Please cite this article in press as: Chondronasiou et al., Deciphering the roadmap of *in vivo* reprogramming toward pluripotency, Stem Cell Reports (2022), https://doi.org/10.1016/j.stemcr.2022.09.009

Stem Cell Reports
Article
ISSCR
OPEN ACCESS

Deciphering the roadmap of *in vivo* reprogramming toward pluripotency

Dafni Chondronasiou,[1] Jaime Martínez de Villarreal,[2,3] Elena Melendez,[1] Cian J. Lynch,[1] Natalia del Pozo,[2] Marta Kovatcheva,[1] Mònica Aguilera,[1] Neus Prats,[1] Francisco X. Real,[2,3,4] and Manuel Serrano[1,5,*]
[1]Institute for Research in Biomedicine (IRB Barcelona), Barcelona Institute of Science and Technology (BIST), Barcelona, Spain
[2]Epithelial Carcinogenesis Group, Molecular Oncology Programme, Spanish National Cancer Research Centre (CNIO), Madrid, Spain
[3]CIBERONC, Madrid, Spain
[4]Departament de Medicina i Ciències de la Vida, Universitat Pompeu Fabra, Barcelona, Spain
[5]Catalan Institution for Research and Advanced Studies (ICREA), 08010 Barcelona, Spain
*Correspondence: manuel.serrano@irbbarcelona.org
https://doi.org/10.1016/j.stemcr.2022.09.009

SUMMARY

Differentiated cells can be converted into pluripotent stem cells by expressing the transcription factors OCT4, SOX2, KLF4, and MYC (OSKM) in a process known as reprogramming. Here, using single-cell RNA sequencing of pancreas undergoing reprogramming, we identify markers along the trajectory from acinar cell identity to pluripotency. These markers allow direct *in situ* visualization of cells undergoing dedifferentiation and acquiring features of early and advanced intermediate reprogramming. We also find that a fraction of cells do not dedifferentiate upon OSKM expression and are characterized by stress markers of the REG3 and AP-1 families. Importantly, most markers of intermediate reprogramming in the pancreas are also observed in stomach, colon, and cultured fibroblasts expressing OSKM. Among them is LY6A, a protein characteristic of progenitor cells and generally upregulated during tissue repair. Our roadmap defines intermediate reprogramming states that could be functionally relevant for tissue regeneration and rejuvenation.

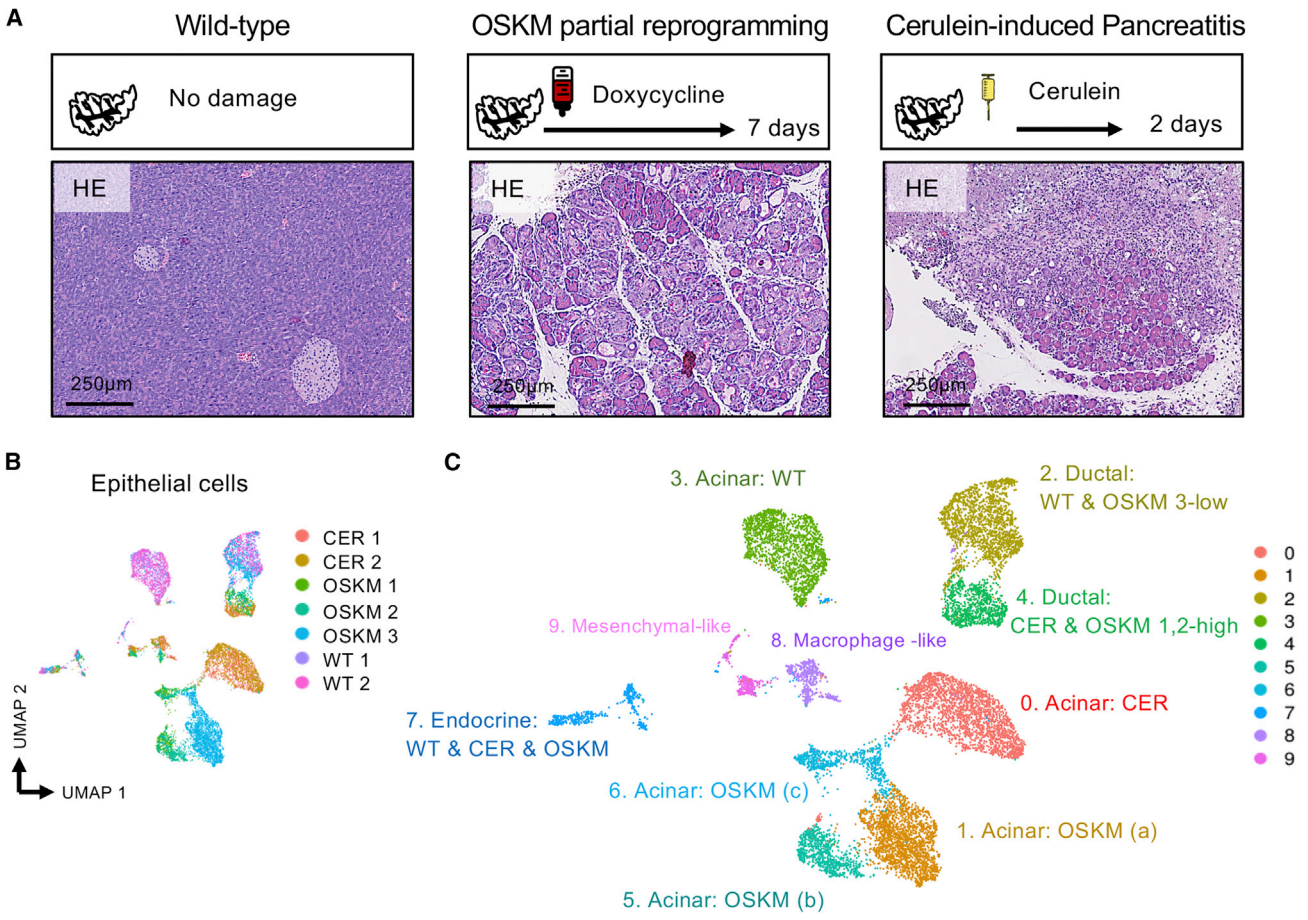
INTRODUCTION

The ability to manipulate cell fate *in vitro* has revolutionized regenerative medicine. The most striking breakthrough in the field occurred when Yamanaka first demonstrated the ability of adult differentiated cells to give rise to pluripotent cells (induced pluripotent stem cells [iPSCs]) upon the simultaneous expression of four transcriptional factors, OCT4, SOX2, KLF4, and MYC (OSKM) (Takahashi and Yamanaka, 2006). During reprogramming, a fraction of adult somatic cells shut down the transcriptional programs linked to their cell identity and progressively activate the transcriptional network of pluripotency (Deng et al., 2021; Takahashi and Yamanaka, 2006). In mice, OSKM expression recapitulates the events of cellular dedifferentiation in multiple tissues and leads to the emergence of pluripotent cells (Abad et al., 2013; Mosteiro et al., 2016; Ohnishi et al., 2014). The ultimate manifestation of complete reprogramming *in vivo* is the formation of teratomas, a tumor originated from iPSCs differentiating into all three germ layers (Abad et al., 2013; Ohnishi et al., 2014).

Cellular reprogramming is an inefficient process due, at least in part, to the existence of multiple cell-autonomous barriers, such as tumor suppressors, chromatin regulators, transcription factors, signaling pathways, and microRNAs (Arabacı et al., 2021; Haridhasapavalan et al., 2020). Various studies have tried to untangle the complex cascade of molecular and epigenetic events occurring during *in vitro* reprogramming as well as to define intermediate states (Brambrink et al., 2008; Chronis et al., 2017; O'Malley et al., 2013; Polo et al., 2012; Stadtfeld et al., 2008; Zviran et al., 2019). The majority of cells do not successfully complete reprogramming and have varying cell fates: some undergo apoptosis, mainly triggered by MYC induction (Kim et al., 2018), while others undergo senescence as a result of the activation of tumor-suppressor pathways driven by p53, and INK4A/ARF (Banito et al., 2009; Hong et al., 2009; Kawamura et al., 2009; Li et al., 2009; Marión et al., 2009; Utikal et al., 2009). Of note, senescent cells, despite failing to reprogram, indirectly promote the reprogramming of neighbor cells by secreting interleukin-6 (IL-6) which acts as a pro-reprogramming factor (Brady et al., 2013; Chiche et al., 2017; Mosteiro et al., 2016). Those cells that successfully dedifferentiate upon OSKM expression enter into a heterogenous phase, characterized by its high plasticity and generally known as intermediate reprogramming. Cells with placental-, neuronal-, or epidermal-like identities have been described as reprogramming intermediates *in vitro* (Kurian et al., 2013; O'Malley et al., 2013; Schiebinger et al., 2017). Trophectoderm stem cells (TSCs) and extraembryonic endoderm (XEN) cells are other alternative fates during OSKM reprogramming (Deng et al., 2021; Liu et al., 2020; Parenti et al., 2016; Schiebinger et al., 2019; Xing et al., 2020). Finally, intermediate-reprogrammed fibroblasts can generate mesodermal progenitors, which can subsequently differentiate into endothelium and smooth muscle (Kurian et al., 2013).

Despite the detailed knowledge about the intermediate states of reprogramming under *in vitro* culture conditions, little is known about the molecular roadmap of *in vivo* reprogramming. Here, we use single-cell transcriptomics to decipher the process of reprogramming *in vivo*. Focusing on pancreas, as the tissue with the highest reprogramming

Please cite this article in press as: Chondronasiou et al., Deciphering the roadmap of *in vivo* reprogramming toward pluripotency, Stem Cell Reports (2022), https://doi.org/10.1016/j.stemcr.2022.09.009



(legend on next page)

efficiency in our mouse model (Abad et al., 2013; Mosteiro et al., 2016), we identify markers of intermediate reprogramming and visualize intermediate-reprogrammed cells within tissues by using immunohistochemistry and RNA fluorescence *in situ* hybridization. Moreover, we demonstrate that other tissues undergoing reprogramming and *in vitro* reprogrammed mouse embryonic fibroblasts (MEFs) share many of the markers identified in the pancreas. To our knowledge, this is the first time that an OSKM-driven roadmap has been generated *in vivo* in the context of an adult organism.

