Texto, Logotipo

Descripción generada automáticamente

**ORIGIN OF CHROMOSOMAL INSTABILITY IN PANCREATIC CANCER**

Master Thesis

Análisis Bioinformático Avanzado, Universidad Pablo de Olavide

**Author**: Francisco Javier Mendoza León

**Tutor**: Daniel Rico

**Tutor**: Valentine Comaills

# ABSTRACT

Pancreatic ductal adenocarcinoma (PDAC) remains one of the cancer types with the poorest prognosis, and its progression involves a wide variety of genomic insults, complex chromosomal rearrangement, and an elevated rate of ploidy alterations. However, the mechanisms behind the generation of this extensive genomic heterogeneity during the initiation of tumorigenesis remain to be elucidated. Focusing on epigenetic reconversion after inflammatory damage, this study investigates whether acinar-to-ductal metaplasia (ADM) constitutes the origin of genomic instability in pancreatic acinar cells, promoting early tumorigenesis. By combining a computational approach based on available scRNA-seq data with experimental techniques, the results obtained suggest that ADM represents a vulnerable state in which cells accumulate genomic damage, such as double-strand DNA breaks and copy number variations, in early stages of PDAC progression. These alterations appear to compromise the integrity of the nuclear envelope, potentially leading to transient ruptures that enhance further damage, as reported in other types of cancer.

# TABLE OF CONTENTS

[ABSTRACT 2](#_Toc184657605)

[TABLE OF CONTENTS 3](#_Toc184657606)

[INTRODUCTION 4](#_Toc184657607)

[I. Pancreatic cancer and chromosomal imbalance 4](#_Toc184657608)

[II. Pancreatic Intraepithelial Neoplasia and Acinar to Ductal Metaplasia 5](#_Toc184657609)

[III. Inflammation and cancer initiation 6](#_Toc184657610)

[IV. Epigenetic reconversion, nuclear envelope and genomic instability 7](#_Toc184657611)

[OBJECTIVES 8](#_Toc184657612)

[METHODS 8](#_Toc184657613)

[Reference pre-processed scRNA-seq dataset 8](#_Toc184657614)

[Data download and initial processing 9](#_Toc184657615)

[CNV analysis 10](#_Toc184657616)

[Differential Expression Analysis and Functional Enrichment 10](#_Toc184657617)

[Immunohistochemistry (IHC) 11](#_Toc184657618)

[RESULTS 12](#_Toc184657619)

[Determining clusters of interests 12](#_Toc184657620)

[CNV patterns in early stages of pancreatic tumorigenesis 13](#_Toc184657621)

[Enriched pathways in early ADM transformation 15](#_Toc184657622)

[DNA breakage and chromosomal instability events following inflammatory damage 16](#_Toc184657623)

[DISCUSSION 17](#_Toc184657624)

[Future directions 18](#_Toc184657625)

[REFERENCES 19](#_Toc184657626)

# INTRODUCTION

## I. Pancreatic cancer and chromosomal imbalance

Pancreatic cancer ranks among the cancers with the highest mortality rates and poorest prognosis to date. In particular, pancreatic ductal adenocarcinoma (PDAC), which accounts for more than 90% of all pancreatic malignancies, has an overall 5-year survival rate of approximately 5%. This is mainly due to its asymptomatic course throughout most of its progression, which greatly hinders early diagnosis, and an increasing resistance to existing treatment options. Similarly to many other cancer types, PDAC exhibits a wide variety of genetic insults, ranging from point mutations to more intricate events, chiefly alterations in the number of copies of specific DNA segments or chromosomes generally termed as copy number variation (CNV)1.

Given the availability of new bioinformatic tools, massive efforts have been made in sequencing to understand the origins of genomic instability linked to the initiation and progression of cancer. Recent research has shown that the chromosomal imbalance resulting from CNV events is critical in certain cancer progressions, with tumors exhibiting an elevated level of CNV being more lethal2,3. Furthermore, comparative analyzes of whole genome sequencing data from primary tumors and metastases have revealed that tumor evolution in PDAC patients features a gradual accumulation of genetic anomalies, with events of loss of heterozygosis (LOH), chromothripsis, and complete genome duplication as early stages and driving events of PDAC onset4,5.

Thus, a significant challenge in the diagnosis of PDAC lies in its extensive genetic heterogeneity. RNA expression profiles of PDAC tissue have been conducted in hopes of generating a new framework for reclassification, personalized diagnosis, and individualized treatment. Nonetheless, the tumor stroma in PDAC accounts for most of the tumor volume, thereby diluting signal from proper tumor cells and complicating its classification in clinic. Therefore, tumor purity presents a major obstacle to the use of traditional bulk RNA seq data in patient diagnosis and tumor classification6,7.

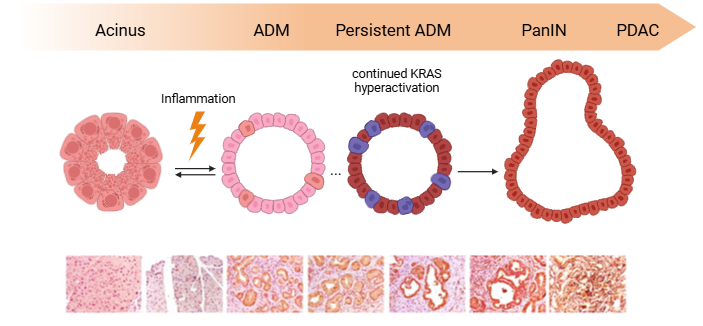
Additionally, given that PDAC is generally not diagnosed until advanced stages, the study of tumor progression is also challenged by the unavailability of tumor samples at different stages of the disease. However, animal models have greatly helped the field in modeling disease and its progression. The classical model of PDAC progression is transgenic mice expressing the *Kras* gene carrying the G12D hyperactive mutation, paired with the induction of pancreatic inflammation. These models mimic the disease from lesions that histologically culminate in a fully invasive and metastatic disease8.

