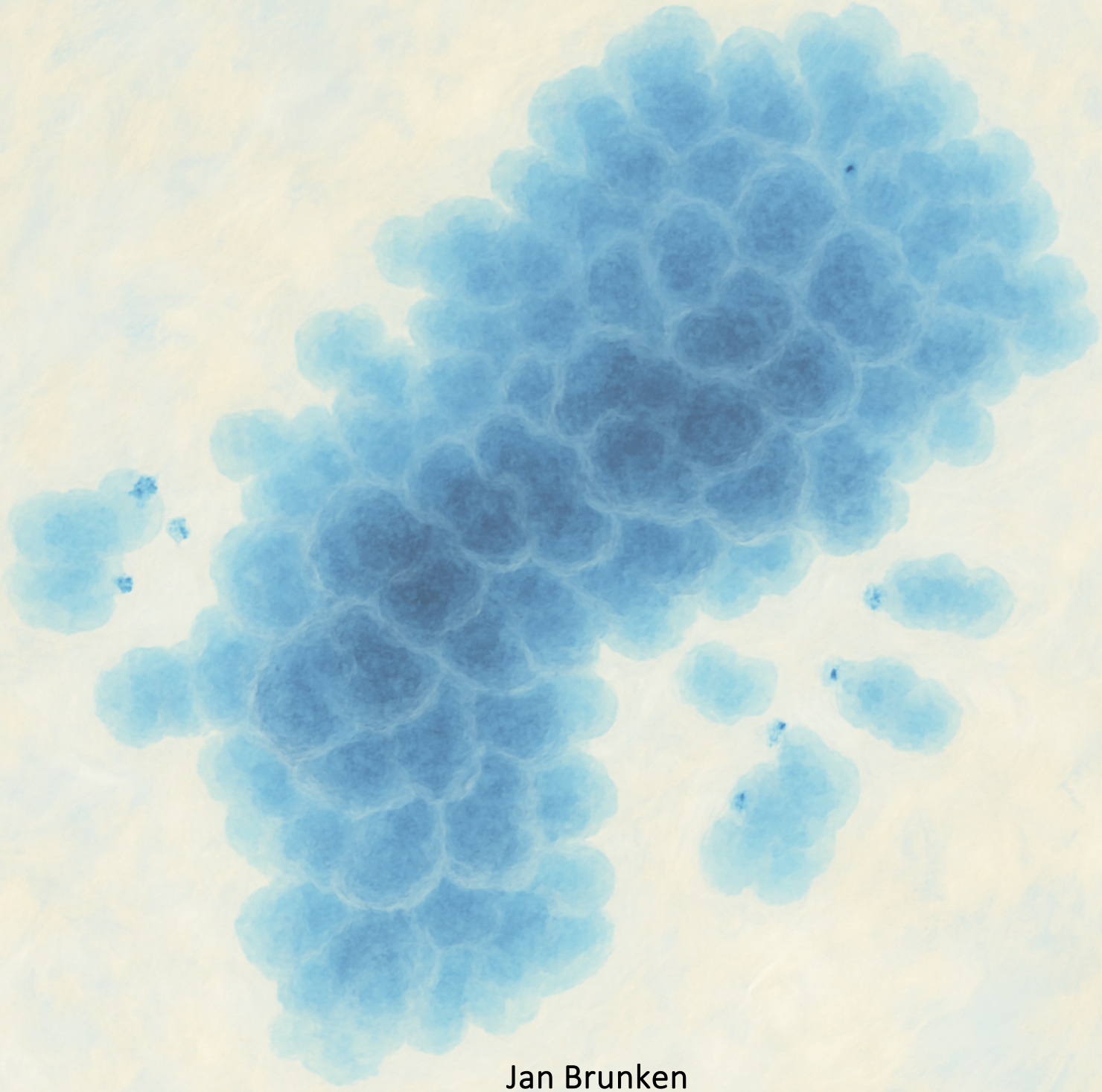
Tissue Remodeling Promotes Polyploid Cell Proliferation and Chromosomal Instability in Exocrine Glands



Jan Brunken

Inaugural dissertation

for

obtaining the doctoral degree

of the

Combined Faculty of Mathematics, Engineering and Natural Sciences

of the

Ruprecht-Karls-University

Heidelberg

Presented by

Jan Brunken, M. Sc.

born in Göttingen, Germany

Oral Examination: June 23, 2025

Tissue Remodeling Promotes Polyploid Cell Proliferation and Chromosomal Instability in Exocrine Glands

Referees: Prof. Dr. Ana Martin-Villalba

Dr. Michael Milsom

“Nachdem nun einerseits gezeigt werden konnte, dass mehrpolige Mitosen zur Entstehung solcher aus ihrem Gleichgewicht geratener Zellen führen, und nachdem andererseits bekannt ist, dass in Geschwülsten simultane Mehrteilungen vorkommen, wird die Hypothese eines Zusammenhanges beider Erscheinungen einer Prüfung wert sein. Freilich müsste dabei noch vorausgesetzt werden, dass nicht nur in der entwickelten, sondern schon in der entstehenden Geschwulst mehrpolige Mitosen vorkämen. Wodurch diese selbst verursacht werden, wäre eine zweite Frage…”

Theodor Boveri, Über Mehrpolige Mitosen als Mittel zur Analyse des Zellkerns, 1902

Summary

Chromosomal instability (CIN) is a hallmark of cancer, implicated in both tumor initiation and progression. Yet, its origins in normal, untransformed tissues remain poorly understood. Polyploidy is a feature of many adult tissues – particularly in exocrine glands such as the pancreas and the mammary gland. Although polyploidy has been linked to CIN in malignancies, the role of polyploidy in mutation acquisition within naïve tissues remains largely unexplored.

In this thesis, I investigate the link between polyploidy and CIN during tissue remodeling, utilizing organoid and *in vivo* models of pancreatic injury. I demonstrate that polyploid acinar cells contribute to regeneration by undergoing acinar-to-ductal metaplasia. This transition involves extensive cellular remodeling, including cell shrinkage, which disrupts proper spindle orientation, especially in binucleated, polyploid acinar cells. These spindle defects result in mitotic errors such as lagging chromosomes, chromatin bridges, and the formation of micronuclei. Crucially, I observe extensive DNA damage within these micronuclei, consistent with chromothripsis, revealing a pathway to CIN in untransformed cells.
To explore the broader relevance of this mechanism, I employed a lactating mammary gland organoid model, that was developed under my supervision. In this model, polyploidy was induced through pregnancy hormone treatment mimicking transitions that occur to prepare the tissue for milk production during lactation. Strikingly, this model recapitulated the mitotic abnormalities and CIN observed in the pancreas, suggesting that polyploidy-associated chromosomal instability is a conserved feature of glandular tissues. Notably, all findings were obtained in primary, untransformed cells without genetic manipulation or genotoxic stress, supporting a physiological origin of CIN linked to polyploid cell division.

To acquire these findings, I developed an imaging and image analysis workflow to quantitatively assess ploidy states, nuclear number, micronuclei frequency, and expression levels of key markers. This workflow is also detailed in the thesis.

Together, my results uncover a novel mechanism by which normal polyploid cells contribute to CIN during tissue regeneration. This work provides new insights into the origins of chromosomal instability in regenerative contexts, highlighting the dual role of polyploidy in promoting both tissue repair and chromosomal mis-segregation.

Zusammenfassung

Chromosomeninstabilität (CIN) ist ein Kennzeichen von Krebs, das sowohl bei der Tumorentstehung als auch beim Fortschreiten des Tumors eine Rolle spielt. Die Ursprünge dieser Instabilität in normalen, nicht transformierten Geweben sind jedoch nach wie vor kaum bekannt. Polyploidie ist ein Merkmal vieler adulter Gewebe - insbesondere in exokrinen Drüsen wie der Bauchspeicheldrüse und der Brustdrüse. Obwohl Polyploidie mit CIN bei bösartigen Erkrankungen in Verbindung gebracht wurde, ist die Rolle der Polyploidie bei der Mutagenese in naiven Geweben noch weitgehend unerforscht.

In dieser Arbeit untersuche ich den Zusammenhang zwischen Polyploidie und CIN während des Gewebeumbaus unter Verwendung von Organoid- und *In-vivo*-Modellen der Pankreasverletzung. Ich zeige, dass polyploide Azinuszellen zur Regeneration beitragen, indem sie eine azinär-duktale Metaplasie durchlaufen. Dieser Übergang ist mit einem umfassenden zellulären Umbau verbunden, zu dem auch die Schrumpfung der Zellen gehört, wodurch die korrekte Spindelausrichtung gestört wird, insbesondere bei zweikernigen, polyploiden Azinuszellen. Diese Spindeldefekte führen zu mitotischen Fehlern wie langsame Folgechromosomen, Chromatinbrücken und der Bildung von Mikronuklei. Entscheidend ist, dass ich innerhalb dieser Mikronuklei umfangreiche DNA-Schäden beobachte, die mit Chromothripsis übereinstimmen und einen Weg zu CIN in nicht transformierten Zellen aufzeigen.
Um die allgemeine Bedeutung dieses Mechanismus zu untersuchen, habe ich ein Organoidmodell der laktierenden Brustdrüse verwendet, das unter meiner Aufsicht entwickelt wurde. In diesem Modell wurde die Polyploidie durch eine Behandlung mit Schwangerschaftshormonen induziert, die die Übergänge nachahmt, die zur Vorbereitung des Gewebes auf die Milchproduktion während der Laktation stattfinden. Auffallend ist, dass dieses Modell die in der Bauchspeicheldrüse beobachteten mitotischen Anomalien und CIN rekapitulierte, was darauf hindeutet, dass die mit Polyploidie verbundene chromosomale Instabilität ein konserviertes Merkmal von Drüsengeweben ist. Bemerkenswert ist, dass alle Ergebnisse in primären, nicht transformierten Zellen ohne genetische Manipulation oder genotoxischen Stress erzielt wurden, was für einen physiologischen Ursprung von CIN in Verbindung mit polyploider Zellteilung spricht.

Um diese Erkenntnisse zu gewinnen, habe ich einen Arbeitsablauf für die Bildgebung und Bildanalyse entwickelt, um den Ploidiezustand, die Kernzahl, die Mikronuklei-Häufigkeit und die

Expressionsniveaus wichtigster Marker quantitativ zu bestimmen. Auch dieser Arbeitsablauf wird in der Dissertation ausführlich beschrieben.

Zusammengenommen decken meine Ergebnisse einen neuartigen Mechanismus auf, durch den normale polyploide Zellen zur CIN während der Geweberegeneration beitragen. Diese Arbeit bietet neue Einblicke in den Ursprung der chromosomalen Instabilität in regenerativen Kontexten und unterstreicht die doppelte Rolle der Polyploidie bei der Förderung der Gewebereparatur und der chromosomalen Fehlaufspaltung.

Table of Contents

List of figures

List of tables

List of abbreviations

2D	two-dimensional
3D	three-dimensional
4D	four-dimensional
α-Amy	α-Amylase
ADM	Acinar-to-ductal metaplasia
AFP	α-fetoprotein
AI	Artificial intelligence
ANKRD26	Ankyrin repeat domain-containing protein 26
Annot	Annotation
ANOVA	Analysis of variance
ATM	Ataxia telangiectasia mutated
ATR	Ataxia telangiectasia and Rad 3 related
AURKA	Aurora kinase A
β-Cas	β-Casein
BFB	Breakage-fusion-bridge
Binuc	Binucleated
BRCA2	Breast cancer 2 gene
BrdU	Bromodeoxyuridine
BSA	Bovine serum albumin
CC	Crick-Crick orientation of DNA reads
CCK	Cholecystokinin
CDK1	Cyclin-dependent kinase 1
CDKN2A	Cyclin-dependent kinase inhibitor 2A gene
CDT1	Chromatin licensing and DNA replication factor 1
cGAS	Cyclic GMP-AMP synthase
Chr	Chromosome
CIN	Chromosomal instability
CiP	Cerulein-induced pancreatitis
CK19	Cytokeratin 19
CNN	Convolutional neural network
CTCF	Corrected total cellular fluorescence
CMnCF	Corrected mean cellular fluorescence

CW	Crick-Watson orientation of DNA reads
d	days (referring to days in cell culture)
DAPI	4’,6-diamidino-2-phenylindole
DCLK1	Doublecortin-like kinase 1
DEN	Diethylnitrosamine
DKFZ	Deutsches Krebsforschungszentrum/German Cancer Research Center
dL	Days of lactation
DMEM	Dulbecco’s modified eagle medium
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid
dP	Days of pregnancy
DPBS	Dulbecco’s phosphate buffered saline
E_2	17β-Estradiol/Estrogen
E-Cad	E-Cadherin
eccDNA	Extrachromosomal circular DNA
EDTA	Ethylenediaminetetraacetate
EGF	Epidermal growth factor
EGFP	Enhanced green fluorescent protein
EGFR	Epidermal growth factor receptor
EMBL	European Molecular Biology Laboratory
ER	Estrogen receptor
F12	F12 nutrient mix
FACS	Fluorescence-activated cell sorting
FBS	Fetal bovine serum
FGF	Fibroblast growth factor
FSC-A	Forward-scatter area
FSC-H	Forward-scatter height
FUCCI2	Fluorescent ubiquitylation-based cell cycle indicator 2
g	gravitational force equivalent
G0	Gap 0 phase/resting state
G1	Gap 1 phase of cell cycle
G2	Gap 2 phase of cell cycle
GFP	Green fluorescent protein
GPU	Graphics processing unit

H2A	Histone H2A
γH2A.X	Serine-139 phosphorylated histone H2A family member X
H2B	Histon H2B
HBSS	Hank's balanced salt solution
HER2/neu	Human epidermal growth factor receptor
HNF1β	Hepatocyte nuclear factor 1 beta
HOG	Histogram of oriented gradients
ICC	Immunocytochemistry
IF	Immunofluorescence
IGF-1	Insulin-like growth factor 1
IHC	Immunohistochemistry
i.p.	Intraperitoneal
ITS-X	Insulin-transferrin-selenium-ethanolamine
K5	Keratin 5/Cytokeratin5
K8	Keratin 8/Cytokeratin 8
kb	Kilobases
KRAS	Kirsten rat sarcoma viral oncogene homologue gene
$KRAS^{G12D}$	Point mutation in codon 12 of the *KRAS* gene (glycine to aspartic acid)
LBP	Local Binary Patterns
M	Mitosis phase of cell cycle
Mb	Megabases
MaxIP	Maximum intensity projection
MCM	Minichromosome maintenance
MDM2	Murine double minute 2
mG	Membrane-bound green fluorescent protein
MLP	Multilayer perceptron
MN	Micronucleus
MOM	Morphogenesis medium
Mononuc	Mononucleated
NA	Numerical aperture
NEB	Nuclear envelope breakdown
NFκB	Nuclear factor kappa B
NK	Natural killer
NSB	Nuclei sorting buffer

P_4	Progesterone
p53	Tumor protein 53
PABC	Pregnancy-associated breast cancer
PanIN	Pancreatic intraepithelial neoplasia
PBS	Phosphate-buffered saline
PDAC	Pancreatic ductal adenocarcinoma
PDX1	Pancreatic and duodenal homeobox 1
PGCC	Polyploid giant cancer cells
PHx	Partial hepatectomy
pHH3	Serine-10 phosphorylated histone H3
PI	Propidium iodide
PIDD	P53-induced death domain protein
PLM	Pregnancy-lactation medium
PMT	Photomultiplier tube
POC	Pancreatic organoid culture (medium)
PPBC	Postpartum breast cancer
PR	Progesterone receptor
Prl	Prolactin
PTEN	Phosphatase and tenson homolog gene
PTF1A	Pancreatic transcription factor subunit alpha
px	Pixel
RAIDD	RIP-associated protein with a death domain
RANKL	Receptor activator of nuclear factor kappa B ligand
RBC	Red blood cell
REG	Regenerating protein family
ReLu	Rectified linear unit
rhEGF	Recombinant human epidermal growth factor
rhFGF2	Recombinant human fibroblast growth factor 2
RT	Room temperature
S	Synthesis phase of cell cycle
SAC	Spindle assembly checkpoint
SMAD4	Mothers against decapentaplegic homolog 4 gene
SOX9	SRY-box transcription factor 9
SSC-A	Side-scatter area

STAT3	Signal transducer and activator of transcription 3
STING	Stimulator of interferon genes
SumIP	Sum intensity projection
SV	Structural variant
SVM	Support vector machine
TAP	Transient amplifying progenitor
TC	Tubular complex
TEB	Terminal end bud
Tff2	Trefoil factor 2
TP53	Tumor protein 53 gene
TNBC	Triple-negative breast cancer
WF	Watson fraction
Wnt	Wingless Int-1
WT	Wildtype
WW	Watson-Watson read orientation

1 Introduction

The question of what causes cancer has fascinated and perplexed humankind for millennia. In ancient times, it was attributed to disbalanced bodily fluids or divine retribution. With the rise of scientific medicine in the 19th century, figures like Theodor Boveri and Rudolf Virchow fundamentally transformed our understanding of cancer, framing it as a disease of cells. Virchow's insight that "omnis cellula e cellula" established the cellular basis of pathology, while Boveri's visionary hypothesis that chromosomal aberrations could drive tumor formation laid the conceptual groundwork for modern cancer genetics (Boveri, 1914; Virchow, 1858). Over the past century, and particularly in recent decades, advances in molecular biology, genomics, and imaging technologies have revolutionized our ability to probe the molecular mechanisms underlying cancer. And yet, despite this technological progress, our fundamental understanding of how chromosomal instability arises in normal tissues remains incomplete. Much of what we know about chromosomal instability is derived from highly artificial models: cancer cell lines grown in vitro, genetically manipulated organisms, or experimental conditions far removed from physiological reality. As a result, key questions remain: Can chromosomal instability emerge in unperturbed tissues? What cellular states or structural features predispose certain cells to genomic derailment? And how do these instabilities unfold in the early, often silent, stages of disease development? This thesis aims to explore these questions by investigating the origins of chromosomal instability in polyploid cells within naïve, unmanipulated glandular tissues, shedding light on a fundamental, yet largely overlooked, aspect of tumor initiation.

1.1 Exocrine glands

In mammals, exocrine glands are epithelial-derived structures specialized for the production and secretion of various substances onto external or internal epithelial surfaces via ductal systems. These secretions – including enzymes, mucus, sweat, lipids, and milk – fulfill vital physiological roles, supporting digestion, thermoregulation, surface protection, and reproductive processes. Exocrine glands range from unicellular forms, such as mucus-secreting goblet cells, to complex multicellular organs like the salivary glands, pancreas, mammary glands, and liver. The regulation of exocrine gland function is mediated through coordinated neural and hormonal signals, ensuring secretion is precisely timed and localized.

In both humans and mice, the developmental architecture, molecular regulation, and secretory functional outputs of exocrine glands are broadly conserved. The fundamental developmental

pathways controlling exocrine gland formation – involving signaling cascades such as fibroblast growth factor (FGF), epidermal growth factor (EGF), and the activity of transcription factors including hepatocyte nuclear factor 1β (HNF1β), and SRY-box transcription factor 9 (SOX9), are highly conserved across species (W. Guo et al., 2012; Jennings et al., 2015; McNally & Stein, 2017; Tanaka et al., 2021). These similarities have positioned the mouse as a central model for elucidating mechanisms of exocrine gland morphogenesis, homeostasis, regeneration, and disease.

1.1.1 The pancreas

1.1.1.1 Anatomy and function of the pancreas

The pancreas functions as a multifunctional organ, fulfilling both endocrine and exocrine roles. The endocrine compartment, composed of hormone-producing cells, is organized into discrete clusters known as pancreatic islets, which are embedded within the surrounding exocrine tissue. Despite comprising only 1-2 % of the total pancreatic volume, the islets, particularly the insulin-secreting beta cells, have been the subject of intense research focus largely due to their critical role in glucose homeostasis and their involvement in the pathogenesis of diabetes mellitus (Ionescu-Tirgoviste et al., 2015). In contrast, many fundamental aspects of the exocrine pancreas, including its development, disease mechanisms, and regenerative capacity, remain relatively underexplored.

The exocrine pancreas is characterized by a highly branched ductal network organized into interconnected lobes and lobules. It produces and transports a digestive fluid rich in enzymes essential for nutrient breakdown. This fluid is conveyed through a hierarchical ductal system comprising the main pancreatic duct, interlobular ducts, intralobular ducts, and intercalated ducts before being discharged into the duodenum via the common bile duct. Along the ductal system, epithelial cells secrete bicarbonate ions into the lumen, neutralizing gastric acid and creating a pH environment favorable for enzymatic activity (Longnecker & Thompson, 2023).

Pancreatic acinar cells sit the termini of the intercalated ducts and are organized into single-layered, grape-like structures known as acini (Figure 1.1). Acinar cells are the predominant cell type of the pancreas and are specialized for the production and apical storage of digestive enzymes within secretory granules. Remarkably, they exhibit the highest rate of protein synthesis among all adult mammalian cell types (Case, 1978). Upon stimulation via the hormone cholecystokinin (CCK), these enzymes are released via exocytosis into the ductal lumen through the acinar cell's apical membrane (Cheng et al., 2020).

Specialized ductal cells, termed centroacinar cells, extend from the terminal ducts into the center of each acinus, forming a critical link between acinar and ductal compartments

(Longnecker & Thompson, 2023). In histological sections, acinar and ductal cells are distinguishable by their size and morphology: acinar cells are approximately 20 µm in diameter with abundant cytoplasm, whereas ductal cells are smaller (approximately 5 µm) with nuclei that occupy the majority of the cell's cross-sectional area (Motta et al., 1997). Molecularly, ductal cells are identified by the expression of markers such as cytokeratin 19 (CK19), HNF1β, and SOX9, while acinar cells are characterized by the expression of digestive enzymes like α-amylase and the transcription factor Pancreas transcription factor 1 subunit alpha (PTF1A).

Finally, pancreatic stellate cells locate between pancreatic lobules and acinar cells and fulfill the role of fibroblast-like cell to produce and regulate extracellular matrix components.

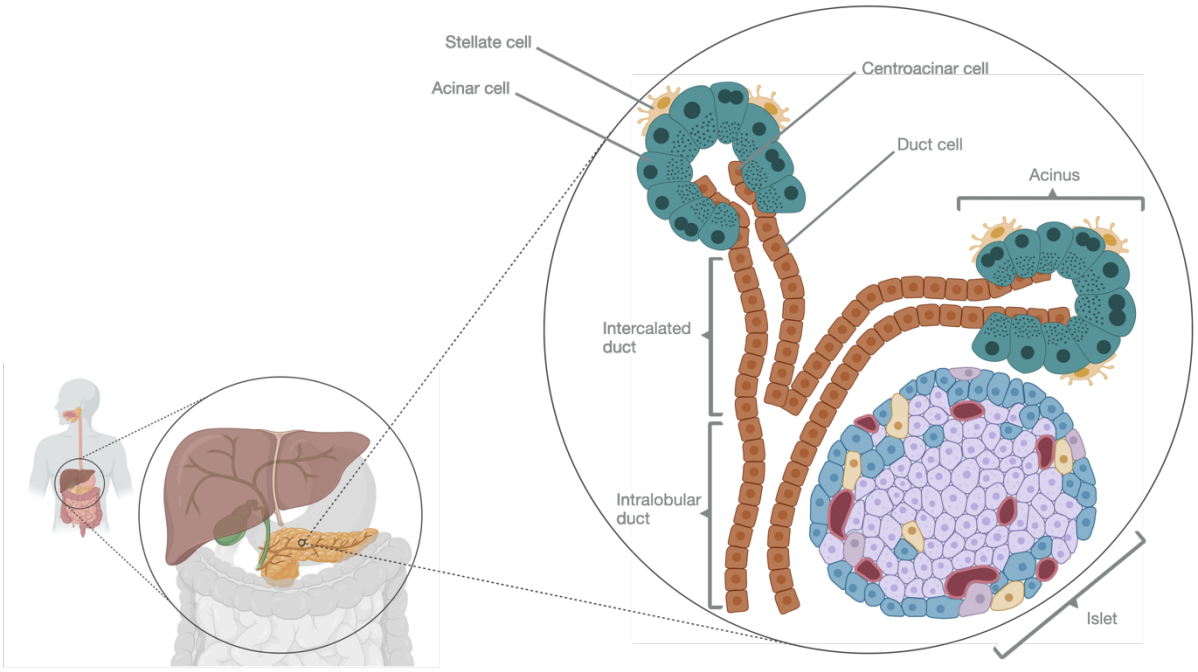


Figure 1.1 | Pancreas anatomy.

Hormone producing endocrine cells are organized in pancreatic islets. The exocrine compartment consists of enzyme producing acinar cells and a ductal network that channels the pancreatic juice into the duodenum via the common bile duct. Several acinar cells form a grape-like acinus. Digestive enzymes are stored in apically located zymogen granules of acinar cells and get secreted into intercalated ducts via the acinar cell's apical membrane. Parts of the figure created with BioRender.com.

1.1.1.2 Clonal and transcriptional heterogeneity of acinar cells

To sustain tissue integrity and ensure recovery from both daily cellular attrition and injury-induced loss, multicellular organisms have evolved robust mechanisms of tissue replenishment and homeostatic maintenance. Typically, this regenerative capacity is mediated by somatic stem cells and progenitor populations, which possess the ability to self-renew and differentiate into

specialized cell types. A prevailing organizational paradigm across many tissues involves the localization of stem cells within specialized microenvironments known as niches, which maintain stem cells in a quiescent and undifferentiated state under homeostatic conditions (Moore & Lemischka, 2006). Upon tissue injury, stem cells are activated, exit quiescence, and proliferate to regenerate the damaged tissue. Prominent examples of adult stem cell niches include the subventricular zone in the brain, the hematopoietic stem cell niche in the bone marrow, and the bulge region of hair follicles (Doetsch, 2003; Orkin & Zon, 2008; Tumbar et al., 2004).

Alternative models of tissue regeneration have also been documented. In certain organs, such as articular cartilage and the kidney, tissue repair, although with rather limited capacities, predominantly occurs through the proliferation of pre-existing, differentiated cells without a requisite stem cell population, (Benigni et al., 2010; Goldring & Marcu, 2009). Moreover, recent studies have highlighted the role of dedifferentiation, wherein mature, post-mitotic somatic cells revert to a progenitor-like or multipotent state. These dedifferentiated cells and their progeny may subsequently either redifferentiate into their original lineage or undergo transdifferentiation into other distinct cell types, thereby contributing to tissue repair and regeneration (Iismaa et al., 2018; Varga & Greten, 2017).

In contrast to organs with well-defined stem cell niches, the adult exocrine pancreas lacks a clearly identifiable somatic stem cell population. Instead, the acinar cell compartment, which constitutes the majority of the exocrine pancreas, is maintained by the replication of differentiated acinar cells. Multiple studies have demonstrated that, under homeostatic conditions, acinar cells are lineage-restricted and exhibit little evidence of multipotency (Desai et al., 2007; Wollny et al., 2016). Utilizing lineage tracing approaches, Wollny and colleagues revealed clonal heterogeneity within the acinar cell population, identifying a subset of long-lived, unipotent cells responsible for acinar maintenance (Wollny et al., 2016). Upon injury, a distinct subset of quiescent cells becomes activated to compensate for cell loss, indicating a functional division of labor within the acinar compartment. Thus, the population of otherwise quiescent regenerative acinar cells can be considered as a pool of facultative progenitor cells (Z. Jiang et al., 2020).

Single-cell transcriptomic profiling facilitated the identification of stathmin-1 (STMN1), a microtubule destabilizer, as a marker associated with both inherent and injury-induced proliferative capacity in acinar cells (Wollny et al., 2016). Additionally, doublecortin-like kinase 1 (DCLK1), another microtubule regulator, was identified as a marker of the facultative progenitor population that would account for injury-induced cell loss (Westphalen et al., 2016). More

recently, high-throughput sequencing technologies gave more insights into the transcriptional heterogeneity within acinar cells. Totsi et al. identified three distinct populations, consisting of acinar-s (secretory), acinar-i (idling) and acinar-Reg+ cells (Tosti et al., 2021). While acinar-s cells possess pronounced secretory capabilities (~50 % of their transcripts encode digestive enzymes), acinar-i cells show gene enrichment in insulin and glucagon signaling and are found near pancreatic islets. Acinar-Reg+ cells are defined by high expression of the regenerating (REG) family of genes which are associated with inflammation and regeneration. Their proximity to macrophages and lack of progenitor gene signature hints at a paracrine function in immune modulation upon injury. Another study identified transient amplifying progenitors (TAPs) marked by trefoil factor 2 (Tff2) which contribute to homeostasis but do not possess regenerative capacities upon injury (Z. Jiang et al., 2023). A link between these progenitors and the distinct proliferative population of acinar cells during homeostasis found by Wollny et al., 2016 seems plausible.

1.1.1.3 Tissue remodeling: acinar cell plasticity and regeneration

Among pancreatic pathologies, pancreatitis, characterized by inflammation of the pancreas, is the most common cause of injury to the exocrine pancreas. Pancreatitis may present as either an acute or chronic condition and can be precipitated by a variety of etiologies, including gallstones, chronic alcohol consumption, traumatic injury, and genetic predispositions. Regardless of the initiating factor, pancreatitis typically results from aberrant intracellular activation of digestive enzymes, leading to autodigestion, inflammation, and acinar cell death (Motta et al., 1997). In response to such injury, acinar tissue regeneration is mediated predominantly by the proliferation of surviving acinar cells, with no evidence for lineage conversion from non-acinar cells (Desai et al., 2007). Importantly, despite the regenerative process, the acinar lineage remains lineage-restricted, showing no multipotent differentiation toward other pancreatic cell types.

Nevertheless, acinar cell regeneration is accompanied by significant cellular plasticity. A proportion of acinar cells transiently dedifferentiates following injury, adopting a less specialized phenotype and forming structures termed tubular complexes (de la Porte et al., 1991). These duct-like structures exhibit molecular features associated with pancreatic ductal cells, including upregulation of CK19 and SOX9, and concomitant downregulation of acinar-specific markers such as those encoding digestive enzymes (Means et al., 2005; Pinho et al., 2011). This process is referred to as acinar-to-ductal metaplasia (ADM). The functional identity of dedifferentiated acinar cells remains a topic of ongoing investigation. The classical definition of metaplasia as

conversion from one differentiated tissue type into another requires the transdifferentiation from one mature cell type to another. In contrast dedifferentiation is rather associated to the loss of a specialized function toward a less differentiated identity. In pancreas biology literature, the terms “ADM”, “acinar cell transdifferentiation” and “acinar cell dedifferentiation” are often used interchangeably, which is probably because ADM can involve processes of both, cell identity loss and cell identity conversion (Grimont et al., 2022; Marstrand-Daucé et al., 2023; Parte et al., 2022). However, for the rest of this thesis, I will stick to the most used term ADM. Rather than fully adopting a ductal phenotype, ADMs exhibit characteristics akin to embryonic pancreatic progenitor cells, expressing developmental markers such as Nestin and pancreatic and duodenal homeobox 1 (PDX1; Jensen et al., 2005; Means et al., 2005). Consequently, ADM is increasingly regarded as a form of injury-induced developmental reprogramming, reactivating latent embryonic gene expression programs to facilitate tissue repair (Storz, 2017).

1.1.1.4 Pancreatic ductal adenocarcinoma

The exocrine compartment of the pancreas serves as the cellular origin for pancreatic ductal adenocarcinoma (PDAC), the most prevalent and aggressive form of pancreatic cancer. With a five-year survival rate ranging between 1% and 4%, PDAC remains among the most lethal malignancies and ranks as the fourth leading cause of cancer-related deaths globally (Siegel et al., 2022). Despite significant advances in cancer therapeutics for other malignancies, the prognosis for unresectable and metastatic PDAC patients has only seen minimal improvement, with median survival times lingering between 4 to 6 months post-diagnosis (Bengtsson et al., 2020; De Dosso et al., 2021). This dismal outlook is attributed to the highly invasive and metastatic nature of PDAC, its asymptomatic early stages leading to delayed diagnosis, and the marked resistance of the tumor to both conventional chemotherapies and emerging therapeutic strategies such as immunotherapies (Ryan et al., 2014).

Given these challenges, there is a critical need to deepen our understanding not only of fully developed tumors but also of the early molecular and cellular events underpinning PDAC initiation and progression. It is widely recognized that PDAC in humans arises through the stepwise malignant transformation of neoplastic precursor lesions (Distler et al., 2014). The most abundant and best characterized of these precursor lesions are pancreatic intraepithelial neoplasias (PanINs), which represent the primary histological precursor to PDAC (Maitra et al., 2005). A striking feature of PanIN and PDAC is the high prevalence of activating point mutations in the *KRAS* gene, particularly within codon 12, observed in over 90% of cases (Kanda et al., 2012).

Preclinical studies utilizing genetically engineered mouse models have been instrumental in elucidating the cellular origins of PDAC. Oncogenic *KRAS* activation in either ductal or acinar cells can drive tumorigenesis (Ferreira et al., 2017; A. Y. L. Lee et al., 2019). However, lineage tracing experiments have demonstrated that ductal cells, upon oncogenic *KRAS* expression, typically fail to initiate PanIN formation (Brembeck et al., 2003; Ferreira et al., 2017). In contrast, acinar cells exhibit a pronounced susceptibility to oncogenic transformation, forming expanded PanIN lesions that subsequently progress to invasive carcinoma (Carrière et al., 2007; Ferreira et al., 2017; Guerra et al., 2007). Thus, current models support acinar cells as the predominant cell-of-origin for PDAC (Kopp et al., 2012; Marstrand-Daucé et al., 2023; Parte et al., 2022).

Under physiological conditions, acinar cells that undergo pancreatitis-induced acinar-to-ductal metaplasia (ADM) typically revert to their differentiated acinar phenotype during tissue regeneration (Jensen et al., 2005). However, the presence of oncogenic *KRAS* mutations, along with aberrant and sustained epidermal growth factor receptor (EGFR) signaling, can "lock" ADMs in their proliferative progenitor-like state (Figure 1.2). This pathological dedifferentiation culminates in persistent ADM lesions, which may then irreversibly progress to PanIN lesions.

PanINs are histologically classified into three grades based on the extent of architectural and cytological atypia: PanIN-1 (low-grade), PanIN-2 (intermediate-grade), and PanIN-3 (high-grade or carcinoma in situ; Pian et al., 2025). Importantly, early PanIN lesions (PanIN-1 and PanIN-2) frequently undergo oncogene-induced senescence, a protective mechanism driven by *KRAS* mutations that halts further neoplastic progression (Collado et al., 2005; Morton et al., 2010; Serrano et al., 1997). However, this senescence barrier can be circumvented by the acquisition of additional genetic alterations, particularly in tumor suppressor genes such as cyclin-dependent kinase inhibitor 2A (*CDKN2A* encoding $p16^{INK4A}$), mothers against decapentaplegic homolog 4 (*SMAD4*), breast cancer 2 (*BRCA2)*, and tumor protein 53 (*TP53)*. These secondary mutations facilitate the progression from senescent PanIN lesions to high-grade PanIN-3 and ultimately to invasive PDAC (Hruban et al., 2000; Morton et al., 2010; Storz, 2017; Wilentz et al., 1998). Another study questioned this gradual model by linking the rapid clinical development of PDAC tumors to single catastrophic events causing the necessary mutations for tumor development – thus exploiting the more recent punctuated equilibrium model of tumorigenesis (Notta et al., 2016).

At present, the role of cellular heterogeneity during the transition from normal acinar cells through ADM and premalignant lesion formation to overt malignancy remains poorly defined. Addressing the molecular, functional, and morphological diversity of individual acinar cells under conditions reflective of early tumorigenesis is crucial to unravel the complex mechanisms

driving PDAC initiation. Such studies hold promise for identifying novel biomarkers and therapeutic targets that could improve early detection and intervention in this devastating disease.

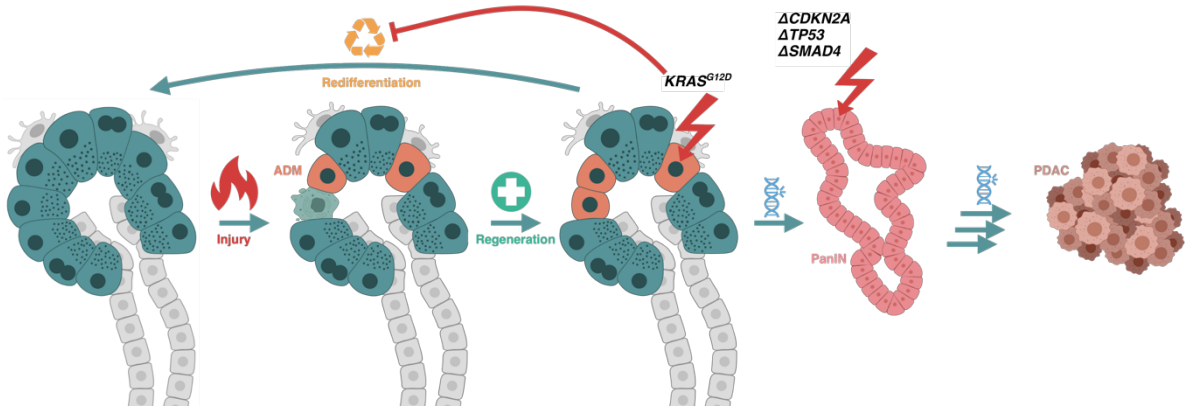


Figure 1.2 | Acinar cell regeneration is linked with carcinogenesis.

During pancreatitis, acinar cells dedifferentiate towards a duct-/progenitor-like phenotype and re-differentiate back into acinar cells as part of the regenerative process. Upon oncogenic EGFR/KRAS signaling, these cells get locked in their dedifferentiated state and progress further into PanIN 1 lesions. *CDKN2A* inactivation is required for progression towards PanIN 2 lesions. Additional mutations in tumor suppressor genes such as *SMAD4*, *BRCA2* and *TP53* may lead to cancerous PanIN 3 lesions and PDAC. Parts of the figure created with BioRender.com.

1.1.1.5 Modelling ADM and PDAC initiation

In order to better understand the processes governing acinar plasticity and ADM formation, the Martin-Villalba lab developed a robust protocol for the isolation and culture of acinar-derived organoids (Wollny et al., 2016). These organoids faithfully recapitulate key features of ADM, including the downregulation of acinar markers (e.g., digestive enzymes) and upregulation of ductal-associated markers (e.g., CK19, SOX9), alongside morphological changes such as the formation of spherical structures with expansive lumina and decreased average cell size (Figure 1.3 A). If cultured for 10 days or longer in presence of EGF, a fraction of these organoids undergoes further morphological transformation including papillary folding, aligned columnar cells accompanied with reduced proliferation – all features of early PanIN lesions (Figure 1.3 B; Brunken, 2019). Thus, acinar-derived organoid models represent a promising platform for dissecting the molecular and cellular mechanisms underlying acinar dedifferentiation, injury-induced plasticity, tissue regeneration, and modelling of early pancreatic cancer development.

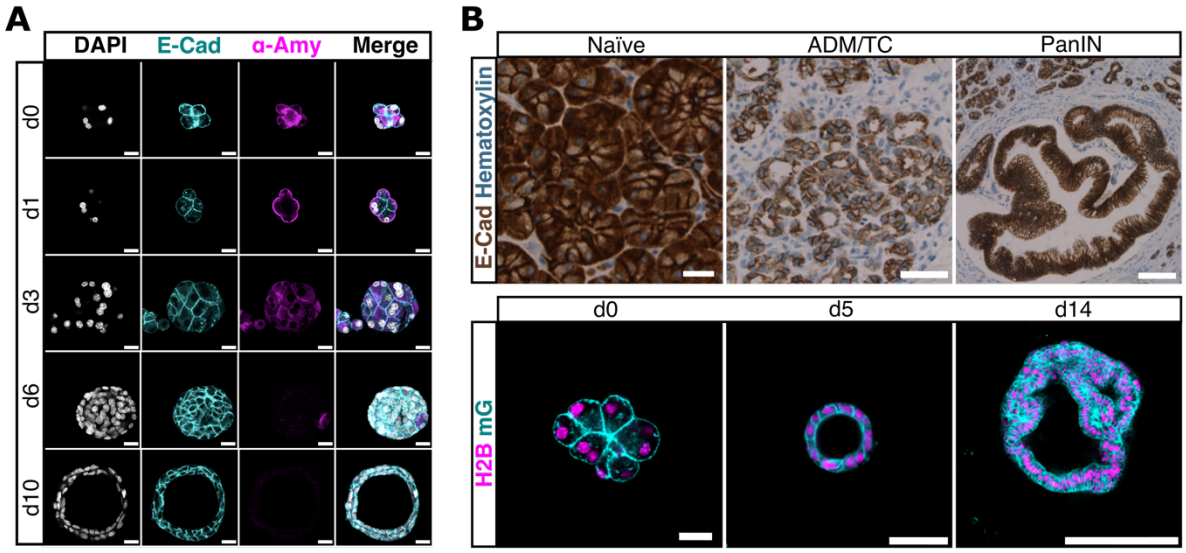


Figure 1.3 | Acinar cell-derived organoids recapitulate features of ADM and PanIN lesions.

(A) Confocal immunofluorescence images of mouse acinar cell derived organoids (mADOs) fixed at different time points in culture (d0, d1, d3, d6, d10) and stained for E-Cadherin (E-Cad, cyan), the acinar cell marker α-Amylase (α-Amy, magenta) and nuclei using 4',6-Diamidino-2-phenylindol (DAPI, white). Scale bars 20 µm. **(B)** Top row: immunohistochemical images of human pancreas tissue sections showing naïve, ADM/TC and PanIN-lesion tissue. E-Cadherin (brown), hematoxylin (blue). Scale bar lengths indicated on images. Bottom row: immunofluorescence images of mouse pancreatic acinar-derived organoids (mADOs). mADO development recapitulates ADM and PanIN formation in vitro. Scale bar 20 µm (left), 50 µm (center), 100 µm (right). All images were adapted from (Brunken, 2019).

1.1.2 The mammary gland

1.1.2.1 Anatomy and function of the mammary gland

The mammary gland is a specialized secretory organ whose primary function is the production and delivery of milk to feed offspring. Structurally, the adult mammary gland consists of a branched ductal tree embedded within a stromal fat pad (Figure 1.4). The epithelial component of the gland is organized into a bilayered structure comprising two distinct cell types: the inner luminal epithelial cells and the outer myoepithelial cells (Biswas et al., 2022).

Luminal cells line the ducts and alveoli and are primarily responsible for milk synthesis and secretion. These cells express characteristic markers such as keratin 8 (K8), which distinguishes them from surrounding cell types, while myoepithelial cells express keratin 5 (Stingl et al., 2006). Myoepithelial cells form a contractile layer surrounding the luminal cells, facilitating the ejection of milk during lactation by compressing the alveoli. Both cell types play crucial roles in maintaining mammary gland architecture and function, and disruptions in their interactions can lead to pathological conditions.

The stromal compartment, composed of adipocytes, fibroblasts, immune cells, and extracellular matrix components, provides mechanical support and biochemical signals essential for glandular homeostasis and remodeling (Biswas et al., 2022).

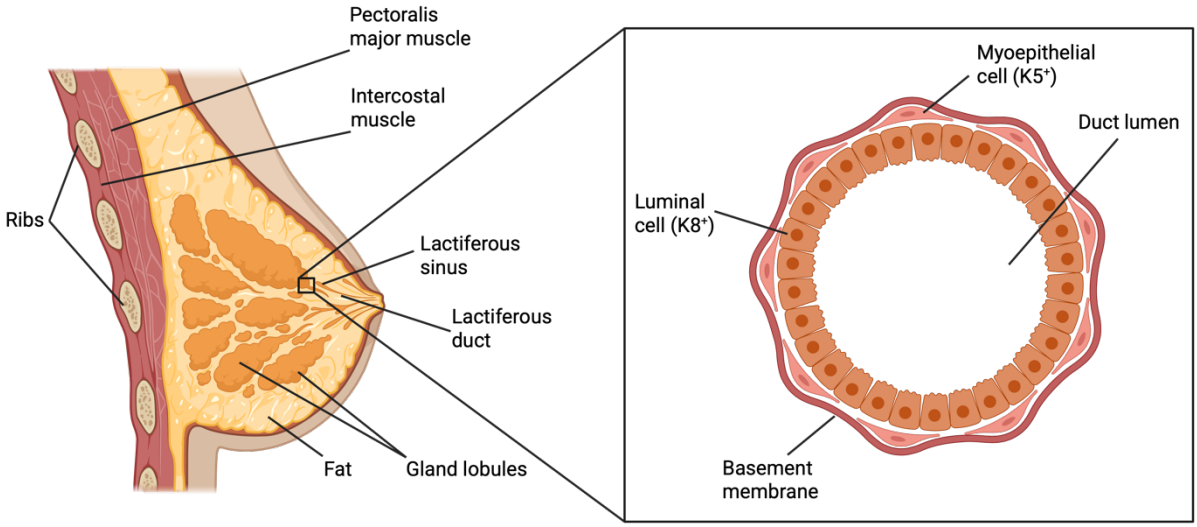


Figure 1.4 | Mammary gland anatomy.

Left: schematic cross-section of human mammary gland tissue with secretory gland lobules embedded in fat tissue. Right: schematic magnification of a secretory duct with keratin 8 (K8)-expressing luminal cells and keratin 5 (K5)-expressing myoepithelial cells. Figure created with Biorender.com.

1.1.2.2 Mammary gland development

Mammary gland development is a highly dynamic process that occurs in distinct stages throughout life: embryonic development, puberty, pregnancy, and adulthood. The initial formation of the mammary anlage during embryogenesis gives rise to a rudimentary ductal structure. However, most postnatal mammary gland development occurs during puberty, characterized by extensive ductal elongation and branching morphogenesis (Macias & Hinck, 2012).

This process is mediated by an involved interplay of systemic hormones and local growth factors (Arendt & Kuperwasser, 2015). In puberty, estrogen, especially the biologically most potent version 17β-estradiol (E_2)[1], produced by the ovaries, plays a central role in ductal elongation, primarily acting on stromal cells to induce paracrine signaling that stimulates epithelial proliferation in the outgrowing terminal end buds (TEBs, Figure 1.5).

[1]Since E_2 was used in different experiments throughout this thesis to assess the effect of estrogen, the terms "estrogen" and "E_2" will be used interchangeably in the following sections.

Progesterone (P_4) promotes side-branching and alveolar bud formation, preparing the gland for potential future lactation. In addition to ovarian hormones prolactin (Prl), secreted by the anterior pituitary, and growth hormone, acting through insulin-like growth factor 1 (IGF-1), are critical regulators of ductal development and epithelial differentiation (Brisken et al., 1998; Neville et al., 2002). Furthermore, paracrine mediators such as amphiregulin, wingless Int-1 (Wnt) proteins, and receptor activator of nuclear factor κB (NFκB) ligand (RANKL) contribute to local signaling pathways that fine-tune epithelial morphogenesis (Brisken et al., 2000; Ciarloni et al., 2007; Hennighausen & Robinson, 2005).

1.1.2.3 Tissue remodeling: pregnancy, lactation, and involution

The mammary gland undergoes its most profound remodeling during the reproductive cycle. Pregnancy induces massive epithelial expansion and alveolar differentiation in preparation for lactation. This transformation is driven by hormonal cues, including elevated levels of estrogen, progesterone, prolactin, and placental lactogens (Hennighausen & Robinson, 2005; Neville et al., 2002). Estrogen and progesterone continue to stimulate ductal and alveolar growth, whereas prolactin plays a pivotal role in terminal differentiation of luminal cells into secretory alveolar cells capable of synthesizing milk components (Figure 1.5; Watson & Khaled, 2008). Oxytocin mediates the milk ejection reflex by stimulating myoepithelial cell contraction during lactation (Nickerson et al., 1954).

Following parturition, the lactational phase is maintained by continued prolactin signaling in response to suckling-induced stimuli, which suppress hypothalamic dopamine release and sustain prolactin secretion. Milk synthesis and secretion require not only hormonal stimulation but also a finely tuned metabolic program within luminal cells to produce proteins, lipids, and carbohydrates (Anderson et al., 2007).

Upon weaning, the mammary gland undergoes involution, a tightly regulated process involving extensive epithelial cell apoptosis, extracellular matrix remodeling, and clearance of milk and cellular debris (Watson, 2006). Involution is marked by the withdrawal of lactogenic hormones, increased expression of pro-apoptotic signals such as signal transducer and activator of transcription 3 (STAT3) activation, and the re-establishment of the adipose-rich stromal environment (Sargeant et al., 2014). This process restores the gland to a near pre-pregnant state, preserving the capacity for future reproductive cycles.

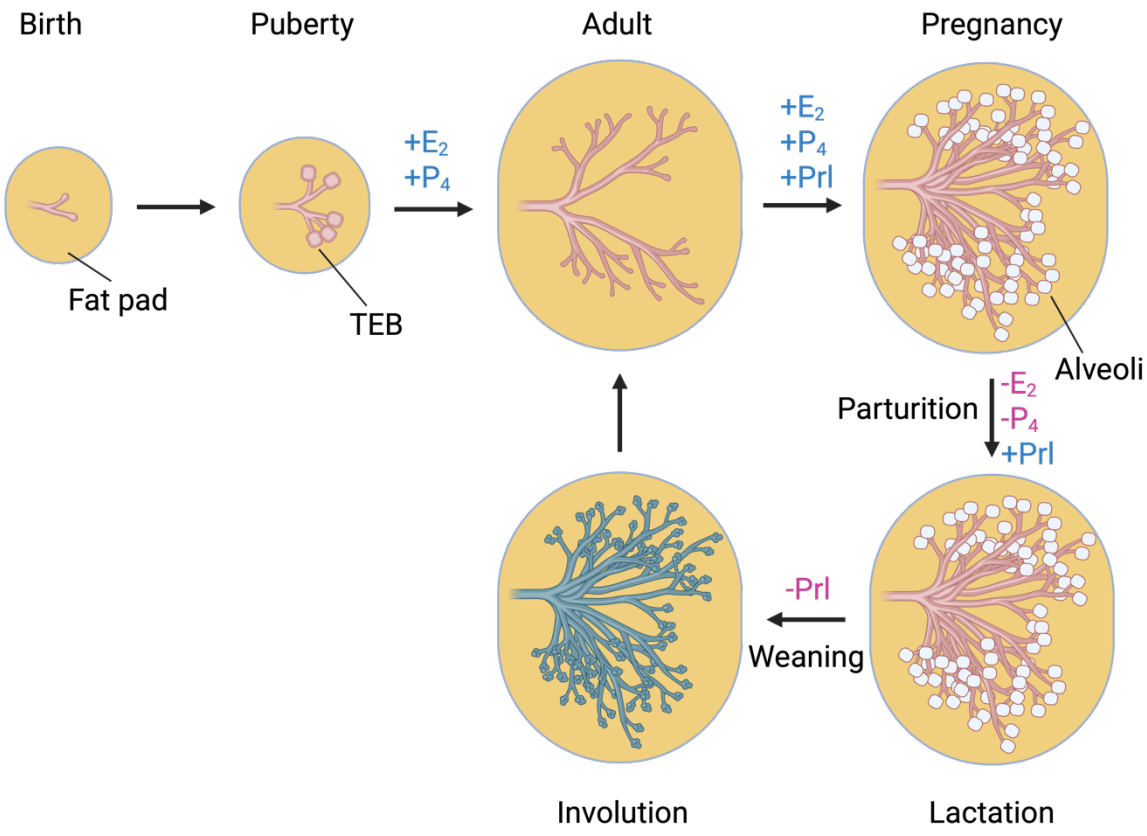


Figure 1.5 | Mammary gland development and remodeling.

Schematic representations of mammary gland growth, branching and tissue remodeling during post-natal development, adulthood and the reproductive cycle. Relevant hormones are highlighted: E_2 (estrogen), P_4 (progesterone), Prl (prolactin). Figure created in Biorender.com.

1.1.2.4 Breast Cancer

Breast cancer arises from the malignant transformation of epithelial cells within the mammary gland and represents the most common cancer in women worldwide (Siegel et al., 2022). The majority of breast cancers are thought to originate from luminal epithelial cells, although basal-like tumors may arise from myoepithelial or progenitor populations (Visvader & Stingl, 2014). Subtypes of breast cancer are classified based on molecular markers such as estrogen receptor (ER), progesterone receptor (PR), and human epidermal growth factor receptor 2 (HER2/neu) expression, with significant implications for prognosis and therapy (Perou et al., 2000). The triple-negative (ER-, PR-, HER2/neu-) breast cancer (TNBC) represents the subtype with highest metastasis and recurrence rate and is accompanied with the poorest prognosis, mainly due to the lack of therapeutic options (Dent et al., 2007).

Importantly, the periods of pregnancy, lactation, and involution profoundly influence breast cancer risk and biology. Pregnancy-associated breast cancer (PABC) refers to breast cancer cases with a diagnosis during pregnancy or up to one year after giving birth (Galati et al., 2023).

Although full-term pregnancy is associated with a long-term reduction in breast cancer risk, there is a transient increase in risk immediately following childbirth (Liu et al., 2002; P. Schedin, 2006). Postpartum breast cancer (PPBC), defined as breast cancer diagnosed up to ten years after delivery, is associated with particularly aggressive tumor biology and poorer clinical outcomes compared to breast cancer in nulliparous women (Borges et al., 2020). Studies indicate that an increase in breast cancer risk after childbirth can even persists for more than two decades for TNBC and ER-positive breast cancer (Liu et al., 2002; Nichols et al., 2018). The involution process itself is thought to create a pro-tumorigenic microenvironment characterized by immune suppression, extracellular matrix remodeling, and increased stromal protease activity, all of which may facilitate tumor initiation and progression (Lyons et al., 2011).

Thus, understanding the unique physiological and molecular changes that occur in the mammary gland during the reproductive cycle is crucial for elucidating the mechanisms underlying breast cancer development and for improving strategies for early detection and treatment.

1.2 Polyploidy

1.2.1 Definition and implications on cell physiology

As a key concept in genetics and cell biology, the term *ploidy* (Greek *-plóos*, "-fold") describes the number of homologous sets of chromosomes in a cell. In addition, ploidy is used to define the "genetic make-up" of each cell in a given organism and thus to classify the organism itself.

