INVESTIGATING THE EFFECTS OF LOSS OF RUNX1 AND ITS ROLE IN MITOTIC GENE BOOKMARKING AND BREAST CANCER

A Thesis Presented

by

Priyanka Chavarkar

to

The Faculty of the Graduate College

of

The University of Vermont

In Partial Fulfillment of the Requirements
for the Degree of Master of Science
Specializing in Pharmacology

January, 2024

Defense Date: 15 November 2023
Thesis Examination Committee:

Gary Stein, PhD, Advisor
Janet Stein, PhD, Advisor
Jason Stumpff, Ph.D., Chairperson
Frances Carr, Ph.D.
Jonathan Gordan, Ph.D.
Holger Hoock, DPhil, Dean of the Graduate College

ABSTRACT

While the transformation of cancer cells through epigenetic modifications has been recognized for some time, recent advances in genome-wide genomic and epigenomic sequencing has revealed the extensive presence of mutations in epigenetic regulators and the comprehensive scope of epigenomic changes in cancer cells. It is now evident that genetic and epigenetic mechanisms are not mutually exclusive and work in tandem to facilitate the acquisition of the characteristic traits of cancer.

Mitosis plays a pivotal role in transmitting lineage specific transcriptional control, imparting phenotypic memory and determination of cell fate. Nuclear organization and three-dimensional architecture of the genome is disrupted during mitosis, yet daughter cells preserve their gene expression patterns and phenotypic state. During mitosis most transcription factors dissociate from promoter sequences as well as from bulk chromatin. However, around 20% of transcription factors and chromatin binding proteins are retained at their specific target sites.

Retention of the RUNX1 tumor suppressor with target gene loci on mitotic chromosomes during cell division, a process we designate “Mitotic gene bookmarking,” which epigenetically sustains competency for gene expression during cell division.

In this study, we are investigating the role of RUNX1 in maintaining the epithelial phenotype and its loss in transformation to the mesenchymal phenotype using degron technology, targeting the RUNX1 factor for rapid degradation, to directly establish the requirement for this tumor suppressor to sustain gene expression that is obligatory for physiological control of normal and tumor phenotypes.

By combing the expression profiles of mature RNA and nascent transcript analysis we were able to measure rapid changes in the core of RUNX1 regulatory network and obtain a list of differentially expressed genes. Furthermore, loss of RUNX1 results in the emergence of a heterogenous population of cells that are resistant to apoptosis and undergo a transition between cellular states exhibiting partial EMT and phenotypic plasticity.

ACKNOWLEDGEMENTS

I would like to thank my thesis committee Drs. Gary and Janet Stein for the support they provided as I prepared for my defense. I would like to express my gratitude to Dr. Jonathan Gordan and Dr Frances Carr for their helpful feedback. A special thank you to Dr Jason Stumpff for taking the time to be a part of my defense committee. Lastly, I would like to thank Dr Andrew Fritz and Dr Louis Dillac for their endless patience and great mentorship throughout my research experience.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1: INTRODUCTION

1.1. Overview

Epigenetics is defined as the study of structural changes to the DNA which reversibly affects gene function but cannot be attributed to alterations in DNA sequence. Several epigenetic modifications may be passed on from parental cell to daughter cell and maintained across multiple cell generations [1,2]. These changes involve DNA methylation, posttranslational histone modification, nucleosome and chromatin remodeling, transcription factor binding and microRNA silencing.

The role of epigenetics in exerting biological control of cell and tissue structure, function, and phenotype can be recognized by the fact that all somatic cells in an organism have the same DNA, but the pattern of gene expression varies depending on the cell type. A study comparing phenotypic variation in monozygotic twins highlights that the differences in observable traits in genetically identical individuals is not merely influenced by the mechanisms underlying translation of genetic code but also on the complex regulatory machinery that interplays between the DNA and nuclear microenvironment [3].

Emerging evidence shows that epigenetic regulation is closely related to cardiovascular and neurological disease-related gene function and can influence the onset and progression of the disease. Epigenetic modifications also serve as important biomarkers for early detection, diagnosis, prognosis and interventional therapy [4,5]. In light of the significant role played by epigenetics in shaping cellular functions, gaining a deeper understanding of both normal and disrupted epigenetic mechanisms can offer valuable insights into potential therapeutic strategies for various diseases, including cancer. Cancer has long been viewed as a disease arising from the mutation

of oncogene or inactivation of tumor-suppressor genes. While the transformation of cancer cells through epigenetic modifications has been recognized for some time, recent advances in genome-wide genomic and epigenomic sequencing has revealed the extensive presence of mutations in epigenetic regulators and the comprehensive scope of epigenomic changes in cancer cells [6]. It is now evident that genetic and epigenetic mechanisms are not mutually exclusive and work in tandem to facilitate the acquisition of the characteristic traits of cancer.

Cancer cells also exhibit morphological alterations within the nuclear microenvironment. Nuclear organization, higher-level genomic configurations, and the spatial positioning of enhancers and promoters and other associated proteins within the cell nucleus, has a significant impact on gene regulation, genomic instability, cell proliferation and differentiation and lineage commitment, contributing to the onset and development of cancer[7,8]. Gaining insight into these fundamental processes is not only essential for comprehending the atypical gene expression associated with tumor initiation and progression but also provides a foundation for innovative strategies in cancer diagnosis and treatment.

1.2. Mechanisms of epigenetic control and pathogenesis of cancer

Eukaryotic genomes are packaged into compact, dense structures within the nucleus which can be influenced by modifications in the chromatin and the positioning and composition of chromatin components depending on the cell's specific needs. Despite the fact that the entirety of the human genome's sequence was finally completed in 2022 through DNA sequencing[9], it is not yet fully understood. Approximately 3.2×10^9 nucleotides of the human genome are distributed across 23 pairs of chromosomes. The nucleosome is the smallest structural component of chromatin, consisting of

octamers of histones (two molecules each of H2A, H2B, H3, and H4) around which the DNA is wrapped in a left-handed spiral. Together they yield a chromatin fiber of approximately 10 nm diameter separated by linker DNA segments averaging about 80 base pairs in length[10]. This loosely packed form of chromatin appears as beads on a string when viewed under a microscope. Multiple nucleosomes can be further condensed into compact 30-nanometer fibers. Thus, the accessibility of the DNA for transcription factors is dependent on the positioning of the nucleosomes and the chromatin state.

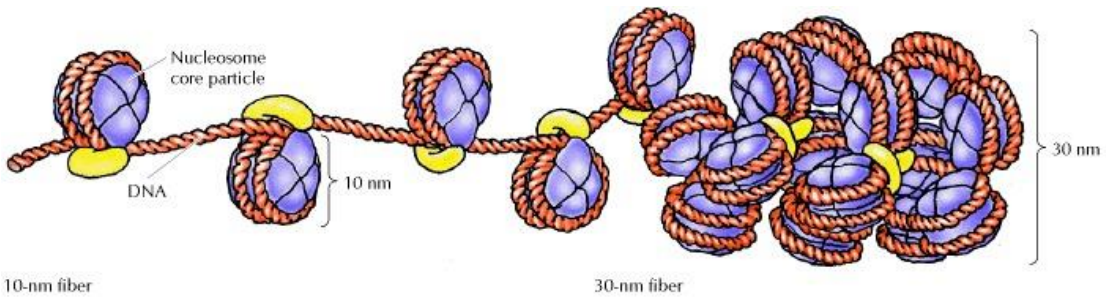


Figure 1: Schematic representation of the Nucleosome. *Lightly packed, euchromatin, appears as a 10nm beaded fiber and is usually under active transcription. Tightly packed, heterochromatin, is condensed into 30nm fibres and is less accessible for transcription. From: The Cell: A Molecular Approach. 2nd edition. Cooper G.M.*

To attain the precise level of regulation needed to coordinate nuclear functions like DNA replication, repair, and transcription, cells have evolved several mechanisms to fine-tune chromatin structure and function in localized and specific ways. The inability to inherit these hierarchal marks of epigenetic and architectural control can give rise to inappropriate initiation or inhibition of gene expression potentially causing pathological alterations, including the development of cancer[11,12].

1.2.1. Histone post-translational modifications

The positioning of nucleosomes along the chromatin is not random. Although histones can bind to any sequence of DNA, they tend to have preferential binding for certain sequences over others. The nucleosome contains hundreds of hydrogen bond-mediated direct histone-DNA interactions and just as many water-mediated ones. The conformation of the DNA within the nucleosome core is dependent on the number, site, and strength of the histone-DNA binding sites around the octamer surface. The bridging hydrogen bonds from the water molecules stabilize the nucleosome as well as enhance DNA motility and conformational variation[13], making the electrostatic and hydrogen bonding interactions between the DNA molecule and histone dimer highly dynamic[14]. The nature of the DNA-wrapping sterically blocks DNA-binding proteins and can inhibit or facilitate the attachment of RNA polymerases and transcription factors in specific locations. Spontaneous, transient site exposure by DNA unwrapping and rewrapping, nucleosome sliding and remodeling, are the major mechanisms that mediate the range of DNA accessibility and occupancy of DNA-binding proteins, through post-translational modifications of histones or histone variant substitution.

Histone proteins consist of a globular C-terminal core domain and an extended NH_2-terminal tail of variable length, both which are subject to extensive post translational modifications (PTMs), including methylation, acetylation, phosphorylation, ubiquitylation, ADP ribosylation, crotonylation, succinylation, malonylation, and biotinylation of specific residues on the amino acid chain. While most of the PTM sites are found within the NH_2 terminus, recent studies using mass spectrometry have identified novel sites residing in the histone globular core domain[15]. The signaling pathways involved in modifying, maintaining, and resetting histone marks are

interconnected in a dynamic balance such that all PTMs are reversible. Modifications to the amino acid residues of histones dramatically changes its characteristic and reactivity and has the potential for altering nucleosome-nucleosome and histone-DNA landscape, yielding different functional consequences. Histone acetylation or phosphorylation is usually associated with an open chromatin state which generally activates transcription, whereas deacetylation does the opposite. Histone methylation often leads to gene silencing, while demethylation activates them.

Linker histones like H1 and its various isoforms play an important role in compacting and regulating chromatin structure and act as binding sites for interacting proteins. They are positioned at the junction where the linker DNA connects with the portion of the DNA that wraps around the histone core[16]. Histone H1 and other non-canonical histone variants affect chromatin accessibility[17] and recurrent, loss-of-function, missense mutations of linker histone genes has been linked with tumor evolution in follicular lymphoma[18].

From a clinical perspective, targeting factors and enzymes involved in modulating epigenetic pathways using small molecule epigenetic modifiers in conjunction with traditional chemotherapeutic agents can have synergistic effects. These epigenetic drugs block the active site. Additionally, epigenetic processes can prepare and improve the receptiveness of treatment-resistant tumors to radiation, cytotoxic chemotherapy, and immunotherapy by reducing the obstacles that hinder the reception of signals from signaling pathways.

1.3. Mitotic Gene Bookmarking

The cell cycle is a tightly regulated series of events where all the genetic information within the cell is replicated and segregated into two identical daughter cells. The cell

progresses through 4 stages consisting of Gap 1 (G1), synthesis (S), Gap 2 (G2), and mitosis (M). G1 phase or Gap 1 phase is the first growth phase of a cell just after it exits cell division. During this phase the chromosomes are decondensed and cell resumes its biosynthetic activities in preparation for the next cell cycle, nearly doubling in size. The G1 phase holds particular significance within the cell cycle as it determines whether a cell will proceed with division or exit the cell cycle. Under favorable conditions and extracellular signaling, cyclin-CDK complexes are activated, driving the cell into S phase[19]. DNA replication origin sites are selected and a pre-replication complex, consisting of ORC (origin recognition complex), CDC6 (cell division control protein 6), CDT1 and MCM2–7 (minichromosome maintenance complex 2–7) are recruited to these sites[20]. During S phase or synthesis phase DNA replication is initiated and MCM proteins dissociate from the origin site so that DNA replication occurs only once during a single cell cycle[10]. After the genome is duplicated the cells enter G2 phase or Gap 2 phase, where cell growth continues, and proteins are synthesized for progression into M phase. Stringent G2 checkpoint machinery ensures that incomplete or damaged chromosomes are not replicated. The two Gap phases offer the cell an opportunity to monitor and repair its internal and external environment before it fully commits to transmitting its genetic and phenotypic identity to progeny cells.

There is an essential need for communication between the mechanisms orchestrating cell growth and proliferation, and these requirements are specific to each particular cellular phenotype. Both normal and cancerous cells can display unrestrained proliferation. Nevertheless, in tumor cells, there is frequently a compromise in the rigorous regulation of the cell cycle. This often results in an imbalance between proliferation, cell survival, and the ability to respond to physiological signals, along

with the development of resistance to therapies aimed at targeting cell cycle checkpoints.

Epigenetic control during mitosis

Mitosis plays a pivotal role in transmitting lineage specific transcriptional control, imparting phenotypic memory and determination of cell fate[20]. Nuclear organization and three-dimensional architecture of the genome is disrupted during mitosis, yet daughter cells preserve their gene expression patterns and phenotypic state.

Several dramatic events take place during the M phase starting with the condensation of duplicated DNA strands into elongated chromosomes. This is triggered by the dissociation of the ORC (origin recognition complex) from the chromatin[21]. Subsequently, the nuclear envelope disintegrates, and the duplicated chromosomes, each comprising a pair of sister chromatids, attach to the microtubules of the mitotic spindle. The sister chromatids are pulled apart towards the opposite poles of the cell followed by the division of the cytoplasm, resulting in completion of cell division. During mitosis most transcription factors dissociate from promoter sequences as well as from bulk chromatin[22], resulting in transcriptional activity being significantly downregulated[21,23]. Protein synthesis and RNA polymerase activity also decreases[24]. Recent DNase-seq and ATAC-seq analyses of mitotic chromosomes suggest that overall chromatin accessibility remains largely unchanged when compared to interphase. Local dynamic reduction in accessibility occurs over narrow regions of transcription factor binding sites and distal regulatory elements, whereas proximal promoters tend to maintain their accessibility[25]. Around 20% of transcription factors and chromatin binding proteins are retained at their specific target sites[21,22], especially those essential for cell fate maintenance. A majority of the

histone acetylation marks are lost, and histone acetyltransferases dissociate from the chromosome[26]. In contrast, histone H3 lysine 4 trimethylation (H3K4me3), a transcriptional activation mark, and histone H3 lysine 9 trimethylation (H3K9me3) and H3K27 trimethylation (H3K27me3), transcriptional repressive marks are retained during mitosis[27]. Similarly, Histone H4 lysine 5 acetylation (H4K5Ac) is a mark for active transcription at promoter regions during interphase and is maintained in mitosis[28]. Other than aging and cancer-associated hypomethylation, DNA methylation patterns are essentially maintained throughout mitosis[29]. Thus, bookmarking factors like histone modifications, DNA methylation patterns and lineage-specific transcription factors facilitates rapid resumption of transcriptional machinery responsible for maintaining fidelity of cell proliferation upon mitotic exit.

Transcription factors binding to chromosomes during mitosis.

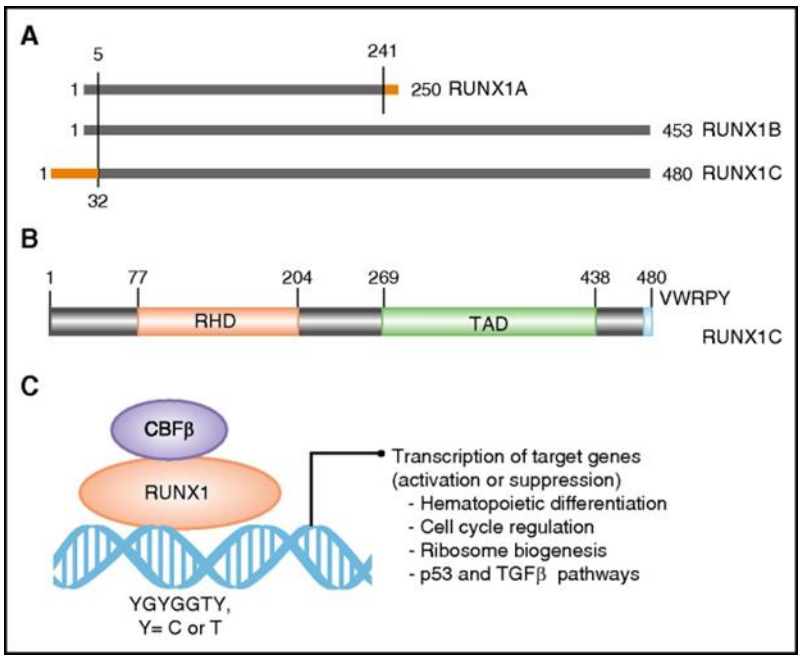
Transcription factors that have been observed to associate with mitotic chromosomes are often involved in cell fate regulation. GATA1, a major hematopoiesis regulator, is partially retained on the mitotic chromatin for rapid recruitment of its coregulators and reactivation of its bookmarked genes after mitotic exit[30]. TATA-binding protein (TBP) binds to RNA polymerase II preinitiation complex at the promoters that are transcribed during interphase and recruits several proteins including promoter-bound activators, TFIIA, and IIB. TBP-TATA complex is regulated by negative cofactor 2 (NC2), which inhibits and stimulates transcription in a context dependent manner[31]. Chromatin immunoprecipitation (ChIP) assays revealed that TBP, TFIID and TFIIB remain associated with active gene promoters during mitosis and NC2 remains associated with only some genes[32]. BRD4 (Bromodomain Containing 4), a chromatin reader, recognizes acetylated lysine residues and strongly binds to H4K5Ac. H4K5Ac

is maintained during cell division and serves as a gene bookmark for post-mitotic transcriptional re-activation[28].

