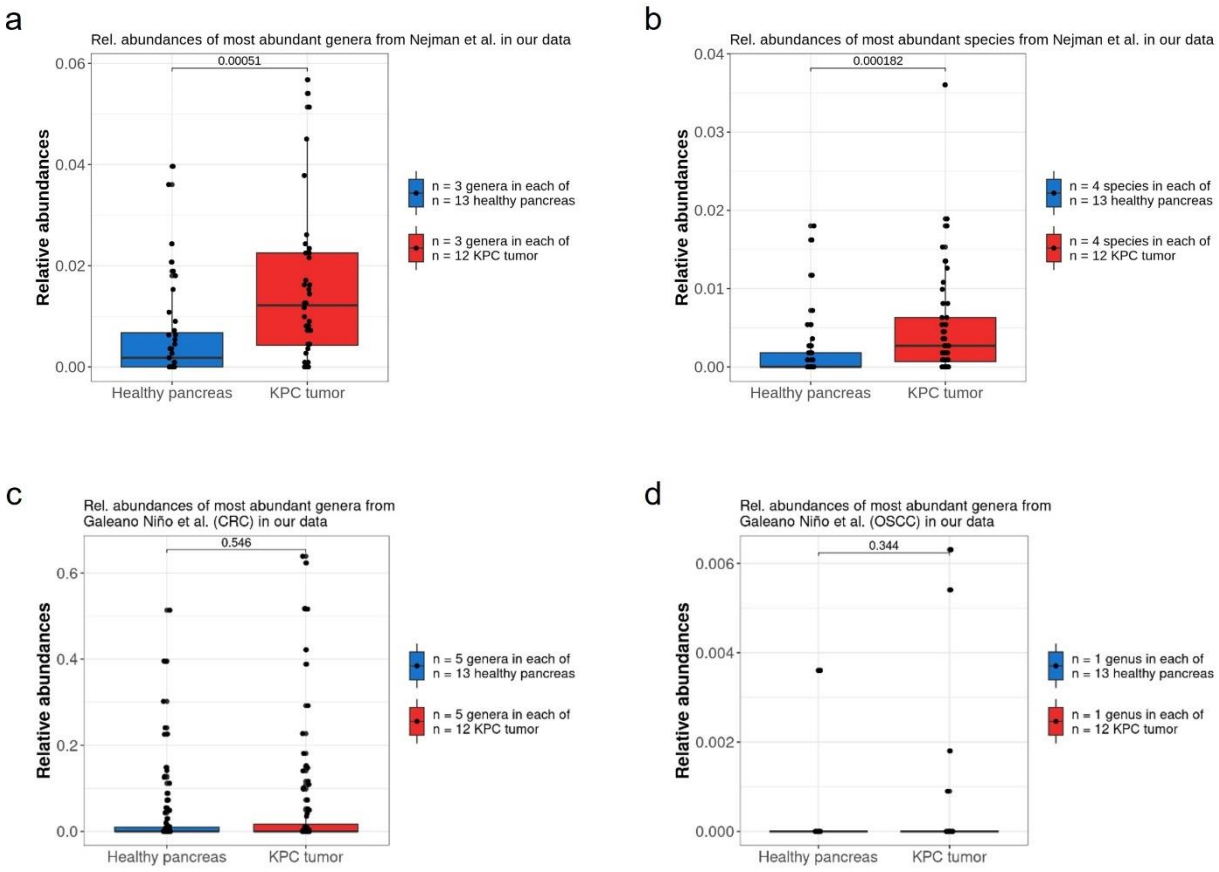

Figure 26: Comparison of the murine KPC tumoral and healthy pancreatic microbiome data to published human PDAC microbiome data. (a) Relative abundances of the most abundant genera from (Nejman et al., 2020) in the murine pancreatic microbiome data; n = 3 genera per sample. (b) Relative abundances of the most abundant species from (Nejman et al., 2020) in the murine pancreatic microbiome data; n = 4 species per sample. (c) Relative abundances of the most abundant genera found in CRC samples from (Galeano Niño et al., 2022) in the murine pancreatic microbiome data; n = 5 genera per sample. (d) Relative abundances of the most abundant genera found in OSCC samples from (Galeano Niño et al., 2022) in the murine pancreatic microbiome data; n = 1 genus per sample. All statistical testing performed with Mann-Whitney *U* test.

Eventually, after demonstrating the presence of human PDAC microbiota in murine KPC tumors, it was interesting to compare the microbiota of other cancer entities to the murine microbiome data. To this end, the 10 most dominant genera from recently published human colorectal carcinoma (CRC) and oral squamous cell carcinoma (OSCC) microbiome data by Galeano Niño et al. were assessed for their prevalence in the murine KPC tumors and healthy pancreas samples, respectively (Galeano Niño et al., 2022). In n = 12 KPC tumors, n = 6 CRC genera were found, and another n = 6 genera in n = 13 healthy pancreas samples. In both KPC tumors and healthy pancreas, an overlap of n = 5 CRC genera (*Bacteroides*, *Escherichia*, *Gemella*, *Shigella*, *Streptococcus*) were found. Regarding the OSCC genera, n = 2 appeared in KPC tumors, and thereof only n = 1 genus (*Alistipes*) was detected in healthy pancreas. These genera present in KPC tumors and healthy pancreas did not differ significantly in their relative abundances between both tissue groups (**figure 26c,d**). In short, genera from other cancer entities are not, if at all, differentially abundant in KPC tumor and healthy pancreas tissue.

4.3 Chapter 3 | The microbiome of tumors derived from orthotopically transplanted human PDAC cell lines of different subtypes

4.3.1 Capan-2 and MiaPaCa2 tumors do not differ in their tumoral but normal adjacent pancreatic microbiomes

Creating the basis for preclinical PDAC microbiome research by characterizing transgenic mouse model systems entailed investigations of xenograft PDAC models with a stronger relation to the human disease. To this end, FFPE tumors with adjacent normal tissue from orthotopically transplanted Capan-2 and MiaPaCa2 human PDAC cells were obtained. The former is a representative of the CLA PDAC subtype, the latter belongs to the BL variety. The comparison of their microbiomes aims at understanding the dynamics and interactions between tumor cells and TME, and the impact of the subtype. With the help of H&E stainings, tumor and histologically normal adjacent pancreatic tissue were delineated from each other and investigated separately. **Figure 27a** illustrates the experimental design with exemplary H&E-stained tissue sections. Tissue samples were obtained in FFPE blocks at which step the workflow was executed within the work frame of this thesis.

Since the samples stemmed from another study, there were only limited animals available that had not received treatment other than non-toxic control treatment such as H_2O. While n = 10 MiaPaCa2 tumor samples were available from 5 of which additionally adjacent normal pancreatic tissue was recovered, only n = 5 Capan-2 tumor samples could be attained from all of which adjacent normal pancreatic tissue was sampled. Isolating DNA from these samples and analyzing their microbiomes via 16S rRNA gene sequencing, revealed no significant differences in alpha diversity metrics OS, SI, and InvS between the four groups (**figure 27b-d**), except for the observed number of species in MiaPaCa2 normal adjacent tissue and MiaPaCa2 tumor tissue which differed significantly from one another (**figure 27b**). MiaPaCa2 normal adjacent tissues showed the highest median number of species with up to around 150 in one sample, whereas the OS numbers of all other groups resided below 50 (**figure 27b**).

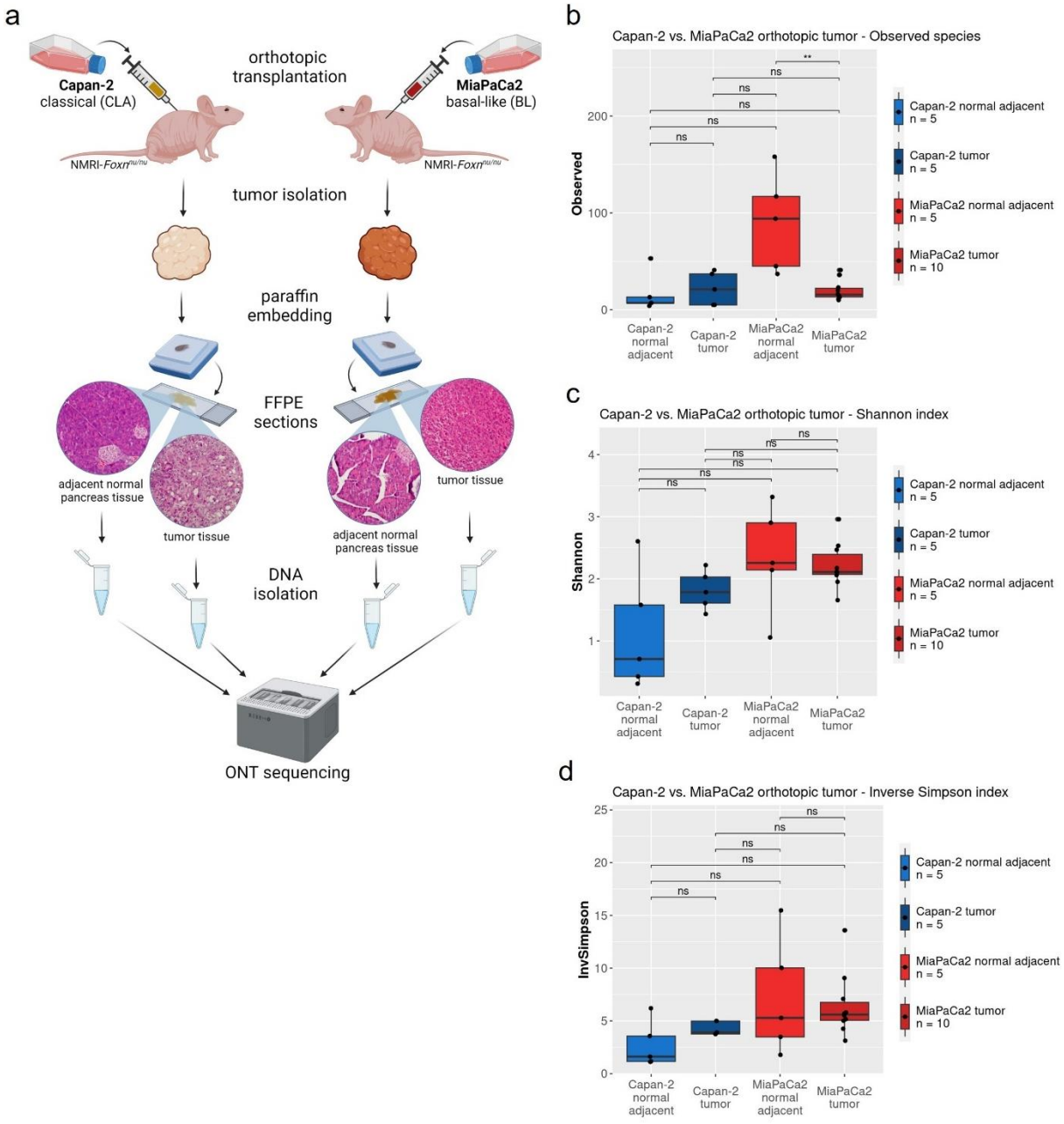

Figure 27: Investigation of the microbiomes in Capan-2 and MiaPaCa2-derived orthotopic tumor and tumor-adjacent histologically normal pancreatic tissue samples. (a) Experimental design of pancreatic microbiome analysis employing xenograft PDAC mouse models with tumors derived from orthotopically transplanted human Capan-2 and MiaPaCa2 cells; with exemplary H&E stainings of the four regarded tissue groups. DNA was extracted from FFPE tissue sections and subject to 16S rRNA gene sequencing. (b-d) Alpha diversity metrics OS, SI, and InvS are not significantly different between the four regarded tissue groups, except for the number of observed species between MiaPaCa2-derived tumor and normal adjacent tissue samples. All metrics calculated with non-rarified library sizes. All statistical testing performed with t-test or Wilcoxon rank test according to normal distribution test results. **, $p \leq 0.01$. ns, not significant.

The decontamination and normalization processes were more difficult for the FFPE data sets as compared to the data from self-sampled, snap-frozen tissue. In the case of long-term stored FFPE tissues from other sources, the contamination issue is even stronger and requires very stringent decontamination approaches. Besides a substantial number of NTCs, an additional filter pipeline developed by the Department of Medical Bioinformatics was applied to the FFPE data phyloseq objects on top of the standard procedure described in the methods **3.2.13**. This procedure resulted in a rigorous removal of alleged contaminants leaving only very few reads and species per sample, some even in the one-digit range. This outcome in turn required an unusually low rarefaction threshold in order to keep the already low number of samples. Calculating beta diversity metrics uwUF, wUF, and BC to assess differences between the groups was done for each tumor and tissue origin groups separately. Neither Capan-2 nor MiaPaCa2 tumors showed significant clustering away from their respective normal adjacent pancreatic counterparts in any of the three beta diversity metrics (**figure 28a-f**). Surprisingly, Capan-2 and MiaPaCa2 tumors did not show significant differences in any of the beta diversity metrics in question either (**figure 28j-l**). Most interestingly however, normal pancreatic tissue samples derived from next to Capan-2 and MiaPaCa2 tumors, respectively, differed significantly from each other regarding uwUF and wUF (**figure 28g,h**). BC could not be plotted properly using PCoA, the multidimensional scaling method of choice to visualize the dissimilarity, due to the data properties. Thus, non-metric multidimensional scaling (NMDS) was applied as the best alternative to plot the data. Independently, the two groups did not reach statistical significance regarding BC (**figure 28i**).

Eventually, the microbial composition was of major interest for an even more in-depth understanding of the data set. **Figure 29** illustrates the low number of species left in each sample after rarefaction. Among these few bacterial species per sample, environmental contaminants, such as *Bradyrhizobium*, *Curtobacterium*, and *Haliangium*, stand out in the color-coded legend. Within one tissue group, great species variation can be observed, MiaPaCa2 tumors being the only group repeatedly presenting *Bradyrhizobium vignae* and *Bradyrhizobium guangxiense* as a discernible pattern but also being the tissue group with the most samples. Overall, many species were identified, however, basically every sample has a different bacterial composition.

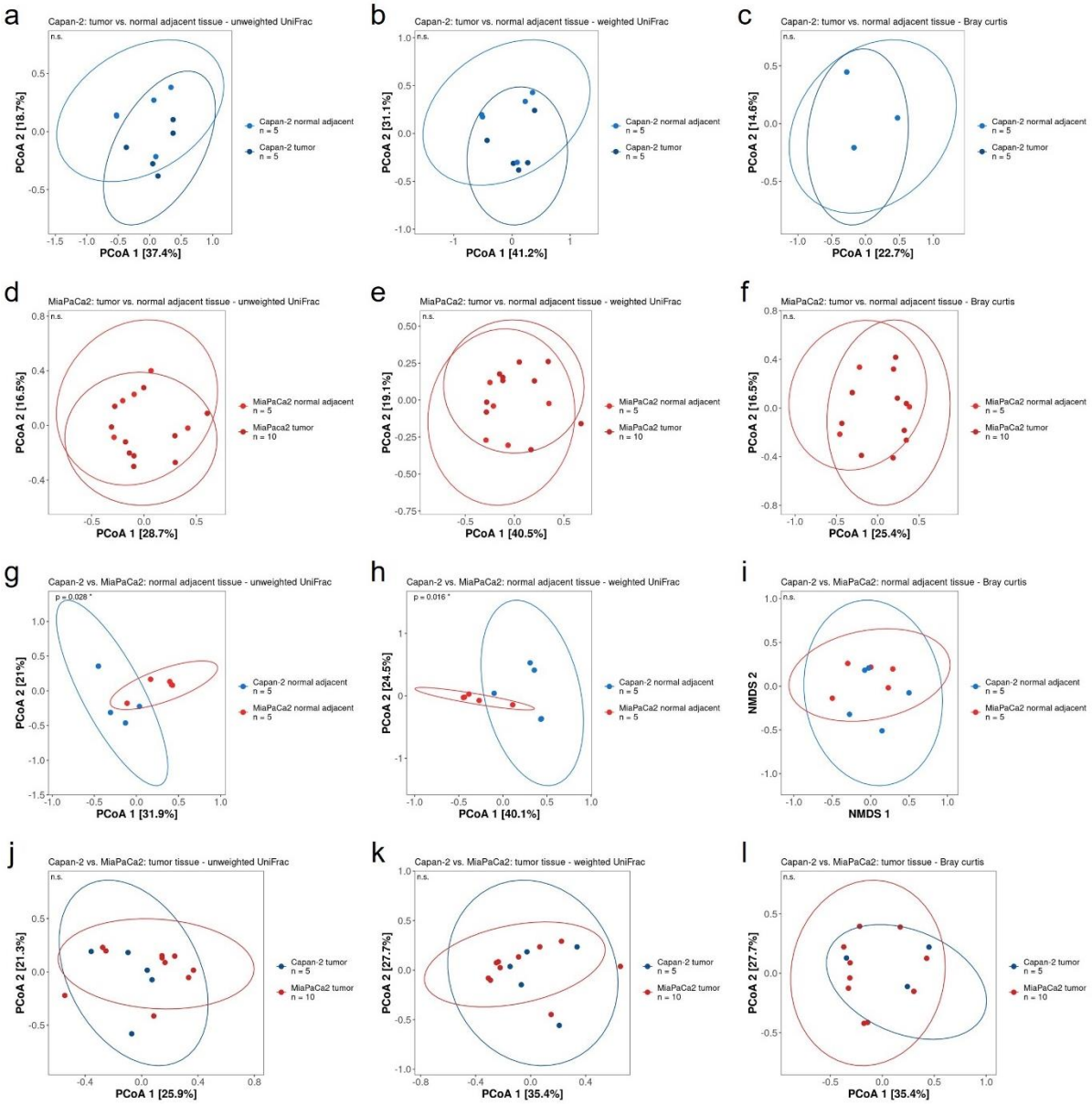

Figure 28: Beta diversity between Capan-2 and MiaPaCa2-derived orthotopic tumor and tumor-adjacent histologically normal pancreatic tissue samples. (a-c) Beta diversity metrics uwUF, wUF, and BC do not show significant differences between Capan-2-derived tumor and normal adjacent tissue samples. (d-f) Beta diversity metrics uwUF, wUF, and BC do not show significant differences between MiaPaCa2-derived tumor and normal adjacent tissue samples. (g-i) Beta diversity metrics uwUF and wUF show significant differences between Capan-2- and MiaPaCa2-derived normal adjacent tissue samples. BC, plotted with NMDS instead of standard PCoA, does not show significant differences between these two tissue groups. (j-l) Beta diversity metrics uwUF, wUF, and BC do not show significant differences between Capan-2- and MiaPaCa2-derived tumor tissue samples. All metrics calculated with rarified library sizes. All statistical testing performed with PERMANOVA. *, $p \leq 0.05$. n.s., not significant.

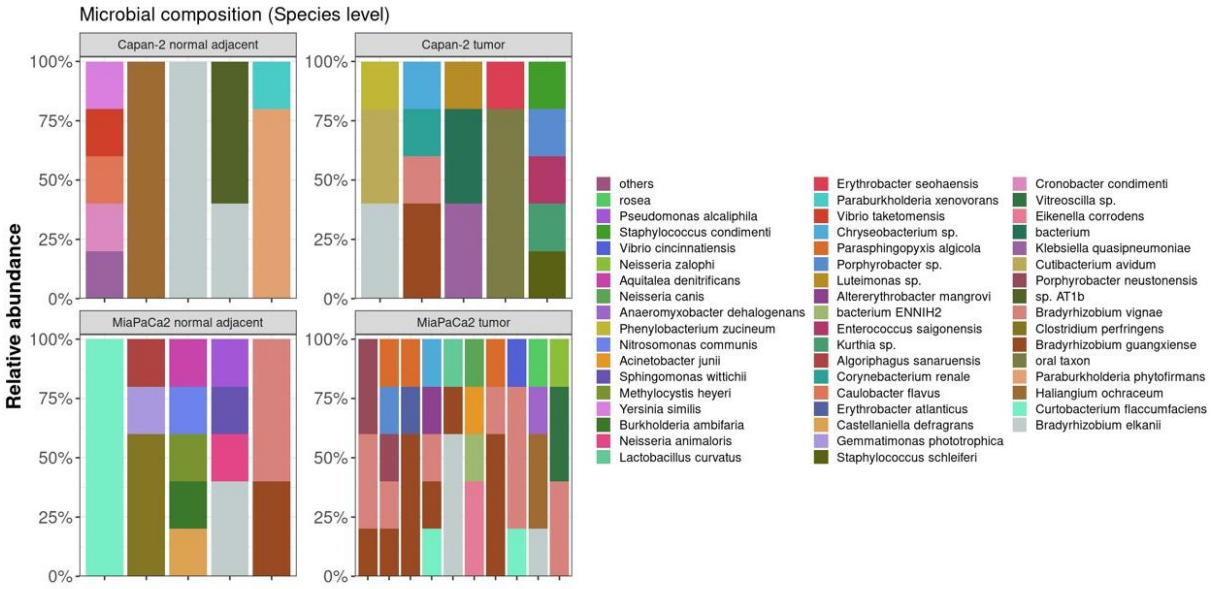

Figure 29: Microbial compositions of Capan-2 and MiaPaCa2-derived orthotopic tumor and tumor-adjacent histologically normal pancreatic tissue samples. Relative abundances of species in percent are plotted. Sequencing data passed extensive filter processes, therefore leaving only one to five different taxa per sample.

4.3.2 2D-cultured Capan-2 and MiaPaCa2 cells do not harbor different microbiomes

Being aware of the issues with the FFPE Capan-2/MiaPaCa2 data set, the following control experiment was even more important. Although cell culture is performed under the highest sterile conditions possible, even in S1 labs with the lowest safety standards, contamination must be considered. Also, the moment the cells are transplanted from the seemingly sterile environment of 2D cell culture flasks to the recipient animal, the introduction of contaminants cannot be excluded. Furthermore, analyzing Capan-2 and MiaPaCa2 cells aimed at clarifying the origin of differences between the tissue groups based on the assumption that the microbes found in tumor samples are mouse-born. For these reasons, DNA from both cell lines was extracted in 5 biological replicates each and tested for bacteria via 16S rRNA gene sequencing. NTCs were introduced in an almost 3:1 ratio, outnumbering the true samples by far.

With no surprise, bacterial DNA was detected by Nanopore sequencing in all samples. Like the FFPE tissue samples, the cell data was subject to contamination correction by applying the additional filter pipeline. Following the standard analysis pipeline next, alpha diversity

metrics were calculated using non-rarified phyloseq objects. There were no significant differences found between Capan-2 and MiaPaCa2 cells in neither OS, SI, nor InvS (**figure 30a-c**).

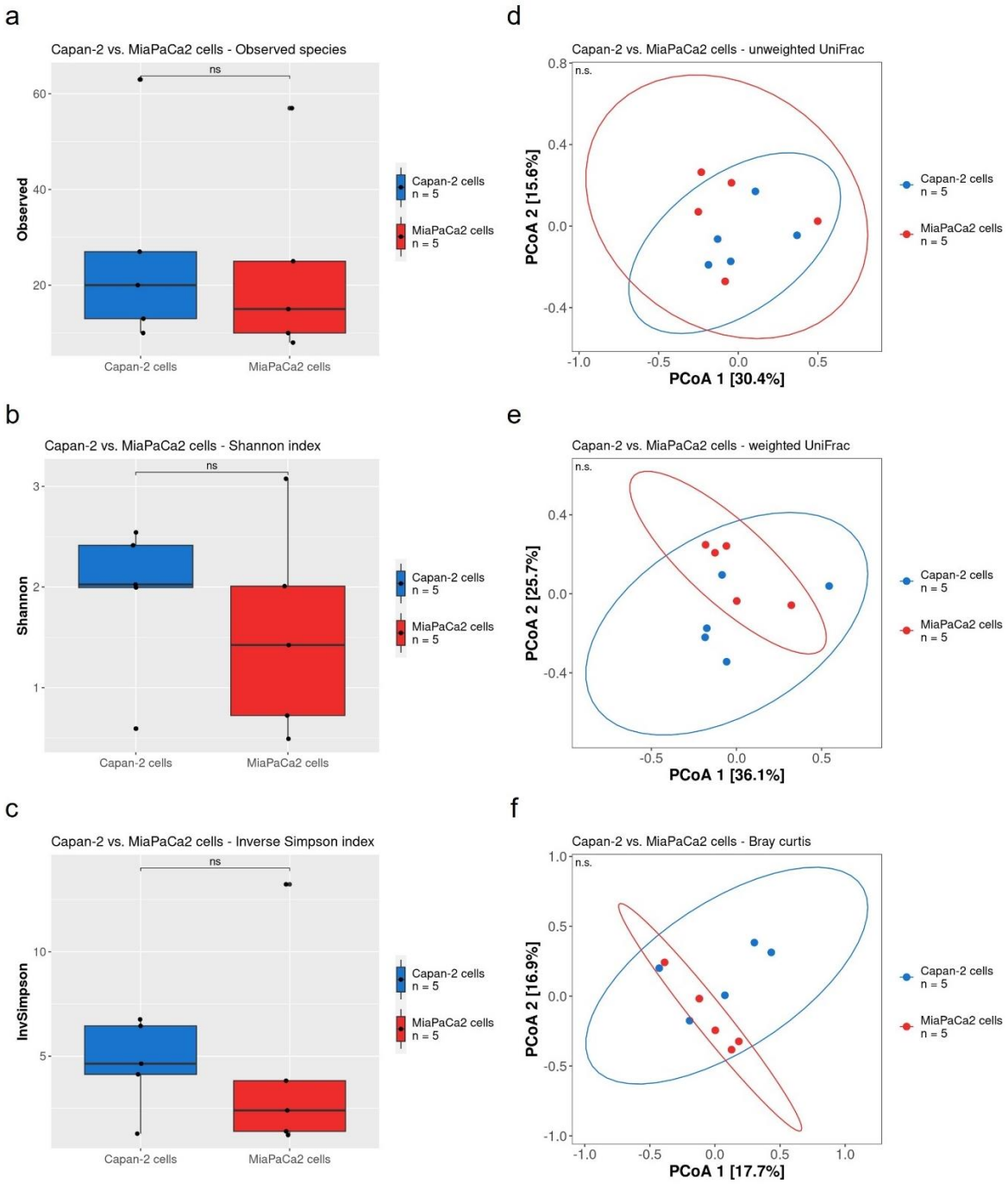

Figure 30: Analysis of the microbiomes of 2D-cultured Capan-2 and MiaPaCa2 cells. (a-c) Alpha diversity metrics OS, SI, and InvS are not significantly different between Capan-2 and MiaPaCa2 cells. Metrics calculated with non-rarified library sizes. Statistical testing performed with t-test or Wilcoxon rank test according to normal distribution test results. ns, not significant. (d-f) Beta diversity metrics uwUF, wUF, and BC do not show significant differences between Capan-2 and MiaPaCa2 cells. Metrics calculated with rarified library sizes. Statistical testing performed with PERMANOVA. n.s., not significant.

