**An Integrative Genetic, and Epigenetic
Characterization of
Pancreatic Neuroendocrine Neoplasms (PanNENs)
defines Distinct Molecular Features of
Endocrine and Exocrine-like Subgroups**

D I S S E R T A T I O N

zur Erlangung des akademischen Grades

**Doctor of Philosophy
(Ph.D.)**
im Fach Biologie

eingereicht an der
Lebenswissenschaftlichen Fakultät
der Humboldt-Universität zu Berlin

von
M.Sc. Tincy Simon

Präsidentin der Humboldt-Universität zu Berlin
Prof. Dr.-Ing. Dr. Sabine Kunst

Dekan der Lebenswissenschaftlichen Fakultät
der Humboldt-Universität zu Berlin
Prof. Dr. Dr. Christian Ulrichs

Gutachter/innen

1. Prof. Dr. Christine Sers
2. Prof. Dr. Nils Blüthgen
3. Prof. Dr. med. Aurel Perren

Eingereicht am: August 30, 2021
Tag der mündlichen Prüfung: November 23, 2021

"It is only in our darkest hours that we may discover the true strength of the brilliant light within ourselves that can never, ever, be dimmed."

Doe Zantamata

Abstract

Pancreatic Neuroendocrine Neoplasm (PanNEN) is a rare form of cancer with a diverse and highly debated tumor classification. PanNENs comprise a heterogeneous set of high-grade PanNECs and NETG3 PanNETs, in addition to the more commonly diagnosed low-grade NETG1 and intermediate grade NETG2 PanNETs. High-grade PanNENs display poor prognosis among patients and, despite maintaining some morphological distinctions, PanNEC and NETG3 remain challenging to diagnose histologically. This study presents a methylation-based classification, which precisely distinguishes PanNETs and PanNECs, exposing the underlying complexity evident in molecular and tumor cell-of-origin features of PanNENs. DNA methylation analysis, in addition to mutation and copy number profiling was performed on a PanNEN cohort, comprised of NETG1, NETG2, NETG3 and PanNEC samples, to define and characterize PanNEN subgroups. My work establishes distinct PanNEN subgroups based on methylation profiles, whereby the PanNETs and PanNECs were separated into two groups: Group A and Group B, respectively. While Group B tumors are enriched for recurring alterations in *KRAS*, *TP53* and *SMAD4*, the mutational spectrum of Group A tumors encompasses alterations in classical PanNEN genes including *MEN1*, *DAXX*, *ATRX* and *VHL*. Recurring whole chromosomal aberrations are evident in Group A tumors, in contrast to Group B tumors, which reveal a significant focal deletion of the *RB1* locus. Using methylation profiles of alpha-, beta-, ductal and acinar pancreatic cell types, in addition to expression patterns of normal cell type marker genes within tumors, the study concluded that the PanNET tumors are alpha-like, beta-like, and intermediate-like tumors of endocrine origin, while PanNECs of Group B are acinar-like tumors with a potential exocrine origin. The proportion of acinar-cell methylation profile and expression of SOX9 protein in Group B tumors are comparable to Pancreatic Ductal adenocarcinoma (PDAC), a tumor entity with an established exocrine cell-of-origin. Together, the novel findings of this thesis establish a comprehensive profile distinctly characterizing PanNET and PanNEC tumors at the genetic and epigenetic molecular level. Importantly, this work provides strong evidence for the emerging yet unproven theory of an exocrine origin of PanNECs, offering a new approach for treating patients with this rare but often fatal disease.

Zusammenfassung

Neuroendokrine Neoplasmen der Bauchspeicheldrüse (PanNEN) sind eine seltene Form von Krebs mit einer vielfältigen und stark umstrittenen Tumorklassifikation. PanNENs umfassen eine heterogene Gruppe von hochgradigen PanNECs und NETG3 PanNETs, sowie die häufiger diagnostizierten niedriggradigen NETG1 und mittelgradigen NETG2 PanNETs. Hochgradige PanNENs weisen eine schlechte Prognose auf, und trotz einiger morphologischer Unterscheidungsmerkmale sind PanNECs und NETG3s histologisch schwierig zu diagnostizieren. In dieser Studie wird eine auf Methylierung basierende Klassifizierung vorgestellt, die PanNETs von PanNECs unterscheidet und die zugrunde liegende Komplexität aufzeigt, die in den molekularen Merkmalen und den Ursprungszellen der PanNENs zum Ausdruck kommt. An einer PanNEN-Kohorte, die sich aus NETG1-, NETG2-, NETG3- und PanNEC-Proben zusammensetzt, wurde neben Mutations- und Kopienzahlprofilen auch eine DNA-Methylierungsanalyse durchgeführt, um PanNEN-Untergruppen zu definieren und zu charakterisieren. In meiner Arbeit wurden PanNENs auf Grundlage ihrer Methylierungsprofile gruppiert, wodurch die PanNETs und PanNECs in die zwei Gruppen A und B aufgeteilt wurden. Während Tumore der Gruppe B häufig wiederkehrenden Veränderungen in *KRAS*, *TP53* und *SMAD4* aufweisen, umfasst das Mutationsspektrum von Tumoren der Gruppe A Veränderungen in klassischen PanNEN-Genen wie *MEN1*, *DAXX*, *ATRX* und *VHL*. Darüberhinaus sind Gruppe A Tumore durch wiederkehrende chromosomale Aberrationen gekennzeichnet, wohingegen Tumore der Gruppe B eine signifikante fokale Deletion des *RB1*-Lokus aufweisen. Anhand von Methylierungsprofilen von alpha-, beta-, duktalen und azinären Pankreaszelltypen sowie von Expressionsmustern normaler Zelltyp-Markergene innerhalb der Tumore kam diese Studie zu dem Schluss, dass die PanNET-Tumore alpha-, beta- und intermediär-ähnliche Tumore endokrinen Ursprungs sind, während die PanNECs azinäre Tumore mit einem potenziellen exokrinen Ursprung sind. Letzteres basiert auf der Beobachtung, dass die Ausprägung des Azinuszellenprofils und die Expression des SOX9-Proteins in den Tumoren der Gruppe B vergleichbar mit denen des duktalen Adenokarzinoms des Pankreas (PDAC) ist, einer Tumorentität mit nachgewiesenem exokrinen Zellursprung. Zusammengenommen ergeben die neuen Erkenntnisse dieser Arbeit ein umfassendes Profil, das PanNET- und PanNEC-Tumore auf

genetischer und epigenetischer Ebene eindeutig charakterisiert. Wichtig ist, dass diese Arbeit starke Beweise für die aufkommende, aber unbewiesene Theorie eines exokrinen Ursprungs von PanNECs liefert und somit einen neuen Ansatz für die Behandlung von Patienten mit dieser seltenen, aber oft tödlichen Krankheit bietet.

Acknowledgements

I want to take this opportunity to acknowledge personally few people who have been instrumental, in ensuring success in my doctoral studies.

First and foremost, I want to thank Prof. Dr. Christine Sers, my supervisor, for giving me the utmost freedom and support to pursue what I was passionate about. Christine saw through my potential in theoretical biology and allowed me to pursue a computational project despite hiring me as an experimental biologist. Her confidence and trust in me helped to not only successfully complete my PhD, but also to realize what I wanted to pursue as a career. I want to also acknowledge Prof. Dr. Nils Blüthgen for giving me the opportunity to join the amazing group that you have. In the 6 years of our acquaintanceship, you maintained a group filled with some of the most intellectual and wonderful scientists I have ever come across. By being part of this group, I was able to develop my analytical and computational skills throughout my PhD, which ultimately led to the completion of this project.

I want to thank Prof. Dr. David Capper for joining the Neuropathology and giving his first presentation at the NGS meeting, which led to the beginning of an idea that became the core of my PhD project. I want to thank the amazing group of technicians of Prof. Sers group, especially Andrea Menne, whose experimental support as well as the thorough and organized work maintained a proper order and constant progress during my PhD project.

I want to thank Dr. Bertram Klinger and Mathurin Dorel, for being patient with me, answering my countless questions, at the same time, pushing me with your encouraging words whenever I felt like I reached my limit.

I want to thank Dr. Pamela Riemer, for supporting me for part of the project and for the many scientific discussions that helped me to understand the topic in-depth. I want to thank Dr. Manuela Benary, for motivating me during the writing process, which was extremely difficult during the pandemic. Additionally, I would like to acknowledge Jenny Wiederspahn for reviewing my thesis extensively for language associated errors.

I would like to also acknowledge my graduate schools, Berlin School of Integrative Oncology (BSIO) and CompCancer for providing me the funding, as well as the academic courses, seminar and conference, all of which helped me improve my scientific knowledge and skills.

Above all, I want to thank Dr. Soulafa Mamlouk, who was the backbone and a key player in the success of this project. Thanks Soulafa, for seeing me as an equal, for believing in me, even when I didn't and for being patient with me, even at times it might have been

"Your support network is the solid ground from which you can propel yourself upwards."

Anna Barnes

Contents

List of Figures

To Germany, for the 7 years of experience worth a lifetime . . .

Preface

Cancer at its core is a population of cells that arose from a mutated cell, while thriving at the expense of their neighbors, and ultimately causing the entire cellular system to perish. Given the explosion of technological innovation and strong advancement of cancer research, the concept that cancer comes from a "mutated cell" expanded to "aberrated cell", encompassing the various genetic, epigenetic, and expression changes that can occur within a genome. This process of progression is unique to the organ from which a tumor manifests, as well as to the specific cell type within that organ. Further variables including cellular morphology, proliferation rate, and/or molecular aberrations characterize the underlying heterogeneity within a given tumor-type of a distinct cell-of-origin.

Pancreatic cancers exemplify the aforementioned variables vividly. Pancreatic cancers can arise from three different cells of the pancreas: Pancreatic Ductal Adenocarcinomas (PDAC) stem from ductal cells or acinar cells, and Pancreatic Neuroendocrine Neoplasms (PanNENs) are thought to originate from the endocrine cells of the pancreas. The specific endocrine cell type of the pancreas from which PanNENs arise, however, has not yet been discovered. Though the tumors are of pancreatic origin, the distinct molecular, morphological diversity and potentially even the different normal precursor to these tumors drive different paths of progression. PDAC is the most aggressive tumor of the pancreas with a 5-year survival rate of 7%, while PanNENs are rare tumors that have a general 5-year survival rate of 54%. Time and time again, the overall good prognosis has mistakenly symbolized PanNEN tumors to have an indolent nature. The heterogeneity within this tumor type, which is defined by diverse histopathological features and genetic characteristics, can be attributed to differences in differentiation stage and tumor grade. It can dramatically decline the 5-year survival rate to <1%.

The pancreatic cancer cell-of-origin, particularly in the case of PanNEN, has been a highly debated and hitherto unresolved issue in the field of Gastroenteropancreatic Neuroendocrine Tumors (GEP-NETs). The predominant factor at the core of this debate is the two fundamentally different groups of PanNENs: the well-differentiated, low, intermediate and high grade Pancreatic Neuroendocrine Tumors (PanNETs), and the poorly differentiated highly aggressive tumors, Pancreatic Neuroendocrine Carcinomas (PanNECs). The differences in genetics, morphology and grade collectively presume diverse cellular origins and could potentially rationalize the differences in prognosis evident in PanNEN tumors.

Investigation of tumor heterogeneity, complemented with in-vitro analysis, is critical

for the advancement in therapeutic approaches. As such, having representative *in-vitro* models is an indispensable criterion to continue down this path. Two well-recognized *in-vitro* models used widely among researchers in PanNEN studies are the BON1 and QGP1 cell lines. However, *in-vitro* models come with the caveat of misrepresenting the underlying tumor heterogeneity. Hence, it becomes imperative to thoroughly characterize the tumor *in-vitro* model to pursue a meaningful investigation of potential targets vulnerable to tumor progression.

Here I present my accomplishment in pursuit of advancing the current knowledge of various aspects involved in PanNEN tumorigenesis during my 4-year thesis project. I utilize computational approaches to characterize and establish the multi-omic heterogeneity, fundamentally distinguishing the well-differentiated low, intermediate, and high grade tumors, PanNETs and the poorly differentiated highly aggressive tumors, PanNECs. I provide deep insight into divergent normal precursors that give rise to PanNEN subtypes and underscore the genetic and epigenetic aberrations characterizing these subtypes. Finally, I characterize the cell line models of PanNENs using multi-omics data as well as mathematical modeling of signaling networks, laying the foundational work required to identify novel therapeutic targets for PanNEN treatment. Collectively, my work has progressed the principal knowledge of tumor heterogeneity present in PanNENs and commenced the crucial steps vital to expand existing therapeutic approaches in accordance with the identified heterogeneity.

"Always walk through life as if you have something new to learn, and you will."

Vernon Howard

Chapter 1

Introduction

Cancer is frequently described as a disease with one of the highest mortality rates in humans. Besides the most frequently occurring cancers such as lung, breast or colorectal cancer, there is also a not insignificant number of so called "rare cancers". Even so, the general population unintentionally disregards the rare cancers and subsumes them as indolent in nature. The rare cancer types, which carry in general good prognosis, nevertheless, comprise a heterogeneous set of tumors with hidden aggressive nature. Pancreatic Neuroendocrine Neoplasms (PanNENs) are among such rare tumors with an incidence of approximately 0.5 in 100,000 cases, encompassing <3% of all pancreatic neoplasms (Yadav, Sharma, and Zakalik, 2018). The mean age of diagnosis of PanNENs is 58 years, equally represented in both genders (Brooks, Shavelle, and Vavra-Musser, 2019). The rich complexity of PanNEN tumors, directly influencing the prognosis and overall survival of patients, are concealed within the histopathological, molecular and tumor cell-of-origin heterogeneity.

On the following pages, I will present the state-of-the-art knowledge in each of these different categories that built the current model of the PanNEN heterogeneity, while accentuating what is lacking or requires substantial improvement. I will begin by elaborating on the histopathological diversity seen within PanNEN tumors and further characterize these differences by associating the various levels of molecular heterogeneity observed in PanNENs. Subsequently, as a major topic explored in my thesis, I will introduce what is known in regard to tumor cell-of-origin of PanNENs and highlight what is still lacking pertaining to the poorly differentiated subset of PanNEN tumors. Ultimately, precise therapeutic intervention, built on the established heterogeneity, paves the future for treatment of PanNENs. Along these lines, I will present the current therapeutic approaches for PanNEN treatment and touch on the bona fide in-vitro models of PanNENs employed for advancing the currently limited treatment options, while drawing attention to what needs to be further investigated before utilizing them to study novel therapeutic strategies.

1.1   Histopathological heterogeneity in PanNENs

Histopathological differences of tissue composition and tumor behavior within PanNENs were among the first characteristics employed for sub-classification of the tumors. Tumor grade classification, as well as hormonal production status of tumors, profoundly dominate the histopathological diversity, allowing us to discern certain aspects of the tumors impacting the path to tumorigenesis.

1.1.1   Tumor grade classification

Tumor grade histological classification of PanNENs has endured a long road of re-evaluation in the past two decades due to its rich complexity. The very first tumor classification came about in 1995, wherein neuron-specific markers such as synaptophysin were used to establish the neuroendocrine nature of the tumor, while site, size, morphological difference (well-differentiated vs. poorly differentiated), hormonal status, angioinvasion and metastatic spread were further applied to determine sub-categories with prognostic implications (Capella et al., 1995). This became the foundation for the future WHO classifications, with higher emphasis on the size of the tumor as well as the proliferation index determined either via the mitotic count or Ki67 index (DeLellis RA., Lloyd RV., Heitz PU., 2004). Each of the successive classification schemes showed dramatic improvement. Nevertheless, the lack of precise stratification in addition to vague grade terminology hindered a clear prognostic assessment. To overcome these limitations, the 2010 WHO classification was introduced. PanNENs were now classified into three morphologically distinct subtypes: the well-differentiated Pancreatic Neuroendocrine Tumors (PanNETs), the poorly differentiated Pancreatic Neuroendocrine Carcinomas (PanNECs) and mixed adeno-neuroendocrine carcinomas (MANNECs) (Bosman et al., 2010). Based on Ki67 and mitotic index, PanNETs were further divided into NETG1 and NETG2 and all PanNECs were G3 grade of small or large cell type (SCNEC and LCNEC, respectively). The WHO 2010 classification was well predictive of prognosis and survival (McCall et al., 2013; Rindi et al., 2012). In addition, the new classification underlined the incidence of metastasis, wherein 50% of the PanNET tumors showed metastasis while close to 100% of PanNECs showed metastasis of liver and/or lymph nodes (Basturk et al., 2014; Brooks, Shavelle, and Vavra-Musser, 2019). The grading scheme of 2010 substantially improved the classification system; however, it lacked the incorporation of subtle differences evident in regard to morphology as well as tumor responsiveness to therapy.

The updated and current standard system of histological grading of PanNENs is the WHO 2017 classification. The new scheme refined the classification system of 2010, accommodating for the subtle details observed within each grading classes. The initial change in the 2017 scheme separated PanNENs with a Ki67 index >20% into NETG3 and NEC tumors based

on various factors. The most significant difference was seen at the morphological level wherein, G3 tumors harboring highly proliferative features, now termed NETG3, maintained morphological features similar to well-differentiated tumors (Basturk et al., 2015; Milione et al., 2016) (Table 1.1.1). Furthermore, NETG3 maintained functional and genetic features consistent with a PanNET tumors and distinct from PanNEC tumors (Basturk et al., 2015; Jiao et al., 2011; Konukiewitz et al., 2017; Yachida et al., 2012). Most notably, NETG3 showed better prognosis than NEC and was also unambiguously responsive towards conventional therapy often used for NETG1/G2 tumors (Basturk et al., 2015; Heetfeld et al., 2015; Milione et al., 2016; Vélayoudom-Céphise et al., 2013). In addition to PanNEC classification changes, the Ki67 index that characterized a NETG1 tumor was modified to <3%. The final adaptation to the grading system involved replacing MANNEC with MiNEN, or mixed neuroendocrine and non-neuroendocrine neoplasm, to include the observation that not all mixed components were of glandular origin and that the neuroendocrine component can in fact be well-differentiated (La Rosa, Sessa, and Uccella, 2016). The major subclasses, PanNET and PanNEC as asserted above, have clear morphological features distinguishing one another. One can observe distinctive architectures defining the overall appearance of the tissue within each subgroup that further add to their individual complexities. NETG1 and NETG2 maintain islet-like small cells with salt-and-pepper chromatin appearance, but can also aggregate into solid, trabecular, glandular or tubulo-acinar patterns within the tumor region (Fukushima, 2017) (Figure 1.1). These features of low grade PanNETs are also often present in NETG3 tumors, while PanNEC tumors show a highly different morphology (Tang et al., 2016; Basturk et al., 2014). SCNEC have small atypical neoplastic cells with a high nucleus-to-cytoplasm ratio and are characterized by hyperchromatic nuclei and nuclear molding. LCNEC, on the other hand, can have round to polygonal cells and their nuclei have either vesicular chromatin or prominent nucleoli (Basturk et al., 2014) (Figure 1.1). Although cell morphology has demonstrated to be a strong means of classification, it has remained significantly challenging to establish a consistent and accurate diagnosis of NETG3 and NEC due to the presence of overlapping histopathological features. Pathologists have experimented on combining molecular and cell morphological features in addition to prior clinical knowledge, including the presence and strength of neuroendocrine markers to eliminate this uncertainty (Tang et al., 2016). The accuracy of correct diagnosis indeed increased significantly as indicated by the stark difference in prognosis (75 vs 11 months for well-differentiated PanNET and Poorly differentiated PanNEC, respectively); nevertheless, the additional features improving the diagnostics are not recurring events and therefore fail to differentiate 5% of ambiguous cases or lead to misclassification (Tang et al., 2016). Taken together, the current grade classification system of PanNENs has dramatically improved to define the morphologically diverse tumor subgroups, yet fails to consistently classify NETG3 and NECs, which is indispensable for accurate therapeutic and

prognostic approaches. Further studies are prompted to characterize the molecular features distinguishing the PanNET and PanNEC subcategories.
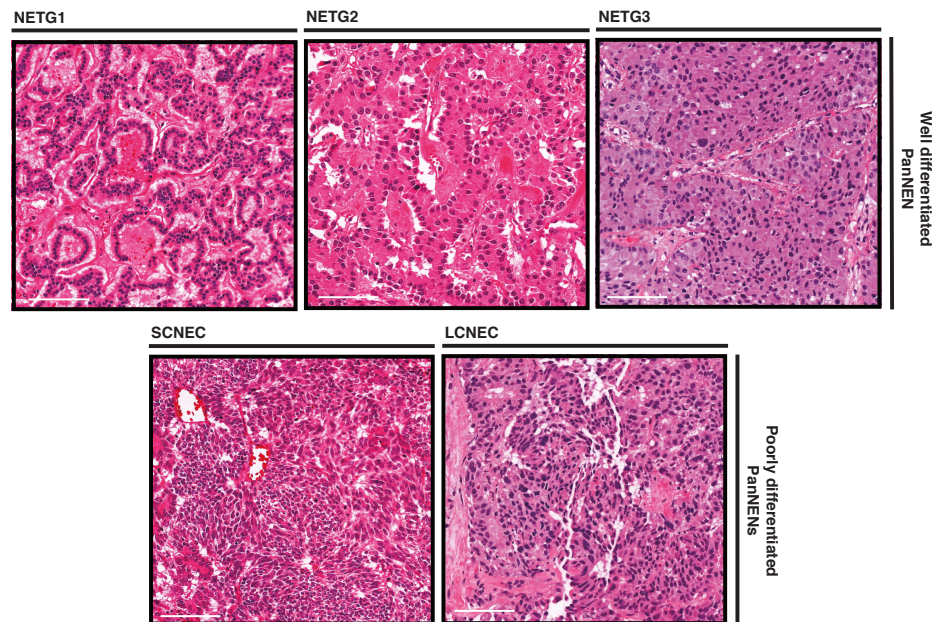
FIGURE 1.1: **Histological diversity in PanNEN.** Representative histological sections of well-differentiated NETG1, NETG2 and NETG3 tumors (top panel) as well as poorly differentiated PanNEC tumors (bottom panel). Small cell NEC (SCNEC) and large-cell NEC (LCNEC) are further denoted in the images. Images are obtained from an in-house PanNEN cohort. Scale bar represent 100 µm.

1.1.2 Functionality of PanNENs

The term "functionality" in the context of PanNEN tumors signifies their potential to exhibit hormone associated symptoms. Consequently, PanNEN tumors can be phenotypically characterized as hormone producing functional PanNEN or non-functional PanNEN that maintains largely asymptomatic behavior and often remains undetected until symptoms arise due to tumor mass (Halfdanarson et al., 2008).

Functional PanNETs

Functional PanNETs are extremely rare in the populations and can manifest as sporadic tumors or as a result of a hereditary syndrome. They are often presented as NETG1 and NETG2 tumors and never occur as PanNEC tumors. Insulinomas represent 30-45% of all identified functional PanNETs and are often harboring solitary, benign small tumors with

WHO 2010 Classification			
	Tumor Grade	Ki67	Mitotic Index: High-Power Field (HPF)
PanNET	NETG1	<= 2%	<2/10
	NETG2	3-20%	2/20/10
PanNEC	NEC	>20%	>20/10
WHO 2017 Classification			
	Tumor Grade	Ki67	Mitotic Index: High-Power Field (HPF)
PanNET	NETG1	<3%	<2/10
	NETG2	3-20%	2/20/10
	NETG3	>20%	>20/10
PanNEC	NEC	>20%	>20/10

TABLE 1.1.1: **WHO Classification of PanNENs.** PanNEN tumor grade classification difference between WHO 2010 and WHO 2017 system. The table is adapted from (Klöppel et al., 2017)

<10% malignancy rate (Halfdanarson et al., 2008; Crippa et al., 2012; Keutgen, Nilubol, and Kebebew, 2016). The majority of the insulinomas are sporadic tumors, with the exception of 10% of cases associated with hereditary Multiple endocrine Neoplasia type 1 (MEN1) syndrome (Krampitz and Norton, 2013). Surgical intervention is often the curative care for insulinoma with a success rate of 98%, frequently dependent on the grade of the tumor at the time of presentation and on whether complete resection was achieved during surgery (Tucker, Crotty, and Conlon, 2006). Glucagonomas constitute about 1/3 of all diagnosed functional PanNETs, and the remaining infrequent functional PanNETs include vaso active peptide (VIP)-omas and somatostatinomas. Although comparatively rarer than insulinomas, these functional PanNETs are often large (>5 cm) lesions exhibiting aggressive malignant behavior in 60-80% of the cases and manifest when the disease has locally advanced and/or metastasized (Jensen et al., 2012; Frankton and Bloom, 1996). Similar to insulinoma, MEN1 syndrome associated glucagonomas and somatostatinomas are also frequent in these functional tumor types. Patients suffer largely from the obstruction caused by excessive hormonal production. As such, the treatment approach therefore entails controlling the symptoms associated with the excessive hormone production and, if possible, resection of the tumor. The major treatment action against glucagonoma symptoms is providing a somatostatin analog (Metz and Jensen, 2008), while VIPoma and somatostatinoma patients are treated with octreotide or lanreotide to improve the hormone associated symptoms (Jensen et al., 2012). Albeit the fact that functional PanNETs are considerably few among the diagnosed cases of PanNENs and can be managed by regulating hormonal associated symptoms, it is vividly apparent that these tumors manifest, more often than not, in a most aggressive and malignant manner.

Non-functional PanNETs

Ninety percent of the diagnosed PanNENs are non-functional neoplasms (Brooks, Shavelle, and Vavra-Musser, 2019; Halfdanarson et al., 2008; Krampitz and Norton, 2013). Granted they are considered "non-functional", the tumors can in fact be hormone producing, however, they do not exhibit any clinical symptoms associated with excessive production analogous to functional PanNETs (Krampitz and Norton, 2013). These tumors are often detected from the sheer tumor mass or complications that arise pertaining to the mass. Therefore, patients are frequently presented with a liver metastasis, or after their disease had advanced regionally (Brooks, Shavelle, and Vavra-Musser, 2019; Franko et al., 2010; Halfdanarson et al., 2008). Prognosis of non-functional PanNETs is significantly dependent on multiple factors including age, the grade and stage of the tumor, localization as well as treatment approach. In fact, tumor grade and localization significantly impinge on the possibility of surgical intervention as advanced metastasized tumors that are higher in grade are usually not resectable (Brooks, Shavelle, and Vavra-Musser, 2019; Franko et al., 2010; Halfdanarson et al., 2008). In such a case, median survival time dramatically decreases from 16 years to <1 year (Brooks, Shavelle, and Vavra-Musser, 2019). Additionally, low and intermediate grade PanNETs metastasize in only about 34% of the cases, while this proportion drastically increases to 80% in the case of NETG3. In the case of PanNEC tumors, the tumors almost always occur as metastases (Basturk et al., 2014; Fang and Shi, 2019). Due to the vast majority of PanNENs being non-functional tumors and exhibiting such a diverse prognostic nature, significant efforts have been made to further understand the molecular characteristics underscoring these heterogeneous attributes.

1.2 Molecular heterogeneity in PanNENs

Initial knowledge of the molecular characterization of PanNETs came from the data gathered on hereditary syndromes linked to PanNENs. In the past decade, however, an upsurge in multi-omics data furthered our understanding of the underlying genetic and epigenetic heterogeneity promoting PanNEN tumorigenesis.

1.2.1 Germline mutations in PanNETs

The majority of the PanNEN tumors occurs sporadically, yet 10% can manifest due to a hereditary syndrome associated with germline alterations.
MEN1 hereditary syndrome, also known as Wermer syndrome, is an autosomal dominantly inherited disease and is characterized by microadenomas of parathyroid, pancreatic and pituitary endocrine glands (Anlauf et al., 2006; Chandrasekharappa et al., 1997; Marx and Simonds, 2005). While the majority of the MEN1 associated PanNETs are non-functional

tumors, a considerable number of functional PanNETs also arise from this particular syndrome (Jensen et al., 2008). Although MEN1 syndrome causes lesions in multiple tissues of endocrine origin, MEN1 associated PanNETs are often malignant metastases causing mortality in more than 50% of these patients (Conemans et al., 2017; Dean et al., 2000; Goudet et al., 2010). Alterations in the Multiple Endocrine Neoplasia type 1 *(MEN1)* gene is the direct causative factor of MEN1 hereditary syndrome. The gene is located in the long arm of chromosome 11 and codes for the 601 amino acid long nuclear protein Menin, which is expressed in various endocrine tissues (Chandrasekharappa et al., 1997). Menin protein interacts with a multitude of transcription factors in order to influence several signaling pathways involved in proliferation, differentiation and survival (Agarwal et al., 1999; Jin et al., 2003; Razmara, Monazzam, and Skogseid, 2018; Sowa et al., 2003; Sowa et al., 2004). In addition, Menin interacts with chromatin remodeling complexes to further influence transcriptional regulations (Huang et al., 2012; Hughes et al., 2004; Murai et al., 2011). MEN1 syndrome patients often carry mutations in the *MEN1* locus and more than 80% of these cases have loss of heterozygosity (LOH) of the chromosomal (11q13) region (Lemos and Thakker, 2008; Perren et al., 2007). Interestingly, 30% of MEN1 patients have no germline alteration in the *MEN1* locus, nevertheless germline alterations of Cyclin Dependent Kinase Inhibitor 1b *CDKN1B)*, a downstream target of Menin, have been reported in *MEN1* mutation-null cases (Belar et al., 2012; Georgitsi et al., 2007).

VHL disease is the second most common autosomal dominant hereditary syndrome, which also manifests in PanNET tumors in more than 20% of the cases. Distant metastases are exhibited in 10-20% of these cases (Blansfield et al., 2007; Yamasaki et al., 2006). VHL associated PanNETs are exclusively non-functional, with 80% of the cases harboring germline alteration in the Von Hippel-Lindau *(VHL)* gene (Blansfield et al., 2007; Hammel et al., 2000). LOH of the *VHL* locus is also evident in VHL disease associated PanNETs (Lott et al., 2002). The gene is localized to the short arm of chromosome 3 and codes for the 213 amino acid long VHL protein (Latif et al., 1993). It has a critical role in the Hypoxia inducible factor (HIF) pathway. Under normoxic condition, VHL is part of the VBC-Cul2 complex along with Elongin B, Elongin C, Cullin-2 and Rbx1 proteins in order to ubiquitinate the HIF transcription factor for its subsequent degradation (Kamura et al., 1999; Pause et al., 1997). Loss of function of VHL prompts the accumulation of HIF, resulting in the transcription of several important genes, including angiogenesis associated Vascular Endothelial Growth Factor *(VEGF)*, Platelet-derived Growth Factor *(PDGF)* and Transforming Growth Factor Alpha *(TGF-α)* (Kaelin and Maher, 1998; Maxwell et al., 1999). Interestingly, *VHL* gene mutational events never appear in sporadic PanNETs, despite their involvement in hereditary diseases (Schmitt et al., 2009; Moore et al., 2001; Scarpa et al., 2017).

PanNETs associated with Neurofibromatosis *(NF-1)* hereditary disease are uncommon and occur in <10% of the cases (Fujisawa et al., 2002). In addition, functional or nonfunctional

PanNETs arising from Tuberous Sclerosis Complex 2 (*TSC2*) associated syndrome are only found in rare occasions ($\approx$1%) (Francalanci et al., 2003; Kim, Kerr, and Morehouse, 1995; Merritt et al., 2006; Schwarzkopf and Pfisterer, 1994).

Recent whole genome analysis of 100 tumors revealed that non-syndrome associated rare germline events act as clear drivers of tumorigenesis in PanNETs (Scarpa et al., 2017). Deleterious mutation in the DNA-repair gene MutY DNA Gycosylase *(MUTYH)*, leading to a mutational signature harboring G:C>T:A transversions, Breast Cancer 2 *(BRCA2)* mutation, resulting in the BRCA deficiency mutational signature, as well as alterations in *CDKN1B* and Checkpoint Kinase 2 *(CHEK2)* were among the detected germline events (Al-Tassan et al., 2002; Shinmura et al., 2000; Scarpa et al., 2017).

Undoubtedly, germline alterations as single events or in concert with an associated hereditary syndrome create an additional layer of genetic heterogeneity and can have a significant impact on multiple cellular processes in order to promote tumor progression.

1.2.2   Somatic mutations in PanNETs

Somatic mutational heterogeneity in PanNETs is extensive in that putative recurrent genes are quite scarce and rare variants are identified in many protein coding genes.

The most frequent alterations in the mutational landscape of non-functional PanNETs, identified as somatic driver alterations, are missense variants and large base-pair deletions detected within the *MEN1* gene (37%) (Corbo et al., 2010; Jiao et al., 2011; Scarpa et al., 2017). The alterations identified spanned the entire Menin protein domain and no hotspot regions were evident (Corbo et al., 2010; Scarpa et al., 2017). Truncating mutations within *MEN1* translated to the lack of nuclear staining within these tumors, while remaining alterations display a strong presence of Menin protein in the cytoplasm (Corbo et al., 2010). *MEN1* alterations were significantly associated with alternative lengthening of telomeres (ALT) positivity, granted it is a transcription factor of the Telomerase Reverse Transcriptase *(TERT)* gene, and thereby influences the telomerase machinery (Lin and Elledge, 2003; Scarpa et al., 2017). Consequently, the role of deregulation in chromatin remodeling processes was speculated. Additional recurring PanNEN mutations detected were also chromatin remodeling associated genes: alterations in Death Domain-Associated Protein 6 *(DAXX)* and Alpha-Thalassemia/Mental Retardation X-linked *(ATRX)* (22% and 11%, respectively) taking place in a mutually exclusive manner (Jiao et al., 2011; Scarpa et al., 2017). The majority of the *DAXX* mutations and a subset of *ATRX* mutations resulted in loss of its protein expression in their respective tumors (Pea et al., 2018). *DAXX* encodes a multifunctional protein that is primarily found in the nucleus as part of the promyelocytic leukemia protein nuclear bodies (PML-NBs). An important role of this complex involves enhancing Fas-mediated apoptosis

by regulating transcription of genes associated with the process (Torii et al., 1999). Nevertheless, in the context of chromatin remodeling processes, PML bound DAXX in cooperation with ATRX modulates chromatin remodeling activity by depositing the H3.3 histone variant in heterochromatin region such as centromeres and telomeres (Drané et al., 2010; Goldberg et al., 2010; Wong et al., 2010). As such, ALT has been reported in the majority of the cases carrying an *ATRX* or *DAXX* mutation, while emphasizing its significant association with metastasis of low grade PanNETs (Pea et al., 2018; Scarpa et al., 2017). In addition, mutations in the *DAXX* gene alone were significantly associated with metastasis among low grade PanNETs (Cives et al., 2019; Scarpa et al., 2017). *MEN1*, *DAXX*, *ATRX* mutations together constitute the most frequently mutated genes. In fact, their mutational status is thus far the sole genetic marker associated with prognosis of PanNET tumors; cases harboring alterations in any of the three genes or in combination showed an overall reduction in disease-free survival in low grade PanNETs compared to the wild type group (Chan et al., 2018; Marinoni et al., 2014; Yuan et al., 2014).

Rare mutational variants have also been identified in other chromatin remodeling associated genes such as AT-Rich Interactive Domain-containing protein 1A *(ARID1A)*, SET Domain containing 2 *(SETD2)*, Chromodomain Helicase DNA binding protein 8 *(CHD8)* and DNA (cytosine-5)-Methyltransferase 1 *(DNMT1)*, as well as in pro-apoptotic gene Cytoplasmic FMR1 Interacting Protein 2 *(CYFIP2)* (Vandamme et al., 2019; Roy et al., 2018). A recent study that specifically investigated well-differentiated distant metastasis revealed that major molecular features characterizing the distant metastases included loss of DAXX, ATRX, ARID1A, p16 proteins and/or H3K36me3 mark deficiency due to disruption of SETD2 function. These aberrations significantly decreased disease-free survival time; 98% of the patient cohort void of these alterations had a disease-free survival of 5 years, while it reduced to 39% in cohorts harboring at least one of these alterations (Roy et al., 2018).

Genes involved in the Mammalian Target of Rapamycin (mTOR)/Phosphoinositide-3-Kinases (PI3K)/Protein Kinase B(AKT) signaling axis are also mutated in 14% of PanNETs (Jiao et al., 2011; Scarpa et al., 2017). Commonly altered genes among these were Phosphatase and Tensin Homolog *(PTEN)*, *TSC2*, Phosphatidylinositol-4,5-bisphosphate 3-Kinase Catalytic Subunit Alpha *(PIK3CA)* and DEP Domain Containing 5, GATOR1 Subcomplex Subunit *(DEPDC5)*. Although mutational occurrence was low, TSC2 and PTEN expressions were significantly downregulated in the majority of PanNET tumors (Missiaglia et al., 2010); low cytoplasmic expression of TSC2 and PTEN protein expression showed significant correlation with aggressiveness of the tumor (based on WHO 2004 classification and metastasis status). In addition, their expression status correlated with shorter disease-free survival (Missiaglia et al., 2010).

Gene fusion events have also been detected as potential genetic aberrations driving the tumorigenesis in PanNETs. Recent studies identified novel *TSC1-TMEM71* fusions which

disrupt the mTOR regulatory activity of TSC Complex Subunit 1 *(TSC1)* and thereby hyper-activate the mTOR signaling. Furthermore, two different fusions with BEN Domain Containing 2 *(BEND2)* gene (*CHD7-BEND2* and *ESWR1-BEND2*) transactivated *BEND2* and promoted oncogenic activity (Scarpa et al., 2017; Williamson et al., 2019). Interestingly, a single *TSC1-CHD7* fusion was specifically identified in NETG3 tumors.

In summary, the mutational landscape of PanNETs evidently lacks a significant number of recurring genes and highlights strong inter-patient heterogeneity, yet clear convergence of alterations to a few molecular pathways is convincingly apparent in PanNETs.

The PanNEC mutational spectrum is not as widely explored as with PanNETs, potentially owing to the fact that the vast majority of PanNENs are PanNETs (92.5%) and only about 10% are PanNECs. In contrast to PanNETs, PanNECs are characterized by recurring mutations in Ki-ras2 Kirsten Rat Sarcoma Viral Oncogene *(KRAS)* (48.7%), Tumor Protein 53 *(TP53)* (66%) and Retinoblastoma *(RB1)* genes (16%) (Hijioka et al., 2017; Hijioka et al., 2015; Konukiewitz et al., 2018; Yachida et al., 2012). Mutations in SMAD Family Member 4 *(SMAD4)* gene, although not as frequent, were also identified in PanNECs (Bartsch et al., 1999). *RB1* mutations in PanNECs are often seen in combination with its loss of expression (50-70%), and PanNECs void of *RB1* mutations showed a loss of expression of p16 protein, suggesting the convergence of the alterations in disrupting cell cycle control (Hijioka et al., 2017; Yachida et al., 2012). Loss of p16 expression was seen in 75% of distant metastases and further associated with poor survival among patients harboring the alterations (Roy et al., 2018). *TP53* alterations also resulted in abnormal immunohistochemistry (IHC) staining, highlighting the direct effect of *TP53* mutation on protein expression. Non-recurring mutations in various genes were also associated with PanNEC, and they seldom overlapped with NETG3 tumors. However, comparison of recurrently and non-recurrently altered genes in regard to biological processes revealed, both shared and exclusively altered signaling axes between the high grade tumors (Konukiewitz et al., 2018); gene mutations converging in cell cycle/DNA repair/transcriptional control, epigenetic regulators, ErbB-, WNT- and mitogen-activated protein kinase (MAPK) signaling pathways were shared between NETG3 and NEC tumors. On the other hand, rare alterations in PI3K/AKT pathway were present in PanNEC tumors while the mTOR pathway was consistently altered in NETG3 when compared to NEC tumors.(Konukiewitz et al., 2018). Yet again, we are confronted by the clear distinction between PanNETs and PanNECs: The histologically different tumors also overtly differ at the genetic level.

1.2.3   Copy Number Aberrations (CNA) in PanNENs

Analogous to mutational profiling, initial copy number changes observed in PanNENs became evident through the analysis of hereditary syndromes, wherein all the altered tumor

suppressors were accompanied by the LOH feature (Perren et al., 2007; Lemos and Thakker, 2008; Lott et al., 2002). Subsequent genetic analysis expanded this list to determine other loci- and/or, region-specific copy number alterations that are associated with sporadic functional and non-functional PanNETs.

The vast majority of PanNETs being sporadic non-functional tumors essentially lead to the identification of unique Copy Number Aberrations (CNA) of this subset of tumors. Studies on non-functional PanNETs revealed that >68% of non-functional PanNETs harbored LOH of 11q in combination with 6q; additional frequent losses in chromosomes 1,3, and 22 were also evident in sporadic non-functional PanNETs (Nagano et al., 2007; Rigaud et al., 2001). With respect to gains, chromosomes 5,7,12,14,17 and 20 were frequently aberrated in PanNETs (Nagano et al., 2007). Remarkably, the aforementioned aberrations were significantly associated with one another, suggesting that they are co-occurring events in PanNETs. The current advancement in next-generation sequencing reaffirmed this theory and demonstrated that CNA identified as localized events were in fact part of global wide changes present in PanNET tumors (Hong et al., 2020; Lawrence et al., 2018; Pea et al., 2018; Scarpa et al., 2017). CNA heterogeneity classifying PanNET tumors can therefore be collapsed into four distinct signatures based on chromosome arm length CNA patterns: I) recurring copy number losses of specific chromosomes, II) limited copy number events with mostly loss of chromosome 11, III) polyploidy tumors and IV) aneuploidy tumors (Scarpa et al., 2017). In addition, recurring mutations were also associated with the CNA signatures; the first two groups were further characterized by *MEN1* mutations in combination with loss of chromosome 11, resulting in a bi-allelic inactivation of this tumor suppressor gene (Lawrence et al., 2018). The aneuploid tumors in non-functional PanNETs also show frequent DAXX/ATRX loss of expression along with an alternative telomerase lengthening phenotype (Pea et al., 2018). An alternative succinct stratification of the above CNA patterns was achieved in recent work while accommodating for aberrations seen in functional PanNETs (Hong et al., 2020). NF-PanNET copy number signatures were redefined as I) CNA deletion groups with high LOH of various tumor suppressor genes, II) CNA amplification groups and III) CNA neutral groups. Identified Insulinomas carried CNA neutral subtype and also CNA amplification groups enriched for Ying Yang 1 *(YY1)* T372R hotspot mutations (Cao et al., 2013; Hong et al., 2020).

Evidently, PanNET copy number signatures are very well established. In contrast, region-specific and/or global changes in chromosomes have yet to be investigated in the context of PanNECs.

1.2.4   Epigenetic changes in PanNENs

Epigenetic analyses performed on PanNEN tumors were primarily focused on DNA methylation status on few candidate genes. Therefore, a clear global signature identifying PanNEN heterogeneity at the epigenetic level is lacking. In a comparatively small study, a 33 gene-based classification of well-differentiated low/intermediate grade PanNET identified three clusters: cluster 1 comprising tumors highly hypomethylated and resembling normal pancreas, cluster 2 harboring tumors with recurring hypermethylation of Caspase-8 *(CASP8)*, Glutathione S-transferase P *(GSTP1)* and Ras association domain family 1 *(RASSF1)*, and cluster 3 consisting of higher frequency of hypermethylation in a subset of genes compared to cluster 1 and 2 (Stefanoli et al., 2014). Based on candidate chromosomal region analysis, cluster 3 also harbored high CNA and tumors were associated with poor prognosis.

Some of the initial studies further evaluated gene promoter hypermethylation in candidate genes that play critical roles in cancer formation and progression. The most recurrently hypermethylated promoters included those of *RASSF1A*(75%), Cyclin Dependent Kinase Inhibitor 2A *(CDKN2A)*(40%), O-6-Methylguanine-DNA Methyltransferase *(O-MGMT)*(40%), Retinoic Acid Receptor Beta *(RAR-β)*(25%) and MutL Homolog 1 *(hMLH1)* (23%) (Arnold et al., 2007; Dammann et al., 2003; House et al., 2003; Liu et al., 2005). Hypermethylation in multiple genes among the aforementioned genes associated significantly with reduced disease-free survival among PanNET patients (House et al., 2003). The most recurring hypermethylation of *RASSF1A* located in 3p21.3 locus, is often found accompanying an LOH, resulting in downregulation of its protein expression. Hypermethylation of *RASSF1A* was, in fact, significantly associated with metastasis (Liu et al., 2005). This multi-functional protein has been extensively studied for its involvement as a tumor suppressor in regulating apoptosis via direct interaction with activated KRAS and by indirectly regulating p53 stability via MDM2 (Dubois et al., 2019). Similar to *RASSF1A*, TIMP Metallopeptidase Inhibitor 3 *(TIMP-3)* is an additional tumor suppressor gene that is frequently hypermethylated (62%). Hypermethylation and subsequent loss of its expression is significantly associated with metastasis in PanNETs. Additional tumor suppressor genes were also identified in lower frequency to be involved in PanNENs via promoter hypermethylation, suggesting a unique mechanism promoting tumorigenesis. Methylation change involving the α-internexin *(INA)* promoter is one of the best characterized clinicopathological aberrations, wherein hypermethylation and reduced expression was significantly associated with advanced stage, metastasis, recurrence and shorter overall survival among PanNET patients (Liu et al., 2014). In addition to these recurring individual methylation changes, few publications have also explored how genomic alterations can work complementary to or in combination with the frequently observed *MEN1* mutations. Pleckstrin Homology Like Domain Family A

Member 3 *(PHLDA3)* hypermethylation and LOH were often seen among patients harboring *MEN1* mutations (Ohki et al., 2014). Furthermore, in the absence of Menin, changes in Histone 3 Lysine 27 tri-methylation (H3K27me3) at the promoters of Menin target genes and subsequent loss of expression of a key target gene, IGF2BP2, were observed in pancreatic islet *MEN1* null cells (Lin et al., 2015). Eventhough, candidate aberrations recurring or otherwise have been identified in PanNEN, a clear global signature identifying PanNEN heterogeneity at the epigenetic level is indeed lacking.

Various studies have highlighted the stark difference in methylation status within multiple genes between histopathological subgroups (Dejeux et al., 2009; How-Kit et al., 2015). Sporadic insulinomas often carry promoter hypermethylation and subsequent loss of expression in *MLH1* gene, which has been significantly attributed to malignancy in these functional PanNETs (Mei et al., 2009). Variation in methylation status was also attributed to grades, whereby gene methylation was more prominent among NETG1 compared to NETG3 well-differentiated PanNETs (How-Kit et al., 2015). Specific to NETG3 tumors were hypermethylation and subsequent lack of protein expression of Homeobox Only Protein homeobox *(HOPX)* gene, emphasizing its potential as a prognostic biomarker for NETG3 tumors (Ushiku et al., 2016). Taken together, epigenetic changes, particularly DNA methylation aberrations, further affirm the histopathological heterogeneity within PanNEN tumors.

1.3 Tumor cell-of-origin heterogeneity in PanNENs

Cell-of-origin is a difficult and seldom investigated area in PanNENs. These tumors are presumed to have an origin from the islet of Langerhans, a cluster of 4 major cell types, aggregated into small islands and dispersed throughout the pancreas. The organogenesis and subsequent differentiation and establishment of different pancreatic cell types have been extensively studied in mouse models (Cano et al., 2014). The final adult pancreas encompasses the exocrine compartment harboring the ductal cells as well as the acinar cells (30% and 70%, respectively), and the endocrine compartment enclosing approximately 70% insulin-producing beta (β) cells, 20% glucagon-producing alpha (α) cells, <10% somatostatin-containing delta (δ) cells and <5% the pancreatic polypeptide producing (PP) cells (Bouwens and Pipeleers, 1998; Stefan et al., 1982; Rahier, Goebbels, and Henquin, 1983; Clark et al., 1988). These islet cells are organized along the blood vessels and maintain unique paracrine interactions between one another (Cabrera et al., 2006). Therefore, PanNEN can potentially arise from any or a combination of these cell types.

1.3.1   Current understanding of tumor cell-of-origin in PanNENs

The first publications exploring tumor cell-of-origin appeared only within the time frame of the past two years, shedding light on potential normal cell identity of well-differentiated PanNETs. Chan *et. al.* performed a study of PanNET G1 and G2 samples, investigating the association of gene expression in samples carrying classical PanNET mutations and gene expression profiles of normal pancreatic cell types. Within the cohort, 58% of the cases harbored *MEN1*, *DAXX*, and/or *ATRX* (A-D-M) mutations (Chan et al., 2018). Unsupervised hierarchical clustering of 3000 most variable genes expressed identified two distinct subtypes and showed enrichment for A-D-M alterations in one subtype. The A-D-M signature exhibited high similarity to the normal α-cell signature, suggesting a singular cell-of-origin for these tumors. The A-D-M expression signature included upregulation of Hepatocyte Nuclear Factor 1 Homeobox A *(HNF1A)* and was enriched for genes with HNF1A transcription factor motifs. They also showed hypermethylation and downregulation of Pancreatic and Duodenal Homeobox 1 *(PDX1)*. Both of these features are consistent with an α-cell origin, suggesting a subgroup of well-differentiated PanNETs harboring an A-D-M mutational signature arising from an α-cell of the pancreas.

Stratification of PanNETs using epigenetic changes in combination with expression has been undertaken as a means to establish biological classification pertaining to normal cell lineages and to further guide treatment. The approach from Cejas *et al.* is in fact a breakthrough work, highly contributing to the cell-of-origin study in PanNETs while at the same time readily applicable in clinical settings. However, gene expression profiles can change depending on tumor condition, and therefore lack the stability to decipher the true normal cell type features. In light of this, epigenetic features have been explored to further validate and advance the cell-of-origin identity.

Utilizing Histone 3 Lysine 27 acetylation (H3K27ac) super enhancer mark, PanNETs can be sub-divided into 3 main groups (Cejas et al., 2019). Group A and B PanNET-restricted enhancer marks enriched for areas of open chromatin in normal α- and β-cells, respectively. In addition, significantly higher H3K27ac marks were present in α-specific gene loci, Aristaless Related Homeobox *(ARX)* and Iroquois Homeobox 2 *(IRX2)* within group A compared to group B, while higher H3K27ac marks specific to group B was observed in β-specific gene loci, *PDX1*. Expression profiling, in comparison to normal α- and β-islet cells showed significant enrichment of α- and β-specific transcripts in group A and group B, respectively, establishing that group A and group B resemble mature α- and β-cells at multiple molecular levels and could very well be the origin of each subtype. IHC further corroborated these finding and therefore, the groups were henceforth identified as ARX+ and PDX1+ subtypes. As for group C, H3K27ac marks at these established loci of Group A and B varied and expression profiling and IHC showed coexpression of PDX1 and ARX protein. Therefore, Group C, thereafter known as the double positive (DP) group, could have either dual

lineage or a lineage indicative of α- β- progenitor cells[1]. Further, mutational, clinical and biomarker characterization of each subtypes revealed that ARX+ and DN cases showed significantly higher ALT, compared to PDX1+ and DP+ tumors, and relapse was significantly associated with ARX+ALT+ tumors. A methylation-based approach has also been employed and came to a similar conclusion as Cejas *et al* (Domenico et al., 2020).

Utilizing differentially methylated sites of α- and β-cells as references, Domenico *et. al.* identified three groups of PanNETs: α-like resembling ARX+ subtype, β-like similar to PDX1+ subtype and intermediate-like tumors representing DP subtype. Based on their analysis, disease-free survival was strongly reduced in the intermediate-like subgroup, however, α-like tumors with loss of DAXX or ATRX had significantly shorter survival compared to patients with β-like or intermediate-like tumors. Undoubtedly, both publications arrived at the same conclusion that PanNETs indeed have an endocrine cell of origin and that they could originate from an α-, β- or a progenitor of both lineages.