The ploidy of an organism as well as a cell's ploidy, are described by the Latin small case letter "n" (2n meaning diploid, 4n meaning tetraploid etc.). Of note, the cell's DNA content (tracked by the capital letter "C") might differ from its ploidy. For instance, a cycling diploid cell in G2 phase of the cell cycle, contains the same of amount of DNA as a tetraploid cell, despite still being diploid (2n/4C). Although these notions of ploidy and DNA content are not as standardized as for example SI-units, I will stick to them throughout this thesis as most of the relevant literature does.

Mammalian reproduction includes the fusion of a sperm and an oocyte, both containing a single or haploid chromosome set, generating a zygote containing two homologous chromosome sets. Therefore, mammals are considered diploid organisms[2].

[2]The Plains Viscacha rat (*Tympanoctomys barrerae*) was initially regarded as the first known tetraploid mammal due to its exceptionally large genome (M. Gallardo et al., 1999). However later studies questioned the tetraploid status and as of now the scientific community has not reach consensus (M. H. Gallardo et al., 2006; Svartman et al., 2005).

Polyploidy, which refers to more than two complete sets of chromosomes, occurs mainly in plants, including most flowering plants as well as many plants bred for agriculture. In the animal kingdom, polyploidy occurs less often but some examples include the autotetraploid Salmonidae fish or amphibians such as the allotetraploid *Xenopus laevis* as well as several insect species. Autotetraploidy refers the spontaneous duplication of the whole genome while allotetraploidy originates from the fusion of two genomes from two different species generating hybrid organisms. Polyploidization usually occurs during times of increased stress such as environmental changes and mass extinction events (Van de Peer et al., 2017). Polyploidy facilitates a certain stress resistance and increases the number of genes available to adapt to such stress and can thus be considered as a genetic playground.

Being polyploid has several implications on cell physiology. First, polyploid cell are often larger than diploids due to the increased DNA content (Darmasaputra, van Rijnberk, et al., 2024). For one thing, the DNA increases the spatial demand per se. For another thing, polyploidy can lead to altered gene expression. Multiple gene copies can lead to higher levels of transcription, causing greater output of cellular building blocks (Choudhury et al., 2024). However, gene expression might also be tightly regulated to prevent such effects, leading to dosage compensation (Darmasaputra, van Rijnberk, et al., 2024). Additionally, polyploidy can cause changes in cell cycle dynamics leading to altered timing and duration of the different cell cycle phases (Comai, 2005).

1.2.2 Polyploidy in human and murine tissues

While *Homo sapiens* and *Mus musculus* are diploid organisms (2n = 46 and 2n = 40, respectively), several tissues in these species harbor polyploid cells with varying functions. Here, it must be mentioned, that polyploidy on the cellular level can occur in different variations. First, nuclear polyploidy describes a cell state, where the cell nucleus itself contains more than 2 sets of chromosomes, so the whole cell's DNA is packed within one compartment (Donne et al., 2020). Such phenotype is normally generated by endoreplication, which refers to completing the S phase of the cell cycle but skipping mitosis and cytokinesis. Mitotic slippage, the premature exit of mitosis before chromosome segregation upon stress, can cause nuclear polyploidy in an erroneous manner as well. Second, cellular polyploidy describes the phenomenon of having more than one nucleus per cell (Donne et al., 2020). Such multinucleated cells can be generated via endomitosis (replication and mitosis but cytokinesis is skipped) or cell fusion. Additionally, failed cytokinesis can generate multinucleated cells due to defective chromatin segregation. Of course, mixed forms with more than two polypoid nuclei per cell can exist as well.

1.2.2.1 Polyploid cardiomyocytes and megakaryocytes

Cardiomyocytes actively proliferate during fetal development. However, after birth hyperplastic growth of the heart switches to hypertrophic growth including multinucleation and polyploidization of most cardiomyocytes in human and mouse (Y. Guo & Pu, 2020). Polyploid cardiomyocytes increase in cell size leading to an enhanced contractile force generation to pump blood efficiently to every organ. On top, polyploidization in the heart works as an "decision point" to limit proliferation. It is thought that this limited proliferative capacity helps to maintain a highly organized structure in the heart needed for proper electrical and mechanical function. While limited proliferation prevents cancer development from the heart, it also limits heart regeneration – one of the main causes for accumulative damage in heart tissue with time and increasing risks of coronary heart diseases with age (Kirillova et al., 2021).

Another cell type that is polyploid but shows limited proliferative capacities is the megakaryocyte (Mazzi et al., 2018). Diploid megakaryocyte progenitor cells divide normally in the bone marrow. However, once a progenitor commits to become a megakaryocyte, it switches to repeating cycles of endomitosis, leading to giant cells with a DNA content of up to 64n. Such high ploidy levels support the function of megakaryocytes in constantly producing blood platelets, but they will not produce daughter cells anymore (Mazzi et al., 2018).

1.2.2.2 Polyploidy in exocrine glands

A third kind of cell that is often polyploid rather describes a whole group of different cell types – secretory exocrine cells. Such cells include hepatocytes, luminal alveolar cells of the mammary gland and pancreatic acinar cells, the latter two will be used as model tissues presented in this thesis (Donne et al., 2020; Rios et al., 2016; Wollny et al., 2016).

Nevertheless, murine hepatocytes, the chief functional cells of the liver, are by far the best studied cells regarding the function and consequences of polyploidy in glandular tissues. Thus, a broad overview of the current knowledge about hepatocyte polyploidy will help grasping parallels to polypoid cells from the pancreas and the mammary glands. Hepatocytes are mostly diploid in the developing liver up to the point of weaning (Donne et al., 2020). Changing diet from a high-fat, non-toxic and constant diet (mother's milk) to more complex solid food, forces hepatocytes to adapt with more functional capacity. Polyploidization increases metabolic activity, blood protein and bile output and helps in detoxifying environmental chemicals. Thus over 90 % of murine hepatocytes and approximately 30-40 % of human hepatocyte become polyploid (Donne et al., 2020). While human polyploid hepatocytes are mostly binucleated and tetraploid, murine hepatocytes can regularly become 8n or 16n. Interestingly, cell fusion has

been discussed as a potential rare mechanism for liver polyploidization, although other studies found that endoreplication and endomitosis represent the principal mechanisms for polyploid hepatocyte generation (Darmasaputra, Geerlings, et al., 2024; Gentric & Desdouets, 2014; Wang et al., 2003). In contrast to the previous examples of polyploid cell types, hepatocytes can still proliferate under certain circumstances (Duncan et al., 2010; Matsumoto et al., 2020). In the naïve liver, homeostasis is maintained by slow hepatocyte turnover. Notably, the pool of diploid hepatocytes can sustain liver turnover without invoking any stem-cell–like precursors. Although diploid and polyploid cells can both proliferate under such normal physiological conditions, diploid hepatocytes are the principal drivers of homeostatic renewal. In humans, diploid hepatocytes show a seven-fold higher annual generation rate (Heinke et al., 2022). Polyploid hepatocytes, by contrast, divide rather infrequently, making only a minor contribution to routine cell replacement. A whole different situation emerges in the injured liver. The liver is well known for its extraordinary regenerative capacities. For example, after a 70 % partial hepatectomy (PHx) in rodents, remnant hepatocytes re-enter cell cycle and proliferate to restore the liver mass within about 7-14 days in mice (Miyaoka et al., 2012). In this scenario, both diploid and polyploid hepatocytes contribute to regrowth via self-renewal of the differentiated cells, again without any need for progenitor activation. Diploid hepatocytes often respond robustly as their smaller size and genomic simplicity can translate to faster cell cycle completion. That said, polyploid hepatocytes are far from inert during regeneration. Recent lineage-tracing in mice revealed that the bulk of polyploid hepatocytes can clonally expand after injury, meaning most regenerating hepatocyte clones originate from pre-existing hepatocytes of any ploidy rather than from a diploid pool of cells exclusively (Matsumoto et al., 2020). Of note, although dedicated stem or progenitor cells are not required for efficient regeneration after most liver injuries, severe or chronic damage of hepatocyte such as alcoholic/non-alcoholic steatohepatitis (NASH), or toxin-induced liver failure that exhausts mature hepatocytes, can activate an additional pool of otherwise dormant bipotent progenitor cells (Lowes et al., 2003). These oval cells express markers of immature liver cells such as α-fetoprotein (AFP) but additionally express hepatocyte (albumin) and cholangiocyte (CK19) markers.

As mentioned above, other exocrine glands where polyploid cells can be found include the pancreas and the mammary gland. Binucleated pancreatic acinar cell have been known for 100 years (Dolley, 1925). Yet, they remain a highly understudied phenomenon, often unnoticed by the scientific community. In the adult mouse pancreas, acinar cells appear to be binucleated at amounts of approximately 40 % (Figure 1.6. A), while around 15 % of human acinar cells are

binucleated (Figure 1.6 B, Wollny et al., 2016). Similar to hepatocytes in the liver, polyploid acinar cell emergence coincides with nutrition shift imposed during weaning, most probably to adapt to changed demands on food digestion and boost enzyme production (Anzi et al., 2018). Both, hepatocyte and pancreatic acinar cell binucleation upon weaning have been shown to depend on the atypical E2F transcription factors E2F7 and E2F8 which repress genes important for proper cytokinesis, leading to omitted cytokinesis[3] (Matondo et al., 2018; Pandit et al., 2012).

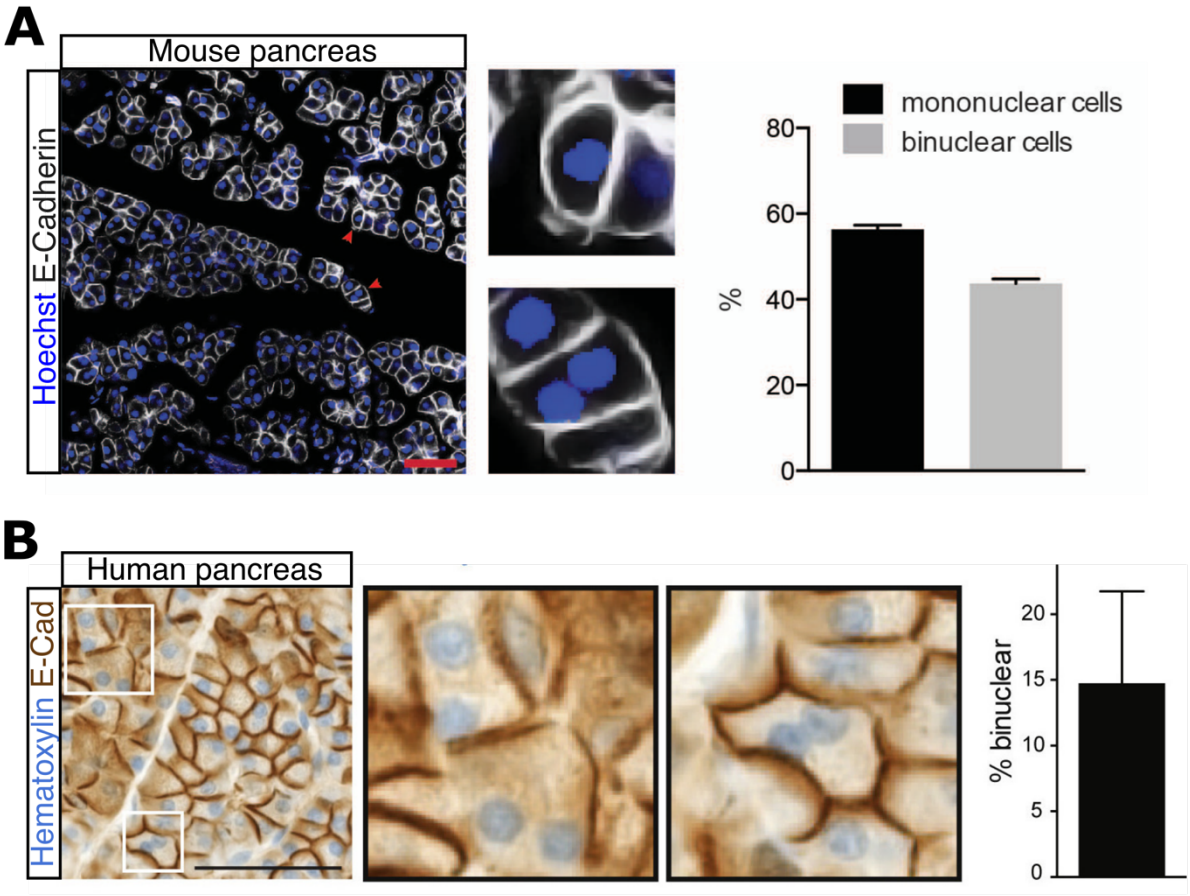


Figure 1.6 | Binucleated acinar cells in the exocrine pancreas.

(A) Left: immunofluorescence images of mouse pancreatic tissue stained for E-Cadherin (grey) and nuclei using DAPI (blue). Scale bar 50 µm. Right: quantification of the abundance of binucleated acinar cells in the mouse pancreas. **(B)** immunohistochemical images of human pancreatic tissue stained for E-Cadherin (brown) and nuclei using hematoxylin (blue). Scale bar 50 µm. Right: quantification of the abundance of binucleated acinar cells in the human pancreas. Adapted from Wollny et al., 2016.

[3]Note: the term "omitted cytokinesis" is not commonly used in literature. Instead "failed cytokinesis" is more regularly used. However, "failed cytokinesis" also refers to an erroneous abortion of cytokinesis. Thus, I use "omitted cytokinesis" to emphasize a scheduled skip of cytokinesis to generate polyploid cells in a targeted manner.

Compared to the pancreas and the liver, the mammary gland is an exocrine tissue with much stronger temporal restrictions regarding it secretory function. The transient nature of milk secretion can also be observed by the presence of polyploid luminal alveolar cells during lactation (Figure 1.7). In the mouse mammary glands, polyploid luminal cells emerge at late pregnancy between 16.5 and 18.5 days of pregnancy (dP) and reach levels of up to 50 % by two days of lactation (dL) at which approximately 50 % of the polyploid cells are binucleated (Molinuevo et al., 2024; Rios et al., 2016). Apart from mice, binucleated luminal cells were detected in various mammalian species including human with an abundance of approximately 30 % (Rios et al., 2016). The resulting increase in DNA content and cell size in polyploid luminal cells is believed to support milk production for adequate offspring nourishment (Fu et al., 2015; Rios et al., 2016). Similar to acinar cells and hepatocytes, binucleated luminal cells emerge through omitted cytokinesis. This is triggered by the upregulation of Aurora kinase A (AURKA), a critical regulator of proper spindle formation during mitosis (Rios et al., 2016). In contrast, mononucleated polyploid luminal cells are generated via physiological DNA damage upon mitotic stress during pregnancy, leading to cyclin-dependent kinase 1 (CDK1) inhibition and mitotic arrest, effectively causing endoreplication (Molinuevo et al., 2024). After weaning, rapid tissue remodeling induced by involution leads to the disappearance of binucleated luminal cells within 48 hours, potentially through apoptosis (Rios et al., 2016).

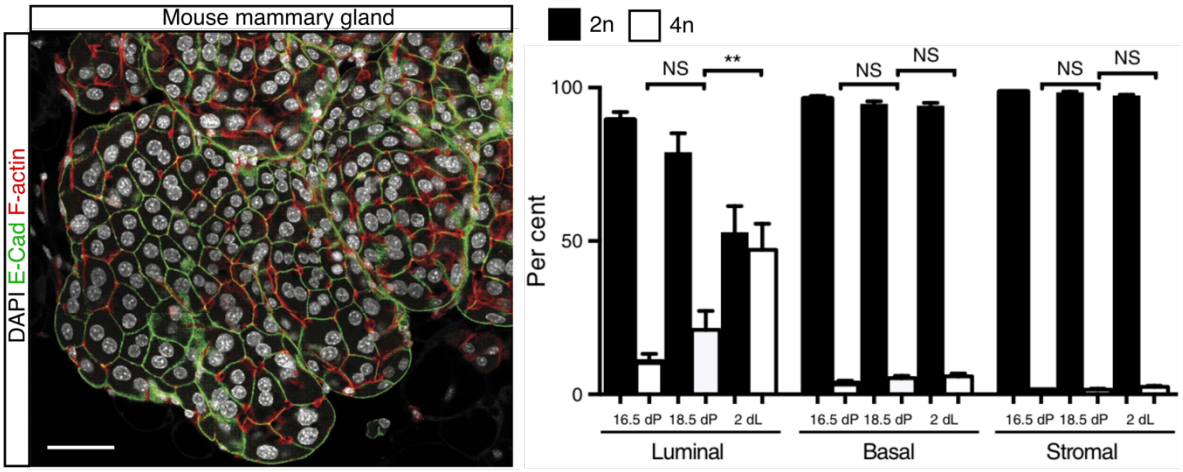


Figure 1.7 | Polyploid luminal cells in the mouse mammary gland.

Left: immunofluorescence image of mouse mammary gland tissue at 4 days of lactation stained for E-Cadherin (E-Cad, green), F-actin (red) and nuclei using DAPI (white). Scale bar 40 µm. Right: flow cytometry-based quantification of ploidy levels in different cell compartments of the mouse mammary gland at different stages of pregnancy (dP) and lactation (dL). Adapted from Rios et al., 2016.

1.2.3 Whole genome duplication and cancer

Apart from physiological polyploidy, as seen in various human tissues, polyploidy is linked to cancer in a complex interplay of stabilizing and destabilizing forces. On the one hand polyploidy provides multiple copies of tumor suppressor genes, such as tumor protein 53 (T*P53*) or phosphatase and tensin homolog (*PTEN)* and thus provides a certain degree of resilience against inactivating mutations. On the other hand, this resilience facilitates an increased resistance to apoptosis. Under which circumstances such aspects would facilitate pro- or anti-tumorigenic effects is matter of active research. Polyploid giant cancer cells (PGCC) for example, are believed to play critical roles in tumor relapse after chemotherapeutic treatments. Their increased DNA content combined with a low turnover rate provide selective adaption to survive such treatments. After the treatment period, these cells can repopulate the tumor via amitotic cell divisions (neosis) generating highly variable daughter cells (Sundaram et al., 2004). On top, the treatment using DNA damaging reagents itself can cause the emergence polyploid cancer cells. Furthermore, polyploidy is inherently linked to chromosomal instability and aneuploidy, both hallmarks of cancer and premalignant lesions, as described thoroughly in the next section (Storchova & Pellman, 2004). Another study showed that unscheduled polyploidy is linked to improper scaling of cell mass and replication factors, leading to replication stress and DNA damage in proliferative polypoid cells (Gemble et al., 2022). Together, it can be said that polyploidy in cancer acts as a reservoir of DNA damage by amplifying its accumulation and alleviating its consequences at the same time, as nicely presented in a recent study (Hayashi et al., 2024).

1.3 Chromosomal instability

Chromosomal instability (CIN) refers to an increased rate of chromosomal mis-segregation and structural alterations, leading to dynamic changes in the karyotype over time. CIN is a hallmark of many cancers and plays a critical role in tumorigenesis by promoting genetic diversity, drug resistance, and aggressive tumor evolution. The following sections provide an overview of the cellular mechanisms underlying CIN, its relationship with polyploidy, and its pathological consequences.

1.3.1 Chromosome Segregation Errors and Aneuploidy

Faithful and symmetric mitotic segregation of chromosomes into two daughter cells is pivotal for sustained genomic integrity (Figure 1.8 A). However, errors during this process are common in

cancer cells and are a major source of aneuploidy, defined as an abnormal number of chromosomes.

One of the key mechanisms underlying chromosome mis-segregation is the formation of merotelic attachments, where a single kinetochore becomes simultaneously attached to microtubules emanating from both spindle poles (Figure 1.8 B, Cimini et al., 2001). Unlike syntelic or monotelic errors, merotelic attachments are often not detected by the spindle assembly checkpoint (SAC) and can persist into anaphase (Cimini et al., 2002). As a consequence, the affected chromosome lags during segregation, forming a lagging chromosome visible in the anaphase spindle midzone (Figure 1.8 B).

Lagging chromosomes are a principal cause of aneuploidy, as they can be mis-segregated into the wrong daughter cell or fail to be incorporated into the main nucleus at all (Cimini et al., 2002).

In addition to lagging chromosomes, mis-segregation events can result in the formation of chromatin bridges, where chromosome arms are stretched between daughter cells due to unresolved sister chromatid entanglements, dicentric chromosomes, or DNA repair defects (Maciejowski et al., 2015)

Chromatin bridges often undergo mechanical rupture during cytokinesis, leading to further chromosomal fragmentation and rearrangements. These breakage events can initiate breakage-fusion-bridge (BFB) cycles, a mutagenic process described by Barbara McClintock in maize, where broken chromosomes fuse, mis-segregate, and break again in successive cell divisions, driving genome instability (McClintock, 1941).

Thus, chromosome segregation errors during mitosis are central to the generation of aneuploidy and complex chromosomal rearrangements observed in cancer cells.

1.3.2 The Link Between Polyploidy and Chromosomal Instability

Polyploidy is closely associated with increased chromosomal instability. One key mechanism linking polyploidy to CIN is the deregulation of centrosome numbers. Normally, cells have two centrosomes that create a bipolar spindle during mitosis. However, polyploid cells frequently harbor supernumerary centrosomes (Ganem et al., 2009).

Extra centrosomes predispose cells to form multipolar spindles, leading to highly abnormal chromosome segregation patterns (Figure 1.8 C). While many polyploid cells cluster their centrosomes into two poles to achieve a pseudo-bipolar mitosis, this clustering process is still error-prone due to transient multipolarity before clustering (Kwon et al., 2008). Even when centrosomes are successfully clustered, the presence of extra microtubule-organizing centers increases the likelihood of merotelic kinetochore attachments, resulting in elevated rates of

lagging chromosomes and subsequent aneuploidy (Ganem et al., 2009). Together, polyploidy does not only reflect a distinguished genomic state but actively promotes ongoing CIN through mechanical and regulatory vulnerabilities during cell division.

1.3.3 Micronuclei and Chromothripsis

Lagging chromosomes that fail to be incorporated into the main daughter nuclei can form micronuclei, small extranuclear bodies surrounded by their own nuclear envelope (Figure 1.8 B, C; Fenech et al., 2011). Similarly, chromatin bridges that break during cytokinesis can result in acentric chromosomal fragments that form micronuclei.

Micronuclei are a hallmark of genome instability and have clinical relevance as biomarkers for cancer diagnosis and prognosis. Elevated frequencies of micronuclei are observed across a wide range of tumor types and correlate with poor clinical outcomes (Fenech et al., 2011; Terradas et al., 2010).

Beyond serving as indicators of instability, micronuclei play a direct mechanistic role in genome rearrangements through a phenomenon known as chromothripsis. Chromothripsis is characterized by massive, localized genomic fragmentation and chaotic reassembly, resulting in complex structural rearrangements (Stephens et al., 2011). As such, it is often confined to a single chromosome.

Micronuclei are particularly prone to chromothripsis due to several vulnerabilities. The micronuclear envelope is often structurally defective and prone to spontaneous rupture during interphase (Hatch et al., 2013). Upon rupture, micronuclear chromatin is exposed to the cytoplasm, leading to extensive DNA damage. Furthermore, DNA replication within micronuclei is asynchronous and incomplete, causing replication stress and further increasing the risk of chromosomal fragmentation (C.-Z. Zhang et al., 2015).

Chromothripsis contributes to tumorigenesis by enabling the rapid accumulation of oncogenic rearrangements and copy number alterations in a single catastrophic event, accelerating tumor evolution and adaptation (C.-Z. Zhang et al., 2013). Genomic studies have revealed that chromothripsis is prevalent in a wide range of human cancers, including PDAC and breast cancer with a prevalence of over 60 % (Cortés-Ciriano et al., 2020; Y. Li et al., 2020; Voronina et al., 2020).

Thus, micronuclei mark chromosomal instability and actively fuel cancer progression by driving complex genomic alterations that.

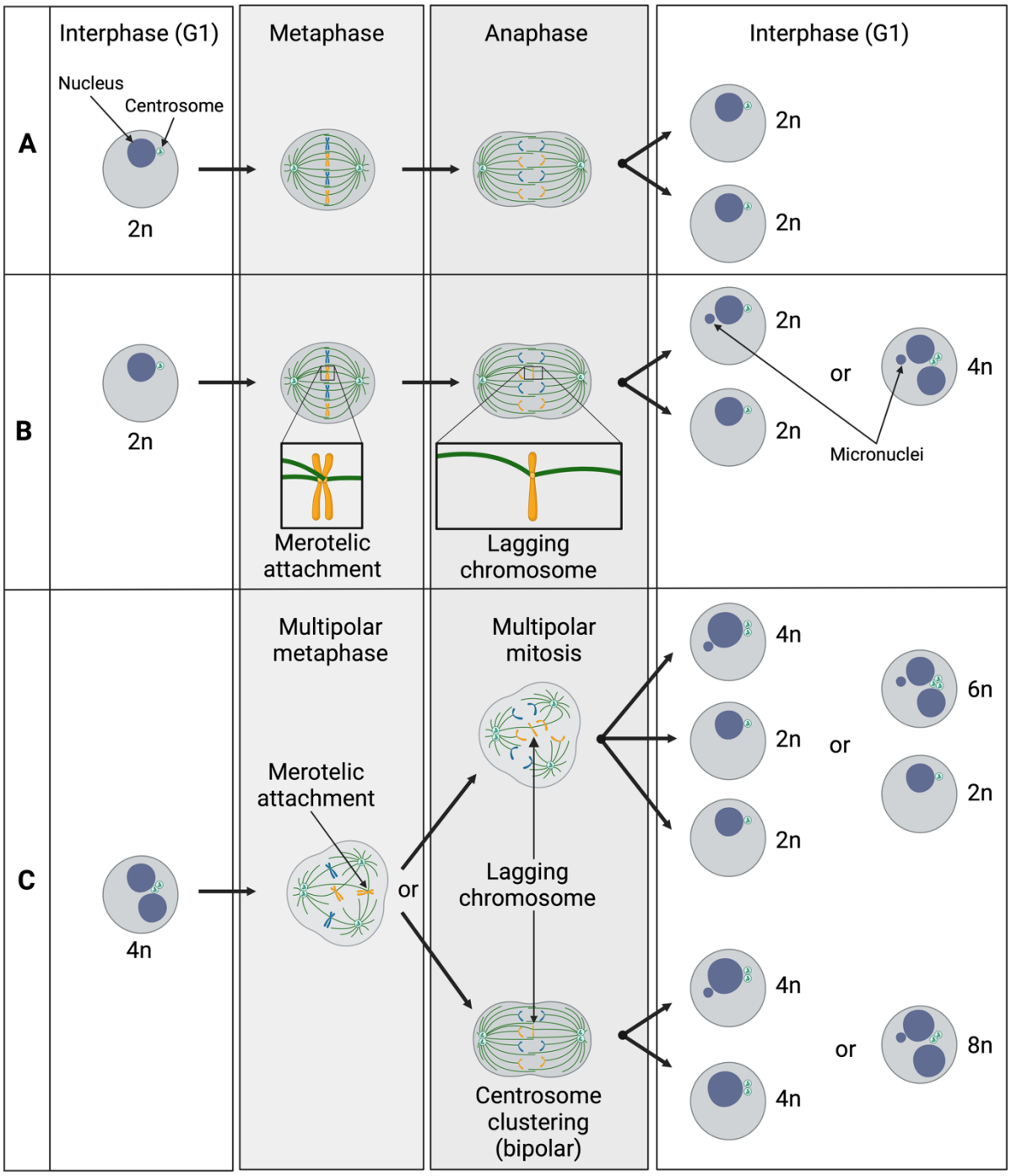


Figure 1.8 | Schematic overview of chromatin segregation errors and their link to polyploidy.

(A) Normal mitosis of a diploid cell generating two diploid daughter cells. **(B)** Diploid cell, that encounters a merotelic attachment and a lagging chromosome highlighted in magnification boxes. Generated daughter cells can exhibit micronuclei. **(C)** A tetraploid binucleated cell in G1 phase undergoing a multipolar spindle arrangement in metaphase due to its increased centrosome numbers (2 in G1 phase), which are highly prone to merotelic attachments, lagging chromosomes and micronuclei. Figure generated in Biorender.com.

1.4 Quantitative image analysis in biomedical research

Advances in imaging technologies, such as confocal, light-sheet and super-resolution microscopy, have enabled researchers to visualize biological structures with high spatial resolution. However, extracting meaningful quantitative information from these complex datasets requires sophisticated computational methods. In this thesis, a customized image analysis pipeline was developed to segment individual cells from confocal images, classify them based on morphological characteristics, and quantify fluorescence intensities at the single-cell level. In the following sections, I introduce key concepts of computer vision and image analysis that form the foundation of this approach.

1.4.1 Digital Representation of Images

In digital systems, images are represented as arrays of numerical values. In grayscale images, each pixel is assigned an intensity value that typically ranges from 0 (black) to 65536 (white) for 16-bit encoding (which was used in this thesis). More formally, an image can be mathematically described as a two-dimensional (2D) function:

$$I: (x, y) \rightarrow v$$

where (x, y) represent the spatial coordinates and v the corresponding pixel intensity value. Modern biomedical imaging systems often allow to additionally acquire data along the optical depth axis z, and eventually along the temporal dimension t creating a three-dimensional (3D) image stack or a four-dimensional (4D) hyperstack, respectively. In fluorescence microscopy, it is furthermore common to capture multiple optical channels corresponding to different fluorophores or biomolecular stains (Figure 1.9). Each optical channel highlights distinct biological structures, such as nuclei, membranes, or specific proteins, thus resulting in multi-channel images. Together, digital fluorescence microscopy images are encoded as multi-dimensional numerical arrays formally represented as:

$$I: (x, y, z, c, t) \rightarrow v$$

where c indexes the imaging channels.

By encoding images as numerical data, computers can perform mathematical operations on them, enabling tasks such as filtering, transformation, segmentation, and classification (Gonzalez & Woods, 2017).

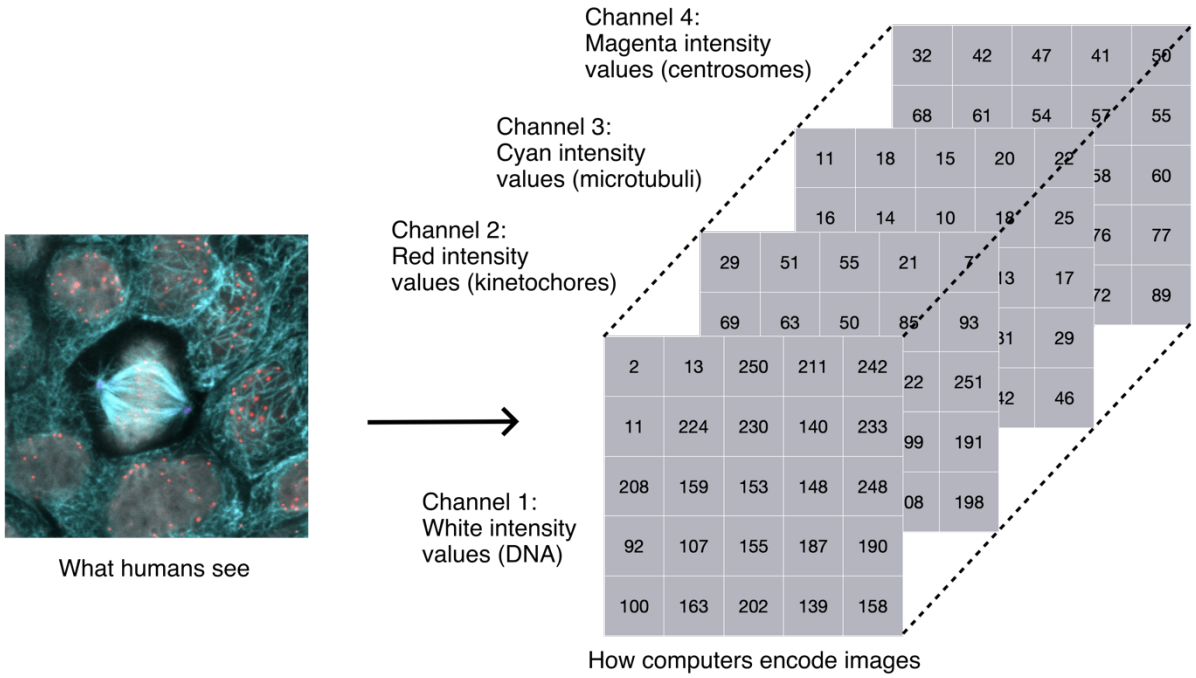


Figure 1.9 | Images are stored as multidimensional arrays of pixel values.

Schematic illustration of how an actual microscopy image, here showing DNA (white), kinetochores (red), microtubules (cyan) and centrosomes (magenta), gets stored as a multidimensional array of values that correspond to the individual pixel intensities.

1.4.2 Classical Feature Extraction and Classification

Image classification aims to assign a label to an image based on its content or meaning. Early computer vision approaches achieved this by manually designing algorithms to extract measurable properties, known as features, from the image data. Features captured aspects such as edge orientation, texture patterns, or shape contours, serving as compact representations of the visual information.

Common examples of classical feature extraction include the use of Sobel filters for edge detection, Histograms of Oriented Gradients (HOG) for shape characterization and Local Binary Patterns (LBP) for capturing texture information (Dalal & Triggs, 2005; Sobel & Feldman, 1973; Yang & Chen, 2013). After extracting features, classification algorithms such as support vector machines (SVMs) or decision trees were used to map feature vectors to class labels (Cortes & Vapnik, 1995).

Although effective in controlled settings, the success of these methods was limited by the quality and relevance of the chosen features, which often required domain expertise and laborious trial-and-error optimization.

1.4.3 Deep Learning and Multilayer Perceptrons

Although many of the core algorithms underlying modern artificial intelligence (AI) technologies, such as transformers and diffusion models, were developed decades ago, it was the combination of abundant data availability and dramatic advances in computational power that fueled the deep learning revolution, enabling computers to excel in tasks that were previously thought to be exclusive to human cognition. The introduction of deep learning offered an alternative approach by enabling models to learn features directly from raw image data. The foundational architecture for deep learning in supervised tasks is the multilayer perceptron (MLP), an artificial neural network composed of an input layer, one or multiple hidden layers, and an output layer (Rosenblatt, 1958).

Each neuron in the network performs a weighted summation of its inputs, applies a non-linear activation function, such as the rectified linear unit (ReLU) or the sigmoid function, and passes the result forward to the next layer (Nair & Hinton, 2010). During the feedforward process, the input is propagated through successive layers, transforming it into increasingly abstract representations.

Training an MLP involves minimizing a loss function, which quantifies the discrepancy between predicted and true class labels. The optimization of the network weights is achieved via gradient descent algorithms, where gradients are calculated efficiently through backpropagation (Rumelhart et al., 1986). During backpropagation, the error is reversely propagated from the output layer to the input layer to update the weights in a manner that reduces the overall loss.

Key hyperparameters, such as the learning rate, which controls the magnitude of weight updates, and the number of epochs, which determines how many complete passes are made over the training data, have a critical influence on the final model performance.

1.4.4 Convolutional Neural Networks

Although MLPs are powerful, they treat all inputs as independent, ignoring the spatial relationships inherent in image data. Convolutional neural networks (CNNs) address this limitation by incorporating local connectivity and parameter sharing through convolutional operations (Figure 1.10 A; Lecun et al., 1998).

In a convolutional layer, a set of small learnable filters is applied across the input image to produce feature maps that preserve spatial hierarchies. These filters, also referred to as convolutional kernels, locally extract patterns such as gradients, textures, edges or corners and are optimized during training. Pooling layers, typically using max pooling operations, downsample feature maps by summarizing local neighborhoods, reducing computational

complexity and imparting translational invariance (Figure 1.10 B). The stride parameter controls the step size of the convolution or pooling operation, directly influencing the size of the resulting feature maps. Dropout layers introduce random deactivation of neurons during training, mitigating overfitting by encouraging redundancy in feature representations (Srivastava et al., 2014). At the final stage, fully connected layers integrate the spatial features into global feature vectors for classification.

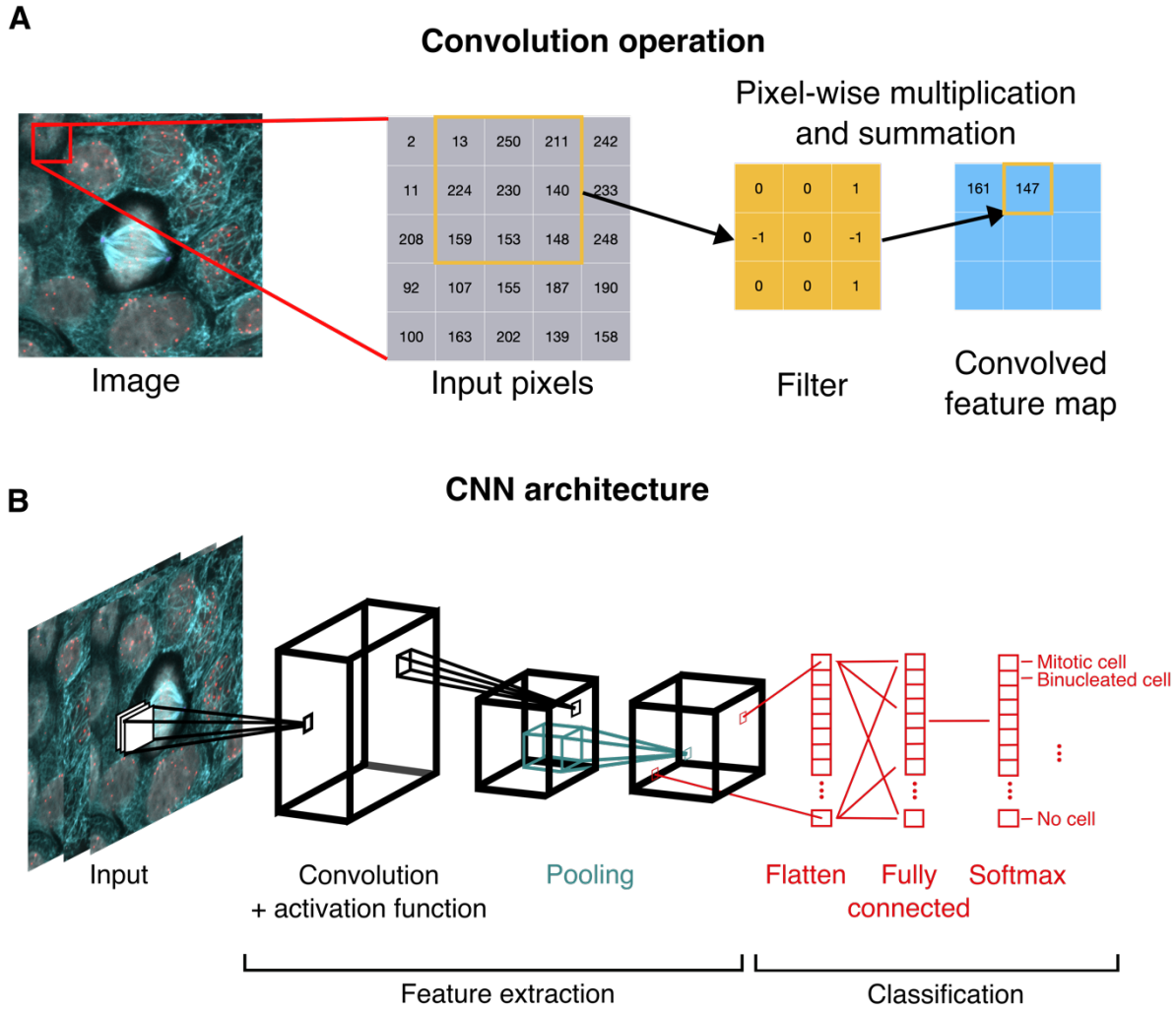


Figure 1.10 | Schematic representation of image classification using CNNs.

(A) Visualization of convolutional operations on images to extract generalizing features. A convolutional filter (here 3x3) slides over the input pixel values with a distinct stride and calculates an element-wise multiplication and summation to convolve image features. **(B)** High-level architecture of a CNN. The feature extractor component with convolutional filters and pooling layers creates highly convolved feature maps. A classification component flattens the final feature map into a 1D vector and classifies the image using fully connected layers and a normalization function (here Softmax) to convert the output into a probability function corresponding to the number of possible classes.

Several landmark CNN architectures, some of which were used as part of this thesis have significantly advanced the field. AlexNet demonstrated that deep networks trained with graphical processor unit (GPU) acceleration could dramatically outperform traditional approaches in large-scale image classification tasks, notably winning the 2012 ImageNet competition by a significant margin (Krizhevsky et al., 2012). Its use of ReLU activations, dropout regularization, and overlapping max pooling introduced key design principles that remain standard today. VGGNet proposed a simple yet powerful architecture composed of sequential 3×33×3 convolutions and demonstrated that deeper networks could achieve better performance when trained appropriately (Simonyan & Zisserman, 2015).

ResNet addressed the challenge of training extremely deep networks by introducing residual connections, also referred to as skip connections, which create shortcuts allowing gradients to flow more easily during backpropagation, thus preventing vanishing gradient problems, which were not only an issue in computer vision but especially problematic in natural language processing tasks (He et al., 2015). The skip connection has since then been used in many architectures including modern generative models such as transformers, and diffusion models for image generation. Interestingly, a recent study mapped the *Drosophila melanogaster* larva brain connectome and found "multilayer shortcut" connections, recapitulating the information flow from skip connections (Winding et al., 2025; Xu et al., 2024). EfficientNet further optimized network design by proposing a compound scaling method that systematically balances network depth, width, and input resolution, achieving state-of-the-art performance with fewer parameters (Tan & Le, 2020).

1.4.5 Image Segmentation

Image segmentation aims at subdividing an image into distinctly labeled regions, each corresponding to a particular class or object. In semantic segmentation, every pixel is classified into a category, but individual objects belonging to the same category are not distinguished (Figure 1.11). In contrast, instance segmentation differentiates individual objects, producing unique labels for each instance. Panoptic segmentation combines both approaches, providing a unified framework that assigns each pixel both a semantic label and an instance label when applicable, ensuring comprehensive scene understanding.

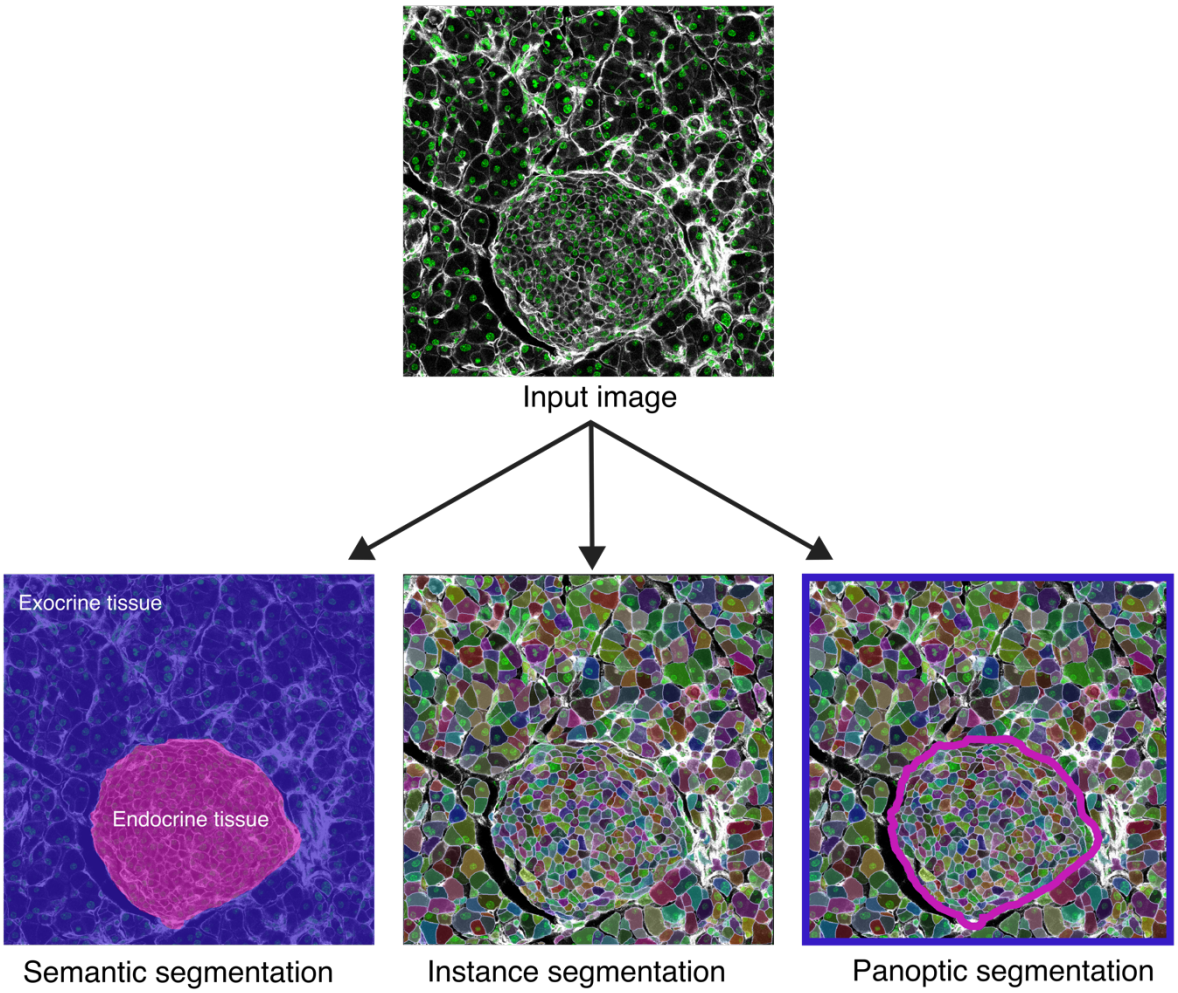


Figure 1.11 | Overview of different segmentation tasks in computer vision.

A confocal immunofluorescence image of a mouse pancreatic tissue section was stained for E-Cadherin (white) and nuclei using DAPI (green). Semantic segmentation: each pixel gets a label assigned corresponding to its semantic class. Instance segmentation: objects in images get detected, located and assigned to an own instance label. Panoptic segmentation: combines semantic segmentation and instance segmentation to label individual objects of multiple classes.

1.4.5.1 Classical segmentation approaches

Classical methods for semantic segmentation include thresholding techniques, where pixels are separated based on intensity values. Otsu's method is a well-known algorithm that automatically determines the optimal threshold by minimizing intra-class variance (Otsu, 1979). Another popular approach is watershed segmentation, which treats the image as a topographic surface and segments regions by simulating the flooding of basins from local minima (Beucher & Lantuéjoul, 1979). Although effective for simple images, these techniques often struggle with overlapping objects, heterogeneous intensities, and noise, which are common challenges in biomedical images.

1.4.5.2 Deep Learning for Semantic Segmentation

Deep learning approaches have dramatically improved semantic segmentation by reframing the task as a pixel-wise classification problem. The U-Net architecture is among the most influential models in this domain, particularly for biomedical applications (Ronneberger et al., 2015).

U-Net consists of an encoder-decoder structure, where the encoder progressively reduces spatial resolution while capturing semantic context through a series of convolutions and pooling operations. The decoder restores spatial resolution through upsampling and convolutional operations. Critically, skip connections are used to concatenate corresponding feature maps from the encoder to the decoder, enabling the network to retain fine-grained spatial information essential for precise segmentation.

Due to its efficiency, robustness, and ability to perform well even with limited training data, U-Net has become the gold standard for tasks such as cell segmentation and tissue structure delineation.

1.4.5.3 Instance Segmentation

Instance segmentation methods combine object detection, which locates objects within an image, and semantic segmentation, which assigns class labels to pixels. This dual challenge is particularly relevant for biological images containing dense and overlapping structures.

In bioimage analysis, specialized deep learning models have been developed for this purpose. StarDist represents cells as star-convex polygons, predicting radial distances from a central point to the object boundary at fixed angles (Schmidt et al., 2018). This strategy allows for efficient and accurate delineation of tightly packed and irregularly shaped cells.

Cellpose adopts a fundamentally different approach to cell segmentation by modeling spatial vector flows rather than directly predicting object masks (Stringer et al., 2021). Specifically, Cellpose transforms a manually annotated input training image into its vector flow representation with spatial gradients pointing to the center of each cell (Figure 1.12 A). This flow-based representation allows the model to handle highly variable cell morphologies, including elongated, rounded, or irregular shapes, without requiring a fixed object template (Figure 1.12 B). A neural network is then trained to predict the flow field vectors for each pixel for the horizontal and vertical axes, creating horizontal and vertical gradients and effectively encoding the direction and relative distance to the nearest cell edge (Figure 1.12 C). In addition, the neural network predicts for each pixel an associated probability that it belongs to a cell or not based on the annotation mask, similar to standard semantic segmentation.

The underlying architecture of Cellpose is based on a modified U-Net (Figure 1.12 D). During inference, predicted flow fields are integrated using a numerical integration process (similar to gradient descent) to reconstruct coherent individual cell masks from the pixel-wise flow information. Effectively, every pixel tracks its gradient to its fixed center and all pixels whose gradients track into the same center are assigned with the same cell label (Figure 1.12 E, F).
This approach enables Cellpose to generalize remarkably well across a wide range of biological datasets, including fluorescent images, histological stains, and even non-biomedical imagery, without requiring extensive retraining. Furthermore, the model can adapt to novel cell types or imaging modalities by fine-tuning on a relatively small number of annotated examples, making it particularly well-suited for heterogeneous datasets. Throughout this thesis, Cellpose was employed as the primary instance segmentation tool, given its robustness and versatility for automated and high-throughput single-cell analysis from image data.

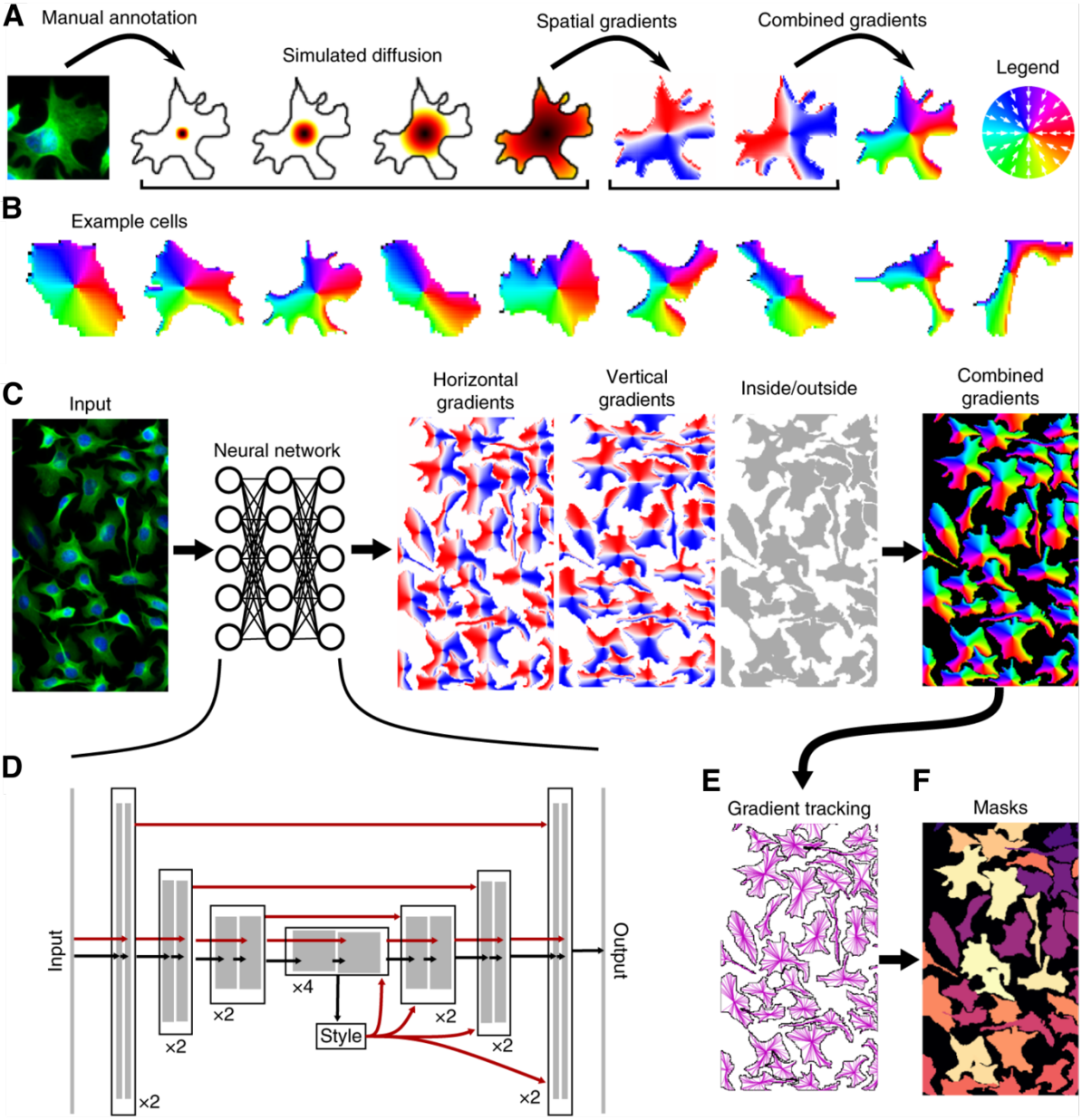


Figure 1.12 | Cellpose algorithm for cell instance segmentation.

(A) Annotation labels from manual annotation of training data are converted into their vector flow representations of spatial gradients via simulated diffusion. **(B)** Examples of vector flow representations for different cell morphologies, Cellpose can segment. **(C)** A neural network based on a modified U-Net architecture **(D)** is trained to predict the horizontal and vertical gradients as well as a binary segmentation mask of cells and background. These three layers of information are integrated to generate a combined gradient map. **(E)** For prediction, each pixel's gradient is tracked to its fixed center point to assign converging gradients to the same cell, thus producing instance segmentation masks **(F)**. Adapted from Stringer et al., 2021.

1.5 Aims of the thesis

Although the role of polyploid cells in exocrine tissues regarding their increased secretory capacities is well established, studies addressing their behavior under non-homeostatic conditions are largely missing.

