

Aus der Klinik für Gastroenterologie, gastrointestinale Onkologie und Endokrinologie
(Prof. Dr. med. V. Ellenrieder)
der Medizinischen Fakultät der Universität Göttingen

**Genetische Polymorphismen in Typ-III-Interferon-Genen
und deren prognostische Signifikanz für das hepatozelluläre
Karzinom und das duktales Pankreasadenokarzinom**

INAUGURAL-DISSERTATION
zur Erlangung des Doktorgrades
der Medizinischen Fakultät der
Georg-August-Universität zu Göttingen

vorgelegt von
Henriette Huschka
aus
Hannover

Göttingen 2020

Dekan:

Herr Prof. Dr. med. W. Brück

Betreuungsausschuss

Betreuerin:

Frau Prof. Dr. rer. nat. S. Mihm

Ko-Betreuerin:

Frau Prof. Dr. H. Bickeböllner

Prüfungskommission

Betreuerin:

Frau Prof. Dr. rer. nat. S. Mihm

Ko-Referent:

Herr Prof. Dr. rer. nat. T. Beißbarth

Promotor:

Herr Prof. Dr. med. M. Oppermann

Datum der mündlichen Prüfung:

16.03.2021

Hiermit erkläre ich, die Dissertation mit dem Titel „Genetische Polymorphismen in Typ-III-Interferon-Genen und deren prognostische Signifikanz für das hepatozelluläre Karzinom und das duktales Pankreasadenokarzinom“ eigenständig angefertigt und keine anderen als die von mir angegebenen Quellen und Hilfsmittel verwendet zu haben.

Göttingen, den 29.07.2020

.....

(Unterschrift)

Inhaltsverzeichnis

Abbildungsverzeichnis	IV
Tabellenverzeichnis	VI
Abkürzungsverzeichnis	VII
1 Einleitung	1
1.1 Das hepatozelluläre Karzinom (HCC)	2
1.1.1 Ätiologie des HCC	3
1.1.2 Klinische Aspekte des HCC	5
1.2 Das duktales Pankreasadenokarzinom (PDAC)	7
1.2.1 Ätiologie des PDAC	8
1.2.2 Klinische Aspekte des PDAC	10
1.3 Interferone (IFNe)	13
1.3.1 Zellspezifische Expressions- und Wirkorte der IFNe	13
1.3.2 Pathogenerkennung sowie Aktivierung der IFNe und ihrer Effektoren	14
1.3.3 Antivirale und antitumorale Wirkungen sowie klinische Nutzung der IFNe	17
1.3.4 Genetische Polymorphismen in Typ-III-IFN-Genen	18
1.4 <i>The Cancer Genome Atlas</i> (TCGA)	22
1.5 Fragestellung	24
2 Material und Methoden	25
2.1 <i>The Cancer Genome Atlas</i> (TCGA)	25
2.1.1 Frei zugängliche Datensätze	25
2.1.2 Zugangsberechtigung	26
2.1.3 Datensätze mit kontrolliertem Zugang	27
2.1.4 Auslesen der polymorphen Positionen	27
2.2 Das 1000-Genome-Projekt	29
2.3 Statistische Verfahren	29
2.3.1 Deskriptive und explorative Statistik	30
2.3.2 Induktive Statistik	32
2.3.3 Bestimmung des Hardy-Weinberg-Gleichgewichts (HWE)	36
2.3.4 Bestimmung des Kopplungsungleichgewichts (LD)	36

2.3.5 Ereigniszeitanalysen: Verfahren nach Kaplan-Meier und Cox-Regression	38
3 Ergebnisse	45
3.1 Beschreibung der TCGA-LIHC- und TCGA-PAAD-Patientenkollektive .	45
3.1.1 Übersicht über die demografischen und klinischen Daten des TCGA-LIHC-Kollektivs	45
3.1.2 Übersicht über die demografischen und klinischen Daten des TCGA-PAAD-Kollektivs	47
3.2 Erheben von Genotypen polymorpher Positionen in den Typ-III-IFN-Genen sowie Identifizierung und Validierung von <i>tag</i> SNPs	49
3.3 Stratifizierung der HCC- und PDAC-Patienten europäischer Herkunft nach ihren <i>IFNL</i> -Genotypen	53
3.3.1 Vergleich der TCGA-LIHC- und TCGA-PAAD-Kollektive mit einem Kontrollkollektiv aus dem 1000-Genome-Projekt hinsichtlich der Genotypverteilung	55
3.3.2 Analyse der Strata hinsichtlich demografischer und klinischer Daten	56
3.4 Analysen zur Tumorprogression	59
3.4.1 Tumorprogression definiert durch die Dauer des progressionsfreien Intervalls (PFI) und des spezifischen progressionsfreien Intervalls (PFI.2)	59
3.4.2 Tumorprogression definiert durch die Gesamtüberlebensdauer (OS)	65
3.4.3 Zusammenfassung der Assoziationen mit Formulierung einer möglichen <i>Follow-Up</i> -Kontrolle	71
4 Diskussion	73
4.1 Assoziation zwischen den Typ-III-IFN-Genen und der PDAC-Progression	74
4.2 Therapie der TCGA-LIHC- und TCGA-PAAD-Patientenkollektive . . .	75
4.3 Duale Wirkungsweise der IFNe und humane Pseudogenisierungsmechanismen	77
4.4 Einfluss der Typ-III-IFNe auf die Tumorprogression des HCC	80
4.5 Prognostik der Tumorprogression	84
4.6 Limitationen der durchgeführten Analysen	85
5 Zusammenfassung	87

6	Anhang	89
6.1	Das Analysekonzept, die Datenressourcen und Software	89
6.2	Deskriptive und explorative Statistik	92
6.3	Stratifizierung der Daten des TCGA-LIHC-Kollektivs nach der Ätiologie	97
6.4	Mediane ereignisfreie Zeiten und mediane Beobachtungszeiträume	99
6.5	Ereigniszeitanalyse des spezifischen progressionsfreien Intervalls (PFI.2) nach dem Kaplan-Meier-Verfahren für nach klinischen Merkmalen strati- fizierte HCC- und PDAC-Patienten	102
7	Literaturverzeichnis	104

Abbildungsverzeichnis

1.1	Die <i>Hallmarks of Cancer</i> mit möglichen medikamentösen Angriffspunkten	2
1.2	Darstellung der Häufigkeit genetischer Mutationen im HCC und die veränderten Zellfunktionen oder Signalwege	4
1.3	Stadieneinteilung und Therapieoptionen von HCC-Patienten	6
1.4	Darstellung der Vielfalt der Genomalterationen des PDAC	9
1.5	Behandlungsstrategien für PDAC-Patienten eingeteilt nach Tumorresektabilität	12
1.6	Mechanismen von der Pathogendetektion bis zur Aktivierung der IFN-stimulierten Gene (ISGe)	15
1.7	IFN-vermittelte Aktivierung der IFN-stimulierten Gene (ISGe) über den JAK-STAT-Signalweg	16
1.8	Darstellung des <i>IFNL</i> -Genlokus mit ausgewählten Polymorphismen . . .	19
1.9	Zusammenhang zwischen den <i>IFNL3/4</i> -rs12979860-Genotypen und der spontanen Viruselimination bei Hepatitis C	20
2.1	Exemplarische Kaplan-Meier-Kurvendarstellung	40
2.2	Zusammenfassung der Arbeitsprozesse	44
3.1	Übersicht über den <i>IFNL</i> -Genlokus auf Chromosom 19 innerhalb der NCBI <i>Genome Workbench</i>	49
3.2	Übersicht über den <i>IFNL3</i> - und den <i>IFNL4</i> -Genlokus mit den fünf markierten polymorphen Positionen (M1–M5) innerhalb der NCBI <i>Genome Workbench</i>	50
3.3	Beispiel für einen <i>minor</i> homozygoten [GG] und einen heterozygoten [CG] <i>IFNL3</i> -rs28416813-Genlokus	51
3.4	Beispielhafte Darstellung des <i>IFNL4</i> -rs368234815-Genlokus	52
3.5	Berechnung des Kopplungsungleichgewichts (<i>Linkage Disequilibrium</i> , LD) der fünf <i>IFNL</i> -Polymorphismen	54
3.6	Ereigniszeitanalyse der progressionsfreien Intervalle nach dem Kaplan-Meier-Verfahren für HCC-Patienten	63
3.7	Ereigniszeitanalyse der progressionsfreien Intervalle nach dem Kaplan-Meier-Verfahren für PDAC-Patienten	64
3.8	Ereigniszeitanalyse der Gesamtüberlebensdauer (OS) nach dem Kaplan-Meier-Verfahren für HCC-Patienten	66

3.9	Ereigniszeitanalyse der Gesamtüberlebensdauer (OS) nach dem Kaplan-Meier-Verfahren für PDAC-Patienten	68
A.1	Altersverteilung der HCC- und der PDAC-Patienten bei Diagnose	92
A.2	Altersverteilung der nach <i>IFNL3</i> -rs4803217- bzw. <i>IFNL3</i> -rs28416813-Genotypen stratifizierten HCC- bzw. PDAC-Patienten	93
A.3	Verteilung der Anteile der nach <i>IFNL3</i> -rs4803217-Genotypen CC, CA, AA stratifizierten HCC-Patienten auf die benannten Risikofaktoren	94
A.4	Verteilung der Anteile der nach <i>IFNL3</i> -rs4803217-Genotypen CC, CA, AA stratifizierten HCC-Patienten auf die klinischen Kriterien Tumordifferenzierungsgrad bzw. Tumorstadium	95
A.5	Verteilung der Anteile der nach <i>IFNL3</i> -rs28416813-Genotypen CC, CG, GG stratifizierten PDAC-Patienten auf die klinischen Kriterien Tumordifferenzierungsgrad bzw. Tumorstadium	96
A.6	Ereigniszeitanalyse des spezifischen progressionsfreien Intervalls (PFI.2) nach dem Kaplan-Meier-Verfahren für nach klinischen Merkmalen stratifizierte HCC-Patienten	102
A.7	Ereigniszeitanalyse des spezifischen progressionsfreien Intervalls (PFI.2) nach dem Kaplan-Meier-Verfahren für nach klinischen und genetischen Merkmalen stratifizierte PDAC-Patienten	103

Tabellenverzeichnis

1.1	Allelvarianten ausgewählter Polymorphismen im Bereich der <i>IFNL3</i> - und <i>IFNL4</i> -Gene sowie beschriebene Phänotypen	22
2.1	Übersicht über die in NCBI <i>Genome Workbench</i> markierten und ausgelesenen <i>IFNL</i> -Polymorphismen auf Chromosom 19	28
3.1	Demografische und klinische Daten des TCGA-LIHC-Kollektivs	46
3.2	Demografische und klinische Daten des TCGA-PAAD-Kollektivs	48
3.3	Vergleich der TCGA-LIHC- und TCGA-PAAD-Kollektive mit einem Kontrollkollektiv aus dem 1000-Genome-Projekt hinsichtlich der Genotypverteilung	55
3.4	Analysen zur Assoziation der <i>IFNL3</i> -rs4803217-Genotypen mit demografischen und klinischen Daten für HCC-Patienten	58
3.5	Analysen zur Assoziation der <i>IFNL3</i> -rs28416813-Genotypen mit demografischen und klinischen Daten für PDAC-Patienten	59
3.6	Übersicht über die einfache und die multiple Cox-Regressionsanalyse für die drei Ereigniszeiten PFI/PFI.2/OS der HCC-Patienten	69
3.7	Übersicht über die einfache und die multiple Cox-Regressionsanalyse für die drei Ereigniszeiten PFI/PFI.2/OS der PDAC-Patienten	70
A.1	Zusammenfassung des durchgeführten Analysekonzepts	90
A.2	Zusammenfassung der Datenressourcen und der Software	91
A.3	Stratifizierung des TCGA-LIHC-Kollektivs nach der Ätiologie	98
A.4	Übersicht über die medianen ereignisfreien Zeiten und die medianen Beobachtungszeiträume der HCC-Patienten für alle drei Ereigniszeiten . . .	100
A.5	Übersicht über die medianen ereignisfreien Zeiten und die medianen Beobachtungszeiträume der PDAC-Patienten für alle drei Ereigniszeiten . .	101

Abkürzungsverzeichnis

AJCC	<i>American Joint Committee on Cancer</i>
DKG	Deutsche Krebsgesellschaft
DKH	Deutsche Krebshilfe
eRA	<i>electronic Research Administration</i>
FOLFIRINOX	Folinsäure/5-Fluoruracil (5-FU)/Irinotecan/Oxaliplatin
GDC	<i>Genomic Data Commons</i>
H_0	Nullhypothese
H_1	Alternativhypothese
HBV	Hepatitis-B-Virus
HCC	Hepatozelluläres Karzinom (<i>hepatocellular carcinoma</i>)
HCV	Hepatitis-C-Virus
HWE	Hardy-Weinberg-Gleichgewicht (<i>Hardy-Weinberg-Equilibrium</i>)
IFN	Interferon
<i>IFNL</i>	<i>IFN lambda</i>
ISG	IFN-stimuliertes Gen
JAK	Januskinase
KI	Konfidenzintervall
LD	Kopplungsungleichgewicht (<i>Linkage Disequilibrium</i>)
LIHC	<i>liver hepatocellular carcinoma</i>
MAF	Häufigkeit des selteneren Allels (<i>minor allele frequency</i>)
NCBI	<i>National Center for Biotechnology Information</i>
NCI	<i>National Cancer Institute</i>
OS	Gesamtüberlebensdauer (<i>overall survival</i>)
PAAD	<i>pancreatic adenocarcinoma</i>
PDAC	Duktales Pankreasadenokarzinom (<i>pancreatic ductal adenocarcinoma</i>)
PFI	Progressionsfreies Intervall (<i>progression free interval</i>)
PFI.2	Spezifisches progressionsfreies Intervall (<i>progression free interval.2</i>)
ROS	Reaktive Sauerstoffspezies (<i>reactive oxygen species</i>)
SNP	Einzelnukleotidpolymorphismus (<i>single nucleotide polymorphism</i>)
STAT	<i>signal transducers and activators of transcription</i>
TCGA	<i>The Cancer Genome Atlas</i>

Abkürzungsverzeichnis

TCGA-CDR	<i>TCGA Pan-Cancer Clinical Data Resource</i>
WXS	<i>whole exome sequencing</i>

1 Einleitung

Tumoren stellen weltweit eine führende Todesursache dar, deren Inzidenz auch aufgrund der demografischen Entwicklung weiter steigt. Neben anderen Studien fasst die Studie GLOBOCAN 2018 die jüngsten Tumorentwicklungen zusammen (Bray et al. 2018). Die Tumorentitäten mit der höchsten Inzidenz weltweit sind Lungentumoren bei beiden Geschlechtern (11,6 %), Brusttumoren bei Frauen (11,6 %), Prostatatumoren bei Männern (7,1 %) und kolorektale Tumoren bei beiden Geschlechtern (6,1 %). Die häufigsten Todesfälle in Bezug zu allen Krebstodesfällen, definiert durch die Kennzahl Mortalität, werden durch Lungentumoren (18,4 %), gefolgt von kolorektalen Tumoren (9,2 %), Lebertumoren (8,2 %) und Magentumoren (8,2 %) verursacht. Zu den tödlichsten Tumoren, deren 5-Jahres-Überlebensrate unter 20 % liegt, zählen jedoch Pankreastumoren mit einer 5-Jahres-Überlebensrate von 8 % und Lebertumoren sowie Lungentumoren mit einer 5-Jahres-Überlebensrate von jeweils 18 % (Bray et al. 2018; Siegel et al. 2018).

Die Maßnahmen zur Bekämpfung von Tumorerkrankungen umfassen neben der Prävention, d. h. dem Vermeiden von bekannten Risikofaktoren, die Tumorfrüherkennung sowie die Entwicklung suffizienter Therapiemöglichkeiten zur Heilung von Tumoren (Bray et al. 2018).

Zu den vermeidbaren Auslösern von Krebserkrankungen zählen unter anderem Zigarettenrauch für Lungentumoren, virale oder bakterielle Infektionen für Zervix-, Magen- und Lebertumoren und ionisierende Strahlung für Schilddrüsenkarzinome (Bray et al. 2018). Screeningmethoden, die zur Früherkennung von Tumoren oder auch bereits ihrer Vorläufer dienen, sind zum Beispiel der Test auf okkultes Blut im Stuhl (*Fecal Occult Blood Test*, FOBT) und die Koloskopie zur Früherkennung von kolorektalen Tumoren (Fedewa et al. 2015).

Für eine Heilung ist die Kenntnis der Tumorgenese und der Tumoreigenschaften entscheidend. Der Tumorgenese liegen häufig Genominstabilitäten und -veränderungen zugrunde. Dazu gehören unter anderem Mutationen, die eine Resistenz gegenüber der Apoptose vermitteln, Mutationen, die zur Überexpression von Wachstumsfaktoren und damit zur unkontrollierten Proliferation führen oder Mutationen, die die Erkennung maligner Zellen durch das Immunsystem unterbinden (Hanahan und Weinberg 2011).

Diese *Hallmarks of Cancer* bilden Angriffspunkte für die therapeutische Intervention (Abb. 1.1).

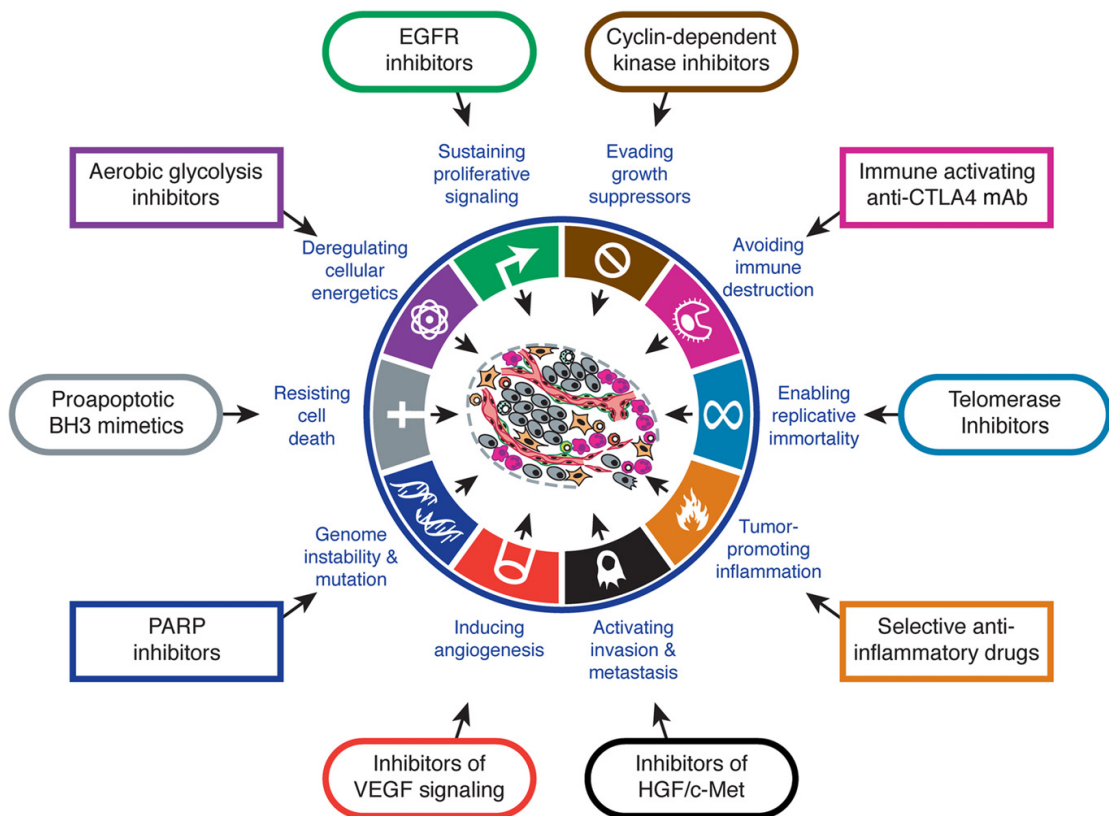


Abb. 1.1: Die *Hallmarks of Cancer* mit möglichen medikamentösen Angriffspunkten.

Durch Genominstabilitäten und -veränderungen können gesunde Zellen zu invasiven und metastasierenden Tumorzellen werden (Hanahan und Weinberg 2011). Die Verwendung der Abbildung erfolgt mit freundlicher Genehmigung von Elsevier.

Neben den spezifischen Eigenschaften des Tumors ist insbesondere die Tumorabwehr des Wirts entscheidend für das Tumorwachstum. Die Balance zwischen Tumorwachstum und -abwehr macht die Tumorprogression aus. Neuere Tumorthera­pieansätze zielen auf eine Stärkung der immunsystemvermittelten Abwehr ab und versuchen, durch den Einsatz sogenannter Immuntherapeutika das patienteneigene Immunsystem gegen den Tumor zu sensibilisieren (Yang 2015).

1.1 Das hepatozelluläre Karzinom (HCC)

Global werden etwa 841.000 neue Fälle des hepatozellulären Karzinoms (*hepatocellular carcinoma*, HCC) pro Jahr diagnostiziert und ca. 782.000 Todesfälle dokumentiert, mit den höchsten Raten in den Ländern Nord- und Westafrikas sowie Südostasiens (Bray et al. 2018).

In Deutschland erkrankten im Jahr 2016 etwa 9.000 Menschen an HCC und ca. 8.000 Patienten verstarben, wobei weniger als 40 % der HCC-Patienten das erste Jahr nach Diagnose überlebten (Robert Koch-Institut 2019). Bei Männern repräsentiert das HCC die fünft häufigste Tumorart und bei Frauen ist es die acht häufigste Tumorart (Bray et al. 2018). Wie die meisten Tumorerkrankungen ist auch das HCC eine Erkrankung des höheren Lebensalters mit einem Inzidenzmaximum von etwa 64 Jahren (NCI 2020b).

1.1.1 Ätiologie des HCC

Studien zeigten, dass in gesunden Lebern fast nie ein HCC entsteht (Marrero et al. 2018). In der chronisch geschädigten Leber, in der sich eine Fibrose bis hin zur Zirrhose entwickeln kann, steigt das Risiko für die Entstehung eines HCC signifikant an. Von den HCC-Patienten weisen 80–90% eine Leberzirrhose auf und 5–30% der Patienten mit Leberzirrhose entwickeln nach fünf Jahren ein HCC (Marrero et al. 2018).

Die chronische Leberschädigung mit der Folge einer Leberzirrhose wird durch verschiedene Faktoren verursacht, die damit gleichzeitig die Risikofaktoren für die Entstehung des HCC darstellen. Dazu zählen Infektionen mit dem Hepatitis-B-Virus (HBV) oder dem Hepatitis-C-Virus (HCV), Intoxikationen durch Alkohol, Medikamente oder Aflatoxin B1 (AFB1), nutritive Intoxikationen oder Leberzellverfettung und genetische Dispositionen im Leberstoffwechsel wie zum Beispiel die hereditäre Hämochromatose (Bray et al. 2018). Das gleichzeitige Vorliegen mehrerer Risikofaktoren innerhalb eines Patienten ist möglich.

Der Tumorgenese des HCC zugrunde liegende Genominstabilitäten und -veränderungen wurden durch die genomweite Sequenzierung aller proteincodierenden Regionen (*whole exome sequencing*, WXS) detektiert. Diese Sequenzierung von insgesamt 243 Tumorgewebeproben zeigt einer französischen Studie entsprechend 161 genetische Veränderungen, die 11 verschiedene Signalwege betreffen (Schulze et al. 2015). Hierbei wurde beobachtet, dass bekannte Treibermutationen (*driver mutations*) anderer Tumorentitäten wie die Mutationen des *EGFR* (*epidermal growth factor receptor*)-Gens, des *BRAF* (*B-Raf proto-oncogene, serine/threonine kinase*)-Gens, des *PIK3CA* (*phosphatidylinositol-4,5-bisphosphate 3-kinase catalytic subunit alpha*)-Gens oder des *KRAS* (*KRAS proto-oncogene, GTPase*)-Gens bei dem HCC nur selten auftreten (Schulze et al. 2016) (Abb. 1.2). Stattdessen zählen zu den drei am häufigsten mutierten Genen innerhalb des HCC: (i) aktivierende Mutationen im Promoterbereich des *TERT* (*telomerase reverse transcriptase*)-Gens, (ii) vorwiegend funktionsbeeinträchtigende *Missense*-Mutationen des Tumorsuppressorgens *TP53* (*tumor protein p53*) und (iii) den WNT (*wingless integrated*)-Signalweg aktivierende Mutationen im *CTNNB1* (*catenin*

beta 1)-Gen (Schulze et al. 2015) (Abb. 1.2). Zusätzlich sind die Gene *AXIN1* (*axin 1*), *ARID1A* (*AT-rich interaction domain 1A*), *CDKN2A* (*cyclin dependent kinase inhibitor 2A*) und *CCND1* (*cyclin D1*) häufig betroffen (Schulze et al. 2015) (Abb. 1.2).

Die Analyse der Daten von Schulze et al. (2015) ergab zusätzlich, dass das Auftreten der Mutationen mit Risikofaktoren assoziiert: Mutationen des *CTNNB1* (*catenin beta 1*)- und des *TERT* (*telomerase reverse transcriptase*)-Gens assoziieren mit der alkoholinduzierten Leberzirrhose, wohingegen Mutationen im *TP53* (*tumor protein p53*)-Gen mit dem HBV-bedingten HCC assoziieren (Abb. 1.2). Dagegen kommt es bei dem HCC unklarer Ätiologie vermehrt zu Mutationen des *IL6ST* (*interleukin 6 signal transducer*)-Gens. Für das HCV-bedingte HCC konnte keine Assoziation mit einer spezifischen Mutation festgestellt werden (Schulze et al. 2015).

Neben diesen molekularen Anomalien in Zellsignalwegen weisen viele HCC-Patienten geänderte Angiogenesemechanismen und epigenetische Veränderungen auf (Schulze et al. 2016).

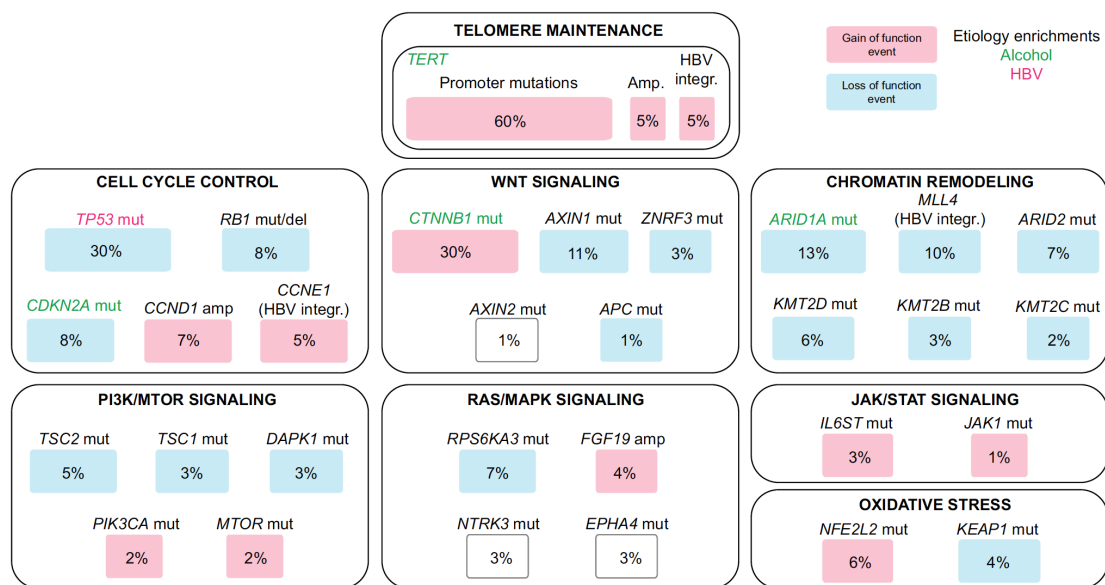


Abb. 1.2: Darstellung der Häufigkeit genetischer Mutationen im HCC und die veränderten Zellfunktionen oder Signalwege.

Mehrere Mutationen können gleichzeitig in einer HCC-Tumorzelle auftreten. Am häufigsten werden Mutationen im Promoter des *TERT* (*telomerase reverse transcriptase*)-Gens, des Tumorsuppressorgens *TP53* (*tumor protein p53*) und des *CTNNB1* (*catenin beta 1*)-Gens detektiert (Schulze et al. 2016). Die Verwendung der Abbildung erfolgt mit freundlicher Genehmigung von Elsevier.

1.1.2 Klinische Aspekte des HCC

Vor der Therapie als sekundärprophylaktische Maßnahme steht die Primärprävention durch Vermeidung oder Verringerung der zum HCC führenden Risikofaktoren. Es wurde nachgewiesen, dass landesweit eingeleitete HBV-Impfungen zu einer deutlichen Reduzierung der Inzidenz des HCC führen, sodass eine HBV-Impfung jedem Neugeborenen und allen Menschen mit hohem Infektionsrisiko empfohlen wird (Bray et al. 2018; Marrero et al. 2018). Liegt eine chronische HBV-Infektion vor, ist die konsequente antivirale Therapie mit Interferonen (IFNen) oder Lamivudin die derzeitige Methode, um das Risiko für die Entwicklung eines HCC signifikant zu reduzieren (Marrero et al. 2018). Auch bei Vorliegen einer HCV-Infektion konnte durch eine Eradikation des Virus mittels IFN-basierter Therapie bei Patienten mit und ohne Leberzirrhose eine eindeutige HCC-Risikoreduktion erreicht werden. Die Voraussetzung für die Risikoreduktion war das dauerhaft virologische Ansprechen auf die Therapie (Marrero et al. 2018).

Bei jeglicher Form der chronischen Lebererkrankung wird zur Alkoholabstinenz geraten. Bei einer nicht-alkoholinduzierten Steatohepatitis wird zur Gewichtsreduktion und bei Diabetes mellitus-induzierten Lebererkrankungen zu einer Therapie mit Metformin geraten (DKG/DKH 2013).

Zur weiteren HCC-Prophylaxe empfehlen die Leitlinien der AASLD (*American Association for the Study of Liver Disease*) die aktive Überwachung von HCC-Risikopatienten mittels Abdomensonografie. Die Bestimmung des Lebertumormarkers Alpha-Fetoprotein (AFP) im Serum ist umstritten (Marrero et al. 2018).

Nicht für die aktive Überwachung, aber für die Diagnose eines HCC wird die Bildgebung mittels Mehrphasen-CT oder -MRT und gegebenenfalls die Biopsie auffälliger Herde in der Leber mit anschließender pathologischer Untersuchung empfohlen (Marrero et al. 2018).

Symptome des HCC treten zumeist erst in fortgeschrittenen Tumorstadien durch die Einschränkung der physiologischen Funktion der Leber auf und umfassen unter anderem Druckschmerzen im rechten Oberbauch, Vergrößerung des Abdomens mit Appetitverlust und Übelkeit, Gewichtsabnahme und Leistungsminderung, Aszites, Ödeme der unteren Extremitäten sowie Skleren- und Hautikterus (Tunissioli et al. 2017).

In Europa werden Patienten mit diagnostiziertem HCC auf der Basis des Tumorstadiums, des Allgemeinzustands der Patienten, der Anzahl und Größe der Tumorherde, der Tumorsymptome und der Leberfunktion, beschrieben durch die Child-Pugh-Klassifikation, in Stadien eingeteilt (nach der *Barcelona Clinic Liver Cancer* (BCLC)-Klassifikation) (Forner et al. 2010) (Abb. 1.3). Mithilfe dieser Einteilung (sehr frühes Stadium, Frühstadium, Intermediärstadium, fortgeschrittenes Stadium und Endstadium) kann die Pro-

gnose der HCC-Patienten individuell abgeschätzt und die jeweilige Therapie danach ausgerichtet werden (Forner et al. 2010) (Abb. 1.3). Bei Patienten im sehr frühen Stadium/Frühstadium des HCC wird eine kurative Therapie mittels Resektion, Ablation oder Transplantation angestrebt, wohingegen Patienten in einem Intermediärstadium eine Chemoembolisation empfohlen wird (Forner et al. 2010). Bei Patienten in einem fortgeschrittenen Tumorstadium erfolgt die Standardbehandlung primär systemisch mittels des Multikinaseinhibitors Sorafenib als Erstlinientherapie. Patienten im Endstadium wird nur noch der Leidensdruck durch eine symptomorientierte Palliativtherapie (*Best Supportive Care*, BSC) genommen (Forner et al. 2010) (Abb. 1.3).

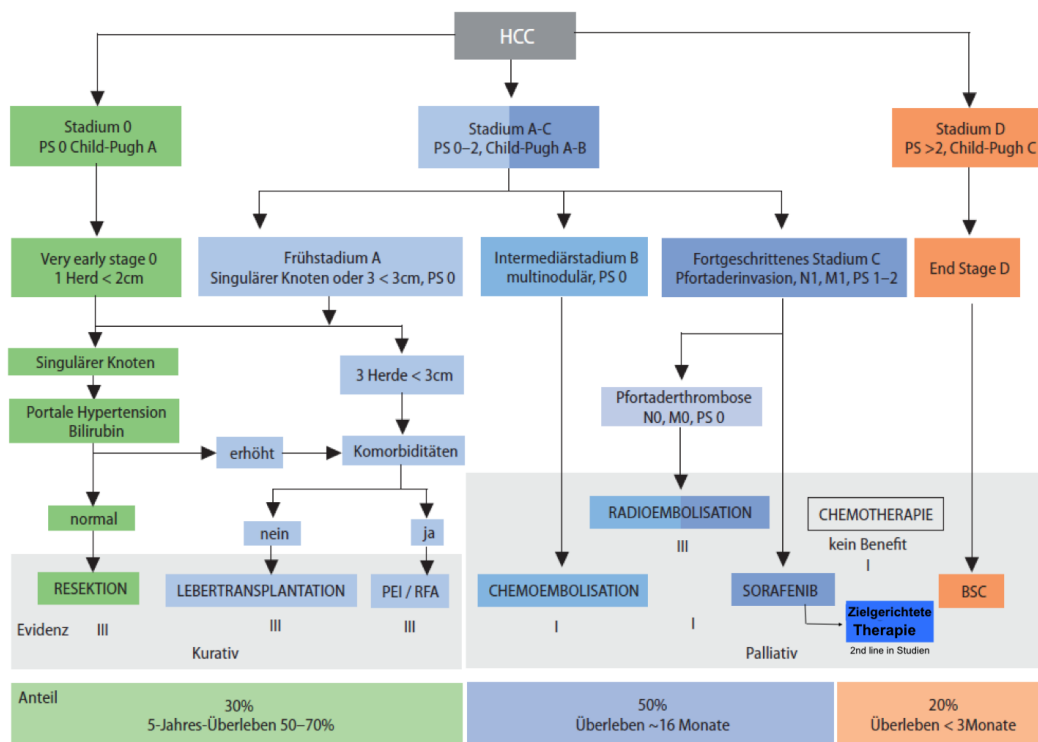


Abb. 1.3: Stadieneinteilung und Therapieoptionen von HCC-Patienten.

Bei Patienten im sehr frühen Stadium 0 und Fröhstadium A wird eine kurative Therapie mittels Transplantation, Resektion oder Ablation (Radiofrequenzablation (RFA)/ Perkutane Ethanolinjektion (PEI)) angestrebt, wohingegen Patienten im Intermediärstadium B eine Chemoembolisation und Patienten im fortgeschrittenen Stadium C eine primär systemisch Therapie mit dem Multikinaseinhibitor Sorafenib als Erstlinienmedikament empfohlen wird. Bei keinem Ansprechen können Patienten im fortgeschrittenen Stadium C auch innerhalb einer Studie in eine zielgerichtete Therapie eingeschlossen werden. Patienten im Endstadium D wird nur noch der Leidensdruck durch eine symptomatische Palliativtherapie genommen (*Best Supportive Care*, BSC) (modifiziert nach: Geier et al. 2012). Die Verwendung der Abbildung erfolgt mit freundlicher Genehmigung von Springer Nature.

Eine aktuelle Therapieoption bilden die sogenannten Immuntherapeutika, die an der Stärkung der wirtseigenen Tumorabwehr ansetzen (vgl. S. 2 oben). Dazu zählen unter anderem Inhibitoren des PD-1 (*programmed cell death protein 1*)-Rezeptors, der auf Immunzellen vorliegt und bei Aktivierung das Immunsystem herunterreguliert. Des Weiteren können als Immuntherapeutika Inhibitoren der auf Tumorzellen exprimierten Liganden PD-L1/2 (*programmed cell death 1 ligand 1/2*) wie Nivolumab eingesetzt werden. Nivolumab wurde bereits als Zweitlinienmedikament für die Behandlung von HCC-Patienten im fortgeschrittenen Stadium zugelassen (Marrero et al. 2018). Die Überexpression der Liganden PD-L1/2 (*programmed cell death 1 ligand 1/2*) auf Tumorzellen ist ein wichtiger Mechanismus der Tumorzellen zur Immunevasion, da durch die Bindung des Liganden an den PD-1 (*programmed cell death protein 1*)-Rezeptor die Apoptose von Immunzellen eingeleitet wird (Lázár-Molnár et al. 2017; Marrero et al. 2018).

Für die Entdeckung und Beschreibung, diesen Mechanismus der Immunevasion zu unterbinden und damit einen neuen Angriffspunkt innerhalb der Tumorthherapie zu ermöglichen, wurde im Jahr 2018 der Nobelpreis an Allison und Honjo verliehen (Fares et al. 2019).

1.2 Das duktales Pankreasadenokarzinom (PDAC)

Das duktales Pankreasadenokarzinom (*pancreatic ductal adenocarcinoma*, PDAC) ist ein Tumor des exokrinen Pankreas (Haeberle und Esposito 2019). Global werden etwa 459.000 neue Fälle des PDAC pro Jahr diagnostiziert und ca. 432.000 Todesfälle dokumentiert mit den höchsten Raten in Europa, Nordamerika und Australien/Neuseeland (Bray et al. 2018). In Deutschland sind im Jahr 2016 insgesamt etwa 18.400 Menschen an PDAC erkrankt, wobei die Todesfallzahl aufgrund der ungünstigen Prognose fast ebenso viele Personen umfasste (Robert Koch-Institut 2019). Aufgrund fehlender spezifischer Symptome wird das PDAC erst sehr spät entdeckt, worin sich die schlechte Prognose von einem medianen Überleben nach Erstdiagnose von nur wenigen Monaten für die PDAC-Patienten begründet (Siegel et al. 2018). Wie auch bereits das HCC ist das PDAC eine Erkrankung des höheren Lebensalters mit 80 % der Patienten zwischen 60 und 80 Jahren und nur sehr wenigen Patienten unter 40 Jahren. Hierbei liegt die Inzidenz und Mortalität bei Männern höher als bei Frauen (Reig et al. 2016). Das PDAC stellt sowohl klinisch als auch ätiologisch und molekular einen sehr vielfältigen Tumor dar, der häufig eine Therapieresistenz aufweist (Raphael et al. 2017).

1.2.1 Ätiologie des PDAC

Das PDAC weist vielfältige, sehr komplexe und interindividuell unterschiedliche Mutationen auf, wobei es sich zumeist um somatische Mutationen, hier vor allem um Punktmutationen, handelt (Raphael et al. 2017).

Bailey et al. (2016) identifizierten in ihren Genomanalysen von 456 PDAC-Tumorgewebe-proben 32 wiederkehrend mutierte Gene, die sich 10 molekularen Pfaden zuordnen lassen: dem RAS/MAPK (*KRAS protooncogene/mitogen-activated protein kinase*)-Signalweg, der G1/S-Transition, dem TGF- β (*transforming growth factor beta 1*)-Signalweg, den Histonmodifizierungen, dem SWI/SNF (*SWItch/sucrose non-fermentable*)-Komplex, dem Tumorsuppressorgen BRCA (*breast cancer*), dem WNT (*wingless integrated*)-Signalweg, dem NOTCH1 (*notch receptor 1*)-Gen, der RNA-Prozessierung und der DNA-Reperatur. Die vier am häufigsten wiederkehrend mutierten Gene sind hierbei das Onkogen *KRAS* (*KRAS proto-oncogene, GTPase*) in 93 % der PDAC-Fälle, das Tumorsuppressorgen *TP53* (*tumor protein p53*) in 72 % der Fälle, das *SMAD4* (*SMAD family member 4*)-Gen in 32 % der Fälle und das *CDKN2A* (*cyclin dependent kinase inhibitor 2A*)-Gen in 30 % der Fälle (Raphael et al. 2017) (Abb. 1.4). Die weniger häufigen *KRAS* (*KRAS proto-oncogene, GTPase*)-Wildtyp-tumoren weisen ein anderes Mutationsspektrum und andere Signalwegveränderungen auf. Unter anderem wurden Mutationen in den *GNAS* (*GNAS complex locus*)-Genen, den *BRAF* (*B-Raf proto-oncogene, serine/threonine kinase*)-Genen, den *CTNNB1* (*catenin beta 1*)-Genen und auch in anderen *RAS*-Genen detektiert (Raphael et al. 2017). Dieser Unterschied der Mutationen nimmt vermutlich Einfluss auf die klinischen Therapieergebnisse der PDAC-Patienten (Raphael et al. 2017).

Wie viele andere Tumoren kann auch das PDAC aus präkanzerösen Läsionen hervorgehen, wobei diese in mikroskopische und makroskopische Läsionen eingeteilt werden können (Haeberle und Esposito 2019). Zu den mikroskopischen Präkanzerosen zählen hauptsächlich die nichtinvasiven pankreatisch intraepithelialen Neoplasien (PanIN), bei denen je nach Grad der Zell- und Kernatypie ein *low-grade* von einem *high-grade* Stadium unterschieden werden kann (Haeberle und Esposito 2019). Die Stadien weisen auch typische Mutationen auf, wobei die *KRAS* (*KRAS proto-oncogene, GTPase*)-Mutation sowohl in dem *low-grade* als auch in dem *high-grade* Stadium vorkommt. Die *CDKN2A* (*cyclin dependent kinase inhibitor 2A*)-, *TP53* (*tumor protein p53*)- und *SMAD4* (*SMAD family member 4*)-Mutationen wurden hingegen nur in dem *high-grade* Stadium detektiert (Haeberle und Esposito 2019).

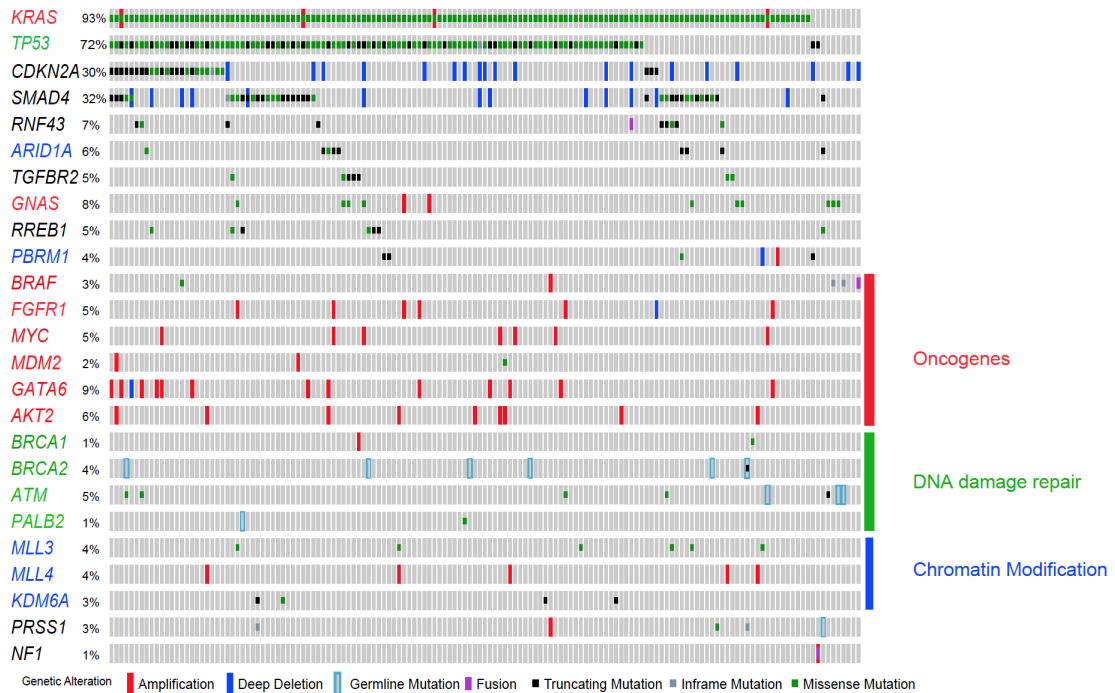


Abb. 1.4: Darstellung der Vielfalt der Genomalalterationen des PDAC.

Die Übersicht zeigt Mutationen in 149 Tumoren, deren Art (*Missense Mutation*, *Inframe Mutation* und andere) farbig gekennzeichnet ist. Auf der Basis der betroffenen Gene konnten die meisten drei funktionellen Gruppen zugeordnet werden: Onkogene, DNA-Reparaturgene und Gene, die die Chromatinmodifikation regulieren. Auf der linken Seite ist die Detektionshäufigkeit der Mutationen innerhalb der 149 Tumoren in Prozent angegeben (modifiziert nach: Raphael et al. 2017). Die Verwendung der Abbildung erfolgt mit freundlicher Genehmigung von Elsevier.

Zu den makroskopischen Präkanzerosen zählen hauptsächlich die intraduktal papillär muzinösen Neoplasien (IPMN) sowie die muzinös zystischen Neoplasien (MCN) (Haeberle und Esposito 2019). Auch die intraduktal papillär muzinösen Neoplasien (IPMN) können in ein *low-grade* und ein *high-grade* Stadium eingeteilt werden, wobei neben den *KRAS* (*KRAS proto-oncogene, GTPase*)-, *CDKN2A* (*cyclin dependent kinase inhibitor 2A*)-, *TP53* (*tumor protein p53*)- und *SMAD4* (*SMAD family member 4*)-Mutationen häufiger als bei den pankreatisch intraepithelialen Neoplasien (PanIN) zusätzlich eine *GNAS* (*GNAS complex locus*)-Mutation vorliegt (Haeberle und Esposito 2019).

Die muzinös zystischen Neoplasien (MCN) stellen zystische Pankreasläsionen vorwiegend bei Frauen dar, die die gleichen Mutationen wie die intraduktal papillär muzinösen Neoplasien (IPMN) aufweisen. Eine Ausnahme hierbei bildet die *GNAS* (*GNAS complex locus*)-Mutation, die deutlich seltener vorliegt (Singhi et al. 2018).

