Anthony Gavalas Group – PhD Project 1

Adult pancreatic centroacinar cells as cells of origin of pancreatic cancer

Pancreatic ductal adenocarcinoma (PDAC) represents over 90% of pancreatic cancer cases and it is among the most lethal forms of cancer. Kras mutations (Kras*) that result in permanent activation of the protein are the initiating event in more than 90% of PDAC cases and are required for tumor maintenance. The cells of origin of PDAC remain unknown and this hinders efforts to identify early molecular markers of the disease and understand the core molecular machinery of the transformed cells. Effective targeting of this machinery will help treat the disease and limit the possibility of relapse.

Strong candidate cells of origin would be adult progenitor cells that are mitotically active, and already express Kras. Previous work in our lab showed that Aldh1b1-expressing centroacinar cells (CACs) fulfill these criteria. Additionally, we have found that a functional Aldh1b1 is necessary in a PDAC mouse model for the development of the disease. To evaluate whether PDAC may originate from these cells and elucidate the implication of Aldh1b1 function, we have generated and validated several new mouse models. Ongoing work with these mouse models indicates that the targeted expression of Kras* in Aldh1b1^{+} CACs results in PDAC. Isolated and expanded in vitro Kras* / Aldh1b1 expressing CACs give rise to after orthotopic transplantation in immunocompromised mice.

The project will develop along three axes (a) conclude the experiments described above which establish that Aldh1b1 expressing CACs are cells of origin of PDAC (b) use both genetic and metabolic analyses of in vitro expanded CACs to identify the molecular mechanisms implicated in the Aldh1b1 mediated oncogenic transformation of CACs (c) analyze the early transcriptome and chromatin changes accompanying this process to identify the gene regulatory networks driving this process. The combination of these experiments will identify regulatory networks that would be potential targets for PDAC therapies.

Project Description
Anthony Gavalas, Prof Dr, Dresden

Adult pancreatic centroacinar cells as cells of origin of PDAC and the key role of a metabolic regulator in the transformation process

Project Description

1 Starting Point

1.1 State of the art

Pancreatic ductal adenocarcinoma (PDAC) represents over 90% of pancreatic cancer cases and it is among the most lethal forms of malignancies, with a dismal five-year survival rate ranging globally between 5-11%. It is projected that by 2030 will become the third leading cause of cancer-related deaths in Europe and the second leading cause in the United States [1-3]. It is often diagnosed at an advanced stage due to its asymptomatic nature and the lack of early molecular markers that could be used for effective screening. Activating KRAS mutations are implicated in more than 90% of PDAC cases as the major initiation event and are also required for tumor maintenance [4-6]. Pharmacologically targeting KRAS was not possible but recent developments [7,8] offer renewed hope, particularly for patients with metastases. Identifying pancreatic adult cells that are particularly sensitive to oncogenic transformation would allow the identification of the early molecular events associated with the Kras-mediated metabolic rewiring [6,9,10], the identification of early molecular markers, as well as metabolic or signaling pathways that could be targeted for effective combinatorial therapies. Self-renewing progenitor cells, which proliferate in response to inflammation and already express Kras, would be amenable to oncogenic transformation because a single event, namely the acquisition of a Kras-activating mutation, would be sufficient to transform them. Mouse models have indicated that such cells are rare and may reside in both the acinar and duct compartments but have not identified the responsible (transformable) acinar or duct subpopulation(s).

Three independent gene expression studies of human PDAC, using primary tumors, developed three classification systems with prognostic value for the survival of patients with resected PDAC [11-13]. These were compared in a separate cohort and it was concluded that high-purity tumors could be classified into ‘classical-like’ and ‘basal-like’ subtypes [14]. A reclassification of subtypes using a combined cohort of patients with both primary and metastatic tumors by single-cell sequencing refined the classification into five groups : ‘basal-like A’, ‘basal-like B’, ‘hybrid’, ‘classical A’ and ‘classical B’. Classical and basal-like expression signatures segregated to distinct cell populations within each individual tumor, suggesting that PDAC evolves as a mixture of both expression phenotypes [15]. A comparison of mouse ductal cell-derived and acinar cell-derived tumor signatures with these data revealed that they were enriched in basal-like and classical subtypes of human PDAC, respectively [16]. The co-existence of both signatures within the same tumor suggested that a single progenitor population could act cell of origin. In the mouse, neoplasias of acinar or duct origin proceed to PDAC via distinct pathophysiologies [16-18], raising the possibility that a progenitor population that can give rise to both duct and acinar cells might be a better-suited starting population to model the human disease. Despite its importance, the identification of the molecular profile of the cell(s) of origin in PDAC remains unresolved.

Genetically engineered mouse models (GEMMs) have contributed significantly to our understanding of the disease. A commonly used conditional *Kras** allele (*Kras* ^{LSLG12D}) enables the Cre-mediated expression of the *Kras* ^{G12D} allele from the endogenous *Kras* locus. Initially, it was activated using a *Pdx1::Cre* driver already active in all embryonic pancreatic progenitor (PP) cells [19]. The weak transcriptional activity of the *Kras* locus requires the simultaneous activation of mutated alleles of tumor suppressors such as *CDKN2A, TP53* and *SMAD4* to increase the

penetrance and severity of the cancer lesions [20-22]. However, these mutations arise late in the normal course of the disease [23] and their early activation introduces additional variables and precludes the elucidation of the early specific events following Kras* activation. To bypass this limitation, a conditional *Kras* ^{G12D} allele was inserted in the *Actb* locus (*Actb* ^{iKrasG12D}) [24]. The *Actb* locus is very active transcriptionally in all cells and the activation of the *Actb* ^{iKrasG12D} allele results in high levels of Kras ^{G12D} expression. Such high levels of Kras expression are found in human cancers and can alone drive PDAC development in mice [25]. The *Pdx1::Cre / Actb* ^{KrasG12D} (PK) animals recapitulate all the stages of PanIN progression before developing PDAC and transcriptome analyses showed a very high concordance of the resulting PDAC with human PDACs, indicating the physiological relevance of the model [24]. The *Actb* ^{iKrasG12D} allele simplifies breeding and allows the analysis of effects resulting exclusively from Kras* activation.

To address the concern that the expression of Kras ^{G12D} in all embryonic PP cells in the models above may reflect developmental defects, newer GEMMs employed either the tamoxifen-inducible form of Cre recombinase or the doxycycline-inducible system to express Cre and broadly induce Kras* expression in either acinar or duct cells. Expression of Kras* in acinar cells in the context of cerulein-induced pancreatitis lead to the formation of PanIN lesions and eventually PDAC [26-28]. On the other hand, Kras* expression in duct cells was not sufficient to induce PDAC [28-30] except in the presence of additional mutations, but the efficiency in the induction and the severity of the lesions was weaker than in acinar cells [11,31]. Acinar and duct cells are highly heterogeneous [32-35] and neoplasias in these models develop focally rather than diffusely, suggesting that only a small subset of the targeted cells is competent to undergo transformation. Indeed, expression of Kras* in rare TERT ^{high} self-renewing acinar cells lead to PDAC [36] , but it is not known whether these cells normally express Kras. Thus, the cell(s) of origin remain elusive and the early events following Kras* activation have not been explored.

Centroacinar cells (CACs) are a fourth, understudied cell population in the adult pancreas. They are small, rare cells with minimal cytoplasm and long cytoplasmic extensions [37,38] and connect the acinar lobules with the ductal tree. Based on morphological and molecular criteria, it was suggested that they may represent adult progenitor cells as well as cells of origin of PDAC [39-42]. Their prospective isolation, based on their high Aldh activity, showed that they can generate pancreatospheres *in vitro* containing endocrine and exocrine cells [43]. The identification of Aldh1b1 as a specific CAC marker in our lab enabled further analysis. *Aldh1b1* encodes a mitochondrial enzyme, initially expressed in all PP cells during development. In the adult mouse pancreas, it is exclusively expressed in the CACs and Aldh1b1^{+} CACs expand upon cerulein-induced pancreatitis [44]. Using an *Aldh1b1* ^{CreERT2} knock-in line generated in our lab, we showed that these cells are necessary and sufficient for the formation of self-renewing adult pancreatic organoids. Additionally, we found that Aldh1b1^{+} CACs self-renew and contribute to all three pancreatic lineages in the adult organ during homeostasis. Consistent with this, single-cell RNA sequencing (scRNA-Seq) and pseudotime analyses indicated two distinct differentiation pathways; one towards acinar cells and one towards bipotent duct / endocrine cells. A key feature of the molecular profile of these cells is the preferential expression of small GTPases, including Kras, the expression of which was confirmed by immunofluorescence. Aldh1b1^{+} cells are strongly proliferative and highly enriched in Kras expression as well as Hes1, a key indicator of active Notch pathway, a pathway activated in early-stage human PanINs [39]. Aldh1b1 is overexpressed in human and mouse pancreatic cancers and genetic ablation of *Aldh1b1* in the *PK* mouse model of PDAC (*Aldh1b1* $^{null/null}$ */ Pdx1::Cre / Actb* $^{iKrasG12D/+}$)(*APK mice*) abolished PDAC development [45].

Taking into account all of the above, we suggest that the Aldh1b1^{+} CACs may serve as cells of origin of PDAC. As such, they could explain the findings of the molecular analyses of the human

tumors and provide a unifying concept on the cellular origin of PDAC. Thus, **our central hypotheses are** : **(a) Aldh1b1^{+} CACs are cells of origin of PDAC** and (**b) Aldh1b1^{+} CACs will give unprecedented access to the very first molecular events following Kras ^{G12D} expression.** To address (a), we will conditionally express Kras ^{G12D} exclusively in CACs and carry out orthotopic transplantations of Kras ^{G12D} expressing Aldh1b1^{+} CACs. Preliminary results presented below strongly support this hypothesis. To address (b), we have formulated several possibilities and set corresponding exploration paths. We postulate that Aldh1b1 mediates the Kras-driven metabolic rewiring of transformed cells, chromatin opening to allow for the expression of genes necessary for the oncogenic process and protection of the transformed cells from excessive reactive oxygen species (ROS) levels. To address these hypotheses, we have generated unique mouse models and established all the necessary experimental procedures. Preliminary results on (a) and (b) and our work on embryonic PP cells, support these hypotheses.

In summary, we will assess the potential of Kras* transformed CACs to give rise to PDAC, identify the early molecular events driving Kras* mediated transformation in pancreatic progenitor cells and their dependence on the function of the metabolic regulator Aldh1b1. These analyses may provide new potential targets for combinatorial PDAC therapies.

