The Microenvironmental Influence on the Fate and Function of Breast Stem Cells

A thesis submitted by
Colin Trepicchio
in partial fulfillment of the requirements for the degree of PhD in
Cell, Molecular, and Developmental Biology
Tufts University
Graduate School of Biomedical Sciences

April 2nd, 2024

Charlotte Kuperwasser, PhD

Abstract:

The human breast can undergo cyclical development and involution over a woman's lifespan from puberty to menopause. This regenerative ability suggests the presence of stem cells capable of reproducing the necessary structures and cell types for tissue function. However, a lack of suitable models has hindered understanding of this epithelial regrowth and its signaling mechanisms. In this study, I enhanced a novel hydrogel model through the incorporation of single primary breast cells to form ductal-lobular organoids, enabling fresh developmental insights and diverse scientific applications. Using this model to observe both morphological and expression-based alterations, I uncovered a novel signaling pathway involving the receptor tyrosine kinase DDR1 and the transcription factor RUNX1.

Dedication:

To my sons:

Henry, always stay curious.

Fox, always stay smiling.

Love you both. Always.

Acknowledgements:
To Charlotte, thank you for giving me the freedom to make mistakes and grow. I am now confident I can make my way through any challenges ahead.

To Gat, thank you for pushing me to always be a better scientist. Your mentorship has allowed be to become one.

To my Thesis Advisory Committee: Drs. Madeleine Oudin, Karl Munger, and Phil Haydon. Thank you for the patience during this growth. I know now that this is the true beginning of my scientific journey.

To my lab mates: Nicole Traugh, Youssof Mal, Dan Fein, Meadow Parrish, and Hahyung Kim. Thank you for always being there to celebrate every success and dust me off after the many failures. There is no group I would have rather been with during these years.

To my friends: Matt, Kevin, Johnson, Alan, Becky, Gab, Kyle and Tanner. Thank you for checking in on me when I was lost in this process. You pulled me out of the weeds many times and gave me the break needed to go back in and be successful.

To my dogs. Opie and Siggy, thanks for keeping me company as I wrote this. Penny, thank you for getting me to this point, we did it girl! I miss you every day.

To Anya. Thank you for being our children's third parent. I can only do what I do because I never worry about them. They are in the best hands.

To the Disstons. Thank you for everything from loving and caring for our children while I write, to fixing the whole house before it falls apart. I could not have gotten this far without your help.

To Andrew and Nicola. Andrew, thank you for always pushing me forward in life. I am so glad we got to be best friends in life. I hope mine have the same bond. Nicola, thank you for bringing a continued love for learning and games to the family. It has been more fun since you joined. Love you both.

To my family. Dad, thank you for the guidance and the outside eyes on everything. I am proud to know this work will sit next to yours. Mom, thank you for helping me stay true to who I am, and the meatballs. I never give up because of you.

Natalie, no words can explain how much I love you. Thank you for being the one I will always know, and my rock when everything is spinning around. We can do anything together.

Table of Contents:

List of Tables:

List of Figures:

List of Copyrighted Materials:

Beghini A. Core Binding Factor Leukemia: Chromatin Remodeling Moves Towards Oncogenic Transcription. *Cancers*. 2019; 11(12):1973. https://doi.org/10.3390/cancers11121973

Goyama, S., Huang, G., Kurokawa, M. *et al.* Posttranslational modifications of RUNX1 as potential anticancer targets. *Oncogene* **34**, 3483–3492 (2015). https://doi-org.ezproxy.library.tufts.edu/10.1038/onc.2014.305

Ichikawa, M., Yoshimi, A., Nakagawa, M. *et al.* A role for RUNX1 in hematopoiesis and myeloid leukemia. *Int J Hematol* **97**, 726–734 (2013). https://doi.org/10.1007/s12185-013-1347-3

Leitinger B. Discoidin domain receptor functions in physiological and pathological conditions. Int Rev Cell Mol Biol. 2014;310:39-87. doi: 10.1016/B978-0-12-800180-6.00002-5. PMID: 24725424; PMCID: PMC4021107.

*Oftedal, O.T., Dhouailly, D. Evo-Devo of the Mammary Gland. J Mammary Gland Biol Neoplasia **18**, 105–120 (2013). https://doi-org.ezproxy.library.tufts.edu/10.1007/s10911-013-9290-8*

*Rauner, G., Jin, D.X., Miller, D.H. et al. Breast tissue regeneration is driven by cell-matrix interactions coordinating multi-lineage stem cell differentiation through DDR1. Nat Commun **12**, 7116 (2021). https://doi.org/10.1038/s41467-021-27401-611*

Sokol ES, Miller DH, Breggia A, Spencer KC, Arendt LM, Gupta PB. Growth of human breast tissues from patient cells in 3D hydrogel scaffolds. Breast Cancer Res. 2016 Mar 1;18(1):19. doi: 10.1186/s13058-016-0677-5. PMID: 26926363; PMCID: PMC4772689

Sokol ES, Sanduja S, Jin DX, Miller DH, Mathis RA, Gupta PB. Perturbation-expression analysis identifies RUNX1 as a regulator of human mammary stem cell differentiation. PLoS Comput Biol. 2015 Apr 20;11(4):e1004161. doi: 10.1371/journal.pcbi.1004161. PMID: 25894653; PMCID: PMC4404314.

List of Abbreviations:

αSMA - Alpha Smooth Muscle Actin
2D - Two Dimensional
3D - Three Dimensional
AML - Acute Myeloid Leukemia
Bgb - Big Brother
BLM - Bloom Syndrome Helicase
BMI - Body Mass Index
BRCA1 - Breast Cancer Gene 1
Bro - Brother
CBF - Core Binding Factor
CD49f - Integrin Alpha 6
CK14 - Cytokeratin 14
CK8/18 - Cytokeratin 8/18
CNS - Central Nervous System
DAPI - 4′,6-diamidino-2-phenylindole (Live Cell Stain)
DBD - DNA Binding Domain
DDR1 - Discoidin Domain Receptor 1
DDR1i - Discoidin Domain Receptor 1 Inhibition
DDR1r - DDR1 Inhibitor Release
DNA - Deoxyribonucleic Acid
DSB - Double Strand Break
ECAD - E-Cadherin
ECM - Extra Cellular Matrix
EGFR - Epidermal Growth Factor Receptor
EMT - Epithelial to Mesenchymal Transition
EpCAM - Epithelal Cellular Adhesion Molecule
ER - Estrogen Receptor
ESR1 - Estrogen Receptor 1
HDAC - Histone Deacetylases
HER2 - Human Epidermal Growth Factor Receptor 2
hPSC - Human Induced Pluripotent Stem Cell
HR - Hormone Receptor
IP - Immunoprecipitation
iPSC - Induced Pluripotent Stem Cell
JAG1 - Jagged 1
MET - Hepatocyte Growth Factor Receptor
mRNA - Messanger Ribonucleic Acid
neg - Negative
PTM - Post Translational Modification
RasGAP - Ras-Specific GTPase-Activating Protein
RNA - Ribonucleic Acid
RTK - Receptor Tyrosine Kinase
RUNX - Runt Related Transcription Factor 1
RUNXi - RUNX Inhibition
SRC - Sarcoma Gene

TDLU - Terminal Ductal Lobular Unit
TF - Trascription Factor
TNBC - Triple Negative Breast Cancer
VEGFR - Vascular Endothelial Growth Factor Receptor

Chapter I: Introduction

1.1. Epithelial Breast Tissue

The mammary gland is a tissue that defines the class Mammalia within the kingdom Animalia (Capuco and Akers, 2009; Gregory, 1910; Oftedal, 2002). The development of this organ changed the way that mothers pass nutrition and immunological factors to their young. Response to pregnancy hormones as well as those released after birth triggers the organogenesis allowing for secretion of proteins and lipids forming milk from specialized cells (Russo and Russo, 2004), and permits for nutrient transfer between mother and offspring, greatly increasing their chances of survival. The ability of the body to form its own highly digestible and nourishing nutrients to feed progeny allows greater safety and increased fitness compared to species whose offspring are either left to defend themselves upon hatching from eggs or whose parents have to immediately hunt for their young, allowing predators increased chances to get into the nest (Oftedal, 2002).

The evolutionary history of the breast predates the emergence of mammals, with proto-mammals, or synapsids, believed to have possessed flat sweat glands capable of producing a nutrient-rich secretion to nourish offspring (Figure 1.1.) (Oftedal, 2002). Over millennia, early mammals further shaped the primitive mammary gland, enhancing its structure, invaginating it within the body, and promoting the development of alveoli for more efficient nutrient delivery to the young (Oftedal and Dhouailly, 2013). Today, among placental mammals, the mammary gland exhibits vast variability in anatomical features, cellular composition, and microenvironmental responses, reflecting millennia of evolutionary adaptation to a diverse ecological population (Capuco and Akers, 2009;

Gregory, 1910). The human breast, in particular, stands as a uniquely refined example of this organ, embodying millennia of selective pressures (Oftedal, 2002).

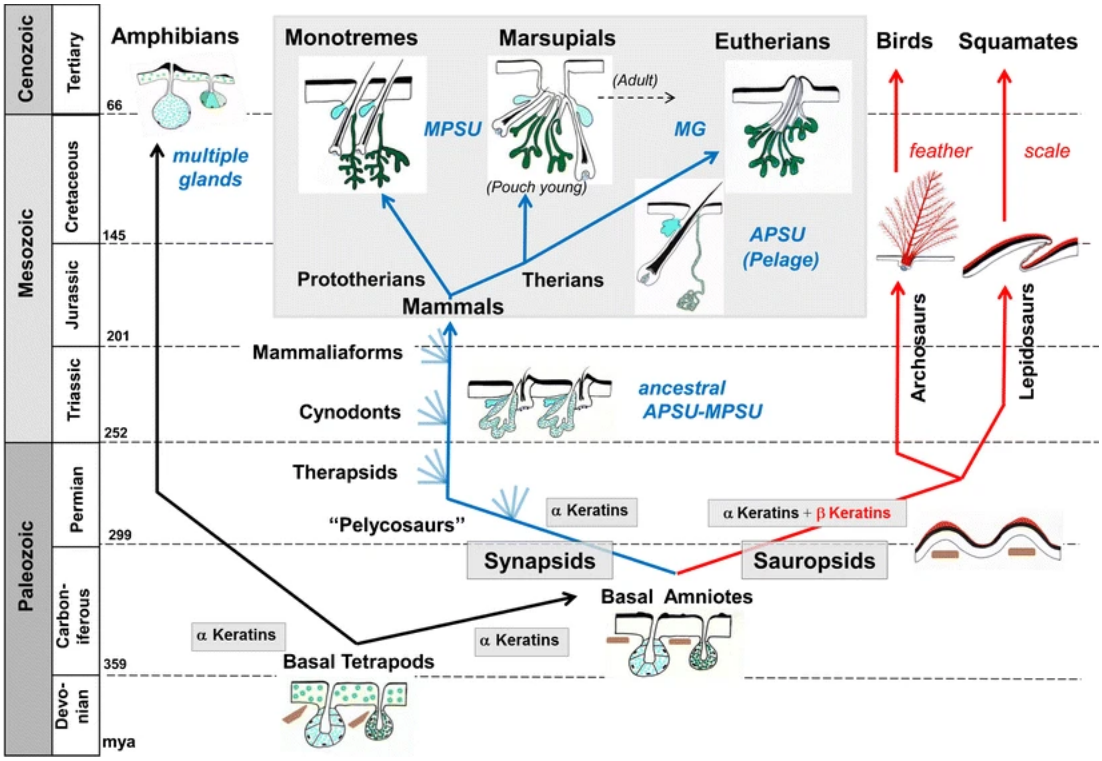


Figure 1.1. Evolutionary History of the Mammary Gland: It is believed that synapsids were the first lactating organisms and that their basic sweat glands formed the basis for today's mammary gland. Reprinted with permission *Oftedal, O.T., Dhouailly, D. Evo-Devo of the Mammary Gland. J Mammary Gland Biol Neoplasia* ***18****, 105–120 (2013).*

Undergoing millions of years of evolution, the mammary gland is a complex organ that is incredibly diverse amongst the class, reflecting the great diversity of both body plan and reproductive strategies evolutionarily entrenched in Mammalia (Capuco and Akers, 2009; Oftedal, 2002; Russo and Russo, 2004). Although with some variation, the mammary gland assumes a compound configuration consisting of two bilayered structures each comprised of luminal and basal cells. These compound structures, lobules and ducts, are responsible for the production of nutrients for the young and subsequent delivery to intended location respectively (Oftedal, 2002; Russo and Russo, 2004). While

these general structures remain similar across species, the networks that they form, the signals they receive from their microenvironment, and their response to said signals widely vary.

This lack of direct comparability can have implications when attempting to understand the development of both healthy and diseased tissue of the breast. As both structures, interlobular stroma, response to hormones, developmental cycles, and microenvironment vary between the human breast and the mammary gland of other organisms, animal models do not fully recapitulate the development of human tissue or its response to external stimuli (Mukherjee et al., 2022; Sugimoto and Sato, 2021). Additionally, 2D cell culture and current humanoid models do not retain physiologically relevant morphology, preventing the proper transduction of signals from the microenvironment to specialized tissue types (Kapałczyńska et al., 2018). As these models do not accurately replicate the intricate signaling seen in human breast tissue, translation of research to a clinical setting may not lead to expected results (Hockney et al., 2023; Kim et al., 2020). To combat these discrepancies, a collagen-based hydrogel scaffolding that allows for the formation of biomimetic breast organoids derived from healthy primary tissue has been developed (Sokol et al., 2016) and has already shown its potential by defining a novel role for the transmembrane tyrosine kinase receptor Discoidin Domain Receptor 1 (DDR1) (Rauner et al., 2021).

This thesis work strives to advance research and make an impact in the field in two ways. The first of these is through the enhancement of a novel 3D hydrogel model used to grow physiologically relevant primary breast organoids. This study reveals that the dissociation of organoid initiating units from an undefined cluster of epithelial cells to

single cells reduces challenges in reproducibility and quantification across primary samples allows for new stages of development to be observed and increases the applications possible through a primary breast organoid model. The second is the discovery of a novel downstream signaling axis from DDR1 to the essential transcription factor RUNX1, observed through use of this single cell derived primary breast organoid model. This investigation highlights the role of a DDR1-RUNX1 axis in cellular differentiation and morphological development of primary breast organoids was defined and may help to inform therapy options for clinical use in specific breast cancer subtypes.

1.1.1. Anatomy and Development of Mammary and Human Breast Tissue

Within a matrix of collagen, adipose tissue, and other stromal cells, exists the functional tissue that derives the mammary gland. This epithelial tissue is developed from progenitor cells that drive the proliferation and maturation of two cell types in distinct layers, the external basal cells that are oriented towards the microenvironment, and the internal luminal cells that are positioned towards the hollow lumen (Muschler and Streuli, 2010). These specialized cell types form both the network of ducts and the lobules that develop along these ductal structures (Figure 1.2.). Each lobule is made of functional alveoli where specialized cells respond to hormones, producing milk that collects and is pulled down the ducts to the teat (Arendt et al., 2014; Russo and Russo, 2004; Wellings et al., 1975).

The structure and function of human breast tissue is comparable to the structures of its anatomical counterpart, which is made up of the same cell types and is in part

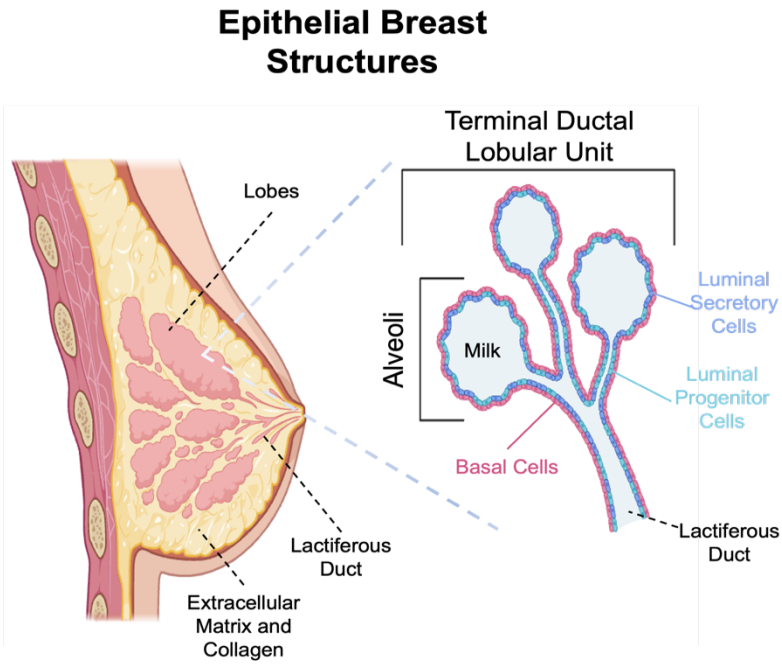


Figure 1.2. Human Breast Anatomy: Schematic of human breast anatomy and cross section of a TDLU with compound ductal alveolar structure. Both structural components are comprised of basal cells (red) on ECM facing layer and luminal cells (blue and green) on the interior.

distinguished from other mammary tissue by the development of the functional terminal ductal lobular unit at the end of its extensive ductal networks that promotes both development and alveologenesis (Russo and Russo, 2004). Breast tissue, like mammary tissue, also goes through a multitude of dynamic developmental phases throughout a woman's life and the duration of the tissues regenerative capacity that change both the composition and the functionality of the organ (Arendt et al., 2014; Russo and Russo, 2004; Williams and Daniel, 1983).

However, breast tissue is also unique in many ways. In contrast to the mammary tissue of other organisms, human breast tissue does not retain the same physiology, cellular makeup, and microenvironment characteristics (Holen et al., 2017). For example, the mouse mammary gland and human breast tissue are incredibly different (Bresslau,

1920; Daniel et al., 1968; Holen et al., 2017). Beyond major anatomical differences between the two species such as location, quantity, and size, the architecture of the functional glandular tissue in humans is more complex with a more extensive ductal lobular network, different classifications of lobules, and different terminal units (Dontu and Ince, 2015). These comparative tissues also respond differently to extracellular cues such as hormones and go through different signaling prompts during reproductive cycles. In addition, the intralobular stroma that surrounds the glandular tissue is different between that of mice and women: with adipose cells primarily encompassing mouse tissue and a more collagenous stromal compartment in humans. Together, these major differences, along with the many subtle distinctions between these tissues, creates a need for a more replicative model of human breast tissue. This need is particularly critical for understanding the intricate developmental phases underlying tissue formation, in which lack of a proper model has impeded a comprehensive understanding of the human breast's regenerative potential**.**

While the intricacies and signaling that drive this development in humans is not understood, the morphological traits of breast development have been characterized. These developmental phases begin during embryogenesis, when nipples connecting to 15-20 simple ductal structures form through the enfolding of a bud on the ectoderm. The developmental process then remains on pause until puberty when these ductal structures begin to network, becoming more complex as branches increase in their secondary and tertiary branching as well as their length (Williams and Daniel, 1983). During this time the functional alveoli also begin to bud to create the proper specialized cells to receive hormones, such as estrogen, progesterone, and prolactin, upon pregnancy to initiate

cyclical breast tissue growth (Williams and Daniel, 1983). In response to these hormones, the tissue enlarges through the proliferation and maturation of both ductal and alveolar structures. Lobules, in preparation for the formation of milk secretion, greatly increase in both number and size (Arendt et al., 2014; Russo and Russo, 2004). The tissue retains this mass and function until post-pregnancy regression where shifts in hormones direct involution of the tissue and apoptosis of the epithelial tissue begins to return close to pre-pregnancy levels (Russo and Russo, 2004). Upon receiving proper hormonal signaling, triggered by a subsequent pregnancy, the tissue will once again begin to proliferate to restore the glandular tissue to a functional orientation. These cycles occur until onset of menopause when the tissue atrophies and loses its regenerative capacity (Arendt et al., 2014; Hutson et al., 1985; Russo and Russo, 2004; Wellings et al., 1975).

1.1.2. Cellular Hierarchy and Stem Cells of the Breast

The cyclical process of human breast development and regeneration strongly suggests the presence of an adult stem cell compartment within the breast. Stem cells possess the remarkable capacity for self-renewal and pluripotency, enabling them to either replicate themselves or asymmetrically differentiate into specialized cell types (Reya et al., 2001). These cells exhibit varying degrees of pluripotency depending on developmental stage and tissue type, broadly classified into embryonic and adult stem cells. Embryonic stem cells, prevalent during fetal development, display high pluripotency, contributing to the formation of the entire body plan (Young, 2011). In contrast, adult stem cells are more tissue-specific and demonstrate a more limited capacity for tissue formation (Ferraro et al., 2010; de Morree and Rando, 2023). While

embryonic stem cells are implicated in breast initiation, the mechanisms governing human epithelial breast tissue regeneration are not as well understood, primarily due to challenges in the modeling of human breast tissue.

Thus, the field currently does not agree upon the types of cells that maintain this heterogeneous population (Arendt et al., 2014; Ferraro et al., 2010; Petersen and Polyak, 2010; Rios et al., 2014; Van Keymeulen et al., 2011; Visvader, 2009; Visvader and Stingl, 2014). Current cellular hierarchy dogma (Figure 1.3.) states that the multipotent breast stem cell does not persist beyond embryogenesis, however, its direct progeny, bipotent progenitors as well as long term unipotent cells, can regenerate the entirety of the adult mammary gland cellular hierarchy and maintain homeostasis within the tissue (Rios et al., 2014; Van Keymeulen et al., 2011; Visvader and Stingl, 2014). Research into this question, through the utilization of lineage tracing, has produced mixed results depending on models and techniques. Some data suggests that long term unipotent progenitor cells, when signaled, drive tissue regeneration during pregnancy through the generation of increasingly differentiated basal and luminal cells (Van Keymeulen et al., 2011).

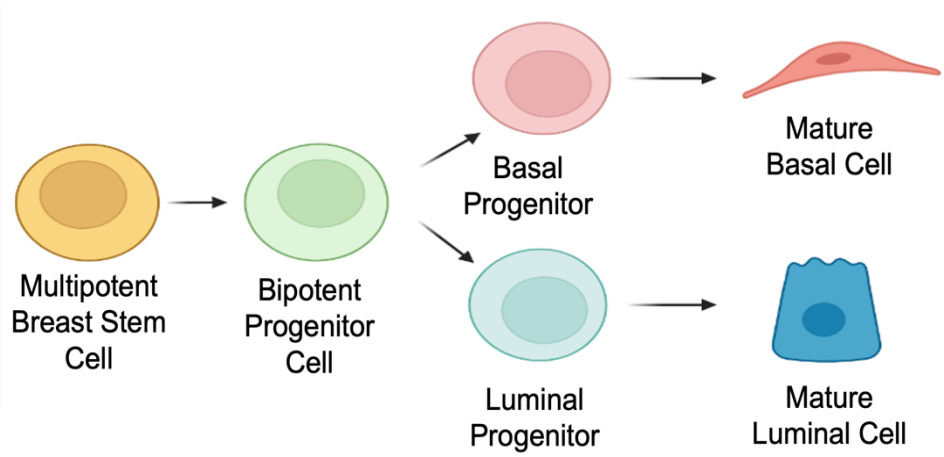


Figure 1.3. Human Breast Cellular Hierarchy: Common differentiation pathway and cellular subtype that makes up the mammary gland and the breast.

Other data from transplantation studies suggests that the cells driving tissue regeneration may be more defined and undergoing plasticity to recapitulate necessary cell types. In this vein, it has been observed that rather than a single cell type serving as the sole stem cell, experiments conducted under various conditions reveal that basal and luminal cell types exhibit plasticity and contribute to repopulating both epithelial lineages (Gupta et al., 2019; Keller et al., 2012; Shackleton et al., 2006). In these cases, cells act as "facultative' stem cells and give rise to both basal and luminal progenitor cells which through signaling differentiate from this progenitor state to that of a mature basal or luminal cell type respectively. Furthermore, it has been shown that there are specialized cell subtypes in both the basal and luminal lineages, such as estrogen receptor positive luminal cells (Jordan and Koerner, 1975; Stender et al., 2010). Each of these diverse cell types serves a variety of different functions within the structure based on their location and the molecular signals that they are exposed to (Arendt et al., 2014).

Basal cells, observed to be spindle-shaped, are the external cell layer of the human breast derived from basal progenitors and are acknowledged to specifically express epithelial markers such as; alpha smooth muscle actin (αSMA), and cytokeratin 14 (CK14), as well as being $EpCAM^{neg/low}$ $CD49f^{high}$ (Arendt et al., 2014; Sarrio et al., 2012). These cells are also known as myoepithelial cells due to their ability to act as a pseudo-muscular tissue through contractility, activated by oxytocin, to drive secreted milk from lobules to the nipple during lactation (Reversi et al., 2005). Basal cells across epithelial tissues are known to exhibit an essential interaction with the specialized extra cellular matrix (ECM) of the basement membrane that supports structural development. In the breast, basal cells are thought to act as the protective layer to the epithelial

glandular tissue and the interacting layer with the collagen-rich ECM microenvironment and immune cells, while also playing a role of dispersing signals from the microenvironment into the tissue (Gieniec and Davis, 2022; Williams and Daniel, 1983).

While basal cells are located on the ECM facing compartment, luminal cells line the apical surface and the lumen of glandular tissue. Luminal cells, derived from luminal progenitor cells, are the primary functional cell type of the organ as they some develop into specialized hormone-sensing cells that respond in turn with a response of proliferation, hollowing of the lumen, the production of milk, or a combination of the three (Brisken and O'Malley, 2010; Brisken et al., 1999; Haslam and Shyamala, 1979). Along with the expression of hormone receptors (HRs), these cells are characterized by the expression of markers such as Cytokeratin 8/18 (CK8/18) and E-Cadherin (E-Cad), as well as being $EpCAM^{high}$ $CD49f^{neg/low}$ (Arendt et al., 2014; Rauner et al., 2021; Sarrio et al., 2012).

While current paradigms may not fully explain the breadth of the breast's regenerative capacity, the ability to recapitulate all required cell types strongly suggests the existence of a facilitative stem cell population within mammary tissue. Until recently, studies investigating the breast stem cells have been limited due to differences between human tissue and the available models (Bresslau, 1920; Gregory, 1910; Holen et al., 2017; Russo and Russo, 2004). Recent advances in the 3D modeling of human tissue has allowed for the development of biologically relevant structures in conditions that allow for regular monitoring during the development of these tissues (Linnemann et al., 2015, 2017; Rauner et al., 2021; Sokol et al., 2016). Through the further enhancement of one of these novel 3D hydrogel models, this thesis aims to contribute clinically translatable

knowledge on the stem cell compartment of the breast and the signaling that occurs during human breast development.

1.2. Enhancement of Scientific Exploration Through 3D Tissue Models

Scientific models have long been fundamental to research, playing a central role in modern scientific practices. These models vary in complexity, aiming to simplify interactions among a limited number of factors for easier comprehension or replicate the intricate cellular context to provide a more comprehensive understanding of real-world phenomena (Chittleborough and Treagust, 2009). The choice between 2D and 3D models, for instance, involves a trade-off between representing cell-cell interactions with simplicity or capturing the contextual nuances of structural expression in response to external stimuli.

Similarly, opting between animal models or primary organoid models offers insights into developmental processes or therapeutic interventions within the holistic environment of a complete organism or the same development in a more restricted yet more human-relevant system. Particularly in the translation of clinically relevant work, selecting an appropriate model can enhance the acquisition of pertinent information, while an inadequate model may obfuscate results, leading to discrepancies and contradictions during the transition from laboratory research to clinical application (Jensen and Teng, 2020; Kapałczyńska et al., 2018).

Understanding the functionalities and limitations of each model allows for the proper interpretation of experimental results. 2D cell culture models have long been indispensable in human cellular research, serving as a primary tool for studying the

effects of experimental factors on human tissue without ethical concerns associated with human testing. Immortalized cell lines, collected from various tissues, offer researchers the flexibility to investigate variables in tissues of interest. These cell lines can be genetically modified and subjected to alterations in their niche, enabling researchers to mimic diverse physiological conditions (Abbas et al., 2021; He et al., 2017; Kapałczyńska et al., 2018).

While 2D models have significantly advanced disease therapeutics and facilitated the discovery of cancer-specific targets, their use raises questions about the relevance of findings to in vivo conditions. Immortalized cell lines, often oncogenic in origin, may not accurately represent normal human tissue, leading to limitations in replicating research outcomes (Abbas et al., 2021; Kapałczyńska et al., 2018; Linnemann et al., 2015). Additionally, the lack of tissue structure in 2D models hinders proper cellular interactions, potentially misregulating physiological signaling observed in healthy tissue (He et al., 2017). Moreover, inconsistencies in culturing techniques and the limited genetic variability within cell lines further highlight the need to understand the model's limitations for better translational research (Abbas et al., 2021; Ben-David et al., 2018).

On the other hand, animal models, such as Planarians, Drosophila, and Mus Musculus, offer an alternative approach to studying biological processes with greater context and translatability compared to 2D models. Unlike 2D models, animal models maintain tissue structure and normal physiological interactions, providing a more comprehensive understanding of biological phenomena (Chan et al., 2021; Hickman et al., 2017; Mukherjee et al., 2022). Animal models also allow for experiments across populations, enhancing the generalizability of research findings. Within the use of animal

models, the selection of an appropriate model organism is crucial for maximizing the information obtained (Chittleborough and Treagust, 2009). While simpler organisms can elucidate specific interactions and components within a system, more complex models offer insights into multifaceted systems resembling human tissue. Despite their advantages, structural and cellular differences among model organisms can complicate results and hinder direct translation to human tissue, emphasizing the importance of careful consideration when choosing a model for research purposes.

Despite their faults, the information gained from these models makes up the foundation of our current scientific understanding. Historically, however, certain questions have proven too complex for existing technologies to address. Until recently, inquiries regarding human stem cell-initiated embryogenesis and tissue regeneration have posed challenges in modeling and observation within the limitations of 2D, 3D, and animal models, primarily due to disparities in both structure and fundamental biology among these models. Efforts to enhance the humanization of these models aim to tackle these questions are ongoing.

In the last decade, improvements in 3D cell culture/organoid models have allowed for their expansion as a major model type used for research into organogenesis and regeneration of multiple tissue types, including the human breast. Through the development of more tunable 3D scaffolds and a greater understanding of stem cell biology (He et al., 2017; Leisten et al., 2012; Sokol et al., 2016), large improvements have been made that have allowed for new and increasingly accurate knowledge regarding cellular responses, and processes such as organogenesis and regeneration. This truer representation of human tissue development also increases the application of the

model and allows for the interpretations of questions surrounding stem cell initiated tissue generation.

1.2.1. 3D Culture and Modeling of Human Breast Development

An emerging focus on stem cells and their microenvironment, coupled with the inherent limitations of traditional 2D cell culture and animal models, drove the creation of novel ex vivo models aimed at addressing inquiries concerning stem cell development into fully formed tissues. Investigations into pluripotent stem cells underscored the significance of the microenvironment, or niche, not only in the shaping of tissue architecture but also in the orchestration of the signaling required for tissue generation (Roskelley et al., 1994). Similarly, studies focusing on the surrounding cellular environment have provided insights into the biophysical characteristics of these macromolecules, which play pivotal roles in tissue formation (Alcaraz et al., 2008; Lee et al., 1984; Li et al., 1987; Rauner et al., 2021; Roskelley et al., 1994). This reciprocal exchange of observations between stem cells and the extracellular matrix has pushed the field forward, leading to the development of powerful biomimetic models.

The exploration of the stem cell niche began in the 1970s when researchers observed that the microenvironment surrounding cells played a crucial role not only in tissue development but also in the onset of diseases like cancer and the aging process (Patel and Lodish, 1987; Schofield, 1978; Vogel et al., 1997). This realization prompted researchers to harness the potential of this matrix to induce the differential expression of various cell types and facilitate tissue formation, moving beyond the confines of traditional 2D cell cultures. Utilizing a basement membrane matrix derived from mouse-

originating Engelbreth-Holm-Swarm tumor cells, researchers created an extract that would later become the commercially available Matrigel (Kibbey, 1994). This basement membrane extract provided a substrate where mouse stem cells could interact with niche signaling cues, fostering tissue development (Li et al., 1987). Subsequent studies have investigated the individual components of Matrigel to gain a deeper understanding of their roles in tissue development, deciphering their structural and signaling contributions during the differentiation process (Lee et al., 1984; Li et al., 1987; Roskelley et al., 1994).

Researchers focusing on the human breast were among the first in utilizing these matrix-based models to gain insights into tissue development. By embedding mouse mammary cells in gels derived from Matrigel basement membrane or purified collagen gels, researchers observed not only cell proliferation but also self-organization and the formation of specific tissue structures resembling those observed in vivo (Lee et al., 1984; Li et al., 1987). Tissues cultured in these gels exhibited distinct milk secretion protein expression patterns compared to those grown in 2D environments. This highlighted the crucial role of the extracellular matrix (ECM) in guiding spontaneous organogenesis, extending beyond its structural support function to actively influence tissue growth and organization.

These findings provided researchers with new methodologies to investigate the initiation of mammary structure development and regeneration. Previously, such investigations were limited to embryogenesis studies or surgical transplantations. However, these new experiments could be conducted more easily without the need for entire animals, enabling more accessible and observable studies. Despite becoming the standard in many breast studies, these techniques still face significant limitations in

translating results to human tissue. While mouse tissue could faithfully replicate morphologically accurate structures when cultured in 3D gels made of Matrigel or collagen or transplanted into cleared mouse fat pads, human tissue failed to achieve comparable levels of tissue growth (Pasic et al., 2011; Yang et al., 1987).

To overcome the limitations of growing relevant human breast structures in Matrigel and collagen gels, researchers sought novel approaches to enhance existing models for studying human biology-related issues. While primary adult tissue failed to form structures in Matrigel, researchers discovered that the immortalized human breast cell line MCF10A could form small acini (Debnath et al., 2003). Since these cells were non-tumorigenic, they provided researchers with a model to investigate the initiation and differentiation of basic human structures and their interactions with the environment. Moreover, this model offered insights into oncogenic transformations and facilitated the study of tumor development (Debnath et al., 2003; Fischbach et al., 2007). For the first time, processes like lumen hollowing could be observed in human tissue ex vivo, highlighting the significant differences in tissue structural mechanics compared to growth on 2D plastic surfaces. The ability to model, observe, and manipulate these functions laid the groundwork for our understanding of 3D breast biology.

Additionally, efforts were made to “humanize” mouse fat pads as another platform for growing human tissue. As primary human tissue and Matrigel are not compatible, studying breast morphogenesis in vivo became preferable to not studying it at all. Xenograft models, involving the injection of immortalized human fibroblasts into cleared mouse fat pads, provided a space for human epithelial morphogenesis to occur (Kuperwasser et al., 2004). This approach enabled studies on growth and development

within a system that could be modulated with experimental variables, thereby offering a deeper understanding of lifelong factors influencing breast structure growth. While these models injected a level of humanization into 3D breast models and facilitated comparisons of therapeutics in more human-like disease models, neither readily enabled the observation of the fundamental process of initiating a biomimetic human breast structure under human conditions.

Concurrently with these 3D modeling advances came another major technological leap with the introduction of induced pluripotent stem cells (iPSCs) (Takahashi and Yamanaka, 2006), soon followed by human pluripotent stem cells derived from somatic (Yu et al., 2007), or from fibroblasts (Takahashi et al., 2007). These cells gave researchers the ability to undefine any cell into a plastic state and then redefine it into a cell type of choice, allowing access to stem cells able to form all tissue types. While previously stem cells were rare and acquired from surgeries, this advancement made them readily available and opened the opportunity for use of human organoids. Standardized procedures allowed the focus of research to be on the factors that drove the differentiation of these cells and not directly on their stem potential, allowing researchers to finally study the interactions between the stem cells themselves and the niche.

Other models around the same time probed adult tissues for long term stem cells that were able to differentiate and recapitulate functional tissues. Early work with intestinal crypt cells showed that certain tissues do not require being attached to a niche for the regeneration of tissue, but more importantly that some adult tissues in humans still retain stem cells (Rosenbluth et al., 2020; Sato et al., 2009). From this time, organoids generated under physiologically defined conditions and with the use of hPSC, embryonic

stem, or adult stem cells as cell sources from tissues such as, the lung, breast, liver, heart, skin, stomach, esophagus, intestine, and colon (Sokol et al., 2016; Zhao et al., 2022).

In recent years, the understanding of the stem-niche relationship has allowed for the development of models that enable the formation of morphologically accurate human breast tissue from primary sources (Linnemann et al., 2015, 2017; Sokol et al., 2016), providing researchers with a new spatial-temporal understanding of normal and disease development across the tissue. The tissue generated from these models, cultured in floating pliable hydrogels, forms physiologically accurate ductal-lobular structures containing basal and luminal cell types in their proper arrangement, facilitating some quantitative analysis and further study of structural development (Linnemann et al., 2015). However, while the tissue formed in these initial models was morphologically accurate, the culture conditions often included components not normally found in the human breast environment (Linnemann et al., 2017), potentially clouding the translation of results from the laboratory to clinical settings.

More recently, efforts to create a more biomimetic 3D hydrogel model without inhibitors and containing only ECM components found in the breast stroma, namely collagen-1, laminin, fibronectin, and hyaluronic acid, seeded with small clusters of primary patient cells, allows for the formation of a biomimetic model (Rauner et al., 2021; Sokol et al., 2016). Compared to previously iterations of breast culture models, this organoid has proper compound structure formation (i.e. the generation of a TDLU that contains both ductal and alveolar structures) (Figure 1.4. A), maintains lineage fidelity (Figure 1.4. B), allows for the expression of HRs, and exists in a biomimetic niche, making it the most representative model of human breast tissue thus far. This novel

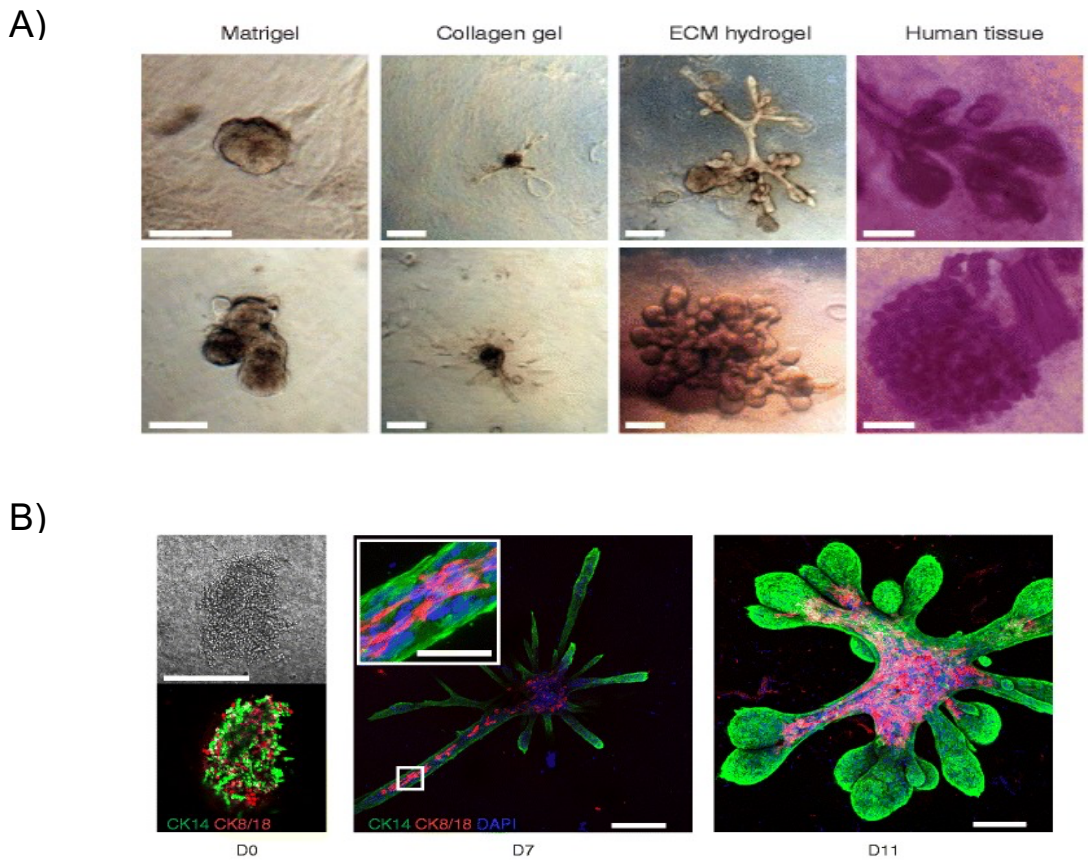


Figure 1.4. Formation of a Biomimetic Breast Organoid: A) Comparison of primary human growth in 3D conditions, including Matrigel, pure collagen gels, and a novel 3D Hydrogel Model to sections of human primary tissue. B) Development of primary breast tissue in novel hydrogel model on day of seeding (D0) when seeded as a small cluster of primary cells, day 7 after seeding (D7) during ductal elongation, and Day 11 after seeding (D11) during alveologenesis. CK14 (green) is a basal marker, while CK8/18 is a luminal marker. Adapted *Sokol ES, Miller DH, Breggia A, Spencer KC, Arendt LM, Gupta PB. Growth of human breast tissues from patient cells in 3D hydrogel scaffolds. Breast Cancer Res. 2016 Mar 1;18(1):19. doi: 10.1186/s13058-016-0677-5. PMID: 26926363; PMCID: PMC4772689*

model has already begun to advance our knowledge of breast tissue and its interaction with the ECM as well as what signaling is essential for breast tissue development of a biomimetic structure at specific times during expansion of the glandular tissue (Rauner et al., 2021). Part of this thesis work aims to enhance this novel 3D model through the continued dissociation of the initiating clusters of primary cells from small clusters to single cells. This advancement allows for more accurate and representative

quantification, a more complete understanding of the developmental timeline if the structure, and new research applications such as lentiviral transduction of tissue that will further enhance research in breast development.

