Long Noncoding RNAs Enforce Pancreatic Cancer Identity and Block Reprogramming

Dmytro Grygoryev[1,2*], Seung-Won Lee[1,2*], Connor Mitchell Frankston[1,3], Shauna Rakshe[4], Mark Berry[1], Alex Hirano[5], Taelor Ekstrom[1], Elise Manalo[1], Julien Tessier[1], Marilynn Chow-Castro[1], Jason Link[2,6], Dove Keith[6], Brett C. Sheppard[6, 7], Suzanne Fei[4], Terry Morgan[1, 7, 8], Helen E. Remotti[10], Wenli Yang[11], Emma E. Furth[12], Sudhir Thakurela[5], Rosalie Sears[2,6,8], and Jungsun Kim[1,2,8,13]

[1]Cancer Early Detection Advanced Research Center, Knight Cancer Institute, Oregon Health & Science (OHSU) School of Medicine, Portland, OR.

[2]Department of Molecular and Medical Genetics, OHSU School of Medicine, Portland, OR.

[3]Biomedical engineering graduate program, Department of Biomedical Engineering, OHSU School of Medicine, Portland, OR.

[4]Biostatistics Shared Resource, Knight Cancer Institute, OHSU, Portland, OR

[5]Lurie Center for Autism, Department of Pediatrics, Harvard Medical School, Boston, MA

[6]Brenden-Colson Center for Pancreatic Care, OHSU School of Medicine, Portland, OR.

[7]Department of Surgery, OHSU School of Medicine, Portland, OR.

[8]Knight Cancer Institute, OHSU School of Medicine, Portland, OR.

[9]Pathology & Laboratory Medicine, OHSU School of Medicine, Portland, OR.

[10]Department of Pathology, Columbia University, New York, NY.

[11]Institute for Regenerative Medicine, UPenn Perelman School of Medicine, Philadelphia, PA.

[12]Pathology and Laboratory Medicine, UPenn Perelman School of Medicine, Philadelphia, PA.

* Equal contribution

[13]corresponding author: Dr. Jungsun Kim (kimjungs@ohsu.edu)

Keywords (max 10): Cancer cell identity, cancer epigenetic memory, long noncoding RNA, master transcription factor-mediated cellular reprogramming, pancreatic Cancer, pluripotency, and plasticity.

Highlight

Summary (150 words or fewer)

Pancreatic ductal adenocarcinoma (PDAC), one of the deadliest malignancies, paradoxically exhibits both transcriptional plasticity and rigidity. Despite dynamic transcriptional states, PDAC maintains a stable pancreatic epithelial identity that resists therapeutic intervention. This rigidity reflects robust transcriptional memory, but the underlying mechanisms remain poorly understood. While the Yamanaka factors (OCT4, SOX2, KLF4, MYC; OSKM) can reprogram normal cells by overriding established cell identity, solid tumors like PDAC remain refractory, suggesting the presence of active epigenetic barriers. Long noncoding RNAs (lncRNAs) regulate chromatin state and cell identity, but whether they actively enforce cancer-specific transcriptional memory and restrict reprogramming remains largely unexplored. Herein, through CRISPR interference screens targeting cancer-associated lncRNAs, we identified PDAC-enriched lncRNAs that restrict OSKM-mediated reprogramming. Knockdown of these lncRNAs enabled the acquisition of pluripotency-associated features, suppressed PDAC identity programs, and reduced tumorigenicity in vivo. Among the most potent repressors of reprogramming were ATXN7L3-AS1, INTS4P2, and AC079921.2, which were upregulated in PDAC but minimally expressed in normal pancreas. While ATXN7L3-AS1 harbors pathogenic somatic variants in PDAC and is associated with mitotic progression-potentially contributing to the faithful inheritance of malignant programs, AC079921.2 is linked to extracellular matrix organization genes that resist suppression during OSKM induction. Thus, these findings uncover a previously unrecognized role for cancer-associated lncRNAs as key enforcers of transcriptional memory in PDAC. Targeting them provides a conceptual and therapeutic framework for erasing cancer cell memory, resetting the epigenetic state, and re-sensitizing PDAC cells to differentiation- or reprogramming-based interventions.

Introduction (no subheading, max 7,000 including main figure legends).

Pancreatic ductal adenocarcinoma (PDAC) is one of the deadliest human cancers, emphasizing the urgent need for deeper biological understanding to develop new therapies[1]. Many cancer cells show remarkable transcriptional plasticity and heterogeneity. Yet, their epigenetic landscape plays a crucial role in limiting this plasticity, maintaining a stable, cancer-specific identity. PDAC exemplifies this paradox; it exhibits transcriptional plasticity with dynamic transitions between classical and basal subtypes, yet it keeps a stable core identity defined by oncogenic transcriptional programs and classical pancreatic epithelial signatures[2-8]. This coexistence of plasticity and rigidity contributes to therapeutic resistance and aggressive clinical behavior. However, the molecular mechanisms that constrain their transcriptional memory and preserve malignant identity remain poorly understood.

Experimental reprogramming through the forced expression of lineage-defining master transcription factors (TFs), such as pioneer factors, has demonstrated the capacity to override established cell identity[9-14]. The best-known example is the induction of pluripotency via Yamanaka factors (OCT4, SOX2, KLF4, and MYC: OSKM), which reprograms normal somatic cells into induced pluripotent stem cells (iPSCs)[13,14]. However, in solid tumors including PDAC, reprogramming is often impeded: reprogrammed cancer cells frequently retain epigenetic memory of their tissue of origin and resist full identity erasure[15-19]. This resistance is unexpected given the inherent plasticity of many cancers, suggesting the presence of robust, cancer-specific transcriptional memory mechanisms.

Cellular identity is maintained by tightly coordinated gene regulatory networks composed of TFs, cofactors, and chromatin regulators[20-25]. Although these networks can be reset during regeneration or experimental reprogramming, core networks encode a transcriptional memory preserved through mitotic bookmarking and chromatin state[23-27]. This memory stabilizes cell identity and acts as a barrier to fate conversion[28-39]. Long noncoding RNAs (lncRNAs) have emerged as critical regulators of cell memory. By scaffolding chromatin remodeling complexes and guiding gene expression in cis and trans, lncRNAs orchestrate cell type-specific

transcriptional programs[40-50]. They are essential for maintaining pluripotency and self-renewal in embryonic stem (ES) cells, and have been shown to both facilitate and constrain cellular reprogramming[51,52]. Dysregulation of lncRNAs is increasingly recognized in cancer, where they contribute to disease progression and therapeutic resistance[53-56]. However, it remains unclear whether specific lncRNAs in solid tumors such as PDAC act as active molecular barriers that preserve malignant cell identity and impede cell fate conversion.

We hypothesized that a subset of cancer-associated lncRNAs enforces transcriptional memory in PDAC, thereby preserving malignant identity and preventing OSKM-induced reprogramming. To test this, we performed CRISPR interference (CRISPRi) screens targeting cancer-associated lncRNAs in both normal pancreatic and PDAC cells under OSKM induction. These screens identified lncRNAs that constrain plasticity, maintain PDAC-specific transcriptional programs, and suppress the emergence of pluripotent states. Remarkably, their suppression, in combination with OSKM, enabled the reprogramming of PDAC cells, erased their malignant identity, and reduced their tumorigenic potential in vivo. These findings uncover a previously unrecognized role for cancer-associated lncRNAs in enforcing transcriptional memory and stabilizing PDAC cell identity. Together, our results demonstrate that depletion of such lncRNAs, alongside OSKM induction, can override transcriptional rigidity and promote functional reprogramming. This work lays the foundation for reprogramming-based therapeutic strategies aimed at erasing malignant transcriptional memory, resetting the epigenetic landscape, and re-sensitizing PDAC cells to differentiation- or reprogramming-based interventions.

RESULTS

Restricted Plasticity of PDAC Cells Toward a Pluripotent State. Primary PDAC samples include a large proportion of normal cells that are more amenable to reprogramming (Table S1)[57]. To accurately assess the reprogramming of human PDAC cells, we first enriched these

cells using patient-derived xenografts (PDXs)[57]. PDAC comprises two major transcriptional subtypes: classical and basal-like (also referred to as squamous)[2-4], and we fully characterized PDX tumors representing both subtypes[57]. To assess the reprogramming potential across all PDAC subtypes, we employed PDX-derived PDAC cells and established PDAC cell lines representing both classical (CFPAC1, ST-00014490 PDX-derived PDAC) and basal (Panc1, ST-00007599 PDX-derived PDAC) subtypes, using BJ human fibroblasts and H6c7 human normal pancreatic ductal epithelial cells (hnPDECs) as controls (Fig. 1A)[57].

To enable transgene-free reprogramming, we utilized Sendai virus–mediated delivery of the polycistronic vector expressing OSKM[58]. Because the Sendai virus is a non-integrating RNA virus, it provides a clean platform for reprogramming without genomic integration[59]. This system achieved efficient and consistent transgene expression across all primary PDAC cells, comparable to that in H6c7 hnPDECs and BJ fibroblasts[59].

We first validated the Sendai virus–mediated OSKM delivery system using human BJ fibroblasts and human acinar cells isolated from pancreatitis tissue of PDAC patients (Fig. S1A-S1B). The reprogramming efficiency of acinar cells (0.4%, 122/30,000 cells) was comparable to that of fibroblasts (Fig. S1C). The resulting iPSC lines expressed the canonical pluripotency marker TRA-1-60, exhibited normal karyotypes, and successfully differentiated into all three germ layers (Fig. S1D-F). These results confirmed that Sendai virus–mediated OSKM delivery effectively reprogrammed human pancreatic acinar cells and fibroblasts into a pluripotent state.

In contrast to pancreatic acinar cells, H6c7 hnPDECs failed to express TRA-1-60 by day 30 post-OSKM induction, although many expressed SSEA-1, a marker associated with the pluripotency of murine cells but with the early embryonic differentiation in human cells (Fig. S2A). TRA-1-60 expression eventually emerged by day 60 (data not shown), indicating that hnPDECs possessed a stronger lineage memory than acinar cells and fibroblasts. Classical-type PDAC cells expressed SSEA-1 but failed to express TRA-1-60, consistent with stalled progression toward pluripotency, after OSKM induction (Fig. S2A). OSKM induction did not change the

SSEA-1 and TRA-1-60 expression pattern of basal-type PDAC cells (Fig.S2A), suggesting a failure to rewire their transcriptional identity.

The persistent expression of SSEA-1 in PDAC cells indicated partial lineage rewiring without successful acquisition of pluripotency. Even after 60 days of OSKM induction, PDAC cells failed to achieve full reprogramming (data not shown). These findings suggested that PDAC cells retained a cancer-specific transcriptional memory that resisted reprogramming, creating a barrier to full cell fate conversion.

A Subset of PDAC Identity Genes Can Be Overridden by OSKM, but Many Others Cannot. Early molecular events following OSKM induction are critical in determining cell fate during reprogramming[29,60-66]. To characterize early transcriptional responses in PDAC versus normal cells, we examined transcriptomic profiling of PDAC cells (ST-00007599, ST-00014490, Panc1, CFPAC1) and control cells (BJ fibroblasts and H6c7 hnPDECs) at day 7 post-OSKM induction (Fig. 1A). Notably, ectopic OSKM expression levels were comparable across all cell types (Fig. S2B), and PCA data show consistent patterns across technical replicates within each sample (Fig. S3A), enabling direct comparison of downstream transcriptional responses.

OSKM typically initiates reprogramming of normal cells by silencing original lineage-specific gene programs[29,60-66]. We therefore hypothesized that OSKM might downregulate a subset of the PDAC-specific transcriptome, particularly those involved in the malignant identity and potential vulnerabilities. To test this, we first defined PDAC-unique genes by identifying transcripts significantly upregulated across all PDAC cells relative to both fibroblasts and hnPDECs before OSKM treatment (Fig. 1B, Table S2A). Only 7 out of 375 of these PDAC-unique genes were significantly downregulated (>2-fold, $q < 0.05$) on day 7 following OSKM induction (Fig. 1B, Table S2B). These OSKM-responsive genes were enriched in pathways related to apoptosis and epithelial cell differentiation (Table S2C). *SPDEF*, encoding a transcription factor essential for classical PDAC tumorigenesis[67], was among OSKM-responsive

genes and whose high expression in PDAC was associated with poor patient prognosis in TCGA datasets (Fig. 1C-D).

To systematically assess transcriptional responses to OSKM across cell types, we compared gene expression changes at day 7 post-induction. BJ fibroblasts exhibited the greatest number of differentially expressed genes (DEGs), followed by H6c7 hnPDECs, whereas PDAC cells showed markedly fewer DEGs overall (Fig. S3B, Table S2D). This reduced DEG count in PDAC may reflect inter-sample heterogeneity, which can mask consistent transcriptional changes at the population level. Indeed, when analyzed at the individual cell level, each PDAC sample exhibited more DEGs than H6c7 hnPDECs (Fig. S3B, Table S2E), suggesting that PDAC cells harbor higher intrinsic transcriptional plasticity. Importantly, this plasticity is not accompanied by a loss of malignant identity. Rather, it is constrained by a robust transcriptional memory that preserves PDAC-specific gene programs despite OSKM-mediated reprogramming.

Collectively, these findings indicate that while OSKM can partially suppress tumor-associated gene programs, the core transcriptional identity of PDAC cells remains largely intact.

Transcriptional Memory and Epigenetic Barriers Limit OSKM Reprogramming in PDAC. Despite OSKM induction, PDAC cells maintain a stable malignant transcriptional state, suggesting the presence of transcriptional memory that actively resists reprogramming. To uncover the molecular programs underlying this resistance, we compared gene expression changes in PDAC cells versus normal cells (BJ fibroblasts and H6c7 hnPDECs) following OSKM induction. We hypothesized that “genes unresponsive to OSKM” in PDAC function as “safeguards” that preserve malignant identity, while “genes aberrantly upregulated in PDAC”—despite being silenced or unchanged in normal cells—reflect “cancer-specific reprogramming barriers” that may antagonize cell fate conversion. These aberrant responses could not only

inhibit effective reprogramming but also enhance PDAC plasticity in maladaptive ways, potentially reinforcing malignancy.

To account for sample heterogeneity, we developed a classification framework that incorporates inter-sample variability across PDAC (n=4) and normal (n=2) cells. Differentially expressed genes before and after OSKM treatment were identified in each cell using likelihood ratio testing (LRT) with $\log_2$(fold change) > 0.2 and FDR < 0.05 (Fig. 1E, Fig. S3C). Each DEG was then ranked based on a weighted Hamming distance[68] relative to predefined expression change patterns, including: (1) downregulation in normal cells with minimal change in cancer ("OSKM unresponsive genes in PDAC"), (2) no change or downregulation in normal cells and upregulation in cancer ("OSKM aberrantly upregulated genes in PDAC") (Fig. 1F, Table S2F-H, method).

Strikingly, OSKM-unresponsive genes were enriched in cell adhesion pathways, while OSKM-aberrantly upregulated genes were associated with mesodermal and morphogenetic lineage programs (Fig. 1G–H). Both categories were predicted targets of PRC2 core components (SUZ12, EZH2) (Fig. 1G-H), and significantly overlapped with known SUZ12 targets in human and mouse embryonic stem (ES) cells (Fig. S3E). Notably, expression of PRC2 subunit genes remained unchanged post-OSKM (Fig. S3F), suggesting that impaired recruitment—not subunit abundance—may underlie the failure to suppress these programs.

In summary, OSKM induction downregulated some PDAC marker genes but did not affect, or even upregulated, the expression of the majority of PDAC-specific genes. Many of these were PRC2-regulated genes involved in mesodermal lineage, indicating a key epigenetic barrier to successful reprogramming in PDAC.

Suppression of Cancer-Associated lncRNAs in PDAC Cells Significantly Enhances the Reprogramming Efficiency. Given the established role of long noncoding RNAs (lncRNAs) in directing PRC2 recruitment and maintaining transcriptional memory across diverse biological

contexts[41,50,69-71], we hypothesized that a subset of cancer-associated lncRNAs in PDAC actively sustains malignant identity. We further reasoned that depleting these lncRNAs could weaken transcriptional memory and facilitate OSKM-mediated reprogramming by disrupting cancer-specific epigenetic barriers. Accordingly, we performed CRISPRi–based screens targeting 691 cancer-associated lncRNAs using a pooled lentiviral library containing 7,087 gRNAs (10 gRNAs per lncRNA, plus non-targeting controls) (Table S3A)[72,73]. The screens were conducted in the context of OSKM-mediated reprogramming of normal ductal cells and PDAC cells, using the H6c7 and HPAFII cell lines engineered to stably express dCas9-KRAB (henceforth referred to as H6c7-dCas9-gRNA and HPAFII-dCas9-gRNA; Fig. 2A, Fig. S3G).

Strikingly, by day 30 post-OSKM induction, a significantly higher frequency of TRA-1-60^{+} pluripotent stem-like cells emerged in the HPAFII-dCas9-gRNA population compared to controls (Fig. 2B, Table S3B). Isolated TRA-1-60^{+} and TRA-1-60^{-} (Hereafter referred to as dCas9-gRNA-OSKM-TP and dCas9-gRNA-OSKM-TN, respectively) populations remained phenotypically stable over multiple passages (Fig. 2C). In contrast, the same gRNA pool targeting lncRNAs did not enhance reprogramming of H6c7 hnPDEC cells into pluripotency (Fig. S3H), suggesting that these lncRNAs function in a PDAC-specific manner to preserve cancer-associated transcriptional memory.

To investigate the transcriptional identity of PDAC cells reprogrammed with OSKM and depleted of lncRNAs, we performed single-cell RNA sequencing (scRNA-seq) on dCas9-gRNA-OSKM-TP and dCas9-gRNA-OSKM-TN populations, along with appropriate controls: dCas9, dCas9-gRNA, and dCas9-OSKM-TP (rare TRA-1-60^{+} cells generated from OSKM-reprogrammed HPAFII cells without lncRNA knockdown) (Fig. 3A).

After filtering out low-quality cells with >10% mitochondrial gene content and regressing out cell cycle–associated transcripts, we retained high-quality transcriptomes for downstream analysis (Fig. S4A-B). We merged all samples and identified 15 transcriptionally distinct clusters using the Louvain algorithm (Fig. 3A–C). HPAFII-dCas9 cells primarily occupied clusters

corresponding to tubular epithelial (Group 0) and invasive epithelial (Group 1) cell types, along with several minor heterogeneous populations (Fig. 3B–C, Fig. S4C). As expected, dCas9-OSKM-TP cells (green) were transcriptionally indistinguishable from dCas9 cells (black), with no DEGs detected between the two groups (Fig. 3B), highlighting that OSKM alone was insufficient to reprogram PDAC cells.

Although dCas9-gRNA cells (light green, Fig. 3B) upregulated pyruvate metabolism genes and downregulated epithelial and oxidative phosphorylation genes compared to parental HPAFII-dCas9 cells (Fig. S4D), they remained restricted to Groups 0 and 1, consistent with persistent PDAC transcriptional identity (Fig. 3B–C).

In contrast, dCas9-gRNA-OSKM-TP (Red, Fig. 3B) cells were markedly depleted from Groups 0 and 1 and instead clustered predominantly in Group 5—a population unique to TRA-1-60^{+} cells (Fig.3B-C, Fig. S4E). Indeed, PCA analysis further confirmed that dCas9-gRNA-OSKM-TP cells were transcriptionally distinct from other subpopulations, reflecting reprogramming-induced identity shifts (Fig. S4F). Group 5 was defined by loss of pancreatic and germ layer markers and increased expression of pluripotency-associated genes (Fig. 3B–J, Table S4A-C), indicating a transition toward a pluripotent state.

To assess the developmental identity of HPAFII-dCas9-gRNA-OSKM-TP cells, we compared their transcriptomic profiles with reference datasets representing human PDAC[74], human and mouse normal pancreas[75], human early embryo before the three germ layers formed[76], and Parkinson's disease patients' fibroblast-derived iPSC[77]. PCA visualization showed that dCas9-gRNA-OSKM-TP cells clustered closely with early human embryos, and distinctly apart from PDAC or normal pancreatic cells (Fig. 3K).

Together, these findings demonstrate that lncRNA knockdown enhances OSKM-induced reprogramming and enables the emergence of cells with transcriptional features characteristic of pluripotent-like states. In contrast, OSKM expression alone was insufficient to erase the transcriptional identity of PDAC cells.

Suppression of Cancer-Associated lncRNAs During OSKM Reprogramming Rewires PDAC Gene Programs. Next, we investigated whether the cell state transition induced by the combination of lncRNA depletion and OSKM induction was accompanied by transcriptional erasure of PDAC identity. Notably, cell clusters corresponding to mucinous (Group 0) and invasive epithelial (Group 1) PDAC phenotypes were markedly depleted in both dCas9-gRNA-OSKM-TP and -TN populations (Fig. 3B–C), suggesting that OSKM induction in combination with lncRNA knockdown broadly disrupts cancer-specific programs—even in TRA-1-60^{-} cells.

To systematically assess the extent to which PDAC identity is suppressed, we curated "PDAC-specific signature genes" from published scRNA-seq datasets by identifying transcripts enriched in PDAC cells but expressed in fewer than 20% of normal pancreatic cells[74,75]. PDAC-specific signature genes were significantly enriched in downregulated genes in dCas9-gRNA-OSKM-reprogrammed PDAC cells, especially TP cell populations, relative to dCas9-gRNA controls (Fig. 3L). By contrast, no significant downregulation was observed in cell populations treated with OSKM or gRNA only, indicating that both OSKM expression and lncRNA knockdown are required to repress PDAC transcriptional programs effectively.