1.2 Preliminary work

Table 1 List of available compound mouse lines, their corresponding abbreviation and use

Abbreviation	Explanation / Function	Genotype
LT	*Aldh1b1* **L**ineage **T**racing	*Aldh1b1* $^{CreERT2/+}$ / *Rosa26* $^{LSLtdTomato/+}$
LTK	***LT*** and ***K****ras* ^{G12D} expression	*Aldh1b1* $^{CreERT2/+}$ / *Rosa26* $^{LSLtdTomato/+}$ / *Actb* $^{iKrasG12D/+}$
ALTK	***A****ldh1b1 null* and ***LTK***	*Aldh1b1* $^{2CreERT2/fl}$ / *Rosa26* $^{LSLtdTomato/+}$ / *Actb* $^{iKrasG12D/+}$
PK	***P****dx1*-dependent ***K****ras* ^{G12D} expression	*Pdx1::Cre* / *Actb* $^{iKrasG12D/+}$
APK	***A****ldh1b1 null* and ***PK***	*Aldh1b1* $^{null/null}$ / *Pdx1::Cre* / *Actb* $^{iKrasG12D/+}$
DCK	**D**ouble **C**onditional* *Kras* ^{G12D} expression	*Aldh1b1* $^{CreERT2/+}$ / *Pdx1* $^{FlpO/+}$ / *Actb* $^{2iKrasG12D/+}$
ADCK	***A****ldh1b1* null and ***DCK***	*Aldh1b1* $^{2CreERT2/fl}$ / *Pdx1* $^{FlpO/+}$ / *Actb* $^{2iKrasG12D/+}$

* *Pdx1* (FlpO) and *Aldh1b1* (CreERT2) dependent

Aldh1b1 marks adult progenitor cells and is necessary for PDAC development [45]
Immunofluorescence analysis of Aldh1b1^{+} CACs showed that they co-express progenitor as well as ductal and acinar markers. Ki67 immunofluorescence showed that they are much more mitotic than the other pancreatic cells (figure 1A) [45]. These findings suggested that Aldh1b1^{+} CACs might be adult pancreas progenitor cells. To address this, we generated an *Aldh1b1* lineage tracing allele (*Aldh1b1* ^{CreERT2}). To ensure specificity, the CreERT2 coding sequence was inserted in frame after the Aldh1b1 coding sequence, separated by the self-cleaving peptide P2A and without deleting any gene elements. Using *Aldh1b1* $^{CreERT2/+}$ */ Rosa26* $^{LSLtdTomato/+}$ lineage tracing (*LT*) mice, we labelled exclusively Aldh1b1^{+} CACs with high (~ 80%) efficiency (figure 1B). Lineage tracing analysis showed that these cells give rise to cells of all three pancreatic lineages (figure 1C) [45]. Accordingly, sc-RNA Seq analyses of labelled cells one day and four weeks after labelling indicated two differentiation paths, one towards the acinar lineage and one towards duct / endocrine progenitors (figure 1D). *Kras* and *Hes1* were among the key genes expressed in these cells (figure 1E). Immunofluorescence analysis confirmed their expression and demonstrated that CACs are 5-fold and 37-fold enriched in Kras and Hes1 expression, respectively, as compared to the other pancreatic cells (figure 1F-H) [45]. To address whether Aldh1b1 activity was functionally implicated in PDAC development, we genetically eliminated *Aldh1b1* in the *PK* model of PDAC [24]. Strikingly, the *Aldh1b1* ^{null} */ PK* (*APK*) mice did not develop PDAC (figure 1H) and their pancreata were histologically normal [45].

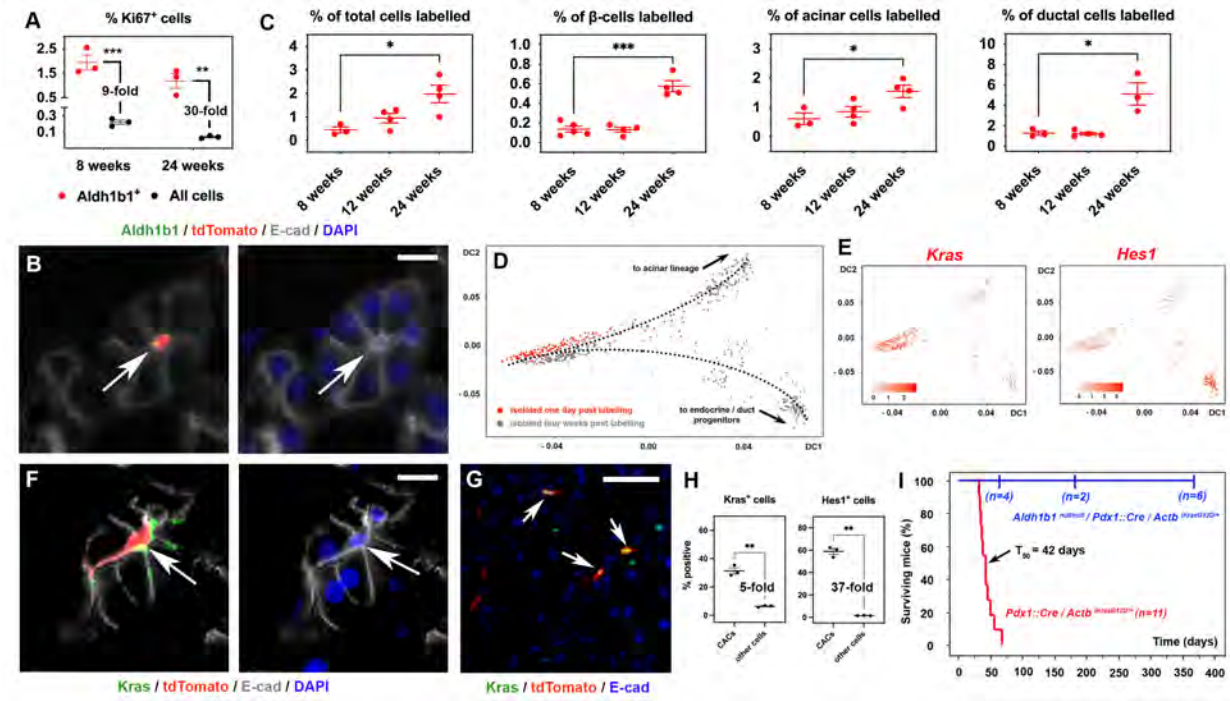

Figure 1. Aldh1b1 marks adult progenitor cells and is necessary for PDAC development. (A) Quantification of Ki67$^+$ Aldh1b1$^+$ and all pancreatic cells. (B) Co-expression of Aldh1b1 and nuclear tdTomato after tamoxifen treatment. (C) Quantification of tdTomato$^+$ cells at three different time points following labeling at 8 weeks of age. (D) Diffusion map and differentiation trajectories of labeled cells isolated at two different time points. (E) Expression of *Kras* and *Hes1* in individual cells. (F-H) Co-expression of tdTomato with Kras (F), Hes1 (G) and quantification of Kras$^+$ and Hes1$^+$ cells in the tdTomato$^+$ and general cell populations (H). Kaplan-Meier survival analysis of *AP* and *APK* mice. Dots in A, B, G correspond to independent experiments and horizontal lines represent the mean ± standard error of the mean (SEM) with $p \leq 0.05$ (*), $p \leq 0.005$ (**) and $p \leq 0.0005$ (***). Scale bars correspond to 20 µm (B, F) and 50 µm (G).

Generation and validation of new mouse alleles. To address whether targeted expression of Kras^{G12D} in Aldh1b1$^+$ cells would result in the development of PDAC, we introduced the *Actb* ^{iKrasG12D} allele in the *LT* line to generate the *LTK* model (Table 1) and activated Kras ^{G12D} expression by tamoxifen administration. Due to the strong Aldh1b1 expression in the intestinal crypt cells and their high proliferation, these mice developed extensive intestinal tumors within 15 days, precluding any meaningful conclusion with regards to the pancreas [45]. To restrict Kras ^{G12D} activation in the pancreatic Aldh1b1$^+$ cells, we had to introduce a second, Flp recombinase-dependent, condition for Kras ^{G12D} activation. We evaluated the efficiency of two available *Pdx1-FlpO* drivers in labeling CACs by crossing them with the *MuX* reporter line [46]. The first (*Pdx1* ^{FlpO}) is a knock-in allele [47] (figure 2A), whereas the second (*Pdx1::FlpO*) is a transgene [48]. The general recombination efficiency for both lines was similar but the *Pdx1::FlpO* allele failed to consistently label Aldh1b1$^+$ cells (7.3% ± 4.1%, n=3). On the other hand, the efficiency of the *Pdx1* ^{FlpO} allele in labeling CACs was similar (20.3% ± 7.5%, n=3) as in the general population (figure 2B, C). This confirmed that the *Pdx1* ^{FlpO} allele can be used to drive Flp recombination in CACs. Then, we generated a double conditional allele (*Actb* ^{2iKrasG12D}) introducing a loxp-flanked Neo cassette as well as an FRT-flanked transcriptional stop (3 x pA) cassette upstream of the Kras^{G12D} coding sequence in the *Actb* locus (figure 2D). Each of these cassettes acts as an efficient transcription termination signal [49] and thus, *Kras* ^{G12D} expression would be initiated only after both FlpO- and Cre-mediated recombination. To assess this, we administered tamoxifen to *Aldh1b1* $^{2CreERT2/+}$ / *Pdx1* $^{FlpO/+}$ / *Actb* $^{dcKrasG12D/+}$ (*DCK*) animals and assessed Kras expression three months later. Only tamoxifen-treated animals showed Kras ^{G12D} expression, confirming the functionality of the allele (figure 2E).