## II. Pancreatic Intraepithelial Neoplasia and Acinar to Ductal Metaplasia

PDACs originate from precursor lesions termed Pancreatic Intraepithelial Neoplasia (PanIN) through a gradual progression. Paralleled sequencing studies of PanIN and PDAC samples have highlighted PanIN as a pretumorigenic lesion capable of progressing into PDAC through the accumulation of advantageous genetic driver events9. Notably, *Kras* mutations play a prominent role among these, and are present in approximately 90% of patients with PDAC. The constitutive expression of *Kras* in acinar cells induces a reversible event known as acinar to ductal metaplasia (ADM) and low-grade PanINs in mice. Accompanied by the accumulation of additional mutations in tumor suppressor genes, PanINs can eventually advance into PDAC. This transition can be accelerated by environmental stresses, most prominently inflammation10.

PanIN lesions are characterized by a disruption in the typical architecture of acinar cells and are categorized into different grades according to the level of cytological and architectural alterations. Lineage-tracing studies in mice have suggested that a prominent route to PDAC involves acinar cells undergoing acinar to ductal metaplasia (ADM), followed by the formation of PanIN lesions11,12 and, subsequently, invasive cancer (Figure 1). Furthermore, recent evidence confirms that inflammation-induced chromatin changes in ADM play an indispensable role in pancreatic tumor initiation13,14*.* ADM is a common and reversible process during pancreatic inflammation (pancreatitis) or injury that facilitates tissue repair. During pancreatic regeneration, acinar cells must cease the production of digestive enzymes to protect the compromised tissue from additional harm. Consequently, acinar cells acquire ductal features and inhibit enzyme secretion12*.*

Additionally, PanIN evolution is associated with alterations in nuclear shape13, underscoring the potential involvement of nuclear envelope protein modification and the probable appearance of nuclear envelope rupture (NER), which can explain the events of chromothripsis15–17.



**Figure 1**. Schematic depiction of acinar-to-ductal metaplasia (ADM) progression towards pancreatic ductal adenocarcinoma (PDAC). Following inflammatory injury, acinar cells begin to exhibit ductal characteristics. ADM is a reversible process, but if the damage is persistent, genetic reprogramming can lead to the formation of PanIN. Paired with sustained oncogenic Kras activation, these lesions can potentially initiate PDAC. Created with BioRender.com. Modified from Marstrand-Daucé, L et al. (Int. J. Mol. Sci., 2023).

## III. Inflammation and cancer initiation

Inflammation is part of the natural reparative response to healing and is essential to protecting against infections and detrimental environmental aggressors. Yet, when inflammation becomes prolonged and sustained, it can become damaging. Chronic inflammation has been established as a primary risk factor in oncogenesis, with a substantial body of research associating inflammatory events and immune repair mechanisms with increased tumor incidence, growth, and progression18.

Upon inflammatory responses, certain cells may undergo transformations that ultimately contribute to tumorigenesis. Still, the mechanisms by which these genomic alterations occur remain largely unknown. Multiple factors are put responsible for DNA damage during inflammation through defects in DNA damage repair pathways19 or increased release of reactive oxygen and nitrogen species20. Such mechanisms, however, result in point mutations and are not responsible for aneuploidy, complex chromosomal rearrangements, or gains in ploidy. As previously mentioned, extensive bioinformatic efforts to trace the oncologic history of patients have underscored the significance of CNVs as early driver events. Consequently, the elevated rate of ploidy alterations and chromosomal rearrangement established early on in many cancer types, such as PDAC, breast, or testicular cancer 4,21–24, remains to be elucidated.

## IV. Epigenetic reconversion, nuclear envelope and genomic instability

Instances of genetic instability subsequent to high plasticity events have been reported in various cancer types. Specifically, Dr. Valentine Comaills demonstrated that epigenetic alterations during Epithelial to Mesenchymal Transition (EMT) – a reversible embryogenic program adopted by epithelial cells in response to inflammatory stimuli – correlates with significant genomic instability, leading to long-term and heritable genetic anomalies. In breast cells, exposure to the inflammatory cytokine TGFß induces changes in ploidy, along with nuclear envelope rupture (NER) during interphase, increasing vastly the genetic diversity of unique clones25. Upon TGFß treatment, several nuclear envelope (NE) proteins are downregulated, such as Lamin A and B, Emerin or other NUPs. This may be attributed to the necessity for enhanced NE flexibility or for detachment of heterochromatin regions that repressed within Lamin associated domains to facilitate epigenetic regulation25–27. Ultimately, this decrease of NE results in several consequences, most notably an enhanced flexibility – albeit reduced resistance – of the NE which may induce transient NER, and the subsequent emergence of mitotic anomalies, including micronuclei formation and binucleated cells28. NER during interphase has gained an increased interest in cancer research, as it can serve as the origin of various simple and complex chromosomal rearrangements29 and trigger the activation of the cGAS/STING immune pathway15.