We next examined whether this transcriptional reset included the reactivation of tumor suppressor genes (TSG) listed in the OncoKB database (Table S4D)[78,79]. As expected, parental HPAFII-dCas9 cells showed minimal or no expression of TSGs (Fig. 3M). In contrast, cells treated with either OSKM or dCas9-gRNA re-expressed distinct subsets of TSGs. Notably, reprogrammed pluripotent cells (dCas9-gRNA-OSKM-TP) exhibited the most robust re-expression of TSGs (Fig. 3M).

We further characterized subpopulations of dCas9-gRNA–OSKM–TP cells re-expressing TSG (Fig. 3B–C, group 5). Using unsupervised k-means clustering, TSGs were grouped into three distinct clusters based on their expression profiles (Fig. 3N). Comparing expression levels of TSGs across cell states defined by NANOG and TRA-1-60 expressions-

NANOG$^-$/TRA-1-60$^-$ (N$^-$/T$^-$), NANOG$^-$/TRA-1-60$^+$ (N$^-$/T$^+$), NANOG$^+$/TRA-1-60$^-$ (N$^+$/T$^-$), and NANOG$^+$/TRA-1-60$^+$ (N$^+$/T$^+$)-revealed that TSG expression was predominantly enriched in NANOG$^+$ states (Fig. 3O).

Together, these findings revealed that OSKM reprogramming in combination with lncRNA depletion not only induced a pluripotent-like state but also erased the PDAC identity and restored TSG expression, suggesting a true reversal of malignant cell fate.

LncRNA Depletion in Combination with OSKM Erases PDAC Identity and Reduces Malignant Tumor Formation In Vivo. We next assessed whether the pluripotency and suppression of the PDAC identity driven by lncRNA depletion in combination with OSKM induction altered the tumorigenic potential in vivo.

To this end, we xenografted HPAFII-dCas9-gRNA-OSKM-TP and -TN cells into immunodeficient NOD-SCID-IL2Rγ^null (NSG) mice, using HPAFII, HPAFII-dCas9-gRNA, and HPAFII-dCas9-OSKM-TP cells as controls. In addition, we included human ES (huES) cells as a pluripotency control (Fig. 4A, S5A-E). As expected, HPAFII cells rapidly formed tumors within one-month post-injection (Fig. 4B–C). All control groups similarly formed tumors, although dCas9-gRNAs produced slightly smaller tumors compared to HPAFII cells. Strikingly, mice injected with dCas9-gRNA-OSKM-TP cells either failed to develop tumors or formed significantly smaller tumors than dCas9-gRNA-OSKM-TN and all controls ($p < 0.005$; Fig. 4B–C).

Histopathological analysis confirmed that tumors from all control groups displayed classical grade 2 adenocarcinoma morphology, with strong human-specific marker TRA-1-85, pancreatic ductal marker CK19, and proliferation marker Ki67 immunostainings and absence of apoptotic marker Cleaved Caspase-3 (CC3) immunostaining, indicating high proliferation and minimal apoptosis (Fig. 4D-G, 5A-P). Extensive central necrosis was observed, consistent with advanced PDAC pathology. In contrast, tumors derived from dCas9-gRNA-OSKM-TP cells lacked typical adenocarcinoma features (Fig. 4H) and showed robust CC3 immunostaining in

addition to Ki67 immunostaining (Fig. 5Q-T), suggesting a transition toward a less aggressive, more benign-like phenotype.

Importantly, tumors from dCas9-OSKM-TP and dCas9-gRNAs cells retained classical adenocarcinoma histology (Fig. 4E-F, 5E-F, and 5I-J), underscoring that neither OSKM expression nor lncRNA knockdown alone was sufficient to reprogram the malignant phenotype. Interestingly, dCas9-OSKM-TP tumors contained more differentiated, non-epithelial cell types in the tumor center, reflecting partial but incomplete identity shifts (Fig. 4F).

Consistent with being pluripotent, human ES cells formed fully mature teratomas within three months post-injection[80,81]. At the one-month point used for PDAC comparisons, ES cell-derived tumors were small and composed predominantly of neuronal epithelial cells, consistent with ectodermal differentiation upon early exit from pluripotency (Fig. S5A-C). These early teratomas possess both proliferation and apoptotic markers, further distinguishing them from HPAFII-derived tumors (Fig. S5D-E).

Together, these findings demonstrate that lncRNA depletion in combination with OSKM expression effectively reprogrammed PDAC cells, resulting in reduced tumorigenicity, loss of classical adenocarcinoma identity, and reactivation of apoptotic programs, in vivo. Importantly, this phenotypic reversal was not achieved by OSKM expression or lncRNA knockdown alone, highlighting the requirement of both to overcome the malignant PDAC state.

PDAC-Enriched lncRNAs Act as Barriers to Cellular Reprogramming and Promote the Tumor Identity. To characterize lncRNAs that maintained PDAC gene expression and were selectively depleted during successful reprogramming, we performed next-generation sequencing of genomic DNA from FACS-sorted dCas9-gRNA-OSKM-TN, dCas9-gRNA-OSKM-TP, and dCas9-gRNA cells (Fig. 6A, S5A, Table S5A-E). We applied Z-score normalization and MAGeCK[82] negative control normalization to gRNA count data from two independently passaged cultures (separated by 3–6 passages) and selected gRNAs with Z-scores > 2 and

positive beta scores ($p < 0.05$) (Table S5F-J, methods). gRNAs consistently enriched in both cultures, as defined by either Z-score or MAGeCK analysis, were retained (Table S5K; Fig. S5A).

This stringent analysis identified several lncRNAs-targeting gRNAs that were enriched in dCas9-gRNA-OSKM-TP cells compared to dCas9-gRNA (Fig. 6A-B; Fig. S5F-I). Among these, ATXN7L3-AS1 (ENSG00000267394), AC079921.2 (ENSG00000249685), and INTS4P2 (ENSG00000273024, located within LINC03006:ENSG00000290926) were validated by qRT-PCR and found to be significantly downregulated in HPAFII-dCas9-gRNA-TP cells compared to control HPAFII-dCas9 (Fig. 6C). Importantly, these lncRNAs were minimally expressed or absent in normal pancreatic tissue (Fig. 6D-F) and their expression remained unchanged in response to OSKM alone in both PDAC and normal control cells (Fig. S6A). —suggesting resistance to OSKM-mediated reprogramming.

To investigate how individual lncRNAs contribute to PDAC transcriptional identity, we established HPAFII PDAC cells expressing CRISPRi gRNAs targeting ATXN7L3-AS1, AC079921.2, or INTS4P2, along with non-targeting control (NT) gRNAs. As a comparison, we introduced the same gRNAs into non-malignant pancreatic ductal epithelial H6c7 cells. In HPAFII cells, knockdown of each lncRNA led to a reduction in multiple PDAC-enriched genes, including basal subtype markers (KRT5, TP63), classical ductal TFs (GATA6, HNF1B) (Fig. 7A). Intriguingly, knockdown of ATXN7L3-AS1 or AC079921.2 modestly increased the expression of pancreatic progenitor markers PDX1, FOXA2, and SOX9, suggesting partial de-repression of developmental programs. In contrast, knockdown of INTS4P2 had minimal effects on endodermal gene expression. These transcriptomic changes were largely absent in H6c7 cells (Fig. 7B), underscoring the context-specific function of these lncRNAs in maintaining PDAC identity. Supporting its functional relevance, multiple pathological somatic variants are clustered within the ATXN7L3-AS1 locus in PDAC samples according to COSMIC[83], further implicating this lncRNA in maintaining malignant identity (Fig. 7C).

To gain insight into the potential functions of these lncRNAs, we examined their co-expression patterns across the TCGA Pan-Cancer RNA-seq dataset. Expression data were available for ATXN7L3-AS1 and AC079921. The expression of ATXN7L3-AS1 was positively correlated with the expression of genes involved in mitotic cell cycle progression (Fig. S6B). These genes were predicted to have UBTF as a potential upstream regulator based on motif analysis (Fig. S6C) and carried active enhancer marks such as H3K27ac and H3K4me2 associated with transcriptional activity and memory (Fig. S6D).

On the other hand, AC079921.2 expression was positively correlated with genes involved in extracellular matrix (ECM) remodeling, cell adhesion, and migration—key hallmarks of invasive PDAC (Fig. S6E). These correlated genes were predicted SUZ12/PRC2 targets (Fig. S6F) and showed promoter enrichment for H3K27me3 (Fig. S6G), suggesting that AC079921.2 may counteract PRC2-mediated transcriptional silencing. Notably, genes that were either aberrantly upregulated or unresponsive to OSKM-induced reprogramming (Fig. 1F–H) significantly overlapped with AC079921.2–associated genes (Fig. 7D–E). This shared gene set was enriched for morphogenesis, migration, and ECM-related pathways (Fig. 7F–G), and similarly predicted to be SUZ12 targets (Fig. 7H–I). Together, these findings indicate that AC079921.2 reinforces malignant PDAC identity by sustaining pro-invasive, morphogenetic transcriptional programs that resist OSKM-mediated resetting.

Intriguingly, genes whose expression negatively correlated with ATXN7L3-AS1 or AC079921.2 were enriched for translational and metabolic pathways (Fig. S6H–I). These genes were marked by active histone modifications (H3K36me3 and H3K79me2; Fig. S6J–K) and predicted to be targets of transcriptional regulators involved in growth and metabolism, including MYC, MAX, TAF1, and GABPA (Fig. S6L–M). These findings suggest that ATXN7L3-AS1 and AC079921.2 may suppress biosynthetic and proliferative gene programs while sustaining malignant transcriptional identity in PDAC.

Taken together, these data identify ATXN7L3-AS1, INTS4P2, and AC079921.2 as key lncRNAs that help preserve PDAC transcriptional identity.

Discussion

PDAC exhibits remarkable transcriptional rigidity despite underlying plasticity, resisting lineage reprogramming and maintaining a stable malignant identity. In this study, we performed CRISPRi screens during OSKM-induced reprogramming and identified a subset of cancer-associated lncRNAs as active enforcers of transcriptional memory, which preserve malignant identity and impede reprogramming. Depletion of these lncRNAs, in combination with OSKM, enabled efficient reprogramming of PDAC cells into pluripotent states, resulting in the loss of adenocarcinoma identity and reduced malignant potential both in vitro and in vivo. These findings uncover a previously unrecognized role for cancer-associated lncRNAs as molecular gatekeepers that preserve PDAC identity and constrain cell fate plasticity. Targeting these lncRNAs provides a conceptual and therapeutic framework for erasing cancer cell memory, resetting the epigenetic landscape, and re-sensitizing PDAC cells to differentiation- or reprogramming-based therapies.

Are ATXN7L3-AS1 and AC079921.2 Gatekeepers of PDAC Identity and Plasticity? Cellular memory is epigenetically inherited across cell divisions and is essential for maintaining cell identity. During mitosis, specific transcription factors act as mitotic bookmarks to preserve lineage fidelity[26,27,84]. Similarly, lncRNAs help maintain transcriptional memory and identity by regulating chromatin architecture and gene expression through interactions with chromatin-modifying enzymes.

ATXN7L3-AS1 is located on chromosome 17, upstream of *ATXN7L3* and *UBTF* in the antisense direction (Fig. 7C). *ATXN7L3* encodes a component of the SAGA histone acetylation complex[85], while UBTF is an HMGB-box transcription factor essential for Pol I/II activity and

known to be oncogenic[85]. Motif analysis suggested that UBTF may act as an upstream regulator of genes positively correlated with ATXN7L3-AS1 expression, implying that ATXN7L3-AS1 could reinforce a UBTF-driven transcriptional program. Although UBTF has not been directly implicated in PDAC, it is known oncogenic driver in other cancers[86]. For example, UBTF::ATXN7L3 fusions define a high-risk subtype of B-ALL[87], and UBTF tandem duplications are recurrent in acute myeloid leukemia[88]. Notably, ATXN7L3-AS1 harbors pathogenic somatic variants in PDAC and is associated with mitotic progression in TCGA cancers (Fig. 7C, S6B). These findings support the hypothesis that ATXN7L3-AS1 may contribute to the faithful transmission of malignant identity in PDAC and represents a promising target for disrupting transcriptional memory and malignant persistence.

In pan-cancer analyses, AC079921.2 expression is positively associated with genes involved in ECM remodeling, cell adhesion, and migration—hallmarks of mesenchymal-like tumor states (Fig. S6E). In normal epithelial tissues, these genes are typically repressed by the PRC2, which includes SUZ12 and EZH2, suggesting that AC079921.2 may antagonize PRC2-mediated silencing. Consistent with this model, genes aberrantly upregulated or unresponsive to OSKM reprogramming significantly overlap with AC079921.2–correlated genes in PDAC (Fig. 7D–I), indicating that AC079921.2 sustains a transcriptional program resistant to OSKM-induced reprogramming and thereby stabilizes malignant identity. By maintaining expression of ECM and morphogenetic genes, AC079921.2 may reinforce a mesenchymal-like transcriptional state that supports invasion and adaptability to external cues. Future studies should investigate whether AC079921.2 directly interacts with PRC2 components to modulate their repressive activity in PDAC cells. Targeting this lncRNA could offer a strategy to limit lineage plasticity and prevent undesired cell-state transitions during therapy or reprogramming.

Intriguingly, levels of both ATXN7L3-AS1 and AC079921.2 are negatively correlated with genes involved in translation and metabolic pathways. This suggests that these lncRNAs may help safeguard PDAC identity by repressing biosynthetic and metabolic gene networks that

are otherwise poised for high transcriptional activity, thereby restricting the potential for reprogramming.

Combining lncRNAs Depletion and OSKM Induction Is Required to Erase PDAC Identity

Although individual knockdown of ATXN7L3-AS1, AC079921.2, and INTS4P2 moderately reduced expression of a subset of PDAC marker genes, our results indicate that neither lncRNA depletion nor OSKM induction alone is sufficient to fully erase PDAC identity or suppress malignancy in vivo. In xenograft models, lncRNA depletion without OSKM induction modestly reduced tumor size, likely due to decreased proliferation without increased apoptosis; however, it did not reprogram PDAC cells or erase the adenocarcinoma identity. Likewise, OSKM induction alone produced tumors with a more differentiated histology and reduced proliferation; yet these tumors retained core features of PDAC. In contrast, combined treatment with lncRNA knockdown and OSKM induction (HPAFII-dCas9-gRNA-OSKM-TP cells) resulted in the loss of the original adenocarcinoma identity, reactivation of apoptotic pathways, and emergence of NANOG+ pluripotent stem-like cells. Notably, the re-expression of TSGs in NANOG^{+} pluripotent cells suggests that the loss of PDAC identity is coupled to a transcriptional reset and a fate conversion toward a pluripotency.

Toward a New Reprogramming Cocktail: lncRNA Depletion Sensitizes PDAC to OSKM-Induced Cell Fate Transitions. We previously attempted to reprogram PDAC using OSKM induction alone; however, full reprogramming into a pluripotent state was impeded[15]. To sustain partially reprogrammed states, we maintained low-level ectopic OSKM expression, which enabled us to model the progression of PDAC from a precursor-like to an invasive phenotype upon withdrawal of reprogramming factors[15]. While this model was informative and valuable for biomarker discovery[89], the requirement for persistent ectopic OSKM expression limits its applicability for translational cancer therapies. Thus, we propose that lncRNA depletion

sensitizes PDAC cells to OSKM-mediated reprogramming by weakening the epigenetic barriers that preserve malignant identity and constrain plasticity. To enhance clinical relevance, we optimized a non-integrating, Sendai virus–mediated system for transient OSKM expression across multiple patient-derived PDAC cell lines, offering efficient reprogramming while minimizing the risks associated with genomic integration[59].

Beyond pluripotency, cancer reprogramming into alternative lineages, such as dendritic cells[90], has also been demonstrated. It will be crucial to determine whether lncRNA depletion can similarly facilitate PDAC reprogramming toward alternative or non-permissive lineages, such as acinar cells, which are incompatible with oncogenic KRAS. We anticipate that targeting lncRNAs regulating transcriptional memory and plasticity could enhance PDAC responsiveness to reprogramming cues and improve lineage fidelity.

Altogether, the selective expression of *ATXN7L3-AS1*, *AC079921.2*, and *INTS4P2* in PDAC underscores their potential roles as regulators of transcriptional memory and plasticity, acting as molecular barriers that preserve malignant identity and constrain cell fate transitions.

Limitations of the Study

Our study focuses specifically on lncRNAs that sustain the cancer identity in PDAC and uses OSKM to force reprogramming of PDAC toward pluripotency. However, lncRNAs that sustain the cancer identity may differ across cancer types, and different lineage-specific transcription factors may be required for the reprogramming of other cancer types.

While we observed a robust reduction in malignancy and activation of apoptosis in tumors derived from HPAFII-dCas9-gRNA-OSKM-TP cells, it remains unclear whether the erasure of malignancy is permanent. Long-term xenograft studies will be required to assess the durability of reprogramming-induced tumor suppression.

Conclusions. Our study demonstrates that the combined depletion of cancer-associated lncRNAs and induction of OSKM factors is required to erase the malignant identity of PDAC cells. These findings reveal that cancer-specific lncRNAs act as key enforcers of transcriptional memory, preserving PDAC identity and resisting reprogramming even in the presence of reprogramming cues. By maintaining the epigenetic rigidity of PDAC cells, these lncRNAs may contribute to therapeutic resistance and disease recurrence. Targeting them offers both a conceptual and therapeutic framework for erasing cancer cell memory, resetting the epigenetic landscape, and re-sensitizing PDAC cells to differentiation-based or reprogramming-based interventions.

RESOURCE AVAILABILITY

Lead Contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Jungsun Kim (kimjungs@ohsu.edu).

Materials availability

- Plasmids, PDX tumors, and reprogrammed cancers generated in this study will be available upon request.
- This study did not generate new unique reagents.

Data and code availability

- Raw data and codes for sequencing data have been deposited at NCBI GEO and are publicly available as of publication. Accession numbers will be listed in the key resource table. The lead contact, upon request, will share any raw data reported in this paper.
- Codes for sequencing data have been deposited on GitHub and are publicly available as of publication.
- Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

ACKNOWLEDGMENTS We thank Drs. Ken Zaret and Theresa Lusardi for comments on the manuscript. We thank the patients who donated their tissue for this study and P. Roberts and R. Armstrong for their technical assistance. We thank the Brenden-Colson Center for Pancreatic Care, the Oregon Pancreas Tissue Registry, the OHSU Knight Biostatistics Shared Resource and Biolibrary (NCI P30CA69533), and the Molecular Virology Core at the Oregon National Primate Research Center (NIH P51OD011092). The IRB protocol is the Oregon Pancreas

Tissue Registry (IRB #3609). The work was funded by the CRUK-OHSU Project Award (2018-CRUK-OHSU-001 to JK), MRF New Investigator Grant (GCNCR1042A to JK), and Knight CEDAR (68182-906-000, 68182-933-000, 68182-939-000, 68182-954-000, Manuscript Prep 2024-1978) from the Cancer Early Detection Advanced Research Center at OHSU Knight Cancer Institute. This article was edited by Life Science Editors.

AUTHOR CONTRIBUTIONS

Conceptualization, JK; Methodology, DG, SL, CMF, SR, MB, AH, TE, EM, JT, MCC, JL, and WY; Validation, DG, SL, CMF, SR, SF, TM, and WY; Formal Analysis, DG, SL, CMF, SR, AH, SF, TM, WY, EF, ST, and JK; Investigation, D.G., SL, CMF, SR, MB, SF, TM, WY, EF, ST, and JK.; Resources, JL, DK, BCS, EF, and RCS; Data Curation, DG, SL, CMF, SR, AH, SF, and ST; Writing-original draft, JK; Writing-Review & Editing, DG, SL, and JK; Visualization, DG, CMF, SR, AH, SF, and ST, Supervision, JK; Project Administration, JK; Funding Acquisition, JK.

DECLARATION OF INTERESTS

The authors declare no competing interests.

MAIN FIGURE TITLES and LEGENDS

Figure 1. OSKM-mediated cellular reprogramming downregulates only a subset of PDAC unique genes.

A. Schematic diagram of reprogramming of normal controls (BJ fibroblast and H6c7 hnPDEC) and human classical (CFPAC1, ST-00014490) and basal (Panc1, ST-00007599) subtypes of PDAC cells using Sendai virus-mediated expression of OSKM factors. Cells were transduced with OSKM-expressing Sendai virus on day 0. For transcriptomic analysis, cells were harvested on days 0 and 7. To assess pluripotency and early differentiation, cells were fixed on day 30 and immunostained for TRA-1-60 and SSEA1, respectively. B. Each Venn diagram shows PDAC-specific genes that are upregulated relative to normal controls on day 0 ($\log_2$[PDAC/normal controls] ≥ 1, FDR < 0.05). The heatmap displays the expression profiles of 375 PDAC-specific genes consistently upregulated across all PDAC samples, with Z-scores indicated by the color key. Seven of the 375 PDAC-unique genes were significantly downregulated by OSKM across all PDAC cells ($\log_2$[$OSKM^+$/$OSKM^-$] < –1, FDR < 0.05). C-D. Example of an OSKM-induced downregulated PDAC-specific gene. (C) *SPDEF* mRNA levels in the indicated cells without and with OSKM induction. (D) Kaplan–Meier survival analysis of *SPDEF* expression in the TCGA PDAC cohort (n = 150). Patients were stratified by SPDEF expression quartiles (Low: <7.611; High: >9.675), based on $\log_2$-transformed normalized RSEM values. P-value was calculated using the log-rank test (χ^2, rho = 0). The high *SPDEF* expression in PDAC is associated with poor overall survival of patients with PDAC. E. Identification of aberrantly increased or unchanged genes in PDAC following OSKM induction. The empirical cumulative distribution function (ECDF) displays the cumulative probability distribution of $\log_2$ fold-change values across samples, before and after OSKM induction (day 7 vs. day 0). The X-axis indicates $\log_2$(FC), and the Y-axis indicates the cumulative fraction of genes. The plot is zoomed in to the range of -1 to 1 to highlight subtle shifts in gene expression. Genes with

$\log_2$(FC) > 0.2 or < –0.2 and FDR < 0.05 were defined as DEGs in each sample, based on inspection of the Mean Absolute Deviation (MAD) heatmap as an empirical measure of deviation from the null hypothesis (Fig. S3C; see Methods). Genes were then queried against pre-defined patterns of expression changes across the samples. A weighted Hamming distance was used to measure how well a gene's pattern of expression changes across samples matched with each query (see Methods). For each query, genes with weighted Hamming distances no greater than 0.35 were selected as hits for downstream analysis. F. Violin plots showing the expression levels of genes corresponding to the indicated gene queries. OSKM-unresponsive PDAC genes were defined by the pattern [-1][-1][0][0][0][0] (left, distance < 0.375, n = 1,466; Table S2F), and OSKM-aberrantly upregulated PDAC genes were defined by the queries [-1][-1][0/+1][0/+1][0/+1][0/+1] (middle, distance < 0.25, n=1815; Table S2G) and [-1/0][-1/0][1][1][1][1] (right, distance = 0, n = 82; Table S2H). G-H. Gene Ontology (GO) and upstream regulator analyses for genes that were either unresponsive to OSKM in PDAC (G, Table S2F) or aberrantly upregulated following OSKM induction (H, Table S2G).