1.3. **Discoidin Domain Receptor 1 and Development**

Functional tissue requires continuous crosstalk between cells and their microenvironment to interpret signals from the extracellular space and transduce them into intracellular signals. Cells utilize Receptor Tyrosine Kinases (RTKs) to facilitate interactions between receptors and ligands with high affinity (Du and Lovly, 2018; Lemmon and Schlessinger, 2010; Schlessinger, 2000). RTKs share consistent structural features, comprised of an extracellular receptor domain for ligand binding, a hydrophobic transmembrane domain, and an intracellular catalytic domain that responds to ligand-receptor binding and undergoes conformational changes allowing for phosphorylation. Upon activation, the catalytic domain often serves as a site for protein scaffolding, where numerous cell-specific proteins interact and are subsequently activated, initiating signaling cascades that drive various intracellular changes, from modulating protein-protein interactions to orchestrating shifts in the transcriptome and cell identity.

RTKs are categorized into 20 classes based on their known ligand families (Wagner et al., 2013). These ligands encompass various signaling molecules, including those mediating cell-cell interactions, paracrine signals, secreted factors such as cytokines and growth factors, endocrine hormones, direct interaction with the ECM, and some receptors lacking known ligands. The diversity of receptor-ligand pairings underscores the broad functional roles maintained by this class of transmembrane receptors. RTK

dysregulation is common in tumors, leading to aggressive malignancies (Du and Lovly, 2018). To counteract this, selective small molecules have been developed to inhibit RTK signaling by binding competitively to their ligand-binding sites. Notable targets include HER2 (Lapatinib), EGFR (Gefitinib), MET (Crizotinib), VEGFR (Sunitinib), among others. Over 20 small molecule inhibitors are approved for cancer treatment (Hojjat-Farsangi, 2014), further emphasizing the importance for further research into these receptors beyond their fundamental role in development.

1.3.1. The Structure and Function of DDR1

One cell surface receptor recently explored in human tissue is Discoidin Domain Receptor 1 (DDR1). DDR1, along with DDR2, located on chromosome 6 and 1 respectively (Leitinger, 2014), are transmembrane receptors that binds their ligand in the extracellular space with their discoidin homology domain and delivers their signal intracellularly (Figure 1.5. A). These receptors are primarily expressed on epithelial cells (Alves et al., 1995). Functional isoforms of these receptors, containing the kinase catalytic domain, typically range from 855 to 913 amino acids in length (Song et al., 2011). Ligand binding triggers kinase activation and autophosphorylation, which can be sustained for several days post stimulation (Vogel et al., 1997), facilitating the recruitment and activation of various proteins, such as SRC family proteins, Stat transcription factors, RasGAP, and others (Leitinger, 2014) (Figure 1.5. B).

For many years DDR1's ligand was unknown, giving it an orphan receptor designation until 1997. At this time it was discovered that, unlike many other RTK ligands that are secreted and sent from neighboring cells to stimulate a response, DDR1's

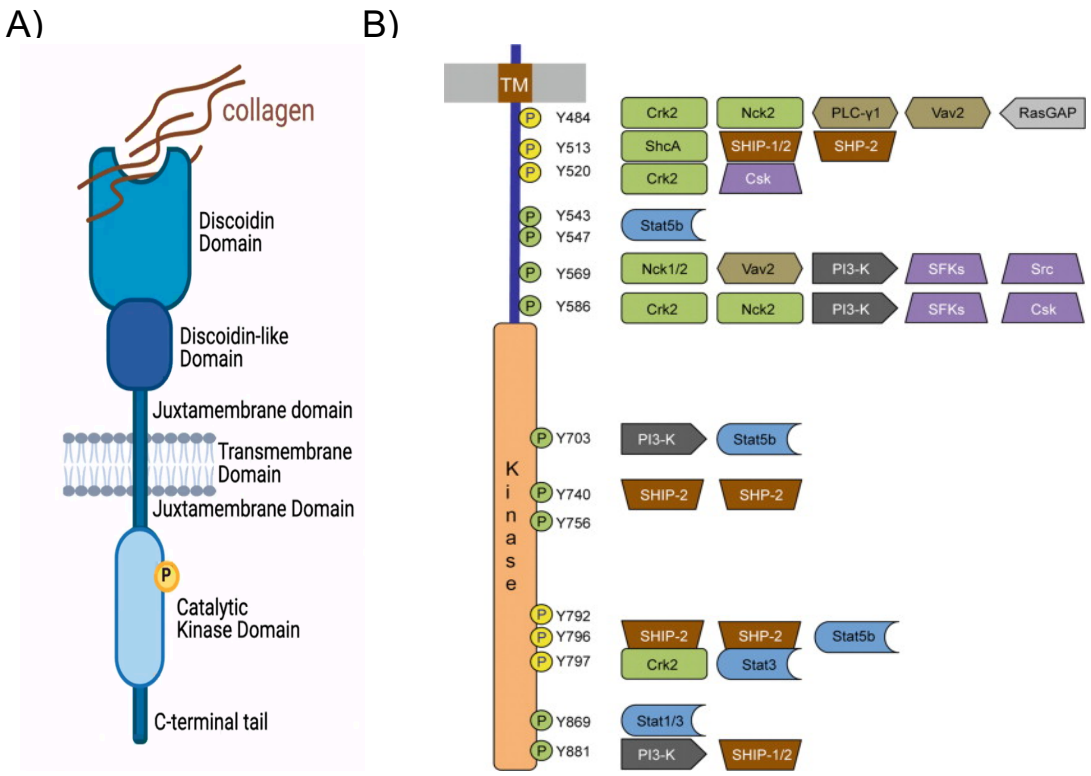


Figure 1.5. Structure of DDR1: A) Structure of transmembrane receptor tyrosine kinase Discoidin Domain Receptor 1 (DDR1). B) Diagram of DDR1's catalytic kinase domain phosphorylation sites and subsequent protein scaffolding that binds to the activated receptor for their own activation. Reprinted with permission from *Leitinger B. Discoidin domain receptor functions in physiological and pathological conditions. Int Rev Cell Mol Biol. 2014;310:39-87. doi: 10.1016/B978-0-12-800180-6.00002-5. PMID: 24725424; PMCID: PMC4021107.*

only known ligands were various fibrillar forms of collagen (Types I - IV) in the ECM (Shrivastava et al., 1997; Vogel et al., 1997). DDR1 exhibits specificity for binding the amino acid sequence GVMGFO across these fibrillar collagens (Xu et al., 2011). This signaling between the cell and the niche is hypothesized to play a crucial role in development, allowing cells to directly respond to cues in their microenvironment and determine their cellular identity and needs.

The developmental role of DDR1 in the formation of tissue has been validated through many model organisms, as its expression is conserved as far back in history as

some of the earliest multicellular organisms, metazoans, highlighting its importance to essential cellular processes and the existence of multicellular life (Chan et al., 2021). DDR1's specific developmental functions include differentiation, cell migration, cell cycle progression, control over EMT, and involvement in anti-apoptotic pathways (Leitinger, 2014). These functions are in conjuncture with synergistic and antagonistic interactions between DDR1 and other cell membrane proteins and receptors.

Experiments in planarians, or flat worms, have shown that beyond its role in neuronal development, DDR1 activation through the binding of collagen can have far-reaching effects that drive the differentiation and development of multiple cell types (Chan et al., 2021). Through the binding of collagen IV, DDR1 stimulates neurons to produce NRG-7, a neuregulin that promotes the asymmetric division, differentiation, and proliferation of neoblast cells, destined to repopulate any damaged tissue within the organism. The ability of DDR1 to direct regeneration and lead to the signaling of an indirect cell type should be investigated in other tissues and organisms.

Knockout of DDR1 in mice led to a variety of issues with organogenesis across tissue types. These mice, which were smaller in size, experienced deficiencies in the formation of reproductive tissues, mammary glands, kidneys, inner ears, gastrointestinal tracts, and central nervous system (CNS) (Barker et al., 1995; Vogel et al., 2001). This loss of collagen signaling reduced both function and regenerative capabilities upon tissue damage in both developing and adult tissues (Leitinger, 2014; Meyer zum Gottesberge et al., 2008). These studies show the importance of interaction between cell and matrix and imply that this receptor could have major implications in human developmental biology.

1.3.2. The Human Physiology of DDR1

Technological advancements in cell culture techniques have enabled recent exploration of DDR1's role in organogenesis across various human tissues, including the colon, ovaries, central nervous system, liver, hematopoietic system, and others (Ben Arfi et al., 2022; Chen et al., 2021; Elkamhawy et al., 2021; Fowler et al., 2020; Leitinger, 2014; Rauner et al., 2021). Insight from model organisms such as the mouse and flatworm has allowed researchers to translate what is known and compare it to human physiology. These researchers have discovered that DDR1's role in organogenesis and homeostasis translates to the human nervous system, liver, and hematopoietic system.

DDR1 has recently been the focus of study in neurodegenerative research. In healthy tissue, activation of DDR1 has been observed to drive expression of n-cadherin, required to form cohesive neuronal tissues (Huang et al., 2016; Shintani et al., 2008), and the myelination and remyelination of oligodendrocytes (Vilella et al., 2019). In CNS disease states such as Alzheimer's and Parkinson's, DDR1 is often upregulated (Fowler et al., 2020), and that this activation can create disease-specific gene signatures. It has also been shown that inhibition or knockdown of DDR1 can reduce the burden of neurotoxic proteins (Fowler et al., 2020). It has been hypothesized that DDR1 can control survival pathways through modulation of neuroinflammatory responses (Fowler et al., 2020).

In liver research, DDR1 has been primarily studied in the context of the fibrotic disease (Manabe Ichiro et al., 2002; Moll et al., 2019; Takai et al., 2018). DDR1 plays a crucial role in liver regeneration by activating hepatic stellate cells, driving tissue regrowth (Romayor et al., 2020). Hepatic stellate cells are also responsible for collagen

synthesis in the liver, which is essential for maintaining tissue integrity during regeneration. However, over-activation of these cells through constant tissue damage leads to the overproduction of DDR1, causing fibrotic disease. DDR1's role in the fibrotic disease state is also being investigated in renal fibrosis, pulmonary fibrosis, and the scarring of the skin (Chen et al., 2016; Moll et al., 2019), in which DDR1 inhibitors are being investigated for therapeutic purposes (Moll et al., 2019).

Despite being known for their non-adherent nature, even hematopoietic cells leverage DDR1's binding to collagen to regulate differentiation, proliferation, and population homeostasis within the bone marrow ECM (Abbonante et al., 2013; Leisten et al., 2012; Zanetti and Krause, 2020). Hematopoietic stem cells and precursors interact with the collagenous matrix of the bone marrow to allow for mobility and development that is crucial for normal blood cell formation (Abbonante et al., 2013; Zanetti and Krause, 2020), further showcasing DDR1's control of stem regulation.

The data from these diverse tissues highlights not only the receptors, but also the microenvironment's role in tissue development through stem cell regulation. Dysregulation or overexpression of DDR1 can contribute to disease states, necessitating inhibitor treatment. While DDR1 drives proliferation and cell survival, inadequate downstream signaling may lead to tissue dysfunction, and thus, DDR1 is often upregulated in cancers and fibrotic diseases (Berestjuk et al., 2022; Moll et al., 2019; Valiathan et al., 2012). A more enriched understanding of DDR1's fundamental regulation during the development and regeneration of these tissues in physiologically healthy conditions will allow further understanding of how it can also contribute to these disease states.

1.3.3. DDR1's Role in the Human Breast Epithelium

Research into DDR1's role in human breast tissue is emerging. In mice, DDR1 plays a significant role in mammary gland development, as evidenced by major defects in ductal tissue branching and breast milk secretion in knockout mice (Barker et al., 1995; Leitinger, 2014; Vogel et al., 2001). DDR1 expression and activity spans throughout the development of the mouse mammary gland from embryogenesis into adulthood, implying that DDR1 expression could be necessary through menopause in women (Vogel et al., 2001).

Compared to the predominantly adipose mouse fat pad, the human breast is contained in a more collagenous matrix, highlighting the potential significance of DDR1 expression in human breast glandular epithelial tissue, present in both lobular and ductal structures. Studies have shown this to be true, as DDR1 is necessary for the differentiation of the luminal lineage which is subsequently required for the formation of fully mature alveolar structures (Rauner et al., 2021) (Figure 1.6. A). DDR1 on "stem" and basal populations drives Jagged1 (JAG1) expression, interacting with NOTCH1 on luminal cells to promote cellular layer maturation (Figure 1.6. B). While DDR1 demonstrates regeneration and differentiation abilities, its nuclear translocation and transcriptional regulation remain poorly understood in healthy human tissue, preventing a comprehensive understanding of breast development. As with many of the other tissues that DDR1 has been studied in, the improper activity of DDR1 in breast tissue is also accompanied by disease states, particularly triple-negative breast cancer (TNBC) (Han et al., 2022; Sun et al., 2018, 2021; Takai et al., 2018; Zhong et al., 2019). While physical characteristics of the protein were implied to play a role in TNBC disease progression, it

is not known what function DDR1's downstream signaling plays in the oncogenesis of this breast cancer subtype.

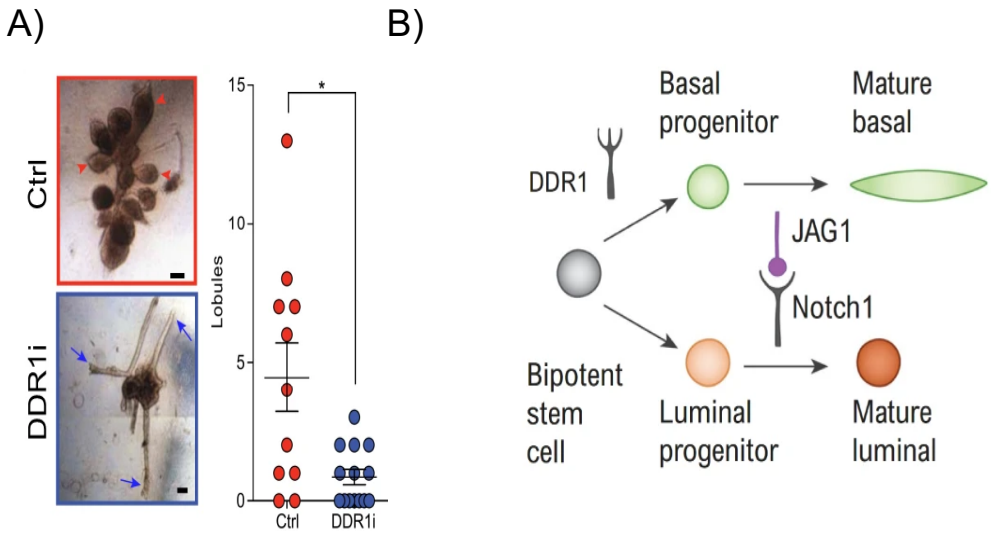


Figure 1.6. Roles of DDR1 in the Human Breast: A) Loss of DDR1 expression in developing breast organoids leads to the loss of alveoli. B) DDR1 activation allows the differentiation of "Stem" cells into Basal Progenitors which express JAG1. The interaction with these JAG1 expressing basal cells with NOTCH1 expressing luminal cells drives both cell types to maturity. Adapted with permission from *Rauner, G., Jin, D.X., Miller, D.H. et al. Breast tissue regeneration is driven by cell-matrix interactions coordinating multi-lineage stem cell differentiation through DDR1. Nat Commun* ***12****, 7116 (2021).*

Collectively, findings from these various organs highlight the pivotal involvement of DDR1 in tissue homeostasis, through its regulatory influence on stem cell dynamics, differentiation processes, and cellular proliferation. An in-depth comprehension of the downstream signaling pathways governed by DDR1 holds promise not only for advancing our understanding of its tissue generating developmental functions but also for understanding how this signaling can be manipulated to promote disease and exploring therapeutic avenues in disease contexts. This thesis project aims to investigate DDR1's downstream network in the breast with the overarching goal of elucidating its regulatory mechanisms throughout the process of breast organogenesis and development.

1.4. RUNX1's Involvement in Tissue Development

RTKs serve as initiators of signaling cascades that drive changes in cellular programming and identity, however it is downstream proteins that transduce these signals into functional changes in cellular machinery and programming. One of the most powerful drivers of cellular change are transcription factors (TFs) (Spitz and Furlong, 2012). Upon activation, these nuclear proteins bind to specific DNA binding domain (DBD) sequences, recruiting cofactors to promoter and enhancer regions, and facilitating DNA conformational changes that promote mRNA synthesis or repress gene transcription (Bauer et al., 2010; Jemc and Rebay, 2007). Diverse post-translational modifications (PTMs), including phosphorylation and acetylation, further modulate transcription factor activity, contributing to tissue-specific and cell-specific responses (Yeung et al., 2018).

Transcription factors often form complexes with various cofactors, amplifying the complexity and diversity of gene expression regulation. Interactions between primary transcription factors and cofactors drive differential transcriptional responses, influenced by cellular context and the dynamic interplay of protein expression and epigenetic regulation (Bauer et al., 2010; Jemc and Rebay, 2007). These modifications, alongside transcription factors' non-canonical roles, extend their functions beyond transcriptional regulation, encompassing DNA repair, cell cycle control, epigenetic modulation, and immune response (Friedman, 2009; Hu et al., 2022; Krishnan and Ito, 2017).

However, misregulated transcription factor activity is often a hallmark of disease progression, swiftly transforming cells from benign to malignant states (Lee and Young, 2013). Tumor suppressor transcription factors, which normally maintain cellular homeostasis, are frequently lost during oncogenesis, while oncogenic transcription

factors promote malignant transformation by driving aberrant gene expression programs (Bushweller, 2019; Xu et al., 2000). Additionally, certain transcription factors exhibit context-dependent roles, acting as either tumor suppressors or oncogenes based on the cellular environment, highlighting the intricate interplay between transcriptional regulation and disease pathogenesis (Shen et al., 2018).

1.4.1. RUNX1 Structure and Interactions

Runt-related transcription factor 1 (RUNX1), also known as core binding factor alpha (CBFα) (Okuda et al., 2001), is a versatile transcription factor with dual roles as an activator or repressor of transcription, depending on the context (Figure 1.7. A and B). RUNX1 belongs to a family of transcription factors acknowledged as master regulators of development and differentiation of many tissue types, and is primarily recognized for its role in hematopoiesis (Bresciani et al., 2014; Friedman, 2009; Growney et al., 2005; Okuda et al., 2001). Initially identified in Drosophila, the Runt proteins, to which RUNX1 is homologous, play crucial roles in the proper development of segmented embryos, specific sections of the central nervous system (CNS), and the regulation of sex determination (Canon and Banerjee, 2000; Walrad et al., 2011).

Runt proteins are characterized by a distinctive Runt domain, containing a highly conserved DNA binding domain essential for their function. This domain facilitates the binding of Runt proteins to specific DNA sequences, including the core binding factor (CBF) DNA sequence of TGTGGT (Tahirov et al., 2001). Interacting with cofactors, Runt forms functional protein complexes typical of more complex organisms' transcription factors (Adya et al., 2000; Canon and Banerjee, 2000; Walrad et al., 2011).

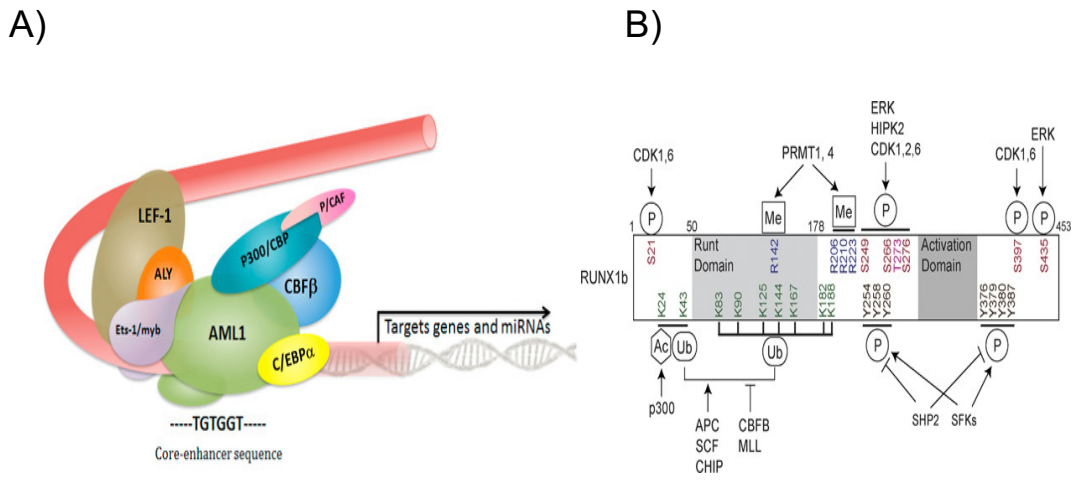


Figure 1.7. RUNX1: Interactions and Post Translational Modifications: A) Diagram of RUNX1's interactome. Reprinted with permission from Beghini A. Core Binding Factor Leukemia: Chromatin Remodeling Moves Towards Oncogenic Transcription. *Cancers*. 2019; 11(12):1973. (2019). B) Diagram of RUNX1's post translational modifications (PTMs) and effectors of these PTMs. Reprinted with permission from from Goyama, S., Huang, G., Kurokawa, M. *et al.* Posttranslational modifications of RUNX1 as potential anticancer targets. *Oncogene* **34**, 3483–3492 (2015)

Typically, Runt interacts with a β-subunit, Brother (Bro) or Big Brother (Bgb), inducing a structural change in Runt that enhances its DNA binding affinity to the target sequence (Adya et al., 2000). Other cofactors allow for increased complexity of the roles that Runt can play. For example, interaction with the protein product of another essential drosophila embryonic gene, *hairless*, allows for the control of specific spatial regulation of Runt target gene expression (Walrad et al., 2011). These interactions control the Runt protein's switch from an activator to a repressor of transcription, depending upon the cellular context that it is given.

As a primary pair-rule gene, a master regulator of drosophila body plan development in the embryo, Runt is activated directly in oocytes premade maternal mRNA transcripts, and then will shift to a zygotic expression (Tsai and Gergen, 1994). Its activation cascades to drive the activation of secondary pair-rule genes which further

shape embryonic development. If mutations in the Runt gene *run* do not cause embryonic lethality, it can cause segmentation defects that can leave the drosophila with abnormalities or even the deletion of entire body segments. Mutations and deletions to Runt-related proteins in other organisms such as vertebrates can cause embryonic lethality (Growney et al., 2005), or the misregulation of tissue (Ariffin, 2022; van Bragt et al., 2014; Gerritsen et al., 2019). The essentialness of this protein for life underscores its importance in stem cell differentiation, organogenesis, and tissue regeneration, warranting the utmost scrutiny in multiple fields of biological research.

RUNX1 stands as a highly conserved transcription factor, preserving not only its Runt-domain but also the TGTGGT Core-enhancer sequence across vertebrates, indicating the persistence of essential interaction sites throughout evolution (Uchida et al., 1997) (Figure 1.7. A). Its critical functions in human development underscore the enduring importance of RUNX proteins for embryonic viability and proper development. However, with the increasing complexity of body plans since Drosophila, the ancestral roles of the Runt protein have diversified into three distinct proteins (Ito et al., 2015; Krishnan, 2023; Lin, 2022; Otálora-Otálora et al., 2019). To compensate for this complexity, the RUNX protein family now comprises RUNX1, RUNX2, and RUNX3, each retaining the core binding factor site while adopting spatially and temporally specific functions in different tissues (Ito et al., 2015).

While these RUNX proteins share some interactomes and transcriptomes to compensate for potential loss of any family member, significant variability exists among them. Notably, their predominant expression sites differ: RUNX1 prevails in the hematopoietic system (Okuda et al., 2001), RUNX2 is prominent in osteogenesis and

skeletal development and is sometimes linked to a pivotal point in vertebrate evolution (Xu et al., 2015), while RUNX3 plays a significant role in neurogenesis (Levanon et al., 2002). This distribution ensures the vital contribution of the RUNX family to the development of body plans in vertebrates. Through the deployment of these three distinct family members, the RUNX family sustains its fundamental role in shaping the blueprint of development.

Like its counterpart Runt in drosophila, RUNX1 exhibits interactions with multiple cofactors for PTMs diversifying its roles within different cellular contexts (Figure 1.7. A and B). In an evolutionary push for simplicity, all tissues and RUNX family members bind to a β-subunit, core binding factor β (CBFβ) (Malik et al., 2019; Qin et al., 2015). Analogous to Runt's β-subunits Bro or Bgb, CBFβ doesn't directly bind to DNA but induces a conformational change in the RUNX protein, enhancing its DNA affinity (Adya et al., 2000; Tahirov et al., 2001). These various other cofactors are further influenced by tissue and cell type, expression, and cellular signals induced by internal and external stimuli (Ito et al., 2015; Krishnan, 2023). These cofactors have the ability to toggle RUNX's transcriptional activity and can change it from an activator to a repressor of transcription.

The ability of RUNX to associate with these cofactors confers non-canonical roles, augmenting the complexity of its functional repertoire. Beyond its transcriptional function, RUNX proteins contribute to DNA repair by interacting with p53 and BLM helicase (Krishnan, 2023; Samarakkody et al., 2020), cell cycle regulation through engagement with CDKs and ERK (Friedman, 2009), epigenetic regulation by partnering with chromatin-modifying enzymes such as HDACs (Beghini, 2019), and participate in

diverse functions such as apoptosis regulation and cellular metabolism (Chen et al., 2023; Kilbey et al., 2017). Given RUNX1's involvement in numerous developmental facets, continued exploration of its roles as a master transcriptional regulator remains imperative.

1.4.2. RUNX1's Function in the Hematopoietic System

RUNX1's initial role in human biology was discovered in relation to its involvement in acute myeloid leukemia, a disease characterized by the disruption of hematopoiesis (Gerritsen et al., 2019; Gonzales et al., 2021; Uchida et al., 1997) (Figure 1.8. A). AML most commonly affects patients over 60, but in rare instances can develop in pediatric patients anywhere from just a few days after birth to young adulthood, pointing to an origin beyond environmental exposure. Before the understanding that this was a homolog of the Runt protein of drosophila, this protein was given the nomenclature AML1, as it was believed to be a major driver of the disease (Uchida et al., 1997). In subsets of hematopoietic cancers, RUNX1 and its crucial binding partner CBFβ can undergo chromosomal breaks and fuse with other proteins, leading to constitutive misregulation of transcriptional activity (Liu et al., 2021) (Figure 1.8. A). This fusion has been shown to directly lead to increased proliferation, but incomplete differentiation of hematopoietic precursor cells, referred to as blasts (Saultz and Garzon, 2016). This accumulation of partially differentiated cells results in patients experiencing anemia, immunodeficiency, and infiltration of these blasts into bone marrow and organs, interrupting their normal physiological functions. Interestingly it has been observed that loss of CBFβ during hematopoiesis phenocopies the loss of RUNX1, cementing the

importance of both major components to this developmental process (Bresciani et al., 2014).

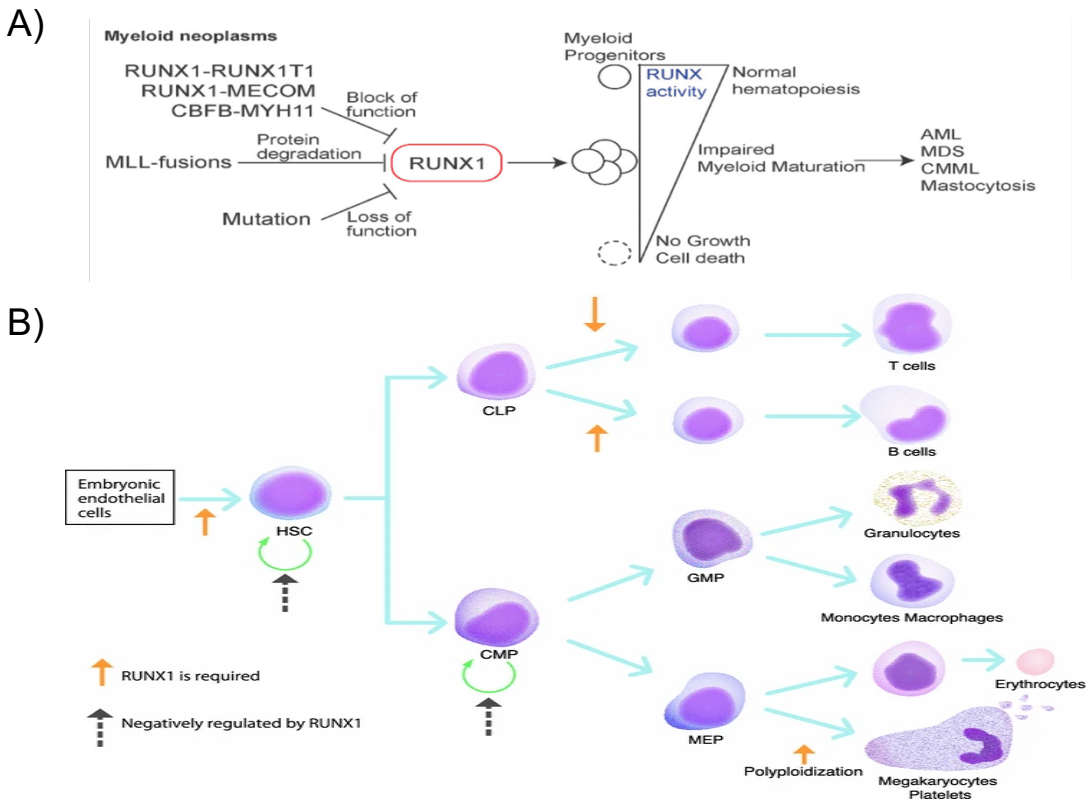


Figure 1.8. RUNX1 Roles in Hematopoiesis: A) Functional mutations to RUNX1 and their effects as seen in hematopoietic cancers, normally resulting in the accumulation of undifferentiated cells. Reprinted with permission from Goyama, S., Huang, G., Kurokawa, M. *et al.* Posttranslational modifications of RUNX1 as potential anticancer targets. *Oncogene* **34**, 3483–3492 (2015). B) Developmental diagram encompassing hematopoiesis as regulated by RUNX1. Reprinted with Permission from Ichikawa, M., Yoshimi, A., Nakagawa, M. *et al.* A role for RUNX1 in hematopoiesis and myeloid leukemia. *Int J Hematol* **97**, 726–734 (2013).

The pathology of diseases involving the misregulation of RUNX1 has offered insights into its pivotal role in hematopoietic system development. By identifying consistently inhibited cell types in AML samples with RUNX1 mutations, such as erythrocytes, leukocytes, and platelets, researchers gained a clearer understanding of the

spatial-temporal regulation of RUNX1 in system development (Bresciani et al., 2014; Friedman, 2009; Okuda et al., 2001) (Figure 1.8. B). Studies have demonstrated RUNX1's essential role in the initial specification of Hematopoietic Stem Cells (HSCs), integrating signals in the bone marrow microenvironment to maintain HSC and progenitor cell populations through self-renewal and multipotency preservation, utilizing its epigenetic and transcriptional functionalities (North et al., 2002) (Figure 1.8. B). These signals enable appropriate responses to cues directing self-renewal or differentiation, ensuring balanced hematopoietic cell production.

Since the discovery of and evolutionary correlation of RUNX, improved sequencing methodologies have shown that it is widely expressed throughout the body including epithelial tissue in the gastrointestinal tract (Usui et al., 2006), lung (Jeong et al., 2022; Tang et al., 2018), breast (Sokol et al., 2015), and brain (Fukui et al., 2018). Tumor sequencing analyses have suggested roles for RUNX1 in solid tissues, where loss of its expression leads to epithelial fidelity loss and increased aggressive mesenchymal phenotypes (Hong et al., 2017; Otálora-Otálora et al., 2019; Rose et al., 2020), emphasizing its role in tissue homeostasis. These findings collectively reinforce RUNX1's status as a master regulator of tissue development and regeneration in human organs.

1.4.3. RUNX1's Role in the Human Breast Epithelium

RUNX1's increased attention in other tissue types has allowed it to become the focus in mammary and breast tissue as well. Much of what we do know about RUNX1 comes from understanding how it functions in a malignant cell state. As with many

aspects of RUNX1 biology, its behavior is highly context-dependent. Notably, this context appears to diverge among breast cancer subtypes, particularly concerning the tumors' hormone receptor expression patterns (van Bragt et al., 2014). Mutations that cause loss of RUNX1 activity such as point, frameshift, or deletions are often noted in HR+ cancers, most of which are ER+ (Stender et al., 2010). Single-cell sequencing has revealed that ER+ luminal cells in the human breast typically lack RUNX1 expression, although chromatin immunoprecipitation experiments have indicated direct binding of RUNX1 to the ERα-encoding gene, ESR1 (Lachmann et al., 2010). Furthermore, studies suggest that RUNX1 contributes to stabilizing the epithelial phenotype and may reduce epithelial-to-mesenchymal transitions (EMT), metastasis, and tumor aggressiveness (Hong et al., 2017; Rose et al., 2020).

Conversely, high levels of RUNX1 expression correlate with a poorer outcome in TNBC. Unlike HR+ tumors, HR- tumors exhibiting elevated RUNX1 expression have significantly lower survival rates, attributed in part to RUNX1's role in driving proliferation, thus facilitating tumor aggressiveness (Fernández et al., 2023; Ferrari et al., 2014). Additionally, oncogenic programming may induce expression of RUNX1 cofactors that reshape its transcriptional network to favor tumorigenesis (Mercado-Matos et al., 2017). The dual nature of RUNX1, both as a tumor suppressor and an oncogene, underscores its pivotal role in maintaining homeostasis within the dynamic breast tissue. Studying RUNX1's involvement in breast diseases provides crucial insights into its normal physiological functions. Given its dual role as both a tumor suppressor and an oncogene, RUNX1 expression appears to be crucial for balancing stem cell self-renewal and differentiation into mature progeny (Otálora-Otálora et al., 2019; Sokol et al., 2015).

Recent investigations have demonstrated that perturbations of RUNX1 in breast progenitor cells prevents the exit from a stem-like state and prevented the formation of TDLUs (Sokol et al., 2015) (Figure 1.9.). These findings underscore RUNX1's necessity for cellular differentiation. Considering its widespread expression across various cell types within the breast (van Bragt et al., 2014), it is reasonable to conclude that it plays additional cell-specific role. Collectively, this data together indicates that there should be a greater focus on understanding RUNX1's role in the development of the breast and other tissues.

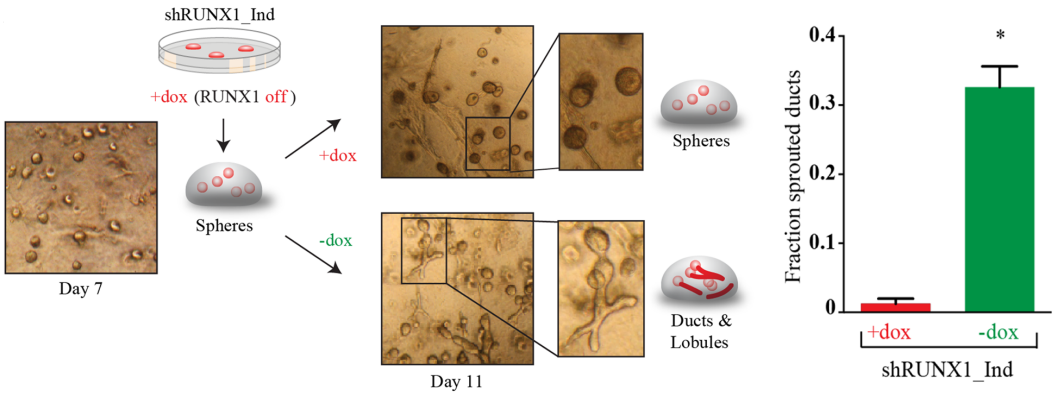


Figure 1.9. RUNX1 Roles in the Breast: Effects and quantification of RUNX1 knock down on primary breast stem cell regenerative capacity. Loss of RUNX1 prevents differentiation and proper morphology of ductal lobular structures in 3D. Reprinted with permission from *Sokol ES, Sanduja S, Jin DX, Miller DH, Mathis RA, Gupta PB. Perturbation-expression analysis identifies RUNX1 as a regulator of human mammary stem cell differentiation. PLoS Comput Biol. 2015 Apr 20;11(4):e1004161. doi: 10.1371/journal.pcbi.1004161. PMID: 25894653; PMCID: PMC4404314.*

A major purpose of the following studies was to help develop a greater understanding of RUNX1's role in the regulation of differentiation from facultative multilineage stem cells into mature basal and luminal cells. Through the utilization of a novel 3D hydrogel model able to grow primary single-cell derived breast structures, we

were able to draw a novel connection between the transmembrane RTK DDR1's binding of collagen and subsequent activation, and a change in RUNX1's transcriptome that drives differentiation.

Chapter II: Materials and Methods:

2.1. 2D and 3D Tissue Culture Reagents, and Assays:

2D Cell Culture of MCF10 Cells:
MCF10A (ATCC CRL-10317) and MCF10F (ATCC CRL) cells were cultured in DMEM/F12 (Corning) supplemented with 10 μg/mL insulin (Sigma), 20 ng/mL hEGF (E9644, Sigma Aldrich), 500 ng/mL hydrocortisone (Sigma), 100 ng/mL cholera toxin (Sigma-Aldrich), 5% Horse Serum (Gibco), and 1x Antibiotic/Antimycotic (Corning). The cells were routinely passaged every 2 to 3 days at a 1:4 to 1:10 ratio.

