Technische Universität München, Fakultät für Medizin

Evolutionary trajectories and *KRAS* dosage define pancreatic cancer phenotypes

Sebastian Müller

Vollständiger Abdruck der von der
Fakultät für Medizin
der Technischen Universität München zur Erlangung des akademischen Grades
eines Doktors der Naturwissenschaften
genehmigten Dissertation.

Vorsitzender: Prof. Dr. Dieter Saur

Prüfende/-r der Dissertation:

1. Prof. Dr. Radu Roland Rad

2. Prof. Dr. Martin Hrabě de Angelis

3. Prof. Dr. Vigo Heissmeyer

Die Dissertation wurde am 08.07.2019 bei der Technischen Universität
München eingereicht und durch die Fakultät für Medizin
am 11.02.2020 angenommen.

To my parents

Give every day the chance
to become the most beautiful day of your life.
Mark Twain

Zusammenfassung

Das duktale Pankreaskarzinom (PDAC) zeichnet sich durch häufige Mutation einiger weniger Gene (sogenannte Signatur-Gene, wie *KRAS*, *CDKN2A*, *TP53*, *SMAD4*), sowie durch beträchtliche Heterogenität darüber hinaus mutierter Krebsgene aus. Die Annahme, dass die phänotypische Diversität des PDACs auf Grundlage heterogen mutierter Krebsgene erklärt werden kann, konnte bis auf wenige Ausnahmen bisher nicht bestätigt werden. Ebenso konnten durch den Vergleich von Primarius und Metastase keine wiederholt mutierten Metastasierungsgene gefunden werden, wodurch die genetische Basis der PDAC-Metastasierung bisher weitgehend unverstanden bleibt.

In der hier vorliegenden Arbeit kann gezeigt werden, dass zentrale Aspekte der PDAC-Biologie durch Variation der Gendosis mutierter Signatur-Gene und deren Entwicklung entlang bestimmter Evolutionswege erklärt werden können. So konnte eine Erhöhung der Gendosis der initiierenden *KRAS* Mutation ($KRAS^{MUT\text{-}iGD}$) bereits in frühen Vorläuferläsionen des PDAC gefunden werden. Durch die Analyse verschiedener $Kras^{MUT}$-induzierten Mausmodelle des PDAC konnte die Bedeutung einer erhöhten $Kras^{MUT}$-Gendosis für die frühe Progression als auch die Metastasierung des PDAC aufgedeckt werden; eine Beobachtung die wiederum die häufige Metastasierung in Patienten bei Erstdiagnose erklären kann. Insgesamt wurden die Genome und Transkriptome von 135 Primarius/Metastase-Zellkulturen (isoliert aus verschiedenen $Kras^{G12D}$-induzierten PDAC Mausmodellen) umfassend charakterisiert. Diese Ressource von Zellkulturen verschiedener Mausmodelle des PDAC wurde genutzt um Limitierungen der Gendosis-Analyse in humanem PDAC-Gewebe, durch die „Kontamination" mit Nicht-Tumorzellen des Stromas, zu umgehen. Mittels der integrativen Analyse von (i) Genom-/Transkriptom-Daten, (ii) Tumorphänotypen, (iii) humanen PDAC-Studien und (iv) funktionellen Experimenten konnte eine Reihe zusätzlicher $Kras^{MUT\text{-}iGD}$-vermittelter Effekte aufgezeigt werden. So definierten verschiedene Stufen der $Kras^{MUT}$-Gendosis und -Expression bestimmte Zellmorphologien *in vitro*, als auch Histopathologien und klinische Krankheitseigenschaften *in vivo*. Verschiedene Stufen der $Kras^{MUT}$-Expression verursachten zelluläre Plastizität, wie z.B. Epitheliale-zu-Mesenchymale Transition; wobei die höchste $Kras^{MUT}$-Expression dem aggressivstem, entdifferenziertestem Phänotyp zu Grunde lag. Mechanistisch entwickelte sich die Variation der onkogenen Dosis entlang bestimmter Evolutionswege, welche durch definierte Tumorsuppressorgene (*Cdkn2a*, *Trp53*, *Tgfbr2*) und deren Dosisstatus (wildtypisch/heterozygoter, homozygoter Funktionsverlust) kontrolliert werden. Durch phyolgenetische Untersuchungen konnte gezeigt werden, dass $Kras^{MUT\text{-}iGD}$ zunächst den bereits vorangegangenen homozygoten

Funktionsverlust von *Cdkn2a* oder *Trp53* voraussetzt, dann aber parallel durch konvergente Evolution entstehen kann. Im Gegensatz dazu zeigten PDAC mit heterozygotem Verlust von *Cdkn2a* eine Amplifikation/Erhöhung der Gendosis alternativer Onkogene (*Myc*, *Yap1*, *Nfkb2*), die mit $Kras^{MUT\text{-}HET}$ kollaborieren um die Progression von Pankreastumoren anzutreiben; allerdings mit geringerer Fähigkeit zur Metastasierung. Weiterhin zeigten diese Tumore eine Onkogen-spezifische/Kontext-abhängige Haploinsuffizienz von *Cdkn2a* auf, ein Evolutionsweg der ebenfalls in einem $Kras^{MUT}$-induzierten Mausmodell mit zusätzlicher Inaktivierung des *Tgfbr2*-Gens bevorzugt wurde. Zusammengefasst wird in der vorliegenden Arbeit gezeigt, wie sich onkogene Dosisvariation entlang definierter Evolutionswege entwickelt und wie damit zentrale Eigenschaften der Krankheit (frühe Progression, Histopathologie, Metastasierung, zelluläre Plastizität und klinische Aggressivität) einhergehen. Dadurch werden bisher unbekannte, grundlegende genetische Prinzipien der Evolution des PDAC und dessen phänotypischer Diversifikation identifiziert.

Circa ein Drittel aller soliden Tumore tragen aktivierende Mutationen in einem der Ras Gene, einschließlich der häufigen allelischen Imbalance der Ras-Mutation, was eine über die Pankreaskarzinogenese hinausgehende Bedeutung der hier identifizierten genetischen Mechanismen der Ras-induzierten Tumorigenese vermuten lässt.

Summary

Pancreatic ductal adenocarcinoma (PDAC) has frequent alterations in a few genes (*KRAS*, *CDKN2A*, *TP53*, *SMAD4*) and extensive heterogeneity of cancer drivers beyond. The expectation that mutational landscapes of rare drivers could explain phenotypic diversity has -with few exceptions- not come true. Likewise, PDAC metastasis is not understood, and comparisons of primary/metastasis pairs did not find recurrently mutated “metastasis genes”.

In this thesis, it could be shown that key aspects of PDAC biology are defined by gene-dosage variation of PDAC signature genes, evolving along distinct evolutionary routes. The increased gene dosage of the initiating *KRAS* mutation ($KRAS^{MUT\text{-}iGD}$) was found to occur already in early stages of human PDAC precursor lesions. The analysis of various $Kras^{G12D}$-driven PDAC mouse models revealed the importance of the acquisition of $Kras^{MUT\text{-}iGD}$ for both, early progression and metastasis, rationalizing the high frequency of human PDAC dissemination in patients at first diagnosis. In total, the genomes and transcriptomes of 135 primary/metastasis cell cultures derived from different $Kras^{G12D}$-driven PDAC mouse models were comprehensively characterized. This resource of mouse PDAC cell cultures was exploited to overcome the limitations of gene dosage analysis caused by the high content of stroma cells in human PDAC tissues. Integrative analyses of the genomic/transcriptomic data and tumour phenotypes, combined with human studies and functional analyses revealed a series of additional $Kras^{MUT}$-dosage effects: different levels of $Kras^{MUT}$ gene dosage and expression defined distinct cellular morphologies *in vitro*, histopathologies and clinical outcomes *in vivo*. The expression level of $Kras^{MUT}$ was also found to induce cellular plasticity, including epithelial-to-mesenchymal transition; with highest $Kras^{MUT}$-specific expression levels underlying the most aggressive undifferentiated phenotypes. Mechanistically, oncogenic dosage-variation is linked to distinct evolutionary routes, governed by defined types (*Cdkn2a*, *Trp53*, *Tgfbr2*) and states (wildtype/heterozygous, homozygous inactivation) of tumour-suppressor gene alterations. Phylogenetic tracking studies revealed convergent evolution of $Kras^{MUT\text{-}iGD}$ and its dependency on prior homozygous loss of *Cdkn2a* or *Trp53*. By contrast, PDAC with heterozygous loss of *Cdkn2a* showed alternative amplifications of known and novel oncogenes (*Myc*, *Yap1*, *Nfkb2*) that collaborate with $Kras^{MUT\text{-}HET}$ to drive pancreatic tumour progression, however with lower metastatic potential. Of note, these PDAC showed oncogene-selective/context-dependent haploinsufficiency of *Cdkn2a*, an evolutionary route that was preferred in a $Kras^{MUT}$-driven PDAC mouse model with combined inactivation of *Tgfbr2*. Taken together, the study presented here identifies genetic

hallmarks of pancreatic cancer evolution and shows how oncogenic dosage-variation is differentially licensed along individual routes to control critical disease characteristics, including early progression, histopathology, metastasis, cellular plasticity and clinical aggressiveness.

Approximately on third of solid cancers show activating mutations of Ras genes (frequently involving the allelic imbalance of mutated Ras) suggesting that the principles of Ras-driven pancreatic carcinogenesis might be also relevant to other cancer types.

List of abbreviations

ADEX	Aberrantly differentiated endocrine exocrine
ATP1	Activating transposon 1
CCLE	Cancer cell line encyclopaedia
Cdkn2a	Cyclin-dependent kinase inhibitor 2A
Cdkn2a$^{\Delta HET}$	Heterozygous inactivation of the *Cdkn2a* locus
Cdkn2a$^{\Delta HOM}$	Homozygous inactivation of the *Cdkn2a* locus
Chr	Chromosome
CI	Confidence interval
CNA	Copy number alteration
COSMIC	Catalogue of somatic mutations in cancer
CRISPR	Clustered regularly interspaced short palindromic repeats
DAVID	Database for annotation, visualization and integrated discovery
dd$H_2$0	Double-distilled H_2O
DMEM	Dulbecco's modified eagle's medium
DNA	Deoxyribonucleic acid
EMT	Epithelial-to-mesenchymal transition
FCS	Fetal calf serum
GFP	Green fluorescence protein
GO	Gene ontology
hPanIN	Human pancreatic intraepithelial neoplasia
hPDAC	Human pancreatic ductal adenocarcinoma
IPMN	Intraductal papillary mucinous neoplasia
Kras	Kirsten rat sarcoma viral oncogene homolog
Kras$^{G12D\text{-}AG}$	Arm-level gain of the *Kras*^{G12D} locus
Kras$^{G12D\text{-}FG}$	Focal gain of the *Kras*^{G12D} locus
Kras$^{G12D\text{-}HET}$	Heterozygous status of the *Kras*^{G12D} locus
Kras$^{G12D\text{-}iGD}$	Homozygous status of the *Kras*^{G12D} locus
Kras$^{G12D\text{-}LOH}$	Copy number-neutral LOH of the *Kras*^{G12D} locus
LOH	Loss of heterozygosity
MAF	Mutation annotation format
MCN	Mucinous cystic neoplasm
M-FISH	Multicolour fluorescence in situ hybridization
mPDAC	Murine pancreatic ductal adenocarcinoma
MSigDB	Molecular signatures database
Myc	V-Myc avian myelocytomatosis viral oncogene homolog
Ncruc	Noncoding region upstream of *Cdkn2a*

Nfkb2	Nuclear factor kappa B subunit 2
NGS	Next generation sequencing
OR	Odds ratio
P/S	Penicillin / streptomycin
PanIN	Pancreatic intraepithelial neoplasia
PCR	Polymerase chain reaction
PDAC	Pancreatic ductal adenocarcinoma
PK	PDAC mouse model driven by $Kras^{G12D}$ expression in pancreatic cells
PKC	PDAC mouse model driven by $Kras^{G12D}$ expression and knockout of *Cdkn2a* in pancreatic cells
PKP	PDAC mouse model driven by $Kras^{G12D}$ expression and knockout of *Trp53* in pancreatic cells
PK-PB	PDAC mouse model with high mutational load driven by $Kras^{G12D}$ expression and PiggyBac-based insertional mutagenesis in pancreatic cells
PKT	PDAC mouse model driven by $Kras^{G12D}$ expression and knockout of *Tgfbr2* in pancreatic cells
QiSeq	Quantitative insertion site sequencing
QM	Quasimesenchymal
qPCR	Quantitative polymerase chain reaction
RelB	V-Rel avian reticuloendotheliosis viral oncogene homolog B
RNA	Ribonucleic acid
SCRB-Seq	Single-cell RNA barcoding and sequencing
sgRNA	Single guide ribonucleic acid
SNP	Single nucleotide polymorphism
SNV	Single nucleotide variation
Tgfbr1	Transforming growth factor beta receptor 1
Tgfbr2	Transforming growth factor beta receptor 2
$Tgfbr2^{\Delta HET}$	Heterozygous inactivation of the *Tgfbr2* gene
$Tgfbr2^{\Delta HOM}$	Homozygous inactivation of the *Tgfbr2* gene
Tgfβ	Transforming growth factor beta
Tp53/*Trp53*	Tumour protein p53
$Trp53^{\Delta HET}$	Heterozygous inactivation of the *Trp53* gene
$Trp53^{\Delta HOM}$	Homozygous inactivation of the *Trp53* gene
VCF	Variant call format
Yap1	Yes-associated protein 1

List of figures

Contents

1. Introduction

1.1 Pancreatic ductal adenocarcinoma

Pancreatic tumours are the fourth leading cause of cancer-related mortality in Western World and have the worst prognosis of all cancers. In 2016, more than 300.000 patients died from a tumour in the pancreas (Ilic *et al.*, 2016). Industrialized countries show the highest incidence and mortality rate. Whilst treatment options have constantly improved in most other cancer types, 5-year survival rates in PDAC stayed around 5% in the last four decades (Rahib *et al.*, 2014; Siegel *et al.*, 2016). The by far most frequent histological subtype of pancreatic cancer is pancreatic ductal adenocarcinoma (PDAC), accounting for about 85% of all pancreatic tumours (Klimstra *et al.*, 2009). PDAC is characterized by its late diagnosis, its early metastatic dissemination (typically to the lymph nodes, liver and lung) and its pronounced desmoplastic stromal reaction. Because efforts to improve drug delivery through the dense/extensive stroma of PDAC and early detection of PDAC have not been successful yet, there is only marginal improvement in the treatment of PDAC that manifests in an unchanged poor prognosis of PDAC patients. In addition, the incidence of pancreatic tumours is rising and thus PDAC is predicted to become the second most frequent cause of cancer-related death by 2030 (Rahib *et al.*, 2014; Ying *et al.*, 2016) (Figure 1a).

Histopathological and genetic analyses have defined three distinct pre-malignant lesions of the pancreas: (i) pancreatic intraepithelial neoplasms (PanIN) (ii) intraductal papillary mucinous neoplasms (IPMN) and (iii) mucinous cystic neoplasms (MCN) (Maitra *et al.*, 2005). PanINs are the best described precursor lesions of human PDAC and are thought to progress through a stepwise accumulation of specific somatic mutations and cellular atypia. According to the extent of cytological dysplasia, PanINs are sub-stratified from grade 1 to grade 3 (carcinoma in situ) (Hezel *et al.*, 2006). The PanIN progression stages are typically paralleled by somatic mutations resulting in *KRAS* activation, as well as inactivation of *CDKNA2*, *TP53* and genes of the canonical transforming growth factor-β (TGFβ) signalling pathway (Hruban *et al.*, 2000) (Figure 1b). The occurrence of *KRAS*-activating mutations is the first and almost universal event (95% of PDAC) during PanIN initiation/progression. *KRAS*-activating mutations are therefore critical during early PanIN progression and for initiating PDAC development. Activating mutations are not randomly distributed along the protein-coding sequence of *KRAS*. There are two main hotspot mutations resulting in changes of amino acids G12 (~80% of cases) or Q61 (~5% of cases)

(data from QCMG study, cBioPortal) (Figure 2a). The intrinsic GTPase-activity gets destroyed through these hotspot mutations resulting in a constitutively activated *KRAS* protein with an approximately 1000-fold lower intrinsic GTPase acitivty (Pylayeva-Gupta *et al.*, 2011).

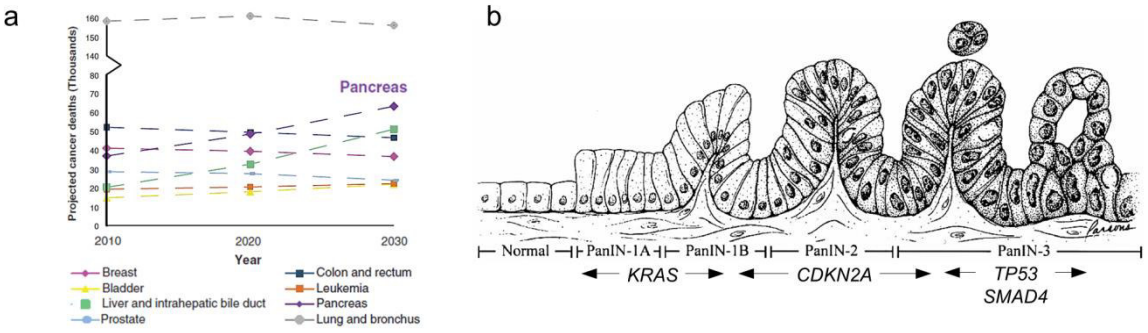

Figure 1 | Clinic and progression of pancreatic ductal adenocarcinoma.
[a] Projected cancer deaths of pancreatic cancer and additional cancer entities by 2020 and 2030 (modified from (Rahib *et al.*, 2014). **[b]** Classical/stepwise progression model of pancreatic cancer. Morphologic and genetic progression of pancreatic epithelial cells trough defined precursor stages (PanINs) to invasive pancreatic cancer is shown from left to right (modified from (Hruban *et al.*, 2000). The stage-specific appearance of activating mutations in *KRAS* or inactivating alterations in *CDKN2A*, *TP53* and *SMAD4* are shown below PanIN stages.

Activating mutations in *KRAS* are typically followed by inactivation of the *CDKN2A* locus during PanIN progression. Loss of function of *CDKN2A* already occurs in low-grade PanIN lesions and is observed in more than 90% of PDAC cases (Schutte *et al.*, 1997; Hruban *et al.*, 2000). The *CDKN2A* locus encodes two physically linked tumour suppressor genes, namely $p16^{INK4A}$ and $p14^{ARF}$ ($p19^{Arf}$ in mice) (Figure 2b). The transcription of *INK4A* and *ARF* starts from two different exons, exon 1α and exon1β respectively. Later on during transcription, exon 1α and 1β are spliced into identical/shared exons 2 and 3. *INK4A* and *ARF* encode for two potent tumour suppressor genes regulating many processes involved in tumourigenesis, such as growth factor signalling and cell cycle progression (Figure 2c). $p16^{INK4A}$ is a cyclin-dependent kinase inhibitor that prevents binding of CDK4/6 to D-type cyclins (Russo *et al.*, 1998) and phosphorylation of retinoblastoma protein (RB). Activation of $p16^{INK4A}$ leads to hypo-phosphorylation of RB thereby preventing progression of the cell cycle from G1 to S phase. Under physiological conditions $p16^{INK4A}$ is not expressed. Oncogenic mutations that deregulate cell cycle progression activate expression of $p16^{INK4A}$. Activation of $p16^{INK4A}$ in turn results in the sequestration of the cell from the cell cycle (a state called oncogene-induced senescence). In comparison to $p16^{INK4A}$ and *RB1*, $p19^{ARF}$ is involved in a distinct anti-cancer pathway (Figure 2c). Like $p16^{INK4A}$, $p19^{ARF}$ is also not expressed under physiological conditions but can be readily induced through the overexpression of oncogenes such as *c-Myc* and *E2F* (Zindy *et al.*, 1998; Zhu *et al.*, 1999; Dimri *et al.*, 2000). The ARF protein directly binds to MDM2. MDM2 is an E3 ubiquitin-

protein ligase regulating TP53 protein stability and degradation. Activation of *ARF* transcription leads to sequestration of MDM2 from TP53, TP53 protein stabilization and induction of TP53 effector signalling programs such as senescence or apoptosis. Of note, the response of *ARF* to oncogenic stress signals seems to be context dependent, as overexpression of RAS activates *ARF* in mouse, but not human, cells (Palmero *et al.*, 1998; Sharpless *et al.*, 2001; Huot *et al.*, 2002). As compared to *INK4A*, the importance of *ARF* in human tumour suppression is more complicated to analyse because of (i) the frequent co-deletion of *ARF* with *INK4A* and (ii) the scarcity of *ARF* specific mutations or promoter methylation (Sharpless, 2005). For example in PDAC, mutations exclusively inactivating *INK4A* but sparing the *ARF* gene have been described (Bardeesy *et al.*, 2006a). *INK4A*-specific germline mutations are also associated with increased risk of PDAC development (Goldstein *et al.*, 1995; Whelan *et al.*, 1995). However, the biochemical interaction of *ARF* and *MDM2* seems to rely only on the first 25 N-terminal amino acids encoded by exon 1β, making the occurrence of *ARF*-specific inactivating mutations extremely unlikely. Of note, *ARF*-specific germline mutations have been reported in a few cases of familial melanoma or astrocytoma (Randerson-Moor *et al.*, 2001; Rizos *et al.*, 2001). While overlapping stretches of open reading frames are common in viruses and bacteria, the shared exon usage of the *CDKN2A* locus is practically unique in the mammalian genome (Sharpless, 2005). Although it is not yet clear why *INK4A* and *ARF* tumour suppressors are juxtaposed within 0.03 Mb of the 3000 Mb genome (making them vulnerable for co-deletion), it is likely that the evolution of *ARF* in the pre-existing *INK4A* locus allows for a more precise and thorough discrimination of physiological and oncogenic growth signalling in complex, long-lived mammalian organisms (Sharpless, 2005).

The third hallmark of PanIN progression is the inactivation of *TP53* protein which typically occurs in more advanced PanIN stages (Hruban *et al.*, 2000). *TP53* governs multiple processes involved in tumour suppression and is mutated in up to 85% of PDAC cases (Yachida *et al.*, 2012). About 66% of all *TP53* mutations are missense mutations that are typically located in the DNA binding domain and result in the functional inactivation of *TP53* (Jones *et al.*, 2008; Yachida *et al.*, 2012). As already outlined above, *TP53* and *ARF* reside in the same anti-cancer pathway (Figure 2c). Although *ARF* was found to have *TP53*-independent tumour suppressive activity, experimental evidence suggests that activation of *TP53* is the major function of *ARF*-mediated tumour suppression *in vivo* (Sharpless, 2005). In addition to sensing oncogenic signalling via *ARF*, *TP53* is well known for its central role in regulating the DNA damage response. It has been estimated that approximately 50 double strand breaks are occurring in a single healthy human cell per

day (Vilenchik *et al.*, 2003). Complex DNA damage, such as double strand breaks or single stranded DNA, are sensed by ATM/ATR proteins leading to the recruitment of CHK1/2 to the site of DNA damage and phosphorylation of MDM2. Phosphorylation of MDM2 reduces its activity, thereby stabilizing TP53 and inducing TP53 target genes to trigger reversible cell cycle arrest and to allow for repair of DNA damage (Figure 2c). In case of extensive or persistent DNA damage signals, such as caused by oncogene-induced replication stress, TP53 activation results in permanent cell cycle arrest (senescence) or programmed cell death (apoptosis). More recent studies have challenged the relative importance of these classical *TP53* functions in tumour suppression and have revealed the role of *TP53* in additional cellular processes, such as metabolism, differentiation, metastasis and modulating the microenvironment (Bieging *et al.*, 2014). Interestingly, somatic inactivation of *TP53* and *ARF* coexist in approximately 40% of human PDACs (Heinmoller *et al.*, 2000; Maitra *et al.*, 2003). The co-occurrence of *TP53* and *ARF* inactivation implies non-overlapping tumour suppressor roles for both genes during PDAC progression. However, the distinct requirements for *TP53* versus *ARF* inactivation in PDAC development have not been genetically explored yet.

The inactivation of *SMAD4* is the fourth hallmark mutation that occurs during late stages of PanIN progression to human PDAC (Hruban *et al.*, 2000). *SMAD4* is inactivated in 55% of human PDAC cases, either through somatic mutation (~25%) or through homozygous deletion (~30%) (Hahn *et al.*, 1996). *SMAD4* is a co-transcription factor and the central effector of the canonical transforming growth factor-β (TGFβ) signalling pathway, which is important for the regulation of cellular growth, differentiation and tissue homeostasis (Shi *et al.*, 2003). The canonical TGFβ signalling pathway is activated through the binding of TGFβ to type I receptor TGFβ receptor 1 (TGFBR1) and type II receptor TGFBR2 hetero-tetrameric receptor complexes (Figure 2d). Type I and type II receptors fulfil distinct functions within the receptor complex: type I receptors act as a signal propagator while type II receptors function as activator of downstream signalling. In the setting of the canonical TGFβ signalling pathway, upon binding of TGFβ to the receptor complex TGFBR2 phosphorylates TGFBR1 thereby inducing a conformational change in TGFBR1 that facilitates the recruitment of receptor SMAD proteins (R-SMADs) (Massague, 2012). R-SMADs are than phosphorylated by TFGBR1 leading to the exposure of the nuclear localisation signal and translocation of R-SMADs into the nucleus. In the nucleus, R-SMADs form hetero-dimeric or –trimeric transcriptional complexes with SMAD4 (the only Co-SMAD) that regulate context-specific expression of broad sets of genes. The TGFβ signalling pathway is notable for its dualistic role in the development and progression of pancreatic cancer: during early PanIN stages it is a potent inhibitor of the cellular growth

and activator of apoptosis in pre-neoplastic cells, whereas it is promoting neoplastic growth in more advanced PanIN stages or human PDAC when the canonical TGFβ signalling pathway is already inactivated (Siegel *et al.*, 2003; Whittle *et al.*, 2015). Alternative to *SMAD4* mutations, about 10% of human PDAC cases harbour inactivating mutations in *TGFBR1*, *TGFBR2*, activin A receptor type 1B (*ACVR1B*) or *SMAD3*. These mutations tend to be mutually exclusive to *SMAD4* inactivation and provide alternative mechanisms to alter the canonical TGFβ signalling pathway (Jones *et al.*, 2008; Biankin *et al.*, 2012). *KRAS*, *CDKN2A*, *TP53* and *SMAD4* alterations are present in more than 50% of human PDAC cases and are thus hallmark mutations of the disease (Waddell *et al.*, 2015). The study of PanIN precursor lesions supported a stepwise progression model with the gradual

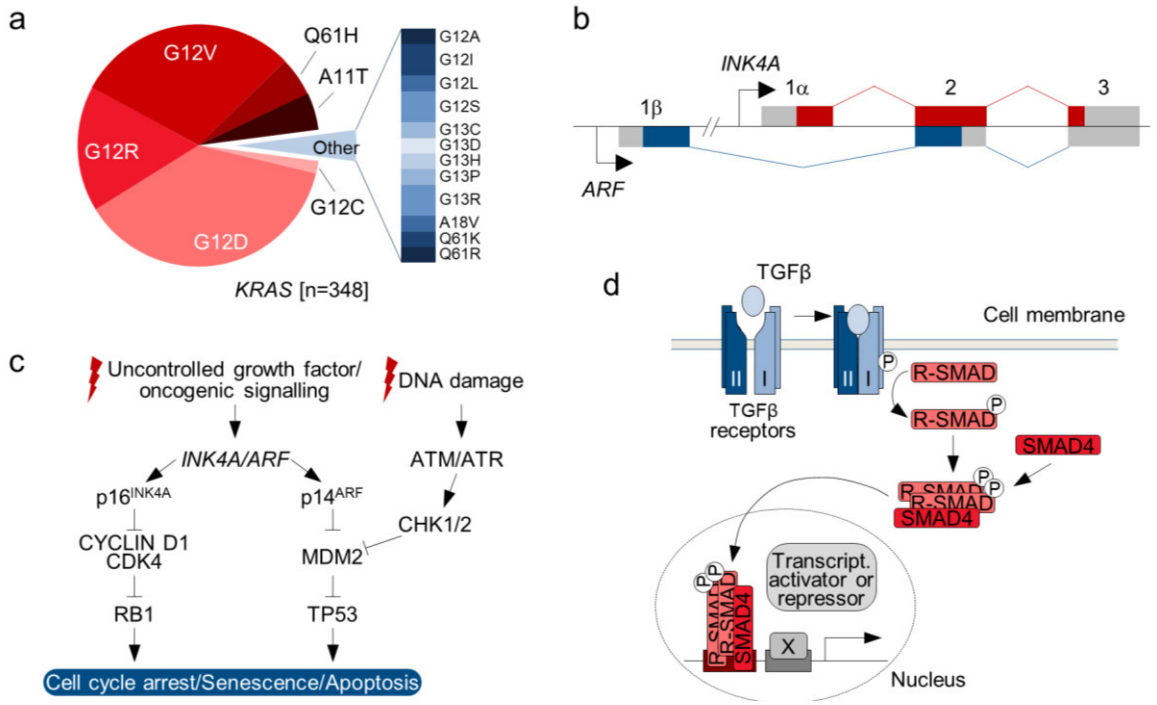

Figure 2 | Signature mutations of pancreatic ductal adenocarcinoma.

[a] Activating mutations in *KRAS* are not randomly distributed along the gene with hotspot alterations of amino acids G12 (~90%) and Q61 (~5%) (data from QCMG study, cBioPortal). **[b]** Structural organization of the *CDKN2A* locus encoding for *INK4A* and *ARF* tumour suppressor genes. *INK4A* and *ARF* coding exons are coloured in red and blue, respectively. Transcription of *INK4A* and *ARF* is induced from distinct promoters upstream of exon 1α and 1β. The transcript is then spliced to shared exons 2 and 3 resulting in different reading frames for *INK4A* and *ARF*. **[c]** *INK4A* and *ARF* are not expressed from the *CDKN2A* locus under physiological conditions but are readily induced by oncogenic signalling. INK4A protein blocks the binding of CYCLIN D1 to CDK4, thereby preventing the phosphorylation of RB1 and cell cycle progression. ARF protein binds to MDM2 E3 ubiquitin-protein ligase which results in stabilization of TP53 protein and activation of TP53 effector pathways. The activity of MDM2 protein can also be reduced through its phosphorylation by CHK1/2 during DNA damage signalling, thereby leading to TP53 stabilization and activation. **[d]** The canonical TGFβ signalling pathway is activated by the binding of TGFβ ligands to type I and type II TGFβ receptors. Ligand binding induces phosphorylation of the type I receptor by the type II receptor. The activated type I receptor induces downstream signalling by phosphorylation of R-SMADs, subsequent hetero-dimerization/trimerization of R-SMADs with SMAD4 and translocation to the nucleus where the SMAD complex can activate or repress gene transcription.

accumulation of somatic genetic alterations along increasing cellular atypia (Hruban *et al.*, 2000). This pattern indicates that the acquisition of necessary mutations is associated with waves of clonal expansion. Recent studies challenged the linear progression model and showed that an alternative model of punctuated evolution might be also relevant during the progression of a subset of human PDAC (Waddell *et al.*, 2015; Notta *et al.*, 2016). In contrast to linear evolution, punctuation describes the accumulation of genomic and phenotypic changes in a small/unobservable niche that then lead to a very fit phenotype that quickly sweeps through the population (Figure 3). For example chromothripsis is defined as a phenomenon by which tens to hundreds of chromosomal alterations occur during a single catastrophic genomic event within a single cell cycle (Stephens *et al.*, 2011; Korbel *et al.*, 2013). Chromothripsis is observed in 10 to 67% of pancreatic cancers (Waddell *et al.*, 2015; Notta *et al.*, 2016). The first study by Waddell *et al.* did not find evidence for the hypothesis that the occurrence of chromothripsis is a major mechanism of driver gene accumulation. By contrast, the study performed by Notta *et al.* observed chromothripsis in approximately two thirds of all analysed cases of pancreatic cancer. In ~16% of cases the occurrence of chromothripsis caused combined genetic alterations in *KRAS*, *CDKN2A*, *TP53* and/or *SMAD4* (predominantly affecting two genes) (Notta *et al.*, 2016). Knockout of multiple tumour suppressor genes within a single genomic catastrophe (punctuated evolution) might be a reasonable model to explain why human PDAC is typically in an advanced/invasive stage when first diagnosed (Figure 3). Although mutational phenomena such as chromothripsis have been linked to aggressive tumour behaviour in other cancer entities (Rausch *et al.*, 2012), punctuated evolution is detected only in the minority of PDAC cases indicating that linear progression is still the most plausible pathway of PDAC carcinogenesis.

Beyond the well-organized accumulation of genetic alterations in a few key driver genes during PanIN-to-PDAC progression, the sequencing of more than 500 PDAC exomes as well as more than 100 PDAC genomes revealed the presence of extensive mutational heterogeneity in PDAC (Jones *et al.*, 2008; Biankin *et al.*, 2012; Bailey *et al.*, 2016) (Figure 4a). These mutations typically occur in genes involved in SWI/SNF mediated chromatin remodelling, DNA damage repair or axon guidance pathway as well as in well-known oncogenes (Waddell *et al.*, 2015). However, only a few genetic alterations, such

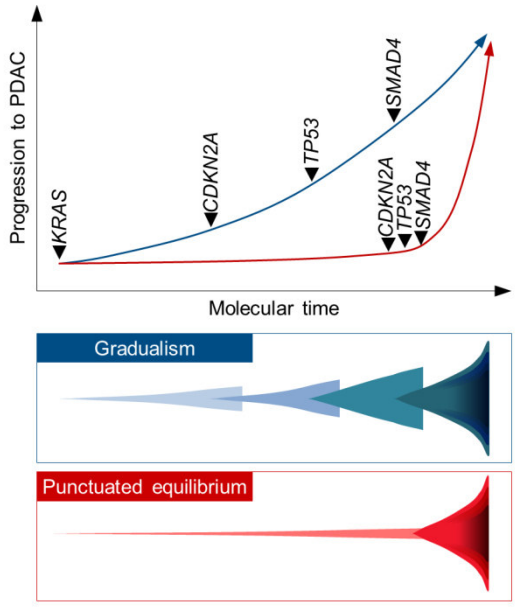

Figure 3 | Progression models of pancreatic ductal adenocarcinoma.
Two theoretical models of PDAC evolution and progression are shown. The classical model (in blue) is characterized by the stepwise accumulation of somatic gene alterations that are interspersed by waves of clone selection and expansion. In contrast, the alternative model (in red) is characterized by an extended phase of preneoplastic growth, simultaneous acquisition of driver gene alterations within a short time frame through genomic catastrophes (punctuation) and subsequent rapid tumour progression/outgrowth. Figure modified from Notta *et al.*, 2016)

as in *BRCA1/2*, *PALB2*, *MYC* or *KDM6A*, have been attributed to certain phenotypes of pancreatic cancer. For example, genetically unstable human PDACs are significantly associated with homozygous inactivating mutations of *BRCA1/2* or *PALB2* and respond well to platinum-based therapy while wildtype tumours do not (Waddell *et al.*, 2015) (Figure 4b). *MYC* amplifications, *KDM6A* mutations or upregulated expression of the *TP63ΔN* transcriptional network were reported to be involved in squamous differentiation of human PDAC, although functional evidence is still missing (Witkiewicz *et al.*, 2015; Bailey *et al.*, 2016). For the majority of the remaining heterogenic mutations it is not clear how they shape fundamental biological and pathological phenotypes of human PDAC, such as de-differentiation or metastatic dissemination. Strikingly, a recent study found that all driver mutations were shared between the primary pancreatic tumour and the corresponding metastases. Genetic alterations which did not have known or predicted functional consequences (passenger mutations) contributed to all the mutational heterogeneity that emerged in the metastases (Makohon-Moore *et al.*, 2017) (Figure 4c). To explain the lack of new driver mutations a complementary work investigated the epigenomic changes (methylation and acetylation of histone residues) that arise from the primary tumour to distant metastatic lesions. They found that large blocks of chromatin modifications were changed from the primary pancreatic tumour to the corresponding metastases and that this was reflected on the transcriptional level by the disproportional dependency of the metastases for glucose that gets metabolised through the oxidative branch of the pentose phosphate pathway (McDonald *et al.*, 2017). The first study suggests a fundamental hypothesis, that a primary pancreatic cancer has acquired already all mutations that are necessary for its metastatic dissemination. The second study suggests that this can be

explained on the basis of the epigenomic and metabolic progression from the primary tumour to the metastatic lesion. However, both studies leave many questions unanswered: (i) is there a genetic determinant of PDAC metastasis that is already present in the primary tumour (ii) and how is that related to the progression during PDAC carcinogenesis, (iii) are the epigenomic changes observed in the metastatic lesions necessary for the dissemination of tumour cells and (iv) are these changes already present in sub-clones of the primary tumour and when are they selected for. From a clinical point of view, metastatic dissemination of tumour cells is the major cause of PDAC associated death (Yachida *et al.*, 2009). Reflected by the limited understanding of the metastatic process, the 5-year survival rates of pancreatic cancer patients have stayed below 5% during the last 30 years. Further, PDAC is expected to become the second leading cause of cancer-related deaths world-wide within the next decade (Rahib *et al.*, 2014). In current treatment strategies, pancreatic cancer patients receive systematic chemotherapy using Gemcitabine or FOLFORINOX (a combination therapy of folinic acid, 5-fluoruracil, irinotecan and oxaliplatin) (Cid-Arregui *et al.*, 2015). Due to the combination of four cytotoxic drugs, FOLFORINOX is a highly aggressive treatment option that is in most cases only tolerated in young pancreatic cancer patients. A new treatment modality is the combinatorial

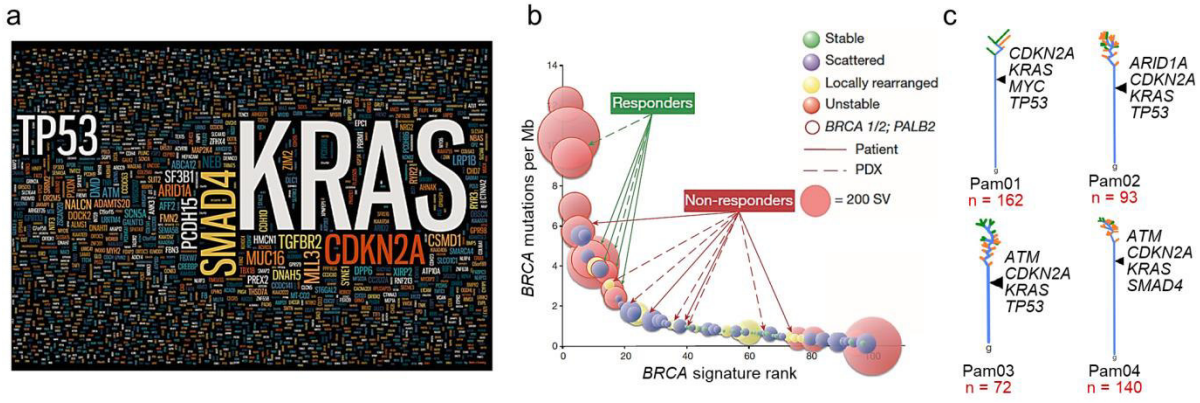

Figure 4 | Mutational landscape and phenotypes of pancreatic ductal adenocarcinoma.
[a] Word cloud of genes mutated in human pancreatic cancer. The frequency of the gene mutation is correlated to the word size. *KRAS*, *CDKN2A*, *TP53* and *SMAD4* are the most frequent alterations that dominate the genetic landscape of pancreatic cancer. Beyond these signature alterations, additional mutated genes show extensive heterogeneity (figure from Cowley *et al.*, 2013). **[b]** *BRCA* mutational signature is a potential biomarker for the response of pancreatic cancer patients to therapy with platinum and PARP inhibitors. Patients are ranked by the prevalence of *BRCA* signature mutations (x axis) and *BRCA* signature mutation load (y axis). The size of the circles corresponds to the structural variant load in each patient. The response of patients to Platinum-based therapy is indicated (modified from Waddell *et al.*, 2015). **[c]** Evolution of driver gene alterations in 4 treatment-naive metastatic human pancreatic cancer patients. Time is inferred from the number of somatic alterations and is indicated by the length of the tree. Divergence of analysed lesions is indicated by the width of the tree. The trunk (blue line) starts at the germline (g) and gives rise to primary pancreatic cancer (green line) and metastasis (orange line). In all four cases, driver gene alterations occur at the trunk of tumour evolution and are shared between all lesions of an individual patient (modified from Makohon-Moore *et al.*, 2017).

treatment using nab-paclitaxel (nanoparticle albumin–bound paclitaxel) plus Gemcitabine that was shown to significantly prolong the median overall survival of pancreatic cancer patients as compared to treatment with Gemcitabine only (Von Hoff *et al.*, 2013). However, applying the most advanced and aggressive systematic treatments can only minimally extend the life span of pancreatic cancer patients that will finally succumb to recurrence of metastatic pancreatic cancer. The poor survival of pancreatic cancer patients is the consequence of the still limited understanding of fundamental processes in PDAC biology and phenotypic diversification. Accordingly, it is not surprising that no early detection strategies were developed yet and that effective drugable targets remain to be discovered for PDAC.

1.2 Mouse models of human PDAC

Early studies in the 1990ths identified the four signature mutations in *KRAS*, *CDKN2A*, *TP53* and genes of the canonical TGFβ signalling pathway that are present in more than 50% of all human PDAC genomes (Hruban *et al.*, 2000; Waddell *et al.*, 2015). The knowledge of the signature mutations provided the basis for the PanIN-to-PDAC progression model through the stepwise accumulation of these genetic alterations in PanINs (Figure 1b). However, functional testing of the observed mutations and progression patterns in an *in vivo* context was not possible before the development of a pancreatic cancer mouse model by the Tuveson laboratory in 2003.

The study cancer in mice began already in 1978 with the generation of inbred mouse strains that showed increased susceptibility to spontaneous or carcinogen-induced tumourigenesis (reviewed in Jonkers *et al.*, 2002). This approach was however very time consuming and required consecutive brother-sister-matings for more than 20 generations making it infeasible for the systematic study of cancer in mice. A major breakthrough was the discovery that mice generated from oocytes injected with DNA fragments encoding for oncogenes were prone to the development of certain tumour types. For example the injection of Myc-coding DNA fragments lead to the development of breast cancers (Stewart *et al.*, 1984) or the injection of simian virus 40 (SV40) -derived large T antigen caused brain tumours (Brinster *et al.*, 1984). A second breakthrough came a few years later through the discovery that specific mutations can be targeted to the endogenous locus of genes in mouse embryonic stem (ES) cells. Targeting of genes in ES cells by homologous recombination (also known as gene targeting) allowed for the transfer of engineered genetic modifications through the mouse germ line (Robertson *et al.*, 1986). One major

disadvantage of these “conventional” mouse models was that the genetic modifications were present throughout all cells of the mouse, which led to unwanted carcinogenesis or embryonic lethality (depending on the gene mutation or knockout). This critical limitation led to the development of conditional mouse models in a third technological advancement. These third generation mouse models typically utilize Cre/LoxP or Flp/FRT site specific recombination systems (Sauer, 1998) for tissue- or time-specific gene activation/inactivation to circumvent undesired carcinogenesis and/or embryonic lethality. In addition they also allow for the study of spontaneous carcinogenesis because not all cells of the mouse body are mutated, such as cells of the microenvironment.

1.2.1 The *Kras^{G12D}*-driven mouse model of pancreatic cancer:

Different genetic or chemical approaches have been tested to establish a pancreatic cancer mouse model in the 1980s (Corbett *et al.*, 1984; Longnecker, 1984). However, it was the generation of the *Kras$^{LSL\text{-}G12D}$* mouse in 2001 (Jackson *et al.*, 2001) that for the first time facilitated the expression of a hotspot activating *Kras* mutation at physiological levels from the endogenous locus. In that mouse, the expression of the *Kras^{G12D}* mutation from the endogenous locus is first blocked by STOP cassette which has been introduced upstream of exon 2 (containing the G12D mutation). The transcriptional STOP cassette is flanked by identically orientated LoxP sites that allow for Cre-mediated recombination of LoxP sites, deletion of the STOP cassette and tissue- or time-specific activation of the *Kras^{G12D}* allele. The G>D mutation in codon 12 is one of the *KRAS* hotspot mutations occurring in human pancreatic cancer. The combination of the *Kras$^{LSL\text{-}G12D}$* allele with pancreas-specific expression of Cre recombinase in transgenic Pdx1-Cre or knock-in *Ptf1a^{Cre}* driver lines was used to establish a reliable mouse models for pancreatic cancer (Hingorani *et al.*, 2003) (Figure 5a). In all mice, the expression of *Kras^{G12D}* in pancreatic cells at physiological levels induced premalignant lesions reminiscent of human PanINs. In the mouse, PanINs progressed at low frequency to invasive and/or metastatic PDAC. The model recapitulated the full spectrum of the human disease with stepwise progression of pancreatic cells to precursor lesions to invasive/metastatic carcinoma and revealed that activating *Kras* mutations are central for pancreatic carcinogenesis (Figure 5a). It also provided the first functional evidence *in vivo* for the stepwise progression model of PDAC that was already proposed earlier on the basis of observations in human patients (Hruban *et al.*, 2000). The development of a reliable pancreatic cancer mouse model provided the basis for the systematic functional analyses of genetic alterations that co-occur with activating *KRAS* mutations during PanIN-to-PDAC progression.

1.2.2 The *Kras^{G12D};Cdkn2a$^{Flox/Flox}$*-driven mouse model of pancreatic cancer:

The most frequent tumour suppressor gene alteration that is observed in 80 to 95% of all human PDAC affect the *CDKN2A* locus and occurs already in early PanIN lesions (Schutte *et al.*, 1997; Hruban *et al.*, 2000) (Figure 1b). Because of the already well described role of the *CDKN2A* locus for tumour suppression in various tissues, it was thought that this locus is also highly relevant for suppressing tumourigenesis in the pancreas as well. First functional evidence for this hypothesis came from a mouse model with homozygous knockout of the *Cdkn2a* locus (comprising *Ink4a* and *Arf* gene products) in the pancreas. Full knockout of *Cdkn2a* in the mouse pancreas did not lead to the development of pancreatic cancer or any preneoplastic lesions. However, full knockout of the *Cdkn2a* locus in combination with expression of the activating *Kras^{G12D}* mutation (PKC mice) resulted in early appearance of PanIN lesions and the rapid progression to highly invasive, micro-metastatic PDAC (Aguirre *et al.*, 2003) (Figure 5b). In comparison to the *Kras^{G12D}* model, the survival of *Kras^{G12D}*;*Cdkn2a$^{Flox/Flox}$* mice was dramatically reduced from more than 12 months to only 2 month suggesting that inactivation of the *Cdkn2a* locus is critical for PanIN progression and PDAC development (Aguirre *et al.*, 2003). Pancreatic tumours of the PKC mouse model showed a strong invasion into surrounding/nearby organs such as duodenum, stomach and spleen. Metastases to the liver and lung were less frequent, however invasion of the lymph node was frequently observed, indicating metastatic capability of the tumour cells. It was speculated by the authors of the study, that distant metastases to the liver and lung were less frequent because of the rapid primary tumourigenesis. The extremely shortened survival of PKC mice is probably not allowing for the outgrowth of micro-metastases at the distant organ. All primary pancreatic tumours derived from the PKC model stained positive for the ductal marker CK19, indicating duct-like differentiation of the tumour cells. In conjunction with the appearance of neoplastic, glandular structures, most tumour cases were characterized by the appearance of tumour cells with spindle-cell morphology or marked cytoplasmic and nuclear pleomorphism representing sarcomatoid/anaplastic pancreatic carcinoma variants (Aguirre *et al.*, 2003; Bardeesy *et al.*, 2006a) (Figure 5b). Interestingly, the frequency of sarcomatoid/anaplastic differentiated pancreatic tumour areas was strongly reduced in *Kras^{G12D}*;*Trp53$^{Flox/Flox}$* and *Kras^{G12D}*;*Trp53$^{R172H/+}$* -driven mouse models (Hingorani *et al.*, 2005; Bardeesy *et al.*, 2006a).

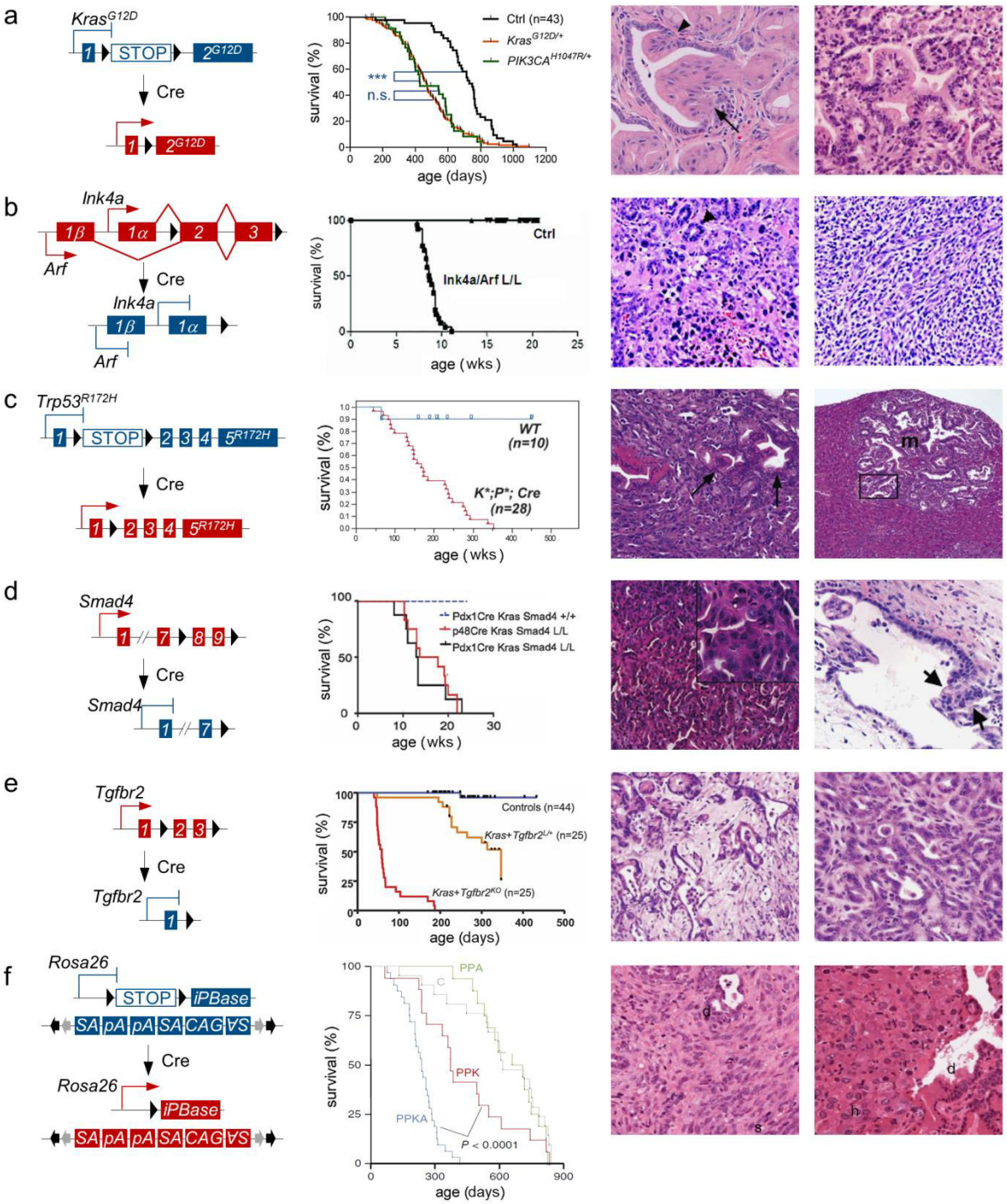

Figure 5 | Overview of human PDAC mouse models analysed in this thesis.

All mouse alleles shown in this overview can be activated through the expression of Cre recombinase. Pdx1-Cre and *Ptf1a*^{Cre} driver lines can be used to express Cre recombinase specifically in cells of the pancreas. **[a]** The *Kras*$^{LSL\text{-}G12D}$ allele. Expression of oncogenic *Kras*^{G12D} is driven from the endogenous locus upon deletion of the STOP cassette by Cre recombinase. The STOP cassette also results in a knockout of wildtype *Kras* before Cre recombination (also known as knockout-first allele). Expression of endogenous levels of oncogenic *Kras*^{G12D} in the pancreas induces PanIN formation with slow progression to PDAC and a median tumour latency of ~67 weeks (Kaplan-Meier survival curves from Eser *et al.*, 2013). H&E of PanIN-2 lesion (arrows) with moderate nuclear atypia and loss of epithelial polarity (left) and well-differentiated PDAC with duct-like structures (right) (H&Es modified from Hingorani *et al.*, 2003). **[b]** The *Cdkn2a*^{FL} allele. Exons 2 and 3 of the *Cdkn2a* locus are flanked by LoxP sites in equal orientation allowing for Cre-mediated

recombination of LoxP sites, deletion of both Exons and functional inactivation of *Cdkn2a*. Homozygous inactivation of *Cdkn2a* dramatically accelerates *Kras^{G12D}*-driven pancreatic carcinogenesis with a median survival of ~9 weeks (Kaplan-Meier survival curves from Aguirre *et al.*, 2003). PDAC from this model is characterized by poor differentiation (left H&E, arrowheads indicate irregular formed glands) and frequently shows areas of sarcomatoid dedifferentiation/spindle shaped cells (right H&E) (H&Es modified from Aguirre *et al.*, 2003). **[c]** The *Trp53$^{LSL-R172H}$* allele. Expression of mutant *Trp53^{R172H}* from the endogenous locus can be activated through Cre-mediated recombination of the STOP cassette. Such as in the *Kras$^{LSL-G12D}$* allele, the STOP cassette leads to a knockout of wild-type *Trp53* before Cre recombination (knockout-first allele). Heterozygous inactivation of *Trp53* through expression of mutant *Trp53^{R172H}* accelerates *Kras^{G12D}*-driven pancreatic carcinogenesis with a median survival of ~40 weeks (Kaplan-Meier survival curves from Hingorani *et al.*, 2005). PDAC from this model is characterized by well/poor differentiation (left H&E, arrowheads indicate well differentiated areas) and frequently develops distant organ metastases (right H&E, well-differentiated liver metastasis, metastasis (m)) (H&Es modified from Hingorani *et al.*, 2005). **[d]** The *Smad4^{FL}* allele. Exons 8 and 9 of *Smad4* are flanked by LoxP sites in equal orientation allowing for Cre-mediated recombination of LoxP sites, deletion of both Exons and functional inactivation of *Smad4*. Homozygous deletion of *Smad4* accelerates *Kras^{G12D}*-driven carcinogenesis in the pancreas with a median survival of ~16 weeks(Kaplan-Meier survival curves from Bardeesy *et al.*, 2006b). Only the minority of mice developed PDAC (left H&E, moderately differentiated PDAC). Tumours arising in this model typically resembled intraductal papillary mucinous neoplasia (IPMN), a precursor lesion of PDAC (right H&E, IPMN, arrowheads indicate carcinoma *in situ*) (H&Es modified from Bardeesy *et al.*, 2006b). **[e]** The *Tgfbr2^{FL}* allele. Exons 2 and 3 of the *Tgfbr2* gene are flanked by equally orientated LoxP sites allowing for Cre-mediated recombination of LoxP sites, deletion of both Exons and functional inactivation of *Tgfbr2*. Homozygous deficiency of *Tgfbr2* dramatically accelerates *Kras^{G12D}*-driven pancreatic carcinogenesis with a median survival of ~8 weeks (Kaplan-Meier survival curves from Ijichi *et al.*, 2006). PDAC from this model is characterized by well differentiation (left H&E) and the absence of distant organ metastasis. Right H&E shows PDAC with well-differentiation and nuclear atypia (both H&Es modified from Ijichi *et al.*, 2006). **[f]** Conditional *piggyBac* transposition system for genome-wide screening in mice. Expression of *piggyBac* transposase (*iPBase*) from the *Rosa26* promoter can be established by Cre-mediated recombination of the STOP cassette. The ATP transposon contains *piggyBac* (dark arrows) and *Sleeping Beauty* inverted terminal repeats (grey arrows) allowing for mobilization by both transposases. Gene insertion of the transposon can activate or inactivate gene expression depending on its orientation (unidirectional activation of gene expression by CAG promoter, bidirectional inactivation of gene expression through two splice acceptors (SA) and Poly(A) sites (pA)). The *piggyBac* transposition system significantly accelerates *Kras^{G12D}* driven pancreatic carcinogenesis with a median survival of ~34 weeks (Kaplan-Meier survival curves from Rad *et al.*, 2015). Mice from this model develop PDAC with a wide range of distinct histological phenotypes, such as (i) poorly differentiated PDAC with areas of sarcomatoid (s) or ductal (d) differentiation (left H&E) or (ii) poorly differentiated PDAC with regions showing hepatoid (h) or ductal (d) differentiation (right H&E) (H&Es modified from Rad *et al.*, 2015).