Figure 2. CRISPRi against cancer-associated lncRNAs significantly improves the efficiency of reprogramming PDAC cells. A. Schematic diagram of reprogramming HPAFII PDAC and H6c7 hnPDEC cells by CRISPRi screening. The reprogramming efficiency was measured by immunostaining for TRA-1-60 and SSEA1. The TRA-1-60+ pluripotent stem cell-like cells were quantified and sorted by flow cytometry at 30 days. B. Comparison of the percentage of TRA-1-60$^+$ cells between HPAFII-dCas9 control cells and HPAFII-dCas9 cells expressing gRNAs targeting lncRNAs, measured by flow cytometry on day 30 after OSKM induction. *indicates significant differences in the % of cells between TRA-1-60$^-$ (91.2%) and TRA-1-60$^+$ (8.8%) expressing cells in HPAFII-dCas9 (**q=1.60E-05, n=4,* Unpaired Welch t-test), whereas there were no significant differences between TRA-1-60$^-$ and TRA-1-60$^+$ cells in HPAFII-dCas9/gRNA

cells. C. Representative images and quantification summary of TRA-1-60^{+} and SSEA1^{+} cells on plates after multiple passages (Scale bar: 284 μm).

Figure 3. Suppression of PDAC gene programs and activation of pluripotency and tumor suppressor signatures in HPAFII-dCas9-gRNA-OSKM-TP cells. A. Schematic diagram of the scRNA-seq workflow. B-J. UMAP visualizations showing distinct clusters by sample (B), 15 Louvain-defined clusters (C), as well as cells expressing BFP (D), pancreatic ductal marker KRT19 (E), *PODXL* encoding TRA-1-60 (F), *NANOG* (G), pluripotency-associated genes (H, Table S5A), endodermal and PDAC-related genes (I, Table S5B), and ectodermal/mesodermal lineage markers (J, Table S5C). Each dot represents a cell, and the color intensity indicates the normalized UMI counts (scaled by the total UMI per cell). K. Principal component analysis (PCA) of HPAFII-dCas9-gRNA-OSKM-TP cells alongside publicly available scRNA-seq datasets from human PDAC (n=6, GSE212966[74]), human and mouse normal pancreas (n=4, GSE84133[75]), human early embryo before the three germ layers formed (D6-D14, n=555, GSE136447[76]), and iPSC derived from Parkinson's disease patients' fibroblast (n=12, GSE183248[77]). L. Volcano plots of DEGs between dCas9-gRNA and either dCas9-gRNA-OSKM-TN (left) or HPAFII-dCas9-gRNA-OSKM-TP (right). PDAC-specific signature genes are annotated. M. Heatmap showing expression of 318 TSGs across the indicated cell populations (Table S4D). The plot shows z-scores of the mean TSG expression across cells in the sample. The distinct subsets show that each sample has a set of TSGs more highly expressed in that sample than in others. N-O. Heatmap (N) and quantification (O) of TSG expression in the dCas9-gRNA-OSKM-TP cells. Each column represents an individual cell, each row corresponds to a TSG, and expression levels are shown as Z-scores normalized across all cells for each gene (N). Expression levels of TSGs in each cluster are shown across the indicated subpopulations; the Y-axis indicates Z-score (O).

Figure 4. Reduced tumor size in xenografts derived from HPAFII-dCas9-gRNA-TP cells.

A. Schematic diagram of xenograft assay of TRA-1-60^{+} and TRA-1-60^{-}cells derived from OSKM-reprogrammed HPAFII-dCas9/gRNA, along with controls. Tumors were harvested from NSG mice one month after being subcutaneously injected with dCas9-gRNA-OSKM-TN and dCas9-gRNA-OSKM-TP, along with controls: Parental HPAFII, huES, dCas9-gRNA, dCas9-OSKM-TP (n=4-8 for each cell type). B. Representative images of tumors from two independent experiments; two to four mice for each cell type were used in each experiment. C. Comparison of tumor size among different cell types. The differences in tumor sizes (width x height) were tested pairwise using a non-parametric Mann-Whitney test. Only tumor sizes between parental HPAFII and dCas9-gRNA-OSKM-TP are significantly different (*** p=0.0002, n=8 each, Mann-Whitney test). D-H. Representative H&E images of the indicated xenografted tumor samples (D; HPAFII, E; dCas9-gRNA, F; dCas9-OSKM-TP, G; dCas9-gRNA-OSKM-TN, H; dCas9-gRNA-OSKM-TP) from three independent mice. Mouse identifiers (e.g., #1-1, #1-2, #2-1, #2-2) denote individual animals associated with each cell type. The scale bar indicates 50 μm.

Figure 5. Loss of ductal identity and reactivation of apoptotic markers in xenograft tumors derived from PDAC cells reprogrammed with OSKM and lncRNA depletion.

Representative images of TRA-1-85 human specific marker, human CK19 pancreatic ductal marker, Cleaved Caspase-3 apoptotic marker, and Ki67 proliferation marker staining in xenograft tumors derived from parental HPAFII cells (A–D), dCas9-gRNA cells (E–H), dCas9-OSKM-TP cells (I–L), dCas9-gRNA-OSKM-TN cells (M–P), and dCas9-gRNA-OSKM-TP cells (Q–T). Mouse identifiers (e.g., #1-1, #1-2, #2-1, #2-2) denote individual animals associated with each cell type. Arrows indicate positive brown signals. Scale bar: 50 μm.

Figure 6. Emergence of ATXN7L3-AS1, INTS4P2, and AC079921 lncRNAs as PDAC cell identity factors.

A. Schematic flow chart of CRISPRi screening B. Significantly enriched lncRNAs in HPAFII-dCas9-gRNA-TP cells in comparison to the HPAFII-dCas9-gRNA reference sample are visualized in the volcano plot. The enrichment score of lncRNAs is shown as the "β scores." Positive and negative β scores indicate positive and negative selections in the TRA-1-60$^+$ compared to the reference condition, respectively. The X-axis shows the β-score of TRA-1-60$^+$ compared to the reference. The Y-axis displays –log10 (p-values). The p-value is calculated by ranking genes with a skewed distribution compared to a uniform null model and permuting to calculate the significance of the skew in distribution. β-score and p-values have been calculated separately from two independent experiments and combined to display the top-ranked lncRNAs in one volcano plot. The horizontal and vertical lines indicate the cutoff threshold. The red dots indicate lncRNAs with a β-score over zero and a permutation p-value <0.05. The candidate lncRNAs that were validated by qRT-PCR were highlighted with names. C. The expression of candidate lncRNAs in normal pancreatic cells and cancers, and HPAFII-dCas9-gRNA-OSKM-TN and TP cells (n=3, y-axis indicates relative expression over GAPDH, error bars=SD). D-F. Expression level of candidate lncRNAs in human normal tissues (GTEx). Expression values are in TPM (Transcripts per Million), calculated from a gene model with isoforms collapsed to a single gene. The Y-axis shows $\log_{10}$(TPM+1). Box plots are shown as median and 25th and 75th percentiles; points are displayed as outliers if they are above or below 1.5 times the interquartile range. Pancreatic tissues (n = 362) were included in the study.

Figure 7. Cancer-Enriched lncRNAs Preserve PDAC Identity, Map to Somatically Mutated Loci, and Correlate with Pro-Invasive Transcriptional Programs. A-B. Expression of basal subtype markers (KRT5, TP63), classical ductal TFs (GATA6, HNF1B, ONECUT1), and pancreatic progenitor markers (PDX1, FOXA2, and SOX9) in HPAFII (A) and H6c7 (B) cells following CRISPRi-mediated knockdown of ATXN7L3-AS1, AC079921.2, or INTS4P2. Y-axis indicates relative expression normalized to GAPDH and scaled to non-targeting controls (NT); error bars represent SD. C. UCSC Genome Browser view of the ATXN7L3-AS1 locus (chr17:44,198,841–44,234,868; hg38), located between the ATXN7L3 and UBTF genes. Pathogenic somatic variants identified in PDAC from the COSMIC database are displayed along the locus. D-E. Overlap between AC079921.2–positively correlated genes across pan-TCGA tumors and OSKM-aberrantly upregulated (D) or unresponsive (E) genes in PDAC reprogramming. Hypergeometric test was used for significance. F-G. GO analysis of overlapping genes from (D) and (E), enriched for ECM organization, muscle contraction, cell adhesion, and migration pathways. H-I. Upstream regulator analysis highlights SUZ12 and chromatin-associated complexes as candidate regulators of AC079921.2–associated transcriptional programs.

SUPPLEMENTAL FIGURE TITLES and LEGENDS

Figure S1 (Relevant to Figure 1). Transgene-free reprogramming of human normal fibroblast and pancreatic acinar cells into pluripotency using Sendai-virus mediated expression of OSKM master TFs. A. Immunostaining of pluripotency markers in iPSCs established from BJ6 human normal fibroblast cells. B. Verification of expression of pancreatic acinar marker amylase and lack of pancreatic ductal marker K19 on pancreatic acinar cells using immunofluorescence (IF). C. Schematic diagram of reprogramming pancreatic acinar cells into iPSC using Sendai-OSKM. Initial numbers of cells and emerged TRA-1-60^{+} cells after 26 days were counted to determine the reprogramming efficiency. Three clones were established after five passages. D.

Confirmation of expression of pluripotency surface markers TRA-1-60 and SSEA4 in established acinar iPS lines through flow cytometry. E. Representative image of cytogenetic analysis on 20 G-banded metaphase cells derived from pancreatic acinar-derived iPSC. F. Confirmation of functional pluripotency of acinar-derived iPSC clones. Their embryoid bodies (EB) differentiated into three germ layer lineages, which were verified by staining for three germ layer markers using flow cytometry.

Figure S2 (Relevant to Figure 1). Reprogramming of human normal pancreatic ductal and cancer cells. A. Representative images of colonies stained for TRA-1-60 and SSEA-1 in indicated cells. Scale bars indicate 100 μm. B. Expression of OSKM, pluripotency, and PDAC genes in the indicated cells before and 7 days after OSKM induction.

Figure S3 (Relevant to Figures 1-2). A-F. RNA-seq of classical (CFPAC1 and ST-00014490) and basal (Panc1 and ST-00007599) subtypes of PDAC and normal controls (human BJ6 fibroblast and H6c7 hnPDECs) before and 7 days after OSKM induction. (A) PCA plot. (B) Proportion of genes changed upon OSKM induction among a total of 302,383 genes in normal cells (BJ, H6c7) and PDAC (CFPAC1, ST-00014490, Panc1, and ST-00007599). A chi-squared test was used to compare the proportion of genes exhibiting expression changes following OSKM induction in PDAC cells relative to other normal cell types. The Y-axis indicates the proportion of DEGs relative to the total number of genes. *** indicates p-value $< 2.2 \times 10^{-13}$. Black and blue bars indicate comparison between PDAC cells and BJ or H6c7 cells, respectively (See Table S2D-E). (C) Mean Absolute Deviation (MAD) heatmap of Q–Q statistics across FDR cut-off points and $\log_2$ (FC) truncated thresholds. The heatmap displays the MAD between observed Q–Q statistics and a truncated standard normal distribution (N(0,1)), computed across combinations of FDR cutoffs (x-axis) and $\log_2$ FC thresholds (y-axis). The color key indicates MAD of Q-Q (stat) vs truncated N (0,1). Lower MAD values (dark blue) indicate better agreement with the expected null distribution, suggesting more reliable gene selection. Red contours mark regions of increasing deviation. White horizontal and vertical lines indicate $\log_2$(FC) = 0.2 and FDR 0.05 thresholds, respectively. (D) Density and Empirical Cumulative Distribution Function (ECDF) plots of log2 FC values across all samples. Red dashed lines indicate |log2 FC|=0.2. (E) Intersection of aberrantly upregulated PDAC genes (downregulated in normal but upregulated in PDAC) and Suz12 target genes in human and mouse ES cells (hypergeometric p-value). F. Expressions of genes encoding PRC2 components (SUZ12 and EZH2) and REST

transcriptional repressor in indicated cells before and after OSKM induction. G-H. CRISPRi against cancer-associated lncRNAs. (G) Schematic diagram of KRAB-dCas9-Blastidin and gRNA-EF1a-Puro2a-BFP vectors. (H) Flow cytometry analysis of TRA-1-60$^+$ cells from H6c7 cells on day 30 after OSKM induction, along with lncRNA depletion.

Figure S4 (Relevant to Figure 3) scRNAseq of HPAFII-dCas9-gRNA-OSKM-TP cells along with controls. A. Quality control density plots showing the distribution of cells by key metrics: percentage of mitochondrial gene expression, complexity (features per count), total transcripts per cell, and number of detected genes per cell. B. Expression levels of *BFP*, *OSKM*, and *NANOG* across samples (dCas9, dCas9-gRNA, dCas9-gRNA-OSKM-TN, dCas9-gRNA-OSKM-TP, dCas0-OSKM-TP). The X-axis represents individual samples, and the Y-axis shows normalized UMI counts. C. Top GO terms enriched in genes from Group 0 and Group 1 clusters. D. Top GO terms associated with DEGs in dCas9-gRNA cells compared to parental dCas9 cells. E. Violin plots showing normalized UMI counts for genes in cells from groups 0, 1, and 5 across samples. F. PCA projection of each sample.

Figure S5 (Relevant to Figure 4-6) A-E. Xenograft of huES cells; Representative images of immunostaining of H&E (A), human-specific marker TRA-1-85 (B), human ductal marker CK19 (C), apoptotic marker Cleaved Caspase-3 (D), and proliferation marker Ki67 (E). F-I. Identification of candidate gRNAs reproducible from two independent experiments. Detailed diagram of selecting candidate lncRNAs (F). Z-score of top 20 gRNAs (G). The estimated β-score plots of TRA-1-60$^+$ and TRA-1-60$^-$ cells compared to references (HPAFII-dCas9-gRNA). The β scores of all gRNAs are normalized based on the negative control normalization (H). Volcano plot of top-ranked lncRNAs showing β scores enriched in TRA-1-60$^+$ cells compared to the reference control. The Y-axis displays $-\log_{10}$ (p-values). The p-value is calculated by ranking genes with a skewed distribution compared to a uniform null model and permuting to

calculate the significance of the skew in distribution. The X-axis indicates the β scores in TRA-1-60^{+}. Red indicates positive enrichment (beta score >0) and blue indicates negative enrichment (beta score <0) (I).

Figure S6 (Relevant to Figure 6-7) Association of ATXN7L3-AS1 with mitotic cell cycle progression and AC079921.2 with cell adhesion and migration in the TCGA dataset. A. Expression of candidate lncRNAs before and after OSKM induction. B-D. Top GO terms (B), potential upstream regulators (C), and histone modifications (D) enriched in the upstream of genes whose expressions are positively correlated with ATXN7L3-AS1 mRNA levels in TCGA Pan-cancer (Spearman correlation >0.4, n=2887). E-G. Top GO terms of (E), potential upstream regulators (F), and histone modifications of the upstream (G) of genes whose expressions are positively associated with AC079921.2 in TCGA Pan-cancer (Spearman correlation >0.4, 2305). H-M. Top GO terms of (H, I), histone modification of the upstream (J, K), potential upstream regulators (L, M) in of genes whose expressions are negatively associated with ATXN7L3-AS1 and AC079921.2 in TCGA Pan-cancer, respectively (Spearman correlation <-0.4, ATXN7L3-AS1, n=1928; AC079921.2, n=667).

SUPPLEMENTAL TABLE TITLES and LEGENDS

Table S1. *KRAS* genotyping in iPS cells.

Table S2. *Related to Figure 1. OSKM-mediated reprogramming is not sufficient to override PDAC identity.*

- **Table S2A.** Regularized $\log_2$-transformed DESeq2 expression values for 375 PDAC-unique genes.
- **Table S2B.** Regularized $\log_2$-transformed DESeq2 expression values for 7 PDAC-unique genes regulated by OSKM induction.
- **Table S2C.** Gene ontology (GO) analysis of 7 PDAC-unique genes downregulated by OSKM induction.
- **Table S2D.** Chi-square test comparing the proportion of genes with expression changes in pooled PDAC vs. normal cells following OSKM induction.
- **Table S2E.** Chi-square test comparing the proportion of genes with expression changes in each PDAC and normal sample following OSKM induction.
- **Table S2F.** List of OSKM-unresponsive PDAC genes (signature: [-1][-1][0][0][0][0], outliers removed, n=1,466, distance < 0.375; related to Fig. 1F).
- **Table S2G.** List of OSKM-aberrantly upregulated PDAC genes (signature: [-1][-1][0/+1][0/+1][0/+1][0/+1], outliers removed, n=1,815, distance < 0.25; related to Fig. 1F).
- **Table S2H.** List of strongly OSKM-aberrantly upregulated PDAC genes (signature: [-1/0][-1/0][1][1][1][1], n=82, distance = 0; related to Fig. 1F).

Table S3. *Related to Figure 2. Depletion of cancer-associated lncRNAs enhances reprogramming of PDAC.*

- **Table S3A.** List of gRNAs targeting cancer-associated lncRNAs (CRISPRIL17_hg38).
- **Table S3B.** Quantification of TRA-1-60$^+$ and TRA-1-60$^-$ cells by flow cytometry.

Table S4. *Related to Figure 3. Single-cell RNA-seq reveals transcriptional reprogramming toward pluripotency, suppression of PDAC identity, and reactivation of tumor suppressor genes (TSGs).*

- **Table S4A.** List of pluripotency marker genes.
- **Table S4B.** List of endodermal and PDAC marker genes.
- **Table S4C.** List of ectodermal and mesodermal marker genes.
- **Table S4D.** List of OncoKB-annotated tumor suppressor genes (total: n=379; excluding oncogenes: n=318).

Table S5. *Related to Figure 6. Identification of enriched gRNAs in HPAFII-dCas9-gRNA-OSKM-TP cells.*

- **Table S5A.** Sample barcodes and Illumina adaptor sequences.
- **Table S5B.** Read counts aligned to gRNAs (Experiment 1).
- **Table S5C.** Read counts aligned to gRNAs (Experiment 2).
- **Table S5D.** Raw hit counts (Experiment 1).
- **Table S5E.** Raw hit counts (Experiment 2).
- **Table S5F.** Z-scores of each gRNA (Experiment 1).
- **Table S5G.** Z-scores of each gRNA (Experiment 2).
- **Table S5H.** MAGeCK-MLE output (aggregated reference x500, Experiment 1).
- **Table S5I.** MAGeCK-MLE output (Experiment 2).
- **Table S5J.** MAGeCK-MLE output (combined Experiment 1 and 2).
- **Table S5K.** Final list of selected candidate gRNAs.

Table S6. Related to Figure 7. COSMIC Pathogenic somatic variants in ATXN7L3-AS1 (GRCh38).

- Table S6A. Pathogenic somatic variants in ATXN7L3-AS1 identified in pan-cancers
- Table S6B. Pathogenic somatic variants in ATXN7L3-AS1 identified in pancreatic cancer

STAR METHODS

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Derivation of Patient-Derived Xenografts (PDX) and Tissue Culture. Human PDAC specimens were obtained under the Oregon Pancreas Tissue Registry study (IRB00003609). Informed consent was obtained from all subjects. All experimental protocols were approved by the OHSU Institutional Review Board. All methods were carried out in accordance with relevant guidelines and regulations. All animal works for PDX tumors were performed with the OHSU Institutional Animal Use and Care Committee (IACUC) approval. The derivation, expansion, culture, and characterization of PDAC PDXs, are described in our protocol paper[57]. Culture conditions for H6c7 and other PDAC cell lines are described[57]. BJ6 fibroblasts were cultured in DMEM supplemented with 10% fetal bovine serum (FBS)[65].

Reprogramming of Pancreatic Cancer Cells and Normal Controls. Reprogramming was performed using the CytoTune-iPS Sendai Reprogramming Kit (Invitrogen), which delivers the non-integrating Sendai virus encoding OCT4, SOX2, KLF4, and c-MYC (OSKM)[91]. Viral titers for Sendai virus and lentivirus were described in our earlier publication[59]. On day 7 post-transduction, cells were harvested for RNA sequencing and replated onto mitomycin C-treated mouse embryonic fibroblasts (MEFs) to support further reprogramming.