RESULTS

scRNA-seq captures partially reprogrammed populations in the pancreas

To capture the cellular heterogeneity generated during partial reprogramming in the pancreas, we performed single-cell RNA sequencing (scRNA-seq). Three reprogrammable (OSKM) and two control (wild-type [WT]) mice were treated with doxycycline (1 mg/mL) for 7 days (Figure 1A). Upon examination of the pancreas, we could appreciate that the extent of reprogramming was high in mice OSKM 1 and 2 and modest in mouse OSKM 3. For comparison, we treated another group of mice with cerulein (CER) for 2 consecutive days (Figure 1A). CER produces severe acinar damage leading to acinar cell plasticity and acinar-to-ductal metaplasia (ADM). At the end of the treatment, each pancreas was dissociated into single-cell suspension, sorted for live ($DAPI^-$) cells, and sequenced using the 10x Genomics system. From each individual pancreas, we sequenced between 6,657 and 14,240 cells, amounting to a total of 52,717 cells, which we analyzed with the R package Seurat. All samples were merged without integration (Figures S1A, S1D, and S2A), clustered, and annotated using the scCATCH tool together with manual annotation based on the expression of defined markers (see experimental procedures; Figures S1B, S1C, S1E, S1F, S2B, and S2C; Table S1). We observed that in all three conditions, even in the context of reprogramming, the main pancreatic cell types could still be identified, which is consistent with the fact that reprogramming and CER-induced ADM produce focal transformations, not affecting the entire pancreas. However, cell-type composition was dramatically altered after OSKM activation or CER damage, as was clearly depicted in the uniform manifold approximation and projection (UMAP) by the abundant infiltration of inflammatory cells and activated fibroblasts likely corresponding to pancreatic stellate cells (Figures S1 and S2). The changes in the cellular composition of the tissues largely reflect their histological changes (Figure 1A). Of note, an analysis of the inflammatory cells present during reprogramming can be found elsewhere (Melendez et al., 2022). Here, we have focused on the epithelial compartment of the pancreas, as it undergoes the most dramatic histological changes during reprogramming (Figure 1A).

To capture the differences among WT, CER-treated, and OSKM-induced pancreas at the epithelial level, we selected acinar, ductal, and endocrine cells coming from all samples, and we reanalyzed them (Figure 1B). Acinar gene expression was detected in 5 independent clusters (WT: cluster 3; CER: cluster 0, OSKM: clusters 1, 5, and 6) (Figures 1C and 1D; Table S2). On the other hand, ductal cells coming from $OSKM^{high}$ pancreas clustered together with CER ductal cells (cluster 4), but not with ductal cells from control mice, suggesting that CER and OSKM affect ductal cells in similar ways. Finally, endocrine cells clustered together (cluster 7) regardless of their origin (Figures 1C and 1D; Table S2). Overall, the epithelial analysis of the scRNA-seq data suggests that acinar cells are the most affected ones in the pancreas after OSKM activation in our mouse model. Moreover, OSKM expression produces distinct transcriptional changes in

Figure 1. Single-cell RNA sequencing reveals the susceptibility of pancreatic acinar cells to *in vivo* OSKM reprogramming

(A) Representative images of hematoxylin and eosin (H&E) staining of pancreas from wild-type (WT), cerulein (CER)-treated (total of 14 intraperitoneal [i.p.] injections in 2 consecutive days) for inducing acute pancreatitis, and OSKM-induced mice (1 mg/mL doxycycline in the drinking water for 7 days) for inducing partial reprogramming.

(B) All epithelial cells (acinar, ductal, endocrine) coming from all samples (two WT, two CER, and three OSKM mice) were selected, merged, and reanalyzed as represented in this uniform manifold approximation and projection (UMAP). The contribution of each sample is depicted in this UMAP by a different color.

(C) Merged cells were clustered using a cluster resolution parameter of 0.2, and each cluster was annotated according to the expression of characteristic markers.

(D) Established marker genes identify acinar cells (*Pnliprp2*), ductal cells (*Spp1*), endocrine cells (*Ppy*), macrophage-like (*Sparc*) cells, and mesenchymal-like (*Sparc*) cells.

(E) All acinar-related clusters coming from WT and OSKM-induced mice (clusters 1,3,5, and 6) were selected, merged, and reanalyzed as represented in this UMAP. The contribution of each sample is depicted in this UMAP by a different color.

(F) Trajectory analysis of the acinar cells coming from WT and OSKM-induced mice using the Slingshot tool (Street et al., 2018).

(G) UMAP visualization of the acinar marker *Pnliprp2* in acinar cells corresponding to WT and OSKM pancreata.

See also Figures S1 and S2.

the acinar compartment compared with CER-induced acute pancreatitis.

To gain further insight into the acinar-associated transcriptional states generated by OSKM activation, we extracted only acinar cells coming from OSKM and WT pancreas (Figure 1E). After reclustering, we identified 5 acinar clusters, one for WT and four for OSKM (Figure 1F and Table S3), and we performed cell lineage and pseudotime trajectory inference analysis using Slingshot (Street et al., 2018), pre-setting WT acinar cells as the starting point. In this way, we identified three main trajectories (or pseudotimes) emerging from a common OSKM cluster (Figures 1F and S2D). We will refer to this precursor cluster as pre-reprogrammed (pre-REP) (Figure 1F). This pre-REP cluster is the hub for the emergence of three different trajectories. Two of them retain acinar identity, and therefore we considered them as "non-reprogrammed" (non-REP) fates (Figure 1F). We noted that one of these clusters was characterized by an abundant presence of the transcription factor AP-1, so we refer to this cluster as non-REP/AP-1 (Figure 1F). Interestingly, a third trajectory emerging from pre-REP gives rise to a cluster of cells that have largely lost their acinar identity, as judged by the strong reduction of acinar markers (Table S3), such as *Pnliprp2* (Figure 1G), and we refer to them as intermediate-reprogrammed (int-REP) cells (Figure 1F).

In vivo identification of pre-REP and non-REP OSKM fates

Upon examination of the pre-REP and non-REP clusters (Tables S2 and S3), we focused on *Reg3a* and related family members because they were absent from normal acinar cells and highly expressed in both clusters (Figures 2A and S3A). The "regenerating" (Reg) family of C-type lectins is known to be upregulated in pancreatitis (Chen et al., 2019), and, indeed, we observed upregulation of several *Reg3* family members in CER-treated pancreas, but the level of upregulation was much lower compared with OSKM-expressing pancreas (Figures 2B and S3A). We visualized the presence of REG3A/G$^+$ cells in pancreas undergoing reprogramming by immunohistochemistry using an antibody that recognizes these two members of the REG3 family (Figure 2C). Interestingly, REG3A/G cells were rare in the dysplastic areas with abnormal histology but abundant in the regions of the exocrine pancreas with apparently normal histology (Figure 2C). This suggests that the pre-REP and non-REP cells retain a normal histology while undergoing transcriptional changes.

The non-REP cluster included two markers, leucine-rich alpha-2-glycoprotein 1 (*Lrg1*) and lipocalin 2 (*Lcn2*), which were absent in the pre-REP and non-REP/AP-1 clusters (Figures 2A and S3B; Tables S2 and S3). Interestingly, qRT-PCR of bulk pancreas RNA revealed that *Lrg1* was dramatically upregulated in pancreas upon OSKM expression but not upon CER treatment (Figure 2B). RNA fluorescence *in situ* hybridization (RNA-FISH) indicated that *Lrg1* expression was restricted to pancreatic areas that maintained their acinar identity (as defined by carboxypeptidase A1, *Cpa1*) in the context of OSKM-expressing pancreas (Figure 2D). Finally, we wondered if non-REP cells were actively expressing the transgenic OSKM cassette. By RNA-FISH, we observed that the expression of transgenic OSKM after 7 days of doxycycline was largely restricted to the regions that had lost acinar identity (*Cpa1* negative) (Figure S3C). Therefore, after 7 days of *in vivo* reprogramming, transgenic expression of OSKM is restricted to cells undergoing reprogramming that co-exist with non-REP cells.

In vivo identification of int-REP

We then focused on the int-REP cluster. This cluster consists of a heterogeneous population of cells expressing genes that are not normally expressed in the pancreas. This cluster also contains a few cells expressing the pluripotency marker *Oct4* (Figure S4), reinforcing the idea that this cluster represents cells undergoing reprogramming. After examining the genes upregulated in the int-REP cluster (from the acinar UMAP in Figure 1F and from the epithelial UMAP cluster 6 in Figure 1C), we selected 23 genes whose expression patterns were restricted to the int-REP cluster (Figure S4; Tables S2 and S3). We evaluated their expression by RT-qPCR in bulk pancreas, and we validated that almost all of them were selectively induced upon OSKM reprogramming but not in CER-treated pancreas (Figure 3A). Based on their expression pattern (Figure S4), we could divide them into three categories that reflect their sublocalization along the reprogramming trajectory: "early," "advanced," or both "early and advanced" (Figures 3A and S4). We noticed that *Muc5ac*, which is normally expressed in the gastric and respiratory tracts, was induced early upon OSKM activation (Figure 3B). RNA-FISH illustrates the expression of *Muc5ac* both in a subset of cells that retain acinar identity (*Cpa1* positive) and in a subset of cells that have lost acinar identity (*Cpa1* negative) (Figure 3C).

Among the early-and-advanced markers of int-REP cells, we noticed *Ly6a* (*Sca-1*) (Figure 4A), a member of the LY6 family associated with stem/progenitor cells in various tissues including pancreas (Holmes and Stanford, 2007; Leinenkugel et al., 2022) and previously reported to be upregulated during *in vitro* reprogramming of MEFs (Schwarz et al., 2018). LY6A immunohistochemistry clearly defined the dysplastic regions of reprogrammed pancreas, and the amount of LY6A$^+$ cells was positively correlated with the extension of tissue dysplasia (Figures 4B, 4C, and S5A). Immunohistochemistry did not reveal LY6A$^+$ cells in normal pancreas or in CER-treated pancreas (Figure 4B). Similarly, *Aldh3a1* was also upregulated in the entire int-REP cluster (early and advanced) (Figure 4D), and