# OBJECTIVES

Within this framework, the project seeks to investigate whether epigenetic reconversion that occurs during ADM cellular plasticity events is the origin of genomic and chromosomal instability in pancreatic acinar cells during PDAC development, promoting increased genomic diversity and tumorigenesis. The primary hypothesis proposes that the acquisition of particular genomic patterns can induce transcriptional alterations and establish distinct dependencies on signaling pathways to endure drastic chromosomal imbalance. These specific pathways could represent therapeutic targets of interest and offer a strategy for the development of new PDAC classifications and personalized therapies.

By monitoring the progression of tumorigenic lesions during the initial stages of pancreatitis in inducible models, and focusing on simpler aberrant karyotypes, the study centers on identifying for genomic dependencies prior to the emergence of more intricate alterations.

# METHODS

This master thesis is part of a broader research project (EMERGIA EMC21\_00318) led by Dr. Valentine Comaills (University of Seville) with support from Dr. Daniel Rico (CSIC). The study integrates wet-laboratory techniques, to support the outcomes from computational analysis. Thus, the current MSc thesis serves as an initial step toward addressing the project's specific objectives.

While the group aims to generate its own scRNA-seq data in the near future, it is also partially reliant on publicly accessible datasets. For this thesis, primary analyses are based on data published by Burdziak et al. (Science, 2023)13.

## Reference pre-processed scRNA-seq dataset

The main computational analyses conducted in this study focused on a publicly accessible single-cell RNA seq data from Burdziak et al. (Science, 2023) 13. As part of their investigation, Burdziak’s team characterized PDAC progression by using genetically engineered murine models (GEMMs). Based on the recombinase Cre-Lox system, the expression of reporter *LSL-mKate2* was induced under the control of acinar cell-specific *Ptf1a*-Cre transcription factor bothindependently and in conjunction with the mutated oncogenic allele *LSL-KrasG12D*. In addition, mutated *p53R172H* alleles were also utilized in certain models to accelerate tumor malignancy.

Expression of reporter mKate2 allows identification of pancreatic epithelial cells and their isolation by FACS sorting. Inflammatory damage was induced via caerulin (CAE) injections, and samples were collected in different stages of PDAC progression, including: healthy (N1) and regenerating normal tissue following CAE injection (N2), pre-neoplasic tissue expressing *KrasG12D* independently (K1) or in combination with CAE inflammatory injury (K2), early (K3) and late (K4) PanIN lesions, as well as *p53R172H*-inducedprimary tumor (K5) and distal metastasis (K6). These samples were used to prepare the scRNA-seq libraries using the 10X Genomics Chromium instrument, following the protocol for the 3’ Reagent Kit v2. Sequencing was performed with an Illumina sequencer.

As detailed in the original publication, the scRNA-seq data from all conditions were pre-processed in accordance with standard protocols with SEQC, including demultiplexing, barcode correction, alignment, and UMI correction. Burdziak and colleagues developed a cluster-based iterative filtering framework to manually identify high-quality populations and mitigate technical noise. Filtered count matrices were generated for each sample and integrated into a single data matrix. Following library size normalization and supplementary filtering steps, including outlier elimination via Phenograph, DoubletDetection, and removal of stromal contaminants, the final dataset consisted of 28131 high-quality mKate2+ epithelial cells and 16828 genes.

## Data download and initial processing

The final processed scRNA-seq dataset data was downloaded as a scanpy anndata object file (h5ad) from Gene Expression Omnibus, under accession number GSE207938. This object comprises raw and un-logged normalized counts, as well as metadata including Phenograph clusters, batch and conditions, and PCA and tSNE coordinates.

In R version 4.4.1, the h5ad file was initially converted into a Seurat object using the schard library30, importing raw counts, metadata and dimensionality reductions. Upon verification of the transformed data, preliminary processing of the dataset was performed, including the addition of tissue state to the metadata and log2 transformation of normalized counts, following author’s recommendations.

It was noted that the gene symbol format within the dataset did not correspond with the conventional representation for mouse genes, as all characters were capitalized. After verification of mouse-specific genes, the symbol format was corrected to ensure the appropriate functionality of subsequent analyses. Finally, the processed dataset was integrated into a new Seurat object and stored as an RDS file for the subsequent analyses.

The script with commented code and additional steps employed for this process is available in Annex I.

## CNV analysis

Analysis of CNV patterns in early stages of the onset of PDAC was performed using the inferCNV package31 of the Trinity Cancer Transcriptome Analysis Toolkit Project (CTAT) designed to detect large-scale somatic alterations in the number of chromosomes from tumor scRNA sequencing data by comparing expression levels across the sample genome with a reference group.

InferCNV requires three main inputs that are grouped in an inferCNV object: i) a raw counts matrix of scRNA-seq data, ii) an annotations file indicating the reference and sample cell populations, and iii) a gene ordering file indicating the chromosomal location for each gene. CNV profiles are detected by a moving of expression data that smooths gene-specific expression patterns while maintaining chromosomic aberrations.

The Gencode genome reference file for the GRCm39 (vM32) assembly, provided by CTAT, was used in this study. The script with commented code and additional steps employed for this process is available in Annex II.