METHOD DETAILS

Tissue immunohistochemistry (IHC), cell immunostaining, antibodies, qRT-PCR, and Primers. Tissue IHC, immunofluorescence staining, and qRT-PCR were performed as previously described in our published work[15].

QUANTIFICATION AND STATISTICAL ANALYSIS

RNA-seq Data Processing and Normalization. Bulk RNA-seq was performed as previously described[59,92]. Raw FASTQ files were processed using fastp to generate clean reads by removing adapter sequences, reads containing poly-N, and low-quality reads. During this step, quality metrics such as Q20, Q30, and GC content were also calculated. All downstream analyses were conducted using the high-quality clean reads. The reference genome and gene annotation files were downloaded directly from the genome database. The genome index was built using HISAT2 (v2.0.5), and paired-end clean reads were aligned to the reference genome using HISAT2 (v2.0.597)[93]. FeatureCounts (v1.5.0-p3) was used to quantify the number of reads mapped to each gene. FPKM (Fragments Per Kilobase of transcript per Million mapped reads) values were then calculated for each gene based on its length and the number of mapped reads. Differential expression analysis was performed using the DESeq2 R package (v1.20.0), which employs statistical routines based on the negative binomial distribution to identify differentially expressed genes from count data[94]. P-values were adjusted for multiple testing using the Benjamini-Hochberg method to control FDR. Genes with an adjusted P-value ≤ 0.05 were considered significantly differentially expressed. RNA-seq data is available on NCBI GSExxxx.

To identify PDAC-specific gene categories, we first performed differential expression analysis using DESeq2 to compare each PDAC cell line to normal controls (BJ fibroblasts and H6c7 pancreatic epithelial cells). Genes that were significantly upregulated in PDAC cells (FDR < 0.05, $\log_2$ FC > 1) were selected. We then intersected the upregulated gene sets across all four PDAC cell lines to define a core set of PDAC-upregulated genes. To assess the impact of OSKM induction, we compared gene expression profiles before and after OSKM treatment and identified differentially expressed genes using the same thresholds (FDR < 0.05, $\log_2$FC> 1 or < –1).

Pattern Classification Following OSKM and Gene Ranking by Weighted Hamming Distance. DEGs before and after OSKM induction (day 7 vs. day 0) were identified for each cell type using the LRT implemented in DESeq2[94]. To establish robust thresholds for DEG calling, we evaluated how closely observed DEG statistics matched the expected null hypothesis distribution. Specifically, we constructed empirical quantile-quantile (Q-Q) plots comparing DEG statistics from genes presumed to be non-differentially expressed (those with simultaneously high adjusted p-values and low log2(FC)) against their expected null distribution. We applied Huber regression to these Q-Q plots and measured the mean absolute deviation (MAD) between the fitted regression line and the identity line over the range of observed statistics. Higher MAD values indicate greater divergence from the null distribution, reflecting genuine biological signals. MAD values were computed across a range of adjusted p-values and log2(FC) threshold combinations, and visualized as a heatmap. Based on this analysis, genes with log2(FC) > 0.2 and FDR < 0.05 were considered significantly differentially expressed, as these thresholds corresponded to a clear elevation in MAD.

To classify gene expression responses to OSKM induction, we analyzed DEGs across six cell types—two normal controls and four PDAC cell lines ([normal#1][normal#2][cancer#1][cancer#2][cancer#3][cancer#4])—using six predefined expression patterns: i) Downregulated in normal, unchanged in cancer: [-1][-1][0][0][0][0], ii) Downregulated in normal, upregulated in cancer: [-1][-1][1][1][1][1], iii) Downregulated in normal, upregulated or unchanged in cancer: [-1][-1][0/+1][0/+1][0/+1][0/+1], iv) Unchanged in normal, upregulated in cancer: [0][0][1][1][1][1], v) Upregulated in normal, downregulated in cancer: [1][1][-1][-1][-1][-1], vi) Upregulated in normal, unchanged in cancer: [1][1][0][0][0][0]. These patterns were used to stratify gene responses to OSKM and identify cancer-specific regulatory behaviors. We classified genes in Class I ([-1][-1][0][0][0][0]) as OSKM-unresponsive genes, and genes in Class II ([-1][-1][1][1][1][1]), III ([-1][-1][0/+1][0/+1][0/+1][0/+1]), and IV [0][0][1][1][1][1] as OSKM-aberrantly upregulated genes (Table S2B-D).

To rank genes within each predefined expression pattern category, we used a weighted Hamming distance metric[68] relative to the query vector. Let q denote the query vector representing a specific regulation profile (e.g., upregulated in normal with minimal change in cancer samples), and let cat(S) denote the categorical expression vector of a gene. The weighted Hamming distance d is calculated as: $d = \sum(i=1 \text{ to } N)\ w_i \times \text{Ind}(q_i \neq \text{cat}(S)_i)$ where N is the number of conditions or replicates, w_i is the weight assigned to the *i*th condition, and Ind is the indicator function that returns 1 if the query and gene vector differ at position *i*, and 0 otherwise. This metric penalizes deviations from the target pattern, with weights chosen to balance representation across biological conditions and to reflect the hierarchical structure of replicates within each condition. Genes with lower values of the weighted Hamming distances are considered better matches to the specified query. For each query of interest, genes were ranked in ascending order based on their weighted Hamming distance.

Gene Ontology (GO) and Upstream Analysis. GO analysis was performed using Enrichment Analysis[95] or the ToppGene Suite[96]. To determine whether gene sets of interest are enriched for putative TF targets, we performed ChIP-X enrichment analysis using datasets from ENCODE or ChEA[97,98], accessed through Enrichr[99]or X2K[100]. ChIP-X refers to genome-wide TF binding data obtained from ChIP-ChIP, ChIP-Seq, ChIP-PET, or DamID experiments[97,98] . The p-value is computed using Fisher's exact test or the hypergeometric test. This is a binomial proportion test that assumes a binomial distribution and independence for the probability of any gene belonging to any set. The q-value is an adjusted p-value using the Benjamini-Hochberg method for correction for multiple hypothesis testing. The combined score is calculated by taking the logarithm of the p-value from the Fisher exact test and multiplying it by the Z-score, which reflects the deviation from the expected rank. The odds ratio (used for rank-based enrichment) is derived by performing the Fisher exact test across numerous random gene sets to estimate

the mean and standard deviation of the expected rank for each term in the gene-set library. A Z-score is then computed to quantify the deviation of the observed rank from expectation.

Single-cell RNAseq analysis. Single-cell RNA sequencing (scRNA-seq) was performed using the Chromium Next GEM Single Cell 5' Kit (v2.0, 10x Genomics). All scRNA-seq samples were processed using Seurat v5. Following quality control, cells were filtered, normalized, scaled, batch-corrected, and clustered in accordance with best practices outlined in the Seurat v5 vignette[101]. Cells expressing gRNA were filtered for BFP. Clustering was performed using the Seurat integration workflow with the CCA (Canonical Correlation Analysis) method, at a resolution of 0.5 using the Louvain algorithm. Harmony integration was also performed to assess potential differences compared to the Seurat CCA-based integration. scRNAseq data set was extracted from human PDAC (n=6, GSE212966[74]), human and mouse normal pancreas (n=4, GSE84133[75]), human early embryo before the three germ layers formed (D6-D14, n=555, GSE136447[76]), and iPSC derived from Parkinson's patients' fibroblast (n=12, GSE183248[77]). NT clusters were manually defined based on NANOG and TRA-1-60 expression profiles. Cells were grouped according to their expression of these markers, with clusters curated based on predefined thresholds rather than unsupervised clustering algorithms.

CRISPRi screen and next-generation sequencing. The lentivirus expressing a sgRNA library with 7087 different sgRNAs against cancer common lncRNAs and dCas9 protein was obtained through Millipore[72,73]. The HPAFII and H6c7 cells expressing dCas9 were generated by transducing HPAFII and H6c7 cells with lentiviruses carrying the dCas9-Bsr gene at an MOI of 1, followed by selecting cells for Blastidin. The CRISPRi efficiency was assessed in the clonal HPAFII and H6c7 cells expressing dCa9 with positive controls. The clonal HPAFII-dCas9 and H6c7-dCas9 were subsequently transduced with a lentiviral library of 7087 sgRNAs at an MOI of 0.1-0.3 with 300- 500x coverage to maintain an average 300-500-fold representation of each

sgRNA. After puromycin selection, cells were collected for the reference controls (300x, 500x). HPAFII and H6c7 cells were reprogrammed with OSKM Sendai viruses using the OSKM Sendai viruses[91]. The CRISPRi-targeted cells were always maintained at 300-500 coverage throughout the procedure. Technical replicate genomic DNA isolated from each group was subjected to nested PCR amplification to create NGS libraries, identifying enriched or depleted sgRNAs in TRA-1-60+ reprogrammed cells compared to reference controls or TRA-1-60^{-}. The hit count corresponding to each sgRNA was computed from raw FastQ files using Millipore's custom pipelines. The raw hit counts were normalized by Z-score for exploratory analysis. The MAGeCK package with negative control sgRNA normalization was used to analyze the enriched lncRNAs[82]. The MAGeCK algorithm first performs normalization and then uses information borrowing across genes to estimate variance[82]. A negative binomial model is then fit and used to test for differential hits on genes across samples. Significance testing uses permutations. We used the MAGeCK mle module, since there were more than two conditions. The MAGeCKFlute R package was used with R version 4.4 for data analysis and visualization[102]. We selected gRNAs with Z-scores > 2 and positive beta scores ($p < 0.05$) from two independently passaged cultures (separated by 3-6 passages) (Table S5F-I). To ensure robustness, raw gRNA counts from both experiments were merged, scaling the second dataset to correct for batch differences (Exp1/Exp2 ≈ 1.86), with counts rounded up to the nearest integer (Table S5J). gRNAs consistently enriched in both cultures, as defined by either Z-score or MAGeCK analysis, were retained (Table S5K; Fig. S5A).

Identification of Disease-Associated Somatic Variants. Disease-associated somatic variant data for ATXN7L3-AS1 were collected from the COSMIC[83] and ClinVar[103] databases via LncBook 2.0[104]. We retained variants labeled as “confirmed somatic mutations” in COSMIC. Variants were classified as pathogenic if they were annotated as having a significant impact on

disease in ClinVar or if they had a FATHMM-MKL score greater than 0.7, a threshold previously shown to predict pathogenic potential of noncoding variants[105] (Table S6A-B).

ADDITIONAL RESOURCES. In vivo amplification and characterization of patient-derived xenograft (PDX) tumors have been described in STAR Protocols[57]. Detailed protocols for transduction of PDAC cells and subtype characterization of PDX tumors were published in Heliyon[59].

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
Human TRA-1-85/CD147 Antibody	R&D Systems	Cat # MAB3195
PE Mouse anti-Human TRA-1-85 Antigen	BD Biosciences	Cat # 563021
human Cytokeratin 19 (Abcam, BA-17)	Abcam	Cat # ab7755
Ki67	Cell Signaling Technology,	Cat # 9449
Cleaved Caspase-3	Cell Signaling Technology	Cat # 12692
FITC conjugate, Anti-TRA-1-60 Antibody, clone TRA-1-60, FITC conjugate	Millipore-Sigma	FCMAB115F
TRA-1-60 Alexa Fluor™ 488 Conjugate Kit for Live Cell Imaging	Thermo Scientific	Cat # A25618
Human/Mouse SSEA-1 Alexa Fluor« 594-conjugated Antibody	R&D systems	IC2155T-050UG
GloLIVE Human/Mouse SSEA-1 NL557 Live Cell Antibody	R&D Systems	Cat # NLLC2155R
Sendai virus HN Monoclonal Antibody (1A6), Alexa Fluor™ 647, eBioscience™	Invitrogen	Cat # 51-6494-82
OCT4 Antibody (OTI9B7) [Alexa Fluor« 405]	Novus Biologicals	NBP2-71225AF405
Anti-Oct4 antibody - ChIP Grade	Abcam	Cat# ab19857
Biological samples		
CF-1 MEFS 2 MILLION	Thermo Fisher Scientific	A34180
Dr4 Mouse Embryonic Fibroblasts, irradiated	Thermo Fisher Scientific	A34966
Chemicals, peptides, and recombinant proteins		
Iscript™ Advanced cDNA Synthesis Kit	Bio-Rad	1725037
RNase-Free DNase Set (50)	Qiagen	79254
Applied Biosystems TaqMan Fast Advanced Master Mix	Thermo Fisher Scientific	44-445-57
SPRIselect, Liquid	Beckman Coulter	B23318
Keratinocyte-SFM (1X)	Invitrogen	17005042
Primocin™	Thermo Fisher Scientific	ant-pm-05
Agilent RNA 6000 Pico Reagents	Agilent Technologies	5067-1514

ACCUMAX™ Cell Detachment Solution, Sterile, Liquid	Thermo Fisher Scientific	07921
InSolution™ Y-27632	Calbiochem	688001-500UG
Epidermal Growth Factor, human, animal component free,	Millipore-Sigma	E5036
Gibco fgf Basic (aa 10 155) Recombinant Human Protein	Invitrogen	PHG0021
Knockout™ Serum Replacement	Invitrogen	10828-028
L Glutamine (200mM)	Invitrogen	25030081
Eagle's Minimum Essential Medium (EMEM)	ATCC	30-2003
0.1% Gelatin in Water	STEMCELL Technologies	07903
Embryonic Stem Cell FBS, Qualified, One Shot™ Format, US Origin	Thermo Fisher Scientific	16141002
Human TruStain FcX™ (Fc Receptor Blocking Solution) 50 tests	Bio Legend	422301
Cholera Toxin from Vibrio cholerae	Thermo Fisher Scientific	C8052-.5MG
Blasticidin, Sterile, ≥95%, 10 mg/mL, Liquid	InvivoGen	ANTBL05
Puromycin, 100 mg (10 x 1 ml tubes)	InvivoGen	ANTPR1
DEFINED KERATINOCYTE-SFM, COMBO COMBINATION	Invitrogen	10744019
Protein Free, Thermo Scientific, Pierce™ Protein-free T20 (PBS) blocking buffer	Pierce	PI37573
Bovine Pituitary Extract (BPE)	Invitrogen	13028014
PowerTrack™ SYBR Green Master Mix	Applied Biosystems	A46109
Collagen I Rat Protein, Tail	Gibco	A1048301
FIX VIABILITY STAIN 660 100UG	Bd Biosciences	BDB564405
Cytotune™ EmGFP Sendai Fluorescence Reporter	Life Science	A16519
Critical commercial assays		
CytoTune™-iPS 2.0 Sendai Reprogramming Kit	*Invitrogen*	A16518
Human CRISPRi Common Cancer Cell Lines Kit	Millipore-Sigma	CRISPRIS10
Human CRISPR Control Kit	Millipore	CRISPRICON-1KT
QIAamp UCP DNA Micro Kit	Qiagen	56204
Lenti-X™ Provirus Quantitation Kit	Takara	631239
Deposited data		
Raw and analyzed data	This paper	GEO: GSE247243
Table S3A. List of gRNAs targeting cancer-associated lncRNAs (CRISPRIL17_hg38).	This paper, Supplementary Table	Table S3A.
Experimental models: Cell lines		
Human: BJ fibroblast cells	ATCC	CRL-2522
Human: H6c7 normal pancreatic ductal cells	KeraFast	ECA001-FP
Human: hTERT-HPNE,	ATCC	CRL-4023
Human: Passage 40 H1 and H9 ES cells	Wicell	WA01
Human: HPAFII PDAC cells	ATCC	CRL-1997
Human: Panc1 PDAC cells	ATCC	CRL-1469
Human: CFPAC1 cells	ATCC	CRL-1918

Human: PDAC cells from PDX (ST-00014490)	Star Protocols[57]	n.a.
Human: PDAC cells from PDX (ST-00007599)	Star Protocols[57]	n.a.
Experimental models: Organisms/strains		
Mouse: NOD-SCID-IL2Rgc null (NSG) (4-6 weeks old, female/male)	The Jackson Laboratory	JAX: #005557
Oligonucleotides		
Primer: AC079921.2 #2	IDT	Forward: AGAGCATGGGGTG AAAGACC Reverse: AGAGACTTTCTCAG GGTGGC
Primer: INTS4P2 #1:	IDT	Forward: AGCAGCAAAACAG ACTTTACCC Reverse: TGTCCACAAAGCT GTCCTGC
PrimerINTS4P2 #2	IDT	Forward: AGATACGGACCCT GGAGCAT Reverse: GAACCACCCAAGA GCCATCA
RAB1A	Life Technologies	Hs00366313_m1
GAPDH	Life Technologies	Hs02758991_g1
GATA6	Life Technologies	Hs00232018_m1
SOX9	Life Technologies	Hs00165814_m1
KRT19	Life Technologies	Hs00761767_s1
Foxa2	Life Technologies	Hs05036278_s1
Hnf1b	Life Technologies	Hs01001602_m1
Onecut1	Life Technologies	Hs00413554_m1
FUT4	Life Technologies	Hs01106466_s1
MNX1	Life Technologies	Hs00907365_m1
MNX1	Life Technologies	Hs03454395_m1
PDX1	Life Technologies	Hs00426216_m1
Sendai-CMYC	Life Technologies	Mr04269876_mr
Sendai-OCT3-4	Life Technologies	Mr04269878_mr
Sendai-KLF4	Life Technologies	Mr04421256_mr
SEV-SOX2	Life Technologies	Mr04269881_mr
Sendai	Life Technologies	Mr04269880_mr
Recombinant DNA		
Software and algorithms		

DESEQ2	Love et al.,[94]	https://bioconductor.org/packages/devel/bioc/vignettes/DESeq2/inst/doc/DESeq2.html
MAGeCK	Li et al., 2014[82] and Wang e al., 2019[102]	https://github.com/liulab-dfci/MAGeCK
Other		

REFERENCE

1. Halbrook, C.J., Lyssiotis, C.A., Pasca di Magliano, M., and Maitra, A. (2023). Pancreatic cancer: Advances and challenges. Cell *186*, 1729-1754. 10.1016/j.cell.2023.02.014.
2. Moffitt, R.A., Marayati, R., Flate, E.L., Volmar, K.E., Loeza, S.G., Hoadley, K.A., Rashid, N.U., Williams, L.A., Eaton, S.C., Chung, A.H., et al. (2015). Virtual microdissection identifies distinct tumor- and stroma-specific subtypes of pancreatic ductal adenocarcinoma. Nat Genet *47*, 1168-1178. 10.1038/ng.3398.
3. Bailey, P., Chang, D.K., Nones, K., Johns, A.L., Patch, A.M., Gingras, M.C., Miller, D.K., Christ, A.N., Bruxner, T.J., Quinn, M.C., et al. (2016). Genomic analyses identify molecular subtypes of pancreatic cancer. Nature *531*, 47-52. 10.1038/nature16965.
4. Chan-Seng-Yue, M., Kim, J.C., Wilson, G.W., Ng, K., Figueroa, E.F., O'Kane, G.M., Connor, A.A., Denroche, R.E., Grant, R.C., McLeod, J., et al. (2020). Transcription phenotypes of pancreatic cancer are driven by genomic events during tumor evolution. Nat Genet *52*, 231-240. 10.1038/s41588-019-0566-9.
5. Raghavan, S., Winter, P.S., Navia, A.W., Williams, H.L., DenAdel, A., Lowder, K.E., Galvez-Reyes, J., Kalekar, R.L., Mulugeta, N., Kapner, K.S., et al. (2021). Microenvironment drives cell state, plasticity, and drug response in pancreatic cancer. Cell *184*, 6119-6137 e6126. 10.1016/j.cell.2021.11.017.
6. Burdziak, C., Alonso-Curbelo, D., Walle, T., Reyes, J., Barriga, F.M., Haviv, D., Xie, Y., Zhao, Z., Zhao, C.J., Chen, H.A., et al. (2023). Epigenetic plasticity cooperates with cell-cell interactions to direct pancreatic tumorigenesis. Science *380*, eadd5327. 10.1126/science.add5327.
7. Baba, T., Finetti, P., Pancreatic Cancer, G., Lillemoe, K.D., Warshaw, A.L., Fernandez-Del Castillo, C., and Liss, A.S. (2022). A Lesson in Transcriptional Plasticity: Classical Identity Is Silenced, but Not Lost, in Pancreatic Ductal Adenocarcinoma Cell Lines. Gastroenterology *163*, 1450-1453 e1453. 10.1053/j.gastro.2022.07.005.
8. Singhal, A., Styers, H.C., Rub, J., Li, Z., Torborg, S.R., Kim, J.Y., Grbovic-Huezo, O., Feng, H., Tarcan, Z.C., Sahin Ozkan, H., et al. (2024). A Classical Epithelial State Drives Acute Resistance to KRAS Inhibition in Pancreatic Cancer. Cancer Discov *14*, 2122-2134. 10.1158/2159-8290.CD-24-0740.
9. Vierbuchen, T., Ostermeier, A., Pang, Z.P., Kokubu, Y., Sudhof, T.C., and Wernig, M. (2010). Direct conversion of fibroblasts to functional neurons by defined factors. Nature *463*, 1035-1041. 10.1038/nature08797.
10. Laiosa, C.V., Stadtfeld, M., Xie, H., de Andres-Aguayo, L., and Graf, T. (2006). Reprogramming of committed T cell progenitors to macrophages and dendritic cells by C/EBP alpha and PU.1 transcription factors. Immunity *25*, 731-744. 10.1016/j.immuni.2006.09.011.
11. Xie, H., Ye, M., Feng, R., and Graf, T. (2004). Stepwise reprogramming of B cells into macrophages. Cell *117*, 663-676. 10.1016/s0092-8674(04)00419-2.
12. Davis, R.L., Weintraub, H., and Lassar, A.B. (1987). Expression of a single transfected cDNA converts fibroblasts to myoblasts. Cell *51*, 987-1000. 10.1016/0092-8674(87)90585-x.
13. Takahashi, K., Tanabe, K., Ohnuki, M., Narita, M., Ichisaka, T., Tomoda, K., and Yamanaka, S. (2007). Induction of pluripotent stem cells from adult human fibroblasts by defined factors. Cell *131*, 861-872. 10.1016/j.cell.2007.11.019.
14. 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. 10.1016/j.cell.2006.07.024.
15. Kim, J., Hoffman, J.P., Alpaugh, R.K., Rhimm, A.D., Reichert, M., Stanger, B.Z., Furth, E.E., Sepulveda, A., Yuan, C.X., Won, K.J., et al. (2013). An iPSC Line from Human Pancreatic Ductal

Adenocarcinoma Undergoes Early to Invasive Stages of Pancreatic Cancer Progression. Cell Reports *3*, 2088-2099. 10.1016/j.celrep.2013.05.036.