Considering the inherent potential of polyploid cells to serve as source for chromosomal instability and DNA damage, polyploidy is a highly understudied feature of many adult human tissue and cell types. Its implications in regeneration have been mainly addressed in cardiomyocytes, where polyploidy acts as a roadblock to proliferation and regeneration. However, in exocrine glands, where polyploid secretory cells occur across several organs, the situation is less clear. Considering the current literature, addressing polyploid hepatocytes, pancreatic acinar cells and mammary gland luminal cells, creates a blurred image regarding their proliferative capacities and their potential to acquire chromosomal instability under non-homeostatic conditions. Thus, the overall goal of this PhD project is to determine whether polyploid cells in exocrine tissues, that undergo tissue remodeling beyond homeostasis, activate proliferative capacities and serve as a source of chromosomal instability, potentially leading to tumor initiation. Specifically, my PhD thesis project aims to:

1. **Characterize the role of binucleated pancreatic acinar cells in regeneration.**
   Assess changes in cellular ploidy levels during the regenerative process. Investigate whether polyploid acinar cells undergo ADM and proliferation in response to injury. Determine whether polyploid acinar cells under de- or re-polyploidization.
2. **Evaluate the contribution of polyploidy to chromosomal instability in pancreatic regeneration.**
   Assess mitotic errors, micronuclei formation, and chromothripsis in regenerating acinar cells. Determine whether polyploid acinar cells are more susceptible to CIN compared to their diploid counterparts. Identify the fate of micronucleated cells and their ability to continue proliferating.
3. **Extend findings to polyploid luminal cells of the mammary gland**
   Investigate whether lactation-induced polyploid luminal cells exhibit similar CIN-related features. Determine the impact of hormone-induced polyploidization on genomic stability. Assess the potential link between polyploid luminal cells and postpartum breast cancer.

2 Results

The following sections of this thesis will present the experimental results and method development outcomes obtained during my PhD work. Some of the data presented in the first results section 2.1 were initially generated during my master's thesis project while working in the same laboratory. All results that were exclusively obtained during my master's thesis have already been described in the introduction. Data originating from my master's work that are mentioned in the results sections are clearly indicated as such in the corresponding figure captions and were further extended during my PhD by including additional time points or data points, thus warranting their inclusion in the results part of this thesis.

2.1 Binucleated acinar cells act as facultative progenitors

2.1.1 Ploidy dynamics under non-homeostatic conditions

Previous experiments performed in the Martin-Villalba lab demonstrated that more than 40 % of pancreatic acinar cells in mice and 15% in humans are binucleated (Wollny et al., 2016). Furthermore, it was concluded that binucleated acinar cells would represent a terminally differentiated cell state, not contributing to proliferation during homeostasis or injury. This finding was based on the lack of proliferating binucleated acinar cells in the naïve or injured mouse pancreas. Additionally, it was observed that single binucleated acinar cells or cell clusters of higher number exclusively consisting of binucleated cells, would not form mouse acinar-derived organoids (mADOs) *in vitro*. Thus, Wollny and colleagues inferred that binucleated acinar cells neither belong to the small group of cells with long-term proliferative capacities nor to the sub-population of facultative progenitors replenishing the pancreas during regeneration (Wollny et al., 2016). However, the effect of pancreatic injury on the number of nuclei per cell and ploidy of the acinar population has not yet been addressed. Thus, I first asked if a pancreatic injury changes the mononucleated-to-binucleated cell ratio and whether this shift is reversed upon regeneration. To examine this, I employed cerulein-induced pancreatitis (CiP) in mice. Cerulein mimics the hormone cholecystokinin, triggering pancreatic acinar cells to release digestive enzymes (Kim, 2008). When administered in supramaximal doses, cerulein overwhelms the secretion capacity, causing early activation of zymogens. This premature activation leads to pancreatic autodigestion and initiates an inflammatory response. I fixed pancreas tissue at distinct days after the last cerulein injection (days post injection, dpi) and assessed the kinetics over the course of an acute inflammatory response (2 dpi, 4 dpi) to the fully regenerated organ (28 dpi to 91 dpi, Figure 2.1 A). By using immunofluorescence (IF) staining for

E-Cadherin to distinguish cell boundaries and thereby being able to track nuclear numbers, I quantified the abundance of binucleated acinar cells. I found that their number significantly drops from approximately 40 % (naïve) to 32 % at 2 dpi and 24 % at 4 dpi (Figure 2.1 B, C). The proportion of binucleated cells rose back to 31 % at 28 dpi and was almost back to original value of 40 % at 91 dpi. Thus, I could show that acute injury of the pancreas leads to fewer binucleated acinar cells and that this transient decrease is restored as regeneration is complete. This result is in line with findings from binucleated hepatocytes, which represent a growth end point when generated in regeneration and thus have been proposed to play a role in organ size control (Wilkinson et al., 2019).

Next, I aimed to examine whether binuclear acinar cells upon injury are generated through cell fusion or endomitosis. For this, I administered bromodeoxyuridine (BrdU) immediately after the last cerulein injection. At 91 days, no binucleated cells containing a mix of $BrdU^+$ and $BrdU^-$ nuclei were detected. Instead, all BrdU-labeled binucleated acinar cells exhibited BrdU incorporation in both nuclei, supporting endomitosis as the predominant mechanism of regeneration (Figure 2.1 D). Considering the short BrdU bioavailability of several hours after the last cerulein injection, this furthermore shows that binucleation is facilitated as an early response to injury, at a time when the overall amount of binucleated cell is still decreasing (Maltsev et al., 2022). Endomitosis has also been proposed as the main mechanism for hepatocyte and acinar cell binucleation during development indicating a potential execution of developmental programs to regenerate lost binucleated cells and to signal growth stop and the resolution of the regenerative process (Brodsky, 1985; Darmasaputra et al., 2024).

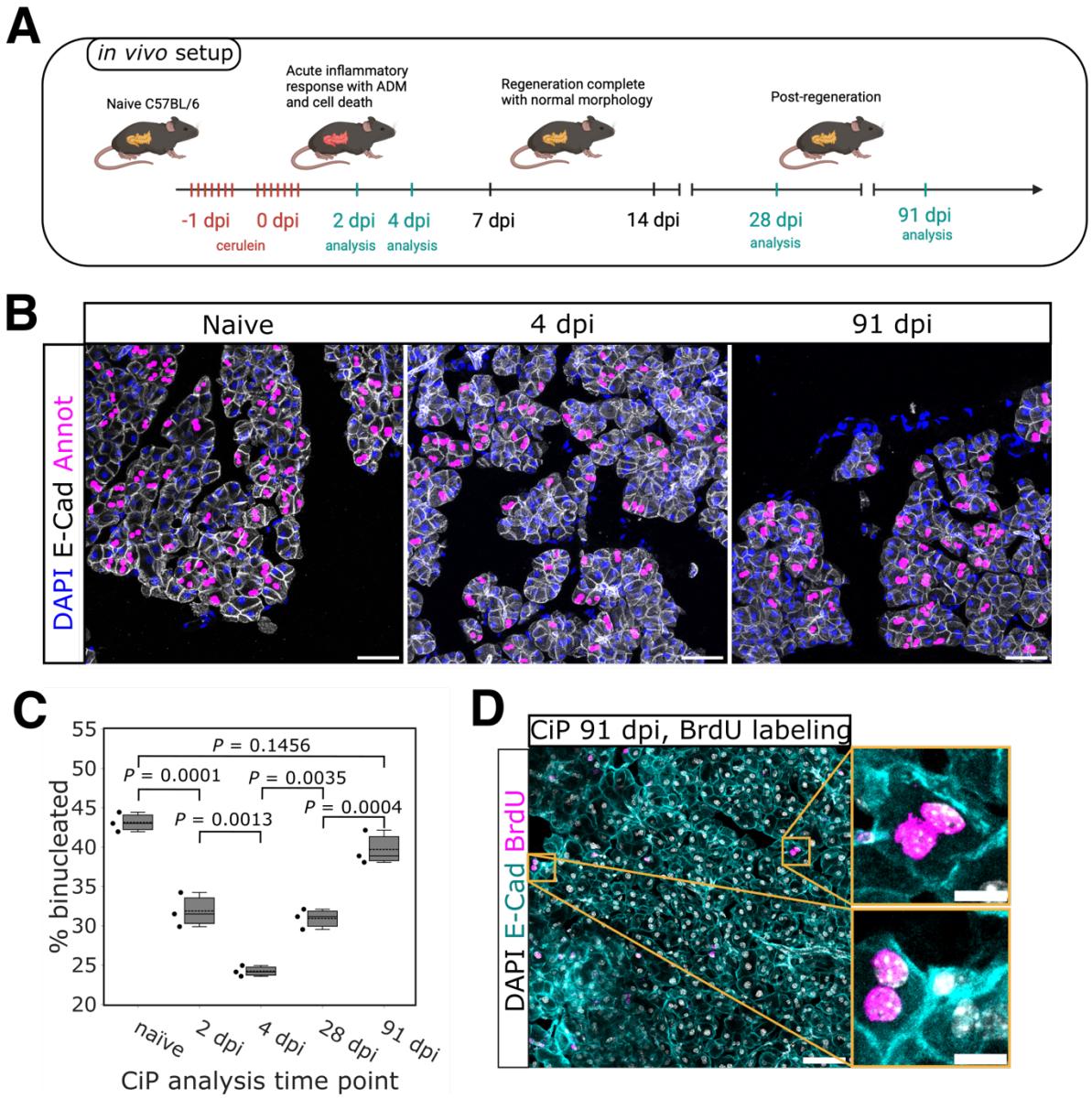


Figure 2.1 | Nuclear number dynamics in the regenerating mouse pancreas.

(A) Experimental setup to assess pancreas regeneration and post-regeneration after acute inflammation using cerulein-induced pancreatitis (CiP). **(B)** IF staining of healthy (naïve), regenerating (4 dpi) and post-regenerative (91 dpi) mouse pancreatic tissue. Injury was mediated by cerulein-induced pancreatitis. DAPI (blue), E-Cadherin (white) and annotations indicating binucleated cells (magenta). Scale bar: 50 µm. **(C)** Quantification of the fraction of binucleated acinar cells in naïve and injured (2 dpi, 4 dpi, 28 dpi and 91 dpi) mouse pancreata. Data from naïve, 2 dpi and 4 dpi were already acquired during my master's thesis (Brunken, 2019) and complemented with 28 dpi and 91 dpi and collectively reanalyzed for this. *P*-values were calculated by one-way ANOVA followed by Tukey's post hoc test (n = 3 mice). **(D)** IF image of post-regenerative mouse pancreatic tissue after cerulein-induced pancreatitis (91 dpi) and subsequent BrdU treatment. BrdU pulse was administered at 0 dpi. DAPI (white), E-Cadherin (cyan), BrdU (magenta). Yellow squares indicated zoomed-regions highlighting binucleated cells with BrdU-positive nuclei. Scale bar 50 µm.

Next, I examined the decline in binucleated cells during the early stages of regeneration. Given that both mononucleated and binucleated acinar cells show comparable susceptibility to apoptosis after pancreatic injury and the acute phase of necrosis and cell death is over, the observed decrease in binucleated cells from 2 dpi to 4 dpi cannot be explained by cell death alone (Lugea et al., 2006; Quan et al., 2018; Wollny et al., 2016). An alternative explanation is that binucleated cells divide to form mononucleated cells. To investigate this possibility, I employed the 3D *in vitro* culture of primary acinar cells supplemented with epidermal growth factor (EGF) and fibroblast growth factor 2 (FGF2), generating mouse acinar-derived organoids (mADOs), which allowed real-time tracking of binucleated acinar cell fates (Figure 2.2 A; Wollny et al., 2016). As described previously, this model recapitulates acinar-to-ductal metaplasia (ADM) and tubular structure formation, processes that acinar cells undergo *in vivo* to support proliferation and pancreatic regeneration (Marstrand-Daucé et al., 2023; Storz, 2017). I fixed mADOs at different time points in culture and stained for E-Cadherin to quantify cell nuclei numbers and ploidy. Additionally, I stained for α-Amylase to assess and validate the progress of ADM (Figure 2.2 B). As reported previously, the number of α-Amylase expressing cells decreased during mADO culture indicating the onset and progression of ADM (Figure 2.2 B; Wollny et al., 2016). In accordance with the data I obtained from *in vivo* pancreatitis, the fraction of binucleated cells in mADOs significantly decreased from ~40 % to approximately 13 % at 13 days (d13) of culture (Figure 2.2 C). I additionally assessed the overall ploidy of the mADO culture using an image analysis pipeline described in section 2.4. Interestingly, the overall number of polyploid cells did not change proportionally but dropped from 42 % at d0 to 30 % at d13.

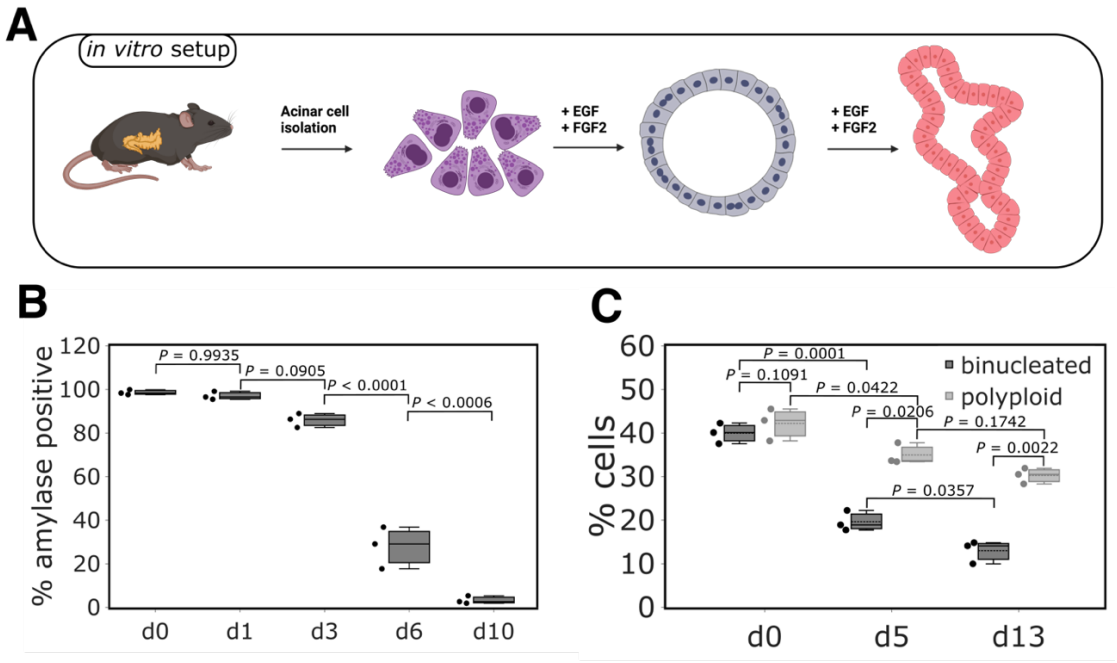


Figure 2.2 | ADM and ploidy dynamics in mADOs.

(A) Experimental setup to assess the role of polyploid pancreatic acinar cells in regeneration and early tumor formation using mouse acinar-derived organoids (mADOs). **(B)** Quantification of α-amylase-expressing cells in mADOs at different times in culture. *P*-values were calculated by one-way ANOVA followed by Tukey's post hoc test (n=3 mice). **(C)** Quantification of binucleated (dark grey) and all polyploid (light grey) cells in mADOs at different times in culture. *P*-values between binucleated and polyploid cells were calculated using a two-sided paired Student's *t* test for each timepoint (d0, d5, d13). Changes over time were compared using a Repeated Measures ANVOA followed by Tukey's post hoc test for binucleated and polyploid cells respectively (n=3 mice).

2.1.2 Binucleated acinar cells divide in mADOs

Next, I followed the fate of binucleated cells in mADOs in real time by live confocal and light-sheet imaging over several hours. To distinguish between mono- and binucleated cells, I took advantage of H2B-mCherry/mG double reporter mice which fluorescently label the cell nuclei as well as plasma membranes in all cells (Strnad et al., 2016). Intriguingly, I observed that binucleated acinar cells do divide (Figure 2.3 A). During mitosis, both nuclei condensate into a single metaphase plate and give rise to two mononucleated cells. As cell divisions performed by polyploid cells might possess altered timings compared to diploid cells, I determined the required time from nuclear envelope breakdown (NEB) to anaphase based on live-cell imaging recordings. While mitoses from binucleated cells exhibited an otherwise normal mitotic phase sequence, the time from NEB to anaphase and therefore the overall duration of cell division was significantly prolonged (Figure 2.3 B). The observed delay is likely due to an extended spindle assembly checkpoint during metaphase, which ensures that all kinetochores are properly attached to microtubule fibers (McAinsh & Kops, 2023).

To further examine whether the proliferative binucleated cells perform a full cell cycle, including an S phase, or rather represent a diploid cell state locked at a late cell cycle stage, I performed short-term BrdU labeling of mADO cultures starting directly after cell plating. I labelled cells in fresh wells every 2 hours and fixed after a pulse time of another 2 hours. As early as 24 hours after cell plating, I began to see BrdU-positive cells, including binucleated cells (Figure 2.3 C). In all cases, both nuclei were always BrdU-positive, as previously shown in cerulein-induced pancreatitis *in vivo*. These results indicate that binucleated cells represent a true polyploid cell state, capable of proliferating with synchronized onset of replication in both nuclei.

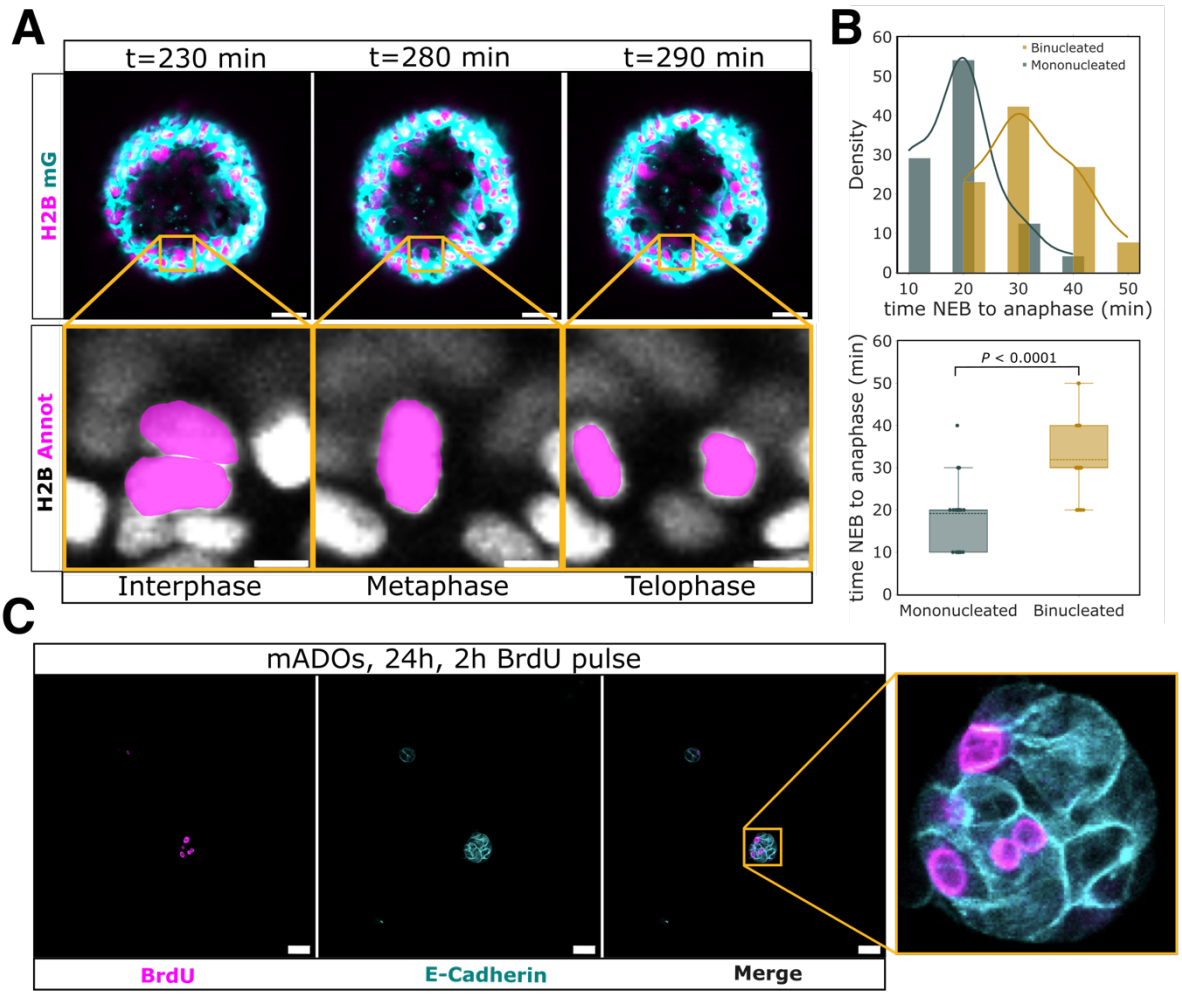


Figure 2.3 | Binucleated acinar cell proliferate in mADOs.

(A) Live-cell image series showing a binucleated acinar cell mitosis in H2B-mCherry/mG mice in interphase, metaphase and telophase. H2B (magenta), plasma membranes/mG (cyan). Scale bar: 20 µm. Yellow squares highlight zoomed regions with H2B (white) and binucleated cell annotation (magenta). Scale bar: 5 µm. **(B)** Duration measurements from nuclear envelope breakdown (NEB) to anaphase derived from mitotic cells from live-cell imaging data. Upper graph shows density distribution and kernel density plot of mononucleated (cyan) and binucleated (yellow) cell durations. Data is shown as histogram with bin size 4 and Kernel Density Estimation. Lower panel shows boxplots of mononucleated (cyan) and binucleated (yellow) cell durations. *P*-value was calculated using a two-sided unpaired Student's *t* test (n = 50 mitotic events from 3 mice). **(C)** IF images of short-term BrdU pulse-chase experiments in mADOs (24h in culture). BrdU-pulse given at 22 h, chase time 2h. BrdU (magenta), E-Cadherin (cyan). Yellow box indicates zoomed region showing a cell cluster with a binucleated BrdU-positive cell. Scale bars: 50 µm

2.1.3 Binucleated acinar cells undergo ADM and proliferate *in vivo*

In contrast to the data I acquired from mADOs, cell divisions of binucleated cells were not detected previously following CiP in mice. Wollny and colleagues had previously quantified prophase cells by phospho-histone H3 (pHH3) staining at 5 days after pancreatitis induction, which can be considered as an advanced stage of regeneration where proliferative binucleated

cells could have divided already (Wollny et al., 2016). Therefore, I quantified proliferating (pHH3-positive) and ADM (CK19-positive) acinar cells at 2 dpi. To distinguish mononucleated from binucleated pHH3-positive cells, I focused the quantification on late G2- and early M-phase nuclei, indicated by speckled pHH3 staining co-localizing with round and not fully condensed nuclear DAPI stain (LI et al., 2005). Notably, I did not only detect binucleated cells that stain positive for CK19 and pHH3, but at 2 dpi approximately 35 % of all CK19-expressing and 33 % of all pHH3-positive acinar cells were binucleated (Figure 2.4 A, B). These findings imply that binucleated acinar cells undergo ADM and proliferate at this early stage of regeneration – although with a slightly reduced propensity or slower dynamics compared to their mononucleated counterparts. As mentioned above, DCLK1 specifically marks the pool of facultative progenitor acinar cells, which possess the ability undergo ADM upon CiP (Bailey et al., 2014; Westphalen et al., 2016). To further validate that binucleated acinar cells exhibit regenerative capacities, I thus stained naïve pancreatic tissue for DCLK1. Indeed, I found that approximately 39 % of all DCLK1 expressing acinar cells are binucleated, supporting a significant contribution of binucleated acinar cells to the subset of facultative progenitor cells (Figure 2.4 C, D).

DCLK1-positive/CK19-positive binucleated cells could potentially get arrested in G1 in an *in vivo* setting by a "tetraploidy checkpoint" – a potential mechanism to prevent cells from cycling after polyploidization that has been proposed in the early 2000s (Andreassen et al., 2001; Margolis et al., 2003). To assess if binucleated acinar cells would become arrested during ADM, I stained mouse pancreatic tissue from cerulein-induced pancreatitis for γ-Tubulin. Polyploidy is often accompanied by an increased number of centrosomes (Bloomfield & Cimini, 2023; Duensing & Duensing, 2010). γ-Tubulin labels nucleation sites for microtubule formation at centrosomes and can thus be used as a proxy for cell ploidy. These extra centrosomes can lead to a multipolar spindle formation with up to four individual spindle poles in case of a tetraploid cell with four centrosomes (Baudoin et al., 2020; Duensing & Duensing, 2010; Faggioli et al., 2011; Maiato & Logarinho, 2014). I screened pancreas tissue after CiP for such metaphase figures that would correspond to polyploid cells. At 2 dpi, I found multipolar metaphases in CiP with up to 8 centrosomes, indicating that tetraploid and even octoploid cells can enter mitosis (Figure 2.4 E). Of note, I found no cell with two metaphase plates, indicating that so called "double mitoses", at which both nuclei undergo mitosis individually, do not play major role in pancreatic regeneration. This supports that binuclear cells divide via nuclear aggregation as observed in organoids.

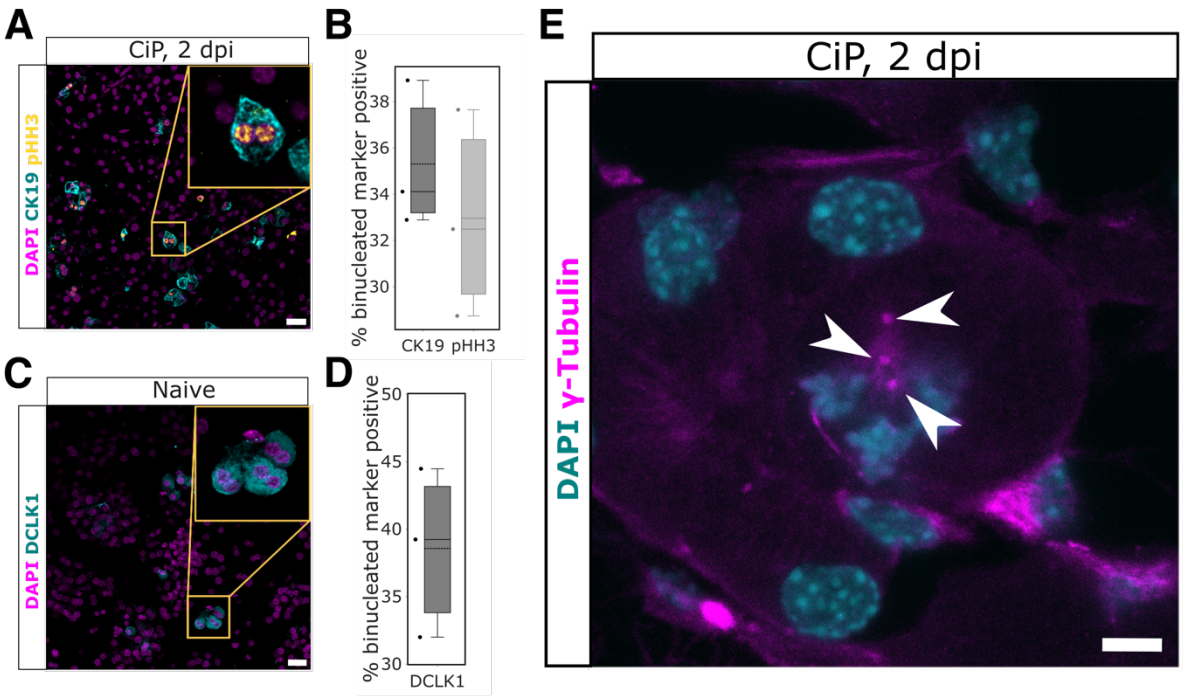


Figure 2.4 | Binucleated acinar cells contribute to regeneration upon inflammation.

(A) IF image of injured mouse pancreas (CiP, 2 dpi) to assess the contribution of binucleated acinar cells to ADM (cytokeratin 19-expressing, CK19) and proliferation (phospho-histone H3-positive, pHH3). DAPI (magenta), CK19 (cyan), pHH3 (yellow). Yellow box indicates zoomed region. Scale bar: 20 µm. Image was adapted from Brunken, 2019 **(B)** Quantification of the binucleated cell fraction in pHH3- and CK19-expressing acinar cells in cerulein-induced pancreatitis at 2 dpi. *P*-value was calculated using a two-sided unpaired Student's *t* test (n = 3 mice). This data was partially presented in Brunken, 2019 and complemented and re-analyzed with more data points per replicate for this thesis. **(C)** IF image of naïve mouse pancreatic tissue to assess the contribution of binucleated acinar cells to the pool of facultative progenitors (Doublecortin-like kinase 1-expressing, DCLK1). DAPI (magenta), DCLK1 (cyan). Scale bar: 20 µm. **(D)** Quantification of binucleated cell fraction in DCLK1-expressing naïve acinar cells. *P*-value was calculated using a two-sided unpaired Student's *t* test (n = 3 mice). **(E)** IF image of injured mouse pancreas (CiP, 2 dpi) stained for γ-Tubulin (magenta) and nuclei using DAPI (cyan), highlighting a multipolar metaphase. White arrow heads indicate centrosomes). The images shows a cell with 8 centrosomes in total visible in other focal planes, which were not projected for better visibility. Scale bar 5 µm.

2.2 ADM promotes multipolar spindle orientations in binucleated cells

2.2.1 Multipolar mitoses originate from binucleated acinar cells

Multipolar mitoses have been observed in hepatocytes as part of the so-called "ploidy conveyor," a process that describes the dynamic de- and re-polyploidization to regulate ploidy levels and facilitate adaptive advantages in regenerating liver tissue (Duncan et al., 2010). However, unlike the regenerating pancreas, hepatocyte proliferation during liver regeneration does not require metaplasia. Hepatocyte-to-cholangiocyte transdifferentiation is not a normal feature of regeneration but rather occurs in response to chronic injury, such as biliary obstruction

or liver fibrosis (Sekiya & Suzuki, 2014). In contrast, acinar-to-ductal metaplasia (ADM) is essential for acinar cell proliferation and pancreatic regeneration.

Thus, I examined whether the additional constrain imposed by metaplasia in the pancreas influences the tendency to undergo multipolar cell division and ploidy reduction in mononucleated or binucleated polyploid acinar cells. To address this question, I first screened H2B-mCherry mADO live-cell imaging data for fully completed multipolar mitoses. I additionally employed live-cell imaging of mADOs of the EGFP-Tuba mouse model, in which fluorescently labels α-Tubulin which enables the visualization of the individual microtubule spindles. I frequently observed mitoses completing with three poles, generating three daughter nuclei, two of which exhibited reduced DNA content (Figure 2.5 A). Notably, multipolar mitoses predominantly occurred in binucleated cells, while mononucleated polyploid cells rarely underwent this process. Instead, their centrosomes tended to cluster and form two poles (Figure 2.5 B).

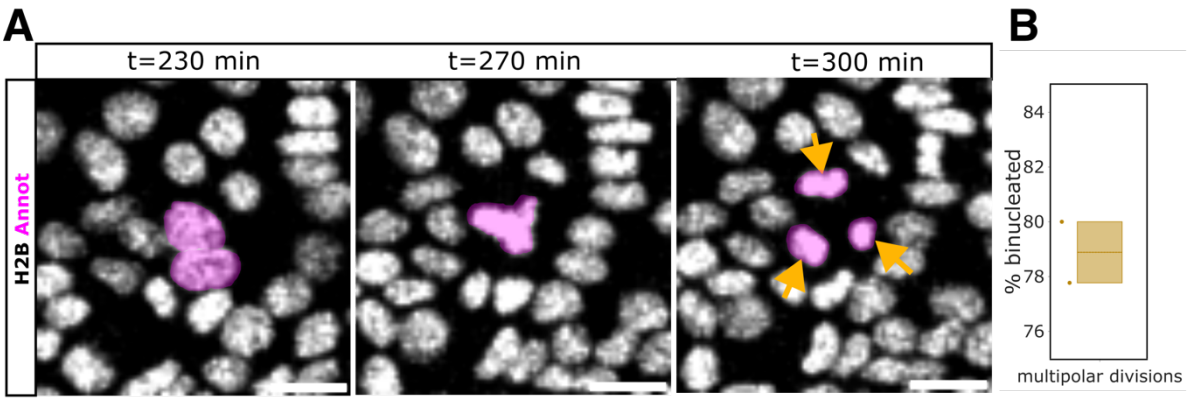


Figure 2.5 | Binucleated acinar cells are prone to multipolar mitoses.

(A) Live-cell image series from H2B-mCherry mADOs showing binucleated cell undergoing multipolar mitosis leading to a ploidy reduction of daughter nuclei. H2B (white), nuclei annotation (magenta).Yellow arrows indicate spindle locations. Scale bar 20 µm. **(B)** Quantification of the fraction of multipolar divisions that was performed by binucleated cells in mADOs (n = 2 mice).

2.2.2 Nuclear positioning impacts spindle geometry

A higher likelihood for binucleated cells to undergo multipolar mitosis may be attributed to the restricted intracellular space resulting from metaplasia, which is associated with a significant decrease in cell size, combined with the increased spatial demands of two nuclei. I quantified the sizes of naïve acinar cells and metaplastic ADM cells from mADOs. As expected, cells that underwent ADM exhibit significantly reduced sizes (Figure 2.6 A, B, C).

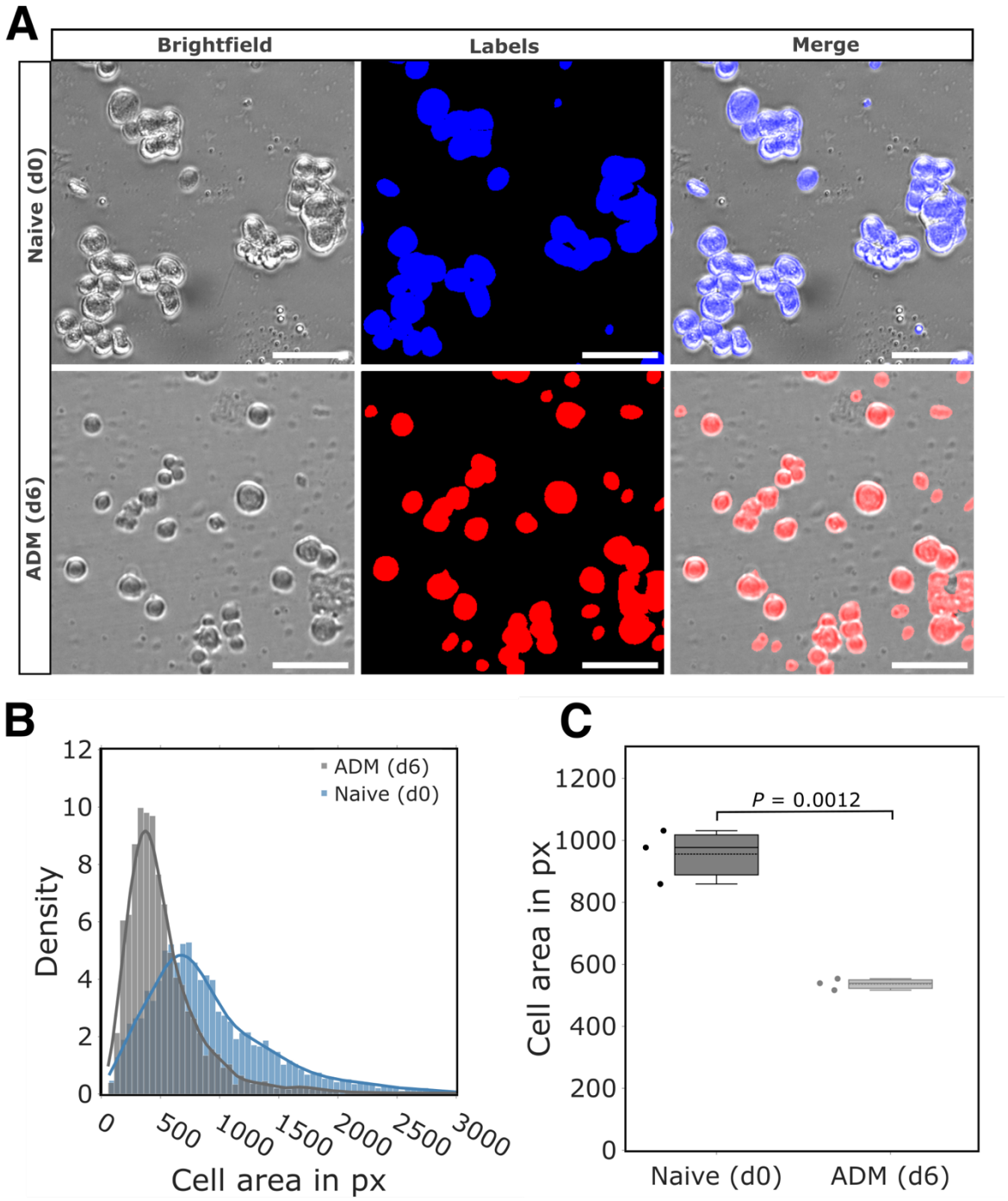


Figure 2.6 | mADOs undergo metaplasia-induced cell shrinking.

(A) Brightfield images of naïve acinar (top row) and ADM (d6 mADOs, bottom row) cells (left column) with corresponding segmentation masks in blue and red (center column) and overlay (right column) used for cell size measurements. Scale bar 50 µm. **(B)** Histogram of cell sizes from naïve (blue) and ADM (red) cells. Data is shown as histogram with bin size 50 and Kernel Density Estimation**. (C)** Cell area measurements of naïve acinar and ADM cells based on dissociated tissue and mADOs in pixels (px). *P*-value was calculated using a two-sided unpaired Student's t test (n = 3 mice).

In addition, it has been proposed that the DNA itself might act as a physical barrier to prevent centrosomes from clustering to form bipolar spindle orientations in polyploid cells (Klisch et al., 2016). An unfavorable positioning of the centrosomes could further facilitate multipolar spindle arrangements. A previous study showed that centrosomes in binucleated G0/G1 cow trophoblasts cluster near the interspace between the two nuclei (Klisch et al., 2016). Such a positioning in combination with limited intracellular space might impair centrosome migration to the cell cortices during prophase. The positioning effect might even be stronger if the nuclei orient along the long axis of the cell, as this is the axis a cell generally tends to divide along (Hertwig, 1884; Middelkoop et al., 2024; Théry et al., 2005). In addition, an orientation of the nuclei parallel to the axis of division, could prevent microtubule fibers from proper attachment to all kinetochores of both nuclei as the nuclei would shield each other in a bipolar spindle geometry. To resolve this issue and pass the spindle assembly checkpoint, a multipolar spindle orientation might be more favorable. Thus, I hypothesized that the cell shrinking induced by ADM, could force the nuclei in binucleated cells to orient along the long/division axis creating an unfavorable initial situation for bipolar mitoses. To test my hypothesis, I first stained naïve acinar cells for the centrosome marker pericentrin to assess centrosome positioning during G0/G1. In accordance with the data from cow trophoblasts, centrosomes in binucleated acinar cells cluster near the interface of the two nuclei (Figure 2.7).

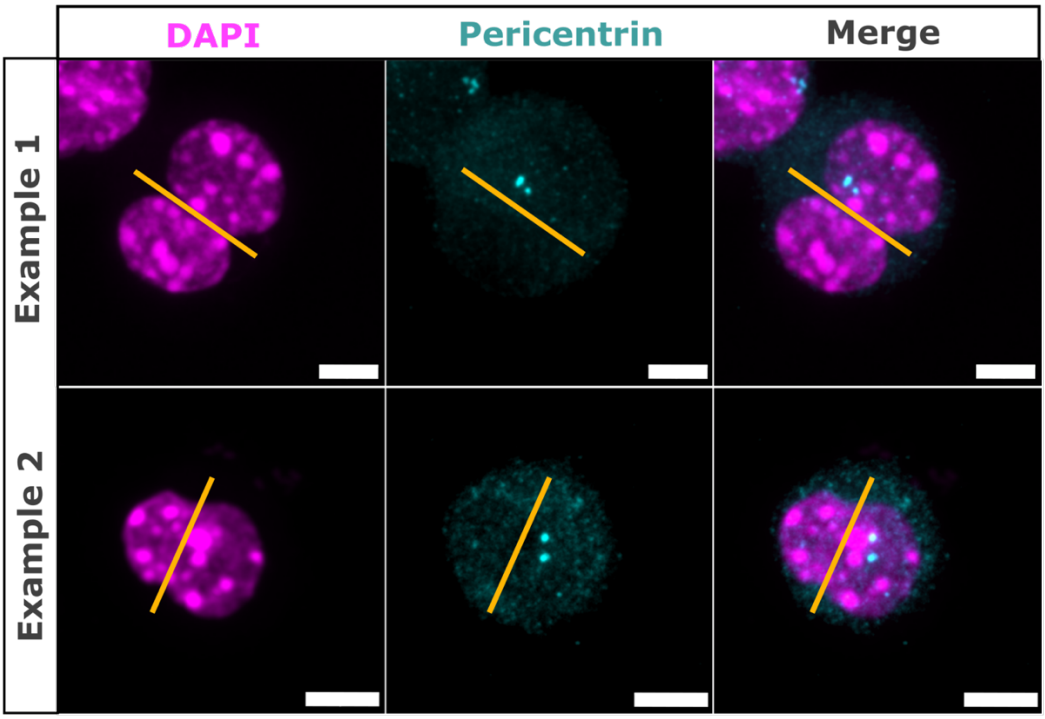


Figure 2.7 | Centrosomes in G1 binucleated acinar cells locate close to the nuclear interface.
IF images of two examples (rows) of the centrosome localization in naïve binucleated acinar cells. DAPI (magenta), pericentrin (cyan), annotation highlighting nuclear interface (yellow). Scale bar 5 µm.

Next, I assessed the nuclear orientation of binucleated cells just before mitosis using live-cell imaging of EGFP-Tuba mADOs. I distinguished orientations along the short and the long/division axis of the cell and classified the subsequent mitosis as bipolar or multipolar (Figure 2.8 A). Intriguingly, I found that multipolar spindle formations were always accompanied by nuclei orientations along the long/division axis of the cell (Figure 2.8 B). In contrast, most bipolar divisions were performed with nuclei positioned along the short axis of the cell, perpendicular to the cell's division axis. Of note, a fraction of binucleated cells exhibited nuclei orientations along the long axis, forming initial multipolar spindles but managed to resolve the multipolar geometry by centrosome clustering and still divided in a bipolar manner. These findings indicate that nuclei orientation and cell size have a pivotal impact on the spindle geometry and the number and ploidy of daughter cells.

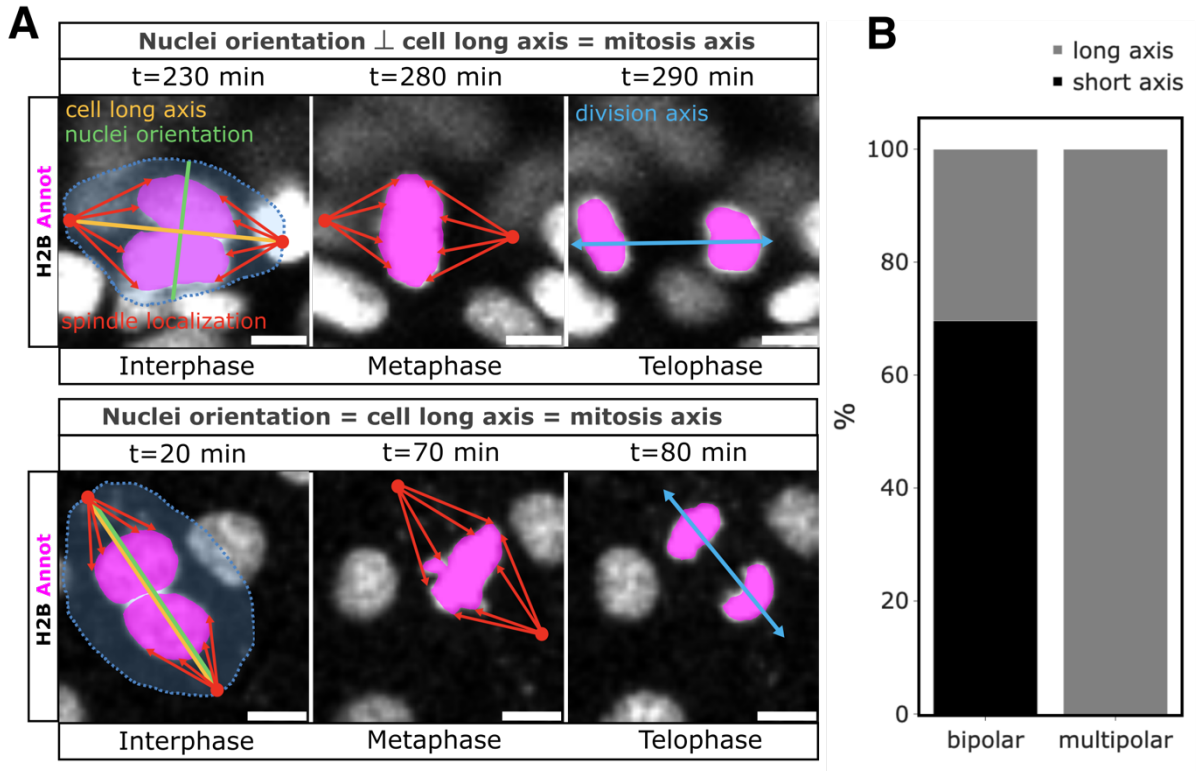


Figure 2.8 | Nuclear orientation dictates mitotic outcome.

(A) Live-cell image series from H2B-mCherry mADOs showcasing mitotic binucleated cells with different nuclear orientations. Top row: prior to mitosis, the nuclei orientation (green) is perpendicular to the cell's long axis (yellow) which corresponds to the division axis (blue). Red annotations indicate the centrosome localizations with red arrows highlighting the microtubule orientation. Bottom row: the nuclei orientation equals the cell's long axis and mitosis axis. H2B (white), nuclei annotation (magenta). Scale bar 5 µm. **(B)** Classification of the nuclei orientation either along the long or the short cell axis before multipolar mitoses based on live-cell imaging data from EGFP-Tuba mADOs. Data is shown as mean from 50 mitotic events (25 from bipolar and 25 from multipolar mitoses). Data collection and analysis for this figure was assisted by Daria Kocherhina, whom I supervised during her lab internship.

Taken together, my findings define a new role for binucleated pancreatic acinar cells as facultative progenitors, that can undergot ADM and divide at early stages of pancreas regeneration replenishing lost pancreatic tissue with two or more than two cells per mitotic passage, creating a highly dynamic change of nuclei numbers per cell and cell ploidies. This suggests that proliferating polyploid glandular cells are not exclusive to the liver. Rather, it underscores a conserved regenerative strategy across various glandular organs, wherein injury triggers proliferation of otherwise quiescent polyploid secretory cells. However, the concurrent occurrence of metaplasia-associated cell shrinkage may elevate the risk of multipolar mitosis, potentially leading to long-term complications.

2.3 mADOs show signs of chromosomal instability

2.3.1 Binucleated acinar cells possess mitotic errors and micronuclei accumulation in mADOs

My next aim was to investigate, whether "scheduled" polyploidy as it occurs in binucleated acinar cells indeed represents a vulnerable state for chromatin segregation errors and CIN. Studies of mammalian cell lines have demonstrated many times that polyploidy in its unscheduled form is perceived as a mutational process and thus is considered as a key step towards aneuploidy and the generation and evolution of cancer genomes (Duensing & Duensing, 2010; Storchova & Pellman, 2004). A major contributor to the vulnerability of a polyploid cell to acquire CIN, lies in its tendency to form multipolar spindle geometries which possess an especially high chance for merotelic microtubule attachments and lagging chromosomes (Cimini et al., 2002; Cosenza & Krämer, 2016; Gisselsson, 2008; Guerrero et al., 2010; Pidoux et al., 2000; Storchova & Pellman, 2004; Thompson & Compton, 2008, 2011). Thus, I acquired live-cell imaging data of H2B-mCherry mADOs and thoroughly screened mitotic events. I especially focused on early stages of mADO formation to capture the first binucleated cell divisions and acquire data as close as possible to an in vivo setting. I observed that the formation of mADOs is initiated by arranging cells from a primary acinus cell cluster into a spherical shape featuring a big lumen. This step lacks any cell divisions and recapitulates the formation of tubular complexes as seen in ADM in tissue. Strikingly, I found that the first binucleated cell divisions that occur after 24-48 h are frequently (~47 % $\pm$ 10.2 %) accompanied by chromosome segregation errors such as lagging chromosomes and DNA anaphase bridges (Figure 2.9 A).

Lagging chromosomes might end up forming micronuclei as they often are not integrated into the main nucleus. Since I frequently observed multipolar spindle formations as well as lagging

chromosomes and chromatin bridges in polyploid acinar cells during ADM, I screened for micronuclei by immunofluorescence staining of wild type mADOs (Figure 2.9 B). I detected micronuclei in almost all organoids with a peak of micronucleus abundance at 4 to 5 days of culture. This probably reflects the time when most of the binucleated cells in a forming mADO divided once. As describe previously, micronuclei are believed to play key roles in tumor initiation and evolution by promoting DNA damage and mutagenesis, which may give growth advantages to genetically unstable cells (Di Bona & Bakhoum, 2024; Mazzagatti et al., 2024). Their tumorigenic effect is linked to chromothripsis - the catastrophic shattering and reassembly of whole or partial chromosomes (Simovic-Lorenz & Ernst, 2025; C.-Z. Zhang et al., 2015). Micronuclei often lack nuclear membrane components such as Lamin A/C or Lamin B, leading to instability and rupture, which contributes to chromothripsis (Crasta et al., 2012; Hatch et al., 2013; Maciejowski et al., 2015, 2020; C.-Z. Zhang et al., 2015). While these events are frequently observed in cancer cells, engineered or treated cultures and tumors, they are rarely observed in normal cells, except in developmental contexts like germline and congenital disorders (Bertelsen et al., 2016; Kloosterman & Cuppen, 2013; Middelkamp et al., 2017). To assess whether micronuclei envelope instability and chromothripsis occur during ADM, I stained mADO cells for the nuclear envelope component Lamin A/C and the DNA damage marker γH2A.X. At d5, approximately 67 % of all micronuclei stained negative for Lamin A/C, indicating differences in micronucleus lamina integrity and composition and potential envelope instability in a majority of micronuclei (Figure 2.9 C, D). Notably, 44 % of the d5 micronuclei in mADOs exhibited massive DNA damage as indicated by strong γH2A.X immunofluorescence signals independent of the main nuclei, which can be used as a readout for the occurrence of chromothripsis events (Figure 2.9 D; Kalsbeek & Golsteyn, 2017; C.-Z. Zhang et al., 2015).

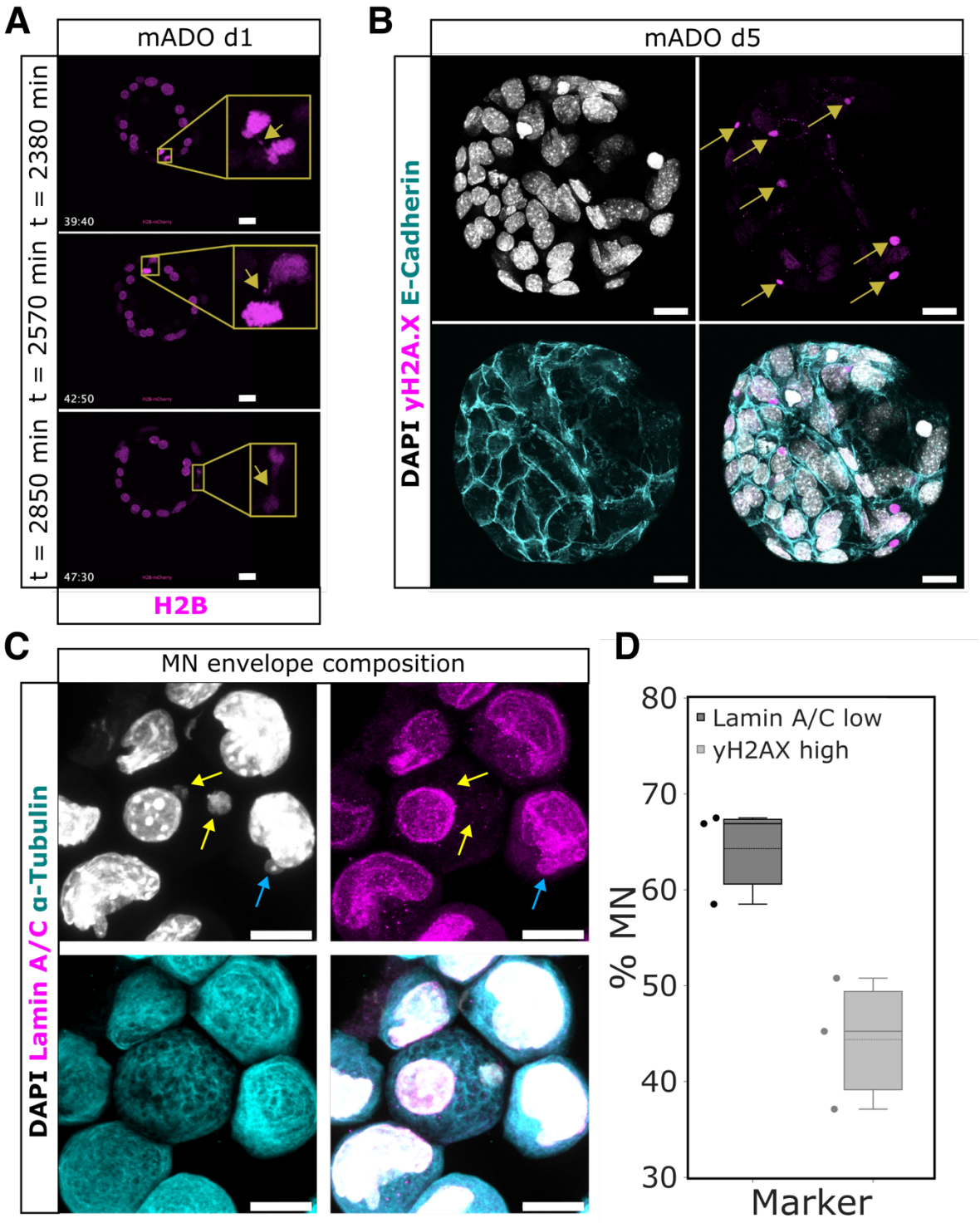


Figure 2.9 | mADOs exhibit signs of chromosomal instability.