## Differential Expression Analysis and Functional Enrichment

Seurat package32 was employed to conduct differential expression analysis across clusters of interest with a customized function based on Seurat’s FindMarkers, identifying differentially expressed genes via a Wilcoxon test. The resultant data frame is organized in descending order based on the average log2FC and adjusted p-value and stored as a text file. The resultant data frames, generated for each comparison of interest, were used to conduct enrichments with clusterProfiler library33.

## Immunohistochemistry (IHC)

Complementary to the computational analysis, immunohistochemistry (IHC) staining was performed on fixed tissue derived from the laboratory’s own mice. As previously mentioned, the models employed for the experimental analyses in the laboratory shares many similarities with the ones described in Burdziak et al.13. Specifically, pancreatic epithelial cells expressed *LSL-GPF* reporter and *LSL-KrasG12D* under *Ptf1a-Cre* regulation activated via tamoxifen ingestion. Control mice only expressed GPF reporter.

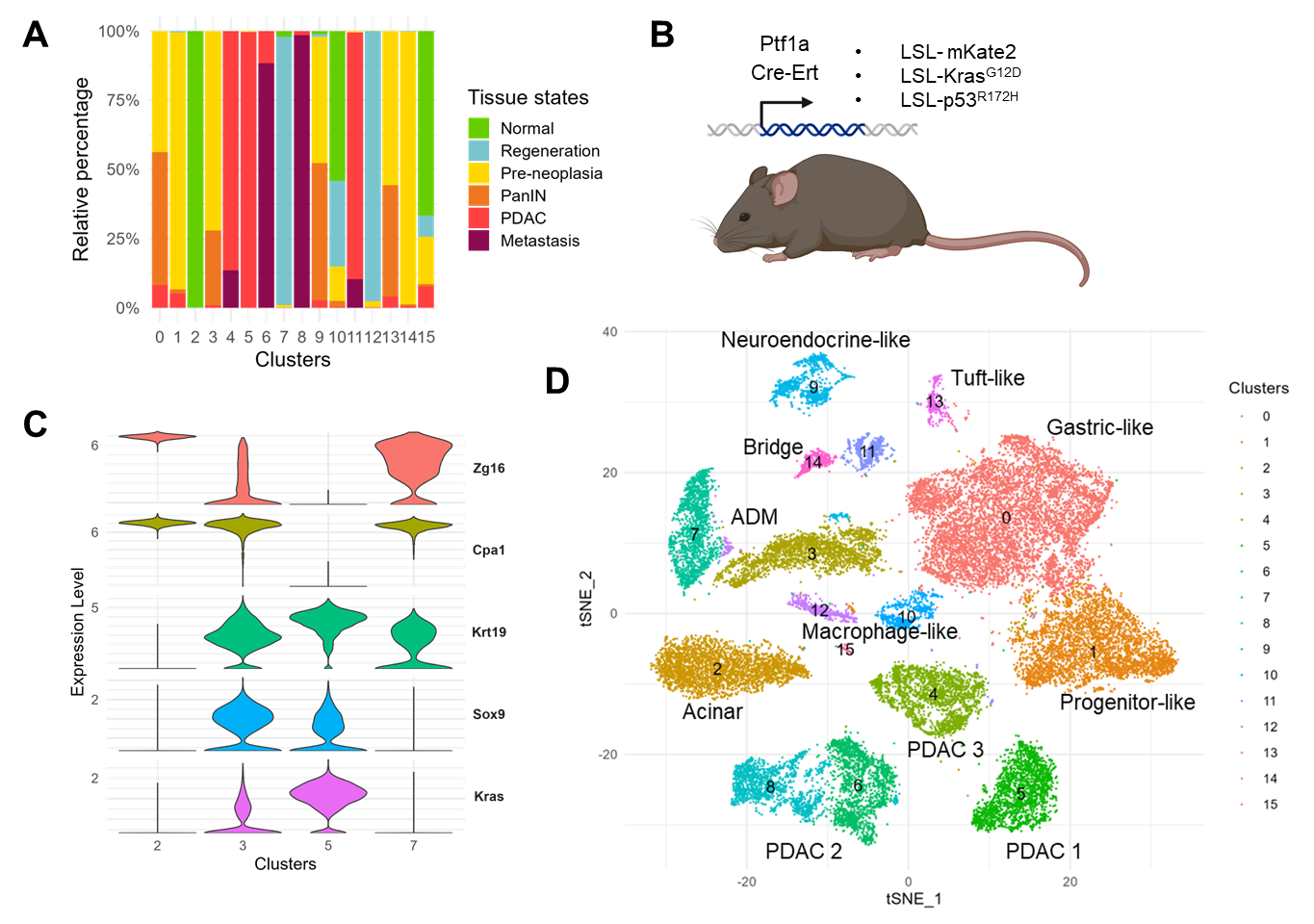
For IHC staining, previously fixed pancreatic tissue slides, provided by CABIMER histology service, were sequentially rinsed with xylene, ethanol series decreasing in concentration (100%, 70%, 50%, 50%, and 30%), distilled H2O and TBST-T for rehydration. Heat induced epitope retrieval was performed by immersing the slides in EDTA solution (1nM, pH=9) for 20' in a pressure cooker set on high. Then, samples were incubated for 1h at room temperature in a TBS-T + 10% fetal bovine serum (FBS) solution for antigen blocking, and tissue was stained at 4ºC overnight using primary antibodies mouse anti-γH2AX (1:200 dilution; Millipore 05-636) and rabbit anti-LaminB1 (1:500 dilution; Abcam, ab16048). Following primary antibody incubation, slides were rinsed with TBS-T and incubated for 2h at room temperature with secondary antibodies (1:200 dilutions) Alexa Flour 647 donkey anti-mouse IgG (Invitrogen, A31571) and Alexa Fluor 647 donkey anti-rabbit IgG (Invitrogen, A31573).Lastly, samples were again washed with TBS-T and mounted with VECTASHIELD Antifade Mounting Medium with DAPI (Vector Laboratories, H-1200-10), cover-slipped and stored at 4°C in the dark.