It is immediately apparent that each of these studies lack the classification of PanNEC and NETG3 tumors with respect to tumor cell-of-origin. In fact, studies analyzing the normal cell identity of NETG3 and PanNEC tumors boil down to one publication, where the authors claim that *KRAS*-mutant NEC tumors are of ductal lineage due to the expression of MUC1 and carcinoembryonic antigen (CEA) (Konukiewitz et al., 2018). Intriguingly, this statement sets in motion a bold new theory: PanNEN tumors, specifically the poorly differentiated PanNECs, could have an origin beyond the known endocrine lineage.

This theory, however, comes with its own controversies since tumors originating from the exocrine lineage are considered to be pancreatic adenocarcinomas (PDAC). PDACs are characterized by ductal glandular morphology and therefore conventionally recognized to have ductal cell-of-origin. Interestingly, these tumors are characterized by highly frequent *KRAS* and *TP53* mutations comparable to PanNEC tumors. Recent evidence from mouse model studies, however, suggests that PDACs can in fact be derived from both ductal and acinar cells depending on the type of genetic aberration, the precursor lesions and the timing of the event in the precursor lesions. Pancreatic Intraepithelial Neoplasms (PanINs) are precursor lesions derived from acinar cells, while aberrations in ductal cells often form Intraductal Papillary Mucinous Neoplasms (IPMNs) (De La O et al., 2008; Habbe et al., 2008). Embryonic acinar cells or those having a progenitor-like phenotype are more susceptible to *KRAS* mutation and readily form PDACs via PanIN lesions while ductal cells are seemingly resistant to *KRAS* mutation and require additional aberrations such as *TP53* to induce PDAC progression from IPMNs (Guerra et al., 2007; Kopp et al., 2012; Ray et al., 2011). Mouse

---

[1]IHC analysis also showed occurrence of double negative (DN) tumors that the authors did not characterize with respect to normal-cell identity or recognize potential similarity to any of the groups

models harboring *KRAS* G12V mutations in targeted precursors cells of acinar lineage read-ily developed PanINs and progressed to PDAC, further demonstrating acinar cells as an-other cell-of-origin of PDACs (Guerra et al., 2007). Interestingly, adult acinar cells can also develop PDAC tumors by first undergoing acinar-to-ductal metaplasia (ADM) to revert the adult cells to a more progenitor-like cells, thereby sensitizing the cells to *KRAS* mutations and subsequent PDAC formation (Liou et al., 2013; Kopp et al., 2012).

SRY-BOX Transcription Factor 9 (SOX9) protein expression induced by SMARCA4 tran-scription activator is at the forefront of this transition of acinar cells to ductal-like cells (Kopp et al., 2012; Tsuda et al., 2018) (Figure 1.2). SOX9 is normally expressed in ductal cells, while adult acinar cells do not show the presence of SOX9 (Shroff et al., 2014). Para-doxically, SOX9 expression in *KRAS* mutation-induced ductal cells resulted in substantial reductions of IPMNs formation (Von Figura et al., 2014; Roy et al., 2015). Furthermore, the progenitor niche within the pancreatic ductal epithelium, under *KRAS* mutations, and in conjunction with additional aberrations can also drive PDAC precursor lesion formation, adding yet another layer of possible cell-of-origin of PDACs (Yamaguchi et al., 2016).

Recently, methylation analysis in combination with RNA sequencing (RNAseq) expression profiling identified two subgroups of PDACs, Methylation^{low}/IFNsign(Interferon signature)^{high} and Methylation^{high}/IFNsign^{low} harboring lineage specific traits of acinar and ductal cells, respectively, further corroborating the aforementioned studies that PDAC in fact have two or more distinct cell-of-origin (Espinet et al., 2020).
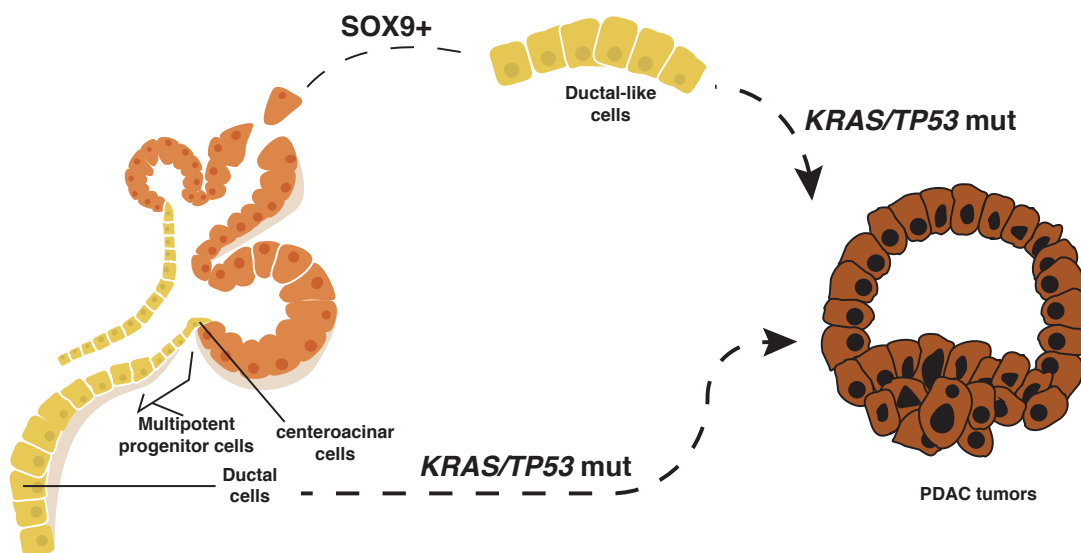
FIGURE 1.2: **Progression model of PDAC formation from acinar cell.** Acinar cells upregulate SOX9 to transition towards a ductal state, thereby, sensitizing them for *KRAS* mutation. Subse-quent aberrations in addition to *KRAS* mutation, including *TP53* mutations are then required to form the PDAC tumor. Adapted from (Kopp et al., 2012)

1.4 Therapy and PanNEN tumors

1.4.1 Current therapeutic approaches for PanNEN tumors

The first line of medical intervention for PanNENs for a given patient is highly dependent on the extent of the tumor progression and hinges on multiple criteria. Patients harboring localized well-differentiated tumors undergo surgery and are reported to have a median survival of 16 years, which dramatically declines to <1 year if the tumor is metastatic and of high grade without a possibility of surgical intervention (Brooks, Shavelle, and Vavra-Musser, 2019). If the tumor has reached the stage of widespread liver metastasis, and therefore cannot be surgically removed, the patient receives radio/chemo-based systemic therapy, which of course carries in itself many side effects (Kulke et al., 2011; Metz and Jensen, 2008). Specifically, the first line of chemotherapy for PanNECs involves a combination of Cisplatin and Etoposide, which have shown marginal anti-tumor activity with major side effects including myelosuppression and gastrointestinal toxicities (Iwasa et al., 2010). The overall response rate of this combination is only 14% with a median progression-free survival of 1.8 months and median overall survival of 5.8 months. In terms of chemotherapy for PanNETs, the patient is given a combination of 5-Flurouracil, with or without Doxorubicin (Pavel et al., 2012). Similar to the poorly differentiated metastasis treatment response, median response is only 6-20 month and for cases with extensive liver involvement, the progression-free survival was further reduced (Kouvaraki et al., 2004). In spite of these treatment interventions, many patients remain unresponsive and require additional novel approaches.

Aggressive and invasive treatments are currently being replaced by targeted therapy approaches in many cancer entities. Limited targeted therapies are in fact available and administered to PanNEN patients on the identification of adequate targets. Yet, the current standing of available targeted therapy in general for PanNEN treatment requires substantial improvement. Sunitinib, a multi-tyrosine kinase receptor inhibitor, was one of the first FDA approved targeted therapy drugs in treating PanNENs. Upregulation of these receptors and its potential towards invasive nature of malignant PanNENs has been established by multiple studies (Casanovas et al., 2005; Fjällskog et al., 2003; Hansel et al., 2003). The second major targeted therapy that is much more prominent in the clinic is Everolimus, which is utilized to inhibit the aberrant mTOR pathway (Kulke et al., 2011; Metz and Jensen, 2008). Deregulation of the pathway has been identified at the genetic and expression level in PanNENs (Jiao et al., 2011; Missiaglia et al., 2010). Multiple clinical trials have established the overall benefit of Everolimus in treating low/intermediate grade tumors that were unresectable, metastasized or were unresponsive to other therapies (Pavel et al., 2011; Yao et al., 2010; Yao et al., 2008; Yao et al., 2011). The treatment indeed increased progression-free survival, resulting in stabilization of the disease with reduced proliferation. Nevertheless,

resistance and relapse were clearly evident in patients and the established therapies have yet to increase the overall survival or result in complete remission, highlighting the importance of advancing the currently limited treatment options.

1.4.2 *In-vtiro* Models for the identification of novel therapeutic approaches for PanNEN tumors

In clinical translational research, many of the insights from tumor cohorts are advanced further in *in-vitro* studies to determine potential targetable anomalies and to identify how they affect tumor progression. It is imperative that we envision this line of approach in PanNENs due to the clear lack of targeted therapy. Although very limited, two *in-vitro* cell line models, BON1 and QGP1, specific for PanNEN tumors have been established to do so. BON1 is a cell line derived from a lymph node metastasis of a PanNET patient with an initial doubling time of 13 days, which has shown a drastic reduction to 2.5 to 1.5 days after years in culture (Evers et al., 1991; Evers et al., 1994; Parekh et al., 1994; Benten et al., 2018). BON1, being not a clonal cell line, displays different morphological cell types, wherein, there are large, rounded cells as well as small cells with dendrite-like extensions, each histological feature being similar to the original tumor (Parekh et al., 1994). The cell line harbors somatostatin and gastrin receptors and also stains for many neuroendocrine markers including serotonin, pancreastatin, bombesin, synaptophysin and Chromogranin A (CgA), further affirming their neuroendocrine origin (Parekh et al., 1994; Hofving et al., 2018). BON1 cells also produce Transforming Growth Factor Beta (TGF-β) 1-3 and Fibroblast Growth Factor (FGF), which can impact the production of serotonin in an autocrine manner (Daniel Beauchamp and Coffey, 1991; Townsend, Ishizuka, and Thompson, 1993). QGP1, with a doubling time of 1.5 days, was generated from a primary PanNEC tumor expressing CarcinoEmbryonic Antigen (CEA) and somatostatin (Drewinko et al., 1976; Kuroki et al., 1984; Benten et al., 2018; Iguchi, Hayashi, and Kono, 1990). However, recent work reported the absence of somatostatin receptor in QGP1, further questioning the integrity of the current line that has undergone more than 20 decades of *in-vitro* culturing. QGP1 does not express CgA but shows the presence of synaptophysin neuroendocrine marker (Benten et al., 2018). In spite of the questionable features, BON1 and QGP1 have a high proliferative index (80.6%+/- 3.3% and 82.6% +/- 1.0%, respectively), suggesting they could be representative of NETG3 and NEC tumors (Benten et al., 2018). However, one needs to take precaution when utilizing them as models to study PanNENs as their genetic representation has been widely debated in the field.

Genetic representation of cell line models

Over the years, the molecular representation of cell lines to PanNEN tumors have come under serious scrutiny based on two publications that performed exome sequencing of these

cell lines. Mutational profiling in fact showed BON1 and QGP1 are void of mutations in recurrently altered genes of PanNET tumors such as *MEN1*, *DAXX*, and *ATRX* (Hofving et al., 2018; Boora et al., 2015; Vandamme et al., 2015). Nevertheless, BON1 and QGP1 both harbor alterations pertaining to NEC tumors including alterations in *TP53* present in both cell lines and *KRAS* as well as Neuroblastoma RAS Viral Oncogene Homolog *(NRAS)* mutations in QGP1 and BON1, respectively. Mutations in mTOR gene combined with LOH have been identified in QGP1, suggesting potential deregulation of the mTOR pathway in this particular cell line (Hofving et al., 2018). Additional alterations identified included genes associated with chromatin remodeling and DNA damage, which are pathways known to be altered in PanNETs. However, from a sequencing point of view, the alterations identified in these cell lines need to be taken with a grain of salt; in fact, re-sequencing of the key variants in each batch must be considered before performing any experiments as many of the variants reported as "mutations"in aforementioned studies were benign Single Nucelotide Polymorphisms (SNPs).

CNA analysis across multiple studies, however, were in agreement that BON1 and QGP1 harbor triploid and tetraploid genomes, respectively (Hofving et al., 2018; Lopez et al., 2010; Vandamme et al., 2015). The specific copy number changes varied slightly depending on how sensitive the methodology was in each of the studies. BON1 showed whole chromosomal gains (copy of 3) in chromosomes 2,5,7,10,12 and regional amplifications of chromosomes (>3 copies) 5,10,14,16, and 20, while long stretches of LOH were present in chromosomes 1,3,4,6,11,17,18,22 and X (Hofving et al., 2018; Vandamme et al., 2015). The region harboring *CDKN2A/B* locus in chromosome 9 was completely lost in BON1 (Lopez et al., 2010; Hofving et al., 2018; Vandamme et al., 2015). With respect to QGP1, regional amplifications were evident in almost all chromosomes, while LOH of nearly the entire chromosome was present in chromosomes 7,8,9,13,17,21,22 and X. CNA of BON1 are consistent with copy number changes identified in PanNET tumors, further affirming integrity of BON1 as a representative model for PanNET tumor studies (Scarpa et al., 2017; Lawrence et al., 2018). Since QGP1 is a PanNEC derived cell line and studies have yet to identify copy number aberrations specific to PanNECs, we lack the knowledge to establish QGP1 as a PanNEC representative model from the copy number changes.

Signaling networks in cell line models

Many of the approved drugs available for targeted therapy are directed towards the three main signaling networks significantly altered in cancer: MAPK and mTOR/PI3K signaling pathways. Various studies in multiple cancer entities have also shown that all three signaling axes are intertwined with one another. As such, both BON1 and QGP1 have been analyzed to determine how targeting some of the proteins involved would impact

their viability. This has been heavily investigated in the context of Everolimus treatment to determine how the drug affects, both directly and indirectly, the functions of these signaling pathways. Everolimus resistance has been evident in PanNEN patients and therefore serves as a means to identify potential mechanism. Multiple studies generated both BON1 and QGP1 Everolimus resistant daughter cell lines and further showed that the resistance can be overturned by a combination of Everolimus and PI3K or dual PI3K-mTOR inhibitors (Prada et al., 2018; Passacantilli et al., 2014). These findings were reaffirmed using different PI3K-AKT-mTOR inhibitors (Vandamme et al., 2016; Zitzmann et al., 2010). Mechanism of resistance in these cell lines have been attributed to various nodes including GSK3 expression and crosstalk to MAPK signaling via PI3K hyperactivity upon mTOR inhibition (Prada et al., 2018; Valentino et al., 2014; Zitzmann et al., 2010). However, evident from this handful of studies is that although we are aware of potential drugs that can hamper the proliferative activity of these cell lines, we clearly lack a comprehensive knowledge of how these signaling pathways interact with one another and what are the different crosstalks, activated from feedback/feedforward loops that one must consider when looking into therapeutical interventions suitable for PanNENs.

Chapter 2

Scope of this study

It is certainly worth appreciating the extensive research that has been undertaken to acknowledge and even comprehend the sheer extent of heterogeneity encompassing PanNEN tumors. In particular, PanNET tumors have been intensively investigated to identify the diverse molecular features characterizing the histopathological heterogeneity, subsequently demonstrating its association towards a distinctive prognosis among patients. Recent findings have also shed light on the potential diverse cell-of-origin of PanNET tumors. On the other hand, we are far from attaining a consensus characterization of PanNEN recapitulating the multi-layer heterogeneity. Specifically, the lack of our understanding in regard to the PanNEC tumor subgroup is evidently emphasized in chapter 1. Genetic and epigenetic aberrations unique to PanNEC tumors have rarely been explored. In addition, there has yet to be a study that identifies or even speculates on the cell-of-origin of the most aggressive fraction of PanNEN tumors. As such, it is imperative that these gaps are filled to advance our perception of the mechanisms driving a particular normal cell to a tumor state of PanNEN tumors. As part of my thesis, I provide a multi-genomic and epigenomic characterization of PanNENs, distinguishing the well-differentiated PanNETs from the poorly differentiated PanNECs. I further characterize the bona fide *in-vitro* models of PanNENs, BON1 and QGP1, at the multi-genomic and signaling axes, priming for future work in determining potential novel therapeutic approaches.

The first part constitutes DNA methylome-based characterization of PanNENs. As a first step, DNA methylation based PanNEN subgroups are classified and the broader methylation features distinguishing the major subgroups are explored.

The second part involves characterization of PanNEN subgroups at the genomic level. Targeted sequencing using commercially available in addition to an in-house PanNET gene specific panel (PanNEN panel) is performed in order to identify recurrent aberrations defining the PanNEN subgroups. The genetic profile is further expanded by identifying CNA of PanNEN subgroups and pioneer in identifying significantly aberrated focal somatic CNA (SCNA) unique to PanNEC specific subgroup.

The third part is comprised of the identification of the tumor cell-of-origin distinctive of the PanNEN subgroups. DNA methylation analysis is emerging as an important platform

that can be exploited to trace back the potential cell-of-origin in cancer. Although tumor and normal total methylomes can differ significantly, several studies have revealed the presence of the remnant footprint of the normal precursors in various types of cancer (Kulis et al., 2012; Oakes et al., 2016; Bormann et al., 2018; Rodríguez-Paredes et al., 2018). As such, three different approaches are employed using tumor and normal DNA methylomes to arrive at a consensus on the potential normal cell precursors of each subgroup. Firstly, utilizing established pancreatic cell type markers, the hypo- and hypermethylation status of marker genes are determined to deduce the representative pattern of each subgroup. The cell marker profile is further expanded by examining the protein expression of normal cell markers in representative sets of tumors from each subgroup. Additionally, DMP based phylo-epigenetic reconstruction is employed to decipher the similarity of the subgroups to the well-known normal pancreatic cell types. Finally, using an algorithm implementing non-negative least squares, normal cell type signatures are deconvoluted in each sample to precisely define the cell of origin for each of the PanNEN subgroups.

The last section deviates from the tumor cohort and rather examines the *in-vitro* models BON1 and QGP1 as potential representatives of PanNEN tumors. The cell line models are investigated for their genomic and methylome based resemblance to the subgroups, as well as the normal pancreatic cell type. Utilizing a mathematical modelling approach, termed modular response analysis (MRA), three cancer associated signaling pathways within BON1 and QGP1 are modeled and evaluated to determine the major differences contributing to their unique behavior as *in-vitro* models.

Chapter 3

Results

3.1 DNA methylation classification identifies a distinct PanNEC subgroup from PanNEN samples

Methylation based tumor classification has demonstrated its potential superior role in identifying tumors of similar origin in some cancer entities compared to conventional histopathological diagnosis. This had been critically evaluated and established in the case of central nervous system (CNS) tumors that carry highly diverse tumor entities and in general lack a standard diagnostic process. Using a comprehensive machine-learning approach, recent work has shown that the CNS tumors, in addition to other non-CNS tumors, can be distinguished and classified into subclasses, further underlining the role of the DNA methylome as an objective and concise diagnostic tool compared to or in addition to traditional pathology (Capper et al., 2018). The strength of methylation marks in precisely classifying tumors and identifying its origin has sparked the experimentation of this approach in identifying clinically relevant subclasses within a tumor entity. This has been demonstrated in colorectal cancer wherein DNA methylation subtypes consistently revealed distinct cell-of-origin features in addition to their prognostic differences (Bormann et al., 2018). Furthermore, as detailed in chapter 1, such measures have been undertaken for well-differentiated PanNETs to identify not only cell-of-origin, but also its association with varying survival rate between the subgroups. The lack of general methylome studies of PanNECs, especially investigating their similarity or dissimilarity to PanNETs, urged us to take an approach analogous to the aforementioned studies.

In order to determine whether histopathological distinction of PanNETs and PanNECs are reflected in the methylation profiles, we initially approached the problem by identifying methylation subgroups within our cohort. We utilized an unsupervised class discovery methodology implemented in a consensus clustering (CC) algorithm. The approach allows for the identification of subgroups within a given dataset based on quantitative evidence and visual stability of the groups upon repeated subsampling and clustering (Wilkerson and Hayes, 2010). The subsampling occurs at the item (samples) and feature (methylation

probes) level. True clusters with a real structure would therefore be robust to sampling vari-
ability. The process essentially iterates on the most informative probes. We first determined
the 10,000 most variable probes (10K) based on standard deviation across the sample set of
57 tumors and then computed the distinct groups. Based on the consensus values, which
correspond to the probability that the sample being assigned to a particular group indeed
belongs to that group, and subsequent hierarchical clustering, we identified that 3 sub-
groups (k=3) stably represented the cohort: Group A, Group B and Group C (Figure 3.1a).
The empirical cumulative distribution function (CDF) of the consensus matrix depicted in
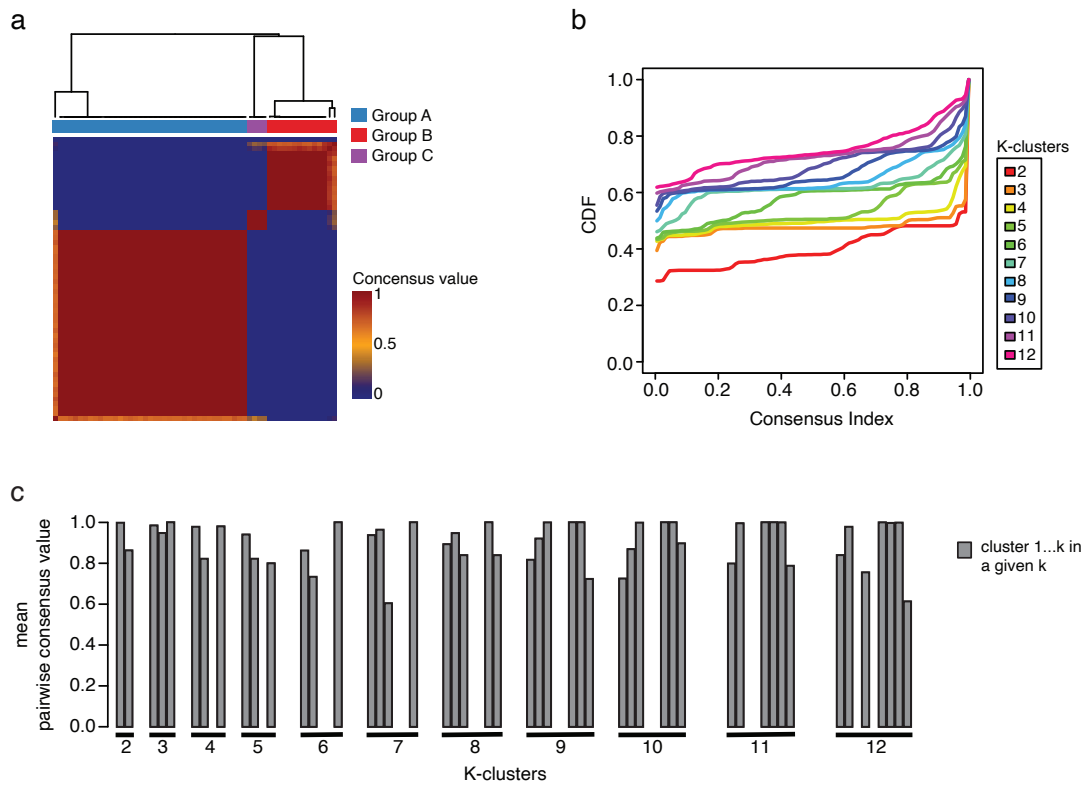
FIGURE 3.1: **PanNEN subgroup discovery.** (a) Unsupervised class discovery using 10000 (10K)
most variable methylation probes. Heatmap displays pairwise consensus values of the samples.
(b) Cumulative distribution function (CDF) curve of the resulting 'k-mer' count for 12-k's. Cluster
count of 'three' resulted in the most stable number of subgroups. (c) mean pairwise consensus
values for each group in a given k-cluster (k: 1-12). Spaces with no bars within a given k-cluster
represents a mean value of 0.

Figure 3.1b further displays a maximum stability at k=3. The CDF represents the distribu-
tion of the consensus matrix within a given k. At k=3, we see the accumulation of consensus
values close to 0 and 1 within the CDF curve. More specifically, the first step of the CDF
curve is around 0, which represents the consensus values between sample pairs clustering

in different groups, a relatively flat line across the 0-1 range, and the second step occurring at a consensus index close to 1, depicting the consensus values for those sample pairs that always cluster together. The curve therefore affirms that at k=3 the samples within a given cluster strongly correlate to one another while maintaining strong dissimilarity to the samples of a different cluster with very few sample pairs showing ambiguous clustering (Figure 3.1b). This is further visualized with respect to the mean pair-wise consensus value for each cluster of k=2 through k=12 (Figure 3.1c). At k=3, all the clusters have relative similar and mean consensus values close to 1, further suggesting a strong within-cluster stability, while as k increases, we see that many of the cluster values are close to 0 (space with no bar), implicating the de-stability of cluster integrity under an additional increase in k.

As a methylation proxy for the downstream analysis, methylation beta values of the probes were investigated, which is computed as follows:

$$beta = \frac{M}{M+U} \tag{3.1}$$

where, M and U corresponds to the methylated and unmethylated signal intensity of a given probe. The beta values from the 10K most variable probes revealed notable distinction between the groups (Figure 3.2a). In particular, distinct histological classifications were enriched in the subgroups. The largest subgroup, Group A, was composed of 39 well-differentiated PanNETs of all grades including 10 NETG3, while Group B was comprised of 13 from a total of 14 NECs and one NETG2 sample. The methylation classification profile therefore also denotes a strict separation between the high-grade NETG3 and NEC tumors, portrayed by the separation of the tumor types into Group A and Group B respectively. Group C consisted of only four samples: two NETG3s, a primary tumor and metastasis of the same patient, one NETG2 and one NEC. In general, there was no enrichment of primary tumor or metastasis status within the subgroup, and they were equally distributed among Group A and Group B. The consistent segregation of the subgroups was further recapitulated in a t-distributed stochastic neighbor embedding (tSNE) analysis using our 10K probes, affirming the presence of distinct methylation patterns in the groups identified (Figure 3.2b).

As evident in Figure 4a, Group B exhibited hypermethylation of the majority of 10K probes across all samples, which led us to explore whether this was a global scale phenomenon (Figure 3.2a). Interestingly, the mean methylation of CpG island probes was significantly higher in Group B and Group C compared to Group A (Two-sample Wilcoxon Test: p-value = 0.0001911 between Group A and B and p = 0.03496 between Group A and C) (Figure 3.2c). Gene Ontology (GO) analysis of the genes associated with the 10K probes significantly enriched for top terms involved in organ development and specifically neurogenesis associated biological processes (Figure 3.2d). Recent work evaluating the transition of the

a

b

c

d

e

f
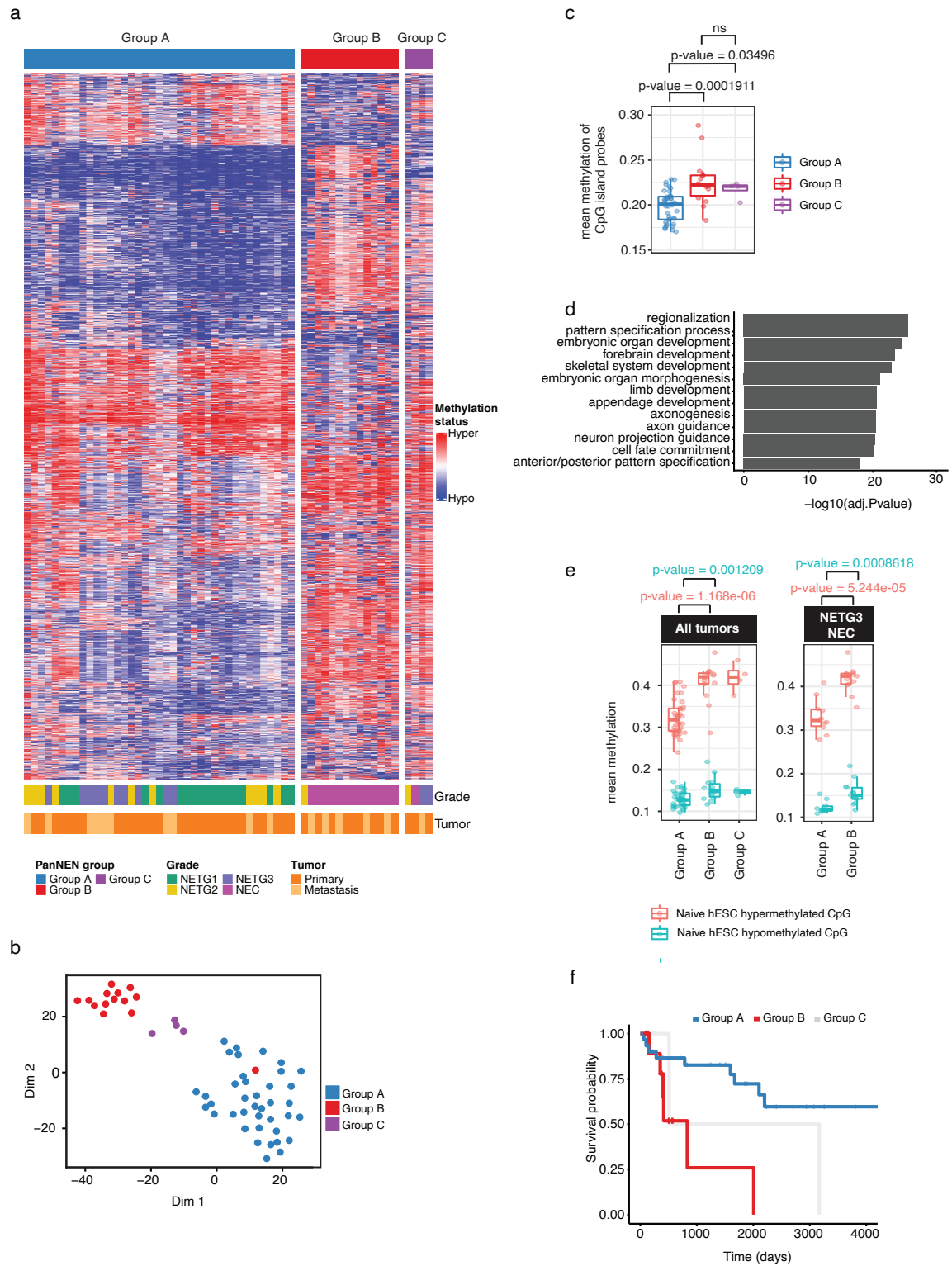
FIGURE 3.2: **Characterization of PanNEN methylome subroups.** (a) Heatmap displaying the methylation status of 10K variable probes in each of the Groups: A, B and C. Methylation beta value was used to perform hierarchical clustering separately on each subgroup, identifying highly similar samples. Hierarchical clustering was performed on the probes and within group samples; color range blue to red represents methylation beta value, columns indicate samples and rows methylation probes. (b) tSNE representation of PanNEN subgroups using 10K most variable probes. Subgroup annotation is used as colors for each sample (dot). (c) Mean methylation of CpG island probes in PanNEN subgroups. Boxplot represents distribution of mean methylation of CpG island associated probes. Each dot depicts the mean value of CpG island associated probes in a given sample. (d) GO ontology analysis of the 10K most variable probes, representing the top 12 GO terms based on -Log_{10}P-value. (e) Mean methylation of hESC associated hypermethylated and hypomethylated probes in PanNEN subgroups. Boxplot represents distribution of hypermethylated (red) and hypomethylated CpG probes of hESC (blue) in all samples (left panel), and only NETG3/NEC samples from Group A and B respectively (right panel). Each dot depicts the mean value of probes in each hESC category for a given sample. Statistical analysis was performed on the difference in distribution between subgroups using a two-sample Wilcoxon test. Box shows 25th and 75th percentiles and sample median as horizontal line, whiskers show maximum and minimum value that is 1.5 times the interquartile range over the 75th and 25th percentile, respectively. (f) Kaplan Meier survival graph from available patient data. Graph shows cohort according to methylation groups A, B and C. Significant difference was found between Group A and B (p = 0.0049) and between all groups (p = 0.011).

human primed Embryonic Stem Cells (hESC) to naïve state demonstrated a gradual and stable acquisition of CpG hypermethylation in genes associated with development which was mirrored in multiple cancer entities (Patani et al., 2020). Given the enrichment of GO developmental terms of the 10K probes and the significantly higher level of methylation of CpG island in Group B tumors, we explored if Group B also showed features reflective of naïve hESC. We first identified differentially methylated probes between primed and naïve hESCs from the Patani *et. al.* dataset and determined hypermethylated and hypomethylated CpGs of naïve hESCs (Patani et al., 2020). We found that the naïve hESC associated and hypermethylated CpGs were significantly more hypermethylated in Group B compared to Group A tumors (Figure 3.2e left panel). A closer analysis of NETG3 samples in Group A compared to NEC samples from Group B further displayed the maintenance of a similar significant difference in the distribution of hypermethylated CpGs of naïve hESCs (Figure 3.2e right panel), highlighting the potential difference in developmental state between these histologically often indistinguishable high-grade PanNEN tumor types. Taken together, we have identified three methylation groups, reflecting the histopathologically distinct PanNETs and PanNECs, and further showed that PanNECs harbor features of significantly more hypermethylation and mirror the epigenetic signature associated with naïve hESCs.

3.2   Distinct recurrent mutations separate PanNENs in Groups A and B

As described in chapter 1, the mutational profiling of PanNET and PanNEC tumors is indeed the most well-investigated molecular aberration. As such, the characterization of PanNEN methylome subgroups was expanded by exploring known and unknown mutational differences that exist between the tumor groups.

Targeted panel sequencing is a highly demanded platform for disease-specific mutation analysis. In fact, over the years, the technology has shown significant impact for discovery as well as validation of mutational aberrations in clinical translational genetics (Rehm, 2013). Multi-gene disease-targeted panel sequencing dominates over both whole genome and whole-exome sequencing with regard to cost, coverage, as well as the time and labor constraints associated with mutational analysis (Bewicke-Copley et al., 2019). Most importantly, panel sequencing remains the single high-throughput platform that can screen for mutations in highly degraded tissues sections of FFPE material. On these grounds, as part of my masters thesis, a novel gene panel, termed PanNEN panel, was designed in collaboration with the Department of Knowledge Management in Bioinformatics at Humboldt University, targeting well-known PanNET associated genes as well as additional genes relevant for therapeutic approaches (Figure 3.3; Table 6.1.1). This in-house PanNEN panel was utilized to identify mutations in 34 PanNEN samples, comprising all NETG1 and NETG2 samples, as well as 2 NETG3 and 1 NEC samples. The mutational screening of the remaining 10 NETG3 and 13 NEC was expanded by including a commercially available Comprehensive Cancer Panel (CCP) in order to screen for genes altered in these high-grade tumor classes that might be distinct from the conventional alterations of PanNET tumors and therefore, not covered by the in-house panel.

The mutation spectrum of the subgroups was limited to few recurrently aberrated genes, while single events within additional genes were evident in Group A and Group B. In total, 41 samples harbored 84 aberrations in 50 genes targeted by the PanNEN and CCP panels (Figure 3.4). These mutations comprised non-synonymous alterations of missense, deletion or indel variant types (Table 6.2.1). Additional sanger sequencing was performed on selected variants to further assert the reliability of our PanNEN in-house panel (Figure 3.5). Three aberrations were detected below 10% allelic frequency among which only a single *SETD2* mutation in PNET8 maintained an allelic frequency of < 6% (Figure 3.5). *MEN1* was the most recurrently mutated gene found within our cohort (25%) with high allelic frequency variants and confined to frameshift truncation, stop-gain and missense variants (Figure 3.4). Two of these *MEN1* mutations in MEN1 syndrome patients (PNET33 and PNET85) were obtained from clinical reports (Table 6.2.1). Mutations in *DAXX* and *ATRX*, genes identified as frequently altered in PanNEN, were also present in 11% and 7%
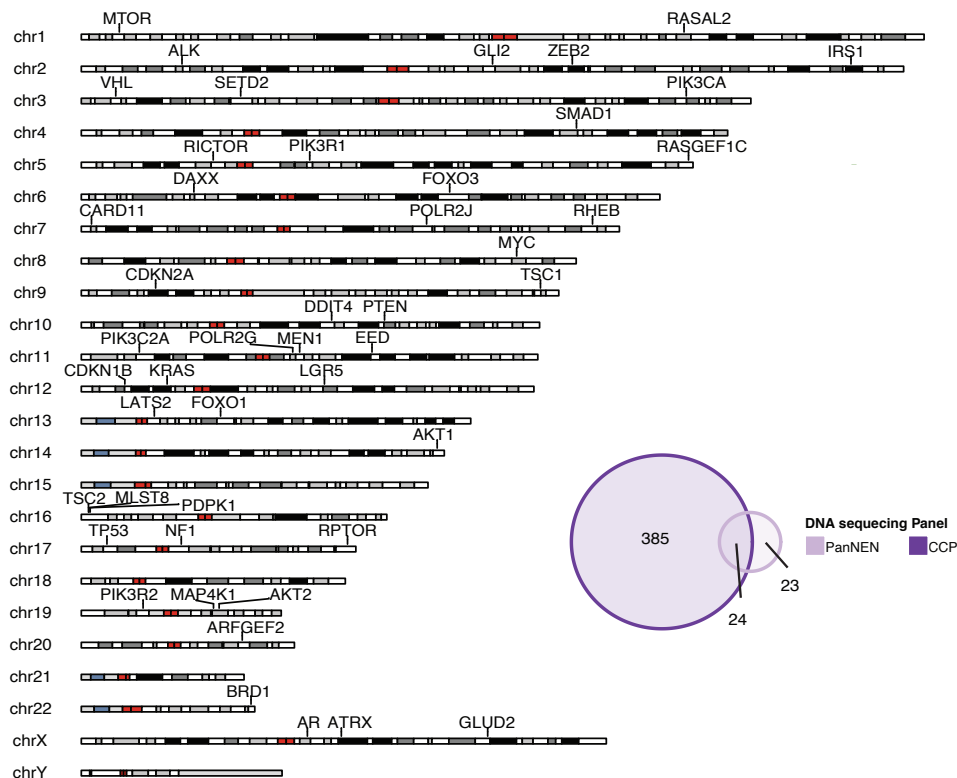
FIGURE 3.3: **PanNEN panel.** Custom PanNEN gene panel for PanNEN mutation analysis. Targeted genes of the PanNEN panel are displayed in their respective chromosomal region. The Venn diagram displays comparison of genes of the PanNEN gene panel to the commercially available CCP panel, which was used for mutational screening of NETG3 and NEC.

of the cohort, respectively. As previously shown, *DAXX*, *ATRX* alterations within the cohort were confirmed to be mutually exclusive (Figure 3.4). Mutations in *MEN1* and *DAXX* (PNET107, PNET79, PNET85) or *MEN1* and *ATRX* genes (PNET50, PNET11, PNET33) co-occurred in six samples and these classical PanNET aberrations were largely enriched in Group A tumors (Figure 3.4). In addition, other recurrent variants identified in Group A samples included aberrations in tumor suppressor gene *VHL* in four samples (PNET5, PNET14, PNET62 and PNET70) and mutations in *PTEN*, a PI3K pathway negative regulator, in two samples (PNET9, PNET108), both of which were COSMIC variants (Table 6.2.1). *VHL* mutations have yet to be identified in sporadic tumors as elaborated in chapter 1 and aberrations of this gene always occurred as part of the VHL hereditary syndrome. However, contrary to the current literature, the *VHL* variant associated with PNET14 was absent in the normal control and showed an allelic frequency of 74% (Figure 3.5; Table 6.2.1). The aberration was in fact a COSMIC variant, further asserting the existence of *VHL* mutations in sporadic PanNET tumors and their potential as driver alterations. The remaining three

FIGURE 3.4: **Mutational profile of the PanNEN cohort.** Identified mutations of targeted genes in
the PanNEN cohort. Barplot depicting mutational frequency in a given sample, colored according
to variant type (top panel); (left) percentage value per row: frequency at which the respective gene
is aberrated in the cohort; (middle panel) mutation variant of each patient, colored according to
variant type; white spacing: no mutation was identified in the regions of the targeted gene; (right)
two row annotations displaying whether the respective gene was covered by the PanNEN panel,
CCP panel or both; PanNEN subgroup, tumor grade and gene panel used for mutational analysis
for each sample is annotated at the bottom.

FIGURE 3.5: **PanNEN mutational variant analysis.** Allelic frequencies and verification status of PanNEN mutations. Mutated genes are displayed at the bottom; heatmap color represents the frequency at which the alterations were identified in a given sample (range from: 0.05 [blue], to 0.93 [red]). Grey: absence of aberration in the respective gene for a given sample.

VHL alterations in PNET5, PNET62 and PNET70 were also established COSMIC variants but, due to the lack of normal controls for these patients, whether their occurrences were sporadic or hereditary in nature was undetermined. Mutations in the major negative regulator of mTOR pathway, *TSC2* (n=4), were also enriched in Group A tumors. Although often associated with TSC2 hereditary syndrome, the variants identified in our cohort were all somatic mutations.

In contrast to the genes associated with Group A, recurring aberrations in *KRAS* (5/14),

SMAD4 (2/14) and *TP53* (3/14) established the mutational profile of Group B PanNEC tumors (Figure 3.4). The *KRAS* alterations identified were all COSMIC variants including G12D, G12R, G12V and Q61D aberrations (Table 6.2.1). *SMAD4* and *TP53* mutations were also COSMIC variants, additionally demonstrating the strong influence these mutations exert in driving Group B tumor progression. Mutations associated with classical PanNET genes *DAXX* and *ATRX* were absent in Group B; however, one sample each of Group B cohort was mutated in *MEN1* (PNET60) and *TSC2* (PNET2) (Figure 3.4; Table 6.2.1). The *TSC2* missense variant identified in PNET2 of Group B was confirmed to have low activity towards inhibiting mTORC1, thereby signifying the deregulation of mTOR pathway in this particular tumor sample (Hoogeveen-Westerveld et al., 2013).

Driver mutations in the targeted genes were not detected in 16 samples, and this included the single non-PanNEC sample in Group B (PNET4), all Group C samples and four Group A tumors (PNET89, PNET91, PNET92 and PNET95) (Figure 3.4). Additionally, multiple samples were collected for two patients, PNET77 and PNET56, both of which displayed interesting profiles. Patient PNET77 had tissues of primary tumor and liver metastasis (PNET77P and PNET77M: NETG3 samples of Group C) surgically removed two years apart. Neither sample displayed mutations picked up by our panel (Figure 3.4; Table 6.2.1). In contrast, patient PNET56 with two liver metastasis samples removed one year apart, PNET56P1 and P2 (both NETG3s in Group A), carried the same alterations in *TSC2* and *BRD3*, suggesting a single ancestor lineage for both tumors.

In addition to the recurrent genetic differences detected between Group A and Group B PanNENs, mutations within genes shared between the groups were also identified during our analysis (Figure 3.4). Single-event alterations, represented in both Group A and Group B, were found in chromatin-remodeling associated genes well-known to be aberrant in PanNENs, including *EP300*, *BRD3*, *AFF1*, *MLLT10*, *PSIP1*, *SETD2* and *KMT2C*. Three alterations targeting PI3K subunits, *PIK3C2B* and *PIK3CG* were found in both groups as well. These mutations suggest potential common mechanisms driving PanNEN development, irrespective of the differences between Group A and Group B methylome subgroups.

Collectively, cohort mutational profiles uncovered key molecular distinctions between the PanNEN methylome groups, which are enriched for aberrations in *MEN1*, *DAXX* and *ATRX* in Group A and for *KRAS*, *TP53* and *SMAD4* in Group B tumors.

3.3   Copy number alterations separate PanNECs from PanNETs

PanNET CNA have been widely established in various studies, as illustrated in chapter 1. PanNECs, however, have rarely been investigated with respect to their whole chromosomal or focal aberrations in the genome. The development of genome-wide methylation array platforms has provided the opportunity to interrogate the CNA, with the same sensitivity

as a SNP array (Feber et al., 2014; Houseman et al., 2009). Due to high-density coverage of the genome, one can readily determine the whole genome and small focal aberrations in addition to the methylation profile within the same sample in a cost-effective manner. Numerous tools have thus far been developed to accurately determine large and small CNA from the methylation signal intensities (*conumeeR*; Mermel et al., 2011). We expanded our characterization to identify unique copy number features utilizing the aforementioned methods in order to discover additional molecular mechanisms distinctively driving the formation of these subgroups.

To explore CNAs in PanNENs, we first inferred methylation $\log_2$ ratios from signal intensities for each genomic segment that carry the same copy number status. For analysis of whole chromosomal copy number changes within each group, we then computed the average $\log_2$ ratio of CNA segments per chromosome and labeled "gain"as a mean $\log_2$ ratio > 0.15 and "loss"as mean $\log_2$ ratio < -0.15 based on the correlation estimation between the $\log_2$ ratio and fluorescence in-situ hybridization (FISH) validation (Figure 3.7c). Subsequently, we performed unsupervised clustering on the mean $\log_2$ ratios in order to determine recurrently aberrated chromosomes as well as subgroup-associated signatures.

The Group A tumors displayed three distinct whole chromosomal CNA profiles: amplification-rich, deletion-rich and low-CNA signatures (Figure 3.6). The amplification-rich signature predominantly harbored copy number gains in chromosomes 14, 5, 7, 19, 18, 17, 13, 20, 12, 4 and 9. Many of these chromosomal gains were synchronous events in the majority of the amplification-rich signature samples, suggesting a concurrent occurrence of these gains. Similarly, samples in the deletion-rich signature carried recurrent deletions of chromosomes 6, 1, 22, 8, 1, 2, 16, 10, 3, and 21, which also presented as synchronous events among many of these chromosomes.

Selective CNA features were further validated in 18 samples using FISH. Specifically, the corresponding whole chromosomal aberrations of chromosomes 5, 9 and 11 were identified using FISH probes targeting the individual centromeres as well as *RICTOR*, *TGFBR1* and *MEN1* genes located in the chromosomes, respectively (Figure 3.7 a and b). Upon analyzing 40 cells per sample, it was evident that most of the chromosomal aberrations observed in chromosome 5 and chromosome 9 involves the gain of an extra copy. Furthermore, copy number losses observed in chromosomes 11 resulted in a single copy loss. Evidently, prediction of whole chromosomal aberrations using mean CNA segments caused misidentification in a few cases (eg: chr5; PNET42 and chr9; PNET91); nevertheless, linear regression analysis of the mean signal count of the centromere probe per sample from FISH versus the mean $\log_2$ ratio from CNA analysis showed a regression coefficient of R^2 =0.6531 and p=6.243 x 10^{-7}, highlighting that our methodology is indeed robust to precisely classify whole chromosomal changes with very few false positives (Figure 3.7c).

We further integrated the identified recurrent mutations and the whole chromosomal CNA

FIGURE 3.6: **Whole chromosomal aberrations in PanNEN subgroups.** Hierarchical clustering of mean log$_2$ ratios of chromosomal segments per autosome is displayed in a heatmap; dotted line represent the cut off used to identify amplification, low-CNA and deletion-rich signatures; column annotations (bottom) display the tumor grade, tumor type, the recurrently aberrated genes as well as FISH verified samples. Clustering was performed within-group and the copy number signature cut-off are determined based on the main branches.

profiles to determine potential co-occurrence of chromosomal aberrations and mutations. Recurrent mutations of tumor suppressor genes *MEN1*, *DAXX*, *TSC2*, and *VHL* were enriched in the amplification-rich and deletion-rich signatures (Figure 3.6, 'Mutation' annotation). The majority of *MEN1* mutations (8/13) were enriched in amplification-rich signature with no notable alteration to the respective chromosome, while 4/5 of the remaining mutations present in deletion-rich and low-CNA signatures were accompanied by chromosomal loss of chromosome 11. In addition, mutations in *DAXX* and *TSC2* within the deletion-rich signature and *VHL* mutations in the low-CNA signature also displayed chromosomal losses
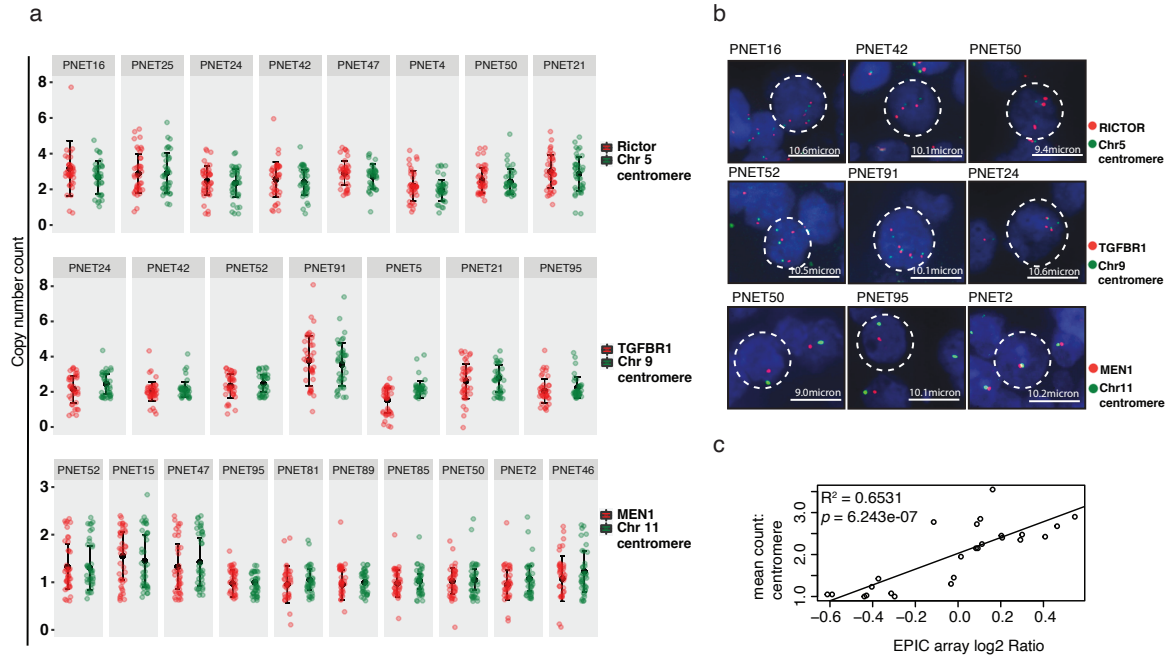
FIGURE 3.7: **FISH verification of whole chromosomal aberrations in the PanNEN cohort.** (a) Quantification of centromere and gene probes. Quantification of chromosomal aberrations for chromosomes 5, 9 and 11 was performed using their corresponding centromere probes (green) and probes in RICTOR, TGFBR1 and MEN1 respectively (red) using FISH. Forty cells randomly selected were counted in each sample; The distribution of signals per sample is depicted; each point represents the mean count of cells for the respective probe; mean and standard deviation represented by error bars. (b) Representative images of FISH validation of whole chromosomal aberrations in PanNEN subgroups. The green dot represents the centromere of chr5 (top panel), chr9 (middle panel) and chr11 (bottom panel). (c) Correlation of $\log_2$ ratio and FISH count. Linear regression of mean copy number count of centromere derived from FISH (y-axis) and mean $\log_2$ ratios of chromosomal segments per autosome (x-axis); diagonal line represents best fit model for the data.

of chromosome 6, 16 and 3, respectively, together highlighting a consistent two-hit tumor suppressor inactivation phenotype taking place particularly in low-CNA and deletion-rich signatures. With respect to the low and intermediate PanNETs, the low-CNA signature was a predominant feature of NETG1 tumors of Group A and contained relatively few aberrations in most chromosomes with no clear recurrences.

