Tracing tumor evolution and subtype specification in pancreatic cancer

Nirakar Rajbhandari

University of California San Diego

Michael Hamilton

Departments of Pharmacology and Medicine, University of California San Diego School of Medicine

Cynthia Quintero

Departments of Pharmacology and Medicine, University of California San Diego School of Medicine

L. Paige Ferguson

Departments of Pharmacology and Medicine, University of California San Diego School of Medicine

Ray Fox

University of California San Diego

Christian M. Schurch

Baxter Laboratory for Stem Cell Biology, Department of Microbiology and Immunology, Stanford University School of Medicine

Jun Wang

Sanford Burnham Prebys Medical Discovery Institute

Mari Nakamura

University of California San Diego

Matthew McDermott

Departments of Pharmacology and Medicine, University of California San Diego School of Medicine

Hannah Pettit

Departments of Pharmacology and Medicine, University of California San Diego School of Medicine

Marcie Kritzik

Departments of Pharmacology and Medicine, University of California San Diego School of Medicine

Haiyong Han

Translational Genomics Research Institute <https://orcid.org/0000-0001-9031-9880>

Derek Cridebring

Molecular Medicine Division, The Translational Genomics Research Institute, Phoenix, Arizona

Susan Tsai

The Medical College of Wisconsin Milwaukee

Andrew M. Lowy

University of California, San Diego Moores Cancer Center

Robert J. Wechsler-Reya

Sanford Burnham Prebys Medical Discovery Institute <https://orcid.org/0000-0002-7463-8352>

Daniel D. Von Hoff

Arizona Cancer Center <https://orcid.org/0000-0003-4723-7681>

Aaron M. Newman

Institute for Stem Cell Biology and Regenerative Medicine, Stanford University School of Medicine, Stanford, CA

Tannishtha Reya (✉ treya@ucsd.edu)

University of California, San Diego <https://orcid.org/0000-0002-5956-8536>

Biological Sciences - Article

Keywords: pancreatic cancer, tumor evolution, Cre, adenocarcinoma, cancer, progenitor cells, stem cells

Posted Date: August 5th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-101578/v1>

License: This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Tracing tumor evolution and subtype specification in pancreatic cancer

Nirakar Rajbhandari^{1,2†}, Michael Hamilton^{1,2†}, Cynthia Quintero^{1,2}, L. Paige Ferguson^{1,2}, Ray Fox^{1,2}, Christian M. Schürch³, Jun Wang⁴, Mari Nakamura^{1,2}, Matthew McDermott^{1,2}, Hannah Pettit^{1,2}, Marcie Kritzik^{1,2}, Haiyong Han⁵, Derek Cridebring⁵, Susan Tsai⁶, Andrew M. Lowy⁷, Robert J. Wechsler-Reya^{4,8}, Daniel D. Von Hoff⁵, Aaron M. Newman⁹ and Tannishtha Reya^{1, 2*}

¹Departments of Pharmacology and Medicine, University of California San Diego School of Medicine La Jolla, CA,

²Moore's Cancer Center, University of California San Diego School of Medicine, La Jolla, CA

³Baxter Laboratory for Stem Cell Biology, Department of Microbiology and Immunology, Stanford University School of Medicine, Stanford, CA

⁴Tumor Initiation and Maintenance Program, NCI-Designated Cancer Center, SBP Medical Discovery Institute, La Jolla, CA

⁵Molecular Medicine Division, The Translational Genomics Research Institute, Phoenix, Arizona

⁶Department of Surgery, The Medical College of Wisconsin, Milwaukee, WI

⁷Department of Surgery, Division of Surgical Oncology, University of California San Diego School of Medicine, La Jolla, CA

⁸Rady Children's Institute for Genomic Medicine, San Diego CA

⁹Institute for Stem Cell Biology and Regenerative Medicine, Stanford University School of Medicine, Stanford, CA

† These two authors contributed equally

*To whom correspondence should be addressed, treya@ucsd.edu

Abstract

Identifying the cells from which cancers arise, and the paths they take towards malignancy, is critical for understanding the molecular basis of tumor initiation and progression. To understand if stem and progenitor cells can serve as cells of origin for cancer, we created a model in which the CreER^{T2} recombinase was knocked into the endogenous locus of the stem cell determinant Msi2. When crossed to a conditional *CAG-LSL-Myc^{T58A}* model, *Msi2-Cre^{ERT2}* mice developed a diversity of tumors across tissues, including multiple pancreatic cancer subtypes: pancreatic ductal adenocarcinoma (PDAC), adenosquamous carcinoma of the pancreas (ASCP), acinar cell carcinoma (ACC), and rare anaplastic tumors. Using single cell analysis, we traced the temporal changes that occurred as normal Msi2⁺ cells evolved through a pre-cancerous stage to PDAC, ASCP or ACC. These results revealed that Msi2⁺ cells were present predominantly in pancreatic ducts and exhibited heterogeneous differentiation states in the normal pancreas. At initiation, Myc triggered oncogenic transformation in the most undifferentiated subset leading to the rise of pre-cancer cells with multi-lineage properties. Subsequently, these pre-cancer cells were driven along different fates by epigenetically activated transcriptional programs with genomic changes amplifying the progression into distinct pancreatic cancer subtypes. Finally, integrating transcriptomic and functional genomic approaches in this new model allowed us to define *Ifne*, *Ifitm3*, *Atf3*, *Hmmr*, and *Raet1e* as novel functional dependencies of ASCP, giving us unique insight into the most lethal of pancreatic malignancies. These data show that multiple pancreatic cancer subtypes can arise from a common pool of pre-malignant cells and provide a powerful molecular framework to understand the programs that shape divergent fates in pancreas cancer and develop approaches for early detection and interception.

Pancreatic cancer has the highest mortality and lowest 5-year survival rate among all major cancers ¹. Pancreatic ductal adenocarcinoma (PDAC) is an aggressive and common form of exocrine cancer ^{2,3}, and has received the most attention in terms of understanding its genesis and its vulnerabilities. However, within the exocrine pancreas, multiple other forms of malignancies can occur. This includes adenosquamous carcinoma of pancreas (ASCP), an aggressive sub-type of PDAC that has the worst clinical prognosis ⁴⁻⁶; acinar cell carcinoma (ACC), which accounts for 15% of all childhood pancreatic cancers and is only moderately responsive to chemotherapies ⁷; and the highly aggressive anaplastic tumors, characterized by their undifferentiated state ⁸. While originally considered rare, recent studies have indicated that a substantially higher fraction of conventional PDAC (18%) harbors admixed adenosquamous features than was previously recognized ⁹. Further, what is considered the 'basal subtype' (40% of PDAC) in fact aligns well with the molecular signature of the adenosquamous subtype. These raise the possibility that a substantive fraction of pancreatic cancers may not have properly matched mouse models and highlight the need for new models that better reflect disease variance and subtypes ^{9,10}. While not as rapidly progressing as PDAC or ASCP, acinar cell carcinomas are nonetheless highly aggressive with an overall 5-year survival rate of less than 10%; 50% of patients present with metastatic disease and have a median survival of 14 months ¹¹. Thus, better models can enable understanding the biological basis of ACC, and the development of new approaches to ensure more durable remissions for this disease as well.

Although there are multiple subtypes of pancreatic cancers, they are currently treated as one disease. Defining where these tumors arise and how they differ during initiation and progression can enable an understanding of their molecular underpinnings and development of subtype specific therapies. While multiple studies have focused on acinar versus ductal cells as an origin for PDAC ¹²⁻¹⁴, the origin for other subtypes and the trajectories cells take has remained relatively unexplored. Here we report a new mouse model in which targeting oncogenic Myc to adult pancreatic precursor cells triggered the development of PDAC, ASCP, ACC and anaplastic tumors. Our current model is particularly exciting and highly relevant to human disease, as deregulated expression or amplification of c-MYC is associated with higher incidence of ASCP (47%) and ACCs (54% in pure and up to 100% in mixed ACCs) in human pancreatic patients, irrespective of *KRAS* mutation status ^{5,15,16}. This work provides an exciting new model for multiple subtypes of pancreatic cancer, and a unique opportunity to trace, at single cell resolution, the origins of the pre-malignant state and its evolution to distinct lethal malignancies.

Results

Msi2 expression marks the cell of origin for cancers across diverse tissues

To understand the tumor initiating potential of normal Msi2⁺ stem/progenitor cells, we developed a new *Msi2-Cre^{ERT2}* knock-in mouse model in which a *CreERT2*-cassette was integrated into the “ATG” codon of the Msi2 locus (**Fig. 1A and 1B**). To assess the fidelity of this new model, we first analyzed endogenous Msi2 expression and found it predominantly expressed in stem/progenitor enriched populations: in Sox2⁺ cells of the subventricular zone (**Fig. 1C**), Clara cells at the bronchoalveolar junction in the lung (**Fig. 1D**), calbindin⁺ cells in the kidneys (**Fig. 1E**) and ductal epithelial cells in the pancreas (**Fig. 1F**). We crossed *Msi2-Cre^{ERT2}* mice to *R26R-LSL-tdTomato* reporters ¹⁷ to compare Cre activity to Msi2 expression. TdTomato (tomato) expression following tamoxifen-driven Cre activation was strongly aligned with Msi2 expression (**Fig. 1G-J**), confirming the fidelity and specificity of this new knock-in model.

To determine if normal Msi2⁺ cells can serve as an origin for cancer, we crossed *Msi2-Cre^{ERT2}* mice with a *CAG-LSL-Myc^{T58A}* model ¹⁸ in which the CAG promoter drives expression of a conditional Myc^{T58A} allele (**Fig. 1K**). *Msi2-Cre^{ERT2}; CAG-LSL-Myc^{T58A}* mice (referred to as *Msi2-Myc*) treated with tamoxifen developed a wide range of cancers, including choroid plexus carcinoma (**Fig. 1L, Fig. S1A**), small-cell lung cancer (**Fig. 1M, Fig. S1B**), gastric carcinoma (**Fig. 1N, Fig. S1C**), and pancreatic cancer (**Fig. 1O**). In other tissues, such as the kidney, c-Myc drove dysplasia (**Fig. S1D-E**) and cooperated with VHL deletion to drive a malignant state. Tumor penetrance varied between tissues; of the mice analyzed, 80% had lung tumors, 68% gastric tumors, 52% brain tumors and 45% pancreatic tumors (**Fig. 1P**). These data identify Msi2⁺ cells as a cell of origin for a range of tumors across tissues.