16. Zhang, X., Cruz, F.D., Terry, M., Remotti, F., and Matushansky, I. (2013). Terminal differentiation and loss of tumorigenicity of human cancers via pluripotency-based reprogramming. Oncogene *32*, 2249-2260, 2260 e2241-2221. 10.1038/onc.2012.237.
17. Bernhardt, M., Novak, D., Assenov, Y., Orouji, E., Knappe, N., Weina, K., Reith, M., Larribere, L., Gebhardt, C., Plass, C., et al. (2017). Melanoma-Derived iPCCs Show Differential Tumorigenicity and Therapy Response. Stem Cell Reports *8*, 1379-1391. 10.1016/j.stemcr.2017.03.007.
18. Papapetrou, E.P. (2016). Patient-derived induced pluripotent stem cells in cancer research and precision oncology. Nat Med *22*, 1392-1401. 10.1038/nm.4238.
19. Kim, J., and Zaret, K.S. (2015). Reprogramming of human cancer cells to pluripotency for models of cancer progression. EMBO Journal *34*, 739-747. 10.15252/embj.201490736.
20. Davidson, E.H., and Erwin, D.H. (2006). Gene regulatory networks and the evolution of animal body plans. Science *311*, 796-800. 10.1126/science.1113832.
21. Shukla, V., Cetnarowska, A., Hyldahl, M., and Mandrup, S. (2022). Interplay between regulatory elements and chromatin topology in cellular lineage determination. Trends Genet *38*, 1048-1061. 10.1016/j.tig.2022.05.011.
22. Almeida, N., Chung, M.W.H., Drudi, E.M., Engquist, E.N., Hamrud, E., Isaacson, A., Tsang, V.S.K., Watt, F.M., and Spagnoli, F.M. (2021). Employing core regulatory circuits to define cell identity. EMBO J *40*, e106785. 10.15252/embj.2020106785.
23. Bulyk, M.L., Drouin, J., Harrison, M.M., Taipale, J., and Zaret, K.S. (2023). Pioneer factors - key regulators of chromatin and gene expression. Nat Rev Genet. 10.1038/s41576-023-00648-z.
24. Zaret, K.S., and Mango, S.E. (2016). Pioneer transcription factors, chromatin dynamics, and cell fate control. Curr Opin Genet Dev *37*, 76-81. 10.1016/j.gde.2015.12.003.
25. Boyer, L.A., Lee, T.I., Cole, M.F., Johnstone, S.E., Levine, S.S., Zucker, J.P., Guenther, M.G., Kumar, R.M., Murray, H.L., Jenner, R.G., et al. (2005). Core transcriptional regulatory circuitry in human embryonic stem cells. Cell *122*, 947-956. 10.1016/j.cell.2005.08.020.
26. Caravaca, J.M., Donahue, G., Becker, J.S., He, X., Vinson, C., and Zaret, K.S. (2013). Bookmarking by specific and nonspecific binding of FoxA1 pioneer factor to mitotic chromosomes. Genes Dev *27*, 251-260. 10.1101/gad.206458.112.
27. Zhu, Z., Chen, X., Guo, A., Manzano, T., Walsh, P.J., Wills, K.M., Halliburton, R., Radko-Juettner, S., Carter, R.D., Partridge, J.F., et al. (2023). Mitotic bookmarking by SWI/SNF subunits. Nature *618*, 180-187. 10.1038/s41586-023-06085-6.
28. Kim, K., Doi, A., Wen, B., Ng, K., Zhao, R., Cahan, P., Kim, J., Aryee, M.J., Ji, H., Ehrlich, L.I., et al. (2010). Epigenetic memory in induced pluripotent stem cells. Nature *467*, 285-290. 10.1038/nature09342.
29. Polo, J.M., Liu, S., Figueroa, M.E., Kulalert, W., Eminli, S., Tan, K.Y., Apostolou, E., Stadtfeld, M., Li, Y., Shioda, T., et al. (2010). Cell type of origin influences the molecular and functional properties of mouse induced pluripotent stem cells. Nat Biotechnol *28*, 848-855. 10.1038/nbt.1667.
30. Bar-Nur, O., Russ, H.A., Efrat, S., and Benvenisty, N. (2011). Epigenetic memory and preferential lineage-specific differentiation in induced pluripotent stem cells derived from human pancreatic islet beta cells. Cell Stem Cell *9*, 17-23. 10.1016/j.stem.2011.06.007.
31. Ohi, Y., Qin, H., Hong, C., Blouin, L., Polo, J.M., Guo, T., Qi, Z., Downey, S.L., Manos, P.D., Rossi, D.J., et al. (2011). Incomplete DNA methylation underlies a transcriptional memory of somatic cells in human iPS cells. Nat Cell Biol *13*, 541-549. 10.1038/ncb2239.
32. Lister, R., Pelizzola, M., Kida, Y.S., Hawkins, R.D., Nery, J.R., Hon, G., Antosiewicz-Bourget, J., O'Malley, R., Castanon, R., Klugman, S., et al. (2011). Hotspots of aberrant epigenomic

reprogramming in human induced pluripotent stem cells. Nature *471*, 68-73. 10.1038/nature09798.

33. Buckberry, S., Liu, X., Poppe, D., Tan, J.P., Sun, G., Chen, J., Nguyen, T.V., de Mendoza, A., Pflueger, J., Frazer, T., et al. (2023). Transient naive reprogramming corrects hiPS cells functionally and epigenetically. Nature *620*, 863-872. 10.1038/s41586-023-06424-7.
34. Hu, Q., Friedrich, A.M., Johnson, L.V., and Clegg, D.O. (2010). Memory in induced pluripotent stem cells: reprogrammed human retinal-pigmented epithelial cells show tendency for spontaneous redifferentiation. Stem Cells *28*, 1981-1991. 10.1002/stem.531.
35. Thiagalingam, S. (2020). Epigenetic memory in development and disease: Unraveling the mechanism. Biochim Biophys Acta Rev Cancer *1873*, 188349. 10.1016/j.bbcan.2020.188349.
36. Cheloufi, S., Elling, U., Hopfgartner, B., Jung, Y.L., Murn, J., Ninova, M., Hubmann, M., Badeaux, A.I., Euong Ang, C., Tenen, D., et al. (2015). The histone chaperone CAF-1 safeguards somatic cell identity. Nature *528*, 218-224. 10.1038/nature15749.
37. Nefzger, C.M., Rossello, F.J., Chen, J., Liu, X., Knaupp, A.S., Firas, J., Paynter, J.M., Pflueger, J., Buckberry, S., Lim, S.M., et al. (2017). Cell Type of Origin Dictates the Route to Pluripotency. Cell Rep *21*, 2649-2660. 10.1016/j.celrep.2017.11.029.
38. Morris, S.A., Cahan, P., Li, H., Zhao, A.M., San Roman, A.K., Shivdasani, R.A., Collins, J.J., and Daley, G.Q. (2014). Dissecting engineered cell types and enhancing cell fate conversion via CellNet. Cell *158*, 889-902. 10.1016/j.cell.2014.07.021.
39. Ghosh, Z., Wilson, K.D., Wu, Y., Hu, S., Quertermous, T., and Wu, J.C. (2010). Persistent donor cell gene expression among human induced pluripotent stem cells contributes to differences with human embryonic stem cells. PLoS One *5*, e8975. 10.1371/journal.pone.0008975.
40. Mattick, J.S., Amaral, P.P., Carninci, P., Carpenter, S., Chang, H.Y., Chen, L.L., Chen, R., Dean, C., Dinger, M.E., Fitzgerald, K.A., et al. (2023). Long non-coding RNAs: definitions, functions, challenges and recommendations. Nat Rev Mol Cell Biol *24*, 430-447. 10.1038/s41580-022-00566-8.
41. Zaratiegui, M., Irvine, D.V., and Martienssen, R.A. (2007). Noncoding RNAs and gene silencing. Cell *128*, 763-776. 10.1016/j.cell.2007.02.016.
42. Flynn, R.A., and Chang, H.Y. (2014). Long noncoding RNAs in cell-fate programming and reprogramming. Cell Stem Cell *14*, 752-761. 10.1016/j.stem.2014.05.014.
43. Li, W., Notani, D., Ma, Q., Tanasa, B., Nunez, E., Chen, A.Y., Merkurjev, D., Zhang, J., Ohgi, K., Song, X., et al. (2013). Functional roles of enhancer RNAs for oestrogen-dependent transcriptional activation. Nature *498*, 516-520. 10.1038/nature12210.
44. Kim, Y.J., Xie, P., Cao, L., Zhang, M.Q., and Kim, T.H. (2018). Global transcriptional activity dynamics reveal functional enhancer RNAs. Genome Res *28*, 1799-1811. 10.1101/gr.233486.117.
45. Kaikkonen, M.U., Lam, M.T., and Glass, C.K. (2011). Non-coding RNAs as regulators of gene expression and epigenetics. Cardiovasc Res *90*, 430-440. 10.1093/cvr/cvr097.
46. Jiao, W., Chen, Y., Song, H., Li, D., Mei, H., Yang, F., Fang, E., Wang, X., Huang, K., Zheng, L., and Tong, Q. (2018). HPSE enhancer RNA promotes cancer progression through driving chromatin looping and regulating hnRNPU/p300/EGR1/HPSE axis. Oncogene *37*, 2728-2745. 10.1038/s41388-018-0128-0.
47. Statello, L., Guo, C.J., Chen, L.L., and Huarte, M. (2021). Gene regulation by long non-coding RNAs and its biological functions. Nat Rev Mol Cell Biol *22*, 96-118. 10.1038/s41580-020-00315-9.
48. Guttman, M., and Rinn, J.L. (2012). Modular regulatory principles of large non-coding RNAs. Nature *482*, 339-346. 10.1038/nature10887.

49. Ng, S.Y., Johnson, R., and Stanton, L.W. (2012). Human long non-coding RNAs promote pluripotency and neuronal differentiation by association with chromatin modifiers and transcription factors. EMBO J *31*, 522-533. 10.1038/emboj.2011.459.
50. Khalil, A.M., Guttman, M., Huarte, M., Garber, M., Raj, A., Rivea Morales, D., Thomas, K., Presser, A., Bernstein, B.E., van Oudenaarden, A., et al. (2009). Many human large intergenic noncoding RNAs associate with chromatin-modifying complexes and affect gene expression. Proc Natl Acad Sci U S A *106*, 11667-11672. 10.1073/pnas.0904715106.
51. Guttman, M., Donaghey, J., Carey, B.W., Garber, M., Grenier, J.K., Munson, G., Young, G., Lucas, A.B., Ach, R., Bruhn, L., et al. (2011). lincRNAs act in the circuitry controlling pluripotency and differentiation. Nature *477*, 295-300. 10.1038/nature10398.
52. Wang, Y., Xu, Z., Jiang, J., Xu, C., Kang, J., Xiao, L., Wu, M., Xiong, J., Guo, X., and Liu, H. (2013). Endogenous miRNA sponge lincRNA-RoR regulates Oct4, Nanog, and Sox2 in human embryonic stem cell self-renewal. Dev Cell *25*, 69-80. 10.1016/j.devcel.2013.03.002.
53. Wang, Z.W., Pan, J.J., Hu, J.F., Zhang, J.Q., Huang, L., Huang, Y., Liao, C.Y., Yang, C., Chen, Z.W., Wang, Y.D., et al. (2022). SRSF3-mediated regulation of N6-methyladenosine modification-related lncRNA ANRIL splicing promotes resistance of pancreatic cancer to gemcitabine. Cell Rep *39*, 110813. 10.1016/j.celrep.2022.110813.
54. Ye, M., Dong, S., Hou, H., Zhang, T., and Shen, M. (2021). Oncogenic Role of Long Noncoding RNAMALAT1 in Thyroid Cancer Progression through Regulation of the miR-204/IGF2BP2/m6A-MYC Signaling. Mol Ther Nucleic Acids *23*, 1-12. 10.1016/j.omtn.2020.09.023.
55. Jiang, M.C., Ni, J.J., Cui, W.Y., Wang, B.Y., and Zhuo, W. (2019). Emerging roles of lncRNA in cancer and therapeutic opportunities. Am J Cancer Res *9*, 1354-1366.
56. Arunkumar, G., Baek, S., Sturgill, D., Bui, M., and Dalal, Y. (2022). Oncogenic lncRNAs alter epigenetic memory at a fragile chromosomal site in human cancer cells. Sci Adv *8*, eabl5621. 10.1126/sciadv.abl5621.
57. Tessier, J., Grygoryev, D., Chow-Castro, M., Ekstrom, T., Manalo, E., Lee, S.W., Berry, M., Link, J.M., Keith, D., Allen-Petersen, B.L., et al. (2025). Protocol to purify and culture human pancreatic cancer cells from patient-derived xenografts. STAR Protoc *6*, 103672. 10.1016/j.xpro.2025.103672.
58. Ban, H., Nishishita, N., Fusaki, N., Tabata, T., Saeki, K., Shikamura, M., Takada, N., Inoue, M., Hasegawa, M., Kawamata, S., and Nishikawa, S. (2011). Efficient generation of transgene-free human induced pluripotent stem cells (iPSCs) by temperature-sensitive Sendai virus vectors. Proc Natl Acad Sci U S A *108*, 14234-14239. 10.1073/pnas.1103509108.
59. Grygoryev, D., Ekstrom, T., Manalo, E., Link, J.M., Alshaikh, A., Keith, D., Allen-Petersen, B.L., Sheppard, B., Morgan, T., Soufi, A., et al. (2024). Sendai virus is robust and consistent in delivering genes into human pancreatic cancer cells. Heliyon *10*, e27221. 10.1016/j.heliyon.2024.e27221.
60. Cacchiarelli, D., Trapnell, C., Ziller, M.J., Soumillon, M., Cesana, M., Karnik, R., Donaghey, J., Smith, Z.D., Ratanasirintrawoot, S., Zhang, X., et al. (2015). Integrative Analyses of Human Reprogramming Reveal Dynamic Nature of Induced Pluripotency. Cell *162*, 412-424. 10.1016/j.cell.2015.06.016.
61. Takahashi, K., Tanabe, K., Ohnuki, M., Narita, M., Sasaki, A., Yamamoto, M., Nakamura, M., Sutou, K., Osafune, K., and Yamanaka, S. (2014). Induction of pluripotency in human somatic cells via a transient state resembling primitive streak-like mesendoderm. Nat Commun *5*, 3678. 10.1038/ncomms4678.
62. Buganim, Y., Faddah, D.A., Cheng, A.W., Itskovich, E., Markoulaki, S., Ganz, K., Klemm, S.L., van Oudenaarden, A., and Jaenisch, R. (2012). Single-cell expression analyses during cellular

reprogramming reveal an early stochastic and a late hierarchic phase. Cell *150*, 1209-1222. 10.1016/j.cell.2012.08.023.

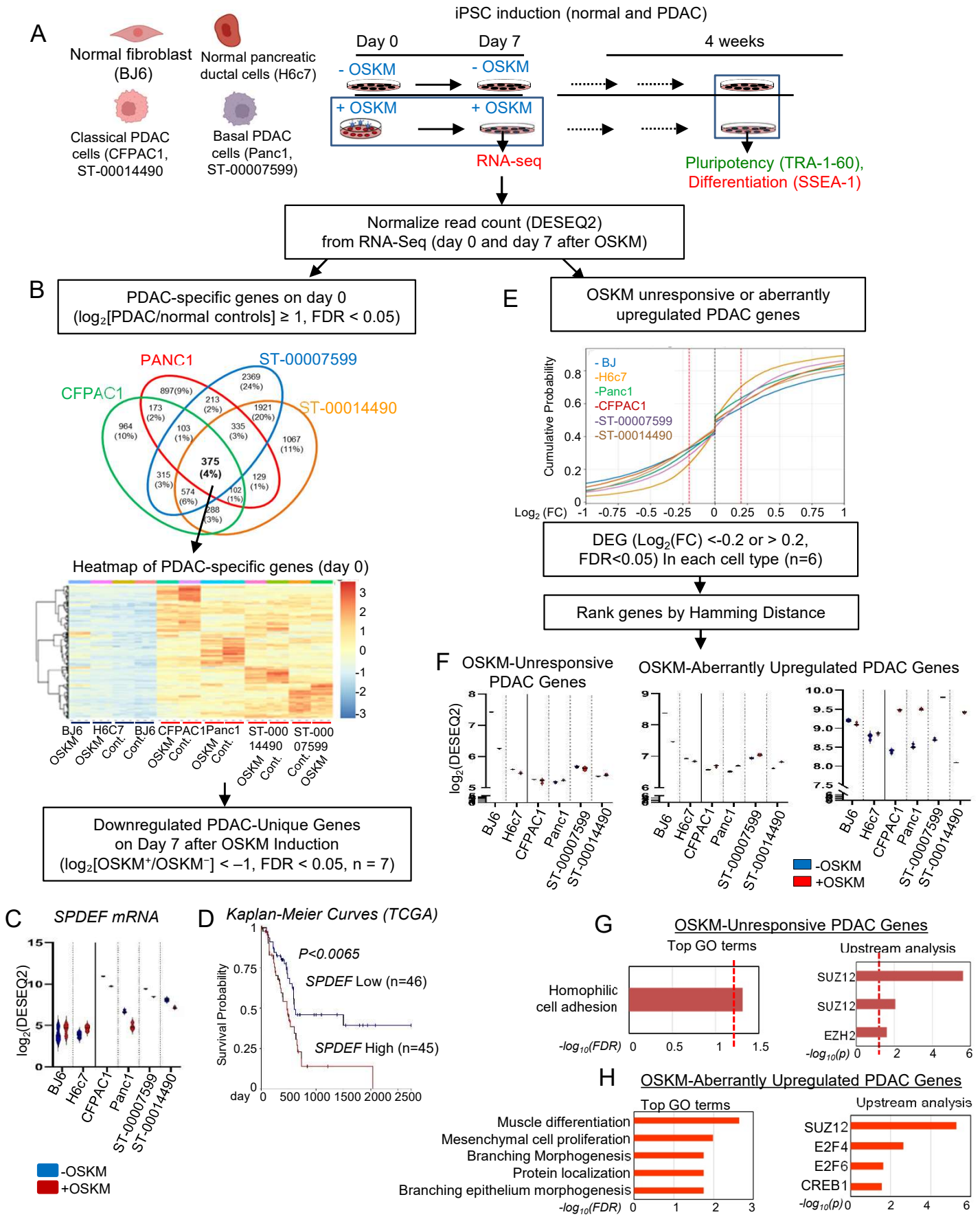
63. Knaupp, A.S., Buckberry, S., Pflueger, J., Lim, S.M., Ford, E., Larcombe, M.R., Rossello, F.J., de Mendoza, A., Alaei, S., Firas, J., et al. (2017). Transient and Permanent Reconfiguration of Chromatin and Transcription Factor Occupancy Drive Reprogramming. Cell Stem Cell *21*, 834-845 e836. 10.1016/j.stem.2017.11.007.
64. 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 e420. 10.1016/j.cell.2016.12.016.
65. Soufi, A., Donahue, G., and Zaret, K.S. (2012). Facilitators and impediments of the pluripotency reprogramming factors' initial engagement with the genome. Cell *151*, 994-1004. 10.1016/j.cell.2012.09.045.
66. Li, D., Liu, J., Yang, X., Zhou, C., Guo, J., Wu, C., Qin, Y., Guo, L., He, J., Yu, S., et al. (2017). Chromatin Accessibility Dynamics during iPSC Reprogramming. Cell Stem Cell *21*, 819-833 e816. 10.1016/j.stem.2017.10.012.
67. Tonelli, C., Yordanov, G.N., Hao, Y., Deschenes, A., Hinds, J., Belleau, P., Klingbeil, O., Brosnan, E., Doshi, A., Park, Y., et al. (2024). A mucus production programme promotes classical pancreatic ductal adenocarcinoma. Gut *73*, 941-954. 10.1136/gutjnl-2023-329839.
68. Hamming, R.W. (1950). Error detecting and error correcting codes. The Bell System Technical Journal *29*, 147-160. 10.1002/j.1538-7305.1950.tb00463.x.
69. Petracovici, A., and Bonasio, R. (2021). Distinct PRC2 subunits regulate maintenance and establishment of Polycomb repression during differentiation. Mol Cell *81*, 2625-2639 e2625. 10.1016/j.molcel.2021.03.038.
70. Zhao, J., Ohsumi, T.K., Kung, J.T., Ogawa, Y., Grau, D.J., Sarma, K., Song, J.J., Kingston, R.E., Borowsky, M., and Lee, J.T. (2010). Genome-wide identification of polycomb-associated RNAs by RIP-seq. Mol Cell *40*, 939-953. 10.1016/j.molcel.2010.12.011.
71. Yang, F., Deng, X., Ma, W., Berletch, J.B., Rabaia, N., Wei, G., Moore, J.M., Filippova, G.N., Xu, J., Liu, Y., et al. (2015). The lncRNA Firre anchors the inactive X chromosome to the nucleolus by binding CTCF and maintains H3K27me3 methylation. Genome Biol *16*, 52. 10.1186/s13059-015-0618-0.
72. Liu, S.J., Horlbeck, M.A., Cho, S.W., Birk, H.S., Malatesta, M., He, D., Attenello, F.J., Villalta, J.E., Cho, M.Y., Chen, Y., et al. (2017). CRISPRi-based genome-scale identification of functional long noncoding RNA loci in human cells. Science *355*. 10.1126/science.aah7111.
73. Horlbeck, M.A., Gilbert, L.A., Villalta, J.E., Adamson, B., Pak, R.A., Chen, Y., Fields, A.P., Park, C.Y., Corn, J.E., Kampmann, M., and Weissman, J.S. (2016). Compact and highly active next-generation libraries for CRISPR-mediated gene repression and activation. Elife *5*. 10.7554/eLife.19760.
74. Chen, J., Liu, Z., Wu, Z., Li, W., and Tan, X. (2023). Identification of a chemoresistance-related prognostic gene signature by comprehensive analysis and experimental validation in pancreatic cancer. Front Oncol *13*, 1132424. 10.3389/fonc.2023.1132424.
75. Baron, M., Veres, A., Wolock, S.L., Faust, A.L., Gaujoux, R., Vetere, A., Ryu, J.H., Wagner, B.K., Shen-Orr, S.S., Klein, A.M., et al. (2016). A Single-Cell Transcriptomic Map of the Human and Mouse Pancreas Reveals Inter- and Intra-cell Population Structure. Cell Syst *3*, 346-360 e344. 10.1016/j.cels.2016.08.011.
76. Xiang, L., Yin, Y., Zheng, Y., Ma, Y., Li, Y., Zhao, Z., Guo, J., Ai, Z., Niu, Y., Duan, K., et al. (2020). A developmental landscape of 3D-cultured human pre-gastrulation embryos. Nature *577*, 537-542. 10.1038/s41586-019-1875-y.