Consistent unique recurrent whole chromosomal signatures were absent in Group B and Group C tumors; nonetheless, the noteworthy features included recurrent copy number loss in chromosome 10 (50%) and chromosome 22 (57%), which were both aberrated in at least half of the samples of Group B (Figure 3.6). This is of course not a distinctive feature of Group B as loss of chromosomes 10 and 22 were among the features of the deletion-rich

signature of Group A. Therefore, tumors of Group B and Group C were defined by few recurrent whole chromosomal aberrations yet not distinct to the groups, likely due to the low number of samples in each of the subgroups.

The approach utilized for the identification of copy number gains and losses would indeed mask the subtle, yet strong changes of high-copy amplifications and complete deletions observed in specific regions of the genome. These focal changes can reveal critical information on the processes inevitable for tumor progression, as often, these CNA aberrations target vital genes instrumental for such mechanisms. Using the log$_2$ ratios of CNA segments as computed in our initial analysis, we investigated focal CNA using GISTIC (Figure 3.8a). Chromosome regional amplification of 1q21.2, 8p23.1 and 14q11.2, as well as deletion of chromosomal region 15q11.2 were significantly associated with both Group A and Group B (Table 6.3.1; Table 6.4.1; Table 6.5.1; Table 6.6.1). These aberrations were enriched in copy number variations of genes in multiple gene families including *DEFB1*, *OR*, *GOLG* and the *FAM90* gene family.

Distinct focal aberrations significantly associated with each group were also clearly evident (Figure 3.8a). Group A had unique significant gains in chromosomal regions 10q11.22 and 16p11.2, which predominantly affected pseudogenes (Table 6.3.1). Nevertheless, an interesting gene relevant for PanNEN within 10q11.22 locus was pancreatic polypeptide receptor 1 *(PPYR1)*. The coded protein has a high affinity to PP, the hormone produced by PP-cells and lower affinity to NPY and PYY proteins (Mannon, 2004). Additional genes also affected by this amplification included the subunits of WASH complex, *FAM21B* and *FAM21C*, and endocytosis associated genes *AGAP4* and *AGAP9*.

Interestingly, Group A had a substantial number of focal deletions associated with it in regions of 1p36.32, 2q37.3, 4q34.3, 6p25.3, 8p23.2, 9p21.3, 10q26.3, 12p11.1, 21q11.2, 22q12.3 and 22q13.32. In total, these deletions affected 73 genes comprised of protein-coding, miRNA genes as well as pseudogenes (Table 6.4.1). Among these, we identified genes associated with multiple signaling pathways that are known to be aberrated in various cancer types. For instance, deletion of 22q12.3, seen in 12 samples (31%), resulted in loss of *DEPDC5*, a major negative regulator of mTOR signaling (Padi et al., 2019). The deletion also entailed the loss of the *SLC5A1* gene, which codes for a direct glucose-galactose transporter (Hummel et al., 2011) as well as the *SLC5A4* gene encoding a mannose transporter (Tazawa et al., 2005), all potentially impinging on the deregulation of mTOR signaling. *DUSP22*, a major component of JNK pathway often downregulated in colorectal cancer and recently identified to regulate EGFR signaling in prostate cancer, was lost due to the 6p25.3 loss in Group A (Yu et al., 2015; Ha et al., 2019; Lin et al., 2019). In addition, the identified loss of 8p23.2 seen in 10 samples (21%) resulted in the deletion of *ARHGEF10*, a Rho guanine nucleotide exchange factor (GEF) (Mohl et al., 2006) known to act as a putative tumor suppressor in breast cancer, urothelial and hepatocellular carcinoma (Xue et al., 2012; Cooke
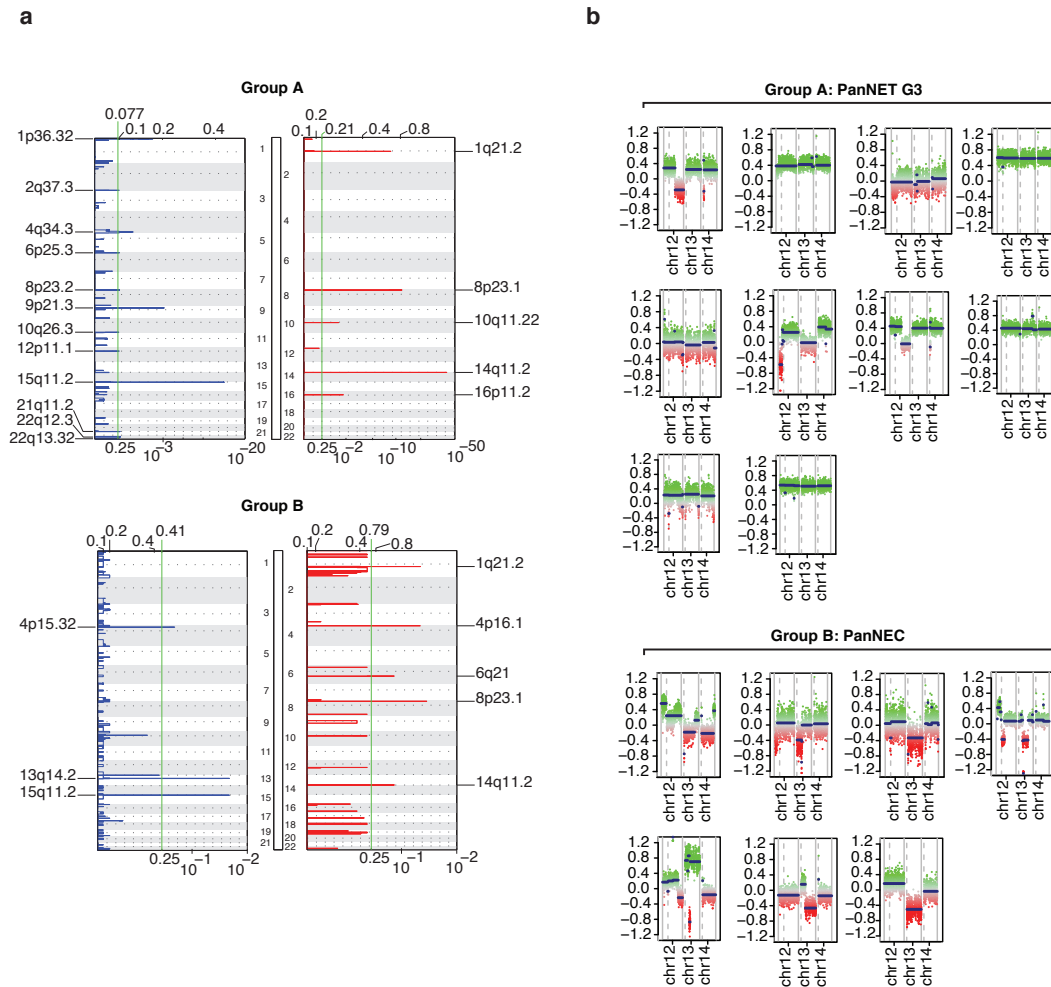
FIGURE 3.8: **Focal copy number aberrations in the PanNEN cohort.** (a)Significant focal copy number aberrations in the PanNEN subgroup. Focal aberrations in Group A (top panel) and Group B (bottom panel) are displayed here; focal copy number losses in blue horizontal lines and gains in red horizontal line. Log$_2$ ratio range is provided at the top and q-value is given at the bottom. Vertical green line represents the q-value cut off at 0.25 to call significant. The significantly aberrated focal regions are explicitly written (left and right). (b). Chromosome 12, 13, and 14 copy number status in NETG3 (top panel) and NEC (bottom panel). Intensity values of each bin are plotted in colored dots; each color indicates 'methylated' and 'unmethylated' channels of each CpG; segments are shown as horizontal blue lines.

et al., 2008; Riehn et al., 2011). The region also carries *MCPH1* coding for a multifunctional protein critical for cell cycle regulation and is often mutated or downregulated in cancer entities. Lastly, focal deletion of 9p21.3 in Group A lead to the loss of CDKN1B and CDKN2A, two key regulators of the cell cycle (Table 6.4.1).

In contrast to the focal aberrations intrinsic to Group A, Group B showed distinct focal gains

in 4p16.1 and 6q21 as well as focal deletions in 4p15.32 and 13q14.2 (Table 6.5.1; Table 6.6.1). The focal gain in the 4p16.1 region resulted in amplification of *DRD5*, a D5 subtype of dopamine receptor (Sidhu, 1998), homeobox protein *HMX1* which has transcriptional antagonistic function (Amendt, Sutherland, and Russo, 1999), and *SLC2A9*, a urate transporter (Ruiz et al., 2018). Gain of the 6q21 region showed paradoxically an amplification of *HACE1*, a tumor suppressor with ubiquitinate activity (Torrino et al., 2011). Interestingly, the deletion of 13q14.2 affecting 50% of Group B led to the deletion of *RB1*, an important cell cycle regulator with similar functional significance as *CDKN2A* and *CDKN2B* (Figure 3.8b upper panel). On closer examination, we found that NETG3s in Group A did not show losses of the *RB1* region (Figure 3.8b lower panel), and rather, the deletion was exclusive to the PanNECs of Group B.

Therefore, while Group A was enriched for recurrent whole chromosomal and many focal copy number aberrations, particularly deletions, Group B specifically showed few significant focal aberrations and exclusively harbored *RB1* loss associated with PanNEC tumors.

3.4   Tumor cell-of-origin characterization of PanNEN subgroups

During development, normal cellular hierarchy, and as a consequence the cell-type heterogeneity within an organ is often established and maintained through distinct transcriptional factor expression profiles. This has been extensively investigated and established in pancreas using mouse model studies. Among the prominent transcription factors critical for adult pancreatic cell type maintenance is the transcription factor PDX1 which is vital for β cell identity and ARX, upstream of IRX2, both of which are well-established α-cell-specific transcription factors (Yang et al., 2011; Dorrell et al., 2011; Muraro et al., 2016). In addition, presence of NKX6-1, NKX2-2 and PAX6 further maintains the endocrine lineage for α- and β-cells (Jennings et al., 2015). On the other hand, SOX9, a downstream target of NOTCH, is required for the establishment of cell fates of endocrine and exocrine cells (Rovira et al., 2010; Lynn et al., 2007). However, while SOX9 is absent in the committed endocrine precursors (Seymour et al., 2007), it is an important player in pancreatic ductal and centro-acinar cell development and maintanance, and is not found in adult acinar cells (Shroff et al., 2014; Qadir et al., 2020).

Recent advancements in single cell RNA sequencing platforms have generated a plethora of data that was able to clearly discern and reaffirm the presence of these transcription factors in the adult pancreatic cell types in order to uphold the tissue complexity. The underlying mechanisms leading to these differential and consistent expression profiles include DNA hypomethylation of cell-specific marker genes combined with enrichment for open chromatin and active histone marks. Interestingly, it has been demonstrated in multiple tissues that DNA methylation marks of a normal precursor is retained in the respective tumors and

can additionally determine profiles associated with tumors developing from different cells of a single tissue type (Bormann et al., 2018). The distinct characteristics of DNA methylation capable of being traced back to its normal precursor presents an opportunity to identify whether our PanNEN methylome subgroups that show genetic disparity are in fact tumors of distinct cell-of-origin.

We undertook three major approaches to define the normal precursor identity of our subgroups. First, we identified differentially methylated pancreatic cell marker genes between the PanNEN Group A and Group B and further characterized them at the protein level. In addition, utilizing published 450K array methylation profiles of α, β, ductal and acinar cells, particularly for the Group B tumors, we further expanded our knowledge on the underlying normal cell profiles defining these PanNEN subgroups and established novel cell markers that play an active role in how the tumors are formed.

3.4.1 PanNETs in Group A harbor endocrine cell-of-origin signatures

Cell-of-origin studies in PanNENs have been thus far restricted to well-differentiated PanNETs. Due to the unique classification grades enriched within each of the methylome subgroups from our analysis, we examined whether the imprinted cell-of-origin pattern persisting in the tumor entity is distinct for each subgroup. The first approach investigated the DNA methylation features of established cell marker genes known to influence the formation and maintenance of differentiated cells of the pancreas. We explored the PanglaoDB database (Franzén, Gan, and Björkegren, 2019), which collects published single cell data, to determine genes specifically expressed in a given cell type. We utilized DNA methylation as the surrogate for the expression status of each of the genes within our tumor samples. A total of 174 marker genes expressed in differentiated cell types of the adult pancreas were curated from PanglaoDB (Table 6.7.1). Upon identification of differentially methylated probes (DMPs) between Group A and Group B, we determined DMPs associated with genes overlapping with the curated list. In total, we detected 122 methylation probes of genes associated with cell-type specificity within the 770 DMPs from our samples. The average difference in beta value ($|\Delta beta|$) between the groups ranged from 0.15 to 0.30. To determine the probes and, thereby, the cell marker genes strongly distinguishing the groups, we restricted the $|\Delta beta|$ value to $|\Delta beta| > 0.25$. Based on the cut-off for the p-value threshold and fold difference ($-\log_{10}P > 5$ and $|\Delta beta| > 0.25$), we uncovered the genes and their associated probes that demonstrated the most significant differences between Group A and Group B (Figure 3.9a; Table 6.8.1). These included 85 significantly enriched probes (red points in Fig. 9a) associated with 23 marker genes of α, β, γ, δ, ductal, acinar and islet Schwann cells (Figure 3.9b). In general, irrespective of the cell type to which the expression profile of the marker genes aligned to, the vast majority of these probes maintained a stro-
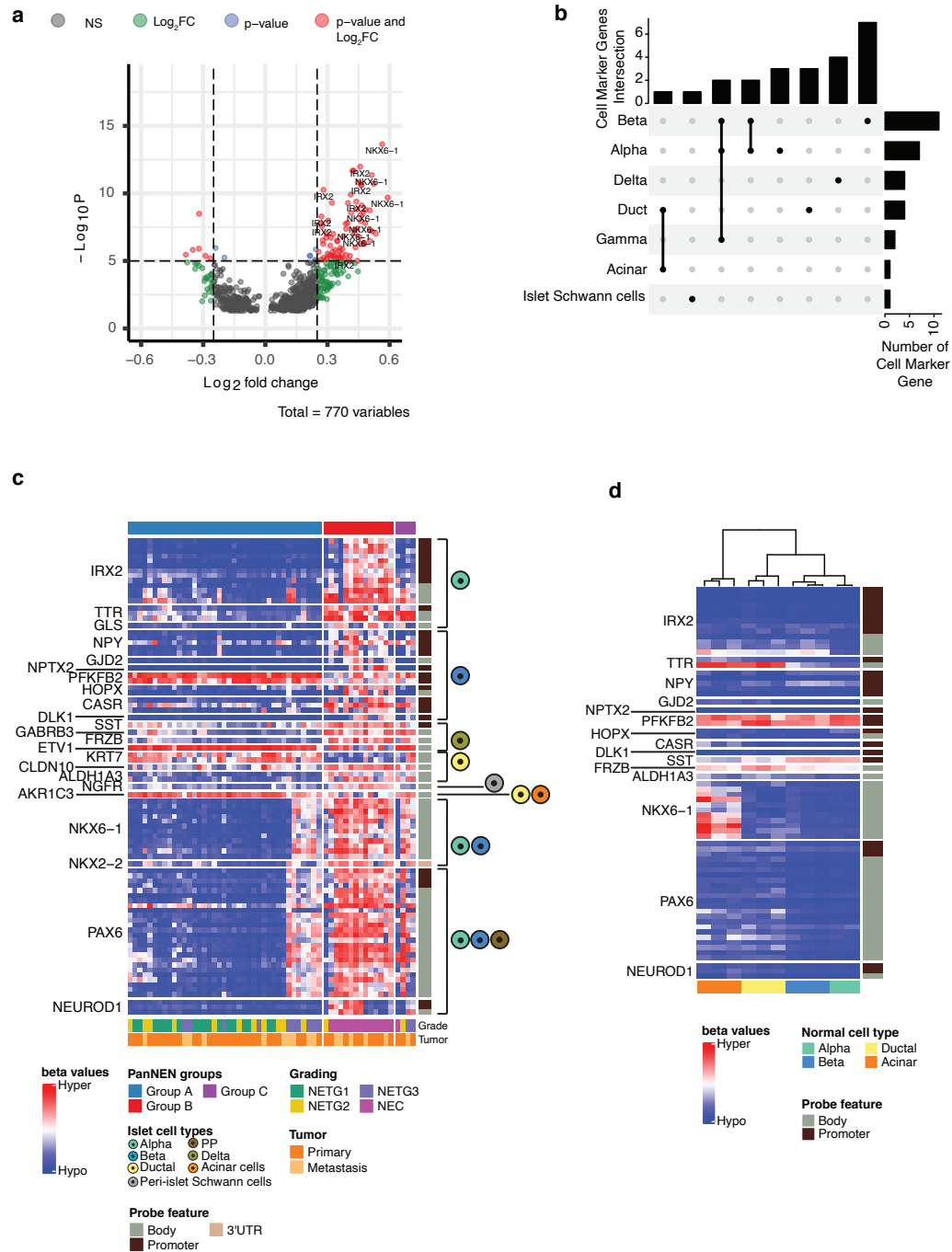
FIGURE 3.9: **Cell marker associated methylation analysis.** (a) Differentially Methylated Probes (DMPs) from our cohort associated with pancreatic cell markers (n=770). Dots represent the intersect between -Log$_{10}$P-value and the log$_2$ fold change (FC) for a given probe. Cut off for significant -Log$_{10}$P-value is 5 (adjusted p-value is 10e-6); cut off for significant log$_2$FC is > |0.25|. Red dots denote the probes that pass both cut offs; green dots represent probes that only surpass the log2FC threshold; blue dots show probes that only have a significant p-value; grey dots represent probes that did not pass either of the cut offs; significantly associated DMP probes of IRX2 and NKX6-1 are labeled in the plot. (b) Cell marker gene count her cell type. Number of marker genes associated with each cell type is represented in the bar plot (row bar plot); marker genes associated with more than one cell type (dots connected with line) or a single cell type are also depicted (column bar plot); the cell type of the respective associations are linked (bottom annotation). (c) Methylation beta values of significant DMPs belonging to pancreatic cell markers of PanNEN subgroups. Heatmap displaying the methylation beta values of DMPs (row) in each sample (column) of the subgroups. The heatmap color represents the methylation beta value; row annotation identifies the genomic region associated with the probe (panel on the right) in addition to which cell type the gene shows specificity in gene expression (according to PangaloDB). Column annotations (bottom) display the tumor grade and tumor type. (d) Methylation beta value of significant DMPs in normal cell type profiles. Heatmap displaying the methylation beta values of DMPs (row) that showed an overlap in normal cell type 450K methylation data. The heatmap color represents the methylation beta value; row annotation identifies the genomic region associated with the probe (panel on the right) in addition to which cell type the gene shows specificity in gene expression (according to PangaloDB); column annotates the normal cell type.

ng hypomethylation across samples of Group A, while Group B tumors showed hypermethylation of all DMPs under investigation, with the exception of *KRT7* associated probes (Figure 3.9c). In addition, genes associated with endocrine cell lineage maintenance, such as *PAX6*, *NKX6-1* and *NKX 2-2* were mainly hypomethylated in Group A, except for 6 NETG3 samples (PNET42, PNET57, PNET61, PNET56P1, PNET56P2 and PNET107), and one NETG2 sample (PNET24), which showed hypermethylation of *PAX6*, *NKX6-1* and *NKX2-2* genes reminiscent of Group B samples.

In order to identify whether the underlying cell marker genes' methylation marks are an indicative feature of a specific pancreatic cell type, we extracted overlapping DMPs from normal pancreatic cell types methylation profiles and compared them to the subgroup profiles. Illumina 450K array methylation profiles of presorted normal pancreatic cell types were obtained: α(n=2), β(n=3), acinar(n=3) and ductal(n=3) cells (Neiman et al., 2017; Espinet et al., 2021). Among the 85 significant DMPs associated with marker genes within the subgroups, 65 DMPs of 15 marker genes were identified in these normal cell type profiles due to the fact that the data was processed in the 450K array instead of the larger EPIC array (Figure 3.9d). Interestingly, neither of the groups accurately embodied any of the normal cell type marker profiles; the closest representation of methylation features of normal cell types, however, were evident in Group A. Hierarchical clustering of tumor samples in addition to the normal cell type using the aforementioned 65 probes, revealed that the cell marker

methylation in α, β, and ductal cells closely resembled Group A while the only NETG2 sample (PNET4) of Group B showed the closest representation to acinar cell marker methylation status (Figure 3.10a). The majority of the Group B tumors (10/14) formed their own cluster, distinct from Group A or Group C tumors. Remarkably, Group A NETG3 samples that showed hypermethylation of *PAX6*, *NKX6-1* and *NKX2-2* clustered independently of the remaining Group A samples, and resided in proximity to Group B samples. In fact, three Group B PanNECs clustered within this independent branch. Multiple probes associated with *IRX2* showed the most significant difference between Groups A and B (Figure 3.9a, red dots). Indeed, Group A tumors were consistently hypomethylated in *IRX2* promoter associated probes, while Group B displayed strong hypermethylation (Figure 3.9c). As detailed in chapter 1, previous work has utilized *IRX2* and *PDX1* acetylation status to characterize cell-of-origin subtypes in well-differentiated PanNETs. Although *PDX1* was not differentially methylated significantly between Group A and Group B, the 10K probes establishing the groups showed variable methylation patterns of this gene within the groups (Figure 3.10b). Incidentally, Group A tumors displayed within-group divergence with respect to *IRX2* and *PDX1* methylation status, whereby one subclass carried hypomethylation of both *IRX2* and *PDX1*, while the second subclass featured hypomethylation of *IRX2* and hypermethylation *PDX1*. The latter subclass was enriched for *MEN1*, *DAXX* and *ATRX* mutations, while recurring *VHL* mutations were seen in samples carrying hypomethylation of *IRX2* and *PDX1*. In addition to the methylation signature of these cell marker genes, presence of ARX protein, together with the PDX1, were previously applied to classify distinct diagnostic attributes pertaining to the subtypes of well-differentiated PanNETs. As such, given the significance of *IRX2* and *PDX1* methylation profiles in classifying our subgroups, we expanded our analysis to identify whether differences in ARX and PDX1 protein expression further characterized the subgroups. We investigated ARX and PDX1 protein status in 18 Group A samples and 12 Group B samples. The stained sections were evaluated for the presence and absence of the respective proteins by Dr. med. Armin Jarosch from the Institute of Tumor Pathology, at the Charite University of Medicine, Berlin, Germany and Prof. Dr. med. Aurel Perren from the Institute of Pathology at the Bern University, Bern, Switzerland (Table 6.14.1 includes pathologist diagnosis. Stainings of samples from Bern cohort was evaluated Dr. Perren). In agreement with the data and with previous studies, Group A tumors were ARX+, which indicated positive nuclear staining and PDX1-, or absence of PDX1, in 14 cases. It is to be noted that the positive staining was defined even if only a handful of cells in the tumor region were positive (min. 1%). This distinctive pattern of expression therefore suggests that at the protein level, Group A tumors portray an α-cell-like phenotype (Figure 3.11). The ARX-PDX1- pattern was also present in 2 samples of Group A tumors, while one sample each showed the ARX-PDX1+ (PNET51) and ARX+PDX1+ (PNET62) phenotype. Consistently, Group B tumors showed a negative pattern of PDX1
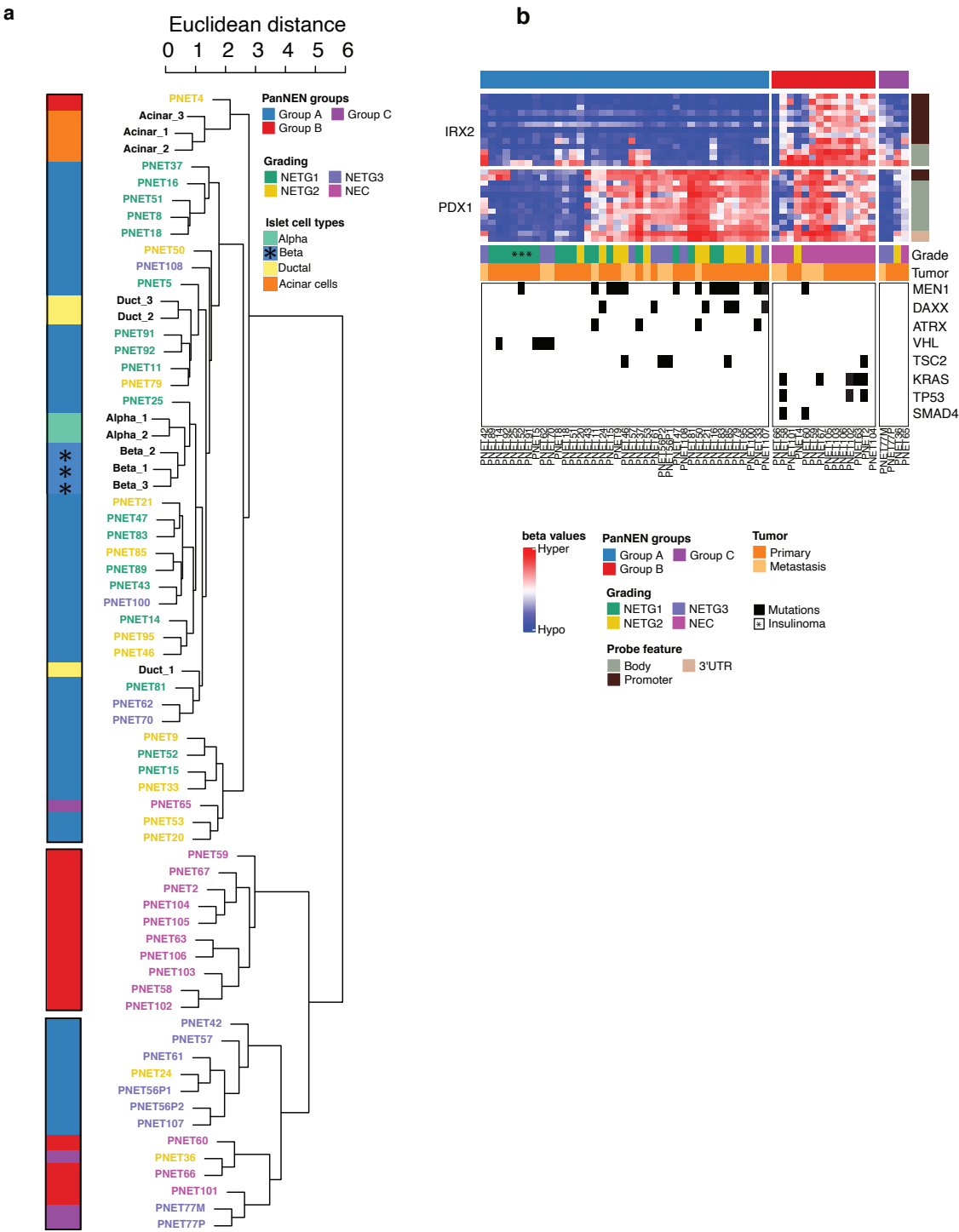
FIGURE 3.10: **Characterization of tumors using cell marker methylation profiles and α- β-cell features.** (a) Euclidean hierarchical clustering of cell marker methylation profiles that showed overlap with normal cell type 450K data. Sample names are annotated with grading color; bottom annotation depicts the methylome group and the normal cell type. Methylation profile of *IRX2* and *PDX1* probes. (b) Methylation beta values of probes associated with IRX2 and PDX1 of PanNEN subgroups. DMP probes of *IRX2* and 10K probes associated with *PDX1* (rows) for each sample (column) are presented here; rows are split at each associated gene (left); columns are split at each subgroup (bottom); row annotation identifies the genomic region associated with the probe (panel on the right); column annotation (bottom) displays the tumor grade, tumor type as well as the recurrently aberrated genes in the context of *IRX2* and *PDX1* methylation state.

protein in addition to either ARX-, displayed in eight samples, or ARX+ phenotype seen in four samples, respectively (Figure 3.12). Additionally, one sample showed the ARX+PDX1+ pattern (PNET58). Nevertheless, the ARX+PDX1- phenotype was significantly more common in Group A compared to Group B (p-value=0.02036: Fisher's exact test) reaffirming the α-like tumor feature in Group A samples[1].

3.4.2 PanNECs in Group B display an exocrine acinar-like cell signature and differ from PDACs in ductal-like cell signatures

The analysis approach taken in sub-section 3.4.1 clearly identified an endocrine-like phenotype for Group A tumors, Apart from eliminating a potential endocrine origin, a distinct normal cell associated signature had yet to be identified in Group B. Our analysis was therefore extended to explore the global differential methylation profiles of normal pancreatic cell types. A total of 46,500 DMPs were identified that distinctively differentiate the normal cell types from one another (adjusted P-value < 0.01, $|\Delta beta| > 0.2$). Based on the beta value of the DMP probes, Pearson distance between the tumor samples and the pancreatic cell types were computed and a phylo-epigenetic tree was constructed (Figure 3.13). The tree split into three main branches from the root (Figure 3.13 black dots); the lower branch consists entirely of Group A tumors, and clustered closely with β- and α-cell profiles. In fact, the insulinomas grouped together and maintained the closest distance to the normal endocrine cell type. The second main branch also harbored Group A tumors but exclusively the NETG3s and 2 NETG2 tumors. The final main branch had multiple splits: the first split separating PNET16 from the remaining samples and a subsequent split resulted in yet another Group A tumor clad comprising all three PanNET grades. This split also lead to the main branch harboring Group B tumors. Strikingly, all tumors from Group B, except for PNET58, clustered with ductal and acinar cells, constructing a separate and compact clad. PNET58 was ARX+ and PDX1+ (Figure 3.12), clearly different from the ARX-PDX1- profile

---

[1]The statistics are computed with only samples carrying all three staining for consistency purposes. In the coming sections, significant enrichment of SOX9 staining patterns are evaluated in regard to ARX and PDX1.
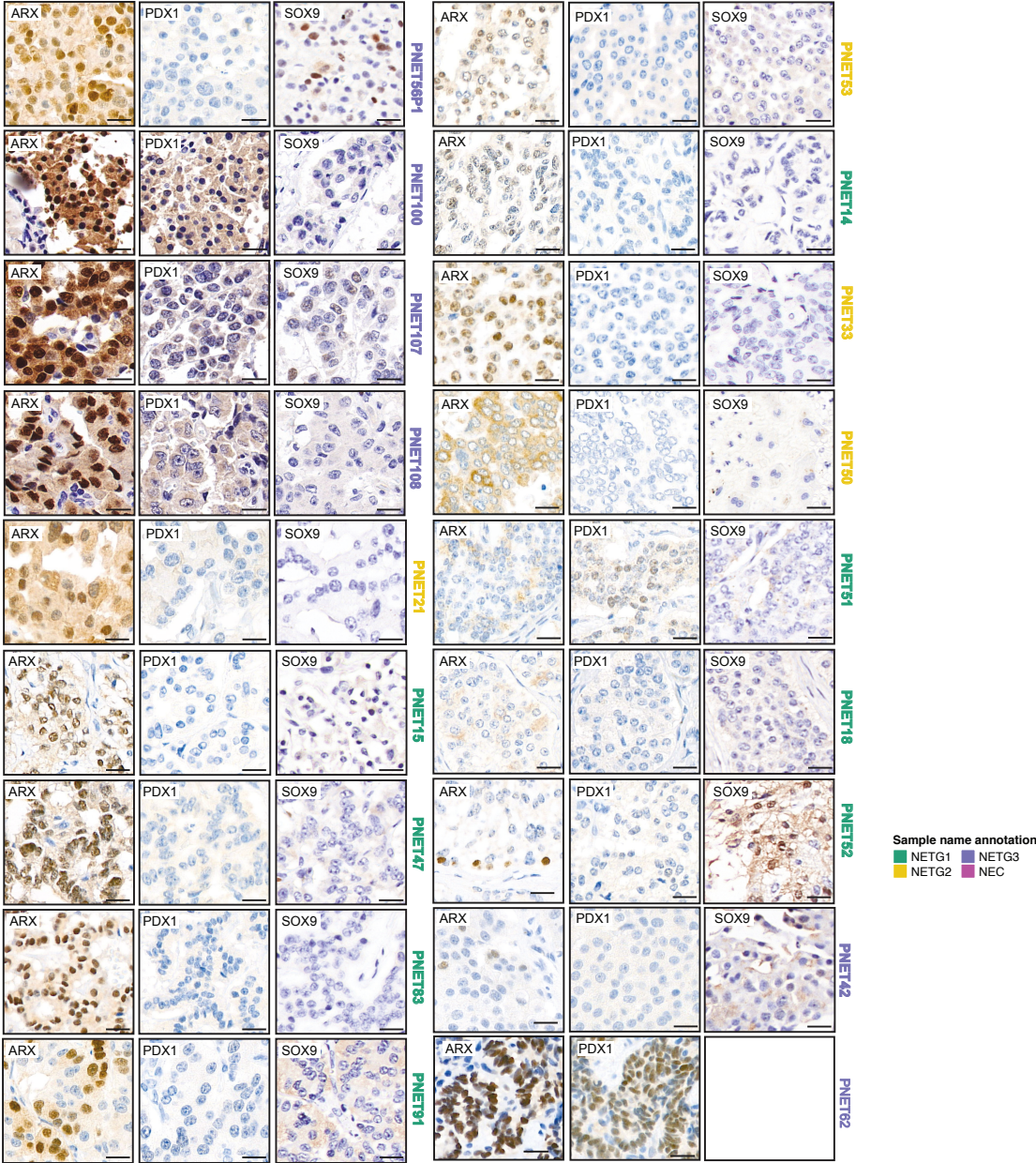
FIGURE 3.11: **Cell marker IHC of Group A.** IHC analysis PDX1, ARX and SOX9 in Group A samples. All samples that underwent IHC for at least one markers are depicted. Scale bars represent 20 μm.
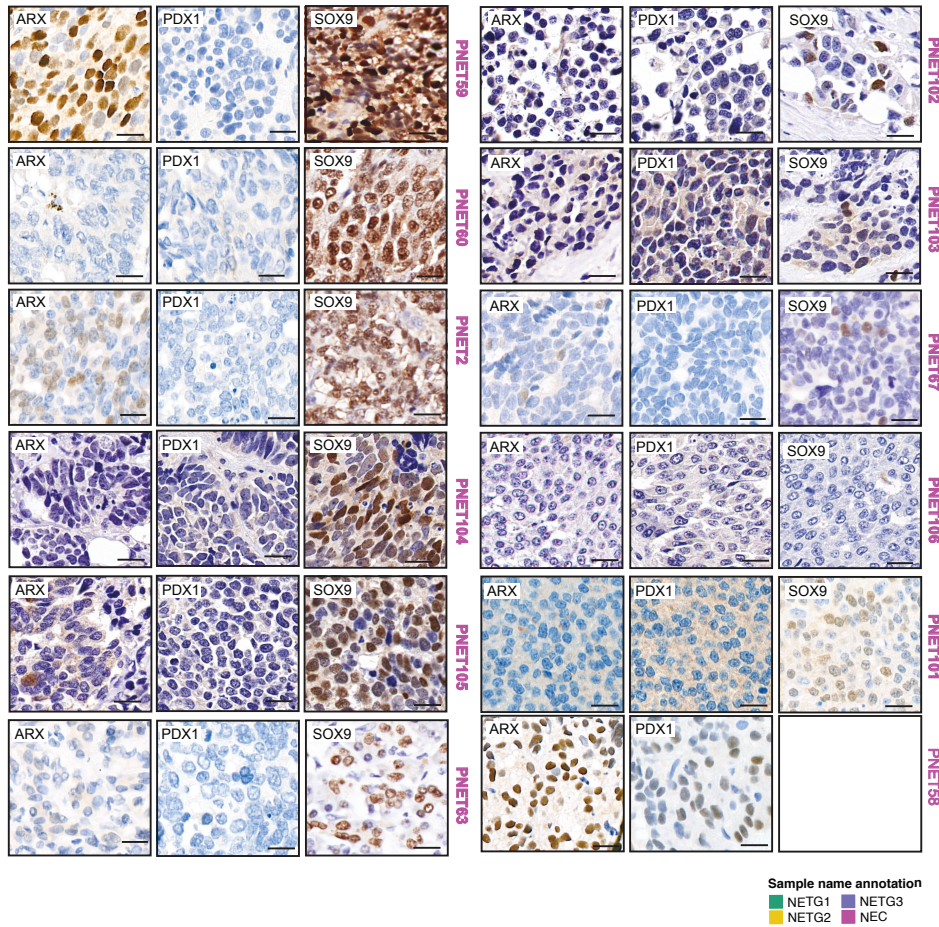
FIGURE 3.12: **Cell marker IHC of Group B.** IHC analysis PDX1, ARX and SOX9 in Group B samples. All the samples that underwent IHC for at least one marker are depicted. Scale bars represent 20 μm.

of the majority of the Group B tumors. The proximity of Group B tumors to the ductal and acinar cell profiles suggests a clear resemblance of this tumor subgroup to the exocrine cells of the pancreas and a rather distant relation to the endocrine cells of the pancreas.

To further corroborate our finding, we utilized an independent method to determine the normal cell signature composition within the PanNEN samples. As a reference, we utilized 6096 CpG probes defined by Moss *et. al.*, which distinguish ductal, acinar and β-cells (Moss et al., 2018). For the missing α-cell reference we utilized the 450K profiles mentioned in the previous sub-section 3.4.1 and determined the methylation values for the CpGs which differentiated β-, acinar and ductal cells. Using Euclidean distance analysis, we confirmed that the reference cell types could be clearly discerned from one another and therefore, did not require additional probe selection (Figure 3.14a). Given the reference and the tumor sam-
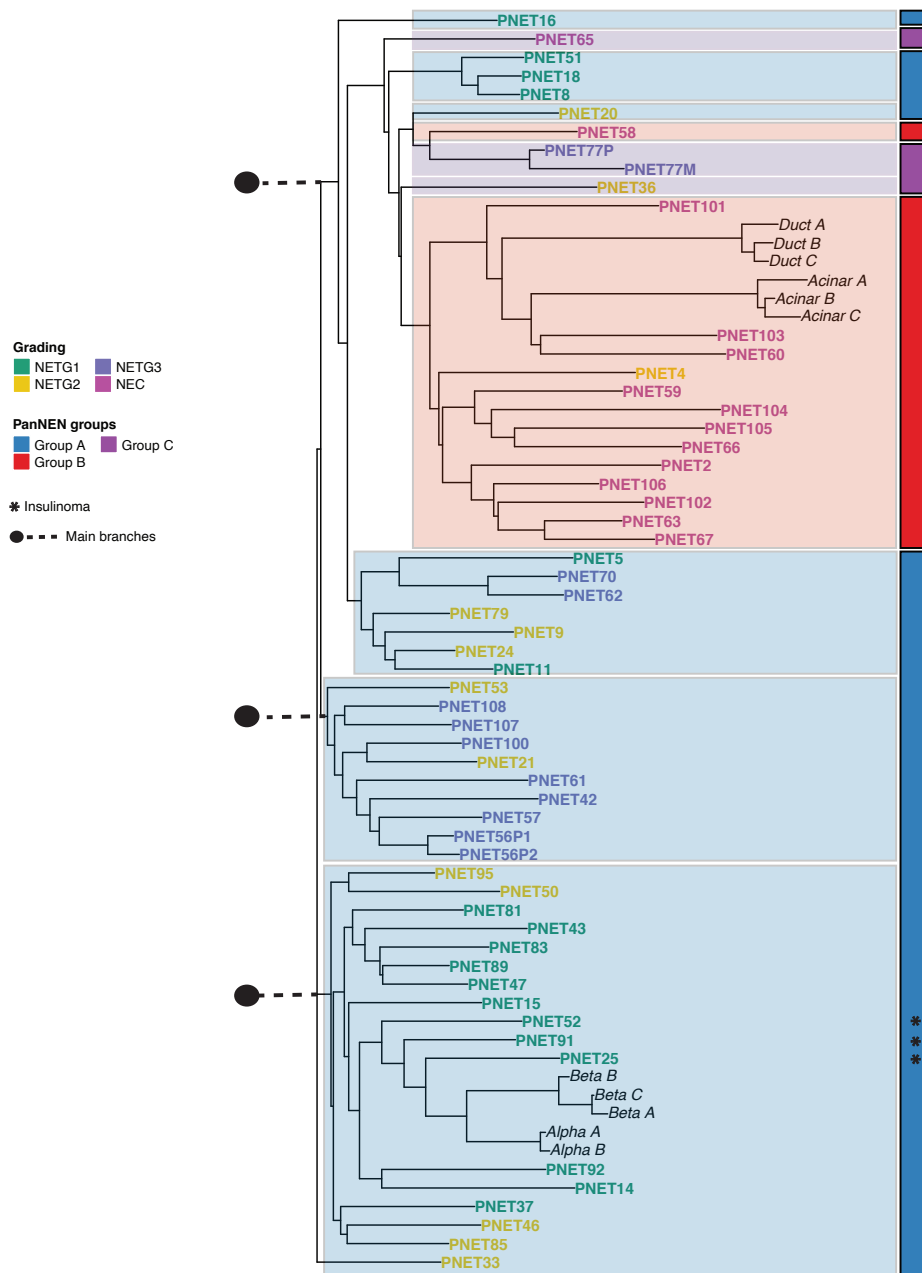
FIGURE 3.13: **Clustering analysis using DMPs of normal cell types.** Phylogenetic analysis of PanNEN tumors and normal pancreatic cell types. Pearson distance between the samples were computed using the differentially methylated CpGs between normal α-, β-, ductal and acinar cell samples (46,500 DMPs) and neighbor-joining tree estimation was performed to determine the clads and the branching. Row annotation identifies the PanNEN subgroups and highlights the main branch from which the closely associated samples of a given subgroup split. The black dots depict the three main ancestral branches.
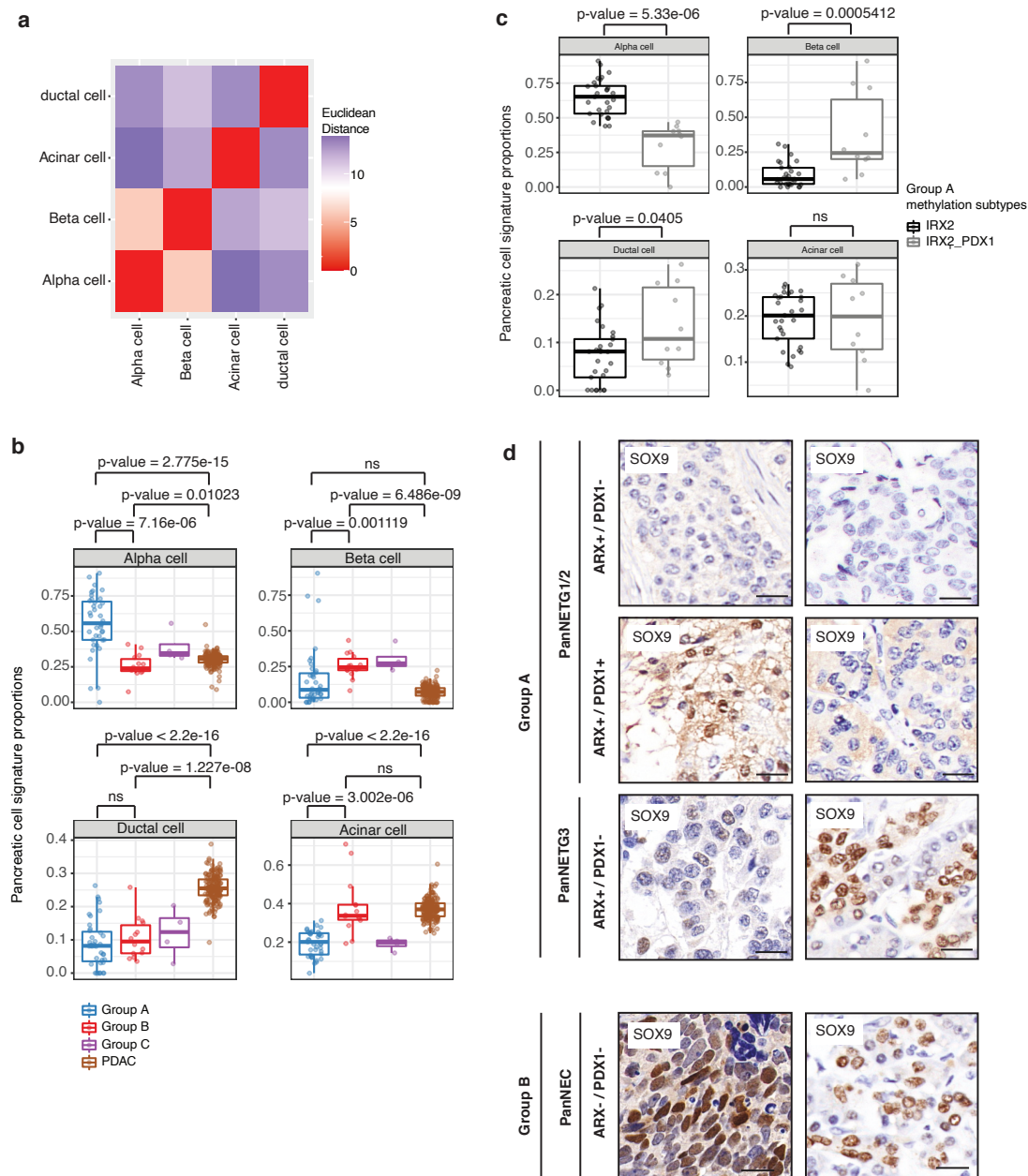
FIGURE 3.14: **Deconvolution of normal cell type composition in PanNEN methylome groups.**
(a) Correlation analysis of the reference pancreatic cell atlas. Euclidean distance between each cell
type was computed and the correlation matrix of the distances is displayed here. The heatmap
color depicts the distance measure between a given normal cell type pair. (b) Proportion of pan-
creatic cell signatures in tumor subgroups. The boxplot represents the distribution of the signature
proportions of α-, β-, ductal and acinar cells (each main box) in the subgroups and PDAC tumors;
each dot depicts the proportion of the respective cell type signature in a given sample; statistical
analysis was performed on the difference in distribution between subgroups using a two-sample
Wilcoxon test. (c) Proportion of pancreatic cell atlas signatures in *IRX2* hypo/ *PDX1* hyper and
IRX2 hypo/ *PDX1* hypo tumors of Group A. the boxplot represents the distribution of the propor-
tion of signature of α-, β-, ductal and acinar cells (each main box) in Group A tumors of respective
IRX2 and *PDX1* methylation profiles; each dot depicts the proportion of the respective cell type
signature in a given sample. (d) IHC analysis of SOX9 in PanNEN subgroups. Representative
images of IHC in each subgroup are displayed here; the scale bar represents 20um.

ples, a deconvolution algorithm was employed which models the profile of the tumors as a linear combination of the methylation profiles of the cell types. Under this assumption, the method determines the methylation signature proportion using non-negative least squares linear regression (NNLS). In addition to the grouped samples from our PanNEN cohort, 167 PDACs were included as an independent pancreatic tumor type (Lomberk et al., 2018).

Group A PanNETs displayed significant enrichment for the α-cell signature composition compared to Group B and to a lesser extent to Group C (Figure 3.14b and Figure 3.15a). The mean proportion of the α-cell signature present in Group A was 0.55, with the maximum proportion present at 0.91. Apart from three cases, all samples carried an α-cell signature proportion of > 0.25 and the three outliers carried the β-cell signature proportion > 0.70 (0.71, 0.74, 0.90) (Table 6.9.1). Given the presence of α-cell and β-cell signature proportions in Group A tumors, we further explored the proportions with respect to the IRX2-PDX1 subclasses of Group A identified in sub-section 3.4.1 (Figure 3.14c). Interestingly, tumors carrying a high α-cell signature proportion were more significantly enriched in samples with hypermethylation of *PDX1*, and tumors harboring a high β-cell signature proportion or an equal representation of both cell type signature proportions displayed an enrichment for hypomethylation of both *PDX1* and *ARX*. With respect to the remaining normal cell signature proportion, Group A tumors exhibited a β-, ductal and acinar cell signature proportion of < 0.40.

Group B NECs, on the other hand, showed a significant increase in the acinar cell signature proportion when compared to Group A and to Group C (Figure 3.14b). The range of the acinar cell signature proportion for Group B tumors fell in between 0.19 to 0.70, with a mean value of 0.38. The two samples carrying the lowest acinar cell signature proportion, PNET59 and PNET66, harbored equal α-cell and β-cell signature proportions (PNET59: 0.38 AND 0.35; PNET66: 0.41, 0.34) (Table 6.9.1). Interestingly, however, the most striking feature of Group B tumors was that the acinar cell signature proportion of Group B was comparable to that found in PDAC tumors (a mean value of 0.37 and range between 0.25 to 0.61) (Figure 3.14b and Figure 3.15b). However, the ductal signature composition of PDAC tumors was significantly higher than other groups of PanNENs, distinguishing this pancreatic exocrine cancer from PanNENs and specifically from exocrine-like Group B (Figure 3.14b). NETG3 samples from Group A contained similar profiles compared to other PanNET samples, and showed no resemblance to the acinar cell similarity of PanNECs, implying yet another characteristic distinguishing high-grade PanNEN tumors (Figure 3.15c).

In light of all three approaches taken for the identification of the cell of origin for the PanNEN methylome groups, the analyses uncovered that Group A tumors carry α-like, β-like and intermediate-like tumors of endocrine origin, while Group B tumors harbor a significantly more exocrine profile and maintain an acinar-like signature similar to PDAC tumors.
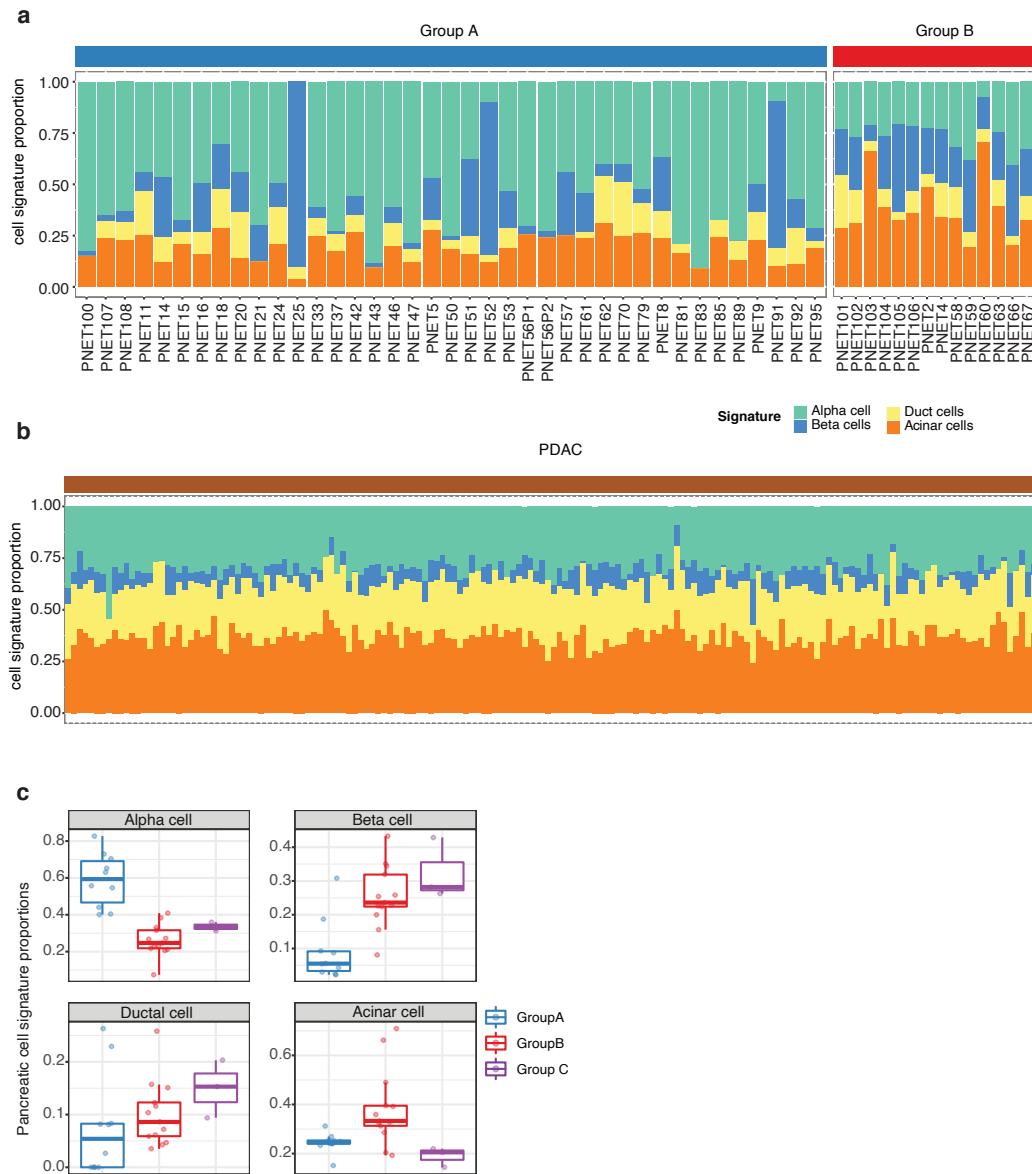
FIGURE 3.15: **Proportion of cell signatures in the PanNEN cohort.** (a) Visualization of cell signature composition in tumors. The stacked bar plot represents the proportion of the pancreatic cell signature summed to 100%; the colors represent cell signatures of α-, β-, ductal and acinar cells; top annotation depicts the PanNEN subgroup for the respective samples. (b) Proportion of pancreatic cell signature in each tumor of PDAC (n=167). Annotations are the same as in figure a. (c) Proportion of pancreatic cell signatures in NETG3 and NEC tumors of Group A Group B as well as PDAC tumors. the boxplot represents the distribution of the proportion of cell signature of α-, β-, ductal and acinar cells (each main box) in the NETG3 and NEC tumors of Group A and Group B subgroups and PDAC tumors; each dot depicts the proportion of the signature of the respective cell type in a given sample.