A new model for multiple pancreatic cancer subtypes

The pancreas tumors that developed in *Msi2-Myc* mice were of particular interest, as they represented a wide diversity of pancreatic cancer subtypes, including PDAC, ASCP, ACC and anaplastic tumors. We classified tumors using histology and canonical markers: cytokeratin19 and absence of amylase for PDAC (**Fig. 2A**), p63 for squamous differentiation (SD) (**Fig. 2B**), and cytokeratin 5/14 for tumors with squamous features (SF) (**Fig. 2C and 2C'**), amylase for ACC (**Fig. 2D**), and the lack of differentiation markers for anaplastic tumors (**Fig. 2E**). While some *Msi2-Myc* mice harbored only one subtype, others presented with two or more within the same pancreas (**Fig. 2F**). Overall, ductal tumors with squamous features (CK5, CK14) or frank ASCP (p63⁺) were found in 44% (11/25), ACC in 88% (22/25), and anaplastic tumors in 8% (2/25) of the mice (**Table 1**); all expressed c-Myc at high levels consistent with their Myc-driven origins (**Fig. 2A-B and 2D-E, bottom**). Similarly, Msi-Cre^{ERT2}-driven *Kras* activation in the presence of

cerulein led to rare cases of PanINs but did not yield the penetrance or diversity of subtypes observed in the *Msi2-Myc* mice (not shown). This suggests that targeting Myc specifically to precursor cells was critical to directing the full diversity of tumor fates.

Additionally, RNA-Sequencing (RNA-Seq) on tumor cells from this model revealed that ACC and PDAC/ASCP subtypes harbored transcriptomic landscapes distinct from each other and from the cell of origin (**Fig. 2G**); ASCP tumors were more closely related to *Kras/p53* driven (KPC) tumors¹⁹ consistent with their ductal nature. In addition, ASCP tumors showed higher expression of classical squamous markers like *Krt5*, *Krt6a* and *Trp63*, along with an enrichment in (Δ N) *Trp63* transcripts (**Fig. 2H and 2I**). These data collectively indicate that tumor subtypes in this new model reflect the diversity of human pancreatic cancers with high fidelity and could be valuable for identifying new subtype specific dependencies.

Single cell map of the cell of origin and pre-cancer states in pancreas cancer subtypes

The fact that multiple subtypes of pancreatic cancer arose from a common pool of *Msi2*⁺ pancreatic cells provided a rare opportunity to trace the trajectory of tumor cells as they diverged into distinct fates. We used single cell sequencing to track the molecular changes that occurred as normal *Msi2*⁺ cells ($t=0$) transitioned to a pre-cancerous state (5 weeks post-tamoxifen; i.e $t=5$) and continued to fully developed PDAC, ASCP or ACC tumors at endpoint (10 to 12 weeks post-tamoxifen; i.e, $t=12$) (**Fig. 3A**). To analyze the heterogeneity in normal *Msi2*⁺ cells at $t=0$, we sorted GFP⁺ cells from *Msi2*-GFP reporter mice (*Msi2*^{tgGFP/+}). *Msi2* was expressed predominantly in duct and islet cells and in few acinar cells (<0.5%) (**Fig. S2A-B**).

Msi2⁺ duct cells at $t=0$ fell into three distinct clusters (D1-D3, **Fig. 3B**). To understand how these demarcated clusters may be related and whether they represent a hierarchy, we used cytoTRACE, a recently developed computational pipeline that can identify the most immature cells within single cell populations by tracing transcriptional diversity²⁰. This approach identified D1 and D2 as the most undifferentiated (dark red) and D3 as a more differentiated cell population (green) (**Fig. 3C**). Consistent with this observation, D1 and D2 clusters highly expressed the stem cell and re-programming genes *Prom1*, *Sox4*, and *Gata6* (**Fig. 3D**) as well as *Oct4*, *Tead2* and *Notch* (not shown). They differed in that D1 was enriched for cell cycle genes (**Fig. 3E**), whereas D3 was more differentiated, with lower expression of most stem cell genes and elevated expression of *Foxj1*, a transcription factor required for ciliated cell differentiation (**Fig. 3E**).

Next, to define the molecular changes directed by Myc in precursor cells, we mapped the transcriptomic diversity of pre-cancer cells. *Msi2-Myc* mice were crossed to *R26R-LSL-YFP* mice²¹, and YFP⁺ (i.e Myc-expressing) epithelial cells were sorted and analyzed at 5 wks ($t=5$).

Following Myc induction, pre-cancer lesions were dominated by multi-lineage uncommitted cells (**Fig. 3F**), with an enrichment of genes with direct Myc-binding sites (**not shown**). These uncommitted cells were also marked by cell cycle and chromatin modifying genes including *Hmgb2* and *Top2a* and a multitude of oncogenic programs including *Aldh1a1*, *Ran*, *Ube2c* and *Klf2* (**Fig. 3G**). Consistent with their undifferentiated proliferative nature, these cells harbored embryonic stem cell and cell cycle signatures (**Fig. 3H**). A metagene analysis using the top 10 genes enriched in normal D1, D2 or D3 cell clusters from $t=0$ showed that gene signatures of D1 and D2 but not D3 were the most enriched within these pre-cancer cells. This identifies the D1/D2 fraction as the cells from which midpoint pre-cancer cells (MDP) arose (**Fig. 3I**). Collectively, these data suggest that the more undifferentiated D1/D2 cells may be the most sensitive fraction to Myc-induced transformation, and that Myc enforces the undifferentiated state of *Msi2*⁺ ductal cells in establishing a pre-cancerous state with a multi-lineage fate.

Next, we experimentally tested our hypothesis that acinar and ductal pancreatic cancer subtypes arose from common ductal cells-of-origin by isolating normal duct cells from *CAG-LSL-Myc*^{T58A} mouse and tracing their fate in the presence or absence of Myc *in vitro*. While Myc activation (+Cre) triggered the growth of organoids with both amylase⁺ acinar fated cells and DBA⁺ ductal fated cells (**Fig. 3J**), only duct fated organoids arose in the absence of Myc (-Cre). Consistent with this, *in vivo* lineage tracing experiment using *Msi2-Cre*^{ERT2}; *R26R-LSL-tdTomato*^{+/-} reporter mice revealed that in the absence of Myc, *Msi2*⁺ ductal cells do not give rise to non-ductal lineages at least in postnatal life. Thus, pancreata from one day old mice (d1) pulsed with tamoxifen showed labeling predominantly in duct cells, along with some occasional acinar cells. We observed a statistically significant expansion of labeled cells only in ducts (from $12.6 \pm 0.5\%$ at pulse to $34.1 \pm 6.4\%$ at chase) but not in acinar compartment at the end of the chase period of 10 weeks (**Fig. S3A-B**). We did not observe any significant expansion of either duct or acinar cells when lineage tracing was initiated in adult mice (**Fig. S3C-D**). These data collectively indicate that acinar cells are not likely to arise from *Msi2*⁺ ductal cells in the absence of Myc, but that these cells have latent potential to form both ductal and acinar lineages, a potential that is re-activated by oncogenic Myc.

Single cell map of developmental trajectories of pancreas cancer subtypes

To understand how the pre-cancer cells evolved into distinct tumor subtypes, we analyzed end stage ($t=12$) acinar and adenosquamous tumors from *Msi2-Myc* mice. Single cell RNA-sequencing (scRNA-Seq) analysis of acinar tumors showed elevated levels of conventional markers lipase (*Cel*) and β -catenin (*Ctnnb1*) (**Fig. 4A**). The population identified by CytoTRACE

as ACC precursors were enriched in *Hmgb2* and *Top2a* and harbored the fewest copy number variations (CNVs) (**Fig. 4B**) - genomic changes which accrue with tumor progression. Further, Monocle, an algorithm that can project single cell relationships based on the kinetics of gene expression²² built a single trajectory originating from this precursor population (**Fig. 4C left**, cells in green), leading to established ACC cells (**Fig. 4C right**).

Similarly, we used endpoint PDAC/ASCP tumors to determine whether they evolve from distinct origins and progress along distinct paths relative to ACC. Much like the human disease, *Msi2-Myc* tumors contained ductal and squamous populations in varying proportions. We analyzed two independent squamous tumors, one with mixed ductal/squamous features (*Krt7+*, SF, **Fig. 4D-F**) and the other with *p63+* squamous differentiation (frank ASCP/SD, **Fig. 4G-I**). Importantly, the most immature cluster in the bi-fated ductal/squamous tumors and in the *p63+* ASCP was similar to the most immature cluster in ACC: all expressed *Hmgb2/Top2a* (**Fig. 4E, 4H left and middle**), and containing the fewest CNVs. We observed that CNVs accumulated progressively with increasing differentiation, highlighting a remarkable alignment of how genomic evolution of the tumors paralleled its epigenetically controlled progression towards a differentiated lineage (**Fig. 4E, 4H right**). In the bi-fated tumor both ductal adenocarcinoma and squamous tumor cells arose from the *Top2a/Hmgb2+* population (**Fig. 4E-F**) but diverged at a branch point either continuing to a PDAC fate (*Sox9*, *Agr2* and *Tff2* at endpoint) or turning towards a squamous fate (*Krt5*, *Krt6a*, *Krt15* and *Krt7* at endpoint) (**Fig. 4F**). Differentiated ASCP tumor cells expressed *p63* in addition to the keratins at endpoint (**Fig. 4I, right**). These results based on tracing of the initiation and evolution of pancreatic cancer subtypes indicate that ACC, PDAC and ASCP subtypes all arise from a common pool of less differentiated pre-cancer cells in response to oncogenic *Myc* (**Fig. 4J**).