(A) Live-cell image series from H2B-mCherry mice showing mitotic errors formed in proliferative binucleated cells at the onset of mADO formation (d1). H2B (magenta), yellow boxes indicate zoomed regions highlighting mitotic errors at different time points. Scale bar 20 µm. **(B)** IF staining from whole-mount d5 mADOs for γH2A.X (magenta), E-Cadherin (cyan) and nuclei using DAPI (white). Yellow arrows highlight micronuclei with strongly active DNA damage response in mADOs. Scale bar 20 µm. **(C)** Immunofluorescence images of dissociated d5 mADOs cells stained to assess micronuclear envelope composition for Lamin A/C. DAPI (white), Lamin A/C (magenta), α-Tubulin (cyan). Scale bar 10

µm. **(D)** Micronuclei (MN) characterization of d5 mADOs based on nuclear envelope composition (Lamin A/C low signal) or DNA damage response (γH2A.X high signal). Data is shown as percentage of all assessed micronuclei (n = 3 mice).

Apart from causing micronuclei and their potential downstream effects, lagging chromosomes can also represent obstacles for cleavage furrow ingression during cytokinesis (Cimini et al., 2002; Pidoux et al., 2000). Revisiting live-cell imaging data from H2B-mCherry/mG mice, I found that lagging chromosomes in mADOs eventually resulted in regressed cleavage furrows leading to failed cytokinesis and additional whole genome doubling of the single binucleated daughter cell (Figure 2.10). As the cleavage furrow is almost closed but then starts to regress, it rather indicates a failed attempt of cytokinesis than a omitted scheduled endomitosis.

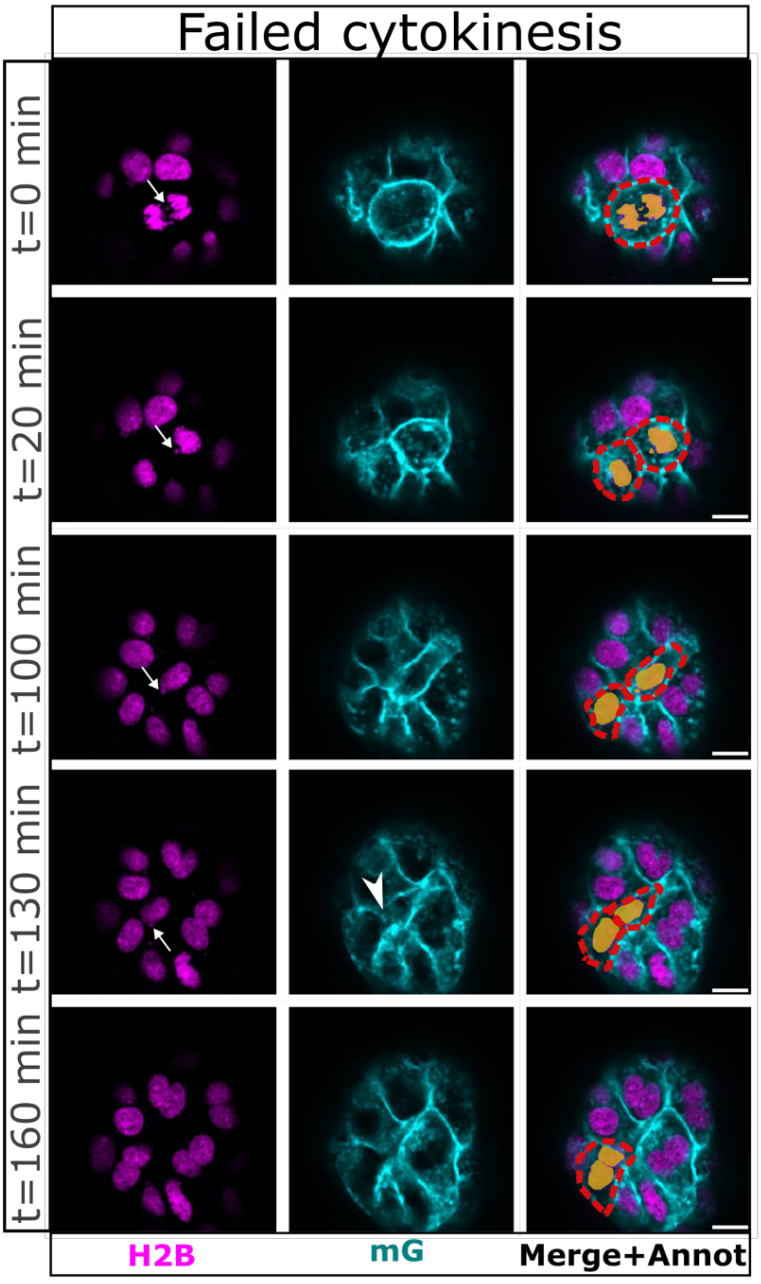


Figure 2.10 | Mitotic errors cause failed cytokinesis and hyperpolyploidization.
Live-cell image series of a d5 H2B-mCherry/mG mADO cell exhibiting a lagging chromosome with micronucleation (white arrow) and a subsequent failed cytokinesis resulting in a binucleated cell with hyperpolyploidization. White arrowhead indicates cleavage furrow regression. H2B (magenta), mG (cyan). Yellow annotation indicates nuclei, and red annotation indicates cell borders. Scale bar 10 µm.

2.4 Excursion: Addressing quantitative image analysis

This section deals with experimental and computational method development to quantitively assess cell properties from confocal images. The graphs in this section display representative data sets that I obtained during my PhD work but for better visibility do not represent the full data that was used in section 2.5 to draw actual biological conclusions. These representative graphs from method development are plotted with dark background.

2.4.1 Experimental layout to enable quantitative assessment of confocal images

As I observed that mADOs exhibit signatures of chromosomal instability, I evaluated different methods to i) quantitatively assess these signatures and ii) assign the underlying ploidy and nuclear number of the corresponding cells. A high throughput would certainly be possible using a flow cytometry-based method. Others reported that mononucleated and binucleated cardiomyocytes could be distinguished cytometrically by measuring the photomultiplier tube (PMT) signal width recorded from a fluorescent DNA dye (Windmueller et al., 2020). As the width of such a signal correlates with the time an object takes to pass the laser, a binucleated cell creates a larger signal width than a mononucleated cell even if the underlying DNA content is same. The signal area in turn, could be used to determine the overall DNA content of the cell. However, I tested this method using a BD LSRFortessa™ analyzer, but I was not able to receive a sufficient separation of the signal widths from mononucleated and binucleated acinar cells (data not shown). Possible reasons for this include a potentially low resolution of the PMTs of the BD LSRFortessa™ or the proximity of the two nuclei in acinar cells in contrast to a larger nuclear distance in other cell types such as cardiomyocytes (qualitative comparison of the images shown in (Yücel et al., 2020) with images from this thesis are sufficient to see the differences). However, apart from the nuclear number, the quantification of CIN itself had to be addressed as well. As chromosome segregation errors are of rather transient nature, I decided to focus on micronuclei as a less transient consequence of such errors. Again, micronuclei can be assessed by flow cytometry. By lysing the cells to obtain a nuclei suspension and measuring the forwards scatter area signal and area signals from a DNA dye it is possible to distinguish the larger and brighter primary nuclei from micronuclei. This method is regularly used to determine the genotoxicity of substances by exposing cells *in vitro* to different chemicals and measuring the relative micronuclei abundance in the aftermath (Rodrigues et al., 2021). To be able to distinguish micronuclei from debris and to acquire statistically relevant numbers of events, this approach requires a high abundance of micronuclei – at least in a control sample to set up the gates. As

the cellular yield of primary mADO cultures is generally low compared to 2D cell cultures from established cell lines and the problem regarding the nuclear number per cell persisted, I decided to develop a reproducible fluorescence imaging-based approach. Nuclear dyes such as DAPI or Hoechst, make the detection of binucleated cells and micronuclei easily accessible. To complement the detection of micronuclei, I aimed at measuring the DNA damage response by γH2A.X signals. Furthermore, I had to consider actively cycling cells, as a diploid cell in G2 would have a doubled DNA content despite still being diploid (2n/4C). This was possible by staining for the replication regulatory factor geminin. Geminin binds and inhibits chromatin licensing and DNA replication factor 1 (CDT1) in early S phase. CDT1 facilitates the loading of the minichromosome maintenance (MCM) helicase to the DNA which licenses an origin for replication. By negatively regulating replication, geminin prevents re-replication of DNA before mitosis. Geminin accumulates and remains bound to CDT1 until the end of mitosis. Thus, I used geminin staining, to fully exclude S/G2/M cells from the quantification as the determination of the underlying ploidy proves to be difficult. Finally, I stained for α-Tubulin as a marker for the cell body to enable instance segmentation of whole cells.

As mADOs are three-dimensional structures, fixing and staining as whole-mount organoids would make quantitative assessment of fluorescence signals difficult for two reasons. First, fluorescence signals from confocal microscopes change along the depth axis (z), with generally decreasing intensities the further away the focal plane locates from the objective. Fluorescent beads added to the culture could serve as a control to correct for depth-dependent signal loss but might create other artifacts and an overall less accessible and reproducible method, as one would ideally add beads corresponding to each optical channel to be recorded. Second, cell segmentation in 3D is much more elaborate regarding data preparation (i.e. labeling of training and testing data sets), more error prone due to algorithm complexity and computationally more intensive. Additionally, robust pre-trained models for 2D cell segmentation are often readily available and need less fine-tuning and adaption to new data sets for transfer learning. Thus, I decided to enzymatically dissociate mADOs prior to fixation and fixate single cells onto standard microscopy slides as a 2D monolayer using a Cytospin™ centrifuge. After fixation with 4 % paraformaldehyde, mADO cells could be stained for γH2A.X, geminin, α-Tubulin and nuclei using standard immunocytochemistry protocols and imaged by confocal microscopy as multi-image tile scans (Figure 2.11, see methods).

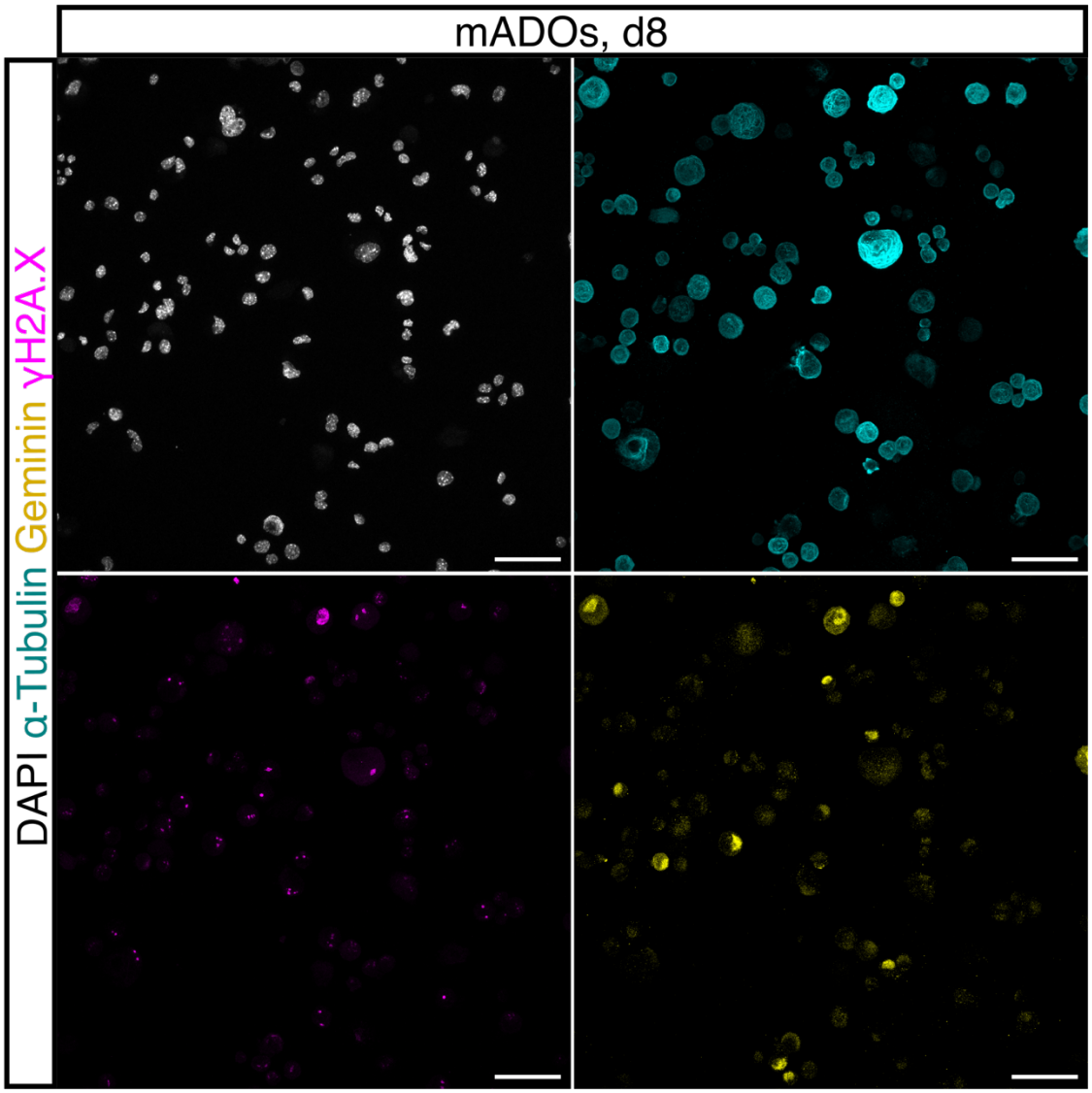


Figure 2.11 | Staining panel to assess ploidy, DNA damage and micronuclei from mADOs.
Representative IF images from mADOs that were dissociated into single cells, fixed at d5 and stained for α-Tubulin (cyan), γH2A.X (magenta), geminin (yellow) and nuclei (white) using DAPI. Scale bar 10 µm.

2.4.2 Computational analysis of ploidy levels and DNA damage response

This section describes the computational approach for single cell analysis of confocal images for cell ploidy, DNA damage response level, nuclear number and micronucleus abundance. A broad overview of the method including the previously describe experimental approach as well as instance cell segmentation, image classification and fluorescence signal analysis is shown in Figure 2.12.

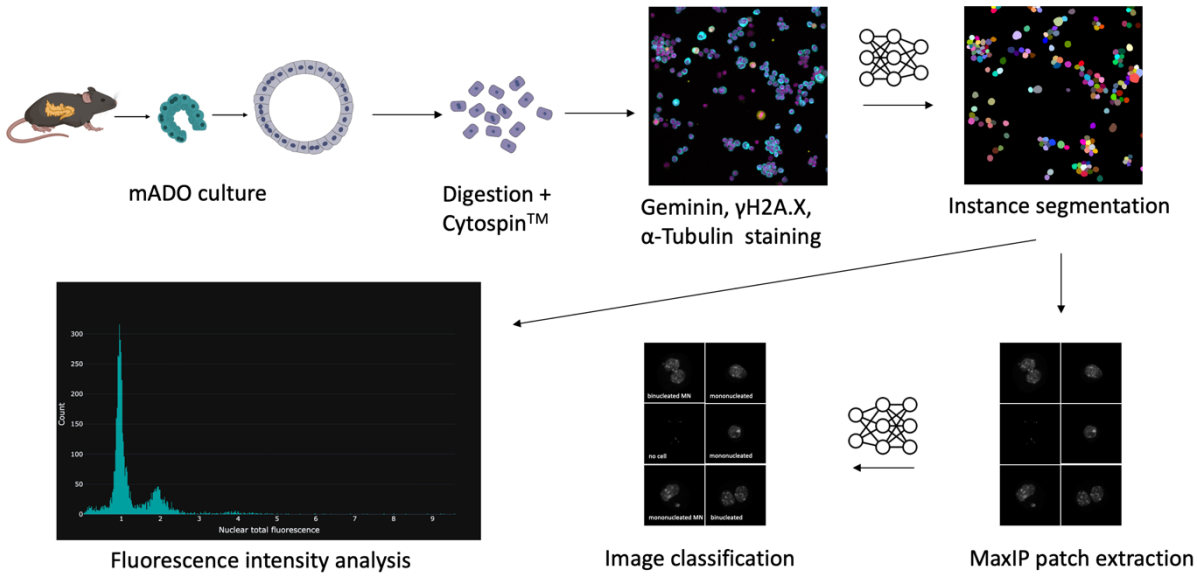


Figure 2.12 | Schematic overview of the experimental and computational workflow.

Schematic experimental and image analysis workflow to classify cells based on confocal images and assess ploidy state as well as γH2A.X signal intensities quantitatively. mADO cultures were dissociated into single cells, fixated onto microscopy slides using a Cytospin™ centrifuge and stained for geminin (cell cycle marker), α-Tubulin (cell body marker) and nuclei using DAPI to measure DNA content. Maximum intensity projections (MaxIP) of α-Tubulin confocal images were instance segmented using a custom Cellpose model. Each cell's MaxIP of the DAPI channel was cropped based on its cell segmentation shape and fed into a convolutional neuronal network to classify it according to either mononucleated, "mononucleated micronucleated", "binucleated", "binucleated micronucleated" or "no cell". Instance segmentation images were further used to extract various numerical features to correct for mis-segmented cells and assess measures such as DAPI, geminin and γH2A.X signal intensities.

As the raw confocal imaging data was acquired in the format of a 3D stack (of a 2D cell monolayer), I transformed those stacks into 2D maximum intensity projections (MaxIPs) and sum intensity projections (SumIPs). MaxIPs of the α-Tubulin images give a good representation of cell morphology with clear outer borders and were thus used to perform instance segmentation. However, as MaxIPs only consider the pixel value with highest intensity from each confocal z-slice per xy-pixel coordinate, I used SumIPs of the DAPI, γH2A.X and geminin channels for fluorescence intensity measurements.

To be able to extract features such as ploidy and DNA damage response from single cells, I needed to segment cells as individual objects. The segmentation mask is used to localize each cell within an image and represents its morphology and the segmentation label is additionally used as an identifier for each cell. I tested different approaches for cell instance segmentation including simple thresholding followed by watershed segmentation, Random forest-based classification using ilastik and deep-learning-based methods such as StarDist and Cellpose. Finally, I decided to stick to Cellpose because of two reasons. First, their pre-trained "Cyto2"

model already yielded decent segmentation results without any fine-tuning as it has been trained on fluorescent data of cell bodies from a 2D monolayer (Figure 2.13). Second, Cellpose 2.0's human-in-the-loop approach made it possible to fine-tune their existing models with minimal labeling effort and without the need of another software tool (Pachitariu & Stringer, 2022). Additionally, it made deep learning-based cell segmentation accessible to students that I supervised during my PhD work with minimal programming requirements.

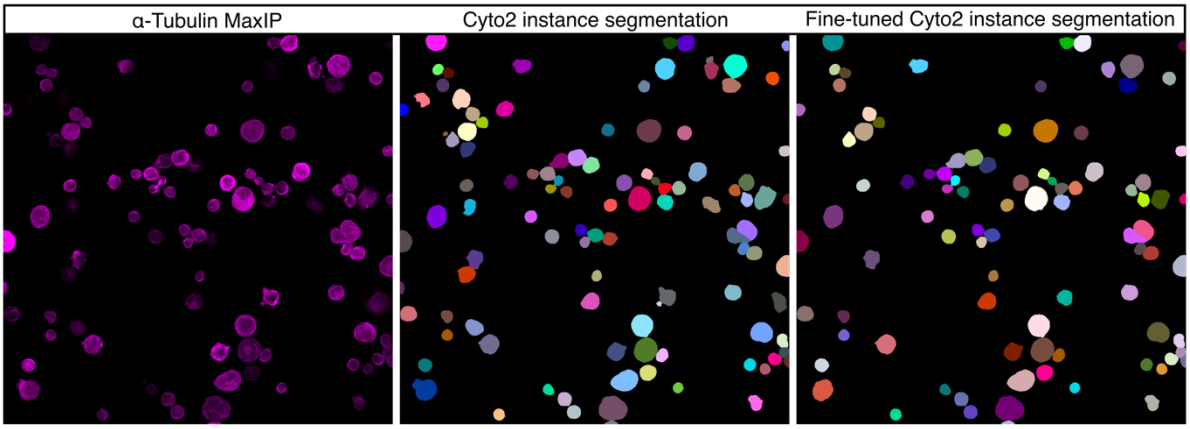


Figure 2.13 | Exemplary instance segmentation using Cellpose 2.0.

Left: α-Tubulin MaxIP input image from a d8 mADO. Center: instance segmentation using the pretrained Cyto2 model. Right: instance segmentation using a custom fine-tuned Cyto2 model, re-trained on 10 α-Tubulin MaxIPs (not including the shown).

Next, I had to extract corresponding cell morphology features such as the area and the perimeter, as well as fluorescence pixel intensity values for each cell to assign DNA content, cell cycle state and DNA damage response. The perimeter was later used to calculate the circularity for each cell to assess mis-segmented cells like doublets or cells touching the image border. The area was used to exclude mis-segmented objects such as debris with a very small area or clusters of multiple cells merged into one label with a very large area (Figure 2.14). Although the fine-tuned Cyto2 model from Cellpose 2.0 achieved very good segmentation results, these data curation steps further improved the overall quality.

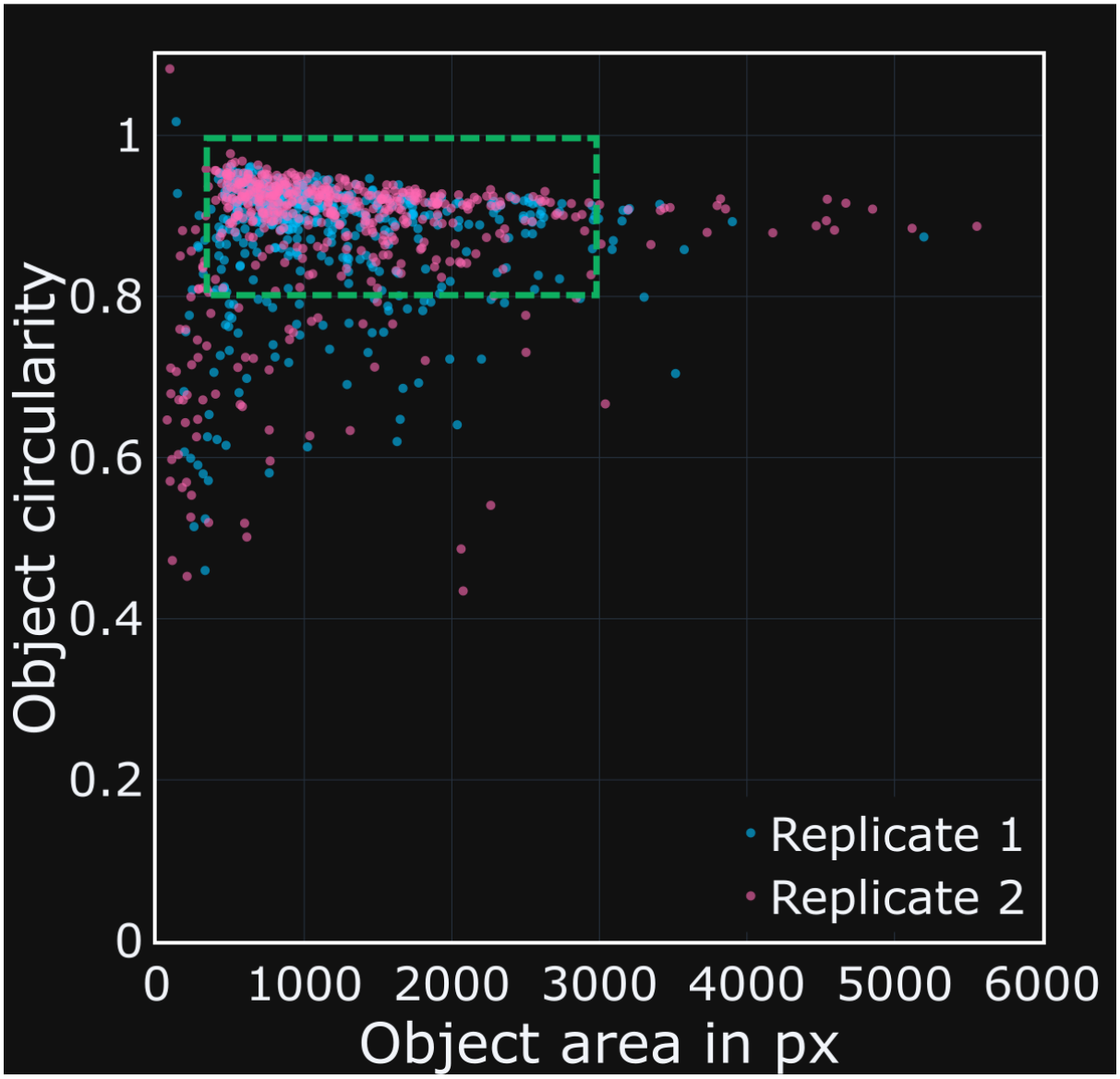


Figure 2.14 | Filtering segmented objects for single cells.

Circularity is plotted against area of segmented objects to exclude mis-segmented objects such as clusters of multiple cells or debris. The data shows two biological replicates from d8 mADOs.

To assess the DNA content of each cell I calculated the background **corrected total cellular fluorescence** (CTCF) of the DAPI channel (see methods). As the geminin expression and γH2A.X intensities correlate with DNA content, I calculated the background **corrected mean cellular fluorescence** (CMnCF) for those channels. Like fluorescence intensity values from flow cytometry experiments, the pixel intensity values from confocal images are relative and not absolute measures, meaning that they are not standardized across experiments or even across different days of recordings from the same experiment. This is because just by shutting off and turning on again the detectors and lasers, even with constant laser powers, detector voltages

and intensity thresholds, can result in shifted fluorescence intensity distributions. Thus, the background corrected intensity values CTCF and CMnCF can only be compared within one recording of multiple tile scans. To achieve comparability between the different recordings, I normalized the data to known distributions within the recorded data. For DNA content, I used the diploid 2n/2C population and for geminin and γH2A.X, I used the marker-negative populations. Figure 2.15 visualizes this process for two biological replicates of d8 mADOs recorded on two different days. To identify the respective populations in the data, I first scaled the CTCF and CMnCF data to values between 0 and 100 for each channel (DAPI, geminin, γH2A.X). Then I smoothed their density distributions using a Gaussian smoothing algorithm. Next, I detected the largest peak of the smoothed data by local maximum detection. This approach assumes that most of the cells in each image have a 2C DNA content, are geminin-negative and γH2A.X-negative which is the case in mADOs and was validated for each experiment by visual inspection of the respective histograms. To further refine the center of each distribution, I fitted a Gaussian distribution with the center of the largest detected peak as an initial guess for the mean. Finally, I normalized the scaled data to mean of the respective normal distribution. For the DNA content, I multiplied the normalized DAPI CTCF values additionally by the factor 2, so they interpretable as DNA content (C) or ploidy (n) levels, respectively. Classification of DNA contents, cell cycle stage and DNA damage response is then possible via simple thresholding of the normalized data (Figure 2.15). Another way of classifying the cell states based on the normalized fluorescence intensities is described in the next section and involves the usage of Plotly Dash apps.

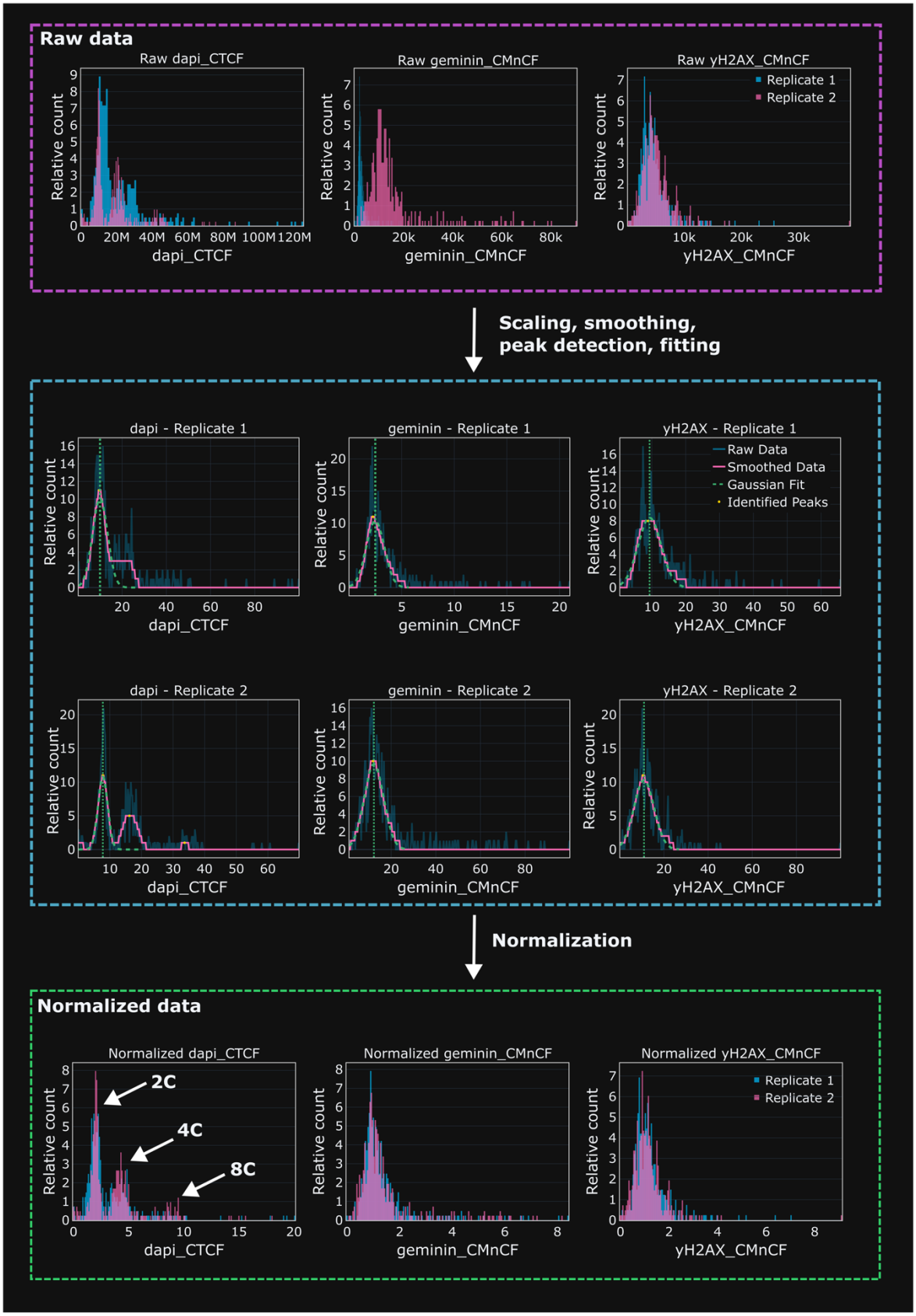


Figure 2.15 | Fluorescence intensity normalization.

Upper panel (magenta box): raw DAPI CTCF and geminin and γH2A.X CMnCF values for two biological replicates of d8 mADOs. Center panel (blue box): after scaling the raw CTCF/CMnCF density distributions (blue curves), Gaussian

smoothing (magenta curves), peak detection (yellow dots) and Gaussian fitting (green dashed curves), the center of the corresponding largest peak is estimated by the mean of the Gaussian fit (green vertical dashed line) for each channel and replicate. Bottom panel (green box): Normalization to the center of the largest peaks leads to aligned distributions. Peaks corresponding to different DNA contents can be easily distinguished (2C, 4C, 8C).

2.4.3 Dash apps to aid quantitative image analysis

Classification of the data that I acquired from image analysis requires a general understanding of the underlying biology. As shown in Figure 2.15, the assignment of DNA content states via the normalized DAPI CTCF is relatively straight forwards. However, for other distributions such as the geminin and yH2A.X CMnCF data, it is less clear where to set thresholds. So, when is a cell in G2 and when does it have activated DNA damage response? The big advantage of an image-based as opposed to for example a flow cytometry-based approach lies in the capability of additionally evaluating the data by visual inspection of the images. For human visual perception it is often easier to judge if a cell is positive for certain marker when observing microscopy images in context, compared to a more abstract data distribution. Additionally, comparing individual cells by visual inspection of the images and their corresponding data points in the graphs serves as a good control to validate the reliability of the method. As I could not find any appropriate software tools that combine an image viewer with the ability to load and plot corresponding numerical data and link data points back to objects in the image defined by a segmentation mask, I employed Plotly Dash to develop such a data analysis app (Figure 2.16). Dash is an open-source Python framework that can be used to build interactive web applications, especially for data visualization and analysis tasks. My dash app loads 2D images (i.e. the SumIPs from my experiments), corresponding instance segmentation masks and the numerical data tables generated by cell feature extraction. It is split into an “Image pane” and a “Plotting pane”. The image pane always displays the currently chosen image file from the “Image file” dropdown menu. It is possible to select individual channels, adjust the displayed contrast range and to overlay the corresponding label masks as yellow cell outlines if “Toggle labels” is active. The plotting pane shows a graph with all data points from the selected filters from the “Sample” and “Replicate” dropdown menus colored in cyan. This works if the file naming scheme of each image file has been given correctly. It always plots the properties selected under the “X” and “Y” dropdown menus. If both, “X” and “Y” have the same property selected, a histogram of the corresponding property is shown. Magenta data point represent cells that are currently visible in the image pane. It is also possible to click or select cells from the image pane that are displayed as yellow data points in the plotting pane. Vice versa, it is possible to click or select data points in the plotting pane to highlight cells in the image pane with blue outlines. Finally, the user can

select pre-defined markers or classes to classify data points selected in the plotting pane and save the class information for each classified cell in an exportable data table. In Figure 2.16, normalized geminin CMnCF (geminin_CMnCF_norm) is plotted against normalized DAPI CTCF (nuc_CTCF_norm) from three replicates of a WT mADO sample. The image pane shows one SumIP tile from replicate 1. The datapoints in the plotting pane were selected using a lasso tool and the magenta data points falling into this area are highlighted in blue in the image pane. It can be observed that most of the highlighted cells show a strong geminin expression. However, one highlighted cell at the upper image border exhibits a weaker signal compared to the other geminin positive cells. This cell probably corresponds to the magenta data point close to the lower border of the selected area. Thus, by visual inspection of the image data and reciprocal comparison to the numerical data, the user can evaluate whether it makes sense to rather strictly classify cells, in this case as G2, or not. In this example, it makes sense to select G2 cells rather generously because these cells would be excluded from the final quantification anyway. Since the classification filter works on the whole data set and not only on the cells that are shown in the image pane, it also allows for an unbiased classification across all replicates and samples. Even if the cell stage classification is done by simple thresholding from the extracted and normalized fluorescence intensities, the dash app proved to greatly enhance exploratory data analysis prior to thresholding and validation of applied thresholds in the aftermath.

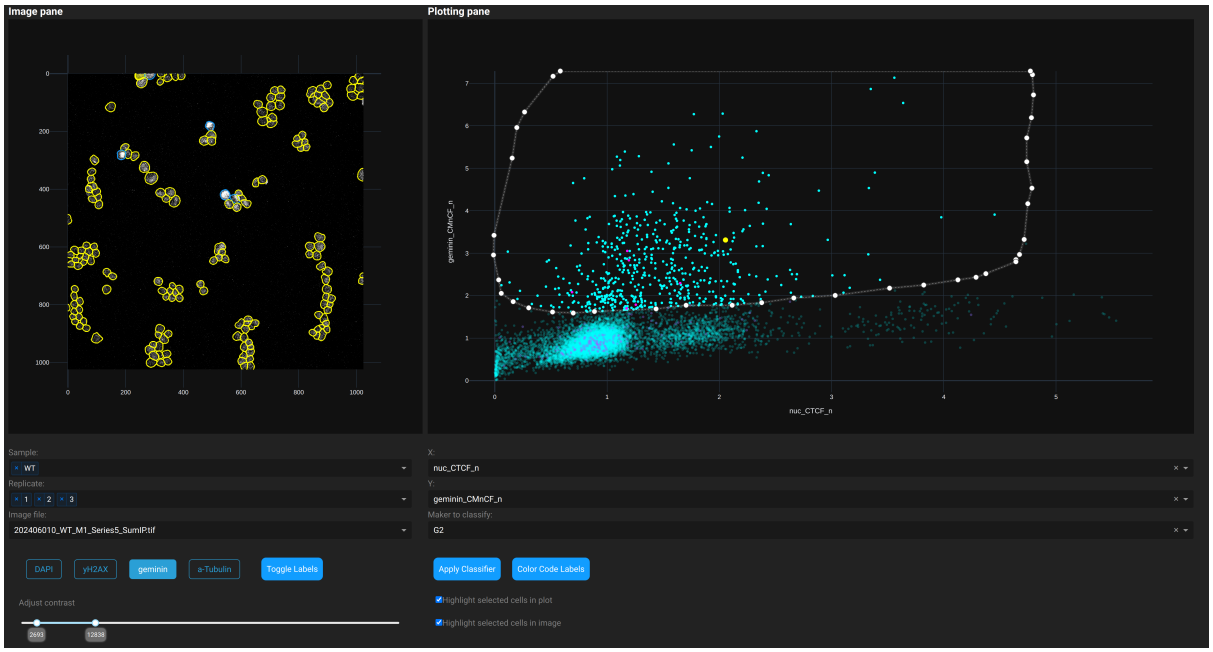


Figure 2.16 | Dash app user interface for cell state classification and exploratory data analysis.

The dash app is split into an image and a plotting plane and provides functionality to classify cell states based on measured image features such as fluorescence intensities by interactive reciprocal validation on images and measured data. Selected data shows geminin-positive (G2) cells.

One drawback of the way I plot the labels in the image pane and link them to the data points in the plotting pane is its high demand on computational power. Even with GPU accelerated plotting, large data sets of several tens of thousands of cells make the dash app much less responsive. To cope with such large data sets, I adapted the app to create a second, light weight version which only contains the plotting and classifications capabilities. Thus, large image data sets could be handled by using smaller subsets of the data for exploratory data analysis in the original dash app and the analysis on the full data set could be performed in the second version.

2.4.4 Nuclear number and micronuclei detection using CNNs

After determination of ploidy and DNA damage response, I aimed on linking the abundance of micronuclei to ploidy and nuclear numbers per cell. As micronuclei often occur as a result from merotelic attachments and lagging chromosomes, I assumed that the ploidy but also the nuclear count might play a role in the process of micronuclei generation in mADOs. To determine the nuclear count and micronuclei presence, I aimed to train convolutional neural networks for image classification. To be able classify each cell separately and link it to the properties I extracted from the previous analysis steps, I cropped each cell from the MaxIP images based on its α-Tubulin instance segmentation and saved it separately at a size of 90x90 pixels. Thus, the segmentation label could also serve as an identifier for the image classification procedure. I stratified these images into 5 different classes including “mononucleated”, “mononucleated micronucleated”, “binucleated”, “binucleated micronucleated”, and “no cell”. Originally, I also included a class for mitotic cells. However, as the number of clearly mitotic cells was relatively low, the classification performance generally did not yield as sufficient results as for the other classes. Furthermore, mitotic cells could already be detected by geminin staining, making the detection via image classification redundant.

Next, I had to find a proper CNN architecture. CNNs for image classification consist of a feature extractor component (convolutional and pooling layers) and a classifier at the end (fully connected layer). Training of a CNN from scratch, where all layers in the feature extractor part are randomly initialized, requires huge amounts of labeled data. Like the transfer learning approach that I used with the Cyto2 model of Cellpose for instance segmentation, it is possible to use pretrained CNNs for image classification. The open-source machine learning framework TensorFlow together with its high-level application programming interface Keras provides pretrained CNNs for various architectures in Python. Although these networks have usually not been trained on biological images, it can be reasonable to re-use (parts of) the feature extractor component. CNNs extract low level features (edges, gradients, textures) in early layers and high-level features (patterns, entire objects) in deeper layers. This means that a network that has been

trained on images showing round objects such as wheels from a vehicle or footballs, are likely able to extract features of other round objects, such as cells or nuclei. Thus, I decided to focus on networks that have been trained on the ImageNet data set (Deng et al., 2009). The ImageNet data set consists of approximately 14 million images from 1000 different classes including animals, vehicles, scenes, and everyday objects. As one of the most structured data sets with high quality labels, ImageNet is widely employed for benchmarking. It is furthermore used in the ImageNet Large Scale Visual Recognition Challenge, one of the most impactful computer vision challenges, as it sparked the deep learning boom in 2012 when AlexNet won the challenge by beating traditional methods by a huge margin (Krizhevsky et al., 2012). I experimented with various CNN architectures that have been trained on the ImageNet data set including VGGs, ResNets and EfficientNets. I received the best results using the EfficientNet B1 architecture consisting of 340 convolutional layers. Because my data set consisting of 1981 labeled images from five classes was much smaller compared to the ImageNet data set, allowing the network to train on all 340 layers poses a high risk of overfitting. Thus, it made sense to perform transfer learning only on a subset of the deeper layers for high-level feature extraction and the final fully connected layer for classification. The rest of network does not get updated by backpropagation, so it remains frozen. This preserves the low-level features already learned in early layers and keeps the network grounded in its generalized knowledge. I achieved best results by unfreezing the last 24 layers and adapting the fully connected layer for classification to 5 classes of the EfficientNet B1 network pre-trained on the ImageNet dataset. For other hyperparameter configurations, such as the optimizer algorithm, pooling method, batch size and classifier activation function, see methods section.

Training the network for 200 epochs at a learning rate of 10^{-5} using Sparse Categorical Cross Entropy as loss function, resulted in an overall model accuracy of 0.9469 (or 94.69 %) on the validation data and 0.9470 (94.70 %) on the training data at epoch 76 (Figure 2.17 A), indicating very good convergence. Accordingly, the training and validation loss functions converged at epoch 76 (Figure 2.17 B).

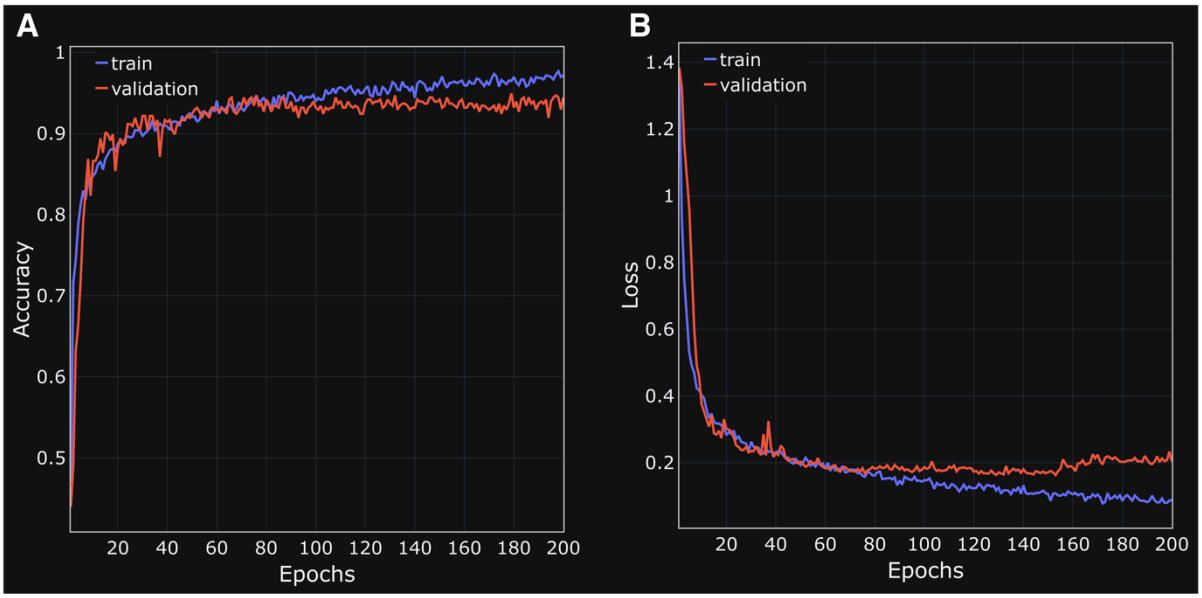


Figure 2.17 | Learning curves from training an EfficientNet B1 model on d8 mADO images.

(A) Accuracy curve for train (blue) and validation (red) sets. **(B)** Corresponding loss curve from Sparse Categorical Cross Entropy loss.

Further training resulted in divergence of the loss and accuracy curves with lower performance on the validation data and higher performance on the training data, indicating overfitting. Thus, the model at epoch 76 was used for further model validation on a test set consisting of another 396 images. The classification report is shown Table 1 for all five classes “mononucleated” (mononuc), “mononucleated micronucleated” (mononuc MN), “binucleated” (binuc), “binucleated micronucleated” (binuc MN) and “no. cell”. The network showed an overall very good performance across all five classes with most metrics reaching 85 % or higher. The overall accuracy on the test set reached 94 %. The “micronucleated” classes showed a slightly lower performance compared to their “non-micronucleated” counterparts. The “no cell” category had a 100 % positive rate for the test set, most probably because these images were either completely black or only showed some debris, making those images quite different from the other classes. The corresponding confusion matrix indicating the number of correctly and incorrectly classified images by comparing the predicted (columns) vs actual (rows) classes is shown in Table 2.

The results from the next sections regarding quantification of ploidy, DNA damage response, and micronuclei abundance were obtained by employing this image analysis pipeline.

Table 1 | Classification report for EfficientNet B1 prediction on 396 test images from mADOs.

Class	Precision	Recall	F1-score	Support
mononuc	0.97	0.95	0.96	177
mononuc MN	0.84	0.97	0.90	39
binuc	0.99	0.92	0.95	88
binuc MN	0.87	0.90	0.88	58
no cell	1.00	1.00	1.00	34

Table 2 | Confusion matrix for EfficientNet B1 prediction on 396 test images from mADOs.

	mononuc	mononuc MN	binuc	binuc MN	no cell
mononuc	169	0	0	8	0
mononuc MN	0	38	1	0	0
binuc	0	7	81	0	0
binuc MN	6	0	0	52	0
no cell	0	0	0	0	34

2.5 Chromosomal instability in mADOs is linked to polyploidy

As mentioned, micronuclei can be more easily detected, compared to transient chromatin segregation errors. Thus, I next addressed the quantification of micronucleated cells and γH2A.X signal intensities in mADO cells as well as their underlying DNA content. At d5 in culture, I found that the amount of γH2A.X high cells is strongly increased in the polyploid fraction of mADO cells (Figure 2.18 A). Additionally, my analysis revealed that the ploidy distribution of γH2A.X high cells is generally shifted towards cells with higher DNA content, with many cells showing an intermediate ploidy, presumably indicating aneuploid chromosome sets (Figure 2.18 B). In contrast, γH2A.X low cells showed a more distinct bimodal distribution of diploid and tetraploid cells.

I further classified segmented cells at d5 by the presence of micronuclei based on the underlying DNA content and nuclear number. Here, I observed the highest abundance of micronuclei in binucleated polyploid cells (9.37 %; Figure 2.18 C). Mononucleated polyploid cells still exhibited a greater proportion of micronucleation compared to mononucleated diploid cells. These results indicate that micronuclei formation often co-occurs with cleavage furrow regression leading to failed cytokinesis as described above. My assay does not consider cells that underwent

multipolar mitoses and ploidy reduction to become diploid while still being able to carry micronuclei generated in the ploidy reducing mitoses or earlier. This implies that the true number of micronuclei linked to an initially polyploid genome could even be higher.

To validate my findings in an in vivo setting, I revisited the cerulein-induced pancreatitis model and stained tissue sections at 4 dpi for α-Amylase and CK19 to detect acinar cells that underwent ADM. I regularly found binucleated α-Amylase/CK19-double-positive cells carrying a micronucleus (Figure 2.18 D). Of note, no micronuclei were detected in naïve freshly isolated acinar cells, indicating that those micronuclei were formed during the inflammatory response against the cerulein treatment.

My results indicate that early cell divisions of polyploid acinar cells undergoing ADM are especially prone to chromatin segregation errors leading to micronuclei formation with extensive DNA damage as seen in chromothripsis.

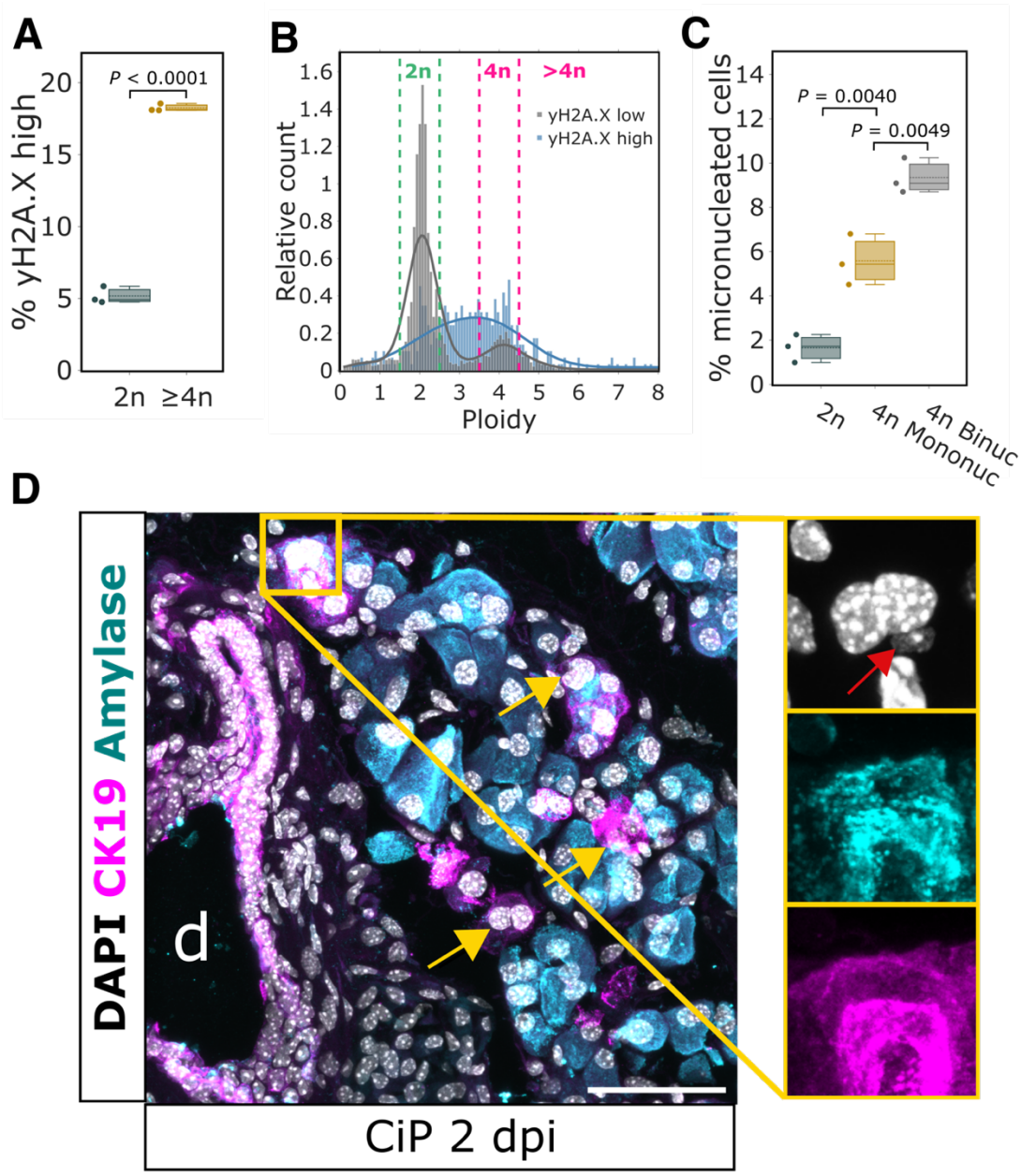


Figure 2.18 | Chromosomal instability is linked to the polyploid state in mADOs.

(A) Box plot showing γH2A.X intensity quantification in d5 mADOs comparing diploid (2N) and polyploid (4N) cells. *P*-value was calculated using a two-sided paired Student's *t* test (n = 3 mice). **(B)** Ploidy distributions by histograms and KDE curves of γH2A.X high and γH2A.X low cells in d5 mADOs (n = 3 mice). Data is shown as histogram with bin size 0.07 and Kernel Density Estimation. **(C)** Micronuclei abundance in d5 mADO based on percent of micronucleated cells according to ploidy and nuclear number: diploid (2n, cyan), polyploid mononucleated (4n mononuc, yellow), polyploid binucleated (4n binuc, grey). *P*-values were calculated by one-way ANOVA followed by Tukey's post hoc test (n = 3 mice). **(D)** IF image from mouse cerulein-induced pancreatitis at 2 dpi. Pancreatitis sections stained against α-Amylase (cyan) and CK19 (magenta). Yellow magnification box shows micronucleated acinar cell with red arrow indicating micronucleus. Yellow arrows indicate CK19-/Amylase-double-positive binucleated acinar cells, d: duct. Scale bar 50 µm. This image was recorded with assistance from Sanzhar Aitbay, a student, whom I supervised for his master's thesis.