The compartmentalization of RNA polymerase I and II machinery, RUNX transcription factors, and chromatin remodeling machinery in nuclear microenvironments is a pre-requisite for physiologically responsive gene expression.

1.4. RUNX1 Transcription factor

Runt-related transcription factor 1 (RUNX1) or Acute myeloid leukemia 1 (AML1) belongs to the RUNX family of genes (consisting of RUNX1, RUNX2, and RUNX3). The RUNX1 protein, also known as the CBFα subunit of the heterodimeric complex Core binding factor (CBF), is encoded by the RUNX1 gene located on chromosome 21. Various isoforms of RUNX1 can be transcribed through alternative splicing via two different promoters, distal promoter 1 or proximal promoter 2. All RUNX proteins contain the highly conserved runt-homology domain (RHD) near the N-terminal, which is responsible for its sequence specific DNA-binding and interaction with its non-DNA-binding partner protein CBFβ. CBFβ is instrumental to the function of all three RUNX factors because it allosterically enhances DNA binding of the CBFα subunit to the core site of various enhancers and promoter proteins. Isoforms 1B and 1C contain a large transactivation domain, which is necessary for protein-protein interactions. These two isoforms also share a five amino acid long VWRPY motif on its C-terminal, which binds Transducin-like Enhancer of split (TLE1)[33]. Through its interactions with different proteins, RUNX1 plays a crucial role in hematopoiesis, ribosome biogenesis, cell cycle regulation, and transforming growth factor β (TGFβ) signaling pathways[34].



Raman Sood,Yasuhiko Kamikubo,Paul Liu, Role of RUNX1 in hematological malignancies, Blood, 2017,

Figure 2: Domain structure of RUNX1 protein. *(A) The 3 major isoforms of RUNX1 (1A, 1B, and 1C). (B) Domain architecture of RUNX1. (C) RUNX1 heterodimerization with its binding partner, CBFβ, and interaction with DNA at promoters of target genes.*

RUNX2 is the master regulator of osteoblast differentiation and bone formation[35] and RUNX3 is involved in neuronal and T-lymphocyte development[34] as well as alveolar cell differentiation[36].

RUNX1 is one of the most commonly mutated genes in a variety of hematological disorders and malignancies. Genetic aberrations of RUNX1 range from somatic point mutations seen in myelodysplastic syndrome (MDS), to chromosomal translocations seen in leukemia. Unlike RUNX1 and RUNX2, RUNX3 inactivation occurs mainly through defects in DNA methylation and histone modification[37] leading to deregulated differentiation and preneoplasia, suggesting a tumor suppressor role for RUNX3.

1.4.1. Role of RUNX1 in Breast Cancer

Among the RUNX family proteins, RUNX1 has the highest expression in the normal human breast epithelium, specifically in basal and mammary duct-lining cells. It is the only member for which point mutations, frame-shift mutations and deletions have been identified in human luminal breast cancer[38]. All of the mutations are loss of function mutations that either affect amino acid residues that directly bind to the DNA, or disrupt the fold pattern of its RHD or abolish its dimerization to CBFβ, both of which would also inhibit its DNA-binding. While we know that RUNX1 is a master regulator of ER+ luminal mammary epithelial cells[39], the exact mechanisms of its transcriptional control and its role in determining cell fate through mitotic gene bookmarking has not yet been established.

1.5. Study Rationale and Hypothesis

Runx1 transcription factor plays a crucial role in mediating breast tissue specific gene expression and stabilizing the normal mammary epithelial phenotype. RUNX1 mutations are often associated with ER+, luminal subtype of breast cancer and indicates a tumor suppressor role for RUNX1. While the RUNX transcription factors have been shown to bookmark target genes during mitosis, there is a compelling requirement to functionally link RUNX1 bookmarking with the control of the breast cell phenotype and consequential loss of RUNX1 tumor suppression with an epithelial to mesenchymal transition, a critical step in tumor initiation. In this study we hypothesize that blocking RUNX1 mitotic bookmarking by degron-mediated RUNX1 proteolysis prior to mitosis prevents epigenetically mediated expression of breast cell phenotypic genes, initiates EMT, and confers stemness in progeny cells.

CHAPTER 2: MATERIALS AND METHODS

2.1. Cell line and culture

To address the role of RUNX1 this study primarily utilized the normal-like human mammary epithelial MCF10A cell line and MCF10A degron cells (clone 7) with FKBP12 ^{F36V} knock-in. Both, parental and degron cell were cultured in Dulbecco's Modified Eagle Medium (DMEM)/F-12 50/50 mixture base medium supplemented with 5% (v/v) horse serum-25ml/500ml media, 10μg/mL human insulin, 20ng/mL human epidermal growth factor (hEGF), 100ng/mL Cholera toxin, 0.5μg/mL of Hydrocortisone, 10ug/ml Penicillin-Streptomycin and 2 mM L-glutamine. Cells were fed every 2 days and grown to confluence before subculturing. Cells were washed once with Phosphate Buffered Saline (PBS) and treated with trypsin for 10-15 minutes before trypsin was neutralized with growth medium. Cells were then centrifuged at 500xg for 5 min and the cell pellet was resuspended in medium and transferred to 100mm cell culture plates.

2.2. RUNX1 Degron System

Using CRISPR-mediated locus-specific knock-in, members from our lab were able to tag the endogenous RUNX1 gene with an FKBP12^{F36V} epitope just prior to the stop codon of the RUNX1 in MCF10A cells. Thus, the transcribed protein will be expressed as RUNX1-FKBP12^{F36V} fusion protein, along with double hemagglutinin (HA) epitope tag. Addition of dTAG1, a small molecular degrader, induces rapid, reversible, and selective ubiquitination and proteasomal degradation of FKBP12^{F36V} fusion proteins by recruitment of E3-ligase, within the span of 1 hour.

2.3. Cell synchronization

Cells were synchronized using RO-3306 which selectively inhibits CDK1/cyclin B1 activity and reversibly arrests human cells at the G2/M border of the cell cycle. MCF10A cell plates at 60-70% confluency were given fresh growth medium supplemented with 5μM RO-3306. The G2/M population was harvested after 20 hours of treatment. For the G1 population, RO-3306 treated cells were washed once with PBS followed by 5 washes of growth factor-free media in order to release them from the RO-3306 block. Then they were cultured in RO-3306-free medium for three hours at 37ºC and harvested. Additionally, both G2/M and G1 populations were treated with DMSO or dTAG1 3 hours prior to being harvested.

2.4. Western blotting

Cells were harvested and lysed in 1x RIPA cell lysis buffer containing complete EDTA-free protease inhibitor cocktail tablets and 5μM MG132 proteasome inhibitor. Lysates were sonicated, cleared by centrifugation, and mixed with Laemmli sample loading buffer, followed by a 5-minute incubation at 100°C. Proteins were subjected to sodium dodecyl sulphate-polyacrylamide gel electrophoresis (SDS-PAGE) and transferred onto PVDF membranes. Membranes were blocked in 5% milk at room temperature for 1 hour and were incubated overnight at 4ºC in 5% bovine serum albumin (BSA) solution containing primary antibodies (1:1000 dilution). On the following day, membranes were washed and incubated in horseradish peroxidase-conjugated anti-mouse or anti-rabbit secondary antibodies. Proteins were visualized using Enhanced chemiluminescence (ECL) assay.

2.5. Flow cytometry and immunostaining

AldeRed ALDH Detection Assay

Cultured cells were detached from plates with trypsin and resuspended in growth media. 1 million cells were pelleted and resuspended in 1ml of AldeRed Assay Buffer supplemented with Verapamil and 5 µL/m l of the prepared AldeRed 588-A dye. The cell suspension was divided into two microcentrifuge tubes, where one tube was treated with 2.5µL of DEAB which served as a control. Cells were incubated for 30 min at 37º C, following which cells were centrifuged and resuspended in 300µL of AldeRed Assay Buffer for flow cytometry.

CD24 and CD44 staining.

Cultured cells were detached from plates with Accutase and resuspended in growth media. 1 million cells were pelleted and resuspended in PBS supplemented with 1% Fetal Bovine Serum and 1:5 dilution of anti-CD44-APC and anti-CD24-PEvio770 antibodies. The cell suspension was incubated for 15 mins at 4º C, following which cells were centrifuged and resuspended in PBS for flow cytometry.

H3pS28 staining.

Cultured cells were detached from plates with trypsin and resuspended in Phosphate-buffered saline (PBS). 1 million cells were pelleted and fixed with 80% ethanol for 30 minutes at 4º C followed by 10 mins of permeabilization using 0.25% Triton. The cells were centrifuged and washed with PBS and incubated overnight at 4º C in PBS supplemented with 0.5% BSA and H3pS28 antibody conjugated with FITC (1:100

dilution). On the following day, cells were centrifuged and resuspended in propidium iodide (PI) for flow cytometry.

2.6. RNA sequencing

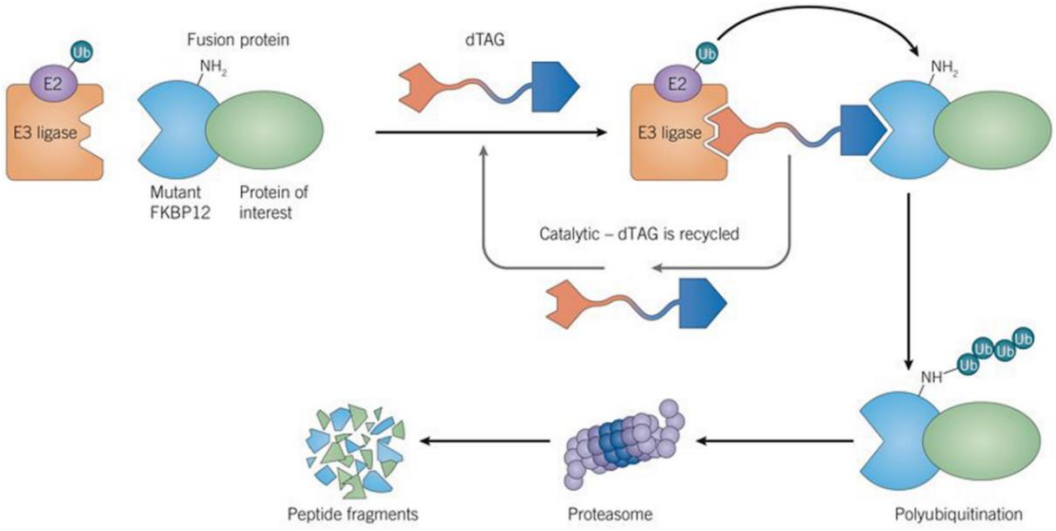
Cells were harvested and lysed using Trizol Reagent. RNA was isolated using Direct-zol RNA MiniPrep including the DNA digestion step. RNA isolates were quantified using Qubit Fluorometer and were checked for integrity using the Bioanalyzer System. Libraries were prepared using Takara Bio, HI Mammalian - SMARTer Stranded Total RNA Sample Prep Kit according to the manufacturer's instructions. The libraries were sequenced using Illumina NovaSeq paired end 150 bp sequencing.

CHAPTER 3: RESULTS

3.1. CRISPR-mediated FKBP12^{F36V} tagging of endogenous RUNX1.

Although numerous genetic techniques such as conditional knockouts and RNA interference, have been implemented to yield comprehensive scientific insight into the functions of target proteins in mammalian tissue, they fall short in terms of lack of reversibility, slow kinetics, off target effects and residual gene expression.

The emergence of the dTAG (degradation TAG) system offers an effective alternative to genetic knockdown/knockout. The system utilizes CRISPR/Cas9-mediated locus-specific knock-in of mutant FKBP12^{F36V} in frame with the protein of interest. This fuses FKBP12^{F36V} with our target protein, RUNX1. The fusion protein was also engineered to express double hemagglutinin (HA) epitope tag. Subsequent treatment with dTAG1, a heterobifunctional small molecule degrader, targets the entire fusion protein for degradation. The dTAG1 molecule is composed of an E3 ligase ligand linked to a FKBP12^{F36V} ligand, that is highly selective for the FBKP12 protein with a single point F36V mutation and does not affect endogenous wild type FKBP12. dTAG1 forms a ternary complex between an E3 ubiquitin ligase, von Hippel Lindau (VHL) protein, and the RUNX1- FKBP12^{F36V} fusion protein, leading to its polyubiquitination and subsequent degradation by the proteasome within 1 hour. This process is rapid and reversible, allowing for the identification of direct targets of RUNX1 and investigating the effects of the loss of epigenetic bookmarking by eliminating RUNX1 just prior to mitosis.



https://www.tocris.com/resources/tocris-blog/dtag-protein-degradation-platform-target-validation

Figure 3: Mechanism of action of dTAG1.

To validate the degron system and the efficacy of dTAG1, we performed Western blot analysis at different timepoints. Treatment of MCF10A RUNX1-FKBP12^{F36V} cells with 500nM of dTAG1 caused rapid RUNX1 degradation within the first 1 hour (Figure 4). A 12kb mobility shift consistent with the addition of FKBP12^{F36V} was also observed.

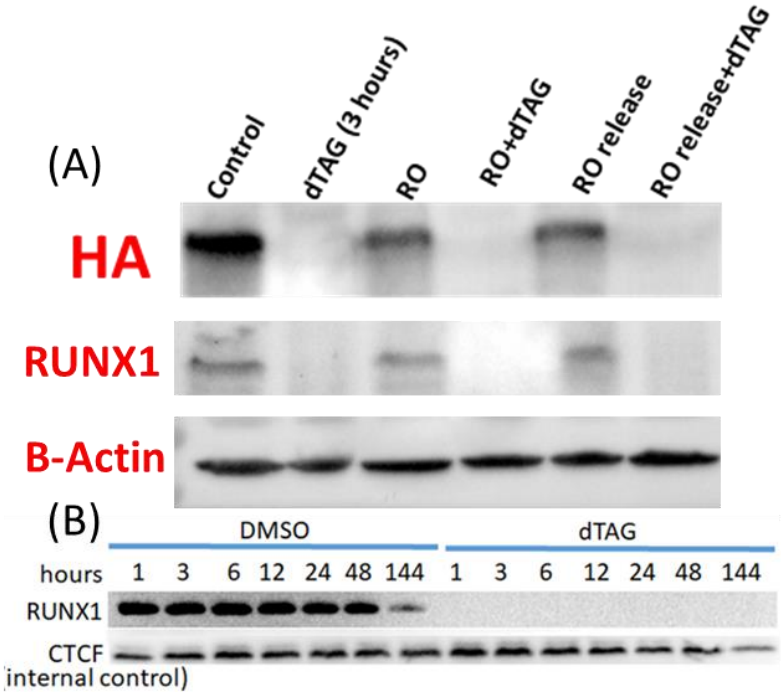


Figure 4: Western blot showing the protein expression of RUNX1 and HA in MCF10A Degron cells treated with DMSO and dTAG1 *(A) Cells in the asynchronous, G2/M and G1 phase of the cell cycle showing complete degradation of RUNX1 and HA after 3hrs of dTAG1 treatment. (B) Short-term and long term timepoints (1, 3, 6, 12, 24, 48, and 144 hours) show complete absence of RUNX1 protein in dTAG1 treated cells.*

3.2. Loss of RUNX1 induces a morphological change in MCF10A cells.

During our validation experiments, actively proliferating MCF10A parental and degron cells were treated with dTAG1 or DMSO-control and were monitored via phase contrast light microscopy. MCF10A cells have a cobble-stone appearance characteristic of epithelial cells. The MCF10A degron cells treated with DMSO maintain this morphology, however dTAG1-treated cells begin to exhibit mesenchymal-like features. After 48 hours of treatment a few cells become elongated, spindle-like, with abundant cytoplasm and indistinguishable cell borders. This is consistent with previous findings in the lab where significant macroscopic changes were first observed at 48 hours[40]. By 6 days, their number has grown significantly.

The parental MCF10A without the FKBP12 do not change morphologically upon the addition of dTAG1, confirming that the cellular transformation is a result of the loss of RUNX1 and not an off-target toxic effect of the dTAG1 molecule.