Genomic and epigenetic changes drive divergent tumor fates

Genomic signatures of the end-stage tumors showed that ACC and PDAC/ASCP isolated from the same mice shared common CNVs, suggesting these cells arose from a common precursor that later diverged (**Fig. 4K, top and bottom**). This finding was further supported by the molecularly overlapping (**Fig. 4L, left**), multilineage (**Fig. 4L, right**) and oncogene enriched common core signature (**Fig. 4M**) between pre-malignant cells from midpoint (*t*=5) and endpoint (*t*=12) lesions. In line with this notion, scRNA-Seq on a human tumor with squamous features (**Fig. S4A**) showed existence of pre-malignant cells with an undifferentiated state (**Fig. S4B, left, red**), low CNV frequency (**Fig. S4B, right**) and the “common pre-cancer” core signature (**Fig. S4C, left**), supporting the conservation of an equivalent ancestral precursor pool in human

disease. We also found that *TOP2A* and *HMGB2*, key markers of pre-malignant cells, were also expressed in a cohort of human pancreatic cancers analyzed (**Fig. S4C**), where their abundance correlated with poor survival (**Fig. S4D-G**; [\[www.proteinatlas.org\]](http://www.proteinatlas.org)²³).

To understand whether genomic changes underlay these divergent transcriptional programs (**Fig. S5A-D**), we mapped CNV evolution. End stage ACC and ASCP tumors displayed distinct patterns of recurrent CNVs: amplifications in chromosomes 6 and 8 were common in ACC (**Fig. 4N**), and amplifications in chromosomes 3 and 7 were common in ASCP (**Fig. 4O**). Network analysis of amplified genes indicated they may drive subtype specific programs: amino acid metabolism, autophagy, and oxidative phosphorylation in ACC (not shown) and Ras signaling, previously linked to a basal phenotype¹⁰, as well as Notch signaling, phospholipid/cholesterol metabolism, and innate immunity in ASCP (not shown). Amplified genes identified in these models were conserved in corresponding human tumor subtypes as well^{24,25} (**Fig. S6A and S6B**).

To define the role of CNVs in fate decisions, we closely analyzed pre-cancer cells ($t=5$) and found that subtype specific CNVs had begun to emerge in pre-cancer cells (**Fig. 4P**). Consistent with this, distinct transcriptional programs- indicating emergence of early acinar (*Zg16*, *Ctrl* and *Prss1*) or early ductal fates (*Reg3a*, *Reg3b*, *Reg3g* and *Dmbt1*) were found in distinct pools of cells (**Fig. 4Q**). However, projection of hallmark CNVs on the single cell map indicated they occurred after fate specific gene expression had already emerged (**Fig. 4R**) suggesting that the transcriptional changes that ultimately dictate fate, were triggered not by genetic changes but more likely by epigenetic mechanisms, with the caveat that CNVs may be present but undetectable in some cells. Consistent with this possibility, single cell ATAC-Seq on pre-cancer cells ($t=5$) and end-stage pancreas ductal tumors ($t=12$) revealed that the loci for both duct and acinar lineage were open in pre-cancer cells (**Fig. S7A**). However, the acinar loci were epigenetically silenced as tumors committed to the ductal lineage (**Fig. S7B**). These findings suggest that multi-lineage pre-cancer cells have both ductal and acinar open epigenetic landscapes, and both epigenetic and transcriptomic changes direct cell fate, while specific genomic changes may work to reinforce or commit pre-cancer cells to distinct lineages.

Identification of novel dependencies in pancreatic adenosquamous cancer

The development of this pancreatic adenosquamous tumor model provided a unique opportunity to not only understand how ASCP tumors develop (**Fig. 4J**) but also to identify their vulnerabilities, an important endeavor because ASCP remains essentially unresponsive to current standard of care. Therefore, to test the novel dependencies of squamous tumor cells, we focused on a selected set of genes that were enriched in ASCP tumors but whose role in pancreatic

cancers are unknown; this included *Mecom* (MDS1 and EVI1 complex locus), *Hmmr* (Hyaluronan mediated motility receptor), *Atf3* (Activating transcription factor 3), *Il18* (Interleukin 18), *Ifne* (Interferon epsilon), *Tspan4* (Tetraspanin 4), *Raet1e* (Retinoic acid early transcript 1e), *Ifitm3* (interferon induced transmembrane protein 3) and *Sprr3* (Small proline-rich protein 3) (**Fig. 5A**). Of these, *Ifne*, *Ifitm3*, *Atf3*, *Hmmr*, *Tspan4*, *Sprr3* and *Raet1e* most significantly decreased the sphere/colony forming potential of mouse ASCP/ductal tumor cells *in vitro* (**Fig. 5B**). In concordance with the findings from the mouse model, inhibition of *TSPAN4*, *RAET1E*, *IFITM3*, *SPRR3* and *IFNE* significantly affected the sphere/colony forming potential of BxPC3, a c-MYC overexpressing, *KRAS* wildtype human squamous pancreatic cancer line, *in vitro* (**Fig. 5C**) highlighting the importance of these genes in maintenance of squamous tumor cells.

Next, we were particularly interested in exploring the impact of *Ifne* inhibition *in vivo*, as it is a novel type I Interferon, studied only in the context of the innate immune response. IFNE expression was highly enriched and specific to squamous tumor cells (**Fig. 5D**) in *Msi2-Myc* mice and in human pancreatic adenosquamous cancers (n = 4/5; i.e, 80%) (**Fig. 5E**) but undetectable in the normal pancreas. Functional inhibition of *Ifne* in pancreatic tumor cells derived from a squamous tumor in the *Msi2-Myc* model significantly blocked tumor initiation and growth *in vivo* (**Fig. 5F**), further indicating that *Ifne* is a critical dependency in ASCP.

Collectively, these data provide an exciting new model for the ductal origin of pancreatic tumor subtypes (**Fig. 5G**), where *Msi2-Cre^{ERT2}* driven overexpression of Myc preferentially targets the most immature subfractions of pancreatic ductal cells leading to the expansion of multilineage cells. These cells subsequently give rise to a common pre-cancer pool which follows distinct trajectories to evolve into differentiated tumor subtypes. Common pre-cancer cells are driven by Myc to have both acinar and ductal loci open, thus driving diverging trajectories at an epigenetic level into either an acinar or ductal fate. Subsequently, unique genomic changes that favor oncogenic growth along distinct lineages promote the growth and establishment of acinar or ductal/ASCP tumors. Finally, creation of this new model allowed us to define *Ifne*, *Ifitm3*, *Atf3*, *Hmmr*, *Sprr3* and *Raet1e* as novel functional dependencies of ASCP, giving us unique insight into the most lethal subtype of pancreatic malignancies. This work thus provides an important understanding of the programs that shape divergent fates in pancreas cancer and provides a foundation to develop approaches to early detection, interception, and treatment of pancreatic cancer subtypes.

Discussion

One of the most important findings of our work is the demonstration that an immature fraction of duct cells can serve as cells-of-origin for multiple subtypes of pancreas cancer. Most of the prior models have focused on driving oncogenic *Kras* and *Trp53* that resulted in classical PDAC^{12,13,26,27-29}. Previous models overexpressing *Myc* in the pancreas have been developed (either together with *Ras*, or embryonically)³⁰⁻³³; however, none have resulted in the broad range of tumor subtypes we report, and therefore, have not been able to trace how these subtypes arise. Our data shows that multiple subtypes likely arise in our model due to *Myc* being targeted to an early progenitor capable of expanding into a multilineage tumor precursor and subsequently differentiating along diverse lineages. Our work brings new insight into the events that drive the pre-malignant state and its transition to diverse lethal malignancies, thus providing an important framework to develop early interception strategies in pancreatic cancer.

The *Msi2-Myc* mice may also serve as a model for pancreatic carcinomas with mixed phenotypes including acinar-ductal adenocarcinoma, PDAC-ASCP, and acinar-endocrine-ductal carcinoma. Such mixed carcinomas are considered aggressive, and are being reported at an increasing prevalence³⁴⁻³⁶, especially with the advent of single cell sequencing strategies³⁷. In addition, observations that some tumors shift fates after therapy may be consistent with our finding that common tumor precursors persist as an ancestral pool in PDAC patient samples and raise the possibility they may adopt a distinct fate in response to epigenetic and genetic events.

The *Msi2-Myc* model is particularly important because currently all pancreatic cancer is treated equivalently, and subtype specific treatments have not been defined. Given recent data showing that a higher fraction of PDACs are admixed with adenosquamous cells^{9,10} than previously realized, effectively modeling this heterogeneity has become more urgent. In this context, the *Msi2-Myc* mice present a unique, timely and important model to better understand this disease as it can open many new avenues of discovery: mapping the transcriptomic and microenvironmental landscape of ASCP, determining its key dependencies, and screening and testing new therapeutic strategies. In this light, combination of single cell transcriptomics with functional genomics allowed us to identify a wide range of potentially novel dependencies in *Msi2-Myc*-induced ASCP cells, including *Ifitm3*, *Tspan4*, *Raet1e*, *Atf3*, *Sprr3* and *Hmnr*. Interestingly, one of the unexpected findings of our work was the identification of innate immunity associated cytokine of type I interferon family- *Ifne*, as a key dependency and potential therapeutic target in pancreatic adenosquamous cancers. Significant growth impact upon short hairpin-mediated genetic ablation of IFNE in mouse and human pancreatic squamous tumor cells suggests that ASCP cells are critically dependent on IFNE signaling for their survival independent of immune cells and other stromal components in the tumor microenvironment. Our findings open a path

forward for future studies to understand the mechanistic basis of its ASCP-specific activity and the potential for therapeutic targeting of IFNE for the treatment of human ASCPs.

The model we have described here provides an interesting and important complement to another model targeting stem/progenitor cells using the Prominin-1 (CD133) promoter^{38,39}; in this model mutations (except Myc) were activated and resulted in tumors in multiple organs except the pancreas. The distinct tumor phenotypes in the two models highlight how the nature of oncogene (Myc) as well as the cell in which it is targeted (Msi2+) are key in directing diverse tumor types. While Msi2+ cells have been shown to be stem and progenitor cells across tissues⁴⁰⁻⁴², whether they harbored the latent capacity to launch tumorigenic programs was unknown. Our data provide the first evidence that normal Msi2+ cells are primed for, and highly sensitive to, Myc-driven transformation across organs. To our knowledge no other reported Myc model has resulted in such a diverse range of tumor types. Thus, in a broader context, this may serve as an important resource for understanding how Myc mutations drive tumor initiation and progression from progenitor cells and allow development of strategies for early detection and interception based on the unique vulnerabilities of diverse cancers.