2.5.1 Micronucleated cells in mADOs remain proliferative

An ongoing question, that is regularly discussed in the light of tumor formation and evolution, addresses the fate of aneuploid/micronucleated cells. Excessive unrepaired DNA damage often leads to cell cycle arrest paired with increased autophagy which ultimately might lead to cell death and clearance (De Zio et al., 2012; Matt & Hofmann, 2016). Such protection can normally be facilitated by cell-intrinsic pathways like ataxia telangiectasia mutated (ATM) and ataxia telangiectasia and Rad 3 related (ATR) mediated p53 stabilization or the cyclic GMP-AMP synthase (cGAS) - Stimulator of interferon genes (cGas-STING) signaling axis – although the contribution of cGA-STING signaling to sensing chromothripsis-induced DNA damage from micronuclei is still under debate (Abuetabh et al., 2022; Flynn et al., 2021; Mackenzie et al., 2017; Takaki et al., 2024). In contrast, it was shown that micronucleated cells can re-enter cell cycle and pass their aneuploid genome to daughter cells (Hintzsche et al., 2017). Importantly, these findings were obtained from either immortalized, or cancer cell lines or primary cells treated with cytotoxic chemicals and/or ionizing radiation. Thus, it is debatable to what extend micronucleated cells re-enter cell cycle and complete further mitoses under physiological conditions. Using live-cell imaging of H2B-mCherry mADOs, I found that micronucleated cells are indeed capable of maintaining proliferative capacities (Figure 2.19 A). In these cells, the micronucleus eventually gets reintegrated into the main nucleus and further mitotic defects can be commonly observed. Those reintegration events would be especially likely if the micronuclear DNA contains centromeric regions, which possess a certain degree of protection against DNA damage and can facilitate microtubule binding.

To further delineate the full process of a mitotic error with micronuclei formation and subsequent proliferation, I subjected H2B-mCherry/mG mADOs to long-term live-cell imaging over the course of 4 days, starting at d3 and using an inverted light-sheet fluorescence microscope. Tracking of the individual nuclei and cells based on the mCherry and GFP signals allowed me to reproduce the lineage relations in the growing organoid (Figure 2.19 B). First, I noticed that the two daughter cells from an aberrant binucleated cell division behaved differently. While the micronucleated sister cell remained proliferative (highlighted in red), the other sister cell stopped proliferating, putatively because of the chromosome loss. Second, the proliferative micronucleated sister cells did not show growth disadvantages compared to other cells and had a similar cell cycle length. Generally, I measured an average doubling rate of 16.2 h. A similar experiment has been performed on mouse intestinal organoids, that acquired mitotic errors due to a highly proliferative phenotype with a rapid doubling rate of approximately 10 h (de Medeiros et al., 2022). While these cells potentially acquired mitotic errors because of their high (error

prone) turnover rate, I propose that the tendency of pancreatic acinar cells to perform aberrant mitoses is inherently linked to the combination of polyploidy/binucleation and metaplasia.

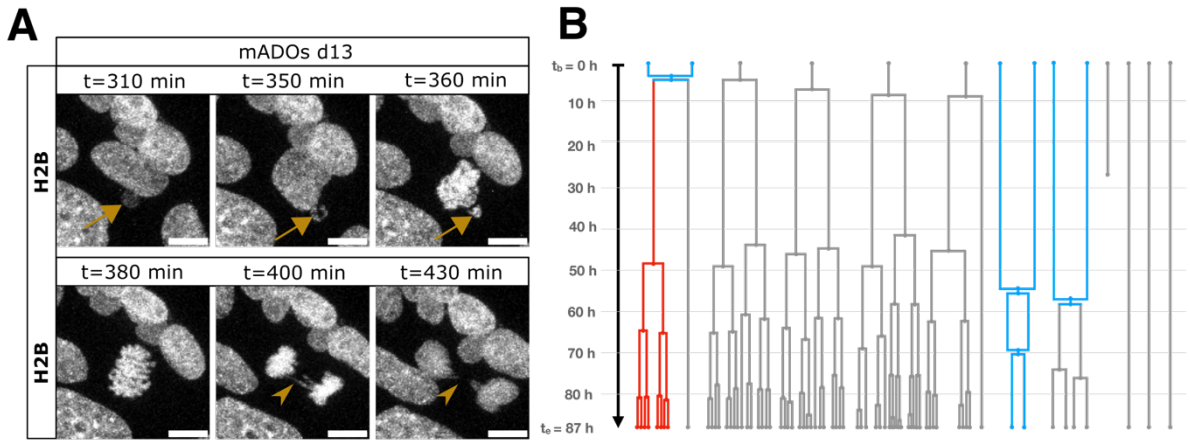


Figure 2.19 | Micronucleated mADO cells remain proliferative.

(A) Live-cell image series of a d13 H2B-mCherry mADO. White arrow indicates micronuclei reintegration upon mitosis. White arrow heads indicate a subsequent chromatin bridge in the same cell. Scale bar 20 µm. **(B)** Reconstructed lineage tree of mADO long-term live-cell imaging over 87 h (d3-d7). Grey tracks indicate mononucleated, blue tracks indicate binucleated cells. Red tracks indicate the lineage of a micronucleated cell from a binucleated mitosis with retained proliferative capacities.

2.5.2 Genomic analyses showcase consequences of mitotic errors in mADOs

To assess the genomic landscape of polyploid acinar cells after proliferation on the molecular level, I subjected cells from day 9 mADOs to single-cell whole genome sequencing. I employed Strand-seq, which gave me two advantages. First, by sequencing only the template strand of a single nucleus, haplotype information enables the distinction of homologous chromosomes and the detection of complex genomic rearrangements. Each read can be assigned to a specific strand orientation defined as either Watson (W) or Crick (C). Second, Strand-seq requires the processing of cells which have BrdU incorporated into the nascent strand from the previous replication phase. This allowed me to enrich for cells that have divided and potentially underwent chromatin segregation errors. To specifically sort for diploid and polyploid nuclei, I used a two-step fluorescence-activated cell sorting (FACS) strategy based on the fluorescent ubiquitylation-based cell cycle indicator 2 (Fucci2) model (Figure 2.20). As this mouse model labels geminin with mVenus, this enabled me to specifically exclude S-, G2- and M-phase cells – similar to my image analysis pipeline – which might possess an increased DNA content despite being diploid. Additionally, the Fucci2 model expresses chromatin licensing and DNA replication factor 1 (CDT1) fused to mCherry, hence labeling cells in G0/G1. After sorting for G0/G1 cells based on mCherry signal, cells were lysed to extract single intact nuclei. BrdU incorporation could be

determined by the detection of Hoechst33342 quenching in the second sorting step, while propidium iodide (PI) labeling was used to resolve the DNA content of the nuclei. This led to four distinct populations in the second sorting step, two for each ploidy level (2n, 4n) and two for BrdU incorporation (quenched vs. unquenched) Hoechst33342 signal.

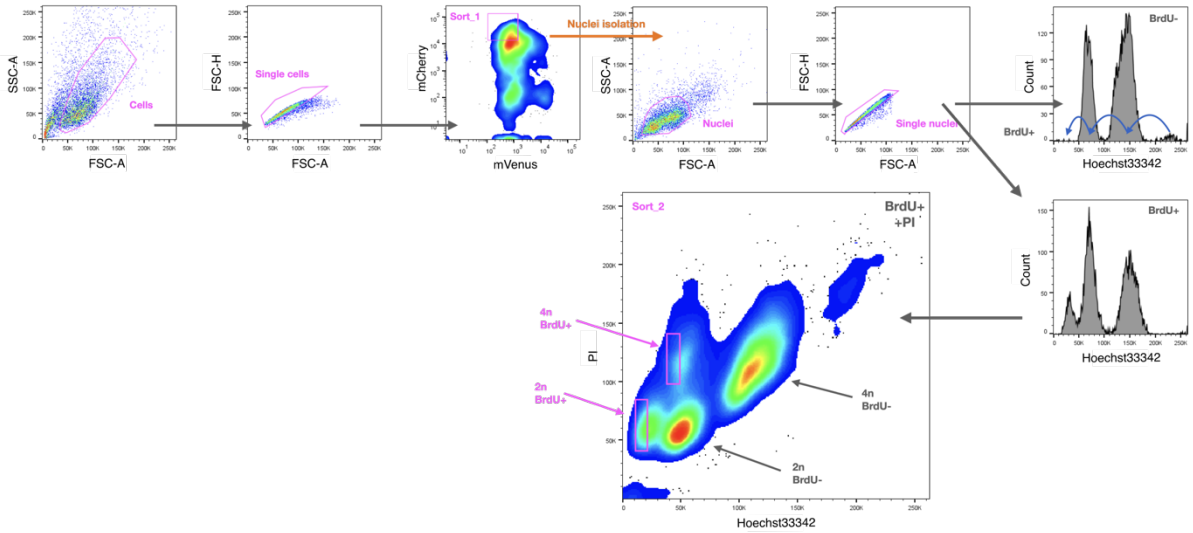


Figure 2.20 | Sorting strategy to sort single nuclei for Strand-seq according to ploidy.

Sorting strategy to sort single nuclei for Strand-seq. An initial sorting step consisted of sorting mCherry$^+$ cells from Fucci2, indicating non-cycling cells (sort_1). After lysis, doublet nuclei were excluded and BrdU incorporation was assessed by quenching of Hoechst33342 (Hoechst 33342 signals roughly half upon BrdU incorporation). Blue arrows in the BrdU$^-$ (no BrdU during mADO culture) control graph indicate, where the fractions of different ploidy levels would be expected to shift if BrdU is incorporated. BrdU$^+$ graph shows the sorting from a BrdU treated sample. Propidium iodide (PI) is used to separate the Hoechst33342 peaks by ploidy, as PI is not affected by BrdU incorporation and to sort BrdU$^+$ nuclei for Strand-seq (sort_2). Magenta annotations indicate gating. 2n: diploid nuclei, 4n: tetraploid nuclei, SSC-A: side scatter area signal, FSC-A: forward scatter area signal, FSC-H: forward scatter height signal.

The library preparation was performed by Katharina Bauer from the single-cell open lab of the German Cancer Research Center (DKFZ) Heidelberg. I sequenced 96 nuclei and obtained 42 high-quality genomes (22 from diploid and 20 from tetraploid nuclei). The computational analysis of the Strand-seq data was performed by Enrico Frigoli, a bioinformatics PhD student in the Martin-Villalba lab. The average ploidy of each cell was assessed by visual inspection of the Strand-seq results as shown in Figure 2.21 A and B. Enrico Frigoli's method of visualizing the Strand-seq data for each chromosome separately, furthermore allowed to identify sister chromatid exchanges and structural abnormalities (Figure 2.22 C).

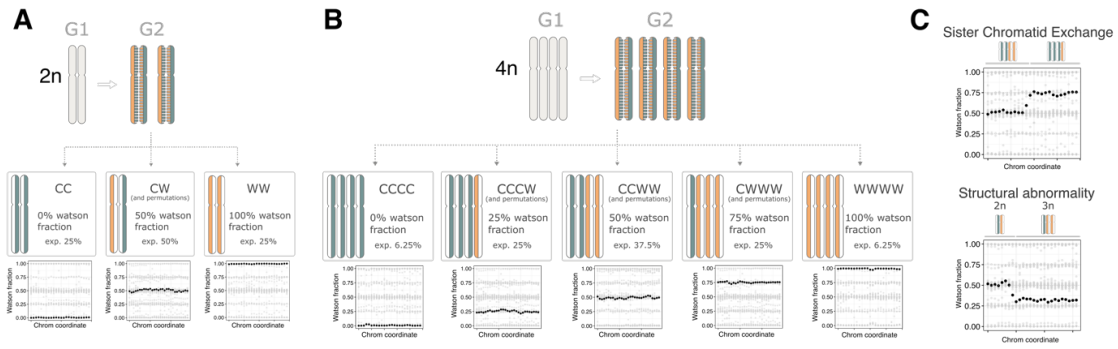


Figure 2.21 | Distinction of ploidy levels and structural rearrangements in Strad-seq data.

(A, B) Theoretical distributions of Watson fractions (WF) - defined as the proportion of sequencing reads aligned to the Watson strand relative to the total number of aligned reads within a genomic bin - are shown for diploid (A) and tetraploid (B) loci. During DNA replication, each strand has a 50 % probability to be used as template, resulting in possible configurations of Crick-Crick (CC), Crick-Watson (CW), Watson-Crick (WC), and Watson-Watson (WW) in the daughter cells with diploid chromosomes, with analogous configurations for tetraploid loci. Plotting the WF of genomic bins across chromosomes allows for the inference of cellular ploidy, as some values can only be explained by a specific ploidy/copy number value (within a reasonable range), e.g. 33 % and 66 % are specific for 3n, 25 % and 75 % for 4n. Additionally, combining WF data with normalized total read counts enables the estimation of chromosome copy numbers when WFs take ambiguous values such as 0 % and 100 %. **(C)** The same type of visualization allows to spot sister chromatid exchanges (top panel) and large structural variants such as duplications and deletions (bottom panel). Plots prepared by Enrico Frigoli.

He identified five cells exhibiting whole-chromosome aneuploidies and/or major structural variants (SVs) of which 3 had a tetraploid basal ploidy. Among these, one cell exhibited a complex ploidy structure characterized by several chromosomes gains or losses, or having larger segments deleted, while other chromosomes remained intact and were present in two or four copies (Figure 2.22 A). Notably, the presence of extensive deletions, especially of chromosome 11 suggests that the deleted regions may have originated from a lagging chromosome that was lost during mitosis. Another tetraploid cell possessed aneuploidies including the loss or gain of a full chromosome copy leading to a pentaploid state, and triploid loci (Figure 2.22 B). These findings demonstrate that organoids derived from primary naïve acinar cells can acquire features typically found in tumors and that polyploid cells are likely to be more susceptible to these transformations. Again, cells that might have been polyploid initially and reduced their ploidy due to multipolar mitoses could not be identified by this method.

Mitotic errors and catastrophic DNA damage has not been linked to the natively occurring polyploidy found in acinar cells so far.

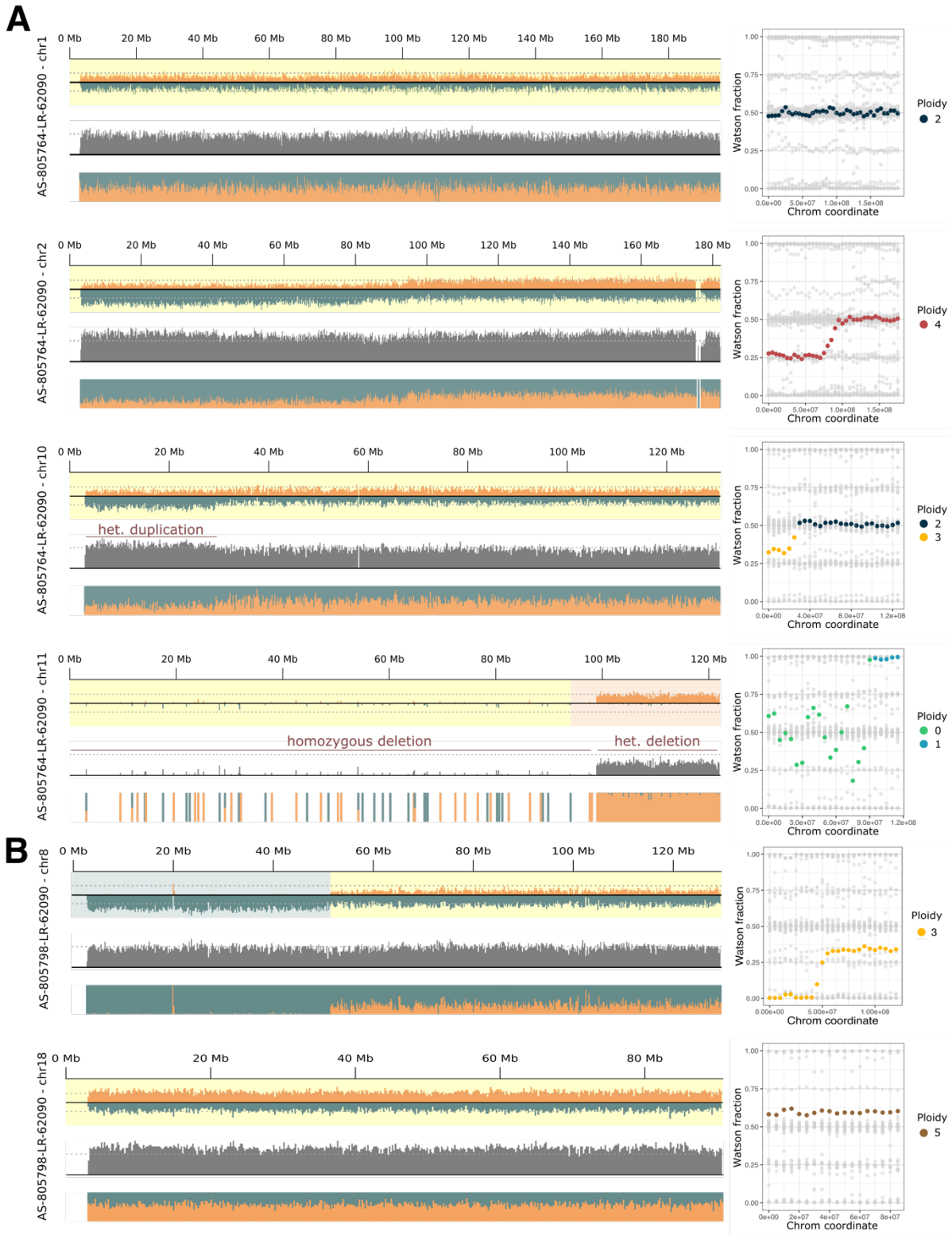


Figure 2.22 | Strand-seq identifies structural abnormalities in mADOs.

(A) Single-cell Strand-seq plot showing: a diploid chromosome (chr1), a tetraploid chromosome (chr2), a heterozygous duplication at the start of chr10, and a deletion in chr11. Left panels display coverage tracks with read depth divided by Watson and Crick strands, total read depths, and normalized Watson/Crick fractions, at a 200 kilobases (kb) bin resolution. The dotted lines represent the median value for the cell. Right panels show Watson fraction versus chromosomal (Chrom.) coordinates, colored by ploidy, to highlight genomic gains or losses across chromosomes, in

5 megabase (Mb) steps (smoothed over a 10 Mb window); gray points in the background represent data from the same chromosome in other cells. **(B)** Single-cell Strand-seq plot showing a triploid chromosome (chr8) and a pentaploid chromosome (chr18) from a tetraploid cell analogous to **(A)**. Plots prepared by Enrico Frigoli.

2.6 Conserved principles of tissue remodeling, polyploidy, and chromosomal instability across glandular organs

2.6.1 Establishment of polyploid lactating mammary gland-derived organoids

Unlike the pancreas and the liver, which retain polyploid cells throughout life, the mammary glands in both humans and mice generate polyploid cells only during pregnancy and lactation, a period when milk production is essential to nourish offspring (Rios et al., 2016). During pregnancy, luminal cells in the mammary gland differentiate into secretory alveolar cells that produce and secrete milk. Following weaning, the gland undergoes post-lactational involution, a process in which these alveolar cells either return to a quiescent state or undergo apoptosis, resulting in a reduction in organ size (Oakes et al., 2006). This dynamic transition makes the mammary gland an especially valuable model for studying naïve polyploidy, as the non-lactating gland provides an ideal physiological control. Moreover, both preclinical and epidemiological evidence suggest that improper regulation of post-lactational involution may play a role in tumorigenesis (Jung et al., 2022; Lyons et al., 2011; Radisky & Hartmann, 2009). To broaden the relevance of my findings from the pancreas, I therefore chose to investigate the mammary gland as a complementary system. To this end, Veronika Lummer, a medical doctor student, whom I supervised for her thesis, established lactating mouse mammary-derived organoids (mMDOs) based on a published protocol (Sumbal et al., 2020). During an initial morphogenesis phase where isolated epithelial mammary gland cells are cultured in medium containing FGF2 (morphogenesis medium, MOM) for 6 days, a known regulator of mammary branching development, the growing organoids undergo branching morphogenesis (Figure 2.23, Figure 2.24 A; Ewald et al., 2008; Rivetti et al., 2020). The lactating capacities of these organoids mainly rely on prolactin (Prl) treatment after the morphogenesis phase (Sumbal et al., 2020). Besides prolactin to induce milk production, Veronika additionally added other known pregnancy-related factors - namely progesterone (P_4), estrogen (E_2) and EGF to stimulate the formation of polyploid cells. Estrogen induces ductal elongation and further branching during pregnancy and additionally increases the expression of prolactin receptors (Arendt & Kuperwasser, 2015). Progesterone is a known mediator of alveolar growth and differentiation in synergy with prolactin (Brisken et al., 1998; H. J. Lee & Ormandy, 2012). EGF is further a critical regulator of cell proliferation and has been shown to play important roles in binucleation of luminal mammary

gland cells (Borellini & Oka, 1989; Rios et al., 2016). After the initial morphogenesis phase, mMDOs were cultured in this hormone-enriched medium (pregnancy-lactation medium, PLM) for another 6 days (Figure 2.23).

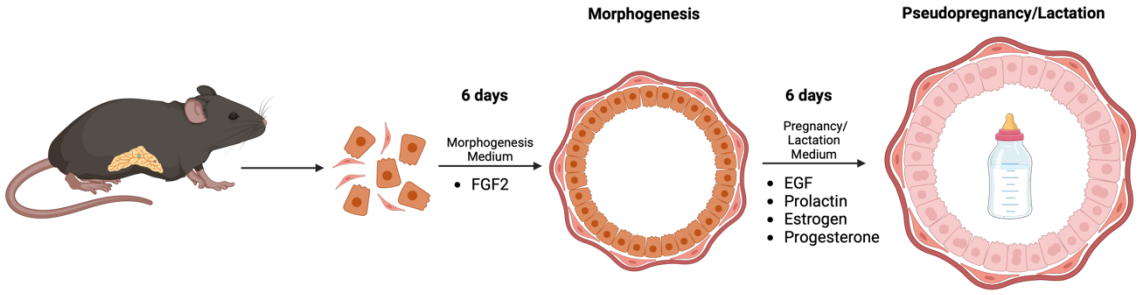


Figure 2.23 | Workflow to generate lactating polyploid mammary gland-derived organoids.
Experimental workflow to generate lactating mouse mammary gland organoids (mMDOs) consisting of diploid and polyploid secretory cells. The procedure includes cells extraction, 6 days morphogenesis supplementing with FGF2 and 6 days of pseudopregnancy/lactation including polyploidization and milk production induced by EGF, prolactin, estrogen and progesterone.

IF staining for the luminal cell marker keratin 8 (K8) and the myoepithelial cell marker keratin 5 (K5) indicates a near-physiologic cell composition and gland architecture, with a branched luminal epithelium supported by myoepithelial cells in mMDOs (Figure 2.24 A; C. M.-C. Li et al., 2020). Using my previously described image analysis pipeline for the quantitative assessment of marker expression and DNA content based on fluorescence intensities from confocal images, I examined mMDO cells for various properties. Another student, Sefa Berkay Cayir, whom I supervised for his master's thesis project, assisted during these wet lab experiments. First, I investigated the cellular composition of mMDOs based on K8. I assessed the abundance of K8-positve cells in mMDOs exclusively cultured in MOM for 12 days or MOM for 6 days and PLM for another 6 days. To isolate the individual effects of key pregnancy hormones and EGF, each factor was selectively omitted from the culture medium in separate cultures. Quantification revealed that 66.24 % of organoid cells were K8-positive following treatment with morphogenesis medium, indicating a predominance of luminal lineage cells (Figure 2.24 B). Subsequent exposure to pregnancy hormones did not significantly alter this composition, with the proportion of luminal cells ranging between 63.31 % and 73.07 %, showing no statistically significant differences. These data confirm the near-physiologic cell composition in mMDOs (C. M.-C. Li et al., 2020) .

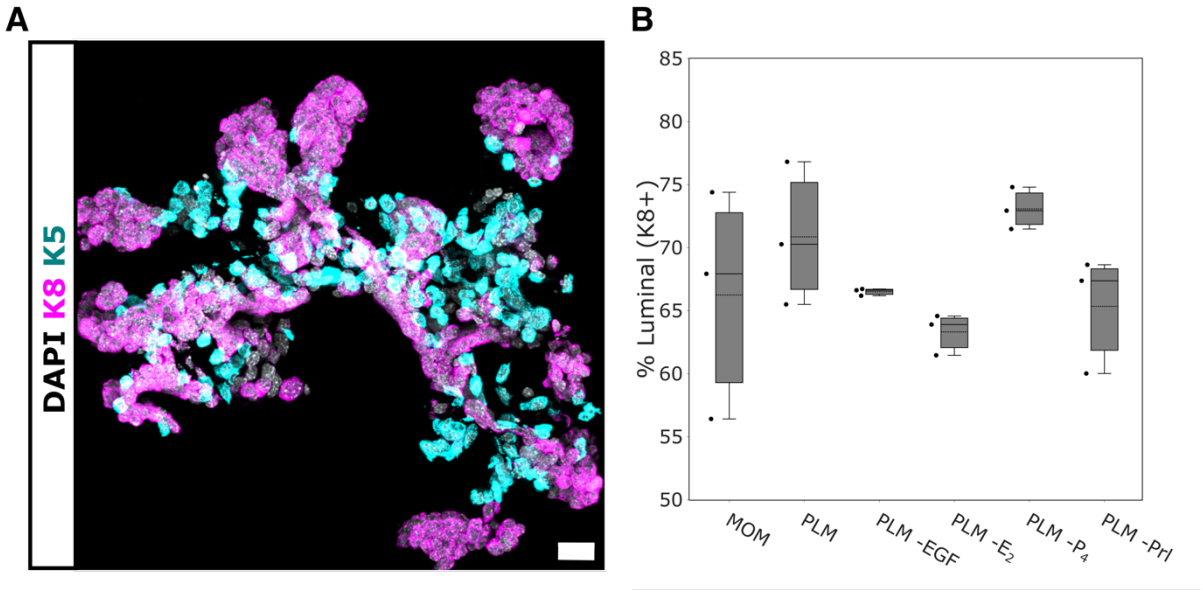


Figure 2.24 | mMDOs resemble *in situ* architecture and cell composition.

(A) IF image of mouse mammary gland-derived organoid (mMDO) stained for the luminal cell marker keratin 8 (K8, magenta) and the myoepithelial cell marker keratin 5 (K5, cyan). Scale bar 20 µm. Image acquired by Veronika Lummer**.** **(B)** Quantification of luminal cells (K8+) in mMDOs treated with morphogenesis medium for 12 days (MOM) or morphogenesis medium for 6 days followed by different combinations of hormones in pregnancy-lactation medium (PLM) for 6 days. *P*-values (not shown) between culture conditions were calculated by one-way ANOVA followed by Tukey's post hoc test (n = 3 mice) but did not yield values below 0.05.

Next, I addressed the lactating capacities of mMDOs. IF staining of whole-mount mMDOs for the human and murine milk protein β-casein, revealed β-casein producing cells and milk storage in mMDO lumina (Figure 2.25 A). Quantitative examination of the abundance of β-casein-positive cells confirmed the lactogenic effect of prolactin (Figure 2.25 B). Culture conditions lacking prolactin failed to induce β-casein expression, whereas all media containing prolactin led to the emergence of β-casein–positive cells, with proportions ranging from 1.82 % to 5.46 %. No significant differences were observed among the prolactin-containing conditions. Since no β-casein–positive cells were detected under prolactin-deficient conditions, a traditional statistical test comparing conditions with and without prolactin was not appropriate. Instead, Cohen's d was calculated to assess the effect size, revealing a large effect (Cohen's d: PLM/MOM = 2.88; PLM/PLM -Prl: 2.86).

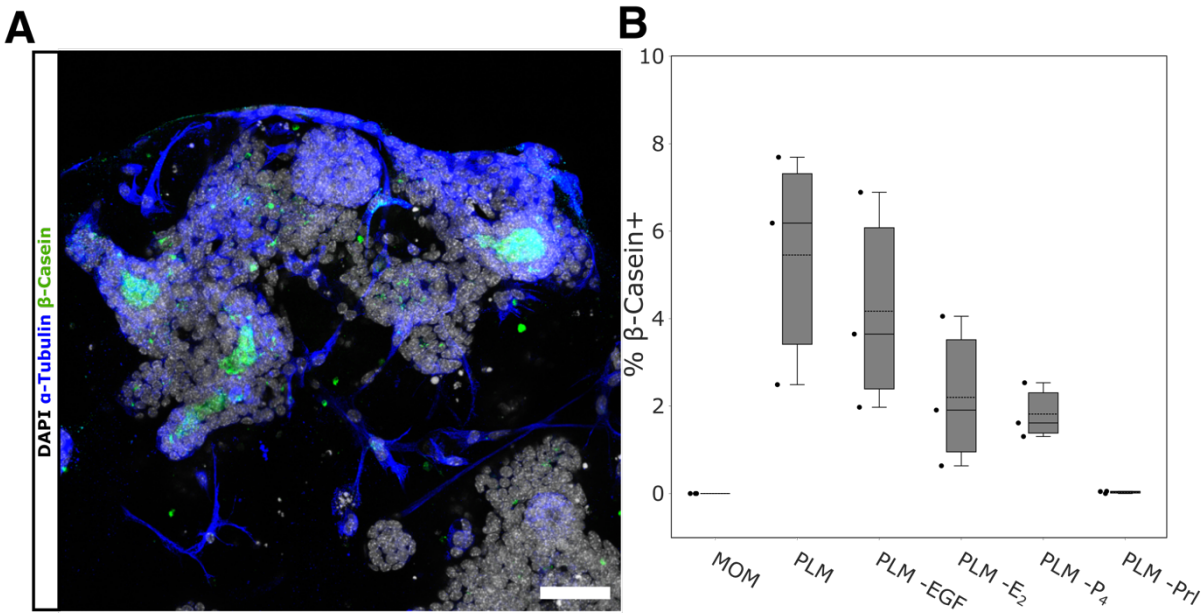


Figure 2.25 | Pregnancy hormones induce lactation in mMDOs.

(A) IF image of lactating mMDO stained against α-Tubulin (blue) and β-Casein (green). Scale bar 50 µm. **(B)** Quantification of β-Casein-positive cells in mMDOs cultured with MOM for 12 days or MOM for 6 days followed by different combinations of hormones in PLM for 6 days. *P*-values (not shown) between culture conditions were calculated by one-way ANOVA followed by Tukey's post hoc test (n = 3 mice) but did not yield values below 0.05 (MOM and PLM -Prl were not considered as no β-Casein-positive cells were detected).

By using geminin staining as a proxy for cycling cells in the mMDO culture, I further aimed to delineate the effect of different pregnancy hormones on the proliferative capacities (Wohlschlegel et al., 2002). Additional staining for β-casein and subsequent analysis of β-casein–positive and –negative cells revealed a markedly higher proportion of proliferating cells among the β-casein–positive fraction (Figure 2.26 A). Since lactating cells under MOM and PLM without prolactin conditions did not develop, no statistical test was performed for these conditions. In β-casein–negative cells, the proportion of cycling cells remained low across all treatments and ranged from 2.69 % to 5.89 %. In contrast, β-casein–positive cells exhibited substantially higher proliferation rates, ranging from 22.30 % to 36.62 %. The highest percentage of proliferating cells was observed in organoids treated with the complete pregnancy-lactation medium. These results indicate a synergistic effect of hormones and factors associated with growth of the mammary gland (estrogen, progesterone and EGF) and prolactin – potentially by a co-expression of the corresponding receptors or in a paracrine manner.

By extending the published protocol for lactating mammary gland organoids with additional pregnancy hormones, Veronika Lummer observed the occurrence of binucleated polyploid cells, which was recorded by us using live-cell imaging of H2B-mCherry/mG mMDOs (Figure 2.26 B). Thus, I investigated the amount of polyploid cells under different culture conditions. mMDOs that

were not treated with pregnancy-associated hormones showed a low level of polyploid cells (5.12 %; Figure 2.26 C). Treatment with pregnancy related hormones led to a notable increase in the proportion of polyploid cells across most conditions, with polyploidy reaching up to 15.41 %. Although the full pregnancy medium (PLM) did not result in a statistically significant increase in polyploid cells compared to the MOM control ($p = 0.088$), the effect size (Cohen's $d = 2.269$) indicated that a larger sample size might confirm this effect. Significant increases in polyploid cell populations were observed in conditions lacking EGF (PLM -EGF) and progesterone (PLM -P_4). Notably, the hormone combination including prolactin, EGF, and progesterone – but lacking estrogen – produced only a minimal increase in polyploid cells (6.06 %). Moreover, both PLM -EGF and PM -P_4 showed significantly higher levels of polyploidy compared to the estrogen-deficient condition. This result indicates a potential role of estrogen in promoting polyploidization of the lactating mammary gland. The polyploidization effect of estrogen can be further seen by specifically addressing β-casein-positive cells (Figure 2.26 D). β-casein-negative and β-casein-positive cells in the culture condition lacking estrogen did not show a significant difference regarding ploidy levels (6.17 % vs 1.85 %). However, all estrogen-containing conditions showed significantly elevated polyploidy levels in the β-casein-positive fraction of mMDO cells with the complete PLM exhibiting highest levels of 40.15 %. These results support the polyploidization effect of estrogen, specifically acting on β-casein-positive cells. Again, such effect could be mediated by a co-expression of the prolactin and the estrogen receptors as it has been reported for breast cancer cells (Murphy et al., 1984). β-casein-positive cells from complete PLM additionally showed a significantly increased polyploid cell number compared to the medium conditions lacking EGF. This result is in line with findings from Rios and colleagues who identified EGF as a critical factor of mammary gland polyploidization upon pregnancy (Rios et al., 2016). They additionally identified Aurora kinase A as an important mediator of cytokinesis "failure" or omission and binucleation. Given that estrogen has been shown to induce Aurora kinase A expression through GATA-3-activation in breast cancer cells, it can be postulated that increasing estrogen levels during pregnancy not only promote expansion of the mammary luminal cells but also contribute to their polyploidization (S. Jiang et al., 2010).

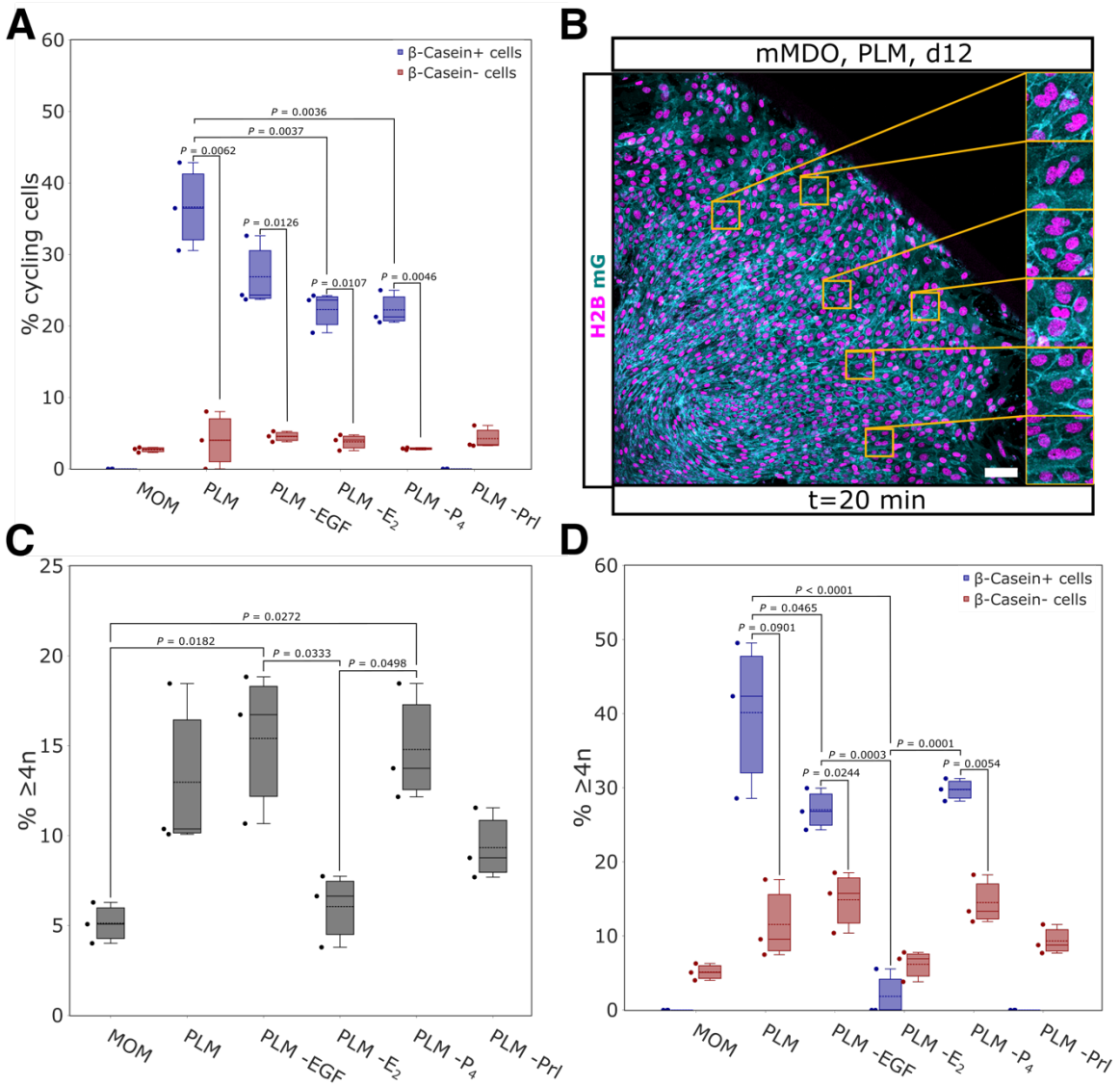


Figure 2.26 | Pregnancy hormones induce polyploidization in mMDOs.

(A) Comparison of the fraction of cycling (geminin-positive) cells among β-Casein-positive and β-Casein-negative cells for different hormone treatments. *P*-values between β-Casein-positive and β-Casein-negative cells within each culture condition were calculated using a two-sided paired Student's t test. *P*-values between culture conditions were calculated by one-way ANOVA followed by Tukey's post hoc test (n = 3 mice). **(B)** Image from live-cell microscopy of H2B-mCherry/mG mMDOs after adding lactation medium. Yellow boxes highlight regions showing binucleated cells. H2B (magenta), mG (cyan). Scale bar 50 µm. Image recorded with assistance from Veronika Lummer. **(C)** Ploidy measurements of mMDOs cultured with MOM for 12 days or MOM for 6 days followed by different combinations of hormones in PLM for 6 days. *P*-values between culture conditions were calculated by one-way ANOVA followed by Tukey's post hoc test (n = 3 mice). **(D)** Comparison of the fraction of polyploid cells among β-Casein-positive and β-Casein-negative cells for different hormone treatments. *P*-values between β-Casein-positive and β-Casein-negative cells within each culture condition were calculated using a two-sided paired Student's t test. *P*-values between culture conditions were calculated by one-way ANOVA followed by Tukey's post hoc test (n = 3 mice).

2.6.2 Polyploid luminal cells in lactating mammary gland organoids acquire mitotic defects and micronuclei

The original protocol of the lactating mammary gland organoids which served as a template for polyploid lactating mMDOs, demonstrated that hormone withdrawal by switching back to MOM after the lactation transformation, led to another remodeling mimicking involution. Thus, to address the behavior of binucleated/polyploid cells under such conditions Veronika Lummer and I performed live-cell imaging of H2B-mCherry/mG mMDOs in a joint effort. Similar to the pancreatic acinar cells upon ADM in mADOs, binucleated mMDO cells were capable of proliferating, eventually undergoing multipolar mitoses (Figure 2.27 A). Notably, quantification of the amount of cycling cells among the β-casein-positive fraction revealed a significantly higher amount of polyploid cells that were actively cycling (Figure 2.27 B). On top, IF staining for γH2A.X and α-Tubulin revealed binucleated/polyploid cells with micronuclei and pronounced DNA damage response (Figure 2.27 C). Image-based quantification of the γH2A.X-high fraction of mMDO cells demonstrates a significantly higher proportion of polyploid cells with active DNA damage response (Figure 2.27 D). Finally, live-cell imaging of H2B-mCherry/mG mMDOs revealed that cells exhibited signs of CIN by undergoing chromatin segregation errors and micronuclei accumulation and eventually remained proliferative, as described for the pancreatic acinar cells (Figure 2.27 E).

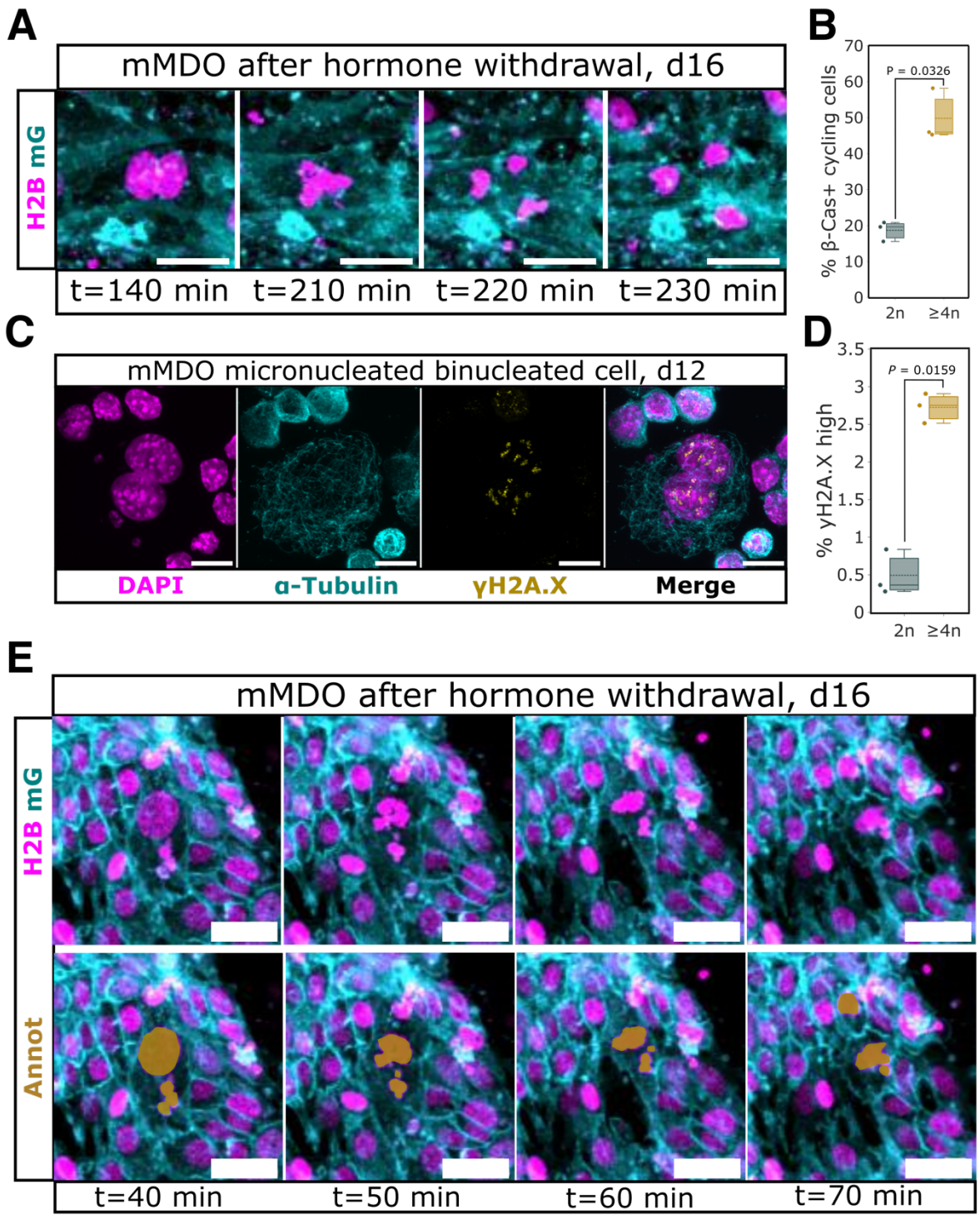


Figure 2.27 | Polyploid cells in lactating mMDOs exhibit sings of chromosomal instability.

(A) Live-cell image series of an H2B-mCherry/mG mMDO after hormone withdrawal showing a dividing binucleated cell undergoing multipolar mitosis: H2B (magenta), mG (cyan). Scale bar 20 µm. **(B)** Quantification of β-Casein-positive cycling cells among the diploid and polyploid fraction in d12 mMDOs (6 days MOM + 6 days PLM). *P*-value was calculated using a two-sided paired Student's t test (n = 3 mice). **(C)** Immunofluorescence images of γH2A.X high binucleated mMDO cell with micronuclei. DAPI (white), α-Tubulin (cyan), γH2A.X (yellow). Scale bar 10 µm. **(D)** γH2A.X measurement of diploid and polyploid mMDO cells at d12 (6 days MOM + 6 days PLM). *P*-value was calculated using a

two-sided paired Student's t test (n = 3 mice). **(E)** Live-cell image series of an H2B-mCherry/mG mMDO after hormone withdrawal showing a micronucleated cell undergoing mitosis. Upper row: H2B (magenta), mG (cyan). Bottom row: yellow annotation indicated nuclei of interest. Scale bar 20 μm.

3 Discussion

3.1 Polyploidy in regenerating glands – a deal with the devil?

My findings reveal a striking plasticity of binucleated polyploid acinar cells in response to injury, offering new insights into pancreatic regeneration and the early stages of tumor initiation. Although these cells remain highly quiescent under homeostatic conditions, they gain proliferative capacities following injury by undergoing acinar-to-ductal metaplasia (ADM) – a pivotal event in both tissue repair and tumorigenesis. Upon ADM, binucleated acinar cells re-enter the full cell cycle and frequently divide, often producing two, and occasionally three, mononucleated daughter cells. This results in a reduction of ploidy and a depletion of the binucleated cell population. The generation of more than two daughter cells in a single division event is driven by the presence of supernumerary centrosomes, which increases the likelihood of forming more than two spindle poles during mitosis. My data show that such multipolar divisions are more prevalent in binucleated than in polyploid mononucleated acinar cells, likely due to unfavorable centrosomal and nuclear positioning that arises after metaplasia-induced cell shrinkage. Moreover, these atypical cell divisions are highly susceptible to mitotic segregation errors, often resulting in lagging chromosomes, micronuclei formation, and extensive DNA damage – hallmarks of chromosomal instability and key drivers of tumor evolution, as exemplified by chromothripsis.

Cell culture experiments previously showed that polyploid cells rapidly eliminate extra centrosomes, likely as an adaptive strategy to mitigate the risk of chromosomal segregation errors (Baudoin et al., 2020; Bloomfield & Cimini, 2023). This process involves asymmetric centrosome clustering, where centrosomes are unequally distributed – for example, one centrosome on one pole and three on the other in a tetraploid cell – ultimately giving rise to polyploid cells with a single centrosome. These single-centrosome polyploid cells may possess a proliferative advantage over their multi-centrosome counterparts due to improved fidelity in chromosome segregation. As a result, cells with multiple centrosomes are selectively lost over time. Interestingly, similar mechanisms have evolved in polyploid organisms such as the allotetraploid *Xenopus laevis*, which employ strategies to maintain accurate chromatin segregation despite increased genomic content (Kubiak Jacek Z.and Prigent, 2012). However, this adaptive mechanism appears to be absent during pancreatic development, polyploidization, and homeostasis, as binucleated acinar cells continue to retain supernumerary centrosomes. This raises the question of whether the absence of mechanisms to reduce excess centrosome

numbers in polyploid acinar cells reflects a lack of selective pressure, given that these cells typically only divide under extreme conditions. Even if adverse outcomes such as cancer were to emerge years later (or within months in mice), the organism might already be beyond reproductive age and thus excluded from the inheritable gene pool. Alternatively, the persistence of supernumerary centrosomes could represent an adaptive feature that enhances regenerative capacity – enabling a rapid proliferative response by producing more than two daughter cells per division and introducing genetic variability through potential chromatin segregation errors. Such variability could confer regenerative advantages, as suggested by the "ploidy conveyor" model proposed for polyploid hepatocytes, which involves dynamic cycles of de- and repolyploidization (Duncan et al., 2010).

Interestingly, larger, longer-lived animals tend to have fewer binucleated polyploid pancreatic acinar cells, with an inverse correlation observed between cell size and both organismal size and lifespan (Anzi et al., 2018; Wollny et al., 2016). In addition, smaller animals like mice exhibit a much higher prevalence of polyploid hepatocytes and luminal mammary gland cells during lactation – likely reflecting an adaptation to meet greater metabolic demands and to enhance regenerative capacity in response to frequent injuries or stress (Donne et al., 2020; Rios et al., 2016).

However, while polyploidy is often linked to tumorigenesis due to its association with chromosomal instability, it has also been proposed as a tumor-suppressive mechanism, especially in the liver, by offering redundancy in tumor suppressor genes like TP53, thereby increasing resistance to oncogenic mutations (discussed in more detail in the following section, (Y.-H. Lin et al., 2020; S. Zhang et al., 2018)).

This leads me to a key question: Why have smaller, short-lived animals evolved to widely utilize such a seemingly advantageous strategy: boosting metabolism, enhancing regeneration, and potentially suppressing tumors through a relatively simple process like endoreplication or endomitosis – while larger, long-lived animals like humans, which are also capable of polyploidy, employ it far more sparingly?

This question ties into Peto's Paradox – the observation that larger animals with vastly more cells and longer lifespans, such as elephants and whales, do not exhibit proportionally higher cancer rates (Tollis et al., 2017). Proposed explanations include increased tumor suppressor gene copies (e.g., elephants have 20 retrocopies of TP53 Callier, 2019), enhanced immune surveillance, and reduced rates of cell division.

In mammals, the capacity to regenerate often represents a trade-off with tumor suppression (Maggiore & Zhu, 2024; Zhuang et al., 2025). While enhanced regeneration relies on stem cell activation, proliferation, and resistance to apoptosis, these same features can increase cancer risk. Of note, amphibians like axolotls and salamanders, which display both high regenerative ability and low tumor incidence, are an exception – and thus a subject of ongoing research (Roy & Gatien, 2008; Vieira et al., 2019).

Based on my findings, I propose that polyploidy represents a similar trade-off. At steady state, polyploid secretory cells may offer tumor-suppressive advantages. However, upon injury and activation, their division can lead to chromosomal mis-segregation, ploidy reduction, and genomic instability – factors that can promote tumorigenesis. In mice, this trade-off skews toward regeneration: their short lifespan (~2 years) means that early reproductive success outweighs the risks of late-onset cancers. Rapid tissue repair is evolutionarily more valuable than long-term tumor suppression.

In contrast, humans require durable organ function well before sexual maturity. For us, maintaining long-term genomic integrity is critical. As a result, we might limit glandular polyploidy and instead rely more heavily on diploid regeneration. Of course, polyploidy would represent only one facet of tissue-level cancer resistance. While it doesn't fully explain species-level differences in cancer incidence, particularly in the context of Peto's Paradox, it may constitute one important layer within a broader, multi-tiered system of tumor suppression in long-lived organisms.

3.2 Scheduled polyploidy as a potential origin of genomic instability

As described in the previous section, the polyploid state of a cell has been perceived controversially in relation to cancer. While many studies show detrimental effects of extra centrosomes, leading to chromatin segregation defects and DNA damage combined with higher resilience to undergo apoptosis, other studies propose a buffering effect of the extra copies of tumor suppressor genes, facilitating a certain degree of protection against loss of heterozygosity which has been primarily proposed for polyploid hepatocytes (H. Lin et al., 2021; Y.-H. Lin et al., 2020; Matsumoto et al., 2021; Storchova & Pellman, 2004; S. Zhang et al., 2018). However, studies about tumor initiation often employ genetically engineered animal models or chemical treatments that would affect almost every cell in the corresponding tissues. For instance, diethylnitrosamine (DEN) mediated tumor induction leads to DNA alkylation adducts and tumorigenesis in hepatocytes (Montesano, 1981). This results in the formation of dozens of independent tumor nodules throughout the liver. Considering that DEN is a mutagen directly

causing DNA damage, a buffering effect in polyploid cells after high dosage treatment seems reasonable. However, “real-world” tumorigenesis is most often not induced by single high dosages of a mutagen affecting almost every cell in the organ as it is commonly performed in the lab but rather the results of multiple regeneration rounds following multiple or chronic injuries over many months and years, counting ageing as a mutagen itself. No matter if a gradual or the more recently proposed punctuated equilibrium model is consulted to explain tumor evolution, when the question of the origin of a tumor is raised, the cells that show the greatest tendency to DNA damage and genetic instability among the known risk factors should be considered (Graham & Sottoriva, 2017). Notta et al. proposed for the pancreas, that the traditional view of mutation acquisition in a gradual manner does not match clinical observations of pancreatic cancer genomes *per se* (Notta et al., 2016). Thus, they suggested that a significant amount of PDACs arise through chromatin segregation error-related catastrophic events, like chromothripsis, that induce multiple mutations simultaneously. My data fits these observations very well, providing the binucleated polyploid state of an acinar cell as a potential origin for this punctuated equilibrium. Of note, the punctuated model of tumor evolution most probably exists side-by-side to the gradual one and represents an alternative route to PDAC formation. Accordingly, the cell of origin for PDAC might vary from case to case. Besides acinar cells, the ductal compartment of the exocrine pancreatic epithelium has been proposed as the origin of PDAC and depending on the cell type of origin, the mode of mutation acquisition might vary as well (Ferreira et al., 2017).

3.3 How do polyploid secretory cells cope with chromosomal instability?

As my data demonstrates, neither micronucleated primary acinar cells in mADOs nor micronucleated primary luminal mammary gland cells in mMDOs are inevitably silenced or killed by internal cues. Instead, a subset retains proliferative capacities and further accumulates chromosomal instability. So how do these cells manage to deal with CIN?

Members of the E2F family of transcription factors have been implicated in regulating hepatocyte polyploidization by inducing cell cycle arrest in cells harboring extra centrosomes (Sladky et al., 2020). Since E2F transcription factors were also proposed as regulators for acinar cell ploidy, it is likely that the pancreas employs a similar mechanism to control for ploidy levels (Matondo et al., 2018). This supports the notion of the binucleated state as a growth endpoint.