In the *PK* and *APK* models, Kras ^{G12D} is first expressed in the embryonic PP cells. Thus, the lack of PDAC development in the *APK* mice [45] could reflect embryonic defects. This raises the question of whether *Aldh1b1* function is necessary in the adult for PDAC development and, if so, whether

it is required in the presumed cells of origin (CACs) or is induced in neighboring stromal cells to promote their proliferation. To address this, it would be necessary to inactivate *Aldh1b1* in CACs while also activating Kras ^{G12D} expression in the same cells. To this end, we generated two new *Aldh1b1* alleles, a conditional loss-of-function (*Aldh1b1* ^{fl}) allele and a knock-in CreERT2 (*Aldh1b1* ^{2CreERT2}) allele where the Aldh1b1 coding sequence is eliminated. To ensure fidelity, all other gene elements are preserved in both alleles. The *Aldh1b1* ^{fl} allele (figure 2F) was generated using the EMMA (https://www.infrafrontier.eu/emma) *Aldh1b1* ^{tm2a} ES line and its functionality was verified in *Aldh1b1* $^{fl/+}$*/ Pdx1::Cre* P1 pancreata for Cre-mediated deletion of the *Aldh1b1* coding region using appropriate primers (figure 2G). The *Aldh1b1* ^{2CreERT2} allele (figure 2H) was generated in the MPI-CBG transgenic core and its functionality was assessed by crossing it with the *Rosa26* $^{LSLtdTomato/+}$ / *Aldh1b1* $^{2CreERT2/+}$ mice treated with tamoxifen. This allele showed the same specificity and efficiency as the original *Aldh1b1* ^{CreERT2} allele [45], labelling 86.9±4.1% (n=3) of Aldh1b1^{+} cells (figure 2I). Thus, using these two new *Aldh1b1* alleles, we will generate *Aldh1b1* $^{2CreERT2/fl}$ / *Pdx1* $^{FlpO/+}$ */ Actb* $^{LSLKrasG12D/+}$ (*ADCK*) mice in which tamoxifen induction would simultaneously eliminate *Aldh1b1* and induce Kras* expression in Aldh1b1^{+} CACs. This is feasible because the same labelling conditions result in labelling CACs that do not express Aldh1b1 (figure 2J).

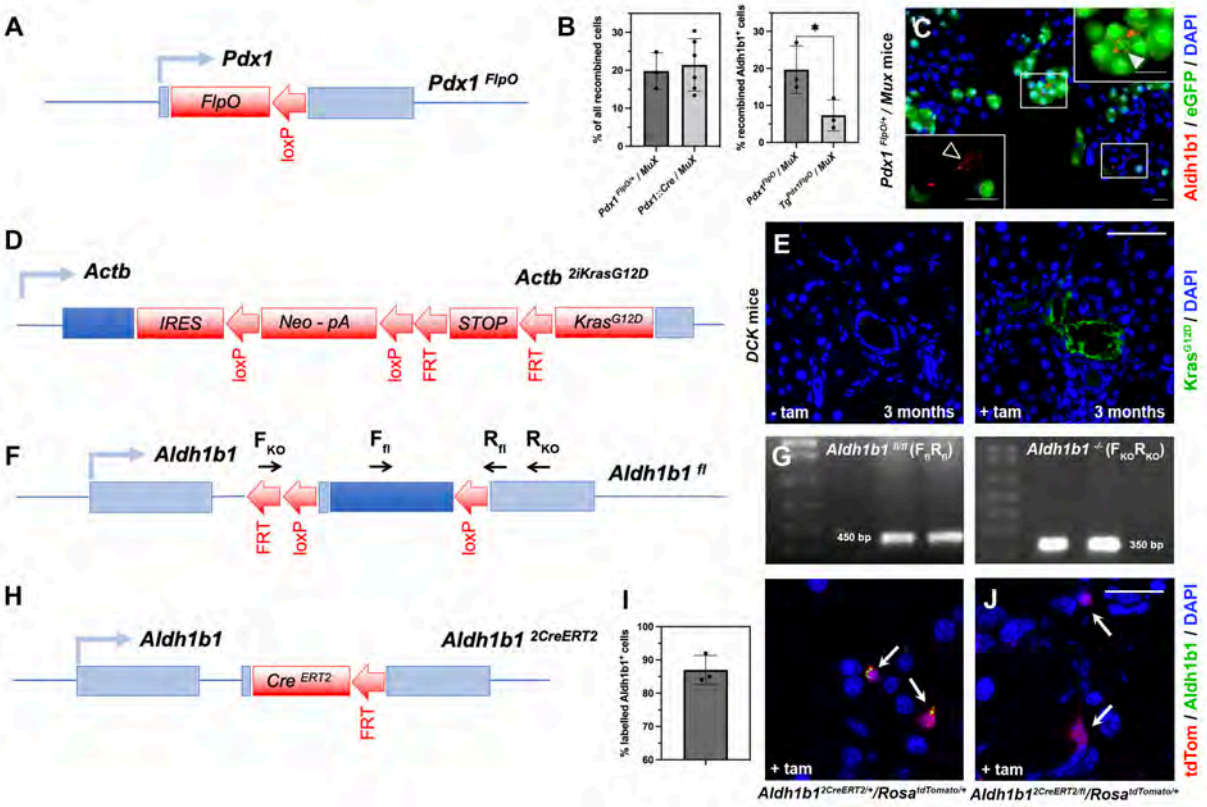

Figure 2. Generation and validation of new alleles. (A-C) Schematic of the *Pdx1* ^{FlpO} allele (A), quantification of the recombination efficiencies of the *Pdx1* ^{FlpO} and the *Pdx1::Cre* alleles (B) as determined by immunofluorescence for colocalization of eGPF and Aldh1b1 (C). (D, E) Schematic of the new *Actb* ^{2iKrasG12D} allele (D) and expression of Kras in DCK animals (E). (F, G) Schematic of the *Aldh1b1* ^{fl} allele and genotyping with specific primers confirming the Cre-mediated elimination of the Aldh1b1 CDS from it (G). (H-J) Schematic of the knock-in *Aldh1b1* ^{2CreERT2} allele and CAC labeling in the absence (I) or presence of the *Aldh1b1* ^{fl} allele (J) as determined by immunofluorescence for colocalization of tdTomato and Aldh1b1. Blue boxes represent coding (dark) or non-coding (pale) exons and lines introns. Added elements are in red. Arrows in F represent the forward (F) and reverse PCR primers for the floxed (fl) and deleted (KO) alleles. Dots on the graphs correspond to independent experiments and horizontal lines represent the mean ± SEM with $p \leq 0.05$ (*). Scale bars correspond to 20 (C), 50 (E) or 25 µm (I, J).

PanINs, PDAC and IPMNs following KrasG12D activation in Aldh1b1^{+} cells. Kras^{G12D} expression was induced in *DCK* mice by tamoxifen injections in two consecutive days and in the context of acute pancreatitis induced by multiple cerulein injections. (figure 3A). Pancreata were analyzed histologically three months later (n=2) by hematoxylin/eosin stainings for the detection

of neoplasias, alcian blue/nuclear fast red staining for the detection of mucins, hematoxylin/picrosirius red for the detection of desmoplasias and immunofluorescence. In contrast to untreated *DCK* pancreata, which appeared normal, treated *DCK* pancreata showed alcian blue positive areas of neoplasias and increased desmoplasia (figure 3B, C). A closer examination of the treated pancreata revealed the presence of acinar to ductal metaplasia (ADM), PanINs 1-3 and PDAC foci as well as intraductal papillary mucinous neoplasms (IPMNs) (figure 3D-F). These findings suggest that Aldh1b1^{+} cells may indeed act as the origin of PDAC but additional analyses are needed to firmly establish that and conduct the same analyses in the absence of pancreatitis.

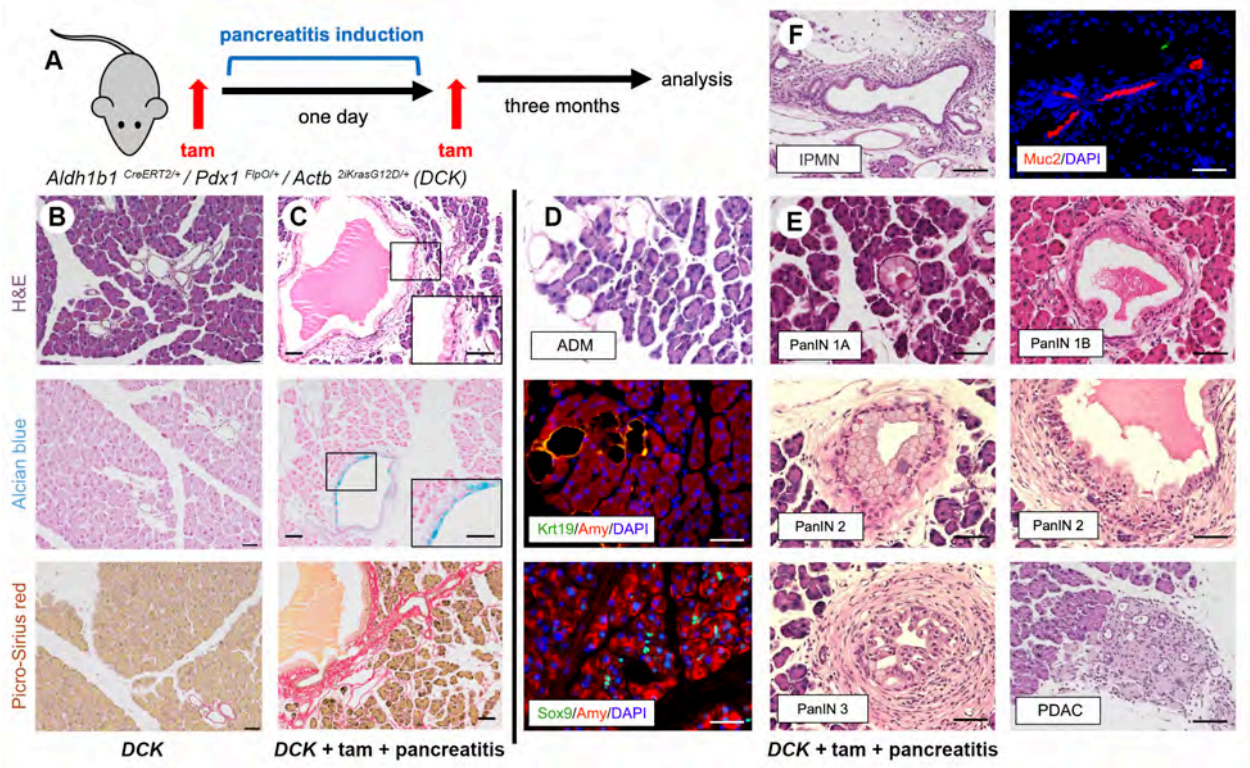

Figure 3. ADM, PanINs, PDAC and IPMNs following Kras ^{G12D} activation in Aldh1b1^{+} cells. (A) *DCK* mice were twice with tamoxifen in the context of acute pancreatitis and pancreata were histologically analyzed three months later. (B, C) Histological analysis of *DCK* control mice (B) and DCK experimental mice treated with both tamoxifen and cerulein (C). (D-F) Detection of ADM by histology and immunofluorescence (D), PanINs and PDAC by histology (E) as well as of IPMNs by histology and immunofluorescence (F). Scale bars correspond to 50 and 20 (inserts) µm (B, C) and 100 µm (histology in F and D PDAC).