Acknowledgements

We are grateful to Michael Karin, Gerard Evan and Anjana Rao for scientific advice and comments on the manuscript; Maïke Sander, Kay-Uwe Wagner and Christopher Wright for sharing reagents and mice; Kristen Jepsen and Eugenia Ricciardelli for help with single cell RNA-Sequencing; and Armin Ahmadi and Kendall Chambers for technical support. N.R. received support from the Tobacco Related Disease Research Program T29FT0280; L.P.F. received support from Ruth L. Kirschstein National Research Service Award F31CA247489 and T32 GM007752; M.H. and C.Q. received support from T32 HL086344; M.M received support from T32 GM007752. We are grateful to the Lee T. Hanley Fund for Pancreatic Cancer Research for supporting this work. This work was also supported by R01 CA186043, to A.M.L and T.R, R35 CA197699 to T.R and an SU2C–CRUK–Lustgarten Foundation Pancreatic Cancer Dream Team Research Grant (SU2C-AACR-DT-20-16) to D.V.H., H.H., A.M.L., and T.R.

Materials and Methods:

Mice

Msi2^{CreERT2/+} knockin mice were generated by conventional gene targeting (genOway Inc, France). Generation of *Msi2e*^{GFP/+} reporter mice was previously described⁴². All of the GFP-reporter mice used in experiments were heterozygous for the *Msi2* allele. *Rosa26-LSL-YFP* mice²¹ was previously described. The *Lox-Stop-Lox (LSL)-Myc*^{T58A}-*IRES-Luciferase* mice were generated as described¹⁸. The *Rosa26-LSL-tdTomato* mice (B6.Cg-Gt(*ROSA*)26Sor^{tm14(CAG-tdTomato)Hze/J}; stock number: 007914) were purchased from Jackson Laboratory¹⁷. Immuno-compromised NOD/SCID (NOD.CB17-Prkdcscid/J; stock No: 001303) and NSG (NOD.Cg-PrkdcscidIL2rgtm1Wji/SzJ, Stock No: 005557) mice were purchased from The Jackson Laboratory. All mice were bred and maintained in the specific-pathogen free animal care facilities at the University of California San Diego. Animals had access to food and water ad libitum and were housed in ventilated cages under controlled temperature and humidity with a 12-hour light-dark cycle. All animal experiments were performed according to protocols approved by the University of California San Diego Institutional Animal Care and Use Committee. Therefore, males and females of each strain were equally used for experimental purposes and both sexes are represented in all datasets. All mice enrolled in experimental studies were treatment-naïve and not previously enrolled in any other experimental study. Only mice heterozygotes for *CAG-LSL-Myc*^{T58A} were used for *Msi2-Myc* tumor studies.

In vivo tamoxifen delivery

Fresh tamoxifen (Sigma T-5648) stock was prepared by dissolving tamoxifen in corn oil (Sigma C8267) at 20 mg/ml concentration. For Cre-reporter labeling and lineage tracing experiments, adult mice were injected intraperitoneally with a single dose of 2 mg per 20 g body weight. Myc induction in *Msi2-Cre*^{ERT2} mice was initiated at 6-8 wks of age with 3 doses of tamoxifen (2 mg per 20 g of body weight) delivered over 5 days. The first day of tamoxifen injection was considered as day-0 (d0).

Tissue dissociation, cell isolation, and FACS analysis

All mouse pancreatic tumors were aseptically taken out from tumor-bearing mice, washed with cold 1x PBS (GIBCO, Life Technologies) and minced into 1-2 mm pieces using scalpel blades immediately following resection. Tumor pieces were collected into a 50 mL Falcon tube containing 10 mL Gey's balanced salt solution (Sigma, #G9779), 10 mg Collagenase P (Roche,

#11249002001), and 0.2 mg DNase I (Roche, #4716728001). Samples were incubated for 20 to 30 minutes at 37°C, then pipetted and returned to 37°C for 15 minutes. Samples were re-pipetted then passed through a 70um nylon mesh (Corning). Red blood cells were lysed using RBC lysis buffer (eBioscience, #00-4333-57) and remaining tumor cells were washed, resuspended in HBSS (GIBCO, Life Technologies) containing 2.5% FBS and 2 mM EDTA for staining, FACS analysis, and cell sorting. Analysis and cell sorting were carried out on a FACS Aria III machine (Becton Dickinson), and data were analyzed with FlowJo software (Tree Star).

For dissociation of normal and pre-cancerous pancreas, pancreata were dissociated using modification of a protocol previously developed by Assi and colleagues ⁴³. In brief, pancreata were minced with scissors and collected into 50 ml Falcon tubes containing 10 ml of EGTA buffer with 0.6 mg/ml Collagenase P. Samples were incubated for 15 minutes in a rotor at 37 °C. Samples were washed once with cold PBS without mixing the samples and centrifuged at 500g for 3 minutes. Pellets were then further dissociated using Calcium Buffer with 0.6 mg/ml of *Collagenase P*, and placed in a rotor at 37 C for 20 minutes. At the end of incubation samples were pipetted, and then neutralized with 10 ml of cold PBS and passed through 70 um nylon mesh. Cells were centrifuged at 500g for 3 minutes, and pellets were processed for FACS staining without RBC lysis step. For analysis of cell surface markers by flow cytometry, 0.5×10^6 to 1×10^6 cells/ml were resuspended in EGTA buffer containing 0.5% BSA (for normal or pre-cancerous pancreas) or HBSS buffer containing 2.5% FBS (for tumors) with DNase I (20ug/ml), then stained with respective antibodies in the presence of Fc block. The following fluorochrome conjugated rat antibodies were used: anti-mouse EpCAM (eBioscience), anti-mouse CD45-PE/Cy7 (eBioscience), anti-mouse CD31-PE (BD Bioscience). Propidium-iodide (Life Technologies) or Sytox Blue (Life Technologies) was used to stain for dead cells.

Immunofluorescence staining

All fresh tissues, including normal and cancerous pancreas were fixed in 4% PFA for at least 24 hours (Fisher Scientific) and paraffin embedded in the Immunohistochemistry Core at Moores Cancer Center, University of California San Diego using standard protocols. 5 um sections were obtained and deparaffinized in xylene. For paraffin embedded mouse and human pancreas tissues, antigen retrieval was performed at 95-100 °C for 40 minutes using 1x Citrate Buffer, pH 6.0 (eBioscience). For frozen tumor sections, OCT embedded flash frozen tissues were freshly sectioned and were fixed in 100% methanol for 30 minutes. Sections from both paraffin embedded and frozen tissues were blocked in PBS containing 0.25%% Tween-20 (Sigma- Aldrich) and 5% Goat Serum (Fisher Scientific). All incubations with primary antibodies were carried out

overnight at 4°C. Incubation with Alexafluor-conjugated secondary antibodies (Molecular Probes) was performed for 1 hour at room temperature. DAPI (Molecular Probes) was used to detect DNA and images were obtained with a Confocal Leica TCS SP5 II (Leica Microsystems).

Human and Mouse Pancreatic Cancer Cell Lines

Mouse pancreatic cancer cell lines and organoids were generated from established end-stage tumors (10-12 week after tamoxifen induction) in *Msi2-Myc* mice. Tumors were aseptically isolated and dissociated into single cell suspension as mentioned above. Dissociated single cells were then either plated in three-dimension (3-D) in organoid culture conditions or plated in 2D in 1x DMEM containing 10% FBS, 1x Pen/Strep, 1x Glutamine and 1x non-essential amino acids (GIBCO, Life Technologies). All cells used the experiments were maintained at low passage condition (less than 10).

Human BxPC3 cells were originally derived from a PDAC metastasis and have been previously validated (ATCC # CRL-1687). BxPC-3 cell lines were cultured in 2D conditions in 1x DMEM (GIBCO, Life Technologies) containing 10% FBS, 1x Pen/Strep, 1x Glutamine and 1x nonessential amino acids (GIBCO, Life Technologies).

Cloning of shRNA lentiviral constructs and production

Short hairpin RNA (shRNA) constructs were designed and cloned into pLV-hU6-EF1a-green vector by Biosettia. ***All the target sequences are listed separately.*** Virus was produced in 293T cells transfected with shRNA constructs along with psPAX2 and pMD2.G constructs mixed at a 4:2:1 proportion in Opti-MEM media with a dropwise addition of X-tremeGENE DNA transfection reagent (Roche). Viral supernatants were collected for two days then concentrated by ultracentrifugation at 20,000 rpm for 2 hours at 4 °C.

In vitro matrigel colony formation assay for mouse and human pancreatic cancer cells

For *in vitro* colony forming assay, 500-1000 of mouse, or 200-500 human pancreatic cancer (BxPC3) cells were resuspended in 100 uL DMEM media with 10% FBS, then mixed with growth factor reduced matrigel (BD Biosciences, 354230) at a 1:1 ratio and plated in triplicates in 96-well ultra-low adhesion culture plates (Costar). After incubation at 37_°C for 5 min, 100 uL of media was placed over the matrigel layer. Colonies were counted 7-10 days after plating.

Organoid culture and in vitro transformation assay

For *in vitro* transformation of ductal organoids, normal ducts were aseptically hand-picked from 6-8 weeks old *CAG-LSL-Myc^{T58A}* mice, and cultured in PancreaCult Organoid Media (StemCell Technologies, Inc.). For maintenance, mature organoids were recovered using Cell Recovery Solution (Corning 354253), dissociated into single cells using Accumax Cell Dissociation Solution (Innovative Cell Technologies), and plated at a density of 1500 to 2000 cells per 20 μ L matrigel (BD Biosciences, 354230) domes in pre-warmed 48-well plates at 37C. All the organoids were passaged every 4 to 5 days and maintained for less than 10 passages. Mature organoids were transduced with retrovirus encoding empty control or “Cre-YFP” at second passage, and the pure infected cells were sorted based on YFP expression.