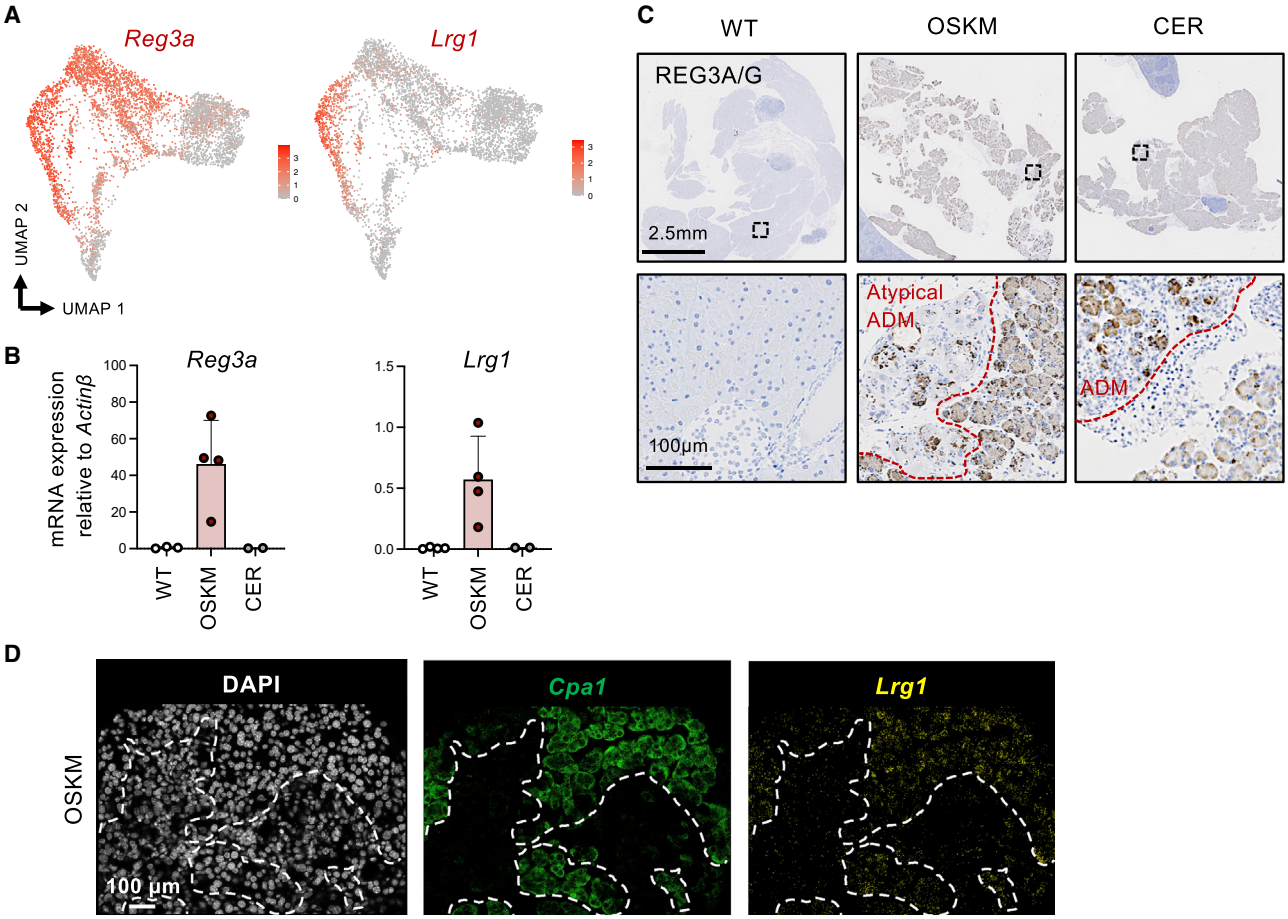


Figure 2. *In vivo* identification of pre-reprogramming and non-reprogramming OSKM fates

(A) UMAP representation of *Reg3a* and *Lrg1* expression in acinar cells coming from WT and OSKM-induced pancreata.

(B) mRNA expression of *Reg3a* and *Lrg1* in bulk pancreas comparing all three conditions. Each dot corresponds to a different mouse. Bars indicate the mean and error bars represent the standard deviation. The values in OSKM mice were significantly different from WT ($p < 0.05$, Student's t-test).

(C) Immunohistochemistry of REG3A/G in the pancreas of WT, CER-treated, and OSKM-induced mice.

(D) RNA fluorescence *in situ* hybridization (RNA-FISH) of *Cpa1* (acinar marker) and *Lrg1* in reprogrammed pancreas. Scale bar, 100 μm. See also Figure S3.

RNA-FISH revealed that it was largely expressed in cells that had lost acinar identity (*Cpa1* negative), although some cells co-expressed *Aldh3a1* and *Cpa1* (Figure 4E). *Aldh3a1* is a member of the aldehyde dehydrogenase (ALDH) superfamily that is upregulated in response to oxidative stress, acting as an antioxidant (Vasiliou and Nebert, 2005) and participating in cell-cycle regulation (Estey et al., 2007). Additionally, *Cd9* was another marker of this cluster (Figure S5B), which has been previously identified as a marker of pancreatic cancer stem cells (Wang et al., 2019). Of note, we observed CD9 to be localized in the nucleus of OSKM dysplastic pancreatic cells (Figure S5B). In this regard, a nuclear CD9 pool has previously been reported in the context of human breast carcinoma, where it participates in mitotic processes (Rappa et al., 2014).

Finally, we analyzed advanced int-REP genes. We noticed that this group of genes is enriched in keratins (Figure 3A), including those characteristic of stratified epithelia, such as *Krt14* (Figure 4F). Regions of dysplastic pancreas were clearly stained with KRT14, whereas WT and CER pancreas were completely negative (Figure 4G). *Krt14* is a keratin that it is not normally expressed in pancreas except in the context of rare squamous pancreatic cancers (Bailey et al., 2016; Real et al., 1993). Notably, the extent of OSKM-induced dysplasia was positively correlated with KRT14 expression in the pancreas (Figure 4H). Another gene in the advanced int-REP category is *Ly6g6c*, which also belongs to the LY6 family, although it is located in a different gene cluster than *Ly6a* (Upadhyay, 2019) (Figure 4I). Using RNA-FISH, we could map the *Ly6g6c*-expressing

Please cite this article in press as: Chondronasiou et al., Deciphering the roadmap of *in vivo* reprogramming toward pluripotency, Stem Cell Reports (2022), https://doi.org/10.1016/j.stemcr.2022.09.009

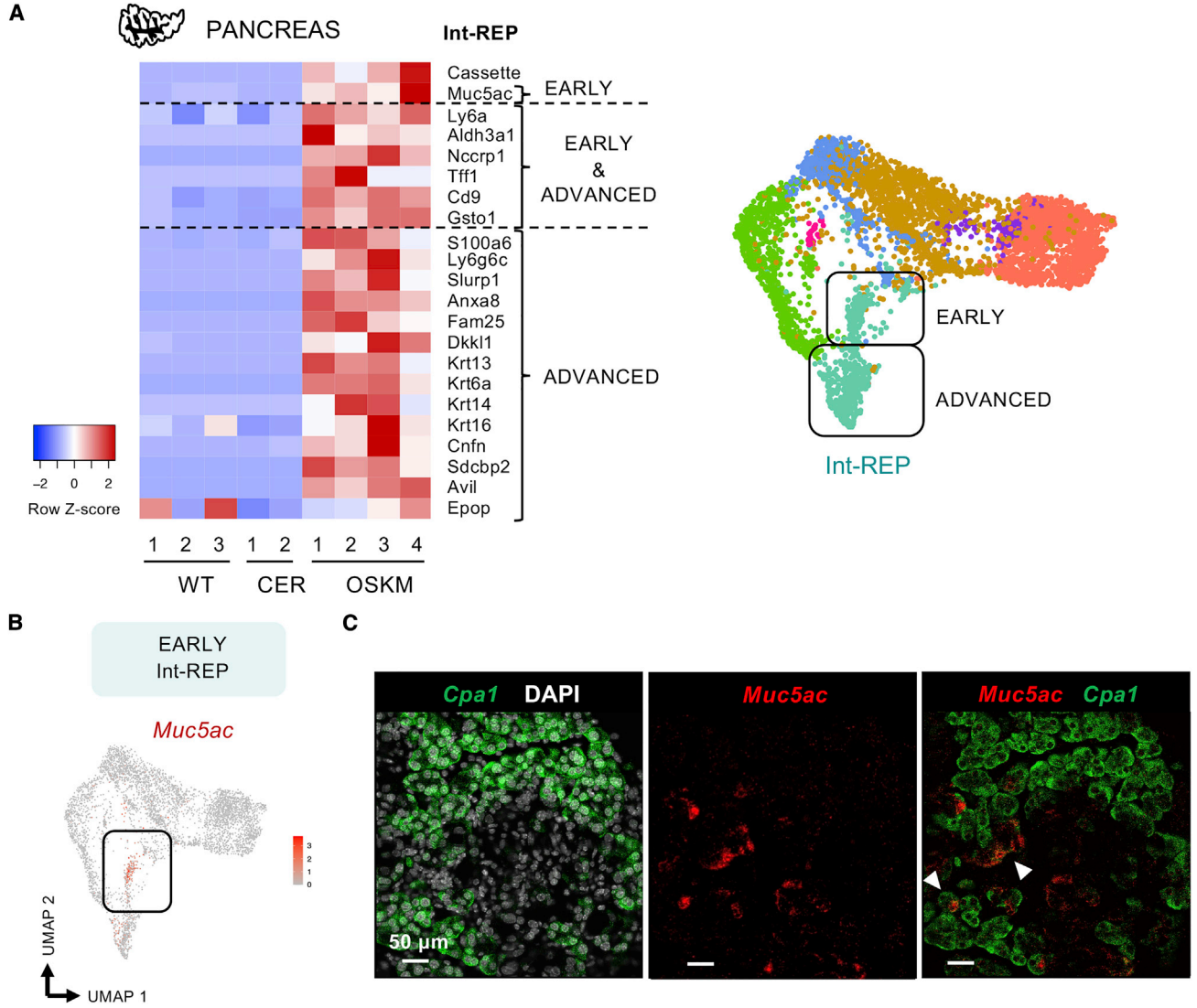


Figure 3. *In vivo* identification of intermediate reprogramming

(A) Heatmap representation of mRNA expression measured by qRT-PCR of selected genes characterizing cells undergoing partial reprogramming in pancreas from WT (n = 3), CER-treated (n = 2), and OSKM-induced (n = 4) mice. Blue represents low expression and red high expression of those genes. Depending on their expression pattern, genes have been divided into three groups: early, advanced, or early and advanced. Genes *Sprr2a3* and *Sprr2b* could not be detected.

(B) UMAP representation of the expression of the "early" marker *Muc5ac* in acinar cells coming from WT and OSKM-induced pancreata.

(C) Images of RNA-FISH of partially reprogrammed pancreas using probes against *Cpa1* and *Muc5ac* together with DAPI. Scale bars, 50 μm.