3.4.3   Group B PanNECs display SOX9 patterns of exocrine cells

The striking similarity in acinar cell signature proportions between PDAC tumors and the Group B subgroup compelled us to identify additional evidence that could implicate an analogous mechanism of tumor formation. An earlier study linked SOX9 expression and formation of precursor lesions of PDACs from acinar cells, together with expression of oncogenic KRAS (Kopp et al., 2012). SOX9 has been identified as a marker for PDACs as well as for the precursor lesions PanIN and IPMNs (Shroff et al., 2014). Given the clear importance of SOX9 in PDAC tumors, particularly with respect to the formation of the tumor from acinar cells, we investigated the presence of SOX9 in PanNEN tumors.

To identify whether the PanNECs in Group B display SOX9, we performed IHC for representative samples of the cohort. We obtained a subset of the tumor samples from our cohort, which included 11 Group B samples and 17 Group A (representative images in Figure 3.14d; full cohort IHC provided in Figure 3.11: Group A and Figure 3.12: Group B, Table 6.14.1 includes pathologist diagnosis). With the exception of four cases, Group A samples were negative for SOX9 expression. Among the four cases that were positive for SOX9 (SOX9+), three NETG3s presented with nuclear staining in a subset of cells within the tumor tissue (PNET42, PNET56P1 and PNET107) and one NETG2 (PNET52), an insulinoma and MEN1 syndrome tumor, exhibited nuclear and cytoplasmic staining in a heterogenous manner. Interestingly, from the 11 PanNEC samples of Group B, 9 samples were positive for SOX9 staining (Figure 3.12). In contrast to Group A, Group B displayed strong staining and presented SOX9 homogeneously in 5 out of the 9 positive samples. The two samples that displayed no SOX9 expression (PNET106 and PNET67) were KRAS mutants. To further link the endocrine/exocrine feature with respect to SOX9 expression and thereby illustrate a representative cell marker signature of Group A and Group B, we additionally evaluated the SOX9 expression in concert with ARX and PDX1 expression of the same samples. Group A SOX9+ NETG3s were additionally ARX+ (Figure 3.11). This pattern was also present in three SOX9+ Group B NEC tumors (PNET59, PNET2, PNET67). Nevertheless, the SOX9- Group B sample were additionally ARX- and PDX1- (Figure 3.12). Collectively, our IHC analyses identified that α-/β-/intermediate-like tumors in Group A harbored significantly more ARX+PDX1-SOX9- phenotype, while Group B acinar-like tumors were enriched for the ARX-PDX1-SOX9+ phenotype (adjusted p-value = 0.001998: Fisher's Exact Test and False Discovery Rate (FDR) corrected).

3.5   Characterization of PanNEN in-vitro cell line models

The genetic and epigenetic characterization of PanNEN entities have indeed advanced over the years. The novel findings in section 3.1 through section 3.4 would certainly make a

significant impact on further expanding this knowledge and potentially become the foundation for future research in characterizing the formation of PanNEN tumors. Regrettably, however, the current knowledge is far from being translated into the clinic. Additional *in-vitro* and *in-vivo* analyses in alignment with the current findings is necessary to translate the insights we gained from high-throughput technology to the bedside.

3.5.1   Genetic and epigenetic features of BON1 and QGP1 represents Group B PanNEN methylome tumors

The bona fide *in-vitro* models BON1 and QGP1 have been introduced in chapter 1, including their genetic composition at the mutational and copy number level. We decided to expand our current knowledge by further analyzing the cell lines with respect to the Pan-NEN methylome groups. Specifically, BON1 and QGP1 cell lines were characterized using three gene panels, CCP, PanNEN and an in-house colorectal cancer panel (CRC5), to determine the underlying driver mutations. Furthermore, the cell lines were investigated in terms of copy number aberrations and methylation profiles to additionally identify features representative of PanNEN methylome groups.

The key driver mutations and chromosomal copy number changes of BON1 and QGP1 have been previously investigated; nonetheless, the current analysis began by first identifying the alterations within our batch of BON1 and QGP1 cell lines to determine whether additional alterations accumulated under the continued pressure of culturing conditions. As part of my masters thesis, the cell lines underwent targeted panel sequencing and key mutations in cancer-associated oncogenes and tumor suppressors were characterized. In total, BON1 had 10 alterations in cancer-associated genes, while QGP1 harbored 12 mutations (Table 6.10.1). The *NRAS* mutation p.Q61R and The *ASXL1* mutation p.V1349A found in BON1 were both COSMIC aberrations. In addition, BON1 also harbored a stop gain mutation in *TP53*, identified as 'pathogenic' in the ClinVar database and a *SMAD4* variant predicted to result in a frameshift truncation. Genes altered in NETG3 tumors including *MYH11* and *HNF1A* were also mutated in the BON1 cell line.
In contrast, QGP1 harbored COSMIC variants in *KRAS* (G12V) and two different variants of *TP53* (p.P98L – COSMIC and p.P98Lfs*75 – 'Pathogenic' - ClinVar). QGP1 also harbored an *mTOR* mutation and surprisingly also carried two different *APC* COSMIC variants (p.R2166Q, p.Y486*), which are unlikely to be expected in PanNEN tumors. Nevertheless, given the mutational profile, lacking the key drivers of PanNET tumors and harboring alterations in the recurrently altered genes of PanNEC tumors, including *TP53*, *KRAS* and *SMAD4*, it is likely that at the mutational level, BON1 and QGP1 are representative of the Group B PanNEN methylome group.

Copy number profiling of BON1 and QGP1 showed aberrations consistent with current literature knowledge and highlighted some key focal events complementing the mutational findings (Figure 3.16a). Due to the limitation of the analysis pipeline in detecting ploidy from the methylation array, the data could not be used to interpret the total chromosome count of the cell lines. Regional amplification in chromosome 5,10,14,16 and 20 were however evident in BON1, along with regional deletions in chromosome 1,6,9 and 12. Whole chromosomal LOH was also evident in chromosomes 4,6,8,9,11,12,13,17,18 and 22.

In contrast, regional gains were present in all chromosomes of QGP1, with higher $\log_2$ ratios, referring to a higher copy number state, strongly displayed in chromosomes 1,12 and 20. Interestingly, however, the copy number losses were as much prevalent as the gains in QGP1 with strong deletions apparent in chromosomes 1,3,4,6,8,9,12,13 and 18. The particular genome copy number profile of QGP1 with significant focal deletions is consistent with Group B tumors, while Group A maintained relatively stable genomes with low levels of strong focal aberrations and high numbers of whole chromosomal gains.

We further investigated the copy number status specifically of *CDKN2A* and *RB1* as the regional loss of these genes was a clear signature mark of Group A and Group B, respectively (Figure 3.16b). *RB1* loss was exclusively present in QGP1 only, while the $\log_2$ ratio close to zero was maintained in BON1 indicative of two copies within the *RB1* locus. Interestingly, BON1 showed a complete loss of the region harboring *CDKN2A*, while QGP1 seems to also show deletion of the region harboring the gene. Taken together, the copy number status of QGP1 is more in line with what is exhibited by Group B tumors.

Given that the genetic aberrations were consistent with Group B tumors, we speculated whether the methylation features of Group B tumors were also preserved in BON1 and QGP1. The methylation of cell marker DMPs distinguishing Group A and B were not consistently maintained in BON1 and QGP1 (Figure 3.16c). More specifically, both QGP1 and BON1 lacked the global hypermethylation feature of Group B. They did, however, show a global pattern closely resembling NETG3 tumors of Group A, particularly due to the hypermethylation of *PAX6* and maintenance of the majority of the Group A cell marker genes' methylation profile in the remaining probes. As it is evident that the cell marker gene signature of Group A tumors can also be seen in rare cases within Group B tumors, we decided to identify how the cell lines cluster within the phylo-epigenetic tree along with the tumors and normal cell type (Figure 3.16d). As previously described, the methylation data of DMP probes distinguishing the normal cell types were extracted and the tree was reconstructed by including the beta values of DMP from the tumor samples as well as BON1 and QGP1. The cell lines clustered farthest from the main branch of Group A and formed a small clad along with 2 NETG3 (PNET62 and PNET70), 1 NETG2 (PNET9) and 2 NETG1 (PNET11 and PNET5). Nonetheless, the most informative feature determining the cell lines' representation of the groups resulted from the fact that the cell lines clad emerged from the main
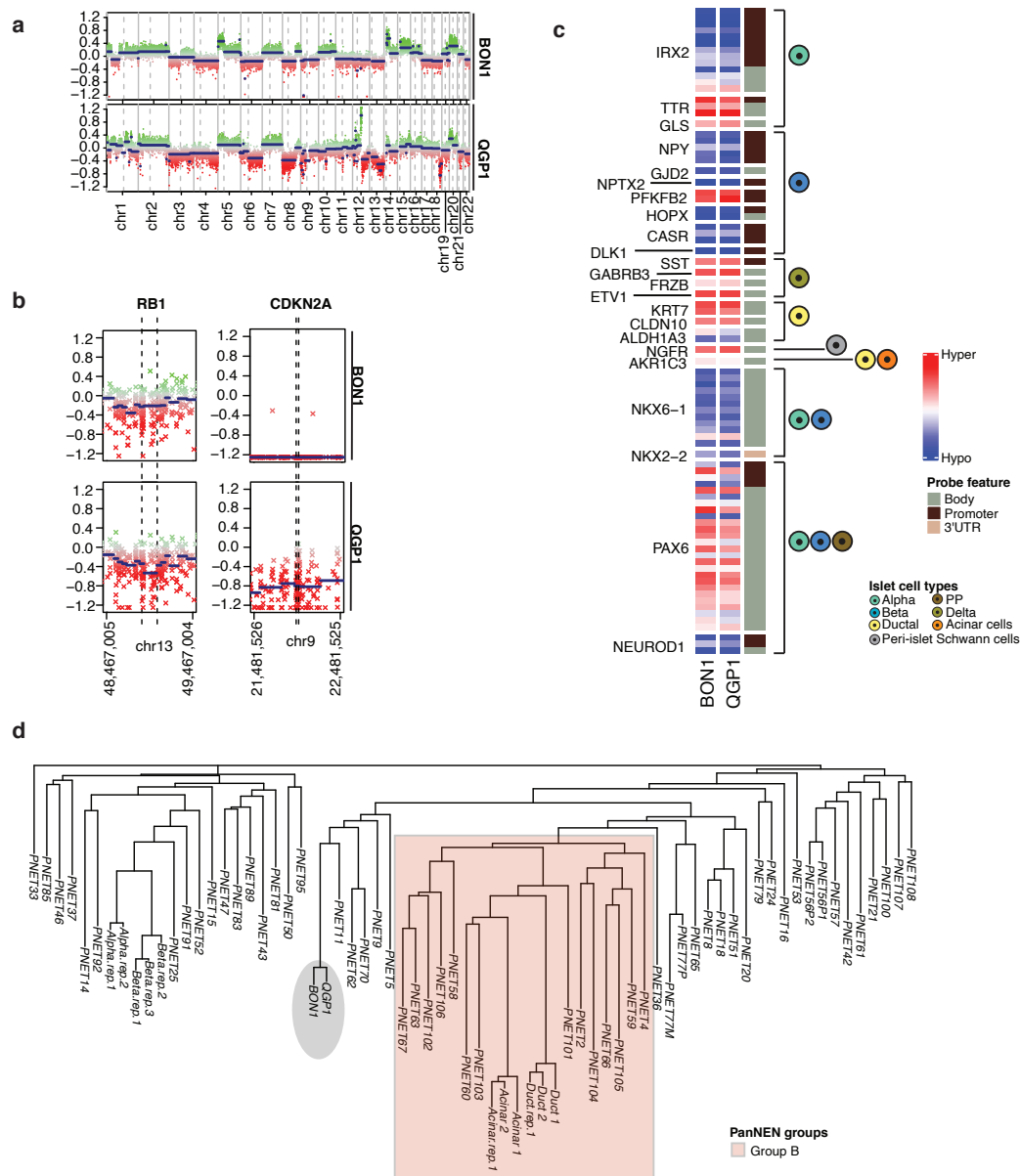
FIGURE 3.16: **Genetic aberrations of cell lines.**(a) Genomic profiling of cell lines. Genome wide copy number aberrations of BON1 and QGP1 are displayed here. Intensity values of each bin are plotted in colored dots. Each color indicates methylated and unmethylated channels of each CpG. Segments are shown as horizontal blue lines. (b) CNA status of *RB1* and *CDKN1A* and *CDKN2B* loci. Color annotations are the same as in figure a. (c) Methylation beta value of significant DMPs initially identified between Group A and B. Heatmap displaying the methylation beta values of DMPs (rows) in BON1 and QGP1 (columns). The heatmap color represents the methylation beta value. row annotation identifies the genomic region associated with the probe (panel on the right) in addition to which cell type the gene shows specificity for gene expression (according to PangaloDB). (d) Phylogenetic analysis of PanNEN tumors, normal pancreatic cell types and PanNEN cell lines. Pearson distances between the samples were computed using the differentially methylated CpGs between normal ductal and acinar cells samples (46500 DMPs) and neighbor joining tree estimation was performed to determine the clads and the branching. The grey circle shows the cell line positioning within the tree and the red box displays the group B clad.

branch that also gave rise to the clad harboring all Group B tumors as well as the exocrine cells.

One can therefore agree that given the genetic and epigenetic features of the cell lines in comparison to PanNEN methylome groups, BON1 and particularly QGP1 due to its more precise consistency of copy number feature are *in-vitro* models representative of Group B tumors.

3.5.2 Mathematical modeling of the MAPK, PI3K/AKT and mTOR signaling networks in BON1 and QGP1

It is imperative that the identification of the cell lines as being representative of PanNEN methylome groups is accompanied by a comprehensive understanding of the biological processes associated with the growth and survival of these models. In fact, the interconnected signaling networks within these *in-vitro* systems are at the core the key factors involved in the overall survival of these cell lines, which makes them the perfect system to exploit and identify effective therapeutic approaches. Consequently, quantitative mathematical models were generated to determine the behavior of key cancer-associated signaling pathways including MAPK, mTOR and PI3K/AKT networks in BON1 and QGP1, while emphasizing the crosstalks, feedbacks and feedforward loops facilitating the proliferation and survival of these cell lines.

The modelling approach we utilized is based on the well-established modular response analysis (MRA) (Bruggeman et al., 2002; Kholodenko et al., 2002) that has been employed to model signaling networks in multiple cancer cell lines (Klinger et al., 2013; Dorel et al., 2018; Hood et al., 2019; Brandt et al., 2019). The method is implemented in the R package STeady-STate Analysis of Signalling Networks (STASNet) for a more efficient utilization of the approach. Briefly, the network system is represented by phosphorylation of phosphosites of few key kinases within a given signaling pathway. These are termed the nodes of the model. Given the data, MRA aims to quantitatively deduce the local response coefficient, r^i_j, which is the direct local influence one node j has on the node i, from the global response coefficient, denoted as R^i_j, which is the overall response of the system on a specific node i upon a specific perturbation of node j. The global response coefficient is therefore determined experimentally and measured upon the system reaching steady state. Mathematically, the local response coefficient can be represented as follows:

$$R = -r^{-1}(s + i) \tag{3.2}$$

where s and i define the simulation and inhibition matrix, respectively. The matrices contain the stimulation strength or inhibition strength on a node under a given perturbation.

Additionally, STASNet incorporates a measurement error accommodating for measuring discrepancies and modifies the equation as follows:

$$R^m - E = R = -r^{-1}(s + i) \qquad (3.3)$$

wherein R^m is the measured global response and E is the error model.

The approach is implemented in STASNet, such that under a given network topology the model tries to fit the data and determine the local response and perturbation coefficients by minimizing the log-likelihood, estimated by the weighted sum of squared differences between data and the fit. The fit, which corresponds to the likelihood of the model, is further improved by removal of a link and/or by addition of a new link. Finally, profile likelihood analysis is implemented to evaluate the identifiability and reliability of the parameters. Computing the profile likelihood, and thereby identifying the parameter value confidence interval, allows us to determine how much a parameter can vary in each cell line without significantly worsening the overall fit of the model.

Evaluation of the data

Given that the cell lines BON1 and QGP1 have substantial resemblance to the Group B PanNEN methylome group, they are perfect candidates for studying drug sensitivity, and combinatorial treatment possibilities for therapeutic approaches in patients. As a first step, we therefore decided to characterize the three main cancer-associated signaling networks, MAPK, mTOR and PI3K/AKT in BON1 and QGP1. The experimental processes for the generation of perturbation data was performed by Dr. Pamela Riemer from the Molecular Tumor Pathology, Group of Prof. Dr. Christine Sers at the Charite University of Berlin and Dr. Anja Sieber from the group of Prof. Dr. Nils Blüthgen of the Integrative Research Institute (IRI) for Life Sciences at Humboldt University, Berlin. Briefly, the perturbation data of the cell lines were generated by first starving them for 24hr after which different inhibitors were added. After 1hr of inhibition, the cells were additionally stimulated with Epidermal Growth Factor (EGF) and Insulin-like Growth Factor (IGF) (Figure 3.17a and b). Key phosphorylation events within these signaling cascades depicted in Figure 3.17c were measured upon reaching a steady state after 30 minutes of stimulation. Initial analysis of the analytes represented as M-A plots revealed a strong agreement among the three replicates. The M-A plot transforms the data onto M (log ratio) and A (mean average) scales. As expected for unbiased replicates, they spread narrowly around their respective mean value (0 in M-axis) (Figure 3.18a). The analytes that deviate from this showed intensity-dependent signal variation, such as in the case of MEK1 and p70S6K as well as intensity-independent signal variations, including AKT_{T308} phospho-sites, and RPS6 readouts. AKT_{T308} discordance cannot be associated with one particular replicate as M-A value for replicate comparison of
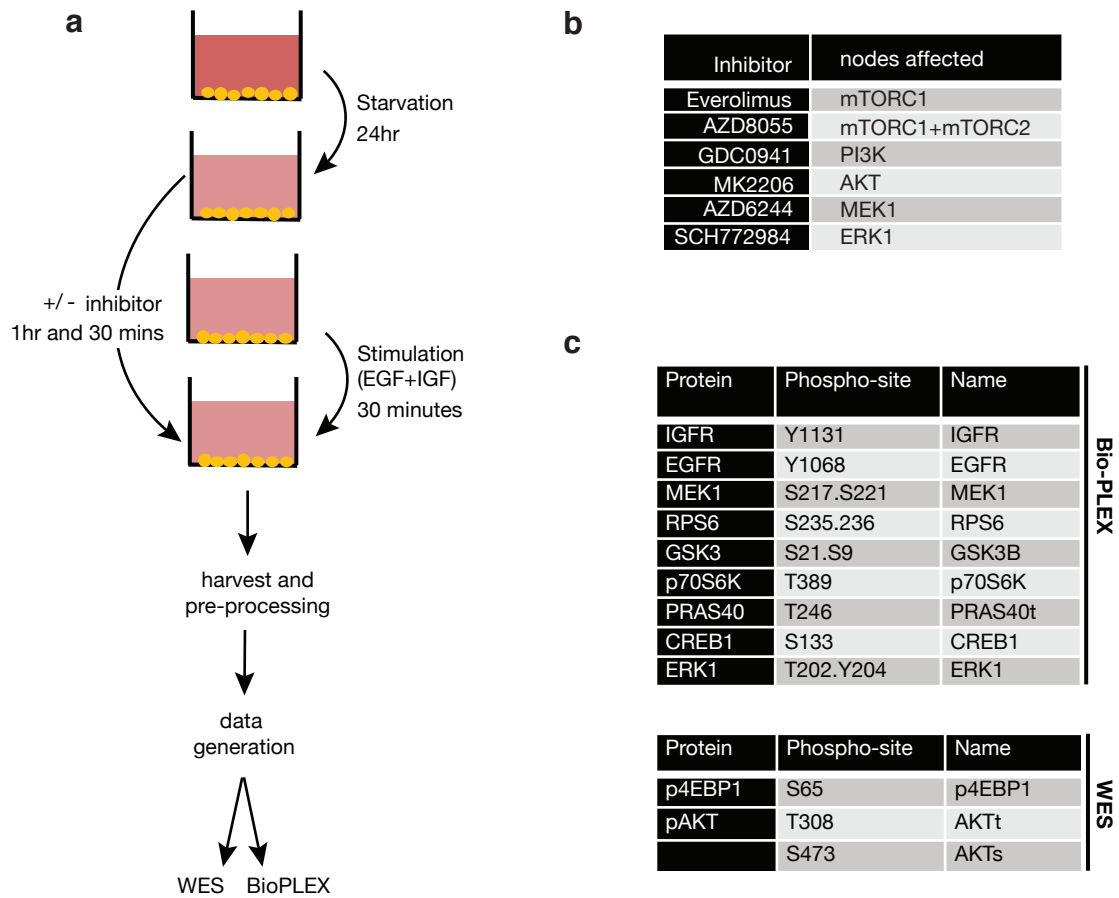
FIGURE 3.17: **Perturbation data generation in BON1 and QGP1.** (a) Experimental setup of perturbation analysis in BON1 and QGP1. Each step was performed in the sequence shown. Data quantification was done using two different platforms: WES and Bio-Plex. (b) Inhibitors and their targets. The table displays the inhibitors that were used for the perturbation of the cell lines (left column) and their respective perturbed targets (column right). (c) Measured phosphosites and their respective proteins. The tables are split according to the platform that was used to measure the phosphosite (WES vs Bio-Plex). the name column depicts the identity of each phosphosite displayed in the steady state model.

A-C has a slope close to zero while the replicate comparison between A-B and B-C is above and below the mean value of 0, respectively. Therefore, while modeling the data, it is necessary to cautiously interpret the parameters associated with the analyte. With respect to RPS6 values, replicate A seems to be an outlier. We decided to nevertheless keep the replicate as the overall M-A plot combining all analytes derived from each method showed that there were no systematic shifts among replicates across the panel of readouts and therefore one can proceed with downstream analysis and modeling (Figure 3.18b). The $\log_2$ fold changes of the stimulation and treatment compared to the non-treated cells showed strong activity among the analytes in BON1 and QGP1 (Figure 3.18c). In general, the fold
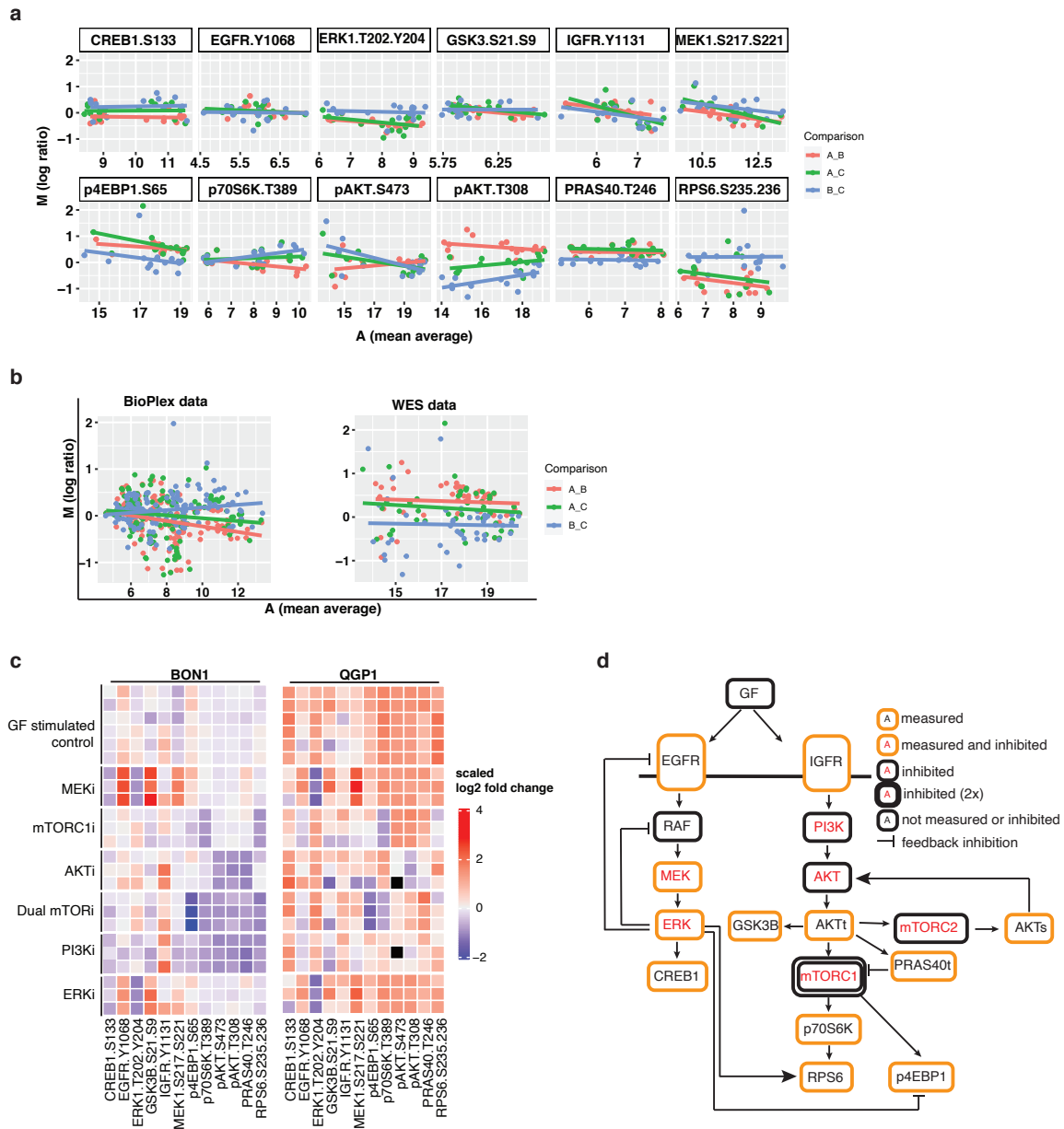
FIGURE 3.18: **General data evaluation of the perturbation dataset.** (a) M-A plots of raw data per readout. The y-axis represents M (log ratio) scale and the x-axis shows A (mean average) scale between two replicates. The colors represent the replicates under comparison. (b) M-A plots of raw data for all antibodies combined in BioPlex and WES datasets. The axis annotation is consistent with figure a. (c) Log transformed fold changes of experimental data. The heatmap displays the log_2 value normalized to non-stimulated controls for BON1 and QGP1. Row annotation depicts the inhibitor and column annotation depicts the phosphosite measured. (d). Literature network topology of PanNEN cell lines. The figure displays a graphical representation of the initial network topology utilized for the modeling of BON1 and QPG1 signalling networks.

change range was much higher for QGP1 compared to BON1, with relative maintenance of strong activity under all conditions. Notwithstanding, we did see clear shifts in activities in expected players under specific perturbations. For instance, MEK1 inhibition showed strong MEK1 phosphorylation while ERK1 phosphorylation was significantly hampered in both cell lines. Additionally, BON1 and QGP1 displayed dephosphorylation of p4EBP1 and p70S6K, the key downstream targets of mTOR under dual mTOR inhibitions, further confirming the reliability of the experiment as well as the data for generating the models.

Development of steady state model of signaling network

We deduced a general network topology of MAPK, mTOR and PI3K signaling in BON1 and QGP1 based on literature analysis. Given the scarcity of information regarding the network configuration in these cell lines, the majority of the topology was inferred from what is known with respect to GEP-NETs in general, PDAC cell lines as well as other KRAS driven cancer cell lines (Zanini et al., 2020; Rozengurt, 2014; Lamberti et al., 2018; Mendoza, Er, and Blenis, 2011; Sarbassov et al., 2005; Zitzmann et al., 2010; Prada et al., 2018; Valentino et al., 2014; Vandamme et al., 2016). The phosphositePlus database was also explored to determine any additional known kinases upstream of our target phosphosites (Hornbeck et al., 2015). We began with the assumption that the topologies are largely similar in BON1 and QGP1. Utilizing the network topology and the measured global responses, we developed a steady state network model that quantified the signaling network. In order to improve the fit of the model, links were either reduced or added iteratively based on the suggestions of STASNet as well as based on a combination of literature knowledge and the evaluation of correlation between responses of the readouts.

BON1 The Initial network topology-based, reduced model resulted in 25 coefficients with a likelihood score of (LS) 326.24. The LS defines the fit of the model given the data, with the value expected to be around 59 (12 readouts x 7 perturbations - 25 coefficients) for an appropriate fit. The network reduction step resulted in removal of three links (Figure 3.19d). The algorithm determined this by iteratively removing the links that do not significantly deteriorate the LS. Provided the LS was vastly larger than the theoretical value, it was evident that the network does not contain all the information (links) necessary to recapitulate the data. This was particularly true in the case of GSK3B (Figure 3.19a; heatmap). Based on the R^2 value, which is the proportion of the variance in the data of the readout explained by the model, the GSK3B variance is not captured by the model. This can be viewed in comparison to EGFR phosphorylation, where the model simulated very well the response of EGFR phosphorylation upon perturbation (Figure 3.19a; barplot). Hence, the R^2 value was close to 1 due to the optimal fit of the model to the data for EGFR readout. None of the responses of GSK3B were simulated under the current model and

FIGURE 3.19: **BON1 model.** (a) Evaluation of the predicted vs experimental output of CREB1 ,EGFR, GSK3 and IGFR in BON1 cell line perturbation data. The heatmap shows $\log_2$ fold changes (the value and its corresponding color) of each measurement under the respective perturbation (row annotation) (left panel). The bars represent R^2 values of each measured node (column annotation). (b) Graphical representation of the BON1 network configuration. The dotted line and solid green links show connections that were removed and added, respectively, to the initial network topology. (c) Comparison of $\log_2$ fold changes of experimental data with perturbation data. The heatmap represents the observed $\log_2$ fold changes of each measurement under a given perturbation in the experimental data (top panel), in the stimulated data based on the initial network topology (middle panel) and in the stimulated data based on the optimized final network topology (bottom panel) of figure b. (d) Quantile-Quantile plot of the model error compared to a normal distribution for BON1. The red line shows the expected line (theoretical distribution) and the dots represent the model error for each data point.

therefore, as per the suggested extension by STASNet algorithm, we added a link from RAF to GSK3B, which further improved the fit of the model as well as captured the overall variations of responses seen in the GSK3B readout. Further down the refinement process, this link was changed to ERK to GSK3B since regulation of GSK3B via ERK1 rather than via RAF is highly supported by literature knowledge (Pal et al., 2017) (Figure 3.19a; heatmap and barplot corresponding to ERK1->GSK3B). Similarly, linking ERK1 to p70S6K further improved not only the response of p70S6K under MAPK pathway inhibition, but also the response of its downstream target RPS6 (Iijima et al., 2002).

The perturbation response of IGFR was not identified using the current model, so an additional link associated with the negative feedback from mTORC2 via p70S6K was added to ensure a better fit of the global response observed in IGFR (Figure 3.19a; heatmap mTORC2 -> IGFR). Regulation of mTORC2 via p70S6K was added to complement the mTORC2 to IGFR link, as it explained the experimental output of AKT_{S473} under mTORC1 inhibition compared to no inhibition, along with the known negative feedback from p70S6K to PI3K (Rozengurt, Soares, and Sinnet-Smith, 2014). Interestingly, the p70S6K to mTORC2 and mTORC2 to IGFR have not been previous identified in literature. Nevertheless, comparing the experimental data, IGFR depicted the strong change in the phosphorylation upon dual mTOR, PI3K and AKT inhibitors (Figure 3.19c). Upon the addition of the mTORC2 to IGFR link, these responses were clearly fitted by the model ($R^2 \approx 1$)(Figure 3.19a; barplot).

With respect to CREB1, however, the fit did not improve dramatically, and in fact, none of the CREB1 responses could be modeled under the current topology. The model was refitted using the regulation via ERK (ERK -> CREB1) and AKT (AKTt -> CREB1), but, both links were removed during the reduction process, which could point to the overall CREB1 signal being too noisy for the model interpretation. The final model additionally included PI3K to RAF crosstalk (Parrales et al., 2013; Bradley et al., 2008; Ebi et al., 2013), to improve MEK1 and ERK1 readout simulations. This resulted in an LS score of 140.81, improving the fit considerably and reaching an optimum close to the theoretical score (Figure 3.19b and c). The Quantile-Quantile (Q-Q) plot represented the model error for each data point compared to a theoretical distribution, which is the identity between quantiles of the data and of the standard normal distribution. One can appreciate that with the exception of a few points in both extremes, deviating away from the normal expected fit, the overall model error was spread along a normal distribution curve (Figure 3.19d). Therefore, our steady state model with 30 parameters (Table 6.11.1) quantitatively represents the signaling connectivity of MAPK/PI3K/AKT/mTOR networks in BON1 cell line.

QGP1 Comparable to the BON1 network, the QGP1 model generated via the initial network topology captured many of the responses displayed in the experimental data.

Comparing the experimental fold change to the simulated fold change, its evident that the model is an appropriate fit of the perturbation data, given the overall trend of the responses are recapitulated well by the model upon removing one link and with only few refinement required(Figure 3.20c). An LS of 371.926 with 28 parameters meant one had to improve the network to fit the data more accurately. Based on the simulation data in comparison to the experimental data along with the R^2 plots of the readout, the main changes not captured by the model again belonged to the CREB1 responses and IGFR feedbacks (Figure 3.20a) Therefore, we extended our QGP1 model by adding a link from AKT_{T308} (Hornbeck et al., 2015) and p70S6K to CREB1, which was able to capture the strong response upon AKT inhibition as well as the phosphorylation response under MEK1 and ERK1 inhibition (Groot, Ballou, and Sassone-Corsi, 1994; Hornbeck et al., 2015). IGFR responses under dual mTOR inhibiton as well as PI3K and AKT inhibition, also displayed in BON1, was modeled by addition of the mTORC2 feedback towards IGFR (Figure 3.20a). The links substantially improved the R^2 value. PI3K to RAF crosstalk was also added as part of the network refinement process, which resulted in an improved simulation of MEK1 and particularly the ERK1 readouts (Parrales et al., 2013; Bradley et al., 2008; Ebi et al., 2013) (Figure 3.20c). Additionally, an ERK to p70S6K connection was added to better predict the p70S6K readout under PI3K and AKT inhibition (Iijima et al., 2002). Together with the newly extended links, the accurate fit of the model is certainly visible in the Q-Q plot, as almost all the data points fall along the identity line (Figure 3.20d). One single data point shifted away from the red line. This datapoint is the AKT_{T308} readout upon AKT inhibition. There was a change in the response sign as simulated data shows a positive response, while experimentally, AKT inhibition resulted in de-phosphorylation of AKT_{T308}. AKT inhibitor MK2006 not only affects the downstream propagation of AKT signaling, but also prevents the phosphorylation of AKT itself by upstream kinases (Hirai et al., 2010). However, the MRA algorithm is not equipped to handle this activity because the approach assumes that inhibitor effects with respect to phosphorylation is only occurring downstream of the node and not on the node. Collectively, our final model with an LS of 155.76 with a total of 35 parameters accurately depicts the signaling network assembly in QGP1 cell line (Figure 3.20b; Table 6.11.1).

Comparison analysis of BON1 and QGP1 parameters   The next step involved determining the parameter estimation and comparison of our cell lines. Although BON1 and QGP1 are neuroendocrine cell lines and their models share many parameters, we cannot compare them directly. They are not isogenic lines and therefore show different basal activity in each cell line. Nevertheless, the feedbacks and inhibitory effects are still comparable between the cell lines (Figure 3.21a). In addition, we looked at the parameters unique to each of the
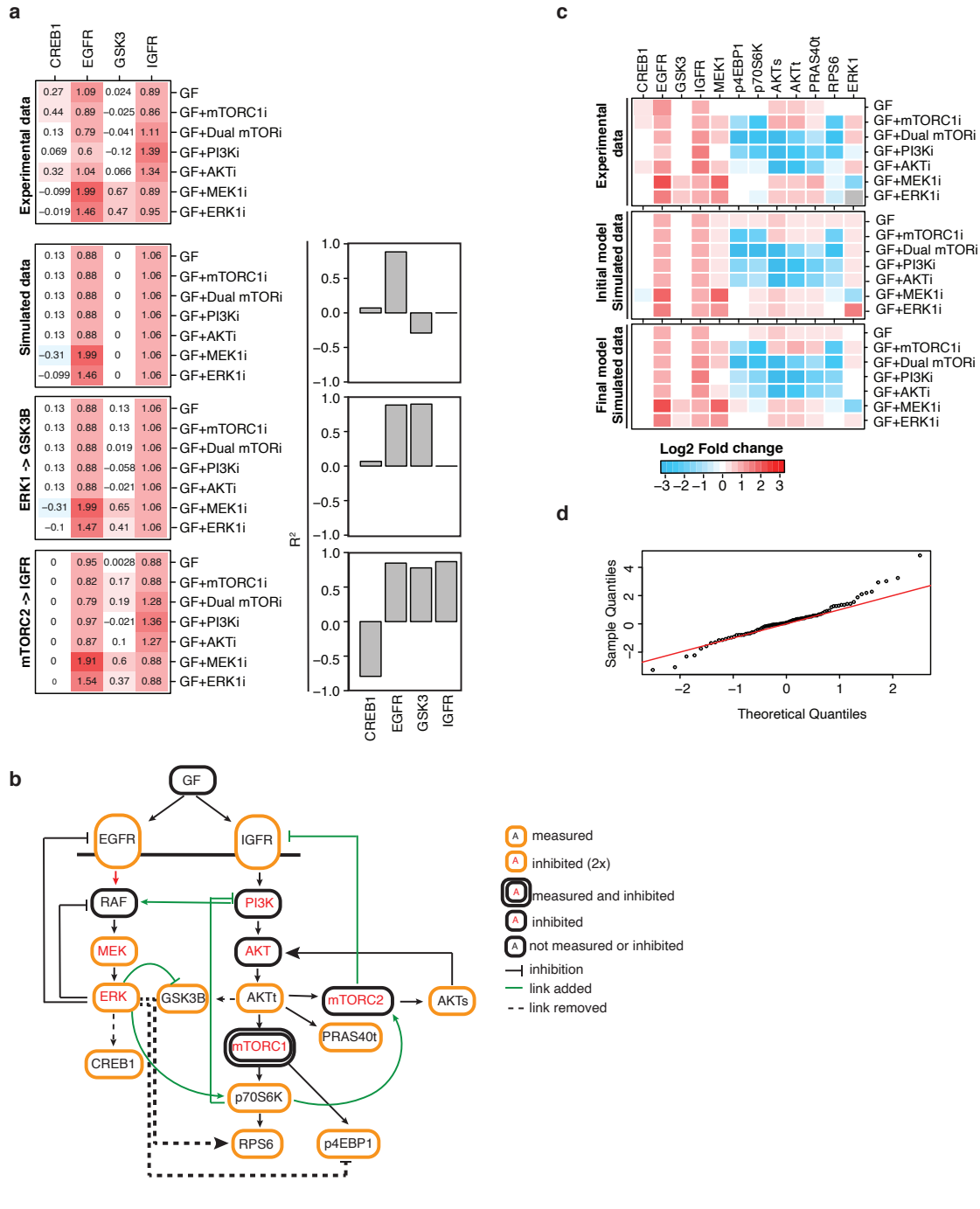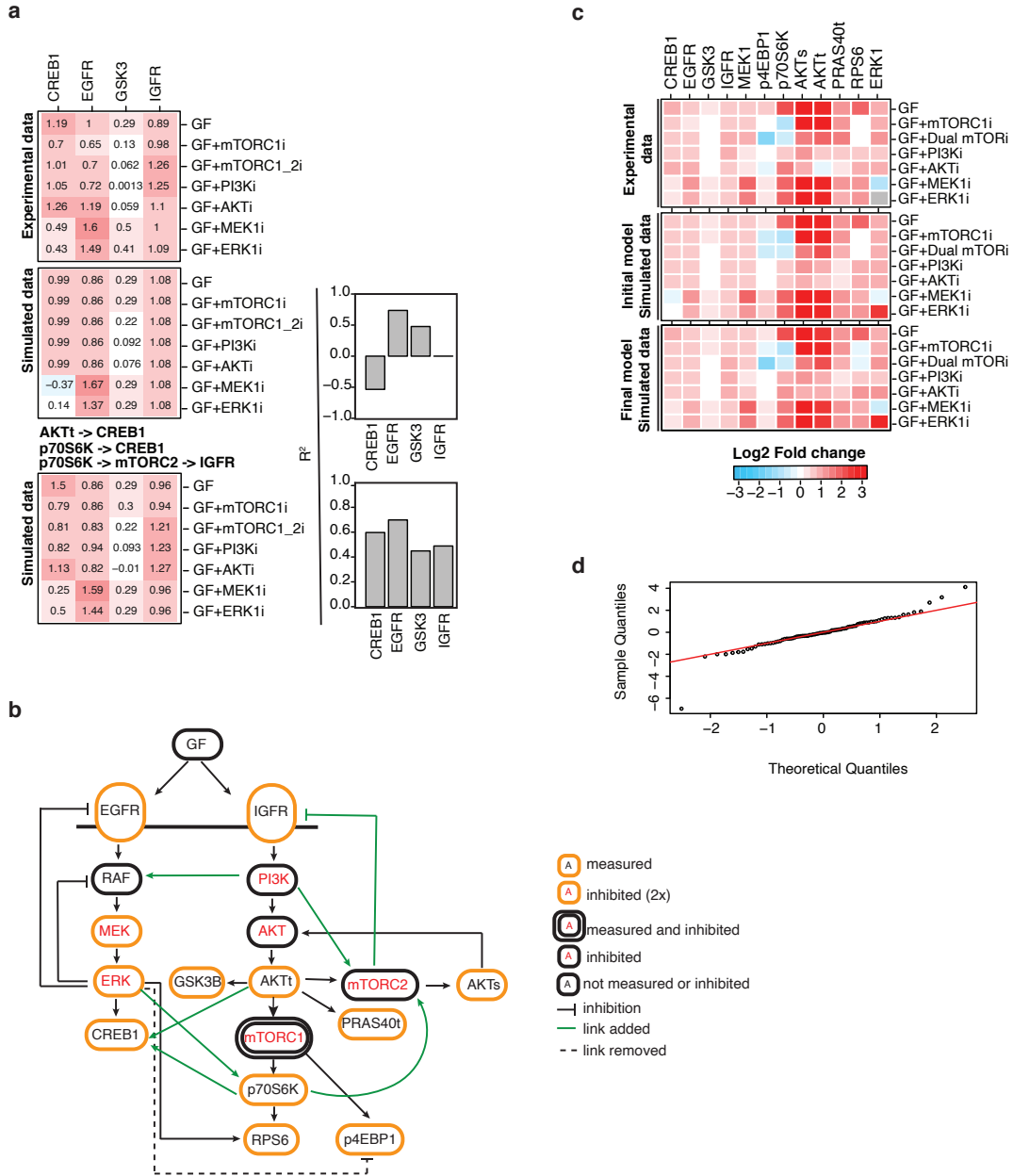
FIGURE 3.20: **QGP1 model.** (a) Evaluation of the predicted vs experimental output of CREB1, EGFR, GSK3 and IGFR in QGP1 cell line perturbation data. The heatmap shows $\log_2$ fold changes (the value and its corresponding color) of each measurement under the respective perturbation (row annotation) (left panel). The bars represent R^2 values of each measured node (column annotation). (b) Graphical representation of QGP1 network configuration. The dotted line and solid green links show connections that were removed and added, respectively, to the initial network topology. (c) Comparison of $\log_2$ fold changes of experimental data with perturbation data. The heatmap represents the observed $\log_2$ fold changes of each measurement under a given perturbation in the experimental data (top panel), in the stimulated data based on the initial network topology (middle panel) and in the stimulated data based on the optimized final network topology (bottom panel) of figure b. (d) Quantile-Quantile plot of the model error compared to a normal distribution for QGP1. The red line shows the expected line and the dots represent the model error for each data point.

cell lines (Figure 3.21a). Thus, we first fixed the inhibitor parameter value such that the difference in the inhibition response is captured in the link coefficients emanating from the inhibited node and then determined the product of combined parameter values such that the effect, not just on the direct node but rather on the entire signaling pathway, is identified. Finally, the confidence interval was computed using the profile-likelihood approach implemented in STASNet to determine the range of values that the parameter could take without significantly altering the fit by the model.

A total of 12 parameter combinations were identified as unique to either BON1 or QGP1 (Figure 3.21a). Among these, the feedback from AKT_{T308} and ERK1 to PI3K via p70S6K were both non-identifiable parameters (Figure 3.21b). Non-identifiability is associated with the presence of a non-finite confidence interval value in one or both directions from the minimum. This means that the true value cannot be determined for a non-identifiable parameter as varying the parameter in either direction from its minimum maintained the best fit LS or if the LS does not increase over a certain threshold (Raue et al., 2009). Such parameter ambiguity arises from either insufficient observation in the model itself for example, from the non-measured nodes of mTORC1, and PI3K or is due to limited amounts and/or quality of the data. Therefore, estimation of the parameter and subsequent interpretation of its effect on the model is difficult.

As initially pointed out, regulation of CREB1 via AKT_{T308} and ERK1, dependent and independent of p70S6K, were unique to QGP1 (Figure 3.21b; yellow bar). In fact, we observed a parameter value of 0.098 with an upper bound at -0.022 and a lower bound at 0.22 for the direct activation of CREB1 phosphorylation via AKT_{T308}, while a similar parameter range was also observed for p70S6K dependent CREB1 activation (Table 6.12.1). Interestingly, the regulation via ERK1 to p70S6K also showed similar parameter values (0.046: -0.002 to 0.16). However, the direct regulation by ERK1 showed the strongest pathway value with a minimum at 0.43, and lower and upper bound around 0.22 and 0.66, respectively (Figure 3.21b). Together, one can suggest that the regulation of CREB1 is more prominent via MAPK signaling than the PI3K/mTOR signaling axis in the QGP1 cell line. Additional interesting parameters involved GSK3B phosphorylation, which showed a negative regulation via ERK1 in BON1 and a positive regulation via AKT in QGP1. Furthermore, our parameter analysis displayed a unique link from EGFR -> RAF -> MEK1 in QGP1 although not evident from our model, this is a potential link that is also present in BON1. The model did not identify this parameter in BON1, since the activity from EGFR to ERK1 upon stimulation was negligible due to the high ERK phosphorylation in ground state. Additionally, the model result holds that BON1 is less responsive to EGF than QGP1 (path value: BON1 = 0.65 vs QGP1 = 0.88) and implies that BON1 generally has have much less activity compared to QGP1.

Combined analysis of both cell lines showed that in general the stimulation as well as the negative and positive feedbacks were stronger in BON1 compared to QGP1 (Figure 3.21c).
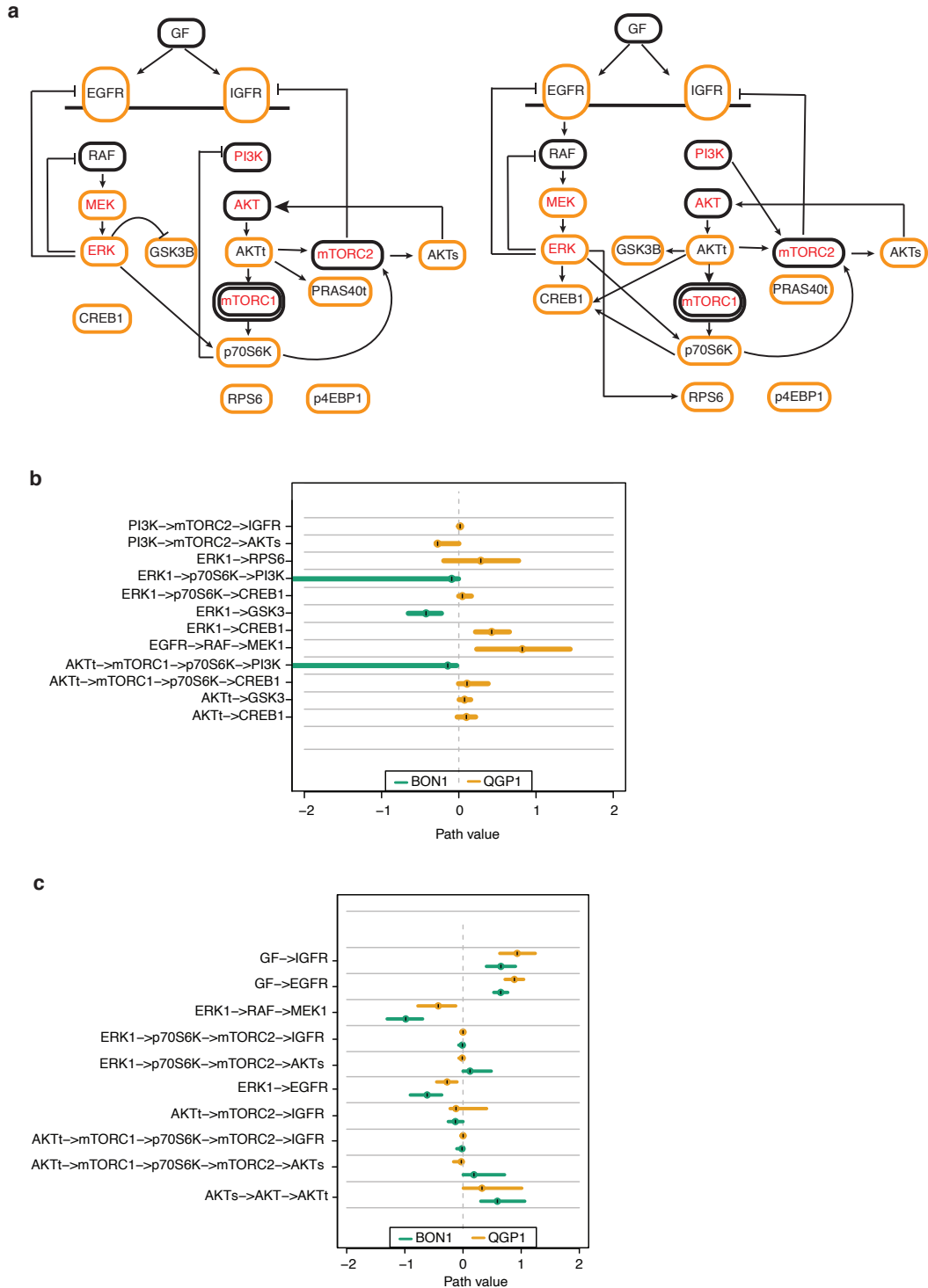
FIGURE 3.21: **Parameter evaluation of cell line models.** (a) Visual representation of shared and unique links of PanNEN cell lines. The left network represents shared and unique links of the BON1 cell line and the right network depicts the shared and unique links of the QGP1 cell line. (b) Path value and confidence interval of parameters(right) with the respective models (left). The parameter values and their upper and lower bound are depicted for unique paths in BON1 and QGP1. The y-axis shows combined paths. The lower bound beyond the x-axis limit is -inf. The colors depict the cell lines. (c) Comparison of negative and positive feedbacks shared between BON1 and QGP1. The annotation for figure c is the same as in figure b.