The median observed number of species was 20 and 15 for Capan-2 and MiaPaCa2 cells, respectively, after the extreme filtering process which again caused the drop of many reads requiring an unusually low rarefaction threshold to keep all samples. None of the beta diversity metrics, neither uwUF, wUF, nor BC, showed significant differences between Capan-2 and MiaPaCa2 cells (**figure 30d-f**). Regarding the species identified in the cell culture samples, the composition exhibits great variation among all samples, also within one cell line, and the vast amount of environmental contaminating bacteria is evident (**figure 31**). For instance, *Bradyrhizobium vignae* and *Bradyrhizobium guangxiense* are again found. In addition, *Sphingomonas*, *Sphingobium* and *Novosphingobium*, known as typical environmental bacteria, appear in the microbial composition.

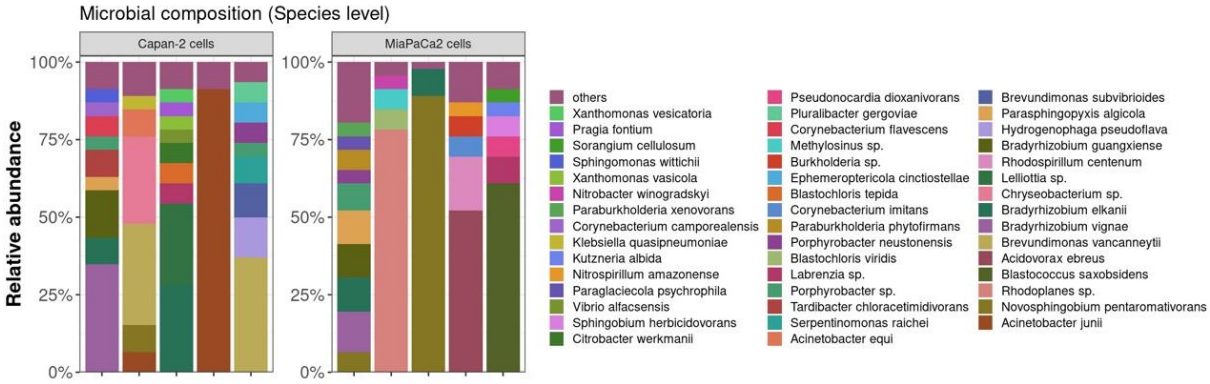

Figure 31: Microbial compositions of 2D-cultured Capan-2 and MiaPaCa2 cells. Relative abundances of species in percent are plotted.

4.3.3 The microbiomes of TNFα-treated Capan-1 tumors differ in parts from vehicle-treated controls

Another available sample set subject to investigation consisted of xenograft mouse model tumors derived from orthotopically transplanted human Capan-1 cells. Like Capan-2, they are representatives of the CLA PDAC subtype. After tumor development, the transplanted mice received either TNFα or H_2O in two cohorts for three weeks followed by sacrifice, tumor isolation and paraffin embedding. The study from which these samples derived found the CLA Capan-1 tumors to become more BL after TNFα treatment (Tu et al., 2021). These tissues were analyzed for their microbiomes via 16S rRNA gene sequencing in order to learn more about the subtype as determinator of the microbiome. **Figure 32** illustrates the

experimental design with exemplary H&E-stained tissue sections of the FFPE blocks at which step the workflow was executed within the work frame of this thesis.

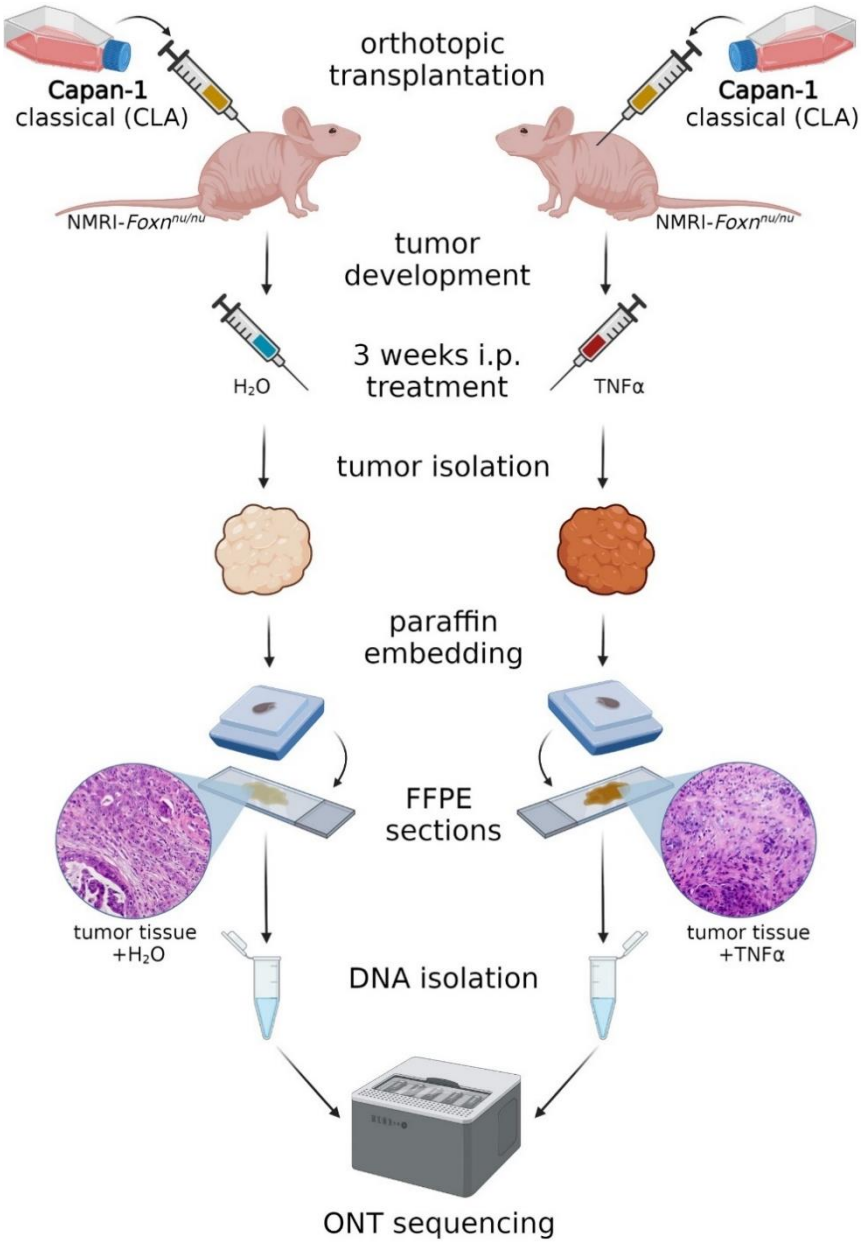

Figure 32: Experimental design of pancreatic microbiome analysis in TNFα/H_2O-treated Capan-1 xenograft PDAC mouse models. Mice orthotopically transplanted with human Capan-1 cells were intraperitoneally (i.p.) treated with TNFα or its vehicle control H_2O, respectively, to introduce lineage reprogramming prior to tumor isolation. Exemplary H&E stainings of the two tissue groups. DNA was extracted from FFPE tissue sections and subject to 16S rRNA gene sequencing.

Following the standard analysis pipeline, the double filtered FFPE samples sequencing data were unrarefied employed to calculate alpha diversity metrics. Both treatment groups were composed of n = 8 tumor samples. According to statistical testing, there were no significant

differences between TNFα- and H_2O-treated Capan-1 tumor microbiomes in OS, SI and InvS (**figure 33a-c**). However, noticeably, the variance in H_2O-treated Capan-1 tumors was always greater than in TNFα-treated tissues, and at least regarding OS, the median diverged visibly between the two groups. Capan-1 tumors +H_2O presented with a median OS number of about 55, Capan-1 tumors +TNFα had a median OS of just under 80 (**figure 33a**).

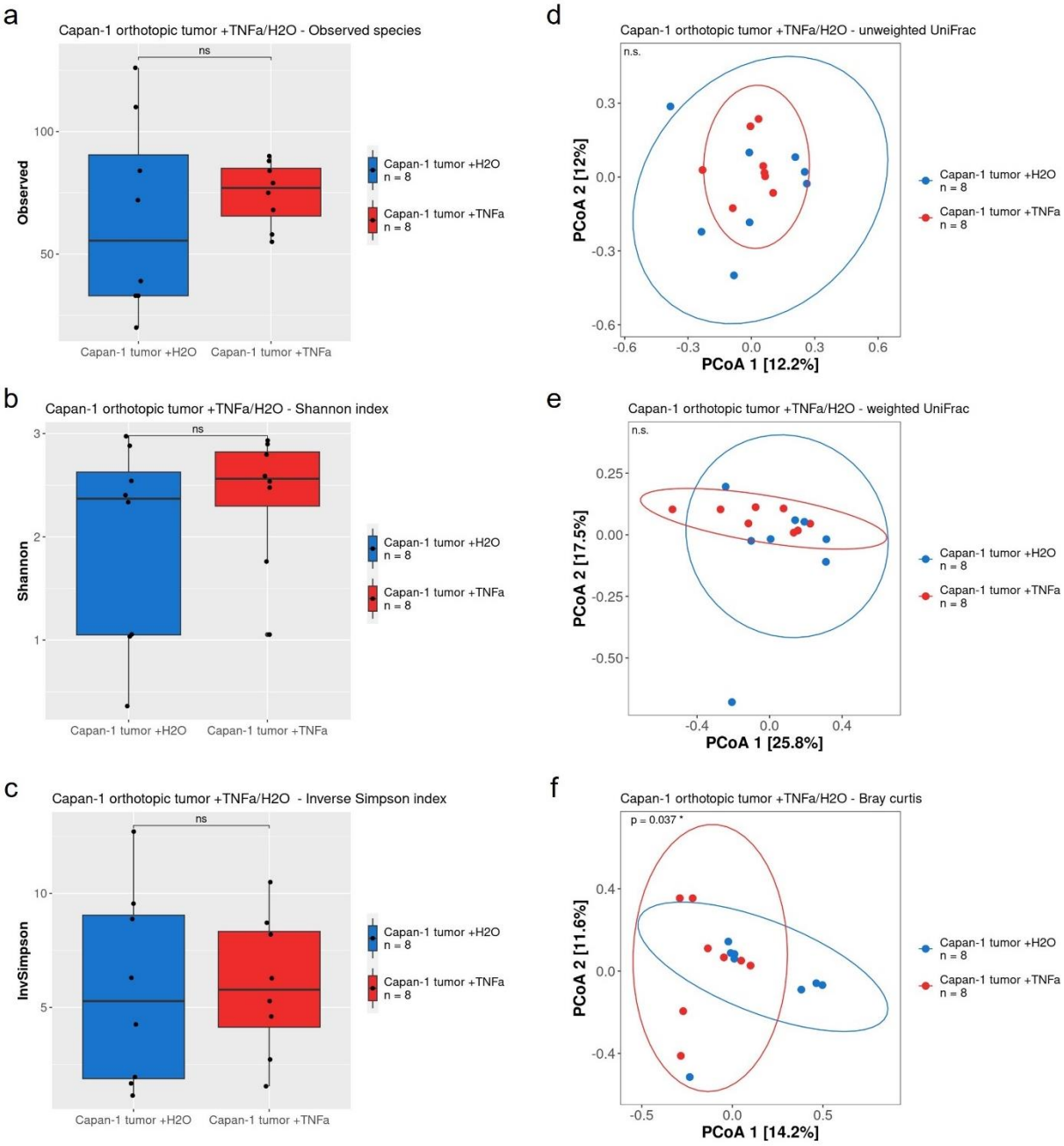

Figure 33: Analysis of the microbiomes of TNFα/H_2O-treated Capan-1-derived orthotopic tumor tissue samples. (a-c) Alpha diversity metrics OS, SI, and InvS are not significantly different between TNFα- and H_2O-treated Capan-1 tumors. Metrics calculated with non-rarified library sizes. Statistical testing performed with t-test or Wilcoxon rank test according to normal distribution test results. ns, not significant. (d-f) Beta diversity metrics are calculated with rarified library sizes. UwUF and wUF do not

show significant differences between TNFα- and H_2O-treated Capan-1 tumors. BC shows a significant difference between TNFα- and H_2O-treated Capan-1 tumors. Statistical testing performed with PERMANOVA. *, $p \leq 0.05$. n.s., not significant.

Rarefying the read numbers, calculating and plotting uwUF, wUF and BC beta diversity metrics with PCoA revealed homogeneity of both Capan-1 tumors treated with TNFα and H_2O in both UniFrac measures (**figure 33d,e**). Interestingly, in BC, indeed a statistically significant difference prevailed between the two differently treated tumor sample populations (**figure 33f**). Generally, the samples are composed of very different combinations of bacteria with contaminating species again dominating in some (**figure 34**). *Streptococcus thermophilus*, *Escherichia coli*, *Curtobacterium flaccumfaciens*, and an unspecified oral bacterium are among the most prominent species.

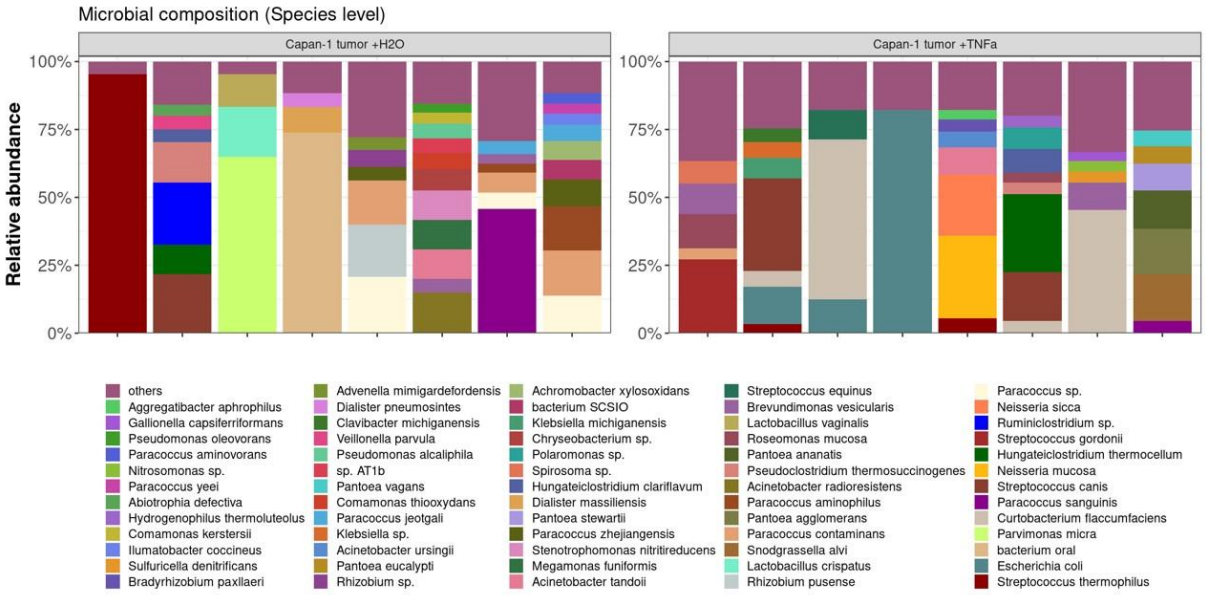

Figure 34: Microbial compositions of TNFα/H_2O-treated Capan-1-derived orthotopic tumor tissue samples. Relative abundances of species in percent are plotted.

4.4 Chapter 4 | The pancreatic microbiome of human PDAC patients with different tumor subtypes

4.4.1 The human PDAC microbiome is distinguishable from its matched healthy pancreatic counterpart

Ensuing orthotopic transplantation experiments towards more translational, patient-oriented microbiome research, human resected PDAC samples as well as healthy tissue samples from the surrounding pancreas of the same patients were examined. The samples taken at surgery by the Department of General, Visceral and Pediatric Surgery were provided as FFPE tissue blocks by the Department Pathology alongside with H&E stainings of the samples which had the tumor areas marked by pathologists. **Figure 35a** illustrates the workflow which started with cutting the blocks for the experiment presented here, and also contains exemplarily H&E stainings of the normal pancreatic and PDAC tissue samples. Further, FFPE human PDAC material was subject to 16S rRNA gene FISH and LPS IHC qualitatively demonstrating the greater abundance of bacterial components in tumor tissue as compared to adjacent normal human pancreas (**figure 35b,c**).

To start with, diseased and healthy pancreatic microbiomes were compared by analyzing n = 30 human PDAC tumor samples and n = 24 matched healthy pancreas tissues via 16S rRNA gene sequencing. Unlike the tissue samples derived from orthotopically transplanted human cells which were scratched from the same slide after histological examination and demarcation of tumor from adjacent normal tissue, the patient-derived healthy pancreas samples were sampled and embedded separately. Moreover, for every sample, a paraffin NTC was taken from the same section and further buffer-only controls were introduced into the data set, again causing the NTCs to outnumber the true samples. As already highlighted in **4.3**, an extra filter pipeline was applied to FFPE sample data in addition to the standard decontamination process. Before rarifying the data, alpha diversity was estimated in the three metrics OS, SI, and InvS none of which unveiled significant differences in the diversity within samples between the two groups normal pancreas and tumor (**figure 36a-c**). The median number of observed species was approximately 12 in both groups.

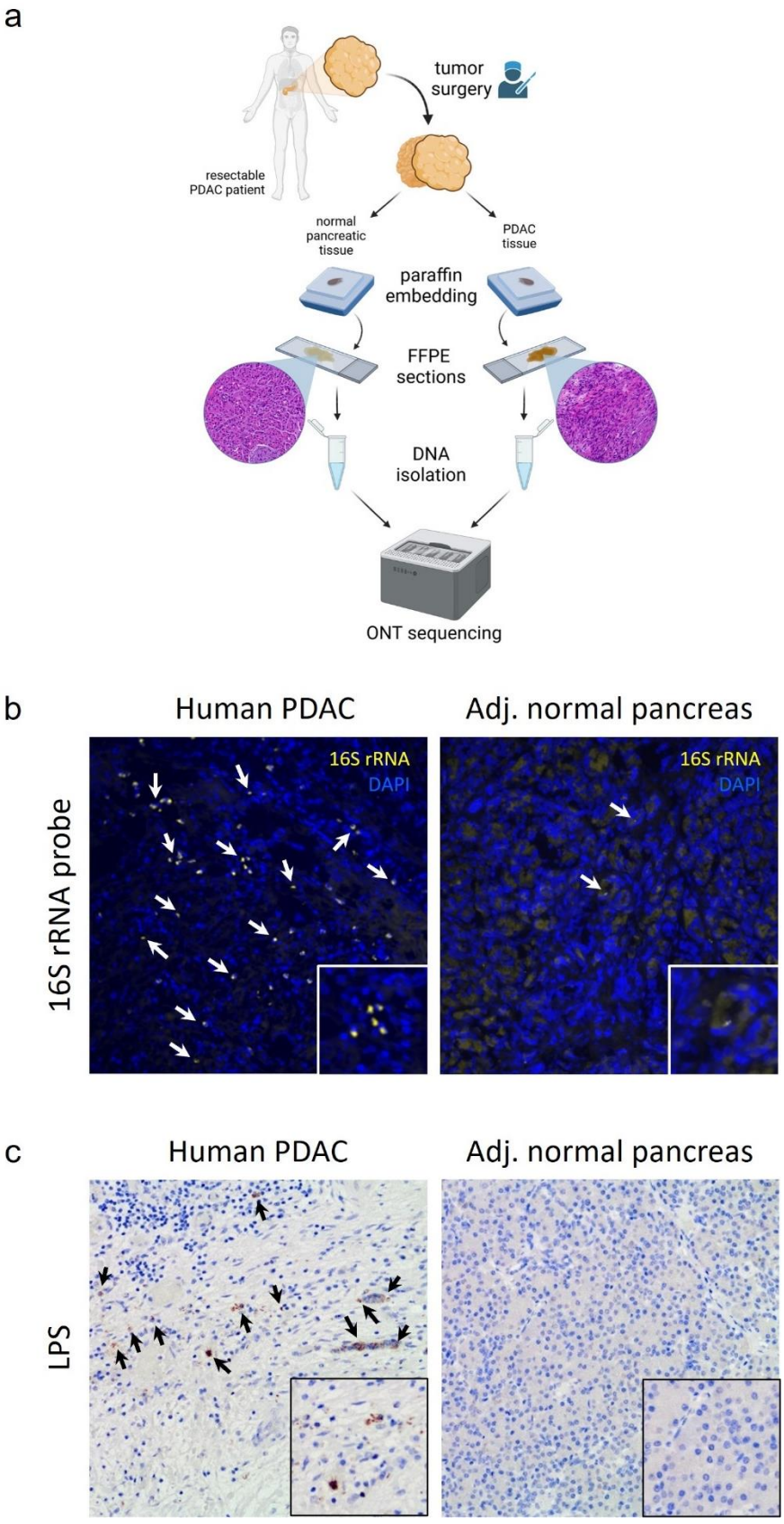

Figure 35: Investigation and visualization of bacterial components in matched tumor and normal pancreatic tissue samples of human PDAC patients. (a) Experimental design of pancreatic microbiome analysis employing tumor and normal pancreas tissue samples from the same human PDAC patients who underwent surgery. Exemplary H&E stainings of the two tissue groups. DNA was extracted from FFPE tissue sections and

subject to 16S rRNA gene sequencing. (b) FISH of the 16S rRNA probe EUB338 to the DNA of human PDAC and adjacent normal pancreas FFPE tissue. Nucleic counterstaining with DAPI. White arrows indicate the yellow, dot-shaped 16S rRNA gene light signals against yellow background autofluorescence. (c) IHC staining of LPS on human PDAC and adjacent normal pancreas FFPE tissue. Black arrows indicate the reddish-brown, dot-shaped LPS antibody signals. Nucleic counterstaining with hematoxylin.

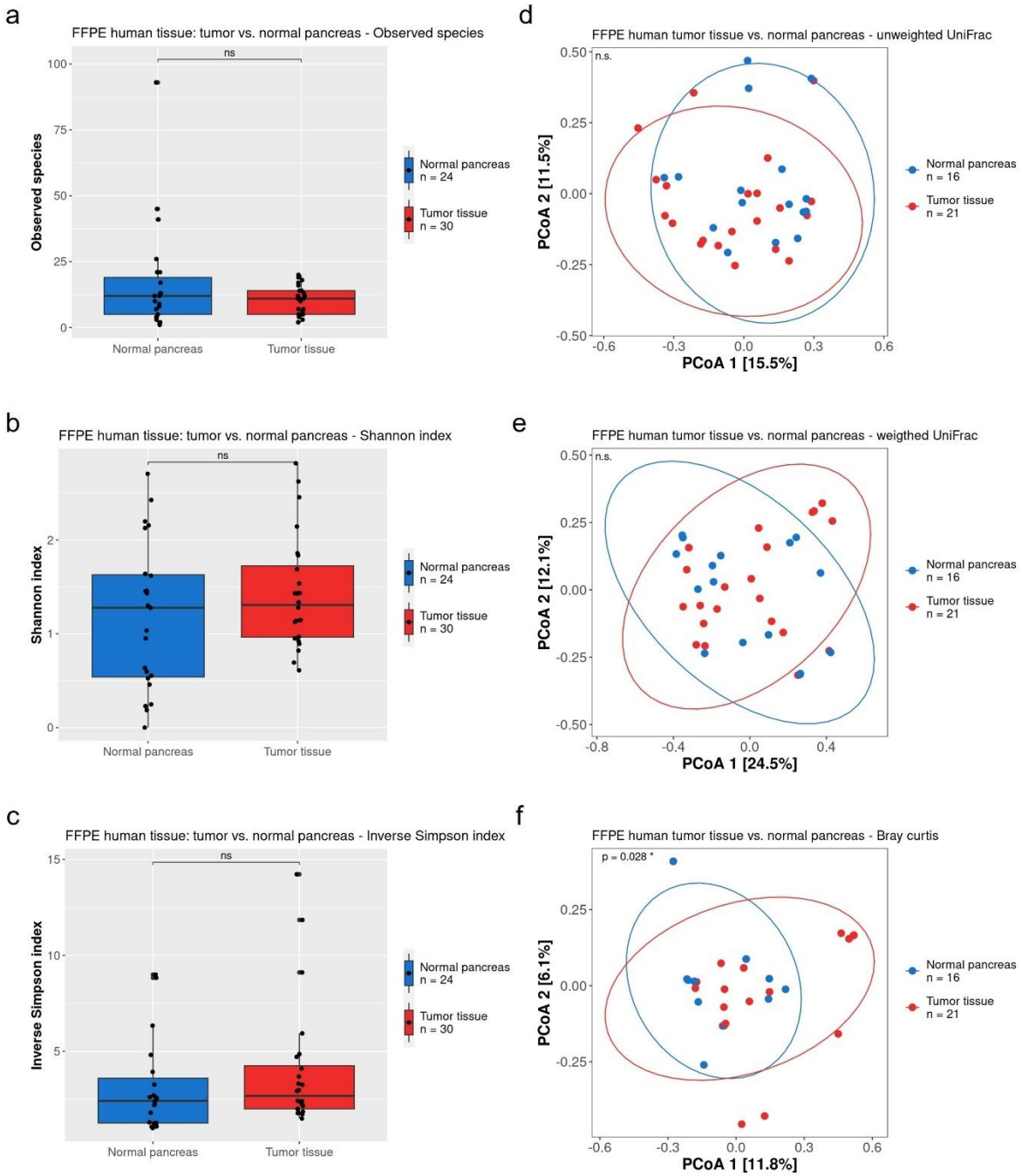

Figure 36: Analysis of the microbiomes in matched tumor and normal pancreatic tissue samples of human PDAC patients. (a-c) Alpha diversity metrics OS, SI, and InvS are not significantly different between human PDAC and normal pancreas samples. Metrics calculated with non-rarified library sizes. Statistical testing performed with t-test or

Wilcoxon rank test according to normal distribution test results. ns, not significant. (d-f) Beta diversity metrics are calculated with rarified library sizes. Eight normal pancreas samples and nine tumor tissue samples dropped by rarefaction. UwUF and wUF do not show significant differences between human PDAC and normal pancreas samples. BC shows a significant difference between human PDAC and normal pancreas samples. Statistical testing performed with PERMANOVA. *, $p \leq 0.05$. n.s., not significant.