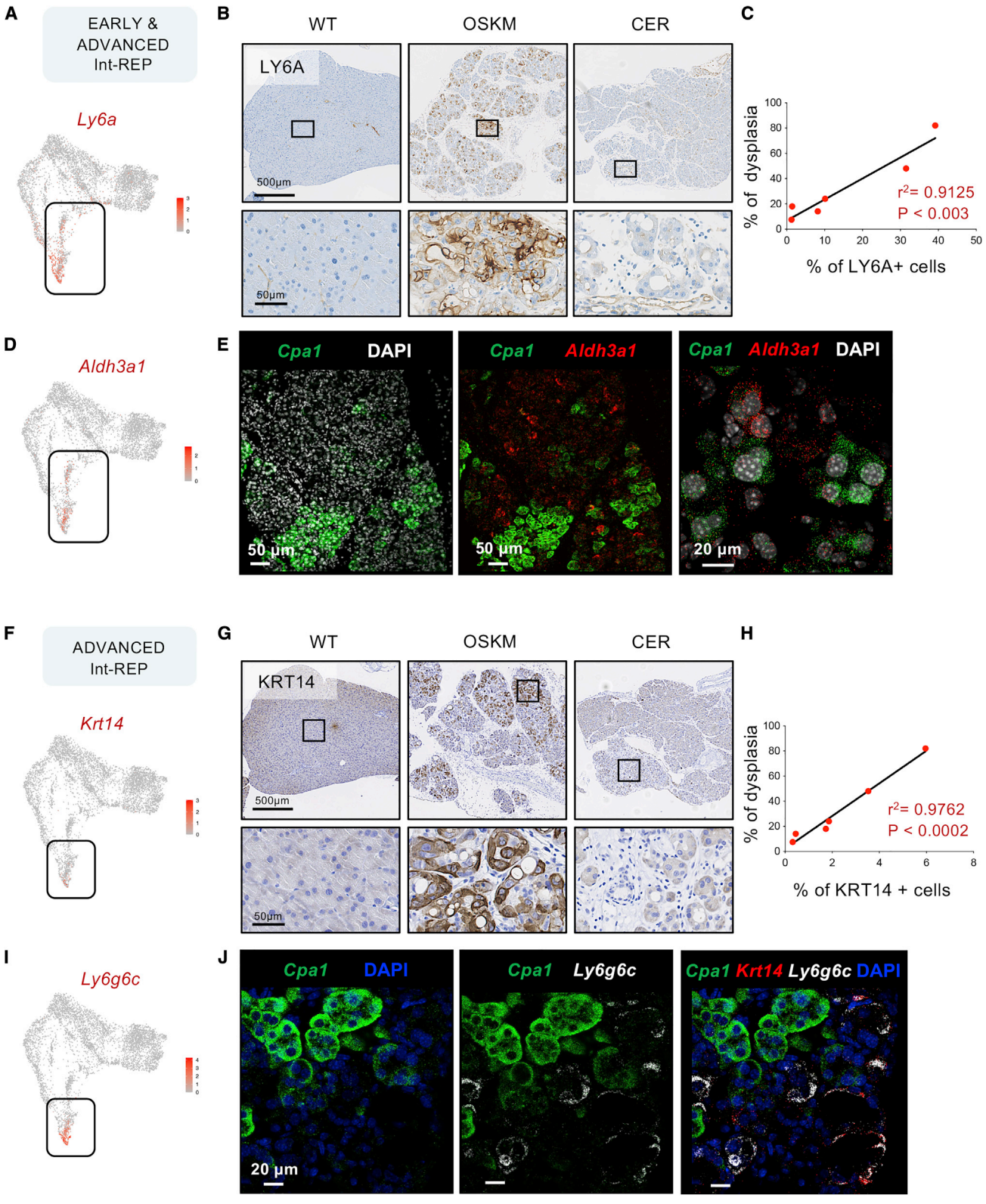
See also Figure S4.

cells into the dedifferentiated areas where acinar cells had completely lost their identity (*Cpa1* negative) (Figure 4J). The partial co-localization of *Krt14* with *Ly6g6c* suggests that they are part of the same transcriptional trajectory (Figure 4J). Other interesting genes in the advanced-intermediate reprogramming were *Slurp1* and *Sdcbp*. SLURP1 is another member of the LY6 family located in the same cluster as LY6A. SLURP1 is a secreted protein that positively regulates cholinergic receptors and stabilizes epithelial cell-cell junctions (Campbell et al., 2019; Vasilyeva et al., 2017).

A roadmap of *in vivo* reprogramming in the pancreas

To further refine and support the existence of early and advanced states of int-REP, we performed additional RNA-FISH staining combining markers representative of

Please cite this article in press as: Chondronasiou et al., Deciphering the roadmap of *in vivo* reprogramming toward pluripotency, Stem Cell Reports (2022), https://doi.org/10.1016/j.stemcr.2022.09.009



(legend on next page)

different states. For example, we observed that *Muc5ac* (early) and *Ly6g6c* (advanced) did not co-localize (Figure 5A). On the other hand, *Krt14* (advanced) partially co-localized with *Aldh3a1* (early and advanced) (Figure 5A). The dynamic nature of the intermediate states was even more clearly illustrated when we simultaneously stained for *Muc5ac* (early), *Ly6a* and *Aldh3a1* (early and advanced), and *Ly6g6c* (advanced) (Figure 5A). To get more quantitative information about the co-existence of markers, we plotted the frequency of cells co-expressing the above markers in our scRNA-seq data (Figure S6A). Consistent with our classification of early and advanced markers, the most frequent combinations involved *Muc5ac*/*Ly6a* and *Ly6a*/*Ly6g6c* doble positive cells, while the *Muc5ac*/*Ly6g6c* combination was very infrequent. As we observed frequent co-expression of markers, we performed a cluster-independent trajectory analysis of WT and OSKM acinar cells using the Monocle 2 tool (Qiu et al., 2017). Interestingly, this analysis provided a trajectory that was essentially linear. The localization of the above-discussed markers was again consistent with the classification of early and advanced markers. For example, *Ly6g6c* is concentrated at the very end of the trajectory, whereas *Ly6a* is more spread and *Muc5ac* is essentially absent from the advanced reprogramming region (Figures 5B, S6B, and S6C).

Finally, we asked if the process of oncogenic neoplasia in the exocrine pancreas could share similarities with int-REP. For this, we extracted from previously reported scRNA-seq data a signature formed by the top 50 genes defining oncogenic *Kras*-driven acinar metaplastic cells (Schlesinger et al., 2020), and we mapped this signature in our acinar UMAP. Interestingly, the advanced int-REP cells were enriched in the oncogene-induced metaplastic signature (Figure 5C). Of note, this signature does not include LY6A and KRT14, and indeed, when we examined oncogenic *Kras*-induced pancreatic intraepithelial neoplasias (PanINs) by immunohistochemistry we could not detect these markers (Figure S6D). We conclude that PanINs and int-REP are distinct processes sharing transcriptional features. On the other hand, we also extracted a signature of the top 50 genes defining CER-induced metaplastic cells (Table S4), and it was enriched in the non-REP cells but not in the int-REP cells (Figure 5D).

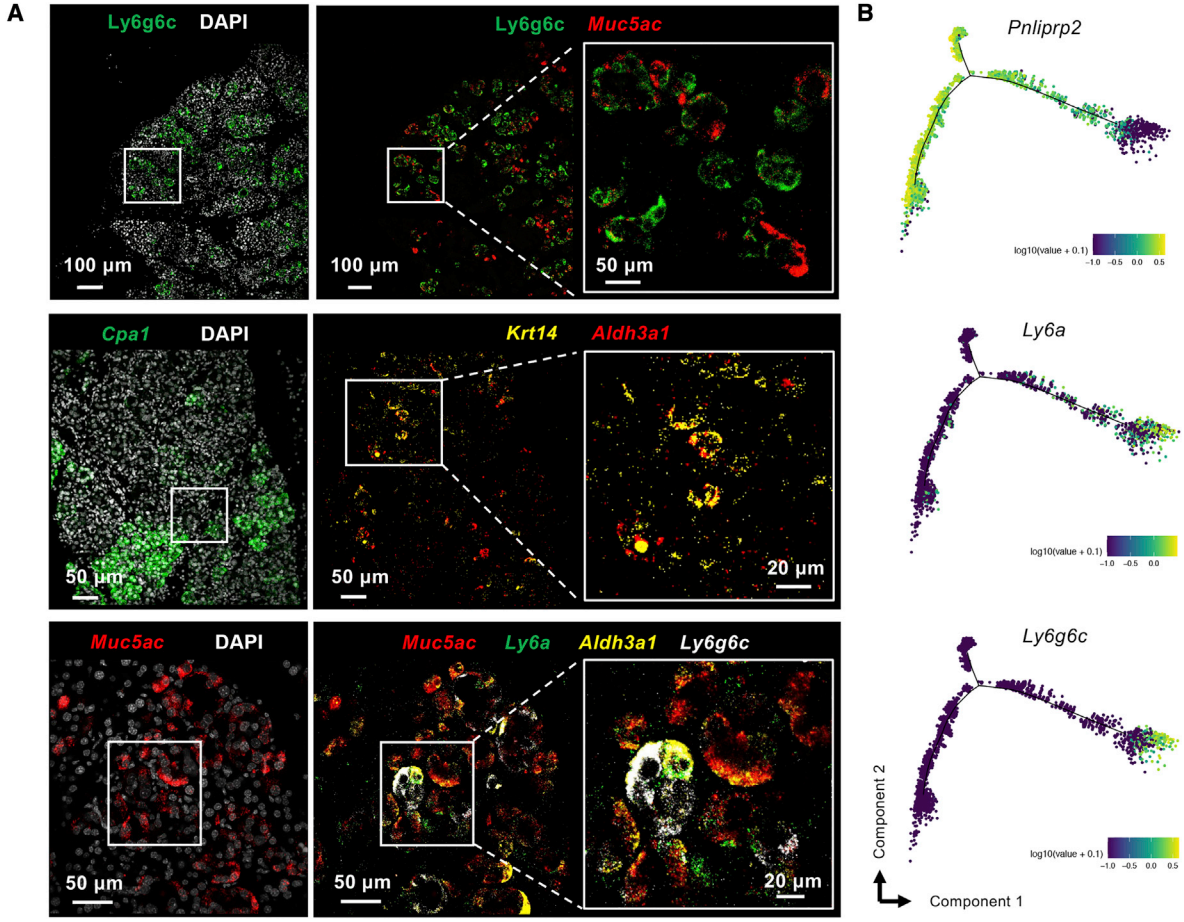
Collectively, our data support a first roadmap for *in vivo* reprogramming in the pancreas that is supported by scRNA-seq data, qRT-PCR in bulk tissues, and RNA-FISH and immunohistochemistry in histological sections (Figure 5E). We also conclude that the process of intermediate reprogramming is distinct from oncogenic-induced or CER-induced metaplasia, although it shares transcriptional features with oncogenic metaplasia.

Markers of int-REP are shared among several tissues

Having identified a signature of int-REP markers in the exocrine pancreas (comprising 23 genes) (Figure S4), we wondered if it would also be upregulated during reprogramming of other tissues. For this, we evaluated the expression of these markers in the stomach and colon of OSKM-induced mice, and, interestingly, many genes of this signature were also upregulated by qRT-PCR in reprogrammed stomach and colon compared with WT samples (Figure 6A). By immunohistochemistry, dysplastic epithelial cells in reprogramming glandular stomach and colon were clearly stained by LY6A and KRT14 after 7 days of treatment with doxycycline, while normal WT tissues did not express these markers (Figure 6B). Of note, it has been previously reported that in the context of ulcerative colitis, the intestinal epithelium undergoing regeneration acquires a fetal-like expression profile marked by LY6A expression (Yui et al., 2018). Another interesting observation is that the non-dysplastic epithelial regions of the reprogramming intestine were positive for REG3A/G, similar to the non-reprogramming areas of the pancreas upon OSKM expression (Figure 6C). In the case of the glandular stomach, reprogramming also induced REG3A/G+ cells, although their number was much lower than in pancreas or colon. We conclude that the process of reprogramming shares similar features among different tissues *in vivo*.