Das klinische Ziel ist die rechtzeitige Erkennung und Elimination dieser präkanzerösen Läsionen, bevor sie sich zu invasiven Karzinomen entwickeln. Zur Diagnose von Pankre-

aszysten und zur Detektion fortgeschrittener Neoplasien gibt es vorläufige Ansätze, präoperativ Pankreaszystenflüssigkeit mit der hochsensitiven und spezifischen NGS (*Next Generation Sequencing*)-Methode zu untersuchen (Singhi et al. 2018).

Meist unterliegt dem PDAC eine sporadische Tumorgenese (*sporadic pancreatic cancer*, SPC), jedoch wird geschätzt, dass 10 % der PDACs einen familiären Hintergrund aufweisen (Matsubayashi et al. 2017). Pankreaskarzinompatienten mit einem familiären Pankreaskarzinom (*familial pancreatic cancer*, FPC) sind definiert als Patienten mit mindestens zwei erstgradigen Verwandten mit Pankreaskarzinom ohne das Vorliegen eines Syndroms. Dieser familiäre Hintergrund führt zu einem signifikanten Anstieg des Risikos, an einem Pankreaskarzinom zu erkranken (Matsubayashi et al. 2017).

Einige wenige hereditäre Pankreaskarzinomfälle stehen im Zusammenhang mit einem Syndrom, unter anderem mit dem Peutz-Jeghers-Syndrom, der hereditären Pankreatitis und dem familiären atypischen multiplen Muttermal- und Melanomsyndrom. Des Weiteren zählen zu den Syndromen das hereditäre Brust- und Ovarialkrebssyndrom, das hereditäre kolorektale Karzinom ohne Polyposis und die familiäre adenomatöse Polyposis (Matsubayashi et al. 2017).

Bei diesen hereditären PDAC-Formen sind einige auslösende Keimbahnmutationen bekannt, unter anderem in den *BRCA2* (*BRCA2 DNA repair associated*)-Genen, den *CDKN2A* (*cyclin dependent kinase inhibitor 2A*)-Genen, den *APC* (*APC regulator of WNT signaling pathway*)-Genen, den *MSH2/6* (*mutS homolog 2/6*)-Genen, den DNA-Mismatch-Reparaturgenen, den *STK11* (*serine/threonine kinase 11*)-Genen sowie den *PRSS1/2* (*serine protease 1/2*)-Genen (Matsubayashi et al. 2017).

Neben den familiären und hereditären Risikofaktoren konnten weitere Risikofaktoren detektiert werden. Hierzu zählen Zigaretten- und Alkoholkonsum sowie das Vorliegen einer chronischen Pankreatitis, einer Adipositas oder eines Diabetes mellitus. Des Weiteren werden die bereits erwähnten intraepithelialen Neoplasien (PanIN) und muzinösen Neoplasien (IPMN/MCN) zu den Risikofaktoren gezählt (Matsubayashi et al. 2017).

1.2.2 Klinische Aspekte des PDAC

Das PDAC tritt vorwiegend im Pankreaskopf auf, wodurch es häufig zur Gallengangsobstruktion und damit zu dem erst recht spät auftretenden Leitsymptom des schmerzlosen Ikterus kommt. Weitere Symptome können Gewichtsverlust, Abdominalschmerzen, Steatorrhö und neu aufgetretener Diabetes mellitus sein. Jedoch sind die meisten Patienten mit einem PDAC asymptomatisch, wenn ihr Tumorleiden, welches sich häufig bereits in einem weit fortgeschrittenen oder metastasierten Stadium befindet, als Zufallsbefund entdeckt wird. Aufgrund dessen ist bei 80 % der Patienten das PDAC bei Diagnose bereits fortgeschritten (Ducreux et al. 2015).

CA19-9 ist bis zum heutigen Stand der aussagekräftigste Tumormarker für Pankreas-tumoren, jedoch ist er für die Diagnose nicht aussagekräftig genug, sondern nur als prognostischer Marker und zum Therapiemonitoring bei bereits diagnostizierten PDAC-Patienten verwendbar (Scarà et al. 2015). Aufgrund des Mangels an einem aussagekräftigen Tumormarker existieren noch keine allgemeingültigen Screeningempfehlungen, wodurch eine Früherkennung erschwert wird (Haeberle und Esposito 2019). Deswegen ist die Primärprävention durch Reduktion der Risikofaktoren und durch Lebensstilmodifikation eine sehr wichtige Maßnahme.

Zur Diagnostik und gleichzeitigen Stadieneinteilung des Tumors wird nach den Leitlinien der ESMO (*European Society for Medical Oncology*) zunächst ein CT empfohlen. Bei Bedarf wird die Diagnostik noch durch eine endoskopische Ultraschalluntersuchung oder seltener durch ein MRT erweitert (Ducieux et al. 2015).

Nach einem Expertenkonsens der amerikanischen Hepato-Pankreato-Biliären Assoziation wurde das PDAC in drei Resektabilitätsstadien unterteilt: resektabel, grenzwertig (*borderline*) resektabel und nicht resektabel/metastasiert (Ducieux et al. 2015) (Abb. 1.5). Von dem grenzwertigen (*borderline*) Resektabilitätsstadium kann das lokal fortgeschrittene Stadium als Subgruppe unterschieden werden. Diese Einteilung bildet die Grundlage für die weiteren Therapieoptionen (Abb. 1.5).

Die einzige Möglichkeit, PDAC-Patienten zu heilen, ist die chirurgische Pankreasade-nektomie, die jedoch nur für die etwa 20 % der diagnostizierten Patienten im resektablen Stadium überhaupt in Frage kommt und nur bei 3–4% zur langfristigen Tumorelimination führt (Ducieux et al. 2015). Anschließend werden die Patienten adjuvant mit einer Chemotherapie nach dem FOLFIRINOX-Regime (Folinsäure/5-Fluoruracil (5-FU)/Irinotecan/Oxaliplatin) behandelt (Ducieux et al. 2015) (Abb. 1.5).

Die Patienten mit einem grenzwertig resektablen Tumor werden einer neoadjuvanten (Radio)-/Chemotherapie auch nach dem FOLFIRINOX-Regime zugeführt. Anschließend wird ihre Resektabilität re-evaluieren (Abb. 1.5).

Im fortgeschrittenen Stadium werden vorwiegend Chemotherapeutika verordnet.

Die Wahl des Chemotherapeutikums richtet sich nach dem Allgemeinzustand des Patienten, wobei für Patienten im guten Allgemeinzustand das FOLFIRINOX-Regime in Betracht kommt. Nach aktueller Studienlage soll hauptsächlich darauf geachtet werden, dass keine Gemcitabin-Monotherapie erfolgt (Conroy et al. 2018).

Bei den metastasierten Tumoren werden nach einem palliativen Behandlungskonzept (*Best Supportive Care*, BSC) zunächst die möglicherweise vorhandenen Symptome kontrolliert. Hierzu zählen die Schmerzkontrolle und gegebenenfalls das Erweitern von Duodenum- und Gallengangsobstruktionen. Die weitere Therapie ist von dem Allgemeinzustand des Patienten abhängig und kann von dem alleinigen palliativen Behandlungskon-

zept bei sehr morbidem Patienten bis hin zu einer FOLFIRINOX-Chemotherapie reichen (Ducreux et al. 2015) (Abb. 1.5).

Es wurden bereits Mutationen in PDAC detektiert, die sich für eine zielgerichtete Therapie eignen würden und die Gegenstand aktueller Forschung darstellen. Jedoch konnten bis jetzt noch keine zielgerichteten Therapien von Studien in die Klinik integriert werden (Ducreux et al. 2015).

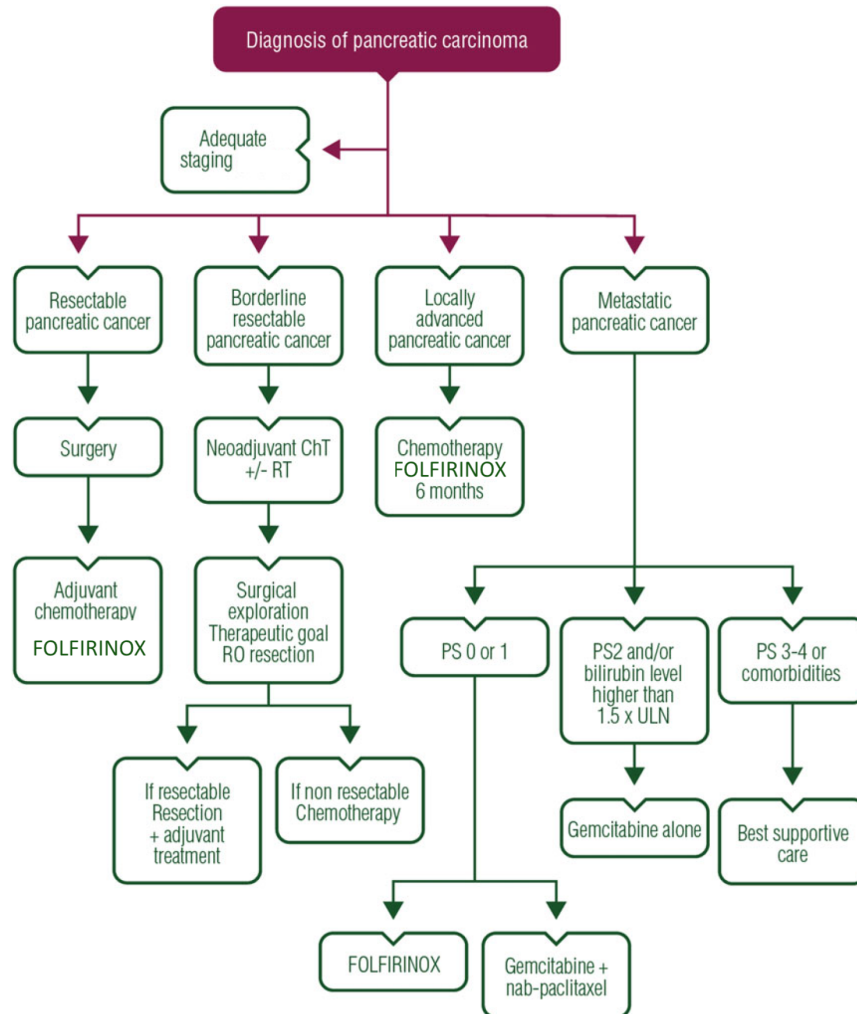


Abb. 1.5: Behandlungsstrategien für PDAC-Patienten eingeteilt nach Tumorresektabilität.

Chemotherapie (ChT), Radiotherapie (RT), FOLFIRINOX (Folinsäure/5-Fluoruracil (5-FU)/Irinotecan/Oxaliplatin), Allgemeinzustand (*performance status*, PS), oberer Grenzwert des Referenzbereichs (*upper limit of normal*, ULN) (modifiziert nach: Ducreux et al. 2015). Die Verwendung der Abbildung erfolgt mit freundlicher Genehmigung von Elsevier.

1.3 Interferone (IFNe)

IFNe wurden 1957 als eine „Aktivität“ entdeckt, die in der Lage war, mit der Replikation von Influenzaviren zu interferieren (Isaacs und Lindemann 1957; Isaacs et al. 1957).

Mittlerweile sind im Menschen 13 sehr ähnliche IFN- α -Proteine und jeweils ein IFN- β -, IFN- κ -, IFN- ω - und IFN- ϵ -Protein bekannt, die zu der Gruppe der Typ-I-IFNe zusammengefasst werden und jeweils von einem eigenen Gen codiert werden. Diesen IFNen ist gemeinsam, dass sie an einen Rezeptor, den IFN- α -Rezeptor (IFNAR), binden (Schneider et al. 2014).

Ein weiteres Protein mit antiviraler Aktivität ist das IFN- γ , das sich als sogenanntes Immun-IFN oder Typ-II-IFN in seiner Wirkungsweise wesentlich von den Typ-I-IFNen und auch den Typ-III-IFNen unterscheidet. Anders als Typ-I- und Typ-III-IFNe ist es kein Vermittler der angeborenen Immunität, sondern vermittelt bei immunologischen Prozessen die erworbene Immunität. Es wird vom *IFNG*-Gen codiert und bindet an einen eigenen Rezeptor: den IFN- γ -Rezeptor (IFNGR) (Schroder et al. 2004).

Typ-III-IFNe wurden erst 2003, Jahre nach der Sequenzierung des humanen Genoms, entdeckt (Collins et al. 2003; Kopenko et al. 2003). Zwei unabhängige Forschungsgruppen haben in der Absicht, weitere antiviral wirksame Proteine zu finden, unter Zuhilfenahme von verschiedenen bioinformatischen Gendatenbanken im humanen Genom nach Typ-I-*IFN*-ähnlichen Genen gesucht und drei Gene identifizieren können: *IL-29*, *IL-28A* und *IL-28B* bzw. *IFNL1*, *IFNL2* und *IFNL3*. Die Proteine werden inzwischen einheitlich als IFN- λ -Proteine bezeichnet. Da sie an einen eigenen Rezeptor (IFN- λ -Rezeptor, IFNLR) binden, wurden sie zu den Typ-III-IFNen zusammengefasst (Kopenko et al. 2003).

Bis zum Jahr 2013 waren nur diese drei IFN- λ -Subtypen bekannt. Erst in Experimenten mit humanen primären Hepatozyten, die mit synthetischer, doppelsträngiger RNA als virales Mimetikum stimuliert wurden, entdeckten Prokunina-Olsson et al. (2013) neben den Transkripten von *IFNL1-3* Transkripte eines weiteren *IFNL*-Gens, das *IFNL4*.

1.3.1 Zellspezifische Expressions- und Wirkorte der IFNe

Die Expression der Typ-I- und der Typ-III-IFNe kann als Teil der angeborenen Immunität in jeder kernhaltigen Körperzelle durch Viren ausgelöst werden. Zusätzlich exprimieren spezialisierte Zellen, die plasmazytoiden dendritischen Zellen (*plasmacytoid dendritic cells*, pDC), große Mengen an Typ-I-IFNen als Antwort auf eine virale Infektion und verknüpfen dadurch die angeborene mit der erworbenen Immunität (Kopenko et al. 2003; Schneider et al. 2014).

Im Unterschied dazu wird das Typ-II-IFN ausschließlich von natürlichen Killerzellen (NK), Makrophagen sowie von T-Lymphozyten exprimiert (Schroder et al. 2004).

Die Zielzellen der IFNe definieren sich über die Expression ihrer Rezeptoren. Die heterodimeren IFN- α -Rezeptoren (IFNAR) sind nahezu ubiquitär exprimiert. Dies erklärt die starken Nebenwirkungen einer Therapie mit IFN- α (O'Brien et al. 2014).

Im Gegensatz zu der auf wenige Zelltypen restringierten Expression des Typ-II-IFN, weist sein homodimerer IFN- γ -Rezeptor (IFNGR) eine weit verbreitete, nahezu ubiquitäre Expression auf (Schneider et al. 2014).

Die heterodimeren IFN- λ -Rezeptoren (IFNLR) werden vorwiegend von epithelialen Zellen exprimiert, vor allem im Respirations- und Gastrointestinaltrakt (Kotenko et al. 2003). Diese restringierte IFN- λ -Rezeptor-Expression lässt weniger unerwünschte Nebenwirkungen unter Therapie erwarten und verspricht somit Vorteile einer Therapie mit Typ-III-IFNen (Mihm 2015).

1.3.2 Pathogenerkennung sowie Aktivierung der IFNe und ihrer Effektoren

Die gemeinsame Wirkung der IFNe besteht in der Kontrolle und Bekämpfung von Pathogenen (Schneider et al. 2014).

Zunächst kommt es zur extra- oder intrazellulären Detektion mikrobieller oder viraler Nukleinsäuren oder anderer Pathogen-assoziiierter molekularer Muster (*pathogen-associated molecular patterns*, PAMP) über eine Vielzahl unterschiedlicher Mustererkennungsrezeptoren (*pattern recognition receptor*, PRR) (Schneider et al. 2014). Zu den Mustererkennungsrezeptoren (*pattern recognition receptor*, PRR) zählen unter anderem die membranständigen TLRs (*toll-like receptors*). Intrazelluläre zytosolische Mustererkennungsrezeptoren (*pattern recognition receptor*, PRR) sind beispielsweise die Doppelstrang-RNA erkennenden RLRs (*retinoic acid-inducible gene 1 (RIG-I)-like receptors*) oder die ALRs (*AIM2-like receptors*), die zytosolische DNA detektieren können (Schneider et al. 2014) (Abb. 1.6).

Nach Erkennung der Pathogen-assoziierten molekularen Muster (*pathogen-associated molecular patterns*, PAMP) kommt es intrazellulär zur Aktivierung des STING (*stimulator of IFN genes*)-Proteins und des MAVS (*mitochondrial antiviral-signaling*)-Proteins (Heim 2013; Schneider et al. 2014) (Abb. 1.6). Daraufhin erfolgt die Translokation der Transkriptionsfaktoren NF- κ B (*nuclear factor 'kappa-light-chain-enhancer' of activated B-cells*) und IFN-regulatorischer Faktoren 3 und 7 (IRF 3/7) in den Nukleus, wo es zur Bindung an aktivierende DNA-Elemente kommt, die die Transkription von IFNen auslösen (Abb. 1.6).

Um die jeweils spezifische antivirale Wirkung ausüben zu können, lösen die IFNe endokrin oder parakrin über ihre spezifischen Rezeptoren die Expression der sogenannten IFN-stimulierten Gene (ISGe) oder IFN-Effektoren aus (Schneider et al. 2014) (Abb. 1.6).

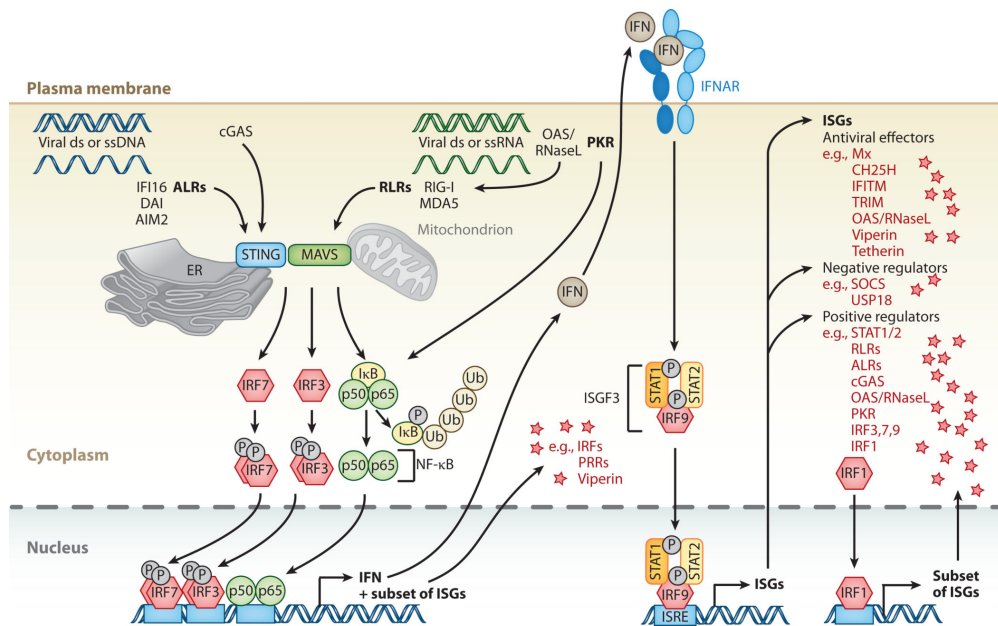


Abb. 1.6: Mechanismen von der Pathogendetektion bis zur Aktivierung der IFN-stimulierten Gene (ISGs).

In dieser Abbildung erkennen intrazelluläre Mustererkennungsrezeptoren die Pathogen-assoziierten molekularen Muster und vermitteln über verschiedene Proteinaktivierungen die IFN-Expression. Die IFNe induzieren über spezifische Signalwege die Translation von einem umfangreichen Spektrum an ISGen, die antivirale, aber auch negative und positive Regulationsfunktionen ausüben (Schneider et al. 2014). Die Verwendung der Abbildung erfolgt mit freundlicher Genehmigung von Annual Reviews.

Die Bindung der IFNe an ihre spezifischen Rezeptoren löst eine Rezeptorkonformationsänderung und dadurch die Aktivierung des JAK-STAT-Signalwegs (Januskinase-*signal transducers and activators of transcription*-Signalweg) aus, wobei es zur Phosphorylierung der intrazellulären IFN-Rezeptorabschnitte kommt. Somit können die extrazellulären Signale der IFNe in den Nucleus transportiert werden (Schneider et al. 2014).

Infolge der Phosphorylierung der IFN-Rezeptorabschnitte kommt es zur Rekrutierung und Phosphorylierung der Transkriptionsfaktoren STAT unter Bildung von STAT-Homo- oder -Heterodimeren (Schneider et al. 2014) (Abb. 1.7). Die intrazelluläre Rekrutierung der STAT1-Homodimere als Transkriptionsfaktoren, die durch alle IFN-Rezeptoren erfolgen kann, führt zur nukleären Bindung der STAT1-Homodimeren an aktivierende DNA-Sequenzelemente (Gamma-aktivierte Sequenzelemente, GAS) der ISGe und löst damit ihre Transkription aus (Abb. 1.7).

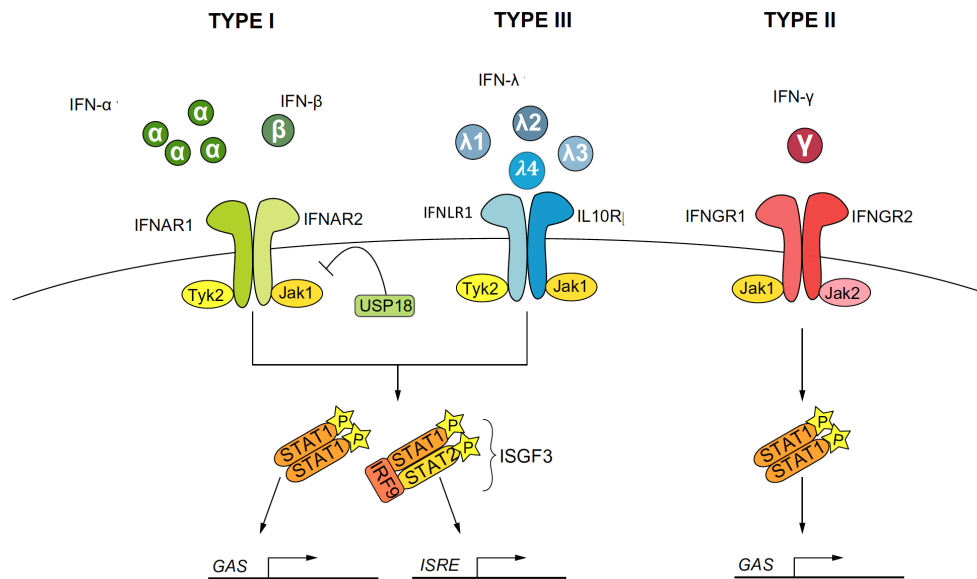


Abb. 1.7: IFN-vermittelte Aktivierung der IFN-stimulierten Gene (ISGe) über den JAK-STAT-Signalweg.

Das Ziel der IFN-Wirkung über den JAK-STAT-Signalweg ist die Induktion von ISGen. Hierbei lösen Typ-I- und Typ-III-IFNe trotz unterschiedlicher Rezeptoren (IFNAR1/2 bzw. IL10R2/IFNLR1) über die Aktivierung von STAT1/2 die gleichen Signalmechanismen aus und führen zu sehr ähnlichen ISG-Expressionen. Das Typ-II-IFN aktiviert über seinen Rezeptor (IFNGR1/2) nur STAT1, wodurch es zu einer teilweise abweichenden ISG-Expression kommt (modifiziert nach: Heim 2013). Die Verwendung der Abbildung erfolgt mit freundlicher Genehmigung von Elsevier.

Typ-I- und Typ-III-IFNe sind zusätzlich in der Lage, durch die Rekrutierung des Heterodimers aus STAT1 und STAT2 sowie durch die zusätzliche Rekrutierung des IFN-regulierenden Faktors 9 (IRF-9) den IFN-stimulierten Genfaktor-3 (ISGF3)-Komplex zu bilden (Heterotrimer) (Abb. 1.7). Dieser Komplex bindet wiederum intranukleär an sogenannte ISRE (*IFN-stimulated response elements*) in den Promotorregionen der ISGe und führt zur deren Transkription (Heim 2013; Schneider et al. 2014) (Abb. 1.7). Anzahl und Art der exprimierten ISGe sind für jeden IFN-Typ und jeden Zelltyp unterschiedlich, wobei zwischen allen IFN-Typen ein gewisses Spektrum an überlappenden ISGen auffindbar ist (Schneider et al. 2014).

Nach der Aktivierung der IFN-Rezeptoren greifen negative Rückkopplungsmechanismen, die einer überschießenden ISG-Antwort entgegenwirken. Vermittelt durch sofortige Endozytose und Abbau der IFN-Rezeptoren sowie der *de novo*-Synthese von Inhibitoren induzieren diese negativen Rückkopplungsmechanismen eine frühe und eine verzögerte, langanhaltende IFN-Desensibilisierung (Heim 2013; Schneider et al. 2014).

Der verzögerte, langanhaltende Rückkopplungsmechanismus wirkt ausschließlich hem-

mend auf die Typ-I-IFNe, hier vor allem auf das IFN- α , und erfolgt über die Ubiquitin-spezifische Peptidase 18 (USP18). Bei diesem Mechanismus kommt es zur Hemmung der Bindung zwischen der IFN- α -Rezeptorkette 2 (IFNAR2) und der JAK 1, wodurch die JAK 1 nicht mehr aktiviert werden kann (Schneider et al. 2014) (Abb. 1.7).

1.3.3 Antivirale und antitumorale Wirkungen sowie klinische Nutzung der IFNe

Die antiviralen Wirkungen der Typ-I- und Typ-III-IFNe beruhen auf den Wirkungen ihrer Effektoren (Kotenko et al. 2003). In verschiedenen Experimenten konnten antivirale Wirkungen von Typ-I- und Typ-III-IFNen *in vivo* in humanen Zellen gegen vielfältige Viren nachgewiesen werden. Unter anderem gegen das Herpes-Simplex-Virus Typ 2 (HSV-2) (Zhou et al. 2015), das Cytomegalievirus (CMV) (Egli et al. 2014b), das Influenzavirus (Egli et al. 2014a), das Dengue-Virus (Palma-Ocampo et al. 2015) sowie *in vitro* gegen HBV und HCV (Ank et al. 2006; Prokunina-Olsson et al. 2013)) konnten die Wirkungen detektiert werden.

Des Weiteren wurden für Typ-I- und Typ-III-IFNe antiproliferative und antitumorale Wirkungen für verschiedene Tumoren beschrieben, darunter für Melanome (Lasfar et al. 2006), Lungenadenokarzinome (Tezuka et al. 2012), Ösophaguskarzinome (Li et al. 2010) und das HCC (Abushahba et al. 2010; Stiff und Carson 2015; Lasfar et al. 2019).

Maher et al. (2008) zeigten eine Abhängigkeit der antiproliferativen und antitumoralen Wirkungen von der Aktivierungsdauer des JAK-STAT-Signalwegs. Hierbei fiel auf, dass die Typ-III-IFNe länger anhaltend aktivieren, wodurch ein effizienterer antiproliferativer Effekt zustande kam als bei den Typ-I-IFNen. Dies zeigt neben der restringierten Rezeptorexpression einen weiteren Vorteil von Typ-III-IFNen gegenüber Typ-I- und Typ-II-IFNen für mögliche antitumorale therapeutische Anwendungen.

Weitere Studien beschrieben, dass das ISG-Spektrum von post-transkriptionalen Modifikationen abhängt. Dadurch wird bestimmt, ob die IFNe antitumoral wirken oder in den Tumoren eine Chemo- und Bestrahlungstherapieresistenz induzieren (duale Wirkung der IFNe durch U-STAT- oder P-STAT-Signalweg) (Minn 2015).

Das Typ-II-IFN ist nicht wie die Typ-I- und Typ-III-IFNe an der direkten Virenbekämpfung über das angeborene Immunsystem beteiligt, sondern wirkt indirekt antiviral über die verstärkte Expression von Proteinen, die in der für die erworbene Immunität obligaten Antigenpräsentation involviert sind, wie zum Beispiel die Proteasomuntereinheiten oder die Haupthistokompatibilitätskomplex-Moleküle (*major histocompatibility complex*, MHC) (Mihm et al. 2004; Schroder et al. 2004).

Auf diese Art kann das Typ-II-IFN die Typ-I-IFN-Wirkungen erhöhen (hier vor allem die Wirkung des IFN- α und des IFN- β), was zu verstärkten antiviralen und antitumo-

ralen Wirkmechanismen führt (Schroder et al. 2004).

Bei einigen Tumorentitäten konnten epigenetische Gen-Stillegungen (*gene silencing*) und somatische Deletionen innerhalb der IFN- α - und IFN- β -Chromosomenloki sowie Keimbahnmutationen der ISGe detektiert werden. Vermutlich sind diese Genänderungen mit einer frühen Immortalität der Tumorzellen assoziiert, was wiederum verdeutlicht, dass aktivierte IFNe antitumorale Wirkungen aufweisen (Kulaeva et al. 2003).

Aufgrund ihrer zahlreichen antiviralen sowie antitumoralen Wirkungen werden die IFNe heutzutage auch im klinischen Alltag verwendet, wobei vorwiegend das IFN- α unter anderem bei viralen Erkrankungen wie dem Hepatitis B und C, bei Tumoren des blutbildenden Systems (unter anderem dem Multiplen Myelom, der chronisch myeloischen Leukämie, dem follikulären Lymphom), bei soliden Tumoren (unter anderem dem renalen Zellkarzinom, dem HCC) und bei Multipler Sklerose (hier vor allem das IFN- β) angewandt wird (Lasfar et al. 2011).

Neben einer Kombination von IFN- α und IFN- λ , bei der in Tiermodellen gezeigt werden konnte, dass feine Unterschiede in der Wirkungsweise von Typ-I- und Typ-III-IFNen existieren, die zum Wirtsvorteil genutzt werden könnten, scheint vor allem die IFN- λ -Monotherapie vielversprechend, da weniger Nebenwirkungen erwartet werden (Lasfar et al. 2008; Maher et al. 2008).

1.3.4 Genetische Polymorphismen in Typ-III-IFN-Genen

Noch vor der Entdeckung des *IFNL4*-Gens wurden durch unabhängige genomweite Assoziationsstudien (GWAS) an Humankollektiven europäischer, afrikanischer, asiatischer und japanischer Abstammung mit bemerkenswerter Übereinstimmung genetische Polymorphismen im Bereich der *IFNL*-Gene identifiziert, die mit der spontanen und der therapieinduzierten Ausheilung einer HCV-Infektion assoziierten (Ge et al. 2009; Supiah et al. 2009; Tanaka et al. 2009; Thomas et al. 2009).

Einer dieser Polymorphismen *IFNL3* rs12979860 befindet sich *upstream* des *IFNL3*-Gens, aufgrund der entgegengesetzten Orientierung von *IFNL3* und *IFNL2* auch *upstream* des *IFNL2* (Abb. 1.8).

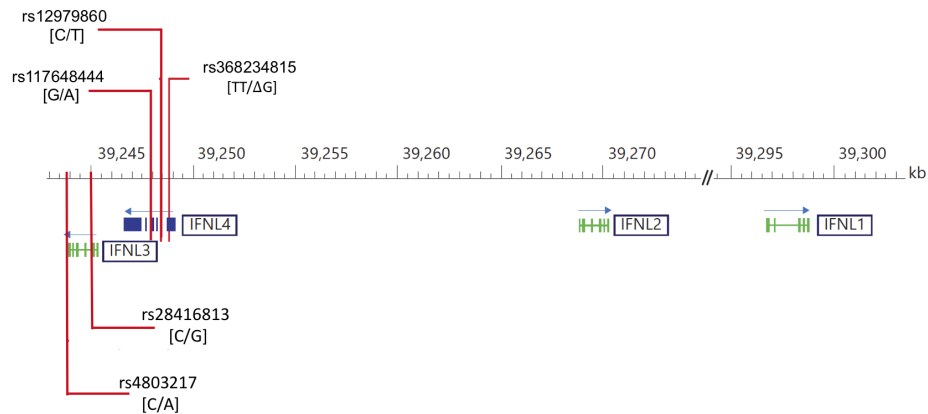


Abb. 1.8: Darstellung des *IFNL*-Genlokus mit ausgewählten Polymorphismen.

Die *IFNL*-Gene befinden sich auf dem Chromosom 19 (19q13.12–13). Die *IFNL3*- und *IFNL4*-Genloki enthalten die in dieser Arbeit betrachteten fünf polymorphen Positionen, über die bereits publiziert wurde: *IFNL3/4* rs12979860 [C/T] (wird mit Fähigkeit der Virenbekämpfung in Zusammenhang gebracht), *IFNL4* rs368234815 [TT/ΔG] (entscheidet über die Expression des IFN- λ_4), *IFNL4* rs11764844 [G/A] (führt über die Aminosäuresubstitution (P70S) zu einem funktionsbeeinträchtigten IFN- λ_4), *IFNL3* rs4803217 [C/A] (beeinflusst die Stabilität der *IFNL3*-mRNA) und *IFNL3* rs28416813 [C/G] (beeinflusst die IFN- λ_3 -Expression) (modifiziert nach: Mihm 2015). Die Verwendung der Abbildung erfolgt mit freundlicher Genehmigung von Karger Publishers.

Die naheliegende Vermutung, dass dieser Polymorphismus die Expression des *IFNL3*-Gens oder auch die der anderen Typ-III-IFNe reguliert, konnte trotz intensiver Bemühungen nicht bestätigt werden. Erst mit der Entdeckung des *IFNL4* wurde klar, dass rs12979860 nicht nur *upstream* des *IFNL3* sondern gleichzeitig auch im Intron 1 des *IFNL4*-Gens lokalisiert ist (Abb. 1.8). Seine Bezeichnung änderte sich von *IFNL3* rs12979860 zu *IFNL3/4* rs12979860.

Eine Analyse der Allelhäufigkeiten bezüglich der Ansprechrate (*sustained virological response*, SVR) auf eine Medikation mit IFN- α bei chronischer Hepatitis C in verschiedenen Bevölkerungsgruppen ergab eine positive Korrelation zwischen der Häufigkeit des *IFNL3/4*-rs12979860-Allels C und der Ausheilung einer chronischen Hepatitis C unter IFN- α -Therapie (Ge et al. 2009). Diese Beobachtung erklärt auch die Ausheilungs- und Chronifizierungstendenzen bei HCV-Infektionen innerhalb verschiedener Ethnizitäten.

Im selben Jahr veröffentlichten Thomas et al. (2009) ihre Ergebnisse zu der Assoziation der *IFNL3/4*-rs12979860-Genotypen mit der spontanen Viruselimination des HCV, wobei auch ein positiver Zusammenhang zwischen dem Auftreten des *IFNL3/4*-rs12979860-Allels C und der spontanen Ausheilung ermittelt wurde (Abb. 1.9).

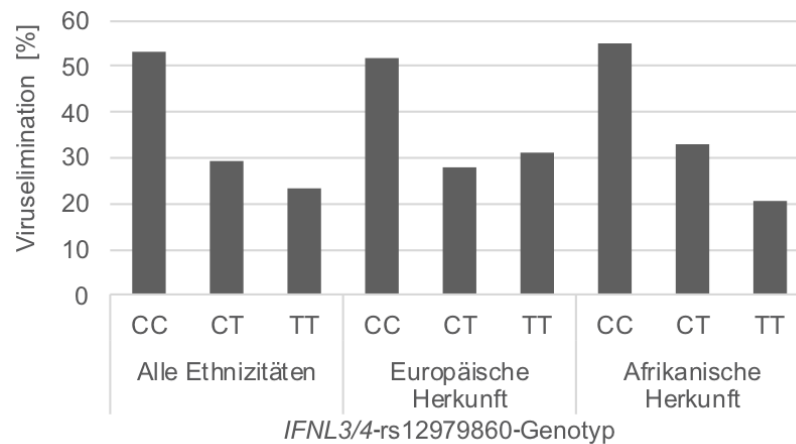


Abb. 1.9: Zusammenhang zwischen den *IFNL3/4*-rs12979860-Genotypen und der spontanen Viruselimination bei Hepatitis C.

Die Daten für diese Abbildung wurden der Publikation von Thomas et al. (2009) entnommen, in der die Genotypisierung von 1006 Patienten, unterteilt nach europäischer und afrikanischer Herkunft, aus sechs unabhängigen HCV-Kollektiven vorgenommen wurde. Hierbei kam es bei 388 Patienten zur spontanen Viruselimination. Bei 620 Patienten persistierte die Infektion.

Die Allelvarianten des Einzelnukleotidpolymorphismus (*single nucleotide polymorphism*, SNP) *IFNL3/4* rs12979860 stehen im Kopplungsungleichgewicht (*Linkage Disequilibrium*, LD) mit den Allelvarianten eines Dinukleotidpolymorphismus im *IFNL4*-Gen, *IFNL4* rs368234815 [TT/ Δ G] (Prokunina-Olsson et al. 2013). Ursprünglich als Dinukleotidpolymorphismus bezeichnet, wird er inzwischen in einigen Datenbanken als zwei aufeinanderfolgende getrennte Polymorphismen mit der Nomenklatur von zwei SNPs geführt (*IFNL4* rs11322783 und *IFNL4* rs74597329) (Prokunina-Olsson 2019).

IFNL4 rs368234815, der innerhalb des Exon 1 des *IFNL4*-Gens liegt, entscheidet über die Expression des IFN- λ_4 -Proteins (Abb. 1.8). Beim Vorliegen der Allelvariante Δ G wird ein funktionstüchtiges IFN- λ_4 codiert (Prokunina-Olsson et al. 2013). Das Allel TT bedingt eine Verschiebung des offenen Leserasters und damit den vorzeitigen Abbruch der Proteinsynthese (Pseudogen) (Prokunina-Olsson et al. 2013) (Tab. 1.1).

Es konnte gezeigt werden, dass die Transkriptmenge des *IFNL4*-Gens unabhängig vom Genotyp ist, jedoch funktionstüchtige IFN- λ_4 -Proteine nur bei Trägern des Allels Δ G synthetisiert werden (Onabajo et al. 2015; Obajemu et al. 2017). Auf Grundlage dessen wurde für diese Arbeit zwischen Trägern des Allels Δ G, die für ein funktionstüchtiges IFN- λ_4 codieren und Allel-TT-Homozygoten, bei denen eine Leserasterverschiebung zu einem funktionalen *knock-out* des IFN- λ_4 führt, unterschieden.

Mihm et al. (2004) zeigten noch vor der Entdeckung des *IFNL4*-Gens, dass bei Patienten mit chronischer Hepatitis C Transkripte der Typ-I- und Typ-III-IFNe zwar in mononukleären Zellen des peripheren Blutes (*peripheral blood mononuclear cell*, PBMC), nicht aber in Lebergewebe nachweisbar waren. Dieses Ergebnis war kompatibel mit dem Nachweis der Aktivität einer HCV-Protease, die in der Lage ist, die Phosphorylierung von Transkriptionsfaktoren und damit die Expression von IFNen zu unterdrücken. Das unerwartete Ergebnis dieser Studie war, dass trotz nicht nachweisbarer Typ-I- und Typ-III-IFN-Expression in dem Lebergewebe die IFN-Effektoren aktiviert waren (Mihm et al. 2004).

Mit der Entdeckung und Beschreibung des *IFNL4* im Jahr 2013 bot sich als Erklärung an, dass die HCV-Protease zwar die Expression von Typ-I- und Typ-III-IFNen in den Hepatozyten unterdrückt, jedoch nicht die des *IFNL4*, das wahrscheinlich für die Aktivierung der ISGe verantwortlich ist (Prokunina-Olsson et al. 2013; Heim 2013; O'Brien et al. 2014).

Paradoxerweise ist die Abwesenheit des IFN- λ_4 bei einer HCV-Infektion prognostisch günstiger, obwohl *in vitro*-Studien ähnliche antivirale Wirkungen des IFN- λ_4 aufzeigen konnten wie für die anderen Typ-III-IFNe (O'Brien et al. 2014; Obajemu et al. 2017). Es konnte gezeigt werden, dass das IFN- λ_4 spezifisch durch die HCV-Infektion aktiviert wird, jedoch reicht die Aktivität nur zur leichten Hemmung des HCV aus und nicht zur kompletten Elimination (Heim 2013; O'Brien et al. 2014). Diese insuffiziente Aktivierung von IFNen und die insuffiziente antivirale Wirkung des IFN- λ_4 könnten der Grund sein, warum eine HCV-Infektion in vielen Fällen zu einer Chronifizierung führt.

Durch das schneller agierende IFN- λ_4 werden ISGe früher hochreguliert, wodurch auch die negativen Rückkopplungsmechanismen früher aktiviert werden als bei anderen IFNen (Obajemu et al. 2017). Diese negativen Rückkopplungsmechanismen hemmen vor allem den IFN- α -Signalweg und machen den Körper refraktär für weitere IFN-Stimulationen. Daraus ergibt sich eine mögliche Erklärung für das schlechte Ansprechen der Patienten mit chronischer Hepatitis C, die gleichzeitig Träger des *IFNL4*-rs368234815-Allels ΔG sind, auf eine exogene IFN- α -Therapie (Obajemu et al. 2017).

Diese Erkenntnisse gehen einher mit der Detektion des Zusammenhangs zwischen der schlechten Prognose für Patienten mit HCV und der Ausprägung des *IFNL4*-rs368234815-Allels ΔG (Prokunina-Olsson et al. 2013). Dieses Allel bewirkt die Synthese eines funktionstüchtigen IFN- λ_4 , was zu einer Desensibilisierung gegenüber IFN- α führt und die Wahrscheinlichkeit der HCV-Chronifizierung erhöht.

Des Weiteren ist im Exon 2 des *IFNL4*-Gens der SNP *IFNL4* rs11764844 lokalisiert, bei welchem das Vorliegen des *minor* Allels A zur einer Aminosäuresubstitution an Position

70 führt (P70S), woraus die Expression eines funktionsbeeinträchtigten IFN- λ_4 -Proteins resultiert (Abb. 1.8, Tab. 1.1). Die *IFNL4*-rs11764844-Allelvariante A assoziiert mit einer besseren Prognose für Patienten mit Hepatitis C, wodurch vorherige Befunde, die die Assoziation eines nicht funktionsfähigen IFN- λ_4 mit einer besseren Prognose beschreiben, gestützt werden (O'Brien et al. 2015).

Der SNP *IFNL3* rs4803217 liegt in der 3'-untranslatierte Region (UTR) vom *IFNL3* positioniert (Abb. 1.8). Es gibt Hinweise, dass das *minor* Allel A der *IFNL3*-mRNA eine erhöhte Instabilität verleiht und damit den Abbau begünstigt (O'Brien et al. 2015) (Tab. 1.1).

Auch der SNP *IFNL3* rs28416813 beeinflusst die IFN- λ_3 -Expression durch eine veränderte Bindung des Transkriptionsfaktors NF- κ B (*Nuclear Factor 'kappa-light-chain-enhancer' of activated B-Cells*) an spezifische Gensequenzen, wobei beim Vorliegen des *minor* Allels G die Expression des IFN- λ_3 abnimmt (Chinnaswamy et al. 2013) (Tab. 1.1). Im Hinblick auf die Überwindung der HCV-Infektion scheint neben dem Fehlen des IFN- λ_4 , eine erhöhte Stabilität und damit erhöhte Expression des IFN- λ_3 günstiger.

Tab. 1.1: Allelvarianten ausgewählter Polymorphismen im Bereich der *IFNL3*- und *IFNL4*-Gene sowie beschriebene Phänotypen

Polymorphismus	Allele	Major Allel Phänotyp	Minor Allel Phänotyp
<i>IFNL3</i> rs4803217	[C/A]	↑ Stabilität <i>IFNL3</i> -mRNA	↓ Stabilität <i>IFNL3</i> -mRNA
<i>IFNL3</i> rs28416813	[C/G]	↑ IFN- λ_3 -Expression	↓ IFN- λ_3 -Expression
<i>IFNL4</i> rs117648444	[G/A]	Wildtyp IFN- λ_4	beeinträchtigtes IFN- λ_4
<i>IFNL3/4</i> rs12979860	[C/T]	↑ HCV-Ausheilung	↓ HCV-Ausheilung
<i>IFNL4</i> rs368234815	[TT/ Δ G]	funktionsloses IFN- λ_4	funktionales IFN- λ_4

1.4 The Cancer Genome Atlas (TCGA)

Dieses Vorhaben greift auf die US-amerikanisch geförderte TCGA (*The Cancer Genome Atlas*)-Datenbank¹ innerhalb der NCI (*National Cancer Institut*) GDC (*Genomic Data Commons*)-Datensammlung zurück.

Die TCGA-Datenbank wurde in die GDC-Datensammlung aufgenommen. Sie gilt als eine unermessliche Wissensquelle und als eines der wichtigsten Projekte, das das umfassende Verständnis der für jede Tumorentität individuellen Genomalterationen mittels innovativer Genomanalysetechnologien stark beschleunigt (Tomczak et al. 2015).

Das Ziel der TCGA-Datenbank ist es, gesammelte Tumordaten in harmonisierter Form

¹<https://portal.gdc.cancer.gov/>, Zugriff am 11.10.2018

Wissenschaftlern weltweit zur Verfügung zu stellen. Außerdem wurde das nicht-maligne Wirtsgenom, auf welchem in dieser Arbeit das Hauptinteresse liegt, ausgelesen und in die Datenbank aufgenommen. Dadurch soll die Krebsforschung und im Speziellen die Forschung an diagnostischen Methoden, an Behandlungsmethoden und letztendlich an der Heilung gefördert werden. Hierbei soll das Prinzip des *data sharing* als globales Kommunikationsmittel zwischen Wissenschaftlern dienen sowie einen regen globalen Wissensaustausch unterstützen (Grossman et al. 2016).

Die TCGA-Datenbank wurde als sehr umfangreiches, internationales Projekt zwischen dem NCI und dem NHGRI (*National Human Genome Research Institute*) im Jahr 2006 mit einem dreijährigen Pilotprojekt gestartet und endete im Jahr 2013 mit der Probensammlung². Insgesamt wurden molekulargenetische Daten von über 10.000 Primärtumoren und korrespondierenden nicht-malignen Gewebeproben von 33 verschiedenen Tumorentitäten erhoben und mit dazugehörigen demografischen und klinischen Daten zur Verfügung gestellt³. Die molekulargenetischen Daten enthalten genomische Daten, das Transkriptom, miRNA-Daten, Daten zu SNPs, Daten der Kopienzahlvariationen und Daten zu der DNA-Methylierung sowie Aufnahmen mikroskopischer Gewebepreparate. Im TCGA-LIHC (*liver hepatocellular carcinoma*)-Kollektiv sind 377 Patienten mit HCC unterschiedlicher Ethnizität zusammengefasst. Das TCGA-PAAD (*pancreatic adenocarcinoma*)-Kollektiv umfasst 185 Patienten mit PDAC hauptsächlich europäischer Herkunft.