77. Novak, G., Kyriakis, D., Grzyb, K., Bernini, M., Rodius, S., Dittmar, G., Finkbeiner, S., and Skupin, A. (2022). Single-cell transcriptomics of human iPSC differentiation dynamics reveal a core molecular network of Parkinson's disease. Commun Biol *5*, 49. 10.1038/s42003-021-02973-7.
78. Suehnholz, S.P., Nissan, M.H., Zhang, H., Kundra, R., Nandakumar, S., Lu, C., Carrero, S., Dhaneshwar, A., Fernandez, N., Xu, B.W., et al. (2024). Quantifying the Expanding Landscape of Clinical Actionability for Patients with Cancer. Cancer Discov *14*, 49-65. 10.1158/2159-8290.CD-23-0467.
79. Chakravarty, D., Gao, J., Phillips, S.M., Kundra, R., Zhang, H., Wang, J., Rudolph, J.E., Yaeger, R., Soumerai, T., Nissan, M.H., et al. (2017). OncoKB: A Precision Oncology Knowledge Base. JCO Precis Oncol *2017*. 10.1200/PO.17.00011.
80. Avior, Y., Biancotti, Juan C., and Benvenisty, N. (2015). TeratoScore: Assessing the Differentiation Potential of Human Pluripotent Stem Cells by Quantitative Expression Analysis of Teratomas. Stem Cell Reports *4*, 967-974. https://doi.org/10.1016/j.stemcr.2015.05.006.
81. Thomson, J.A., Itskovitz-Eldor, J., Shapiro, S.S., Waknitz, M.A., Swiergiel, J.J., Marshall, V.S., and Jones, J.M. (1998). Embryonic stem cell lines derived from human blastocysts. Science *282*, 1145-1147. 10.1126/science.282.5391.1145.
82. Li, W., Xu, H., Xiao, T., Cong, L., Love, M.I., Zhang, F., Irizarry, R.A., Liu, J.S., Brown, M., and Liu, X.S. (2014). MAGeCK enables robust identification of essential genes from genome-scale CRISPR/Cas9 knockout screens. Genome Biol *15*, 554. 10.1186/s13059-014-0554-4.
83. Sondka, Z., Dhir, N.B., Carvalho-Silva, D., Jupe, S., Madhumita, McLaren, K., Starkey, M., Ward, S., Wilding, J., Ahmed, M., et al. (2024). COSMIC: a curated database of somatic variants and clinical data for cancer. Nucleic Acids Res *52*, D1210-D1217. 10.1093/nar/gkad986.
84. Palozola, K.C., Donahue, G., Liu, H., Grant, G.R., Becker, J.S., Cote, A., Yu, H., Raj, A., and Zaret, K.S. (2017). Mitotic transcription and waves of gene reactivation during mitotic exit. Science *358*, 119-122. 10.1126/science.aal4671.
85. Li, W., Atanassov, B.S., Lan, X., Mohan, R.D., Swanson, S.K., Farria, A.T., Florens, L., Washburn, M.P., Workman, J.L., and Dent, S.Y.R. (2016). Cytoplasmic ATXN7L3B Interferes with Nuclear Functions of the SAGA Deubiquitinase Module. Mol Cell Biol *36*, 2855-2866. 10.1128/MCB.00193-16.
86. Zhang, J., Zhang, J., Liu, W., Ge, R., Gao, T., Tian, Q., Mu, X., Zhao, L., and Li, X. (2021). UBTF facilitates melanoma progression via modulating MEK1/2-ERK1/2 signalling pathways by promoting GIT1 transcription. Cancer Cell International *21*, 543. 10.1186/s12935-021-02237-8.
87. Passet, M., Kim, R., Gachet, S., Sigaux, F., Chaumeil, J., Galland, A., Sexton, T., Quentin, S., Hernandez, L., Larcher, L., et al. (2022). Concurrent CDX2 cis-deregulation and UBTF::ATXN7L3 fusion define a novel high-risk subtype of B-cell ALL. Blood *139*, 3505-3518. 10.1182/blood.2021014723.
88. Umeda, M., Ma, J., Huang, B.J., Hagiwara, K., Westover, T., Abdelhamed, S., Barajas, J.M., Thomas, M.E., Walsh, M.P., Song, G., et al. (2022). Integrated Genomic Analysis Identifies UBTF Tandem Duplications as a Recurrent Lesion in Pediatric Acute Myeloid Leukemia. Blood Cancer Discov *3*, 194-207. 10.1158/2643-3230.BCD-21-0160.
89. Kim, J., Bamlet, W.R., Oberg, A.L., Chaffee, K.G., Donahue, G., Cao, X.-j., Chari, S., Garcia, B.A., Petersen, G.M., and Zaret, K.S. (2017). Detection of early pancreatic ductal adenocarcinoma with thrombospondin-2 and CA19-9 blood markers. Science translational medicine *9*, 1-13.
90. Ascic, E., Akerstrom, F., Sreekumar Nair, M., Rosa, A., Kurochkin, I., Zimmermannova, O., Catena, X., Rotankova, N., Veser, C., Rudnik, M., et al. (2024). In vivo dendritic cell reprogramming for cancer immunotherapy. Science *386*, eadn9083. 10.1126/science.adn9083.
91. Fusaki, N., Ban, H., Nishiyama, A., Saeki, K., and Hasegawa, M. (2009). Efficient induction of transgene-free human pluripotent stem cells using a vector based on Sendai virus, an RNA virus

that does not integrate into the host genome. Proc Jpn Acad Ser B Phys Biol Sci *85*, 348-362. 10.2183/pjab.85.348.

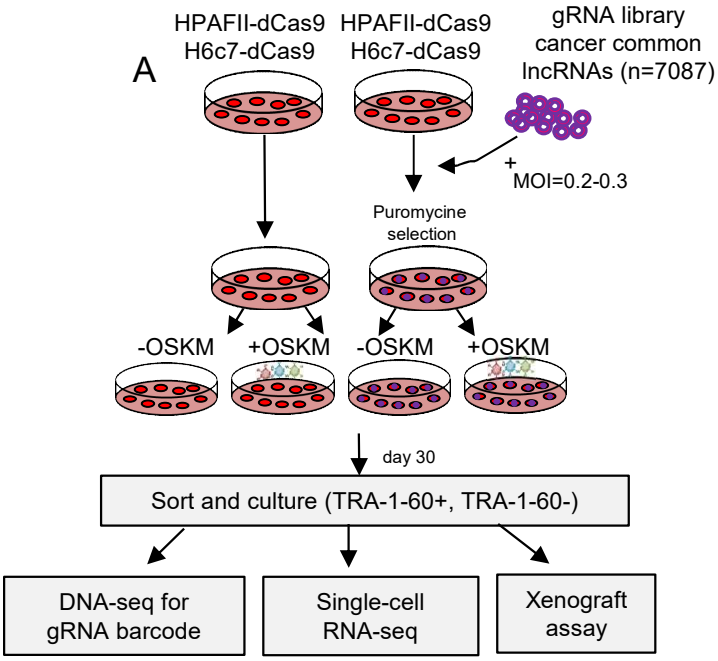
92. Kim, J., Ekstrom, T., Yang, W., Donahue, G., Grygoryev, D., Ngo, T., Muschler, J.L., Morgan, T., and Zaret, K.S. (2021). Longitudinal analysis of human pancreatic adenocarcinoma development reveals transient gene expression signatures. Mol Cancer Res. 10.1158/1541-7786.MCR-21-0483.
93. Kim, D., Paggi, J.M., Park, C., Bennett, C., and Salzberg, S.L. (2019). Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. Nature Biotechnology *37*, 907-915. 10.1038/s41587-019-0201-4.
94. Love, M.I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol *15*, 550. 10.1186/s13059-014-0550-8.
95. Thomas, P.D., Ebert, D., Muruganujan, A., Mushayahama, T., Albou, L.P., and Mi, H. (2022). PANTHER: Making genome-scale phylogenetics accessible to all. Protein Sci *31*, 8-22. 10.1002/pro.4218.
96. Chen, J., Bardes, E.E., Aronow, B.J., and Jegga, A.G. (2009). ToppGene Suite for gene list enrichment analysis and candidate gene prioritization. Nucleic Acids Res *37*, W305-311. 10.1093/nar/gkp427.
97. Lachmann, A., Xu, H., Krishnan, J., Berger, S.I., Mazloom, A.R., and Ma'ayan, A. (2010). ChEA: transcription factor regulation inferred from integrating genome-wide ChIP-X experiments. Bioinformatics *26*, 2438-2444. 10.1093/bioinformatics/btq466.
98. Keenan, A.B., Torre, D., Lachmann, A., Leong, A.K., Wojciechowicz, M.L., Utti, V., Jagodnik, K.M., Kropiwnicki, E., Wang, Z., and Ma'ayan, A. (2019). ChEA3: transcription factor enrichment analysis by orthogonal omics integration. Nucleic Acids Res *47*, W212-W224. 10.1093/nar/gkz446.
99. Xie, Z., Bailey, A., Kuleshov, M.V., Clarke, D.J.B., Evangelista, J.E., Jenkins, S.L., Lachmann, A., Wojciechowicz, M.L., Kropiwnicki, E., Jagodnik, K.M., et al. (2021). Gene Set Knowledge Discovery with Enrichr. Curr Protoc *1*, e90. 10.1002/cpz1.90.
100. Clarke, D.J.B., Kuleshov, M.V., Schilder, B.M., Torre, D., Duffy, M.E., Keenan, A.B., Lachmann, A., Feldmann, A.S., Gundersen, G.W., Silverstein, M.C., et al. (2018). eXpression2Kinases (X2K) Web: linking expression signatures to upstream cell signaling networks. Nucleic Acids Res *46*, W171-W179. 10.1093/nar/gky458.
101. Hao, Y., Stuart, T., Kowalski, M.H., Choudhary, S., Hoffman, P., Hartman, A., Srivastava, A., Molla, G., Madad, S., Fernandez-Granda, C., and Satija, R. (2024). Dictionary learning for integrative, multimodal and scalable single-cell analysis. Nature Biotechnology *42*, 293-304. 10.1038/s41587-023-01767-y.
102. Wang, B., Wang, M., Zhang, W., Xiao, T., Chen, C.H., Wu, A., Wu, F., Traugh, N., Wang, X., Li, Z., et al. (2019). Integrative analysis of pooled CRISPR genetic screens using MAGeCKFlute. Nat Protoc *14*, 756-780. 10.1038/s41596-018-0113-7.
103. Landrum, M.J., Chitipiralla, S., Kaur, K., Brown, G., Chen, C., Hart, J., Hoffman, D., Jang, W., Liu, C., Maddipatla, Z., et al. (2025). ClinVar: updates to support classifications of both germline and somatic variants. Nucleic Acids Res *53*, D1313-D1321. 10.1093/nar/gkae1090.
104. Ma, L., Cao, J., Liu, L., Du, Q., Li, Z., Zou, D., Bajic, V.B., and Zhang, Z. (2019). LncBook: a curated knowledgebase of human long non-coding RNAs. Nucleic Acids Res *47*, D128-D134. 10.1093/nar/gky960.
105. Shihab, H.A., Rogers, M.F., Gough, J., Mort, M., Cooper, D.N., Day, I.N., Gaunt, T.R., and Campbell, C. (2015). An integrative approach to predicting the functional effects of non-coding and coding sequence variation. Bioinformatics *31*, 1536-1543. 10.1093/bioinformatics/btv009.

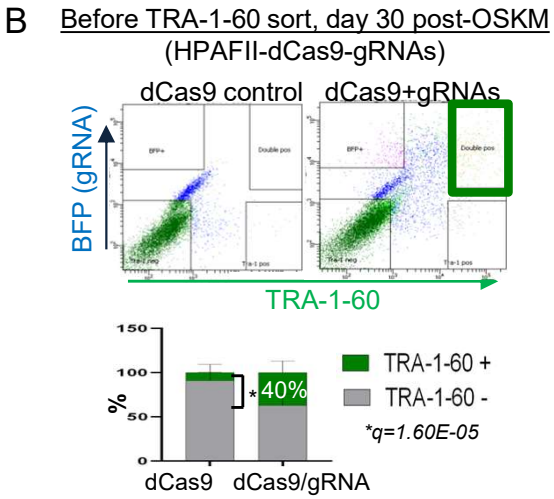
Grygoryev & Lee/Kim et al.,2025
Figure 1

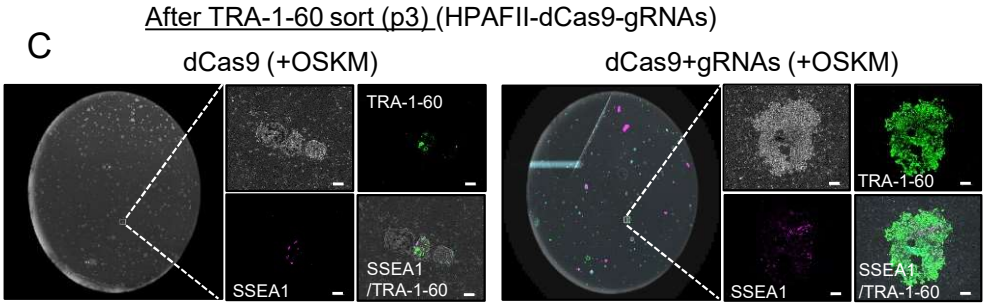


Grygoryev & Lee/Kim et al.,2025

Figure 2

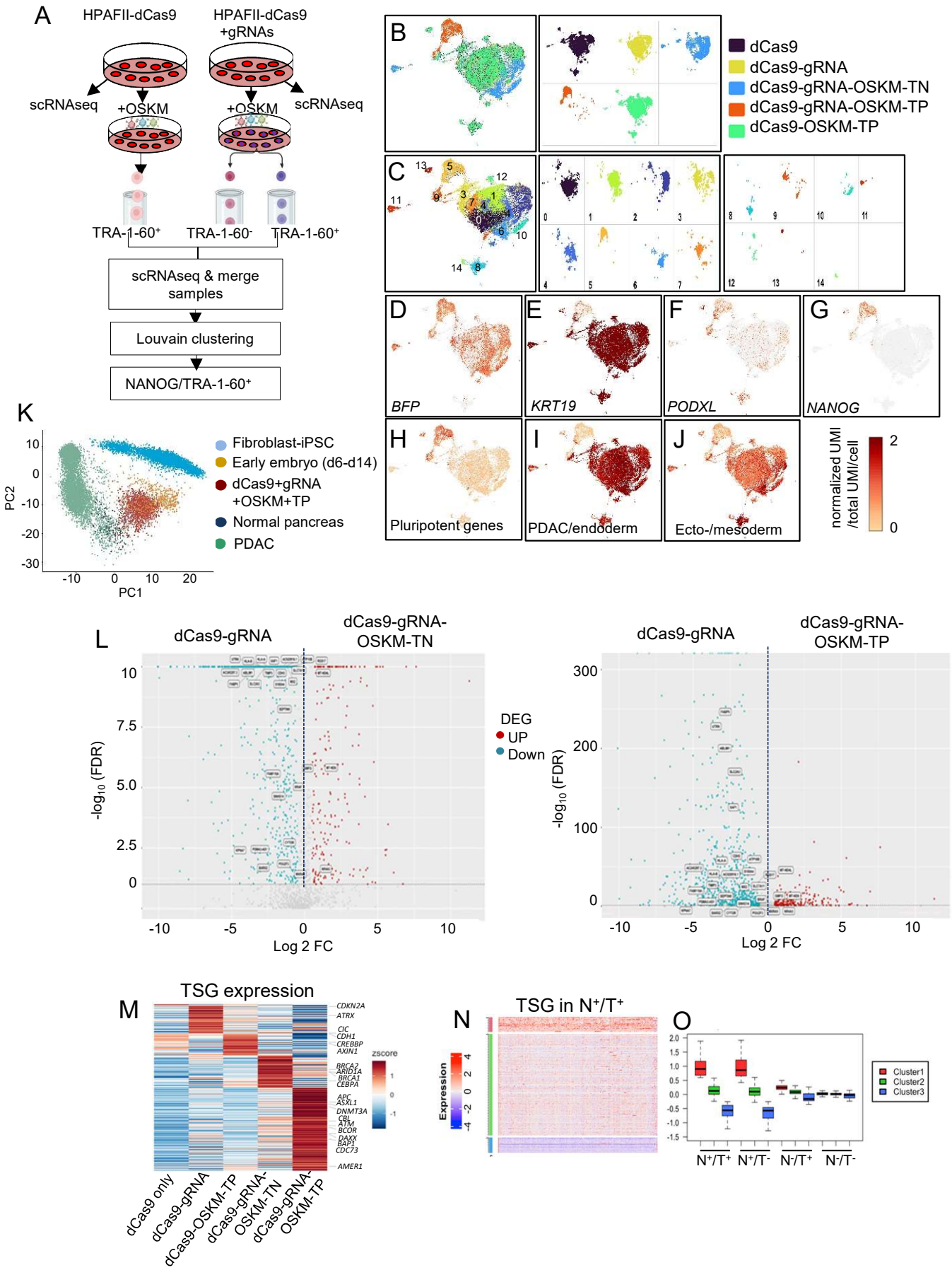






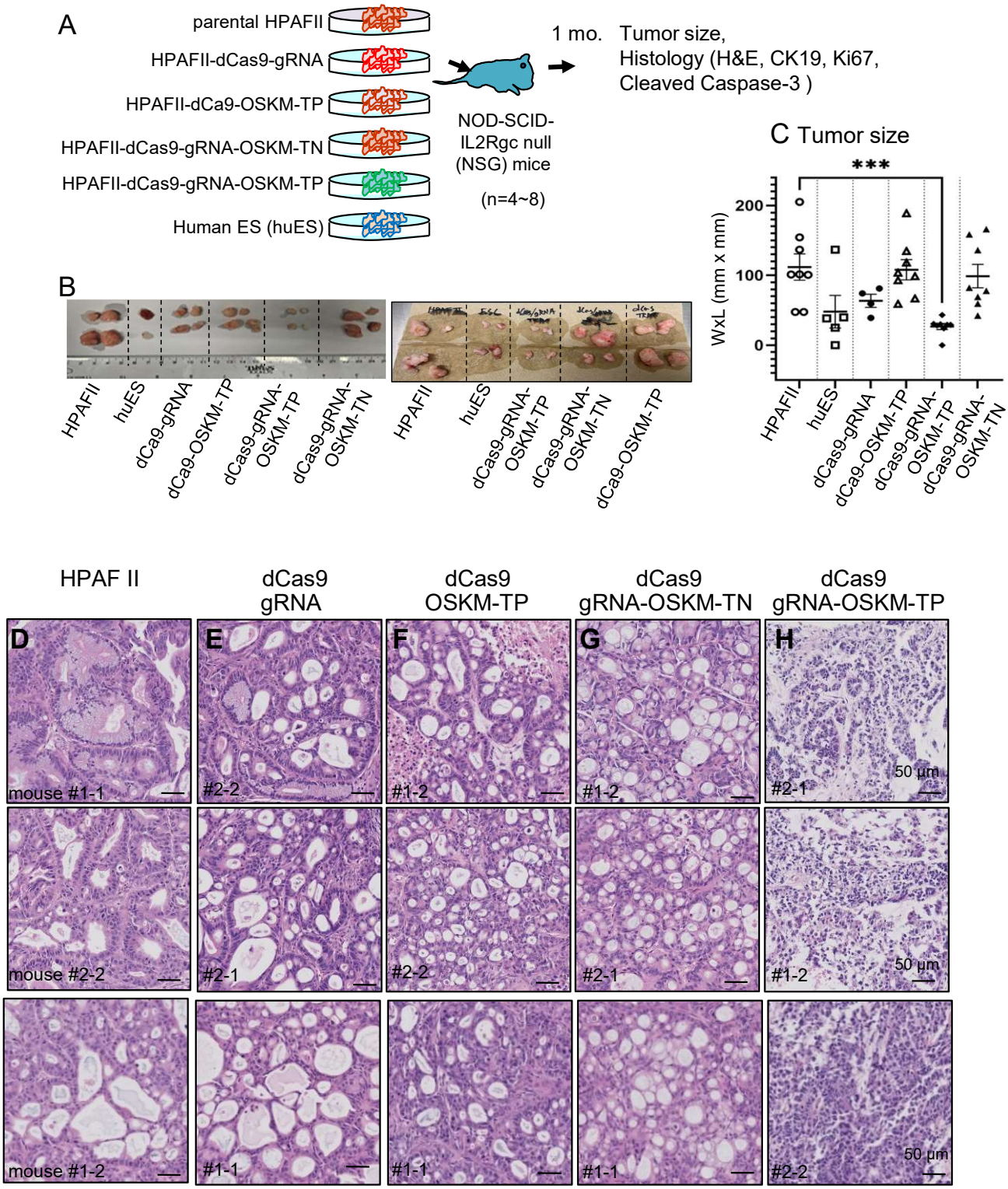
	Total cell #	Total SSEA1+	Total TRA-1-60+	TRA-1-60+ SSEA1+	TRA-1-60+ only	SSEA1+ only
dCas9+OSKM	809	29 (3.6%)	35 (4.3%)	20 (2.5%)	15 (1.9.%)	9 (1%)
dCas9+gRNAs	2665	23 (1%)	0	0	0	23 (1%)
dCas9+gRNAs+OSKM	212	47 (22.2%)	96 (45.3%)	32 (15.1%)	64 (30.2%)	15 (7.1%)