Primary Sample Preparation:
Primary tissue samples from healthy women undergoing reduction mammoplasties were dissociated using a 3 mg/mL collagenase (Roche Life Science) solution and sorted via density gradient to separate epithelial and stromal fractions. Single cells were obtained from the dissociation of epithelial clusters using 0.25% trypsin-EDTA (Gibco) and filtered through a 40 μm mesh filter after fibroblast removal.

Collagen Stimulation Assay:
MCF10A and MCF10F cells, seeded at densities of 1e6 or 2.5e6 cells for 10 or 15 cm plates respectively, were immediately treated with 2 μM DDR1 inhibitor DDR1-in-1 (Tocris 5077) or 5 μM RUNX1 inhibitor AI-10-104 (Aobious AOB17076). A solution of media, collagen type I (0.05 mg/mL), and 0.1 N NaOH for polymerization was applied during media changes 24 hours after initial seeding. After 24 hours of collagen stimulation, collagen and media were removed, and cells are detached with 0.25% trypsin-EDTA, and prepared for downstream assays. Cells were also grown in control and treatment conditions without the addition of soluble collagen for cellular baselines.

3D Hydrogel Model Culture:

A suspension of single-cell primary tissue samples or MCF10 cells was prepared with 1.7 mg/mL rat tail collagen I (Corning), 40 μg/mL laminin (Thermo Fisher Scientific), 20 μg/mL fibronectin (Gibco), and 10 μg/mL hyaluronic acid (Sigma), adjusted to pH 7.3 with 0.1 N NaOH. 200uL hydrogels were each plated in a four-chamber slide (Falcon) as a mold, polymerized for 1 h at 37 °C, and overlaid with MEGM medium (Lonza CC-3150) supplemented with 1x Antibiotic/Antimitotic, 1x Glutamax (Gibco). Organoids were grown from 18-28 days depending on regenerative rate of patient samples. Upon controls reaching maturity, structures were dissociated from hydrogel with collagenase and further dissociated using trypsin-EDTA, or were lysed and examined from mRNA or protein expression.

Inhibitor Time course:

Single-cell primary samples were seeded into polymerizing hydrogel mixture under three conditions: control, chronic DDR1, or chronic RUNX inhibitor treatment beginning either on day 0 (induction), or on day 7 (patterning). Samples were matured for up to 28 days and scored blindly for structure types. Upon reaching maturity, control structures were dissociated from the hydrogel using collagenase. Subsequently, the dissociation process involved trypsin-EDTA, or the structures were lysed for examination of mRNA or protein expression.

BRCA1 pγH2AX Assay:

Control and BRCA1 Mut. Primary samples were dissociated to single cell levels as seen above and were seeded into hydrogels of normal ratios. Samples were grown for 11 days and then were treated with 10uM, 60uM or 100uM concentrations of Doxorubicin. At

day 14 DXR treatment was removed, and structures were grown for another 4-7 days until growth had ceased. Structures were fixed and stained for Immunofluorescence.

Appendix Organoid Growth:
Appendix/Intestinal crypt cells were dissociated through rounds of EDTA chelation and vortexing to shake the cells loose from surrounding tissue. Crypt cells were collected and frozen back in "Basic Intestinal Media" formulated with DMEM/F12 (Corning), Glutamax (Gibco), and HEPES (Fisher Scientific) to support cell viability. Upon seeding in hydrogels, cells were supported with Intesticult media (Stem Cell Technologies) formulated with WNT3 and other necessary factors for intestinal development.

Primary Cell Transduction:
Primary epithelial cells, isolated as according to above protocol, were placed into a non-adherent 12 well plate at a volume of 100uL at varying quantities depending on viral MOI and scope of experiment. Virus in solution is added to a final concentration of 200uL per well. Cells are spun down at 2000 RPM and 30°C for 2 hours and then incubated overnight. Cells were washed and recounted prior to seeding into gel.

2.2. RNA Isolation and PCR:

RNA Isolation and Quantitative RT-PCR:
RNA extraction was performed on pelleted cells with the RNAeasy kit (Qiagen). cDNA synthesis was carried out using the iScript cDNA kit (Bio-Rad), and qRT-PCR was performed with Sybr green (Bio-Rad). The primer sequences for target genes are provided below.

qRT-PCR Primers:
ABCD4 Forward TGCGTACTACACCCTCAACG

Reverse TGAACGGGGAGATGATGAG

BCAR3 Forward TCAGGGATCCACATCTTCTG

Reverse CCAGCTCCTTCTTCAGTTTCTC

ZEB1 Forward TAAGAACTGCTGGGAGGATGAC

Reverse TCTGCATCTGACTCGCATTC

2.3. Protein Isolation and Assays:

Western Blot:
Cells grown in 2D or isolated from 3D hydrogels were pelleted by centrifugation and were fractionated using the NE-PER Nuclear and Cytoplasmic Extraction Reagents (Thermo, 78833), containing both 1x protease inhibitor cocktail (Cell Signaling Technology) and 1x phosphatase inhibitor (Cell Signaling Technology) according to manufacturer's protocol. Nuclear fraction samples were separated via NuPAGE gel (Invitrogen) and transferred to PDVF (BioRad) to be blocked in 5% BSA (Rockland). Blots were next incubated with primary antibody overnight at 4°C and with secondary antibody for 1 hour at room temperature. Immunoblot membranes were developed using a chemiluminescent substrate (Thermo Fisher Scientific) and imaged with the Chemidoc XRS+ with Image Lab 6.0.1 software (BioRad, Hercules, CA). ImageJ2 (Version 2.8.0/1.53t) was used to densitometry quantifications. Primary Antibodies used were: RUNX1 (4336, Cell Signaling Technology, Clone D33G6, 1:1000), CBFβ **(A303-549A,** Bethyl Laboratories, 1:1000), HDAC1 (5356, Cell Signaling Technology, Clone 10E2, 1:1000), H3 (9715, Cell Signaling Technology, 1:1000). Secondary antibodies used were

Goat anti-Rabbit (7074, Cell Signaling Technology, 1:1000) and Goat anti-Mouse (7076, Cell Signaling Technology, 1:1000).

Co-Immunoprecipitation:
MCF10A cells from 2D collagen stimulation assays were pelleted by centrifugation and lysed using 1X RIPA Buffer containing both 1x protease inhibitor cocktail and 1x phosphatase inhibitor. Cells were then incubated in immunoprecipitative antibody overnight at 4°C. Antibodies and attached proteins were conjugated to the magnetic beads at room temperature for 40 minutes. Samples were separated via SDS-PAGE gel and transferred to PDVF to be blocked in 5% BSA. Blots were then incubated with primary antibody overnight at 4°C and with secondary antibody for 1 hour at room temperature. Immunoprecipitative antibody was RUNX1 (HPA004176, SIGMA, 5ug/mg lysate). Primary antibodies were RUNX1 Ms (sc-365644, Santa Cruz, A-2, 1:1000), CBFβ Ms (67885-1, ProteinTech,1D7F2, 1:1000), and secondary antibody: Goat anti-Mouse (7076, Cell Signaling Technology, 1:1000).

2.4. Cell Staining:

Immunofluorescence:
Cells and hydrogels were fixed with 4% Paraformaldehyde (Fisher Scientific) for 30 minutes, permeabilized with 0.1% Triton 100X for 24 hours, and incubated at 4°C for 18 hours with primary antibodies: E-Cad (13-1700, Thermo Fisher, HECD-1, 1:100) CK-14 (RB-9020, Thermo Fisher, 1:300) p γH2AX (JBW301, Millipore, 1:500), LGR5(C-16, Santa Cruz, 1:200). Samples were then incubated at 4°C for 18 hours with secondary antibodies: DAPI (D1306, Life Technologies, 1:1000), AF488 (A11008,

Invitrogen,1:1000), AF555 (A21424, Invitrogen,1:1000), Phalloidin-AF647 (A22289, Invitrogen,1:500), and placed on slides with anti-fade reagent.

Flow Cytometry:
Structures derived from patient samples, either control or inhibited during patterning at Day 7, were dissociated from hydrogels and from structures and then were pelleted by centrifugation. Samples were washed and stained with the antibodies CD49f-FITC (555736, BD Biosciences, GoH3, 1:20) and EPCAM-PE (347198, BD Biosciences, 1:20). Samples were run on LSRII. FlowJo (Version 10.9.0) was used for visualization and quantification.

2.5. Microscopy and Long Term Multipoint Live Imaging

Microscopy:
Immunofluorescence images were captured using Nikon AXR (Nikon Microscopy) using NIS-Elements software. Brightfield images were obtained using Nikon AXR and Eclipse Ti-U (Nikon Microscopy), utilizing SPOT 5.6 software.

Live Imaging:
Primary single cells, isolated as previously described, underwent incubation with the cell tracking dye Cytopainter Green (1:500, cat# ab138891, Abcam) for 30 minutes. Following washing, cells were seeded at a concentration of 100 cells per 20 uL hydrogel. The gel fabrication process involved depositing 20 uL hydrogel drops onto the center wells of a 96-well plate (Corning, #3603). After allowing the gels to incubate for one hour at 37 °C until fully polymerized, 80 uL of MEGM was added to each well, and the gels were gently lifted off the well surface with a pipette tip. The cultures were promptly

placed in a pre-warmed incubator chamber (Okolab Inc) at 37°C and containing 5% CO_2 enclosed over a Nikon Eclipse Ti2-AX confocal microscope (Nikon Microscopy). Images were collected from selected points immediately after the addition of media, and every 30-45 minutes thereafter, using both brightfield as well as AF488 or AF555 laser, using 4x magnification and 2.5x zoom across nine z-positions. To maintain optimal growth factors and liquid volume, 20-40 uL of MEGM was added to the culture twice a week to prevent hydrogels from drying out. The cultures underwent live imaging for 18-21 days, and the analysis and production of videos across locations and timepoints were conducted using NIS-Elements (Nikon) and Premiere Pro (Adobe) software.

2.6. Sequencing

RNA-Seq:
mRNA isolated from MCF10A cells in a collagen stimulation assay was run on the Illumina NextSeq 6000.

SC-RNA Seq:
Previously published scRNA-seq data from organoids (Rauner et al., 2021). scRNA-seq data were analyzed using Seurat v3 (Stuart et al., 2019) for data integration, normalization, and feature selection. Briefly, raw data was loaded and integrated into one Seurat object using the merge function. Filtering removed cells with < 200 or >2500 genes and mitochondrial content greater than 7.5%. Genes detected in less than 3 cells were dropped from analysis. The data was normalized by multiplying transcripts by a factor of 10,000 followed by log-transforming the data. Variable features used for analysis were identified by using the FindVariableFeatures function, with a low cutoff of

0.0125 and a high cutoff of 5 for dispersion and a low cutoff of 0.1 and a high cutoff of 0.8 for average expression. The data was integrated by the FindIntegrationAnchors and IntegrateData functions, which identify the anchors to integrate the two datasets, and then integrates them together. Cells were then clustered using K-nearest neighbor (KNN) graphs and the Louvain algorithm using the first 10 dimensions from principal component analysis. Clustered cells were visualized by tSNE embedding using the default settings in Seurat. Clusters were called using the FindClusters function with a resolution of 1. To identify differentially expressed genes between cell clusters, we utilized the FindAllMarkers function to identify features detected in >10% of a cell cluster compared to all other cells. Pathway analysis to identify enriched biological pathways associated with differentially expressed genes was done using established databases, such as PanglaoDB (Franzén et al., 2019). The top 15 differentially expressed markers were used to determine gene expression location.

2.7. Statistics and Data Analysis

RNA-Seq Analysis:

Read Alignment to the human genome conducted using STAR (Dobin et al., 2013) with the CRCh37/hg19 assembly. Library normalization and differential expression analysis was performed using R with DESeq2 (3.17) (Love et al., 2014). Sample SSR107 was removed as it was deemed an outlier by PCA. Differential gene analysis was conducted with significance determined by log fold change greater than or less than 0 with a p value of less than 0.05. Gene Ontological analysis was performed using the ChEA dataset (Lachmann et al., 2010) on Harmonize (Version 3.0) and with the TRRUST (Han et al., 2015) dataset from Enrichr (Xie et al., 2021). Heat maps and Venn Diagrams were

created in R using Pheatmap (Kolde, 2023) and ggVennDiagram (Gao et al., 2021), respectively. Protein-Protein Interactions generated in STRINGDB (Szklarczyk et al., 2022).

Mutational Analysis:

RUNX1, CBFβ, and DDR1 were queried for alteration frequency in breast tumors on cBioPortal using primary BC data from 20 studies broken down by cancer type (Cerami1 et al., 2012). Mutational status across breast cancer subtype by PAM50 as well as exploration into mutational mutual exclusivity were also conducted using these studies on cBioPortal. Kaplan Meier survival curves were produced using METABRIC (Curtis et al., 2012) data from the Breast Cancer Integrative Platform (BCIP) (Wu et al., 2017), plotting overall survival based upon transcriptome analysis of either a triple-negative status or non-triple-negative (i.e., the expression of at least one receptor ER/PR/HER2).

Statistics:

All statistical analyses were performed using GraphPad Prism 8-10. Student's t-tests (two-sided) were used as a determinant of significance unless otherwise stated. Data were expressed as Mean ± SD. Significance levels were indicated as follows: * = p-value < 0.05, ** = p-value < 0.01, *** = p-value < 0.001, **** = p-value < 0.0001.

Chapter III: Utilization of a Primary Single Cell Derived 3D Breast Model

3.1. Introduction:

The human breast, and its anatomical analog the mammary gland, are comprised of ductal and lobular structures responsible for the milk production that is vital for offspring nutrition (Russo and Russo, 2004). Unlike many other tissues, the breast does not reach full development during fetal stages but undergoes cyclic phases of growth and involution throughout a woman's life (Russo and Russo, 2004). This regenerative ability suggests the presence of facultative stem cells that respond to signaling cues by recapitulating tissue, generating both basal and luminal cell types essential for functionality. Transplantation studies have shown that mammary cells are able to reconstitute the mammary gland when put into the proper environment (Kuperwasser et al., 2004; Lee et al., 1984; Li et al., 1987), however, the growth of human tissue in biomimetic conditions has historically been difficult. Thus, little is understood about the stem cell compartment of the human breast and how they can differentiate and form tissue. This lack of knowledge impedes a comprehensive understanding of breast development.

As 2D cell culture and animal models do not recapitulate the structure and cellular intricacies of breast tissue, few investigative steps have been able to be made towards understanding breast stem cell differentiation. Large steps have been taken over the last four decades, however, to create a model able to answer such questions. Initial studies focusing on the ex vivo growth of mouse mammary tissue proved that 3D tissue can develop in the proper extracellular matrix conditions (Lee et al., 1984; Li et al., 1987), as well as that each of these ECM components plays a role in both structural and signaling aspects of tissue development (Roskelley et al., 1994). Although primary human tissue

did not form with morphological accuracies within these matrices, immortalized human cell lines demonstrated the capability to develop hollowed acinar structures, providing a partial representation of human tissue growth and a model of 3D human tissue interacting with a matrix (Debnath et al., 2003). The humanization of mouse mammary fat pads provided another avenue for the growth of human breast tissue (Kuperwasser et al., 2004). However, due to limitations inherent in each of these models, it was not possible to study the initiation and development of human breast tissue.

Recent advances in 3D modeling in the last decade have allowed for the growth of primary human tissue in hydrogels, allowing for larger control over variables and observational ability when studying tissue development. Patient derived cells cultured in these conditions were able to differentiate and reconstitute breast structures that replicated those seen in human tissue, containing both ductal lobular structures and basal and luminal layers in the proper orientation (Linnemann et al., 2015, 2017; Sokol et al., 2016). While some of these structures relied on components not normally localized in the breast, one hydrogel model, one was able to reconstitute this biomimetic tissue utilizing minimal factors. Containing only collagen type I, hyaluronic acid, laminin and fibronectin, in concentrations that best mimic the ECM conditions of the human breast, and then immersed in a basic epithelial medium containing minimal supplements, this model supported the formation of miniaturized breast tissue (Sokol et al., 2016). Utilizing clusters of breast cells taken from healthy donors as organoid initiating factors retained not only representative structure, but also mRNA expression, and expression of hormone receptors like that of the human tissue it was taken from.

While addressing numerous questions regarding human breast development, there are still opportunities for further refinement of the model to enhance the depth of knowledge gained from each experiment. This thesis work advances the model by generating breast tissue organoids from single cells rather than cell clusters, as first demonstrated by Sokol et al., thereby improving its reproducibility across experiments and patient samples, and enabling the recapitulation of earlier developmental stages. Moreover, this work illustrates the feasibility of genetically modifying primary cells that generate organoids, furthering our understanding of the processes by which stem cells regenerate breast epithelial glandular tissue. Collectively, this data underscores the potential of the 3D hydrogel model, especially when combined with primary single cell derived organoids, to accommodate experiments of increasing complexity aimed at unraveling tissue development.

3.2. Results:

3.2.1. Characterization of a Primary Single Cell Derived Organoids:

The ability to seed primary human tissue in 3D cell culture matrix, allowing for it to autonomously regenerate into biomimetic tissue models that can be replicated across primary samples and requires minimal factors for growth was an incredible leap in the field that has already proven to allow for new discoveries important in human tissue (Rauner et al., 2021; Sokol et al., 2015). However, the initial units of these biomimetic structures, formed from small cell clusters, can introduce inconsistencies among replicas due to variations in starting populations. For instance, differing cell counts, i.e. more or less progenitor cells at various stages of differentiation, may lead to disparate growth

rates among replicas. What's more, the use of primary samples from multiple patients of different ages, BMIs, and parity, further confounds the ability to create accurate quantifications. The reduction of the primary breast organoids starting unit from a cluster of cells to a single cell has increased the utility of this 3D hydrogel (Figure 3.1.). Primarily, the earlier stages of TDLU organoid development can now be examined and quantified across conditions and patient samples through more controlled starting conditions.

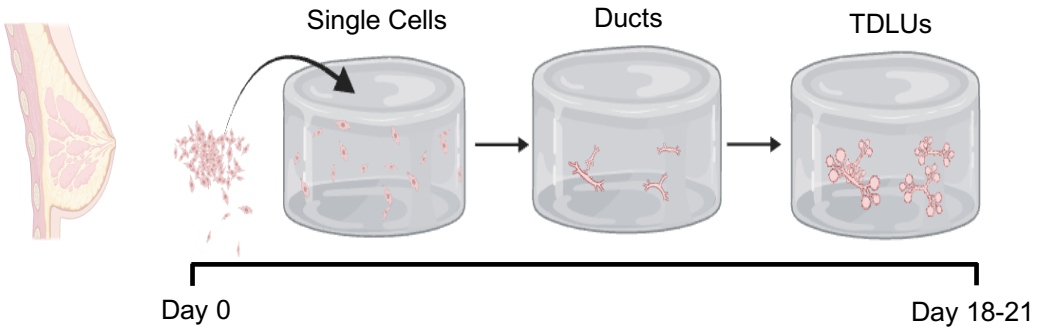


Figure 3.1. Schematic Representation of Human Breast Organogenesis in a 3D Hydrogel Organoid Model: A) Graphical Representation and summary of human breast organogenesis in a 3D hydrogel TDLU organoid model.

When single primary breast cells are seeded in the hydrogel, they form three different structure types; alveolar organoid structures comprised of irregular or rudimentary clusters of cells that do not expand outward into the matrix and lack distinct architecture, ductal organoid structures made of extended branched structures that have clear architecture but do not form alveoli, and compound organoid structure that contain elongated branched structures that form defined alveoli at the ends and along the branching ductal structures (Figure 3.2. A). Seeding 1000 single cells into a 200 uL hydrogel allows for the formation of an average of 11 total structures per gel across multiple patient samples, leading to an average of 1 initiating cell per 91 cells seeded

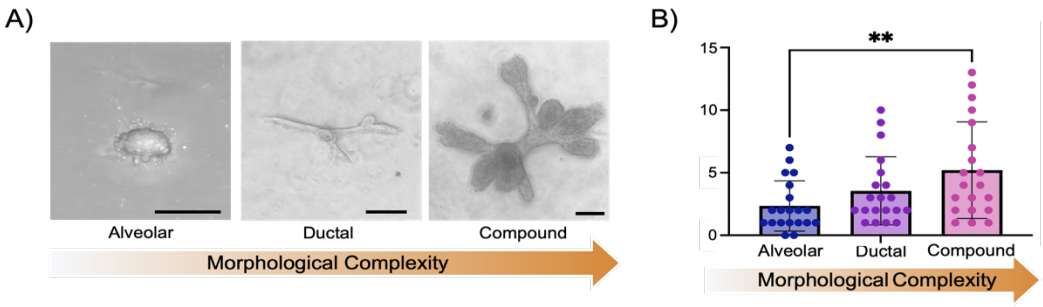


Figure 3.2. Characterization and Quantifications of Structures Derived from Single Cell Breast Organoids: A) Representative brightfield images of structures formed during organogenesis in 3D hydrogels when initiated with single cells. Scale bar = 100 µm. B) Quantification of 3D organoid morphologies from 5 primary patients. Mean ± SD (n= 4 gels/primary patient samples).

(Figure 3.2. B). These cells do not form each structure type at an equal rate, forming a mean of 2.35 ± 2 alveoli structures, 3.55 ± 2.8 ductal structures, and 5.2 ± 3.8 compound structures**.** While there is a significant difference in the number of alveolar and compound organoids formed, neither is significantly different from ductal organoids.

Through use of immunofluorescence staining, characteristics of these structures can be further identified as they develop into complex TDLU structures. Interestingly, these structures, despite being formed from a single cell, can produce structures of multiple defined cell types (Figure 3.3.). The expression of CK14 (a marker of basal cells) and E-Cad (a marker of luminal cells) shows that these two cell types not only exist within these structures but also create two clearly defined layers in which the $CK14^{+}$ basal cells form an outer layer while the densely packed $E\text{-}Cad^{+}$ layer is interior to the basal population, lining the hollow lumen of the structures. Interface between these two layers can be seen by the dual staining of the structure cross-section. This data indicates that single primary breast cells can generate both basal and luminal layers and structures

of a compound ductal-alveolar organoid, and together this implies that there is an epithelial cell in adult breast tissue capable of acting as a facultative stem cell.

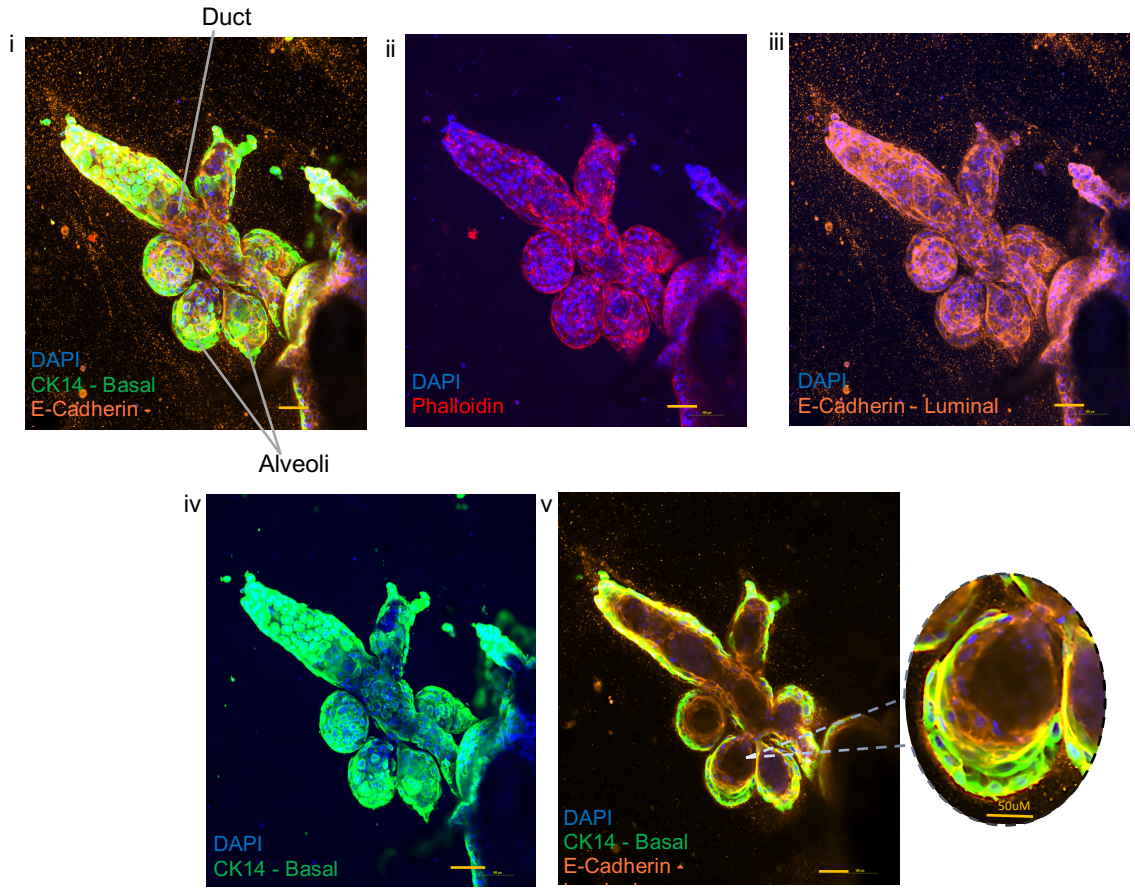


Figure 3.3. Immunofluorescent Characterization of Single Cell Derived Breast Organoids: i-iv) Immunostaining of a Three-dimensional Maximum Intensity Projection (MIP) TDLU organoids, at day 21 of development, stained with phalloidin (red) for actin cytoskeleton, with CK14 (green), E-cadherin (orange) for cell lineage, and with DAPI (blue) for nuclei. Scale bars = 100 μm v) Two-dimensional cross-section of MIP (i) with offset highlighting a lobule showing distinct $CK14^{+}$ (green) and $E\text{-}Cad^{+}$ (orange) layers with an interacting layer displaying yellow. Scale bar = 100 μm, scale bar of offset = 50 μm.

The ability for this hydrogel to support the growth and maturation of primary healthy breast tissue provokes inquiry into if it can also support the growth of tissue harboring mutations. Using samples containing a well-known driver of many TNBCs, the

BRCA1 mutation (Friedman et al., 1994; Miki et al., 1994), could allow researchers to study how these cells mitigate the double strand break (DSB) repair the BRCA1 protein is responsible for in both healthy and mutated tissue. As these single-cell-derived structures are directly comparable, time course analysis during development could allow for the quantification of DSB repair across developmental time points and even after maturation. Some preliminary work has already been conducted attempting to monitor differences in DSB repair across TDLU development. BRCA1- mutant primary samples grown from single cells form structures in the same general timeline of physiologically normal tissue (Appendix Figure 6.1. A, i-ii). DNA damage was induced through doxorubicin (DXR) treatment starting at day 11 until day 17, upon reaching morphogenesis, and was fluorescently visualized for integration into the structure on day 17 (Appendix Figure 6.1. B, i-iii) (Karukstis et al., 1998). DXR treatment promoted DSB, as shown by DNA damage marker phospho-γH2AX in both healthy and mutated tissue (Appendix Figure 6.1. C). 5-11 days after treatment, it was determined that while control tissue retained sites of DSBs it continued to develop, unlike BRCA1 mutant tissue which was not able to withstand the treatment (Appendix Figure 6.1. D). While further work needs to be done to optimize this assay for both treatment and timing, this offers some hope that only small adjustments are necessary for the development of an informative assay to compliment other known DNA-damage assays through visualization and quantification.

This work also raises the question of this hydrogels support of organoids from other tissue sources. As this model is easily tunable, allowing for the addition of other ECM components or signaling molecules, it is possible that this simple matrix may

support the life and morphogenesis of other tissue types. Some preliminary work in the growth of appendix tissue has already been completed. Here it was shown that without any modulation of the tissue or model, single crypt cells from healthy appendix tissue was able to reconstitute hollow tube like that of intestinal or appendix tissue (Appendix Figure 6.2. A-C, Video 6.1.) (Berry, 1900). Further work needs to be done to optimize conditions for the acquisition and culture of this tissue in 3D. This data, supported by the growth of both healthy and mutant primary breast samples, underscores the considerable potential for experimental versatility afforded by this hydrogel, and will hopefully allow for a deeper understanding of tissue development.

3.2.2. Long Term Multipoint Live Imaging:

To better understand the process of organogenesis in the context of human breast development, a novel confocal-based long-term multipoint live imaging technique was utilized to gain increased resolution of cellular motility by following the initiation and morphogenesis of epithelial breast tissue. By taking images of single cell structures immediately post-polymerization of the hydrogel (+1 hour post seeding), and every 30-45 minutes thereafter for up to 28 days, the true origin of the organoid can be observed. Utilization of this technique allows for over 150 different individual points to be monitored over this time, with adjustments to resolution, number of gels, and number of images taken along the z axis allowing for more points or higher resolution during this time.

Initially it was believed that this organogenesis would occur akin to that of a tree, with the main body developing, elongating, branching from this main body, forming

secondary and tertiary branches, and finally forming leaves. What was discovered instead was that these structures go through 4 distinct periods of development while forming a complex TDLU organoid containing both ducts and alveoli: induction, patterning, morphogenesis, and maturation (Figure 3.4., Video 6.2.).

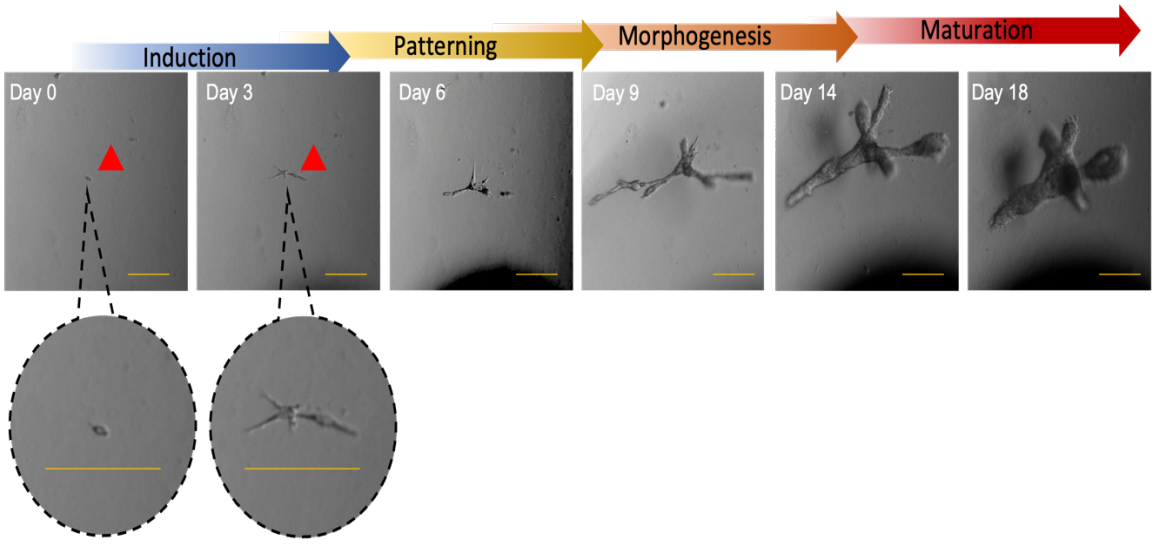


Figure 3.4. Breakdown of Breast Organoid Developmental Stages: Timeline with brightfield images of TDLU formation from single cells, starting on day 0 immediately after hydrogel polymerization to the formation of complex ductal-lobular structures from day 18 onward. Four phases of organogenesis showing induction, patterning, morphogenesis, and maturation. Scale bar = 100 µm.

During the induction phase, which occurs during days 0-4, initiating cells seemingly begin to sense their environment by moving through the ECM and interacting with the matrix, which likely signals for proliferation and the initiation of cellular differentiation. Two different types of induction have been observed: Type I with single cells that divide and move fluidly throughout the matrix (Video 6.3.), and Type II where tightly packed cells form a spheroid structure and eventually release individual cells into their microenvironment (Video 6.4.). This motility continues into the next phase, the patterning phase, which occurs during days 5-8. During this timeframe, proliferative cells

take an almost mesenchymal approach to organogenesis through a loose adherence to one another as they rapidly move through the matrix. As the cells continue to divide, differentiate, and reach a critical mass, they begin to rapidly invade the matrix, moving back and forth from the loosely adhered main cell mass. The paths the cells take will eventually become the areas in which developing ductal branches will emerge.

The progression through patterning occurs before the formation of an interconnected and unified tissue, which begins to occur around day 9. At this point, cells begin to go through the process of morphogenesis, in which cell-cell adhesion begins to increase until the once mesenchymal-like cells are now beginning to form a cohesive tissue. As the cells have now settled into an identity of not just an individual cell but that of a tissue, they begin to form the elongated ductal structures. These structures will continue to elongate and branch, eventually leading to the formation of alveoli buds along and on the ends of the branched ductal structures. Finally, around day 14, organoids enter the final stage of development, maturation, in which the compound structure contracts as it creates cell-cell adhesion and forms a cohesive tissue. However, these structures remain dynamic as coordinated cell movements still are occurring throughout the structure, with movement of cells throughout the ductal structure, and a rotational movement within the alveoli (Video 6.5.). The process of maturation finalizes with the eventual hollowing of the lumen for a functional tissue.

During these live-imaging experiments, we observed two previously unreported cell types. The first of these cell types was dubbed a type of leader cells, which were previously described as the cells involved in the invasion of the matrix. While cells involved in the elongation of the ductal network have been previously identified as

"leader cells" (Sokol et al., 2016; Varner and Nelson, 2014; Wang et al., 2017), the cells dubbed leader cells in our experiments were observed to be more mobile and with less cell-cell adhesion during early stages of development. During patterning, it was observed that these cells not only invade into the matrix but also completely detached themselves from the main structure body at the time (Video 6.6.). The ability of these cells to realign themselves with the main cell body in repetition implies a cell type with both epithelial and mesenchymal characteristics that can help direct where the structures should be developing in a 3D space. Further work is needed to understand if these cells are signaling for other cells to follow or modify the matrix through the release of MMPs and other matrix-modulating enzymes.

The second cell type seen that has not been previously described in the breast tissue were highly motile cells that seem to form a mesenchyme. These cells seemingly develop at a later developmental time point towards the end of morphogenesis and the beginning of maturation and are ejected at mass volume from the tissue structure itself and are derived from the preliminary epithelial population initially seeded into each gel (Video 6.7.). This mesenchyme divides rapidly and causes major changes to the microenvironment, including gel contraction as they infiltrate through and reach the exterior of the hydrogel. It can be noted that their presence causes major tension on the hydrogel itself seemingly pulling collagen fibers inward during expansion (Video 6.8.). While seen often, their presence is not necessary for the formation of a TDLU as many compound organoids do not require their presence for full maturation. Further work needs to be done to confirm if this cell population is a remnant of the leader cells who

have received incorrect signaling, or if they are truly an important player in organogenesis.

3.2.3. Lentiviral Transduction of Organoid-Forming Cells.

One of the most important benefits of the generation of organoids from dissociated single cells is the ability to genetically modify cells before they go on to form an entire structure. This contrasts with organoid generation from clusters of cells that already have a 3D architecture and therefore may not be able to undergo complete viral transduction, due to viral penetrance through tissue as well as the low probability that every cell would receive the vector. If all cells were not infected by a virus, there is a possibility that this initiating cluster may not generate a fully transfected organoid. In theory, if these structures are derived from a single cell, any genetic modifications made to the cell prior to its seeding would allow for all cells in the organoid to carry the same changes through its intrinsic clonality. Indeed, this was seen to be true in initial experiments using the Lentiviral Gene Ontology vectors (LeGO). Using multiple vectors, allowing for the expression of fluorescent proteins Venus, mCherry, and Cerulean, primary cells were infected either with an individual vector, or all three at a MOI = 1.5, 7.5, or 15, to determine if a single cell was responsible for the derivation of an entire structure. It was observed that organoids can form as one color but can also form a structure with multiple fluorescent proteins, either with multiple fluorescent types in one cell or with different cells in the structure exhibiting different fluorescent colors. Use of a non-LeGO GFP vector confirms that these structures can form and retain genetically driven fluorescence.

This data points to multiple potential ways in which breast organoids can be formed, either initiated through one single cell that develops a one-colored structure (Figure 3.5. A), through structures initiated with multiple cells (Figure 3.5. B), or even

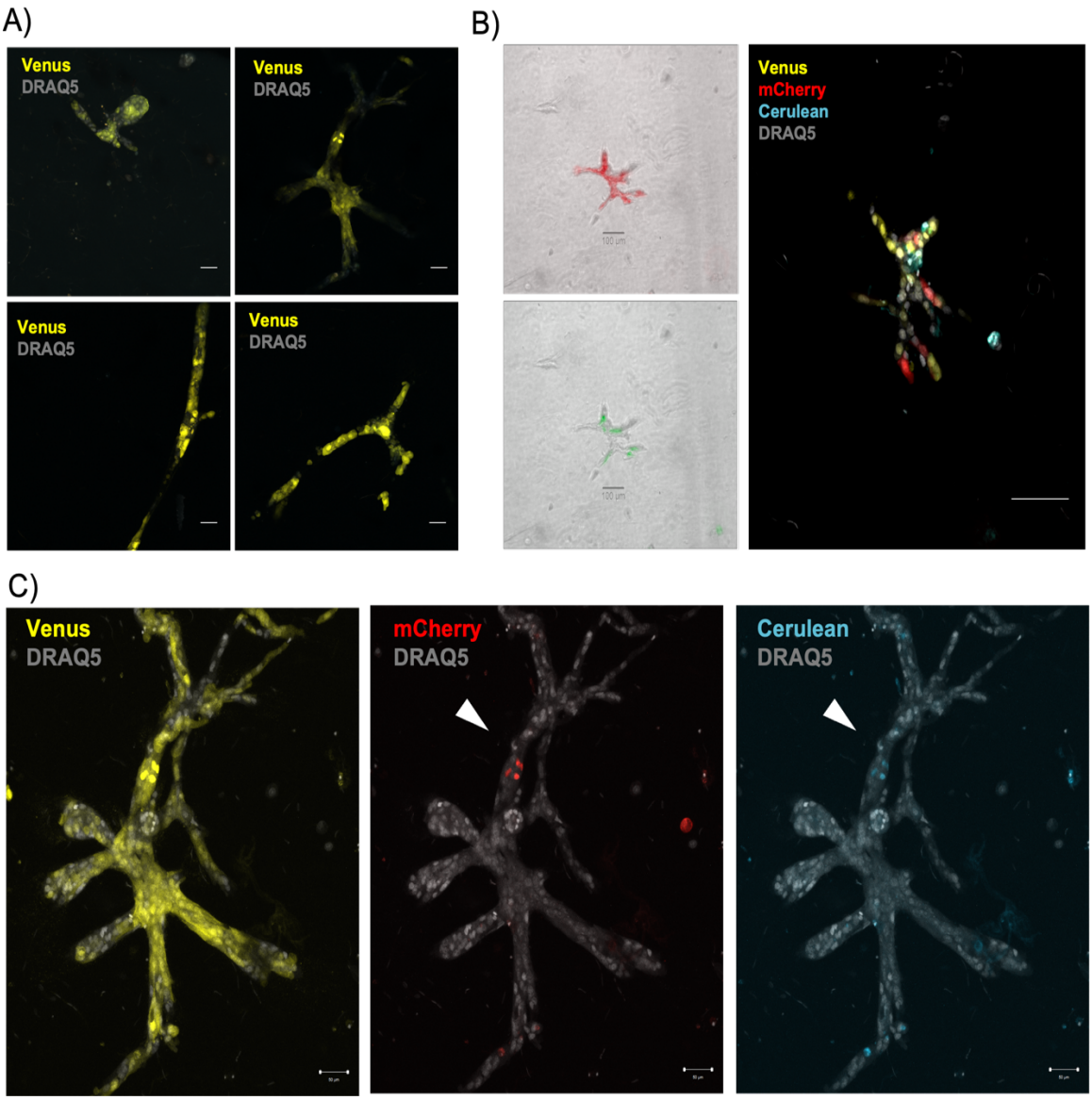


Figure 3.5. Fluorescent Patterns from Lentiviral Transduction of Single-Cell-Derived Organoid: A) Representative images of structures stained with live cell stain DRAQ5 and transfected to express Venus fluorescent protein with MOI of 7.5 showing fully infected structures. Scale bar = 50 µm B) Representative brightfield and fluorescent images showing a structure expressing all three vectors and stained with DRAQ5. Scale bar = 100 µm. C) Representative images of structures stained with live cell stain DRAQ5 and primarily expressing Venus, but also containing some cells expressing all three vectors (White Arrows) Scale bar = 50 µm.

potentially through the engulfment of non-dividing cells during the ductal expansion process (Figure 3.5. C). Live imaging gives us hints that structures can arise in all three ways, as it was seen that some structures arise from small cluster of cells while others generate through the efforts of a single cell. While done with primary tumor tissue (Hanna et al., 2016; Hines et al., 2015), and primary tissue ex vivo (Pastor et al., 2023), however this data acts as proof of concept that healthy primary tissue can both be infected after being dissociated down to single cells and then still regenerate a biomimetic structure while retaining the expression of their genetic modification.