In polyploid hepatocytes, extra centrosomes lead to PIDDosome-mediated Caspase 2 activation leading to p53 dependent cell cycle arrest thus facilitating the prevention of further polyploidization (Burigotto et al., 2021; Rizzotto et al., 2024). This system works by the accumulation of p53-induced death domain (PIDD) accumulation via ankyrin repeat domain-containing protein 26 (ANKRD26), a protein associated to centrosomes. PIDD accumulation further leads to PIDDosome-assembly, Caspase 2 activation and finally p53 stabilization through murine double minute 2 (MDM2) cleavage (Yao et al., 2024). Thus, the more centrosomes a cell contains, the stronger the p53 response will be, hence limiting further growth and polyploidization. Interestingly, the same signaling axis has been shown to lead to p53-induced apoptosis in hematopoietic cell lines (Nalm6, BaF3) cells upon unscheduled polyploidization (Rizzotto et al., 2024). As scheduled polyploidization in epithelial glandular cells but also cardiomyocytes does not lead to apoptosis but to cell cycle arrest, it is likely that these cells exhibit a higher tolerance to p53-associated activity. This effect could also lead to a generally increased resilience towards chromosomal instability and DNA damage – eventually even as an adaptive mechanism to aid regeneration as discussed in the previous section. Since the PIDDosome axis is active in polyploid hepatocytes to limit extensive polyploidization of levels higher than 8n, a factor that links p53 to this pathway and the regular DNA damage response pathway might be differentially regulated in polyploid secretory cells. Regular DNA damage response functions by the detection of damaged sites by kinases such as ATM (double-strand breaks) or ATR (single-strand breaks) which phosphorylate H2A.X histones to provide a platform for DNA damage repair (Abuetabh et al., 2022). In parallel, ATM/ATR start to phosphorylate p53 directly or indirectly through checkpoint kinase phosphorylation to prevent p53 from degradation by MDM2. Stabilized p53 induces the expression of cell cycle repressors such as p21 (cyclin-dependent kinase 2 inhibitor) to give the cell enough time to repair the damage. If DNA damage repair takes too long, it additionally induces mediators of apoptosis such as p53 upregulated modulator of apoptosis (PUMA) and NOXA. However, ATM (responsible for double strand break detection), can also phosphorylate MDM2, thereby inhibiting its activity directly (Maya et al., 2001). Since MDM2 links both centrosome/PIDDosome-mediated and DNA damage-mediated cell cycle arrest/apoptosis, it might be worth checking its expression and regulation in polyploid exocrine cells.

Another potentially involved pathway is the cGAS/STING pathway for cytosolic DNA sensing and subsequent interferon type I response. Commonly understood as a pathway to detect pathogenic DNA to induce inflammation and cell death, repress cell cycle and activate innate immunity, studies showed that cGAS can also bind to self-DNA when presented as micronuclei

with ruptured envelopes (Mackenzie et al., 2017). Other studies demonstrated that the binding of cGAS to self-DNA does not lead to forwarding of the signal to induce an interferon response because of the interaction with an acidic patch in the histone H2A-H2B hetero-dimer that inhibits its synthase activity (Flynn et al., 2021; Pathare et al., 2020; Takaki et al., 2024). Accordingly, mitotic nuclear envelope breakdown does not lead to cGas/STING activation. Yet, the activation of the cGAS/STING pathway by small, nucleosome-lacking fragments of DNA, such as small extrachromosomal circular DNAs (eccDNA), which can emerge from chromothripsis, is still under debate. Interestingly, eccDNA has additionally been shown to activate STING in an cGAS-independent manner (Wu et al., 2023, 2024). Hence a closer look on the cGAS/STING signaling activity in polyploid secretory cells could further shed light on the intricate capacities of these cells to withstand chromosomal instability.

3.4 Polyploid luminal cells of the mammary gland and postpartum breast cancer

Postpartum breast cancer (PPBC) is characterized by an increased risk of breast cancer following childbirth, with incidence peaking around six years postpartum and remaining elevated particularly among women who give birth after the age of 30 (Albrektsen et al., 2005). While polyploid cells in the mammary gland, critical for milk production during pregnancy and lactation, have not yet been directly implicated in PPBC, my findings suggest they may play a contributory role. Specifically, I observed that polyploid mammary cells retain proliferative potential, show signs of mitotic errors, and exhibit elevated DNA damage following exposure to pregnancy-associated hormones. Interestingly, polyploid mammary gland cells showed increased proliferative capacities compared to their diploid counterparts. This result stands in contrast to my data on binucleated cells from the mouse pancreas, where polyploidy evoked slightly limited proliferative capacities. These data were collected *in vivo* (CiP) though, which makes comparability difficult. Yet, a potential explanation might be given by a differential activation of polyploid cell proliferation levels depending on the strength of the underlying causative effect (injury/ADM or involution). Such effects were reported for the liver, where a relatively mild injury (30 % partial hepatectomy) did not lead to an activation of polyploid hepatocytes and hypertrophy as principal driver of regeneration whereas a severe injury (70 % partial hepatectomy) activated binucleated hepatocytes to contribute to regeneration (Miyaoka et al., 2012).

Notably, estrogen-driven activation of Aurora kinase A, which can induce polyploidization in mammary-derived organoids (mMDOs), has also been implicated in breast cancer pathogenesis (S. Jiang et al., 2010). These observations support the hypothesis that polyploidization may contribute to the development of PPBC. This is further reinforced by recent evidence showing that DNA damage during lactation can itself induce polyploidy in mammary gland cells (Molinuevo et al., 2024).

The post-lactation involution process creates a tumor-promoting microenvironment through extensive tissue remodeling, epithelial cell apoptosis, immune cell infiltration, and extracellular matrix remodeling - conditions that facilitate tumor growth and metastasis (P. Schedin, 2017). The inflammation and immune suppression associated with involution resemble wound healing and may allow emerging malignant cells to evade immune detection (Clarkson et al., 2000, 2003; Hennigar et al., 2015). Polyploid cells are particularly well-suited to survive and adapt within such pro-tumorigenic environments, suggesting that polyploidization could enhance susceptibility to PPBC by reinforcing these effects(X. Li et al., 2024).

3.5 Polyploid cells as potential origin for human cancers

Polyploid cells, while generally genetically buffered under homeostasis, undergo ploidy reduction and multipolar divisions during regeneration, which are highly error prone. These events often result in lagging chromosomes, micronuclei, and chromothripsis, mechanisms known to drive genomic instability and oncogenesis. Despite this, most regenerative events do not lead to cancer, suggesting the existence of intrinsic control systems capable of eliminating genomically aberrant cells. A strong candidate for such a system is the immune surveillance network, particularly natural killer (NK) cells, which are adept at detecting and eliminating aneuploid or chromosomally unstable cells (Chan et al., 2014).

This balance, however, appears fragile. Many chronic disease states, including hepatitis, fatty liver disease, and chronic pancreatitis, are characterized by immune compromising or immune exhaustion, either by medication or the disease itself (Bosch et al., 2024; Ebrahimi et al., 2023; Glaubitz et al., 2023). Additionally, chronic CIN itself, induced by polo-like kinase 1 overexpression has been shown to facilitate immune evasion by NK cell suppression in a breast cancer model (Kandala et al., 2023). Interestingly, polo-like kinase 1 is also a substrate for Aurora kinase A, and both have been shown to induce polyploidy in the mammary gland during lactation (Rios et al., 2016).

When immune oversight is diminished, polyploid cells undergoing aberrant mitosis may accumulate mutations unchecked, increasing the likelihood of malignant transformation. This

model aligns well with epidemiological data: chronic pancreatitis is one of the strongest risk factors for pancreatic ductal adenocarcinoma, while chronic liver inflammation significantly increases the risk for hepatocellular carcinoma (Campbell et al., 2021; Gandhi et al., 2022). Similarly, proliferative phases of the mammary gland during pregnancy and lactation involve polyploid alveolar cells, and when paired with inflammatory conditions or hormonal dysregulation, may contribute to breast cancer development.

By contrast, tissues like the skin, which regenerate primarily through diploid progenitor cells and exhibit relatively low polyploidy, are less prone to inflammation-driven cancers. Chronic inflammatory conditions such as atopic dermatitis or psoriasis, though intense, rarely predispose to skin cancer unless coupled with external mutagens such as UV radiation (Crisafulli et al., 2021; Wan et al., 2023). This suggests that polyploidy creates a unique vulnerability: the tendency to accumulate chromosome segregation errors during regeneration that, under immune surveillance failure, can translate into clonal expansion of precancerous cells.

In essence, polyploidy in regenerating secretory tissues may represent a biological "deal with the devil" – an evolutionary compromise where enhanced regenerative plasticity and metabolic efficiency are tolerated at the cost of potential genome instability, buffered in healthy states by robust immune clearance. When this buffer is weakened by chronic injury or inflammation, the risk of transformation sharply rises. This model provides a unifying framework for understanding why certain polyploidy-enriched organs are hotspots for inflammation-associated cancers, while others remain comparatively resistant.

4 Conclusion

Over a century ago, Theodor Boveri proposed that errors in mitotic chromosome segregation and multipolar mitoses could serve as the foundation for cancer development (Boveri, 1902, 1914). However, he also argued that for such mechanisms to be causative, they would need to be observed not only in malignant but also in premalignant tissues. Despite this early insight, the occurrence and relevance of chromosomal instability (CIN) in normal tissues has remained largely unexplored. Instead, the field has often taken for granted that CIN is an inevitable feature of cancer progression, with many studies focusing on its dramatic consequences such as chromothripsis, micronuclei formation, and genomic chaos – without investigating the origin of CIN in physiologically normal or regenerative contexts.

My work addresses this long-standing gap by demonstrating that naturally occurring polyploid cells in exocrine glands, such as the pancreas and mammary gland, can serve as a physiological source of CIN. These cells, while normally quiescent, retain the capacity to re-enter the cell cycle in response to injury or hormonal stimuli. Upon doing so, their supernumerary centrosomes predispose them to multipolar mitoses, mitotic errors, and DNA damage, creating a cellular environment ripe for CIN and its associated phenomena including micronuclei formation, chromothripsis, and potentially tumorigenic clonal expansion. Notably, these cells are often resistant to apoptosis and capable of continued proliferation, providing a selective advantage for propagation of damaged genomes. Thus, my findings reveal a previously underappreciated source of chromosomal instability within normal regenerating tissues, offering a new lens through which to understand the earliest steps of tumor initiation – one rooted not only in mutation but in tissue architecture, ploidy, and regenerative demand.

Looking ahead, an important challenge will be to unravel how polyploid cells tolerate or respond to CIN, both through intrinsic mechanisms such as the PIDDosome complex, ATM/ATR-mediated DNA damage response, and cGAS-STING signaling and through extrinsic factors, including interactions with immune surveillance cells, like natural killer cells.

Indeed, a suppressed or immature immune system may further enable the expansion of these CIN-prone cells, contributing to tumor emergence. This could help explain epidemiological observations linking cancer risk to chronic inflammation, repetitive tissue injury, and even physiological processes like post-lactational involution, where transient immunosuppression and tissue remodeling might create a permissive environment for transformation.

From a broader perspective, understanding how CIN arises in non-malignant tissues opens new avenues for preventive medicine – an area of increasing importance as traditional medicine, focused on treating disease after onset, approaches its therapeutic limits. By identifying the

cellular origins and molecular pathways that permit or suppress CIN under normal conditions, we may uncover novel targets for early intervention, long before pathological transformation occurs.

In sum, this work revisits Boveri's foundational insights through the lens of modern regenerative biology, showing that the roots of chromosomal instability – and thus cancer – may lie not only in mutation, but in the very cells designed to repair and sustain our tissues. By recognizing the dual nature of polyploid cells as both, regenerative agents and potential seeds of instability, we move closer to a future where cancer prevention is proactive, targeted, and grounded in the biology of normal tissue homeostasis.

5 Materials and methods

5.1 Materials

5.1.1 Animals

Table 3 | Mouse strains

Strain	Full name	Function/Phenotype	Source
C57BL/6N	-	Wildtype	DKFZ Heidelberg, Central animal laboratory
H2B-mCherry	B6-Gt(ROSA)26Sortm3Sia	Constitutive histone 2B (H2B)-mCherry	Prof. Dr. Jan Ellenberg, EMBL Heidelberg
mG	B6-Gt(ROSA)26Sortm4.1(ACTB-EGFP)Luo/Amv)	Constitutive membrane-localized GFP expression	Prof. Dr. Takashi Hiragi, EMBL Heidelberg
H2B-mCherry/mG	B6-Gt(ROSA)26Sortm3.1Sia Gt(ROSA)26Sortm4.1(ACTB-EGFP)Luo/Amv	Constitutive H2B-mCherry and membrane-localized GFP expression	Crossed from H2B-mCherry and mG mice at DKFZ Heidelberg, Central animal laboratory
EGFP-Tuba	B6-Gt(ROSA)26Sortm12.1Sia	Constitutively EGFP-α-Tubulin expression	Prof. Dr. Jan Ellenberg, EMBL Heidelberg
Fucci2	B6-Tg(Gt(ROSA)26Sor-Fucci2)#Sia	mCherry-hCdt1 expression in G0/G1 and mVenus-hGeminin expression in S/G2/M phase	Dr. Michael Milsom, DKFZ Heidelberg

5.1.2 Chemicals & reagents

Table 4 | Chemicals, reagents and Kits

Reagent	Supplier
B27 Serum-Free Supplement	Gibco
Boric acid (H_3BO_3)	Sigma Aldrich
Bovine Serum Albumin (BSA)	Sigma Aldrich
BrdU (5-Bromo-2'-deoxyuridine)	Thermo Fisher
Calcium chloride ($CaCl_2$) 1M	Thermo Fisher
DAPI (4',6-diamidino-2-phenylindole)	Sigma Aldrich
Dulbecco's Modified Eagle Medium/Nutrient Mixture F-12 (DMEM/F12)	Gibco
Dulbecco's Modified Eagle Medium/Nutrient Mixture F-12 (DMEM/F12) + Gluta- MAX™	Gibco
Dulbecco's Modified Eagle Medium/Nutrient Mixture F-12 (DMEM/F12) Gibco with: L-glutamine without: HEPES, phenol red	Gibco
Dulbecco's Phosphate-Buffered Saline (DPBS)	Gibco
Ethylenediaminetetraacetic acid (EDTA)	Merck
Ethanol, absolute	Sigma Aldrich
Fetal Bovine Serum (FBS)	Biological Industries
Fluoromount G	eBioscience
Gentamicin Reagent Solution	Gibco
Goat Serum	Merck
Hank's Balanced Salt Solution (HBSS)	Gibco
Hydrogen chloride (HCl)	Sigma Aldrich
Hoechst33342	Invitrogen
Horse Serum	Biochrom
Insulin-Transferrin-Selenium (ITS-X)	Gibco

Ketavet	Pfizer
Matrigel®, Growth Factor Reduced	Corning
Magnesium chloride ($MgCl_2$) 1M	Thermo Fisher
N-2 Supplement	Gibco
NextSeq 500/550 Mid Output Kit v2.5 (150 cycles)	Illumina
NP-40 Surfactant-Amps™	Thermo Fisher
Paraformaldehyde (Roti®-Histofix 4%)	Carl Roth
Penicillin-Streptomycin	Gibco
ProFreeze™-CDM	Lonza
Propidium iodide, 1 mg/ml	Thermo Fisher
RBC Lysis Buffer	Invitrogen
Rompun	Bayer
Sodium Azide (NaN_3)	Sigma Aldrich
Sodium chloride (NaCl) 0.9 %	Sigma Aldrich
Sodium chloride (NaCl) 5M RNase-free	Thermo Fisher
Sodium hydroxide (NaOH)	Sigma Aldrich
Sodium tetraborate ($Na_2B_4O_7$)	Sigma Aldrich
Sucrose	Sigma Aldrich
Tissue-Tek® O.C.T.™ Compound	Sakura
Triton™-X 100	Sigma Aldrich
TrypLE ™ Express	Gibco
UltraPure™ 1M Tris-HCl-Buffer, pH 7.5	Invitrogen

Table 5 | Recombinant proteins and hormones

Protein/Hormone	Supplier
Collagenase A	Sigma-Aldrich
Collagenase Type IV	Gibco
DNAse I	Sigma Aldrich
Dispase® II	Sigma Aldrich
rhEGF	Active Biosciences
β-Estradiol	Sigma-Aldrich
bFGF2	Pelobiotech
rhFGF2	Miltenyi Biotec
Progesterone	Sigma Aldrich
Murine Prolactin	Peprotech

Table 6 | Primary Antibodies

Antibody	Host	Dilution	Supplier	Catalog number
Anti-α-Amylase	Rabbit	1:200	Sigma-Aldrich	A8273
Anti-α-Tubulin	Rat	1:500	abcam	ab6160
Anti-β-Casein	Mouse	1:200	Santa Cruz	sc-166530
Anti-BrdU	Rat	1:100	abcam	ab6362
Anti-Cytokeratin 19	Rat	1:50	Hybridoma Bank	Troma III Conc.
Anti-DCLK1	Rabbit	1:200	Abcepta	AP7219B
Anti-E-Cadherin	Rat	1:1000	Thermo Fisher	13-1900
Anti-γ-H2A.X	Mouse	1:200	Novus Biologicals	05-636
Anti-γ-Tubulin	Rabbit	1:200	abcam	ab11317
Anti-Geminin	Rabbit	1:500	abcam	ab195047

Anti-Keratin 5 (K5)	Rabbit	1:500	Biolegend	905501
Anti-Keratin 8 (K8)	Mouse	1:500	Biolegend	904805
Anti-Lamin A/C	Mouse	1:200	Cell Signaling	4777S
Anti-Pericentrin	Rabbit	1:100	abcam	ab4448
Anti-Phospho-Histone 3	Rabbit	1:250	Merck	06-570

Table 7: Secondary antibodies

Antibody	Conjugate	Dilution	Supplier	Catalog number
Anti-Mouse	Alexa Fluor® 488	1:200	Life Technologies	A-21202
Anti-Rabbit	Alexa Fluor® 568	1:200	Life Technologies	A-10042
Anti-Rat	Alexa Fluor®647	1:200	Life Technologies	A-21247

5.1.3 Solutions and media

Table 8 | Solutions and media

Solution/Medium	Components
Perfusion Solution	Ketavet, 5.71 mg/ml Rompun 2.80, mg/ml in 0.9 % NaCl
Pancreatic organoid culture (POC) medium	DMEM/F12 supplemented with: B27 Serum-Free Supplement (1x) N-2 Supplement (1x) 20 ng/ml rHu EGF 20 ng/ml rhFGF2 1% Penicillin-Streptomycin
Separation Buffer	0.5 % BSA 2 mM EDTA in PBS
Solution D	1 mg/ml Collagenase Type IV in POC1
Solution R	1 % BSA in PBS
Solution S	4 % BSA in PBS
Solution P	0.1 % BSA in PBS
Solution W	2 % PenStrep in PBS
Digestion Solution I	2mg/ml Collagenase A 0.25% Trypsin 5% FBS 10 µL/ml Gentamicin in Basal DMEM/F12
Digestion Solution II	5 mg/ml Dispase 2 0.5 mg/ml DNAse 1 in Basal DMEM/F12
Wash Medium 1	5% FBS in DMEM/F12 + GlutaMAX™

Wash Medium 2	0.1% FBS in PBS
Basal Mammary Media (BMM)	500 ml DMEM/F12 + GlutaMAX™ 5 ml 100X ITS-X 5 ml 100X Pen-Strep
Morphogenesis Organoid Medium (MOM)	1X BMM 45 ng/ml bFGF2
Pregnancy Organoid Medium (PLM)	1X BMM 25 ng/ml rhEGF 40 ng/ml b-Estradiol 120 ng/ml Progesterone 120 ng/ml murine Prolactin
Wash/Block Buffer (WBB)	150 mg BSA 1.5 ml Horse Serum 1.5 ml Goat Serum in 27 ml PBS/0.1%TritonX-100
Nuclei staining buffer (NSB)	1.0 ml 1 M Tris-HCl, pH 7.5 308 µl 5 M NaCl 10 µl 1 M $CaCl_2$ 5 µl 1M $MgCl_2$ 266.5 µl 7.5 % BSA 100 µl 10 % (vol/vol) NP40 10 µl 10 mg/ml Hoechst33342 100 µl 1 mg/ml PI
Freeze buffer	425 µl 2x ProFreeze-CDM 75 µl 100 % DMSO 500 µl 1x PBS
0.15 M sodium borate buffer, pH 8.5	3.1 g boric acid 4.75 g sodium tetraborate filled to 1000 ml with ddH_2O

5.1.4 Devices and equipment

Table 9 | Devices and equipment

Category	Type / Model	Supplier
Biological Safety Cabinet	NU-437-600E	Nuaire
Cell strainer	100µm	Corning
Centrifuge	Heraeus® Multifuge™ 3S-R	Thermo Fisher
Cytospin	Cytospin 3 Shandon	Shandon
Confocal Microscope	A1R	Nikon
Confocal Microscope	TCS SP8	Leica

Counting chamber	Neubauer Improved	Brand
Cryomicrotome	CM 1950	Leica
FACS Sorter	FACSAria™ Fusion	BD Biosciences
Flow Cytometer	LSR Fortessa™	BD Biosciences
Graphics processing unit	RTX4090	Nvidia Corporation
Incubator	Heracell™ 240i	Thermo Fisher
Lightsheet Microscope	InVi-SPIM	Luxendo
Magnetic Stirrer	Yellow MAG HS 7	YellowLine
Microscopy Slides	Superfrost® Plus	Thermo Fisher
Microscopy/Cell Culture slides	CELLview™	Greiner Bio-One
Microscopy/Cell Culture slides	Lab-Tek® Chamber SlideTM system	Thermo Fisher
Microscopy/Cell Culture multiwell plate	96 Well Sensoplate	Greiner Bio-One
Orbital Mixer	Intelli Mixer	NeoLab
PVDF filter	0.2 µm	Cytvia
Scale	Secura	Sartorius
Sequencer	NextSeq 550	Illumina

5.2 Methods

5.2.1 Animal models

All animals used in this project were housed under specific pathogen-free conditions at a 12 h light dark circle at 22 °C and fed *ad libitum*. All procedure and experiments were in accordance with the German Cancer Research Center (DKFZ) guidelines and approved by the Regierungspräsidium Karlsruhe. For this study, C57BL/6N mice are referred to as wild type (WT). The following transgenic mouse lines were used:

- **H2B-mCherry (B6-Gt(ROSA)26Sortm3Sia):** This mouse line constitutively expresses histone 2B (H2B) fused to mCherry, which endogenously labels cell nuclei.
- **mG (B6-Gt(ROSA)26Sortm4.1(ACTB-EGFP)Luo/Amv):** The mG construct specifically labels plasma membranes by utilizing a modified part of the membrane-bound domain of the MARCKS protein fused to GFP.
- **EGFP-Tuba (B6-Gt(ROSA)26Sortm12.1Sia):** These mice constitutively express EGFP fused to α-Tubulin, labeling microtubule fibers as part of the cytoskeleton as well as the mitotic spindle.
- **Fucci2 (B6-Tg(Gt(ROSA)26Sor-Fucci2)#Sia):** The Fucci2 mouse model expresses mCherry-hCdt1 in G1 and mVenus-hGeminin in S/G2/M phase.

H2B-mCherry and EGFP-Tuba mouse lines were kindly provided by Dr. Jan Ellenberg (EMBL Heidelberg) and the mG mouse line was kindly provided by Dr. Takashi Hiragi (former EMBL Heidelberg) and transferred to the DKFZ animal housings via embryo transfer. Fucci2 mice were a kind gift from Dr. Michael Milsom (DKFZ Heidelberg). Material transfer agreements with the original owners of the mouse lines were conducted.

R26-H2B-mCherry and mG mice were crossed to generate H2B-mCherry/mG double reporter mouse lines. All mice were aged 8-16 weeks when used in experiments.

5.2.2 Cerulein-induced pancreatitis

Acute pancreatitis in mice was induced using the oligopeptide cerulein (Sigma-Aldrich, C9026) in WT mice. Cerulein was dissolved in phosphate-buffered saline (PBS) to a concentration of 5 µg/ml and administered hourly (50 µg/kg body weight) by 7 intraperitoneal (i.p.) injections on two consecutive days (14 injections in total). The day of the last injection was defined as 0 dpi. Three replicates of cerulein-treated mice were sacrificed and employed for pancreas extraction at 2

dpi, 4 dpi, 28 dpi and 91 dpi, respectively. To control for injection-induced injuries, three WT mice were i.p.-injected with comparable volumes of 0.9 % NaCl and otherwise treated identically.

5.2.3 Mouse pancreas extraction

To obtain pancreatic tissue, mice were anesthetized by an i.p. injection of 700 µl – 800 µl Ketavet (5.71 mg/ml)/Rompun (2.80 mg/ml) in 0.9 % NaCl. Blood cells were cleared out by a left-ventricular perfusion using 20 ml Hank's balanced salt solution (HBSS). In case of tissue fixation for cryo-sectioning, perfusion was continued with another 20 ml of 4 % Roti-Histofix (Carl Roth). The pancreas was exposed by opening the abdominal cavity via a mid-abdominal vertical incision. Pancreatic tissue was extracted and attached adipose tissue was removed, if necessary. Until further usage, pancreatic tissue was kept in PBS on ice.

5.2.4 Mouse acinar-derived organoid culture

The following solutions were prepared and sterile-filtered: S (4 % bovine serum albumin (BSA) in PBS), R (1 % BSA in PBS), D (1 mg/ml Collagenase IV in 0.25 % BSA in PBS) and W (2 % Penicillin-Streptomycin in PBS). Pancreatic tissue was extracted and collected in 10 ml solution W on ice. The tissue was separated from remaining fat and rinsed once more in 10 ml solution W. The tissue was then minced into small pieces with a volume of approximately 1 mm^3. Tissue pieces were collected, rinsed in 10 ml solution W and then digested in solution D for 30 min at 37 °C and 5 % CO_2. Every 5 min, the suspension was pipetted up and down using serological pipettes (5 ml) to further facilitate dissociation. The digestion product was filtered through a 100 µm cell strainer to remove pancreatic islets and the cell strainer was rinsed with 10 ml Solution R. Four centrifugation tubes (15 ml) were prepared, each containing 6 ml of Solution C. For each tube, 5 ml of the filtered cell suspension were gently transferred on top to achieve a BSA gradient by layer separation. Acini were isolated by a single centrifugation step at 50 x g and 4 °C for 2 min. The supernatants were, removed and the pellets were resuspended and washed with Solutions C and I, successively (50 x g, 4 °C for 2 min). After the last washing step, the pellets were resuspended in 500 µl Pancreatic organoid culture (POC) medium (1:1 DMEM/F12 + GlutaMAX™ 2 % (v/v) B27 serum free supplement, 1 % (v/v) N2 supplement, 1 % penicillin-streptomycin, 2 mM L-glutamine 20 ng/ml rhEGF, 20 ng/ml rhFGF2) each and pooled. At this point, acinar cells build clusters of approximately 4-10 cells per cluster. Since cell-cell contacts facilitate organoid formation for this cell type, no further dissociation steps were performed. Acinar cells were counted and mixed with ice cold unpolymerized Matrigel® according to a density of 250 cells/µl. For 3D organoid culture, a 20 µl-droplet of this mixture was placed in each well of the respective cell culture

dishes and incubated for 20 min at 37 °C and 5 % CO_2 to facilitate Matrigel® polymerization, before POC medium was added. Acinar cells were cultured at 37 °C and 5 % CO_2 in either 24-well plastic plates for fixed-cell imaging experiments and Strand-seq, 10-well CELLview™ Slides (Greiner Bio-One) for confocal live-cell imaging or Luxendo TruLife dishes for light sheet live-cell imaging. The time point of cell seeding in Matrigel® was defined as day 0 (d0).

5.2.5 Mouse mammary gland extraction

Female animals were euthanized via cervical dislocation. Following sacrifice, the
mice were pinned in supine position and the ventral area was disinfected with 70% ethanol. A nick was cut above the pubis with small surgical scissors and continued towards the thoracic nipples. Then, two lateral incisions were made, and the skin was peeled to reveal the inguinal mammary fat pads. For thoracic pads, the same procedure was repeated in the thoracic area. To remove the fat pads, they were squeezed and pulled with the forceps while making small incisions to remove them from the skin. Lastly, collected mammary glands were washed in 1x cold PBS and transferred to a 50 mL tube containing 1x Dulbecco's Modified Eagle Medium/Nutrient Mixture F-12 (DMEM/F12 (1:1) + GlutaMAX™) at room temperature (RT).

5.2.6 Mouse mammary gland derived organoid culture

The protocol for mouse mammary gland derived organoids was adapted from
After extraction, glands were quickly rinsed in 1x DPBS and transferred to a 100 mm petri dish for mincing with 2 sterile disposable scalpels. Then, minced mammary tissue was transferred to a 50 mL falcon containing 10 mL of pre-warmed digestion medium I (2 mg/ml Collagenase A, 0.25 % trypsin, 5 % fetal bovine serum (FBS), 10 µg/ml gentamicin in 1:1 DMEM/F12 + GlutaMAX™) and incubated on an orbital shaker at 200 rpm at 37 °C for 40 min. After digestion, the tube was briefly vortexed and the content was transferred to a 15 mL falcon, which was centrifuged at 450 x g for 10 min. Resulting fat layer contains trapped mammary cells inside, so together with the supernatant it was taken to a new falcon and pipetted up-down to release the epithelial cells from the fat. The fat mixture was centrifuged again, and the pellet was combined with the main pellet to maximize the yield. The combined pellet was resuspended in 5 mL wash medium I (5 % FBS in 1:1 DMEM/F12 + GlutaMAX™) and inverted 5-6 times, followed by centrifugation for 5 min at 450 x g. The supernatant was discarded, and the pellet was resuspended with red blood cell lysis buffer (RBC) and incubated for 1 min at RT to get rid of red blood cells. After the incubation, 8 mL PBS/1%FBS was added into the tube, and it was centrifuged for 2 min at 200 x g. The supernatant was aspirated, and the pellet was resuspended with 4 mL pre-warmed digestion

medium II (5 mg/ml Dispase 2, 0.5 mg/ml DNAse I in 1:1 DMEM/F12 + GlutaMAX™). The cells were incubated in digestion medium II for 5 min in a 37 °C water bath. 6 mL wash medium II (0.1 % FBS in PBS) was added to stop the reaction. Then the tube was centrifuged at 450 x g for 5 min and the pellet was resuspended in 1 mL Morphogenesis Organoid Medium (MOM, 1:1 DMEM/F12 + GlutaMAX™ supplemented with 1 % (v/v) ITS-X, 2 %, 1 % penicillin-streptomycin, 45 ng/ml rhFGF2). Freshly isolated primary mammary epithelial cells were mixed with growth factor reduced Matrigel® and plated in domes in 8 µL Matrigel® pre-coated 24-well culture plates for fixed-cell experiments or 10-well CELLview™ Slides (Greiner Bio-One) for live-cell imaging (one dome per well, 20 µL of Matrigel® per dome, ~1500 organoids per dome) to increase the stability of the domes. After the Matrigel® polymerization in upside down position for 35 min at 37 °C, the 3D organoid cultures were covered with pre-warmed MOM and incubated at 37 °C in 5% CO2 incubator. Fresh media change was done once in every 2 days periodically, all media were prepared freshly every time. For the experiments, the organoids were cultured in MOM for 6 days, followed by either another 6 days in MOM, 6 days in Pregnancy-Lactation Medium (PLM) or 6 days in PLM lacking one of the factors E_2, P_4, Prl and EGF.

5.2.7 Fixation and single-cell dissociation for mADOs and mMDOs

At given time points, mADO and mMDO cultures were dissociated using TrypLE™ for 20-30 min with visual validation of dissociation under the microscope. After dissociation, single cells were washed once in PBS and fixed for 15 min in 4 % PFA for further usage.

5.2.8 *In vivo* BrdU-assay

BrdU (15 mg/ml in 0.9 % NaCl) was administered i.p. according to 150 mg/kg body weight to three C57BL/6N WT mice at 1 dpi after cerulein-induced pancreatitis. Mice were sacrificed and pancreas tissue was extracted at 91 dpi.

5.2.9 *In vitro* BrdU-assay

mADOs from one C57BL/6N WT mouse were cultured in 24-well plates as described above. Starting at 0 h after cell plating, BrdU was added to POC medium to a final concentration of 10 µM for three wells. After 2 h incubation the cells from these three wells were dissociated and fixed and BrdU was added to another three wells. This procedure was followed 24 h consecutively.

5.2.10 Immunohistochemistry (IHC) staining

After extraction, pancreatic tissue was post-fixed in 4 % PFA for 2 h at 4 ◦C. For tissue dehydration, the pancreas was immersed in 30 % sucrose in PBS and incubated at 4 °C until it settling to the bottom of the vial. The tissue was embedded in Tissue-Tek® O.C.T.™ Compound and stored at −80 °C. Prior to antibody staining, the tissue was cut into sections of 20 µm thickness using a cryomicrotome (−20 °C) and applied to microscope slides (Superfrost®Plus, Thermo Fisher). The tissue was incubated in PBS at RT until the surrounding Tissue-Tek® O.C.T.™ Compound was completely removed. Tissue sections were bordered using a hydrophobic pen. In case of BrdU staining, tissue sections were incubated in 2 M HCl for 30 min at 37 °C followed by neutralization with 0.15 M sodium borate buffer for 30 min at RT and three times washing in PBS for 5 min at RT. For permeabilization, the tissue was incubated in 0.1 % Triton™-X 100 in PBS for 30 min at RT. Blocking was performed by incubation in wash/block buffer (0.1 % Triton™-X 100 in PBS) for 1 h at RT. Primary antibodies were diluted in wash/block buffer according to Table S1 and applied to the tissue sections for primary staining over night at 4 °C. Afterwards, the tissue was washed three times with wash/block buffer for 5 min at RT. Secondary antibodies were diluted in wash/block buffer according to Table S1. Tissue sections were incubated in secondary antibody solution for 1 h at RT. After washing three times in wash/block buffer for 5 min at RT, the tissue was mounted using Fluoromount G, with DAPI (4',6'-diamidino-2-phenylindole) and stored at 4 °C until data acquisition. For BrdU staining, DNA hydrolysis was performed before permeabilization using 2M HCl for 30 min at 37 °C. HCl was neutralized using 0.15 M borate buffer for 30 min at RT. The tissue was washed three times for 5 min in PBS before continuation with the standard IHC protocol.

5.2.11 Immunocytochemistry (ICC) staining

After dissociation and fixation at given time points, single mADO and mMDO cells were fixed on microscopy slides by Cytospin™ centrifugation (10 min at 8000 rpm). Cell areas were bordered using a hydrophobic pen. For permeabilization, cells were incubated in 0.1 % Triton™-X 100 in PBS for 30 min at RT. Blocking was performed by incubation in wash/block buffer for 1 h at RT. Primary antibodies were diluted in wash/block buffer according to Table S1 and applied to the tissue sections for primary staining over night at 4 °C. Afterwards, the cells were washed three times with wash/block buffer for 5 min at RT. Secondary antibodies were diluted in wash/block buffer according to Table S1. Cells were incubated in secondary antibody solution for 1 h at RT. After washing three times in wash/block buffer for 5 min at RT, cells were mounted using Fluoromount G, with DAPI (4',6'-diamidino-2-phenylindole) and stored at 4 °C until data

acquisition. For BrdU staining, DNA hydrolysis was performed before permeabilization using 2M HCl for 30 min at 37 °C. HCl was neutralized using 0.15 M borate buffer for 30 min at RT. Cells were washed three times for 5 min in PBS before continuation with the standard IHC protocol.

5.2.12 Fixed-sample confocal imaging

IHC/ICC-stained tissue sections and cells were imaged as xyz stacks using a Leica TCS SP8 confocal microscope. The xy resolution was set to 1024 x 1024 pixels at a line frequency of 200 Hz (tissue sections) or 400 Hz (Cytospin™ slides) and a z-step size of 2 µm (Δz = 2 µm). Tunable spectral photomultiplier tubes were used as detectors. Images were acquired using a 40x immersion oil-based objective (Leica 40x Plan Apo NA 1.30). The following laser and filter setup was used to image fixed samples:

Table 10 | Optical setup for fixed-sample imaging.

Fluorophore	Excitation wavelength	Detection filter
DAPI	UV-diode (405 nm)	412 nm – 489 nm
Alexa Fluor® 488	Laser (488 nm)	504 nm – 543 nm
Alexa Fluor® 568	Laser (552 nm)	575 nm – 622 nm
Alexa Fluor® 647	Laser (638 nm)	641 nm – 763 nm

5.2.13 Live-cell imaging

Time-lapse imaging of mADOs and mMDOs was performed using a Nikon A1R confocal microscope as well as a Luxendo InVi SPIM light sheet microscope (long-term imaging for lineage tree reconstruction of mADOs only). Live-cell imaging using a Nikon A1R confocal microscope was performed with a 20x air objective (Nikon 20x Plan Apo λ NA 0.75) at z-step sizes of Δz = 2 µm. The xy resolution was set to 1024 x 1024 pixels at a line frequency of 200 Hz. Light sheet imaging using a Luxendo InVi SPIM was performed with a 10x illumination objective (10x CFI Plan Fl NA 0.3) and a 25x detection objective (Nikon 25x CFI Apo NA 1.1) at Δz = 1 µm and 2048 x 2048 pixels (Δx = 204 nm, Δy = 204 nm). The light sheet diameter was set to 1.7 µm.

Double-fluorescent H2B-mCherry/mG mice were used to detect single cells within organoids and to determine the number of nuclei per cell as well as NEB to anaphase durations. EGFP-Tuba mice were used to visualize spindle formations in mADOs to quantify and characterize multipolar division events. Time periods of imaging ranged from 8 h to 84 h depending on the experiment at

a temporal resolution of 6 frames per hour (Δt = 10 min). The following laser and filter setup was used to image live cell samples imaged by Nikon A1R:

Table 11 | Optical setup for live-cell imaging (Nikon A1R).

Fluorophore	Excitation Wavelength	Detection filter
GFP/EGFP	Laser (488 nm)	500 nm – 550 nm
mCherry	Laser (561 nm)	600 nm – 650 nm

The following laser and filter setup was used to image live cell samples imaged by Luxendo InVi SPIM:

Table 12 | Optical setup for live-cell imaging (Luxendo InVi SPIM).

Fluorophore	Excitation Wavelength	Detection filter
GFP	Laser (488 nm)	497 nm – 554 nm
mCherry	Laser (561 nm)	610 nm – 651 nm

5.2.14 Manual quantitative image analysis

Manual cell counting was performed for all experiments involving the analysis of tissue sections using the built-in Cellcounter tool in Fiji ImageJ (v2.14.0/1.54f). This includes the quantification of mononucleated and binucleated cells as well as CK19-, pHH3- and DCLK1-positive. Additionally, the quantification of amylase-expressing cells in mADOs was performed likewise. Manual image annotations for visualization purposes were drawn using Napari (v0.4.18) or Inkscape (v1.3.2). NEB to anaphase durations were determined by visual frame-to-frame inspection and quantification of time lapse movies of H2B-mCherry/mG mADOs using Fiji ImageJ (v2.14.0/1.54f). Bipolar and multipolar divisions in mADOs were assessed by screening EGFP-Tuba time-lapse movies for the corresponding type of division using Fiji ImageJ (v2.14.0/1.54f). The nuclear orientation axis of binucleated cells before bipolar or multipolar division was assessed by manually annotating and measuring the cell's long and short axes in the last time frame before NEB becomes visible using Fiji ImageJ (v2.14.0/1.54f).

For the lineage tree reconstruction from long-term live-cell imaging experiment, raw data was cropped using the Fiji plugin BigDataProcessor2 (v1.7.1). Cropped images were subjected to denoising via Noise2Void (v0.3.3, https://github.com/juglab/n2v) and deconvolution via flowdec (v1.1.0, https://github.com/hammerlab/flowdec). After conversion to BigDataViewer (v6.3.0) xml/h5 files, the lineage tree was reconstructed using Mastodon (v1.0.0-beta-30).

5.2.15 Automated quantitative image analysis

Dissociated organoids were fixed and adhered to microscopy slides using Cytospin and cells were subsequently stained for α-tubulin, geminin, γH2A.X, and nuclei using DAPI as described previously. Alpha-tubulin staining enabled the segmentation of cell bodies, while geminin was used to identify cells in G2/M (exhibiting doubled DNA content) for exclusion from further analysis, and γH2A.X to detect DNA damage; DAPI staining provided a measure of DNA content. Confocal images were exported as both maximum and sum intensity projections for each of the four channels using Fiji/ImageJ (v2.14.0/1.54f). The maximum projection of the alpha-tubulin channel was employed for instance segmentation with Cellpose (v2.2.3), whose "Cyto 2" model was fine-tuned on a subset of α-tubulin images. For each segmented object, features including area and perimeter (derived from the masked labels) and the median and integrated pixel intensities for DAPI, γH2A.X, and geminin (obtained from the masked sum intensity projections) were extracted using numpy (v1.24.3) and scikit-image (v0.21.0). Corrected total cellular fluorescence (CTCF) and corrected mean cellular fluorescence (CMnCF) values were then calculated for each channel by subtracting the background fluorescence intensity ($\mathrm{I}^{\mathrm{integrated}}_{\mathrm{background}}$, $\mathrm{I}^{\mathrm{mean}}_{\mathrm{background}}$) value multiplied with area of each object (A) from the object's mean and integrated fluorescence intensities ($\mathrm{I}^{\mathrm{integrated}}_{\mathrm{object}}$, $\mathrm{I}^{\mathrm{mean}}_{\mathrm{object}}$) respectively:

$$CTCF = I^{integrated}_{object} - I^{integrated}_{background} \cdot A$$

$$CTMnF = I^{mean}_{object} - I^{mean}_{background} \cdot A$$

Each objects circularity was calculated using the objects area and perimeter (P) according to:

$$\mathrm{C} = 4\pi \cdot \frac{\mathrm{A}}{\mathrm{P}}$$

Area and circularity were used to flag mis-segmented objects (e.g., doublets or cells at image borders). Subsequent analyses focused on DAPI CTCF and geminin and γH2A.X CMnCF, with the corresponding histograms scaled between 0 and 100 and smoothed via a Gaussian smoothing algorithm. For each channel, the largest peak was detected (local maximum approach via scipy.signal.find_peaks (v1.10.1)) and defined as diploid (2n) for DAPI, G0/G1 for geminin, and γH2A.X low for γH2A.X – an assumption verified by visual inspection. After Gaussian fitting around the largest peak with the peak center as an initial guess for the mean, the histograms were normalized according to the fitted mean of the normal distribution. In case of DNA content via DAPI CTCF, the values were additionally multiplied by the factor 2 to enable readability as

ploidy values. Normalized fluorescence thresholds were established as follows: DAPI CTCF values of 1.5–2.5 (diploid), 3.5–4.5 (tetraploid), and >4.5 (higher polyploidy). For the normalized geminin and γH2A.X values the thresholds were selected using the Dash app as presented in section 2.4.3 by plotting the corresponding CMnCF fluorescence intensity values against the normalized DAPI CTCF. To assess micronuclei presence and determine nuclear number, the maximum projection of the DAPI channel was cropped for each cell, based on the α-tubulin segmentation label, generating individual 90x90 pixels-sized cell patches. A classifier built on an ImageNet pre-trained EfficientNet B1 convolutional neural network (CNN) using Tensorflow (v2.13.0) was re-trained for 200 epochs on 1985 labeled images with a validation split of 0.2 and at a learning rate of 10^{-5} using an Nvidia RTX 4090 GPU, to categorize each cell patch into one of 5 classes (“mononucleated”, “binucleated”, “mononucleated micronucleated”, “binucleated micronucleated”, and “no cell” - representing mis-segmentation or debris):

```
base_model = tf.keras.applications.EfficientNetB1(include_top=False,
        input_shape = (90,90,3),
        pooling = 'avg',
        weights = 'imagenet')
```

The training data was augmented using a random flip in horizontal and vertical orientations, random rotation as well as random contrast. The batch size was set to 16. The last 24 layers were additionally unfrozen to enable feature leaning from high-level features. The top of the architecture was complemented by adding a flatten layer, a dense layer with 500 neuron (ReLu activation) and another dense layer with 5 neurons (Softmax activation) for class output. The final model consisted of 7233676 parameters. For the loss function, Sparse categorical cross entropy was used. The model performance was tracked by validation loss and best weights were saved. The model performance was tested on test set consisting of 396 images from all 5 classes. The model performance was assessed by precision (positive predictive value), recall (sensitivity) and F1-score (harmonic mean of precision and recall).

The resulting cell classification was integrated with the extracted features for downstream analysis. Full code for the image analysis pipeline is available at https://github.com/janbrunken/PloidyAnalysis_2D. In case of changes to the repository, all source code is of course also available upon request.

All classification results that went into the final quantifications were assessed by visual inspection of the classified image patches and manually corrected if necessary.

For the data analysis and visualization app development, Plotly Dash (v2.15.0) was used.

5.2.16 Fluorescence-activated cell sorting (FACS) for Strand-seq

mADOs generated from Fucci2 mice were cultured in POC medium as described above. BrdU was added to the medium at 40 µM and incubated for 40 h. At d9, the cells were dissociated into single cells as described above. An initial sorting step was performed with intact cells to exclude G2 (mVenus$^+$) cells and enrich for G0/G1 (mCherry$^+$). At this, single cells were sorted into nuclei staining buffer containing Hoechst33258 and propidium iodide (PI) as described by the author of the Strand-seq protocol (Sanders et al., 2017). After cell lysis, nuclei were sorted by Hoechst33258 signal to detect BrdU incorporation and PI signal to distinguish different DNA contents of nuclei (see Fig. S4a). For further details regarding the selection process of sorting gates, I refer to the original Strand-seq protocol (Sanders et al., 2017). Single BrdU-containing nuclei were sorted into 96-well plates containing 5 µl freeze buffer (42.5 % (v/v) 2x Profreeze-CDM, 7.5 % DMSO (v/v), 50 % (v/v) PBS) and stored at -80 °C until further processing.

5.2.17 Strand-seq

Strand-seq was used to detect structural variants in mADO cultures. The library preparation of sorted nuclei was performed by Katharina Bauer according to the original protocol as described in (Sanders et al., 2017). Sequencing was performed using a NextSeq 550 sequencing system using a NextSeq 500/550 Mid Output Kit v2.5 (150 cycles) according to the manufacturers instructions.

5.2.18 Processing and analysis of Strand-seq data

Processing and analysis of the Strand-seq data was performed by Enrico Frigoli. Raw sequencing files were processed using the MosaiCatcher pipeline (v2.3.5, (Weber et al., 2023)) to align reads on mm10 genome, filter low quality cells, and obtain raw read counts at 200kb genomic bins. The strandtools package (https://git.embl.de/cosenza/strandtools) was then used to normalize binned counts and to generate, in addition to custom python code, coverage track plots. For each genomic bin, the number of reads aligned to each strand defined the Watson Fraction (WF), that is, the number of reads aligned to the Watson strand over the total number of reads (sequencing depth) for each bin. For the ploidy of each nucleus, the presence of SVs was assessed by visually inspecting WFs and normalized total counts across each chromosome in each high-quality nucleus.

5.2.19 Statistical analysis

Statistical analyses were performed using scipy (v1.10.1). *P*-values less that 0.05 are considered as statistically significant. Comparisons between two groups were performed using a paired Student's *t* test in case of paired biological replicates or an unpaired Student's *t* test in case of unpaired biological replicates. Comparisons between more than two groups were performed using analysis of variance (ANOVA) followed by a Tukey's post hoc test. Cohen's d values were calculated using numpy (v1.23.5) according to:

$$d = \frac{\mu_1 - \mu_2}{\sqrt{\frac{s_1^2 + s_2^2}{2}}}$$

with μ_1 and μ_2 as the mean values of group 1 and 2 and s_1 and s_2 as their corresponding standard deviations.

5.2.20 Plotting

Graph visualization apart from Strand-seq was done using Plotly (v5.17.0). For box plots shown in this thesis, whiskers represent minimum and maximum values, the box covers the data between the q1 and q3 quartiles, the horizontal line represents the median and the dashed horizontal line represents the mean. Each data point is shown separately.
Figures were prepared using Inkscape (v1.3.2).

6 Appendix

6.1 Statistical analysis output

Statistical analyses were performed using scipy (v1.10.1). Note that the hormones estrogen (E_2) and progesterone (P_4) are labeled as E2 and P4 in the statistical analysis, respectively.

Figure 2.1 C: Nuclear number in vivo

Statistical analysis of the nuclear number in vivo was performed using one-way ANOVA followed by Tukey's post hoc test.

F_onewayResult(statistic=64.10323603339648, pvalue=4.3308487198860804e-07)

Multiple Comparison of Means - Tukey HSD, FWER=0.05

group1	group2	meandiff	p-adj	lower	upper	reject
2 dpi	28 dpi	-0.9433	0.9492	-5.3086	3.422	False
2 dpi	4 dpi	-7.6533	0.0013	-12.0186	-3.288	True
2 dpi	91 dpi	7.8267	0.0011	3.4614	12.192	True
2 dpi	naïve	11.2633	0.0001	6.898	15.6286	True
28 dpi	4 dpi	-6.71	0.0035	-11.0753	-2.3447	True
28 dpi	91 dpi	8.77	0.0004	4.4047	13.1353	True
28 dpi	naïve	12.2067	0.0	7.8414	16.572	True
4 dpi	91 dpi	15.48	0.0	11.1147	19.8453	True
4 dpi	naïve	18.9167	0.0	14.5514	23.282	True
91 dpi	naïve	3.4367	0.1456	-0.9286	7.802	False

Descriptive Statistics
naïve: min=41.96, max=44.44, mean=43.13666666666666, std=1.244842694212136, n=3
2 dpi: min=29.89, max=34.23, mean=31.873333333333335, std=2.1939538129444123, n=3
4 dpi: min=23.58, max=24.95, mean=24.22, std=0.6894200461257278, n=3
28 dpi: min=29.53, max=32.1, mean=30.930000000000003, std=1.3003461077728498, n=3
91 dpi: min=38.07, max=42.15, mean=39.699999999999996, std=2.1600694433281524, n=3

Figure 2.2 B: Amylase kinetics

Statistical analysis of the amylase kinetics was performed using one-way ANOVA followed by Tukey's post hoc test.

Multiple Comparison of Means - Tukey HSD, FWER=0.05

group1	group2	meandiff	p-adj	lower	upper	reject
d0	d1	-1.5567	0.9935	-14.2449	11.1315	False
d0	d3	-12.68	0.0502	-25.3682	0.0082	False
d0	d6	-70.7433	0.0	-83.4315	-58.0551	True
d0	d10	-95.3333	0.0	-108.0215	-82.6451	True
d1	d3	-11.1233	0.0935	-23.8115	1.5649	False
d1	d6	-69.1867	0.0	-81.8749	-56.4985	True
d1	d10	-93.7767	0.0	-106.4649	-81.0885	True
d3	d6	-58.0633	0.0	-70.7515	-45.3751	True
d3	d10	-82.6533	0.0	-95.3415	-69.9651	True
d6	d10	-24.59	0.0006	-37.2782	-11.9018	True

Descriptive statistics
d0: min=97.62, max=99.93, mean=98.62333333333333, std=1.184497080339728, n=3
d1: min=95.5, max=99.16, mean=97.06666666666666, std=1.885983386282425, n=3
d3: min=82.54, max=88.96, mean=85.94333333333333, std=3.227418989429989, n=3
d6: min=17.69, max=36.84, mean=27.88, std=9.63406975270576, n=3
d10: min=1.9, max=5.34, mean=3.2899999999999996, std=1.8124844826921969, n=3

Figure 2.2 C: Ploidy kinetics in vitro

1. Compare Polyploid vs Binucleated at each time point using paired t-test.

d0: Paired t-test between Binucleated & Polyploid: p-value = 0.1091
d5: Paired t-test between Binucleated & Polyploid: p-value = 0.0206
d13: Paired t-test between Binucleated & Polyploid: p-value = 0.0022

2. Compare changes over time for Binucleated and Polyploid using Repeated Measures ANOVA for normally distributed data

Repeated Measures ANOVA for Binucleated Cells:

Anova

	F Value	Num DF	Den DF	Pr > F
day	152.7068	2.0000	4.0000	0.0002

Repeated Measures ANOVA for Polyploid Cells:

Anova

	F Value	Num DF	Den DF	Pr > F
day	10.9743	2.0000	4.0000	0.0238

Post-hoc comparisons (pairwise comparisons across time points)

Post-hoc pairwise comparisons for Binucleated:
Multiple Comparison of Means - Tukey HSD, FWER=0.05

group1	group2	meandiff	p-adj	lower	upper	reject
d0	d5	-20.3433	0.0001	-26.4558	-14.2308	True
d0	d13	-27.0067	0.0	-33.1192	-20.8942	True
d5	d13	-6.6633	0.0357	-12.7758	-0.5508	True

Post-hoc pairwise comparisons for Polyploid:
Multiple Comparison of Means - Tukey HSD, FWER=0.05

group1	group2	meandiff	p-adj	lower	upper	reject
d0	d5	-7.2467	0.0422	-14.182	-0.3113	True
d0	d13	-11.95	0.0045	-18.8854	-5.0146	True
d5	d13	-4.7033	0.1742	-11.6387	2.232	False

Descriptive statistics
Descriptive statistics for Binucleated:
d0: min=37.55, max=42.27, mean=39.98, std=2.363112354502006, n=3
d5: min=17.75, max=22.24, mean=19.636666666666667, std=2.329213028757423, n=3
d13: min=9.98, max=14.83, mean=12.973333333333334, std=2.6171804166570807, n=3

Descriptive statistics for Polyploid:
d0: min=38.19, max=45.5, mean=42.199999999999996, std=3.706359399734463, n=3
d5: min=33.43, max=37.77, mean=34.95333333333333, std=2.4420141959729365, n=3
d13: min=28.32, max=31.92, mean=30.25, std=1.814028665705149, n=3

Figure 2.3 B: NEB to anaphase

Statistical analysis of the time from NEB to anaphase time of mononucleated and binucleaeted cells was performed using t-Test.

Ttest_indResult(statistic=-5.36459924475027, pvalue=2.304851485588728e-06)

Descriptive statistics
mononucleated : min=10, max=40, mean=19.166666666666668, std=7.755316082290385, n=24
binuncleated: min=20, max=50, mean=31.923076923076923, std=8.952868040680938, n=26

Figure 2.6 C Cell size

Statistical analysis of the cell size was performed using t test.