Real-Time Quantitative PCR analysis

RNA was isolated using RNeasy Plus Mini kits (QIAGEN) and converted to cDNA using Superscript Mix (Quantabio). Quantitative real-time PCR was performed using an iCycler (BioRad) by mixing cDNAs, iQ SYBR Green Supermix (BioRad) and gene specific primers. ***All primer sequences are listed separately.*** All real time data was normalized to B2M or Gapdh.

RNA extraction and library preparation for RNA Sequencing

Tumors from 10-12 weeks old *Msi2-Myc* mice were harvested and dissociated into a single cell suspension as described above. Tumor cells were stained with anti-CD45 and anti-EpCAM antibodies (eBioscience) in the presence of Fc block. 50,000-100,000 CD45-/EpCAM+ cells were sorted, washed with 1x PBS and the pellet were lysed in RLT-Plus Buffer (Qiagen # 74034) in the presence of β -mercaptoethanol. Total RNA was isolated using RNeasy Micro Plus kit (Qiagen # 74034). Total RNA was assessed for quality using an Agilent Tapestation, and all samples had RIN >7.0. RNA libraries were generated from >50 to 100 ng of RNA using Illumina's TruSeq Stranded mRNA Sample Prep Kit following manufacturer's instructions. RNA libraries were multiplexed and sequenced with 75 basepair (bp) single end reads (SR75) to a depth of approximately 30 million reads per sample on an Illumina HiSeq4000 using V4 sequencing chemistry.

Single cell capture and sequencing

Pancreas dissociations were performed using the protocol as described above. For the determination of heterogeneity in *Msi2* expressing cells in normal adult pancreas at baseline (time, t=0), pancreata from *Msi2^{eGFP/+}* reporter mice were used. All viable GFP+ cells were collected after negative selection for CD45/CD31+ cells. For the analysis of pancreata at mid-point of tumor

progression after the onset of Myc expression following tamoxifen delivery, CD45-/EpCAM+/Cre-reporter+ (YFP+) cells were collected at 5 weeks after tamoxifen treatment. Cre-reporter+, epithelial cells were used to identify Myc expressing cells in *Msi2-Cre^{ERT2}; LSL-Myc^{T58A}; R26-LSL-YFP* mice. In most cases, viable epithelial (EpCAM+, CD45-) cells were sorted from the tumor, while in cases with high tumor cell content, all viable cells were sorted in preparation for single cell sequencing.

Sorted cells were washed once and resuspended in 1x PBS with 0.05% BSA. Sorted cells were manually counted and viability was assessed using Trypan Blue staining. Up to 16,000 cells were loaded into one well of the 10X chip to produce Gel Bead-in-Emulsions (GEMs) and capture individual cells. We utilized the Chromium Single Cell 3' Reagent Version 1 Kit (10X Genomics). GEMs underwent reverse transcription to barcode RNA before cleanup and cDNA amplification. Libraries were prepared with the Chromium Single Cell 3' Reagent Version 3 Kit. Each sample was sequenced on HiSeq4000 (Illumina) in Rapid Run Mode with paired-end sequencing parameters for an average depth of 50,000 reads per cell.

Computational analyses of sequencing data

Bulk RNA-seq analysis

The mouse GRCm38 assembly was acquired from the Ensembl ftp site (<http://uswest.ensembl.org/info/data/ftp/index.html>) and indexed using kallisto using default settings⁴⁴. Transcript quantifications was performed using 50 bootstrap samples with the remaining options on default settings. The sleuth R package was used for pairwise differential expression using the Wald test, with an FDR <0.05 and a beta value >1⁴⁵. Gene and transcript quantifications were also generated using sleuth. Gene enrichment analysis was performed with differentially expressed genes using the Enrichr software⁴⁶ (<https://maayanlab.cloud/Enrichr/>).

Single cell RNA sequencing (scRNA-Seq) analysis

Mouse and human single cell RNA sequencing was aligned to the 10x genomics pre-built mm10 and hg38 references, respectively, and feature-barcode matrices were generated using cellranger v3 (<https://support.10xgenomics.com/single-cell-gene-expression/software/pipelines/latest/what-is-cell-ranger>). Secondary analysis was performed using the Seurat v3.1 R package⁴⁷. Cells were filtered for a minimum of 500 features, a maximum of 2,500 features and a mitochondrial percentage less than 10% per cell. Read counts were normalized using log normalization and 2,000 variable features were identified using a vst selection method. PCA dimensionality reduction was performed and elbow plots were used to

determine dimensionality. Cluster resolutions were adjusted between 0.3-0.6 accordingly to obtain discrete gene signatures among the clusters. Uniform Manifold Approximation (UMAP) was used to render final single cell composition plots.

The FindAllMarkers function within Seurat was used to identify the top 10 differentially expressed genes, ranked accordingly to average log fold change. Gene set enrichment of scRNA-seq data was performed using PRESTO v1.0⁴⁸ and fgsea v1.1 R packages⁴⁹. The Wilcoxauc function within the presto package was used to develop pre-ranked gene lists, and fgsea was to perform gene set enrichment using 1,000 permutations and the C2 curated gene sets from MSigDB (<https://www.gsea-msigdb.org/gsea/msigdb>). Significantly enriched gene sets were defined by an FDR <0.05 and an NES >2.5. Metagene analysis was performed with AUCCell R package⁵⁰. AUCCell was used to determine the proportion of genes, of the top 20 differentially expressed genes identified from t=0, expressed at t=5. A minimum AUC cutoff of 0.1 was used to determine the cells expressing a minimum of 2 genes of the original t=0 signatures. Single cell differentiation status was determined using non-normalized read count matrices with the cytoTRACE R package²⁰ or the web-based software (<https://cytotrace.stanford.edu>). Genes correlated with 'common pre-cancer' cells, or immature cells, were identified using the cytoTRACE software.

For trajectory analysis Monocle v2.90 was used^{22,51}. Single cell matrices files were initially processed, as described above, and then read in as a cell dataset with a lower detection limit of 0.7 and filtered using a minimum expression of 0.5. Dimensionality was determined using elbow plots and reduced using a tSNE reduction method. Unsupervised clustering of the cells was performed, and the rho and delta values were adjusted to the number of clusters previously identified by Seurat⁴⁷. Differential expression analysis according to the clusters was performed, and then ranked according to q-value. Dimensionality reduction was performed again using a DDRTree method and cells were ordered along a trajectory. The number of end-point states were not specified in the ordering of the cells.

Genomic alterations from scRNA-seq data was determined using InferCNV⁵². As the 'common pre-cancer' cells served as a nature "spike-in" control and were determined as common precursors for the different tumor fates, 'common pre-cancer' cells were used as reference cells in the InferCNV analysis. A cutoff of 0.1 and a HMM-based CNV prediction method was used in the InferCNV analysis. The total CNVs and amplified regions/genes were determined individually in each of the biological replicates.

Single cell assay for transposase accessible chromatin sequencing (scATAC-seq) analysis

Mouse pancreas or tumors were dissociated and processed for single cell capture using protocol mentioned above. Captured single cells were then further processed for single cell ATAC sequencing using 10x Genomics' instruction. Mouse single cell ATAC sequencing was aligned and accessibility counts were generated using Cell Ranger ATAC v2.0.0 with the Cell Ranger provided mm10-1.2.0 mouse assembly. Signac v1.1.1 was used for secondary analysis of mouse scATAC-seq results (**Stuart et al., 2020**). Cells were filtered for outliers with the following QC metrics: peak region fragment length of >3,000 and <100,000, reads in peaks >40, blacklist ration <0.025, nucleosome signal >4 and TSS enrichment > 2. The first LSI component was not used, as it captured sequencing depth variation. UMAP non-linear reduction was used and clustering was perform using the SLM algorithm at a resolution of 0.2-1.2.

STRING interactome network analysis

In order to understand the broader pathways that might be driven by amplified genes identified in mouse ACC and ASCP samples, we used network analysis in Cytoscape⁵³. Briefly, we used a core set of amplified genes that were conserved across three replicate ACC and ASCP mouse tumors to seed and propagate independent ACC and ASCP networks within the mouse STRING interactome, including only high confidence interactions (>0.8)⁵⁴. The STRING interactome contains known and predicted functional protein-protein interactions. The interactions are assembled from a variety of sources, including genomic context predictions, high throughput lab experiments, and co-expression databases; interaction confidence is based on weighted combination of all lines of evidence, with higher quality experiments contributing to a higher confidence score (score is out of 1.0 as the highest). To explore neighboring genes, we integrated 66 conserved ACC amplified genes or 87 conserved ASCP amplified genes into the STRING interactome and used them to run the diffusion algorithm within Cytoscape. The diffusion algorithm ranks all nodes within the interactome using a heat value, indicating how closely each node interacts with our set of amplified seed genes. Extracting the top 500 ranking "hot" genes, we assembled a propagated "hot" network for both ACC and ASCP. In order to identify the larger programs enriched within these networks, we next wanted to cluster closely related nodes and determine their functional annotations. We used the community clustering algorithm in clusterMaker to cluster each network, extracting the largest connected, clustered component for further functional enrichment analysis and visualization. Functional enrichment was run within Cytoscape using STRING, and representative terms were chosen from the significant results and

used to annotate each cluster. Nodes in each network are colored by cluster, with seed genes colored in cyan and identified by a triangular node shape and gene name annotation.

Figures and Legends:

Figure 1 (Part I): Generation of *Msi2-Cre^{ERT2}* knock-in mice

(A) Schematic representation of wildtype and recombined *Msi2* alleles, with flanking BamH1 restriction sites and the position of southern blot 3' probe (P). The CreER^{T2} coding sequence (CDS) was cloned into the ATG site of *Msi2*. **P**= Southern blot probe. **(B)** Southern blot of 5 different mouse tissues heterozygous for the targeted allele, compared with one homozygous for the wildtype allele (WT) from C57BL/6 mouse. **(C-F)** Immunofluorescence images showing expression of *Msi2* (green, upper panel) and tissue specific markers (red, lower panel) in brain (Sox2, **C**), lung (CCSP, **D**), kidney (Calbindin, **E**) and pancreas (DBA, **F**). Scale= 25 μ m.