With the success of the simple fluorescent vectors, more complex viral vectors were used to further show the utility of the transfection of primary single-cell organoids. The FUCCI vector has been previously used to model cell cycle changes in cell lines (Zielke and Edgar, 2015). As the cell cycle progresses, the FUCCI vector changes the expression of its fluorescent protein from red (G1) to yellow (G1/S) to green (S/G2/M). This is accomplished by fusing fluorescent vectors to DNA replication factors like DNA replication initiator Cdt1 (Red) (Nishitani et al., 2001) or inhibitors like Geminin (green) (Wohlschlegel et al., 2002) that are universal in their activation at specific timepoints during the cell cycle. Through transfection of initiating single primary cells in combination with live imaging techniques, it could be observed where and when certain cells, and eventually certain areas of the tissue, are dividing and going through the cell cycle compared to more quiescent areas of tissue (Video 6.9.). While no discernable pattern could be determined from the pilot study, as there seemed to be dividing and non-dividing cells throughout the structure at all times, it could be a tool of great importance

for understanding both the development of organoids through multiple stages and the effects of perturbations on the cell cycle.

A NOTCH1 activity reporter (Hansson et al., 2006), which create fluorescence in response to the cleavage of NOTCH1 intracellular domain, was also used to show the functionality of both the 3D organoid and the transfection of primary single cell. NOTCH 1 is a transmembrane signaling receptor that binds its ligands expressed on the cell surface of adjacent cells. NOTCH1 regulates polarity in tissue layers, and was shown to be necessary for the differentiation and proliferation of the luminal population and alveologenesis in TDLU organoids initiated with clusters of cells (Rauner et al., 2021). As previously reported, NOTCH1 activity occurs only in the luminal layer of tissue as it is activated by JAGGED1 (JAG1), the most commonly expressed NOTCH ligand in the breast (Rauner et al., 2021), which is exclusively expressed in basal cells. The interaction between NOTCH1 and JAG1 is believed to be the driver of both luminal and basal cell maturation.

Initial testing with the NOTCH1 reporter shows that single cells transfected with this vector can indeed derive structures with NOTCH1 Activity based fluorescence. What's more, the activity of the NOTCH reporter seems to be more internally expressed, supporting the idea that NOTCH1 activity takes place in the luminal population (Figure 3.6.), adding evidence to previously research. Combining this method with live imaging techniques shows that the swirling seen in the alveoli during maturation is occurring in the luminal layer (Video 6.10.). Together these results show the great utility in being able to address organogenesis-based questions through transfection of primary cells.

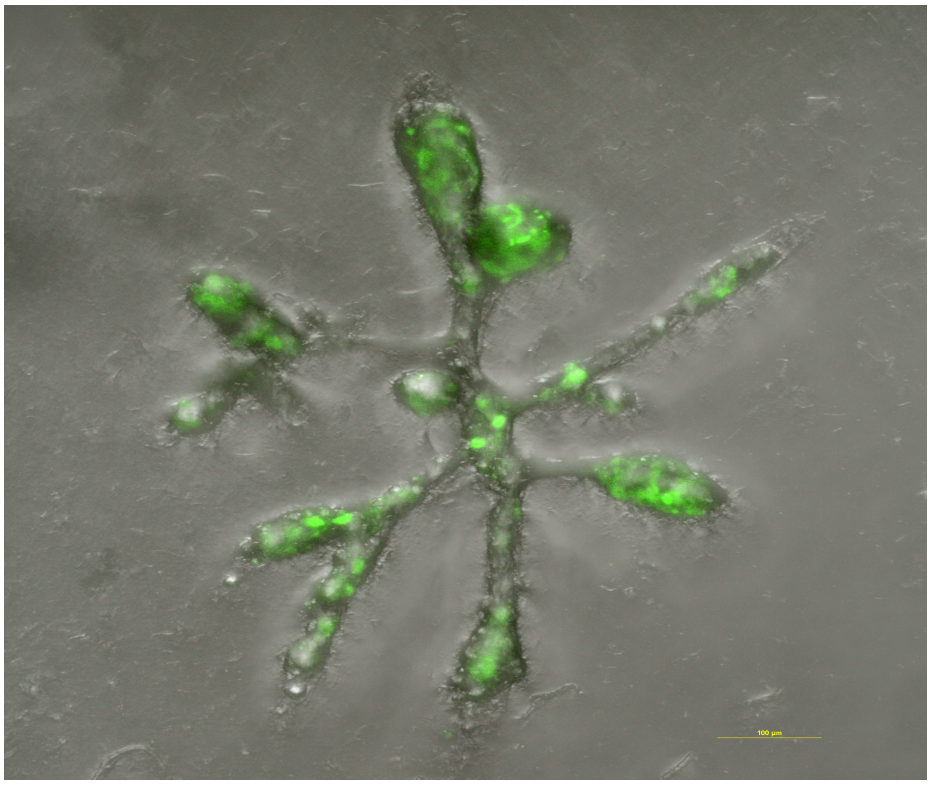


Figure 3.6. Fluorescent Pattern of a Notch Activity Reporter in a Single-Cell-Derived Organoid: Representative images of a structure transfected with a NOTCH1 activity reporter fluorescing green where NOTCH1 is transcriptionally active. Scale bar = 100 μm.

3.3. Discussion:

The use of primary single cells, over initiating clusters of cells, to form organoids does not create a different or better breast TDLU than one initiated with cluster of cells but does greatly add to the utility of the model in multiple ways. One of the most important ways this model has improved is through increased replicability. The use of single cells allows for controlled quantities of starting units being embedded in the hydrogel. Compared to clusters of cells, which greatly varied in the number of cells and differentiated states, single cells were much to quantify in exact concentrations. This ability to control for growth biases in gels allows for better comparison across conditions

with experiments that include drug treatments. This aspect is particularly crucial when comparing various patient samples, as heterogeneity in numerous factors can lead to variations in the initiating units for the TDLU. For example, it has been previously shown that luminal differentiation decreases with age (Garbe et al., 2012), which could create an unseen imbalance when using clusters of cells from older patients. The capacity to compare findings and results across phenotypic levels, encompassing morphological and expression-based assessments, facilitates studies aimed at mitigating biases from the use of a singular cell line or animal model. By enhancing genetic diversity, such an approach increases the likelihood that findings will be applicable across multiple patient cohorts and populations.

Because of this single-cell model, the true initiating unit of this biomimetic breast organoid can also be further studied, allowing researchers to see the true triggering of the developmental processes through to the facultative stem-like nature seen in these single cells. Reduction of starting material allowed for the quantification of epithelial cells in the adult breast retaining the ability to regenerate a structure. About 1/91 cells were able to form an organoid, however, this means that around 1/182 were able to form complex ductal lobular structures (5.2± 3.8 compound organoids per 1000 cells seeded). This retention of stemness only further shows how vital the ability to regenerate the TDLU is, further corroborated by the fact that structures trend towards becoming complex compound structures compared to alveolar or ductal structures. The internal cellular programming of a single cell retains the regenerative capacity to from a compound structure more often than an incomplete one, showing the intrinsic nature driving the development of mature TDLUs. This also provokes the question, what is happening that

prevents other structures from being able to form compound structures. Research shows that this could be due in part to issues with signaling between the matrix and epithelial cells (Rauner et al., 2021), but there are other cell innate factors such as that could play a role in the morphological discrepancy. While not done in a fully biomimetic model, it has been shown that primary luminal cells grown in other 3D hydrogels produce alveoli, while basal and progenitor cells produce ductal and ductal lobular structures (Linnemann et al., 2015). Future work through knockout models and lineage tracing should be aimed to determine what factors are involved in the generation of all structure types in this model.

Through live imaging experiments, we observed the remarkable capability of a single cell to function as a facultative stem cell, undergoing division and progression through multiple developmental stages immediately after seeding. This pioneering approach yielded invaluable insights into the growth dynamics of these structures; a phenomenon previously unexplored by live imaging experiments utilizing a single cell as the initiating factor. By accurately defining developmental stages of this breast organoid, including induction, patterning, morphogenesis, and maturation, we were able to pinpoint precise moments of structural programming shifts, significantly enhancing our understanding and experimental reproducibility in studying these organoids. Furthermore, this methodology unveiled previously unknown cell types and functions, expanding our knowledge base. While seen in mice (Cowin and Wysolmerski, 2010; Veltmaat et al., 2006), this live imaging provided the first visual documentation of human primary breast cells undergoing induction and patterning, revealing novel cellular behaviors and previously uncharacterized cell types involved in tissue regeneration. Contrary to

expectations, initiating cells exhibited dynamic behavior within the hydrogel, suggesting environmental sensing, matrix clearance, and intercellular interactions. Patterning analyses unveiled a departure from the anticipated contiguous duct formation, with initial branching facilitated by loosely adhered cells clearing the collagen matrix, paving the way for subsequent morphogenesis and elongation of ductal structures.

The defining discovery during induction and patterning was the activity of these "leader cells". Although similar to cell types that have been described before (Sokol et al., 2016; Varner and Nelson, 2014; Wang et al., 2017), the nature of the cells we observe in experiments utilizing single cell initiated structures seems to be much more active at the earlier stages of induction and patterning, before the sprouting of ducts when canonical leader cells push forward ductal elongation. In reality, these cells seem more akin to trailblazer and opportunist cell states described in the invasion of cancer, in which specific cells (trailblazers) contain the capacity for movement through the ECM, and opportunists follow in close succession (Pearson, 2019). As cancer often hijacks developmental processes for their own advantage, it is possible that these cells share traits with EMT-induced metastasizing cells, that also allow them to find their way through the matrix. However, as this is a controlled process in development, the invasion is coordinated and does not extend indefinitely. Future work needs to be conducted to quantify the motion and characterize the functions of the motile cells in early stages of TDLU organoid development.

Unlike the processes of induction and patterning, morphogenesis and maturation occurred in a way that was much more expected and canonical to the development previously seen in organoids, including the elongation of ducts, formation of lobular

buds, and expansion of cohesive tissue. Yet live imaging still helped us define two traits of these organoids grown in 3D. The first was the almost cyclonic motion of cells within the lobules. This rotation has been described in the breast (Fernández et al., 2021), however our live imaging confirms this constant motion through alveoli development and demonstrates how even with a defined tissue structure the cells are still highly motile as development occurs. The use of the NOTCH1 activity reporter further showed that the moving cells include those that seem to be luminal in nature, as NOTCH1 is only active in the luminal population of the TDLU. The purpose of this movement is unknown, yet it has may have implications for the formation of functional tissue. It is possible the cells are using this to increase the quantity of cellular signals received from the matrix throughout development, or that they are still going through the process of maturation. By constantly cycling through the structure, these cells may be able to form more cell-matrix and cell-cell interactions, potentially signaling to continue proliferation, development, and differentiation. Future work needs to be done to understand the cells undergoing and mechanics of this motion as well as if it occurs in vivo.

The formation of the mysterious mesenchyme was also a novel discovery through live imaging efforts. These cells initiate within the structure and proliferate at a high rate to quickly expand. It is hypothesized that this cell type plays a supporting role in the formation of TDLU's through the modulation of the ECM at a high rate, (Rauner et al., 2023), eventually leading to enough tension that it can cause an entire hydrogel to collapse upon itself. Other bilayered organoid structures, such as the salivary gland, also seem to form a mesenchymal layer such as this one that surrounds and encapsulates the glandular tissue (Harunaga et al., 2011), although there is little information on either cell

type to draw any major comparisons besides the sheer number and great motility. In breast tissue, it was noted that these cells stain for CK14, meaning that they may retain some basal epithelial features (Rauner et al., 2023).

Perhaps the largest benefit from primary single-cell derived organoids is the ability to transfect the individual cells before being seeded, allowing for entire structures to carry genetic modifications. Not only does this help with being able to identify live cells of interest upon seeding in the hydrogels, but it also creates a method by which new information can be gathered during and after organogenesis without the need to fix and stain the organoids. Basic LeGO and GFP fluorescent models allowed us to predict how these models were forming before we were able to use live imaging to confirm our hypothesis. As many structures only produced a singular color, it was assumed that they developed from a single cell that retained the initial vector. This did not happen in all cases though, as some structures were made up of cells expressing different specific individual vectors. It is possible that these cells did not fully dissociate from each other, ended up grouped together during the infection process, or located each other in the hydrogel during induction and patterning and formed a cohesive tissue during morphogenesis, processes which have all been observed through brightfield or live imaging techniques. There were also instances in which only a couple of cells was seen expressing different fluorescence from the rest of the main structure. In these cases, it is hypothesized that the cells were absorbed into the structure during its elongation, as it is well known that these ductal structures will form together to create a network.

This proof of concept allowed for the use of more complex vectors, used to understand other features of the forming TDLU organoid. While full videos of structures

forming after being transfected with either FUCCI or NOTCH1 vectors were not obtained due to technical issues with the microscope itself, the live imaging that was captured shows that it is not only possible for these vectors to make it into a cell that will eventually derive an entire TDLU organoid, but also that it can be expressed in specific cell types based off of our knowledge of development (i.e. NOTCH1 activity is only seen in luminal cells as previously described (Rauner et al., 2021). This enhancement of model will be critical to gathering novel information about human breast development for things such as proliferative capacity of cell populations, lineage tracing, and a further understanding of the novel cell types described above.

This was also the clearing of a large hurdle for virally induced gene therapy in human breast tissue, which, in combination with new gene editing techniques, could allow for the editing of mutated breast tissue. One of the major roadblock researchers have had is the ability to successfully transfer genetic material in vivo, especially virally where it causes immunogenicity issues as well as safety concerns (Sheikh-Hosseini et al., 2021). The ability to create an ex vivo platform in which viral transduction can occur and be selected for provides a workaround for these issues. While it is not a direct way to fight tumors and metastasis in the breast at the moment, the ability to edit oncogenic tissue and regrow breast tissue in an ex vivo biomimetic environment could allow for the implantation of functional tissue after cancer treatment. As the onset of cancer is trending towards occurring to earlier in a woman's life (Fernandes et al., 2023; Giaquinto et al., 2022), this may be the reality of treatment in the future.

With a stronger understanding of how these structures develop, and with a foundation now laid out for how to best visualize structure formation, future directions

include increased quantification of visualized phenomena and structural generation occurring during development. This includes work not only understanding the function of Leader Cells, but also the quantification of how far they leave the main tissue body, how often, and how often a duct will form in this location. This information will also be vital to collect for the mesenchymal cell population derived from the structure. The discovery of these cell types is proof that this work piloting the utilization of single cell derived organoids has drastically improved an already strong model by increasing its reproducibility and its applications.

*In this chapter, all research presented was conducted by the author, encompassing experimental design, data collection, and analysis. Interpretation of novel cell types and stages of development observed, regarding Leader and Mesenchymal Cells, was conducted as a lab with Gat Rauner, Nicole Traugh, and Meadow Parrish. The findings elucidate the author's comprehensive understanding of the subject matter and reflect their expertise in the field.

Chapter IV: DDR1 Regulates RUNX1 to Drive Breast Stem Cell Differentiation

4.1. Introduction:

The use of a novel hydrogel model to grow human breast tissue allows for a deeper investigation into the development and regulation of stem and progenitor cells as they generate tissue through triggering of protein signaling pathways. Previously, activation of DDR1, through the binding of one of its ligands, collagen type I, was implicated as a necessary factor for the differentiation of stem cells and the formation of biomimetic breast organoids grown initiated from clusters of primary breast cells (Rauner et al., 2021). Along with morphological defects to the tissue, DDR1 inhibition lead to significant expression changes seen at both the protein and RNA level (Rauner et al., 2021). While DDR1 is known to activate proteins at its catalytic kinase domain, it is not known to have transcriptional activities of its own. Thus, DDR1 is likely using proteins that bind and form activating scaffold networks on its catalytic kinase domain when DDR1 becomes activated by binding collagen I, culminating with modulation of mRNA transcription, through alterations to transcription factors.

One such transcription factor may be RUNX1, known to be a master regulator of stem and progenitor cell fate decisions during the development of the hematopoietic system and other tissues (Bresciani et al., 2014; Elagib et al., 2003; Friedman, 2009; Growney et al., 2005). RUNX1 and its essential binding partner CBFβ are known for their role in the differentiation of many blood cell types (Liu et al., 2021) and have been noted to act as both activators and repressors of transcription depending on their cellular context. In recent years RUNX1 has also been observed playing a role in differentiation and development of other tissues including the brain (Fukui et al., 2018), lung (Jeong et al., 2022; Tang et al., 2018), intestine (Usui et al., 2006), and hair follicle (Hoi et al.,

2010) as well as others. RUNX1's stem and developmental functions across tissues designates significance in understanding its role in the breast, especially as these roles can often be hijacked by oncogenic programs to increase tumorigenicity and survival.

In healthy breast tissue it has been shown that RUNX1 is expressed in a majority of cell types (Janes, 2011) and that its expression is required for the exit from a stem cell state (Sokol et al., 2015). Beyond that much of what is known about RUNX1's role in the breast comes from its activities in breast cancer, where it is in the top 30 most mutated breast cancer genes with a frequency of around 3.3% (Ariffin, 2022). It has been seen that RUNX1 can stabilize the epithelial phenotype (Hong et al., 2017) and its loss is associated with metastasis (Hong et al., 2017; Ramaswamy et al., 2003). The similarities between the roles of RUNX1 and DDR1 in the development of the human breast indicates that these two proteins may be involved in a common signaling pathway.

Thus, it was hypothesized that DDR1 has a regulatory effect on the transcription factor RUNX1 that modulates it transcriptional capabilities allowing for differentiation to occur. The work done in this thesis aimed to fill this gap in knowledge regarding DDR1's downstream regulation of transcription, while also clarifying thro role in development that this transcriptional changes control. It was discovered through use of a primary single cell derived breast organoid model that DDR1 controls both the interaction and expression of RUNX1 and its essential binding partner CBFβ, and that this interaction shifts cellular programing from that of a stem cell state to the differentiation of cells through multiple developmental time points.

4.2. Results:

4.2.1. DDR1 Directs Breast Organogenesis Throughout Development

In prior studies, DDR1 has been shown to be expressed and was characterized within 3D hydrogels through the utilization of cell clusters containing primary epithelial breast cells as units for initiating organoid formation. However, the utilization of primary single cells to establish functional tissue has expanded the range of developmental timepoints observable during the generation of the Terminal Ductal Lobular Unit (TDLU). Utilizing a DDR1 inhibitor (DDR1i) and an improved comprehension of breast organoid developmental timepoints, we investigated the essential role of DDR1 at earlier stages of development. This involved DDR1i treatment initiation either on day 0, during induction, or on day 7, during patterning. (Figure 4.1.).

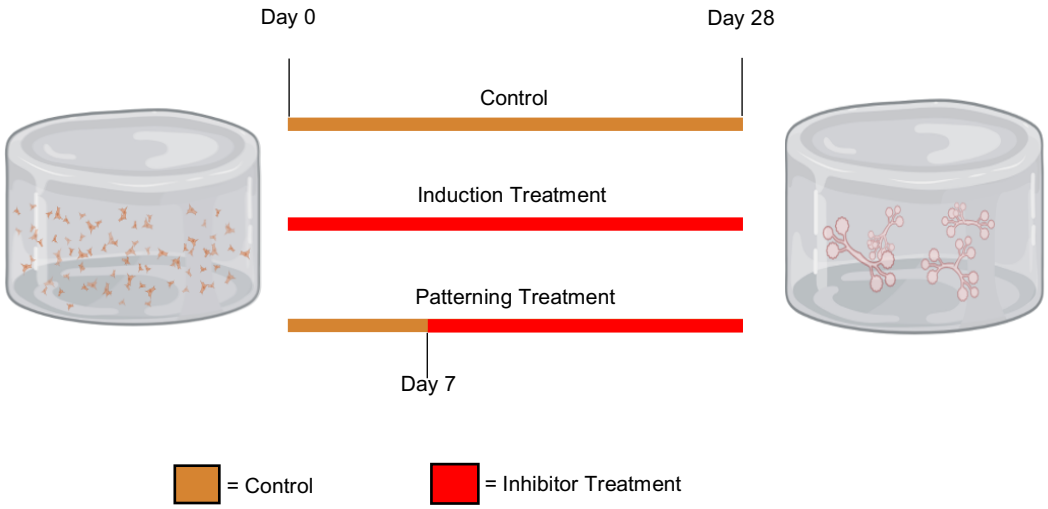


Figure 4.1. Schematic Representation of Methodology Used to Test Inhibitor Effects Across Early Developmental Timelines: Schematic of the strategy to test the effects of DDR1i on breast TDLU organogenesis. Inhibition during induction began with DDR1i treatment starting at day 0 and concluded when control structures were fully formed no later than day 28. Inhibition during patterning started with DDR1i beginning day 7 and concluding around day 28.

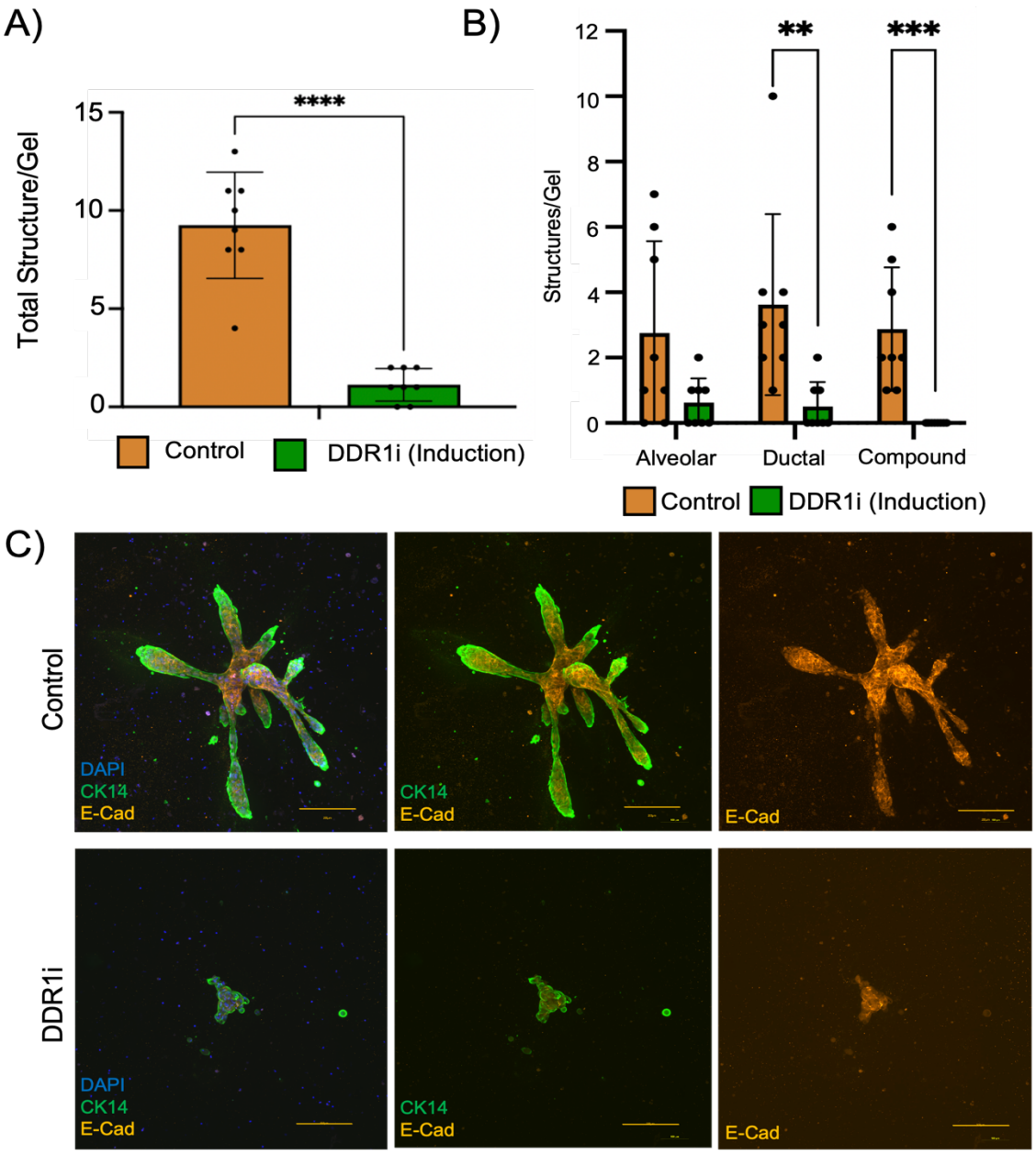


Figure 4.2. Structural Quantification and Morphological Characterization of DDR1 Inhibition During Induction: A) Quantification of the total number of organoids that formed following DDR1i treatment during induction. Data presented as Mean ± SD (n= 4 gels/primary patient samples). B) Quantification of the types of organoids that formed following DDR1i treatment during induction. Data presented as Mean ± SD (n= 4 gels/primary patient samples). C) Representative immunofluorescence staining of organoids from control or DDR1 inhibitor-treated gels, with treatment initiated during the induction phase of organoid formation. CK14 (green), and E-Cadherin (orange), and DAPI (blue) staining. Scale bar = 200 µm. Statistical significance was determined through multiple t-tests, with significance levels indicated as follows: *p-value < 0.05, **p-value < 0.01, ***p-value < 0.001, ****p-value < 0.0001.

Persistent inhibition of DDR1 during induction, beginning on day 0, led to a dramatic decrease in the number of organoids formed (Figure 4.2. A). Furthermore, those organoids that did form were primarily rudimentary alveolar structures. No structures were able to fully mature and form a complex compound structure (Figure 4.2. B). Immunofluorescence shows that the alveolar structures that did form during DDR1i had irregular characteristics including both the shape of the structure and cells that stain positive for both basal and luminal markers, CK14 and E-Cad respectively, suggesting that these cells are retained in a bipotent state and are unable to differentiate and form proper structures without DDR1 activity (Figure 4.2. C).

Chronic inhibition of DDR1 at a later timepoint in organoid development, during the patterning phase, did not lead to a significant difference in the total number of organoids (Figure 4.3. A) but did lead to a significant reduction in the number of compound ductal-alveolar organoids formed and a significant increase in the number of ductal only structures formed, implying that they are being restricted from undergoing alveologenesis (Figure 4.3. B). Immunofluorescent staining shows that ductal structures under DDR1i treatment do form both basal and luminal lineages, implying that DDR1 inhibition during patterning permits stem-cell differentiation, but impairs the morphogenesis of the tissue itself (Figure 4.3. C).

It has been previously seen that the generation of alveolar structures during development requires the differentiation and expansion of the luminal population (Rauner et al., 2021)., Consequently, we examined if a change in the luminal population may be responsible for the loss of alveolar structures. Through flow cytometry, the frequency of

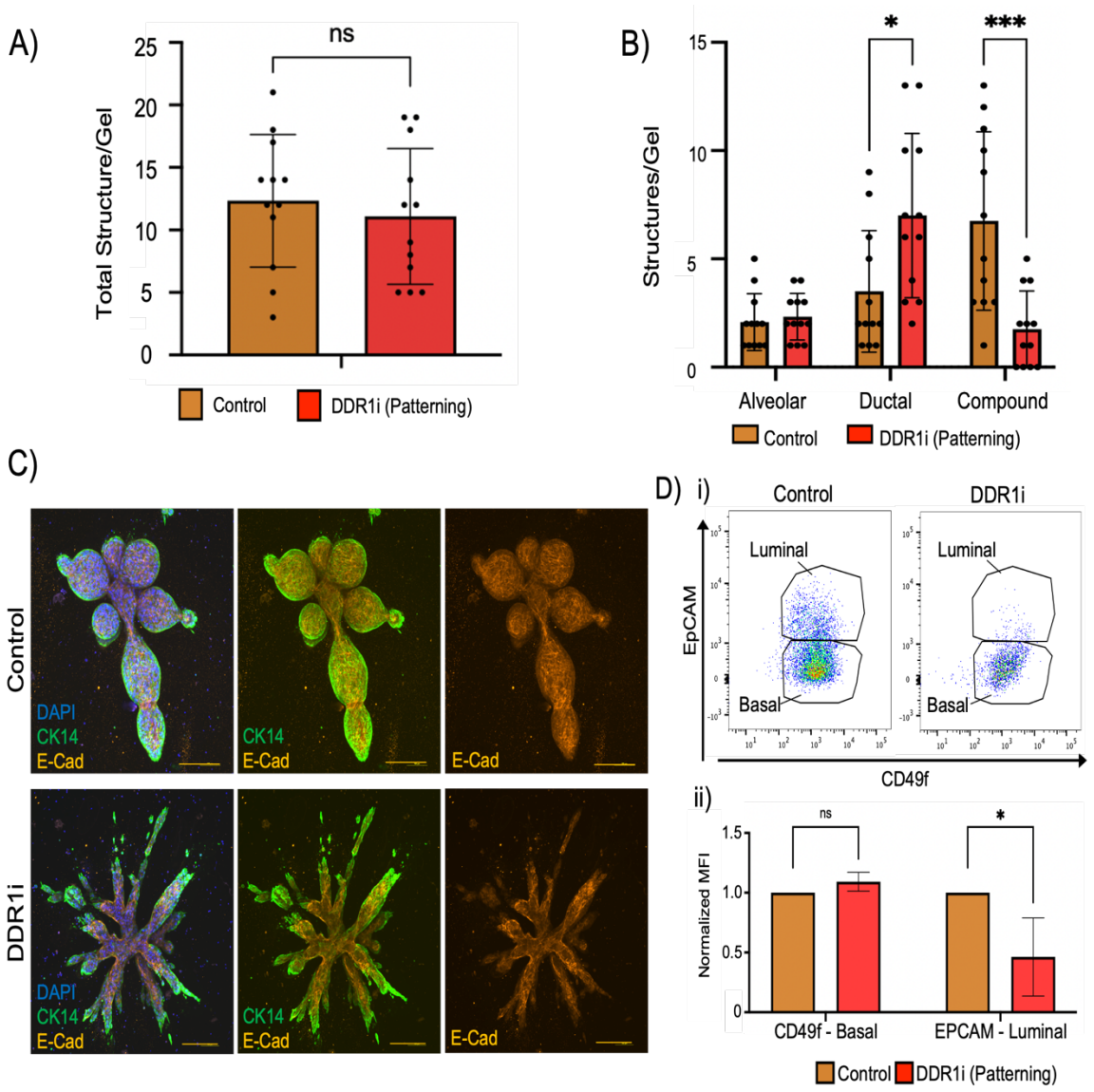


Figure 4.3. Structural Quantification and Morphological Characterization of DDR1 Inhibition During Patterning: A) Quantification of the total number of organoids that formed following DDR1i treatment during Patterning. Data presented as Mean ± SD (n= 4 gels/primary patient samples). B) Quantification of the types of organoids formed in DDR1i treated cultures beginning during patterning. Data presented as Mean ± SD (n= 4 gels/primary patient samples). C) Representative immunofluorescence staining of organoids from control or DDR1i treated gels initiated during the patterning phase of organoid formation. CK14 (green), and E-Cadherin (orange), DAPI (blue) staining. Scale bar = 200 µm. D) Representative flow cytometry plot analysis of basal (CD49f) and luminal (Epcam) cells (FACS) (i) and quantification of mean fluorescent intensity (ii) from primary patient samples cultured in 3D, treated with DDR1i during patterning (n=3 primary samples, and values are expressed as Mean ± SD). Statistical significance was determined via multiple t-tests, with significance levels indicated as follows: *p-value < 0.05, **p-value < 0.01, ***p-value < 0.001, ****p-value < 0.0001.

luminal ($EpCAM^{high}$) and basal ($EpCAM^{neg/low}/CD49f^{pos}$) cells was quantified in TDLU organoids that underwent DDR1i treatment during patterning (Figure 4.3. D). Compared to the control, there was a significant reduction in the number of luminal cells present when DDR1 was inhibited, in line with previous findings. Together these results suggest that DDR1 is necessary for compound ductal-alveolar formation and that during induction it is responsible for stem cell differentiation, while later during patterning it is responsible for morphogenesis through control of luminal differentiation.

4.2.2. DDR1 Inhibition Modulates the Expression of RUNX1 and its Target Genes

Studies in multiple human tissues have shown that RUNX1 is a master regulator of development and differentiation (Bresciani et al., 2014; Elagib et al., 2003; Hoi et al., 2010; Liu et al., 2022; Sokol et al., 2015). In breast tissue it has been shown that RUNX1 expression is required for cells to be able to exit the stem cell state, similar to the phenotype seen in DDR1 inhibited structures (Sokol et al., 2015). Accordingly, a connection between DDR1 and RUNX1 was probed using scRNA-Seq data of RUNX1 and RUNX1 target genes, in which clusters of cells grown for 14 days in a 3D hydrogel and treated with DDR1 from day 0-14 (DDR1i) or from day 0-12 when DDR1 inhibition was released (DDR1r) (Figure 4.4.).

It was observed that RUNX1 expression is lost in bipotent progenitor and basal cell types upon treatment with a DDR1i (Figure 4.5. A). Interestingly, these are the same cell types in which DDR1 is expressed (Appendix Figure 6.3. A). While in basal cells RUNX1 expression returns upon release of DDR1i, this loss of RUNX1 in progenitor cells appears to be irreversible. Looking at other Core binding factor family members

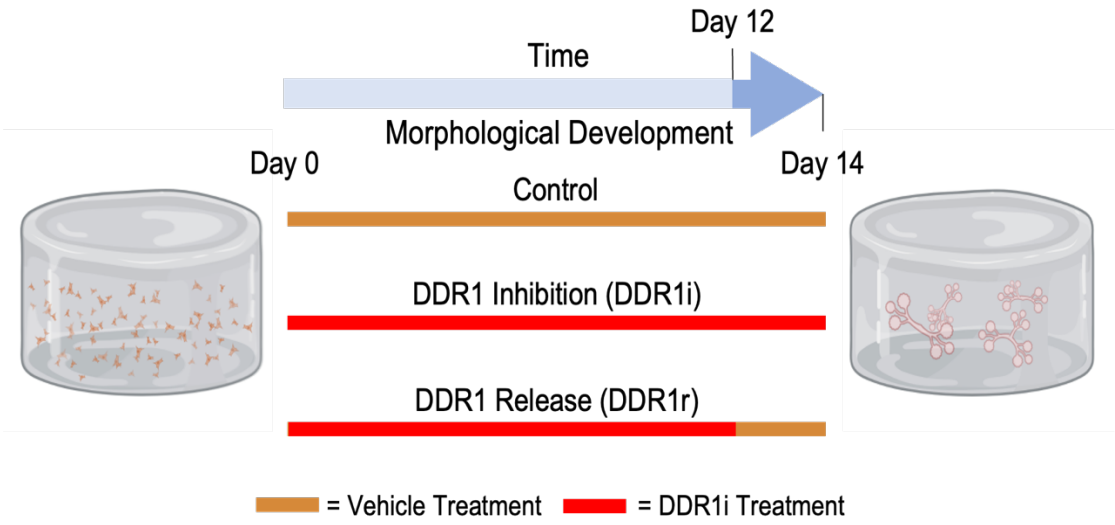


Figure 4.4. Schematic Representation of Methodology Used to DDR1 Inhibitor Effects During Organoid Initiated From Clusters of Primary Cells: A) Schematic representation of DDR1 inhibitor time course design for scRNA-seq. DDR1i treatment starting on day 0 and concluding on day 14, and DDR1r treatment with DDR1i initiation on day 0 and cessation on day 12, with the last two days free from inhibition.

RUNX2, RUNX3 and CBFβ, it is seen that RUNX1's response to DDR1 inhibition specific as these other proteins was limited and not affected by DDR1 inhibition (Appendix Figure 6.3. B-D)

Loss of RUNX1 likely causes changes in the expression of RUNX1 target genes. As RUNX family members are known to be both activators and repressors of transcription based on the context, expression levels of various known RUNX1 target genes both increase and decrease in response to DDR1 inhibition. For example, RUNX1 target gene DUT, an essential nucleotide metabolism enzyme upregulated in many breast cancer subtypes (Davison et al., 2021) gains expression through DDR1i (Figure 4.5. B). Inversely, MYC, a known RUNX1 target gene and interactor, irreversibly loses expression in all cell types upon exposure to DDR1 inhibition (Figure 4.5. C). These trends are seen in many other probed target genes (Appendix Figure 6.3. E-I).

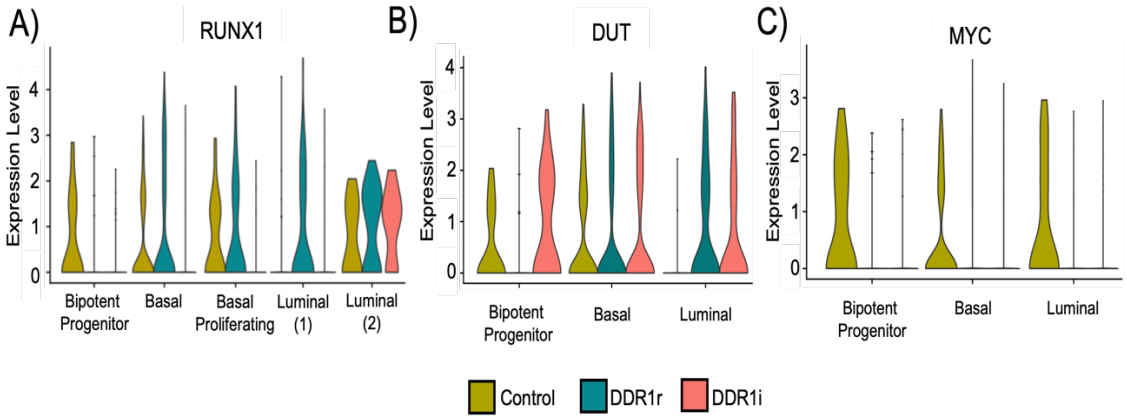


Figure 4.5. Changes in Expression of RUNX1 and RUNX1 Target Genes in Organoids Initiated as Clusters: A-C) Violin plots showing the distribution of RUNX1, DUT, and MYC expression in primary tissue organoids under control, DDR1r or DDR1i conditions.