Images were acquired with 20X and 60X objectives on a Leica DMi8 microscope with a W-View Gemini (Hamamatsu) camera attached. Compositions were processed using the Leica LAS X Office software (v3.10.0), and DNA damage was quantified using Image J software (v2.14.0). Graph results were obtained in GraphPad Prism (v8.4.2).

# RESULTS

## Determining clusters of interests

As an initial step prior to performing any analysis, cells of interest were identified. The source metadata did not contain annotations for cell types; however, such information was available in the original paper. After examining the tSNE representation of cell types available in the paper, the following cell types of interest were selected as initial groups of interest: committed acinar cells grouped in cluster 2, ADM cells grouped in clusters 3 and 7, and PDAC cells grouped in clusters 4,5,6 and 8. This project is centered on the initial phases of PDAC progression before accumulation of aberrant genetic imbalances. Hence, tissue state distribution across all clusters was explored (Figure 2). In this context, cells in cluster 7 originated from regenerative pancreatic tissue after CAE injury (condition N2), depicting an earlier stage of ADM, whereas ADM cells grouped in cluster 3 represented a later ADM state derived from samples taken after CAE injection coupled with *KrasG12D* activation (conditions K1-K5)*.* Similarly, among multiple clusters of cells identified as PDAC cells, only cluster 5 comprised predominantly cells from non-metastatic PDAC tissue (K5), representing an earlier stage of the PDAC primary tumor. tumor. This was corroborated by analyzing the expression of characteristic marker genes for acinar, ductal, and tumor cells: acinar markers, including *Zg16* or *Cpa1*, were expressed in committed acinar cells and ADM cells (clusters 2, 3, and 7), whereas ductal markers such as *Krt19* or *Sox9* were exclusively expressed in ADM cells (clusters 3 and 7). PDAC cells were validated by the expression of *Kras*.



**Figure 2:** Exploration of cell origin within each cluster. **A)** Barplot depicting the relative percentage of cells from each tissue state within each cluster. Cluster 3 comprises ADM cells derived from various conditions after CAE injection and Kras activation, whereas ADM cells within cluster 5 originate from recovering normal pancreatic tissue. Similarly, among tumorigenic cells, cluster 5 uniquely does not include cells from metastatic samples. **B)** Burdziak et al. GEMMs expressing mKate2+ exocrine pancreatic cells under Ptf1a-Cre regulation, adapted from from Burdziak et al. (Science 2023). **C)** Violin plots of normalized gene expression in acinar cells (cluster 2), ADM cells (clusters 3 and 7), and PDAC cells (cluster 5), of acinar (Zg16, Cpa1), ductal (Krt19, Sox9), and tumorigenic (Kras) marker genes. **D)** tSNE visualization of cell clusters and cell type annotations, adapted from from Burdziak et al. (Science 2023).

## CNV patterns in early stages of pancreatic tumorigenesis

To examine the initial patterns of duplications or aneuploidies in PDAC, inferCNV analysis was conducted with acinar cells as the reference group in comparison to both ADM and early PDAC cells (Figure 3). The resulting heatmap demonstrates variations in the CNV profiles across ADM and PDAC states, identifying potential subpopulations of cells that exhibit relatively established patterns of chromosomal aberrations. While this imbalance is more pronounced in PDAC cells, early ADM states already display a marked variation in their profiles, which is relatively preserved in the subsequent stages of PDAC progression.

Gráfico

Descripción generada automáticamente con confianza baja

**Figure 3:** Copy Number Variations (CNV) analysis with inferCNV using data from Burdziak et al. (Science 2023). Chromosomal regions with duplication (red) or deletion (blue) events are indicated in the heatmap, thus establishing possible identifying patterns of specific subgroups within cells that have undergone ADM and from PDAC.

The observed patterns are not uniformly maintained across all cells within each cluster, thereby suggesting potential subpopulations in each group. For instance, both ADM clusters include subgroups of cells exhibiting expression profiles more closely aligned with those shown in the control group, which may represent acinar cells still undergoing ADM transformation. Supplementary outputs from the inferCNV run provide a more detailed depiction of subpopulations within each state, which can be extracted in the future to enable more comprehensive analyses.

## Enriched pathways in early ADM transformation

To further examine the affected processes during ADM transformation, a more detailed examination of the differentially expressed genes within each cluster of interest was conducted.