1.2.3 *Kras^{G12D};Trp53$^{Flox/Flox}$*- and *Kras^{G12D};Trp53$^{LSL-R172H/LSL-R172H}$*-driven mouse models of pancreatic cancer:

Like *CDKN2A*, loss of *TP53* function is a hallmark of pancreatic carcinogenesis. Its common role in tumour suppression and its inactivation in 50 to 80% of PanIN cases suggested a critical role of *TP53* in controlling PanIN progression and PDAC development (Bardeesy *et al.*, 2006a) (Figure 1b). In contrast to *CDKN2A*, *TP53* is typically inactivated through homozygous deletion and/or somatic mutations. In most cases, somatic mutations occur in the DNA-binding domain (DBD) of *TP53* either through mutations affecting TP53 residues that are directly involved in DNA binding or through mutations that alter the global

structure of the TP53-DBD (Olive *et al.*, 2004). It is important to note, that *TP53* deletions and *TP53* mutations are considered to be not functionally redundant. For example, a dominant-negative effect was shown for two *Trp53* hotspot mutant alleles, namely codon mutations R172H and R270H (R175H and R273H in humans), on the function of the remaining wildtype *Trp53* allele. When thymocytes were irradiated, the apoptotic response was reduced in cells derived from mice heterozygous for mutant and wildtype *Trp53* alleles in comparison to cells derived from mice heterozygous for knockout and wildtype *Trp53* alleles. Further a gain of function property was suggested for both mutant *Trp53* alleles. While TRP53 protein is completely absent in knockout cells it was shown that mutant TRP53 protein can bind to family-member proteins TRP63 and TRP73, thereby reducing their transcriptional activity and potential tumour suppressive functions (Olive *et al.*, 2004). In principal, the additional functions described for mutant *Trp53* alleles can be ascribed either to the dominant-negative or the gain of function model, or both in combination. To study the role of *Trp53* inactivation during PanIN progression and PDAC development in mice, conditional *Trp53* knockout (*Trp53*^{Flox}) and/or mutant (*Trp53*$^{LSL\text{-}R172H}$) alleles were crossed into the *Pdx1-Cre;Kras*$^{LSL\text{-}G12D}$ background (Hingorani *et al.*, 2005; Bardeesy *et al.*, 2006a) (Figure 5c). Without expression of oncogenic *Kras*^{G12D}, homozygous deletion of *Trp53* did not result in any morphological changes or pancreatic carcinogenesis. As for the full knockout of *Cdkn2a*, complete deletion of *Trp53* in combination with *Kras*^{G12D} mutation (PKP mice) resulted in rapid tumourigenesis in the pancreas and a severely reduced median survival of 2 month as compared to 12 month of the *Kras*^{G12D} model. The histological analysis of pancreatic cancers from the PKP model revealed ductal adenocarcinomas with a predominantly well differentiated morphology (Bardeesy *et al.*, 2006a). Heterozygous expression of the dominant-negative *Trp53* hotspot mutant R172H at physiological levels from the endogenous *Trp53* locus alone did also not result in any changes of pancreatic morphology or any signs of carcinogenesis (Hingorani *et al.*, 2005). In the *Pdx1-Cre;Kras*$^{LSL\text{-}G12D}$ background, the presence of the heterozygous *Trp53*$^{LSL\text{-}R172H}$ allele (PKP^{R172H} mice) induced PDAC with frequent metastases to the liver or lung and reduced the survival of the mice to ~5 months. Similar to the PKP model with *Trp53* deletion, pancreatic tumours derived from the PKP^{R172H} mouse model were predominantly characterized by a well-differentiated morphology and represented ductal adenocarcinoma (Hingorani *et al.*, 2005) (Figure 5c). Interestingly, it was observed that the metastatic burden is increased in PKP^{R172H} as compared to PKP mice, despite similar kinetics of primary pancreatic cancer development (Morton *et al.*, 2010). Later on it was shown that the R172H missense mutation of *Trp53* inactivates growth arrest function of *Trp53* but at the same time also promotes metastatic dissemination through its gain of function activity. The sustained expression of the mutant *Trp53* allele was necessary for

elevated activity of the platelet-derived growth factor receptor β (PDGFRb) which was identified to confer metastatic capability (Weissmueller *et al.*, 2014).

1.2.4 *Kras^{G12D};Smad4$^{Flox/Flox}$*- and *Kras^{G12D};Tgfbr2$^{Flox/Flox}$*-driven mouse models of pancreatic cancer:

Approximately 65% of all human PDAC cases carry alterations in genes involved in the canonical transforming growth factor β (TGFβ) pathway. In most cases, *SMAD4* (~55%) or *TFGBR2* (~5%) are inactivated either through deletion and/or somatic mutations (Hahn *et al.*, 1996; Siegel *et al.*, 2003; Whittle *et al.*, 2015). *SMAD4* or *TGFBR2* inactivation are observed in late PanIN stages when loss of *CDKN2A* and *TP53* function already occurred (Figure 1b). Accordingly, the canonical TGFβ pathway was thought to play a central role for the progression of late PanIN stages to pancreatic cancer (Hruban *et al.*, 2000). Two different models were developed to investigate the role of *Smad4* or *Tgfbr2* deletion for PDAC initiation and progression in the mouse. As already observed for complete inactivation of *Cdkn2a* or *Trp53*, the pancreas-specific homozygous knockout of either *Smad4* (Bardeesy *et al.*, 2006b) or *Tgfbr2* (Ijichi *et al.*, 2006) alone did not result in morphological abnormalities or the induction of any malignant lesions in the mouse pancreas. The combination of pancreas-specific expression of oncogenic *Kras^{G12D}* and complete deletion of *Smad4* or *Tgfbr2* resulted in a dramatically reduced median survival of ~80 days and 59 days, respectively (Figure 5d,e). *Kras^{G12D}*;*Smad4^{Flox}* mice predominantly developed pancreatic lesions resembling intraductal papillary mucinous neoplasia (IPMN, a PDAC precursor in humans) and high-grade PanINs that infrequently progressed to pancreatic carcinomas (Bardeesy *et al.*, 2006b) (Figure 5d). In contrast, *Kras^{G12D}*;*Tgfbr2^{Flox}* mice (PKT mouse model) uniformly developed well-differentiated PDAC, indicating *Smad4*-independent functions of TGFβ signalling downstream of *Tgfbr2* during *Kras^{G12D}*-driven pancreatic carcinogenesis (Figure 5e). Although tumours from *Kras^{G12D}*;*Tgfbr2^{Flox}* mice were very large, they typically showed minimal invasion into the duodenum or lymph nodes and no distant metastasic lesions in the liver or lung were observed. Only three mice with extraordinary long survival of 170 to 190 days developed invasive PDAC with metastatic dissemination to the liver, lung and/or diaphragm (Ijichi *et al.*, 2006). *Kras$^{G12D/+}$*;*Tgfbr2$^{Flox/+}$* mice with heterozygous knockout of *Tgfbr2* showed an increased median survival of ~300 days, developed PDAC with similar histology but with more frequent metastatic spread. Notably, PDAC cell lines derived from these mice retained the wildtype allele of *Tgfbr2* suggesting that complete *Tgfbr2* inactivation is not required for PDAC formation but can further accelerate rapid primary pancreatic carcinogenesis (Ijichi *et al.*, 2006).

1.2.5 Mouse models of pancreatic cancer using transposon-based insertional mutagenesis and CRISPR/Cas9:

The technological progress in pancreatic cancer mouse models during the last decades led to the establishment of a *Kras*^{G12D}-driven pancreatic cancer mouse model that closely resembles the pathobiology of the human disease. The long latency of the PK mouse model also allowed for studying the cooperation of additional genetic alterations with oncogenic *Kras*^{G12D}. At the same time, the development of next-generation sequencing technologies revolutionized the analysis of the cancer genome but also revealed a previously underestimated genetic complexity of cancer. However, the study of genetic alterations found in human PDAC and their involvement in disease progression is challenging and time-consuming when engineered in the classical PK mouse model. For instance, the long-standing debate whether *TP53* acts oncogenic or tumour suppressive during tumorigenesis could only be resolved after extensive studies in mouse models with various different engineered *Trp53* alterations (reviewed in Attardi *et al.*, 2005). In principle, insertional mutagenesis strategies can be applied to the PK mouse model to accelerate and simplify the discovery and functional analysis of cancer genes *in vivo*. Transposable elements that were used for insertional mutagenesis in invertebrates were inactivated during vertebrate evolution and are therefore not available in mice (Copeland *et al.*, 2010). The reconstruction and optimization of the Tc1/Mariner-type transposon Sleeping Beauty (SB) was the first transposable element allowing for insertional mutagenesis in the mouse genome. The application of the SB system in the PK mouse model (PK-SB mice) confirmed many genes and pathways previously identified in human PDAC studies. *Usp9x* was identified as the top hit in the screen and inactivation of *USP9X* was previously thought to improve survival in human cancer. However, the insertion pattern of SB transposons in the *Usp9x* gene and additional experiments revealed that *Usp9x* acts as tumour suppressor through acceleration of tumorigenesis and blocking anoikis of pancreatic cancer cells (Perez-Mancera *et al.*, 2012). The PiggyBac (PB) transposon system was isolated from the cabbage looper moth Trichoplusia ni (Ding *et al.*, 2005) and represents a complementary tool to SB for insertional mutagenesis in mice (Rad *et al.*, 2010). Application of the PB transposon system in PK mice (PK-PB mice) confirmed previously known tumorigenic processes and identified several new pancreatic cancer genes (Rad *et al.*, 2015) which were not found in the PK-SB mouse model. For example, PB transposon insertional mutagenesis identified (i) *Foxp1* as an oncogenic transcription factor driving pancreatic carcinogenesis and invasion, (ii) a non-coding region as a *Cdkn2a* cis-regulatory region and (iii) *Fign* as a driver of hepatoid differentiation in pancreatic cancer (Rad *et al.*, 2015) (Figure 5f). The results demonstrate that insertional mutagenesis

is a valuable tool to screen for and functionally catalogue cancer drivers *in vivo*, such as genes or non-coding genomic regions that are up-/downstream of frequently mutated cancer genes. These genes are otherwise difficult to identify by conventional next-generation sequencing approaches.

The recent progress in cancer genome analyses and *in vivo* screening approaches dramatically increased the list of identified, putative cancer genes. Although these genes are found to be mutated more frequently than expected by chance, they still require functional validation experiments as *bona fide* cancer drivers. Classical mouse models represent the gold-standard for the functional validation of candidate cancer genes. The long time frames of allele generation and model intercrossing however limits there utility for high-throughput cancer gene validation. The discovery of the Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR)/Cas system as a mechanism of adaptive immunity in certain bacteria and archaea opened a new era of genome engineering (Barrangou *et al.*, 2007). The CRISPR-associated Cas9 protein is a dsDNA endonuclease that can by targeted to a genomic region through a programmable single-guide RNA (sgRNA) (Mali *et al.*, 2013). CRISPR/Cas9 was recently established for genome manipulation in human and mouse cells *in vitro* (Jinek *et al.*, 2012; Wiedenheft *et al.*, 2012; Cong *et al.*, 2013). Later on, CRISPR/Cas9 was shown to be also functional in mice thereby dramatically reducing the time frames required for model generation and allowing for high-throughput cancer gene validation *in vivo* (Platt *et al.*, 2014; Sanchez-Rivera *et al.*, 2014; Xue *et al.*, 2014; Weber *et al.*, 2015). On the basis of the PK pancreatic cancer mouse model, CRISPR/Cas9 was used for the knockout of single tumour suppressor genes in the pancreas (Chiou *et al.*, 2015; Mazur *et al.*, 2015). The use of a transfection-based method for the delivery of CRISPR/Cas9 components to the adult pancreas of mice, allowed for the simultaneous knockout of multiple tumour suppressor genes *in vivo*. The multiplexed knockout of tumour suppressor genes dramatically accelerated pancreatic carcinogenesis and allows for the rapid generation of complex genotypes, e.g. for gene cooperation analysis, engineering of chromosomal rearrangements such as deletions or translocations, and high-throughput cancer gene validation *in vivo* (Maresch *et al.*, 2016).

The technological advances in next-generation sequencing, cancer gene screening in mice and the discovery of CRISPR/Cas9 for genome engineering opened a new era of molecular biology that will propel the understanding of the genetic complexity of (pancreatic) cancer.

1.3 Genetic instability, aneuploidy and gene dosage in cancer evolution

It has been almost more than 100 years ago, when Boveri observed abnormal chromosome complements in the genome of cancer cells (reviewed in Schvartzman *et al.*, 2010). Until today, the role of chromosomal alterations as a passenger or driver of cancer initiation and progression is still a central issue in the field of cancer research. This is remarkable, since the alteration of chromosomes is one of the major genetic phenotypes that distinguish a malignant from a healthy cell (Figure 6a,b). From a genetic point of view, two important distinctions have to be made: (i) chromosomal instability (CIN) is the un-physiological high frequency of chromosome gain or loss whereas (ii) aneuploidy is the presence of an abnormal number of chromosomes as compared to the normal/diploid state. While CIN is generating aneuploidy, aneuploid cells can maintain a uniform, stable karyotype and therefore not all aneuploid cells show signs of a CIN phenotype. CIN can be further sub-stratified into whole chromosome instability (W-CIN) and segmental chromosome instability (S-CIN, including translocations, focal deletions/amplifications) (Schvartzman *et al.*, 2010). Human tumour karyotypes often show combinations or all features of aneuploidy, W-CIN and/or S-CIN. It was shown that on average about 25% of a typical cancer genome is affected by whole arm or whole chromosome somatic copy number alterations (CNAs). In addition, typically 10% of the cancer genome was found to be altered by focal somatic CNAs (Gordon *et al.*, 2012). In the past years, detailed studies of cancer genomes and the use of mouse models have clearly shown that structural or segmental rearrangements can drive tumourigenesis through the activation of oncogenes (such as recurrent *IGH-MYC* translocations in Burkitt lymphoma (Mitelman *et al.*, 2007) or the inactivation of tumour suppressor genes (such as translocations interrupting the sequence of *CDKN2A* or *SMAD4* in PDAC (Waddell *et al.*, 2015)). However, the role of establishing an aneuploid karyotype during tumour development has received considerably less attention than CIN. Alterations affecting genes of the DNA damage response, such as inactivating mutations of *TP53*, *ATM*, *BRCA1/2*, result in the acquisition of a mutator phenotype of the cancer cell with high rates of somatic mutations and/or structural chromosomal alterations (Gordon *et al.*, 2012). There are also multiple pathways of mitotic failure by which a cell can in principal establish an aneuploid karyotype with gains and/or losses of chromosomes. However, and in contrast to frequent alterations of DNA damage pathway genes, mutations in the mitotic checkpoint machinery are extremely rare making failure of the mitotic checkpoint unlikely to be compromised in cancer (Schvartzman *et al.*, 2010). Interestingly, it was found that the mitotic checkpoint is often

over-activated, rather than reduced, in chromosomally unstable tumour cells and that this hyperactivation may come along with the inactivation of major tumour suppressor pathways leading to the acquisition of an aneuploid and/or CIN phenotype (reviewed in Schvartzman *et al.*, 2010). This also raises the important question if aneuploidy is just a bystander of tumourigenesis or if it is an active driver of cellular transformation. Notably, regions of whole arm or whole chromosome somatic CNAs were recurrent across different cancer types and showed preference for either genetic gain or loss (but not both at the same time). This observation indicates selection of CNAs during tumourigenesis rather than being random/passanger events (Beroukhim *et al.*, 2010). However, it is much more difficult to pinpoint the underlying genes in whole chromosome aneuploidy as compared to translocations where exact/recurrent breakpoints or fusion genes are known.

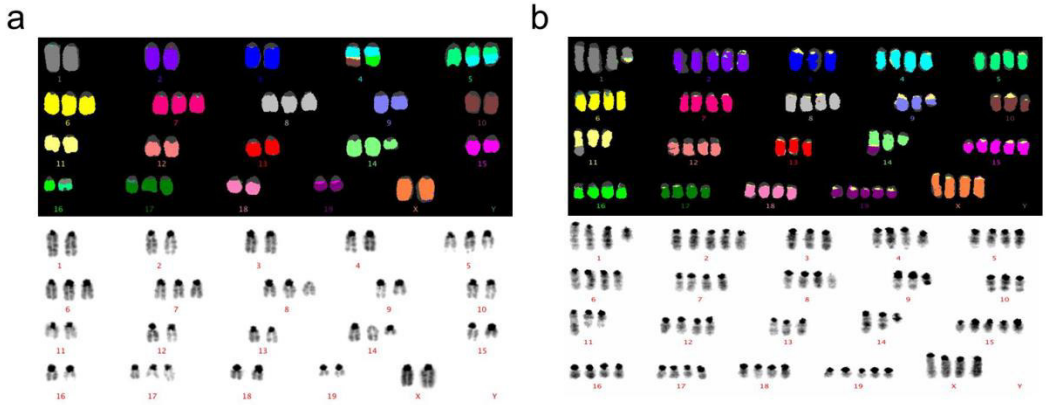

Figure 6 | Chromosome alterations in a *Kras*^{G12D}-driven mouse model of human PDAC.
M-FISH karyotypes of a cancer cell cultures established from primary PDACs of the PK mouse model. **[a]** Diploid primary mouse PDAC cell culture with complex karyotypic alterations including chromosome copy number changes and translocations. **[b]** Tetraploid mouse PDAC cell culture with an additional copy of chr14. (The whole figure was modified from Mueller *et al.*, 2018).

At the organismal level, the effects of aneuploidy are usually lethal (Gordon *et al.*, 2012). In humans, it is only the trisomy of chromosome 21 in patients with Down's syndrome that is viable for many years. All other patients with different autosomal trisomies do not survive the first few years after birth (Gordon *et al.*, 2012). One major reason for the detrimental effects of aneuploidy at the organismal level might be the wide transcriptional and proteomic imbalances that are caused by the gain or loss of whole chromosomes. For example, it was shown that the alteration of protein stoichiometry and gene dosage in aneuploid cells can interfere with the assembly of large protein complexes which are composed of genes located on different chromosomes, such as the ribosome, thereby compromising cellular homeostasis and triggering a proteotoxic stress response (Torres *et al.*, 2007; Torres *et al.*, 2010). In healthy cells, unfolded, misfolded or aggregated protein

is salvaged by the ubiquitin-proteasome and chaperone pathways that mitigate the proteotoxic burden of unfolded proteins. In aneuploid cells, the effects of aneuploidy might overwhelm the capacity of these detoxification pathways leading to protein aggregation thereby generating an energetic and proteotoxic burden on the cells that is incompatible with live (Gordon *et al.*, 2012). The importance of protein balance for cellular fitness can also be illustrated based on the phenomenon of X chromosome inactivation in female mammals. Each cell of a female mammal carries two copies of the X chromosome, while cells from males carry only one X chromosome but an additional Y chromosome. Unlike the gene-poor Y chromosome, which contains less than 100 genes in humans, more than 1000 genes with important functions for organismal development and cell viability are located on the X chromosome. In females, this would therefore result in a double dosing of X-linked genes. To compensate for the increased doses of genes located on the X chromosome and to prevent cells from proteotoxic stress, mammalian females have developed a unique mechanism of gene dosage regulation called X-chromosome inactivation. During this process, female mammals transcriptionally inactivate one of the two copies of the X chromosomes in a highly coordinated manner (Lyon, 1961). The silenced X chromosome gets condensed into an extremely compact structure, the Barr body, and remains in a stably inactivated state (Boumil *et al.*, 2001). Accordingly, gene expression remains active on only one of both X chromosomes to ensure proper gene dosage and correct protein stoichiometry which are critical for the maintenance of cellular homeostasis/fitness, e.g. through the accurate assembly of large protein complexes.

Dosage or activity alterations of oncogenes and tumour suppressor genes are also an established mechanism for the initiation and progression of malignancies, as exemplified by high-level amplification of *MYC*, hyperactivity of RAS proteins or haploinsufficiency of *PTEN* tumour suppression in many different types of human cancer (Berger *et al.*, 2011). It was proposed by Alfred Knudson more than 40 years ago that both alleles of a tumour suppressor gene need to be inactivated, accompanied by complete loss of gene function, to initiate cancer growth (two-hit hypothesis) (Knudson *et al.*, 1976) (Figure 7a). However, it was reported in later studies that major tumour suppressor mechanisms, such as tumour suppression by *PTEN*, seem to be exquisitely sensitive to dosage changes. Using a hypermorphic *Pten* (*Pten*^{hy}) allele in the mouse, it was shown that expressing 80% of normal *Pten* levels can already induce tumours of different spectrum (Alimonti *et al.*, 2010). Breast tumours were most frequent in this model, retained both *Pten* copies and *Pten* expression levels above heterozygosity. In combination with a *Pten* null allele (*Pten*$^{-}$) an allelic series with decreasing levels of *Pten* expression was created with *Pten*$^{+/+}$ > *Pten*$^{hy/+}$ > *Pten*$^{-/+}$ > *Pten*$^{hy/-}$. Critically, activation of Akt signalling was correlated to *Pten* levels in

mammary tumours and distinct levels of reduced *Pten* expression induced hyperproliferation in a tissue-specific manner. In line with these observations, the survival of the mice as well as the incidence of tumours scaled with decreasing levels of *Pten* expression suggesting a continuum model for *Pten*-regulated carcinogenesis (Alimonti *et al.*, 2010) (Figure 7b). Like for *Pten*, the expression level/gene dosage of oncogenes can also be a critical determinant for the biological effects of oncogenic signalling. Although the activation of oncogenes is essential for the initiation of tumorigenesis, it was found that forced overexpression of unphysiological high levels of *MYC* or oncogenic $HRAS^{G12V}$ provoked apoptosis or premature senescence in human cells, but not proliferation (Evan *et al.*, 1992; Serrano *et al.*, 1997). To test the effects of different levels of Ras hyper-activation in the mammary gland a titratable $Hras^{G12V}$ allele was generated (Sarkisian *et al.*, 2007). Low expression of $Hras^{G12V}$, comparable to endogenous, physiological Ras levels, stimulated cellular proliferation and hyperplasia of the mammary epithelium. By contrast, expression of increased $Hras^{G12V}$ levels, similar to those found in established tumours, induced premature senescence in mammary epithelial cells *in vivo* that was *Cdkn2a*-dependent and irreversible upon later $Hras^{G12V}$ downregulation. The study suggested a dosage-dependent model of mammary tumourigenesis in which only low levels of $Hras^{G12V}$ expression are tolerated to induce low-level cellular proliferation that allows for initiating tumour growth. Further upregulation of oncogenic signalling is governed by the *Cdkn2a* tumour suppressor pathway which needs to be bypassed for

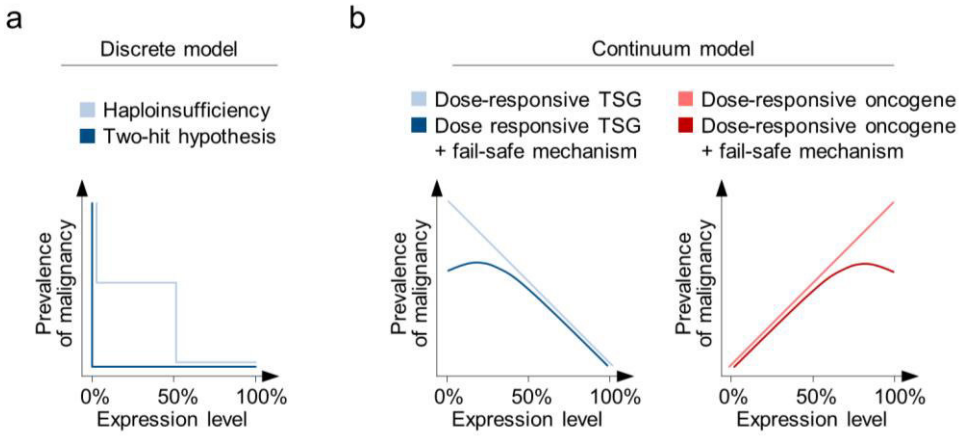

Figure 7 | Models of tumour suppression and oncogenic signalling in tumourigenesis.
[a] The discrete model of tumour suppression with discrete copy number levels of tumour suppressor genes. Tumourigenesis can be induced by loss of one copy of a tumour suppressor gene (haploinsufficiency) or requires loss of both copies of the tumour suppressor gene (two-hit hypothesis). **[b]** The continuum progression model of tumourigenesis. In contrast to the discrete model with distinct tumour suppressor gene levels, a continuum of decreasing tumour suppressor gene levels (left) or increasing oncogene levels (right) is associated with tumourigenesis in the continuum model. Of note, this correlation has not to be linear due to the induction of additional fail-safe mechanisms when tumour suppressor gene or oncogene expression falls below/exceeds a certain critical threshold. (The whole figure was modified from Berger *et al.*, 2011).

progression to full-blown cancer (Sarkisian *et al.*, 2007). This model is also in line with observations from other studies in *Kras*^{G12D}-driven mouse models of lung cancer. Restoration of functional *Trp53* in already established lung tumours caused significant but incomplete tumour regression (Feldser *et al.*, 2010; Junttila *et al.*, 2010). Tumour cell loss was specifically located to regions of high-grade adenocarcinomas, but not to low grade adenomas. Although *Trp53* was re-activated on the genetic level in all tumour cells, the selective induction of *Trp53*-regulated effector pathways was strongly correlated with marked upregulation of *Kras*^{G12D} signalling flux in high-grade adenocarcinomas. Amplification of oncogenic signalling was required for activation of *Arf* in high-grade lesions which lead to stabilization of *TRP53* protein and senescence or apoptosis. Consequently, these findings suggest that the *TRP53* pathway is not engaged by low levels of oncogenic *Kras*^{G12D} activity which are sufficient to drive cellular proliferation during early stages of lung tumour development. Therefore, activation of the *Arf*/*Trp53* tumour surveillance pathway is not triggered by the abnormal persistence of growth signals conferred by oncogenic *Kras*^{G12D}, but gets activated when *Kras*^{G12D} dosage exceeds a certain critical level (Feldser *et al.*, 2010; Junttila *et al.*, 2010). Similar to Ras, *MYC* expression is deregulated and/or elevated in many different cancer types. Since activity of *MYC* is also required for cellular proliferation it remained unclear how cells are able to discriminate between physiological and oncogenic *MYC* (Murphy *et al.*, 2008). A mouse model expressing *MycER*^{T2} from the constitutive and ubiquitous active *Rosa26* locus (*R26*$^{MycERT2/+}$ mice) with *Myc* dosages comparable to physiological *Myc* levels was generated to study the effects of *Myc* deregulation without overexpression. Low levels of deregulated were sufficient to drive ectopic proliferation of somatic cells of these mice. *R26*$^{MycERT2/MycERT2}$ mice were used to test the cellular effects of concomitant deregulation plus overexpression of *Myc* in comparison to *R26*$^{MycERT2/+}$ mice. *Myc* overexpression in homozygous mice revealed that engagement of the *Arf*/*Trp53* tumour suppressor pathway requires elevated *Myc* levels suggesting that distinct threshold levels of *Myc* governs its biological output (Murphy *et al.*, 2008). Similar observations were made in a *Braf*^{V600E}-driven mouse model of intestinal cancer. Physiologic expression levels of *Braf*^{V600E} were sufficient to drive low-level proliferation and hyperplasia. In low and high grade dysplasia, BRAF/MEK/ERK signalling gets amplified driving both tumour progression but also triggering intrinsic tumour suppression. Inactivation of the *Ink4a*/*Rb1* and *Arf*/*Trp53* tumour suppressor pathways occurred specifically in late stages during the progression to invasive carcinoma. The findings of the study suggested that the late-stage specificity of tumour suppressor gene inactivation results from the inability of low BRAF/MEK/ERK signalling levels to trigger both tumour suppressor pathways during early stages of the disease (Rad *et al.*, 2013).

It has been suggested by different studies in mice and humans that gene dosage alterations of driver genes play an important role during pancreatic cancer evolution and phenotypic diversification. For example, high level amplifications of *MYC* are associated with adenosquamous differentiation and poor survival in human PDAC (Witkiewicz *et al.*, 2015). Using fluorescence *in situ* hybridization, it was shown that the *MYC* amplifications occurred already in high-grade PanIN lesions, suggesting that the overexpression of *MYC* plays an important role during PanIN-to-PDAC progression as well as for establishing histological and aggressive phenotypes. In line with these observations, the tumours derived from a *Myc*-overexpressing mouse model of pancreatic cancer showed areas of adenosquamous differentiation that also stained positively for *Trp63*, an established marker for squamous differentiation in human tumours (Witkiewicz *et al.*, 2015). In another descriptive study of human PDAC, it was observed that the mutant-specific imbalance of *KRAS* is associated with poor prognosis and progression of PDAC to undifferentiated carcinomas of the pancreas (Krasinskas *et al.*, 2013). Furthermore, loss of the wildtype allele of *Kras* was associated with increased incidence of metastasis in a $Kras^{G12D}$-driven mouse model of PDAC with complete knockout of *Ink4a* (Qiu *et al.*, 2011). Since activating mutation of *KRAS* is an almost universal event during PDAC initiation which occurs in more than 90% of cases (Makohon-Moore *et al.*, 2016), this suggests that mutant *KRAS* has important functions beyond PDAC initiation. As already described above, it is hypothesized that intensification of oncogenic signalling is a general phenomenon during tumour evolution for the progression to invasive cancer. Notably, wildtype *Kras* was reported to have tumour suppressor activity in a lung adenocarcinoma mouse model with inverse correlation of wildtype *Kras* expression and mutant $Kras^{G12D}$ activity (Zhang *et al.*, 2001). This might explain why wildtype *Kras* was lost in more than 67% of *Kras*-mutant mouse lung adenocarcinomas and argues that allelic imbalance of mutant *Kras* represents a plausible mechanism for increasing mutant $Kras^{G12D}$ dosage during lung cancer progression. It is not known yet, if this is also relevant for the evolution of metastatic PDAC and if $Kras^{G12D}$-activity is driving different biological, histological and clinical phenotypes of PDAC.

1.4 Aim of the study

During the last thirty years the 5-year survival rate of pancreatic ductal adenocarcinoma (PDAC) stayed below 5% reflecting our still limited molecular understanding of the disease (Rahib *et al.*, 2014; Siegel *et al.*, 2016). Although the genetic alterations of more than 500 PDAC genomes and exomes have been profiled to date, only a few genetic drivers could be associated with the histological differentiation status of pancreatic tumours. Likewise, no recurrently mutated metastasis driver has been identified yet that can explain early and frequent metastasis which is typically observed in PDAC patients (Makohon-Moore *et al.*, 2017). Accordingly, while significant progress has been made in characterizing the genetics of pancreatic cancer in the last decade, the understanding of PDAC evolution and phenotypic diversification is still limited. For instance, important conceptual questions yet to be answered are: (i) How can PanIN-to-PDAC progression be explained from a genetic point of view, e.g. in a setting where all four pancreatic cancer signature mutations have been already acquired? (ii) What is, or is there any, genetic mechanism/alteration that drives metastatic dissemination? (iii) Are early progression and early/frequent metastatic dissemination connected/linked through the same molecular events? (iv) What is the molecular basis of cellular plasticity and different histopathological PDAC grading? and (v) Why is clinical aggressiveness a characteristic feature of undifferentiated pancreatic carcinomas?

The aim of this study is to comprehensively characterize primary tumours and corresponding metastasis from different *Kras*^{G12D}-driven pancreatic cancer mouse models to identify genetic drivers and molecular mechanisms that define central aspects of PDAC evolution and phenotypic diversification. In cooperation with the teams of Prof. Dr. Dieter Saur and PD Dr. Günter Schneider at TU Munich, the genomes and transcriptomes of cell cultures derived from *Kras*^{G12D}-driven pancreatic cancer mouse models are characterized. The comprehensive analysis of these cell cultures allows for overcoming critical limitations of human PDAC genome analysis, such as: (i) the high content of stroma cells in bulk samples of primary hPDAC especially complicating transcriptome and gene dosage analyses, (ii) the scarcity of publically available cell culture-based hPDAC resources, (iii) the limited availability of published metastatic hPDAC cohorts and (iv) the molecular alterations that result from chemotherapeutic treatment of patients that bias (evolutionary) analyses of the naive PDAC disease.

2 Materials & Methods

The “Materials” section contains all materials that were used by myself to perform experiments. Reagents that were used for methods conducted by collaboration partners are not included in the “Materials” section. If a method was performed by a collaboration partner or core facility, this information can be found in the “Methods” section for each individual method.

2.1 Materials

2.1.1 Reagents and enzymes

Material	Supplier
1000bp DNA Ladder	New England Biolabs
100bp DNA Ladder	New England Biolabs
20 gauche cannula	Seidel Medipool
2-Mercaptoethanol, 98%	Sigma-Aldrich
2-Propanol (isopropanol)	Carl Roth
ABgene Storage Plate, 96-well, 2.2 mL, square well, conical	Thermo Fisher Scientific
Acetic acid	Sigma-Aldrich
Adhesive PCR Plate Foils	Thermo Fisher Scientific
Agarose	Sigma-Aldrich
Ampicillin	Sigma-Aldrich
Biopsy/tissue embedding cassettes	Simport Scientific
Cell culture dishes (100 mm)	Greiner Bio-One
Cell culture flasks (50 mL, 250 mL, 550 mL)	Greiner Bio-One
Cell culture plates (6-well, 12-well, 24-well, 96-well)	Corning
Cell scrapers	Sarstedt
Collagenase Type II	Worthington Biochemical
Conical tubes (15 mL, 50 mL)	Greiner Bio-One
Cover slips	Gerhard Menzel B.V.
Cresol Red	Sigma-Aldrich
Cryotubes (1.6 mL)	Sarstedt
Deoxynucleotide (dNTP) Mix (10mM each)	Fermentas
Dimethyl sulfoxide (DMSO)	Carl Roth
DirectPCR Lysis Reagent (Mouse Tail)	Viagen Biotech
DNA LoBind Tubes (1.5 mL)	Eppendorf
Doxycycline	Sigma-Aldrich
Dulbecco's Modified Eagle Medium (DMEM), high-glucose	Sigma-Aldrich
Dulbecco's phosphate-buffered saline (DPBS), no calcium, no magnesium	Thermo Fisher Scientific
Eosine	Waldeck
Ethanol absolute	Carl Roth
Ethidium bromide	Sigma-Aldrich
Ethylenediaminetetraacetic acid (EDTA)	Sigma-Aldrich
Exonuclease I	New England Biolabs
Fetal Calf Serum Superior	Biochrom
Forene® isoflurane	Abbott
Formalin	Carl Roth

Material	Supplier
Gel Loading Dye, Purple (6x)	New England Biolabs
Glass slides SuperFrost™ Plus	Thermo Fisher Scientific
Glycerol	Sigma-Aldrich
Hard-Shell® 96-Well PCR Plates, high profile, semi skirted	Bio-Rad Laboratories
Hard-Shell® Low-Profile Thin-Wall 96-Well Skirted PCR Plate	Bio-Rad Laboratories
Hematoxylin	Merck
Hydrigen Chloride (HCl)	Carl Roth
KAPA Mouse Genotyping Kit	Sigma-Aldrich
KaryoMAX™ Colcemid™ Solution in PBS	Gibco
LB-Agar (Luria/Miller)	Carl Roth
LB-Medium (Luria/Miller)	Carl Roth
Lipofectamine® 2000	Thermo Fisher Scientific
MicroAmp® optical 96-well reaction plate	Thermo Fisher Scientific
MicroAmp® Optical Adhesive Film	Thermo Fisher Scientific
microTUBE AFA Fiber Snap-Cap 6x16mm Case	Covaris
Pasteur pipettes	Brand
PCR stripes (8 tubes)	Sarstedt
PEG6000	Sigma-Aldrich
Penicillin-Streptomycin (5,000 U/ml)	Thermo Fisher Scientific
Petri dishes (100 mm)	Greiner Bio-One
Phosphate buffered saline	Sigma-Aldrich
Pipette tips with filter (10 µL, 100 µL, 200 µL, 300 µL, 1250 µL)	Biozym Scientific
Polybrene	Sigma-Aldrich
Potassium chloride (KCl)	Carl Roth
Proteinase K	Sigma-Aldrich
Puromycin	Sigma-Aldrich
Q5® High-Fidelity DNA Polymerase	New England Biolabs
Reaction tubes safe-seal (0.5 mL, 1.5 mL, 2 mL)	Sarstedt
Reservoirs	Integra Biosciences
RNAlater	Sigma-Aldrich
Roti®-Histofix 4 %	Carl Roth
Scalpels	B. Braun Melsungen
Serological pipettes (5 mL, 10 mL, 25 mL, 50 mL)	Greiner Bio-One
Sodium hydroxide solution (NaOH)	Merck
Sucrose	Sigma-Aldrich
SuperScript II	Thermo Fisher Scientific
SurePrint G3 Custom CGH 60K microarray	Agilent Technologies
SurePrint G3 Mouse CGH 240K microarray	Agilent Technologies
SYBR™ Select Master Mix	Thermo Fisher Scientific
Syringes (1 mL, 30 mL)	B. Braun Melsungen
TaqMan™ Fast Advanced Master Mix	Thermo Fisher Scientific
TransIT®-LT1 Transfection Reagent	Mirus Bio LCC
TRIS PUFFERAN	Carl Roth
Trypsin-EDTA (0.5%)	Thermo Fisher Scientific
Tween® 20	Carl Roth
Xylene	Carl Roth

2.1.2 Buffers and solutions

Buffer / Solution	Composition
50x Tris acetate EDTA (TAE) buffer, pH 8.48	2 M TRIS PUFFERAN 5.71 % Acetic acid 50 mM EDTA

Buffer / Solution	Composition
Suc Rot	0.01 M TRIS PUFFERAN 0.04 M Cresol Red 30 % Sucrose

2.1.3 Primers

Primer	Sequence
AmpliconSeq_Apc-gDNA_F	5'-GCGAATAAGCACCACTCCTC-3'
AmpliconSeq_Apc-gDNA_R	5'-AAGAATGAACCAACACCAAGG-3'
AmpliconSeq_Arid1a-gDNA_F	5'-GTTCTGATTCCTGTGCTCGC-3'
AmpliconSeq_Arid1a-gDNA_R	5'-TCCATCACCTACCTGCTGTG-3'
AmpliconSeq_Arid1b-gDNA_F	5'-AGTTCTGGGGTACTTGGAATCA-3'
AmpliconSeq_Arid1b-gDNA_R	5'-GGTACTGCAAGCCTCCCA-3'
AmpliconSeq_Arid5b-gDNA_F	5'-TGGCTTGCACGGACCTTATA-3'
AmpliconSeq_Arid5b-gDNA_R	5'-ATCAGCAGTTGGACGGTCTT-3'
AmpliconSeq_Atm-gDNA_F	5'-TCCTTTTCAACTGTTCCTGTTACA-3'
AmpliconSeq_Atm-gDNA_R	5'-GACAATGGAAAGGCGAGTCA-3'
AmpliconSeq_Brca1-gDNA_F	5'-AGCGTGAGAACTCCTCCAAA-3'
AmpliconSeq_Brca1-gDNA_R	5'-CTGCCATGAGGAAGAACACA-3'
AmpliconSeq_Brca2-gDNA_F	5'-TCACGAGTTTCTCCGTGTCA-3'
AmpliconSeq_Brca2-gDNA_R	5'-GCTCTGGCTGTCTCGAACTT-3'
AmpliconSeq_Cdkn2a-ex1-gDNA_F	5'-TCTCACCTCGCTTGTCACAG-3'
AmpliconSeq_Cdkn2a-ex1-gDNA_R	5'-AAGTACTCCATCTCCCGGGA-3'
AmpliconSeq_Cdkn2a-ex2-gDNA_F	5'-TCAACTACGGTGCAGATTCG-3'
AmpliconSeq_Cdkn2a-ex2-gDNA_R	5'-CGGGTGGGTAAAATGGGAAC-3'
AmpliconSeq_Cdkn2b-gDNA_F	5'-CCGAAGCTACTGGGTCTCC-3'
AmpliconSeq_Cdkn2b-gDNA_R	5'-CACTTGCCCAGCTTGTACG-3'
AmpliconSeq_hKras-gDNA_F1	5'-ACGATACACGTCTGCAGTCAA-3'
AmpliconSeq_hKras-gDNA_F2	5'-TCGTCGGCAGCGTCAGATGTGTATAAGAGACAG AAAGGTACTGGTGGAGTATTTGATAGTG-3'
AmpliconSeq_hKras-gDNA_R1	5'-TCCCAAGGAAAGTAAAGTTCCCATA-3'
AmpliconSeq_hKras-gDNA_R2	5'-GTCTCGTGGGCTCGGAGATGTGTATAAGAGACA GGGTCCTGCACCAGTAATATGCA-3'
AmpliconSeq_mKras-cDNA_R	5'-GTCTCGTGGGCTCGGAGATGTGTATAAGAGACA GCCCTCATTGCACTGTACTCCT-3'
AmpliconSeq_mKras-gDNA_F	5'-TCGTCGGCAGCGTCAGATGTGTATAAGAGACAG AAGGCCTGCTGAAAATGACTGA-3'
AmpliconSeq_mKras-gDNA_R	5'-GTCTCGTGGGCTCGGAGATGTGTATAAGAGACA GACACCCAGTTTAAAGCCTTGGA-3'
AmpliconSeq_Pten-gDNA_F	5'-TGCGAGGATTATCCGTCTTC-3'
AmpliconSeq_Pten-gDNA_R	5'-CATCCGTCTACTCCCACGTT-3'
AmpliconSeq_Rosa26.1-gDNA_F	5'-TCTGATGCCCTCTTCTGGTG-3'
AmpliconSeq_Rosa26.1-gDNA_R	5'-GGCTAAACTCTGGCCCTACA-3'
AmpliconSeq_Rosa26.2-gDNA_F	5'-GGAAGGATTGTCTGTGCCCT-3'
AmpliconSeq_Rosa26.2-gDNA_R	5'-ATTTTCCAAAGCCCTCCCCA-3'
AmpliconSeq_Smad4-gDNA_F	5'-TGCAGTGTCACAGATGCTCA-3'
AmpliconSeq_Smad4-gDNA_R	5'-CTCAGGAACTGGAGGAAGCA-3'
AmpliconSeq_Trp53-gDNA_F	5'-ACATAGCAAGTTGGAGGCCA-3'
AmpliconSeq_Trp53-gDNA_R	5'-CCACTCACCGTGCACATAAC-3'
Genotyping_Cdkn2atm4Rdp_1	5'-CCAAGTGTGCAAACCCAGGCTCC-3'
Genotyping_Cdkn2atm4Rdp_2	5'-TTGTTGGCCCAGGATGCCGACATC-3'
Genotyping_Krastm1Dsa_1	5'-CACCAGCTTCGGCTTCCTATT-3'
Genotyping_Krastm1Dsa_2	5'-AGCTAATGGCTCTCAAAGGAATGTA-3'
Genotyping_Krastm1Dsa_3	5'-GCGAAGAGTTTGTCCTCAACC-3'
Genotyping_Krastm4Tyj_1	5'-CACCAGCTTCGGCTTCCTATT-3'
Genotyping_Krastm4Tyj_2	5'-AGCTAATGGCTCTCAAAGGAATGTA-3'
Genotyping_Krastm4Tyj_3	5'-CCATGGCTTGAGTAAGTCTGC-3'

Primer	Sequence
Genotyping_Ptf1atm1(cre)Hnak_1	5'-CCTCGAAGGCGTCGTTGATGGACTGCA-3'
Genotyping_Ptf1atm1(cre)Hnak_2	5'-CCACGGATCACTCACAAAGCGT-3'
Genotyping_Ptf1atm1(cre)Hnak_3	5'-GCCACCAGCCAGCTATCAA-3'
Genotyping_Tg(Pdx1-cre)6Tuv_1	5'-GCT CAT TGG GAG CGG TTT TG-3'
Genotyping_Tg(Pdx1-cre)6Tuv_2	5'-ACATCTTCAGGTTCTGCGGG-3'
Genotyping_Tg(Pdx1-cre)6Tuv_3	5'-CAC GTG GTT TAC CCT GGA GC-3'
Genotyping_Tg(Pdx1-flpo)#Dsa_1	5'-AGAGAGAAAATTGAAACAAGTGCAGGT-3'
Genotyping_Tg(Pdx1-flpo)#Dsa_2	5'-CGTTGTAAGGGATGATGGTGAACT-3'
Genotyping_Tg(Pdx1-flpo)#Dsa_3	5'-AACACACACTGGAGGACTGGCTAGG-3'
Genotyping_Tg(Pdx1-flpo)#Dsa_4	5'-CAATGGTAGGCTCACTCTGGGAGATGATA-3'
Genotyping_Tgfbr2tm1.2Hlm_1	5'-TAAACAAGGTCCGGAGCCCA-3'
Genotyping_Tgfbr2tm1.2Hlm_2	5'-ACTTCTGCAAGAGGTCCCCT-3'
Genotyping_Trp53tm1.1Dgk_1	5'-CAAGAGAACTGTGCCTAAGAG-3'
Genotyping_Trp53tm1.1Dgk_2	5'-CTTTCTAACAGCAAAGGCAAGC-3'
Genotyping_Trp53tm1Brn_1	5'-CACAAAAAACAGGTTAAACCCAGC-3'
Genotyping_Trp53tm1Brn_2	5'-GCACCTTTGATCCCAGCACATA-3'
IlluminaLibraryQuantification_F	5'-ATGATACGGCGACCACCGAG-3'
IlluminaLibraryQuantification_R	5'-CAAGCAGAAGACGGCATACGAG-3'
MycoplasmaPCR	5´-CACCTGAGTAGTATGCTCGC-3´
MycoplasmaPCR	5´-CGCCTGAGTAGTACGTACGC-3´
MycoplasmaPCR	5´-CGCCTGAGTAGTACGTTCGC-3´
MycoplasmaPCR	5´-CGCCTGAGTAGTATGCTCGC-3´
MycoplasmaPCR	5´-CGCCTGGGTAGTACATTCGC-3´
MycoplasmaPCR	5´-TGCCTGAGTAGTACATTCGC-3´
MycoplasmaPCR	5´-TGCCTGGGTAGTACATTCGC-3´
NexteraIndexPrimer_NGS_i5_S502	5'-AATGATACGGCGACCACCGAGATCTACACCTCTC TATTCGTCGGCAGCGTC-3
NexteraIndexPrimer_NGS_i5_S503	5'-AATGATACGGCGACCACCGAGATCTACACTATCC TCTTCGTCGGCAGCGTC-3'
NexteraIndexPrimer_NGS_i5_S505	5'-AATGATACGGCGACCACCGAGATCTACACGTAA GGAGTCGTCGGCAGCGTC-3'
NexteraIndexPrimer_NGS_i5_S506	5'-AATGATACGGCGACCACCGAGATCTACACACTG CATATCGTCGGCAGCGTC-3'
NexteraIndexPrimer_NGS_i5_S507	5'-AATGATACGGCGACCACCGAGATCTACACAAGG AGTATCGTCGGCAGCGTC-3'
NexteraIndexPrimer_NGS_i5_S508	5'-AATGATACGGCGACCACCGAGATCTACACCTAA GCCTTCGTCGGCAGCGTC-3'
NexteraIndexPrimer_NGS_i5_S510	5'-AATGATACGGCGACCACCGAGATCTACACCGTC TAATTCGTCGGCAGCGTC-3'
NexteraIndexPrimer_NGS_i5_S511	5'-AATGATACGGCGACCACCGAGATCTACACTCTCT CCGTCGTCGGCAGCGTC-3'
NexteraIndexPrimer_NGS_i7_N701	5'-CAAGCAGAAGACGGCATACGAGATTCGCCTTAG TCTCGTGGGCTCGG-3'
NexteraIndexPrimer_NGS_i7_N702	5'-CAAGCAGAAGACGGCATACGAGATCTAGTACGG TCTCGTGGGCTCGG-3'
NexteraIndexPrimer_NGS_i7_N703	5'-CAAGCAGAAGACGGCATACGAGATTTCTGCCTG TCTCGTGGGCTCGG-3'
NexteraIndexPrimer_NGS_i7_N704	5'-CAAGCAGAAGACGGCATACGAGATGCTCAGGAG TCTCGTGGGCTCGG-3'
NexteraIndexPrimer_NGS_i7_N705	5'-CAAGCAGAAGACGGCATACGAGATAGGAGTCCG TCTCGTGGGCTCGG-3'
NexteraIndexPrimer_NGS_i7_N706	5'-CAAGCAGAAGACGGCATACGAGATCATGCCTAG TCTCGTGGGCTCGG-3'
NexteraIndexPrimer_NGS_i7_N707	5'-CAAGCAGAAGACGGCATACGAGATGTAGAGAGG TCTCGTGGGCTCGG-3'
NexteraIndexPrimer_NGS_i7_N710	5'-CAAGCAGAAGACGGCATACGAGATCAGCCTCGG TCTCGTGGGCTCGG-3'
NexteraIndexPrimer_NGS_i7_N711	5'-CAAGCAGAAGACGGCATACGAGATTGCCTCTTG TCTCGTGGGCTCGG-3'
NexteraIndexPrimer_NGS_i7_N712	5'-CAAGCAGAAGACGGCATACGAGATTCCTCTACG TCTCGTGGGCTCGG-3'

Primer	Sequence
NexteraIndexPrimer_NGS_i7_N714	5‘-CAAGCAGAAGACGGCATACGAGATTCATGAGCG TCTCGTGGGCTCGG-3‘
NexteraIndexPrimer_NGS_i7_N715	5‘-CAAGCAGAAGACGGCATACGAGATCCTGAGATG TCTCGTGGGCTCGG-3‘
QiSeq_PCR1_PB3pr_1	5’-GACGGATTCGCGCTATTTAGAAAGAGAG-3’
QiSeq_PCR1_PB5pr_1	5’-GATATACAGACCGATAAAACACATGCGTCA-3’
QiSeq_PCR1_SplAP1	5’-GTTCCCATGGTACTACTCATA-3’
QiSeq_PCR2_PB3pr_2	5’-AATGATACGGCGACCACCGAGATCTACACATGC GTCAATTTTACGCAGACTATC-3’
QiSeq_PCR2_PB5pr_2	5’-AATGATACGGCGACCACCGAGATCTACACCACG CATGATTATCTTTAACGTACGTCAC-3’
QiSeq_PCR2_SplAP2_V1.1-96	5’-CAAGCAGAAGACGGCATACGAGATCGGTNNNNN NNNTAATACGACTCACTATAGG-3’
QiSeq_qPCR_Library_qPCR2.1	5’-AATGATACGGCGACCACCGAGATC-3’
QiSeq_qPCR_Library_qPCR2.2	5’-CAAGCAGAAGACGGCATACGAGAT-3’
QiSeq_qPCR-AdapterLigation	5’-ATACGACTCACTATAGGTGACAGCG-3’
QiSeq_SplinkeretteAdapter_BottomStrand	5’-PHOS-GCGCTCGCTGTCACCTATAGTGAGTCGTA TTATAATTTTTTTTTCAAAAAAA-3’
QiSeq_SplinkeretteAdapter_TopStrand	5’-GTTCCCATGGTACTACTCATATAATACGACTCACT ATAGGTGACAGCGAGCGCT-3’
qPCR_hCDH1_F	5'-CTTTGACGCCGAGAGCTACA-3'
qPCR_hCDH1_R	5'-TCGACCGGTGCAATCTTCAA-3'
qPCR_hGAPDH_F	5'-CTCCTCCGGGTGATGCTTTT-3'
qPCR_hGAPDH_R	5'-ATGAAGGGGTCATTGATGGCA-3'
qPCR_hMMP1_F	5'-CCAGGTATTGGAGGGGATGC-3'
qPCR_hMMP1_R	5'-GTCCAAGAGAATGGCCGAGT-3'
qPCR_hPPIA_F	5'-GCCGAGGAAAACCGTGTACT-3'
qPCR_hPPIA_R	5'-CTTGGTATCCAGGCCCCTTA-3'
qPCR_hVIM_F	5'-CGGGAGAAATTGCAGGAGGA-3'
qPCR_hVIM_R	5'-AAGGTCAAGACGTGCCAGAG-3'
qPCR_mGapdh_F	5'-TGTGTCCGTCGTGGATCTGA-3'
qPCR_mGapdh_Probe	5'-[FAM]TGCCGCCTGGAGAAACCTGCC[TAM]-3'
qPCR_mGapdh_R	5'-CACCACCTTCTTGATGTCATCATAC-3'
qPCR_mKras_F	5'-GAGAACTGGGGAGGGCTTTC-3'
qPCR_mKras_Probe	5'-[FAM]TGAAGATGTGCCTATGGTCCTGG[TAM]-3'
qPCR_mKras_R	5'-TCCTGAGCCTGTTTCGTGTC-3'
SCRB-Seq_PCR-1_SINGV6	5’-/5Biosg/ACACTCTTTCCCTACACGACGC-3’
SCRB-Seq_PCR-2_i7.1-96	5’-CAAGCAGAAGACGGCATACGAGATNNNNNNNNG TCTCGTGGGCTCGG-3’
SCRB-Seq_PCR-2_P5NEXTPT5	5’-AATGATACGGCGACCACCGAGATCTACACTCTTT CCCTACACGACGCTCTTCCG*A*T*C*T*-3’
SCRB-Seq_RT-TemplateSwitch_E3V6NEXT	5'-/5Biosg/ACACTCTTTCCCTACACGACGCTCTTCCG ATCT[BC6]N10T30VN-3'
SCRB-Seq_RT-TemplateSwitch_E5V6NEXT	5’-iCiGiCACACTCTTTCCCTACACGACGCrGrGrG-3’