Grygoryev & Lee/Kim et al.,2025
Figure 3

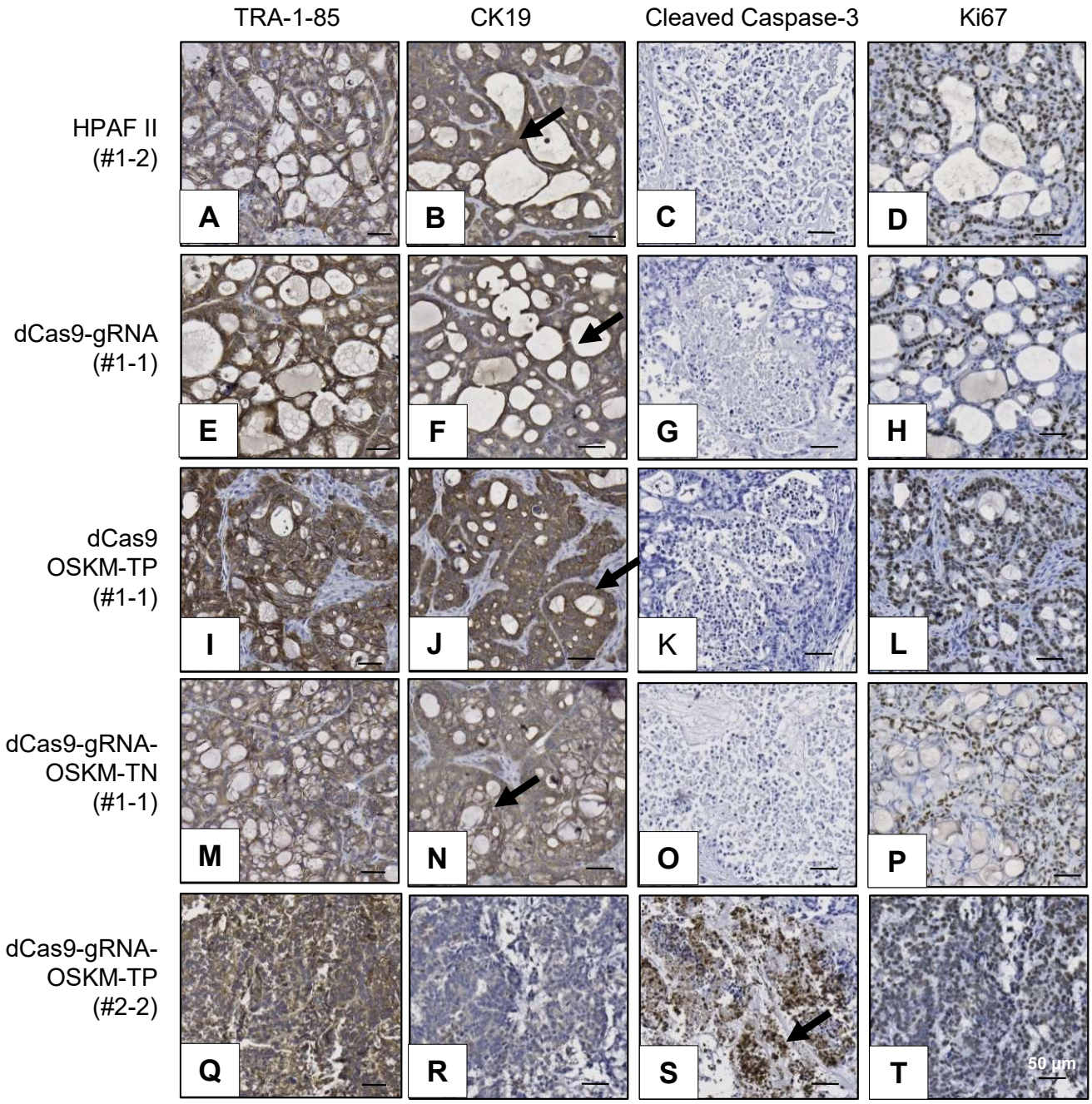


Grygoryev & Lee/Kim et al.,2025

Figure 4

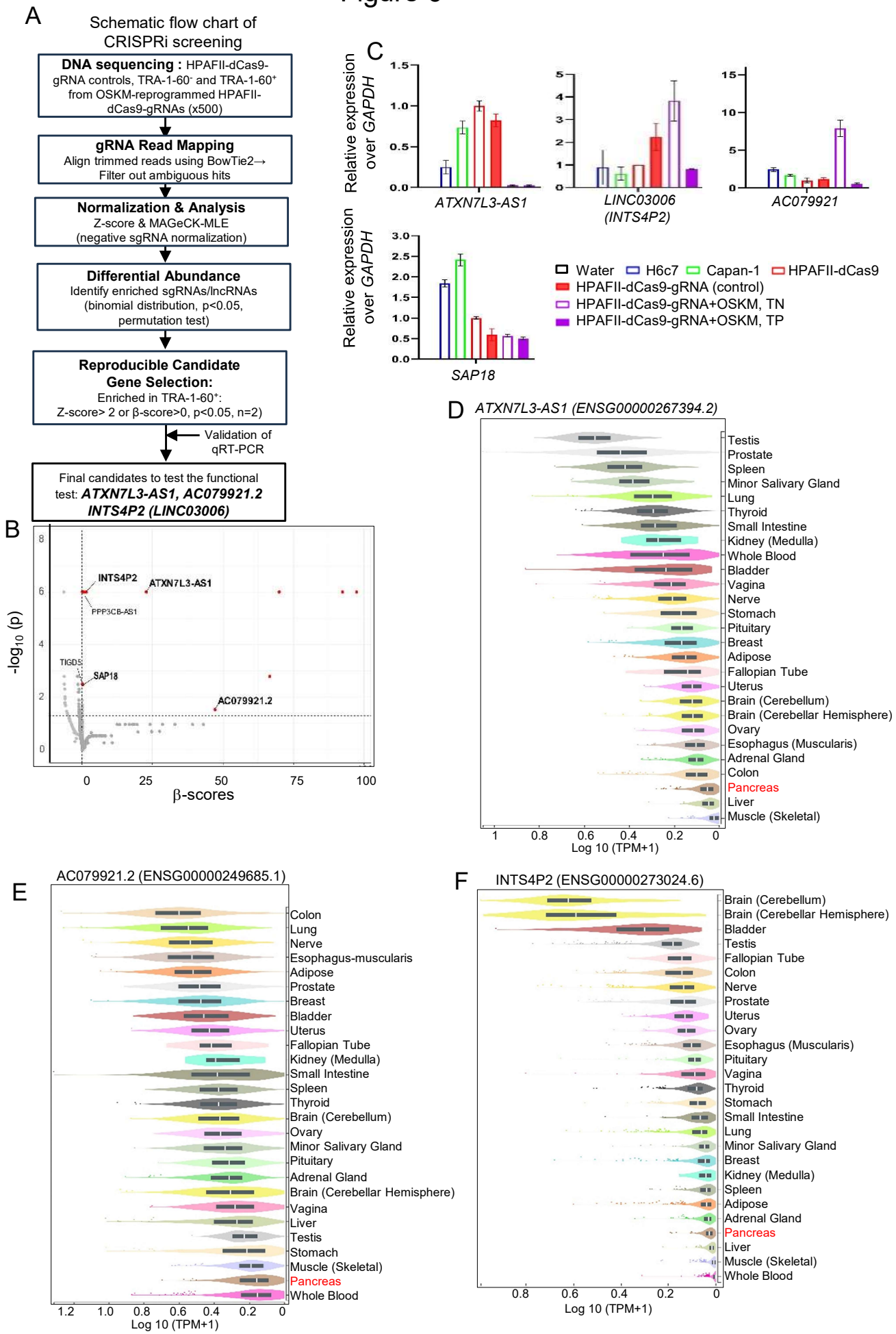


Grygoryev & Lee/Kim et al.,2025
Figure 5



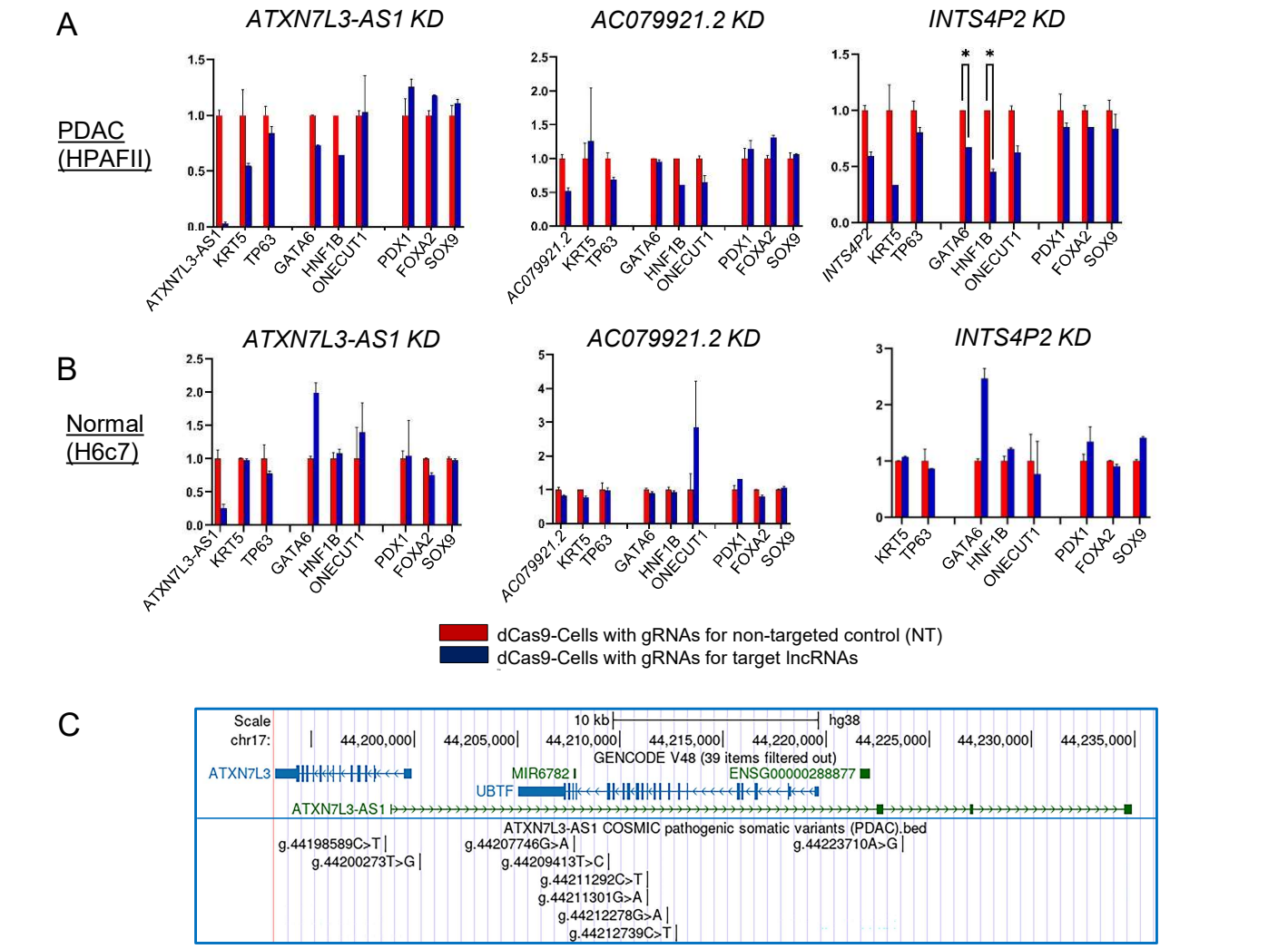
Grygoryev & Lee/Kim et al.,2025

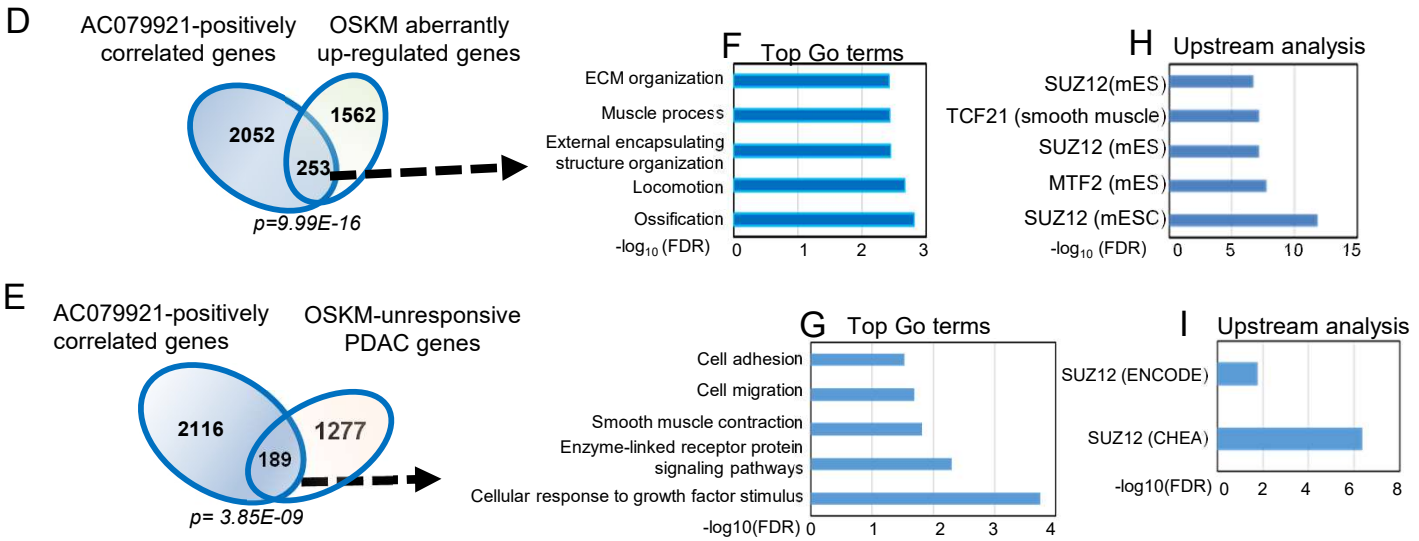
Figure 6



Grygoryev & Lee/Kim et al.,2025

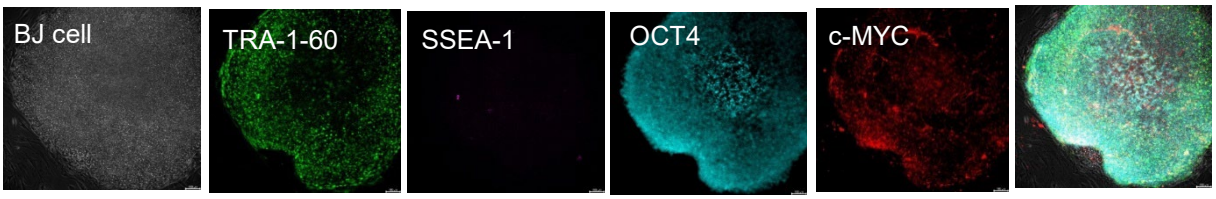
Figure 7



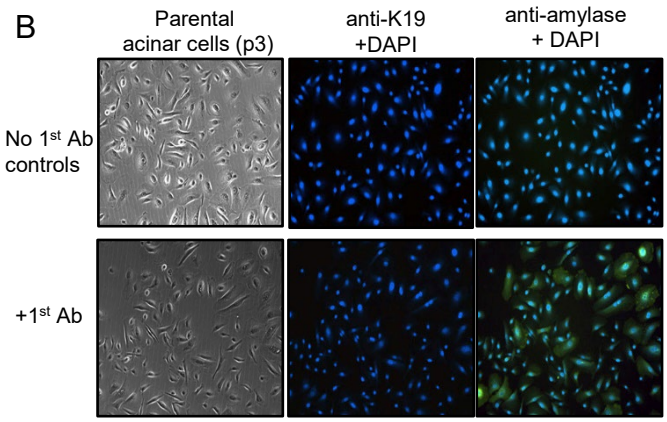


Grygoryev & Lee/Kim et al.,2025
Figure S1

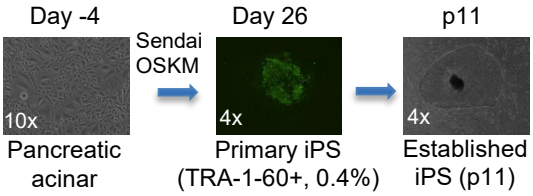
A. Immunostaining of pluripotency markers in BJ6 established iPSC