Bei der Projektkonzeption der TCGA-Datenbank war das Zusammentragen demografischer und klinischer Daten und insbesondere der Ereigniszeiten inklusive Nachbeobachtungen nicht das primäre Ziel, weswegen die demografischen und klinischen Daten sowie die Ereigniszeiten für alle Tumorentitäten im Nachhinein kuratiert, validiert und als TCGA-CDR (*Pan-Cancer Clinical Data Resource*)-Datensatz publiziert wurden (Liu et al. 2018). Bei den HCC- und PDAC-Patienten wird ausdrücklich die Nutzung des progressionsfreien Intervalls (*progression free interval*, PFI), des spezifischen progressionsfreien Intervalls (*progression free interval.2*, PFI.2) und der Gesamtüberlebensdauer (*overall survival*, OS) als Ereigniszeiten empfohlen⁴ (Weinstein et al. 2013; Liu et al. 2018).

²<https://www.cancer.gov/about-nci/organization/ccg/research/structural-genomics/tcga/history/timeline>, Zugriff am 25.08.2019

³<https://portal.gdc.cancer.gov/>, Zugriff am 25.08.2019

⁴<https://gdc.cancer.gov/about-data/publications/pancanatlas>, Zugriff am 15.03.2019

1.5 Fragestellung

Diese Arbeit beschäftigt sich mit der Bedeutung genetischer Keimbahnvariationen in Typ-III-IFN-Genen für die Progression der beiden tödlichsten Tumorentitäten HCC und PDAC. Typ-III-IFN-Gene enthalten genetische Polymorphismen, die mit einem überzeugenden antiviralen Phänotyp assoziieren. Eine der Variationen, *IFNL4* rs368234815 [TT/ Δ G], entscheidet über die Fähigkeit des Wirts, IFN- λ_4 zu exprimieren. Während die *IFNL4*-rs368234815-Allelvariante Δ G für ein funktionsfähiges IFN- λ_4 codiert, führt das Vorliegen des Allels TT über eine Verschiebung des Leserasters zu einem funktionalen *knock-out* des *IFNL4*-Gens.

Die Frage nach der klinischen Evidenz für oder gegen eine prognostische Signifikanz dieser *IFNL*-Polymorphismen für die Tumorentitäten HCC und PDAC sollte unter Nutzung der Tumordatenbank TCGA adressiert werden. Die TCGA-Datenbank enthält neben frei zugänglichen demografischen und klinischen Datensätzen auch eingeschränkt zugängliche molekulargenetische und genomische Daten. Für letztere war zunächst für die Verwendung im Rahmen der hier definierten Fragestellung eine Zugangsberechtigung einzuholen (*Data Access Request Approval Process*).

Die TCGA-Datenbank enthält Daten von insgesamt 377 HCC-Patienten (TCGA-LIHC-Kollektiv), davon 187 Patienten europäischer Herkunft, und 185 PDAC-Patienten (TCGA-PAAD-Kollektiv) mit insgesamt 162 Patienten europäischer Herkunft.

Zunächst sollten die frei zugänglichen Daten deskriptiv und explorativ analysiert werden. Im Anschluss sollten mittels Daten des 1000-Genome-Projekts im Zusammenhang mit den Keimbahnvariationen stehende und für den *IFNL*-Genlocus repräsentative polymorphe Positionen (*tag*SNP) definiert werden. Die Genotypen der fünf ausgewählten Keimbahnvariationen sollten aus dem nicht-malignen Wirtsgewebe in Form von nicht-malignem, dem Tumor benachbarten Gewebe (Code: 11A) sowie Proben aus peripheren Blutzellen (Code: 10A) ausgelesen werden. Die dadurch ermittelten Genotypverteilungen der HCC- und PDAC-Patienten sollten mit den Genotypverteilungen eines Referenzgenoms aus dem 1000-Genome-Projekt verglichen werden. Anschließend sollten die Patientengruppen nach Genotypen stratifiziert und hinsichtlich der Tumorprogression analysiert werden. Als Kriterien für die Tumorprogression sollten die Ereigniszeiten PFI, PFI.2 und OS herangezogen werden. Hierbei sollten die Tumorprogressionskriterien durch Ereigniszeitanalysen nach dem Kaplan-Meier-Verfahren und der Cox-Regression ausgewertet werden.

Die Ergebnisse sollen eine Einschätzung der Eignung der Typ-III-IFNe als prognostische Wirtsfaktoren für das HCC und das PDAC ermöglichen.

2 Material und Methoden

Diese Arbeit entspricht einer genetischen Assoziationsstudie, bei der die Progression zweier Tumorentitäten (HCC bzw. PDAC) vor dem Hintergrund genetischer Polymorphismen im Bereich der Typ-III-IFNe betrachtet wird. Als Kriterien für die Progression dienen die Ereigniszeiten PFI, PFI.2 sowie OS. Verglichen werden Patienten mit dem genetischen Hintergrund, der eine Expression des *IFNL4*-Gens zu einem funktionstüchtigen IFN- λ_4 -Protein zulässt, mit denen, die aufgrund einer Leserasterverschiebung kein funktionstüchtiges IFN- λ_4 exprimieren können (Fall-Kontroll-Ansatz).

Zur Analyse standen demografische, klinische und genomische Daten der zwei Patientenkollektive TCGA-LIHC und TCGA-PAAD sowie kuratierte Ereigniszeiten des PFI, des PFI.2 und der OS zur Verfügung.

Das verwendete Analysekonzept (Anhang Tab. A.1) und alle in dieser Arbeit verwendeten Datenquellen (Anhang Tab. A.2) wurden tabellarisch zusammengefasst.

2.1 *The Cancer Genome Atlas (TCGA)*

Die Analysen und das Erheben der Ergebnisse in dieser Arbeit wurden durch Daten der durch NCI und NHGRI (*National Human Genome Research Institute*) geführten, renommierten TCGA-Datenbank (phs000178.v10.p8) ermöglicht (Informationen zu der TCGA-Datenbank ¹).

2.1.1 **Frei zugängliche Datensätze**

Für die vorliegende Arbeit wurden die kuratierten, validierten und publizierten demografischen und klinischen Daten der HCC- und PDAC-Patienten dem TCGA-CDR-Datensatz entnommen. Alle Daten wurden gesichtet und die für diese Arbeit relevanten Kriterien ausgelesen und deskriptiv zusammengefasst. Die ausgewählten Kriterien setzen sich aus dem Patientenalter zum Zeitpunkt der Diagnose, dem Geschlecht, der Ethnizität, dem Differenzierungsgrad des Tumors und dem Stadium der Tumorerkrankung jeweils zum Zeitpunkt der Diagnose zusammen. Das Stadium der Tumorerkrankung ist innerhalb der TCGA-Datenbank nach den Kriterien des AJCC (*American Joint Committee on Cancer*) eingeteilt. Für das TCGA-LIHC-Kollektiv wurden außerdem auf der Basis von serologischen und anamnestischen Daten mutmaßliche Risikofaktoren (Ätiologie) kategorisiert. Zu den Risikofaktoren zählen Alkoholkonsum, Ausprägung einer nicht-alkoholinduzierten Steatose oder Steatohepatitis, hereditäre Hämochromatose, Le-

¹<http://cancergenome.nih.gov>, Zugriff am 27.10.2018

bertoxinexposition, α 1-Antitrypsinmangel, Autoimmunhepatitis sowie HBV- und HCV-Infektionen. Die HCC-Patienten wurden in eine Gruppe ohne jegliche Risikofaktoren bei nicht vorhandener Hepatitisserologie („Kein Risiko“), in eine Gruppe mit nicht-viral bedingten Lebererkrankungen bei nicht vorhandener Serologie („Kein virales Risiko“) sowie in jeweils eine Gruppe mit serologischem Nachweis einer HBV- oder HCV-Infektion („HBV“ bzw. „HCV“) und eine Gruppe mit serologischem Nachweis einer HBV-/HCV-Koinfektion („HBV-/HCV-Koinfektion“) unterteilt.

Zusätzlich enthalten die Daten verschiedene Ereigniszeiten, von denen das PFI und das PFI.2 sowie die OS für Analysen in dieser Arbeit genutzt wurden. Das PFI wird als Zeitspanne von der Diagnose bis zur Progression der Tumorerkrankung, bis zum Auftreten eines weiteren Tumors jeglicher Entität oder bis zum Tod durch einen Tumor jeglicher Entität definiert. Hierbei wird die Progression des Ausgangstumors, ein lokoregionales Tumorrezidiv, Metastasierung, ein neuer Primärtumor und Tod durch einen Tumor jeweils jeglicher Tumorentität als Ereignis gewertet und in dieser Arbeit als PFI-Ereignis zusammengefasst (Liu et al. 2018). Das PFI.2 wird als Zeitspanne von der Diagnose bis zur Progression der ursprünglichen Tumorerkrankung oder bis zum Tod durch selbige Tumorentität definiert, wobei hier im Unterschied zu dem PFI nur die Ausgangstumorentität betrachtet wird (in dieser Arbeit HCC bzw. PDAC). Als Ereignis wird die Progression des Ausgangstumors, ein lokoregionales Tumorrezidiv, Metastasierung des Ausgangstumors, ein neuer Primärtumor der gleichen Entität wie der Ausgangstumor und Tod durch die Ausgangstumorentität als Ereignis gewertet und in dieser Arbeit als PFI.2-Ereignis zusammengefasst. Zusätzlich wurde jeder neue Tumor, dessen Entität unbekannt war, im überarbeiteten Datensatz des TCGA-CDR bei der Ereigniszeit PFI.2 ausgeschlossen (Liu et al. 2018). Die OS ist definiert als die Zeitspanne zwischen der Diagnose und dem Tod durch jegliche Ursache (Liu et al. 2018).

2.1.2 Zugangsberechtigung

Die Sequenzdateien der HCC- und PDAC-Patienten fallen in die Kategorie Datensätze mit kontrolliertem Zugang, für deren Nutzung eine Genehmigung eingeholt werden muss (*Controlled Data Access Request*) (Übersicht über das Prozedere²). Zunächst ist die Erstellung eines eRA (*electronic Research Administration*) *Commons Accounts* über die Institution erforderlich, deren eigene eRA-Registrierung Voraussetzung ist. Die Universitätsmedizin Göttingen (UMG) verfügt über eine institutionelle eRA-Registrierung, sodass über den *Signing Official* für den *Principal Investigator* ein eRA-Account eingerichtet wurde.

Der eRA-Account ermöglicht den Zugang zum Datenportal dbGaP (*database of Ge-*

²<https://gdc.cancer.gov/access-data/obtaining-access-controlled-data>, Zugriff am 29.10.2018

notypes and Phenotypes) und das Einreichen eines *Controlled Data Access Request*, in dem das Vorhaben beschrieben wird (*Research Use Statement*). Zusätzlich müssen sich Antragssteller und Institution verpflichten, die Datenschutzrichtlinien gemäß *Genomic Data Use Code of Conduct* zu wahren (*Data Use Certificate Agreement*).

Das Projekt wurde über die Universitätsmedizin Göttingen (UMG) über das dbGaP (*database of Genotypes and Phenotypes*)-Datenportal NIH (*National Institutes of Health*) angemeldet und genehmigt (*project ID* 20041). Der Zugang zu allen geschützten Daten des TCGA wurde gewährt.

Die Datennutzung genügt US-amerikanischen und deutschen Datenschutzrichtlinien.

Für genehmigte Projekte gewähren die NIH (*National Institutes of Health*) Zugriff auf die Datensätze mit kontrolliertem Zugang für einen Zeitraum von einem Jahr.

2.1.3 Datensätze mit kontrolliertem Zugang

Proben des primären Tumorgewebes (Code: 01A), des nicht-malignen, dem Tumor benachbarten Gewebes (Code: 11A) sowie Proben aus peripheren Blutleukozyten (Code: 10A) stellen das Gewebe zur Ermittlung der WXS-DNA-Sequenzierdaten. Das Gewebe für die Erstellung der RNA-Sequenzierdaten entstammt entweder Proben des primären Tumorgewebes (Code: 01A) oder Proben des nicht-malignen, dem Tumor benachbarten Gewebes (Code: 11A).

Detaillierte Informationen über die TCGA-Datenstruktur, DNA- und RNA-Sequenzanalysen, die durchgeführten Experimente mit Protokollen sowie über die benutzten Software werden in dem TCGA-Datenportal zur Verfügung gestellt³. Darüber hinausgehende Informationen finden sich in wissenschaftlichen Artikeln zu den einzelnen Tumorentitäten (für HCC: Raphael et al. 2017; für PDAC: Ally et al. 2017), wobei die DNA und RNA sowohl der HCC- als auch der PDAC-Tumorentität mittels des Illumina HiSeq 2000 sequenziert wurden (Raphael et al. 2017; Ally et al. 2017).

2.1.4 Auslesen der polymorphen Positionen

Aufgrund der Größe von bis zu 80 GB der im BAM-Format vorliegenden Dateien wurde das *BAM slicing tool* genutzt, um die Dateien auf den Lokus von Interesse, in dieser Arbeit der *IFNL*-Lokus auf Chromosom 19 (Chr19: 39.230.000–39.300.000), zu begrenzen. Alle in TCGA verfügbaren Sequenzdateien, pro Patienten durchschnittlich vier, wurden auf diese Art zugeschnitten und auf ein Gruppenlaufwerk der Universitätsmedizin Göttingen (UMG) heruntergeladen. Insgesamt wurden auf diese Weise etwa 2.300 Dateien generiert.

Zur Darstellung und zum Auslesen von genomischen Sequenzen wurde die Software NCBI

³<https://tcga-data.nci.nih.gov/tcga/>, Zugriff am 29.10.2018

(*National Center for Biotechnology Information*) *Genome Workbench*⁴ genutzt, deren Funktionen und Anwendung in verschiedenen Tutorials und Artikeln beschrieben werden (Zhang et al. 2007).

Die Bearbeitung der Dateien innerhalb der NCBI *Genome Workbench* verlangt sogenannte *indexfiles*, die für jede der zugeschnittenen BAM-Dateien mittels IGV (*Integrative Genomics Viewer*) angelegt wurden (Robinson et al. 2017). Zum Auslesen polymorpher Positionen im *IFNL*-Lokus erfolgte die Ausrichtung der Sequenzdaten (*alignment*) an Chromosom 19 eines Referenzgenoms des Menschen (*Human Genome Reference Assembly*, GRCh38.p12). Von jedem Patienten wurde mindestens eine Sequenzdatendatei in die NCBI *Genome Workbench* eingelesen. Sofern vorhanden, wurde das Material der peripheren Blutleukozyten (Code: 10A) gewählt. Ansonsten wurde nicht-malignes Gewebe (Code: 11A) verwendet. Zu Kontrollzwecken wurden von zufällig ausgewählten Patienten Daten aller verfügbaren Materialien eingelesen. Pro eingelesener Sequenzdatendatei wurde an fünf polymorphen Positionen innerhalb der *IFNL3*- und *IFNL4*-Gene die Genotypen ausgelesen (Tab. 2.1).

Tab. 2.1: Übersicht über die in NCBI *Genome Workbench* markierten und ausgelesenen *IFNL*-Polymorphismen auf Chromosom 19

Marker	Position auf Chromosom 19	Polymorphismen	Allele
M1	39.243.580	<i>IFNL3</i> rs4803217	[C/A]
M2	39.245.004	<i>IFNL3</i> rs28416813	[C/G]
M3	39.247.938	<i>IFNL4</i> rs117648444	[G/A]
M4	39.248.147	<i>IFNL3/4</i> rs12979860	[C/T]
M5	39.248.514	<i>IFNL4</i> rs368234815	[TT/ΔG]

Zur Bestimmung der Genotypen der polymorphen Positionen wurden die Kriterien der biotechnischen Firma Illumina angewendet⁵. Es wurde davon ausgegangen, dass die Genotypen ab einer Sequenziertiefe von mindestens 20 (besser 30) *reads* belastbar sind, wobei von einer Heterozygotie ausgegangen wird, wenn beide Allelvarianten eine Sequenziertiefe von ≥ 10 *reads* haben und das Verhältnis ihrer *reads* ≤ 3 beträgt⁵.

Bei Nichterreichen der Mindestanzahl an *reads* oder Uneindeutigkeit wurden die verfügbaren Daten auf Konsistenz mit den anderen, im LD befindlichen Genompositionen überprüft (siehe Kapitel 3.2 Abb. 3.5).

⁴<https://www.ncbi.nlm.nih.gov/tools/gbench/>, Zugriff am 20.01.2019

⁵https://www.illumina.com/Documents/products/technotes/technote_snpcaller_sequencing.pdf, Zugriff am 08.04.2019

2.2 Das 1000-Genome-Projekt

Für diese Arbeit wurde ein gesundes, nicht-tumordiagnostiziertes Kontrollkollektiv benötigt, welches dem 1000-Genome-Projekt (Phase 3, Version 5) entnommen wurde.

Das internationale 1000-Genome-Projekt⁶ hat sich zum Ziel gesetzt, sämtliche Varianten von Genen, die mit einer Häufigkeit von $\geq 1\%$ im Genom der gesamten Menschheit vorkommen, in einem Katalog zu erfassen (Via et al. 2010). Dieses Projekt startete mit einem Pilotprojekt zur Methodenanalyse im Jahr 2008 und endete mit der dritten Projekt- und Analysephase im Jahr 2015 nach Sequenzierung des Genoms von 2.504 Probanden aus 26 verschiedenen Populationen, wobei 84,4 Millionen genetische Variationen erfasst wurden (Altshuler et al. 2010; Altshuler et al. 2012; Auton et al. 2015).

Die Daten des 1000-Genome-Projekts werden unter anderem von dem *Ensembl Genome Browser* zur öffentlichen Verwendung zur Verfügung gestellt⁷ (Zheng-Bradley und Flicek 2016).

Als Kontrollkollektiv für die TCGA-Kollektive europäischer Herkunft diente die im 1000-Genome-Projekt enthaltene europäische Population (EUR), zu deren Subpopulationen die Einwohner aus Utah mit nord- und westeuropäischer Abstammung (CEU), die Finnen in Finnland (FIN), die Briten in England und Schottland (GBR), die iberische Population in Spanien (IBS) und die toskanische Population in Italien (TSI) gezählt werden⁸.

2.3 Statistische Verfahren

Alle in dieser Arbeit verwendeten statistischen Analyseprogramme wurden im Anhang Tab. A.2 zusammengefasst.

Die statistischen Analysen sowie die einfachen und multiplen Cox-Regressionsmodelle und auch die Überlebensanalysen/Ereigniszeitanalysen nach dem Verfahren von Kaplan und Meier wurden in der Programmiersprache R mithilfe des Programms RStudio (Version 3.5.2)⁹ ausgeführt. Das Programm RStudio wird kontinuierlich aktualisiert und es werden regelmäßig überarbeitete und neue Funktionspakete zur Verfügung gestellt. Tutorials halfen bei der korrekten Anwendung von R, hier vor allem STHDA (*Statistical tools for high-throughput data analysis*)¹⁰.

⁶<https://www.internationalgenome.org/>, Zugriff am 22.01.2019

⁷http://grch37.ensembl.org/Homo_sapiens, Zugriff am 12.04.2019

⁸<https://www.internationalgenome.org/category/population/>, Zugriff am 09.07.2019

⁹<http://www.rstudio.com/>, Zugriff am 01.02.2019

¹⁰<http://www.sthda.com/english/>, Zugriff am 28.05.2019

2.3.1 Deskriptive und explorative Statistik

Die deskriptive und explorative Statistik soll die Daten, die sich aus dem demografischen und klinischen, durch die ausgelesenen Wirtstgenotypen ergänzten Datensatz zusammensetzen, zusammenfassen und beschreiben. Dies dient dazu, einen ersten Überblick über die Spannbreite der Werteverteilung und ein Verständnis für eventuelle Zusammenhänge von den Daten zu erhalten.

Die ausgewählten Kriterien dieser Arbeit weisen verschiedene Skalenniveaus und Merkmalstypen auf, die unterschiedliche statistische Anwendungen erfordern. Dabei können die Daten in die Nominal-, Ordinal-, Intervall- und Verhältnisskala eingeteilt werden (Gaus und Muche 2017, S. 82ff.). Zu den nominalen Skalenniveaus zählen verbal beschreibbare Kriterien, deren Ausprägungen keiner Reihenfolge unterliegen, wie in dieser Arbeit Geschlecht, Ethnizität, Ätiologie und Genotyp. Die Kriterien Differenzierungsgrad und Stadium der Tumorerkrankung sind auch verbal beschreibbar, jedoch ist bei ihren Ausprägungen eine Reihenfolge definiert, wodurch sie zu den Ordinalskalen zählen. Bei diesen beiden ersten Skalenniveaus sind die Abstände zwischen den einzelnen Ausprägungen nicht definiert und nicht von Interesse, wohingegen bei der Intervallskala die Ausprägungen in Zahlen angegeben werden, die eine definierte Reihenfolge und definierte Abstände aufweisen (Gaus und Muche 2017, S. 82ff.). Liegt der Intervallskala ein Nullpunkt vor, so wird sie auch Verhältnisskala genannt (Gaus und Muche 2017, S. 82ff.). Die verhältnisskalierten Kriterien dieser Arbeit sind das Alter bei Diagnose in Jahren, das PFI, das PFI.2 sowie die OS jeweils in Tagen.

Neben den Skalenniveaus können die klinischen Kriterien auch in verschiedene Merkmalstypen unterteilt werden, wobei häufig die erste Unterteilung in qualitative (oder kategoriale) und quantitative (oder metrische) Merkmale erfolgt. Weiter werden bei den qualitativen Merkmalen nominale (Geschlecht, Ethnizität, Ätiologie, Genotyp) und ordinal/geordnete (Differenzierungsgrad, Tumorstadium) Merkmale unterschieden, wobei das nominale Merkmal Geschlecht durch das Vorliegen von nur zwei Ausprägungen auch als dichotomes oder Alternativmerkmal bezeichnet wird (Gaus und Muche 2017, S. 82ff.). Die quantitativen Merkmale sind numerischer Art und werden in die stufenlosen stetigen Merkmale (Alter bei Diagnose, PFI, PFI.2, OS) zusammengefasst (Gaus und Muche 2017, S. 82ff.).

Zur Deskription der Patientenkollektive wurden Tabellen mit den ausgewählten Kriterien und deren Verteilungen unter den Patienten als absolute und relative Häufigkeiten erstellt. Diese Verteilungen in die Häufigkeiten können sowohl für quantitative als auch qualitative Merkmale betrachtet werden und wurden zur Veranschaulichung auch grafisch in Form von Histogrammen, Boxplots und Säulendiagrammen dargestellt (An-

hang Kapitel 6.2). Zur weiteren Beschreibung des quantitativen Merkmals Alter wurde als Maß der zentralen Tendenz der arithmetische Mittelwert (\bar{x}) berechnet. Hierbei ist darauf zu achten, dass der arithmetische Mittelwert nur für quantitative, stetige Merkmale, die normalverteilt sind, sinnvoll und aussagekräftig berechenbar ist (Gaus und Muche 2017, S. 126ff.). Außerdem wird der arithmetische Mittelwert stark von extremen Werten (Ausreißern), die teilweise richtig, aber auch falsch entstanden sein können, beeinflusst.

Zur Betrachtung der Streuung der Werte um den arithmetischen Mittelwert wurde die Standardabweichung (s) berechnet, welche die Wurzel aus der Varianz (s^2) ist. Die Standardabweichung berechnet sich wie folgt (Gaus und Muche 2017, S. 126ff.)

$$s = \sqrt{s^2} = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2} \text{ mit } \bar{x} = \frac{1}{n} \sum_{i=1}^n x_i . \quad (2.1)$$

Hierbei sei n die Patientenzahl, x_i das jeweilige Alter der Patienten und \bar{x} der arithmetische Mittelwert des Alters.

Für Daten, die eine kleine Streuung aufweisen, wird eine kleine Standardabweichung errechnet und umgekehrt, wobei Ausreißer auch die Standardabweichung stark beeinflussen. Am aussagekräftigsten wird die Standardabweichung aus normalverteilten, verhältnisskalierten Werten berechnet (Gaus und Muche 2017, S. 126ff.).

Um die Robustheit der zentralen Tendenz und der Abweichung sicherzustellen, wurden zusätzlich der Median als zentrale Tendenz und der Quartilsabstand als Abweichung errechnet, da beide Schätzfunktionen von der Bedingung des Vorliegens einer Normalverteilung unabhängig sind und nicht so stark von Ausreißern beeinflusst werden. Jedoch müssen mindestens ordinalskalierte Daten mit nur einem Maximum der Berechnung zugrunde liegen. Der Median ist der mittlere Wert von Daten, die der Größe nach sortiert wurden, sodass jeweils 50 % der Daten über und unter dem Median liegen. Liegt eine gerade Datenanzahl vor, so wird der arithmetische Mittelwert der beiden mittleren Daten errechnet und als Median angegeben (Gaus und Muche 2017, S. 126ff.).

Beim Quartilsabstand wird die Differenz zwischen dem dritten und ersten Quartil berechnet, wodurch der Bereich mit den mittleren 50 % der Daten angegeben wird. Beim Vergleich des Mittelwerts und der Standardabweichung mit dem Median und dem Quartilsabstand kann der Stichprobenfehler der Schätzfunktionen bewertet werden. Außerdem entspricht der Vergleich einem Kriterium zur Einschätzung des Vorliegens einer Normalverteilung (Gaus und Muche 2017, S. 210ff.).

Normalverteilte Werte sind stetige Werte, die unter anderem nur ein Maximum sowie eine symmetrische Verteilung aufweisen, wobei der arithmetische Mittelwert das Maxi-

mum der Verteilungskurve, auch Gaußsche Kurve genannt, ist. Außerdem dürfen bei den Werten keine großen Ausreißer vorliegen und die Fallzahl sollte $n \geq 70$ sein.

Beim Vergleich des Mittelwerts mit dem Median sollten die Werte etwa gleich groß sein, sonst ist von einer schiefen, nicht normalverteilten Kurve auszugehen (Gaus und Muche 2017, S. 210ff.). In dieser Arbeit wurde das Kriterium Alter auf Normalverteilung untersucht, wobei die grafische Darstellung der Altersverteilungen durch Histogramme die Beurteilung unterstützt (Anhang Abb. A.1). Alleine das Alter der PDAC-Patienten war normalverteilt (Anhang Abb. A.1 (B)).

Die Teststärke parametrischer Tests, die bei normalverteilten Daten angewandt werden können, ist mit der Teststärke von nicht-parametrischen Tests, die keine Normalverteilung der Daten voraussetzen, vergleichbar. Außerdem muss bei den parametrischen Tests noch auf weitere Annahmen getestet werden, sodass in dieser Arbeit auch bei dem normalverteilten Kriterium Alter der PDAC-Patienten mit nicht-parametrischen Tests gearbeitet wurde (van Hecke 2010).

Als Kennzahl für die Allelhäufigkeiten wird die Häufigkeit des selteneren Allels (*minor allele frequency*, MAF) berechnet. Die MAF berechnet sich aus dem Quotienten aus der Auftretenshäufigkeit des *minor* Allels (d. h. die Summe der *minor* Allele (aa, Aa)) geteilt durch die Gesamtauftrittshäufigkeit aller Allele (d. h. die Summe der *major* und *minor* Allele (AA, Aa, aa))

$$MAF = \frac{2aa + Aa}{(2aa + Aa) + (2AA + Aa)} . \quad (2.2)$$

2.3.2 Induktive Statistik

Das statistische Testen dient zur Überprüfung von Vermutungen (Hypothesen) und liefert ein Maß für die statistische Plausibilität beobachteter Unterschiede. Hierbei werden vor dem Testen eine Nullhypothese (H_0), die die Gleichheit der beobachteten Werte postuliert, und eine Alternativhypothese (H_1) festgelegt, die das Gegenteil besagt und von einem nicht zufälligen Unterschied der beobachteten Werte ausgeht (Gaus und Muche 2017, S. 264ff.). Außerdem wird zwischen einem ein- oder zweiseitigen Testverfahren unterschieden, wobei die Tendenzrichtung der analysierten Werte bei dem einseitigen Test bereits bekannt ist und bei dem zweiseitigen nicht (Gaus und Muche 2017, S. 264ff.). In dieser Arbeit wurden ausschließlich zweiseitige Tests durchgeführt.

Zusätzlich wird vor dem Testen das Signifikanzniveau α festgelegt, welches die Grenzwahrscheinlichkeit bildet, ab der nicht mehr von einem zufälligen Unterschied ausgegan-

gen wird und somit die Entscheidungsgrenze für die später berechneten Überschreitungswahrscheinlichkeiten (p-Werte) darstellt. In dieser Arbeit wurde ein $\alpha \leq 0,05$ festgelegt, ab dem die zufällige Entstehung der beobachteten Unterschiede unwahrscheinlich ist (H_0 abgelehnt) und von einem nicht zufälligen, signifikanten Unterschied ausgegangen wird (H_1 angenommen). Bei Werten von $\alpha > 0,05$ reicht die Evidenz nicht aus, um H_0 zu verwerfen (Bickeböller und Fischer 2007, S. 56ff.; Gaus und Muche 2017, S. 264ff.).

Gleichzeitig bildet das Signifikanzniveau α die maximale Grenze für das α -Risiko. Der α -Fehler, auch als Fehler erster Art bezeichnet, besagt, dass H_0 fälschlicherweise abgelehnt wird, obwohl die Hypothese richtig ist. Mit dem Wert des α -Risikos wird festgelegt, mit welcher maximalen Wahrscheinlichkeit dieser Fehler auftritt. Der β -Fehler, oder Fehler zweiter Art, wird definiert als die fälschliche Annahme von H_0 , obwohl ein nicht zufälliger Unterschied vorliegt. Das Komplement ($1 - \beta$) entspricht der Teststärke des Tests und gibt die Wahrscheinlichkeit an, einen wirklichen, nicht rein zufälligen Unterschied zu erkennen. Mit Festlegung des $\alpha \leq 0,05$ ist das α -Risiko höher als bei einem $\alpha \leq 0,01$, jedoch steigt im Gegenzug die Teststärke des Tests ($1 - \beta$) an (Bickeböller und Fischer 2007, S. 56ff.; Gaus und Muche 2017, S. 264ff.).

Bei der Durchführung mehrerer Hypothesentests sollte beachtet werden, dass das Signifikanzniveau α vor dem Testen mithilfe verschiedener Regeln auf das multiple Testen adjustiert wird. Diese Adjustierungsverfahren wurden jedoch im Rahmen dieser Arbeit nicht angewandt, da der betrachtete Umfang an Hypothesentests dies nicht erforderte. Mit der Überschreitungswahrscheinlichkeit (p-Wert) wird die Wahrscheinlichkeit berechnet, „mit der die tatsächlich gewonnen Daten – oder noch extremere, noch unwahrscheinlichere Daten – rein zufällig unter der Nullhypothese entstanden sein können“ (Gaus und Muche 2017, S. 266). Je kleiner der p-Wert ist und je näher er an Null liegt, desto wahrscheinlicher sind die beobachteten Unterschiede nicht per Zufall entstanden und die Annahme von H_0 wird unwahrscheinlich (Gaus und Muche 2017, S. 264ff.).

Die ausschließliche Betrachtung des p-Werts enthält keine Aussage über die Richtungstendenz der getesteten Daten. Deswegen werden zusätzlich tendenzweisende Werte betrachtet, wie zum Beispiel in dieser Arbeit die medianen progressionsfreien Intervalle, PFI bzw. PFI.2, und die medianen Überlebenszeiten für die nach den *IFNL*-Genotypen stratifizierten TCGA-Kollektive. Außerdem bietet auch die Angabe der Konfidenzintervalle (KI) eine Möglichkeit, die Genauigkeit der errechneten Werte und die Richtung des Einflusses einzuschätzen und Rückschlüsse auf die Signifikanz zu ziehen. Das KI stärkt die statistische Plausibilität und erhärtet die klinische Relevanz der errechneten Werte. Dabei gibt es einen Vertrauensbereich an, in dem die wahren, nicht mittels Stichproben ermittelbaren Werte mit einer gewissen Wahrscheinlichkeit liegen. Diese Wahrscheinlichkeit wird im Vorhinein festgelegt, in dieser Arbeit auf 95 %. Ein KI von 95 % besagt,

dass mit 95 %iger Wahrscheinlichkeit die wahren Werte in dem berechneten Vertrauensbereich liegen. Hierbei wird zur Darstellung des Intervalls von einer Normalverteilung ausgegangen und der Mittelwert \pm des Standardfehlers (SE) berechnet (Gaus und Muche 2017, S. 247ff.). Bei einem 95 % KI gilt

$$95 \% KI = \bar{x} \pm 1,96 SE . \quad (2.3)$$

Umgekehrt gilt auch die Annahme der Irrtumswahrscheinlichkeit, die bei einem 95 % KI besagt, dass zu 5 % der wahre Wert nicht im KI liegt (Gaus und Muche 2017, S. 247ff.).

Die Wahl der Hypothesentests hängt von verschiedenen Faktoren ab, unter anderem von der Verteilung der Werte, der Art des vorliegenden Merkmaltyps und der Anzahl der betrachteten Gruppen sowie von der jeweiligen Fragestellung. Zur statistischen Auswertung der qualitativen Kriterien Ätiologie, Tumordifferenzierungsgrad, Tumorstadium und Genotyp mit ≥ 2 Parallelgruppen wurde der vielseitige χ^2 -Unabhängigkeitstest angewandt, der im RStudio als „chisq.test“ zu finden ist (Gaus und Muche 2017, S. 281ff.). Hierbei wird auf die Unabhängigkeit zwischen den beobachteten Kriterien getestet. H_0 lautet in diesem Fall, dass die Kriterien voneinander unabhängig sind. Bei H_1 wird hingegen eine Abhängigkeit erwartet. Die beobachteten Werte der absoluten Häufigkeit, die sich in eine Kontingenztafel schreiben lassen, werden mit erwarteten, ausgerechneten Häufigkeiten, die H_0 unterliegen, verglichen. Die erwartete Häufigkeit (b_i) wird durch das Produkt der betrachteten Randsummen dividiert durch die Gesamtsumme innerhalb einer Kontingenztafel berechnet. Anschließend werden die beobachteten und erwarteten Häufigkeiten verglichen. Hierzu wird der Quotient aus der quadrierten Differenz zwischen der einzelnen beobachteten (a_i) und der erwarteten (b_i) Häufigkeit und der erwarteten Häufigkeit gebildet und anschließend werden die einzelnen Quotienten, die sogenannten χ^2 -Anteile, aufsummiert (Gaus und Muche 2017, S. 281ff.)

$$\chi^2 = \sum_{i=1}^n \frac{(a_i - b_i)^2}{b_i} \quad \text{mit} \quad b_i = \frac{\text{Randsummenprodukt}}{\text{Gesamtsumme}} . \quad (2.4)$$

Hierbei sei n die Anzahl der χ^2 -Anteile. Die Freiheitsgrade (Fg) berechnen sich aus dem Produkt der jeweiligen um eins verkleinerten Anzahl (N) der verschiedenen Ausprägungen der Kriterien

$$Fg = ((N(\text{Ausprägung 1}) - 1)(N(\text{Ausprägung 2}) - 1)) . \quad (2.5)$$

Die Voraussetzungen für die Durchführung des χ^2 -Unabhängigkeitstests sind das Vorliegen von absoluten Häufigkeiten und die ausreichend große Anzahl der erwarteten Häufigkeiten, wobei die erwartete Häufigkeit jeder Zelle der Kontingenztafel ≥ 10 und die beobachtete Häufigkeit ≥ 2 sein muss (Gaus und Muche 2017, S. 281). Wenn die Bedingung der ausreichend großen Werte der erwarteten Häufigkeit nicht erfüllt wird, kann der exakte Fisher-Yates-Test (auch exakter χ^2 -Test genannt) angewandt werden.

Der exakte Fisher-Yates-Test wurde von Fisher beschrieben und ermöglicht das Testen auf Merkmalsunabhängigkeit, auch wenn die Bedingung der erwarteten Häufigkeit des χ^2 -Unabhängigkeitstests nicht erfüllt ist (Gaus und Muche 2017, S. 291). Der Test ist im RStudio unter dem Befehl „fisher.test“ aufzurufen. Zunächst werden die Einzelwahrscheinlichkeiten der gegebenen sowie aller noch extremeren Kontingenzfelder berechnet und anschließend zu dem p-Wert des Tests aufsummiert, wobei in dieser Gleichung a, b, c, d die absoluten Häufigkeiten der Kontingenztafeln darstellen (Bickeböller und Fischer 2007, S. 294; Gaus und Muche 2017, S. 291)

$$p = \frac{(a+b)!(c+d)!(a+c)!(b+d)!}{(a+b+c+d)!a!b!c!d!} . \quad (2.6)$$

Die statistische Auswertung des verhältnisskalierten Kriteriums Alter erfolgte mittels des auf einer Varianzanalyse beruhenden Kruskal-Wallis-Tests, bei dem mehr als zwei Parallelgruppen betrachtet werden können (Gaus und Muche 2017, S. 302ff.).

Im Programm RStudio ist der Test unter „kruskal.test“ zu finden.

Die Voraussetzung des Kruskal-Wallis-Tests ist, dass die Werte unabhängig und mindestens ordinal skaliert sind. H_0 des Tests besagt, dass es keine Unterschiede zwischen den zentralen Tendenzen der betrachteten Gruppen gibt. Bei dem Test werden die Daten in Ränge umgewandelt. Anschließend werden die Rangsummen der einzelnen Gruppen betrachtet und daraus die Testgröße H berechnet. Hierbei entspricht I der Anzahl der Gruppen, n_i der Anzahl der Fälle in der i -ten Gruppe, R_i der Rangsumme der i -ten Gruppe und n der Gesamtfallzahl in allen Gruppen zusammen

$$H = \frac{12}{n(n+1)} \sum_{i=1}^I \frac{R_i^2}{n_i} - 3(n+1) . \quad (2.7)$$

Zur Berechnung der Freiheitsgrade wird die Gesamtanzahl der Gruppen um eins vermindert (Gaus und Muche 2017 S. 311/ S. 328ff.).

2.3.3 Bestimmung des Hardy-Weinberg-Gleichgewichts (HWE)

Des Weiteren wurde untersucht, ob sich die erhobenen Genotypverteilungen im Hardy-Weinberg-Gleichgewicht (*Hardy-Weinberg-Equilibrium*, HWE) befinden, welches von Hardy (1908) und Weinberg (1908) zur gleichen Zeit unabhängig voneinander beschrieben wurde. Hierbei wird von einer idealisierten Allelvererbung nach der Mendelschen Vererbungslehre ohne Mutationen, Migrationen und Selektionen sowie ohne Einschränkung der Panmixie innerhalb einer unendlich großen Population ausgegangen, sodass das Auftreten eines bestimmten Genotyps allein von der Allelhäufigkeit bestimmt wird und die Wahrscheinlichkeit für das Auftreten eines Genotyps in jeder Generation gleich ist (Stern 1943; Bickeböller und Fischer 2007, S. 76ff.). Unter den genannten Voraussetzungen stellt sich für zwei Allele (A,a) mit ihren Allelhäufigkeiten (p,q) ein Gleichgewicht für die Genotyphäufigkeiten ein

$$(p + q)^2 = p^2 + 2pq + q^2 = 1 . \quad (2.8)$$

Ob die Genotypverteilung der in dieser Arbeit betrachteten Kohorten sich im HWE befinden und keiner äußeren Beeinflussung unterliegt, wurde mittels des χ^2 -Tests geprüft. Bei dem Vorliegen von H_0 wird davon ausgegangen, dass die Genotypverteilung im HWE liegt. Dies wurde bis zu einem $p \geq 0,05$ angenommen. Bei dem Testen auf ein HWE gilt, dass für den betrachteten Genlokus mit n Allelen und $n(n + 1)/2$ Genotypen $n - 1$ Allelhäufigkeiten geschätzt werden müssen, sodass die Teststatistik eine χ^2 -Verteilung mit folgendem Freiheitsgrad (Fg) hat (Bickeböller und Fischer 2007, S. 81f.)

$$Fg = \frac{n(n + 1)}{2} - 1 - (n - 1) . \quad (2.9)$$

Inwieweit die Verteilung der Genotypen dem HWE folgt, wurde mit Hilfe eines online verfügbaren Berechnungstools bestimmt¹¹.

2.3.4 Bestimmung des Kopplungsungleichgewichts (LD)

Zur Identifizierung von *tagSNPs* bzw. zur Bestätigung der Nutzbarkeit der genetischen Polymorphismen als gegenseitige Surrogate wurde das LD ermittelt, welches die nicht zufällige Allelassoziation verschiedener Genloki auf einem Haplotyp in einer Population angibt (Bickeböller und Fischer 2007, S. 89ff.; Slatkin 2008). Es wird von Haplotypblöcken mit mehreren assoziierten Genloki ausgegangen, zwischen denen Bereiche mit hoher Rekombinationsaktivität (*hot spots*) liegen (Bickeböller und Fischer 2007, S. 98ff.; Slat-

¹¹<http://www.dr-petrek.eu/documents/HWE.xls>, Zugriff am 13.01.2019

kin 2008). Wenn sich die beobachtete Auftretenshäufigkeit beider Allele zusammen von dem Produkt der erwarteten, unabhängigen Einzelauftrittshäufigkeit der Allele unterscheidet, befinden sie sich im LD (Slatkin 2008). Dies wird mittels des Koeffizienten D des LD dargestellt

$$D_{ab} = p_{ab} - p_a p_b . \quad (2.10)$$

Bei der Koeffizientenberechnung D_{ab} gibt p_{ab} die gemeinsame Auftretenshäufigkeit der Allele a und b (beobachtete Allelhäufigkeit), p_a die Einzelauftrittshäufigkeit des Allels a und p_b die Einzelauftrittshäufigkeit des Allels b (erwartete Allelhäufigkeit) an.

Würden die Allele a und b rein zufällig und unabhängig voneinander auftreten, so wäre $D = 0$ und die beiden Allele befänden sich im Kopplungsgleichgewicht (*Linkage Equilibrium*, LE) (Slatkin 2008).

Jedoch ist der LD-Koeffizient D abhängig von der Allelhäufigkeit, sodass die Vergleichbarkeit der D -Werte verschiedener Allelpaare nicht gegeben ist. Lewontin definierte einen standardisierten Wert D' , wobei der Koeffizient D durch den theoretischen Maximalwert von D , D_{\max} , dividiert wird

$$D' = \frac{D}{D_{\max}} \quad (2.11)$$

mit der Berechnungsvorschrift

$$D_{\max} = \begin{cases} \max\{-p_a p_b, -(1 - p_a)(1 - p_b)\}, & \text{wenn } D < 0 \\ \min\{p_a(1 - p_b), (1 - p_a)p_b\}, & \text{wenn } D > 0 \end{cases} . \quad (2.12)$$

Wenn $D' = 1$, dann entspricht die Allelassoziation der maximal erwarteten Assoziation. Hierbei wird von dem Vorliegen eines kompletten LD gesprochen (Carlson et al. 2004). Ein perfektes LD liegt vor, wenn der Korrelationsfaktor $r^2 = 1$, wobei als Voraussetzung für das Auftreten eines perfekten LD die MAFs der betrachteten Genloki übereinstimmen müssen (Carlson et al. 2004)

$$r^2 = \frac{D^2}{p_a(1 - p_a)p_b(1 - p_b)} . \quad (2.13)$$

Das LD wird durch viele Faktoren beeinflusst, unter anderem durch Rekombinationen, Mutationen, Selektionen und genetische Drifts, wodurch sich der Koeffizient D von Generation zu Generation verändert (Bickeböller und Fischer 2007, S. 89ff.; Slatkin 2008).

Die ausschließliche Betrachtung des LD gibt keine Information über die Kopplung, d. h. über die gemeinsame Vererbung der Allele der betrachteten Genloki. Jedoch bleibt ein LD bei starker Kopplung über mehrere Generationen hinweg bestehen, sodass ein starkes LD ein indirekter Hinweis auf das Vorliegen einer starken Kopplung und der engen Nachbarschaft der betrachteten Genloki ist (Bickeböller und Fischer 2007, S. 89ff.)

Das LD mit seinen standardisierten Koeffizienten D' und Korrelationskoeffizienten r^2 wurde unter der Verwendung der Webseite LDlink¹², welche sich auf Daten des 1000-Genome-Projekts (Phase 3, Version 5) bezieht, errechnet.

2.3.5 Ereigniszeitanalysen: Verfahren nach Kaplan-Meier und Cox-Regression

Die Ereigniszeitanalyse entspricht einer Analyse nach der Methode von Kaplan und Meier, die als erstes ein Verfahren für die Handhabung inkompletter Beobachtungen entwickelten und sogenannte zensierte Patienten in ihren Analysen berücksichtigten (Gaus und Muche 2017, S. 198). Ereignisse (*events*) werden in dieser Arbeit als Progression oder als Tod definiert. Jedoch erleben nicht alle Patienten ein Ereignis im Beobachtungszeitraum. Diese Patienten werden als zensiert eingestuft. Weitere Gründe für eine inkomplette Beobachtung und damit Zensierung sind das vorzeitige Ausscheiden der Patienten aus der Studie ohne Ereigniseintritt durch Verweigerung oder Unmöglichkeit der Nachbeobachtung (Dudley et al. 2016; Gaus und Muche 2017, S. 188ff.). Im RStudio wurden „survival“, das zur Berechnung der Tests diente, und „survminer“, welches die Ergebnisse visualisierte, als *packages* geladen und benutzt.

Die Ereigniszeit entspricht der Dauer vom Beobachtungsstart bis zum Ereigniseintritt. Jeder neu diagnostizierte Patient fängt bei dem Zeitpunkt $t = 0$ an, sodass alle beobachteten Patienten für die Berechnungen mit dem gleichen Anfangszeitpunkt beginnen und bis zum Eintritt des Ereignisses oder der Zensierung beobachtet werden. Grafisch können diese Beobachtungen mithilfe der sogenannten Kaplan-Meier-Kurven dargestellt werden (Abb. 2.1). Hierbei gibt jede Stufe in der Kurve mindestens einen Ereigniseintritt an und bildet somit das Ende einer horizontalen Linie und damit das Ende des jeweiligen Ereignisintervalls und gleichzeitig den Anfang des neuen Ereignisintervalls. Die zensierten Patienten werden als senkrechte Striche auf den Kurven vermerkt und beeinflussen nicht die Länge der jeweiligen Intervalle (Gaus und Muche 2017, S. 188ff.).