In vitro expanded Kras ^{G12D} expressing CACs give rise to PanINs and PDAC after transplantation. To directly assess whether Kras* expressing CACs may act as tumor-initiating cells, we labelled CACs in *LT* and *LTK* mice by tamoxifen administration and isolated them by FACS one day later. PDAC cells were isolated by FACS from *PK* animals based on their high Aldh activity [45,50]. All three types of cells were expanded in vitro as organoids under identical conditions. *LT* cells (n=4) grew significantly more slowly (T_d 3.7±0.3 days) than *LTK* (n=5) or *PK* (n=5) cells, which expanded with very similar doubling time (T_d) (2.8±0.2 vs 2.9±0.1 days), suggesting that CACs are sensitive to the expression of Kras ^{G12D}. After ten passages, one million each of *LT*, *LTK* and *PK* cells were orthotopically transplanted in *Foxn1* $^{null/null}$ (*Nu/Nu*) mice (n=2 for each) and the transplanted pancreata were collected three months later for histological analyses (figure 4A). As expected, there were no abnormalities in the *LT* transplanted pancreata (figure 4B) but *PK*-transplanted pancreata showed strong evidence of PDAC development, mucin production, strong desmoplastic reaction and extensive ectopic expression of Krt19 (figure 4C).

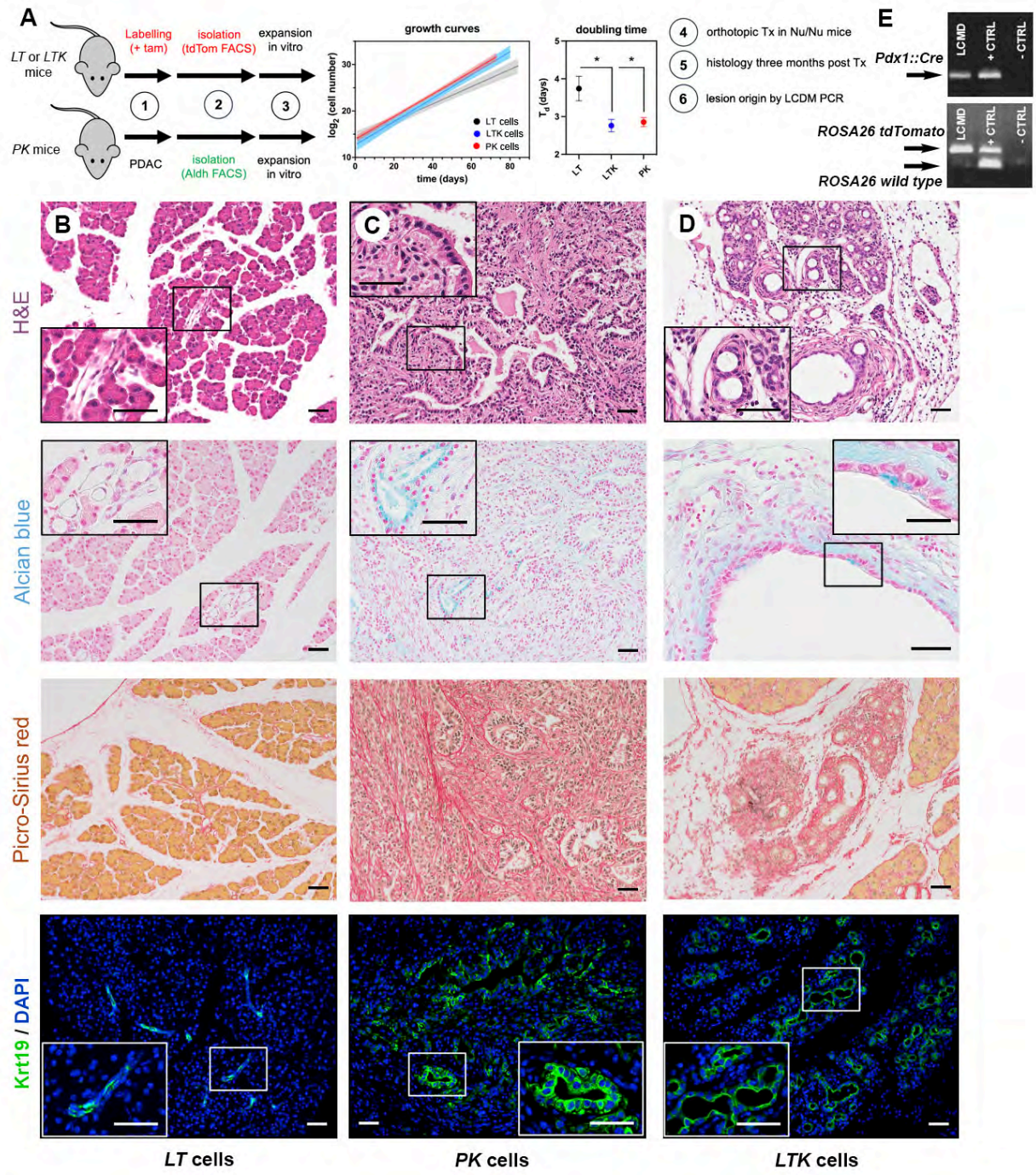

Figure 4. In vitro expanded Kras ^{G12D} expressing CACs result in PanINs and PDAC after transplantation. (A) Experimental design, growth curves with corresponding T_d for isolated and in vitro expanded LT (n=4), LTK (n=5) as well as PK cells (n=5). (B-C) Histological analysis by H&E, alcian blue/nuclear fast red and picro Sirius red as well as Krt19 immunofluorescence of pancreata transplanted with *LT* (B), *PK* (C), or *LTK* (D) cells. (D) PCR genotyping on lesions isolated by LCMD for the Pdx1::Cre and ROSA26 wild type and tdTomato alleles. Positive controls are from mouse ear clips. Shaded areas in the growth curves correspond to 95% confidence interval and horizontal lines represent the mean ± SEM with $p \leq 0.05$ (*). Scale bars correspond to 50 µm (B, C, D).

LTK-transplanted pancreata also contained PanINs, PDAC foci, extensive mucin-positive and desmoplastic areas, accompanied by widespread upregulation of Kr19 (figure 4D). The origins of the lesions were confirmed through their isolation by Laser Capture Micro Dissection (LCMD) and PCR for alleles specific to the transplanted cells (figure 4E). These findings strengthened the hypothesis that Aldh1b1^{+} cells act as the origin of PDAC but additional experiments are needed.

Aldh1b1-mediated metabolic regulation determines chromatin structure and gene expression in embryonic PP cells. We identified Aldh1b1 as a downstream Ngn3 effector [51]

and we then showed that Aldh1b1 regulates the timing of endocrine differentiation and is necessary for the generation of fully functional beta cells [52,53]. These and other studies in our lab [54,55] aimed at increasing the differentiation efficiency of human pluripotent stem (hPS) cells into pancreatic endocrine cells [56,57] but also lead to the identification of Aldh1b1^{+} CACs [45]. We found that FACS-isolated *Aldh1b1* null embryonic PP cells contain significantly higher levels of reactive oxygen species (ROS) (figure 5A). Importantly, *Aldh1b1* null PP cells could not be expanded in vitro as organoids unless the ROS scavenger N-acetyl-cysteine (NAC) was present, suggesting that in the absence of Aldh1b1, ROS levels were toxic for cell survival and expansion (figure 5B).

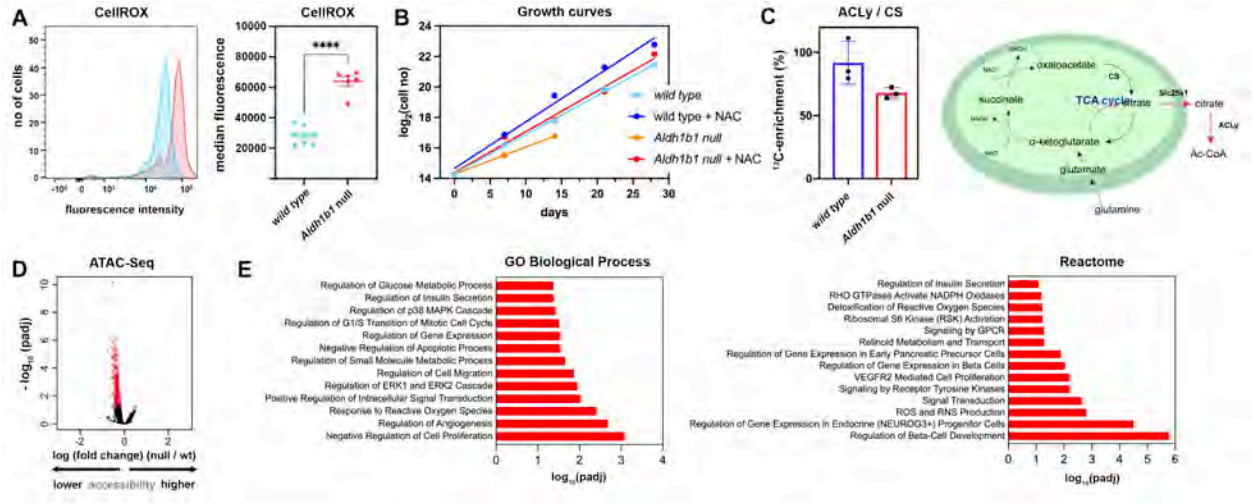

Figure 5. Aldh1b1 promotes Ac-CoA generation and reduces ROS levels in embryonic PP cells. (A) Flow cytometry determination of ROS levels through Cell-ROX staining of E-cadherin^{+} embryonic PP cells. (B) Wild type embryonic PPs expand as spheroids irrespective of the presence of NAC but *Aldh1b1 null* PPs expanded only in its presence (each curve from n=3). (C) Reduced ACLy flux relative to citrate synthase (CS) in *Aldh1b1* ^{null} PP cells isolated by FACS and cultured in vitro as organoids. (D) Functional gene categories regulated by Aldh1b1 expression in embryonic PP cells. (D) Volcano plot showing the number of differentially regulated loci as suggested by ATAC-Seq. Dots on the graphs A-C correspond to independent experiments and horizontal lines represent the mean ± SEM with $p \leq 0.0005$ (****).