The notably reduced read numbers throughout the data set after strict filtering demanded a rarefaction threshold considerably lower than usual and still about a third of the samples were eliminated. Normal pancreas remained with n = 16 samples and n = 21 tumor samples were left. Concerning beta diversity, no differences between healthy and diseased pancreas were identified calculating uwUF and wUF (**figure 36d,e**). However, BC indeed demonstrated statistical significance of the difference between human PDAC and corresponding samples from still healthy areas of the pancreas (**figure 36f**). The separation of the data points may not be as clear compared to the murine KPC data in **4.2**, however, this is a commonly observed data property in microbiome studies by now, especially with human samples which do not derive from such standardized conditions and need to be checked for confounders.

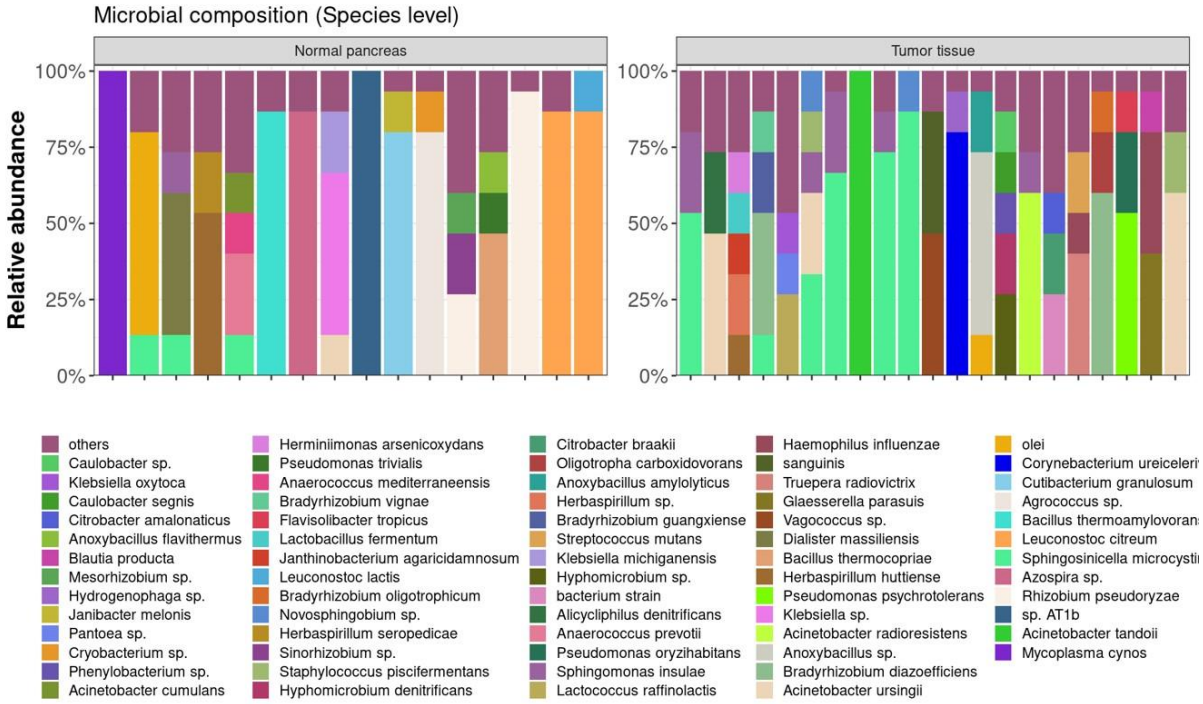

Figure 37: Microbial compositions of matched tumor and normal pancreatic tissue samples of human PDAC patients. Relative abundances of species in percent are plotted.

Yet, looking at the microbial composition, it displays a subtle divergence between normal pancreas and tumor tissue samples, although contaminants happen to appear here as well (**figure 37**). For instance, one normal pancreas sample supposedly consisted exclusively of *Mycoplasma cynos*, the only species left after the filter process.

4.4.2 Stratifying by classical, basal-like and hybrid PDAC subtypes reveals significantly different microbiome in hybrid tumors

In view of personalized medicine with more effective treatment impact in the battle against cancer, stratifying patients has been a successful concept so far. Also regarding PDAC, this seems to be a quite promising approach, at least in the laboratory-based research context to better understand tumor biology. As elucidated in detail in the introduction, CLA and BL forms of PDAC are discussed for several years now. More recently, the HYB subtype was additionally proposed, a mixed form of the two established ones. Its inclusion is consequential considering the extremely heterogeneous histology of PDAC masses which also applied to the samples processed here, wherefore the HYB subtype was covered as well within the scope of this thesis.

To begin with, the resected tumor samples had to be subtyped. This was done by IHC staining for the CLA tumor marker protein GATA6 and the BL marker CK5 as previously published (O'Kane et al., 2020). Staining serial tissue sections allowed the direct comparison of the markers' presence in the respective area. Of the 30 PDAC tumor samples sequenced, n = 6 samples derived from pre-experiments in which two different DNA extraction methods for the pathology samples had been tested (data not presented within the scope of this thesis) but the other 24 paraffin blocks were immediately available. These were cut and the sections stained, however, technical issues arose scanning 4 of the slides which could not be solved in due time leaving n = 20 tumor samples for subtyping. Employing QuPath, the area to be measured was pre-defined whereby staining anomalies were excluded (yellow lines in **figure 38**). Within this marked area, positively stained cells were then automatically detected by QuPath (displayed in red in the QuPath overlays in **figure 38**). The percentage of positive cells was determined and thereof z-scores were calculated. The latter were divided into three quantiles (lowerQ, mid, upperQ) and by comparing those for GATA6 and CK5, the three

subtypes (CLA, HYB, BL) were defined. The groups sizes were about equal with n = 6 samples accounting for BL tumors, n = 7 CLA and n = 7 HYB specimen. **Figure 38a-c** shows examples of GATA6 and CK5 stainings of samples classified as CLA, BL and HYB, respectively. While in CLA and BL tumor samples, the section was uniformly characterized by either GATA6 or CK5 staining for the greatest part (**figure 38a,b**), in hybrid tumors, both GATA6 and CK5 were pronounced depending on the tissue region (**figure 38c**).

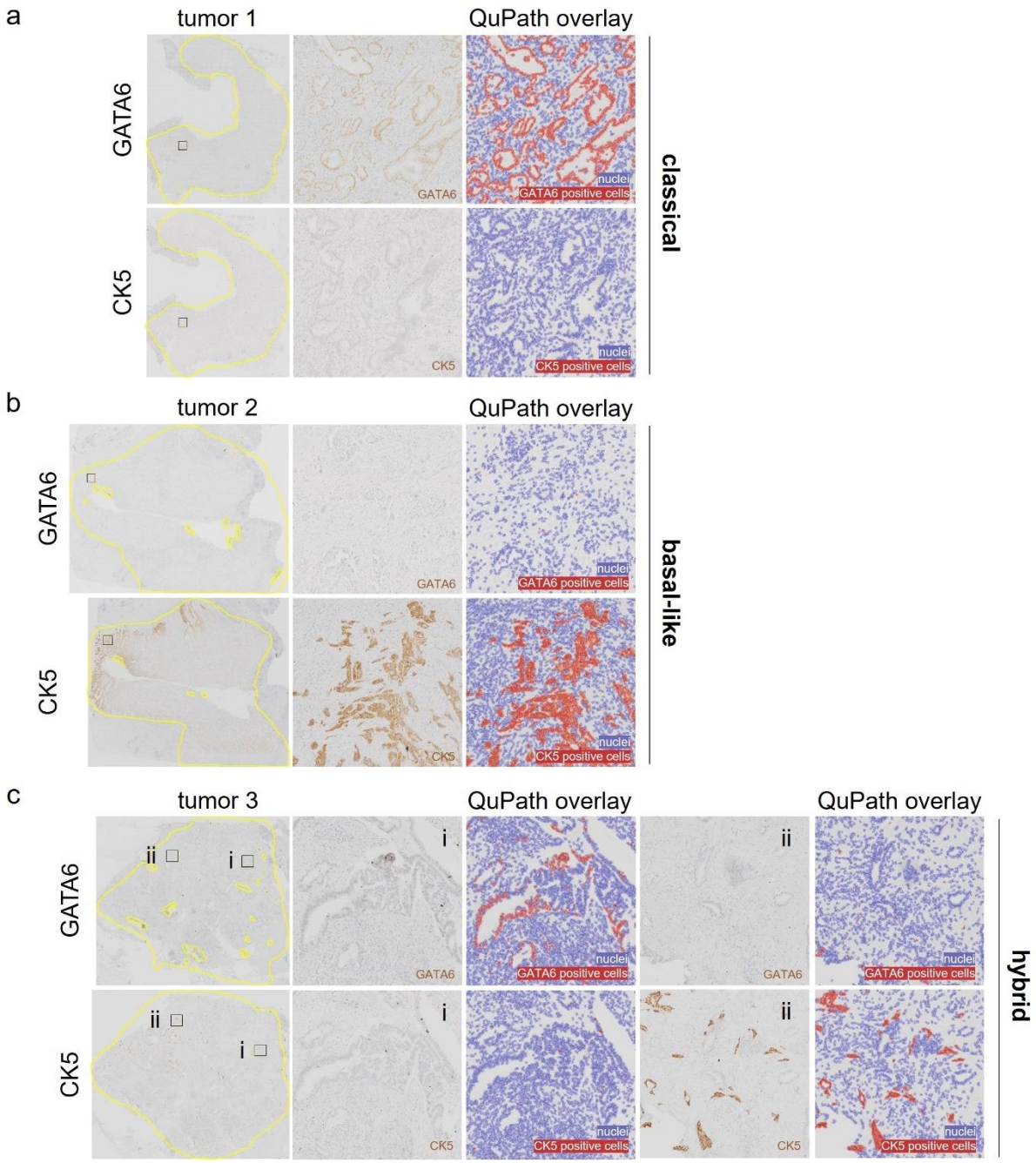

Figure 38: Human PDAC subtyping (CLA, BL, HYB) via IHC targeting marker proteins GATA6 (CLA) and CK5 (BL). (a) Two consecutive FFPE sections of a human PDAC sample are stained for GATA6 and CK5, respectively, via IHC. Employing QuPath, GATA6/CK5-positive cells (red) are separated from hematoxylin only-stained cells (blue).

Quantifying the GATA6- and CK5-positive stained areas and evaluating the difference between these fractions via z-scores and quantiles leads to the classification of this sample as classical PDAC subtype. (b) Following the procedure described in (a), the tumor sample of this human PDAC patient is subtyped as basal-like. (c) Following the procedure described in (a) unravels simultaneous appearance of GATA6 and CK5 to about equal extents in the here presented human PDAC sample. Two image details (i) and (ii) illustrate the predominant occurrence of GATA6 and CK5 in different areas of the tumor sample. This is confirmed by quantification and results in the assignment to the hybrid PDAC subtype.

This subtyping information was added to the metadata of each of the 20 samples and joined with the sequencing data in the adjusted phyloseq objects. Stratified by CLA, BL and HYB, the data was analyzed again. Alpha diversity did not return significant differences between the three groups in any of the considered metrics OS, SI and InvS (**figure 39a-c**). Except the number of observed species was significantly different between BL and HYB tumor samples with about 12 and 5 median species, respectively. Although rarefying the data with a low threshold, one CLA and one HYB sample were excluded leaving n = 6 samples in each group. Calculating and plotting beta diversity metrics uwUF, wUF and BC using PCoA next recapitulated and emphasized the finding of the hybrid tumor form being different from the two others which in turn were found to be rather similar in their microbial composition (**figure 39d-f**). This was also the case when trying to evaluate the data only stratified by CLA and BL based on the z-score alone where no significant difference between the two subtypes was found (data not shown). Regarding uwUF, HYB and BL tumors were significantly different from each other, while HYB vs. CLA almost reached statistical significance ($p = 0.051$) (**figure 39d**). Concerning wUF, the groups did not display any differences (**figure 39e**). Eventually, BC clearly demonstrated a significant difference between HYB and BL as well as HYB and CLA (**figure 39f**). The origin of this separation of hybrid tumors illustrated in rather abstract distances becomes concrete in the microbial composition at species level in **figure 40**. While the samples of BL and CLA tumors greatly vary in their microbial composition, the HYB tumors comprise samples with mostly similar bacteria. Here, the dominant species are *Sphingosinicella microcystinivorans* and *Sphingomonas insulae*. However, this simultaneously points out the problem commonly shared by the FFPE data, namely the retention of contaminating species in the data set.

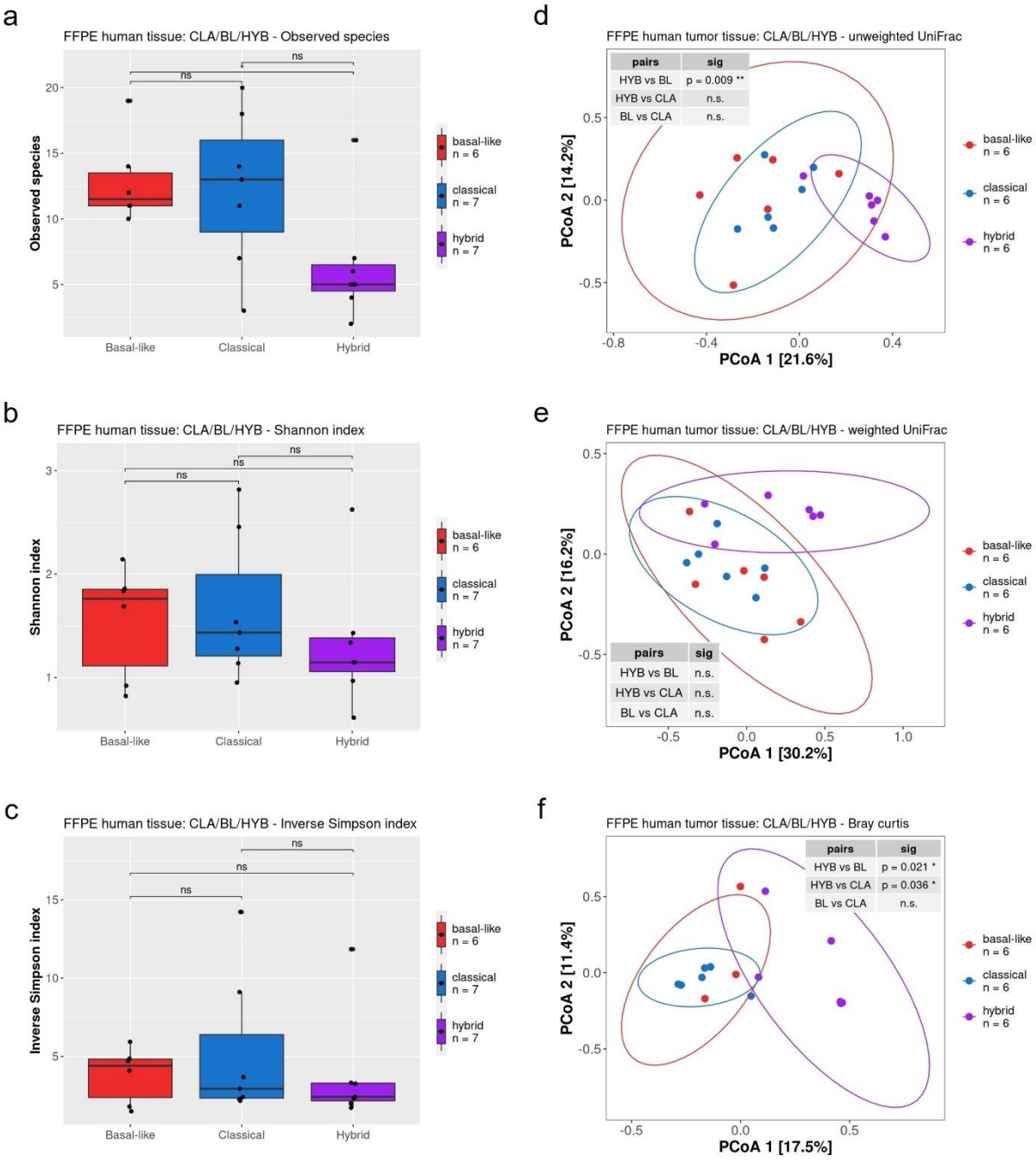

Figure 39: Analysis of the microbiomes in tumor samples of human PDAC patients stratified by PDAC subtypes BL, CLA, and HYB. (a-c) Alpha diversity metrics OS, SI, and InvS are not significantly different between the three regarded tumor subtypes, except for the number of observed species between basal-like and hybrid tumors. Metrics calculated with non-rarified library sizes. Statistical testing performed with t-test or Wilcoxon rank test according to normal distribution test results. *, $p \leq 0.05$; ns, not significant. (d-f) Beta diversity metrics are calculated with rarified library sizes. One classical tumor sample and one hybrid tumor sample dropped by rarefaction. UwUF shows a significant difference between hybrid and basal-like tumor samples. BC shows significant differences between hybrid and basal-like tumor samples as well as between hybrid and classical tumor samples. The other group comparisons and wUF remain insignificant. Statistical testing performed with pairwise PERMANOVA. sig, significance; *, $p \leq 0.05$; **, $p \leq 0.01$; n.s., not significant.

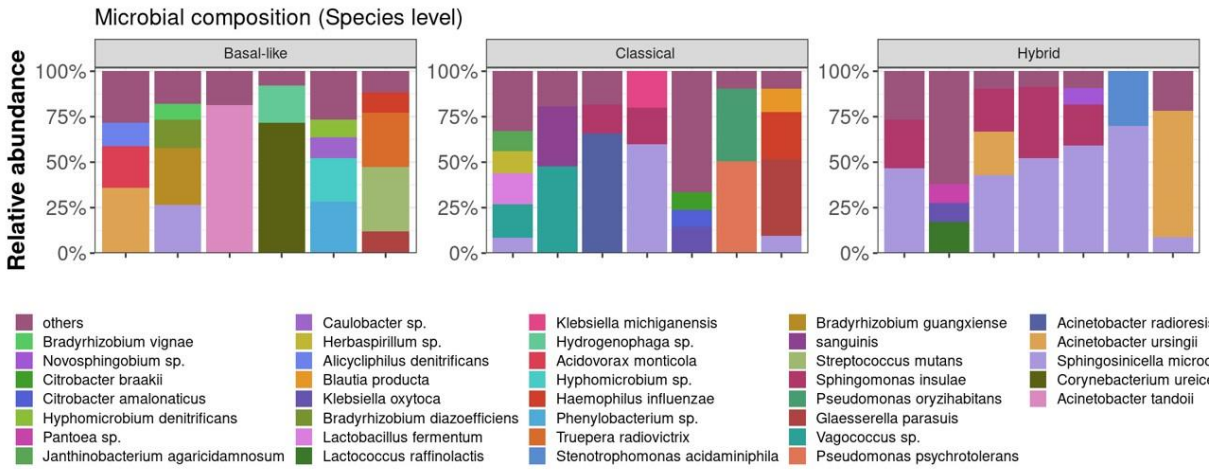

Figure 40: Microbial compositions of tumor samples from resected human PDAC patients stratified by PDAC subtypes BL, CLA, and HYB. Relative abundances of species in percent are plotted.

5 Discussion

5.1 Chapter 1 | Chances and challenges of third generation sequencing in the context of high and low microbial biomass samples

5.1.1 Oxford Nanopore sequencing outcompetes NGS technology for microbiome research

Scientific progress most often goes hand in hand with technical advances. This is also mirrored in the clinical setting where, for instance, modern imaging procedures allow to detect and distinguish what formerly had escaped from being discovered at all. For laboratories of different specializations, it usually poses a challenge to implement new techniques outside their commonly used toolboxes in terms of money, time, and expertise. Establishing a sequencing pipeline customized to the laboratory's needs also turned out to be a time-consuming, labor-intensive project itself within this thesis requiring many benchmarking experiments and the acquisition of new competences. Yet, methodological advancements are often inevitable in order to realize certain project goals.

After observing its involvement in a plethora of diseases, the microbiome also attracted great attention in the oncology field. Therefore, it was only a matter of time until pancreatology picked up on this trend and implemented microbiology as a joint discipline. Studying the microbiome in cancer is highly compelling for various reasons. It was shown to be able of inducing neoplastic transformation, certain pathogens may cause and promote cancer, dysbiosis is capable of influencing cancer progression, and some microbes have proven to serve as target for cancer prevention (Schwabe and Jobin, 2013). As introduced in **1.4.2**, significant findings about tumoral microbiomes were made that pushed PDAC research to a new vantage point which clearly offers innovative perspectives for the persistent problems with this tumor entity.

To date, the vast majority of studies addressing the microbiome in cancer has employed amplicon sequencing using an Illumina platform, the most popular representative of NGS (del Castillo et al., 2019; Geller et al., 2017; Nejman et al., 2020; Pushalkar et al., 2018; Riquelme et al., 2019). Only very few studies can be found that chose a different approach

like whole-genome sequencing as performed by Guo and colleagues, however, it was still NGS-based (Guo et al., 2021). Although NGS is well-established and most papers rely on it for its low error rate of below <1% and comparably low costs, it also comes along with disadvantages. Short fragment sequencing of usually 150 to 250 bases is not suitable for a 1,500 bp long region of interest if the integrity of this gene decides on the specificity of the detected microbe. In most cases, the experiments of the aforementioned studies were performed only sequencing one or two of the nine variable regions (V) of the bacterial marker gene 16S rRNA (**figure 41**), most often single V4 or a combination of V3-V4 which are known to give genus level resolution at best. This constitutes a major drawback since depending on the selected variable region, the sequencing results vary significantly (Bukin et al., 2019). Therefore, it is difficult to compare studies with different sequencing approaches. The only study that covered multiple Vs was conducted by Nejman et al. and included 5 amplicons covering 68% of the 16S rRNA gene (Nejman et al., 2020). NGS has been the best methodology available for a long time, however, with the advent of third generation sequencing better options are now on the market.

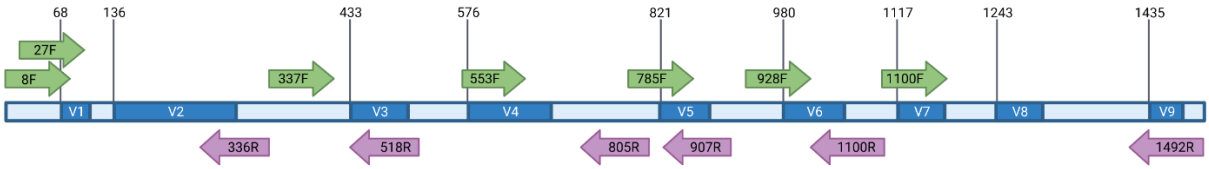

Figure 41: Schematic illustration of 16S rRNA gene with variable regions and established primers for NGS. The 16S rRNA gene, simplified depicted in blue, contains nine variable regions (V1-V9) of different length with varying spacing. Black numbers at the top indicate the number of the DNA base at which the regarded V starts. The 16S rRNA gene has a total length of ~1500 bp. Green arrows mark forward primers (F), purple arrows mark reverse primers (R). All primers contain the DNA base number at which they start. Most studies employing NGS amplify V4 or V3-V4. The ONT SQK-16S024 16S rRNA gene sequencing kit used in this thesis includes the barcode-coupled 16S primers 27F and 1492R. Own illustration adapted from (help.ezbiocloud.net/16s-rrna-and-16s-rrna-gene, accessed on 2023-07-24 22:58:00).

Besides offering a portable sequencing platform, rendering the own work independent of sequencing units as well as the option of real-time sequencing, the biggest advantage of ONT sequencing is the possibility to analyze long fragments. Several 100 kb in one stretch are not uncommon and the longest read officially reported in a publication was about 2.3 Mb long (Payne et al., 2019). The ability to sequence long fragments allows the resolution of bacterial

communities on species level. For these reasons, it was decided to establish third generation sequencing by ONT in our laboratory. This enabled independence from in-house NGS-based sequencing facilities in view of future real-time bed-side sequencing.

It goes without saying that every method comes along with pitfalls and weaknesses, especially on the level of such high complexity. Nanopore sequencing is not yet broadly established, although it would render microbiome studies a lot more comparable. This may partially be attributed to the complex bioinformatics that need to be mastered. Also, the comparably high error rate of 10% surely still deters users from relying on this technology. This is due to the physical properties of the nanopores which, despite their small size, are passed by more than one nucleic base at the same time, wherefore inaccurate voltage changes along the membrane may occur and it is difficult to precisely match them to a single DNA base (**figure 42**).