The stimulatory response of both IGFR and EGFR was significantly different between the cell lines, and the response coefficient of the stimulation was higher for QGP1 than BON1. In addition, parameters associated with MAPK signaling showed significant differences between the cell lines. Particularly, the coefficient associated with the negative feedbacks of ERK1 to EGFR and to RAF were higher in BON1 compared to QGP1, with the highest feedback value close to -1 (-0.98: -1.30 to -0.69) present in the ERK1 to RAF feedback (Figure 3.21c; Table 6.12.1). Major differences were also observed in the positive feedbacks on mTORC2 activity originating from AKT_{T308} and ERK and disseminated via p70S6K (Figure 3.21c). Both feedbacks were stronger in BON1 in comparison to QGP1, whose optimal parameter spread was around 0 for both the links. Together, the difference in cell lines can be attributed to their unique parameters, suggesting additional and distinct regulations within the respective cell line and to their feedback parameter potential, which can further define the vulnerability of the cell lines to an overall perturbation.

Chapter 4

Discussion and outlook

The advent of multi-platform high throughput data has generated the prime opportunity to study the divergence of cancer entities in an in-depth manner. In fact, an integrative approach of molecular sequencing data has revealed that one of the major factors influencing distinct aberrations, leading to diverse oncogenic processes, can very well be attributed to the differences in the originating cell (Hoadley et al., 2014; Hoadley et al., 2018). This can be clearly seen in Squamous tumors that clustered together due to their consistent 3p loss and 3q gain. In addition, clustering based on hypermethylation, despite the exclusion of tissue-specific loci, resulted in the aggregation of organ-specific tumors (Hoadley et al., 2014; Hoadley et al., 2018).

Based on the current literature knowledge of molecular aberrations of PanNENs, comprising of minimal features associated with the PanNEC subtype, a clear divergence is seen at the mutational level between the well-differentiated PanNETs including NETG1, NETG2, and NETG3, and poorly differentiated PanNECs. However, the lack of alterations associated with DNA methylation and copy numbers in PanNECs remains a major gap in providing a comprehensive profile distinguishing these sub-classifications of PanNENs. The knowledge gap hampers precise clinical diagnosis of either NEC or NETG3 tumors, which is of high clinical relevance due to their direct influence towards poor prognosis among PanNENs. They show histopathological differences compared to low- and intermediate PanNETs yet maintain low morphological distinction from one another.

Leveraging on the idea that the cell-of-origin has a major influence on the diverse molecular features associated with distinct cancer entities, we explored two major questions: 1) does the cell-of-origin features distinguish the well-differentiated PanNETs from poorly differentiated PanNECs? 2) Can we subsequently identify unique genetic and methylation characteristics that can impact the precision in diagnostic and therapeutic approaches of these subtypes?

4.1   PanNEN tumors have three distinct methylation subgroups

DNA methylation at CpG dinucleotides is a mechanism of cell-type specific gene regulation, inherited in a continuous manner throughout development; hence it is a stable marker of cell identity (Dor and Cedar, 2018). Methylation patterns are now considered as a robust method to identify the tumor cell-of-origin across tissues types (Espinet et al., 2020), and in different cancer types (Bormann et al., 2018; Fernandez et al., 2012; Moran et al., 2016; Holm et al., 2016; Gaiti et al., 2019). In addition, they have been used for characterization of subgroups within a tumor entity, for example in central nervous system tumors (Capper et al., 2018), breast cancer (Holm et al., 2016), ovarian cancer (Lo Riso et al., 2020), as well as in the tissue of origin of cell-free DNA (cfDNA) in disease (Moss et al., 2018). Given the paramount influence of the DNA methylome on patient subgrouping within a cancer entity and across cancer entities, we employed an unsupervised class discovery approach to identify PanNEN tumor subgroups within our cohort and subsequently demonstrated distinct cell-of-origin features defining the identified subtypes.

Tumor DNA methylation-based classification modeled PanNEN tumors into three methylome subgroups. The largest two groups, Groups A and B, harbored the majority of well-differentiated PanNETs and poorly differentiated PanNECs, respectively (Figure 3.1a). Our classification feature included the 10K methylation probes that displayed the strongest methylation variability between Group A and Group B, thereby outlining an unbiased methodology that can discern PanNETs, particularly NETG3, from PanNEC tumors. Interestingly, our approach additionally captured a "mixed"group, Group C, carrying 4 samples, comprising of 1 NETG2, 2 NETG3 and 1 NEC tumor. This further defines a set of histologically indistinguishable tumors that shows methylation features unlike the tumors of similar grade. The survival among patients of distinct groups supported our methylation-based classification of PanNET tumors of Group A and PanNEC tumors of group B, but also identified our Group C tumors with a prognosis in-between Group A and Group B (Figure 3.1f). Nevertheless, given the low number of samples in Group C, the identity of these tumors is still a major speculation and cannot be reliably stated. In fact, to establish the methylation-based classification system for improving diagnosis of PanNENs, a validation cohort encompassing all 4 grades and an overall larger number is required. Training our classifier using our current cohort, and subsequently predicting the subgroups of the validation cohort, is vital to build such methodology as a diagnostic backbone in a clinical setting.

4.2   Group A and Group C are comprised of endocrine-like tumors

Within chapter 1, the current progress in identifying cell-of-origin of PanNEN tumors was clearly demonstrated. Three recent studies in fact explored the potential endocrine origin

of PanNET tumors based on the normal cell epigenetic features retained in the tumor population (Cejas et al., 2019; Domenico et al., 2020; Lakis et al., 2021). Cejas *et. al.* investigated the histone acetylation associated with normal cell identify in combination with transcriptomes in a cohort of NETG1 and NETG2 samples. On the other hand, Di Domenico *et. al.* and Lakis *et. al.*, utilized the DNA methylation 450K platform and independently determined the cell-of-origin of PanNET tumors from cohorts of 125 and 84 well-differentiated PanNET samples, respectively. Although both cohorts were largely comprised of NETG1 and NETG2, Lakis *et. al.* additionally included 2 NETG3 tumors. All three studies unanimously confirmed α- and β-cells as potential normal cell precursors of PanNET tumors. In fact, they identified α-like, β-like and intermediate-like tumors within their cohort, confirming the pancreatic endocrine origin for PanNET tumors.

Evidently lacking from their approach are potential cell-of-origin characteristics defining the most aggressive subtypes of PanNENs: NETG3 and NEC tumors. The current thesis aim to close this gap by including not only these tumor subtypes, but also by interrogating the methylation profile using the 850K EPIC beadchip platform, thereby significantly increasing the number of CpGs investigated.

The identification of α-like, β-like and intermediate-like PanNET tumors was reaffirmed by our study. The top DMPs associated with pancreatic cell markers belonged to *IRX2* and *NKX6-1*, which were strictly hypomethylated in Group A, indicating an endocrine origin for these tumors (Figure 3.9a). As detailed in chapter 1, and reiterated in section 3.4, IRX2 and PDX1 expression characterized α-like and β-like tumors, respectively (Cejas et al., 2019; Domenico et al., 2020). In fact, A closer look at their methylation pattern in Group A clearly separated the samples into an α-like subgroup enriched for *PDX1* hypermethylation and *IRX2* hypomethylation, carrying mean proportions of >55% α-cell signature, and β-like tumors that carried >75% β-cell signature (*PDX1* hypomethylation and *IRX2* hypomethylation) (Figure 3.10b; Figure 3.14c).

Interestingly, the intermediate subgroup of endocrine-like tumors of Group A showed a cell marker methylation profile analogous to β-like tumors; they were *PDX1* hypomethylated and *IRX2* hypomethylated tumors. With respect to the normal cell signature proportion, however, intermediate-like tumors showed equal proportions of α-, β-cell signature, thereby diverging from the remaining Group A tumors. Given the proportion of α-, β-cell signatures in intermediate-like tumors, there are two possibilities of origin: during development, α- and β-cells diverge from a common ancestor that shows properties of both cell types. Such a progenitor can indeed be the actual normal precursor of intermediate-like tumors as they would theoretically carry an equal representation of both α and β cell methylation signatures. Adult pancreas does not carry an endocrine progenitor in the islet of Langerhans though, which therefore points to the second possible origin: α-cells themselves. The plasticity of α-cells to transdifferentiate and regenerate β-cell populations has

been extensively studied (Xiao et al., 2018; Lu, 2014). Particularly, their de-differentiation to a naïve endocrine progenitor state unfolds the possibility of this cell type, with both α- and β-cell features, giving rise to intermediate-like tumors. These pro-α-cells have the capacity to switch between the production of glucagon and glucagon-like peptide 1 (GLP-1) transcribed from the same gene *Gcg* depending on the abundance of prohormone convertase PC2 or PC1/3. Under certain conditions, including beta-cell deficiency, adult α-cells can tip the balance towards PC1/3 abundance and result in the production of GLP-1, thereby maintaining an un-differentiated pro-α-cells state. In fact, the earliest endocrine progenitor cells are GLP-1 producing pro-α-cells that can give rise to α-cells and β-cells depending on the relative expression of ARX and PAX4, respectively (Habener and Stanojevic, 2012; Zhang et al., 2019; Lee et al., 2018). Further, single cell analysis has identified a population of proliferating α-cells within the adult pancreas (Segerstolpe et al., 2016). This could be the subpopulation carrying the GLP-1 protein. At the RNA expression level, however, they are indistinguishable with respect to Glucagon or GLP-1 production.

Group C tumors are also reflective of an endocrine feature (Figure 3.10b). The mixed Group C tumors included all three cell-of-origin characteristics of endocrine precursors, suggesting both Group A and Group C are tumors of endocrine cell-of-origin.

The cell marker expression associated with the normal cell type was absent in Group A tumors (Figure 3.11). The α-like tumors showed ARX^+ $PDX1^-$ expression in 12/13 stained samples. Although the expression profile resembles that of an α-cell, this was not a unique feature of α-like tumors. The β-like tumors were diagnosed as insulinomas (samples PNET 25, 52, 91), but on the protein level showed an ARX^+ $PDX1^-$ profile. In addition, the intermediate-like tumors included PDX1 and ARX double negative (PNET 50 and 18) tumors, PDX1 positive only (PNET 51) and ARX positive only (PNET 14) further highlighting that methylation is a more consistent marker of lineage, and more likely to expose the cell of origin than expression of lineage specific markers. The discrepancy of methylation and protein presence in these tumors can be attributed to the additional regulations taking place within the genome including cell-type-specific chromatin marks or transcriptional programs (Hoadley et al., 2018).

4.3 Group B tumors display exocrine-like properties

It is important to emphasize that unlike the aforementioned studies which showed distinct clusters for α-like, β-like and intermediate-like tumors, our study clustered the endocrine tumors together, irrespective of the cell type from which they originated from. This is highly influenced by the presence of Group B tumors which were lacking endocrine features. Phylo-epigenetic analysis exposed the tight clustering of PanNEC samples with exocrine cells and additionally highlighted their strict separation from endocrine cells and

endocrine-like PanNET tumors (Figure 3.13). PanNECs have been repeatedly compared to PDACs in terms of the mutational spectrum (Konukiewitz et al., 2018). Therefore, there have been speculations in the field as to whether PanNECs are in fact of ductal origin. The key difference between these two tumor entities became apparent at the cell-type similarity level whereby PDACs carried equal proportions of acinar and ductal-cell signatures, whereas NEC tumors showed the highest proportion of the acinar cell signature. Given their lack of strong endocrine features, and enrichment of the acinar cell signature, we hypothesize that NEC tumors of Group B are acinar cell-like tumors of exocrine origin.

A recent study by Kopp *et al.* (Kopp et al., 2012) showed that SOX9 accelerated the formation of precursor lesions of PDAC when co-expressed with oncogenic KRAS. By lineage tracing, their study also suggested that upon the expression of *SOX9*, PanIN lesions and subsequently pancreatic ductal adenocarcinomas arise from ductal metaplasia of the pancreatic acinar cells, a phenomenon known as acinar-to-ductal metaplasia (Kopp et al., 2012). Apart from its critical role in tumor formation from acinar cells, SOX9 is a crucial factor, initially maintained in the multipotent progenitor state (Cano et al., 2014) in order to regulate pancreatic cell development. At more advanced developmental stages, the expression is subsequently restricted to NKX6.1$^+$ bipotent progenitor cells and later, in adult pancreatic cell types, SOX9 is constrained to ductal and centro-acinar cells of the pancreas (Seymour et al., 2007; Arda, Benitez, and Kim, 2013). We detected SOX9 in 81% (9 /11) of PanNECs available for IHC in Group B and 60% (3/5) of NETG3 samples in Group A (Figure 3.11 and (Figure 3.12). Interestingly, Group B samples carried a SOX9$^+$, ARX$^-$ and PDX1$^-$ phenotype. In line with the aforementioned findings in PDAC, our data led us to hypothesize that Group B samples are acinar-like tumors, which develop by a mechanism similar to that of PDAC formation via the expression of *SOX9*. Although SOX9 expression was largely absent in the Group A tumors, presence of SOX9 was evident in NETG3 tumors. Regardless, the distinction between the NETG3 Group A tumors and the Group B tumors is based on the fact that NETG3 SOX9$^+$ samples are α-like tumors while Group B tumors are acinar-like tumors. Since SOX9 plays a critical role in the multipotent and bipotent state of pancreatic development, it is likely that SOX9$^+$ NETG3 tumors may originate from endocrine cells, given their similarity to α-cells and, in the course of tumor progression, revert to expressing SOX9 as a mechanism to move towards a progenitor-like phenotype.

Evidently, the acinar cell signature of NEC tumors is significantly more abundant when compared to the endocrine-like tumors (mean of 0.38 of acinar-cell proportion vs 0.25 of α- and β-cell proportion). However, it is also noteworthy that the mean proportion of α- and β-cell signatures is maintained around 0.25 and not far behind the observed mean proportion of the acinar cell signature (0.38). This certainly indicates the possibility of a potential progenitor resembling an early ancestor of pancreatic development as the source of the precursor cell. Group B tumors also mirrored the naïve hESC methylation profile, thus

further establishing the feasibility of the progenitor-like precursor. Recently, novel, multipotent progenitor-like cells were identified in the major pancreatic ducts (MPDs), which were PDX1$^+$/ALK3$^+$ cells that maintained their de-differentiated state under the presence of BMP7 all the while maintaining the expression of the ductal cell marker SOX9 (Qadir et al., 2018). The single-cell RNAseq analysis of the ALK3$^+$ sorted cells determined subpopulations within these progenitors that carry lineage differentiation potential including pro-ductal, pro-acinar and ducto-endocrine axes (Qadir et al., 2020). Evidently, multilineage potential could also be embedded in the methylation profile of these progenitor-like cells and could therefore include relatively equal distribution of acinar, α- and β-cell signatures. Since the ductal component is absent within the tumors included in the cohort studied here, it is likely that the pro-acinar or ducto-endocrine SOX9$^+$ subpopulations might indeed be the normal precursors to the exocrine-like NEC tumors. Validation of this hypothesis would require first obtaining ALK3$^+$ sorted progenitor-like cells and perform single-cell DNA methylation profiling combined with transcriptome (scNMT-seq) analysis to determine the methylation signature associated with each of the progenitor axes (Clark et al., 2018). Subsequently, the profiles could then be deconvoluted to see the proportion of these progenitor signatures captured within the PanNEN cohort.

4.4 Distinct mutational and copy number profiles characterize endocrine-like and exocrine-like tumors

The mutational characteristics of α-like tumors identified by Chan *et. al.* is reproduced in our results, where most samples with *ATRX*, *DAXX* and *MEN1* (A-D-M) mutations belonged to α-like tumors (Chan et al., 2018). The strict enrichment of the *MEN1* mutation in α-like tumors might be attributed to the specificity of *MEN1* alterations in the formation of tumors from α-cells. Various mouse model studies have identified that MEN1 depletion from α-cells surprisingly resulted in the trans differentiation to β-cells and subsequent formation insulinomas (Lu et al., 2010; Shen et al., 2010). One of our β-like tumors diagnosed as insulinoma indeed carries a *MEN1* mutation. Therefore, based on the mouse model studies, it might be possible that this tumor in fact progressed from an α-cell and not from a β-cell. This also has additional positive implications towards the intermediate-like tumors, as the methylation profile embodying the final trans-differentiated cell type additionally affirms the theory of a pro-α-cell giving rise to intermediate-like tumors. Since the majority of the *MEN1* tumors maintained their α-cell signature, one can speculate that the presence of additional aberrations prior to *MEN1* is in fact forcing the maintenance of α-like profiles within these tumor entities.

Interestingly, the four VHL mutations identified within our cohort belonged the β-like

and/or intermediate-like tumor group carrying hypomethylation of *IRX2* and *PDX1*. Although the function of VHL protein in regulating β-cell function and maintainance of the β-cell mass has been explored, its direct influence on tumor formation from β-cells is yet to be investigated (Choi et al., 2011; Puri, Cano, and Hebrok, 2009).

Alpha-like tumors showed distinct CNA profiles including an amplification-rich signature, a low-CNA signature and a deletion-rich signature. The amplification-rich signature and deletion-rich signature were present largely in NETG2 and NETG3 tumors. Low CNA profiles, on the other hand, were enriched not only in NETG1 α-like tumors but also in β-like and the intermediate-like tumors of Group A. With respect to molecular timing of the events, a recent study has revealed that clonal copy number losses particularly of chromosomes 2, 6, 11 and 16 are among the earliest events followed by late clonal synchronous events of chromosomal gains in chromosome 5,7 and 19, as well as *MEN1* and *DAXX* mutations (Gerstung et al., 2020). Given the copy number profile as well as recurrent mutations, it is likely that the deletion-rich signature associated α-like tumors are early in the tumor progression timeline while the amplification-rich signature tumors of an α-cell of origin, with enrichment of *MEN1* mutations, are later in the molecular clock of cancer formation.

Acinar-like tumors were characterized by recurring mutational aberrations in *KRAS*, *SMAD4* and *TP53*. This is in line with the previous findings of the NEC mutation spectrum detailed in chapter 1. Mutations in *DAXX* and *ATRX* were not present in the Group B tumors, however, a single *MEN1* aberration was identified within this subgroup. The sample also carried a *SMAD4* mutation; therefore, it is likely that the tumor progression to the NEC grade is rather influenced by the latter aberration. Using focal DNA copy number analysis, we found that Rb1 loss, evident in NEC tumors at the protein level (Yachida et al., 2012), was due to DNA copy number loss of chromosome 13. This recurrent focal aberration was present in 50% of the PanNECs but not in any NETG3 samples, additionally highlighting yet another genomic feature distinguishing these high-grade tumors. In contrast, cell cycle regulators *CDKN1B* and *CDKN2A* were lost in Group A PanNET samples. These important driver events in NETs and NECs indicate a dependency of both tumor entities on cell cycle regulation pathways.

Group C tumors lacked mutational aberrations associated with the targeted panel genes. This might suggest that they are not driven by mutations but rather the tumor progression is influenced by other molecular landscapes that are not explored in the context of this project. However, one must also rule out a possibility of technicality. That is, the targeted sequencing only covered a total of 482 genes while the human genome encodes 50,000 protein coding genes. Since the Group C tumor subgroup showed a strong divergence at the methylation level in comparison to tumors of similar grade, it is also probable that they

might carry non-conventional mutations. Additional mutational profiling using whole exome sequencing of these tumors can indeed provide more clarity. Group C also displayed a methylation pattern of CpGs associated with cell markers resembling a mixed Group A and B profile. Although they remain close to Group B samples in the phylo-epigenetic analysis, their similarity to acinar cell signatures is not comparable to that found in Group B and they seem to show an increased proportion of endocrine cell signatures. Further identification of the Group C cell of origin and their genomic characteristics are limited due to the small number of samples.

4.5   Therapeutic strategies adapted for unique genomic characteristics are essential for endocrine-like and exocrine-like tumor

The distinct molecular landscapes of the endocrine-like PanNET tumors and exocrine-like PanNEC tumors is strongly indicative of separate directions of drug treatment (reviewed by (Scarpa, 2019)). The current state-of-the art targeted treatment strategy approved for well-differentiated PanNET tumors includes the FDA approved drugs Sunitinib and Everolimus, targeting angiogenesis and mTOR signaling, respectively (see chapter 1 for details). The molecular knowledge informing the use of these drug targets for treating low- and intermediate-grade PanNETs comprises few genetic aberrations within the components of these signaling axes. Our cohort harbored 37% mTOR/PI3K axis associated aberrations including mutations in *TSC2*, *PTEN*, and *PI3K* component genes and copy number aberration of 22q12.3 resulting in the deletion of *DEPDC5* gene. Recurring mutations in *VHL* and a single event in *HINF1A* were also present in our cohort, further promoting the involvement of angiogenesis in PanNET tumors. Due to the relatively low number of aberrations, the benefits of treatment with Sunitinib are potentially exclusively effective for a handful of patients and much less than those that could benefit from Everolimus. Therefore, it is imperative that the patients are initially screened for these recurrent aberrations prior to use of Everolimus and/or Sunitinib, which is currently not undertaken.

The recurrent mutational and focal copy number aberrations of Group B tumors suggest the prospect of different treatment regimens unlike that of the Group A tumor subgroup. Group B tumors were defined by *KRAS* mutations, and it was altered in 36% of the subgroup. Similarly, loss of *RB1* and mutations in *TP53*, alterations that impinge on cell cycle regulation, were found in 71% of Group B tumors. Given the highly proliferative nature of this aggressive tumor subgroup, the optimal therapeutic strategy would therefore include targeting these highly recurring molecular aberrations. Conventional targeted inhibitors are yet to be developed for *TP53* mutations and its associated pathways. However, targeting loss of *RB1* is certainly being investigated. The canonical pathway of the RB1 protein

results in the regulation of G1 to S phase transition. The protein acts as a checkpoint of the transition state and under mitogenic signals, it is repressed by CDK4/6, thereby allowing for the expression of E2F target genes and subsequently the progression of the cell into the S phase of the cell cycle (Sherr, 1996). Since RB1 is deleted in Group B tumors, one can indirectly interfere with the process by targeting the CDK4/6 proteins using the highly specific inhibitor PD0332991, thus forcing the cell to an early cell cycle exit and thereby reducing the proliferative activity of the tumor cells (Fry et al., 2004; Knudsen et al., 2019).

4.6 Optimal therapy for exocrine-like tumors is deciphered by exploring signaling dependencies of in-vitro models

A common assumption that vastly flawed in the context of cancer, is being made when considering such linear approaches: Alteration in gene A results in the aberration of protein B and thereby deregulate process C. Compensatory mechanisms including proteins with similar function, or crosstalk to another signaling pathway can in fact counteract the treatment regimen based on theoretical linear relationships and lead to therapy resistance. For instance, CDK2 in a normal cell type can compensate for the inhibition of CDK4/6, thus releasing the Group B tumor cells back to a more proliferative state. This can be overcome by combinatorial treatment targeting the oncogenic signal propagated through oncogenic signaling pathways such as MAPK or mTOR signaling (Knudsen et al., 2019). Such a combinatorial treatment, however, again generates the paradigm that the targeted pathways are linear and are not interconnected with additional signaling axes.

We therefore explored the potential of a cell line model to study signaling crosstalk and the cells' vulnerabilities to combinatorial therapy. Based on the genomic aberrations, the QGP1 cell line was deemed an appropriate candidate, as it carries the signature aberrations of Group B tumors including *KRAS*, and *TP53* mutations as well as RB1 loss. BON1 cell lines also carry mutational profiles indicative of Group B tumors, but lack the copy number aberrations associated with them. Nevertheless, the DNA methylation profile suggests that both cell lines can be utilized as genetic models of Group B exocrine-like tumors.
We modeled the signaling axes of MAPK and mTOR/PI3K pathways as the initial step to identifying crosstalk and vulnerabilities reflective of Group B tumors. Interestingly, both cell lines displayed a crosstalk between mTOR signaling and MAPK signaling via PI3K. However, the differences were largely retained in the feedback potential of the cell lines. The negative feedbacks in the respective pathways were highly diminished and the positive feedback mechanisms were amplified in the QGP1 cell lines. As a result, we observed a strong activation of the downstream signaling in QGP1, in contrast to BON1 where the feedback has much more control over the activity of the signaling nodes. In the

context of inhibitors, combinatorial inhibitor treatment targeting PI3K and ERK signaling should hamper the overall activity of all three signaling axes. The combination of PI3K and ERK, however, is not ideal as it might cause toxicity in patients. Therefore, one can consider MEK1 inhibition in combination with IGFR. In the future, the models need to be systematically and quantitively assessed to determine best combinatorial treatments with the strongest impact on the signaling pathways.

In order to translate these novel insights to the clinical setting, it is crucial to validate the presence of such vulnerabilities in the tumor cohort. The cell lines cultured in a two-dimensional environment for many years can certainly modify the dependencies of the signaling pathways. Therefore, RNA sequencing profiles of the Group B tumors in comparison to Group A need to be generated and differential gene expression, in the context of pathways needs to be investigated. The well established tools including SPEED2 or Enrichr can be utilized to determine differentially activated pathways (Rydenfelt et al., 2021; Kuleshov et al., 2016). Additionally, one could also utilize the PROGENy tool to determine pathway activity from basal expression (Schubert et al., 2018). This can further validate if the pathway regulations within the in-vitro cell lines are indeed representative models of the tumors. Further, additional vulnerabilities can be deduced from combinatorial treatment simulations as well as by generating additional knockout experiments of key nodes in the pathway to eventually decipher the optimal targeted approach for therapy.

4.7   Concluding remark

The advancement of next generation sequencing and high-throughput data analysis has certainly impacted how tumor entities are viewed in clinical translational research. Comprehensive profiling and accordingly strategizing the diagnostic and therapeutic interventions are increasingly becoming the norm. For instance, methylation-based profiling underlined by cell-of-origin features have dramatically improved diagnosis of CNS tumors (Capper et al., 2018) and identified diversity in cell of origin for colorectal cancers (Bormann et al., 2018). The distinction in prognosis among patients along with specificity in genetic and methylation features have also proposed more stratified therapeutic approaches for these cancers.

Unfortunately, such comprehensive profiles have been largely confined to the most frequently occurring solid and blood cancer entities. The neglected rare tumor entities, on the other hand, are lagging several years behind in such genetic profiling. In fact, molecular characterization of PanNEN tumors have only recently taken a minor step forward and mainly encompassed the low and intermediate NETG1 and NETG2 tumor grades. This is of course reflected in the few clinical trials that have investigated the potential of targeted

therapy wherein, the efficacy and responses of Everolimus and Sunitinib are explored for NETG1 and NETG2 tumor grades. Even with the current genomic knowledge attained for these grades, a trial is yet to be performed that takes stratification measures into consideration prior to testing. Recent evidence shows that apart from the recurrence of key genetic drivers, the differences within and between the low and intermediate grade tumors that are directly associated with patient prognosis are also attributed to cell-of-origin from the specific pancreatic endocrine cell type.

Alarmingly, the most aggressive PanNEN tumors with a much higher mortality rate, which includes NETG3 and PanNECs, are overlooked in these comprehensive studies. The strong similarity between these tumors additionally often lead to misdiagnosis, highlighting the importance of bridging this gap and defining the molecular profiles that can then be further exploited for accurate diagnosis as well as the appropriate treatment regimens entailed towards targeting tumor characteristics.

Our study has moved the field a step forward in expanding the molecular characterization of the highly understudied PanNEN subtypes, NETG3 and PanNEC. We have demonstrated that NETG3 tumors have genetic and methylation features characteristic of the low and intermediate NETG1 and NETG2 PanNET tumors and additionally, originate from the endocrine cells of the pancreas. In contrast, PanNEC tumors harbor genetic, and methylation features highly dissimilar to PanNET tumors and demonstrate exocrine-like cells as a potential normal precursor. They show significant resemblance with the PDAC, an exocrine tumor, compared to the PanNET tumors, emphasizing not only the need to treat these PanNEN subtypes as a distinct entity but also consider exploring the treatment strategies more in line with what is currently available and investigated in PDAC patients.

Ultimately, cancer is a disease state that cannot be looked at in parts but rather as a melting pot of disrupted molecular hierarchies. Understanding the major changes in these molecular states is essential to identify how the tumor developed from a normal cell and what led to its progression, thereby establishing adequate diagnosis and treatment approaches. This project has shed light on the endocrine and exocrine cell of origin for PanNEN tumors distinguishing the well-differentiated PanNETs, and particularly the high grade NETG3 from the poorly differentiated and aggressive PanNEC tumors. The intrinsic composition of the methylation profiles, in combination with diversity in genomic aberrations, advocates for treating each of these subgroups as its own entity and tailoring therapeutic approaches towards the distinct molecular profiles which drive tumor progression.

Chapter 5

Materials and Methods

5.1 Patient cohort representation

The patient cohort associated with section 3.1 through section 3.4 consisted of 57 Pan-NEN samples collected from 55 patients. Primary tumor and metastasis of a single patient (PNET77) were obtained and two metastases, arising one year apart, were also obtained for another patient (PNET56). The Institute of Pathology at Charite University of Berlin provided 48 samples of all grades and the University of Bern provided 9 NETG3/NEC samples. Together, these included 41 primary tumors from pancreas and 16 metastases. The metastases of the cohort comprised of 11 liver metastasis, 2 lymph node metastasis and one metastasis each in the bladder, papillary and peritoneal regions. The resected primary tumors were located in the head (29.8%), tail (24.6%) and body (3.5%) of the pancreas and two samples had lesions in multiple sections of the pancreas. Clinical reports on the patients were collected together with tumor associated reports, which are presented in Table 6.13.1 and Table 6.14.1, respectively. All tumor samples were obtained as Formalin-fixed paraffin-embedded (FFPE) blocks. With the exception of 10 cases, normal controls for the respective patients were additionally acquired. Normal tissue sections were either tissue adjacent to the tumor (as per the pathologist's examination; normal adjacent n= 15), a separate block containing only normal tissue (normal distant n = 29) or peripheral blood samples (n = 3). All patients provided signed consent as part of the clinical documentation protocol of the Charité University in Berlin. Samples of the University of Bern were provided by the Tissue Biobank Bern (TBB) according to the relevant ethics approvals.

5.2 PanNEN panel design

The first part of the PanNEN gene panel design was based on a text-mining approach to extract high-quality information regarding genes associated with PanNENs from GeneView (Thomas et al., 2012), the Catalogue of Somatic Mutations in Cancer (COSMIC) (Bamford et al., 2004) and few PanNEN publications (Cao et al., 2013; Jiao et al., 2011; Yuan et al.,

2014; Banck et al., 2013). This was done in collaboration with the Department of Knowl-
edge Management in Bioinformatics at Humboldt University. From this preliminary list,
47 PanNEN-likely driver genes were then curated with further focus on MAPK and mTOR
signaling pathways. Subsequently, amplicons were designed with GRCh37 genome refer-
ence using the Ion AmpliSeq Designer tool (Life Technologies) for the 47 candidate genes
provided under the criteria "DNA Gene design (multi-pool)". The panel was designed to
generate primers targeting 125-bp stretches of exon regions of the selected genes. The final
panel design included 1175 amplicons, which were divided into two primer pools.

5.3   DNA isolation

DNA isolation and the subsequent DNA sequencing was performed by Andrea Menne the
Molecular Tumor Pathology, Group of Prof. Dr. Christine Sers at the Charite University
of Berlin. Briefly, FFPE Tissue samples, including embedded biopsies, were sectioned and
stained with Hematoxylin and Eosin (H & E). Pathologists demarcated tumor and healthy
tissue areas in the H &E slides. Depending on the size of the marked area, 12 sections
of 5 μm each from tumor samples and 6 sections of 5 μm for adjacent or distant normal
tissue were used for DNA isolation. The tissues were macro-dissected from the slides
and DNA was prepared using the GeneRead DNA FFPE kit (Qiagen). Quality and quan-
tity of DNA was determined by RNAseP quantification as per manufacturer's instructions
(Thermo Fisher Scientific).

5.4   DNA sequencing

DNA libraries using 20ng of DNA were prepared with Ion Ampliseq Library kit (Thermo
Fisher Scientific). Libraries included regions targeted by the panel primers, distributed into
two amplicon pools per DNA sample for the PanNEN panel and 4 amplicon pools per DNA
sample for CCP. Upon ligation to Ion Xpress Barcode Adapters (Thermo Fisher Scientific)
and purification using Agencourt AMPure beads (Beckman Coulter), two samples were
mixed at equal ratio on a 318v2 sequencing chip. Using the Ion Torrent PGM (Thermo
Fisher Scientific), raw intensity data were generated upon sequencing at an average depth
of 1158 reads using PanNEN panel and an average read depth of 217.03 reads for CCP
panel.

5.5   Sanger sequencing

To validate results from the panel sequencing we performed Sanger sequencing using 5-
10ng/ul DNA, as per the manufacturer's protocol. The sample preparation was performed

by Kerstin Wanke-Möhr the Molecular Tumor Pathology, Group of Prof. Dr. Christine Sers at the Charite University of Berlin. Primers were obtained from the panel primer pool and 10 µM were added to generate the Sanger mix, which was subsequently, sequenced at Eurofins Scientific facility (Eurofins Scientific). Validation was performed for specific mutations which showed sub-optimal amplicon performance, were located within long nucleotide repeat areas or harbored low allelic frequency. The resulting signal intensity images were manually scanned to identify the targeted mutations.

5.6 Fluorescence In-Situ Hybridization (FISH)

Fluorescence in situ hybridization (FISH) was performed by Andrea Menne from the Molecular Tumor Pathology, Group of Prof. Dr. Christine Sers at the Charite University of Berlin. Three µm tumor sections from 23 samples were obtained. We used commercially available, standardized centromere and gene probes for detecting chromosome 5 and *RICTOR* gene, chromosome 9 and *TGFBR1* gene, and chromosome 11 and *MEN1* gene (all FISH probes obtained from Empire Genomics). Hybridization was performed according to manufacturer's instructions. We scored a maximum of 40 cells per sample using an Olympus microscope (Olympus) upon manually evaluating cell structure and signal intensities. Analysis was conducted using BioView solo image analysis (Abbott Molecular).

5.7 DNA methylation

DNA methylation profiling of all PanNEN cohort samples was performed with 200–500ng of DNA using the Infinium MethylationEPIC BeadChip array (850K; Illumina) according to the protocols provided by the manufacturer. The data was generated by the Department of Neuropathology, Group of Prof. Dr. David Capper at the Charite University of Berlin. Normal cell type HumanMethylation450K data of Pancreatic α, β, acinar and ductal cells were obtained directly from the published lab or GEO ((Neiman et al., 2017), GSE122126 and GSE134217). Published 167 PDAC Infinium HumanMethylation450K data and Infinium MethylationEPIC data of human Embryonic Stem Cells (hESC) were additionally acquired from GSE49149 and GSE128130, respectively.

5.8 Cell culture and phosphoprotein assay

The phosphoprotein data generation was performed by Dr. Pamela Riemer and Kerstin Wanke-Möhr from the Molecular Tumor Pathology, Group of Prof. Dr. Christine Sers at the Charite University of Berlin and Dr. Anja Sieber from the group of Prof. Dr. Nils

Blüthgen of the Integrative Research Institute (IRI) for Life Sciences at Humboldt University, Berlin. Briefly, the QGP1 cell line was cultured in D10 medium containing Gibco Dulbecco's Modified Eagle Medium (DMEM) and 2mM Ultraglutamine, supplemented with 10U/mg/ml Penicillin and Streptomycin, and 10% Fetal calf serum (FCS). To culture BON1 cells, 4.5g/l glucose were additionally added to the aforementioned mix. Each of the cell lines was seeded at 2×10^6 cells/10cm dish. The following day, the media was removed, and the cells were washed twice with PBS prior to 24 hr serum starvation using D10 medium lacking FCS and glucose. The cells were then treated with either an inhibitor (PI3K: GDC0941(Axon) 1 μM, AKT: MK2206(LKT Laboratories) 1 μM, Dual mTOR: AZD8055(Cayman Chemical) 100nM, mTORC1: Everolimus (Cayman Chemical) 100nM, MEK1: AZD6244(Selleckchem) 5 μM, ERK1: SCH772984(Selleckchem) 1 μM) in the appropriate final concentration, or in the case of controls, with 0.1% DMSO. After 1h of treatment, the cells were stimulated for 30 mins with both 30ng/ml EGF and 100nM IGF. For the non-stimulated controls, PBS/0.1%BSA was used instead. At least 3 replicates were prepared per condition/readout (see next paragraph for readouts) in BON1 and QGP1 including 5 inhibitor, control-stimulated and control-non stimulated conditions. The cells were then harvested, lysed with Bio-Plex Cell Lysis Kit (Bio-Rad) and the protein concentrations were quantified using a standard BCA assay.

For the Bio-Plex run, 15 μg of protein were obtained and additional preparations were performed according to the manufacturer's protocol (ProtATonce; custom generated protocol based on the phosphoprotein antibody panel). Phosphoprotein levels were assessed using the magnetic beads coupled antibodies, specific for each target site: $IGFR_{Y1131}$, $EGFR_{Y1068}$, $MEK1_{S217.221}$, $RPS6_{S235.236}$, $GSK3B_{S9.21}$, $p70S6K_{T389}$, $PRAS40_{T246}$, $CREB1_{S133}$, and $ERK1_{T202.Y204}$. The phosphoproteins were quantified with the the latest Bio-Plex system termed MAGPIX protein assay system (Bio-Rad) as per manufacturer's instructions.

For the WES capillary western system (ProteinSimple/Biotechne), the majority of the primary antibodies (pAKTt: Cell Signaling, rabbit - #2965, pAKTs: Cell Signaling, rabbit - #4060, p4EBP1 S65: Cell Signaling, rabbit - #9451) were diluted in 1:50 concentrations, while Vinculin (R&D Systems, mouse – MAB68969) was diluted in 1:30 concentrations and standard concentrations suggested by manufacturer were utilized for secondary antibodies. 1 μg/μL protein was loaded into the capillary and the run was executed using the manufacturer's guidelines (proteinSimple) under the default setting with a single change: primary antibody incubation time was increased by 90mins. Due to the conflict in the protein size for each of the AKT phosphosites, two separate runs were performed in parallel, each detecting one of the pAKT site, in addition to Vinculin and p4EBP1 phosphorylation as technical controls.

5.9 Immunohistochemistry (IHC)

Representative samples from Group A and B described in chapter 3 were subjected to Immunohistochemistry (IHC) for ARX, PDX1 and SOX9 expressions. The IHCs were performed in Institute of Pathology at Charite University of Berlin as well as Institute of Pathology at University of Bern, Switzerland. Briefly, 2.5 µm FFPE sections were used for ARX (1:1500, R&D Systems, sheep - AF7068), PDX1 (1:100, R&D Systems, mouse - MAB2419) and SOX9 (1:100, Cell Signaling, rabbit - #82630) immunostainings. Antigen retrieval was performed by heating the Tris30 buffer at 95°C for 30 minutes. The primary antibodies were incubated for 30 minutes at the specified dilutions and upon secondary antibody incubation, the staining was visualized utilizing Bond Polymer Refine Detection kit, with DAB (3,3'-Diaminobenzidine) as chromogen (Leica Biosystems). The immunostainings for each marker were performed on an automated staining system (Leica Bond RX; Leica Biosystems).

Evaluation of the staining was performed by two pathologists: Dr. med. Armin Jarosch (Institute of Pathology, Charite University, Germany: SOX9 staining) and Prof. Dr. med. Aurel Perren (Institute of Pathology, Bern University, Switzerland: PDX1 and ARX staining). Cases were identified as strong positive when a large section of the tissue was positively stained, single cell positive if few single cells within the tissue were positively stained or negative, when the entire tissue was negative for the targeted marker. The assessment for the analyzed sample is reported in Table 6.14.1.

5.10 Data processing and analysis

5.10.1 Mutational analysis

The raw DNA sequencing reads were first aligned to the GRCh37 reference genome using the Torrent Mapping and Alignment Program (TMAP; Life Technologies). A cut-off of quality > 50 nucleotides in aligned reads and a mapping quality of >4 was placed using an in-house python script. The processed bam files were then utilized for variant calling using the TS Variant Caller plugin under the "strict" setting per Ion Suite (Ion Torrent platform) parameter profiles. The generated variant call format (VCF) files of the tumor and normal samples per patient were merged and, read counts for the reference and alterated allele in all variants were extracted to determine the representation of the variants in both cases. The merging process was excluded in the 10 cases that did not have a paired normal tissue. The merged or single VCF files were subsequently annotated using the SoFIA (Childs et al., 2016) annotation framework and additionally filtered in a semi-automated manner. Variants that were positive for the following set of parameters were removed using custom R scripts: intronic and synonymous variants, 1000Genome variants with frequency

greater than 1% in the population, variants within homopolymer regions of >4 nucleotides and tumor variant allelic frequency <5%. For the tumor-normal pair samples, representation of variant allelic frequencies in matched normal tissue were manually assessed and variants carrying comparable allelic frequencies to that of the corresponding tumor were subsequently removed.

5.10.2   DNA methylation preprocessing

Raw idat files were preprocessed using a subset-quantile within array normalization (SWAN) approach provided through the R package minfi (Aryee et al., 2014; Maksimovic, Gordon, and Oshlack, 2012). Probes performing poorly in the analysis were further filtered out if they met the following criteria: probes with a detection p-value >0.01 in at least one sample, probes cross reactive to multiple sites in the genome (Chen et al., 2013), sex chromosome probes and probes containing SNPs with an allele frequency >0.01.
When comparing 450K and EPIC samples, the processing was done by first converting the EPIC platform to 450K using combineArray() and then performing the aforementioned steps. When tumor and normal data were analyzed together, normalization was done using the preprocessFunnorm() function in order to accommodate for the global variation between normal and tumor data. This also entailed removing cross reactive probes of both EPIC and 450K array. hESC data normalization was performed as depicted in Patani, *et al* (Patani et al., 2020). Briefly, we normalized the data using the preprocessNoob() function and performed the aforementioned filtering approach to remove ambiguous probes. Methylation beta values for each dataset were used as surrogates for methylation and for all downstream analysis including statistics and visualization.

5.10.3   Subgroup identification and associated analysis

Using the 10K most variable probes, determined by row(probe)-wise standard deviation (σ), DNA methylation-based subgroups of PanNEN were identified with the R package ConsensusClusterPlus() under the following parameters: maxK=12, reps=1000, pItem=0.8 and pFeature=1. The function executed agglomerative hierarchical clustering after performing 1-Pearson correlation distance. A consensus matrix carrying pair-wise consensus values was finally generated for 12 clusters. The most stable number of clusters was determined based on the cumulative distribution score curve (CDF) that reached an approximate maximum (k=3) in combination with the correlation heatmap for each cluster in the 3-k heatmap. Hierarchical clustering using the R package stats was done by first obtaining a dissimilarity matrix using euclidean algorithm and subsequently performing complete linkage clustering. tSNE Dimension reduction analysis was performed using the Rtsne package under the perplexity=8. Genes associated with the 10K most variable probes were evaluated for term enrichment of GO pathway biological processes using the enrichGO() function in the

clusterprofiler R package. The following parameters were defined to identify significantly enriched terms: pAdjustMethod = "BH" (Benjamini and Hochberg), pvalueCutoff = 0.01, qvalueCutoff = 0.05. All genes represented in the methylationEPIC array were used as background. In order to reduce generality of GO terms, the simplify() function was used. The final set of terms was curated by filtering only those that showed an adjusted p-value < 0.01 and fold enrichment of greater than 1.5. -Log_{10}P-value and a barplot was generated for the 12 most significant terms using the ggplot2 package.

5.10.4 Differentially Methylated Probes (DMP) and associated analysis

Upon extracting and assigning the samples to the identified Group A and Group B clusters, differentially methylated probes (DMP) were identified using CHAMP package (Tian et al., 2017) function champ.DMP() under the following parameters: adjPVal = 0.05, and adjust.method = "BH", arraytype= "EPIC".

Differentiated pancreatic cell markers showing sensitivity_human > 0.05 for α, β, γ, δ, ε, acinar, ductal or islet Schwann cells were curated from PangaloDB (Franzén, Gan, and Björkegren, 2019). DMP associated genes that overlapped with curated Islet cell markers were extracted and significantly enriched DMPs of Islet cell markers were extracted using the following criteria: $|\Delta beta| > 0.25$ and - Log_{10}P > 5 (Figure 3.9a). To determine closely related samples within each group with respect to the DMPs, hierarchical clustering with complete linkage was performed using beta values prior to visualization (Figure 3.9c). Probes overlapping to 450K platform run of α, β, ductal and acinar cell were additionally extracted and visualized (Figure 3.9d). Hierarchical clustering with complete linkage was performed to determine closely related cell types based on the DMP probes. Methylation values of these DMPs were obtained for tumor and normal samples to compute a euclidean distance matrix and subsequently perform complete linkage hierarchical clustering using stats R package (Figure 3.10a).

DMP between α, β, ductal and acinar cell types were identified using CHAMP package function champ.DMP under the following parameters: adjPVal = 0.05, and adjust.method = "BH", arraytype= "450K". Significant probes exhibiting $|\Delta beta|$ value > 0.2 and adjusted p-value < 0.01 were obtained. A total of 46,500 unique methylation probes were collected and defined as DMPs of normal cell types. Upon preprocessing and downstream filtering (see sub-section 5.10.2) of tumor and normal data combined, 38892 DMPs of normal cell type remained in tumor-normal matrix. Methylation beta values were obtained to calculate pearson distance using the function get_dist() from factoExtra R package. Finally, neighbor-joining tree estimation was performed using nj() function in the ape package to generate phylo-epigenetic trees (Figure 3.13).

hESC probe identification was performed using an adaptation of the method in Patani *et. al.* (Patani et al., 2020). Briefly, probes carrying a mean beta value < 0.3 across the primed

cells were defined as background matrix. Unmethylated probes of hESC were then defined as those that carry mean beta < 0.3 in both primed and naive hESC. Hypermethylated probes of hESC were determined by first calling DMPs using the background matrix. CHAMP.DMP() ran with the adjPval=0.05, adjust.method = "BH", arraytype= "EPIC". Probes carrying Δbeta < -0.1 in naive state compared to primed state were then extracted. Accordingly, upon normalization, preprocessing and filtering (see sub-section 5.10.2), the methylation values of tumor samples for unmethylated probes of hESC and hypermethylated probes of hESC were determined. Mean values of each Probe-type (unmethylated/hypermthylated probes of hESC) in each sample of the PanNEN patient cohort were computed and the distribution of these computed means in the PanNEN methylome group was visualized. In addition, mean values for NETG3 and NEC samples of Group A and Group B were extracted and additionally visualized (Figure 3.2e).

5.10.5   Normal cell signature analysis

In order to determine cell signature proportion in each sample, the methodology provided by Moss *et. al.* was utilized (Moss et al., 2018). Briefly, The reference atlas was first filtered for β, ductal and acinar cell profiles. In order to add an α-cell profile, the normal cell types were preprocessed and normalized (Neiman *et. al.*, GSE122126 and GSE134217) (see section 5.10.2), and subsequently, the alpha cells (n=3) were extracted. The mean value of each probe was calculated, and the overlap of the probes compared to the featured CpGs of reference atlas were obtained. The final matrix, defined as the pancreatic normal reference atlas, contained the methylation values for α, β, ductal and acinar cells of the overlapping probes. Euclidean distance between each sample, given the pancreatic normal reference atlas, was computed using get_dist() function from the FactoExtra R package. PanNEN and PDAC data were normalized separately as mentioned in sub-section 5.10.2, and the methylation beta value matrix was combined for subsequent analysis. The program developed by Moss *et. al*. was then employed to identify the normal cell signature proportion in samples of the PanNEN and PDAC cohorts.

5.10.6   Copy Number Aberrations (CNA)

Copy Number Aberrations (CNA) were identified from EPIC array data using the R package conumee (*conumeeR*). Upon raw preprocessing, mean value of similar CNA segments per autosomal region was obtained and a mean value per autosome was computed for each sample to determine the $\log_2$ ratio of intensities across each chromosome. A cut-off of x > 0.15 and x < -0.15 was placed based on linear regression analysis in Figure 3.7c to limit the number of false positives. To determine the whole chromosomal signature within each subgroup, euclidean distance-based complete linkage hierarchical clustering was also performed on the data. To determine focal aberrations, $\log_2$ values of chromosomal segments,

determined by conumee were obtained for the samples and GISTIC software was executed separately for each group under the following parameters: -genegistic 1, -smallmem 1, -broad 1, -brlen 0.5, -conf 0.90 -armpeel 1 and -gcm extreme. GISTIC was only performed in Group A and Group B, and not for Group C, due to the limited number of samples.

5.10.7 Phosphoprotein data analysis

Signals obtained from Bio-Plex manager software (BioRad), acquired as .lxb files, were processed with an in-house custom script using lxb R package. The resulting file consisted of the median Fluorescence Intensity (FI) for each readout under all perturbation/treatment conditions. Quantification of the phosphoproteins for WES runs was based on the observed area under the curve at each of the expected sizes: $p4EBP1_{S65}$ (29kDa), pAKT (60kDa) and Vinculin (107.5 kDa), which was obtained and evaluated using ProteinSimple Compass Software (ProteinSimple/BioTechne). The resulting file contained the phosphoprotein quantity for each replicate of the readout for all inhibitory conditions. Quality of the replicates was assessed by transforming the data to M (log_2 ratio) and A (mean average) scales as follows:

$$M_{xy} = log_2x - log_2y \tag{5.1}$$

$$A_{xy} = \frac{log_2x - log_2y}{2} \tag{5.2}$$

Where x and y refer to two replicate pairs. The values were then visualized per readout and as all readouts combined. The overall response of the perturbations was also evaluated by log transforming the data upon normalization with non-stimulated control and subsequent scaling of the data.

To prepare the perturbation data input for MRA modeling, the raw data from both assays were combined to generate the complete input dataset in MIDAS format as described in Dorel *et. al.* (Dorel et al., 2018). Network structure deduced from literature data and from additional expertise in the field were used as a prior-knowledge based starting network. This was formatted into .dat file as described in Dorel *et. al.* (Dorel et al., 2018). The final basal.dat file required for the model generation enlisted all the nodes that have a basal activity, which, in this study included all the nodes with the exception of the Growth Factor (GF) and GF receptors. The modeling and subsequent evaluations were performed using STASNet R package. Briefly, the initial models with parameters quantified as local response coefficients (*r*) were generated using the createModel() function, which was then evaluated in comparison to the experimental data using plotModelAccuracy(). The individual goodness of fit for each analyte was also determined using plotModelScores()function. Next, Reduction of the model was performed using selectMinimalModel() function which iteratively removed links that do not significantly deteriorate the LS, as tested by a likelihood ratio test with p<0.05. Once the reduction was complete, the reduced model was again

evaluated using the aforementioned functions, to compare the simulated model to the experimental data output and the overall response of each analyte. Finally, potential novel links were additionally determined by either suggestExtension() function of STASNet or by evaluating the model error and identifying links that could explain the difference. Each new link was then iteratively added to the network.dat file and a new model was generated and evaluated subsequently, until the model had reached a reasonable fit quality, when compared to the theoretical value, and the data points were well explained by the model.

