Dissertation
submitted to the
Combined Faculties for the Natural Sciences and for Mathematics
of the Ruperto-Carola University of Heidelberg, Germany
for the degree of
Doctor of Natural Sciences

presented by

Master of Science Cornelia Jäkel, née Siebenkäs
born in: Nürnberg, Germany
Oral-examination: 08.05.2017

Genome-wide genetic and epigenetic analyses of pancreatic acinar cell carcinomas reveal aberrations in genome stability and cell cycle control

Referees: PD Dr. Odilia Popanda

Prof. Dr. Christoph Plass

Contributions

Sections of this thesis are based on a draft of the submitted manuscript “Genome-wide genetic and epigenetic analyses of pancreatic acinar cell carcinomas reveal aberrations in cell cycle control and genome stability”[1] and are put between quotation marks throughout the thesis. The draft was originally written by me as first author; however parts of the text in these sections might contain suggestions and corrections from co-authors.

450K arrays, library preparation and whole exome sequencing were run at the Genomics and Proteomics Core Facility at the DKFZ.

The Data Management and Genomics IT (eilslabs) at the DKFZ aligned sequencing data and called SNVs.

Dr. Reka Toth (Division of Epigenomics and Cancer Risk Factors, DKFZ) wrote the R script for the integrative analysis for Figure 48.

Dr. Yassen Assenov (Division of Epigenomics and Cancer Risk Factors, DKFZ) wrote the R script for calculating cell type contributions and plotting the DNA methylation phylogenetic tree for Figure 14 and Figure 17.

PD Dr. Frank Bergmann (Heidelberg University Hospital) provided human tissue samples and performed immunohistochemical stainings in Figure 51a and Supplementary Figure 15a.

Dr. med. Henrik Einwächter (Technical University of Munich) provided an unpublished ACC mouse model and cell lines used for experiments in section 4.4.7.

Under my supervision, Daniel van der Duin performed qPCR analysis and MassARRAY in Figure 22, Figure 50c, d and Supplementary Figure 10.

Declaration

Declarations according to § 8 (3) b) and c) of the doctoral degree regulations:

a) I hereby declare that I have written the submitted dissertation myself and in this process have used no other sources or materials than those expressly indicated,

b) I hereby declare that I have not applied to be examined at any other institution, nor have I used the dissertation in this or any other form at any other institution as an examination paper, nor submitted it to any other faculty as a dissertation.

Heidelberg, 08.03.2017

_____

(Cornelia Jäkel)

An meine Großmütter Hilde und Mari, die so stolz auf mich gewesen wären

Table of Contents

Table of Contents

Summary

Pancreatic acinar cell carcinomas (ACC) are rare pancreatic cancers, which affect mainly adult patients in their sixth decade of life, however may also affect young children. Due to late arising clinical symptoms, tumors are often already large at the time of diagnosis but ACC-specific therapeutic options are lacking. This results in poor survival rates of ACC patients. Due to the rarity of this tumor, collections of tumor tissues and patient data are generally scarce. Consequently, studies published up to date are often case studies or molecular studies using only few tumor samples and genome-wide data sets are very limited. Most studies focusing on target genes known in the much more common pancreatic ductal adenocarcinomas have not been successful in identifying frequent recurrent aberrations in ACC. To this end, this thesis aims to unravel the genome- and epigenome-wide molecular aberrations in ACC, thereby investigating point mutations, mutational signatures, DNA methylation, and copy number aberrations (CNA). Two independent cohorts with a total of 73 tumors are included, representing one of the largest tissue-based collections for ACC. The analyses reveal that, although ACC show a high mutational load per tumor, the mutated genes are not frequently recurring. Somatic signatures of mutational processes are calculated based on the sequence context of point mutations and regardless of the genes affected and are similar among different tumors. Mainly signatures due to tobacco consumption and, more interestingly, defective DNA repair mechanisms are identified. DNA methylation patterns of ACC are compared to normal pancreatic tissues, including sorted pancreatic cell types, and to other pancreatic cancers. Analyses demonstrate that acinar cells are the likely cell of origin of ACC. Further, ACC display a distinct methylation pattern compared to normal pancreatic tissues, pancreatic ductal adenocarcinoma, and neuroendocrine tumors. Differentially methylated genes are enriched in pathways involved in embryonic development and cell adhesion pathways. As example the protocadherin cluster is depicted with wide-spread hypermethylation which influences RNA expression. Massive CNA are detected in ACC and many CNA identified are shared among the tumors. Furthermore, many cancer-relevant genes map to these regions. Mixed acinar-neuroendocrine carcinomas display a very similar molecular pattern compared to pure ACC, suggesting these tumor types belong to the same tumor entity. ACC metastases do not display additional molecular events, thus primary tumors already harbor the potential to metastasize. An integrative analysis identifies aberrations in the four tumor suppressor genes *ARID1A*, *APC*, *CDKN2A*, and *ID3*, which are confirmed by immunohistochemistry. Taken together, this thesis shows that ACC harbor numerous genomic and epigenomic aberrations which mainly concern

genome stability and cell cycle control. These can be exploited by targeted therapies in basket trials which are the most promising approach for patients with rare cancers in which traditional clinical trial designs are not feasible.

Zusammenfassung

Azinuszellkarzinome der Bauspeicheldrüse (ACC) sind seltene Tumore, welche vor allem erwachsene Patienten in ihrem sechsten Lebensjahrzent betreffen, aber auch in Kindern vorkommen können. Aufgrund der spät auftretenden Symptome sind die Tumore zum Zeitpunkt der Diagnose oft sehr groß und es gibt keine ACC-spezifischen therapeutischen Optionen, was zu den schlechten Überlebensraten von ACC-Patienten beiträgt. Wegen der Seltenheit dieses Tumors sind Gewebe- und Datenkollektive in der Regel sehr beschränkt. Folglich sind die meisten bisher veröffentlichten Studien Fallstudien oder molekulare Studien mit nur wenigen Tumoren und Daten auf der genomweiten Ebene sind sehr begrenzt. Die meisten Studien konzentrieren sich auf solche Zielgene, die in den viel häufiger auftretenden pankreatischen duktalen Adenokarzinomen bekannt sind, finden diese Aberrationen in ACC aber nicht. Daher ist das Ziel der hier vorgelegten Studie, molekulare Aberrationen im ACC auf der genom- und epigenomweiten Ebene zu untersuchen, und dabei Punktmutationen, Mutations-Signaturen, DNA-Methylierung, und Kopienzahl-Aberrationen (CNA) zu erforschen. Diese Arbeit umfasst zwei unabhängige Kohorten mit insgesamt 73 Tumoren, welche eine der größten Gewebe-basierten Sammlungen für ACC darstellen. Die Analysen zeigen, dass ACC eine hohe Mutationsrate pro Tumor aufweisen, die mutierten Gene aber nicht häufig wiederkehrend sind. Somatische Signaturen von Mutationsprozessen werden auf der Grundlage des Sequenzkontextes von Punktmutationen und unabhängig von den betroffenen Genen berechnet und sind unter den verschiedenen ACC ähnlich. Dazu gehören Signaturen, welche aufgrund von Tabakkonsum und fehlerhaften DNA-Reparaturmechanismen entstehen. Darüber hinaus werden DNA-Methylierungsmuster von ACC mit anderen Pankreaskarzinomen und mit normalen Pankreasgeweben verglichen, einschließlich reiner Pankreaszelltypen. Die Ähnlichkeit der Muster zeigt, dass Azinuszellen wahrscheinlich die Ursprungszellen von ACC sind. Zudem zeigen ACC ein einzigartiges Methylierungsmuster im Vergleich zu normalem Pankreasgewebe und den anderen Bauspeicheldrüsenkrebsarten, duktales Adenokarzinom und neuroendokrine Tumoren. Differentiell methylierte Gene sind für Signalwege der embryonalen Entwicklung und Zelladhäsion angereichert. Ein Beispiel in dieser Studie ist der Protocadherin-Cluster mit weit verbreiteter Hypermethylierung, welche auch Einfluss auf die RNA Expression hat. CNA werden in ACC sehr häufig nachgewiesen und mehrere identifizierte Regionen kommen in vielen der Tumoren vor. Viele Krebs-relevante Gene liegen in diesen Regionen. Gemischte azinär-neuroendokrine Karzinome zeigen ein sehr ähnliches molekulares Muster wie reine ACC, was darauf hindeutet, dass diese Tumore einer Tumorentität angehören. Darüber hinaus zeigen ACC-Metastasen keine zusätzlichen molekularen Ereignisse, Primärtumore

haben somit bereits das Potenzial zur Metastasierung. Eine integrative Analyse identifiziert Aberrationen in den vier Tumorsuppressorgenen *ARID1A*, *APC*, *CDKN2A* und *ID3*, die durch Immunhistochemie bestätigt werden. Zusammengenommen zeigt diese Arbeit, dass ACC zahlreiche genomische und epigenetische Aberrationen aufweisen, die vor allem Prozesse der Genomstabilität und der Zellzykluskontrolle zuzuordnen sind. Dies kann durch gezielte Therapien in sogenannten Basket-Studien ausgenutzt werden. Dies ist der vielversprechendste Therapieansatz für Patienten mit seltenen Krebsarten, bei denen traditionelle klinische Studiendesigns nicht möglich sind.

List of abbreviations

°C	degrees Celsius
µl	microliter
µM	µmol/l
450K	Illumina's 450K BeadChip Array
A	adenine
ACC	acinar cell carcinoma
aCGH	array comparative genomic hybridization
ADM	acinar-to-ductal metaplasia
ATAC	Assay for Transposase Accessible Chromatin
AWD	alive with disease
BER	base excision repair
bHLH	basic helix-loop helix protein
BLCA	bladder urothelial carcinoma (TCGA)
bp	base pairs
BRCA	breast cancer (TCGA)
C	cytosine
CDK	cyclin-dependent kinase
CNA	copy number aberrations
COAD	colon adenocarcinoma (TCGA)
COSMIC	Catalogue of somatic mutations in cancer
CpG	base C followed by base G
CpGi	CpG island
DAC	5-Aza-2'-deoxycytidine
dCas	deactivated Cas
DMR	differentially methylated region
DMS	differentially methylated site

DOC	dead of other cause
DOD	dead of disease
DSB	double-strand break
DTT	Dithiothreitol
EMA	European Medicines Agency
ENCODE	Encyclopedia of DNA Elements
EU	European Union
F	forward primer
FANTOM	functional annotation of the mammalian genome
FDA	Food and Drug Administration
FF	fresh frozen
FFPE	formalin fixed paraffin embedded
FLAGS	FrequentLy mutAted GeneS
G	guanine
GBM	glioblastoma multiforme (TCGA)
GOI	gene of interest
h	hours
H3K27ac	acetylation of the lysine 27 in histone 3
H3K27me3	triple methylation of the lysine 27 in histone 3
H3K4me1	mono methylation of the lysine 4 in histone 3
H3K4me3	triple methylation of the lysine 4 in histone 3
H3K9me3	triple methylation of the lysine 9 in histone 3
HMEC	human mammary epithelial cells
HMM	Hidden Markov Model
HNSC	head and neck squamous cell carcinomas (TCGA)
HOMER	Hypergeometric Optimization of Motif EnRichment
HR	homologous recombination

IHC	immunohistochemistry
kb	kilo base pair
KIRC	kidney renal clear cell carcinoma (TCGA)
l	liter
lncRNA	long non-coding RNA
LTFU	long -term follow-up
LUAD	lung adenocarcinoma (TCGA)
LUMP	leukocytes unmethylation for purity
LUSC	lung squamous cell carcinoma (TCGA)
MACNEC	mixed acinar-neuroendocrine carcinoma
Mb	mega base pair
min	minutes
miRNA	micro RNA
ml	milliliter
mM	millimol/l
MMR	mismatch repair
MSP	methylation-specific polymerase chain reaction
NED	no evidence of disease
NER	nucleotide excision repair
NHEJ	non-homologous end joining
No.	number
OV	ovarian serous cystadenocarcinoma (TCGA)
PAAD-TCGA	PDAC dataset from TCGA
PanIN	pancreatic intraepithelial neoplasia
PCA	principal component analysis
PCDH	protocadherin family, human
pcdh	protocadherin family, mouse

pcdhg@	constant exons of mouse pcdhg genes
PDAC	pancreatic ductal adenocarcinoma
PFTC	Primary fallopian tube carcinoma
PNET	pancreatic neuroendocrine tumors
PPC	primary peritoneal cancer
PRAD	prostate adenocarcinoma (TCGA)
q-PCR	quantitative polymerase chain reaction
qRT-PCR	quantitative reverse transcriptase polymerase chain reaction
R	reverse primer
RCC	renal cell carcinoma
rcf	relative centrifugal force
READ	rectal adenocarcinoma (TCGA)
rpm	revolutions per minute
RT	room temperature
sec	seconds
SEGA	subependymal giant cell astrocytoma
siRNA	Small interfering RNA
SKCM	skin cutaneous melanoma (TCGA)
ss	single strand
STAD	Stomach Adenocarcinoma (TCGA)
T	thymine
TCGA	The Cancer Genome Atlas
THCA	Thyroid carcinoma (TCGA)
UCEC	uterine corpus endometrial carcinoma (TCGA)
USA	United States of America
WES	whole exome sequencing

1 Introduction

1.1 The hallmarks of cancer

Cancer is the uncontrolled spread and growth of cells. Hallmarks that cells need to achieve to initiate and maintain tumor growth were defined by Hanahan and Weinberg[2,3] and include sustaining proliferative signaling, evading growth suppressors, avoiding immune destruction, enabling replicative immortality, activating invasion and metastasis, inducing angiogenesis, resisting cell death and deregulating cellular energetics[2,3]. Two aspects are considered to enable these hallmarks, namely genome instability and mutations and tumor-promoting inflammation (Figure 1)[2,3]. Genome instability leads to a higher rate of new random aberrations in cells and therefore to a higher chance of malignant transformation whereas inflammation triggers a pro-proliferative environment, including growth factors, pro-angiogenic factors and activation of tissue stem cells. For this thesis, the hallmark of evading growth suppressors and the enabling characteristic of genome instability will be further addressed.

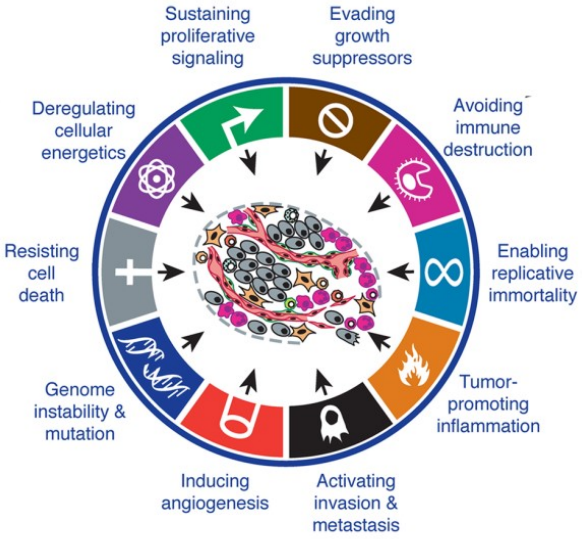


Figure 1 The hallmarks of cancer. Adapted from Hanahan and Weinberg[2,3].

1.1.1 The hallmark of evading growth suppressors

Besides activating pro-proliferative signaling, a cancer cell needs to escape growth inhibitors which are tightly regulating cell proliferation. These growth inhibitors act to either force a cell from a proliferative state into the quiescent G0-state or to force a cell into a final post mitotic state. Cancer cells must evade these two scenarios to survive. Two of the most prominent growth suppressors are Retinoblastoma 1 (RB1) and Tumor Protein 53 (TP53). Whereas TP53

can push a cell to undergo apoptosis, both RB1 and TP53 are negative regulators of the cell cycle[4-6]. The cell cycle is regulated by a tight interplay of cyclins (CCN) and cyclin-dependent

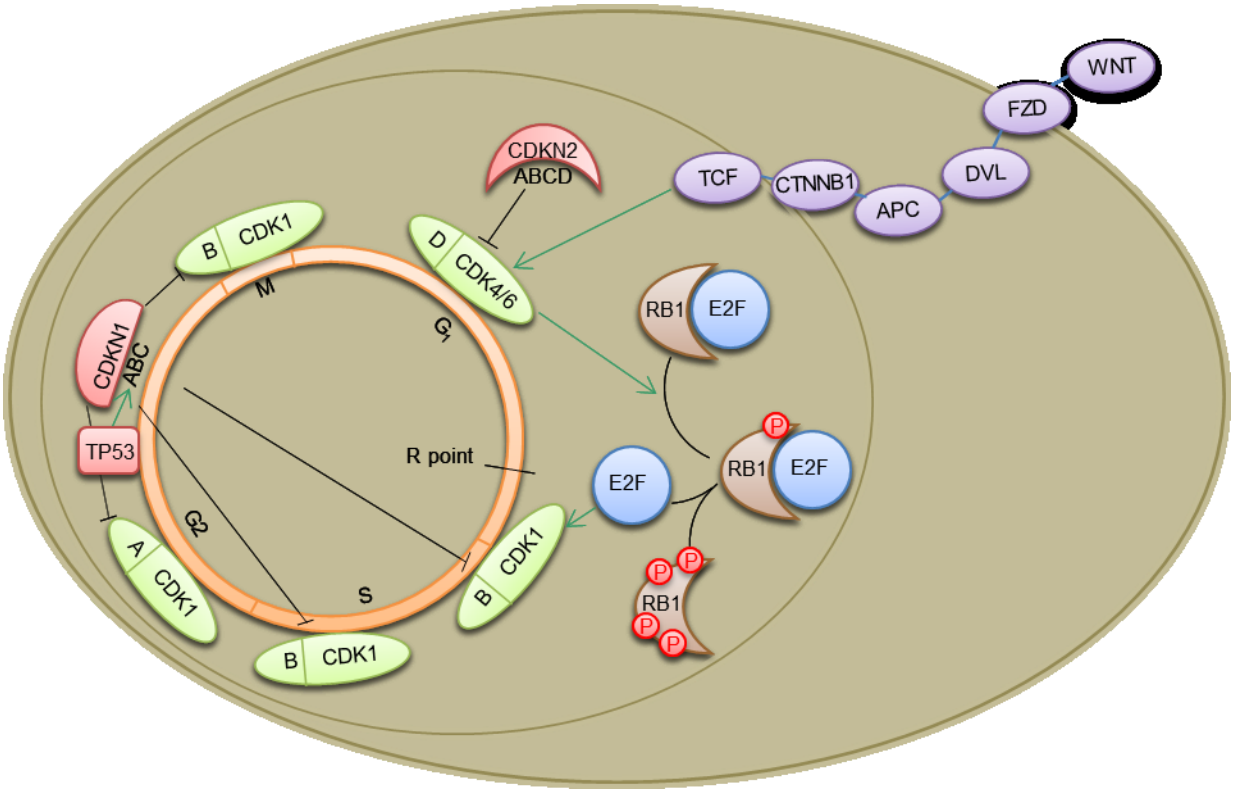


Figure 2 Regulators of the cell cycle including RB1 and the WNT pathway.

kinases (CDK). The cyclin D-CDK4/6 complex is active during G1- Phase, whereas cyclin E, A and B together with CDK2 and CDK1 regulate the rest of the cycle (Figure 2)[7,8]. The cell cycle is additionally regulated by a number of CDK inhibitors CDKN1A, CDKN1B, CDKN1C, CDKN2A, CDKN2B, CDKN2C, CDKN2D, formerly known as $p21^{Cip1}$, $p27^{Kip1}$, $p57^{Kip2}$, $p16^{INK4A}$, $p15^{INK4B}$, $p18^{INK4C}$, $p19^{INK4D}$. While CDKN1A, -B, and -C inhibit the CCN-CDK complexes of E-CDK2, A-CDK2, A-CDK1 and B-CDK1; CDKN2A, -B, -C, -D inhibit the CCND-CDK4/6 complex (Figure 2)[7,8]. When RB1 is unphosphorylated it binds and inhibits the transcription factor E2F. CCND/CDK4/6 can hypo-phosphorylate RB1, a prerequisite for passing the restriction point (R point; checkpoint towards the end of G1 phase where a cell still responds to external growth factors)[6,9]. CCNE/CDK2 then hyper-phosphorylates and inactivates RB1, releasing E2F, which in a positive feedback loop activates its own transcription and a number of target genes, allowing the cell to pursue the cell cycle and surmount the R point (Figure 2)[6,7,9]. TP53 impacts the cell cycle by activating CDKN1A (Figure 2)[7,10]. A vast number of signaling cascades crosstalks with the cell cycle, including the WNT pathway (Figure 2). Inactivation of any of the

cell cycle suppressors CDKN1A, -B, -C, CDKN2A, -B, -C, -D, RB1, and TP53 therefore leads to a constant firing of the cell cycle by a loss of control of the pro-proliferating signals of cyclins. Indeed, inactivating events of these genes have been described in many cancer entities. Inactivation of CDKN2A is amongst the most prominent alterations in cancer[11,12], but CDKN2B, CDKN1A, and CDKN1B alterations are also frequently detected in a number of tumors[12,13]. In addition, RB1 is a very frequent target in many cancer entities, including the eponymous cancer retinoblastoma[6,13-15].

1.1.2 The enabling characteristics of genome instability

For cancer cells to acquire the above mentioned hallmarks their genome and epigenome needs to be altered permanently. Otherwise they do not obtain a growth advantage. Therefore, cancer cells undergo a number of selective clonal states acquiring more aberrations in their genome and epigenome, thus cells are growing from a normal tissue state to precursor lesions to cancer and potentially to metastatic cancer. In the following the different modes of aberrations in the genome and epigenome of cancer and the machinery that protects the genome from these aberrations will be introduced.

1.1.2.1 Point mutations

Point mutations contribute to tumorigenesis in most cancers, *e.g.* activating *KRAS* mutations which are found in nearly all pancreatic ductal adenocarcinomas (PDAC) and are thought to be drivers of tumor development in these cancers[16,17]. The number of mutations per tumor can range from 0.001 mutations per mega base pair (Mb) to more than 400 mutations per Mb in different tumor entities and is correlated with patient age[18-20]. In the beginnings of elucidating cancer genomes, candidate genes had to be preselected for Sanger sequencing[21]. This has vigorously changed since high throughput sequencing has enabled sequencing of whole genomes or exomes in an unbiased way. Today, cancer genomes can be sequenced relatively cheap and fast. The major challenge in recent years is thus not to detect a point mutation, but to identify driver mutations which give a cell a growth advantage, in contrast to bystander mutations which do not contribute to the malignancy of tumors[20].

1.1.2.2 Somatic signatures of mutational processes

Another quite recent approach of investigating mutations based on whole genome or whole exome sequencing (WGS and WES) in tumors is by investigating the occurrence of somatic signatures of mutational processes (also termed mutational signatures)[19,22-26]. These signatures are independent from genomic location, as the kind of mutation (*i.e.* C to A, C to T, C to G, T to

A, T to C, and T to G) is considered in the sequence context of one base upstream and one base downstream. Frequency plots of these base triplets reveal different frequencies amongst tumor entities. Employing non-negative matrix factorization reveals that these mutational patterns can be explained by 30 reoccurring somatic signatures across tumor entities[19,22-26]. Each tumor and each tumor type exhibits different numbers and combinations of these signatures. Some signatures, *e.g.* signature 1 and signature 5 occur in nearly every tumor type and every tumor sample. Other signatures occur specifically only in one tumor type, *e.g.* signature 12 in liver cancer and signature 28 in stomach cancer. Until today, 16 out of these 30 signatures have a likely or assumed cause of the patterns (depicted in Figure 3 and reviewed in Helleday, et al. [24] and Alexandrov [27]). For example, signature 1 is due to a deamination of 5-

Figure 3 Somatic signatures in human cancers. Signatures were depicted by type of base pair substitution and the preceding and following base pair, resulting in 96 frequency plots of base pair triplets. Likely causes of signatures were depicted below the signatures. 5-mC: deamination of 5-methylcytosine. AID/APOBEC: hyperactivity of AID/APOBEC, BRCA1/2: loss of *BRCA1* or *BRCA2*, MMR: DNA mismatch repair deficiency, poly: polymerase. Figure adapted from the Cosmic website[26].

methylcytosine[19], which is a natural process in a cell and which is increasing with age and cell division rate[19] (refer to 1.1.2.4). Signatures 2 and 13 are generated by a hyperactivity of the AID/APOBEC family proteins[19,22], while signatures 9 and 10 are due to defective DNA polymerases η and ε[19], respectively. Signatures 3, 6, 15, 20 and 26 create a pattern caused by defects in DNA repair[19]. More specific, signature 3 is caused by mutations in *BRCA1* or *BRCA2*[22], whereas the remaining three signatures are due to defects in the mismatch repair

machinery (refer to 1.1.2.5). A number of exogenous factors can also contribute to the signatures. Signature 4 and 29 occur after tobacco consumption[28], signature 7 after UV light exposure[19], and signature 11 after consumption of alkylating agents, *e.g.* the cytostatic agent temozolomide[19]. Signature 22 occurs upon aristolochid acid exposure and signatures 24 upon aflatoxin exposure. Thus, theses somatic signatures give rise to the possibility to trace back the potential reason of tumor development (reviewed in Hollstein, et al. [29]). These studies are still in the early phases of development; however can give rise to the unique possibility to identify potential risk factors in rare tumors.

1.1.2.3 Copy number changes

Copy number aberrations (CNA) which include gains and losses of DNA stretches or more complex rearrangements can inactivate tumor suppressor genes or activate oncogenes[30,31]. These aberrations are usually longer than one kilo base pair (kb), however can affect a whole chromosome arm, whole chromosomes or even the whole cancer genome[12]. Tumors exhibiting CNA are present in every tumor entity[12,32,33], however are highly variably between different tumor types and tumor samples[12]. However some amplified and deleted genes are common in many tumor types, including deletions of *CDKN2A*, *STK11*, and *PDE4D* and amplifications in *CCND1*, *EGRF*, and *MYC*[12,32]. In tumors with many small sized CNA, the DNA repair machinery can be impaired, including *BRCA1/2* mutations and aberrations in DNA mismatch repair (refer to 1.1.2.5). Additionally, when telomeres are lost in tumors, breakage-fusion-bridge cycles lead to many large deletions and amplifications by improper fusion to sister chromatids[34-36]. CNA have massively been investigated locus-specific by fluorescence in situ hybridization[37], and genome-wide by different array techniques including comparative genome hybridization arrays[37], and SNP arrays[38,39]. In addition, whole genome sequencing was employed to obtain high resolution maps of copy numbers[40].

1.1.2.4 Epigenetic alterations

Conrad Waddington was the first to introduce the term epigenetics in 1942. It can be defined by the study of “changes in gene function that are heritable and that do not entail a change in DNA sequence”[41]. This means it encompasses the study of modifications ‘on top of’ (epi-) DNA, *i.e.* everything that is not part of the DNA bases cytosine (C), guanine (G), thymine (T) or adenine (A). This includes chemical modifications of bases, *e.g.* methylation. Additionally histones which are the protein scaffold for DNA can be altered at their histone tails, *e.g.* by methylation or acetylation. And non-coding transcripts, *e.g.* micro RNAs (miRNA) or long non-coding RNA (lncRNAs) can regulate expression of their target genes.

The existence of DNA methylation in human has been first proposed in 1975[42,43]. The most prominent DNA modification is the methylation of the 5' carbon atom of cytosine. This modification usually only takes place when the C is followed by a G (CpG). This modification is fragile and spontaneous deamination of 5-methylcytosine is common and when it is not repaired this leads to a C to T mutation[44]. In fact, this is the reason why CpG sites are statistically underrepresented in the genome, as evolution has filtered out unimportant CpG sites. Only important CpG sites are maintained. This led to the development of so-called CpG islands, where CpG sites are much more common than expected. They are usually defined by a defined minimum length (usually 200 base pairs (bp)), a GC content (usually 50%) and a distinct ratio of CpG sites to the number of Cs and Gs in the sequence (usually greater than 0.6)[45]. CpG islands mainly occur in gene-rich regions and are often found in promoter regions. CpG islands in promoter regions are often unmethylated in actively transcribed genes[46,47] to create an open chromatin to make room for the transcription machinery[48], whereas they are methylated at inactive DNA regions and genes. In contrast, intragenic DNA methylation stabilizes the DNA transcription machinery and prevents the formation of undesirably transcription start sites within a gene, i.e. intragenic methylation is associated with higher transcription rates[49]. Additionally DNA methylation also controls activity of distal regulatory sites, including (i) enhancers, which are non-coding regions which activate transcription of target genes and (ii) insulators, which are non-coding regions which inactivate transcription of target genes[50,51]. The location of DNA methylation has been shown to be highly tissue-specific, and therefore tissues of unknown origin can be predicted based on their DNA methylome[52-55]. It has been shown that cancer methylomes are highly aberrant[56,57] and that the differential methylation of DNA and histones contributes to differential expression of target genes, e.g. promoter methylation of *CDKN2A* downregulates this gene in PDAC[58]. In fact, tumor suppressor genes are frequently targeted by promoter hypermethylation in cancer[59-62].

Another layer of epigenetic control is the chemical modification of histone tails. A wide variety of different modifications have been identified and for some of these modifications functional consequences of nearby genes have been elucidated. For example the triple methylation of the lysine 4 in histone 3 (H3K4me3), H3K4me1, and acetylation of H3K27 (H3K27ac) are marks of active regions with open chromatin, whereas H3K27me3, and H3K9me3 are marks of inactive regions with closed chromatin[63]. The histone marks are not mutually exclusive and display a complicated interplay that we only begin to understand. To dissect the functional role of these modifications, a Hidden Markov Model (HMM) was able to stably predict functional roles of

combinations of chromatin modifications[64]. The knowledge of these modifications has thus been very useful in identifying regulatory elements in the genome. Whereas genes can be predicted based on their reading frame; promoter regions, enhancers, and insulators are more difficult to anticipate. The HMM has been successful in identifying dozens of functional regions in the genome based on chromatin modifications[64] in a wide variety of normal and tumor tissues[65]. Like aberrant DNA methylation, aberrations in chromatin structure are frequent events in cancer[66-68].

Thirdly, expression of genes might be regulated by non-coding transcripts, *e.g.* miRNA or lncRNA. These small RNAs, usually ~22 bp long, interfere with the expression level of their target genes, by complementary binding of miRNA to the 3' untranslated regions of mRNAs[69,70]. Subsequently the targeted mRNAs are degraded.. In contrast to that lncRNAs are usually >200 bp long and interfere with transcription in multiple ways, *e.g.* by altering the chromatin structure or DNA methylation of a target locus[71]. MiRNA and lncRNA deregulation has been described in all investigated cancer entities[72-77].

In summary, epigenetic aberrations are a common feature in cancer. Although, studies in the last decades revealed many aspects of epigenetics and cancer, much remains to be investigated. For example, genetics and epigenetics are closely interlinked[67], e.g. *IDH1*-mutated cancer harbor specific epigenomes[78]. However, it remains unknown how this interplay is exactly regulated and whether the genetic or epigenetic event is the initial one.

1.1.2.5 DNA repair machinery defects

Besides regulators of the cell cycle addressed in 1.1.1, cells are protected by a DNA repair machinery, which jumps in when DNA is altered. Exogenous agents, *e.g.* UV light, X-rays and alkylating agents can cause a number of DNA lesions[79-83]. When an error occurs, DNA is generally repaired by one of the following pathways: mismatch repair (MMR), base excision repair (BER), nucleotide excision repair (NER), homologous recombination (HR), or non-homologous end joining (NHEJ)[7].

MMR, BER, and NER all repair aberrations on one strand of the DNA. DNA polymerases are highly efficient enzymes and, also owing to their proofreading ability, work with extremely low error rates of one error per 10^7 bp[7]. This rest error rate is fixed by the MMR machinery which follows polymerases during replication. Tumors deficient in MMR are termed microsatellite-instable (MSI) tumors and harbor many amplifications and deletions in repetitive sequences, *e.g.* microsatellites[84]. When single bases have been chemically altered by endogenous or

exogenous agents, BER or NER machinery repairs these lesions. BER mainly occurs after endogenous alterations resulting in minor lesions of the DNA[85]. As the name suggests, BER removes only one nucleotide by subsequent steps of a DNA glycosylase, apurinic/apyrimidinic (AP) endonucleases and an AP lyase. The correct nucleotide is replaced by polymerase β and DNA ligases[85]. In contrast, NER occurs mainly after errors due to exogenous agents, *e.g.* UV light, and repairs lesions which result in bulky lesions of the DNA. It removes a whole stretch of the DNA strand (approximately 25-30 bp long) containing the aberrant base by a large complex of different enzymes and this gap is then filled in by specific polymerases and closed by a ligase[86].

Double-strand breaks (DSB) are repaired by either HR or NHEJ[87]. During late S phase and G2 phase when a sister chromatid is available, HR is carried out[88]. The DSB has to be recognized by ATM, and the histone H2AX is subsequently phosphorylated by ATM or the DNA-dependent protein kinase PRKDC[89]. The resulting γ-H2AX enables an open chromatin state for the repair process[90,91]. An exonuclease then removes the stretches of DNA strands around the break, resulting in free 3' single strands (ss) of both strands. BRCA1 and BRCA2 serve as scaffold proteins to recruit DNA repair machinery to DSB, *e.g.* BRCA2 has eight binding pockets for RAD51, which coats these ssDNAs[92,93]. These ss are then mapped to the respective regions of the unwound sister chromatid and repaired by a DNA polymerase and ligase[94]. When the sister chromatid is not available after a DSB, NHEJ is employed. As the name suggests, the ends of the DSB are repaired without a template resulting in a high error rate of this repair process. There the region around the DSB is resected by exonucleases, the 3' ss are then brought into proximity and ligated, usually leading to small deletions[95].

Until today it is not clear how a cell repairs epigenetic alterations. Epigenetic maintenance has been examined in the context of replication and during DNA damage repair. During replication, DNMT1 ensures that DNA methylation is copied onto the new DNA[96], to make sure daughter cells harbor the same methylome. DNA damage repair processes are accompanied by changes in DNA methylation and histone marks[97-101], and sometimes these loci are permanently altered[102]. Further studies are needed to evaluate epigenetic changes after DNA repair. It is currently unknown how faulty epigenetic modifications are repaired, and more studies are needed to investigate which epigenetic repair mechanisms exist.

Defects in the DNA machinery give rise to DNA alterations in a fast paced manner. This enables a cell to give rise to many clones and increase the chance of malignant transformation.

Therefore it is not surprising, that genetic and epigenetic defects in the DNA repair machinery have been frequently described in sporadic and hereditary cancers[103-107]. For example xeroderma pigmentosum, a hereditary skin cancer syndrome, is caused by germline defects in *XPA*, -*B*, -*C*, -*D*, -*E*, -*F*, -*G*, and -*V*, mostly members of the NER complex[108]. *BRCA1* or *BRCA2* are often mutated in sporadic and familial cancer cases, the most prominent one being breast cancer[109,110].

In summary, defects in repair, cell cycle control, and epigenetic control are of main interest for tumor development and should play a role in pancreatic acinar cell carcinomas (ACC), which are the focus of this work.

1.2 Pancreas

The pancreas consists of two main compartments with vitally important functions for the human body. The endocrine compartment resides in Langerhans islands consisting of different endocrine cell types, mainly regulating glucose homeostasis by releasing hormones into the blood stream. There are five different endocrine cells types in the pancreas, namely α-cells, which produce glucagon, β-cells which produce insulin, δ-cells which produce somatostatin, PP-cells which produce pancreatic polypeptide, and ε-cells which produce ghrelin. The exocrine compartment consists of acinar cells producing pancreatic juice containing enzymes, while duct cells build up the ducts for delivering the digestive enzymes into the duodenum (Figure 4).

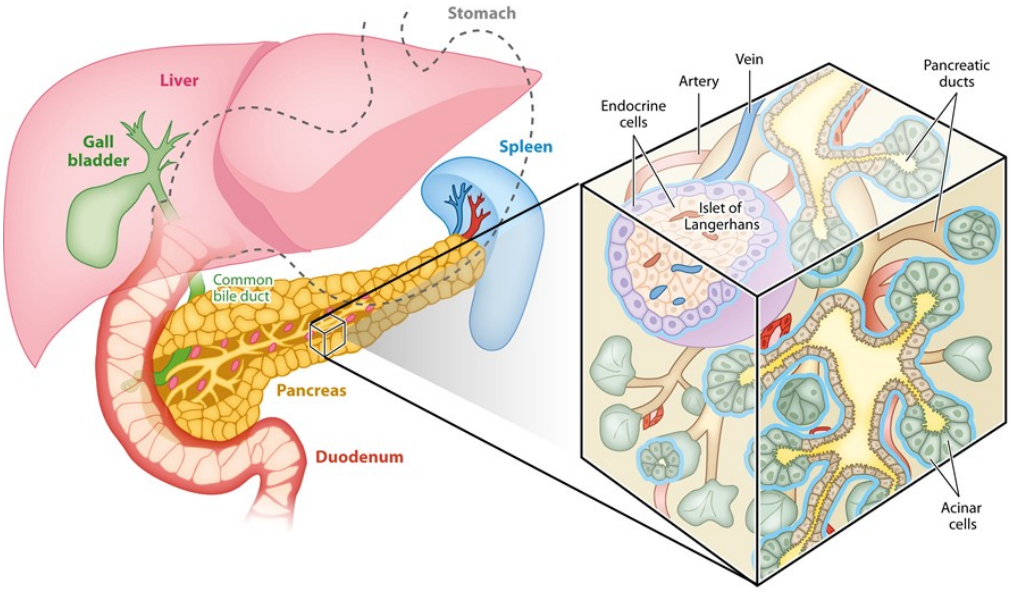


Figure 4 Anatomy of the pancreas. Figure adapted from Shih, et al. [111].

1.2.1 Embryonic development of the pancreas

The pancreas arises from a dorsal and a ventral bud from the distal foregut endoderm. Exocrine and endocrine pancreatic cells all develop from the same pancreatic progenitor cell (Figure 5). During development pancreatic buds emerge and the progenitor cells differentiate to tip cells at the protrusion and trunk cells at the inside of the pancreatic buds. Tip cells later give rise to acinar cells, while trunk cells give rise to duct cells and endocrine precursor cells. The latter are differentiating into the different mature endocrine cells (reviewed by Shih, et al. [111] and Jennings, et al. [112]). These differentiation steps are highly determined by a cascade of transcription factors[113]. The TF PDX1, PTF1A, and SOX9 are specifically active in the pancreatic progenitor cells, NKX6-1/2 define the trunk cell lineage while PTF1A defines the tip

cell lineage[114-116]. Notch signaling is crucial for two steps: It decides whether a pancreatic progenitor cell gives rise to a tip or a trunk cell and then whether a trunk cell gives rise to endocrine precursor or duct cells[117-119]. PTF1A, NR5A2, and MIST1 are important TF defining

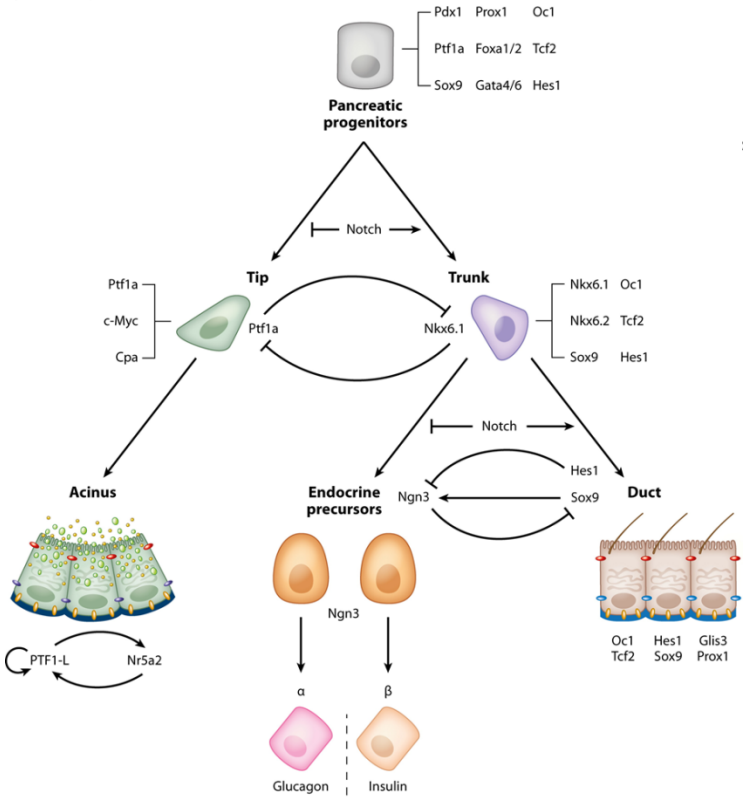


Figure 5 Generation of pancreatic cell types during embryonic development. Figure adapted from Shih, et al. [111].

the acinar cell fate, while duct cells are defined by expression of SOX9 and endocrine cells by the expression of PAX6 and NEUROG3[116,120,121]. Endocrine cells are then differentiated into 5 different cell types. The underlying mechanism of this differentiation process is unknown, however it seems to be time-dependent as endocrine cells are created in a step-wise manner from α- to β- to δ- and finally to PP-cells[122].

1.2.2 Pancreatic plasticity

Upon tissue damage, the pancreas can, although at slow rates, replace damaged tissue. The most straightforward hypothesis to explain this tissue regeneration was that pancreatic pluripotent stem cells exist[123]. The search for pancreatic stem cells however has been unsuccessful. All pancreatic cell types further have a different RNA expression profile and

specific histone modification signature[124,125], suggesting that these cell types are regulated on the epigenetic level.

Studies are emerging that show the pancreas as a highly plastic tissue, meaning that upon tissue damage, a cell type can dedifferentiate into a more multipotent, proliferating cell and replace cell types of its own kind and other pancreatic cell types (Figure 6). Studies suggest

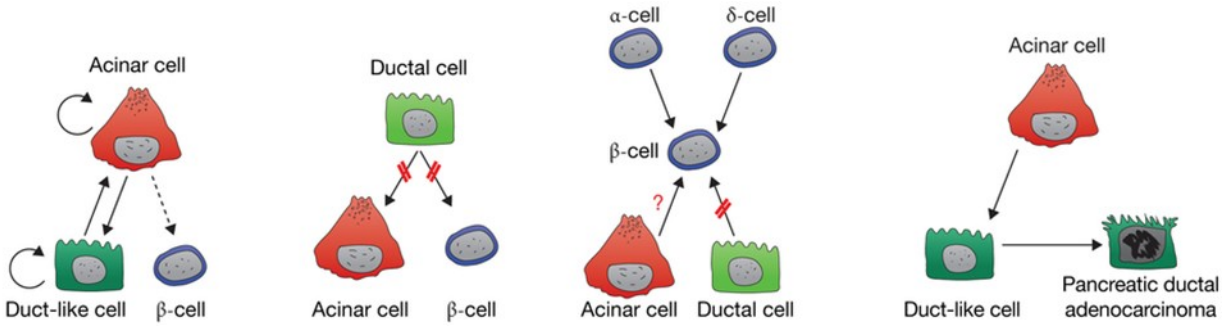


Figure 6 Pancreatic tissue is highly plastic upon tissue damage. Figure adapted from Kopp, et al. [126].

different scenarios are possible. Acinar cells seem to be able to dedifferentiate into multipotent progenitor cells[127] which can – similar to pancreatic development – give rise to acinar cells and to another precursor cell which gives rise to new duct and endocrine cells[128-130]. Differentiated endocrine cells can transdifferentiate into other endocrine cells by tissue manipulation, *e.g.* α-cells can replace β-cells[131,132] or *vice versa*[133-135], and δ-cells can be turned into β-cells[132] and *vice versa*[136]. Recently, a study revealed that it might not be all acinar cells that can be dedifferentiated. Indeed it was shown that only a subpopulation of mononuclear acinar cells in the pancreas proliferates in homeostatic conditions and that these cells are often marked with STMN1. Upon tissue damage additional acinar cells acquire the possibility to proliferate[137]. This high plasticity seems to be relevant for tumor initiation in the pancreas. For the development of PDAC (see 1.2.3.2), the most commonly agreed hypothesis is that these tumors arise from acinar cells which dedifferentiate into ductal-like structures, a process come to known as acinar-to-ductal metaplasia (ADM)[138-142]. If there is no further stress on these cells, ADM is reversible. However another hit, *e.g.* a *KRAS* mutation on top of ADM leads to precursor lesions of PDAC[141,143].

1.2.3 Pancreatic cancer

Pancreatic cancer is an aggressive cancer with one of the highest mortality rates[144] and an incidence especially high in the western world (Figure 7). As symptoms develop only late at an

advanced tumor state, these cancers have often spread to other organs. There are several pancreatic tumor types which differ in histopathological, molecular, and clinical behavior. Here, the rare pancreatic ACC which are the topic of this thesis are introduced and the two most common pancreatic cancers PDAC and pancreatic neuroendocrine tumors (PNET) are presented.

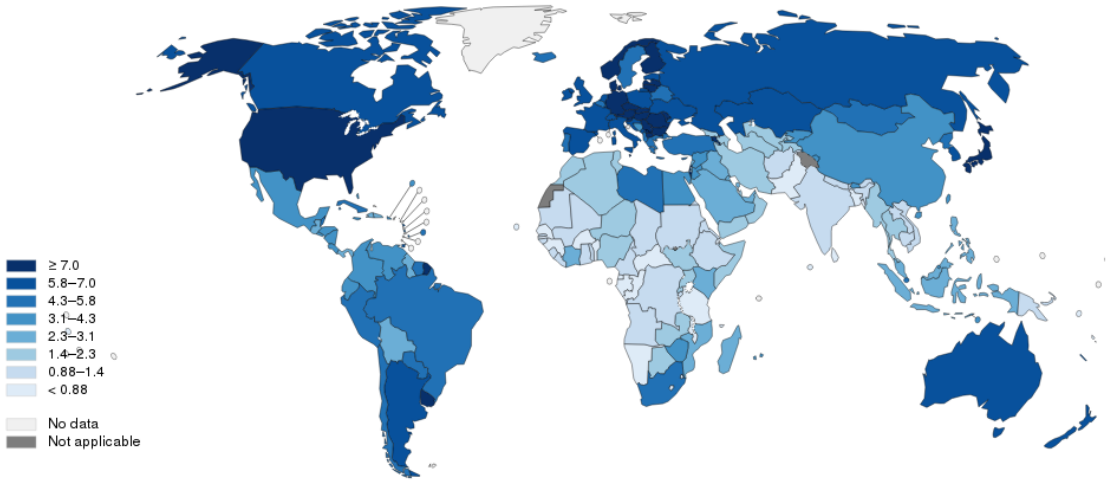


Figure 7 World-wide age-standardized rates of incidences of pancreatic cancers, estimated by the Globocan project[145]. Scale referring to incidences per 100,000 people.

1.2.3.1 Pancreatic acinar cell carcinoma

ACC is an aggressive cancer accounting for only about one percent of all pancreatic cancers. Diagnosis usually comes late due to late arising and unspecific clinical symptoms, *e.g.* abdominal pain, nausea, and weight loss. This results in large tumor sizes with an average of 8 cm[146-149] and metastases in about 50% of cases[146,148,150-152] consequently leading to a poor overall survival of around 45 months and a 5 year survival between 20 and 70% of patients[146,149,152-155]. Apart from surgical removal which is not an option for a majority of ACC patients with advances disease[153,156], standardized recommended treatments are lacking and patients' treatment plans are consequently very heterogeneous[151,155]. ACC affects predominantly males (about two third)[146,147] and older patients with a mean age around 65[146,147,149,153]. Occasionally it can occur at earlier ages, including childhood[152-154,157-160]. Tumor cells resemble acinar cells morphologically containing eosinophilic granules and prominent nucleoli[146,161-164]; however it is unknown whether ACC arise from acinar cells or other pancreatic cell types. Tumors are usually well circumscribed, sometimes by capsules[146]. The architecture of tumors varies from densely to acinar, glandular and/or trabecular growth patterns[146,149].

Necrotic cells are found in roughly 50% of tumors, while vascular infiltration and perineural invasion occurs in about 50% and 30% respectively[146]. Expression of digestive enzymes can be detected in most tumors including trypsin (95% of cases), carboxyl ester lipase (90%), lipase (26%), and amylase (9%)[146,149]. Patients with tumors highly expressing lipase usually show a generalized subcutaneous fat necrosis and polyarthralgia when this enzyme is secreted into the blood stream (about 5% of cases)[151]. Some ACC (~20%) additionally express neuroendocrine markers such as synaptophysin and chromogranin A[146,147,165] in at least 25% of tumor cells, thus showing similarities to PNET (refer to 1.2.3.3). However, as clinical presentations are similar to ACC these mixed acinar-neuroendocrine carcinomas (MACNEC) are grouped together with pure ACC[146]. The expression of the aforementioned enzymes is used to distinct ACC from other pancreatic cancers. Atypical acinar cell lesions have been described in the literature[166-170], however there is no data supporting that these lesions actually progress to ACC.

Not only studies on the clinicopathological side but investigations on the genome and epigenome of pancreatic ACC are lacking, primarily due to the rarity of this tumor. Consequently there is a lack of larger tissue collections and less interest for research in pharma industry but also in funding agencies. Most published studies are performed with small sample sizes and are focusing on a few selected genes. However, these studies show that molecular alterations known in PDAC are usually not affected in ACC[171-178]. For example, TP53 has long thought to play no role in ACC[171,172,176]. However, recently a study revealed that 13% of primary tumors and 31% of metastases harbor *TP53* mutations, while 53% of primary tumors and 50% of metastases harbor *TP53* deletions[179]. In addition, *BRAF* fusions have recently been reported in two studies, revealing that 11 out of 47 tumors harbor fusions[180,181]. A few mutations in *BRAF* have also been reported[180,182]. One of the more closely examined pathways in ACC is the WNT pathway (Figure 2). “Mutations in *APC* have been reported in about 10-20% of cases[171,173,182]. Activation of the CTNNB1 protein has been reported in about 12-15%[146,171]. One study investigated alterations in *APC*”[1] and found 7% of cases have mutation, 48% deletions, and 56% are hypermethylated[177]. “However, data on APC protein expression are still lacking.”[1]

Further epigenetic studies are scarce, with one study investigating promoter hypermethylation in a panel of tumor suppressor genes[183]. This revealed many and recurrent hypermethylation; however with the limit of the applied method methylation-specific polymerase chain reaction (MSP) which is not quantitative and not very specific. However, a mouse model revealed that the incidence of ACC arising form *Tp53*$^{-/-}$ and *Apc*$^{+/-}$ mice[184] can be reduced by an additional *Dnmt1* knockout, suggesting that DNA methylation plays a role in the development of this

tumor[185]. In addition, another study of four ACC, compared microRNA profiles of ACC with that of PNET (refer to 1.2.3.3) and found, that a set of ten microRNAs can differentiate between the tumor entities. In summary, epigenetic data on ACC are limited but suggest that the epigenome is an important component in ACC.

Recently, two genome-wide studies have shed some light on ACC in a gene-unbiased way with exome sequencing. They found a high mutational load in ACC with the most common mutation in *SMAD4* occurring in 25% of tumors[182,186]. Thus, no shared mutation was found in the majority of tumors, suggesting other mechanism of tumor development. A classical and array-based comparative genomic hybridization analysis revealed many chromosomal imbalances in ACC[147,176,187]. However, resolution of this array is limited thus making it difficult to pinpoint distinct important aberrations[147]. Microsatellite instability and DNA mismatch abnormalities have been reported in a small subset of tumors[171,180,182,188]. These studies suggest CNA might play a role in this tumor entity.

1.2.3.2 Pancreatic ductal adenocarcinoma

PDAC is the most common (~90%) and with a 5 year survival of 2% the most aggressive of all pancreatic cancers[189,190]. Therefore, it is one of the most lethal tumors in general. This is caused by the lack of symptoms until the tumors reach quite a large size, and due to the high infiltrating nature of PDAC, most tumors have metastasized when they are diagnosed so that surgery is not possible for all cases[190,191]. In unresectable tumors, treatment of patients usually includes chemotherapy with the standard chemotherapeutic regimen gemcitabine in monotherapy or in combination with paclitaxel[192]. However, resistance to therapies is the rule[193]. Developing efficient therapies for targeting the frequent *KRAS* mutations so far have failed in the clinics (reviewed in Baines, et al. [194] and Stephen, et al. [195]) and of the targeted therapies tested including VEGF-, IGFR-, and EGFR-inhibitors only the latter one showed minor improvements in survival[193,196,197].

PDAC have been shown to arise through a sequence of precursor lesions termed pancreatic intraepithelial neoplasia (PanIN). PanIN-1A lesions are characterized by duct cell elongation, PanIN-1B by formation of papillae. PanIN-2 lesions then harbor aberrations in nuclear structure, while PanIN-3 show severe nuclear dysplasia and finally develop to invasive growing adenocarcinomas. The molecular events during these steps are commonly thought to be *KRAS* mutations that are occurring before or in PanIN-1, followed by *CDKN2A* mutations which are accumulating through PanIN-1B to PanIN3 lesions. *TP53, SMAD4,* and *BRCA2* alterations are

obtained in late PanIN2 and PanIN3 (Figure 8)[16,17,58,198-200]. While this sequence of events is supported by numerous studies, one recent study challenged this view and suggested that the molecular events occur all at the same time[201]. As this would highly change the way of searching for drugs and methods for early detection of PDAC, more studies are needed to evaluate this alternative hypothesis.

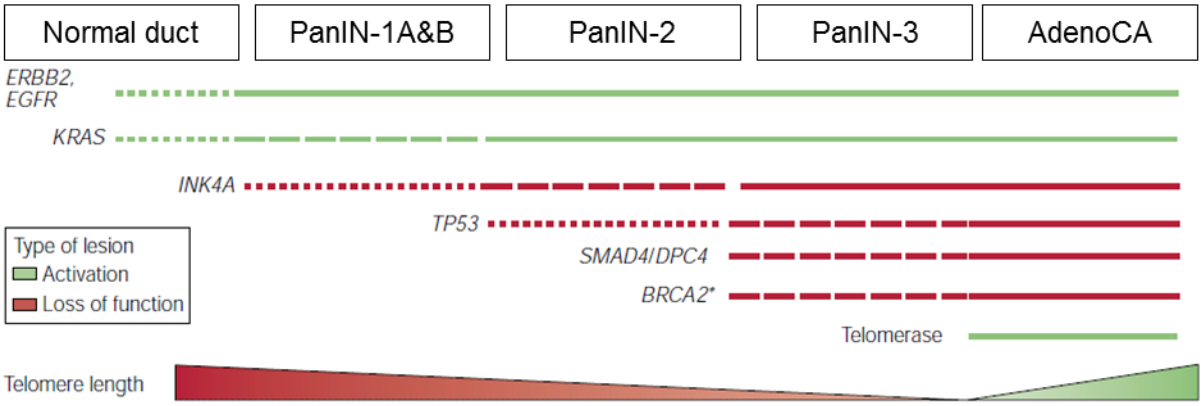


Figure 8 PDAC carcinogenesis and molecular alterations. AdenoCA = Adenocarcinoma. Figure adapted from Bardeesy and DePinho [199].

The cell of origin has been highly debated in PDAC. Due to the morphological similarities, the expression of ductal markers mucin and cytokeratins[190], and the fact that precursor lesions resemble ductal cells, it has been assumed that PDAC arise from ductal cells. However, it has been difficult to prove this[202]. In fact, studies are accumulating which show that PDAC might arise from acinar cells through a process called acinar ductal metaplasia (refer to Figure 8)[138,140,141,143,203-206]. There, acinar cells first dedifferentiate to a multipotent progenitor cell and can redifferentiate into acinar cells. However, when a harming stimulus continues these cells can dedifferentiate into ductal-like cells. If another oncogenic event occurs, *e.g. KRAS* mutations, these ductal-like cells give rise to PanIN lesions and ultimately lead to cancer. There is still some debate on whether all acinar cells are capable of ADM. While some support the hypothesis that only centroacinar cells – cells which are at the border of acinar cells and duct cells – have the capability to do so there is evidence that only a fraction of acinar cells has the capability to self-replicate during tissue homeostasis and stress stimuli[137].

1.2.3.3 Pancreatic neuroendocrine tumors

PNET are the second most common pancreatic tumors (~10%)[190]. In 10-30% of PNET, so-called functional PNET, expression of endocrine enzymes are secreted by the tumor, and can include the hormones glucagon, insulin, somatostatin, gastrin, vasoactive intestinal polypeptide

(VIP), growth hormone-releasing factor (GHRH) and adrenocorticotrophic hormone (ACTH) and are accompanied by massive impairments of the systemic impact of these hormones on the body. Based on the expression of these hormones PNET are classified into glucagomas insulinomas, somatostatinoma, gastrinoma, VIPoma, GHRH- and ACTH- secreting PNET. The remaining 70-90% are non-functional PNET which do not express hormones or secrete them into the blood stream. Survival of PNET patients is better than those of ACC and PDAC patients, however still quite low with a five year overall survival of 33%[189]. However, survival and metastatic potential highly depend on the subtype of PNET, *e.g.* insulinomas are much less likely to spread (10%) in contrast to somatostatinoma (50-70%) and VIPomas (40-70%)[207]. Standard recommended therapies highly depend on the subtype and symptoms and include surgery, chemotherapy, and radiotherapy[208]. Approved targeted therapies include mTOR inhibitors and the tyrosine kinase inhibitor sunitinib[207,209-211] (reviewed in Phan and Dave [212] and Valle, et al. [213]).

There is quite some data underlining the molecular changes in PNET, however not as detailed as in PDAC. A recent study integrated for the first time different genome-wide data sets to unravel the molecular causes of PNET[214]. About 17% of these tumors occur due to hereditary aberrations in a number of genes, which is in contrast to ACC and PDAC. Germline mutations in *MEN1* in patients with Multiple Endocrine Neoplasia Type 1 and *VHL* mutations account for the most common heritable forms of PNET[215-217]. Recently germline mutations in *MUTYH*, *CHEK2* and *BRCA2* were identified[214]. Somatic point mutations are frequent events, namely aberrations occur in *MEN1* (40%), *DAXX* (25%), *ATRX* (18%), *TSC2* (9%) and *PTEN* (7%)[209,211,217-219]. Further, PNET harbor instable genomes[220-222] with reported imbalances in 6p22 and 17p13.1, as accessed by an array comparative genomic hybridization (aCGH study)[223] and several regions containing tumor suppressor genes (1q25.1, 3p21, 18q13.3, 9q21.3, and 11q13.1)[214].

The cell of origin in PNET is unknown. Deletion of a tumor suppressor gene in endocrine α-cells however produces insulinomas und glucagomas, suggesting that endocrine α-cells might be the cell of origin[224], however there is supporting data that these tumors arise from the exocrine system[225].

In summary, molecular aberrations and tumor development in the pancreatic cancers PDAC, and PNET are well known. This is in contrast to ACC, were comprehensive data sets are lacking and therefore this tumor entity needs profound additional investigation.

2 Aims

This study aimed to investigate the molecular aberrations in ACC that contribute to tumorigenesis and to identify potential new treatment options for patients. Two independent ACC cohorts were used to generate whole genome and epigenome-wide data sets. Throughout the thesis, comparison to the pancreatic ductal adenocarcinoma and pancreatic neuroendocrine tumors were made when possible. Further it was delineated whether MACNEC exhibit the same molecular aberrations as pure ACC or whether they represent a separate tumor entity. In addition, ACC metastases were examined to dissect whether primary tumors harbor different aberrations compared to metastases.

Molecular aberrations in ACC were scrutinized on multiple levels.

First, whole exome sequencing aimed to elaborate the mutational landscape of ACC, by dissecting the most commonly mutated genes and to investigate potential mutational signatures, thus discovering potential risk factors for ACC.

Second, DNA methylation was examined by 450K arrays with the aim of identifying epigenetic alterations in this tumor. Differential methylation of promoters and gene bodies was called for the purpose of investigating deregulated genes, and gene enrichment analyses revealed affected pathways. Differential methylation of the main pathway was studied in more detail. DNA methylation data also served as a tool to discover the cell of origin of ACC.

Third, chromosomal imbalances of ACC were determined based on the 450K array data, common CNA were identified and genes mapping to CNA were followed up.

Finally, the data sets were combined in an integrated analysis to identify recurrent aberrations. These were overlapped with previously published cancer gene lists. It was aimed to confirm strong differential candidate regions with independent methods and by immunohistochemistry. To identify drugs that exploit these molecular aberrations, approved targeted therapies and drugs currently being tested in clinical trials were screened. Therefore, it was aimed to propose new therapeutic interventions for ACC patients on the basis of the newly identified molecular aberrations.