Gráfico

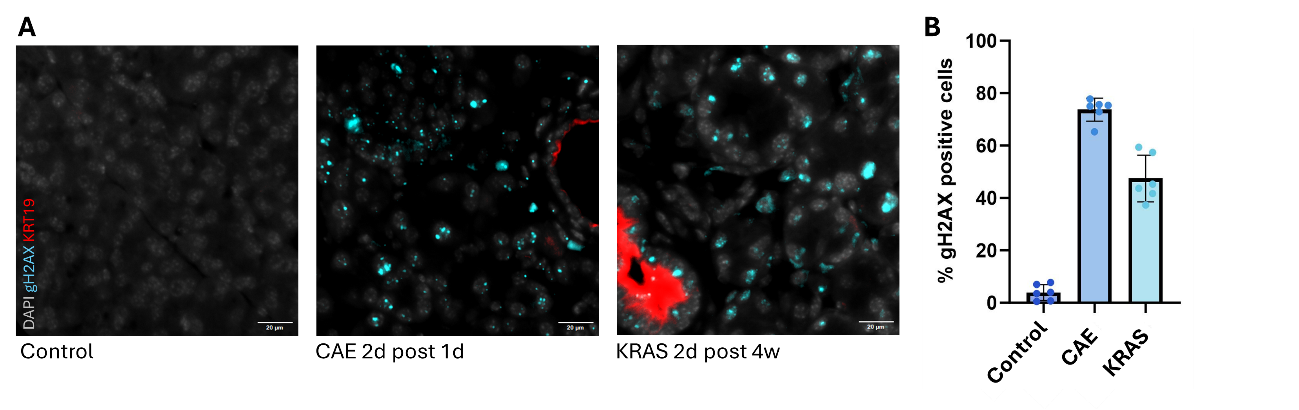
Descripción generada automáticamente

**Figure 4:** Functional enrichment analysis of differentially expressed genes in ADM cells in relation to acinar and PDAC cells. Cells that have undergone ADM show alterations in terms related to nuclear envelope integrity and reparative pathways.

Functional enrichment analysis of these genes generated expected results, such as the upregulation of digestive enzymes and protein production pathways in acinar cells, or anaerobic respiration in PDAC cells. However, particularly significant findings were observed in relation to ADM transformation (Figure 4). Notably, cells that had undergone this epigenetic reconversion displayed a downregulation of genes associated to the nuclear matrix, stress responses and wound healing pathways. This suppression was noted from early stages of ADM transformation, highlighting their potential role in the progression of this state.

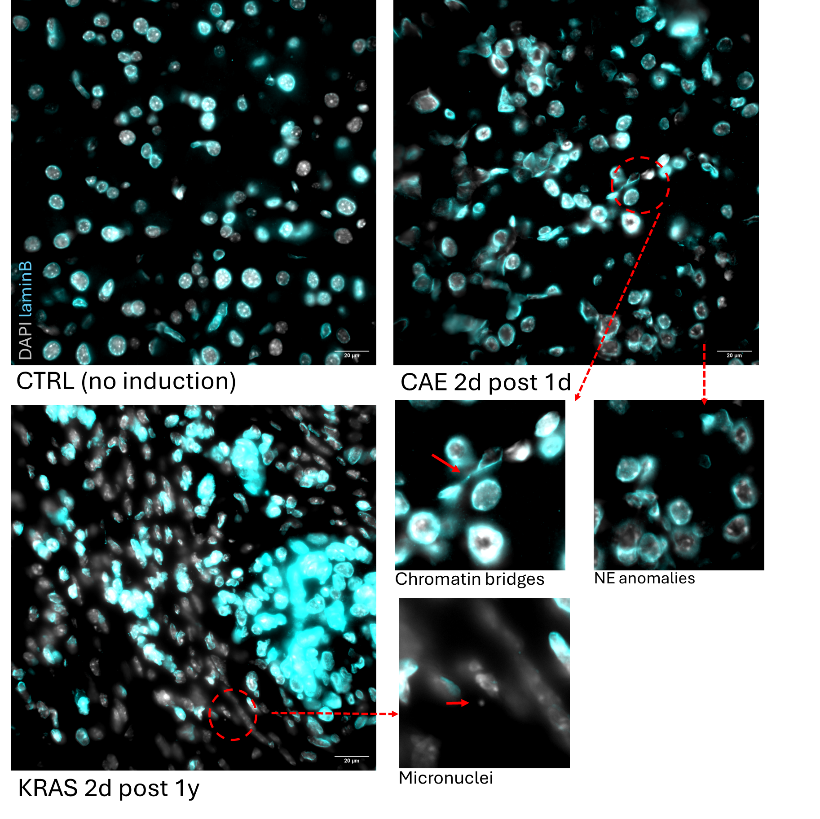
## DNA breakage and chromosomal instability events following inflammatory damage

Supportive experimental results derived from IHC revealed a substantial increase in DNA damage in the models employed after acute inflammatory injury or constitutive *Kras* mutation. Both mice subjected to CAE injection for two days and mice expressing the *KrasG12D* mutation exhibited a significant increase of γH2AX signal, an indicator of double-strand DNA break (DSBs) (Figure 5).



**Figure 5. A)** Immunofluorescence images from γH2AX staining to analyze inflammatory damage and Kras activation impact in pancreatic cells. **B)** Graph depicting a significant increase of DNA double-strand damage in caerulein (CAE) treated and Kras-activated pancreas. CAE treatment and Kras induction significantly increase DNA double-strand breakage.

In addition, staining for LaminB1, a structural protein of the nuclear membrane, allowed identification of multiple indicators of instability, including alterations of the nuclear envelope, formation of micronuclei or apparition of chromatin bridges (Figure 6). Altogether, these events can lead to complex chromosomal rearrangements and genomic imbalance, both of which are critical factors in cancer progression and contribute to tumor heterogeneity.