Figure 4. *In vivo* identification of early and advanced markers of intermediate reprogramming

(A) UMAP representation of *Ly6a* expression in acinar cells coming from WT and OSKM-induced pancreata.
(B) Immunohistochemistry of LY6A in paraffin-embedded sections of pancreas from WT, CER-treated, and OSKM-induced mice.
(C) The extent of OSKM-induced dysplasia is highly correlated with the presence of $LY6A^+$ cells in the reprogrammed pancreas.
(D) UMAP representation of *Aldh3a1* expression in acinar cells coming from WT and OSKM-induced pancreata.
(E) Images of RNA-FISH of partially reprogrammed pancreas using probes against *Cpa1* and *Aldh3a1* together with DAPI.
(F) UMAP representation of *Krt14* expression in acinar cells coming from WT and OSKM-induced pancreata.
(G) Immunohistochemistry of KRT14 in paraffin-embedded sections of pancreas from WT, CER-treated, and OSKM-induced mice.
(H) The extent of OSKM-induced dysplasia is highly correlated with the presence of $KRT14^+$ cells in the reprogrammed pancreas.
(I) UMAP representation of *Ly6g6c* expression in acinar cells coming from WT and OSKM-induced pancreata. See also Figure S5.
(J) Images of RNA-FISH of partially reprogrammed pancreas using probes against *Cpa1*, *Ly6g6c*, and *Krt14* together with DAPI. Scale bars, 500, 50, and 20 μm, as indicated on each image separately.
See also Figure S5.



(legend on next page)

In vivo and *in vitro* reprogramming share markers of int-REP

Having identified markers of intermediate OSKM reprogramming in several tissues, we wondered if these markers would also be upregulated during *in vitro* reprogramming. We examined the expression of the int-REP signature (comprising 23 genes) by qRT-PCR in primary MEFs undergoing *in vitro* reprogramming. Remarkably, all int-REP markers were upregulated during *in vitro* reprogramming of MEFs, and the majority were absent in the final state of iPSCs (Figure 7A). To gain insight into the regulation of the int-REP markers, we extracted from a previous report (Chronis et al., 2017) the assay for transposase-accessible chromatin (ATAC)-seq profiles of these genes during *in vitro* reprogramming. All of them presented low or undetectable ATAC-seq peaks in the parental fibroblast cells, in agreement with their lack of expression. Interestingly, all the loci coding for these markers presented prominent ATAC-seq peaks in the intermediate state of reprogramming (pre-iPSC), suggesting an accessible chromatin compatible with their expression (Figures 7B and S7A). At the pluripotent stage, the majority of the open chromatin peaks associated to the intermediate markers were no longer detectable. These observations link chromatin accessibility with the expression of int-REP markers.

To further characterize the int-REP states, we wondered about their capacity to progress into pluripotency. We focused on markers LY6A (early and advanced) and KRT14 (advanced). Previous investigators have found that LY6A$^+$ *in vitro* intermediates purified by cell sorting are less efficient in generating iPSC colonies compared with LY6A intermediates (Schwarz et al., 2018). We confirmed these observations using antibody-linked magnetic beads to isolate LY6A$^+$ intermediates (Figure S7B). We also used fibroblasts carrying a lineage tracing allele that, upon tamoxifen administration, converts *Krt14*-expressing cells from Tomato$^+$ into GFP$^+$ cells. By the end of the reprogramming process in the presence of tamoxifen, all the iPSC colonies were Tomato$^+$, indicating that *Krt14*$^+$ intermediates do not produce iPSCs (Figure S7C). Reprogramming intermediates with low efficiency of conversion into pluripotent cells can be of interest in future applications in regeneration because of their reduced risk of teratoma formation.

We conclude that the intermediate reprogramming stages are characterized by a common set of markers that can be useful for future applications in regeneration or rejuvenation.

DISCUSSION

In this work, we report the first roadmap of *in vivo* partial reprogramming. scRNA-seq of reprogrammed pancreas suggests that pancreatic cell types are not equally affected by OSKM activation, with acinar cells being the most susceptible ones to reprogramming. This probably reflects the known plasticity of acinar cells, which can rapidly dedifferentiate upon inflammatory or oncogenic damages (van Roey et al., 2021). Also, expression of OSKM in acinar cells results in a general shut down of enhancers and dedifferentiation (Shibata et al., 2018). As a comparison, we included a model of acute pancreatitis induced by CER because it represents another process of acinar plasticity known as acinar-to-ductal metaplasia (ADM). Our scRNA-seq analysis indicates that OSKM-reprogrammed acinar cells undergo changes that are distinct from CER-induced ADM. Regarding the rest of the pancreatic cell types, endocrine cells appear to be resistant to OSKM reprogramming, while ductal cells acquire a reactive molecular profile similar to ductal cells in CER-treated pancreas. It is possible that each cell type requires different periods of OSKM activation to surrender their somatic identity, probably reflecting the potency of epigenetic barriers (Arabacı et al., 2021).

Focusing on acinar cells, we found that a fraction of acinar cells retain their somatic identity in the context

Figure 5. A roadmap of *in vivo* reprogramming in the pancreas

(A) RNA-FISH visualizes the expression of multiple intermediate-reprogramming markers in OSKM-induced pancreas using probes against *Cpa1*, *Ly6g6c*, *Aldh3a1*, *Krt14*, and *Ly6a* together with DAPI. Section bars, 100 and 50 μm, as indicated on each image separately.

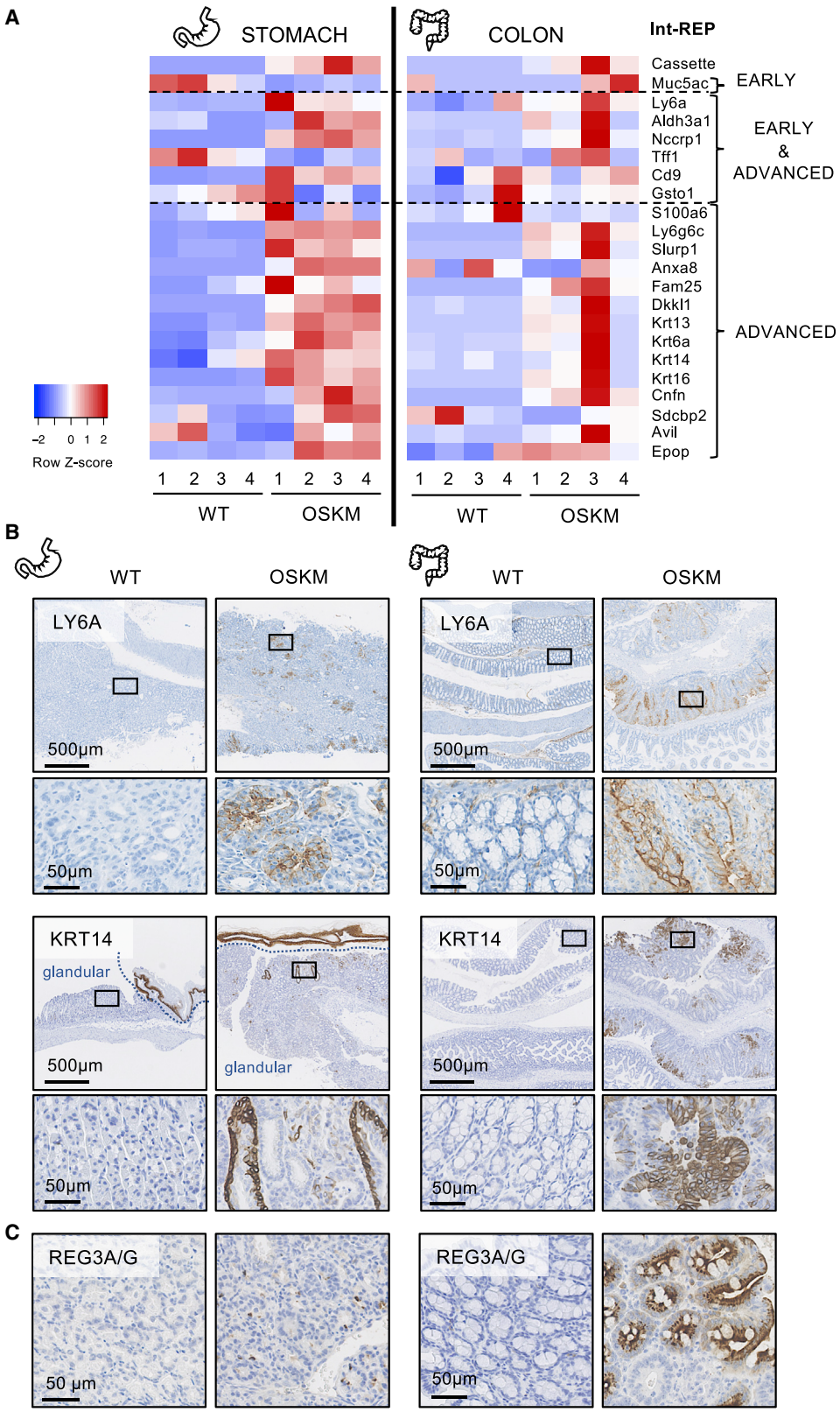
(B) Trajectory analysis of the acinar cells coming from WT and OSKM-induced mice using the Monocle 2 tool (Qiu et al., 2017) while visualizing the gene expression of *Pnliprp2*, *Ly6a*, and *Ly6g6c*.

(C) An oncogene-induced metaplastic signature (top 50 differentially upregulated genes versus normal acinar compartment ranked by log fold change [FC] from scRNA-seq data reported in Schlesinger et al., 2020) plotted on our acinar UMAP.

(D) A CER signature (top 50 differentially upregulated genes versus normal acinar compartment ranked by logFC from our epithelial UMAP (Figure 1C) plotted on our acinar UMAP.

(E) Our working model illustrates the molecular roadmap induced by OSKM activation in pancreas. Briefly, upon OSKM activation *in vivo*, whole acini react by upregulating *Reg3a* (pre-reprogramming state). Subsequently, while a fraction of acinar cells are refractory to reprogramming and retain their identity (*Reg3a*$^+$ *Lrg1*$^+$: non-reprogramming state), another fraction loses their acinar identity and acquires distinct gene expression profiles along the process of dedifferentiation (intermediate reprogramming). Ultimately, few cells successfully make it to the pluripotent state, marked by the expression of *Nanog* and endogenous *Oct4*.

See also Figure S6.