3 Materials and Methods

3.1 Materials

3.1.1 Instruments

Instruments	Manufacturer
Airclean 600 PCR workstation	Starlab International, Hamburg, Germany
Bioanalyzer 2100	Agilent Technologies, Santa Clara, CA, USA
Cell culture incubator	Sanyo, Taipei, Taiwan
Centrifuge 5415R	Eppendorf, Hamburg, Germany
Centrifuge 5424	Eppendorf, Hamburg, Germany
Centrifuge 5810 R	Eppendorf, Hamburg, Germany
Concentrator 5301	Eppendorf, Hamburg, Germany
DM IL light microscope	Leica, Wetzlar, Germany
Duomax 1030 mixing platform	Heidolph, Schwabach, Germany
Epson Perfection V700 photo scanner	Epson, Suwa, Japan
Gel electrophoresis chamber model 192	Bio-Rad, Hercules, CA, USA
Geneamp PCR system 9700	Applied Biosystems, Foster City, CA, USA
Herasafe laminar Flow workbench	Thermo Fisher Scientific, Waltham, MA, USA
Hypercassette Amersham	GE Healthcare, Chicago, IL, United States
iScan array system	Illumina, San Diego, CA, USA
Leitz Diavert light microscope	Leica, Wetzlar, Germany
LightCycler 480 RT-qPCR device	Roche Diagnostics, Mannheim, Germany
MassARRAY Compact System	Agena Bioscience, San Diego, CA, USA
MassARRAY Nanodispensor	Agena Bioscience, San Diego, CA, USA
Mastercycler epgradient PCR cycler	Eppendorf, Hamburg, Germany
Mastercycler pro 384 PCR cycler	Eppendorf, Hamburg, Germany
Mettler AT 261 Deltarange microgram scale	Mettler Toledo, Zwingenberg, Germany
Microwave oven	Bosch, Stuttgart, Germany

Nanodrop spectrophotometer ND1000	Peqlab, Erlangen, Germany
Novex Mini-Cell electrophoresis chamber	Thermo Fisher Scientific, Waltham, MA, USA
PowerPac Basic electrophoresis power supply	Bio-Rad, Hercules, CA, USA
QTRAP 6500 mass spectrometer	AB SCIEX, Darmstadt, Germany
Qubit® 2.0 Fluorometer	Invitrogen, Life Technologies, Darmstadt, Germany
SPECTRAmax M5 microplate reader	Molecular Devices, Sunnyvale, CA, USA
Thermomixer Comfort	Eppendorf, Hamburg, Germany
Tissue lyser	Qiagen, Hilden, Germany
Ventana BenchMark Ultra	Ventana Medical Systems, Oro Valley, AZ, USA

3.1.2 Laboratory consumables and reagents

Consumables/Reagents	**Manufacturer**
0.05% Trypsin-EDTA	Thermo Fisher Scientific, Waltham, MA, USA
100 bp DNA ladder	Fermentas, Life Technologies, Darmstadt, Germany
10x PCR Buffer	Qiagen, Hilden, Germany
2-mercaptoethanol	Sigma Aldrich, St. Louis, MO, USA
5-aza-2'-Deoxycytidine	Sigma Aldrich, St. Louis, MO, USA
5x first strand synthesis buffer	Invitrogen, Life Technologies, Darmstadt, Germany
8-well single cap PCR strips	Biozym, Hessisch Oldendorf, Germany
96 and 384 well PCR plates	Steinbrenner, Wiesenbach, Germany
anti-APC (HPA013349)	Sigma Aldrich, St. Louis, MO, USA
anti-Arid1A (HPA005456)	Sigma Aldrich, St. Louis, MO, USA
anti-CDKN2A (P16) (551153)	BD Bioscience, San Jose, CA, USA
anti-Histone H1 (SAB4501366)	Sigma Aldrich, St. Louis, MO, USA

anti-ID3 (SAB1412646)	Sigma Aldrich, St. Louis, MO, USA
anti-Jak1 (SAB4300393)	Sigma Aldrich, St. Louis, MO, USA
anti-PCDHG (N159/5)	Antibodies Incorporated, Davis, CA, USA
anti-SOX2 (ab97959)	Abcam, Cambridge, United Kingdom
Calcein-AM	Sigma Aldrich, St. Louis, MO, USA
Cell culture dishes (10cm, 15 cm)	TPP, Trasadingen, Switzerland
Cell culture multiwell plates (6, 24, 96, 384 well)	Greiner Bio-One, Frickenhausen, Germany
CELLSTAR cell culture flasks (T25, T75)	Greiner Bio-One, Frickenhausen, Germany
Complete Protease inhibitor cocktail, EDTA free	Roche Diagnostics, Mannheim, Germany
Dharmafect siRNA transfection reagent	Dharmacon, Lafayette, CO, USA
Dimethylsulfoxide (DMSO)	Merck, Darmstadt, Germany
DMEM medium 1X	Thermo Fisher Scientific, Waltham, MA, USA
DNA oligonucleotides 100µM	Sigma Aldrich, St. Louis, MO, USA
DNase 1	Fermentas, Life Technologies, Darmstadt, Germany
Dneasy blood & tissue kit	Qiagen, Hilden, Germany
dNTPs (dATP, dTTP, dCTP, dGTP)	Fermentas, Life Technologies, Darmstadt, Germany
Dulbecco's modified Eagle Medium (DMEM)	Gibco, LifeTechnologies, Darmstadt, Germany
Ethidiumbromide	Sigma Aldrich, St. Louis, MO, USA
EZ DNA Methylation kit	Zymo Research Europe GmbH, Freiburg, Germany
EZ-96 DNA Methylation kit	Zymo Research Europe GmbH, Freiburg, Germany
falcon tubes 15ml, 50ml	Greiner Bio-One, Frickenhausen, Germany
Fetal bovine serum (FCS)	Biochrome, Berlin, Germany

Filter tips and normal tips for pipettes (10µl, 20µl, 200µl, 1000µl)	Biozym, Hessisch Oldendorf, Germany
Gene Ruler DNA ladder mix	Fermentas, Life Technologies, Darmstadt, Germany
Goat anti-mouse antibody	Santa Cruz Biotechnology, Dallas, TX, USA
Goat anti-rabbit antibody	Santa Cruz Biotechnology, Dallas, TX, USA
High sensitivity DNA kit for bioanalyzer	Agilent Technologies
HotstarTaq DNA polymerase kit	Qiagen, Hilden, Germany
HPLC-grade ethanol (abs.) and methanol	Sigma Aldrich, St. Louis, MO, USA
Hyperfilm ECL chemiluminescence film	GE Healthcare, Chicago, IL, United States
Infinium HD DNA Restoration kit	Illumina, San Diego, CA, USA
Infinium HD FFPE QC kit	Illumina, San Diego, CA, USA
Infinium HumanMethylation450 Bead Chip Array (Illumina 450K)	Illumina, San Diego, CA, USA
Invisorb Genomic DNA kit II	Stratec, Birkenfeld, Germany
Invitrolon PVDF membrane	Invitrogen, Life Technologies, Darmstadt, Germany
Loading Dye (6x)	Fermentas, Life Technologies, Darmstadt, Germany
M.SssI CpG methyltransferase	Thermo Fisher Scientific, Waltham, MA, USA
MassCLEAVE T7 kit (T Cleavage)	Sequenom, San Diego, CA, USA
MEM Non-Essential Amino Acids Solution (100X)	Thermo Fisher Scientific, Waltham, MA, USA
MS-grade trypsin	Sigma Aldrich, St. Louis, MO, USA
Nonidet P-40 (NP-40)	Sigma Aldrich, St. Louis, MO, USA
Nupage 10% Bis-Tris gel	Thermo Fisher Scientific, Waltham, MA, USA
Nupage antioxidant	Thermo Fisher Scientific, Waltham, MA, USA
Nupage MOPS SDS running buffer	Thermo Fisher Scientific, Waltham, MA, USA
Nupage transfer buffer	Thermo Fisher Scientific, Waltham, MA, USA

Phosphate buffered saline	Thermo Fisher Scientific, Waltham, MA, USA
Qubit dsDNA HS assay kit	Invitrogen, Life Technologies, Darmstadt, Germany
Reaction tubes (1.5ml, 2ml, 5ml)	Eppendorf AG, Hamburg, Germany
Shrimp alkaline phosphatase	Sequenom, San Diego, CA, USA
siGenome siRNA pools	Dharmacon, Lafayette, CO, USA
Sodium dodecyl sulfate (SDS)	Carl Roth, Karlsruhe, Germany
Sterile serological pipettes (5ml, 10ml, 25ml)	BD Bioscience, San Jose, CA, USA
SuperScript III Reverse Transcriptase	Invitrogen, Life Technologies, Darmstadt, Germany
SureSelect Automated Library Prep and Capture System SureSelectXT Automated Target Enrichment for Illumina Paired-End Multiplexed Sequencing	Illumina, San Diego, CA, USA
SureSelectXT Target Enrichment System for Illumina Paired-End Multiplexed Sequencing Library	Illumina, San Diego, CA, USA
TRIzol	Thermo Fisher Scientific, Waltham, MA, USA
Triton X-100	Sigma Aldrich, St. Louis, MO, USA
LightCycler 480 Probes Master	Roche Diagnostics, Mannheim, Germany
Universal ProbeLibrary Probes	Roche Diagnostics, Mannheim, Germany
Venor GeM mycoplasma detection kit	Minerva Biolabs, Berlin, Germany
Western Lightning Plus-ECL chemiluminescence substrate	Perkin-Elmer, Waltham, MA, USA

3.1.3 Patient material

“Two independent ACC cohorts were investigated (for details refer to”[1] Table 10). Cohort I was “a collection of 34 tumors (22 primary tumors and 12 metastases) and 24 normal pancreatic tissues which were cases from the Institute of Pathology, University of Heidelberg, Germany. Primary tumors from this cohort were used as discovery cohort, as matched germline controls were available for sequencing.”[1] Study cohort II “was used as validation cohort and consisted of

39 tumors (38 primary tumors and 1 metastasis) and 10 normal pancreatic tissues which were cases from the Consultation Center for Pancreatic and Endocrine Tumors of the Department of Pathology, Technical University Munich, Germany. Further, 20 and 38 primary tumors, and 3 and 1 metastases from cohort I and II, respectively and 8 normal pancreatic samples were combined on a tissue microarray to facilitate immunohistochemical stainings. Most tissues were available as formalin fixed and paraffin embedded (FFPE) tissues, fresh frozen (FF) material was available for five tumors. In addition, DNA from 17 pancreatic PNET obtained from FFPE tissue was investigated. Tumors were micro-dissected by a pathologist with a tumor purity of >90%. All tissue samples were provided by the tissue bank of the National Center for Tumor Diseases (NCT, Heidelberg, Germany) in accordance with the regulations of the tissue bank and the approval of the ethics committee of Heidelberg University (no. 207/2005). The results shown on PDAC are based upon data generated by"[1] The Cancer Genome Atlas (TCGA) Research Network: http://cancergenome.nih.gov.

3.2 Methods

3.2.1 Laboratory methods

3.2.1.1 DNA extraction

DNA was extracted with the Invisorb DNA kit II according to the manufacturer's instructions. In case of FFPE tissue, paraffin was dissolved by two cycles of one ml Octane. Tissue (FFPE and FF) was washed with pure ethanol and digested with 200 µl Lysis Buffer G and 20 µl Proteinase K at 60°C overnight. 400 µl binding buffer containing silica particles was added, the whole mixture was centrifuged at 10,000 revolutions per minute (rpm), 30 seconds (sec) and supernatant was discarded. Pellets were washed two times with wash buffer and dried at 60°C. Pre-warmed (60°C) elution buffer was added and incubated at 60°C for 3 minutes (min). Samples were centrifuged at 14,000 rpm for 2 min and supernatant was collected. The elution step was carried out twice.

3.2.1.2 Bisulfite conversion

DNA was bisulfite converted with the EZ DNA methylation kit. One µg of DNA and 5 µl of dilution buffer in a total volume of 50 µl were incubated at 37°C for 15 min. Bisulfite conversion reagent was prepared according to the manufacturer's instruction and 100 µl was added per sample. Then, samples were incubated at 50°C for 16 hours and 4°C for at least 10 min. 400 µl binding buffer was added to a Zymo spin column and the sample was added and mixed. After a 30 sec centrifugation step at 12,000 relative centrifugal force (rcf), the column was washed once with 100 µl wash buffer. Desulphonation buffer (200 µl) was added and incubated for 15 min at room temperature (RT) and the column was washed twice with 200 µl wash buffer. Samples were then eluted twice with 25 µl of 1: 1 diluted elution buffer. Bisulfite-converted DNA was stored at -80°C.

3.2.1.3 RNA extraction

RNA was extracted according to the TRIzol Reagent protocol with slight changes to the manufacturer's instructions. Instructions are based on 1 ml TRIzol starting volume, but volumes were adjusted in case RNA was extracted from smaller cell numbers. Briefly, cells were incubated with 1 ml TRIzol for 5 min at RT. Samples were then either frozen at -80°C or immediately processed by adding 0.2 ml chloroform, shaken and centrifuged at 12,000 rcf for 15 min at 4°C. The upper phase was removed, placed in a new tube and precipitated with 0.5 ml 100% ethanol and 10 µg glycogen (10 min at RT, followed by 10 min 12,000 rcf at 4°C). Pellet

was washed twice with 75% ethanol; pellet was dried at RT and dissolved in pure water at 55°C, 12 min. RNA was stored at -80°C.

3.2.1.4 cDNA synthesis

1 µg of RNA in a total volume of 8 µl was DNAse digested (1µl buffer, 1µl DNAse) and incubated at RT for 10 min. 1 µl EDTA (12.5 mM) was added for 10 min, 65°C. For cDNA synthesis, 1 µl deoxynucleotides and 1 µl random hexamers were incubated for 5 min at 65°C. After 1 min on ice a mixture of 4 µl 5X buffer, 0.3 µl reverse transcriptase III, 1 µl 0.1 mol/l DTT and 1.7 µl H_2O was added and incubated for 5 min at 25°C, 60 min at 50°C and 15 min at 70°C. cDNA was stored at -20°C and diluted 1: 7 for quantitative real-time polymerase chain reaction (qRT-PCR).

3.2.1.5 Whole exome sequencing

“FF samples were checked and selected with H&E stainings for high tumor cellularity. For DNA from FF samples, library preparation was performed according to Agilent’s „SureSelectXT Target Enrichment System for Illumina Paired-End Multiplexed Sequencing Library” kit, whereas for DNA from FFPE tissue the “SureSelect Automated Library Prep and Capture System SureSelectXT Automated Target Enrichment for Illumina Paired-End Multiplexed Sequencing” kit was”[1] employed at the DKFZ genomics and proteomics core facility. “Samples were run paired-end (125 bp) on a HiSeq 2000 v4”[1]. WES data from FFPE tissue has been shown to reproduce data from FF tissue[226].

3.2.1.6 Sanger sequencing

Sanger sequencing was performed by GATC Biotech by delivering DNA samples and respective primers (Table 1).

Table 1 Primer sequences used for Sanger sequencing.

primer	sequence
Seq_P16_Ex1_F	agtcctccttccttgccaac
Seq_P16_Ex1_R	agaatcgaagcgctacctga
Seq_P16_Ex2_F	cctggctctgaccattctgt
Seq_P16_Ex2_R	tggaagctctcagggtacaaa
Seq_P16_Ex3_F	tgccacacatctttgacctc
Seq_P16_Ex3_R	cgatcttgagacacggcttt
Seq_P14_Ex1_F	ctcagagccgttccgagat
Seq_P14_Ex1_R	caaacaaaacaagtgccgaat
Seq_P15_Ex1_F	cttggcccagctgaaaac

Seq_P15_Ex1_R	acatcggcgatctaggttcc
Seq_P15_Ex2_F	ggctctgaccactctgctct
Seq_P15_Ex2_R	gctttaacagcaaaacacccttt

3.2.1.7 Genome-wide analysis of DNA methylation

"DNA was isolated from FFPE tissue using the Invisorb Genomic DNA kit II. Quality was evaluated by real-time PCR analysis on Light Cycler 480 Real-Time PCR System using the Infinium HD FFPE QC kit. DNA samples were bisulfite treated with the EZ-96 DNA Methylation kit and subsequently treated with the Infinium HD DNA Restoration kit. Each sample was whole genome amplified and enzymatically fragmented following the instructions in the Infinium HD FFPE Methylation Guide. The DNA was applied to the Illumina's 450K BeadChip Array (450K array) and hybridization was performed for 16-24 h at 48°C. Microarray scanning was done using an iScan array scanner."[1] Except for DNA isolation, all steps were performed at the DKFZ core facility genomics and proteomics.

3.2.1.8 MassARRAY analysis of DNA methylation in candidate regions

"MassARRAY is a mass spectrometry based method which quantifies methylation levels of specific candidate regions[227]. DNA was bisulfite treated as described above. Regions of interest were amplified using conventional PCR with primers specific for bisulfite-converted DNA, followed by in vitro transcription and fragmentation in a base specific way and run on the MassARRAY which separates fragments due to their mass to charge ratio. The shift from an unmethylated to a methylated fragment will result in quantitative methylation values."[1] A two-sided, unpaired Welch t-test was performed to evaluate significances. Primer sequences can be found in Table 2 and Table 3.

Table 2 Primer used for MassARRAY in human samples. Forward primers (F) had a 5' tag sequence aggaagagag, while reverse primer (R) had a 5'cagtaatacgactcactatagggagaaggct.

primer	sequence
MA_CDKN2A_1F	TTTTAGTTAGTTTTGGTGTTGGAGG
MA_CDKN2A_1R	AAAAAACAATAAAAAACCCTACCCA
MA_CDKN2A_2F	AAATTTGAGGTTAAAGATGTGTGGT
MA_CDKN2A_2R	AACAATAAACTAACTACTAAACCAAAAAAA
MA_CDKN2A_3F	TGTTTTTTAAATTTTTTGGAGGGAT
MA_CDKN2A_3R	AAAACCCAATCCTCCTTCCTTAC
MA_CDKN2A_4F	GTTTTAGTTGTTAGTTAAGTTTTTGGGA
MA_CDKN2A_4R	TAAAAATACAAAAACCCTAAACCC
MA_CDKN2A_5F	TTAGAAGATTAGGTAGGAAAGGTTTT
MA_CDKN2A_5R	TATCTCCTCCTCCTCCTAACCTAAA
MA_CDKN2B_1F	GTTTGGATTGTTTTTGGGAA

MA_CDKN2B_1R	ACCCCTTAACCCAACTAAAAAC
MA_FOXD2_1F	GTGTTAGGAGTAAGTGTATTGGGGAT
MA_FOXD2_1R	ACTAAAACCCAAAACTTCTAAAAAA
MA_FOXD3_1F	TTTGTAGGATGGTTATGGTGATGAG
MA_FOXD3_1R	ACCCCTAACATTACCCAAAAAAAC
MA_FOXD4_1F	TTTTAGTAGTTGTTTTTGTTTGGG
MA_FOXD4_1R	ACCATAACCATCCTACAAAACCC
MA_FOXD4_2F	GTTAGTGGTTATTTTTTGGGATGTT
MA_FOXD4_2R	CCACCACACATAATCTTAAAAAAACTT
MA_H19_IC1_1_F	GGTTTTTATGAGTGTTTTATTTTTAGATGA
MA_H19_IC1_1_R	CACATAAATATTTCTAAAAACTTCTCCTTC
MA_H19_IC1_2_F	GGGTTGTGATGTGTGAGTTTGTA
MA_H19_IC1_2_R	AAAACAATAAAATATCCCAATTCCAT
MA_HIST1H1B_1F	GATATGGTGGTAAGAAATTGTTAGAAGA
MA_HIST1H1B_1R	CTCCACCAATCACAAAACAAC
MA_HIST1H2BI_1F	GTTTGGATTTTTTTGGAAGTGAT
MA_HIST1H2BI_1R	TCCAAAACTATAAAAATTATAAACTCCTTC
MA_KCNQ1_IC2_1_F	GTTTAAATTTTTTTAGAGAGATGGGG
MA_KCNQ1_IC2_1_R	AAACTCCTCAACATAATTCTCCTCC
MA_KCNQ1_IC2_2_F	GGTATTTTTTTGGGGGTTTTTTA
MA_KCNQ1_IC2_2_R	AAACACACAACTCACCTCAACAA
MA_SOX17_1F	ATTGGTTATATTTGTGTAGAAAAGGTT
MA_SOX17_1R	AATAAAAAATAAAACACTAAAATACCCC
MA_SOX17_2F	GGTTTAGGGGTAGGGGTTTAGT
MA_SOX17_2R	AACATCTCAATACCTCACTCCCC
MA_Twist1_1F	AGTTGTAGTTTGTTATTTTGGAGTTTA
MA_Twist1_1R	CTAAACAAAATTCAAACCCTCA
MA_Twist1_2F	GGAGTGGTTGTGATAGTAGTAATGG
MA_Twist1_2R	TTAACACTTTTCTTAACATACCCCC
PCDHA9-CpGi1-1F	GTTTGTTTTAGTTGGATTTTAAAGG
PCDHA9-CpGi1-1R	TCACCTCCAAATAAATACTACACTCC
PCDHAC1_CELL_F	TAGTTGTGTAGGGTTAAAGTTGTTTG
PCDHAC1_CELL_R	TCATTTATATCCACCACAATAATAATAA
PCDHAC1-CpGi1-1F	GGGGGTTTTTGTTTTTTTAAGTTTAG
PCDHAC1-CpGi1-1R	AAAACTACCCAAATCTTAACCTCCA
PCDHAC1-CpGi1-2F	AGTTGTGTAGGGTTAAAGTTGTTTG
PCDHAC1-CpGi1-2R	ATAACTTAATATCCCATTACTTCCCTC
PCDHAC2_CELL_F	GTTAGGAGTTTTTGGGAGGGTTTA
PCDHAC2_CELL_R	CAACACTTCCAAACTAAACAAACAA
PCDHAC2-CpGi1-1F	GATATTTTGTGTTAGAGGAGTAGGTATT
PCDHAC2-CpGi1-1R	CAACACTTCCAAACTAAACAAACAAC
PCDHB10-CpGi-1F	TTATTTGGTGATTAAGGTGGTGG
PCDHB10-CpGi-1R	TCCACCAAAAACAAATACAAC
PCDHB14_CELL_F	GGATAATGGTTATTTGTTTGTTTTT
PCDHB14_CELL_R	TTAACCAACACCACCAACCTATACT

PCDHB14-CpGi-1F	GGATAATGGTTATTTGTTTGTTTTT
PCDHB14-CpGi-1R	CCACCACCTTAATCACCAAAT
PCDHB15_CELL_F	ATTTGGTGATTAAGGTGGTGG
PCDHB15_CELL_R	TCCACCAAATACCCTAAAAAAA
PCDHB15-CpGi-1F	TTAGTGGTGTTGGTTAAGGATAATG
PCDHB15-CpGi-1R	CCAATACCACCACCAAATAAAC
PCDHB16-CpGi-1F	GGGTTATTTGGTGATTAAGGTGG
PCDHB16-CpGi-1R	TAACCAACACCACCAACCTCTACTT
PCDHB4_CELL_F	ATTTGGTGATTAAGGTGGTGG
PCDHB4_CELL_R	TCCACCAAATACCCTAAAAAAA
PCDHB4-CpGi-1F	GTTTATTAGATGTTTTGGAAAGGGG
PCDHB4-CpGi-1R	TAAATAACTTCTCCCAACCCTACCT
PCDHGA12_CELL_F	AAAAATTAGTGAAAATGTAGTTATTGAGA
PCDHGA12_CELL_R	CACTTCCATCTAATAAAATCCTAACTCC
PCDHGA12-CpGi-1F	TAGATATAGTTTTGGATAGGGAATAGG
PCDHGA12-CpGi-1R	AAAAAACTCCTCTAAAATTATTCTCTAAAA
PCDHGA12-CpGi-2F	TTTTAGATTTTATTTTGTATTTGGTGG
PCDHGA12-CpGi-2R	AATCTACAAAAAAACCTACACCCC
PCDHGB4/GA8-CpGi1-1F	GTTGTTGGTTAAAGTGGAGAGTTT
PCDHGB4/GA8-CpGi1-1R	TCCTCAACTTACTCACCAACCTA
PCDHGB7_CELL_F	TGGGGTAGAATAAAGGTAGTAGTAAAGG
PCDHGB7_CELL_R	AAATCCCTCAACCTCTAACCTAAAA
PCDHGB7-CpGi-1F	GTTAGGTTATTTGGTGATTAAGGTGG
PCDHGB7-CpGi-1R	CAACCCCAAACTAAAAAACCC
PCDHGC3_CELL_F	GGGATGAGGTAGAGATTGAATAGT
PCDHGC3_CELL_R	ACAAAAAAACAATATCCCACACAAC
PCDHGC3-CpGi1-1F	TATTGGTTGGGATTTTGTGTGT
PCDHGC3-CpGi1-1R	AAACCATCTCACTCAAAAACAATTC
PCDHGC3-CpGi1-2F	TTTATTTAGGAAATGAAATTGGAGA
PCDHGC3-CpGi1-2R	TCAACACTAACTAAAAACTAAACTCCC

Table 3 Primer used for MassARRAY in mouse samples. Forward primers (F) had a 5' tag sequence aggaagagag, while reverse primer (R) had a 5'cagtaatacgactcactatagggagaaggct.

primer	**sequence**
mPcdha8_CpGi_long_1F	GGGTAGGTGATTTGTTTTTTGATTA
mPcdha8_CpGi_long_1R	AAACTACAACAACTCCAACTCCTCA
mPcdhac1_CpGi_long_1F	TAATATTTGGAGTTTGTAATGTAAGTAGG
mPcdhac1_CpGi_long_1R	AAAAAATCCTCAAACATTAACTCAT
mPcdhac1_CpGi_short_1F	AATATTTGGAGTTTGTAATGTAAGTAGG
mPcdhac1_CpGi_short_1R	TACAACACCACAAAAAACCCAC
mPcdhac2_CpGi_long_1F	TTTGTTTTGGTTGTTGTTTTTGTTT
mPcdhac2_CpGi_long_1R	TCTCCTACAAAAATTTAAAATCCAACT

mPcdhac2_CpGi_short_1F	ATTTGGAATTGGATTTGATTAATGG
mPcdhac2_CpGi_short_1R	CCACAAAATTATAAACCAACACTTC
mPcdhb17_CpGi_long_1R	AAAAAAACCATCCACCACTAACAC
mPcdhb17_CpGi_short_1F	TGGGTGTAATAGATTAAGGTTTATTTG
mPcdhb17_CpGi_short_1R	ACAACCACCACCTTAATAACCAA
mPcdhb17+20_CpGi_short_1F	GTTGGATTATGAGGTTTTGTAGGTTT
mPcdhb18_CpGi_long_1F	AATTAAGGTTTTGGATTATGAAGTTT
mPcdhb18_CpGi_long_1R	TCCACCAACAACACCTACAATATAA
mPcdhb18_CpGi_short_1F	GAGGTTTATTTGAGTTTAGTAGTTAGGTTT
mPcdhb18_CpGi_short_1R	ATCCACAACCACCACCTTAATAAC
mPcdhb20_CpGi_long_1F	GGTTTAGGATTTAGAGTTGGTTTT
mPcdhb20_CpGi_long_1R	AAAAAAACCATCCACCACTAACA
mPcdhb20_CpGi_short_1R	TAACCAAATTCTACCACCCTAAACA
mPcdhb21/2_CpGi_long_1F	GGTAGGTAGGGTTGAGAGAAGTTATTT
mPcdhb21/2_CpGi_long_1R	CACTACTACCACCCCAAAACCAA
mPcdhb21/2_CpGi_short_1F	ATTTGGTTATTAAGGTGGTGGT
mPcdhb21/2_CpGi_short_1R	CCATTATCCTTAACTAACAACAACAA
mPcdhga12_CpGi1_long_1F	AAGGATTTAGGGTGGGTAATATTTTT
mPcdhga12_CpGi1_long_1R	AACTCTCCAACACCAATTCTAAATA
mPcdhga12_CpGi1_short_1F	AGGATTTAGGGTGGGTAATATTTTT
mPcdhga12_CpGi1_short_1R	TAATAAACCCCATACACAACTCCTC
mPcdhgb4_CpGi_long_1F	ATTTGGGGGTTAATGGTTAGGTTAT
mPcdhgb4_CpGi_long+short_1R	CTAAATCCCTATCTCCCAAAACAC
mPcdhgb4_CpGi_short_1F	TATTTGGTTATTAAGGTGGTGGTTG
mPcdhgb7_CpGi1_short_1F	TTTTTTTATTGTGAAGATTTGAAGA
mPcdhgb7_CpGi1_short_1R	AACAAAACAAAAACAACAACAAAAT
mPcdhgc3_CpGi_long_1F	GTTGGATAGGAAATTTTGGAAAGTA
mPcdhgc3_CpGi_long_1R	CAACTAACAAAAACAAAACTCCCAC
mPcdhgc3_CpGi_short_1F	TATTGTAATTTTGGAGTTGGTGGTAG
mPcdhgc3_CpGi_short_1R	TTTATAAAAAATTACTCCCCACATC
mLINE1_2_F	TTTAATGATATTTTGGTTAAGGAAGGA
mLINE1_2_R	ATTTCCACTTAATTAATTTCACCCC
mLINE1_4_F	AAATGGGGTTTAGAATTGAATAAAGA
mLINE1_4_R	TCCATTCCTCTATTAAAAAACATCT
mLINE1_5_F	TTTTTGGTTTTATGTGAAGTTTTTTG
mLINE1_5_R	AAAATAAACCCACACACCTATAATCA

3.2.1.9 Quantitative (reverse transcriptase) PCR

“Copy numbers were evaluated by performing quantitative PCR on the target regions by the LightCycler 480 System. Actin β (ACTB) and Albumin (ALB) were used as housekeeping genes.”[1] The master mix consisted of 2 µl of DNA (for quantitative polymerase chain reaction (q-PCR)) or cDNA (for qRT-PCR)), 2.5 µl probe master mix, 0.04 µl of the respective probe and

0.5 µl 10 µM primer mix (forward and reverse). Primer sequences can be found in Table 4, Table 5 and Table 6.

Relative expressions for the gene of interest (GOI) were calculated with the following formula:

$$\text{relative expression} = 2^{-[\text{case}(C_{T(GOI)} - C_{T(\text{average HK})}) - \text{control}(C_{T(GOI)} - C_{T(\text{average HK})})]}$$

CNA were calculated with the following formula:

$$Y = \text{Eff}_{GOI}{}^{-C_{T(GOI)}} - \text{average}(\text{Eff}_{ACTB}{}^{-C_{T(ACTB)}} + \text{Eff}_{ALB}{}^{-C_{T(ALB)}})$$

$$\text{relative DNA content} = \frac{Y_{\text{Sample}}}{Y_{\text{average normal tissues}}} \text{ x } 2$$

Table 4 qRT-PCR primer human.

primer	sequence	probe
Exp-PCDHA9-1F	tacagagcgaacgggagaac	38
Exp-PCDHA9-1R	cagtcagggttgggctgt	38
Exp-PCDHAC1-1F	aggggatcacgctaatgtca	16
Exp-PCDHAC1-1R	ctgtgcatgcctgctctc	16
Exp-PCDHAC2-1F	ccagcttcaggtaagcgaat	1
Exp-PCDHAC2-1R	cctgcgcactctctatgtga	1
Exp-PCDHA@1,2-1F	cctgactggcgttactctgc	75
Exp-PCDHA@1,2-1R	ctggaccagcccgtagaat	75
Exp-PCDHA@2,3-1F	ccaacagtatccagtgcaacac	2
Exp-PCDHA@2,3-1R	aaggtccagctgttgctgtt	2
Exp-PCDHB4-1F	tggtgcaggacaaaccact	68
Exp-PCDHB4-1R	gccacgagggttaagctg	68
Exp-PCDHB16-1F	cggaaaatacaactgcctacg	1
Exp-PCDHB16-1R	cttccaggctttgggaaac	1
Exp-PCDHB10-1F	tcattttccaactctgttgctg	2
Exp-PCDHB10-1R	aaccatctttccattttctcca	2
Exp-PCDHB14-1F	ttccagagaatgcctcagaga	88
Exp-PCDHB14-1R	gggaggttatcttgaatagagcaa	88
Exp-PCDHB15-1F	aatgcatttccaagtgttactgaa	2
Exp-PCDHB15-1R	tctgggattttcagggtcat	2
Exp-PCDHGB4-1F	aacgctccggttttctcac	57
Exp-PCDHGB4-1R	cagaagccctgacttgtgaa	57
Exp-PCDHGC3-1F	ctgtcatcgctttgctcagt	85
Exp-PCDHGC3-1R	aggtcaccagcccgttct	85
Exp-PCDHGA12-1F	cccagatacgctattcagttcc	78
Exp-PCDHGA12-1R	agatgtcgcccaccctaga	78
Exp-PCDHG@3-1F	ggtgtgcttttacgtgatgg	44

Exp-PCDHG@3-1R	ggcagatcaaggacagacg	44
Exp-GAPDH-F	agccacatcgctcagacac	60
Exp-GAPDH-R	gcccaatacgaccaaatcc	60
Exp-HPRT-F	tgaccttgatttattttgcatacc	73
Exp-HPRT-R	cgagcaagacgttcagtcct	73
Exp-ACTB-F	attggcaatgagcggttc	11
Exp-ACTB-R	ggatgccacaggactccat	11

Table 5 qRT-PCR mouse primer.

primer	sequence	probe
Exp-mpcdha8-1F	ttctttggactcctccgaga	13
Exp-mpcdha8-1R	ctgtgcatgcctgctcttag	13
Exp-mpcdhac1-1F	ggatcattcaaatgtggaagc	13
Exp-mpcdhac1-1R	ctgtgcatgcctgctcttag	13
Exp-mpcdhac2-1F	aaaaatgatgctggctctcaa	13
Exp-mpcdhac2-1R	ctgtgcatgcctgctcttag	13
Exp-mpcdhb18-1F	attgaccaagccttttgcat	22
Exp-mpcdhb18-1R	gcctggctggagatagagc	22
Exp-mpcdhb20-1F	aggcgagtgctgaatctttg	6
Exp-mpcdhb20-1R	gccacaaaagagccaatttc	6
Exp-mpcdhb22-1F	aaaaggaatcccgcttgc	48
Exp-mpcdhb22-1R	gcggtccagtttttcatttaag	48
Exp-mpcdhga12-1F	ccagatacgctattcggttcc	78
Exp-mpcdhga12-1R	gagatgttgcccaccctaga	78
Exp-mpcdhgb4-1F	acttcgactgcgacactcct	4
Exp-mpcdhgb4-1R	ccaccacttccgactgaaag	4
Exp-mpcdhgb7-1F	cagcgtgcctttgaccat	12
Exp-mpcdhgb7-1R	gtctcgcgcttgtaaagtcag	12
Exp-mpcdhgc3-1F	acaggtgttgggtgcagag	56
Exp-mpcdhgc3-1R	gcttgagagaaacgccagtc	56
Exp-mpcdhb17-1F	tcaacgataatgcacccaag	66
Exp-mpcdhb17-1R	tgttctcgggaatttgcact	66
Exp-mpcdhg@1,2-1F	cgtttctctcaagcccagag	9
Exp-mpcdhg@1,2-1R	gcatctctgtatcaaactggttgt	9
Exp-mpcdhg@2,3-1F	atgctgcaagccatgatct	53
Exp-mpcdhg@2,3-1R	cagggtagagctcccatcag	53
Exp-mpcdha@1,2-1F	actctgcctcgctaagagca	25
Exp-mpcdha@1,2-1R	gaccagcccgtagaatgc	25
Exp-mpcdhg@2,3-1F	ccaacagtatccagtgcaacac	51
Exp-mpcdhg@2,3-1R	aaggtccagctgttgctgtt	51
Exp-mmap4k1-1F	taccgggccagctcacta	74
Exp-mmap4k1-1R	ggcctctggaactccattt	74

Exp_mHPRT1_F	tcctcctcagaccgctttt	95
Exp_mHPRT1_R	cctggttcatcatcgctaatc	95
Exp_mTBP_F	cggtcgcgtcattttctc	107
Exp_mTBP_R	gggttatcttcacacaccatga	107
Exp_mActb_F	aaggccaaccgtgaaaagat	56
Exp_mActb_R	gtggtacgaccagaggcatac	56

Table 6 q-PCR primer human.

primer	sequence	probe
qCDKN2B_Ex1_1F	aaccgttacaattgctctcactc	22
qCDKN2B_Ex1_1R	ttccgcaggcagactacac	22
qCDKN2A4_P14_Int2_1F	gggtctccttcatttggtga	25
qCDKN2A4_P14_Int2_1R	cccagggaggagagtctga	25
qCDKN2A_Int1_2F	ggagccgaagtctccttctt	55
qCDKN2A_Int1_2R	gccaattccctccagttaca	55
qSOX2_DS_1F	ttgcaaactagacatgcaaagtg	3
qSOX2_DS_1R	gggccagagtaatccaaaca	3
qAPC_Int6_1F	ggctctagcaaccctccttc	75
qAPC_Int6_1R	tgcgatgtacacgcctaaag	75
qAPC_Int1_1F	ggccattagcagtattcgatg	77
qAPC_Int1_1R	ccctaggcttttggctatttg	77
qTET_Int2_F	gtaaatggaggcatagaggcata	29
qTET_Int2_R	tgggatgactattccccaaa	29
qTET_Int2_F	ttacacactgaacctttgtaccg	55
qTET_Int2_R	tgtagttaagagaaaagggaagttca	55
qID3_Int2_F	cgggggaggaaagaagac	68
qID3_Int2_R	gctggaggtacagaccaagttt	68
qID3_Int1_F	aagagttacgcgaggcaatc	89
qID3_Int1_R	tccaggtaagcctcgaagtc	89
qARID1A_Int1_F	ccctggactgaaggaactca	2
qARID1A_Int1_R	cacaatggcacccaactgt	2
qFOXD1_Ex1_F	caattggaaatcctagcagtaaagt	47
qFOXD1_Ex1_R	gactctgcaccaagggactg	47
qFOXD1_Ex1_F	gagaggttgtggcggatg	7
qFOXD1_Ex1_R	gacgctgagcgagatctgt	7
qACTB_A1_F	gctacgagctgcctgacg	54
qACTB_A1_R	ggctggaagagtgcctca	54
qALB_A1_F	acaaagatgacaacccaaacct	9
qALB_A1_R	aagcagtgcacatcacatcaa	9

3.2.1.10 Immunohistochemistry

Immunohistochemistry (IHC) was performed with the avidin-biotin complex method on a Ventana BenchMark Ultra, for antibodies refer to Table 7. “Staining intensities were evaluated by a pathologist. Proteins were considered downregulated when the value of the tumor was at least one staining intensity lower than the lowest staining intensity of the normal tissues whereas proteins were considered upregulated when the value of the tumor was at least one staining intensity higher than the highest staining intensity of the normal tissues.”[1]

Table 7 Antibodies and concentrations used for IHC. poly: polyclonal, mono: monoclonal

gene	lot number	clone	source	clonality	dilution/time	universal linker/HRP multimer
APC	A57860	-	rabbit	poly IgG	1: 50/24'	8'/8'
ARID1A	G105716	-	rabbit	poly IgG	1: 50/24'	8'/8'
CDKN2A (P16)	3294873	G175-405	mouse	mono IgG1	1: 50/24'	8'/8'
HIST1H	210948	-	rabbit	poly IgG	1: 50/36'	12'/12'
ID3	SAB1412646	3D3	mouse	mono IgG2aκ	1: 1000/24'	8'/8'
JAK1	871521119	-	rabbit	poly IgG	1: 50/60'	12'/12'
PCDHG	443-3KS-15	N159/5	mouse	mono IgG1	1: 100/24'	8'/8'
SOX2	GR196138-7	-	rabbit	poly IgG	1: 100/24'	8'/8'

3.2.1.11 Cell culture

Cell culture was performed according to standard cell culture regulations. Cells were grown without antibiotics and were subjected to regular mycoplasma tests. For media refer to Table 8.

Table 8 Cell lines and media used.

cell line	organism	medium
266-6	Mouse	DMEM+10% FCS
T510558	Mouse	DMEM+10%FCS+non-essential amino acids
T510586	Mouse	DMEM+10%FCS+non-essential amino acids
T510677	Mouse	DMEM+10%FCS+non-essential amino acids

3.2.1.12 5-Aza-2'-deoxycytidine treatments

Cells were treated for three subsequent days with 0.5, 1.0 or 2.0 µM of 5-Aza-2'-deoxycytidine (DAC) and harvested on day four.

3.2.1.13 Small interfering RNA treatments

For small interfering RNA (siRNA) treatments, cells were transfected in 12 well plates with 10 mM siRNAs (Table 9) with Dharmafect transfection reagent according to the manufacturer's protocol. For sequences refer to Table 9.

Table 9 siRNA sequences for knockdown experiments.

siRNA	sequence
siPCDHG@	gcacctggcccaacaacca
siPCDHG@	cagagatgctgcaagccat

3.2.1.14 Cell viability assay

Cell viability was assessed with Calcein-AM assays. One hour before read-out, medium was replaced with fresh medium containing 0.8 µM Calcein-AM. Just before read out cells were washed with PBS and lysed with PBS with 10% Triton-X 100 for 5 min on a shaker. Readout was performed with a Spectramax at 494 nm excitation wavelength and detected at 530 nm.

3.2.2 Bioinformatical methods

3.2.2.1 Whole exome sequencing

"Alignment of data was processed by the following parameters: reference genome: hs37d5, alignment program: bwa-0.7.8 mem, alignment parameter: -T 0, duplication marking program: picard-1.125. Default duplication marking program parameters were used (https://broadinstitute.github.io/picard/command-line-overview.html#MarkDuplicates). Alterations that are likely to be benign, so called FrequentLy mutAted GeneS (FLAGS)[228] were excluded. Previously published WES studies of ACC[182,186] were utilized as a comparison to the here generated WES results."[1] The Data Management and Genomics IT (eilslabs), DKFZ aligned sequencing data and called SNVs.

3.2.2.2 Mutational signatures

"Mutational frequencies were plotted using the SomaticSignatures package (version 2.8.4) available on Bioconductor[229]. WES data from ACC and publicly available WES from TCGA (available in the SomaticCancerAlterations Bioconductor package 1.8.2[230]) were used to assess frequency of point mutations in the context of three nucleotides (X-A/T-X). For investigating the occurrences of published mutational signatures summarized by the catalogue of somatic mutations in cancer (COSMIC)[26] in ACC, deconstructSigs (version 1.8.0)[231] was employed."[1]

3.2.2.3 Tumor purity (LUMP)

“In addition to the high tumor purity assessment of the pathologist, tumor purity was estimated by the LUMP (leukocytes unmethylation for purity) method[232], which correlates well with the other tumor purity methods ESTIMATE and ABSOLUTE[232]. To that end, 44 immune-specific CpG site methylation levels were averaged and divided by 0.85[232].”[1]

3.2.2.4 DNA methylation phylogenetic tree

“Phylogenetic trees were generated calculating Euclidean distance matrices based on all CpG sites on the array by the minimal evolution method[233] using the fastme.bal function”[1] and plotted with the plot.phylo function “from the R package ape[234].”[1]

3.2.2.5 Analysis of genome-wide DNA methylation by 450K

“Analysis was run with the R Package RnBeads[235]. In summary, data was subjected to quality control. Normalization with BMIQ[236] and differential methylation was then called between groups using RnBeads rank cutoff, which implements the difference in mean methylation, the quotient in mean methylation and the p-value obtained by limma test (for multiple comparisons) or Student’s t-test (for paired analysis). Region annotations were implemented to analyze the average differential methylation in the following regions: gene bodies (defined by Ensembl[237]), promoter regions defined by RnBeads[235] (regions spanning 1.5 kb upstream and 0.5 kb downstream of the transcription start site for every gene[235]), CpG Islands (defined according to the definition of the UCSC browser[45]), and promoter segments (intersection of promoters and CpG islands). For investigation of recurring methylation changes from primary tumors to metastases, paired t-test was used to calculate significances.”[1]

3.2.2.6 Cell type contributions

“For cell type contributions, the method published by Houseman, et al. [238] was used which is a method similar to regression calibration. 450K data from normal hematopoietic cell types were obtained from Reinius, et al. [239]. In addition to the blood sample methylation data from this work DNA methylation profiles from sorted pancreatic cells (acinar, duct, α- and β-cells)“[1] obtained from Yuval Dor (The Hebrew University-Hadassah Medical School, Jerusalem, Israel) were added.

3.2.2.7 Gene enrichment analysis

“The function annotation tool from the DAVID website was used[240] for gene enrichment analysis.”[1]

3.2.2.8 Enrichment of transcription factors, chromatin states and histones at differentially methylated sites

Data from the Encyclopedia of DNA Elements (ENCODE, https://www.encodeproject.org/) was downloaded, including transcription factor binding sites and histone marks from Panc1, and chromatin states from human mammary epithelial cells (HMEC), and overlapped with differentially methylated sites (DMS). These were defined by RnBeads rank cutoff and were used as input, whereas all other sites remaining after quality control and normalization served as background input. P-values were calculated using Fisher's exact test.

3.2.2.9 Transcription factor binding site enrichment (HOMER)

Hypergeometric Optimization of Motif EnRichment (HOMER)[241] was used to calculate known transcription factor binding motives in the differentially methylated regions (DMRs). Inputs were the same sites used for 3.2.2.8.

3.2.2.10 Analysis of genome-wide copy numbers by 450K

"The Bioconductor package conumee[242] was used to calculate CNA from the intensities obtained from the 450K array (bin probe size was set to five, rest of parameters set to default). Gistic[243] was then employed to investigate frequently deleted/amplified regions and genes (with default parameters)."[1]

3.2.2.11 Circos plots

"The R package circlize was used to create circos plots[244]. For CNA the total number of tumors out of all 41 primary tumors that were subjected to 450K analysis which were amplified (red) or deleted (blue) was depicted. For aberrations in methylation the difference of the promoter region means (tumor minus normal tissue) is depicted (hypermethylation dark green, hypomethylation yellow)."[1] The top or bottom end of each circle represents all tumors, *i.e.* 100%.

3.2.2.12 Calculation of integrative categories

"To integrate changes in methylation and CNA within the genes, data was categorized into the following nine groups: deleted and promoter hypermethylated, deleted only, hypermethylated only, deleted and promoter hypomethylated, unaltered, amplified and promoter hypermethylated, hypomethylated only, amplified only, amplified and hypomethylated. The following cutoffs were used: for CNA the deletion and amplification thresholds from Gistic were employed[243]. For differential methylation, a tumor was called to be hypo-/hypermethylated at a gene if the promoter methylation was below or above 20% of the mean methylation level of all normal pancreatic tissues."[1]

3.2.2.13 Calculation of intertumor heterogeneity

"iCluster Plus[245] was used to investigate molecular subgroups of ACC. Three clusters were determined to explain the highest variation while showing the least complexity. Definition of these clusters was only based on CNA and promoter DNA methylation, as point mutations did not add any information to these clusters."[1]

4 Results

4.1 Study design of the thesis

Two independent cohorts for ACC were employed. Cohort I consisted of 21 primary tumors (18 pure ACC, 3 MACNEC), 12 metastases (6 lymph node metastases, 5 liver metastases, and 1 peritoneal metastasis), 12 adjacent normal pancreatic tissues, and 6 healthy normal pancreatic tissues. Cohort II consisted of 38 primary tumors (30 pure ACC, 8 MACNEC), 1 peritoneal metastasis, 2 adjacent normal pancreatic tissues, and 8 healthy normal pancreatic tissues. The mean age of patients was 56.5 years (range 29-75) and 58.2 (range 7-79) in cohort I and cohort II, respectively. Males were predominantly affected, with only 26.1% and 28.2% female patients in cohort I and cohort II, respectively. Mean survival in cohort I was 27 months, whereas survival data was not available for cohort II. For a summary which tumors were employed for the different experiments refer to Table 10. In the following, all primary tumors are labelled with a preceding “T”, all lymph node metastases with a preceding “LK”, and all remaining metastases with a preceding “M”, while normal tissues are labelled with a preceding “N”. Each patient was given a unique number, while patients from cohort II have a “K” preceding the patient number. For example, 6T was the primary tumor of patient 6 of cohort I, while 6LK was the lymph node metastasis of the same patient, and sample K4T was the primary tumor of patient 4 in cohort II.

Table 10 Patient characteristics and summary of samples used for molecular analyses.

	cohort I	cohort II
total number of tumors	34	39
general characteristics		
tissue collected in	Heidelberg	all over the world
year of collection	2003-2013	1990-2004
female: male (%female)	6: 17 (26.1 %)	11: 28 (28.2 %)
mean age in years (range)	56.5 (29-75)	58.2 (7-79)
median survival in months	27	NA
mutational analysis – whole exome sequencing (FFPE+FF)		
primary tumors (pure ACC, MACNEC)	22 (18, 4); 5FF+17FFPE	1 FFPE
normal tissues (adjacent healthy pancreas, healthy liver)	22 (19, 3)	1
methylation analysis and copy number aberrations- 450K Array (FFPE)		
primary tumors (pure ACC, MACNEC)	22 (18, 4)	19 (15, 4)
metastases	12 (6 lymph nodes, 5 liver, 1 peritoneal)	NA
normal tissues (adjacent)	20 (14)	10 (2)
protein analysis - immunohistochemistry (FFPE)		
primary tumors (pure ACC, MACNEC)	20 (16, 4)	38 (30, 8)
metastases	3 (liver)	1 (peritoneal)
normal tissues	8	

4.2 ACC harbor high mutational loads, however no frequently recurrent events occur

Up to date, no precursor lesions or frequent molecular events during carcinogenesis were identified in ACC. This is in contrast to PDAC, where the sequential development of cancer by mutations of driver genes are well described (refer to 1.2.3.2). Previously performed WES revealed that there were no frequent mutations in ACC[182,186]. To examine the ACC cohorts, 22 tumors for which matched normal tissues were available, were sequenced. "A median number of 137 mutations/tumor were identified with a wide-spread range from 40 to 1023 mutations/tumor"[1] (Figure 9a), "similar to previous reports by Jiao, et al. [182] and, Furukawa, et al. [186]."[1] FLAGS[228] are genes that are often mutated in public exomes and are unlikely to contribute

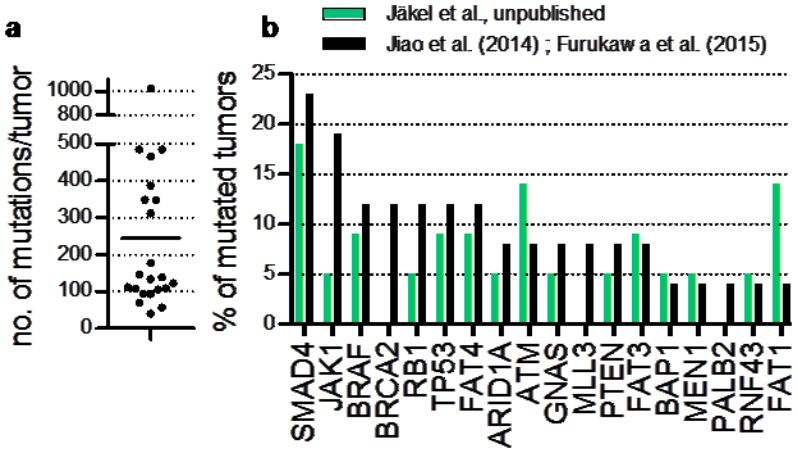


Figure 9 ACC harbor many mutations but no frequent events. a. number of mutations per tumor determined by WES. Line represents average of all tumors (244 mutations). **b.** comparison of WES (in green) with previously published mutations in ACC (in black).

to tumorigenesis and are probably only non-coding variations. These genes have in general longer coding sequences and are less evolutionary conserved. After excluding FLAGS from the WES data, no mutations occurring frequently were found. "Genes mutated in more than one sample were *COL12A1, FRY, FRYL,* and *PLB1* which occurred in 23% of carcinomas"[1] (5 out of 22), "followed by *CACNA1A*, *CCDC57*, *COL23A1*, *MKL2*, *RAP1GAP*, *SMAD4*, and *SRCAP* occurring in 18% of carcinomas"[1] (4 out of 22), and 68 genes mutated in 14% of tumors (3 out of 22) (Table 11). These mutations were then compared "to published WES data on ACC."[1] This revealed "similar low frequency mutations"[1] in the genes "reported by Jiao, et al. [182] and Furukawa, et al. [186], *e.g. SMAD4* was amongst the most frequently mutated genes"[1], followed by *JAK1* and *BRAF* mutations (23%, 19%, 12% in the published studies and 18%, 5%, 19% in the here presented study, respectively). "Genes which are often mutated in other tumors, *e.g. TP53*, *BRCA1*, and *MEN1* showed a very low frequency in ACC"[1] (Figure 9b). Overall, no frequent "mutations were identified suggesting that point mutations play a minor role in ACC."[1]

Table 11 Most frequent mutations in ACC. Most frequent mutations detected in WES after excluding FLAGS (5 out of 22) are listed on top, followed by 4 and 3 out of 22. No. = Number.

Mutation frequency (No. out of 22)	Affected genes
23% (5)	*COL12A1, FRY, FRYL, PLB1*
18% (4)	*CACNA1A, CCDC57, COL23A1, MKL2, RAP1GAP, SMAD4, SRCAP*
14% (3)	*ABI3BP, ADAMTS2, ADCY4, AFF3, ALK, APOE, ATP1A2, AXIN1, C3orf20, CCBE1, CHD1, COL4A2, CSMD3, CTTNBP2, DNMT3A, EPS8, FAM129B, FANCA, GLI3, GPR112, HCFC1, HK2, HLA-B, HLA-DQB1, HLA-DRB1, HLA-DRB6, HMHA1, IL15RA, ITGA7, KCNT1, KIAA1109, KMT2A, LGR6, LLGL2, MAST3, MYH11, NUP98, OTOG, PDE10A, PLCB1, PLCD4, PLEKHG3, POLR1A, PTK7, RASGRF1, RECQL, RHPN1, RNF44, RP11-112H10.4, RP11-161M6.5,SSTR5-AS1, RP11-188C12.3, RP11-319G6.1, RRBP1, RREB1, SCN10A, SHC4, SHPRH, SNAP91, SPTA1, SPTBN4, TDO2, TMEM132C, TTC40, UNC13B, VAV2, WBSCR17, WDFY4, WDR24*

4.3 ACC exhibit distinct mutational signatures associated with defective DNA repair, smoking and deamination of 5-methylcytosine

As large ACC cohorts are lacking both on the tissue level but also on the level of clinical data, studies assessing risk factors for ACC so far had not been possible. “Mutational signatures as an alternative approach to tackle this challenge”[1] were generated based on the WES results. “Different carcinogens such as UV light or tobacco have specific impacts on the mutational patterns of cancer[19]”[1] (refer to 1.1.2.2). “The mutational frequencies of point mutations (*i.e.* C>A, C>G, C>T, T>A, T>C, T>G,”[1] in respect to the pyrimidine base) “in the context of one base up- and one base downstream”[1] were calculated, *i.e.* it is not relevant which gene is mutated but only in which base pair context a base is mutated. Using this approach frequent C>A and C>T point mutations in ACC were identified (Figure 10, first row), which “is similar to what was observed in other tumor entities”[1] from TCGA (Figure 10, rows two until nine): Glioblastoma

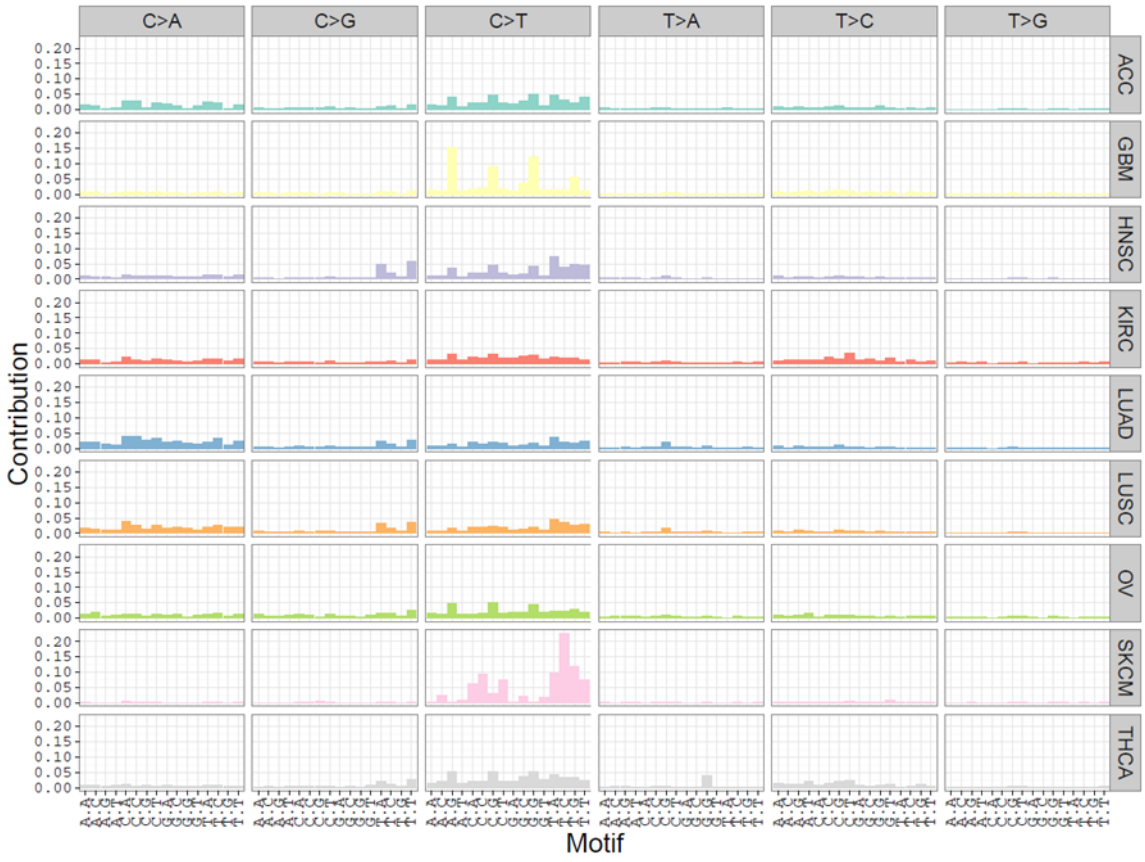


Figure 10 Mutational signatures of ACC compared to published datasets from TCGA. GBM: glioblastoma multiforme, HNSC: head and neck squamous cell carcinoma, KIRC: kidney renal clear cell carcinoma, LUAD: lung adenocarcinoma, LUSC: lung squamous cell carcinoma, OV: ovarian serous cystadenocarcinoma, SKCM: skin cutaneous melanoma, THCA: thyroid carcinoma.

multiforme (GBM), head and neck squamous cell carcinoma, (HNSC), kidney renal clear cell carcinoma (KIRC), lung adenocarcinoma (LUAD), lung squamous cell carcinoma (LUSC), ovarian serous cystadenocarcinoma (OV), skin cutaneous melanoma (SKCM), and thyroid carcinoma (THCA) all harbor many C>T mutations, while KIRC, LUAD, LUSC, and OV additionally harbor frequent C>A mutations. Comparing the signatures detected in ACC to previously "published mutational signatures as depicted in the COSMIC database[26] using the deconstructSigs package[231], 18 out of the previously published 30 signatures were identified in ACC"[1] (Figure 11). "Each tumor displayed on average four signatures (range: one to six; for details refer to Supplementary Table 1). Signature 1, corresponding to deamination of 5-methylcytosine"[1], a process associated with age is found in most tumor types and samples. "This mutational signature was found in 20 out of 22 ACC. Signature 4 which is related to tobacco smoking was present in 12 tumors. In combination with tobacco chewing (signature 29) which was detectable in two patients, almost two third (14/22) of ACC had mutational signatures associated with tobacco-associated carcinogens. Signatures associated with defective DNA repair (signatures 3, 6, 15, and 20) were identified in 15/22 tumors. Signature 3 is specifically associated with *BRCA1* and *BRCA2* mutations, whereas the other three are associated with defective DNA mismatch repair. Thus, ACC revealed distinct mutational signatures. This suggests defects in DNA double-strand break and DNA mismatch repair which may contribute to the high genomic instability observed in this disease"[1] (refer to 4.5).

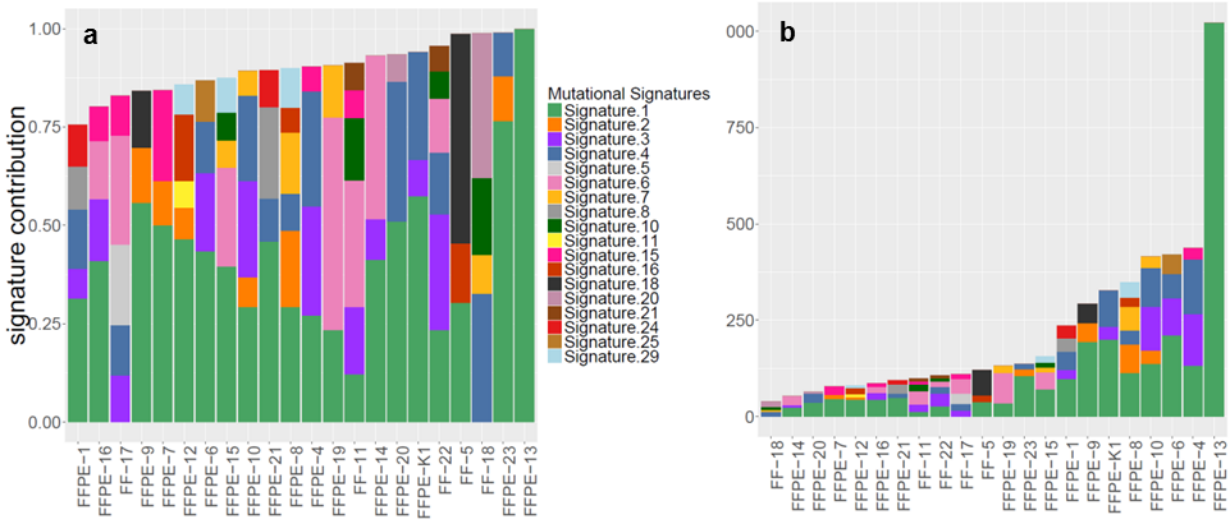


Figure 11 Contribution of published signatures to the mutational profile of each of the sequenced tumors a. relative contribution **b.** in relation to absolute numbers.

4.4 ACC harbor unique methylomes

Methylomes of ACC were analyzed and to that end all tissues from cohort I, and 19 primary tumors, 2 adjacent normal tissues, and 8 healthy normal tissues from cohort II were subjected to 450K array analysis. To compare ACC to other pancreatic cancers, 17 PNET were analyzed by 450K and data of a total of 146 PDAC and 10 adjacent normal tissues were downloaded from TCGA.