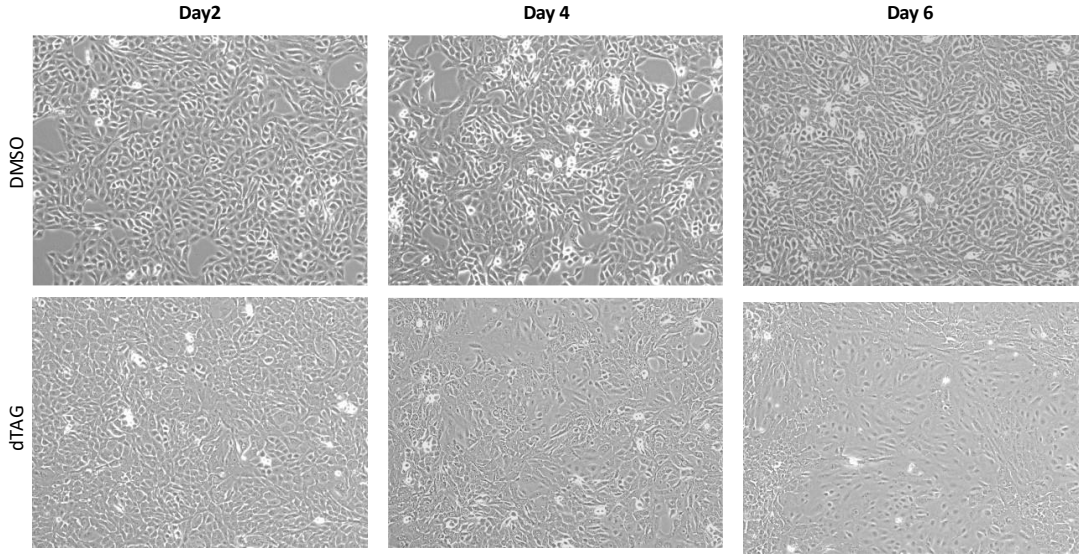


Figure 5: Phase contrast microscopy images of morphological changes in MCF10A degron cells after treatment with DMSO and dTAG1 for 2, 4 and 6 days. *Cells were photographed at 10x magnification. By Day 6, dTAG1 treated cells have transformed into elongated, spindle-like cells, with abundant cytoplasm and indistinguishable cell borders.*

3.3. MCF10A cells treated with dTAG1 survive in growth factor depleted media.

In addition to the morphological changes, we observed that MCF10A degron cells treated with dTAG1 and left in continuous culture for prolonged periods without re-feeding were resistant to cell death and maintained confluency longer than the DMSO treated cells. 1 x 10^6 cells were plated in 100mm tissue culture plates and treated with 500nM of DMSO or dTAG1 on the following day. Within two days in culture, cells had grown to confluency and were subcultured at 7 x 10^5 cells/plate with fresh growth media supplemented with DMSO or dTAG1. On Day 4, cells were refed with DMSO/dTAG1 supplemented media and left in continuous culture without media change.

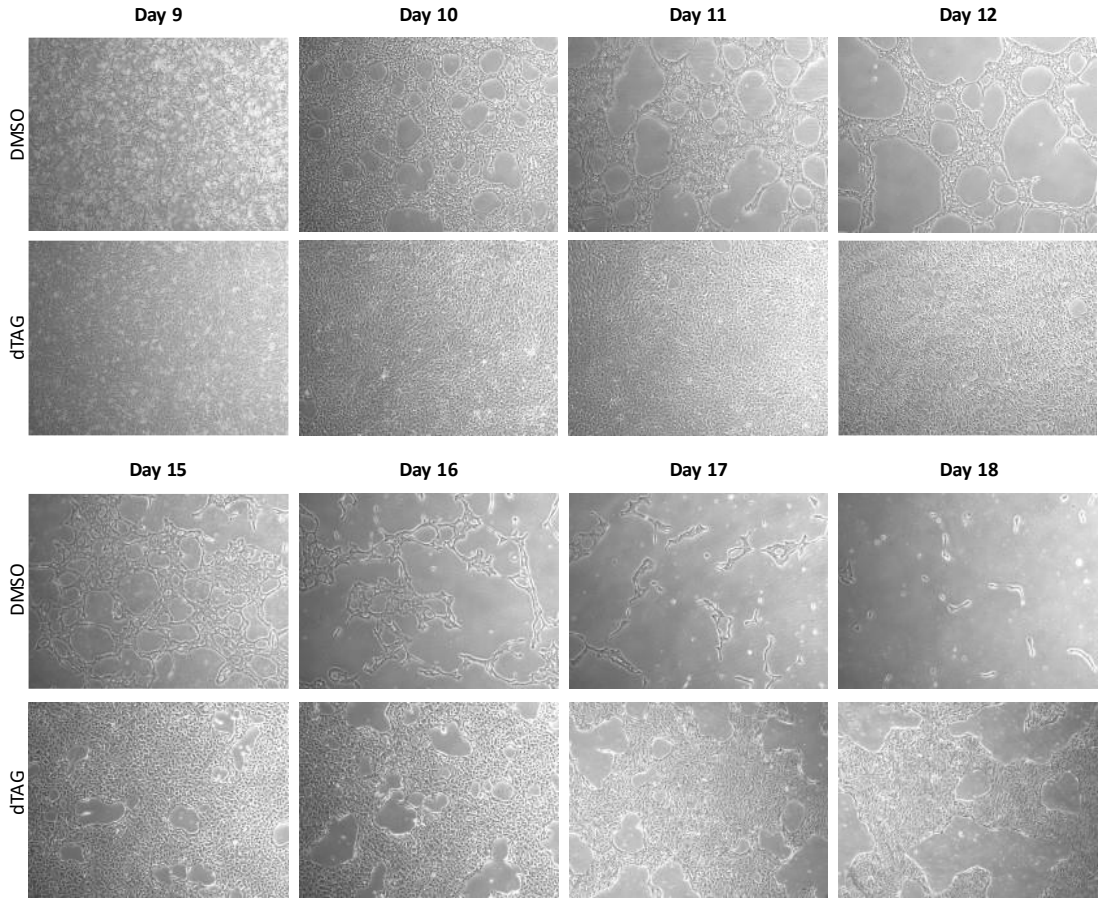


Figure 6: Survival of MCF10A degron cells in the absence of growth factors. *Images were taken at 4x magnification at different time points (9, 10, 11, 12, 15, 16, 17, 18 days) of treatment with DMSO or dTAG1 without refeeding the cells with fresh growth media.*

Cells in the DMSO plates begin to visibly die by Day 9 of treatment (5 days since last media change), observed as small, empty, cell-free pockets within the 2D culture. By Day 12, around 50% cells are still alive, and by Day 18, nearly all of the cells have died. In contrast, dTAG1-treated cells continue to survive even after growth factor depletion. They remain confluent until approximately Day14-15 of dTAG1 treatment (10-11 days since last media change) at which point we begin to see the appearance of empty pockets. By Day 18 (12 days since refeeding), the confluency of dTAG1 treated cells has gone down to around 50-60%. By Day 27-28, nearly all of these cells are dead. Notably, both DMSO and dTAG1-treated cells recover and return to active proliferation once all the cells receive fresh media on Day 18.

3.4. Identification of the core set of genes dysregulated by RUNX1 in normal-like mammary epithelial cells

Based on these observations (Figure 5 and 6), it is resonable to speculate that RUNX1 stabilizes the phenotype of normal breast epithelial cells and its absence modifies the epigenome. To support our hypothesis we designed a series of experiments to explore the role of RUNX1 in mitotic gene bookmarking. We were able to synchronize the cells using RO-3306, a reversible inhibitor of cyclin-dependent kinase-1 (CDK1), which is crucial for the G2 to M phase transition. After achieving effective cell cycle arrest at the G2/M border, with a treatment duration of 20 hours, we removed RO-3306 from the culture medium by washing the cells several times. At this point we added dTAG1 to the fresh RO-3306-free growth media. The reactivation of CDK1 allows cells to progress from the G2/M phase, into early G1 phase. The rapid onset of action of dTAG1 ensures that the cells proceed through mitosis in the absence of RUNX1, thereby providing a comprehensive picture on the consequences of loss of epigenetic bookmarking. We performed RNA-seq and PRO-seq on asynchronous, G2/M arrested and G1 cells that were depleted of RUNX1 after 3 hours of treatment with dTAG1. Each of these treatment conditions were compared with their respective DMSO control. While RNA-seq is a popular high-throughput sequencing technique for analyzing and quantifying the RNA content, the assay relies on accumulation of mature RNA over hours, making it inappropriate for measurement of rapid changes in gene expression at a specific time point. An advantage of PRO-seq over RNA-seq is that it captures a genome-wide snapshot of actively transcribing RNA polymerases, and detects nascent transcripts. By comparing the expression profiles with the DMSO-control and combining PRO-seq and RNA-seq measurements (using a $\log_2$Fold-

change cutoff of 0.5 and adjusted p-value < 0.001), we were able to obtain a list of differentially expressed genes.

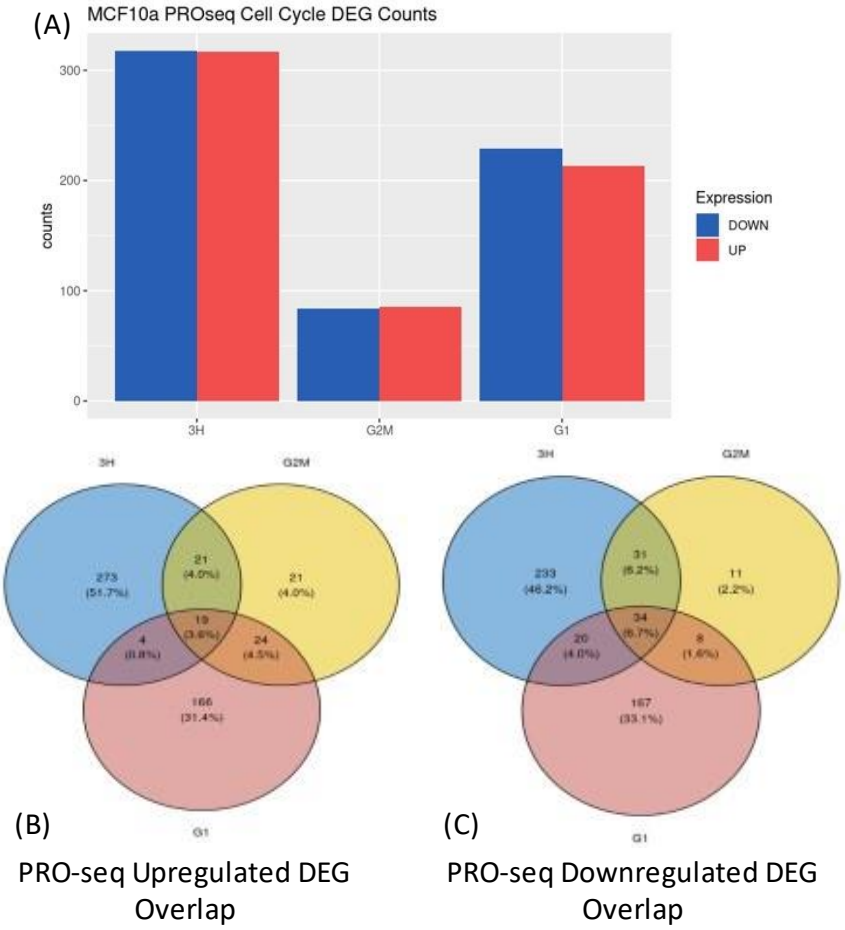


Figure 7: Differentially expressed genes from PRO-seq analysis of the cell cycle. *(A) A graph representing the total counts of upregulated and downregulated genes in asynchronous, G2/M and G1 populations treated with dTAG1 for 3 hours in comparison with its corresponding DMSO control. (B) An overlap of the upregulated genes between the phases of the cell cycle. (C) An overlap of the downregulated genes between the phases of the cell cycle.*

In accordance with previous research[21–23], the rate of transcription slows down during the G2/M phase and picks back up as the cells enter G1. Through PRO-seq, we observed an approximately equal number of genes were repressed and activated upon RUNX1 degradation. Out of the total of 169 differentially initiated genes that were identified in late G2 and early mitosis, 11 were transcription factors. GLIS3, VAX2

TBX3, ID1, BHLHE41, IRF1, and FOSL1 have decreased expression and NR3C2, RUNX2, BCL6, and POU2F3 have increased expression.

Because these key transcriptional regulators were already induced prior to cell division, we can infer that they may also contribute to the mitotic bookmarking profile of the cell. Thus, the subsequent alteration in its postmitotic gene regulation is indirectly mediated by the loss of RUNX1.

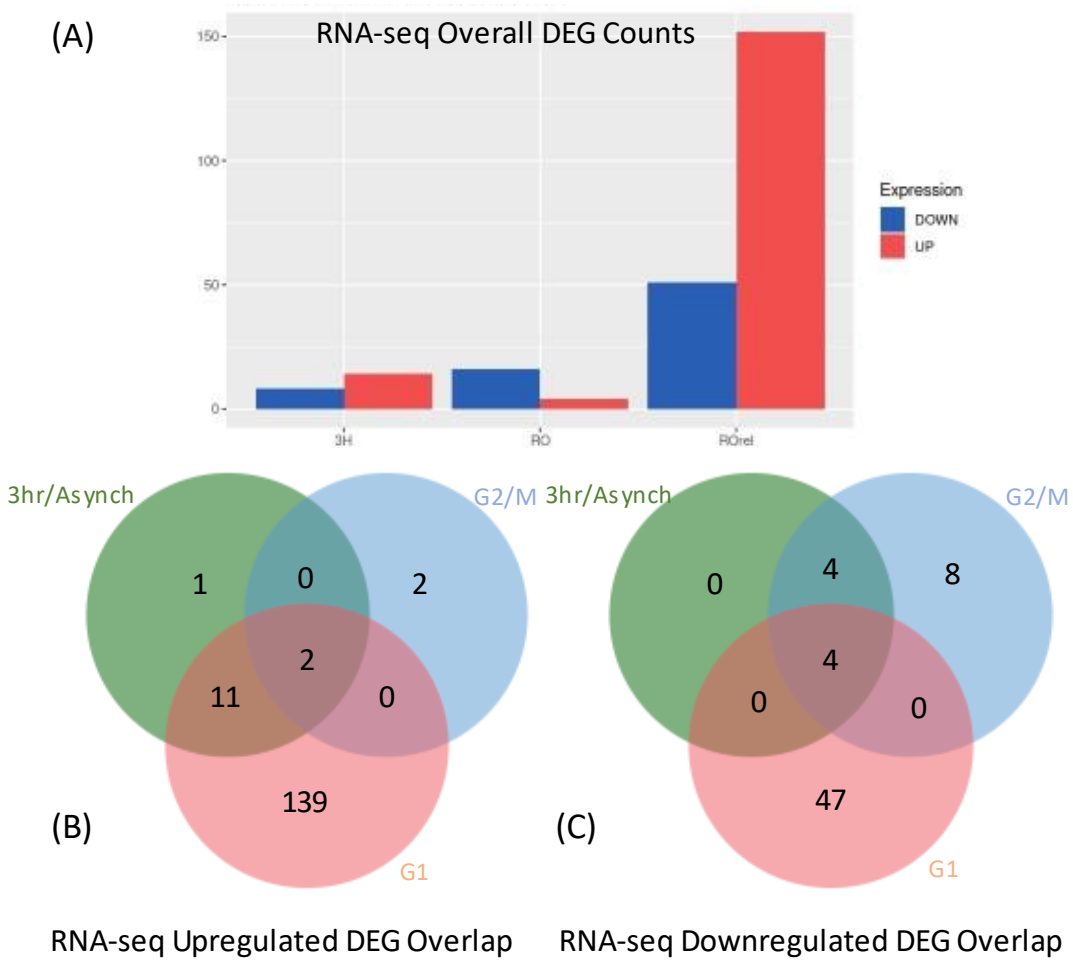


Figure 8: Differentially expressed genes from RNA-seq analysis of the cell cycle. *(A) A graph representing the total counts of upregulated and downregulated genes in asynchronous, G2/M and G1 populations treated with dTAG1 for 3 hours in comparison with its corresponding DMSO control. (B) An overlap of the upregulated genes between the phases of the cell cycle. (C) An overlap of the downregulated genes between the phases of the cell cycle.*

Unlike PRO-seq, where 635 genes were differentially initiated after 3 hours of dTAG1 treatment of the asynchronous cells, only 22 stable mRNA transcripts were differentially expressed in the RNA-seq results. Additionally, more genes are activated than repressed due to the loss of RUNX1 in G1 phase of the cell cycle. Apart from the 6 genes that are differentially expressed across the cell cycle, the upregulated genes overlap between the asynchronous cells and early G1, while there is no overlap with G2/M phase genes, implying that most of the genes that are upregulated (98%) upon the loss of RUNX1 are associated with G1 phase of the cell cycle. These genes are normally suppressed by RUNX1, and are now activated as the cells exit mitosis. Similarly, downregulated genes are shared between the asynchronous and G2/M phase, while there is no overlap with G1. This indicates that these genes are highly specific to certain phases of the cell cycle and RUNX1 is necessary for their transcription. NKX3-1, a transcription factor and diagnostic biomarker for prostate cancer, is downregulated in G2/M. It may also have a cascading effect on the transcriptional changes brought about by degradation of RUNX1.

In order to understand the functional profile and underlying biological processes of the transcriptomic remodeling following the disruption of RUNX1 bookmarking, we performed a Gene Set Enrichment Analysis (GSEA), overlapping the differentially expressed genes in the dTAG1 treated cells with genome-wide expression profiles of gene sets from Molecular Signature Database (MSigDB).