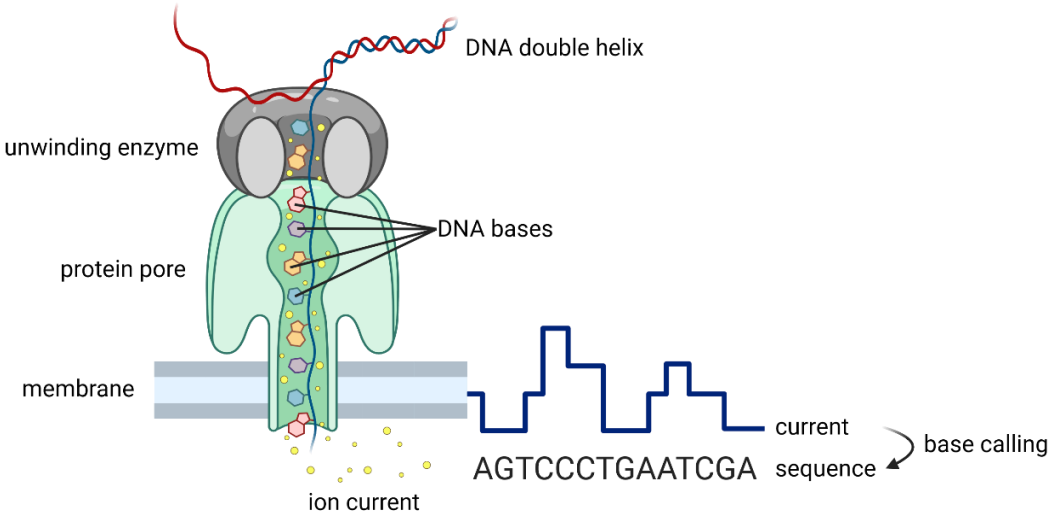

Figure 42: Schematic illustration of nanopore sequencing mode of operation. Lipid membrane-integrated protein channels (nanopores) bond with unwinding enzymes (helicases) which feed the DNA strands through the pores. While one nucleotide at a time is pushed into the nanopore, the latter technically holds several bases at the same time. Simultaneously, the electrostatic potential applied at the membrane causes an ion current to pass through the nanopore. Changes of this current caused by different DNA bases are translated to the DNA sequence in a computational process called base calling. Own illustration created with (biorender.com).

At the time of running sequencing experiments for this thesis, the QS of base caller Guppy was 9, however, ONT recently developed new sequencing chemistry reaching a QS of 20 and higher representing a 1 in 100 chance of an incorrectly annotated base, thereby promising

single molecule accuracy of over 99% (nanoporetech.com). Taken together, this means ONT sequencing is not only capable of keeping up with NGS options but in fact outcompetes them regarding microbiome studies, even though this technology is overall still in its infancy compared to its predecessors. To the present day, the established pipeline in our laboratory stands the test, runs multiple clinical trials such as P-MAPS (NCT04777812) and is continuously refined. For instance, the enormously reduced error rate pledging ONT Q20+ chemistry is currently being established.

P-MAPS (NCT04777812) is only one example for the growing number of clinical studies registered at ClinicalTrials.gov that aim at investigating the multiple impact layers of the microbiome in the context of PDAC. However, most other trials still employ NGS techniques. The choice of ONT sequencing does not only serve as reliable platform with high specificity but also may directly be employed for deployment in clinical practice. This sets new perspectives to the elucidation of the microbiome-PDAC-axis and the translation of the obtained insights.

5.1.2 Optimized workflow for sequencing bacteria in murine fecal and pancreatic tumor samples

Besides sequencing, the workflow requires other crucial steps that had to be established according to the specific needs in examining murine samples of different origins. The most efficient DNA preparation method is a key element in investigating microbial communities. As mentioned before, the way of sampling, extracting and sequencing DNA has a strong impact on the outcome of the analysis (Bjerre et al., 2019; Teng et al., 2018). Hence, testing different DNA extraction kits was necessary to set the gold standard for the here presented samples to be sequenced by nanopores.

The International Human Microbiome Standards project aims at a precise characterization of the human microbiome via the coordinated development of standard operating procedures (SOP) to foster the comparability between studies in human microbiome research (human-microbiome.org). Incorporating the instructions and suggestions from these SOPs into the original manual of the different manufacturer's kits resulted in two protocols for each kit except for the QIAamp DNA Microbiome Kit which is conceptualized in a way not offering

feasible implementation of the IHMS recommendations, and it is anyway designed specifically for the extraction of microbial DNA. Therefore, it was generally expected to return lower DNA concentrations than the other kits which pick up host DNA as well. Surprisingly, the Invitrogen IHMS protocol still performed the best regarding the chosen readout parameters DNA concentration, DNA purity, OS, and read counts, while the Qiagen microbiome kit achieved the worst values for most parameters besides DNA concentration. Regarding the A260/A280 values for pure DNA, all protocols achieved median ratios close to 1.8 arguing against the presence of contaminants absorbing around 280 nm such as proteins and phenols. Contrary, most kits failed to reach acceptable A260/A230 ratios and resided at appreciably lower values than 2.0. This indicates contamination with substances that absorb close to 230 nm such as carbohydrates or phenol. Carbohydrate carryover often happens with plants, phenol is used in conventional DNA extraction with TRIzol reagent, for instance. In column-based kits such as the ones used here, guanidinehydrochlorid (guanidine-HCl) is often employed and denatures proteins. Guanidine is known to disturb ONT library preparation and sequencing as stated on the company's website. The conclusion may be drawn that DNA samples isolated with the Invitrogen IHMS protocol achieve a higher throughput in sequencing due to more intense washing. The sequencing depth (read count) is remarkably higher using this protocol, likely due to lower amounts of substances interfering with downstream sequencing removed by these additional washing steps.

Following the ONT protocols for library preparation which were directly taken from the company, the data processing was the most advanced step in establishing the pipeline. It has been published before that Centrifuge as a classifier needed to assign taxa to reads is a high sensitivity tool (Leidenfrost et al., 2020; Sanderson et al., 2018). However, previously, it was only used with preformatted indexes pvh and p which failed to annotate all 17 species in the more complex mock community used here. Assuming this was due to incomplete databases, not due to incomplete sequencing, new libraries were built according to official Centrifuge instructions. First, an index containing a large set of complete NCBI RefSeq genomes was applied to the data. Still, three species were not found in the data including *Candida albicans*, the most common human fungal pathogen (Sagué et al., 1993). The NCBI complete genomes index (abv library) only comprises 18 fungi. Consequently, it is not advisable to employ this index in case of analyzing diverse microbiomes containing many different fungi. Because of this result, the nt library (complete and incomplete genomes) was applied for classification in

the pre-experiments of the method establishment (**4.1**) as well as the KPC experiments (**4.2**). Yet, the disadvantage of using such comprehensive libraries is that it takes a high computational memory capacity, and the analyses are a lot more time-consuming. Within the scope of this thesis, the focus was set on bacteria which were filtered from the sequencing data at an early step in the analysis R script. Therefore, it was decided to classify the FFPE sample data using the abv library. Technically, it is possible to detect specifically fungi by performing 18S rRNA gene sequencing. NGS-based studies performed amplicon sequencing of the internal transcribed spacer between the eukaryotic 18S and 5.8S rRNA genes for this purpose (Aykut et al., 2019). Beyond that, ONT metagenomic sequencing also provides the opportunity to perform *de novo* sequencing. Regarding the pre-defined mock community, however, classification with the nt database returned the best recall and AUPR values. Solely, the precision remained low as for the other indexes. This can be attributed to the mock community containing some very low abundant species of < 0.1%. Thus, only taxa with abundances below this prevalence filter threshold could be removed by calculating overall precision and recall which in turn led to the annotation of many incorrect species.

Feces are high microbial biomass samples that are comparably straightforward to investigate. The workflow, established with rectal swabs among others, could directly be transferred to these murine specimens. Metagenomic sequencing was chosen to analyze the microbiome of these DNA samples due to extremely high concentrations of several hundred ng/µL and almost unexceptionally excellent purity ratios after Invitrogen IHMS isolation and the application of a purification kit. Concerning tumor samples, the methodology was not that trivial. In their identity as low microbial biomass samples, it was set to apply 16S rRNA gene sequencing since this technique works with amplifying PCR and can therefore deal with low template material and highly host contaminated samples (in this case murine DNA). Still, it was necessary to test the two best performing DNA isolation kits as well as the PCR parameters DNA input and number of cycles. The strongly varying read numbers among samples that were prepared the same way underline the general feasibility of the pipeline, yet also the great heterogeneity even among tumors of the same genetic origin. This strongly argues for other factors influencing the abundance of a tumoral microbiome when comparing animals of the same genetic background. It is important to mention that in experiments performed on snap-frozen bulk tumor tissue, a random piece was cut off without paying attention to the exact tumor area. It is likely that this approach contributes to

the varying sequencing depth, speculating about an uneven distribution of the microbiome across the pancreas and subsequently the tumor. While different PCR parameters were tested, the PCR reaction mix was left unchanged. Higher amounts of input DNA and an increased number of PCR cycles do not necessarily result in an elevated amount of PCR product since too much DNA template may cause steric hindrance and at some point, the reagents needed for amplification are used up. In this benchmarking experiment, however, it worked, presumably because the microbial to murine DNA ratio is very low in tumors so more template and amplification time in fact results in more PCR product potentially to be sequenced. For these samples, the Invitrogen IHMS protocol, 50 ng input DNA and 35 PCR cycles resulted in considerably higher read counts compared to 10 ng and 25 cycles suggested by the original ONT sequencing protocol.

5.2 Chapter 2 | Dynamics of intestinal and tumoral microbiomes in KPC mice and human PDAC patients

5.2.1 Locally distinct microbiomes in KPC mice are unique in health and disease

Regarding human PDAC, the involvement of the microbiome in various aspects of the disease from inception to clinical prognosis due to its impact on therapeutics and the immune system is acknowledged by now. Convincing association studies have connected changes in the microbial composition with distinct tumor stages and various publications discuss how to use microbiota in therapy and diagnostics although this is not yet put into practice (Pfisterer et al., 2022). Moreover, attempts have been undertaken to characterize the tumoral microbiome (del Castillo et al., 2019; Geller et al., 2017; Guo et al., 2021; Nejman et al., 2020; Pushalkar et al., 2018; Riquelme et al., 2019). This will also be further discussed in **5.4** regarding humans. On the contrary, murine models of PDAC, such as the widely deployed transgenic KPC model, are far less studied regarding their disease-related microbiome. Yet, it is crucial for basic and translational science to reproducibly define the exact intestinal and intratumoral microbial composition in order to implement the microbiome as a reliable diagnostic and therapeutic tool in the future.

The comprehensive characterization of fecal and tumoral microbiomes of mice with a KPC background is fundamental for further functional studies focusing on the mechanisms behind the microbiome-disease-relation. As main readouts, alpha and beta diversity were estimated. Although alpha diversity is an important measure to evaluate single samples or one group, beta diversity is more relevant in the context of group comparison. Alpha diversity metrics not showing any significant differences, which was mostly the case in the experiments presented in this thesis, does not necessarily mean that there is no difference between the regarded groups overall but states that the groups themselves have an about equal level of diversity. For instance, the same number of observed species does not necessarily imply the same kind of species. A large degree of variance in alpha diversity within one group may indicate that confounding variables were not considered causing inhomogeneity within the group, and it will most likely be difficult to identify a common microbial factor to be associated with the group's main uniting property. Nevertheless, beta diversity measures may show differences between groups depending on the composition of

their individual samples. Typical measures consider phylogeny and species abundance as elucidated in the introduction (**1.4.1**). This knowledge about quantitative and qualitative diversity metrics enables the interpretation of microbial ecology in the context of PDAC. In addition, the familiarization with the normalization method rarefaction for beta diversity estimation is important. Although it is critically reviewed, rarefaction is currently the best normalization technique available for microbial sequencing data (Weiss et al., 2017). In order to handle the at times greatly varying library sizes, samples are rarefied which means random subsampling of reads from the original pool (Lin and Peddada, 2020). The larger the number of species subsampled is, the more species will be found in the sample. This draws rarefaction curves that start out growing fast, then approach a plateau. In the results chapter 2 (**4.2**) discussed here, the optimized ONT sequencing-based analysis workflow established in **4.1** proved its quality and suitability for murine microbiome research via the application of rarefaction and evaluating measures of diversity. It was not only a technical success, but also significant differences were found between both fecal and tumor samples of KPC mice and their respective healthy controls.

The gut is one of the best-studied microbial niches in humans as well as in mouse models. Especially compared to other internal organs and tumors that are low microbial biomass sites, it harbors an exceptional load of microbiota, additionally facilitating the technical aspects that form part of microbiome studies. Typical gut microbes as well as their functional implications in metabolizing nutrients and drugs are described. In the case of PDAC, positive correlations between gut and pancreas microbiome were drawn. For instance, Ren et al. first declared the stool microbiota to be unique in PDAC, and others additionally stated that it can serve as a non-invasive biomarker (Kartal et al., 2022; Nagata et al., 2022; Ren et al., 2017). The finding that the fecal microbiome of tumor-bearing KPC mice mimicking human PDAC differs significantly from the one of healthy controls was in line with these human studies. Feces of two more time points during pancreatic tumorigenesis were investigated and even at an early age of only 6 weeks, accurately distinguishable microbial compositions were revealed comparing KPC and CTRL mice. Notably, invasive tumors have not formed yet at this age, even tumor lesions have usually not developed in these tissues (Westphalen and Olive, 2012). At 12 weeks, when tumors start to manifest, this difference persists, and regarding uwUF which considers presence/absence of species as elucidated before, it becomes clearer and statistically more significant over time. This finding hints

towards the fecal microbiome being an early sensor and indicator of changes in the seemingly healthy pancreatic tissue and underlines its biomarker potential. Since other confounders were excluded as potential reasons for the changes in the gut microbiome, it can be concluded with great certainty that the genotype of the mice, i.e., *Kras* and *Trp53* mutations responsible for pancreatic carcinogenesis, is the determinant of these microbial differences. Confounder control is particularly important as it was recently shown that, comparably to humans, factors such as sex have major impact on the gut microbiome of mice that carry established tumors (Kaune et al., 2023).

In contrast to these results, a study by Pushalkar and colleagues found the feces of their PDAC mouse model to have similar bacterial profiles at a young age compared with the controls (Pushalkar et al., 2018). They studied the *LSL-Kras*$^{G12D/+}$*;P48-Cre* (KC) transgenic mouse model that does not harbor a *Trp53* mutation and is therefore characterized by, unlike the KPC model, a slowly progressive tumor development. Interrelating these observations, the gut microbiome really seems to be influenced by the time-dependent tumor onset and simultaneous disease severity. This fits in with the hypothesis of this thesis that the tumor (subtype) impinges on the microbiome. However, functional questions remain unanswered, such the exact mechanisms how PC causes changes in the fecal microbiome. ONT metagenomic sequencing enables functional analyses and may be exploited in this regard. Bacterial translocation via the digestive route, the duodenum and pancreatic ducts has been reported before, however, bacteria found here in feces and pancreas were not necessarily the same (del Castillo et al., 2019; Glaubitz et al., 2023; Pushalkar et al., 2018). Only *Lactobacillales* were found in both healthy compartments, a fact that allows the consideration of the bacteria migrating from the gut to the pancreas perceiving the latter as an initially sterile organ.

The pancreatic microbiome is of special interest because of possible direct implications on pathogenesis and impact on treatment such as sequestering of chemotherapy (Geller et al., 2017). Also, the exocrine pancreatic function affected by PDAC has been reported to influence the composition and maintenance of gut microbiota (Ahuja et al., 2017; Frost et al., 2019). This may at least partially explain the shift in the fecal microbiome and argue for its potential implication as biomarker. The presence of bacterial components in PDAC tissue had been shown before but mainly in humans (Geller et al., 2017; Nejman et al., 2020; Pushalkar et al., 2018). In this thesis, this finding was reproduced and confirmed for KPC

mice. As opposed to healthy pancreatic tissue, cancerous pancreas harbors a more abundant microbiome and the intratumoral bacterial particles were mainly found intracellularly. Like the fecal microbiome, the microbiome of KPC tumors also proved to be significantly different from its healthy counterpart in beta diversity metrics. Yet, the difference was a little less clear between tumor and healthy pancreas than for the fecal sample communities. This may be traced back to the fecal microbiome being more robust and high microbial biomass samples generally being less prone to contamination issues. Tumors and most internal organs are low biomass samples, wherefore contamination poses a severe problem. *Escherichia coli*, for instance, was found in both KPC tumor and healthy pancreas samples. Since contaminants are usually systematically introduced and therefore present in all samples handled together, they may mask the species in fact causing differences between groups. The detection of contaminants mostly distorted the results in preceding studies with the exception of the Nejman study that contained a large set of NTCs and a striking filter pipeline (Nejman et al., 2020). Together with their unique methodological approach of sequencing multiple Vs with NGS elucidated in **5.1.1**, this marks the study as pioneering work. Contamination as one of the biggest pitfalls in microbiome research and ways to deal with it will be further discussed in **5.3.1**.

Regarding the tumor areas, no difference between tumor center and tumor periphery was found regarding beta diversity. This was rather unexpected since these sub-compartments, also referred to as reactive and deserted sub-TMEs in human PDAC in a study by Grünwald et al., differ greatly in their cellular and structural composition potentially offering habitats to various species of differing specializations due to an oxygen gradient and differing nutrients from the periphery to the center of the tumor (Grünwald et al., 2021). Looking at the microbial composition detected in both groups unveils the uniform presence of *Escherichia coli* in these tumor samples as well. While it is a common gut bacterium comprising not only pathogenic but also commensal strains, it appears to be the main contaminant in this setting (Martinson and Walk, 2020). Therefore, it is highly likely that the missing differences in beta diversity are due to methodological pitfalls rather than given biological conditions. Yet, surprisingly, the number of observed species was significantly different between center and periphery with a higher median OS in KPC tumor centers, although a median OS of 500 admittedly also argues for the presence of contaminating species. It is possible that most of these species are only low abundant, hence not causing

differences in abundance-dependent measures such as wUF. Galeano Niño and colleagues published a study where they showed that bacteria preferably colonize less vascularized tumor areas (Galeano Niño et al., 2022). This matches with the higher number of OS in KPC tumor centers that are also less vascularized in their fibrotic and necrotic properties.

5.2.2 Similarities between murine KPC and human PDAC microbiomes present the prospect of clinically relevant microbiome studies

Finding the locally distinct fecal and tumoral microbiomes to be unique in the context of modelled PDAC and systematically building up the fundamental knowledge about these microbiomes in one of the most important mouse models of PDAC with characterizing data, the comparison to human PDAC was more than encouraged. Both fecal and tumoral microbiomes were assessed for consistency with the complex human disease in two different approaches.

The fecal microbiome was analyzed by applying stool classifiers lately developed from human data and testing their predictive value on the KPC data (Kartal et al., 2022; Nagata et al., 2022). These two publications found stool samples of human PDAC patients to be unique and distinguishable from healthy subjects. Being able to predict tumor cases from stool samples would be a milestone in PDAC research. Making use of preselected genera and species from the Kartal and Nagata data, the accuracy of predicting PDAC from KPC fecal samples reached AUC values of up to 80% from plotting the applied Ridge regression which is generally regarded as good performance (D'Agostino et al., 2013). The model calculated here based on the Nagata et al. features performed even slightly better than the accuracy achieved by Nagata and colleagues themselves. In summary, it was possible to predict tumors from murine KPC feces employing metagenomic classifiers built from recent publications. This means human and murine fecal microbiomes are highly similar in their dysbiosis to a degree that allows drawing conclusions from one species to another. This adds another point to the list in which the KPC mouse model recapitulates human PDAC, an extraordinarily complex malignant disease. Also, it is extremely remarkable since the published human data was ascertained with NGS and still compares to the here presented ONT sequencing results of the murine microbiome.

Since there is less biological relevance to predict cancer cases directly from PDAC tissue as it is not routinely available, there was no regression model calculated at this point. However, it is still highly relevant to learn about the comparability between the tumoral microbiomes of KPC mice and human PDAC patients to better understand tumor biology and the pronounced TME. As already mentioned, Nejman et al. published an outstanding study with reliable human PDAC microbiome data which was used to compare the KPC tumoral microbiome data to (Nejman et al., 2020). Finding their most abundant tumoral genera and species in the KPC tumors, additionally also with a higher abundance in the tumor samples than in the healthy organ, highlights the validity of the comparison of human and murine pancreatic tumors. Comparing the tumoral microbiome of KPC mice to the ones of other human cancer entities such as CRC and OSCC demonstrated the clear distinction between them (Galeano Niño et al., 2022). This is of particular interest as it loosens species boundaries in microbiome research and strengthens the specificity of the pancreatic tumoral microbiome, although oral cavity, gut and pancreas all form part of the digestive tract and are therefore highly interconnected. Inversely, regarding the oral cavity, for instance, a pronounced heterogeneity and rather low specificity of the microbial pattern to PDAC is known, meaning the microbial signatures may also be found in the context of other diseases (Stasiewicz et al., 2021). Current literature has mainly investigated the pancreatic tumoral microbiome in humans, some in orthotopic transplantation mouse models, whereas the data presented here on KPC tumors are novel and of essential importance in contributing to an integral picture of this mouse model.

5.3 Chapter 3 | PDAC subtype-microbiome-relations and the nature of contamination in microbiome research

5.3.1 FFPE samples are extremely prone to harbor contaminants rendering data analysis challenging

The advent of NGS offered an entirely new dimension to microbiome research allowing high throughput studies and an extreme sensitivity even to exceptionally low abundant species. However, at the same time these advantages come with the risk of also detecting contaminants that do not form part of the original sample community. This mainly poses a problem in low biomass samples in which contaminants can quickly become the predominant microbes (Davis et al., 2018; Eisenhofer et al., 2019; Salter et al., 2014). Tumor and healthy organ samples are popular examples for low biomass specimen, and they require particular attention regarding this important topic. The FFPE data sets discussed in the following serve as prime example for this aspect of microbiome research. Fortunately, the awareness for contamination being a serious challenge rose with an increasing number of publications addressing this issue and proposing ways of dealing with it (Glassing et al., 2016; Rampelotto et al., 2019). Yet, interestingly, studies still publish environmental bacteria as part of tumoral microbiomes without explaining their decontamination approaches.

As part of the results in **4.3**, the sequencing data of Capan-2- and MiaPaCa2-derived tumors revealed the presence of environmental bacteria in their microbial composition as well. Regardless of the efforts to keep everything as clean as possible, the introduction of contamination during every step of microbiome analysis workflows cannot be avoided entirely. Regarding FFPE samples, this carries even more weight than concerning snap-frozen tissue samples collected at animal sacrifice since they are only touched once with sterile tweezers, then directly frozen in sterile tubes until DNA isolation. Tissue-holding paraffin blocks, especially when derived from other groups or departments, cannot be controlled and have been touched by various people before entering the DNA extraction process. This is a structurally inevitable event, and generally, FFPE samples are the most common form of specimen in clinics, wherefore it is even more important to focus on contamination avoidance strategies and decontamination methods when working with this sample type. Receiving FFPE tissues, not readily isolated DNA, was essential in doing so

because it allowed to control number and kind of NTCs introduced to the sample pool (**figure 43**). Their extraordinary relevance is also described by Eisenhofer and colleagues as the most essential step in controlling for contamination (Eisenhofer et al., 2019). Further ways of dealing with contamination besides taking along NTCs at every step of the workflow are running samples and NTCs on separate FCs to avoid computational cross-over, and eventually, the bioinformatic removal of contamination via filter pipelines and R packages such as decontam.

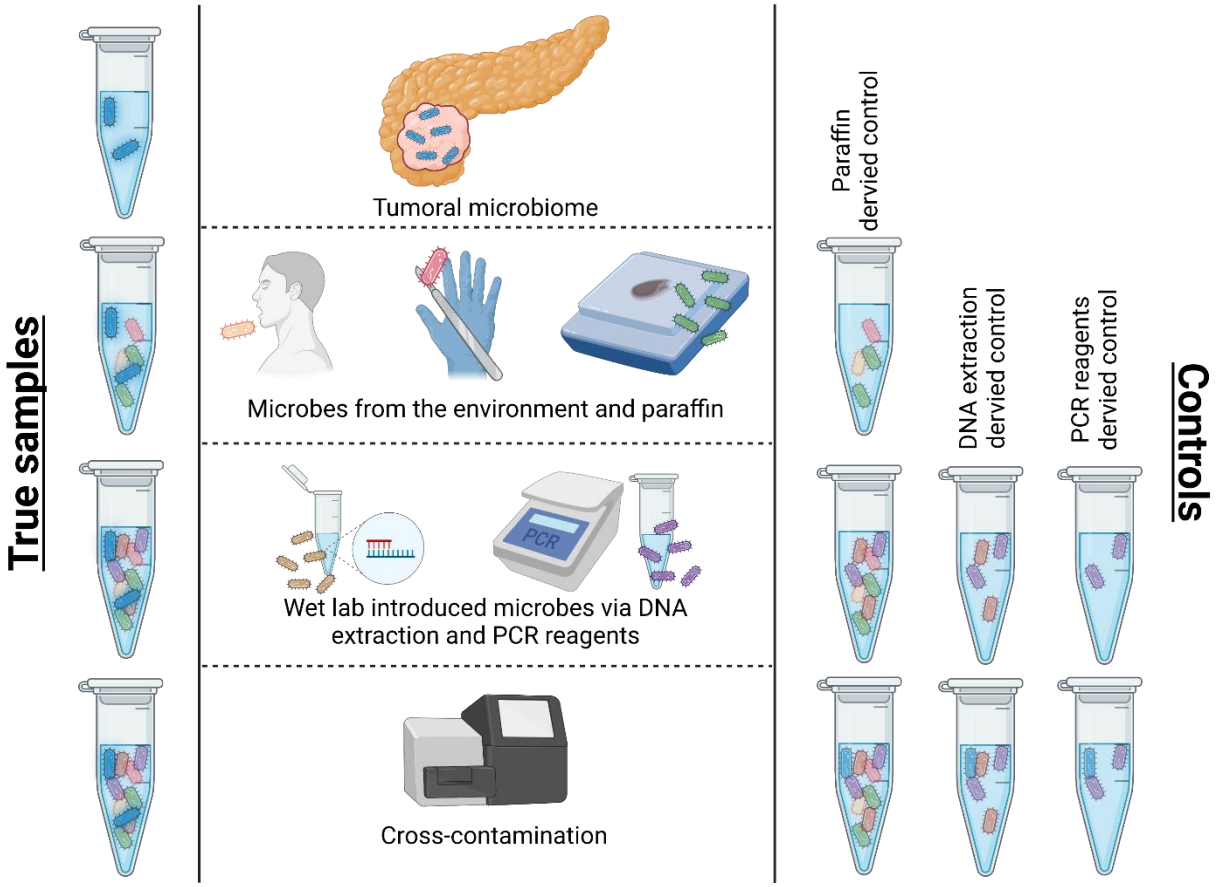

Figure 43: Sources of microbial contamination along the wet- and dry-lab workflow with adequate negative controls. The originally sampled microbiome is enriched with contaminants from the environment and paraffin as well as DNA extraction kit and PCR reagents. Because this cannot be entirely avoided, it is crucial to carry along NTCs from every putative contaminating step. These controls only hold contaminating microbiota which may be removed computationally from the true samples. However, sequencing errors may cause computational cross-contamination via index switches in multiplex sequencing approaches (barcode leakage). Therefore, NTCs should be run on separate FCs in order to prevent the assignment of those sequences to true samples. Illustration taken from (Pfisterer et al., 2022) (CC BY 4.0).