4.4.1 Generated data on ACC is of high quality

In a first step, the quality of 450K data was assessed. All samples were checked for quality control probes of staining, hybridization, extension, target removal, bisulfite conversion,

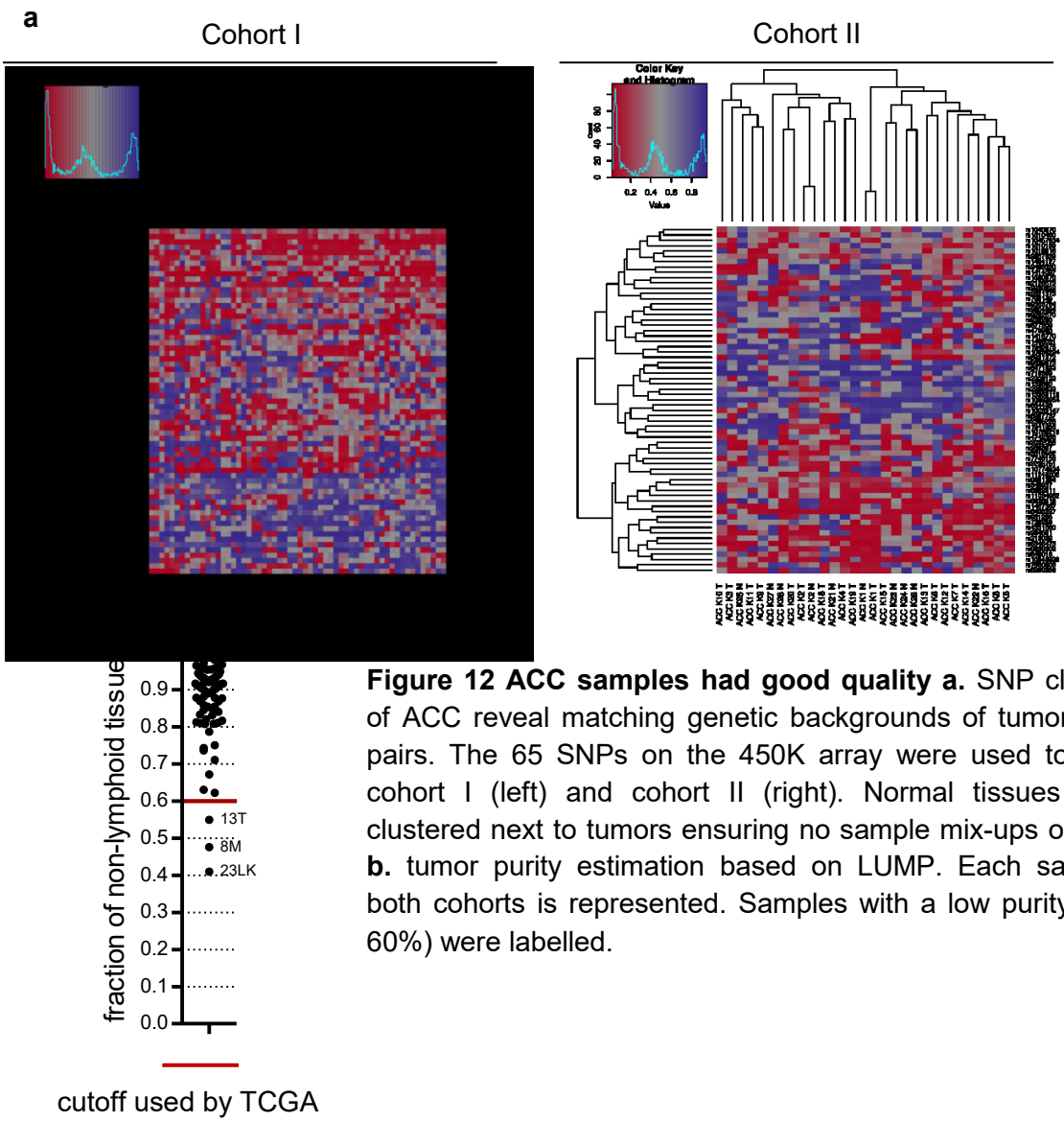


Figure 12 ACC samples had good quality a. SNP clustering of ACC reveal matching genetic backgrounds of tumor-normal pairs. The 65 SNPs on the 450K array were used to cluster cohort I (left) and cohort II (right). Normal tissues always clustered next to tumors ensuring no sample mix-ups occurred. **b.** tumor purity estimation based on LUMP. Each sample of both cohorts is represented. Samples with a low purity (below 60%) were labelled.

specificity and negative control. Cluster analysis of the 65 SNPs on the array confirmed that samples were not mixed up as primary tumors always clustered together with their adjacent normal tissues and their metastases in cohort I and cohort II (Figure 12a).

Next, tumor purity was estimated with the LUMP algorithm, which measures infiltration of leukocytes[232]. The LUMP value highly correlates with other tumor purity estimates, i.e. ESTIMATE (based on expression[246]) and ABSOLUTE (based on copy number data[247])[232]. Nearly all tumor samples were above the purity cutoff used by TCGA of 60%. Only three samples, namely 13T, 8M, and 23LK were below the threshold (Figure 12b). It was also assessed whether the two cohorts behave similar. In an unsupervised clustering including all investigated ACC, tumors from both cohorts intermingle (Figure 13, line 1, cohort I: green, cohort II: orange). This was observed for all tissue types within the cohorts. Adjacent normal tissue, pancreatic normal tissue, and ACC of the two cohorts are mixed, *i.e.* there was no batch effect (Figure 13).

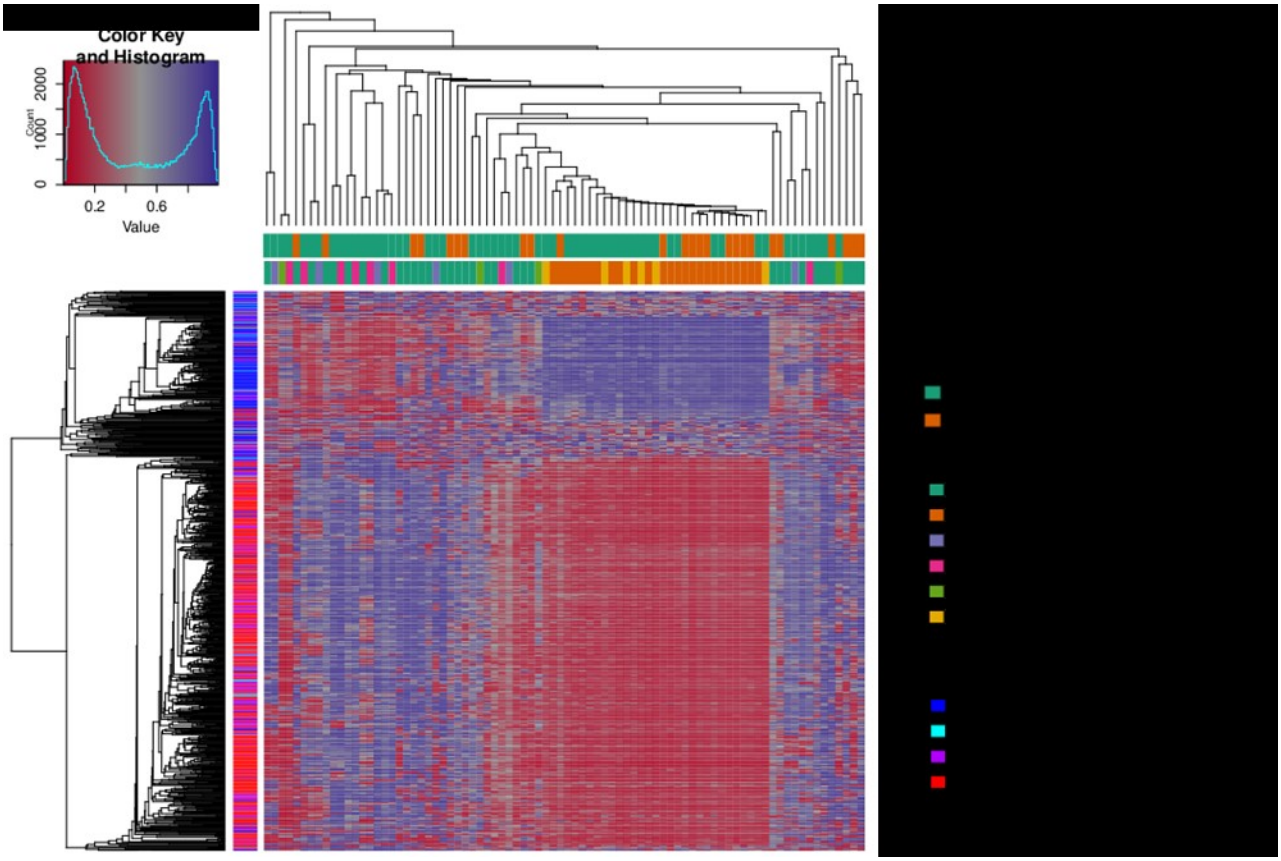


Figure 13 "Correlation-based clustering of all single CpG-sites (agglomeration strategy: average) with annotations of cohort (lane 1) and tissue type (lane 2)."[1]

4.4.2 Methylation patterns of ACC suggests acinar cells as the cell of origin

The knowledge of the cell of origin is crucial in order to use the appropriate normal control for methylation analysis. As the cell of origin is unknown for ACC, the methylation pattern of ACC was first compared to sorted healthy pancreatic cells, *i.e.* duct cells, acinar cells, endocrine α-cells, and endocrine β-cells using the algorithm by Houseman, et al. [238].

Figure 14a showed that the highest enriched cell type for ACC cohort I and cohort II were acinar cells (35.4%), suggesting this tumor entity arises from acinar cells. It also showed enrichment (compared to normal tissue) for endocrine α-cells which was due to the endocrine component in MACNEC (enrichment 15.0% versus 2.6%, Figure 14b). Endocrine β-cells and duct cells were depleted in ACC compared to normal tissue (6.2% versus 10.2%, and 15.3% versus 26.3%, respectively). In addition to pancreatic cell types, infiltrations of previously published blood cell types[239] were calculated. These were generally quite low in ACC (21.7%). The most common blood cell type methylation pattern was that of CD8+ T cells (4.9%) (Figure 14a, b).

In contrast to ACC, PNET showed the highest enrichment for endocrine α-cells (38.8%). They were enriched for endocrine β-cells (16.0%), but depleted for acinar (15.2%) and duct cells (3.3%), suggesting that PNET arose from endocrine α- and β-cells. These tumors, too, showed low infiltration of immune cells (24.6%), with the most common ones natural killer cells (6.7%) and monocytes (6.8%) (Figure 14a, b).

Duct cell methylation patterns were most common in PDAC (23.5%), with endocrine α-cells also enriched (6.9%, due to some outliers, see Figure 14, same outliers in Figure 18). However, enrichment for acinar cells was still quite strong (17.9%), whereas endocrine β-cell methylation patterns were scarce (1.6%). Infiltrations of immune cells were generally much higher in PDAC (46.3%), with the most common blood cell types being monocytes (11.3%) and natural killer cells (18.6%) (Figure 14a, c).

To sum up, pancreatic cancers seem to be combatted by the immune system by different proportions of blood cell types. In addition, they generally harbored the methylomes from the cells they resembled phenotypically. Therefore, pancreatic normal tissues were subsequently used as a normal control for ACC, as they mostly contain acinar cells (~90%).

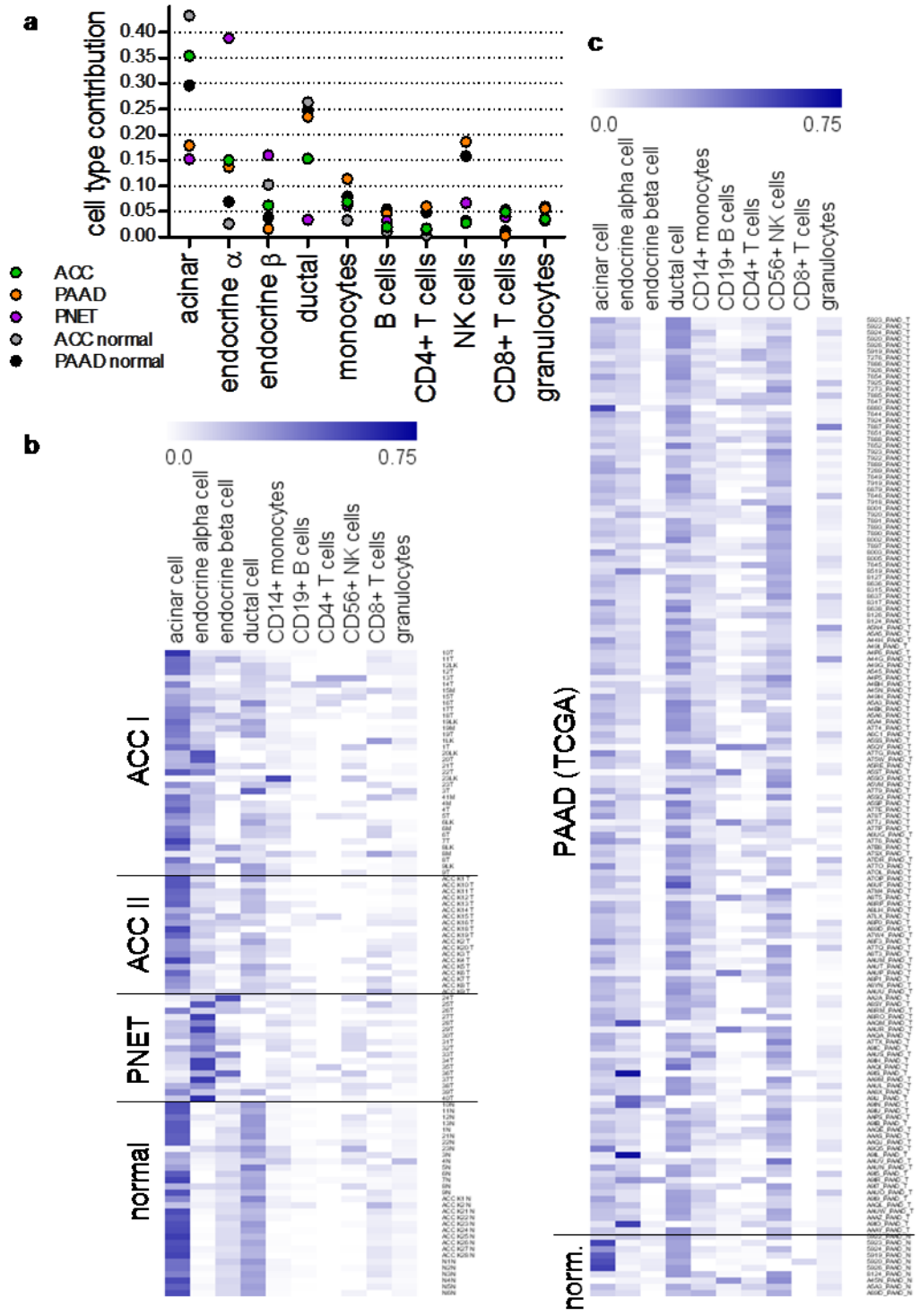


Figure 14 Cell type contributions of sorted pancreatic and blood cells to pancreatic tumors. a. percentages of the methylomes contributing to ACC, PAAD, PNET, ACC normal (= normal tissue from ACC cohorts), PAAD normal (= normal tissue from PAAD cohort) **b.** cell type contributions per sample from 450K data generated for the present study (each row represents one tumor). **c.** cell type contributions per sample from PAAD data set. PAAD = PDAC dataset from TCGA. Dr. Yassen Assenov wrote the script for the analysis.

4.4.3 Global methylation patterns of ACC

In a next step, the global methylation of ACC was compared to normal pancreatic tissue, employing unsupervised clustering. Primary tumors did not form a separate cluster in cohort I. However, normal tissue samples were next to each other, while primary tumors were split to the left and right of the normal tissues (Figure 15). MACNEC did not form a separate cluster, but intermingled with pure ACC, suggesting that their molecular patterns were similar. The

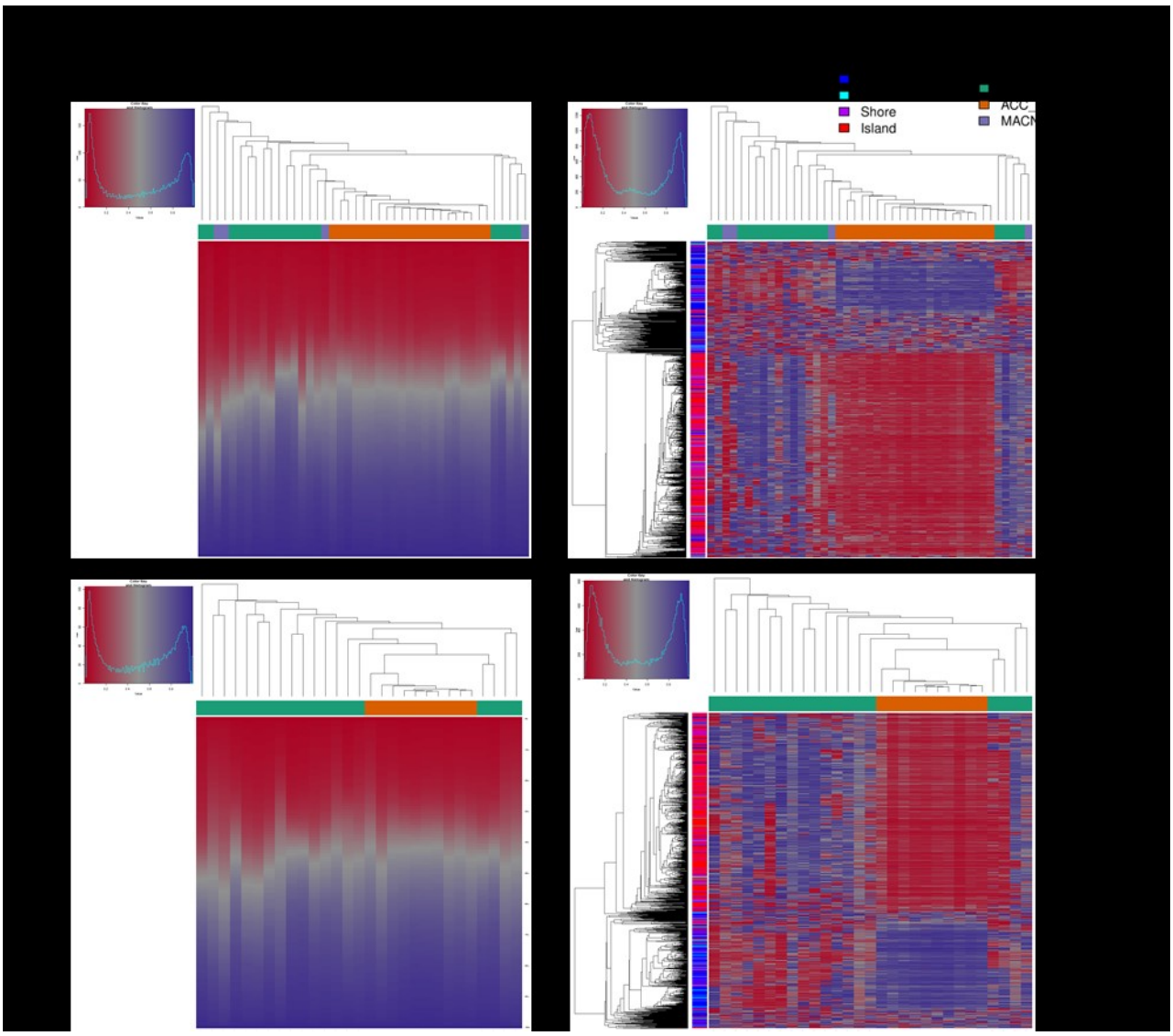


Figure 15 Correlation-based clustering of ACC samples. Tumors and normal tissues were clustered based on either all sites (left panel) or the 1000 most differential methylated sites (right panel) for cohort I (top panel) and cohort II (lower panel). The annotation on top of each clustering depicts the pure ACC tumors (green), MACNEC (purple) or normal tissues (orange). The annotation on the left of the right panel corresponds to the location of the differential CpG sites: open sea (dark blue), shelf (turquoise), shore (purple), and island (red).

observations were very similar for clustering based on all sites (Figure 15 left panel) and clustering based on the 1,000 most differentially methylated sites (DMS) (Figure 15 right panel), suggesting that the differential methylation between normal tissues and tumors was were stable. This was confirmed by cohort II (Figure 15 lower panel).

“Principal Component analysis (PCA) of all CpG sites of cohort I and cohort II”[1] showed that “normal pancreatic tissues clustered very closely together while the tumors were distinct from normal tissues and formed a wide-spread cluster, *i.e.* showed a different global methylation pattern”[1] (Figure 16 green dots versus the remaining dots). Interestingly, MACNEC did not form a separate cluster but intermingled with pure ACC (Figure 16 blue dots versus black and brown dots). In addition, cohort I and cohort II tumors again were mixed (Figure 16 black versus brown dots).

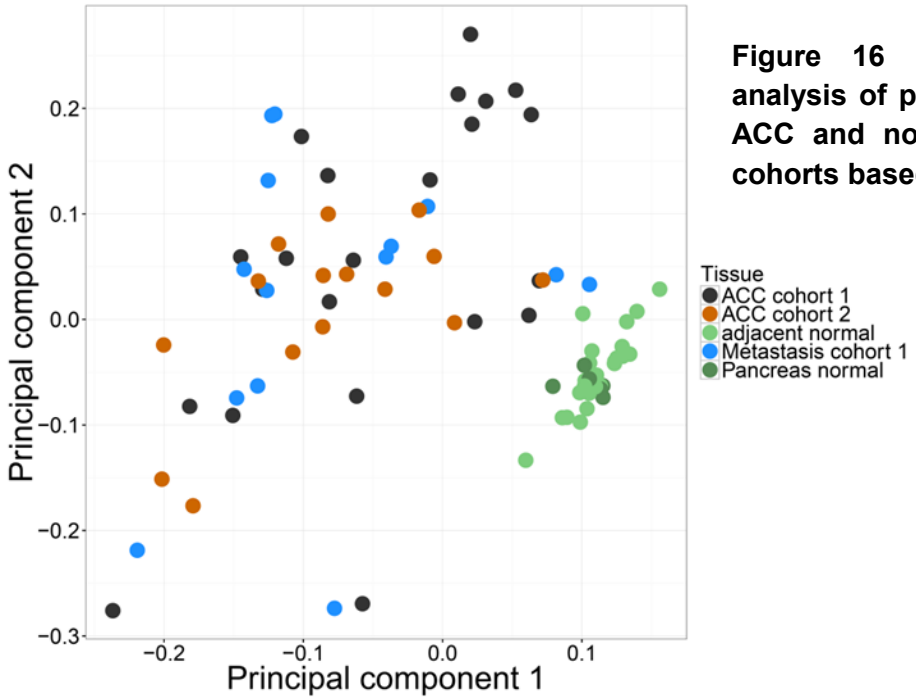


Figure 16 Principal component analysis of primary and metastatic ACC and normal tissues of both cohorts based on all CpG sites.

4.4.4 Distinction of ACC from other pancreatic cancers based on the methylomes

Further, it was investigated whether ACC form a separate subgroup within pancreatic cancers on the molecular level. To that end, a DNA methylation phylogenetic tree was generated, including ACC, PNET, PDAC, and sorted pancreatic cells (acinar cells, duct cells, endocrine α-cells, and endocrine β-cells). This revealed that the majority of tumors form specific branches according to their histopathological classification (Figure 17), *i.e.* one branch for ACC, PNET

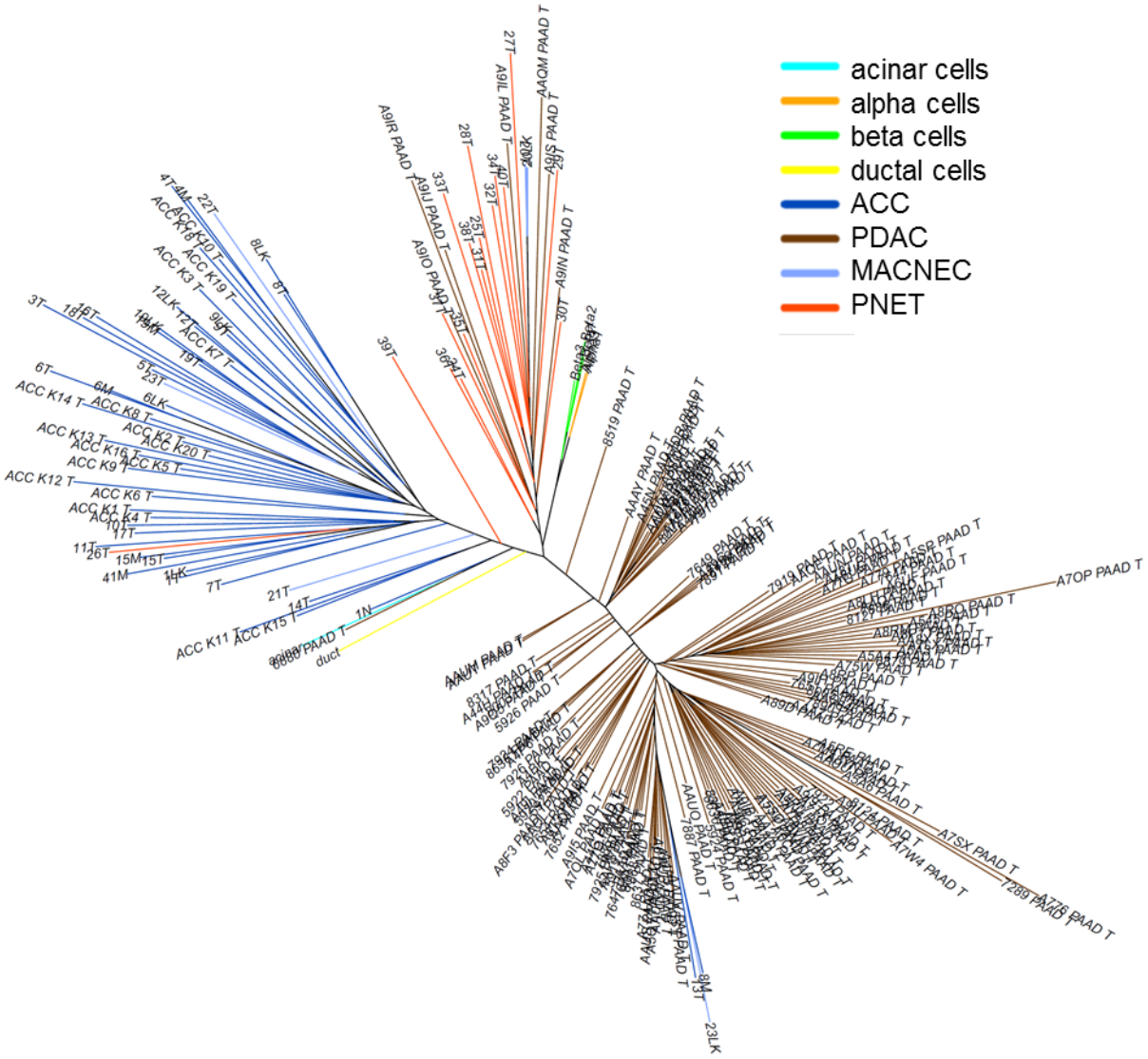


Figure 17 DNA methylation phylogenetic tree of pancreatic cancers. Dr. Yassen Assenov wrote the script for the analysis.

and PDAC each. A few outliers were detected especially in the PNET branch, however these were the same tumors that stood out in the PCA (refer to Figure 18), suggesting that these tumors were misclassified and that methylation analysis might support diagnosing mixed tumors correctly. Three ACC clustering with PDAC were actually the same three tumors with low purity (Figure 12b), suggesting that the tumor content in these samples was too low to correctly map them. Both α- and β- cells were within the branch of PNET, whereas acinar and duct cells were between the branches of ACC and PNET, confirming observations from Figure 14. In addition, MACNEC were distributed throughout the branch of ACC, supporting the idea that they form one tumor entity with pure ACC (refer to Figure 16).

In addition to the phylogenetic tree, PCA revealed that (i) each tumor entity forms their own cluster, (ii) ACC were closest to acinar cells, (iii) PDAC were closest to duct cells, and (iv) PNET were closest to endocrine cells (Figure 18). A few outliers can be detected; however they are in line with previously reported outliers.

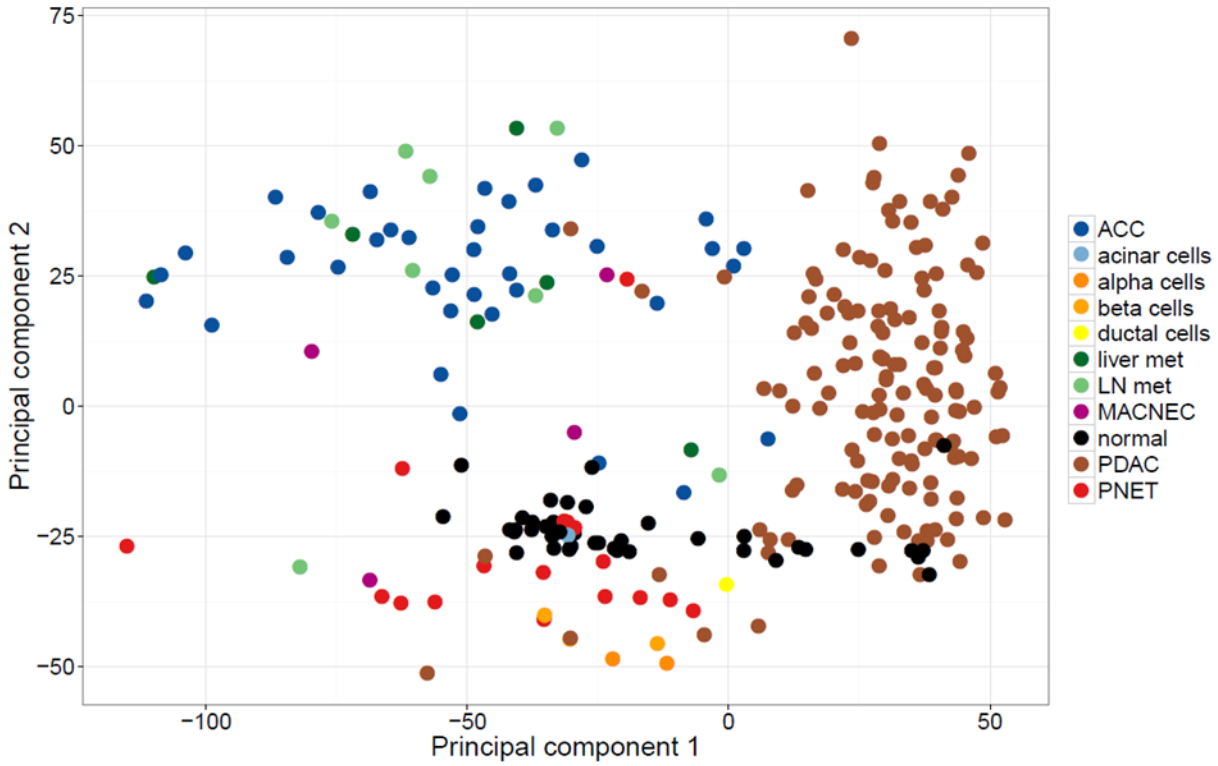


Figure 18 Principal component analysis of ACC, PDAC, PNET, and mixed and sorted normal pancreatic cells. Analysis was based on all sites. LN met: lymph node metastases from ACC, liver met: liver metastases from ACC.

4.4.5 ACC harbored many differentially methylated sites and regions

In addition, differential methylation analysis between different groups was performed (for memberships of each sample to groups refer to Supplementary Table 2). In addition to differential methylation calling of single CpG sites, biological relevant regions were defined, namely CpG islands, promoters, the overlap of CpG islands with promoters (from now on termed promoter CpG islands), gene bodies, enhancers, and lncRNAs. The number of regions, length per region, number of sites per region and the site density between these regions differ. This has to be kept in mind when comparing the occurrences of differential methylation. While promoters have by definition a constant size of 2,000 bp, most CpG islands were between 200 and 1,000 bp, most enhancers between 200 and 10,000 bp, most genes between 200 and

50,000 bp, most lncRNAs promoters between 2300 and 2500, and most promoter CpG islands between 200 and 1,000 bp (Supplementary Figure 1). The number of sites per region was highest in CpG islands, followed by promoters and promoter CpG islands, and lowest in enhancers (Supplementary Figure 2). The distribution of sites within each region was highest at the 5' and 3' borders in the case of CpG islands, enhancers and promoter CpG islands, while

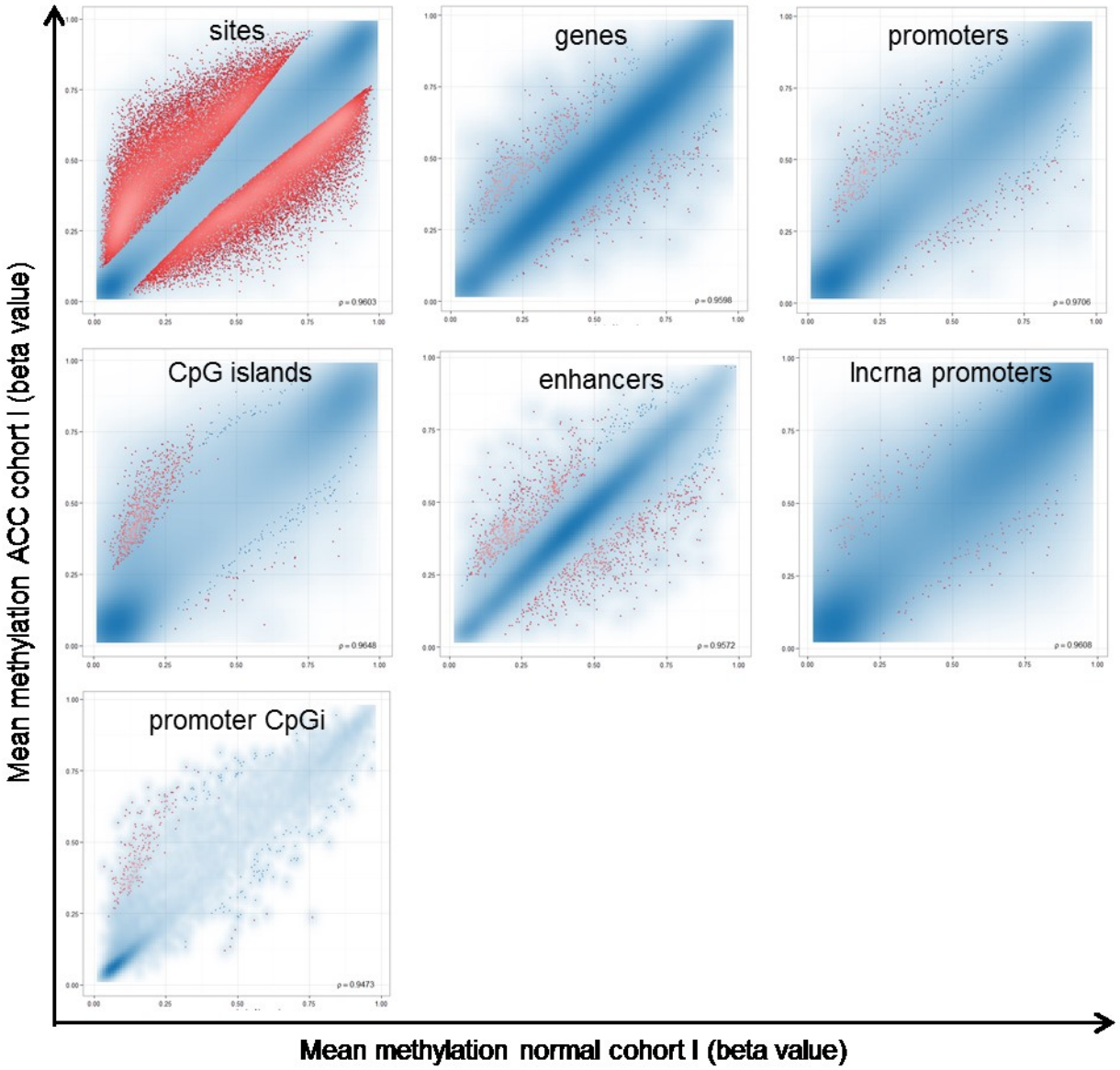


Figure 19 Scatterplots of cohort I showed massive aberrations in DNA methylation. Scatterplots of the mean methylation of all normal tissues (x axis) versus all tumors (y axis) in cohort I, at the site level and all region annotations. Red dots were sites or regions above the selected rank cutoff.

gene bodies show a peak at the 5' border, and lncRNAs and promoters peak within the region (Supplementary Figure 3).

To evaluate the use of normal tissues, differential methylation between adjacent normal tissues and normal pancreatic tissues from healthy individuals were compared. Not many DMRs were identified, with the exception of lncRNAs and some hypomethylated promoters (Supplementary Figure 4 and Supplementary Figure 5). Therefore, comparison of tumors to normal tissues was subsequently performed using all normal tissues, *i.e.* adjacent normal tissues and normal pancreatic tissues combined. This revealed differential methylation at 44193 CpG sites in cohort I. Of these, 22602 sites were hypermethylated (Figure 19, Figure 20a). Cohort II had 26959 differentially methylated CpG sites, including 6925 hypermethylated sites (Supplementary Figure 6 and Supplementary Figure 7a). About 43% of sites, *i.e.* 18852 were

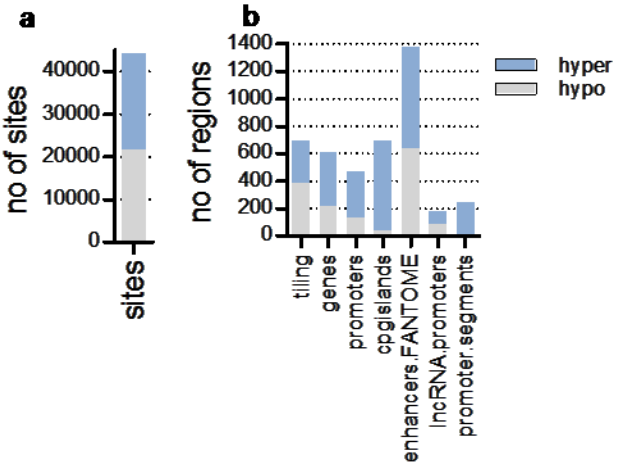


Figure 20 Number of sites in cohort I showed massive aberrations in DNA methylation. Number of sites per region that were hypermethylated (blue) or hypomethylated (grey) in the tumors.

confirmed in cohort II (Figure 21a, b) and about half of these CpG sites (8436) were hypermethylated sites (Figure 21b). "After mapping the CpG sites to either"[1] CpG islands, enhancers, genes, lncRNA promoters, promoters or promoter CpG islands (for region definitions refer to methods), a total number of 690, 1373, 608, 181, 466, and 245 DMRs were detected in cohort I, respectively (Figure 19, Figure 20b). Of these, there were 657, 737, 388, 96, 332, and 237 hypermethylated regions, respectively. In cohort II a total number of 543, 1268, 440, 230, 534, and 390 DMRs were found, respectively. Of these 491, 476, 189, 78, 2210, and 363 regions were hypermethylated, respectively (Supplementary Figure 6, Supplementary Figure 7b).

The validation rate of these regions were 60%, 56%, 44%, 58%, 53%, and 77%, respectively (Figure 21a) and therefore 411, 766, 270, 105, 245, and 189 validated regions were obtained, respectively (Figure 21c). Of these regions the majority displayed hypermethylation (389 CpG islands, 166 genes, 51 lncRNA promoters, 166 promoters, and 185 promoter CpG islands) with the exception of enhancers where only about half of the regions (374) were hypermethylated. Promoter segments were even almost exclusively hypermethylated (Figure 21c).

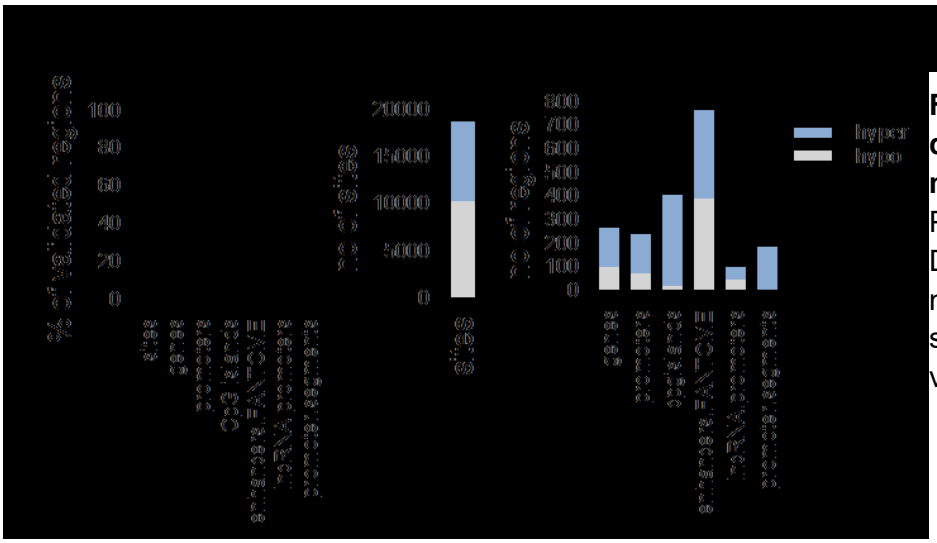


Figure 21 Validated differential methylation in ACC a. Percentage of validated DMS and DMR **b.** total number of validated sites **c.** total number of validated regions.

“A total number of 364 genes that show differential methylation either at their“[1] promoters, CpG islands promoters, and/or gene bodies “were identified in cohort I and validated in cohort II”[1] (Supplementary Table 3). Some of these (*TWIST1*, *HIST1H1B*, *FOXD2*, *FOXD4*, *SOX17*, and *KCNQ1*) were confirmed by MassARRAY (Figure 22 and Supplementary Figure 10).

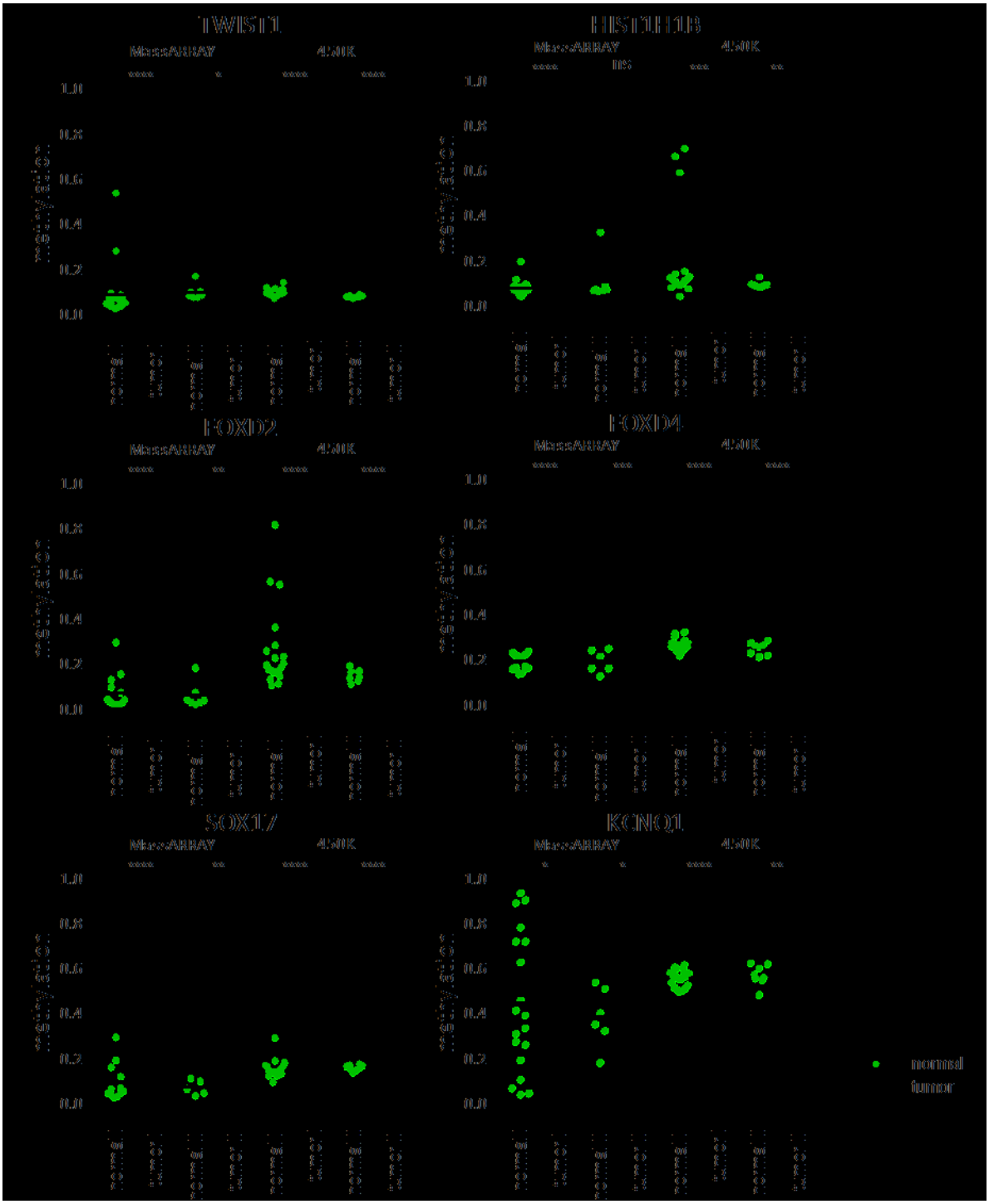


Figure 22 Validation of 450K methylation by MassARRAY technology. Average methylation per region was depicted for each tumor. Daniel van der Duin performed MassARRAY under my supervision.

Gene set enrichment analysis (Figure 23) using the DAVID tool[240] revealed that these 364 genes were enriched for cell adhesion pathways (hemophilic cell adhesion and cell-cell adhesion) and embryonic development pathways (including embryonic morphogenesis, pattern specification process, regionalization, and anterior/posterior pattern formation).

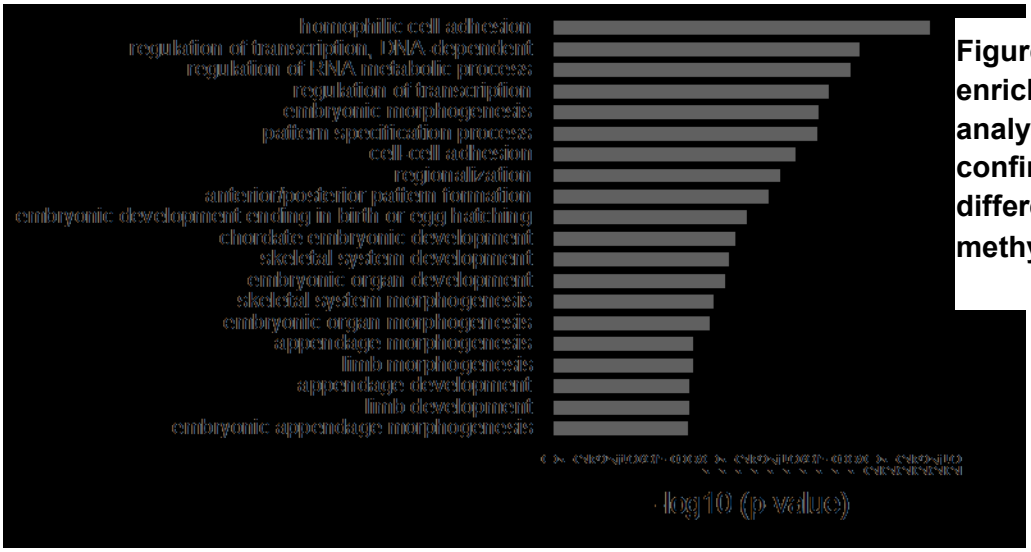


Figure 23 Gene set enrichment analysis of confirmed differentially methylated genes

Enrichment analysis of transcription factor binding sites at DMRs using the HOMER tool[241] revealed that many of the transcription factors playing a role during pancreatogenesis (*e.g. FOXA1*, *FOXA2*, *NR5A2*, and *NKX6-1*, also refer to Figure 5) were enriched at the DMS in both

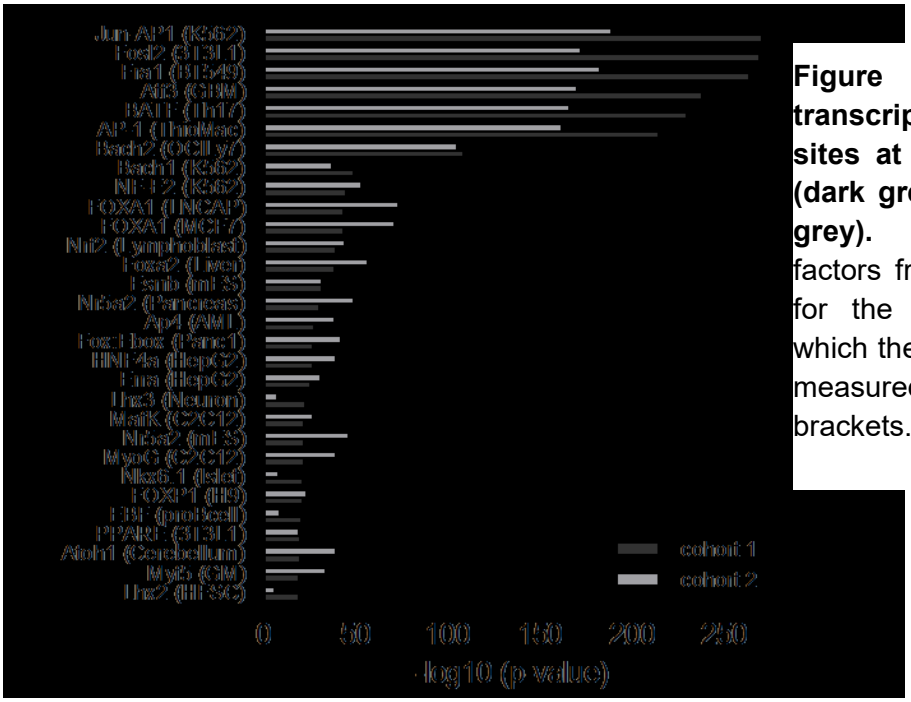


Figure 24 Enrichment of transcription factor binding sites at DMS in ACC cohort I (dark grey) and cohort II (light grey). Known transcription factors from HOMER were used for the analysis. Cell type in which the transcription factor was measured in is written in brackets.

cohorts (Figure 24). Thus, differential methylation occurred at regions which play a role during the development of the pancreas.

The DMS were depleted for the chromatin states promoters (active and weak), and transcription (transcription elongation and weak transcription). They were enriched for enhancers (weak and strong), poised promoters, insulators, and repressed states. This was observed in both cohorts (Figure 25a). DMS were enriched for the histone mark H3K4me1 and depleted for histone marks H3K4me3 and H3K27ac in both cohorts (Figure 25b). Thus, differential methylation preferentially occurred outside of genes and promoters, within non-coding regions of the genome.

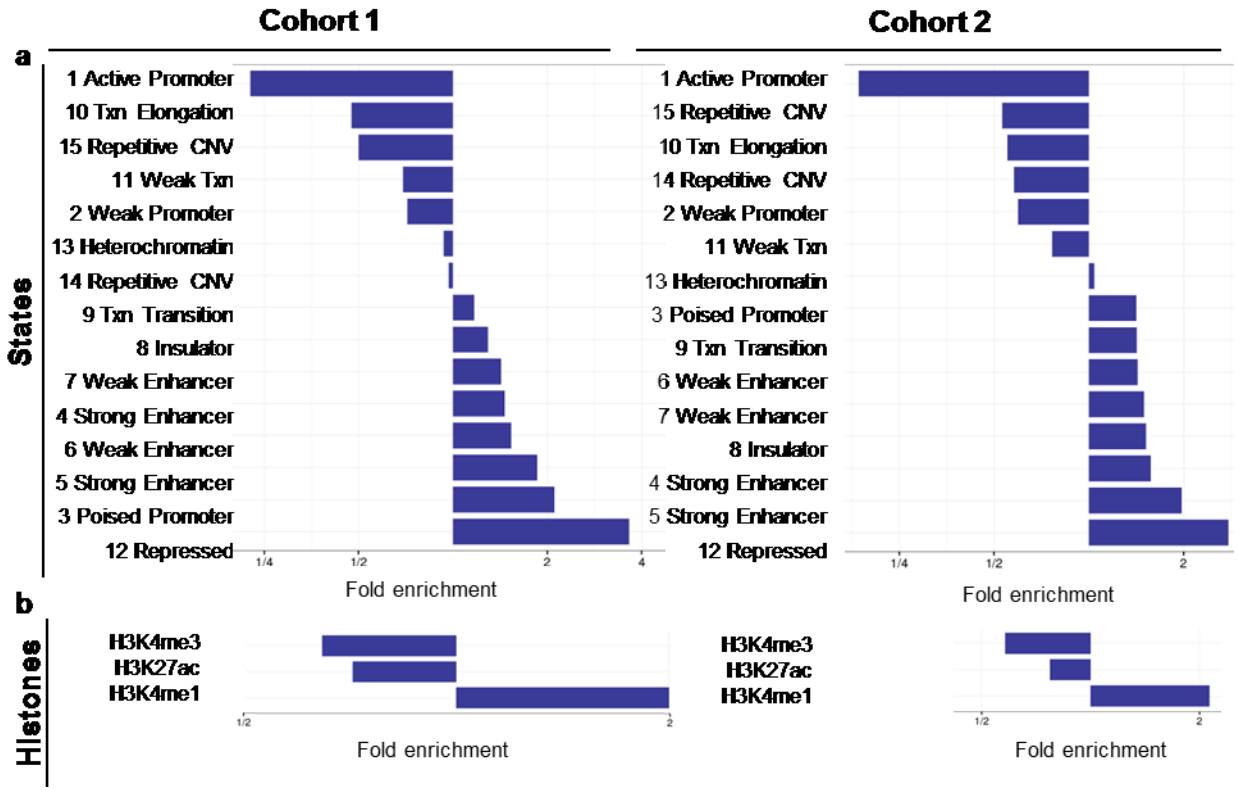


Figure 25 Enrichment of chromatin states and histone marks at DMS in ACC cohort I (left panel) and cohort II (right panel). a. Enrichment of DMS at chromatin states from ENCODE‘s HMEC **b.** Enrichment of DMS at histone marks from ENCODE‘s PANC1 cells.

4.4.6 Aberrations in methylation predominantly occurred during tumorigenesis und not during metastases formation

Cohort I contained besides primary tumors also 12 metastases. A paired analysis of primary tumors with their metastases was performed to discover changes in methylation which occur

during metastases formation (Figure 26). Roughly 10,000 DMS were identified, mainly hypomethylation (9351 sites, Figure 27a). However, when mapping these sites to regions as defined above, only a few of these specific regions were identified (Figure 27b), *i.e.* 9 genes, 31 promoters, 3 CpG islands, 1 enhancer, 15 lncRNAs and 3 promoter segments, suggesting that metastases exhibited a very similar epigenome as their primary tumors.

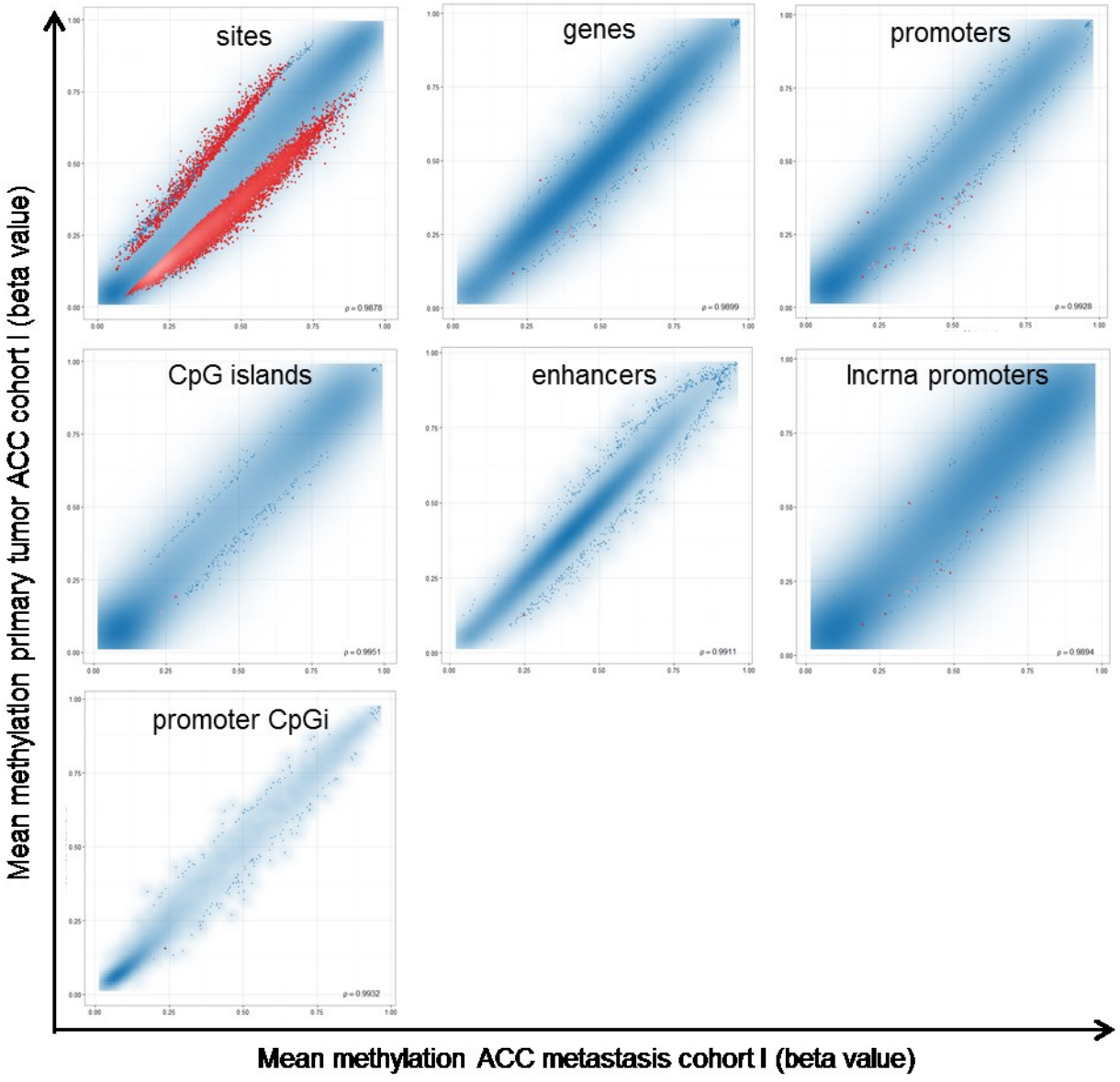


Figure 26 Metastases showed few aberrations in DNA methylation compared to their primary tumors. a. Scatterplots of the paired analysis of the mean methylation of all metastases (x axis) versus all primary tumors (y axis) in cohort I at the site level and all region annotations used. Red dots are sites or regions above the selected rank cutoff.

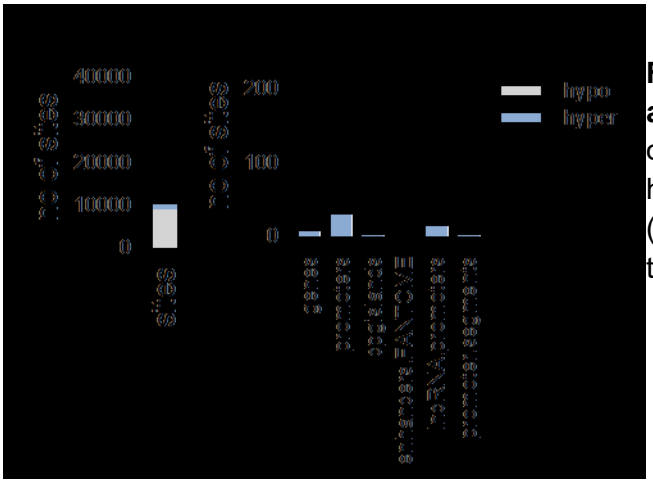


Figure 27 ACC metastases showed few aberrations in DNA methylation. a. Number of sites and **b.** number of regions that were hypermethylated (blue) or hypomethylated (grey) in the metastases compared to primary tumors.