Auf der x-Achse der Kaplan-Meier-Kurven wird der Beobachtungszeitraum vom Beginn $t = 0$ bis zum Ende $t = t_{\max}$ dargestellt und auf der y-Achse wird die Überlebensrate ($S(t_i)$) bzw. allgemein die Wahrscheinlichkeit der Abwesenheit des Ereigniseintritts aufgetragen. Diese Überlebensrate entspricht der Wahrscheinlichkeit, bis zum Zeitpunkt t zu

¹²<https://ldlink.nci.nih.gov/>, Zugriff am 20.02.2019

überleben oder anders formuliert „dem Anteil der am Ende des [Zeitpunkts] t Lebenden bezogen auf die ursprüngliche Fallzahl unter Berücksichtigung von Zensierungen“ (Gaus und Muche 2017, S. 198). Die Überlebensrate ($S(t_i)$) als kumulierte Wahrscheinlichkeit kann mit dem Kaplan-Meier-Schätzer errechnet werden. Hierbei wird das Produkt aus den einzelnen Überlebenschancen ($U(t_i)$) der jeweiligen Intervalle ermittelt

$$S(t_i) = U(t_1) * U(t_2) * \dots * U(t_i) . \quad (2.14)$$

Die Überlebenschancen ($U(t_i)$) geben den Anteil der am Ende des Intervalls noch Lebenden zu dem Anteil der am Beginn des Intervalls Lebenden an, wobei n_i die Anzahl der Lebenden am Anfang des Intervalls t_i und d_i die Anzahl der innerhalb des Intervalls t_i Verstorbenen sei (Gaus und Muche 2017, S. 188ff.)

$$U(t_i) = \frac{n_i - d_i}{n_i} . \quad (2.15)$$

Es ist erkennbar, dass die zensierten Patienten die Fallzahl und damit die Überlebensraten ($S(t_i)$) beeinflussen. Diese Beeinflussung wird grafisch durch die unterschiedliche Länge der vertikalen Striche, die die horizontalen Ereignisintervalle verbinden und die Änderungen der Überlebensraten ($S(t_i)$) zwischen den Ereignisintervallen darstellen, angezeigt (Dudley et al. 2016) (Abb. 2.1).

Die beobachtete Zeit bei der Überlebensrate $S(t_i) = 50\%$ entspricht der medianen Überlebenszeit bzw. allgemein der medianen ereignisfreien Zeit.

Aus den Kaplan-Meier-Kurven kann die aktuelle Fallzahl zu den verschiedenen Zeitpunkten nicht abgelesen werden, weswegen zumeist unter den Kurven eine Tabelle mit den Fallzahlen der Patienten, die noch unter Risiko des Ereigniseintritts stehen, angegeben ist. Hierbei entsprechen die Patienten unter Risiko denjenigen Patienten, die an dem jeweiligen Beobachtungszeitpunkt noch kein Ereignis erlebt haben und noch nicht zensiert wurden. Außerdem kann mit der Angabe der Anzahl der Patienten unter Risiko die Aussagekraft des Kaplan-Meier-Schätzers beurteilt werden. Eine kleine Fallzahl noch lebender Patienten am Ende der Kurven bedeutet eine hohe Wahrscheinlichkeit für größere Zufallsfehler. Daher wird empfohlen, die Kurvendarstellung bei einer Fallzahl von 10–20 % der Patienten bei einer großen Ausgangsfallzahl oder bei einer Fallzahl ≤ 10 Patienten bei einer kleinen Ausgangsfallzahl zu beenden (Dudley et al. 2016).

Aufgrund der schlechteren Prognose von Pankreaskarzinompatienten werden kleine Fallzahlen vermutlich früher erreicht als bei Leberkarzinompatienten, weswegen die x-Achse bei den Kaplan-Meier-Kurven in dieser Arbeit bei den PDAC-Patienten auf drei Jahre

($t = 1095$ Tage) und bei den HCC-Patienten auf vier Jahre ($t = 1461$ Tage) limitiert wurde. Zusätzlich sind teilweise Diagramme unter den Kaplan-Meier-Kurven dargestellt, die die Anzahl der zensierten Patienten zu verschiedenen Zeitpunkten verdeutlichen (Abb. 2.1).

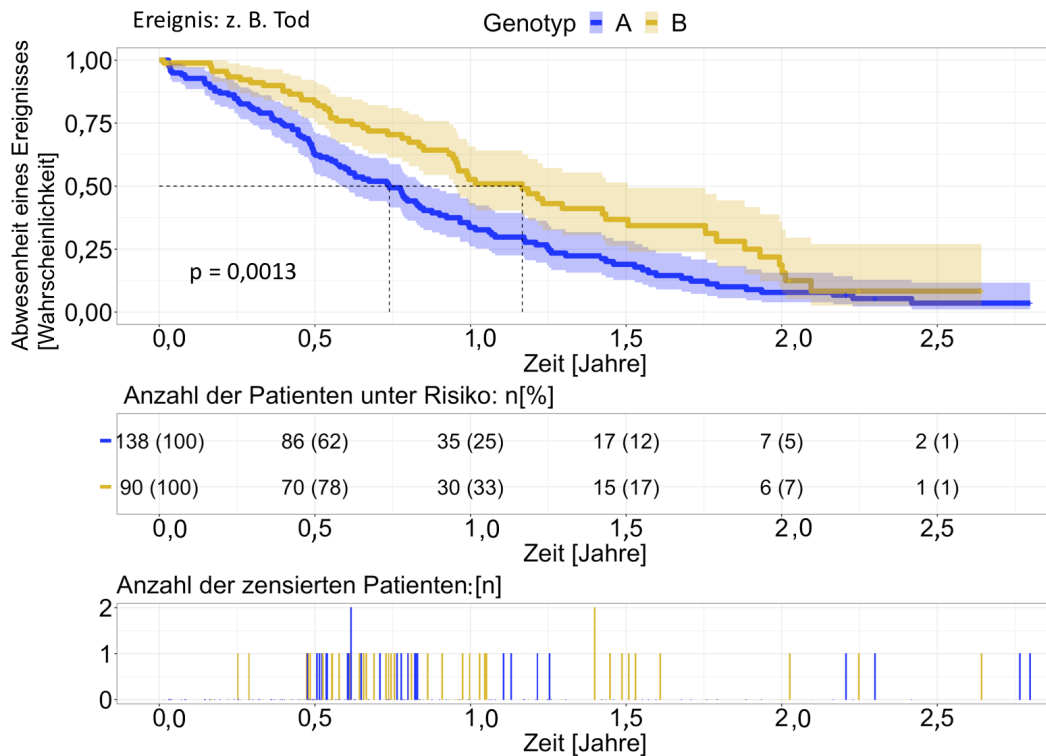


Abb. 2.1: Exemplarische Kaplan-Meier-Kurvendarstellung.

Darstellung der Wahrscheinlichkeit der Abwesenheit eines Ereignisses in Abhängigkeit vom Genotyp der Patienten für einen Ereigniszeitraum von 2,5 Jahren. Jede Stufe der Kurven gibt mindestens einen Ereigniseintritt an. Außerdem ist das Ergebnis des Log-Rank-Tests mit $p = 0,0013$ dargestellt. Die 95 % Konfidenzintervalle (KI), die den Vertrauensbereich der Werte angeben, sind dargestellt. In diesem Beispiel überschneiden sich die KI kaum und liegen kompakt um die Kurven. Die KI können zur Bestätigung der Aussagekraft der Werte und des Log-Rank-Testergebnisses herangezogen werden, wobei eine kleinere Spanne und kein Überschneiden für eine kleine Streuung der Werte und eine höhere Genauigkeit sprechen. Außerdem sind unter der Kaplan-Meier-Kurve der absolute und relative Anteil der Patienten unter Risiko angegeben. Die Patientenfallzahl nimmt beim Eintreten eines Ereignisses und bei Zensierung ab, sodass diese Anteile die Anzahl der zu dem Zeitpunkt lebenden, nicht zensierten Patienten angibt, bei denen das Risiko noch existiert, einem beobachtbaren Ereignis zu erliegen. Die unterste grafische Darstellung gibt die Anzahl der zensierten Patienten an, was auch bereits in der Kaplan-Meier-Kurve als vertikale Striche auf der Kurve verdeutlicht wird (in dieser Abbildung nicht erkennbar). Der Datensatz, der dieser Darstellung zugrunde liegt, stand innerhalb des Programms RStudio in dem *package* „survival“ für Lehrzwecke zur Verfügung.

Weiterhin wurden zur Validierung der TCGA-CDR-Daten die medianen Beobachtungszeiträume für die verschiedenen Ereigniszeiten (PFI/PFI.2/OS) errechnet und tabellarisch dargestellt (Anhang Kapitel 6.4).

Ein weiterer wichtiger Wert bei der Ereigniszeitanalyse ist die Angabe des Risikos innerhalb eines Zeitintervalls t_i zu versterben oder allgemein das definierte Ereignis zu erleben. Dieses Risiko wird auch als Hazard ($H(t_i)$) bezeichnet. Hierbei wird die Anzahl der Todesfälle/Ereigniseintritte d_i innerhalb eines Zeitintervalls t_i durch die Anzahl der Lebenden/Patienten ohne Ereignis am Anfang n_i des Zeitintervalls t_i dividiert (Bickel und Fischer 2007, S. 59ff.; Gaus und Muche 2017, S. 202ff.). Das Hazard kann sich im Verlauf der Zeit verändern (Gaus und Muche 2017, S. 205)

$$H(t_i) = \frac{d_i}{n_i} . \quad (2.16)$$

Werden zwei oder mehr Gruppen miteinander verglichen, kann das Verhältnis der einzelnen Hazards jeder Gruppe zueinander durch das Hazard Ratio ($HR(t_i)$) ermittelt werden. Das Hazard Ratio gibt an, um wie viel sich die Sterbewahrscheinlichkeit bzw. Wahrscheinlichkeit eines Ereigniseintritts einer Patientengruppe im Vergleich zu der anderen Patientengruppe unterscheidet. Hierfür werden die einzelnen Hazard-Werte der jeweiligen Gruppen dividiert, wobei $H_1(t_i)$ das Hazard der einen Gruppe und $H_2(t_i)$ das Hazard der anderen Gruppe darstellt (Gaus und Muche 2017, S. 202ff.)

$$HR(t_i) = \frac{H_1(t_i)}{H_2(t_i)} . \quad (2.17)$$

Um die im Nachfolgenden beschriebenen Analysen, d. h. die Log-Rank-Tests und die einfachen sowie multiplen Cox-Regressionsanalysen, durchführen zu können, ist das Vorliegen eines proportionalen Hazards erforderlich. Ein proportionales Hazard bedeutet, dass sich die einzelnen Hazards der verschiedenen Gruppen zwar im Verlauf der Zeit ändern können, jedoch gleichmäßig, sodass das Hazard Ratio über die Zeit konstant bleibt. Grafisch kann dies an gleichmäßig, proportional verlaufenden Kaplan-Meier-Kurven beurteilt werden, die sich nicht überschneiden dürfen (Gaus und Muche 2017, S. 205ff.). Auf das Vorliegen eines proportionalen Hazards kann mittels des Schoenfeld-Residuums getestet werden, das im RStudio mit der Funktion „cox.zph“ durchgeführt werden kann. Die Schoenfeld-Residuen $r_i(\beta)$ werden aus der Differenz zwischen dem beobachteten Variablenwert Z_i mit einem Ereignis zum Zeitpunkt t_i und dessen mittels des Regressionskoeffizienten β angepassten Mittelwerts zum Zeitpunkt t_i errechnet (Grambsch und Therneau 1994)

$$r_i(\beta) = Z_i - M(\beta, t_i) . \quad (2.18)$$

Die Residuen sind als zeitunabhängige Variablen definiert, sodass sie sich rein zufällig über die Zeitachse verteilen, wenn sie in Abhängigkeit der Zeit t grafisch dargestellt werden. Wenn diese zufällige Verteilung auftritt, kann von einem proportionalen Hazard ausgegangen werden (Grambsch und Therneau 1994).

Der Log-Rank-Test vergleicht die kompletten Überlebenskurven verschiedener Gruppen miteinander, wobei Todesfälle wie auch Eintritt eines anderen Ereignisses den Kurven zugrunde liegen können. Gleichzeitig gilt auch hier die Voraussetzung des proportionalen Hazards. H_0 geht vom Fehlen eines Unterschieds zwischen den verschiedenen Kaplan-Meier-Kurven aus, während H_1 einen Unterschied annimmt. Auch dem Log-Rank-Test liegt der χ^2 -Test als Berechnungshintergrund zugrunde, wobei die erwarteten Häufigkeiten der Todesfälle bzw. Ereigniseintritte mit den beobachteten verglichen werden (Gaus und Muche 2017, S. 296ff.). Die Freiheitsgrade (Fg) berechnen sich wie folgt

$$Fg = (N(\text{Gruppen}) - 1) . \quad (2.19)$$

Das multiple Cox-Regressionsmodell wurde von Cox beschrieben und ist eine Methode zur Schätzung der unabhängigen Einflüsse mehrerer freier Kriterien auf ein abhängiges Kriterium (bzw. das Zielereignis) bei gleichzeitigem Auftreten von zensierten Patienten (Gaus und Muche 2017, S. 384). Hierbei ist der ermittelte Einfluss eines freien Kriteriums auf das abhängige Kriterium für alle anderen betrachteten Einflussgrößen adjustiert. In dieser Arbeit sind die abhängigen Kriterien/Zielereignisse die Ereigniszeiten PFI, PFI.2 und OS.

Der Einfluss mehrerer Kriterien auf die Hazardfunktion, also die Wahrscheinlichkeit innerhalb eines Zeitintervalls das Zielereignis zu erfahren, wird abgeschätzt

$$H(t) = b_0(t)e^{(b_1x_1+b_2x_2+\dots+b_nx_n)} . \quad (2.20)$$

Hierbei sei $b_0(t)$ das *Baseline*-Hazard, was gilt, wenn alle freien Variablen null sind, b_n seien die zu schätzenden Regressionskoeffizienten der freien Variablen, die die Effektstärke angeben, und x_n seien die Werte der Einflusskriterien.

Die Parameter e^b sind die jeweiligen Hazard Ratios der freien Kriterien (Ziegler et al. 2007; Gaus und Muche 2017, S. 384f.). Dabei dienen die mittels des geschätzten Regres-

sionskoeffizienten berechneten, adjustierten Hazard Ratios als Maß für die Stärke des Zusammenhangs zwischen freiem und abhängigen Kriterium (Ziegler et al. 2007).

Zur Absicherung, dass das errechnete Cox-Regressionsmodell aussagekräftig bleibt, wurde als Grenze für die Anzahl der freien Kriterien die Bedingung gesetzt, dass pro Kriterium ≥ 10 Ereignisse eingetreten sein müssen (Peduzzi et al. 1995).

Zunächst wurde der allgemeine, nicht für die jeweils anderen Kriterien adjustierte Einfluss der einzelnen freien Kriterien auf das abhängige Kriterium durch ein einfaches Cox-Regressionsmodell abgeschätzt. Aufgrund der Modellierung der freien Kriterien Ätiologie, Differenzierungsgrad und Tumorstadium als Faktoren lagen zu viele Kriterien für ein multiples Cox-Regressionsmodell vor. In dieser Arbeit wurde für die modellierten Faktorkriterien jeweils die Ausprägung mit der größten Fallzahl zum Vergleich mit den anderen Ausprägungen genommen, wobei die Differenzierungsgrade G3 und G4 sowie Tumorstadien III und IV aufgrund ihrer geringen Fallzahlen zusammengefasst wurden. Alle Kriterien mit einem ermittelten signifikanten Einfluss ($p \leq 0,05$) oder Tendenz zum signifikanten Einfluss ($p \leq 0,1$) innerhalb des einfachen Cox-Regressionsmodells wurden in der multiplen Cox-Regressionsanalyse zur Schätzung der unabhängigen, für die jeweils anderen betrachteten Kriterien adjustierten Einflusstärke verwendet. Hierbei wurde im RStudio die Funktion „coxph“ genutzt. Zur Vermittlung eines Eindrucks der Einflusstärke der signifikanten Kriterien auf das Zielereignis wurden die einzelnen Kriterien unter Zuhilfenahme der RStudiofunktion „ggadjustedcurves“ grafisch innerhalb von Kaplan-Meier-Kurven dargestellt (Anhang Kapitel 6.5).

Die Arbeitsprozesse dieser Arbeit wurden zu Übersichtszwecken innerhalb eines Flussdiagramms dargestellt (Abb. 2.2).

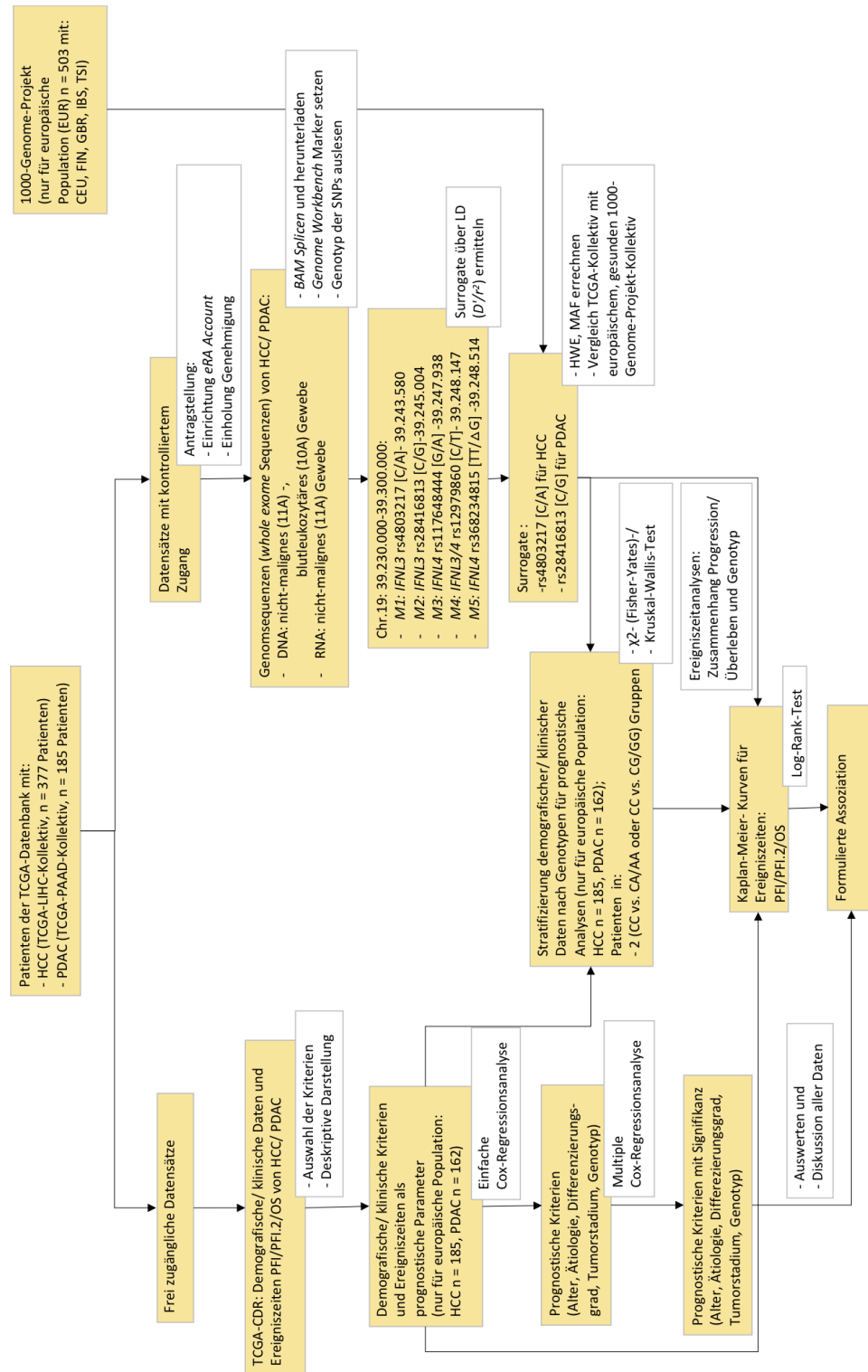


Abb. 2.2: Zusammenfassung der Arbeitsprozesse.

Die Abbildung illustriert die verwendeten Materialien (gelb hinterlegte Flächen) und die vorgegenommenen Arbeitsschritte (weiß hinterlegte Flächen).

3 Ergebnisse

3.1 Beschreibung der TCGA-LIHC- und TCGA-PAAD-Patientenkollektive

3.1.1 Übersicht über die demografischen und klinischen Daten des TCGA-LIHC-Kollektivs

Die TCGA-Datenbank enthält Daten von 377 Leberkrebspatienten, die mit einem HCC (C22.0) diagnostiziert wurden. Das Gesamtkollektiv der HCC-Patienten weist einen Altersdurchschnitt von 59,5 Jahren auf und eine Geschlechterverteilung von etwa 2,1 : 1 (m : f) (Tab. 3.1). Der Großteil der Patienten ist mit 51,0 % ($n = 187$) europäischer oder mit 43,9 % ($n = 161$) asiatischer Herkunft. Nur 4,6 % ($n = 17$) Patienten sind afrikanischer Herkunft und 0,5 % ($n = 2$) HCC-Patienten sind amerikanische Ureinwohner (Tab. 3.1).

Wie in dem Abschnitt *Material und Methoden* (Kapitel 2.1.1) beschrieben, wurde die mutmaßliche Ätiologie der HCC-Erkrankungen auf der Basis der serologischen und anamnestischen Daten der TCGA-Datenbank kategorisiert. Innerhalb des Gesamtkollektivs stellen HCC-Patienten mit einer HBV-Infektion mit 38,2 % ($n = 144$) den größten Anteil dar, gefolgt von 23,3 % ($n = 88$) der Patienten mit einer HBV-/HCV-Koinfektion. Der hohe Anteil HBV-Infizierter unter den HCC-Patienten ist auf die hohe Prävalenz einer HBV-Infektion unter Asiaten zurückzuführen (El-Serag und Rudolph 2007; El-Serag 2011). Diese Annahme lässt sich durch Stratifizierung der einzelnen Ethnizitäten nach den Risikofaktoren bestätigen. Die Patienten asiatischer Herkunft sind häufiger als Patienten europäischer Herkunft mit HBV oder einer Kombination aus HBV und HCV infiziert (Anhang Kapitel 6.3). Das Gesamtkollektiv der HCC-Patienten weist vorwiegend einen Tumordifferenzierungsgrad G2 (48,4 %, $n = 180$) und G3 (33,3 %, $n = 124$) sowie ein Tumorstadium I (49,6 %, $n = 175$) auf (Tab. 3.1).

Die europäischen HCC-Patienten sind im Mittel 3,8 Jahre älter als das Gesamtkollektiv und weisen bei der Geschlechterverteilung einen geringeren Anteil an männlichen Patienten auf (1,3 : 1 (m : f)) (Tab. 3.1). Die Risikofaktoren sind homogener verteilt und es überwiegt nicht wie in der asiatischen Population die HBV-Infektion (Tab. 3.1). Die Gruppe europäischer Patienten unterscheidet sich nicht wesentlich im Tumordifferenzierungsgrad oder im Tumorstadium von dem Gesamtkollektiv (Tab. 3.1).

Diese Arbeit beschränkt sich auf HCC- und PDAC-Patienten europäischer Herkunft, um genetisches *Confounding* zu verhindern.

Tab. 3.1: Demografische und klinische Daten des TCGA-LIHC-Kollektivs

	Patienten aller Ethnizitäten <i>n</i> = 377	Patienten europäischer Herkunft <i>n</i> = 187
¹ Alter bei Diagnose M ± SD	59,5 ± 13,5	63,3 ± 13,8
Median Quartilsabstand [Jahre]	61,0 18,0	66,0 15,0
Geschlecht m/f [n(m %)]	255 (67,6)/122	105 (56,1)/82
² Ethnizität [n(%)]		
Europäische Herkunft	187 (51,0)	187 (100,0)
Asiatische Herkunft	161 (43,9)	
Afrikanische Herkunft	17 (4,6)	
Amerikanische Ureinwohner	2 (0,5)	
Ätiologie [n(%)]		
Kein Risiko	54 (14,3)	38 (20,3)
Kein virales Risiko	43 (11,4)	35 (18,7)
HBV	144 (38,2)	42 (22,5)
HCV	48 (12,7)	31 (16,6)
HBV-/HCV-Koinfektion	88 (23,3)	41 (21,9)
³ Differenzierungsgrad [n(%)]		
G1	55 (14,8)	34 (18,6)
G2	180 (48,4)	102 (55,7)
G3	124 (33,3)	46 (25,1)
G4	13 (3,5)	1 (0,5)
⁴ Stadium der Tumorerkrankung [n(%)]		
I	175 (49,6)	80 (47,3)
II	87 (24,6)	44 (26,0)
III	86 (24,4)	41 (24,3)
IV	5 (1,4)	4 (2,4)
⁵ IFNL3 rs4803217 [C/A]		
CC:CA:AA	227:123:26	79:89:19
MAF	0,233	0,340
HWE	<i>p</i> = 0,10	<i>p</i> = 0,40

¹Altersangaben waren für 376/377 bzw. 187/187 Patienten verfügbar

²Angaben zur Ethnizität waren für 367/377 bzw. 187/187 Patienten verfügbar

³Daten zum Tumordifferenzierungsgrad waren für 372/377 bzw. 183/187 Patienten verfügbar

⁴Daten zum Stadium der Tumorerkrankungen waren für 353/377 bzw. 169/187 Patienten verfügbar

⁵Genotypen konnten für 376/377 bzw. 187/187 Patienten ausgelesen werden

3.1.2 Übersicht über die demografischen und klinischen Daten des TCGA-PAAD-Kollektivs

Die TCGA-Datenbank enthält Daten von 185 Pankreaskrebspatienten, die mit einer bösartigen Neubildung des Pankreas innerhalb verschiedener Lokalisationen (C25.-) diagnostiziert wurden.

Im Vergleich zu dem Gesamtkollektiv der HCC-Patienten weist das Gesamtkollektiv der PDAC-Patienten einen etwa 5,2 Jahre höheren Altersdurchschnitt von 64,7 Jahren und bei der Geschlechterverteilung einen geringeren Anteil an Männern auf (1,2 : 1 (m : f)) (Tab. 3.2).

Mit 90,0 % ($n = 162$) ist der Großteil der PDAC-Patienten europäischer Herkunft und nur 6,1 % ($n = 11$) sind asiatischer und 3,9 % ($n = 7$) sind afrikanischer Herkunft (Tab. 3.2).

Bei den europäischen Patienten liegt ein Altersdurchschnitt von 65,4 Jahren und eine Geschlechterverteilung von 1,3 : 1 (m : f) vor (Tab. 3.2).

Wie die HCC-Patienten weisen die PDAC-Patienten europäischer Herkunft bei Diagnose vorwiegend einen Tumordifferenzierungsgrad G2 (52,2 %, $n = 84$) und G3 (29,2 %, $n = 47$) auf. Im Vergleich zu den HCC-Patienten weist die Mehrheit der PDAC-Patienten jedoch zum Zeitpunkt der Diagnose das fortgeschrittenere Tumorstadium II (81,8 %, $n = 130$) auf. Dies spiegelt das verhältnismäßig späte Erkennen einer PDAC-Erkrankung wider. Aus Gründen der Vergleichbarkeit und um auch in dieser Kohorte Verzerrungen aufgrund verschiedener Ethnizitäten zu vermeiden, werden im Folgenden auch bei den PDAC-Patienten nur diejenigen europäischer Herkunft betrachtet.

Tab. 3.2: Demografische und klinische Daten des TCGA-PAAD-Kollektivs

	Patienten aller Ethnizitäten <i>n</i> = 185	Patienten europäischer Herkunft <i>n</i> = 162
¹ Alter bei Diagnose M ± SD	64,7 ± 11,1	65,4 ± 10,8
Median Quartilsabstand [Jahre]	65,0 16,0	65,0 15,5
Geschlecht m/f [n(m %)]	102 (55,1)/83	92 (56,8)/70
² Ethnizität [n(%)]		
Europäische Herkunft	162 (90,0)	162 (100,0)
Asiatische Herkunft	11 (6,1)	
Afrikanische Herkunft	7 (3,9)	
³ Differenzierungsgrad [n(%)]		
G1	32 (17,6)	28 (17,4)
G2	97 (53,3)	84 (52,2)
G3	51 (28,0)	47 (29,2)
G4	2 (1,1)	2 (1,2)
⁴ Stadium der Tumorerkrankung [n(%)]		
I	21 (11,5)	20 (12,6)
II	152 (83,5)	130 (81,8)
III	4 (2,2)	4 (2,5)
IV	5 (2,7)	5 (3,1)
<i>IFNL3</i> rs28416813 [C/G]		
CC:CG:GG	86:79:20	76:69:17
MAF	0,322	0,318
HWE	<i>p</i> = 0,77	<i>p</i> = 0,82

¹Altersangaben waren für 182/185 bzw. 159/162 Patienten verfügbar

²Angaben zur Ethnizität waren für 180/185 bzw. 162/162 Patienten verfügbar

³Daten zum Tumordifferenzierungsgrad waren für 182/185 bzw. 161/162 Patienten verfügbar

⁴Daten zum Stadium der Tumorerkrankungen waren für 182/185 bzw. 159/162 Patienten verfügbar

3.2 Erheben von Genotypen polymorpher Positionen in den Typ-III-IFN-Genen sowie Identifizierung und Validierung von *tagSNPs*

Die WXS-DNA- und RNA-Sequenzdaten der TCGA-Datenbank des blutleukozytären Gewebematerials (Code: 10A) wurden bei den HCC- und den PDAC-Patienten an fünf polymorphen Positionen im *IFNL*-Genlokus ausgelesen. Die Phänotypen der Polymorphismen wurden bereits in der Einleitung ausführlich vorgestellt (Tab. 1.1 in Kapitel 1.3.4).

Tab. 2.1 in Kapitel 2.1.4 listet ihre exakten Positionen auf Chromosom 19 auf, die in der NCBI *Genome Workbench* mit M1–M5 markiert wurden.

Abb. 3.1 zeigt ein repräsentatives *alignment* der *reads* des *IFNL*-Genlokus an Chromosom 19. Das *alignment* lässt erkennen, dass die Methode des WXS angewandt wurde, denn die codierenden Regionen wurden wesentlich häufiger abgelesen als intergenische Abschnitte.



Abb. 3.1: Übersicht über den *IFNL*-Genlokus auf Chromosom 19 innerhalb der NCBI *Genome Workbench*.

Das repräsentative *alignment* der *reads* (grau hinterlegt) verdeutlicht die Anwendung der Methode des WXS. Die Marker, die die fünf polymorphen Positionen markieren, sind in Abb. 3.2 deutlicher unterscheidbar.

Eine Übersicht über den *IFNL3*- und den *IFNL4*-Lokus zeigt die unterschiedlichen Sequenzierertiefen an diesen fünf polymorphen Positionen (Abb. 3.2). Die Methodik des WXS hat für das *IFNL3*-Gen wesentlich mehr *reads* generiert als für das *IFNL4*-Gen.



Abb. 3.2: Übersicht über den *IFNL3*- und den *IFNL4*-Genlocus mit den fünf markierten polymorphen Positionen (M1–M5) innerhalb der NCBI *Genome Workbench*. Der *IFNL3*- und der *IFNL4*-Genlocus wurden auf Chromosom 19 im Referenzgenom sowie die daran ausgerichteten *reads* (grau hinterlegt) dargestellt.

Abb. 3.3 zeigt in einer höheren Auflösung für zwei Beispiele die *reads* an der Position M2. Im ersten Beispiel gibt es bei ungefähr 24 *reads* bis auf eine Ausnahme nur *reads* des Allels G, sodass hier vom homozygoten Genotyp GG ausgegangen wird (Abb. 3.3 (A)). Im zweiten Beispiel finden sich bei 189 *reads* insgesamt 116 *reads* für das Allel C und 73 *reads* für das Allel G, sodass hier der heterozygote Genotyp CG vorliegt (Abb. 3.3 (B)). Diese Annahmen genügen den in *Material und Methoden* (Kapitel 2.1.4) beschriebenen Kriterien, die für das Auslesen von Genotypen angelegt wurden.

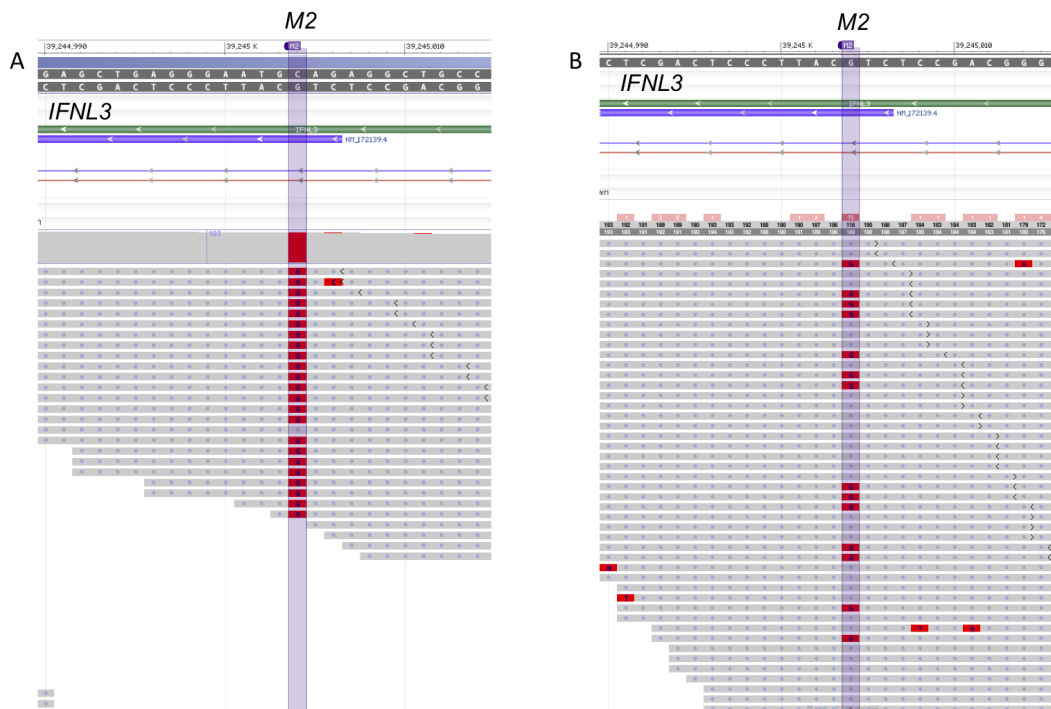


Abb. 3.3: Beispiel für einen *minor* homozygoten [GG] und einen heterozygoten [CG] *IFNL3*-rs28416813-Genlocus.

Die grauen Felder stellen die *major* Allelvariante C und die roten Felder die *minor* Allelvariante G dar. Der homozygote Genotyp (A) weist bei 23/24 *reads* das Allel G auf. Der heterozygote Genotyp (B) weist in diesem Beispiel bei insgesamt $n = 189$ sequenzierten *reads* bei $n = 73$ *reads* das Allel G und bei $n = 116$ *reads* das Allel C auf (nicht vollständig dargestellt).

Abb. 3.4 zeigt in einer noch höheren Auflösung die Position M5 im Exon 1 des *IFNL4*-Gens mit insgesamt nur vier *reads*. Alle diese vier *reads* sind mit dem Allel TT identisch. Als Grund für diese geringe Sequenziertiefe an der Position M5 des *IFNL4*-rs368234815-Polymorphismus ist anzunehmen, dass die Position aufgrund ihrer Pseudogenisierung in der Bevölkerung beim Sequenzieren nach der WXS-Methode nicht als Exon erkannt und deswegen seltener ausgelesen wurde (Prokunina-Olsson et al. 2013).

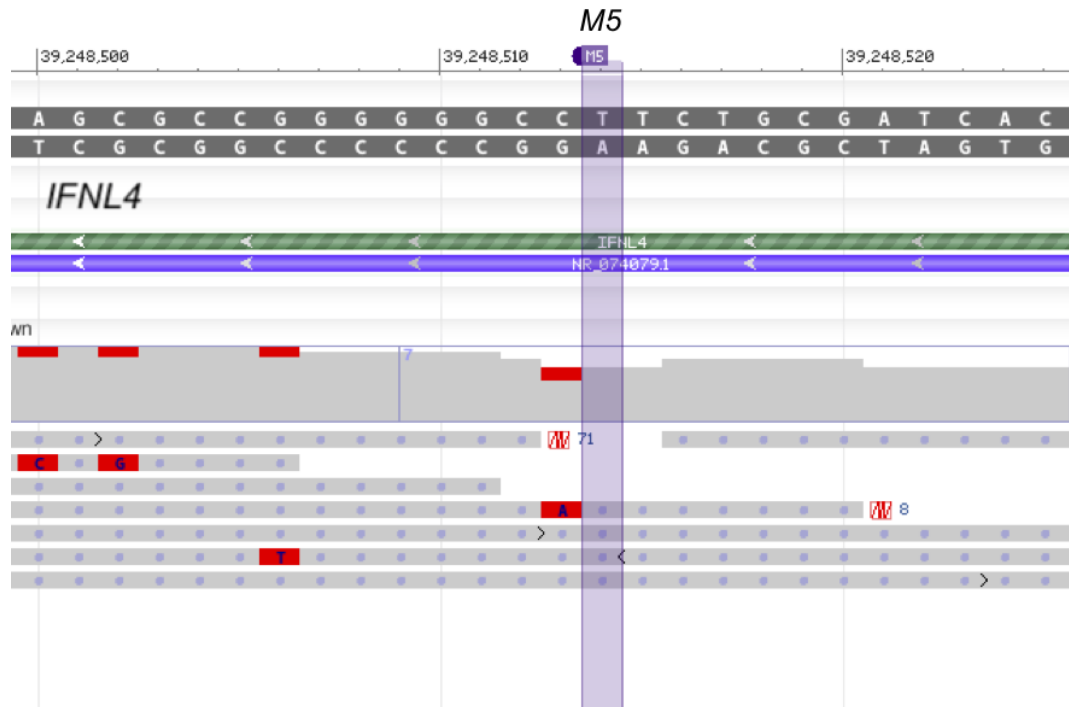


Abb. 3.4: Beispielhafte Darstellung des *IFNL4*-rs368234815-Genlokus.

Deutlich erkennbar ist, dass die erforderliche Mindestanzahl an *reads* mit den hier dargestellten $n = 4$ *reads* an Position M5 nicht erreicht wird.

Die Überprüfung des LD zwischen den fünf polymorphen Positionen von Interesse innerhalb der *IFNL3*- und *IFNL4*-Genloki diente der Identifizierung und Validierung von *tag*SNPs bzw. der Sicherstellung der gegenseitigen Nutzbarkeit als Surrogate. Hierbei stellen *tag*SNPs einzelne repräsentative Polymorphismen innerhalb eines Haplotypblocks dar, die im LD mit anderen Polymorphismen desselben Haplotypblocks stehen und zu ihnen redundante Informationen tragen. Die Typisierung alleine dieser informativen *tag*SNPs sollte ausreichen, um häufige Haplotypen identifizieren zu können (Bickeböller und Fischer 2007, S. 98ff.).

Als Grundlage für die Berechnungen dienten Daten des 1000-Genome-Projekts¹. Der standardisierte LD-Koeffizient D' und der Korrelationskoeffizient r^2 wurden sowohl für die Gesamtheit aller im 1000-Genome-Projekt vertretenen Individuen ($n = 2504$) sowie für diejenigen europäischer Herkunft ($n = 503$) berechnet.

Für beide betrachteten Gruppen zeigt sich, dass die fünf polymorphen Positionen im LD stehen, wobei der standardisierte LD-Koeffizient D' anzeigt, dass bei allen Patienten nahezu ein komplettes LD ($D' = 1$) vorliegt (Abb. 3.5). Vier der fünf Positionen befinden

¹<https://ldlink.nci.nih.gov/>, Zugriff am 20.02.2019

sich zusätzlich noch in einem nahezu perfekten LD, was durch den Korrelationsfaktor $r^2 = 1$ verdeutlicht wird. Die Abweichung des errechneten r^2 -Wertes bei dem SNP *IFNL4* rs117648444 lässt sich mit seiner gegenüber den anderen Polymorphismen geringeren MAF erklären (0,118 vs. 0,310) (Abb. 3.5).

Bis auf den SNP *IFNL4* rs117648444 sind die betrachteten polymorphen Positionen als *tag*SNPs und zur gegenseitigen Validierung verwendbar.

Das LD zwischen den *IFNL*-Polymorphismen erlaubt Rückschlüsse von den Allelen mit genügender Sequenziertiefe auf Allele, deren Sequenziertiefe zum direkten Auslesen des Genotyps nicht aussagekräftig genug ist. Bei den HCC-Patienten fand sich eine ausreichende Sequenziertiefe des SNP *IFNL3* rs4803217. Dieser SNP steht im LD mit dem *IFNL4* rs368234815 und dient bei den HCC-Patienten als Surrogat ($D' = 0,991$, $r^2 = 0,968$) (Abb. 3.5). Bei den PDAC-Patienten weist dieser SNP *IFNL3* rs4803217 mit durchschnittlich etwa 10 *reads* nur eine ungenügende Sequenziertiefe auf. Jedoch zeigt der SNP am 5'-Terminus *IFNL3* rs28416813 eine belastbare Sequenziertiefe von bis zu 200 *reads* bei den PDAC-Patienten auf (Abb. 3.3). Dieser SNP steht auch im LD mit *IFNL4* rs368234815 ($D' = 0,977$, $r^2 = 0,954$) und dient in dieser Arbeit bei den PDAC-Patienten als Surrogat für *IFNL4* rs368234815 (Abb. 3.5).

3.3 Stratifizierung der HCC- und PDAC-Patienten europäischer Herkunft nach ihren *IFNL*-Genotypen

Mit der beschriebenen Methodik (Kapitel 2.1.4, Kapitel 3.2) konnten für nahezu alle Tumorpatienten die Genotypen der *IFNL*-Keimbahnvariationen ausgelesen werden. Die sogenannte *call rate* der HCC-Patienten aller Ethnizitäten beträgt 99,7 % (376/377), die der HCC-Patienten europäischer Herkunft 100 % (187/187). Die *call rate* der PDAC-Patienten aller Ethnizitäten beträgt ebenfalls 100 % (185/185).

Auf Basis der Sequenziertiefe wurde für die HCC-Patienten der SNP *IFNL3* rs4803217 und für die PDAC-Patienten der SNP *IFNL3* rs28416813 als Surrogate gewählt. Diese spiegeln die genetische Disposition, das funktionsfähige Protein IFN- λ_4 zu codieren oder aufgrund einer Leserasterverschiebung nur noch eine funktionale *knock-out*-Variante zu codieren (*IFNL4* rs368234815), wider.

Für das Gesamtkollektiv der HCC-Patienten ergibt sich eine Genotypverteilung des Surrogats *IFNL3* rs4803217 von 227 : 123 : 26 (CC : CA : AA), für die Patienten europäischer Herkunft eine Verteilung von 79 : 89 : 19 (CC : CA : AA) (Tab. 3.1).

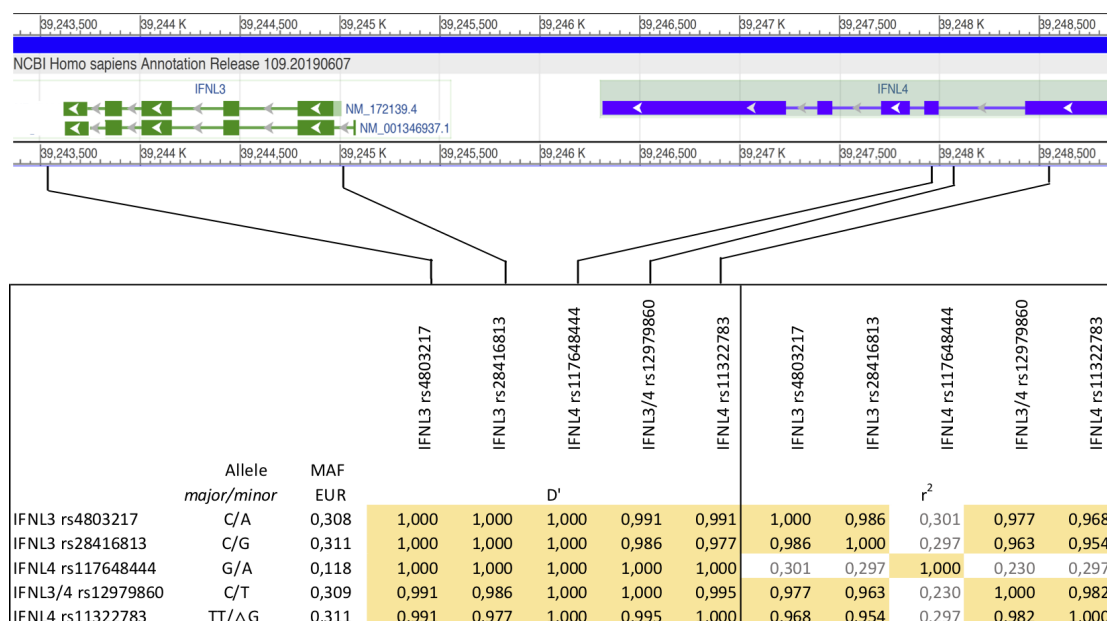


Abb. 3.5: Berechnung des Kopplungsungleichgewichts (*Linkage Disequilibrium*, LD) der fünf *IFNL*-Polymorphismen.

Das LD wurde mit dem standardisierten Koeffizienten D' und dem Korrelationskoeffizienten r^2 auf der Basis der Daten des 1000-Genome-Projekts ermittelt. Für diese Darstellung wurde auf die Population europäischer Herkunft ($n = 503$) zurückgegriffen. Der standardisierte Koeffizient D' zeigt, dass alle fünf SNPs in einem nahezu kompletten LD zueinander stehen. Der Korrelationskoeffizient r^2 zeigt, dass sich die Polymorphismen bis auf *IFNL4* rs117648444, aufgrund seiner geringeren MAF, als *tag*SNPs bzw. als gegenseitige Surrogate eignen. Für den Dinukleotidpolymorphismus *IFNL4* rs368234815 wurde für diese Berechnungen der in den Daten enthaltene SNP *IFNL4* rs11322783 genutzt.

Bei einer MAF von 0,340 für die HCC-Patienten europäischer Herkunft sind insgesamt 108 Patienten Träger des selteneren Allels A, d. h. diese Patienten sind aufgrund ihrer Disposition in der Lage, ein funktionstüchtiges IFN- λ_4 zu exprimieren. Bei den 79 für das häufigere Allel C Homozygoten wird angenommen, dass sie aufgrund der die Translation beendenden Variante TT an Position *IFNL4* rs368234815 kein funktionstüchtiges IFN- λ_4 -Protein ausbilden können (*knock-out*-Variante).