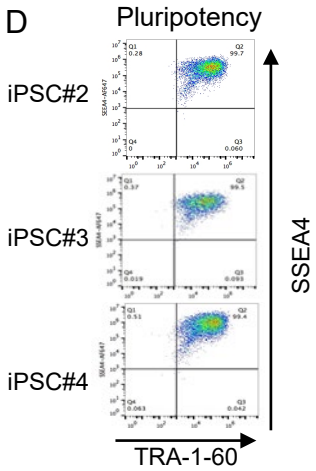
B



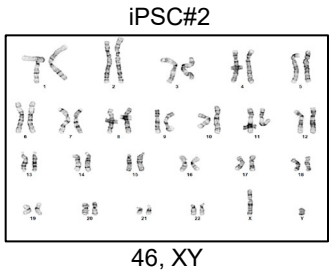
C



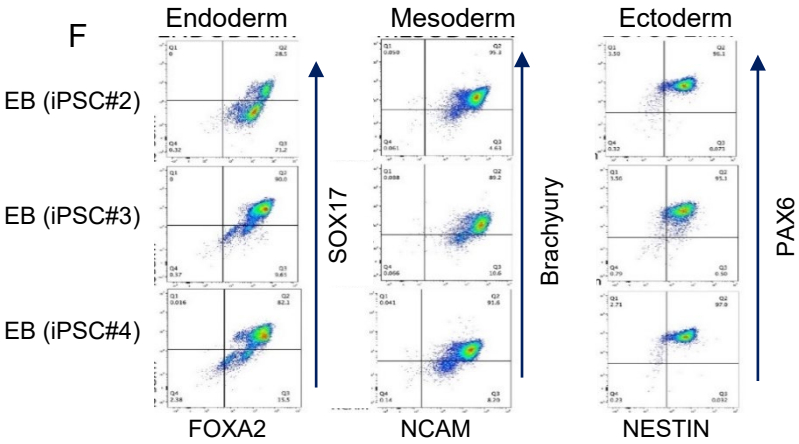
D



E

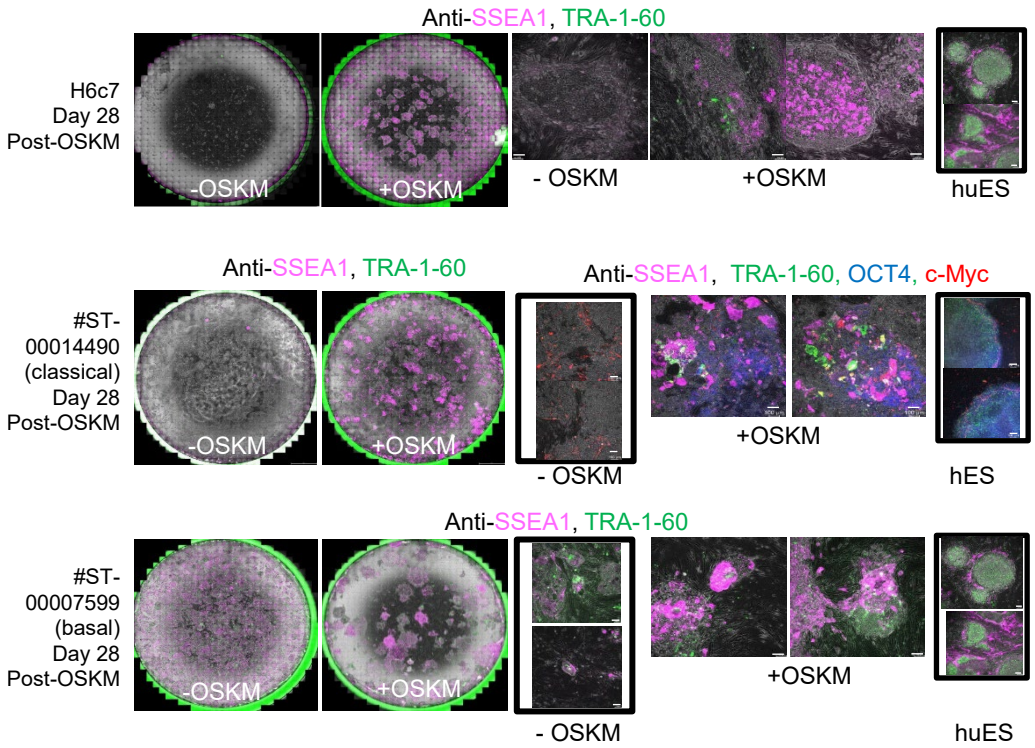


F

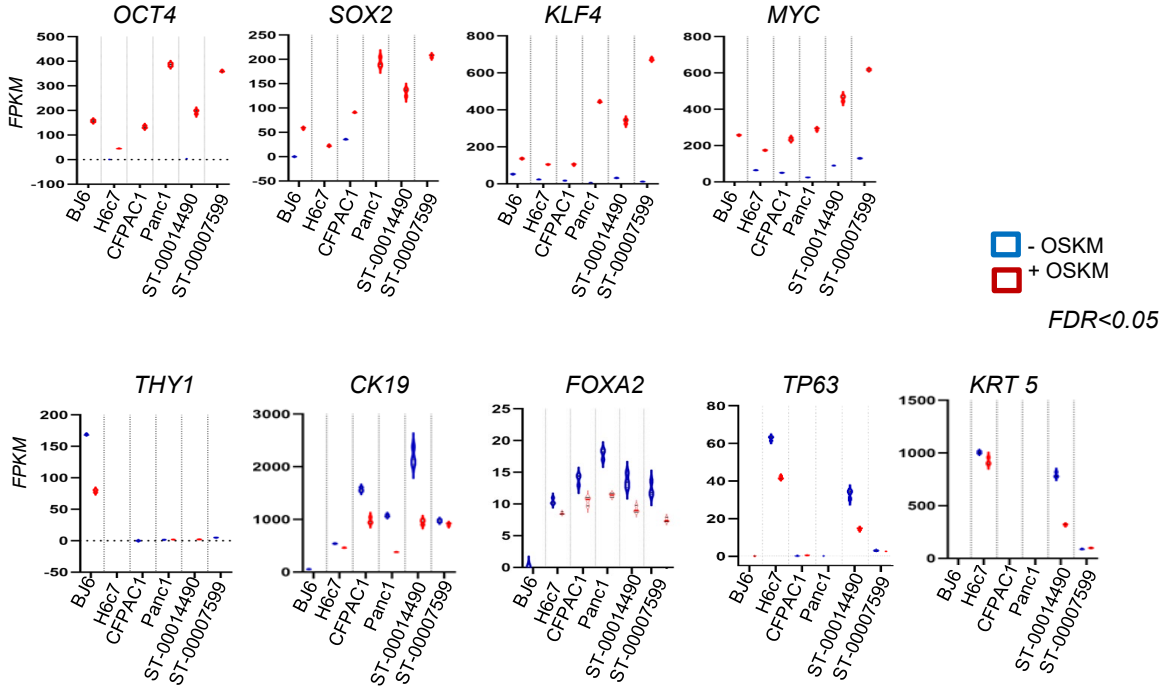


Grygoryev & Lee/Kim et al.,2025
Figure S2

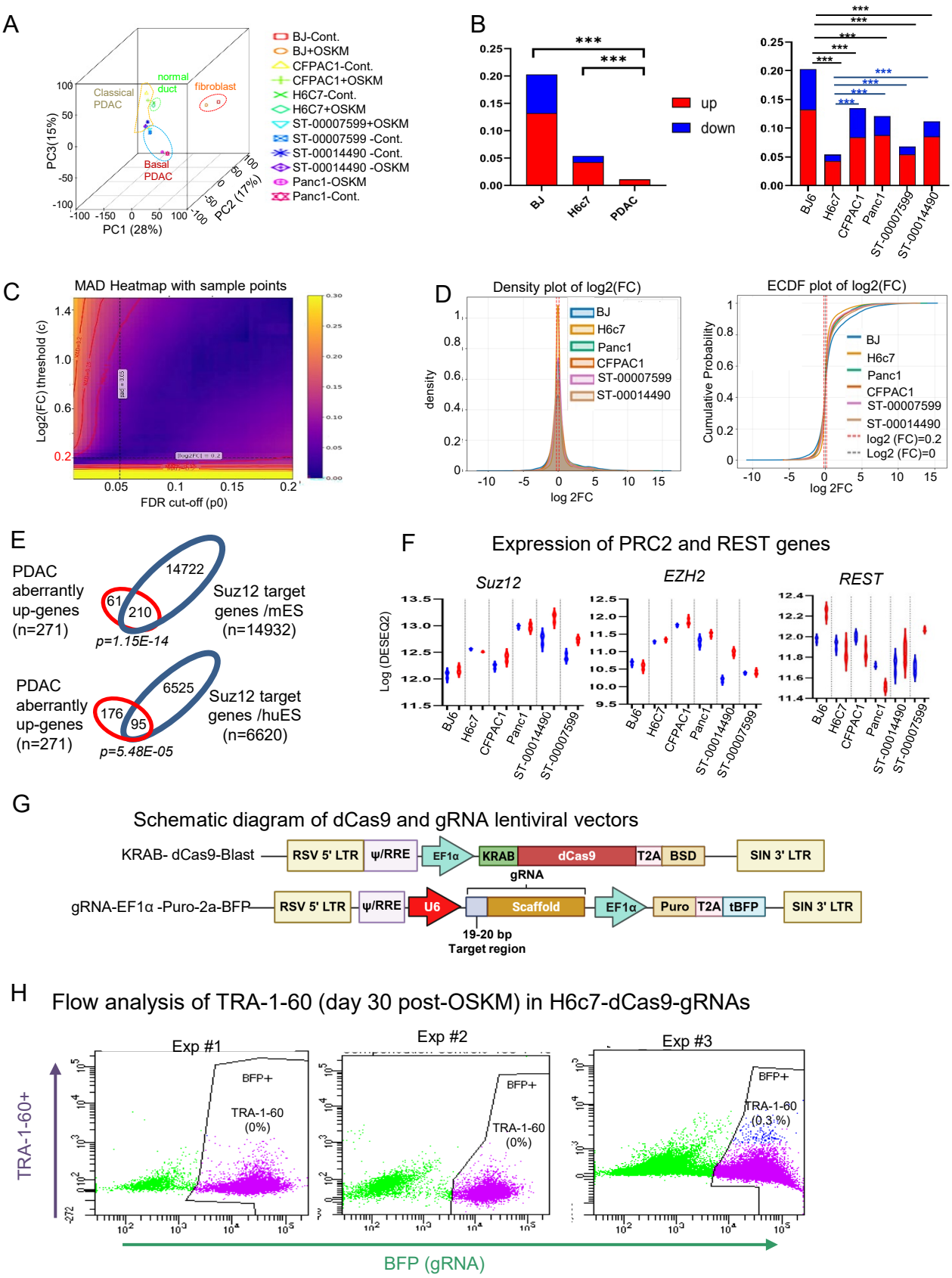
A. Immunostaining of pluripotency and early-differentiation markers



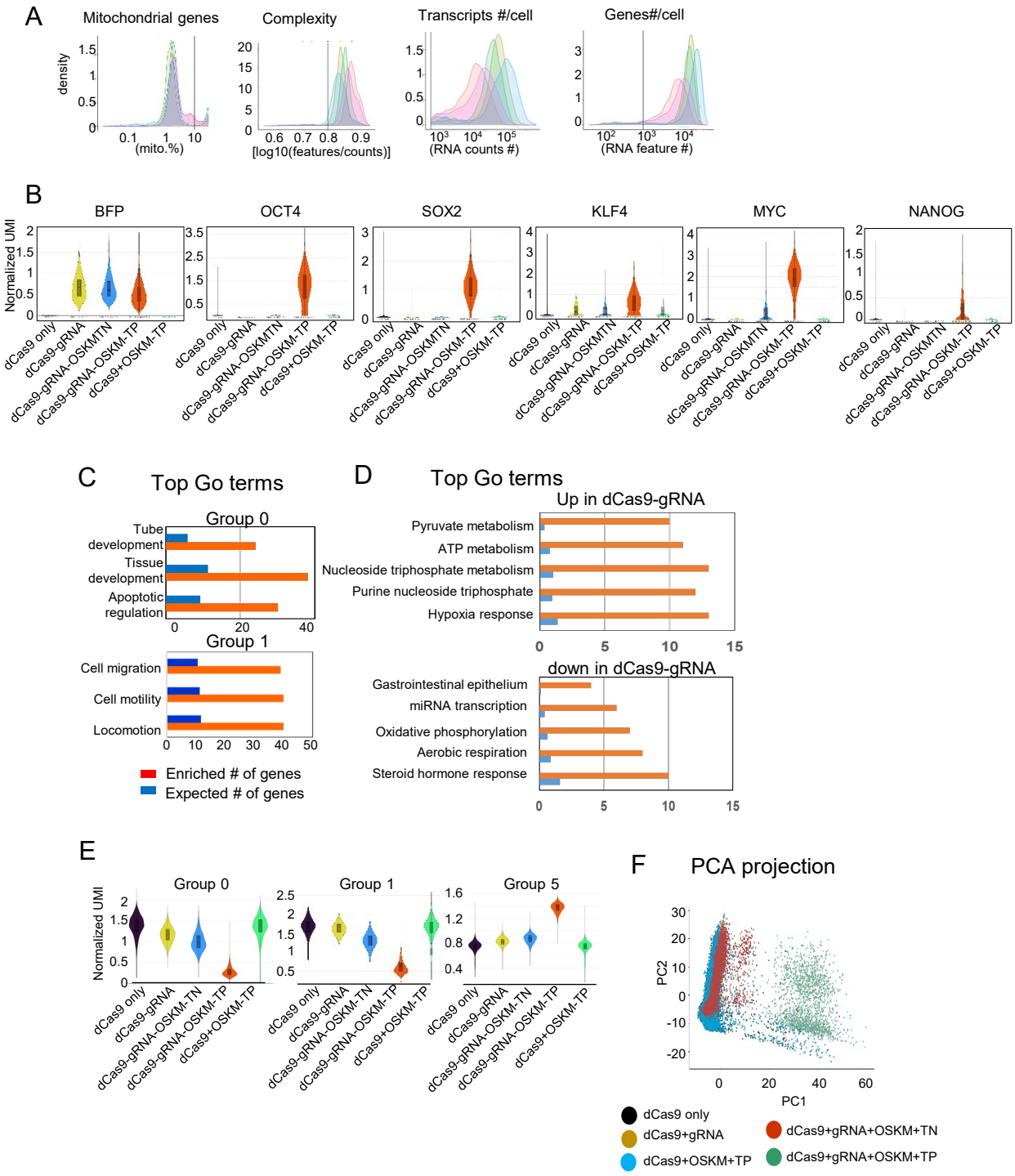
B. Expression of OSKM, pluripotency, and PDAC genes on day 7 after OSKM induction



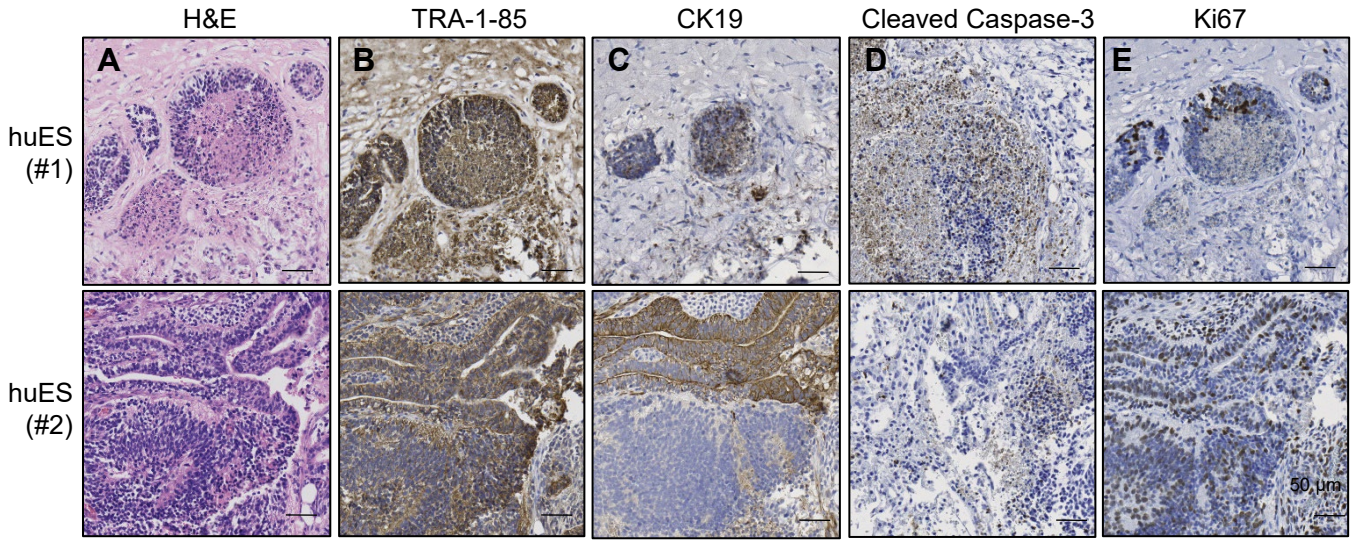
Grygoryev & Lee/Kim et al.,2025
Figure S3

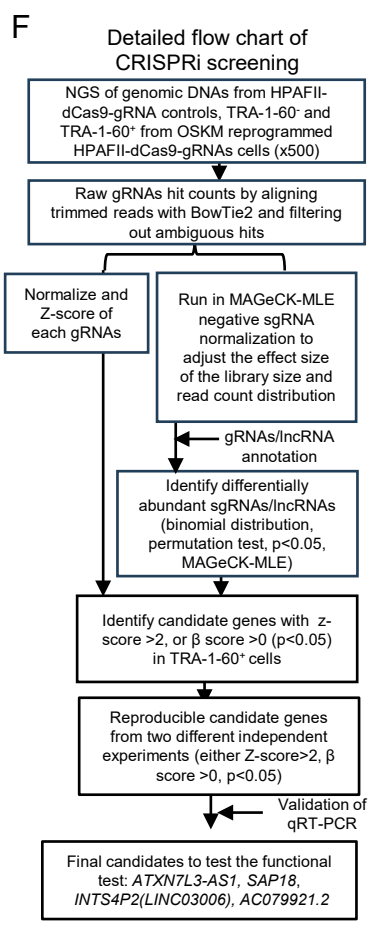


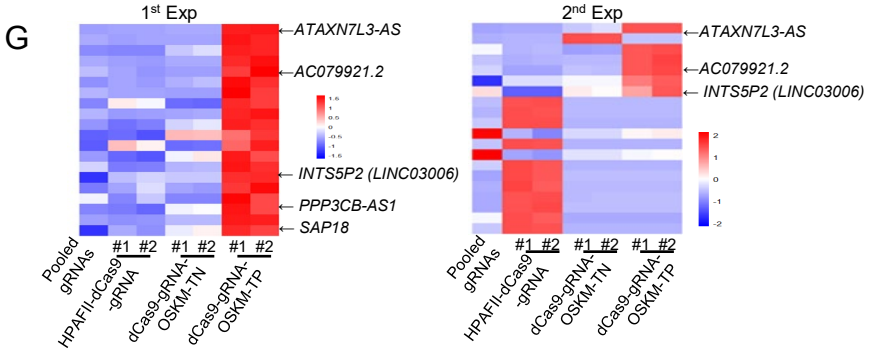
Grygoryev & Lee/Kim et al.,2025
Figure S4

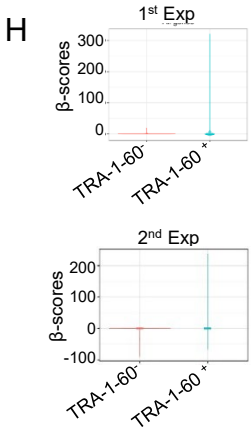


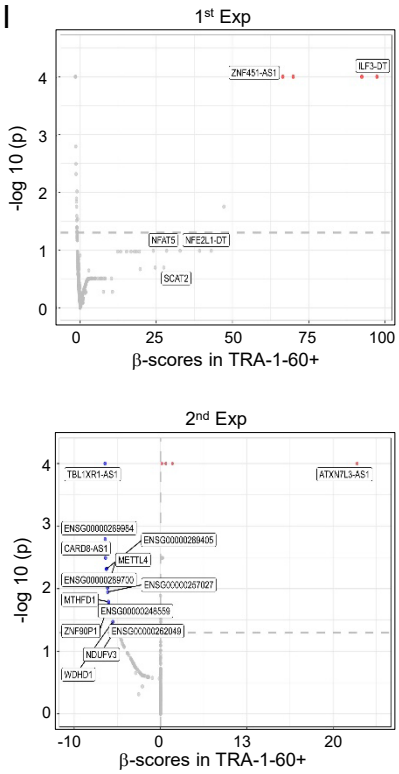
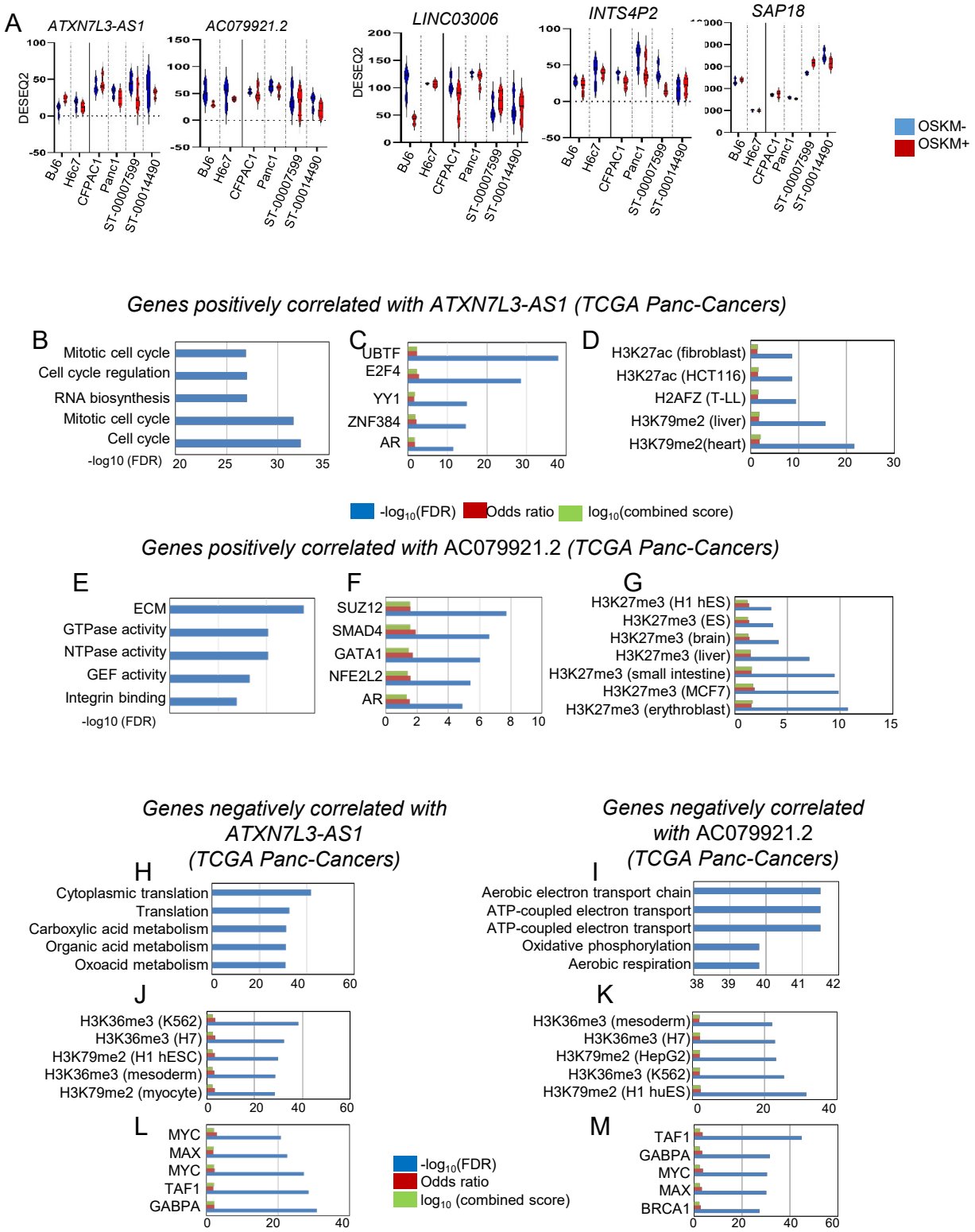
Grygoryev & Lee/Kim et al.,2025
Figure S5











Grygoryev & Lee/Kim et al.,2025
Figure S6