The main issue with low biomass samples is that they may contain as much true species DNA as the NTCs hold DNA concentrations of contaminating species. As discussed in **5.1.2**, samples containing very low abundant but true species provoke the prevalence filter threshold to be set below these abundances leading to the persistence of contaminants of higher abundances. For this reason, it is extremely difficult to find the right thresholds and to filter out the correct species without eradicating all true species in the samples. This can lead to very low read counts and subsequently species numbers. Consequently, in workflows applying rarefaction new issues open up with low rarefaction thresholds only discovering few species in the samples. This may putatively prevent the unraveling of differences in the microbial compositions of distinct groups. In contrast, removing contaminants efficiently may open a chance to render the data significant by unveiling differences that had been covered up by contaminants evenly distributed across all groups to be compared.

In this data set comparing tumors derived from orthotopically transplanted human PDAC cell lines of different subtypes, also no differences between both tumoral microbiomes were encountered. While there is a chance that this might be a biologically true finding, it is de facto debatable whether this really represents the situation in the *in vivo* tumor or whether there are in fact microbial differences hidden between Capan-2- and MiaPaCa2-based tumors. These differences between CLA and BL PDAC subtypes have been reported recently by Guo and colleagues who found the BL subtype of PDAC, characterized by a higher inflammatory response, to have a more diverse microbiome as compared to CLA tumors (Guo et al., 2021). They also hypothesized a functional contribution of the microbiome in shaping the tumor subtype. However, this study worked with human material, whereas the data presented here derives from samples of a xenograft mouse model without a fully functional immune system (extremely low T cell count due to absence of thymus), a key factor in tumor biology, although even in individuals with a functional immune system tumors develop and escape immune surveillance (Gonzalez et al., 2018). Putatively, the immunogenic landscape in the tumor also plays a major role in microbiome formation. Conversely, immunogenic reprogramming of the TME via immune suppression by the PDAC microbiome has been described (Pushalkar et al., 2018). In any case, the impaired immune system in the xenograft mouse model is a serious drawback in this experiment and needs to be kept in mind. It is possible that the environmental contaminants were transferred to the mice during the transplantation procedure or even ingested via the plant-based chow

without the animals being capable of controlling them. *Rhizobium*, a symbiont of legumes, and *Agrobacterium* are associated with mouse chow and have been found to be the most abundant genera in both wild type and KC mouse pancreas before (Pushalkar et al., 2018). Eventually, in the course of establishing the biologically and statistically most legitimate way of analyzing the data, some preliminary results reached statistical significance. This shows how much the results also depend on the analysis strategy and that they always need to be critically reviewed. Although the primary focus in this experiment was set on differences between the two tumor subtypes as well as between the tumor and its corresponding tumor-surrounding histologically normal pancreatic tissue, it is still an interesting finding of the two different tumor-adjacent normal tissues being significantly different from each other in their microbiomes. It argues for the tumor subtype to influence the microbiome in surrounding non-malignant cells that do not (yet) exhibit a transformed phenotype. If this holds true in future repetition experiments, it might give hints towards the microbial role in tumor development and possibly even be exploited for early PDAC subtype discrimination.

5.3.2 First evidence suggests that the tumoral microbiome is impacted by the PDAC subtype

Finally, despite knowing about the pitfalls of contamination and the low immune system in the xenograft mouse model, an attempt was made to approach more functional microbiome studies. A recent study by Tu and Klein et al. investigated the influence of TNFα produced by macrophages on PDAC subtype identity (Tu et al., 2021). These macrophages are actively recruited via MCP1 by the BL neoplastic cells to maintain their pronounced stromal immune landscape and associated aggressiveness. Lineage reprogramming even enabled a subtype switch from CLA to BL triggered by these TNFα-producing macrophages (Tu et al., 2021).

The authors demonstrated this subtype switch by treating orthotopically Capan-1-transplanted nude mice with TNFα and subsequent tissue analysis. Receiving the FFPE tumor tissues from this study, they could be investigated regarding their microbiomes with the aim to get an idea about the impact of the PDAC subtype on the bacterial composition and to find out what drives the tumoral microbiome. If the microbiome changed with the subtype, the tumor subtype, determined by exogenous TNFα recruiting macrophages in turn

secreting more TNFα switching the CLA to a more BL phenotypic state via lineage reprogramming, would also be responsible for an altered microbiome. If the microbiome is not influenced by these cellular processes and alternations, it is more likely that other factors act on the microbial TME component.

The results were not uniform in all three beta diversity metrics evaluated; however, BC indeed showed a significant difference between TNFα- and H_2O-treated Capan-1 tumors. This argues for a change in the bacterial composition, at least focusing on dissimilarities taking abundances into account, in tumors undergoing a forced subtype switch towards a mesenchymal phenotype accompanied by higher invasiveness and chemoresistance (Moffitt et al., 2015). Tu and Klein et al. had stated that cJUN, a transcription factor of the BL gene signature, leads inflammatory immune cells enriched in poorly differentiated BL tumors into TNFα production, thus causing the subtype switch via a TNFα-mediated cytokine storm (Tu et al., 2021). The infiltration of macrophages and general change in the immune landscape may possibly also alter the microbial composition in the TME regarding abundance and types of bacteria. It is conceivable that this only happens as side effect, still it may end up in a positive feedback loop in which the altered microbiome further shapes the tumor subtype identity via immunogenic TME reprogramming (Chakladar et al., 2020; Neesse et al., 2015; Pushalkar et al., 2018). Although a higher number of observed species in TNFα-treated tumors does not seem to fit in this picture expecting the pronounced inflammatory immune environment to eradicate some species, it needs to be kept in mind that in this data set a high number of contaminating species were found that most likely distort the findings to a certain degree.

In view of improving the methodology applied in this thesis, more profound differences between the regarded tissue groups can be expected to be found. This will also facilitate future functional analysis of the subtype-specific microbiome. For instance, the evaluation of chemosensitivity and therapy resistance, respectively, upon bacterial ablation in subtype-specific orthotopic mouse models, as well as the identification of microbiota-derived metabolites, e.g., in bEVs, in the blood of mice with CLA/BL pancreatic tumors are major goals.

5.4 Chapter 4 | The significance of human PDAC tumor microbiomes in patient stratification

5.4.1 Adjacent microbial niches 'tumor' and 'normal' in human PDAC patients differ from each other

Studies on transgenic and xenograft mouse models are indispensable and very useful due to their genetic similarity with humans rendering them a valid and valuable reflection of complex human PDAC. Yet, their lifestyle and other extrinsic factors underlie stringent control und influence modifiable according to the experimental need on all individuals of the regarded group at the same time. This is essential in basic microbiome research since any kind of intrinsic or extrinsic factor can cause microbial changes, so-called dysbiosis, even in healthy subjects. However, translational science requests the transfer and validation of knowledge gained from *in vivo* studies to the human subject. Patients are by far less uniform in their living habits and prerequisites, therefore automatically presenting with diverse microbiomes due to a plethora of reasons other than the disease status of interest. Confounders such as sex, age, diet, or smoking may heavily impact on the microbial composition. For instance, in lung tumors of current smokers, bacteria metabolizing nicotinate and phenolic compounds are found to be significantly enriched (Nejman et al., 2020). Hence, in human studies, it is important to register potential microbiome manipulators such as antibiotics uptake in the metadata and to stratify patients subsequently in order to narrow the influential factors down and to detect true differences between patient groups.

In this experimental setting, the patient background was not incorporated in detail yet, what might give an explanation why the data points do not show such a strict separation and continuous significant difference throughout all beta diversity metrics in comparison with the murine KPC data. However, this is commonly seen in microbiome studies and can probably be traced back to the heterogeneity among the assessed individuals. Nonetheless, human PDAC specimen significantly differed from healthy pancreas samples of the same patients regarding BC. Also in uwUF, a tendency might be observed of the two groups diverging apart. With larger group sizes and patient stratification or outlier elimination, it is likely to achieve more specific results reaching statistical significance. The median OS

number of around 12 is comparable to numbers found in the literature where Nejman et al. for example reported a median OS of approximately 4 for pancreatic tumor samples with maximum OS numbers around 30 in their sample cohort (Nejman et al., 2020).

The differing microbiome of PDAC tumors as opposed to healthy pancreas has been described before and is of exceptional interest with regard to its future employment as a biomarker, for instance (del Castillo et al., 2019; Geller et al., 2017; Pushalkar et al., 2018). Several clinical studies are currently ongoing to elucidate the role of the microbiome in PDAC and different application options. Here, the most interesting finding is that the microbiomes of PDAC tumors tend to differ from histologically healthy pancreas of the same patients, not only pancreas samples of other healthy subjects or even benign pancreatic diseases. Finding locally distinct sub-microbiomes within the same organ is remarkable. As mentioned before in the murine context, this contributes to the understanding of the microbial role in therapy resistance. Although the microbiome might not be different in tumor sub-compartments, the gradient between diseased and healthy parts of one organ calls questions about chemoresistance to mind and why therapeutics often do not reach the tumor.

5.4.2 The biological relevance of a hybrid PDAC form is highlighted by its outstanding microbiome

Arguing that patient stratification is the key to successful PDAC management, the human tumor sequencing data were subdivided according to PDAC subtypes CLA, BL, and HYB in the final analysis approach (Chan-Seng-Yue et al., 2020; Moffitt et al., 2015). Subtyping the samples was performed after IHC staining for markers GATA6 and CK5 as elucidated before. GATA6 is a vital protein for pancreas development and has been recognized as potential oncogene in PC (Decker et al., 2006; Fu et al., 2008). *GATA6* is an unambiguous marker gene for CLA subtype tumors due to its high expression in the latter and low expression in the BL counterparts (Collisson et al., 2011). Its expression was lately tested successfully as surrogate biomarker to discriminate CLA and BL forms of advanced PDAC (O'Kane et al., 2020). The same study also successfully evaluated the possibility of identifying BL tumors by CK5.

The human sample set assessed in this thesis was subtyped according to O'Kane et al. as elucidated in **4.4.2**, with the addition of a third subtype, the recently described HYB form (Chan-Seng-Yue et al., 2020; O'Kane et al., 2020). Surprisingly, despite the small number of 20 samples technically successfully subtyped and the CLA subtype usually being more frequent (Juiz et al., 2020), the three subtypes were evenly distributed. This facilitated statistical analysis which returned unexpected, yet intriguing results. It was assumed to find microbial differences between CLA and BL tumors as published before (Guo et al., 2021). Instead, CLA and BL presented to be quite congruent, while HYB tumors were significantly different from both established PDAC subtypes, at least regarding uwUF and BC. In wUF, this did not reach statistical significance, still the data points tended to cluster. On the one hand, this fits the finding of Capan-2 (CLA) and MiaPaCa2 (BL) xenograft tumors not differing from each other regarding their microbial compositions. On the other hand, it emphasizes the eligibility of the HYB PDAC subtype to be regarded as its independent subtype also based on the properties of its microbiome. Of course, it needs to be considered that the sample numbers are comparably small in this pioneering experiment and yet again, contaminating species are present in the sequenced microbial composition.

The expansion of PDAC subtyping and patient stratification is essential in exploring new vantage points for personalized medicine. Besides transcriptional profiles, also different types of metabolic pathways were employed to discriminate between sub-forms of PDAC. Microbiome-based PDAC subtyping is surely still speculative but might bear profitable chances in the future.

The clinical trial PDA-MAPS (NCT04922515), "PDAC microbiome as predictor of subtypes", tackles this research question employing buccal and rectal swabs. PDAC patients with different tumor subtypes indeed show significant differences in beta diversity in preliminary data. The possible future application of these promising results has the potential to greatly improve PDAC prognosis by differential diagnosis. The study by Riquelme et al. on LTS and STS of PDAC exhibiting different microbiomes further adds to this concept. The answer to the question why some patients outlive PDAC for an extraordinarily long time might implicate instructions in dealing with this deadly disease in order to change the clinical situation for the better. However, the predictive value of the microbiome so far restricts to research studies (Guo et al., 2021). Even the transcriptome-based PDAC subtyping is not

(yet) applied in the clinics regarding patient stratification for personalized medicine, a concept that PDAC would certainly benefit from considering its massive heterogeneity (Topham et al., 2021). Further research efforts are crucial on the way to efficient and successful PDAC management.

6 Conclusion

The thesis at hand unites the fields of oncology and microbiology in a joint research interest and interdisciplinary project. It addresses PDAC, one of the worst human malignancies regarding survival rate, in multiple regards. Divided in four main result chapters, it covers (i) the implementation and optimization of third generation sequencing by Oxford Nanopore Technologies for low biomass samples, (ii) the characterization of murine KPC fecal and tumoral microbiomes and their comparison to appropriate human data with predictive value estimation of stool classifiers, (iii) the investigation of the impact of transcriptional and phenotypic PDAC subtypes on their microbiomes in samples of xenograft mouse models including a functional subtype switch experiment, and (iv) the microbiome analysis of human PDAC with regard to the discrimination from surrounding normal pancreas and considering the subtypes CLA, BL, and HYB after their determination via marker protein IHC.

Modern microbiome research requires efficient high throughput technology. For various reasons, in particular the ability of long fragment sequencing, nanopore sequencing by ONT was chosen to be established. This comprehensive project included the evaluation of seven different DNA extraction protocols, testing of several PCR parameters, and benchmarking of multiple settings in the bioinformatic data processing resulting in a workflow specified for murine (low biomass) samples. It comprises the IHMS-modified Invitrogen protocol for DNA extraction, 50 ng input DNA and 35 PCR cycles in 16S rRNA gene sequencing, or 1 µg input DNA in the metagenomic sequencing protocol, Centrifuge classification using the nt database for snap-frozen samples, or the abv library for FFPE samples, and the Minimap2 alignment parameters AS1000 and Cov50. The successful establishment of this method in our laboratory was the basis for the following experiments.

In the second main project of this thesis, evidence for the presence of bacteria in murine PC was provided via FISH and IHC staining methods. Subsequently, the fecal and tumoral microbiomes of the widely used transgenic KPC mouse model were successfully characterized. Doing so for the first time using Nanopore sequencing, both microbial niches were proven to differ significantly from their corresponding healthy controls and, most interestingly, simultaneously to resemble the situation in human PDAC. Thus, it can be

concluded that the KPC model is a suitable GEMM to study the microbiome in the context of PDAC and the data adds to a profound understanding of this model system which is indispensable for future preclinical studies.

Eventually, FFPE pancreatic tissue samples from xenograft mouse models orthotopically transplanted with immortalized human PDAC cell lines Capan-2 (CLA), MiaPaCa2 (BL), and Capan-1 (CLA) which were additionally treated with TNFα/H_2O, as well as from resected human PDAC patients were studied. The idea of subtyping tumors, in other words stratifying patients by different aspects to maximize treatment efficiency, is no new concept and quite a success so far in other cancer entities. However, in PDAC, this is still under investigation and the missing variety of treatment options remains a problem even if there were reliable tumor subtypes in clinical practice. Connecting the current standard of knowledge on the microbiome and PDAC subtypes, however, is a new vantage point for this research field. The pioneering data presented here are undoubtedly rather preliminary and need further attention with modulation and repetition of the experiments due to contamination issues, as well as the correlation with patient information and orointestinal data. Still, there are promising hints for PDAC subtypes to impinge on the intratumoral microbiome potentially offering new therapeutical options.

After decades of studying PDAC tumor biology and reaching a continuously better understanding of the major molecular mechanisms in this highly complex disease, an age of new perspectives began with the discovery of microbes in and around PDAC. Today, it can no longer be claimed that bacteria and other microbiota are chance discoveries and mere contamination in these tissues. There is an extensive organ-microbiome-crosstalk, yet, the mechanisms are still mostly unknown. In particular, the prime question for the dysbiosis to be contributing cause or consequent outcome of PDAC pathogenesis remains open. Functional studies focusing on subtype-specific microbiome targeting are the future of PDAC (microbiome) research.

BIBLIOGRAPHY

Adamska, A., Domenichini, A., Falasca, M., 2017. Pancreatic Ductal Adenocarcinoma: Current and Evolving Therapies. International Journal of Molecular Sciences 18, 1338. https://doi.org/10.3390/ijms18071338

Adelman, M.W., Woodworth, M.H., Langelier, C., Busch, L.M., Kempker, J.A., Kraft, C.S., Martin, G.S., 2020. The gut microbiome's role in the development, maintenance, and outcomes of sepsis. Crit Care 24, 278. https://doi.org/10.1186/s13054-020-02989-1

Ahuja, M., Schwartz, D.M., Tandon, M., Son, A., Zeng, M., Swaim, W., Eckhaus, M., Hoffman, V., Cui, Y., Xiao, B., Worley, P.F., Muallem, S., 2017. Orai1-Mediated Antimicrobial Secretion from Pancreatic Acini Shapes the Gut Microbiome and Regulates Gut Innate Immunity. Cell Metab 25, 635–646. https://doi.org/10.1016/j.cmet.2017.02.007

Alam, A., Levanduski, E., Denz, P., Villavicencio, H.S., Bhatta, M., Alhorebi, L., Zhang, Y., Gomez, E.C., Morreale, B., Senchanthisai, S., Li, J., Turowski, S.G., Sexton, S., Sait, S.J., Singh, P.K., Wang, J., Maitra, A., Kalinski, P., DePinho, R.A., Wang, H., Liao, W., Abrams, S.I., Segal, B.H., Dey, P., 2022. Fungal mycobiome drives IL-33 secretion and type 2 immunity in pancreatic cancer. Cancer Cell 40, 153-167.e11. https://doi.org/10.1016/j.ccell.2022.01.003

Almoguera, C., Shibata, D., Forrester, K., Martin, J., Arnheim, N., Perucho, M., 1988. Most human carcinomas of the exocrine pancreas contain mutant c-K-ras genes. Cell 53, 549–554. https://doi.org/10.1016/0092-8674(88)90571-5

Amann, R.I., Binder, B.J., Olson, R.J., Chisholm, S.W., Devereux, R., Stahl, D.A., 1990. Combination of 16S rRNA-targeted oligonucleotide probes with flow cytometry for analyzing mixed microbial populations. Applied and Environmental Microbiology 56, 1919–1925. https://doi.org/10.1128/aem.56.6.1919-1925.1990

Ammer-Herrmenau, C., Asendorf, T., Beyer, G., Buchholz, S.M., Cameron, S., Damm, M., Frost, F., Henker, R., Jaster, R., Phillip, V., Placzek, M., Ratei, C., Sirtl, S., van den Berg, T., Weingarten, M.J., Woitalla, J., Mayerle, J., Ellenrieder, V., Neesse, A., 2021a. Study protocol P-MAPS: microbiome as predictor of severity in acute pancreatitis-a prospective multicentre translational study. BMC Gastroenterol 21, 304. https://doi.org/10.1186/s12876-021-01885-4

Ammer-Herrmenau, C., Pfisterer, N., van den Berg, T., Gavrilova, I., Amanzada, A., Singh, S.K., Khalil, A., Alili, R., Belda, E., Clement, K., Abd El Wahed, A., Gady, E.E., Haubrock, M., Beißbarth, T., Ellenrieder, V., Neesse, A., 2021b. Comprehensive Wet-Bench and Bioinformatics Workflow for Complex Microbiota Using Oxford Nanopore Technologies. mSystems 6, e00750-21. https://doi.org/10.1128/mSystems.00750-21

Ammer-Herrmenau, C., Pfisterer, N., Weingarten, M.F., Neesse, A., 2020. The microbiome in pancreatic diseases: Recent advances and future perspectives. United European Gastroenterology Journal 8, 878–885. https://doi.org/10.1177/2050640620944720

Aune, D., Greenwood, D.C., Chan, D.S.M., Vieira, R., Vieira, A.R., Navarro Rosenblatt, D.A., Cade, J.E., Burley, V.J., Norat, T., 2012. Body mass index, abdominal fatness and pancreatic cancer risk: a systematic review and non-linear dose–response meta-analysis of prospective studies. Annals of Oncology 23, 843–852. https://doi.org/10.1093/annonc/mdr398

Aykut, B., Pushalkar, S., Chen, R., Li, Q., Abengozar, R., Kim, J.I., Shadaloey, S.A., Wu, D., Preiss, P., Verma, N., Guo, Y., Saxena, A., Vardhan, M., Diskin, B., Wang, W., Leinwand, J., Kurz, E., Kochen Rossi, J.A., Hundeyin, M., Zambrinis, C., Li, X., Saxena, D., Miller, G., 2019. The fungal mycobiome promotes pancreatic oncogenesis via activation of MBL. Nature 574, 264–267. https://doi.org/10.1038/s41586-019-1608-2

Bailey, P., Chang, D.K., Nones, K., Johns, A.L., Patch, A.-M., Gingras, M.-C., Miller, D.K., Christ, A.N., Bruxner, T.J.C., Quinn, M.C., Nourse, C., Murtaugh, L.C., Harliwong, I., Idrisoglu, S., Manning, S., Nourbakhsh, E., Wani, S., Fink, L., Holmes, O., Chin, V., Anderson, M.J., Kazakoff, S., Leonard, C., Newell, F., Waddell, Nick, Wood, S., Xu, Q., Wilson, P.J., Cloonan, N., Kassahn, K.S., Taylor, D., Quek, K., Robertson, A., Pantano, L., Mincarelli, L., Sanchez, L.N., Evers, L., Wu, J., Pinese, M., Cowley, M.J., Jones, M.D., Colvin, E.K., Nagrial, A.M., Humphrey, E.S., Chantrill, L.A., Mawson, A., Humphris, J., Chou, A., Pajic, M., Scarlett, C.J., Pinho, A.V., Giry-Laterriere, M., Rooman, I., Samra, J.S., Kench, J.G., Lovell, J.A., Merrett, N.D., Toon, C.W., Epari, K., Nguyen, N.Q., Barbour, A., Zeps, N., Moran-Jones, K., Jamieson, N.B., Graham, J.S., Duthie, F., Oien, K., Hair, J., Grützmann, R., Maitra, A., Iacobuzio-Donahue, C.A., Wolfgang, C.L., Morgan, R.A., Lawlor, R.T., Corbo, V., Bassi, C., Rusev, B., Capelli, P., Salvia, R., Tortora, G., Mukhopadhyay, D., Petersen, G.M., Munzy, D.M., Fisher, W.E., Karim, S.A., Eshleman, J.R., Hruban, R.H., Pilarsky, C., Morton, J.P., Sansom, O.J., Scarpa, A., Musgrove, E.A., Bailey, U.-M.H., Hofmann, O., Sutherland, R.L., Wheeler, D.A., Gill, A.J., Gibbs, R.A., Pearson, J.V., Waddell, Nicola, Biankin, A.V., Grimmond, S.M., 2016. Genomic analyses identify molecular subtypes of pancreatic cancer. Nature 531, 47–52. https://doi.org/10.1038/nature16965

Baquero, F., Nombela, C., 2012. The microbiome as a human organ. Clin Microbiol Infect 18 Suppl 4, 2–4. https://doi.org/10.1111/j.1469-0691.2012.03916.x

Batterham, R.L., Le Roux, C.W., Cohen, M.A., Park, A.J., Ellis, S.M., Patterson, M., Frost, G.S., Ghatei, M.A., Bloom, S.R., 2003. Pancreatic polypeptide reduces appetite and food intake in humans. J Clin Endocrinol Metab 88, 3989–3992. https://doi.org/10.1210/jc.2003-030630

Baum, J., Simons, B.E., Unger, R.H., Madison, L.L., 1962. Localization of glucagon in the alpha cells in the pancreatic islet by immunofluorescent technics. Diabetes 11, 371–374.

Beaugerie, L., Petit, J.-C., 2004. Antibiotic-associated diarrhoea. Best Practice & Research Clinical Gastroenterology 18, 337–352. https://doi.org/10.1016/j.bpg.2003.10.002

Ben, Q., Xu, M., Ning, X., Liu, J., Hong, S., Huang, W., Zhang, H., Li, Z., 2011. Diabetes mellitus and risk of pancreatic cancer: A meta-analysis of cohort studies. European Journal of Cancer 47, 1928–1937. https://doi.org/10.1016/j.ejca.2011.03.003

Berg, G., Rybakova, D., Fischer, D., Cernava, T., Vergès, M.-C.C., Charles, T., Chen, X., Cocolin, L., Eversole, K., Corral, G.H., Kazou, M., Kinkel, L., Lange, L., Lima, N., Loy, A., Macklin, J.A., Maguin, E., Mauchline, T., McClure, R., Mitter, B., Ryan, M., Sarand, I., Smidt, H., Schelkle, B., Roume, H., Kiran, G.S., Selvin, J., Souza, R.S.C. de, van Overbeek, L., Singh, B.K., Wagner, M., Walsh, A., Sessitsch, A., Schloter, M., 2020. Microbiome definition re-visited: old concepts and new challenges. Microbiome 8, 103. https://doi.org/10.1186/s40168-020-00875-0

Biffi, G., Oni, T.E., Spielman, B., Hao, Y., Elyada, E., Park, Y., Preall, J., Tuveson, D.A., 2019. IL1-Induced JAK/STAT Signaling Is Antagonized by TGFβ to Shape CAF Heterogeneity in Pancreatic Ductal Adenocarcinoma. Cancer Discov 9, 282–301. https://doi.org/10.1158/2159-8290.CD-18-0710

Bjerre, R.D., Hugerth, L.W., Boulund, F., Seifert, M., Johansen, J.D., Engstrand, L., 2019. Effects of sampling strategy and DNA extraction on human skin microbiome investigations. Sci Rep 9, 17287. https://doi.org/10.1038/s41598-019-53599-z

Bockman, D.E., 1993. Anatomy of the Pancreas. In: Go VLW, DiMagno EP, Gardner JD, Lebenthal E, Reber HA, Scheele GA, eds. The Pancreas: Biology, Pathobiology, and Disease, New York: Raven Press, 1993:1-8.