4.4.7 The protocadherin cluster is hypermethylated in ACC

As one of the enriched pathways was cell adhesion, the protocadherin (*PCDH*) cluster was more closely examined. This cluster contains 52 genes, which separate into three subclusters

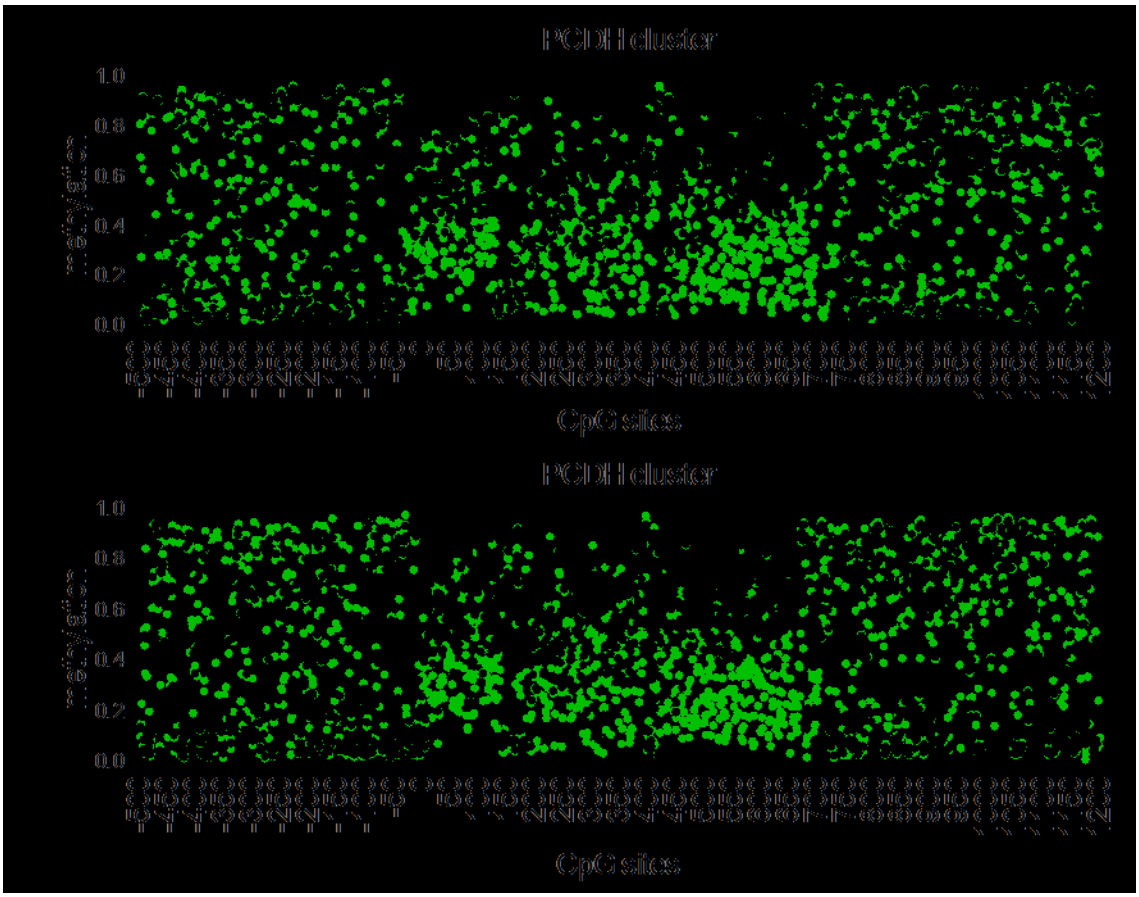


Figure 28 DNA methylation of the *PCDH* cluster of **a.** cohort I and **b.** cohort II. Tumor = green, normal = black. Mean methylation of each CpG site.

(14 in α-, 22 in β-, and 22 in *γ-PCDH*) encoded all in one cluster on chromosome 5q31.3. *PCDH-β* (*PCDHB*) genes are encoded by one exon, while *PCDH-α* (*PCDHA*) and *PCDH-γ* (*PCDHG*) genes are encoded by one unique exon per gene and three constant exons shared by all genes within the subclusters[248]. Each gene is controlled by its own promoter and nearly all members have one CpG island in its promoter and one intragenic CpG island[249,250]. The vast majority of studies on *PCDH* focused on the role of these genes in neuronal context. It has been shown that expression of *PCDH* is controlled by differential promoter activation, *i.e.* each cell expresses a different number of different *PCDH* members. This gives rise to a total number of about 3×10^{10} different combinations[251]. As PCDH members interact in hemophilic tetramers[252], this suggests that neuronal cells use PCDH expression on their cell surface as a cell recognition mechanism (reviewed in Chen and Maniatis [248]).

In other tissues the role of *PCDH* is not investigated, however a few studies revealed that this cluster is hypermethylated in some cancer entities[253-257]. In ACC, hypermethylation spread specifically over this 700 kb *PCDH* region but not to neighboring areas. This was observed for both cohorts (Figure 28 a, b). When comparing the hypermethylation of the *PCDH* cluster to other pancreatic cancers, it revealed that the hypermethylation of *PCDH* was strongest in ACC, while PNET did not show any hypermethylation, and PDAC only showed a slight hypermethylation (Figure 29). To examine whether *PCDH* hypermethylation is a general feature of tumors, 14 tumor entities from TCGA where 450K data was available were screened. This revealed that the majority of cancers (12 out of 14) were hypermethylated at this locus, *e.g.* uterine corpus endometrial carcinoma (UCEC) was hypermethylated to a similar extent as ACC,

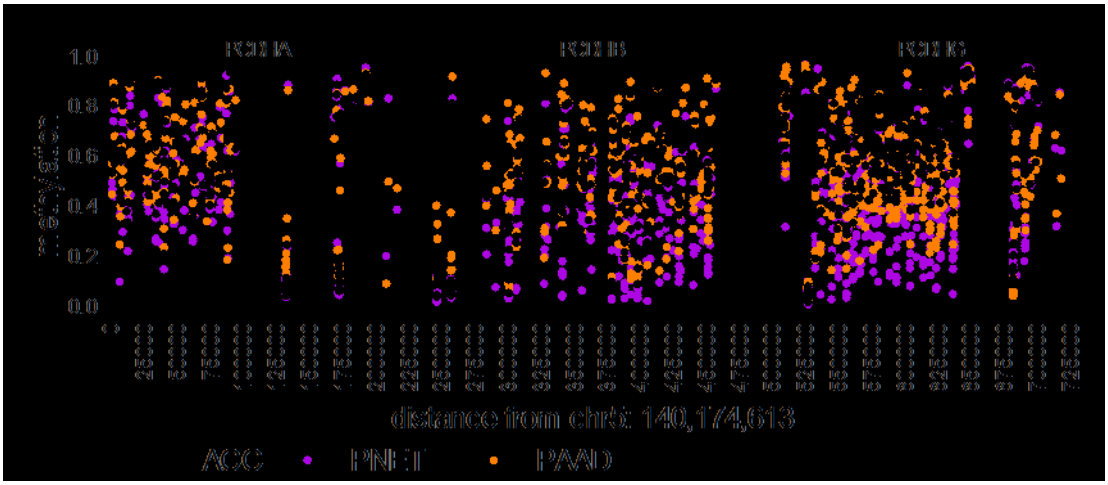


Figure 29 DNA methylation of the *PCDH* cluster in pancreatic cancers. ACC: black, PNET: purple, PAAD-TCGA: orange. Mean methylation of each CpG site.

while stomach adenocarcinoma (STAD) exhibited normal methylation (Figure 30) at this cluster. For a list of tumors hypermethylated at this region refer to Supplementary Table 4.

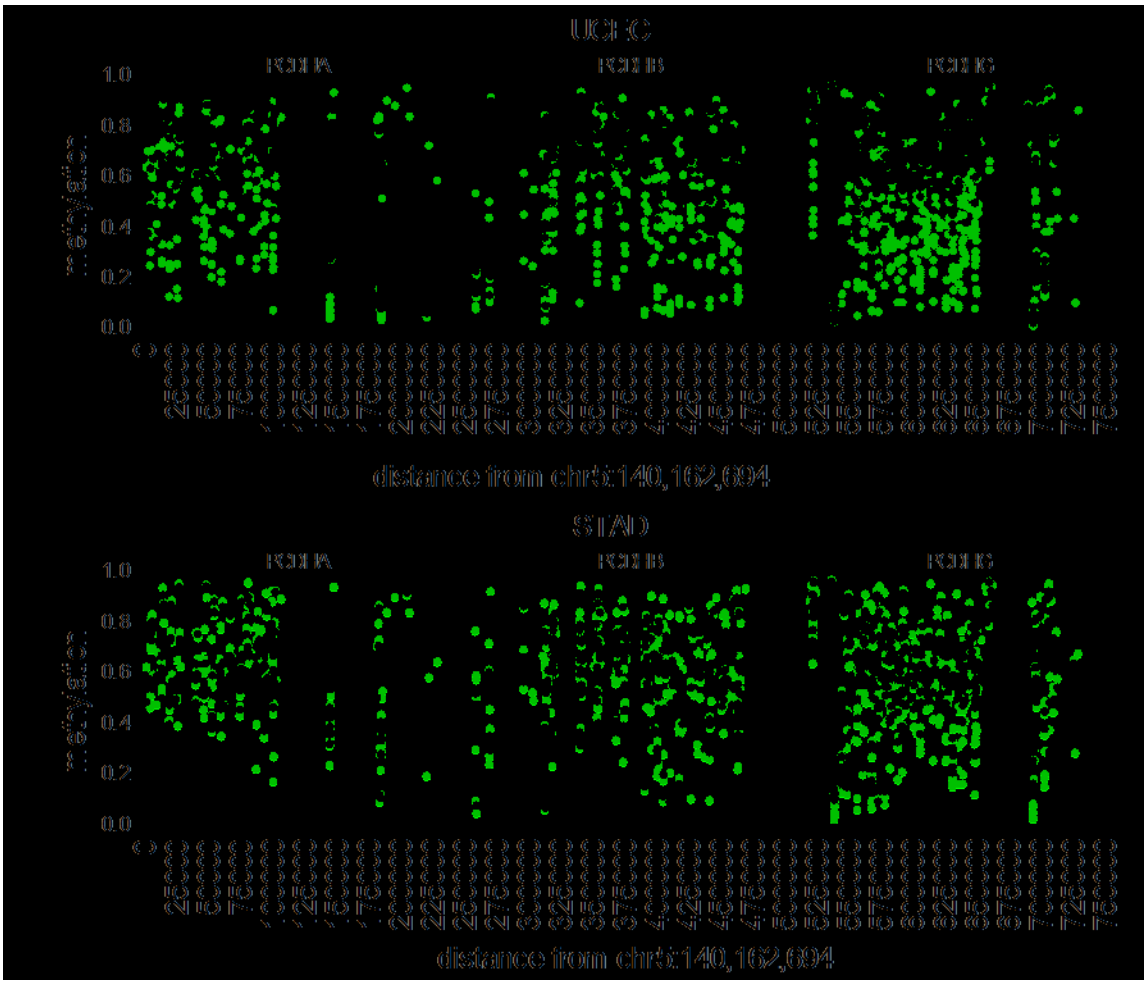


Figure 30 DNA methylation of the *PCDH* cluster of a. the hypermethylated tumor entity UCEC and **b.** the tumor entity STAD not hypermethylated at this cluster. Tumor = black, normal = green. Mean methylation of each CpG site. Source of 450K data: TCGA.

In a next step, it was evaluated whether the hypermethylation can be reversed by DAC treatment. To identify whether this was a likely event, the literature was screened for available 450K data on DAC treated cell lines. Published breast, colon and ovarian cancer cell lines[258] were identified and screened. A total of 38 out of 54 cell lines showed a significant reduction of methylation upon DAC treatment, revealing that DAC can demethylate this locus in the majority of cell lines. Figure 31 shows three examples (breast cancer cell line BT20, colon cancer cell line Colo320, and ovarian cancer cell line A2780) with a reduction of methylation upon DAC

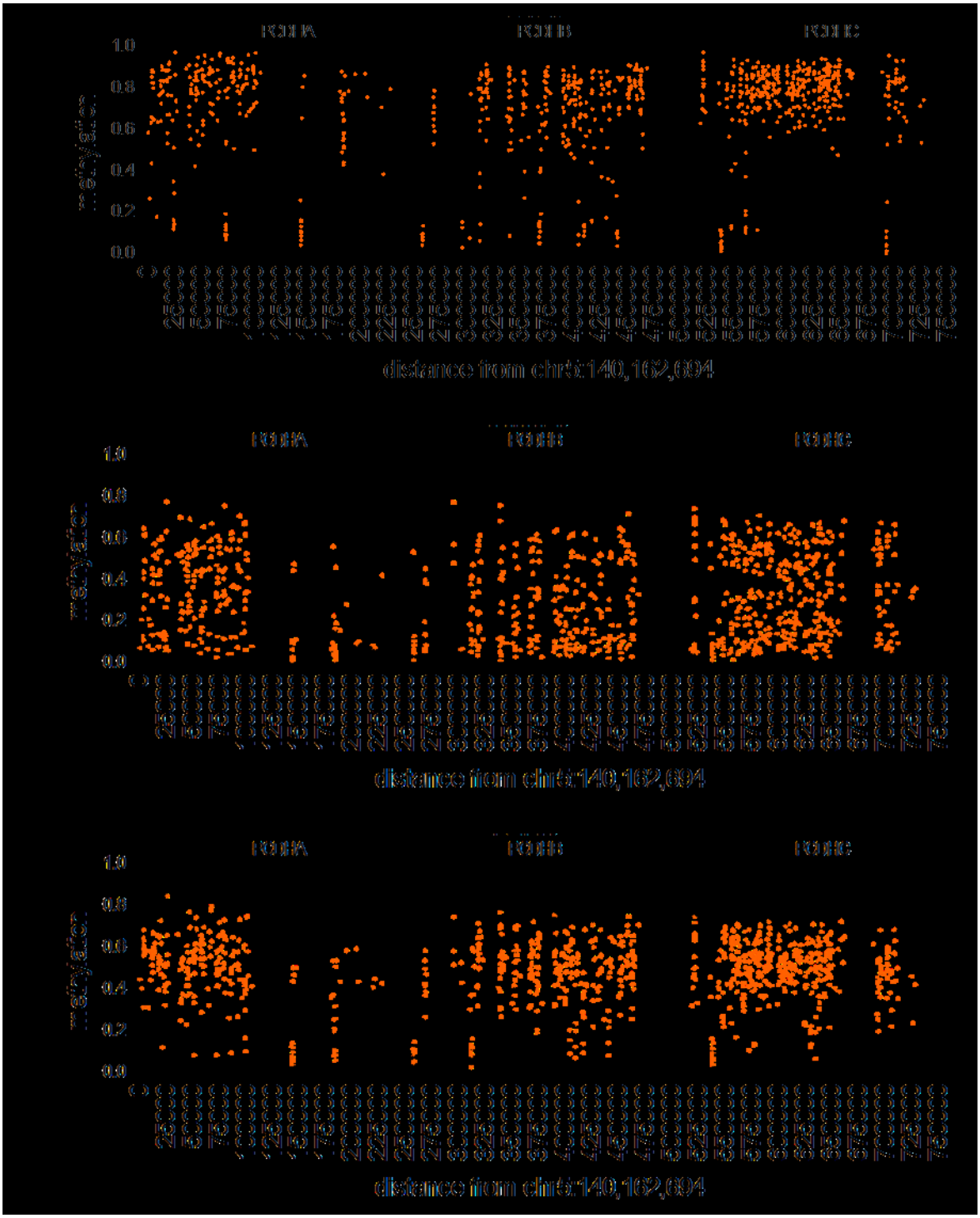


Figure 31 *PCDH* cluster demethylation by DAC treatment was cell line dependent. Cell line data by depicting three cell lines (BT20: breast cancer, Colo320: colon cancer, A2780: ovarian cancer) that changed methylation upon DAC treatment. Black: mock treatment, orange: DAC treatment.

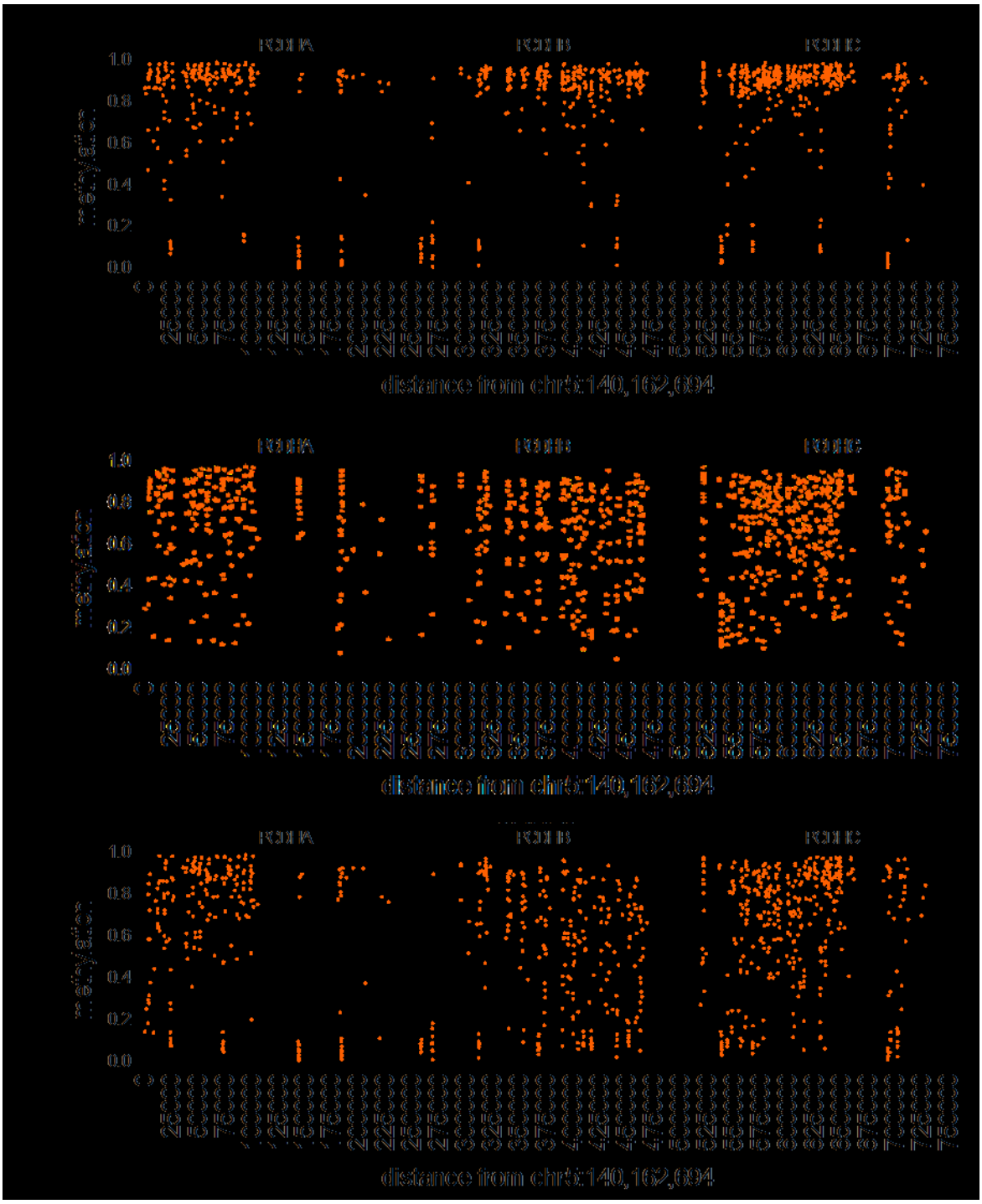


Figure 32 *PCDH* cluster demethylation by DAC treatment was cell line dependent. Cell line data depicting three cell lines (HCC1954: breast cancer, Caco2: colon cancer, OAW28: ovarian cancer) that did not change methylation upon DAC treatment. Black: mock treatment, orange: DAC treatment.

treatment. In contrast, Figure 32 shows three examples (breast cancer cell line HCC1954, colon cancer cell line Caco2, and ovarian cancer cell line OAW28) that did not change methylation. Significantly altered cell lines upon DAC treatment were summarized in Supplementary Table 5.

To investigate *PCDH* hypermethylation in ACC, an unpublished mouse model provided by Dr. med. Henrik Einwächter (Technische Universität München) was used (from now on termed T510). As in the human situation, mouse ACC showed a promoter hypermethylation throughout the different α-, β-, and γ-family members of *Pcdh*, as assessed by MassARRAY technology (Figure 33, green and black dots). Cell lines obtained from these mouse tumors maintained the methylation pattern of the primary tumors (Figure 33, red and black dots), thus were a suitable model for further experiments.

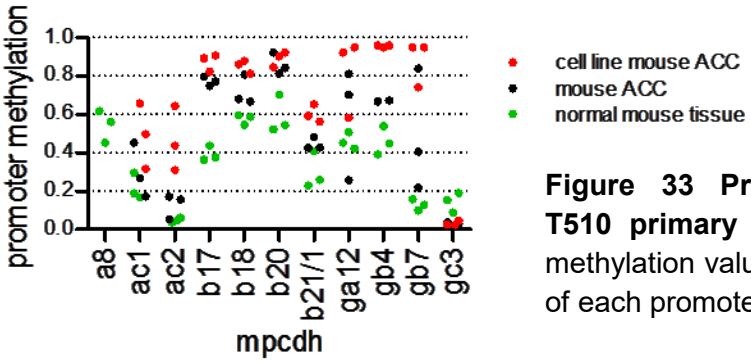


Figure 33 Promoter hypermethylation of T510 primary tumors and cell lines. Mean methylation value as measured by MassARRAY of each promoter was depicted.

To test whether the methylation level of the *Pcdh* cluster can be altered in mouse ACC, three T510 cell lines were subsequently treated with increasing concentrations of DAC (0.5, 1.0, and 2.0 µM). Two out of the three cell lines responded well to the treatment and revealed a decreased methylation in most of the *Pcdh* genes. These were the same cell lines where *Line1* methylation, a measurement for global DNA methylation, was decreased (Figure 34a). This suggests that the third cell line should be treated with a higher dose of DAC to obtain global demethylation. DAC concentration did not change the degree of demethylation significantly from 0.5 µM to 2.0 µM (Figure 34a). Indeed statistical analysis revealed that the single treatment dosages did not lead to a significant effect, while when combining the treatment dosages, results were significant (Supplementary Table 6). RNA expression of these isoforms surprisingly decreased in the *Pcdhb*-members, while it increased in *Pcdhg*-members (Figure 34b and Supplementary Table 7), suggesting that the methylation of the *Pcdh* cluster is functional relevant. This has previously been demonstrated for *Pcdha* in mouse neuroblastoma cell

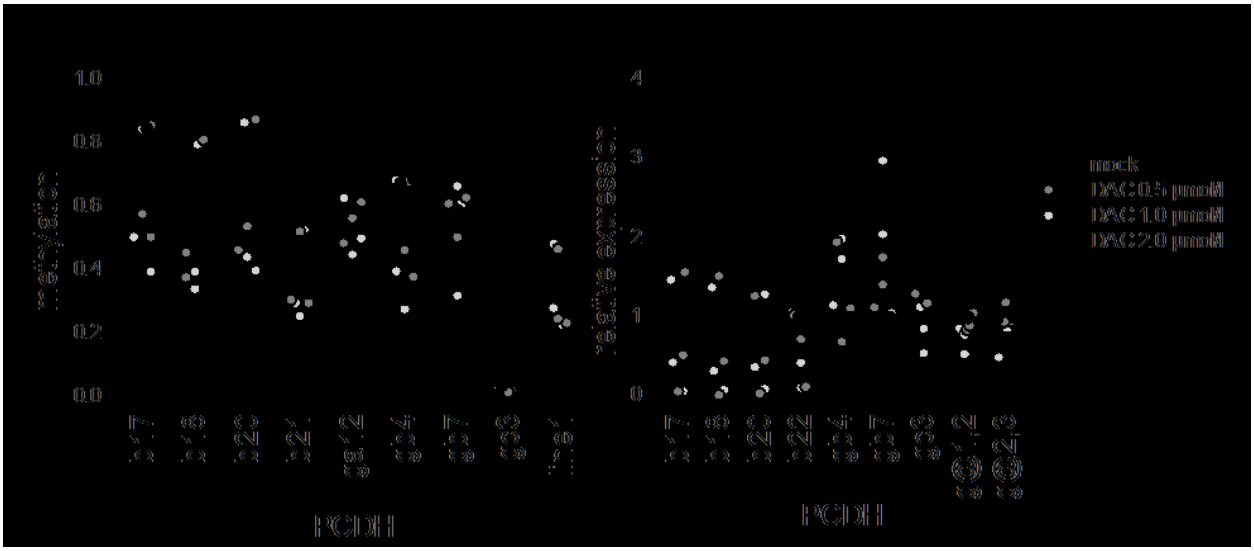


Figure 34 Influence of DAC treatment in T510 cell lines on *Pcdh* methylation and expression. a. Promoter methylation of *Pcdh* measured by MassARRAY **b.** relative expression of *Pcdh* genes measured by qPCR. g@ refers to the common exons of *Pcdhg* members.

lines[259]. *Pcdha* genes were not expressed in the mouse ACC T510 cell lines and were therefore not included.

To investigate the functional impact of a loss of *Pcdh* genes, *Pcdhg* members were knocked down using two siRNAs targeting the common exons, which should lead to a downregulation of all *Pcdhg* genes. Up to a concentration of 10 nM siRNA, cell viability was not affected by siRNA transfection (Supplementary Figure 11). *Pcdhg* expression decreased after 24h and 48h, especially in the higher expressed variant *Pcdhgc3* and the constant exons (termed *Pcdhg@1,2* and *Pcdhg@2,3*) (Figure 35).

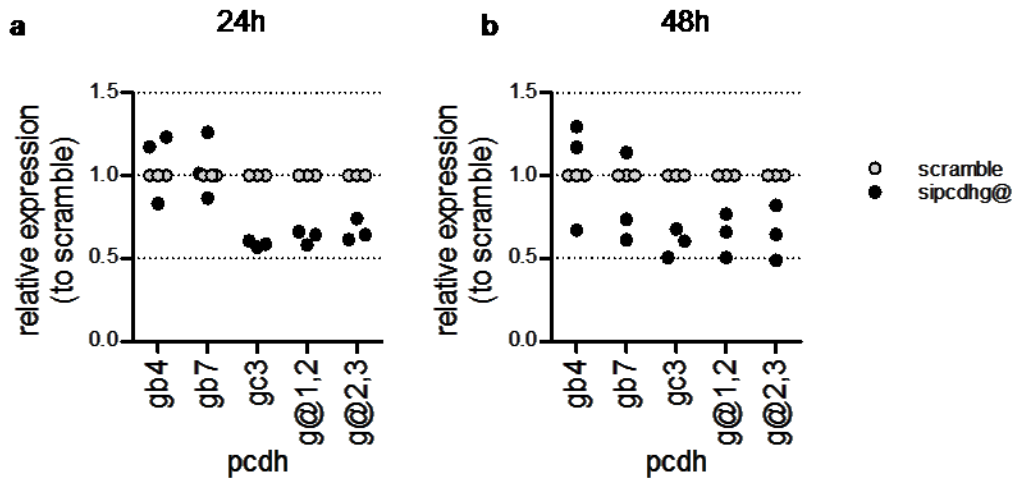


Figure 35 Expression *Pcdhg* genes after siRNA knockdown of *Pcdhg@*.

To reveal whether knockdown of *Pcdhg@* lead to a phenotype, a cell migration assay was performed. A slightly increased invasion of the cell lines was detected, although not significant (Figure 36). Due to the ambiguous results and weak effects the work on the *Pcdh* cluster in ACC was not further continued (refer to 5.2 and 5.6).

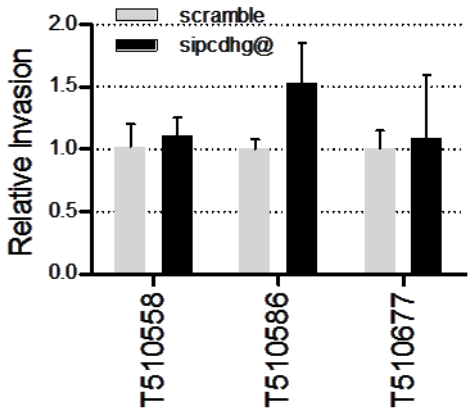


Figure 36 Invasion potential of T510 cell lines after siPcdhg@ knockdown. Numbers following T510 correspond to single cell line identifiers.

4.5 ACC harbor vastly instable genomes

“As the 450K array was build using the SNP array technologies, signal intensities can be used to generate”[1] maps of CNA[38]. The 450K array harbors many probes in gene promoters and bodies, therefore data for coding loci are in high resolution[260]. CNA were calculated for ACC employing the Bioconductor package conumee[242]. Plotting CNA for each tumor and each chromosome, nearly all tumors showed massive chromosomal aberrations in cohort I (Figure 37) and cohort II (Supplementary Figure 13). Some tumors showed many changes on the chromosomal arm level (*e.g.* 4T and 17T in Figure 37, and K4T and K18T in Supplementary Figure 13), while others had many smaller deletions and amplifications (*e.g.* 5T and 23T in Figure 37, and K8T and K12T in Supplementary Figure 13). Basically, only tumors with low tumor purity (see Figure 12b) showed a mainly stable genotype (*i.e.* 13T and 14T in Figure 37), suggesting that the tumor content in these samples was too low to detect chromosomal changes. The tumors with aberrations tended to have more deletions than amplifications and some broad range regions occurred in many tumors, *e.g.* loss of chromosome arm 1p, amplification of chromosome arm 1q.

GISTIC 2.0[243] was then used to identify commonly amplified and deleted regions amongst the different ACC and below the chromosomal arm level. To catch smaller regions, the bin size was set to five probes, meaning that at least five probes had to be within a defined region. The left panel of Figure 38 shows amplified and deleted regions of cohort I (top) and cohort II (bottom). Significantly enriched or deleted regions spanning less than one chromosomal arm level were depicted as peaks (middle and right panel of Figure 38, respectively). Q-values below 0.25 were considered as significant. With this approach it was possible to pinpoint aberrations, *e.g.* deletions to 1p36, 4q35, 9p21.3, 16p13.3, and 18q21.2, and amplifications to 1q42, 3q26.33, and 7p22.3. To investigate whether MACNEC exhibited different CNA than ACC, all MACNEC from both cohorts were combined and MACNEC-specific CNA were examined. This revealed that MACNEC were not significantly different to ACC, and regions identified in MACNEC were overlapping with regions altered in ACC (Figure 39). Using 450K data to calculate CNA in ACC thus revealed that (i) these tumors have highly instable genomes, (ii) many of the occurring CNA are shared amongst the tumors, and (iii) MACNEC harbor the same alterations as pure ACC.

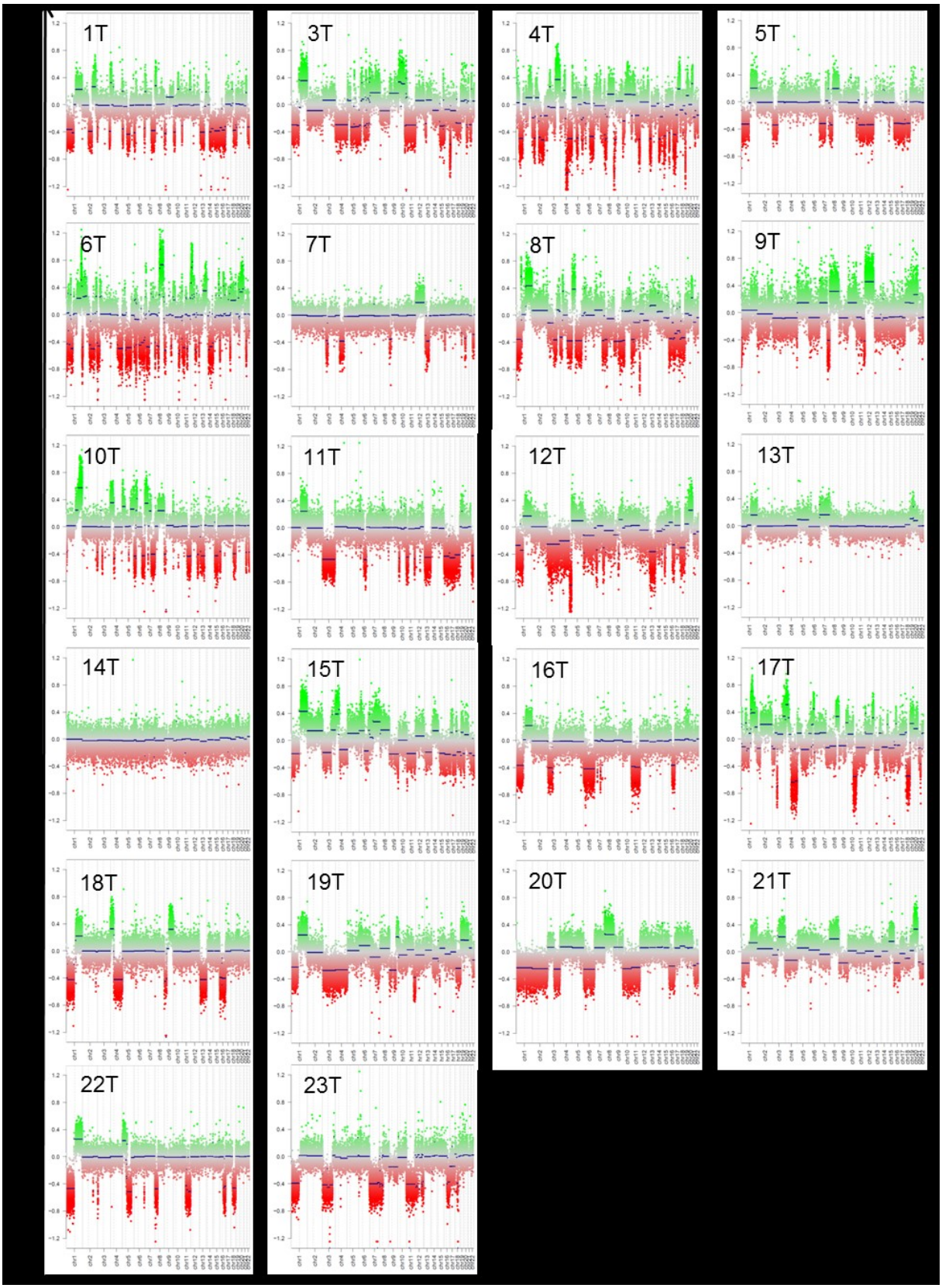

1T
3T
4T
5T
6T
7T
8T
9T
10T
11T
12T
13T
14T
15T
16T
17T
18T
19T
20T
21T
22T
23T

Figure 37 ACC from cohort I had highly instable genomes. Copy number profiles were calculated for each tumor and CNA were depicted for each tumor and each chromosome. Amplifications were depicted in green while deletions were depicted in red.

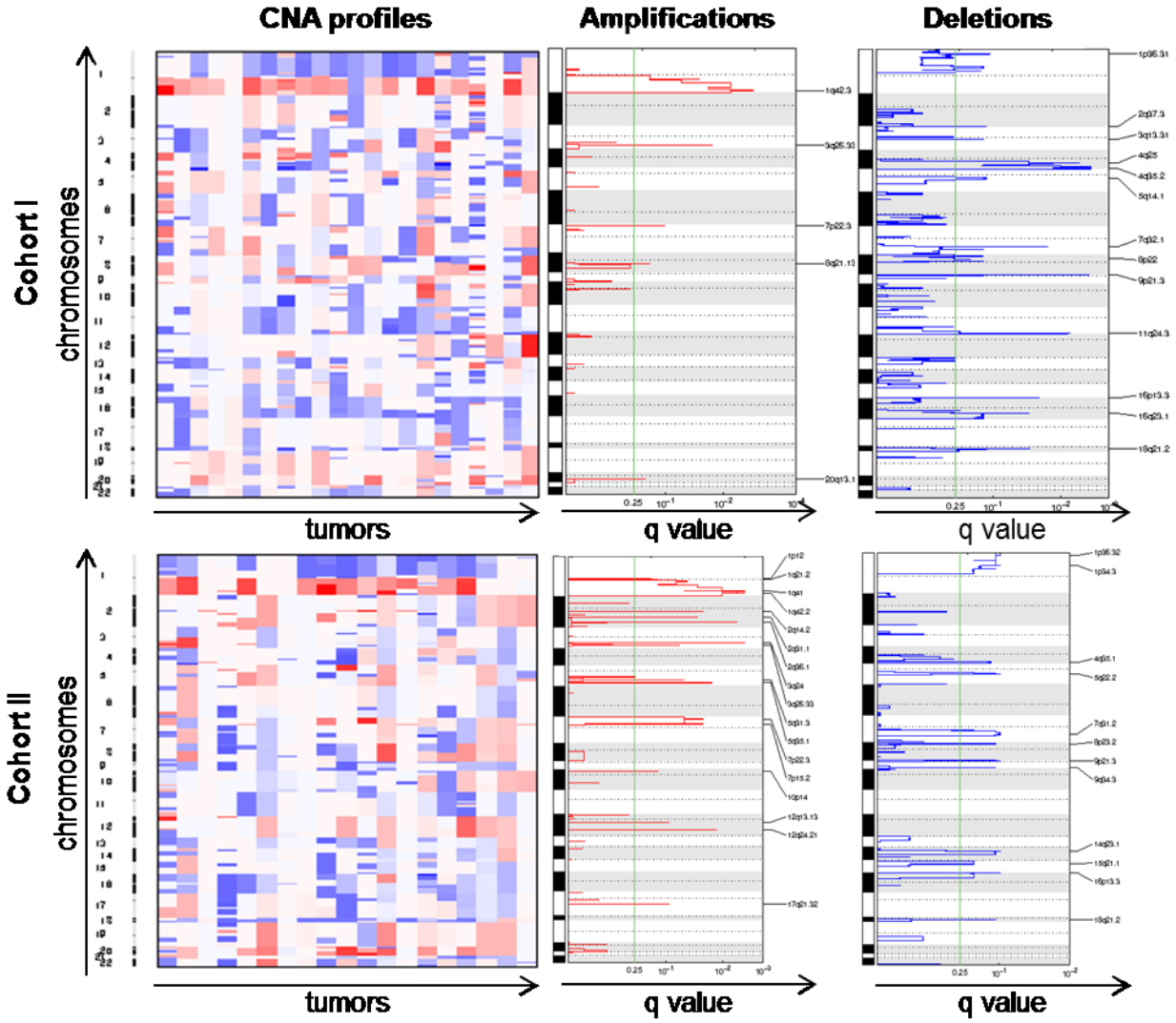


Figure 38 ACC show commonly shared CNA. Copy numbers are depicted in a heatmap (left panel) for each ACC from cohort I (x axis) and each chromosome (y axis). Commonly amplified regions (center panel) and deleted regions (right panel) are depicted with a q-value threshold of 0.25 (green line). Cohort I was depicted on top, cohort II on the bottom. Amplifications = red, deletions = blue.

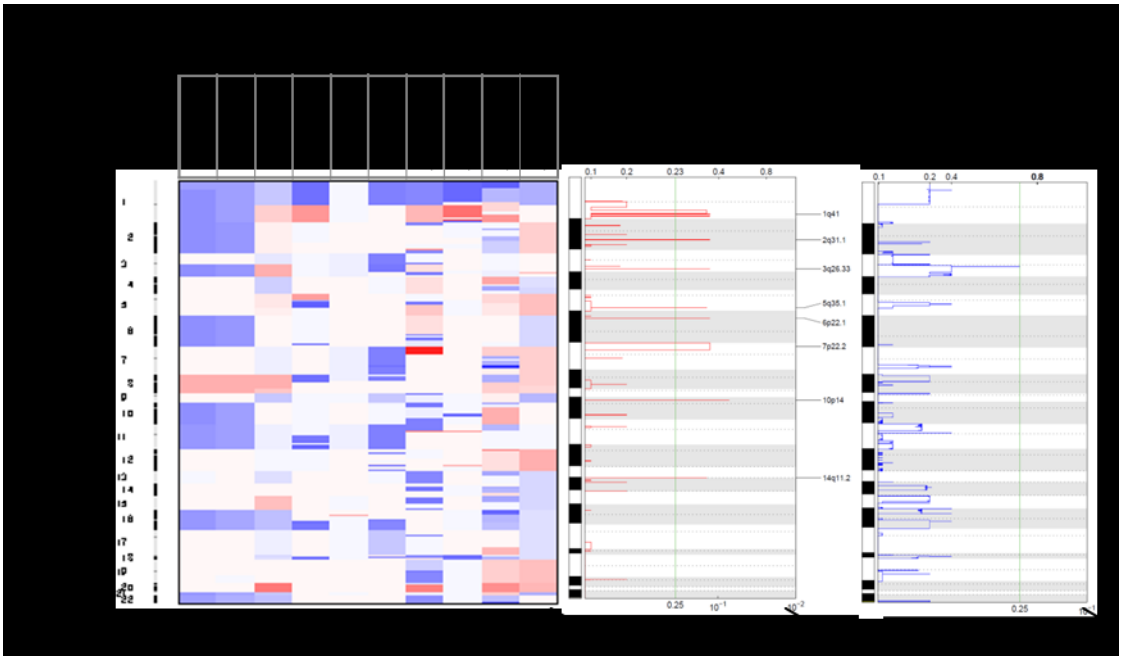


Figure 39 MACNEC show commonly shared CNA. Copy numbers are depicted in a heatmap (left) for each MACNEC of cohort I and cohort II (x axis) and each chromosome (y axis). Commonly amplified regions (center) and deleted regions (right) were depicted with a q-value threshold of 0.25 (green line). Amplifications = red, deletions = blue.

4.5.1 Copy number aberrations lead to many deleted genes in ACC

Next, amplified and deleted regions identified in Figure 38 were mapped to genes, and 2324 deleted and 323 amplified genes in cohort I were identified. Of these, 62% of deleted and 11% of amplified genes were validated in cohort II (Figure 40). In addition, some of these regions were confirmed with quantitative PCR (Figure 50c).

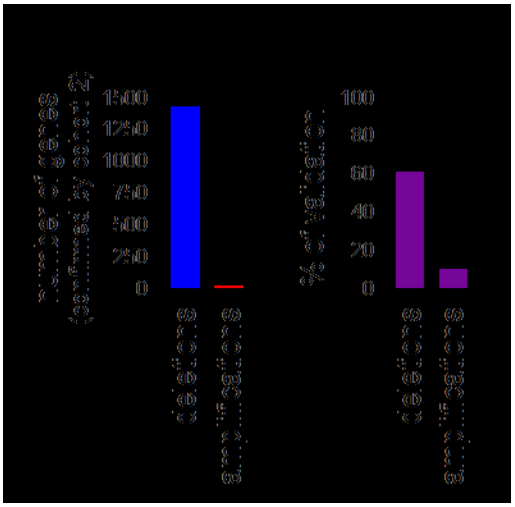


Figure 40 Number of genes mapping to CNA in ACC. a. total number of genes **b.** confirmation rate of identified genes in cohort II in %.

Gene set enrichment analysis of these 1441 deleted and 35 amplified genes revealed enrichment of pathways associated with negative regulation of the cell cycle and cell growth and mitosis (Figure 41).

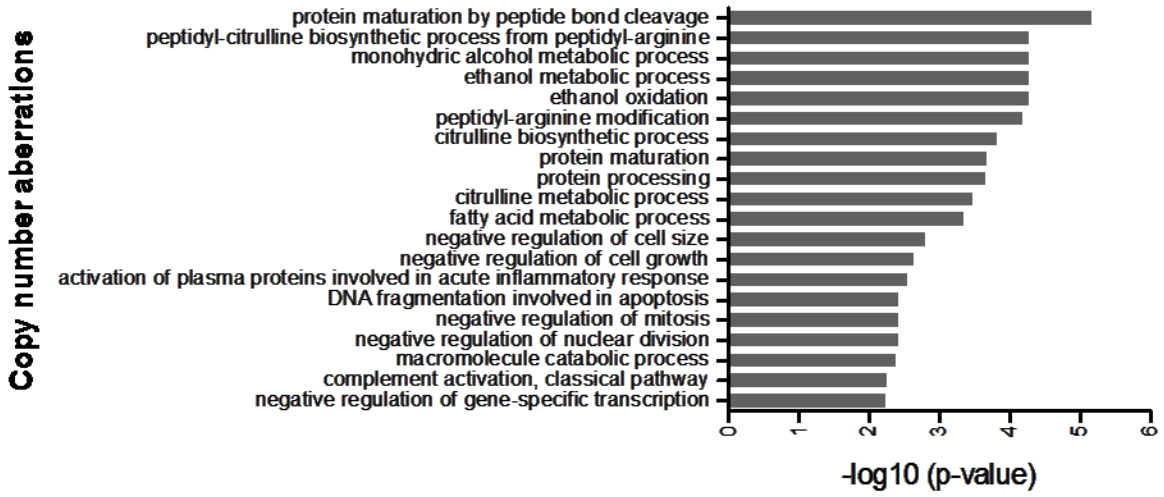


Figure 41 Gene set enrichment analysis of genes significantly deleted or amplified in ACC.

4.5.2 Aberrations in copy numbers predominantly occur during tumorigenesis und not during metastases formation

To investigate whether tumor cells acquired additional aberrations upon metastases formation, a set of ACC metastases was examined on the CNA level (Figure 42). Most CNA of metastases were very similar to their primary tumors (compare metastases from Figure 42 with Figure 37; numbers refer to individual patients). Exception, where aberrations did not recapitulate the aberrations found in the primary counterpart were 8M and 23LK (Figure 42). These were two of the three samples below a tumor purity of 60% (Figure 12b), which was likely the reason for the normal copy numbers found in these two samples. When plotting primary tumors next to metastases in a heatmap, similarities of primary tumors and metastases became even more evident (Figure 43). An additional evaluation of CNA on the gene basis did not reveal any recurrent changes from primary tumors to metastases.

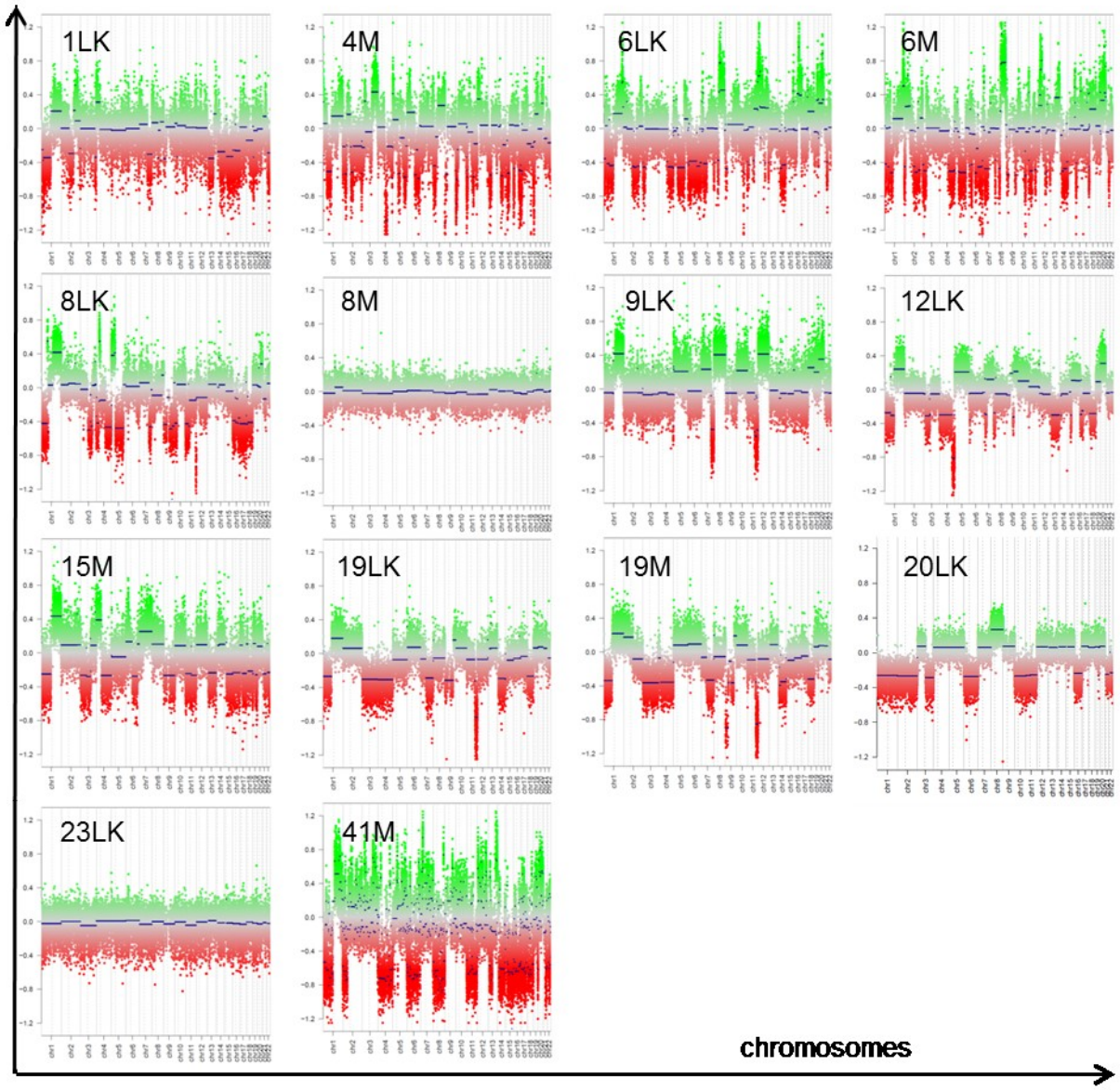


Figure 42 Metastases from cohort I had highly instable genomes. Copy number profiles were calculated for each tumor and CNA are depicted for each tumor and each chromosome. Amplifications are depicted in green while deletions are depicted in red. Numbers correspond to patients, CNA of primary tumors of these patients can be found in Figure 37.

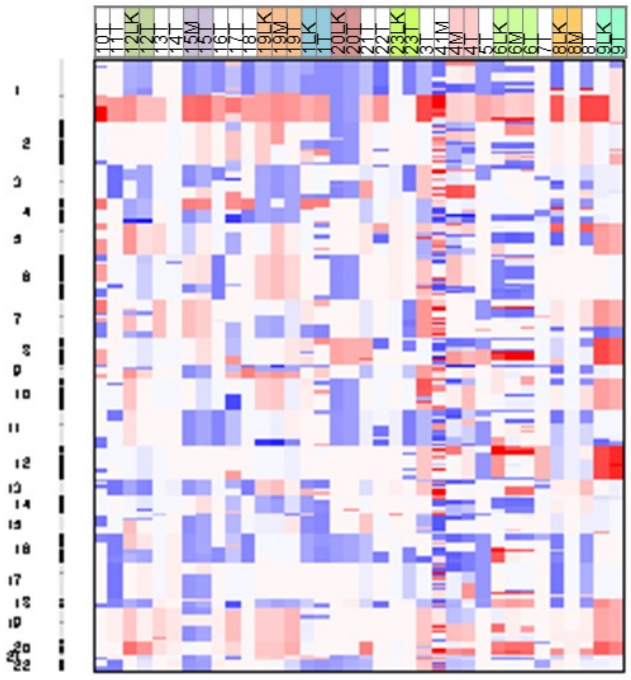


Figure 43 Metastases resemble their primary tumors. Copy numbers are depicted in a heatmap for each metastasis (x axis) and each chromosome (y axis). Metastases are plotted next to their primary tumors and colors correspond to patients (white: tumors without metastases). Amplifications = red, deletions = blue.

4.5.3 Copy number landscape from ACC differ vastly from other pancreatic tumors

Most of the here reported CNA were not previously reported in ACC. There are a number of studies investigating CNA in other pancreatic cancers, however not based on 450K arrays. To compare CNA of different pancreatic cancers based on the same method, 450K data were used to comparatively analyze CNA in ACC, PNET, and PDAC.

In contrast to ACC, PNET showed many chromosomal arm level aberrations, mainly amplifications (Figure 44 and Figure 45). Localized aberrations were nearly absent, with the exception of three amplified regions that cannot be detected in ACC, namely 8p23.1, 14q32.31, and 22q11.1 (Figure 45 middle and right panel). This genomic picture of PNET seems quite unique and very recently a study employing GISTIC analyses based on SNP arrays revealed many amplifications and only few deletions in PNET[214]. This however, needs further validation.

Next, CNA of ACC were compared with PDAC. PDAC harbored many small amplifications and deletions. The deletions of 1p36, 9p21.3, and 18q21.2 and amplifications in 7p22.3 identified in ACC were present in PDAC (Figure 46). However the deletions in 4q35 and 16p13.3, and the amplifications in 1q42 and 3q26.33 were unique to ACC.

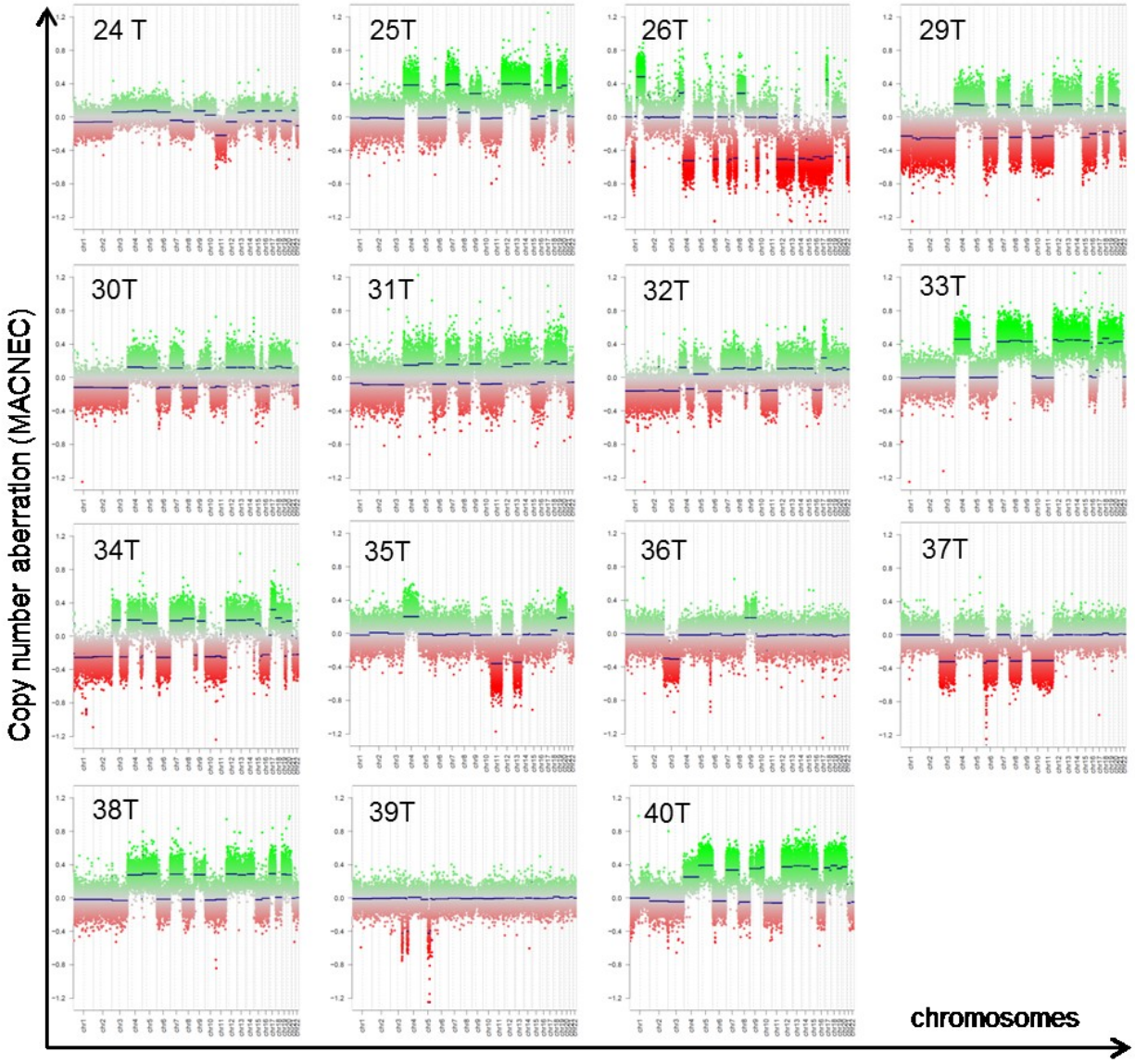


Figure 44 PNET had highly instable genomes. Copy number profiles were calculated for each PNET and CNA are depicted for each tumor and each chromosome. Amplifications are depicted in green while deletions are depicted in red.

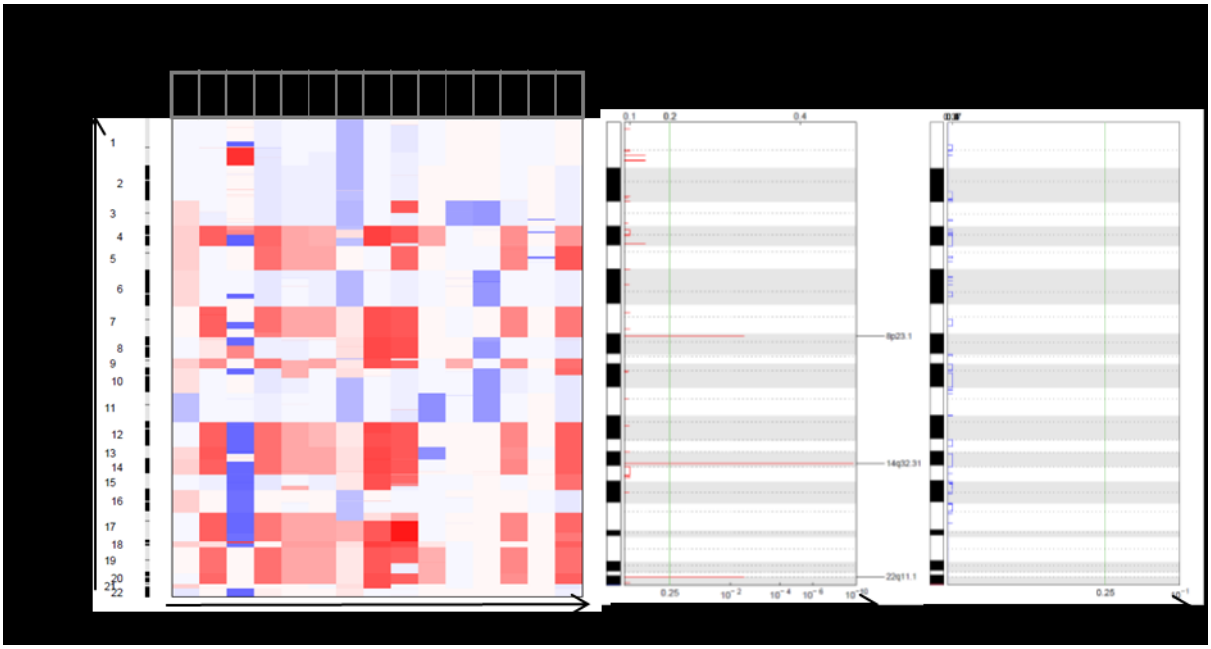


Figure 45 PNET show commonly shared CNA. Copy numbers are depicted in a heatmap (left) for each tumor (x axis) and each chromosome (y axis). Commonly amplified regions (center) and deleted regions (right) are depicted with a q-value threshold of 0.25 (green line). Amplifications = red, deletions = blue.

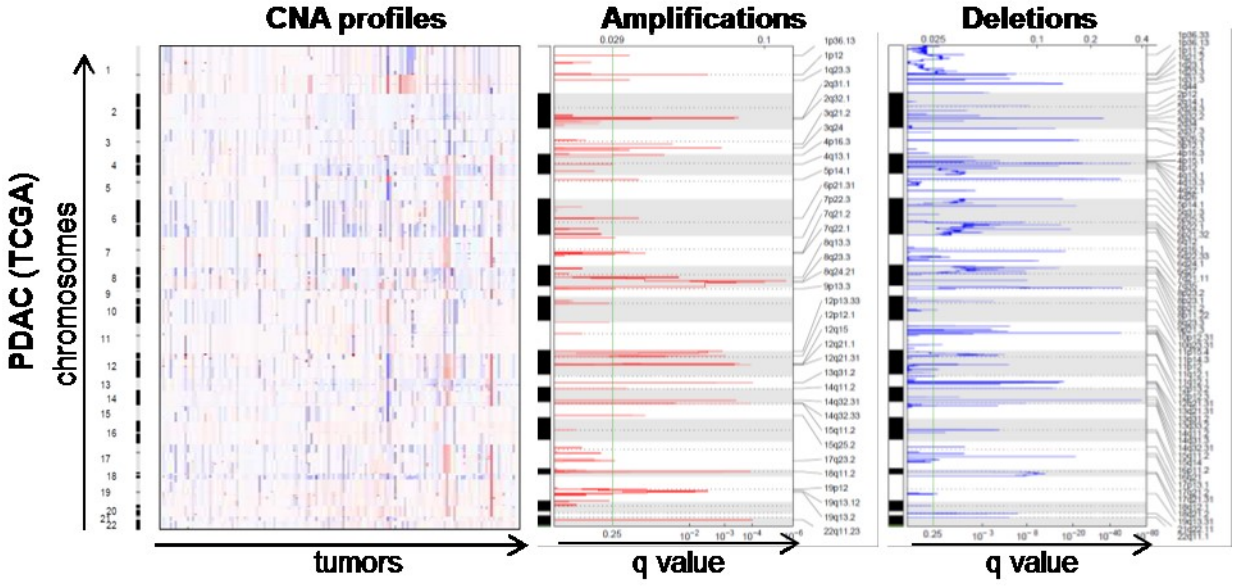


Figure 46 PDAC show commonly shared CNA. Copy numbers are depicted in a heatmap (left) for each tumor (x axis) and each chromosome (y axis). Commonly amplified regions (center) and deleted regions (right) are depicted with a q-value threshold of 0.25 (green line). Amplifications = red, deletions = blue.