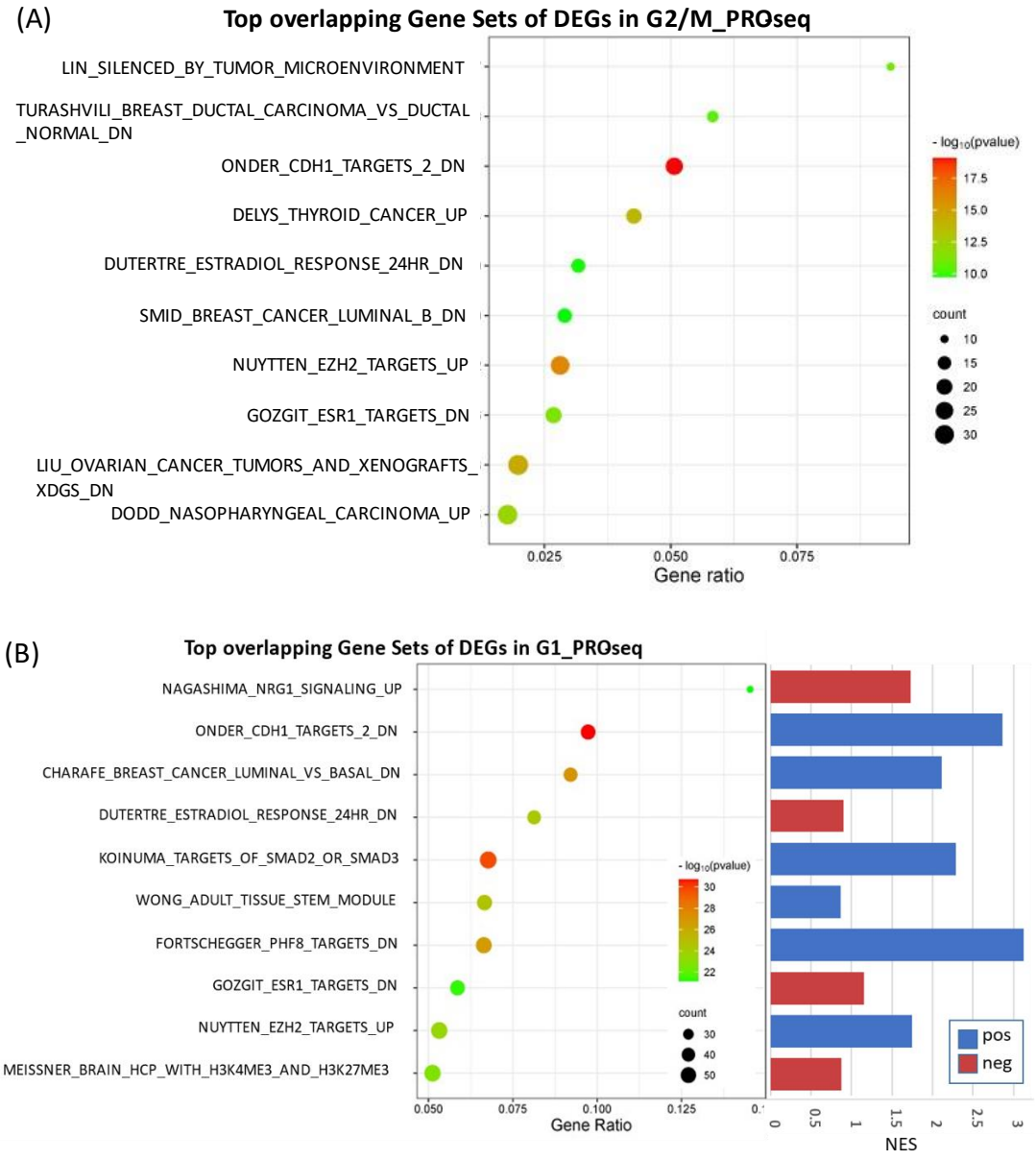


Figure 7: Pathway analysis of the differentially initiated genes based on the PRO-seq results post RUNX1 protein degradation across the different phases of the cell cycle. *(A) Pathways with the most statistically significant overlaps of gene sets in the G2/M phase of the cell cycle based on the gene ratios using an FDR q-value cutoff of 0.05 and p-value cutoff of 0.001. (B) Pathways with the most statistically significant overlap of gene sets in the G1 phase of the cell cycle along with their normalized enrichment scores.*

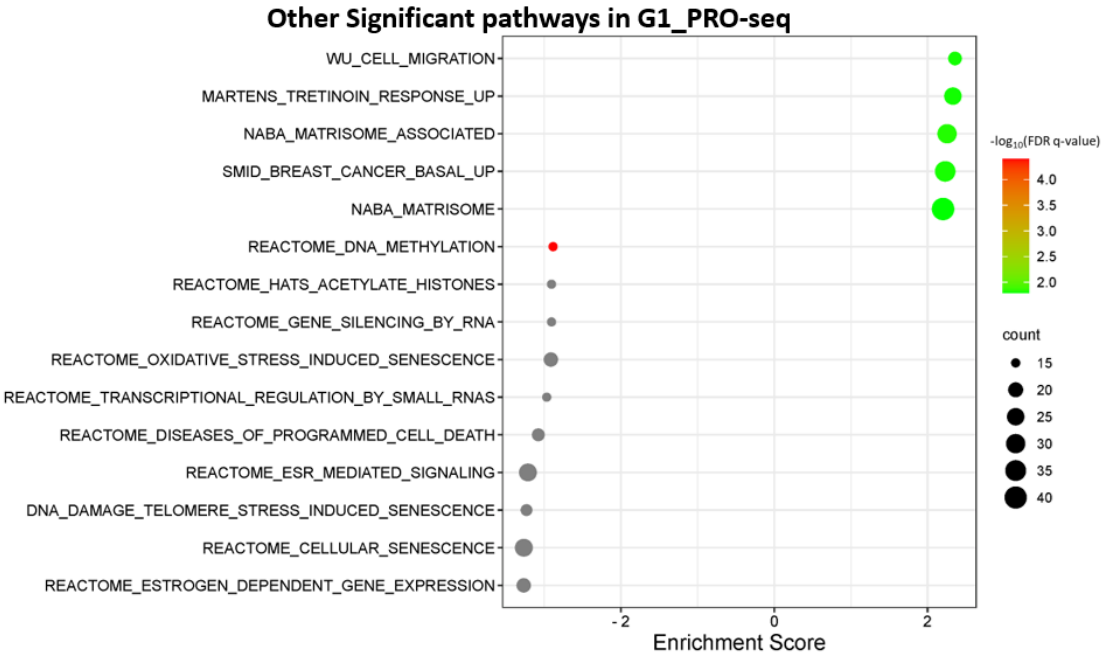
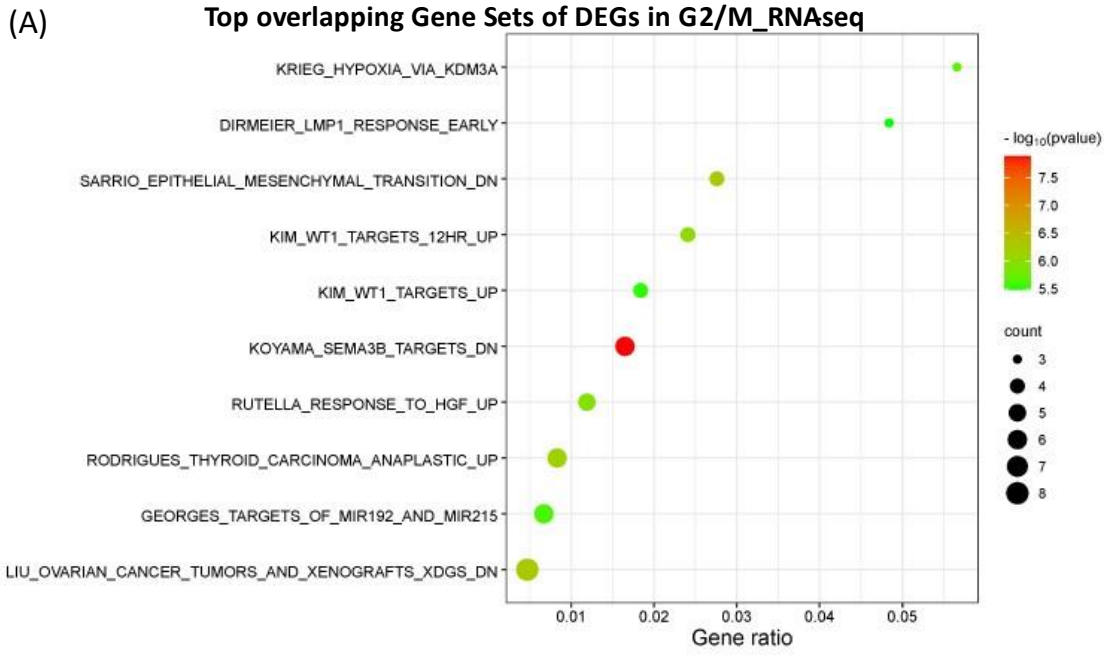


Figure 8: Pathway analysis of the differentially initiated genes based on the PRO-seq post RUNX1 protein degradation across the different phases of the cell cycle. *Pathways were picked based on a statistically significant Normalized Enrichment score with FDR q-value cutoff of <0.05 and p-value of <0.001.*

(A)



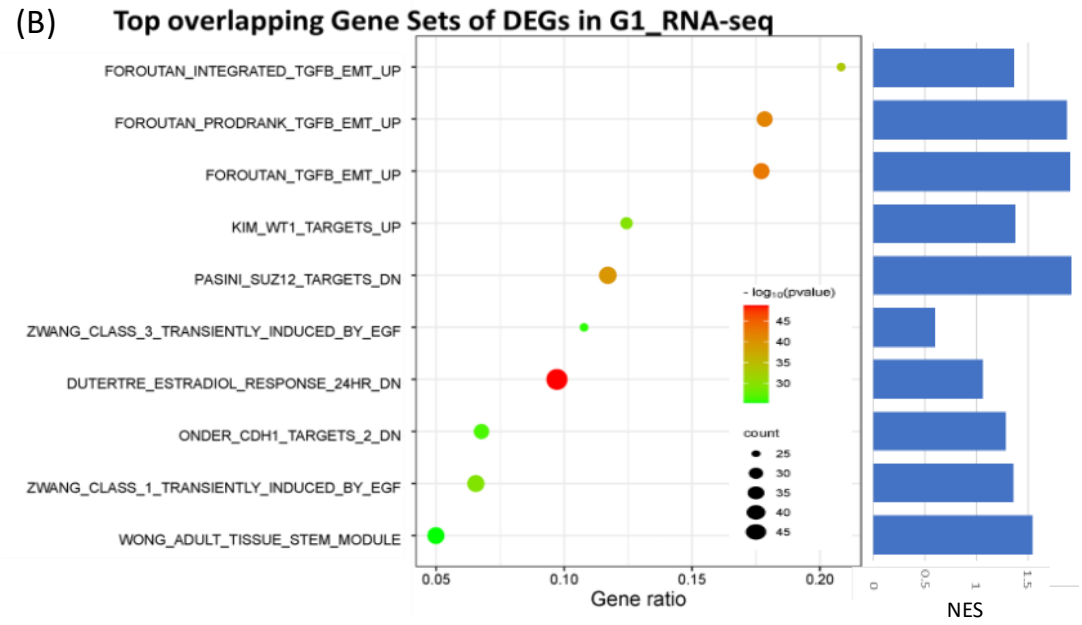


Figure 9: Pathway analysis of the differentially expressed genes based on RNA-seq results post RUNX1 protein degradation across the different phases of the cell cycle. *(A) Pathways with the most statistically significant overlaps of gene sets in the G2/M phase of the cell cycle based on the gene ratios using an FDR q-value cutoff of 0.05 and p-value cutoff of 0.001. (B) Pathways with the most statistically significant overlap of gene sets in the G1 phase of the cell cycle along with their normalized enrichment scores.*

Assessment of the gene set enrichment shows that differentially expressed genes in G2/M phase are negatively enriched in the Gozgit ESR1 targets DN gene set. This gene set compares the gene expression profile between TMX2-28 cells, an estrogen receptor 1 negative (ESR1 or ER) variant of the ER-positive MCF-7 cell line, and the parental MCF-7 cell line. This indicates that the genes that are down regulated in ER-negative cells are also down regulated in our differentially expressed gene sets obtained from PRO-seq analysis, further corroborating the ER negative nature of the MCF10A cells. Despite being an ER-negative cell line, genes involved in estradiol response are expressed in G2/M indicating that RUNX1 plays a role in regulating estrogen receptor target genes and estrogen response. In terms of the phenotypic profile, the DEGs overlap with gene sets associated with basal-like breast cancer cell

lines. In addition, genes that are upregulated after the knockdown of EZH2 are also upregulated in our gene set, whereas many of the genes that are downregulated upon the knockdown of CDH1 are upregulated upon the loss of RUNX1. These characteristics are maintained as the cells transition from mitosis to G1 in the absence of RUNX1.

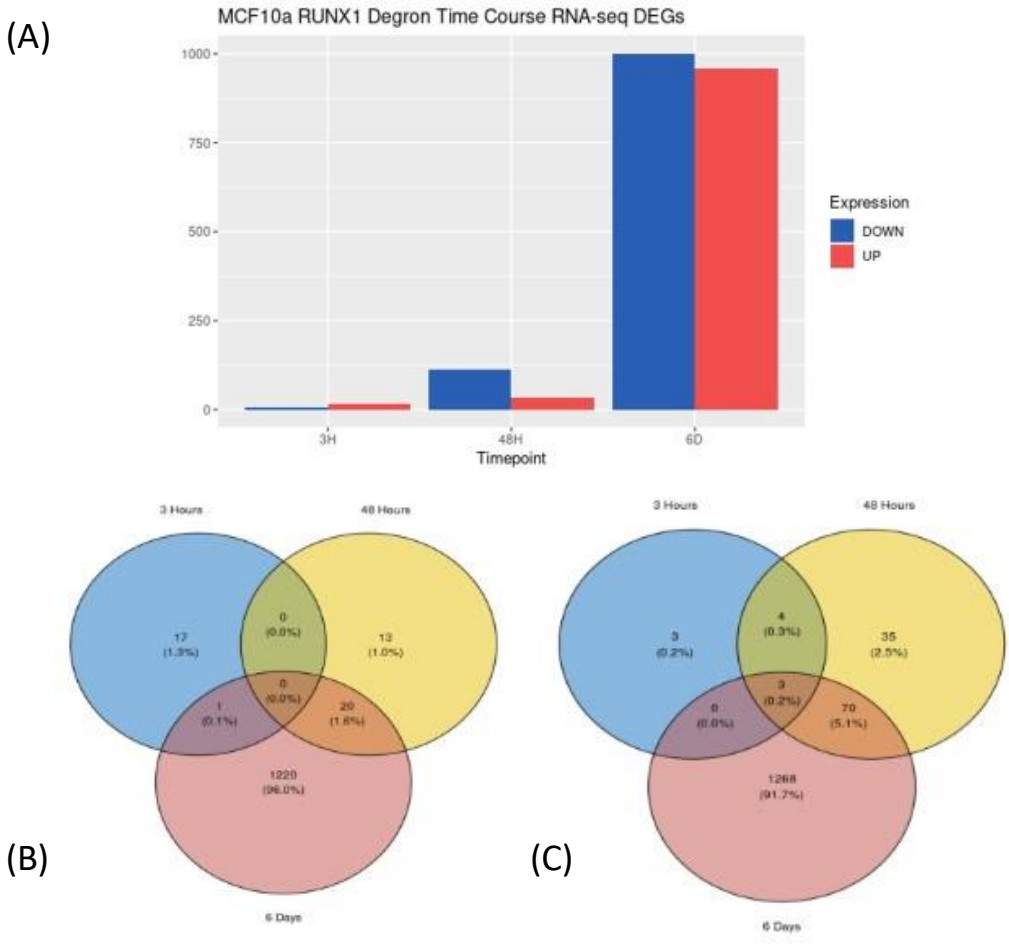
In G1, gene sets associated with Transforming growth factor beta (TGFβ) related EMT as well as Adult stem tissue module are positively enriched.

While the cells maintain their basal-like gene expression pattern throughout cell division, results from the cell cycle analysis of RUNX1 degradation indicate that the cells are poised to undergo epithelial to mesenchymal transformation and are taking on stem-like properties as they exit mitosis. There is a significant overlap with gene sets related to cell migration, extracellular matrix regulation, and escape from programmed cell death and cellular senescence. Particularly, differentially initiated genes in G2/M are enriched in pathways related to prostate, thyroid, nasopharyngeal and breast ductal carcinoma. Through GSEA analysis we can deduce that these genes are regulated by a complex balance between RUNX1 mediated DNA methylation, histone acetylation and gene silencing by RNA.

To assess if this change in the pattern of gene expression is sustained through multiple cell cycles, we performed high-throughput next-generation sequencing of the transcriptome in asynchronous cells at long term time points.

MCF10A degron cells were plated at a density of 1 million cells per 100mm tissue culture plate. They were given fresh growth media containing 500nM of either dTAG1 or DMSO (control) on the following day and cultured for 48 hours. Considering that the doubling time for MCF10A cells is 24 hours, the cells reached a confluency of approximately 90% after 2 days of treatment, at which point, a set of

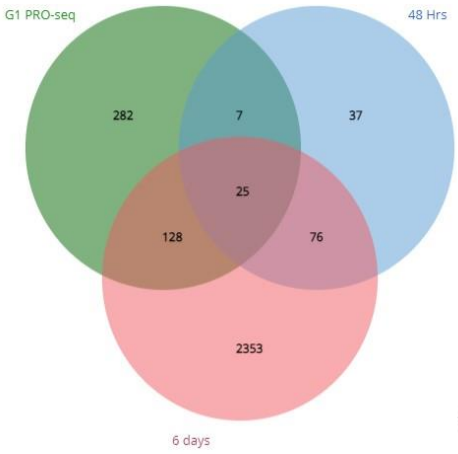
plates were harvested for the 48-hour mark. The cells from the remaining sets of plates were dissociated using trypsin and replated at a density of $6x10^5$ – $7x10^5$ cells per plate along with fresh media supplemented with 500nM of DMSO or dTAG1. The cells were refed on Day 4 and finally harvested on Day 6 of treatment. An additional set of asynchronous cells were treated for 3 hours to investigate the immediate changes upon the loss of RUNX1.