Bosetti, C., Lucenteforte, E., Silverman, D.T., Petersen, G., Bracci, P.M., Ji, B.T., Negri, E., Li, D., Risch, H.A., Olson, S.H., Gallinger, S., Miller, A.B., Bueno-de-Mesquita, H.B., Talamini, R., Polesel, J., Ghadirian, P., Baghurst, P.A., Zatonski, W., Fontham, E., Bamlet, W.R., Holly, E.A., Bertuccio, P., Gao, Y.T., Hassan, M., Yu, H., Kurtz, R.C., Cotterchio, M., Su, J., Maisonneuve, P., Duell, E.J., Boffetta, P., La Vecchia, C., 2012. Cigarette smoking and pancreatic cancer: an analysis from the International Pancreatic Cancer Case-Control Consortium (Panc4). Annals of Oncology 23, 1880–1888. https://doi.org/10.1093/annonc/mdr541

Bray, J.R., Curtis, J.T., 1957. An Ordination of the Upland Forest Communities of Southern Wisconsin. Ecological Monographs 27, 325–349. https://doi.org/10.2307/1942268

Brody, H., 2020. The gut microbiome. Nature 577, S5. https://doi.org/10.1038/d41586-020-00194-2

Buechler, M.B., Pradhan, R.N., Krishnamurty, A.T., Cox, C., Calviello, A.K., Wang, A.W., Yang, Y.A., Tam, L., Caothien, R., Roose-Girma, M., Modrusan, Z., Arron, J.R., Bourgon, R., Müller, S., Turley, S.J., 2021. Cross-tissue organization of the fibroblast lineage. Nature 593, 575–579. https://doi.org/10.1038/s41586-021-03549-5

Bukin, Y.S., Galachyants, Y.P., Morozov, I.V., Bukin, S.V., Zakharenko, A.S., Zemskaya, T.I., 2019. The effect of 16S rRNA region choice on bacterial community metabarcoding results. Sci Data 6, 190007. https://doi.org/10.1038/sdata.2019.7

Caglar, V., Kumral, B., Uygur, R., Alkoc, O.A., Ozen, O.A., Demirel, H., 2014. Study of Volume, Weight and Size of Normal Pancreas, Spleen and Kidney in Adults Autopsies. Forensic Medicine and Anatomy Research 02, 63. https://doi.org/10.4236/fmar.2014.23012

Chakladar, J., Kuo, S.Z., Castaneda, G., Li, W.T., Gnanasekar, A., Yu, M.A., Chang, E.Y., Wang, X.Q., Ongkeko, W.M., 2020. The Pancreatic Microbiome is Associated with Carcinogenesis and Worse Prognosis in Males and Smokers. Cancers 12, 2672. https://doi.org/10.3390/cancers12092672

Chan-Seng-Yue, M., Kim, J.C., Wilson, G.W., Ng, K., Figueroa, E.F., O'Kane, G.M., Connor, A.A., Denroche, R.E., Grant, R.C., McLeod, J., Wilson, J.M., Jang, G.H., Zhang, A., Dodd, A., Liang, S.-B., Borgida, A., Chadwick, D., Kalimuthu, S., Lungu, I., Bartlett, J.M.S., Krzyzanowski, P.M., Sandhu, V., Tiriac, H., Froeling, F.E.M., Karasinska, J.M., Topham, J.T., Renouf, D.J., Schaeffer, D.F., Jones, S.J.M., Marra, M.A., Laskin, J., Chetty, R., Stein, L.D., Zogopoulos, G., Haibe-Kains, B., Campbell, P.J., Tuveson, D.A., Knox, J.J., Fischer, S.E., Gallinger, S., Notta, F., 2020. Transcription phenotypes of pancreatic cancer are driven by genomic events during tumor evolution. Nat Genet 52, 231–240. https://doi.org/10.1038/s41588-019-0566-9

Chiang, Y.-T., Chien, Y.-C., Lin, Y.-H., Wu, H.-H., Lee, D.-F., Yu, Y.-L., 2021. The Function of the Mutant p53-R175H in Cancer. Cancers (Basel) 13, 4088. https://doi.org/10.3390/cancers13164088

Cicenas, J., Kvederaviciute, K., Meskinyte, I., Meskinyte-Kausiliene, E., Skeberdyte, A., Cicenas, J., 2017. KRAS, TP53, CDKN2A, SMAD4, BRCA1, and BRCA2 Mutations in Pancreatic Cancer. Cancers (Basel) 9, 42. https://doi.org/10.3390/cancers9050042

Collisson, E.A., Sadanandam, A., Olson, P., Gibb, W.J., Truitt, M., Gu, S., Cooc, J., Weinkle, J., Kim, G.E., Jakkula, L., Feiler, H.S., Ko, A.H., Olshen, A.B., Danenberg, K.L., Tempero, M.A., Spellman, P.T., Hanahan, D., Gray, J.W., 2011. Subtypes of pancreatic ductal adenocarcinoma and their differing responses to therapy. Nat Med 17, 500–503. https://doi.org/10.1038/nm.2344

Conroy, T., Desseigne, F., Ychou, M., Bouché, O., Guimbaud, R., Bécouarn, Y., Adenis, A., Raoul, J.-L., Gourgou-Bourgade, S., de la Fouchardière, C., Bennouna, J., Bachet, J.-B., Khemissa-Akouz, F., Péré-Vergé, D., Delbaldo, C., Assenat, E., Chauffert, B., Michel, P., Montoto-Grillot, C., Ducreux, M., 2011. FOLFIRINOX versus Gemcitabine for Metastatic Pancreatic Cancer. N Engl J Med 364, 1817–1825. https://doi.org/10.1056/NEJMoa1011923

Covell, W.P., 1928. A microscopic study of pancreatic secretion in the living animal. The Anatomical Record 40, 213–223. https://doi.org/10.1002/ar.1090400204

Cullin, N., Azevedo Antunes, C., Straussman, R., Stein-Thoeringer, C.K., Elinav, E., 2021. Microbiome and cancer. Cancer Cell 39, 1317–1341. https://doi.org/10.1016/j.ccell.2021.08.006

Daemen, A., Peterson, D., Sahu, N., McCord, R., Du, X., Liu, B., Kowanetz, K., Hong, R., Moffat, J., Gao, M., Boudreau, A., Mroue, R., Corson, L., O'Brien, T., Qing, J., Sampath, D., Merchant, M., Yauch, R., Manning, G., Settleman, J., Hatzivassiliou, G., Evangelista, M., 2015. Metabolite profiling stratifies pancreatic ductal adenocarcinomas into subtypes with distinct sensitivities to metabolic inhibitors. Proceedings of the National Academy of Sciences 112, E4410–E4417. https://doi.org/10.1073/pnas.1501605112

D'Agostino, R.B., Pencina, M.J., Massaro, J.M., Coady, S., 2013. Cardiovascular Disease Risk Assessment: Insights from Framingham. Glob Heart 8, 11–23. https://doi.org/10.1016/j.gheart.2013.01.001

Davis, N.M., Proctor, D.M., Holmes, S.P., Relman, D.A., Callahan, B.J., 2018. Simple statistical identification and removal of contaminant sequences in marker-gene and metagenomics data. Microbiome 6, 226. https://doi.org/10.1186/s40168-018-0605-2

de Martel, C., Ferlay, J., Franceschi, S., Vignat, J., Bray, F., Forman, D., Plummer, M., 2012. Global burden of cancers attributable to infections in 2008: a review and synthetic analysis. The Lancet Oncology 13, 607–615. https://doi.org/10.1016/S1470-2045(12)70137-7

Decker, K., Goldman, D.C., L. Grasch, C., Sussel, L., 2006. Gata6 is an important regulator of mouse pancreas development. Developmental Biology 298, 415–429. https://doi.org/10.1016/j.ydbio.2006.06.046

del Castillo, E., Meier, R., Chung, M., Koestler, D.C., Chen, T., Paster, B.J., Charpentier, K.P., Kelsey, K.T., Izard, J., Michaud, D.S., 2019. The Microbiomes of Pancreatic and Duodenum Tissue Overlap and Are Highly Subject Specific but Differ between Pancreatic Cancer and Noncancer Subjects. Cancer Epidemiology, Biomarkers & Prevention 28, 370–383. https://doi.org/10.1158/1055-9965.EPI-18-0542

Dolenšek, J., Rupnik, M.S., Stožer, A., 2015. Structural similarities and differences between the human and the mouse pancreas. Islets 7, e1024405. https://doi.org/10.1080/19382014.2015.1024405

Duell, E.J., Lucenteforte, E., Olson, S.H., Bracci, P.M., Li, D., Risch, H.A., Silverman, D.T., Ji, B.T., Gallinger, S., Holly, E.A., Fontham, E.H., Maisonneuve, P., Bueno-de-Mesquita, H.B., Ghadirian, P., Kurtz, R.C., Ludwig, E., Yu, H., Lowenfels, A.B., Seminara, D., Petersen, G.M., La Vecchia, C., Boffetta, P., 2012. Pancreatitis and pancreatic cancer

risk: a pooled analysis in the International Pancreatic Cancer Case-Control Consortium (PanC4). Annals of Oncology 23, 2964–2970. https://doi.org/10.1093/annonc/mds140

Eisenhofer, R., Minich, J.J., Marotz, C., Cooper, A., Knight, R., Weyrich, L.S., 2019. Contamination in Low Microbial Biomass Microbiome Studies: Issues and Recommendations. Trends in Microbiology 27, 105–117. https://doi.org/10.1016/j.tim.2018.11.003

Elyada, E., Bolisetty, M., Laise, P., Flynn, W.F., Courtois, E.T., Burkhart, R.A., Teinor, J.A., Belleau, P., Biffi, G., Lucito, M.S., Sivajothi, S., Armstrong, T.D., Engle, D.D., Yu, K.H., Hao, Y., Wolfgang, C.L., Park, Y., Preall, J., Jaffee, E.M., Califano, A., Robson, P., Tuveson, D.A., 2019. Cross-Species Single-Cell Analysis of Pancreatic Ductal Adenocarcinoma Reveals Antigen-Presenting Cancer-Associated Fibroblasts. Cancer Discov 9, 1102–1123. https://doi.org/10.1158/2159-8290.CD-19-0094

Eresen, A., Yang, J., Shangguan, J., Li, Y., Hu, S., Sun, C., Yaghmai, V., Benson III, A.B., Zhang, Z., 2020. Prediction of therapeutic outcome and survival in a transgenic mouse model of pancreatic ductal adenocarcinoma treated with dendritic cell vaccination or CDK inhibitor using MRI texture: a feasibility study. Am J Transl Res 12, 2201–2211.

Erstad, D.J., Sojoodi, M., Taylor, M.S., Ghoshal, S., Razavi, A.A., Graham-O'Regan, K.A., Bardeesy, N., Ferrone, C.R., Lanuti, M., Caravan, P., Tanabe, K.K., Fuchs, B.C., 2018. Orthotopic and heterotopic murine models of pancreatic cancer and their different responses to FOLFIRINOX chemotherapy. Dis Model Mech 11, dmm034793. https://doi.org/10.1242/dmm.034793

Espiau-Romera, P., Courtois, S., Parejo-Alonso, B., Sancho, P., 2020. Molecular and Metabolic Subtypes Correspondence for Pancreatic Ductal Adenocarcinoma Classification. Journal of Clinical Medicine 9, 4128. https://doi.org/10.3390/jcm9124128

Fan, X., Alekseyenko, A.V., Wu, J., Peters, B.A., Jacobs, E.J., Gapstur, S.M., Purdue, M.P., Abnet, C.C., Stolzenberg-Solomon, R., Miller, G., Ravel, J., Hayes, R.B., Ahn, J., 2016. Human oral microbiome and prospective risk for pancreatic cancer: a population-based nested case-control study. Gut 67, 120–127. https://doi.org/10.1136/gutjnl-2016-312580

Farrell, J.J., Zhang, L., Zhou, H., Chia, D., Elashoff, D., Akin, D., Paster, B.J., Joshipura, K., Wong, D.T.W., 2012. Variations of oral microbiota are associated with pancreatic diseases including pancreatic cancer. Gut 61, 582–588. https://doi.org/10.1136/gutjnl-2011-300784

Ferlay, J., Partensky, C., Bray, F., 2016. More deaths from pancreatic cancer than breast cancer in the EU by 2017. Acta Oncologica 55, 1158–1160. https://doi.org/10.1080/0284186X.2016.1197419

Fischer, C.G., Wood, L.D., 2018. From somatic mutation to early detection: insights from molecular characterization of pancreatic cancer precursor lesions. The Journal of Pathology 246, 395–404. https://doi.org/10.1002/path.5154

Flory, C.M., Thal, A., 1947. A transparent chamber for the observation of the pancreas of the living mouse. The Anatomical Record 97, 33–39. https://doi.org/10.1002/ar.1090970105

Frost, F., Kacprowski, T., Rühlemann, M., Bülow, R., Kühn, J.-P., Franke, A., Heinsen, F.-A., Pietzner, M., Nauck, M., Völker, U., Völzke, H., Aghdassi, A.A., Sendler, M., Mayerle, J., Weiss, F.U., Homuth, G., Lerch, M.M., 2019. Impaired Exocrine Pancreatic Function Associates With Changes in Intestinal Microbiota Composition and Diversity. Gastroenterology 156, 1010–1015. https://doi.org/10.1053/j.gastro.2018.10.047

Fu, B., Luo, M., Lakkur, S., Lucito, R., Iacobuzio-Donahue, C.A., 2008. Frequent genomic copy number gain and overexpression of GATA-6 in pancreatic carcinoma. Cancer Biology & Therapy 7, 1593–1601. https://doi.org/10.4161/cbt.7.10.6565

Galeano Niño, J.L., Wu, H., LaCourse, K.D., Kempchinsky, A.G., Baryiames, A., Barber, B., Futran, N., Houlton, J., Sather, C., Sicinska, E., Taylor, A., Minot, S.S., Johnston, C.D., Bullman, S., 2022. Effect of the intratumoral microbiota on spatial and cellular heterogeneity in cancer. Nature 611, 810–817. https://doi.org/10.1038/s41586-022-05435-0

Geller, L.T., Barzily-Rokni, M., Danino, T., Jonas, O.H., Shental, N., Nejman, D., Gavert, N., Zwang, Y., Cooper, Z.A., Shee, K., Thaiss, C.A., Reuben, A., Livny, J., Avraham, R., Frederick, D.T., Ligorio, M., Chatman, K., Johnston, S.E., Mosher, C.M., Brandis, A., Fuks, G., Gurbatri, C., Gopalakrishnan, V., Kim, M., Hurd, M.W., Katz, M., Fleming, J., Maitra, A., Smith, D.A., Skalak, M., Bu, J., Michaud, M., Trauger, S.A., Barshack, I., Golan, T., Sandbank, J., Flaherty, K.T., Mandinova, A., Garrett, W.S., Thayer, S.P., Ferrone, C.R., Huttenhower, C., Bhatia, S.N., Gevers, D., Wargo, J.A., Golub, T.R., Straussman, R., 2017. Potential role of intratumor bacteria in mediating tumor resistance to the chemotherapeutic drug gemcitabine. Science 357, 1156–1160. https://doi.org/10.1126/science.aah5043

Geng, F., Zhang, Y., Lu, Z., Zhang, S., Pan, Y., 2020. Fusobacterium nucleatum Caused DNA Damage and Promoted Cell Proliferation by the Ku70/p53 Pathway in Oral Cancer Cells. DNA and Cell Biology 39, 144–151. https://doi.org/10.1089/dna.2019.5064

Ghiorzo, P., 2014. Genetic predisposition to pancreatic cancer. World Journal of Gastroenterology 20, 10778–10789. https://doi.org/10.3748/wjg.v20.i31.10778

Gibson, G.R., Roberfroid, M.B., 1995. Dietary Modulation of the Human Colonic Microbiota: Introducing the Concept of Prebiotics. The Journal of Nutrition 125, 1401–1412. https://doi.org/10.1093/jn/125.6.1401

Glassing, A., Dowd, S.E., Galandiuk, S., Davis, B., Chiodini, R.J., 2016. Inherent bacterial DNA contamination of extraction and sequencing reagents may affect interpretation of microbiota in low bacterial biomass samples. Gut Pathogens 8, 24. https://doi.org/10.1186/s13099-016-0103-7

Glaubitz, J., Wilden, A., Frost, F., Ameling, S., Homuth, G., Mazloum, H., Rühlemann, M.C., Bang, C., Aghdassi, A.A., Budde, C., Pickartz, T., Franke, A., Bröker, B.M., Voelker, U., Mayerle, J., Lerch, M.M., Weiss, F.-U., Sendler, M., 2023. Activated regulatory T-cells promote duodenal bacterial translocation into necrotic areas in severe acute pancreatitis. Gut. https://doi.org/10.1136/gutjnl-2022-327448

Gonzalez, H., Hagerling, C., Werb, Z., 2018. Roles of the immune system in cancer: from tumor initiation to metastatic progression. Genes Dev 32, 1267–1284. https://doi.org/10.1101/gad.314617.118

Gopinathan, A., Morton, J.P., Jodrell, D.I., Sansom, O.J., 2015. GEMMs as preclinical models for testing pancreatic cancer therapies. Disease Models & Mechanisms 8, 1185–1200. https://doi.org/10.1242/dmm.021055

Gorkiewicz, G., Moschen, A., 2018. Gut microbiome: a new player in gastrointestinal disease. Virchows Arch 472, 159–172. https://doi.org/10.1007/s00428-017-2277-x

Grünwald, B.T., Devisme, A., Andrieux, G., Vyas, F., Aliar, K., McCloskey, C.W., Macklin, A., Jang, G.H., Denroche, R., Romero, J.M., Bavi, P., Bronsert, P., Notta, F., O'Kane, G., Wilson, J., Knox, J., Tamblyn, L., Udaskin, M., Radulovich, N., Fischer, S.E., Boerries, M., Gallinger, S., Kislinger, T., Khokha, R., 2021. Spatially confined sub-

tumor microenvironments in pancreatic cancer. Cell 184, 5577-5592.e18. https://doi.org/10.1016/j.cell.2021.09.022

Guarner, F., Malagelada, J.-R., 2003. Gut flora in health and disease. The Lancet 361, 512–519. https://doi.org/10.1016/S0140-6736(03)12489-0

Guenther, M., Gil, L., Surendran, S.A., Palm, M.A., Heinemann, V., von Bergwelt-Baildon, M., Mayerle, J., Engel, J., Werner, J., Boeck, S., Ormanns, S., 2022. Bacterial Lipopolysaccharide as a Negative Predictor of Adjuvant Gemcitabine Efficacy in Pancreatic Cancer. JNCI Cancer Spectrum 6, pkac039. https://doi.org/10.1093/jncics/pkac039

Guo, J., Xie, K., Zheng, S., 2016. Molecular Biomarkers of Pancreatic Intraepithelial Neoplasia and Their Implications in Early Diagnosis and Therapeutic Intervention of Pancreatic Cancer. International Journal of Biological Sciences 12, 292–301. https://doi.org/10.7150/ijbs.14995

Guo, W., Zhang, Yuchao, Guo, S., Mei, Z., Liao, H., Dong, H., Wu, K., Ye, H., Zhang, Yuhang, Zhu, Y., Lang, J., Hu, L., Jin, G., Kong, X., 2021. Tumor microbiome contributes to an aggressive phenotype in the basal-like subtype of pancreatic cancer. Commun Biol 4, 1–13. https://doi.org/10.1038/s42003-021-02557-5

Half, E., Keren, N., Reshef, L., Dorfman, T., Lachter, I., Kluger, Y., Reshef, N., Knobler, H., Maor, Y., Stein, A., Konikoff, F.M., Gophna, U., 2019. Fecal microbiome signatures of pancreatic cancer patients. Sci Rep 9, 16801. https://doi.org/10.1038/s41598-019-53041-4

Han, K., Nam, J., Xu, J., Sun, X., Huang, X., Animasahun, O., Achreja, A., Jeon, J.H., Pursley, B., Kamada, N., Chen, G.Y., Nagrath, D., Moon, J.J., 2021. Generation of systemic antitumour immunity via the in situ modulation of the gut microbiome by an orally administered inulin gel. Nat Biomed Eng 5, 1377–1388. https://doi.org/10.1038/s41551-021-00749-2

Heinemann, V., Haas, M., Boeck, S., 2013. Neoadjuvant treatment of borderline resectable and non-resectable pancreatic cancer. Annals of Oncology 24, 2484–2492. https://doi.org/10.1093/annonc/mdt239

Henderson, J.R., Daniel, P.M., Fraser, P.A., 1981. The pancreas as a single organ: the influence of the endocrine upon the exocrine part of the gland. Gut 22, 158–167. https://doi.org/10.1136/gut.22.2.158

Hingorani, S.R., Petricoin, E.F., Maitra, A., Rajapakse, V., King, C., Jacobetz, M.A., Ross, S., Conrads, T.P., Veenstra, T.D., Hitt, B.A., Kawaguchi, Y., Johann, D., Liotta, L.A., Crawford, H.C., Putt, M.E., Jacks, T., Wright, C.V.E., Hruban, R.H., Lowy, A.M., Tuveson, D.A., 2003. Preinvasive and invasive ductal pancreatic cancer and its early detection in the mouse. Cancer Cell 4, 437–450. https://doi.org/10.1016/S1535-6108(03)00309-X

Hingorani, S.R., Wang, L., Multani, A.S., Combs, C., Deramaudt, T.B., Hruban, R.H., Rustgi, A.K., Chang, S., Tuveson, D.A., 2005. Trp53R172H and KrasG12D cooperate to promote chromosomal instability and widely metastatic pancreatic ductal adenocarcinoma in mice. Cancer Cell 7, 469–483. https://doi.org/10.1016/j.ccr.2005.04.023

Hosein, A.N., Dougan, S.K., Aguirre, A.J., Maitra, A., 2022. Translational advances in pancreatic ductal adenocarcinoma therapy. Nat Cancer 3, 272–286. https://doi.org/10.1038/s43018-022-00349-2

Hruban, R.H., Adsay, N.V., Albores–Saavedra, J., Compton, C., Garrett, E.S., Goodman, S.N., Kern, S.E., Klimstra, D.S., Klöppel, G., Longnecker, D.S., Lüttges, J., Offerhaus, G.J.A.,

2001. Pancreatic Intraepithelial Neoplasia: A New Nomenclature and Classification System for Pancreatic Duct Lesions. The American Journal of Surgical Pathology 25, 579.

IARC Working Group on the Evaluation of Carcinogenic Risks to Humans, 2012. Biological agents. Volume 100 B. A review of human carcinogens. IARC Monogr Eval Carcinog Risks Hum 100, 1–441.

Innes, J.T., Carey, L.C., 1994. Normal pancreatic dimensions in the adult human. The American Journal of Surgery 167, 261–263. https://doi.org/10.1016/0002-9610(94)90088-4

Juiz, N., Elkaoutari, A., Bigonnet, M., Gayet, O., Roques, J., Nicolle, R., Iovanna, J., Dusetti, N., 2020. Basal-like and classical cells coexist in pancreatic cancer revealed by single-cell analysis on biopsy-derived pancreatic cancer organoids from the classical subtype. FASEB J 34, 12214–12228. https://doi.org/10.1096/fj.202000363RR

Kanda, M., Matthaei, H., Wu, J., Hong, S.-M., Yu, J., Borges, M., Hruban, R.H., Maitra, A., Kinzler, K., Vogelstein, B., Goggins, M., 2012. Presence of Somatic Mutations in Most Early-Stage Pancreatic Intraepithelial Neoplasia. Gastroenterology 142, 730-733.e9. https://doi.org/10.1053/j.gastro.2011.12.042

Kartal, E., Schmidt, T.S.B., Molina-Montes, E., Rodríguez-Perales, S., Wirbel, J., Maistrenko, O.M., Akanni, W.A., Alhamwe, B.A., Alves, R.J., Carrato, A., Erasmus, H.-P., Estudillo, L., Finkelmeier, F., Fullam, A., Glazek, A.M., Gómez-Rubio, P., Hercog, R., Jung, F., Kandels, S., Kersting, S., Langheinrich, M., Márquez, M., Molero, X., Orakov, A., Rossum, T.V., Torres-Ruiz, R., Telzerow, A., Zych, K., Investigators, M.S., Investigators, P.S., Benes, V., Zeller, G., Trebicka, J., Real, F.X., Malats, N., Bork, P., 2022. A faecal microbiota signature with high specificity for pancreatic cancer. Gut 71, 1359–1372. https://doi.org/10.1136/gutjnl-2021-324755

Kaune, T., Griesmann, H., Theuerkorn, K., Hämmerle, M., Laumen, H., Krug, S., Plumeier, I., Kahl, S., Junca, H., Gustavo Dos Anjos Borges, L., Michl, P., Pieper, D.H., Rosendahl, J., 2023. Gender-specific changes of the gut microbiome correlate with tumor development in murine models of pancreatic cancer. iScience 26, 106841. https://doi.org/10.1016/j.isci.2023.106841

Kesh, K., Mendez, R., Mateo-Victoriano, B., Garrido, V.T., Durden, B., Gupta, V.K., Oliveras Reyes, A., Merchant, N., Datta, J., Banerjee, Santanu, Banerjee, Sulagna, 2022. Obesity enriches for tumor protective microbial metabolites and treatment refractory cells to confer therapy resistance in PDAC. Gut Microbes 14, 2096328. https://doi.org/10.1080/19490976.2022.2096328

Kim, J.R., Han, K., Han, Y., Kang, N., Shin, T.-S., Park, H.J., Kim, H., Kwon, W., Lee, S., Kim, Y.-K., Park, T., Jang, J.-Y., 2021. Microbiome Markers of Pancreatic Cancer Based on Bacteria-Derived Extracellular Vesicles Acquired from Blood Samples: A Retrospective Propensity Score Matching Analysis. Biology 10, 219. https://doi.org/10.3390/biology10030219

Kim, S.I., Kang, N., Leem, S., Yang, J., Jo, H., Lee, M., Kim, H.S., Dhanasekaran, D.N., Kim, Y.-K., Park, T., Song, Y.S., 2020. Metagenomic Analysis of Serum Microbe-Derived Extracellular Vesicles and Diagnostic Models to Differentiate Ovarian Cancer and Benign Ovarian Tumor. Cancers 12, 1309. https://doi.org/10.3390/cancers12051309

Kiss, B., Mikó, E., Sebő, É., Toth, J., Ujlaki, G., Szabó, J., Uray, K., Bai, P., Árkosy, P., 2020. Oncobiosis and Microbial Metabolite Signaling in Pancreatic Adenocarcinoma. Cancers 12, 1068. https://doi.org/10.3390/cancers12051068

Kleeff, J., Korc, M., Apte, M., La Vecchia, C., Johnson, C.D., Biankin, A.V., Neale, R.E., Tempero, M., Tuveson, D.A., Hruban, R.H., Neoptolemos, J.P., 2016. Pancreatic cancer. Nat Rev Dis Primers 2, 1–22. https://doi.org/10.1038/nrdp.2016.22

Konstantinidis, I.T., Warshaw, A.L., Allen, J.N., Blaszkowsky, L.S., Castillo, C.F., Deshpande, V., Hong, T.S., Kwak, E.L., Lauwers, G.Y., Ryan, D.P., Wargo, J.A., Lillemoe, K.D., Ferrone, C.R., 2013. Pancreatic Ductal Adenocarcinoma: Is There a Survival Difference for R1 Resections Versus Locally Advanced Unresectable Tumors? What Is a "True" R0 Resection? Annals of Surgery 257, 731. https://doi.org/10.1097/SLA.0b013e318263da2f

Kuziel, G.A., Rakoff-Nahoum, S., 2022. The gut microbiome. Curr Biol 32, R257–R264. https://doi.org/10.1016/j.cub.2022.02.023

Kyriazis, A.A., Kyriazis, A.P., Sternberg, C.N., Sloane, N.H., Loveless, J.D., 1986. Morphological, biological, biochemical, and karyotypic characteristics of human pancreatic ductal adenocarcinoma Capan-2 in tissue culture and the nude mouse. Cancer Res 46, 5810–5815.