4.6 Integrative analysis reveals many aberrations in cancer-related genes

To obtain a complete view of chromosomal aberrations in ACC, DNA methylation was integrated with CNA data in all tumors from both cohorts. Point mutations were not included, as they were not highly recurrent and as data was only available for cohort I. Depicting the data in a circos plot revealed that CNA occurred at distinct chromosomes or chromosome arms, *e.g.*

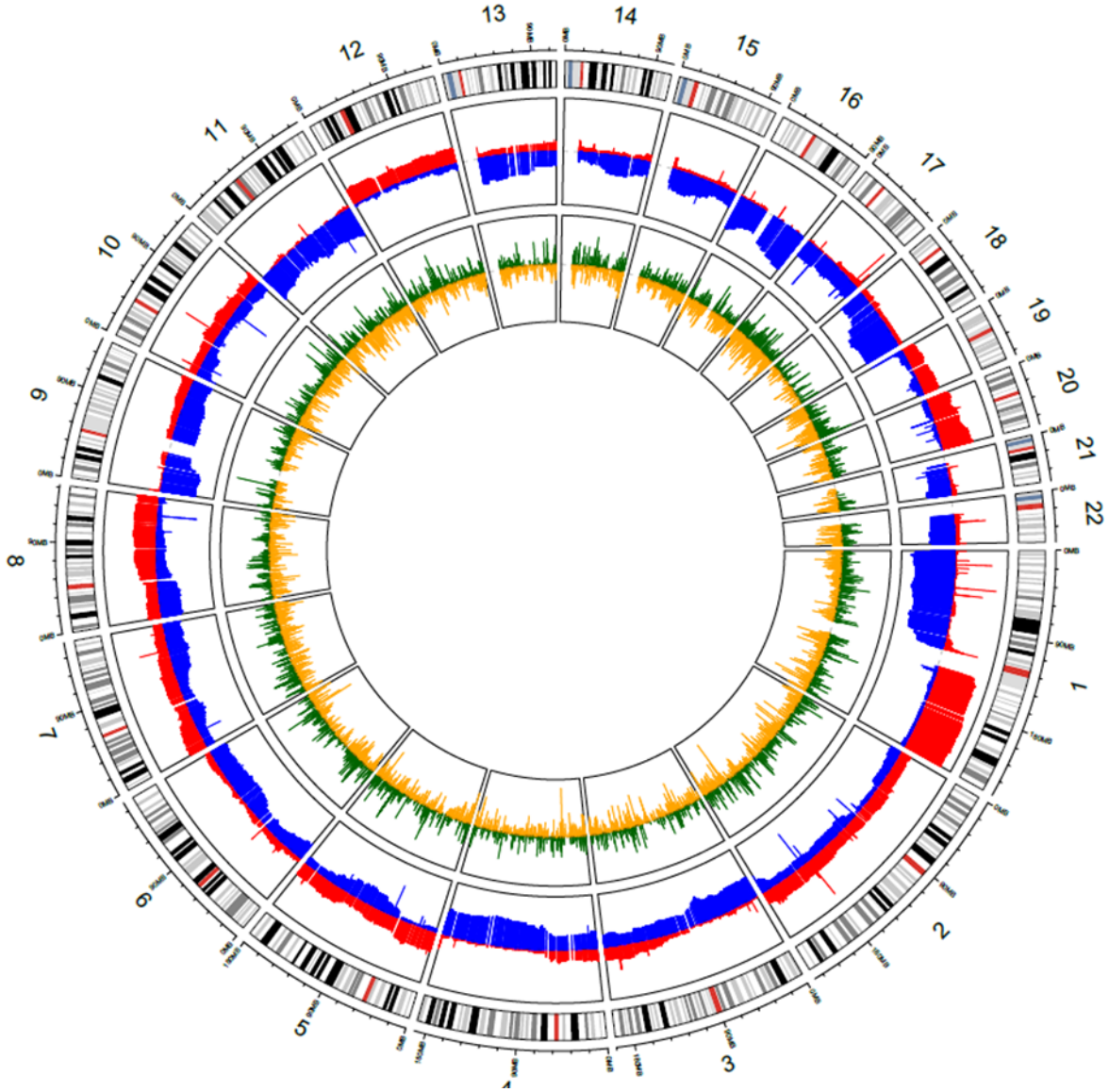


Figure 47 Circos plot displaying CNA and differential methylation in ACC. Outer circle: all autosomes are depicted (centromere in red). Middle circle: CNA are depicted in relation to their recurrences (top or bottom of circle represents all tumors) (red: amplifications, blue: deletions) Inner circle: Differential methylation was depicted in relation to recurrences (top or bottom of circle represents all tumors) (green: hypermethylation, yellow; hypomethylation).

amplification of 1q and deletion of chromosome 11. In contrast to this, aberrant methylation was quite evenly distributed with hyper- and hypomethylation occurring in every chromosome (Figure 47 and Supplementary Figure 14). In a next step, frequently altered genes were identified by putting each gene in each tumor into one of nine categories: deleted and promoter

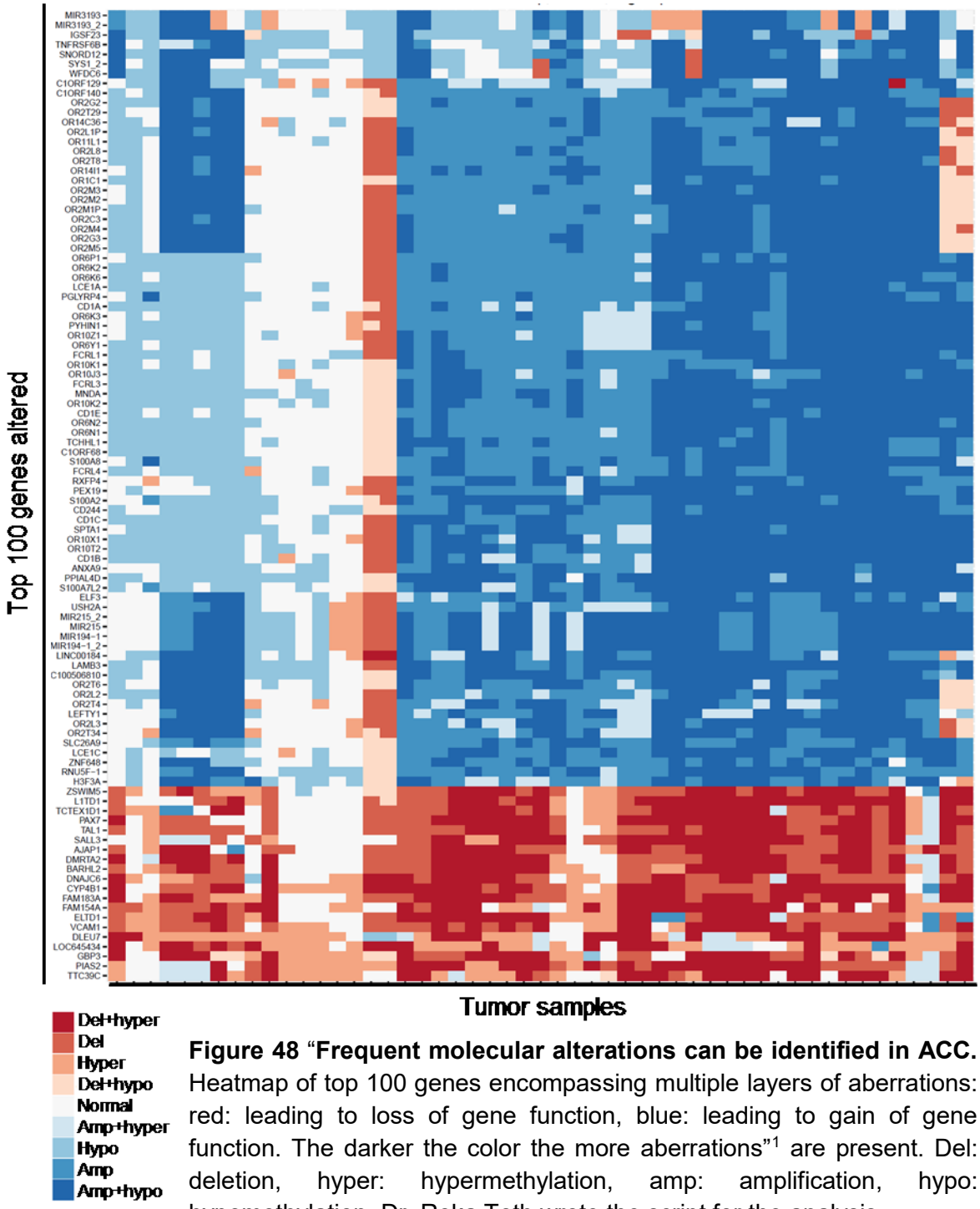


Figure 48 "Frequent molecular alterations can be identified in ACC. Heatmap of top 100 genes encompassing multiple layers of aberrations: red: leading to loss of gene function, blue: leading to gain of gene function. The darker the color the more aberrations"[1] are present. Del: deletion, hyper: hypermethylation, amp: amplification, hypo: hypomethylation. Dr. Reka Toth wrote the script for the analysis.

hypermethylated, deleted only, hypermethylated only, deleted and promoter hypomethylated, unaltered, amplified and promoter hypermethylated, hypomethylated only, amplified only, amplified and hypomethylated. The top 100 genes were depicted in Figure 48 and - in contrast to the current literature - many genes were altered in the majority of ACC. As the goal of this study was to identify targets that lead to cancer, genes that were associated with tumorigenesis were extracted from the literature. To obtain a list as comprehensive as possible, multiple published lists with cancer-related genes were used. These lists were not mutually exclusive and not of the same lengths as they served as a resource list of potentially relevant genes in tumorigenesis. Top hits from the present study were overlapped with known and candidate cancer genes from King's college[261], tumor suppressor genes from TSGene[262], driver genes (mutated, CNA, and predisposition driver genes) from Vogelstein, et al. [20], and epigenetic regulators from Plass, et al. [67]. A total number of 292 genes were identified in ACC (Figure 49, for the whole list refer to Supplementary Table 8).

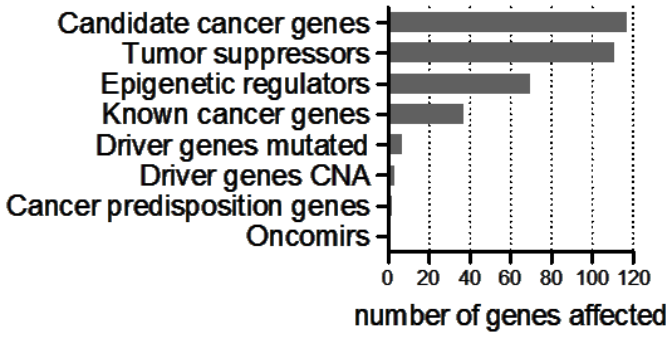


Figure 49 Genes identified by integrative approach that re-occur in previously published gene lists associated with tumorigenesis: Oncomirs (oncogenic microRNAs, King's college)[261], Cancer predisposition genes (Vogelstein et al.)[20], Driver genes CNA (Vogelstein et al.)[20], Driver genes mutated (Vogelstein et al.)[20], Known cancer genes (King's college)[261], Epigenetic regulators (Plass et al.)[67], Tumor suppressor genes (TSGene)[262], Candidate cancer genes (King's college)[261].

4.7 Findings from integrative analysis can be confirmed on the protein level

Next it was "investigated whether the observed epigenetic and genetic aberrations in ACC come along with a loss of the respective protein expression. Immunohistochemical stainings were performed on a tissue microarray including 23 ACC from cohort I and 39 ACC from cohort II as well as 8 normal pancreatic samples. The protein expression of eight aberrant genes for which high quality antibodies were available (ARID1A, APC, CDKN2A, HIST1H, ID3, JAK1, PCDHG, and SOX2) was evaluated."[1] Four of these genes that showed very frequent deletions and some promoter hypermethylation (Figure 50) were identified that lead to a decrease or loss of expression on the protein level (Figure 51), "while the remaining four display protein changes

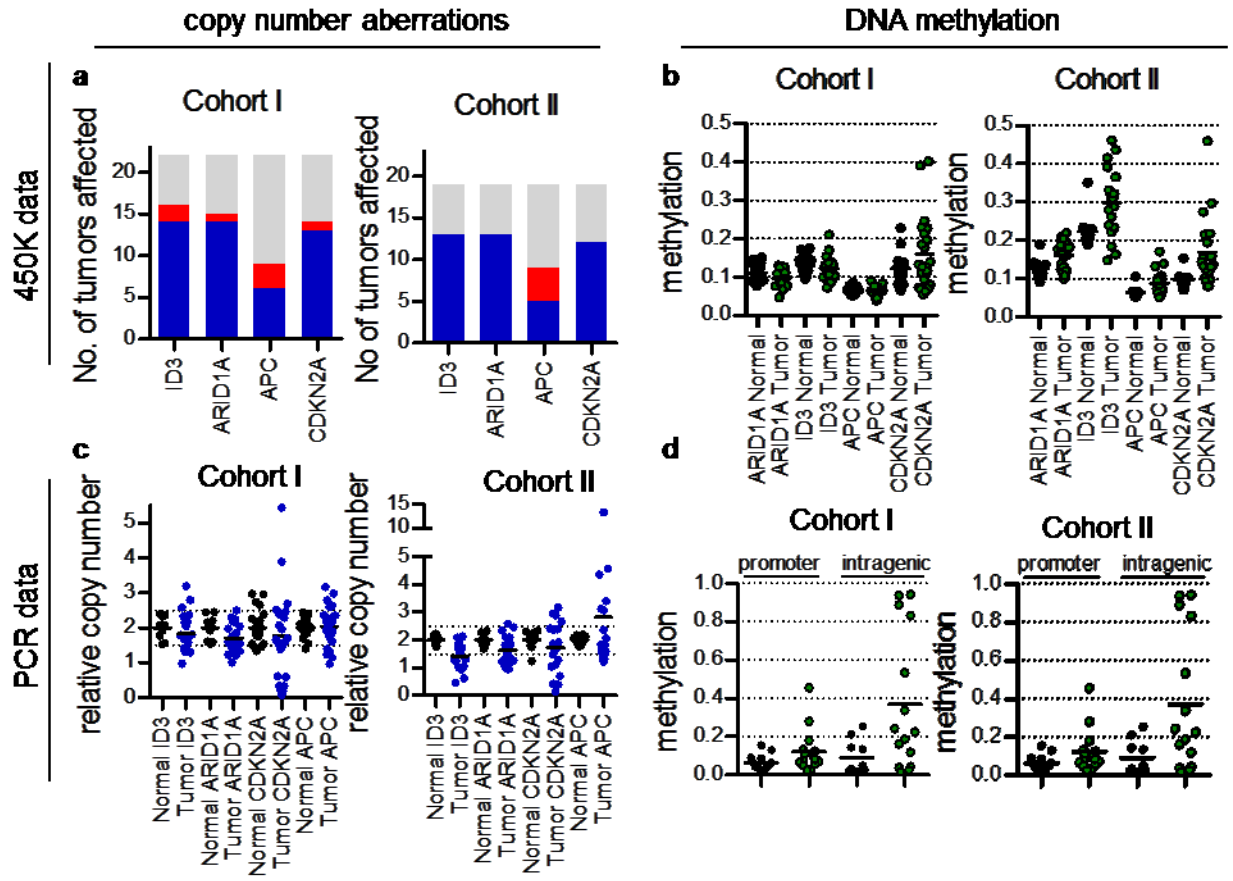


Figure 50 CNA, DNA methylation of *ID3*, *ARID1A*, *CDKN2A*, and *APC* by 450K and PCR of cohort I and cohort II. a. CNA as measured by the 450K array (blue: deletion, red: amplification, grey: normal). **b.** Promoter methylation as measured by the 450K array (black: normal, green: tumor). **c.** CNA as measured by qPCR (black: normal, blue: tumor). **d.** promoter and intragenic methylation of *CDKN2A* as measured by MassARRAY (black: normal, green: tumor). Daniel van der Duin performed qPCR analysis and MassARRAY under my supervision.

only in a few tumors"[1] (Supplementary Figure 15). In fact, "ID3 was down-regulated in 89% and 94%, ARID1A in 68% and 74%, APC in 71% and 62%, and CDKN2A in 53% and 52% of samples from cohort"[1] I and II, respectively (Figure 51), for statistical analysis refer to Supplementary Table 9. "Sixty out of 62 investigated tumors had a downregulation in one or more of these four proteins"[1] (Figure 52a). "Strikingly, most tumors showed protein alterations in more than one of these four tumor suppressor genes. Nineteen cancers show alterations in all four genes, 21 cancers in three genes and 13 in two genes"[1] (Figure 52b). "Interestingly, ID3 and ARID1A alterations were evident in 60 out of 61 tumors, suggesting that these two tumors suppressors were more important events. It was possible to predict the majority of protein losses based on the aberrant DNA methylation and copy numbers"[1] (Figure 53). "As expected from the WES results, point mutations did not add any additional value to this"[1] (mutations are marked as "X" in Figure 53). "Taken together, the loss of several tumor suppressor genes"[1] was identified and affected "a majority of ACC, suggesting an important role of these genes during the initiation and progression of these tumors."[1]

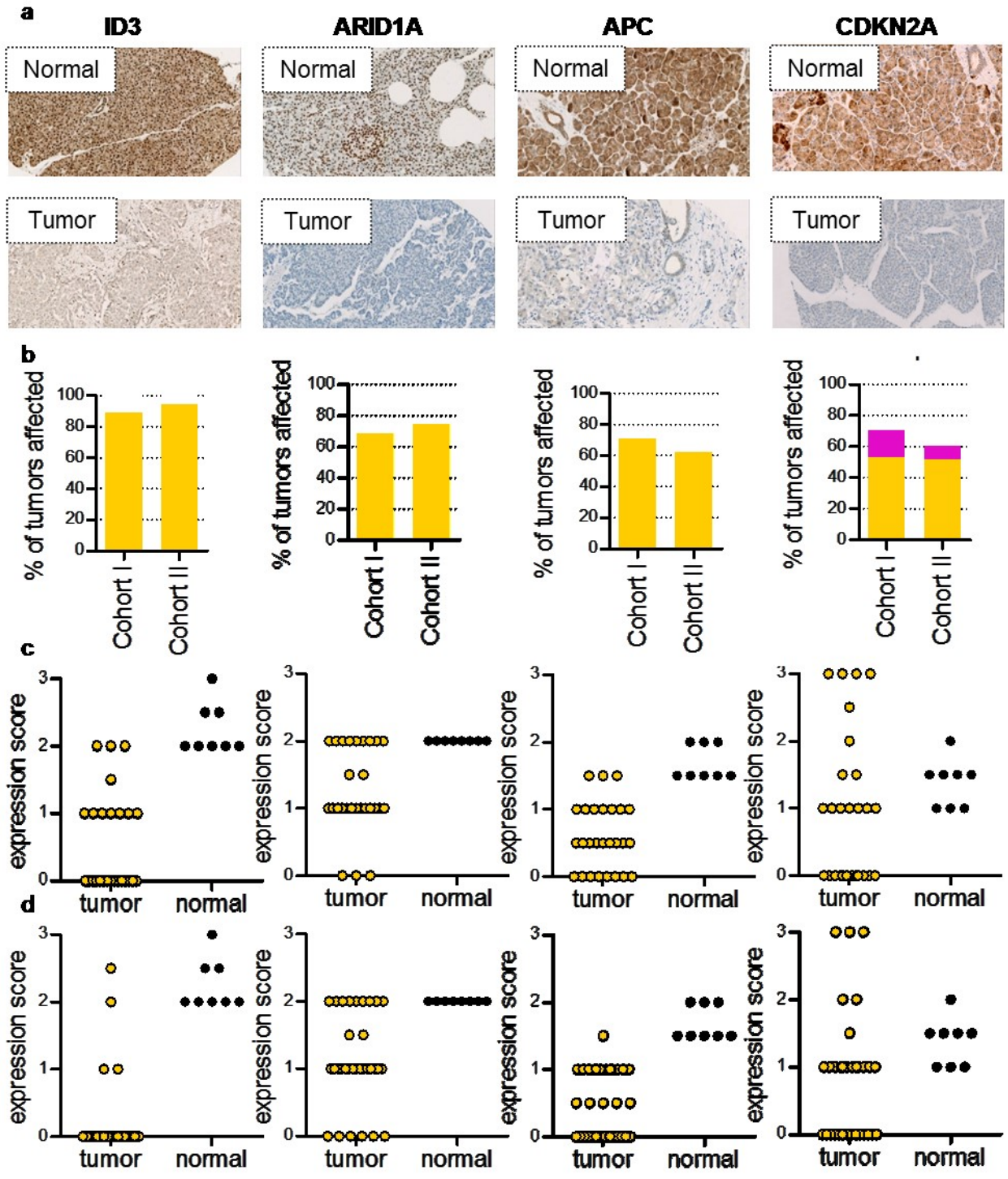


Figure 51 ID3, ARID1A, APC, and CDKN2A are down-regulated in the majority of ACC. a. representative IHC figures of normal and tumor tissue. PD Dr. Frank Bergmann performed IHC stainings. **b.** percentages of tumors affected. Yellow: downregulated, pink: upregulated. **c.** dot plot of pathology score of protein expression of cohort I **d.** dot plot of pathology score of protein expression of cohort II. yellow: tumor, black: normal.

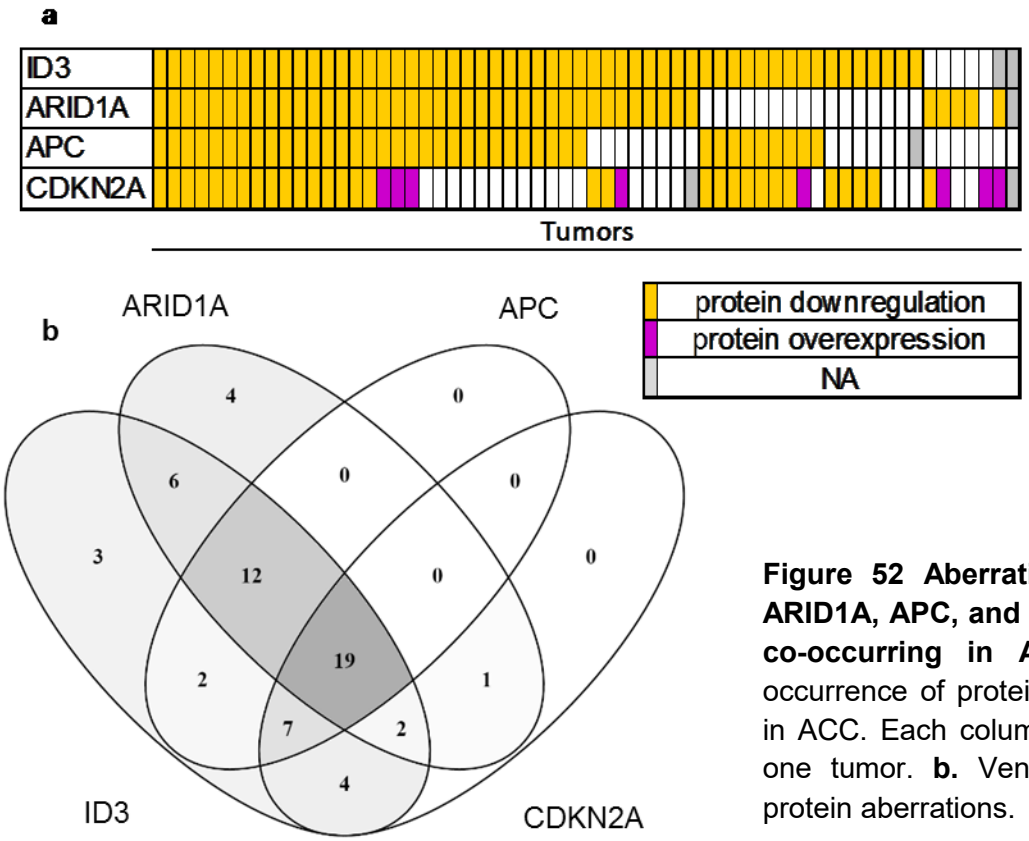


Figure 52 Aberrations in ID3, ARID1A, APC, and CDKN2A are co-occurring in ACC. a. co-occurrence of protein aberrations in ACC. Each column represents one tumor. **b.** Venn-diagram of protein aberrations.

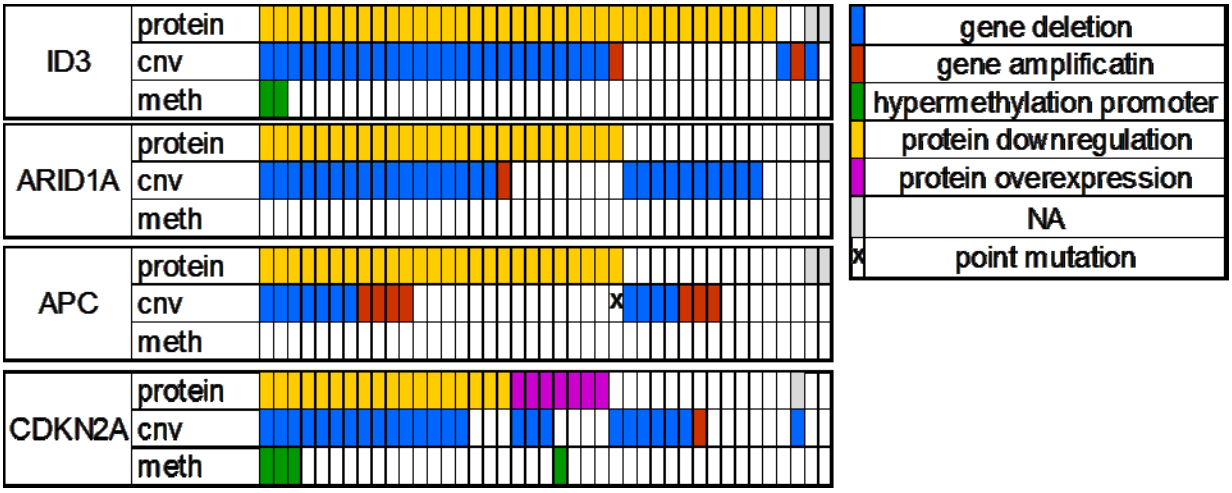


Figure 53 Aberrations in *ID3*, *ARID1A*, *APC,* and *CDKN2A* are co-occurring in ACC. Co-occurrances of protein aberrations with molecular aberrations in copy numbers and DNA methylation. Each column represents one tumor. Columns are sorted within each gene.

4.8 Molecular subgroups can be identified in ACC based on DNA methylation and copy number aberration data

Bergmann, et al. [147] has previously identified three molecular subgroups based on aCGH data. To investigate whether ACC form subclusters iCluster analysis[245], which builds molecular clusters on the basis of multiple genome-wide layers, was employed. Data for DNA methylation and CNA was used as input and 517 promoters from the DNA methylation data and 324 genomic regions of the CNA data revealed three molecular subgroups (refer to Supplementary Table 10 for regions and promoters defining the clusters). One cluster revealed a high methylation, one a low methylation and one an intermediate methylation phenotype (Figure 54). These clusters did not correlate with any of the clinical data that was available (age, gender, pure versus mixed ACC, metastases, smoking), nor with the allocation to the cohorts, tumor purity, or survival of patients (Figure 54 and Figure 55), and therefore remain to be investigated.

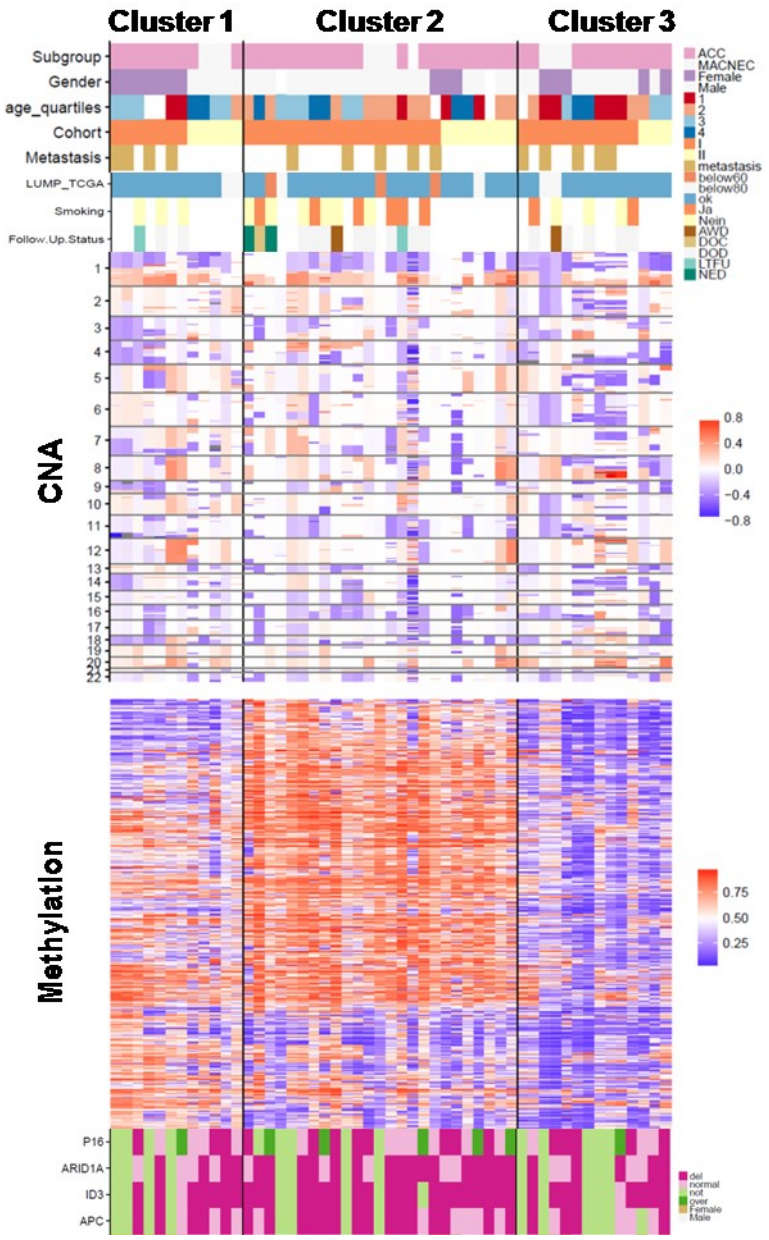


Figure 54 ACC form three molecular subclusters. CNA (red: amplifications, blue: deletions) and differential methylation (red: hypermethylation, blue: hypomethylation) define three subclusters. Top annotation panel displays clinical characteristics: Subgroup: mixed versus pure ACC, Gender, age quartiles (patients are grouped into quartiles based on their age), cohort (I or II), metastasis (tumors which were not primary tumors), LUMP_TCGA (tumor purity, based on cutoff used by TCGA), smoking (yes or no), and follow-up status (AWD: alive with disease, DOC: dead of other cause, DOD: dead of disease, LTFU: long-term follow-up, NED: no evidence of disease). Bottom annotation table displays aberrant protein expression in ID3, ARID1A, CDKN2A, and APC.

4.9 Survival in ACC correlates with age but not with any other clinical or molecular parameters

Data for survival was available for cohort I and survival analysis was performed according to different clinical and molecular groups. The overall survival of all patients was 27 months. Survival correlated with age, especially patients in the last age quartile displayed short survival times (nine months). The status of smoking, gender or subgroup of ACC (pure versus mixed) did not correlate with survival. Neither did molecular patterns, *e.g.* molecular clusters based on DNA methylation and CNA data, or protein expression of ARID1A, ID3, CDKN2A, and APC alone or in combination (Figure 55). However, as the sample size for ACC with available survival data was small, results should be interpreted with precaution and validated in larger cohorts.

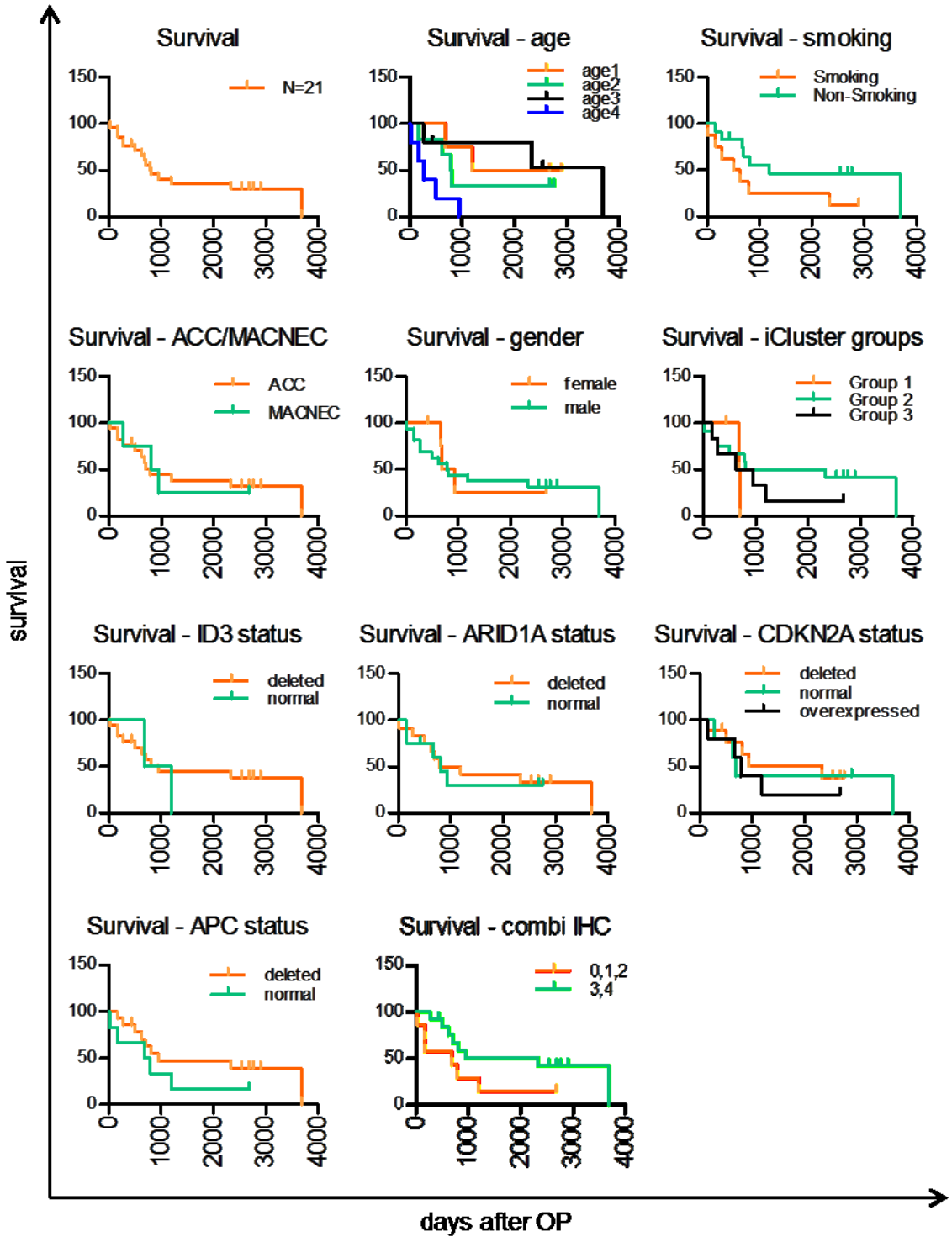


Figure 55 Survival curves of cohort I depending on different clinical and molecular parameters. iCluster groups: groups defined above (see 4.8); combi IHC: combination of protein aberrations: 0, 1, 2 refers to zero, one, or two aberrantly expressed tumor suppressor genes. 3, 4 refer to three or four aberrantly expressed tumors suppressors. Time was depicted in days from date of operation (OP).

5 Discussion

5.1 Genome-wide screen identifies frequent molecular aberrations in ACC

In this work, a genome-wide approach was undertaken to unravel the molecular alterations of ACC. Up to date most studies focused on investigating single genes or gene sets and did not identify molecular aberrations in the vast majority of ACC[182,186]. The only two studies investigating ACC genome-wide were WES studies, in which no frequently recurrent point mutations were identified. This study presented a genome- and epigenome wide approach, identifying numerous aberrations also in driver genes, validating them on the protein level and identifying potential drug targets which might allow ACC patients to participate in so-called "basket trials".

5.1.1 WES reveals high mutational load and specific mutational signatures

PDAC "are characterized by a surprisingly high recurrent mutational load in *KRAS* and *TP53* (~90% and ~50%, respectively), and by only few genes that show frequent localized deletions[16,17],"[1] whereas PNET are characterized by frequent mutations in *MEN1*, *DAXX*, *ATRX*, *TSC2*, and *PTEN*, often involving germline mutations[214]. However hyper-mutation in PDAC is rather rare (~1%)[263] and has not been described in PNET. "In striking contrast, ACC did"[1] show a high mutational burden per tumor, but did "not show any recurrent point mutations."[1] Somatic signatures of these "tumors highly differed. ACC exhibited two signatures (4 and 29) associated with tobacco consumption which were not present in PDAC[26,28]."[1] In addition, "ACC harbored the mutational signatures 15 and 20, which did not occur in PDAC[26] and which are caused by defective DNA repair mechanisms."[1] These data suggest that (i) smoking is a risk factor for ACC, and (ii) these tumors might be caused by defects in DNA repair mechanisms. As there was no detailed smoking history of ACC patients, this remains to be validated. In addition, the DNA repair machinery should be further investigated to elucidate, which genes or pathways are altered and therefore contribute to the high chromosomal instability in ACC. As an additional layer of information, whole genome sequencing could be performed in order (i) to be able to call insertions, deletions, and fusions of chromosomes that might contribute to tumorigenesis, and (ii) to improve the calculations of the mutational signatures, which are more accurate with increasing number of mutations. Furthermore, the point mutations and mutational signatures should be confirmed in cohort II, which was not possible due to the lack of matched normal tissues. Sequencing without matched tissue is not recommended as cells accumulate an

increasing number of non-functional mutations during their lifetime which are irrelevant for tumorigenesis but cannot be filtered out.

5.1.2 The ACC methylome is highly aberrant

Methylation analysis in ACC revealed that there were many aberrant methylation events detectable in ACC (Figure 15 and Figure 16) and that these aberrations were distinct from normal pancreatic tissue and the other pancreatic cancers PDAC and PNET (Figure 17 and Figure 18). It revealed that the methylation profile of MACNEC was similar to ACC (Figure 18), suggesting that MACNEC form one common tumor entity with pure ACC, as clinical parameters previously suggested[146,264]. These data hold the potential to diagnose ACC based on their methylome. Although in most cases IHC is clearly distinguishing the different tumor entities, there are cases where diagnosis was not always unambiguous. In these cases methylation analysis would be highly beneficial, as clinical parameters, *e.g.* expected survival, differ between these pancreatic cancers. In addition, the aberrations occurring during tumorigenesis in ACC revealed changes in pathways involved in development (for further discussion refer to 5.5) and homophilic cell adhesion (for further discussion refer to 5.2). Previously reported differential methylation in ACC was not reproducible[173,177,183]. This might be due to the non-specific methods used, *e.g.* methylation-specific PCR[183], which is not a quantitative measurement of DNA methylation. Thus, the here presented results of the 450K array are more reliable.

“A further interesting finding”[1] based on the methylome data “was that ACC contain only few infiltrating immune cells, consisting mainly of CD8+ T-cells.”[1] In contrast, “PDAC, as previously shown[265-267], are infiltrated by many immune cells with CD4+ T-cells prevailing,”[1] whereas PNET are mostly infiltrated by natural killer cells and monocytes (Figure 14). This suggests that these tumor types were not only different on the intrinsic level, but in the way the immune system combatted them.

In the differential methylation calling process, the normal tissue samples were both from adjacent normal tissue of cancer patients and from pancreatic tissues of healthy individuals. This was done to compensate for the following two scenarios: (i) adjacent normal tissue might already contain differential methylation leading to tumorigenesis and these relevant changes would be missed when only using adjacent tissue. (ii) Normal pancreatic tissue from healthy individuals might contain differential methylation due to other factors, *e.g.* genetic background, environmental factors, and thus differential methylation might be called that occurred not due to

tumorigenesis. These points however would probably only account for small changes, as the differential methylation calling between adjacent normal tissue and healthy normal tissue did not reveal many aberrant regions (Supplementary Figure 4 and Supplementary Figure 5).

Mainly promoter sites and gene bodies were investigated for differential methylation in this study. However, aberrant DNA methylation also plays a role outside of these CpG dense regions, *e.g.* in enhancers as shown in Figure 20b and Supplementary Figure 5b. These hits were not further followed, as enhancers are not equally well covered on the 450K array as the coding regions. To better evaluate the methylome of ACC, Infinium MethylationEPIC array with more than 850 k CpG sites and a higher coverage of enhancers (ENCODE[268] and functional annotation of the mammalian genome (FANTOM5) enhancers[269] are included) should be performed[270]. Another alternative is whole genome bisulfite sequencing[271], which would provide an even higher coverage of CpGs of the genome than the array-based approaches, but would request the availability of fresh tissue samples.

To examine the molecular differences between ACC and the other pancreatic cancers in more detail, differential methylation between the tumor entities should be compared. However, as the tumor purity of PDACs was by far lower than that of ACC (Figure 14) this comparison was not performed in the presented study, as this would likely lead to false positive hits due to differential methylation between blood cell types and pancreatic tissue. For differential methylation calling of PNET, tumors should be compared to the respective neuroendocrine cells. However, only too few sorted neuroendocrine normal cells were available for comparison and those were from a different background and acquisition method which could bias the analysis (sorted cells were from fresh tissues whereas ACC were from FFPE tissue, and 450K array for sorted cells was run in the USA). Endocrine cells should be isolated from adjacent healthy tissue from PNET patients to call proper differential methylation, though.

5.1.3 ACC harbor highly instable genomes

Aberrant copy numbers in ACC were previously described in two studies employing array comparative genome hybridization arrays[147] and fluorescence-based PCR[176,187]. Here, it was confirmed that ACC harbor instable genomes. The 450K array provided the opportunity to not only identify broad-ranged, but additionally focal chromosomal gains and losses. This made it possible to pinpoint deletions and amplifications to distinct genes and chromosomal sites thus overcoming the limited resolution of the previously used assays. Bergmann, et al. [147] described deletions of 1p, 9p, 16q, and chromosome 18, and amplifications of 1q and chromosome 7. This

work confirmed these alterations and located them to deletions of 1p36, 9p21.3, 16p13.3, and 18q21.2 and amplifications of 1q42 and 7p22.3. The deletion in 9p21.3 was most striking and contained *CDKN2A* which was one of the tumor suppressor genes identified in this work.

The downside of using copy number calculations based on the 450K array is directly linked to the disadvantage arising for methylation analysis. As the coverage of probes is only dense in gene-rich regions, the high resolution of copy numbers is restricted to those regions. In gene deserts, where only few probes were present, the resolution of CNA is limited. SNP arrays on the contrary would provide a more equally distributed CNA profile[39,272], however would add additional costs as another array has to be run.

The results from the CNA depend on the input parameters. Here, a bin size of five was used, which means that each selected segment in the genome had to contain at least five probes. If this was not the case, the segment was fused to the next segment to achieve a number of five. Increasing this bin size number leads to larger fragments and consequently lowers resolution, but results in a higher quality. On the contrary decreasing this bin size leads to smaller fragments and a higher resolution, at the expense of quality.

5.1.4 ACC did not acquire additional aberrations upon metastases formation

Although metastases are frequent in ACC patients[148,150-152], there is not much known about the molecular events during metastases formation in ACC. La Rosa, et al. [179] reported *TP53* alterations that occurred in 31% of ACC metastases, but only in 13% of primaries, suggesting metastases acquire additional molecular alterations. To that end this study encompassed 13 metastases. They harbored a very similar global methylation profile (Figure 16, Figure 26, and Figure 27) and are similar on the copy number level to primary tumors sites (Figure 42 and Figure 43). This suggests that metastases did not acquire any additional molecular changes upon metastases formation in this ACC cohort. Thus, the model of parallel progression[273] might apply here that metastases did not evolve during the late stages of tumorigenesis, but metastases already spread early from the primary tumors. This suggests that metastases harbor the same aberrations as the primary tumor, comparable to PDAC[201,274] but in contrast to PNET where metastases acquire additional lesions[220,221,223].

5.2 The protocadherin cluster is hypermethylated in ACC

The functional role of the *PCDH* cluster has up to now mainly been investigated in neuronal context and only little in other tissues. Previous studies showed hypermethylation in a number of tumor entities[253-257], and investigation of TCGA datasets revealed that this hypermethylation occurs in the vast majority of cancer entities (Figure 30 and Supplementary Table 4). However the functional impact of this remains unknown. Here it was shown, that hypermethylation is functional, *i.e.* leads to changes in the expression pattern. The result of PCDH downregulation was difficult to determine in the setting here, as the T510 cell lines have doubling times of approximately 24 hours and cells with transient expression are overgrown before the potential phenotype can fully manifest. To elucidate the functional relevance of PCDH expression outside of neurons, the best option would be to generate an isogenic pair with and without PCDH expression. To that end one could delete the constant exons of *PCDH* by the CRISPR-Cas method[275-277]. To address the functional role of methylation at this locus further, one could target the methylation of these sites with demethylating enzymes, *e.g.* Tet Methylcytosine Dioxygenase (*TET*) fused to a deactivated Cas variant which does not cut the DNA anymore (dCas)[278,279].

5.3 ACC harbor aberrations in genome stability and cell cycle control

"Up to date, this is the first study which showed in two independent cohorts, aberrations of"[1] the tumor suppressor genes *ARID1A*, *APC*, *CDKN2A,* and *ID3* "on the protein level in the majority of patients. These proteins play a role – amongst others – in chromosomal stability and cell cycle control"[1] (Figure 56). Thus, not only the mode of alterations in ACC versus other pancreatic cancers "is different, also the most common targets differ. ACC harbored frequent aberrations in *APC* which only occurred in less than 20% of PDAC[200,280] and a loss of ID3, which has been reported to be overexpressed in PDAC[281,282]." Conversely, *APC* and *ID3* mutations have not been described in PNET. ACC were "therefore not only different on the clinical and pathological side, but as this study showed also on many molecular levels."[1]

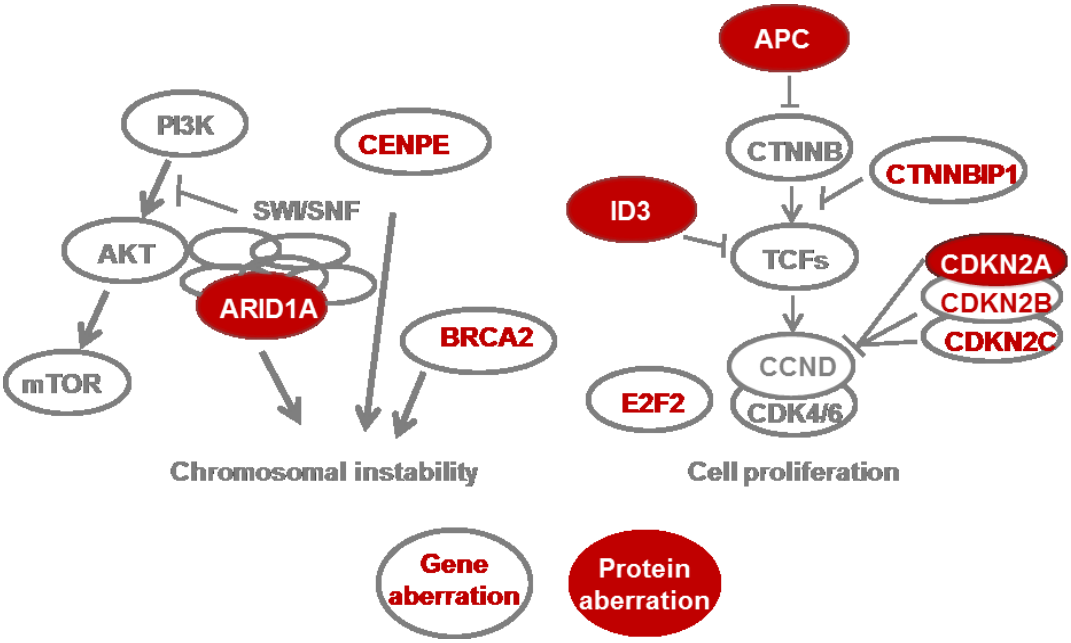


Figure 56 Graphical summary of aberrations in ACC. Affected genes are red, while genes additionally investigated on the protein level have a red background.

5.3.1 Molecular aberrations in genome stability

With this study it was shown "that ACC had an imbalanced genome"[1] (Figure 37, Figure 38 and Figure 56). "This might be explained by"[1] the here identified "(i) losses of *ARID1A*, *BRCA1/2,* and *CENPE*, and by (ii) mutational signatures associated with defective DNA repair. ARID1A was lost in approximately 70% of cases. ARID1A as part of the SWI/SNF complex is involved in multiple pathways to ensure chromosomal stability. It is involved in NHEJ and its depletion leads to an impaired NHEJ at DNA DSB and less accumulation of SMARCA2 (the catalytic subunit of the BAF SWI/SNF complexes) at laser-irradiated sites[283]. Due to interactions with DNA

topoisomerase II α (TOP2A) ARID1A plays a role in maintaining proper chromosome segregation and the prevention of anaphase bridges during mitosis[284]. This could lead to the massive chromosomal gains and losses observed in this study."[1] Jiao, et al. [182] and Furukawa, et al. [186] reported *ARID1A* mutations in 8% of ACC cases. Other than that no other genetic or epigenetic alterations have been reported in this gene in ACC. Here it was shown "that integrating different datasets and confirming hits on the protein level led to the discovery that ARID1A was very frequently abrogated in ACC. *ARID1A* deletions have been reported in a number of other tumor entities including endometrial, gastric, hepatocellular, breast, and pancreatic (PDAC) cancers (COSMIC website[285]). Many cancers have been shown to harbor besides an ARID1A loss also an activation of the PI3K pathway, including ovarian clear cell carcinoma[286,287], endometrioid ovarian carcinomas[287], and gastric cancers[288]."[1]

"BRCA2 is important in maintaining genomic stability as it is involved in"[1] HR. *BRCA2* mutation have been reported to occur in 3-12% of ACC[151,182,186] and in a few case studies[289,290], gene loss was reported in 40%[291] and loss of protein expression has been reported in 45% of cases[186]. Intriguingly a mouse model has been published were a heterozygous deletion of *Brca2* (in combination with *Kras*^{G12D} mutations) led to PDAC while a homozygous deletion led to ACC in the majority of cases[292], supporting that BRCA2 plays a role in ACC. "These aberrations and probably other alterations not identified in this study lead to mutational signatures associated with DNA repair defects in ~70% tumors. Mutational signature 3, which has been proposed to be caused by a loss of BRCA1 and/or BRCA2, was detected in 10 out of the 22 sequenced tumors. Signatures 6, 15, and 20 all associated with DNA mismatch repair occur in seven, five, and two tumors."[1] In the top list of recurrently aberrant genes the centromeric protein E (*CENPE*) was identified to be deleted. CENPE is responsible for capturing spindle microtubules during cell division. This needs further validation on the protein level. Overall, multiple hits were identified that could lead to the observed chromosomal instability and these give rise to new potential therapeutic interventions (refer to 5.4).

5.3.2 Molecular aberrations in cell cycle control

"Cell cycle control was impaired in ACC by"[1] aberrations of APC, CDKN2A, and ID3, which are all "negative regulators of the cell cycle"[1] (Figure 56).

APC was previously reported to be altered on the gene level in ACC in multiple studies. Mutations were identified in 7-18% of carcinomas[171,177,182,186], a gene loss was reported in 19-48%[177,291], a gain in 10-25%[177,291] and hypermethylation in 50%-67%[177,183]. A mouse model with

knockouts in *Tp53* (-/-) and *Apc* (+/-) revealed that 22% of mice develop ACC and that in these cases the additional *APC* allele has been lost, supporting the importance of this gene for ACC development[184]. “APC inhibits CTNNB1, which in turn activates Transcription Factors (TCFs),”[1] that activate transcription of a number of target genes involved in cell proliferation. “One of the target genes of TCFs is CCND, a member of the cell cycle[293,294]. *APC* was deleted in 71% and 62%”[1] (cohort I and II, respectively). No recurrent hypermethylation of *APC* was detected in this study. This might be due to the higher specificity of the 450K array compared to MSP and methylation-specific multiplex ligation probe amplification that were previously used[177,183].

“CDKN2A aberrations have been well studied as they are common in many tumor entities.”[1] In ACC, deletions were previously reported in 14-25%[177,180,291], amplifications in 25%[291] and hypermethylation in 58%[183]. No mutations or protein aberrations in ACC have been reported so far[174,177,182,186]. “Two different proteins encoded by *CDKN2A* activate the cell cycle by inhibiting both Cyclin D1 and the TP53 inhibitor MDM2.”[1] In the presented dataset, it was shown that indeed CDKN2A is a common target in ACC. CDKN2A protein expression was absent in >50% of cases and additionally cases with CDKN2A overexpression were detected, “suggesting other members of this pathway were abrogated, as reported for”[1] a number of other cancer entities[295,296].

Nearly all ACC “(~90%) exhibited aberrations in ID3.”[1] No previous reports in ACC on alterations in this gene exist. ID3 “is a basic Helix-loop Helix protein (bHLH) that inhibits other bHLH transcription factors by forming non-functional heterodimers, thereby preventing transcription of target genes. This is how ID3 inhibits the cyclin inhibitor CDKN1A[297] and activates cyclin D3 and cyclin E *via* TCF3[281,298].”[1] In addition, “ID3 activates Caspases 3 and 9, and inhibition therefore might lead to impaired apoptosis[299]. TCF3 further regulates acinar cell identity by activating expression of the pancreatic transcription factor *PTF1A*[282] and acinar transcription factor *BHLHA15* (= *Mist-1*)[300]. A loss of ID3 therefore should lead to an overexpression of TCF3, resulting in activation of the acinar cell program, whereas an activation of ID3 as shown by Kim, et al. [300] leads to an inactivation of the acinar cell program in PDAC. Thus, ID3 expression might be the switch that distinguishes PDACs from ACC”[1] as ID3 has been reported to be overexpressed in PDAC[281,282]. “Therefore, ID3 and TCF3 so far seem to be important players in the pancreas, but their physiologic functions and what happens during tumorigenesis remains to be elucidated.”[1]

There were “tumors with downregulated protein expressions in one of the four tumor suppressor”[1] genes “which were neither deleted, mutated, nor did they exhibit promoter hypermethylation”[1] (Figure 53). “Other epigenetic mechanisms, *e.g.* miRNAs or lncRNAs or upstream regulatory pathways/transcription factors, were not investigated in the present study and could shed further light in these cases.”[1]

5.4 New potential treatment options for ACC

The main treatment for ACC patients is surgical resection. Due to the lack of a standard therapy, every ACC patient receives a unique sequence of treatments with only limited success. Most commonly chemotherapies were used for treatment and only rarely targeted drugs or radiotherapy. Often, nucleoside analogs, *e.g.* gemcitabine or thymidylatsynthase inhibitors 5-fluoruracil and its derivates capecitabine, floxuridine, and tegafur were used. Cytostatic drugs that interfere with microtubules, *e.g.* paclitaxel and docletaxel have been reported. Platinum-based therapies, *e.g.* oxaliplatin and cisplatin were so far the most promising candidates for ACC patients. In addition, combination treatments of 5-FU, folinic acid, and oxaliplatin and/or irinotecan (a topoisomerase inhibitor) *e.g.* FOLFIRINOX, FOLFOX, FOLFIRI, and combinations of gemcitabine with irinotecan, capecitabine, cisplatin, oxaliplatin, and docetaxel were frequently used. The most commonly used targeted therapy reported is the EGFR inhibitor erlotinib. Other targeted therapies were scarce with one case receiving pyroxamide, an HDAC-inhibitor in combination with MEDI-522, an antibody against αVβ3 integrin[151,155,186,301-311].

Recent advances in therapies for other cancer entities were generally due to extensive research for the most common types of cancer, *e.g.* breast, lung or prostate cancer. Preclinical and clinical research for rare cancers (in the EU defined as less than five cases per 10,000 people[312]) is often scarce due to the lack of available patient tissues and data, and the small potential market. However, all rare diseases add up to 6-8% of all diseases corresponding to 30 million people in the EU[312]. Rare cancers even add up to 22% of all cancers[313], revealing the necessity of putting effort into research for these diseases. To overcome the limitations of clinical trials in rare cancers, one could elaborate international cooperation to increase patient numbers, elongate the time frame these studies are conducted in and increase the awareness of these studies in the affected patient population[314]. However, one of the most promising approaches is the concept of biomarker driven therapies[315], where not the organ that gave rise to the cancer, but the molecular alterations are defining therapeutic interventions. That way, rare cancers can be included in basket trials based on their molecular alterations[316]. The alterations that were identified in this study in ACC give rise to a number of therapies which are already in clinical trials or which are EMA[317]- or FDA[318]-approved for other cancer entities (Table 12) and are discussed in the following section. Due to the similarities of molecular aberrations of metastatic lesions to primary tumors, these approaches are promising to target all cancer sites.

Table 12 Available drugs for molecular alterations in ACC. OC: ovarian cancer, PPC: primary peritoneal cancer, PFTC: primary fallopian tube carcinoma, SEGA: subependymal giant cell astrocytoma, RCC: renal cell carcinoma, BC: breast cancer, mut: mutated

exploits loss of	signaling pathway	molecular mechanism	drug	No. of clinical studies ongoing		Approved	
				cancer.gov[319]	clinicaltrials.gov[320]	FDA[318] approved for	EMA[317] approved for
CDKN2A/ID3/APC	Cyclin D	inhibits CDK4/6	Palbociclib (PD-0332991)	27	102	BC	HR+ HER2- BC
			Ribociclib	18	49	/	/
			abemaciclib	12	35	/	/
ARID1A	mTOR	inhibits mTOR	Sirolimus	38	1942	BC	/
			Everolimus	51	2006	BC, PNET, SEGA, RCC	BC, PNET, RCC
			AZD8055	NA	5	/	/
		inhibits mTOR+PI3K	BEZ235	/	22	/	/
	single stand breaks	inhibits PARP	Olaparib	21	118	OC + BRCA mut.	OC, PFTC, PPC + BRCA mut.
	ARID1B	synthetic lethality	NA	/	/	/	/
	PRC2	inhibits EZH2; synthetic lethality	GSK126	1	1	/	/
			Tazemetostat	3	6	/	/
			CPI-1205	1	1	/	/
		inhibits EED; synthetic lethality	mak683	1	1	/	/
	ATR	synthetic lethality	VX-970	13	9		
			AZD6738	1	4	/	/
	PI3K	PI3K family	Buparlisib	15	89	/	/
		inhibits AKT	perifosine	/	44	/	/
			MK-2206	/	50	/	/
APC	WNT	inhibits the recruiting of CTNNB with CBP	PRI-724	1	6	/	/
		inhibits PORCN, blocks acetylation of WNT ligands	WNT974	3	2	/	/
		competes with FZD8 receptor	OMP-54F28	2	4	/	/
		binds and inhibits FZD	OMP-18R5	1	4	/	/
		inhibits CTNNB	CWP232291	1	2	/	/
CDKN2A	mitotic checkpoint	inhibits AURKs, VEGFRs, and PDGFRs	Ilorasertib	2	2	/	/

“PI3K inhibition, ARID1B, EZH2, and PARP inhibition might all lead to effective therapies in tumors lacking ARID1A[321]. An *ARID1A* model has shown that *ARID1A* loss is not sufficient for tumorigenesis, but requires PI3K pathway activation both of which lead to IL-6 overproduction[322]. This mouse model is responding to PI3K inhibitor treatment. Further evaluation of this pathway in human ACC on the protein level is needed to address this

question. If this hypothesis holds true there are a number of drugs already being tested in clinical trials for other cancer entities which would be useful for treating ACC patients with *ARID1A* deletion and PI3K activation, including mTor inhibitors (everolimus, Sirolimus, AZD8055), AKT inhibitors (perifosine, MK2206), and PI3K inhibitors (Buparlisib).”[1] Indeed, a partial response to everolimus treatment in an ACC patient was previously reported[173,323] and in an ACC mouse model rapamycin treatment stopped tumor growth[324]. “As ARID1A is mutually exclusive with ARID1B in SWI/SNF complexes, *ARID1A* mutated cancer cells get dependent on ARID1B. Inhibition leads to synthetic lethality, *i.e.* leading to cancer cell death whereas healthy cells can compensate ARID1B inhibition[325,326]. Further, EZH2 inhibition seems to be another mechanism of synthetic lethality, as cancer cells which harbor SWI/SNF defects are prone to react to disturbances in EZH2[327]. A number of EZH2 inhibitors are being tested preclinically, four of them are already in clinical trials (GSK126, Tazemetostat, CPI-1205, and MAK683)[319,320]. Finally, ARID1A loss can be exploited by PARP inhibition. Upon DNA damage, ARID1A is recruited to DNA breaks by ATR and processes DNA DSB to ss breaks. Cancer cells lacking ARID1A cannot tolerate a loss of PARP, which is a sensor of DNA single-strand breaks and is involved in base excision repair. PARP inhibition has been” shown to be”[1] effective in the absence of ARID1A *in vitro* for a number of cancer cell lines and *in vivo* with cell line engraftments by Shen, et al. [328].”[1] Whereas “the interaction of ARID1A with ATR can be exploited by ATR inhibitors, leading to apoptosis in ARID1A deficient cells[329].”[1]

The observed genomic instability can also be exploited for therapies. “The PARP inhibitor Olaparib is already approved for *BRCA1/2*-mutated ovarian, fallopian tube, and primary peritoneal cancers (EMA[317]) and breast cancer (FDA[318])”[1], and might therefore be effective in ACC with a loss of BRCA (Figure 11 and [151,182,186,289,290]). In addition to PARP inhibition which has been successful in *BRCA*-mutated cancers, platinum-based chemotherapies have shown promising results. These drugs lead to crosslinking of DNA and subsequent to DNA double-strand breaks, thus cancer cells with a defective DNA DSB repair are more prone to react to these drugs[330]. In fact, two cases of ACC with a mutated *BRCA2* showed a complete and one case a partial response to platinum-based therapies[186,289], while other ACC where the status of *BRCA1/2* was not investigated responded to platinum-based therapies[151,301-303,331-334]. And a xenograft mouse model of ACC with a lack of *Brca2* responded best to oxaliplatin therapy[290]. Mismatch-repair deficient ACC patients might benefit from immunotherapies, as tumors deficient in mismatch repair have been reported to be more prone to present non-self-antigens at their cell surface[335,336].