To understand RUNX1 activity modulation in the context of DDR1 inhibition, a select preliminary group of RUNX1 target genes was used to probe RUNX1 transcriptional activity when treated with DDR1i (Figure 4.6.). Primary samples, treated with DDR1i beginning on day 7 during patterning, all had the induction of the RUNX1 target genes. Similarly, MCF10A cells, an immortalized normal breast cell line grown in 2D with collagen I stimulation and in 3D collagen-based hydrogel, saw the induction of the same genes. This was also seen in the MCF10A derivative cell line MCF10F under collagen stimulation conditions (Appendix Figure 6.3. J). Intriguingly, when either MCF10A or MCF10F cell lines were treated in 2D without the addition of collagen I to stimulate activation of DDR1, it was seen that expression of RUNX1 and RUNX1 target genes were not significantly modulated in the presence of a DDR1 inhibitor (Appendix

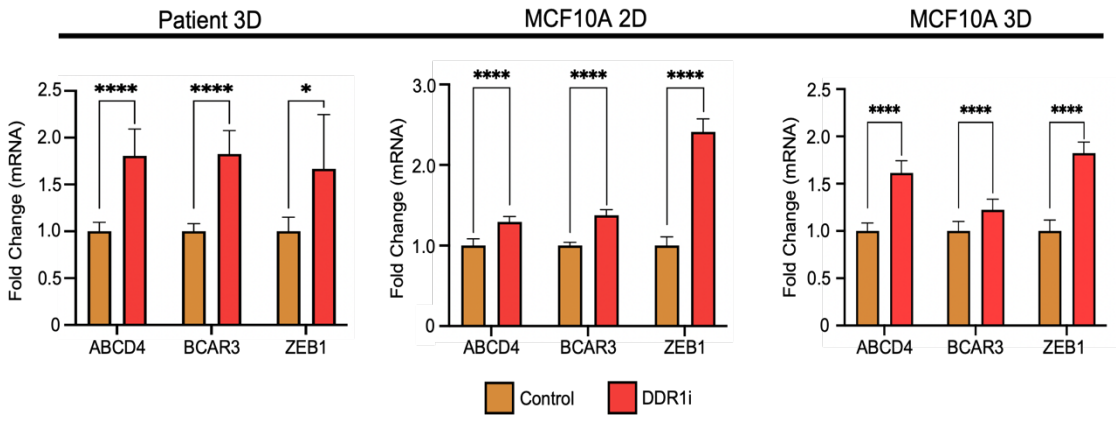


Figure 4.6. Changes in Expression of RUNX1 and RUNX1 Target Genes in Single Cell Derived Organoids and Cell Lines: Compilation of qRT-PCR quantification data derived from three primary patient samples cultured in a 3D environment, MCF10A cells grown in 2D or MCF10A cells grown in 3D, comparing RUNX1 target gene response to DDR1 inhibition. Values expressed as Mean ± SD. Statistical significance was determined via multiple t-tests, with significance levels indicated as follows: *p-value < 0.05, **p-value < 0.01, ***p-value < 0.001, ****p-value < 0.0001.

Figure 6.3. K, L). Taken together, this data suggests that RUNX1 transcriptional activity is modulated by DDR1's activity when binding fibrillar collagen type I.

4.2.3. RUNXi Phenocopies DDR1 inhibition:

If RUNX1-based transcription is downstream of DDR1 and is affected by its activation through binding of collagen type I, we would expect that inhibition of RUNX1 at different developmental time points would generate a similar phenotype in both cell populations and structure. Using a small molecule inhibitor of RUNX (AI-10-104), that functions by preventing interactions between RUNX family members and its cofactor CBFβ, primary human cells were treated with the RUNX inhibitor (RUNXi) on day 0 or day 7 during induction or patterning respectively. As RUNX1 is the primary RUNX family member expressed in the breast tissue (Mercado-Matos et al., 2017), it is also likely the family member driving a majority of the effects in the tissue.

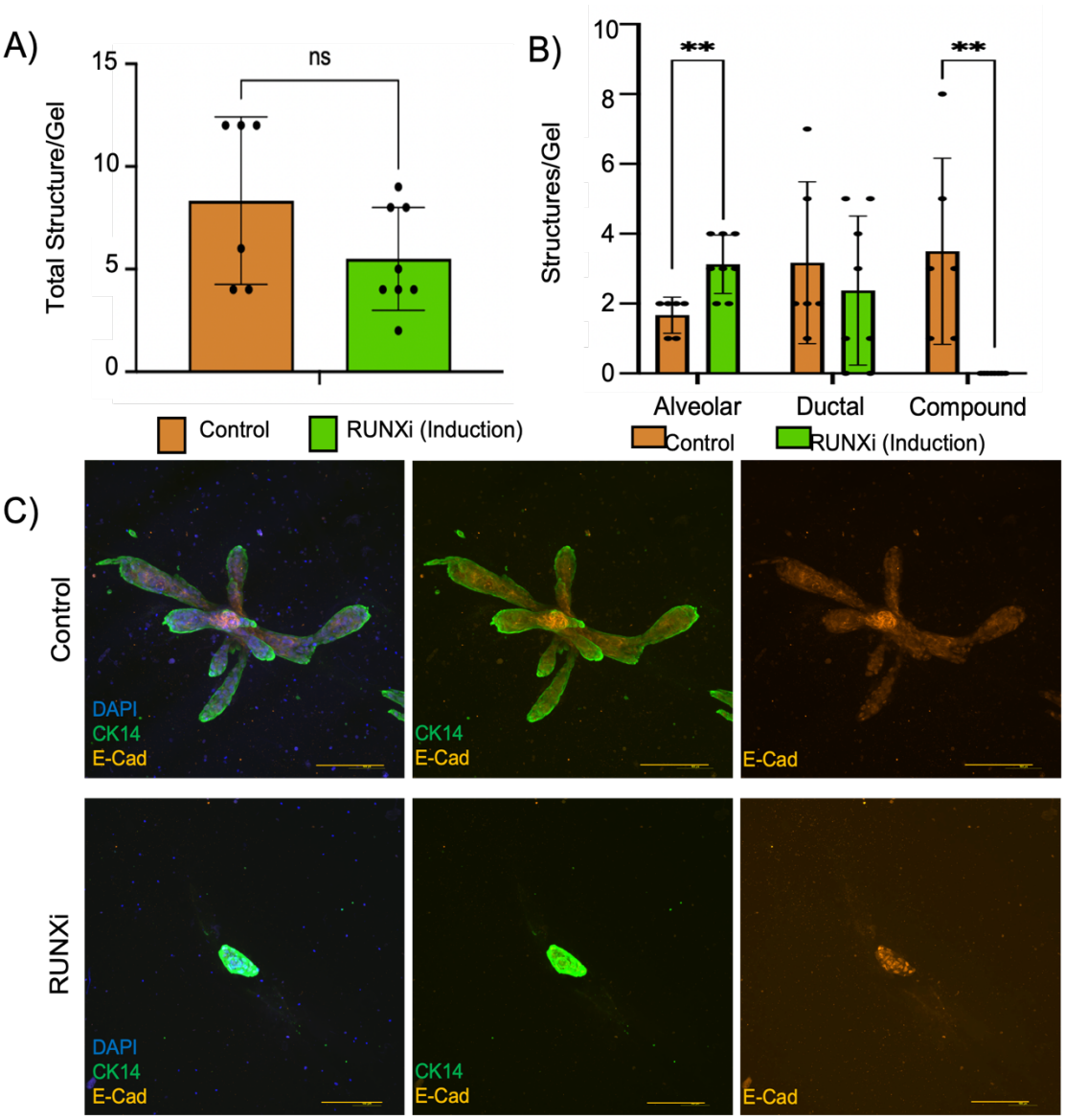


Figure 4.7. Structural Quantification and Morphological Characterization of RUNX Inhibition During Induction: A) Quantification of the total number of organoids that formed following RUNXi treatment during induction. Data presented as Mean ± SD (n= 4 gels/primary patient samples). B) Quantification of the types of organoids that formed following RUNXi treatment during induction. Data presented as Mean ± SD (n= 4 gels/primary patient samples). C) Representative immunofluorescence staining of organoids from control or RUNX inhibitor-treated gels, with treatment initiated during the induction phase of organoid formation. CK14 (green), and E-Cadherin (orange), and DAPI (blue) staining. Scale bar = 200 µm. Statistical significance was determined through multiple t-tests, with significance levels indicated as follows: *p-value < 0.05, **p-value < 0.01, ***p-value < 0.001, ****p-value < 0.0001.

Inhibition of RUNX during induction, beginning at day 0, led to a slight but not significant decrease in the total number of organoids formed (Figure 4.7. A), and a significant increase in alveolar structures and a complete loss of compound TDLU organoids, showing a blockage in morphogenesis (Figure 4.7. B). Immunofluorescent staining shows that the alveolar structures that did develop under inhibitor conditions were rudimentary and double positive for basal (CK14) and luminal (E-Cad) markers, implying that they are limited to their bipotent stem state and cannot differentiate (Figure 4.7. C). These effects are reminiscent of, albeit not identical to, the effects of DDR1 inhibition at this developmental time point.

Inhibition of RUNX during patterning, beginning on day 7, did not result in a significant change to the number of structures formed (Figure 4.8. A) yet did have a significant increase in the number of ductal structures formed and a significant loss of compound structures (Figure 4.8. B). Ductal structures that formed had both luminal and basal cells, implying that during patterning RUNX1 is not influencing stem cells, but the morphology of these structures (Figure 4.8. C). To investigate whether this change in morphology is associated with a loss of the luminal population, we performed flow cytometry population analysis. The results show that RUNXi causes a loss of $EpCAM^{high}$ luminal cells while the $EpCAM^{neg/low}/CD49f^{pos}$ basal population remains consistent across conditions (Figure 4.8. D). Together this data suggests that RUNX1 is responsible for stem cell differentiation during induction and alveologenesis through the expansion of the luminal population during patterning. Inhibition of RUNX1 at both the induction and patterning time points recapitulates DDR1 inhibition, suggesting that DDR1 may act

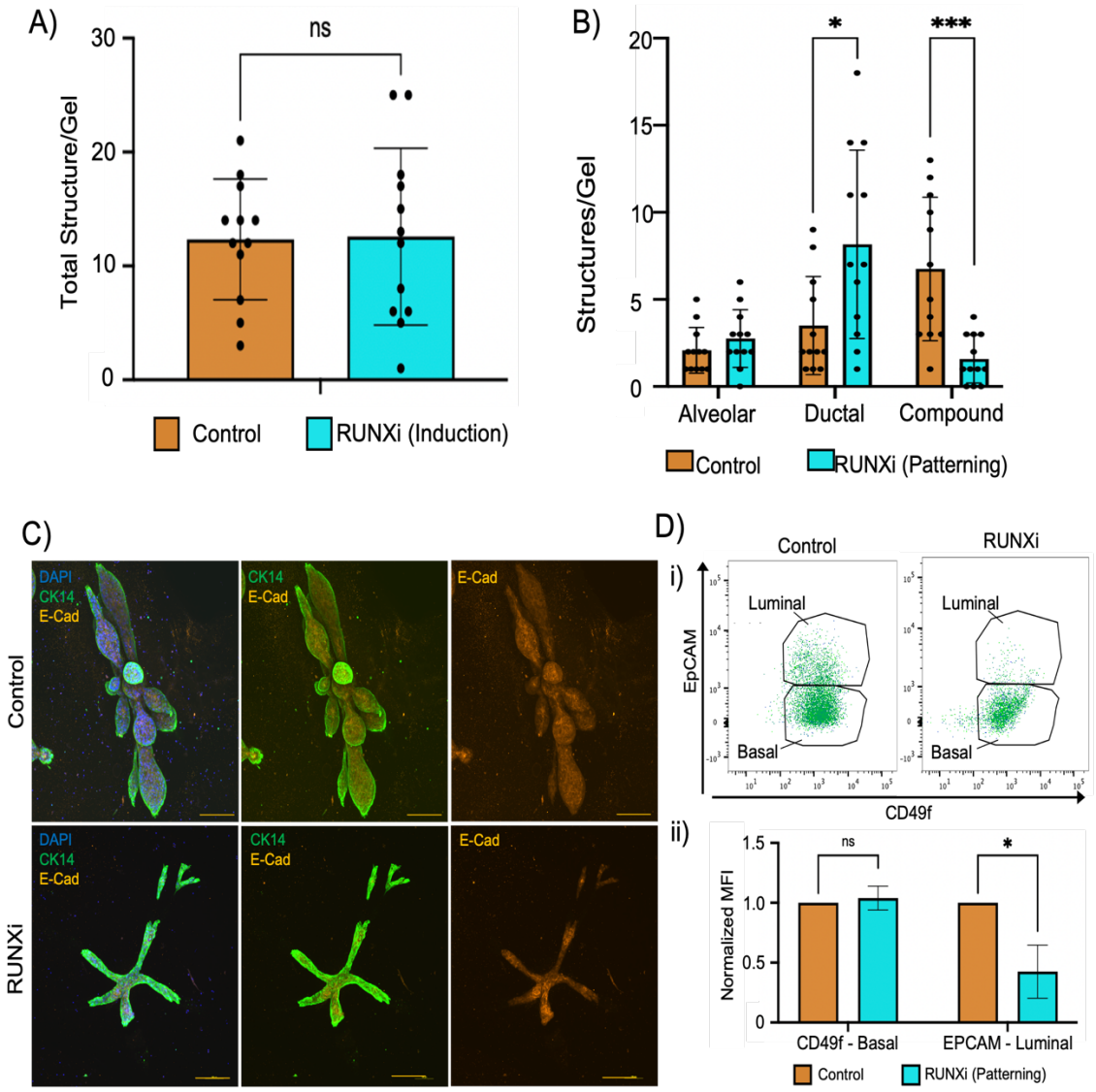


Figure 4.8. Structural Quantification and Morphological Characterization of RUNX Inhibition During Patterning: A) Quantification of the total number of organoids that formed following RUNXi treatment during Patterning. Data presented as Mean ± SD (n= 4 gels/primary patient samples). B) Quantification of the types of organoids that formed following RUNXi treatment during patterning (n= 4 gels/primary patient samples). Data presented as Mean ± SD. C) Representative immunofluorescence staining of organoids from control or RUNX inhibitor-treated gels, with treatment initiated during the induction phase of organoid formation. CK14 (green), and E-Cadherin (orange), and DAPI (blue) staining. Scale bar = 200 µm. D) Representative flow cytometry plot analysis of myoepithelial/basal (CD49f) and luminal (Epcam) cells (FACS) (i) and quantification of mean fluorescent intensity (ii) from primary patient samples cultured in 3D, treated with RUNX inhibitor during patterning (n=3 primary samples). values are expressed as Mean ± SD. Statistical significance was assessed using multiple t-tests, with significance levels indicated as follows: *p-value < 0.05, **p-value < 0.01, ***p-value < 0.001, ****p-value < 0.0001.

through RUNX1 to drive the differentiation and development of a physiologically normal TDLU organoid in this 3D hydrogel model.

4.2.4. DDR1 Controls RUNX1 Interactions to Regulate its Gene Set:

To further probe DDR1's influence over RUNX1's transcriptional network in the breast, bulk RNA sequencing was done, comparing differential expression in cells treated with either DDR1i or RUNXi compared to the control. When looking at the 576 genes differentially expressed when treated with DDR1i it was observed that over half of these genes (316/576) overlapped with differentially expressed genes from cells treated with RUNXi (Figure 4.9. A). Heatmap analysis of all the genes shows that these treatment groups are more similar to one another than they are to the control, as they nest together as one treatment group (Figure 4.9. B).

To gain a comprehensive understanding of the regulation of this gene set and compare the roles of other transcription factors known to regulate these genes, potentially shedding light on an overarching function of these genes, we input the overlapping gene set into ChEA ChIP-seq database (Lachmann et al., 2010), and visualized the results using Harmonizome (Figure 4.9. C) (Rouillard et al., 2016), as well as the TRRUST transcription factor interaction database (Han et al., 2015) visualized with ENRICHR (Appendix Figure 6.4.) (Xie et al., 2021) These results not only continues pointing to this gene set being highly regulated by RUNX1, but it also suggested potential functions of this gene set based on well-known functions of other transcription factors known to regulate this gene set. The most stand out hits were those been previously characterized to induced pluripotent stem cell transcription factors; SOX2, NANOG, and OCT4

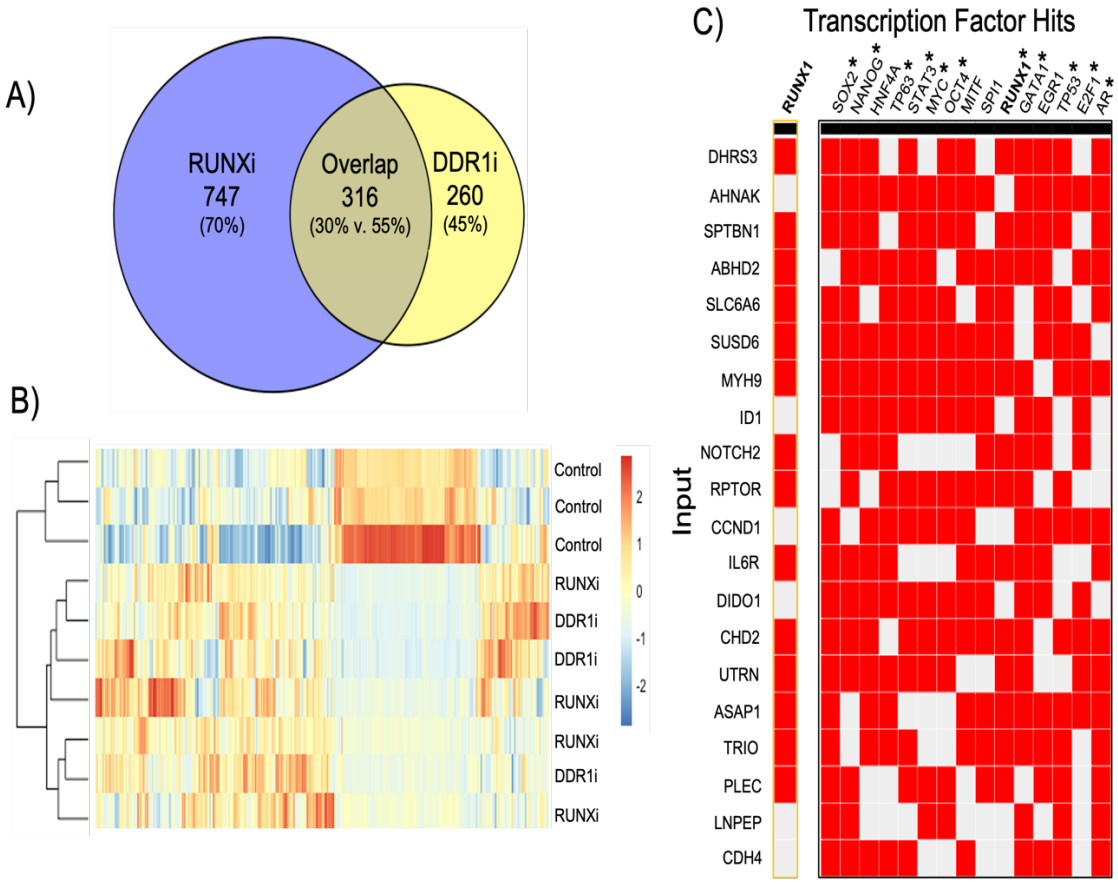


Figure 4.9. DDR1-RUNX1 Transcriptional Network: A) Venn diagram depicting the overlapping differentially expressed genes from bulk RNAseq data obtained from MCF10A cells treated with Col-1 and DDR1 inhibitor (DDR1i) or Col-1 and RUNX inhibitor (RUNX1i), compared to control. B) Heatmap of the overlapping differentially expressed genes, as compared to the control. C) Hierarchically clustered heatmap visualization of the associations between DDR1-RUNX1 overlapping gene set (input) and Enriched transcription factors, utilizing data from the ChEA database (Lachmann et al., 2010). * Represents transcription factor hits of interest.

(Allouba et al., 2015; Luo et al., 2013; Shi and Jin, 2010; Zhang and Cui, 2014). While not as well known, other top transcription factor hits such as TP63, STAT3, MYC, GATA1, TP53, E2F1, and AR are also reported to regulate stemness and differentiation across tissue types (Barbieri and Pietenpol, 2006; Chang et al., 2013; Craig et al., 2023; Fang et al., 2020; Hu et al., 2011; Huang et al.; Jorgez et al., 2021; Murashima et al., 2015; Portal et al., 2022; Wu et al., 2017; Xu et al., 2023). Using the protein-protein

interaction database STRINGDB (Szklarczyk et al., 2022), we found that these transcription factor hits form a close interacting network in which each protein is at most 2 interactions away from any other protein, showing a tightly related network of transcription (Appendix Figure 6.5.). GSEA (Subramanian et al., 2005) of the overlapping gene set further shows these genes are implicated in known RUNX1 processes such as MYC signaling (Choi et al., 2017), epithelial-mesenchymal transition (EMT) (Khawaled and Aqeilan, 2017; Lu et al., 2020) and tissue development (Coffman, 2003), estrogen response (Stender et al., 2010), and other functions related to both the breast and development (Mootha et al., 2003; Subramanian et al., 2005) (Table 4.1.).

Table 4.1. DDR1-RUNX1 Transcriptome-Associated Activities Identified Through GSEA:

Hallmark Gene Set	# of Bookmarked Genes	p-value
Mitotic Spindle	21	$3.03\ e^{-18}$
Myc Targets	18	$1.34\ e^{-14}$
Oxidative Phosphorylation	14	$3.48\ e^{-10}$
Apical Junction	11	$3.21\ e^{-7}$
Epithelial Mesenchymal Transition	11	3.21 e-7
IL2 Stat5 Signaling	10	$2.5\ e^{-6}$
E2F Targets	9	$1.93\ e^{-5}$
Estrogen Response Early	9	$1.93\ e^{-5}$
Estrogen Response Late	9	$1.93\ e^{-5}$
NOTCH Signaling	4	$8.75\ e^{-5}$
Myogenesis	8	$1.27\ e^{-4}$
TGF Beta Signaling	4	$6.78\ e^{-4}$
UV Response (Down)	6	$7.17\ e^{-4}$
G2M Checkpoint	7	$7.4\ e^{-4}$
MTORC1 Signaling	7	$7.4\ e^{-4}$
TNFA Signaling vis NFkB	7	$7.4\ e^{-4}$
Complement	6	$3.77\ e^{-3}$
Myc Targets V2	3	$9.01\ e^{-3}$
Heme Metabolism	5	$1.65\ e^{-2}$
P53 Pathway	5	$1.65\ e^{-2}$

This overlapping transcriptional network based on DDR1 activity through RUNX1 suggests that DDR1 may be physically regulating RUNX1 as a mechanism for this change. Looking at the transcription factor concentrations specifically in the nucleus of MCF10A cells by western blot, there is a significant loss of RUNX1 and its essential binding partner CBFβ when treated with DDR1i under collagen stimulation conditions in both cell lines and primary samples (Figure 4.10. A and B). Similar to what was seen with the mRNA expression of cells exposed to DDR1i but not collagen I, DDR1i does not have an effect on cell protein expression compared to the control and there is no loss of either CBF components in MCF10A cells (Appendix Figure 6.6.). Use of 3D hydrogel cultured primary single cell derived TDLU organoids also showed that a loss of nuclear RUNX1 and an even greater loss of nuclear CBFβ was induced through DDR1i.

As it is hypothesized that RUNX1 and CBFβ bind together in the cytoplasm prior to entry into the nucleus (Liu et al., 2021; Zhen et al., 2020), it was important to determine if DDR1i interferes with this interaction. We used immunoprecipitation to see the interacting relationship between RUNX1 and CBFβ, in the presence and absence of active DDR1. Indeed, there was a significant reduction in CBFβ bound to RUNX1 in DDR1i conditions (Figure 4.10. C). Inhibition of RUNX1 using the AI-10-104 inhibitor served as a positive control, as this inhibitor's mechanism is through blockage of interaction between RUNX1 and CBFβ. The quantity of RUNX1-CBFβ binding did not differ between inhibitor treatments of DDR1 and RUNX (Figure 4.10. C).

To evaluate potential downstream targets of DDR1 that may act as intermediates between the collagen receptor and RUNX1, a phosphor-kinase array was utilized to query proteins known to be activated by DDR1 in other organisms and tissues, specifically

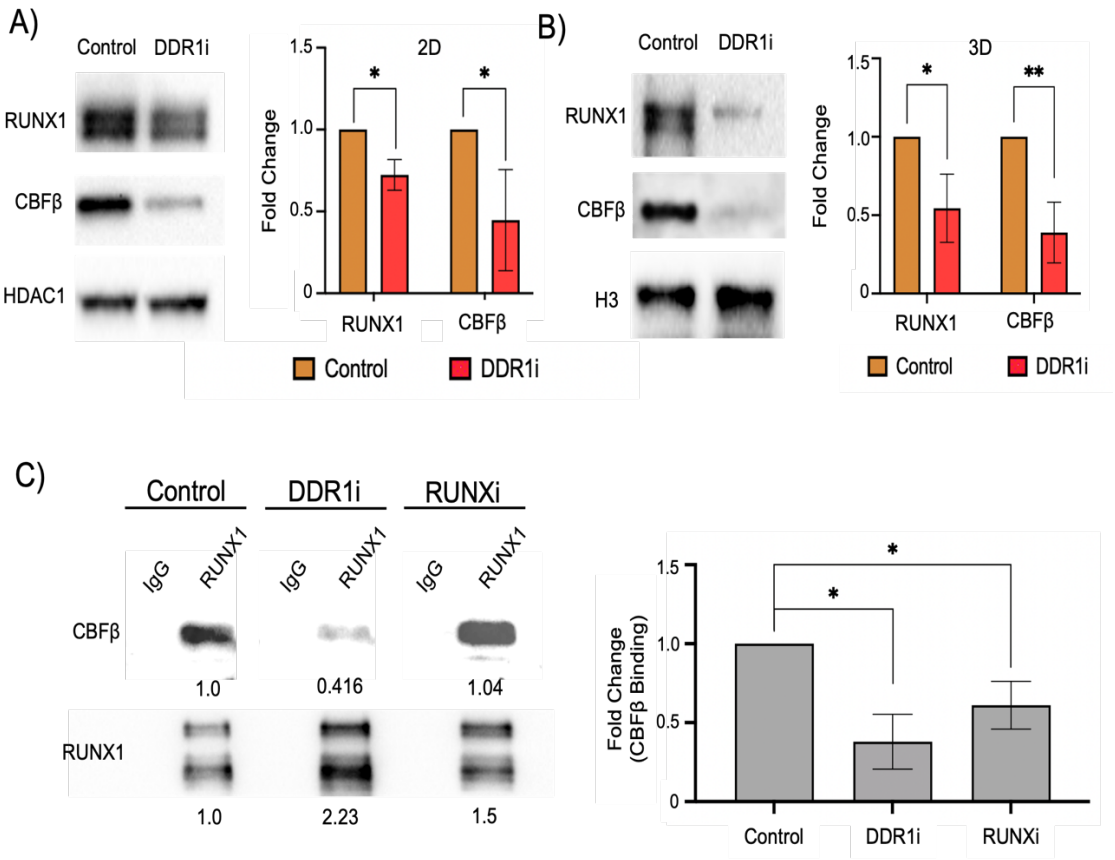


Figure 4.10. RUNX1 and CBFβ Protein Expression and Interaction are Decreased by DDR1i: A) Representative western blot analysis and quantification of RUNX1 and DDR1 expression from nuclear lysates of MCF10A cells treated with collagen I and DDR1i (n=3). Data presented as Mean ± SD. B) Representative western blot and quantification of nuclear protein from fractionated lysate obtained from primary patient samples cultured in 3D hydrogels (n=3). Data expressed as Mean ± SD. C) Co-IP of RUNX1 and subsequent blotting and quantification of CBFβ from the lysate of MCF10A samples cultured in 2D treated with collagen I and DDR1i. Quantification is derived from n=3 independent experiments and is presented as Mean ± SD. Statistical significance was determined through multiple t-tests, with significance levels indicated as follows: *p-value < 0.05, **p-value < 0.01, ***p-value < 0.001, ****p-value < 0.0001.

members of the SRC family kinases (SFKs), which are known to be both activated by DDR1 and modulators of RUNX1 activity. It was seen that a majority of changes with DDR1i occurred with the loss of activation phosphorylation on kinases including SRC family kinases WNK1, YES, and LYN as well as kinase MSK1/2, transcription factors of the STAT family (STAT1,2,3, and 5), and other proteins such as HSP27 known to be lost

with DDR1i (Table 4.2. and Appendix Figure 6.7.) (Wantoch von Rekowski et al., 2020). Many of the proteins activated by DDR1, directly or indirectly, are known to be RUNX1 interactors and form a network that could feasibly drive changes to its structure through PTMs (Appendix Figure 6.8.). Together this data shows that RUNX1 drives a majority of DDR1's transcriptional network through signaling that allows RUNX1 to bind CBFβ, thereby allowing it to drive differentiation programming driving the differentiation of basal and luminal lineages.

Table 4.2. Potential DDR1-RUNX1 Intermediaries Identified Through Phosphokinase Array:

Hit	Protein Function
WNK1	Kinase
HSP27	HSP
YES	Kinase
STAT2	Transcription Factor
STAT3	Transcription Factor
MSK1/2	Kinase
STAT3	Transcription Factor
LYN	Kinase
ERK1/2 (Gained Phosphorylation with DDR1i)	Kinase

4.2.5. The DDR1-RUNX Axis in Breast Cancer:

Breast cancer is caused by the aberration of development and differentiation. As the breast is a tissue that goes through cycles of growth and involution, there are many opportunities for this misregulation to occur. Once oncogenesis begins, impacted breast cells no longer properly respond to the physiological developmental signals and allow for the cells to deviate from normal differentiation programming, often reverting into a more

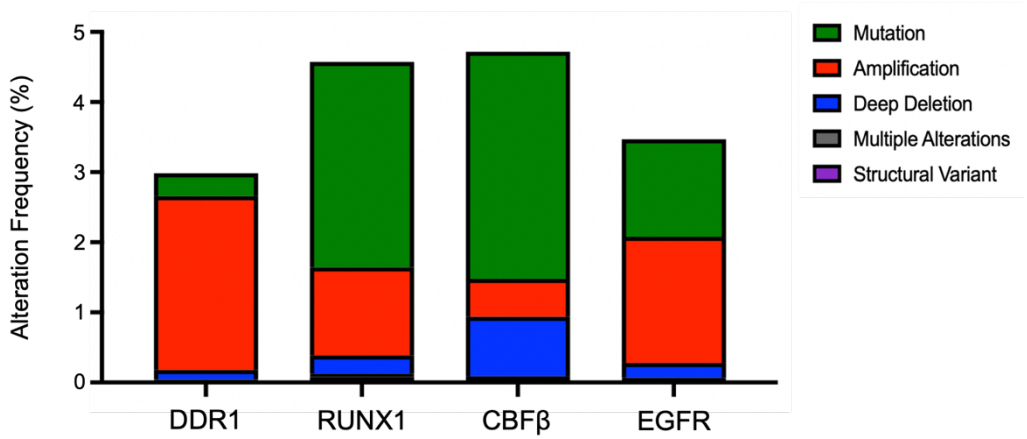


Figure 4.11. Mutational Analysis of DDR1, RUNX1 and CBFβ: Genomic alteration frequencies of RUNX1, DDR1, CBFb, and EGFR in human breast cancer. Mutational frequency from 10,363 breast cancer (BC) samples in 9,776 patients, categorized by mutation type, of genes of interest and well-known regulators in BC from the cBioPortal database.

stem-like states (Fillmore and Kuperwasser, 2008; Zhang et al., 2020). We hypothesized that since the DDR1-RUNX1 axis is responsible for pushing cells from a stem cell state to a differentiated one, it may be frequently mutated in breast cancer. Using the cBioPortal comprehensive collection of breast cancer data (de Bruijn et al., 2023; Cerami1 et al., 2012; Gao et al., 2013), mutation status of DDR1, RUNX1, and CBFβ were quantified per mutation type. Overall, these mutations occurred at rates between 3~5% across all breast tumors, a rate similar to that of EGFR which is a common mutation (Masuda et al., 2012) (Figure 4.11.). By looking at co-occurring mutations across breast tumors, mutual exclusivity between gene mutations can be probed. Interestingly, there is a high rate of co-occurrence between DDR1 and RUNX1 mutations as well as DDR1 and CBFβ mutations. RUNX1 and CBFβ mutations tend to be mutually exclusive in breast cancer (Table 4.3) as well as generally across tissues (Table 4.4). This data suggests that the DDR1-RUNX1 axis is a target of frequent mutation in breast cancers.

Table 4.3. Mutual Exclusivity or Co-Expression of DDR1, RUNX1 and CBFβ mutations in Breast Cancer:

A	B	Neither	A Not B	B Not A	Both	Log2 Odds Ratio	p-Value	q-Value	Tendency
DDR1	RUNX1	3427	57	184	19	2.634	<0.001	<0.001	Co-occurrence
DDR1	CBFβ	3456	67	155	9	1.583	0.006	0.009	Co-occurrence
RUNX1	CBFβ	8237	447	401	21	-0.051	1.000	1.000	Mutual exclusivity

Table 4.4. Mutual Exclusivity or Co-Expression of DDR1 and RUNX1 mutations in Pan-Cancer Studies:

A	B	Neither	A Not B	B Not A	Both	Log2 Odds Ratio	p-Value	q-Value	Tendency
DDR1	RUNX1	9708	234	239	17	1.561	<0.001	<0.001	Co-occurrence

To determine the relevance of DDR1, RUNX1, and CBFβ expression to patient care, METABRIC data from the Breast Cancer Integrative Platform (Wu et al., 2017) was used to produce Kaplan-Meier curves showing overall survival in both TNBC and Non-TNBC (i.e. tumors expressing one or more expressed receptor of either ER, PR, or HER2). In non-TNBC breast cancers, expression of DDR1 or RUNX1 indicated a better patient outcome (Figure 4.12., panels i-ii). In contrast, DDR1 and RUNX1 expression in TNBC tumors is associated with a poor patient outcome (Figure 4.12., panels iv-v). In both cases, higher CBFβ expression is associated with a negative outcome (Figure 4.12., panels iii, vi). This data underscores an ongoing interplay through the interconnected phenotypes resulting from the modulation of DDR1 and RUNX1, which persists beyond normal tissue development and extends into cancer. These findings may hold clinically relevant implications for treatment strategies.

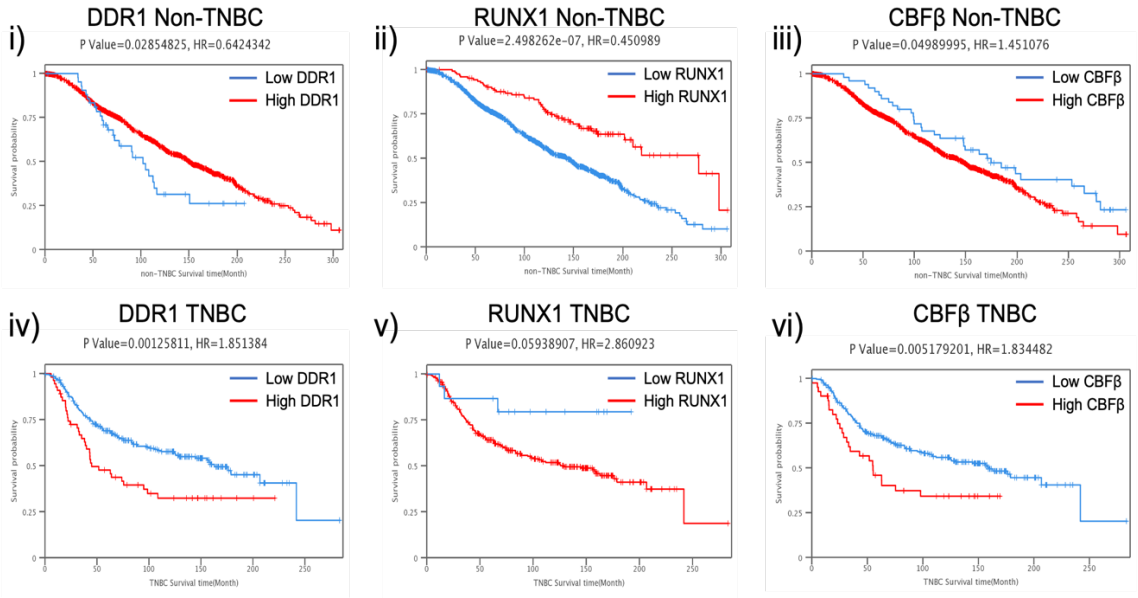


Figure 4.12. Breast Cancer Survival Based on Expression of DDR1, RUNX1 and CBFβ: Kaplan-Meier OS curve based on high (red) and low (blue) DDR1 (i), RUNX1 (ii), or CBFβ (iii) in triple negative breast cancer (TNBC) or non-TNBC (iv-vi) from METABRIC data on the Breast Cancer Integrative Platform.

4.3. Discussion:

The identification of a novel DDR1-RUNX1 signaling axis has unveiled a new pathway with implications in differentiation, development, and disease within breast tissue. Utilizing a primary single-cell-derived 3D breast organoid model, we observed that the DDR1-RUNX1 axis influences the transition out of the stem state and the formation of ductal structures during induction, while during patterning, it governs luminal differentiation and alveologenesis. Additionally, DDR1's transcriptional network is primarily regulated by RUNX1's expression and its interaction with CBFβ, holding potential implications for the clinical management of breast cancers.

While morphologically phenocopying one another, there were some differences seen between the structures under inhibitor conditions. The first was the total number of structures formed during induction, where DDR1i conditions created a significant

difference in the number of structures formed and RUNXi conditions did not. This may be due to DDR1's position upstream of RUNX1 and its RUNX1-independent functions in the signaling for proliferation and structural development, thus preventing the formation of any organoid. The second major difference observed was the quantity of differentially expressed genes caused by the inhibitors in comparison to the control population, where RUNXi caused the differential expression of twice as many genes as DDR1i. This could be due to DDR1 being more nuanced and only regulating a specific subset of genes by RUNX1, and that when DDR1 is inhibited, only that subsection is affected, compared to when RUNX is indiscriminately affected by its inhibitor. This subtle shift in programming could allow for DDR1-independent or other DDR1-RUNX1 processes to occur, allowing for the maintenance of long-term stem cells (North et al., 2002).

It is still not known if the DDR1-RUNX1 axis is a direct interaction between the two proteins, based on what is known about RUNX1 it seems more likely that it is being modulated by an intermediate between the RTK and the transcription factor. There are many possibilities for what this intermediate could be, but we can hypothesize that it is a member of the SFK proteins, proteins known to interact with RUNX1 (Han et al., 2022; Lu et al., 2011). Evidence for this was uncovering in the same knockout experiment that revealed DDR1's major role in the differentiation of breast stem cells. In this experiment the second highest target was a SRC family kinase FYN, implying that this protein was found to also be necessary for exit from a stem cell state and the development of a TDLU (Rauner et al., 2021). DDR1 is also known to interact with SFK proteins (Huang et al., 2012). Activation phosphorylation of other SRK, including YES and LYN were also seen to be affected by the loss of DDR1 activation through a phosphokinase array and are

known to be downstream of FYN activity. Further work is needed to determine which SRC family kinase is involved, and if there are any other proteins involved in this pathway.