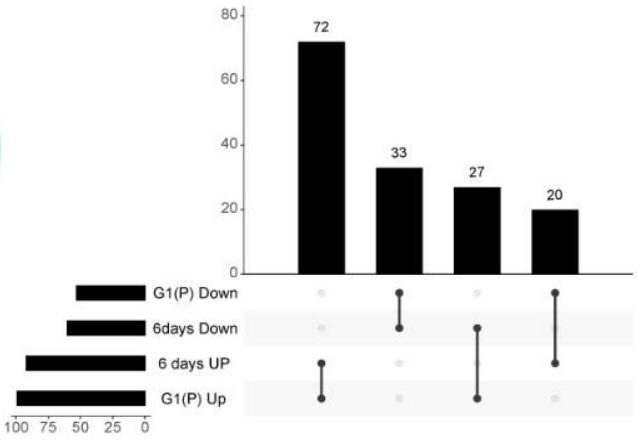
RNA-seq Upregulated DEG Overlap RNA-seq Downregulated DEG Overlap

Figure 10: Differentially expressed genes from RNA-seq analysis over the course of 3 hours, 48 hours, and 6 days. *(A) A graph representing the total counts of upregulated and downregulated genes in asynchronous cells treated with dTAG1 for 3, 48 and 144 hours compared to its corresponding DMSO control. (B) An overlap of the upregulated genes between the different timepoints. (C) An overlap of the downregulated genes between the different timepoints.*

Long term RNA-seq analysis of RUNX1 degradation showed a gradual rise in the number of differentially expressed genes from 3 hours to 48 hours of RUNX1 degradation, followed by a sudden burst at 6 days. There is very little overlap of genes between the 3 timepoints. Thus, a wave of changes in the steady state mRNA concentration, beginning 3 hours after addition of dTAG1 was observed; genes that were upregulated at 3 hours begin to wane by the 48-hour mark and new sets of genes are activated. Only 1.6% of these genes are maintained up to 6 days of dTAG1 treatment. A higher percentage (5.1%) of downregulated genes at 48 hours continue to remain suppressed by 6 days, likely due to the loss of their transcriptional activation by RUNX1. A little over one third of the 28 differentially expressed genes (DEGs) at 3 hours of dTAG1 treatment are downregulated. This is reversed at 48 hours, where 77.24% of DEGs are downregulated. By 6 days, the number or upregulated and downregulated genes is roughly equal (48.06% and 51.93% respectively). However, when compared with the PRO-seq results there is a greater overlap, indicating that the genes that were differentially initiated in early G1, upon the loss of RUNX1, are being fully transcribed at a later stage.



(A) Overall DEGs



(B) Up and down regulated genes

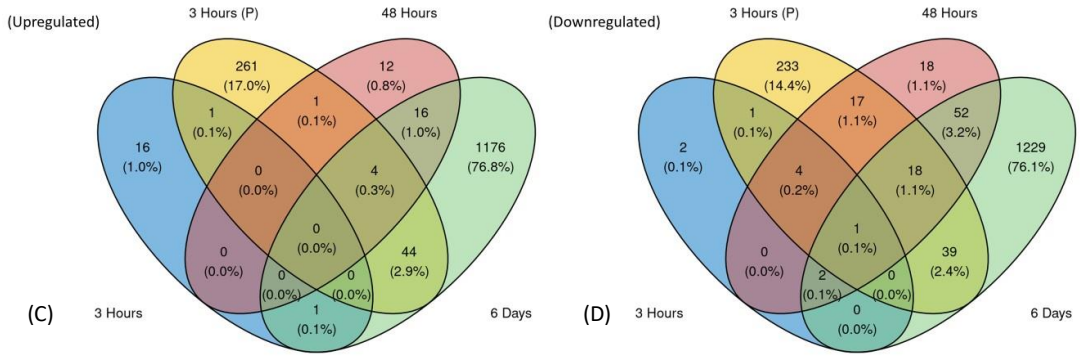


Figure 11: Overlap of differentially initiated genes from the PRO-seq results with the long-term treatment RNA-seq results. *(A) Overlap of overall differentially initiated genes between early G1 phase of the cell cycle and, 48 hours and 6 days of dTAG1 treatment in asynchronous cells. (B) An upset plot depicting the changes in direction of regulation in the 153 genes that are in common between G1 phase (PRO-seq) and 6 days. (C) Overlap of the upregulated genes between 3 hours (PRO-seq), 3 hours (RNA-seq), 48 hours, and 6 days of dTAG1 treatment.*

FOLD CHANGE			
Gene	G1	48 Hr	6 Days
SCEL	2.340982	5.547962	7.51419
SERPINB7	2.055648	2.625162	6.155352
ADGRF1	1.686192	1.839869	6.043246
MAP2	1.515384	1.561517	5.158765
NDRG1	1.753362	1.884301	4.3637
RETREG1	1.960892	1.728629	3.170361
GDPD3	1.706199	1.644852	2.393014
FGFBP1	1.760765	2.386227	2.114392
PSCA	0.595516	0.51911	1.523943
TGFBR2	1.56658	0.578531	0.666568
IGFL2-AS1	0.666266	0.401241	0.551951
DPYD	0.611448	0.54977	0.492667
PPARGC1A	0.435718	0.185489	0.404518
CDC42EP4	0.582891	0.637594	0.332971
KDR	1.692773	0.483678	0.332447
TNS4	1.548147	0.639813	0.303223
GJB2	0.34457	0.538956	0.267007
AREG	1.720883	0.557983	0.241427
TSC22D3	0.498494	0.498052	0.233799
MMP7	0.348771	0.182598	0.232355
TFCP2L1	2.090199	0.452125	0.1687
GPAM	0.466179	0.271828	0.1618
CCDC102B	0.656834	0.315698	0.146068
RP1	0.556967	0.382562	0.085886
ZBTB16	1.722614	0.391496	0.08576

FOLD CHANGE		
Gene	G1	48 Hr
RNF150	0.63875	0.632737
THBS1	1.596271	0.621227
PTTG1IP	2.07129	0.602995
GLIPR1	0.604941	0.506047
TGFB2	0.666446	0.498688
FBXO32	0.471447	0.376795
PLEKHS1	0.244648	0.195124

FOLD CHANGE		
Gene	G1	6 days
ARRDC4	0.474401	2.25102
SCNN1A	1.524974	0.57529
PLEC	1.555185	0.555946
H1-4	0.565665	0.543598
TACC2	1.617675	0.510432
STMN1	0.631062	0.505238
PYGB	2.01214	0.485853
PXK	0.386469	0.48283
COL4A6	1.514474	0.480657
LINC0304(	1.564575	0.479079
AVPI1	0.565473	0.477902
ITGA6	1.597031	0.462505
H2BC15	0.611152	0.460844
PDLIM2	1.690698	0.420193
SLC26A2	1.681398	0.413495
UNC5B	1.638929	0.411832
H1-3	0.313398	0.404401
PIK3R1	1.654424	0.396893
H2AC7	0.535127	0.378954
GALNT10	1.571348	0.365794
DDIT4	1.556335	0.361881
CHRM3	0.616592	0.345768
UBE2C	0.64353	0.331719
FLRT3	1.644579	0.311589
UNC5C	0.654008	0.298848
FHOD3	0.505794	0.298755
PER1	0.61171	0.293471
RAD51	0.662141	0.287482
TRNP1	1.572593	0.286547
COPS8	0.651835	0.27633
H3C2	0.409621	0.253709
E2F1	0.667484	0.243248
STC2	1.957681	0.242029
AURKB	0.642388	0.24042
TYMS	0.66893	0.239747
H2AC4	0.336656	0.233306
H1-5	0.405206	0.223484
ADAT2	1.545106	0.223222
NPR3	0.513328	0.220815
FKBP5	1.516444	0.092627
CCNA1	0.535909	0.043339
PID1	2.795658	0.040568

FOLD CHANGE		
Gene	G1	6 days
TNFSF15	2.639621	43.1134
SPNS2	1.959726	16.11442
GLYATL2	0.450152	14.12527
THSD7A	1.68367	13.69725
CPA4	1.6716	12.3573
CXCL1	3.950954	11.95342
CCL22	2.66735	10.57026
PLAU	3.276027	10.52394
RGL1	1.517131	8.358986
KITLG	1.63162	7.219563
IGFBP3	1.694469	6.71826
HS3ST1	1.797232	6.599775
PLCXD2	1.988908	6.404854
N4BP3	1.728598	5.672625
NCF2	2.112558	4.944397
NR3C2	1.927433	4.870388
DAPK2	1.595031	4.555157
HID1	1.58063	4.516772
SLPI	1.673421	4.434712
DUSP8	1.513985	4.11163
PIM1	1.535128	4.038709
NOS1AP	1.699137	3.969923
PLK2	2.054552	3.83964
A2ML1	1.62689	3.529853
TLR1	0.386295	3.339005
TMEM45B	1.554438	3.262323
WNT7B	2.247507	2.996626
TP53I11	1.545349	2.994952
KRT17	1.562998	2.945022
PMEPA1	1.554389	2.93338
TTC9	1.614514	2.761342
PRDM1	1.664263	2.72933
PDE4B	2.600407	2.71435
FLRT2	1.614561	2.600911
PHACTR3	5.406467	2.551147
CREBRF	0.6681	2.546897
ALDH1A3	1.987648	2.531396
TLR6	0.603938	2.46912
MYLIP	0.620694	2.373657
EDN1	1.853069	2.318806
NCEH1	1.63703	2.31829
GABARAPL	0.66468	2.25419

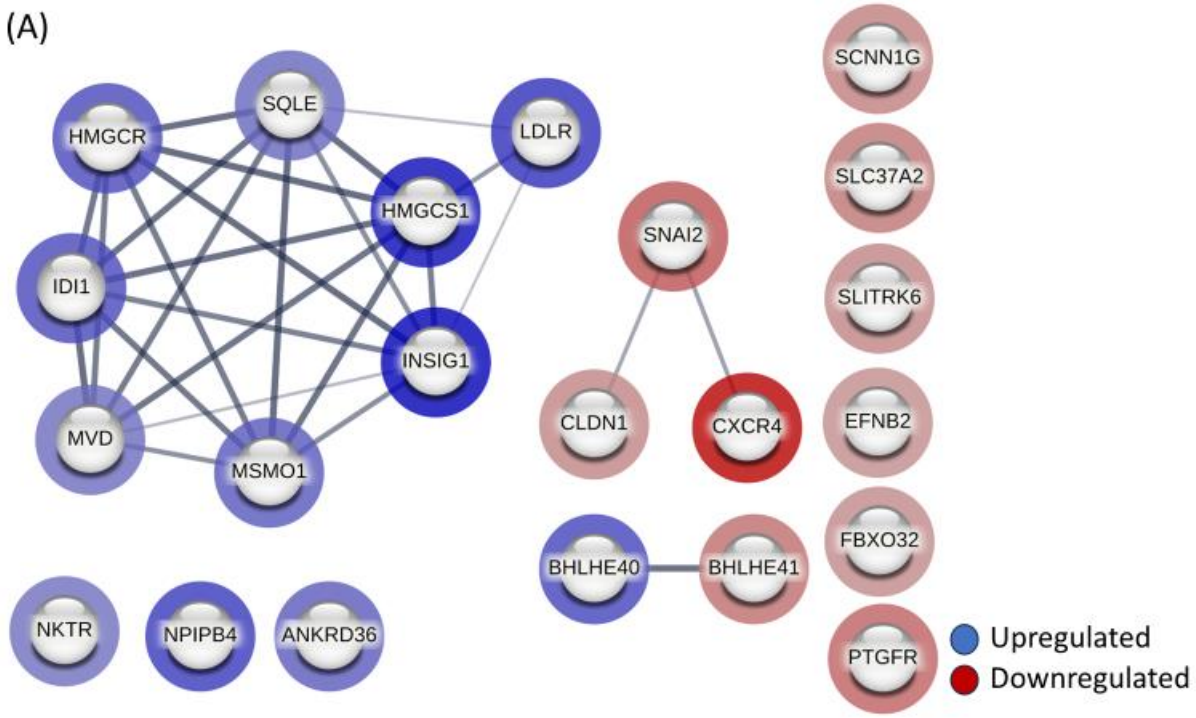
FOLD CHANGE		
Gene	G1	6 days
SLITRK6	0.431533	2.243054
TMPRSS13	1.603314	2.203806
FRMD6	1.737687	2.198639
PDLIM1	1.761569	2.139723
JUN	0.472552	2.125859
C1orf116	1.973585	1.978137
ANXA11	1.882346	1.96285
CRIM1	1.53109	1.890659
SQSTM1	0.628335	1.887984
PRSS22	0.595063	1.881345
CEACAM1	1.940261	1.848751
GDA	1.771927	1.848641
CLMN	1.593763	1.818489
BTG2	2.000537	1.814693
TNFRSF21	2.024146	1.814344
DGKA	1.618963	1.800149
RNF24	1.619209	1.79106
KRT15	2.255904	1.776814
PEX13	0.659283	1.772086
SLIT2	1.828577	1.729109
AMOTL2	2.343687	1.711729
JUNB	0.666252	1.696906
ARHGAP25	1.658854	1.6892
DENND11	0.663975	1.687105
NCCRP1	1.93071	1.679393
CLSTN1	1.531273	1.673955
IER5	1.606373	1.654114
GABARAPL	0.645526	1.610896
CDC42EP1	1.556666	1.610882
FGF2	1.847345	1.610183
PHLDB2	1.56769	1.588898
PADI2	1.70669	1.588288
VSIR	1.709591	1.587925
PMAIP1	0.467054	1.582744
MIR205HG	1.506437	1.55997
SMIM14	0.385127	1.556777
EMP1	1.782143	1.547515
RMND5B	1.830187	1.531814
TGFBR1	0.652269	1.517783
LITAF	0.590633	1.515413
ERI1	0.634165	0.662179
KLHDC4	1.651579	0.624847
TEX2	1.520325	0.576053

Table 1: List of overlapping differentially expressed genes between G1 phase (PRO-seq) and 48 hours and 6days of dTAG1 treatment along with their fold change values.

Of the 442 differentially initiated genes (Figure B), 32 continue to be expressed at 48 hours, and 153 (34.61%) are expressed at 6 days of dTAG1 treatment. From the genes that are shared between G1 phase of the cell cycle and 6 days, 105 gene maintain the directionality of their expression, and 47 switch their direction of expression.

Similarly, 45 nascent transcripts expressed in asynchronous cells that were depleted of RUNX1 have reached steady state expression at 48 hours, and 106 overlap with the 6-days timepoint.

This overlap is reflected in the gene set enrichment analysis. Pathways that were statistically significant in the early stages of the G1 are still enriched after the prolonged absence of RUNX1. Some of these common pathways are involved in adult tissue stemness, TGFβ signaling, estradiol response, and regulating targets of ESR1, CDH1 and EZH2. Thus, these shared genes play a key role in regulating the genotypic and phenotypic transformation of MCF10A upon the loss of RUNX1.



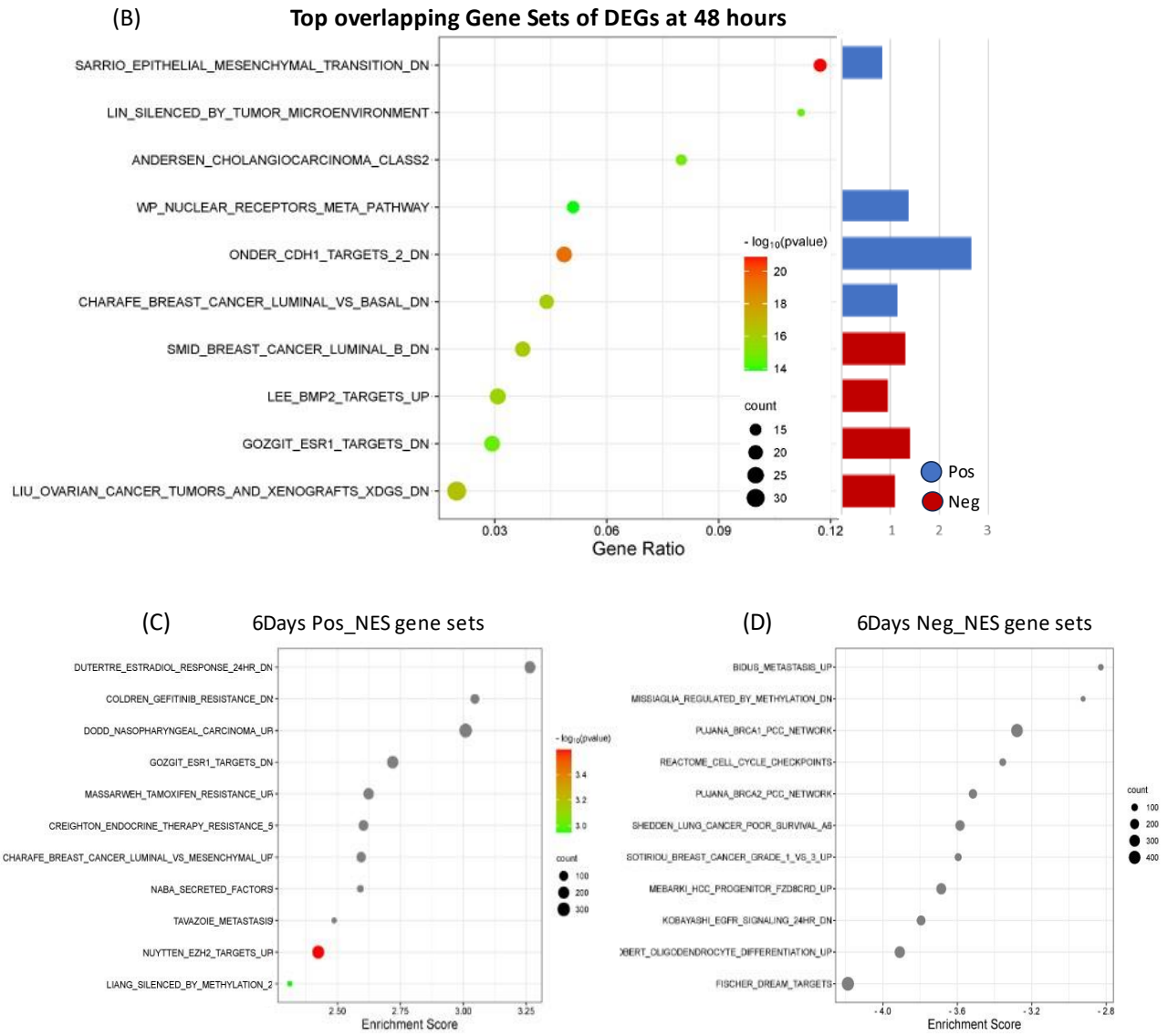


Figure 12: Pathway analysis of the differentially expressed genes based on RNA-seq results post RUNX1 protein degradation across the 3 hours, 48 hours, and 6 days of dTAG1 treatment. *(A) Protein-Protein interaction network of differentially expressed gene after 3 hours of dTAG1 treatment. (B) Pathways with the most statistically significant overlap of gene sets based on the gene ratios using an FDR q-value cutoff of 0.05 and p-value cutoff of 0.001 after 48 hours of dTAG1 treatment. (C) Statistically significant overlapping gene sets with a positive normalized enrichment score after 6 days of dTAG1 treatment. (D) Statistically significant overlapping gene sets with a negative normalized enrichment score after 6 days of dTAG1 treatment.*