“The identified aberrations in cell cycle control and chromosomal instability can be exploited therapeutically in different ways with drugs already tested in other tumors in clinical trials or which are already EMA[317]- and/or FDA[318]-approved”[1] (Table 12). “APC, CDKN2A, and ID3 are all negative regulators of the cell cycle”[1] and thus can “be all exploited by molecules interfering with CDK4/6 (*e.g.* Palbociclib, Ribociclib, Abemaciclib)[337]”[1], reviewed in Asghar, et al. [338]. “Palbociclib is a CDK4/6 inhibitor that has finally made its way into the clinics in 2016 and is approved by the EMA[317] and FDA[318] for a subtype of breast cancer. Other CDK4/6 inhibitors (Ribociclib, Abemaciclib) are still being tested in clinical trials. Deletions in *APC* are frequent in colorectal cancer and aberrations on the gene level have previously been reported in ACC[171,177,180,182,183,291], mostly at low frequency and not on the protein level. Here,”[1] it was shown “that the *APC* gene was frequently deleted and this leads to a loss of protein. APC is a member of the WNT pathway, for which currently a number of targeted therapies are being tested in clinical trials (PRI-724, WNT974, OMP-54F28, OMP-185, and CWP232291, for details see Table 1). CDKN2A aberrations have been well studied as they are common in many tumor entities. However it has been difficult to target this aberration therapeutically. Ilorasertib, an Aurora kinase inhibitor is currently being tested in *CDKN2A*-mutated cancers and might therefore be beneficial in ACC. For ID3, therapeutic options apart from CDK4/6 inhibition are lacking so far. However, as ~90% of tumors exhibit aberrations in ID3, such therapies would benefit most ACC patients.”[1]

5.5 Acinar cells might be the cell of origin of ACC

As the cell of origin is unknown in ACC, this issue was further addressed. “Previously it has been shown that tumors resemble the”[1] epigenetic “profile of their cell of origin[339-342], thus one has to be careful when calculating differential methylation to actually identify cancer-related methylation changes.”[1] Cell type contributions based on the methylome of sorted pancreatic cells revealed that the three investigated pancreatic cancers all exhibited the highest cell type contribution of the cell they phenotypically resembled. This hints towards the hypothesis that these cancers arose from these cell types. PDAC samples showed high enrichment of acinar cell methylation patterns in addition to ductal cell methylation patterns, *i.e.* supporting previous studies on the cell of origin where both acinar and duct cells were shown to give rise to PDAC[138,140,141,343]. However, it could be possible that the tumor samples were contaminated with high loads of respective healthy cells, *i.e.* many acinar cells within the ACC samples. This scenario is however unlikely as ACC are very densely growing tumors and were pathologically assessed to contain more than 90% tumor cells. The molecular distances between the tumors and healthy cells were evaluated employing DNA methylation phylogenetic trees and support the hypothesis that PNET arise from endocrine cells. Acinar and duct cells were branching between ACC and PDAC so that a clear statement is not possible.

When considering the cell of origin, one should account for the fact that it might not be a differentiated cell giving rise to tumors in the pancreas. So far, no pluripotent stem cells were identified in the pancreas. Instead, the scientific community supports the hypothesis that differentiated pancreatic cells can dedifferentiate into pluripotent cells which give rise to new cells thus leading to tissue repair within the pancreas[126]. This is mainly described for acinar cells which were shown to give rise to new acinar cells, duct cells, and probably to endocrine cells[127]. Additionally, α- and δ-cells seem to give rise to new β-cells[131,344], whereas ductal cells do not seem to have the capability of regenerating other pancreatic cell types[115,345] (refer to 1.2.2). Thus, these dedifferentiated pluripotent cells are still part of the normal physiological function of the pancreas and this plasticity is reminiscent of the development of pancreatic cells during embryogenesis (see Figure 5). However these pluripotent cells would be more vulnerable for additional hits leading to tumorigenesis, as they already exhibit similar features as cancer cells, *i.e.* an increased proliferation rate. It has been reported that in addition to acinar cell markers and PDX1 transcription factor, most ACC stain positive for the duct cell markers KRT7 and KRT19[146,346], another hint that ACC are immature acinar cells. In fact, transcription factor binding site analysis of DMS in ACC revealed that many transcription factors of early pancreatic

development were enriched (Figure 24), *e.g. FOXA1*, *FOXA2*, *NR5A2*, and *NKX6-1*. This could either be interpreted that acinar cells were the cell of origin and transcription factor programs from embryogenic pathways were activated. However, it could also mean that these transcription factor programs were already activated in the cell of origin, resulting in a pluripotent dedifferentiated acinar cell and stay activated in ACC. Therefore, the enrichment of these transcription factors would only be visible due to the fact that acinar cells were taken as a reference. To dissect this issue more thoroughly, a larger number of sorted pancreatic cells and the isolation of dedifferentiated pluripotent cells would benefit the analyses. For the latter option this would require the establishment of protocols for isolating these cell types. In addition, traditional lineage tracing[347], where the putative cell of origin is marked and traced could be employed for ACC as it was done for PDAC[138,141,348]. In summary, this work showed that acinar cells are the most likely cell of origin for ACC; however this still needs to be validated.

5.6 Further research in ACC is challenged by a number of obstacles

Although this is the first study that encompassed multiple layers of genome-wide data, the generated datasets are not exhaustive. RNA sequencing would be greatly beneficial for correlating the changes of copy numbers, methylation, and point mutations with gene expression. However, “as pancreatic RNA starts to be fragmented by pancreatic enzymes the second the tissue sample is collected, high quality RNA sequencing approaches were not feasible retrospectively.”[1] In addition, a proteomics approach as performed for normal pancreatic tissue[349] is still pending for ACC. This would enable the possibility to screen for even more functional relevant events and give rise to even more potential therapeutic targets. The 3D chromatin architecture of ACC should be investigated as this often dictates whether a gene is activated or not. Calculating open and closed chromatin in cohort I based on the work of Fortin and Hansen [350] generated bins in which the chromatin in ACC is open or closed. However, when trying to confirm these results with cohort II, only 25% of regions showed the same pattern (either stayed closed/open from tumor to normal or changed in the same direction, Supplementary Figure 16). This is merely what was expected from pure chance, thus revealing that this method was not suitable for this dataset. Thus, a more sophisticated approach, *e.g.* Assay for Transposase Accessible Chromatin (ATAC) sequencing should be performed to reveal open and closed chromatin sites[351]. In addition, 5C sequencing could be performed to unravel the interactions of the genome. These two sequencing approaches were not possible with the patient material from this study, as it requires immediately snap frozen tissue to remain the architecture of the chromatin.

The role of molecular alterations of *ARID1A*, *APC*, *CDKN2A*, and *ID3* should be investigated in ACC models to examine whether they play a similar role as in other cancer entities. Only a murine (266-6)[352] and a rat cell line (AR42J)[353] of ACC are currently available, which might not fully reflect the epigenetic and genetic changes of human disease. Thus, a human ACC cell line needs to be established, and 3D cell culture might resolve “this issue as it has been shown for primary pancreatic acinar cells[354].”[1] So far, “published mouse models have the hindrance that they only occur sporadically and do not consistently form”[1] ACC[292,324,355-357].

6 Conclusions and Outlook

“The integration of epigenetic and genetic alterations revealed that pancreatic acinar cell carcinomas are characterized by numerous copy number aberrations and aberrantly methylated sites and regions. They did not show recurrent mutations but displayed distinct mutational signatures. The protein expression of the four tumor suppressor genes *ID3*, *ARID1A*, *APC*, and *CDKN2A* were affected in ~90%, ~70%, ~60-70%, and ~60-70% of acinar cell carcinomas”[1], respectively. “The latter three have been reported as driver genes in tumorigenesis[20], therefore these aberrations have to be considered for the development and progression of this disease.”[1] Several drugs which target these genes or the according pathways “are already on the market or in clinical trials, thus offering new treatment options for ACC patients.”[1] Further studies are needed to evaluate the role of the identified aberrations in tumor initiation and progression, and to address the role of defective DNA repair in these tumors.

This study included sorted pancreatic cell types to investigate the potential cell of origin. Calling of differential methylation has to be carefully considered when working with mixed normal tissue. Here, it was shown that global methylation patterns of ACC were most similar to acinar cells, making the acinar cell the most likely cell of origin. However this hypothesis has to be corroborated, first of all with a greater number of sorted pancreatic cells, ideally from the same patients of the here presented cohorts, and secondly with additional studies, including lineage tracing studies in mice.

A next step for elucidating tumor development of ACC would be to identify precursor lesions. Besides the above mentioned lineage tracing studies, this study gives rise to another approach of identifying lesions. As field cancerization might also exist in ACC, staining of whole slides instead of tumor tissue punches on microarrays might identify non-cancerous regions that already lack one of the here identified tumor suppressor genes. This is of course based on the assumption, that these protein losses are early events already visible during tumor initiation, which is currently unknown.

The occurrence of one hallmark and one enabling characteristic from Weinberg and Hanahan’s proposal[3] were identified in ACC in this study. ACC evade growth suppressors on multiple levels, and genome instability enables these cancers to achieve the basis for obtaining the other hallmarks. However, a cancer cell needs to achieve further hallmarks as mentioned in 1.1.1[3]. Such molecular aberrations still remain to be identified in ACC. The presented genomic and epigenomic landscape can now serve as a basis for this and for further hypothesis building for

research of ACC. This will provide the possibility to dissect the differences between ACC and other pancreatic cancer subtypes. Understanding why acinar cell carcinomas are rare although acinar cells are the most abundant cell type within the pancreas will not only add more insight into ACC development, but add knowledge about pancreatic homeostasis, pancreatic plasticity, and other pancreatic cancers.

Results of this study and studies to come will have to be translated into the clinic. Basket trials, in which not the cancer entity but the underlying molecular alterations determine the inclusion into a trial is a promising approach for ACC. It is however unlikely that all ACC patients will benefit from this, therefore an ACC consortium should be built to recruit enough patients for ACC specific clinical trials.

7 References

1. Submitted Manuscript Jäkel, C. ,Bergmann, F., Toth, R., Assenov, Y., van der Duin, D., Strobel, O., Hank, T., Klöppel, G, Dorrell, C., Grompe, M., Moss, J., Dor, Y., Schirmacher, P., Plass, C., Popanda, O., Schmezer, P. Genome-wide genetic and epigenetic analyses of pancreatic acinar cell carcinomas reveal aberrations in cell cycle control and genome stability. (2017).
2. Hanahan, D. & Weinberg, R.A. The hallmarks of cancer. *Cell* **100**, 57-70 (2000).
3. Hanahan, D. & Weinberg, R.A. Hallmarks of cancer: the next generation. *Cell* **144**, 646-74 (2011).
4. Tarapore, P. & Fukasawa, K. Loss of p53 and centrosome hyperamplification. *Oncogene* **21**, 6234-40 (2002).
5. Giono, L.E. & Manfredi, J.J. The p53 tumor suppressor participates in multiple cell cycle checkpoints. *J Cell Physiol* **209**, 13-20 (2006).
6. Chinnam, M. & Goodrich, D.W. RB1, development, and cancer. *Curr Top Dev Biol* **94**, 129-69 (2011).
7. Weinberg, R.A. The biology of cancer Second Edition. *Garland Science* (2014).
8. Schafer, K.A. The cell cycle: a review. *Vet Pathol* **35**, 461-78 (1998).
9. Zetterberg, A., Larsson, O. & Wiman, K.G. What is the restriction point? *Curr Opin Cell Biol* **7**, 835-42 (1995).
10. el-Deiry, W.S. *et al.* WAF1, a potential mediator of p53 tumor suppression. *Cell* **75**, 817-25 (1993).
11. Liggett, W.H., Jr. & Sidransky, D. Role of the p16 tumor suppressor gene in cancer. *J Clin Oncol* **16**, 1197-206 (1998).
12. Zack, T.I. *et al.* Pan-cancer patterns of somatic copy number alteration. *Nat Genet* **45**, 1134-40 (2013).
13. Kandoth, C. *et al.* Mutational landscape and significance across 12 major cancer types. *Nature* **502**, 333-9 (2013).
14. Sherr, C.J. Cancer cell cycles. *Science* **274**, 1672-7 (1996).
15. Malumbres, M. & Barbacid, M. To cycle or not to cycle: a critical decision in cancer. *Nat Rev Cancer* **1**, 222-31 (2001).
16. Biankin, A.V. *et al.* Pancreatic cancer genomes reveal aberrations in axon guidance pathway genes. *Nature* **491**, 399-405 (2012).
17. Waddell, N. *et al.* Whole genomes redefine the mutational landscape of pancreatic cancer. *Nature* **518**, 495-501 (2015).
18. Lawrence, M.S. *et al.* Mutational heterogeneity in cancer and the search for new cancer-associated genes. *Nature* **499**, 214-8 (2013).
19. Alexandrov, L.B. *et al.* Signatures of mutational processes in human cancer. *Nature* **500**, 415-21 (2013).
20. Vogelstein, B. *et al.* Cancer genome landscapes. *Science* **339**, 1546-58 (2013).
21. Sanger, F. & Coulson, A.R. A rapid method for determining sequences in DNA by primed synthesis with DNA polymerase. *J Mol Biol* **94**, 441-8 (1975).
22. Nik-Zainal, S. *et al.* Mutational processes molding the genomes of 21 breast cancers. *Cell* **149**, 979-93 (2012).
23. Alexandrov, L.B., Nik-Zainal, S., Wedge, D.C., Campbell, P.J. & Stratton, M.R. Deciphering signatures of mutational processes operative in human cancer. *Cell Rep* **3**, 246-59 (2013).
24. Helleday, T., Eshtad, S. & Nik-Zainal, S. Mechanisms underlying mutational signatures in human cancers. *Nat Rev Genet* **15**, 585-98 (2014).
25. Alexandrov, L.B. & Stratton, M.R. Mutational signatures: the patterns of somatic mutations hidden in cancer genomes. *Curr Opin Genet Dev* **24**, 52-60 (2014).

26. http://cancer.sanger.ac.uk/cosmic/signatures.
27. Alexandrov, L.B. Understanding the origins of human cancer. *Science* **350**, 1175 (2015).
28. Alexandrov, L.B. *et al.* Mutational signatures associated with tobacco smoking in human cancer. *Science* **354**, 618-622 (2016).
29. Hollstein, M., Alexandrov, L.B., Wild, C.P., Ardin, M. & Zavadil, J. Base changes in tumour DNA have the power to reveal the causes and evolution of cancer. *Oncogene* **36**, 158-167 (2017).
30. Stephens, P.J. *et al.* Complex landscapes of somatic rearrangement in human breast cancer genomes. *Nature* **462**, 1005-10 (2009).
31. Stratton, M.R., Campbell, P.J. & Futreal, P.A. The cancer genome. *Nature* **458**, 719-24 (2009).
32. Beroukhim, R. *et al.* The landscape of somatic copy-number alteration across human cancers. *Nature* **463**, 899-905 (2010).
33. Ye, K. *et al.* Systematic discovery of complex insertions and deletions in human cancers. *Nat Med* **22**, 97-104 (2016).
34. Artandi, S.E. & DePinho, R.A. Telomeres and telomerase in cancer. *Carcinogenesis* **31**, 9-18 (2010).
35. Murnane, J.P. Telomere loss as a mechanism for chromosome instability in human cancer. *Cancer Res* **70**, 4255-9 (2010).
36. Maciejowski, J. & de Lange, T. Telomeres in cancer: tumour suppression and genome instability. *Nat Rev Mol Cell Biol* **18**, 175-186 (2017).
37. Popescu, N.C. & Zimonjic, D.B. Molecular cytogenetic characterization of cancer cell alterations. *Cancer Genet Cytogenet* **93**, 10-21 (1997).
38. Feber, A. *et al.* Using high-density DNA methylation arrays to profile copy number alterations. *Genome Biol* **15**, R30 (2014).
39. Van Loo, P. *et al.* Analyzing cancer samples with SNP arrays. *Methods Mol Biol* **802**, 57-72 (2012).
40. Abel, H.J. & Duncavage, E.J. Detection of structural DNA variation from next generation sequencing data: a review of informatic approaches. *Cancer Genet* **206**, 432-40 (2013).
41. Wu, C. & Morris, J.R. Genes, genetics, and epigenetics: a correspondence. *Science* **293**, 1103-5 (2001).
42. Holliday, R. & Pugh, J.E. DNA modification mechanisms and gene activity during development. *Science* **187**, 226-32 (1975).
43. Riggs, A.D. X inactivation, differentiation, and DNA methylation. *Cytogenet Cell Genet* **14**, 9-25 (1975).
44. Ehrlich, M., Norris, K.F., Wang, R.Y., Kuo, K.C. & Gehrke, C.W. DNA cytosine methylation and heat-induced deamination. *Biosci Rep* **6**, 387-93 (1986).
45. CpG island track UCSC browser https://genome.ucsc.edu/cgi-bin/hgTrackUi?hgsid=383138943_cuVjAzBaxUkx2fiCOkIfa81Sjr6U&c=chr21&g=cpgIslandSuper&cpgIslandSuper=show.
46. Naveh-Many, T. & Cedar, H. Active gene sequences are undermethylated. *Proc Natl Acad Sci U S A* **78**, 4246-50 (1981).
47. Baubec, T. & Schubeler, D. Genomic patterns and context specific interpretation of DNA methylation. *Curr Opin Genet Dev* **25**, 85-92 (2014).
48. Watt, F. & Molloy, P.L. Cytosine methylation prevents binding to DNA of a HeLa cell transcription factor required for optimal expression of the adenovirus major late promoter. *Genes Dev* **2**, 1136-43 (1988).
49. Neri, F. *et al.* Intragenic DNA methylation prevents spurious transcription initiation. *Nature* **543**, 72-77 (2017).
50. Aran, D. & Hellman, A. DNA methylation of transcriptional enhancers and cancer predisposition. *Cell* **154**, 11-3 (2013).

51. Charlet, J. *et al.* Bivalent Regions of Cytosine Methylation and H3K27 Acetylation Suggest an Active Role for DNA Methylation at Enhancers. *Mol Cell* **62**, 422-31 (2016).
52. Roadmap Epigenomics, C. *et al.* Integrative analysis of 111 reference human epigenomes. *Nature* **518**, 317-30 (2015).
53. Schultz, M.D. *et al.* Human body epigenome maps reveal noncanonical DNA methylation variation. *Nature* **523**, 212-6 (2015).
54. Ziller, M.J. *et al.* Charting a dynamic DNA methylation landscape of the human genome. *Nature* **500**, 477-81 (2013).
55. Moran, S. *et al.* Epigenetic profiling to classify cancer of unknown primary: a multicentre, retrospective analysis. *Lancet Oncol* **17**, 1386-1395 (2016).
56. Witte, T., Plass, C. & Gerhauser, C. Pan-cancer patterns of DNA methylation. *Genome Med* **6**, 66 (2014).
57. Nones, K. *et al.* Genome-wide DNA methylation patterns in pancreatic ductal adenocarcinoma reveal epigenetic deregulation of SLIT-ROBO, ITGA2 and MET signaling. *Int J Cancer* **135**, 1110-8 (2014).
58. Schutte, M. *et al.* Abrogation of the Rb/p16 tumor-suppressive pathway in virtually all pancreatic carcinomas. *Cancer Res* **57**, 3126-30 (1997).
59. Costello, J.F. & Plass, C. Methylation matters. *J Med Genet* **38**, 285-303 (2001).
60. Esteller, M. CpG island hypermethylation and tumor suppressor genes: a booming present, a brighter future. *Oncogene* **21**, 5427-40 (2002).
61. Goeppert, B. *et al.* Down-regulation of tumor suppressor A kinase anchor protein 12 in human hepatocarcinogenesis by epigenetic mechanisms. *Hepatology* **52**, 2023-33 (2010).
62. Dutruel, C. *et al.* Early epigenetic downregulation of WNK2 kinase during pancreatic ductal adenocarcinoma development. *Oncogene* **33**, 3401-10 (2014).
63. Zhou, V.W., Goren, A. & Bernstein, B.E. Charting histone modifications and the functional organization of mammalian genomes. *Nat Rev Genet* **12**, 7-18 (2011).
64. Ernst, J. & Kellis, M. Discovery and characterization of chromatin states for systematic annotation of the human genome. *Nat Biotechnol* **28**, 817-25 (2010).
65. Ernst, J. & Kellis, M. Large-scale imputation of epigenomic datasets for systematic annotation of diverse human tissues. *Nat Biotechnol* **33**, 364-76 (2015).
66. Wolffe, A.P. Chromatin remodeling: why it is important in cancer. *Oncogene* **20**, 2988-90 (2001).
67. Plass, C. *et al.* Mutations in regulators of the epigenome and their connections to global chromatin patterns in cancer. *Nat Rev Genet* **14**, 765-80 (2013).
68. Feinberg, A.P., Koldobskiy, M.A. & Gondor, A. Epigenetic modulators, modifiers and mediators in cancer aetiology and progression. *Nat Rev Genet* **17**, 284-99 (2016).
69. Lujambio, A. & Lowe, S.W. The microcosmos of cancer. *Nature* **482**, 347-55 (2012).
70. Ha, M. & Kim, V.N. Regulation of microRNA biogenesis. *Nat Rev Mol Cell Biol* **15**, 509-24 (2014).
71. Sotillo, E. & Thomas-Tikhonenko, A. The long reach of noncoding RNAs. *Nat Genet* **43**, 616-7 (2011).
72. Baer, C. *et al.* Extensive promoter DNA hypermethylation and hypomethylation is associated with aberrant microRNA expression in chronic lymphocytic leukemia. *Cancer Res* **72**, 3775-85 (2012).
73. Baer, C., Claus, R. & Plass, C. Genome-wide epigenetic regulation of miRNAs in cancer. *Cancer Res* **73**, 473-7 (2013).
74. Croce, C.M. Causes and consequences of microRNA dysregulation in cancer. *Nat Rev Genet* **10**, 704-14 (2009).
75. Goeppert, B. *et al.* Cadherin-6 is a putative tumor suppressor and target of epigenetically dysregulated miR-429 in cholangiocarcinoma. *Epigenetics* **11**, 780-790 (2016).

76. Arab, K. *et al.* Long noncoding RNA TARID directs demethylation and activation of the tumor suppressor TCF21 via GADD45A. *Mol Cell* **55**, 604-14 (2014).
77. Weiss, M., Plass, C. & Gerhauser, C. Role of lncRNAs in prostate cancer development and progression. *Biol Chem* **395**, 1275-90 (2014).
78. Sturm, D. *et al.* Hotspot mutations in H3F3A and IDH1 define distinct epigenetic and biological subgroups of glioblastoma. *Cancer Cell* **22**, 425-37 (2012).
79. Setlow, R.B. & Carrier, W.L. Pyrimidine dimers in ultraviolet-irradiated DNA's. *J Mol Biol* **17**, 237-54 (1966).
80. Howard, B.D. & Tessman, I. Identification of the Altered Bases in Mutated Single-Stranded DNA. Ii. In Vivo Mutagenesis by 5-Bromodeoxyuridine and 2-Aminopurine. *J Mol Biol* **9**, 364-71 (1964).
81. Greenman, C. *et al.* Patterns of somatic mutation in human cancer genomes. *Nature* **446**, 153-8 (2007).
82. Parsons, D.W. *et al.* An integrated genomic analysis of human glioblastoma multiforme. *Science* **321**, 1807-12 (2008).
83. Ravanat, J.L. *et al.* Radiation-mediated formation of complex damage to DNA: a chemical aspect overview. *Br J Radiol* **87**, 20130715 (2014).
84. Boland, C.R. & Goel, A. Microsatellite instability in colorectal cancer. *Gastroenterology* **138**, 2073-2087 e3 (2010).
85. Krokan, H.E. & Bjoras, M. Base excision repair. *Cold Spring Harb Perspect Biol* **5**, a012583 (2013).
86. Scharer, O.D. Nucleotide excision repair in eukaryotes. *Cold Spring Harb Perspect Biol* **5**, a012609 (2013).
87. Khanna, K.K. & Jackson, S.P. DNA double-strand breaks: signaling, repair and the cancer connection. *Nat Genet* **27**, 247-54 (2001).
88. Branzei, D. & Foiani, M. Regulation of DNA repair throughout the cell cycle. *Nat Rev Mol Cell Biol* **9**, 297-308 (2008).
89. Stiff, T. *et al.* ATM and DNA-PK function redundantly to phosphorylate H2AX after exposure to ionizing radiation. *Cancer Res* **64**, 2390-6 (2004).
90. Paull, T.T. *et al.* A critical role for histone H2AX in recruitment of repair factors to nuclear foci after DNA damage. *Curr Biol* **10**, 886-95 (2000).
91. Fernandez-Capetillo, O., Lee, A., Nussenzweig, M. & Nussenzweig, A. H2AX: the histone guardian of the genome. *DNA Repair (Amst)* **3**, 959-67 (2004).
92. Genois, M.M. *et al.* Interactions between BRCA2 and RAD51 for promoting homologous recombination in Leishmania infantum. *Nucleic Acids Res* **40**, 6570-84 (2012).
93. Jensen, R.B., Carreira, A. & Kowalczykowski, S.C. Purified human BRCA2 stimulates RAD51-mediated recombination. *Nature* **467**, 678-83 (2010).
94. Li, X. & Heyer, W.D. Homologous recombination in DNA repair and DNA damage tolerance. *Cell Res* **18**, 99-113 (2008).
95. Davis, A.J. & Chen, D.J. DNA double strand break repair via non-homologous end-joining. *Transl Cancer Res* **2**, 130-143 (2013).
96. Schermelleh, L. *et al.* Dynamics of Dnmt1 interaction with the replication machinery and its role in postreplicative maintenance of DNA methylation. *Nucleic Acids Res* **35**, 4301-12 (2007).
97. Alabert, C. & Groth, A. Chromatin replication and epigenome maintenance. *Nat Rev Mol Cell Biol* **13**, 153-67 (2012).
98. Dabin, J., Fortuny, A. & Polo, S.E. Epigenome Maintenance in Response to DNA Damage. *Mol Cell* **62**, 712-27 (2016).
99. Hyun, Y. *et al.* The catalytic subunit of Arabidopsis DNA polymerase alpha ensures stable maintenance of histone modification. *Development* **140**, 156-66 (2013).

100. Lowe, M., Hostager, R. & Kikyo, N. Preservation of Epigenetic Memory During DNA Replication. *J Stem Cell Res Ther (Edmond)* **1**(2016).
101. Luijsterburg, M.S. *et al.* PARP1 Links CHD2-Mediated Chromatin Expansion and H3.3 Deposition to DNA Repair by Non-homologous End-Joining. *Mol Cell* **61**, 547-62 (2016).
102. Russo, G. *et al.* DNA damage and Repair Modify DNA methylation and Chromatin Domain of the Targeted Locus: Mechanism of allele methylation polymorphism. *Sci Rep* **6**, 33222 (2016).
103. Hoeijmakers, J.H. Genome maintenance mechanisms for preventing cancer. *Nature* **411**, 366-74 (2001).
104. Jeggo, P.A., Pearl, L.H. & Carr, A.M. DNA repair, genome stability and cancer: a historical perspective. *Nat Rev Cancer* **16**, 35-42 (2016).
105. Chae, Y.K. *et al.* Genomic landscape of DNA repair genes in cancer. *Oncotarget* **7**, 23312-21 (2016).
106. Chaisaingmongkol, J. *et al.* Epigenetic screen of human DNA repair genes identifies aberrant promoter methylation of NEIL1 in head and neck squamous cell carcinoma. *Oncogene* **31**, 5108-16 (2012).
107. Kuhmann, C. *et al.* Altered regulation of DNA ligase IV activity by aberrant promoter DNA methylation and gene amplification in colorectal cancer. *Hum Mol Genet* **23**, 2043-54 (2014).
108. Thielmann, H.W., Popanda, O., Edler, L. & Jung, E.G. Clinical symptoms and DNA repair characteristics of xeroderma pigmentosum patients from Germany. *Cancer Res* **51**, 3456-70 (1991).
109. Hartman, A.R. *et al.* Prevalence of BRCA mutations in an unselected population of triple-negative breast cancer. *Cancer* **118**, 2787-95 (2012).
110. Cavanagh, H. & Rogers, K.M. The role of BRCA1 and BRCA2 mutations in prostate, pancreatic and stomach cancers. *Hered Cancer Clin Pract* **13**, 16 (2015).
111. Shih, H.P., Wang, A. & Sander, M. Pancreas organogenesis: from lineage determination to morphogenesis. *Annu Rev Cell Dev Biol* **29**, 81-105 (2013).
112. Jennings, R.E., Berry, A.A., Strutt, J.P., Gerrard, D.T. & Hanley, N.A. Human pancreas development. *Development* **142**, 3126-37 (2015).
113. Arda, H.E., Benitez, C.M. & Kim, S.K. Gene regulatory networks governing pancreas development. *Dev Cell* **25**, 5-13 (2013).
114. Schaffer, A.E., Freude, K.K., Nelson, S.B. & Sander, M. Nkx6 transcription factors and Ptf1a function as antagonistic lineage determinants in multipotent pancreatic progenitors. *Dev Cell* **18**, 1022-9 (2010).
115. Kopp, J.L. *et al.* Sox9+ ductal cells are multipotent progenitors throughout development but do not produce new endocrine cells in the normal or injured adult pancreas. *Development* **138**, 653-65 (2011).
116. Zhou, Q. *et al.* A multipotent progenitor domain guides pancreatic organogenesis. *Dev Cell* **13**, 103-14 (2007).
117. Apelqvist, A. *et al.* Notch signalling controls pancreatic cell differentiation. *Nature* **400**, 877-81 (1999).
118. Afelik, S. & Jensen, J. Notch signaling in the pancreas: patterning and cell fate specification. *Wiley Interdiscip Rev Dev Biol* **2**, 531-44 (2013).
119. Shih, H.P. *et al.* A Notch-dependent molecular circuitry initiates pancreatic endocrine and ductal cell differentiation. *Development* **139**, 2488-99 (2012).
120. Masui, T. *et al.* Replacement of Rbpj with Rbpjl in the PTF1 complex controls the final maturation of pancreatic acinar cells. *Gastroenterology* **139**, 270-80 (2010).
121. Delous, M. *et al.* Sox9b is a key regulator of pancreaticobiliary ductal system development. *PLoS Genet* **8**, e1002754 (2012).

122. Johansson, K.A. *et al.* Temporal control of neurogenin3 activity in pancreas progenitors reveals competence windows for the generation of different endocrine cell types. *Dev Cell* **12**, 457-65 (2007).
123. Ziv, O., Glaser, B. & Dor, Y. The plastic pancreas. *Dev Cell* **26**, 3-7 (2013).
124. Bramswig, N.C. *et al.* Epigenomic plasticity enables human pancreatic alpha to beta cell reprogramming. *J Clin Invest* **123**, 1275-84 (2013).
125. Dorrell, C. *et al.* Transcriptomes of the major human pancreatic cell types. *Diabetologia* **54**, 2832-44 (2011).
126. Kopp, J.L., Grompe, M. & Sander, M. Stem cells versus plasticity in liver and pancreas regeneration. *Nat Cell Biol* **18**, 238-45 (2016).
127. Pan, F.C. *et al.* Spatiotemporal patterns of multipotentiality in Ptf1a-expressing cells during pancreas organogenesis and injury-induced facultative restoration. *Development* **140**, 751-64 (2013).
128. Li, W. *et al.* In vivo reprogramming of pancreatic acinar cells to three islet endocrine subtypes. *Elife* **3**, e01846 (2014).
129. Zhou, Q., Brown, J., Kanarek, A., Rajagopal, J. & Melton, D.A. In vivo reprogramming of adult pancreatic exocrine cells to beta-cells. *Nature* **455**, 627-32 (2008).
130. von Figura, G., Morris, J.P.t., Wright, C.V. & Hebrok, M. Nr5a2 maintains acinar cell differentiation and constrains oncogenic Kras-mediated pancreatic neoplastic initiation. *Gut* **63**, 656-64 (2014).
131. Thorel, F. *et al.* Conversion of adult pancreatic alpha-cells to beta-cells after extreme beta-cell loss. *Nature* **464**, 1149-54 (2010).
132. Chera, S. *et al.* Diabetes recovery by age-dependent conversion of pancreatic delta-cells into insulin producers. *Nature* **514**, 503-7 (2014).
133. Dhawan, S., Georgia, S., Tschen, S.I., Fan, G. & Bhushan, A. Pancreatic beta cell identity is maintained by DNA methylation-mediated repression of Arx. *Dev Cell* **20**, 419-29 (2011).
134. Gao, T. *et al.* Pdx1 maintains beta cell identity and function by repressing an alpha cell program. *Cell Metab* **19**, 259-71 (2014).
135. Papizan, J.B. *et al.* Nkx2.2 repressor complex regulates islet beta-cell specification and prevents beta-to-alpha-cell reprogramming. *Genes Dev* **25**, 2291-305 (2011).
136. Schaffer, A.E. *et al.* Nkx6.1 controls a gene regulatory network required for establishing and maintaining pancreatic Beta cell identity. *PLoS Genet* **9**, e1003274 (2013).
137. Wollny, D. *et al.* Single-Cell Analysis Uncovers Clonal Acinar Cell Heterogeneity in the Adult Pancreas. *Dev Cell* **39**, 289-301 (2016).
138. Habbe, N. *et al.* Spontaneous induction of murine pancreatic intraepithelial neoplasia (mPanIN) by acinar cell targeting of oncogenic Kras in adult mice. *Proc Natl Acad Sci U S A* **105**, 18913-8 (2008).
139. Maitra, A. & Leach, S.D. Disputed paternity: the uncertain ancestry of pancreatic ductal neoplasia. *Cancer Cell* **22**, 701-3 (2012).
140. De La, O.J. *et al.* Notch and Kras reprogram pancreatic acinar cells to ductal intraepithelial neoplasia. *Proc Natl Acad Sci U S A* **105**, 18907-12 (2008).
141. Kopp, J.L. *et al.* Identification of Sox9-dependent acinar-to-ductal reprogramming as the principal mechanism for initiation of pancreatic ductal adenocarcinoma. *Cancer Cell* **22**, 737-50 (2012).
142. Pin, C.L., Ryan, J.F. & Mehmood, R. Acinar cell reprogramming: a clinically important target in pancreatic disease. *Epigenomics* **7**, 267-81 (2015).

143. Morris, J.P.t., Cano, D.A., Sekine, S., Wang, S.C. & Hebrok, M. Beta-catenin blocks Kras-dependent reprogramming of acini into pancreatic cancer precursor lesions in mice. *J Clin Invest* **120**, 508-20 (2010).
144. http://globocan.iarc.fr/.
145. World wide incidence of pancreatic cancer, International Agency for Research on Cancer, World Health Organization http://gco.iarc.fr/today/online-analysis-map?mode=cancer&mode_population=continents&population=900&sex=0&cancer=9&type=0&statistic=0&prevalence=0&color_palette=default&projection=natural-earth.
146. La Rosa, S. *et al.* Clinicopathologic study of 62 acinar cell carcinomas of the pancreas: insights into the morphology and immunophenotype and search for prognostic markers. *Am J Surg Pathol* **36**, 1782-95 (2012).
147. Bergmann, F. *et al.* Acinar cell carcinomas of the pancreas: a molecular analysis in a series of 57 cases. *Virchows Arch* **465**, 661-72 (2014).
148. Wood, L.D. & Klimstra, D.S. Pathology and genetics of pancreatic neoplasms with acinar differentiation. *Semin Diagn Pathol* **31**, 491-7 (2014).
149. Klimstra, D.S., Heffess, C.S., Oertel, J.E. & Rosai, J. Acinar cell carcinoma of the pancreas. A clinicopathologic study of 28 cases. *Am J Surg Pathol* **16**, 815-37 (1992).
150. Klimstra D, H.R., Klöppel G, Morohoshi T, and Ohike N. . Acinar cell neoplasms of the pancreas. *WHO classification of tumours of the digestive system, Bosman, FT, et al. (Eds.)* **Lyon**, 314-318 (2010).
151. Lowery, M.A. *et al.* Acinar cell carcinoma of the pancreas: new genetic and treatment insights into a rare malignancy. *Oncologist* **16**, 1714-20 (2011).
152. Wisnoski, N.C., Townsend, C.M., Jr., Nealon, W.H., Freeman, J.L. & Riall, T.S. 672 patients with acinar cell carcinoma of the pancreas: a population-based comparison to pancreatic adenocarcinoma. *Surgery* **144**, 141-8 (2008).
153. Schmidt, C.M. *et al.* Acinar cell carcinoma of the pancreas in the United States: prognostic factors and comparison to ductal adenocarcinoma. *J Gastrointest Surg* **12**, 2078-86 (2008).
154. Kitagami, H. *et al.* Acinar cell carcinoma of the pancreas: clinical analysis of 115 patients from Pancreatic Cancer Registry of Japan Pancreas Society. *Pancreas* **35**, 42-6 (2007).
155. Butturini, G. *et al.* Aggressive approach to acinar cell carcinoma of the pancreas: a single-institution experience and a literature review. *Langenbecks Arch Surg* **396**, 363-9 (2011).
156. Hartwig, W. *et al.* Acinar cell carcinoma of the pancreas: is resection justified even in limited metastatic disease? *Am J Surg* **202**, 23-7 (2011).
157. Ellerkamp, V., Warmann, S.W., Vorwerk, P., Leuschner, I. & Fuchs, J. Exocrine pancreatic tumors in childhood in Germany. *Pediatr Blood Cancer* **58**, 366-71 (2012).
158. Dall'igna, P. *et al.* Pancreatic tumors in children and adolescents: the Italian TREP project experience. *Pediatr Blood Cancer* **54**, 675-80 (2010).
159. Shorter, N.A., Glick, R.D., Klimstra, D.S., Brennan, M.F. & Laquaglia, M.P. Malignant pancreatic tumors in childhood and adolescence: The Memorial Sloan-Kettering experience, 1967 to present. *J Pediatr Surg* **37**, 887-92 (2002).
160. Lack, E.E., Cassady, J.R., Levey, R. & Vawter, G.F. Tumors of the exocrine pancreas in children and adolescents. A clinical and pathologic study of eight cases. *Am J Surg Pathol* **7**, 319-27 (1983).
161. Hackeng, W.M., Hruban, R.H., Offerhaus, G.J. & Brosens, L.A. Surgical and molecular pathology of pancreatic neoplasms. *Diagn Pathol* **11**, 47 (2016).
162. Klimstra, D.S. & Adsay, V. Acinar neoplasms of the pancreas-A summary of 25 years of research. *Semin Diagn Pathol* **33**, 307-18 (2016).

163. Burns, W.A. *et al.* Lipase-secreting acinar cell carcinoma of the pancreas with polyarthropathy. A light and electron microscopic, histochemical, and biochemical study. *Cancer* **33**, 1002-9 (1974).
164. La Rosa, S., Sessa, F. & Capella, C. Acinar Cell Carcinoma of the Pancreas: Overview of Clinicopathologic Features and Insights into the Molecular Pathology. *Front Med (Lausanne)* **2**, 41 (2015).
165. Stelow, E.B., Shaco-Levy, R., Bao, F., Garcia, J. & Klimstra, D.S. Pancreatic acinar cell carcinomas with prominent ductal differentiation: Mixed acinar ductal carcinoma and mixed acinar endocrine ductal carcinoma. *Am J Surg Pathol* **34**, 510-8 (2010).
166. Tanaka, T., Mori, H. & Williams, G.M. Atypical and neoplastic acinar cell lesions of the pancreas in an autopsy study of Japanese patients. *Cancer* **61**, 2278-85 (1988).
167. Longnecker, D.S., Shinozuka, H. & Dekker, A. Focal acinar cell dysplasia in human pancreas. *Cancer* **45**, 534-40 (1980).
168. Oertel, J.E. The pancreas. Nonneoplastic alterations. *Am J Surg Pathol* **13 Suppl 1**, 50-65 (1989).
169. Shinozuka, H., Lee, R.E., Dunn, J.L. & Longnecker, D.S. Multiple atypical acinar cell nodules of the pancreas. *Hum Pathol* **11**, 389-91 (1980).
170. Stamm, B.H. Incidence and diagnostic significance of minor pathologic changes in the adult pancreas at autopsy: a systematic study of 112 autopsies in patients without known pancreatic disease. *Hum Pathol* **15**, 677-83 (1984).
171. Abraham, S.C. *et al.* Genetic and immunohistochemical analysis of pancreatic acinar cell carcinoma: frequent allelic loss on chromosome 11p and alterations in the APC/beta-catenin pathway. *Am J Pathol* **160**, 953-62 (2002).
172. Hoorens, A. *et al.* Pancreatic acinar cell carcinoma. An analysis of cell lineage markers, p53 expression, and Ki-ras mutation. *Am J Pathol* **143**, 685-98 (1993).
173. de Wilde, R.F. *et al.* Analysis of LKB1 mutations and other molecular alterations in pancreatic acinar cell carcinoma. *Mod Pathol* **24**, 1229-36 (2011).
174. Moore, P.S. *et al.* Pancreatic tumours: molecular pathways implicated in ductal cancer are involved in ampullary but not in exocrine nonductal or endocrine tumorigenesis. *Br J Cancer* **84**, 253-62 (2001).
175. Terhune, P.G., Memoli, V.A. & Longnecker, D.S. Evaluation of p53 mutation in pancreatic acinar cell carcinomas of humans and transgenic mice. *Pancreas* **16**, 6-12 (1998).
176. Rigaud, G. *et al.* Allelotype of pancreatic acinar cell carcinoma. *Int J Cancer* **88**, 772-7 (2000).
177. Furlan, D. *et al.* APC alterations are frequently involved in the pathogenesis of acinar cell carcinoma of the pancreas, mainly through gene loss and promoter hypermethylation. *Virchows Arch* **464**, 553-64 (2014).
178. Pellegata, N.S. *et al.* K-ras and p53 gene mutations in pancreatic cancer: ductal and nonductal tumors progress through different genetic lesions. *Cancer Res* **54**, 1556-60 (1994).
179. La Rosa, S. *et al.* TP53 alterations in pancreatic acinar cell carcinoma: new insights into the molecular pathology of this rare cancer. *Virchows Arch* **468**, 289-96 (2016).
180. Chmielecki, J. *et al.* Comprehensive genomic profiling of pancreatic acinar cell carcinomas identifies recurrent RAF fusions and frequent inactivation of DNA repair genes. *Cancer Discov* **4**, 1398-405 (2014).
181. Ross, J.S. *et al.* The distribution of BRAF gene fusions in solid tumors and response to targeted therapy. *Int J Cancer* **138**, 881-90 (2016).
182. Jiao, Y. *et al.* Whole-exome sequencing of pancreatic neoplasms with acinar differentiation. *J Pathol* **232**, 428-35 (2014).
183. Guo, M. *et al.* Epigenetic changes associated with neoplasms of the exocrine and endocrine pancreas. *Discov Med* **17**, 67-73 (2014).

184. Clarke, A.R., Cummings, M.C. & Harrison, D.J. Interaction between murine germline mutations in p53 and APC predisposes to pancreatic neoplasia but not to increased intestinal malignancy. *Oncogene* **11**, 1913-20 (1995).
185. Oghamian, S. *et al.* Reduction of pancreatic acinar cell tumor multiplicity in Dnmt1 hypomorphic mice. *Carcinogenesis* **32**, 829-35 (2011).
186. Furukawa, T. *et al.* Whole exome sequencing reveals recurrent mutations in BRCA2 and FAT genes in acinar cell carcinomas of the pancreas. *Sci Rep* **5**, 8829 (2015).
187. Taruscio, D. *et al.* Pancreatic acinar carcinoma shows a distinct pattern of chromosomal imbalances by comparative genomic hybridization. *Genes Chromosomes Cancer* **28**, 294-9 (2000).
188. Liu, W. *et al.* DNA mismatch repair abnormalities in acinar cell carcinoma of the pancreas: frequency and clinical significance. *Pancreas* **43**, 1264-70 (2014).
189. Zell, J.A., Rhee, J.M., Ziogas, A., Lipkin, S.M. & Anton-Culver, H. Race, socioeconomic status, treatment, and survival time among pancreatic cancer cases in California. *Cancer Epidemiol Biomarkers Prev* **16**, 546-52 (2007).
190. Bosman, F.T., Carneiro, F., Hruban, R.H. & Theise, N.D. WHO Classification of Tumours of the Digestive System, Fourth Edition. *International Agency for Research on Cancer (IARC)* (2010).
191. Yeo, C.J. *et al.* Pancreaticoduodenectomy for cancer of the head of the pancreas. 201 patients. *Ann Surg* **221**, 721-31; discussion 731-3 (1995).
192. Ducreux, M. *et al.* Cancer of the pancreas: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. *Ann Oncol* **26 Suppl 5**, v56-68 (2015).
193. Ghaneh, P., Costello, E. & Neoptolemos, J.P. Biology and management of pancreatic cancer. *Postgrad Med J* **84**, 478-97 (2008).
194. Baines, A.T., Xu, D. & Der, C.J. Inhibition of Ras for cancer treatment: the search continues. *Future Med Chem* **3**, 1787-808 (2011).
195. Stephen, A.G., Esposito, D., Bagni, R.K. & McCormick, F. Dragging ras back in the ring. *Cancer Cell* **25**, 272-81 (2014).
196. Moore, M.J. *et al.* Erlotinib plus gemcitabine compared with gemcitabine alone in patients with advanced pancreatic cancer: a phase III trial of the National Cancer Institute of Canada Clinical Trials Group. *J Clin Oncol* **25**, 1960-6 (2007).
197. Philip, P.A. Targeted therapies for pancreatic cancer. *Gastrointest Cancer Res* **2**, S16-9 (2008).
198. Brosens, L.A., Hackeng, W.M., Offerhaus, G.J., Hruban, R.H. & Wood, L.D. Pancreatic adenocarcinoma pathology: changing "landscape". *J Gastrointest Oncol* **6**, 358-74 (2015).
199. Bardeesy, N. & DePinho, R.A. Pancreatic cancer biology and genetics. *Nat Rev Cancer* **2**, 897-909 (2002).
200. Bailey, P. *et al.* Genomic analyses identify molecular subtypes of pancreatic cancer. *Nature* **531**, 47-52 (2016).
201. Notta, F. *et al.* A renewed model of pancreatic cancer evolution based on genomic rearrangement patterns. *Nature* **538**, 378-382 (2016).
202. Brembeck, F.H. *et al.* The mutant K-ras oncogene causes pancreatic periductal lymphocytic infiltration and gastric mucous neck cell hyperplasia in transgenic mice. *Cancer Res* **63**, 2005-9 (2003).
203. Houbracken, I. *et al.* Lineage tracing evidence for transdifferentiation of acinar to duct cells and plasticity of human pancreas. *Gastroenterology* **141**, 731-41, 741 e1-4 (2011).
204. Strobel, O. *et al.* In vivo lineage tracing defines the role of acinar-to-ductal transdifferentiation in inflammatory ductal metaplasia. *Gastroenterology* **133**, 1999-2009 (2007).

205. Shi, C. *et al.* KRAS2 mutations in human pancreatic acinar-ductal metaplastic lesions are limited to those with PanIN: implications for the human pancreatic cancer cell of origin. *Mol Cancer Res* **7**, 230-6 (2009).
206. Shi, G. *et al.* Maintenance of acinar cell organization is critical to preventing Kras-induced acinar-ductal metaplasia. *Oncogene* **32**, 1950-8 (2013).
207. Ramage, J.K. *et al.* Guidelines for the management of gastroenteropancreatic neuroendocrine (including carcinoid) tumours (NETs). *Gut* **61**, 6-32 (2012).
208. Plockinger, U. *et al.* Guidelines for the diagnosis and treatment of neuroendocrine gastrointestinal tumours. A consensus statement on behalf of the European Neuroendocrine Tumour Society (ENETS). *Neuroendocrinology* **80**, 394-424 (2004).
209. Jiao, Y. *et al.* DAXX/ATRX, MEN1, and mTOR pathway genes are frequently altered in pancreatic neuroendocrine tumors. *Science* **331**, 1199-203 (2011).
210. Corbo, V. *et al.* MEN1 in pancreatic endocrine tumors: analysis of gene and protein status in 169 sporadic neoplasms reveals alterations in the vast majority of cases. *Endocr Relat Cancer* **17**, 771-83 (2010).
211. Missiaglia, E. *et al.* Pancreatic endocrine tumors: expression profiling evidences a role for AKT-mTOR pathway. *J Clin Oncol* **28**, 245-55 (2010).
212. Phan, A.T. & Dave, B. The pivotal role of mammalian target of rapamycin inhibition in the treatment of patients with neuroendocrine tumors. *Cancer Med* **5**, 2953-2964 (2016).
213. Valle, J.W. *et al.* A systematic review of non-surgical treatments for pancreatic neuroendocrine tumours. *Cancer Treat Rev* **40**, 376-89 (2014).
214. Scarpa, A. *et al.* Whole-genome landscape of pancreatic neuroendocrine tumours. *Nature* (2017).
215. Levy-Bohbot, N. *et al.* Prevalence, characteristics and prognosis of MEN 1-associated glucagonomas, VIPomas, and somatostatinomas: study from the GTE (Groupe des Tumeurs Endocrines) registry. *Gastroenterol Clin Biol* **28**, 1075-81 (2004).
216. Thakker, R.V. *et al.* Clinical practice guidelines for multiple endocrine neoplasia type 1 (MEN1). *J Clin Endocrinol Metab* **97**, 2990-3011 (2012).
217. Tonelli, F., Giudici, F., Fratini, G. & Brandi, M.L. Pancreatic endocrine tumors in multiple endocrine neoplasia type 1 syndrome: review of literature. *Endocr Pract* **17 Suppl 3**, 33-40 (2011).
218. Cupisti, K. *et al.* Lack of MEN1 gene mutations in 27 sporadic insulinomas. *Eur J Clin Invest* **30**, 325-9 (2000).
219. Zhuang, Z. *et al.* Somatic mutations of the MEN1 tumor suppressor gene in sporadic gastrinomas and insulinomas. *Cancer Res* **57**, 4682-6 (1997).
220. Zhao, J. *et al.* Genomic imbalances in the progression of endocrine pancreatic tumors. *Genes Chromosomes Cancer* **32**, 364-72 (2001).
221. Speel, E.J. *et al.* Genetic differences in endocrine pancreatic tumor subtypes detected by comparative genomic hybridization. *Am J Pathol* **155**, 1787-94 (1999).
222. Stumpf, E. *et al.* Chromosomal alterations in human pancreatic endocrine tumors. *Genes Chromosomes Cancer* **29**, 83-7 (2000).
223. Gebauer, N. *et al.* Genomic landscape of pancreatic neuroendocrine tumors. *World J Gastroenterol* **20**, 17498-506 (2014).
224. Lu, J. *et al.* Alpha cell-specific Men1 ablation triggers the transdifferentiation of glucagon-expressing cells and insulinoma development. *Gastroenterology* **138**, 1954-65 (2010).
225. Vortmeyer, A.O., Huang, S., Lubensky, I. & Zhuang, Z. Non-islet origin of pancreatic islet cell tumors. *J Clin Endocrinol Metab* **89**, 1934-8 (2004).

226. Van Allen, E.M. *et al.* Whole-exome sequencing and clinical interpretation of formalin-fixed, paraffin-embedded tumor samples to guide precision cancer medicine. *Nat Med* **20**, 682-8 (2014).
227. Ehrich, M. *et al.* Quantitative high-throughput analysis of DNA methylation patterns by base-specific cleavage and mass spectrometry. *Proc Natl Acad Sci U S A* **102**, 15785-90 (2005).
228. Shyr, C. *et al.* FLAGS, frequently mutated genes in public exomes. *BMC Med Genomics* **7**, 64 (2014).
229. Gehring, J.S., Fischer, B., Lawrence, M. & Huber, W. SomaticSignatures: inferring mutational signatures from single-nucleotide variants. *Bioinformatics* **31**, 3673-5 (2015).
230. Gehring, J. SomaticCancerAlterations: Somatic Cancer Alterations. *R package version 1.8.2.* (2016).
231. Rosenthal, R., McGranahan, N., Herrero, J., Taylor, B.S. & Swanton, C. DeconstructSigs: delineating mutational processes in single tumors distinguishes DNA repair deficiencies and patterns of carcinoma evolution. *Genome Biol* **17**, 31 (2016).
232. Aran, D., Sirota, M. & Butte, A.J. Systematic pan-cancer analysis of tumour purity. *Nat Commun* **6**, 8971 (2015).
233. Desper, R. & Gascuel, O. Fast and accurate phylogeny reconstruction algorithms based on the minimum-evolution principle. *J Comput Biol* **9**, 687-705 (2002).
234. Paradis, E., Claude, J. & Strimmer, K. APE: Analyses of Phylogenetics and Evolution in R language. *Bioinformatics* **20**, 289-90 (2004).
235. Assenov, Y. *et al.* Comprehensive analysis of DNA methylation data with RnBeads. *Nat Methods* **11**, 1138-40 (2014).
236. Teschendorff, A.E. *et al.* A beta-mixture quantile normalization method for correcting probe design bias in Illumina Infinium 450 k DNA methylation data. *Bioinformatics* **29**, 189-96 (2013).
237. Ensemble 73; http://sep2013.archive.ensembl.org/.
238. Houseman, E.A. *et al.* DNA methylation arrays as surrogate measures of cell mixture distribution. *BMC Bioinformatics* **13**, 86 (2012).
239. Reinius, L.E. *et al.* Differential DNA methylation in purified human blood cells: implications for cell lineage and studies on disease susceptibility. *PLoS One* **7**, e41361 (2012).
240. Huang da, W., Sherman, B.T. & Lempicki, R.A. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc* **4**, 44-57 (2009).
241. http://homer.salk.edu/homer/.
242. Hovestadt, V. & Zapatka, M. conumee: Enhanced copy-number variation analysis using Illumina 450k methylation arrays. R package version 0.99.4 http://www.bioconductor.org/packages/release/bioc/html/conumee.html. (2016).
243. Mermel, C.H. *et al.* GISTIC2.0 facilitates sensitive and confident localization of the targets of focal somatic copy-number alteration in human cancers. *Genome Biol* **12**, R41 (2011).
244. https://cran.r-project.org/web/packages/circlize/index.html.
245. Mo, Q. & Shen, R. iClusterPlus: Integrative clustering of multi-type genomic data. R package version 1.9.0. (2013).
246. Yoshihara, K. *et al.* Inferring tumour purity and stromal and immune cell admixture from expression data. *Nat Commun* **4**, 2612 (2013).
247. Carter, S.L. *et al.* Absolute quantification of somatic DNA alterations in human cancer. *Nat Biotechnol* **30**, 413-21 (2012).
248. Chen, W.V. & Maniatis, T. Clustered protocadherins. *Development* **140**, 3297-302 (2013).
249. Wang, X., Su, H. & Bradley, A. Molecular mechanisms governing Pcdh-gamma gene expression: evidence for a multiple promoter and cis-alternative splicing model. *Genes Dev* **16**, 1890-905 (2002).

250. Toyoda, S. *et al.* Developmental epigenetic modification regulates stochastic expression of clustered protocadherin genes, generating single neuron diversity. *Neuron* **82**, 94-108 (2014).
251. Yagi, T. Molecular codes for neuronal individuality and cell assembly in the brain. *Front Mol Neurosci* **5**, 45 (2012).
252. Schreiner, D. & Weiner, J.A. Combinatorial homophilic interaction between gamma-protocadherin multimers greatly expands the molecular diversity of cell adhesion. *Proc Natl Acad Sci U S A* **107**, 14893-8 (2010).
253. Abe, M. *et al.* CpG island methylator phenotype is a strong determinant of poor prognosis in neuroblastomas. *Cancer Res* **65**, 828-34 (2005).
254. Novak, P. *et al.* Agglomerative epigenetic aberrations are a common event in human breast cancer. *Cancer Res* **68**, 8616-25 (2008).
255. Dallosso, A.R. *et al.* Frequent long-range epigenetic silencing of protocadherin gene clusters on chromosome 5q31 in Wilms' tumor. *PLoS Genet* **5**, e1000745 (2009).
256. Dallosso, A.R. *et al.* Long-range epigenetic silencing of chromosome 5q31 protocadherins is involved in early and late stages of colorectal tumorigenesis through modulation of oncogenic pathways. *Oncogene* **31**, 4409-19 (2012).
257. Wang, K.H. *et al.* Global methylation silencing of clustered proto-cadherin genes in cervical cancer: serving as diagnostic markers comparable to HPV. *Cancer Med* **4**, 43-55 (2015).
258. Li, H. *et al.* Immune regulation by low doses of the DNA methyltransferase inhibitor 5-azacitidine in common human epithelial cancers. *Oncotarget* **5**, 587-98 (2014).
259. Kawaguchi, M. *et al.* Relationship between DNA methylation states and transcription of individual isoforms encoded by the protocadherin-alpha gene cluster. *J Biol Chem* **283**, 12064-75 (2008).
260. Morris, T.J. *et al.* ChAMP: 450k Chip Analysis Methylation Pipeline. *Bioinformatics* **30**, 428-30 (2014).
261. An, O., Dall'Olio, G.M., Mourikis, T.P. & Ciccarelli, F.D. NCG 5.0: updates of a manually curated repository of cancer genes and associated properties from cancer mutational screenings. *Nucleic Acids Res* **44**, D992-9 (2016).
262. Zhao, M., Kim, P., Mitra, R., Zhao, J. & Zhao, Z. TSGene 2.0: an updated literature-based knowledgebase for tumor suppressor genes. *Nucleic Acids Res* **44**, D1023-31 (2016).
263. Humphris, J.L. *et al.* Hypermutation In Pancreatic Cancer. *Gastroenterology* **152**, 68-74 e2 (2017).
264. Ohike, N., Kosmahl, M. & Kloppel, G. Mixed acinar-endocrine carcinoma of the pancreas. A clinicopathological study and comparison with acinar-cell carcinoma. *Virchows Arch* **445**, 231-5 (2004).
265. Clark, C.E. *et al.* Dynamics of the immune reaction to pancreatic cancer from inception to invasion. *Cancer Res* **67**, 9518-27 (2007).
266. Shibuya, K.C. *et al.* Pancreatic ductal adenocarcinoma contains an effector and regulatory immune cell infiltrate that is altered by multimodal neoadjuvant treatment. *PLoS One* **9**, e96565 (2014).
267. Zhang, Y., McAllister, F. & Pasca di Magliano, M. Immune cells in pancreatic cancer: Joining the dark side. *Oncoimmunology* **3**, e29125 (2014).
268. Consortium, E.P. An integrated encyclopedia of DNA elements in the human genome. *Nature* **489**, 57-74 (2012).
269. Andersson, R. *et al.* An atlas of active enhancers across human cell types and tissues. *Nature* **507**, 455-61 (2014).
270. Moran, S., Arribas, C. & Esteller, M. Validation of a DNA methylation microarray for 850,000 CpG sites of the human genome enriched in enhancer sequences. *Epigenomics* **8**, 389-99 (2016).

271. Wang, Q. *et al.* Tagmentation-based whole-genome bisulfite sequencing. *Nat Protoc* **8**, 2022-32 (2013).
272. Chiang, C.W. *et al.* Rapid assessment of genetic ancestry in populations of unknown origin by genome-wide genotyping of pooled samples. *PLoS Genet* **6**, e1000866 (2010).
273. Klein, C.A. Parallel progression of primary tumours and metastases. *Nat Rev Cancer* **9**, 302-12 (2009).
274. Makohon-Moore, A.P. *et al.* Limited heterogeneity of known driver gene mutations among the metastases of individual patients with pancreatic cancer. *Nat Genet* (2017).
275. Cong, L. *et al.* Multiplex genome engineering using CRISPR/Cas systems. *Science* **339**, 819-23 (2013).
276. Tsai, S.Q. *et al.* Dimeric CRISPR RNA-guided FokI nucleases for highly specific genome editing. *Nat Biotechnol* **32**, 569-76 (2014).
277. Gilbert, L.A. *et al.* CRISPR-mediated modular RNA-guided regulation of transcription in eukaryotes. *Cell* **154**, 442-51 (2013).
278. Xu, X. *et al.* A CRISPR-based approach for targeted DNA demethylation. *Cell Discov* **2**, 16009 (2016).
279. Liu, X.S. *et al.* Editing DNA Methylation in the Mammalian Genome. *Cell* **167**, 233-247 e17 (2016).
280. Franko, J. *et al.* Loss of heterozygosity predicts poor survival after resection of pancreatic adenocarcinoma. *J Gastrointest Surg* **12**, 1664-72; discussion 1672-3 (2008).
281. Lee, S.H. *et al.* The Id3/E47 axis mediates cell-cycle control in human pancreatic ducts and adenocarcinoma. *Mol Cancer Res* **9**, 782-90 (2011).
282. Dufresne, M. *et al.* Id3 modulates cellular localization of bHLH Ptf1-p48 protein. *Int J Cancer* **129**, 295-306 (2011).
283. Watanabe, R. *et al.* SWI/SNF factors required for cellular resistance to DNA damage include ARID1A and ARID1B and show interdependent protein stability. *Cancer Res* **74**, 2465-75 (2014).
284. Dykhuizen, E.C. *et al.* BAF complexes facilitate decatenation of DNA by topoisomerase IIalpha. *Nature* **497**, 624-7 (2013).
285. http://cancer.sanger.ac.uk/cosmic/gene/analysis?ln=ARID1A.
286. Yamamoto, S., Tsuda, H., Takano, M., Tamai, S. & Matsubara, O. Loss of ARID1A protein expression occurs as an early event in ovarian clear-cell carcinoma development and frequently coexists with PIK3CA mutations. *Mod Pathol* **25**, 615-24 (2012).
287. Samartzis, E.P., Noske, A., Dedes, K.J., Fink, D. & Imesch, P. ARID1A mutations and PI3K/AKT pathway alterations in endometriosis and endometriosis-associated ovarian carcinomas. *Int J Mol Sci* **14**, 18824-49 (2013).
288. Zang, Z.J. *et al.* Exome sequencing of gastric adenocarcinoma identifies recurrent somatic mutations in cell adhesion and chromatin remodeling genes. *Nat Genet* **44**, 570-4 (2012).
289. Ploquin, A. *et al.* Prolonged Survival in a Patient with a Pancreatic Acinar Cell Carcinoma. *Case Rep Oncol* **8**, 447-50 (2015).
290. Hall, J.C. *et al.* Novel patient-derived xenograft mouse model for pancreatic acinar cell carcinoma demonstrates single agent activity of oxaliplatin. *J Transl Med* **14**, 129 (2016).
291. Dewald, G.W. *et al.* Fluorescence in situ hybridization to visualize genetic abnormalities in interphase cells of acinar cell carcinoma, ductal adenocarcinoma, and islet cell carcinoma of the pancreas. *Mayo Clin Proc* **84**, 801-10 (2009).
292. Skoulidis, F. *et al.* Germline Brca2 heterozygosity promotes Kras(G12D) -driven carcinogenesis in a murine model of familial pancreatic cancer. *Cancer Cell* **18**, 499-509 (2010).
293. Schmitt-Graeff, A. *et al.* Coordinated expression of cyclin D1 and LEF-1/TCF transcription factor is restricted to a subset of hepatocellular carcinoma. *Liver Int* **25**, 839-47 (2005).