While the intermediate between these two proteins is still unknow, it has been shown that this is just one of three ways in which DDR1i affects RUNX1. Beyond the direct interaction between RUNX1 and CBFβ, it was shown that DDR1i decreased the expression of both proteins and the mRNA of RUNX1. Regulation of this transcription factor at every level serves to highlight the important connection between these two proteins in the role of controlling differentiation. While these regulatory events do not have to be mutually exclusive, it is also possible that all of the changes to RUNX1 expression occurs through the disruption of its interaction with CBFβ. It has been previously shown that the RUNX1-CBFβ interaction stabilizes RUNX1 protein (Qin et al., 2015) and allows it to control its own transcription (Martinez et al., 2016). By disrupting this interaction, RUNX1 would be marked for degradation and its mRNA would naturally decrease. Further work needs to be done to determine if these are controlled through a single action or 3 individual regulations.

Additional research is also required to understand the breakdown of what cellular processes the DDR1-RUNX1 axis is controlling through RUNX1's transcriptional activity. Identifying known functions of other transcriptional regulators that control the same gene set that is controlled by the DDR1-RUNX1 axis could reveal the overarching roles of this axis. One of the most telling hits on this list were the embryonic and induced pluripotent stem cell transcription factors SOX2, NANOG, and OCT4. Although the breast epithelial cells responsible for reconstituting the TDLU are not pluripotent stem

cells, this suggests that the differentiated genes they possess may still regulate breast stem cell behavior. This observation aligns with the data demonstrating that inhibition of both DDR1 and RUNX1 prevents these cells from transitioning out of a stem state. Alternatively, other transcription factor hits including TP63, STAT3, MYC, GATA1, TP53, E2F1, and AR are known to regulate differentiation in multiple solid tissue types including stratified epithelial tissues (Barbieri and Pietenpol, 2006; Craig et al., 2023; Fang et al., 2020; Portal et al., 2022; Wu et al., 2003) and both male and female reproductive tissues (Chang et al., 2013; Hu et al., 2011; Huang et al.; Jorgez et al., 2021; Murashima et al., 2015; Xu et al., 2023). Misregulation of transcription factors also has implications in cancers of these tissues. Some transcription factor hits (SOX2, OCT4, Nanog, E2F1, MYC) are observed acting like oncogenes (Dang, 2012; Narayan et al., 2017; Noh et al., 2012; Schaefer and Lengerke, 2020; Yang and Sladek, 2018), while others (TP53, STAT3, EGR1, and GATA1) are known to act as tumor suppressors (Aigner et al., 2019; Aubrey et al., 2016; Wang et al., 2021; Zheng and Blobel, 2010). Together the implication of these transcription factors regulating this overlapping gene set paints a picture that the DDR1-RUNX1 axis plays a role in the transition from stem cell programming to that of differentiated tissue.

Finally, evidence from pan-cancer studies suggests that this signaling axis may exist within other tissue types. With growing interest into the roles of DDR1 and RUNX1 during the development of human tissues, it is probable that more and more research will find that DDR1 may play some role in the regulation of RUNX1. This warrants further research to fill the new gaps of knowledge presented above, as common pathways controlling the development of tissues are important to researchers focused on both

fundamental and disease-oriented biology. As small molecule inhibitors of RUNX1 and DDR1 are entering clinical trials (Elkamhawy et al., 2021; Illendula et al., 2016), it is important to understand exactly how they can be used therapeutically.

In this chapter, all research presented was conducted by the author, encompassing experimental design, data collection, analysis, and interpretation except for the scRNA Sequencing Analysis which was conducted by Nicole Traugh at the request of the Author. The findings elucidate the author's comprehensive understanding of the subject matter and reflect their expertise in the field.

Chapter V: Discussion and Significance

5.1. Discussion:

Past research has made a monumental impact on the groundwork for our understanding of breast development; however, a lack of a replicable model has made it hard to predict if discoveries made in the mammary glands of animal models, or 2D cell culture, will translate to human tissue in a clinical setting. 3D models offer potential solutions to these challenges. Human breast tissue, in particular, stands to benefit from enhanced 3D modeling to better understand its physiology and develop personalized treatments (Dai et al., 2017; Holen et al., 2017). As 3D culturing techniques advance (Baptista et al., 2022; Corrò et al., 2020; Rauner et al., 2023), the potential for therapeutic and regenerative applications becomes increasingly promising.

The regenerative potential of breast tissue provides a rich source of cells for initiating ductal lobular structure development (Mohan et al., 2021; Visvader and Stingl, 2014). Despite this benefit, current 3D models face challenges in achieving biomimetic fidelity, with previous iterations falling short in accurately representing human tissue physiology or extracellular matrix composition. Recent advancements, such as the development of a novel 3D hydrogel model incorporating breast stroma components and primary patient cells, offer promising avenues for more representative and informative breast tissue modeling (Rauner et al., 2021, 2023; Sokol et al., 2016). Further enhancements, including the refinement of initiating cell clusters dissociation, hold potential for improving the accuracy and versatility of 3D breast tissue models for future research applications.

This enhancement of the organoid initiating units to single cells, marks the true ability to make associations between human organoid model conditions. Previous efforts

could not truly be directly compared to one another due to inequalities in initiating units, both in quantity and developmental stage. As clusters of epithelial cells previously used for organoid development (Rauner et al., 2021; Sokol et al., 2016) may have contained cells at different developmental stages, as well as all in different quantities, each organoid structure itself may be in slightly different stages of development, which could modulate their phenotypic response to perturbation. Enhancement of this 3D model to generate single-cell derived organoid structures removes these differences, allowing for more comparable and replicable data across patient samples.

Part of this inability to compare previous conditions came from the discovery of new initial stages of organogenesis during breast TDLU development. While previous work understanding part of DDR1's role in the breast was successful due to the visually drastic morphological phenotype that occurs with the loss of alveologenesis (Rauner et al., 2021), nuances in developmental timing easily could have obfuscated important results. The splitting of an "early" timepoint into two distinct developmental stages, Induction and Patterning, has allowed for a more fine-tuned understanding of how organogenesis initiating signaling drives development through different avenues at each stage.

While DDR1i had been previously predicted to retain a stem cell state in breast structures through its knockout (Rauner et al., 2021), there was no morphological data to confirm this as structures were shown to still develop ducts. Use of single cells allows the model to observe earlier time points and interrogate these timepoints as two distinct developmental phases. As seen from experiments initiated with single cells, DDR1i during induction does indeed prevent the exit of cells from the bipotent stem state.

Alternatively, it was seen that during patterning DDR1 is responsible for luminal differentiation and expansion, as well as the morphogenesis of alveoli structures on the ends of ductal branches. Additionally, inhibition of RUNX1, the transcription factor believed to be controlled by DDR1 signaling, phenocopied the effects of DDR1i at two different stages of development. The consistency in the traits affected by the interruption of these two proteins normal functions across separate developmental stages strengthens the probability of their association, a detail not discernible in previous iterations of this model.

As seen from this work, perturbation of the same protein can have drastically different effects depending on the developmental stages. These stages offer further considerations for researchers using 3D models to understand both morphological and expression phenotypes and can dictate the proper planning and execution of experiments. Thus, an overlooked aspect of organoids systems is potential issues caused by using pre-differentiated cellular clusters as organoid initiating units. As seen from work in our own lab (Rauner et al., 2021), the use of clusters pushes initiation from the induction phase to the patterning phase, completely removing the stem cell context from the experiment, leaving important results without supporting evidence. Use of single cells would avoid instances where one cluster may contain more cells at stages ready to produce a structure, while another structure requires more proliferation before being able to sprout ducts. Without the additional context coming from direct observations of breast development from a single cell, the impact of this work would be greatly lessened.

Additional definition of these phases as well as the phases of morphogenesis and maturation would also allow for a deeper characterization of development. A better

understanding of these developmental phases can also answer new questions introduced by this single-cell enhancement. Some remaining questions include those pertaining to developmental cues that control structure morphology and why some are directed to be compound, while others retain either ductal or alveolar features. While some data is exists pointing to certain unipotent stem cells driving these ductal or acinar structures (Linnemann et al., 2015), it impresses new questions on the predicted ability of these cells to undergo plasticity at different stages of their development, and may offer new clues on which cells are the de-facto stem cell in the adult human breast.

Other unanswered questions from these single cell experiments revolve around new cell types discovered, including their identity, the roles that they play in breast development, and most importantly if they are an artifact of transplantation, or cells that exist within the human body. As the ability to live image developing human tissue is a technique in its infancy, it is possible that other tissues may utilize these cell types as well. A better understanding of the activities of these cells in the breast could give researchers a leg up in understanding their role in these other tissues.

Finally, it would be important to understand how all these traits may relate to other patient-specific factors such as age, race, body mass index, hormonal exposure, and parity as well as others. It is well understood that each of these characteristics can have drastic effects on breast biology. As grouping organoid growth together in this way may showcase drastically different phenotypes that would be hard to characterize when interspersed with all other breast organoids, studies focused on the growth of tissues from these cohorts could make important observations about breast biology that could help alleviate issues in women's health.

In the short term, future directions for this model include further enhancements such as the inclusion of resident stroma and modulating the niche that makes up the model, the addition and integration of disease models into the novel 3D hydrogel model, and the growth of organoids from other primary tissue sources. The most readily accomplishable is the addition of supporting stroma to the niche environment to additionally increase the relevance of the model. The ability to add immune cells, adipose tissue, neuronal tissue, vasculature, and fibroblasts (Fein et al., 2023) would allow for associations between the growth and signaling of these cell types that could augment the way organogenesis occurs in this model. It has been observed that signaling between all four cell types and the epithelium in mammary tissues plays a role in the development of the mammary gland (Couldrey et al., 2002; Fein et al., 2023; Kothari et al., 2020; Liu et al., 2012; Reed and Schwertfeger, 2010). The ability to add each cell type to the model individually or in concert allows for more enriched comprehension of these interactions in the likely reciprocal proliferation, differentiation, and development of the various tissues.

Also feasible in the short term is the ability to convert this model to study breast cancer, focusing on the dire need for more precision medicine-based therapies. This model would likely support primary tumor tissue and could allow for a deeper understanding of metastasis and tumor growth, as well as any interactions with other cell types found in the breast. Growth in a 3D hydrogel could be used to one day help inform personalized treatment strategies by through the ability to probe a patient's tumor for drug resistance through secondary mutations.

Use of this model to help determine potentially important pathways in cancer has already been modeled by DDR1 and RUNX1 in this thesis. Understanding that effects to DDR1 could have a major effect on RUNX1 signaling gives it clinical implications. This is seen through the use of an archive that collects patient tumor data that indicates expression and mutational status of the DDR1-RUNX1 axis may be useful prognostic tools. Overall survival shows the dual nature of these proteins as moderators of both stem and differentiation programming. High expression of RUNX1 and DDR1 in HR+ cancers, which indicates a more positive outcome, would likely have an anti-ER effect as RUNX1 is known to have an inverse relationship with the expression of ERα (Stender et al., 2010). It is also possible that higher expression of these proteins allows for the stability of an epithelial phenotype (Hong et al., 2017). Conversely, TNBC with high expression of DDR1 or RUNX1 had more negative patient outcomes. This has been attributed to RUNX1's driving of proliferation (Fernández et al., 2023) and DDR1's ability to create an immune-exclusive environment through the alignment of collagen fibers (Sun et al., 2021). It was also interesting to see that both DDR1 and RUNX1 or DDR1 and CBFβ are often co-mutated, not just in breast cancer but across many cancer types, increasing the evidence that this is a vital target that can give tumors increased fitness. This data shows that the DDR1-RUNX1 axis is an interesting target for specific breast tumors and that additional work must be done to understand more of the roles that these two genes play.

Use of this gel could also be used to directly support the growth of primary tissue with oncogenic mutations can give glimpses into cancers onset. Our use of structures initiated from single cell primary samples harboring a BRCA1 mutation shows that these

structures do form on a normal timeline compared to healthy tissue, and form structures with similar morphological traits as wild type. While initial testing shows that these structures respond to doxorubicin treatment through the induction of γH2AX at the site of double strand breaks, future work needs to be done to determine both the best timing and best way to visualize and quantify the accumulation and subsequent response of double strand breaks. With the right conditions, it is possible that work with samples containing mutations such as this could lead to direct observations of oncogenesis. Observations such as these made through a biomimetic model are the kind necessary for defining new classes of personalized therapy.

Ideally this model would be used to not only observe how cancer becomes metastatic, but also the initiation of oncogenesis. Breast Cancers such as ductal carcinomas and lobular carcinomas in-situ (DCIS or LCIS respectively) act as ticking timebombs. While a majority remain benign, many still eventually progress and metastasize, forcing women to make life-changing treatment decisions that are potentially unnecessary (Ugai et al., 2022). Future work utilizing the hydrogel should focus on both the initiation of oncogenesis and drivers of metastasis to help advise clinical decisions. The ability to further study what factors drives these tumors to become malignant may not only inform the best kind of treatment but could also allow for the development of new treatments.

In the short term it is likely that this model will also be used to broaden our understanding of the development of tissues other than the breast. The tunable nature of the model gives a wide range of conditions able to be achieved through modulation of the matrix components, base growth media, supplements, and other factors like stiffness and

density. Through the right conditions, likely determined through analysis of tissue resident matrix, it is possible that this 3D hydrogel may support the life and biomimetic morphogenesis of other tissue types leading to new models and therapies. As collagen type I makes up 90% of human protein mass (Naomi et al., 2021), it is likely that even small adjustments may allow support. This was seen with appendix organoids where the only adjustment of any parameters was the change in base media, crypt cells from health appendix tissue were able to be cultured and formed 3D structures. Other epithelial tissues, especially glandular tissue like salivary gland or thyroid, or tissue in which collagen I is a major component of the matrix such as skin or cornea may also find success forming organoids in 3D hydrogel culture. While the development of a single-cell-derived biomimetic model for use in other organs is necessary for translatable research, it is hard to obtain the healthy patient samples required for its development. This model may also support the growth of stem cells derived from HPSCs which could circumvent the need for patient biopsies and could allow for the quick development of new tissue models without the worry of using rare samples. Continued investigation into the growth of all stem cell types in this hydrogel system could add new context to the study of many tissue systems.

One long term goal for this model would be to harness the power of its diverse structural phenotypes across development for large scale screens. While few pathways have currently been probed in this manner, this work exists as a pilot study in which common signaling pathways may be identified through use of these different morphologies across the breast's development. A comprehensive database containing these phenotypes from single cell derived organoids exposed to other small molecules,

could allow for high-throughput screens in which exposure to small molecule library compounds could quickly be screened for function and common pathways, which could also quickly expand the knowledge of signaling occurring during the different stages of breast development. As the model has already been miniaturized for growth in 96 well plates, high-throughput screens only require a more comprehensive library of structural phenotypes. In theory it is possible that this screening technique could be used for the understanding of developmental pathways in organoids from other tissues as well, making it a powerful technique that should be further investigated.

A long-term goal in the study of any patient derived regenerative tissue is for its re-implantation to help with consequences of surgery and disease. The use of resident adult stem cells residing in a patient's own tissue to drive recovery is the ideal situation for surgical reimplantation, as it is sustainably sourced and would prevent rejection. One day this technique may be required for breast tissue, as increasing breast cancer rates continue to impact younger women who may still want to have children and breast feed (Ugai et al., 2022). The ability to biopsy areas of healthy breast tissue, screen it for oncogenic markers, regrow structures, and reimplant them into the breast is closer than ever. Even more ideal would be the ability to genetically edit out mutations, edit in protective qualities, allowing for peace of mind that the tissue being added back in would not only function properly, but would be cancer free. While some of these processes may be years away in development, it has been shown that the premise of gene therapy could also be used in a clonally derived breast TDLU.

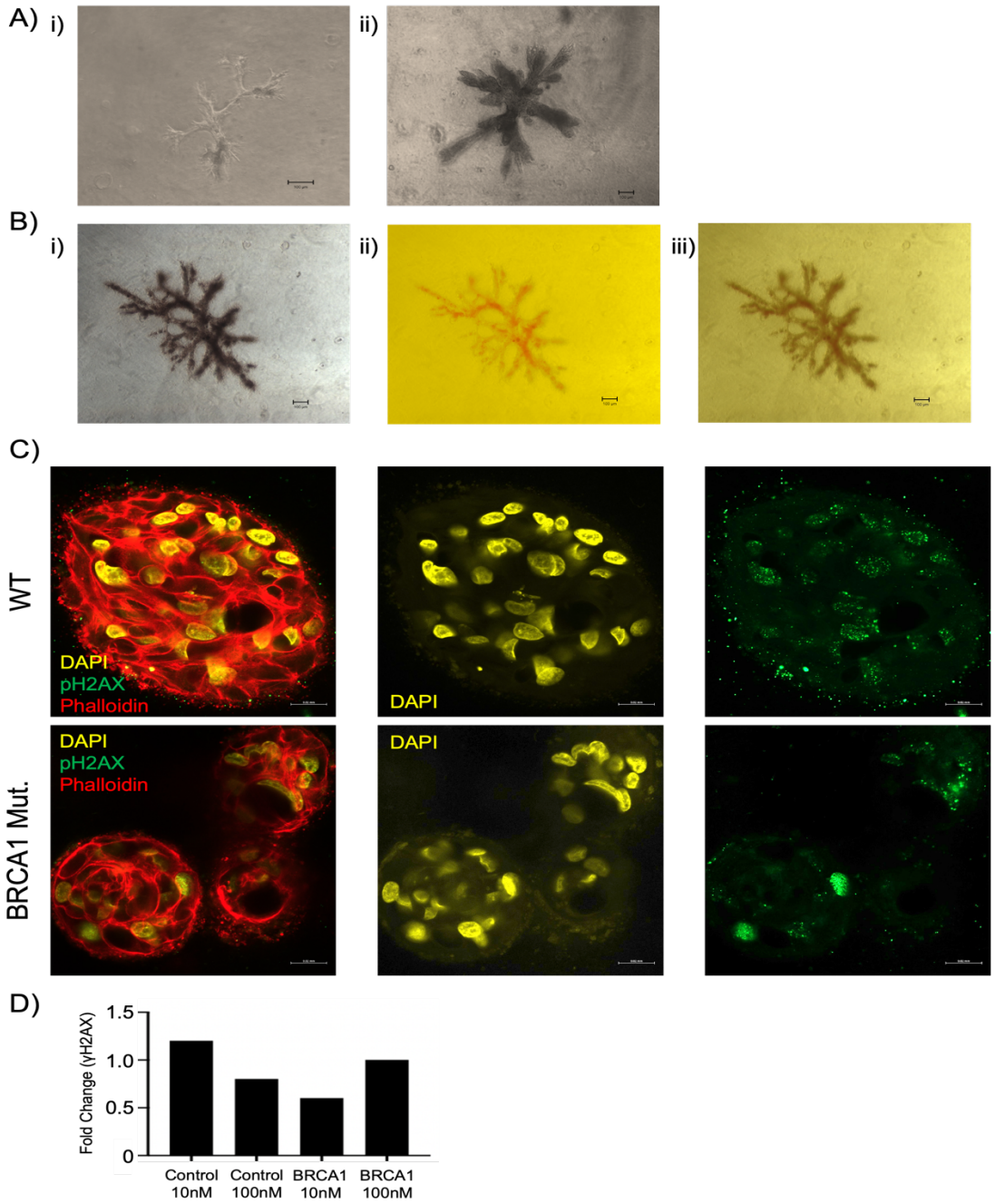
To conclude this thesis, I believe that this work makes two statements to fill gaps in scientific knowledge. The first is on the relevance of the DDR1-RUNX1 axis to human

biology. While DDR1 is clearly an important signaling protein as it gathers information about the cellular niche and converts it into cellular activity, there has been a lack of understanding in the regulation of DDR1's downstream signaling, potentially due to lack of collagen in 2D culture experimental systems preventing it from standing out as a target in many human experiments. As models adopt more biomimetic conditions, it is likely that DDR1 will continue to emerge as an essential receptor for regeneration and organogenesis and that this activity will be in part through RUNX1, as their relationship is hinted at as being concurrent across tissues (Table 4.4.).

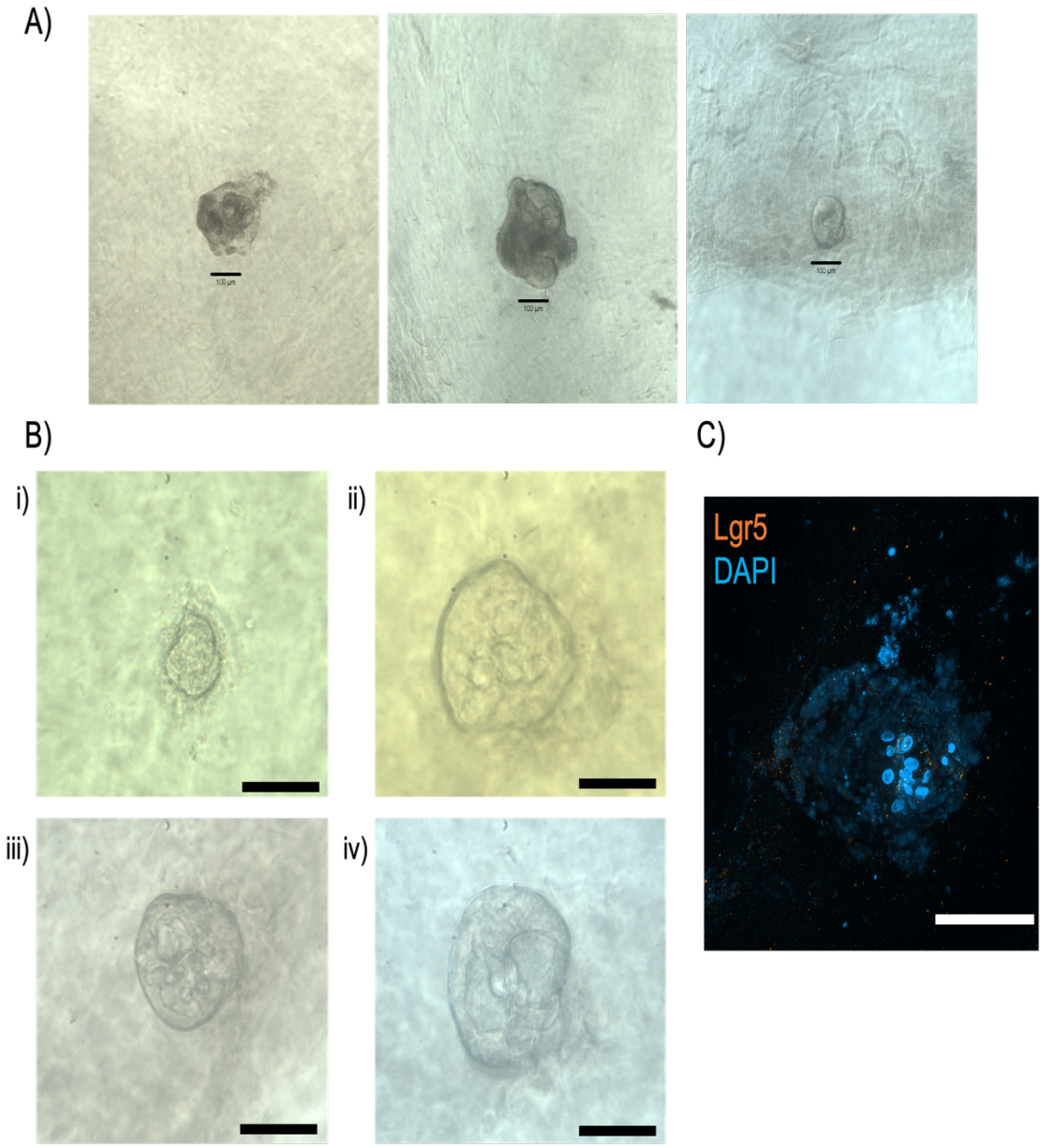
The second and more important statement this work makes is through the enhancement and promotion of this organoid model derived from primary single cells. The use of readily available tissue and easily acquirable materials allows for utilization of this model to be easily ascertainable. Use of this model allows for quantifiable data across treatment groups and allows for biomimetic knowledge of not just cellular expression-based responses, but also for lineage and morphological changes as well. This clear increase in knowledge from every experiment run brings more readily translatable data from the bench to the clinic. This thesis work, which would not have been as impactful without such a model, is just the beginning of what can be accomplished with a single cell derived organoid.

Chapter VI: Appendices

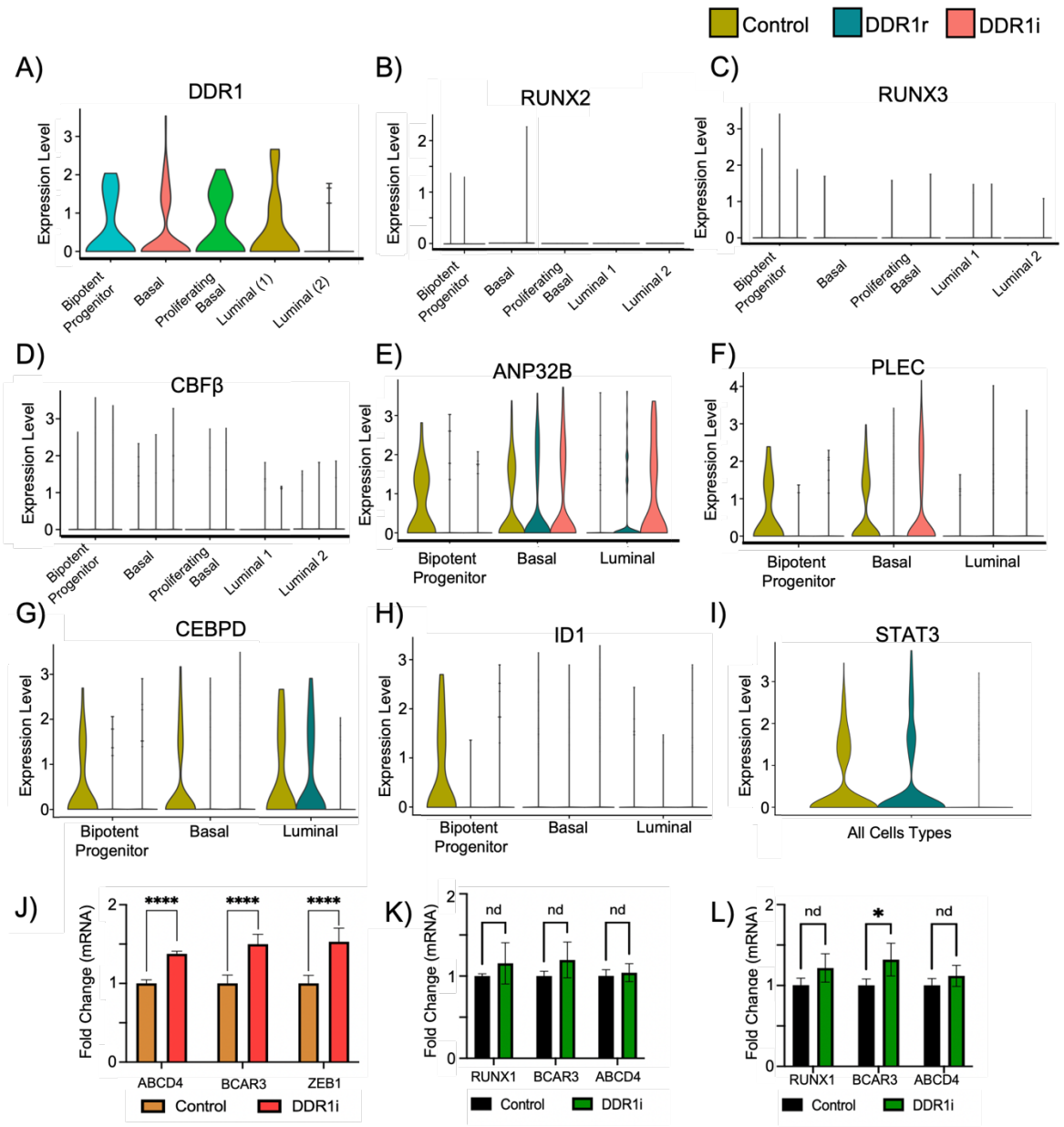
6.1. Appendix Figures



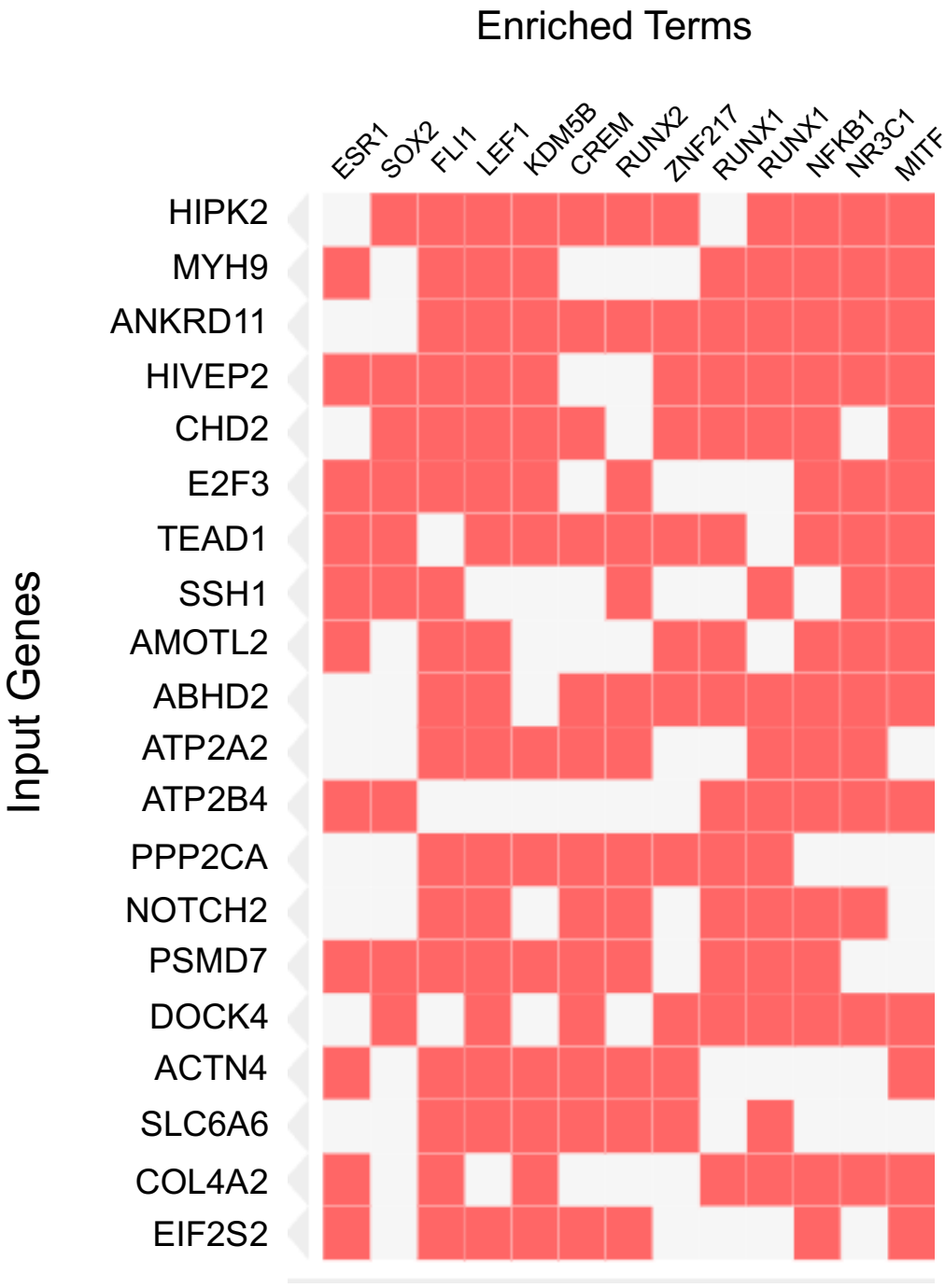
Appendix Figure 6.1. Preliminary Characterization of Doxorubicin Treatment in BRCA1 Mutant Primary Single Cell Derived Organoids: A) Panel of primary appendix organoids on day 17 Scale bar = 100um. B) Appendix organoid i) on day 7, ii) day 12, iii) day 14, and again iv) on day 17. Scale bar = 50um. C) Immunofluorescent staining of Control and BRCA1 mutant primary samples grown for 22 days and stained for DSB damage marker pH2AX. Scale bar = 20um. D) Fold change of pH2AX across timepoints and treatments with control and BRCA1 mut primary organoids.



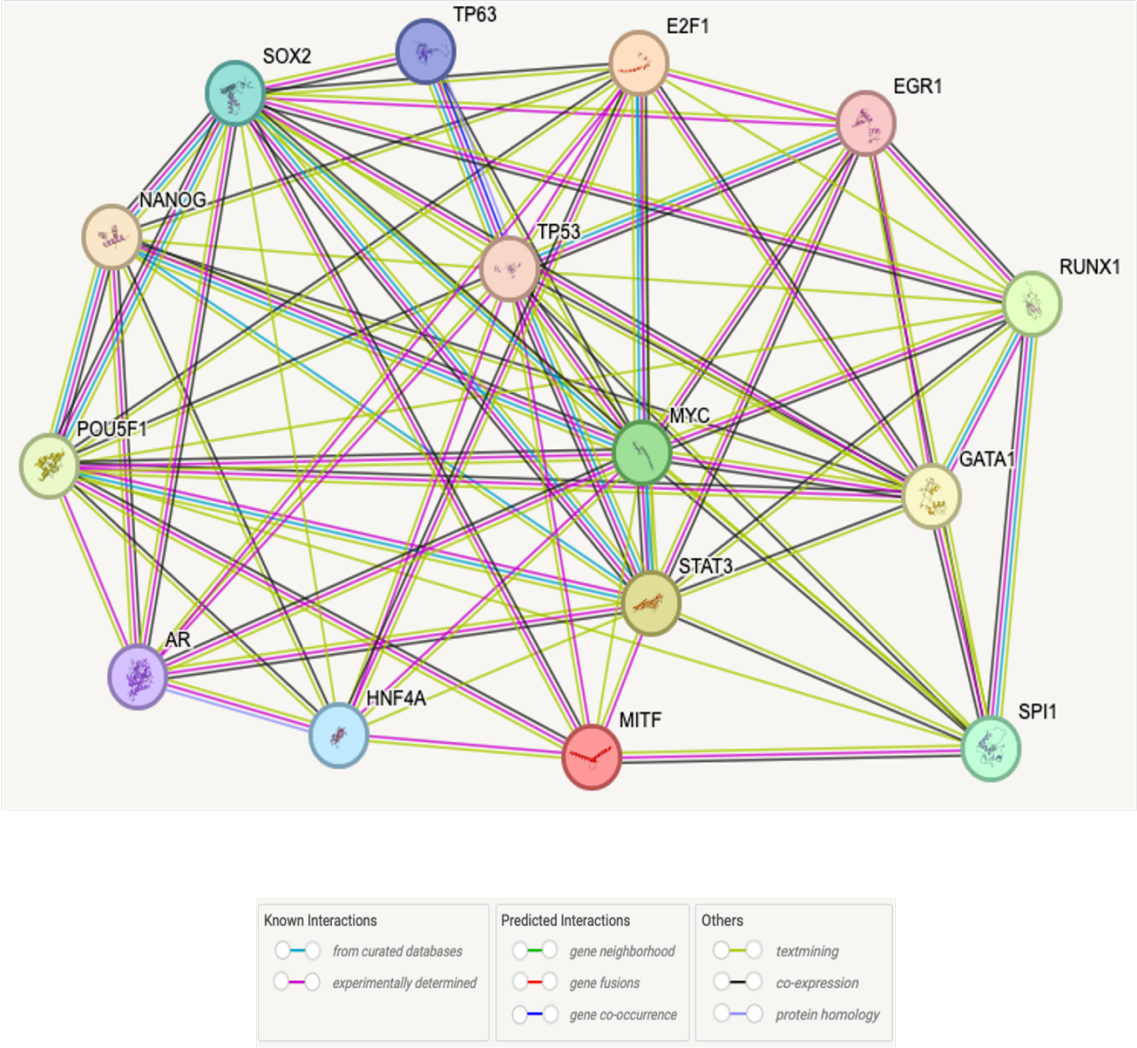
Appendix Figure 6.2. Preliminary Characterization of Appendix Organoids: A) Development of organoid carrying BRCA1 Mutation on day 10 and day 22 after seeding. B) i) Brightfield image of BRCA1 mutant organoid on day 17. Scale bar = 100um. ii) Fluorescent image of DXR localization in BRCA1 mut. structure. iii) overlap of DXR fluorescence over brightfield. Scale bar = 100um. C) Immunofluorescence of appendix organoid with DAPI and intestinal crypt cell marker LGR5. Scale bar = 50um.



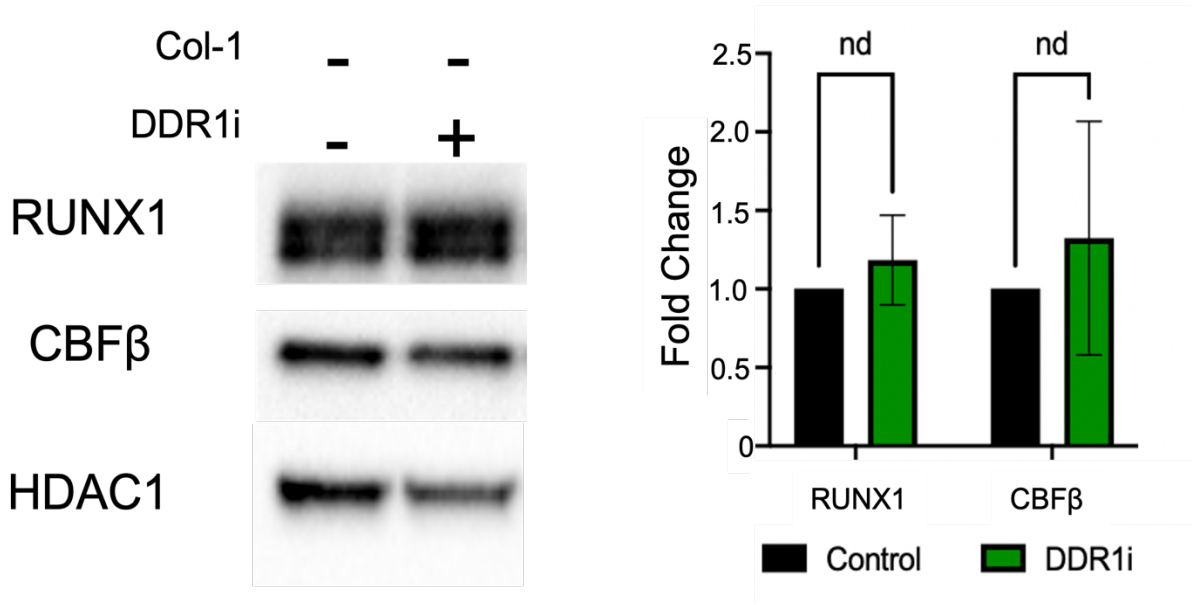
Appendix Figure 6.3. Differential Expression Analysis in Response to DDR1 Inhibition: A) Violin plots display the distribution of DDR1 expression across various epithelial breast cell types. B-I) These plots depict the distribution of expression in RUNX2, RUNX3, and RUNX1 target genes (ANP32B, PLEC, ID1, and STAT3) within epithelial clusters in primary tissue organoids grown for 14 days under either control DDR1i or DDR1r conditions. J) Quantification of the expression of RUNX1 target genes from MCF10F cells in a 2D collagen stimulation assay. K) mRNA expression levels of RUNX1 target genes in MCF10A cells without collagen I stimulation under control or DDR1i conditions. L) mRNA expression levels of RUNX1 target genes in MCF10F cells without collagen I stimulation under control or DDR1i conditions. The data are expressed as Mean ± SD. Statistical significance was determined through multiple t-tests, with significance levels indicated as follows: *p-value < 0.05, **p-value < 0.01, ***p-value < 0.001, ****p-value < 0.0001.