Metabolic flux analyses (Alves lab, TU Dresden) [58] on FACS-isolated and expanded embryonic PP cells revealed a reduction in the cytosolic production of acetyl-CoA (Ac-CoA) via ATP citrate lyase (ACLy) in *Aldh1b1* ^{null} cells (figure 5C). Ac-CoA is a rate-limiting substrate for histone acetyltransferases [59,60] and we then asked whether chromatin structure and gene expression were affected. We established sensitive methods that allowed ATAC-Seq and RNA-Seq analyses from 10K cells, FACS-isolated from a single embryonic pancreas. In agreement with the reduction of metabolic flow to Ac-CoA, ATAC-Seq analyses showed that the chromatin of *Aldh1b1* ^{null} embryonic PP cells is more compact (figure 5D). RNA-Seq analyses showed that ROS, signal transduction and endocrine differentiation processes were affected (figure 5E). By analogy to these findings, we hypothesize that Aldh1b1 may protect from deleterious ROS levels, mediate chromatin remodeling and facilitate signal transduction pathways mediating the oncogenic transformation of Kras* expressing CACs.

2 Objectives and work programme

2.1 Anticipated total duration of the project

Three years / thirty-six months

2.2 Objectives

Gene expression analyses of human PDAC samples revealed that behind the cellular heterogeneity there are two gene expression signatures, the classical (acinar-like) and the basal (duct-like) which co-exist. GEMMs have shown that neoplasias can be induced from either duct

or acinar cells but with distinct pathophysiologies Additonally, neither the analysis of human samples, nor the current GEMMS can be used to identify the very first events following Kras* expression and the cellular requirements for transformation. Based on their distinct features, we suggest that CACs may act as cells of origin for PDAC and we have developed unique animal models that allow us to address this hypothesis and analyse the cellular events immediately following Kras* activation. Our preliminary data fully support this line or research. The project is structured into four independent aims.

AIM 1 Assess the effectiveness of Aldh1b1^{+} CACs as PDAC tumor initiating cells.
AIM 2 Determine whether Aldh1b1 function is required in CACs for tumor initiation.
AIM 3 Determine the Kras^{G12D}-driven metabolic changes in CACs and the implication of Aldh1b1.
AIM 4 Determine the Kras^{G12D}-driven transcriptome and chromatin changes in CACs and the implication of Aldh1b1 in these changes.

2.3 Work programme including proposed research methods

AIM 1 Assess the effectiveness of Aldh1b1^{+} CACs as PDAC tumor initiating cells. The preliminary data support the **hypothesis that Aldh1b1^{+} CACs can give rise to PDAC** but a more rigorous assessment of the tumors and the role of inflammation remain necessary.

Aim 1.1 In vivo induction of Kras ^{G12D} expression in CACs. To induce expression of Kras ^{G12D} only in the pancreatic *Aldh1b1* expressing CACs, we will use the *DCK* model and assess the development of PanINs and PDAC in the absence and presence of acute pancreatitis. Seven-week-old *DCK* mice will be tamoxifen-treated to activate expression of the *Actb ^{2iKras}* allele (model 1). The same will be done also in the context of acute pancreatitis (model 2) (figure 3A). Untreated *DCK* mice or subjected only to acute pancreatitis of the same age will be used as the respective controls. For each of the two experimental and controls group, five animals will be scored at 12 and 24 weeks for the development of PanINs, IPMNs and PDAC. A set of five 5 um paraffin sections 50 um apart will be scored for each of the hematoxylin / eosin, alcian blue and picro sirius histological stainings. The number and total area of the foci of mucin production (alcian blue) as well as the total area of desmoplastic reaction (picro sirius red) will be scored and normalized for total tissue area. One-way ANOVA and Tukey's test will be used to determine statistical significance in comparing lesion frequency and severity in the absence and presence of pancreatitis. Immunofluorescence of adjacent paraffin sections will be used to confirm expression of Kras ^{G12D} and assess the extent of Aldh1b1, Sox9, Muc2, Muc5Ac, c-Myc, p-ERK, Ki67 and α-SMA expression in areas corresponding to the lesions. Of particular interest will be whether Aldh1b1 expression is maintained in Kras^{G12D} cells, whether it is expressed in stromal cells (α-SMA^{+} areas) and whether Aldh1b1^{+} cells are more proliferative.

Aim 1.2 Transplantation of in vitro expanded Kras ^{G12D} expressing CACs. Despite the very high specificity of the *Aldh1b1 ^{CreeRT2}* driver [45], we cannot completely rule out the possibility that Kras ^{G12D} expression may still be induced in non-CACs. This is particularly true in the context of pancreatitis, where other pancreatic cells, particularly acinar cells, may dedifferentiate and upregulate *Aldh1b1* in response to the inflammation. To address this, we will administer tamoxifen to seven-week-old *LT* and *LTK* mice to induce tdTomato expression in the CACs of both while also activating Kras ^{G12D} expression in the latter (see also preliminary results). Labelled CACs will be isolated by FACS from a single cell suspension as already established in the lab. As a positive control, PDAC cells will be FACS isolated from *PK* mice as Aldh^{+} cells, using the live Aldefluor assay [45]. All three cell populations will be expanded in vitro as organoids and expanded *LTK* cells will be genotyped to ensure that recombination in the *Actb* locus, and thus Kras ^{G12D} expression, has taken place. One million expanded *LT*, *LTK* and *PK* cells will be orthotopically transplanted in *Nu/Nu* mice and the pancreata of transplanted mice will be assessed 12 and 24 weeks after transplantation by histology and immunofluorescence as described in aim 1.1. The origin of the

lesions will be confirmed independently of tdTomato expression by Laser Capture Microdissection (LCMD) of affected areas and PCR genotyping as we have already established for the *ROSA26* ^{LSLtdTomato} (for LTK cells) and *Pdx1::Cre* alleles (for PK cells).

Experiments in this aim will address the extent that Kras ^{G12D} expression in CACs can drive the formation of IPMNs, PanINs and PDAC and if this is promoted in the context of pancreatitis as well as whether Kras ^{G12D} expressing CACs can initiate tumor formation cell autonomously.

AIM 2 Determine whether Aldh1b1 function is required in CACs for tumor initiation. Since *Aldh1b1* is already expressed in Kras$^+$ CACs and because *Aldh1b1* is necessary for cancer development in the *APK* mice [45], we **hypothesise that Aldh1b1 function is required during the initial stages of Kras ^{G12D} mediated oncogenic transformation of CACs.** The immunofluorescence experiments in 1.1 and 1.2 will give a first indication but to definitively address this, we will generate the *ADCK* (*Aldh1b1* $^{2CreERT2/fl}$ / *Pdx1* $^{FlpO/+}$/ *Actb* $^{2iKrasG12D/+}$) mouse model. In these mice, which carry a single funtional Aldh1b1 (*Aldh1b1^{fl}*) allele, tamoxifen treatment will induce KrasG12D expression and inactivate the *Aldh1b1^{fl}* allele in the Aldh1b1$^+$ CACs. *ADCK* mice will be treated and analyzed exactly as *DCK* mice in Aim 1.1 and lesions will be isolated by LCMD and genotyped for both the recombined and non-recombined *Aldh1b1^{fl}* alleles. The frequency and severity of lesions and the expression pattern of Aldh1b1 will be compared to the results obtained from the *DCK* mice. If Aldh1b1 function is necessary only in stroma cells, its inactivation in the CACs will have no effect on the appearance of PanINs and PDAC development. If Aldh1b1 is required downstream of the Kras ^{G12D} activation, in already transformed cells, it is expected that lesions will be detected but with reduced severity. But if Aldh1b1 function mediates the Kras ^{G12D} – induced transformation, we expect that the frequency and severity of the lesions will be dramatically reduced.

These experiments will determine whether Aldh1b1 is indeed required in CACs in a cell autonomous manner for the Kras^{G12D}-driven oncogenic transformation.

AIM 3 Determine the Kras^{G12D} driven metabolic changes in CACs and the implication of Aldh1b1. Our finding that Aldh1b1 function is necessary for PDAC development in the PK model raised the important question of how this enzymatic function enables the Kras ^{G12D} driven oncogenic transformation. In this aim, we will address the **hypotheses that Aldh1b1 may be a necessary component** for (a) the Kras ^{G12D} driven metabolic reprogramming and (b) ROS detoxification necessary for cancer cell survival.

Aim 3.1 Metabolic reprogramming following Kras^{G12D} induction in CACs and its dependence upon Aldh1b1 function. Kras ^{G12D} driven metabolic reprogramming consists mainly of increased glycolysis, facilitated diversion of glutamine into the TCA cycle through the action of aspartate transaminase and increased lipid synthesis [6,9,61]. To analyze the Kras ^{G12D} driven metabolic effects in CACs we will use the *LT* and *LTK* models and to analyze the implication of Aldh1b1 we will generate the *Aldh1b1* $^{2CreERT2/fl}$ / *ROSA26* $^{LSLtdTomato/+}$ / *Actb* ^{iKrasG12D} *(ALTK)* mice. ALTK mice carry a single funtional Aldh1b1 (*Aldh1b1^{fl}*) allele and tamoxifen administration will inactivate the functional *Aldh1b1^{fl}* allele while also inducing tdTomato and Kras ^{G12D} expression in the Aldh1b1$^+$ CACs. Thus, *LT*, *LTK* and *ALTK* eight-week adult mice will be treated with tamoxifen and, two days later, wild type (LT), Kras expressing (LTK) and Kras expressing / *Aldh1b1* null (ALTK) tdTomato$^+$ cells will be isolated by FACS as established in the lab. Cells will be expanded in vitro as organoids, labeled with either ^{13}C-labeled glucose or ^{13}C-labeled glutamine, and analyzed for glycolytic, TCA and lipid biosynthesis metabolic fluxes in the Alves lab as described previously [58] and as already been done for embryonic PP cells. Genotyping will be used to confirm the elimination of *Aldh1b1* and the activation of Kras^{G12D} expression. Based on our experience, four passages of each isolated sample will provide enough cells for all these

analyses and five independently isolated samples from each genotype will be sufficient to maximize the statistical significance of each analysis. The mathematical analysis of the resulting ^{13}C-labeled time courses will provide a quantitative estimate of the flux through individual reactions and metabolic pathways. This elaborate analyses will assess flux alterations in multiple metabolic pathways and thus pinpoint how exactly Kras ^{G12D} drives metabolic reprogramming in CACs and what is the contribution of Aldh1b1 in this process.