2.1.4 Library preparation and sequencing

Material	Supplier
Agencourt AMPure XP magnetic beads	Beckman Coulter
Agilent High Sensitivity DNA Kit	Agilent Technologies
EB buffer	Qiagen
KAPA DNA Standards and Primers for Illumina	Kapa Biosystems
KAPA HiFi HotStart ReadyMix (2x)	Kapa Biosystems
KAPA Library Quantification Kit	Kapa Biosystems
KAPA SYBR Fast qPCR ABI Mix (2x)	Kapa Biosystems
MiSeq Reagent Nano Kit v2 (300-cycles)	Illumina
NEBNext® Ultra DNA Library Prep Kit for Illumina®	New England Biolabs
Nextera XT Kit	Illumina
Sodium hydroxide (NaOH)	Carl Roth

Material	Supplier
TG NextSeq® 500/550 High Output Kit v2 (300 cycles)	Illumina

2.1.5 Plasmids

Plasmid	Supplier
pINDUCER20	Addgene, #44012
pMD2.G	Addgene, #12259
psPAX	Addgene, #12260

2.1.6 Bacteria

Bacteria	Supplier
One Shot® Stbl3™ chemically competent E. coli	Thermo Fisher Scientific

2.1.7 Cell lines

Cell Line	Provided By / Supplier
13886 PrimaryPancreaticTumour	Group of Prof. Dr. Dieter Saur/PD Dr. Günter Schneider
14390 PrimaryPancreaticTumour	Group of Prof. Dr. Dieter Saur/PD Dr. Günter Schneider
16990 PrimaryPancreaticTumour	Group of Prof. Dr. Dieter Saur/PD Dr. Günter Schneider
16992 LungMet1	Group of Prof. Dr. Dieter Saur/PD Dr. Günter Schneider
16992 PrimaryPancreaticTumour	Group of Prof. Dr. Dieter Saur/PD Dr. Günter Schneider
2259 LymphNodeMet1	Group of Prof. Dr. Dieter Saur/PD Dr. Günter Schneider
2259 PrimaryPancreaticTumour	Group of Prof. Dr. Dieter Saur/PD Dr. Günter Schneider
3202 LiverMet1	Group of Prof. Dr. Dieter Saur/PD Dr. Günter Schneider
3202 LymphNodeMet1	Group of Prof. Dr. Dieter Saur/PD Dr. Günter Schneider
3202 PrimaryPancreaticTumour	Group of Prof. Dr. Dieter Saur/PD Dr. Günter Schneider
3250 PrimaryPancreaticTumour	Group of Prof. Dr. Dieter Saur/PD Dr. Günter Schneider
4072 PrimaryPancreaticTumour	Group of Prof. Dr. Dieter Saur/PD Dr. Günter Schneider
4706 PrimaryPancreaticTumour	Group of Prof. Dr. Dieter Saur/PD Dr. Günter Schneider
4900 PrimaryPancreaticTumour	Group of Prof. Dr. Dieter Saur/PD Dr. Günter Schneider
5123 Blood	Group of Prof. Dr. Dieter Saur/PD Dr. Günter Schneider
5123 LiverMet1	Group of Prof. Dr. Dieter Saur/PD Dr. Günter Schneider
5123 PrimaryPancreaticTumour	Group of Prof. Dr. Dieter Saur/PD Dr. Günter Schneider
5320 LiverMet1	Group of Prof. Dr. Dieter Saur/PD Dr. Günter Schneider
5320 LiverMet3	Group of Prof. Dr. Dieter Saur/PD Dr. Günter Schneider
5320 LungMet1	Group of Prof. Dr. Dieter Saur/PD Dr. Günter Schneider
5320 PrimaryPancreaticTumour	Group of Prof. Dr. Dieter Saur/PD Dr. Günter Schneider
53578 PrimaryPancreaticTumour	Group of Prof. Dr. Dieter Saur/PD Dr. Günter Schneider
53631 Blood	Group of Prof. Dr. Dieter Saur/PD Dr. Günter Schneider
53631 LiverMet1	Group of Prof. Dr. Dieter Saur/PD Dr. Günter Schneider
53631 LiverMet2	Group of Prof. Dr. Dieter Saur/PD Dr. Günter Schneider
53631 LungMet2	Group of Prof. Dr. Dieter Saur/PD Dr. Günter Schneider
53631 LungMet3	Group of Prof. Dr. Dieter Saur/PD Dr. Günter Schneider
53631 PrimaryPancreaticTumour	Group of Prof. Dr. Dieter Saur/PD Dr. Günter Schneider
53646 LiverMet1	Group of Prof. Dr. Dieter Saur/PD Dr. Günter Schneider
53646 LiverMet2	Group of Prof. Dr. Dieter Saur/PD Dr. Günter Schneider
53646 LiverMet3	Group of Prof. Dr. Dieter Saur/PD Dr. Günter Schneider
53646 PrimaryPancreaticTumour	Group of Prof. Dr. Dieter Saur/PD Dr. Günter Schneider
53704 LiverMet2	Group of Prof. Dr. Dieter Saur/PD Dr. Günter Schneider
53704 LiverMet3	Group of Prof. Dr. Dieter Saur/PD Dr. Günter Schneider
53704 PrimaryPancreaticTumour	Group of Prof. Dr. Dieter Saur/PD Dr. Günter Schneider
5671 PrimaryPancreaticTumour	Group of Prof. Dr. Dieter Saur/PD Dr. Günter Schneider
5748 PrimaryPancreaticTumour	Group of Prof. Dr. Dieter Saur/PD Dr. Günter Schneider

Cell Line	Provided By / Supplier
6075 Blood	Group of Prof. Dr. Dieter Saur/PD Dr. Günter Schneider
6075 LiverMet1	Group of Prof. Dr. Dieter Saur/PD Dr. Günter Schneider
6075 LungMet1	Group of Prof. Dr. Dieter Saur/PD Dr. Günter Schneider
6075 PrimaryPancreaticTumour	Group of Prof. Dr. Dieter Saur/PD Dr. Günter Schneider
6554 PrimaryPancreaticTumour	Group of Prof. Dr. Dieter Saur/PD Dr. Günter Schneider
8028 Blood	Group of Prof. Dr. Dieter Saur/PD Dr. Günter Schneider
8028 LiverMet1	Group of Prof. Dr. Dieter Saur/PD Dr. Günter Schneider
8028 PrimaryPancreaticTumour	Group of Prof. Dr. Dieter Saur/PD Dr. Günter Schneider
8182 PrimaryPancreaticTumour	Group of Prof. Dr. Dieter Saur/PD Dr. Günter Schneider
8248 PrimaryPancreaticTumour	Group of Prof. Dr. Dieter Saur/PD Dr. Günter Schneider
8296 PrimaryPancreaticTumour	Group of Prof. Dr. Dieter Saur/PD Dr. Günter Schneider
8305 LiverMet1	Group of Prof. Dr. Dieter Saur/PD Dr. Günter Schneider
8305 PrimaryPancreaticTumour	Group of Prof. Dr. Dieter Saur/PD Dr. Günter Schneider
8349 PrimaryPancreaticTumour	Group of Prof. Dr. Dieter Saur/PD Dr. Günter Schneider
8442 LungMet1	Group of Prof. Dr. Dieter Saur/PD Dr. Günter Schneider
8442 PrimaryPancreaticTumour	Group of Prof. Dr. Dieter Saur/PD Dr. Günter Schneider
8513 LiverMet1	Group of Prof. Dr. Dieter Saur/PD Dr. Günter Schneider
8513 PrimaryPancreaticTumour	Group of Prof. Dr. Dieter Saur/PD Dr. Günter Schneider
8570 Blood	Group of Prof. Dr. Dieter Saur/PD Dr. Günter Schneider
8570 LiverMet1	Group of Prof. Dr. Dieter Saur/PD Dr. Günter Schneider
8570 PrimaryPancreaticTumour	Group of Prof. Dr. Dieter Saur/PD Dr. Günter Schneider
8661 LiverMet1	Group of Prof. Dr. Dieter Saur/PD Dr. Günter Schneider
8661 LiverMet2	Group of Prof. Dr. Dieter Saur/PD Dr. Günter Schneider
8661 LungMet1	Group of Prof. Dr. Dieter Saur/PD Dr. Günter Schneider
8661 PrimaryPancreaticTumour	Group of Prof. Dr. Dieter Saur/PD Dr. Günter Schneider
9091 LiverMet1	Group of Prof. Dr. Dieter Saur/PD Dr. Günter Schneider
9091 PrimaryPancreaticTumour	Group of Prof. Dr. Dieter Saur/PD Dr. Günter Schneider
9203 LungMet1	Group of Prof. Dr. Dieter Saur/PD Dr. Günter Schneider
9203 LungMet2	Group of Prof. Dr. Dieter Saur/PD Dr. Günter Schneider
9203 LungMet3	Group of Prof. Dr. Dieter Saur/PD Dr. Günter Schneider
9203 LungMet4	Group of Prof. Dr. Dieter Saur/PD Dr. Günter Schneider
9203 LungMet5	Group of Prof. Dr. Dieter Saur/PD Dr. Günter Schneider
9203 LymphNodeMet1	Group of Prof. Dr. Dieter Saur/PD Dr. Günter Schneider
9203 PrimaryPancreaticTumour	Group of Prof. Dr. Dieter Saur/PD Dr. Günter Schneider
9591 PrimaryPancreaticTumour	Group of Prof. Dr. Dieter Saur/PD Dr. Günter Schneider
AA1116 PrimaryPancreaticTumour	Group of Prof. Dr. Dieter Saur/PD Dr. Günter Schneider
AA1165 PrimaryPancreaticTumour	Group of Prof. Dr. Dieter Saur/PD Dr. Günter Schneider
AA1261 PrimaryPancreaticTumour	Group of Prof. Dr. Dieter Saur/PD Dr. Günter Schneider
AA1370 PrimaryPancreaticTumour	Group of Prof. Dr. Dieter Saur/PD Dr. Günter Schneider
AA1377 PrimaryPancreaticTumour	Group of Prof. Dr. Dieter Saur/PD Dr. Günter Schneider
AA1467 PrimaryPancreaticTumour	Group of Prof. Dr. Dieter Saur/PD Dr. Günter Schneider
AA168 PrimaryPancreaticTumour	Group of Prof. Dr. Dieter Saur/PD Dr. Günter Schneider
AA169 PrimaryPancreaticTumour	Group of Prof. Dr. Dieter Saur/PD Dr. Günter Schneider
AA172 PrimaryPancreaticTumour	Group of Prof. Dr. Dieter Saur/PD Dr. Günter Schneider
AA199 PrimaryPancreaticTumour	Group of Prof. Dr. Dieter Saur/PD Dr. Günter Schneider
AA267 PrimaryPancreaticTumour	Group of Prof. Dr. Dieter Saur/PD Dr. Günter Schneider
AA348 PrimaryPancreaticTumour	Group of Prof. Dr. Dieter Saur/PD Dr. Günter Schneider
AA456 PrimaryPancreaticTumour	Group of Prof. Dr. Dieter Saur/PD Dr. Günter Schneider
AA728 PrimaryPancreaticTumour	Group of Prof. Dr. Dieter Saur/PD Dr. Günter Schneider
AA821 PrimaryPancreaticTumour	Group of Prof. Dr. Dieter Saur/PD Dr. Günter Schneider
AA852 PrimaryPancreaticTumour	Group of Prof. Dr. Dieter Saur/PD Dr. Günter Schneider
AA966 PrimaryPancreaticTumour	Group of Prof. Dr. Dieter Saur/PD Dr. Günter Schneider
B590 PrimaryPancreaticTumour	Group of Prof. Dr. Dieter Saur/PD Dr. Günter Schneider
C065 PrimaryPancreaticTumour	Group of Prof. Dr. Dieter Saur/PD Dr. Günter Schneider
C2532 PrimaryPancreaticTumour	Group of Prof. Dr. Dieter Saur/PD Dr. Günter Schneider
C2810 PrimaryPancreaticTumour	Group of Prof. Dr. Dieter Saur/PD Dr. Günter Schneider
C2922 PrimaryPancreaticTumour	Group of Prof. Dr. Dieter Saur/PD Dr. Günter Schneider

Cell Line	Provided By / Supplier
C3356 PrimaryPancreaticTumour	Group of Prof. Dr. Dieter Saur/PD Dr. Günter Schneider
C3443 PrimaryPancreaticTumour	Group of Prof. Dr. Dieter Saur/PD Dr. Günter Schneider
C4340 PrimaryPancreaticTumour	Group of Prof. Dr. Dieter Saur/PD Dr. Günter Schneider
C4466 PrimaryPancreaticTumour	Group of Prof. Dr. Dieter Saur/PD Dr. Günter Schneider
C4468 PrimaryPancreaticTumour	Group of Prof. Dr. Dieter Saur/PD Dr. Günter Schneider
C4604 PrimaryPancreaticTumour	Group of Prof. Dr. Dieter Saur/PD Dr. Günter Schneider
C4617 PrimaryPancreaticTumour	Group of Prof. Dr. Dieter Saur/PD Dr. Günter Schneider
C4692 PrimaryPancreaticTumour	Group of Prof. Dr. Dieter Saur/PD Dr. Günter Schneider
C5040 PrimaryPancreaticTumour	Group of Prof. Dr. Dieter Saur/PD Dr. Günter Schneider
C5081 PrimaryPancreaticTumour	Group of Prof. Dr. Dieter Saur/PD Dr. Günter Schneider
C5315 PrimaryPancreaticTumour	Group of Prof. Dr. Dieter Saur/PD Dr. Günter Schneider
HEK293FT	ATCC
HUPT3	COSMIC Cell Lines Project
KG471 PrimaryPancreaticTumour	Group of Prof. Dr. Dieter Saur/PD Dr. Günter Schneider
KG513 PrimaryPancreaticTumour	Group of Prof. Dr. Dieter Saur/PD Dr. Günter Schneider
KG526 PrimaryPancreaticTumour	Group of Prof. Dr. Dieter Saur/PD Dr. Günter Schneider
KG537 PrimaryPancreaticTumour	Group of Prof. Dr. Dieter Saur/PD Dr. Günter Schneider
KG564 PrimaryPancreaticTumour	Group of Prof. Dr. Dieter Saur/PD Dr. Günter Schneider
mPDAC021 PrimaryPancreaticTumour	Group of Prof. Dr. Roland Rad
mPDAC901 PrimaryPancreaticTumour	Group of Prof. Dr. Roland Rad
NS418 PrimaryPancreaticTumour	Group of Prof. Dr. Dieter Saur/PD Dr. Günter Schneider
NS490 PrimaryPancreaticTumour	Group of Prof. Dr. Dieter Saur/PD Dr. Günter Schneider
PANC0327	COSMIC Cell Lines Project
PPAA_4_1h PrimaryPancreaticTumour	Group of Prof. Dr. Roland Rad
PPAB_11_5c PrimaryPancreaticTumour	Group of Prof. Dr. Roland Rad
PPAB_17_2c PrimaryPancreaticTumour	Group of Prof. Dr. Roland Rad
PPAB_24_1f PrimaryPancreaticTumour	Group of Prof. Dr. Roland Rad
PPAB_24_1h PrimaryPancreaticTumour	Group of Prof. Dr. Roland Rad
PPAB_25_4b PrimaryPancreaticTumour	Group of Prof. Dr. Roland Rad
PPAB_25_4g PrimaryPancreaticTumour	Group of Prof. Dr. Roland Rad
PPAB_26_2i PrimaryPancreaticTumour	Group of Prof. Dr. Roland Rad
PPAB_3_10i PrimaryPancreaticTumour	Group of Prof. Dr. Roland Rad
PPAB_3_3g PrimaryPancreaticTumour	Group of Prof. Dr. Roland Rad
PPAB_3_5b PrimaryPancreaticTumour	Group of Prof. Dr. Roland Rad
PPAB_3_5c PrimaryPancreaticTumour	Group of Prof. Dr. Roland Rad
PPAB_4_5c PrimaryPancreaticTumour	Group of Prof. Dr. Roland Rad
PPAB_4_9a PrimaryPancreaticTumour	Group of Prof. Dr. Roland Rad
PPAB_4_9i PrimaryPancreaticTumour	Group of Prof. Dr. Roland Rad
PPAB_9_3k PrimaryPancreaticTumour	Group of Prof. Dr. Roland Rad
PPAB_9_4h PrimaryPancreaticTumour	Group of Prof. Dr. Roland Rad
R1035 PrimaryPancreaticTumour	Group of Prof. Dr. Dieter Saur/PD Dr. Günter Schneider
R254 PrimaryPancreaticTumour	Group of Prof. Dr. Dieter Saur/PD Dr. Günter Schneider
S134 LiverMet1	Group of Prof. Dr. Dieter Saur/PD Dr. Günter Schneider
S134 LiverMet2	Group of Prof. Dr. Dieter Saur/PD Dr. Günter Schneider
S134 LiverMet3	Group of Prof. Dr. Dieter Saur/PD Dr. Günter Schneider
S134 PrimaryPancreaticTumour	Group of Prof. Dr. Dieter Saur/PD Dr. Günter Schneider
S302 PrimaryPancreaticTumour	Group of Prof. Dr. Dieter Saur/PD Dr. Günter Schneider
S411 PrimaryPancreaticTumour	Group of Prof. Dr. Dieter Saur/PD Dr. Günter Schneider
S559 PrimaryPancreaticTumour	Group of Prof. Dr. Dieter Saur/PD Dr. Günter Schneider
S821 PrimaryPancreaticTumour	Group of Prof. Dr. Dieter Saur/PD Dr. Günter Schneider
S914 PrimaryPancreaticTumour	Group of Prof. Dr. Dieter Saur/PD Dr. Günter Schneider
W22 PrimaryPancreaticTumour	Group of Prof. Dr. Dieter Saur/PD Dr. Günter Schneider

2.1.8 Mice

Mouse Strain	Source
Cdkn2a^{tm4Rdp}	In-house mouse facility of Klinikum r. d. Isar, TUM
Gt(ROSA)26Sor$^{tm1(pb)Brd}$	In-house mouse facility of Wellcome Trust Sanger Institute, UK
Kras^{tm1Dsa}	In-house mouse facility of Klinikum r. d. Isar, TUM
Kras^{tm4Tyj}	In-house mouse facility of Klinikum r. d. Isar, TUM
Ptf1a$^{tm1(cre)Hnak}$	In-house mouse facility of Klinikum r. d. Isar, TUM
Tg(Pdx1-cre)6Tuv	In-house mouse facility of Klinikum r. d. Isar, TUM
Tg(Pdx1-flpo)#Dsa	In-house mouse facility of Klinikum r. d. Isar, TUM
Tgfbr2$^{tm1.2Hlm}$	In-house mouse facility of Klinikum r. d. Isar, TUM
TgTn(pb-sb-ATP1)S2Brd	In-house mouse facility of Wellcome Trust Sanger Institute, UK
Trp53$^{tm1.1Dgk}$	In-house mouse facility of Klinikum r. d. Isar, TUM
Trp53^{tm1Brn}	In-house mouse facility of Klinikum r. d. Isar, TUM

2.1.9 Kits

Kit	Supplier
DNeasy Blood & Tissue Kit	Qiagen
MinElute Reaction Cleanup Kit	Qiagen
NucleoBond® Xtra Midi EF	Macherey-Nagel
QIAprep Spin Miniprep Kit	Qiagen
QIAquick Gel Extraction Kit	Qiagen
QIAquick PCR Purification	Qiagen
QIAshredder	Qiagen
Qubit® dsDNA BR Assay Kit	Thermo Fisher Scientific
RNase-free DNase set	Qiagen
RNeasy Mini Kit	Qiagen
ZytoLight® SPEC KRAS/CEN12 Dual Color Probe	ZytoVision

2.1.10 Databases

Databases	Source
Cancer Cell Line Encyclopedia	https://portals.broadinstitute.org/ccle
cBioPortal	http://www.cbioportal.org/
DAVID	https://david.ncifcrf.gov/
Ensembl	http://www.ensembl.org/index.html
Gene Expression Omnibus	https://www.ncbi.nlm.nih.gov/geo/
International Cancer Genome Consortium	https://icgc.org/
Molecular Signatures Database	http://software.broadinstitute.org/gsea/msigdb/annotate.jsp
PubMed	https://www.ncbi.nlm.nih.gov/pubmed

2.1.11 Software

Software	Company
Agilent Genomic Workbench v7.0	Agilent Technologies
AxioVision v4.8	Carl Zeiss
EndNote X7	Thomson Reuters
GraphPad Prism v5.01	GraphPad Software
Illumina Sequence Analysis Viewer v2.4.5	Illumina
ImageJ (v1.50i)	Open source
Microsoft Office 365	Microsoft
R v3.2.2	The R Project, The R Foundation
Snapgene 3.1	GSL Biotech
StepOne™ v2.3	Applied Biosystems

2.1.12 Technical equipment

Equipment	Company
Agilent Bioanalyzer 2100	Agilent Technologies
Analytical balance A 120 S	Sartorius
Axio Imager.A1	Carl Zeiss
AxioCam ICc5	Carl Zeiss
AxioCam MRc	Carl Zeiss
Centrifuge 5424	Eppendorf
Centrifuge 5810 R	Eppendorf
Class II Biological Safety Cabinet	Thermo Fisher Scientific
CO_2-incubator Heracell™ VIOS 250i	Thermo Fisher Scientific
DynaMag™-96 Side Skirted Magnet	Thermo Fisher Scientific
Electrophoresis power supply Power Pac 200	Bio-Rad Laboratories
Heated paraffin embedding module EG1150 H	Leica Microsystems
HiSeq 1500 System	Illumina
Homogenisator Precellys® 24	Bertin Instruments
Horizontal gel electrophoresis system	Biozym Scientific
Incubator NCU-Line® IL 23	VWR International
Leica TCS SP8; DMi8 CS	Leica Microsystems
M220 Focused-ultrasonicator	Covaris
Magnetic stirrer D-6010	neoLab Migge
Microwave	Thermo Fisher Scientific
MiSeq System	Illumina
Neubauer hemocytometer, improved	LO-Laboroptik
NextSeq 550 System	Illumina
pH meter 521	WTW
Pipettes Reference®, Research®	Eppendorf
Primovert Microscope	Carl Zeiss
Qubit® 2.0 Fluorometer	Thermo Fisher Scientific
StepOne Plus Real-Time PCR System	Applied Biosystems
Thermocycler Tpersonal 48	Biometra
Thermocycler TProfessional Basic 96	Biometra
Thermocycler TProfessional Basic Gradient 96	Biometra
ThermoMixer® comfort 5355	Eppendorf
Tissue processor ASP300	Leica Microsystems
Ultra Low-Temperature Freezer Innova® U725	Eppendorf
UVsolo 2 Gel Documentation System	Analytik Jena
Vortex-Genie 2	Scientific Industries
Weighing Scale A120S	Sartorius

2.1.13 Manufacturers

Company	Location of Headquarter
Abbott	Ludwigshafen, Germany
Addgene	Cambridge, Massachusetts, USA
Agilent Technologies	Santa Clara, CA, USA
Analytik Jena	Jena, Germany
Applied Biosystems	Carlsbad, CA, USA
ATCC	Manassas, VA, USA
B. Braun Melsungen	Melsungen, Germany
Beckman Coulter	Munich, Germany
Bertin Instruments	Montigny-le-Bretonneux, France
Biochrom	Berlin, Germany
Biometra	Göttingen, Germany

Company	Location of Headquarter
Bio-Rad Laboratories	Hercules, CA, USA
Biozym Scientific	Hessisch Oldendorf, Germany
Brand	Wertheim, Germany
Carl Roth	Karlsruhe, Germany
Carl Zeiss	Oberkochen, Germany
Corning	Corning, NY, USA
Covaris	Woburn, MA, USA
Eppendorf	Hamburg, Germany
Eurofins Genomics	Ebersberg, Germany
Fermentas	St. Leon-Rot, Germany
Gerhard Menzel B.V.	Braunschweig, Germany
Gibco	Carlsbad, CA, USA
GraphPad Software	San Diego, CA, USA
Greiner Bio-One	Kremsmünster, Austria
GSL Biotech	Chicago, IL, USA
Illumina	San Diego, CA, USA
Integra Biosciences	Biebertal, Germany
Kapa Biosystems	Wilmington, MA, USA
Leica Microsystems	Wetzlar, Germany
LO-Laboroptik	Friedrichsdorf, Germany
Macherey-Nagel	Düren, Germany
Merck	Darmstadt, Germany
Microsoft	Redmond, Washington, USA
Mirus Bio LCC	Madison, WI, USA
neoLab Migge	Heidelberg, Germany
New England Biolabs	Ipswich, MA, USA
Qiagen	Hilden, Germany
Sarstedt	Nümbrecht, Germany
Sartorius	Göttingen, Germany
Scientific Industries	Bohemia, NY, USA
Seidel Medipool	Gauting-Buchendor, Germany
Sigma-Aldrich	St. Louis, MO, USA
Simport Scientific	Beloeil, QC, Canada
The R Project, The R Foundation	Vienna, Austria
Thermo Fisher Scientific	Waltham, MA, USA
Thomson Reuters	Carlsbad, CA, USA
Viagen Biotech	Los Angeles, CA, USA
VWR International	Darmstadt, Germany
Waldeck	Münster, Germany
Worthington Biochemical	Lakewood, NJ, USA
WTW	Weilheim, Germany
ZytoVision	Bremerhaven, Germany

2.2 Methods

2.2.1 Isolation of primary mPDAC cell cultures

Cell cultures of primary pancreatic tumours and metastases from mice were in large parts generated in the laboratory of Prof. Dr. med. Dieter Saur and PD Dr. med. Günter Schneider at the University Hospital, TU Munich. Primary pancreatic tumours or metastasis of the liver, lung or lymph node were cut into small pieces and digested for 18-24 h in DMEM (Sigma-Aldrich) containing 10% fetal calf serum (FCS, Biochrom), 1x Penicillin/Streptomycin (P/S, Thermo Fisher Scientific Inc.) and 200 U/ml of collagenase type II (Worthington Biochemical). Cells were seeded in 6 well (Greiner Bio-One) and expanded to T175 (Greiner Bio-One) for freezing. Cell cultures were passaged as long as all contaminating stromal cells/fibroblasts were removed from the primary cancer cell culture. Cells were frozen in freezing medium containing 40% DMEM, 50% FCS and 10% dimethyl sulfoxide (Roth) as early as possible to allow for analysis of the cells with passage lower than 10. Re-genotyping was performed of all cell cultures to (i) validate the correct genotype of the cell cultures and (ii) to test for contaminating stromal cells as indicated by the presence of an un-recombined PCR band in the genotyping PCR of the *Kras*^{G12D}-allele. Pancreatic cancer cell cultures were routinely tested by PCR for infection with mycoplasma (see below). Primary cells were cultured in DMEM containing 10% FCS and 1x P/S for expansion of the cells as well as for all experiments that were performed in this work.

2.2.2 Mycoplasma PCR

For testing mycoplasma contamination, cells were cultured for 3 days to 100% confluency in DMEM supplemented with 10% FCS without P/S. 1 ml of cell culture supernatant was harvested in a 1.5 ml Eppendorf tube and stored at -20°C. 5.5 µl of 2x KAPA Genotyping Mix (Sigma-Aldrich), 0.22 µl of primer mix (containing 5 µM of each primer; 7x forward plus 3x reverse primer) and 4.28 µl ddH_2O were prepared for each PCR reaction resulting in a total volume of 10 µl MasterMix per sample. Harvested cell culture supernatant was defrosted on ice and 1 µl of added to the MasterMix. PCR reaction was performed with the following conditions:

Temperature	Time	Cycles
98 °C	180 sec	1x
65 °C	60 sec	1x
72 °C	60 sec	1x
98 °C	15 sec	35x
65 °C	30 sec	35x
72 °C	20 sec	35x
72 °C	300 sec	1x
16 °C	Pause	---

PCR reaction was analysed on a 1.5% agarose gel (presence of a PCR band at 500 bp indicates mycoplasma contamination).

2.2.3 Isolation of genomic DNA and RNA from primary cell cultures

Genomic DNA (gDNA) was isolated using the DNeasy Blood & Tissue Kit (Qiagen) from murine primary cell cultures. Briefly, cells were cultured to 60-80% confluency in a T175 cell culture flask (Greiner Bio-One), trypsinized, split to two 1.5 ml tubes (Sarstedt). After centrifugation, the supernatant was removed from the pellet. Later on, pellets were used for DNA isolation according to manufacturer's instructions of the DNeasy Blood & Tissue Kit.

For RNA isolation, primary cell cultures were grown on a 10 cm dish (Greiner Bio-One) in DMEM supplemented with 10% FCS, but not containing P/S, to 60-80% confluency. Medium was removed from cultured cells and cells were washed twice with ice-cold DPBS (Thermo Fisher Scientific). Any excess of DPBS was removed and 500 µl of RLT buffer (Qiagen) supplemented with 1:100 ß-mercaptoethanol (Sigma-Aldrich) were pipetted per 10 cm dish. Cell solution was immediately transferred to a QiaShredder column (Qiagen) and centrifuged for 2min with full speed at RT. Homogenized cells were stored for up to one year at -80°C. RNA was isolated with the RNeasy Mini kit (Qiagen) according to manufacturer's instructions.

2.2.4 Histological characterization of mouse PDAC and micro-metastases screening

2 µm thick sections were cut from formalin-fixed paraffin embedded (FFPE) material and H&E stained at the pathological core facility. Analysis of tumour histology was performed by two independent veterinary pathologists experienced in comparative PDAC pathology at the Institute of Pathology, TU Munich. Histopathologic grading was performed according to the most recent consensus report of genetically engineered mouse models (Hruban *et al.*, 2006). Histopathologic examination of the micro-metastatic status was performed by

analysing three H&E-stained liver sections (each section separated by approximately 200µm). Each specimen was screened for the presence of micro-metastatic lesions by a veterinary pathologist.

2.2.5 Animal experiments

Pancreas-specific Cre driver lines (Hingorani *et al.*, 2003; Nakhai *et al.*, 2007; Schonhuber *et al.*, 2014) were crossed with *Kras*$^{G12D\text{-}Panc}$ mice (Jackson *et al.*, 2001; Schonhuber *et al.*, 2014) only, or in combination with *Ink4a/Arf*$^{\Delta HOM\text{-}Panc}$ (Aguirre *et al.*, 2003), *Trp53*$^{\Delta HOM\text{-}Panc}$ (Jonkers *et al.*, 2001; Lee *et al.*, 2012), *Tgfbr2*$^{\Delta HET\text{-}Panc}$ or *Tgfbr2*$^{\Delta HOM\text{-}Panc}$ mice (Chytil *et al.*, 2002) for the generation of PK, PKC, PKP or PKT cohorts, respectively. All animals with signs of sickness were sacrificed in compliance with the European guidelines for the care and use of laboratory animals. Prism (GraphPad Software v5.01) was used for the generation of Kaplan-Meier survival curves. All mice were maintained on C57Bl/6;129S6/SvEv mixed background and housed under specific-pathogen-free conditions. Female and male mice were randomly submitted to tumour/mouse model cohorts. For necropsy of tumour-bearing mice, the abdominal cavity was macroscopically checked for pancreatic cancer and for metastases at the main metastatic routes (liver, lung, lymph nodes). Animal studies were approved by the Institutional Animal Care and Use Committees (IACUC) of Technische Universität München (Regierung von Oberbayern, Munich, Germany).

2.2.6 Genotyping

Ear punches of 3 weeks old pups were used for genotyping. Ear punches were lysed in DirectPCR Lysis Reagent (Mouse Tail) (Viagen Biotech) supplemented with proteinase K (Sigma-Aldrich) (20 µg/ml) at 55°C overnight. Proteinase K was inactivated at 95°C for 15 min. Ear punches were diluted one in five for genotyping and 1 µl of the diluted lysis solution was added to 4 µl of H2O (including mouse allele-specific genotyping primers) and 5 µl of 2x genotyping master mix. The 2x concentrated genotyping mix was prepared from peqGOLD Taq-DNA-Polymerase kit (VWR), supplemented with dNTPs (Fermentas GmbH), sucrose (Sigma-Aldrich) and SucRot buffer, was used for genotyping. The identical PCR program was used for genotyping:

Temperature	Time	Cycles
95 °C	180 sec	1x
95 °C	45 sec	40x
XX °C	60 sec	40x
72 °C	90 sec	40x
72 °C	300 sec	1x
16 °C	Pause	---

Of note, the annealing temperatures were allele-specific for each genotyping PCR:

Mouse strain	Annealing temperature
Ptf1atm1(cre)Hnak	60°C
Tg(Pdx1-cre)6Tuv	64°C
Tg(Pdx1-flpo)#Dsa	56°C
Krastm4Tyj	55°C
Krastm1Dsa	55°C
Cdkn2atm4Rdp	58°C
Trp53tm1Brn	64°C
Trp53tm1.1Dgk	57°C
Tgfbr2tm1.2Hlm	59°C

2.2.7 Amplicon-based deep sequencing at the *Kras* locus or of *Kras* mRNA

For amplicon-based deep sequencing of the *Kras* locus, either fifty ng of high-quality genomic DNA (gDNA) or 1 µl of reversely transcribed mRNA (cDNA) was used. Q5® High-Fidelity DNA Polymerase (New England Biolabs) and primers with Nextera adapter overhangs were used for amplification of the *Kras* locus using the following protocol:

Temperature	Time	Cycles
98 °C	30 sec	1x
98 °C	10 sec	40x
60 °C	20 sec	40x
72 °C	15 sec	40x
72 °C	120 sec	1x
16 °C	Pause	---

PCR products were purified by solid phase reversible immobilization. For this, 0.8x volume (20 µl) of Agencourt® AMPure® XP beads (Beckman Coulter GmbH) were added to the PCR reaction (25 µl). Cleanup was performed according to manufacturer's instructions. After cleanup of the first PCR reaction, Nextera index primers (Illumina) were added to the PCR amplicon in a second Q5® PCR step (15 cycles) for barcoding of up to 96 samples using the following PCR protocol:

Temperature	Time	Cycles
98 °C	30 sec	1x
98 °C	10 sec	15x
65 °C	30 sec	15x
72 °C	60 sec	15x
72 °C	120 sec	1x
16 °C	Pause	---

PCR products were cleaned up using the Agencourt® AMPure® XP kit (Beckman Coulter GmbH) and applied on a 1.5% agarose gel. The relative quantity of each PCR product was estimated from the agarose gel by using ImageJ (v1.50i). Equal amounts of samples were pooled according to the agarose quantification/normalization. The pooled library was quantified by KAPA SYBR® Fast qPCR ABI Mix and the KAPA Library Quantification Kit (Kapa Biosystems) by using the following qPCR protocol:

Temperature	Time	Cycles
95 °C	300 sec	1x
95 °C	15 sec	35x
60 °C	45 sec	35x

The pooled library was adjusted to a final concentration of 4 nM. Twenty to fifty percent of PhiX DNA was spiked into the pooled library depending on the heterogeneity of amplicons present in the final library (e.g. spike in of 50% PhiX if only one type of PCR amplicon is present in the pooled library). The spiked library was denatured and further diluted for sequencing according to manufacturer's instructions. Sequencing of the final library was performed in 300bp paired end mode on a MiSeq system (Illumina). Sequencing raw reads were analysed by Thomas Engleitner, a bioinformatician in our group. Sequencing raw reads were mapped to the mouse reference genome (Ensemble release GRCm38p4, Genome Reference Consortium). For the calculation of *Kras*^{G12D} over *Kras*^{WT} ratios, variant allele calling was performed at the *Kras* locus on chr6 at position 145246771.

2.2.8 *KRAS*^{G12} status analysis in micro-dissected human PanINs

Nineteen patients were included into hPanIN lesion analysis, according to approval by the Ethics Committee of the Faculty of Medicine of the Technische Universität München. FFPE material was cut in 10 µm thick serial sections and air-dried overnight. Short incubation with xylene (Carl Roth) was used for removing paraffin of embedded tissues. The first section was subjected to H&E staining and used for screening and grading of PanIN lesions. Another three sections were stained with hematoxylin (Merck) and kept wet for the micro-dissection procedure. Micro-dissection of human PanINs and control samples was performed by using a 20 gauche cannula (Seidel Medipool) and an Axio Imager microscope (Zeiss). Micro-dissection was documented by taking pictures pre- and post-sampling. The photo documentation was also used for the re-identification of each micro-dissected specimen on the corresponding H&E-stained section. Histopathological grading and micro-dissection of PanIN lesions was performed by an experienced pathologist at the Institute of Pathology, TU Munich. Micro-dissected samples were collected in 1.5 ml low-bind Eppendorf tubes. Genomic DNA (gDNA) was extracted with MinElute spin columns

(Qiagen) for higher sample concentration. gDNA isolation was performed according to manufacturer's instructions. Five µl of gDNA isolated from each hPanIN lesion or normal tissue reference control were used for amplicon-based deep sequencing of *KRAS* exon-2 hotspot mutations. The *KRAS* exon-2 was amplified by a two-step/nested PCR protocol. First, outer primers (hKras gDNA amplicon seq_F1 and R1) were used to amplify *KRAS* exon-2 using Q5® High-Fidelity DNA Polymerase (New England Biolabs) with the following PCR protocol:

Temperature	Time	Cycles
98 °C	30 sec	1x
98 °C	10 sec	50x
60 °C	20 sec	50x
72 °C	30 sec	50x
72 °C	120 sec	1x
4 °C	Pause	---

PCR reactions were cleaned up by adding 0.8x volume (20 µl) of Agencourt® AMPure® XP beads (Beckman Coulter GmbH) to each PCR reaction (25 µl). Cleanup was performed according to manufacturer's instructions. Next, 15 µl of purified PCR product were subjected to a second round of PCR amplification using an inner pair of primers (hKras gDNA amplicon seq_F2 and R2). Q5® High-Fidelity DNA Polymerase (New England Biolabs) and the following PCR protocol were used for amplification:

Temperature	Time	Cycles
98 °C	30 sec	1x
98 °C	10 sec	25x
60 °C	20 sec	25x
72 °C	15 sec	25x
72 °C	120 sec	1x
4 °C	Pause	---

Clean up of PCR reaction was performed by adding 1.0x volume (25 µl) of Agencourt® AMPure® XP beads (Beckman Coulter GmbH) to each PCR reaction (25 µl).

As described in the section above, sequencing adapters and barcodes were added to each individual amplicon by PCR with Illumina Nextera primers. PCR steps, library quantification and sequencing were also performed as described above. Sequencing raw reads were mapped to the human reference genome (GRCh38.p10) by Thomas Engleitner, a bioinformatician in our group. For the calculation of $KRAS^{G12}$ ratios, variant allele calling was performed for $KRAS^{G12}$ hotspot mutations (positions 25398284 and 25398285 on chr12).

2.2.9 Whole genome sequencing (WGS)

One µg of genomic DNA (gDNA) from mPDAC tumour cell line S821 as well as from the corresponding tail control was subjected to shearing using a M220 focused ultrasonicator (Covaris). High-molecular gDNA was sheared to a fragment size of approximately 500bp. 500ng of sheared gDNA was subjected to library preparation using the NEBNext® Ultra™ II DNA Library Prep Kit (New England Biolabs). Samples were prepared according to manufacturer's instructions. After library preparation, the final concentration of each sample was quantified using the Kapa Biosystems library quantification kit (Kapa Biosystems) as described above. Prepared DNA from primary mPDAC S821 and the corresponding tail control were pooled in equimolar amounts, denatured and diluted to a final concentration of 1.8 pM. Whole genome sequencing was performed on a NextSeq 550 machine (Illumina) in 300 bp paired end mode which resulted in a ~20x coverage of tumour and tail sample, respectively. Next-Generation-Sequencing raw data was analysed in large parts by Thomas Engleitner, a bioinformatician in our group. bcl2fastq software v2.18.0.12 (Illumina) was used to convert sequencing raw data to the fastq format. Next, trimming of raw reads was performed with Trimmomatic v0.3648 to preserve an average base Phred quality of 25. Processed reads were mapped to GRCm38.p5 reference mouse genome using the BWA-MEM algorithm v0.7.1249 with alternative contig handling.

2.2.10 Inference of chromothripsis

For the inference of chromothripsis in cancer genomes, WGS data from primary mPDAC S821 was analysed by Thomas Engleitner and Maximilian Zwiebel, two bioinformaticians in our group, according to the criteria proposed previously by Korbel *et al.* (Korbel *et al.*, 2013). Bioinformatic analysis strategies were developed by Thomas Engleitner, Maximilian Zwiebel, Roman Maresch and myself. Copy number states were estimated on the basis of WGS data by using the Bioconductor HMMcopy package 1.16.0 and the Bioconductor DNAcopy package 1.48.0. For the analysis of LOH patterns, a catalogue of SNPs was generated by computing variant positions in tumour versus control using samtools mpileup v1.3.150. SNPs: (i) in regions with mapping quality below 60, (ii) in regions with an average phredscore lower than 20, (iii) in regions with strand bias and (iv) with variant allele frequency below 20% or higher than 85% in the tail control sample (likely homozygous germline SNPs) were excluded from LOH analyses. Furthermore, the minimal read coverage support for each individual SNP was set to eight reads. To account for SNPs that are the result of reads falsely mapped to genomic regions with large sequence homology, SNPs in genomic regions with mouse line specific variations (Mouse Genomes Project, REL-1505) or segmental duplications (UCSC Genome Browser) were

also excluded from LOH analyses. For the remaining set of SNPs the variant allele frequency was calculated in the mPDAC S821 sample and subjected to LOH analysis. For the analysis of structural variations (SVs), DELLY v0.7.651 was used. The DELLY output also contained the classification of SVs according to the mapping orientation of the paired-end reads relative to the reference genome: (i) deletion-type (3'to5'), (ii) duplication-type (5'to3') and (iii) inversion-type (5'to5' and 3'to3'). SVs predicted by DELLY were filtered for further downstream analyses based on mapping quality, variant frequency and distance between forward and reverse read of mapped SVs. Finally, the filtered SNP, CNV and SV data was used for testing the six criteria for inference of chromothripsis as proposed by Korbel *et al.* (Korbel *et al.*, 2013):

(i) Clustering of SV breakpoints was analysed with a χ^2-goodness-of-fit test.

(ii) Regularity of oscillating copy number states was tested by comparing the quantity of observed distinct copy number states on chromosome 4 versus the quantity of distinct copy number states that would be expected in theory from a progressively rearranged chromosome. In the theoretical progressive model, SVs are acquired sequentially one after another. For this, a virtual chromosome was generated by Monte Carlo simulation with sequential introduction of SVs as described by Stephens *et al.* (Stephens *et al.*, 2011). 100 chromosome simulation runs were computed for each distinct number of SV breakpoints. Mean values of distinct copy number states for each simulated chromosome as well as 95% CI for each distinct number of SV breakpoints were calculated.

(iii) For the analysis of interspersed loss and retention of heterozygosity, the Jaccard index was calculated between heterozygously deleted segments and regions comprising LOH and SNP information.

(iv) Randomness of observed DNA segment order was analysed by using Monte Carlo simulation as described by Korbel *et al.* (Korbel *et al.*, 2013). A χ^2-goodness-of-fit test was applied to test for the uniform distribution of all four possible SV-types.

(v) The alternating 5'/3' sequence of mapped paired end reads was tested by applying the right-sided Wald-Wolfowitz runs test against the null hypothesis of randomly distributed 5'-to-3' breakpoint joints as implemented in R package randtests 1.0.

(vi) The prevalence of SVs affecting a specific haplotype was tested by analysing the karyotypes on the basis of M-FISH data from mPDAC S821 instead of WGS data.

2.2.11 FISH analyses

Metaphase preparations of pancreatic cancer cell lines were generated by adding 10 µl of Karyomax Colcemid (Thermo Fisher Scientific) to exponentially growing adherent cells for 45 min at 37°C (total of 10 ml culturing medium). Cells were then washed once with DPBS (Thermo Fisher Scientific) and detached by using trypsin (Thermo Fisher Scientific). Next, cells were spun down for 5min at 350g, RT and the cell pellet was resuspended in 4 ml of hypotonic 0.56% KCl solution for 12min at RT. 500µL of freshly prepared fixative (3 units methanol (Roth) plus 1 unit acetic acid (AppliChem) were gently added to the cell suspension and centrifuged for 5min at 350g, RT. The cell pellet was resuspended in 3 ml of fixative and centrifuged for 5 min at 350g, RT. The whole process was repeated once, and the cell final pellet was resuspended in 250 µl fixative for long term storage of the metaphase preparation at -20°C. Later, already prepared metaphase preparations were sent to the team of Dr. Fengtang Yang at the Cytogenetics Core Facility of the Wellcome Sanger Institute, Cambridge, UK for multicolor fluorescence in situ hybridization (M-FISH). M-FISH was performed according to the protocol developed by Jentsch *et al*. (Jentsch *et al.*, 2001). M-FISH data was used for the analysis of whole chromosome or large structural alterations in mouse or human PDAC cell cultures.

Interphase fluorescence in situ hybridization (Interphase-FISH) for *KRAS* gene copy detection in human PanIN sections was performed by the team of Dr. Katja Steiger, Institute of Pathology of the TU Munich, Germany using the ZytoLight® SPEC KRAS/CEN12 Dual Color Probe kit according to manufacturer's instructions (ZytoVision GmbH). Birefly, 2 µm thick FFPE sections were deparaffinized, pre-incubated in CC 2 buffer (at 95°C for 24 min), treated with pepsin solution (at 37°C for 8 min) and denatured by a heat treatment step at 80°C for 8 min on an automated Discovery XT system (Ventana Medical Systems). Co-denaturing at 75°C for 10 min and by incubating at 37°C overnight was performed for hybridization of KRAS/CEN12 dual colour probes in a ThermoBrite system (Abbott Laboratories). Slides were washed to remove un-hybridized probes. Then, nuclei were stained with 4',6-diamidino-2-phenylindole (DAPI) and covered by an antifade mounting medium. Finally, coverslips were applied on top of the mounting medium and slides were stored at 4°C in the dark for confocal laser scanning microscopy (LSM) analyses. A Leica TCS SP8; DMi8 CS microscope equipped with a 63x/1.4 oil immersion objective (Leica Microsystems GmbH) was used for the detection of ZytoLight® SPEC KRAS/CEN12 Dual Color Probes and DAPI nuclear staining with the following excitation/emission settings: DAPI 405 nm/415-490 nm; ZyGreen, 503 nm/510-540 nm; ZyOrange, 547 nm/560-650 nm. Z-stacked images with a magnification factor of 3 and a frame size of 2048 x 2048 pixels were collected to cover the whole nucleus of a human

PanIN cell. Huygens Essential software (Scientific Volume Imaging) was used for processing and deconvolution of generated images. Leica LAS X software was then used for merging of images and converting maximum projections.

2.2.12 aCGH analysis

For the analysis of copy number profiles either a SurePrint G3 Mouse CGH 240K or a SurePrint G3 Custom CGH 60K microarray (Agilent Technologies) was used. Labelling and hybridization of tumour and control DNA was performed according to manufacturer's instructions by the Genetics Core Facility of the Wellcome Sanger Institute, Cambridge, UK or by the team of Dr. Kristian Unger at the Research Unit Radiation Cytogenetics, Helmholtz Centre Munich, Germany. Pre-processing of aCGH data and manual curation of copy number segments was performed by using Agilent Genomic Workbench (AGWB, Agilent) software v7.0.4.0. Raw log ratios were re-centralized to the most common ploidy state by using the legacy centralization option of AGWB. Copy number variations were called by applying the ADM-2 algorithm with threshold setting "5". Curated data was imported into R for CNV analysis and overlay of copy number segments for the identification of recurrently altered genomic regions.

2.2.13 Whole-exome sequencing (WES) analysis in mouse PDAC

SureSelect Mouse All Exon Kit (Agilent) was used for pull down of coding exons according to manufacturer's instructions and conducted by the Genetics Core Facility of the Wellcome Sanger Institute, Cambridge, UK. Whole exome sequencing of the final libraries was performed on a HiSeq 2000 system (Illumina). Analysis of Next-Generation-Sequencing raw read data was conducted by Thomas Engleitner, a bioinformatician in our group. Raw sequencing data was trimmed first using Trimmomatic v0.3348. Leading and trailing bases with phred scores lower than 25 and read lengths shorter than 50 nucleotides were removed. In addition, reads with an average base quality lower than 25 in a sliding window of 10 nucleotides were excluded from further downstream analysis. BWA-MEM 0.7.1249 with default settings was used for mapping of processed reads to the mouse GRCm38.p3 reference genome. Reads resulted from PCR duplicates were flagged with Picard tools v1.130 and excluded from variant calling. Reads containing indels were realigned to improve mapping and indel detection using GATK toolkit v3.4.46. Somatic nucleotide variations (SNVs) were called by comparing the tumour to the corresponding tail sample using Mutect v1.1.755 software. In addition to filtering out SNVs that were detected in the corresponding tail, SNVs were further filtered for germ line variants listed

in release 1505 of the Mouse Genome Project SNP database (Keane *et al.*, 2011). SNVs were filtered for: (i) variant allele frequency of at least 10% (ii) read coverage of at least 10 in both the control and the tumour sample and (iii) read count support of the variant nucleotide of at least 3 reads in the tumour sample and the absence of any read supporting the variant in the control sample (read count support = 0). "DKFZBiasFilter" tool (https://github.com/eilslabs/DKFZBiasFilter, using default settings) was used to filter out SNVs with oxidation artefacts introduced during library preparation (FOXOG-Score = 1), strand bias or PCR bias. SNVs were annotated using SNPeff v4.157. The final catalogue of SNVs was generated by removing SNVs affecting splice sites or located upstream/downstream of a gene, all not directly effecting the translated exonic regions. Detection of indels was performed by using Pindel tool (Ye *et al.*, 2009). Read coverage was re-calculated for indels by using bedtools v2.17.059. Indels were than filtered for: (i) total coverage of indel positions in both tumour and control of at least 20 reads each and (ii) variant allele frequency of the indel of at least 10%. Analyis of LOH patterns was performed as described in section “whole-genome sequencing (WGS)”.

2.2.14 WES data analysis from human PDAC

The download and storage of mapped BAM files from the study of Witkiewicz *et al.* (Sequence Read Archive accession number PRJNA278883) was approved by the Ethics Committee of the Faculty of Medicine of the Technische Universität München. BAM files were processed with similar settings as described in section “Whole-exome sequencing (WES) analysis in mouse PDAC”. Importantly, SNVs of each human PDAC were filtered against germ line variants reported with a prevalence of at least 1% in the 1000 Genomes Project (dbSNP build 146) (instead of using the mouse SNP database).

For the pancreatic cancer cohort of the cancer genome atlas consortium (TCGA-PPAD), variant call format (VCF) files generated by Mutect2 were downloaded from the NIH Genomic Data Commons portal. VCF files were filtered as described above. Of note, filtering of variant positions with strand bias as identified by DKFZBiasFilter tool, was not possible on the basis of VCF files. Again, SNVs were filtered against germ line variants reported with a prevalence of at least 1% in the 1000 Genomes Project.

Furthermore, the mutation annotation format (MAF) files were downloaded from additional published sequencing studies of human pancreatic cancer cohorts; (i) Bailey *et al.* (Bailey *et al.*, 2016), (ii) pancreatic cancer cell lines from the Cancer Cell Line Encyclopedia (Barretina *et al.*, 2012) and (iii) SNV data from pancreatic cancers as analysed by Alexandrov *et al.* (Alexandrov *et al.*, 2013a). Potential SNPs were removed from SNVs by

filtering for germ line variants reported with a prevalence of at least 1% in the 1000 Genomes project. Remaining SNVs were annotated and filtered with SNPeff as described above.

Raw data from all human PDAC cohorts was analysed by Thomas Engleitner, a bioinformatician in our group.

2.2.15 Analysis of mutational signatures

Mutational signatures were analysed with help from Sebastian Lange from our group. A list of 21 mutational signatures previously identified by Alexandrov *et al.* (Alexandrov *et al.*, 2013a) was used as the reference data set of known mutational signature. Due to the high similarity of mutational signature 1A and 1B (on the mutational and biological level), signature 1B was excluded from further downstream analyses. “deconstructSigs” v1.8.064 tool was used with standard settings to decompose the contribution of each mutational signature to the mutational spectra of the individual pancreatic cancer cohorts.

2.2.16 qRT-PCR analysis

First, 1 µg of RNA was reversely transcribed using random hexamers and the SuperScript II protocol (Thermo Fisher Scientific) according to manufacturer’s instructions. Synthesized cDNA was diluted one in ten in H_2O before qPCR. TaqMan real-time qPCR was performed for detection of total (wild-type and mutant) *Kras* mRNA levels in mouse PDAC tumours. For this, TaqMan qPCR chemistry (Thermo Fisher Scientific) was used in combination with *Kras*-specific primers and probes. *Gapdh* was used as housekeeping gene for normalization of *Kras* transcript levels. Quantification of *VIM*, *CDH1* and *MMP1* mRNA levels in human PDAC cell lines was performed by using SYBR™ Select master mix (Thermo Fisher Scientific) and *VIM*-, *CDH1*- or *MMP1*-specific primers. The combination of housekeeping genes *GAPDH* plus *PPIA* was used as for normalization of transcript levels in human PDAC cell lines.

The following reagent mix and PCR program were used for TaqMan real-time qPCR:

Reagent	Volume
qPCR MasterMix	5.0 µl
Primer fwd (10 µM)	0.4 µl
Primer rev (10 µM)	0.4 µl
TaqMan probe (10 µM)	0.2 µl
cDNA (1:10)	2.5 µl
H_2O	ad 10µl

Temperature	Time	Cycles
50 °C	120 sec	1x
95 °C	20 sec	1x
95 °C	1 sec	40x
60 °C	20 sec	40x

The following reagent mix and PCR program were used for SYBR™ real-time qPCR:

Reagent	Volume
qPCR MasterMix	5.0 µl
Primer fwd (10 µM)	0.2 µl
Primer rev (10 µM)	0.2 µl
cDNA (1:10)	2.5 µl
H_2O	ad 10µl

Temperature	Time	Cycles
50 °C	120 sec	1x
95 °C	120 sec	1x
95 °C	15 sec	40x
60 °C	60 sec	40x

The quantification of *Kras*^{G12D}-specific transcript levels was performed in two steps. First, the absolute (wild-type plus mutant) *Kras* mRNA levels were determined by qRT-PCR as described in this section above. Second, amplicon-based deep sequencing of the *Kras* locus was performed on the identical cDNA for the analysis of mutant to wild-type *Kras* mRNA proportions according to the protocol described in section “Amplicon-based deep sequencing at the *Kras* locus or of *Kras* mRNA”. The absolute *Kras* mRNA level (qPCR) was then multiplied with the proportion of mutant *Kras*^{G12D}-specific transcripts (NGS) to calculate the mutant *Kras*^{G12D}-specific mRNA level.