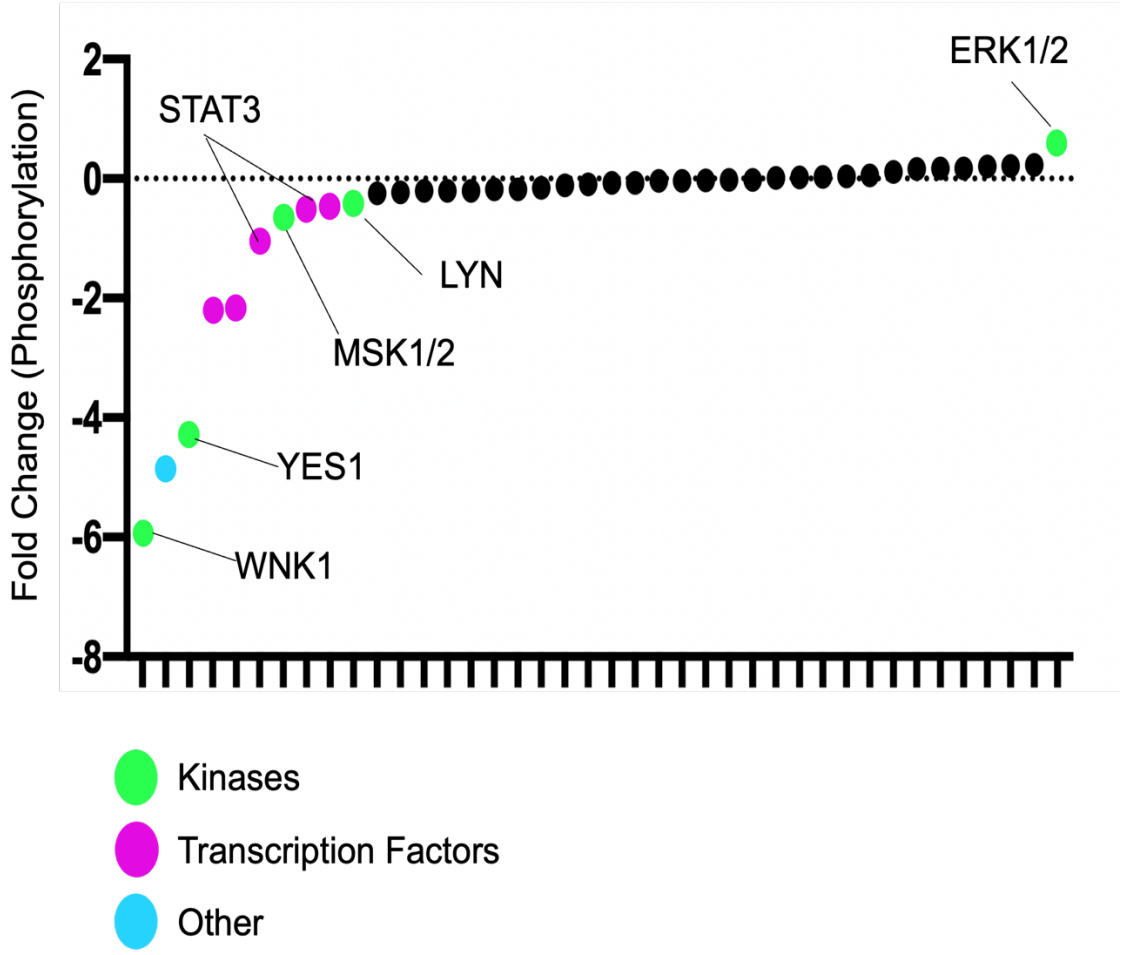
Appendix Figure 6.4. Transcription Factor Enrichment for Core DDR1-RUNX1 Gene Set: A hierarchically clustered heatmap illustrating the association between different transcription factor hits and genes differentially expressed from the experimentally determined overlapping gene set (input). The data utilized for this analysis are sourced from the TRRUST database. In the heatmap, red squares signify that the associated transcription factor is known to regulate the gene in this dataset. This visualization was generated using ENRICHR.



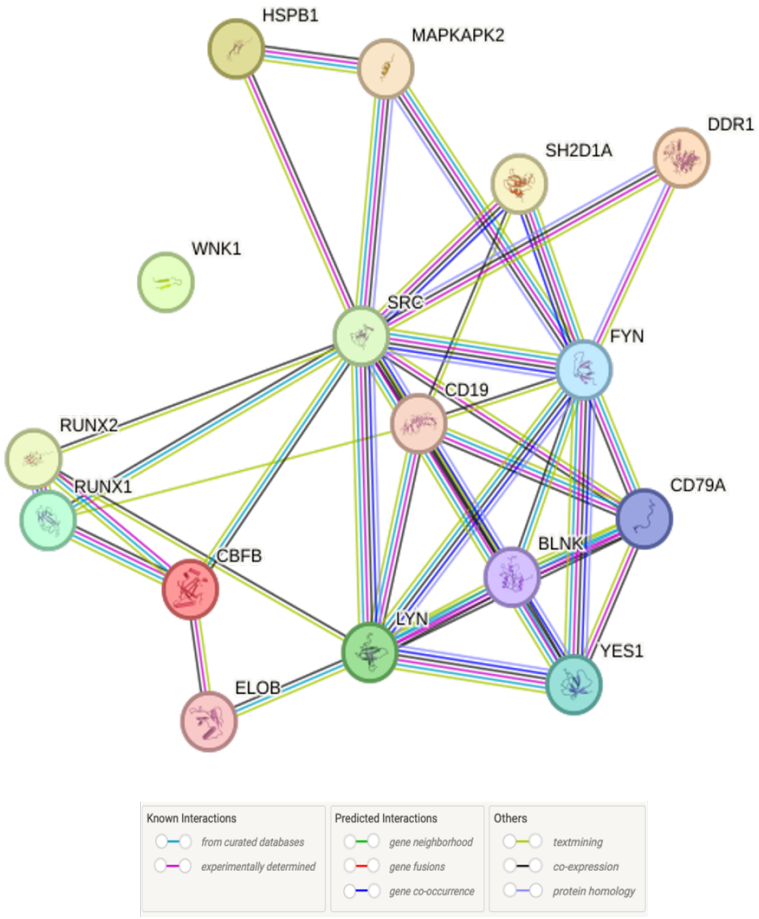
Appendix Figure 6.5. String Analysis of Transcription Factor Hits: STRINGDB network of protein-protein interactions between the transcription factor hits from ChEA database known to regulate the overlapping DDR1-RUNX1 transcriptome.



Appendix Figure 6.6. Effects of DDR1i on RUNX1 and CBFβ Expression Without Collagen-1 Stimulation: Representative western blot and quantification of nuclear RUNX1 and CBFβ without collagen stimulation in control or DDR1i conditions.



Appendix Figure 6.7. Potential DDR1-RUNX1 Protein Intermediates through a Phospho-Kinase Array: Phosphokinase Array showing loss or gain of phosphorylation with DDR1i treatment.



Appendix Figure 6.8. String Analysis of Potential DDR1-RUNX1 Protein Intermediates: STRINGDB network of protein-protein interactions between DDR1 and RUNX1 with hits from the phosphokinase array.

Chapter VII: Bibliography

Abbas, M., Moradi, F., Hu, W., Regudo, K.L., Osborne, M., Pettipas, J., Atallah, D.S., Hachem, R., Ott-Peron, N., and Stuart, J.A. (2021). Vertebrate cell culture as an experimental approach – limitations and solutions. Comparative Biochemistry and Physiology Part B: Biochemistry and Molecular Biology *254*, 110570. https://doi.org/10.1016/j.cbpb.2021.110570.

Abbonante, V., Gruppi, C., Rubel, D., Gross, O., Moratti, R., and Balduini, A. (2013). Discoidin Domain Receptor 1 Protein Is a Novel Modulator of Megakaryocyte-Collagen Interactions *. Journal of Biological Chemistry *288*, 16738–16746. https://doi.org/10.1074/jbc.M112.431528.

Adya, N., Castilla, L.H., and Liu, P.P. (2000). Function of CBFβ/Bro proteins. Seminars in Cell & Developmental Biology *11*, 361–368. https://doi.org/10.1006/scdb.2000.0189.

Aigner, P., Mizutani, T., Horvath, J., Eder, T., Heber, S., Lind, K., Just, V., Moll, H.P., Yeroslaviz, A., Fischer, M.J.M., et al. (2019). STAT3β is a tumor suppressor in acute myeloid leukemia. Blood Advances *3*, 1989–2002. https://doi.org/10.1182/bloodadvances.2018026385.

Alcaraz, J., Xu, R., Mori, H., Nelson, C.M., Mroue, R., Spencer, V.A., Brownfield, D., Radisky, D.C., Bustamante, C., and Bissell, M.J. (2008). Laminin and biomimetic extracellular elasticity enhance functional differentiation in mammary epithelia. EMBO J *27*, 2829–2838. https://doi.org/10.1038/emboj.2008.206.

Allouba, M.H., ElGuindy, A.M., Krishnamoorthy, N., Yacoub, M.H., and Aguib, Y.E. (2015). NaNog: A pluripotency homeobox (master) molecule. Global Cardiology Science & Practice *2015*. https://doi.org/10.5339/gcsp.2015.36.

Alves, F., Vogel, W., Mossie, K., Millauer, B., Höfler, H., and Ullrich, A. (1995). Distinct structural characteristics of discoidin I subfamily receptor tyrosine kinases and complementary expression in human cancer. Oncogene *10*, 609–618. .

Arendt, L.M., Keller, P.J., Skibinski, A., Goncalves, K., Naber, S.P., Buchsbaum, R.J., Gilmore, H., Come, S.E., and Kuperwasser, C. (2014). Anatomical localization of progenitor cells in human breast tissue reveals enrichment of uncommitted cells within immature lobules. Breast Cancer Research *16*, 453. https://doi.org/10.1186/s13058-014-0453-3.

Ariffin, N.S. (2022). RUNX1 as a Novel Molecular Target for Breast Cancer. Clinical Breast Cancer *22*, 499–506. https://doi.org/10.1016/j.clbc.2022.04.006.

Aubrey, B.J., Strasser, A., and Kelly, G.L. (2016). Tumor-Suppressor Functions of the TP53 Pathway. Cold Spring Harb Perspect Med *6*, a026062. https://doi.org/10.1101/cshperspect.a026062.

Baptista, L.S., Porrini, C., Kronemberger, G.S., Kelly, D.J., and Perrault, C.M. (2022). 3D organ-on-a-chip: The convergence of microphysiological systems and organoids. Frontiers in Cell and Developmental Biology *10*. .

Barbieri, C.E., and Pietenpol, J.A. (2006). p63 and epithelial biology. Exp Cell Res *312*, 695–706. https://doi.org/10.1016/j.yexcr.2005.11.028.

Barker, K.T., Martindale, J.E., Mitchell, P.J., Kamalati, T., Page, M.J., Phippard, D.J., Dale, T.C., Gusterson, B.A., and Crompton, M.R. (1995). Expression patterns of the novel receptor-like tyrosine kinase, DDR, in human breast tumours. Oncogene *10*, 569–575. .

Bauer, D.C., Buske, F.A., and Bailey, T.L. (2010). Dual-functioning transcription factors in the developmental gene network of Drosophila melanogaster. BMC Bioinformatics *11*, 366. https://doi.org/10.1186/1471-2105-11-366.

Beghini, A. (2019). Core Binding Factor Leukemia: Chromatin Remodeling Moves Towards Oncogenic Transcription. Cancers *11*, 1973. https://doi.org/10.3390/cancers11121973.

Ben Arfi, K., Schneider, C., Bennasroune, A., Bouland, N., Wolak-Thierry, A., Collin, G., Le, C.C., Toussaint, K., Hachet, C., Lehrter, V., et al. (2022). Discoidin Domain Receptor 1 Expression in Colon Cancer: Roles and Prognosis Impact. Cancers (Basel) *14*, 928. https://doi.org/10.3390/cancers14040928.

Ben-David, U., Siranosian, B., Ha, G., Tang, H., Oren, Y., Hinohara, K., Strathdee, C.A., Dempster, J., Lyons, N.J., Burns, R., et al. (2018). Genetic and transcriptional evolution alters cancer cell line drug response. Nature *560*, 325–330. https://doi.org/10.1038/s41586-018-0409-3.

Berestjuk, I., Lecacheur, M., Carminati, A., Diazzi, S., Rovera, C., Prod'homme, V., Ohanna, M., Popovic, A., Mallavialle, A., Larbret, F., et al. (2022). Targeting Discoidin Domain Receptors DDR1 and DDR2 overcomes matrix-mediated tumor cell adaptation and tolerance to BRAF-targeted therapy in melanoma. EMBO Mol Med *14*, e11814. https://doi.org/10.15252/emmm.201911814.

Berry, R.J.A. (1900). The True Cæcal Apex, or the Vermiform Appendix: Its Minute and Comparative Anatomy. J Anat Physiol *35*, 83-100.9. .

van Bragt, M.P., Hu, X., Xie, Y., and Li, Z. (2014). RUNX1, a transcription factor mutated in breast cancer, controls the fate of ER-positive mammary luminal cells. ELife *3*, e03881. https://doi.org/10.7554/eLife.03881.

Bresciani, E., Carrington, B., Wincovitch, S., Jones, M., Gore, A.V., Weinstein, B.M., Sood, R., and Liu, P.P. (2014). CBFβ and RUNX1 are required at 2 different steps during the development of hematopoietic stem cells in zebrafish. Blood *124*, 70–78. https://doi.org/10.1182/blood-2013-10-531988.

Bresslau, E. (1920). The Mammary Apparatus of the Mammalia in the Light of Ontogenesis and Phylogenesis (Methuen).

Brisken, C., and O'Malley, B. (2010). Hormone Action in the Mammary Gland. Cold Spring Harb Perspect Biol *2*, a003178. https://doi.org/10.1101/cshperspect.a003178.

Brisken, C., Kaur, S., Chavarria, T.E., Binart, N., Sutherland, R.L., Weinberg, R.A., Kelly, P.A., and Ormandy, C.J. (1999). Prolactin Controls Mammary Gland Development via Direct and Indirect Mechanisms. Developmental Biology *210*, 96–106. https://doi.org/10.1006/dbio.1999.9271.

de Bruijn, I., Kundra, R., Mastrogiacomo, B., Tran, T.N., Sikina, L., Mazor, T., Li, X., Ochoa, A., Zhao, G., Lai, B., et al. (2023). Analysis and Visualization of Longitudinal Genomic and Clinical Data from the AACR Project GENIE Biopharma Collaborative in cBioPortal. Cancer Res *83*, 3861–3867. https://doi.org/10.1158/0008-5472.CAN-23-0816.

Bushweller, J.H. (2019). Targeting transcription factors in cancer — from undruggable to reality. Nat Rev Cancer *19*, 611–624. https://doi.org/10.1038/s41568-019-0196-7.

Canon, J., and Banerjee, U. (2000). Runt and Lozenge function in Drosophila development. Seminars in Cell & Developmental Biology *11*, 327–336. https://doi.org/10.1006/scdb.2000.0185.

Capuco, A.V., and Akers, R.M. (2009). The origin and evolution of lactation. J Biol *8*, 37. https://doi.org/10.1186/jbiol139.

Cerami1, E., Gao, J., Dogrusoz, U., Gross, B.E., Sumer, S.O., Aksoy, B.A., Jacobsen, A., Byrne, C.J., Heuer, M.L., Larsson, E., et al. (2012). The cBio Cancer Genomics Portal: An Open Platform for Exploring Multidimensional Cancer Genomics Data. Cancer Discov *2*, 401–404. https://doi.org/10.1158/2159-8290.CD-12-0095.

Chan, A., Ma, S., Pearson, B.J., and Chan, D. (2021). Collagen IV differentially regulates planarian stem cell potency and lineage progression. Proc Natl Acad Sci U S A *118*, e2021251118. https://doi.org/10.1073/pnas.2021251118.

Chang, C., Lee, S.O., Wang, R.-S., Yeh, S., and Chang, T.-M. (2013). Androgen Receptor (AR) Physiological Roles in Male and Female Reproductive Systems: Lessons Learned from AR-Knockout Mice Lacking AR in Selective Cells. Biol Reprod *89*, 21. https://doi.org/10.1095/biolreprod.113.109132.

Chen, C., Deng, J., Yu, X., Wu, F., Men, K., Yang, Q., Zhu, Y., Liu, X., and Jiang, Q. (2016). Identification of novel inhibitors of DDR1 against idiopathic pulmonary fibrosis by integrative transcriptome meta-analysis, computational and experimental screening. Mol. BioSyst. *12*, 1540–1551. https://doi.org/10.1039/C5MB00911A.

Chen, L., Kong, X., Fang, Y., Paunikar, S., Wang, X., Brown, J.A.L., Bourke, E., Li, X., and Wang, J. (2021). Recent Advances in the Role of Discoidin Domain Receptor Tyrosine Kinase 1 and Discoidin Domain Receptor Tyrosine Kinase 2 in Breast and Ovarian Cancer. Frontiers in Cell and Developmental Biology *9*. .

Chen, Y., He, Z., Yang, S., Chen, C., Xiong, W., He, Y., and Liu, S. (2023). RUNX1 knockdown induced apoptosis and impaired EMT in high-grade serous ovarian cancer cells. Journal of Translational Medicine *21*, 886. https://doi.org/10.1186/s12967-023-04762-8.

Chittleborough, G.D., and Treagust, D.F. (2009). Why Models are Advantageous to Learning Science. Educación Química *20*, 12–17. https://doi.org/10.1016/S0187-893X(18)30003-X.

Choi, Ah., Illendula, A., Pulikkan, J.A., Roderick, J.E., Tesell, J., Yu, J., Hermance, N., Zhu, L.J., Castilla, L.H., Bushweller, J.H., et al. (2017). RUNX1 is required for oncogenic Myb and Myc enhancer activity in T-cell acute lymphoblastic leukemia. Blood *130*, 1722–1733. https://doi.org/10.1182/blood-2017-03-775536.

Coffman, J.A. (2003). Runx transcription factors and the developmental balance between cell proliferation and differentiation. Cell Biology International *27*, 315–324. https://doi.org/10.1016/S1065-6995(03)00018-0.

Corrò, C., Novellasdemunt, L., and Li, V.S.W. (2020). A brief history of organoids. Am J Physiol Cell Physiol *319*, C151–C165. https://doi.org/10.1152/ajpcell.00120.2020.

Couldrey, C., Moitra, J., Vinson, C., Anver, M., Nagashima, K., and Green, J. (2002). Adipose tissue: A vital in vivo role in mammary gland development but not differentiation. Developmental Dynamics *223*, 459–468. https://doi.org/10.1002/dvdy.10065.

Cowin, P., and Wysolmerski, J. (2010). Molecular Mechanisms Guiding Embryonic Mammary Gland Development. Cold Spring Harb Perspect Biol *2*, a003251. https://doi.org/10.1101/cshperspect.a003251.

Craig, D.J., Crawford, E.L., Chen, H., Grogan, E.L., Deppen, S.A., Morrison, T., Antic, S.L., Massion, P.P., and Willey, J.C. (2023). TP53 mutation prevalence in normal airway epithelium as a biomarker for lung cancer risk. BMC Cancer *23*, 783. https://doi.org/10.1186/s12885-023-11266-7.

Curtis, C., Shah, S.P., Chin, S.-F., Turashvili, G., Rueda, O.M., Dunning, M.J., Speed, D., Lynch, A.G., Samarajiwa, S., Yuan, Y., et al. (2012). The genomic and transcriptomic architecture of 2,000 breast tumours reveals novel subgroups. Nature *486*, 346–352. https://doi.org/10.1038/nature10983.

Dai, X., Cheng, H., Bai, Z., and Li, J. (2017). Breast Cancer Cell Line Classification and Its Relevance with Breast Tumor Subtyping. J Cancer *8*, 3131–3141. https://doi.org/10.7150/jca.18457.

Dang, C.V. (2012). MYC on the Path to Cancer. Cell *149*, 22–35. https://doi.org/10.1016/j.cell.2012.03.003.

Daniel, C.W., De Ome, K.B., Young, J.T., Blair, P.B., and Faulkin, L.J. (1968). The in vivo life span of normal and preneoplastic mouse mammary glands: a serial transplantation study. Proc Natl Acad Sci U S A *61*, 53–60. .

Davison, C., Morelli, R., Knowlson, C., McKechnie, M., Carson, R., Stachtea, X., McLaughlin, K.A., Prise, V.E., Savage, K., Wilson, R.H., et al. (2021). Targeting nucleotide metabolism enhances the efficacy of anthracyclines and anti-metabolites in triple-negative breast cancer. Npj Breast Cancer *7*, 1–13. https://doi.org/10.1038/s41523-021-00245-5.

Debnath, J., Muthuswamy, S.K., and Brugge, J.S. (2003). Morphogenesis and oncogenesis of MCF-10A mammary epithelial acini grown in three-dimensional basement membrane cultures. Methods *30*, 256–268. https://doi.org/10.1016/S1046-2023(03)00032-X.

Dontu, G., and Ince, T.A. (2015). Of Mice and Women: A Comparative Tissue Biology Perspective of Breast Stem Cells and Differentiation. J Mammary Gland Biol Neoplasia *20*, 51–62. https://doi.org/10.1007/s10911-015-9341-4.

Du, Z., and Lovly, C.M. (2018). Mechanisms of receptor tyrosine kinase activation in cancer. Molecular Cancer *17*, 58. https://doi.org/10.1186/s12943-018-0782-4.

Elagib, K.E., Racke, F.K., Mogass, M., Khetawat, R., Delehanty, L.L., and Goldfarb, A.N. (2003). RUNX1 and GATA-1 coexpression and cooperation in megakaryocytic differentiation. Blood *101*, 4333–4341. https://doi.org/10.1182/blood-2002-09-2708.

Elkamhawy, A., Lu, Q., Nada, H., Woo, J., Quan, G., and Lee, K. (2021). The Journey of DDR1 and DDR2 Kinase Inhibitors as Rising Stars in the Fight Against Cancer. Int J Mol Sci *22*, 6535. https://doi.org/10.3390/ijms22126535.

Fang, Z., Lin, M., Li, C., Liu, H., and Gong, C. (2020). A comprehensive review of the roles of E2F1 in colon cancer. Am J Cancer Res *10*, 757–768. .

Fein, D.E.C., Traugh, N., Rauner, G., Mal, Y., Trepicchio, C., and Kuperwasser, C. (2023). Selective inhibition of fibroblast-specific Domain Discoidin Receptor 1 (DDR1) reduces collagen deposition and modulates fibroblast-specific cytokine

release within the breast microenvironment. 2023.08.11.552937. https://doi.org/10.1101/2023.08.11.552937.
Fernandes, U., Guidi, G., Martins, D., Vieira, B., Leal, C., Marques, C., Freitas, F., Dupont, M., Ribeiro, J., Gomes, C., et al. (2023). Breast cancer in young women: a rising threat: A 5-year follow-up comparative study. Porto Biomed J *8*, e213. https://doi.org/10.1097/j.pbj.0000000000000213.
Fernández, N.B., Sosa, S.M., Roberts, J.T., Recouvreux, M.S., Rocha-Viegas, L., Christenson, J.L., Spoelstra, N.S., Couto, F.L., Raimondi, A.R., Richer, J.K., et al. (2023). RUNX1 Is Regulated by Androgen Receptor to Promote Cancer Stem Markers and Chemotherapy Resistance in Triple Negative Breast Cancer. Cells *12*, 444. https://doi.org/10.3390/cells12030444.
Fernández, P.A., Buchmann, B., Goychuk, A., Engelbrecht, L.K., Raich, M.K., Scheel, C.H., Frey, E., and Bausch, A.R. (2021). Surface-tension-induced budding drives alveologenesis in human mammary gland organoids. Nat. Phys. *17*, 1130–1136. https://doi.org/10.1038/s41567-021-01336-7.
Ferrari, N., Mohammed, Z.M.A., Nixon, C., Mason, S.M., Mallon, E., McMillan, D.C., Morris, J.S., Cameron, E.R., Edwards, J., and Blyth, K. (2014). Expression of RUNX1 Correlates with Poor Patient Prognosis in Triple Negative Breast Cancer. PLOS ONE *9*, e100759. https://doi.org/10.1371/journal.pone.0100759.
Ferraro, F., Celso, C.L., and Scadden, D. (2010). ADULT STEM CELLS AND THEIR NICHES. Adv Exp Med Biol *695*, 155–168. https://doi.org/10.1007/978-1-4419-7037-4_11.
Fillmore, C.M., and Kuperwasser, C. (2008). Human breast cancer cell lines contain stem-like cells that self-renew, give rise to phenotypically diverse progeny and survive chemotherapy. Breast Cancer Research *10*, R25. https://doi.org/10.1186/bcr1982.
Fischbach, C., Chen, R., Matsumoto, T., Schmelzle, T., Brugge, J.S., Polverini, P.J., and Mooney, D.J. (2007). Engineering tumors with 3D scaffolds. Nat Methods *4*, 855–860. https://doi.org/10.1038/nmeth1085.
Fowler, A.J., Hebron, M., Balaraman, K., Shi, W., Missner, A.A., Greenzaid, J.D., Chiu, T.L., Ullman, C., Weatherdon, E., Duka, V., et al. (2020). Discoidin Domain Receptor 1 is a therapeutic target for neurodegenerative diseases. Human Molecular Genetics *29*, 2882–2898. https://doi.org/10.1093/hmg/ddaa177.
Franzén, O., Gan, L.-M., and Björkegren, J.L.M. (2019). PanglaoDB: a web server for exploration of mouse and human single-cell RNA sequencing data. Database *2019*, baz046. https://doi.org/10.1093/database/baz046.
Friedman, A.D. (2009). Cell Cycle and Developmental Control of Hematopoiesis by Runx1. J Cell Physiol *219*, 520–524. https://doi.org/10.1002/jcp.21738.
Friedman, L.S., Ostermeyer, E.A., Szabo, C.I., Dowd, P., Lynch, E.D., Rowell, S.E., and King, M.-C. (1994). Confirmation of BRCA1 by analysis of germline mutations linked to breast and ovarian cancer in ten families. Nat Genet *8*, 399–404. https://doi.org/10.1038/ng1294-399.
Fukui, H., Rünker, A., Fabel, K., Buchholz, F., and Kempermann, G. (2018). Transcription factor Runx1 is pro-neurogenic in adult hippocampal precursor cells. PLoS One *13*, e0190789. https://doi.org/10.1371/journal.pone.0190789.

Gao, J., Aksoy, B.A., Dogrusoz, U., Dresdner, G., Gross, B., Sumer, S.O., Sun, Y., Jacobsen, A., Sinha, R., Larsson, E., et al. (2013). Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal. Sci Signal *6*, pl1. https://doi.org/10.1126/scisignal.2004088.

Garbe, J.C., Pepin, F., Pelissier, F., Sputova, K., Fridriksdottir, A.J., Guo, D.E., Villadsen, R., Park, M., Petersen, O.W., Borowsky, A.D., et al. (2012). Accumulation of multipotent progenitors with a basal differentiation bias during aging of human mammary epithelia. Cancer Res *72*, 3687–3701. https://doi.org/10.1158/0008-5472.CAN-12-0157.

Gerritsen, M., Yi, G., Tijchon, E., Kuster, J., Schuringa, J.J., Martens, J.H.A., and Vellenga, E. (2019). RUNX1 mutations enhance self-renewal and block granulocytic differentiation in human in vitro models and primary AMLs. Blood Adv *3*, 320–332. https://doi.org/10.1182/bloodadvances.2018024422.

Giaquinto, A.N., Sung, H., Miller, K.D., Kramer, J.L., Newman, L.A., Minihan, A., Jemal, A., and Siegel, R.L. (2022). Breast Cancer Statistics, 2022. CA Cancer J Clin *72*, 524–541. https://doi.org/10.3322/caac.21754.

Gieniec, K.A., and Davis, F.M. (2022). Mammary basal cells: Stars of the show. Biochimica et Biophysica Acta (BBA) - Molecular Cell Research *1869*, 119159. https://doi.org/10.1016/j.bbamcr.2021.119159.

Gonzales, F., Barthélémy, A., Peyrouze, P., Fenwarth, L., Preudhomme, C., Duployez, N., and Cheok, M.H. (2021). Targeting RUNX1 in acute myeloid leukemia: preclinical innovations and therapeutic implications. Expert Opinion on Therapeutic Targets *25*, 299–309. https://doi.org/10.1080/14728222.2021.1915991.

Goyama, S., Huang, G., Kurokawa, M., and Mulloy, J.C. (2015). Posttranslational modifications of RUNX1 as potential anticancer targets. Oncogene *34*, 3483–3492. https://doi.org/10.1038/onc.2014.305.

Gregory, W.K. (1910). The Orders of Mammals (The Trustees).

Growney, J.D., Shigematsu, H., Li, Z., Lee, B.H., Adelsperger, J., Rowan, R., Curley, D.P., Kutok, J.L., Akashi, K., Williams, I.R., et al. (2005). Loss of Runx1 perturbs adult hematopoiesis and is associated with a myeloproliferative phenotype. Blood *106*, 494–504. https://doi.org/10.1182/blood-2004-08-3280.

Gupta, P.B., Pastushenko, I., Skibinski, A., Blanpain, C., and Kuperwasser, C. (2019). Phenotypic Plasticity: Driver of Cancer Initiation, Progression, and Therapy Resistance. Cell Stem Cell *24*, 65–78. https://doi.org/10.1016/j.stem.2018.11.011.

Han, H., Shim, H., Shin, D., Shim, J.E., Ko, Y., Shin, J., Kim, H., Cho, A., Kim, E., Lee, T., et al. (2015). TRRUST: a reference database of human transcriptional regulatory interactions. Sci Rep *5*, 11432. https://doi.org/10.1038/srep11432.

Han, Q., Xiao, F., Ma, L., Zhou, J., Wang, L., Cheng, H., Zhu, J., Yao, F., Lyu, J., and Du, L. (2022). DDR1 promotes migration and invasion of breast cancer by modulating the Src-FAK signaling. Neo *69*, 1154–1164. https://doi.org/10.4149/neo_2022_220316N289.

Hanna, C., Kwok, L., Finlay-Schultz, J., Sartorius, C.A., and Cittelly, D.M. (2016). Labeling of Breast Cancer Patient-derived Xenografts with Traceable Reporters for Tumor Growth and Metastasis Studies. J Vis Exp 54944. https://doi.org/10.3791/54944.

Hansson, E.M., Teixeira, A.I., Gustafsson, M.V., Dohda, T., Chapman, G., Meletis, K., Muhr, J., and Lendahl, U. (2006). Recording Notch Signaling in Real Time. Developmental Neuroscience *28*, 118–127. https://doi.org/10.1159/000090758.

Harunaga, J., Hsu, J.C., and Yamada, K.M. (2011). Dynamics of salivary gland morphogenesis. J Dent Res *90*, 1070–1077. https://doi.org/10.1177/0022034511405330.

Haslam, S.Z., and Shyamala, G. (1979). Effect of oestradiol on progesterone receptors in normal mammary glands and its relationship with lactation. Biochem J *182*, 127–131. .

He, J., Xiong, L., Li, Q., Lin, L., Miao, X., Yan, S., Hong, Z., Yang, L., Wen, Y., and Deng, X. (2017). 3D modeling of cancer stem cell niche. Oncotarget *9*, 1326–1345. https://doi.org/10.18632/oncotarget.19847.

Hickman, D.L., Johnson, J., Vemulapalli, T.H., Crisler, J.R., and Shepherd, R. (2017). Commonly Used Animal Models. Principles of Animal Research for Graduate and Undergraduate Students 117–175. https://doi.org/10.1016/B978-0-12-802151-4.00007-4.

Hines, W., Yaswen, P., and Bissell, M. (2015). Modelling breast cancer requires identification and correction of a critical cell lineage-dependent transduction bias. Nature Communications *6*, 6927. https://doi.org/10.1038/ncomms7927.

Hockney, S., Parker, J., Turner, J.E., Todd, X., Todryk, S., Gieling, R.G., Hilgen, G., Simoes, D.C.M., and Pal, D. (2023). Next generation organoid engineering to replace animals in cancer drug testing. Biochemical Pharmacology *213*, 115586. https://doi.org/10.1016/j.bcp.2023.115586.

Hoi, C.S.L., Lee, S.E., Lu, S.-Y., McDermitt, D.J., Osorio, K.M., Piskun, C.M., Peters, R.M., Paus, R., and Tumbar, T. (2010). Runx1 directly promotes proliferation of hair follicle stem cells and epithelial tumor formation in mouse skin. Mol Cell Biol *30*, 2518–2536. https://doi.org/10.1128/MCB.01308-09.

Hojjat-Farsangi, M. (2014). Small-Molecule Inhibitors of the Receptor Tyrosine Kinases: Promising Tools for Targeted Cancer Therapies. Int J Mol Sci *15*, 13768–13801. https://doi.org/10.3390/ijms150813768.

Holen, I., Speirs, V., Morrissey, B., and Blyth, K. (2017). In vivo models in breast cancer research: progress, challenges and future directions. Dis Model Mech *10*, 359–371. https://doi.org/10.1242/dmm.028274.

Hong, D., Messier, T.L., Tye, C.E., Dobson, J.R., Fritz, A.J., Sikora, K.R., Browne, G., Stein, J.L., Lian, J.B., and Stein, G.S. (2017). Runx1 stabilizes the mammary epithelial cell phenotype and prevents epithelial to mesenchymal transition. Oncotarget *8*, 17610–17627. https://doi.org/10.18632/oncotarget.15381.

Hu, W., Zheng, T., and Wang, J. (2011). Regulation of Fertility by the p53 Family Members. Genes Cancer *2*, 420–430. https://doi.org/10.1177/1947601911408892.

Hu, Y., Pan, Q., Zhou, K., Ling, Y., Wang, H., and Li, Y. (2022). RUNX1 inhibits the antiviral immune response against influenza A virus through attenuating type I interferon signaling. Virology Journal *19*, 39. https://doi.org/10.1186/s12985-022-01764-8.

Huang, C., Zhao, S., Yang, Y., Guo, T., Ke, H., Mi, X., Qin, Y., Chen, Z.-J., and Zhao, S. TP63 gain-of-function mutations cause premature ovarian insufficiency by

inducing oocyte apoptosis. J Clin Invest *133*, e162315. https://doi.org/10.1172/JCI162315.

Huang, H., Woo, A.J., Waldon, Z., Schindler, Y., Moran, T.B., Zhu, H.H., Feng, G.-S., Steen, H., and Cantor, A.B. (2012). A Src family kinase–Shp2 axis controls RUNX1 activity in megakaryocyte and T-lymphocyte differentiation. Genes Dev. *26*, 1587–1601. https://doi.org/10.1101/gad.192054.112.

Huang, H., Svoboda, R.A., Lazenby, A.J., Saowapa, J., Chaika, N., Ding, K., Wheelock, M.J., and Johnson, K.R. (2016). Up-regulation of N-cadherin by Collagen I-activated Discoidin Domain Receptor 1 in Pancreatic Cancer Requires the Adaptor Molecule Shc1. J Biol Chem *291*, 23208–23223. https://doi.org/10.1074/jbc.M116.740605.

Hutson, S.W., Cowen, P.N., and Bird, C.C. (1985). Morphometric studies of age related changes in normal human breast and their significance for evolution of mammary cancer. J Clin Pathol *38*, 281–287. .

Ichikawa, M., Yoshimi, A., Nakagawa, M., Nishimoto, N., Watanabe-Okochi, N., and Kurokawa, M. (2013). A role for RUNX1 in hematopoiesis and myeloid leukemia. Int J Hematol *97*, 726–734. https://doi.org/10.1007/s12185-013-1347-3.

Illendula, A., Gilmour, J., Grembecka, J., Tirumala, V.S.S., Boulton, A., Kuntimaddi, A., Schmidt, C., Wang, L., Pulikkan, J.A., Zong, H., et al. (2016). Small Molecule Inhibitor of CBFβ-RUNX Binding for RUNX Transcription Factor Driven Cancers. EBioMedicine *8*, 117–131. https://doi.org/10.1016/j.ebiom.2016.04.032.

Ito, Y., Bae, S.-C., and Chuang, L.S.H. (2015). The RUNX family: developmental regulators in cancer. Nat Rev Cancer *15*, 81–95. https://doi.org/10.1038/nrc3877.

Janes, K.A. (2011). RUNX1 and its understudied role in breast cancer. Cell Cycle *10*, 3461–3465. https://doi.org/10.4161/cc.10.20.18029.

Jemc, J., and Rebay, I. (2007). Identification of transcriptional targets of the dual-function transcription factor/phosphatase eyes absent. Developmental Biology *310*, 416–429. https://doi.org/10.1016/j.ydbio.2007.07.024.

Jensen, C., and Teng, Y. (2020). Is It Time to Start Transitioning From 2D to 3D Cell Culture? Front. Mol. Biosci. *7*. https://doi.org/10.3389/fmolb.2020.00033.

Jeong, E.-M., Pereira, M., So, E.-Y., Wu, K.Q., Del Tatto, M., Wen, S., Dooner, M.S., Dubielecka, P.M., Reginato, A.M., Ventetuolo, C.E., et al. (2022). Targeting RUNX1 as a novel treatment modality for pulmonary arterial hypertension. Cardiovasc Res *118*, 3211–3224. https://doi.org/10.1093/cvr/cvac001.

Jordan, V.C., and Koerner, S. (1975). Tamoxifen (ICI 46,474) and the human carcinoma 8S oestrogen receptor. Eur J Cancer (1965) *11*, 205–206. https://doi.org/10.1016/0014-2964(75)90119-x.

Jorgez, C.J., Seth, A., Wilken, N., Bournat, J.C., Chen, C.H., and Lamb, D.J. (2021). E2F1 regulates testicular descent and controls spermatogenesis by influencing WNT4 signaling. Development *148*, dev191189. https://doi.org/10.1242/dev.191189.

Kapałczyńska, M., Kolenda, T., Przybyła, W., Zajączkowska, M., Teresiak, A., Filas, V., Ibbs, M., Bliźniak, R., Łuczewski, Ł., and Lamperska, K. (2018). 2D and 3D cell cultures – a comparison of different types of cancer cell cultures. Arch Med Sci *14*, 910–919. https://doi.org/10.5114/aoms.2016.63743.

Karukstis, K.K., Thompson, E.H.Z., Whiles, J.A., and Rosenfeld, R.J. (1998). Deciphering the fluorescence signature of daunomycin and doxorubicin. Biophysical Chemistry *73*, 249–263. https://doi.org/10.1016/S0301-4622(98)00150-1.

Keller, P.J., Arendt, L.M., Skibinski, A., Logvinenko, T., Klebba, I., Dong, S., Smith, A.E., Prat, A., Perou, C.M., Gilmore, H., et al. (2012). Defining the cellular precursors to human breast cancer. Proceedings of the National Academy of Sciences *109*, 2772–2777. https://doi.org/10.1073/pnas.1017626108.

Khawaled, S., and Aqeilan, R.I. (2017). RUNX1, a new regulator of EMT in breast cancer. Oncotarget *8*, 17407–17408. https://doi.org/10.18632/oncotarget.15623.

Kibbey, M.C. (1994). Maintenance of the EHS sarcoma and Matrigel preparation. Journal of Tissue Culture Methods *16*, 227–230. https://doi.org/10.1007/BF01540656.