**Figure 6.** Inflammatory damage and constitutive activation of Kras lead to the emergence of chromosomal instability events, indicated by the presence of chromatin bridges, the formation of micronuclei, or the emergence of nuclear envelope anomalies.

# DISCUSSION

In all, the collection of data obtained throughout this project supports the hypothesis that acinar-to-ductal metaplasia during pancreatic tissue regeneration represents a highly susceptible phase in which epithelial cells are predisposed to genomic damage, ultimately facilitating the onset of PDAC and contributing to tumor heterogeneity.

This is supported by the IHC findings relative to γH2AX analysis, which demonstrate significant damage subsequent to an acute inflammatory event. The rapid double-stranded DNA breakage observed following CAE injection would potentially provide the basis for the occurrence of inaccuracies in DNA repair processes, ultimately resulting in chromosomal rearrangements, deletions or duplications of DNA segments, or the formation of micronuclei. Such events are characteristic of the progression of various cancer types, and contribute to the extensive genomic heterogeneity observed among tumor subclones15,34,35. In this sense, results from CNV pattern analysis with inferCNV suggest that, in early stages of PDAC, there is already a considerable heterogeneity present in the genomic landscape of ADM transformed cells under inflammatory response or *Kras* activation. Even preceding neoplastic lesions, pancreatic tissue exhibits specific patterns of deletions and duplications that are relatively consistent across the experimental samples. Many of the patterns initially detected in ADM cells persist into the tumorigenic stages, suggesting that these alterations originate from the epigenetic plasticity occurring during ADM and progressively accumulate throughout the progression of the disease16,25,36.

Additionally, results derived from IHC and functional enrichment may suggest that the accumulation of these genetic insults could originate from an enhanced vulnerability of the nuclear envelope. In particular, enrichment analysis of differentially expressed genes in cells subjected to ADM transformation, in comparison to acinar and tumorigenic cells, indicates a suppression of genes involved in maintaining the nuclear envelope integrity15,16,25. These findings are consistent with qualitative analysis of IHC with LaminB1, in which nuclear envelope deformations and indicators of mitotic anomalies such as micronuclei and chromatin bridges were observed.

Nevertheless, while these results seem to indicate a tendency attributed to nuclear envelope weakness following the major epigenetic reconversion associated with ADM, the findings remain preliminary. Thus, further analyses and corroboration with additional datasets are necessary to reach definitive conclusions.

## Future directions

As for the forthcoming directions of the project, it is necessary to contrast these findings with additional datasets to determine whether ADM events generate distinct and diverse patterns of CNV within the genomic landscape, and whether these patterns are comparable to those identified in this study.

Also, analysis based on scRNA-seq is limited in certain aspects. Due to the inherent nature of the data, possible CNVs in non-coding regions are undetected. Moreover, gene expression levels are highly regulated by other events, such as epigenetic variation, which makes it difficult to differentiate between CNV-driven expression changes and other sources. Thus, alternative approaches are being explored, including the use of scATAC-seq data. This alternative facilitates higher coverage, takes into consideration epigenetic diversity and can be integrated with scRNA-seq 37,38.

# REFERENCES

1. Badowska-Kozakiewicz, A. *et al.* Diabetes Mellitus and Pancreatic Ductal Adenocarcinoma—Prevalence, Clinicopathological Variables, and Clinical Outcomes. *Cancers* **14**, 2840 (2022).

2. Jamal-Hanjani, M. *et al.* Tracking the Evolution of Non–Small-Cell Lung Cancer. *N. Engl. J. Med.* **376**, 2109–2121 (2017).

3. Davoli, T., Uno, H., Wooten, E. C. & Elledge, S. J. Tumor aneuploidy correlates with markers of immune evasion and with reduced response to immunotherapy. *Science* **355**, eaaf8399 (2017).

4. Notta, F. *et al.* A renewed model of pancreatic cancer evolution based on genomic rearrangement patterns. *Nature* **538**, 378–382 (2016).

5. Chan-Seng-Yue, M. *et al.* Transcription phenotypes of pancreatic cancer are driven by genomic events during tumor evolution. *Nat. Genet.* **52**, 231–240 (2020).

6. Pervin, J. *et al.* Clinically impactful metabolic subtypes of pancreatic ductal adenocarcinoma (PDAC). *Front. Genet.* **14**, (2023).

7. Fang, Y. *et al.* Single-Cell Transcriptome Reveals the Metabolic and Clinical Features of a Highly Malignant Cell Subpopulation in Pancreatic Ductal Adenocarcinoma. *Front. Cell Dev. Biol.* **10**, (2022).

8. Herreros-Villanueva, M., Hijona, E., Cosme, A. & Bujanda, L. Mouse models of pancreatic cancer. *World J. Gastroenterol.* **18**, 1286–1294 (2012).

9. Makohon-Moore, A. P. *et al.* Precancerous neoplastic cells can move through the pancreatic ductal system. *Nature* **561**, 201–205 (2018).

10. Guerra, C. *et al.* Pancreatitis-Induced Inflammation Contributes to Pancreatic Cancer by Inhibiting Oncogene-Induced Senescence. *Cancer Cell* **19**, 728–739 (2011).

11. Kopp, J. L. *et al.* Identification of Sox9-Dependent Acinar-to-Ductal Reprogramming as the Principal Mechanism for Initiation of Pancreatic Ductal Adenocarcinoma. *Cancer Cell* **22**, 737–750 (2012).