Aim 3.2 Implication of Aldh1b1 in attenuating ROS levels in Kras ^{G12D} expressing CACs. Metabolic reprogramming in cancer cells is accompanied by elevated ROS production, which is counteracted by elevated antioxidant defenses [62]. *Aldh1b1* null embryonic PP cells have significantly higher levels of ROS than their wild type counterparts and cannot expand as organoids in vitro, unless NAC is present. To determine whether, following Kras ^{G12D} expression, a similar phenomenon is taking place in *Aldh1b1* null CACs, TdTomato$^+$ cells will be FACS isolated from *LT*, *LTK* and *ALTK* animals and stained with 2 μM CellROX orange to detect intracellular ROS levels by flow cytometry, as already established in the lab. At least three biological replicates for each genotype will be analysed. If the absence of Aldh1b1 function leads to increased ROS levels, FACS isolated *LT*, *LTK* and *ALTK* cells will be expanded as organoids in the presence or absence of NAC to determine whether Aldh1b1 mediated ROS detoxification is necessary for cell survival and expansion (n=3 for each genotype / condition).

We will then address the role of Aldh1b1 in reducing ROS levels in cancer cells *in vivo*. *APK* mice are resistant to the development of pancreatic cancer [45] and, to substitute the hypothesized function of Aldh1b1 in ROS detoxification, NAC (10 g/lt) will be provided in the drinking water of pregnant mice carrying *APK* embryos, dams of *APK* neonates as well as *APK* mice post weaning. Mice will be kept until postantal week 12 and the survival rate will be determined. Histological analysis and scoring of the pancreata will be done on hematoxylin / eosin stainings. A total of twelve NAC-treated and twelve non-treated *APK* animals, as controls, will be used for these analyses. If Aldh1b1 function is necessary to reduce ROS levels, we expect that cancer will reappear in the treated APK mice.

To complement the approach above, we will use the *DCK* and *ADCK* models in which Kras^{G12D} expression is induced in adult CACs. In analogy with the lack of PDAC development in the *APK* model, we anticipate that, in the *ADCK* model, lesions will be significantly reduced or even absent. Following Kras ^{G12D} activation in the context of acute pancreatitis, *ADCK* mice will be provided with NAC in the drinking water (10 g/lt) and the analyses will be carried out as described for the DCK and ADCK models in Aims 1 and 2. We anticipate that the effects of ROS detoxification will be easier to appreciate in the context of inflammation.

These experiments will determine whether Kras^{G12D} expression in CACs raises their ROS levels and whether Aldh1b1 function is necessary to reduce their levels to promote survival and proliferation of the Kras^{G12D} expressing CACs in vitro and in vivo.

AIM 4 Determine the Kras ^{G12D} driven chromatin and transcriptome changes in CACs and the implication of Aldh1b1. Variations in the concentration of metabolic intermediates, such as acetyl-CoA, S-adenosylmethionine, α-ketoglutarate and the NAD+/NADH ratio, regulate the activity of histone and DNA modifying enzymes [37,38]. In turn, histone and DNA modifications regulate gene expression. Accordingly, *Aldh1b1* loss of function results in chromatin structure and gene expression changes in embryonic PPs. By analogy, we **hypothesise that Aldh1b1 mediates the Kras^{G12D} driven changes in chromatin structure and gene expression** which are necessary for the oncogenic transformation of CACs.

LT, *LTK* and *ALTK* eight-week-old adult mice will be treated with tamoxifen and, two days later, wt, Kras^{G12D} expressing and Kras^{G12D} expressing / *Aldh1b1* null CACs, respectively, will be isolated by FACS as tdTomato$^+$ cells and three independent isolations from each genotype will

be carried out. This procedure yields an average of 25000 tdTomato$^+$ cells per pancreas, a number sufficient for each of the analyses described below including genotyping by qPCR to confirm recombination at the *Aldh1b1* and *Actb* loci in the *LTK* and *ALTK* samples.

Aim 4.1 Histone and DNA modification changes in Kras^{G12D} expressing CACs and their dependence upon Aldh1b1 function. Histone acetylations are associated with transcriptional activation, whereas histone methylations may mark either transcribed or repressed chromatin. To assess whether there is a general shift to active chromatin following the expression of Kras ^{G12D} in CACs and whether this is regulated by Aldh1b1, we will use epigenetics cytometry by time of flight (EpiCyTOF). This is a multiplex approach, allowing the simultaneous use of up to fifty antibodies and the scoring of individual cells, and as such is clearly superior to immunofluorescence or ChIP. The Albert lab (letter of collaboration) and the Mass Cytometry Core Facility in CRTD (https://tu-dresden.de/cmcb/technologie-plattform/facilities/mass-cytometry/expertise_team) have already established a large panel of antibodies for EpiCyTOF including the key H3K9ac, H3K27ac, H4K5ac, H4K16ac, H3K36me2 and H3K36me3 active transcription modifications as well as the key heterochromatin-associated H4K20me3 and H3K9me2 modifications [63]. To eliminate batch effects, cells from each isolation will be uniquely barcoded using the Palladium barcoding system (Standard BioTools). To minimize cell loss, barcoded samples will be mixed with Tantalum-labeled human peripheral blood monocytes (PBMCs) (BioLegend). Labeling of purified antibodies and measurements with the CyTOF XT will be performed in the Mass Cytometry Core Facility in CRTD using their standardized methodology. Conversion of raw counts into flow cytometry standards and data normalization will be performed with the CyTOF software v9.2. Debarcoding and exclusion of the PBMCs will be done with the Catalyst script in R. The flow cytometry analysis platform OMIQ (Dotmatics) will be used for gating, dimensionality reduction and cluster analysis. These analyses will allow us to determine whether there are shifts toward transcriptionally active or primed chromatin upon Kras ^{G12D} expression and if these changes are reduced or abolished in the absence of Aldh1b1 function.

The histone modification profiling will be complemented by the analysis of DNA methylation patterns and the identification of differentially methylated sites using bulk EM-seq. The EM-Seq method uses an enzymatic reaction, rather than bisulfite, to convert unmethylated cytosines to uracils, introducing fewer biases and having significantly higher sensitivity [64]. Extracted DNA from the isolated cells will be processed by the Dresden concept Genome Center (DcGC) (https://dresden-concept.de/genome-center/) using their established experimental procedure. The sequencing data will be processed with an existing pipeline for DNA methylation data. This includes quality control, mapping, removing duplicates and detection of methylation levels of CpGs. Peaks will first be classified based on their proximity or overlap with genes and then according to the gene feature that they overlap with, i.e. transcription start sites (TSS), promoter, exons, introns as well as 5' and 3' untranslated regions (UTRs). Distal peaks will be assigned to the nearest gene. Functional enrichment analyses will be carried out using databases such as GO, KEGG and Reactome.

Aim 4.2 Chromatin structure changes in Kras ^{G12D} expressing CACs and their dependence upon Aldh1b1 function. Independently of the epigenetic changes in Kras ^{G12D} expressing CACs in the presence or absence of Aldh1b1 function, it will be important to independently determine changes in chromatin accessibility because they may take place even in the absence of changes in the abundance of specific histone modifications if modifications are just differentially redistributed. To identify regions of differentially accessible chromatin, we will use the ATAC-Seq assay, a relatively simple procedure, compared to alternative methods, which requires a smaller number of cells and has a high signal/noise ratio and sensitivity [51-53]. In collaboration with the Albert lab (TU Dresden, CRTD), we have set up the method in our lab and can receive reliable results from as few as 10000 cells. Isolated cells will be used to prepare nuclei, carry out genomic

DNA labeling by activated transposase and isolate genomic DNA. Genomic sequencing, sequencing read mapping and normalization will be carried out in the DcGC. Peak calling and identification of regulated peaks will be carried out in DcGC using a pipeline we have developed, in collaboration with the Scientific Computing core in MPI-CBG, specifically for the analysis of ATAC-Seq data. Regulated peaks will be classified and analysed as described for peaks obtained from the EM-Seq analysis.

Aim 4.3 Gene expression changes in Kras ^{G12D} expressing CACs, their dependence upon Aldh1b1 function and cross-validation with human data. We will perform both bulk RNA sequencing and single-cell RNA sequencing (scRNA-seq) on *LT*, *LTK* and *ALTK* CACs. Bulk RNA-seq provides higher sensitivity and robust quantification whereas scRNA-seq resolves cellular heterogeneity and allows the identification of subpopulations and transitional states. To enhance the resolution of CAC differentiation trajectories, samples will be collected at both two days and four weeks following tamoxifen administration [45]. RNA isolated with standard procedures will be used for bulk RNA Seq with the SMART Seq2 procedure and Illumina sequencing. Cells for scRNA-seq will be processed using the 10X chromium platform. Library preparations and subsequent analyses will be carried out in the DcGC.

We will use bulk and scRNA-seq data for differential gene expression analysis between CACs and other pancreatic cell populations (acinar, ductal, endocrine) from publicly available reference datasets (Tabula Muris, Human Cell Atlas, and pancreas-specific atlases) for integrative comparison via cell type annotation, clustering projection, and gene signature scoring. CAC-specific transcriptional programs will be explored through pathway enrichment and regulatory network analyses (SCENIC, DoRothEA). This will allow us to derive a robust CAC gene expression signature cross-validated at both population and single-cell resolution, and benchmarked against known pancreatic cell types.

Differential gene expression analysis between *LT* and *LTK* CACs will identify transcriptional programs activated by oncogenic KRAS. ScRNA-seq will resolve cellular heterogeneity and identify changes in the differentiation trajectories, as determined by pseudotime analyses, of acinar and duct / endocrine progenitors upon Kras^{G12D} expression. Regulatory network analysis (SCENIC) will define key transcriptional regulators driving Kras-mediated reprogramming. Similar analyses and comparisons between *ALTK* and *LTK* cells will determine which aspects of the Kras^{G12D} driven transformation depend upon Aldh1b1 function. Particularly important will be upregulated signaling and metabolic pathways in *LTK* relative to *LT* and *ALTK* samples as these may provide potential therapeutic targets.

To link murine CAC transformation to human pancreatic cancer, the LT and LTK CAC gene signatures will first be projected onto human bulk tumor datasets (eg TCGA-PAAD) using single-sample gene set enrichment analysis (ssGSEA) and digital deconvolution methods (CIBERSORTx) to assess the extent to which CAC-like or transformed CAC-like transcriptional programs are active across diverse healthy and PDAC human samples. Cross-species integration of our murine single-cell datasets with human PDAC scRNA-seq atlases using canonical correlation analysis and ortholog mapping will allow positioning the transformed CACs along early neoplastic progression axes in humans. We will then correlate CAC-derived signatures and clinical variables, tumor subtypes, and patient outcomes. This integrative approach will determine whether CACs constitute a candidate cell-of-origin for human PDAC and will highlight conserved oncogenic pathways relevant to early tumorigenesis and disease progression.