2.2.17 RNA-Sequencing analysis

Libraries for bulk 3’-prime transcript end RNA-Seq (SCRB-Seq) were prepared in cooperation with Christoph Ziegenhain at the Anthropology and Human Genomics Research Unit of the LMU Munich, Germany as described by Parekh *et al.* (Parekh *et al.*, 2016). Briefly, oligo-dT primers containing sample barcodes, unique molecular identifiers (UMIs) and sequencing adapters (Integrated DNA Technologies) were used for reverse transcription of mRNA into cDNA. In the next step, barcoded samples were pooled. Un-incorporated primers were digested with ExonucleaseI (New England Biolabs). KAPA HiFi HotStart ReadyMix (2x) (Kapa Biosystems) was then used for the amplification of the cDNA pool. 0.8 ng of cDNA was subjected to tagmentation using the Nextera XT Kit (Illumina). 3’ ends of tagmented cDNA were amplified by using KAPA HiFi HotStart ReadyMix (2x) (Kapa Biosystems) with primers specific for the Nextera XT adapter and the oligo-dT primer overhang. The final library was subjected to paired-end sequencing on

a HiSeq 1500 machine (Illumina) with the following sequencing settings: Read1 with 16 cylces for sequencing sample barcodes and UMIs plus Read2 with 50 cycles for sequencing the cDNA fragment. Analysis of raw read Next-Generation-Sequencing data was conducted by Thomas Engleitner, a bioinformatician in our group. Sample- and gene-wise UMI-filtered expression tables were generated by processing raw sequencing data with the published Drop-seq pipeline v1.0 (Macosko *et al.*, 2015). STAR 2.5.1b67 with default parameters was used to calculate alignments to the reference genome (GRCm38). ENSEMBL annotation release 75 was used for the definition of genes and transcripts. All further analyses were performed with R version 3.2.2. The top 10% of genes with the highest variability in expression levels across the sequenced mPDAC cohort were used for an initial hierarchical clustering of the samples based on correlation as distance measure. Cluster stability was tested by bootstrapping with pvclust package v2.068. This approach resulted in the identification of 4 distinct clusters in the PK cohort (C2a, C2b, C2c and C1). Differential gene expression between defined clusters was computed with DEseq269. A gene was considered to be differentially expressed between clusters if the absolute log2-foldchange was above 0.8 and the adjusted P-value was ≤0.05. DAVID 6.870 or the "Molecular Signature Database" (MSigDB) v6.0 homepage (http://software.broadinstitute.org/gsea/msigdb/annotate, hypergeomtric test) were used for gene set enrichment analysis of differentially expressed genes. MSigDB output was sorted by false discovery rate (FDR) and filtered for top30 enriched terms with a FDR of $P \leq 10^{-4}$. In addition, the hallmark EMT gene set was downloaded from MSigDB v5.272 and published PDAC classifier genes (Collisson *et al.*, 2011) were used for the cross-species analyses of PDAC transcriptional subtypes.

2.2.18 Human PDAC subtyping

Cross-comparison of human PDAC classification systems proposed by Collisson e*t al.*, Moffitt *et al.* or Bailey *et al.* (Collisson *et al.*, 2011; Moffitt *et al.*, 2015; Bailey *et al.*, 2016) was performed on the basis of normalized RNAseq data from the study of Bailey *et al.* (Bailey *et al.*, 2016). Classifier gene lists proposed by Collisson *et al.* or Moffitt *et al.* (Collisson *et al.*, 2011; Moffitt *et al.*, 2015) were used for unbiased hierarchical clustering (ward method) of transcriptomes from pancreatic tumours from the cohort of Bailey *et al.* (Bailey *et al.*, 2016). Clustering analysis was conducted with help from Thomas Engleitner, a bioinformatician in our group. Only samples with a histological classification of "PDA-Adenosquamous carcinoma" or "Pancreatic Ductal Adenocarcinoma" in the Bailey cohort were used for clustering.

2.2.19 Microarray data analysis

Affymetrix expression array data from the Cancer Cell Line Encyclopedia (CCLE) was downloaded in version 2.17 from the Broad Institute. Tumours with hematopoietic or lymphoid origin were excluded from further downstream analysis since (i) the expression patterns of all liquid tumours was shown to be distinct from the expression patterns of all solid tumours (Barretina *et al.*, 2012) and (ii) the main goal of the study described here was to investigate the expression profiles of pancreatic cancers (solid tumours). RMA74 was used for normalization of microarray-based expression profiles. For genes that were represented by two or more probe sets on the expression microarray, the probe that showed the highest median expression level across the cohort was kept but all other probes of the corresponding gene were filtered out. For testing of TP63ΔN transcriptional network activity, pathway target genes were download from the "Pathway Interaction Database" (PID) (Schaefer *et al.*, 2009) and used for unbiased hierarchical clustering by ward method. For all pancreatic cancer cohorts, gene set enrichment analysis of differentially expressed genes was conducted with DAVID or MSigDB v6.0.

The microarray data sets described in the same section below are Illumina-based and were VST-transformed and subjected to quantile normalization as implemented in lumi (Du *et al.*, 2008). Microarray-based expression profiles of human PDAC cell lines were downloaded from the Gene Expression Omnibus database (accession number GSE17891). Analysis of EMT hallmark gene set and PDAC classifier gene expression was conducted as already described above. Limma (Ritchie *et al.*, 2015) was applied to human wild-type pancreatic tissue and human PDAC cell lines for the identification of genes differentially expressed between both groups. The detection of differentially expressed genes was performed with a 5% alpha level threshold.

Microarray-based expression data of the Australian pancreatic cancer cohort of the International Cancer Genome Consortium (PACA-AU cohort of the ICGC) was downloaded from the repository https://dcc.icgc.org/repositories. The PACA-AU cohort contained different types and classes of pancreatic cancers. Therefore, the data set was filtered for samples with the following properties: (i) transcriptional PDAC subtype information available from Bailey *et al.*, (ii) Bailey HistoSubtypes is "PDA-Adenosquamous carcinoma" or "Pancreatic Ductal Adenocarcinoma" or ICGC WHO Grading is "Undifferentiated carcinoma". The filtering resulted in a data set of 75 human pancreatic cancer samples which were categorized into six subgroups: (i-iv) PDACs of pancreatic progenitor, immunogenic, squamous or aberrantly differentiated endocrine exocrine (ADEX) transcriptional subtypes (v) adenosquamous pancreatic carcinoma and (vi) undifferentiated pancreatic carcinoma. Analysis of variance (ANOVA) was performed to

identify genes that were differentially expressed across the six defined groups. Genes with an adjusted P-value of ≤0.05 were used for hierarchical clustering of pancreatic cancer transcriptomes. Analysis of the gene cluster tree identified five sub-clusters of co-regulated gene expression. Genes from each expression cluster were subjected to gene set enrichment analyses to test for cluster-specific pathway deregulation.

The seventeen primary pancreatic cancer cell cultures from PK-PB mice were already established and published elsewhere (Rad *et al.*, 2015). Microarray-based gene expression profiling was performed by the Genetics Core Facility of the Wellcome Sanger Institute, Cambridge, UK. For this, RNA was extracted from primary cell cultures by using RNeasy Mini kit (Qiagen) according to manufacturer's instructions. High-quality RNA was subjected to gene expression profiling on a MouseWG-6 v2.0 Expression BeadChip (Illumina). Samples were clustered on the basis of the top 5% genes that showed the highest gene expression variability across the cohort. The ward method was used for aggregation of samples for unbiased hierarchical clustering. Limma was used as described already in this section above. A gene was considered to be differentially regulated between two groups if the log2-fold was at least 0.8 and if the adjusted P-value was ≤0.05.

All microarray raw data was processed by Thomas Engleitner, a bioinformatician in our group.

2.2.20 Quantitative transposon insertion site sequencing (QiSeq)

The seventeen primary pancreatic cancer cell cultures from PK-PB mice were already established and published elsewhere (Rad *et al.*, 2015). gDNA was isolated from the primary pancreatic cancer cell lines by using the DNeasy Blood & Tissue Kit (Qiagen) according to manufacturer's instructions. Isolated high-quality gDNA was subjected to genome-wide, quantitative transposon integration site sequencing as described by Friedrich *et al.* (Friedrich *et al.*, 2017). Next-Generation-Sequencing raw data was analysed by Thomas Engleitner, a bioinformatician in our group. Only transposon integration sites that were located in an intragenic region and had a read support ≥20 were used for the calculation of the mutational burden. Transposon integration sites that are supported by at least 20 reads and reside in intragenic regions were counted. The *Cdkn2a/Ncruc* locus was considered for homozygous inactivation if: (i) *Cdkn2a* was the top hit in the transposon insertion list of an individual primary pancreatic cancer and (ii) the read coverage of the *Cdkn2a* transposon insertion was approximately twice as high as the second hit (when the *Cdkn2a* locus was not affected by additional CNA).

2.2.21 Lentiviral transduction of human PDAC cell lines and overexpression of GFP or *KRAS^{G12D}*

The pINDUCER20 vector system (Meerbrey *et al.*, 2011) was used for doxycycline-inducible overexpression of *KRAS^{G12D}* and GFP control in human PDAC cell lines. First, cDNA of GFP or oncogenic *KRAS^{G12D}* (CCDS 8702.1, 35G>A) were cloned into the pINDUCER20 lentiviral vector. Cloned vectors were verified using Sanger sequencing. Lentiviral production was performed in HEK293FT cells. For this, HEK293FT cells were transfected with psPAX and pMD2.G virus packaging plasmids and the pINDUCER20 GFP or *KRAS^{G12D}* vector by using TransIT®-LT1 (Mirus Bio LLC) according to manufacturer's instructions. 48h and 72h post transfection, virus-containing supernatant of HEK293FT cells was pooled. Polyethylene glycol 6000 precipitation (Kutner *et al.*, 2009) was used for virus purification and concentration. Finally, purified virus was shock-frozen in liquid N_2 and stored at -80°C.

For transduction of human PDAC cell lines, $1x10^5$ HUPT3 or PANC0327 hPDAC were incubated with virus in the presence of 1 µg/mL polybrene. Two days post transduction, successfully infected cells were selected with puromycin antibiotic. HUPT3 and PANC0327 human PDAC cell lines were obtained from the COSMIC Cell Lines Project and were routinely tested for mycoplasma contamination by PCR.

Overexpression of *KRAS^{G12D}* or GFP control was induced by culturing cells for 1, 3 or 5 days in P/S-free culturing medium supplemented with 10% FCS and 100 ng/µL doxycycline. RNA isolation, qRT-PCR as well as SCRB-Seq were performed as described in the corresponding sections above. Raw sequencing data was analysed by Thomas Engleitner, a bioinformatician in our group, and mapped to the human reference genome (GRCh 38p10). ENSEMBL annotation release 88 was used for the definition of transcript and gene definitions. DESeq2 was used for the analysis of genes differentially expressed in *KRAS^{G12D}* vs. GFP groups.

2.2.22 Somatic CRISPR/Cas9 gene-editing for tumour clone tracking in mice

Multiplexed pancreatic genome engineering and cancer induction by transfection-based CRISPR-Cas9 delivery in PK mice was performed together with Roman Maresch as described by Maresch *et al.* (Maresch *et al.*, 2016). Primary pancreatic cancer cell cultures were generated as described in section "Isolation of primary mPDAC cell cultures". All isolated mPDAC cell cultures were routinely monitored for the simultaneous presence of cells with epithelial or mesenchymal morphology. For the separation of both phenotypes, cells were incubated with trypsin for different times (Thermo Fisher Scientific). Short-term

incubation with trypsin for 2 to 3 minutes resulted in detachment of cells with mesenchymal morphology, while cells with epithelial morphology remained adherent. Both cell fractions were then grown to 80% confluency. Enrichment of either morphology was repeated for 3 to 6 rounds of differential trypsinization until morphologically pure cell cultures could be established. Targeted amplicon-based deep sequencing of CRISPR/Cas9-targeted loci was performed as described elsewhere (Weber *et al.*, 2015; Maresch *et al.*, 2016) to determine the clonal origin of epithelial and mesenchymal cells within a single mPDAC. The genomic ratio and expression levels of oncogenic $Kras^{G12D}$ as well as analysis of RNA-Seq data for gene set enrichment analysis of differentially expressed gene were carried out as described in the corresponding sections above.

3 Results

Written contents and figures of this chapter have been previously published in the research article "Evolutionary routes and KRAS dosage define pancreatic cancer phenotypes", Mueller, Engleitner, Maresch et al. (Nature. 2018 Feb 1;554(7690):62-68). This study includes the molecular characterization of 135 primary and metastatic pancreatic cancer cell cultures that were previously generated in the laboratory of Prof. Dr. med. Dieter Saur and PD Dr. med. Günter Schneider at the University Hospital, TU Munich. Contributions of authors other than myself to the analysis presented in this work are indicated in the figure legends and the methods section.

Mouse models are a valuable to overcome most limitations possessed by studying human PDAC genomes: (i) cell cultures can be established reliably from the primary tumour or metastatic lesions of PDAC mouse models with different genotypes, (ii) cell cultures derived from PDAC mouse models can be used to study "natural" tumour evolution of end-stage PDAC without the confounding effects of therapy-induced molecular alterations and (iii) mouse models are a versatile tool to study the causal relationships of the genetic alterations observed in human PDAC genomes and their impact on tumour initiation and/or progression.

Several hundred cell lines of murine primary and metastatic tumours with isolated activating have been already collected by the team of Prof. Dr. Dieter Saur, by the group of PD Dr. Günter Schneider and by our group of Prof. Dr. Roland Rad. In a joint effort, my task was to characterize the genomes and transcriptomes of this primary pancreatic cancer cell cultures and their corresponding metastasis with single activating *Kras*^{G12D} mutations or in combination with complete inactivation of *Cdkn2a*, *Tp53* or *Tgfbr2*. As compared to human PDAC genomes, the genomes of mouse PDAC cell cultures appeared to be less complex which allowed to exploit the lower mutational burden and therefore easier interpretability in mice to unravel a key importance of oncogenic gene-dosage variation in PDAC evolution and phenotypic diversification.

3.1 The genetic landscape of mouse and human PDAC

First, 38 primary pancreatic cancer cell cultures derived from mice expressing mutant *Kras*^{G12D} conditionally in the pancreas (PK mice) (Jackson *et al.*, 2001; Schonhuber *et al.*, 2014) were subjected to comprehensive genetic characterization by multiplex FISH (M-FISH), whole-exome sequencing (WES) and array comparative genomic hybridization (aCGH). A previously published whole-exome sequencing study of human PDAC using micro-dissection (Witkiewicz *et al.*, 2015) to reduce/avoid dilution of tumour cells by stromal cells served as a reference data set for cross species comparisons. Importantly, we established a bioinformatics analysis pipeline that allowed for analysing WES data derived from mouse or human PDAC using identical parameter settings. In addition to the primary pancreatic cancer cell cultures, the tail DNA of each corresponding mouse was subjected for WES and served as the reference control to distinguish germline variants (e.g. SNPs) from somatic mutations that occurred during pancreatic cancer evolution. 319 synonymous and 606 non-synonymous somatic mutations (SNVs) were identified in 38 primary mPDACs based on WES (Figure 8a). Surprisingly, recurrently mutated genes were infrequent in mice with only one tumour carrying a homozygous SNV in the DNA-binding domain of *Trp53* and the absence of Tgfβ pathway alterations in any of the mPDACs. Recurrently mutated genes overlapped only for a subset of mPDACs with common insertions sites identified in pancreas-specific transposon screens in PDAC mouse models (Mann *et al.*, 2012; Perez-Mancera *et al.*, 2012; Rad *et al.*, 2015) and/or with genes recurrently mutated in human cancers (Tate *et al.*, 2019) such as *Camta1*, *Csmd1*, *Fndc3a*, *Lrp1b* or *Pten* (Figure 8a). These findings were in line with the observation that the mutational burden was significantly reduced by 3.3 and 1.5 fold for the median number of SNVs and small insertions/deletions (indels) in murine PDAC as compared to human PDAC (Figure 8b). Further, the rates of copy number alterations (CNAs) as well as structural variations were also decreased by 2.8 and 2.6 fold respectively in murine versus human PDAC (Figure 8b).

The different rate of genetic alterations in both species raised the question that different mutational processes might be at play during mouse and human PDAC evolution. Somatic mutations that are found in cancer genomes might be the result of different mutational processes, such as defective DNA mismatch repair, infidelity of DNA replication, enzymatic DNA modification and mutagen exposure. These processes

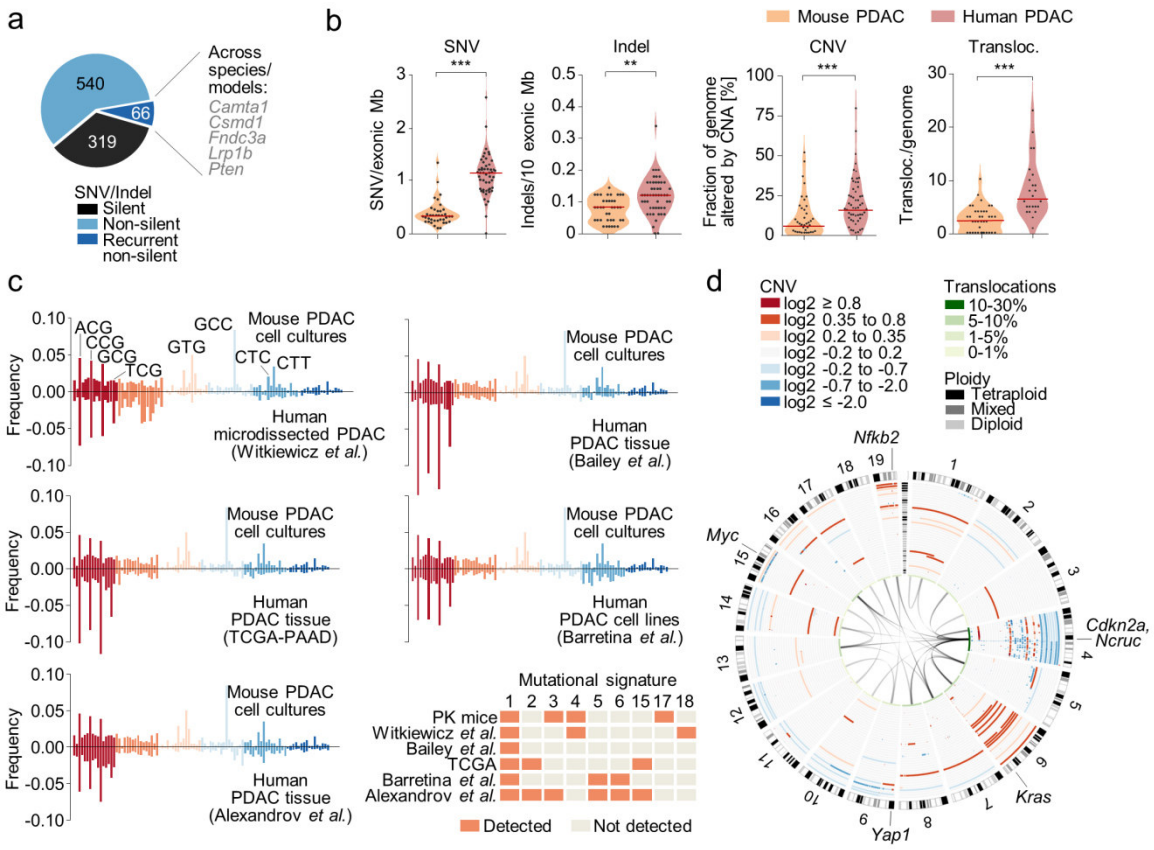

Figure 8 | The genetic landscape of mouse PDAC and cross-species comparison to the human disease.

[a] Single nucleotide variants (SNV) and insertions/deletions (Indels) as identified on the basis of WES in mPDAC derived from PK mice (n=38). Genes that were recurrently mutated and altered in genome-wide transposon screens in the pancreas or human cancers are shown. **[b]** Comparison of the mutational burden in mouse and human PDAC. SNV, Indel and CNA load were calculated on the basis of WES data of PDAC cell cultures from PK mice (n=38) and micro-dissected human PDAC (n=51, data used from Witkiewicz *et al.*, 2015). Translocation load was quantified on the basis of M-FISH karyotypes of mPDAC cell cultures (n=38) and human PDAC cell lines (n=24). **P=0.002, ***P≤0.001, Mann-Whitney test; bars, median. **[c]** Cross-species comparison of the trinucleotide-specific mutational spectrum from primary mouse PDAC cell cultures to different human PDAC cohorts with distinct biological properties. The mutational spectrum of each cohort was decomposed on the basis of 21 mutational signatures that were previously established by Alexandrov *et al.* (Alexandrov *et al.*, 2013a). Only mutational signature 1 was found to contribute to the mutational spectrum of all analysed PDAC cohorts. **[d]** Circos plot showing copy number profiles as detected by aCGH as well as ploidy state and translocations as detected by M-FISH for each individual primary mPDAC cell culture from PK mice (n=38). Mixed ploidy indicates that at least 3 out of 10 karyotypes showed diploidy or tetraploidy. Minimal regions of overlapping CNAs and containing cancer genes are shown. WES raw data was analysed by Thomas Engleitner. Mutational signatures were decomposed by Sebastian Lange. (The whole figure was modified from Mueller *et al.*, 2018).

generate distinct combinations of genetic mutations that can be further refined on the sequence context immediately 3-prime and 5-prime to the mutated base (termed trinucleotide context-specific mutation) ((Alexandrov *et al.*, 2013b). For example when considering the trinucleotide context, a “T to A” mutation can be summarized as ApTpG (“p” indicating the phosphate group between two nucleosides) to ApApG thereby

generating 96 possible trinucleotide-specific mutation types. Based on the frequency distribution of 96 trinucleotide-specific base substitutions, Alexandrov *et al.* identified 21 distinct mutational signatures to be operative across the spectrum of human cancer (Alexandrov *et al.*, 2013a). We used WES data of each primary PDAC cell culture derived from PK mice (n=38) to determine trinucleotide-specific substitutions. The substitutions found in individual tumours were finally pooled to derive a merged mutation spectrum that was compared to human PDAC cohorts, each with distinct biological characteristics such as bulk tissues, micro-dissected tissue or cell lines. The mutational spectra from murine and all human PDAC cohorts showed similar patterns of peaks and frequency distribution (Figure 8c). However, multiple mutational signatures are contributing to the mutation spectrum of each PDAC cohort. The "deconstructSigs" tool was used to decompose the contribution of the 21 mutational signatures to each mutation spectrum. The reliable deconvolution of a mutational spectrum into its mutational signatures is strongly depending on the number of substitutions per sample and requires >50 SNVs for their reliable identification (Rosenthal *et al.*, 2016). The low mutation load of primary mPDAC from PK mice with only a median of 18 SNVs per tumour did not allow for the testing of individual tumours. Therefore, the pooled mutational spectrum of each cohort was used for the deconstruction of mutational signatures. Only Signature 1 was identified across all cohorts of murine and human PDAC which is characterized by C to T transitions in NpCpG trinucleotide context (Figure 8c) and associated with spontaneous deamination of 5-methyl-cytosine during aging (Alexandrov *et al.*, 2013a). All other identified mutational signatures were detected in only two or one of the six PDAC cohorts indicating that these signature might not be uniformly relevant for pancreatic carcinogenesis. C to G substitutions at GCC trinucleotides were specifically observed in mouse but not human PDAC cohorts (Figure 8c). Of note, this peak is not a general phenomenon of all mPDAC since only four mPDAC were predominantly contributing to this peak. Although the mutational load and the frequency of recurrently altered genes is significantly lower in murine as compared to human PDAC, the analysis of mutational spectra and the identification of Signature 1 in all PDAC cohorts indicates that the mutational processes operating during PDAC evolution are similar in both species.

Additional evidence that mutational processes operating during human PDAC formation are reflected in PDAC from PK mice comes from the analyses of CNA patterns. Although PDAC from PK mice had an overall decreased CNA burden, there was however a substantial variation of CNA patterns between individual mPDACs. Some primary mPDACs had almost no CNA with only a few focal alterations, whereas others showed partial/whole chromosome alterations (aneuploidy) and/or extensive clustered

chromosomal rearrangements defined by ten or more alterations per affected chromosome (Figure 8d). One third of primary PDAC from PK mice (14/38) showed signs of such complex rearrangements located on chromosome 4 in 12 cases and invariably involving alteration of the *Cdkn2a* locus. One of the two remaining tumours had a massive rearrangements on chr15 with strong amplification of *Myc* locus while the other tumour showed a clustered rearrangements on chr1 (Figure 9a-n). The predominant association of clustered rearrangements with well-known cancer genes, such as *Cdkn2a* or *Myc*, reflects the selection of these events during tumour evolution. Notably, a specific feature of complex rearrangements in most cancers was the regularity of copy number alterations predominantly oscillating between two states. This observation is highly incompatible with the sequential accumulation of CNAs in a progress scenario over time that would result in complex CNA patterns with more than two states and suggests chromothripsis as the underlying mutational process (Stephens *et al.*, 2011). Chromothripsis is characterized by tens to hundreds of rearrangements which are typically located on one or a few chromosomes and are thought to be acquired in a single genomic catastrophe during tumour evolution (Stephens *et al.*, 2011). More than one cancer-causing lesion can be acquired during chromothripsis. In ~16% of pancreatic cancer, chromothripsis caused simultaneous knockout of canonical preneoplastic cancer drivers that are likely to initiate malignant cancer growth (Notta *et al.*, 2016). These analyses revealed that some pancreatic cancers evolve through a punctuated equilibrium while others progress through the stepwise accumulation of cancer gene alterations. To test whether chromothripsis is also an evolutionary trajectory in PDAC from PK mice, the genome of one PDAC S821 was sequenced. Based on whole-genome sequencing data, Korbel *et al.* proposed six criteria for the statistical inference of chromothripsis in cancer genomes: (i) clustering of breakpoints, (ii) regularity of oscillating copy-number states, (iii) overlapping CNA and loss of heterozygosity (LOH) patterns, (iv) randomness of DNA segment order/joints, (v) alternating head-tail sequences of rearrangement reads and (vi) rearrangements affecting only one haplotype. All hallmarks defining the one-off nature of chromothripsis could be confirmed by rearrangement analysis and computational simulations (Figure 9q-w). In addition, M-FISH confirmed that the massive loss of genetic DNA on chr4 affected only one of both haplotypes but not the other (Figure 9t). Clusters of somatic hypermutation in association with breakpoint junctions, which are characteristic for complex rearrangement types arising through replication-based mechanisms (e.g. chromoanasynthesis) (Forment *et al.*, 2012), were not observed further supporting that the complex rearrangements translocation of a ~200kb region from chr6 was inserted into chr4 while a far smaller fragment from chr4 was in turn integrated into chr6 (Figure 9p). In mPDAC S821,

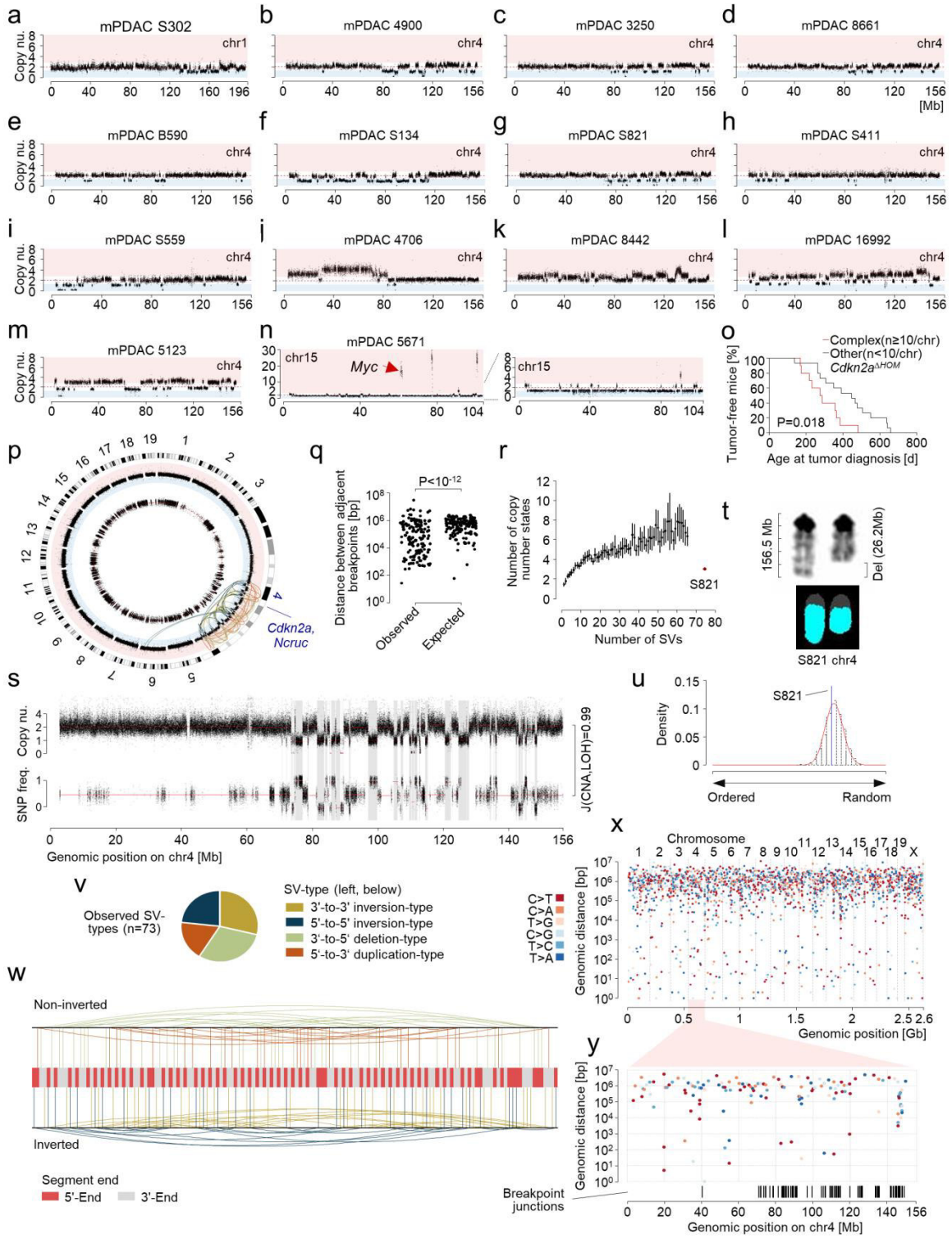

Figure 9 | Complex rearrangements and statistical inference of chromothripsis in primary PDAC cell cultures from PK mice.

[a-n] mPDACs with chromosomes showing complex/clustered rearrangements as detected by aCGH. Nine primary mPDAc show copy number patterns oscillating around a few states indicative of chromothripsis [a-i] whereas five primary mPDAC display more complex rearrangements with multiple copy number states that were likely acquired through progressive chromosome rearrangements [j-m]. Primary mPDAC 5671 showed oscillating copy number profile with high-level amplification of *Myc* suggesting the occurrence of chromothripsis with accompanied double minute chromosome formation [n]. **[o]** Survival of PK mice from mPDAC with homozygous deletion of

Cdkn2a (*Cdkn2a*$^{\Delta HOM}$) and complex rearranged chromosomes or tumour genomes without complex rearrangements (complex rearrangements as defined by ≥10 CNA segments per chromosome). Two-sided Log-rank test. **[p]** Circos plot generated from WGS data showing SNPs (inner circle), CNAs (middle circle) and structural variations (outer circle) for primary mPDAc cell culture S821 from PK mice. Chromosome 4 shows massive rearrangements with homozygous deletion of the *Cdkn2a* locus and a balanced translocation of a ~200kb segment from chr6 and a far smaller segment from chr4. **[q-w]** Statistical inference of chromothripsis on the basis of structural variants identified by WGS on chr4 of primary mPDAC S821. Criteria for statistical testing of chromothripsis were previously established by Korbel *et al.* (Korbel *et al.*, 2013). **[q]** The chromothriptic model is characterized by clustering of DNA breakpoints. The distance of adjacent breakpoints showed significantly shorter distances as expected in a progressive model when tested against an exponential distribution ($P<10^{-12}$; χ^2-goodness-of-fit test). **[r]** On a chromothriptic chromosome, copy number patterns tend to oscillate around a few copy number states whereas higher states can be observed in the progressive model. The number of observed copy number states was simulated for the progressive model by using Monte Carlo approach with sequential accumulation of structural variants observed on chr4. The number of copy number states observed on chr4 was decreased as compared to the progressive model. Black points indicate mean, black lines show 95% CI. **[s]** Chromosomes affected by chromothripsis typically show interspersed loss and retention of heterozygosity. Chr4 display a strong overlap of heterozygous deletions and regions of LOH (Jaccard index (J)=0.99). **[t]** Chromothripsis typically affects only one of both haplotypes. In M-FISH, only a single copy of chr4 shows loss of chromosomal content indicating that only one haplotype is affected by chromothripsis. **[u]** In a chromothriptic model, chromosome shattering and re-joining of fragment occurs randomly. For this, a background probability distribution was generated by re-ordering of observed chr4 fragments in running Monte Carlo simulations. The observed chr4 segment order was found to lie within the null model of chromothripsis. Two-sided P=0.78. **[v]** In the chromothriptic model, all 4 types of structural variations occur with similar frequency. **[w]** Due to the one-off nature of chromothripsis, the read-orientation of structural variants detected by paired-end sequencing show an alternating sequence of 3'-/5'-read orientation when mapped and ordered to their position on the reference chromosome. The alternating sequence of 3'-/5'-read orientation was tested by applying the right-sided Wald-Wolfowitz runs test. $P<10^{-12}$. **[x]** Rainfall plot of primary PDAC genome from PK mouse S821 for the detection of clusters with high mutation rates on the basis of WGS. Dots represent SNVs and dot colours indicate nucleotide substation type. The genomic position of the SNV is shown on the x-axis and the distance of the SNV to the previous SNV is shown on the y-axis. **[y]** Rainfall plot with Chr4 zoom-in from [x]. Breakpoint junctions on chr4 as detected by WGS are plotted below the SNVs. No clusters with increased mutation rates could be identified on chr4. This provides further evidence for chromothripsis as the underlying mutational process. Chromothripsis does involve end joining DNA repair mechanism that are not producing clusters of hypermutation. In contrast, other mutational processes that can generate complex rearrangements through replication-based processes, such as chromoanasynthesis, show breakpoint-associated clusters of hypermutation. Analysis of WGS raw data in panels [p-w] was performed by Thomas Engleitner and Maximilian Zwiebel. (The whole figure was modified from Mueller *et al.*, 2018).

chromothripsis did not cause knockout of multiple tumour suppressor genes through intra- or interchromosomal rearrangements. Of note, the detection of chromothripsis in a pancreatic cancer mouse model allows for the experimental interrogation of various hypotheses associated with this genetic phenomenon. For example, a recent study by Notta *et al.* proposed that complex rearrangements in human PDAC could trigger accelerated evolution of preneoplastic cancer lesions. This idea was based on the finding that PDAC patients with chromothripsis have a shorter survival in comparison to patients without chromothripsis. This analysis can be confounded by multiple factors such as different responses of PDAC with/without chromothripsis during chemotherapy or different on-set of tumour development making it almost impossible to estimate the time of tumour

evolution from a preneoplastic lesion to full-blown PDAC. However, the “synchronized nature” of tumour initiation through the activation of the mutant *Kras*^{G12D}-allele in the PK pancreatic cancer mouse model allows for unbiased validation of this observation. Animals with PDACs having complex rearrangements showed indeed a shortened time to tumour development supporting the finding that tumour suppressor loss (here *Cdkn2a*) through genomic catastrophes triggers more rapid pancreatic cancer formation than inactivation of *Cdkn2a* through whole chromosome loss or focal deletion (Figure 9o).

The comprehensive genetic analysis of primary cell cultures derived from a *Kras*^{G12D}-driven mouse model of pancreatic cancer revealed similarities and differences in comparison to the genetics of human PDAC. Genomic alterations that are generated during cellular crisis, such as chromothripsis, as well as overlapping mutational signatures were observed in mouse and human PDAC indicating that the mutational processes that operate during tumour evolution are similar in both species. As already observed in other mouse models of human cancer (such as non-small-cell lung cancer [NSCLC] or lung adenocarcinoma) (McFadden *et al.*, 2014; McFadden *et al.*, 2016; Chung *et al.*, 2017), the mutational burden detected in *Kras*^{G12D}-driven primary pancreatic cancer cell cultures from mice was significantly reduced as compared to hPDAC. Besides deletion of *Cdkn2a*, recurrently mutated genes (such as *TP53* or *SMAD4* in hPDAC) were almost absent in mPDAC revealing significant cross-species differences in the need for the acquisition of specific gene alterations. However, at the same time the *Kras*^{G12D}-driven mouse model of pancreatic cancer develops similar biological and clinical phenotypes as the human disease, most relevant: (i) progression of pancreatic cells through a PanIN-PDAC-sequence, (ii) ductal histology of tumours with different grades of cellular dedifferentiation and (iii) metastatic dissemination of tumour clones to lymph nodes, liver and/or lung. How the mouse is able to resemble major human PDAC phenotypes in a background of reduced mutational load and if these findings are relevant to the understanding of the human disease will be investigated in the remaining work of my thesis.

3.2 Increased gene dosage of *Kras^{G12D}* links early progression and metastasis of PDAC

As shown above, the mutational burden as well as the incidence of recurrently mutated genes was significantly decreased in mPDAC as compared to hPDAC. The overlay of the copy number profiles derived from aCGH data of primary pancreatic cell cultures from PK mice (n=38) revealed two frequent genetic alterations that occurred in more than 30% of tumours from PK mice: (i) amplification of the *Kras* locus and (ii) deletion of the *Cdkn2a* locus (Figure 10a). Amplification of the *KRAS* locus also frequently occurs in human PDAC (Yamada *et al.*, 1986; Heidenblad *et al.*, 2002). In addition LOH analysis based on SNP information extracted from WES data of the same mouse pancreatic cancer cell cultures revealed that chr4 (harbouring the *Cdkn2a* locus) and chr6 (harbouring the *Kras* locus) were most frequently affected by LOH as compared to all other chromosomes of the mouse genome (Figure 10b). The high frequency of CNA and LOH at the *Cdkn2a* or *Kras* locus indicated that both genetic alterations are functionally relevant to mouse PDAC evolution and progression. Therefore, genetic alterations involving the *Kras* locus were investigated in more detail on the basis of WES, aCGH and M-FISH data. The combined analysis of these genetic data revealed four distinct "states" of *Kras^{G12D}* gene dosage in mPDAC from PK mice: (i) focal gain (*Kras$^{G12D\text{-}FG}$*, 7.9% of cases), (ii) arm-level gain (*Kras$^{G12D\text{-}AG}$*, 23.7% of cases), (iii) copy-number neutral loss of the *Kras* wild-type allele (*Kras$^{G12D\text{-}LOH}$*, in 36,8% of cases) or no change/remaining the heterozygous state of the *Kras^{G12D}* allele (*Kras$^{G12D\text{-}HET}$*, 31,6% of cases) (Figure 10c). In total, about two thirds of primary pancreatic cancer cell cultures had allelic imbalances at the *Kras* locus and resulting in increased gene dosage of the mutant *Kras^{G12D}* allele (hereafter together designated to as *Kras$^{G12D\text{-}iGD}$*). Of note, gains affecting less than 50% of the chromosome length were defined as "focal", whereas gains involving the whole chromosome or more than 50% of the chromosome length were defined as "arm-level". WES, aCGH and M-FISH data were interpreted in combination to identify the genetic mechanisms leading to increased *Kras^{G12D}* dosage in mPDAC (Figure 11). Focal gains of *Kras^{G12D}* were observed in 3 primary mPDAC from PK mice, either arising through replication-based mechanisms (2 cases, one with high-level amplification of the *Kras^{G12D}* locus and one with low-level amplification) or translocation which was followed by amplification of the translocated part of chr6 (one case). For example mPDAC S134 shows high-level amplification of a 600kb segment involving the *Kras^{G12D}* locus in aCGH. The small size of the amplified region, its sharp borders and the strong increase in Kras copy number (~8 copies) suggest that

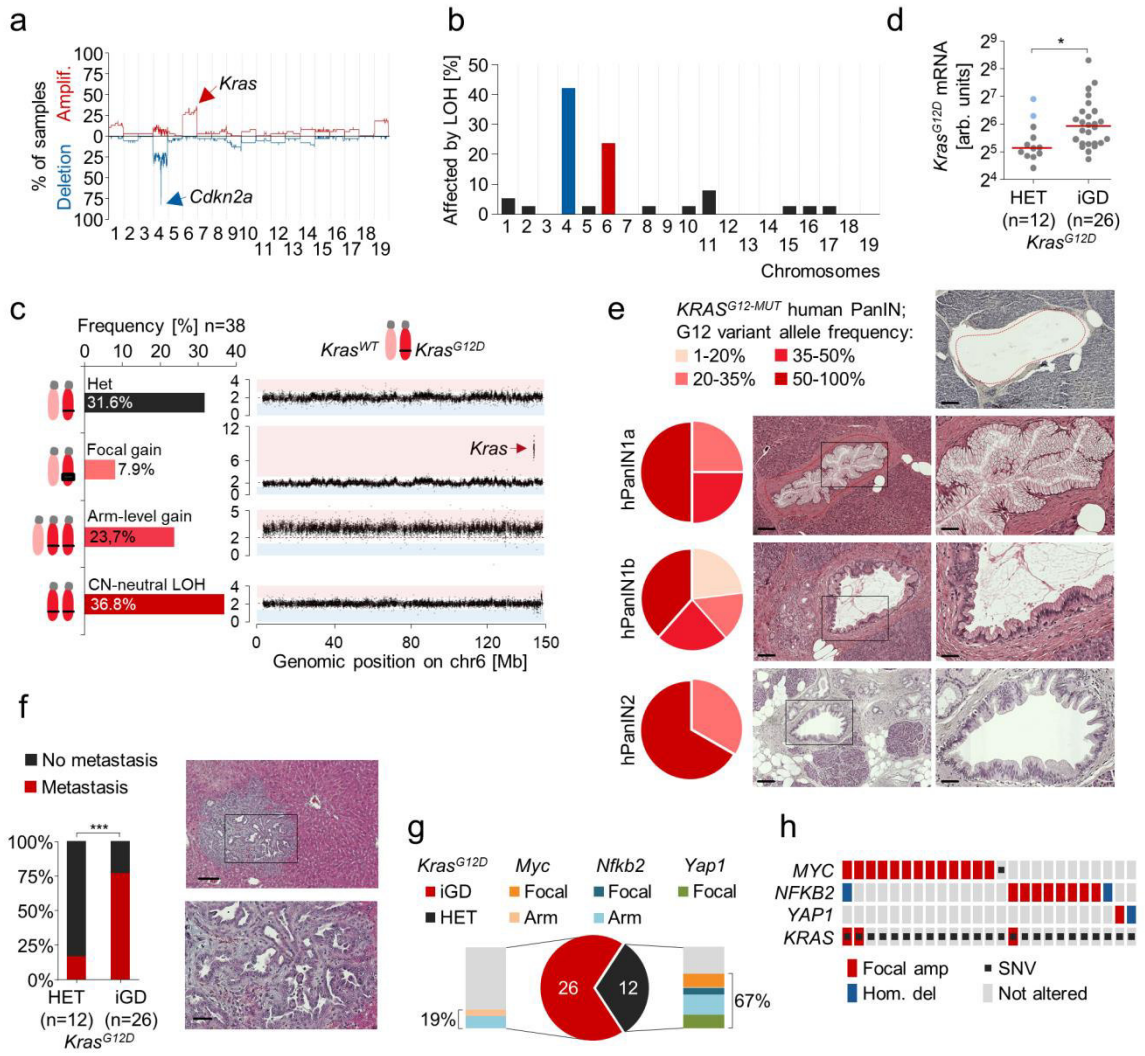

Figure 10 | Increased gene dosage of *Kras*^{G12D} is associated with early progression and metastasis of PDAC.

[a] Copy number overlay of CNA profiles from primary mPDAC cell cultures derived from PK mice (n=38) as detected on the basis of aCGH. Chromosomes are shown on the x-axis and prevalence of chromosomal alterations is given on the y-axis. **[b]** LOH-frequency graph for primary mPDAC cell cultures derived from PK mice (n=38). LOH was detected by the extraction of SNPs from WES data. If the allele frequency of a SNP was changed to ≤0.1 or ≥0.9 in a genomic region with a size of ≥200kb, the chromosome was considered to be affected by LOH. Chr4 (containing the *Cdkn2a* locus) and chr6 (containing the *Kras* locus) were most frequently affected by LOH. **[c]** States of *Kras*^{G12D} gene dosage in primary mPDAC cell cultures from PK mice (n=38) as detected by the combined analysis of WES, aCGH and M-FISH data. The prevalence of each state is indicated on the left while an exemplary CNA plot is given on the right. X-axis in the CNA plot represents the genomic position on chr6 and y-axis shows copy number relative to a diploid genome. **[d]** *Kras*^{G12D} gene dosage state influences mutant-specific *Kras*^{G12D} transcript levels as detected by the combined analysis of qRT-PCR and amplicon-based RNA-Seq data. *P=0.02, Mann-Whitney test; bars, median. **[e]** Stage-specific G12 variant allele frequencies in microdissected human PanIN with *KRAS*^{G12} mutation as detected on the basis of amplicon-based deep sequencing (n=24). For each PanIN stage, representative H&E stains are shown. H&E stains of consecutively cut specimens were used for the histopathologic diagnosis of corresponding PanIN lesions. PanIN1a lesion after microdissection is shown. Scale bars, 150 µm (left and top right) and 50 µm; squares indicate the area of zoom-in. **[f]** Frequency of distant organ metastasis of mPDAC with *Kras*$^{G12D\text{-}HET}$ as compared mPDAC with Kras$^{G12D\text{-}iGD}$ status. (***P=0.001, Fisher's exact test). Metastasis was assessed macroscopically and by analysing serially-cut tissue sections for the presence of micro-metastases. H&E stained sections show liver metastasis. Scale bars, 150 µm (top) and 50 µm

(bottom); square indicate the area of zoom-in. **[g]** Identification of alternative oncogenic amplifications in *Kras*$^{G12D\text{-}HET}$ mouse PDAC. mPDAC without increased gene dosage of *Kras*^{G12D} show gain of either *Myc*, *Nfkb2* or *Yap1* to intensify partial aspects of Ras signaling for progression to PDAC. Arm, arm-level amplification; Focal, focal amplification. **[h]** Human PDAC show frequent amplification of *MYC*, *NFKB2* and *YAP1* in a cohort of microdissected hPDAC tissues (data from Witkiewicz *et al*., 2015). Of note, alternative oncogenic amplifications can also cooperate with *KRAS*$^{MUT\text{-}iGD}$ status. (The whole figure was modified from Mueller *et al*., 2018).

Kras^{G12D} was amplified by multiple cycles of repeated template-switching of a replication-based DNA repair mechanism (Figure 11a). Pancreatic cancer S134 had a mutant allele frequency of *Kras*^{G12D} of 89.1% as detected by amplicon-based deep sequencing of the *Kras* locus. mPDAC 4706 carried a focal amplification of the *Kras*^{G12D} locus which involved a translocation with chr4 as detected by M-FISH. Most likely, a reciprocal translocation of chr4 and chr6 occurred first thereby generating two rearranged chromosomes: Der(4)T(4;6) and Der(6)T(4;6). In a second step, the rearranged Der(6)T(4;6) was missegregated during mitosis leading to the focal gain of the distal chr6 region carrying the mutant *Kras*^{G12D} allele (Figure 11b). The mutant allele frequency of *Kras*^{G12D} in tumour 4706 was 72.2%. Arm-level gain of the *Kras* locus occurred in 9 of 38 mPDACs (~24%), with 7 cases showing amplification of the whole-chromosome and 2 tumours with concomitant intra-chromosomal deletions or translocations that however did not involving the *Kras* locus. For example, primary pancreatic tumour R1035 harboured three copies of chr6 ("classical" trisomy) that was most likely the result of chromosome missegregation during mitosis (Figure 11c). The mutant allele frequency of this tumour was 69.8%. mPDAC 8442 also showed gain of an additional copy of chr6. The observed intra-chromosomal deletion resulted in the deletion of a19.6Mb segment on one of the three copies of chr6, but did not involve the *Kras* locus which is located on the distal end of the chromosome arm. The reduced length of one of the three chr6 copies harbouring the 19.6Mb deletion could be also visualised by M-FISH (Figure 11d). The mutant allele frequency of *Kras*^{G12D} in mPDAC 8442 was 66.4%. Copy-number neutral LOH of *Kras*^{G12D} (*Kras*$^{G12D\text{-}LOH}$) occurred in 14 of 38 PDAC (~36.8%) from PK mice. In principal, two mechanism can generate *Kras*$^{G12D\text{-}LOH}$ (homozygosity of *Kras*^{G12D}): (i) mitotic recombination of the *Kras* locus or (ii) mitotic error resulting in gain of an additional copy of chr6 with *Kras*^{G12D} and subsequent loss of the chr6 copy carrying the wild-type *Kras* locus. For example, mPDAC 16992 and B590 showed copy-number neutral LOH of the *Kras*^{G12D} locus with *Kras* mutant allele frequencies of 99.2% and 96.3%, respectively. However, both tumours differed in their chr6 SNP patterns as detected by WES. While the whole chr6 was affected by CN-neutral LOH in mPDAC 16992 (indicating chromosome missegregation as the underlying mechanism), only a partial region of chr6 containing the *Kras* locus underwent CN-neutral LOH in mPDAC B590 (therefore most likely resulting

from mitotic recombination) (Figure 11e,f). Taken together, approximately two thirds of all primary pancreatic cancer cell cultures from PK mice showed increased dosage of *Kras*^{G12D} (hereafter referred to as *Kras*$^{G12D\text{-}iGD}$). Notably, allelic imbalances at the *Kras* locus always affected the mutant *Kras*^{G12D} allele suggesting selective pressure to increase *Kras*^{G12D} dosage during pancreatic tumour evolution/progression. *Kras*$^{G12D\text{-}iGD}$ also affected the transcriptional output of the mutant *Kras*^{G12D} allele. The allele-specific expression of mutant *Kras*^{G12D} mRNA was investigated by the combined interpretation of amplicon-based RNA-seq and qRT-PCR. RNA-seq data was used to determine the proportion of mutant over wild-type *Kras* transcripts while qRT-PCR was used to determine the absolute quantity of total (wild-type plus mutant) *Kras* transcripts. The mutant-specific expression of *Kras*^{G12D} can then be calculated by multiplying the proportion of *Kras*^{G12D} mRNA levels with the absolute expression of *Kras*, revealing that primary mPDAC with *Kras*$^{G12D\text{-}iGD}$ had significantly increased expression of *Kras*^{G12D} in comparison to cancers with *Kras*$^{G12D\text{-}HET}$ (Figure 10d). This analysis also revealed, that two out of 12 mPDAC with *Kras*$^{G12D\text{-}HET}$ status lost expression of the *Kras* wild-type allele suggesting non-genetic mechanisms

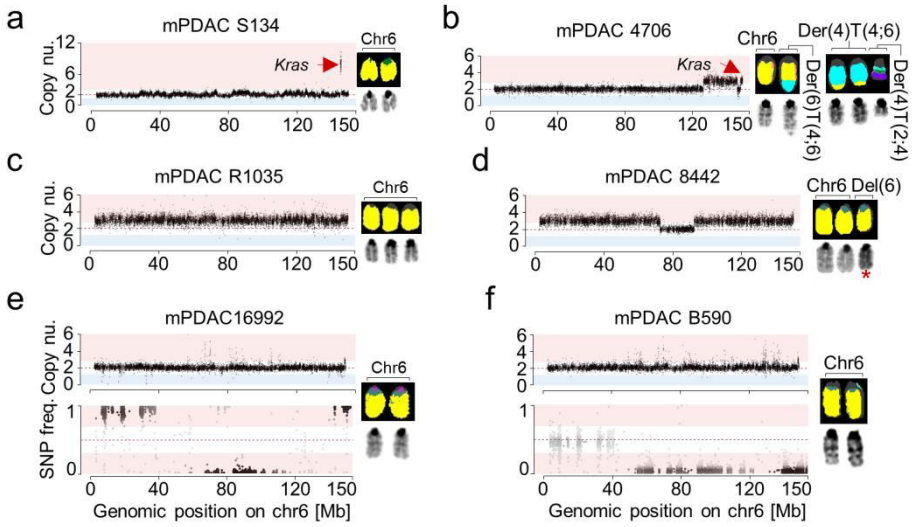

Figure 11 | Genetic mechanisms of *Kras*^{G12D} gene dosage alterations

[a-f] Genetic mechanisms leading to alterations in *Kras*^{G12D} gene dosage were analysed on the basis of aCGH, M-FISH and whole exome sequencing (WES) in primary mPDAC cell cultures from PK mice (n=38). For each genetic mechanism, a representative tumour is shown. **[a]** High-level focal amplification of the *Kras*^{G12D} locus likely through multiple cycles of repeated template-switching of a replication-based DNA repair mechanism. **[b]** Focal amplification of *Kras*^{G12D} locus due to translocation of chr4 and chr6 and subsequent missegregation of the translocated chromosome resulting in gain of an additional copy of *Kras*^{G12D}. **[c-d]** Whole chromosome amplification of chr6 (trisomy) involving the *Kras*^{G12D} locus through mitotic missegregation. **[e-f]** Copy-number neutral LOH of the *Kras*^{G12D} locus. LOH of the whole chr6 in mPDAC 16992 suggests chromosome missegregation as the underlying mechanism [e], whereas LOH of only a distal region of chr6 in mPDAC B590 indicates that this alteration results from mitotic recombination. (The whole figure was modified from Mueller *et al.*, 2018).

of silencing the *Kras* wild-type expression and further supporting the importance *Kras*^{G12D} gene dosage increase for driving pancreatic carcinogenesis.