Für das Gesamtkollektiv der PDAC-Patienten ergibt sich für das Surrogat *IFNL3* rs28416813 eine Verteilung der Genotypen von 86 : 79 : 20 (CC : CG : GG) und für die Patienten europäischer Herkunft von 76 : 69 : 17 (CC : CG : GG) (Tab. 3.2). Hierbei sind mit einer MAF von 0,318 86 der europäischen PDAC-Patienten Träger des selteneren Allels G und damit zur Expression eines funktionstüchtigen IFN- λ_4 -Proteins disponiert. Demgegenüber unterliegt das IFN- λ_4 bei den 76 Allel-C-Homozygoten einer Pseudogenisierung (Tab. 3.2).

Die Genotypverteilung sowohl der Gesamtkollektive als auch der HCC- und PDAC-Patienten europäischer Herkunft befinden sich im HWE ($p > 0,05$) (Tab. 3.1, Tab. 3.2).

3.3.1 Vergleich der TCGA-LIHC- und TCGA-PAAD-Kollektive mit einem Kontrollkollektiv aus dem 1000-Genome-Projekt hinsichtlich der Genotypverteilung

Nach der Ermittlung der Genotypverteilung beider Patientenkollektive wurden diese mit der Genotypverteilung eines Referenzkollektivs aus dem 1000-Genome-Projekt, hier einer Gruppe europäischer Herkunft ($n = 503$), verglichen² (Tab. 3.3).

Zwischen dem Kollektiv der HCC-Patienten und dem Kontrollkollektiv unterscheidet sich die Verteilung der Genotypen CC und CA um jeweils etwa 6 % (Tab. 3.3). Dieser relativ große Unterschied konnte jedoch nicht als signifikant gesichert werden (χ^2 -Test: $p = 0,34$).

Die absoluten und relativen Häufigkeiten der Genotypverteilungen der PDAC-Patienten und des Referenzkollektivs des 1000-Genome-Projekts stimmen nahezu überein (Tab. 3.3). Die MAFs der Populationen entsprechen sich mit etwa 0,310 bis auf die MAF des Kollektivs der HCC-Patienten, das eine etwas höhere MAF von 0,340 aufweist (Tab. 3.3). Die Verteilung der Genotypen in dem Kontrollkollektiv befindet sich an den beiden betrachteten polymorphen Positionen (*IFNL3* rs4803217, *IFNL3* rs28416813) im HWE ($p > 0,05$) (Tab. 3.3). Diese Ergebnisse sprechen für das Vorliegen einer repräsentativen Stichprobe der europäischen Population aus den TCGA-LIHC- und TCGA-PAAD-Kollektiven. Des Weiteren zeigen sich keine Anhaltspunkte, dass die Allele Risikofaktor für die Entstehung von HCC- oder PDAC-Tumoren sind.

Tab. 3.3: Vergleich der TCGA-LIHC- und TCGA-PAAD-Kollektive mit einem Kontrollkollektiv aus dem 1000-Genome-Projekt hinsichtlich der Genotypverteilung

Patienten europäischer Herkunft (<i>IFNL3</i> rs4803217 [C/A]), [n(%)]							
	Total	CC	CA	AA	MAF	HWE	<i>p</i>
1000-Genome-Projekt	503 (100,0)	243 (48,3)	210 (41,7)	50 (9,9)	0,308	$p = 0,64$	0,34
TCGA-LIHC-Kollektiv	187 (100,0)	79 (42,2)	89 (47,6)	19 (10,2)	0,340	$p = 0,40$	
Patienten europäischer Herkunft (<i>IFNL3</i> rs28416813 [C/G]), [n(%)]							
	Total	CC	CG	GG	MAF	HWE	<i>p</i>
1000-Genome-Projekt	503 (100,0)	243 (48,3)	207 (41,2)	53 (10,5)	0,311	$p = 0,37$	0,95
TCGA-PAAD-Kollektiv	162 (100,0)	76 (46,9)	69 (42,6)	17 (10,5)	0,318	$p = 0,82$	

²http://grch37.ensembl.org/Homo_sapiens, Zugriff am 03.08.2019

3.3.2 Analyse der Strata hinsichtlich demografischer und klinischer Daten

Die Fallzahlen sowohl der HCC- als auch der PDAC-Patienten mit Homozygotie für das *minor* Allel sind sehr gering ($n = 19$ für HCC-Patienten, $n = 17$ für PDAC-Patienten). Daher wurden für beide Kollektive alle nachfolgenden Analysen nicht in Bezug auf drei Genotypen, sondern in Bezug auf die Träger des selteneren, für ein funktionstüchtiges IFN- λ_4 codierenden Allels, gegenüber den für das häufigere Allel Homozygoten durchgeführt (*IFNL3* rs4803217 CA/AA vs. CC für HCC Patienten, *IFNL3* rs28416813 CG/GG vs. CC für PDAC Patienten) (Tab. 3.4, Tab. 3.5).

Ein Zusammenhang zwischen den Genotypen der Patienten und ihrem Alter bei Tumordiagnose wurde bei der Betrachtung der Altersmittelwerte und -mediane sowie der grafischen Darstellung innerhalb von Boxplots (Anhang Abb. A.2) nicht erwartet und konnte bei beiden Patientenkollektiven nicht gezeigt werden (Kruskal-Wallis-Test: $p(\text{HCC}) = 0,31$, $p(\text{PDAC}) = 0,98$) (Tab. 3.4, Tab. 3.5). Innerhalb der gebildeten Strata verteilen sich Männer und Frauen in beiden Kollektiven etwa gleichmäßig (56 % : 44 %) (Tab. 3.4, Tab. 3.5).

Obwohl sich die Verteilung der Risikofaktoren der HCC-Patienten in den für die *IFNL*-Genotypen gebildeten Gruppen deskriptiv und explorativ unterscheidet, konnte der beobachtete Unterschied nicht gesichert werden (χ^2 -Test: $p(\text{HCC}) = 0,59$) (Tab. 3.4, Anhang Abb. A.3, Anhang Kapitel 6.3).

Der Differenzierungsgrad der Tumorzellen wird histologisch durch den Vergleich mit organgleichen nicht-malignen Zellen ermittelt. Je undifferenzierter die Tumorzellen sind, d. h. je höher der Differenzierungsgrad, desto aggressiver ist der Tumor auch in seinem Metastasierungsverhalten und desto schlechter ist höchstwahrscheinlich die Prognose im Vergleich zu einem niedrigeren Differenzierungsgrad (Edge et al. 2010).

Das Stadium der Tumorerkrankung, das sich innerhalb der TCGA-Datenbank nach dem AJCC-Klassifizierungssystem richtet, definiert sich über die Größe, die Ausdehnung und gegebenenfalls über das Vorliegen von Metastasen des Tumors, wobei die lokale Ausdehnung, die Verbreitung in regionale Lymphknoten und die Metastasierung in andere Gewebe betrachtet werden. Es gilt zumeist, je größer und ausgedehnter der Tumor ist (d. h. je höher das Stadium der Tumorerkrankung), desto schlechter ist die Prognose (Edge et al. 2010).

Angaben zum Tumordifferenzierungsgrad standen für annähernd alle HCC- und PDAC-Patienten zur Verfügung (Tab. 3.1, Tab. 3.2). Angaben zum Stadium der Tumorerkrankung fehlten bei 18 HCC-Patienten und bei 3 PDAC-Patienten (Tab. 3.1, Tab. 3.2). Aufgrund der geringen Fallzahl von Patienten beider Kollektive mit dem höchsten Differenzierungsgrad G4 und mit den fortgeschrittenen Tumorstadien III und

IV wurden Patienten mit den Differenzierungsgraden G3 und G4 sowie Patienten mit den Tumorstadien III und IV für weitere Berechnungen zusammengefasst (Tab. 3.4, Tab. 3.5).

Bei der Betrachtung der HCC-Patienten europäischer Herkunft wurden die meisten Patienten mit einem mittleren Differenzierungsgrad G2 (55,7 %, $n = 102$) bzw. G3 (25,1 %, $n = 46$) diagnostiziert (Tab. 3.4). In diesen Fällen ist die Dedifferenzierung der Tumorzellen bereits fortgeschritten. Patienten mit dem *IFNL3*-rs4803217-Genotyp CC unterscheiden sich nicht signifikant im Tumordifferenzierungsgrad von den Trägern des Allels A (CA/AA) (χ^2 -Test: $p = 0,87$) (Tab. 3.4).

Bei den insgesamt 169 HCC-Patienten europäischer Herkunft mit Angaben zum Stadium der Tumorerkrankung wurden die meisten Patienten in einem Stadium I (47,3 %, $n = 80$) diagnostiziert. Stadium I bedeutet, dass ein nicht metastasierter, kleiner Tumor ohne Infiltration der Nachbargewebe und ohne Lymphknotenbefall vorliegt (Edge et al. 2010). Auch hier unterscheiden sich Patienten mit dem *IFNL3*-rs4803217-Genotyp CC nicht signifikant im Tumorstadium von den Trägern des Allels A (CA/AA) (χ^2 -Test: $p = 0,69$) (Tab. 3.4).

Genau wie bei den HCC-Patienten wurden die meisten PDAC-Patienten mit einem mittleren Differenzierungsgrad G2 (52,2 %, $n = 84$) bzw. G3 (29,2 %, $n = 47$) diagnostiziert (Tab. 3.5). Für die Patienten mit dem *IFNL3*-rs28416813-Genotyp CC konnte kein signifikanter Unterschied im Tumordifferenzierungsgrad zu den Trägern des Allels G (CG/GG) aufgezeigt werden (χ^2 -Test: $p = 0,42$) (Tab. 3.5).

Im Unterschied zu den HCC-Patienten wurden deutlich mehr PDAC-Patienten zum Zeitpunkt der Diagnose in einem Tumorstadium II (81,8 %, $n = 130$) diagnostiziert (Tab. 3.5). Jedoch weisen auch hier die PDAC-Patienten mit dem *IFNL3*-rs28416813-Genotyp CC keinen signifikanten Unterschied im Tumorstadium zu den Trägern des Allels G (CG/GG) auf (exakter Fisher-Yates-Test: $p = 0,86$) (Tab. 3.5).

Tab. 3.4: Analysen zur Assoziation der *IFNL3*-rs4803217-Genotypen mit demografischen und klinischen Daten für HCC-Patienten

	Total <i>n</i> = 187	CC <i>n</i> = 79	CA <i>n</i> = 89	AA <i>n</i> = 19	CA/AA <i>n</i> = 108	<i>p</i>
Alter bei Diagnose M ± SD	63,3 ± 13,8	65,1 ± 11,6	61,4 ± 15,7	64,8 ± 12,4	62,0 ± 15,1	0,31 ³
Median Quartilsabstand[Jahre]	66,0 15,0	66,0 14,0	65,0 17,0	66,0 13,0	65,0 16,0	
Geschlecht m/f [n(m %)]	105(56,1)/82	42(53,2)/37	51(57,3)/38	12(63,2)/7	63(58,3)/45	n.t.
Ätiologie [n(%)]						
Kein Risiko	38 (20,3)	14 (17,7)	21 (23,6)	3 (15,8)	24 (22,2)	
Kein virales Risiko	35 (18,7)	18 (22,8)	15 (16,9)	2 (10,5)	17 (15,7)	
HBV	42 (22,5)	20 (25,3)	17 (19,1)	5 (26,3)	22 (20,4)	0,59 ⁴
HCV	31 (16,6)	11 (13,9)	13 (14,6)	7 (36,8)	20 (18,5)	
HBV-/HCV-Koinfektion	41 (21,9)	16 (20,3)	23 (25,8)	2 (10,5)	25 (23,1)	
¹Differenzierungsgrad [n(%)]						
G1	34 (18,6)	16 (20,5)	15 (17,4)	3 (15,8)	18 (17,1)	
G2	102 (55,7)	42 (53,8)	46 (53,5)	14 (73,7)	60 (57,1)	
G3	46 (25,1)	20 (25,6)	24 (27,9)	2 (10,5)	26 (24,8)	0,87 ^{4,5}
G4	1 (0,5)		1 (1,2)		1 (1,0)	
²Stadium der Tumorerkrankung [n(%)]						
I	80 (47,3)	30 (43,5)	41 (50,6)	9 (47,4)	50 (50,0)	
II	44 (26,0)	20 (29,0)	18 (22,2)	6 (31,6)	24 (24,0)	
III	41 (24,3)	16 (23,2)	21 (25,9)	4 (21,1)	25 (25,0)	0,69 ^{4,6}
IV	4 (2,4)	3 (4,3)	1 (1,2)		1 (1,0)	

¹Daten zum Tumordifferenzierungsgrad waren für 183/187 Patienten verfügbar

²Daten zum Stadium der Tumorerkrankungen waren für 169/187 Patienten verfügbar

³ Kruskal-Wallis-Test, getestet wurden Träger des selteneren Allels A gegen die für das häufigere Allel C Homozygoten (*IFNL3* rs4803217 CA/AA vs. CC)

⁴ χ^2 -Text/exakter Fisher-Yates-Test, getestet wurden Träger des selteneren Allels A gegen die für das häufigere Allel C Homozygoten (*IFNL3* rs4803217 CA/AA vs. CC)

⁵ getestet wurde für G1, G2 vs. G3/G4

⁶ getestet wurde für I, II vs. III/IV

n.t. nicht getestet

Tab. 3.5: Analysen zur Assoziation der *IFNL3*-rs28416813-Genotypen mit demografischen und klinischen Daten für PDAC-Patienten

	Total <i>n</i> = 162	CC <i>n</i> = 76	CG <i>n</i> = 69	GG <i>n</i> = 17	CG/GG <i>n</i> = 86	<i>p</i>
¹ Alter bei Diagnose M ± SD	65,4 ± 10,8	65,3 ± 10,5	65,1 ± 10,6	66,8 ± 13,1	65,4 ± 11,0	0,98 ⁵
Median Quartilsabstand[Jahre]	65,0 15,5	65,0 17,0	65,0 12,5	67,0 20,0	65,5 14,3	
Geschlecht m/f [n(m %)]	92(56,8)/70	44(57,9)/32	40(58,0)/29	8(47,1)/9	48(55,8)/38	n.t
² Differenzierungsgrad [n(%)]						
G1	28 (17,4)	10 (13,2)	14 (20,6)	4 (23,5)	18 (21,2)	
G2	84 (52,2)	41 (53,9)	37 (54,4)	6 (35,3)	43 (50,6)	
G3	47 (29,2)	24 (31,6)	16 (23,5)	7 (41,2)	23 (27,1)	0,42 ^{6,7}
G4	2 (1,2)	1 (1,3)	1 (1,5)		1 (1,2)	
³ Stadium der Tumorerkrankung [n(%)]						
I	20 (12,6)	8 (10,8)	10 (14,7)	2 (11,8)	12 (14,1)	
II	130 (81,8)	62 (83,8)	54 (79,4)	14 (82,4)	68 (80,0)	
III	4 (2,5)	1 (1,4)	3 (4,4)		3 (3,5)	0,86 ^{6,8}
IV	5 (3,1)	3 (4,1)	1 (1,5)	1 (5,9)	2 (2,4)	

¹ Altersangaben waren für 159/162 Patienten verfügbar

² Daten zum Tumordifferenzierungsgrad waren für 161/162 Patienten verfügbar

³ Daten zum Stadium der Tumorerkrankungen waren für 159/162 Patienten verfügbar

⁵ Kruskal-Wallis-Test, getestet wurden Träger des selteneren Allels G gegen die für das häufigere Allel C Homozygoten (*IFNL3* rs28416813 CG/GG vs. CC)

⁶ χ^2 -Test/exakter Fisher-Yates-Test, getestet wurden Träger des selteneren Allels G gegen die für das häufigere Allel C Homozygoten (*IFNL3* rs28416813 CG/GG vs. CC)

⁷ getestet wurde für G1, G2 vs. G3/G4

⁸ getestet wurde für I, II vs. III/IV

n.t. nicht getestet

3.4 Analysen zur Tumorprogression

3.4.1 Tumorprogression definiert durch die Dauer des progressionsfreien Intervalls (PFI) und des spezifischen progressionsfreien Intervalls (PFI.2)

Als erste Kriterien zur Analyse der Tumorprogression wurden die Ereigniszeiten PFI und PFI.2 ausgewählt, welche einen wichtigen Aspekt zur Einschätzung der Lebensqualität der Tumorpatienten darstellen.

Das PFI ist als Zeitspanne von der Diagnose bis zur Progression der Tumorerkrankung, bis zum Auftreten eines weiteren Tumors jeglicher Entität oder bis zum Tod durch einen Tumor jeglicher Entität definiert (PFI-Ereignis). Das PFI.2 wird als Zeitspanne von der Diagnose bis zur Progression der ursprünglichen Tumorerkrankung oder bis zum Tod durch selbigen definiert (PFI.2-Ereignis) (ausführlich in Kapitel 2.1.1).

Der Vergleich der Angaben des PFI mit denen des PFI.2 bei den HCC-Patienten zeigt,

dass insgesamt vier Patienten mit dem *IFNL3*-rs4803217-Genotyp CA ein unspezifisches Ereignis erlebten und deswegen bei PFI.2 zensiert wurden (Anhang Tab. A.4). Die weiteren Werte sind zwischen dem PFI und dem PFI.2 nahezu identisch, weswegen nachfolgend, wenn nicht explizit anders erwähnt, nur die Werte für das PFI.2 angegeben werden.

HCC-Patienten mit dem *IFNL3*-rs4803217-Genotyp CA weisen mit 21,0 Monaten das längste mediane PFI.2 auf (Abb. 3.6, Anhang Tab. A.4). Auffällig ist, dass das mediane PFI.2 der für das häufigere Allel C Homozygoten mit 18,4 Monaten zwischen den medianen PFI.2 der Heterozygoten und der Allel-A-Homozygoten liegt (Anhang Tab. A.4). Diese kontraintuitive Verteilung kommt möglicherweise durch die geringe Fallzahl der Allel-A-Homozygoten zustande.

Insgesamt besteht ein geringer Unterschied des medianen PFI.2 zwischen den HCC-Patienten, die Träger des selteneren Allels A sind, und den für das häufigere Allel C Homozygoten, was auf eine geringe Relevanz bzw. auf das Fehlen einer klinischen Relevanz hindeutet.

Diese Annahme konnte durch die grafische Darstellung der Ereigniszeit innerhalb von Kaplan-Meier-Kurven bestätigt werden (Abb. 3.6). Die Kaplan-Meier-Kurven verlaufen alle bis zu dem Zeitpunkt $t = 1,75$ Jahre ungekreuzt eng parallel. Danach schneiden sich die Kurven, wodurch kein proportionales Hazard vorliegt. Erneut fällt auf, dass die Kurve der Patienten, bei denen das häufigere Allel C homozygot vorliegt, zwischen den Kurven der heterozygoten und den Allel A homozygoten Patienten liegt (Abb. 3.6).

Die Patienten mit dem *IFNL3*-rs4803217-Genotyp CC weisen keinen signifikanten Unterschied in der Länge des PFI bzw. des PFI.2 zu den Patienten mit den *IFNL3*-rs4803217-Genotypen CA und AA auf (Log-Rank-Test: $p(\text{PFI}) = 0,83$, $p(\text{PFI.2}) = 0,65$).

Aus der einfachen Cox-Regressionsanalyse wurden die Kriterien Ätiologie, Differenzierungsgrad und Tumorstadium in die multiple Cox-Regressionsanalyse übernommen (Tab. 3.6). Innerhalb der multiplen Cox-Regressionsanalyse ist eine von den anderen getesteten Kriterien unabhängige Assoziation des PFI.2 mit der Ätiologie HCV feststellbar (multiple Cox-Regressionsanalyse: $p = 0,01$) (Tab. 3.6). Hierbei führt das Vorliegen einer HCV-Infektion bei HCC-Patienten zu einem etwa 2,3-fach erhöhten Risiko ein PFI.2-Ereignis zu erleiden (multiple Cox-Regressionsanalyse: HR (95 % KI) = 2,26 (1,20–4,29)).

Des Weiteren zeigt sich die Tendenz zur Assoziation zwischen den Patienten mit dem niedrigsten Differenzierungsgrad G1 im Vergleich zu den Patienten mit einem Differenzierungsgrad G2 und einem längeren PFI.2 (multiple Cox-Regressionsanalyse: $p = 0,08$) (Tab. 3.6). Eine deutlich unabhängige Assoziation wird zwischen den Pati-

enten mit einem höheren Tumorstadium II oder III/IV im Vergleich zu den Patienten mit dem niedrigsten Tumorstadium I und einem kürzeren PFI.2 deutlich (multiple Cox-Regressionsanalyse: $p(\text{II vs. I}) = 0,04$, $p(\text{III/IV vs. I}) < 0,01$) (Tab. 3.6). Hierbei erhöht sich das Risiko, ein Ereignis zu erleiden, um das bis zu 2-fache für die Patienten mit einem höheren Stadium (Tab. 3.6). Zu beachten ist, dass bei dem Kriterium Stadium der Tumorerkrankung die Testannahme des proportionalen Hazards verletzt wurde (Schoenfeld Residuum: $p = 0,02$) (Tab. 3.6). Die Einflussstärken der in der multiplen Analyse ermittelten signifikanten unabhängigen Assoziationsgrößen (HCV-Infektion, Tumordifferenzierungsgrad und Stadium des Tumors) wurden in Kaplan-Meier-Kurven grafisch dargestellt (Anhang Abb. A.6).

Bei dem Vergleich der Angaben des PFI mit denen des PFI.2 bei den PDAC-Patienten erlebten zwei Patienten mit dem *IFNL3*-rs28416813-Genotyp CC bei dem PFI ein unspezifisches Ereignis und wurden deswegen bei dem PFI.2 zensiert (Anhang Tab. A.5). Die weitere Ereignisverteilung und auch die medianen progressionsfreien Intervalle, PFI bzw. PFI.2, sind nahezu identisch und werden deswegen in der nachfolgenden Beschreibung, wenn nicht explizit anders erwähnt, immer nur für das PFI.2 angegeben (Anhang Tab. A.5).

Das mediane PFI.2 der PDAC-Patienten mit dem *IFNL3*-rs28416813-Genotyp CC ist mit 12,8 Monaten um 6,7 Monate kürzer als bei den Patienten, die Träger des selteneren *IFNL3*-rs28416813-Allels G sind (CG/GG) (Anhang Tab. A.5).

Durch die grafische Darstellung der Ereigniszeiten PFI und PFI.2 innerhalb von Kaplan-Meier-Kurven konnte dieser beschriebene Unterschied veranschaulicht werden. Alle drei Kurven verlaufen stetig sinkend, bis sie ab etwa $t = 2$ Jahre nahezu parallel verlaufen (Abb. 3.7). Hierbei muss erneut darauf hingewiesen werden, dass die Kurve der Patienten, die homozygot für das *IFNL3*-rs28416813-Allel G sind, die beiden anderen Kurven schneidet und somit kein proportionales Hazard vorliegt.

Die beiden Kurven der Patienten, die Träger des selteneren *IFNL3*-rs28416813-Allels G (CG/GG) sind, liegen ab $t = 0,5$ Jahre über der Kurve der Allel-C-Homozygoten (Abb. 3.7). Die Länge des PFI und des PFI.2 unterscheidet sich signifikant zwischen den Allel-C-Homozygoten und den Trägern des selteneren Allels G (*IFNL3*-rs28416813-Genotypen CC vs. CG/GG) (Log-Rank-Test: $p(\text{PFI}) < 0,01$, $p(\text{PFI.2}) = 0,01$).

Aus der einfachen Cox-Regressionsanalyse wurden die Kriterien Differenzierungsgrad, Tumorstadium und *IFNL3*-rs28416813-Genotyp in die multiple Cox-Regressionsanalyse übernommen (Tab. 3.7). Innerhalb der multiplen Cox-Regressionsanalyse zeigt sich ein von den anderen untersuchten Kriterien unabhängiger Zusammenhang zwischen den Patienten mit dem niedrigsten Tumorstadium I im Vergleich zu den Patienten mit dem

höheren Stadium II und einem längeren PFI.2 (Tab. 3.7). Das Risiko, ein PFI.2-Ereignis zu erfahren, steigt um das bis zu 2,7-fache bei den PDAC-Patienten im Stadium II (multiple Cox-Regerssionsanalyse: $p = 0,03$, HR (95 % KI) = 0,37 (0,15–0,89)) (Tab. 3.7). Die Verletzung der Proportionalität des Hazards muss in diesem Fall beachtet werden (Schoenfeld Residuum: $p = 0,05$) (Tab. 3.7).

Außerdem konnte gezeigt werden, dass auch der Polymorphismus *IFNL3* rs28416813 unabhängig von den anderen untersuchten Kriterien mit dem PFI.2 assoziiert (multiple Cox-Regressionsanalyse: $p = 0,02$) (Tab. 3.7). Hierbei korreliert der Genotyp des *IFNL3* rs28416813, der die genetische Disposition zur Expression eines funktionsfähigen IFN- λ_4 -Proteins spiegelt, mit einem verlängerten PFI.2. Das Vorliegen des selteneren Allels G, das stellvertretend für die Fähigkeit IFN- λ_4 exprimieren zu können steht, senkt die Wahrscheinlichkeit der PDAC-Patienten ein Ereignis zu erleiden um 39 % im Vergleich zu den Patienten, die für das häufigere Allel C bzw. für die *knock-out*-Variante des IFN- λ_4 homozygot sind (multiple Cox-Regressionsanalyse: HR (95 % KI) = 0,61 (0,40–0,93)) (Tab. 3.7).

Die Einflussstärken der in der multiplen Analyse ermittelten signifikanten unabhängigen Assoziationsgrößen (Tumordifferenzierungsgrad (tendenziell), Stadium des Tumors und *IFNL3*-rs28416813-Genotyp) wurden in Kaplan-Meier-Kurven grafisch dargestellt (Anhang Abb. A.7).

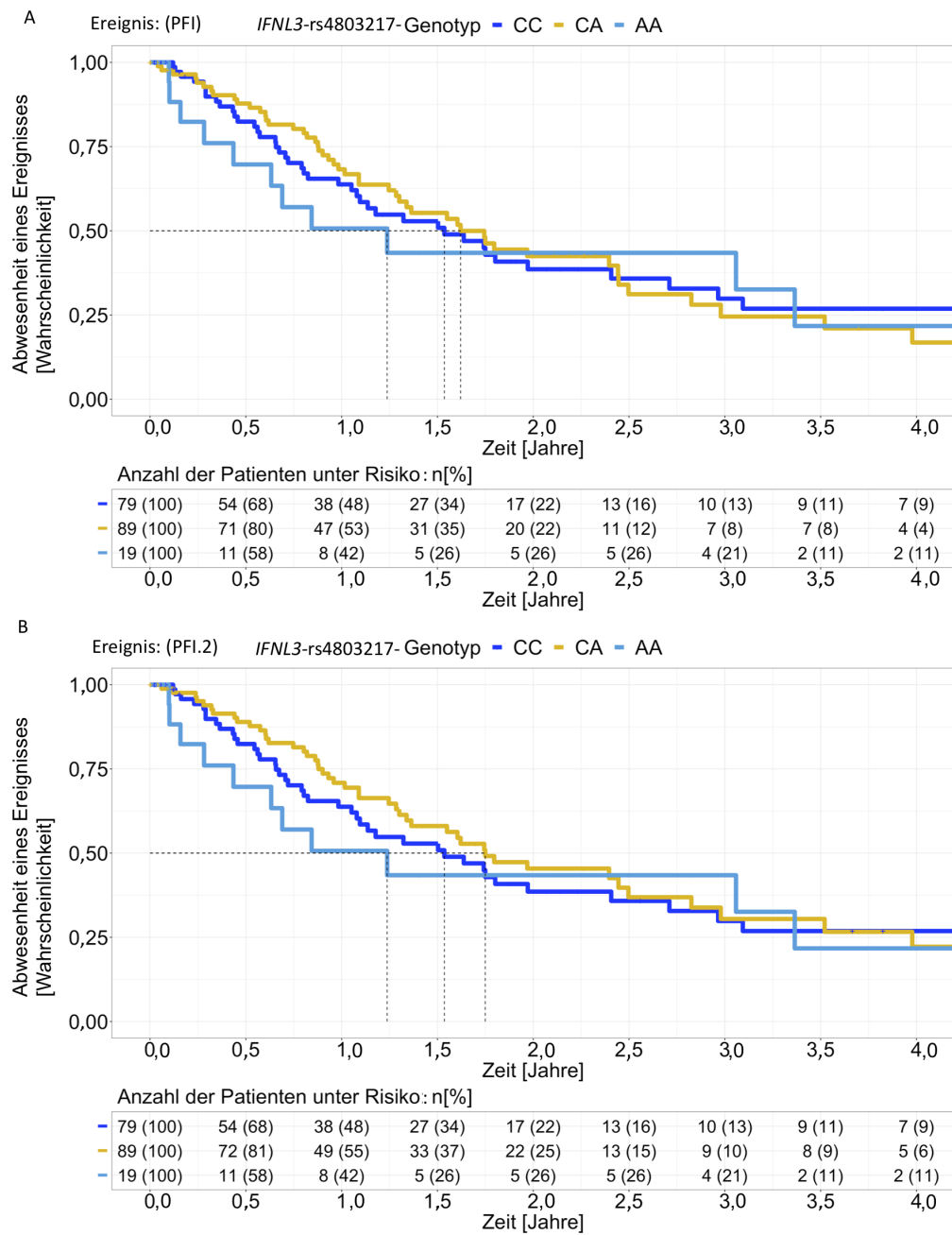


Abb. 3.6: Ereigniszeitanalyse der progressionsfreien Intervalle nach dem Kaplan-Meier-Verfahren für HCC-Patienten.

Die Patienten europäischer Herkunft des TCGA-LIHC-Kollektivs wurden nach *IFNL*-Genotypen (hier: *IFNL3*-rs4803217-Genotyp CC, CA, AA) stratifiziert. Dargestellt sind das progressionsfreie Intervall (PFI) (A) und das spezifische progressionsfreie Intervall (PFI.2) (B). Die gestrichelten Linien entsprechen den medianen progressionsfreien Intervallen. Unterhalb der Kurven sind zusätzlich die absolute und relative Anzahl der Patienten unter Risiko angeben.

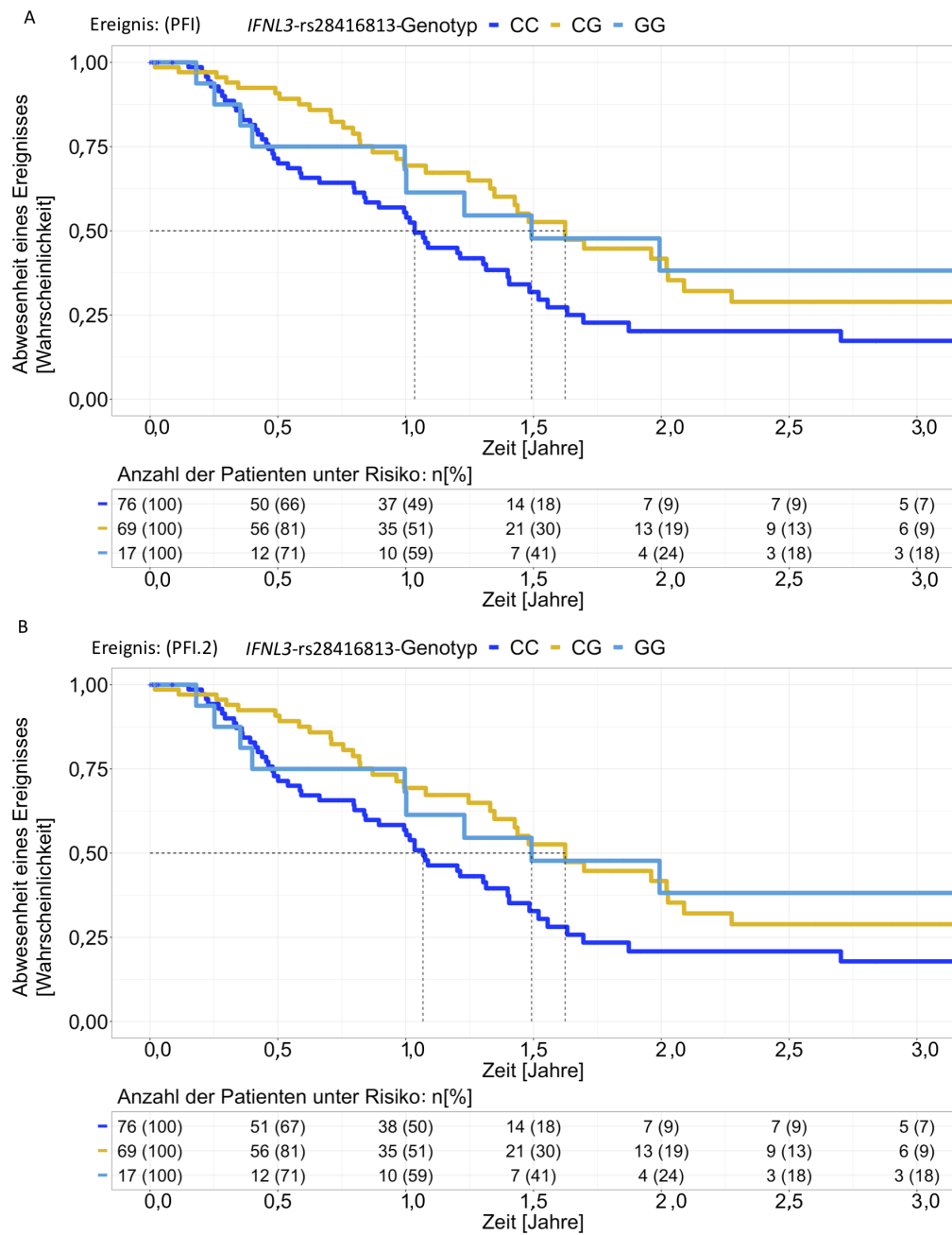


Abb. 3.7: Ereigniszeitanalyse der progressionsfreien Intervalle nach dem Kaplan-Meier-Verfahren für PDAC-Patienten.

Die Patienten europäischer Herkunft des TCGA-PAAD-Kollektivs wurden nach *IFNL*-Genotypen (hier: *IFNL3*-rs28416813-Genotyp CC, CG, GG) stratifiziert. Dargestellt sind das progressionsfreie Intervall (PFI) (A) und das spezifische progressionsfreie Intervall (PFI.2) (B). Die gestrichelten Linien entsprechen den medianen progressionsfreien Intervallen. Unterhalb der Kurven sind zusätzlich die absolute und relative Anzahl der Patienten unter Risiko angeben.

3.4.2 Tumorprogression definiert durch die Gesamtüberlebensdauer (OS)

Als ein weiteres Kriterium zur Analyse der Tumorprogression wurde die Ereigniszeit OS mit dem Ereignis Tod betrachtet. Die OS ist definiert als die Zeitspanne zwischen der Diagnose und dem Tod. Dieses Kriterium bringt einen weiteren Zeitaspekt in die Analysen.

Der mediane Beobachtungszeitraum der HCC-Patienten europäischer Herkunft beträgt 22,1 Monate, in welchem 41,2 % (77/187) der Patienten verstarben und 58,8 % (110/187) der Patienten zensiert wurden, weil das Ereignis im Beobachtungszeitraum nicht eintrat (Anhang Tab. A.4). Die mediane Überlebensdauer ist bei den HCC-Patienten mit dem *IFNL3*-rs4803217-Genotyp CC mit 46,8 Monaten am höchsten (Abb. 3.8, Anhang Tab. A.4). Alle Kaplan-Meier-Kurven verlaufen relativ flach und überschneiden sich erstmalig etwa ab dem Zeitpunkt $t = 2,5$ Jahre (Abb. 3.8). Das bedeutet, dass anschließende Analysen aufgrund der Missachtung des proportionalen Hazards nur eingeschränkt beurteilbar sind. Auffällig ist, dass die Kurve der Allel-C-Homozygoten zwischen den Kurven der Heterozygoten und Allel-A-Homozygoten liegt (Abb. 3.8). Vermutlich liegt dieser kontraintuitive Kurvenverlauf an der geringen Fallzahl der Allel-A-Homozygoten. Es zeigt sich kein Hinweis auf eine Korrelation des Genotyps mit der OS. Die Analysen zur Assoziation wurden nur für die Träger des selteneren Allels A, die die genetische Disposition zur Expression eines funktionstüchtigen IFN- λ_4 spiegeln, gegenüber denen für das häufigere Allel C Homozygoten durchgeführt. Innerhalb der OS bei den Patienten mit dem *IFNL3*-rs4803217-Genotyp CC konnte kein signifikanter Unterschied zu den Trägern des Allels A (CA/AA) gezeigt werden (Log-Rank-Test: $p = 0,87$).

Aus der einfachen Cox-Regressionsanalyse wurden die Kriterien Alter und Differenzierungsgrad in die multiple Cox-Regressionsanalyse übernommen (Tab. 3.6). Ein von den anderen betrachteten Kriterien unabhängiger, signifikanter Zusammenhang mit der OS ist innerhalb der multiplen Cox-Regressionsanalyse für das Alter feststellbar, wobei ein höheres Alter mit einer kürzeren OS assoziiert ist (multiple Cox-Regressionsanalyse: $p = 0,02$) (Tab. 3.6). Diese Korrelation scheint zunächst klinisch nur schwach relevant (95 % KI = 1,00–1,04). Da das Alter aber als stetiger Wert modelliert wurde, wird die Risikozunahme pro Altersjahr betrachtet. Deswegen ist auch eine durch das KI relativ schwach wirkende Korrelation trotzdem klinisch relevant. Des Weiteren besteht eine Tendenz zu einer unabhängigen Korrelation des Differenzierungsgrads mit der OS, wobei das Vorliegen des Differenzierungsgrads G2 bei HCC-Patienten im Vergleich zu Patienten mit einem Differenzierungsgrad G1 das Sterberisiko nahezu verdoppeln würde (multiple Cox-Regressionsanalyse: $p = 0,08$, HR = 0,53). Die klinische Bedeutung dieser Beobachtung ist aber nur eingeschränkt relevant (95 % KI = 0,27–1,07) (Tab. 3.6).

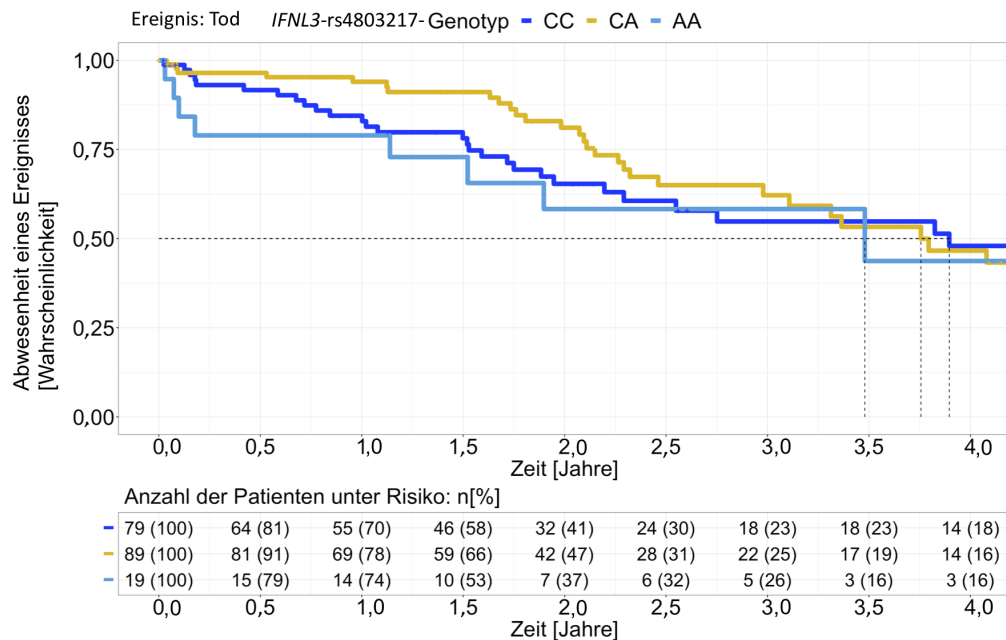


Abb. 3.8: Ereigniszeitanalyse der Gesamtüberlebensdauer (OS) nach dem Kaplan-Meier-Verfahren für HCC-Patienten.

Die Patienten europäischer Herkunft des TCGA-LIHC-Kollektivs wurden nach *IFNL*-Genotypen (hier: *IFNL3*-rs4803217-Genotyp CC, CA, AA) stratifiziert. Die gestrichelten Linien entsprechen den medianen Überlebenszeiten. Unterhalb der Kurven sind zusätzlich die absolute und relative Anzahl der Patienten unter Risiko angeben.

Bei den PDAC-Patienten beträgt der mediane Beobachtungszeitraum für das gesamte Kollektiv 15,3 Monate, in welchem 53,1 % (86/162) der Patienten verstarben und 46,9 % (76/162) der Patienten zensiert wurden, da das Ereignis im Beobachtungszeitraum nicht eintrat (Anhang Tab. A.5).

Die mediane Überlebenszeit der PDAC-Patienten mit dem *IFNL3*-rs28416813-Genotyp CC liegt mit 17,0 Monaten unter den medianen Überlebenszeiten von den Patienten mit den *IFNL3*-rs28416813-Genotypen CG bzw. GG (24,3 bzw. 20,2 Monate) (Abb. 3.9, Anhang Tab. A.5). Diese Tendenz wurde bereits bei der Betrachtung der medianen progressionsfreien Intervalle, PFI bzw. PFI.2, erkannt (Kapitel 3.4.1) und setzt sich bei den medianen Überlebenszeiten fort.

Die Kaplan-Meier-Kurven verlaufen stetig sinkend und am Anfang schon fast identisch übereinander, wobei die Kurve der Patienten mit dem *IFNL3*-rs28416813-Genotyp CC unter den Kurven der Träger des selteneren Allels G (CG/GG) verläuft (Abb. 3.9).

Ab dem Zeitpunkt $t = 1,25$ Jahre trennen sich die Kurven und die Kurve der Allel-C-Homozygoten sinkt steiler als die Kurve der Träger des Allels G bis alle drei Kurven ab der Zeit $t = 2$ Jahre nahezu parallel verlaufen (Abb. 3.9). Zu beachten ist, dass die

Kurve der für das seltenere Allel G Homozygoten die beiden anderen Kurven schneidet, wodurch das proportionale Hazard nicht vorliegt (Abb. 3.9). Dies muss bei allen weiteren Analysen beachtet werden.

Anhand der Betrachtung der medianen Überlebenszeiten und der Kaplan-Meier-Kurven konnte deskriptiv ein Unterschied zwischen den Patienten, die durch das Vorliegen des selteneren *IFNL3*-rs28416813-Allels G befähigt sind, ein funktionsfähiges IFN- λ_4 zu exprimieren, und den für das häufigere Allel C Homozygoten erkannt werden. Die Berechnung des Log-Rank-Tests ergab eine Tendenz zur signifikanten Assoziation (Log-Rank-Test: $p = 0,05$).

Aus der einfachen Cox-Regressionsanalyse wurden die Kriterien Alter, Differenzierungsgrad, Tumorstadium und *IFNL3*-rs28416813-Genotyp in die multiple Cox-Regressionsanalyse übernommen (Tab. 3.7). Innerhalb der multiplen Cox-Regressionsanalyse bei den PDAC-Patienten konnten für die ausgewählten Kriterien zum Teil Tendenzen des Zusammenhangs mit der OS ermittelt werden (Tab. 3.7). Das Alter spielt auch bei PDAC-Patienten eine Rolle für die Tumorprogression (NCI 2019). Jedoch konnte die erwartete Alterskorrelation in dieser Arbeit wahrscheinlich aufgrund der geringeren Fallzahl und Ereigniszahl nicht gezeigt werden.

Der Zusammenhang zwischen der OS und den Patienten mit dem Tumorstadium II, die ein 2,5-fach erhöhtes Sterberisiko aufweisen im Vergleich zu Patienten mit einem Tumorstadium I, zeigt in der multiplen Cox-Regressionsanalyse die Tendenz zu einem signifikanten Zusammenhang (multiple Cox-Regressionsanalyse: $p = 0,06$). Jedoch ist die klinische Relevanz durch das 95 % KI und die Verletzung des proportionalen Hazards nur eingeschränkt beurteilbar (multiple Cox-Regressionsanalyse: HR (95 % KI) = 0,40 (0,16–1,04), Schoenfeld Residuum: $p = 0,01$) (Tab. 3.7).

Zusätzlich zeigt sich eine Tendenz zur Assoziation der Patienten, die durch die Ausprägung der *IFNL3*-rs28416813-Genotypen CG bzw. GG ein funktionsfähiges IFN- λ_4 exprimieren können, mit einer verlängerten OS und einem um 32 % reduzierten Sterberisiko im Vergleich zu den für das häufigere Allel C Homozygoten (multiple Cox-Regressionsanalyse: $p = 0,09$, HR (95 % KI) = 0,68 (0,44–1,07)) (Tab. 3.7).

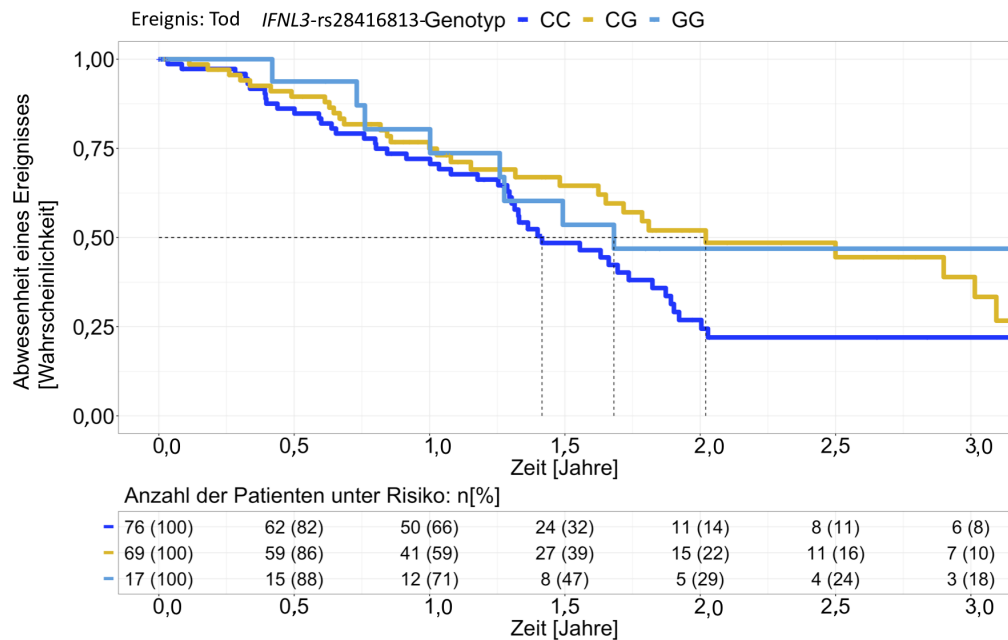


Abb. 3.9: Ereigniszeitanalyse der Gesamtüberlebensdauer (OS) nach dem Kaplan-Meier-Verfahren für PDAC-Patienten.