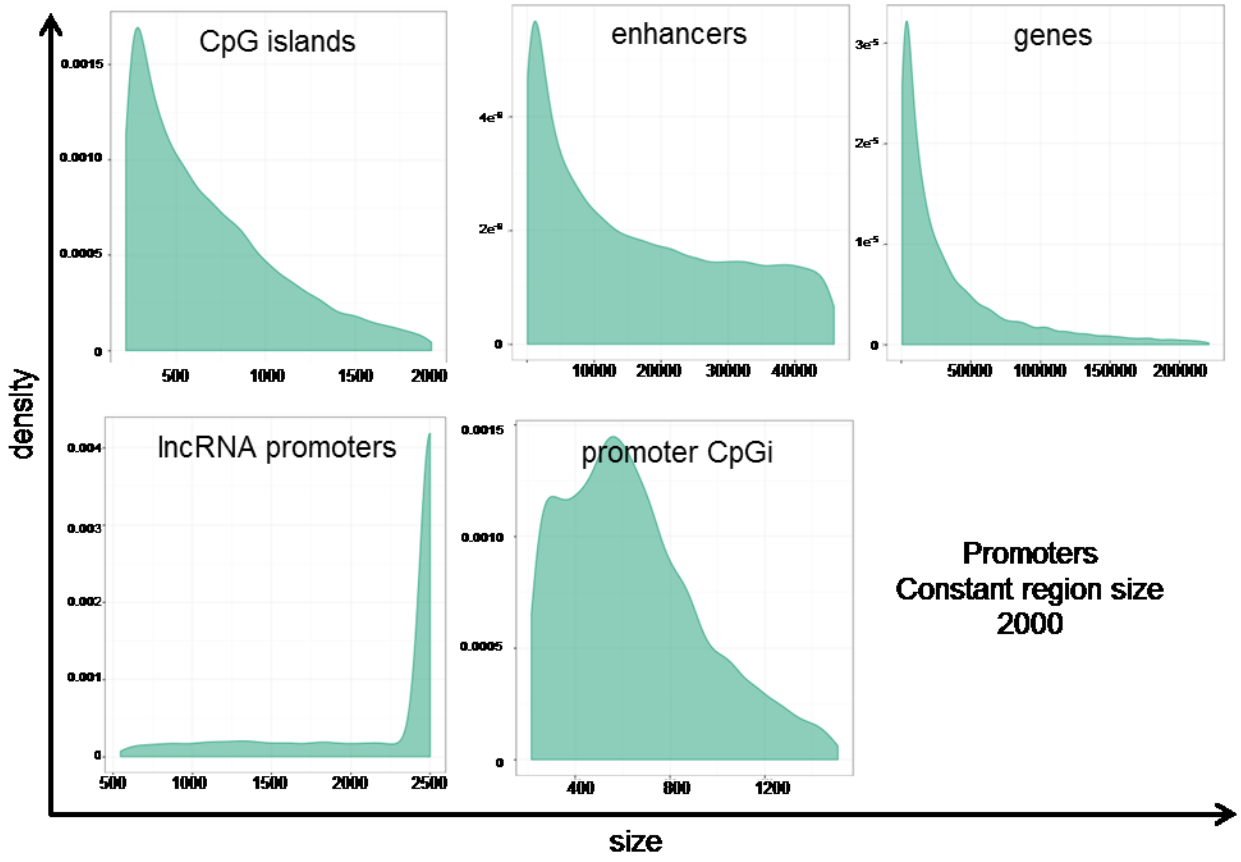
294. Klein, E.A. & Assoian, R.K. Transcriptional regulation of the cyclin D1 gene at a glance. *J Cell Sci* **121**, 3853-7 (2008).
295. Romagosa, C. *et al.* p16(Ink4a) overexpression in cancer: a tumor suppressor gene associated with senescence and high-grade tumors. *Oncogene* **30**, 2087-97 (2011).
296. Miyasaka, Y. *et al.* Senescence in intraductal papillary mucinous neoplasm of the pancreas. *Hum Pathol* **42**, 2010-7 (2011).
297. O'Brien, C.A. *et al.* ID1 and ID3 regulate the self-renewal capacity of human colon cancer-initiating cells through p21. *Cancer Cell* **21**, 777-92 (2012).
298. Schmitz, R. *et al.* Burkitt lymphoma pathogenesis and therapeutic targets from structural and functional genomics. *Nature* **490**, 116-20 (2012).
299. Simbulan-Rosenthal, C.M. *et al.* Id3 induces a caspase-3- and -9-dependent apoptosis and mediates UVB sensitization of HPV16 E6/7 immortalized human keratinocytes. *Oncogene* **25**, 3649-60 (2006).
300. Kim, S. *et al.* The basic helix-loop-helix transcription factor E47 reprograms human pancreatic cancer cells to a quiescent acinar state with reduced tumorigenic potential. *Pancreas* **44**, 718-27 (2015).
301. Kruger, S. *et al.* Acinar cell carcinoma of the pancreas: a rare disease with different diagnostic and therapeutic implications than ductal adenocarcinoma. *J Cancer Res Clin Oncol* **142**, 2585-2591 (2016).
302. Holen, K.D. *et al.* Clinical characteristics and outcomes from an institutional series of acinar cell carcinoma of the pancreas and related tumors. *J Clin Oncol* **20**, 4673-8 (2002).
303. Yoo, C. *et al.* Efficacy of Chemotherapy in Patients with Unresectable or Metastatic Pancreatic Acinar Cell Carcinoma: Potentially Improved Efficacy with Oxaliplatin-containing Regimen. *Cancer Res Treat* (2016).
304. Seki, Y., Okusaka, T., Ikeda, M., Morizane, C. & Ueno, H. Four cases of pancreatic acinar cell carcinoma treated with gemcitabine or S-1 as a single agent. *Jpn J Clin Oncol* **39**, 751-5 (2009).
305. Matos, J.M. *et al.* Pancreatic acinar cell carcinoma: a multi-institutional study. *J Gastrointest Surg* **13**, 1495-502 (2009).
306. Seth, A.K. *et al.* Acinar cell carcinoma of the pancreas: an institutional series of resected patients and review of the current literature. *J Gastrointest Surg* **12**, 1061-7 (2008).
307. Glazer, E.S. *et al.* Systematic Review and Case Series Report of Acinar Cell Carcinoma of the Pancreas. *Cancer Control* **23**, 446-454 (2016).
308. Lee, J.L. *et al.* Locally advanced acinar cell carcinoma of the pancreas successfully treated by capecitabine and concurrent radiotherapy: report of two cases. *Pancreas* **27**, e18-22 (2003).
309. Chen, C.P. *et al.* Concurrent chemoradiation is effective in the treatment of alpha-fetoprotein-producing acinar cell carcinoma of the pancreas: report of a case. *Pancreas* **22**, 326-9 (2001).
310. Kobayashi, S. *et al.* Acinar cell carcinoma of the pancreas successfully treated by en bloc resection and intraperitoneal chemotherapy for peritoneal relapse: a case report of a 15-year survivor. *Pancreas* **23**, 109-12 (2001).
311. Antoine, M., Khitrik-Palchuk, M. & Saif, M.W. Long-term survival in a patient with acinar cell carcinoma of pancreas. A case report and review of literature. *JOP* **8**, 783-9 (2007).
312. EURORDIS - Rare Diseases Europe - Council Recommendation of 8 June 2009 on “An action in the field of Rare Diseases”.
313. http://www.rarecancerseurope.org/.
314. Panageas, K.S. Clinical trial design for rare cancers: why a less conventional route may be required. *Expert Rev Clin Pharmacol* **8**, 661-3 (2015).
315. Simon, R. & Roychowdhury, S. Implementing personalized cancer genomics in clinical trials. *Nat Rev Drug Discov* **12**, 358-69 (2013).

316. Billingham, L., Malottki, K. & Steven, N. Research methods to change clinical practice for patients with rare cancers. *Lancet Oncol* **17**, e70-80 (2016).
317. http://www.ema.europa.eu/ema/.
318. https://www.fda.gov/.
319. https://www.cancer.gov/publications/dictionaries/cancer-drug.
320. https://www.clinicaltrials.gov/.
321. Samartzis, E.P. *et al.* Loss of ARID1A expression sensitizes cancer cells to PI3K- and AKT-inhibition. *Oncotarget* **5**, 5295-303 (2014).
322. Chandler, R.L. *et al.* Coexistent ARID1A-PIK3CA mutations promote ovarian clear-cell tumorigenesis through pro-tumorigenic inflammatory cytokine signalling. *Nat Commun* **6**, 6118 (2015).
323. Klumpen, H.J. *et al.* mTOR inhibitor treatment of pancreatic cancer in a patient With Peutz-Jeghers syndrome. *J Clin Oncol* **29**, e150-3 (2011).
324. Ding, L. *et al.* Neurogenin 3-directed cre deletion of Tsc1 gene causes pancreatic acinar carcinoma. *Neoplasia* **16**, 909-17 (2014).
325. Helming, K.C. *et al.* ARID1B is a specific vulnerability in ARID1A-mutant cancers. *Nat Med* **20**, 251-4 (2014).
326. Witkiewicz, A.K. *et al.* Whole-exome sequencing of pancreatic cancer defines genetic diversity and therapeutic targets. *Nat Commun* **6**, 6744 (2015).
327. Bitler, B.G. *et al.* Synthetic lethality by targeting EZH2 methyltransferase activity in ARID1A-mutated cancers. *Nat Med* **21**, 231-8 (2015).
328. Shen, J. *et al.* ARID1A Deficiency Impairs the DNA Damage Checkpoint and Sensitizes Cells to PARP Inhibitors. *Cancer Discov* **5**, 752-67 (2015).
329. Williamson, C.T. *et al.* ATR inhibitors as a synthetic lethal therapy for tumours deficient in ARID1A. *Nat Commun* **7**, 13837 (2016).
330. Helleday, T., Petermann, E., Lundin, C., Hodgson, B. & Sharma, R.A. DNA repair pathways as targets for cancer therapy. *Nat Rev Cancer* **8**, 193-204 (2008).
331. Simon, M., Bioulac-Sage, P., Trillaud, H. & Blanc, J.F. FOLFOX regimen in pancreatic acinar cell carcinoma: case report and review of the literature. *Acta Oncol* **51**, 403-5 (2012).
332. Schempf, U. *et al.* FOLFIRINOX as first-line treatment for unresectable acinar cell carcinoma of the pancreas: a case report. *Z Gastroenterol* **52**, 200-3 (2014).
333. Callata-Carhuapoma, H.R. *et al.* Pancreatic acinar cell carcinoma with bilateral ovarian metastases, panniculitis and polyarthritis treated with FOLFIRINOX chemotherapy regimen. A case report and review of the literature. *Pancreatology* **15**, 440-4 (2015).
334. Pfrommer, S. *et al.* Successful Salvage Chemotherapy with FOLFIRINOX for Recurrent Mixed Acinar Cell Carcinoma and Ductal Adenocarcinoma of the Pancreas in an Adolescent Patient. *Case Rep Oncol* **6**, 497-503 (2013).
335. Le, D.T. *et al.* PD-1 Blockade in Tumors with Mismatch-Repair Deficiency. *N Engl J Med* **372**, 2509-20 (2015).
336. Blocking PD-1 in Tumors with Faulty DNA Repair. *Cancer Discov* **6**, OF6 (2016).
337. O'Leary, B., Finn, R.S. & Turner, N.C. Treating cancer with selective CDK4/6 inhibitors. *Nat Rev Clin Oncol* **13**, 417-30 (2016).
338. Asghar, U., Witkiewicz, A.K., Turner, N.C. & Knudsen, E.S. The history and future of targeting cyclin-dependent kinases in cancer therapy. *Nat Rev Drug Discov* **14**, 130-46 (2015).
339. Oakes, C.C. *et al.* Evolution of DNA methylation is linked to genetic aberrations in chronic lymphocytic leukemia. *Cancer Discov* **4**, 348-61 (2014).
340. Sproul, D. *et al.* Tissue of origin determines cancer-associated CpG island promoter hypermethylation patterns. *Genome Biol* **13**, R84 (2012).

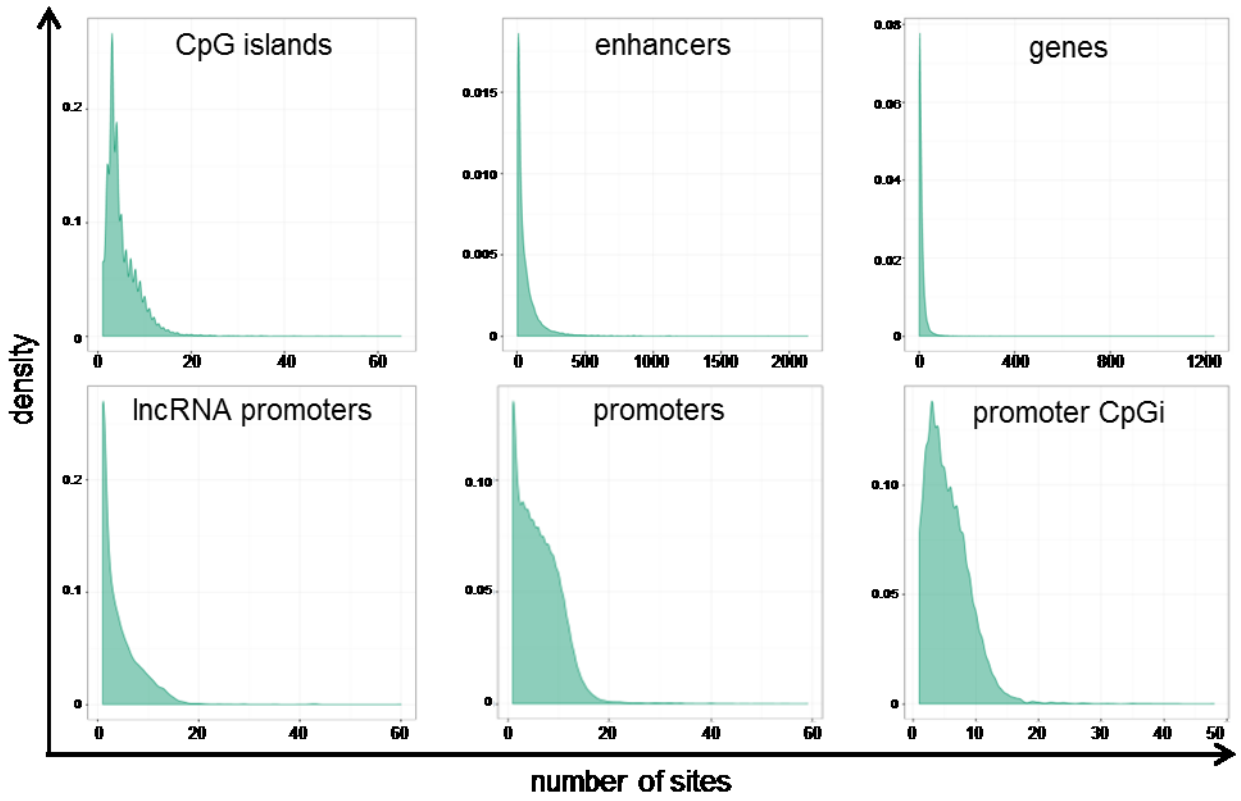
341. Lehmann-Werman, R. *et al.* Identification of tissue-specific cell death using methylation patterns of circulating DNA. *Proc Natl Acad Sci U S A* **113**, E1826-34 (2016).
342. Polak, P. *et al.* Cell-of-origin chromatin organization shapes the mutational landscape of cancer. *Nature* **518**, 360-4 (2015).
343. Ray, K.C. *et al.* Epithelial tissues have varying degrees of susceptibility to Kras(G12D)-initiated tumorigenesis in a mouse model. *PLoS One* **6**, e16786 (2011).
344. Collombat, P. *et al.* The ectopic expression of Pax4 in the mouse pancreas converts progenitor cells into alpha and subsequently beta cells. *Cell* **138**, 449-62 (2009).
345. Solar, M. *et al.* Pancreatic exocrine duct cells give rise to insulin-producing beta cells during embryogenesis but not after birth. *Dev Cell* **17**, 849-60 (2009).
346. Park, J.Y. *et al.* Pdx1 expression in pancreatic precursor lesions and neoplasms. *Appl Immunohistochem Mol Morphol* **19**, 444-9 (2011).
347. Blanpain, C. Tracing the cellular origin of cancer. *Nat Cell Biol* **15**, 126-34 (2013).
348. Bailey, J.M. *et al.* DCLK1 marks a morphologically distinct subpopulation of cells with stem cell properties in preinvasive pancreatic cancer. *Gastroenterology* **146**, 245-56 (2014).
349. Kim, M.S. *et al.* A draft map of the human proteome. *Nature* **509**, 575-81 (2014).
350. Fortin, J.P. & Hansen, K.D. Reconstructing A/B compartments as revealed by Hi-C using long-range correlations in epigenetic data. *Genome Biol* **16**, 180 (2015).
351. Buenrostro, J.D., Giresi, P.G., Zaba, L.C., Chang, H.Y. & Greenleaf, W.J. Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. *Nat Methods* **10**, 1213-8 (2013).
352. Ornitz, D.M. *et al.* Elastase I promoter directs expression of human growth hormone and SV40 T antigen genes to pancreatic acinar cells in transgenic mice. *Cold Spring Harb Symp Quant Biol* **50**, 399-409 (1985).
353. Longnecker, D.S., Lilja, H.S., French, J., Kuhlmann, E. & Noll, W. Transplantation of azaserine-induced carcinomas of pancreas in rats. *Cancer Lett* **7**, 197-202 (1979).
354. Qu, C. & Konieczny, S.F. Pancreatic Acinar Cell 3-Dimensional Culture. *Bio Protoc* **3**(2013).
355. Zhang, N., Lyons, S., Lim, E. & Lassota, P. A spontaneous acinar cell carcinoma model for monitoring progression of pancreatic lesions and response to treatment through noninvasive bioluminescence imaging. *Clin Cancer Res* **15**, 4915-24 (2009).
356. Azzopardi, S., Pang, S., Klimstra, D.S. & Du, Y.N. p53 and p16Ink4a/p19Arf Loss Promotes Different Pancreatic Tumor Types from PyMT-Expressing Progenitor Cells. *Neoplasia* **18**, 610-617 (2016).
357. Kong, B. *et al.* Pancreas-specific activation of mTOR and loss of p53 induce tumors reminiscent of acinar cell carcinoma. *Mol Cancer* **14**, 212 (2015).

8 Appendix

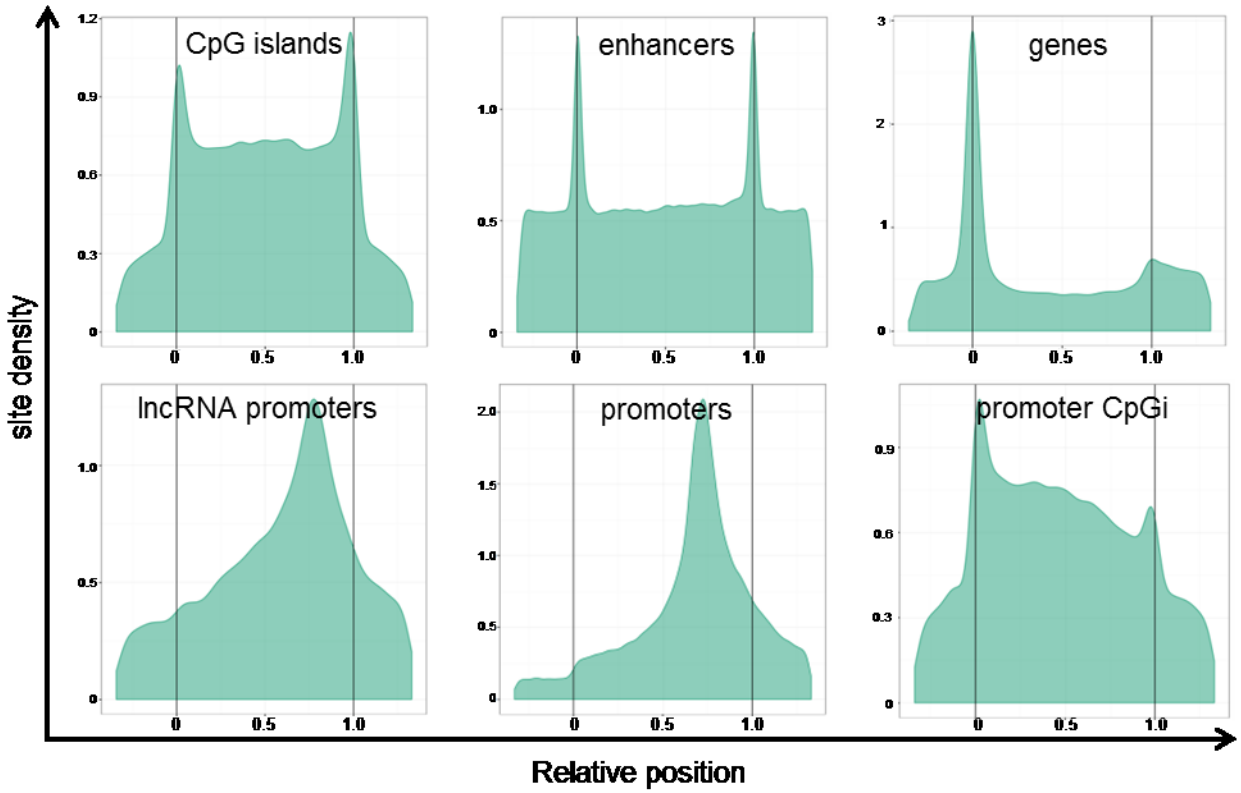
8.1 Supplementary Figures



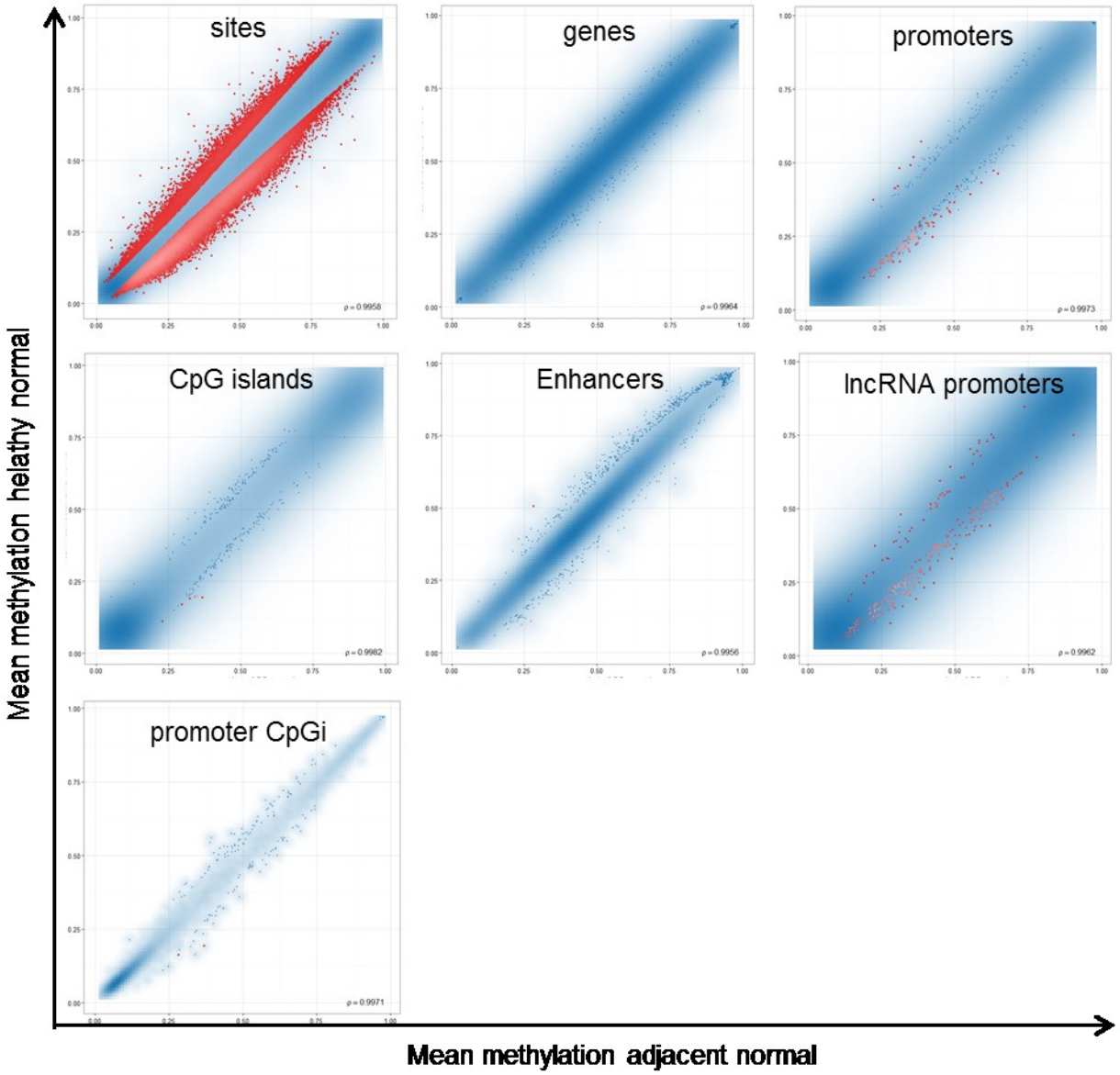
Supplementary Figure 1 Region length distributions of region annotations used for methylation analysis. For definition of region annotations refer to 3.2.2.5.



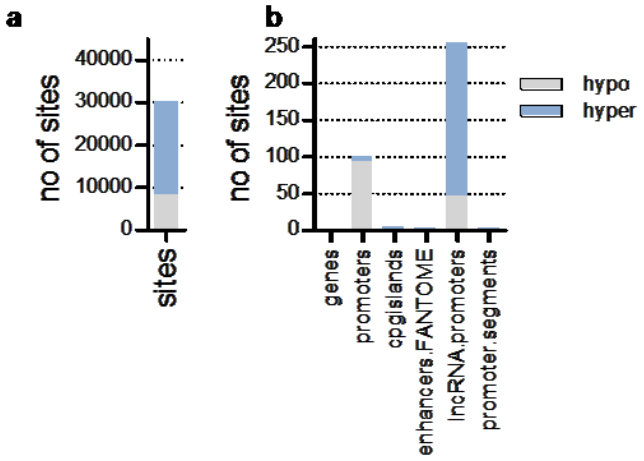
Supplementary Figure 2 Number of CpG sites per region annotation used for methylation analysis. For definition of region annotations refer to 3.2.2.5.



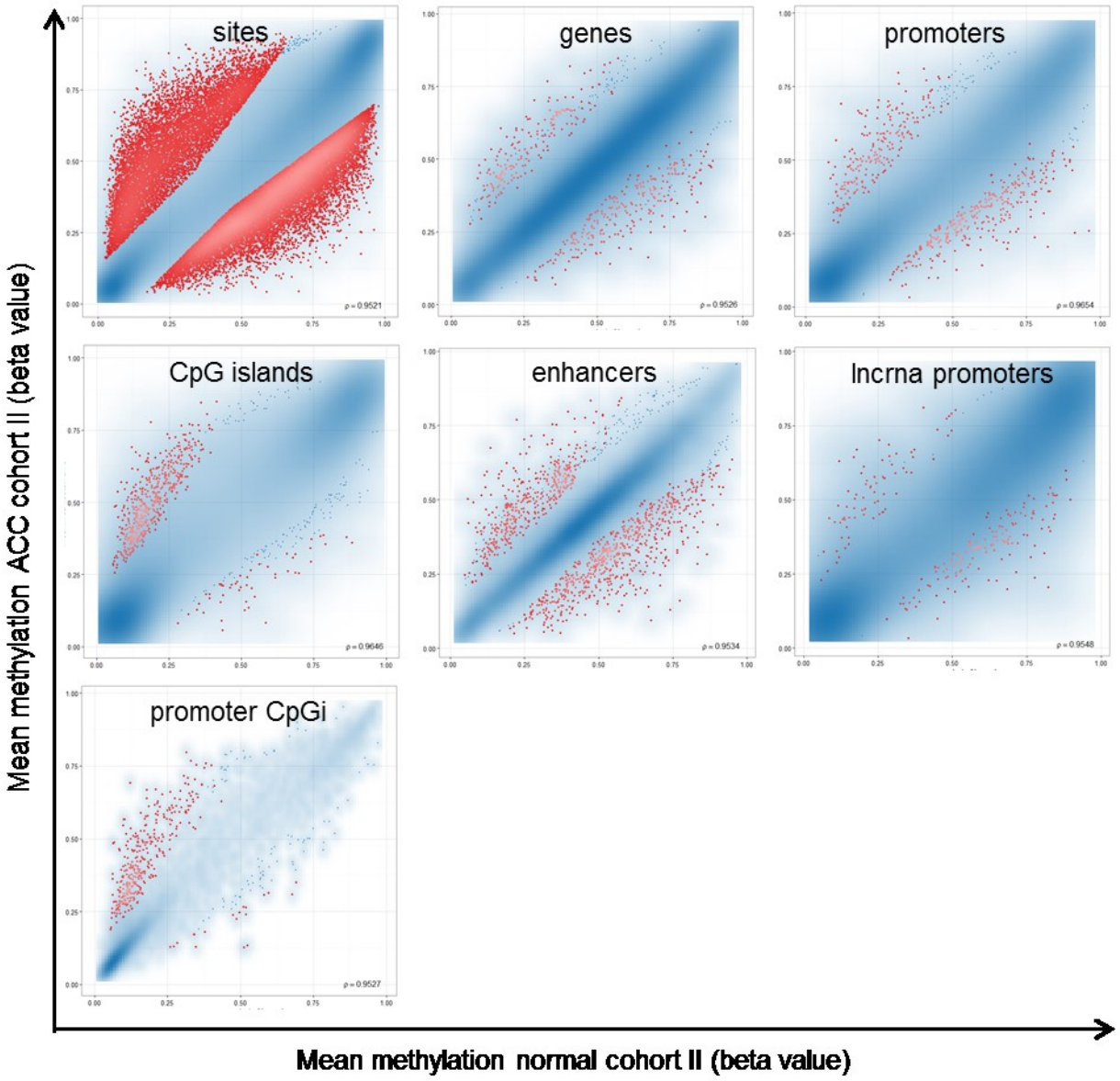
Supplementary Figure 3 Distributions of CpG sites in region annotation used for methylation analysis. For definition of region annotations refer to 3.2.2.5.



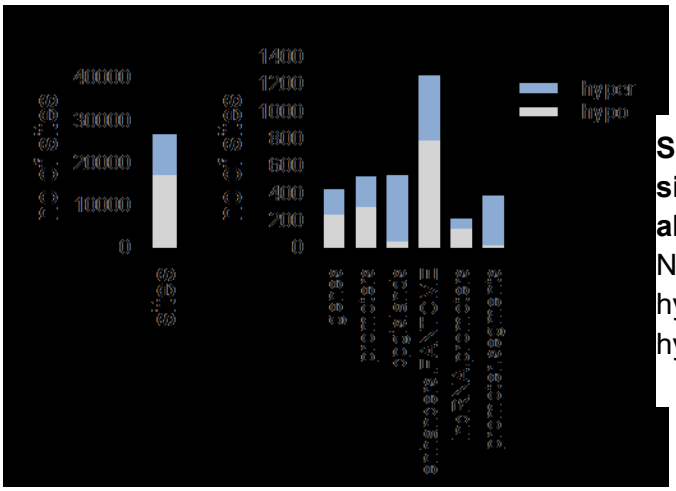
Supplementary Figure 4 Adjacent normal pancreatic tissues show few aberrations in DNA methylation with the exception of promoter hypomethylation and lncRNA promoters. Scatterplots of the paired analysis of the mean methylation of all adjacent normal tissues (x axis) versus all healthy normal tissues (y axis) at the site level and all region annotations used. Red dots are sites or regions above the selected rank cutoff.



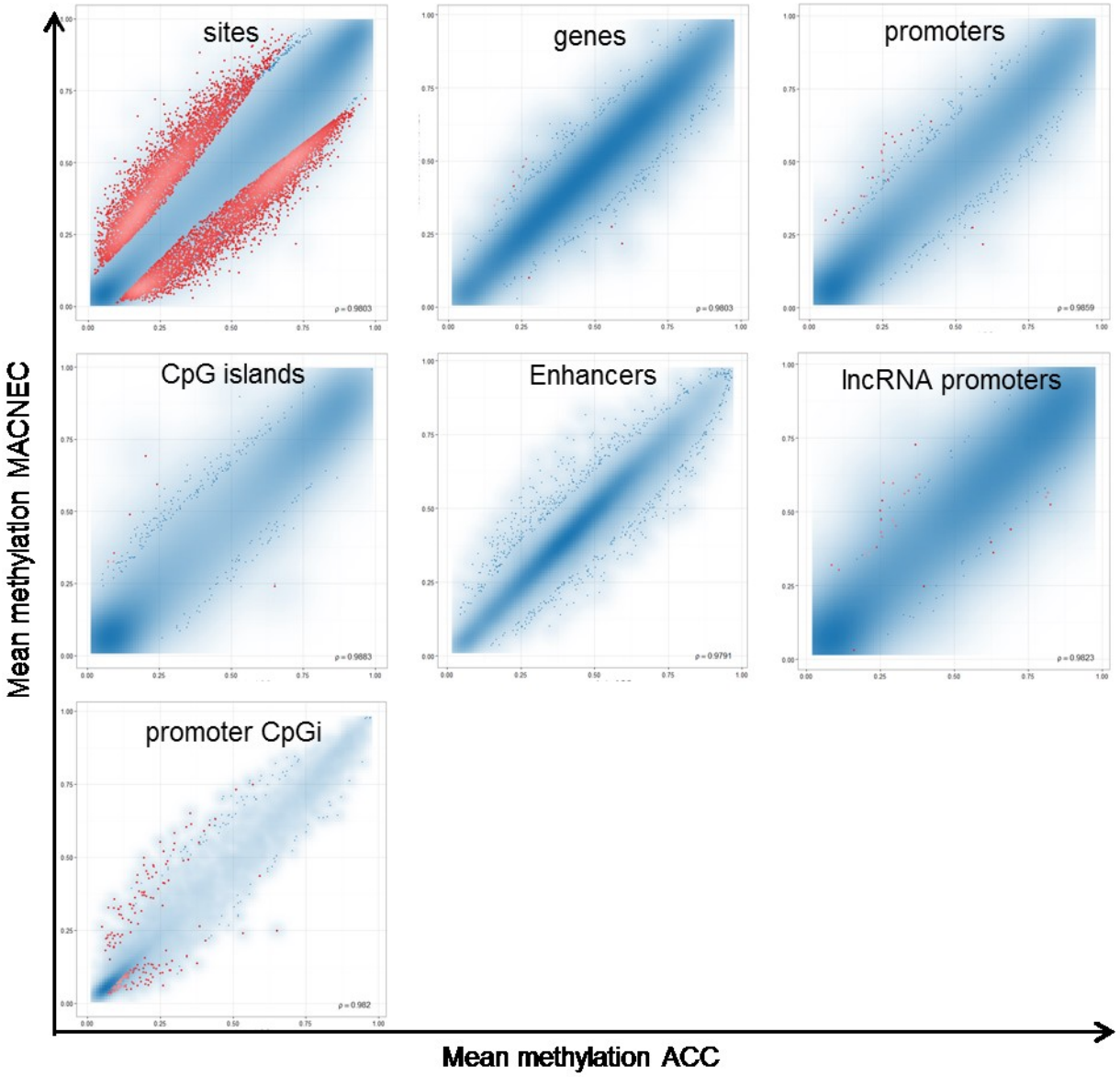
Supplementary Figure 5 Adjacent normal pancreatic tissues show few aberrations in DNA methylation with the exception of promoter hypomethylation and lncRNA promoters. a. Number of sites and **b.** number of regions that are hypermethylated (blue) or hypomethylated (grey) in the adjacent normal tissue compared to normal tissues.



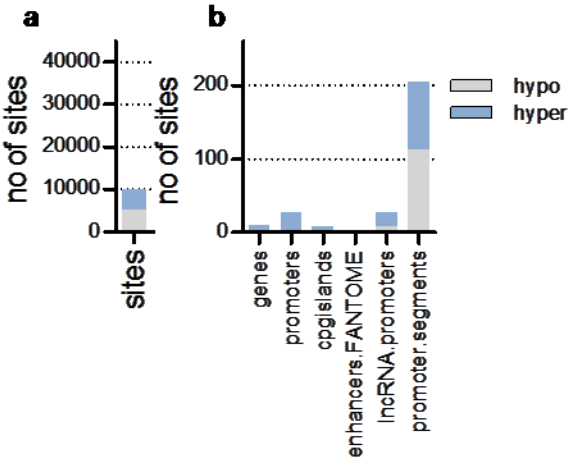
Supplementary Figure 6 Scatterplots of cohort II showing massive aberrations in DNA methylation. Scatterplots of the mean methylation of all normal tissues (x axis) versus all tumors (y axis) in cohort II at the site level and all region annotations used. Red dots are sites or regions above the selected rank cutoff.



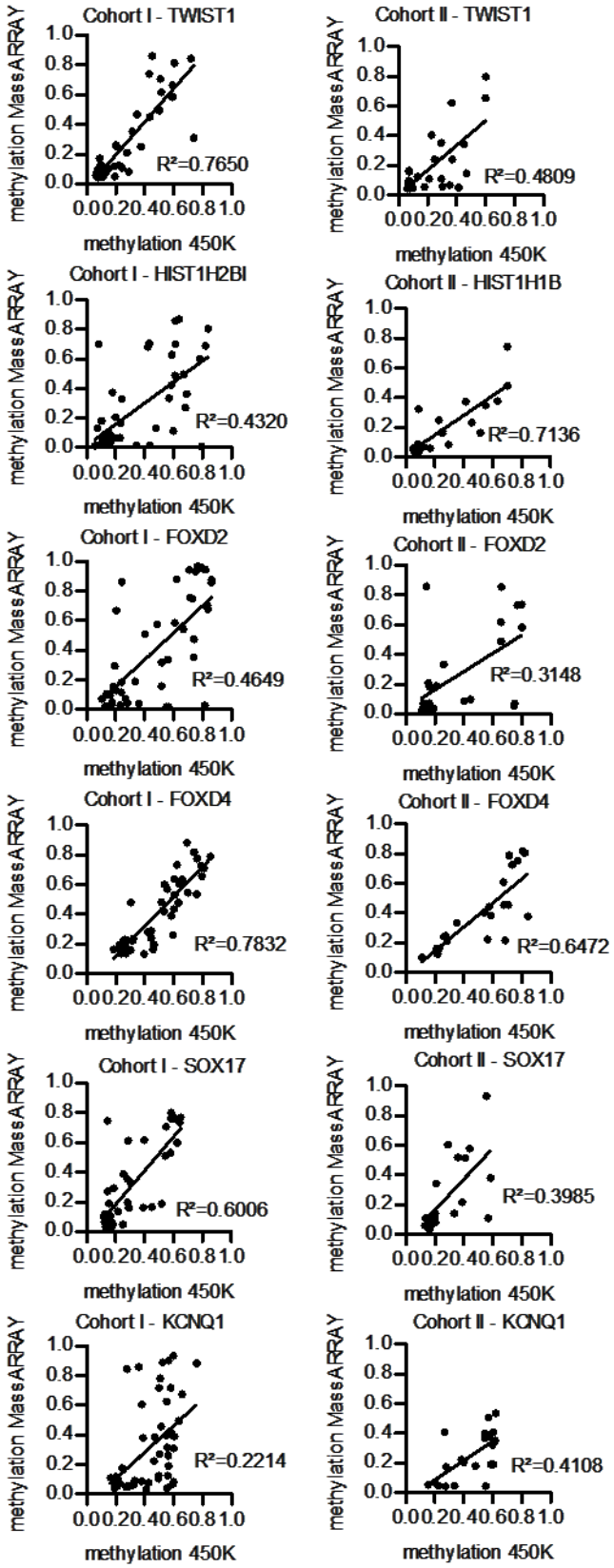
Supplementary Figure 7 Number of sites in cohort II showing massive aberrations in DNA methylation. Number of sites per region that are hypermethylated (blue) or hypomethylated (grey) in the tumors.



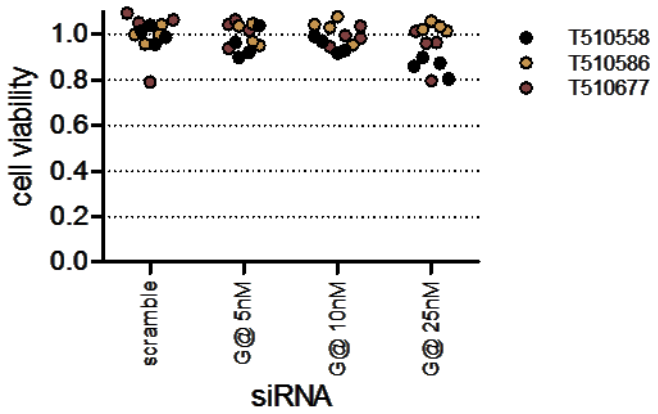
Supplementary Figure 8 MACNEC show few aberrations in DNA methylation compared to pure ACC. Scatterplots of the mean methylation of all pure ACC (x axis) versus all MACNEC (y axis) at the site level and all region annotations used. Red dots are sites or regions above the selected rank cutoff.



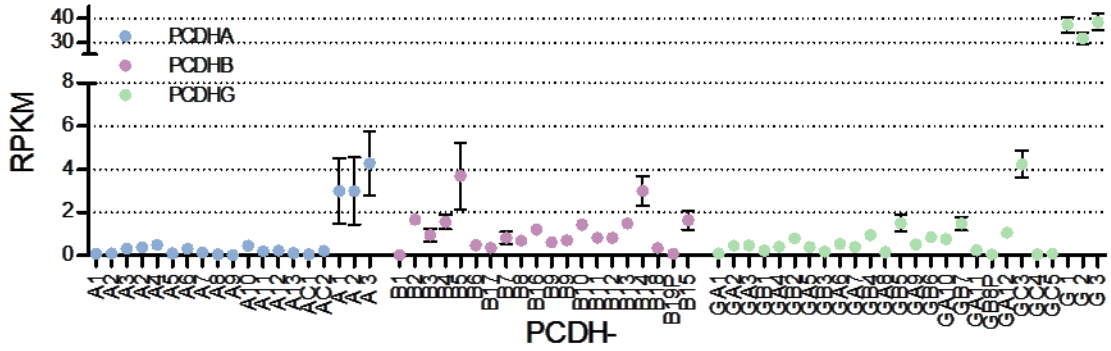
Supplementary Figure 9 MACNEC show few aberrations in DNA methylation. a. Number of sites and **b.** number of regions that are hypermethylated (blue) or hypomethylated (grey) in the MACNEC compared to pure ACC.



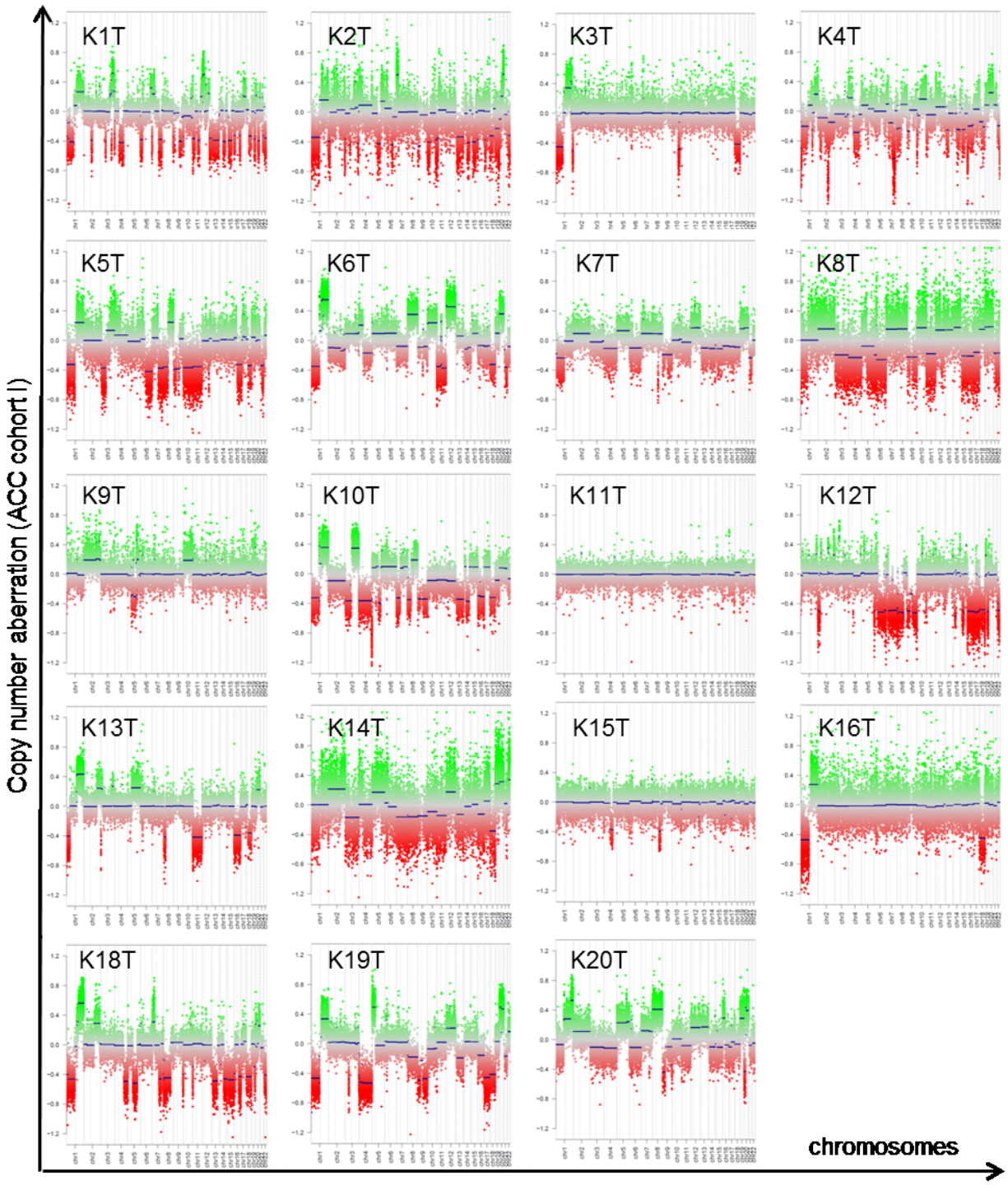
Supplementary Figure 10 Correlation plots of MassARRAY and 450K results in cohort I (left panel) and cohort II (right panel). Daniel van der Duin performed MassARRAY under my supervision.



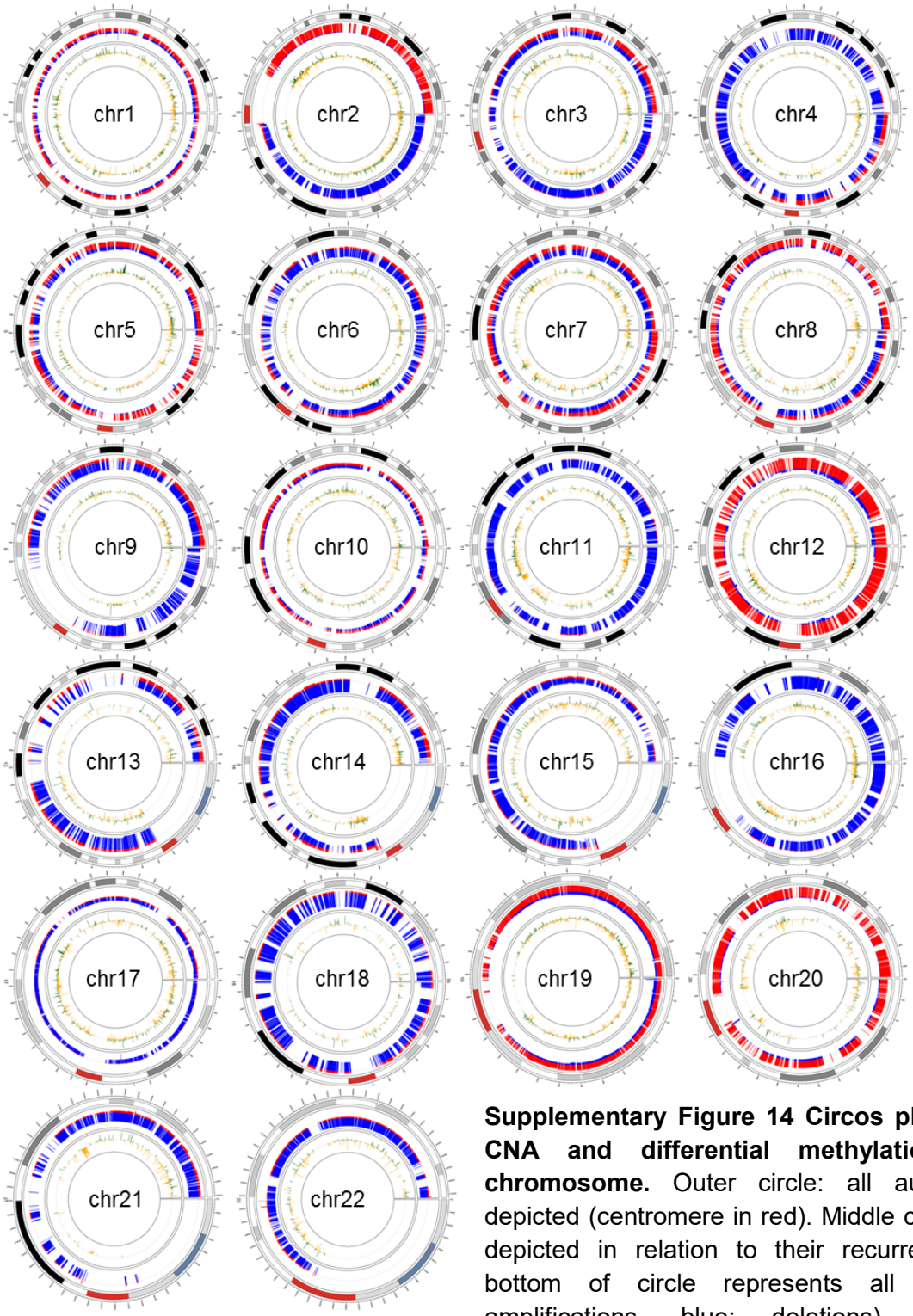
Supplementary Figure 11 Cell viability of T510 cell lines after siRNA treatment



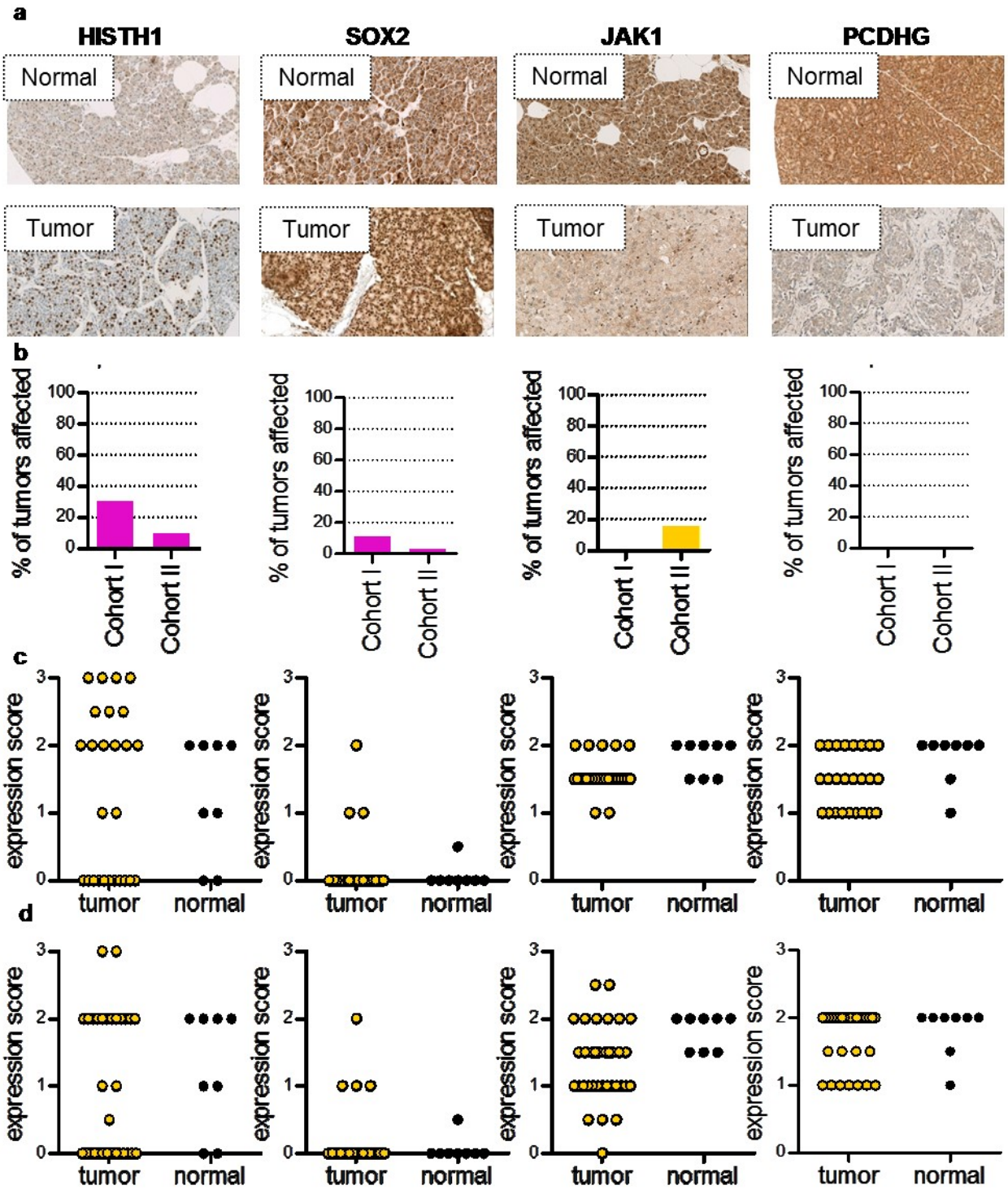
Supplementary Figure 12 Expression of *PCDH* members in normal pancreas. Mean methylation +/- SEM of three normal pancreatic tissues from TCGA was depicted. RPKM: reads per kilobase per million mapped reads.



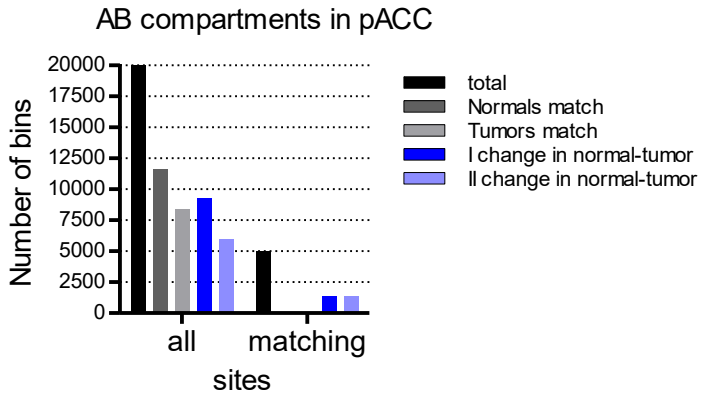
Supplementary Figure 13 ACC from cohort II had highly instable genomes. Copy number profiles were calculated for each tumor and CNA are depicted for each tumor (y axis) and each chromosome (x axis). Amplifications are depicted in green while deletions are depicted in red.



Supplementary Figure 14 Circos plot displaying CNA and differential methylation in each chromosome. Outer circle: all autosomes are depicted (centromere in red). Middle circle: CNA are depicted in relation to their recurrences (top or bottom of circle represents all tumors; red: amplifications, blue: deletions) Inner circle: Differential methylation was depicted in relation to recurrences (top or bottom of circle represents all tumors) (green: hypermethylation, yellow: hypomethylation).



Supplementary Figure 15 HISTH1 and SOX2 are up-regulated, and JAK1 and PCDHG are down-regulated in some ACC cases. a. representative IHC figures of normal and tumor tissue. PD Dr. Frank Bergmann performed immunohistochemical stainings **b.** percentages of tumors affected. Yellow: downregulated, pink: upregulated. **c.** dot plot of pathology score of protein expression of cohort I. **d.** dot plot of pathology score of protein expression of cohort II. yellow: tumor, black: normal.



Supplementary Figure 16 AB compartments of the two cohorts (all) and their overlap (matching)

8.2 Supplementary Tables

Supplementary Table 1 Somatic signatures calculated for each ACC.

		Signature																	
		1	2	3	4	5	6	7	8	10	11	15	16	18	20	21	24	25	29
Tumors	FF-11	0.12	0	0.17	0	0	0.32	0	0	0.16	0	0.07	0	0	0	0.07	0	0	0
	FF-17	0	0	0.12	0.13	0.21	0.28	0	0	0	0	0.1	0	0	0	0	0	0	0
	FF-18	0	0	0	0.33	0	0	0.1	0	0.19	0	0	0	0	0.37	0	0	0	0
	FF-22	0.23	0	0.3	0.16	0	0.14	0	0	0.07	0	0	0	0	0	0.06	0	0	0
	FF-5	0.3	0	0	0	0	0	0	0	0	0	0	0.15	0.53	0	0	0	0	0
	FFPE-1	0.31	0	0.07	0.15	0	0	0	0.11	0	0	0	0	0	0	0	0.11	0	0
	FFPE-10	0.29	0.07	0.25	0.22	0	0	0.06	0	0	0	0	0	0	0	0	0	0	0
	FFPE-12	0.46	0.08	0	0	0	0	0	0	0	0.07	0	0.17	0	0	0	0	0	0.08
	FFPE-13	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	FFPE-14	0.41	0	0.1	0	0	0.41	0	0	0	0	0	0	0	0	0	0	0	0
	FFPE-15	0.4	0	0	0	0	0.25	0.07	0	0.07	0	0	0	0	0	0	0	0	0.09
	FFPE-16	0.41	0	0.16	0	0	0.15	0	0	0	0	0.09	0	0	0	0	0	0	0
	FFPE-19	0.23	0	0	0	0	0.54	0.13	0	0	0	0	0	0	0	0	0	0	0
	FFPE-20	0.51	0	0	0.36	0	0	0	0	0	0	0	0	0	0.07	0	0	0	0
	FFPE-21	0.46	0	0	0.11	0	0	0	0.23	0	0	0	0	0	0	0	0.09	0	0
	FFPE-23	0.77	0.11	0	0.11	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	FFPE-4	0.27	0	0.28	0.29	0	0	0	0	0	0	0.06	0	0	0	0	0	0	0
	FFPE-6	0.43	0	0.2	0.13	0	0	0	0	0	0	0	0	0	0	0	0	0.1	0
	FFPE-7	0.5	0.11	0	0	0	0	0	0	0	0	0.23	0	0	0	0	0	0	0
	FFPE-8	0.29	0.19	0	0.09	0	0	0.16	0	0	0	0	0.06	0	0	0	0	0	0.1
	FFPE-9	0.56	0.14	0	0	0	0	0	0	0	0	0	0	0.15	0	0	0	0	0
	FFPE-K1	0.57	0	0.09	0.27	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Supplementary Table 2 Group memberships for differential methylation calling.

tumor	cohort I	cohort II	metastasis	MACNEC	adjacent normal
10T	primary tumor			ACC	
11T	primary tumor			ACC	
12T	primary tumor		primary tumor	ACC	
13T	primary tumor			ACC	
14T	primary tumor			ACC	
15T	primary tumor			ACC	
16T	primary tumor			ACC	
17T	primary tumor			ACC	
18T	primary tumor			ACC	
19T	primary tumor		primary tumor	ACC	
1T	primary tumor		primary tumor	ACC	
3T	primary tumor			ACC	
4T	primary tumor		primary tumor	ACC	
5T	primary tumor			ACC	
6T	primary tumor		primary tumor	ACC	
7T	primary tumor			ACC	
8T	primary tumor		primary tumor	ACC	
9T	primary tumor		primary tumor	ACC	
10N	normal				adjacent normal
11N	normal				adjacent normal
12N	normal				adjacent normal
1N	normal				adjacent normal
21N	normal				adjacent normal
22N	normal				adjacent normal
23N	normal				adjacent normal
3N	normal				adjacent normal
4N	normal				adjacent normal
5N	normal				adjacent normal
8N	normal				adjacent normal
9N	normal				adjacent normal
19M			metastasis		
20LK			metastasis		
41M					
4M			metastasis		
8M			metastasis		
12LK			metastasis		
19LK					
1LK			metastasis		
6LK			metastasis		
8LK					
9LK			metastasis		
20T	primary tumor		primary tumor	MACNEC	
21T	primary tumor			MACNEC	
22T	primary tumor			MACNEC	
23T	primary tumor			MACNEC	
6N	normal				adjacent normal
7N	normal				adjacent normal
N1N	normal				healthy normal
N2N	normal				healthy normal
N3N	normal				healthy normal
N4N	normal				healthy normal
N5N	normal				healthy normal
N6N	normal				healthy normal

6M					
ACC K17 T		primary tumor		ACC	
ACC K18 T		primary tumor		ACC	
ACC K19 T		primary tumor		ACC	
ACC K20 T		primary tumor		ACC	
ACC K1 T		primary tumor		ACC	
ACC K5 T		primary tumor		ACC	
ACC K6 T		primary tumor		ACC	
ACC K8 T		primary tumor		ACC	
ACC K9 T		primary tumor		ACC	
ACC K10 T		primary tumor		ACC	
ACC K11 T		primary tumor		ACC	
ACC K12 T		primary tumor		ACC	
ACC K13 T		primary tumor		ACC	
ACC K14 T		primary tumor		ACC	
ACC K15 T		primary tumor		ACC	
ACC K2 N		normal			adjacent normal
ACC K1 N		normal			adjacent normal
ACC K2 T		primary tumor		MACNEC	
ACC K3 T		primary tumor		MACNEC	
ACC K4 T		primary tumor		MACNEC	
ACC K7 T		primary tumor		MACNEC	
ACC K22 N		normal			healthy normal
ACC K23 N		normal			healthy normal
ACC K24 N		normal			healthy normal
ACC K26 N		normal			healthy normal
ACC K27 N		normal			healthy normal
ACC K28 N		normal			healthy normal

Supplementary Table 3 364 genes differentially methylated either at their promoter site or gene body, identified in cohort I and validated in cohort II. Numbers refers to Ensemble annotation 'ENSG'.