Amplification of *Ras*/*Raf* signalling has been also reported to occur at different stages during Ras-driven tumourigenesis in the mammary gland, lung or intestine (Sarkisian *et al*., 2007; Feldser *et al*., 2010; Junttila *et al*., 2010; Rad *et al*., 2013). To identify at which stage during pancreatic cancer progression gene dosage of *Kras*^{G12D} is being amplified, low-grade human pancreatic intraepithelial neoplasias (hPanIN; precursors of hPDAC) from were micro-dissected from 19 human individuals. The allele frequencies of *KRAS* hotspot mutations was determined by amplicon-based deep sequencing of *KRAS* exon-2. Interestingly, human PanIN that harboured a *KRAS* mutation frequently showed *KRAS* mutant allele frequencies of >50% indicating the occurrence of *KRAS*$^{MUT\text{-}iGD}$ in these precursor lesions. *KRAS*^{MUT} allele frequencies of >50% were detected in 50%, 38% and 67% of *KRAS* exon2-mutated hPanIN1a, hPanIN1b and hPanIN2, respectively (Figure 10e). Although all hPanIN samples were micro-dissected to remove “contaminating” stromal cells, it is important to note that an estimated fraction of 10% to 60% of non-tumour cells still remained in the analysed samples suggesting that the true frequency of human low-grade PanIN with *KRAS*$^{MUT\text{-}iGD}$ has to be considered even higher. Five hPanIN samples had *KRAS*^{MUT}-allele frequencies of close to 100% that can in principle result from deletion of the *KRAS* wild-type copy without gaining an additional copy of mutant *KRAS* (hemizygosity of *KRAS*^{MUT}: on copy of *KRAS*^{MUT} per cell) or with gaining a second *KRAS*^{MUT} copy (homozygosity/CN-LOH of *KRAS*^{MUT}: two copies of *KRAS*^{MUT} per cell). Interphase FISH of the *KRAS* locus on chr12 and a chr12 centromere probe was used to verify that these precursor lesions indeed gained a second *KRAS*^{MUT} copy. No loss of one KRAS allele nor monosomy of chr12 were detected by interphase FISH (Figure 12a-c). Moreover, the detection of *KRAS*$^{MUT\text{-}iGD}$ in human PanINs might be also biased through adjacent tumours that migrate into pancreatic ducts thereby cross-contaminating micro-dissected PanINs. The false-positive detection of *KRAS*$^{MUT\text{-}iGD}$ in human PanINs can however be excluded due to multiple reseons: (i) *KRAS*$^{MUT\text{-}iGD}$ in IMPN-related hPanINs in patients without invasive PDAC, (ii) the large distance of analysed samples to associated pancreatic cancers and (iii) the presence of different *KRAS* exon-2 mutations in PanINs and associated cancers within the same patient. The high frequency of *KRAS*$^{MUT\text{-}iGD}$ in low-grade PDAC precursor lesions further supports the notion that increasing the dosage of the mutant *KRAS* allele is critical during early PDAC evolution and progression.

No recurrently mutated metastasis driver has been identified yet in human pancreatic cancer (Campbell *et al*., 2010; Yachida *et al*., 2010; Makohon-Moore *et al*., 2017). Therefore, the impact of increased dosage of mutant *Kras*^{G12D} in mouse PDAC on the

capability of tumour cells to form metastasis in the liver, lung or lymph nodes was analysed. Strikingly, the vast majority of metastasized primary PDAC from PK mice were *Kras*$^{G12D\text{-}iGD}$ (20 of 22 primary mPDAC) whereas primary mPDAC that did not metastasize were predominantly *Kras*$^{G12D\text{-}HET}$ (10 of 16 primary mPDAC). Thus, *Kras*$^{G12D\text{-}iGD}$ is associated with a markedly increased metastatic potential (OR 16.7, 95% CI 2.8-98.0, Figure 10f), implicating increased dosage of *Kras*^{G12D} in driving both: early progression and acquisition of metastatic capability. The results from mice also rationalize why analysing the genomes of human PDAC primary/metastases pairs could not identify metastasis-specific gene alterations (Campbell *et al.*, 2010; Yachida *et al.*, 2010; Makohon-Moore *et al.*, 2017). Since dosage increase of *KRAS*^{MUT} occurs during early steps of PDAC evolution (and therefore being already usually present in the primary tumour) its contribution to the metastatic dissemination of the tumour cells could not be detected by analysing primary/metastases pairs. The dual role of increased dosage of mutant *KRAS* for early PDAC progression as well as for metastatic dissemination of tumour cell clones to distant organs explains the early as well as high incidence of PDAC metastasis at first diagnosis in patients.

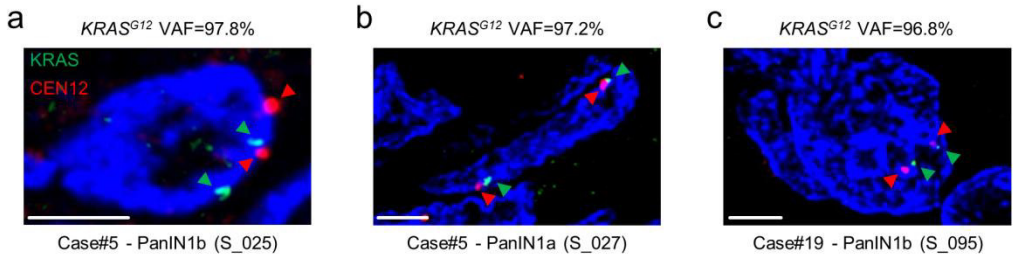

Figure 12 | Interphase FISH of the *KRAS* locus in human PanIN

[a-c] Interphase FISH for chr12 (red) and the *KRAS* locus on chr12 (green) of human PanIN samples with *KRAS*^{G12} variant allele frequencies (VAFs) of ~100%. Scale bars, 2.5µm; CEN12, centromere probe chr12. (The whole figure was modified from Mueller *et al.*, 2018).

3.3 Amplification of "alternative" oncogenes in *Kras$^{G12D-HET}$* primary PDACs

As outlined in the section before, the process of increasing the dosage of oncogenic mutated *KRAS* is important during early steps of PDAC carcinogenesis. However, only two thirds of all primary PDAC cell cultures from PK mice had *Kras$^{G12D-iGD}$* (Figure 10c), raising the question if there are distinct genetic events/alterations that are driving progression of mPDAC with *Kras$^{G12D-HET}$* status. WES, aCGH and M-FISH data from these tumours were screened for additional oncogenic alterations that might cooperate with *Kras$^{G12D-HET}$*. Interestingly, high-level amplification of *Myc* could be detected in two cases and amplification of *Yap1* was observed in another two cancers out of 12 primary mPDAC with *Kras$^{G12D-HET}$* status (Figure 10g and 13a-d). *Myc* and *Yap1* are well known oncogenes and were reported to cooperate with mutant *Kras* in pancreatic cancer mouse models (Kapoor *et al.*, 2014; Shao *et al.*, 2014; Stellas *et al.*, 2014; Zhang *et al.*, 2014; Diersch *et al.*, 2016). cBioPortal, a platform for visualization, analysis and download of large-scale cancer genomics data sets (Cerami *et al.*, 2012; Gao *et al.*, 2013), was used to investigate the frequency of *MYC* and *YAP1* amplifications in the data set of micro-dissected primary human PDAC from Witkiewicz *et al.* (Witkiewicz *et al.*, 2015). High-level amplifications of *MYC* and *YAP1* were detected in 12% (13 out of 109) and 1% of primary hPDAC cases (1 out of 109), respectively (Figure 10h). Since only low-level amplifications of *Yap1* were found in the mouse PDAC cohort and because the analysis using cBioPortal reports only high-level amplifications, the frequency of biologically relevant *YAP1* gains is likely to be underestimated in human PDAC by this analysis. Furthermore, gain of a third copy of chr19 (trisomy of chr19) was enriched in mPDAC with *Kras$^{G12D-HET}$* (3 of 12 cases, 25%) in comparison to primary tumours with *Kras$^{G12D-iGD}$* (4 of 26 cases, 15%), although this observation was not statistically significant (Figure 10g). The high frequency of chr19 trisomy and its enrichment in *Kras$^{G12D-HET}$* tumours suggested that chr19 contains an alternative oncogene that is cooperating with oncogenic *Kras* signalling. However, the identification of the underlying gene is almost impossible based on whole chromosome alterations and requires smaller/focal amplifications to reduce the set of candidates to a handful of genes. Primary mPDAC 4072 carried such a focal amplification on chr19 containing only 20 genes (Figure 13e). Using the data set from Witkiewicz *et al.*, cross-species analyses revealed that the syntenic region was also frequently amplified in human PDAC. Only two genes were in the minimal syntenic peak region: *NFKB2* and *PSD* which were both amplified in 7% of human PDAC (8 of 109 cases) (Figure 13f). Interestingly, mining data from the human protein atlas revealed that *PSD* was not expressed in acinar

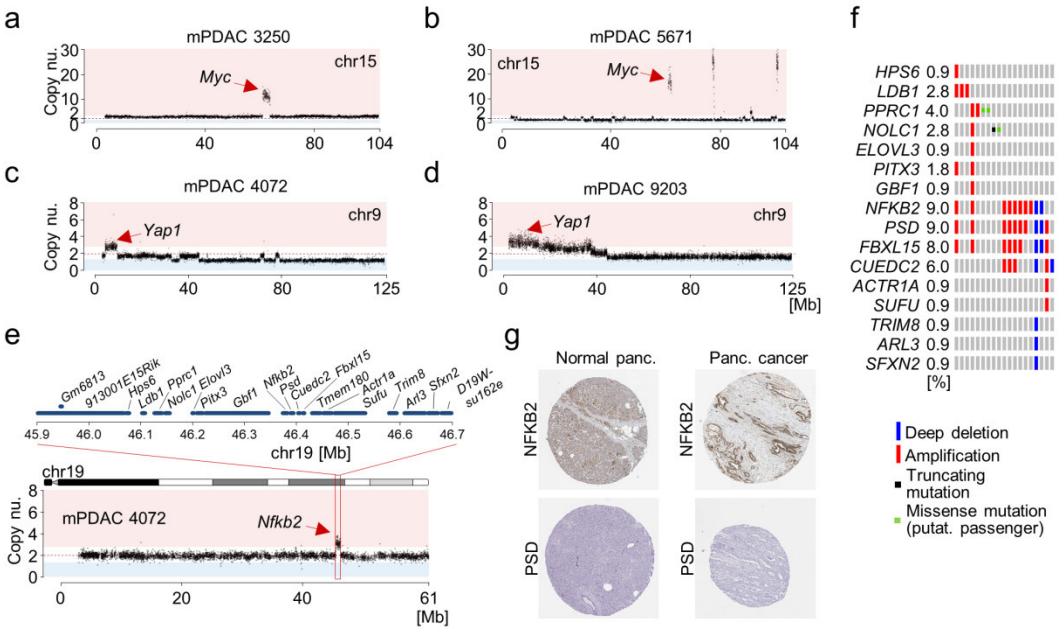

Figure 13 | Amplification of alternative oncogenes in primary mPDAC with *Kras*$^{G12D\text{-}HET}$ status.

[a-b] High-level amplification of *Myc* (chr 15) as detected on the basis of aCGH in primary cultures of mPDAC 3250 [a] and 5671 [b]. **[c-d]** Focal amplification of *Yap1* (chr9) as identified by aCGH in primary cell cultures of mPDAC 4072 [c] and mPDAC 9203 [d] **[e]** Focal amplification of a genomic region on chr19 as detected by aCGH in primary cell culture of mPDAC 4072. The amplified segment contains 20 genes, including *Nfkb2*. **[f]** Cross-species analysis of the chr19 segment from [e] shows recurrent amplification of the syntenic region on chr10 in human PDAC (8 out of 109 cases). Sixteen out of 20 genes involved in the chr19 amplification on mouse PDAC could be assigned to human orthologues. Only *NFKB2* and *PSD* reside in the most frequently amplified minimal region (data from Witkiewicz *et al.*, 2015; oncoplot from cBioPortal (Cerami *et al.* 2012; Gao *et al.*, 2013)). **[g]** Expression of NFKB2 and PSD protein in human pancreas or human PDAC as detected by IHC (data from TheHumanProteinAtlas Uhlen *et al.*, 2015). (The whole figure was modified from Mueller *et al.*, 2018).

or ductal cells of the human pancreas and was also absent human PDAC samples while *NFKB2* showed strong expression in normal tissue as well as pancreatic cancers (Figure 13g). Accordingly, *Nfkb2* was the most likely target gene of the amplified region on chr19 in primary mPDAC 4072. *Nfkb2* is a component of the non-canonical Nfkb signalling pathway, but its role *in vivo* for PDAC development and progression has not been investigated yet. However, *in vitro* it was demonstrated that *Nfkb2* accelerated cell cycle progression in hPDAC cell lines (Schneider *et al.*, 2006). Furthermore, the knockout of the *Nfkb2* interaction partner RelB delayed the formation of PanIN lesions in a *Kras*^{G12D}-driven pancreatic cancer mouse model (Hamidi *et al.*, 2012). These results suggest that *Myc*, *Yap1* and *Nfkb2* are “alternative” oncogenes that can drive primary PDAC development by cooperating with *Kras*$^{G12D\text{-}HET}$ and thereby they obviate the need for increasing *Kras*^{G12D} gene dosage. Altogether, the amplification of partial aspects of Kras downstream signalling

is sufficient to drive early PDAC progression whereas the capability to metastasise to distant organs is linked to the amplification of the full *Kras*^{G12D} signalling program in *Kras*$^{G12D\text{-}iGD}$ PDAC.

3.4 Evolutionary trajectories and tumour suppressor genes license oncogenic dosages

The most frequent gene inactivation occurring in primary PDAC from PK mice involved the *Cdkn2a* locus and/or *Ncruc*, a non-coding *Cdkn2a*-regulatory region located close to the *Cdkn2a* locus (Figure 8d). The *Cdkn2a* locus encodes for the two potent tumour suppressor genes *Ink4a* and *Arf*, which were shown to restrain *Kras*^{G12D}-driven pancreatic carcinogenesis in the mouse through the Rb1 or Trp53 pathway, respectively (Bardeesy *et al.*, 2006a). *Cdkn2a/Ncruc* were inactivated through different types of chr4 alterations: (i) loss of one copy of chr4 (arm-level), (ii) complex rearrangements with ≥10 CNAs on chr4 (complex), (iii) focal deletion (focal) or (iv) translocation of chr4 with other chromosomes (translocation) (Figure 14a,b). Interestingly, the *Cdkn2a* locus was not altered in all primary mPDAC to the same extent: 27 of 38 mPDACs showed complete/homozygous inactivation of the *Cdkn2a* locus while the other 11 tumours did not lose both copies of the *Cdkn2a* locus (10 cases with heterozygous deletions and one cancer with no detectable genetic alterations of *Cdkn2a*). One possible explanation for this observation was that the dosage of oncogenic *Kras*^{G12D} and the extent of *Cdkn2a* deletion might be interconnected. Indeed, the majority of mPDAC with homozygous *Cdkn2a* deletions (hereafter referred to as *Cdkn2a*$^{\Delta HOM}$) were *Kras*$^{G12D\text{-}iGD}$ (23 of 27 cases, 85%) whereas mPDAC with incomplete *Cdkn2a* inactivation (hereafter referred to as *Cdkn2a*$^{\Delta HET/WT}$) predominantly remained the *Kras*$^{G12D\text{-}HET}$ status and did not increase *Kras*^{G12D} gene dosage (8 of 11 cases, 73%) (Figure 14c). In addition the allele-specific expression of mutant *Kras*^{G12D} mRNA in *Cdkn2a*$^{\Delta HOM}$ and *Cdkn2a*$^{\Delta HET/WT}$ cohorts were compared by combining amplicon-based RNA-seq with qRT-PCR of the *Kras* locus. As already described above, RNA-seq data was used to determine the proportion of mutant over wild-type *Kras* transcripts while qRT-PCR was used to determine the absolute quantity of total (wild-type plus mutant) *Kras* transcripts. In line with the genetic data, allele-specific expression of oncogenic *Kras*^{G12D} was significantly increased in primary mPDAC with *Cdkn2a*$^{\Delta HOM}$ as compared to tumours with *Cdkn2a*$^{\Delta HET/WT}$ status further supporting the proposed functional interaction of both loci (Figure 14d). To verify this observation in

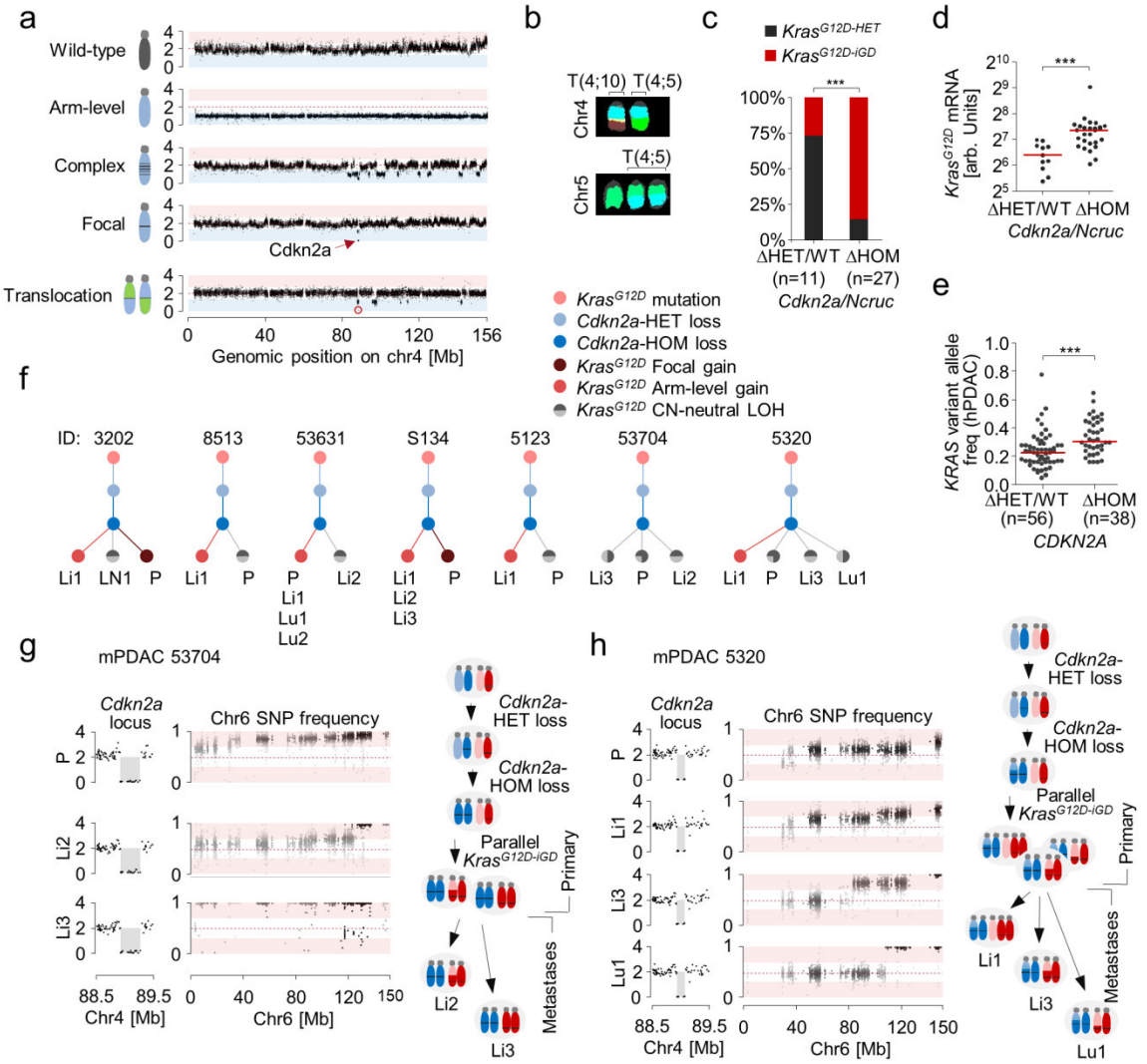

Figure 14 | *Cdkn2a* tumour suppressor gene states define distinct evolutionary trajectories and *Kras*^{G12D} dosage in PDAC.

[a] Types of chr4 alterations involving the *Cdkn2a* locus in primary mPDAC cell cultures from PK mice (n=38) as detected by the combined analysis of aCGH and M-FISH data. The prevalence of each chr4 alteration type is shown on the left while an exemplary CNA plot is given on the right. X-axis in the CNA plot represents the genomic position on chr4 and y-axis shows copy number relative to a diploid genome. **[b]** Inter-chromosomal translocations involving the *Cdkn2a* locus on chr4 as identified by M-FISH. **[c]** Frequency of *Kras*$^{G12D\text{-}iGD}$ in cell cultures of primary mPDAC with heterozygous/no inactivation of the *Cdkn2a*/*Ncruc* locus (ΔHET/WT) as compared to homozygous inactivation of the *Cdkn2a*/*Ncruc* locus (ΔHOM). ***P=0.001, Fisher's exact test, OR 15.3, 95% CI 2.8-83.9. **[d]** Mutant-specific *Kras*^{G12D} transcript levels as detected by the combined analysis of qRT-PCR and amplicon-based RNA-Seq data in cell cultures of primary mPDAC with *Cdkn2a*/*Ncruc*$^{\Delta HET/WT}$ as compared to mPDAC with *Cdkn2a*/*Ncruc*$^{\Delta HOM}$ status. ** P=0.003, two-tailed Mann–Whitney test; bars, median. **[e]** Variant allele frequency of *KRAS*^{MUT} in human PDAC with heterozygous/no inactivation of the *CDKN2A* locus (ΔHET/WT) in comparison to homozygous inactivation of the *CDKN2A* locus (ΔHOM). Data from Witkiewicz *et al.*, 2015. ***P≤0.001, Mann-Whitney test; bars, median. **[f]** Evolution of *Cdkn2a* and *Kras*^{G12D} states based on chr4 and chr6 CNA/LOH patterns in metastatic mPDAC from PK mice. In total, thirteen metastatic mPDAC with a total of 25 corresponding metastasis were analysed by aCGH and WES. Seven cases and 16 associated metastases are shown, for which the order of genetic events (represented by dots) could be reconstructed. Bifurcations indicate divergent evolution following a shared genetic event. The distance of dots does not correspond to evolutionary distances. P, primary mPDAC; Li, liver metastasis; Lu, lung metastasis; LN, lymph node metastasis. **[g,h]** Detailed exemplification of the evolution of *Cdkn2a* (chr4) and *Kras*^{G12D} (chr6) states in metastatic mPDAC 53704 [g] and

metastatic mPDAC 5320 [h]. CNA plots are based on aCGH data and show identical *Cdkn2a* deletions in the primary mDPAC and the corresponding metastases (y-axis shows copy number relative to a diploid genome). SNP profiles were analysed on the basis of WES data and show distinct patterns on chr6 in the metastases and the primary tumour (y-axis shows SNP allele frequency). Schemes were established from the combined interpretation of aCHG and WES data. (The whole figure was modified from Mueller *et al.*, 2018).

human PDAC, the data set of microdissected hPDAC was analysed. As in the mouse, $KRAS^{MUT}$ variant allele frequencies (indicative of $Kras^{G12D\text{-}iGD}$) were significantly higher in hPDAC with $CDKN2A^{\Delta HOM}$ versus $CDKN2A^{\Delta HET/WT}$ (Figure 14e). Accordingly, the status of *CDKN2A* deletion and dosage of mutated *KRAS* are tightly linked in mouse and human PDAC. In principle there are two possible scenarios how both genes interact during pancreatic tumour evolution *in vivo*: (i) $KRAS^{MUT\text{-}iGD}$ occurs first, but activates the *CDKN2A* tumour suppressor locus, thereby inducing senescence and preventing cell cycle progression until *CDKN2A* is lost (as proposed in a $Hras^{G12V}$-driven mammary cancer mouse model) (Sarkisian *et al.*, 2007); or (ii) homozygous deletion of *CDKN2A* occurs first, thereby allowing for increasing $KRAS^{MUT}$ gene dosage in a second event. Both scenarios describe two fundamental different principles of tumour evolution. In the first setting, the senescence state can be bypassed and reverted by complete inactivation of *CDKN2A*, while in the second setting the senescence state needs to be prevented because pancreatic carcinogenesis would be otherwise irreversibly arrested *in vivo*.

To address the interaction of oncogenic $Kras^{G12D}$ and the tumour suppressor locus *Cdkn2a in vivo*, the CNA and LOH profiles of primary PDACs and their corresponding metastases from PK mice were compared. Thirteen primary mPDAC with $Cdkn2a^{\Delta HOM}$, $KRAS^{G12D\text{-}iGD}$ status and 25 associated metastases were available in the PK cohort. The combined analysis of CNA and LOH profiles of these tumours allowed for the simultaneous detection of copy number alterations and allelic imbalances (CN-neutral alterations) at the *Cdkn2a* and *Kras* locus. Interestingly, deletion patterns involving the *Cdkn2a* locus were identical in all 13 primary/metastases pairs whereas CNA and LOH patterns involving the *Kras* locus on chr6 were variable in metastases of the same mouse when compared to the corresponding primary tumour in 7 of 13 cases (Figure 14f). Thus, homozygous deletion of *Cdkn2a* preceded dosage increase of $Kras^{G12D}$ in all genetically discordant cases. Of note, the sequential order of *Cdkn2a* deletion and acquisition of $Kras^{G12D\text{-}iGD}$ could not be reconstructed in the remaining 6 mPDAC/metastases pairs either because the SNP density was too low in 4 cases (due to homozygous genetic background of the mice) or chr6 CNA and LOH patterns were similar in the primary mPDAC and the corresponding metastases (in 2 cases). The detailed analyses of two mPDAC/metastases pairs with genetically discordant $Kras^{G12D\text{-}iGD}$ patterns are shown in Figure 14g,h. Primary mPDAC

53704 had two liver metastases. All three cancer cell cultures had identical patterns of homozygous *Cdkn2a* deletion on chr4 as identified by aCGH. In contrast, SNP patterns on chr6 were distinct: liver metastasis 2 (Li2) showed partial chr6 LOH at a distal region of the chromosome (probably resulting from mitotic recombination of a partial chromosome arm region of chr6) while in liver metastasis 3 (Li3) the whole chromosome arm of chr6 underwent LOH (most likely through chromosome missegregation during mitosis). These observations confirm that *Cdkn2a* was deleted homozygously before diversification of chr6 CNA/LOH patterns and $Kras^{G12D\text{-}iGD}$ occurred. When considering that both liver metastases arose from two distinct clones in the primary tumour the findings also explain the gradual SNP pattern on chr6 in the primary pancreatic tumour (Figure 14g). The combined analysis of CNA/LOH patterns in the primary mPDAC 53704 and its corresponding liver metastases suggests the following sequence of genetic alterations: (i) the initial step of Cre-mediated activation of $Kras^{G12D}$ in pancreas cells of the mouse was followed by a heterozygous deletion of *Cdkn2a*, (ii) in the next step the second copy of *Cdkn2a* was lost due to chromosome missegregation during mitosis resulting in CN-neutral LOH of chr4 and homozygous *Cdkn2a* deletion and (iii) complete loss of *Cdkn2a* allowed for convergent evolution of tumour clones with distinct patterns of $Kras^{G12D\text{-}iGD}$ that later on gave rise to independent metastasis in the liver (Figure 14g). In another metastatic mPDAC from PK mouse 5320, the primary tumour showed the identical deletion of *Cdkn2a* as both liver metastases and the lung metastasis revealing that all cells derive from the same ancestral cell. By contrast, CNA and SNP patterns of chr6 were discordant in all four cancer cell cultures. Liver metastasis 1 (Li1) showed whole chromosome arm amplification of chr6 (trisomy, $Kras^{G12D\text{-}AG}$) as identified by the combined interpretation of aCGH and WES. Liver metastasis 3 (Li3) and lung metastases 1 (Lu1) did not show chr6 CNAs but were affected by distinct patterns of $Kras^{G12D\text{-}LOH}$. While in Lu1 a small region (involving the *Kras* locus) of the distal arm of chr6 underwent LOH, a larger region was affected by LOH in Li3. Thus, all three metastases of mPDAC 5320 harbour distinct patterns of $Kras^{G12D\text{-}iGD}$ explaining the complex gradual SNP pattern of the corresponding primary mPDAC (Figure 14h). Again, these analyses also show that convergent evolution of allelic imbalances at the $Kras^{G12D}$ locus occurs after homozygous deletion of *Cdkn2a*.

Altogether, these results provide experimental evidence for several key evolutionary principles in PDAC. First, increase of $Kras^{G12D}$ gene dosage is contingent on complete deletion of *Cdkn2a*. Second, amplification of alternative oncogenes such as *Myc*, *Yap1* or *Nfkb2a* con occur on a background with incomplete/heterozygous deletion of *Cdkn2a*, suggesting a so far unappreciated context-dependent role of *Cdkn2a* haploinsufficiency. Notably, wild-type *Cdkn2a* status was observed only in one out of 38 mPDAC. Third,

parallel acquisition of distinct patterns of increased *Kras*^{G12D} gene dosage after homozygous loss of *Cdkn2a* demonstrates strong functional convergence towards increased gene dosage of oncogenic *Kras*^{G12D} upon complete barrier loss.

However, the finding that *Cdkn2a* is controlling oncogenic levels of *Kras*^{G12D} *in vivo* is mainly derived from correlation-based analyses. To provide additional evidence for functional convergence towards *Kras*$^{G12D\text{-}iGD}$ in a *Cdkn2a*$^{\Delta HOM}$ context *in vivo*, mice with pancreas-specific expression of *Kras*^{G12D} and homozygous knockout of *Cdkn2a* (*Kras*$^{G12D\text{-}Panc}$;*Cdkn2a*$^{\Delta HOM\text{-}Panc}$, hereafter referred to as PKC mice) were generated. The genetic ratio of mutant *Kras*^{G12D} in 16 mPDAC from PKC mice was analysed using aCGH and amplicon-based deep sequencing of the murine *Kras* locus. The combined interpretation of copy number profiles and *Kras*^{G12D} ratio revealed that 100% (16 of 16) of mPDAC derived from PKC mice had *Kras*$^{G12D\text{-}iGD}$ status (Figure 15a). This proof of principle experiment is confirming that acquisition of increased *Kras*^{G12D} gene dosage is the preferred evolutionary trajectory upon homozygous loss of *Cdkn2a*, which might also provide a reasonable explanation for the drastically reduced survival of PKC as compared to PK mice (8.5 vs. 57 weeks as reported by Bardeesy *et al.*, 2006a).

Besides inactivation of *CDKN2A*, additional genetic hallmarks of human PDAC are mutations in the tumour suppressor gene *TP53* and genes of the canonical TGFβ signalling pathway (Jones *et al.*, 2008). The functional relevance of *TP53*, *SMAD4* or *TGFBR2* knockout for pancreatic carcinogenesis *in vivo* were tested and published already before in *Kras*^{G12D}-driven mouse models of pancreatic cancer (Bardeesy *et al.*, 2006a; Bardeesy *et al.*, 2006b; Ijichi *et al.*, 2006). Notably, the knockout of *Trp53*, *Smad4* or *Tgfbr2* was associated with a strongly reduced median survival of the mice of 6.2 (Bardeesy *et al.*, 2006a), 15.7 (Bardeesy *et al.*, 2006b) or 8.4 weeks (Ijichi *et al.*, 2006) respectively in comparison to 57 weeks median survival of PK mice. The shortened median survival of these mice indicates that *Trp53*, *Smad4* and *Tgfbr2* are cooperating with *Kras*^{G12D} signalling in the pancreas, such as knockout of *Cdkn2a* in PKC mice.

To test whether *Trp53* is also a barrier for the acquisition of *Kras*$^{G12D\text{-}iGD}$, 16 primary mPDAC cell cultures from mice with conditional expression of mutant *Kras*^{G12D} in the pancreas with concomitant homozygous deletion of *Trp53* (*Kras*$^{G12D\text{-}Panc}$;*Trp53*$^{\Delta HOM\text{-}Panc}$, hereafter referred to as PKP mice) were analysed. The combined interpretation of copy number data from aCGH with the genetic ratio of *Kras*^{G12D} derived from amplicon-based deep sequencing of the *Kras* locus revealed that all tumours (16 of 16) were *Kras*$^{G12D\text{-}iGD}$ (Figure 15a). Thus, homozygous loss of *Trp53* predisposes for increasing gene dosage of *Kras*^{G12D} as the preferred evolutionary route. This finding makes also sense in the light of

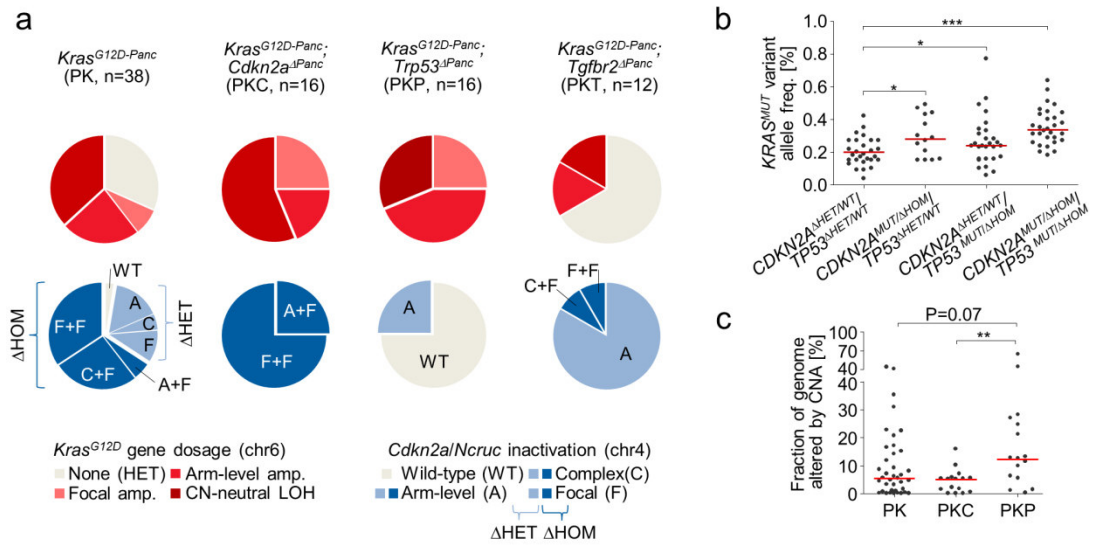

Figure 15 | Types and states of hallmark PDAC tumour suppressor gene alterations differentially licence oncogenic dosage variation in PDAC.

[a] Types and frequencies of *Kras* (chr6) and *Cdkn2a* (chr4) alteration states in mice expressing pancreas-specific $Kras^{G12D}$ (PK) alone or in combination with engineered *Cdkn2a* deletion (PKC), *Trp53* deletion (PKP) or *Tgfbr2* deletion (PKT). **[b]** Variant allele frequency of $KRAS^{MUT}$ in human PDAC with homozygous inactivation of the *CDKN2A* and/or *TP53* locus ($CDKN2A^{MUT/\Delta HOM}$, $TP53^{MUT/\Delta HOM}$) versus human PDAC with heterozygous/no inactivation of *CDKN2A* and *TP53* ($CDKN2A^{\Delta HET/WT}$/$TP53^{\Delta HET/WT}$). Data from Witkiewicz *et al.*, 2015. *P≤0.05, ***P≤0.001, Rank-based ANOVA ($P=5.8*10^{-6}$), P-values for group wise comparisons are shown; bars, median. **[c]** Fraction of the genome altered by CNA in cell cultures of primary mPDAC derived from PK (n=38), PKC (n=16) or PKP mice (n=16). Two-sided rank-based ANOVA (P=0.01); post hoc testing with two-sided Tukey honest significant difference test, ** adjusted P=0.009, adjusted P values for group-wise comparisons are shown; bars, median. (The whole figure was modified from Mueller *et al.*, 2018).

a previous publication from Bardeesy *et al.* showing that inactivation of *Cdkn2a* is bypassed in PKP mice (Bardeesy *et al.*, 2006a). The analysis of microdissected PDAC from the data set of Witkiewicz *et al.* revealed that the allelic ratio of $KRAS^{MUT}$ is significantly increased in tumours with homozygous inactivation of *TP53* when compared to tumours that remained genetic copies of *CDKN2A* or *TP53* (Figure 15b), providing evidence that *TP53* is also a barrier for increasing $KRAS^{MUT}$ dosage in human PDAC. The comparison of copy number profiles from primary mPDAC cell cultures of PK, PKC and PKP mice also showed that PKP tumours had significantly increased CNA levels in comparison to PK and PKC tumours (Figure 15c). mPDAC from PKP mice also had a tendency to amplify $Kras^{G12D}$ through arm-level gain (trisomy), while CN-neutral LOH of $Kras^{G12D}$ predominated in mPDAC from PKC mice (Figure 15a). Interestingly, complex rearrangements were frequently observed on chr4 (and involving the *Cdkn2a* locus) in mPDAC from PK mice but were completely absent in PKC and PKP tumours (Figure 15a). The absence of complex rearrangements in PKC and PKP mPDAC provides functional

evidence *in vivo* that chromothripsis is the consequence of natural selection for *Cdkn2a* inactivation during early stages of pancreatic carcinogenesis in PK mice.

To address the role of *TGFBR2* alterations for $KRAS^{MUT\text{-}iGD}$, 12 primary mPDAC cell cultures from mice with conditional expression of mutant $Kras^{G12D}$ in the pancreas plus concomitant heterozygous or homozygous deletion of *Tgfbr2* ($Kras^{G12D\text{-}Panc}$;$Tgfbr2^{\Delta HET/HOM\text{-}Panc}$, hereafter referred to as PKT mice) were subjected to aCGH and amplicon-based deep sequencing of the murine *Kras* locus. Of note, all primary PKT tumours had genetic alterations involving the *Cdkn2a* locus on chr4. Two mPDAC were $Cdkn2a^{\Delta HOM}$ and had $Kras^{G12D\text{-}iGD}$, while the other ten pancreatic cancers were $Cdkn2a^{\Delta HET}$ and had $Kras^{G12D\text{-}HET}$ in most cases (8 out of 10 cases) (Figure 15a). $Kras^{G12D\text{-}HET}$ mPDACs from PKT mice showed trisomy of chr19 in 50% of cases (4 out of 8 tumours) indicating amplification of the alternative oncogene *Nfkb2*, such as observed in $Kras^{G12D\text{-}HET}$ primary cancers from PK mice. Thus, the frequency of $Kras^{G12D}$ gene dosage gain was significantly reduced in the PKT cohort (4 out of 12 cases) versus the PK cohort (26 out of 38 cases) (P=0.04, Fisher's exact test, odds ratio 0.23, CI 0.06-0.92). Interestingly, mPDAC from PKT mice with heterozygous deletion of the *Tgfbr2* did not inactivate the second functional allele of the *Tgfbr2* gene, as observed already before (Ijichi *et al.*, 2006). This is in contrast to previous publications studying PKC and PKT mice with heterozygous deletions of *Cdkn2a* or *Trp53*, respectively. Almost all primary mPDAC from these mice lost the second allele of the heterozygously inactivated tumour suppressor gene (Aguirre *et al.*, 2003; Hingorani *et al.*, 2005; Bardeesy *et al.*, 2006a) indicating that the cooperation of oncogenic $Kras^{G12D}$ with the inactivation of *Cdkn2a* or *Trp53* is different from the cooperation with *Tgfbr2* inactivation. Contrary to genetic *Cdkn2a* or *Trp53* alterations that favour pancreatic cancer evolution through trajectories with increased $Kras^{G12D}$ gene dosage, *Tgfbr2* alterations direct carcinogenesis through routes with $Kras^{G12D\text{-}HET}$ and *Cdkn2a* haploinsufficiency. Overall, these findings provide functional evidence *in vivo* for evolutionary contingencies and convergence at early stages of pancreatic cancer evolution: different tumour suppressor genes and pathways (*Cdkn2a*, *Trp53* or Tgfβ), their alteration types (ΔHOM, ΔHET) and their combinations (e.g. $Cdkn2a^{\Delta HET}$ plus *Tgfbr2* deletion) license tumour evolution and oncogenic dosage variation into different trajectories.

3.5 Integrating genomes and transcriptomes with pancreatic cancer phenotypes

In the next step, the goal was to integrate principles of genetic evolution with the expression patterns and biological or clinical phenotypes of PDAC. Using a 3-prime end RNA sequencing technology (SCRB-Seq) (Macosko *et al.*, 2015; Parekh *et al.*, 2016), the transcriptomes were generated for all primary mPDAC cell cultures from the PK cohort (n=38). The transcriptomes were sorted into groups by unbiased hierarchical clustering and generated two major clusters C1 and C2. C2 could be further stratified into 3 sub-clusters C2a, C2b and C2c (Figure 16a). No sub-clusters could be identified by visual inspection of the C1 cluster tree. Further, the transcriptomes of C1- and C2-clustered primary mPDAC cell cultures were compared for differential gene expression. This analysis identified 2179 genes to be significantly regulated (adjusted P value ≤0.05) with a log2 fold change ≥0.8 between mPDAC cell cultures from C1 versus C2. The expression of 1064 genes was significantly increased in C1 whereas 1115 genes were upregulated in C2. The lists of differentially regulated genes were then subjected to pathway analysis using the DAVID online tool (Huang da *et al.*, 2009a, b). The top GO term of C2 clustered mPDACs was “epithelial cell differentiation” while C1 mPDACs were defined by terms related to “mesenchymal cell differentiation” (Figure 16b). This suggested that distinct cell morphologies are defining the largest difference in expression patterns of primary mPDAC cell cultures, with C1 cells being mesenchymal and C2 cells being epithelial. Indeed, the comparison of cell culture photos from these primary mPDAC revealed, that all C1 lines had mesenchymal cell morphology and that all C2 lines were epithelial (Figure 16c).

Three previous studies have classified human PDAC tissues or cell lines on the basis of their transcriptional profiles (Collisson *et al.*, 2011; Moffitt *et al.*, 2015; Bailey *et al.*, 2016). All three studies used different experimental approaches to investigate subtypes of human PDAC: (i) Collisson *et al.* microdissected human PDAC from tissue sections to avoid/reduce contamination of cancer cells with pancreas stroma cells and defined 3 subtypes: classical, quasimesenchymal (QM) and exocrine-like (ii) Moffitt *et al.* identified 2 subtypes based on virtual/bioinformatical microdissection, patient-derived xenografts and human PDAC cell lines: classical and basal-like and (iii) Bailey *et al.* used bulk tumour tissues and proposed 4 subtypes of human PDAC: classical, immunogenic, squamous and

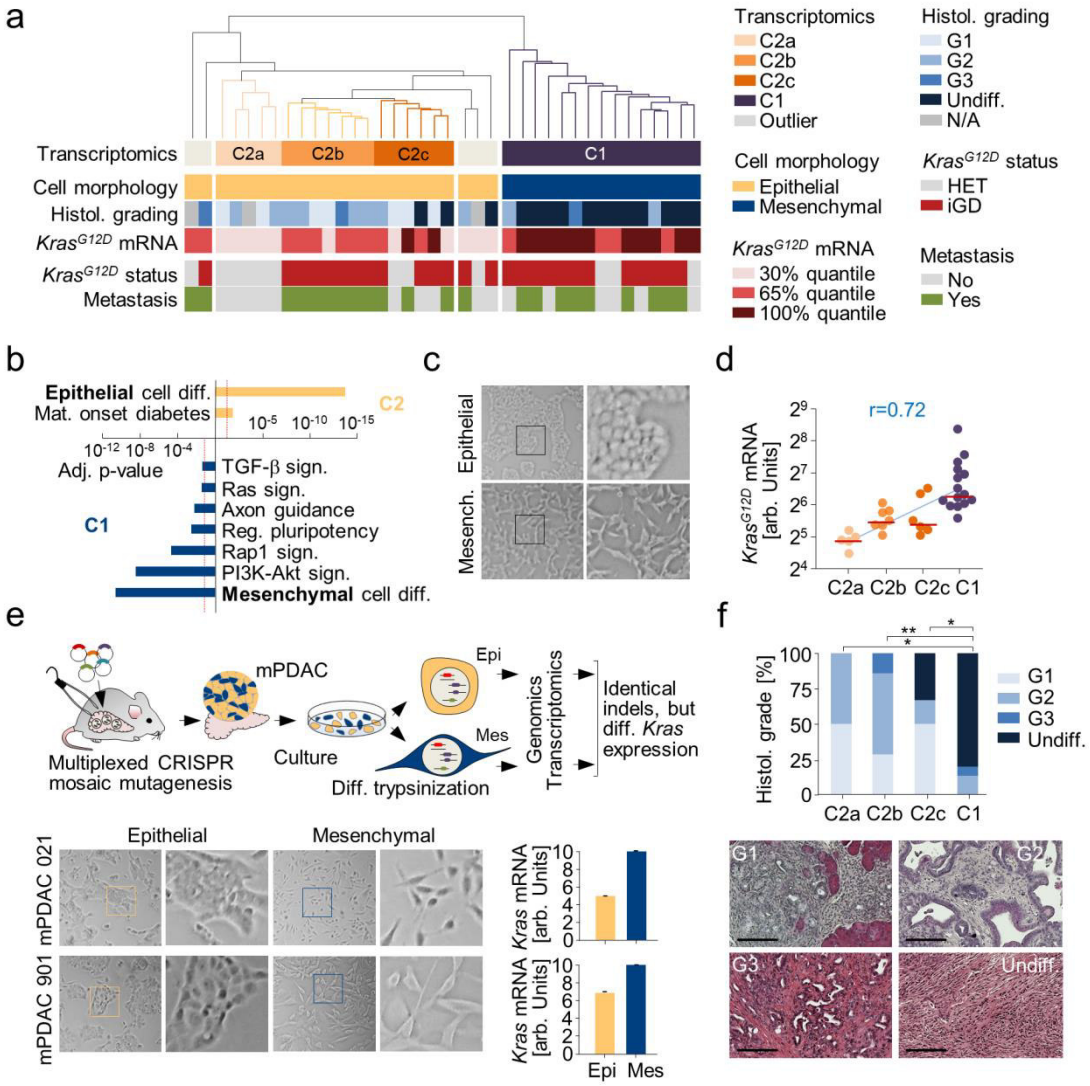

Figure 16 | Integration of transcriptome profiles with genomic data, cellular morphology and histopathology links molecular, morphologic and clinical PDAC phenotypes.

[a] Unbiased hierarchical clustering of transcriptome profiles of primary mPDAC cell cultures derived from PK mice (n=38). Two major transcriptional clusters emerged. C2 could be further sub-stratified into C2a/b/c. For each mPDAC the (i) cell morphology, (ii) histopathological grading, (iii) *Kras*^{G12D} mRNA expression, (iv) genetic *Kras*^{G12D} status and (v) capability of distant organ metastasis is annotated. **[b]** Gene set enrichment analysis of genes differentially regulated between primary mPDAC cell cultures of transcriptional clusters C2 versus C1. Selected gene sets are shown. **[c]** Representative pictures of primary mPDAC cell cultures with epithelial (from C2) or mesenchymal phenotype (from C1) *in vitro*. 100x magnification; squares indicate the area of zoom-in. **[d]** Mutant-specific *Kras*^{G12D} transcript levels in indicated groups of primary mPDAC transcriptional clusters (C2a/b/c/C1, n=5/7/6/15 PK mice) as detected by the combined analysis of qRT-PCR and amplicon-based RNA-Seq data. $P=1.9\times10^{-6}$, two-sided Pearson correlation; bars, median. **[e]** Somatic cancer modelling using CRISPR/Cas9-based multiplexed inactivation of tumour suppressor genes relevant to human PDAC (see also Fig. 18d-g). CRISPR/Cas9 delivery was performed using electroporation-based transfection to achieve low-frequency mosaicism and clonal tumour outgrowth. Primary mPDAC cultures were screened for parallel presence of epithelial and mesenchymal cell morphologies, which were separated/enriched by differential trypsinization. Identical CRISPR/Cas9-induced indel signatures in epithelial/mesenchymal cell pairs of individual mice indicate a common cell of origin for both phenotypes. Representative microscopic images of enriched epithelial and mesenchymal cell populations are shown for two mPDACs. 100x magnification; squares indicate the area of zoom-in. Expression of total *Kras* mRNA normalized to *Gapdh* in epithelial/mesenchymal cell pairs as detected by qRT-PCR.

[f] Histophathology of mPDAC from PK mice in indicated groups of transcriptional clusters (C2a/b/c/C1, n=4/7/6/15 PK mice). Representative H&E-stained sections of different histopathological grades are shown. RNA-Seq raw data was analysed by Thomas Engleitner. Transfection-based delivery of CRISPR/Cas9 components to the murine pancreas *in vivo* was conducted by Roman Maresch and myself. (The whole figure was modified from Mueller *et al.*, 2018).

aberrantly differentiated endocrine exocrine (ADEX). An independent cross-comparison of the proposed human PDAC subtypes was performed to improve the integration of the primary PDAC cell cultures from PK mice into the three human PDAC classification systems. The RNA-seq data from PDAC and adenosquamous pancreatic carcinoma from the data set of Bailey *et al.* was used for this analysis. In the cohort of Bailey *et al.*, RNA-seq data from other histological subentities of pancreatic cancer, such as IPMN (intraductal papillary mucinous neoplasm), MCN (mucinous cystic neoplasm) or acinar cell carcinoma, were excluded from the cross-comparison. Subtyping information according to Bailey *et al.* was already available from that study (Bailey *et al.*, 2016). The subtypes proposed by Collisson *et al.* and Moffitt *et al.* were assigned to the dataset from Bailey *et al.* by using the classifier gene lists from the corresponding study for unbiased hierarchical clustering of RNA-seq data (Figure 17a,b). This subtyping information was then used for the cross-comparison of Bailey subtypes, Collisson subtypes and Moffitt subtypes for each individual hPDAC. The consensus clustering revealed that all three human PDAC classification systems are largely overlapping (Figure 17c). On exception to this was the lack of the exocrine-like subtype from Collisson *et al.* (ADEX in Bailey *et al.*) in the classification system of Moffitt *et al.*, who proposed that the expression of exocrine-like marker genes is merely a technical artefact that stems from pancreas or stroma cells rather than from tumour cells. When considering that the transcriptomic data from the Collisson data set was generated by microdissection of human PDAC tissues, the “contamination” of the exocrine-like signature with genes expressed in stromal cells is only conceivable if these marker genes are highly expressed in the pancreas stroma, but not in the tumour cells. This possibility was further investigated through the comparison of transcriptomic profiles of human PDAC cell lines to human wild-type pancreas tissue. In line with the assumption, the expression of exocrine-like genes was strongly upregulated by 13 to 241-fold (median of 183 fold) in human wild-type pancreas tissue when compared to hPDAC cell lines. Strikingly, 15 out of the 19 marker genes of the exocrine-like signature were represented within the top50 genes upregulated in human wild-type pancreas tissue (Figure 17d). Although these findings do not exclude the existence of an exocrine-like subtype of human PDAC, they provide a reasonable explanation that the contamination of microdissected human PDAC tissues with a few healthy pancreas cells in the dataset from

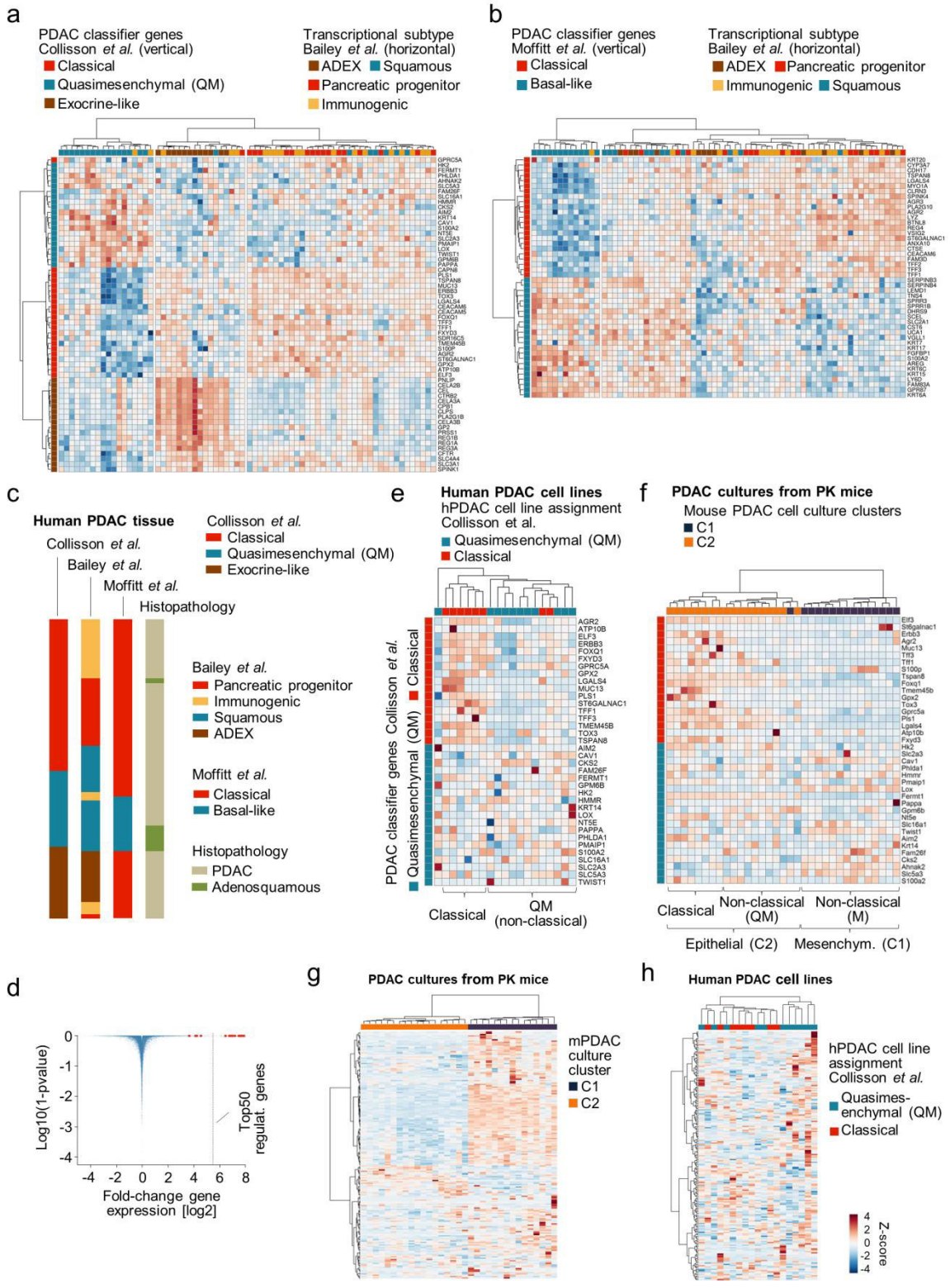

Figure 17 | Transcriptional subtyping of human and mouse PDAC.