Ttest_indResult(statistic=8.08020717194209, pvalue=0.0012745821380656093)

Descriptive statistics
d0: min=65, max=5077, mean=923.8992048643593, std=564.9820610012046, n=4276
d6: min=57, max=4896, mean=536.6985149901933, std=378.119701757392, n=3569

Figure 2.18 A: yH2AX high cells

Statistical analysis of the yH2AX high cells was performed using t test.

Ttest_indResult(statistic=-34.93234958417755, pvalue=4.007477638810532e-06)

Descriptve statistics
2n: min=4.74898236092266, max=5.847953216374268, mean=5.170488497275614, std=0.5925118837153398, n=3
≥4n: min=18.088737201365188, max=18.55670103092784, mean=18.25223102784067, std=0.263919018306324, n=3

Figure 2.18 C: Micronucleated cells

Statistical analysis of the micronuclei in vitro was performed using one-way ANOVA followed by Tukey's post hoc test.

F_onewayResult(statistic=56.170059795990376, pvalue=0.00013033400466231704)

Multiple Comparison of Means - Tukey HSD, FWER=0.05

group1	group2	meandiff	p-adj	lower	upper	reject
2n	4n_binuc	7.69	0.0001	5.4637	9.9163	True
2n	4n_mononuc	3.9267	0.004	1.7004	6.153	True
4n_binuc	4n_mononuc	-3.7633	0.0049	-5.9896	-1.537	True

Descriptive statistics
2n: min=0.99, max=2.25, mean=1.6533333333333333, std=0.632639971337042, n=3
4n_mono: min=4.51, max=6.8, mean=5.579999999999999, std=1.152345434320803, n=3
4n_bin: min=8.7, max=10.24, mean=9.343333333333334, std=0.8006455728556388, n=3

Figure 2.24 B: K8-positive cells, all samples

One-way ANOVA between samples
ANOVA result: F=1.694244524611953, p=0.21040956612437076

Descriptive statistics
PLM -Prl: min=60.01825650387951, max=68.64437913299045, mean=65.34524360096022, std=4.656891063845407, n=3
MOM: min=56.4129856306546, max=74.395329441201, mean=66.244809194117, std=9.108307079881683, n=3

PLM -EGF min=66.18754277891856, max=66.72318779143704, mean=66.51132394728371, std=0.28482101476723515, n=3
PLM -P_4: min=71.47937411095306, max=74.80314960629921, mean=73.07278660366103, std=1.6661144821090481, n=3
PLM: min=65.49865229110512, max=76.81288173237091, mean=70.86005143069877, std=5.680254402490224, n=3
PLM -E2: min=61.46226415094339, max=64.57818796308332, mean=63.314098443220246, std=1.6390022083236957, n=3

Cohen's d for all sample pairs

- PLM -PRL vs MOM: -0.12436053770714725
- PLM -PRL vs PLM -EGF: -0.353457034343611
- PLM -PRL vs PLM -P4: -2.209557822706405
- PLM -PRL vs PLM: -1.0617980198463701
- PLM -PRL vs PLM -E2: 0.5818376204279868
- MOM vs PLM -PRL: 0.12436053770714725
- MOM vs PLM -EGF: -0.041360554886662265
- MOM vs PLM -P4: -1.0428514206763289
- MOM vs PLM: -0.6080415014028898
- MOM vs PLM -E2: 0.4478477448634707
- PLM -EGF vs PLM -PRL: 0.353457034343611
- PLM -EGF vs MOM: 0.041360554886662265
- PLM -EGF vs PLM -P4: -5.4897926635813015
- PLM -EGF vs PLM: -1.0813446270477443
- PLM -EGF vs PLM -E2: 2.717992643125063
- PLM -P4 vs PLM -PRL: 2.209557822706405
- PLM -P4 vs MOM: 1.0428514206763289
- PLM -P4 vs PLM -EGF: 5.4897926635813015
- PLM -P4 vs PLM: 0.5286336216936766
- PLM -P4 vs PLM -E2: 5.905001701348681
- PLM vs PLM -PRL: 1.0617980198463701
- PLM vs MOM: 0.6080415014028898
- PLM vs PLM -EGF: 1.0813446270477443
- PLM vs PLM -P4: -0.5286336216936766
- PLM vs PLM -E2: 1.8050758192031167
- PLM -E2 vs PLM -PRL: -0.5818376204279868
- PLM -E2 vs MOM: -0.4478477448634707
- PLM -E2 vs PLM -EGF: -2.717992643125063
- PLM -E2 vs PLM -P4: -5.905001701348681
- PLM -E2 vs PLM: -1.8050758192031167

Figure 2.25 B: β-Casein-positive cells, all samples

One-way ANOVA
ANOVA result: F=5.168473580615222, p=0.0092743326978943

Tukey HSD
Multiple Comparison of Means - Tukey HSD, FWER=0.05

group1	group2	meandiff	p-adj	lower	upper	reject
PLM -PRL	MOM	-0.0323	1.0	-4.6226	4.5579	False
PLM -PRL	PLM -EGF	4.1379	0.0865	-0.4524	8.7282	False
PLM -PRL	PLM -P4	1.7853	0.7766	-2.805	6.3755	False
PLM -PRL	PLM	5.4238	0.0179	0.8335	10.0141	True
PLM -PRL	PLM -E2	2.1667	0.6218	-2.4236	6.757	False
MOM	PLM -EGF	4.1702	0.0832	-0.4201	8.7605	False
MOM	PLM -P4	1.8176	0.7643	-2.7727	6.4079	False
MOM	PLM	5.4561	0.0172	0.8658	10.0464	True
MOM	PLM -E2	2.1991	0.6082	-2.3912	6.7894	False
PLM -EGF	PLM -P4	-2.3526	0.5439	-6.9429	2.2377	False
PLM -EGF	PLM	1.2859	0.9278	-3.3044	5.8762	False
PLM -EGF	PLM -E2	-1.9712	0.7033	-6.5614	2.6191	False
PLM -P4	PLM	3.6385	0.155	-0.9518	8.2288	False
PLM -P4	PLM -E2	0.3815	0.9997	-4.2088	4.9718	False
PLM	PLM -E2	-3.2571	0.2356	-7.8473	1.3332	False

Descriptive statistics
PLM -PRL: min=0.0, max=0.0516173434273916, mean=0.03234353469983593, std=0.02818156912537332, n=3
MOM: min=0.0, max=0.0, mean=0.0, std=0.0, n=3
PLM -EGF: min=1.973859695918912, max=6.890198968312454, mean=4.17022506852825, std=2.499644450240362, n=3
PLM -P4: min=1.3045659809332664, max=2.533692722371968, mean=1.8175964934327233, std=0.639229927744965, n=3
PLM: min=2.4904775857017287, max=7.6923076923076925, mean=5.4561174294395665, std=2.676533510489137, n=3
PLM -E2: min=0.6343452650657001, max=4.054054054054054, mean=2.199067005277458, std=1.72823392872268, n=3

Cohen's d for each pair of samples
- PLM -PRL vs MOM: 1.6230702138728714
- PLM -PRL vs PLM -EGF: -2.3409234502628817
- PLM -PRL vs PLM -P4: -3.9458085800697873
- PLM -PRL vs PLM: -2.8656280325758905
- PLM -PRL vs PLM -E2: -1.7727938070824172
- MOM vs PLM -EGF: -2.3593710895538837
- MOM vs PLM -P4: -4.021197225546492
- MOM vs PLM: -2.882876390814753
- MOM vs PLM -E2: -1.7994961976756816
- PLM -EGF vs PLM -P4: 1.289538743982284
- PLM -EGF vs PLM: -0.49656026862423047
- PLM -EGF vs PLM -E2: 0.9173132121038143
- PLM -P4 vs PLM: -1.8699145814666842
- PLM -P4 vs PLM -E2: -0.29277232638683753
- PLM vs PLM -E2: 1.4457496852803202

Figure 2.26 A: cycling cells

Paired t-Tests between b-Casein+ and b-Casein- cells per sample
PLM -PRL: t=-4.605122285461162, p=0.044060845016483295
PLM -EGF: t=8.810689515069122, p=0.012638219453633895
PLM -P4: t=14.691648412402264, p=0.0046010142992528895
PLM: t=6.337386482206637, p=0.024005919071619978
PLM -E2: t=9.592828105460802, p=0.010692938394270574
MOM: t=-12.624719751781495, p=0.0062157361417392734

Cohen's d for each sample between b-Casein+ and b-Casein-
PLM -PRL: -3.7600666008331145
PLM -EGF: 6.282240317704417
PLM -P4: 11.424565925918445
PLM: 6.2811846394055095
PLM -E2: 8.594251273449673
MOM: -10.308040512500321

One way ANOVA between samples for b-casein+ cells
ANOVA result: F=52.32485462399954, p=9.849859854635167e-08
Tukey HSD for b-casein+ cells
Multiple Comparison of Means - Tukey HSD, FWER=0.05

group1	group2	meandiff	p-adj	lower	upper	reject
PLM -PRL	MOM	0.0	1.0	-9.7872	9.7872	False
PLM -PRL	PLM -EGF	26.8853	0.0	17.0982	36.6725	True
PLM -PRL	PLM -P4	22.2631	0.0001	12.476	32.0503	True
PLM -PRL	PLM	36.6278	0.0	26.8406	46.4149	True
PLM -PRL	PLM -E2	22.3004	0.0001	12.5132	32.0875	True
MOM	PLM -EGF	26.8853	0.0	17.0982	36.6725	True
MOM	PLM -P4	22.2631	0.0001	12.476	32.0503	True
MOM	PLM	36.6278	0.0	26.8406	46.4149	True
MOM	PLM -E2	22.3004	0.0001	12.5132	32.0875	True
PLM -EGF	PLM -P4	-4.6222	0.6213	-14.4093	5.165	False
PLM -EGF	PLM	9.7424	0.0513	-0.0447	19.5296	False
PLM -EGF	PLM -E2	-4.5849	0.6287	-14.3721	5.2022	False
PLM -P4	PLM	14.3646	0.0036	4.5775	24.1518	True
PLM -P4	PLM -E2	0.0372	1.0	-9.7499	9.8244	False
PLM	PLM -E2	-14.3274	0.0037	-24.1145	-4.5402	True

Descriptive statistics for b-casein+ cells
PLM -PRL: min=0.0, max=0.0, mean=0.0, std=0.0, n=3
MOM: min=0.0, max=0.0, mean=0.0, std=0.0, n=3
PLM -EGF: min=23.711340206185564, max=32.62032085561498, mean=26.885328462041624, std=4.976096934147361, n=3
PLM -P4: min=20.51282051282051, max=25.0, mean=22.263138752500453, std=2.40075931056185, n=3
PLM: min=30.55555555555556, max=42.85714285714285, mean=36.627762215997514, std=6.152299592819014, n=3

PLM -E2: min=19.047619047619047, max=24.24242424242424, mean=22.3003848003848, std=2.834608059188502, n=3

Cohen's d for each pair of samples for b-casein+ cells
PLM -PRL vs PLM -EGF: 1.6230702138728714
PLM -PRL vs PLM -P4: -2.3409234502628817
PLM -PRL vs PLM: -3.9458085800697873
PLM -PRL vs PLM -E2: -2.8656280325758905
PLM -PRL vs MOM: -1.7727938070824172
PLM -EGF vs PLM -P4: -2.3593710895538837
PLM -EGF vs PLM: -4.021197225546492
PLM -EGF vs PLM -E2: -2.882876390814753
PLM -EGF vs MOM: -1.7994961976756816
PLM -P4 vs PLM: 1.289538743982284
PLM -P4 vs PLM -E2: -0.49656026862423047
PLM -P4 vs MOM: 0.9173132121038143
PLM vs PLM -E2: -1.8699145814666842
PLM vs MOM: -0.29277232638683753
PLM -E2 vs MOM: 1.4457496852803202

One way ANOVA between samples for b-casein- cells
ANOVA result: F=0.5019075253116787, p=0.769432740443051
Descriptive statistics for b-casein- cells
PLM -PRL: min=3.2485233984552475, max=6.079027355623101, mean=4.239613377542958, std=1.5945778025344526, n=3
MOM: min=2.2857142857142856, max=3.007518796992481, mean=2.6918523805421906, std=0.3693091950747171, n=3
PLM -EGF: min=3.782312925170068, max=5.263157894736842, mean=4.5368112513228525, std=0.74082376035301, n=3
PLM -P4: min=2.6548672566371683, max=2.9743589743589745, mean=2.8254004399880377, std=0.1608348175370093, n=3
PLM: min=0.0, max=8.022836538461538, mean=4.007205179080179, std=4.011424906576742, n=3
PLM -E2: min=2.560819462227913, max=4.756756756756757, mean=3.7818479984141526, std=1.1184660742105752, n=3

Cohen's d for each pair of samples for b-casein- cells
MOM vs PLM -EGF: 1.3372944928026835
MOM vs PLM: -0.23904309107349278
MOM vs PLM -P4: 1.2479181819361014
MOM vs PLM -E2: 0.07613966661193738
MOM vs PLM -PRL: 0.3323759664398203
PLM -EGF vs PLM: -3.1520283700061826
PLM -EGF vs PLM -P4: -0.4688681128448978
PLM -EGF vs PLM -E2: -0.46177012755827224
PLM -EGF vs PLM -PRL: -1.3087173127492544
PLM vs PLM -P4: 3.192664804133393
PLM vs PLM -E2: 0.1836059412489871
PLM vs PLM -PRL: 0.795848408252707
PLM -P4 vs PLM -E2: -0.4163065701052043
PLM -P4 vs PLM -PRL: -1.1970406746800497

PLM -E2 vs PLM -PRL: 0.07652980299155368

Figure 2.26 C: Ploidy all cells, all samples

One way ANOVA between samples
ANOVA result: F=5.83908881083062, p=0.0058271833167736495

Tukey HSD between samples

Multiple Comparison of Means - Tukey HSD, FWER=0.05

group1	group2	meandiff	p-adj	lower	upper	reject
PLM -PRL	MOM	-4.2122	0.6014	-12.9407	4.5164	False
PLM -PRL	PLM -EGF	6.0754	0.2515	-2.6531	14.804	False
PLM -PRL	PLM -P4	5.4554	0.3485	-3.2732	14.1839	False
PLM -PRL	PLM	3.636	0.7272	-5.0926	12.3645	False
PLM -PRL	PLM -E2	-3.2799	0.799	-12.0085	5.4486	False
MOM	PLM -EGF	10.2876	0.0182	1.559	19.0162	True
MOM	PLM -P4	9.6676	0.0272	0.939	18.3961	True
MOM	PLM	7.8481	0.0876	-0.8804	16.5767	False
MOM	PLM -E2	0.9322	0.999	-7.7963	9.6608	False
PLM -EGF	PLM -P4	-0.62	0.9999	-9.3486	8.1085	False
PLM -EGF	PLM	-2.4394	0.9285	-11.168	6.2891	False
PLM -EGF	PLM -E2	-9.3553	0.0333	-18.0839	-0.6268	True
PLM -P4	PLM	-1.8194	0.9784	-10.548	6.9091	False
PLM -P4	PLM -E2	-8.7353	0.0498	-17.4639	-0.0068	True
PLM	PLM -E2	-6.9159	0.1553	-15.6445	1.8127	False

Descriptive statistics
PLM -PRL: min=7.690984170681349, max=11.55015197568389, mean=9.335298485361227, std=1.9918409198526368, n=3
MOM: min=4.010025062656641, max=6.285714285714287, mean=5.123129700957249, std=1.1386512015587462, n=3
PLM -EGF: min=10.669511869831954, max=18.83458646616541, mean=15.410724991881215, std=4.238947076082041, n=3
PLM -P4: min=12.159223950493391, max=18.46619576185671, mean=14.790683480037636, std=3.2805499492436265, n=3
PLM: min=10.080183276059564, max=18.461538461538463, mean=12.971276129546055, std=4.7569465174401575, n=3
PLM -E2: min=3.79096813170216, max=7.741251325556734, mean=6.055375363721843, std=2.037696775982308, n=3

Cohen's d for each pair of samples
PLM -PRL vs MOM: 2.5963581758489633
PLM -PRL vs PLM -EGF: -1.8344760046441226
PLM -PRL vs PLM -P4: -2.01023722589662
PLM -PRL vs PLM: -0.9970760430584129
PLM -PRL vs PLM -E2: 1.627834759325887
MOM vs PLM -EGF: -3.314683912835187
MOM vs PLM -P4: -3.937171465010857

MOM vs PLM: -2.269109825533668
MOM vs PLM -E2: -0.5648037059660016
PLM -EGF vs PLM -P4: 0.1635922816796867
PLM -EGF vs PLM: 0.5414503128396695
PLM -EGF vs PLM -E2: 2.8130264515083523
PLM -P4 vs PLM: 0.4452799790273633
PLM -P4 vs PLM -E2: 3.1988436238375373
PLM and PLM -E2: 1.8899590608982675

Figure 2.26 D: Ploidy β-Casein+ and β-Casein- cells, all samples

Paired t-Tests between b-Casein+ and b-Casein- cells per sample
PLM -PRL: t=-8.129045894014114, p=0.014797789656658898
PLM -EGF: t=6.279066457470026, p=0.024437634677580134
PLM -P4: t=13.56902239806769, p=0.005387437040695935
PLM: t=3.101840869919939, p=0.09010772976728443
PLM -E2: t=-3.1389499157384257, p=0.08826185796025521
MOM: t=-7.793010645995699, p=0.016070203691524936

Cohen's d for each sample between b-Casein+ and b-Casein-
PLM -PRL: -6.637338178667094
PLM -EGF: 3.4282567348852147
PLM -P4: 5.9045217091123785
PLM: 3.3920165443922423
PLM -E2: -1.5964369997769094
MOM: -6.362966547588858

One way ANOVA between samples for b-casein+ cells
ANOVA result: F=43.094884536565075, p=2.948370332704509e-07

Tukey HSD for b-casein+ cells
Multiple Comparison of Means - Tukey HSD, FWER=0.05

group1	group2	meandiff	p-adj	lower	upper	reject
PLM -PRL	MOM	0.0	1.0	-12.9638	12.9638	False
PLM -PRL	PLM -EGF	27.025	0.0002	14.0612	39.9888	True
PLM -PRL	PLM -P4	29.7475	0.0001	16.7837	42.7112	True
PLM -PRL	PLM	40.1538	0.0	27.19	53.1176	True
PLM -PRL	PLM -E2	1.8519	0.996	-11.1119	14.8156	False
MOM	PLM -EGF	27.025	0.0002	14.0612	39.9888	True
MOM	PLM -P4	29.7475	0.0001	16.7837	42.7112	True
MOM	PLM	40.1538	0.0	27.19	53.1176	True
MOM	PLM -E2	1.8519	0.996	-11.1119	14.8156	False
PLM -EGF	PLM -P4	2.7225	0.9777	-10.2413	15.6862	False
PLM -EGF	PLM	13.1288	0.0465	0.165	26.0926	True
PLM -EGF	PLM -E2	-25.1731	0.0003	-38.1369	-12.2094	True
PLM -P4	PLM	10.4063	0.1471	-2.5574	23.3701	False
PLM -P4	PLM -E2	-27.8956	0.0001	-40.8594	-14.9318	True
PLM	PLM -E2	-38.302	0.0	-51.2657	-25.3382	True

Descriptive statistics for b-casein+ cells
PLM -PRL: min=0.0, max=0.0, mean=0.0, std=0.0 n=3
MOM: min=0.0, max=0.0, mean=0.0, std=0.0 n=3
PLM -EGF: min=24.324324324324326, max=29.94652406417112, mean=27.024990699945217, std=2.8175998915166995 n=3
PLM -P4: min=28.205128205128204, max=31.25, mean=29.747454082560466, std=1.5228256287947792 n=3
PLM: min=28.57142857142857, max=49.53703703703704, mean=40.1538022616454, std=10.654404889918641 n=3
PLM -E2: min=0.0, max=5.555555555555555, mean=1.8518518518518519, std=3.2075014954979206 n=3

Cohen's d for each pair of samples for b-casein+ cells
PLM -PRL vs PLM -EGF: 1.6230702138728714
PLM -PRL vs PLM -P4: -2.3409234502628817
PLM -PRL vs PLM: -3.9458085800697873
PLM -PRL vs PLM -E2: -2.8656280325758905
PLM -PRL vs MOM: -1.7727938070824172
PLM -EGF vs PLM -P4: -2.3593710895538837
PLM -EGF vs PLM: -4.021197225546492
PLM -EGF vs PLM -E2: -2.882876390814753
PLM -EGF vs MOM: -1.7994961976756816
PLM -P4 vs PLM: 1.289538743982284
PLM -P4 vs PLM -E2: -0.49656026862423047
PLM -P4 vs MOM: 0.9173132121038143
PLM vs PLM -E2: -1.8699145814666842
PLM vs MOM: -0.29277232638683753
PLM -E2 vs MOM: 1.4457496852803202

One way ANOVA between samples for b-casein- cells
ANOVA result: F=4.616185599626204, p=0.013976812802048547
Tukey HSD for b-casein- cells
Multiple Comparison of Means - Tukey HSD, FWER=0.05

group1	group2	meandiff	p-adj	lower	upper	reject
PLM -PRL	MOM	-4.2148	0.6412	-13.339	4.9093	False
PLM -PRL	PLM -EGF	5.5545	0.3738	-3.5696	14.6787	False
PLM -PRL	PLM -P4	5.1731	0.444	-3.9511	14.2973	False
PLM -PRL	PLM	2.213	0.9592	-6.9111	11.3372	False
PLM -PRL	PLM -E2	-3.1669	0.8444	-12.2911	5.9573	False
MOM	PLM -EGF	9.7693	0.0335	0.6452	18.8935	True
MOM	PLM -P4	9.3879	0.0425	0.2638	18.5121	True
MOM	PLM	6.4279	0.2415	-2.6963	15.552	False
MOM	PLM -E2	1.0479	0.9986	-8.0762	10.1721	False
PLM -EGF	PLM -P4	-0.3814	1.0	-9.5056	8.7428	False
PLM -EGF	PLM	-3.3415	0.8146	-12.4657	5.7827	False
PLM -EGF	PLM -E2	-8.7214	0.064	-17.8456	0.4027	False
PLM -P4	PLM	-2.9601	0.8764	-12.0842	6.1641	False
PLM -P4	PLM -E2	-8.34	0.0807	-17.4642	0.7842	False
PLM	PLM -E2	-5.3799	0.4051	-14.5041	3.7442	False

Descriptive statistics for b-casein- cells
PLM -PRL: min=7.694956102599415, max=11.55015197568389, mean=9.337949853142746, std=1.98963123041085, n=3
MOM: min=4.010025062656641, max=6.285714285714287, mean=5.123129700957249, std=1.1386512015587462, n=3
PLM -EGF: min=10.394557823129253, max=18.532969176746, mean=14.892476022778625, std=4.13640159098359, n=3
PLM -P4: min=11.947127605490596, max=18.25641025641026, mean=14.511061293199992, std=3.316408966365437, n=3
PLM: min=7.478632478632479, max=17.61904761904762, mean=11.550989519739518, std=5.356689447809045, n=3
PLM -E2: min=3.815169478644171, max=7.783783783783783, mean=6.1710552701477726, std=2.086069732888949, n=3

Cohen's d for each pair of samples for b-casein- cells
PLM -PRL vs MOM: 2.6001662432496007
PLM -PRL vs PLM -EGF: -1.7113771001949143
PLM -PRL vs PLM -P4: -1.8916538515320298
PLM -PRL vs PLM: -0.5477019486592535
PLM -PRL vs PLM -E2: 1.5536019155473508
MOM vs PLM -EGF: -3.2203029372878635
MOM vs PLM -P4: -3.786334424273017
MOM vs PLM: -1.659925004761544
MOM vs PLM -E2: -0.6235767054859247
PLM -EGF vs PLM -P4: 0.10174064810876596
PLM -EGF vs PLM: 0.6982376947146032
PLM -EGF vs PLM -E2: 2.6623929174997056
PLM -P4 vs PLM: 0.6644494890427349
PLM -P4 vs PLM -E2: 3.0103955253215213
PLM vs PLM -E2: 1.3235299494262656

Figure 2.27 B: β-Casein+ cycling cells by ploidy
Paired t-test between 2n and ≥4n
t=-5.400689209860441, p=0.032616685409400034

Descriptive statistics
2N: min=15.606936416184972, max=20.869565217391305, mean=18.70989360810924, std=2.7552052995443157, n=3
≥4N: min=45.30386740331492, max=58.2089552238806, mean=49.8376075423985, std=7.258150360197397, n=3

Cohen's d for 2n and ≥4n
Cohen's d: -5.670283497400286

Figure 2.27 D: yH2A.X high by ploidy, all cells

Paired t-test between 2n and ≥4n
t=-7.823436880639164, p=0.01594841759372563

Descriptive statistics
2n: min=0.2783576896311761, max=0.8373205741626795, mean=0.49310487581007306, std=0.30113360540514267, n=3
≥4n: min=2.515723270440252, max=2.9069767441860463, mean=2.7249978642026504, std=0.1970497708005941, n=3

Cohen's d for 2N and ≥4n
Cohen's d: -8.770747585670701

Acknowledgements

Looking back at this journey that began with a simple internship, it's hard to believe how far things have come. What started as a few tentative weeks in the lab slowly grew into a Master's project and eventually became the backbone of this doctoral thesis.

I'm immensely grateful for the trust and guidance that accompanied me at every stage of this evolution. I want to express my deep appreciation to my supervisor Prof. Ana Martin-Villalba. Your unwavering support made all the difference. You created a space where curiosity was welcome, ideas were challenged, and mistakes were part of the process. Thank you for believing in me, pushing me, and for always making time, no matter how packed your schedule.

To my second referee, Dr. Michael Milsom, thank you for your valuable feedback, your thoughtful critiques, and for being willing to engage deeply with the work, especially during my thesis advisory committee meetings. I'm equally grateful to the other two members of my TAC – Prof. Dr. Gislene Pereira and Prof. Dr. Jan Lohmann. Your perspectives were sharp, honest, and constructive.

To our collaborators and all external supporters - thank you for sharing your expertise, your patience, and occasionally your optimism when things didn't go as planned. Many thanks go to Prof. Dr. Eli Pikarsky, who brought a completely new perspective to the project and supported us from far away. Thank you. A heartfelt thanks goes to the Nikon Imaging Center Heidelberg, especially Dr. Ulrike Engel for teaching me the intricate secrets of microscopy. I'm equally thankful to Manuela Brom and all other members of the DKFZ Light Microscopy Facility. Your technical expertise and problem-solving mindset saved the day more times than I can count. A huge thanks to Katharina Bauer from the Single-cell Open Lab of the DKFZ for preparing the Strand-seq library.

To the students I had the pleasure of mentoring – Andrea, Sanzhar, Veronika, Berkay and Dasha, thank you for your energy, your questions, and for reminding me what it means to see a problem with fresh eyes. It's been a joy to watch you grow, and I've learned just as much from you as you may have from me.

To Enrico, the person who tackled the Strand-seq analysis – your quiet focus and dedication made a huge difference in this project. Thank you for your careful work and your time.
To our secretary Irmgard and our technicians Katrin, Aylin, Aleksa, Lisa and most importantly Sonja: thank you for help and your encouragement making sure the lab and this project kept functioning.

And finally, to the entire lab team – Suse, Axel, Hadil, Greta, Irene, Zeynep, Masha, Santi, Alena, Leo, Andre, Xiaoyu and all Alumni who shared my path – thank you for making the time in the lab so enjoyable. Science is better when it's shared with people who care, challenge each other, and laugh a lot.

Einen besonders großen Dank möchte ich an meine Eltern richten – Mama, Papa – ohne euch wäre das alles nicht möglich gewesen. Danke für euren Rückhalt, für euren Glauben an mich und für eure Liebe.

Und zu guter Letzt: Vielen Dank Lara, für deine unendliche Geduld, für deine Liebe, und für deine immerwährende Motivation, die sichtlich auf mich abgefärbt hat. Dieses Kapitel endet mit deinem Namen zwischen den Zeilen.

Contribution statement

This thesis is the result of my independent research and writing. I have developed the structure, arguments, and content of this work on my own. To enhance readability and improve language flow in certain sections, I made occasional use of ChatGPT (model 4o). However, this assistance was always based on my own templates, ideas, and formulations, and did not substitute for my own academic input or judgment. Additionally, I have submitted a manuscript that includes key results and figures from this thesis. As the manuscript is currently under review and has not yet been published, I have chosen not to cite it within this document.
The data presented in this thesis were collected and analyzed by me unless stated otherwise in the text or figure cations.

External contributions to this thesis can be summarized as follows.
Figure 2.7: Experimental work and image analysis were assisted by Daria Kocherhina, whom I supervised during a lab internship.
The Strand-seq library was prepared by Katharina Bauer. The Strand-seq analysis was performed by Enrico Frigoli.
Figure 2.18 was acquired with assistance from Sanzhar Aitbay, a master student, whom I supervised for his thesis.
Figure 2.21 and Figure 2.22 were prepared by Enrico Frigoli.
The image recorded in Figure 2.24 A was recorded by Veronika Lummer, an MD student, whom I supervised for her thesis project.
The experimental work in the lab for the quantifications in Figure 2.24 - Figure 2.27 was assisted by Sefa Berkay Cayir, whom I supervised for his master's thesis project.
The live-cell imaging experiments in Figure 2.26 and Figure 2.27 were performed by me and Veronika Lummer in a joint effort at the Nikon Imaging Center Heidelberg.

Bibliography

Abuetabh, Y., Wu, H. H., Chai, C., Al Yousef, H., Persad, S., Sergi, C. M., & Leng, R. (2022). DNA damage response revisited: the p53 family and its regulators provide endless cancer therapy opportunities. *Experimental & Molecular Medicine*, *54*(10), 1658–1669. https://doi.org/10.1038/s12276-022-00863-4

Albrektsen, G., Heuch, I., Hansen, S., & Kvåle, G. (2005). Breast cancer risk by age at birth, time since birth and time intervals between births: exploring interaction effects. *British Journal of Cancer*, *92*(1), 167–175. https://doi.org/10.1038/sj.bjc.6602302

Anderson, S. M., Rudolph, M. C., McManaman, J. L., & Neville, M. C. (2007). Key stages in mammary gland development. Secretory activation in the mammary gland: it's not just about milk protein synthesis! *Breast Cancer Research*, *9*(1), 204. https://doi.org/10.1186/bcr1653

Andreassen, P. R., Lohez, O. D., Lacroix, F. B., & Margolis, R. L. (2001). Tetraploid state induces p53-dependent arrest of nontransformed mammalian cells in G1. *Molecular Biology of the Cell*, *12*(5), 1315–1328.

Anzi, S., Stolovich-Rain, M., Klochendler, A., Fridlich, O., Helman, A., Paz-Sonnenfeld, A., Avni-Magen, N., Kaufman, E., Ginzberg, M. B., Snider, D., Ray, S., Brecht, M., Holmes, M. M., Meir, K., Avivi, A., Shams, I., Berkowitz, A., Shapiro, A. M. J., Glaser, B., … Dor, Y. (2018). Postnatal Exocrine Pancreas Growth by Cellular Hypertrophy Correlates with a Shorter Lifespan in Mammals. *Developmental Cell*, *45*(6), 726-737.e3. https://doi.org/https://doi.org/10.1016/j.devcel.2018.05.024

Arendt, L. M., & Kuperwasser, C. (2015). Form and Function: how Estrogen and Progesterone Regulate the Mammary Epithelial Hierarchy. *Journal of Mammary Gland Biology and Neoplasia*, *20*(1), 9–25. https://doi.org/10.1007/s10911-015-9337-0

Bailey, J. M., Alsina, J., Rasheed, Z. A., McAllister, F. M., Fu, Y., Plentz, R., Zhang, H., Pasricha, P. J., Bardeesy, N., Matsui, W., Maitra, A., & Leach, S. D. (2014). DCLK1 Marks a Morphologically Distinct Subpopulation of Cells With Stem Cell Properties in Preinvasive Pancreatic Cancer. *Gastroenterology*, *146*(1), 245–256. https://doi.org/https://doi.org/10.1053/j.gastro.2013.09.050

Baudoin, N. C., Nicholson, J. M., Soto, K., Martin, O., Chen, J., & Cimini, D. (2020). Asymmetric clustering of centrosomes defines the early evolution of tetraploid cells. *ELife*, *9*, e54565. https://doi.org/10.7554/eLife.54565

Bengtsson, A., Andersson, R., & Ansari, D. (2020). The actual 5-year survivors of pancreatic ductal adenocarcinoma based on real-world data. *Scientific Reports*, *10*(1), 16425. https://doi.org/10.1038/s41598-020-73525-y

Benigni, A., Morigi, M., & Remuzzi, G. (2010). Kidney regeneration. *The Lancet*, *375*(9722), 1310–1317. https://doi.org/10.1016/S0140-6736(10)60237-1

Bertelsen, B., Nazaryan-Petersen, L., Sun, W., Mehrjouy, M. M., Xie, G., Chen, W., Hjermind, L. E., Taschner, P. E. M., & Tümer, Z. (2016). A germline chromothripsis event stably segregating in 11 individuals through three generations. *Genetics in Medicine*, *18*(5), 494–500. https://doi.org/https://doi.org/10.1038/gim.2015.112

Beucher, S., & Lantuéjoul, C. (1979). Use of Watersheds in Contour Detection. In *In International Workshop on Image Processing: Real-time Edge and Motion Detection/Estimation* (Vol. 132).

Biswas, S. K., Banerjee, S., Baker, G. W., Kuo, C.-Y., & Chowdhury, I. (2022). The Mammary Gland: Basic Structure and Molecular Signaling during Development. *International Journal of Molecular Sciences*, *23*(7). https://doi.org/10.3390/ijms23073883

Bloomfield, M., & Cimini, D. (2023). The fate of extra centrosomes in newly formed tetraploid cells: should I stay, or should I go? *Frontiers in Cell and Developmental Biology*, *11*. https://doi.org/10.3389/fcell.2023.1210983

Borellini, F., & Oka, T. (1989). Growth control and differentiation in mammary epithelial cells. *Environmental Health Perspectives*, *80*, 85–99. https://doi.org/10.1289/ehp.898085

Borges, V. F., Lyons, T. R., Germain, D., & Schedin, P. (2020). Postpartum Involution and Cancer: An Opportunity for Targeted Breast Cancer Prevention and Treatments? *Cancer Research*, *80*(9), 1790–1798. https://doi.org/10.1158/0008-5472.CAN-19-3448

Bosch, M., Kallin, N., Donakonda, S., Zhang, J. D., Wintersteller, H., Hegenbarth, S., Heim, K., Ramirez, C., Fürst, A., Lattouf, E. I., Feuerherd, M., Chattopadhyay, S., Kumpesa, N., Griesser, V., Hoflack, J.-C., Siebourg-Polster, J., Mogler, C., Swadling, L., Pallett, L. J., … Knolle, P. A. (2024). A liver immune rheostat regulates CD8 T cell immunity in chronic HBV infection. *Nature*, *631*(8022), 867–875. https://doi.org/10.1038/s41586-024-07630-7

Boveri, T. (1902). Über mehrpolige Mitosen als Mittel zur Analyse des Zellkerns. *Verhandlungen Der Physikalisch-Medizinischen Gesellschaft Zu Würzburg*, *35*, 67 – 90.

Boveri, T. (1914). *Zur Frage der Entstehung maligner Tumoren*. Fischer.

Brembeck, F. H., Schreiber, F. S., Deramaudt, T. B., Craig, L., Rhoades, B., Swain, G., Grippo, P., Stoffers, D. A., Silberg, D. G., & Rustgi, A. K. (2003). The mutant K-ras oncogene causes pancreatic periductal lymphocytic infiltration and gastric mucous neck cell hyperplasia in

transgenic mice. *Cancer Research*, *63*(9), 2005–2009. http://europepmc.org/abstract/MED/12727809

Brisken, C., Heineman, A., Chavarria, T., Elenbaas, B., Tan, J., Dey, S. K., McMahon, J. A., McMahon, A. P., & Weinberg, R. A. (2000). Essential function of Wnt-4 in mammary gland development downstream of progesterone signaling. *Genes & Development*, *14*(6), 650–654. https://doi.org/10.1101/gad.14.6.650

Brisken, C., Park, S., Vass, T., Lydon, J. P., O'Malley, B. W., & Weinberg, R. A. (1998). A paracrine role for the epithelial progesterone receptor in mammary gland development. *Proceedings of the National Academy of Sciences*, *95*(9), 5076–5081. https://doi.org/10.1073/pnas.95.9.5076

Brodsky, V. Y. (1985). Uryvaeva IV. Genome multiplication in growth and development. Biology of polyploid and polytene cells. *Dev. Cell Biol. Ser*, *15*.

Brunken, J. (2019). *Addressing Acinar Cell Heterogeneity in Regeneration and Tumor Initiation in the Adult Pancreas*. DKFZ Heidelberg.

Burigotto, M., Mattivi, A., Migliorati, D., Magnani, G., Valentini, C., Roccuzzo, M., Offterdinger, M., Pizzato, M., Schmidt, A., Villunger, A., Maffini, S., & Fava, L. L. (2021). Centriolar distal appendages activate the centrosome-PIDDosome-p53 signalling axis via ANKRD26. *The EMBO Journal*, *40*(4), e104844. https://doi.org/https://doi.org/10.15252/embj.2020104844

Callier, V. (2019). Solving Peto's Paradox to better understand cancer. *Proceedings of the National Academy of Sciences*, *116*(6), 1825–1828. https://doi.org/10.1073/pnas.1821517116

Campbell, C., Wang, T., McNaughton, A. L., Barnes, E., & Matthews, P. C. (2021). Risk factors for the development of hepatocellular carcinoma (HCC) in chronic hepatitis B virus (HBV) infection: a systematic review and meta-analysis. *Journal of Viral Hepatitis*, *28*(3), 493–507. https://doi.org/https://doi.org/10.1111/jvh.13452

Carrière, C., Seeley, E. S., Goetze, T., Longnecker, D. S., & Korc, M. (2007). The Nestin progenitor lineage is the compartment of origin for pancreatic intraepithelial neoplasia. *Proceedings of the National Academy of Sciences*, *104*(11), 4437–4442. https://doi.org/10.1073/pnas.0701117104

Case, R. M. (1978). Synthesis, intracellular transport and discharge of exportable proteins in the pancreatic acinar cell and other cells. *Biological Reviews*, *53*(2), 211–347. https://doi.org/https://doi.org/10.1111/j.1469-185X.1978.tb01437.x

Chan, C. J., Smyth, M. J., & Martinet, L. (2014). Molecular mechanisms of natural killer cell activation in response to cellular stress. *Cell Death & Differentiation*, *21*(1), 5–14. https://doi.org/10.1038/cdd.2013.26

Cheng, H. M., Mah, K. K., & Seluakumaran, K. (2020). Pancreatic Exocrine Function: Pancreatic Acinar Cell in Pancreatic Secretion, Cholecystokinin (CCK), Auto-digestion. In H. M. Cheng, K. K. Mah, & K. Seluakumaran (Eds.), *Defining Physiology: Principles, Themes, Concepts. Volume 2: Neurophysiology and Gastrointestinal Systems* (pp. 53–55). Springer International Publishing. https://doi.org/10.1007/978-3-030-62285-5_15

Choudhury, D., Ghosh, D., Mondal, M., Singha, D., Pothuraju, R., & Malakar, P. (2024). Polyploidy and mTOR signaling: a possible molecular link. *Cell Communication and Signaling*, *22*(1), 196. https://doi.org/10.1186/s12964-024-01526-9

Ciarloni, L., Mallepell, S., & Brisken, C. (2007). Amphiregulin is an essential mediator of estrogen receptor α function in mammary gland development. *Proceedings of the National Academy of Sciences*, *104*(13), 5455–5460. https://doi.org/10.1073/pnas.0611647104

Cimini, D., Fioravanti, D., Salmon, E. D., & Degrassi, F. (2002). Merotelic kinetochore orientation versus chromosome mono-orientation in the origin of lagging chromosomes in human primary cells. *Journal of Cell Science*, *115*(3), 507–515. https://doi.org/10.1242/jcs.115.3.507

Cimini, D., Howell, B., Maddox, P., Khodjakov, A., Degrassi, F., & Salmon, E. D. (2001). Merotelic Kinetochore Orientation Is a Major Mechanism of Aneuploidy in Mitotic Mammalian Tissue Cells. *Journal of Cell Biology*, *153*(3), 517–528. https://doi.org/10.1083/jcb.153.3.517

Clarkson, R. W. E., Heeley, J. L., Chapman, R., Aillet, F., Hay, R. T., Wyllie, A., & Watson, C. J. (2000). NF-κB Inhibits Apoptosis in Murine Mammary Epithelia *. *Journal of Biological Chemistry*, *275*(17), 12737–12742. https://doi.org/10.1074/jbc.275.17.12737

Clarkson, R. W. E., Wayland, M. T., Lee, J., Freeman, T., & Watson, C. J. (2003). Gene expression profiling of mammary gland development reveals putative roles for death receptors and immune mediators in post-lactational regression. *Breast Cancer Research*, *6*(2), R92. https://doi.org/10.1186/bcr754

Collado, M., Gil, J., Efeyan, A., Guerra, C., Schuhmacher, A. J., Barradas, M., Benguría, A., Zaballos, A., Flores, J. M., Barbacid, M., Beach, D., & Serrano, M. (2005). Senescence in premalignant tumours. *Nature*, *436*(7051), 642. https://doi.org/10.1038/436642a

Comai, L. (2005). The advantages and disadvantages of being polyploid. *Nature Reviews Genetics*, *6*(11), 836–846. https://doi.org/10.1038/nrg1711

Cortes, C., & Vapnik, V. (1995). Support-vector networks. *Machine Learning*, *20*(3), 273–297. https://doi.org/10.1007/BF00994018

Cortés-Ciriano, I., Lee, J. J.-K., Xi, R., Jain, D., Jung, Y. L., Yang, L., Gordenin, D., Klimczak, L. J., Zhang, C.-Z., Pellman, D. S., Akdemir, K. C., Alvarez, E. G., Baez-Ortega, A., Beroukhim, R., Boutros, P. C., Bowtell, D. D. L., Brors, B., Burns, K. H., Campbell, P. J., … Consortium, P. (2020). Comprehensive analysis of chromothripsis in 2,658 human cancers using whole-genome sequencing. *Nature Genetics*, *52*(3), 331–341. https://doi.org/10.1038/s41588-019-0576-7

Cosenza, M. R., & Krämer, A. (2016). Centrosome amplification, chromosomal instability and cancer: mechanistic, clinical and therapeutic issues. *Chromosome Research*, *24*(1), 105–126. https://doi.org/10.1007/s10577-015-9505-5

Crasta, K., Ganem, N. J., Dagher, R., Lantermann, A. B., Ivanova, E. V, Pan, Y., Nezi, L., Protopopov, A., Chowdhury, D., & Pellman, D. (2012). DNA breaks and chromosome pulverization from errors in mitosis. *Nature*, *482*(7383), 53–58. https://doi.org/10.1038/nature10802

Crisafulli, S., Bertino, L., Fontana, A., Calapai, F., Ingrasciotta, Y., Berretta, M., Trifirò, G., & Guarneri, C. (2021). Incidence of Skin Cancer in Patients With Chronic Inflammatory Cutaneous Diseases on Targeted Therapies: A Systematic Review and Meta-Analysis of Observational Studies. *Frontiers in Oncology*, *Volume 11-2021*. https://www.frontiersin.org/journals/oncology/articles/10.3389/fonc.2021.687432

Dalal, N., & Triggs, B. (2005). Histograms of oriented gradients for human detection. *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)*, *1*, 886–893 vol. 1. https://doi.org/10.1109/CVPR.2005.177

Darmasaputra, G. S., Geerlings, C. C., Chuva de Sousa Lopes, S. M., Clevers, H., & Galli, M. (2024). Binucleated human hepatocytes arise through late cytokinetic regression during endomitosis M phase. *Journal of Cell Biology*, *223*(8), e202403020. https://doi.org/10.1083/jcb.202403020

Darmasaputra, G. S., van Rijnberk, L. M., & Galli, M. (2024). Functional consequences of somatic polyploidy in development. *Development*, *151*(5), dev202392. https://doi.org/10.1242/dev.202392

De Dosso, S., Siebenhüner, A. R., Winder, T., Meisel, A., Fritsch, R., Astaras, C., Szturz, P., & Borner, M. (2021). Treatment landscape of metastatic pancreatic cancer. *Cancer Treatment Reviews*, *96*. https://doi.org/10.1016/j.ctrv.2021.102180

de la Porte, P. L., lovanna, J., Odaira, C., Choux, R., Sarles, H., & Berger, Z. (1991). Involvement of Tubular Complexes in Pancreatic Regeneration After Acute Necrohemorrhagic Pancreatitis. *Pancreas*, *6*(3). https://journals.lww.com/pancreasjournal/fulltext/1991/05000/involvement_of_tubular_complexes_in_pancreatic.7.aspx

de Medeiros, G., Ortiz, R., Strnad, P., Boni, A., Moos, F., Repina, N., Challet Meylan, L., Maurer, F., & Liberali, P. (2022). Multiscale light-sheet organoid imaging framework. *Nature Communications*, *13*(1), 4864. https://doi.org/10.1038/s41467-022-32465-z

De Zio, D., Cianfanelli, V., & Cecconi, F. (2012). New Insights into the Link Between DNA Damage and Apoptosis. *Antioxidants & Redox Signaling*, *19*(6), 559–571. https://doi.org/10.1089/ars.2012.4938

Deng, J., Dong, W., Socher, R., Li, L.-J., Li, K., & Fei-Fei, L. (2009). ImageNet: A large-scale hierarchical image database. *2009 IEEE Conference on Computer Vision and Pattern Recognition*, 248–255. https://doi.org/10.1109/CVPR.2009.5206848

Dent, R., Trudeau, M., Pritchard, K. I., Hanna, W. M., Kahn, H. K., Sawka, C. A., Lickley, L. A., Rawlinson, E., Sun, P., & Narod, S. A. (2007). Triple-Negative Breast Cancer: Clinical Features and Patterns of Recurrence. *Clinical Cancer Research*, *13*(15), 4429–4434. https://doi.org/10.1158/1078-0432.CCR-06-3045

Desai, B. M., Oliver-Krasinski, J., De Leon, D. D., Farzad, C., Hong, N., Leach, S. D., & Stoffers, D. A. (2007). Preexisting pancreatic acinar cells contribute to acinar cell, but not islet β cell, regeneration. *The Journal of Clinical Investigation*, *117*(4), 971–977. https://doi.org/10.1172/JCI29988

Di Bona, M., & Bakhoum, S. F. (2024). Micronuclei and Cancer. *Cancer Discovery*, *14*(2), 214–226. https://doi.org/10.1158/2159-8290.CD-23-1073

Distler, M., Aust, D., Weitz, J., Pilarsky, C., & Grützmann, R. (2014). Precursor Lesions for Sporadic Pancreatic Cancer: PanIN, IPMN, and MCN. *BioMed Research International*, *2014*(1), 474905. https://doi.org/https://doi.org/10.1155/2014/474905

Doetsch, F. (2003). A niche for adult neural stem cells. *Current Opinion in Genetics & Development*, *13*(5), 543–550. https://doi.org/https://doi.org/10.1016/j.gde.2003.08.012

Dolley, D. H. (1925). The general morphology of pancreatic cell function in terms of the nucleo-cytoplasmic relation. *American Journal of Anatomy*, *35*(2), 153–197. https://doi.org/https://doi.org/10.1002/aja.1000350202

Donne, R., Saroul-Aïnama, M., Cordier, P., Celton-Morizur, S., & Desdouets, C. (2020). Polyploidy in liver development, homeostasis and disease. *Nature Reviews Gastroenterology & Hepatology*, *17*(7), 391–405. https://doi.org/10.1038/s41575-020-0284-x

Duensing, A., & Duensing, S. (2010). Centrosomes, polyploidy and cancer. *Adv Exp Med Biol*, *676*(1), 93–103.

Duncan, A. W., Taylor, M. H., Hickey, R. D., Hanlon Newell, A. E., Lenzi, M. L., Olson, S. B., Finegold, M. J., & Grompe, M. (2010). The ploidy conveyor of mature hepatocytes as a source of genetic variation. *Nature*, *467*(7316), 707–710. https://doi.org/10.1038/nature09414

Ebrahimi, F., Simon, T. G., Hagström, H., Söderling, J., Wester, A., Roelstraete, B., & Ludvigsson, J. F. (2023). Risk of Severe Infection in Patients With Biopsy-proven Nonalcoholic Fatty Liver Disease – A Population-based Cohort Study. *Clinical Gastroenterology and Hepatology*, *21*(13), 3346-3355.e19. https://doi.org/10.1016/j.cgh.2023.05.013

Ewald, A. J., Brenot, A., Duong, M., Chan, B. S., & Werb, Z. (2008). Collective Epithelial Migration and Cell Rearrangements Drive Mammary Branching Morphogenesis. *Developmental Cell*, *14*(4), 570–581. https://doi.org/10.1016/j.devcel.2008.03.003

Faggioli, F., Vezzoni, P., & Montagna, C. (2011). Single-cell analysis of ploidy and centrosomes underscores the peculiarity of normal hepatocytes. *PloS One*, *6*(10), e26080.

Fenech, M., Kirsch-Volders, M., Natarajan, A. T., Surralles, J., Crott, J. W., Parry, J., Norppa, H., Eastmond, D. A., Tucker, J. D., & Thomas, P. (2011). Molecular mechanisms of micronucleus, nucleoplasmic bridge and nuclear bud formation in mammalian and human cells. *Mutagenesis*, *26*(1), 125–132. https://doi.org/10.1093/mutage/geq052

Ferreira, R. M. M., Sancho, R., Messal, H. A., Nye, E., Spencer-Dene, B., Stone, R. K., Stamp, G., Rosewell, I., Quaglia, A., & Behrens, A. (2017). Duct- and Acinar-Derived Pancreatic Ductal Adenocarcinomas Show Distinct Tumor Progression and Marker Expression. *Cell Reports*, *21*(4), 966–978. https://doi.org/10.1016/j.celrep.2017.09.093

Flynn, P. J., Koch, P. D., & Mitchison, T. J. (2021). Chromatin bridges, not micronuclei, activate cGAS after drug-induced mitotic errors in human cells. *Proceedings of the National Academy of Sciences*, *118*(48), e2103585118. https://doi.org/10.1073/pnas.2103585118

Fu, N. Y., Rios, A. C., Pal, B., Soetanto, R., Lun, A. T. L., Liu, K., Beck, T., Best, S. A., Vaillant, F., Bouillet, P., Strasser, A., Preiss, T., Smyth, G. K., Lindeman, G. J., & Visvader, J. E. (2015). EGF-mediated induction of Mcl-1 at the switch to lactation is essential for alveolar cell survival. *Nature Cell Biology*, *17*(4), 365–375. https://doi.org/10.1038/ncb3117

Galati, F., Magri, V., Arias-Cadena, P. A., Moffa, G., Rizzo, V., Pasculli, M., Botticelli, A., & Pediconi, F. (2023). Pregnancy-Associated Breast Cancer: A Diagnostic and Therapeutic Challenge. *Diagnostics*, *13*(4). https://doi.org/10.3390/diagnostics13040604

Gallardo, M., Bickham, J., Honeycutt, R., Ojeda, R., & Nelida, K. (1999). Discovery of tetraploidy in a mammal. *Nature*, *401*, 341. https://doi.org/10.1038/43815

Gallardo, M. H., González, C. A., & Cebrián, I. (2006). Molecular cytogenetics and allotetraploidy in the red vizcacha rat, Tympanoctomys barrerae (Rodentia, Octodontidae). *Genomics*, *88*(2), 214–221. https://doi.org/https://doi.org/10.1016/j.ygeno.2006.02.010

Gandhi, S., de la Fuente, J., Murad, M. H., & Majumder, S. (2022). Chronic Pancreatitis Is a Risk Factor for Pancreatic Cancer, and Incidence Increases With Duration of Disease: A Systematic Review and Meta-analysis. *Clinical and Translational Gastroenterology*, *13*(3). https://journals.lww.com/ctg/fulltext/2022/03000/chronic_pancreatitis_is_a_risk_factor_for.5.aspx

Ganem, N. J., Godinho, S. A., & Pellman, D. (2009). A mechanism linking extra centrosomes to chromosomal instability. *Nature*, *460*(7252), 278–282. https://doi.org/10.1038/nature08136

Gemble, S., Wardenaar, R., Keuper, K., Srivastava, N., Nano, M., Macé, A.-S., Tijhuis, A. E., Bernhard, S. V., Spierings, D. C. J., Simon, A., Goundiam, O., Hochegger, H., Piel, M., Foijer, F., Storchová, Z., & Basto, R. (2022). Genetic instability from a single S phase after whole-genome duplication. *Nature*, *604*(7904), 146–151. https://doi.org/10.1038/s41586-022-04578-4

Gentric, G., & Desdouets, C. (2014). Polyploidization in Liver Tissue. *The American Journal of Pathology*, *184*(2), 322–331. https://doi.org/https://doi.org/10.1016/j.ajpath.2013.06.035

Gisselsson, D. (2008). Classification of chromosome segregation errors in cancer. *Chromosoma*, *117*(6), 511–519. https://doi.org/10.1007/s00412-008-0169-1

Glaubitz, J., Asgarbeik, S., Lange, R., Mazloum, H., Elsheikh, H., Weiss, F. U., & Sendler, M. (2023). Immune response mechanisms in acute and chronic pancreatitis: strategies for therapeutic intervention. *Frontiers in Immunology*, *Volume 14-2023*. https://www.frontiersin.org/journals/immunology/articles/10.3389/fimmu.2023.1279539

Goldring, M. B., & Marcu, K. B. (2009). Cartilage homeostasis in health and rheumatic diseases. *Arthritis Research & Therapy*, *11*(3), 224. https://doi.org/10.1186/ar2592

Gonzalez, R., & Woods, R. (2017). *Digital Image Processing Global Edition*. Pearson Deutschland. https://elibrary.pearson.de/book/99.150005/9781292223070

Graham, T. A., & Sottoriva, A. (2017). Measuring cancer evolution from the genome. *The Journal of Pathology*, *241*(2), 183–191. https://doi.org/https://doi.org/10.1002/path.4821

Grimont, A., Leach, S. D., & Chandwani, R. (2022). Uncertain Beginnings: Acinar and Ductal Cell Plasticity in the Development of Pancreatic Cancer. *Cellular and Molecular Gastroenterology and Hepatology*, *13*(2), 369–382. https://doi.org/https://doi.org/10.1016/j.jcmgh.2021.07.014

Guerra, C., Schuhmacher, A. J., Cañamero, M., Grippo, P. J., Verdaguer, L., Pérez-Gallego, L., Dubus, P., Sandgren, E. P., & Barbacid, M. (2007). Chronic Pancreatitis Is Essential for Induction of Pancreatic Ductal Adenocarcinoma by K-Ras Oncogenes in Adult Mice. *Cancer Cell*, *11*(3), 291–302. https://doi.org/https://doi.org/10.1016/j.ccr.2007.01.012

Guerrero, A. A., Martínez-A, C., & van Wely, K. H. M. (2010). Merotelic attachments and non-homologous end joining are the basis of chromosomal instability. *Cell Division*, *5*(1), 13. https://doi.org/10.1186/1747-1028-5-13

Guo, W., Keckesova, Z., Donaher, J. L., Shibue, T., Tischler, V., Reinhardt, F., Itzkovitz, S., Noske, A., Zürrer-Härdi, U., Bell, G., Tam, W. L., Mani, S. A., van Oudenaarden, A., & Weinberg, R. A. (2012). Slug and Sox9 Cooperatively Determine the Mammary Stem Cell State. *Cell*, *148*(5), 1015–1028. https://doi.org/https://doi.org/10.1016/j.cell.2012.02.008

Guo, Y., & Pu, W. T. (2020). Cardiomyocyte Maturation. *Circulation Research*, *126*(8), 1086–1106. https://doi.org/10.1161/CIRCRESAHA.119.315862

Hatch, E. M., Fischer, A. H., Deerinck, T. J., & Hetzer, M. W. (2013). Catastrophic Nuclear Envelope Collapse in Cancer Cell Micronuclei. *Cell*, *154*(1), 47–60. https://doi.org/10.1016/j.cell.2013.06.007

Hayashi, K., Horisaka, K., Harada, Y., Ogawa, Y., Yamashita, T., Kitano, T., Wakita, M., Fukusumi, T., Inohara, H., Hara, E., & Matsumoto, T. (2024). Polyploidy mitigates the impact of DNA damage while simultaneously bearing its burden. *Cell Death Discovery*, *10*(1), 436. https://doi.org/10.1038/s41420-024-02206-w

He, K., Zhang, X., Ren, S., & Sun, J. (2015). *Deep Residual Learning for Image Recognition*. https://arxiv.org/abs/1512.03385

Heinke, P., Rost, F., Rode, J., Trus, P., Simonova, I., Lázár, E., Feddema, J., Welsch, T., Alkass, K., Salehpour, M., Zimmermann, A., Seehofer, D., Possnert, G., Damm, G., Druid, H., Brusch, L., & Bergmann, O. (2022). Diploid hepatocytes drive physiological liver renewal in adult humans. *Cell Systems*, *13*(6), 499-507.e12. https://doi.org/10.1016/j.cels.2022.05.001

Hennigar, S. R., Seo, Y. A., Sharma, S., Soybel, D. I., & Kelleher, S. L. (2015). ZnT2 is a critical mediator of lysosomal-mediated cell death during early mammary gland involution. *Scientific Reports*, *5*(1), 8033. https://doi.org/10.1038/srep08033

Hennighausen, L., & Robinson, G. W. (2005). Information networks in the mammary gland. *Nature Reviews Molecular Cell Biology*, *6*(9), 715–725. https://doi.org/10.1038/nrm1714

Hertwig, O. (1884). *Das Problem der Befruchtung und der Isotropie des Eies: eine Theorie der Vererbung* (Vol. 18). Verlag von Gustav Fischer.