Die Patienten europäischer Herkunft des TCGA-PAAD-Kollektivs wurden nach *IFNL3*-Genotypen (hier: *IFNL3*-rs28416813-Genotyp CC, CG, GG) stratifiziert. Die gestrichelten Linien entsprechen den medianen Überlebenszeiten. Unterhalb der Kurven sind zusätzlich die absolute und relative Anzahl der Patienten unter Risiko angeben.

Bei den Kollektiven der HCC- und PDAC-Patienten konnte wenn auch zum Teil nur andeutungsweise gezeigt werden, dass ein höheres Alter, ein höherer Differenzierungsgrad und ein höheres Stadium der Tumorerkrankung im Zusammenhang mit einer Verkürzung des PFI bzw. des PFI.2 und der OS stehen (Tab. 3.6, Tab. 3.7). Der Differenzierungsgrad und das Stadium der Tumorerkrankung sind unabhängige prognostische Faktoren, die jedoch stark positiv miteinander korrelieren (Derwinger et al. 2010). Wenn eines der beiden Kriterien eine signifikante Assoziation innerhalb eines multiplen Cox-Regressionsmodells aufweist, ist zumeist das jeweils andere Kriterium nicht mehr signifikant, da es keine zusätzlichen Informationen liefert.

Tab. 3.6: Übersicht über die einfache und die multiple Cox-Regressionsanalyse für die drei Ereigniszeiten PFI/PFI.2/OS der HCC-Patienten

	Einfache Analyse		Multiple Analyse		S.R. ⁴
	HR (95 % KI)	<i>p</i>	HR (95 % KI)	<i>p</i>	<i>p</i>
¹Progressionsfreies Intervall (PFI)					
Alter	0,99 (0,98-1,00)	0,16			
Geschlecht (m/f)	0,94 (0,64-1,39)	0,75			
Ätiologie (Kein Risiko vs. HBV)	1,01 (0,54-1,89)	0,98	0,82 (0,41-1,66)	0,59	
Ätiologie (Kein vir. Risiko vs. HBV)	1,00 (0,53-1,87)	1,00	0,88 (0,44-1,73)	0,71	
Ätiologie (HCV vs. HBV)	1,92 (1,07-3,43)	0,03*	2,26 (1,20-4,27)	0,01*	0,74
Ätiologie (HBV/HCV vs. HBV)	1,15 (0,65-2,02)	0,63	1,23 (0,68-2,23)	0,50	
Differenzierungsgrad (G1 vs. G2)	0,60 (0,34-1,05)	0,07(*)	0,62 (0,33-1,16)	0,14	0,28
Differenzierungsgrad (G3/G4 vs. G2)	1,23 (0,79-1,93)	0,36	1,33 (0,83-2,13)	0,23	
Tumorstadium (II vs. I)	1,67 (1,01-2,77)	0,05*	1,64 (0,98-2,74)	0,06(*)	0,02
Tumorstadium (III/IV vs. I)	1,72 (1,05-2,81)	0,03*	1,81 (1,09-3,01)	0,02*	
<i>IFNL3</i> rs4803217 (CA/AA vs. CC)	0,98 (0,66-1,45)	0,92			
²Spezifisches progressionsfreies Intervall (PFI.2)					
Alter	0,99 (0,98-1,00)	0,15			
Geschlecht (m/f)	0,87 (0,58-1,29)	0,48			
Ätiologie (Kein Risiko vs. HBV)	0,89 (0,47-1,70)	0,73	0,68 (0,33-1,39)	0,29	
Ätiologie (Kein vir. Risiko vs. HBV)	1,03 (0,55-1,94)	0,93	0,88 (0,44-1,74)	0,71	
Ätiologie (HCV vs. HBV)	1,95 (1,08-3,50)	0,03*	2,26 (1,20-4,29)	0,01*	0,91
Ätiologie (HBV/HCV vs. HBV)	1,07 (0,60-1,92)	0,82	1,14 (0,62-2,11)	0,67	
Differenzierungsgrad (G1 vs. G2)	0,54 (0,30-0,97)	0,04*	0,54 (0,28-1,07)	0,08(*)	0,14
Differenzierungsgrad (G3/G4 vs. G2)	1,14 (0,73-1,80)	0,56	1,25 (0,78-2,02)	0,35	
Tumorstadium (II vs. I)	1,78 (1,06-3,00)	0,03*	1,73 (1,01-2,93)	0,04*	0,02
Tumorstadium (III/IV vs. I)	1,96 (1,19-3,24)	0,01*	2,12 (1,26-3,55)	<0,01*	
<i>IFNL3</i> rs4803217 (CA/AA vs. CC)	1,10 (0,74-1,63)	0,65			
³Gesamtüberlebensdauer (OS)					
Alter	1,03 (1,01-1,05)	0,02*	1,02 (1,00-1,04)	0,02*	0,93
Geschlecht (m/f)	1,10 (0,70-1,73)	0,67			
Ätiologie (Kein Risiko vs. HBV)	1,70 (0,89-3,25)	0,11			
Ätiologie (Kein vir. Risiko vs. HBV)	1,69 (0,84-3,41)	0,14			
Ätiologie (HCV vs. HBV)	1,50 (0,68-3,29)	0,31			
Ätiologie (HBV/HCV vs. HBV)	1,13 (0,57-2,25)	0,72			
Differenzierungsgrad (G1 vs. G2)	0,50 (0,25-1,00)	0,05*	0,53 (0,27-1,07)	0,08(*)	0,25
Differenzierungsgrad (G3/G4 vs. G2)	0,91 (0,53-1,55)	0,72	1,01 (0,58-1,75)	0,96	
Tumorstadium (II vs. I)	0,94 (0,50-1,79)	0,85			
Tumorstadium (III/IV vs. I)	1,31 (0,75-2,26)	0,34			
<i>IFNL3</i> rs4803217 (CA/AA vs. CC)	1,11 (0,71-0,74)	0,66			

¹PFI-Progressionsfreies Intervall: Zeitspanne von der Diagnose bis zur Progression der Tumorerkrankung, bis zum Auftreten eines weiteren Tumors jeglicher Entität oder bis zum Tod durch einen Tumor jeglicher Entität

²PFI.2-Spezifisches progressionsfreies Intervall: Zeitspanne von der Diagnose bis zur Progression der Tumorerkrankung oder bis zum Tod durch selbige Tumorentität

³OS-Gesamtüberlebensdauer: Zeitspanne von der Diagnose bis zum Tod durch jegliche Ursache

⁴S.R. Schoenfeld Residuum

*signifikanter p-Wert ($p \leq 0,05$)

(*) Tendenz zum signifikanten p-Wert ($p \leq 0,1$)

¹Für 167 HCC Patienten standen die vollständigen Angaben der Kriterien zur Verfügung und ermöglichten die Cox-Regressionsanalyse

Tab. 3.7: Übersicht über die einfache und die multiple Cox-Regressionsanalyse für die drei Ereigniszeiten PFI/PFI.2/OS der PDAC-Patienten

	Einfache Analyse		Multiple Analyse		S.R. ⁴
	HR (95 % KI)	<i>p</i>	HR (95 % KI)	<i>p</i>	<i>p</i>
¹Progressionsfreies Intervall (PFI)					
Alter	1,01 (0,99-1,03)	0,28			
Geschlecht (m/f)	1,15 (0,77-1,72)	0,51			
Differenzierungsgrad (G1 vs. G2)	0,58 (0,31-1,11)	0,10(*)	0,95 (0,49-1,82)	0,87	0,23
Differenzierungsgrad (G3/G4 vs. G2)	1,49 (0,96-2,32)	0,08(*)	1,25 (0,80-1,96)	0,33	
Tumorstadium (I vs. II)	0,30 (0,13-0,69)	0,01*	0,35 (0,14-0,84)	0,02*	0,04
Tumorstadium (III/IV vs. II)	0,97 (0,45-2,11)	0,94	1,06 (0,48-2,32)	0,89	
<i>IFNL3</i> rs28416813 (CG/GG vs. CC)	0,58 (0,39-0,87)	0,01*	0,58 (0,38-0,88)	0,01*	0,56
²Spezifisches progressionsfreies Intervall (PFI.2)					
Alter	1,01 (0,99-1,03)	0,27			
Geschlecht (m/f)	1,13 (0,75-1,70)	0,56			
Differenzierungsgrad (G1 vs. G2)	0,63 (0,33-1,20)	0,16	0,99 (0,51-1,90)	0,97	0,24
Differenzierungsgrad (G3/G4 vs. G2)	1,57 (1,01-2,45)	0,05*	1,34 (0,85-2,12)	0,20	
Tumorstadium (I vs. II)	0,32 (0,14-0,74)	0,01*	0,37 (0,15-0,89)	0,03*	0,05
Tumorstadium (III/IV vs. II)	0,99 (0,46-2,16)	0,98	1,09 (0,50-2,39)	0,83	
<i>IFNL3</i> rs28416813 (CG/GG vs. CC)	0,60 (0,40-0,91)	0,02*	0,61 (0,40-0,93)	0,02*	0,60
³Gesamtüberlebensdauer (OS)					
Alter	1,03 (1,01-1,05)	0,01*	1,01 (1,00-1,04)	0,13	0,67
Geschlecht (m/f)	1,26 (0,83-1,93)	0,28			
Differenzierungsgrad (G1 vs. G2)	0,76 (0,24-1,00)	0,05*	0,76 (0,60-1,60)	0,47	0,24
Differenzierungsgrad (G3/G4 vs. G2)	1,35 (0,85-2,13)	0,20	1,13 (0,70-1,81)	0,61	
Tumorstadium (I vs. II)	0,37 (0,16-0,86)	0,02*	0,40 (0,16-1,04)	0,06(*)	0,01
Tumorstadium (III/IV vs. II)	0,90 (0,36-2,23)	0,82	0,95 (0,38-2,40)	0,92	
<i>IFNL3</i> rs28416813 (CG/GG vs. CC)	0,66 (0,43-1,00)	0,05*	0,68 (0,44-1,07)	0,09(*)	0,66

¹PFI-Progressionsfreies Intervall: Zeitspanne von der Diagnose bis zur Progression der Tumorerkrankung, bis zum Auftreten eines weiteren Tumors jeglicher Entität oder bis zum Tod durch einen Tumor jeglicher Entität

²PFI.2-Spezifisches progressionsfreies Intervall: Zeitspanne von der Diagnose bis zur Progression der Tumorerkrankung oder bis zum Tod durch selbige Tumorentität

³OS-Gesamtüberlebensdauer: Zeitspanne von der Diagnose bis zum Tod durch jegliche Ursache

⁴S.R. Schoenfeld Residuum

*signifikanter p-Wert ($p \leq 0,05$)

(*) Tendenz zum signifikanten p-Wert ($p \leq 0,1$)

¹Für 158 PDAC Patienten standen die vollständigen Angaben der Kriterien zur Verfügung und ermöglichten die Cox-Regressionsanalyse

3.4.3 Zusammenfassung der Assoziationen mit Formulierung einer möglichen *Follow-Up*-Kontrolle

Bei Betrachtung des PFI oder des PFI.2 erlebten jeweils um die 55 % der HCC- und PDAC-Patienten ein PFI- bzw. PFI.2-Ereignis und etwa 40 % der HCC- und 50 % der PDAC-Patienten haben bei Betrachtung der OS ein Ereignis im Sinne des Todes erlitten (Anhang Tab. A.4, Anhang Tab. A.5).

Innerhalb des Beobachtungszeitraums der progressionsfreien Intervalle, PFI bzw. PFI.2, zeigte sich bei HCV-bedingtem HCC ein erhöhtes Risiko, eine HCC-Progression zu erleiden, im Vergleich zu den Patienten mit HBV-bedingtem HCC (Tab. 3.6). Des Weiteren zeigte sich die Tendenz einer positiven Korrelation sowohl zwischen dem Differenzierungsgrad des Tumors als auch dem Tumorstadium und dem Risiko, eine HCC-Progression zu erleiden (Tab. 3.6). Für die HCC-Patienten konnte keine Assoziation zwischen dem PFI bzw. PFI.2 und den *IFNL3*-rs4803217-Genotypen CC oder CA/AA detektiert werden (Tab. 3.6).

Bei den HCC-Patienten stellte sich innerhalb des Beobachtungszeitraums der OS eine positive Assoziation zwischen dem Alter der Patienten und dem Sterberisiko heraus. Außerdem zeigte sich eine Verdopplung des Sterberisikos der HCC-Patienten mit dem Differenzierungsgrad G2 gegenüber Patienten mit dem Differenzierungsgrad G1 (Tab. 3.6). Für die HCC-Patienten konnte keine Assoziation zwischen der OS und den *IFNL3*-rs4803217-Genotypen CC oder CA/AA ermittelt werden (Tab. 3.6).

Innerhalb des Beobachtungszeitraums der progressionsfreien Intervalle, PFI bzw. PFI.2, zeigte sich signifikant die positive Korrelation zwischen dem Tumorstadium der PDAC-Patienten und der Progression der Tumorerkrankung. Für die PDAC-Patienten ergab sich eine signifikante Assoziation zwischen der Progression und den *IFNL3*-rs28416813-Genotypen CC oder CG/GG. Hierbei sinkt das Progressionsrisiko bei Patienten, die das seltenere Allel G ausprägen und damit befähigt sind, ein funktionsfähiges IFN- λ_4 zu exprimieren (Tab. 3.7).

Bei den PDAC-Patienten stellte sich innerhalb des Beobachtungszeitraums der OS eine positive Korrelation zwischen dem Tumorstadium und dem Sterberisiko heraus (Tab. 3.7). Außerdem konnte bei den PDAC-Patienten die Tendenz einer Assoziation zwischen der OS und den *IFNL3*-rs28416813-Genotypen CC oder CG/GG ermittelt werden. Hierbei scheinen auch die Patienten, die Träger des selteneren Allels G sind, einen Vorteil im Überleben zu haben (Tab. 3.7).

Zum Zeitpunkt der Analysen innerhalb dieser Arbeit waren in beiden TCGA-Kollektiven ≤ 54 % der Patienten verstorben. Die in dieser Arbeit beobachteten Korrelationen und

Tendenzen ließen sich auf Grundlage des kontinuierlichen *Follow-ups* der Patienten durch TCGA zu einem späteren Zeitpunkt überprüfen. Dieses kontinuierliche *Follow-up* der Patienten durch TCGA birgt somit die Möglichkeit, bei Vorliegen einer höheren Zahl an Ereigniseintritten die Assoziationen erneut zu analysieren.

Nach Informationen des *GDC Help Desk* werden die Daten zum *Follow-up* kontinuierlich ergänzt, jedoch unregelmäßig, weswegen ein Zeitpunkt für eine nächste Analyse nicht festgelegt werden kann und die aktuellen Versionshinweise innerhalb des TCGA verfolgt werden müssten.

Um die in der vorliegenden Arbeit erhobenen Analysen im Hinblick auf ihre Verwendbarkeit als prognostische Parameter zu stützen, muss ihre Güte, abgesehen von den oben genannten Überlegungen, an unabhängigen und im besten Fall prospektiven Daten validiert werden (Bickeböller und Fischer 2007, S.63).

4 Diskussion

Die vorliegende Arbeit zeigt einen signifikanten Zusammenhang der progressionsfreien Intervalle, PFI bzw. PFI.2, von Patienten mit PDAC und Polymorphismen im Bereich der *IFNL*-Gene. Wird die OS betrachtet, zeigt sich die Tendenz einer Assoziation. Das Vorliegen von Varianten der Polymorphismen, die sich im nahezu kompletten und perfekten LD mit einer Variante befinden, die für ein funktionsfähiges IFN- λ_4 codiert, erweisen sich als prognostisch günstiger als die Varianten, die das Verschieben des Leserasters zu einem funktionalen *knock-out* des IFN- λ_4 spiegeln.

Für die HCC-Patienten dagegen konnte kein Zusammenhang zwischen den Ereigniszeiten (PFI, PFI.2, OS) und genetischen Polymorphismen in *IFNL*-Genen festgestellt werden. Auch aufgrund geringer Fallzahlen der für die selteneren Allele homozygoten HCC- und PDAC-Patienten, multiplen Testens, das nicht komplett korrigiert wurde, sowie durch teilweise nur eingeschränkt erfüllte Bedingungen (zum Beispiel des proportionalen Hazards) sind weitere Analysen an unabhängigen und gegebenenfalls größeren Kollektiven für die Validierung der Aussagen erforderlich.

Die medianen progressionsfreien Intervalle, PFI und PFI.2, sowie die medianen Überlebenszeiten für die in dieser Arbeit betrachteten TCGA-Kollektive, d. h. HCC- und PDAC-Patienten europäischer Herkunft, sind vergleichbar mit Angaben aus der Literatur. Jedoch sind oft die Rahmenbedingungen der Studien nicht vergleichbar. Die progressionsfreien Intervalle bzw. die Überlebenszeiten werden entscheidend von den klinischen Kriterien der Patienten wie dem Tumordifferenzierungsgrad oder dem Stadium der Tumorerkrankung, der Therapierbarkeit und auch der Länge des *Follow-ups* bestimmt. Deutlich sind hierbei die längsten Überlebenszeiten bei chirurgisch behandelten Patienten mit einem niedrigen Differenzierungsgrad und einem niedrigen Tumorstadium erkennbar (Greten et al. 2005; Desai et al. 2015; Adamska et al. 2017; Tannus et al. 2018; Moletta et al. 2019; Wang und Li 2019).

Diese Arbeit hat sich auf die Betrachtung der europäischen Population beschränkt. Aus diesem Grund können über die Bedeutung der Polymorphismen in den *IFNL*-Genen auf die Tumorprogression anderer Populationen keine Aussagen getroffen werden.

Es sei darauf hingewiesen, dass der TCGA-CDR-Datensatz die Nutzung der Daten zu Analyse Zwecken für die gesamten Kollektive validiert und gesichert hat, nicht aber für einzelne Ethnizitäten. Diese Aufteilung nach Ethnizitäten kann zu Fehlern führen, was im Umfang dieser Arbeit nicht weiter betrachtet werden konnte. Des Weiteren sorgt die integrative Analyse der Ereigniszeiten PFI, PFI.2 und OS mit den genetischen Merkmalen des Wirts dafür, dass die Patientenkollektive europäischer Herkunft weiter stra-

tifiziert wurden, wofür auch keine Validierung innerhalb der TCGA-CDR-Datensätze durchgeführt wurde. Hierdurch kann es zur Kompromittierung statistisch signifikanter Überlebensunterschiede kommen. Die Validierung der Konsistenz war innerhalb dieser Arbeit nicht möglich.

Die in dieser Arbeit betrachteten Kriterien üben jeweils einen unterschiedlich starken Einfluss auf die Progression der beiden Tumorentitäten HCC und PDAC aus. Hierbei ist darauf zu achten, dass weitere, in dieser Arbeit nicht berücksichtigte Kriterien die progressionsfreien Intervalle und das Überleben ebenfalls beeinflussen. Dadurch könnten die Ergebnisse dieser Arbeit in einer nicht abschätzbaren Weise moduliert worden sein.

4.1 Assoziation zwischen den Typ-III-IFN-Genen und der PDAC-Progression

Die in dieser Arbeit aufgezeigte Assoziation zwischen den *IFNL*-Polymorphismen und der PDAC-Progression wurde bis jetzt noch nicht beschrieben oder publiziert. Die Frequenz des IFN- λ_4 -generierenden *IFNL4*-rs368234815-Allels ΔG zeigt globale Unterschiede auf. Die afrikanische Population weist eine Frequenz von etwa 70 % und die europäische Population eine Frequenz von nahezu 30 % auf, wohingegen die asiatische Population das IFN- λ_4 -codierende Allel nahezu eliminiert hat ($\approx 0\text{--}5\%$) (Prokunina-Olsson et al. 2013). Prokunina-Olsson et al. (2013) haben gezeigt, dass das *IFNL4*-Gen im Menschen einer Pseudogenisierung unterliegt, wobei diese Tendenz von Afrika ausgehend nach Osten hin verstärkt ausgeprägt ist.

Vor diesem Hintergrund geben die in dieser Arbeit gewonnenen Ergebnisse zusammen mit der global unterschiedlichen Genotypverteilung Anlass zu der Überlegung, dass die asiatische Population aufgrund ihrer geringen MAF des *IFNL4* rs368234815 eine Tendenz zu einer schnelleren Progression der PDAC-Tumorerkrankung mit kürzeren PFI bzw. PFI.2 und einer höheren Mortalität aufweisen müsste. Epidemiologische Auswertungen geben bisher keinen Hinweis. Aktuelle Tumorstatistiken zeigen, dass in Ländern mit einem hohen Lebensstandard und mit Zunahme des Durchschnittsalters die Inzidenz und Mortalität, die sehr ähnliche Zahlenwerte aufweisen, für PDAC am höchsten ist und weiter steigt. Europa, Nordamerika und Australien/Neuseeland weisen die höchste Inzidenz und Mortalität auf, wohingegen in den Regionen Asiens und Afrikas die niedrigsten Raten detektiert wurden (Bray et al. 2018). Zur Beurteilung der ethnischen Einflüsse und nicht der regionalen (Umwelt-)Unterschiede wurden verschiedene ethnische Populationen innerhalb der Vereinigten Staaten von Amerika auf die Inzidenz und Mortalität des PDAC untersucht. Hierbei wies die afrikanische Population mit der höchsten MAF des *IFNL4* rs368234815 die höchste Inzidenz und Mortalität für PDAC mit steigender

Tendenz auf, wohingegen die weiße und asiatische Population geringere Werte aufzeigte und damit einen bis jetzt noch nicht bekannten genetischen Vorteil haben müsste (NCI 2019).

Es ist anzunehmen, dass bei ethnischen Tumoruntersuchungen andere Risikofaktoren als der Einfluss der Ausprägung des *IFNL4*-rs368234815-Polymorphismus eine entscheidendere Rolle spielen. Aufgrund dieser weiteren unbekannt Einflüsse in unterschiedlichen Regionen und Populationen kann die genaue Effektstärke des in dieser Arbeit ermittelten, positiven Einflusses der Expression des IFN- λ_4 bei den PDAC-Patienten europäischer Herkunft nicht abgeschätzt werden (Matsubayashi et al. 2017). Hierbei darf der Einfluss der Diagnose- und Behandlungsmethoden nicht vergessen werden, der entscheidend die Dauer des PFI bzw. des PFI.2 und der OS prägt und regional stark unterschiedlich ist. Jedoch kann, wie in dieser Arbeit gezeigt wurde, innerhalb einer Population (Patienten europäischer Herkunft) und einer Region (Vereinigte Staaten von Amerika) und der Annahme einer bestmöglichen Therapie der Zusammenhang des *IFNL4*-rs368234815-Genotyps (positiv bei IFN- λ_4 -Ausprägung) mit der PDAC-Tumorprogression detektiert werden.

4.2 Therapie der TCGA-LIHC- und TCGA-PAAD-Patientenkollektive

Für das TCGA-LIHC-Kollektiv stellt die TCGA-Datenbank Informationen zu den angewandten chirurgischen Methoden und den Chemotherapien sowie zu den Ablations- und adjuvanten Strahlentherapien zur Verfügung¹. Teilweise wurde der Zeitpunkt des Therapiebeginns, die Dauer und Art der Therapie, die verwendeten Medikamente sowie die Therapieantwort dokumentiert. Bei den HCC-Patienten europäischer Herkunft wurden alle Patienten chirurgisch behandelt, wobei vorwiegend Lobektomien und einfache oder multiple Segmentektomien der Leber durchgeführt wurden. Ein Fall einer totalen Hepatektomie mit Transplantation ist auch verzeichnet. Eine Ablationstherapie, wobei zwischen Thermo-, Radiofrequenz- und Ethanolinjektionsablation unterschieden wurde, wurde bei 40 HCC-Patienten europäischer Herkunft durchgeführt. Adjuvant wurden 7 HCC-Patienten bestrahlt. Eine medikamentöse Therapie erhielten 34 HCC-Patienten. Häufig wurde den Patienten eine Kombination aus verschiedenen Medikamenten verabreicht, entweder gleichzeitig oder zeitlich versetzt aufgrund von Nichtansprechen der ersten Therapie oder aufgrund einer Tumorprogression. Als Erstlinienmedikation wurde vorwiegend der Multikinaseinhibitor Sorafenib verwendet. 12/34 Patienten erhielten eine

¹<https://portal.gdc.cancer.gov/projects/TCGA-LIHC>, Zugriff am 18.02.2020

Chemotherapie, zumeist aus Cisplatin und Gemcitabin. Angaben der Therapieantwort nach dem ersten Behandlungsregime fehlen für die HCC-Patienten.

Für das TCGA-PAAD-Kollektiv stellt die TCGA-Datenbank ebenfalls Informationen zu den angewandten chirurgischen Methoden sowie zu den adjuvanten Chemo- und den adjuvanten Strahlentherapien zur Verfügung². Hierbei sind die Informationen wie bei dem TCGA-LIHC-Kollektiv in den Zeitpunkt des Therapiebeginns, die Dauer und Art der Therapie, die verwendeten Medikamente sowie die Therapieantwort gegliedert. Bei den PDAC-Patienten europäischer Herkunft wurden alle Patienten vorwiegend einer partiellen Duodenopankreatektomie oder einer distalen bzw. totalen Pankreatektomie unterzogen. Eine adjuvante Strahlentherapie erhielten 37 PDAC-Patienten. Dokumentiert ist weiterhin die adjuvante Chemotherapie bei 109 PDAC-Patienten, wobei als erstes vor allem die Gemcitabin-Monotherapie erfolgte. Zusätzlich weist die TCGA-Datenbank für die PDAC-Patienten Angaben der Therapieantwort nach dem ersten Behandlungsregime auf. Hierbei konnten von den 109 chemotherapierten PDAC-Patienten folgende Therapieantworten für 75 Patienten ermittelt werden: 39 Mal kam es zur Progression, 29 vollständige Remissionen wurden dokumentiert, bei 4 Patienten kam es zu einer partiellen Remission und bei 3 Patienten stellte sich eine stabile Erkrankung ein. Für eine Stratifizierung der 75 chemotherapierten PDAC-Patienten mit dokumentierter Therapieantwort nach *IFNL*-Genotypen wurden die Therapieantworten „vollständige Remission“ und „stabile Erkrankung“ zusammengefasst. Die Patienten wurden in eine Gruppe mit für das häufigere Allel C homozygoten Patienten und in eine Gruppe mit Trägern des selteneren Allels G (CG/GG) unterteilt. Hierbei ist deskriptiv und explorativ die Tendenz eines Zusammenhangs zwischen den Trägern des Allels G, die für ein funktionsfähiges *IFN-λ₄* codieren, und einer vollständigen Remission zu erkennen. Dies könnte ein Hinweis dafür sein, dass der Genotyp des Polymorphismus *IFNL4* rs368234815 entweder neben einem Einfluss auf die Überlebenswahrscheinlichkeit auch einen Einfluss auf das Ansprechen einer Chemotherapie hat oder über einen Einfluss auf das Ansprechen auf eine Chemotherapie Einfluss auf die Überlebenswahrscheinlichkeit nimmt.

Die Therapieinformationen wurden innerhalb dieser Arbeit nicht differenzierter berücksichtigt und analysiert. Vorwiegend wurde auf Progressions-/Endparameter eingegangen, in der Annahme, dass sowohl alle HCC- als auch alle PDAC-Patienten jeweils nach Evaluierung ihres Allgemeinzustands bestmöglich nach aktuellen Leitlinien therapiert wurden. Die Frage des Therapieeinflusses bei differenzierter Betrachtung der Therapie soll Gegenstand einer gesonderten Analyse werden.

²<https://portal.gdc.cancer.gov/projects/TCGA-PAAD>, Zugriff am 18.02.2020

4.3 Duale Wirkungsweise der IFNe und humane Pseudogenisierungsmechanismen

Die Assoziation der *IFNL4*-rs368234815-Genvarianten mit verschiedenen viralen und tumoralen Erkrankungen konnte durch vorherige genetische Assoziationstudien belegt werden (Lazear et al. 2015; Lasfar et al. 2019). Jedoch ist die biologische Funktionsweise des IFN- λ_4 -Proteins noch nicht verstanden. Ein Grund dafür ist unter anderem, dass das *IFNL4*-Gen im Maus- und im Rattengenom fehlt und deswegen murine Tiermodelle zur Untersuchung der endogenen *IFNL4*-Wirkweise nicht genutzt werden können (Obajemu et al. 2017). Weiterführende Untersuchungen im Hinblick auf die exakte *IFNL4*-Funktion werden benötigt.

Dennoch wurde eine duale Rolle des IFN- λ_4 -Proteins im Menschen bei viralen und tumoralen Erkrankungen erkannt, die die differenzierte Betrachtung der IFN- λ_4 -Wirkung auf verschiedene Organe und bei verschiedenen Erregern sowie verschiedenen Tumorentitäten erforderlich macht (Lasfar et al. 2019).

Antivirale und antitumorale Wirkungen aller Typ-III-IFNe konnten für verschiedene Entitäten aufgezeigt werden. Antivirale Eigenschaften der Typ-III-IFNe wurden unter anderem bei dem HCV (Prokunina-Olsson et al. 2013), dem HBV (Robek et al. 2005), dem Cytomegalievirus (CMV) (Egli et al. 2014b), dem Herpes-Simplex-Virus 1/2 (HSV1/2) (Zhou et al. 2015), dem Dengue-Virus (Palma-Ocampo et al. 2015), dem Influenzavirus (Egli et al. 2014a) und einigen anderen detektiert (Lazear et al. 2015; Obajemu et al. 2017; Bibert et al. 2018).

Antitumorale Aktivitäten wurden unter anderem bei Melanomzell- und Melanommausmodellen (Lasfar et al. 2006), in humanen Zellkulturen von Lungenadenokarzinomen (Tezuka et al. 2012) und Ösophaguskarzinomen (Li et al. 2010) und *in vivo* hepatozellulären Karzinommausmodellen (Abushahba et al. 2010) ermittelt (Lasfar et al. 2019).

Erst kürzlich wurden neue Erkenntnisse zu den Typ-III-IFNen und deren Wirkung auf Tumoren veröffentlicht, wobei neue Signalwege entdeckt und Unterschiede zu den Typ-I- und Typ-II-IFNen und der ISG-Expression beschrieben wurden. Unter anderem wurde gezeigt, dass die Typ-III-IFNe auch unabhängig vom JAK-STAT-Signalweg JAK2 induzieren und dadurch nicht nur ein anderes Spektrum von IFN-Effektoren exprimieren können, sondern auch die Freisetzung der reaktiven Sauerstoffspezies (*reactive oxygen species*, ROS) und die Degranulation von neutrophilen Granulozyten herabregulieren können. Hierdurch regulieren die Typ-III-IFNe im Gegensatz zu den Typ-I- und Typ-II-IFNen Entzündungsreaktionen herunter und sind nicht an der Exazerbation von Entzündungen oder der Überstimulation des Immunsystems beteiligt (Lasfar et al. 2019).

Die ROS sind wichtig für die Aktivität des Immunsystems, jedoch können sie in zu hoher

Konzentration zu vermehrten Entzündungen und Tumorprogressionen führen. Vermehrt befinden sich ROS in den Fibroblasten um solide Tumoren wie dem PDAC, weswegen die Hypothese aufgestellt wurde, dass sich die Unterdrückung der ROS-Freisetzung durch die Typ-III-IFNe und damit eventuell auch speziell durch IFN- λ_4 hemmend auf die PDAC-Progression auswirke und damit die Überlebenszeit verlängere (Öhlund et al. 2017; Lasfar et al. 2019).

Die Ergebnisse dieser Arbeit führen zur Bestärkung des bisher vermuteten hemmenden Einflusses der Typ-III-IFNe auf die Tumorprogression des PDAC durch direkte und indirekte antitumorale Wirkungen. Zu diesen Wirkungen, die bereits in vorherigen Studien ermittelt wurden, gehören unter anderem die Hemmung der Zellproliferation, die gesteigerte Apoptose und der erhöhte Zellzyklusarrest sowie die Immunzellaktivierung unter anderem der T-Zellen, die Hemmung der Angiogenese sowie die ROS-Inhibition in der Tumorumgebung (Lasfar et al. 2019).

Des Weiteren sind bereits seit längerem apoptotische und antiproliferative Effekte von vorwiegend IFN- β und IFN- γ , aber auch IFN- α , auf das PDAC nachgewiesen worden sowie ihre chemosensibilisierende Wirkung zum Beispiel auf das Chemotherapeutikum 5-Fluoruracil (5-FU) (Booy et al. 2015). Dieser hemmende Einfluss der Typ-I- sowie Typ-II-IFNe auf das Wachstum des PDAC bestärkt die Annahme der möglichen antiproliferativen und apoptotischen Wirkung auch der Typ-III-IFNe auf das PDAC. Aufgrund der dualen Wirkung der IFNe wird der IFN-Gebrauch in der PDAC-Therapie derzeit noch kontrovers diskutiert.

Neben den antitumoralen Wirkungen der Typ-III-IFNe wurde diesen zusätzlich protumorale Wirkungen unter anderem im Blasenkarzinom zugeschrieben, wobei das IFN- λ_2 die Tumordinvasion fördert (Lee et al. 2012). Des Weiteren wurden dem IFN- λ_2 noch die Unterstützung der Angiogenese, der epithelial-mesenchymalen Transition und der Metastasierung als protumorale Effekte zugeschrieben (Pingwara et al. 2017).

Dank weiterer Studien konnten spezifisch für das *IFNL4* Assoziationen mit verschiedenen biologischen Prozessen und Wechselwirkungen mit Pathogenen detektiert werden. Neben dem Nachteil einer genetischen Disposition zur IFN- λ_4 -Expression bei HCV-Infektionen wurden unter anderem noch die verminderte Elimination des Cytomegalievirus (CMV) (Manuel et al. 2015) und vermutlich des Humanen Immundefizienz-Virus (HIV) (Real et al. 2015) in Assoziation mit der IFN- λ_4 -Expression beschrieben (Minas et al. 2018). Des Weiteren wurde eine Assoziation der Disposition zur IFN- λ_4 -Expression mit einem erhöhten Risiko an einem aggressiven Prostatakarzinom zu erkranken und früher daran zu versterben bei Männern mit sexuell übertragbaren Infektionen detektiert (Minas et al. 2018). Die Ausprägung einer IFN-bezogenen DNA-Schaden-resistenten Signatur (*IFN-*

related DNA damage-resistant signature, IRDS) bei afrikanischen Männern ist auch mit dem polymorphen *IFNL4*-Genlocus assoziiert. Diese Signatur zeigt einen Zusammenhang mit einer kürzeren OS und einer Resistenzentwicklung gegenüber Chemotherapien und ionisierender Strahlung. Außerdem assoziiert sie mit einem gesteigerten Tumorstadium und einer gesteigerten Metastasierung sowie einer Erhöhung des Risikos der Entwicklung eines Prostatakarzinoms (Tang et al. 2018). Daneben wurden weitere Korrelationen zwischen den *IFNL4*-Genvarianten und unter anderem dem Kaposi Sarkom (Bibert et al. 2018) und der Entwicklung von Leberfibrose (Eslam et al. 2017) innerhalb verschiedener Studien beschrieben (Prokunina-Olsson 2019).

Es fällt auf, dass bei allen Tumoren viraler Ätiologie die Expression eines funktionsfähigen IFN- λ_4 als prognostisch ungünstig beschrieben wurde. Bei den PDAC-Tumoren, für die kein viraler Infektionszusammenhang angenommen wird, konnte in dieser Arbeit ein prognostisch vorteilhafter Zusammenhang mit der Ausprägung des funktionsfähigen IFN- λ_4 -Proteins gezeigt werden.

Die für den Menschen häufig als unvorteilhaft beschriebenen Assoziationen der *IFNL4*-rs368234815-Allelvariante ΔG mit der Progression verschiedener Krankheiten sowie die verschiedenen detektierten, vor allem human entwickelten Mechanismen, um die Transkription, Translation und Funktionalität des *IFNL4* zu verringern, deuten auf einen insgesamt negativen Einfluss der *IFNL4*-rs368234815-Allelvariante ΔG im Menschen und auf ein Vorteil bei der Abwesenheit des funktionstüchtigen IFN- λ_4 -Proteins hin.

Zu den evolutionären Mechanismen gehören neben der Pseudogenisierung des *IFNL4*-Gens durch den Austausch des ursprünglichen Allels ΔG durch das Allel TT, das die Translation eines funktionstüchtigen IFN- λ_4 aufgrund einer Leserasterverschiebung verhindert, noch die Ausprägung verschiedener weiterer genetischer Polymorphismen innerhalb der *IFNL4*-Genloki wie P70S und 154K, die die Sekretion und Stabilität des IFN- λ_4 -Proteins reduzieren (Bamford et al. 2018). Weitere Mechanismen zur Inhibition der *IFNL4*-Expression auf der Ebene der Transkription und Translation im Menschen wurden beschrieben (Hong et al. 2016; Bamford et al. 2018; Prokunina-Olsson 2019).

Die Ursache der populationsspezifischen Elimination des *IFNL4*, die vor etwa 60.000 Jahren mit der Migration des Menschen aus Afrika begann, ist noch nicht endgültig geklärt. Jedoch wird angenommen, dass der wahrscheinlichste Selektionstreiber die Exposition gegenüber vor allem viralen Pathogenen sei, wobei die HCV-Infektion durch ihren doch relativ niedrigen Anteil an tödlichen Verläufen als alleinige Ursache unwahrscheinlich ist (Key et al. 2014). Außerdem werden noch spezifische Umwelteinflüsse, die dem eigentlichen Selektionstreiber entgegenwirken, angenommen, wodurch die geografischen Unterschiede in den MAFs von den relevanten *IFNL4*-Polymorphismen bestimmt

werden (Key et al. 2014).

Es lässt sich vermuten, dass die bis jetzt noch unbekannte Ursache für die Pseudogenisierung des *IFNL4* häufiger auftritt und/oder einen gesundheitlich noch stärkeren Einfluss auf die Menschen ausübt als das PDAC, sodass auch der in dieser Arbeit ermittelte positive Einfluss des *IFNL4* auf die Tumorprogression der PDAC-Patienten die Elimination der Fähigkeit zur Expression eines funktionsfähigen IFN- λ_4 nicht stoppt. Außerdem besteht die Möglichkeit, dass der Tumor PDAC mit der Elimination des antitumoral wirkenden IFN- λ_4 eine Lücke in der Tumorabwehr nutzt.

4.4 Einfluss der Typ-III-IFNe auf die Tumorprogression des HCC

In dieser Arbeit wurde für Patienten mit einem HCV-bedingten HCC ein etwa 2,3-fach erhöhtes Risiko für eine Progression des HCC detektiert. Diese frühere Progression bei HCC-Patienten mit einer HCV-Infektion wurde auch von Reig et al. (2016) beschrieben. Die Autoren behandelten in kompletter Tumorresponse befindliche HCC-Patienten, deren Tumoren HCV-bedingt waren, mit direkt antiviral wirkenden Substanzen (*direct acting antivirals*, DAA) gegen die HCV-Infektion. Hierbei detektierten Reig et al. (2016) bei den Patienten eine erstaunlich hohe Rate an HCC-Progressionen im Vergleich zu Patienten, deren HCCs nicht HCV-bedingt waren.

Diese publizierten Ergebnisse bestärken die in der vorliegenden Arbeit ermittelte Assoziation zwischen dem Risikofaktor einer HCV-Infektion und der HCC-Progression. Ein Zusammenhang mit dem *IFNL4*-Genotyp lässt sich vermuten und diskutieren. Bereits bekannt ist, dass es nur bei einer HCV-Infektion zu einer Aktivierung und Expression von IFN- λ_4 kommt (Mihm 2015), welches wiederum zu einer nicht ausreichenden Eradikation des HCV führt (Prokunina-Olsson et al. 2013). Transkriptomanalysen von Daten der *IFNL4*-Expression aus der TCGA-Datenbank, stratifiziert nach dem Kriterium Ätiologie, zeigen eine signifikant erhöhte *IFNL4*-Expression bei HCC-Patienten innerhalb der Ätiologie HCV (aufgrund des Umfangs der Daten wurden diese nicht dargestellt). Die nicht ausreichende Viruseradikation bei dem Vorliegen eines funktionstüchtigen IFN- λ_4 führt vermutlich vermehrt zur Chronifizierung und zu vermehrten Leberschäden, die die Entstehung eines HCC begünstigen. Wenn bei dem gleichzeitigen Vorliegen einer HCV-Infektion und der Fähigkeit zur Expression eines funktionstüchtigen IFN- λ_4 die Inzidenzrate für das HCC erhöht ist, liegt die Vermutung nahe, dass auch nach Therapie des HCC sowohl die Rezidiv- als auch die Progressionsrate erhöht sind, da vermutlich nie der komplette Tumor entfernt werden kann und somit die oben genannten Vorausset-

zungen (HCV/funktionsfähiges IFN- λ_4) die Tumorprogression eventuell fördern könnten. Zur Überprüfung dieser Hypothesen sind weiterführende Studien notwendig.

Keine Assoziation wurde zwischen dem *IFNL4*-rs368234815-Genotyp und der Tumorprogression bei den HCC-Patienten europäischer Herkunft detektiert. Zum einen sei darauf hingewiesen, dass mit dem Pankreas und der Leber zwei verschiedene Organsysteme betrachtet wurden, die unterschiedliche biologische Funktionsweisen sowohl im gesunden Organ als auch innerhalb ihrer jeweiligen Tumorentitäten aufzeigen und an denen sich die Wirkungen unter anderem von IFNen unterschiedlich ausprägen können. Hierbei sei auf die Eigenheit der Leber als immuntolerantes Organ aufmerksam gemacht. Zum Beispiel konnte bei der chronischen Hepatitis C nur das IFN- λ_4 und keines der anderen IFNe innerhalb der Leber detektiert werden (Mihm et al. 2004; Mihm 2015). Jedoch wurde, wie bereits erwähnt, die Expression eines funktionstüchtigen IFN- λ_4 mit der schlechteren Ausheilung einer HCV-Infektion innerhalb der Leber beschrieben, was zumeist zu einer Chronifizierung der HCV-Infektion mit anschließender chronischen Hepatitis führt. Diese chronische Entzündung ist als ein eindeutiger Risikofaktor für die Entstehung des HCC bekannt. Dies lässt vermuten, dass IFNe als antivirale Immunsystemmediatoren keine oder nur unzureichende Wirkungen in der Leber, vermutlich im Rahmen der lebereigenen Immuntoleranz, ausüben.

Jedoch wurde innerhalb anderer Studien ein antitumoraler Einfluss von IFN- λ_4 auf *in vivo* hepatozellulären Karzinommodellen ermittelt (Prokunina-Olsson et al. 2013), so dass auch ein dosisabhängiger karzinombekämpfender Effekt des IFN- λ_4 auf das HCC in Betracht gezogen werden kann, was bereits durch andere Arbeitsgruppen vermutete wurde (Onabajo et al. 2015).

Zusätzlich wurde von Petta et al. (2017) für ein europäisches Kollektiv beschrieben, dass das Vorliegen des homozygoten *major IFNL4*-rs368234815-Genotyps TT/TT verstärkt zur Fibrose und Nekroinflammation in nichtalkohol-induzierten Fettlebern führt im Vergleich zu den Trägern des *IFNL4*-rs368234815-Allels ΔG . Daraus kann geschlossen werden, dass Patienten, die homozygot für das Allel TT sind und damit kein funktionstüchtiges IFN- λ_4 exprimieren, womöglich an stärkeren Fibroseschäden der Leber leiden, was auch zu der Entwicklung eines HCC führen kann.

Andererseits wurde innerhalb eines Kongressbeitrags von Onabajo et al. (2018) die Untersuchung des Einflusses des *IFNL4*-rs368234815-Polymorphismus auf die Tumorprogression bei HCC-Patienten der TCGA-Datenbank beschrieben, wobei der SNP *IFNL3/4* rs12980275 als Surrogat für *IFNL4* rs368234815 genutzt wurde. Die Autoren ermittelten eine Assoziation zwischen dem Genotyp und der OS. Hierbei wiesen Patienten mit dem *IFNL4*-rs368234815-Allel ΔG eine verkürzte OS und eine immunreduzierte Tumormi-

kroumgebung auf.

Zusätzlich wurde noch der Zusammenhang zwischen dem *IFNL4*-rs368234815-Allel ΔG und einem reduzierten Risiko für Leberzirrhose erwähnt, was durch die vermutete Induktion von Metallothioneine durch das funktionsfähige IFN- λ_4 erklärt wurde. Diesen Zusammenhang beschrieben bereits Eslam et al. (2017), indem sie herausfanden, dass vorwiegend IFN- λ_3 und nicht IFN- λ_4 durch vermehrte Entzündungszellinfiltration in die Leber zu einer erhöhten Leberentzündung und -fibrose führen.

Die Betrachtung der Populationsauswahl bei der Ergebnisinterpretation ist sehr wichtig, wie bei dem Vergleich der Ergebnisse von dieser Arbeit und von der Arbeit von Onabajo et al. (2018) deutlich wird. In dem veröffentlichten Kongressbeitrag von Onabajo et al. (2018) lagen keine Angaben zu der Ethnizität der Patienten vor.

Die regional verschiedenen Stärken der genetischen Einflüsse, die in dieser Arbeit durch die unterschiedlichen Auftretenshäufigkeiten des Allels ΔG des *IFNL4*-rs368234815-Polymorphismus bestimmt wurden, können die Stärke der beobachtbaren Assoziationen entscheiden. Vermutlich wird bei der afrikanischen Population aufgrund der hohen MAF die Stärke der beobachtbaren Assoziationen im Zusammenhang mit dem *IFNL4*-rs368234815-Polymorphismus höher sein als bei den in dieser Arbeit analysierten Patienten europäischer Herkunft. Außerdem spricht die ethnische Verteilung der Inzidenz des HCC für einen noch nicht endgültig ermittelten Vorteil der europäischen Population im Gegensatz zu der asiatischen und afrikanischen Population (McGlynn et al. 2015). Im Unterschied zu dem PDAC ist die Inzidenz und Mortalität des HCC in weniger entwickelten Ländern und vor allem bei Asiaten und Afrikanern höher als in weiter entwickelten Ländern Europas und Nord- bzw. Südamerikas. Jedoch nehmen nicht nur der Entwicklungsstand und die Umwelt Einfluss auf die Tumorzinzidenz. In einer SEER (*The Surveillance, Epidemiology and End Results*)-Studie zwischen 2006 und 2012 konnte bei untersuchten amerikanischen Patienten verschiedener Ethnizität Unterschiede innerhalb einzelner Ethnizitäten detektiert werden, was das Einwirken eines genetischen Einflusses auf die Entwicklung von HCC vermuten lässt. Hierbei war die Inzidenz für asiatische Amerikaner am höchsten, gefolgt von hispanischen und afrikanischen Amerikanern. Die niedrigste Inzidenz wiesen weiße Amerikaner auf (McGlynn et al. 2015).