(legend on next page)

of OSKM-reprogrammed tissue, while simultaneously acquiring a molecular profile that we have called pre-REP and non-REP, with the latter one divided in two clusters: non-REP and non-REP/AP-1. The pre-REP cluster, according to the trajectory analysis, is a hub that originates three different fates, namely, non-REP, non-REP/AP-1, and int-REP, with the latter one giving rise to cells with pluripotency features. This situation is reminiscent of *in vitro* reprogramming in the sense that a substantial fraction of cells resist dedifferentiation. In contrast to normal acinar cells in non-reprogrammable mice, the three clusters that retain acinar identity during OSKM expression (pre-REP, non-REP, and non-REP/AP-1) express high levels of the *Reg3* family, which are generally upregulated upon injury or stress in the pancreas (Chen et al., 2019). Regarding the non-REP/AP-1 cluster, it is worth mentioning that the AP-1 family of transcription factors plays a critical role during the early stages of *in vitro* reprogramming, and successful progression toward pluripotency requires their subsequent downregulation (Chronis et al., 2017). Indeed, consistent with our observations, constitutive high expression of AP-1 factors, such as FRA1 or JUN, blocks progression of reprogramming (Chronis et al., 2017). In the case of the non-REP cluster, we identified two markers that were exclusive of this cluster, *Lrg1* and *Lcn2*, and absent in the pre-REP and the non-REP/AP-1 clusters. LRG1 has been reported as a biomarker for several malignancies, including pancreatic cancer (Fukamachi et al., 2019; Xie et al., 2019), LCN2 is a biomarker of tissue injury and metabolic disorders (Jaberi et al., 2021).

Senescent cells have been described to be generated during *in vivo* reprogramming and to play a key role in this process by providing secreted factors that promote reprogramming, such as IL-6 (Mosteiro et al., 2016). However, we were not able to detect senescent cells (*p16/Cdkn2a*$^{+}$ cells) in our scRNA-seq analysis. We speculate that senescent cells may be underrepresented due to their large size, which can interfere with the process of cell isolation for 10x Genomics sequencing.

Regarding the fraction of acinar cells that succumb to OSKM-induced dedifferentiation, it consists of a heterogeneous population of cells that we have called int-REP, which are the main focus of this study. This heterogeneous cluster includes very few cells expressing *Oct4*, suggesting that it contains cells on their way to pluripotency. Of note, our analysis has been done after 7 days of exposure to doxycycline, while full reprogramming requires longer periods of OSKM expression (Abad et al., 2013). However, it provides a broad view of the process because of its asynchronous nature. Based on the trajectory of cell identity change, we were able to distinguish early and advanced markers of int-REP, as well as markers that were present at both stages (early and advanced). We briefly discuss below three of these markers of int-REP: MUC5AC (early), LY6A (early and advanced), and KRT14 (advanced).

One of the most prominent markers present along early-and-advanced int-REP is LY6A (also known as SCA1), a member of the LY6 family that is associated with stem/progenitor cells in the hematopoietic system and in solid tissues such as the pancreas (Dzierzak and Bigas, 2018; Holmes and Stanford, 2007; Leinenkugel et al., 2022). It is remarkable that the large majority of dysplastic tissue is LY6A^{+}, and indeed, there is an excellent positive correlation between dysplasia and LY6A positivity. Until now, we could only estimate reprogramming efficiency by evaluating histological alterations. Here, we identify LY6A as an optimal marker for visualizing reprogrammed areas *in vivo*. Of note, *Ly6a* has been previously reported to be rapidly and transiently upregulated during *in vitro* reprogramming of MEFs (Schwarz et al., 2018).

We found intriguing a possible speculative connection between MUC5AC (marker of early int-REP) and KRT14 (advanced int-REP). On one hand, MUC5AC corresponds to a protein that is normally absent in the pancreas, but it is characteristic of other tissues of endodermal origin, such as lung and stomach. On the other hand, KRT14 and several other keratins characteristic of stratified epithelia were found upregulated in advanced intermediates. These keratins are not normally expressed in pancreas, except in the context of rare squamous pancreatic cancers (Bailey et al., 2016; Real et al., 1993). Although highly speculative at present, this could suggest a trajectory of developmental regression, from endoderm-like features to ectoderm-like features.

It is well documented that *in vitro*, partially reprogrammed cells are heterogeneous and can acquire a variety

Figure 6. Markers of intermediate reprogramming are shared among several tissues|

(A) Heatmap representation of mRNA expression by qRT-PCR of the identified markers of the intermediate-reprogramming state in stomach and colon samples of WT (n = 4) and OSKM-induced (n = 4) mice. Samples do not correspond to the same group of mice. Blue represents low expression and red high expression. Genes *Sprr2a3* and *Sprr2b* could not be detected.

(B) Immunohistochemistry of LY6A and KRT14 in intestine and stomach samples of WT and OSKM-treated mice.

(C) Immunohistochemistry of REG3A/G in colon and stomach samples of WT and OSKM-treated mice. Scale bars, 500 and 50 μm, as indicated on each image.

See also Figure S7.

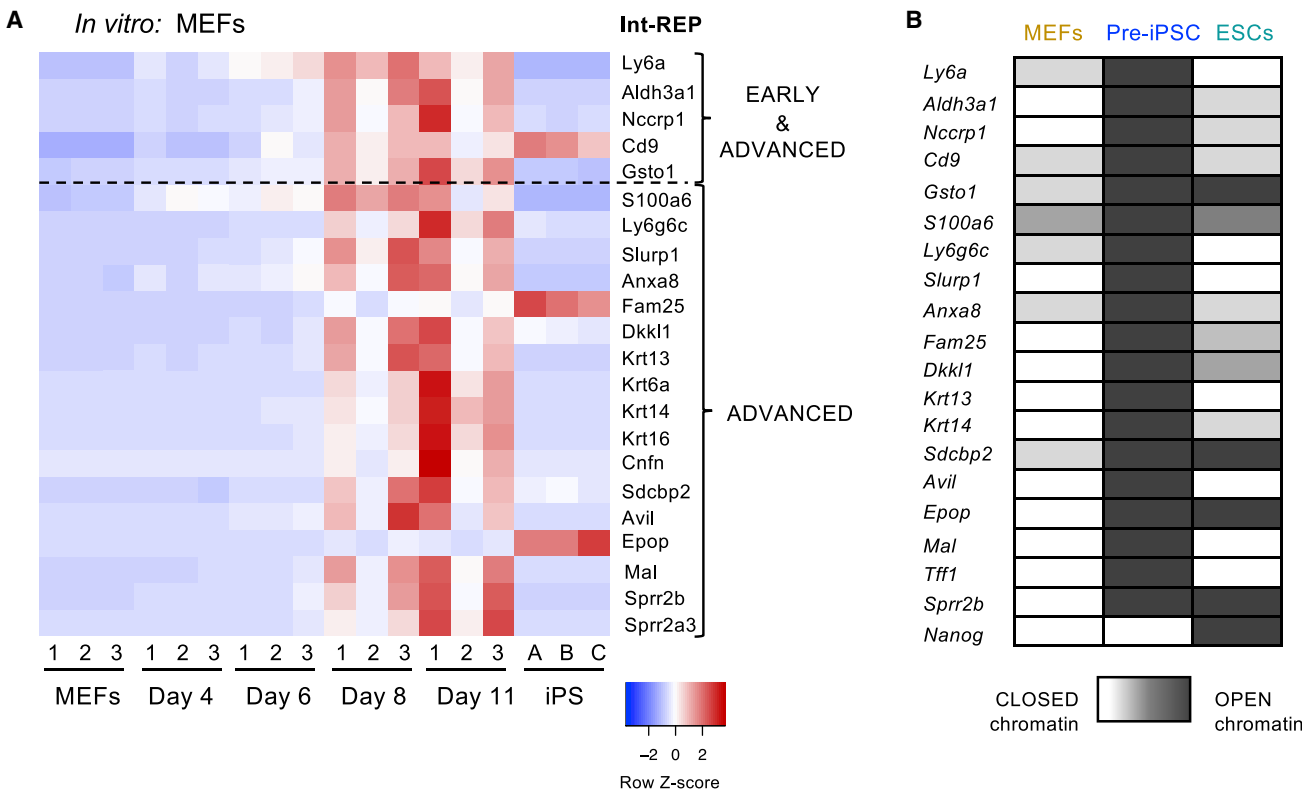


Figure 7. Markers of *in vivo* intermediate reprogramming are also present during *in vitro* reprogramming of MEFs

(A) Heatmap representation of mRNA expression of the identified markers for intermediate-reprogramming state in mouse embryonic fibroblasts (MEFs) undergoing reprogramming *in vitro,* as well as in induced pluripotent stem cells (iPSCs). Blue represents low expression and red high expression. Gene *Muc5ac* was not included.

(B) Semi-quantitively grading of the accessibility state of the chromatin for the genes of intermediate reprogramming (int-REP) signature, as well as *Nanog* (as a marker of pluripotency) using previously published ATAC-seq data (Chronis et al., 2017). The accessibility state of the chromatin is illustrated through a color gradient, with a white color to represent closed chromatin and dark gray to represent open chromatin.

of fates depending on the clues from the extracellular medium (see introduction). In this regard, we wondered if the markers that we have identified in pancreas could be extrapolated to other tissues. Interestingly, most mRNA markers of pancreas int-REP (signature comprising 23 genes) were upregulated in colon and stomach undergoing reprogramming, as well as in reprogrammed fibroblasts (MEFs) *in vitro*. We also visualized partially reprogrammed colon and stomach using immunohistochemistry for LY6A and KRT14. This demonstrates the conservation of the inermediate reprogramming stages among different tissue contexts and even compared to *in vitro* culture conditions. Of note, in agreement with a previous report (Schwarz et al., 2018), we have observed that LY6A^{+} intermediates isolated from *in vitro* reprogramming fibroblasts had a low efficiency of conversion into iPSC colonies. In the case of KRT14, none of the intermediates upregulating this gene produced iPSC colonies. Reprogramming intermediates with low efficiency of conversion into pluripotent cells can be of interest in future applications in regeneration because of their reduced risk of teratoma formation.

Regarding cancer, we show that a signature of oncogene-induced metaplasia is enriched in the advanced int-REP cells. However, some markers of advanced reprogramming, in particular LY6A and KRT14, are absent from the earliest preneoplastic lesions of the pancreas, known as pancreatic intraepithelial neoplasias (PanIN). It has been reported that OSKM expression in the pancreas accelerates oncogenic *Kras*-driven neoplasia (Shibata et al., 2018). Also, some markers of int-REP have been connected with pancreatic cancer. In particular, MUC5AC is abundantly expressed in PanINs (Ganguly et al., 2021). Moreover, CD9 has been reported to be expressed and functionally relevant for pancreatic cancer stem cells (Wang et al., 2019). Finally, a fraction of advanced pancreatic ductal adenocarcinomas express KRT14 (Bailey et al., 2016; Real et al., 1993), and this has been mechanistically connected with the loss of GATA6, a key transcription

factor for acinar identity (Martinelli et al., 2017). Together, these observations support the concept that int-REP and oncogenic metaplasia share some similarities.

Overall, this work identifies and visualizes intermediate states of OSKM reprogramming *in vivo* and provides the first molecular roadmap of this process. Defining the states of *in vivo* reprogramming could be relevant for future applications in tissue regeneration and rejuvenation.