Aim 4.4 Regulatory networks in Kras ^{G12D} driven transformation and their dependence on Aldh1b1 function. To decipher gene regulatory networks implicated in Kras ^{G12D} driven transformation and the implication of Aldh1b1, we will correlate changes in chromatin accessibility and DNA

methylation patterns with changes in gene expression. The integration of ATAC-Seq and EM-Seq data will reveal gene regulatory elements that may be actively controlling gene expression following Kras ^{G12D} activation and which of those might be lost when Aldh1b1 function is not present. These gene regulatory elements will be used for the identification of transcription factor binding sites by interrogating databases of transcription factor binding motifs such as TRANSFAC and HOCOMOCO. Finally, the superimposition of the information from the RNA-Seq on whether genes are regulated (activated or repressed) will allow us to characterize biologically relevant regulatory networks. All putative regulatory regions will be ranked according to the magnitude and direction of change and these changes will be correlated with the differential expression of the associated genes. We will also examine whether this correlation endures following the stratification of putative regulatory regions into proximal and distal elements based on their distance from neighboring TSS. To construct regulatory networks, the identified transcription factor binding motifs will be used to build regulatory networks predicting the regulation of specific genes. Computational tools that implement machine learning algorithms to reliably construct regulatory networks are available [65-67] and our collaborators in the DcGC have the capacity and knowledge to implement them. The validity of the suggested regulatory networks will then be assessed against the actual data obtained from the RNA-seq experiments.

These experiments will identify the immediate Kras ^{G12D}-driven gene regulatory networks and their dependence on the metabolism of the progenitors, thus providing rational therapeutic targets.

Risk Management

Our preliminary results strongly suggest that CACs act as PDAC tumor-initiating cells. for PDAC. All the technical requirements and expertise to assess this (Aims 1.1 and 1.2) are in place. The assessment of Aldh1b1 implication in CACs (Aim 2), by eliminating *Aldh1b1* while inducing Kras* expression using an *Aldh1b1* Cre driver, is technically challenging but all mouse lines are in place and the preliminary results indicate that it is feasible. Mice in Aims 1.1 and 2 will be *Pdx1* heterozygotes but, since Pdx1 is not expressed in CACs [45] and mice will be provided with standard fodder, an interference is highly unlikely. *Pdx1* heterozygosity and glucose intolerance may change the frequency of neoplasias but this is not an objective of our investigations. Furthermore, as an alternative approach, Aim 1.2 employs transplantation of *Pdx1* wild type CACs to assess whether CACs are tumor initiating cells. In aims 3 and 4 we will address the early molecular events following the activation of Kras*, something that has never been addressed before. Now we have all the mouse lines and the techniques in place to understand these events in unprecedented detail. Aims 3 and 4 will also examine the possible implication of Aldh1b1 in these early molecular events. Conceptually, the grant carries a good balance of novel insights to be gained with well-established approaches and mouse models as well as innovative approaches and ideas that are experimentally feasible.

Outlook / Perspective

We identified a rare adult pancreatic cell population (Aldh1b1^{+} CACs) as a strong candidate for the origin of PDAC because these cells are progenitor cells, remain proliferative and already express Kras. Moreover, Aldh1b1 is indispensable for PDAC development in a mouse PDAC model [45]. Preliminary experiments support this hypothesis and the mouse models and approaches we have established will allow us to understand the very early molecular events following Kras* expression and its interaction with metabolic regulators. This is a unique approach that could define the basic attributes of PDAC cells of origin and may suggest new pathways as targets for PDAC therapies. We do not include validation of these pathways in human samples in this grant because it would have been both overly ambitious and speculative at this point. However, clinicians in the National Center for Tumor Diseases (NCT) where one of the target diseases is pancreatic cancer, will take advantage of these findings. Future work in our lab will include the

analysis of the secretome and the exosomes produced from these cells in search of early biomarkers.

Time Schedule

MONTHS		1 - 6	7 - 12	13-18	19-24	25-30	31-36
AIM 1	1.1						
	1.2						
AIM 2							
AIM 3	3.1						
	3.2						
AIM 4	4.1						
	4.2						
	4.3						
	4.4						

2.4 Handling of research data

As per DFG guidelines, we adhere to the FAIR (Findable, Accessible, Interoperable, Reusable) data concept. A systematic collection, archiving and publication of research data will protect them and increase their traceability and transparency. We will use the infrastructure of the ZIH (Centre for Information Services and High-Performance Computing) at TU Dresden for all data and standard operating procedure (SOP) storage and backup for access, management and sharing remotely. Open Access publication of the results in peer-reviewed journals and deposition of data sets to public databases such as the Gene Expression Omnibus (GEO) will follow.

2.5 Relevance of sex, gender and/or diversity

The relevance of sex and gender is an important issue in medicine. PDAC is slightly more often in men and this is attributed mostly to higher risk factors, such as smoking, diabetes, obesity and alcohol, in men [68]. A role for sex hormones has been postulated but a recent association study did not validate this hypothesis [69]. Here, we explore the early function of Kras* in progenitor cells in otherwise healthy mice under standard conditions. Thus, we don't expect a sex-related variability but results will also be analyzed by sex to directly assess this. Transcriptome analyses would reveal differential regulation of sex hormone pathways following induction of Kras* expression in CACs. If there are such indications, we will expand the number of samples to determine any statistical significance but, to keep experimental and animal costs low, we will not pursue this from the outset. The gender of researchers is not relevant to the research project. There are no persons under study or individuals who might be affected by the implementation of the research results. Diversity of researchers in any aspect is not relevant to the study.

3 Project- and subject-related list of publications

1 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).

2 Park, W., Chawla, A. & O'Reilly, E. M. Pancreatic Cancer: A Review. *JAMA* 326, 851-862, doi:10.1001/jama.2021.13027 (2021).

3 Siegel, R. L., Miller, K. D., Fuchs, H. E. & Jemal, A. Cancer statistics, 2022. *CA Cancer J Clin* 72, 7-33, doi:10.3322/caac.21708 (2022).

4 Kanda, M. *et al.* Presence of somatic mutations in most early-stage pancreatic intraepithelial neoplasia. *Gastroenterology* 142, 730-733 e739, doi:10.1053/j.gastro.2011.12.042 (2012).

5 Collins, M. A. *et al.* Oncogenic Kras is required for both the initiation and maintenance of pancreatic cancer in mice. *The Journal of clinical investigation* 122, 639-653, doi:10.1172/JCI59227 (2012).

6 Ying, H. *et al.* Oncogenic Kras maintains pancreatic tumors through regulation of anabolic glucose metabolism. *Cell* 149, 656-670, doi:10.1016/j.cell.2012.01.058 (2012).

7 Aguirre, A. J., Stanger, B. Z. & Maitra, A. Hope on the Horizon: A Revolution in KRAS Inhibition Is Creating a New Treatment Paradigm for Patients with Pancreatic Cancer. *Cancer research* 84, 2950-2953, doi:10.1158/0008-5472.CAN-24-1926 (2024).

8 Duan, X. *et al.* A pancreatic cancer organoid platform identifies an inhibitor specific to mutant KRAS. *Cell stem cell* 31, 71-88 e78, doi:10.1016/j.stem.2023.11.011 (2024).

9 Son, J. *et al.* Glutamine supports pancreatic cancer growth through a KRAS-regulated metabolic pathway. *Nature* 496, 101-105, doi:10.1038/nature12040 (2013).