[a] Unbiased hierarchical clustering of transcriptome profiles derived from primary pancreatic cancer tissues (n=71) from the study of Bailey *et al.* (Bailey *et al.*, 2016). Transcriptional subtype information according to Bailey *et al.* was available for all samples. Clustering was performed on the basis of PDAC classifier genes defined by Collisson *et al.* (Collisson *et al.*, 2011). On the x-axis three distinct sub-clusters could be identified with: quasimesenchymal and exocrine-like genes generating clusters, which are almost exclusively characterized by Bailey subtypes squamous and ADEX, respectively. Classical Collisson classifier genes formed a third distinct cluster that is

enriched for Bailey subtypes pancreatic progenitor and immunogenic. **[b]** Unbiased hierarchical clustering of transcriptome profiles of primary pancreatic cancer tissues (n=71) from Bailey *et al.* using classifier genes defined by Moffitt *et al.* resulted in perfect separation of classifier genes on the y-axis. Two clusters could be identified on the x-axis with: (i) upregulated expression of Moffitt classical genes enriched for in Bailey subtypes pancreatic progenitor and immunogenic and (ii) increased expression of basal-like genes was associated with squamous Bailey pancreatic cancer subtypes. **[c]** Consensus clustering showing the overlap of the three transcriptome-based classification systems. Assignment of the Bailey subtype, Collisson subtype and Moffitt subtype was extracted for each cancer from the analyses in [a,b]. Samples are sorted according to classification system proposed by Collisson *et al.* **[d]** Volcano plot of differential gene expression in human pancreas tissues as compared to human pancreatic cancer cell lines. Classifier genes of the exocrine transcriptional PDAC subtype from Collisson *et al.* are indicated by enlarged red dots. **[e]** Hierarchical clustering of transcriptome profiles of human pancreatic cancer cell lines (GEO series GSE17891) by using classifier genes proposed by Collisson *et al.* (Collisson *et al.*, 2011). **[f]** Hierarchical clustering of transcriptome profiles of primary PDAC cell cultures derived from PK mice (n=33 from C2a/b/c/C1) by using classifier genes proposed by Collisson *et al.* (Collisson *et al.*, 2011). **[g]** Unbiased hierarchical clustering of transcriptome profiles of primary PDAC cell cultures derived from PK mice (n=33 from C2a/b/c/C1) on the basis of an EMT hallmark gene set (Subramanian *et al.*, 2005). **[h]** Unbiased hierarchical clustering of transcriptome profiles of human pancreatic cancer cell lines (GEO series GSE17891) on the basis of an EMT hallmark gene set (Subramanian *et al.*, 2005). Analysis of transcriptome raw data and clustering of samples was performed by Thomas Engleitner. (The whole figure was modified from Mueller *et al.*, 2018).

Collisson *et al.* can impose an exocrine-like signature on these tumours. Overall, there is a large overlap of at least two human PDAC subtypes that are already covered by the initial subtyping by Collisson *et al.*: (i) classical with pancreatic progenitor/immunogenic and (ii) quasimesenchymal (QM) with basal-like and squamous. As described already by Collisson *et al.*, the expression of classical and QM signature genes did also allow for the stratification of human PDAC cell lines (Figure 17e). Importantly, the most prominent change in the gene expression pattern of QM subtype hPDAC cell lines was down-regulation/extinction of classical signature genes whereas the expression of QM assigner genes was variable and did not show a significant enrichment across both subtypes (Figure 17e). Accordingly, the QM subtype is hereafter termed non-classical (QM). Projection of the Collisson classifier gene list on primary mouse PDAC cell cultures from PK mice did also result in separation of classical-equivalent and non-classical (QM)-equivalent mPDAC. Of note, classical- and non-classical (QM)-equivalent subtypes were both covered within the C2 cluster of cells with epithelial morphology. Mesenchymal cells within C1 formed a third group which has not been described in collections of human PDAC cell lines so far (hereafter referred to as non-classical (M) subtype) (Figure 17f). This observation was further supported by using an EMT hallmark gene set (Subramanian *et al.*, 2005) for unbiased hierarchical clustering of primary mouse PDAC cell culture transcriptomes which resulted in a clear separation of mPDAC with epithelial (C2) and mesenchymal morphology *in vitro* (C1) (Figure 17g). By contrast, applying the same procedure to human PDAC cell lines did not result in a clear clustering of the cell lines, further illustrating the underrepresentation of mesenchymal human PDAC cell lines in

available cell line collections (Figure 17h). As outlined further below in detail, the EMT signature was however enriched in transcriptomes derived from tissues of human undifferentiated pancreatic carcinoma and which was found to be equivalent to the non-classical (M) subtype of mouse PDAC (Figure 19b). Finally, expression of exocrine-like signature genes was also absent in RNA-seq data of primary PDAC cell cultures derived from PK mice further supporting the conclusion that these genes are primarily expressed in pancreas stroma cells.

The pathway analysis of genes upregulated in C1 mPDAC using the DAVID online tool (Subramanian *et al*., 2005) also revealed an significant enrichment of these genes in Ras downstream signalling, such as PI3K-Akt signalling, Rap1 signalling and TGFβ signalling. Notably, this cannot be explained on the basis of genetic dosage of $Kras^{G12D}$ alone since only mPDACs in C2a were $Kras^{G12D\text{-}HET}$. By contrast, mPDACs in C2b, C2c and C1 were all predominantly characterized by having a $Kras^{G12D\text{-}iGD}$ status. To further investigate the upregulation of Ras downstream signalling in C1, the mutant-specific expression of $Kras^{G12D}$ was tested by combining *Kras* qPCR (providing the total expression of mutant plus wild-type *Kras* transcripts) with amplicon-based deep sequencing of *Kras* mRNA (providing the proportions of mutant and wild-type *Kras* transcripts). In line with the observations from the gene set enrichment analysis, expression of mutant $Kras^{G12D}$ increased gradually from C2a to C2b/c and was further elevated significantly in mPDACs of C1 (Figure 16d). Accordingly, the mesenchymal phenotype of C1-clustered mPDAC is defined by the expression of mutant $Kras^{G12D}$ above a certain threshold.

In principle, the mesenchymal phenotype can be cause or consequence of the upregulated expression of $Kras^{G12D}$ in C1. To further analyse the causality between $KRAS^{MUT}$ expression and induction of an EMT program, human PDAC cell lines were transduced using lentivirus that contained an doxycycline inducible $KRAS^{G12D}$ expression cassette. Control cell were transduced with an empty vector containing an inducible GFP expression cassette. Two human PDAC cell lines (HUPT3 and PANC0327) with homozygous loss of *CDKN2A* ($CDKN2A^{\Delta HOM}$) and heterozygous $KRAS^{MUT}$ status ($KRAS^{MUT\text{-}HET}$) were chosen to allow for (i) simulation of increased gene dosage of $KRAS^{MUT}$ and (ii) overexpression of $KRAS^{MUT}$ without the induction of *CDKN2A*-mediated oncogene-induced senescence. Overexpression of $KRAS^{G12D}$ and GFP was induced for 1, 3 and 5 days (Figure 18a). The RNAseq-derived transcriptomes were analysed individually for each specific cell line. All three time points were combined by group ($KRAS^{G12D}$ or GFP) for differential gene expression analysis. Gene set enrichment analysis of genes upregulated in the $KRAS^{G12D}$ versus the GFP group was performed by using the “Molecular Signature

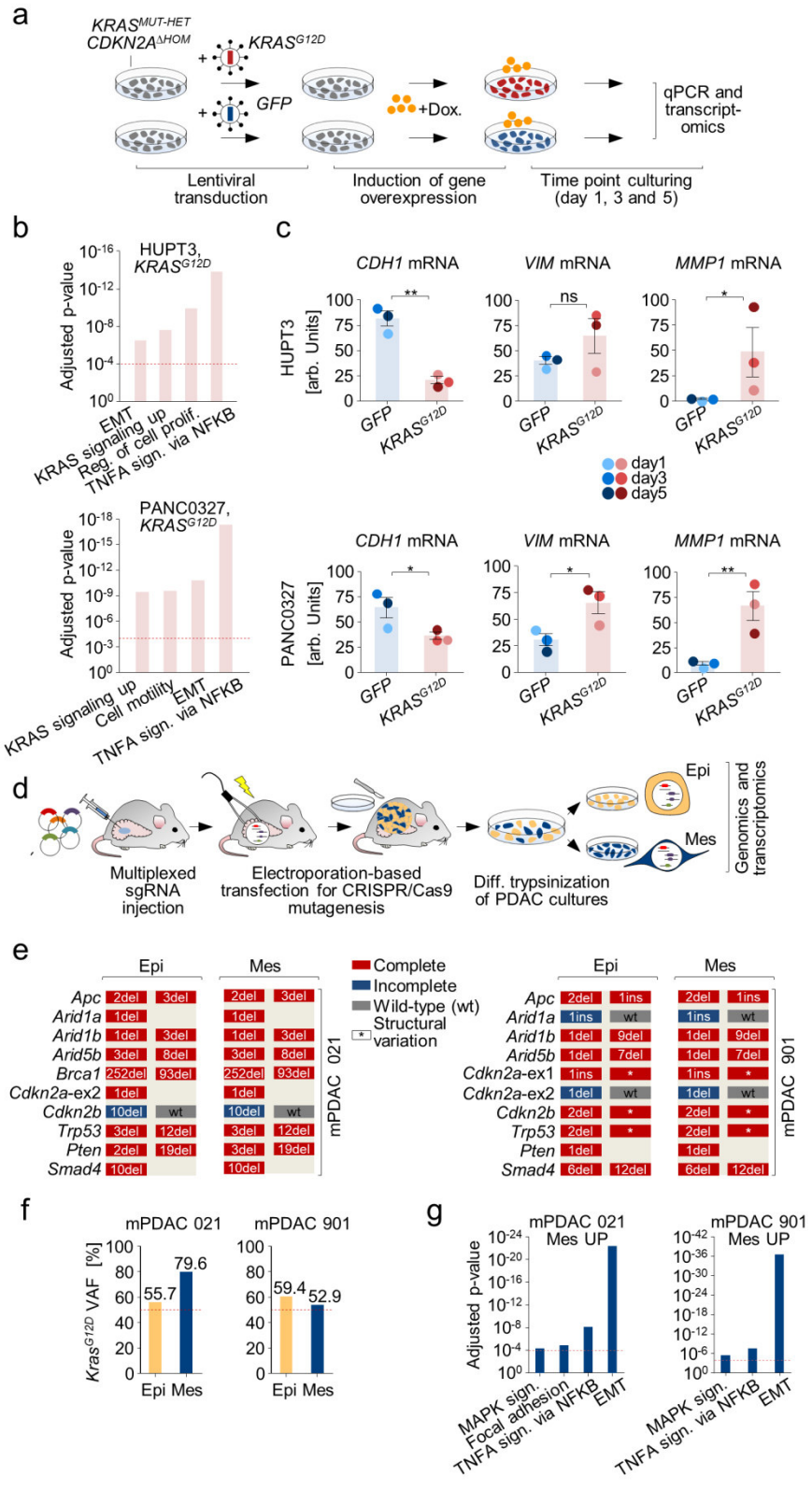

Figure 18 | Functional analyses of *Kras*^{G12D} gene dosage and EMT in human and mouse PDAC.

[a-d] *KRAS*^{G12D} overexpression in human PDAC cell lines induces an EMT-like transcriptional program. **[a]** Graphical illustration of the experimental workflow. Two human pancreatic cancer cell lines (PANC0327 and HUPT3) with heterozygous *KRAS*^{MUT} (*KRAS*$^{MUT-HET}$) status and homozygous deletion of *CDKN2A* (*CDKN2A*$^{\Delta HOM}$) were transduced with lentivirus harbouring a doxycycline-

inducible expression cassette of *KRAS*^{G12D} or GFP-control. Culturing medium containing doxycycline was added to cells for 1, 3 and 5 days to induce expression of *KRAS*^{G12D} or GFP control. **[b]** Gene set enrichment analysis of genes significantly upregulated in human pancreatic cancer cell lines overexpressing *KRAS*^{G12D} as compared to control cell lines overexpressing GFP. The graphs show selected gene sets for each human pancreatic cancer cell line. Y-axis shows false discovery rate-adjusted P values. **[c]** Validation of *KRAS*^{G12D}-induced EMT-like transcriptional programs. Expression levels of marker gens for epithelial (*CDH1*) or mesenchymal (*VIM1*) cellular differentiation as well as for matrix disassembly/invasion (*MMP1*) were quantified by qPCR. Gene expression was normalized to the expression levels of housekeeper genes *PPIA* and *GAPDH*. *P ≤ 0.05, **P ≤ 0.005, ns=not significant, two-tailed t-test; bars=mean; error bars=SEM. **[d-g]** Application of transfection-based CRISPR-Cas9 delivery in mice for multiplexed pancreatic genome engineering and phylogenetic tracking of epithelial and mesenchymal cancer clones *in vivo*. **[d]** Graphical illustration of principle steps of the experimental work flow. Multiple CRISPR/Cas9 vectors, each targeting a distinct tumour suppressor gene, were delivered to cells of the murine pancreas *in vivo* by electroporation-based transfection as previously described by Maresch *et al.* (Maresch *et al.*, 2016). This approach allowed for the delivery of multiple CRISPR/Cas9 vectors to only a few pancreatic cells (average of ~120 cells per pancreas) and induction of clonal tumours. Primary PDAC cell cultures generated from this mouse model were screened for the simultaneous appearance of epithelial and mesenchymal cancer cells. Both cellular morphologies were observed in mPDAC cell cultures derived from mouse 021 and 901. Finally, each morphology was enriched for by differential trypsinization. **[e]** Analysis of CRISPR/Cas9-mediated indel formation at all sgRNA-target sites by amplicon-based deep sequencing in epithelial and mesenchymal primary mPDAC cell cultures 021 and 901. **[f]** Variant allele frequency (VAF) of *Kras*^{G12D} in epithelial and mesenchymal mPDAC cell cultures 021 and 901 as detected by amplicon-based deep sequencing of the *Kras* locus. **[g]** Gene set enrichment analysis of genes significantly upregulated in mesenchymal versus epithelial primary mPDAC cell cultures 021 or 901. The graphs show selected gene sets for mPDAC cell cultures of each mouse. Y-axis shows false discovery rate-adjusted P values. Raw data of amplicon-based deep sequencing of CRISPR/Cas9 target sites and RNA-Seq was analysed by Thomas Engleitner. Transfection-based delivery of CRISPR/Cas9 components to the murine pancreas *in vivo* was conducted by Roman Maresch and myself. (The whole figure was modified from Mueller *et al.*, 2018).

Database" (MSigDB) (Subramanian *et al.*, 2005). Upon induction of *KRAS*^{G12D} expression, both human PDAC cell lines, HUPT3 and PANC0327, showed an upregulation of transcription for genes involved in the hallmark gene set "*KRAS* signalling up" and "EMT" (epithelial-to-mesenchymal-transition) (Figure 18b). These findings were further validated by testing the expression of EMT marker genes Vimentin (*VIM*) and E-cadherin (*CDH1*). Both genes are widely used as markers for epithelial (*CDH1*) or mesenchymal (*VIM*) cell differentiation/morphology (Nieto *et al.*, 2016). qPCR of *VIM* and *CDH1* mRNA levels confirmed the upregulation of the mesenchymal and downregulation of the epithelial marker gene transcripts further supporting the finding that overexpression of *KRAS*^{G12D} induced an EMT-like transcriptional program in both human PDAC cell lines (Figure 18c).

To further substantiate the finding that pancreatic cancer cells can undergo an epithelial-to-mesenchymal-transition upon upregulation of *KRAS*^{MUT} expression, the emergence of epithelial and mesenchymal clones within the same tumour was tracked *in vivo* using indel patterns introduced by CRISPR/Cas9. The method of transfection-based, multiplexed CRISPR/Cas9 somatic mutagenesis in the pancreata of PK mice *in vivo* was already previously established in our group (Maresch *et al.*, 2016). This approach allows for: (i)

low-frequency/mosaic CRIPSR/Cas9 delivery to few cells in order to induce tumours arising from a single clone (ii) inactivation of multiple tumour suppressor genes in individual cells and (iii) exploitation of tumour suppressor gene indel patterns for phylogenetic tracking of tumour clones and metastases. From each mPDAC that was induced by CRISPR/Cas9, multiple cancer cell cultures were generated and monitored for the simultaneous appearance of epithelial and mesenchymal cell morphologies. Differential trypsinization was used to enrich for either morphology (Figure 18d). Such mPDAC were identified in PK mouse 021 and 901 (Figure 16e). All tumour loci targeted by CRISPR/Cas9 were amplified by PCR and subjected to amplicon-based deep sequencing. This analysis revealed that indel patterns were identical in epithelial/mesenchymal cell culture pairs derived from PK mouse 021 or 901, providing evidence for that: (i) epithelial and mesenchymal cell morphologies can arise from the same tumour clone in vivo and (ii) the tumour suppressor genes inactivated by CRISPR/Cas9 are not contributing to the different phenotypes (Figure 18e). Importantly however, mesenchymal cells in mPDAC 021 showed increased dosage of $Kras^{G12D}$, upregulation of $Kras^{G12D}$-specific mRNA levels and intensification of Ras-related downstream signalling pathways when compared to the epithelial cells of the same primary mPDAC. In mPDAC 901 the gene dosage of epithelial and mesenchymal cell clones did not vary with both clones being $Kras^{G12D\text{-}HET}$. However, the expression of mutant-specific $Kras^{G12D}$ transcripts was increased in mesenchymal cells as compared to epithelial cells further supporting the role of increased $Kras^{G12D}$ dosage in defining the cellular phenotype, even when all major pancreatic cancer tumour suppressor pathways were inactivated through CRISPR/Cas9 (Figure 16e, 18f,g).

PDAC from humans and mice can be scored histologically on the basis of their cellular atypia, cellular differentiation, stroma invasion, necrotic areas and other additional criteria. In the next step, the histological tumour grade of the mPDAC from PK was annotated to the clustering of the transcriptomes to correlate the histological *in vivo* phenotype of the tumour cells with (i) the defined expression clusters and (ii) epithelial/mesenchymal cell morphology *in vitro*. The integration of the mPDAC histology revealed that the histological grade scores of the tumours increased from C2a to C2b/c to C1. mPDAC from expression cluster C2a showed a well or moderately differentiated histology (also referred to as G1 or G2) whereas mPDAC from C1 displayed almost exclusively undifferentiated histology with spindle shaped tumour cells and sparsely distributed areas with ductal differentiation (Figure 16f). Dedifferentiation can occur during disease progression or selection of drug resistant clones during tumour therapy and is associated with a shortened patient survival as compared to pancreatic tumours with ductal differentiation (Winter *et al.*, 2008; Iacobuzio-Donahue *et al.*, 2009). This observation was also reflected in the PK cohort

where mice with undifferentiated pancreatic carcinomas showed the tendency towards reduced survival when compared to mice with well, moderately or poorly differentiated (G1, G2 or G3) mPDAC (Figure 19a). The strong enrichment of undifferentiated tumours within the C1 expression cluster revealed that (i) *in vivo* and *in vitro* phenotypes of the tumour cells were strongly correlated and (ii) that this histological PDAC subtype was characterized by the highest levels of *Kras*^{G12D} expression and increased activation of Ras downstream signalling pathways providing and explanation for the aggressive behaviour of undifferentiated pancreatic carcinomas.

Human pancreatic carcinoma with undifferentiated histology are underrepresented in publically available human surgical series or cell line collections (1-3% of all cases),

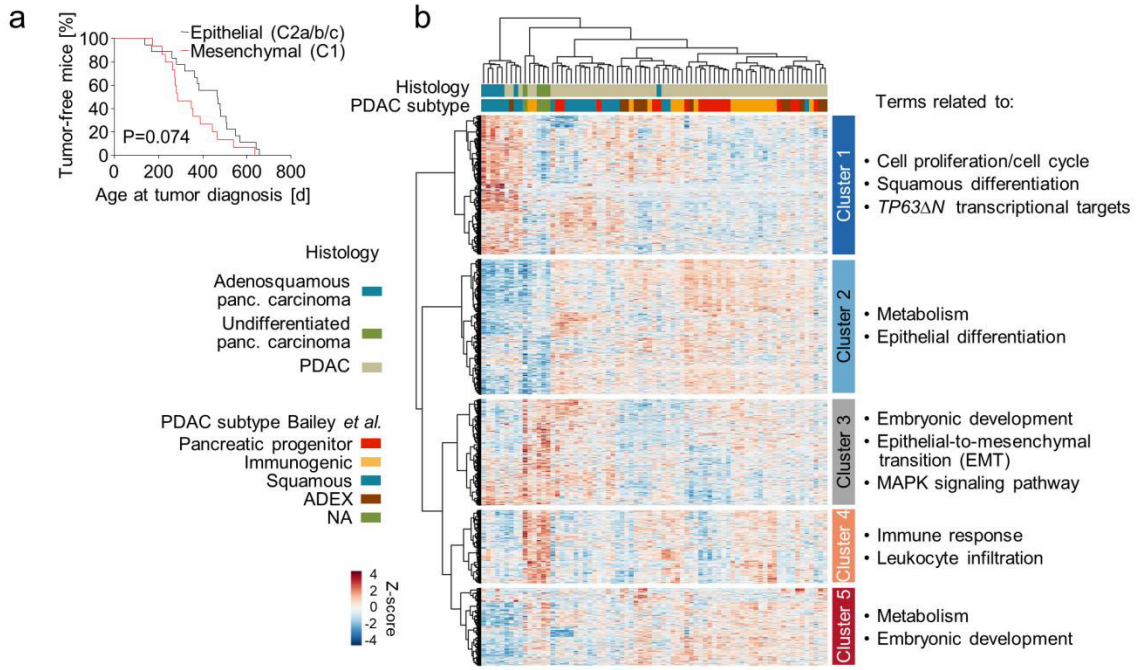

Figure 19 | Undifferentiated human PDAC show upregulation of Ras downstream signalling and transcriptional programs related to EMT.

[a] Undifferentiated/mesenchymal pancreatic carcinomas from PK mice (n=15, C1) show a tendency towards reduced survival when compared to epithelial mPDAC with histopathological grade 1–3 (G1–G3) (n=18, C2a/b/c). P-value calculated by log-rank test. **[b]** Analysis of expression profiles from human PDAC tissues with World Health Organization (WHO) grade 1 to 3 (n=64), adenosquamous pancreatic carcinoma (n=7) and human undifferentiated pancreatic carcinoma (n=4). Samples matching the above mentioned histopathological criteria and with available transcriptional subtype information according to Bailey *et al.* were supplemented with expression profiles of undifferentiated pancreatic carcinomas from the Australian pancreatic cancer cohort of the International Cancer Genome Consortium (ICGC PACA-AU). Undifferentiated pancreatic carcinoma cluster together and show upregulated expression of genes in Cluster 3 and 4, but not of Cluster 1 genes which are specifically upregulated in adenosquamous pancreatic carcinoma. Analysis of transcriptome raw data and clustering of samples was performed by Thomas Engleitner. (The whole figure was modified from Mueller *et al.*, 2018).

probably due to the typically advanced stage of these tumours at diagnosis (Morohoshi *et al.*, 1983). Accordingly, the frequency of human PDAC with focal areas of undifferentiated pancreatic carcinoma increases to up to 16% of cases in autopsy series (Iacobuzio-Donahue *et al.*, 2009). To compare undifferentiated tumours from PK mice with human undifferentiated carcinoma, publically available transcriptome datasets were screened for pancreatic carcinomas with undifferentiated histological grading. Four such pancreatic cancers were identified in the Australian pancreatic cancer dataset of the International Cancer Genome Consortium (PACA-AU cohort of the ICGC). Subtyping information according to the publication of Bailey *et al.* was also available. Human PDAC with WHO grade 1 to 3 (corresponding to G1 to G3, n=64), adenosquamous pancreatic carcinoma (n=7) and undifferentiated pancreatic carcinoma (n=4) were included into the analysis of expression profiles (total of n=75 expression profiles). Other histological subtypes of pancreatic cancer, such as acinar cell carcinoma, IPMN or MCN, were excluded. This dataset was further sub stratified into defined groups of pancreatic cancer: (i-iv) PDAC with WHO grade 1 to 3 subdivided into pancreatic progenitor, immunogenic, squamous and ADEX subtypes, (v) adenosquamous pancreatic carcinoma and (vi) undifferentiated pancreatic carcinoma. ANOVA was performed to identify genes which showed significantly altered gene expression in at least one of the six defined groups. The list of differentially regulated genes was in turn used for unbiased hierarchical clustering of the 75 pancreatic cancer transcriptomic profiles which resulted in five clusters of co-regulated gene expression (Figure 19b). The list of genes contributing to each cluster of co-regulated gene expression was analysed for enriched gene sets using the “Molecular Signature Database” (MSigDB) (Subramanian *et al.*, 2005). Significantly enriched gene sets were evaluated for gene sets/pathways that could be condensed into a few predominating molecular categories/processes. Importantly, the four samples of undifferentiated pancreatic carcinoma clustered together in the cluster tree indicating that their transcriptomes were distinct from the other pancreatic cancer subtypes. Undifferentiated pancreatic carcinomas showed upregulation of genes in cluster 3 of co-regulated gene expression and downregulation of genes in clusters 2 and 5. Cluster 3 was enriched for genes sets of the MAPK signalling pathway and gene sets related to EMT and embryonic development, whereas cluster 2 and 5 were characterized by gene sets of epithelial cell differentiation or gene sets of embryonic development and metabolic signatures, respectively (Figure 19b). Similar patterns of gene set enrichment were also observed in undifferentiated PDAC from PK mice further supporting the link of upregulated Ras-related downstream signalling pathways with EMT and undifferentiated tumour histology. The expression of cluster 4 genes was upregulated in PDAC of the immunogenic Bailey subtype and showed a strong enrichment for gene sets associated with processes of immune response and leukocyte

infiltration. Of note, the expression of cluster 4 genes was also strongly upregulated in undifferentiated pancreatic carcinomas indicating an increased rate of immune cell infiltration in this subtype of pancreatic cancer. Cluster 1 of co-regulated gene expression contained almost all pancreatic carcinoma samples with adenosquamous histology and was enriched for gene sets related to the cell cycle/proliferation, squamous cell differentiation and activation of *TP63ΔN* transcriptional targets. This pattern of gene set enrichment was previously reported by Bailey *et al.* for the squamous PDAC subtype (Bailey *et al.*, 2016). Notably, in our analysis the upregulated expression of *TP63ΔN* transcriptional target genes was predominantly driven by adenosquamous pancreatic carcinomas which were only a subset of all samples in the squamous PDAC subtype. In contrast to the adenosquamous pancreatic carcinomas, the expression of cluster 1 genes was downregulated in undifferentiated pancreatic carcinomas suggesting that activation of *TP63ΔN* transcriptional network in pancreatic carcinoma with adenosquamous histology is not linked to EMT or increased dosages of $KRAS^{MUT}$ (Figure 19b).

3.6 Biological impact of $Kras^{G12D}$ gene dosage in a mouse model with high mutational load

The lower mutational load observed primary cell cultures derived from $Kras^{G12D}$-driven mouse models of pancreatic cancer, was exploited to discover new principles in PDAC evolution and phenotypic diversification. However, the lack of additional genetic alterations may also influence the relevance of the conclusions made in mouse tumours when cross-compared to the human setting. To account for the decreased mutational burden in mouse PDAC, the PK mouse model was inter-crossed with mice carrying components of the *PiggyBac* transposon-based insertional mutagenesis system (Figure 20a). The *PiggyBac* transposase is required for the activation and mobilization/transposition of the genetically engineered ATP1 transposon (Rad *et al.*, 2010). Insertion of the ATP1 transposon into gene-encoding regions of the mouse genome allows for activation or inactivation of that gene, depending on the position and orientation of the transposon insertion. Thereby, transposon-based insertional mutagenesis enables for the genome-wide screening of cancer genes in an *in vivo* setting. PK-PB mice were generated in a previous study by Rad *et al.* and showed reduced survival due to transposon mutagenesis and accelerated pancreatic tumour formation when compared to PK mice (Rad *et al.*, 2015). Seventeen primary PDAC cell cultures derived from PK-PB mice of that study were subjected to QiSeq which allows for genome-wide transposon screening and quantitative sequencing of

transposon insertion sites (Friedrich *et al.*, 2017). Primary mPDACs from PK-PB mice showed extensive mutational burden with a median of 494 transposon insertion sites per tumour as identified by QiSeq. The reduced survival of PK-PB mice also show that these transposon insertions are accelerating pancreatic cancer evolution and are therefore biologically relevant. To test if the findings in primary PDAC from PK mice are also relevant in the setting of high mutational burden, primary tumour cell cultures from PK-PB mice were subjected to aCGH, amplicon-based deep sequencing of the *Kras* locus and microarray-based gene expression profiling, in addition to QiSeq. First, the transcriptomic profiles of the pancreatic cancers were used for unbiased hierarchical clustering. Like in

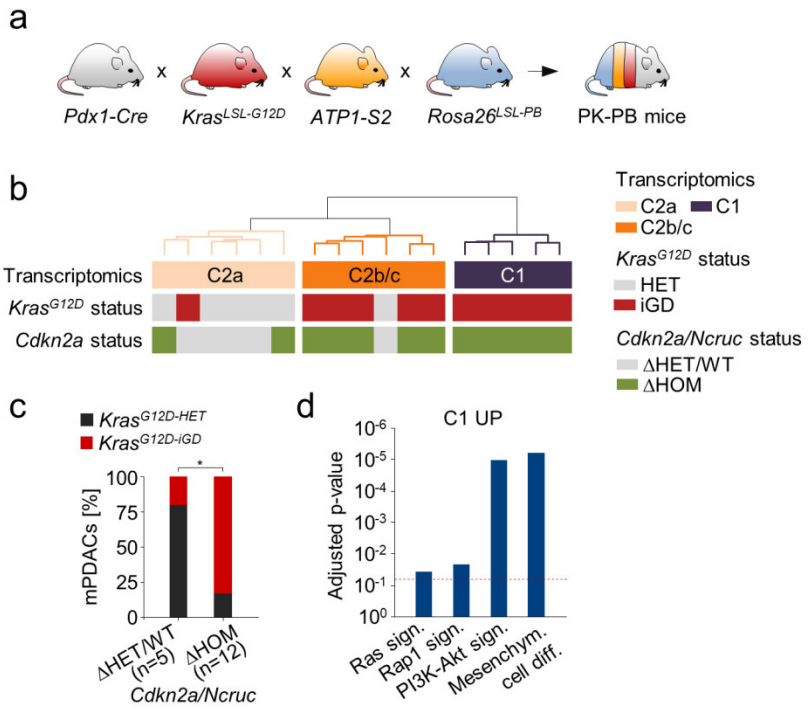

Figure 20 | Gene dosage of *Kras*^{G12D} defines PDAC biology in a pancreatic cancer mouse model with high mutation load.

[a] Mouse lines used for the generation of a PDAC mouse model with high mutational load (PK-PB mice). **[b]** Unbiased hierarchical clustering of array-based expression profiles of primary mPDAC cell cultures derived from PK-PB mice (n=17) results in two major clusters C1 and C2 (with sub-clusters C2a and C2b/c). Genetic status of *Kras*^{G12D} and *Cdkn2a* is annotated for each mPDAC as defined by the combined analysis of amplicon-based deep sequencing of the *Kras* locus, aCGH and QiSeq data. *Kras*^{G12D} gene dosage was significantly associated with expression clusters (P=0.01, two-sided Fisher's exact test). **[c]** Gene dosage of *Kras*^{G12D} in primary mPDAC cell cultures of PK-PB mice with *Cdkn2a*/*Ncruc*$^{\Delta HET/WT}$ versus *Cdkn2a*/*Ncruc*$^{\Delta HOM}$ status. *P=0.03, Fisher's exact test, OR 20.0, 95% CI 1.4-287.8. **[d]** Gene set enrichment analysis of genes significantly upregulated in mesenchymal (C1, n=5) versus epithelial (C2, n=12) primary mPDAC cell cultures of PK-PB mice. Selected gene sets upregulated in C1-clustered mPDACs are shown in the graph. Y-axis shows false discovery rate-adjusted P values. Analysis of QiSeq and amplicon-based deep sequencing raw data as well as analysis of expression profiles and clustering of samples was performed by Thomas Engleitner. (The whole figure was modified from Mueller *et al.*, 2018).

primary mPDAC of PK mice, two major clusters could be observed (termed C1 and C2 in analogy to PK-derived mPDAC) (Figure 20b). Again, cluster C2 could be further stratified into sub-clusters: one sub-cluster that was characterized by mPDAC with *Kras*$^{G12D\text{-}HET}$ and *Cdkn2a*$^{\Delta HET/WT}$ status (most likely resembling cluster C2a in PK mice) and another cluster with mPDAC predominantly having *Kras*$^{G12D\text{-}iGD}$ and *Cdkn2a*$^{\Delta HOM}$ status (resembling cluster C2b/c in PK mice). *Kras*^{G12D} status was analysed by the combined interpretation of aCGH and amplicon-based deep sequencing of the *Kras* locus whereas the status of the *Cdkn2a* locus was determined by combining aCGH data with quantitative transposon insertion site sequencing (QiSeq). Statistical testing revealed that the *Kras*^{G12D} status was significantly associated with expression clusters. As in PK mice, there was a significant association of homozygous inactivation of *Cdkn2a* (*Cdkn2a*$^{\Delta HOM}$) with increased gene dosage of *Kras*^{G12D} (*Kras*$^{G12D\text{-}iGD}$) in primary mPDAC cell cultures from PK-PB mice. Accordingly, *Cdkn2a* is also an important barrier for *Kras*^{G12D} gene dosage variation in PK-PB mice (Figure 20c). In a last step, gene set enrichment analysis using the DAVID tool was performed for genes differentially regulated between C1- and C2-clustered primary mPDAC from PK-PB mice. Again, this analysis revealed striking similarities to mPDAC expression clusters of PK mice. C1-clustered PDAC from PK-PB mice showed a strong enrichment for gene sets related to mesenchymal differentiation as well as an upregulation of Ras-related downstream signalling pathways (Figure 20d). Overall, the analyses of primary PDAC cell cultures derived from PK-PB mice shows that the findings made in PK mice are equally valid in a pancreatic cancer model with high mutational burden.

3.7 Conclusion

A major challenge of therapeutic targeting of PDAC is our limited molecular understanding the disease. The lower mutational burden and the easier interpretability of mouse PDAC genomes were exploited for the discovery of principles defining pancreatic cancer evolution and phenotypic diversification. The dosage variation of oncogenes and tumour suppressor genes was identified as a critical determinant of PDAC biology. The three major PDAC tumour suppressor pathways (*CDKN2A*, *TP53* and TGFβ) were found to differentially license dosage variation of oncogenes along distinct evolutionary routes, thereby controlling critical disease characteristics including: early progression,

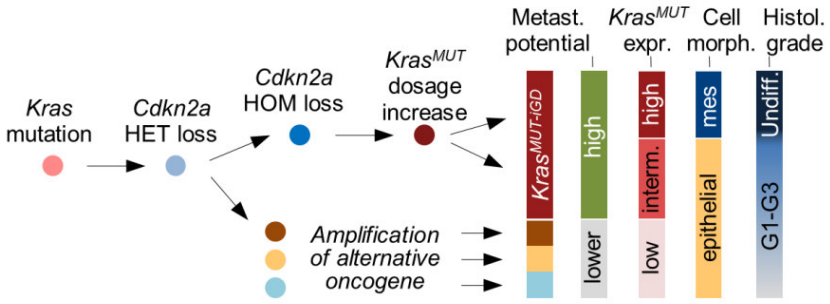

Figure 21 | Simplified conceptual framework for the evolution of molecular, morphological and clinical disease characteristics in PDAC.

Oncogenic gene dosages are critical determinants of PDAC biology and develop along distinct evolutionary trajectories licensed by different types (*Cdkn2a*, *Trp53*, *Tgfbr2*) and states (wildtype/heterozygous, homozygous) of tumour suppressor gene inactivation. For simplicity, only the interaction of $Kras^{G12D}$ and *Cdkn2a* gene dosage during tumour evolution is shown. *Trp53* loss, which can also promote $Kras^{MUT}$ gene dosage increase, is not shown. Likewise, the effect of *Tgfbr2* inactivation, which supports evolution predominantly through $Cdkn2a^{HET}/Kras^{MUT\text{-}HET}$ trajectories, is not visualized. Of note, depicted trajectories are not completely exclusive, e.g. *Myc* or *Nfkb2* amplification can cooperate with both, amplified and non-amplified $Kras^{MUT}$ to drive pancreatic carcinogenesis (see also Figure 10h). Major aspects of biological and clinical PDAC phenotypes are linked to the evolution of oncogenic dosage along different trajectories. (Figure modified from Mueller *et al.*, 2018).

histopathology, metastasis as well as cellular plasticity. Gene dosage gain of $KRAS^{MUT}$ was found to be a genetic hallmark of early disease progression in PanIN. *MYC*, *YAP1* and *NFKB2* were identified as "alternative" oncogenes that can drive early PDAC progression in $KRAS^{MUT}$ non-amplified contexts. The type and nature of oncogene amplification is dictated by historical constraints and contingencies that direct PDAC evolution through distinct trajectories: homozygous loss of *CDKN2A* or *TP53* allows for acquisition of a $Kras^{G12D\text{-}iGD}$ status whereas gene dosage increase of alternative oncogenes is feasible in genetic background with heterozygous inactivation of *CDKN2A* (referred to as context-dependent haploinsufficiency of *CDKN2A* such as in the context of

TGFβ pathway alterations). These evolutionary trajectories and further modulation of *KRAS*$^{MUT\text{-}iGD}$ expression also identified critical determinants of clinical PDAC behaviour including: *KRAS*$^{MUT\text{-}iGD}$ (high metastatic potential), chromothripsis (repid progression) and highest *KRAS*^{MUT} expression (aggressive, undifferentiated pancreatic cancers with reduced survival). Overall, this study provides a new conceptual framework for the understanding of PDAC evolution and phenotypic diversification on the basis of three genetic hallmarks: (i) gene mutation, (ii) gene dosage and (iii) evolutionary trajectories (Figure 21).

4 Discussion

4.1 A new conceptual framework for the understanding of pancreatic cancer evolution and phenotypic diversification

Pancreatic cancer is the fourth leading cause of cancer-related mortality in Western World and is predicted to become the second by 2030 (Rahib *et al.*, 2014). Pancreatic ductal adenocarcinoma (PDAC), commonly referred to as pancreatic cancer, accounts for 95% of all pancreatic tumours. While new treatment options have improved the prognosis and outcome of many other cancer types, the 5-year survival rate of patients with PDAC stayed around 5% in the last 30 years (Rahib *et al.*, 2014). The very poor survival rate of PDAC patients can be related to two key aspects: (i) treatment response rates of PDAC patients are low due to the lack of effective treatment options and (ii) PDAC is typically diagnosed late when the disease is already locally invasive or metastasised to distant organs due to the lack of early detection methods.

The current model of pancreatic cancer evolution considers pancreatic intraepithelial neoplasms (PanINs) as the most common precursor lesions of pancreatic cancer. PanINs are the best described precursor lesions of pancreatic cancer and are thought to progress through a stepwise accumulation of activating or inactivating alterations in *KRAS*, *CDKN2A*, *TP53* and genes involved in the canonical TGFβ signalling pathway. Activation of *KRAS* through somatic mutation is the first and an almost universal event during early PanIN progression. As PanINs progress to invasive pancreatic cancer, they acquire extensive intra- as well as intertumoural mutational heterogeneity with low frequency of recurrently altered genes. Through the revolution in next generation sequencing (NGS), the genetic characterization of more than 500 pancreatic cancer exomes/genomes revealed that there are only a few pancreatic cancer signature mutations. These mutations occur in *KRAS*, *CDKN2A*, *TP53* and genes of the canonical TGFβ pathway, each genetically altered in more than 50% of PDAC patients, and were already well known from the genetic study of PanIN progression (Waddell *et al.*, 2015). Beyond these signature mutations, the frequency of recurrent gene alterations in PDAC drops below 10% with a long tail of genes involved in DNA damage repair, chromatin remodelling, axon guidance pathway or carcinogenesis in general (Campbell *et al.*, 2010; Biankin *et al.*, 2012; Waddell *et al.*, 2015). Of note, only few of the many heterogenic genetic alterations could be attributed to certain phenotypes of pancreatic cancer so far. For instance (i) genetically unstable human PDACs are significantly associated with homozygous inactivating

mutations of *BRCA1/2* or *PALB2* and can be targeted by platinum-based therapy (Waddell *et al.*, 2015) and (ii) *MYC* amplifications, inactivating mutations of *KDM6A* or upregulated expression of the TP63ΔN transcriptional network are associated with squamous differentiation of human PDAC (Witkiewicz *et al.*, 2015; Bailey *et al.*, 2016). However, it is still not possible to broadly link these heterogenic gene or pathway alterations to clinical, morphological or biological characteristics of the disease. For instance, PDAC is a disease which is very well known for its aggressive growth as well as for its early and frequent metastasis to distant organs, typically to the lymph nodes, liver or lung. The metastatic process has been linked to the epigenetic and metabolic reprogramming of metastatic pancreatic cancer cells (McDonald *et al.*, 2017; Roe *et al.*, 2017). Strikingly however, no frequently mutated metastasis driver gene has been identified yet (Makohon-Moore *et al.*, 2017) obscuring the genetic understanding of metastatic tumour cell dissemination. Accordingly, although significant progress has been made in the genetic characterization of pancreatic cancer during the last years, the understanding of PDAC evolution and phenotypic diversification is still limited with many conceptional questions that are yet to be answered, e.g.: (i) How can PanIN-to-PDAC progression be explained from a genetic point of view, e.g. in a setting where all four pancreatic cancer signature mutations have been already acquired? (ii) What is, or is there any, genetic mechanism/alteration that drives metastatic dissemination? (iii) Are early progression and early/frequent metastatic dissemination connected/linked through the same molecular events? (iv) What is the molecular basis of cellular plasticity and different histopathological PDAC grading? and (v) Why is clinical aggressiveness a characteristic feature of undifferentiated pancreatic carcinomas?

In the work at hand, the traditional way to understand cancer phenotypes on the basis of “mutations” through the characterization of pancreatic cancer genomes is being extended by two additional cornerstones: “dosage variation of oncogenic mutations” and their “evolution along distinct trajectories”, controlled by different types and states of tumour suppressor alterations (Figure 21). This integrative “3-dimensional view” provides an improved framework of the molecular and mechanistic understanding of different aspects of pancreatic cancer biology, including early progression, metastasis, cellular plasticity, histopathology and clinical aggressiveness. The increased gene dosage of oncogenic *KRAS* (*KRAS*$^{MUT\text{-}iGD}$) is central to all the aforementioned processes and will be discussed in detail below.

4.2 *KRAS*$^{MUT\text{-}iGD}$ in primary pancreatic cancer progression

Chromosomal alterations, aneuploidy and genetic instability are genetic hallmark phenotypes of cancer cells that distinguish them from healthy cells. Although this observation is already more than 100 years old (reviewed in Schvartzman *et al*., 2010), the role of these genetic alterations as a driver or passenger of human tumorigenesis is still controversial in cancer research. It is now clear, that structural or segmental rearrangements can drive tumorigenesis through the activation of oncogenes, such as *IGH-MYC* translocations in Burkitt lymphoma (Mitelman *et al*., 2007), or inactivation of tumour suppressor genes, such as *CDKN2A* and *SMAD4* in pancreatic cancer (Waddell *et al*., 2015). However, the role of whole arm or whole chromosome somatic CNAs during tumour evolution is much less well understood. Genomic regions affected by these alterations were found to be recurrent across different types of human cancers and had a preference for either genetic amplification or loss (but not both at the same time), suggesting that they were selected during tumour progression (Beroukhim *et al*., 2010). The general difficulty of investigating the functional role of these alterations is that they affect larger regions of the cancer cell genome, complicating the identification of the underlying genes and associated molecular mechanisms. Here in this study, it could be shown that large-sized chromosomal alterations (aneuploidy) play a critical role in driving pancreatic cancer progression and phenotypic diversification when altering the oncogenic dosage of mutant/oncogenic *KRAS*. In a *Kras*^{G12D}-driven mouse model of pancreatic cancer approximately two thirds of all primary tumour cell cultures showed increased gene dosage of the mutant *Kras*^{G12D} locus either through (i) arm-level gain (23.7% of all mPDAC), (ii) focal amplification (7.9% of all mPDAC) or copy-number-neutral loss of the *Kras* wild-type allele (36.8% of all mPDAC). The gene dosage increase always involved the mutant *Kras*^{G12D} allele (*Kras*$^{G12D\text{-}iGD}$) and also affected the transcriptional output of the *Kras* locus by increasing *Kras*^{G12D}-specifc mRNA levels in *Kras*$^{G12D\text{-}iGD}$ as compared to *Kras*$^{G12D\text{-}HET}$ mPDAC. The specific gene dosage increase of the mutant *Kras*^{G12D} allele indicates that there is selection for oncogenic signalling intensification during pancreatic cancer progression. Twelve primary pancreatic cancer cell cultures (31.6% of all mPDAC) did not increase the dosage of *Kras*^{G12D} and remained in a state with *Kras*$^{G12D\text{-}HET}$ status. Interestingly, these mPDAC were enriched for amplification of other oncogenes such as *Myc*, *Yap1* or *Nfkb2* indicating that oncogenic signalling is increased in these tumours through the amplification of alternative oncogenes that are cooperating with *Kras*$^{G12D\text{-}HET}$ to drive pancreatic cancer progression in this genetic background. The idea of Ras/Raf signalling amplification during pancreatic cancer progression is supported by other studies

in the context of mammary, intestinal or lung tumorigenesis (Sarkisian *et al.*, 2007; Murphy *et al.*, 2008; Feldser *et al.*, 2010; Junttila *et al.*, 2010; Rad *et al.*, 2013). In mammary tumorigenesis, a titratable *Hras*^{G12V}-driven mouse model was used to show that low expression levels of *Hras*^{G12V} (comparable to physiological *Hras* levels found in normal mouse tissues) induce proliferation and hyperplasia of mammary epithelial cells. Interestingly, chronic induction of low *Hras*^{G12V} levels induced mammary tumours only after spontaneous upregulation of *Hras*^{G12V} expression (Sarkisian *et al.*, 2007). In a *Braf*^{V600E}-driven mouse model of colorectal cancer, physiological levels of *Braf*^{V600E} expression were sufficient to induce intestinal hyperplasia. Progression of dysplastic lesions to invasive colorectal cancer were accompanied by stage-specific activation of the Wnt pathway and intensification/hyperactivation of the Braf/Mek/Erk signalling pathway. These findings are also in line with observations in a *Kras*^{G12D}-driven mouse model of lung cancer. Low levels of oncogenic *Kras*^{G12D} signalling were sufficient to drive the development of early lung tumour stages, such as lung adenoma, whereas amplification of MAPK signalling was strongly associated with the progression to malignant lung adenocarcinoma (Feldser *et al.*, 2010; Junttila *et al.*, 2010). Although, these studies support the importance of oncogenic signalling intensification during pancreatic cancer progression it is not clear how *Kras*$^{G12D\text{-}iGD}$ is influencing the signalling output of Ras-mediated downstream pathways at the molecular level. A study by Zhang *et al.* (Zhang *et al.*, 2001) observed that mice with one functional copy of *Kras* were much more susceptible to the chemical induction of lung tumorigenesis in comparison to mice with two functional copies. In this model, lung tumours typically acquire activating mutations in one copy of *Kras*. Such as in the *Kras*^{G12D}-driven pancreatic cancer mouse model, chemically induced lung tumours of mice with two functional *Kras* copies lost the wild-type allele in more than two thirds of cases during progression to lung tumours. Moreover, the expression of wild-type *Kras* was inversely correlated to the activity of the MAPK pathway. The authors of the study concluded that wild-type *Kras* acts as a tumour suppressor during lung tumour progression through reducing signalling output of the oncogenic/mutated *Kras* allele and explaining the increased lung cancer susceptibility of mice with only on functional Kras allele (Zhang *et al.*, 2001).

Through the micro-dissection of low-grade PanINs, it could be shown that the gene dosage of mutant *KRAS* gets increased already very early during progression to PDAC. More than one third of all *KRAS*-mutant low-grade PanIN lesions showed *KRAS*$^{MUT\text{-}iGD}$, which is probably even underestimated due to the imperfect removal of contaminating stromal cells by micro-dissection. The high frequency of *KRAS*$^{MUT\text{-}iGD}$ in very early pre-malignant lesions

further supports the model that the intensification/over-activation of Ras effector pathways is of critical importance for the progression to PDAC.

More recent studies further investigated the molecular basis of the phenomenon that *Kras*^{WT} can act as a tumour suppressor in more detail. Nan *et al.* found that instead of the existing view of Ras signalling through GTP-bound monomers, activation of Ras effector pathways is dependent on the dimerization of GTP-bound Ras at the cellular membrane. At physiological expression levels, *Kras*^{G12D} formed dimers that activated MAPK signalling, while at low expression levels *Kras*^{G12D} remained monomeric and was only able to activate the Ras effector pathway when artificially dimerized (Nan *et al.*, 2015). Another study by Ambrogio *et al.* is further supporting these findings. When introducing a D154Q mutation into *Kras*, dimerization of *Kras* monomers could be abrogated without interfering with other fundamental biochemical properties of *Kras*. In human *KRAS*-mutant lung adenocarcinoma cell lines, the tumour suppressive function of wild-type *KRAS* could be abrogated when replaced with *KRAS*^{D154Q}. Strikingly, the oncogenic activity of various *KRAS* mutants (*KRAS*$^{G12C/D/V}$) was also abrogated *in vivo* when the D154Q mutation was introduced to create a double mutant *Kras*$^{G12/D154Q}$ variant (Ambrogio *et al.*, 2018). Overall, these studies highlight a critical role of *Kras* dimerization for the activation of Ras effector pathways which might be also relevant for the intensification of oncogenic *Kras*^{G12D} signalling during tumour progression. In the study at hand, two thirds of *Kras*^{G12D}-driven primary mPDAC cell cultures gained specifically an extra copy of mutant *Kras*^{G12D} (*Kras*$^{G12D\text{-}iGD}$) during pancreatic cancer progression. When considering the *Kras* dimerization model for Ras signalling, one plausible explanation is that the increased proportion of *Kras*^{G12D} monomers in *Kras*$^{G12D\text{-}iGD}$ (i) favours the dimerization of *Kras*^{G12D} monomers and (ii) reduces the inhibitory effects of wild-type *Kras*, thereby increasing the oncogenic signalling flux through the Ras effector pathways and driving the progression to full-blown pancreatic cancers. Because *Kras*$^{G12D\text{-}HET}$ cancers remain in low signalling status of *Kras* they require the amplification of alternative oncogenes to over-activate Ras downstream pathways as required for cancer progression.

4.3 *KRAS*$^{MUT\text{-}iGD}$ in primary pancreatic cancer phenotypes

The sequencing of more than 500 pancreatic cancer exomes and more than 100 genomes has revealed that there is extensive mutational heterogeneity beyond the four signature mutations in *KRAS*, *CDKN2A*, *TP53* or genes of the canonical TGFβ pathway (Jones *et al.*, 2008; Biankin *et al.*, 2012; Waddell *et al.*, 2015; Bailey *et al.*, 2016). However, the

majority of heterogenic mutations could so far not be assigned to a broad range of biological or clinical phenotypes of pancreatic cancer, such as de-differentiation and metastatic dissemination. This is illustrating that the current understanding of fundamental processes of pancreatic cancer biology is still limited. Here, it could be shown that the dosage of mutant *KRAS* adds a previously unrecognized mechanism to the understanding of pancreatic cancer progression. In addition to the importance of *KRAS*$^{MUT\text{-}iGD}$ for pancreatic cancer progression, the gene dosage of *KRAS*^{MUT} also appeared to be a critical determinant for various biological and clinical PDAC phenotypes.

Primary PDAC cell cultures from PK mice with highest *Kras*^{G12D}-specific mRNA levels and strongest activation of Ras downstream pathways showed mesenchymal cell morphology *in vitro*. Notably, *Kras*$^{G12D\text{-}iGD}$ occurred to the same extent in epithelial and mesenchymal primary PDAC cell cultures indicating that the transcriptional upregulation of *Kras*^{G12D} mRNA levels above a certain threshold is required for epithelial-to-mesenchymal transition (EMT). Mesenchymal-like transcriptional programs could also be induced in human PDAC cell cultures through the overexpression of *KRAS*^{G12D}, indicating that *KRAS*^{MUT} is not simply a bystander of EMT. The mesenchymal *in vitro* phenotype of primary pancreatic cancer cell cultures from PK mice was also correlated with the *in vivo* phenotype of the corresponding primary tumours. While almost all tumours of epithelial mPDAC cell cultures were classified as PDAC grade 1 to 3 (ductal morphology), tumours from mesenchymal mPDAC cell cultures were almost exclusively classified as undifferentiated pancreatic carcinomas (spindle shape morphology). Interestingly, the survival of PK mice with primary pancreatic tumours that gave rise to mesenchymal cell cultures was reduced when compared to mice with tumours that gave rise to epithelial cell cultures. This is also reflected in the setting of human pancreatic cancer, where dedifferentiation is associated with poor patient survival (Winter *et al.*, 2008; Iacobuzio-Donahue *et al.*, 2009). Here in this study, it could be shown that human undifferentiated pancreatic carcinomas also display over-activation of Ras downstream signalling through the MAPK pathway and upregulation of genes involved in EMT, like in mesenchymal pancreatic cancer cell cultures from PK mice. The finding that undifferentiated pancreatic carcinomas show highest levels of *KRAS*^{MUT}-specific expression and oncogenic Ras signalling appears to make sense in the light of the previously mentioned fact that these tumours also show the most aggressive clinical behaviour/phenotype.