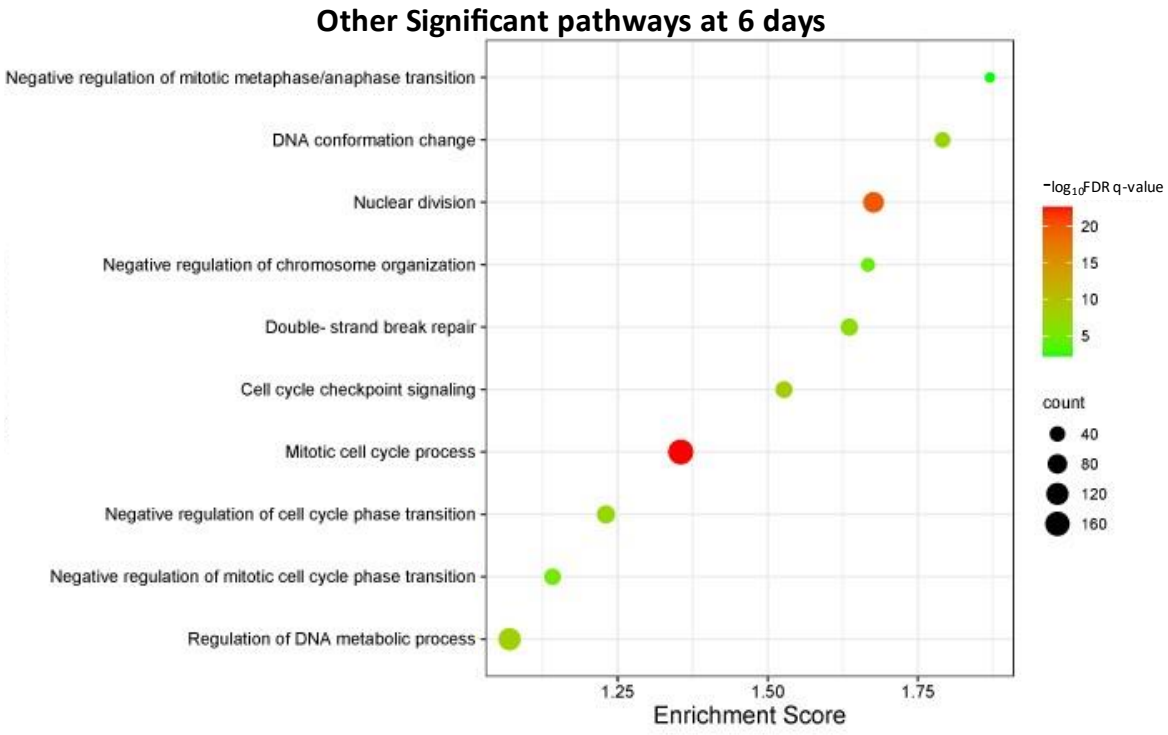


Figure 13: String database analysis of the differentially expressed genes based on the RNA-seq result after RUNX1 loss for 6 days. *Pathways were picked based on a statistically significant Normalized Enrichment score with FDR q-value cutoff of <0.05 and p-value of <0.001.*

After 3 hours of dTAG1 treatment, 64.3% of DEGs are upregulated and 35.7% are downregulated. Short-term loss of RUNX1 results in an upregulation of cholesterol biosynthesis pathways, which is not only required for G1 to S phase transition[41] but also plays an important role in maintaining tumor stem cells[42]. Increased cholesterol metabolism is necessary for rapid cell growth and tumorigenesis by activating downstream signaling of hedgehog, Notch, and receptor tyrosine kinases pathways. Dysregulated expression of tumor promoters and suppressors can be detected as early as 3 hours of RUNX1 depletion. Long noncoding RNAs (lncRNAs), NEAT1 and MALAT1 (also known as NEAT2), are overexpressed in various types of cancer[43,44]. NEAT1 is a target of BRCA1 and is also known to promote EMT, cellular proliferation and resistance to apoptosis[44]. MALAT1 regulates cell proliferation and metastasis in a context specific manner. Elevated levels of MALAT1 are linked with

decreased disease specific survival in ER-negative, lymph node negative, HER2 overexpressing subtype and triple negative subtype of breast cancer[45].

Loss of RUNX1 activates tumor promoter BHLHE40 and inhibits tumor suppressor BHLHE41. In fact, BHLH41 is also downregulated in all phases of the cell cycle. Thus, early stage of RUNX1 degradation is characterized by the loss of the tumor suppressive role of RUNX1 and increased transcriptional activation, which can be corelated to the overall genomic and epigenomic changes observed when RUNX1 is eliminated during mitosis.

Several of the GSEA identified gene sets associated with cell cycle expression data are also enriched at 48 hours of dTAG1 treatment. These include the target genes regulated by CDH1, EZH2 and ESR1, and TGFβ signaling, However, the pattern of gene expression is associated with both luminal and basal subtype of breast cancer. Targets of bone morphogenetic protein 2 (BMP2) are dysregulated upon the loss of RUNX1. BMP2 belongs to the TGFβ super family and plays an important role in cellular differentiation. This may be the first indication that the dTAG1 treated cells are no longer maintaining their normal mammary epithelial phenotype and have gained the ability to differentiate into both basal and luminal cells. This is supported by the morphological changes that are visible on phase contrast microscopy.

By 6 days of dTAG1 treatment, the gene expression pattern of the cells is similar to the luminal subtype of breast cancer, indicating that the cells have transformed from a basal/epithelial genotype to a luminal/mesenchymal genotype. GSEA also indicates that the cells have developed a resistance to estradiol and hormone therapy. Metastasis gene sets are both positively and negatively enriched, which would suggest that there are heterogenous populations of cells with more metastatic potential than others. Other pathways indicate that loss of RUNX1 causes a conformational change

in the DNA and chromosomal structure, and organization. RUNX1 also appears to be necessary for cell cycle progression and cell cycle check point signaling, and its loss results in metaphase/anaphase transition arrest.

Most importantly, loss of RUNX1 downregulates the expression of BRCA1 and BRCA2, and their downstream targets. Coldren Gefitinib Resistance Down gene set is positively enriched, which is related to the gene expression pattern in involved in HER pathway signaling. Thus, loss of RUNX1 results in slowly proliferating, tamoxifen resistant, HER2/HER3-positive, luminal subtype of breast cells that may have metastatic potential and bipotent stem-like properties. These effects may be attributed to the disruption of mitotic gene bookmarking of RUNX1 in MCF10A cells.

3.5. MCF10A cells undergo partial EMT upon the loss of RUNX1.

To verify our findings, we performed western blot analysis to detect markers of epithelial to mesenchymal transformation, both in the cell cycle and at long term timepoints.

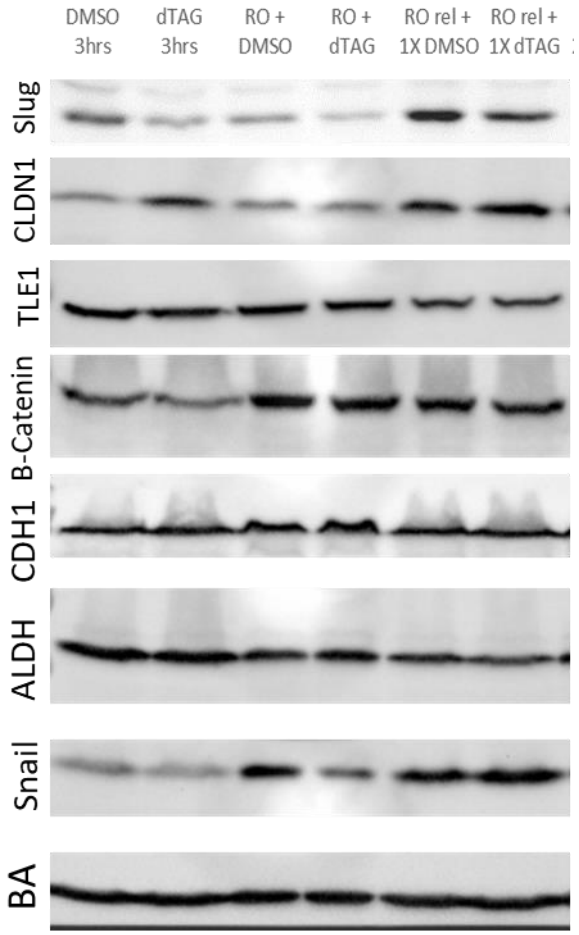


Figure 14: Western blots analysis of markers of EMT across the cell cycle after 3 hours of treatment with either DMSO or dTAG1.

At the protein level, most of the markers of EMT do not change across the cell cycle. Snail and Slug, also known as SNAI1 and SNAI2 respectively, belong to the zinc finger transcription factor family and participate in events associated with epithelial mesenchymal transition (EMT). Slug protein appears to be slightly decreased across the phases of the cell cycle after the loss of RUNX1. Snail is decreased in G2/M but

otherwise remains unchanged. Elimination of RUNX1 for 3 hours increases the expression of Claudin1 protein in asynchronous cells. However, its nascent transcripts and steady state mRNA are downregulated. In the absence of RUNX1, the amount of CLDN1 and SNAI1 proteins continue to drop over the period of 6 days. CDH1 is elevated after 3 hours of RUNX1 loss and then comes back down to baseline.

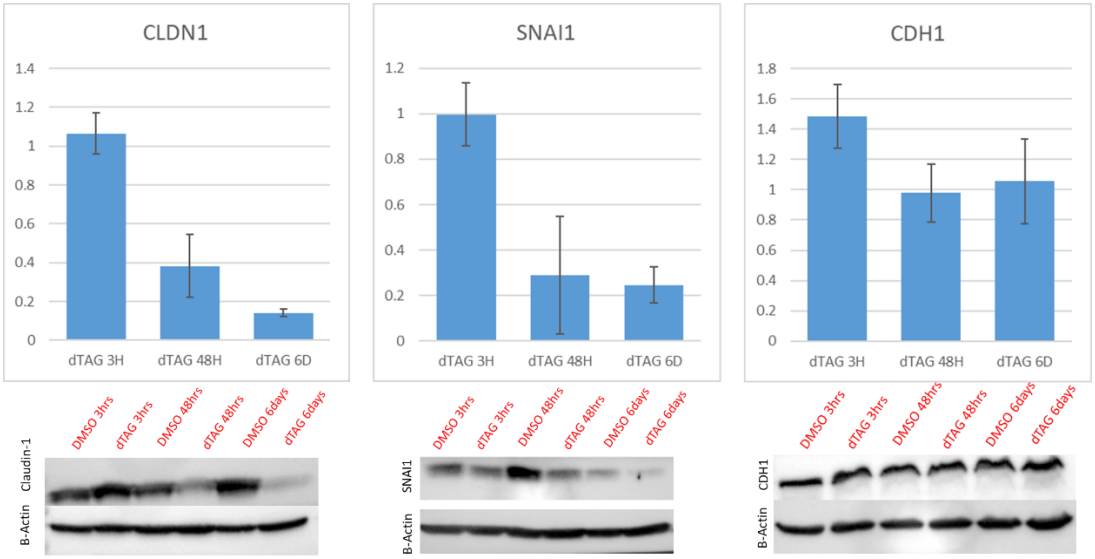


Figure 15: Western blot analysis of CLDN1, SNAI1 and CDH1 after 3, 48 and 144 hours of treatment with either DMSO or dTAG1

Because SNAI1 and SNAI2 are down regulated, their suppressive action on CDH1 is lost, resulting in cells that retain their epithelial features. However, the decreased expression of Claudin-1 (CLDN1), a major component of tight junction complexes, results in loss of cell polarity, cellular framework and cell to cell communication, all of which leads to EMT and tumorigenesis[46]. In addition, mRNA expression of CLDN4 is upregulated upon the loss of RUNX1. MCF7 cells overexpressing CLDN4 have an increased rate of proliferation and resistance to apoptosis[47].

While the level of CDH1 may not be elevated, the loss of its usual suppressors may explain the reason for the activation of its downstream targets. Taking a deeper look at some of the individual differentially expressed genes may reveal the underlying

mechanisms of genetic and epigenetic remodeling taking place after the loss of RUNX1.

T-box 3 gene (TBX3), a transcription factor with a highly conserved DNA binding domain called T-Box, is known to activate and/or repress its target genes through direct DNA binding or by recruiting other co-factors[48]. Depletion of TBX3 in early stages of embryogenesis inhibits cell differentiation and organogenesis. TBX3 mRNA and nascent transcripts are down regulated in G2/M post RUNX1 loss. Though very little is known about the role of TBX3 in mitosis, several T-Box factors have been implicated as tumor promoters and/or suppressors[49]. TBX3 is necessary for the expression of Inhibitor of DNA binding 1 (ID1)[50], a transcription regulator that inhibits binding of basic helix-loop-helix transcription factors to the DNA[51]. Thus, transcription factors that are usually inhibited from mitotic bookmarking may now bind to the chromatin and contribute to the altered pattern of gene expression as the cells emerge into G1. Additionally, previous studies conducted by Lee et al. shows that ID1 can interact with the basic helix-loop-helix L zipper transcription factor, MITF[52], which binds and activates E-cadherin[53]. This, in turn, may result in upregulation of downstream targets of E-cadherin (CDH1).

Epithelial cell adhesion molecule, E-cadherin, is a single-span transmembrane glycoprotein that promotes cell-cell adhesion through interactions with adjacent E-cadherin molecules of neighboring cells. While E-cadherin itself is not differentially expressed, 46 differentially initiated genes in G1 are positively associated (NES: 2.8653083, p-value:0, FDR q-value:0) with the Onder CDH1 Targets 2 DN gene set. The number of enriched genes in this gene set goes down to 23 genes at 48 hours of dTAG1 treatment and then rises to 173 genes at 6 days.