Figure 1 (Part I): Generation of *Msi2*-Cre^{ERT2} knock-in mice

Figure 1 (Part II): Generation of *Msi2-Cre^{ERT2}* knock-in mice

(G-J) Upper Panel. Immunofluorescence images showing expression of tdTomato representing Msi-driven Cre activity (red). Middle Panel, expression of Sox 2 in brain (green, **G**), CCSP in lung (green, **H**), calbindin 1 in kidney (green, **I** and DBA (white, **J**) and Amylase (green, **J**) in pancreas after 7 days of tamoxifen treatment. Lower Panel shows merged image for each tissue (*du* = duct and *ac*= acinar cells). Scale= 25 μ m.

Figure 1 (Part II): Generation of *Msi2Cre^{ERT2}* knock-in mice

Figure 1 (Part III): *Msi2* expression marks the cell-of-origin for cancers across multiple tissues

(K) Strategy for crossing *Msi2-Cre^{ERT2}* mice to *CAG-LSL-Myc^{T58A}* (*Msi2-Myc*) mice. **(L-O)**, H&E staining of normal tissue sections of wildtype (upper panel) and tumor tissue sections of *Msi2-Myc* mice (middle, and lower panel at higher magnification). Brain **(L, L' and L'')**, lung **(M, M' and M'')**, stomach **(N, N' and N'')** and pancreas **(O, O' and O'')**. Scale= 50 μ m. **(P)** Bar graph showing the overall tumor frequency/penetrance in multiple organs in the *Msi2-Myc* model.

K

Organ specific cancers

P

Organ specific tumor frequency

Figure 1 (Part III): Msi2 expression marks the cell-of-origin for cancers across multiple tissues

Figure 2: Targeted expression of Myc in Msi2+ cells drives multiple pancreatic cancer subtypes in Msi2-Myc mice

(A-D), (top row) H&E and (middle row) immunofluorescence images showing distinct lineage specific markers (Amy= green, CK19= white, p63= red) in representative PDAC (A), ASCP-SD (B), ASCP (SF) (C), ACC (D), and anaplastic (E) tumors in *Msi2-Myc* mice. (Bottom row) Immunofluorescence images showing Myc expression (red) in representative PDAC (A), ASCP-SD (B), ACC (D), and anaplastic (E) tumors. C', Immunofluorescence images showing CK5 (red) and CK14 (green) expression in representative ASCP-SF. E'-E'', Immunofluorescence images showing lack of expression of Chromogranin A (CHGA= red), Amylase (green) and Synaptophysin (SYP= red) in Anaplastic tumor. (F) H&E image of a representative tumor with acinar-ductal mixed histology (top: low magnification; bottom: higher magnification); arrow= Ductal tumor, asterisk (*) = ACC, arrowhead = ACC/PDAC admix. E-Cadherin (ECADH= green), Myc (MYC= red). Scale= 50 μ m. (G) Principal component analysis comparing transcriptome of Msi2+ pancreatic cells from 8-week old Msi2-GFP mice ($t=0$), ACC and PDAC/ASCP tumor cells from *Msi2-Myc* mice and KPC tumor cells¹⁹. (H) Heatmap comparing expression of known classification markers in ACC and PDAC/ASCP from *Msi2-Myc* mice and KPC tumors. (I) Bar-graphs showing relative expression of total and delta-N-terminal (Δ N) Trp63 transcripts in ACC, ASCP and KPC pancreatic tumors.

Pancreatic cancer subtypes in *Msi2-Myc* mice

Figure 2: Targeted expression of Myc in *Msi2*⁺ cells drive multiple pancreatic cancer subtypes in *Msi2-Myc* mice

Rajbhandari, Hamilton *et al*

Figure 3: Immature Msi2+ ductal cells are the target of transformation by Myc

(A) Schematic of strategy used to obtain mouse pancreas cells for scRNA-seq analyses before and after tamoxifen induction of oncogenic Myc. **(B)** UMAP plot showing clusters of different Msi2-expressing cell types from normal adult mouse pancreas cells sorted at 8 weeks of age (t=0). **(C)** UMAP plot generated by cytoTRACE analysis shows identification of ductal cell cluster D1 as the most undifferentiated Msi2-expressing pancreas cells (red), D2 as more differentiated (yellow/orange) and D3 as the most differentiated (green) in normal adult pancreas. **(D)** Violin-plot showing expression of prominent stem cell associated genes in different Msi2-expressing pancreatic cell clusters. **(E)** Heatmap showing top 10 genes enriched in each cluster-(highlighting 3 genes from each cluster) of Msi2-expressing cells in adult pancreas. **(F)** UMAP plot showing most undifferentiated population identified by CytoTRACE in the pre-cancerous population from *Msi2-Myc* mice 5-weeks post tamoxifen (t=5) **(G)** Violin plots showing enriched expression of oncogenic and stemness-associated signature genes in cells at t=5. **(H)** GSEA map showing enrichment of stem cell and cell cycle genes in pre-cancerous undifferentiated cells. **(I)** Metagene analysis showing projection of ductal gene signatures (blue) from normal Msi2-expressing pancreatic cells at t=0 on the undifferentiated cluster from three different tumors at midpoint (left, MDP1-3), with the D1 signature being most aligned with MDP lesions (highest frequency of blue across samples). The total number of cells expressing the indicated gene signatures is indicated (right). **(J)** Immunofluorescence staining shows co-expression of 2 distinct lineage specific markers within an organoid following *in vitro* activation of *CAG-LSL-Myc^{T58A}* (+Cre) and lack of acinar lineage marker in control organoids (-Cre) (Amy= green, DBA= white, DAPI= blue).

Figure 3: Immature Msi2+ ductal cells are the target of transformation by Myc

Figure 4 (Part I): Tracing the evolution of divergent pancreatic cancer subtypes

(A) UMAP plot showing expression of common classification markers in ACC tumors at t=12. **(B)** UMAP plot showing differentiation status and the high concordance of the most undifferentiated cells (left, red) with *Hmg2b* and *Top2a* expression (middle). Increase in CNVs (indicated by increasing purple, right) parallels increasing differentiation in end-stage ACC. **(C)** Trajectory map showing pseudotime progression of ACC cells, where the green cells are *Top2a/Hmg2b* cells at origin and the red cells are the most advanced ACC tumor cells (left); heatmap showing the gene expression profile changes with ACC progression (right) **(D)** UMAP plot showing expression of common classification markers in a PDAC/ASCP tumor. **(E)** UMAP plot showing differentiation status and the high concordance of undifferentiated cells (left, red) with *Hmg2b* and *Top2a* expression (middle). Increasing number of CNVs (right) parallels rising differentiation in an end-stage PDAC/ASCP tumor. **(F)** Trajectory map of a PDAC/SF tumor (left) and heatmap showing gene expression profile changes during PDAC/SF progression (right). **(G)** UMAP plot showing expression of common classification markers in an ASCP tumor. **(H)** UMAP plot showing differentiation status and the high concordance of undifferentiated cells (left) with *Hmg2b* and *Top2a* expression (middle). Increasing number of CNVs (right) parallels rising differentiation in an end-stage ASCP tumor. **(I)** Trajectory map of ASCP tumor (left) and heatmap showing gene expression profile changes during ASCP progression (right).

Figure 4 (Part I): Tracing the evolution of divergent pancreatic cancer subtypes

Figure 4 (Part II): Tracing the evolution of divergent pancreatic cancer subtypes

(J) Trajectory map showing pseudo-time progression of a tri-lineage tumor showing ACC, PDAC and ASCP tumor cells arising from *Top2a/Hmg2b* expressing common pool of pre-cancer cells. **(K)** InferCNV analysis of multi-lineage pancreatic tumors (t=12) showing a shared CNV (arrow) signature among different fates (ACC/ASCP) in 2 independent tumor samples (Tumor #1 and Tumor #2), indicating a shared origin. **(L)** Integration of 'common pre-cancer' cells from t=5 and t=12 shows multilineage states - expression of both ductal (*Cdh1*) and acinar (*Cpa1*) lineage markers. **(M)** Common core signature genes of 'common pre-cancer' cells across time (t=5 and t=12) from *Msi2-Myc* mice. **(N)** Three biological replicates of ACC tumors display both highly conserved (bold arrows, 3/3) and less frequently conserved (thin arrow) CNVs. **(O)** Three independent mouse ASCP tumors showing highly conserved (3/3, bold arrows), and less frequently conserved (2/3, thin arrows) CNVs. **(P)** InferCNV analysis of three independent mouse midpoint pre-cancerous lesions (t=5) showing recurrent and conserved fate specific (Fate #1 vs Fate #2) CNV profiles in Chromosome 6 (arrows). **(Q)** UMAP plots of early stage (t=5) tumors showing early transcriptional divergence towards ACC and PDAC/ASCP fates preceding emergence of CNVs. **(R)** UMAP plot showing emergent pancreatic tumor fates (Pre-ACC and Pre-Ductal) and their corresponding CNV footprints in a pre-cancer lesion (t=5).

Figure 4 (Part II): Tracing the evolution of divergent pancreatic cancer subtypes

Figure 5: Identification of dependencies of pancreatic adenosquamous cancers.

(A) Venn diagram showing results of pairwise differential expression analysis and genes uniquely enriched in ACC and ASCP tumors (left). Heatmaps showing selective gene sets enriched in adenosquamous cancers in *Msi2-Myc* mice. (B) Relative colony count showing the 3D growth impact of shRNA-mediated inhibition of corresponding genes in mouse ASCP cells *in vitro* (n=3). (C) Relative colony count showing impact of shRNA-mediated inhibition of corresponding genes in human ASCP cells (BxPC3) *in vitro* (n=3). (D) Immunofluorescence images showing increased expression of cytokine- IFN γ in representative adenosquamous tumor sample in *Msi2-Myc* mice (n=3). (CK14= green, IFN γ = red, DAPI= blue). (E) Immunofluorescence images showing increased expression of IFN γ in a representative human adenosquamous tumor sample (n=4/5). (CK14= green, IFN γ = red, DAPI= blue). (F) Effect of shRNA-mediated *Ifn γ* inhibition on the growth of *Msi2-Myc* derived ASCP tumor cells transplanted into the flank of recipient mice. (G) Summary schematic of pancreatic cancer subtype initiation and evolution in *Msi2-Myc* tumor model where *Msi2-Cre^{ERT2}* drives the overexpression of mutant Myc (*CAG-LSL-Myc^{T58A}*), preferentially targeting the most immature subfractions of pancreatic ductal cells. This leads to the expansion of multilineage cells that subsequently give rise to a pool of common pre-cancer cells that subsequently follow distinct trajectories to evolve into differentiated tumor subtypes.