12. Wei, D. *et al.* KLF4 Is Essential for Induction of Cellular Identity Change and Acinar-to-Ductal Reprogramming during Early Pancreatic Carcinogenesis. *Cancer Cell* **29**, 324–338 (2016).

13. Alonso-Curbelo, D. *et al.* A gene-environment-induced epigenetic program initiates tumorigenesis. *Nature* **590**, 642–648 (2021).

14. Del Poggetto, E. *et al.* Epithelial memory of inflammation limits tissue damage while promoting pancreatic tumorigenesis. *Science* **373**, eabj0486 (2021).

15. Gauthier, B. R. & Comaills, V. Nuclear Envelope Integrity in Health and Disease: Consequences on Genome Instability and Inflammation. *Int. J. Mol. Sci.* **22**, 7281 (2021).

16. Gauthier, B. R., Lorenzo, P. I. & Comaills, V. Physical Forces and Transient Nuclear Envelope Rupture during Metastasis: The Key for Success? *Cancers* **14**, 83 (2021).

17. Marstrand-Daucé, L. *et al.* Acinar-to-Ductal Metaplasia (ADM): On the Road to Pancreatic Intraepithelial Neoplasia (PanIN) and Pancreatic Cancer. *Int. J. Mol. Sci.* **24**, 9946 (2023).

18. Greten, F. R. & Grivennikov, S. I. Inflammation and Cancer: Triggers, Mechanisms, and Consequences. *Immunity* **51**, 27–41 (2019).

19. Ioannidou, A., Goulielmaki, E. & Garinis, G. A. DNA Damage: From Chronic Inflammation to Age-Related Deterioration. *Front. Genet.* **7**, (2016).

20. Ohnishi, S. *et al.* DNA damage in inflammation-related carcinogenesis and cancer stem cells. *Oxid. Med. Cell. Longev.* **2013**, 387014 (2013).

21. Zack, T. I. *et al.* Pan-cancer patterns of somatic copy number alteration. *Nat. Genet.* **45**, 1134–1140 (2013).

22. Yates, L. R. *et al.* Genomic Evolution of Breast Cancer Metastasis and Relapse. *Cancer Cell* **32**, 169-184.e7 (2017).

23. Shen, H. *et al.* Integrated Molecular Characterization of Testicular Germ Cell Tumors. *Cell Rep.* **23**, 3392–3406 (2018).

24. Durante, M. A. *et al.* Single-cell analysis reveals new evolutionary complexity in uveal melanoma. *Nat. Commun.* **11**, 496 (2020).

25. Comaills, V. *et al.* Genomic Instability Is Induced by Persistent Proliferation of Cells Undergoing Epithelial-to-Mesenchymal Transition. *Cell Rep.* **17**, 2632–2647 (2016).

26. Guerrero-Martínez, J. A., Ceballos-Chávez, M., Koehler, F., Peiró, S. & Reyes, J. C. TGFβ promotes widespread enhancer chromatin opening and operates on genomic regulatory domains. *Nat. Commun.* **11**, 6196 (2020).

27. McDonald, O. G., Wu, H., Timp, W., Doi, A. & Feinberg, A. P. Genome-scale epigenetic reprogramming during epithelial-to-mesenchymal transition. *Nat. Struct. Mol. Biol.* **18**, 867–874 (2011).

28. Hatch, E. M., Fischer, A. H., Deerinck, T. J. & Hetzer, M. W. Catastrophic nuclear envelope collapse in cancer cell micronuclei. *Cell* **154**, 47–60 (2013).

29. Ly, P. *et al.* Selective Y centromere inactivation triggers chromosome shattering in micronuclei and repair by non-homologous end joining. *Nat. Cell Biol.* **19**, 68–75 (2017).

30. Mazin, P. *Schard: Reticulate-Free Single Cell Object Conversion*. (2024).

31. Tickle, T., Tirosh, I., Georgescu, C., Brown, M. & Haas, B. *inferCNV of the Trinity CTAT Project.* (Klarman Cell Observatory, Broad Institute of MIT and Harvard, Cambridge, MA, USA, 2019).

32. Hao, Y. *et al.* Dictionary learning for integrative, multimodal and scalable single-cell analysis. *Nat. Biotechnol.* (2023) doi:10.1038/s41587-023-01767-y.

33. Xu, S. *et al.* Using clusterProfiler to characterize multiomics data. *Nat. Protoc.* **19**, 3292–3320 (2024).

34. Huang, R. & Zhou, P.-K. DNA damage repair: historical perspectives, mechanistic pathways and clinical translation for targeted cancer therapy. *Signal Transduct. Target. Ther.* **6**, 1–35 (2021).

35. Mackenzie, K. J. *et al.* cGAS surveillance of micronuclei links genome instability to innate immunity. *Nature* **548**, 461–465 (2017).

36. Gao, Q. *et al.* The genomic landscape of acute lymphoblastic leukemia with intrachromosomal amplification of chromosome 21. *Blood* **142**, 711–723 (2023).

37. Ramakrishnan, A. *et al.* epiAneufinder identifies copy number alterations from single-cell ATAC-seq data. *Nat. Commun.* **14**, 5846 (2023).

38. Moore, T. W. & Yardımcı, G. G. Robust CNV detection using single-cell ATAC-seq. 2023.10.04.560975 Preprint at https://doi.org/10.1101/2023.10.04.560975 (2023).