258952	189134	146666	230124	232230	257126	254245
234854	136099	258346	253302	154478	106004	184302
229964	125816	226965	235529	186766	265075	081853
198327	261934	229627	188694	234531	181201	174963
270182	235608	260647	264577	183734	081842	253537
207716	234952	154263	221116	109851	239389	204956
238230	141748	158488	254129	187811	271071	216753
234437	187714	261275	259362	149201	131016	230756
267570	249505	228022	233532	183644	268723	253731
262096	132840	202031	207405	236437	204963	251493
207715	232742	180483	215458	258343	130226	253953
179909	176435	255148	268194	139343	272311	178919
256316	129596	209480	226745	089116	187626	253767
228925	253305	254818	237424	089225	233502	196963
170178	152977	185686	186564	189238	269067	186047
257056	269543	253288	230798	151952	169594	229195
159182	176165	143878	201142	139800	255408	256018
248802	156150	204941	234283	224243	113140	164616
196570	253508	169908	224540	129467	179008	168779
260738	164438	225058	223519	198807	134802	181195
251621	207816	256137	224128	248550	232973	243981
253187	272163	255087	230090	260053	164458	255966
105996	253846	147434	163064	159556	141485	196109
202167	152192	231394	136732	255571	267882	149346
078399	162722	205030	231147	248441	224092	260426
232709	272191	259143	224076	103351	185504	256546
226063	128713	168995	144290	134398	185177	253616
008196	250103	225285	217236	166501	072201	264630
259711	198914	251205	216193	259952	110514	196578
182968	257935	232551	128710	261614	263776	226935
228478	164093	143546	175892	259725	273179	272068
175879	237152	259905	128652	177294	259372	201754
229637	253405	263370	119042	262786	187474	172464
226740	254369	264424	135903	186075	261603	181935
165588	187140	199004	163508	263602	099622	230301
248449	176046	231188	241472	182446	158481	201689
128709	255562	261633	170893	267336	234977	094661
264304	120328	232788	121853	166342	211804	199673
231951	254122	122133	157005	263958	238297	263499
106540	271956	205971	250238	256463	252367	234715
122592	138083	167644	163132	105131	164500	174697
142700	102924	234771	169676	160321	140931	180053
225280	184492	237243	250467	196350	224579	253471
198558	263146	266824	109182	215998	197364	104332
187559	184344	249737	179059	249233	269575	254254
148826	197576	269066	142319	121413	215621	231191
235988	189332	246526	259757	125878	240470	250850
270475	253293	172487	186493	132671	225528	225337
151615	197459	255550	170549	101017	221468	077327
170122	107807	184247	251573	124227	242282	165606
237380	253873	224037	248918	198768	224093	155254
257184	069011	263986	249149	231290	173991	257582
113248	148704	225905	247993	224269	271798	203809
092607	168875	251909	113319	238099	180210	271090
235051	232638	221957	204962	271100	229953	244301
261168	187372	143954	253159	154025	252531	106038
263426	188779	269758	242419	261267	100023	263089
182853	228400	248206	255373	117586	263644	120156
253910	170128	131068	124827	078043	249417	250133
272382	171540	232263	218690	270935	108679	271776
253485	213435	236635	168242	236989	238649	232869
262576	259439	200418	182572	227817	212309	248719
215474	255399	168899	235863	204262	226416	229150
146001	165862	267895	112837	117477	232987	118526
238284	267142	181698	112238	231453	162458	250120
240990	204779	238917	228010	183729	142327	178199
162618	143869	225750	122691	272801	163818	242385
254221	203818	254598	146618	234595	238578	256466
164736	006377	255351	207584	253308	253128	128298
163406	227700	230563	005073	255886	261845	231172
262209	120322	270174	164532	257657	205537	226756
163081	121075	209582	185811	105997	224104	196878
231764	120324	261741	106236	273433	226087	256374
161381	227006	227748	243478	168421		

Supplementary Table 4 Cancers from TCGA that are hypermethylated at the PCDH cluster.

cancer	p-value	significance
BRCA	< 0.0001	****
COAD	0.0033	**
BLCA	< 0.0001	****
GBM	< 0.0001	****
HNSC	< 0.0001	****
KIRC	< 0.0001	****
LUAD	< 0.0001	****
LUSC	< 0.0001	****
PRAD	< 0.0001	****
READ	< 0.0001	****
STAD	0.2705	ns
SKCM	0.0164	*
THCA	0.0616	ns
UCEC	< 0.0001	****

Supplementary Table 5 Breast, colon, and ovarian cell lines which are significantly hypomethylated at the *PCDH* cluster after DAC treatment.

Cell line	p-value	signifi.	tissue	Cell line	p-value	signifi.	tissue
MDA 231	0.0781	ns	breast	Colo201	< 0.0001	****	colon
MDA 468	0.842	ns	breast	Colo320	< 0.0001	****	colon
MDA 453	0.0002	***	breast	DLD1	< 0.0001	****	colon
HCC1569	0.0034	**	breast	HCT116	0.0005	***	colon
T47D	0.0035	**	breast	HT29	< 0.0001	****	colon
EFM19	0.0189	*	breast	Lovo	0.002	**	colon
MCF7	0.0105	*	breast	RKO	0.197	ns	colon
BT20	< 0.0001	****	breast	SNUC1	0.0409	*	colon
HCC38	0.1826	ns	breast	SW48	< 0.0001	****	colon
SUM149	< 0.0001	****	breast	SW480	0.0247	*	colon
SUM159	< 0.0001	****	breast	Caco2	0.7114	ns	colon
ZRF530	< 0.0001	****	breast	SW620	< 0.0001	****	colon
HCC1150	0.4447	ns	breast	SKCO1	0.0012	**	colon
MDA415	0.9569	ns	breast	Colo205	0.0818	ns	colon
MDA175	0.6776	ns	breast	A2780	< 0.0001	****	ovarian
MDA436	< 0.0001	****	breast	CAOV3	< 0.0001	****	ovarian
EFM192	< 0.0001	****	breast	EF027	0.0219	*	ovarian
HCC1419	0.875	ns	breast	ES2	0.0634	ns	ovarian
MDA175	0.0006	***	breast	Hey	< 0.0001	****	ovarian
HCC1954	< 0.0001	****	breast	Kuramochi	< 0.0001	****	ovarian
HCC1187	0.9782	ns	breast	OAW28	0.8344	ns	ovarian
SKBR3	0.6978	ns	breast	OV2008	< 0.0001	****	ovarian
CAMA1	< 0.0001	****	breast	OVCAR3	< 0.0001	****	ovarian
BT474	< 0.0001	****	breast	OVCAR5	< 0.0001	****	ovarian
MDA361	0.0673	ns	breast	Ovkate	0.0027	**	ovarian
ZR751	0.0005	***	breast	SKOV3	0.341	ns	ovarian

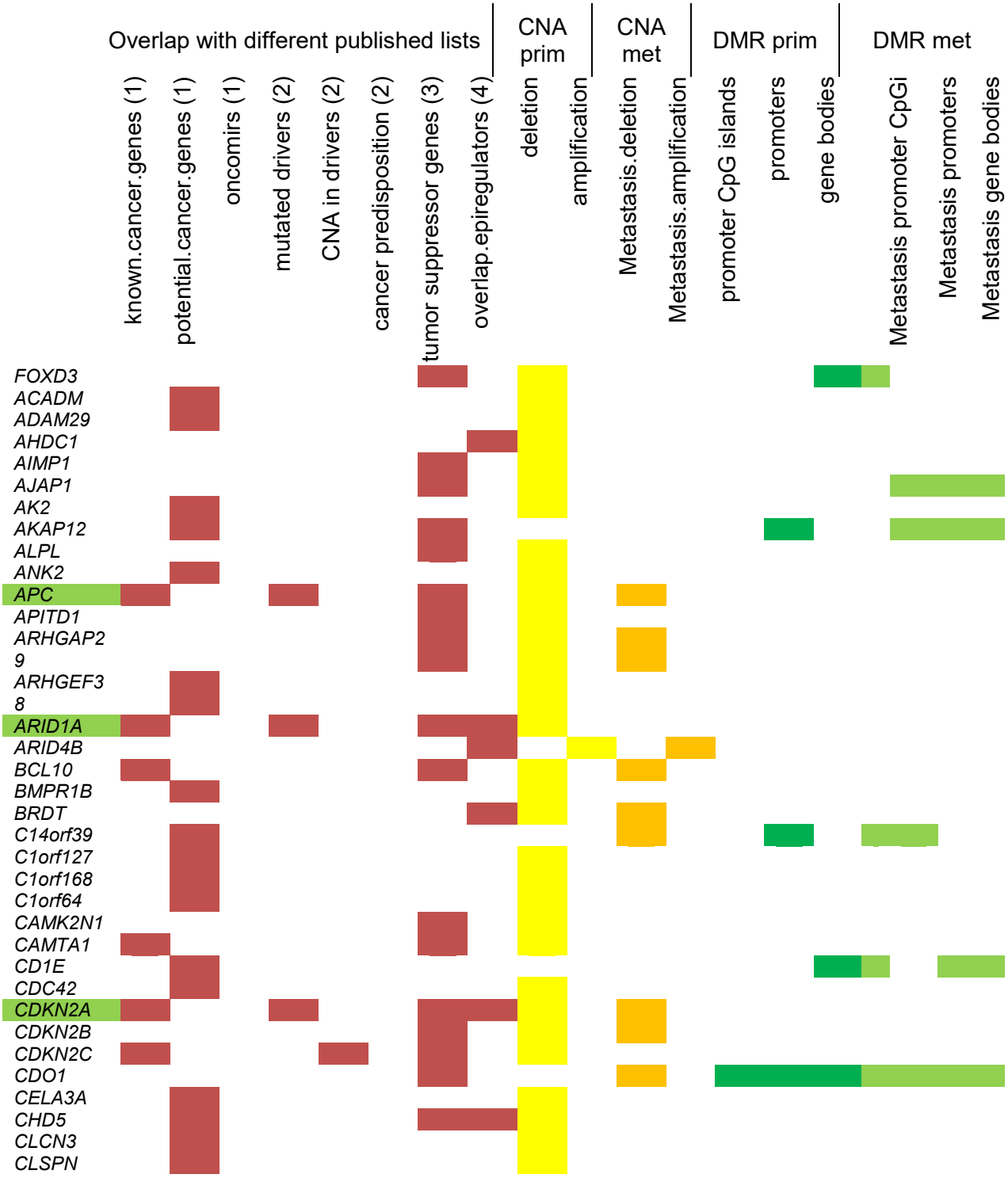
Supplementary Table 6 Significances of DAC treatment in T510 cell lines on mpcdh methylation. signific. : significance.

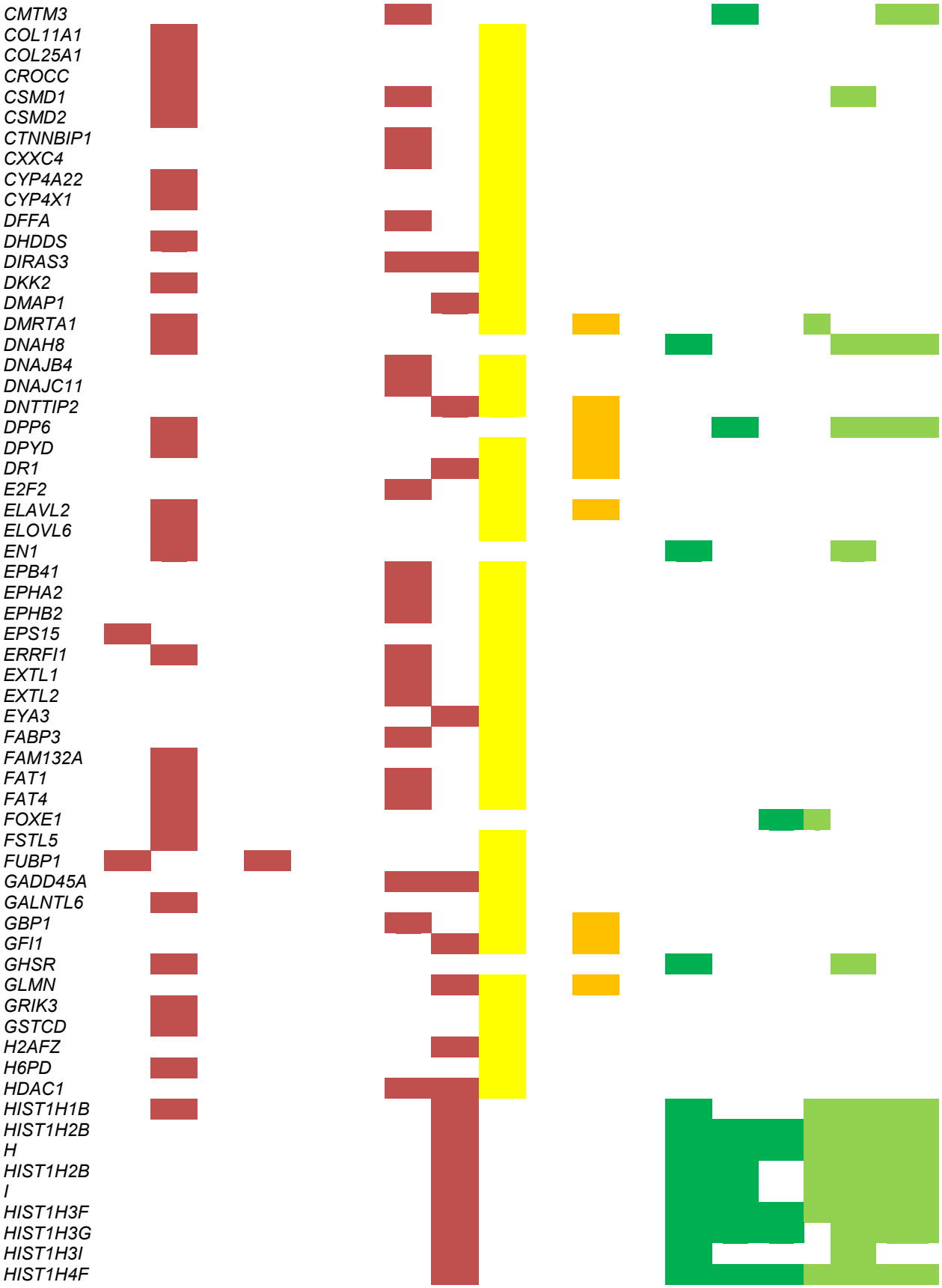
	DAC 0.5 µmol/l		DAC 1.0 µmol/l		DAC 2.0 µmol/l		DAC 0.5-2.0 µmol/l	
	p-value	signific.	p-value	signific.	p-value	signific.	p-value	signific.
b17	0.187	ns	0.179	ns	0.177	ns	0.034	*
b18	0.182	ns	0.168	ns	0.187	ns	0.036	*
b20	0.153	ns	0.149	ns	0.155	ns	0.033	*
b21	0.140	ns	0.160	ns	0.189	ns	0.021	*
ga12	0.135	ns	0.149	ns	0.136	ns	0.001	***
gb4	0.478	ns	0.489	ns	0.447	ns	0.110	ns
gb7	0.225	ns	0.295	ns	0.200	ns	0.012	*
gc3	0.137	ns	0.151	ns	0.103	ns	0.030	*
line1	0.155	ns	0.169	ns	0.165	ns	0.010	**

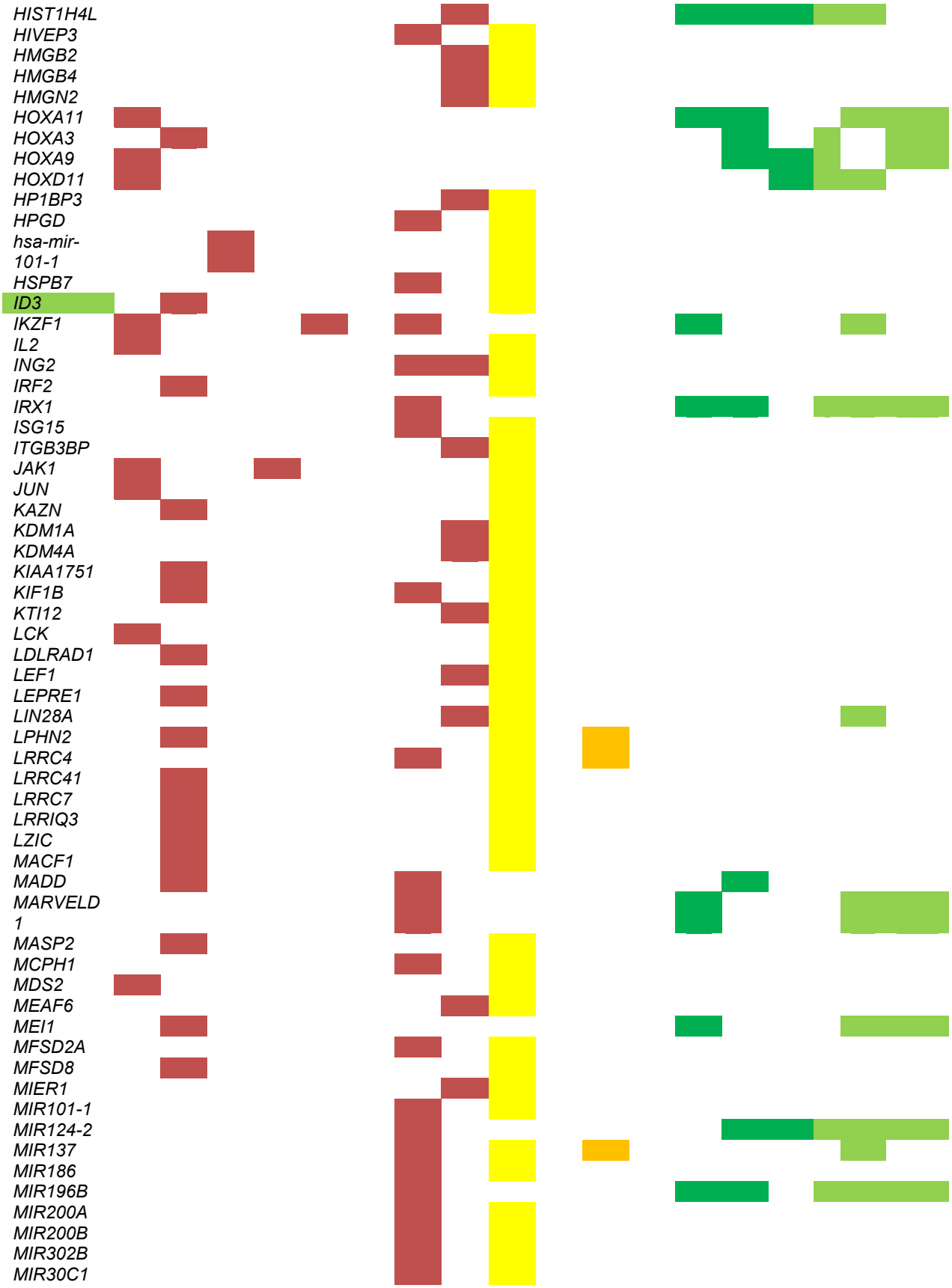
Supplementary Table 7 Significances of DAC treatment in T510 cell lines on mpcdh RNA expression. signific. : significance.

	DAC 0.5 µmol/l		DAC 1.0 µmol/l		DAC 2.0 µmol/l		DAC 0.5-2.0 µmol/l	
	p-value	signific.	p-value	signific.	p-value	signific.	p-value	signific.
b17	0.436	ns	0.415	ns	0.697	ns	0.083	ns
b18	0.387	ns	0.338	ns	0.302	ns	0.025	*
b20	0.325	ns	0.310	ns	0.281	ns	0.016	*
b22	0.222	ns	0.234	ns	0.234	ns	0.183	ns
gb4	0.590	ns	0.034	*	0.194	ns	0.633	ns
gb7	0.261	ns	0.176	ns	0.196	ns	0.209	ns
gc3	0.207	ns	0.572	ns	0.775	ns	0.997	ns
g@1,2	0.518	ns	0.023	*	0.118	ns	0.407	ns
g@2,3	0.646	ns	0.184	ns	0.192	ns	0.621	ns

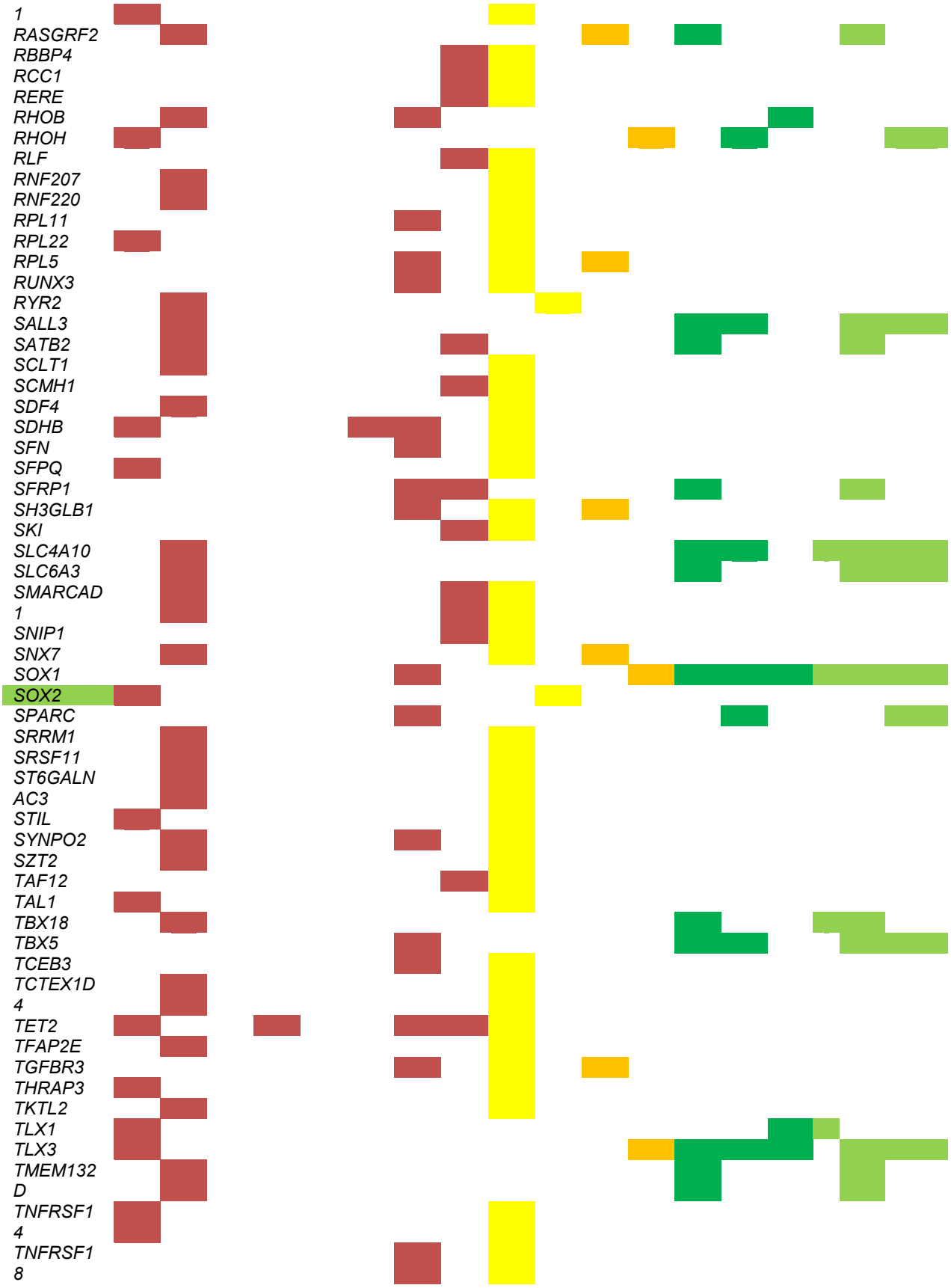
Supplementary Table 8 Top altered genes in ACC overlapped with genes known to play a role in tumorigenesis. (1) Top hits from the present study were overlapped with known and candidate cancer genes from (1) King's college[261], driver genes (mutated, CNA, and predisposition driver genes) from (2) Vogelstein, et al. [20], tumor suppressor genes from (3) TSGene[262], and epigenetic regulators from (4) Plass, et al. [67]. prim = primary tumors, met = metastases

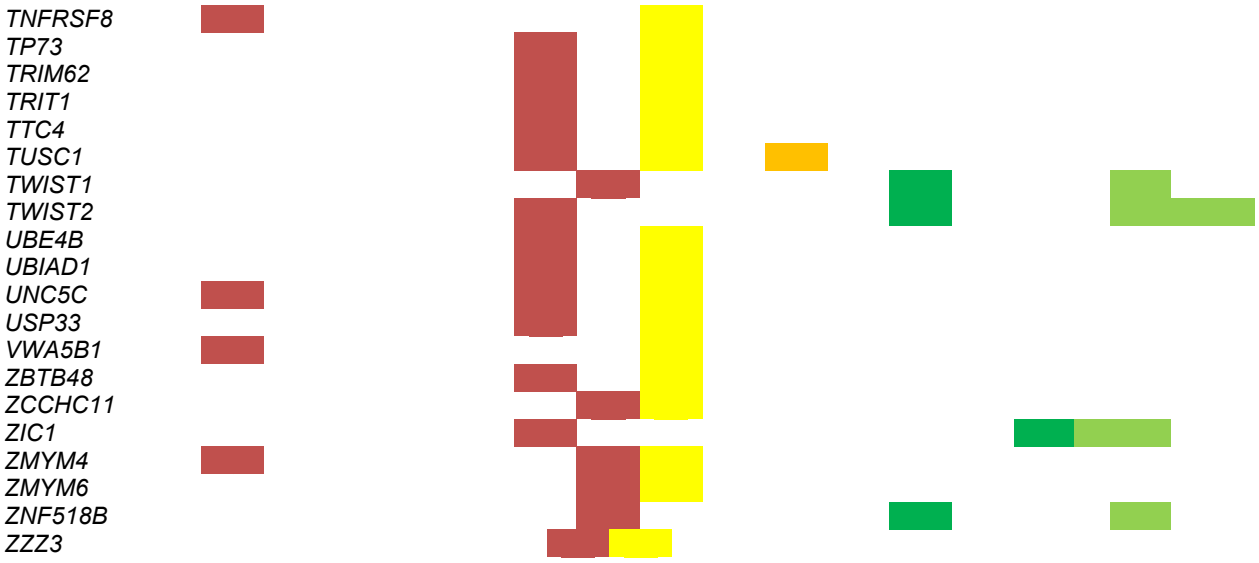



CMTM3
COL11A1
COL25A1
CROCC
CSMD1
CSMD2
CTNNBIP1
CXXC4
CYP4A22
CYP4X1
DFFA
DHDDS
DIRAS3
DKK2
DMAP1
DMRTA1
DNAH8
DNAJB4
DNAJC11
DNTTIP2
DPP6
DPYD
DR1
E2F2
ELAVL2
ELOVL6
EN1
EPB41
EPHA2
EPHB2
EPS15
ERRFI1
EXTL1
EXTL2
EYA3
FABP3
FAM132A
FAT1
FAT4
FOXE1
FSTL5
FUBP1
GADD45A
GALNTL6
GBP1
GFI1
GHSR
GLMN
GRIK3
GSTCD
H2AFZ
H6PD
HDAC1
HIST1H1B
HIST1H2BH
HIST1H2BI
HIST1H3F
HIST1H3G
HIST1H3I
HIST1H4F


HIST1H4L
HIVEP3
HMGB2
HMGB4
HMGN2
HOXA11
HOXA3
HOXA9
HOXD11
HP1BP3
HPGD
hsa-mir-101-1
HSPB7
ID3
IKZF1
IL2
ING2
IRF2
IRX1
ISG15
ITGB3BP
JAK1
JUN
KAZN
KDM1A
KDM4A
KIAA1751
KIF1B
KTI12
LCK
LDLRAD1
LEF1
LEPRE1
LIN28A
LPHN2
LRRC4
LRRC41
LRRC7
LRRIQ3
LZIC
MACF1
MADD
MARVELD 1
MASP2
MCPH1
MDS2
MEAF6
MEI1
MFSD2A
MFSD8
MIER1
MIR101-1
MIR124-2
MIR137
MIR186
MIR196B
MIR200A
MIR200B
MIR302B
MIR30C1


MIR31
MIR34A
MIR367
MIR551A
MOS
MPL
MTAP
MUTYH
MYCL1
MYH4
MYOM3
MYSM1
NASP
NDST4
NEUROG2
NFKB1
NOC2L
NPHP4
NPY5R
NR0B2
NTNG1
NUPR1
OR5L2
ORC1
OSCP1
PADI4
PAEP
PAQR7
PARK7
PAX3
PAX7
PAX9
PCDH10
PCDH18
PCDH8
PCDHA13
PCDHA6
PCDHB15
PCDHB8
PDE4B
PDGFC
PGRMC2
PHACTR4
PHC2
PKDREJ
PLA2G2A
PLCH2
PLEKHG5
PNLIPRP2
PPP3CA
PRDM16
PRDM2
PRDM5
PRKAA2
PRKCB
PRMT6
PTCH2
PTGFR
PTPRU
RAD54L
RAP1GAP
RAP1GDS


1
RASGRF2
RBBP4
RCC1
RERE
RHOB
RHOH
RLF
RNF207
RNF220
RPL11
RPL22
RPL5
RUNX3
RYR2
SALL3
SATB2
SCLT1
SCMH1
SDF4
SDHB
SFN
SFPQ
SFRP1
SH3GLB1
SKI
SLC4A10
SLC6A3
SMARCAD 1
SNIP1
SNX7
SOX1
SOX2
SPARC
SRRM1
SRSF11
ST6GALN AC3
STIL
SYNPO2
SZT2
TAF12
TAL1
TBX18
TBX5
TCEB3
TCTEX1D 4
TET2
TFAP2E
TGFBR3
THRAP3
TKTL2
TLX1
TLX3
TMEM132 D
TNFRSF1 4
TNFRSF1 8



Supplementary Table 9 Statistical analysis of protein expressions *via* Mann Whitney U.

	cohort I		cohort II	
	p-value	significance	p-value	significance
ID3	< 0.0001	****	< 0.0001	****
ARID1A	0.0021	**	0.0009	***
APC	< 0.0001	****	< 0.0001	****
CDKN2A	0.1077	ns	0.0129	*
HISTH1	0.8486	ns	0.6004	ns
SOX2	0.9716	ns	0.9773	ns
JAK1	0.0215	*	0.0145	*
PCDHG	0.0537	ns	0.6303	ns

Supplementary Table 10 Regions from iCluster analysis.

	CNA regions		Methylated promoters
1	chr1.1606361-1684095	1	ENSG00000116251
2	chr1.1684095-2990490	2	ENSG00000009709
3	chr1.2990490-3331154	3	ENSG00000162367
4	chr1.3331154-4000600	4	ENSG00000162374
5	chr1.9970659-10221203	5	ENSG00000142700
6	chr1.33938904-34098842	6	ENSG00000118473
7	chr1.46859774-49748480	7	ENSG00000116745
8	chr1.49748480-50513645	8	ENSG00000162654
9	chr1.50513645-50892674	9	ENSG00000238081
10	chr1.50892674-54665944	10	ENSG00000143032
11	chr1.72748601-72884990	11	ENSG00000224445
12	chr1.72884990-76077832	12	ENSG00000162639
13	chr1.76077832-76138023	13	ENSG00000182898
14	chr1.76138023-77746028	14	ENSG00000143631
15	chr1.77746028-79801791	15	ENSG00000185966
16	chr1.79801791-81968855	16	ENSG00000198854
17	chr1.81968855-86913263	17	ENSG00000203786
18	chr1.86913263-96500991	18	ENSG00000172155
19	chr1.96500991-98387605	19	ENSG00000196734

20	chr1.98387605-99190917	20	ENSG00000186844
21	chr1.99190917-103573989	21	ENSG00000163518
22	chr1.143907550-145001340	22	ENSG00000160856
23	chr1.145001340-145205174	23	ENSG00000158477
24	chr1.145205174-145250914	24	ENSG00000158485
25	chr1.145250914-145382340	25	ENSG00000186306
26	chr1.145382340-145705805	26	ENSG00000180708
27	chr1.145705805-146503900	27	ENSG00000173285
28	chr1.146503900-147374589	28	ENSG00000186400
29	chr1.147374589-147800744	29	ENSG00000198967
30	chr1.147800744-147851370	30	ENSG00000196171
31	chr1.147851370-147994007	31	ENSG00000203757
32	chr1.147994007-149132021	32	ENSG00000180433
33	chr1.149132021-149144161	33	ENSG00000197403
34	chr1.149144161-149212246	34	ENSG00000188340
35	chr1.149212246-149333124	35	ENSG00000213088
36	chr1.149333124-149670057	36	ENSG00000179639
37	chr1.149670057-149718671	37	ENSG00000196266
38	chr1.149718671-149815218	38	ENSG00000203740
39	chr1.149815218-152880108	39	ENSG00000224286
40	chr1.152880108-153029939	40	ENSG00000116132
41	chr1.153029939-153232060	41	ENSG00000117501
42	chr1.153232060-154155697	42	ENSG00000007933
43	chr1.154155697-158120045	43	ENSG00000221390
44	chr1.158120045-158259914	44	ENSG00000094963
45	chr1.158259914-159037843	45	ENSG00000082482
46	chr1.159037843-159174001	46	ENSG00000196482
47	chr1.159174001-161375390	47	ENSG00000234754
48	chr1.161375390-161450473	48	ENSG00000168243
49	chr1.161450473-161492997	49	ENSG00000177489
50	chr1.161492997-161654727	50	ENSG00000177476
51	chr1.161654727-163289415	51	ENSG00000169214
52	chr1.163289415-165414583	52	ENSG00000162722
53	chr1.165414583-170040953	53	ENSG00000224227
54	chr1.170040953-171306733	54	ENSG00000203663
55	chr1.171306733-172795457	55	ENSG00000162727
56	chr1.172795457-176430983	56	ENSG00000171180
57	chr1.176430983-177980492	57	ENSG00000196539
58	chr1.177980492-180165648	58	ENSG00000186487
59	chr1.180165648-181290986	59	ENSG00000237401
60	chr1.181290986-190444666	60	ENSG00000176887
61	chr2.45264-2111623	61	ENSG00000163032
62	chr2.2111623-2940276	62	ENSG00000115138
63	chr2.2940276-3012844	63	ENSG00000158089
64	chr2.112417362-119566221	64	ENSG00000205221
65	chr2.119566221-119590523	65	ENSG00000170820
66	chr2.119590523-119610605	66	ENSG00000204640
67	chr2.119610605-119613877	67	ENSG00000115602
68	chr2.119613877-119698443	68	ENSG00000115604
69	chr2.119698443-120300987	69	ENSG00000225765
70	chr2.120300987-121453143	70	ENSG00000233639
71	chr2.121453143-127315703	71	ENSG00000198914
72	chr2.127315703-129050050	72	ENSG00000135973
73	chr2.129050050-130680984	73	ENSG00000196228
74	chr2.144223682-145823182	74	ENSG00000224655
75	chr2.149215431-150444834	75	ENSG00000152253
76	chr2.150444834-150624507	76	ENSG00000231453
77	chr2.153032696-154154620	77	ENSG00000217236
78	chr2.155433424-161789811	78	ENSG00000174279
79	chr2.161789811-162930666	79	ENSG00000128714
80	chr2.162930666-163520456	80	ENSG00000170178
81	chr2.175300689-176718034	81	ENSG00000128713
82	chr2.176718034-176931864	82	ENSG00000128710
83	chr2.176931864-176937704	83	ENSG00000237380
84	chr2.176937704-176954533	84	ENSG00000128709
85	chr2.176954533-176962928	85	ENSG00000175879
86	chr2.176962928-176986460	86	ENSG00000128652
87	chr2.176986460-177028714	87	ENSG00000187944
88	chr2.177028714-177039801	88	ENSG00000155754
89	chr2.177134307-179245704	89	ENSG00000168530

90	chr2.179245704-189168448	90	ENSG00000236445
91	chr2.189168448-190028935	91	ENSG00000135903
92	chr2.242878496-242887053	92	ENSG00000163081
93	chr2.242887053-243030735	93	ENSG00000263828
94	chr3.32279948-32611876	94	ENSG00000144460
95	chr3.32611876-32904291	95	ENSG00000072080
96	chr3.32904291-34767621	96	ENSG00000168505
97	chr3.34767621-37332133	97	ENSG00000163359
98	chr3.37332133-38306594	98	ENSG00000233608
99	chr3.38306594-38358115	99	ENSG00000134115
100	chr3.38358115-38561466	100	ENSG00000207625
101	chr3.38561466-39821421	101	ENSG00000184345
102	chr3.39821421-41240643	102	ENSG00000196578
103	chr3.42572745-43184996	103	ENSG00000233412
104	chr3.93646378-96494778	104	ENSG00000196098
105	chr3.122458846-123398346	105	ENSG00000206536
106	chr3.123398346-124684891	106	ENSG00000231861
107	chr3.124684891-125053815	107	ENSG00000177707
108	chr3.125053815-126668519	108	ENSG00000243197
109	chr3.126668519-129001312	109	ENSG00000187715
110	chr3.129001312-129035211	110	ENSG00000206384
111	chr3.129035211-129690199	111	ENSG00000182447
112	chr3.130064583-135684445	112	ENSG00000144962
113	chr3.135684445-136002143	113	ENSG00000157005
114	chr3.136002143-138724228	114	ENSG00000187527
115	chr3.138724228-143315342	115	ENSG00000169676
116	chr3.143315342-144899243	116	ENSG00000250819
117	chr3.147074517-147079053	117	ENSG00000248228
118	chr3.147232332-148048753	118	ENSG00000168214
119	chr3.151867537-153839136	119	ENSG00000154277
120	chr4.645440-726697	120	ENSG00000109132
121	chr4.726697-888127	121	ENSG00000215203
122	chr4.1742857-3278748	122	ENSG00000163453
123	chr4.5021111-5167599	123	ENSG00000248505
124	chr4.6021538-8098270	124	ENSG00000205678
125	chr4.8098270-8253235	125	ENSG00000152591
126	chr4.8253235-9352827	126	ENSG00000145358
127	chr4.9922462-11369349	127	ENSG00000186867
128	chr4.11427700-13968277	128	ENSG00000227145
129	chr4.13968277-16142263	129	ENSG00000250341
130	chr4.16142263-17310680	130	ENSG00000196951
131	chr4.17310680-18321009	131	ENSG00000145416
132	chr4.18321009-20253130	132	ENSG00000249106
133	chr4.24583375-36333625	133	ENSG00000179059
134	chr4.36333625-37224078	134	ENSG00000179046
135	chr4.37224078-38615922	135	ENSG00000205097
136	chr4.38615922-40999951	136	ENSG00000113430
137	chr4.90117697-91570154	137	ENSG00000249326
138	chr4.91570154-92034737	138	ENSG00000170549
139	chr4.92034737-94078925	139	ENSG00000216077
140	chr4.94078925-95078953	140	ENSG00000145526
141	chr4.107017800-109933679	141	ENSG00000259663
142	chr4.109933679-110563530	142	ENSG00000016082
143	chr4.110563530-111508999	143	ENSG00000249352
144	chr4.111508999-111559066	144	ENSG00000164326
145	chr4.111559066-113261845	145	ENSG00000171540
146	chr4.113261845-113430656	146	ENSG00000145777
147	chr4.113430656-113558105	147	ENSG00000172901
148	chr4.113558105-114416067	148	ENSG00000145794
149	chr4.114416067-116894979	149	ENSG00000265691
150	chr4.116894979-121079963	150	ENSG00000131435
151	chr4.121079963-121843035	151	ENSG00000251380
152	chr4.131032074-132984152	152	ENSG00000146013
153	chr4.132984152-134589655	153	ENSG00000146001
154	chr4.134589655-135123103	154	ENSG00000204956
155	chr4.135123103-139550400	155	ENSG00000241956
156	chr4.139550400-142253700	156	ENSG00000253236
157	chr4.142253700-144435124	157	ENSG00000145934
158	chr4.144435124-147097364	158	ENSG00000164438
159	chr4.147097364-151051929	159	ENSG00000161055

160	chr4.151051929-153586344	160	ENSG00000248717
161	chr4.153586344-154951775	161	ENSG00000263572
162	chr4.154951775-155875702	162	ENSG00000124827
163	chr4.155875702-157892489	163	ENSG00000124575
164	chr4.157892489-158281194	164	ENSG00000198327
165	chr4.158281194-158817904	165	ENSG00000256316
166	chr4.158817904-160277142	166	ENSG00000197459
167	chr4.160277142-170022074	167	ENSG00000168242
168	chr4.170022074-171663610	168	ENSG00000124657
169	chr4.171663610-176297391	169	ENSG00000204704
170	chr4.176297391-181480012	170	ENSG00000204701
171	chr4.181480012-183178361	171	ENSG00000204695
172	chr4.183178361-184016270	172	ENSG00000213886
173	chr4.184016270-184060895	173	ENSG00000124721
174	chr4.184060895-185090043	174	ENSG00000187871
175	chr4.185090043-186652454	175	ENSG00000112175
176	chr4.186652454-187250670	176	ENSG00000146143
177	chr4.187250670-189548707	177	ENSG00000256980
178	chr4.189548707-189578842	178	ENSG00000112837
179	chr4.189578842-190228298	179	ENSG00000135355
180	chr4.190228298-190239020	180	ENSG00000112214
181	chr4.190239020-190475000	181	ENSG00000112238
182	chr4.190475000-190966117	182	ENSG00000152034
183	chr5.71654404-72471841	183	ENSG00000229315
184	chr5.72471841-72516057	184	ENSG00000112246
185	chr5.72562402-72593048	185	ENSG00000203808
186	chr5.72593048-72595434	186	ENSG00000146352
187	chr5.72595434-72750474	187	ENSG00000186439
188	chr5.72750474-72795181	188	ENSG00000135547
189	chr5.72795181-72858664	189	ENSG00000231023
190	chr5.72858664-74657448	190	ENSG00000118526
191	chr5.74657448-74965039	191	ENSG00000177468
192	chr5.74965039-75464408	192	ENSG00000112038
193	chr5.75464408-76921025	193	ENSG00000229720
194	chr5.76921025-76925446	194	ENSG00000223485
195	chr5.76925446-77043612	195	ENSG00000227508
196	chr5.77043612-79951536	196	ENSG00000171243
197	chr5.79951536-86567376	197	ENSG00000122585
198	chr5.86567376-89352028	198	ENSG00000197576
199	chr5.89352028-92641165	199	ENSG00000078399
200	chr5.94890411-95469201	200	ENSG00000207584
201	chr5.106606747-107569937	201	ENSG00000240990
202	chr5.107569937-108746499	202	ENSG00000243766
203	chr5.110847639-112074043	203	ENSG00000205628
204	chr5.122110639-122847966	204	ENSG00000266168
205	chr5.122847966-122950706	205	ENSG00000185037
206	chr5.123804109-124091260	206	ENSG00000183166
207	chr5.124091260-124919308	207	ENSG00000006128
208	chr6.99151099-100037332	208	ENSG00000106236
209	chr6.108280315-110418140	209	ENSG00000130427
210	chr6.110418140-110680256	210	ENSG00000234715
211	chr6.127439206-129203075	211	ENSG00000205174
212	chr6.129203075-134700557	212	ENSG00000154438
213	chr6.134700557-136657129	213	ENSG00000128610
214	chr6.136657129-140392531	214	ENSG00000230316
215	chr6.140392531-142277658	215	ENSG00000106331
216	chr6.142277658-144652488	216	ENSG00000224865
217	chr6.144652488-145271257	217	ENSG00000221938
218	chr6.145271257-146056419	218	ENSG00000130226
219	chr6.149068439-149803087	219	ENSG00000240247
220	chr6.149803087-150213172	220	ENSG00000239839
221	chr6.150213172-150347012	221	ENSG00000187082
222	chr6.150347012-151054883	222	ENSG00000186599
223	chr6.151711584-152379044	223	ENSG00000186562
224	chr6.152379044-153023279	224	ENSG00000186579
225	chr6.153023279-153241931	225	ENSG00000250305
226	chr6.153552436-156722340	226	ENSG00000164741
227	chr6.156722340-157713044	227	ENSG00000199127
228	chr6.168045602-168952719	228	ENSG00000003987
229	chr6.168952719-170038733	229	ENSG00000180053

230	chr8.29111-1004106	230	ENSG00000104722
231	chr8.1004106-1250400	231	ENSG00000183729
232	chr8.1250400-1879639	232	ENSG00000164736
233	chr8.1879639-1972405	233	ENSG00000104237
234	chr8.1972405-2196764	234	ENSG00000254254
235	chr8.2196764-2418843	235	ENSG00000167910
236	chr8.2418843-2421548	236	ENSG00000254300
237	chr8.2421548-2545988	237	ENSG00000133742
238	chr8.2545988-2586911	238	ENSG00000170289
239	chr8.2586911-2793467	239	ENSG00000176571
240	chr8.2793467-3416209	240	ENSG00000254318
241	chr8.3416209-3965218	241	ENSG00000180543
242	chr8.30515826-30853662	242	ENSG00000104375
243	chr8.47015425-49037460	243	ENSG00000154188
244	chr8.79716756-80239356	244	ENSG00000174417
245	chr8.80239356-80780259	245	ENSG00000245164
246	chr8.80780259-80992750	246	ENSG00000250400
247	chr8.108144638-109493386	247	ENSG00000155897
248	chr9.72961-176893	248	ENSG00000254083
249	chr9.176893-7478187	249	ENSG00000253288
250	chr9.7478187-7800256	250	ENSG00000155886
251	chr9.7800256-10613328	251	ENSG00000188921
252	chr9.10613328-15810627	252	ENSG00000228083
253	chr9.15810627-16729515	253	ENSG00000120242
254	chr9.16729515-19049082	254	ENSG00000225626
255	chr9.19049082-20996718	255	ENSG00000207935
256	chr10.8094142-8097331	256	ENSG00000188386
257	chr10.8097331-8099018	257	ENSG00000186881
258	chr10.89580388-91295045	258	ENSG00000030304
259	chr11.46330260-46450256	259	ENSG00000173077
260	chr11.118661599-118993093	260	ENSG00000204148
261	chr11.118993093-119177430	261	ENSG00000078725
262	chr11.128169380-128693677	262	ENSG00000235865
263	chr11.128693677-129088588	263	ENSG00000085265
264	chr11.129088588-130184046	264	ENSG00000236990
265	chr11.130184046-130207578	265	ENSG00000224382
266	chr11.130207578-131416251	266	ENSG00000225269
267	chr11.131416251-131685205	267	ENSG00000197308
268	chr11.131685205-132641831	268	ENSG00000243350
269	chr11.134145889-134373284	269	ENSG00000107485
270	chr11.134373284-134832162	270	ENSG00000229240
271	chr12.12416946-13251495	271	ENSG00000233968
272	chr12.13431373-16759029	272	ENSG00000204033
273	chr12.16759029-19002078	273	ENSG00000138109
274	chr12.19002078-20091056	274	ENSG00000148798
275	chr12.25707564-26110518	275	ENSG00000108018
276	chr12.26110518-26822563	276	ENSG00000175535
277	chr12.26822563-27088580	277	ENSG00000154478
278	chr12.27088580-27234924	278	ENSG00000148848
279	chr12.27234924-27500572	279	ENSG00000188722
280	chr12.27500572-29299659	280	ENSG00000186766
281	chr12.29299659-29381810	281	ENSG00000188069
282	chr14.47872288-49934863	282	ENSG00000176895
283	chr14.49934863-50320632	283	ENSG00000205495
284	chr14.51240822-51443494	284	ENSG00000176787
285	chr14.51443494-52413159	285	ENSG00000171944
286	chr14.52413159-52622696	286	ENSG00000176742
287	chr14.52622696-54080601	287	ENSG00000213934
288	chr14.54080601-54430831	288	ENSG00000213931
289	chr14.54430831-57046431	289	ENSG00000183251
290	chr14.57046431-57284528	290	ENSG00000167355
291	chr14.57284528-60558039	291	ENSG00000184881
292	chr14.60558039-60952097	292	ENSG00000176239
293	chr14.60952097-60973773	293	ENSG00000184698
294	chr14.60973773-61116048	294	ENSG00000167359
295	chr14.61116048-61119122	295	ENSG00000175520
296	chr14.61119122-61120234	296	ENSG00000181023
297	chr14.61120234-61122774	297	ENSG00000181001
298	chr14.61122774-61793064	298	ENSG00000180988
299	chr14.61793064-66970921	299	ENSG00000183269

300	chr14.66970921-67916110
301	chr14.67916110-71374302
302	chr14.71374302-71581031
303	chr14.71581031-72887638
304	chr14.72887638-73603844
305	chr14.73957626-74752123
306	chr14.74752123-78352846
307	chr14.78352846-84867542
308	chr14.84867542-85667733
309	chr14.85667733-86204203
310	chr16.56518581-57223804
311	chr16.57223804-59786279
312	chr17.64412745-64567021
313	chr17.76801065-76836835
314	chr17.76836835-77021038
315	chr19.7459819-8934543
316	chr19.8934543-9203303
317	chr20.47001709-47895337
318	chr20.47895337-48330123
319	chr20.48330123-48448154
320	chr20.48448154-50419348
321	chr20.54581276-56725695
322	chr20.56725695-60925179
323	chr20.60925179-61506981
324	chr20.61506981-62948235

300	ENSG00000184478
301	ENSG00000180919
302	ENSG00000184933
303	ENSG00000166368
304	ENSG00000170743
305	ENSG00000175868
306	ENSG00000199077
307	ENSG00000254519
308	ENSG00000182053
309	ENSG00000214891
310	ENSG00000205035
311	ENSG00000166007
312	ENSG00000181767
313	ENSG00000181761
314	ENSG00000167822
315	ENSG00000181752
316	ENSG00000181698
317	ENSG00000181689
318	ENSG00000150261
319	ENSG00000172487
320	ENSG00000150269
321	ENSG00000255012
322	ENSG00000172459
323	ENSG00000174914
324	ENSG00000172457
325	ENSG00000181273
326	ENSG00000172381
327	ENSG00000197786
328	ENSG00000172362
329	ENSG00000176495
330	ENSG00000172324
331	ENSG00000166884
332	ENSG00000176200
333	ENSG00000172742
334	ENSG00000166900
335	ENSG00000077498
336	ENSG00000214414
337	ENSG00000223417
338	ENSG00000165323
339	ENSG00000134640
340	ENSG00000082175
341	ENSG00000149968
342	ENSG00000255484
343	ENSG00000137634
344	ENSG00000255248
345	ENSG00000207971
346	ENSG00000259571
347	ENSG00000109943
348	ENSG00000213184
349	ENSG00000196661
350	ENSG00000171053
351	ENSG00000254607
352	ENSG00000254960
353	ENSG00000255087
354	ENSG00000273409
355	ENSG00000047617
356	ENSG00000177575
357	ENSG00000172243
358	ENSG00000134533
359	ENSG00000134538
360	ENSG00000069431
361	ENSG00000256321
362	ENSG00000187950
363	ENSG00000139209
364	ENSG00000139567
365	ENSG00000167768
366	ENSG00000135426
367	ENSG00000197706
368	ENSG00000185821
369	ENSG00000090382

370	ENSG00000127329
371	ENSG00000111046
372	ENSG00000180318
373	ENSG00000180219
374	ENSG00000189238
375	ENSG00000181234
376	ENSG00000175664
377	ENSG00000133105
378	ENSG00000139800
379	ENSG00000182346
380	ENSG00000041515
381	ENSG00000182968
382	ENSG00000176281
383	ENSG00000255582
384	ENSG00000221977
385	ENSG00000100842
386	ENSG00000092009
387	ENSG00000198807
388	ENSG00000189139
389	ENSG00000258949
390	ENSG00000248550
391	ENSG00000186910
392	ENSG00000201672
393	ENSG00000225746
394	ENSG00000260792
395	ENSG00000212380
396	ENSG00000166206
397	ENSG00000228740
398	ENSG00000104044
399	ENSG00000232431
400	ENSG00000137766
401	ENSG00000171956
402	ENSG00000260305
403	ENSG00000259170
404	ENSG00000185551
405	ENSG00000184140
406	ENSG00000182854
407	ENSG00000155719
408	ENSG00000122254
409	ENSG00000166501
410	ENSG00000103375
411	ENSG00000150394
412	ENSG00000263785
413	ENSG00000183024
414	ENSG00000133020
415	ENSG00000108688
416	ENSG00000108700
417	ENSG00000181374
418	ENSG00000172716
419	ENSG00000236320
420	ENSG00000197262
421	ENSG00000132130
422	ENSG00000204897
423	ENSG00000186393
424	ENSG00000171446
425	ENSG00000173908
426	ENSG00000187242
427	ENSG00000171431
428	ENSG00000212899
429	ENSG00000212900
430	ENSG00000221880
431	ENSG00000214518
432	ENSG00000240871
433	ENSG00000212722
434	ENSG00000212721
435	ENSG00000198271
436	ENSG00000244537
437	ENSG00000198443
438	ENSG00000240542
439	ENSG00000204873

440	ENSG00000241595
441	ENSG00000131738
442	ENSG00000234859
443	ENSG00000171360
444	ENSG00000108759
445	ENSG00000180336
446	ENSG00000260027
447	ENSG00000210741
448	ENSG00000229637
449	ENSG00000136457
450	ENSG00000154975
451	ENSG00000141198
452	ENSG00000121075
453	ENSG00000186407
454	ENSG00000266714
455	ENSG00000265369
456	ENSG00000152214
457	ENSG00000101542
458	ENSG00000166634
459	ENSG00000166396
460	ENSG00000179676
461	ENSG00000166342
462	ENSG00000263958
463	ENSG00000256463
464	ENSG00000268119
465	ENSG00000269037
466	ENSG00000225872
467	ENSG00000267640
468	ENSG00000197446
469	ENSG00000105732
470	ENSG00000131408
471	ENSG00000142511
472	ENSG00000180279
473	ENSG00000215998
474	ENSG00000179709
475	ENSG00000198046
476	ENSG00000268182
477	ENSG00000127903
478	ENSG00000152467
479	ENSG00000125878
480	ENSG00000125900
481	ENSG00000089012
482	ENSG00000232528
483	ENSG00000125816
484	ENSG00000125820
485	ENSG00000132671
486	ENSG00000100987
487	ENSG00000180383
488	ENSG00000233354
489	ENSG00000131059
490	ENSG00000168703
491	ENSG00000124157
492	ENSG00000101446
493	ENSG00000243543
494	ENSG00000249139
495	ENSG00000101448
496	ENSG00000149651
497	ENSG00000101017
498	ENSG00000054803
499	ENSG00000215386
500	ENSG00000232263
501	ENSG00000197683
502	ENSG00000186980
503	ENSG00000198390
504	ENSG00000184351
505	ENSG00000206106
506	ENSG00000186930
507	ENSG00000186924
508	ENSG00000187005
509	ENSG00000182591

510	ENSG00000205927
511	ENSG00000233316
512	ENSG00000160200
513	ENSG00000237989
514	ENSG00000225637
515	ENSG00000100181
516	ENSG00000230922
517	ENSG00000205634

9 Publications and Presentations

Publications

Jäkel C, Bergmann F, Toth R, Assenov Y, van der Duin D, Strobel O, Hank T, Klöppel G, Dorrell C, Grompe M, Moss J, Dor Y, Schirmacher P, Plass C, Popanda O, Schmezer P. Genome-wide genetic and epigenetic analyses of pancreatic acinar cell carcinomas reveal aberrations in genome stability and cell cycle control (manuscript submitted)

Oral Presentations

Genome-wide genetic and epigenetic analyses in pancreatic acinar cell carcinomas reveal aberrations in cell cycle control and genome stability Jäkel C *9th Epigenetics@dkfz workshop,* Heidelberg, Germany, January 2017

(Epi-) genetic genome-wide screens in pancreatic acinar cell carcinomas reveal aberrations in cell cycle control and genome stability Siebenkäs C, Bergmann F, Toth R, Plass C, Popanda O, Schmezer P *3rd German-Catalan Workshop on Epigenetics & Disease,* Freiburg, Germany, October 2016

Poster Presentations

Genome-wide genetic and epigenetic screens in pancreatic acinar cell carcinomas reveal druggable aberrations in cell cycle control and genome stability Jäkel C, Bergmann F, Toth R, Plass C, Popanda O, Schmezer P *Helmholtz International Graduate School for Cancer Research 2016 Ph.D. Poster Presentation* Heidelberg, Germany, November 2016
Awarded with poster prize

Epigenetic and genetic characterization of pancreatic acinar cell carcinomas Siebenkäs C, Bergmann F, Plass C, Popanda O, Schmezer P *21th Annual DKFZ PhD Retreat,* Weil der Stadt, Germany, July 2016

Epigenetic and genetic characterization of pancreatic acinar cell carcinomas Siebenkäs C, Bergmann F, Plass C, Popanda O, Schmezer P *24th Biennial Congress of the European Association for Cancer Research,* Manchester, United Kingdom, July 2016

10 Acknowledgements

First of all I would like to thank Peter Schmezer and Odilia Popanda for guiding me through the PhD thesis, for great input and discussions and always having an open door for questions, but also for the time outside of the lab at group events. Then I would like to thank Christoph Plass for giving me the great opportunity to perform my PhD thesis in his division and for his inputs and advices. I also thank Frank Bergmann for the collaboration on this project, for collecting the tumors and finding time in his very busy schedule.

I would like to thank all current and former Schmezer group members for their continuous support, for all answered questions and for an always nice atmosphere. Especially I would like to thank Peter W. and Reini for their support, for always listening and forgiving me when I was in bad moods in stressful times. Christoph W. who helped me a lot in the beginning of my PhD with everything in the lab, and he always had great inputs for my project. Thanks to all hard-working students during that time that made life in the lab cheerful, especially Laura and Clara.

Thank you to all the other current and former members of the division (Moni, Oli, Marion, Tania, Mélanie, Justyna, Annika, Sim, Reini II, Anders, Maria, Kathi, Karin, Caro, Anna-Lena, Clarissa F., Sina, Marina, Mridul, Annette, Wolfgang, Susanna, Yassen, Mischka, Olga, Pavlo, Nicole, Michael, David, Reka, Clarissa G., Dieter, and Daniel) for the nicest atmosphere in and outside the lab, for all your help in the lab, inputs at meetings and answering my questions. Without you my work would not have been possible. Thank you for great lunch times at the Mensa and in the kitchen, and for continuous sportive motivation for ”Unisport” and running groups.

In particular, thank you to all the bioinformaticians who helped along the way, Olga, Yassen and especially Reka: thank you for making bioinformatics understandable! Thank you, Clarissa G. for recruiting the people of the following section. Maria and Kathi, thank you for everything you were great, funny and supportive neighbors at the beginning of my PhD and thank you Clarissa F. you were the best possible follower for them with your kindness, and endless-seeming happiness. Also thanks to all the great and friendly students: Kevin, Anna-Lena, Caro and Sarah. Special thanks to all the technicians in our division for providing all the support during the years and for being such nice people. Thank you, Marion, Moni, Oli, and Annette. And Karin, thank you for the uncountable answers, tips and tricks you taught me, especially during the beginning of my time in the division.

Thank you ”best-office-in-the-lab” Sim, Reini II and Clara for you-know-what. I wish you would have been there from the start and it will be a hard time ever finding colleagues that can keep up with you.

Am Schluss möchte ich meiner Familie und meinen Freunden für ihre Unterstützung danken. Besonderer Dank gilt hierbei meinen Eltern, welche nie an mir gezweifelt haben und wahrscheinlich immer noch denken ich könnte alles. Danke für eure Unterstützung während des Studiums und der Doktorarbeit. Danke, dass ihr immer da seid und es mir nie nachgetragen habt, wenn ich der Forschung höhere Priortät gegeben habe. Danke auch an Sonja, Franzi, Anna und Conny. Zeiten mit euch waren immer wie Urlaub in dem ich den Laboralltag ausschalten konnte. Danke, dass auch ihr mir es verziehen habt, wenn ich nicht immer Zeit hatte und/oder zerstreut war.

Henning, du bist mein Fels in der Brandung und mein Zuhause. Danke, dass du immer für mich da bist, mir den Rücken stärkst, mich ablenkst, alles mitmachst und meine Launen vorallem gegen Ende so tapfer ertragen hast. Ohne dich hätte alles nicht (so gut) funktioniert.