Kyriazis, A.P., Kyriazis, A.A., Scarpelli, D.G., Fogh, J., Rao, M.S., Lepera, R., 1982. Human pancreatic adenocarcinoma line Capan-1 in tissue culture and the nude mouse: morphologic, biologic, and biochemical characteristics. Am J Pathol 106, 250–260.

Lacy, P.E., 1959. Electron microscopic and fluorescent antibody studies on islets of Langerhans. Experimental Cell Research, The Cytochemistry of Enzymes and Antigens 7, 296–308. https://doi.org/10.1016/0014-4827(59)90250-2

Langerhans, P., 1869. Beiträge zur mikroskopischen Anatomie der Bauchspeicheldrüse: Inaugural-Dissertation, zur Erlangung der Doctorwürde in der Medicine und Chirurgie vorgelegt der Medicinischen Facultät der Friedrich-Wilhelms-Universität zu Berlin und öffentlich zu vertheidigen am 18. Februar 1869. Buchdruckerei von Gustav Lange, Berlin.

Larsson, S.C., Wolk, A., 2012. Red and processed meat consumption and risk of pancreatic cancer: meta-analysis of prospective studies. Br J Cancer 106, 603–607. https://doi.org/10.1038/bjc.2011.585

Lee, J.W., Komar, C.A., Bengsch, F., Graham, K., Beatty, G.L., 2016. Genetically Engineered Mouse Models of Pancreatic Cancer: The KPC Model (*LSL-Kras* $^{G12D/+}$ *;LSL-Trp53* $^{R172H/+}$ *;Pdx-1-Cre*), Its Variants, and Their Application in Immuno-oncology Drug Discovery. CP Pharmacology 73. https://doi.org/10.1002/cpph.2

Leidenfrost, R.M., Pöther, D.-C., Jäckel, U., Wünschiers, R., 2020. Benchmarking the MinION: Evaluating long reads for microbial profiling. Sci Rep 10, 5125. https://doi.org/10.1038/s41598-020-61989-x

Li, Q., Jin, M., Liu, Y., Jin, L., 2020. Gut Microbiota: Its Potential Roles in Pancreatic Cancer. Front Cell Infect Microbiol 10, 572492. https://doi.org/10.3389/fcimb.2020.572492

Liang, C., Qin, Y., Zhang, B., Ji, S., Shi, S., Xu, W., Liu, J., Xiang, J., Liang, D., Hu, Q., Liu, L., Liu, C., Luo, G., Ni, Q., Xu, J., Yu, X., 2016. Energy sources identify metabolic phenotypes in pancreatic cancer. ABBS 969–979. https://doi.org/10.1093/abbs/gmw097

Ligat, L., Saint-Laurent, N., El-Mrani, A., Gigoux, V., Al Saati, T., Tomasini, R., Nigri, J., Dejean, S., Pont, F., Baer, R., Guillermet-Guibert, J., Cordelier, P., Lopez, F., Dufresne, M., 2015. Pancreatic preneoplastic lesions plasma signatures and biomarkers based on proteome profiling of mouse models. Br J Cancer 113, 1590–1598. https://doi.org/10.1038/bjc.2015.370

Lin, H., Peddada, S.D., 2020. Analysis of microbial compositions: a review of normalization and differential abundance analysis. npj Biofilms Microbiomes 6, 1–13. https://doi.org/10.1038/s41522-020-00160-w

Liu, F., Li, J., Guan, Y., Lou, Y., Chen, H., Xu, M., Deng, D., Chen, J., Ni, B., Zhao, L., Li, H., Sang, H., Cai, X., 2019. Dysbiosis of the Gut Microbiome is associated with Tumor Biomarkers in Lung Cancer. Int J Biol Sci 15, 2381–2392. https://doi.org/10.7150/ijbs.35980

Liu, J., Liu, C., Yue, J., 2021. Radiotherapy and the gut microbiome: facts and fiction. Radiation Oncology 16, 9. https://doi.org/10.1186/s13014-020-01735-9

Liu, P., Wang, Y., Yang, G., Zhang, Q., Meng, L., Xin, Y., Jiang, X., 2021. The role of short-chain fatty acids in intestinal barrier function, inflammation, oxidative stress, and colonic carcinogenesis. Pharmacological Research 165, 105420. https://doi.org/10.1016/j.phrs.2021.105420

Liu, X.-Y., Xue, L., Zheng, X., Yan, S., Zheng, S.-S., 2010. Pancreas transplantation in the mouse. Hepatobiliary Pancreat Dis Int 9, 254–258.

Loukopoulos, P., Kanetaka, K., Takamura, M., Shibata, T., Sakamoto, M., Hirohashi, S., 2004. Orthotopic transplantation models of pancreatic adenocarcinoma derived from cell lines and primary tumors and displaying varying metastatic activity. Pancreas 29, 193–203. https://doi.org/10.1097/00006676-200410000-00004

Lozupone, C., Knight, R., 2005. UniFrac: a New Phylogenetic Method for Comparing Microbial Communities. Applied and Environmental Microbiology 71, 8228–8235. https://doi.org/10.1128/AEM.71.12.8228-8235.2005

Lozupone, C.A., Hamady, M., Kelley, S.T., Knight, R., 2007. Quantitative and Qualitative β Diversity Measures Lead to Different Insights into Factors That Structure Microbial Communities. Applied and Environmental Microbiology 73, 1576–1585. https://doi.org/10.1128/AEM.01996-06

Luo, J., 2021. KRAS mutation in Pancreatic Cancer. Semin Oncol 48, 10–18. https://doi.org/10.1053/j.seminoncol.2021.02.003

Maisonneuve, P., Lowenfels, A.B., 2015. Risk factors for pancreatic cancer: a summary review of meta-analytical studies. International Journal of Epidemiology 44, 186–198. https://doi.org/10.1093/ije/dyu240

Mallya, K., Gautam, S.K., Aithal, A., Batra, S.K., Jain, M., 2021. Modeling pancreatic cancer in mice for experimental therapeutics. Biochim Biophys Acta Rev Cancer 1876, 188554. https://doi.org/10.1016/j.bbcan.2021.188554

Martincorena, I., Campbell, P.J., 2015. Somatic mutation in cancer and normal cells. Science 349, 1483–1489. https://doi.org/10.1126/science.aab4082

Martinson, J.N.V., Walk, S.T., 2020. Escherichia coli residency in the gut of healthy human adults. EcoSal Plus 9, 10.1128/ecosalplus.ESP-0003–2020. https://doi.org/10.1128/ecosalplus.ESP-0003-2020

Michaud, D.S., Izard, J., Wilhelm-Benartzi, C.S., You, D.-H., Grote, V.A., Tjønneland, A., Dahm, C.C., Overvad, K., Jenab, M., Fedirko, V., Boutron-Ruault, M.C., Clavel-Chapelon, F., Racine, A., Kaaks, R., Boeing, H., Foerster, J., Trichopoulou, A., Lagiou, P., Trichopoulos, D., Sacerdote, C., Sieri, S., Palli, D., Tumino, R., Panico, S., Siersema, P.D., Peeters, P.H., Lund, E., Barricarte, A., Huerta, J.-M., Molina-Montes, E., Dorronsoro, M., Quirós, J.R., Duell, E.J., Ye, W., Sund, M., Lindkvist, B., Johansen, D., Khaw, K.-T., Wareham, N., Travis, R.C., Vineis, P., Bueno-de-Mesquita, H.B., Riboli, E., 2013. Plasma antibodies to oral bacteria and risk of pancreatic cancer in a large

European prospective cohort study. Gut 62, 1764–1770. https://doi.org/10.1136/gutjnl-2012-303006

Michaud, D.S., Joshipura, K., Giovannucci, E., Fuchs, C.S., 2007. A Prospective Study of Periodontal Disease and Pancreatic Cancer in US Male Health Professionals. JNCI: Journal of the National Cancer Institute 99, 171–175. https://doi.org/10.1093/jnci/djk021

Milan, M., Diaferia, G.R., Natoli, G., 2021. Tumor cell heterogeneity and its transcriptional bases in pancreatic cancer: a tale of two cell types and their many variants. The EMBO Journal 40, e107206. https://doi.org/10.15252/embj.2020107206

Moffitt, R.A., Marayati, R., Flate, E.L., Volmar, K.E., Loeza, S.G.H., Hoadley, K.A., Rashid, N.U., Williams, L.A., Eaton, S.C., Chung, A.H., Smyla, J.K., Anderson, J.M., Kim, H.J., Bentrem, D.J., Talamonti, M.S., Iacobuzio-Donahue, C.A., Hollingsworth, M.A., Yeh, J.J., 2015. Virtual microdissection identifies distinct tumor- and stroma-specific subtypes of pancreatic ductal adenocarcinoma. Nat Genet 47, 1168–1178. https://doi.org/10.1038/ng.3398

Mohajeri, M.H., La Fata, G., Steinert, R.E., Weber, P., 2018. Relationship between the gut microbiome and brain function. Nutr Rev 76, 481–496. https://doi.org/10.1093/nutrit/nuy009

Morton, J.P., Timpson, P., Karim, S.A., Ridgway, R.A., Athineos, D., Doyle, B., Jamieson, N.B., Oien, K.A., Lowy, A.M., Brunton, V.G., Frame, M.C., Evans, T.R.J., Sansom, O.J., 2010. Mutant p53 drives metastasis and overcomes growth arrest/senescence in pancreatic cancer. Proceedings of the National Academy of Sciences 107, 246–251. https://doi.org/10.1073/pnas.0908428107

Müller, T.D., Nogueiras, R., Andermann, M.L., Andrews, Z.B., Anker, S.D., Argente, J., Batterham, R.L., Benoit, S.C., Bowers, C.Y., Broglio, F., Casanueva, F.F., D'Alessio, D., Depoortere, I., Geliebter, A., Ghigo, E., Cole, P.A., Cowley, M., Cummings, D.E., Dagher, A., Diano, S., Dickson, S.L., Diéguez, C., Granata, R., Grill, H.J., Grove, K., Habegger, K.M., Heppner, K., Heiman, M.L., Holsen, L., Holst, B., Inui, A., Jansson, J.O., Kirchner, H., Korbonits, M., Laferrère, B., LeRoux, C.W., Lopez, M., Morin, S., Nakazato, M., Nass, R., Perez-Tilve, D., Pfluger, P.T., Schwartz, T.W., Seeley, R.J., Sleeman, M., Sun, Y., Sussel, L., Tong, J., Thorner, M.O., van der Lely, A.J., van der Ploeg, L.H.T., Zigman, J.M., Kojima, M., Kangawa, K., Smith, R.G., Horvath, T., Tschöp, M.H., 2015. Ghrelin. Mol Metab 4, 437–460. https://doi.org/10.1016/j.molmet.2015.03.005

Nagata, N., Nishijima, S., Kojima, Y., Hisada, Y., Imbe, K., Miyoshi-Akiyama, T., Suda, W., Kimura, M., Aoki, R., Sekine, K., Ohsugi, M., Miki, K., Osawa, T., Ueki, K., Oka, S., Mizokami, M., Kartal, E., Schmidt, T.S.B., Molina-Montes, E., Estudillo, L., Malats, N., Trebicka, J., Kersting, S., Langheinrich, M., Bork, P., Uemura, N., Itoi, T., Kawai, T., 2022. Metagenomic Identification of Microbial Signatures Predicting Pancreatic Cancer From a Multinational Study. Gastroenterology 163, 222–238. https://doi.org/10.1053/j.gastro.2022.03.054

Neesse, A., Algül, H., Tuveson, D.A., Gress, T.M., 2015. Stromal biology and therapy in pancreatic cancer: a changing paradigm. Gut 64, 1476–1484. https://doi.org/10.1136/gutjnl-2015-309304

Neesse, A., Bauer, C.A., Öhlund, D., Lauth, M., Buchholz, M., Michl, P., Tuveson, D.A., Gress, T.M., 2019. Stromal biology and therapy in pancreatic cancer: ready for clinical translation? Gut 68, 159–171. https://doi.org/10.1136/gutjnl-2018-316451

Neesse, A., Michl, P., Frese, K.K., Feig, C., Cook, N., Jacobetz, M.A., Lolkema, M.P., Buchholz, M., Olive, K.P., Gress, T.M., Tuveson, D.A., 2011. Stromal biology and therapy in pancreatic cancer. Gut 60, 861–868. https://doi.org/10.1136/gut.2010.226092

Nejman, D., Livyatan, I., Fuks, G., Gavert, N., Zwang, Y., Geller, L.T., Rotter-Maskowitz, A., Weiser, R., Mallel, G., Gigi, E., Meltser, A., Douglas, G.M., Kamer, I., Gopalakrishnan, V., Dadosh, T., Levin-Zaidman, S., Avnet, S., Atlan, T., Cooper, Z.A., Arora, R., Cogdill, A.P., Khan, M.A.W., Ologun, G., Bussi, Y., Weinberger, A., Lotan-Pompan, M., Golani, O., Perry, G., Rokah, M., Bahar-Shany, K., Rozeman, E.A., Blank, C.U., Ronai, A., Shaoul, R., Amit, A., Dorfman, T., Kremer, R., Cohen, Z.R., Harnof, S., Siegal, T., Yehuda-Shnaidman, E., Gal-Yam, E.N., Shapira, H., Baldini, N., Langille, M.G.I., Ben-Nun, A., Kaufman, B., Nissan, A., Golan, T., Dadiani, M., Levanon, K., Bar, J., Yust-Katz, S., Barshack, I., Peeper, D.S., Raz, D.J., Segal, E., Wargo, J.A., Sandbank, J., Shental, N., Straussman, R., 2020. The human tumor microbiome is composed of tumor type–specific intracellular bacteria. Science 368, 973–980. https://doi.org/10.1126/science.aay9189

Ochi, A., Nguyen, A.H., Bedrosian, A.S., Mushlin, H.M., Zarbakhsh, S., Barilla, R., Zambirinis, C.P., Fallon, N.C., Rehman, A., Pylayeva-Gupta, Y., Badar, S., Hajdu, C.H., Frey, A.B., Bar-Sagi, D., Miller, G., 2012. MyD88 inhibition amplifies dendritic cell capacity to promote pancreatic carcinogenesis via Th2 cells. Journal of Experimental Medicine 209, 1671–1687. https://doi.org/10.1084/jem.20111706

Offield, M.F., Jetton, T.L., Labosky, P.A., Ray, M., Stein, R.W., Magnuson, M.A., Hogan, B.L., Wright, C.V., 1996. PDX-1 is required for pancreatic outgrowth and differentiation of the rostral duodenum. Development 122, 983–995. https://doi.org/10.1242/dev.122.3.983

O'Kane, G.M., Grünwald, B.T., Jang, G.-H., Masoomian, M., Picardo, S., Grant, R.C., Denroche, R.E., Zhang, A., Wang, Y., Miller, J.K., Lam, B., Krzyzanowski, P.M., Lungu, I.M., Bartlett, J.M.S., Peralta, M., Vyas, F., Khokha, R., Biagi, J., Chadwick, D., Ramotar, S., Hutchinson, S., Dodd, A., Wilson, J.M., Notta, F., Zogopoulos, G., Gallinger, S., Knox, J.J., Fischer, S.E., 2020. GATA6 Expression Distinguishes Classical and Basal-like Subtypes in Advanced Pancreatic Cancer. Clinical Cancer Research 26, 4901–4910. https://doi.org/10.1158/1078-0432.CCR-19-3724

Olson, S.H., Satagopan, J., Xu, Y., Ling, L., Leong, S., Orlow, I., Saldia, A., Li, P., Nunes, P., Madonia, V., Allen, P.J., O'Reilly, E., Pamer, E., Kurtz, R.C., 2017. The oral microbiota in patients with pancreatic cancer, patients with IPMNs, and controls: a pilot study. Cancer Causes Control 28, 959–969. https://doi.org/10.1007/s10552-017-0933-8

O'Reilly, E.M., Lee, J.W., Zalupski, M., Capanu, M., Park, J., Golan, T., Tahover, E., Lowery, M.A., Chou, J.F., Sahai, V., Brenner, R., Kindler, H.L., Yu, K.H., Zervoudakis, A., Vemuri, S., Stadler, Z.K., Do, R.K.G., Dhani, N., Chen, A.P., Kelsen, D.P., 2020. Randomized, Multicenter, Phase II Trial of Gemcitabine and Cisplatin With or Without Veliparib in Patients With Pancreas Adenocarcinoma and a Germline BRCA/PALB2 Mutation. JCO 38, 1378–1388. https://doi.org/10.1200/JCO.19.02931

Orr, S., Huang, L., Moser, J., Stroopinsky, D., Gandarilla, O., DeCicco, C., Liegel, J., Tacettin, C., Ephraim, A., Cheloni, G., Torres, D., Kufe, D., Rosenblatt, J., Hidalgo, M., Muthuswamy, S.K., Avigan, D., 2023. Personalized tumor vaccine for pancreatic cancer. Cancer Immunol Immunother 72, 301–313. https://doi.org/10.1007/s00262-022-03237-x

Ott, P.A., Bang, Y.-J., Piha-Paul, S.A., Razak, A.R.A., Bennouna, J., Soria, J.-C., Rugo, H.S., Cohen, R.B., O'Neil, B.H., Mehnert, J.M., Lopez, J., Doi, T., van Brummelen, E.M.J.,

Cristescu, R., Yang, P., Emancipator, K., Stein, K., Ayers, M., Joe, A.K., Lunceford, J.K., 2019. T-Cell–Inflamed Gene-Expression Profile, Programmed Death Ligand 1 Expression, and Tumor Mutational Burden Predict Efficacy in Patients Treated With Pembrolizumab Across 20 Cancers: KEYNOTE-028. JCO 37, 318–327. https://doi.org/10.1200/JCO.2018.78.2276

Panebianco, C., Villani, A., Pisati, F., Orsenigo, F., Ulaszewska, M., Latiano, T.P., Potenza, A., Andolfo, A., Terracciano, F., Tripodo, C., Perri, F., Pazienza, V., 2022. Butyrate, a postbiotic of intestinal bacteria, affects pancreatic cancer and gemcitabine response in in vitro and in vivo models. Biomedicine & Pharmacotherapy 151, 113163. https://doi.org/10.1016/j.biopha.2022.113163

Park, W., Chawla, A., O'Reilly, E.M., 2021. Pancreatic Cancer: A Review. JAMA 326, 851–862. https://doi.org/10.1001/jama.2021.13027

Payne, A., Holmes, N., Rakyan, V., Loose, M., 2019. BulkVis: a graphical viewer for Oxford nanopore bulk FAST5 files. Bioinformatics 35, 2193–2198. https://doi.org/10.1093/bioinformatics/bty841

Peng, J., Sun, B.-F., Chen, C.-Y., Zhou, J.-Y., Chen, Y.-S., Chen, H., Liu, L., Huang, D., Jiang, J., Cui, G.-S., Yang, Y., Wang, W., Guo, D., Dai, M., Guo, J., Zhang, T., Liao, Q., Liu, Y., Zhao, Y.-L., Han, D.-L., Zhao, Y., Yang, Y.-G., Wu, W., 2019. Single-cell RNA-seq highlights intra-tumoral heterogeneity and malignant progression in pancreatic ductal adenocarcinoma. Cell Res 29, 725–738. https://doi.org/10.1038/s41422-019-0195-y

Pereira, B.A., Vennin, C., Papanicolaou, M., Chambers, C.R., Herrmann, D., Morton, J.P., Cox, T.R., Timpson, P., 2019. CAF Subpopulations: A New Reservoir of Stromal Targets in Pancreatic Cancer. Trends in Cancer 5, 724–741. https://doi.org/10.1016/j.trecan.2019.09.010

Pfisterer, N., Lingens, C., Heuer, C., Dang, L., Neesse, A., Ammer-Herrmenau, C., 2022. The Microbiome in PDAC—Vantage Point for Future Therapies? Cancers 14, 5974. https://doi.org/10.3390/cancers14235974

Poore, G.D., Kopylova, E., Zhu, Q., Carpenter, C., Fraraccio, S., Wandro, S., Kosciolek, T., Janssen, S., Metcalf, J., Song, S.J., Kanbar, J., Miller-Montgomery, S., Heaton, R., Mckay, R., Patel, S.P., Swafford, A.D., Knight, R., 2020. Microbiome analyses of blood and tissues suggest cancer diagnostic approach. Nature 579, 567–574. https://doi.org/10.1038/s41586-020-2095-1

Prudent, E., Raoult, D., 2019. Fluorescence in situ hybridization, a complementary molecular tool for the clinical diagnosis of infectious diseases by intracellular and fastidious bacteria. FEMS Microbiology Reviews 43, 88–107. https://doi.org/10.1093/femsre/fuy040

Puleo, F., Nicolle, R., Blum, Y., Cros, J., Marisa, L., Demetter, P., Quertinmont, E., Svrcek, M., Elarouci, N., Iovanna, J., Franchimont, D., Verset, L., Galdon, M.G., Devière, J., Reyniès, A. de, Laurent-Puig, P., Laethem, J.-L.V., Bachet, J.-B., Maréchal, R., 2018. Stratification of Pancreatic Ductal Adenocarcinomas Based on Tumor and Microenvironment Features. Gastroenterology 155, 1999-2013.e3. https://doi.org/10.1053/j.gastro.2018.08.033

Pushalkar, S., Hundeyin, M., Daley, D., Zambirinis, C.P., Kurz, E., Mishra, A., Mohan, N., Aykut, B., Usyk, M., Torres, L.E., Werba, G., Zhang, K., Guo, Y., Li, Q., Akkad, N., Lall, S., Wadowski, B., Gutierrez, J., Rossi, J.A.K., Herzog, J.W., Diskin, B., Torres-Hernandez, A., Leinwand, J., Wang, W., Taunk, P.S., Savadkar, S., Janal, M., Saxena, A., Li, X., Cohen, D., Sartor, R.B., Saxena, D., Miller, G., 2018. The Pancreatic Cancer

Microbiome Promotes Oncogenesis by Induction of Innate and Adaptive Immune Suppression. Cancer Discov 8, 403–416. https://doi.org/10.1158/2159-8290.CD-17-1134

Qiu, W., Su, G.H., 2013a. Development of Orthotopic Pancreatic Tumor Mouse Models. Methods Mol Biol 980, 215–223. https://doi.org/10.1007/978-1-62703-287-2_11

Qiu, W., Su, G.H., 2013b. Challenges and advances in mouse modeling for human pancreatic tumorigenesis and metastasis. Cancer Metastasis Rev 32, 10.1007/s10555-012-9408–2. https://doi.org/10.1007/s10555-012-9408-2

Quigley, E.M.M., 2013. Gut bacteria in health and disease. Gastroenterol Hepatol (N Y) 9, 560–569.