The model parameters/local response coefficients are evaluated, and point-wise confidence intervals are computed using the profileLikelihood() function. A likelihood profile is computed by setting the value of a parameter constant at various values while allowing the others to relax. For those models, the inhibitory parameters were also held constant (but not varied) and the confidence interval was calculated using a X^2 test with 7 degrees of freedom (1 fixed parameter + 6 inhibition parameters). The difference in parameter values and their respective confidence interval between the cell lines were then evaluated by first combining consecutive parameters using aggregatePath() function with few nodes designated as non_stop_nodes: p70S6K, mTORC1, mTORC2, iAZD (dummy node for dual mTOR inhibition), and AKT. Non_stop_nodes parameter allowed for aggregation such that the paths did not stop with these nodes and were evaluated with respect to downstream links. Finally, the parameters and their confidence intervals were visualized using plotParameters() function.

5.10.8   Data visualization and statistics

All data analysis, statistics and visualization were performed in R (version 4.0.0). Visualization was done using the base R plotting function, ggplot2 package, ComplexHeatmap package or using specific packages mentioned in each of the subsections. The appropriate statistics mentioned above were all performed using respective R packages or base R functions. For survival analysis and visualization, the survival and survminer packages were used.

Chapter 6

Supplementary Tables

Further data required for assessment of the results was additionally provided as tables in the respective sections. These tables are displayed below.

6.1 PanNEN panel gene coverage

TABLE 6.1.1: PanNEN panel gene coverage. Number of Amplicons refer to the number of primer pairs that were utilized to cover the exonic regions of the given gene.

Name	Chromosome	Number of Amplicons	Total Bases	Covered Bases	Missed Bases	Overall Coverage	Number of Exons	Number of Exons 100% Coverage
RASAL2	chr1	50	4067	3938	129	0.968	20	19
MTOR	chr1	113	8220	8111	109	0.987	57	52
DDIT4	chr10	8	719	655	64	0.911	2	1
PTEN	chr10	18	1302	1223	79	0.939	9	7
MEN1	chr11	23	1938	1710	228	0.882	10	8
POLR2G	chr11	10	599	578	21	0.965	8	7
PIK3C2A	chr11	75	5381	5286	95	0.982	32	28
EED	chr11	25	1450	1450	0	1	13	13
KRAS	chr12	10	737	681	56	0.924	5	3
CDKN1B	chr12	7	617	595	22	0.964	2	1
LGR5	chr12	39	2904	2864	40	0.986	18	14
FOXO1	chr13	20	1988	1567	421	0.788	2	0
LATS2	chr13	32	3337	2767	570	0.829	7	5
AKT1	chr14	26	1573	1470	103	0.935	13	9
MLST8	chr16	14	1061	922	139	0.869	9	4
PDPK1	chr16	22	1811	1604	207	0.886	15	10
TSC2	chr16	89	5834	5405	429	0.926	42	30
RPTOR	chr17	67	4348	4060	288	0.934	34	28
TP53	chr17	22	1383	1351	32	0.977	15	14
NF1	chr17	134	9161	8961	200	0.978	59	55
PIK3R2	chr19	26	2337	1667	670	0.713	15	8
MAP4K1	chr19	52	2884	2670	214	0.926	33	26
AKT2	chr19	27	1576	1536	40	0.975	14	11
GLI2	chr2	46	4891	3716	1175	0.76	13	6
IRS1	chr2	32	3739	3232	507	0.864	1	0
ALK	chr2	68	5153	4713	440	0.915	29	26
ZEB2	chr2	45	3735	3735	0	1	9	9
ARFGEF2	chr20	81	5748	5713	35	0.994	39	38
BRD1	chr22	38	3297	3076	221	0.933	12	6
VHL	chr3	6	672	491	181	0.731	3	2
PIK3CA	chr3	50	3407	3282	125	0.963	20	17
SETD2	chr3	90	7905	7643	262	0.967	21	18
SMAD1	chr4	18	1458	1458	0	1	6	6
RASGEF1C	chr5	20	1531	1237	294	0.808	13	8

Name	Chromosome	Number of Amplicons	Total Bases	Covered Bases	Missed Bases	Overall Coverage	Number of Exons	Number of Exons 100% Coverage
RICTOR	chr5	88	5589	5359	230	0.959	41	33
PIK3R1	chr5	34	2467	2429	38	0.985	18	16
FOXO3	chr6	18	2042	1583	459	0.775	2	1
DAXX	chr6	29	2356	2207	149	0.937	10	6
POLR2J	chr7	7	394	350	44	0.888	4	3
CARD11	chr7	53	3705	3418	287	0.923	24	17
RHEB	chr7	11	635	589	46	0.928	8	7
MYC	chr8	15	1395	1318	77	0.945	3	1
CDKN2A	chr9	8	962	507	455	0.527	7	3
TSC1	chr9	49	3705	3603	102	0.972	22	20
GLUD2	chrX	16	1687	1545	142	0.916	1	0
AR	chrX	31	2873	2641	232	0.919	9	8
ATRX	chrX	113	7829	7632	197	0.975	35	30

6.2   Mutations of PanNEN cohort

TABLE 6.2.1: Mutations of PanNEN cohort. AA - amino acid change. A/D - Adjacent normal tissue/Distance normal tissue. Allele % - proportion of alteration in the sample (Healthy, Primary and/or Metastasis). Quality - Reliability of the variant being an actual variant in the given sample (Healthy, Primary and/or Metastasis). If no value is indicated, it means that the variant was absent in that particular sample. Annotation - COSMIC refers to the Catalogue of Somatic Mutations in Cancer database and LOVD stands for Leiden Open Variation Database. Panel refers to which panel was used to sequence the sample: CCP - Commercial Cancer Panel and PanNEN - Pancreatic NeuroEndocrine Neoplasms panel

Patient	Gene	AA	A/D	Healthy sample allele %	Quality	Primary allele %	Quality	Metastasis allele %	Quality	Annotation	Panel
PNET2	KRAS	p.G12V	D	0		0.574	6902.35			COSMIC	PanNEN
	TSC2	p.A889P		0		0.365	5163.51			LOVD	
	TP53	p.P278L		0.001		0.541	5220.39			COSMIC	
PNET5	VHL	p.P86L	N/A			0.782	5151.74			COSMIC	PanNEN
	LATS2	p.D739N				0.325	2372.55				
PNET8	SETD2	p.K2385N	A	0.015		0.059	51.37				PanNEN
PNET9	RASAL2	p.R769H	D	0.002		0.062	70.55				PanNEN
	MEN1	p.R360W		0.002		0.448	6774.38				
	PTEN	p.R233*		0.004		0.477	1203.87			COSMIC	
PNET11	MEN1	p.H139Y	A	0				0.272	772.3	COSMIC	PanNEN
	ATRX	p.1265_1266delND		0				0.233	507.92		
PNET14	VHL	p.I109S	A	0		0.74	6408.54				PanNEN
PNET15	MEN1	p.Q64*		0.561	1908	0.93	3718.04				PanNEN
PNET16	MEN1	p.K427*	A	0.006		0.672	15826.9				PanNEN
PNET21	DAXX	p.T262I	A	0.333	3582.8	0.819	8023.01				PanNEN
PNET24	DAXX	p.R291P	A	0.001		0.33	3132.51				PanNEN
PNET25	PIK3R2	p.P4S	N/A			0.694	11823.8			COSMIC	PanNEN
PNET33	ATRX MEN1(syndrome)	p.N1860S	D	0.611	704.91	0.685				COSMIC	PanNEN
PNET37	ATRX	p.ET723DCfs*7	N/A			0.138	153.25			COSMIC	PanNEN

Patient	Gene	AA	Healthy sample			Primary		Metastasis		Annotation	Panel
			A/D	allele %	Quality	allele %	Quality	allele %	Quality		
PNET42	AFF1	p.T399S	blood	0				0.583	2329.32		CCP
	PSIP1	p.400dupT		0.087				0.15	1570.3		
	PSIP1	p.TL399N*fs*4		0				0.37	1570.3		
	PSIP1	p.L400Efs*4		0				0.86	1570.3		
	RNF213	p.G4337R		0				0.384	584.48		
	LPHN3	p.C719S		0				0.268	344.02		
PNET43	TSC1	p.S905I	D	0.078		0.401	921.86				PanNEN
PNET47	MEN1	p.C235Lfs*59	D	0		0.716	5710.04				PanNEN
PNET50	MEN1	p.297_298insRS		0				0.738	13136.4		PanNEN
	ATRX	p.K2261E		0.001				0.861	24141.8		
PNET52	MEN1	p.84delS	A	0.171	50	0				COSMIC	PanNEN
PNET79	DAXX	p.D349A	D	0.001		0.517	6449.61				PanNEN
	MEN1	p.79_84delPVADLS		0		0.543					
PNET83	MEN1	p.84delS	D	0		0.878	9831.74				PanNEN
PNET85	DAXX	p.S102*	D	0		0.514	3702.28				PanNEN
	TSC2	p.1254_1257delSVPA		0		0.219	896.16				
	NF1	p.A1610S		0.004		0.311	1868.4				
	MEN1	p.Q453X									
PNET56P1	BRD3	p.Y348C	D	0				0.371	1441.44		CCP
	TSC2	p.855_856delHL		0				0.577	50		
PNET56P2	BRD3	p.Y348C	D	0				0.317			CCP
	TSC2	p.855_856delHL		0				0.343			
PNET58	KRAS	p.G12V	N/A					0.33	279.07		CCP
	MLLT10	p.I564V						0.509	1441.59		
	TP53	p.P177R						0.548	641.52	COSMIC	
	SMAD4	p.R361H						0.454	617.43	COSMIC	
PNET59	IDH1	p.G150V	N/A					0.265	389.01		CCP
	PIK3CG	p.R849*						0.315	376.24		
	SMO	p.P641A						0.287	370.21		
	EP300	p.1626_1629delGRDA						0.476	955.8		
PNET60	MSH6	p.V250Hfs*5	D	0				0.259	514.96		CCP
	MEN1	p.322delH		0				0.9	2013.4		
	SMAD4	p.G365D		0.023				0.858	1286.34	COSMIC	
PNET61	DAXX	p.Q21*	A	0		0.579	642.17				CCP
	PTCH1	p.V932I		0		0.394	820.33			COSMIC	
PNET62	VHL	p.R161*	N/A			0.317	296.95			COSMIC	CCP
	KMT2C	p.R284Q				0.123	136.58			COSMIC	
	CYP2C19	p.R124Q				0.276	504.11				
	IGF2R	p.V1456A				0.148	150.18				
	PLEKHG5	p.M680V				0.079	68.61				
PNET63	KRAS	p.G12R	A	0		0.66	799.01			COSMIC	CCP
	PIK3C2B	p.Q688R		0		0.227	530.63				
	FANCF	p.336_338delDGD		0		0.453	1084.3				
PNET66	KEAP1	p.R169H	D	0				0.424	886.25		CCP
PNET67	PTPRD	p.R141H	A	0				0.413	388.55	COSMIC	CCP
	KRAS	p.Q61R		0				0.672	2512.5	COSMIC	
PNET70	VHL	p.D143E	N/A					0.171	142.42	COSMIC	CCP
	MYH11	p.E1561K						0.47	1296.18	COSMIC	
	CYP2D6	p.Y355C						0.27	201.75	COSMIC	
	CYP2D6	p.H352R						0.196	158.09	COSMIC	
PNET100	ACVR2A	p.L279F	D	0		0.664	925.06				CCP
	PTPRD	p.W1733R		0		0.321	1167.12				

Patient	Gene	AA	A/D	Healthy sample			Primary		Metastasis		Annotation	Panel
				allele %	Quality		allele %	Quality	allele %	Quality		
	PER1	p.P309S		0.037			0.377	672.44				
PNET101	CDKN2C	p.L107F	D	0.014			0.583	837.73				CCP
PNET102	MAFB	p.HP160QA	D	0			0.217	151.24				CCP
PNET103	MDM4	p.P388T	N/A				0.514	555.12				CCP
	SYNE1	p.A7648G					0.308	714.35				
PNET104	LRP1B	p.G619C	N/A				0.208	132.2				CCP
	ERBB4	p.T757A					0.632	2361.61				
	EPHB1	p.N241H					0.226	224.74				
	SAMD9	p.R553*					0.572	2578.2				
	PTPRD	p.L1398P					0.321	702.38				
PNET105	NUP214	p.V562L	D	0			0.242	188.6				CCP
PNET106	KDM6A	p.A205Afs*21	D	0			0.618	696.16				CCP
	TP53	p.R273C		0			0.589	1340.88			COSMIC	
	NIN	p.Y98C		0			0.254	514.73				
	KRAS	p.G12D		0			0.497	824.58			COSMIC	
PNET107	DAXX	p.R306*	D	0.007			0.709	1577.29			COSMIC	CCP
	MEN1	p.Y307Pfs*6		0			0.735	1229.33				
	HIF1A	p.Q397E		0			0.411	1289.54				
PNET108	PTEN	p.Y88*	D	0			0.35	1094.64			COSMIC	CCP

6.3   Focal copy number amplifications in Group A

TABLE 6.3.1: Focal copy number amplifications in Group A. The rows are as follow: Cytoband - cytoband of the chromosome to which the significantly amplified region belong to, q-value - the q-value of the peak region, residual q-value - The q-value of the peak region after removing ("peeling off") amplifications that overlap other more significant peak regions in the same chromosome, wide peak boundaries - The actual chromosomal location(range) in which the significant amplification reside in, and genes in wide peak - the genes residing within the wide peak boundaries.

		GISTIC: Group A Significant Amplification			
cytoband	14q11.2	8p23.1	1q21.2	16p11.2	10q11.22
q-value	3.35E-41	1.36E-12	2.23E-09	0.015495	0.03299
residual q-value	3.35E-41	1.36E-12	2.23E-09	0.015495	0.03299
wide peak boundaries	chr14: 1-20599999	chr8: 7050001-7446939	chr1: 148917247-149699999	chr16: 32000001-33924999	chr10: 46175001-48227853
genes in wide peak	OR4K5	SPAG11B	FAM91A2	TP53TG3	ANXA8L2
	OR11H2	DEFB103B	PPIAL4A	SLC6A10P	CTSL1P2
	OR4K1	DEFB104A	LOC388692	LOC390705	PPYR1
	OR4K15	DEFB105A	HIST2H2BF	HERC2P4	GPRIN2
	OR4K14	DEFB106A	LOC645166	TP53TG3C	PTPN20B
	OR4L1	DEFB107A	PPIAL4B	LOC729264	FAM21B
	OR4N2	LOC349196	PPIAL4C	TP53TG3B	SYT15
	OR4K2	DEFB103A	LOC728855		AGAP4
	OR4K13	FAM90A13	FCGR1C		ANTXRL
	OR4K17	FAM90A5			FAM21C
	POTEG	FAM90A7			BMS1P5
	OR11H12	DEFB107B			FAM35B
	OR4Q3	DEFB104B			FAM35B2
	OR4M1	DEFB106B			AGAP9
	POTEM	DEFB105B			LOC642826
	LOC642426	DEFB109P1B			LOC643650
		FAM90A14			FAM25C
		FAM90A20			PTPN20A

GISTIC: Group A Significant Amplification	
FAM66B	ANXA8
ZNF705G	BMS1P1
DEFB4B	ANXA8L1
	LOC728643
	FAM25B
	FAM25G

6.4 Focal copy number deletions in Group A

TABLE 6.4.1: Focal copy number deletions in Group A. The rows are as follow: Cytoband - cytoband of the chromosome to which the significantly deleted region belong to, q-value - the q-value of the peak region, residual q-value - The q-value of the peak region after removing ("peeling off") deletions that overlap other more significant peak regions in the same chromosome, wide peak boundaries - The actual chromosomal location(range) in which the significant deletion reside in, and genes in wide peak - the genes residing within the wide peak boundaries.

						GISTIC: Group A Significant Deletion						
cytoband	15q11.2	9p21.3	1p36.32	4q34.3	21q11.2	22q12.3	2q37.3	6p25.3	8p23.2	12p11.1	22q13.32	10q26.3
q-value	2.10E-13	0.00083978	0.005574	0.064882	0.21748	0.21258	0.22502	0.22937	0.22502	0.22502	0.22502	0.24045
residual q-value	2.10E-13	0.00083978	0.005574	0.064882	0.21748	0.21748	0.22502	0.22502	0.22502	0.22502	0.22502	0.24045
wide peak boundaries	chr15: 1-22824999	chr9: 21555001-23698648	chr1: 4005001-4724999	chr4: 182075001-183246982	chr21: 15345395-16334999	chr22: 32025018-32754999	chr2: 242825001-243199373	chr6: 1-524999	chr8: 490126-7224999	chr12: 33585001-38714994	chr22: 48915001-49264999	chr10: 131905187-133609506
genes in wide peak	hsa-mir-1268	CDKN2A	LOC284661	hsa-mir-1305	HSPA13	SLC5A1	LOC728323	IRF4	hsa-mir-596	ALG10	FAM19A5	hsa-mir-378c
	hsa-mir-3118-6	CDKN2B		MGC45800	RBM11	SLC5A4		DUSP22	ANGPT2		LOC284933	GLRX3
	hsa-mir-3118-4	MTAP		MIR1305	SAMSN1	YWHAH			DEFA1		MIR4535	TCERG1L
	NBEAP1	C9orf53			LIPI	DEPDC5			DEFA3			MIR378C
	OR4N4	DMRTA1			ABCC13	RFPL2			DEFA4			
	HERC2P3	FLJ35282			LOC388813	AP1B1P1			DEFA5			
	GOLGA6L1	CDKN2B-AS1				C22orf24			DEFA6			
	POTEB					C22orf42			DEFB1			
	LOC348120					PRR14L			CLN8			
	OR4M2								MYOM2			
	OR4N3P								DLGAP2			
	NF1P2								ARHGEF10			
	CHEK2P2								KBTBD11			
	LOC646214								AGPAT5			
	CXADRP2								CSMD1			
	REREP3								MCPH1			
	GOLGA6L6								ERICH1			
	LOC727924								DEFT1P			
	GOLGA8C								LOC286083			
	GOLGA8DP								LOC349196			
	MIR4509-1								XKR5			
	MIR4509-2								FAM90A13			
	MIR4509-3								FAM90A5			
									DEFA10P			
									DEFB109P1B			
									FAM90A14			
									MIR596			
									DEFA1B			
									FAM90A20			
									FAM66B			
									ZNF705G			
									LOC100287015			

GISTIC: Group A Significant Deletion
DEFT1P2
MIR4659A
MIR4659B
LOC100652791

6.5   Focal copy number amplifications in Group B

TABLE 6.5.1: Focal copy number amplifications in Group B. The rows are as follow: Cytoband - cytoband of the chromosome to which the significantly altered region belong to, q-value - the q-value of the peak region, residual q-value - The q-value of the peak region after removing ("peeling off") amplifications that overlap other more significant peak regions in the same chromosome, wide peak boundaries - The actual chromosomal location(range) in which the significant amplification reside in, and genes in wide peak - the genes residing within the wide peak boundaries.

GISTIC: Group B Significant Amplification				
cytoband	8p23.1	1q21.2	4p16.1	6q21	14q11.2
q-value	0.039216	0.05047	0.05047	0.12643	0.12643
residual q-value	0.039216	0.05047	0.05047	0.12643	0.12643
wide peak boundaries	chr8: 7050001-7784999	chr1: 149265113-149799999	chr4: 8714604-9909983	chr6: 105200001-105959999	chr14: 1-20599999
genes in wide peak	DEFB4A	FCGR1A	hsa-mir-548i-2	PREP	OR4K5
	SPAG11B	FAM91A2	DRD5	BVES	OR11H2
	DEFB103B	PPIAL4A	HMX1	HACE1	OR4K1
	DEFB104A	LOC388692	SLC2A9	POPDC3	OR4K15
	DEFB105A	HIST2H2BF	USP17L6P	BVES-AS1	OR4K14
	DEFB106A	PPIAL4B	USP17	LIN28B	OR4L1
	DEFB107A	PPIAL4C	DEFB131		OR4N2
	LOC349196	HIST2H3D	LOC650293		OR4K2
	DEFB103A	LOC728855	LOC728369		OR4K13
	FAM90A13	FCGR1C	LOC728373		OR4K17
	FAM90A5		LOC728379		POTEG
	FAM90A7		USP17L5		OR11H12
	FAM90A8		LOC728393		OR4Q3
	FAM90A18		LOC728400		OR4M1
	FAM90A9		LOC728405		POTEM
	FAM90A10		MIR548I2		LOC642426
	DEFB107B				
	DEFB104B				
	DEFB106B				
	DEFB105B				
	DEFB109P1B				
	FAM90A14				
	SPAG11A				
	FAM90A20				
	FAM90A19				
	FAM66B				
	ZNF705G				
	LOC100132396				
	DEFB4B				

6.6 Focal copy number deletions in Group B

TABLE 6.6.1: Focal copy number deletions in Group B. The rows are as follow: Cytoband - cytoband of the chromosome to which the significantly deleted region belong to, q-value - the q-value of the peak region, residual q-value - The q-value of the peak region after removing ("peeling off") deletions that overlap other more significant peak regions in the same chromosome, wide peak boundaries - The actual chromosomal location(range) in which the significant deletion reside in, and genes in wide peak - the genes residing within the wide peak boundaries.

	GISTIC: Group B Significant Deletion		
cytoband	13q14.2	15q11.2	4p15.32
q-value	0.023493	0.023493	0.17855
residual q-value	0.023493	0.023493	0.17855
wide peak boundaries	chr13: 48665001-49884999	chr15: 1-22743091	chr4: 16085001-17524999
genes in wide peak	RCBTB2	hsa-mir-1268	QDPR
	MLNR	hsa-mir-3118-6	LDB2
	RB1	hsa-mir-3118-4	TAPT1
	ITM2B	NBEAP1	FLJ39653
	LPAR6	OR4N4	
	FNDC3A	HERC2P3	
	CYSLTR2	POTEB	
	CDADC1	LOC348120	
		OR4M2	
		OR4N3P	
		NF1P2	
		CHEK2P2	
		LOC646214	
		CXADRP2	
		REREP3	
		GOLGA6L6	
		LOC727924	
		GOLGA8C	
		GOLGA8DP	
		MIR4509-1	
		MIR4509-2	
		MIR4509-3	

6.7   Pancreatic cell marker genes curated from PangloDB

TABLE 6.7.1: Pancreatic cell marker genes curated from PangloDB. The Ubiquitousness Index refers to the expression status of the gene in the cell clusters. The value ranges between 0 and 1. An ubiquitousness Index close to 0 indicate that the given gene is not expressed in any cell cluster of that particular cell type, while an index close to 1 suggest that it is expressed in almost all cell clusters of the given cell type. Gene type can either be protein-coding gene or non-coding RNA gene. Canonical Marker indicate whether the expression of the particular marker gene defines the presence of the cell type. Sensitivity Human and Sensitivity Mouse columns refer to how frequently this marker is expressed in cells of the given cell type in humans and mouse, respectively. Specificity Human and Specificity Mouse refers to how frequently this marker is expressed in cells not of the given cell type in humans and mouse, respectively. For the sensitivity and specificity columns, NA means no data was available to compute on.

Official Gene Symbol	Cell Type	Ubiquitousness Index	Gene Type	Canonical Marker	Sensitivity Human	Sensitivity Mouse	Specificity Human	Specificity Mouse
CTRB1	Acinar cells	0.017	protein-coding	1	1	0.957143	0.00062893	0.0159201
KLK1	Acinar cells	0.013	protein-coding	1	0.833333	0.314286	0.00503145	0.0128263
CELA3A	Acinar cells	0.001	protein-coding	1	0.833333	0.128571	0	0
PRSS1	Acinar cells	0.002	protein-coding	1	1	0.0285714	0.00597484	0
SPINK1	Acinar cells	0.029	protein-coding	1	1	0	0.0352201	0
CEL	Acinar cells	0.002	protein-coding	1	0.166667	0.442857	0	0.0008379
CELA2A	Acinar cells	0.005	protein-coding	1	0.333333	0.628571	0	0.00290042
CPB1	Acinar cells	0.005	protein-coding	1	0.833333	0.5	0.00786164	0.00232034
RNASE1	Acinar cells	0.022	protein-coding	1	0.75	0	0.0877358	0
AMY2B	Acinar cells	0.003	protein-coding	1	0.0833333	0.528571	0	0.00148244
CPA2	Acinar cells	0.004	protein-coding	1	0.916667	0.457143	0	0.00277151
CPA1	Acinar cells	0.006	protein-coding	1	0.833333	0.585714	0	0.00438286
CELA3B	Acinar cells	0.007	protein-coding	1	0.75	0.8	0.00125786	0.00502739
PNLIP	Acinar cells	0.005	protein-coding	1	0.833333	0.585714	0.00031447	0.00354496
CTRB2	Acinar cells	0.001	protein-coding	1	0.916667	0	0.00125786	0
PLA2G1B	Acinar cells	0.007	protein-coding	1	0.833333	0.228571	0.00157233	0.00638092
PRSS2	Acinar cells	0.01	protein-coding	1	1	0.7	0.0103774	0.00696101
CLPS	Acinar cells	0.021	protein-coding	1	0.583333	0.957143	0.00031447	0.0215276
REG1A	Acinar cells	0.005	protein-coding	1	1	0	0.0273585	0
SYCN	Acinar cells	0.008	protein-coding	1	0.833333	0	0.00157233	0
PNLIPRP1	Acinar cells	0.008	protein-coding	1	0.5	0.628571	0	0.00663874
CTRC	Acinar cells	0.003	protein-coding	1	0.833333	0.414286	0	0.00135353
REG3A	Acinar cells	0.005	protein-coding	1	0.916667	0	0.0106918	0
PRSS3	Acinar cells	0.006	protein-coding	1	0.833333	0.0285714	0.0289308	0
REG1B	Acinar cells	0.002	protein-coding	1	0.916667	0	0.0110063	0
GDF15	Acinar cells	0.014	protein-coding	1	0.833333	0	0.054088	0.00309378
MUC1	Acinar cells	0.016	protein-coding	1	0.916667	0.185714	0.0295597	0.0117306
C15ORF48	Acinar cells	0.018	protein-coding	1	1	0	0.109119	0
AKR1C3	Acinar cells	0.007	protein-coding	1	1	0	0.0389937	0
OLFM4	Acinar cells	0.013	protein-coding	1	0.583333	0	0.0132075	0.0126974
GSTA1	Acinar cells	0.012	protein-coding	1	0.916667	0	0.0267296	0.00837899
LGALS2	Acinar cells	0.032	protein-coding	1	0.916667	0.0142857	0.0345912	0.0308733
PDZK1IP1	Acinar cells	0.025	protein-coding	1	1	0.114286	0.0396226	0.0228166
RARRES2	Acinar cells	0.03	protein-coding	1	1	0	0.0751572	0
CXCL17	Acinar cells	0.007	protein-coding	1	0.75	0	0.0248428	0.00367386
GSTA2	Acinar cells	0.007	protein-coding	1	0.916667	0.0142857	0.00628931	0.00650983
ANPEP	Acinar cells	0.015	protein-coding	1	0.75	0.0714286	0.0100629	0.0166291
ANGPTL4	Acinar cells	0.031	protein-coding	1	0.166667	0	0.0295597	0.0329359
ALDOB	Acinar cells	0.031	protein-coding	1	0.583333	0.0428571	0.00408805	0.0370609
GCG	Alpha cells	0.024	protein-coding	1	0.83871	0.75	0.0439734	0.0180766
TTR	Alpha cells	0.055	protein-coding	1	0.83871	0	0.065802	0
ARX	Alpha cells	0.019	protein-coding	1	0.129032	0.35	0.00094907	0.0227726
NKX2-2	Alpha cells	0.002	protein-coding	NA	0.290323	0.075	0.00759253	0.00045031
PCSK2	Alpha cells	0.032	protein-coding	1	0.83871	0.85	0.0167668	0.0327436
TM4SF4	Alpha cells	0.017	protein-coding	1	0.806452	0	0.0132869	0
CRYBA2	Alpha cells	0.009	protein-coding	1	0.967742	0.975	0.0113888	0.00488903
NKX6-1	Alpha cells	0.006	protein-coding	1	0.0322581	0.3	0.0101234	0.00450306

Official Gene Symbol	Cell Type	Ubiquitousness Index	Gene Type	Canonical Marker	Sensitivity Human	Sensitivity Mouse	Specificity Human	Specificity Mouse
KCNK16	Alpha cells	0.007	protein-coding	NA	0.225806	0.875	0.0056944	0.00443873
PCSK1	Alpha cells	0.008	protein-coding	NA	0.0645161	0.225	0.00885796	0.00791251
PRRG2	Alpha cells	0.009	protein-coding	NA	0.0645161	0.475	0.00031636	0.0100354
IRX2	Alpha cells	0.012	protein-coding	1	0.0967742	0.275	0.00537805	0.0124799
ALDH1A1	Alpha cells	0.039	protein-coding	1	0.903226	0	0.080038	0.0288196
PEMT	Alpha cells	0.013	protein-coding	1	0.709677	0	0.0385954	0.00276616
CHGA	Alpha cells	0.057	protein-coding	1	0.967742	1	0.0287884	0.0598263
SMIM24	Alpha cells	0.037	protein-coding	1	0.741935	0	0.0281556	0
F10	Alpha cells	0.011	protein-coding	1	0.645161	0	0.00917431	0.00836282
SCGN	Alpha cells	0.014	protein-coding	1	0.83871	0	0.0218285	0
SLC30A8	Alpha cells	0.005	protein-coding	1	0.677419	0	0.00885796	0
SH3GL2	Alpha cells	0.025	protein-coding	1	0.0322581	0	0.00221449	0
SCGB2A1	Alpha cells	0.003	protein-coding	1	0.806452	0	0.00980702	0
MAFB	Alpha cells	0.053	protein-coding	1	0.709677	0.7	0.0528314	0.0512062
PAX6	Alpha cells	0.049	protein-coding	1	0.258065	0.85	0.0120215	0.0569315
NEUROD1	Alpha cells	0.038	protein-coding	1	0.806452	1	0.0155014	0.0405275
LOXL4	Alpha cells	0	protein-coding	1	0.129032	0	0	0.00032165
PLCE1	Alpha cells	0.005	protein-coding	1	0.0322581	0	0.00822525	0.00456739
GC	Alpha cells	0.009	protein-coding	1	0.774194	0.375	0.00664347	0.00694757
KLHL41	Alpha cells	0.001	protein-coding	1	0.0322581	0	0.00031636	0.00051464
FEV	Alpha cells	0.003	protein-coding	1	0.16129	0.475	0.00189813	0.00225153
RFX6	Alpha cells	0.001	protein-coding	1	0.225806	0	0.00126542	0
SMARCA1	Alpha cells	0.015	protein-coding	1	0.0967742	0	0.0408099	0
UCP2	Alpha cells	0.013	protein-coding	1	0.16129	0	0.00347991	0
FXYD5	Alpha cells	0.213	protein-coding	1	0.741935	0	0.374881	0.173882
RGS4	Alpha cells	0.052	protein-coding	1	0.387097	0	0.0056944	0
GLS	Alpha cells	0.061	protein-coding	1	0.0967742	0.25	0.0189813	0.0718559
PDX1	Beta cells	0.005	protein-coding	1	0.681818	0.690476	0.00126183	0.00289519
FXYD2	Beta cells	0.025	protein-coding	1	0.363636	0	0.00189274	0.0306891
NPY	Beta cells	0.015	protein-coding	NA	0.136364	0.238095	0.00473186	0.0164704
INS	Beta cells	0.009	protein-coding	1	1	0	0.0470032	0
MAFA	Beta cells	0.002	protein-coding	1	0.5	0.666667	0.00031546	0.00019301
NKX2-2	Beta cells	0.002	protein-coding	1	0.727273	0.119048	0.00536278	0.00032169
NKX6-1	Beta cells	0.006	protein-coding	1	0.772727	0.785714	0.00504732	0.00315254
IAPP	Beta cells	0.023	protein-coding	1	1	1	0.0353312	0.0179502
PCSK2	Beta cells	0.032	protein-coding	1	0.909091	0.857143	0.018612	0.0326192
G6PC2	Beta cells	0.004	protein-coding	1	0.863636	0.785714	0.00536278	0.00012868
SLC30A8	Beta cells	0.005	protein-coding	1	0.909091	0	0.00914827	0
PCSK1	Beta cells	0.008	protein-coding	1	0.818182	0.785714	0.00378549	0.00636943
GJD2	Beta cells	0.005	protein-coding	1	0.590909	0.452381	0.00189274	0.00379592
SCGN	Beta cells	0.014	protein-coding	1	1	0	0.0230284	0
IGF2	Beta cells	0.013	protein-coding	1	0.0454545	0	0.0463722	0.00090073
SYT13	Beta cells	0.029	protein-coding	1	0.590909	0	0.00283912	0
NPTX2	Beta cells	0.008	protein-coding	1	0.818182	0	0.0126183	0.00688413
PFKFB2	Beta cells	0.001	protein-coding	1	0.272727	0.0714286	0.00315457	0.00019301
EDARADD	Beta cells	0.001	protein-coding	1	0.0454545	0	0.0022082	0.00090073
HOPX	Beta cells	0.033	protein-coding	1	0.863636	0.333333	0.0864353	0.0214244
SH3GL2	Beta cells	0.025	protein-coding	1	0.0454545	0	0.0022082	0
ADCYAP1	Beta cells	0.003	protein-coding	1	0.772727	0	0.00094637	0.00257351
SCGB2A1	Beta cells	0.003	protein-coding	1	0.818182	0	0.0119874	0
CASR	Beta cells	0.001	protein-coding	1	0.272727	0.0952381	0.00031546	0.00025735
MAFB	Beta cells	0.053	protein-coding	1	0.772727	0.357143	0.0542587	0.0520492
PAX6	Beta cells	0.049	protein-coding	1	0.909091	0.785714	0.00820189	0.0570032
NEUROD1	Beta cells	0.038	protein-coding	1	0.909091	0.785714	0.0170347	0.0409831
ISL1	Beta cells	0.028	protein-coding	1	0.681818	0.738095	0.00883281	0.0294666
SIX3	Beta cells	0.014	protein-coding	1	0.0454545	0	0.00283912	0
BMP5	Beta cells	0.002	protein-coding	1	0.0909091	0	0.00599369	0.00077205
DLK1	Beta cells	0.055	protein-coding	1	0.681818	0.333333	0.0416404	0.0572605
MEG3	Beta cells	0.288	protein-coding	1	0.909091	0.833333	0.109779	0.333591
RGS16	Beta cells	0.01	protein-coding	1	0.136364	0	0.0195584	0
PDX1	Delta cells	0.005	protein-coding	1	0.125	0.7	0.00565327	0.00430177
FFAR4	Delta cells	0.005	protein-coding	1	0.25	0.9	0.00408291	0.00430177
SST	Delta cells	0.027	protein-coding	1	1	0	0.0452261	0
RBP4	Delta cells	0.032	protein-coding	1	0.875	0	0.03549	0
HHEX	Delta cells	0.037	protein-coding	1	1	1	0.0204146	0.0389727
ISL1	Delta cells	0.028	protein-coding	1	1	1	0.0109925	0.0307544
SCGN	Delta cells	0.014	protein-coding	1	0.875	0	0.0276382	0
KCNK16	Delta cells	0.007	protein-coding	1	0.125	1	0.00753769	0.00603531
IAPP	Delta cells	0.023	protein-coding	1	0.75	1	0.040201	0.0199679
GPC5-AS1	Delta cells	0.001	protein-coding	1	0.5	0	0.00282663	0

Official Gene Symbol	Cell Type	Ubiquitousness Index	Gene Type	Canonical Marker	Sensitivity Human	Sensitivity Mouse	Specificity Human	Specificity Mouse
UCP2	Delta cells	0.013	protein-coding	1	0.125	0	0.00471106	0
GABRB3	Delta cells	0.025	protein-coding	1	0.125	0.2	0.00408291	0.0303692
EDN3	Delta cells	0.003	protein-coding	1	0.125	0	0.00219849	0.00378812
FRZB	Delta cells	0.012	protein-coding	1	0.125	0.8	0.0301508	0.00744783
PCSK1	Delta cells	0.008	protein-coding	1	0.75	0.4	0.00753769	0.0082183
ETV1	Delta cells	0.037	protein-coding	1	0.125	0.8	0.00785176	0.0445586
CFTR	Ductal cells	0.003	protein-coding	1	0.257143	0	0.00380108	0.00288739
MUC1	Ductal cells	0.016	protein-coding	1	0.228571	0	0.0307254	0.012512
AMBP	Ductal cells	0.013	protein-coding	1	0.171429	0	0.0120367	0.0121912
ANXA4	Ductal cells	0.06	protein-coding	1	0.342857	0	0.0706367	0.0565287
SPP1	Ductal cells	0.057	protein-coding	1	0.285714	0	0.050681	0
GDF15	Ductal cells	0.014	protein-coding	1	0.628571	0	0.050681	0.00307988
AKR1C3	Ductal cells	0.007	protein-coding	1	0.371429	0	0.038961	0
MMP7	Ductal cells	0.007	protein-coding	1	0.657143	0	0.0180551	0.00365736
DEFB1	Ductal cells	0.009	protein-coding	1	0.4	0	0.0218562	0.00532563
SERPING1	Ductal cells	0.066	protein-coding	1	0.285714	0	0.0972442	0
TSPAN8	Ductal cells	0.047	protein-coding	1	0.485714	0	0.0430789	0
CLDN10	Ductal cells	0.028	protein-coding	1	0.342857	0	0.0202724	0.0302214
SLPI	Ductal cells	0.048	protein-coding	1	0.885714	0	0.0861577	0
SERPINA5	Ductal cells	0.005	protein-coding	1	0.257143	0	0.0237567	0
PIGR	Ductal cells	0.031	protein-coding	1	0.2	0	0.0424454	0.0291306
CLDN1	Ductal cells	0.01	protein-coding	1	0.228571	0	0.0262908	0.00564646
LGALS4	Ductal cells	0.032	protein-coding	1	0.228571	0	0.036427	0.0307347
PERP	Ductal cells	0.077	protein-coding	1	0.971429	0	0.158695	0.0519731
PDLIM3	Ductal cells	0.013	protein-coding	1	0.114286	0	0.0183719	0.0120629
WFDC2	Ductal cells	0.056	protein-coding	1	0.685714	0	0.0877415	0
ALDH1A3	Ductal cells	0.008	protein-coding	1	0.485714	0	0.0297751	0.00327238
KRT19	Ductal cells	0.088	protein-coding	1	0.942857	0	0.162179	0.0654475
TFF1	Ductal cells	0.006	protein-coding	1	0.0571429	0	0.0199557	0
KRT7	Ductal cells	0.056	protein-coding	1	0.914286	0	0.0921761	0.0474174
CLDN4	Ductal cells	0.047	protein-coding	1	1	0	0.118784	0.0286814
LAMB3	Ductal cells	0.017	protein-coding	1	0.4	0	0.0497308	0.00943215
TACSTD2	Ductal cells	0.033	protein-coding	1	0.971429	0	0.0886918	0
CCL2	Ductal cells	0.037	protein-coding	1	0.371429	0	0.0614507	0.0318896
CXCL2	Ductal cells	0.041	protein-coding	1	0.285714	0	0.0541653	0.0381136
CTSH	Ductal cells	0.147	protein-coding	1	0.542857	0	0.104213	0.154187
S100A10	Ductal cells	0.374	protein-coding	1	1	0	0.526449	0
PPY	Gamma (PP) cells	0.019	protein-coding	1	1	0	0.0386185	0.0153994
SCGN	Gamma (PP) cells	0.014	protein-coding	1	1	0	0.0276295	0
FXYD2	Gamma (PP) cells	0.025	protein-coding	1	0.285714	0	0.00376766	0.0306064
SCGB2A1	Gamma (PP) cells	0.003	protein-coding	1	1	0	0.0153846	0
GPC5-AS1	Gamma (PP) cells	0.001	non-coding RNA	1	0.857143	0	0.0021978	0
CMTM8	Gamma (PP) cells	0.035	protein-coding	1	1	0	0.0241758	0.0371511
ARX	Gamma (PP) cells	0.019	protein-coding	1	0.142857	0	0.00188383	0.0236124
NEUROD1	Gamma (PP) cells	0.038	protein-coding	1	1	0	0.0210361	0.0429901
PAX6	Gamma (PP) cells	0.049	protein-coding	1	0.285714	0	0.0138148	0.058967
ISL1	Gamma (PP) cells	0.028	protein-coding	1	0.428571	0	0.0125589	0.0313763
AQP3	Gamma (PP) cells	0.025	protein-coding	1	0.714286	0	0.0697017	0.0159127
MPZ	Peri-islet Schwann cells	0.002	protein-coding	NA	1	0.6	0.00188029	0.00141252
OLFML2A	Peri-islet Schwann cells	0.003	protein-coding	NA	1	0.5	0.0050141	0.0023114
GULP1	Peri-islet Schwann cells	0.01	protein-coding	NA	1	0.9	0.0294579	0.00404494
GFRA3	Peri-islet Schwann cells	0.005	protein-coding	1	1	1	0.00219367	0.00494382
INSC	Peri-islet Schwann cells	0.005	protein-coding	NA	1	0.5	0.00125353	0.00603531
SLITRK2	Peri-islet Schwann cells	0.006	protein-coding	NA	1	0	0.00125353	0
FIGN	Peri-islet Schwann cells	0.007	protein-coding	NA	1	1	0.00376058	0.0059069
SEMA3B	Peri-islet Schwann cells	0.009	protein-coding	NA	1	0	0.0175494	0
NGFR	Peri-islet Schwann cells	0.007	protein-coding	NA	1	1	0.00188029	0.00680578

6.8 Significantly enriched cell marker genes' associated methylation probes

TABLE 6.8.1: Significantly enriched cell marker genes' associated methylation probes. logFC refers to the difference in fold changes between Group A and Group B in log scale. the $\log_{10}P$ is the log transformed estimate of the computed P-value column. ΔBeta indicates the absolute value of logFC.

Probes	Cell Marker	logFC	P-value	$\log_{10}P$	ΔBeta
cg17444738	NKX6-1	0.56226048	2.22E-14	13.6537594	0.56226048
cg05903444	IRX2	0.4594854	1.08E-12	11.9669731	0.4594854
cg18371475	IRX2	0.42404587	1.89E-12	11.7236725	0.42404587
cg01735082	FRZB	0.42399681	2.36E-12	11.6272781	0.42399681
cg24275315	NKX6-1	0.51217435	4.40E-12	11.3560733	0.51217435
cg12606400	TTR	0.44866134	1.62E-11	10.7898706	0.44866134
cg04504066	PAX6	0.52008956	2.04E-11	10.6907322	0.52008956
cg09524455	IRX2	0.46463463	2.10E-11	10.6771657	0.46463463
cg08235864	IRX2	0.46734255	2.12E-11	10.6734954	0.46734255
cg15941948	IRX2	0.28283732	5.23E-11	10.2811243	0.28283732
cg15666789	TTR	0.41059132	1.25E-10	9.90157165	0.41059132
cg25830182	NKX6-1	0.58938648	1.92E-10	9.71626844	0.58938648
cg04992127	IRX2	0.4408783	4.00E-10	9.3980383	0.4408783
cg08204280	IRX2	0.40159511	4.90E-10	9.30997379	0.40159511
cg18037921	ALDH1A3	0.31698465	6.12E-10	9.21335488	0.31698465
cg22256677	NKX6-1	0.4621334	7.38E-10	9.13190771	0.4621334
cg08521225	TTR	0.48477969	1.27E-09	8.89773864	0.48477969
cg16865446	PAX6	0.44364988	1.50E-09	8.82269288	0.44364988
cg22557091	PAX6	0.50167476	1.94E-09	8.71259199	0.50167476
cg10457539	NKX6-1	0.47270136	2.46E-09	8.60864117	0.47270136
cg10417567	NKX6-1	0.41629544	2.88E-09	8.54092774	0.41629544
cg06119874	PFKFB2	-0.3212771	3.39E-09	8.47024774	0.3212771
cg07688460	NKX6-1	0.45542433	4.16E-09	8.38123869	0.45542433
cg11552694	IRX2	0.27208043	5.19E-09	8.28511488	0.27208043
cg13891702	PAX6	0.45999145	1.13E-08	7.94666382	0.45999145
cg18297736	NKX6-1	0.29911374	1.17E-08	7.93023805	0.29911374
cg01904582	PAX6	0.39936253	1.28E-08	7.89150928	0.39936253
cg13596833	PAX6	0.26940245	1.43E-08	7.84532599	0.26940245
cg22598426	NKX6-1	0.48164795	1.70E-08	7.76933953	0.48164795
cg15889872	PAX6	0.38423528	1.72E-08	7.76351794	0.38423528
cg21093166	IRX2	0.27261484	2.90E-08	7.53710952	0.27261484
cg08499046	PAX6	0.44453749	3.95E-08	7.40305266	0.44453749
cg08116462	PAX6	0.38646221	4.52E-08	7.34510766	0.38646221
cg26578682	IRX2	0.50641281	5.21E-08	7.28294158	0.50641281
cg02489958	NKX6-1	0.42323707	5.85E-08	7.23247695	0.42323707
cg27011060	PAX6	0.51636973	6.40E-08	7.19401171	0.51636973
cg18082638	PAX6	0.43044918	9.05E-08	7.04335934	0.43044918
cg14439629	PAX6	0.32834294	9.25E-08	7.03399341	0.32834294
cg24470575	NGFR	0.29563142	1.00E-07	6.99884465	0.29563142
cg09537620	PAX6	0.53405544	1.02E-07	6.9900789	0.53405544
cg16616455	GABRB3	0.37580253	1.30E-07	6.88569064	0.37580253
cg12076463	NKX6-1	0.45418853	1.72E-07	6.76466914	0.45418853
cg11469061	PAX6	0.31824485	1.97E-07	6.70525922	0.31824485
cg22651463	GLS	0.2906471	2.16E-07	6.66496132	0.2906471
cg16640855	NEUROD1	0.34607882	2.92E-07	6.53405666	0.34607882
cg15929698	NPY	0.27841676	2.98E-07	6.52644523	0.27841676
cg11128216	PAX6	0.50317899	3.07E-07	6.51221848	0.50317899
cg27406678	NKX6-1	0.34430664	3.78E-07	6.42302948	0.34430664
cg05321361	IRX2	0.47595166	4.22E-07	6.37503389	0.47595166
cg03905867	PAX6	0.4600715	5.29E-07	6.27651308	0.4600715
cg09224689	PAX6	0.47879883	5.77E-07	6.23888859	0.47879883
cg00792849	GJD2	0.27688391	5.88E-07	6.23049751	0.27688391
cg21748401	CASR	0.31292273	7.59E-07	6.11978568	0.31292273
cg13245152	PAX6	0.43237399	1.02E-06	5.99242961	0.43237399
cg25884711	NPY	0.36288164	1.07E-06	5.96995242	0.36288164
cg12305238	ETV1	-0.3225731	1.16E-06	5.93498117	0.32257311

Probes	gene	logFC	P-value	$\log_{10}P$	ΔBeta
cg18952796	NPTX2	0.37357208	1.28E-06	5.89422344	0.37357208
cg06482260	KRT7	-0.3500031	1.54E-06	5.81255376	0.35000312
cg04938549	PAX6	0.3381367	1.56E-06	5.80633456	0.3381367
cg13570972	PAX6	0.35653857	1.82E-06	5.73895	0.35653857
cg27652350	ALDH1A3	0.32427075	2.01E-06	5.69622303	0.32427075
cg02770983	PAX6	0.25325547	2.48E-06	5.60475435	0.25325547
cg24852548	HOPX	0.36945	2.85E-06	5.54450154	0.36945
cg21526749	SST	0.29836754	3.04E-06	5.51783227	0.29836754
cg22740492	PAX6	0.38772033	3.67E-06	5.43485552	0.38772033
cg18147181	KRT7	-0.3825187	3.72E-06	5.42965129	0.38251867
cg26862822	NKX2-2	0.34888142	3.82E-06	5.41844283	0.34888142
cg05930013	CLDN10	0.31848072	3.91E-06	5.40735254	0.31848072
cg06714480	NEUROD1	0.30801175	3.96E-06	5.40259481	0.30801175
cg02771142	PAX6	0.40695922	3.97E-06	5.40156323	0.40695922
cg18077866	CASR	0.3337274	4.03E-06	5.39492593	0.3337274
cg21471367	AKR1C3	-0.2904125	4.14E-06	5.38282332	0.29041253
cg01050885	NPY	0.32611203	4.34E-06	5.36217419	0.32611203
cg01380361	PFKFB2	-0.2658821	5.30E-06	5.27582462	0.26588212
cg15215534	PAX6	0.35654741	5.89E-06	5.23019907	0.35654741
cg00355281	NPY	0.2797044	6.20E-06	5.20741813	0.2797044
cg09382096	PAX6	0.35525127	6.35E-06	5.19702231	0.35525127
cg07629454	CASR	0.36863321	6.38E-06	5.19547671	0.36863321
cg22976218	NEUROD1	0.29647646	6.71E-06	5.17338261	0.29647646
cg09853371	HOPX	0.33453146	6.77E-06	5.16930247	0.33453146
cg16964348	NPY	0.2971393	7.22E-06	5.14174689	0.2971393
cg26504021	IRX2	0.37997004	7.67E-06	5.11503733	0.37997004
cg17412258	DLK1	0.2700869	7.88E-06	5.10352612	0.2700869
cg08784129	PAX6	0.40289185	9.00E-06	5.04571854	0.40289185
cg11173941	IRX2	0.44314966	9.21E-06	5.03564956	0.44314966

6.9   Proportion of normal cell signature in PanNEN tumors

TABLE 6.9.1: Proportion of normal cell signature in PanNEN tumors.