EXPERIMENTAL PROCEDURES

Mouse model and treatments

Animal experimentation was performed at the Spanish National Cancer Research Centre (CNIO) in Madrid and at the Institute for Research in Biomedicine (IRB) in Barcelona, according to protocols approved by the CNIO-ISCIII Ethical Committee for Research and Animal Welfare (CEIyBA) in Madrid, and by the Animal Care and Use Ethical Committee of animal experimentation of Barcelona Science Park (CEEAPCB) and the Catalan Government in Barcelona. We used the reprogrammable mice known as i4F-B, which carry a ubiquitous doxycycline-inducible OSKM transgene, abbreviated as i4F, and we inserted it into the *Pparg* gene (Abad et al., 2013). To activate the four Yamanaka factors, 1 mg/mL doxycycline hyclate BioChemica (PanReac, A2951) was administered in the drinking water supplemented with 7.5% sucrose for a period of 7 days, and mice were sacrificed directly after. Acute pancreatitis was induced as previously described (Carrière et al., 2011). In brief, adult WT mice were intraperitoneally injected with CER (Bachem), an analog of the pancreatic secretagogue cholecystokinin for mice, dissolved in saline solution (100 μg/kg body weight) hourly for 7 h on 2 consecutive days. Mice were sacrificed the day after the last injection.

Tissue processing and scRNA-seq

Mouse primary pancreatic cells were obtained by digesting the whole pancreas with 1 mg/mL Collagenase P (Sigma, 11213865001) supplemented with 2 U/mL Dispase II (Life Technologies, 17105041), 0.1 mg/mL Soybean Trypsin Inhibitor (Life Technologies, 17075-029), and 0.1 mg/mL DNase I (Sigma, D4513) in HBSS with Ca^{2+}/Mg^{2+} (Life Technologies, 14025050). Tissue was dissociated using the gentle MACS Octo Dissociator (Miltenyi Biotech) and further processed, and single-cell suspension of pancreatic cells was resuspended in fluorescence-activated cell sorting (FACS) buffer (10 mM EDTA, 2% FBS in Ca^{2+}/Mg^{2+}-free PBS). $DAPI^-$ cells were selected by cell sorting using FACSAria Fusion (BD Biosciences), and sorted cells were loaded onto a 10x Chromium Single Cell Controller chip B (10× Genomics) as described in the manufacturer's protocol (Chromium Single Cell 3' GEM, Library & Gel Bead Kit v.3, PN-1000075). Libraries were loaded at a concentration of 1.8 pM and sequenced in an asymmetrical pair-end format in a NextSeq500 instrument (Illumina).

RNA-FISH

In situ 4-color RNA-FISH was performed using primer-padlock 2-oligo hybridization of the RNA targets (Table S5), essentially as described (Wang et al., 2018), except that hybridization was performed in 30% formamide buffer (Merck). Briefly, single-stranded DNA (ssDNA) oligo probes were designed against the coding regions of target RNA transcripts (with the exception of the OSKM cassette, where the non-coding regions were deliberately selected), using Picky2.0 to identify loci without RNA secondary structure or repetitive sequence, followed by blast searching to confirm target specificity. Probe hybridization was performed overnight at 40°C with rocking in a humid chamber. Probes then underwent rolling circle amplification, polyacrylamide gel mounting, and proteinase K (Sigma-Aldrich) sample clarification before imaging, and images were acquired in a Zeiss LSM880 microscope, with processing in ImageJ using background subtraction (50 pixels) and a 2-pixel median-filter.

More details for the experimental procedures are included in supplemental information.

Resource availability

Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the corresponding author, Manuel Serrano (manuel.serrano@irbbarcelona.org).

Materials availability

For enquiries about materials and methods, contact with the corresponding author.

DATA AND CODE AVAILABILITY

The scRNA-seq datasets have been deposited in the NCBI Gene Expression Omnibus (GEO) database and are accessible through GEO: GSE188819.

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.stemcr.2022.09.009.

AUTHOR CONTRIBUTIONS

D.C. designed most experiments, performed animal experimentation, collected samples, performed RNA analysis, prepared the figures, and co-wrote the manuscript. J.M.d.V. performed all the bioinformatic analysis related to the scRNA-seq. E.M. contributed to the animal experimentation and collection of samples. C.J.L. performed the RNA-FISH. N.d.P. and M.K. contributed to the animal experimentation and collection of samples. M.A. developed all the histological staining. N.P. supervised and interpreted the histological analyses. F.X.R. supervised the analysis of scRNA-seq analysis and provided his pancreas expertise throughout the study. M.S. designed and supervised the study and co-wrote the manuscript. All authors discussed the results and commented on the manuscript.

ACKNOWLEDGMENTS

We are grateful to Maria Isabel Muñoz for her assistance with the animal protocols. We thank the IRB Core Facilities for their support, the Barcelona Science Park Animal Facility for animal maintenance, and the Spanish National Center for Genomic Analysis

(CNAG) for the scRNA sequencing. D.C. was recipient of a fellowship from the "la Caixa" Foundation. E.M. was funded by a Future Fellowship from the IRB. M.K. is supported by a fellowship from the Spanish Association Against Cancer (AECC). Work in the laboratory of F.X.R. was funded by grants from the Spanish Ministry of Science co-funded by the ERDF-EU (SAF2015-70553-R and RTI2018-101071-B-I00). Work in the laboratory of M.S. was funded by the IRB and by grants from the Spanish Ministry of Science co-funded by the ERDF-EU (SAF2017-82613-R), European Research Council (ERC-2014-AdG/669622), Secretaria d'Universitats i Recerca del Departament d'Empresa i Coneixement of Catalonia (Grup de Recerca consolidat 2017 SGR 282), "la Caixa" Foundation, and the Milky Way Research Foundation. The CNIO and IRB are supported by the Spanish Ministry of Science as Centres of Excellence "Severo Ochoa."

CONFLICT OF INTERESTS

M.S. is shareholder and advisor of Senolytic Therapeutics, Inc., Life Biosciences, Inc, Rejuveron Senescence Therapeutics, AG, and Altos Labs, Inc. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Received: September 7, 2022
Revised: September 20, 2022
Accepted: September 21, 2022
Published: October 20, 2022

REFERENCES

Abad, M., Mosteiro, L., Pantoja, C., Cañamero, M., Rayon, T., Ors, I., Graña, O., Megías, D., Domínguez, O., Martínez, D., et al. (2013). Reprogramming in vivo produces teratomas and iPS cells with totipotency features. Nature *502*, 340–345.

Arabacı, D.H., Terzioğlu, G., Bayırbaşı, B., and Önder, T.T. (2021). Going up the hill: chromatin-based barriers to epigenetic reprogramming. FEBS J. *288*, 4798–4811.

Bailey, P., Chang, D.K., Nones, K., Johns, A.L., Patch, A.M., Gingras, M.C., Miller, D.K., Christ, A.N., Bruxner, T.J.C., Quinn, M.C., et al. (2016). Genomic analyses identify molecular subtypes of pancreatic cancer. Nature *531*, 47–52.

Banito, A., Rashid, S.T., Acosta, J.C., Li, S., Pereira, C.F., Geti, I., Pinho, S., Silva, J.C., Azuara, V., Walsh, M., et al. (2009). Senescence impairs successful reprogramming to pluripotent stem cells. Genes Dev. *23*, 2134–2139.

Brady, J.J., Li, M., Suthram, S., Jiang, H., Wong, W.H., and Blau, H.M. (2013). Early role for IL-6 signalling during generation of induced pluripotent stem cells revealed by heterokaryon RNA-Seq. Nature. Nat. Cell Biol. *15*, 1244–1252.

Brambrink, T., Foreman, R., Welstead, G.G., Lengner, C.J., Wernig, M., Suh, H., and Jaenisch, R. (2008). Sequential expression of pluripotency markers during direct reprogramming of mouse somatic cells. Cell Stem Cell *2*, 151–159.

Campbell, G., Swamynathan, S., Tiwari, A., and Swamynathan, S.K. (2019). The secreted Ly-6/uPAR related protein-1 (SLURP1) stabilizes epithelial cell junctions and suppresses TNF-α-induced cytokine production. Biochem. Biophys. Res. Commun. *517*, 729–734.

Carrière, C., Young, A.L., Gunn, J.R., Longnecker, D.S., and Korc, M. (2011). Acute pancreatitis accelerates initiation and progression to pancreatic cancer in mice expressing oncogenic Kras in the nestin cell lineage. PLoS One *6*, e27725.

Chen, Z., Downing, S., and Tzanakakis, E.S. (2019). Four decades after the discovery of regenerating islet-derived (Reg) proteins: current understanding and challenges. Front. Cell Dev. Biol. *7*, 235.

Chiche, A., Le Roux, I., von Joest, M., Sakai, H., Aguín, S.B., Cazin, C., Salam, R., Fiette, L., Alegria, O., Flamant, P., et al. (2017). Injury-induced senescence enables in vivo reprogramming in skeletal muscle. Cell Stem Cell *20*, 407–414.e4.

Chronis, C., Fiziev, P., Papp, B., Butz, S., Bonora, G., Sabri, S., Ernst, J., and Plath, K. (2017). Cooperative binding of transcription factors orchestrates reprogramming. Cell *168*, 442–459.e20.

Deng, W., Jacobson, E.C., Collier, A.J., and Plath, K. (2021). The transcription factor code in iPSC reprogramming. Curr. Opin. Genet. Dev. *70*, 89–96.

Dzierzak, E., and Bigas, A. (2018). Blood development: hematopoietic stem cell dependence and independence. Cell Stem Cell *22*, 639–651.

Estey, T., Piatigorsky, J., Lassen, N., and Vasiliou, V. (2007). ALDH3A1: a corneal crystallin with diverse functions. Exp. Eye Res. *84*, 3–12.

Fukamachi, K., Hagiwara, Y., Futakuchi, M., Alexander, D.B., Tsuda, H., and Suzui, M. (2019). Evaluation of a biomarker for the diagnosis of pancreas cancer using an animal model. J. Toxicol. Pathol. *32*, 135–141.

Ganguly, K., Krishn, S.R., Rachagani, S., Jahan, R., Shah, A., Nallasamy, P., Rauth, S., Atri, P., Cox, J.L., Pothuraju, R., et al. (2021). Secretory mucin 5AC promotes neoplastic progression by augmenting KLF4-mediated pancreatic cancer cell stemness. Cancer Res. *81*, 91–102.

Haridhasapavalan, K.K., Raina, K., Dey, C., Adhikari, P., and Thummer, R.P. (2020). An insight into reprogramming barriers to iPSC generation. Stem Cell Rev. Rep. *16*, 56–81.

Holmes, C., and Stanford, W.L. (2007). Concise review: stem cell antigen-1: expression, function, and enigma. Stem Cell. *25*, 1339–1347.

Hong, H., Takahashi, K., Ichisaka, T., Aoi, T., Kanagawa, O., Nakagawa, M., Okita, K., and Yamanaka, S. (2009). Suppression of induced pluripotent stem cell generation by the p53-p21 pathway. Nature *460*, 1132–1135.