Des Weiteren spielt die Therapie sowie der Screening- und Behandlungsstandard einer Population eine große Rolle sowohl auf eine Verlängerung des PFI bzw. des PFI.2 als auch auf eine Verlängerung der OS und wird bei Patienten europäischer Herkunft als vergleichsweise hoch angenommen (Li et al. 2019). Der europäische Therapiestandard vermag möglicherweise, die Unterschiede der Tumorprogression zwischen IFN- λ_4 -exprimierenden Patienten und den Patienten, die kein funktionstüchtiges IFN- λ_4 expri-

mieren, zu verringern.

Es lässt sich vermuten, dass der Vorteil der europäischen Population mit einem hohen Screening- und Therapiestandard, unter anderem der weit verbreiteten HBV-Impfung, zusammenhängt.

Das Vorliegen der mittleren MAF kann bedeutet, dass die Vor- und Nachteile der IFN- λ_4 -Wirkung etwa gleichmäßig in der Bevölkerung verteilt vorliegen. Die verschiedenen Therapiestandards und die Unterschiede in der MAF könnten mitverantwortlich dafür sein, dass eine mögliche Richtungstendenz des Genotypeinflusses in dieser Arbeit nicht detektiert werden konnte. Wenn die Annahme besteht, dass bei den afrikanischen und asiatischen Populationen im Durchschnitt ein etwas niedriger Screening- und Therapiestandard vorliegt, werden die Einflüsse des IFN- λ_4 voraussichtlich nicht wie bei der europäischen Population durch effektive und rechtzeitige Therapien reduziert. Für genauere Erkenntnisse sind zusätzliche Studien notwendig.

Des Weiteren gibt das Wissen über die MAF Anlass zu der Vermutung, dass die afrikanische Population mit einer Frequenz des Allels ΔG von etwa 70 % vermehrt an einem HCV-bedingten HCC leidet und die asiatische Population vermehrt aufgrund ihrer sehr geringen MAF von etwa 0–5 % eher an einem *IFNL3*-vermittelten, Leberfibrose-induzierten HCC leidet. Zusätzlich kommt bei Asiaten noch die hohe Frequenz der HBV-Infektionen als Ätiologie des HCC hinzu. Diese Annahmen konnten mittels zusätzlicher Betrachtung aller nach Ätiologie stratifizierte Ethnizitäten des TCGA-LIHC-Kollektivs bestärkt werden (Anhang Tab. A.3). Hierbei wurde eine Tendenz in die vermutete Richtung der vermehrten HCV-Infektion bei der afrikanischen Population und einer vermehrten Fibrose (innerhalb der Ätiologie „Kein virales Risiko“) bei der asiatischen Population als Grund für das HCC gezeigt (Anhang Tab. A.3). Jedoch waren sehr wenig afrikanische Amerikaner in dem Kollektiv enthalten und die Fibrose war als Ätiologie nicht einzeln aufgegliedert, weswegen eine Bestätigung nicht vorgenommen werden konnte.

Neben der Betrachtung der IFN- λ_4 -Wirkung lässt sich vermuten, dass der Einfluss des stabileren und vermehrt exprimierten IFN- λ_3 bei den für ein nicht funktionsfähiges IFN- λ_4 codierenden Patienten auch einen Einfluss auf die progressionsfreien Intervalle, PFI bzw. PFI.2, und die OS der Patienten hat. Dies würde die direkte Funktion der Surrogat SNPs (*IFNL3* rs4803217 und *IFNL3* rs28416813) widerspiegeln. Aufgrund der bekannten antiviralen und antitumoralen Wirkung des IFN- λ_3 kann eine positive Assoziation zwischen dem *IFNL4*-rs368234815-Allel TT, d. h. keine Expression eines funktionstüchtigen IFN- λ_4 und gleichzeitig eine verstärkte und stabilere Expression des IFN- λ_3 , und einem verlängerten PFI bzw. PFI.2 sowie einer verlängerten OS der HCC-Patienten vermutet werden. Dadurch könnte das Vorliegen eines *IFNL4*-rs368234815-Genotyps TT/TT, was

vor allem auf die europäische und asiatische Population zutrifft, einen Überlebensvorteil vermuten lassen. Dieser Vermutung wirken aber Studien von O'Brien et al. (2015) entgegen, bei denen gezeigt werden konnte, dass die Ausprägung eines stabileren IFN- λ_3 zusammen mit dem Auftreten eines funktionstüchtigen IFN- λ_4 mit einer schlechteren Ausheilung des HCV assoziiert war im Vergleich zu allen anderen Allelvariationen an diesen zwei polymorphen Genloki (*IFNL4* rs368234815 und *IFNL3* rs4803217). Diese Beobachtung führte zu der Annahme, dass die höhere *IFNL3*-Stabilität einen nicht so starken positiven Effekt habe, vor allem nicht wenn sie gemeinsam mit dem Genotyp des funktionstüchtigen IFN- λ_4 auftrete (Prokunina-Olsson 2019). Aus dieser Beobachtung und Behauptung wurde in dem Artikel außerdem gefolgert, dass die Wirkung des funktionstüchtigen IFN- λ_4 den Einfluss des stabileren IFN- λ_3 überwäge und somit der *IFNL4*-rs368234815-Polymorphismus der einflussreichste dieses Haplotypblocks sei und als Kandidatenpolymorphismus agiere (O'Brien et al. 2015).

Die Betrachtung der Funktionsweise der Leber und ihres immuntoleranten Status sowie der positiven und der negativen Wirkungen des funktionstüchtigen IFN- λ_4 auf die Leber und auf die Entwicklung eines HCC lässt keine Aussage über eine Tendenz der Assoziation des *IFNL4*-rs368234815-Genotyps mit der Tumorprogression des HCC innerhalb der europäischen Population zu.

In Bezug auf die MAF zeigen sich wahrscheinlich eindeutigere Wirkrichtungen bei der Untersuchung der afrikanischen Population mit einer sehr hohen MAF für den *IFNL4*-rs368234815-Polymorphismus oder bei der asiatischen Population, deren MAF des *IFNL4*-rs368234815 sehr gering ist.

4.5 Prognostik der Tumorprogression

Der Einfluss des Tumordifferenzierungsgrads und auch des Stadiums der Tumorerkrankung auf die Tumorprogression verhielt sich wie erwartet und bestätigt erneut die analysierten Werte. Hierbei geht ein höherer Differenzierungsgrad und ein höheres Tumorstadium mit einer schlechteren Tumorprognose, kürzeren progressionsfreien Intervallen, PFI bzw. PFI.2, und einem kürzeren Überleben einher. Diese Ergebnisse sind konsistent zu vorherigen Analysen: ein undifferenzierterer Tumor, d. h. ein Tumor mit höherem Differenzierungsgrad, metastasiert früher, verhält sich aggressiver und führt zu einer schlechteren Prognose als ein differenzierterer Tumor. Ein vom Volumen größerer Tumor, d. h. ein Tumor mit höherem Tumorstadium, deutet ebenfalls auf eine schlechtere Prognose hin als ein Tumor mit niedrigerem Tumorstadium. Diese Einflüsse sind unter anderem in der siebten Auflage des *AJCC Cancer Staging Manual* zusammengefasst (Edge et al. 2010). Eine Literaturanalyse mit Betrachtung von 114 wissenschaftlichen

Artikeln zeigte diese Korrelation der Höhe des Differenzierungsgrads mit der Tumorprogression für HCC-Patienten (Martins-Filho et al. 2017). Des Weiteren wurde von mehreren unabhängigen Arbeitsgruppen der Differenzierungsgrad bei PDAC-Patienten als einer der stärksten unabhängigen Einflussfaktoren auf das Überleben beschrieben (Rochefort et al. 2013).

Bei Betrachtung der OS innerhalb der multiplen Cox-Regressionsanalyse weisen die HCC-Patienten mit höherem Alter ein leicht erhöhtes Sterberisiko auf. Da das Sterberisiko bei dem stetigen Kriterium Alter in Risikoerhöpfung pro Altersjahr betrachtet wird, steigt das Risiko mit jedem Jahr an und schon eine geringe signifikante Tendenz ist meist klinisch relevant. Aus Tumorstatistiken, unter anderem dem von den NIH (*National Institutes of Health*) öffentlich zur Verfügung gestellten Programm SEER (*The Surveillance, Epidemiology and End Results*), ist eindeutig bekannt, dass Tumoren, hierbei sowohl das HCC als auch das PDAC, vorwiegend Erkrankungen des höheren Lebensalters sind und dass das höhere Alter somit einen relevanten Risikofaktor darstellt (NCI 2020a).

In der vorliegenden Arbeit wurde für das PDAC der Polymorphismus *IFNL4* rs368234815 als ein Wirtsfaktor für die Tumorprognose ermittelt, welcher bei Trägern des *minor* Allels ΔG auf einen günstigeren Verlauf der Erkrankung mit einer erhöhten Lebensqualität in Form von verlängerten progressionsfreien Intervallen, PFI bzw. PFI.2, und einer höheren OS hindeutet. Diese Entdeckung und die bis jetzt unbekannte Interaktion auf molekularer Ebene erfordern weitere Bestätigungen und Untersuchungen. Die Assoziationen lassen sich aufgrund des kontinuierlichen *Follow-ups* der Patienten und der ergänzenden Dokumentation in der TCGA-Datenbank zu einem späteren Zeitpunkt mit einer höheren Fallzahl erneut analysieren. Hierbei ist das Aktualisieren der TCGA-Daten abzuwarten und regelmäßig zu überprüfen.

4.6 Limitationen der durchgeführten Analysen

Diese Arbeit weist Limitationen auf, die berücksichtigt werden müssen. Bei Fehlen oder Inkonsistenz der TCGA-Daten wurden durch das TCGA-CDR-Netzwerk standardisierte Annahmen getroffen, um die Daten zu harmonisieren und nutzen zu können (Liu et al. 2018). Diese Inkonsistenz der TCGA-Daten kam aber erst dadurch zustande, dass bei der Projektplanung das Zusammentragen der demografischen und klinischen Daten und insbesondere der Ereigniszeiten nicht das primäre Ziel der TCGA-Datenbank war und diese Daten erst später integriert wurden. Zusätzlich existiert kein konkreter *Follow-up*-Plan von der TCGA-Datenbank, sodass die *Follow-up*-Zeiten nicht einheitlich gesammelt wurden und in Zukunft deswegen auch nicht komplett konsistent vervollständigt werden

können.

Das TCGA-CDR-Netzwerk deutet auf die potentielle Fehlerquelle der sogenannten konkurrierenden Risiken hin, die bei Überlebensanalysen und vor allem bei den spezifischen Ereigniszeitanalysen, die von einer konkreten Krankheit ausgehen, eine Rolle spielt. Hierbei kann die Beobachtung des Eintritts des definierten Ereignisses durch den früheren Eintritt einer anderen Krankheiten verhindert werden (Liu et al. 2018).

In dieser Arbeit und auch innerhalb des TCGA-CDR-Datensatzes wurden die jeweilig angewandten Therapien der HCC- bzw. PDAC-Patienten nicht betrachtet, die aber einen erheblichen Einfluss auf die Tumorprogression (PFI bzw. PFI.2) sowie auf die OS haben. Hierdurch besteht die Möglichkeit, dass es zu Veränderungen der Ergebnisse und damit zu einer falschen Interpretation kommt. Das Gleiche gilt für weitere nicht beachtete Faktoren, zum Beispiel weitere klinische Kriterien, die in dieser Arbeit einen nicht abschätzbaren und nicht beachteten Einfluss auf die Ergebnisse ausüben. Eine allgemeine, aber für diese Arbeit irrelevante Limitation der TCGA-Datenbank und auch anderer Datenquellen ist, dass die Tumorproben nicht die ganze genetische Vielfalt jedes einzelnen Tumors widerspiegeln können, da die Gewebeproben für die Genomdatensequenzierung zumeist nur aus einzelnen Bereichen des neu diagnostizierten Tumors gewonnen wurden.

5 Zusammenfassung

Die beiden Tumorentitäten hepatozelluläres Karzinom (*hepatocellular carcinoma*, HCC) und duktales Pankreasadenokarzinom (*pancreatic ductal adenocarcinoma*, PDAC) zählen zu den tödlichsten Tumoren mit der geringsten 5-Jahres-Überlebenswahrscheinlichkeit ($< 20\%$), deren Inzidenz global betrachtet weiter steigt. Interferone (IFNe) (Typ I-III) sind sowohl antiviral als auch antitumoral wirkende Proteine. Innerhalb der Typ-III-IFN-Gene gibt es verschiedene genetische Polymorphismen, die mit einem überzeugenden antiviralen Phänotyp, der spontanen und der therapieinduzierten Elimination des Hepatitis-C-Virus (HCV) im Menschen, assoziieren (Kapitel 1.3.4). Das Allel ΔG einer dieser Variationen, *IFNL4* rs368234815, ist die evolutionsgeschichtlich ältere Variante und erlaubt die Expression eines funktionstüchtigen IFN- λ_4 -Proteins, während die jüngere Allelvariante TT über eine Verschiebung des Leserasters zu einem funktionalen *knock-out* des *IFNL4*-Gens führt. TCGA (*The Cancer Genome Atlas*), eine US-amerikanisch geförderte Datenbank, stellt Genome von Tumorpatienten in großer Zahl und kuratierter Form gemeinsam mit umfangreichen demografischen und klinischen Daten zur Verfügung. Die vorliegende Arbeit greift auf die Sequenzdaten von nicht-malignem Gewebe (Keimbahn) zu, um klinische Evidenz für oder gegen eine prognostische Signifikanz von genetischen Polymorphismen in *IFN lambda* (*IFNL*)-Genen bei Tumorerkrankungen, hier HCC und PDAC, zu finden.

Die TCGA-Datenbank enthält Daten von insgesamt 377 HCC-Patienten (TCGA-LIHC (*liver hepatocellular carcinoma*)-Kollektiv), davon 187 Patienten europäischer Herkunft, und 185 PDAC-Patienten (TCGA-PAAD (*pancreatic adenocarcinoma*)-Kollektiv) mit insgesamt 162 Patienten europäischer Herkunft.

Als Maß für die Tumorprogression werden in dieser Arbeit die Ereigniszeiten progressionsfreies Intervall (*progression free interval*, PFI), spezifisches progressionsfreies Intervall (*progression free interval.2*, PFI.2) und Gesamtüberlebensdauer (*overall survival*, OS) herangezogen.

Durch Berechnung des standardisierten Kopplungsungleichgewicht (*Linkage Disequilibrium*, LD)-Koeffizienten D' und des Korrelationsfaktors r^2 aus den Daten des 1000-Genome-Projekts für die insgesamt fünf betrachteten polymorphen Positionen in *IFNL*-Genen konnte deren Nutzbarkeit als *tagSNPs* bzw. von *IFNL3* rs4803217 für die HCC-Patienten und *IFNL3* rs28416813 für die PDAC-Patienten als Surrogate für *IFNL4* rs368234815 belegt werden.

Anschließend wurden die Genotypen dieser fünf polymorphen Positionen innerhalb der NCBI (*National Center for Biotechnology Information*) *Genome Workbench* dargestellt

und ausgelesen. Ein Vergleich der Genotypverteilung jeweils der HCC- und der PDAC-Patienten mit einem für Individuen europäischer Abstammung adjustierten Kontrollkollektiv aus dem 1000-Genome-Projekt zeigte keine signifikanten Unterschiede.

Für die Auswertungen wurden, um Verzerrungen zu vermeiden, nur die Patienten europäischer Herkunft stratifiziert: in eine Gruppe mit Patienten, die Träger des selteneren Allels ΔG sind und damit die Disposition zur Bildung eines funktionstüchtigen IFN- λ_4 besitzen, und eine Vergleichsgruppe mit Patienten, die durch das Vorliegen des Allels TT kein funktionstüchtiges IFN- λ_4 bilden können.

Weder die nach Genotypen stratifizierten HCC- noch die PDAC-Patienten unterscheiden sich signifikant im Tumordifferenzierungsgrad oder im Tumorstadium. Innerhalb der Ereigniszeitanalysen nach dem Verfahren von Kaplan-Meier und der multiplen Cox-Regressionsanalyse zeigte sich, dass erwartungsgemäß ein höheres Alter, ein höherer Tumordifferenzierungsgrad und ein höheres Tumorstadium sowohl bei den HCC- als auch bei den PDAC-Patienten mit einer früheren Progression (PFI bzw. PFI.2) und einer schlechteren OS assoziieren oder die Tendenz zur Assoziation aufweisen.

Für die Patienten mit einem HCV-bedingten HCC stellte sich ein erhöhtes Risiko für eine Progression im Vergleich zu den Patienten mit einem HBV-bedingten HCC dar.

Für die PDAC-Patienten konnte im Vergleich zu den HCC-Patienten ein signifikanter Zusammenhang zwischen der genetischen Disposition, ein IFN- λ_4 zu exprimieren, und einer Verzögerung der Progression der Tumorerkrankung sowie der Tendenz zu einer längeren OS ermittelt werden.

Diese Arbeit stellt die Polymorphismen in den Typ-III-IFN-Genen, die die genetische Disposition zur Expression von IFN- λ_4 spiegeln, als aussagekräftige prognostische Wirtsfaktoren für die Tumorprogression des PDAC heraus. Hierbei können Aussagen über die progressionsfreien Intervalle, PFI bzw. PFI.2, und damit über die Lebensqualität sowie über die OS getroffen werden. Diese ersten Ergebnisse bedürfen weiterer Bestätigung durch voneinander unabhängige Studien. Die Art des Einflusses und eine Kausalität sind hierbei nicht bekannt und erfordern experimentelle, mechanistische Untersuchungen. Diese Arbeit liefert erste klinische Evidenz für eine prognostische Signifikanz von Polymorphismen in Typ-III-IFN-Genen des Wirtsgenoms, die die genetische Disposition zur Expression eines funktionsfähigen IFN- λ_4 repräsentieren, für die Progression des PDAC.

6 Anhang

6.1 Das Analysekonzept, die Datenressourcen und Software

Das Analysekonzept dieser Arbeit wurde übersichtlich tabellarisch dargestellt. Außerdem wurden die verwendeten Datenressourcen, statistischen Software und ihre jeweiligen Quellen zu Übersichtszwecken ebenfalls tabellarisch zusammengefasst.

Tab. A.2: Zusammenfassung der Datenressourcen und der Software

Datenressourcen	Datenportale	Quellen
a) Hinterlegte Daten		
Fallregistrierungs-/ Probenverarbeitungsdocumentation/ Zusammenfassung der gesammelten Datentypen	<i>National Cancer Institute</i>	https://www.cancer.gov/about-nci/organization/ccg/research/structural-genomics/tcga/using-tcga/types , Zugriff am 15.10.2019
TCGA-LIHC-Daten (377 Patienten; 10.832 Dateien)	<i>Genomic Data Commons data portal</i>	https://portal.gdc.cancer.gov/projects/TCGA-LIHC , Zugriff am 30.07.2019
TCGA-PAAD-Daten (185 Patienten; 5.307 Dateien)	<i>Genomic Data Commons data portal</i>	https://portal.gdc.cancer.gov/projects/TCGA-PAAD , Zugriff am 30.07.2019
<i>Tissue Source Site Codes</i>	<i>National Cancer Institute/ Genomic Data Commons</i>	https://gdc.cancer.gov/resources-tcga-users/tcga-code-tables/tissue-source-site-codes , Zugriff am 30.07.2019
TCGA-interne HCC-Methodik und Analyse	<i>The Cancer Genome Atlas Research Network</i>	http://dx.doi.org/10.1016/j.cell.2017.05.046 , Zugriff am 15.10.2019
TCGA-interne PDAC-Methodik und Analyse	<i>The Cancer Genome Atlas Research Network</i>	http://dx.doi.org/10.1016/j.jcell.2017.07.007 , Zugriff am 15.10.2019
<i>TCGA Pan-Cancer Clinical Data Resource (TCGA-CDR)</i>	<i>National Cancer Institute/ Genomic Data Commons</i>	https://gdc.cancer.gov/about-data/publications/pancanatlas , Zugriff am 03.06.2019
1000-Genome-Projekt (Phase 3, Version 5)	<i>Ensembl GRCh37 release 98</i>	http://grch37.ensembl.org/Homo_sapiens , Zugriff am 12.04.2019 https://www.internationalgenome.org/ , Zugriff am 22.01.2019
NCBI <i>Genome Workbench</i> (humanes Vergleichs-genom)	<i>National Center for Biotechnology Information Human Genome Reference Assembly</i>	https://www.ncbi.nlm.nih.gov/tools/gbench/ , Zugriff am 20.01.2019
b) Statistische Software		
R Version 3.5.2	<i>R Foundation for Statistical Computing</i>	https://www.r-project.org/about.html , Zugriff am 01.02.2019
HWE	<i>Laboratory of Immunogenomics and Immunoproteomics</i>	http://www.dr-petrek.eu/documents/HWE.xls , Zugriff am 13.01.2019
LDlink	<i>National Cancer Institute Division of Cancer Epidemiology & Genetics</i>	https://ldlink.nci.nih.gov/ , Zugriff am 20.02.2019
IGV Version 2.4.15	<i>Broad Institute</i>	https://software.broadinstitute.org/software/igv/ , Zugriff am 03.06.2019

6.2 Deskriptive und explorative Statistik

Um einen ersten Überblick über die Verteilung der Daten zu erhalten, wurde die Altersverteilung der HCC- und PDAC-Patienten zunächst grafisch durch Histogramme dargestellt. Zusätzlich wurden Boxplots und Säulendiagramme aus den klinischen Patientendaten erstellt.

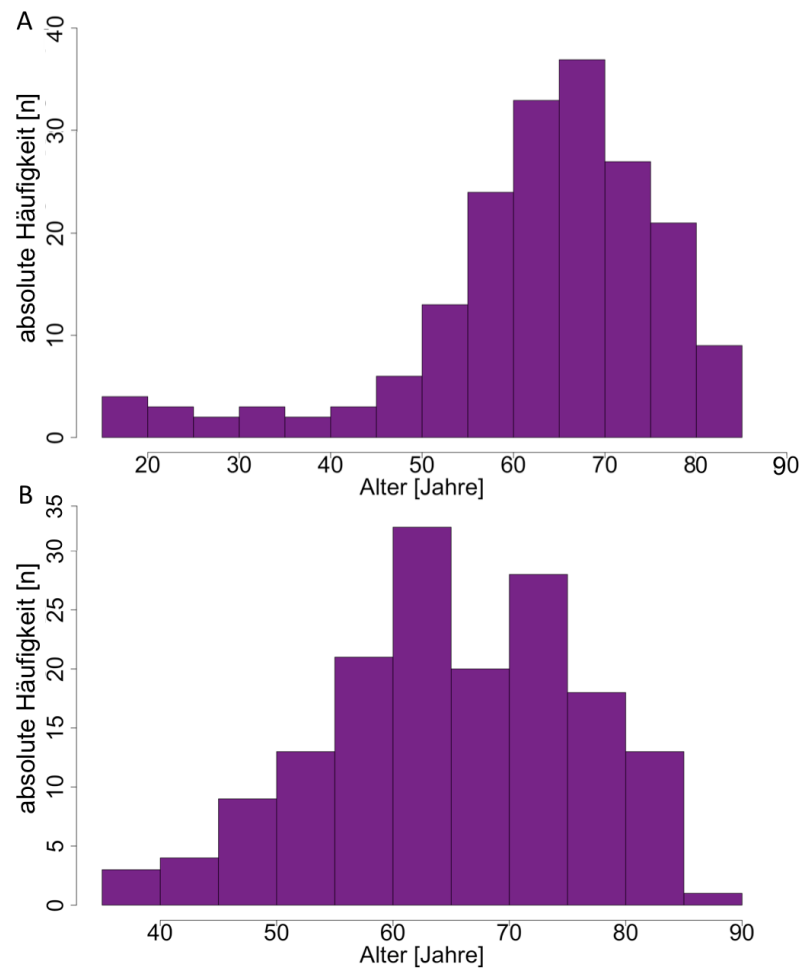


Abb. A.1: Altersverteilung der HCC- und der PDAC-Patienten bei Diagnose.

Das Histogramm für die HCC-Patienten (A) zeigt, dass das Alter bei Diagnose ein Maximum bei etwa 65–70 Jahren aufweist. Es wird deutlich, dass das HCC im höheren Alter eintritt als das PDAC (B). Die Altersspannbreite ist bei den HCC-Patienten größer (16–90 Jahre) als bei den PDAC-Patienten (35–88 Jahre).

Dem Histogramm (A) nach zu beurteilen liegt keine Normalverteilung des Alters bei den HCC-Patienten vor. Bei den PDAC-Patienten (B) scheint eine Normalverteilung des Alters vorzuliegen. Die Klassenbreite der Histogramme wurde mittels der Formel von Freedman und Diaconis (1981) ermittelt.

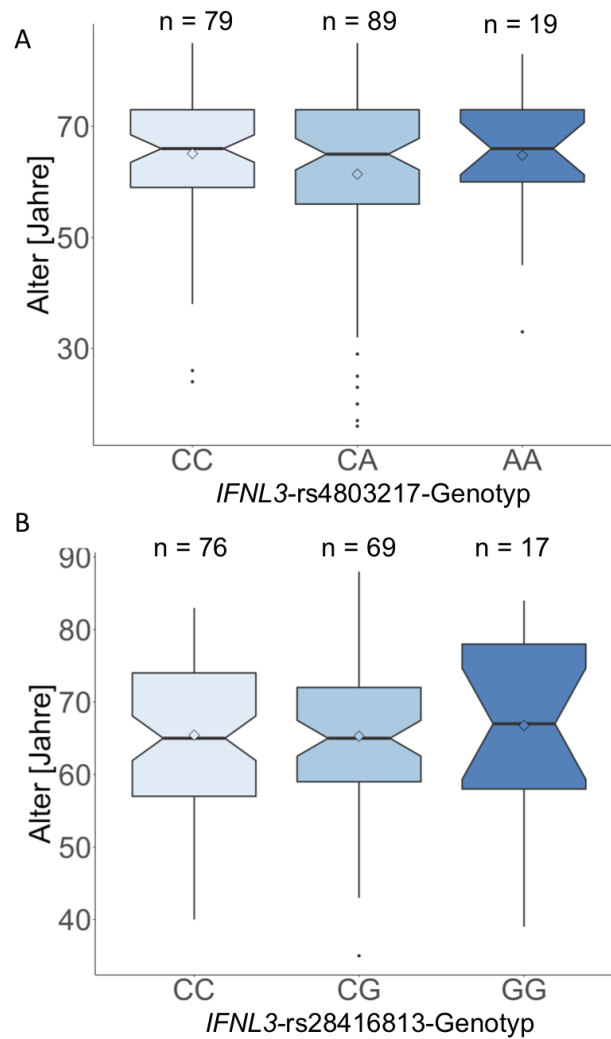


Abb. A.2: Altersverteilung der nach *IFNL3*-rs4803217- bzw. *IFNL3*-rs28416813-Genotypen stratifizierten HCC- bzw. PDAC-Patienten.

Dargestellt sind die Altersverteilungen in Form von Boxplots. Die Kästen der Boxplots stellen den Quartilsabstand mit dem Median als Mittellinie dar. Der Median liegt bei allen HCC-Patienten (A) bei etwa 65 Jahren. Der Median der PDAC-Patienten (B) mit den Genotypen CC und CG liegt auch bei etwa 65 Jahren und weicht bei Patienten mit dem Genotyp GG auf etwa 67 Jahre ab. Die Raute in den Boxen gibt den Mittelwert an und weicht bei den HCC-Patienten mit dem Genotyp CA mit etwa 61 Jahren am stärksten vom Median ab, wohingegen die Mittelwerte bei den PDAC-Patienten nicht stark von den Medianen abweichen. Die *whisker* decken das 1,5-fache des Quartilsabstands ab und alle extremeren Werte sind als Einzelpunkte dargestellt. Die Werte der HCC-Patienten mit dem Genotyp CA bzw. der PDAC-Patienten mit dem Genotyp GG weisen das größte Intervall auf. Insgesamt betrachtet ist die Altersverteilung zwischen den Strata weder bei den HCC- noch bei den PDAC-Patienten signifikant verschieden (Kruskal-Wallis-Test: $p(\text{HCC}) = 0,31$, $p(\text{PDAC}) = 0,98$).

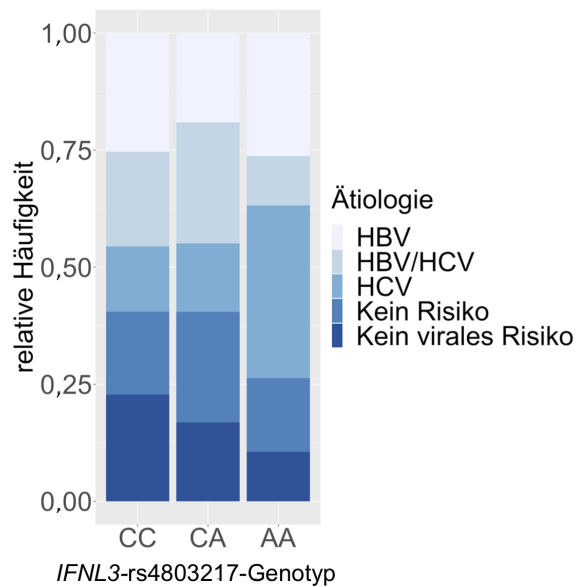


Abb. A.3: Verteilung der Anteile der nach *IFNL3*-rs4803217-Genotypen CC, CA, AA stratifizierten HCC-Patienten auf die benannten Risikofaktoren.

Dargestellt sind die relativen Häufigkeiten in Säulendiagrammen. Innerhalb der Patientengruppe mit dem *IFNL3*-rs4803217-Genotyp CC ist der Anteil der Patienten mit dem Risikofaktor einer HBV-Infektion mit etwa 25 % am größten. Der größte Anteil der Patienten mit einer Koinfektion des HBV und des HCV mit etwa 25 % liegt bei den heterozygoten Patienten vor und bei den Patienten mit dem Genotyp AA ist der Anteil der Patienten mit dem Risikofaktor einer HCV-Infektion mit etwa 40 % am größten. Insgesamt betrachtet ist die Risikoverteilung zwischen den Strata nicht signifikant verschieden (χ^2 -Test: $p(\text{HCC}) = 0,59$).

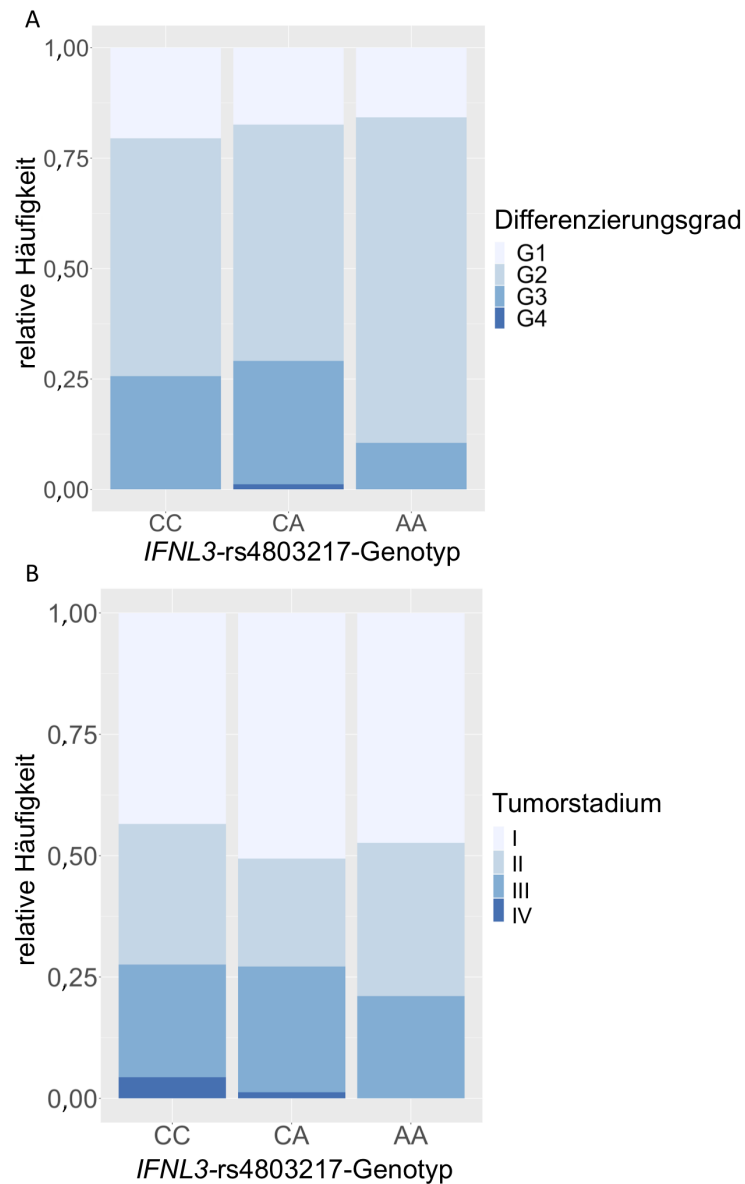


Abb. A.4: Verteilung der Anteile der nach *IFNL3*-rs4803217-Genotypen CC, CA, AA stratifizierten HCC-Patienten auf die klinischen Kriterien Tumordifferenzierungsgrad bzw. Tumorstadium.

Dargestellt sind die relativen Häufigkeiten in Säulendiagrammen. Alle HCC-Patienten unabhängig vom Genotyp weisen vorwiegend ein Differenzierungsgrad G2 (A) mit etwa 55–70 % und ein Tumorstadium I (B) mit 50 % auf. Insgesamt betrachtet ist die Verteilung der Kriterien zwischen den Strata nicht signifikant verschieden (χ^2 -Test: $p(\text{Grad}) = 0,87$, $p(\text{Stadium}) = 0,69$).

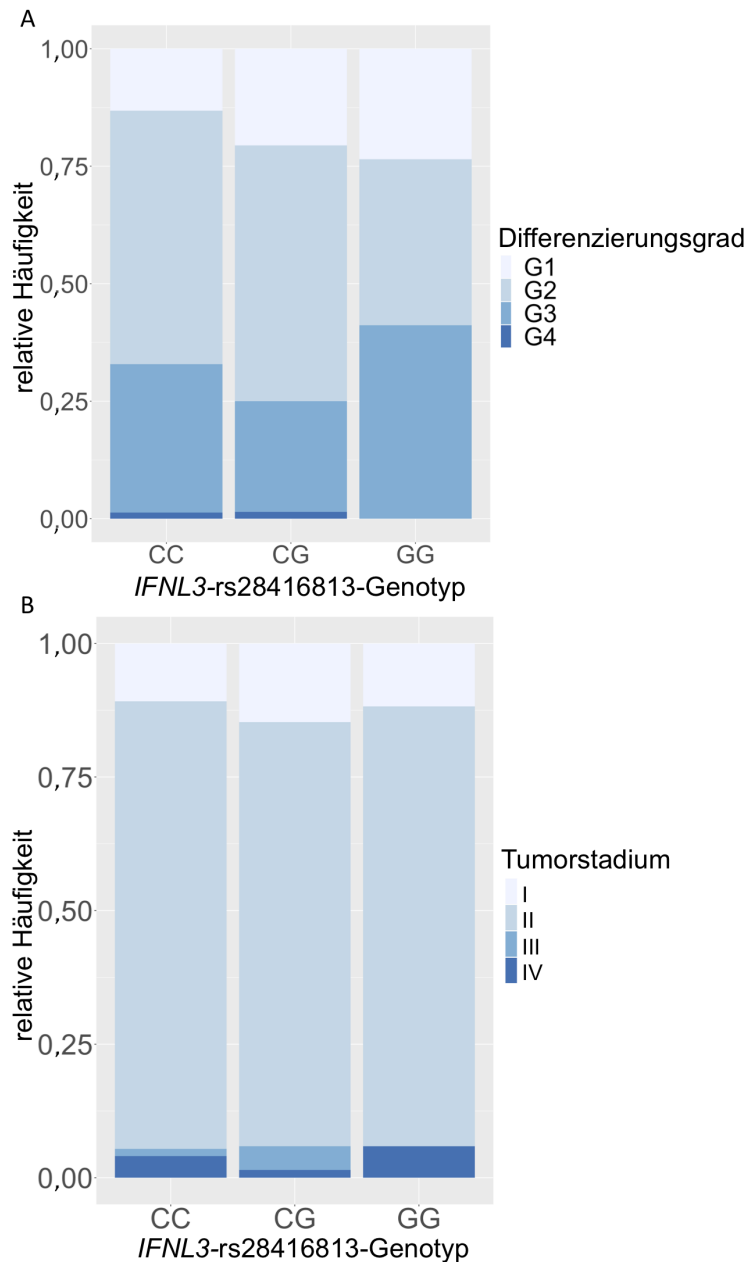


Abb. A.5: Verteilung der Anteile der nach *IFNL3*-rs28416813-Genotypen CC, CG, GG stratifizierten PDAC-Patienten auf die klinischen Kriterien Tumordifferenzierungsgrad bzw. Tumorstadium.

Dargestellt sind die relativen Häufigkeiten in Säulendiagrammen. Innerhalb der PDAC-Patientengruppe mit den *IFNL3*-rs28416813-Genotypen CC und CG ist der Anteil der Patienten mit einem Differenzierungsgrad G2 (A) mit etwa 50 % am größten. Bei den Patienten mit dem Genotyp GG liegt der größte Anteil der PDAC-Patienten mit nahezu 40 % innerhalb des Differenzierungsgrads G3 (A). Alle PDAC-Patienten unabhängig vom Genotyp weisen vorwiegend ein Tumorstadium II (B) mit etwa 80 % auf. Insgesamt betrachtet ist die Verteilung der Kriterien zwischen den Strata nicht signifikant verschieden (χ^2 -Test: $p(\text{Grad}) = 0,42$, exakter Fisher-Yates-Test: $p(\text{Stadium}) = 0,86$).

6.3 Stratifizierung der Daten des TCGA-LIHC-Kollektivs nach der Ätiologie

Aus den anamnestischen und serologischen Daten der TCGA-Datenbank für das TCGA-LIHC-Kollektiv konnten innerhalb dieser Arbeit verschiedene Risikofaktoren (mutmaßliche Ätiologie) unterschieden werden. Anschließend wurde das TCGA-LIHC-Kollektiv nach dieser Ätiologie stratifiziert. Hierbei zeigt sich, dass der größte Anteil aus HBV-infizierten HCC-Patienten besteht, die mit 57,5 Jahren bei Diagnose am jüngsten sind (Tab. A.3). Am ältesten sind die Patienten bei Diagnose mit einem HCC, welches nicht mit einem viralen Risiko, jedoch mit anderen Risikofaktoren, wie beispielsweise Alkoholkonsum, Ausprägung einer nicht-alkoholinduzierten Steatose oder Steatohepatitis und Lebertoxinexpositionen, assoziiert ist. Die Geschlechterverteilung entspricht in allen Risikostrata, bis auf dem Stratum „Kein Risiko“, etwa einem Verhältnis von 2,6 : 1 (m : f). Innerhalb des Stratums „Kein Risiko“ sind mehr Frauen als Männer vertreten (0,6 : 1 (m : f)) (Tab. A.3). Die Risikofaktoren sind innerhalb der Gruppen der Patienten europäischer Herkunft homogen verteilt. Unter den HCC-Patienten asiatischer Herkunft ist der Anteil der HBV-Infizierten am höchsten und unter den Patienten afrikanischer Herkunft überwiegt der Anteil der HCV-Infizierten (Tab. A.3). Im Durchschnitt weisen Patienten mit einem HBV-bedingten HCC einen höheren Differenzierungsgrad (G3) auf als alle anderen Patienten, deren HCC nicht HBV-bedingt ist (Tab. A.3). Das höhere Tumorstadium III weisen hingegen vor allem Patienten auf, bei denen gleichzeitig eine HBV- und HCV-Infektion vorkommt (Tab. A.3). Ein deutlicher Unterschied ist bei Betrachtung der Genotypverteilung zwischen den Patienten, die homozygot für das häufigere Allel C sind, und den Trägern des selteneren Allels A innerhalb der jeweiligen Strata erkennbar. Bei den HCC-Patienten, bei denen kein bzw. kein viraler Risikofaktor detektiert werden konnte sowie bei Patienten mit einer HBV-/HCV-Koinfektion, sind mehr Allel-C-Homozygote als Träger des Allels A vertreten. Auffällig ist, dass bei Patienten mit einem HBV-bedingten HCC die Allel-C-Homozygoten überproportional häufig vertreten sind. Das vermehrte Auftreten der Allel-C-Homozygoten unter den HBV-infizierten HCC-Patienten lässt sich damit erklären, dass vorwiegend Patienten asiatischer Herkunft eine HBV-Infektion aufweisen und gleichzeitig ihre MAF für das Allel A sehr gering ist. Des Weiteren sind die HCV-infizierten HCC Patienten 1,4-fach so häufig Träger des Allels A. Erklärbar ist der höhere Anteil der Träger des Allels A mit der bereits bekannten Tatsache, dass die Expression eines funktionsfähigen IFN- λ_4 schlecht für die spontane und therapieinduzierte Ausheilung einer HCV-Infektion ist und wahrscheinlich vermehrt zu einem HCC führt (Ge et al. 2009). Die Genotypverteilungen innerhalb der Strata befinden sich im HWE ($p > 0,05$).

Tab. A.3: Stratifizierung des TCGA-LIHC-Kollektivs nach der Ätiologie

	Total <i>n</i> = 377	Kein Risiko <i>n</i> = 54	Kein virales Risiko <i>n</i> = 43	HBV <i>n</i> = 144	HCV <i>n</i> = 48	HBV/HCV <i>n</i> = 88
¹ Alter bei Diagnose M ± SD	59,5 ± 13,5	61,1 ± 17,1	67,0 ± 9,5	57,5 ± 13,0	60,1 ± 8,6	57,7 ± 14,4
Geschlecht m/f [n(m %)]	255 (67,6)/122	20 (37,0)/34	31 (72,1)/12	104 (72,2)/40	38 (79,2)/10	62 (70,5)/26
² Ethnizität [n(%)]						
Europäische Herkunft	187 (51,0)	38 (74,5)	35 (83,3)	42 (29,6)	31 (66,0)	41 (48,2)
Asiatische Herkunft	161 (43,9)	11 (21,6)	5 (11,9)	97 (68,3)	7 (14,9)	41 (48,2)
Afrikanische Herkunft	17 (4,6)	1 (2,0)	2 (4,8)	3 (2,1)	8 (17,0)	3 (3,5)
Amerikanische Ureinwohner	2 (0,5)	1 (2,0)			1 (2,1)	
³ Differenzierungsgrad [n(%)]						
G1	55 (14,8)	8 (15,4)	6 (14,0)	11 (7,6)	10 (21,7)	20 (23,0)
G2	180 (48,4)	30 (57,7)	22 (51,2)	55 (38,2)	26 (56,5)	47 (54,0)
G3	124 (33,3)	13 (25,0)	15 (34,9)	67 (46,5)	10 (21,7)	19 (21,8)
G4	13 (3,5)	1 (1,9)		11 (7,6)		1 (1,1)
⁴ Stadium der Tumorerkrankung [n(%)]						
I	175 (49,6)	24 (50,0)	15 (44,1)	84 (59,6)	26 (59,1)	26 (30,2)
II	87 (24,6)	8 (16,7)	11 (32,4)	35 (24,8)	12 (27,3)	21 (24,4)
III	86 (24,4)	15 (31,3)	7 (20,6)	19 (13,5)	6 (13,6)	39 (45,3)
IV	5 (1,4)	1 (2,1)	1 (2,9)	3 (2,1)		
⁵ <i>IFNL3</i> rs4803217 [C/A]						
CC:CA:AA	227:123:26	29:21:4	25:16:2	100:35:8	19:19:10	54:32:2
MAF	0,233	0,269	0,233	0,178	0,406	0,205
¹ Altersangaben waren für 376/377 Patienten verfügbar						
² Angaben zur Ethnizität waren für 367/377 Patienten verfügbar						
³ Daten zum Tumordifferenzierungsgrad waren für 372/377 Patienten verfügbar						
⁴ Daten zum Stadium der Tumorerkrankungen waren für 353/377 Patienten verfügbar						
⁵ Genotypen konnten für 376/377 Patienten ausgelesen werden						

6.4 Mediane ereignisfreie Zeiten und mediane Beobachtungszeiträume

Die medianen ereignisfreien Zeiten und die medianen Beobachtungszeiträume wurden innerhalb Tab. A.4 für die HCC-Patienten und innerhalb Tab. A.5 für die PDAC-Patienten von der Diagnose bis zum Eintritt des Ereignisses oder der Zensierung für alle drei Ereigniszeiten (PFI/PFI.2/OS) in Monaten angegeben. Dabei wurde zwischen dem Beobachtungszeitraum für jeweils das gesamte Kollektiv und für die Patienten mit Ereignis bzw. diejenigen, die zensiert wurden, unterschieden (Gesamt/Ereignis/Zensiert) (Tab. A.4, Tab. A.5). Mit dieser Darstellung sollte festgestellt werden, ob die Beobachtungszeiträume vergleichbar sind. Bei den Analysen nach dem Kaplan-Meier-Verfahren ist hauptsächlich die Betrachtung des Beobachtungszeitraums sowie die Anzahl der Ereigniseintritte und nicht die Betrachtung der Anzahl der Patienten entscheidend.

In dieser Arbeit soll zur Abschätzung der Suffizienz der Beobachtungszeit nach Kriterien aus dem TCGA-CDR-Datensatz vorgegangen werden. Das TCGA-CDR-Netzwerk formuliert unter anderem die Bedingung, dass die Kaplan-Meier-Kurven keine großen Sprünge bei den Zensierungen aufweisen und in dem Intervall von dem letzten Ereignis bis zum Beobachtungsende konstant oder flach seien, sozusagen in einem *steady state* enden (Liu et al. 2018). Des Weiteren soll die maximale und mediane Ereigniszeit kleiner sein, als die maximale und mediane zensierte Zeit, damit eine ausreichend lange Beobachtungszeit auch noch nach Ereigniseintritt gewährleistet ist. Als letztes wird von den Daten erwartet, dass eine Ereignisrate von $\geq 30\%$ vorliegt und dass die Bedingung nach Peduzzi et al. (1995) erfüllt sei, die besagt, dass für jedes getestete Merkmal ≥ 10 Ereignisse eingetreten sein müssen (Liu et al. 2018).