	Signature Proportion			
Sample	α-cells	Acinar cells	β-cells	Ductal cells
PNET100	0.827	0.151	0.022	0
PNET101	0.229	0.286	0.226	0.258
PNET102	0.272	0.313	0.258	0.157
PNET103	0.211	0.662	0.081	0.047
PNET104	0.268	0.392	0.254	0.086
PNET105	0.206	0.327	0.433	0.035
PNET106	0.218	0.359	0.319	0.104
PNET107	0.652	0.241	0.025	0.082
PNET108	0.63	0.233	0.054	0.083
PNET11	0.44	0.252	0.094	0.213
PNET14	0.466	0.121	0.292	0.121
PNET15	0.677	0.21	0.057	0.056
PNET16	0.497	0.161	0.235	0.107
PNET18	0.304	0.287	0.22	0.188
PNET2	0.227	0.49	0.224	0.059
PNET20	0.439	0.139	0.198	0.224
PNET21	0.7	0.125	0.175	0
PNET24	0.496	0.213	0.113	0.177
PNET25	0	0.039	0.904	0.057
PNET33	0.611	0.249	0.055	0.084
PNET36	0.557	0.191	0.225	0.028
PNET37	0.727	0.175	0.017	0.081
PNET4	0.232	0.341	0.261	0.165
PNET42	0.557	0.269	0.093	0.081
PNET43	0.883	0.096	0.021	0
PNET46	0.61	0.201	0.077	0.112
PNET47	0.785	0.121	0.033	0.061
PNET5	0.469	0.277	0.208	0.045

		Signature Proportion		
Sample	α-cells	Acinar cells	β-cells	Ductal cells
PNET50	0.754	0.188	0.017	0.041
PNET51	0.379	0.16	0.375	0.086
PNET52	0.1	0.124	0.744	0.032
PNET53	0.53	0.189	0.185	0.095
PNET56P1	0.704	0.254	0.043	0
PNET56P2	0.73	0.24	0.03	0
PNET57	0.44	0.252	0.308	0
PNET58	0.316	0.333	0.2	0.151
PNET59	0.384	0.193	0.351	0.072
PNET60	0.074	0.709	0.156	0.061
PNET61	0.546	0.241	0.187	0.027
PNET62	0.404	0.312	0.056	0.229
PNET63	0.247	0.395	0.236	0.123
PNET65	0.333	0.144	0.429	0.094
PNET66	0.409	0.204	0.344	0.043
PNET67	0.331	0.325	0.229	0.116
PNET70	0.401	0.249	0.087	0.263
PNET77M	0.36	0.205	0.282	0.153
PNET77P	0.313	0.22	0.263	0.203
PNET79	0.525	0.263	0.066	0.145
PNET8	0.366	0.238	0.268	0.128
PNET81	0.792	0.169	0	0.039
PNET83	0.91	0.09	0	0
PNET85	0.676	0.244	0	0.081
PNET89	0.776	0.132	0	0.092
PNET9	0.5	0.228	0.138	0.133
PNET91	0.097	0.104	0.712	0.087
PNET92	0.574	0.112	0.142	0.172
PNET95	0.716	0.191	0.062	0.031

6.10 Identified driver mutations in PanNEN cell lines

TABLE 6.10.1: Identified driver mutations in PanNEN cell lines. Allelic frequency are computed as the number of altered reads for a given variant divided by the total reads at that position. Three different gene panels were used to determine the variants. if the variant was identified in multiple panels, they are indicated in the Gene Panel column, separated by commas: CCP - Commercial Cancer Panel, CRC5 - ColoRectal Cancer 5 panel and PanNEN - Pancreatic NeuroEndocrine Neoplasms panel

Chromosome	Position	Gene	Amino Acid	Variant Effect	Allelic Frequency	Gene Panel
				BON1		
chr17	7574003	TP53	p.R342*	stop_gained	0.997	CCP,CRC5,PanNEN
chr18	48604787	SMAD4	p.537delD	frameshift_truncation	1	CCP,CRC5
chr1	115256529	NRAS	p.Q61R	missense_variant	0.982	CRC5
chr2	60688270	BCL11A	p.E593K	missense_variant	0.345	CCP
chr2	140990778	LRP1B	p.G4593C	missense_variant	0.527	CCP
chr9	133748331	ABL1	p.N350S	missense_variant	1	CCP
chr12	121416795	HNF1A	p.D75G	missense_variant	0.472	CCP
chr16	15835455	MYH11	p.E915D	missense_variant	0.317	CCP
chr20	31024561	ASXL1	p.V1349A	missense_variant	0.48	CCP
chr20	40713353	PTPRT	p.G1388R	missense_variant	0.54	CCP
				QGP1		
chr12	25398284	KRAS	p.G12V	missense_variant	0.742	CCP,CRC5,PanNEN
chr17	7579393	TP53	p.P98Lfs*75	frameshift_truncation	0.995	CCP,CRC5
chr5	112162854	APC	p.Y486*	stop_gained	0.982	CRC5
chr1	11190767	MTOR	p.R1811H	missense_variant	0.683	PanNEN
chr1	186315299	TPR	p.K1022Q	missense_variant	0.538	CCP
chr2	141946083	LRP1B	p.I307N	missense_variant	0.165	CCP
chr4	106157435	TET2	p.T800S	missense_variant	0.267	CCP
chr4	106196248	TET2	p.Q1550Sfs*49	frameshift_truncation	0.588	CCP
chr5	112177788	APC	p.R2166Q	missense_variant	0.406	CCP
chr9	36966682	PAX5	p.P215L	missense_variant	0.515	CCP

Chromosome	Position	Gene	Amino Acid	Variant Effect	Allelic Frequency	Gene Panel
chr11	102248276	BIRC2	p.L472F	missense_variant	0.443	CCP
chr11	108206666	ATM	p.K2749I	missense_variant	0.65	CCP

6.11   Response coefficients of the respective parameters of BON1 and QGP1 models

TABLE 6.11.1: Response coefficients of the respective parameters of BON1 and QGP1 models. The Parameter column refers to the direct link and the Path Value column represents the response coefficient in log scale.

Parameter	Path Value (r)
QGP1	
AKT->AKTt	-0.1231819
AKTs->AKT	-2.6514043
AKTt->CREB1	0.09819726
AKTt->GSK3	0.07441752
AKTt->PRAS40t	0.46862546
AKTt->mTORC1	0.8740396
AKTt->mTORC2/p70S6K->mTORC2	-47.664468
AZD->mTORC1	0.67935377
AZD->mTORC2/p70S6K->mTORC2	-8.0679253
AZD->p4EBP1	0.26933531
EGFR->RAF->MEK1	0.82193823
ERK1->CREB1	0.4258024
ERK1->EGFR	-0.2732524
ERK1->RAF->MEK1	-0.4254177
ERK1->RPS6	0.28498783
ERK1->p70S6K	0.31350978
GF->EGFR	0.88314364
GF->IGFR	0.93200283
IGFR->PI3K	4.44934968
MEK1->ERK1	1.82778252
PI3K->AKT	-3.4910009
PI3K->RAF->MEK1	0.1750874
PI3K->mTORC2/p70S6K->mTORC2	6.81958436
mTORC1->p4EBP1	0.13568001
mTORC1->p70S6K	0.83144356
p70S6K->mTORC2->AKTs	-0.0402448
p70S6K->mTORC2->IGFR	0.00252605
p70S6K->CREB1	0.14649423
p70S6K->RPS6	0.68450306
iAKT	-8.560621
iAZD	-4.6170035
iERK1	-1.233815
iMEK1	-1.0263373
iPI3K	-0.7872372
imTORC1	-1.8214136
BON1	
AKT->AKTt	1.09861593
AKTs->AKT	0.53790108
AKTt->PRAS40t	0.66215989
AKTt->mTORC1	1.17904843
AKTt->mTORC2/p70S6K->mTORC2	4.36340928
AZD->mTORC1	0.58836534
AZD->mTORC2/p70S6K->mTORC2	0.28242377
AZD->p4EBP1	0.4904269
ERK1->EGFR	-0.6162642
ERK1->GSK3	-0.4245638
ERK1->RAF->MEK1	-0.9847298
ERK1->p70S6K	0.3542028
GF->EGFR	0.64911927
GF->IGFR	0.65195276

Parameter	Path Value (r)
IGFR->PI3K	0.06939884
MEK1->ERK1	0.31650945
PI3K->AKT	0.81012999
PI3K->RAF->MEK1	1.12996237
mTORC1->p4EBP1	0.25318094
mTORC1->p70S6K	0.46922502
p70S6K->mTORC2->AKTs	0.33996021
p70S6K->mTORC2->IGFR	-0.0299503
p70S6K->PI3K	-0.2585075
p70S6K->RPS6	0.96300903
iAKT	-0.9387223
iAZD	-4.2934331
iERK1	-0.7956741
iMEK1	-3.525921
iPI3K	-0.286181
imTORC1	-3.4195615

6.12 All path comparison of BON1 and QGP1 models

TABLE 6.12.1: All path comparison of BON1 and QGP1 models. The aggregated path - signalling path being evaluated, value - minimum parameter value for the parameter. upper bound and lower bound refers to the maximum values in the positive and negative direction, respectively, between which the parameter lay without diminishing the fit of the model.

aggregated path	celline	value	upper bound	lower bound
AKTs->AKT->AKTt	BON1	0.59094669	1.06087183	0.30583311
AKTs->AKT->AKTt	QGP1	0.3266051	1.00989384	0.00514328
AKTt->CREB1	QGP1	0.09819726	0.21819726	-0.0218027
AKTt->GSK3	QGP1	0.07441752	0.15441752	0.00441752
AKTt->mTORC1->p4EBP1	BON1	0.29851259	0.72937385	0.08817344
AKTt->mTORC1->p4EBP1	QGP1	0.1185897	0.38009123	-0.0292822
AKTt->mTORC1->p70S6K->CREB1	QGP1	0.10645949	0.38063484	-0.0043538
AKTt->mTORC1->p70S6K->mTORC2->AKTs	BON1	0.18807925	0.7141351	0.0027874
AKTt->mTORC1->p70S6K->mTORC2->AKTs	QGP1	-0.0292465	0	-0.1617509
AKTt->mTORC1->p70S6K->mTORC2->IGFR	BON1	-0.0165697	0	-0.1065407
AKTt->mTORC1->p70S6K->mTORC2->IGFR	QGP1	0.00183572	0.01555608	-0.0092819
AKTt->mTORC1->p70S6K->PI3K	BON1	-0.1430164	-0.02756769	-Inf
AKTt->mTORC1->p70S6K->RPS6	BON1	0.53277417	1.36760628	0.19394113
AKTt->mTORC1->p70S6K->RPS6	QGP1	0.49743836	1.24749123	0.15466568
AKTt->mTORC2->AKTs	BON1	1.48338552	1.76545722	1.20265027
AKTt->mTORC2->AKTs	QGP1	1.91824844	2.05755847	0
AKTt->mTORC2->IGFR	BON1	-0.1306854	0	-0.2532577
AKTt->mTORC2->IGFR	QGP1	-0.1204027	0.40530625	-0.2220646
AKTt->PRAS40t	BON1	0.66215989	0.95215989	0.46215989
AKTt->PRAS40t	QGP1	0.46862546	0.61862546	0.32862546
EGFR->RAF->MEK1	QGP1	0.82193823	1.44193823	0.23193823
ERK1->CREB1	QGP1	0.4258024	0.6558024	0.2158024
ERK1->EGFR	BON1	-0.6162642	-0.36626425	-0.9062642
ERK1->EGFR	QGP1	-0.2732524	-0.10325235	-0.4532524
ERK1->GSK3	BON1	-0.4245638	-0.22456376	-0.6545638
ERK1->p70S6K->CREB1	QGP1	0.04592737	0.16351767	-0.0018704
ERK1->p70S6K->mTORC2->AKTs	BON1	0.12041486	0.48518706	0.00034067
ERK1->p70S6K->mTORC2->AKTs	QGP1	-0.0126171	0	-0.0694869
ERK1->p70S6K->mTORC2->IGFR	BON1	-0.0106085	0	-0.0723843
ERK1->p70S6K->mTORC2->IGFR	QGP1	0.00079194	0.00668277	-0.0039874
ERK1->p70S6K->PI3K	BON1	-0.0915641	-0.00336923	-Inf
ERK1->p70S6K->RPS6	BON1	0.34110049	0.92915873	0.02370285
ERK1->p70S6K->RPS6	QGP1	0.2145984	0.53591221	0.04250976
ERK1->RAF->MEK1	BON1	-0.9847298	-0.69472975	-1.3047298
ERK1->RAF->MEK1	QGP1	-0.4254177	-0.12541772	-0.7754177
ERK1->RPS6	QGP1	0.28498783	0.77498783	-0.1950122
GF->EGFR	BON1	0.64911927	0.76911927	0.52911927
GF->EGFR	QGP1	0.88314364	1.04314364	0.72314364

aggregated path	celline	value	upper bound	lower bound
GF->IGFR	BON1	0.65195276	0.90195276	0.40195276
GF->IGFR	QGP1	0.93200283	1.24200283	0.63200283
iAKT	BON1	-0.9387223	NA	0
iAKT	QGP1	-8.560621	0	NA
iAZD->mTORC1->p4EBP1	BON1	-0.6395622	0	-Inf
iAZD->mTORC1->p4EBP1	QGP1	-0.425571	Inf	-Inf
iAZD->mTORC1->p70S6K->CREB1	QGP1	-0.3820406	Inf	-Inf
iAZD->mTORC1->p70S6K->mTORC2->AKTs	BON1	-0.4029592	0	-Inf
iAZD->mTORC1->p70S6K->mTORC2->AKTs	QGP1	0.10495402	Inf	0
iAZD->mTORC1->p70S6K->mTORC2->IGFR	BON1	0.03550047	Inf	0
iAZD->mTORC1->p70S6K->mTORC2->IGFR	QGP1	-0.0065876	Inf	-Inf
iAZD->mTORC1->p70S6K->PI3K	BON1	0.30641225	Inf	0
iAZD->mTORC1->p70S6K->RPS6	BON1	-1.1414668	0	-Inf
iAZD->mTORC1->p70S6K->RPS6	QGP1	-1.7851074	0	-Inf
iAZD->mTORC2->AKTs	BON1	-0.4122247	0	-Inf
iAZD->mTORC2->AKTs	QGP1	-1.4991054	0	-Inf
iAZD->mTORC2->IGFR	BON1	0.03631676	Inf	0
iAZD->mTORC2->IGFR	QGP1	0.09409435	Inf	0
iAZD->p4EBP1	BON1	-2.1056151	0	-Inf
iAZD->p4EBP1	QGP1	-1.2435221	0	-Inf
iERK1	BON1	-0.7956741	-0.21567408	-1.4356741
iERK1	QGP1	-1.233815	-0.81381503	-1.753815
IGFR->PI3K	BON1	0.06939884	0.57939884	0.02939884
IGFR->PI3K	QGP1	4.44934968	10.9231535	2.71410331
iMEK1	BON1	-3.525921	-1.2622797	NA
iMEK1	QGP1	-1.0263373	-0.85633731	-1.2363373
imTORC1	BON1	-3.4195615	0	NA
imTORC1	QGP1	-1.8214136	NA	0
iPI3K	BON1	-0.286181	0	NA
iPI3K	QGP1	-0.7872372	0	-1.8072372
MEK1->ERK1	BON1	0.31650945	0.43650945	0.23650945
MEK1->ERK1	QGP1	1.82778252	2.48778252	1.36778252
PI3K->AKT->AKTt	BON1	0.89002172	2.8026066	0.05390303
PI3K->AKT->AKTt	QGP1	0.43002822	1.32281723	0.0783017
PI3K->mTORC2->AKTs	QGP1	-0.274453	0.00166538	-0.274453
PI3K->mTORC2->IGFR	QGP1	0.01722659	0.0277745	-0.0001327
PI3K->RAF->MEK1	BON1	1.12996237	2.33996237	0.01996237
PI3K->RAF->MEK1	QGP1	0.1750874	0.4050874	0.0050874

6.13 Clinical characteristics of patient cohort

TABLE 6.13.1: Clinical characteristics of patient cohort. PNETID - ⋆: samples belong to one patient. Normal - ND: normal distant, NA: normal adjacent. Sample type - FFPE: Formaldehyde Fixed Parafine Embedded. Primary - pancreatic locations are either head, body or tail of the pancreas. Multiple lesions refer to multiple tumors spread within the pancreas. NEC/NET grade WHO 2017 - NETG1: well-differentiated low grade, NETG2: well-differentiated intermediate grade, NETG3: well-differentiated high grade, LCNEC: large cell neuroendocrine carcinoma (poorly differentiated), SCNEC: small cell neuroendocrine carcinoma (poorly differentiated). Metastasis - LN: lymph node metastasis.

PNETID	PanNEN Group	Gender	Normal	Sample type	Primary	Metastasis	NEC/NET grade WHO2017	Ki67 [own analysis or diagnostic report]	Additional diagnosis
					Bern Cohort				
PNET100	Group A	female	ND	FFPE	Papillary region		NETG3	40	ALT+, glucagon+
PNET101	Group B	female	ND	FFPE	Pancreas		NEC	40	
PNET102	Group B	male	ND	FFPE	tail		NEC	50	
PNET103	Group B	female	No normal	FFPE	head		NEC	90	ALT+
PNET104	Group B	male	No normal	FFPE	retro peritoneal		NEC	70	

PNETID	PanNEN Group	Gender	Normal	Sample type	Primary	Metastasis	NEC/NET grade WHO2017	Ki67 [own analysis or diagnostic report]	Additional diagnosis
PNET105	Group B	Male	ND	FFPE	head		NEC	80	
PNET106	Group B	male	ND	FFPE	head		NEC	40	
PNET107	Group A	male	ND	FFPE	head		NETG3	no info table 25% report	
PNET108	Group A	male	ND	FFPE	tail		NETG3	no info table >20% report	
Charite Cohort									
*PNET56P1	Group A	female	ND	FFPE		liver	NETG3	40	
*PNET56P2	Group A		ND	FFPE		liver	NETG3	20	
PNET57	Group A	male	NA	biopsy		liver	NETG3	25	
PNET58	Group B	female	No normal	biopsy		liver	LCNEC	25	
PNET59	Group B	male	No normal	FFPE		bladder	SCNEC	70	
PNET60	Group B	male	ND	FFPE		Peritoneal	LCNEC	40	
PNET61	Group A	male	NA	FFPE	tail		NETG3	25	
PNET62	Group A	female	No normal	biopsy		liver	NETG3	20	
PNET63	Group B	female	No normal	FFPE	head		LCNEC	60	
PNET65	Group C	male	ND	FFPE			LCNEC	40	
PNET66	Group B	male	ND	FFPE		LN	LCNEC	47	
PNET67	Group B	female	NA	FFPE		liver	LCNEC	30	
PNET70	Group A	female	No normal	biopsy		liver	NETG3	25	
PNET2	Group B	male	ND	FFPE	head		SCNEC	47.8670	
PNET4	Group B	male	NA	FFPE	head		NETG2	4.842	
PNET5	Group A	female	No normal	FFPE	head		NETG1	0.23128	
PNET8	Group A	female	NA	FFPE	head		NETG1	1.52856	MEN1 syndrome
PNET9	Group A	female	ND	FFPE	head		NETG2	4.1304	
PNET11	Group A	female	NA	FFPE		LN	NETG1		
PNET14	Group A	female	NA	FFPE	head		NETG1	1.98962	
PNET15	Group A	male	NA	FFPE	multiple lesion		NETG1	0.51052	MEN1 syndrome
PNET16	Group A	female	NA	FFPE	tail		NETG1	2.3538	
PNET18	Group A	male	NA	FFPE	head		NETG1	0.12532	
PNET20	Group A	female	NA	FFPE	head		NETG2	6.6778	
PNET21	Group A	female	NA	FFPE	multiple lesion		NETG2	3.151	
PNET24	Group A	female	NA	FFPE	tail		NETG2	3.6172	
PNET25	Group A	male	No normal	FFPE	body		NETG1	0.70462	Insulinoma
PNET33	Group A	female	ND	FFPE	head		NETG2	3.2214	MEN1 syndrome
PNET36	Group C	female	ND	FFPE	tail		NETG2	4.9006	
PNET37	Group A	male	No normal	FFPE	head		NETG1	0.06298	
PNET42	Group A	male	PB	FFPE		liver	NETG3	52.078	
PNET43	Group A	female	ND	FFPE	body		NETG1	0.39086	
PNET46	Group A	female	ND	FFPE		liver	NETG2	7.6058	
PNET47	Group A	male	ND	FFPE	tail		NETG1	0.59126	
PNET50	Group A	male	ND	FFPE		liver	NETG2	3.4384	
PNET51	Group A	female	ND	FFPE	head		NETG1	0.2888	
PNET52	Group A	female	NA	FFPE	tail		NETG1	0.87594	Insulinoma; MEN1 syndrome
PNET53	Group A	female	ND	FFPE		liver	NETG2	4.4964	
*PNET77P	Group C	female	PB	FFPE	head		NETG3	26	
*PNET77M2	Group C		PB	FFPE		liver	NETG3	25	
PNET79	Group A	male	ND	FFPE	tail and corpus		NETG2	15	
PNET81	Group A	male	ND	FFPE	tail		NETG1	0.76548	
PNET83	Group A	male	ND	FFPE	tail		NETG1	0.50074	
PNET85	Group A	female	ND	FFPE	tail		NETG2	3.3904	
PNET89	Group A	male	ND	FFPE	tail		NETG1	0.74766	
PNET91	Group A	female	ND	FFPE	tail		NETG1	1.922	Insulinoma
PNET92	Group A	male	ND	FFPE	head		NETG1	1.73	
PNET95	Group A	male	NA	FFPE	tail		NETG2	7.7164	

6.14 Molecular characteristics of patient cohort

TABLE 6.14.1: Molecular characteristics of patient cohort. Tumor cell content - pathologist estimation of the tumor purity. T-stage - pathology classification of the tumor stage for a given samples.

PNETID	PanNEN Group	Tumor cell content	T-Stage	Chromogranin A [own analysis or diag. report]	Synaptophysin [own analysis or diag. report]	Staining		
						PDX1	ARX	SOX9
Bern Cohort								
PNET100	Group A	70%	pT3	positive	positive	negative	postive	negative
PNET101	Group B	80%	pT3	strong	strong	negative	negative	positive
PNET102	Group B	80%	pT2			negative	negative	negative
PNET103	Group B	60%	pT3	80%	80%	negative	negative	positive
PNET104	Group B	70%	pT3	strong		negative	negative	Positive
PNET105	Group B	80%	pT3	strong	strong	negative	single cells (negative)	Positive
PNET106	Group B	50%	pT3	focal	strong	negative	negative	negative
PNET107	Group A	60%	pT2			negative	positive	Positive
PNET108	Group A	60	pT3	positive	positive	negative	positive	negative
Charite Cohort								
*PNET56P1	Group A	95%		strong	strong	negative	positive	positive
*PNET56P2	Group A	95%		moderate	strong			
PNET57	Group A	95%		strong	strong			
PNET58	Group B	80%		strong	strong	positive	positive	
PNET59	Group B	95%		strong	strong	negative	positive	positive
PNET60	Group B	95%		partially	partially	negative	negative	positive
PNET61	Group A	90%	pT3	positivity	positivity			
PNET62	Group A	75%		positivity	no staining	positive	positive	
PNET63	Group B	80%	pT1	strong	partially and moderate	negative	negative	positive
PNET65	Group C	90%	pT2	strong	strong			
PNET66	Group B	98%		strong	strong	negative		positive
PNET67	Group B	85%		strong	strong	negative	positive	negative
PNET70	Group A	80%		positivity	positivity			
PNET2	Group B	90%	pT3	3+	1+	negative	positive	positive
PNET4	Group B	80%	pT3	2+	1+			
PNET5	Group A	85%	pT3	1+	2+			
PNET8	Group A	80%		3+	3+			
PNET9	Group A	95%	pT3	2+	3+			
PNET11	Group A	70%		positivity	positivity			
PNET14	Group A	98%	pT3	1+	3+	negative	positive	negative
PNET15	Group A	95%	pT2	2+	3+	negative	positive	negative
PNET16	Group A	95%	pT4	2+	3+			
PNET18	Group A	95%	pT1	2+	2+	negative	negative	negative
PNET20	Group A	95%	pT3	1+	3+			
PNET21	Group A	90%	pT3	2+	3+	negative	positive	negative
PNET24	Group A	80%	pT3	1+	2+			
PNET25	Group A	98%	pT1	positivity	positivity			
PNET33	Group A	75%	pT3	1+	3+	negative	positive	negative
PNET36	Group C	90%	pT3	2+	1+			
PNET37	Group A	95%	pT3	1+	2+			
PNET42	Group A	75%		3+	2+	negative	single cell positive	single cell positive
PNET43	Group A	90%	pT3	2+	2+			
PNET46	Group A	90%		1+	2+			
PNET47	Group A	95%	pT1	2+	3+	negative	positive	negative
PNET50	Group A	95%		1+	3+	negative	negative	negative
PNET51	Group A	90%	pT1	3+	2+	negative	negative	negative
PNET52	Group A	90%	pT3	1+	3+	negative	positive	positive
PNET53	Group A	75%		3+	3+	negative	positive	negative
*PNET77P	Group C	80%	pT3	strong	strong			
*PNET77M2	Group C	90%						
PNET79	Group A		pT3	strong	strong			
PNET81	Group A	95%		2+	3+			
PNET83	Group A	95%	pT2	2+	3+	negative	positive	negative
PNET85	Group A	90%	pT2	2+	3+			
PNET89	Group A	90%	pT1	3+	3+			
PNET91	Group A	95%	pT1	2+	3+	negative	positive	negative
PNET92	Group A	90%	pT2	1+	3+			

PNETID	PanNEN Group	Tumor cell content	T-Stage	Chromogranin A [own analysis or diag. report]	Synaptophysin [own analysis or diag. report]	Staining		
						PDX1	ARX	SOX9
PNET95	Group A	90%	pT3	2+	3+			

Chapter 7

Key Resources

Here you can find the resources that were utilized for the completion of this project. This includes external datasets, experimental resources and key programming functions that were utilized for the appropriate analysis of the data.

7.1 Experimental resources

TABLE 7.1.1: Inhibitors and antibodies

Target	Product	Manufacturer
Inhibitors		
PI3K	GDC0941	Axon
AKT	MK2206	LKT Laboratories
mTORC1 and mTORC2	AZD8055	Cayman Chemical
mTORC1	Everolimus	Cayman Chemical
ERK1	AZD6244	Selleckchem
MEK1	SCH772984	Selleckchem
Antibodies		
$pAKT_{T308}$	#2965	Cell Signaling
$pAKT_{S473}$	#4060	Cell Signaling
$p4EBP1_{S65}$	#9451	Cell Signaling
Vinculin	MAB68969	R&D System
ARX	AF7068	R&D System
PDX1	MAB2419	R&D System
SOX9	#82630	Cell Signaling

7.2 Data analysis resources

TABLE 7.2.1: External dataset

Sample ID	New sample annotation if changed	Dataset
Alpha rep 1	Alpha A	(Neiman et al., 2017)
Alpha rep 2	Alpha B	(Neiman et al., 2017)
Beta rep 1	Beta A	GSE122126
Beta rep 2	Beta B	GSE122126
Beta rep 3	Beta C	(Neiman et al., 2017)
Ductal 1	Ductal A	GSE134217
Ductal 2	Ductal B	GSE134217
Ductal rep 1	Ductal C	GSE122126
Acinar 1	Acinar A	GSE134217
Acinar 2	Acinar B	GSE134217
Acinar rep 1	Acinar C	GSE122126

Sample ID	New sample annotation if changed	Dataset
PDAC (n=167)		GSE49149
hESC		GSE128130

TABLE 7.2.2: Main R packages and their respective functions that were utilized for data processing and analysis. Additional parameter changes deviating from the default setting of the function is detailed in chapter 5

R package (version)	Function	Description
Ape (5.4-1)	`nj()`	performs the neighbor-joining tree estimation from a distance matrix
biomaRt (2.44.1)	`getBM()`	obtaining genomic annotation for a given set of genes
	`useEnsembl()`	define the reference genome to utilize for obtaining annotation
CHAMP (2.18.3)	`champ.DMP()`	Differential methylation analysis between specific groups or all groups
ComplexHeatmap (2.4.3)	`Heatmap()`	generating heatmaps
	`HeatmapAnnotation()`	preparing column-wise annotation data to add to Heatmap()
	`rowAnnotation()`	preparing row-wise annotation data to add to Heatmap()
	`UpSet()`	visualize intersections of multiple sets
ConcensusClusterPlus (1.52.0)	`ConsensusClusterPlus()`	Performs unsupervised class discovery
Conumee	`CNV.bin()`	Combine probes within predefined genomic bins
	`CNV.create_anno()`	Create annotation object required for CNV analysis using 450k or EPIC arrays
	`CNV.detail()`	Analyze copy number profile from a specific regions
	`CNV.fit()`	Normalize a sample to a set of control samples by multiple linear regression. Resulting object contains log2-ratio of probe intensities of the query sample versus the combination of control samples
	`CNV.genomeplot()`	Plots copy number profile of complete genome or specific region
	`CNV.load()`	Combine intensity values from methylated and unmethylated channels
	`CNV.segment()`	Segment the genome into regions of the same copy-number state
dichromat (2.0-0)	`colorRampPalette()`	Generate sequential color scheme
EnhancedVolcano (1.6.0)	`EnhancedVolcano()`	Volcano plot visualization
factoextra (1.0.7)	`get_dist()`	Computes distance matrix
GenomeInfoDb (1.24.2)	`seqlevelsStyle()`	Rename seqlevel in Granges object
ggplot2 (3.3.2)	`geom_boxplot()`	Boxplot visualization of data
	`geom_jitter()`	Evaluation of point distribution without overlap
Iranges (2.22.2)	`subsetByOverlaps()`	Extract overlaping probes
minfi (1.34.0)	`combineArrays()`	Combine EPIC and 450K methylation arrays. Only overlapping probes remain
	`detectionP()`	Generate a detection p-value for every CpG per sample
	`dropLociWithSnps()`	Remove probes with common SNPs
	`getAnnotation()`	Illumina manifest loaded as an R object. Contains information on methylation probes in the respective array
	`getBeta()`	Calculate and extract beta values for CpG probes
	`mapToGenome()`	Convert intesitiy information into M values and beta values, together with associated genomic coordinates; final object generated in a GenomicRatioSet (GRSet)
	`pData()`	Extract phenotype data from RGSet
	`preprocessFunnorm()`	Data normalization of tumor-normal combined RGSet and generate MethylSet (Mset)
	`preprocessSWAN()`	Data normalization of tumor or normal RGSet and generate MethylSet (Mset)
	`read.metharray.exp()`	Read IDAT raw data and generate RGChannelSet (RGSet)
	`read.metharray.sheet()`	Read annotation file containing intensity data (IDAT) file information and addition patient features
regioneR (1.20.1)	`toGRanges()`	Convert getBM() output into Granges object

R package (version)	Function	Description
Rtsne (0.15)	Rtsne()	Dimension reduction analysis of methylation data
RVAideMemoire (0.9-79)	fisher.multcomp()	Performs fisher exact test on multiple variables
STASNet (1.0.2)	aggregateDirectPath()	Aggregated paths from multiple model
	createModel()	Creates parameterized model
	plotModelAccuracy().	Heatmaps of model error and log2 fold changes of experimental data, and simulation data
	plotModelScores()	Scores of measured nodes fit
	plotParameters()	Visualize the aggregatedDirectPath() output
	profileLikelihood()	Computes the profile likelihood and confidence interval
	selectMinimalModel()	Iteratively remove model links and reduce the model without significantly impacting the fit
	suggestExtension()	Additional links are suggested that could improve the model fit
stats (base R)	dist()	Computes distance matrix
	hclust()	Hierarchical clustering of a matrix
	wilcox.test()	Performs wilcoxon rank sum statistical testing
survival (3.2-11)	survfit()	Computes an estimate of a survival curve for censored data
survminer (0.4.9)	ggsurvplot()	Plot survival curves

Chapter 8

Appendix

8.1 List of Abbreviations

A-D-M	**A**TRX,**D**AXX and/or **M**EN1 mutations
AKT	**AK** strain **T**ransforming
ARX	**A**ristaless **R**elated **H**omeobox
ATRX	**A**lpha-**T**halassemia/Mental **R**etardation Syndrome, **X**-Linked
CNA	**C**opy **N**umber **A**berrations
CCP	**C**omprehensive **C**ancer **P**anel
CDKN1B	**C**yclin **D**ependent **K**inase In**h**ibitor **1B**
CDKN2A	**C**yclin **D**ependent **K**inase In**h**ibitor **2A**
COSMIC	**C**atalogue **O**f **S**omatic **M**utations **I**n **C**ancer
DAB	3,3'-**D**i**A**mino**B**enzidine
DAXX	Death Domain Associated Protein 6
DMEM	**D**ulbecco's **M**odified **E**agle **M**edium
DMP	**D**ifferentially **M**ethylated **P**robes
FCS	**F**etal **C**alf **S**erum
FISH	**F**luorescence **I**n-**S**itu **H**ybridization
hESC	**h**uman **E**mbryonic **S**tem **C**ells
IHC	**I**mmuno**H**isto**C**hemistry
IRX2	**IR**oquois Homeobo**X 2**
KRAS	Ki-ras2 **K**irsten **RA**t **S**arcoma viral oncogene
LCNEC	**L**arge **C**ell **N**euro**E**ndocrine **C**arcinoma
MAPK	**M**itogen-**A**ctivated **P**rotein **K**inase
MEN1	**M**ultiple **E**ndocrine **N**eoplasia type 1

MRA	**M**odular **R**esponse **A**nalysis
mTOR	**M**echanistic **T**arget **O**f **R**apamycin
NETG1	**N**euro**E**ndocrine **T**umor **G**rade **1**
NETG2	**N**euro**E**ndocrine **T**umor **G**rade **2**
NETG3	**N**euro**E**ndocrine **T**umor **G**rade **3**
PanNEC	**Pan**creatic **N**euro**E**ndocrine **C**arcinoma
PanNENs	**Pan**creatic **N**euro**E**ndocrine **N**eoplasms
PanNET	**Pan**creatic **N**euro**E**ndocrine **T**umor
PDAC	**P**ancreatic **D**uctal **A**deno**C**arcinoma
PDX1	**P**ancreatic and **D**uodenal Homeobo**X 1**
PI3K	**P**hospho**I**nositide-**3**-**K**inases
RB1	**R**etino**B**lastoma
SCNEC	**S**mall **C**ell **N**euro**E**ndocrine **C**arcinoma
SMAD4	**SMAD** family **M**ember **4**
SOX9	**S**RY-B**OX** Transcription Factor **9**
STASNet	STeady-**ST**ate **A**nalysis of **S**ignalling **Net**works
TP53	**T**umor **P**rotein **53**
TSC2	**T**uberous **S**clerosis **C**omplex **2**
VHL	**V**on **H**ippel-**L**indau

8.2   List of Symbols

α	pancreatic alpha cell
β	pancreatic beta cell
γ	pancreatic gamma cell
δ	pancreatic delta cell
ε	pancreatic epsilon cell
Δ beta	difference in methylation beta values
r	local response coefficient

Bibliography

Agarwal, Sunita K. et al. (1999). "Menin interacts with the AP1 transcription factor JunD and represses JunD-activated transcription". In: *Cell*. ISSN: 00928674. DOI: 10.1016/S0092-8674(00)80967-8.

Al-Tassan, Nada et al. (2002). "Inherited variants of MYH associated with somatic G:C→T:A mutations in colorectal tumors". In: *Nature Genetics*. ISSN: 10614036. DOI: 10.1038/ng828.

Amendt, Brad A., Lillian B. Sutherland, and Andrew F. Russo (1999). "Transcriptional antagonism between Hmx1 and Nkx2.5 for a shared DNA- binding site". In: *Journal of Biological Chemistry* 274.17, pp. 11635–11642. ISSN: 00219258. DOI: 10.1074/jbc.274.17.11635. URL: http://dx.doi.org/10.1074/jbc.274.17.11635.

Anlauf, Martin et al. (2006). "Microadenomatosis of the endocrine pancreas in patients with and without the multiple endocrine neoplasia type 1 syndrome". In: *American Journal of Surgical Pathology*. ISSN: 01475185. DOI: 10.1097/01.pas.0000194044.01104.25.

Arda, H. Efsun, Cecil M. Benitez, and Seung K. Kim (2013). "Gene regulatory networks governing pancreas development". In: *Developmental Cell* 25.1, pp. 5–13. ISSN: 15345807. DOI: 10.1016/j.devcel.2013.03.016. URL: http://dx.doi.org/10.1016/j.devcel.2013.03.016.

Arnold, Christian N. et al. (2007). "Analysis of molecular pathways in sporadic neuroendocrine tumors of the gastro-entero-pancreatic system". In: *International Journal of Cancer*. ISSN: 00207136. DOI: 10.1002/ijc.22569.

Aryee, Martin J. et al. (2014). "Minfi: A flexible and comprehensive Bioconductor package for the analysis of Infinium DNA methylation microarrays". In: *Bioinformatics* 30.10, pp. 1363–1369. ISSN: 14602059. DOI: 10.1093/bioinformatics/btu049.

Bamford, S. et al. (2004). "The COSMIC (Catalogue of Somatic Mutations in Cancer) database and website". In: *British Journal of Cancer* 91.2, pp. 355–358. ISSN: 00070920. DOI: 10.1038/sj.bjc.6601894.

Banck, Michaela S. et al. (2013). "The genomic landscape of small intestine neuroendocrine tumors". In: *Journal of Clinical Investigation* 123.6, pp. 2502–2508. ISSN: 00219738. DOI: 10.1172/JCI67963.

Bartsch, Detlef et al. (1999). "Mutations of the DPC4/Smad4 gene in neuroendocrine pancreatic tumors". In: *Oncogene*. ISSN: 09509232. DOI: 10.1038/sj.onc.1202585.

Basturk, Olca et al. (2014). "Poorly differentiated neuroendocrine carcinomas of the pancreas: A clinicopathologic analysis of 44 cases". In: *American Journal of Surgical Pathology*. DOI: 10.1097/PAS.0000000000000169.

Basturk, Olca et al. (2015). "The high-grade (WHO G3) pancreatic neuroendocrine tumor category is morphologically and biologically heterogenous and includes both well differentiated and poorly differentiated neoplasms". In: *American Journal of Surgical Pathology*. DOI: 10.1097/PAS.0000000000000408.

Belar, Oihana et al. (2012). "Novel mutations in MEN1, CDKN1B and AIP genes in patients with multiple endocrine neoplasia type 1 syndrome in Spain". In: *Clinical Endocrinology*. ISSN: 03000664. DOI: 10.1111/j.1365-2265.2011.04269.x.

Benten, Daniel et al. (2018). "Establishment of the first well-differentiated human pancreatic neuroendocrine tumor model". In: *Molecular Cancer Research*. ISSN: 15573125. DOI: 10.1158/1541-7786.MCR-17-0163.

Bewicke-Copley, Findlay et al. (2019). "Applications and analysis of targeted genomic sequencing in cancer studies". In: *Computational and Structural Biotechnology Journal* 17, pp. 1348–1359. ISSN: 20010370. DOI: 10.1016/j.csbj.2019.10.004. URL: https://doi.org/10.1016/j.csbj.2019.10.004.

Blansfield, Joseph A. et al. (2007). "Clinical, genetic and radiographic analysis of 108 patients with von Hippel-Lindau disease (VHL) manifested by pancreatic neuroendocrine neoplasms (PNETs)". In: *Surgery*. ISSN: 00396060. DOI: 10.1016/j.surg.2007.09.012.

Boora, Ganesh K. et al. (2015). "Exome-level comparison of primary well-differentiated neuroendocrine tumors and their cell lines". In: *Cancer Genetics*. ISSN: 22107770. DOI: 10.1016/j.cancergen.2015.04.002.

Bormann, Felix et al. (2018). "Cell-of-Origin DNA Methylation Signatures Are Maintained during Colorectal Carcinogenesis". In: *Cell Reports* 23.11, pp. 3407–3418. ISSN: 22111247. DOI: 10.1016/j.celrep.2018.05.045.

Bosman, FT et al. (2010). "WHO classification of tumors of digestive system". In: *WHO press*. ISSN: 0007-1447.

Bouwens, L. and D. G. Pipeleers (1998). "Extra-insular beta cells associated with ductules are frequent in adult human pancreas". In: *Diabetologia*. ISSN: 0012186X. DOI: 10.1007/s001250050960.

Bradley, Elizabeth W. et al. (2008). "Pathway crosstalk between Ras/RAF and PI3K in promotion of M-CSF-induced MEK/ERK-mediated osteoclast survival". In: *Journal of Cellular Biochemistry* 104.4, pp. 1439–1451. ISSN: 07302312. DOI: 10.1002/jcb.21719.

Brandt, Raphael et al. (2019). "Cell type-dependent differential activation of ERK by oncogenic KRAS in colon cancer and intestinal epithelium". In: *Nature Communications* 10.1. ISSN: 20411723. DOI: 10.1038/s41467-019-10954-y. URL: http://dx.doi.org/10.1038/s41467-019-10954-y.

Brooks, Jordan C., Robert M. Shavelle, and Kate N. Vavra-Musser (2019). "Life expectancy in pancreatic neuroendocrine cancer". In: *Clinics and Research in Hepatology and Gastroenterology*. ISSN: 2210741X. DOI: 10.1016/j.clinre.2018.08.005.

Bruggeman, Frank J. et al. (2002). "Modular response analysis of cellular regulatory networks". In: *Journal of Theoretical Biology* 218.4, pp. 507–520. ISSN: 00225193. DOI: 10.1016/S0022-5193(02)93096-1.

Cabrera, Over et al. (2006). "The unique cytoarchitecture of human pancreatic islets has implications for islet cell function". In: *Proceedings of the National Academy of Sciences of the United States of America*. ISSN: 00278424. DOI: 10.1073/pnas.0510790103.

Cano, David A. et al. (2014). *Transcriptional control of mammalian pancreas organogenesis*. DOI: 10.1007/s00018-013-1510-2.

Cao, Yanan et al. (2013). "Whole exome sequencing of insulinoma reveals recurrent T372R mutations in YY1". In: *Nature Communications*. ISSN: 20411723. DOI: 10.1038/ncomms3810.

Capella, C. et al. (1995). "Revised classification of neuroendocrine tumours of the lung, pancreas and gut". In: *Virchows Archiv* 425.6, pp. 547–560. ISSN: 09456317. DOI: 10.1007/BF00199342.

Capper, David et al. (2018). "DNA methylation-based classification of central nervous system tumours". In: *Nature* 555.7697, pp. 469–474. ISSN: 14764687. DOI: 10.1038/nature26000.

Casanovas, Oriol et al. (2005). "Drug resistance by evasion of antiangiogenic targeting of VEGF signaling in late-stage pancreatic islet tumors". In: *Cancer Cell*. ISSN: 15356108. DOI: 10.1016/j.ccr.2005.09.005.

Cejas, Paloma et al. (2019). "Enhancer signatures stratify and predict outcomes of nonfunctional pancreatic neuroendocrine tumors". In: *Nature Medicine*. ISSN: 1546170X. DOI: 10.1038/s41591-019-0493-4.

Chan, Chang S. et al. (2018). "ATRX, DAXX or MEN1 mutant pancreatic neuroendocrine tumors are a distinct alpha-cell signature subgroup". In: *Nature Communications* 9.1, pp. 1–10. ISSN: 20411723. DOI: 10.1038/s41467-018-06498-2. URL: http://dx.doi.org/10.1038/s41467-018-06498-2.

Chandrasekharappa, Settara C. et al. (1997). "Positional cloning of the gene for multiple endocrine neoplasia-type 1". In: *Science*. ISSN: 00368075. DOI: 10.1126/science.276.5311.404.

Chen, Yi An et al. (2013). "Discovery of cross-reactive probes and polymorphic CpGs in the Illumina Infinium HumanMethylation450 microarray". In: *Epigenetics* 8.2, pp. 203–209. ISSN: 15592308. DOI: 10.4161/epi.23470.

Childs, Liam Harold et al. (2016). "SoFIA: A data integration framework for annotating high-throughput datasets". In: *Bioinformatics* 32.17, pp. 2590–2597. ISSN: 14602059. DOI: 10.1093/bioinformatics/btw302.

Choi, Diana et al. (2011). "Vhl is required for normal pancreatic B cell function and the maintenance of B cell mass with age in mice". In: *Laboratory Investigation* 91.4, pp. 527–538. ISSN: 00236837. DOI: 10.1038/labinvest.2010.207.

Cives, Mauro et al. (2019). "DAXX mutations as potential genomic markers of malignant evolution in small nonfunctioning pancreatic neuroendocrine tumors". In: *Scientific Reports*. ISSN: 20452322. DOI: 10.1038/s41598-019-55156-0.

Clark, A. et al. (1988). "Islet amyloid, increased A-cells, reduced B-cells and exocrine fibrosis: Quantitative changes in the pancreas in type 2 diabetes". In: *Diabetes Research*. ISSN: 02655985.

Clark, Stephen J. et al. (2018). "ScNMT-seq enables joint profiling of chromatin accessibility DNA methylation and transcription in single cells e". In: *Nature Communications* 9.1, pp. 1–9. ISSN: 20411723. DOI: 10.1038/s41467-018-03149-4. URL: http://dx.doi.org/10.1038/s41467-018-03149-4.

Conemans, Elfi B. et al. (2017). "Prognostic factors for survival of MEN1 patients with Duodenopancreatic tumors metastatic to the liver: results from the DMSG". In: *Endocrine practice : official journal of the American College of Endocrinology and the American Association of Clinical Endocrinologists*. ISSN: 1530891X. DOI: 10.4158/EP161639.OR.

Cooke, Susanna L. et al. (2008). "High-resolution array CGH clarifies events occurring on 8p in carcinogenesis". In: *BMC Cancer* 8, pp. 1–15. ISSN: 14712407. DOI: 10.1186/1471-2407-8-288.

Corbo, Vincenzo et al. (2010). "MEN1 in pancreatic endocrine tumors: Analysis of gene and protein status in 169 sporadic neoplasms reveals alterations in the vast majority of cases". In: *Endocrine-Related Cancer*. ISSN: 13510088. DOI: 10.1677/ERC-10-0028.

Crippa, Stefano et al. (2012). "Surgical management of insulinomas: Short- and long-term outcomes after enucleations and pancreatic resections". In: *Archives of Surgery*. ISSN: 00040010. DOI: 10.1001/archsurg.2011.1843.

Dammann, Reinhard et al. (2003). "Frequent RASSF1A promoter hypermethylation and K-ras mutations in pancreatic carcinoma". In: *Oncogene*. ISSN: 09509232. DOI: 10.1038/sj.onc.1206582.

Daniel Beauchamp, R. and Robert J. Coffey (1991). "Human Carcinoid Cell Production of Paracrine Growth Factors That Can Stimulate Fibroblast and Endothelial Cell Growth". In: *Cancer Research*. ISSN: 15387445.

De La O, Jean Paul et al. (2008). "Notch and Kras reprogram pancreatic acinar cells to ductal intraepithelial neoplasia". In: *Proceedings of the National Academy of Sciences of the United States of America*. ISSN: 00278424. DOI: 10.1073/pnas.0810111105.

Dean, Patrick G. et al. (2000). "Are patients with multiple endocrine neoplasia type I prone to premature death?" In: *World Journal of Surgery*. DOI: 10.1007/s002680010237.

Dejeux, Emelyne et al. (2009). "Hypermethylation of the IGF2 differentially methylated region 2 is a specific event in insulinomas leading to loss-of-imprinting and overexpression". In: *Endocrine-Related Cancer*. ISSN: 13510088. DOI: 10.1677/ERC-08-0331.

DeLellis RA., Lloyd RV., Heitz PU., Eng C (2004). *WHO Classification of Tumors Pathology and Genetics of Tumors of Endocrine Organs*.

Domenico, Annunziata Di et al. (2020). "Epigenetic landscape of pancreatic neuroendocrine tumours reveals distinct cells of origin and means of tumour progression". In: *Nature Communications Biology*, pp. 1–11. ISSN: 2399-3642. DOI: 10.1038/s42003-020-01479-y. URL: http://dx.doi.org/10.1038/s42003-020-01479-y.

Dor, Yuval and Howard Cedar (2018). "Principles of DNA methylation and their implications for biology and medicine". In: *The Lancet* 392.10149, pp. 777–786. ISSN: 1474547X. DOI: 10.1016/S0140-6736(18)31268-6. URL: http://dx.doi.org/10.1016/S0140-6736(18)31268-6.

Dorel, Mathurin et al. (2018). "Modelling signalling networks from perturbation data". In: *Bioinformatics* 34.23, pp. 4079–4086. ISSN: 14602059. DOI: 10.1093/bioinformatics/bty473.

Dorrell, C. et al. (2011). "Transcriptomes of the major human pancreatic cell types". In: *Diabetologia* 54.11, pp. 2832–2844. ISSN: 14320428. DOI: 10.1007/s00125-011-2283-5.

Drané, Pascal et al. (2010). "The death-associated protein DAXX is a novel histone chaperone involved in the replication-independent deposition of H3.3". In: *Genes and Development*. ISSN: 08909369. DOI: 10.1101/gad.566910.

Drewinko, B. et al. (1976). "Establishment of a Human Carcinoembryonic Antigen-producing Colon Adenocarcinoma Cell Line". In: *Cancer Research*. ISSN: 15387445.

Dubois, Fatéméh et al. (2019). *RASSF1A, puppeteer of cellular homeostasis, fights tumorigenesis, and metastasis—an updated review*. DOI: 10.1038/s41419-019-2169-x.

Ebi, Hiromichi et al. (2013). "PI3K regulates MEK/ERK signaling in breast cancer via the Rac-GEF, P-Rex1". In: *Proceedings of the National Academy of Sciences of the United States of America* 110.52, pp. 21124–21129. ISSN: 00278424. DOI: 10.1073/pnas.1314124110.

Espinet, Elisa et al. (2020). "Aggressive PDACs show hypomethylation of repetitive elements and the execution of an intrinsic IFN program linked to a ductal cell-of-origin". In: *Cancer Discovery*. ISSN: 2159-8274. DOI: 10.1158/2159-8290.cd-20-1202.

Espinet, Elisa et al. (2021). "Aggressive pdacs show hypomethylation of repetitive elements and the execution of an intrinsic ifn program linked to a ductal cell of origin". In: *Cancer Discovery* 11.3, pp. 638–659. ISSN: 21598290. DOI: 10.1158/2159-8290.CD-20-1202.

Evers, B. Mark et al. (1991). "Establishment and characterization of a human carcinoid in nude mice and effect of various agents on tumor growth". In: *Gastroenterology*. ISSN: 00165085. DOI: 10.1016/0016-5085(91)90004-5.

Evers, B. Mark et al. (1994). "The Human Carcinoid Cell Line, BON: A Model System for the Study of Carcinoid Tumors". In: *Annals of the New York Academy of Sciences*. ISSN: 17496632. DOI: 10.1111/j.1749-6632.1994.tb17289.x.

Fang, Jiayun M. and Jiaqi Shi (2019). "A clinicopathologic and molecular update of pancreatic neuroendocrine neoplasms with a focus on the New World Health Organization classification". In: *Archives of Pathology and Laboratory Medicine*. DOI: 10.5858/arpa.2019-0338-RA.

Feber, Andrew et al. (2014). "Using high-density DNA methylation arrays to profile copy number alterations". In: *Genome Biology* 15.2. ISSN: 1474760X. DOI: 10.1186/gb-2014-15-2-r30.

Fernandez, Agustin F. et al. (2012). "A DNA methylation fingerprint of 1628 human samples". In: *Genome Research* 22.2, pp. 407–419. ISSN: 10889051. DOI: 10.1101/gr.119867.110.

Fjällskog, Marie Louise H. et al. (2003). "Expression of molecular targets for tyrosine kinase receptor antagonists in malignant endocrine pancreatic tumors". In: *Clinical Cancer Research* 9.4, pp. 1469–1473. ISSN: 10780432.

Francalanci, Paola et al. (2003). "Malignant pancreatic endocrine tumor in a child with tuberous sclerosis". In: *American Journal of Surgical Pathology*. ISSN: 01475185. DOI: 10.1097/00000478-200310000-00012.

Franko, Jan et al. (2010). "Non-functional neuroendocrine carcinoma of the pancreas: Incidence, tumor biology, and outcomes in 2,158 patients". In: *Journal of Gastrointestinal Surgery*. ISSN: 1091255X. DOI: 10.1007/s11605-009-1115-0.

Frankton, Sarah and Stephen R. Bloom (1996). "Glucagonomas". In: *Bailliere's Clinical Gastroenterology*. ISSN: 09503528. DOI: 10.1016/S0950-3528(96)90019-6.

Franzén, Oscar, Li Ming Gan, and Johan L.M. Björkegren (2019). "PanglaoDB: A web server for exploration of mouse and human single-cell RNA sequencing data". In: *Database* 2019.1, pp. 1–9. ISSN: 17580463. DOI: 10.1093/database/baz046.

Fry, David W. et al. (2004). "Specific inhibition of cyclin-dependent kinase 4/6 by PD 0332991 and associated antitumor activity in human tumor xenografts". In: *Molecular Cancer Therapeutics* 3.11, pp. 1427–1437. ISSN: 15357163.