A

Differential Expression Analysis & Identification of ASCP Enriched Genes

B

C

D

E

Primary patient samples

F

Flank tumor growth (Mouse ASCP cells)

Figure 5: Identification of dependencies of Pancreatic Adenosquamous Cancers

G

Figure 5 (Part II): Identification of dependencies of Pancreatic Adenosquamous Cancers

Rajbhandari, Hamilton *et al*

Supplemental Figure S1: *Msi2*⁺ cells mark the cell-of-origin for multiple cancers

(A) Immunofluorescence (IF) staining of serial sections showing MYC (red) expression in AQP1 (red) expressing tumor cells in the brain of an *Msi2-Myc* mouse. **(B-C)** IF images showing expression of MYC (red) in E-cadherin (ECADH= green) expressing epithelial tumor lesion in lung **(B)** and stomach **(C)** in *Msi2-Myc* mice; scale = 25 μ m. **(D-E)** Hematoxylin and Eosin (H&E) staining showing kidney dysplasia from age matched adult wildtype **(D)** and *Msi2-Myc* **(E)** mice; scale = 50 μ m.

Supplemental Figure S1: Msi2⁺ cells mark the cell-of-origin for multiple cancers

Supplemental Figure S2: Msi2-GFP reporter mouse recapitulates Msi2 expression in adult pancreas

(A) Immunofluorescence images showing detectable expression of Msi2 in islet (*is*) and ducts (*du*) but not in acinar (*ac*) cells in adult wildtype pancreas. **(B)** Immunofluorescence images showing predominant co-expression of Musashi-2 (MSI2= red) and GFP (green) in islets (*is*) and ducts (*du*; DBA+, white), with some rare GFP+ acinar (*ac*) cells (arrowhead) in pancreata from adult Msi2-GFP reporter mice. Scale= 50 μ m.

Supplemental Figure S3: *Msi2-Cre^{ERT2}* deletion predominantly marks duct cells in neonatal and adult pancreas

(A) Immunofluorescence image showing recombination, marked by tdTomato+ cells, in the pancreas 3 days after Tamoxifen induction in 1-day old (d1) *Msi2CreERT2; R26R-LSL-tdTomato* mice (left). Bar graph showing quantification of cell-type specific recombination in the pancreas (right). **(B)** Immunofluorescence image showing recombination, marked by tdTomato+ cells, in the pancreas 10 weeks after Tamoxifen induction in 1 day old (d1) *Msi2CreERT2; R26R-LSL-tdTomato* mice (left). Bar graph showing quantification of cell-type specific recombination in the pancreas (right). **(C)** Immunofluorescence image showing recombination, marked by tdTomato+ cells, in the pancreas 7 days after Tamoxifen induction in adult (60 days old, i.e d60) *Msi2CreERT2; R26R-LSL-tdTomato* mice (left). Bar graph showing quantification of cell-type specific recombination in the pancreas (right). **(D)** Immunofluorescence image showing recombination, marked by tdTomato+ cells, in the pancreas 60 days after Tamoxifen induction (chase) in adult (60 days old, i.e d60) *Msi2CreERT2; R26R-LSL-tdTomato* mice (left). Bar graph showing quantification of cell-type specific recombination in the pancreas (right).

Supplemental Figure S4: HMGB2 and TOP2a expression is conserved in human pancreatic cancer and associated with poor prognosis

(A) scRNA-Seq showing conventional gene markers of PDAC and ASCP cell clusters in human pancreatic cancer. (B) cytoTRACE-generated UMAP plot showing relative immature status of ASCP cluster (left) and increase in CNVs with increasing differentiation in scRNA-seq of human pancreatic squamous tumor sample. (C) Metagene analysis showing conservation of mouse common core 'common pre-cancer' signature onto the *TOP2A/HMGB2* expressing immature cell clusters in human pancreatic cancer. (D) Representative IHC image showing *TOP2A* expression in human pancreatic tumor sample. (E) Kaplan-Meier Survival plot showing poor survival in pancreatic cancer patients with high expression of *TOP2A*. (F) Representative IHC image showing *HMGB2* expression in human pancreatic sample. (G) Kaplan-Meier Survival plot showing poor survival in pancreatic cancer patients with high *HMGB2* expression (Source: The Human Protein Atlas, www.proteinatlas.org).

A**Human Pancreatic Tumor Gene Markers** **B****Human Pancreatic Immature Tumor Cells & CNVs** **C****Mouse “Common Pre-Cancer” (CPC) Signature Expressed in Human PDAC/ASCP**

Human Cells Expressing >25%
Mouse CPC Core Signature

TOP2A *HMGB2* **D** **E** **F** **G**

Supplemental Figure S5: Fate specific shifts in genes and pathways as ‘common pre-cancer’ cells progress to ACC or ASCP tumors

(A) GSEA showing top enriched pathways in differentiated mouse ACC tumor cells. **(B)** Heatmap showing relative gene expression changes as ‘common pre-cancer’ cells progress to end-stage ACC cells. **(C)** GSEA showing top enriched pathways in differentiated mouse ASCP tumor cells. **(D)** Heatmap showing relative gene expression changes as ‘common pre-cancer’ cells progress to end-stage ASCP cells.

A **Enriched Molecular Programs in differentiated ACC Cells**
ACC Tumor #1 (T=12)

B **Differential gene expression**

C **Enriched Molecular Programs in differentiated ASCP Cells**
Ductal Tumor (T=12)

D **Differential gene expression**

Supplemental Figure S5 : Fate specific shifts in genes and pathways as ‘common pre-cancer cells’ progress to ACC or ASCP tumors

Supplemental Figure S6: Gene amplifications in human and mouse pancreatic cancer subtypes are conserved

(A) Venn diagram showing gene-level comparison of known human ACC amplified genes and all identified mouse ACC amplified genes. **(B)** Venn diagram showing gene-level comparison of known human amplified ASCP genes and all identified mouse amplified ASCP genes.

A

**Conservation of Amplified Genes
between Mouse and Human ACC Tumors (T=12)**

B

**Conservation of Amplified Genes
between Mouse and Human ASCP Tumors (T=12)**

Supplemental Figure S7: Epigenetic and transcriptional priming occurs in pre-cancer and ASCP cells.

(A) scATAC-seq of pre-cancer cells at $t=5$ showing two distinct epigenetic cell clusters, one uncommitted cell cluster with open chromatin at ductal and acinar lineage genes and another ductal-fated cell cluster with open chromatin only at ductal lineage genes. **(B)** scATAC-seq of ductal/ASCP tumor at $t=12$ showing epigenetic priming of ASCP-associated lineage genes.

A**Epigenetic Priming of Pre-Cancer Cells t=5**

Midpoint #4
of cells = 199

B**Open Chromatin of Lineage-Specific Loci in ASCP Tumor t=12**

of cells = 4,481

References:

- 1 Siegel, R. L., Miller, K. D. & Jemal, A. Cancer statistics, 2020. *CA Cancer J Clin* **70**, 7-30, doi:10.3322/caac.21590 (2020).
- 2 Collisson, E. A., Bailey, P., Chang, D. K. & Biankin, A. V. Molecular subtypes of pancreatic cancer. *Nat Rev Gastroenterol Hepatol* **16**, 207-220, doi:10.1038/s41575-019-0109-y (2019).
- 3 Kleeff, J. *et al.* Pancreatic cancer. *Nat Rev Dis Primers* **2**, 16022, doi:10.1038/nrdp.2016.22 (2016).
- 4 Boyd, C. A., Benarroch-Gampel, J., Sheffield, K. M., Cooksley, C. D. & Riall, T. S. 415 patients with adenosquamous carcinoma of the pancreas: a population-based analysis of prognosis and survival. *J Surg Res* **174**, 12-19, doi:10.1016/j.jss.2011.06.015 (2012).
- 5 Lenkiewicz, E. *et al.* Genomic and Epigenomic Landscaping Defines New Therapeutic Targets for Adenosquamous Carcinoma of the Pancreas. *Cancer Research*, doi:10.1158/0008-5472.CAN-20-0078 (2020).
- 6 Borazanci, E. *et al.* Adenosquamous carcinoma of the pancreas: Molecular characterization of 23 patients along with a literature review. *World J Gastrointest Oncol* **7**, 132-140, doi:10.4251/wjgo.v7.i9.132 (2015).
- 7 Al-Hader, A., Al-Rohil, R. N., Han, H. & Von Hoff, D. Pancreatic acinar cell carcinoma: A review on molecular profiling of patient tumors. *World J Gastroenterol* **23**, 7945-7951, doi:10.3748/wjg.v23.i45.7945 (2017).
- 8 Clark, C. J., Graham, R. P., Arun, J. S., Harmsen, W. S. & Reid-Lombardo, K. M. Clinical outcomes for anaplastic pancreatic cancer: a population-based study. *J Am Coll Surg* **215**, 627-634, doi:10.1016/j.jamcollsurg.2012.06.418 (2012).
- 9 Hayashi, A. *et al.* A unifying paradigm for transcriptional heterogeneity and squamous features in pancreatic ductal adenocarcinoma. *Nature Cancer* **1**, 59-74, doi:10.1038/s43018-019-0010-1 (2020).
- 10 Chan-Seng-Yue, M. *et al.* Transcription phenotypes of pancreatic cancer are driven by genomic events during tumor evolution. *Nat Genet* **52**, 231-240, doi:10.1038/s41588-019-0566-9 (2020).
- 11 Holen, K. D. *et al.* Clinical characteristics and outcomes from an institutional series of acinar cell carcinoma of the pancreas and related tumors. *J Clin Oncol* **20**, 4673-4678, doi:10.1200/jco.2002.02.005 (2002).
- 12 Ferreira, R. M. M. *et al.* Duct- and Acinar-Derived Pancreatic Ductal Adenocarcinomas Show Distinct Tumor Progression and Marker Expression. *Cell Rep* **21**, 966-978, doi:10.1016/j.celrep.2017.09.093 (2017).
- 13 Lee, A. Y. L. *et al.* Cell of origin affects tumour development and phenotype in pancreatic ductal adenocarcinoma. *Gut* **68**, 487-498, doi:10.1136/gutjnl-2017-314426 (2019).
- 14 Guerra, C. *et al.* Pancreatitis-induced inflammation contributes to pancreatic cancer by inhibiting oncogene-induced senescence. *Cancer cell* **19**, 728-739, doi:10.1016/j.ccr.2011.05.011 (2011).
- 15 La Rosa, S. *et al.* c-MYC amplification and c-myc protein expression in pancreatic acinar cell carcinomas. New insights into the molecular signature of these rare cancers. *Virchows Archiv* **473**, 435-441, doi:10.1007/s00428-018-2366-5 (2018).
- 16 Witkiewicz, A. K. *et al.* Whole-exome sequencing of pancreatic cancer defines genetic diversity and therapeutic targets. *Nat Commun* **6**, 6744, doi:10.1038/ncomms7744 (2015).
- 17 Madisen, L. *et al.* A robust and high-throughput Cre reporting and characterization system for the whole mouse brain. *Nat Neurosci* **13**, 133-140, doi:10.1038/nn.2467 (2010).
- 18 Mollaoglu, G. *et al.* MYC Drives Progression of Small Cell Lung Cancer to a Variant Neuroendocrine Subtype with Vulnerability to Aurora Kinase Inhibition. *Cancer Cell* **31**, 270-285, doi:10.1016/j.ccell.2016.12.005 (2017).