SYMBOL	RANK IN GENE LIST	RANK METRIC SCORE	FOLD CHANGE	CORE ENRICHMENT	SYMBOL	RANK IN GENE LIST	RANK METRIC SCORE	FOLD CHANGE	CORE ENRICHMENT
CXCL1	1	1.982	3.9509542	Yes	ITGB4	79	0.8	1.740722	Yes
PLAU	2	1.712	3.276027	Yes	AREG	85	0.783	1.7208835	Yes
IL1RN	8	1.244	2.3692038	Yes	SLPI	108	0.743	1.6734209	No
SCEL	10	1.227	2.3409821	Yes	FLRT3	118	0.718	1.6445788	No
KRT16	12	1.181	2.2670615	Yes	KIAA0040	119	0.717	1.6438427	No
KRT15	13	1.174	2.2559037	Yes	KLK10	137	0.678	1.5995086	No
PTGES	15	1.147	2.2146007	Yes	ITGA6	138	0.675	1.5970314	No
NCF2	17	1.079	2.1125575	Yes	SDC1	140	0.674	1.5957248	No
TFCP2L1	18	1.064	2.090199	Yes	KRT17	159	0.644	1.5629977	No
SERPINB7	24	1.04	2.0556478	Yes	PPL	169	0.637	1.5551513	No
PLK2	25	1.039	2.0545517	Yes	SYBU	190	0.609	1.5252682	No
TNFRSF21	28	1.017	2.0241458	Yes	SCNN1A	191	0.609	1.5249736	No
LAMC2	29	1.015	2.0203635	Yes	FERMT1	205	0.583	1.4982428	No
ALDH1A3	33	0.991	1.9876485	Yes	COBLL1	206	0.583	1.498049	No
C1orf116	34	0.981	1.9735852	Yes	CYP27B1	236	-0.608	0.6560593	No
TIMP3	35	0.977	1.9683845	Yes	PTHLH	248	-0.623	0.6491892	No
CEACAM1	40	0.956	1.9402609	Yes	RASSF9	289	-0.709	0.6115543	No
HS3ST2	50	0.901	1.8677437	Yes	LYPD3	301	-0.729	0.6035103	No
EDN1	55	0.89	1.8530691	Yes	CCNA1	349	-0.9	0.535909	No
HS3ST1	67	0.846	1.7972319	Yes	FHOD3	366	-0.983	0.5057942	No
LAMB3	73	0.82	1.7652363	Yes	CAMK2N1	373	-1.033	0.4887574	No
FGFBP1	76	0.816	1.7607647	Yes	POPDC3	386	-1.153	0.4496707	No
NDRG1	78	0.81	1.7533624	Yes	FOS	411	-1.742	0.298869	No

Table 2: Differentially initiated genes from G1_dTAG1 that are positively enriched with ONDER_CDH1_TARGETS_DN

From the DEGs within this gene set, 3 Type I cytokeratins (KRT15, KRT16, KRT17) are upregulated at 6 days, indicating that the cells are acquiring epithelial characteristics. While KRT15 acts as a tumor promoter in other types of cancer, KRT15 expression in breast cancers is usually lower compared to normal breast tissue. High KRT15 expression is associated with tumor suppressive pathways and better overall survival in patients with breast cancer[54]. On the other hand, KRT17 plays a dual role depending on the type of breast tumor. It is a marker of triple-negative breast cancers[55], and overexpression of KRT17 is significantly associated with poor prognosis in ER−/HER2− type of breast cancer[56]. However, high KRT17 expression in HER2+ breast cancer predicts a favorable prognosis[57]. Overexpression of high molecular weight cytokeratins is associated with basal-like features, whereas the downregulation of KRT5 and KRT6, as seen with the loss of RUNX1, is widely

accepted as a non-basal subclass of breast cancer. Combined with the overexpression of HER2, the RUNX1 depleted cells may be classified as a luminal subtype with basal-like features, implying that the cells have undergone partial EMT or that the loss of RUNX1 gives rise to a small population of stem cells which later differentiate into a heterogenous population, containing cells from the different subtypes of breast cancer.

While some of the genes may confer an epithelial phenotype, other genes within this pathway that were upregulated upon the loss of RUNX1 tend to drive the cells towards EMT. PLAU overexpression predicts a poorer clinical outcome in head and neck squamous cell carcinoma. A study by Chen et al. indicates that PLAU and its associated genes ITGA5, SERPINE1, and TNFRSF12A may be linked with epithelial-to-mesenchymal transition (EMT) process and tumor migration. Particularly, the downregulation of PLAU decreases the expression of TNFRSF12A, leading to inhibition of cell proliferation and migration[58]. Moreover, PLAU mRNA expression is a powerful prognostic marker in HER2-positive tumors[59].

Similarly, EZH2 is a Polycomb group (PcG) protein that is involved in gene silencing through trimethylation of H3K27 and is linked with epithelial to mesenchymal transition (EMT)[60]. Nuytten et al. conducted a study on PC3 cells (prostate cancer) to examine the effects of RNAi mediated knockdown of EZH2 and mapped out a set of genes that were subsequently upregulated[61]. ChIP-seq analysis in prostate cancer cell lines revealed that the enhancer of EZH2 binds to the promoter of RUNX1[62]. Several differentially initiated genes from the cell cycle analysis and DEGs from 6days are positively enriched with this gene set, implying that RUNX1 may indirectly alter histone methylation through EZH2 and that the histone methylation marks that were

altered during G2/M are retained as the cells progress into early G1. This is also supported by the evidence of downregulation of EZH2 mRNA at 6 days.
Additionally, we found enrichment for transcripts related to the Targets of SMAD2 or SMAD3. The SMAD2 and SMAD3 proteins are involved in the intracellular transmission of transforming growth factor β (TGF-β) signaling from the cell membrane to the nucleus and regulating the transcription of target genes. PLAU, LAMC2, and LAMB3 are strongly linked with TGFβ signaling, particularly LAMC2. Inhibiting TGFβ signaling using Vactosertib in LAMC2-expressing pancreatic cells suppress CSC-related metastasis in Pancreatic ductal adenocarcinoma patients[63]. Thus, TGFβ signaling may be a contributing factor for the transformation of dTAG1 treated cells.

3.6. Loss of RUNX1 confers stem-like properties to MCF10A.

A critical step in tumor initiation is the transformation of cells from epithelial to mesenchymal-like phenotype. Cancer stem cells (CSC) are the driving force behind the capacity for tumor cells to become invasive, metastatic, and resistant to therapy. Based on the variable expression pattern showcasing both basal/epithelial and luminal/mesenchymal characteristics, we hypothesize that the loss of RUNX1 gives rise to a small population of stem cells that later differentiate into the different subclasses of breast cancer cells.
CD44, CD24 and ALDH1 are the most frequently used markers to identify CSCs. Cells with high CD44 ($CD44^{high}$) but low CD24 expression ($CD24^{-/low}$) exhibit enhanced cell proliferation and tumorigenesis. High ALDH1 activity ($ALDH^{High}$) is linked with cancer metastasis. Moreover, ALDH1A3 gene expression is upregulated

upon the loss of RUNX1. This isoform is considered to be primarily responsible for the ALDH activity in breast cancer cells[64].

In order to validate our hypothesis and the sequencing results, we performed the AldeRed ALDH Detection Assay. AldeRed reagent AldeRed 588-A diffuses freely into intact live cells where it is oxidized by intracellular ALDH enzymes into its corresponding acid causing the cells to become brightly fluorescent. The amount of fluorescence produced depends on the ALDH activity within the cell. Verapamil, an efflux inhibitor, traps the AldeRed reaction product inside the cell and prevents the loss of fluorescence. The Aldered Assay kit is provided with an ALDH inhibitor, diethylamino benzaldehyde (DEAB) as a negative control. All test samples were compared against their corresponding DEAB-treated control sample to eliminate background fluorescence.

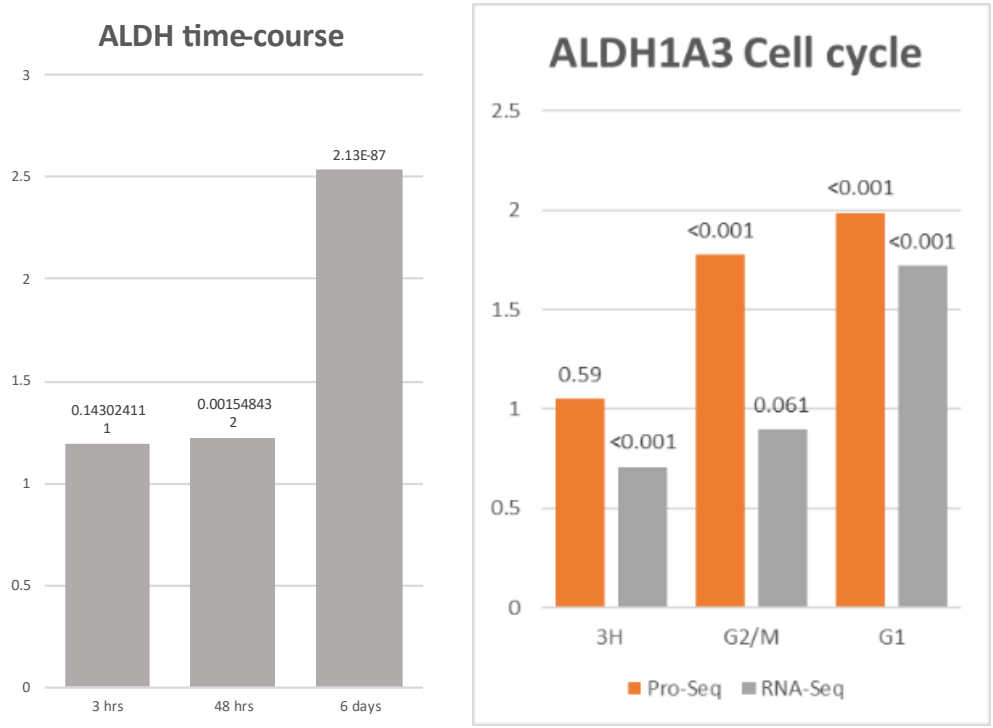


Figure 16: ALDH1A3 expression in dTAG1 treated cells along with p-values. *The graph on the left shows that ALDH1A3 mRNA is upregulated after 6days of dTAG1 treatment. The graph on the right shows that ALDH1A3 is elevated in G2/M and G1 phase of the cell cycle after the loss of RUNX1.*

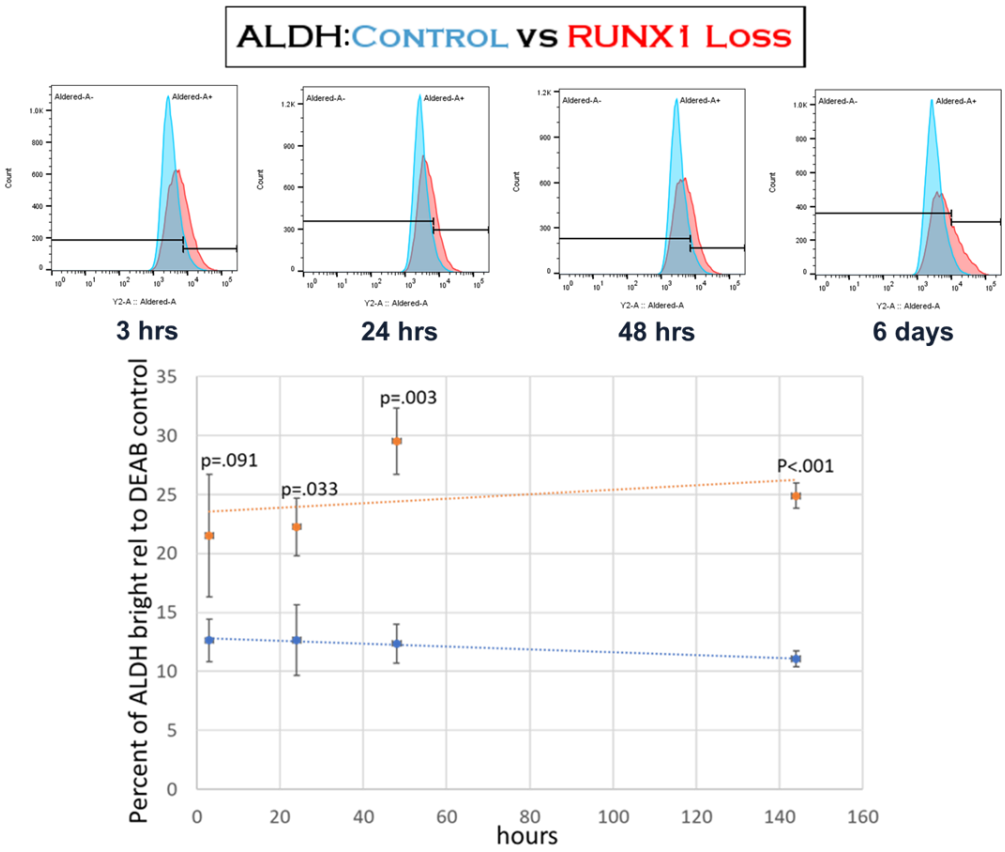


Figure 17: Flow cytometry analysis of ALDH enzyme detection assay. *(A) Comparative Flow cytometry analysis of ALDH activity in DMSO and dTAG1 treated samples at different timepoints (B) Percentage of ALDH bright cells relative to DEAB-treated control*

Flow cytometry analysis shows that MCF10A cells with dTAG1 treatment and subsequent RUNX1 loss have increased ALDH activity. The ALDH activity progressively increases over the course of 6 days.

Double staining of CD44 and CD24 in MCF10A Degron Cells showed that 99.86% of both DMSO and dTAG1 treated cells are CD44+. Therefore, we focused on the CD24 expression in the RUNX1 depleted cells at various timepoints.

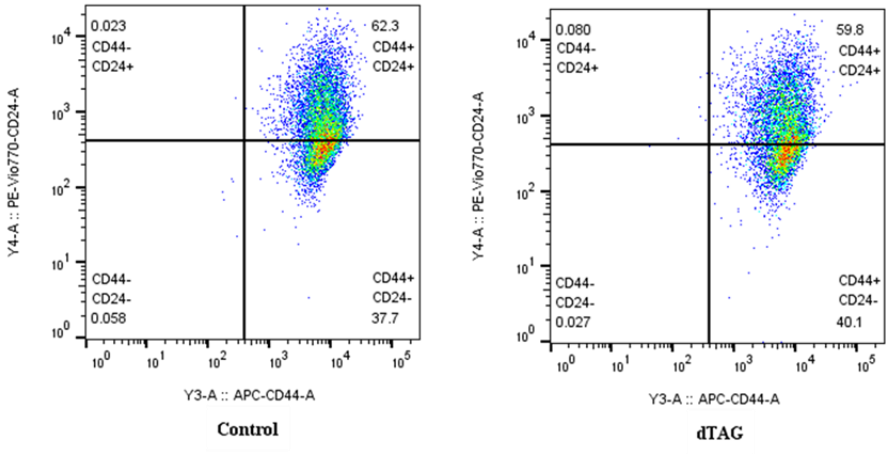


Figure 18: Flow cytometry analysis for double staining of CD44 and CD24 in MCF10A Degron Cells. *99.86% of both DMSO and dTAG1 treated cells are CD44+.*

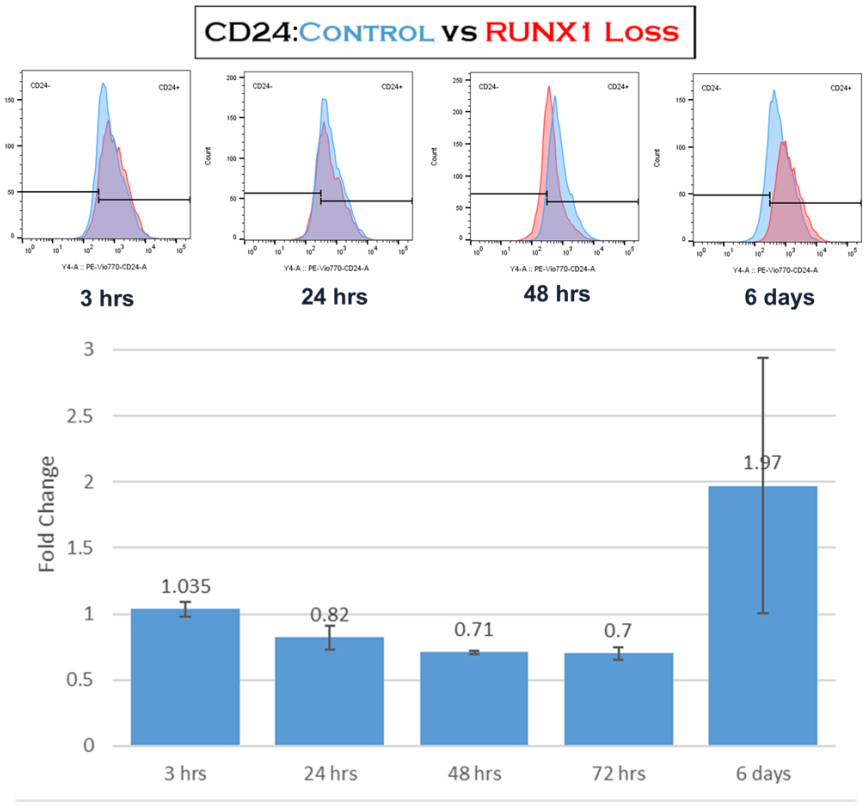


Figure 19: Flow cytometry analysis for CD24 expression. *(A) Comparative Flow cytometry analysis of CD24 expression in DMSO and dTAG1 treated cells at different timepoints (B) Ratio of dTAG1/DMSO calculated based on percentage CD24+ cells*

At 3 hours, there is no significant change in CD24 expression upon the loss of

RUNX1. As treatment time increases (up to 72 hours), there is a gradual decrease in

the $CD24^{+}$ subpopulation. This is similar to the $CD44^{+}$ / $CD24^{-/low}$ expression pattern found in basal mesenchymal, triple negative breast cancer subtype. However, after 6 days of treatment, there is nearly a 2-fold increase in $CD24^{+}$ subpopulation upon RUNX1 loss, which is similar to the expression pattern of $CD44^{+}$ / $CD24^{+}$ seen in basal epithelial type of breast tumors[65].

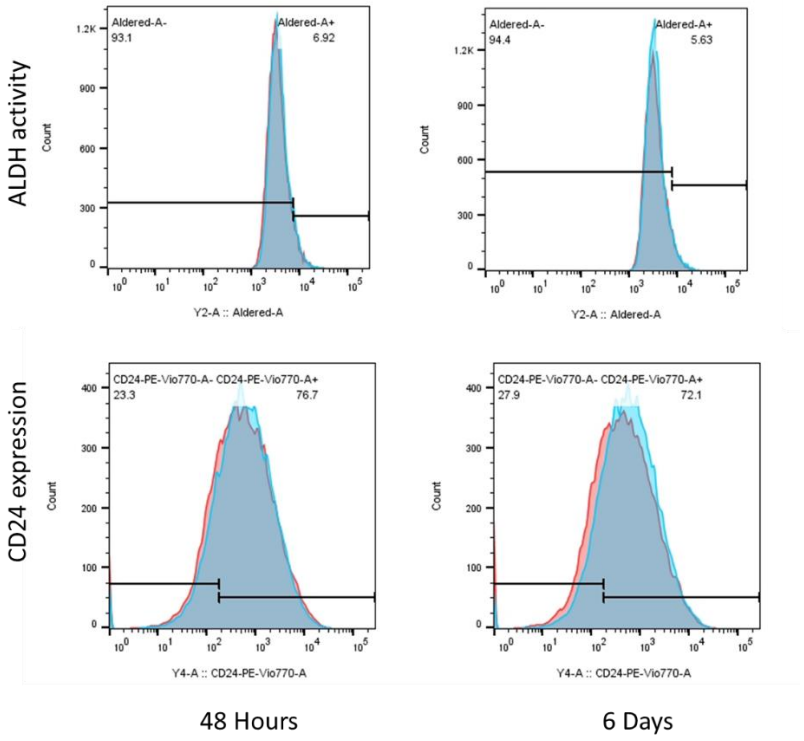


Figure 20: Flow cytometry analysis for ALDH activity and CD24 expression in parental MCF10A cells treated with DMSO or dTAG1 for 48 hour and 6 days.