Fujisawa, Takashi et al. (2002). "Malignant endocrine tumor of the pancreas associated with von Recklinghausen's disease". In: *Journal of Gastroenterology*. ISSN: 09441174. DOI: 10.1007/s535-002-8135-x.

Fukushima, Noriyoshi (2017). "Neuroendocrine Neoplasms of the Pancreas: The Pathological Viewpoint". In: *JOP. Journal of the Pancreas*. ISSN: 1590-8577.

Gaiti, Federico et al. (2019). "Epigenetic evolution and lineage histories of chronic lymphocytic leukaemia". In: *Nature* 569.7757, pp. 576–580. ISSN: 14764687. DOI: 10.1038/s41586-019-1198-z. URL: http://dx.doi.org/10.1038/s41586-019-1198-z.

Georgitsi, Marianthi et al. (2007). "Brief report: Germline CDKN1B/p27Kip1 mutation in multiple endocrine neoplasia". In: *Journal of Clinical Endocrinology and Metabolism* 92.8, pp. 3321–3325. ISSN: 0021972X. DOI: 10.1210/jc.2006-2843.

Gerstung, Moritz et al. (2020). "The evolutionary history of 2 , 658 cancers". In: 578.August 2017.

Goldberg, Aaron D. et al. (2010). "Distinct Factors Control Histone Variant H3.3 Localization at Specific Genomic Regions". In: *Cell* 140.5, pp. 678–691. ISSN: 00928674. DOI: 10.1016/j.cell.2010.01.003. URL: http://dx.doi.org/10.1016/j.cell.2010.01.003.

Goudet, Pierre et al. (2010). "Risk factors and causes of death in men1 disease. a gte (groupe d'etude des tumeurs endocrines) cohort study among 758 patients". In: *World Journal of Surgery*. ISSN: 03642313. DOI: 10.1007/s00268-009-0290-1.

Groot, Rolf P. de, Lisa M. Ballou, and Paolo Sassone-Corsi (1994). "Positive regulation of the cAMP-responsive activator CREM by the p70 S6 kinase: An alternative route to mitogen-induced gene expression". In: *Cell* 79.1, pp. 81–91. ISSN: 00928674. DOI: 10.1016/0092-8674(94)90402-2.

Guerra, Carmen et al. (2007). "Chronic Pancreatitis Is Essential for Induction of Pancreatic Ductal Adenocarcinoma by K-Ras Oncogenes in Adult Mice". In: *Cancer Cell*. ISSN: 15356108. DOI: 10.1016/j.ccr.2007.01.012.

Ha, Jain et al. (2019). "Phosphorylation dynamics of jnk signaling: Effects of dual-specificity phosphatases (dusps) on the jnk pathway". In: *International Journal of Molecular Sciences* 20.24, pp. 1–19. ISSN: 14220067. DOI: 10.3390/ijms20246157.

Habbe, Nils et al. (2008). "Spontaneous induction of murine pancreatic intraepithelial neoplasia (mPanIN) by acinar cell targeting of oncogenic Kras in adult mice". In: *Proceedings of the National Academy of Sciences of the United States of America*. ISSN: 00278424. DOI: 10.1073/pnas.0810097105.

Habener, Joel F and Violeta Stanojevic (2012). "α -cell role in β -cell generation and regeneration © 2012 Landes Bioscience . Do not distribute ." In: *Islets* 4.3, pp. 188–198.

Halfdanarson, Thorvardur R. et al. (2008). "Pancreatic neuroendocrine tumors (PNETs): Incidence, prognosis and recent trend toward improved survival". In: *Annals of Oncology* 19.10, pp. 1727–1733. ISSN: 09237534. DOI: 10.1093/annonc/mdn351.

Hammel, P R et al. (2000). "Pancreatic involvement in von Hippel-Lindau disease. The Groupe Francophone d'Etude de la Maladie de von Hippel-Lindau." In: *Gastroenterology*. ISSN: 0016-5085.

Hansel, Donna E. et al. (2003). "Liver metastases arising from well-differentiated pancreatic endocrine neoplasms demonstrate increased VEGF-C expression". In: *Modern Pathology* 16.7, pp. 652–659. ISSN: 08933952. DOI: 10.1097/01.MP.0000077416.68489.50.

Heetfeld, M. et al. (2015). "Characteristics and treatment of patients with G3 gastroenteropancreatic neuroendocrine neoplasms". In: *Endocrine-Related Cancer*. ISSN: 14796821. DOI: 10.1530/ERC-15-0119.

Hijioka, Susumu et al. (2015). "Does the WHO 2010 classification of pancreatic neuroendocrine neoplasms accurately characterize pancreatic neuroendocrine carcinomas?" In: *Journal of Gastroenterology*. ISSN: 14355922. DOI: 10.1007/s00535-014-0987-2.

Hijioka, Susumu et al. (2017). "Rb loss and KRAS mutation are predictors of the response to platinum-based chemotherapy in pancreatic neuroendocrine neoplasm with grade 3: A Japanese multicenter pancreatic NEN-G3 study". In: *Clinical Cancer Research* 23.16, pp. 4625–4632. ISSN: 15573265. DOI: 10.1158/1078-0432.CCR-16-3135.

Hirai, Hiroshi et al. (2010). "MK-2206, an allosteric akt inhibitor, enhances antitumor efficacy by standard chemotherapeutic agents or molecular targeted drugs in vitro and in vivo". In: *Molecular Cancer Therapeutics* 9.7, pp. 1956–1967. ISSN: 15357163. DOI: 10.1158/1535-7163.MCT-09-1012.

Hoadley, Katherine A. et al. (2014). "Multiplatform analysis of 12 cancer types reveals molecular classification within and across tissues of origin". In: *Cell* 158.4, pp. 929–944. ISSN: 10974172. DOI: 10.1016/j.cell.2014.06.049. URL: http://dx.doi.org/10.1016/j.cell.2014.06.049.

Hoadley, Katherine A. et al. (2018). "Cell-of-Origin Patterns Dominate the Molecular Classification of 10,000 Tumors from 33 Types of Cancer". In: *Cell* 173.2, 291–304.e6. ISSN: 10974172. DOI: 10.1016/j.cell.2018.03.022.

Hofving, Tobias et al. (2018). "The neuroendocrine phenotype, genomic profile and therapeutic sensitivity of GEPNET cell lines". In: *Endocrine-Related Cancer*. ISSN: 14796821. DOI: 10.1530/ERC-17-0445.

Holm, Karolina et al. (2016). "An integrated genomics analysis of epigenetic subtypes in human breast tumors links DNA methylation patterns to chromatin states in normal mammary cells". In: *Breast Cancer Research* 18.1, pp. 1–20. ISSN: 1465542X. DOI: 10.1186/s13058-016-0685-5. URL: http://dx.doi.org/10.1186/s13058-016-0685-5.

Hong, Xiafei et al. (2020). "Whole-genome sequencing reveals distinct genetic bases for insulinomas and non-functional pancreatic neuroendocrine tumours: Leading to a new classification system". In: *Gut*. ISSN: 14683288. DOI: 10.1136/gutjnl-2018-317233.

Hood, Fiona E. et al. (2019). "Isoform-specific Ras signaling is growth factor dependent". In: *Molecular Biology of the Cell* 30.9, pp. 1108–1117. ISSN: 19394586. DOI: 10.1091/mbc.E18-10-0676.

Hoogeveen-Westerveld, Marianne et al. (2013). "Functional Assessment of TSC2 Variants Identified in Individuals with Tuberous Sclerosis Complex". In: *Human Mutation* 34.1, pp. 167–175. ISSN: 10981004. DOI: 10.1002/humu.22202.

Hornbeck, Peter V. et al. (2015). "PhosphoSitePlus, 2014: Mutations, PTMs and recalibrations". In: *Nucleic Acids Research* 43.D1, pp. D512–D520. ISSN: 13624962. DOI: 10.1093/nar/gku1267.

House, Michael G. et al. (2003). "Aberrant Hypermethylation Tumor Suppressor Genes in Pancreatic Endocrine Neoplasms". In: *Annals of Surgery*. DOI: 10.1097/01.sla.0000086659.49569.9e.

Houseman, E. Andrés et al. (2009). "Copy number variation has little impact on bead-array-based measures of DNA methylation". In: *Bioinformatics* 25.16, pp. 1999–2005. ISSN: 13674803. DOI: 10.1093/bioinformatics/btp364.

Hovestadt, Volker and Marc Zapatka. *conumeeR*. R package version 1.24.0.

How-Kit, Alexandre et al. (2015). "DNA methylation profiles distinguish different subtypes of gastroenteropancreatic neuroendocrine tumors". In: *Epigenomics*. ISSN: 1750192X. DOI: 10.2217/epi.15.85.

Huang, Jing et al. (2012). "The same pocket in menin binds both MLL and JUND but has opposite effects on transcription". In: *Nature*. ISSN: 00280836. DOI: 10.1038/nature10806.

Hughes, Christina M. et al. (2004). "Menin associates with a trithorax family histone methyltransferase complex and with the Hoxc8 locus". In: *Molecular Cell*. ISSN: 10972765. DOI: 10.1016/S1097-2765(04)00081-4.

Hummel, Charles S. et al. (2011). "Glucose transport by human renal Na+/D-glucose cotransporters SGLT1 and SGLT2". In: *American Journal of Physiology - Cell Physiology* 300.1, pp. 14–21. ISSN: 15221563. DOI: 10.1152/ajpcell.00388.2010.

Iguchi, Haruo, Itsuro Hayashi, and Akira Kono (1990). "A Somatostatin-secreting Cell Line Established from a Human Pancreatic Islet Cell Carcinoma (Somatostatinoma): Release Experiment and Immunohistochemical Study". In: *Cancer Research*. ISSN: 15387445.

Iijima, Yoshihiro et al. (2002). "c-Raf/MEK/ERK pathway controls protein kinase C-mediated p70S6K activation in adult cardiac muscle cells". In: *Journal of Biological Chemistry* 277.25, pp. 23065–23075. ISSN: 00219258. DOI: 10.1074/jbc.M200328200. URL: http://dx.doi.org/10.1074/jbc.M200328200.

Iwasa, Satoru et al. (2010). "Cisplatin and etoposide as first-line chemotherapy for poorly differentiated neuroendocrine carcinoma of the hepatobiliary tract and pancreas". In: *Japanese Journal of Clinical Oncology*. ISSN: 03682811. DOI: 10.1093/jjco/hyp173.

Jennings, Rachel E. et al. (2015). "Human pancreas development". In: *Development (Cambridge)* 142.18, pp. 3126–3137. ISSN: 14779129. DOI: 10.1242/dev.120063.

Jensen, Robert T. et al. (2008). *Inherited pancreatic endocrine tumor syndromes: Advances in molecular pathogenesis, diagnosis, management, and controversies*. DOI: `10.1002/cncr.23648`.

Jensen, Robert T. et al. (2012). *ENETS consensus guidelines for the management of patients with digestive neuroendocrine neoplasms: Functional pancreatic endocrine tumor syndromes*. DOI: `10.1159/000335591`.

Jiao, Yuchen et al. (2011). "DAXX/ATRX, MEN1, and mTOR pathway genes are frequently altered in pancreatic neuroendocrine tumors". In: *Science*. ISSN: 00368075. DOI: `10.1126/science.1200609`.

Jin, Shenghao et al. (2003). "Menin associates with FANCD2, a protein involved in repair of DNA damage". In: *Cancer Research* 63.14, pp. 4204–4210. ISSN: 00085472.

Kaelin, William G. and Eamonn R. Maher (1998). *The VHL tumour-suppressor gene paradigm*. DOI: `10.1016/S0168-9525(98)01558-3`.

Kamura, T. et al. (1999). "Rbx1, a component of the VHL tumor suppressor complex and SCF ubiquitin ligase". In: *Science*. ISSN: 00368075. DOI: `10.1126/science.284.5414.657`.

Keutgen, Xavier M., Naris Nilubol, and Electron Kebebew (2016). "Malignant-functioning neuroendocrine tumors of the pancreas: A survival analysis". In: *Surgery (United States)*. ISSN: 15327361. DOI: `10.1016/j.surg.2015.11.010`.

Kholodenko, Boris N. et al. (2002). "Untangling the wires: A strategy to trace functional interactions in signaling and gene networks". In: *Proceedings of the National Academy of Sciences of the United States of America* 99.20, pp. 12841–12846. ISSN: 00278424. DOI: `10.1073/pnas.192442699`.

Kim, H., A. Kerr, and H. Morehouse (1995). "The association between tuberous sclerosis and insulinoma". In: *American Journal of Neuroradiology*. ISSN: 01956108.

Klinger, Bertram et al. (2013). "Network quantification of EGFR signaling unveils potential for targeted combination therapy". In: *Molecular Systems Biology* 9.1. ISSN: 17444292. DOI: `10.1038/msb.2013.29`.

Klöppel, Günter et al. (2017). "The ENETS and AJCC/UICC TNM classifications of the neuroendocrine tumors of the gastrointestinal tract and the pancreas: A statement". In: *AJSP: Reviews & Reports* 22.5, pp. 233–239. ISSN: 09456317. DOI: `10.1097/PCR.0000000000000211`.

Knudsen, Erik S. et al. (2019). "Cell Cycle and Beyond: Exploiting New RB1 Controlled Mechanisms for Cancer Therapy". In: *Trends in Cancer* 5.5, pp. 308–324. ISSN: 24058033. DOI: `10.1016/j.trecan.2019.03.005`. URL: `https://doi.org/10.1016/j.trecan.2019.03.005`.

Konukiewitz, Björn et al. (2017). "Somatostatin receptor expression related to TP53 and RB1 alterations in pancreatic and extrapancreatic neuroendocrine neoplasms with a Ki67-index above 20%". In: *Modern Pathology*. ISSN: 15300285. DOI: 10.1038/modpathol.2016.217.

Konukiewitz, Björn et al. (2018). "Pancreatic neuroendocrine carcinomas reveal a closer relationship to ductal adenocarcinomas than to neuroendocrine tumors G3". In: *Human Pathology*. ISSN: 15328392. DOI: 10.1016/j.humpath.2018.03.018.

Kopp, Janel L. et al. (2012). "Identification of Sox9-Dependent Acinar-to-Ductal Reprogramming as the Principal Mechanism for Initiation of Pancreatic Ductal Adenocarcinoma". In: *Cancer Cell*. ISSN: 15356108. DOI: 10.1016/j.ccr.2012.10.025.

Kouvaraki, Maria A. et al. (2004). "Fluorouracil, doxorubicin, and streptozocin in the treatment of patients with locally advanced and metastatic pancreatic endocrine carcinomas". In: *Journal of Clinical Oncology*. ISSN: 0732183X. DOI: 10.1200/JCO.2004.04.024.

Krampitz, Geoffrey W. and Jeffrey A. Norton (2013). "Pancreatic neuroendocrine tumors". In: *Current Problems in Surgery* 50.11, pp. 509–545. ISSN: 00113840. DOI: 10.1067/j.cpsurg.2013.08.001.

Kuleshov, Maxim V. et al. (2016). "Enrichr: a comprehensive gene set enrichment analysis web server 2016 update". In: *Nucleic acids research* 44.W1, W90–W97. ISSN: 13624962. DOI: 10.1093/nar/gkw377.

Kulis, Marta et al. (2012). "Epigenomic analysis detects widespread gene-body DNA hypomethylation in chronic lymphocytic leukemia". In: *Nature Genetics*. ISSN: 10614036. DOI: 10.1038/ng.2443.

Kulke, Matthew H. et al. (2011). *Evolving diagnostic and treatment strategies for pancreatic neuroendocrine tumors*. DOI: 10.1186/1756-8722-4-29.

Kuroki, Motomu et al. (1984). "Identification and partial characterization of the unglycosylated peptide of carcinoembryonic antigen synthesized by human tumor cell lines in the presence of tunicamycin". In: *Molecular Immunology*. ISSN: 01615890. DOI: 10.1016/0161-5890(84)90028-2.

La Rosa, Stefano, Fausto Sessa, and Silvia Uccella (2016). *Mixed Neuroendocrine-Nonneuroendocrine Neoplasms (MiNENs): Unifying the Concept of a Heterogeneous Group of Neoplasms*. DOI: 10.1007/s12022-016-9432-9.

Lakis, Vanessa et al. (2021). "DNA methylation patterns identify subgroups of pancreatic neuroendocrine tumors with clinical association". In: *Communications Biology* 4.1, pp. 1–11. ISSN: 23993642. DOI: 10.1038/s42003-020-01469-0.

Lamberti, Giuseppe et al. (2018). "The role of mTOR in neuroendocrine tumors: Future cornerstone of a winning strategy?" In: *International Journal of Molecular Sciences* 19.3. ISSN: 14220067. DOI: 10.3390/ijms19030747.

Latif, Farida et al. (1993). "Identification of the von Hippel-Lindau disease tumor suppressor gene". In: *Science*. ISSN: 00368075. DOI: 10.1126/science.8493574.

Lawrence, Ben et al. (2018). "Recurrent loss of heterozygosity correlates with clinical outcome in pancreatic neuroendocrine cancer". In: *npj Genomic Medicine*. ISSN: 20567944. DOI: 10.1038/s41525-018-0058-3.

Lee, Young Sun et al. (2018). "Glucagon-like peptide 1 increases β-cell regeneration by promoting α- to β-cell transdifferentiation". In: *Diabetes* 67.12, pp. 2601–2614. ISSN: 1939327X. DOI: 10.2337/db18-0155.

Lemos, Manuel C. and Rajesh V. Thakker (2008). "Multiple endocrine neoplasia type 1 (MEN1): Analysis of 1336 mutations reported in the first decade following identification of the gene". In: *Human Mutation*. ISSN: 10597794. DOI: 10.1002/humu.20605.

Lin, Hsiu Ping et al. (2019). "DUSP22 suppresses prostate cancer proliferation by targeting the EGFR-AR axis". In: *FASEB journal : official publication of the Federation of American Societies for Experimental Biology* 33.12, pp. 14653–14667. ISSN: 15306860. DOI: 10.1096/fj.201802558RR.

Lin, Shiaw Yih and Stephen J. Elledge (2003). "Multiple tumor suppressor pathways negatively regulate telomerase". In: *Cell*. ISSN: 00928674. DOI: 10.1016/S0092-8674(03)00430-6.

Lin, Wenchu et al. (2015). "Dynamic epigenetic regulation by menin during pancreatic islet tumor formation". In: *Molecular Cancer Research*. ISSN: 15573125. DOI: 10.1158/1541-7786.MCR-14-0457.

Liou, Geou Yarh et al. (2013). "Macrophage-secreted cytokines drive pancreatic acinar-to-ductal metaplasia through NF-KB and MMPs". In: *Journal of Cell Biology*. ISSN: 00219525. DOI: 10.1083/jcb.201301001.

Liu, Bei et al. (2014). "α-Internexin: A novel biomarker for pancreatic neuroendocrine tumor aggressiveness". In: *Journal of Clinical Endocrinology and Metabolism*. ISSN: 19457197. DOI: 10.1210/jc.2013-2874.

Liu, Lixia et al. (2005). "Epigenetic alterations in neuroendocrine tumors: Methylation of RAS-association domain family 1, isoform a and p16 genes are associated with metastasis". In: *Modern Pathology*. ISSN: 08933952. DOI: 10.1038/modpathol.3800490.

Lo Riso, Pietro et al. (2020). "A cell-of-origin epigenetic tracer reveals clinically distinct subtypes of high-grade serous ovarian cancer". In: *Genome Medicine* 12.1, pp. 1–20. ISSN: 1756994X. DOI: 10.1186/s13073-020-00786-7.

Lomberk, Gwen et al. (2018). "Distinct epigenetic landscapes underlie the pathobiology of pancreatic cancer subtypes". In: *Nature Communications* 9.1, pp. 1–10. ISSN: 20411723. DOI: 10.1038/s41467-018-04383-6. URL: http://dx.doi.org/10.1038/s41467-018-04383-6.

Lopez, Juan R. et al. (2010). "Spectral karyotypic and comparative genomic analysis of the endocrine pancreatic tumor cell line BON-1". In: *Neuroendocrinology*. ISSN: 00283835. DOI: 10.1159/000254483.

Lott, Steven T. et al. (2002). "High frequency loss of heterozygosity in von Hippel-Lindau (VHL)-associated and sporadic pancreatic islet cell tumors: Evidence for a stepwise mechanism for malignant conversion in VHL tumorigenesis". In: *Cancer Research* 62.7, pp. 1952–1955. ISSN: 00085472.

Lu, Jieli (2014). "Transdifferentiation of pancreatic *α*-cells into insulin-secreting cells: From experimental models to underlying mechanisms". In: *World Journal of Diabetes* 5.6, p. 847. ISSN: 1948-9358. DOI: 10.4239/wjd.v5.i6.847.

Lu, Jieli et al. (2010). "*α* Cell-Specific Men1 Ablation Triggers the Transdifferentiation of Glucagon-Expressing Cells and Insulinoma Development". In: *Gastroenterology* 138.5, 1954–1965.e8. ISSN: 00165085. DOI: 10.1053/j.gastro.2010.01.046. URL: http://dx.doi.org/10.1053/j.gastro.2010.01.046.

Lynn, F. C. et al. (2007). "Sox9 coordinates a transcriptional network in pancreatic progenitor cells". In: *Proceedings of the National Academy of Sciences of the United States of America* 104.25, pp. 10500–10505. ISSN: 00278424. DOI: 10.1073/pnas.0704054104.

Maksimovic, Jovana, Lavinia Gordon, and Alicia Oshlack (2012). "SWAN: Subset-quantile within array normalization for illumina infinium HumanMethylation450 BeadChips." In: *Genome biology* 13.6, pp. 1–12. ISSN: 14656914. DOI: 10.1186/gb-2012-13-6-r44.

Mannon, Peter J. (2004). "Pancreatic Polypeptide Family". In: *Encyclopedia of Gastroenterology*. Ed. by Leonard R. Johnson. New York: Elsevier, pp. 78–83. ISBN: 978-0-12-386860-2. DOI: https://doi.org/10.1016/B0-12-386860-2/00564-5. URL: https://www.sciencedirect.com/science/article/pii/B0123868602005645.

Marinoni, Ilaria et al. (2014). "Loss of DAXX and ATRX are associated with chromosome instability and reduced survival of patients with pancreatic neuroendocrine tumors". In: *Gastroenterology*. ISSN: 15280012. DOI: 10.1053/j.gastro.2013.10.020.

Marx, Stephen J. and William F. Simonds (2005). *Hereditary hormone excess: Genes, molecular pathways, and syndromes*. DOI: 10.1210/er.2003-0037.

Maxwell, Patrick H. et al. (1999). "The tumour suppressor protein VHL targets hypoxia-inducible factors for oxygen-dependent proteolysis". In: *Nature*. ISSN: 00280836. DOI: 10.1038/20459.

McCall, Chad M. et al. (2013). "Grading of well-differentiated pancreatic Neuroendocrine tumors is improved by the inclusion of both ki67 Proliferative index and mitotic rate". In: *American Journal of Surgical Pathology*. ISSN: 01475185. DOI: 10.1097/PAS.0000000000000089.

Mei, Mei et al. (2009). "Clinical implications of microsatellite instability and MLH1 gene inactivation in sporadic insulinomas". In: *Journal of Clinical Endocrinology and Metabolism*. ISSN: 0021972X. DOI: 10.1210/jc.2009-0173.

Mendoza, Michelle C., E. Emrah Er, and John Blenis (2011). "The Ras-ERK and PI3K-mTOR pathways: Cross-talk and compensation". In: *Trends in Biochemical Sciences* 36.6, pp. 320–328. ISSN: 09680004. DOI: 10.1016/j.tibs.2011.03.006. URL: http://dx.doi.org/10.1016/j.tibs.2011.03.006.

Mermel, Craig H. et al. (2011). "GISTIC2.0 facilitates sensitive and confident localization of the targets of focal somatic copy-number alteration in human cancers". In: *Genome Biology* 12.4, pp. 1–14. ISSN: 14747596. DOI: 10.1186/gb-2011-12-4-r41.

Merritt, J. Lawrence et al. (2006). "Extensive acrochordons and pancreatic islet-cell tumors in tuberous sclerosis associated with TSC2 mutations". In: *American Journal of Medical Genetics, Part A*. ISSN: 15524825. DOI: 10.1002/ajmg.a.31351.

Metz, David C. and Robert T. Jensen (2008). *Gastrointestinal Neuroendocrine Tumors: Pancreatic Endocrine Tumors*. DOI: 10.1053/j.gastro.2008.05.047.

Milione, Massimo et al. (2016). "The clinicopathologic heterogeneity of grade 3 gastroenteropancreatic neuroendocrine neoplasms: Morphological differentiation and proliferation identify different prognostic categories". In: *Neuroendocrinology*. ISSN: 14230194. DOI: 10.1159/000445165.

Missiaglia, Edoardo et al. (2010). "Pancreatic endocrine tumors: Expression profiling evidences a role for AKT-mTOR pathway". In: *Journal of Clinical Oncology*. ISSN: 0732183X. DOI: 10.1200/JCO.2008.21.5988.

Mohl, M. et al. (2006). "Gef10 - The third member of a Rho-specific guanine nucleotide exchange factor subfamily with unusual protein architecture". In: *Naunyn-Schmiedeberg's Archives of Pharmacology* 373.5, pp. 333–341. ISSN: 00281298. DOI: 10.1007/s00210-006-0083-0.

Moore, Patrick S. et al. (2001). "Role of disease-causing genes in sporadic pancreatic endocrine tumors: MEN1 and VHL". In: *Genes Chromosomes and Cancer*. ISSN: 10452257. DOI: 10.1002/gcc.1180.

Moran, Sebastian et al. (2016). "Epigenetic profiling to classify cancer of unknown primary: a multicentre, retrospective analysis". In: *The Lancet Oncology* 17.10, pp. 1386–1395. ISSN: 14745488. DOI: 10.1016/S1470-2045(16)30297-2.

Moss, Joshua et al. (2018). "Comprehensive human cell-type methylation atlas reveals origins of circulating cell-free DNA in health and disease". In: *Nature Communications* 9.1. ISSN: 20411723. DOI: 10.1038/s41467-018-07466-6. URL: http://dx.doi.org/10.1038/s41467-018-07466-6.

Murai, Marcelo J. et al. (2011). "Crystal structure of menin reveals binding site for Mixed Lineage Leukemia (MLL) protein". In: *Journal of Biological Chemistry*. ISSN: 00219258. DOI: 10.1074/jbc.M111.258186.

Muraro, Mauro J. et al. (2016). "A Single-Cell Transcriptome Atlas of the Human Pancreas". In: *Cell Systems* 3.4, 385–394.e3. ISSN: 24054720. DOI: 10.1016/j.cels.2016.09.002.

Nagano, Yasuhiko et al. (2007). "Allelic alterations in pancreatic endocrine tumors identified by genome-wide single nucleotide polymorphism analysis". In: *Endocrine-Related Cancer*. ISSN: 13510088. DOI: 10.1677/ERC-06-0090.

Neiman, Daniel et al. (2017). "Islet cells share promoter hypomethylation independently of expression, but exhibit cell-type-specific methylation in enhancers". In: *Proceedings of the National Academy of Sciences of the United States of America* 114.51, pp. 13525–13530. ISSN: 10916490. DOI: 10.1073/pnas.1713736114.

Oakes, Christopher C. et al. (2016). "DNA methylation dynamics during B cell maturation underlie a continuum of disease phenotypes in chronic lymphocytic leukemia". In: *Nature Genetics*. ISSN: 15461718. DOI: 10.1038/ng.3488.

Ohki, Rieko et al. (2014). "PHLDA3 is a novel tumor suppressor of pancreatic neuroendocrine tumors". In: *Proceedings of the National Academy of Sciences of the United States of America*. ISSN: 10916490. DOI: 10.1073/pnas.1319962111.

Padi, Sathish K.R. et al. (2019). "Phosphorylation of DEPDC5, a component of the GATOR1 complex, releases inhibition of mTORC1 and promotes tumor growth". In: *Proceedings of the National Academy of Sciences of the United States of America* 116.41, pp. 20505–20510. ISSN: 10916490. DOI: 10.1073/pnas.1904774116.

Pal, Rituraj et al. (2017). "Inhibition of ERK1/2 Restores GSK3β Activity and Protein Synthesis Levels in a Model of Tuberous Sclerosis". In: *Scientific Reports* 7.1, pp. 1–10. ISSN: 20452322. DOI: 10.1038/s41598-017-04528-5.

Parekh, Dilipkumar et al. (1994). "Characterization of a human pancreatic carcinoid in vitro: Morphology, amine and peptide storage, and secretion". In: *Pancreas*. ISSN: 15364828. DOI: 10.1097/00006676-199401000-00013.

Parrales, Alejandro et al. (2013). "ERK1/2-dependent activation of mTOR/mTORC1/p70S6K regulates thrombin-induced RPE cell proliferation". In: *Cellular Signalling* 25.4, pp. 829–838. ISSN: 08986568. DOI: 10.1016/j.cellsig.2012.12.023. URL: http://dx.doi.org/10.1016/j.cellsig.2012.12.023.

Passacantilli, Ilaria et al. (2014). "Combined therapy with RAD001 e BEZ235 overcomes resistance of PET immortalized cell lines to mTOR inhibition". In: *Oncotarget*. ISSN: 19492553. DOI: 10.18632/oncotarget.2111.

Patani, Hemalvi et al. (2020). "Transition to naïve human pluripotency mirrors pan-cancer DNA hypermethylation". In: *Nature Communications* 11.1, pp. 1–17. ISSN: 20411723. DOI:

10.1038/s41467-020-17269-3. URL: http://dx.doi.org/10.1038/s41467-020-17269-3.

Pause, Arnim et al. (1997). "The von Hippel-Lindau tumor-suppressor gene product forms a stable complex with human CUL-2, a member of the Cdc53 family of proteins". In: *Proceedings of the National Academy of Sciences of the United States of America*. ISSN: 00278424. DOI: 10.1073/pnas.94.6.2156.

Pavel, Marianne et al. (2012). *ENETS consensus guidelines for the management of patients with liver and other distant metastases from neuroendocrine neoplasms of foregut, midgut, hindgut, and unknown primary*. DOI: 10.1159/000335597.

Pavel, Marianne E. et al. (2011). "Everolimus plus octreotide long-acting repeatable for the treatment of advanced neuroendocrine tumours associated with carcinoid syndrome (RADIANT-2): A randomised, placebo-controlled, phase 3 study". In: *The Lancet* 378.9808, pp. 2005–2012. ISSN: 01406736. DOI: 10.1016/S0140-6736(11)61742-X. URL: http://dx.doi.org/10.1016/S0140-6736(11)61742-X.

Pea, Antonio et al. (2018). "Genetic Analysis of Small Well-differentiated Pancreatic Neuroendocrine Tumors Identifies Subgroups With Differing Risks of Liver Metastases". In: *Annals of Surgery*. ISSN: 0003-4932. DOI: 10.1097/sla.0000000000003022.

Perren, Aurel et al. (2007). "Multiple endocrine neoplasia type 1 (MEN1): Loss of one MEN1 allele in tumors and monohormonal endocrine cell clusters but not in islet hyperplasia of the pancreas". In: *Journal of Clinical Endocrinology and Metabolism*. ISSN: 0021972X. DOI: 10.1210/jc.2006-1944.

Prada, Elke Tatjana Aristizabal et al. (2018). "The role of GSK3 and its reversal with GSK3 antagonism in everolimus resistance". In: *Endocrine-Related Cancer*. ISSN: 14796821. DOI: 10.1530/ERC-18-0159.

Puri, Sapna, David A. Cano, and Matthias Hebrok (2009). "A role for von hippel-lindau protein in pancreatic β-cell function". In: *Diabetes* 58.2, pp. 433–441. ISSN: 00121797. DOI: 10.2337/db08-0749.

Qadir, Mirza Muhammad Fahd et al. (2018). "P2RY1/ALK3-Expressing Cells within the Adult Human Exocrine Pancreas Are BMP-7 Expandable and Exhibit Progenitor-like Characteristics". In: *Cell Reports* 22.9, pp. 2408–2420. ISSN: 22111247. DOI: 10.1016/j.celrep.2018.02.006.

Qadir, Mirza Muhammad Fahd et al. (2020). "Single-cell resolution analysis of the human pancreatic ductal progenitor cell niche". In: *Proceedings of the National Academy of Sciences of the United States of America* 117.20, pp. 10876–10887. ISSN: 10916490. DOI: 10.1073/pnas.1918314117.

Rahier, J., R. M. Goebbels, and J. C. Henquin (1983). "Cellular composition of the human diabetic pancreas". In: *Diabetologia*. ISSN: 0012186X. DOI: 10.1007/BF00251826.

Raue, Andreas et al. (2009). "Structural and practical identifiability analysis of partially observed dynamical models by exploiting the profile likelihood". In: *Bioinformatics* 25.15, pp. 1923–1929. ISSN: 13674803. DOI: 10.1093/bioinformatics/btp358.

Ray, Kevin C. et al. (2011). "Epithelial tissues have varying degrees of susceptibility to KrasG12D-initiated tumorigenesis in a mouse model". In: *PLoS ONE*. ISSN: 19326203. DOI: 10.1371/journal.pone.0016786.

Razmara, Masoud, Azita Monazzam, and Britt Skogseid (2018). "Reduced menin expression impairs rapamycin effects as evidenced by an increase in mTORC2 signaling and cell migration". In: *Cell Communication and Signaling*. ISSN: 1478811X. DOI: 10.1186/s12964-018-0278-2.

Rehm, Heidi L. (2013). "Disease-targeted sequencing: A cornerstone in the clinic". In: *Nature Reviews Genetics* 14.4, pp. 295–300. ISSN: 14710056. DOI: 10.1038/nrg3463.

Riehn, Mathias et al. (2011). "A BACH2-BCL2L1 Fusion Gene Resulting from a Lymphoma Cell Line BLUE-1". In: *Cancer* 396.January, pp. 389–396. DOI: 10.1002/gcc.

Rigaud, Gildas et al. (2001). "High resolution allelotype of nonfunctional pancreatic endocrine tumors: Identification of two molecular subgroups with clinical implications". In: *Cancer Research*. ISSN: 00085472.

Rindi, G. et al. (2012). "TNM staging of neoplasms of the endocrine pancreas: Results from a large international cohort study". In: *Journal of the National Cancer Institute*. ISSN: 00278874. DOI: 10.1093/jnci/djs208.

Rodríguez-Paredes, Manuel et al. (2018). "Methylation profiling identifies two subclasses of squamous cell carcinoma related to distinct cells of origin". In: *Nature Communications*. ISSN: 20411723. DOI: 10.1038/s41467-018-03025-1.

Rovira, Meritxell et al. (2010). "Isolation and characterization of centroacinar/terminal ductal progenitor cells in adult mouse pancreas". In: *Proceedings of the National Academy of Sciences of the United States of America* 107.1, pp. 75–80. ISSN: 00278424. DOI: 10.1073/pnas.0912589107.

Roy, Nilotpal et al. (2015). "Brg1 promotes both tumor-suppressive and oncogenic activities at distinct stages of pancreatic cancer formation". In: *Genes and Development*. ISSN: 15495477. DOI: 10.1101/gad.256628.114.

Roy, Somak et al. (2018). "Loss of Chromatin-Remodeling Proteins and/or CDKN2A Associates With Metastasis of Pancreatic Neuroendocrine Tumors and Reduced Patient Survival Times". In: *Gastroenterology*. ISSN: 15280012. DOI: 10.1053/j.gastro.2018.02.026.

Rozengurt, Enrique (2014). "Mechanistic target of rapamycin (mTOR): a point of convergence in the action of insulin/ IGF-1 and G protein-coupled receptor agonists in pancreatic cancer cells". In: *Frontiers in Physiology* 5.September, pp. 1–8. ISSN: 1664-042X. DOI: 10.3389/fphys.2014.00357.

Rozengurt, Enrique, Heloisa P. Soares, and James Sinnet-Smith (2014). "Suppression of feedback loops mediated by pi3k/mtor induces multiple overactivation of compensatory pathways: An unintended consequence leading to drug resistance". In: *Molecular Cancer Therapeutics* 13.11, pp. 2477–2488. ISSN: 15388514. DOI: 10.1158/1535-7163.MCT-14-0330.

Ruiz, Anne et al. (2018). "Human mutations in SLC2A9 (Glut9) affect transport capacity for urate". In: *Frontiers in Physiology* 9.JUN. ISSN: 1664042X. DOI: 10.3389/fphys.2018.00476.

Rydenfelt, Mattias et al. (2021). "SPEED2: Inferring upstream pathway activity from differential gene expression". In: *Nucleic Acids Research* 48.W1, W307–W312. ISSN: 13624962. DOI: 10.1093/NAR/GKAA236.

Sarbassov, Dos D. et al. (2005). "Phosphorylation and regulation of Akt/PKB by the rictor-mTOR complex". In: *Science* 307.5712, pp. 1098–1101. ISSN: 00368075. DOI: 10.1126/science.1106148.

Scarpa, Aldo (2019). "The landscape of molecular alterations in pancreatic and small intestinal neuroendocrine tumours". In: *Annales d'Endocrinologie* 80.3, pp. 153–158. ISSN: 00034266. DOI: 10.1016/j.ando.2019.04.010. URL: https://doi.org/10.1016/j.ando.2019.04.010.

Scarpa, Aldo et al. (2017). "Whole-genome landscape of pancreatic neuroendocrine tumours". In: *Nature* 543.7643, pp. 65–71. ISSN: 14764687. DOI: 10.1038/nature21063.

Schmitt, A. M. et al. (2009). "VHL inactivation is an important pathway for the development of malignant sporadic pancreatic endocrine tumors". In: *Endocrine-Related Cancer*. ISSN: 13510088. DOI: 10.1677/ERC-08-0297.

Schubert, Michael et al. (2018). "Perturbation-response genes reveal signaling footprints in cancer gene expression". In: *Nature Communications* 9.1. ISSN: 20411723. DOI: 10.1038/s41467-017-02391-6. URL: http://dx.doi.org/10.1038/s41467-017-02391-6.

Schwarzkopf, G. and J. Pfisterer (1994). "Metastasizing gastrinoma and tuberous sclerosis complex. Association or coincidence?" In: *Zentralblatt für Pathologie*. ISSN: 08634106.

Segerstolpe, Åsa et al. (2016). "Single-Cell Transcriptome Profiling of Human Pancreatic Islets in Health and Type 2 Diabetes". In: *Cell Metabolism* 24.4, pp. 593–607. ISSN: 19327420. DOI: 10.1016/j.cmet.2016.08.020.

Seymour, Philip A. et al. (2007). "SOX9 is required for maintenance of the pancreatic progenitor cell pool". In: *Proceedings of the National Academy of Sciences of the United States of America* 104.6, pp. 1865–1870. ISSN: 00278424. DOI: 10.1073/pnas.0609217104.

Shen, H. C.Jennifer et al. (2010). "Multiple endocrine neoplasia type 1 deletion in pancreatic α-cells leads to development of insulinomas in mice". In: *Endocrinology* 151.8, pp. 4024–4030. ISSN: 00137227. DOI: 10.1210/en.2009-1251.

Sherr, Charles J. (1996). "Cancer cell cycles". In: *Science* 274.5293, pp. 1672–1674. ISSN: 00368075. DOI: 10.1126/science.274.5293.1672.

Shinmura, Kazuya et al. (2000). "Adenine excisional repair function of MYH protein on the adenine:8-hydroxyguanine base pair in double-stranded DNA". In: *Nucleic Acids Research*. ISSN: 03051048. DOI: 10.1093/nar/28.24.4912.

Shroff, Stuti et al. (2014). "SOX9: A useful marker for pancreatic ductal lineage of pancreatic neoplasms". In: *Human Pathology* 45.3, pp. 456–463. ISSN: 00468177. DOI: 10.1016/j.humpath.2013.10.008. URL: http://dx.doi.org/10.1016/j.humpath.2013.10.008.

Sidhu, Anita (1998). "Coupling of D1 and D5 dopamine receptors to multiple G proteins". In: *Molecular Neurobiology* 16.2, pp. 125–134. ISSN: 0893-7648. DOI: 10.1007/bf02740640.

Sowa, Hideaki et al. (2003). "Inactivation of menin, the product of the multiple endocrine neoplasia type 1 gene, inhibits the commitment of multipotential mesenchymal stem cells into the osteoblast lineage". In: *Journal of Biological Chemistry*. ISSN: 00219258. DOI: 10.1074/jbc.M302044200.

Sowa, Hideaki et al. (2004). "Menin is required for bone morphogenetic protein 2- and transforming growth factor β-regulated osteoblastic differentiation through interaction with Smads and Runx2". In: *Journal of Biological Chemistry*. ISSN: 00219258. DOI: 10.1074/jbc.M401312200.

Stefan, Y. et al. (1982). "Quantitation of endocrine cell content in the pancreas of nondiabetic and diabetic humans". In: *Diabetes*. ISSN: 00121797. DOI: 10.2337/diab.31.8.694.

Stefanoli, Michele et al. (2014). "Prognostic relevance of aberrant DNA methylation in G1 and G2 pancreatic neuroendocrine tumors". In: *Neuroendocrinology*. ISSN: 14230194. DOI: 10.1159/000365449.

Tang, Laura H. et al. (2016). "Well-differentiated neuroendocrine tumors with a morphologically apparent high-grade component: A pathway distinct from poorly differentiated neuroendocrine carcinomas". In: *Clinical Cancer Research*. ISSN: 15573265. DOI: 10.1158/1078-0432.CCR-15-0548.

Tazawa, Shigeki et al. (2005). "SLC5A9/SGLT4, a new Na +-dependent glucose transporter, is an essential transporter for mannose, 1,5-anhydro-D-glucitol, and fructose". In: *Life Sciences* 76.9, pp. 1039–1050. ISSN: 00243205. DOI: 10.1016/j.lfs.2004.10.016.

Thomas, Philippe et al. (2012). "GeneView: A comprehensive semantic search engine for PubMed". In: *Nucleic Acids Research* 40.W1, pp. 585–591. ISSN: 03051048. DOI: 10.1093/nar/gks563.

Tian, Yuan et al. (2017). "ChAMP: Updated methylation analysis pipeline for Illumina BeadChips". In: *Bioinformatics* 33.24, pp. 3982–3984. ISSN: 14602059. DOI: 10.1093/bioinformatics/btx513.

Torii, Seiji et al. (1999). "Human Daxx regulates Fas-induced apoptosis from nuclear PML oncogenic domains (PODs)". In: *EMBO Journal*. ISSN: 02614189. DOI: 10.1093/emboj/18.21.6037.

Torrino, Stéphanie et al. (2011). "The E3 ubiquitin-ligase HACE1 catalyzes the ubiquitylation of active Rac1". In: *Developmental Cell* 21.5, pp. 959–965. ISSN: 15345807. DOI: 10.1016/j.devcel.2011.08.015.

Townsend, Courtney M., Jin Ishizuka, and James C. Thompson (1993). "Studies of growth regulation in a neuroendocrine cell line". In: *Acta Oncologica*. ISSN: 0284186X. DOI: 10.3109/02841869309083900.

Tsuda, Motoyuki et al. (2018). "The BRG1/SOX9 axis is critical for acinar cell-derived pancreatic tumorigenesis". In: *Journal of Clinical Investigation*. ISSN: 15588238. DOI: 10.1172/JCI94287.

Tucker, O. N., P. L. Crotty, and K. C. Conlon (2006). *The management of insulinoma*. DOI: 10.1002/bjs.5280.

Ushiku, Hideki et al. (2016). "Homeobox-only protein expression is a critical prognostic indicator of pancreatic neuroendocrine tumor and is regulated by promoter DNA hypermethylation". In: *Pancreas*. ISSN: 15364828. DOI: 10.1097/MPA.0000000000000646.

Valentino, Joseph D. et al. (2014). "Cotargeting the PI3K and RAS pathways for the treatment of neuroendocrine tumors". In: *Clinical Cancer Research*. ISSN: 15573265. DOI: 10.1158/1078-0432.CCR-13-1897.

Vandamme, T. et al. (2019). "Hotspot DAXX, PTCH2 and CYFIP2 mutations in pancreatic neuroendocrine neoplasms". In: *Endocrine-Related Cancer* 26.1, pp. 1–12. ISSN: 14796821. DOI: 10.1530/ERC-18-0120.

Vandamme, Timon et al. (2015). "Whole-exome characterization of pancreatic neuroendocrine tumor cell lines BON-1 and QGP-1". In: *Journal of Molecular Endocrinology*. ISSN: 14796813. DOI: 10.1530/JME-14-0304.

Vandamme, Timon et al. (2016). "Long-term acquired everolimus resistance in pancreatic neuroendocrine tumours can be overcome with novel PI3K-AKT-mTOR inhibitors". In: *British Journal of Cancer*. ISSN: 15321827. DOI: 10.1038/bjc.2016.25.

Vélayoudom-Céphise, Fritz Line et al. (2013). "Are G3 ENETS neuroendocrine neoplasms heterogeneous?" In: *Endocrine-Related Cancer*. ISSN: 13510088. DOI: 10.1530/ERC-13-0027.

Von Figura, Guido et al. (2014). "The chromatin regulator Brg1 suppresses formation of intraductal papillary mucinous neoplasm and pancreatic ductal adenocarcinoma". In: *Nature Cell Biology*. ISSN: 14764679. DOI: 10.1038/ncb2916.

Wilkerson, Matthew D. and D. Neil Hayes (2010). "ConsensusClusterPlus: A class discovery tool with confidence assessments and item tracking". In: *Bioinformatics* 26.12, pp. 1572–1573. ISSN: 13674803. DOI: 10.1093/bioinformatics/btq170.

Williamson, Laura M. et al. (2019). "Genomic characterization of a welldifferentiated grade 3 pancreatic neuroendocrine tumor". In: *Cold Spring Harbor Molecular Case Studies*. ISSN: 23732873. DOI: 10.1101/mcs.a003814.

Wong, Lee H. et al. (2010). "ATRX interacts with H3.3 in maintaining telomere structural integrity in pluripotent embryonic stem cells". In: *Genome Research*. ISSN: 10889051. DOI: 10.1101/gr.101477.109.

Xiao, Xiangwei et al. (2018). "Endogenous Reprogramming of Alpha Cells into Beta Cells, Induced by Viral Gene Therapy, Reverses Autoimmune Diabetes". In: *Cell Stem Cell* 22.1, 78–90.e4. ISSN: 18759777. DOI: 10.1016/j.stem.2017.11.020. URL: https://doi.org/10.1016/j.stem.2017.11.020.

Xue, Wen et al. (2012). "A cluster of cooperating tumor-suppressor gene candidates in chromosomal deletions". In: *Proceedings of the National Academy of Sciences of the United States of America* 109.21, pp. 8212–8217. ISSN: 00278424. DOI: 10.1073/pnas.1206062109.

Yachida, Shinichi et al. (2012). "Small cell and large cell neuroendocrine carcinomas of the pancreas are genetically similar and distinct from well-differentiated pancreatic neuroendocrine tumors". In: *American Journal of Surgical Pathology*. ISSN: 01475185. DOI: 10.1097/PAS.0b013e3182417d36.

Yadav, Siddhartha, Prabin Sharma, and Dana Zakalik (2018). "Comparison of Demographics, Tumor Characteristics, and Survival between Pancreatic Adenocarcinomas and Pancreatic Neuroendocrine Tumors: A Population-based Study". In: *American Journal of Clinical Oncology: Cancer Clinical Trials*. ISSN: 1537453X. DOI: 10.1097/COC.0000000000000305.

Yamaguchi, Junpei et al. (2016). "Loss of Trefoil Factor 2 From Pancreatic Duct Glands Promotes Formation of Intraductal Papillary Mucinous Neoplasms in Mice". In: *Gastroenterology*. ISSN: 15280012. DOI: 10.1053/j.gastro.2016.07.045.

Yamasaki, Ichiro et al. (2006). "Clinical characteristics of pancreatic neuroendocrine tumors in Japanese patients with von Hippel-Lindau disease". In: *Pancreas* 33.4, pp. 382–385. ISSN: 08853177. DOI: 10.1097/01.mpa.0000240604.26312.e4.

Yang, Yu Ping et al. (2011). "Context-specific αa-to-β-cell reprogramming by forced Pdx1 expression". In: *Genes and Development* 25.16, pp. 1680–1685. ISSN: 08909369. DOI: 10.1101/gad.16875711.

Yao, James C. et al. (2008). "Efficacy of RAD001 (everolimus) and octreotide LAR in advanced low- to intermediate-grade neuroendocrine tumors: Results of a phase II study". In: *Journal of Clinical Oncology* 26.26, pp. 4311–4318. ISSN: 0732183X. DOI: 10.1200/JCO.2008.16.7858.

Yao, James C. et al. (2010). "Daily oral everolimus activity in patients with metastatic pancreatic neuroendocrine tumors after failure of cytotoxic chemotherapy: A phase II trial".

In: *Journal of Clinical Oncology* 28.1, pp. 69–76. ISSN: 0732183X. DOI: 10.1200/JCO.2009.24.2669.

Yao, James C. et al. (2011). "Everolimus for Advanced Pancreatic Neuroendocrine Tumors". In: *New England Journal of Medicine*. ISSN: 0028-4793. DOI: 10.1056/nejmoa1009290.

Yu, Dan et al. (2015). "Decreased expression of dual specificity phosphatase 22 in colorectal cancer and its potential prognostic relevance for stage IV CRC patients". In: *Tumor Biology* 36.11, pp. 8531–8535. ISSN: 14230380. DOI: 10.1007/s13277-015-3588-7.

Yuan, Fei et al. (2014). "KRAS and DAXX/ATRX gene mutations are correlated with the clinicopathological features, advanced diseases, And poor prognosis in Chinese patients with pancreatic neuroendocrine tumors". In: *International Journal of Biological Sciences* 10.9, pp. 957–965. ISSN: 14492288. DOI: 10.7150/ijbs.9773.

Zanini, Sara et al. (2020). "mTOR Pathway in Gastroenteropancreatic Neuroendocrine Tumor (GEP-NETs)". In: *Frontiers in Endocrinology* 11.November. ISSN: 16642392. DOI: 10.3389/fendo.2020.562505.

Zhang, Zhen et al. (2019). "A new way for beta cell neogenesis: Transdifferentiation from alpha cells induced by glucagon-like peptide 1". In: *Journal of Diabetes Research* 2019. ISSN: 23146753. DOI: 10.1155/2019/2583047.

Zitzmann, Kathrin et al. (2010). "Compensatory activation of Akt in response to mTOR and Raf inhibitors - A rationale for dual-targeted therapy approaches in neuroendocrine tumor disease". In: *Cancer Letters*. ISSN: 03043835. DOI: 10.1016/j.canlet.2010.02.018.

Erklärung:

Hiermit erkläre ich, die Dissertation selbstständig und nur unter Verwendung der angegebenen Hilfen und Hilfsmittel angefertigt zu haben.

Ich habe mich anderwärts nicht um einen Doktorgrad beworben und besitze keinen entsprechenden Dok- torgrad.

Ich erkläre, dass ich die Dissertation oder Teile davon nicht bereits bei einer anderen wissenschaftlichen Einrichtung eingereicht habe und dass sie dort weder angenommen noch abgelehnt wurde.

Ich erkläre die Kenntnisnahme der dem Verfahren zugrunde liegenden Promotionsord- nung der Lebenswis- senschaftlichen Fakultät der Humboldt-Universität zu Berlin vom 5. März 2015.

Weiterhin erkläre ich, dass keine Zusammenarbeit mit gewerblichen Promotionsbearbeiterinnen/Promoti- onsberatern stattgefunden hat und dass die Grund- sätze der Humboldt-Universität zu Berlin zur Sicherung guter wissenschaftlicher Praxis eingehalten wurden.

Berlin, August 30, 2021
Tincy Simon