Jaberi, S.A., Cohen, A., D'Souza, C., Abdulrazzaq, Y.M., Ojha, S., Bastaki, S., and Adeghate, E.A. (2021). Lipocalin-2: structure, function, distribution and role in metabolic disorders. Biomed. Pharmacother. *142*, 112002.

Kawamura, T., Suzuki, J., Wang, Y.V., Menendez, S., Morera, L.B., Raya, A., Wahl, G.M., and Izpisúa Belmonte, J.C. (2009). Linking the p53 tumour suppressor pathway to somatic cell reprogramming. Nature *460*, 1140–1144.

Kim, E.J.Y., Anko, M.-L., Flensberg, C., Majewski, I.J., Geng, F.-S., Firas, J., Huang, D.C.S., van Delft, M.F., and Heath, J.K. (2018).

BAK/BAX-Mediated apoptosis is a myc-induced roadblock to reprogramming. Stem Cell Rep. *10*, 331–338.

Kurian, L., Sancho-Martinez, I., Nivet, E., Aguirre, A., Moon, K., Pendaries, C., Volle-Challier, C., Bono, F., Herbert, J.-M., Pulecio, J., et al. (2013). Conversion of human fibroblasts to angioblast-like progenitor cells. Nat. Methods *10*, 77–83.

Leinenkugel, G., Kong, B., Raulefs, S., Miller, K., Roth, S., Jiang, H., Istvánffy, R., Heikenwälder, H., Maeritz, N., Regel, I., et al. (2022). Sca-1 is a marker for cell plasticity in murine pancreatic epithelial cells and induced by IFN-β in vitro. Pancreatology *22*, 294–303.

Li, H., Collado, M., Villasante, A., Strati, K., Ortega, S., Cãamero, M., Blasco, M.A., and Serrano, M. (2009). The Ink4/Arf locus is a barrier for iPS cell reprogramming. Nature *460*, 1136–1139.

Liu, X., Ouyang, J.F., Rossello, F.J., Tan, J.P., Davidson, K.C., Valdes, D.S., Schröder, J., Sun, Y.B.Y., Chen, J., Knaupp, A.S., et al. (2020). Reprogramming roadmap reveals route to human induced trophoblast stem cells. Nature *586*, 101–107.

Marión, R.M., Strati, K., Li, H., Murga, M., Blanco, R., Ortega, S., Fernandez-Capetillo, O., Serrano, M., and Blasco, M.A. (2009). A p53-mediated DNA damage response limits reprogramming to ensure iPS cell genomic integrity. Nature *460*, 1149–1153.

Martinelli, P., Carrillo-De Santa Pau, E., Cox, T., Sainz, B., Dusetti, N., Greenhalf, W., Rinaldi, L., Costello, E., Ghaneh, P., Malats, N., et al. (2017). GATA6 regulates EMT and tumour dissemination, and is a marker of response to adjuvant chemotherapy in pancreatic cancer. Gut *66*, 1665–1676.

Melendez, E., Chondronasiou, D., Mosteiro, L., Martinez de Villareal, J., Fernández-Alfara, M., Lynch, C., Grimm, D., Real, F.X., Alcamí, J., Climent, N., et al. (2022). Natural killer cells act as an extrinsic barrier for in vivo reprogramming. Development *149*, dev200361.

Mosteiro, L., Pantoja, C., Alcazar, N., Marión, R.M., Chondronasiou, D., Rovira, M., Fernandez-Marcos, P.J., Muñoz-Martin, M., Blanco-Aparicio, C., Pastor, J., et al. (2016). Tissue damage and senescence provide critical signals for cellular reprogramming in vivo. Science *354*, aaf4445.

O'Malley, J., Skylaki, S., Iwabuchi, K.A., Chantzoura, E., Ruetz, T., Johnsson, A., Tomlinson, S.R., Linnarsson, S., and Kaji, K. (2013). High-resolution analysis with novel cell-surface markers identifies routes to iPS cells. Nature *499*, 88–91.

Ohnishi, K., Semi, K., Yamamoto, T., Shimizu, M., Tanaka, A., Mitsunaga, K., Okita, K., Osafune, K., Arioka, Y., Maeda, T., et al. (2014). Premature termination of reprogramming in vivo leads to cancer development through altered epigenetic regulation. Cell *156*, 663–677.

Parenti, A., Halbisen, M.A., Wang, K., Latham, K., and Ralston, A. (2016). OSKM induce extraembryonic endoderm stem cells in parallel to induced pluripotent stem cells. Stem Cell Rep. *6*, 447–455.

Polo, J.M., Anderssen, E., Walsh, R.M., Schwarz, B.A., Nefzger, C.M., Lim, S.M., Borkent, M., Apostolou, E., Alaei, S., Cloutier, J., et al. (2012). A molecular roadmap of reprogramming somatic cells into iPS cells. Cell *151*, 1617–1632.

Qiu, X., Mao, Q., Tang, Y., Wang, L., Chawla, R., Pliner, H.A., and Trapnell, C. (2017). Reversed graph embedding resolves complex single-cell trajectories. Nat. Methods *14*, 979–982.

Rappa, G., Green, T.M., and Lorico, A. (2014). The nuclear pool of tetraspanin CD9 contributes to mitotic processes in human breast carcinoma. Mol. Cancer Res. *12*, 1840–1850.

Real, F.X., Vilá, M.R., Skoudy, A., Ramaekers, F.C., and Corominas, J.M. (1993). Intermediate filaments as differentiation markers of exocrine pancreas. II. Expression of cytokeratins of complex and stratified epithelia in normal pancreas and in pancreas cancer. Int. J. Cancer *54*, 720–727.

van Roey, R., Brabletz, T., Stemmler, M.P., and Armstark, I. (2021). Deregulation of transcription factor networks driving cell plasticity and metastasis in pancreatic cancer. Front. Cell Dev. Biol. *9*, 753456.

Schiebinger, G., Shu, J., Tabaka, M., Cleary, B., Subramanian, V., Solomon, A., Liu, S., Lin, S., Berube, P., Lee, L., et al. (2017). Reconstruction of developmental landscapes by optimal-transport analysis of single-cell gene expression sheds light on cellular reprogramming. Preprint at bioRxiv. https://doi.org/10.1101/191056.

Schiebinger, G., Shu, J., Tabaka, M., Cleary, B., Subramanian, V., Solomon, A., Gould, J., Liu, S., Lin, S., Berube, P., et al. (2019). Optimal-Transport analysis of single-cell gene expression identifies developmental trajectories in reprogramming. Cell *176*, 1517.

Schlesinger, Y., Yosefov-Levi, O., Kolodkin-Gal, D., Granit, R.Z., Peters, L., Kalifa, R., Xia, L., Nasereddin, A., Shiff, I., Amran, O., et al. (2020). Single-cell transcriptomes of pancreatic preinvasive lesions and cancer reveal acinar metaplastic cells' heterogeneity. Nat. Commun. *11*, 4516.

Schwarz, B.A., Cetinbas, M., Clement, K., Walsh, R.M., Cheloufi, S., Gu, H., Langkabel, J., Kamiya, A., Schorle, H., Meissner, A., et al. (2018). Prospective isolation of poised iPSC intermediates reveals principles of cellular reprogramming. Cell Stem Cell *23*, 289–305.e5.

Shibata, H., Komura, S., Yamada, Y., Sankoda, N., Tanaka, A., Ukai, T., Kabata, M., Sakurai, S., Kuze, B., Woltjen, K., et al. (2018). In vivo reprogramming drives Kras-induced cancer development. Nat. Commun. *9*, 2081.

Stadtfeld, M., Maherali, N., Breault, D.T., and Hochedlinger, K. (2008). Defining molecular cornerstones during fibroblast to iPS cell reprogramming in mouse. Cell Stem Cell *2*, 230–240.

Street, K., Risso, D., Fletcher, R.B., Das, D., Ngai, J., Yosef, N., Purdom, E., and Dudoit, S. (2018). Slingshot: cell lineage and pseudotime inference for single-cell transcriptomics. BMC Genom. *19*, 477.

Takahashi, K., and Yamanaka, S. (2006). Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. Cell *126*, 663–676.

Upadhyay, G. (2019). Emerging role of lymphocyte antigen-6 family of genes in cancer and immune cells. Front. Immunol. *10*, 819.

Utikal, J., Maherali, N., Kulalert, W., and Hochedlinger, K. (2009). Sox2 is dispensable for the reprogramming of melanocytes and melanoma cells into induced pluripotent stem cells. J. Cell Sci. *122*, 3502–3510.

Vasiliou, V., and Nebert, D.W. (2005). Analysis and update of the human aldehyde dehydrogenase (ALDH) gene family. Hum. Genom. *2*, 138–143.

Vasilyeva, N.A., Loktyushov, E.V., Bychkov, M.L., Shenkarev, Z.O., and Lyukmanova, E.N. (2017). Three-finger proteins from the ly6/uPAR family: functional diversity within one structural motif. Biochemistry. *82*, 1702–1715.

Wang, V.M.Y., Ferreira, R.M.M., Almagro, J., Evan, T., Legrave, N., Zaw Thin, M., Frith, D., Carvalho, J., Barry, D.J., Snijders, A.P., et al. (2019). CD9 identifies pancreatic cancer stem cells and modulates glutamine metabolism to fuel tumour growth. Nat. Cell Biol. *21*, 1425–1435.

Wang, X., Allen, W.E., Wright, M.A., Sylwestrak, E.L., Samusik, N., Vesuna, S., Evans, K., Liu, C., Ramakrishnan, C., Liu, J., et al. (2018). Three-dimensional intact-tissue sequencing of single-cell transcriptional states. Science *361*, eaat5691.

Xie, Z.B., Zhang, Y.F., Jin, C., Mao, Y.S., and Fu, D.L. (2019). LRG-1 promotes pancreatic cancer growth and metastasis via modulation of the EGFR/p38 signaling. J. Exp. Clin. Cancer Res. *38*, 1–12.

Xing, Q.R., El Farran, C.A., Gautam, P., Chuah, Y.S., Warrier, T., Toh, C.X.D., Kang, N.Y., Sugii, S., Chang, Y.T., Xu, J., et al. (2020). Diversification of reprogramming trajectories revealed by parallel single-cell transcriptome and chromatin accessibility sequencing. Sci. Adv. *6*, eaba1190.

Yui, S., Azzolin, L., Maimets, M., Pedersen, M.T., Fordham, R.P., Hansen, S.L., Larsen, H.L., Guiu, J., Alves, M.R.P., Rundsten, C.F., et al. (2018). YAP/TAZ-Dependent reprogramming of colonic epithelium links ECM remodeling to tissue regeneration. Cell Stem Cell *22*, 35–49.e7.

Zviran, A., Mor, N., Rais, Y., Gingold, H., Peles, S., Chomsky, E., Viukov, S., Buenrostro, J.D., Scognamiglio, R., Weinberger, L., et al. (2019). Deterministic somatic cell reprogramming involves continuous transcriptional changes governed by myc and epigenetic-driven modules. Cell Stem Cell *24*, 328–341.e9.