- 19 Johnson, B. L. *et al.* Desmoplasia and oncogene driven acinar-to-ductal metaplasia are concurrent events during acinar cell-derived pancreatic cancer initiation in young adult mice. *PLoS One* **14**, e0221810, doi:10.1371/journal.pone.0221810 (2019).
- 20 Gulati, G. S. *et al.* Single-cell transcriptional diversity is a hallmark of developmental potential. *Science* **367**, 405-411, doi:10.1126/science.aax0249 (2020).
- 21 Srinivas, S. *et al.* Cre reporter strains produced by targeted insertion of EYFP and ECFP into the ROSA26 locus. *BMC Dev Biol* **1**, 4, doi:10.1186/1471-213x-1-4 (2001).
- 22 Trapnell, C. *et al.* The dynamics and regulators of cell fate decisions are revealed by pseudotemporal ordering of single cells. *Nat Biotechnol* **32**, 381-386, doi:10.1038/nbt.2859 (2014).
- 23 Uhlen, M. *et al.* A pathology atlas of the human cancer transcriptome. *Science* **357**, eaan2507, doi:10.1126/science.aan2507 (2017).
- 24 Jakel, C. *et al.* Genome-wide genetic and epigenetic analyses of pancreatic acinar cell carcinomas reveal aberrations in genome stability. *Nat Commun* **8**, 1323, doi:10.1038/s41467-017-01118-x (2017).
- 25 Miyabayashi, K. *et al.* Intraductal transplantation models of human pancreatic ductal adenocarcinoma reveal progressive transition of molecular subtypes. *Cancer Discov*, doi:10.1158/2159-8290.Cd-20-0133 (2020).
- 26 Guerra, C. *et al.* Chronic pancreatitis is essential for induction of pancreatic ductal adenocarcinoma by K-Ras oncogenes in adult mice. *Cancer Cell* **11**, 291-302, doi:10.1016/j.ccr.2007.01.012 (2007).
- 27 Hingorani, S. R. *et al.* Trp53R172H and KrasG12D cooperate to promote chromosomal instability and widely metastatic pancreatic ductal adenocarcinoma in mice. *Cancer Cell* **7**, 469-483, doi:10.1016/j.ccr.2005.04.023 (2005).
- 28 Hingorani, S. R. *et al.* Preinvasive and invasive ductal pancreatic cancer and its early detection in the mouse. *Cancer Cell* **4**, 437-450, doi:10.1016/s1535-6108(03)00309-x (2003).
- 29 Aguirre, A. J. *et al.* Activated Kras and Ink4a/Arf deficiency cooperate to produce metastatic pancreatic ductal adenocarcinoma. *Genes Dev* **17**, 3112-3126, doi:10.1101/gad.1158703 (2003).
- 30 Lin, W. C. *et al.* Dormant cancer cells contribute to residual disease in a model of reversible pancreatic cancer. *Cancer Res* **73**, 1821-1830, doi:10.1158/0008-5472.CAN-12-2067 (2013).
- 31 Sodir, N. M. *et al.* MYC Instructs and Maintains Pancreatic Adenocarcinoma Phenotype. *Cancer Discovery* **10**, 588, doi:10.1158/2159-8290.CD-19-0435 (2020).
- 32 Farrell, A. S. *et al.* MYC regulates ductal-neuroendocrine lineage plasticity in pancreatic ductal adenocarcinoma associated with poor outcome and chemoresistance. *Nature Communications* **8**, 1728, doi:10.1038/s41467-017-01967-6 (2017).
- 33 Sandgren, E. P., Quaife, C. J., Paulovich, A. G., Palmiter, R. D. & Brinster, R. L. Pancreatic tumor pathogenesis reflects the causative genetic lesion. *Proc Natl Acad Sci U S A* **88**, 93-97, doi:10.1073/pnas.88.1.93 (1991).
- 34 Anderson, M. J., Kwong, C. A., Atieh, M. & Pappas, S. G. Mixed acinar-neuroendocrine-ductal carcinoma of the pancreas: a tale of three lineages. *BMJ Case Rep* **2016**, doi:10.1136/bcr-2015-213661 (2016).
- 35 Bosman, F. T., World Health, O. & International Agency for Research on, C. *WHO classification of tumours of the digestive system*. (International Agency for Research on Cancer, 2010).
- 36 International Agency for Research on, C. *WHO classification of tumours of the digestive system*. (International Agency for Research on Cancer, 2019).

- 37 Juiz, N. *et al.* Basal-like and Classical cells coexistence in pancreatic cancer revealed by single cell analysis. *bioRxiv*, 2020.2001.2007.897454, doi:10.1101/2020.01.07.897454 (2020).
- 38 Zhu, L. *et al.* Multi-organ Mapping of Cancer Risk. *Cell* **166**, 1132-1146 e1137, doi:10.1016/j.cell.2016.07.045 (2016).
- 39 Zhu, L. *et al.* Prominin 1 marks intestinal stem cells that are susceptible to neoplastic transformation. *Nature* **457**, 603-607, doi:10.1038/nature07589 (2009).
- 40 Fox, R. G., Park, F. D., Koechlein, C. S., Kritzik, M. & Reya, T. Musashi signaling in stem cells and cancer. *Annu Rev Cell Dev Biol* **31**, 249-267, doi:10.1146/annurev-cellbio-100814-125446 (2015).
- 41 Yousefi, M. *et al.* Msi RNA-binding proteins control reserve intestinal stem cell quiescence. *J Cell Biol* **215**, 401-413, doi:10.1083/jcb.201604119 (2016).
- 42 Fox, R. G. *et al.* Image-based detection and targeting of therapy resistance in pancreatic adenocarcinoma. *Nature* **534**, 407-411, doi:10.1038/nature17988 (2016).
- 43 Assi, M., Danguet, N. & Jacquemin, P. DIE-RNA: A Reproducible Strategy for the Digestion of Normal and Injured Pancreas, Isolation of Pancreatic Cells from Genetically Engineered Mouse Models and Extraction of High Quality RNA. *Front Physiol* **9**, 129, doi:10.3389/fphys.2018.00129 (2018).
- 44 Bray, N. L., Pimentel, H., Melsted, P. & Pachter, L. Near-optimal probabilistic RNA-seq quantification. *Nat Biotechnol* **34**, 525-527, doi:10.1038/nbt.3519 (2016).
- 45 Pimentel, H., Bray, N. L., Puente, S., Melsted, P. & Pachter, L. Differential analysis of RNA-seq incorporating quantification uncertainty. *Nat Methods* **14**, 687-690, doi:10.1038/nmeth.4324 (2017).
- 46 Kuleshov, M. V. *et al.* Enrichr: a comprehensive gene set enrichment analysis web server 2016 update. *Nucleic Acids Res* **44**, W90-97, doi:10.1093/nar/gkw377 (2016).
- 47 Stuart, T. *et al.* Comprehensive Integration of Single-Cell Data. *Cell* **177**, 1888-1902.e1821, doi:10.1016/j.cell.2019.05.031 (2019).
- 48 Browning, B. L. PRESTO: Rapid calculation of order statistic distributions and multiple-testing adjusted P-values via permutation for one and two-stage genetic association studies. *BMC Bioinformatics* **9**, 309, doi:10.1186/1471-2105-9-309 (2008).
- 49 Korotkevich, G., Sukhov, V. & Sergushichev, A. Fast gene set enrichment analysis. *bioRxiv*, 060012, doi:10.1101/060012 (2019).
- 50 Aibar, S. *et al.* SCENIC: single-cell regulatory network inference and clustering. *Nat Methods* **14**, 1083-1086, doi:10.1038/nmeth.4463 (2017).
- 51 Qiu, X. *et al.* Single-cell mRNA quantification and differential analysis with Census. *Nat Methods* **14**, 309-315, doi:10.1038/nmeth.4150 (2017).
- 52 Tickle Timothy, T. I., Georgescu Christophe, Brown Maxwell, Haas Brian. (Klarman Cell Observatory, Broad Institute of MIT and Harvard, Cambridge, MA, USA, 2019).
- 53 Morris, J. H. *et al.* clusterMaker: a multi-algorithm clustering plugin for Cytoscape. *BMC Bioinformatics* **12**, 436, doi:10.1186/1471-2105-12-436 (2011).
- 54 Szklarczyk, D. *et al.* The STRING database in 2017: quality-controlled protein-protein association networks, made broadly accessible. *Nucleic Acids Res* **45**, D362-d368, doi:10.1093/nar/gkw937 (2017).