Frequent and early metastasis is a hallmark phenotype of pancreatic cancer. However, the understanding of the underlying processes of metastatic dissemination is still rudimentary from a genetic point of view. On the basis of whole genome sequencing, a recent study by Makohon-Moore *et al.* (Makohon-Moore *et al.*, 2017) found that all driver

mutations were shared between the primary pancreatic tumour and the corresponding metastatic lesions. All the mutations that emerged in the metastases were predicted to be passenger mutations without any functional consequences. Thereby, the authors concluded that there is no frequently mutated metastases driver gene in pancreatic cancer which is in line with observations made in other cancer entities. Besides the identification of metabolic and epigenetic alterations that are associated with pancreatic cancer metastasis (McDonald *et al.*, 2017; Roe *et al.*, 2017) it remained elusive if there is a genetic basis/factor that is driving metastatic dissemination. Here, it could be shown that *Kras*$^{G12D\text{-}iGD}$ is strongly associated with metastasis of pancreatic cancer to the lymph nodes, liver and/or lung. In PK mice, 76.9% of primary pancreatic tumours with *Kras*$^{G12D\text{-}iGD}$ were metastasized, whereas only 12.5% primary tumours with *Kras*$^{G12D\text{-}HET}$ status were disseminated to distant organs. It appears that gaining a single additional copy of the mutant *Kras*^{G12D} allele is already sufficient to dramatically increase the metastatic potential of primary pancreatic cancer cells. Accordingly, increased gene dosage of *Kras*^{G12D} is a strong metastatic driver, but not the activating mutation of *Kras* itself which might also explain the lack of additional driver gene mutations observed in primary-metastases comparison in human PDAC. The reanalysis of primary and metastatic PDACs from the cohort of Makohon-Moore *et al.* revealed that all PDAC lesions (primary and metastases) showed amplification of the *KRAS* locus, further supporting the notion of *KRAS*$^{MUT\text{-}iGD}$ as a driver of PDAC metastatic dissemination. These findings provide a model in which two distinct stages of pancreatic cancer evolution, (i) early PDAC progression and (ii) metastatic dissemination, are linked through the single genetic event of *KRAS*^{MUT} gene dosage increase. *KRAS*$^{MUT\text{-}iGD}$ occurs in the majority of PDAC, thereby also explaining its early and frequent metastasis in patients. In contrast to *Kras*$^{G12D\text{-}iGD}$, the amplification of alternative oncogenes, such as *Myc* and *Yap1*, in *Kras*$^{G12D\text{-}HET}$ background was not associated with high metastatic capability of primary pancreatic tumour cells in PK mice. Amplification of alternative oncogenes was reported to cooperate with mutant *Kras* and thus might amplify partial aspects of *Kras*^{G12D} downstream signalling that are not sufficient for driving colonization of distant organs. Both options to intensify oncogenic signalling pathways for providing the tumour cell with a metastatic advantage have been already observed in other cancer entities (reviewed in Vanharanta *et al.*, 2013a). For example, *VCAM-1* has been identified as a pro-metastatic gene in breast cancer cells that acts as a quantitative amplifier of oncogenic signalling. Mechanistically, *VCAM-1* was found to quantitatively amplify the output of the PI3K-Akt signalling pathway through hyper-sensitizing the pathway for the activation by low-levels of external growth stimuli (Chen *et al.*, 2011). One possible conclusion from these observations is that the signalling pathway activity level conferred by initial oncogenic events enabled the primary cancer cells to

progress locally. However, the output/robustness of oncogenic signalling needed to be increased when cancer cells were infiltrating and colonizing distant organ sites to enable survival of the tumour cells in a foreign microenvironment with limiting amounts of trophic factors. In renal cell carcinoma, it was found that the output of the hypoxia-inducible signalling pathway, which is initiating renal cell carcinogenesis, can be qualitatively expanded during metastatic progression by the epigenetic activation of previously inaccessible target genes. Specifically, through the demethylation of DNA and repressive histone marks primary renal cell carcinomas were able to engage new target genes of hypoxia-inducible transcription factors, such as the chemokine receptor *CXCR4*. Transcriptional activation of new hypoxia target genes resulted in a qualitative amplification of oncogenic signalling output and conferred the capability of metastatic colonization of the lung and bones (Vanharanta *et al.*, 2013b). It would be interesting to test if there is a certain Ras effector pathway in pancreatic cancer that needs to be over-activated to equip primary pancreatic cancer cells with high metastatic capabilities (qualitative model) or if the competence to colonize distant organs is the result of the a more or less general over-activation of all Ras effector pathways through *Kras*$^{G12D\text{-}iGD}$ (quantitative model).

Of note, the genetic status of *Kras*$^{G12D\text{-}iGD}$ was a better predictor for the occurrence of metastatic dissemination then the expression level of *Kras*^{G12D}. Interestingly however, the transcriptional upregulation of *Kras*^{G12D} expression above a certain level and not the genetic status of *Kras*$^{G12D\text{-}iGD}$ was a better marker for dedifferentiation of primary pancreatic tumour cells, indicating that both processes are governed by distinct molecular mechanisms. It appears that the acquisition of genetic *Kras*$^{G12D\text{-}iGD}$ is required for PDAC metastasis but not sufficient for dedifferentiation which requires additional upregulation of *Kras*^{G12D} expression above a certain expression level. It could be shown that overexpression of *KRAS*^{G12D} in human PDAC cell lines induces an EMT-like transcriptional program supporting the idea that upregulation of *KRAS*^{MUT} mRNA levels above a certain critical threshold plays a causal role in inducing de-differentiation of pancreatic cancer cells.

The involvement of *KRAS*$^{MUT\text{-}iGD}$ in metastasis and de-differentiation of tumour cells might be also helpful for the understanding of the still controversial connection between EMT and metastatic capability of (pancreatic) cancer cells. In general, it is thought that EMT allows a cell to infiltrate blood vessels at the primary tumour site and to survive in the circulation. The reversal process of MET is then required for extravasation at the metastatic site and colonization of the distant organ. *TWIST1*, *SNAI1* and *ZEB1* are the three key transcription factors responsible for EMT (Nieto *et al.*, 2016). However, knockout

of *Twist1* or *Snai1* in a *Kras*^{G12D}-driven mouse model of pancreatic cancer did not have any influence on the formation of invasive primary PDAC nor did it reduce or suppress the occurrence of distant metastasis (Zheng *et al.*, 2015). In another study, the knockout of *Zeb1* in a *Kras*^{G12D}-driven pancreatic cancer mouse model did also not influence the formation of primary mPDAC but substantially reduced the frequency of metastasis to the liver and/or lung (Krebs *et al.*, 2017). One possible interpretation of the observation that EMT suppression does not fully block colonization of distant organs in the *Kras*^{G12D}-driven pancreatic cancer mouse model might be that EMT is an epiphenomenon of molecular processes driving cancer cell metastasis. Specifically in the setting of pancreatic cancer, the acquisition of *Kras*$^{G12D\text{-}iGD}$ is conferring high metastatic competence to primary pancreatic cancer cells. At the same time *Kras*$^{G12D\text{-}iGD}$ is provides the basis for further upregulation of *Kras*^{G12D} mRNA levels that is associated with de-differentiation of tumour cells when exceeding a certain critical threshold. In addition, studying the role of EMT transcription factors for metastasis is further complicated due to their involvement in oncogenic signalling. For instance, EMT-inducing transcription factors were shown to attenuate side-effects of oncogenic insults through mitigating oncogenic induction of TP53-tumoursuppressive pathways or by conferring stemness-related properties to cancer cells (reviewed in Puisieux *et al.*, 2014).

For the first time, the incorporation of *KRAS*^{MUT} gene dosage provides a conceptual framework that can explain pancreatic cancer progression, early/frequent metastasis and dedifferentiation of pancreatic cancer within a single model.

4.4 *KRAS*$^{MUT\text{-}iGD}$ in primary pancreatic cancer evolution

To determine if there is a genetic dependency for the acquisition of *Kras*$^{G12D\text{-}iGD}$ in primary pancreatic cancers from PK mice, the genome of these tumours were screened for other somatic mutations that are acquired during pancreatic cancer progression. The most frequent deletion that could be identified involved the *Cdkn2a* locus and the nearby regulatory *Ncruc* (noncoding region upstream from *Cdkn2a*), jointly referred to as *Cdkn2a/Ncruc*. Notably, 85.2% of mPDAC with homozygous deletion of *Cdkn2a/Ncruc* (*Cdkn2a/Ncruc*$^{\Delta HOM}$) showed *Kras*$^{G12D\text{-}iGD}$ whereas 72.7% of mPDAC with heterozygous deletion or wild-type status (one tumour) of *Cdkn2a/Ncruc* (*Cdkn2a/Ncruc*$^{\Delta HET/WT}$) were *Kras*$^{G12D\text{-}HET}$. The inter-connection of *KRAS*$^{MUT\text{-}iGD}$ and *CDKN2A*$^{\Delta HOM}$ could be also confirmed in a data set of micro-dissected human pancreatic cancer from Witkiewicz *et al.* (Witkiewicz *et al.*, 2015) further supporting the relevance of these observations in human

PDAC. These findings suggest that *CDKN2A*$^{\Delta HOM}$ and *KRAS*$^{MUT\text{-}iGD}$ are tightly linked during pancreatic cancer tumour evolution with two possible evolutionary scenarios: (i) *KRAS*$^{MUT\text{-}iGD}$ occurs first, thereby inducing senescence which has to be escaped by acquisition of *CDKN2A*$^{\Delta HOM}$ or (ii) *CDKN2A*$^{\Delta HOM}$ is acquired first to bypass senescence before *KRAS*$^{MUT\text{-}iGD}$ can occur. Since senescence is thought to represent an irreversible road-block to the development of cancer (Sharpless *et al.*, 2015), both scenarios are fundamentally different in terms of the underling biological principles. To resolve the sequential order of both genetic events, CNA and SNP patterns at the *Cdkn2a* and *Kras* locus were compared in primary mPDAC and corresponding metastases. In all cases, where the sequential order of both events could be inferred from CNA and SNP patterns, *Cdkn2a*$^{\Delta HOM}$ occurred first whereas *Kras*$^{G12D\text{-}iGD}$ was acquired in a second step with heterogenic patterns in different metastases of the same primary mPDAC. These findings support the previous notion that once senescence is acquired it represents a permanent block in pancreatic cancer evolution that cannot be escaped. The observation that *Kras*$^{G12D\text{-}HET}$ and amplification of alternative oncogenes such as *Myc*, *Yap1* or *Nfkb2* can occur in a *Cdkn2a*$^{\Delta HET}$ background suggests context-dependent haploinsufficiency of *Cdkn2a*-mediated tumour suppression. The acquisition of multiple independent *Kras*$^{G12D\text{-}iGD}$ gains in different metastatic lesions of the same primary tumour indicates parallel evolution and functional convergence towards *Kras*$^{G12D\text{-}iGD}$ once the *Cdkn2a*$^{\Delta HOM}$ status is acquired. This finding could be also confirmed in a proof-of-principle experiment by conditional knockout of *Cdkn2a* in the pancreas of PK mice (PKC mice). *Kras*$^{G12D\text{-}iGD}$ was found to occur in all primary pancreatic cancer cell cultures derived from PKC mice providing strong *in vivo* evidence that *Cdkn2a*-mediated tumour suppression is licensing the acquisition of *Kras*$^{G12D\text{-}iGD}$. That *Kras*$^{G12D\text{-}iGD}$ occurs in 100% of all PKC tumours makes also sense in the light of the phenotypes that were described for the PKC model by Guerra *et al.* in 2003. When characterizing the phenotypes of PKC mice this pancreatic cancer model showed: (i) strongly reduced survival due to the rapid onset of pancreatic carcinogenesis, (ii) pronounced local invasion and frequent metastasis to the lymph nodes and (iii) appearance of undifferentiated/spindle-shaped cancer cells in the histology of primary pancreatic tumours. All three phenotypes can be explained from *Cdkn2a*-mediated licensing of *Kras*$^{G12D\text{-}iGD}$ acquisition that was identified as a driver of pancreatic cancer progression, metastasis and dedifferentiation in PK mice.

The critical role of increasing oncogenic dosage or activity for driving tumour initiation and progression has been established already earlier in various murine cancer models and is supporting the conclusions that were made in this work. For example, in a study by Altomonti *et al.* different *Pten* alleles with distinct levels of *PTEN* inactivation were used to modulate the activity of Akt signalling *in vivo*. Through the combination of different *Pten*

alleles a decreasing series of *PTEN* inactivation levels could be established which was associated with decreasing survival of mice and increasing tumour incidence suggesting that over-activation of oncogenic signalling is a critical step for tumour induction and/or progression (Alimonti *et al.*, 2010; Berger *et al.*, 2011). However in some tissues such as prostate, full inactivation of *Pten* (which translates in most strongest activation of Akt signalling) resulted in oncogene-induced senescence and blockade of tumourigenesis that could be bypassed through the inactivation of *Trp53* (Chen *et al.*, 2005). Like for *Pten*, similar observations were made for the activity levels of other oncogenes. For example, the activation of the oncogenes *Myc* or mutant *Ras* is critical for driving tumourigenesis in many tissues. By contrast, it was also found that increasing expression of *MYC* or *HRAS*^{G12V} from low to unphysiological high levels induced apoptosis or premature senescence instead of proliferation in non-malignant cells/fibroblasts *in vitro* (Evan *et al.*, 1992; Serrano *et al.*, 1997). In a breast cancer model that allowed for the titration of distinct levels of oncogenic *Hras*^{G12V} it could be shown *in vivo* that low levels of oncogenic *Hras*^{G12V} expression were sufficient to stimulate the hyper-proliferation of the mammary epithelium but needed to be further increased to drive the formation of mammary tumours. While the expression of low oncogenic *Hras*^{G12V} levels was tolerated in mammary epithelial cells, the expression of *Hras*^{G12V} at similar levels as observed in mouse mammary tumours resulted in oncogene-induced senescence of mammary epithelial cells that was shown to be *Cdkn2a*-dependent (Sarkisian *et al.*, 2007). These findings are also supported by the restoration of functional *Trp53* in a *Kras*^{G12D}-driven mouse model of lung cancer. The restoration of functional *Trp53* resulted in a significant regression of high grade tumours. In contrast, low grade tumours were not affected by restoration of functional *Trp53*. It was found that *Trp53* effector pathways were only activated in high grade lung tumours due to the marked upregulation of oncogenic *Kras*^{G12D} signalling in comparison to low-grade tumours (Feldser *et al.*, 2010; Junttila *et al.*, 2010). Similar observation were made in a *Myc*-driven cancer model in which low levels of *Myc* were sufficient to induce proliferation of somatic cells *in vivo* while an increase of *Myc* expression by 2-fold induced oncogene-induced senescence or apoptosis through the activation of *Arf*/*Trp53* effector pathways (Murphy *et al.*, 2008). Altogether, these studies from different cancer entities and oncogene activations/tumour suppressor gene knockouts support the model of pancreatic cancer progression that is proposed here: (i) low levels of oncogenic *Kras*^{G12D} are sufficient to drive proliferation of epithelial cells of the pancreas, (ii) increased oncogenic signalling is established through *Kras*$^{G12D\text{-}iGD}$ or the amplification of alternative oncogenes that is required for driving pancreatic cancer progression (iii) dosage increase of alternative oncogenes or *Kras*^{G12D} is controlled by the *Cdkn2a*/*Trp53* tumour suppressor pathways that need to be inactivated for tumourigenesis heterozygously or homozygously,

respectively. Additional support for this model comes from studies investigating the mutual exclusivity of activating mutations in *KRAS* and *BRAF* or of activating mutations in *KRAS* and *EGFR*. Interestingly, when each of the oncogenes was activated individually in the lungs of mice, the single oncogene was able to drive lung tumourigenesis. Remarkably, the concomitant activation of mutual exclusive oncogenes (*Kras*^{G12D} plus *Braf*^{V600E} or *Kras*^{G12V} plus *Egfr*^{L858R}) in the lung of mice resulted in the block of lung tumourigenesis that could be attributed to oncogene-induced senescence or cellular toxicity caused through hyper-activation of oncogenic signalling (Cisowski *et al.*, 2016; Ambrogio *et al.*, 2017). While activating mutations of *EGFR* or *BRAF* are mutually exclusive with activating *KRAS* mutations, it is worth noting that activating mutations of *PIK3CA* (upstream of Akt signalling and downstream of *KRAS*) co-occur with activating *KRAS* mutations in cancer, such as in pancreatic cancer. It would be interesting for future investigations, if *KRAS* and *PIK3CA* mutations can occur within the same tumour cell due to: (i) the more modest increase of *KRAS*^{MUT} downstream signalling (as compared to *KRAS*^{MUT} plus *EGFR*^{MUT}/*BRAF*^{MUT}), (ii) the particular sensitivity of oncogenic senescence programs to the MAPK pathway and/or (iii) additional oncogenic mechanisms.

TP53 represents the second hallmark tumour suppressor gene that needs to be inactivated during the progression to pancreatic cancer. Knockout of *Trp53* in the PK model (PKP mice) strongly reduced the survival of the mice due to early formation of metastatic PDAC, such as in PKC mice (Hingorani *et al.*, 2005; Bardeesy *et al.*, 2006a). Interestingly, homozygous knockout of *Trp53* also bypassed the need for inactivation of the *Cdkn2a* locus in primary pancreatic cancers of PKP mice indicating that *Trp53* is also a barrier for the acquisition of *Kras*$^{G12D\text{-}iGD}$. Indeed, all primary pancreatic cancer cell cultures from PKP mice displayed *Kras*$^{G12D\text{-}iGD}$. In contrast to PKC mice, mPDAC from the PKP model showed a well-differentiated phenotype *in vivo* without the appearance of de-differentiated/spindle-shaped cells. It would be interesting to investigate the molecular basis for this difference in future experiments. Maybe, inactivation of *Arf*-mediated *Trp53*-tumour suppression allows for the activation of a Ras signalling program that rely on *Arf*, whereas loss of *Arf* plus *Ink4a* in the full *Cdkn2a* knockout might allow also for the activation of an additional *Ink4a*-dependent Ras signalling network. Both components of Ras signalling might be required for dedifferentiation of pancreatic cancer cells.

Inactivation of the canonical Tgfβ pathway is the third hallmark of pancreatic cancer progression. Inactivation of the canonical Tgfβ pathway typically occurs through the genetic alteration of *SMAD4* in the vast majority of cases or less frequent through the inactivation of *TGFBR2* (Hahn *et al.*, 1996; Siegel *et al.*, 2003; Whittle *et al.*, 2015). The analysis of *Kras*$^{G12D\text{-}iGD}$ in the PK mouse model with heterozygous or homozygous

knockout of *Tgfbr2* in the pancreas (PKT mice) revealed that two thirds of primary pancreatic cancer cell cultures acquired a *Cdkn2a*^{HET} background with *Kras*$^{G12D\text{-}HET}$ status. In contrast to inactivation of *Cdkn2a* or *Trp53*, the inactivation of *Tgfbr2* in the PK mouse model favoured pancreatic carcinogenesis through the evolutionary trajectory with alternative *Kras*$^{G12D\text{-}HET}$-cooperating oncogenic amplifications. The predominance of *Kras*$^{G12D\text{-}HET}$ mPDAC of PKT mice is also supported by the intial phenotypic characterization of the mouse model by Ijichy *et al.* who observed: (i) a strongly reduced survival, (ii) the emergence of well-differentiated primary PDAC and (iii) minimal local invasion and absence of distant metastasis (Ijichi *et al.*, 2006). PKT mice with homozygous knockout of *Tgfbr2* also had a significantly reduced survival as compared to heterozygous knockout mice. Interestingly and similar to findings in the work at hand, the wild-type allele in PKT mice with heterozygous *Tgfbr2* inactivation was not lost during tumour progression. This indicates that *Tgfbr2*$^{\Delta HET}$ is sufficient to drive pancreatic carcinogenesis in a *Kras*$^{G12D\text{-}HET}$ background and also shows that *Tgfbr2*$^{\Delta HOM}$ deletion is not a rate-limiting factor for *Kras*$^{G12D\text{-}HET}$ pancreatic cancer evolution. The absence of local invasion and distant organ metastases in the PKT model further supports the role of *Kras*$^{G12D\text{-}iGD}$ in driving metastases formation.

So far there is only one study available that investigated the genetics of treatment-naive human metastatic PDAC tissues (Makohon-Moore *et al.*, 2017). Critically, the size of that study was very limited due to the fact that the initial cohort of 150 patients was reduced to a set of 4 patients when applying all criteria for selecting untreated metastatic PDAC. The limited availability of human metastatic PDAC cohorts, and specifically the lack of such cell line based resources, highlights the importance of mouse models in studying the "natural" evolution of metastatic pancreatic cancer in an untreated setting. The use of pancreatic cancer cell cultures derived from *Kras*^{G12D}-driven mouse models provided the unique opportunity to study "natural" evolution of metastatic PDAC at large scale in a treatment-naive setting. The key finding here is that oncogenic dosage of *Kras*^{G12D} and alternative oncogenes, is differentially licensed by PDAC hallmark tumour suppressor genes, thereby also controlling phenotypic diversification. Although the all mentioned combinations of the *Kras*^{G12D}-driven mouse model with distinct tumour suppressor gene knockouts accelerate the development of PDAC, the pancreatic cancers of each mouse model show distinct/specific phenotypes that can be attributed to the evolutionary trajectories the tumours evolve along: (i) frequent metastases and local invasion of PDAC from PKC or PKP mice is caused by the acquisition of *Kras*$^{G12D\text{-}iGD}$ (ii) absence of metastases and local invasion in PDAC of PKT mice can be explained by the predominance of *Kras*$^{G12D\text{-}HET}$ in this model.

4.5 The genetic landscape of mouse and human PDAC

Tremendous progress has been made in the genetic characterization of human PDAC genomes in the last decade through the development of next-generation sequencing approaches. To date, more than 500 pancreatic cancer genomes have been sequenced revealing extensive heterogeneity of mutated genes beyond the well-known hallmark mutations occurring in *KRAS*, *CDKN2A*, *TP53* and genes of the canonical TGFβ signalling pathway (such as *SMAD4* or *TGFBR2*). The study of human PDAC genetics revealed: (i) that a typical human PDAC genome carries a median of approximately 60 exonic SNVs, 6 exonic indels, 7 translocations and that in median 15.5% of the human pancreatic cancer genome is affected by copy-number alterations, (ii) that the spectrum of trinucleotide-specific base substitutions (caused by distinct mutational processes in the cell, Alexandrov *et al.*, 2013a) show similar patterns in-between individual hPDAC samples with a predominance of mutational signature 1 associated with cellular processes involved in aging and (iii) that 10 to 67% of hPDAC genomes were affected by chromothripsis (Waddell *et al.*, 2015; Notta *et al.*, 2016). In the study of Notta *et al.*, chromothripsis was frequently associated with the inactivation of tumour suppressor genes suggesting its involvement in the progression of pancreatic cancer (Notta *et al.*, 2016). The comprehensive genetic characterization of primary pancreatic cancer cell cultures derived from a $Kras^{G12D}$-driven mouse model that was performed in the study presented here, revealed substantial similarities as well as differences as compared to hPDAC genetics. The genomes of mPDACs derived from the $Kras^{G12D}$-driven mouse model carried a median of approximately 18 exonic SNVs, 4 exonic indels, 2.5 translocations and in median 5.5% of the cancer genome was affected by copy-number alterations. Overall this translated into a significantly reduced mutational load when compared to human PDAC genomes (SNV, Indel, translocation and CNV load decreased by 3.3, 1.5, 2.6 and 2.8 fold, respectively). The amount of recurrently altered genes was also substantially reduced in mouse primary pancreatic cancer cell cultures which might be linked to the overall reduced mutational load. For example, mutation of *Trp53* occurred in one out of 38 primary mPDACs (up to 85% of human PDAC cases show *TP53* inactivation (Yachida *et al.*, 2012) and no genetic alterations of *SMAD4* were detected in mPDAC (inactivated in 55% of human PDAC cases, Hahn *et al.*, 1996). The biological or genetic reason(s) for the lower mutational burden of mouse PDAC as compared to human PDAC are yet unknown but might be associated with cross-species differences in chromosome architecture (Bardeesy *et al.*, 2006a) or the activation of oncogenic $Kras^{G12D}$ during embryogenesis in the mouse whereas in humans *KRAS* mutations occur in the adult pancreas. Of note, the time point

of activation of oncogenic *Kras* in the mouse has a dramatic influence on the induction and the incidence of pancreatic cancer in this model. When oncogenic $Kras^{G12V}$ was activated in the adult pancreas, the mice were refractory to the induction of PanINs as well as PDAC. However, the block in pancreatic carcinogenesis could be rescued when the adult activation of oncogenic $Kras^{G12V}$ was combined with a mild form of pancreatitis (Guerra *et al.*, 2007). These results confirm that pancreatitis is a risk factor for pancreatic cancer and indicate that additional cell-extrinsic or -intrinsic factors/stimuli are required to drive pancreatic cancer formation from cells of the adult pancreas. Such “co-factors” might be bypassed in the PDAC mouse model with embryonic $Kras^{G12D}$ activation. Importantly however, the lower mutational burden in the primary mouse pancreatic cancer cell cultures allowed for the identification of fundamental processes in PDAC genetics, biology and phenotypic diversification. Because the $Kras^{G12D}$-driven mouse model resembles key features of the human disease in a background of significantly lower mutational burden, one can speculate that only a small number of genetic alterations are contributing to pancreatic cancer evolution and phenotypic diversification.

Although there are substantial differences in mouse and human PDAC genomes there are also significant similarities. For example the mutational spectra overlapped very well when primary mouse PDAC cell cultures were compared to different human PDAC cohorts with distinct properties (bulk tissues, micro-dissected tissues or cell lines). The mutational spectrum is defined by the frequency at which a trinucleotide context-specific base substitution/mutation is occurring and which is indicating the mutational processes that are operative in the (cancer) genome. Several mutational processes/signatures might be at play at the same time. The mutational spectrum can be used to decipher individual mutational signatures that in combination make up the mutational spectrum. Deconstructing the mutational spectrum of the mPDAC cohort revealed that mutational signature 1 was predominantly contributing to the mutational spectrum. Signature 1 is associated with C>T transversion at NCG trinucleotides that are accumulated during aging. Mutational signature 1 was also the only signature that was consistently detected throughout all human PDAC cohorts. All other signatures were identified in 2 or less of the six PDAC cohorts suggesting that they are less relevant to pancreatic cancer or that they are only operative in a subset of pancreatic cancer cases. Overall the similarity of mutational spectra and the predominance of mutational signature 1 in mouse and human PDAC indicate that similar mutational processes are operative during pancreatic carcinogenesis in both species. Chromothripsis was identified as another mutational process that can be detected in 10 to 67% of human pancreatic cancer cases and is thought to drive pancreatic carcinogenesis through the simultaneous rather than

sequential acquisition of preneoplastic genetic drivers (Notta *et al.*, 2016). Complex chromosomal rearrangements with copy-number patterns indicative of chromothripsis were also identified in one third of primary mouse PDAC cell cultures. The detection of at least 10, 20, or 50 copy-number alterations identified in copy-number profiling data has been used in a variety of studies for the assessment of chromothripsis. However, there exist significant differences in the number/amount and distribution of copy-number alterations between cancer entities that limit the specificity of this criterion for the detection of chromothripsis (reviewed by Korbel *et al.*, 2013). Therefore, whole genome sequencing (WGS) and M-FISH was performed for the statistical inference of chromothripsis in mPDAC. All six statistical criteria for the assessment of chromothripsis, as proposed by Korbel and Campbell (Korbel *et al.*, 2013), could be confirmed on the basis of rearrangement sequencing data. These results show that chromothripsis is operative in the *Kras*^{G12D}-driven pancreatic cancer mouse model as well as in human pancreatic cancer.

The occurrence of chromothripis in PDAC of PK mice was associated with reduced survival as compared to mice with PDAC that did not show chromothripsis. In addition, the majority of chromothriptic events occurred on chromosome 4 invariably involving the *Cdkn2a*/*Ncruc* locus. This further supports the role of chromothripsis as a driver of pancreatic carcinogenesis through the inactivation of the *Cdkn2a*/*Ncruc* locus. Although there is a strong association of chromothripsis with preneoplastic genetic drivers in multiple studies and different human cancer entities this hypothesis could not been functionally tested yet due to the lack of appropriate model systems. For this, homozygous inactivation of *Cdkn2a* was engineered in the *Kras*^{G12D}-driven mouse model of pancreatic cancer (PKC mice). In PKC mice, no chromothripsis or other complex chromosomal rearrangements were observed on chromosome 4 providing the first *in vivo* functional evidence that chromothripsis is a *bona fide* mutational process driving and accelerating pancreatic carcinogenesis through the acquisition of cancer-driving genetic alterations, such as inactivation of *Cdkn2a*. Notably, a strong association of *Trp53* and chromothripsis was found in Sonic-Hedgehog medulloblastoma and acutemyeloid leukemia (Rausch *et al.*, 2012). However, complex rearrangements indicative of chromothripsis were absent in a *Kras*^{G12D}-driven mouse model of pancreatic cancer with homozygous, engineered knockout of *Trp53* (PKP mice). Complete inactivation of *Trp53* bypassed the need for *Cdkn2a* inactivation and therefore obviates the need for acquisition of additional cancer driving genetic alterations through chromothripsis in mouse PDAC. The strong link of *TP53* inactivation with chromothripsis in Sonic-Hedgehog medulloblastoma and acutemyeloid

leukemia might thus be explained by additional bottlenecks that need to be overcome in these cancer entities after *TP53* is already inactivated.

4.6 Outlook

Genetic alterations observed in human PDAC genomes could, with very few exceptions, not be broadly linked to different phenotypes of the disease. Since the PDAC signature genes *KRAS*, *CDKN2A*, *TP53* and *SMAD4* occur in 50 to 90% of cases, they were thought not to contribute to the phenotypic diversity of PDAC. However, rare drivers could also not be overtly correlated to specific disease characteristics. So far, the classification of biological PDAC subtypes is still largely transcriptome-based and shows only little correlation with mutated genes (Bailey *et al.*, 2016). Likewise, it is not known if there is a genetic basis for PDAC metastasis since no recurrently mutated “metastasis driver” has been identified yet (Makohon-Moore *et al.*, 2017). Despite the sequencing of more than 500 human PDAC exomes/genomes, mechanisms of key aspects of the disease remain unresolved. The work presented in this thesis shows that the dosage of mutant *KRAS* differs between individual PDACs. $KRAS^{MUT}$ dosage was found to define major aspects of PDAC biology, including early progression, metastasis, histopathology, cellular plasticity and clinical aggressiveness. Importantly, the type and state of tumour suppressor gene alterations controlled $KRAS^{MUT}$ dosage, thereby licencing distinct evolutionary trajectories along which the tumour cells were able to evolve. Neither the mutation of major PDAC signature genes nor the mutational heterogeneity could broadly explain PDAC phenotypes yet. This work extends the traditional approach of “capturing mutations” for the understanding of pancreatic cancer biology by two additional cornerstones: (i) “dosage variation of oncogenes and tumour suppressor genes” and (ii) “evolution of the cancer genome” (see Figure 21).

These advances were mainly possible through establishing a resource of >100 primary cell cultures derived from distinct pancreatic cancer mouse models. Studying mouse models at large scale overcame key limitations of human PDAC genome analyses: (i) genetic complexity of human tumours, (ii) stromal contamination of human bulk tumours (iii) lack of cell culture-based human metastatic PDAC cohorts and (iv) scarcity of treatment naive human metastatic PDAC studies. The analysis of pancreatic cancer mouse models also facilitated many other aspects which could not be investigated in humans, such as (i) mechanistic analysis of oncogene dosage and tumour suppressor gene dosage interaction using different gene knockouts or (ii) phylogenetic tracking of tumour phenotypes *in vivo*

using CRISPR/Cas9. Altogether this illustrates the key importance of mouse models for the understanding of principles in cancer biology.

In future, the analysis of the mechanisms identified in this study should be further extended through the in-depth analysis of the mouse models used here, as well as through the characterization of additional *Kras*^{G12D}- or *Braf*^{V600E} and *Pik3ca*^{H1047R}-driven pancreatic cancer mouse models. For instance, it would be interesting to investigate the epigenetic landscape of primary and metastatic pancreatic cancer cell lines (e.g. histone modifications and DNA methylation) to identify transcription factors and epigenetic patterns that are associated with *Kras*$^{G12D\text{-}iGD}$ status, cancer cell metestasis or organ tropism of metastasis. Likewise, the *in vivo* knockout of genes involved in Ras-downstream signalling pathways at large scale using AAV and CRISPR/Cas9 might reveal if the quantitative or qualitative amplification of Ras signalling (or both) through *Kras*$^{G12D\text{-}iGD}$ is critical for conferring metastatic capability to pancreatic cancer cells. The comparison of the phospho-proteomes from *Kras*$^{G12D\text{-}HET}$ versus *Kras*$^{G12D\text{-}iGD}$ tumours would also help to elucidate the pathways involved in driving pancreatic carcinogenesis in a *Kras*$^{G12D\text{-}HET}$ background.

Phylogenetic tracking of cancer cell clones using CRISPR/Cas9 indel patterns revealed that epithelial pancreatic cancer cells (C2 cluster) can undergo EMT to acquire a mesenchymal phenotype *in vivo* (C1 cluster). However, EMT occurred only in the minority of cases raising the question if epithelial (C2) and mesenchymal mPDAC (C1) evolve from the same or different cell(s) of origin. For example, are pancreatic cancers that evolve from acinar cells more likely to form C1-type tumours whereas C2-type tumours are predominantly evolving from ductal cells, or vice versa? The characterization of the epigenetic landscape described above might be also helpful for investigating if the cell of origin is influencing the evolution of pancreatic cancer cells into distinct clusters (C2a/b/c, C1).

Another important goal would be to develop a *Kras*^{G12D} mouse model that allows for modelling *Kras*^{G12D} gene dosage *in vivo*. The widely used *Kras*$^{LSL\text{-}G12D}$ prevents expression of *Kras*^{G12D} through an upstream lox-stop-lox (LSL) cassette that can be excised by Cre recombinase for activation of *Kras*^{G12D} expression. Of note, the LSL cassette leads to a knockout of the wildtype Kras locus before the LSL cassette is recombined (also known as knockout-first alleles). Attempts to breed the *Kras*$^{LSL\text{-}G12D}$ homozygous would thus result in complete knockout of wildtype *Kras* which is lethal during mouse embryogenesis. Accordingly, increased gene dosage of *Kras*^{G12D} (*Kras*$^{G12D\text{-}iGD}$) and its influence on pancreatic cancer evolution and phenotypic diversification cannot be modelled *in vivo*

using the traditional *Kras*$^{LSL\text{-}G12D}$-allele. The *Kras*$^{LSL\text{-}G12D}$ allele also results in a heterozygous knockout of wildtype *Kras* in “non-recombined” cells, such as in the stroma. Therefore, it would be important to develop a *Kras*^{G12D} mouse allele that overcomes both limitations of (i) heterozygous knockout in non-recombined cells and (ii) inability of modelling *Kras*$^{G12D\text{-}iGD}$ *in vivo*.

RAS gene mutations affect more than 30% of human cancers, often involving their allelic imbalance, such as in PDAC. It would be interesting to investigate if the mechanisms found in the *Kras*^{G12D}-driven pancreatic cancer mouse model can be extrapolated to *Kras*^{G12D}-driven mouse models of other cancer entities, such as the lung, intestine or liver. Such studies would reveal if *Kras*$^{G12D\text{-}iGD}$ is a general phenomenon of tumour progression and whether tumour suppressive mechanisms are similarly licencing oncogenic dosage in distinct tissue backgrounds.

5 Bibliography

1. Ilic, M., and Ilic, I. (2016). Epidemiology of pancreatic cancer. World journal of gastroenterology *22*, 9694-9705.
2. Rahib, L., Smith, B.D., Aizenberg, R., Rosenzweig, A.B., Fleshman, J.M., and Matrisian, L.M. (2014). Projecting cancer incidence and deaths to 2030: the unexpected burden of thyroid, liver, and pancreas cancers in the United States. Cancer research *74*, 2913-2921.
3. Siegel, R.L., Miller, K.D., and Jemal, A. (2016). Cancer statistics, 2016. CA: a cancer journal for clinicians *66*, 7-30.
4. Klimstra, D.S., Pitman, M.B., and Hruban, R.H. (2009). An algorithmic approach to the diagnosis of pancreatic neoplasms. Archives of pathology & laboratory medicine *133*, 454-464.
5. Ying, H., Dey, P., Yao, W., Kimmelman, A.C., Draetta, G.F., Maitra, A., and DePinho, R.A. (2016). Genetics and biology of pancreatic ductal adenocarcinoma. Genes & development *30*, 355-385.
6. Maitra, A., Fukushima, N., Takaori, K., and Hruban, R.H. (2005). Precursors to invasive pancreatic cancer. Advances in anatomic pathology *12*, 81-91.
7. Hezel, A.F., Kimmelman, A.C., Stanger, B.Z., Bardeesy, N., and Depinho, R.A. (2006). Genetics and biology of pancreatic ductal adenocarcinoma. Genes & development *20*, 1218-1249.
8. Hruban, R.H., Goggins, M., Parsons, J., and Kern, S.E. (2000). Progression model for pancreatic cancer. Clinical cancer research : an official journal of the American Association for Cancer Research *6*, 2969-2972.
9. Pylayeva-Gupta, Y., Grabocka, E., and Bar-Sagi, D. (2011). RAS oncogenes: weaving a tumorigenic web. Nature reviews Cancer *11*, 761-774.
10. Schutte, M., Hruban, R.H., Geradts, J., Maynard, R., Hilgers, W., Rabindran, S.K., Moskaluk, C.A., Hahn, S.A., Schwarte-Waldhoff, I., Schmiegel, W., Baylin, S.B., Kern, S.E., and Herman, J.G. (1997). Abrogation of the Rb/p16 tumor-suppressive pathway in virtually all pancreatic carcinomas. Cancer research *57*, 3126-3130.
11. Russo, A.A., Tong, L., Lee, J.O., Jeffrey, P.D., and Pavletich, N.P. (1998). Structural basis for inhibition of the cyclin-dependent kinase Cdk6 by the tumour suppressor p16INK4a. Nature *395*, 237-243.
12. Zindy, F., Eischen, C.M., Randle, D.H., Kamijo, T., Cleveland, J.L., Sherr, C.J., and Roussel, M.F. (1998). Myc signaling via the ARF tumor suppressor regulates p53-dependent apoptosis and immortalization. Genes & development *12*, 2424-2433.
13. Zhu, J.W., DeRyckere, D., Li, F.X., Wan, Y.Y., and DeGregori, J. (1999). A role for E2F1 in the induction of ARF, p53, and apoptosis during thymic negative selection. Cell growth & differentiation : the molecular biology journal of the American Association for Cancer Research *10*, 829-838.
14. Dimri, G.P., Itahana, K., Acosta, M., and Campisi, J. (2000). Regulation of a senescence checkpoint response by the E2F1 transcription factor and p14(ARF) tumor suppressor. Molecular and cellular biology *20*, 273-285.
15. Palmero, I., Pantoja, C., and Serrano, M. (1998). p19ARF links the tumour suppressor p53 to Ras. Nature *395*, 125-126.
16. Sharpless, N.E., Bardeesy, N., Lee, K.H., Carrasco, D., Castrillon, D.H., Aguirre, A.J., Wu, E.A., Horner, J.W., and DePinho, R.A. (2001). Loss of p16Ink4a with retention of p19Arf predisposes mice to tumorigenesis. Nature *413*, 86-91.
17. Huot, T.J., Rowe, J., Harland, M., Drayton, S., Brookes, S., Gooptu, C., Purkis, P., Fried, M., Bataille, V., Hara, E., Newton-Bishop, J., and Peters, G. (2002). Biallelic mutations in p16(INK4a) confer resistance to Ras- and Ets-induced senescence in human diploid fibroblasts. Molecular and cellular biology *22*, 8135-8143.
18. Sharpless, N.E. (2005). INK4a/ARF: a multifunctional tumor suppressor locus. Mutation research *576*, 22-38.
19. Bardeesy, N., Aguirre, A.J., Chu, G.C., Cheng, K.H., Lopez, L.V., Hezel, A.F., Feng, B., Brennan, C., Weissleder, R., Mahmood, U., Hanahan, D., Redston, M.S., Chin, L., and Depinho, R.A. (2006a). Both p16(Ink4a) and the p19(Arf)-p53 pathway constrain progression of pancreatic adenocarcinoma in the mouse. Proceedings of the National Academy of Sciences of the United States of America *103*, 5947-5952.
20. Goldstein, A.M., Fraser, M.C., Struewing, J.P., Hussussian, C.J., Ranade, K., Zametkin, D.P., Fontaine, L.S., Organic, S.M., Dracopoli, N.C., Clark, W.H., Jr., and et al. (1995). Increased risk of pancreatic cancer in melanoma-prone kindreds with p16INK4 mutations. The New England journal of medicine *333*, 970-974.
21. Whelan, A.J., Bartsch, D., and Goodfellow, P.J. (1995). Brief report: a familial syndrome of pancreatic cancer and melanoma with a mutation in the CDKN2 tumor-suppressor gene. The New England journal of medicine *333*, 975-977.
22. Randerson-Moor, J.A., Harland, M., Williams, S., Cuthbert-Heavens, D., Sheridan, E., Aveyard, J., Sibley, K., Whitaker, L., Knowles, M., Bishop, J.N., and Bishop, D.T. (2001). A germline deletion of p14(ARF) but not CDKN2A in a melanoma-neural system tumour syndrome family. Human molecular genetics *10*, 55-62.
23. Rizos, H., Puig, S., Badenas, C., Malvehy, J., Darmanian, A.P., Jimenez, L., Mila, M., and Kefford, R.F. (2001). A melanoma-associated germline mutation in exon 1beta inactivates p14ARF. Oncogene *20*, 5543-5547.
24. Yachida, S., White, C.M., Naito, Y., Zhong, Y., Brosnan, J.A., Macgregor-Das, A.M., Morgan, R.A., Saunders, T., Laheru, D.A., Herman, J.M., Hruban, R.H., Klein, A.P., Jones, S., Velculescu, V., Wolfgang, C.L., and Iacobuzio-Donahue, C.A. (2012). Clinical significance of the genetic landscape of pancreatic cancer and implications for identification of potential long-term survivors. Clinical cancer research : an official journal of the American Association for Cancer Research *18*, 6339-6347.
25. Jones, S., Zhang, X., Parsons, D.W., Lin, J.C., Leary, R.J., Angenendt, P., Mankoo, P., Carter, H., Kamiyama, H., Jimeno, A., Hong, S.M., Fu, B., Lin, M.T., Calhoun, E.S., Kamiyama, M., Walter, K., Nikolskaya, T., Nikolsky, Y., Hartigan, J., Smith, D.R., Hidalgo, M., Leach, S.D., Klein, A.P., Jaffee, E.M., Goggins, M., Maitra, A., Iacobuzio-Donahue, C., Eshleman, J.R., Kern, S.E., Hruban, R.H., Karchin, R., Papadopoulos, N., Parmigiani, G., Vogelstein, B., Velculescu, V.E., and Kinzler, K.W. (2008). Core signaling pathways in human pancreatic cancers revealed by global genomic analyses. Science (New York, NY) *321*, 1801-1806.
26. Vilenchik, M.M., and Knudson, A.G. (2003). Endogenous DNA double-strand breaks: production, fidelity of repair, and induction of cancer. Proceedings of the National Academy of Sciences of the United States of America *100*, 12871-12876.
27. Bieging, K.T., Mello, S.S., and Attardi, L.D. (2014). Unravelling mechanisms of p53-mediated tumour suppression. Nature reviews Cancer *14*, 359-370.

28. Heinmoller, E., Dietmaier, W., Zirngibl, H., Heinmoller, P., Scaringe, W., Jauch, K.W., Hofstadter, F., and Ruschoff, J. (2000). Molecular analysis of microdissected tumors and preneoplastic intraductal lesions in pancreatic carcinoma. The American journal of pathology *157*, 83-92.
29. Maitra, A., Adsay, N.V., Argani, P., Iacobuzio-Donahue, C., De Marzo, A., Cameron, J.L., Yeo, C.J., and Hruban, R.H. (2003). Multicomponent analysis of the pancreatic adenocarcinoma progression model using a pancreatic intraepithelial neoplasia tissue microarray. Modern pathology : an official journal of the United States and Canadian Academy of Pathology, Inc *16*, 902-912.
30. Hahn, S.A., Schutte, M., Hoque, A.T., Moskaluk, C.A., da Costa, L.T., Rozenblum, E., Weinstein, C.L., Fischer, A., Yeo, C.J., Hruban, R.H., and Kern, S.E. (1996). DPC4, a candidate tumor suppressor gene at human chromosome 18q21.1. Science (New York, NY) *271*, 350-353.
31. Shi, Y., and Massague, J. (2003). Mechanisms of TGF-beta signaling from cell membrane to the nucleus. Cell *113*, 685-700.
32. Massague, J. (2012). TGFbeta signalling in context. Nature reviews Molecular cell biology *13*, 616-630.
33. Siegel, P.M., and Massague, J. (2003). Cytostatic and apoptotic actions of TGF-beta in homeostasis and cancer. Nature reviews Cancer *3*, 807-821.
34. Whittle, M.C., Izeradjene, K., Rani, P.G., Feng, L., Carlson, M.A., DelGiorno, K.E., Wood, L.D., Goggins, M., Hruban, R.H., Chang, A.E., Calses, P., Thorsen, S.M., and Hingorani, S.R. (2015). RUNX3 Controls a Metastatic Switch in Pancreatic Ductal Adenocarcinoma. Cell *161*, 1345-1360.
35. Biankin, A.V., Waddell, N., Kassahn, K.S., Gingras, M.C., Muthuswamy, L.B., Johns, A.L., Miller, D.K., Wilson, P.J., Patch, A.M., Wu, J., Chang, D.K., Cowley, M.J., Gardiner, B.B., Song, S., Harliwong, I., Idrisoglu, S., Nourse, C., Nourbakhsh, E., Manning, S., Wani, S., Gongora, M., Pajic, M., Scarlett, C.J., Gill, A.J., Pinho, A.V., Rooman, I., Anderson, M., Holmes, O., Leonard, C., Taylor, D., Wood, S., Xu, Q., Nones, K., Fink, J.L., Christ, A., Bruxner, T., Cloonan, N., Kolle, G., Newell, F., Pinese, M., Mead, R.S., Humphris, J.L., Kaplan, W., Jones, M.D., Colvin, E.K., Nagrial, A.M., Humphrey, E.S., Chou, A., Chin, V.T., Chantrill, L.A., Mawson, A., Samra, J.S., Kench, J.G., Lovell, J.A., Daly, R.J., Merrett, N.D., Toon, C., Epari, K., Nguyen, N.Q., Barbour, A., Zeps, N., Kakkar, N., Zhao, F., Wu, Y.Q., Wang, M., Muzny, D.M., Fisher, W.E., Brunicardi, F.C., Hodges, S.E., Reid, J.G., Drummond, J., Chang, K., Han, Y., Lewis, L.R., Dinh, H., Buhay, C.J., Beck, T., Timms, L., Sam, M., Begley, K., Brown, A., Pai, D., Panchal, A., Buchner, N., De Borja, R., Denroche, R.E., Yung, C.K., Serra, S., Onetto, N., Mukhopadhyay, D., Tsao, M.S., Shaw, P.A., Petersen, G.M., Gallinger, S., Hruban, R.H., Maitra, A., Iacobuzio-Donahue, C.A., Schulick, R.D., Wolfgang, C.L., Morgan, R.A., Lawlor, R.T., Capelli, P., Corbo, V., Scardoni, M., Tortora, G., Tempero, M.A., Mann, K.M., Jenkins, N.A., Perez-Mancera, P.A., Adams, D.J., Largaespada, D.A., Wessels, L.F., Rust, A.G., Stein, L.D., Tuveson, D.A., Copeland, N.G., Musgrove, E.A., Scarpa, A., Eshleman, J.R., Hudson, T.J., Sutherland, R.L., Wheeler, D.A., Pearson, J.V., McPherson, J.D., Gibbs, R.A., and Grimmond, S.M. (2012). Pancreatic cancer genomes reveal aberrations in axon guidance pathway genes. Nature *491*, 399-405.
36. Waddell, N., Pajic, M., Patch, A.M., Chang, D.K., Kassahn, K.S., Bailey, P., Johns, A.L., Miller, D., Nones, K., Quek, K., Quinn, M.C., Robertson, A.J., Fadlullah, M.Z., Bruxner, T.J., Christ, A.N., Harliwong, I., Idrisoglu, S., Manning, S., Nourse, C., Nourbakhsh, E., Wani, S., Wilson, P.J., Markham, E., Cloonan, N., Anderson, M.J., Fink, J.L., Holmes, O., Kazakoff, S.H., Leonard, C., Newell, F., Poudel, B., Song, S., Taylor, D., Waddell, N., Wood, S., Xu, Q., Wu, J., Pinese, M., Cowley, M.J., Lee, H.C., Jones, M.D., Nagrial, A.M., Humphris, J., Chantrill, L.A., Chin, V., Steinmann, A.M., Mawson, A., Humphrey, E.S., Colvin, E.K., Chou, A., Scarlett, C.J., Pinho, A.V., Giry-Laterriere, M., Rooman, I., Samra, J.S., Kench, J.G., Pettitt, J.A., Merrett, N.D., Toon, C., Epari, K., Nguyen, N.Q., Barbour, A., Zeps, N., Jamieson, N.B., Graham, J.S., Niclou, S.P., Bjerkvig, R., Grutzmann, R., Aust, D., Hruban, R.H., Maitra, A., Iacobuzio-Donahue, C.A., Wolfgang, C.L., Morgan, R.A., Lawlor, R.T., Corbo, V., Bassi, C., Falconi, M., Zamboni, G., Tortora, G., Tempero, M.A., Gill, A.J., Eshleman, J.R., Pilarsky, C., Scarpa, A., Musgrove, E.A., Pearson, J.V., Biankin, A.V., and Grimmond, S.M. (2015). Whole genomes redefine the mutational landscape of pancreatic cancer. Nature *518*, 495-501.
37. Notta, F., Chan-Seng-Yue, M., Lemire, M., Li, Y., Wilson, G.W., Connor, A.A., Denroche, R.E., Liang, S.B., Brown, A.M., Kim, J.C., Wang, T., Simpson, J.T., Beck, T., Borgida, A., Buchner, N., Chadwick, D., Hafezi-Bakhtiari, S., Dick, J.E., Heisler, L., Hollingsworth, M.A., Ibrahimov, E., Jang, G.H., Johns, J., Jorgensen, L.G., Law, C., Ludkovski, O., Lungu, I., Ng, K., Pasternack, D., Petersen, G.M., Shlush, L.I., Timms, L., Tsao, M.S., Wilson, J.M., Yung, C.K., Zogopoulos, G., Bartlett, J.M., Alexandrov, L.B., Real, F.X., Cleary, S.P., Roehrl, M.H., McPherson, J.D., Stein, L.D., Hudson, T.J., Campbell, P.J., and Gallinger, S. (2016). A renewed model of pancreatic cancer evolution based on genomic rearrangement patterns. Nature *538*, 378-382.
38. Stephens, P.J., Greenman, C.D., Fu, B., Yang, F., Bignell, G.R., Mudie, L.J., Pleasance, E.D., Lau, K.W., Beare, D., Stebbings, L.A., McLaren, S., Lin, M.L., McBride, D.J., Varela, I., Nik-Zainal, S., Leroy, C., Jia, M., Menzies, A., Butler, A.P., Teague, J.W., Quail, M.A., Burton, J., Swerdlow, H., Carter, N.P., Morsberger, L.A., Iacobuzio-Donahue, C., Follows, G.A., Green, A.R., Flanagan, A.M., Stratton, M.R., Futreal, P.A., and Campbell, P.J. (2011). Massive genomic rearrangement acquired in a single catastrophic event during cancer development. Cell *144*, 27-40.
39. Korbel, J.O., and Campbell, P.J. (2013). Criteria for inference of chromothripsis in cancer genomes. Cell *152*, 1226-1236.
40. Rausch, T., Jones, D.T., Zapatka, M., Stutz, A.M., Zichner, T., Weischenfeldt, J., Jager, N., Remke, M., Shih, D., Northcott, P.A., Pfaff, E., Tica, J., Wang, Q., Massimi, L., Witt, H., Bender, S., Pleier, S., Cin, H., Hawkins, C., Beck, C., von Deimling, A., Hans, V., Brors, B., Eils, R., Scheurlen, W., Blake, J., Benes, V., Kulozik, A.E., Witt, O., Martin, D., Zhang, C., Porat, R., Merino, D.M., Wasserman, J., Jabado, N., Fontebasso, A., Bullinger, L., Rucker, F.G., Dohner, K., Dohner, H., Koster, J., Molenaar, J.J., Versteeg, R., Kool, M., Tabori, U., Malkin, D., Korshunov, A., Taylor, M.D., Lichter, P., Pfister, S.M., and Korbel, J.O. (2012). Genome sequencing of pediatric medulloblastoma links catastrophic DNA rearrangements with TP53 mutations. Cell *148*, 59-71.
41. Bailey, P., Chang, D.K., Nones, K., Johns, A.L., Patch, A.M., Gingras, M.C., Miller, D.K., Christ, A.N., Bruxner, T.J., Quinn, M.C., Nourse, C., Murtaugh, L.C., Harliwong, I., Idrisoglu, S., Manning, S., Nourbakhsh, E., Wani, S., Fink, L., Holmes, O., Chin, V., Anderson, M.J., Kazakoff, S., Leonard, C., Newell, F., Waddell, N., Wood, S., Xu, Q., Wilson, P.J., Cloonan, N., Kassahn, K.S., Taylor, D., Quek, K., Robertson, A., Pantano, L., Mincarelli, L., Sanchez, L.N., Evers, L., Wu, J., Pinese, M., Cowley, M.J., Jones, M.D., Colvin, E.K., Nagrial, A.M., Humphrey, E.S., Chantrill, L.A., Mawson, A., Humphris, J., Chou, A., Pajic, M., Scarlett, C.J., Pinho, A.V., Giry-Laterriere, M., Rooman, I., Samra, J.S., Kench, J.G., Lovell, J.A., Merrett, N.D., Toon, C.W., Epari, K., Nguyen, N.Q., Barbour,

A., Zeps, N., Moran-Jones, K., Jamieson, N.B., Graham, J.S., Duthie, F., Oien, K., Hair, J., Grutzmann, R., Maitra, A., Iacobuzio-Donahue, C.A., Wolfgang, C.L., Morgan, R.A., Lawlor, R.T., Corbo, V., Bassi, C., Rusev, B., Capelli, P., Salvia, R., Tortora, G., Mukhopadhyay, D., Petersen, G.M., Munzy, D.M., Fisher, W.E., Karim, S.A., Eshleman, J.R., Hruban, R.H., Pilarsky, C., Morton, J.P., Sansom, O.J., Scarpa, A., Musgrove, E.A., Bailey, U.M., Hofmann, O., Sutherland, R.L., Wheeler, D.A., Gill, A.J., Gibbs, R.A., Pearson, J.V., Waddell, N., Biankin, A.V., and Grimmond, S.M. (2016). Genomic analyses identify molecular subtypes of pancreatic cancer. Nature *531*, 47-52.