Rahib, L., Smith, B.D., Aizenberg, R., Rosenzweig, A.B., Fleshman, J.M., Matrisian, L.M., 2014. Projecting Cancer Incidence and Deaths to 2030: The Unexpected Burden of Thyroid, Liver, and Pancreas Cancers in the United States. Cancer Research 74, 2913–2921. https://doi.org/10.1158/0008-5472.CAN-14-0155

Rampelotto, P.H., Sereia, A.F.R., de Oliveira, L.F.V., Margis, R., 2019. Exploring the Hospital Microbiome by High-Resolution 16S rRNA Profiling. International Journal of Molecular Sciences 20, 3099. https://doi.org/10.3390/ijms20123099

Rawla, P., Sunkara, T., Gaduputi, V., 2019. Epidemiology of Pancreatic Cancer: Global Trends, Etiology and Risk Factors. World J Oncol 10, 10–27. https://doi.org/10.4021/wjon.v10i1.1166

Ray, P., Rialon-Guevara, K.L., Veras, E., Sullenger, B.A., White, R.R., 2012. Comparing human pancreatic cell secretomes by in vitro aptamer selection identifies cyclophilin B as a candidate pancreatic cancer biomarker. J Clin Invest 122, 1734–1741. https://doi.org/10.1172/JCI62385

Ren, Z., Jiang, J., Xie, H., Li, A., Lu, H., Xu, S., Zhou, L., Zhang, H., Cui, G., Chen, X., Liu, Y., Wu, L., Qin, N., Sun, R., Wang, Wei, Li, L., Wang, Weilin, Zheng, S., 2017. Gut microbial profile analysis by MiSeq sequencing of pancreatic carcinoma patients in China. Oncotarget 8, 95176–95191. https://doi.org/10.18632/oncotarget.18820

Riquelme, E., Zhang, Y., Zhang, L., Montiel, M., Zoltan, M., Dong, W., Quesada, P., Sahin, I., Chandra, V., San Lucas, A., Scheet, P., Xu, H., Hanash, S.M., Feng, L., Burks, J.K., Do, K.-A., Peterson, C.B., Nejman, D., Tzeng, C.-W.D., Kim, M.P., Sears, C.L., Ajami, N., Petrosino, J., Wood, L.D., Maitra, A., Straussman, R., Katz, M., White, J.R., Jenq, R., Wargo, J., McAllister, F., 2019. Tumor Microbiome Diversity and Composition Influence Pancreatic Cancer Outcomes. Cell 178, 795-806.e12. https://doi.org/10.1016/j.cell.2019.07.008

Rodriguez-Arrastia, M., Martinez-Ortigosa, A., Rueda-Ruzafa, L., Folch Ayora, A., Ropero-Padilla, C., 2021. Probiotic Supplements on Oncology Patients' Treatment-Related Side Effects: A Systematic Review of Randomized Controlled Trials. International Journal of Environmental Research and Public Health 18, 4265. https://doi.org/10.3390/ijerph18084265

Routy, B., Le Chatelier, E., Derosa, L., Duong, C.P.M., Alou, M.T., Daillère, R., Fluckiger, A., Messaoudene, M., Rauber, C., Roberti, M.P., Fidelle, M., Flament, C., Poirier-Colame, V., Opolon, P., Klein, C., Iribarren, K., Mondragón, L., Jacquelot, N., Qu, B., Ferrere, G., Clémenson, C., Mezquita, L., Masip, J.R., Naltet, C., Brosseau, S., Kaderbhai, C., Richard, C., Rizvi, H., Levenez, F., Galleron, N., Quinquis, B., Pons, N., Ryffel, B., Minard-Colin, V., Gonin, P., Soria, J.-C., Deutsch, E., Loriot, Y., Ghiringhelli, F., Zalcman, G., Goldwasser, F., Escudier, B., Hellmann, M.D., Eggermont, A., Raoult, D., Albiges, L., Kroemer, G., Zitvogel, L., 2018. Gut microbiome influences efficacy of

PD-1–based immunotherapy against epithelial tumors. Science 359, 91–97. https://doi.org/10.1126/science.aan3706

Royal, R.E., Levy, C., Turner, K., Mathur, A., Hughes, M., Kammula, U.S., Sherry, R.M., Topalian, S.L., Yang, J.C., Lowy, I., Rosenberg, S.A., 2010. Phase 2 Trial of Single Agent Ipilimumab (Anti-CTLA-4) for Locally Advanced or Metastatic Pancreatic Adenocarcinoma. Journal of Immunotherapy 33, 828. https://doi.org/10.1097/CJI.0b013e3181eec14c

Ryan, D.P., Hong, T.S., Bardeesy, N., 2014. Pancreatic Adenocarcinoma. N Engl J Med 371, 1039–1049. https://doi.org/10.1056/NEJMra1404198

Sagué, C.M.B., Jarvis, W.R., National Nosocomial Infections Surveillance System, 1993. Secular Trends in the Epidemiology of Nosocomial Fungal Infections in the United States, 1980–1990. The Journal of Infectious Diseases 167, 1247–1251. https://doi.org/10.1093/infdis/167.5.1247

Salter, S.J., Cox, M.J., Turek, E.M., Calus, S.T., Cookson, W.O., Moffatt, M.F., Turner, P., Parkhill, J., Loman, N.J., Walker, A.W., 2014. Reagent and laboratory contamination can critically impact sequence-based microbiome analyses. BMC Biology 12, 87. https://doi.org/10.1186/s12915-014-0087-z

Sanderson, N.D., Street, T.L., Foster, D., Swann, J., Atkins, B.L., Brent, A.J., McNally, M.A., Oakley, S., Taylor, A., Peto, T.E.A., Crook, D.W., Eyre, D.W., 2018. Real-time analysis of nanopore-based metagenomic sequencing from infected orthopaedic devices. BMC Genomics 19, 714. https://doi.org/10.1186/s12864-018-5094-y

Scholten, J., 2020. Pancreatic Cancer Hits Milestone with Double-Digit Survival Rates. Pancreatic Cancer Action Network. URL https://pancan.org/press-releases/pancreatic-cancer-hits-milestone-with-double-digit-survival-rates/ (accessed 3.21.23).

Schwabe, R.F., Jobin, C., 2013. The microbiome and cancer. Nat Rev Cancer 13, 800–812. https://doi.org/10.1038/nrc3610

Sears, C.L., 2005. A dynamic partnership: Celebrating our gut flora. Anaerobe 11, 247–251. https://doi.org/10.1016/j.anaerobe.2005.05.001

Shapira, M., 2016. Gut Microbiotas and Host Evolution: Scaling Up Symbiosis. Trends in Ecology & Evolution 31, 539–549. https://doi.org/10.1016/j.tree.2016.03.006

Shi, N., Li, N., Duan, X., Niu, H., 2017. Interaction between the gut microbiome and mucosal immune system. Mil Med Res 4, 14. https://doi.org/10.1186/s40779-017-0122-9

Sidhu, M., van der Poorten, D., 2017. The gut microbiome. Aust Fam Physician 46, 206–211.

Siegel, R.L., Miller, K.D., Fuchs, H.E., Jemal, A., 2021. Cancer Statistics, 2021. CA: A Cancer Journal for Clinicians 71, 7–33. https://doi.org/10.3322/caac.21654

Silva-García, O., Valdez-Alarcón, J.J., Baizabal-Aguirre, V.M., 2019. Wnt/β-Catenin Signaling as a Molecular Target by Pathogenic Bacteria. Frontiers in Immunology 10.

Simpson, E.H., 1949. Measurement of Diversity. Nature 163, 688–688. https://doi.org/10.1038/163688a0

Sivan, A., Corrales, L., Hubert, N., Williams, J.B., Aquino-Michaels, K., Earley, Z.M., Benyamin, F.W., Man Lei, Y., Jabri, B., Alegre, M.-L., Chang, E.B., Gajewski, T.F., 2015. Commensal Bifidobacterium promotes antitumor immunity and facilitates anti–PD-L1 efficacy. Science 350, 1084–1089. https://doi.org/10.1126/science.aac4255

Smirnova, E., Huzurbazar, S., Jafari, F., 2019. PERFect: PERmutation Filtering test for microbiome data. Biostatistics 20, 615–631. https://doi.org/10.1093/biostatistics/kxy020

Söling, H.D., Unger, K.O., 1972. The Role of Insulin in the Regulation of α-Amylase Synthesis in the Rat Pancreas. European Journal of Clinical Investigation 2, 199–212. https://doi.org/10.1111/j.1365-2362.1972.tb00645.x

Sonbol, M.B., Mazza, G.L., Mi, L., Oliver, T., Starr, J., Gudmundsdottir, H., Cleary, S.P., Hobday, T., Halfdanarson, T.R., 2022. Survival and Incidence Patterns of Pancreatic Neuroendocrine Tumors Over the Last 2 Decades: A SEER Database Analysis. The Oncologist 27, 573–578. https://doi.org/10.1093/oncolo/oyac049

Standring, S. (Ed.), 2016. Gray's anatomy: the anatomical basis of clinical practice, Forty-first edition. ed. Elsevier Limited, Philadelphia.

Stasiewicz, M., Kwaśniewski, M., Karpiński, T.M., 2021. Microbial Associations with Pancreatic Cancer: A New Frontier in Biomarkers. Cancers 13, 3784. https://doi.org/10.3390/cancers13153784

Stephen, A.M., Cummings, J.H., 1980. THE MICROBIAL CONTRIBUTION TO HUMAN FAECAL MASS. Journal of Medical Microbiology 13, 45–56. https://doi.org/10.1099/00222615-13-1-45

Stott, M.C., Oldfield, L., Hale, J., Costello, E., Halloran, C.M., 2022. Recent advances in understanding pancreatic cancer. Fac Rev 11, 9. https://doi.org/10.12703/r/11-9

Sung, H., Ferlay, J., Siegel, R.L., Laversanne, M., Soerjomataram, I., Jemal, A., Bray, F., 2021. Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. CA: A Cancer Journal for Clinicians 71, 209–249. https://doi.org/10.3322/caac.21660

Tanoue, T., Morita, S., Plichta, D.R., Skelly, A.N., Suda, W., Sugiura, Y., Narushima, S., Vlamakis, H., Motoo, I., Sugita, K., Shiota, A., Takeshita, K., Yasuma-Mitobe, K., Riethmacher, D., Kaisho, T., Norman, J.M., Mucida, D., Suematsu, M., Yaguchi, T., Bucci, V., Inoue, T., Kawakami, Y., Olle, B., Roberts, B., Hattori, M., Xavier, R.J., Atarashi, K., Honda, K., 2019. A defined commensal consortium elicits CD8 T cells and anti-cancer immunity. Nature 565, 600–605. https://doi.org/10.1038/s41586-019-0878-z

Teng, F., Darveekaran Nair, S.S., Zhu, P., Li, S., Huang, S., Li, X., Xu, J., Yang, F., 2018. Impact of DNA extraction method and targeted 16S-rRNA hypervariable region on oral microbiota profiling. Sci Rep 8, 16321. https://doi.org/10.1038/s41598-018-34294-x

Thomas, R.M., Gharaibeh, R.Z., Gauthier, J., Beveridge, M., Pope, J.L., Guijarro, M.V., Yu, Q., He, Z., Ohland, C., Newsome, R., Trevino, J., Hughes, S.J., Reinhard, M., Winglee, K., Fodor, A.A., Zajac-Kaye, M., Jobin, C., 2018. Intestinal microbiota enhances pancreatic carcinogenesis in preclinical models. Carcinogenesis 39, 1068–1078. https://doi.org/10.1093/carcin/bgy073

Topham, J.T., Karasinska, J.M., Lee, M.K.C., Csizmok, V., Williamson, L.M., Jang, G.H., Denroche, R.E., Tsang, E.S., Kalloger, S.E., Wong, H., O'Kane, G.M., Moore, R.A., Mungall, A.J., Notta, F., Loree, J.M., Wilson, J.M., Bathe, O., Tang, P.A., Goodwin, R., Knox, J.J., Gallinger, S., Laskin, J., Marra, M.A., Jones, S.J.M., Renouf, D.J., Schaeffer, D.F., 2021. Subtype-Discordant Pancreatic Ductal Adenocarcinoma Tumors Show Intermediate Clinical and Molecular Characteristics. Clinical Cancer Research 27, 150–157. https://doi.org/10.1158/1078-0432.CCR-20-2831

Tu, M., Klein, L., Espinet, E., Georgomanolis, T., Wegwitz, F., Li, X., Urbach, L., Danieli-Mackay, A., Küffer, S., Bojarczuk, K., Mizi, A., Günesdogan, U., Chapuy, B., Gu, Z., Neesse, A., Kishore, U., Ströbel, P., Hessmann, E., Hahn, S.A., Trumpp, A., Papantonis, A., Ellenrieder, V., Singh, S.K., 2021. TNF-α-producing macrophages determine subtype identity and prognosis via AP1 enhancer reprogramming in pancreatic cancer. Nat Cancer 2, 1185–1203. https://doi.org/10.1038/s43018-021-00258-w

Van Ameringen, M., Turna, J., Patterson, B., Pipe, A., Mao, R.Q., Anglin, R., Surette, M.G., 2019. The gut microbiome in psychiatry: A primer for clinicians. Depress Anxiety 36, 1004–1025. https://doi.org/10.1002/da.22936

Vétizou, M., Pitt, J.M., Daillère, R., Lepage, P., Waldschmitt, N., Flament, C., Rusakiewicz, S., Routy, B., Roberti, M.P., Duong, C.P.M., Poirier-Colame, V., Roux, A., Becharef, S., Formenti, S., Golden, E., Cording, S., Eberl, G., Schlitzer, A., Ginhoux, F., Mani, S., Yamazaki, T., Jacquelot, N., Enot, D.P., Bérard, M., Nigou, J., Opolon, P., Eggermont, A., Woerther, P.-L., Chachaty, E., Chaput, N., Robert, C., Mateus, C., Kroemer, G., Raoult, D., Boneca, I.G., Carbonnel, F., Chamaillard, M., Zitvogel, L., 2015. Anticancer immunotherapy by CTLA-4 blockade relies on the gut microbiota. Science 350, 1079–1084. https://doi.org/10.1126/science.aad1329

Voigt, R.M., Forsyth, C.B., Green, S.J., Engen, P.A., Keshavarzian, A., 2016. Circadian Rhythm and the Gut Microbiome. Int Rev Neurobiol 131, 193–205. https://doi.org/10.1016/bs.irn.2016.07.002

Von Hoff, D.D., Ervin, T., Arena, F.P., Chiorean, E.G., Infante, J., Moore, M., Seay, T., Tjulandin, S.A., Ma, W.W., Saleh, M.N., Harris, M., Reni, M., Dowden, S., Laheru, D., Bahary, N., Ramanathan, R.K., Tabernero, J., Hidalgo, M., Goldstein, D., Van Cutsem, E., Wei, X., Iglesias, J., Renschler, M.F., 2013. Increased Survival in Pancreatic Cancer with nab-Paclitaxel plus Gemcitabine. N Engl J Med 369, 1691–1703. https://doi.org/10.1056/NEJMoa1304369

Waks, A.G., Winer, E.P., 2019. Breast Cancer Treatment: A Review. JAMA 321, 288–300. https://doi.org/10.1001/jama.2018.19323

Wallner, G., Amann, R., Beisker, W., 1993. Optimizing fluorescent in situ hybridization with rRNA-targeted oligonucleotide probes for flow cytometric identification of microorganisms. Cytometry 14, 136–143. https://doi.org/10.1002/cyto.990140205

Wang, L., Xie, D., Wei, D., 2019. Pancreatic Acinar-to-Ductal Metaplasia and Pancreatic Cancer, in: Su, G.H. (Ed.), Pancreatic Cancer: Methods and Protocols, Methods in Molecular Biology. Springer, New York, NY, pp. 299–308. https://doi.org/10.1007/978-1-4939-8879-2_26

Wang, Y.-T., Gou, Y.-W., Jin, W.-W., Xiao, M., Fang, H.-Y., 2016. Association between alcohol intake and the risk of pancreatic cancer: a dose–response meta-analysis of cohort studies. BMC Cancer 16, 212. https://doi.org/10.1186/s12885-016-2241-1

Weersma, R.K., Zhernakova, A., Fu, J., 2020. Interaction between drugs and the gut microbiome. Gut 69, 1510–1519. https://doi.org/10.1136/gutjnl-2019-320204

Weiss, S., Xu, Z.Z., Peddada, S., Amir, A., Bittinger, K., Gonzalez, A., Lozupone, C., Zaneveld, J.R., Vázquez-Baeza, Y., Birmingham, A., Hyde, E.R., Knight, R., 2017. Normalization and microbial differential abundance strategies depend upon data characteristics. Microbiome 5, 27. https://doi.org/10.1186/s40168-017-0237-y

Westphalen, C.B., Olive, K.P., 2012. Genetically Engineered Mouse Models of Pancreatic Cancer. The Cancer Journal 18, 502–510. https://doi.org/10.1097/PPO.0b013e31827ab4c4

Whittaker, R.H., 1972. Evolution and Measurement of Species Diversity. TAXON 21, 213–251. https://doi.org/10.2307/1218190

Whittaker, R.H., 1960. Vegetation of the Siskiyou Mountains, Oregon and California. Ecological Monographs 30, 279–338. https://doi.org/10.2307/1943563

Yachida, S., Iacobuzio-Donahue, C.A., 2009. The Pathology and Genetics of Metastatic Pancreatic Cancer. Archives of Pathology & Laboratory Medicine 133, 413–422. https://doi.org/10.5858/133.3.413

Young, B., O'Dowd, G., Woodford, P., 2014. Wheater's functional histology: a text and colour atlas, Sixth edition. ed, ClinicalKey. Churchill Livingston/Elsevier, Philadelphia, PA.

Yunis, A.A., Arimura, G.K., Russin, D.J., 1977. Human pancreatic carcinoma (MIA PaCa-2) in continuous culture: sensitivity to asparaginase. Int J Cancer 19, 128–135. https://doi.org/10.1002/ijc.2910190118

Zhang, Xiaozhen, Lao, M., Xu, J., Duan, Y., Yang, H., Li, M., Ying, H., He, L., Sun, K., Guo, C., Chen, W., Jiang, H., Zhang, Xiaoyu, Bai, X., Liang, T., 2022. Combination cancer immunotherapy targeting TNFR2 and PD-1/PD-L1 signaling reduces immunosuppressive effects in the microenvironment of pancreatic tumors. J Immunother Cancer 10, e003982. https://doi.org/10.1136/jitc-2021-003982

Zheng, D.-W., Dong, X., Pan, P., Chen, K.-W., Fan, J.-X., Cheng, S.-X., Zhang, X.-Z., 2019. Phage-guided modulation of the gut microbiota of mouse models of colorectal cancer augments their responses to chemotherapy. Nat Biomed Eng 3, 717–728. https://doi.org/10.1038/s41551-019-0423-2

Zhou, Q., Melton, D.A., 2018. Pancreas regeneration. Nature 557, 351–358. https://doi.org/10.1038/s41586-018-0088-0

ACKNOWLEDGEMENTS

So, this is it. I have always kept this last bit as a treat to myself for the very end. During the last couple of months, I came back to it whenever I needed a glimpse of this feeling of being done. And now that I am finally allowed to enjoy this moment, I ask myself: Is it really over? Like ever? I feel like a PhD is just the entry ticket to the unlimited world of science. It means you are a certified scientist, although everybody can be a scientist. Just be curious. Never restless. Always alert. And for this reason, I'd like to start this final part with a disclaimer (because this is what we do in science papers and presentations) – I apologize for unintentional incompleteness of the following. However, I can't have the acknowledgments exceed the page number of the actual thesis. (Although this would be the much easier part.) And please do not pin me down on the order of appearance. (This is worse than setting up the perfect seating arrangement at a wedding.) Alright, here we go.

Prof. Dr. Dr. Albrecht Neeße. Thank you for being willing to serve as my doctoral supervisor. Thank you for trusting my work in progress until the very end. Thank you for teaching me independence and soft skills such as writing and presenting, and for unleashing many abilities that may only be acquired during a PhD. Thank you for showing your support in helping provide the Werner-Creutzfeldt scholarship. Thank you for always ensuring enough funding for research and salary. And finally, thank you for speeding up things in the ER after my bike accident.

Prof. Dr. Matthias Dobbelstein and Prof. Dr. Argyris Papantonis. Thank you for joining my Thesis Advisory Committee and for accompanying me on my PhD journey. Thank you for always speaking true words and for making me feel comfortable in situations that meant a lot of mental load to me. In particular, Prof. Dobbelstein, thank you for taking my interest in forensic sciences seriously.

Prof. Dr. Michael Zeisberg, Prof. Dr. mult. Thomas Meyer, and Prof. Dr. Ralf Dressel. Thank you for spontaneously being willing to complete my examination board. Thank you for your kind responses that lit up my daily writing routine, and for making this organizational process enjoyably uncomplicated.

Christoph. Thank you for joining the crew and completing the lab structures. Every decent research team needs a dedicated postdoc. Thank you for providing the base to my project and for all the dry-lab work you accomplished. Thank you for pushing me although it was not always comfortable but comfortable is not what a PhD is supposed to feel like I guess. Thank you for teaching me the term "dekompensieren" so I could describe my mental breakdowns in a sophisticated way.

Prof. Dr. Philipp Ströbel, Dr. Hanibal Bohnenberger and your team. Thank you for providing the FFPE patient samples and sharing your expertise about DNA extraction from FFPE tissue. Dr. Stefan Küffer, Juliane Schnalke and Laura Lukat. Thank you for teaching me FISH and providing the resources. Christoph Petrynowski. Thank you for your support with the

cleanup of sequencing data. All of these cooperations are very valuable to foster research exchange and scientific progress.

Jutta and Ulrike. The two of you could not be any more different. Maybe that's why you are such a well-fitting team. Always supportive, always complementing each other. Thank you for welcoming me as a part of this team. Thank you for teaching me lab skills and life lessons. Thank you for building me up in my darkest moments when tears hit the lab floor. You are my lab moms for a good reason. In many situations, your support went beyond work.

Catharina, Cathleen, and Charly. Thank you for being an enrichment to the lab crew, for lighting up the daily routine, and for making me aware of how much fun teaching can be. Thank you for listening when I felt the urge to pass on knowledge, for asking for my advice, and for enjoyable moments outside the lab at different conferences and the lab Christmas dinner.

Kristina and Jessica. Thank you for everything lab-related but particularly, thank you for taking care of me that day of my bike accident, just right after being promoted to my foster lab moms. This is what I call irony of fate. Thank you for being colleagues that take their responsibility beyond work.

All other wonderful colleagues, fellow PhDs, and coffee mates. Thank you for creating an enjoyable working atmosphere against all odds. Waltraut, Sercan, Umair, Mengyu, Alice, and everyone else involved. Thank you for taking care of my mice during Christmas and New Year's vacation. Thank you for this favor without complaint, this does not go without saying. Lukas. Thank you for always helping out and sharing your expertise. Laura. Thank you for not only being a lab friend but also for sharing passions outside the scientific world. Lennart. Not sure why it was always science-related dance floors that made us open up and shout out about miserable conditions in academia but I will miss this therapy.

Lissy and Shiv. Thank you for being great PIs with open doors, also to PhDs who are not part of your group. Thank you for answering questions and helping out when I was in need. Thank you for all your undercover work for the department.

Michi, the other animal care takers and vets, and Mrs. Eckermann. Your working areas could not be any more different. But you are united in making me feel comfortable and welcome in a place of science that can be harsh at times. Thank you for helping out and taking over extra tasks that may not even be part of your duties. Thank you for never failing to make me smile in lovely conversations apart from the daily science talk.

Lisa. Thank you for being an irreplaceable lab companion and an even better friend outside the lab. Thank you for being the epitome of reliability and for re-teaching me so many times what a brave woman looks like. Thank you for reminding me that nobody on this earth can take such good care of you than yourself.

Geske. First fellow PhD, then caring postdoc. Nothing can replace a friend who is not only there for you but who is literally there with you. Nothing can replace someone going through exactly the same thing the same time. Nobody can understand the pain, the fear, the

anxiety, the hopelessness, the exhaustion, the defeat than someone who resides together with you at the very bottom. Thank you for sticking with me through the very worst days and nights which we could not tell apart anymore in the windowless seminar room. (Rumor has it this room was built exclusively to keep PhD students in there to work forever.)

Friedi. Thank you for your extraordinary friendship. Thank you for starting the PhD together with me and for making this so much easier. Thank you for proving that quitting does not mean losing but in fact winning and staying true to yourself. Admitting weaknesses is a true strength.

Shilpa. Thank you for making me aware of the PhD movies. Sure, I knew about the comics. Every decent student does. But the movies changed everything. They were the 100% guarantee to make me laugh in any situation. They were my ultimate go-to when I needed a cheer up. What a paradox happy place. (At this point, also thanks to Jorge Cham for making every PhD student feel understood. I am still devastated to date that I missed your talk at the UTK.)

Prof. Dr. Frank Löffler. My mentor at the UTK. Thank you for telling me in our final talk before I left that if I wished to do so, I'd be more than capable to do a PhD. Back then, I didn't even hold a bachelor's degree. Your strong belief in me from that early on is still a mystery to me. But I love how everything turned full cycle in the end. From chicken poo and dechlorinating wastewater bacteria to mouse poo and the fight against contaminating environmental microbiota.

Dr. Mathias Rhein. Thank you for being a wonderful, compassionate human being and for being the first to make me aware of the Impostor phenomenon. It somehow makes a huge difference when you have a name for what is wrong with you. I still treasure this Subway invitation until today.

Marah. During my PhD I lived with a couple of flat mates. But only one of them sticked through this with me until the very end, enduring all of my terribly annoying traits. Never complaining even when I basically vanished the last couple of months. Living together exceeds any regular friendship. Thank you.

Laura. My foster flat mate. Thank you for never being tired of wishing me best of luck, strength, and endurance, and for sprinkling virtual fairy dust on my work. Over and over again. Thank you for making breaks much more fun and less lonely during the last couple of weeks when I was about to mutate into an eremite.

Friends. Every single one. Thank you for enduring my constant venting, thank you for just listening and confirming. Thank you for not being upset when I did not call or text back. Thank you for all of your understanding and for not putting additional pressure on me. Thank you for sharing doubts. Particularly Ronja. Thank you for being the second example that quitting is the epitome of power. No PhD certificate is worth the sacrifice of mental health. Knowing there is no shame in doing so helped me a lot to stick through it. Knowing there is always an exit.

Martin. I once told you I have already come to the point at which I could easily write an entire page of acknowledgements just about you. And at the same time, no words may describe this interpersonal relationship. Thank you for taking care of me when I was not able to do so myself. Thank you for enduring the endless hours of me crying. Thank you for making me laugh again and pulling me out of my darkest moments. You made an example of how a PhD is done in exceptional situations. And you finally holding a PhD against all odds was the ultimate motivator fueling my competitiveness. I could not have put up with that. (Thanks ego.)

Mom, Dad, Bö. Your mostly silent and invisible support is something that I appreciate from the bottom of my heart. At times when I just vanished for months, you showed all the love and understanding and no pressure to send signs of life. This took a lot of weight off my shoulders. Thank you for raising and growing up with me without breaking my stubbornness without which I would not have sticked through this.