Hintzsche, H., Hemmann, U., Poth, A., Utesch, D., Lott, J., & Stopper, H. (2017). Fate of micronuclei and micronucleated cells. *Mutation Research/Reviews in Mutation Research*, *771*, 85–98. https://doi.org/https://doi.org/10.1016/j.mrrev.2017.02.002

Hruban, R. H., Goggins, M., Parsons, J., & Kern, S. E. (2000). Progression Model for Pancreatic Cancer. *Clinical Cancer Research*, *6*(8), 2969–2972.

Iismaa, S. E., Kaidonis, X., Nicks, A. M., Bogush, N., Kikuchi, K., Naqvi, N., Harvey, R. P., Husain, A., & Graham, R. M. (2018). Comparative regenerative mechanisms across different mammalian tissues. *Npj Regenerative Medicine*, *3*(1), 6. https://doi.org/10.1038/s41536-018-0044-5

Ionescu-Tirgoviste, C., Gagniuc, P. A., Gubceac, E., Mardare, L., Popescu, I., Dima, S., & Militaru, M. (2015). A 3D map of the islet routes throughout the healthy human pancreas. *Scientific Reports*, *5*(1), 14634. https://doi.org/10.1038/srep14634

Jennings, R. E., Berry, A. A., Strutt, J. P., Gerrard, D. T., & Hanley, N. A. (2015). Human pancreas development. *Development*, *142*(18), 3126–3137. https://doi.org/10.1242/dev.120063

Jensen, J. N., Cameron, E., Garay, M. V. R., Starkey, T. W., Gianani, R., & Jensen, J. (2005). Recapitulation of elements of embryonic development in adult mouse pancreatic regeneration. *Gastroenterology*, *128*(3), 728–741. https://doi.org/10.1053/j.gastro.2004.12.008

Jiang, S., Katayama, H., Wang, J., Li, S. A., Hong, Y., Radvanyi, L., Li, J. J., & Sen, S. (2010). Estrogen-Induced Aurora Kinase-A (AURKA) Gene Expression is Activated by GATA-3 in Estrogen Receptor-Positive Breast Cancer Cells. *Hormones and Cancer*, *1*(1), 11–20. https://doi.org/10.1007/s12672-010-0006-x

Jiang, Z., White, R. A., & Wang, T. C. (2020). Adult Pancreatic Acinar Progenitor-like Populations in Regeneration and Cancer. *Trends in Molecular Medicine*, *26*(8), 758–767. https://doi.org/10.1016/j.molmed.2020.04.003

Jiang, Z., Wu, F., Laise, P., Takayuki, T., Na, F., Kim, W., Kobayashi, H., Chang, W., Takahashi, R., Valenti, G., Sunagawa, M., White, R. A., Macchini, M., Renz, B. W., Middelhoff, M.,

Hayakawa, Y., Dubeykovskaya, Z. A., Tan, X., Chu, T. H., … Wang, T. C. (2023). Tff2 defines transit-amplifying pancreatic acinar progenitors that lack regenerative potential and are protective against Kras-driven carcinogenesis. *Cell Stem Cell*, *30*(8), 1091-1109.e7. https://doi.org/10.1016/j.stem.2023.07.002

Jung, A. Y., Ahearn, T. U., Behrens, S., Middha, P., Bolla, M. K., Wang, Q., Arndt, V., Aronson, K. J., Augustinsson, A., Beane Freeman, L. E., Becher, H., Brenner, H., Canzian, F., Carey, L. A., Consortium, C. T. S., Czene, K., Eliassen, A. H., Eriksson, M., Evans, D. G., … Chang-Claude, J. (2022). Distinct Reproductive Risk Profiles for Intrinsic-Like Breast Cancer Subtypes: Pooled Analysis of Population-Based Studies. *JNCI: Journal of the National Cancer Institute*, *114*(12), 1706–1719. https://doi.org/10.1093/jnci/djac117

Kalsbeek, D., & Golsteyn, R. M. (2017). G2/M-Phase Checkpoint Adaptation and Micronuclei Formation as Mechanisms That Contribute to Genomic Instability in Human Cells. *International Journal of Molecular Sciences*, *18*(11). https://doi.org/10.3390/ijms18112344

Kanda, M., Matthaei, H., Wu, J., Hong, S., Yu, J., Borges, M., Hruban, R. H., Maitra, A., Kinzler, K., Vogelstein, B., & Goggins, M. (2012). Presence of Somatic Mutations in Most Early-Stage Pancreatic Intraepithelial Neoplasia. *Gastroenterology*, *142*(4), 730-733.e9. https://doi.org/https://doi.org/10.1053/j.gastro.2011.12.042

Kandala, S., Ramos, M., Voith von Voithenberg, L., Diaz-Jimenez, A., Chocarro, S., Keding, J., Brors, B., Imbusch, C. D., & Sotillo, R. (2023). Chronic chromosome instability induced by Plk1 results in immune suppression in breast cancer. *Cell Reports*, *42*(12). https://doi.org/10.1016/j.celrep.2023.113266

Kim, H. (2008). Cerulein Pancreatitis: Oxidative Stress, Inflammation, and Apoptosis. *Gut and Liver*, *2*(2), 74–80. https://doi.org/10.5009/gnl.2008.2.2.74

Kirillova, A., Han, L., Liu, H., & Kühn, B. (2021). Polyploid cardiomyocytes: implications for heart regeneration. *Development*, *148*(14), dev199401. https://doi.org/10.1242/dev.199401

Klisch, K., Schraner, E. M., & Boos, A. (2016). Centrosome Clustering in the Development of Bovine Binucleate Trophoblast Giant Cells. *Cells Tissues Organs*, *203*(5), 287–294. https://doi.org/10.1159/000452271

Kloosterman, W. P., & Cuppen, E. (2013). Chromothripsis in congenital disorders and cancer: similarities and differences. *Current Opinion in Cell Biology*, *25*(3), 341–348. https://doi.org/https://doi.org/10.1016/j.ceb.2013.02.008

Kopp, J. L., von Figura, G., Mayes, E., Liu, F.-F., Dubois, C. L., Morris IV, J. P., Pan, F. C., Akiyama, H., Wright, C. V. E., Jensen, K., Hebrok, M., & Sander, M. (2012). Identification of Sox9-Dependent Acinar-to-Ductal Reprogramming as the Principal Mechanism for Initiation of

Pancreatic Ductal Adenocarcinoma. *Cancer Cell*, *22*(6), 737–750. https://doi.org/10.1016/j.ccr.2012.10.025

Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). ImageNet Classification with Deep Convolutional Neural Networks. In F. Pereira, C. J. Burges, L. Bottou, & K. Q. Weinberger (Eds.), *Advances in Neural Information Processing Systems* (Vol. 25). Curran Associates, Inc. https://proceedings.neurips.cc/paper_files/paper/2012/file/c399862d3b9d6b76c8436e924a68c45b-Paper.pdf

Kubiak Jacek Z. and Prigent, C. (2012). The Centrosome Life Story in Xenopuslaevis. In H. Schatten (Ed.), *The Centrosome: Cell and Molecular Mechanisms of Functions and Dysfunctions in Disease* (pp. 347–363). Humana Press. https://doi.org/10.1007/978-1-62703-035-9_20

Kwon, M., Godinho, S. A., Chandhok, N. S., Ganem, N. J., Azioune, A., Thery, M., & Pellman, D. (2008). Mechanisms to suppress multipolar divisions in cancer cells with extra centrosomes. *Genes & Development*, *22*(16), 2189–2203. https://doi.org/10.1101/gad.1700908

Lecun, Y., Bottou, L., Bengio, Y., & Haffner, P. (1998). Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, *86*(11), 2278–2324. https://doi.org/10.1109/5.726791

Lee, A. Y. L., Dubois, C. L., Sarai, K., Zarei, S., Schaeffer, D. F., Sander, M., & Kopp, J. L. (2019). Cell of origin affects tumour development and phenotype in pancreatic ductal adenocarcinoma. *Gut*, *68*(3), 487. https://doi.org/10.1136/gutjnl-2017-314426

Lee, H. J., & Ormandy, C. J. (2012). Interplay between progesterone and prolactin in mammary development and implications for breast cancer. *Molecular and Cellular Endocrinology*, *357*(1), 101–107. https://doi.org/https://doi.org/10.1016/j.mce.2011.09.020

Li, C. M.-C., Shapiro, H., Tsiobikas, C., Selfors, L. M., Chen, H., Rosenbluth, J., Moore, K., Gupta, K. P., Gray, G. K., Oren, Y., Steinbaugh, M. J., Guerriero, J. L., Pinello, L., Regev, A., & Brugge, J. S. (2020). Aging-Associated Alterations in Mammary Epithelia and Stroma Revealed by Single-Cell RNA Sequencing. *Cell Reports*, *33*(13). https://doi.org/10.1016/j.celrep.2020.108566

LI, D. W., YANG, Q., CHEN, J. T., ZHOU, H., LIU, R. M., & HUANG, X. T. (2005). Dynamic distribution of Ser-10 phosphorylated histone H3 in cytoplasm of MCF-7 and CHO cells during mitosis. *Cell Research*, *15*(2), 120–126. https://doi.org/10.1038/sj.cr.7290276

Li, X., Zhang, L., Wei, X., Datta, T., Wei, F., & Xie, Z. (2024). Polyploidization: A Biological Force That Enhances Stress Resistance. *International Journal of Molecular Sciences*, *25*(4). https://doi.org/10.3390/ijms25041957

Li, Y., Roberts, N. D., Wala, J. A., Shapira, O., Schumacher, S. E., Kumar, K., Khurana, E., Waszak, S., Korbel, J. O., Haber, J. E., Imielinski, M., Akdemir, K. C., Alvarez, E. G., Baez-Ortega, A., Beroukhim, R., Boutros, P. C., Bowtell, D. D. L., Brors, B., Burns, K. H., … Consortium, P. (2020). Patterns of somatic structural variation in human cancer genomes. *Nature*, *578*(7793), 112–121. https://doi.org/10.1038/s41586-019-1913-9

Lin, H., Huang, Y.-S., Fustin, J.-M., Doi, M., Chen, H., Lai, H.-H., Lin, S.-H., Lee, Y.-L., King, P.-C., Hou, H.-S., Chen, H.-W., Young, P.-Y., & Chao, H.-W. (2021). Hyperpolyploidization of hepatocyte initiates preneoplastic lesion formation in the liver. *Nature Communications*, *12*(1), 645. https://doi.org/10.1038/s41467-020-20572-8

Lin, Y.-H., Zhang, S., Zhu, M., Lu, T., Chen, K., Wen, Z., Wang, S., Xiao, G., Luo, D., Jia, Y., Li, L., MacConmara, M., Hoshida, Y., Singal, A. G., Yopp, A., Wang, T., & Zhu, H. (2020). Mice With Increased Numbers of Polyploid Hepatocytes Maintain Regenerative Capacity But Develop Fewer Hepatocellular Carcinomas Following Chronic Liver Injury. *Gastroenterology*, *158*(6), 1698-1712.e14. https://doi.org/https://doi.org/10.1053/j.gastro.2020.01.026

Liu, Q., Wuu, J., Lambe, M., Hsieh, S.-F., Ekbom, A., & Hsieh, C.-C. (2002). Transient increase in breast cancer risk after giving birth: postpartum period with the highest risk (Sweden). *Cancer Causes & Control*, *13*(4), 299–305. https://doi.org/10.1023/A:1015287208222

Longnecker, D. S., & Thompson, E. D. (2023). Anatomy, Histology, and Fine Structure of the Pancreas. In *The Pancreas* (pp. 9–22). wiley.

Lowes, K. Y. M. N., Croager, E. J., Olynyk, J. K., Abraham, L. J., & Yeoh, G. C. T. (2003). Oval cell-mediated liver regeneration: Role of cytokines and growth factors. *Journal of Gastroenterology and Hepatology*, *18*(1), 4–12. https://doi.org/https://doi.org/10.1046/j.1440-1746.2003.02906.x

Lugea, A., Nan, L., French, S. W., Bezerra, J. A., Gukovskaya, A. S., & Pandol, S. J. (2006). Pancreas Recovery Following Cerulein-Induced Pancreatitis Is Impaired in Plasminogen-Deficient Mice. *Gastroenterology*, *131*(3), 885–899. https://doi.org/https://doi.org/10.1053/j.gastro.2006.06.023

Lyons, T. R., O'Brien, J., Borges, V. F., Conklin, M. W., Keely, P. J., Eliceiri, K. W., Marusyk, A., Tan, A.-C., & Schedin, P. (2011). Postpartum mammary gland involution drives progression of ductal carcinoma in situ through collagen and COX-2. *Nature Medicine*, *17*(9), 1109–1115. https://doi.org/10.1038/nm.2416

Macias, H., & Hinck, L. (2012). Mammary gland development. *WIREs Developmental Biology*, *1*(4), 533–557. https://doi.org/https://doi.org/10.1002/wdev.35

Maciejowski, J., Chatzipli, A., Dananberg, A., Chu, K., Toufektchan, E., Klimczak, L. J., Gordenin, D. A., Campbell, P. J., & de Lange, T. (2020). APOBEC3-dependent kataegis and TREX1-driven chromothripsis during telomere crisis. *Nature Genetics*, *52*(9), 884–890. https://doi.org/10.1038/s41588-020-0667-5

Maciejowski, J., Li, Y., Bosco, N., Campbell, P. J., & de Lange, T. (2015). Chromothripsis and Kataegis Induced by Telomere Crisis. *Cell*, *163*(7), 1641–1654. https://doi.org/10.1016/j.cell.2015.11.054

Mackenzie, K. J., Carroll, P., Martin, C.-A., Murina, O., Fluteau, A., Simpson, D. J., Olova, N., Sutcliffe, H., Rainger, J. K., Leitch, A., Osborn, R. T., Wheeler, A. P., Nowotny, M., Gilbert, N., Chandra, T., Reijns, M. A. M., & Jackson, A. P. (2017). cGAS surveillance of micronuclei links genome instability to innate immunity. *Nature*, *548*(7668), 461–465. https://doi.org/10.1038/nature23449

Maggiore, G., & Zhu, H. (2024). Relationships Between Regeneration, Wound Healing, and Cancer. *Annual Review of Cancer Biology*, *8*(Volume 8, 2024), 177–197. https://doi.org/https://doi.org/10.1146/annurev-cancerbio-062822-123558

Maiato, H., & Logarinho, E. (2014). Mitotic spindle multipolarity without centrosome amplification. *Nature Cell Biology*, *16*(5), 386–394. https://doi.org/10.1038/ncb2958

Maitra, A., Fukushima, N., Takaori, K., & Hruban, R. H. (2005). Precursors to Invasive Pancreatic Cancer. *Advances in Anatomic Pathology*, *12*(2). https://journals.lww.com/anatomicpathology/fulltext/2005/03000/precursors_to_invasive_pancreatic_cancer.6.aspx

Maltsev, D. I., Mellanson, K. A., Belousov, V. V, Enikolopov, G. N., & Podgorny, O. V. (2022). The bioavailability time of commonly used thymidine analogues after intraperitoneal delivery in mice: labeling kinetics in vivo and clearance from blood serum. *Histochemistry and Cell Biology*, *157*(2), 239–250. https://doi.org/10.1007/s00418-021-02048-y

Margolis, R. L., Lohez, O. D., & Andreassen, P. R. (2003). G1 tetraploidy checkpoint and the suppression of tumorigenesis. *Journal of Cellular Biochemistry*, *88*(4), 673–683. https://doi.org/https://doi.org/10.1002/jcb.10411

Marstrand-Daucé, L., Lorenzo, D., Chassac, A., Nicole, P., Couvelard, A., & Haumaitre, C. (2023). Acinar-to-Ductal Metaplasia (ADM): On the Road to Pancreatic Intraepithelial Neoplasia (PanIN) and Pancreatic Cancer. *International Journal of Molecular Sciences*, *24*(12). https://doi.org/10.3390/ijms24129946

Matondo, R. B., Moreno, E., Toussaint, M. J. M., Tooten, P. C. J., van Essen, S. C., van Liere, E. A., Youssef, S. A., Bongiovanni, L., & de Bruin, A. (2018). Atypical E2f functions are critical for pancreas polyploidization. *PloS One*, *13*(1), e0190899.

Matsumoto, T., Wakefield, L., Peters, A., Peto, M., Spellman, P., & Grompe, M. (2021). Proliferative polyploid cells give rise to tumors via ploidy reduction. *Nature Communications*, *12*(1), 646. https://doi.org/10.1038/s41467-021-20916-y

Matsumoto, T., Wakefield, L., Tarlow, B. D., & Grompe, M. (2020). Lineage Tracing of Polyploid Hepatocytes Reveals Extensive Proliferation during Liver Regeneration. *Cell Stem Cell*, *26*(1), 34-47.e3. https://doi.org/10.1016/j.stem.2019.11.014

Matt, S., & Hofmann, T. G. (2016). The DNA damage-induced cell death response: a roadmap to kill cancer cells. *Cellular and Molecular Life Sciences*, *73*(15), 2829–2850. https://doi.org/10.1007/s00018-016-2130-4

Maya, R., Balass, M., Kim, S.-T., Shkedy, D., Leal, J.-F. M., Shifman, O., Moas, M., Buschmann, T., Ronai, Z., Shiloh, Y., Kastan, M. B., Katzir, E., & Oren, M. (2001). ATM-dependent phosphorylation of Mdm2 on serine 395: role in p53 activation by DNA damage. *Genes & Development*, *15*(9), 1067–1077. https://doi.org/10.1101/gad.886901

Mazzagatti, A., Engel, J. L., & Ly, P. (2024). Boveri and beyond: Chromothripsis and genomic instability from mitotic errors. *Molecular Cell*, *84*(1), 55–69. https://doi.org/10.1016/j.molcel.2023.11.002

Mazzi, S., Lordier, L., Debili, N., Raslova, H., & Vainchenker, W. (2018). Megakaryocyte and polyploidization. *Experimental Hematology*, *57*, 1–13. https://doi.org/10.1016/j.exphem.2017.10.001

McAinsh, A. D., & Kops, G. J. P. L. (2023). Principles and dynamics of spindle assembly checkpoint signalling. *Nature Reviews Molecular Cell Biology*, *24*(8), 543–559. https://doi.org/10.1038/s41580-023-00593-z

McClintock, B. (1941). THE STABILITY OF BROKEN ENDS OF CHROMOSOMES IN ZEA MAYS. *Genetics*, *26*(2), 234–282. https://doi.org/10.1093/genetics/26.2.234

McNally, S., & Stein, T. (2017). Overview of Mammary Gland Development: A Comparison of Mouse and Human. In F. Martin, T. Stein, & J. Howlin (Eds.), *Mammary Gland Development: Methods and Protocols* (pp. 1–17). Springer New York. https://doi.org/10.1007/978-1-4939-6475-8_1

Means, A. L., Meszoely, I. M., Suzuki, K., Miyamoto, Y., Rustgi, A. K., Coffey Jr, R. J., Wright, C. V. E., Stoffers, D. A., & Leach, S. D. (2005). Pancreatic epithelial plasticity mediated by acinar

cell transdifferentiation and generation of nestin-positive intermediates. *Development*, *132*(16), 3767–3776. https://doi.org/10.1242/dev.01925

Middelkamp, S., van Heesch, S., Braat, A. K., de Ligt, J., van Iterson, M., Simonis, M., van Roosmalen, M. J., Kelder, M. J. E., Kruisselbrink, E., Hochstenbach, R., Verbeek, N. E., Ippel, E. F., Adolfs, Y., Pasterkamp, R. J., Kloosterman, W. P., Kuijk, E. W., & Cuppen, E. (2017). Molecular dissection of germline chromothripsis in a developmental context using patient-derived iPS cells. *Genome Medicine*, *9*(1), 9. https://doi.org/10.1186/s13073-017-0399-z

Middelkoop, T. C., Neipel, J., Cornell, C. E., Naumann, R., Pimpale, L. G., Jülicher, F., & Grill, S. W. (2024). A cytokinetic ring-driven cell rotation achieves Hertwig's rule in early development. *Proceedings of the National Academy of Sciences*, *121*(25), e2318838121. https://doi.org/10.1073/pnas.2318838121

Miyaoka, Y., Ebato, K., Kato, H., Arakawa, S., Shimizu, S., & Miyajima, A. (2012). Hypertrophy and Unconventional Cell Division of Hepatocytes Underlie Liver Regeneration. *Current Biology*, *22*(13), 1166–1175. https://doi.org/10.1016/j.cub.2012.05.016

Molinuevo, R., Menendez, J., Cadle, K., Ariqat, N., Choy, M. K., Lagousis, C., Thomas, G., Strietzel, C., Bubolz, J. W., & Hinck, L. (2024). Physiological DNA damage promotes functional endoreplication of mammary gland alveolar cells during lactation. *Nature Communications*, *15*(1), 3288. https://doi.org/10.1038/s41467-024-47668-9

Montesano, R. (1981). Alkylation of DNA and tissue specificity in nitrosamine carcinogenesis. *Journal of Supramolecular Structure and Cellular Biochemistry*, *17*(3), 259–273. https://doi.org/https://doi.org/10.1002/jsscb.380170307

Moore, K. A., & Lemischka, I. R. (2006). Stem Cells and Their Niches. *Science*, *311*(5769), 1880–1885. https://doi.org/10.1126/science.1110542

Morton, J. P., Timpson, P., Karim, S. A., Ridgway, R. A., Athineos, D., Doyle, B., Jamieson, N. B., Oien, K. A., Lowy, A. M., Brunton, V. G., Frame, M. C., Evans, T. R. J., & Sansom, O. J. (2010). Mutant p53 drives metastasis and overcomes growth arrest/senescence in pancreatic cancer. *Proceedings of the National Academy of Sciences*, *107*(1), 246–251. https://doi.org/10.1073/pnas.0908428107

Motta, P. M., Macchiarelli, G., Nottola, S. A., & Correr, S. (1997). Histology of the exocrine pancreas. *Microscopy Research and Technique*, *37*(5–6), 384–398. https://doi.org/https://doi.org/10.1002/(SICI)1097-0029(19970601)37:5/6<384::AID-JEMT3>3.0.CO;2-E

Murphy, L. J., Murphy, L. C., Vrhovsek, E., Sutherland, R. L., & Lazarus, L. (1984). Correlation of Lactogenic Receptor Concentration in Human Breast Cancer with Estrogen Receptor Concentration. *Cancer Research*, *44*(5), 1963–1968.

Nair, V., & Hinton, G. E. (2010). Rectified linear units improve restricted boltzmann machines. *Proceedings of the 27th International Conference on International Conference on Machine Learning*, 807–814.

Neville, M. C., McFadden, T. B., & Forsyth, I. (2002). Hormonal Regulation of Mammary Differentiation and Milk Secretion. *Journal of Mammary Gland Biology and Neoplasia*, *7*(1), 49–66. https://doi.org/10.1023/A:1015770423167

Nichols, H. B., Schoemaker, M. J., Cai, J., Xu, J., Wright, L. B., Brook, M. N., Jones, M. E., Adami, H.-O., Baglietto, L., Bertrand, K. A., Blot, W. J., Boutron-Ruault, M.-C., Dorronsoro, M., Dossus, L., Eliassen, A. H., Giles, G. G., Gram, I. T., Hankinson, S. E., Hoffman-Bolton, J., … Sandler, D. P. (2018). Breast Cancer Risk After Recent Childbirth. *Annals of Internal Medicine*, *170*(1), 22–30. https://doi.org/10.7326/M18-1323

Nickerson, K., Bonsnes, R. W., Douglas, R. G., Condliffe, P., & du Vigneaud, V. (1954). Oxytocin and milk ejection. *American Journal of Obstetrics & Gynecology*, *67*(5), 1028–1034. https://doi.org/10.1016/0002-9378(54)90261-6

Notta, F., Chan-Seng-Yue, M., Lemire, M., Li, Y., Wilson, G. W., Connor, A. A., Denroche, R. E., Liang, S.-B., Brown, A. M. K., Kim, J. C., Wang, T., Simpson, J. T., Beck, T., Borgida, A., Buchner, N., Chadwick, D., Hafezi-Bakhtiari, S., Dick, J. E., Heisler, L., … Gallinger, S. (2016). A renewed model of pancreatic cancer evolution based on genomic rearrangement patterns. *Nature*, *538*(7625), 378–382. https://doi.org/10.1038/nature19823

Oakes, S. R., Hilton, H. N., & Ormandy, C. J. (2006). Key stages in mammary gland development - The alveolar switch: coordinating the proliferative cues and cell fate decisions that drive the formation of lobuloalveoli from ductal epithelium. *Breast Cancer Research*, *8*(2), 207. https://doi.org/10.1186/bcr1411

Orkin, S. H., & Zon, L. I. (2008). Hematopoiesis: An Evolving Paradigm for Stem Cell Biology. *Cell*, *132*(4), 631–644. https://doi.org/10.1016/j.cell.2008.01.025

Otsu, N. (1979). A Threshold Selection Method from Gray-Level Histograms. *IEEE Transactions on Systems, Man, and Cybernetics*, *9*(1), 62–66. https://doi.org/10.1109/TSMC.1979.4310076

Pachitariu, M., & Stringer, C. (2022). Cellpose 2.0: how to train your own model. *Nature Methods*, *19*(12), 1634–1641. https://doi.org/10.1038/s41592-022-01663-4

Pandit, S. K., Westendorp, B., Nantasanti, S., van Liere, E., Tooten, P. C. J., Cornelissen, P. W. A., Toussaint, M. J. M., Lamers, W. H., & de Bruin, A. (2012). E2F8 is essential for polyploidization in mammalian cells. *Nature Cell Biology*, *14*(11), 1181–1191. https://doi.org/10.1038/ncb2585

Parte, S., Nimmakayala, R. K., Batra, S. K., & Ponnusamy, M. P. (2022). Acinar to ductal cell trans-differentiation: A prelude to dysplasia and pancreatic ductal adenocarcinoma. *Biochimica et Biophysica Acta (BBA) - Reviews on Cancer*, *1877*(1), 188669. https://doi.org/https://doi.org/10.1016/j.bbcan.2021.188669

Pathare, G. R., Decout, A., Glück, S., Cavadini, S., Makasheva, K., Hovius, R., Kempf, G., Weiss, J., Kozicka, Z., Guey, B., Melenec, P., Fierz, B., Thomä, N. H., & Ablasser, A. (2020). Structural mechanism of cGAS inhibition by the nucleosome. *Nature*, *587*(7835), 668–672. https://doi.org/10.1038/s41586-020-2750-6

Perou, C. M., Sørlie, T., Eisen, M. B., van de Rijn, M., Jeffrey, S. S., Rees, C. A., Pollack, J. R., Ross, D. T., Johnsen, H., Akslen, L. A., Fluge, Ø., Pergamenschikov, A., Williams, C., Zhu, S. X., Lønning, P. E., Børresen-Dale, A.-L., Brown, P. O., & Botstein, D. (2000). Molecular portraits of human breast tumours. *Nature*, *406*(6797), 747–752. https://doi.org/10.1038/35021093

Pian, L., Song, M., Wang, T., Qi, L., Peng, T., & Xie, K. (2025). Identification and analysis of pancreatic intraepithelial neoplasia: opportunities and challenges. *Frontiers in Endocrinology*, *Volume 15-2024*. https://doi.org/10.3389/fendo.2024.1401829

Pidoux, A. L., Uzawa, S., Perry, P. E., Cande, W. Z., & Allshire, R. C. (2000). Live analysis of lagging chromosomes during anaphase and their effect on spindle elongation rate in fission yeast. *Journal of Cell Science*, *113*(23), 4177–4191. https://doi.org/10.1242/jcs.113.23.4177

Pinho, A. V, Rooman, I., Reichert, M., De Medts, N., Bouwens, L., Rustgi, A. K., & Real, F. X. (2011). Adult pancreatic acinar cells dedifferentiate to an embryonic progenitor phenotype with concomitant activation of a senescence programme that is present in chronic pancreatitis. *Gut*, *60*(7), 958. https://doi.org/10.1136/gut.2010.225920

Quan, S., Principe, D. R., Dean, A. E., Park, S.-H., Grippo, P. J., Gius, D., & Horikoshi, N. (2018). Loss of Sirt2 increases and prolongs a caerulein-induced pancreatitis permissive phenotype and induces spontaneous oncogenic Kras mutations in mice. *Scientific Reports*, *8*(1), 16501. https://doi.org/10.1038/s41598-018-34792-y

Radisky, D. C., & Hartmann, L. C. (2009). Mammary Involution and Breast Cancer Risk: Transgenic Models and Clinical Studies. *Journal of Mammary Gland Biology and Neoplasia*, *14*(2), 181–191. https://doi.org/10.1007/s10911-009-9123-y

Rios, A. C., Fu, N. Y., Jamieson, P. R., Pal, B., Whitehead, L., Nicholas, K. R., Lindeman, G. J., & Visvader, J. E. (2016). Essential role for a novel population of binucleated mammary epithelial cells in lactation. *Nature Communications*, *7*(1), 11400. https://doi.org/10.1038/ncomms11400

Rivetti, S., Chen, C., Chen, C., & Bellusci, S. (2020). Fgf10/Fgfr2b Signaling in Mammary Gland Development, Homeostasis, and Cancer. *Frontiers in Cell and Developmental Biology*, *Volume 8-2020*. https://doi.org/10.3389/fcell.2020.00415

Rizzotto, D., Vigorito, V., Rieder, P., Gallob, F., Moretta, G. M., Soratroi, C., Riley, J. S., Bellutti, F., Veli, S. L., Mattivi, A., Lohmüller, M., Herzog, S., Bornhauser, B. C., Jacotot, E. D., Villunger, A., & Fava, L. L. (2024). Caspase-2 kills cells with extra centrosomes. *Science Advances*, *10*(44), eado6607. https://doi.org/10.1126/sciadv.ado6607

Rodrigues, M. A., Probst, C. E., Zayats, A., Davidson, B., Riedel, M., Li, Y., & Venkatachalam, V. (2021). The in vitro micronucleus assay using imaging flow cytometry and deep learning. *Npj Systems Biology and Applications*, *7*(1), 20. https://doi.org/10.1038/s41540-021-00179-5

Ronneberger, O., Fischer, P., & Brox, T. (2015). U-Net: Convolutional Networks for Biomedical Image Segmentation. In N. Navab, J. Hornegger, W. M. Wells, & A. F. Frangi (Eds.), *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015* (pp. 234–241). Springer International Publishing.

Rosenblatt, F. (1958). The perceptron: A probabilistic model for information storage and organization in the brain. *Psychological Review*, *65*(6), 386–408. https://doi.org/10.1037/h0042519

Roy, S., & Gatien, S. (2008). Regeneration in axolotls: a model to aim for! *Experimental Gerontology*, *43*(11), 968–973. https://doi.org/https://doi.org/10.1016/j.exger.2008.09.003

Rumelhart, D. E., Hinton, G. E., & Williams, R. J. (1986). Learning representations by back-propagating errors. *Nature*, *323*(6088), 533–536. https://doi.org/10.1038/323533a0

Ryan, D. P., Hong, T. S., & Bardeesy, N. (2014). Pancreatic Adenocarcinoma. *New England Journal of Medicine*, *371*(11), 1039–1049. https://doi.org/10.1056/NEJMra1404198

Sanders, A. D., Falconer, E., Hills, M., Spierings, D. C. J., & Lansdorp, P. M. (2017). Single-cell template strand sequencing by Strand-seq enables the characterization of individual homologs. *Nature Protocols*, *12*(6), 1151–1176. https://doi.org/10.1038/nprot.2017.029

Sargeant, T. J., Lloyd-Lewis, B., Resemann, H. K., Ramos-Montoya, A., Skepper, J., & Watson, C. J. (2014). Stat3 controls cell death during mammary gland involution by regulating uptake of milk fat globules and lysosomal membrane permeabilization. *Nature Cell Biology*, *16*(11), 1057–1068. https://doi.org/10.1038/ncb3043

Schedin, P. (2006). Pregnancy-associated breast cancer and metastasis. *Nature Reviews Cancer*, *6*(4), 281–291. https://doi.org/10.1038/nrc1839

Schedin, Q. G. A. N. D. J. M. A. N. D. J. B. A. N. D. K. C. A. N. D. P. S. A. N. D. P. (2017). Physiologically activated mammary fibroblasts promote postpartum mammary cancer. *JCI Insight*, *2*(6). https://doi.org/10.1172/jci.insight.89206

Schmidt, U., Weigert, M., Broaddus, C., & Myers, G. (2018). Cell Detection with Star-Convex Polygons. In *Medical Image Computing and Computer Assisted Intervention – MICCAI 2018* (pp. 265–273). Springer International Publishing. https://doi.org/10.1007/978-3-030-00934-2_30

Sekiya, S., & Suzuki, A. (2014). Hepatocytes, Rather than Cholangiocytes, Can Be the Major Source of Primitive Ductules in the Chronically Injured Mouse Liver. *The American Journal of Pathology*, *184*(5), 1468–1478. https://doi.org/https://doi.org/10.1016/j.ajpath.2014.01.005

Serrano, M., Lin, A. W., McCurrach, M. E., Beach, D., & Lowe, S. W. (1997). Oncogenic ras Provokes Premature Cell Senescence Associated with Accumulation of p53 and p16INK4a. *Cell*, *88*(5), 593–602. https://doi.org/https://doi.org/10.1016/S0092-8674(00)81902-9

Siegel, R. L., Miller, K. D., Fuchs, H. E., & Jemal, A. (2022). Cancer statistics, 2022. *CA: A Cancer Journal for Clinicians*, *72*(1), 7–33. https://doi.org/https://doi.org/10.3322/caac.21708

Simonyan, K., & Zisserman, A. (2015). *Very Deep Convolutional Networks for Large-Scale Image Recognition*. https://arxiv.org/abs/1409.1556

Simovic-Lorenz, M., & Ernst, A. (2025). Chromothripsis in cancer. *Nature Reviews Cancer*, *25*(2), 79–92. https://doi.org/10.1038/s41568-024-00769-5

Sladky, V. C., Knapp, K., Soratroi, C., Heppke, J., Eichin, F., Rocamora-Reverte, L., Szabo, T. G., Bongiovanni, L., Westendorp, B., Moreno, E., van Liere, E. A., Bakker, B., Spierings, D. C. J., Wardenaar, R., Pereyra, D., Starlinger, P., Schultze, S., Trauner, M., Stojakovic, T., … Villunger, A. (2020). E2F-Family Members Engage the PIDDosome to Limit Hepatocyte Ploidy in Liver Development and Regeneration. *Developmental Cell*, *52*(3), 335-349.e7. https://doi.org/10.1016/j.devcel.2019.12.016

Sobel, I., & Feldman, G. (1973). A 3×3 isotropic gradient operator for image processing. *Pattern Classification and Scene Analysis*, 271–272.

Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., & Salakhutdinov, R. (2014). Dropout: A Simple Way to Prevent Neural Networks from Overfitting. *Journal of Machine Learning Research*, *15*(56), 1929–1958. http://jmlr.org/papers/v15/srivastava14a.html

Stephens, P. J., Greenman, C. D., Fu, B., Yang, F., Bignell, G. R., Mudie, L. J., Pleasance, E. D., Lau, K. W., Beare, D., Stebbings, L. A., McLaren, S., Lin, M.-L., McBride, D. J., Varela, I., Nik-Zainal, S., Leroy, C., Jia, M., Menzies, A., Butler, A. P., … Campbell, P. J. (2011). Massive Genomic Rearrangement Acquired in a Single Catastrophic Event during Cancer Development. *Cell*, *144*(1), 27–40. https://doi.org/10.1016/j.cell.2010.11.055

Stingl, J., Eirew, P., Ricketson, I., Shackleton, M., Vaillant, F., Choi, D., Li, H. I., & Eaves, C. J. (2006). Purification and unique properties of mammary epithelial stem cells. *Nature*, *439*(7079), 993–997. https://doi.org/10.1038/nature04496

Storchova, Z., & Pellman, D. (2004). From polyploidy to aneuploidy, genome instability and cancer. *Nature Reviews Molecular Cell Biology*, *5*(1), 45–54. https://doi.org/10.1038/nrm1276

Storz, P. (2017). Acinar cell plasticity and development of pancreatic ductal adenocarcinoma. *Nature Reviews Gastroenterology & Hepatology*, *14*(5), 296–304. https://doi.org/10.1038/nrgastro.2017.12

Stringer, C., Wang, T., Michaelos, M., & Pachitariu, M. (2021). Cellpose: a generalist algorithm for cellular segmentation. *Nature Methods*, *18*(1), 100–106. https://doi.org/10.1038/s41592-020-01018-x

Strnad, P., Gunther, S., Reichmann, J., Krzic, U., Balazs, B., de Medeiros, G., Norlin, N., Hiiragi, T., Hufnagel, L., & Ellenberg, J. (2016). Inverted light-sheet microscope for imaging mouse pre-implantation development. *Nature Methods*, *13*(2), 139–142. https://doi.org/10.1038/nmeth.3690

Sumbal, J., Chiche, A., Charifou, E., Koledova, Z., & Li, H. (2020). Primary Mammary Organoid Model of Lactation and Involution. *Frontiers in Cell and Developmental Biology*, *8*. https://doi.org/10.3389/fcell.2020.00068

Sundaram, M., Duane L., G., Murali M., R., & and Rajaraman, R. (2004). Neosis: A Novel Type of Cell Division in Cancer. *Cancer Biology & Therapy*, *3*(2), 207–218. https://doi.org/10.4161/cbt.3.2.663

Svartman, M., Stone, G., & Stanyon, R. (2005). Molecular cytogenetics discards polyploidy in mammals. *Genomics*, *85*(4), 425–430. https://doi.org/https://doi.org/10.1016/j.ygeno.2004.12.004

Takaki, T., Millar, R., Hiley, C. T., & Boulton, S. J. (2024). Micronuclei induced by radiation, replication stress, or chromosome segregation errors do not activate cGAS-STING. *Molecular Cell*, *84*(11), 2203-2213.e5. https://doi.org/10.1016/j.molcel.2024.04.017

Tan, M., & Le, Q. V. (2020). *EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks*. https://arxiv.org/abs/1905.11946

Tanaka, J., Takamatsu, K., Yukimori, A., Kujiraoka, S., Ishida, S., Takakura, I., Yasuhara, R., & Mishima, K. (2021). Sox9 function in salivary gland development. *Journal of Oral Biosciences*, *63*(1), 8–13. https://doi.org/https://doi.org/10.1016/j.job.2021.01.005

Terradas, M., Martín, M., Tusell, L., & Genescà, A. (2010). Genetic activities in micronuclei: Is the DNA entrapped in micronuclei lost for the cell? *Mutation Research/Reviews in Mutation Research*, *705*(1), 60–67. https://doi.org/https://doi.org/10.1016/j.mrrev.2010.03.004

Théry, M., Racine, V., Pépin, A., Piel, M., Chen, Y., Sibarita, J.-B., & Bornens, M. (2005). The extracellular matrix guides the orientation of the cell division axis. *Nature Cell Biology*, *7*(10), 947–953. https://doi.org/10.1038/ncb1307

Thompson, S. L., & Compton, D. A. (2008). Examining the link between chromosomal instability and aneuploidy in human cells. *Journal of Cell Biology*, *180*(4), 665–672. https://doi.org/10.1083/jcb.200712029

Thompson, S. L., & Compton, D. A. (2011). Chromosome missegregation in human cells arises through specific types of kinetochore–microtubule attachment errors. *Proceedings of the National Academy of Sciences*, *108*(44), 17974–17978. https://doi.org/10.1073/pnas.1109720108

Tollis, M., Boddy, A. M., & Maley, C. C. (2017). Peto's Paradox: how has evolution solved the problem of cancer prevention? *BMC Biology*, *15*(1), 60. https://doi.org/10.1186/s12915-017-0401-7

Tosti, L., Hang, Y., Debnath, O., Tiesmeyer, S., Trefzer, T., Steiger, K., Ten, F. W., Lukassen, S., Ballke, S., Kühl, A. A., Spieckermann, S., Bottino, R., Ishaque, N., Weichert, W., Kim, S. K., Eils, R., & Conrad, C. (2021). Single-Nucleus and In Situ RNA–Sequencing Reveal Cell Topographies in the Human Pancreas. *Gastroenterology*, *160*(4), 1330-1344.e11. https://doi.org/https://doi.org/10.1053/j.gastro.2020.11.010

Tumbar, T., Guasch, G., Greco, V., Blanpain, C., Lowry, W. E., Rendl, M., & Fuchs, E. (2004). Defining the Epithelial Stem Cell Niche in Skin. *Science*, *303*(5656), 359–363. https://doi.org/10.1126/science.1092436

Van de Peer, Y., Mizrachi, E., & Marchal, K. (2017). The evolutionary significance of polyploidy. *Nature Reviews Genetics*, *18*(7), 411–424. https://doi.org/10.1038/nrg.2017.26

Varga, J., & Greten, F. R. (2017). Cell plasticity in epithelial homeostasis and tumorigenesis. *Nature Cell Biology*, *19*(10), 1133–1141. https://doi.org/10.1038/ncb3611

Vieira, W. A., Wells, K. M., & McCusker, C. D. (2019). Advancements to the Axolotl Model for Regeneration and Aging. *Gerontology*, *66*(3), 212–222. https://doi.org/10.1159/000504294

Virchow, R. (1858). *Die Cellularpathologie in ihrer Begründung auf physiologische und pathologische Gewebelehre: Zwanzig Vorlesungen gehalten während der Monate Februar, März und April 1858 im pathologischen Institute zu Berlin*. Verlag von August Hirschwald.

Visvader, J. E., & Stingl, J. (2014). Mammary stem cells and the differentiation hierarchy: current status and perspectives. *Genes & Development*, *28*(11), 1143–1158. https://doi.org/10.1101/gad.242511.114

Voronina, N., Wong, J. K. L., Hübschmann, D., Hlevnjak, M., Uhrig, S., Heilig, C. E., Horak, P., Kreutzfeldt, S., Mock, A., Stenzinger, A., Hutter, B., Fröhlich, M., Brors, B., Jahn, A., Klink, B., Gieldon, L., Sieverling, L., Feuerbach, L., Chudasama, P., … Ernst, A. (2020). The landscape of chromothripsis across adult cancer types. *Nature Communications*, *11*(1), 2320. https://doi.org/10.1038/s41467-020-16134-7

Wan, J., Shin, D. B., Syed, M. N., Abuabara, K., Lemeshow, A. R., Fuxench, Z. C. C., & Gelfand, J. M. (2023). Malignancy risk in patients with atopic dermatitis: a population-based cohort study. *British Journal of Dermatology*, *189*(1), 53–61. https://doi.org/10.1093/bjd/ljad072

Wang, X., Willenbring, H., Akkari, Y., Torimaru, Y., Foster, M., Al-Dhalimy, M., Lagasse, E., Finegold, M., Olson, S., & Grompe, M. (2003). Cell fusion is the principal source of bone-marrow-derived hepatocytes. *Nature*, *422*(6934), 897–901. https://doi.org/10.1038/nature01531

Watson, C. J. (2006). Key stages in mammary gland development - Involution: apoptosis and tissue remodelling that convert the mammary gland from milk factory to a quiescent organ. *Breast Cancer Research*, *8*(2), 203. https://doi.org/10.1186/bcr1401

Watson, C. J., & Khaled, W. T. (2008). Mammary development in the embryo and adult: a journey of morphogenesis and commitment. *Development*, *135*(6), 995–1003. https://doi.org/10.1242/dev.005439

Weber, T., Cosenza, M. R., & Korbel, J. (2023). MosaiCatcher v2: a single-cell structural variations detection and analysis reference framework based on Strand-seq. *Bioinformatics*, *39*(11), btad633. https://doi.org/10.1093/bioinformatics/btad633

Westphalen, C. B., Takemoto, Y., Tanaka, T., Macchini, M., Jiang, Z., Renz, B. W., Chen, X., Ormanns, S., Nagar, K., Tailor, Y., May, R., Cho, Y., Asfaha, S., Worthley, D. L., Hayakawa, Y., Urbanska, A. M., Quante, M., Reichert, M., Broyde, J., … Wang, T. C. (2016). Dclk1 Defines Quiescent Pancreatic Progenitors that Promote Injury-Induced Regeneration and Tumorigenesis. *Cell Stem Cell*, *18*(4), 441–455. https://doi.org/10.1016/j.stem.2016.03.016

Wilentz, R. E., Geradts, J., Maynard, R., Offerhaus, G. J. A., Kang, M., Goggins, M., Yeo, C. J., Kern, S. E., & Hruban, R. H. (1998). Inactivation of the p16 (INK4A) tumor-suppressor gene in pancreatic duct lesions: Loss of intranuclear expression. *Cancer Research*, *58*(20), 4740–4744.

Wilkinson, P. D., Delgado, E. R., Alencastro, F., Leek, M. P., Roy, N., Weirich, M. P., Stahl, E. C., Otero, P. A., Chen, M. I., Brown, W. K., & Duncan, A. W. (2019). The Polyploid State Restricts Hepatocyte Proliferation and Liver Regeneration in Mice. *Hepatology (Baltimore, Md.)*, *69*(3), 1242–1258. https://doi.org/10.1002/hep.30286

Winding, M., Pedigo, B. D., Barnes, C. L., Patsolic, H. G., Park, Y., Kazimiers, T., Fushiki, A., Andrade, I. V, Khandelwal, A., Valdes-Aleman, J., Li, F., Randel, N., Barsotti, E., Correia, A., Fetter, R. D., Hartenstein, V., Priebe, C. E., Vogelstein, J. T., Cardona, A., & Zlatic, M. (2025). The connectome of an insect brain. *Science*, *379*(6636), eadd9330. https://doi.org/10.1126/science.add9330

Windmueller, R., Leach, J. P., Babu, A., Zhou, S., Morley, M. P., Wakabayashi, A., Petrenko, N. B., Viatour, P., & Morrisey, E. E. (2020). Direct Comparison of Mononucleated and Binucleated Cardiomyocytes Reveals Molecular Mechanisms Underlying Distinct Proliferative Competencies. *Cell Reports*, *30*(9), 3105-3116.e4. https://doi.org/10.1016/j.celrep.2020.02.034

Wohlschlegel, J. A., Kutok, J. L., Weng, A. P., & Dutta, A. (2002). Expression of Geminin as a Marker of Cell Proliferation in Normal Tissues and Malignancies. *The American Journal of Pathology*, *161*(1), 267–273. https://doi.org/https://doi.org/10.1016/S0002-9440(10)64178-8

Wollny, D., Zhao, S., Everlien, I., Lun, X., Brunken, J., Brüne, D., Ziebell, F., Tabansky, I., Weichert, W., Marciniak-Czochra, A., & Martin-Villalba, A. (2016). Single-Cell Analysis Uncovers Clonal Acinar Cell Heterogeneity in the Adult Pancreas. *Developmental Cell*, *39*(3), 289–301. https://doi.org/10.1016/j.devcel.2016.10.002

Wu, Z., Zhang, W., Wang, L., Fan, L., Li, J., & Jin, H. (2023). Extrachromosomal Circular DNA Drives the Progression of DLBCL through Activating the Sting Pathway. *Blood*, *142*(Supplement 1), 1628. https://doi.org/10.1182/blood-2023-190771

Wu, Z., Zhang, W., Wang, L., Leng, J., Li, Y., Fan, Z., Zhan, M., Cao, L., Jiang, Y., Jiang, Y., Sun, B., Fu, J., Li, J., Shi, W., & Jin, H. (2024). Multi-omics integration reveals the oncogenic role of eccDNAs in diffuse large B-cell lymphoma through STING signalling. *Clinical and Translational Medicine*, *14*(8), e1815. https://doi.org/https://doi.org/10.1002/ctm2.1815

Xu, G., Wang, X., Wu, X., Leng, X., & Xu, Y. (2024). *Development of Skip Connection in Deep Neural Networks for Computer Vision and Medical Image Analysis: A Survey*. https://arxiv.org/abs/2405.01725

Yang, B., & Chen, S. (2013). A comparative study on local binary pattern (LBP) based face recognition: LBP histogram versus LBP image. *Neurocomputing*, *120*, 365–379. https://doi.org/https://doi.org/10.1016/j.neucom.2012.10.032

Yao, Y., Zhang, Q., Li, Z., & Zhang, H. (2024). MDM2: current research status and prospects of tumor treatment. *Cancer Cell International*, *24*(1), 170. https://doi.org/10.1186/s12935-024-03356-8

Yücel, D., Solinsky, J., & van Berlo, J. H. (2020). Isolation of Cardiomyocytes from Fixed Hearts for Immunocytochemistry and Ploidy Analysis. *JoVE*, *164*, e60938. https://doi.org/doi:10.3791/60938

Zhang, C.-Z., Leibowitz, M. L., & Pellman, D. (2013). Chromothripsis and beyond: rapid genome evolution from complex chromosomal rearrangements. *Genes & Development*, *27*(23), 2513–2530. https://doi.org/10.1101/gad.229559.113

Zhang, C.-Z., Spektor, A., Cornils, H., Francis, J. M., Jackson, E. K., Liu, S., Meyerson, M., & Pellman, D. (2015). Chromothripsis from DNA damage in micronuclei. *Nature*, *522*(7555), 179–184. https://doi.org/10.1038/nature14493

Zhang, S., Zhou, K., Luo, X., Li, L., Tu, H.-C., Sehgal, A., Nguyen, L. H., Zhang, Y., Gopal, P., Tarlow, B. D., Siegwart, D. J., & Zhu, H. (2018). The Polyploid State Plays a Tumor-Suppressive Role in the Liver. *Developmental Cell*, *44*(4), 447-459.e5. https://doi.org/10.1016/j.devcel.2018.01.010

Zhuang, X., Wang, Q., Joost, S., Ferrena, A., Humphreys, D. T., Li, Z., Blum, M., Krause, K., Ding, S., Landais, Y., Zhan, Y., Zhao, Y., Chaligne, R., Lee, J.-H., Carrasco, S. E., Bhanot, U. K., Koche, R. P., Bott, M. J., Katajisto, P., … Tammela, T. (2025). Ageing limits stemness and tumorigenesis by reprogramming iron homeostasis. *Nature*, *637*(8044), 184–194. https://doi.org/10.1038/s41586-024-08285-0