Im Rahmen dieser Arbeit wurden die gestellten Anforderungen größtenteils erfüllt.

Tab. A.4: Übersicht über die medianen ereignisfreien Zeiten und die medianen Beobachtungszeiträume der HCC-Patienten für alle drei Ereigniszeiten

	Total		CC		CA		AA		CA/AA	
	Gesamt Ereignis Zensiert (n = 187 n = 103 n = 84)	Gesamt Ereignis Zensiert (n = 79 n = 42 n = 37)	Gesamt Ereignis Zensiert (n = 89 n = 49 n = 40)	Gesamt Ereignis Zensiert (n = 19 n = 12 n = 7)	Gesamt Ereignis Zensiert (n = 89 n = 49 n = 40)	Gesamt Ereignis Zensiert (n = 19 n = 12 n = 7)	Gesamt Ereignis Zensiert (n = 89 n = 49 n = 40)	Gesamt Ereignis Zensiert (n = 19 n = 12 n = 7)	Gesamt Ereignis Zensiert (n = 108 n = 61 n = 47)	Gesamt Ereignis Zensiert (n = 108 n = 61 n = 47)
¹ Progressionsfreies Intervall (PFI)	586 19,3	561 18,4	592 19,5	452 14,8	592 19,5	452 14,8	592 19,5	452 14,8	592 19,5	452 14,8
Mediane ereignisfreie Zeit (gesamt) [Tage Monate]	11,9 10,6 12,9	11,3 9,6 12,7	12,7 11,8 13,2	8,3 7,9 12,0	12,7 11,8 13,2	8,3 7,9 12,0	12,7 11,8 13,2	8,3 7,9 12,0	12,1 11,2 13,0	8,3 7,9 12,0
Medianer Beobachtungszeitraum [Monate]										
² Spezifisches progressionsfreies Intervall (PFI.2)	598 19,7	561 18,4	639 21,0	452 14,9	639 21,0	452 14,9	639 21,0	452 14,9	637 20,9	452 14,9
Mediane ereignisfreie Zeit (gesamt) [Tage Monate]	12,0 10,5 13,5	11,3 9,6 12,7	13,1 11,5 14,4	8,3 7,9 12,0	13,1 11,5 14,4	8,3 7,9 12,0	13,1 11,5 14,4	8,3 7,9 12,0	12,5 10,8 14,0	8,3 7,9 12,0
Medianer Beobachtungszeitraum [Monate]										
³ Gesamtüberlebensdauer (OS)	1397 45,9	1423 46,8	1372 45,1	1271 41,8	1372 45,1	1271 41,8	1372 45,1	1271 41,8	1372 45,1	1271 41,8
Mediane ereignisfreie Zeit (gesamt) [Tage Monate]	22,1 22,6 22,1	20,6 18,7 22,4	23,8 26,5 21,8	18,3 13,7 25,4	23,8 26,5 21,8	18,3 13,7 25,4	23,8 26,5 21,8	18,3 13,7 25,4	22,9 25,2 22,1	18,3 13,7 25,4
Medianer Beobachtungszeitraum [Monate]										

¹PFI-Progressionsfreies Intervall: Zeitspanne von der Diagnose bis zur Progression der Tumorerkrankung, bis zum Auftreten eines weiteren Tumors jeglicher Entität oder bis zum Tod durch einen Tumor jeglicher Entität

²PFI.2-Spezifisches progressionsfreies Intervall: Zeitspanne von der Diagnose bis zur Progression der ursprünglichen Tumorerkrankung oder bis zum Tod durch selbige Tumorentität

³OS-Gesamtüberlebensdauer: Zeitspanne von der Diagnose bis zum Tod durch jegliche Ursache

Tab. A.5: Übersicht über die medianen ereignisfreien Zeiten und die medianen Beobachtungszeiträume der PDAC-Patienten für alle drei Ereigniszeiten

	Total		CC		CG		GG		CG/GG	
	Gesamt Ereignis Zensiert (n = 162 n = 93 n = 67)	Gesamt Ereignis Zensiert (n = 76 n = 52 n = 24)	Gesamt Ereignis Zensiert (n = 69 n = 34 n = 35)	Gesamt Ereignis Zensiert (n = 17 n = 9 n = 8)	Gesamt Ereignis Zensiert (n = 69 n = 34 n = 35)	Gesamt Ereignis Zensiert (n = 17 n = 9 n = 8)	Gesamt Ereignis Zensiert (n = 69 n = 34 n = 35)	Gesamt Ereignis Zensiert (n = 17 n = 9 n = 8)	Gesamt Ereignis Zensiert (n = 86 n = 43 n = 43)	Gesamt Ereignis Zensiert (n = 86 n = 43 n = 43)
¹ Progressionsfreies Intervall (PFI)										
Mediane ereignisfreie Zeit (gesamt) [Tag Monat]	486 16,0	378 12,4	593 19,5	545 17,9	593 19,5	545 17,9	593 19,5	545 17,9	593 19,5	593 19,5
Medianer Beobachtungszeitraum [Monat]	12,0 10,1 15,2	11,3 9,6 15,7	13,8 11,8 12,8	14,8 12,0 23,2	13,8 11,8 12,8	14,8 12,0 23,2	13,8 11,8 12,8	14,8 12,0 23,2	12,5 11,9 13,7	12,5 11,9 13,7
Total										
	Gesamt Ereignis Zensiert (n = 162 n = 93 n = 69)	Gesamt Ereignis Zensiert (n = 76 n = 50 n = 26)	Gesamt Ereignis Zensiert (n = 69 n = 34 n = 35)	Gesamt Ereignis Zensiert (n = 17 n = 9 n = 8)	Gesamt Ereignis Zensiert (n = 69 n = 34 n = 35)	Gesamt Ereignis Zensiert (n = 17 n = 9 n = 8)	Gesamt Ereignis Zensiert (n = 69 n = 34 n = 35)	Gesamt Ereignis Zensiert (n = 17 n = 9 n = 8)	Gesamt Ereignis Zensiert (n = 86 n = 43 n = 43)	Gesamt Ereignis Zensiert (n = 86 n = 43 n = 43)
² Spezifisches progressionsfreies Intervall (PFI.2)										
Mediane ereignisfreie Zeit (gesamt) [Tag Monat]	492 16,2	390 12,8	593 19,5	545 17,9	593 19,5	545 17,9	593 19,5	545 17,9	593 19,5	593 19,5
Medianer Beobachtungszeitraum [Monat]	12,1 10,1 15,2	12,0 9,6 15,7	12,3 11,8 12,8	14,8 12,0 23,2	12,3 11,8 12,8	14,8 12,0 23,2	12,3 11,8 12,8	14,8 12,0 23,2	12,5 11,9 13,7	12,5 11,9 13,7
Total										
	Gesamt Ereignis Zensiert (n = 162 n = 86 n = 76)	Gesamt Ereignis Zensiert (n = 76 n = 47 n = 29)	Gesamt Ereignis Zensiert (n = 69 n = 31 n = 38)	Gesamt Ereignis Zensiert (n = 17 n = 8 n = 9)	Gesamt Ereignis Zensiert (n = 69 n = 31 n = 38)	Gesamt Ereignis Zensiert (n = 17 n = 8 n = 9)	Gesamt Ereignis Zensiert (n = 69 n = 31 n = 38)	Gesamt Ereignis Zensiert (n = 17 n = 8 n = 9)	Gesamt Ereignis Zensiert (n = 86 n = 39 n = 47)	Gesamt Ereignis Zensiert (n = 86 n = 39 n = 47)
³ Gesamtüberlebensdauer (OS)										
Mediane ereignisfreie Zeit (gesamt) [Tag Monat]	614 20,2	517 17,0	738 24,3	614 20,2	738 24,3	614 20,2	738 24,3	614 20,2	738 24,3	738 24,3
Medianer Beobachtungszeitraum [Monat]	15,3 12,7 16,7	15,3 14,1 15,9	13,8 12,0 16,4	17,9 13,6 24,2	13,8 12,0 16,4	17,9 13,6 24,2	13,8 12,0 16,4	17,9 13,6 24,2	15,2 12,0 19,2	15,2 12,0 19,2

¹PFI-Progressionsfreies Intervall: Zeitspanne von der Diagnose bis zur Progression der Tumorerkrankung, bis zum Auftreten eines weiteren Tumors, jeglicher Entität oder bis zum Tod durch einen Tumor jeglicher Entität

²PFI.2-Spezifisches progressionsfreies Intervall: Zeitspanne von der Diagnose bis zur Progression der ursprünglichen Tumorerkrankung oder bis zum Tod durch selbige Tumorentität

³OS-Gesamtüberlebensdauer: Zeitspanne von der Diagnose bis zum Tod durch jegliche Ursache

6.5 Ereigniszeitanalyse des spezifischen progressionsfreien Intervalls (PFI.2) nach dem Kaplan-Meier-Verfahren für nach klinischen Merkmalen stratifizierte HCC- und PDAC-Patienten

Durch die multiplen Cox-Regressionsanalysen konnten die freien Kriterien mit einem unabhängigen signifikanten Einfluss auf das PFI.2 ermittelt werden. Hierbei sind die ermittelten Einflüsse der freien Kriterien auf das abhängige Kriterium (PFI.2) für alle anderen betrachteten Einflussgrößen adjustiert. Diese einzelnen signifikanten Einflüsse wurden grafisch durch Kaplan-Meier-Kurven dargestellt.

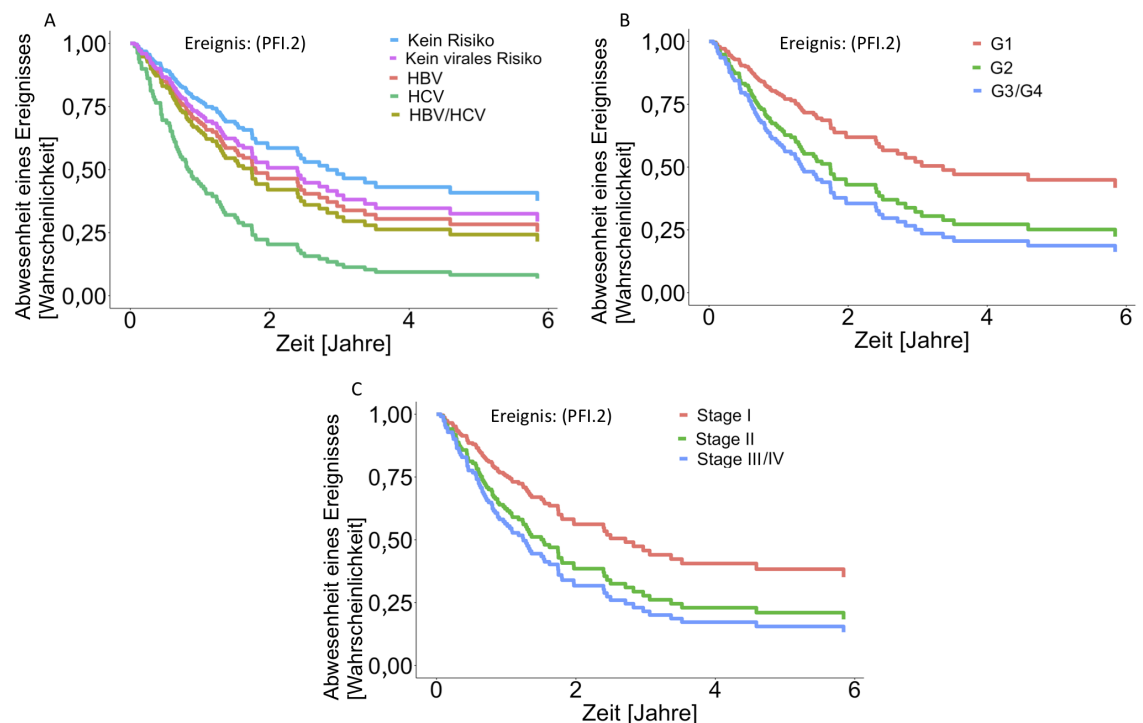


Abb. A.6: Ereigniszeitanalyse des spezifischen progressionsfreien Intervalls (PFI.2) nach dem Kaplan-Meier-Verfahren für nach klinischen Merkmalen stratifizierte HCC-Patienten.

Dargestellt sind in Abhängigkeit von der Zeit die Auswertungen für die Merkmale mit einem von den anderen getesteten Merkmalen unabhängigen, signifikanten Einfluss auf das PFI.2: Risikofaktoren (A), Tumordifferenzierungsgrad (B) und Tumorstadium (C).

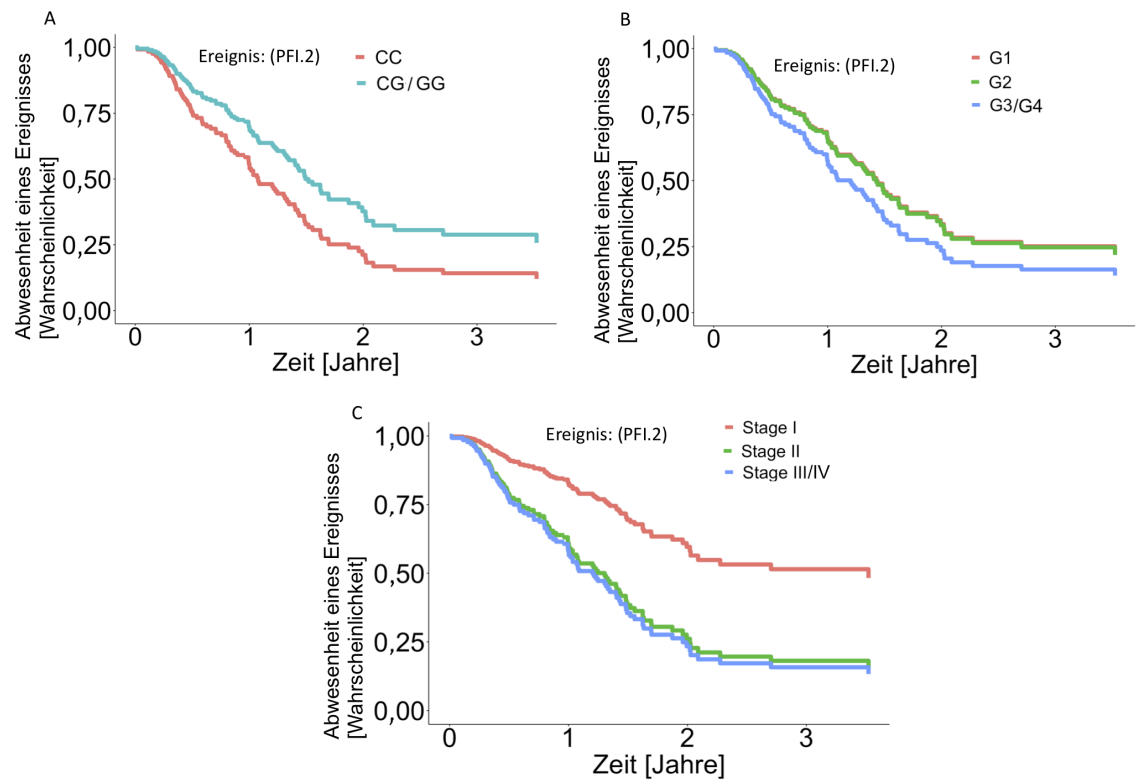


Abb. A.7: Ereigniszeitanalyse des spezifischen progressionsfreien Intervalls (PFI.2) nach dem Kaplan-Meier-Verfahren für nach klinischen und genetischen Merkmalen stratifizierte PDAC-Patienten.

Dargestellt sind in Abhängigkeit von der Zeit die Auswertungen für die Merkmale mit einem von den anderen getesteten Merkmalen unabhängigen, signifikanten Einfluss auf das PFI.2: Genotyp (*IFNL3* rs28416813) (A), Tumordifferenzierungsgrad (B) und Tumorstadium (C).

7 Literaturverzeichnis

- Abushahba W, Balan M, Castaneda I, Yuan Y, Reuhl K, Raveche E, de la Torre A, Lasfar A, Kotenko SV (2010): Antitumor activity of type I and type III interferons in BNL hepatoma model. *Cancer Immunol Immunother* 59, 1059–71
- Adamska A, Domenichini A, Falasca M (2017): Pancreatic ductal adenocarcinoma: Current and evolving therapies. *Int J Mol Sci* 18, e1338
- Ally A, Balasundaram M, Carlsen R, Chuah E, Clarke A, Dhalla N, Holt RA, Jones SJM, Lee D, Ma Y et al. (2017): Comprehensive and integrative genomic characterization of hepatocellular carcinoma. *Cell* 169, 1327–41
- Altshuler D, Durbin RM, Abecasis GR, Bentley DR, Chakravarti A, Clark AG, Collins FS, de la Vega FM, Donnelly P, Egholm M et al. (2010): A map of human genome variation from population-scale sequencing. *Nature* 467, 1061–73
- Altshuler DM, Durbin RM, Abecasis GR, Bentley DR, Chakravarti A, Clark AG, Donnelly P, Eichler EE, Flicek P, Gabriel SB et al. (2012): An integrated map of genetic variation from 1,092 human genomes. *Nature* 491, 56–65
- Ank N, West H, Bartholdy C, Eriksson K, Thomsen AR, Paludan SR (2006): Lambda interferon (IFN- λ), a type III IFN, is induced by viruses and IFNs and displays potent antiviral activity against select virus infections in vivo. *J Virol* 80, 4501–9
- Auton A, Abecasis GR, Altshuler DM, Durbin RM, Abecasis GR, Bentley DR, Chakravarti A, Clark AG, Donnelly P, Eichler EE et al. (2015): A global reference for human genetic variation. *Nature* 526, 68–74
- Bailey P, Chang DK, Nones K, Johns AL, Patch A, Gingras M, Miller KD, Christ AN, Bruxner TJC, Quinn MC et al. (2016): Genomic analyses identify molecular subtypes of pancreatic cancer. *Nature* 531, 47–52
- Bamford CGG, Aranday-Cortes E, Filipe IC, Sukumar S, Mair D, da Silva Filipe A, Mendoza JL, Garcia KC, Fan S, Tishkoff SA et al. (2018): A polymorphic residue that attenuates the antiviral potential of interferon lambda 4 in hominid lineages. *PLoS Pathog* 14, e1007307
- Bibert S, Wojtowicz A, Taffé P, Tarr PE, Bernasconi E, Furrer H, Günthard HF, Hoffmann M, Kaiser L, Osthoff M et al. (2018): Interferon lambda 3/4 polymorphisms are associated with AIDS-related Kaposi’s sarcoma. *AIDS* 32, 2759–65
- Bickeböller H, Fischer C: Einführung in die Genetische Epidemiologie. Springer Verlag, Berlin 2007
- Booy S, Hofland L, van Eijck C (2015): Potentials of interferon therapy in the treatment of pancreatic cancer. *J Interferon Cytokine Res* 35, 327–39

- Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A (2018): Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin* 68, 394–424
- Carlson CS, Eberle MA, Rieder MJ, Yi Q, Kruglyak L, Nickerson DA (2004): Selecting a maximally informative set of single-nucleotide polymorphisms for association analyses using linkage disequilibrium. *Am J Hum Genet* 74, 106–20
- Chinnaswamy S, Chatterjee S, Boopathi R, Mukherjee S, Bhattacharjee S, Kundu TK (2013): A single nucleotide polymorphism associated with hepatitis C virus infections located in the distal region of the IL28B promoter influences NF- κ B-mediated gene transcription. *PLoS One* 8, e75495
- Collins FS, Morgan M, Patrions A (2003): The human genome project: Lessons from large-scale biology. *Science* 300, 286–90
- Conroy T, Hammel P, Hebbar M, Ben Abdelghani M, Wei AC, Raoul J, Chone L, Francois E, Artru P, Biagi JJ et al. (2018): FOLFIRINOX or gemcitabine as adjuvant therapy for pancreatic cancer. *N Engl J Med* 379, 2395–406
- Derwinger K, Kodeda K, Bexe-Lindskog E, Taflin H (2010): Tumour differentiation grade is associated with TNM staging and the risk of node metastasis in colorectal cancer. *Acta Oncol* 49, 57–62
- Desai NV, Sliesoraitis S, Hughes SJ, Trevino JG, Zlotecki RA, Ivey AM, George TJ (2015): Multidisciplinary neoadjuvant management for potentially curable pancreatic cancer. *Cancer Med* 4, 1224–39
- DKG/DKH (2013): Diagnostik und Therapie des hepatozellulären Karzinoms (Leitlinienprogramm Onkologie) S3-Leitlinie der Deutschen Krebsgesellschaft/ Deutschen Krebshilfe.
<http://leitlinienprogramm-onkologie.de/Leitlinien70.html> (Zugriff am 16.04.2019)
- Ducreux M, Cuhna AS, Caramella C, Hollebecque A, Burtin P, Goere D, Seufferlein T, Haustermans K, van Laethem JL, Conroy T et al. (2015): Cancer of the pancreas: ESMO clinical practice guidelines for diagnosis, treatment and follow-up. *Ann Oncol* 26, 56–68
- Dudley WN, Wickham R, Coombs N (2016): An introduction to survival statistics: Kaplan-Meier analysis. *J Adv Pract Oncol* 7, 91–100
- Edge SB, Byrd DR, Compton CC, Fritz AG, Greene FL, Trotti A: *AJCC cancer staging manual*. 7. Auflage. Springer Verlag, Chicago 2010
- Egli A, Santer DM, O’Shea D, Barakat K, Syedbasha M, Vollmer M, Baluch A, Bhat R, Groenendyk J, Joyce MA et al. (2014a): IL-28B is a key regulator of B- and T-cell vaccine responses against influenza. *PLoS Pathog* 10, e1004556

- Egli A, Levin A, Santer DM, Joyce M, O’Shea D, Thomas BS, Lisboa LF, Barakat K, Bhat R, Fischer KP et al. (2014b): Immunomodulatory function of interleukin 28B during primary infection with cytomegalovirus. *J Infect Dis* 210, 717–27
- El-Serag HB (2011): Hepatocellular carcinoma. *N Engl J Med* 365, 1118–27
- El-Serag HB, Rudolph KL (2007): Hepatocellular carcinoma: Epidemiology and molecular carcinogenesis. *Gastroenterology* 132, 2557–76
- Eslam M, McLeod D, Kelaeng KS, Mangia A, Berg T, Thabet K, Irving WL, Dore GJ, Sheridan D, Grønbaek H et al. (2017): IFN- λ 3, not IFN- λ 4, likely mediates IFNL3-IFNL4 haplotype-dependent hepatic inflammation and fibrosis. *Nat Genet* 49, 795–800
- Fares J, Fares MY, Fares Y (2019): Immune checkpoint inhibitors: Advances and impact in neuro-oncology. *Surg Neurol Int* 10, 1–5
- Fedewa SA, Sauer AG, Siegel RL, Jemal A (2015): Prevalence of major risk factors and use of screening tests for cancer in the United States. *Cancer Epidemiol Biomarkers Prev* 24, 637–52
- Forner A, Reig ME, de Lope CR, Bruix J (2010): Current strategy for staging and treatment: The BCLC update and future prospects. *Semin Liver Dis* 30, 61–74
- Freedman D, Diaconis P (1981): On the histogram as a density estimator: L 2 theory. *Z Wahrscheinlichkeitstheor Verw Geb* 57, 453–76
- Gaus W, Muche R: *Medizinische Statistik: Angewandte Biometrie für Ärzte und Gesundheitsberufe*. 2. Auflage. Schattauer, Stuttgart 2017
- Ge D, Fellay J, Thompson AJ, Simon JS, Shianna KV, Urban TJ (2009): Genetic variation in IL28B predicts hepatitis C treatment-induced viral clearance. *Nature* 461, 399–401
- Geier A, Schadde E, Müllhaupt B (2012): Hepatozelluläres Karzinom – Früherkennung, Staging und aktuelle Therapie. *MMW Fortschr Med* 154, 63–7
- Grambsch PM, Therneau TM (1994): Proportional hazards tests and diagnostics based on weighted residuals. *Biometrika* 81, 515–26
- Greten TF, Papendorf F, Bleck JS, Kirchhoff T, Wohlberedt T, Kubicka S, Klempnauer J, Galanski M, Manns MP (2005): Survival rate in patients with hepatocellular carcinoma: A retrospective analysis of 389 patients. *Br J Cancer* 92, 1862–8
- Grossman RL, Heath AP, Ferretti V, Varmus HE, Lowy DR, Kibbe WA, Staudt LM (2016): Toward a shared vision for cancer genomic data. *N Engl J Med* 375, 1109–12
- Haeberle L, Esposito I (2019): Pathology of pancreatic cancer. *Transl Gastroenterol Hepatol* 4, 1–12
- Hanahan D, Weinberg RA (2011): Hallmarks of cancer: The next generation. *Cell* 144, 646–74

- Hardy GH (1908): Mendelian proportions in a mixed population. *Science* 28, 49–50
- Heim MH (2013): Innate immunity and HCV. *J Hepatol* 58, 564–74
- Hong M, Schwerk J, Lim C, Kell A, Jarret A, Pangallo J, Loo Y, Liu S, Hagedorn CH, Gale M et al. (2016): Interferon lambda 4 expression is suppressed by the host during viral infection. *J Exp Med* 213, 2539–52
- Isaacs A, Lindemann J (1957): Virus interference I. The interferon. *Proc R Soc Lond B Biol Sci* 147, 258–67
- Isaacs A, Lindemann J, Valentine RC (1957): Virus interference II. Some properties of interferon. *Proc R Soc Lond B Biol Sci* 147, 268–73
- Key FM, Peter B, Dennis MY, Huerta-Sanchez E, Tang W, Prokunina-Olsson L, Nielsen R, Andres AM (2014): Selection on a variant associated with improved viral clearance drives local, adaptive pseudogenization of interferon lambda 4 (IFNL4). *PLoS Genet* 10, e1004681
- Kotenko SV, Gallagher G, Baurin VV, Lewis-Antes A, Shen M, Shah NK, Langer JA, Sheikh F, Dickensheets H, Donnelly RP (2003): IFN- λ s mediate antiviral protection through a distinct class II cytokine receptor complex. *Nat Immunol* 4, 69–77
- Kulaeva OI, Draghici S, Tang L, Kraniak JM, Land SJ, Tainsky MA (2003): Epigenetic silencing of multiple interferon pathway genes after cellular immortalization. *Oncogene* 22, 4118–27
- Lasfar A, Lewis-Antes A, Smirnov SV, Anantha S, Abushahba W, Tian B, Reuhl K, Dickensheets H, Sheikh F, Donnelly RP et al. (2006): Characterization of the mouse IFN-lambda ligand-receptor system: IFN-lambdas exhibit antitumor activity against B16 melanoma. *Cancer Res* 66, 4468–77
- Lasfar A, Abushahba W, de la Torre A, Castaneda I, Kotenko SV, Reuhl K (2008): IFN therapy in TIB 75 HCC model: Combination of IFN-lambda and IFN-alpha induces complete remission. *Hepatology* 48, 191
- Lasfar A, Abushahba W, Balan M, Cohen-Solal KA (2011): Interferon lambda: A new sword in cancer immunotherapy. *Clin Dev Immunol* 2011, 1–11
- Lasfar A, Zloza A, Silk AW, Lee LY, Cohen-Solal KA (2019): Interferon lambda: Toward a dual role in cancer. *J Interferon Cytokine Res* 39, 22–9
- Lázár-Molnár E, Scandiuzzi L, Basu I, Quinn T, Sylvestre E, Palmieri E, Ramagopal UA, Nathenson SG, Guha C, Almo SC (2017): Structure-guided development of a high-affinity human programmed cell death-1: Implications for tumor immunotherapy. *EBioMedicine* 17, 30–44
- Lazear HM, Nice TJ, Diamond MS (2015): Interferon- λ : Immune functions at barrier surfaces and beyond. *Immunity* 43, 15–28

- Lee S, Lee E, Kim S, Jeong P, Cho Y, Yun SJ, Kim S, Kim G, Choi YH, Cha E et al. (2012): Identification of pro-inflammatory cytokines associated with muscle invasive bladder cancer; The roles of IL-5, IL-20, and IL-28A. *PLoS One* 7, e40267
- Li D, Sedano S, Allen R, Gong J, Cho M, Sharma S (2019): Current treatment landscape for advanced hepatocellular carcinoma: Patient outcomes and the impact on quality of life. *Cancers* 11, 841–59
- Li Q, Kawamura K, Ma G, Iwata F, Numasaki M, Suzuki N, Shimada H, Tagawa M (2010): Interferon-lambda induces G1 phase arrest or apoptosis in oesophageal carcinoma cells and produces anti-tumour effects in combination with anti-cancer agents. *Eur J Cancer* 46, 180–90
- Liu J, Lichtenberg T, Hoadley KA, Poisson LM, Lazar AJ, Cherniack AD, Kovatich AJ, Benz CC, Levine DA, Lee AV et al. (2018): An integrated TCGA pan-cancer clinical data resource to drive high-quality survival outcome analytics. *Cell* 73, 400–16
- Maher SG, Sheikh F, Scarzello AJ, Romero-Weaver AL, Baker DP, Donnelly RP, Gamero AM (2008): IFN- α and IFN- λ differ in their antiproliferative effects and duration of JAK/STAT signaling activity. *Cancer Biol Ther* 7, 1109–15
- Manuel O, Wojtowicz A, Bibert S, Mueller NJ, van Delden C, Hirsch HH, Steiger J, Stern M, Egli A, Garzoni C et al. (2015): Influence of IFNL3/4 polymorphisms on the incidence of cytomegalovirus infection after solid-organ transplantation. *J Infect Dis* 211, 906–14
- Marrero JA, Kulik LM, Sirlin CB, Zhu AX, Finn RS, Abecassis MM, Roberts LR, Heimbach JK (2018): Diagnosis, staging, and management of hepatocellular carcinoma: 2018 practice guidance by the American association for the study of liver diseases. *Hepatology* 68, 723–50
- Martins-Filho SN, Paiva C, Azevedo RS, Alves VAF (2017): Histological grading of hepatocellular carcinoma - A systematic review of literature. *Front Med (Lausanne)* 10, 193
- Matsubayashi H, Takaori K, Morizane C, Maguchi H, Mizuma M, Takahashi H, Wada K, Hosoi H, Yachida S, Suzuki M et al. (2017): Familial pancreatic cancer: Concept, management and issues. *World J Gastroenterol* 23, 935–48
- McGlynn JKA, Petricka JL, London WT (2015): Global epidemiology of hepatocellular carcinoma: An emphasis on demographic and regional variability. *Clin Liver Dis* 19, 223–38
- Mihm S (2015): Activation of type I and type III Interferons in chronic hepatitis C. *J Innate Immun* 7, 251–9
- Mihm S, Frese M, Meier V, Wietzke-Braun P, Scharf J, Bartenschlager R, Ramadori G (2004): Interferon type I gene expression in chronic hepatitis C. *Lab Invest* 84, 1148–59

- Minas TZ, Tang W, Smith CJ, Onabajo OO, Obajemu A, Dorsey TH, Jordan SV, Obadi OM, Ryan BM, Prokunina-Olsson L et al. (2018): IFNL4- Δ G is associated with prostate cancer among men at increased risk of sexually transmitted infections. *Commun Biol* 1, 191
- Minn AJ (2015): Interferons and the immunogenic effects of cancer therapy. *Trends Immunol* 36, 725–37
- Moletta L, Serafini S, Valmasoni M, Pierobon ES, Ponzoni A, Sperti C (2019): Surgery for recurrent pancreatic cancer: Is it effective. *Cancers (Basel)* 11, e991
- NCI (Hrsg.) (2020a): *SEER cancer stat facts: Cancer of any site*.
<https://seer.cancer.gov/statfacts/html/all.html> (Zugriff am 10.01.2020)
- NCI (Hrsg.) (2020b): *SEER cancer stat facts: Liver and intrahepatic bile duct cancer*.
<https://seer.cancer.gov/statfacts/html/livibd.html> (Zugriff am 20.01.2020)
- NCI (Hrsg.) (2019): *SEER cancer stat facts: Pancreatic cancer*.
<https://seer.cancer.gov/statfacts/html/pancreas.html> (Zugriff am 11.10.2019)
- O'Brien TR, Prokunina-Olsson L, Donnelly RP (2014): IFN- λ 4: The paradoxical new member of the interferon lambda family. *J Interferon Cytokine Res* 34, 829–38
- O'Brien TR, Pfeiffer RM, Paquin A, Lang Kuhs KA, Chen S, Bonkovsky HL, Edlin BR, Howell CD, Kirk GD, Kuniholm MH et al. (2015): Comparison of functional variants in IFNL4 and IFNL3 for association with HCV clearance. *J Hepatol* 63, 1103–10
- Obajemu AA, Rao N, Dille KA, Vargas JM, Sheikh F, Donnelly RP, Shabman RS, Meissner EG, Prokunina-Olsson L, Onabajo OO (2017): IFN- λ 4 attenuates antiviral responses by enhancing negative regulation of IFN signaling. *J Immunol* 199, 3808–20
- Öhlund D, Handly-Santana A, Biffi G, Elyada E, Almeida AS, Ponz-Sarvisé M, Corbo V, Oni TE, Hearn SA, Lee EJ et al. (2017): Distinct populations of inflammatory fibroblasts and myofibroblasts in pancreatic cancer. *J Exp Med* 214, 579–96
- Onabajo OO, Porter-Gill P, Paquin A, Rao N, Liu L, Tang W, Brand N, Prokunina-Olsson L (2015): Expression of interferon lambda 4 is associated with reduced proliferation and increased cell death in human hepatic cells. *J Interferon Cytokine Res* 35, 888–900
- Onabajo OO, Wang F, Banday R, Prokunina-Olsson L (2018): Tumor transcriptome analysis and immune infiltrate profiling suggests a role for the IFNL3/IFNL4 genomic locus in liver cancer. *J Immunol* 200, 178
- Palma-Ocampo HK, Flores-Alonso JC, Vallejo-Ruiz V, Reyes-Leyva J, Flores-Mendoza L, Herrera-Camacho I, Rosas-Murrieta NH, Santos-Lopez G (2015): Interferon lambda inhibits dengue virus replication in epithelial cells. *Virology* 12, 150

- Peduzzi P, Concato JJ, Feinstein AR, Holford TR (1995): Importance of events per independent variable in proportional hazards regression analysis II. Accuracy and precision of regression estimates. *J Clin Epidemiol* 48, 1503–10
- Petta S, Valenti L, Tuttolomondo A, Dongiovanni P, Pipitone RM, Camma C, Cabibi D, Di Marco V, Fracanzani AL, Badiali S et al. (2017): Interferon lambda 4 rs368234815 TT>dG variant is associated with liver damage in patients with nonalcoholic fatty liver disease. *Hepatology* 66, 1885–93
- Pingwara R, Witt-Jurkowska K, Ulewicz K, Mucha J, Tonecka K, Pilch Z, Taciak B, Zabielska-Koczywas K, Mori M, Berardozi S et al. (2017): Interferon lambda 2 promotes mammary tumor metastasis via angiogenesis extension and stimulation of cancer cell migration. *J Physiol Pharmacol* 68, 573–83
- Prokunina-Olsson L (2019): Genetics of the human interferon lambda region. *J Interferon Cytokine Res* 39, 599–608
- Prokunina-Olsson L, Muchmore B, Tang W, Pfeiffer RM, Park H (2013): A variant upstream of IFNL3 (IL28B) creating a new interferon gene IFNL4 is associated with impaired clearance of hepatitis C virus. *Nat Genet* 45, 164–73
- Raphael BJ, Hruban RH, Aguirre AJ, Moffitt RA, Yeh JJ, Stewart C, Robertson AG, Cherniack AD, Gupta M, Getz G et al. (2017): Integrated genomic characterization of pancreatic ductal adenocarcinoma. *Cancer Cell* 32, 185–203
- Real LM, Herrero R, Rivero-Juárez A, Camacho Á, Macías J, Vic S, Soriano V, Viedma S, Guardiola JM, Fibla J et al. (2015): IFNL4 rs368234815 polymorphism is associated with innate resistance to HIV-1 infection. *AIDS* 29, 1895–7
- Reig M, Mariño Z, Perello C, Iñarrairaegui M, Ribeiro A, Lens S, Díaz A, Vilana R, Darnell A, Varela M et al. (2016): Unexpected high rate of early tumor recurrence in patients with HCV-related HCC undergoing interferon-free therapy. *J Hepatol* 65, 719–26
- Robek MD, Boyd BS, Chisari FV (2005): Lambda interferon inhibits hepatitis B and C virus replication. *J Virol* 79, 3851–4
- Robert Koch-Institut und die Gesellschaft der epidemiologischen Krebsregister in Deutschland e.V. (Hrsg.): Krebs in Deutschland für 2015/2016. 12. Ausgabe. Berlin 2019
- Robinson JT, Thorvaldsdóttir H, Wenger AM, Zehir A, Mesirov JP (2017): Variant review with the Integrative Genomics Viewer. *Cancer Res* 77, e31–e4
- Rochefort MM, Ankeny JS, Kadera BE, Donald GW, Isacoff W, Wainberg ZA, Hines OJ, Donahue TR, Reber HA, Tomlinson JS (2013): Impact of tumor grade on pancreatic cancer prognosis: Validation of a novel TNMG staging system. *Ann Surg Oncol* 20, 4322–9

- Scarà S, Bottoni P, Scatena R (2015): CA 19-9: Biochemical and clinical aspects. *Adv Exp Med Biol* **867**, 247–60
- Schneider WM, Chevillotte MD, Rice CM (2014): Interferon-stimulated genes: A complex web of host defenses. *Annu Rev Immunol* **32**, 513–45
- Schroder K, Hertzog PJ, Ravasi T, Hume DA (2004): Interferon- γ : An overview of signals, mechanisms and functions. *J Leukoc Biol* **75**, 163–89
- Schulze K, Imbeaud S, Letouze E, Alexandrov LB, Calderaro J, Rebouissou S, Couchy G, Meiller C, Shinde J, Soysouvanh F et al. (2015): Exome sequencing of hepatocellular carcinomas identifies new mutational signatures and potential therapeutic targets. *Nat Genet* **47**, 505–11
- Schulze K, Nault J, Villanueva A (2016): Genetic profiling of hepatocellular carcinoma using next-generation sequencing. *J Hepatol* **65**, 1031–42
- Siegel RL, Miller KD, Jemal A (2018): Cancer statistics, 2018. *CA Cancer J Clin* **68**, 7–30
- Singhi AD, McGrath K, Brand RE, Khalid A, Zeh HJ, Chennat JS, Fasanella KE, Papachristou GI, Slivka A, Bartlett DL et al. (2018): Preoperative next-generation sequencing of pancreatic cyst fluid is highly accurate in cyst classification and detection of advanced neoplasia. *Gut* **67**, 2131–41
- Slatkin M (2008): Linkage disequilibrium - understanding the evolutionary past and mapping the medical future. *Nat Rev Genet* **9**, 477–85
- Stern C (1943): The Hardy-Weinberg law. *Science* **97**, 137–8
- Stiff A, Carson W (2015): Investigations of interferon-lambda for the treatment of cancer. *J Innate Immun* **7**, 243–50
- Suppiah V, Moldovan M, Ahlenstiel G, Berg T, Weltman M, Abate ML, Bassendine M, Spengler U, Dore GJ, Powell E et al. (2009): IL28B is associated with response to chronic hepatitis C interferon- α and ribavirin therapy. *Nat Genet* **41**, 1100–4
- Tanaka Y, Nishida N, Sugiyama M, Kurosaki M, Matsuura K, Sakamoto N, Nakagawa M, Korenaga M, Hino K, Hige S et al. (2009): Genome-wide association of IL28B with response to pegylated interferon-alpha and ribavirin therapy for chronic hepatitis C. *Nat Genet* **41**, 1105–9
- Tang W, Wallace TA, Yi M, Magi-Galluzzi C, Dorsey TH, Onabajo OO, Obajemu A, Jordan SV, Loffredo CA, Stephens RM et al. (2018): IFNL4- Δ G allele is associated with an interferon signature in tumors and survival of African-American men with prostate cancer. *Clin Cancer Res* **24**, 5471–81
- Tannus RK, Almeida-Carvalho SR, Loureiro-Matos CA, Miziara-Gonzalez A, Salzedas-Netto AA, Szejnfeld D, D'Ippolito G, Pereira-Lanzoni V, Souza-Silva I (2018): Eva-

- luation of survival of patients with hepatocellular carcinoma: A comparative analysis of prognostic systems. *PLoS One* **13**, e0194922
- Tezuka Y, Endo S, Matsui A, Sato A, Saito K, Semba K, Takahashi M, Murakami T (2012): Potential anti-tumor effect of IFN- λ 2 (IL-28A) against human lung cancer cells. *Lung Cancer* **78**, 185–92
- Thomas DL, Thio CL, Martin MP, Qi Y, Ge D, O’huigin C, Kidd J, Kidd K, Khakoo SI, Alexander G et al. (2009): Genetic variation in IL28B and spontaneous clearance of hepatitis C virus. *Nature* **461**, 798–801
- Tomczak K, Czerwinska P, Wiznerowicz M (2015): The Cancer Genome Atlas (TCGA): An immeasurable source of knowledge. *Contemp Oncol (Pozn)* **19**, A68–A77
- Tunissiolli NM, Castanhole-Nunes MMU, Biselli-Chicote PM, Pavarino EC, da Silva RF, de Cássia R, da Silva MA, Goloni-Bertollo EM (2017): Hepatocellular carcinoma: A comprehensive review of biomarkers, clinical aspects, and therapy. *Asian Pac J Cancer Prev* **18**, 863–72
- van Hecke T (2010): Power study of anova versus Kruskal-Wallis test. *J Stat Manag Syst* **15**, 241–7
- Via M, Gignoux C, Burchard EG (2010): The 1000 Genomes Project: New opportunities for research and social challenges. *Genome Med* **2**, 3
- Wang C, Li S (2019): Clinical characteristics and prognosis of 2887 patients with hepatocellular carcinoma. A single center 14 years experience from China. *Medicine (Baltimore)* **98**, e14070
- Weinberg W: Über den Nachweis der Vererbung beim Menschen. Gesellschaft für Naturkunde in Württemberg e.V. (Hrsg.): Jahreshefte des Vereins für vaterländische Naturkunde in Württemberg. **64**, Klett&Hartmann, Stuttgart 1908, 369–82
- Weinstein JN, Collisson EA, Mills GB, Shaw KRM, Ozenberger BA, Ellrott K, Shmulevich I, Sander C, Stuart JM (2013): The Cancer Genome Atlas Pan-Cancer analysis project. *Nat Genet* **45**, 1113–20
- Yang Y (2015): Cancer immunotherapy: Harnessing the immune system to battle cancer. *J Clin Invest* **125**, 3335–7
- Zhang J, Finney RP, Rowe W, Edmonson M, Yang SH, Dracheva T, Jen J, Struewing JP, Buetow KH (2007): Systematic analysis of genetic alterations in tumors using Cancer Genome WorkBench (CGWB). *Genome Research* **17**, 1111–7
- Zheng-Bradley X, Flicek P (2016): Applications of the 1000 Genomes Project resources. *Brief Funct Genomics* **16**, 163–70
- Zhou L, Li J, Zhou Y, Liu J, Zhuang K, Gao J, Liu S, Sang M, Wu J, Ho W (2015): Induction of interferon- λ contributes to TLR3 and RIG-I activation-mediated inhibi-

- tion of herpes simplex virus type 2 replication in human cervical epithelial cells. *Mol Hum Reprod* 21, 917–29
- Ziegler A, Lange S, Bender R (2007): Überlebenszeitanalyse: Die Cox-Regression. *Dtsch Med Wochenschr* 132, e42–4

Danksagung

Mein ganzer Dank gebührt meiner Doktormutter Frau Prof. Dr. rer. nat. Sabine Mihm (Arbeitsgruppenleiterin innerhalb der Klinik für Gastroenterologie, gastrointestinale Onkologie und Endokrinologie). Unter ihrer exzellenten Betreuung entstand diese Dissertation. Sie war die treibende Kraft: von der Themenfindung über die anfänglichen technischen Probleme bis hin zur Verschriftlichung der Arbeit stand Frau Prof. Mihm mir zur Seite. Ihre hohe wissenschaftliche Kompetenz wie ihre kritische Betrachtung waren für mich ein großes Vorbild und ermutigten mich immer wieder aufs Neue, diese Arbeit bestmöglich zu durchdenken und zu verfassen. Bei unseren regelmäßigen Treffen brachte sie mir die wissenschaftliche Vorgehensweise und die enorme Wichtigkeit, stets auf dem aktuellsten Forschungsstand zu bleiben, bei. Ohne das große Engagement von Frau Prof. Mihm würde diese Dissertation nicht in der aktuellen Form bestehen. Dafür, dass sie sich unentwegt Zeit für meine Anliegen nahm, kann ich ihr kaum in ausreichendem Maße danken.

Meiner Ko-Betreuerin Frau Prof. Dr. Heike Bickeböller (Direktorin des Instituts für Genetische Epidemiologie) möchte ich ebenfalls ein großes Dankeschön aussprechen. Mit ihrer außerordentlichen Fachkompetenz ermöglichte sie, das Studiendesign dieser Arbeit in Form einer genetischen Assoziationsstudie zu entwerfen, zu überdenken und zu bearbeiten. Bei Treffen konnten durch ihre kritische Betrachtungsweise Schwierigkeiten der Arbeit erkannt und behoben werden. Ohne das große Engagement von Frau Prof. Bickeböller würde diese Arbeit und vorwiegend das Studiendesign nicht in ihrer aktuellen Form existieren.

Auch bei Herrn Dr. Andreas Leha und bei Herrn Dr. John Wiedenhöft aus der wissenschaftlichen Serviceeinheit Medizinische Biometrie und statistische Bioinformatik möchte ich mich vielmals für ihre Beratung zu statistischen Unklarheiten bedanken. Insbesondere für die Hilfe bei der Modellierung der betrachteten Kriterien für die Ereigniszeitanalysen innerhalb des Programms RStudio möchte ich mich bedanken.

Ein sehr großes Dankeschön auch an meinen Freund Manuel Andersch. Neben der technischen Hilfe bei Programmierproblemen mit der TCGA-Datenbank unterstützte er mich tatkräftig bei technischen Fragen, die bei der Verschriftlichung meiner Arbeit mit dem Programm LaTeX aufkamen.