Kilbey, A., Terry, A., Wotton, S., Borland, G., Zhang, Q., Mackay, N., McDonald, A., Bell, M., Wakelam, M.J.O., Cameron, E.R., et al. (2017). Runx1 Orchestrates Sphingolipid Metabolism and Glucocorticoid Resistance in Lymphomagenesis. J Cell Biochem *118*, 1432–1441. https://doi.org/10.1002/jcb.25802.

Kim, J., Koo, B.-K., and Knoblich, J.A. (2020). Human organoids: model systems for human biology and medicine. Nat Rev Mol Cell Biol *21*, 571–584. https://doi.org/10.1038/s41580-020-0259-3.

Kothari, C., Diorio, C., and Durocher, F. (2020). The Importance of Breast Adipose Tissue in Breast Cancer. Int J Mol Sci *21*, 5760. https://doi.org/10.3390/ijms21165760.

Krishnan, V. (2023). The RUNX Family of Proteins, DNA Repair, and Cancer. Cells *12*, 1106. https://doi.org/10.3390/cells12081106.

Krishnan, V., and Ito, Y. (2017). A Regulatory Role for RUNX1, RUNX3 in the Maintenance of Genomic Integrity. In RUNX Proteins in Development and Cancer, Y. Groner, Y. Ito, P. Liu, J.C. Neil, N.A. Speck, and A. van Wijnen, eds. (Singapore: Springer), pp. 491–510.

Kuperwasser, C., Chavarria, T., Wu, M., Magrane, G., Gray, J.W., Carey, L., Richardson, A., and Weinberg, R.A. (2004). Reconstruction of functionally normal and malignant human breast tissues in mice. Proceedings of the National Academy of Sciences *101*, 4966–4971. https://doi.org/10.1073/pnas.0401064101.

Lachmann, A., Xu, H., Krishnan, J., Berger, S.I., Mazloom, A.R., and Ma'ayan, A. (2010). ChEA: transcription factor regulation inferred from integrating genome-wide ChIP-X experiments. Bioinformatics *26*, 2438–2444. https://doi.org/10.1093/bioinformatics/btq466.

Lee, T.I., and Young, R.A. (2013). Transcriptional Regulation and its Misregulation in Disease. Cell *152*, 1237–1251. https://doi.org/10.1016/j.cell.2013.02.014.

Lee, E.Y., Parry, G., and Bissell, M.J. (1984). Modulation of secreted proteins of mouse mammary epithelial cells by the collagenous substrata. Journal of Cell Biology *98*, 146–155. https://doi.org/10.1083/jcb.98.1.146.

Leisten, I., Kramann, R., Ventura Ferreira, M.S., Bovi, M., Neuss, S., Ziegler, P., Wagner, W., Knüchel, R., and Schneider, R.K. (2012). 3D co-culture of hematopoietic stem and progenitor cells and mesenchymal stem cells in collagen scaffolds as a model of the hematopoietic niche. Biomaterials *33*, 1736–1747. https://doi.org/10.1016/j.biomaterials.2011.11.034.

Leitinger, B. (2014). Chapter Two - Discoidin Domain Receptor Functions in Physiological and Pathological Conditions. In International Review of Cell and Molecular Biology, K.W. Jeon, ed. (Academic Press), pp. 39–87.

Lemmon, M.A., and Schlessinger, J. (2010). Cell Signaling by Receptor Tyrosine Kinases. Cell *141*, 1117–1134. https://doi.org/10.1016/j.cell.2010.06.011.

Levanon, D., Bettoun, D., Harris-Cerruti, C., Woolf, E., Negreanu, V., Eilam, R., Bernstein, Y., Goldenberg, D., Xiao, C., Fliegauf, M., et al. (2002). The Runx3 transcription factor regulates development and survival of TrkC dorsal root ganglia neurons. EMBO J *21*, 3454–3463. https://doi.org/10.1093/emboj/cdf370.

Li, M.L., Aggeler, J., Farson, D.A., Hatier, C., Hassell, J., and Bissell, M.J. (1987). Influence of a reconstituted basement membrane and its components on casein gene expression and secretion in mouse mammary epithelial cells. Proceedings of the National Academy of Sciences *84*, 136–140. https://doi.org/10.1073/pnas.84.1.136.

Lin, T.-C. (2022). RUNX1 and cancer. Biochimica et Biophysica Acta (BBA) - Reviews on Cancer *1877*, 188715. https://doi.org/10.1016/j.bbcan.2022.188715.

Linnemann, J.R., Miura, H., Meixner, L.K., Irmler, M., Kloos, U.J., Hirschi, B., Bartsch, H.S., Sass, S., Beckers, J., Theis, F.J., et al. (2015). Quantification of regenerative potential in primary human mammary epithelial cells. Development *142*, 3239–3251. https://doi.org/10.1242/dev.123554.

Linnemann, J.R., Meixner, L.K., Miura, H., and Scheel, C.H. (2017). An Organotypic 3D Assay for Primary Human Mammary Epithelial Cells that Recapitulates Branching Morphogenesis. In 3D Cell Culture: Methods and Protocols, Z. Koledova, ed. (New York, NY: Springer), pp. 125–137.

Liu, P., Liu, J.-P., Sun, S.-J., Gao, Y., Ai, Y., Chen, X., Sun, Y., Zhou, M., Liu, Y., Xiong, Y., et al. (2021). CBFB-MYH11 Fusion Sequesters RUNX1 in Cytoplasm to Prevent DNMT3A Recruitment to Target Genes in AML. Front Cell Dev Biol *9*, 675424. https://doi.org/10.3389/fcell.2021.675424.

Liu, Y., Rutlin, M., Huang, S., Barrick, C.A., Wang, F., Jones, K.R., Tessarollo, L., and Ginty, D.D. (2012). Sexually Dimorphic BDNF Signaling Directs Sensory Innervation of the Mammary Gland. Science *338*, 1357–1360. https://doi.org/10.1126/science.1228258.

Liu, Y., Huang, C., Bai, M., Pi, C., Zhang, D., and Xie, J. (2022). The roles of Runx1 in skeletal development and osteoarthritis: A concise review. Heliyon *8*, e12656. https://doi.org/10.1016/j.heliyon.2022.e12656.

Lu, C., Yang, Z., Yu, D., Lin, J., and Cai, W. (2020). RUNX1 regulates TGF-β induced migration and EMT in colorectal cancer. Pathol Res Pract *216*, 153142. https://doi.org/10.1016/j.prp.2020.153142.

Lu, K.K., Trcka, D., and Bendeck, M.P. (2011). Collagen stimulates discoidin domain receptor 1-mediated migration of smooth muscle cells through Src. Cardiovasc Pathol *20*, 71–76. https://doi.org/10.1016/j.carpath.2009.12.006.

Luo, W., Li, S., Peng, B., Ye, Y., Deng, X., and Yao, K. (2013). Embryonic Stem Cells Markers SOX2, OCT4 and Nanog Expression and Their Correlations with Epithelial-Mesenchymal Transition in Nasopharyngeal Carcinoma. PLOS ONE *8*, e56324. https://doi.org/10.1371/journal.pone.0056324.

Malik, N., Yan, H., Moshkovich, N., Palangat, M., Yang, H., Sanchez, V., Cai, Z., Peat, T.J., Jiang, S., Liu, C., et al. (2019). The transcription factor CBFB suppresses breast cancer through orchestrating translation and transcription. Nat Commun *10*, 2071. https://doi.org/10.1038/s41467-019-10102-6.

Manabe Ichiro, Shindo Takayuki, and Nagai Ryozo (2002). Gene Expression in Fibroblasts and Fibrosis. Circulation Research *91*, 1103–1113. https://doi.org/10.1161/01.RES.0000046452.67724.B8.

Martinez, M., Hinojosa, M., Trombly, D., Morin, V., Stein, J., Stein, G., Javed, A., and Gutierrez, S.E. (2016). Transcriptional Auto-Regulation of RUNX1 P1 Promoter. PLoS One *11*, e0149119. https://doi.org/10.1371/journal.pone.0149119.

Masuda, H., Zhang, D., Bartholomeusz, C., Doihara, H., Hortobagyi, G.N., and Ueno, N.T. (2012). Role of Epidermal Growth Factor Receptor in Breast Cancer. Breast Cancer Res Treat *136*, 10.1007/s10549-012-2289–9. https://doi.org/10.1007/s10549-012-2289-9.

Mercado-Matos, J., Matthew-Onabanjo, A.N., and Shaw, L.M. (2017). RUNX1 and breast cancer. Oncotarget *8*, 36934–36935. https://doi.org/10.18632/oncotarget.17249.

Meyer zum Gottesberge, A.M., Gross, O., Becker-Lendzian, U., Massing, T., and Vogel, W.F. (2008). Inner ear defects and hearing loss in mice lacking the collagen receptor DDR1. Laboratory Investigation *88*, 27–37. https://doi.org/10.1038/labinvest.3700692.

Miki, Y., Swensen, J., Shattuck-Eidens, D., Futreal, P.A., Harshman, K., Tavtigian, S., Liu, Q., Cochran, C., Bennett, L.M., Ding, W., et al. (1994). A Strong Candidate for the Breast and Ovarian Cancer Susceptibility Gene BRCA1. Science *266*, 66–71. https://doi.org/10.1126/science.7545954.

Mohan, S.C., Lee, T.-Y., Giuliano, A.E., and Cui, X. (2021). Current Status of Breast Organoid Models. Front Bioeng Biotechnol *9*, 745943. https://doi.org/10.3389/fbioe.2021.745943.

Moll, S., Desmoulière, A., Moeller, M.J., Pache, J.-C., Badi, L., Arcadu, F., Richter, H., Satz, A., Uhles, S., Cavalli, A., et al. (2019). DDR1 role in fibrosis and its pharmacological targeting. Biochimica et Biophysica Acta (BBA) - Molecular Cell Research *1866*, 118474. https://doi.org/10.1016/j.bbamcr.2019.04.004.

Mootha, V.K., Lindgren, C.M., Eriksson, K.-F., Subramanian, A., Sihag, S., Lehar, J., Puigserver, P., Carlsson, E., Ridderstråle, M., Laurila, E., et al. (2003). PGC-1α-responsive genes involved in oxidative phosphorylation are coordinately downregulated in human diabetes. Nat Genet *34*, 267–273. https://doi.org/10.1038/ng1180.

de Morree, A., and Rando, T.A. (2023). Regulation of adult stem cell quiescence and its functions in the maintenance of tissue integrity. Nat Rev Mol Cell Biol *24*, 334–354. https://doi.org/10.1038/s41580-022-00568-6.

Mukherjee, P., Roy, S., Ghosh, D., and Nandi, S.K. (2022). Role of animal models in biomedical research: a review. Laboratory Animal Research *38*, 18. https://doi.org/10.1186/s42826-022-00128-1.

Murashima, A., Kishigami, S., Thomson, A., and Yamada, G. (2015). Androgens and mammalian male reproductive tract development. Biochimica et Biophysica Acta

(BBA) - Gene Regulatory Mechanisms *1849*, 163–170. https://doi.org/10.1016/j.bbagrm.2014.05.020.
Muschler, J., and Streuli, C.H. (2010). Cell–Matrix Interactions in Mammary Gland Development and Breast Cancer. Cold Spring Harb Perspect Biol *2*, a003202. https://doi.org/10.1101/cshperspect.a003202.
Naomi, R., Ridzuan, P.M., and Bahari, H. (2021). Current Insights into Collagen Type I. Polymers (Basel) *13*, 2642. https://doi.org/10.3390/polym13162642.
Narayan, S., Bryant, G., Shah, S., Berrozpe, G., and Ptashne, M. (2017). OCT4 and SOX2 Work as Transcriptional Activators in Reprogramming Human Fibroblasts. Cell Rep *20*, 1585–1596. https://doi.org/10.1016/j.celrep.2017.07.071.
Nishitani, H., Taraviras, S., Lygerou, Z., and Nishimoto, T. (2001). The human licensing factor for DNA replication Cdt1 accumulates in G1 and is destabilized after initiation of S-phase. J Biol Chem *276*, 44905–44911. https://doi.org/10.1074/jbc.M105406200.
Noh, K.H., Kim, B.W., Song, K.-H., Cho, H., Lee, Y.-H., Kim, J.H., Chung, J.-Y., Kim, J.-H., Hewitt, S.M., Seong, S.-Y., et al. (2012). Nanog signaling in cancer promotes stem-like phenotype and immune evasion. J Clin Invest *122*, 4077–4093. https://doi.org/10.1172/JCI64057.
North, T.E., de Bruijn, M.F.T.R., Stacy, T., Talebian, L., Lind, E., Robin, C., Binder, M., Dzierzak, E., and Speck, N.A. (2002). Runx1 expression marks long-term repopulating hematopoietic stem cells in the midgestation mouse embryo. Immunity *16*, 661–672. https://doi.org/10.1016/s1074-7613(02)00296-0.
Oftedal, O.T. (2002). The Mammary Gland and Its Origin During Synapsid Evolution. J Mammary Gland Biol Neoplasia *7*, 225–252. https://doi.org/10.1023/A:1022896515287.
Oftedal, O.T., and Dhouailly, D. (2013). Evo-Devo of the Mammary Gland. J Mammary Gland Biol Neoplasia *18*, 105–120. https://doi.org/10.1007/s10911-013-9290-8.
Okuda, T., Nishimura, M., Nakao, M., and Fujitaa, Y. (2001). RUNX1/AML1: A Central Player in Hematopoiesis. Int J Hematol *74*, 252–257. https://doi.org/10.1007/BF02982057.
Otálora-Otálora, B.A., Henríquez, B., López-Kleine, L., and Rojas, A. (2019). RUNX family: Oncogenes or tumor suppressors. Oncol Rep *42*, 3–19. https://doi.org/10.3892/or.2019.7149.
Pasic, L., Eisinger-Mathason, T.S.K., Velayudhan, B.T., Moskaluk, C.A., Brenin, D.R., Macara, I.G., and Lannigan, D.A. (2011). Sustained activation of the HER1–ERK1/2–RSK signaling pathway controls myoepithelial cell fate in human mammary tissue. Genes Dev. *25*, 1641–1653. https://doi.org/10.1101/gad.2025611.
Pastor, S., Wicinski, J., Charafe-Jauffret, E., Verhoeyen, E., Guittard, G., and Ginestier, C. (2023). Production of CRISPRi-engineered primary human mammary epithelial cells with baboon envelope pseudotyped lentiviral vectors. STAR Protocols *4*, 102055. https://doi.org/10.1016/j.xpro.2023.102055.
Patel, V.P., and Lodish, H.F. (1987). A fibronectin matrix is required for differentiation of murine erythroleukemia cells into reticulocytes. Journal of Cell Biology *105*, 3105–3118. https://doi.org/10.1083/jcb.105.6.3105.

Pearson, G.W. (2019). Control of Invasion by Epithelial-to-Mesenchymal Transition Programs during Metastasis. Journal of Clinical Medicine *8*, 646. https://doi.org/10.3390/jcm8050646.

Petersen, O.W., and Polyak, K. (2010). Stem Cells in the Human Breast. Cold Spring Harb Perspect Biol *2*, a003160. https://doi.org/10.1101/cshperspect.a003160.

Portal, C., Wang, Z., Scott, D.K., Wolosin, J.M., and Iomini, C. (2022). The c-Myc Oncogene Maintains Corneal Epithelial Architecture at Homeostasis, Modulates p63 Expression, and Enhances Proliferation During Tissue Repair. Invest Ophthalmol Vis Sci *63*, 3. https://doi.org/10.1167/iovs.63.2.3.

Qin, X., Jiang, Q., Matsuo, Y., Kawane, T., Komori, H., Moriishi, T., Taniuchi, I., Ito, K., Kawai, Y., Rokutanda, S., et al. (2015). Cbfb regulates bone development by stabilizing Runx family proteins. J Bone Miner Res *30*, 706–714. https://doi.org/10.1002/jbmr.2379.

Ramaswamy, S., Ross, K.N., Lander, E.S., and Golub, T.R. (2003). A molecular signature of metastasis in primary solid tumors. Nat Genet *33*, 49–54. https://doi.org/10.1038/ng1060.

Rauner, G., Jin, D.X., Miller, D.H., Gierahn, T.M., Li, C.M., Sokol, E.S., Feng, Y.-X., Mathis, R.A., Love, J.C., Gupta, P.B., et al. (2021). Breast tissue regeneration is driven by cell-matrix interactions coordinating multi-lineage stem cell differentiation through DDR1. Nat Commun *12*, 7116. https://doi.org/10.1038/s41467-021-27401-6.

Rauner, G., Traugh, N.C., Trepicchio, C.J., Parrish, M.E., Mushayandebvu, K., and Kuperwasser, C. (2023). Advancements in Human Breast Organoid Culture: Modeling Complex Tissue Structures and Developmental Insights. 2023.10.02.560364. https://doi.org/10.1101/2023.10.02.560364.

Reed, J.R., and Schwertfeger, K.L. (2010). Immune Cell Location and Function During Post-Natal Mammary Gland Development. J Mammary Gland Biol Neoplasia *15*, 329–339. https://doi.org/10.1007/s10911-010-9188-7.

Reversi, A., Cassoni, P., and Chini, B. (2005). Oxytocin receptor signaling in myoepithelial and cancer cells. J Mammary Gland Biol Neoplasia *10*, 221–229. https://doi.org/10.1007/s10911-005-9583-7.

Reya, T., Morrison, S.J., Clarke, M.F., and Weissman, I.L. (2001). Stem cells, cancer, and cancer stem cells. Nature *414*, 105–111. https://doi.org/10.1038/35102167.

Rios, A.C., Fu, N.Y., Lindeman, G.J., and Visvader, J.E. (2014). In situ identification of bipotent stem cells in the mammary gland. Nature *506*, 322–327. https://doi.org/10.1038/nature12948.

Romayor, I., Badiola, I., Benedicto, A., Márquez, J., Herrero, A., Arteta, B., and Olaso, E. (2020). Silencing of sinusoidal DDR1 reduces murine liver metastasis by colon carcinoma. Sci Rep *10*, 18398. https://doi.org/10.1038/s41598-020-75395-w.

Rose, J.T., Moskovitz, E., Boyd, J.R., Gordon, J.A., Bouffard, N.A., Fritz, A.J., Illendula, A., Bushweller, J.H., Lian, J.B., Stein, J.L., et al. (2020). Inhibition of the RUNX1-CBFβ transcription factor complex compromises mammary epithelial cell identity: a phenotype potentially stabilized by mitotic gene bookmarking. Oncotarget *11*, 2512–2530. https://doi.org/10.18632/oncotarget.27637.

Rosenbluth, J.M., Schackmann, R.C.J., Gray, G.K., Selfors, L.M., Li, C.M.-C., Boedicker, M., Kuiken, H.J., Richardson, A., Brock, J., Garber, J., et al. (2020).

Organoid cultures from normal and cancer-prone human breast tissues preserve complex epithelial lineages. Nat Commun *11*, 1711. https://doi.org/10.1038/s41467-020-15548-7.

Roskelley, C.D., Desprez, P.Y., and Bissell, M.J. (1994). Extracellular matrix-dependent tissue-specific gene expression in mammary epithelial cells requires both physical and biochemical signal transduction. Proc Natl Acad Sci U S A *91*, 12378–12382. .

Russo, J., and Russo, I.H. (2004). Development of the human breast. Maturitas *49*, 2–15. https://doi.org/10.1016/j.maturitas.2004.04.011.

Samarakkody, A.S., Shin, N.-Y., and Cantor, A.B. (2020). Role of RUNX Family Transcription Factors in DNA Damage Response. Mol Cells *43*, 99–106. https://doi.org/10.14348/molcells.2019.0304.

Sarrio, D., Franklin, C.K., Mackay, A., Reis-Filho, J.S., and Isacke, C.M. (2012). Epithelial and mesenchymal subpopulations within normal basal breast cell lines exhibit distinct stem cell/progenitor properties. Stem Cells *30*, 292–303. https://doi.org/10.1002/stem.791.

Sato, T., Vries, R.G., Snippert, H.J., van de Wetering, M., Barker, N., Stange, D.E., van Es, J.H., Abo, A., Kujala, P., Peters, P.J., et al. (2009). Single Lgr5 stem cells build crypt-villus structures in vitro without a mesenchymal niche. Nature *459*, 262–265. https://doi.org/10.1038/nature07935.

Saultz, J.N., and Garzon, R. (2016). Acute Myeloid Leukemia: A Concise Review. Journal of Clinical Medicine *5*, 33. https://doi.org/10.3390/jcm5030033.

Schaefer, T., and Lengerke, C. (2020). SOX2 protein biochemistry in stemness, reprogramming, and cancer: the PI3K/AKT/SOX2 axis and beyond. Oncogene *39*, 278–292. https://doi.org/10.1038/s41388-019-0997-x.

Schlessinger, J. (2000). Cell Signaling by Receptor Tyrosine Kinases. Cell *103*, 211–225. https://doi.org/10.1016/S0092-8674(00)00114-8.

Schofield, R. (1978). The relationship between the spleen colony-forming cell and the haemopoietic stem cell. Blood Cells *4*, 7–25. .

Shackleton, M., Vaillant, F., Simpson, K.J., Stingl, J., Smyth, G.K., Asselin-Labat, M.-L., Wu, L., Lindeman, G.J., and Visvader, J.E. (2006). Generation of a functional mammary gland from a single stem cell. Nature *439*, 84–88. https://doi.org/10.1038/nature04372.

Sheikh-Hosseini, M., Larijani, B., Gholipoor Kakroodi, Z., Shokoohi, M., Moarefzadeh, M., Sayahpour, F.A., Goodarzi, P., and Arjmand, B. (2021). Gene Therapy as an Emerging Therapeutic Approach to Breast Cancer: New Developments and Challenges. Human Gene Therapy *32*, 1330–1345. https://doi.org/10.1089/hum.2020.199.

Shen, L., Shi, Q., and Wang, W. (2018). Double agents: genes with both oncogenic and tumor-suppressor functions. Oncogenesis *7*, 1–14. https://doi.org/10.1038/s41389-018-0034-x.

Shi, G., and Jin, Y. (2010). Role of Oct4 in maintaining and regaining stem cell pluripotency. Stem Cell Research & Therapy *1*, 39. https://doi.org/10.1186/scrt39.

Shintani, Y., Fukumoto, Y., Chaika, N., Svoboda, R., Wheelock, M.J., and Johnson, K.R. (2008). Collagen I–mediated up-regulation of N-cadherin requires cooperative

signals from integrins and discoidin domain receptor 1. J Cell Biol *180*, 1277–1289. https://doi.org/10.1083/jcb.200708137.

Shrivastava, A., Radziejewski, C., Campbell, E., Kovac, L., McGlynn, M., Ryan, T.E., Davis, S., Goldfarb, M.P., Glass, D.J., Lemke, G., et al. (1997). An Orphan Receptor Tyrosine Kinase Family Whose Members Serve as Nonintegrin Collagen Receptors. Molecular Cell *1*, 25–34. https://doi.org/10.1016/S1097-2765(00)80004-0.

Sokol, E.S., Sanduja, S., Jin, D.X., Miller, D.H., Mathis, R.A., and Gupta, P.B. (2015). Perturbation-Expression Analysis Identifies RUNX1 as a Regulator of Human Mammary Stem Cell Differentiation. PLOS Computational Biology *11*, e1004161. https://doi.org/10.1371/journal.pcbi.1004161.

Sokol, E.S., Miller, D.H., Breggia, A., Spencer, K.C., Arendt, L.M., and Gupta, P.B. (2016). Growth of human breast tissues from patient cells in 3D hydrogel scaffolds. Breast Cancer Res *18*, 19. https://doi.org/10.1186/s13058-016-0677-5.

Song, S., Shackel, N.A., Wang, X.M., Ajami, K., McCaughan, G.W., and Gorrell, M.D. (2011). Discoidin Domain Receptor 1. Am J Pathol *178*, 1134–1144. https://doi.org/10.1016/j.ajpath.2010.11.068.

Spitz, F., and Furlong, E.E.M. (2012). Transcription factors: from enhancer binding to developmental control. Nat Rev Genet *13*, 613–626. https://doi.org/10.1038/nrg3207.

Stender, J.D., Kim, K., Charn, T.H., Komm, B., Chang, K.C.N., Kraus, W.L., Benner, C., Glass, C.K., and Katzenellenbogen, B.S. (2010). Genome-Wide Analysis of Estrogen Receptor α DNA Binding and Tethering Mechanisms Identifies Runx1 as a Novel Tethering Factor in Receptor-Mediated Transcriptional Activation. Mol Cell Biol *30*, 3943–3955. https://doi.org/10.1128/MCB.00118-10.

Stuart, T., Butler, A., Hoffman, P., Hafemeister, C., Papalexi, E., Mauck, W.M., Hao, Y., Stoeckius, M., Smibert, P., and Satija, R. (2019). Comprehensive Integration of Single-Cell Data. Cell *177*, 1888-1902.e21. https://doi.org/10.1016/j.cell.2019.05.031.

Subramanian, A., Tamayo, P., Mootha, V.K., Mukherjee, S., Ebert, B.L., Gillette, M.A., Paulovich, A., Pomeroy, S.L., Golub, T.R., Lander, E.S., et al. (2005). Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. Proceedings of the National Academy of Sciences *102*, 15545–15550. https://doi.org/10.1073/pnas.0506580102.

Sugimoto, S., and Sato, T. (2021). Organoid vs In Vivo Mouse Model: Which is Better Research Tool to Understand the Biologic Mechanisms of Intestinal Epithelium? Cell Mol Gastroenterol Hepatol *13*, 195–197. https://doi.org/10.1016/j.jcmgh.2021.06.027.

Sun, X., Gupta, K., Wu, B., Zhang, D., Yuan, B., Zhang, X., Chiang, H.-C., Zhang, C., Curiel, T.J., Bendeck, M.P., et al. (2018). Tumor-extrinsic discoidin domain receptor 1 promotes mammary tumor growth by regulating adipose stromal interleukin 6 production in mice. J Biol Chem *293*, 2841–2849. https://doi.org/10.1074/jbc.RA117.000672.

Sun, X., Wu, B., Chiang, H.-C., Deng, H., Zhang, X., Xiong, W., Liu, J., Rozeboom, A.M., Harris, B.T., Blommaert, E., et al. (2021). Tumour DDR1 promotes

collagen fibre alignment to instigate immune exclusion. Nature *599*, 673–678. https://doi.org/10.1038/s41586-021-04057-2.

Szklarczyk, D., Kirsch, R., Koutrouli, M., Nastou, K., Mehryary, F., Hachilif, R., Gable, A.L., Fang, T., Doncheva, N.T., Pyysalo, S., et al. (2022). The STRING database in 2023: protein–protein association networks and functional enrichment analyses for any sequenced genome of interest. Nucleic Acids Res *51*, D638–D646. https://doi.org/10.1093/nar/gkac1000.

Tahirov, T.H., Inoue-Bungo, T., Morii, H., Fujikawa, A., Sasaki, M., Kimura, K., Shiina, M., Sato, K., Kumasaka, T., Yamamoto, M., et al. (2001). Structural Analyses of DNA Recognition by the AML1/Runx-1 Runt Domain and Its Allosteric Control by CBFβ. Cell *104*, 755–767. https://doi.org/10.1016/S0092-8674(01)00271-9.

Takahashi, K., and Yamanaka, S. (2006). Induction of Pluripotent Stem Cells from Mouse Embryonic and Adult Fibroblast Cultures by Defined Factors. Cell *126*, 663–676. https://doi.org/10.1016/j.cell.2006.07.024.

Takahashi, K., Tanabe, K., Ohnuki, M., Narita, M., Ichisaka, T., Tomoda, K., and Yamanaka, S. (2007). Induction of pluripotent stem cells from adult human fibroblasts by defined factors. Cell *131*, 861–872. https://doi.org/10.1016/j.cell.2007.11.019.

Takai, K., Drain, A.P., Lawson, D.A., Littlepage, L.E., Karpuj, M., Kessenbrock, K., Le, A., Inoue, K., Weaver, V.M., and Werb, Z. (2018). Discoidin domain receptor 1 (DDR1) ablation promotes tissue fibrosis and hypoxia to induce aggressive basal-like breast cancers. Genes Dev *32*, 244–257. https://doi.org/10.1101/gad.301366.117.

Tang, X., Sun, L., Wang, G., Chen, B., and Luo, F. (2018). RUNX1: A Regulator of NF-κB Signaling in Pulmonary Diseases. Curr Protein Pept Sci *19*, 172–178. https://doi.org/10.2174/1389203718666171009111835.

Tsai, C., and Gergen, J.P. (1994). Gap gene properties of the pair-rule gene runt during Drosophila segmentation. Development *120*, 1671–1683. https://doi.org/10.1242/dev.120.6.1671.

Uchida, H., Zhang, J., and Nimer, S.D. (1997). AML1A and AML1B can transactivate the human IL-3 promoter. J Immunol *158*, 2251–2258. .

Ugai, T., Sasamoto, N., Lee, H.-Y., Ando, M., Song, M., Tamimi, R.M., Kawachi, I., Campbell, P.T., Giovannucci, E.L., Weiderpass, E., et al. (2022). Is early-onset cancer an emerging global epidemic? Current evidence and future implications. Nat Rev Clin Oncol *19*, 656–673. https://doi.org/10.1038/s41571-022-00672-8.

Usui, T., Aoyagi, K., Saeki, N., Nakanishi, Y., Kanai, Y., Ohki, M., Ogawa, K., Yoshida, T., and Sasaki, H. (2006). Expression status of RUNX1/AML1 in normal gastric epithelium and its mutational analysis in microdissected gastric cancer cells. International Journal of Oncology *29*, 779–784. https://doi.org/10.3892/ijo.29.4.779.

Valiathan, R.R., Marco, M., Leitinger, B., Kleer, C.G., and Fridman, R. (2012). Discoidin domain receptor tyrosine kinases: new players in cancer progression. Cancer Metastasis Rev *31*, 295–321. https://doi.org/10.1007/s10555-012-9346-z.

Van Keymeulen, A., Rocha, A.S., Ousset, M., Beck, B., Bouvencourt, G., Rock, J., Sharma, N., Dekoninck, S., and Blanpain, C. (2011). Distinct stem cells

contribute to mammary gland development and maintenance. Nature *479*, 189–193. https://doi.org/10.1038/nature10573.

Varner, V.D., and Nelson, C.M. (2014). Cellular and physical mechanisms of branching morphogenesis. Development *141*, 2750–2759. https://doi.org/10.1242/dev.104794.

Veltmaat, J.M., Relaix, F., Le, L.T., Kratochwil, K., Sala, F.G., van Veelen, W., Rice, R., Spencer-Dene, B., Mailleux, A.A., Rice, D.P., et al. (2006). Gli3-mediated somitic Fgf10 expression gradients are required for the induction and patterning of mammary epithelium along the embryonic axes. Development *133*, 2325–2335. https://doi.org/10.1242/dev.02394.

Vilella, E., Gas, C., Garcia-Ruiz, B., and Rivera, F.J. (2019). Expression of DDR1 in the CNS and in myelinating oligodendrocytes. Biochimica et Biophysica Acta (BBA) - Molecular Cell Research *1866*, 118483. https://doi.org/10.1016/j.bbamcr.2019.04.010.

Visvader, J.E. (2009). Keeping abreast of the mammary epithelial hierarchy and breast tumorigenesis. Genes Dev *23*, 2563–2577. https://doi.org/10.1101/gad.1849509.

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

Vogel, W., Gish, G.D., Alves, F., and Pawson, T. (1997). The Discoidin Domain Receptor Tyrosine Kinases Are Activated by Collagen. Molecular Cell *1*, 13–23. https://doi.org/10.1016/S1097-2765(00)80003-9.

Vogel, W.F., Aszódi, A., Alves, F., and Pawson, T. (2001). Discoidin Domain Receptor 1 Tyrosine Kinase Has an Essential Role in Mammary Gland Development. Molecular and Cellular Biology *21*, 2906–2917. https://doi.org/10.1128/MCB.21.8.2906-2917.2001.

Wagner, J.P., Wolf-Yadlin, A., Sevecka, M., Grenier, J.K., Root, D.E., Lauffenburger, D.A., and MacBeath, G. (2013). Receptor Tyrosine Kinases Fall into Distinct Classes Based on Their Inferred Signaling Networks. Science Signaling *6*, ra58–ra58. https://doi.org/10.1126/scisignal.2003994.

Walrad, P.B., Hang, S., and Gergen, J.P. (2011). Hairless is a cofactor for Runt-dependent transcriptional regulation. Mol Biol Cell *22*, 1364–1374. https://doi.org/10.1091/mbc.E10-06-0483.

Wang, B., Guo, H., Yu, H., Chen, Y., Xu, H., and Zhao, G. (2021). The Role of the Transcription Factor EGR1 in Cancer. Frontiers in Oncology *11*. .

Wang, S., Sekiguchi, R., Daley, W.P., and Yamada, K.M. (2017). Patterned cell and matrix dynamics in branching morphogenesis. Journal of Cell Biology *216*, 559–570. https://doi.org/10.1083/jcb.201610048.

Wantoch von Rekowski, K., König, P., Henze, S., Schlesinger, M., Zawierucha, P., Januchowski, R., and Bendas, G. (2020). Insight into Cisplatin-Resistance Signaling of W1 Ovarian Cancer Cells Emerges mTOR and HSP27 as Targets for Sensitization Strategies. International Journal of Molecular Sciences *21*, 9240. https://doi.org/10.3390/ijms21239240.

Wellings, S.R., Jensen, H.M., and Marcum, R.G. (1975). An atlas of subgross pathology of the human breast with special reference to possible precancerous lesions. J Natl Cancer Inst *55*, 231–273. .

Williams, J.M., and Daniel, C.W. (1983). Mammary ductal elongation: Differentiation of myoepithelium and basal lamina during branching morphogenesis. Developmental Biology *97*, 274–290. https://doi.org/10.1016/0012-1606(83)90086-6.

Wohlschlegel, J.A., Kutok, J.L., Weng, A.P., and Dutta, A. (2002). Expression of Geminin as a Marker of Cell Proliferation in Normal Tissues and Malignancies. Am J Pathol *161*, 267–273. .

Wu, J., Hu, S., Chen, Y., Li, Z., Zhang, J., Yuan, H., Shi, Q., Shao, N., and Ying, X. (2017). BCIP: a gene-centered platform for identifying potential regulatory genes in breast cancer. Sci Rep *7*, 45235. https://doi.org/10.1038/srep45235.

Wu, R., Sun, S., and Steinberg, B.M. (2003). Requirement of STAT3 Activation for Differentiation of Mucosal Stratified Squamous Epithelium. Mol Med *9*, 77–84. .

Xie, Z., Bailey, A., Kuleshov, M.V., Clarke, D.J.B., Evangelista, J.E., Jenkins, S.L., Lachmann, A., Wojciechowicz, M.L., Kropiwnicki, E., Jagodnik, K.M., et al. (2021). Gene Set Knowledge Discovery with Enrichr. Current Protocols *1*, e90. https://doi.org/10.1002/cpz1.90.

Xu, H., Raynal, N., Stathopoulos, S., Myllyharju, J., Farndale, R.W., and Leitinger, B. (2011). Collagen binding specificity of the discoidin domain receptors: Binding sites on collagens II and III and molecular determinants for collagen IV recognition by DDR1. Matrix Biology *30*, 16–26. https://doi.org/10.1016/j.matbio.2010.10.004.

Xu, J., Li, Z., Hou, Y., and Fang, W. (2015). Potential mechanisms underlying the Runx2 induced osteogenesis of bone marrow mesenchymal stem cells. Am J Transl Res *7*, 2527–2535. .

Xu, X., Brodie, S.G., Yang, X., Im, Y.-H., Parks, W.T., Chen, L., Zhou, Y.-X., Weinstein, M., Kim, S.-J., and Deng, C.-X. (2000). Haploid loss of the tumor suppressor Smad4/Dpc4 initiates gastric polyposis and cancer in mice. Oncogene *19*, 1868–1874. https://doi.org/10.1038/sj.onc.1203504.

Xu, Y., Wu, F., Qin, C., and Lin, Y. (2023). Paradoxical role of phosphorylated STAT3 in normal fertility and the pathogenesis of adenomyosis and endometriosis†. Biology of Reproduction ioad148. https://doi.org/10.1093/biolre/ioad148.

Yang, X.-H., and Sladek, T.L. (2018). Overexpression of the E2F-1 Transcription Factor Gene Mediates Cell Transformation. Gene Expr *4*, 195–204. .

Yang, J., Balakrishnan, A., Hamamoto, S., Elias, J.J., Rosenau, W., Beattie, C.W., Das Gupta, T.K., Wellings, S.R., and Nandi, S. (1987). Human breast epithelial cells in serum-free collagen gel primary culture: Growth, morphological, and immunocytochemical analysis. Journal of Cellular Physiology *133*, 228–234. https://doi.org/10.1002/jcp.1041330205.

Yeung, J., Mermet, J., Jouffe, C., Marquis, J., Charpagne, A., Gachon, F., and Naef, F. (2018). Transcription factor activity rhythms and tissue-specific chromatin interactions explain circadian gene expression across organs. Genome Res. *28*, 182–191. https://doi.org/10.1101/gr.222430.117.

Young, R.A. (2011). Control of the Embryonic Stem Cell State. Cell *144*, 940–954. https://doi.org/10.1016/j.cell.2011.01.032.

Yu, J., Vodyanik, M.A., Smuga-Otto, K., Antosiewicz-Bourget, J., Frane, J.L., Tian, S., Nie, J., Jonsdottir, G.A., Ruotti, V., Stewart, R., et al. (2007). Induced pluripotent

stem cell lines derived from human somatic cells. Science *318*, 1917–1920. https://doi.org/10.1126/science.1151526.

Zanetti, C., and Krause, D.S. (2020). "Caught in the net": the extracellular matrix of the bone marrow in normal hematopoiesis and leukemia. Experimental Hematology *89*, 13–25. https://doi.org/10.1016/j.exphem.2020.07.010.

Zhang, S., and Cui, W. (2014). Sox2, a key factor in the regulation of pluripotency and neural differentiation. World J Stem Cells *6*, 305–311. https://doi.org/10.4252/wjsc.v6.i3.305.

Zhang, X., Powell, K., and Li, L. (2020). Breast Cancer Stem Cells: Biomarkers, Identification and Isolation Methods, Regulating Mechanisms, Cellular Origin, and Beyond. Cancers (Basel) *12*, 3765. https://doi.org/10.3390/cancers12123765.

Zhao, Z., Chen, X., Dowbaj, A.M., Sljukic, A., Bratlie, K., Lin, L., Fong, E.L.S., Balachander, G.M., Chen, Z., Soragni, A., et al. (2022). Organoids. Nat Rev Methods Primers *2*, 1–21. https://doi.org/10.1038/s43586-022-00174-y.

Zhen, T., Cao, Y., Ren, G., Zhao, L., Hyde, R.K., Lopez, G., Feng, D., Alemu, L., Zhao, K., and Liu, P.P. (2020). RUNX1 and CBFβ-SMMHC transactivate target genes together in abnormal myeloid progenitors for leukemia development. Blood *136*, 2373–2385. https://doi.org/10.1182/blood.2020007747.

Zheng, R., and Blobel, G.A. (2010). GATA Transcription Factors and Cancer. Genes & Cancer *1*, 1178–1188. https://doi.org/10.1177/1947601911404223.

Zhong, X., Zhang, W., and Sun, T. (2019). DDR1 promotes breast tumor growth by suppressing antitumor immunity. Oncol Rep *42*, 2844–2854. https://doi.org/10.3892/or.2019.7338.

Zielke, N., and Edgar, B.A. (2015). FUCCI sensors: powerful new tools for analysis of cell proliferation. WIREs Developmental Biology *4*, 469–487. https://doi.org/10.1002/wdev.189.