10 Padanad, M. S. *et al.* Fatty Acid Oxidation Mediated by Acyl-CoA Synthetase Long Chain 3 Is Required for Mutant KRAS Lung Tumorigenesis. *Cell reports* 16, 1614-1628, doi:10.1016/j.celrep.2016.07.009 (2016).
11 Bailey, P. *et al.* Genomic analyses identify molecular subtypes of pancreatic cancer. *Nature* 531, 47-52, doi:10.1038/nature16965 (2016).
12 Collisson, E. A. *et al.* Subtypes of pancreatic ductal adenocarcinoma and their differing responses to therapy. *Nature medicine* 17, 500-503, doi:10.1038/nm.2344 (2011).
13 Moffitt, R. A. *et al.* Virtual microdissection identifies distinct tumor- and stroma-specific subtypes of pancreatic ductal adenocarcinoma. *Nature genetics* 47, 1168-1178, doi:10.1038/ng.3398 (2015).
14 Cancer Genome Atlas Research Network. Electronic address, a. a. d. h. e. & Cancer Genome Atlas Research, N. Integrated Genomic Characterization of Pancreatic Ductal Adenocarcinoma. *Cancer cell* 32, 185-203 e113, doi:10.1016/j.ccell.2017.07.007 (2017).
15 Chan-Seng-Yue, M. *et al.* Transcription phenotypes of pancreatic cancer are driven by genomic events during tumor evolution. *Nature genetics* 52, 231-240, doi:10.1038/s41588-019-0566-9 (2020).
16 Flowers, B. M. *et al.* Cell of Origin Influences Pancreatic Cancer Subtype. *Cancer Discov* 11, 660-677, doi:10.1158/2159-8290.CD-20-0633 (2021).
17 Ferreira, R. M. M. *et al.* Duct- and Acinar-Derived Pancreatic Ductal Adenocarcinomas Show Distinct Tumor Progression and Marker Expression. *Cell reports* 21, 966-978, doi:10.1016/j.celrep.2017.09.093 (2017).
18 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).
19 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).
20 Bardeesy, N. *et al.* Smad4 is dispensable for normal pancreas development yet critical in progression and tumor biology of pancreas cancer. *Genes & development* 20, 3130-3146, doi:10.1101/gad.1478706 (2006).
21 Aguirre, A. J. *et al.* Activated Kras and Ink4a/Arf deficiency cooperate to produce metastatic pancreatic ductal adenocarcinoma. *Genes & development* 17, 3112-3126, doi:10.1101/gad.1158703 (2003).
22 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).
23 Storz, P. Acinar cell plasticity and development of pancreatic ductal adenocarcinoma. *Nat Rev Gastroenterol Hepatol* 14, 296-304, doi:10.1038/nrgastro.2017.12 (2017).
24 Stellas, D. *et al.* Therapeutic effects of an anti-Myc drug on mouse pancreatic cancer. *J Natl Cancer Inst* 106, doi:10.1093/jnci/dju320 (2014).
25 Ji, B. *et al.* Ras activity levels control the development of pancreatic diseases. *Gastroenterology* 137, 1072-1082, 1082 e1071-1076, doi:10.1053/j.gastro.2009.05.052 (2009).
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 Habbe, N. *et al.* Spontaneous induction of murine pancreatic intraepithelial neoplasia (mPanIN) by acinar cell targeting of oncogenic Kras in adult mice. *Proceedings of the National Academy of Sciences of the United States of America* 105, 18913-18918, doi:10.1073/pnas.0810097105 (2008).
28 Kopp, J. L. *et al.* Identification of Sox9-dependent acinar-to-ductal reprogramming as the principal mechanism for initiation of pancreatic ductal adenocarcinoma. *Cancer cell* 22, 737-750, doi:10.1016/j.ccr.2012.10.025 (2012).
29 Brembeck, F. H. *et al.* The mutant K-ras oncogene causes pancreatic periductal lymphocytic infiltration and gastric mucous neck cell hyperplasia in transgenic mice. *Cancer research* 63, 2005-2009 (2003).
30 Ray, K. C. *et al.* Epithelial tissues have varying degrees of susceptibility to Kras(G12D)-initiated tumorigenesis in a mouse model. *PloS one* 6, e16786, doi:10.1371/journal.pone.0016786 (2011).
31 von Figura, G., Morris, J. P. t., Wright, C. V. & Hebrok, M. Nr5a2 maintains acinar cell differentiation and constrains oncogenic Kras-mediated pancreatic neoplastic initiation. *Gut* 63, 656-664, doi:10.1136/gutjnl-2012-304287 (2014).
32 Fernandez, A. *et al.* A Single-Cell Atlas of the Murine Pancreatic Ductal Tree Identifies Novel Cell Populations With Potential Implications in Pancreas Regeneration and Exocrine Pathogenesis. *Gastroenterology* 167, 944-960 e915, doi:10.1053/j.gastro.2024.06.008 (2024).
33 Muraro, M. J. *et al.* A Single-Cell Transcriptome Atlas of the Human Pancreas. *Cell Syst* 3, 385-394 e383, doi:10.1016/j.cels.2016.09.002 (2016).
34 Tosti, L. *et al.* Single-Nucleus and In Situ RNA-Sequencing Reveal Cell Topographies in the Human Pancreas. *Gastroenterology* 160, 1330-1344 e1311, doi:10.1053/j.gastro.2020.11.010 (2021).
35 Wollny, D. *et al.* Single-Cell Analysis Uncovers Clonal Acinar Cell Heterogeneity in the Adult Pancreas. *Developmental cell* 39, 289-301, doi:10.1016/j.devcel.2016.10.002 (2016).
36 Neuhofer, P. *et al.* Acinar cell clonal expansion in pancreas homeostasis and carcinogenesis. *Nature* 597, 715-719, doi:10.1038/s41586-021-03916-2 (2021).
37 Cleveland, M. H., Sawyer, J. M., Afelik, S., Jensen, J. & Leach, S. D. Exocrine ontogenies: on the development of pancreatic acinar, ductal and centroacinar cells. *Seminars in cell & developmental biology* 23, 711-719, doi:10.1016/j.semcdb.2012.06.008 (2012).
38 Beer, R. L., Parsons, M. J. & Rovira, M. Centroacinar cells: At the center of pancreas regeneration. *Developmental biology* 413, 8-15, doi:10.1016/j.ydbio.2016.02.027 (2016).
39 Miyamoto, Y. *et al.* Notch mediates TGF alpha-induced changes in epithelial differentiation during pancreatic tumorigenesis. *Cancer cell* 3, 565-576, doi:10.1016/s1535-6108(03)00140-5 (2003).
40 Stanger, B. Z. *et al.* Pten constrains centroacinar cell expansion and malignant transformation in the pancreas. *Cancer cell* 8, 185-195, doi:10.1016/j.ccr.2005.07.015 (2005).
41 Spears, E., Serafimidis, I., Powers, A. C. & Gavalas, A. Debates in Pancreatic Beta Cell Biology: Proliferation Versus Progenitor Differentiation and Transdifferentiation in Restoring beta Cell Mass. *Front Endocrinol (Lausanne)* 12, 722250, doi:10.3389/fendo.2021.722250 (2021).

42 Esposito, I. *et al.* Hypothetical progression model of pancreatic cancer with origin in the centroacinar-acinar compartment. *Pancreas* 35, 212-217, doi:10.1097/mpa.0b013e31805d0190 (2007).
43 Rovira, M. *et al.* Isolation and characterization of centroacinar/terminal ductal progenitor cells in adult mouse pancreas. *Proceedings of the National Academy of Sciences of the United States of America* 107, 75-80, doi:10.1073/pnas.0912589107 (2010).
44 Ioannou, M. *et al.* ALDH1B1 is a potential stem/progenitor marker for multiple pancreas progenitor pools. *Developmental biology* 374, 153-163, doi:10.1016/j.ydbio.2012.10.030 (2013).
45 Mameishvili, E. *et al.* Aldh1b1 expression defines progenitor cells in the adult pancreas and is required for Kras-induced pancreatic cancer. *Proceedings of the National Academy of Sciences of the United States of America* 116, 20679-20688, doi:10.1073/pnas.1901075116 (2019).
46 Karimova, M. *et al.* A single reporter mouse line for Vika, Flp, Dre, and Cre-recombination. *Scientific reports* 8, 14453, doi:10.1038/s41598-018-32802-7 (2018).
47 Wu, J. *et al.* Generation of a pancreatic cancer model using a Pdx1-Flp recombinase knock-in allele. *PloS one* 12, e0184984, doi:10.1371/journal.pone.0184984 (2017).
48 Schonhuber, N. *et al.* A next-generation dual-recombinase system for time- and host-specific targeting of pancreatic cancer. *Nature medicine* 20, 1340-1347, doi:10.1038/nm.3646 (2014).
49 Azzoni, A. R., Ribeiro, S. C., Monteiro, G. A. & Prazeres, D. M. The impact of polyadenylation signals on plasmid nuclease-resistance and transgene expression. *J Gene Med* 9, 392-402, doi:10.1002/jgm.1031 (2007).
50 Kim, M. P. *et al.* ALDH activity selectively defines an enhanced tumor-initiating cell population relative to CD133 expression in human pancreatic adenocarcinoma. *PloS one* 6, e20636, doi:10.1371/journal.pone.0020636 (2011).
51 Serafimidis, I., Rakatzi, I., Episkopou, V., Gouti, M. & Gavalas, A. Novel effectors of directed and Ngn3-mediated differentiation of mouse embryonic stem cells into endocrine pancreas progenitors. *Stem Cells* 26, 3-16, doi:10.1634/stemcells.2007-0194 (2008).
52 Anastasiou, V. *et al.* Aldehyde dehydrogenase activity is necessary for beta cell development and functionality in mice. *Diabetologia* 59, 139-150, doi:10.1007/s00125-015-3784-4 (2016).
53 Giannios, I. *et al.* Protein Methyltransferase Inhibition Decreases Endocrine Specification Through the Upregulation of Aldh1b1 Expression. *Stem Cells* 37, 640-651, doi:10.1002/stem.2979 (2019).
54 Serafimidis, I., Heximer, S., Beis, D. & Gavalas, A. G protein-coupled receptor signaling and sphingosine-1-phosphate play a phylogenetically conserved role in endocrine pancreas morphogenesis. *Molecular and cellular biology* 31, 4442-4453, doi:10.1128/MCB.05702-11 (2011).
55 Serafimidis, I. *et al.* Pancreas lineage allocation and specification are regulated by sphingosine-1-phosphate signalling. *PLoS Biol* 15, e2000949, doi:10.1371/journal.pbio.2000949 (2017).
56 Jarc, L. *et al.* Regulation of multiple signaling pathways promotes the consistent expansion of human pancreatic progenitors in defined conditions. *Elife* 12, doi:10.7554/eLife.89962 (2024).
57 Zanfrini, E. *et al.* Generation and application of novel hES cell reporter lines for the differentiation and maturation of hPS cell-derived islet-like clusters. *Scientific reports* doi: 10.1038/s41598-024-69645-4 (2024).
58 Alves, T. C. *et al.* Integrated, Step-Wise, Mass-Isotopomeric Flux Analysis of the TCA Cycle. *Cell metabolism* 22, 936-947, doi:10.1016/j.cmet.2015.08.021 (2015).
59 Ghosh-Choudhary, S., Liu, J. & Finkel, T. Metabolic Regulation of Cell Fate and Function. *Trends in cell biology* 30, 201-212, doi:10.1016/j.tcb.2019.12.005 (2020).
60 Ito, K. & Suda, T. Metabolic requirements for the maintenance of self-renewing stem cells. *Nature reviews. Molecular cell biology* 15, 243-256, doi:10.1038/nrm3772 (2014).
61 Yun, J. *et al.* Glucose deprivation contributes to the development of KRAS pathway mutations in tumor cells. *Science* 325, 1555-1559, doi:10.1126/science.1174229 (2009).
62 DeNicola, G. M. *et al.* Oncogene-induced Nrf2 transcription promotes ROS detoxification and tumorigenesis. *Nature* 475, 106-109, doi:10.1038/nature10189 (2011).
63 Ditzer, D. *et al.* Epigenome profiling identifies H3K27me3 regulation of extra-cellular matrix composition in human corticogenesis. *BioRxiv*, doi:10.1101/2024.10.01.616076 (2024).
64 Feng, S., Zhong, Z., Wang, M. & Jacobsen, S. E. Efficient and accurate determination of genome-wide DNA methylation patterns in Arabidopsis thaliana with enzymatic methyl sequencing. *Epigenetics Chromatin* 13, 42, doi:10.1186/s13072-020-00361-9 (2020).
65 Pineda, S., Van Steen, K. & Malats, N. Integrative eQTL analysis of tumor and host omics data in individuals with bladder cancer. *Genet Epidemiol* 41, 567-573, doi:10.1002/gepi.22053 (2017).
66 Omranian, N., Eloundou-Mbebi, J. M., Mueller-Roeber, B. & Nikoloski, Z. Gene regulatory network inference using fused LASSO on multiple data sets. *Scientific reports* 6, 20533, doi:10.1038/srep20533 (2016).
67 Guan, D. *et al.* CMGRN: a web server for constructing multilevel gene regulatory networks using ChIP-seq and gene expression data. *Bioinformatics* 30, 1190-1192, doi:10.1093/bioinformatics/btt761 (2014).
68 Collaborators, G. B. D. P. C. The global, regional, and national burden of pancreatic cancer and its attributable risk factors in 195 countries and territories, 1990-2017: a systematic analysis for the Global Burden of Disease Study 2017. *Lancet Gastroenterol Hepatol* 4, 934-947, doi:10.1016/S2468-1253(19)30347-4 (2019).
69 Peduzzi, G. *et al.* Common variability in oestrogen-related genes and pancreatic ductal adenocarcinoma risk in women. *Scientific reports* 12, 18100, doi:10.1038/s41598-022-22973-9 (2022).