42. Witkiewicz, A.K., McMillan, E.A., Balaji, U., Baek, G., Lin, W.C., Mansour, J., Mollaee, M., Wagner, K.U., Koduru, P., Yopp, A., Choti, M.A., Yeo, C.J., McCue, P., White, M.A., and Knudsen, E.S. (2015). Whole-exome sequencing of pancreatic cancer defines genetic diversity and therapeutic targets. Nature communications *6*, 6744.
43. Makohon-Moore, A.P., Zhang, M., Reiter, J.G., Bozic, I., Allen, B., Kundu, D., Chatterjee, K., Wong, F., Jiao, Y., Kohutek, Z.A., Hong, J., Attiyeh, M., Javier, B., Wood, L.D., Hruban, R.H., Nowak, M.A., Papadopoulos, N., Kinzler, K.W., Vogelstein, B., and Iacobuzio-Donahue, C.A. (2017). Limited heterogeneity of known driver gene mutations among the metastases of individual patients with pancreatic cancer. Nature genetics *49*, 358-366.
44. McDonald, O.G., Li, X., Saunders, T., Tryggvadottir, R., Mentch, S.J., Warmoes, M.O., Word, A.E., Carrer, A., Salz, T.H., Natsume, S., Stauffer, K.M., Makohon-Moore, A., Zhong, Y., Wu, H., Wellen, K.E., Locasale, J.W., Iacobuzio-Donahue, C.A., and Feinberg, A.P. (2017). Epigenomic reprogramming during pancreatic cancer progression links anabolic glucose metabolism to distant metastasis. Nature genetics *49*, 367-376.
45. Yachida, S., and Iacobuzio-Donahue, C.A. (2009). The pathology and genetics of metastatic pancreatic cancer. Archives of pathology & laboratory medicine *133*, 413-422.
46. Cid-Arregui, A., and Juarez, V. (2015). Perspectives in the treatment of pancreatic adenocarcinoma. World journal of gastroenterology *21*, 9297-9316.
47. Cowley, M.J., Chang, D.K., Pajic, M., Johns, A.L., Waddell, N., Grimmond, S.M., and Biankin, A.V. (2013). Understanding pancreatic cancer genomes. Journal of hepato-biliary-pancreatic sciences *20*, 549-556.
48. Von Hoff, D.D., Ervin, T., Arena, F.P., Chiorean, E.G., Infante, J., Moore, M., Seay, T., Tjulandin, S.A., Ma, W.W., Saleh, M.N., Harris, M., Reni, M., Dowden, S., Laheru, D., Bahary, N., Ramanathan, R.K., Tabernero, J., Hidalgo, M., Goldstein, D., Van Cutsem, E., Wei, X., Iglesias, J., and Renschler, M.F. (2013). Increased survival in pancreatic cancer with nab-paclitaxel plus gemcitabine. The New England journal of medicine *369*, 1691-1703.
49. Jonkers, J., and Berns, A. (2002). Conditional mouse models of sporadic cancer. Nature reviews Cancer *2*, 251-265.
50. Stewart, T.A., Pattengale, P.K., and Leder, P. (1984). Spontaneous mammary adenocarcinomas in transgenic mice that carry and express MTV/myc fusion genes. Cell *38*, 627-637.
51. Brinster, R.L., Chen, H.Y., Messing, A., van Dyke, T., Levine, A.J., and Palmiter, R.D. (1984). Transgenic mice harboring SV40 T-antigen genes develop characteristic brain tumors. Cell *37*, 367-379.
52. Robertson, E., Bradley, A., Kuehn, M., and Evans, M. (1986). Germ-line transmission of genes introduced into cultured pluripotential cells by retroviral vector. Nature *323*, 445-448.
53. Sauer, B. (1998). Inducible gene targeting in mice using the Cre/lox system. Methods (San Diego, Calif) *14*, 381-392.
54. Corbett, T.H., Roberts, B.J., Leopold, W.R., Peckham, J.C., Wilkoff, L.J., Griswold, D.P., Jr., and Schabel, F.M., Jr. (1984). Induction and chemotherapeutic response of two transplantable ductal adenocarcinomas of the pancreas in C57BL/6 mice. Cancer research *44*, 717-726.
55. Longnecker, D.S. (1984). Lesions induced in rodent pancreas by azaserine and other pancreatic carcinogens. Environmental health perspectives *56*, 245-251.
56. Jackson, E.L., Willis, N., Mercer, K., Bronson, R.T., Crowley, D., Montoya, R., Jacks, T., and Tuveson, D.A. (2001). Analysis of lung tumor initiation and progression using conditional expression of oncogenic K-ras. Genes & development *15*, 3243-3248.
57. Hingorani, S.R., Petricoin, E.F., Maitra, A., Rajapakse, V., King, C., Jacobetz, M.A., Ross, S., Conrads, T.P., Veenstra, T.D., Hitt, B.A., Kawaguchi, Y., Johann, D., Liotta, L.A., Crawford, H.C., Putt, M.E., Jacks, T., Wright, C.V., Hruban, R.H., Lowy, A.M., and Tuveson, D.A. (2003). Preinvasive and invasive ductal pancreatic cancer and its early detection in the mouse. Cancer cell *4*, 437-450.
58. Aguirre, A.J., Bardeesy, N., Sinha, M., Lopez, L., Tuveson, D.A., Horner, J., Redston, M.S., and DePinho, R.A. (2003). Activated Kras and Ink4a/Arf deficiency cooperate to produce metastatic pancreatic ductal adenocarcinoma. Genes & development *17*, 3112-3126.
59. Hingorani, S.R., Wang, L., Multani, A.S., Combs, C., Deramaudt, T.B., Hruban, R.H., Rustgi, A.K., Chang, S., and Tuveson, D.A. (2005). Trp53R172H and KrasG12D cooperate to promote chromosomal instability and widely metastatic pancreatic ductal adenocarcinoma in mice. Cancer cell *7*, 469-483.
60. Eser, S., Reiff, N., Messer, M., Seidler, B., Gottschalk, K., Dobler, M., Hieber, M., Arbeiter, A., Klein, S., Kong, B., Michalski, C.W., Schlitter, A.M., Esposito, I., Kind, A.J., Rad, L., Schnieke, A.E., Baccarini, M., Alessi, D.R., Rad, R., Schmid, R.M., Schneider, G., and Saur, D. (2013). Selective requirement of PI3K/PDK1 signaling for Kras oncogene-driven pancreatic cell plasticity and cancer. Cancer cell *23*, 406-420.
61. Bardeesy, N., Cheng, K.H., Berger, J.H., Chu, G.C., Pahler, J., Olson, P., Hezel, A.F., Horner, J., Lauwers, G.Y., Hanahan, D., and DePinho, R.A. (2006b). Smad4 is dispensable for normal pancreas development yet critical in progression and tumor biology of pancreas cancer. Genes & development *20*, 3130-3146.
62. Ijichi, H., Chytil, A., Gorska, A.E., Aakre, M.E., Fujitani, Y., Fujitani, S., Wright, C.V., and Moses, H.L. (2006). Aggressive pancreatic ductal adenocarcinoma in mice caused by pancreas-specific blockade of transforming growth factor-beta signaling in cooperation with active Kras expression. Genes & development *20*, 3147-3160.
63. Rad, R., Rad, L., Wang, W., Strong, A., Ponstingl, H., Bronner, I.F., Mayho, M., Steiger, K., Weber, J., Hieber, M., Veltkamp, C., Eser, S., Geumann, U., Ollinger, R., Zukowska, M., Barenboim, M., Maresch, R., Cadinanos, J., Friedrich, M., Varela, I., Constantino-Casas, F., Sarver, A., Ten Hoeve, J., Prosser, H., Seidler, B., Bauer, J., Heikenwalder, M., Metzakopian, E., Krug, A., Ehmer, U., Schneider, G., Knosel, T., Rummele, P., Aust, D., Grutzmann, R., Pilarsky, C., Ning, Z., Wessels, L., Schmid, R.M., Quail, M.A., Vassiliou, G., Esposito, I., Liu, P., Saur, D., and Bradley, A. (2015). A conditional piggyBac transposition system for genetic screening in mice identifies oncogenic networks in pancreatic cancer. Nature genetics *47*, 47-56.
64. Olive, K.P., Tuveson, D.A., Ruhe, Z.C., Yin, B., Willis, N.A., Bronson, R.T., Crowley, D., and Jacks, T. (2004). Mutant p53 gain of function in two mouse models of Li-Fraumeni syndrome. Cell *119*, 847-860.

65. Morton, J.P., Timpson, P., Karim, S.A., Ridgway, R.A., Athineos, D., Doyle, B., Jamieson, N.B., Oien, K.A., Lowy, A.M., Brunton, V.G., Frame, M.C., Evans, T.R., and Sansom, O.J. (2010). Mutant p53 drives metastasis and overcomes growth arrest/senescence in pancreatic cancer. Proceedings of the National Academy of Sciences of the United States of America *107*, 246-251.
66. Weissmueller, S., Manchado, E., Saborowski, M., Morris, J.P.t., Wagenblast, E., Davis, C.A., Moon, S.H., Pfister, N.T., Tschaharganeh, D.F., Kitzing, T., Aust, D., Markert, E.K., Wu, J., Grimmond, S.M., Pilarsky, C., Prives, C., Biankin, A.V., and Lowe, S.W. (2014). Mutant p53 drives pancreatic cancer metastasis through cell-autonomous PDGF receptor beta signaling. Cell *157*, 382-394.
67. Attardi, L.D., and Donehower, L.A. (2005). Probing p53 biological functions through the use of genetically engineered mouse models. Mutation research *576*, 4-21.
68. Copeland, N.G., and Jenkins, N.A. (2010). Harnessing transposons for cancer gene discovery. Nature reviews Cancer *10*, 696-706.
69. Perez-Mancera, P.A., Rust, A.G., van der Weyden, L., Kristiansen, G., Li, A., Sarver, A.L., Silverstein, K.A., Grutzmann, R., Aust, D., Rummele, P., Knosel, T., Herd, C., Stemple, D.L., Kettleborough, R., Brosnan, J.A., Li, A., Morgan, R., Knight, S., Yu, J., Stegeman, S., Collier, L.S., ten Hoeve, J.J., de Ridder, J., Klein, A.P., Goggins, M., Hruban, R.H., Chang, D.K., Biankin, A.V., Grimmond, S.M., Wessels, L.F., Wood, S.A., Iacobuzio-Donahue, C.A., Pilarsky, C., Largaespada, D.A., Adams, D.J., and Tuveson, D.A. (2012). The deubiquitinase USP9X suppresses pancreatic ductal adenocarcinoma. Nature *486*, 266-270.
70. Ding, S., Wu, X., Li, G., Han, M., Zhuang, Y., and Xu, T. (2005). Efficient transposition of the piggyBac (PB) transposon in mammalian cells and mice. Cell *122*, 473-483.
71. Rad, R., Rad, L., Wang, W., Cadinanos, J., Vassiliou, G., Rice, S., Campos, L.S., Yusa, K., Banerjee, R., Li, M.A., de la Rosa, J., Strong, A., Lu, D., Ellis, P., Conte, N., Yang, F.T., Liu, P., and Bradley, A. (2010). PiggyBac transposon mutagenesis: a tool for cancer gene discovery in mice. Science (New York, NY) *330*, 1104-1107.
72. Barrangou, R., Fremaux, C., Deveau, H., Richards, M., Boyaval, P., Moineau, S., Romero, D.A., and Horvath, P. (2007). CRISPR provides acquired resistance against viruses in prokaryotes. Science (New York, NY) *315*, 1709-1712.
73. Mali, P., Esvelt, K.M., and Church, G.M. (2013). Cas9 as a versatile tool for engineering biology. Nature methods *10*, 957-963.
74. Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J.A., and Charpentier, E. (2012). A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. Science (New York, NY) *337*, 816-821.
75. Wiedenheft, B., Sternberg, S.H., and Doudna, J.A. (2012). RNA-guided genetic silencing systems in bacteria and archaea. Nature *482*, 331-338.
76. Cong, L., Ran, F.A., Cox, D., Lin, S., Barretto, R., Habib, N., Hsu, P.D., Wu, X., Jiang, W., Marraffini, L.A., and Zhang, F. (2013). Multiplex genome engineering using CRISPR/Cas systems. Science (New York, NY) *339*, 819-823.
77. Platt, R.J., Chen, S., Zhou, Y., Yim, M.J., Swiech, L., Kempton, H.R., Dahlman, J.E., Parnas, O., Eisenhaure, T.M., Jovanovic, M., Graham, D.B., Jhunjhunwala, S., Heidenreich, M., Xavier, R.J., Langer, R., Anderson, D.G., Hacohen, N., Regev, A., Feng, G., Sharp, P.A., and Zhang, F. (2014). CRISPR-Cas9 knockin mice for genome editing and cancer modeling. Cell *159*, 440-455.
78. Sanchez-Rivera, F.J., Papagiannakopoulos, T., Romero, R., Tammela, T., Bauer, M.R., Bhutkar, A., Joshi, N.S., Subbaraj, L., Bronson, R.T., Xue, W., and Jacks, T. (2014). Rapid modelling of cooperating genetic events in cancer through somatic genome editing. Nature *516*, 428-431.
79. Xue, W., Chen, S., Yin, H., Tammela, T., Papagiannakopoulos, T., Joshi, N.S., Cai, W., Yang, G., Bronson, R., Crowley, D.G., Zhang, F., Anderson, D.G., Sharp, P.A., and Jacks, T. (2014). CRISPR-mediated direct mutation of cancer genes in the mouse liver. Nature *514*, 380-384.
80. Weber, J., Ollinger, R., Friedrich, M., Ehmer, U., Barenboim, M., Steiger, K., Heid, I., Mueller, S., Maresch, R., Engleitner, T., Gross, N., Geumann, U., Fu, B., Segler, A., Yuan, D., Lange, S., Strong, A., de la Rosa, J., Esposito, I., Liu, P., Cadinanos, J., Vassiliou, G.S., Schmid, R.M., Schneider, G., Unger, K., Yang, F., Braren, R., Heikenwalder, M., Varela, I., Saur, D., Bradley, A., and Rad, R. (2015). CRISPR/Cas9 somatic multiplex-mutagenesis for high-throughput functional cancer genomics in mice. Proceedings of the National Academy of Sciences of the United States of America *112*, 13982-13987.
81. Chiou, S.H., Winters, I.P., Wang, J., Naranjo, S., Dudgeon, C., Tamburini, F.B., Brady, J.J., Yang, D., Gruner, B.M., Chuang, C.H., Caswell, D.R., Zeng, H., Chu, P., Kim, G.E., Carpizo, D.R., Kim, S.K., and Winslow, M.M. (2015). Pancreatic cancer modeling using retrograde viral vector delivery and in vivo CRISPR/Cas9-mediated somatic genome editing. Genes & development *29*, 1576-1585.
82. Mazur, P.K., Herner, A., Mello, S.S., Wirth, M., Hausmann, S., Sanchez-Rivera, F.J., Lofgren, S.M., Kuschma, T., Hahn, S.A., Vangala, D., Trajkovic-Arsic, M., Gupta, A., Heid, I., Noel, P.B., Braren, R., Erkan, M., Kleeff, J., Sipos, B., Sayles, L.C., Heikenwalder, M., Hessmann, E., Ellenrieder, V., Esposito, I., Jacks, T., Bradner, J.E., Khatri, P., Sweet-Cordero, E.A., Attardi, L.D., Schmid, R.M., Schneider, G., Sage, J., and Siveke, J.T. (2015). Combined inhibition of BET family proteins and histone deacetylases as a potential epigenetics-based therapy for pancreatic ductal adenocarcinoma. Nature medicine *21*, 1163-1171.
83. Maresch, R., Mueller, S., Veltkamp, C., Ollinger, R., Friedrich, M., Heid, I., Steiger, K., Weber, J., Engleitner, T., Barenboim, M., Klein, S., Louzada, S., Banerjee, R., Strong, A., Stauber, T., Gross, N., Geumann, U., Lange, S., Ringelhan, M., Varela, I., Unger, K., Yang, F., Schmid, R.M., Vassiliou, G.S., Braren, R., Schneider, G., Heikenwalder, M., Bradley, A., Saur, D., and Rad, R. (2016). Multiplexed pancreatic genome engineering and cancer induction by transfection-based CRISPR/Cas9 delivery in mice. Nature communications *7*, 10770.
84. Schvartzman, J.M., Sotillo, R., and Benezra, R. (2010). Mitotic chromosomal instability and cancer: mouse modelling of the human disease. Nature reviews Cancer *10*, 102-115.
85. Gordon, D.J., Resio, B., and Pellman, D. (2012). Causes and consequences of aneuploidy in cancer. Nature reviews Genetics *13*, 189-203.
86. Mitelman, F., Johansson, B., and Mertens, F. (2007). The impact of translocations and gene fusions on cancer causation. Nature reviews Cancer *7*, 233-245.
87. Beroukhim, R., Mermel, C.H., Porter, D., Wei, G., Raychaudhuri, S., Donovan, J., Barretina, J., Boehm, J.S., Dobson, J., Urashima, M., Mc Henry, K.T., Pinchback, R.M., Ligon, A.H., Cho, Y.J., Haery, L., Greulich, H., Reich, M., Winckler, W., Lawrence, M.S., Weir, B.A., Tanaka, K.E., Chiang, D.Y., Bass, A.J., Loo, A., Hoffman, C.,

Prensner, J., Liefeld, T., Gao, Q., Yecies, D., Signoretti, S., Maher, E., Kaye, F.J., Sasaki, H., Tepper, J.E., Fletcher, J.A., Tabernero, J., Baselga, J., Tsao, M.S., Demichelis, F., Rubin, M.A., Janne, P.A., Daly, M.J., Nucera, C., Levine, R.L., Ebert, B.L., Gabriel, S., Rustgi, A.K., Antonescu, C.R., Ladanyi, M., Letai, A., Garraway, L.A., Loda, M., Beer, D.G., True, L.D., Okamoto, A., Pomeroy, S.L., Singer, S., Golub, T.R., Lander, E.S., Getz, G., Sellers, W.R., and Meyerson, M. (2010). The landscape of somatic copy-number alteration across human cancers. Nature *463*, 899-905.

88. Mueller, S., Engleitner, T., Maresch, R., Zukowska, M., Lange, S., Kaltenbacher, T., Konukiewitz, B., Ollinger, R., Zwiebel, M., Strong, A., Yen, H.Y., Banerjee, R., Louzada, S., Fu, B., Seidler, B., Gotzfried, J., Schuck, K., Hassan, Z., Arbeiter, A., Schonhuber, N., Klein, S., Veltkamp, C., Friedrich, M., Rad, L., Barenboim, M., Ziegenhain, C., Hess, J., Dovey, O.M., Eser, S., Parekh, S., Constantino-Casas, F., de la Rosa, J., Sierra, M.I., Fraga, M., Mayerle, J., Kloppel, G., Cadinanos, J., Liu, P., Vassiliou, G., Weichert, W., Steiger, K., Enard, W., Schmid, R.M., Yang, F., Unger, K., Schneider, G., Varela, I., Bradley, A., Saur, D., and Rad, R. (2018). Evolutionary routes and KRAS dosage define pancreatic cancer phenotypes. Nature *554*, 62-68.
89. Torres, E.M., Sokolsky, T., Tucker, C.M., Chan, L.Y., Boselli, M., Dunham, M.J., and Amon, A. (2007). Effects of aneuploidy on cellular physiology and cell division in haploid yeast. Science (New York, NY) *317*, 916-924.
90. Torres, E.M., Dephoure, N., Panneerselvam, A., Tucker, C.M., Whittaker, C.A., Gygi, S.P., Dunham, M.J., and Amon, A. (2010). Identification of aneuploidy-tolerating mutations. Cell *143*, 71-83.
91. Lyon, M.F. (1961). Gene action in the X-chromosome of the mouse (Mus musculus L.). Nature *190*, 372-373.
92. Boumil, R.M., and Lee, J.T. (2001). Forty years of decoding the silence in X-chromosome inactivation. Human molecular genetics *10*, 2225-2232.
93. Berger, A.H., Knudson, A.G., and Pandolfi, P.P. (2011). A continuum model for tumour suppression. Nature *476*, 163-169.
94. Knudson, A.G., Jr., Meadows, A.T., Nichols, W.W., and Hill, R. (1976). Chromosomal deletion and retinoblastoma. The New England journal of medicine *295*, 1120-1123.
95. Alimonti, A., Carracedo, A., Clohessy, J.G., Trotman, L.C., Nardella, C., Egia, A., Salmena, L., Sampieri, K., Haveman, W.J., Brogi, E., Richardson, A.L., Zhang, J., and Pandolfi, P.P. (2010). Subtle variations in Pten dose determine cancer susceptibility. Nature genetics *42*, 454-458.
96. Evan, G.I., Wyllie, A.H., Gilbert, C.S., Littlewood, T.D., Land, H., Brooks, M., Waters, C.M., Penn, L.Z., and Hancock, D.C. (1992). Induction of apoptosis in fibroblasts by c-myc protein. Cell *69*, 119-128.
97. Serrano, M., Lin, A.W., McCurrach, M.E., Beach, D., and Lowe, S.W. (1997). Oncogenic ras provokes premature cell senescence associated with accumulation of p53 and p16INK4a. Cell *88*, 593-602.
98. Sarkisian, C.J., Keister, B.A., Stairs, D.B., Boxer, R.B., Moody, S.E., and Chodosh, L.A. (2007). Dose-dependent oncogene-induced senescence in vivo and its evasion during mammary tumorigenesis. Nature cell biology *9*, 493-505.
99. Feldser, D.M., Kostova, K.K., Winslow, M.M., Taylor, S.E., Cashman, C., Whittaker, C.A., Sanchez-Rivera, F.J., Resnick, R., Bronson, R., Hemann, M.T., and Jacks, T. (2010). Stage-specific sensitivity to p53 restoration during lung cancer progression. Nature *468*, 572-575.
100. Junttila, M.R., Karnezis, A.N., Garcia, D., Madriles, F., Kortlever, R.M., Rostker, F., Brown Swigart, L., Pham, D.M., Seo, Y., Evan, G.I., and Martins, C.P. (2010). Selective activation of p53-mediated tumour suppression in high-grade tumours. Nature *468*, 567-571.
101. Murphy, D.J., Junttila, M.R., Pouyet, L., Karnezis, A., Shchors, K., Bui, D.A., Brown-Swigart, L., Johnson, L., and Evan, G.I. (2008). Distinct thresholds govern Myc's biological output in vivo. Cancer cell *14*, 447-457.
102. Rad, R., Cadinanos, J., Rad, L., Varela, I., Strong, A., Kriegl, L., Constantino-Casas, F., Eser, S., Hieber, M., Seidler, B., Price, S., Fraga, M.F., Calvanese, V., Hoffman, G., Ponstingl, H., Schneider, G., Yusa, K., Grove, C., Schmid, R.M., Wang, W., Vassiliou, G., Kirchner, T., McDermott, U., Liu, P., Saur, D., and Bradley, A. (2013). A genetic progression model of Braf(V600E)-induced intestinal tumorigenesis reveals targets for therapeutic intervention. Cancer cell *24*, 15-29.
103. Krasinskas, A.M., Moser, A.J., Saka, B., Adsay, N.V., and Chiosea, S.I. (2013). KRAS mutant allele-specific imbalance is associated with worse prognosis in pancreatic cancer and progression to undifferentiated carcinoma of the pancreas. Modern pathology : an official journal of the United States and Canadian Academy of Pathology, Inc *26*, 1346-1354.
104. Qiu, W., Sahin, F., Iacobuzio-Donahue, C.A., Garcia-Carracedo, D., Wang, W.M., Kuo, C.Y., Chen, D., Arking, D.E., Lowy, A.M., Hruban, R.H., Remotti, H.E., and Su, G.H. (2011). Disruption of p16 and activation of Kras in pancreas increase ductal adenocarcinoma formation and metastasis in vivo. Oncotarget *2*, 862-873.
105. Makohon-Moore, A., and Iacobuzio-Donahue, C.A. (2016). Pancreatic cancer biology and genetics from an evolutionary perspective. Nature reviews Cancer *16*, 553-565.
106. Zhang, Z., Wang, Y., Vikis, H.G., Johnson, L., Liu, G., Li, J., Anderson, M.W., Sills, R.C., Hong, H.L., Devereux, T.R., Jacks, T., Guan, K.L., and You, M. (2001). Wildtype Kras2 can inhibit lung carcinogenesis in mice. Nature genetics *29*, 25-33.
107. Hruban, R.H., Adsay, N.V., Albores-Saavedra, J., Anver, M.R., Biankin, A.V., Boivin, G.P., Furth, E.E., Furukawa, T., Klein, A., Klimstra, D.S., Kloppel, G., Lauwers, G.Y., Longnecker, D.S., Luttges, J., Maitra, A., Offerhaus, G.J., Perez-Gallego, L., Redston, M., and Tuveson, D.A. (2006). Pathology of genetically engineered mouse models of pancreatic exocrine cancer: consensus report and recommendations. Cancer research *66*, 95-106.
108. Nakhai, H., Sel, S., Favor, J., Mendoza-Torres, L., Paulsen, F., Duncker, G.I., and Schmid, R.M. (2007). Ptf1a is essential for the differentiation of GABAergic and glycinergic amacrine cells and horizontal cells in the mouse retina. Development (Cambridge, England) *134*, 1151-1160.
109. Schonhuber, N., Seidler, B., Schuck, K., Veltkamp, C., Schachtler, C., Zukowska, M., Eser, S., Feyerabend, T.B., Paul, M.C., Eser, P., Klein, S., Lowy, A.M., Banerjee, R., Yang, F., Lee, C.L., Moding, E.J., Kirsch, D.G., Scheideler, A., Alessi, D.R., Varela, I., Bradley, A., Kind, A., Schnieke, A.E., Rodewald, H.R., Rad, R., Schmid, R.M., Schneider, G., and Saur, D. (2014). A next-generation dual-recombinase system for time- and host-specific targeting of pancreatic cancer. Nature medicine *20*, 1340-1347.
110. Jonkers, J., Meuwissen, R., van der Gulden, H., Peterse, H., van der Valk, M., and Berns, A. (2001). Synergistic tumor suppressor activity of BRCA2 and p53 in a conditional mouse model for breast cancer. Nature genetics *29*, 418-425.

111. Lee, C.L., Moding, E.J., Huang, X., Li, Y., Woodlief, L.Z., Rodrigues, R.C., Ma, Y., and Kirsch, D.G. (2012). Generation of primary tumors with Flp recombinase in FRT-flanked p53 mice. Disease models & mechanisms *5*, 397-402.
112. Chytil, A., Magnuson, M.A., Wright, C.V., and Moses, H.L. (2002). Conditional inactivation of the TGF-beta type II receptor using Cre:Lox. Genesis (New York, NY : 2000) *32*, 73-75.
113. Jentsch, I., Adler, I.D., Carter, N.P., and Speicher, M.R. (2001). Karyotyping mouse chromosomes by multiplex-FISH (M-FISH). Chromosome research : an international journal on the molecular, supramolecular and evolutionary aspects of chromosome biology *9*, 211-214.
114. Keane, T.M., Goodstadt, L., Danecek, P., White, M.A., Wong, K., Yalcin, B., Heger, A., Agam, A., Slater, G., Goodson, M., Furlotte, N.A., Eskin, E., Nellaker, C., Whitley, H., Cleak, J., Janowitz, D., Hernandez-Pliego, P., Edwards, A., Belgard, T.G., Oliver, P.L., McIntyre, R.E., Bhomra, A., Nicod, J., Gan, X., Yuan, W., van der Weyden, L., Steward, C.A., Bala, S., Stalker, J., Mott, R., Durbin, R., Jackson, I.J., Czechanski, A., Guerra-Assuncao, J.A., Donahue, L.R., Reinholdt, L.G., Payseur, B.A., Ponting, C.P., Birney, E., Flint, J., and Adams, D.J. (2011). Mouse genomic variation and its effect on phenotypes and gene regulation. Nature *477*, 289-294.
115. Ye, K., Schulz, M.H., Long, Q., Apweiler, R., and Ning, Z. (2009). Pindel: a pattern growth approach to detect break points of large deletions and medium sized insertions from paired-end short reads. Bioinformatics (Oxford, England) *25*, 2865-2871.
116. Barretina, J., Caponigro, G., Stransky, N., Venkatesan, K., Margolin, A.A., Kim, S., Wilson, C.J., Lehar, J., Kryukov, G.V., Sonkin, D., Reddy, A., Liu, M., Murray, L., Berger, M.F., Monahan, J.E., Morais, P., Meltzer, J., Korejwa, A., Jane-Valbuena, J., Mapa, F.A., Thibault, J., Bric-Furlong, E., Raman, P., Shipway, A., Engels, I.H., Cheng, J., Yu, G.K., Yu, J., Aspesi, P., Jr., de Silva, M., Jagtap, K., Jones, M.D., Wang, L., Hatton, C., Palescandolo, E., Gupta, S., Mahan, S., Sougnez, C., Onofrio, R.C., Liefeld, T., MacConaill, L., Winckler, W., Reich, M., Li, N., Mesirov, J.P., Gabriel, S.B., Getz, G., Ardlie, K., Chan, V., Myer, V.E., Weber, B.L., Porter, J., Warmuth, M., Finan, P., Harris, J.L., Meyerson, M., Golub, T.R., Morrissey, M.P., Sellers, W.R., Schlegel, R., and Garraway, L.A. (2012). The Cancer Cell Line Encyclopedia enables predictive modelling of anticancer drug sensitivity. Nature *483*, 603-607.
117. Alexandrov, L.B., Nik-Zainal, S., Wedge, D.C., Aparicio, S.A., Behjati, S., Biankin, A.V., Bignell, G.R., Bolli, N., Borg, A., Borresen-Dale, A.L., Boyault, S., Burkhardt, B., Butler, A.P., Caldas, C., Davies, H.R., Desmedt, C., Eils, R., Eyfjord, J.E., Foekens, J.A., Greaves, M., Hosoda, F., Hutter, B., Ilicic, T., Imbeaud, S., Imielinski, M., Jager, N., Jones, D.T., Jones, D., Knappskog, S., Kool, M., Lakhani, S.R., Lopez-Otin, C., Martin, S., Munshi, N.C., Nakamura, H., Northcott, P.A., Pajic, M., Papaemmanuil, E., Paradiso, A., Pearson, J.V., Puente, X.S., Raine, K., Ramakrishna, M., Richardson, A.L., Richter, J., Rosenstiel, P., Schlesner, M., Schumacher, T.N., Span, P.N., Teague, J.W., Totoki, Y., Tutt, A.N., Valdes-Mas, R., van Buuren, M.M., van 't Veer, L., Vincent-Salomon, A., Waddell, N., Yates, L.R., Zucman-Rossi, J., Futreal, P.A., McDermott, U., Lichter, P., Meyerson, M., Grimmond, S.M., Siebert, R., Campo, E., Shibata, T., Pfister, S.M., Campbell, P.J., and Stratton, M.R. (2013a). Signatures of mutational processes in human cancer. Nature *500*, 415-421.
118. Parekh, S., Ziegenhain, C., Vieth, B., Enard, W., and Hellmann, I. (2016). The impact of amplification on differential expression analyses by RNA-seq. Scientific reports *6*, 25533.
119. Macosko, E.Z., Basu, A., Satija, R., Nemesh, J., Shekhar, K., Goldman, M., Tirosh, I., Bialas, A.R., Kamitaki, N., Martersteck, E.M., Trombetta, J.J., Weitz, D.A., Sanes, J.R., Shalek, A.K., Regev, A., and McCarroll, S.A. (2015). Highly Parallel Genome-wide Expression Profiling of Individual Cells Using Nanoliter Droplets. Cell *161*, 1202-1214.
120. Collisson, E.A., Sadanandam, A., Olson, P., Gibb, W.J., Truitt, M., Gu, S., Cooc, J., Weinkle, J., Kim, G.E., Jakkula, L., Feiler, H.S., Ko, A.H., Olshen, A.B., Danenberg, K.L., Tempero, M.A., Spellman, P.T., Hanahan, D., and Gray, J.W. (2011). Subtypes of pancreatic ductal adenocarcinoma and their differing responses to therapy. Nature medicine *17*, 500-503.
121. Moffitt, R.A., Marayati, R., Flate, E.L., Volmar, K.E., Loeza, S.G., Hoadley, K.A., Rashid, N.U., Williams, L.A., Eaton, S.C., Chung, A.H., Smyla, J.K., Anderson, J.M., Kim, H.J., Bentrem, D.J., Talamonti, M.S., Iacobuzio-Donahue, C.A., Hollingsworth, M.A., and Yeh, J.J. (2015). Virtual microdissection identifies distinct tumor- and stroma-specific subtypes of pancreatic ductal adenocarcinoma. Nature genetics *47*, 1168-1178.
122. Schaefer, C.F., Anthony, K., Krupa, S., Buchoff, J., Day, M., Hannay, T., and Buetow, K.H. (2009). PID: the Pathway Interaction Database. Nucleic acids research *37*, D674-679.
123. Du, P., Kibbe, W.A., and Lin, S.M. (2008). lumi: a pipeline for processing Illumina microarray. Bioinformatics (Oxford, England) *24*, 1547-1548.
124. Ritchie, M.E., Phipson, B., Wu, D., Hu, Y., Law, C.W., Shi, W., and Smyth, G.K. (2015). limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic acids research *43*, e47.
125. Friedrich, M.J., Rad, L., Bronner, I.F., Strong, A., Wang, W., Weber, J., Mayho, M., Ponstingl, H., Engleitner, T., Grove, C., Pfaus, A., Saur, D., Cadinanos, J., Quail, M.A., Vassiliou, G.S., Liu, P., Bradley, A., and Rad, R. (2017). Genome-wide transposon screening and quantitative insertion site sequencing for cancer gene discovery in mice. Nature protocols *12*, 289-309.
126. Meerbrey, K.L., Hu, G., Kessler, J.D., Roarty, K., Li, M.Z., Fang, J.E., Herschkowitz, J.I., Burrows, A.E., Ciccia, A., Sun, T., Schmitt, E.M., Bernardi, R.J., Fu, X., Bland, C.S., Cooper, T.A., Schiff, R., Rosen, J.M., Westbrook, T.F., and Elledge, S.J. (2011). The pINDUCER lentiviral toolkit for inducible RNA interference in vitro and in vivo. Proceedings of the National Academy of Sciences of the United States of America *108*, 3665-3670.
127. Kutner, R.H., Zhang, X.Y., and Reiser, J. (2009). Production, concentration and titration of pseudotyped HIV-1-based lentiviral vectors. Nature protocols *4*, 495-505.
128. Mann, K.M., Ward, J.M., Yew, C.C., Kovochich, A., Dawson, D.W., Black, M.A., Brett, B.T., Sheetz, T.E., Dupuy, A.J., Chang, D.K., Biankin, A.V., Waddell, N., Kassahn, K.S., Grimmond, S.M., Rust, A.G., Adams, D.J., Jenkins, N.A., and Copeland, N.G. (2012). Sleeping Beauty mutagenesis reveals cooperating mutations and pathways in pancreatic adenocarcinoma. Proceedings of the National Academy of Sciences of the United States of America *109*, 5934-5941.
129. Tate, J.G., Bamford, S., Jubb, H.C., Sondka, Z., Beare, D.M., Bindal, N., Boutselakis, H., Cole, C.G., Creatore, C., Dawson, E., Fish, P., Harsha, B., Hathaway, C., Jupe, S.C., Kok, C.Y., Noble, K., Ponting, L., Ramshaw, C.C., Rye, C.E., Speedy, H.E., Stefancsik, R., Thompson, S.L., Wang, S., Ward, S., Campbell, P.J., and Forbes, S.A. (2019). COSMIC: the Catalogue Of Somatic Mutations In Cancer. Nucleic acids research *47*, D941-D947.

130. Alexandrov, L.B., Nik-Zainal, S., Wedge, D.C., Campbell, P.J., and Stratton, M.R. (2013b). Deciphering signatures of mutational processes operative in human cancer. Cell reports *3*, 246-259.
131. Rosenthal, R., McGranahan, N., Herrero, J., Taylor, B.S., and Swanton, C. (2016). DeconstructSigs: delineating mutational processes in single tumors distinguishes DNA repair deficiencies and patterns of carcinoma evolution. Genome biology *17*, 31.
132. Forment, J.V., Kaidi, A., and Jackson, S.P. (2012). Chromothripsis and cancer: causes and consequences of chromosome shattering. Nature reviews Cancer *12*, 663-670.
133. McFadden, D.G., Papagiannakopoulos, T., Taylor-Weiner, A., Stewart, C., Carter, S.L., Cibulskis, K., Bhutkar, A., McKenna, A., Dooley, A., Vernon, A., Sougnez, C., Malstrom, S., Heimann, M., Park, J., Chen, F., Farago, A.F., Dayton, T., Shefler, E., Gabriel, S., Getz, G., and Jacks, T. (2014). Genetic and clonal dissection of murine small cell lung carcinoma progression by genome sequencing. Cell *156*, 1298-1311.
134. McFadden, D.G., Politi, K., Bhutkar, A., Chen, F.K., Song, X., Pirun, M., Santiago, P.M., Kim-Kiselak, C., Platt, J.T., Lee, E., Hodges, E., Rosebrock, A.P., Bronson, R.T., Socci, N.D., Hannon, G.J., Jacks, T., and Varmus, H. (2016). Mutational landscape of EGFR-, MYC-, and Kras-driven genetically engineered mouse models of lung adenocarcinoma. Proceedings of the National Academy of Sciences of the United States of America *113*, E6409-E6417.
135. Chung, W.J., Daemen, A., Cheng, J.H., Long, J.E., Cooper, J.E., Wang, B.E., Tran, C., Singh, M., Gnad, F., Modrusan, Z., Foreman, O., and Junttila, M.R. (2017). Kras mutant genetically engineered mouse models of human cancers are genomically heterogeneous. Proceedings of the National Academy of Sciences of the United States of America *114*, E10947-E10955.
136. Yamada, H., Sakamoto, H., Taira, M., Nishimura, S., Shimosato, Y., Terada, M., and Sugimura, T. (1986). Amplifications of both c-Ki-ras with a point mutation and c-myc in a primary pancreatic cancer and its metastatic tumors in lymph nodes. Japanese journal of cancer research : Gann *77*, 370-375.
137. Heidenblad, M., Jonson, T., Mahlamaki, E.H., Gorunova, L., Karhu, R., Johansson, B., and Hoglund, M. (2002). Detailed genomic mapping and expression analyses of 12p amplifications in pancreatic carcinomas reveal a 3.5-Mb target region for amplification. Genes, chromosomes & cancer *34*, 211-223.
138. Campbell, P.J., Yachida, S., Mudie, L.J., Stephens, P.J., Pleasance, E.D., Stebbings, L.A., Morsberger, L.A., Latimer, C., McLaren, S., Lin, M.L., McBride, D.J., Varela, I., Nik-Zainal, S.A., Leroy, C., Jia, M., Menzies, A., Butler, A.P., Teague, J.W., Griffin, C.A., Burton, J., Swerdlow, H., Quail, M.A., Stratton, M.R., Iacobuzio-Donahue, C., and Futreal, P.A. (2010). The patterns and dynamics of genomic instability in metastatic pancreatic cancer. Nature *467*, 1109-1113.
139. Yachida, S., Jones, S., Bozic, I., Antal, T., Leary, R., Fu, B., Kamiyama, M., Hruban, R.H., Eshleman, J.R., Nowak, M.A., Velculescu, V.E., Kinzler, K.W., Vogelstein, B., and Iacobuzio-Donahue, C.A. (2010). Distant metastasis occurs late during the genetic evolution of pancreatic cancer. Nature *467*, 1114-1117.
140. Kapoor, A., Yao, W., Ying, H., Hua, S., Liewen, A., Wang, Q., Zhong, Y., Wu, C.J., Sadanandam, A., Hu, B., Chang, Q., Chu, G.C., Al-Khalil, R., Jiang, S., Xia, H., Fletcher-Sananikone, E., Lim, C., Horwitz, G.I., Viale, A., Pettazzoni, P., Sanchez, N., Wang, H., Protopopov, A., Zhang, J., Heffernan, T., Johnson, R.L., Chin, L., Wang, Y.A., Draetta, G., and DePinho, R.A. (2014). Yap1 activation enables bypass of oncogenic Kras addiction in pancreatic cancer. Cell *158*, 185-197.
141. Shao, D.D., Xue, W., Krall, E.B., Bhutkar, A., Piccioni, F., Wang, X., Schinzel, A.C., Sood, S., Rosenbluh, J., Kim, J.W., Zwang, Y., Roberts, T.M., Root, D.E., Jacks, T., and Hahn, W.C. (2014). KRAS and YAP1 converge to regulate EMT and tumor survival. Cell *158*, 171-184.
142. Stellas, D., Szabolcs, M., Koul, S., Li, Z., Polyzos, A., Anagnostopoulos, C., Cournia, Z., Tamvakopoulos, C., Klinakis, A., and Efstratiadis, A. (2014). Therapeutic effects of an anti-Myc drug on mouse pancreatic cancer. Journal of the National Cancer Institute *106*.
143. Zhang, W., Nandakumar, N., Shi, Y., Manzano, M., Smith, A., Graham, G., Gupta, S., Vietsch, E.E., Laughlin, S.Z., Wadhwa, M., Chetram, M., Joshi, M., Wang, F., Kallakury, B., Toretsky, J., Wellstein, A., and Yi, C. (2014). Downstream of mutant KRAS, the transcription regulator YAP is essential for neoplastic progression to pancreatic ductal adenocarcinoma. Science signaling *7*, ra42.
144. Diersch, S., Wirth, M., Schneeweis, C., Jors, S., Geisler, F., Siveke, J.T., Rad, R., Schmid, R.M., Saur, D., Rustgi, A.K., Reichert, M., and Schneider, G. (2016). Kras(G12D) induces EGFR-MYC cross signaling in murine primary pancreatic ductal epithelial cells. Oncogene *35*, 3880-3886.
145. Cerami, E., Gao, J., Dogrusoz, U., Gross, B.E., Sumer, S.O., Aksoy, B.A., Jacobsen, A., Byrne, C.J., Heuer, M.L., Larsson, E., Antipin, Y., Reva, B., Goldberg, A.P., Sander, C., and Schultz, N. (2012). The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data. Cancer discovery *2*, 401-404.
146. Gao, J., Aksoy, B.A., Dogrusoz, U., Dresdner, G., Gross, B., Sumer, S.O., Sun, Y., Jacobsen, A., Sinha, R., Larsson, E., Cerami, E., Sander, C., and Schultz, N. (2013). Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal. Science signaling *6*, pl1.
147. Uhlen, M., Fagerberg, L., Hallstrom, B.M., Lindskog, C., Oksvold, P., Mardinoglu, A., Sivertsson, A., Kampf, C., Sjostedt, E., Asplund, A., Olsson, I., Edlund, K., Lundberg, E., Navani, S., Szigyarto, C.A., Odeberg, J., Djureinovic, D., Takanen, J.O., Hober, S., Alm, T., Edqvist, P.H., Berling, H., Tegel, H., Mulder, J., Rockberg, J., Nilsson, P., Schwenk, J.M., Hamsten, M., von Feilitzen, K., Forsberg, M., Persson, L., Johansson, F., Zwahlen, M., von Heijne, G., Nielsen, J., and Ponten, F. (2015). Proteomics. Tissue-based map of the human proteome. Science (New York, NY) *347*, 1260419.
148. Schneider, G., Saur, D., Siveke, J.T., Fritsch, R., Greten, F.R., and Schmid, R.M. (2006). IKKalpha controls p52/RelB at the skp2 gene promoter to regulate G1- to S-phase progression. The EMBO journal *25*, 3801-3812.
149. Hamidi, T., Algul, H., Cano, C.E., Sandi, M.J., Molejon, M.I., Riemann, M., Calvo, E.L., Lomberk, G., Dagorn, J.C., Weih, F., Urrutia, R., Schmid, R.M., and Iovanna, J.L. (2012). Nuclear protein 1 promotes pancreatic cancer development and protects cells from stress by inhibiting apoptosis. The Journal of clinical investigation *122*, 2092-2103.
150. Huang da, W., Sherman, B.T., and Lempicki, R.A. (2009a). Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nature protocols *4*, 44-57.
151. Huang da, W., Sherman, B.T., and Lempicki, R.A. (2009b). Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. Nucleic acids research *37*, 1-13.

152. Subramanian, A., Tamayo, P., Mootha, V.K., Mukherjee, S., Ebert, B.L., Gillette, M.A., Paulovich, A., Pomeroy, S.L., Golub, T.R., Lander, E.S., and Mesirov, J.P. (2005). Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proceedings of the National Academy of Sciences of the United States of America *102*, 15545-15550.
153. Nieto, M.A., Huang, R.Y., Jackson, R.A., and Thiery, J.P. (2016). EMT: 2016. Cell *166*, 21-45.
154. Winter, J.M., Ting, A.H., Vilardell, F., Gallmeier, E., Baylin, S.B., Hruban, R.H., Kern, S.E., and Iacobuzio-Donahue, C.A. (2008). Absence of E-cadherin expression distinguishes noncohesive from cohesive pancreatic cancer. Clinical cancer research : an official journal of the American Association for Cancer Research *14*, 412-418.
155. Iacobuzio-Donahue, C.A., Fu, B., Yachida, S., Luo, M., Abe, H., Henderson, C.M., Vilardell, F., Wang, Z., Keller, J.W., Banerjee, P., Herman, J.M., Cameron, J.L., Yeo, C.J., Halushka, M.K., Eshleman, J.R., Raben, M., Klein, A.P., Hruban, R.H., Hidalgo, M., and Laheru, D. (2009). DPC4 gene status of the primary carcinoma correlates with patterns of failure in patients with pancreatic cancer. Journal of clinical oncology : official journal of the American Society of Clinical Oncology *27*, 1806-1813.
156. Morohoshi, T., Held, G., and Kloppel, G. (1983). Exocrine pancreatic tumours and their histological classification. A study based on 167 autopsy and 97 surgical cases. Histopathology *7*, 645-661.
157. Roe, J.S., Hwang, C.I., Somerville, T.D.D., Milazzo, J.P., Lee, E.J., Da Silva, B., Maiorino, L., Tiriac, H., Young, C.M., Miyabayashi, K., Filippini, D., Creighton, B., Burkhart, R.A., Buscaglia, J.M., Kim, E.J., Grem, J.L., Lazenby, A.J., Grunkemeyer, J.A., Hollingsworth, M.A., Grandgenett, P.M., Egeblad, M., Park, Y., Tuveson, D.A., and Vakoc, C.R. (2017). Enhancer Reprogramming Promotes Pancreatic Cancer Metastasis. Cell *170*, 875-888 e820.
158. Nan, X., Tamguney, T.M., Collisson, E.A., Lin, L.J., Pitt, C., Galeas, J., Lewis, S., Gray, J.W., McCormick, F., and Chu, S. (2015). Ras-GTP dimers activate the Mitogen-Activated Protein Kinase (MAPK) pathway. Proceedings of the National Academy of Sciences of the United States of America *112*, 7996-8001.
159. Ambrogio, C., Kohler, J., Zhou, Z.W., Wang, H., Paranal, R., Li, J., Capelletti, M., Caffarra, C., Li, S., Lv, Q., Gondi, S., Hunter, J.C., Lu, J., Chiarle, R., Santamaria, D., Westover, K.D., and Janne, P.A. (2018). KRAS Dimerization Impacts MEK Inhibitor Sensitivity and Oncogenic Activity of Mutant KRAS. Cell *172*, 857-868 e815.
160. Vanharanta, S., and Massague, J. (2013a). Origins of metastatic traits. Cancer cell *24*, 410-421.
161. Chen, Q., Zhang, X.H., and Massague, J. (2011). Macrophage binding to receptor VCAM-1 transmits survival signals in breast cancer cells that invade the lungs. Cancer cell *20*, 538-549.
162. Vanharanta, S., Shu, W., Brenet, F., Hakimi, A.A., Heguy, A., Viale, A., Reuter, V.E., Hsieh, J.J., Scandura, J.M., and Massague, J. (2013b). Epigenetic expansion of VHL-HIF signal output drives multiorgan metastasis in renal cancer. Nature medicine *19*, 50-56.
163. Zheng, X., Carstens, J.L., Kim, J., Scheible, M., Kaye, J., Sugimoto, H., Wu, C.C., LeBleu, V.S., and Kalluri, R. (2015). Epithelial-to-mesenchymal transition is dispensable for metastasis but induces chemoresistance in pancreatic cancer. Nature *527*, 525-530.
164. Krebs, A.M., Mitschke, J., Lasierra Losada, M., Schmalhofer, O., Boerries, M., Busch, H., Boettcher, M., Mougiakakos, D., Reichardt, W., Bronsert, P., Brunton, V.G., Pilarsky, C., Winkler, T.H., Brabletz, S., Stemmler, M.P., and Brabletz, T. (2017). The EMT-activator Zeb1 is a key factor for cell plasticity and promotes metastasis in pancreatic cancer. Nature cell biology *19*, 518-529.
165. Puisieux, A., Brabletz, T., and Caramel, J. (2014). Oncogenic roles of EMT-inducing transcription factors. Nature cell biology *16*, 488-494.
166. Sharpless, N.E., and Sherr, C.J. (2015). Forging a signature of in vivo senescence. Nature reviews Cancer *15*, 397-408.
167. Chen, Z., Trotman, L.C., Shaffer, D., Lin, H.K., Dotan, Z.A., Niki, M., Koutcher, J.A., Scher, H.I., Ludwig, T., Gerald, W., Cordon-Cardo, C., and Pandolfi, P.P. (2005). Crucial role of p53-dependent cellular senescence in suppression of Pten-deficient tumorigenesis. Nature *436*, 725-730.
168. Cisowski, J., Sayin, V.I., Liu, M., Karlsson, C., and Bergo, M.O. (2016). Oncogene-induced senescence underlies the mutual exclusive nature of oncogenic KRAS and BRAF. Oncogene *35*, 1328-1333.
169. Ambrogio, C., Barbacid, M., and Santamaria, D. (2017). In vivo oncogenic conflict triggered by co-existing KRAS and EGFR activating mutations in lung adenocarcinoma. Oncogene *36*, 2309-2318.
170. Guerra, C., Schuhmacher, A.J., Canamero, M., Grippo, P.J., Verdaguer, L., Perez-Gallego, L., Dubus, P., Sandgren, E.P., and Barbacid, M. (2007). Chronic pancreatitis is essential for induction of pancreatic ductal adenocarcinoma by K-Ras oncogenes in adult mice. Cancer cell *11*, 291-302.

6 Publications

Mueller, S.*, Engleitner, T.*, Maresch, R.*, Zukowska, M., Lange, S., Kaltenbacher, T., Konukiewitz, B., Ollinger, R., Zwiebel, M., Strong, A., Yen, H.Y., Banerjee, R., Louzada, S., Fu, B., Seidler, B., Gotzfried, J., Schuck, K., Hassan, Z., Arbeiter, A., Schonhuber, N., Klein, S., Veltkamp, C., Friedrich, M., Rad, L., Barenboim, M., Ziegenhain, C., Hess, J., Dovey, O.M., Eser, S., Parekh, S., Constantino-Casas, F., de la Rosa, J., Sierra, M.I., Fraga, M., Mayerle, J., Kloppel, G., Cadinanos, J., Liu, P., Vassiliou, G., Weichert, W., Steiger, K., Enard, W., Schmid, R.M., Yang, F., Unger, K., Schneider, G., Varela, I., Bradley, A., Saur, D., and Rad, R. (2018). Evolutionary routes and KRAS dosage define pancreatic cancer phenotypes. Nature 554, 62-68.

Maresch, R.*, **Mueller, S.***, Veltkamp, C., Ollinger, R., Friedrich, M., Heid, I., Steiger, K., Weber, J., Engleitner, T., Barenboim, M., Klein, S., Louzada, S., Banerjee, R., Strong, A., Stauber, T., Gross, N., Geumann, U., Lange, S., Ringelhan, M., Varela, I., Unger, K., Yang, F., Schmid, R.M., Vassiliou, G.S., Braren, R., Schneider, G., Heikenwalder, M., Bradley, A., Saur, D., and Rad, R. (2016). Multiplexed pancreatic genome engineering and cancer induction by transfection-based CRISPR/Cas9 delivery in mice. Nature communications 7, 10770.

Weber, J.*, Ollinger*, R., Friedrich, M., Ehmer, U., Barenboim, M., Steiger, K., Heid, I., **Mueller, S.**, Maresch, R., Engleitner, T., Gross, N., Geumann, U., Fu, B., Segler, A., Yuan, D., Lange, S., Strong, A., de la Rosa, J., Esposito, I., Liu, P., Cadinanos, J., Vassiliou, G.S., Schmid, R.M., Schneider, G., Unger, K., Yang, F., Braren, R., Heikenwalder, M., Varela, I., Saur, D., Bradley, A., and Rad, R. (2015). CRISPR/Cas9 somatic multiplex-mutagenesis for high-throughput functional cancer genomics in mice. Proceedings of the National Academy of Sciences of the United States of America 112, 13982-13987.

Eissmann K., **Mueller S.**, Sticht H., Jung S., Zou P., Jiang S., Groß A., Eichler J., Fleckenstein B., and Reil H. (2013) HIV-1 Fusion is Blocked by GB Virus C E2D Peptides Binding the gp41 Disulfide Loop. PLoS ONE 8(1): e54452.doi:10.1371/journal.pone.0054452

*shared first authorship

7 Acknowledgments

First, I am very grateful to my thesis advisor and referee Prof. Roland Rad for giving me the opportunity to work on this fascinating research project. I also would like to thank him for his valuable scientific guidance and the many insightful discussions, as well as for his trust in me when providing all the resources that were essential for the success of this project. Prof. Roland Rad created a very stimulating research atmosphere that allowed for thinking outside the box.

I am also very thankful to Prof. Dieter Saur and PD Günter Schneider for the generous sharing of the hundreds of pancreatic cancer cell cultures that they have been collecting over the many years. This research project could not have been conducted without their pancreatic cancer mouse model resource and their extensive support.

In addition, I would like to thank Prof. Allan Bradley for providing the resources and access to the genetics core facilities at the Wellcome Trust Sanger Institute.

I would like to thank Prof. Roland Schmid for giving me the opportunity to work in the excellent research environment at the University Hospital, Department of Medicine II of TU Munich. In addition, I would like to thank Prof. Hrabě de Angelis and Prof. Zuzana Storchová for their comments and helpful advices as members of my thesis advisory committee.

Furthermore, I would like to thank Dr. Rupert Öllinger for his scientific guidance and his helpful advices for all kinds of Next-Generation-Sequencing applications, Dr. Björn Konukiewitz for his help with microdissecting human PanINs, Dr. Sebastian Lange and Maximilian Zwiebel for their bioinformatic analyses, Dr. Katja Steiger for her longstanding expertise in mouse tumour pathology, Dr. Fengtang Yang and his team for the possibility to do M-FISH stainings, Dr. Julia Heß and Dr. Kristian Unger for performing aCGH, Dr. Christoph Ziegenhain for his support with RNA-Sequencing and Dr. Ignacio Varela for his help with establishing our Whole-Exome-Sequencing analysis pipeline. I also thank Thorsten Kaltenbacher, Alexander Strong, Julia Eichinger, Magdalena Zukowska, Dr. Barbara Seidler, Juliana Götzfried, Dr. Kathleen Schuck, Dr. Zonera Hassan, Dr. Andreas Arbeiter, Dr. Nina Schönhuber, Dr. Sabine Klein and Dr. Christian Veltkamp for contributing to my work, for providing me with cell lines as well as for the many discussions.

In particular, I would to thank Thomas Engleitner and Roman Maresch for their fantastic work, the great scientific debates and their endless support, especially when it was needed the most. Thank you so much for your essential input to this work. It's a pleasure to work with you.

I also would like to thank my friends and colleagues of the research teams from Prof. Roland Rad, Prof. Dieter Saur, PD Günter Schneider and Dr. Maximilian Reichert for sharing their experiences and the great moments that we experienced together.

Finally, I owe eternal gratitude to my parents and grandparents. They guided me through the entirety of my life and their endless support gave me the freedom to follow my interests at any time. Without you I would never have achieved this. I am also deeply thankful to Julia for always being there for me and for showing me what's important in life.