Performing the same ALDH/CD24/CD44 assay on parental MCF10A at 48 hours and 6 days of dTAG1 treatment showed that there is no difference between in the ALDH activity and CD24 expression between the dTAG1 treated and DMSO control samples, confirming that the changes observed in the MCF10A degron cells are an effect of the loss of RUNX1 and not an off target toxic effect of dTAG1.

Al-Hajj et al. determined that breast cells expressing $CD44^{+/high}/CD24^{-/low}$ markers gave rise to visible tumors within 12 weeks of injection into mammary pads of mice[66]. It is widely accepted that $CD44^{+}$/ $CD24^{-}$ cells exhibit undifferentiated basal/mesenchymal properties. Thus, loss of RUNX1 lead to an increase in the undifferentiated population of cells that can be detected after 2-3 days of dTAG1 treatment. This is also correlated to the simultaneous basal and luminal gene expression pattern identified through GSEA analysis of DEGs at 48 hours. Several studies suggest that $CD44^{+/}CD24^{+}$ expressing cells can readily give rise to a $CD44^{+/}CD24^{-}$ population during tumor initiation and CD24 negative cells are interconverted back to CD24 positive phenotype after metastasis [67]. This subpopulation may not exhibit metastatic properties but may have the ability to colonize and survive in foreign tissue.

3.7. Elimination of RUNX1 Leads to an increase in the mitotic population.

We observed that the rate of cellular proliferation decreases after the loss of RUNX1. This is confirmed by the String database analysis of differentially expressed genes at 6 days. Genes contributing to the negative regulation of the cell cycle and metaphase to anaphase transition are overexpressed at 6 days. Therefore, we performed a flow cytometric analysis of the cells at various timepoints using H3pS28 antibody staining. The serine 28 residue on the N-terminal tail of histone H3 is phosphorylated during prophase to anaphase of mitosis.

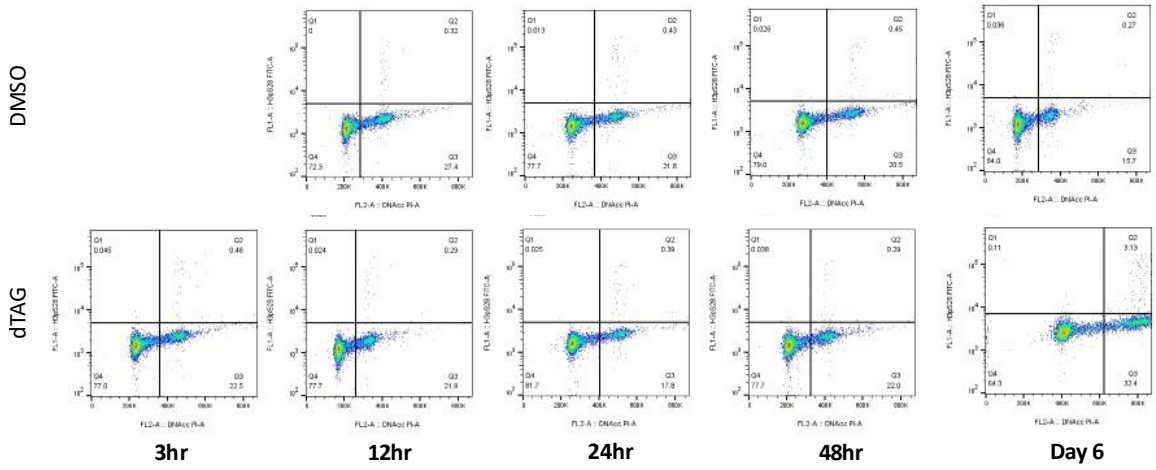


Figure 21: Flow cytometry analysis for H3pS28 antibody in MCF10A degron cells treated with DMSO or dTAG1 at various timepoints. *There is no observable difference between the percentage of mitotic cells at early timepoints. By 6 days there is a 10-fold increase in the mitotic population.*

There is no observable difference between the percentage of mitotic cells in dTAG1-treated and DMSO-control samples at early timepoints (3, 12, 24, 48 hours). However, after prolonged loss of RUNX1 the percentage of mitotic cells increased from 0.27% to 3.13% (more than 10-fold increase), suggesting that prolonged loss of RUNX1 arrests the cells in metaphase/anaphase transition.

3.8. Loss of RUNX1 in MCF10A cells shows oncogenic transformation potential.

Several of the differentially initiated genes overlap with pathways involved in the process of development of other cancers such as prostate, thyroid, and nasopharyngeal carcinoma. Both BRCA1 and BRCA2 and the network of their downstream targets are downregulated after loss of RUNX1. ITGB4 is another gene that is upregulated by the loss of RUNX1. Integrin beta 4 (ITGB4) belongs to the family of heterodimeric transmembrane receptors and plays a key role in tumorigenesis and tumor invasion. Integrins are composed of α and β subunits. The integrin beta 4 subunit tends to associate with alpha 6 subunit. Both the ITGA6 and ITGB4 genes are

regulated by RUNX1 in myeloid cells. RUNX1 binds to ITGA6 through a RUNX1 binding motif in its promoter. Whereas, RUNX1 indirectly activates ITGB4 by interactions with upstream regulatory elements[68]. Previous studies show that ITGB4 is upregulated in various types of tumors [69] and is linked to poor prognosis or aggressive behavior[70,71]. It is also associated with a worse 5-year probability of relapse-free survival in patients with triple-negative breast cancer who received chemotherapy[72].

Integrins act as receptor for laminins like LAMB3 and LAMC2, both of which are upregulated in G1. Additionally, the overexpression of ITGB4 and LAMC2 promotes epithelial–mesenchymal transition [73–75]. High LAMC2 may be associated with stemness, tumor initiation and self-renewal, and plays a critical role in tumor migration and invasion[63]. LAMB3 is also involved in invasion and metastasis[76]. Gene set enrichment analysis also shows that LAMC2, LAMB3, ITGB4, ITGA6, ITGA2B, LAMA5, THBS1, and DDIT4 are major contributors to the activation of PI3K/Akt signaling pathway. As a result, the expression of some downstream targets of the Akt pathway, such as BCL2L1, is upregulated. Aberrant activation of this pathway promotes survival, antiapoptotic signaling and proliferation of tumor cells in many types of cancers.

Other genes that are differentially expressed by the loss of RUNX1 include BCL6, BHLHE40, and NKX3-1. Miyazaki et al. conducted IP experiments in COS7 cells and found that BHLHE40 is bound to RUNX1 in vitro[77]. Class E basic helix-loop-helix protein 40, BHLHE40, is a pro-metastasis transcription factor in breast carcinoma and downstream target of TGFβ signaling[78], which promotes tumor cell survival and migration[79].

Transcriptional repressor BCL6 is highly expressed in breast cancer cell lines and is expressed in 68% of histologically high-grade ductal breast carcinomas[80]. It is upregulated in all phases of the cell cycle and at 6 days. NKX3-1 is a transcription factor and a prostatic tumor suppressor gene[81]. Loss of 1 or both alleles of NKX3-1 leads to the development of preinvasive lesions in mice[82]. IGFBP5, SERPINB2, PTGS2, SOX4, CD24, LAMA4, HPGD, COL1A1, and TSPAN12 were found to be positively enriched in the Tavazoie Metastasis gene set, which defines a group of genes associated with risk of distal metastasis. Downstream targets of SOX4, an oncogene overexpressed in prostate tumor, are also elevated.

MCF10A cells that have been depleted of RUNX1 begin to express ERBB2 and ERBB3 (also known as HER2 and HER3) by 6 days of dTAG1 treatment. ERBB3 is a member of the ERBB protein-tyrosine kinase receptors and usually heterodimerizes with ERBB2, forming a prooncogenic signaling complex. Activation of ERBB2/ERBB3 dimer by ligand binding (neuregulins, NRG1 and NRG2), strongly triggers PI3K/Akt signaling, which promotes tumor progression, invasion, and metastasis[83] as well as tamoxifen resistance[84]. Over expression of ERBB2 is also associated with increased signaling of MAPK, ERK and Jak/Stat signaling[85].

Loss of RUNX1 alters Estrogen response and ESR1 target expression.

By genome-wide ChIP-seq analysis of AR-binding sites, RUNX1 was identified as a target gene for Androgen receptor signaling[62]. RUNX1 also has a binding motif in the ESR1 control region upstream of its transcription start site[39]. Gene sets involved in estrogen mediated signaling and estrogen dependent gene expression are enriched in early G1. By 6 days of dTAG1 treatment, targets of ESR1 (estrogen receptor 1 or Erα) are upregulated and cells become resistant to estradiol and hormone therapy.

Thus, RUNX1 plays a vital role in the regulation of estrogen receptor target genes and hormone response even in ER-negative cell lines.

Loss of RUNX1 alters the epigenetic landscape of MCF10A cells.

RUNX1 degradation upregulates targets of PHF8, a histone demethylase that removes the transcriptional repressive marks of histone H3 dimethyl lysine 9 (H3K9me2). PHF8 also acts as a transcriptional coactivator by binding to H3K4me3 and RNA polymerase II [86]. As discussed earlier, loss of RUNX1 downregulates transcriptional repressor EZH2, thus inhibiting H3K27me3 methylation. RUNX1 degradation is also associated with decreased expression Protein arginine N-methyltransferase 1 (PRMT1), which is responsible for methylating arginine residues of histone (specifically H4R3me2a) and non-histone proteins. It is known to stabilize EZH2[87], thus, the downregulation of PRMT1 may also exacerbate the effects of EZH2 suppression.

Histone-lysine N-methyltransferase, Suv39H1, is found to be downregulated after 6 days of dTAG1 treatment. It functions as a histone methyltransferase responsible for H3K9me3 methylation. It also acts as a cofactor with SNAI1, necessary for the repression of CDH1[75].

Histone deacetylases, HDAC5, HDAC9, and HDAC11, are upregulated after 6 days of dTAG1 treatment and subsequent loss of RUNX1. Nascent transcripts of HDAC3 and HDAC7 are also elevated in asynchronous cells and cells synchronized in G1, respectively. Histone acetyl transferase, HAT1, is downregulated by 6 days. Elimination of RUNX1 may result in chromatin condensation and decreased accessibility to transcription factors. TET2 and TET3 are also upregulated at 6 days, suggesting that loss of RUNX1 results in hypermethylation of DNA.

CHAPTER 4: DISCUSSION

Derangement of epigenetic processes play a pivotal role in the initiation and advancement of numerous cancers. These deviations often act as key drivers of resistance to therapy and the sustenance of cancer cell survival. The phenotypic transformations observed during tumor initiation and progression stem from shifts in gene expression patterns, regulated by the inherent regulatory states encoded within the nucleoprotein structure of chromatin. Changes in chromatin structure that result in varying accessibility to transcription factor binding have been recognized as among the most pertinent genomic features associated with the regulation of epithelial to mesenchymal transformation (EMT).

EMT is a reversible and plastic process that has also been involved in generation of cancer stem cells (CSCs)[88]. Understanding the mechanisms behind the capacity for differentiated cells to transform into a more progenitor-like state that can later produce other cell types is essential for developing preventative and therapeutic approaches during the onset and progression of breast cancer.

Breast cancer is a highly heterogeneous disease, reflected by vastly different disease manifestations among different patients as well as intratumor heterogeneity within each individual. Carcinoma cells exist in both epithelial and mesenchymal cell states concurrently. Conversions between these states have been shown to occur as a result of constantly changing combination of intrinsic and extrinsic cell-biological programs that give rise to epithelial to mesenchymal (EMT) and mesenchymal-to-epithelial (MET) transitions. Activation of an EMT program drives both normal and cancerous mammary epithelial cells into stem-like states[89]

In the current study, we investigated the role of RUNX1 during mitosis and its function in transmitting lineage specific transcriptional control, imparting phenotypic memory and determination of cell fate.

Loss of RUNX1 resulted in a morphological shift. RUNX1 depleted cells presented a mesenchymal-like morphology with elongated spindle-like cells compared with RUNX1 expressing MCF10A cells that exhibit typical cobblestone-like epithelial features. This change was accompanied by an alteration in the pattern of gene expression. Transcriptomic analysis revealed the existence of approximately 3500 differentially expressed genes (DEGs) between MCF10A degron cells treated with dTAG1 and DMSO-control. Analysis of the gene sets and pathway enrichment reinforces the notion that loss of RUNX1 promotes an epithelio-mesenchymal phenotype.

Gene set enrichment analysis of the DEGs reveals that the effects of the loss of RUNX1 can be roughly characterized into 3 phases. An early tumor initiation phase accompanied by increased transcriptional activation, where the tumor suppressive role of RUNX1 is lost. These early changes can be correlated to the loss of RUNX1 epigenetic bookmarking. An intermediate transitional phase where there is an increase in undifferentiated cells that express both basal and luminal genes and a late phase characterized by slowly proliferating, estrogen resistant, HER2/HER3-positive, luminal subtype of breast cells with basal features that may have metastatic potential and bipotent stem-like properties.

To further validate the transcriptomic data, we assessed expression of epithelial and mesenchymal proteins. The partial loss of epithelial phenotype was confirmed by the decreased levels of CLDN1 protein. High SNAI1 expression is usually accompanied

by suppression of CDH1 but in our study SNAI1 protein levels progressively decrease with increased duration of dTAG1 treatment. MDA-MB-231 cells with CRISPR/Cas9 mediated SNAI1 knockout acquired an epithelio-mesenchymal phenotype with stem/progenitor properties and decreased tumor invasion[90]. These cells express Forkhead box A1 (FOXA1), which is also overexpressed in our data set. ChIP analysis in T47D cells determined that RUNX1 has a binding motif in the ESR1 control region, as well as several RUNX-binding motifs in the FOXA1 control region upstream of their transcription start site[39]. FOXA1 is only expressed in luminal subtype of breast cancer cells and is absent in the basal subtype[91]. It functions as an oncogenic transcription factor that can bind condensed chromatin, inducing genome-wide enhancer reprogramming[92]. FOXA1 is a key regulator of estrogen receptor 1 (ESR1 or Erα) function but is also present in ER negative tumors[91]. FOXA1 exerts its ER regulating activity by binding to and activating enhancers enriched in histone H3 lysine 4 mono/di-methylation (H3K4me1/me2)[93]. This mechanism may be responsible for the heightened ESR1 mediated activity seen after 6 days of dTAG1 treatment. While MCF10A cells do not express ER, the loss of RUNX1 can disrupt ESR1 signaling through FOXA1 and produce aggressive, endocrine resistant tumors[92].

Another notable finding is the upregulation of HER2 and HER3 in RUNX1 depleted MCF10A degron cells. This is accompanied by the downregulation of CK5/6. Evidence suggests that FOXA1 is capable of directly modulating HER2 and HER3 gene transcription, via Junctional Adhesion Molecule-A (JAM-A) activation. JAM-A interacts with FOXA1, translocating it to the nucleus, where it binds to the promoter region of HER2[94]. Overexpression of FOXA1 and HER2 is strongly associated with ER negative breast tumors[95] and overall poor patient survival[94].

Previous research in transformed human mammary epithelial cells showed increased tumor-initiating ability through EMT[96,97]. Both BRCA1 and BRCA2 are downregulated after loss of RUNX1. Mutations in BRCA1 and 2 tumor suppressive genes are associated with about 20% of familial breast cancer[98]. BRCA genes play a crucial role in maintaining genomic stability, regulating genetic and epigenetic transcriptional machinery, and regulating cell cycle progression. Results from our study suggest that loss of RUNX1 may directly or indirectly affect the transcription of BRCA genes. Another hallmark of tumor cells is their ability to resist apoptosis. MCF10A cells treated with dTAG1 were able to survive for longer in the absence of growth factors.

A critical step in tumor initiation is for the cells to acquire the ability of self-renewal and differentiation. Cancer stem cells (CSC) are the driving force behind the capacity for tumor cells to become invasive, metastatic, and resistant to therapy. Numerous cell-surface and intracellular markers have been used for the identification of different types of cancer stem cells in breast cancer. Notably, depletion of CK18 is known to promote partial EMT and confers stemness in breast cancer MCF-7 cells through the activation of the Wnt/β-catenin signaling[99]. CK18 is downregulated after 6 days of dTAG1 treatment. Secreted frizzled related protein 1 (SFRP1), a Wnt antagonist, is also downregulated by RUNX1 loss, thus, further enhancing Wnt mediated EMT and acquisition of stem cell-like properties, including the $CD44^{+/high}/CD24^{-/low}$ signature.

Increased ALDH1 activity has been found to be an effective marker to characterize malignant progenitor breast cells. In accordance with the RNA-sequencing data, the expression of the ALDH enzyme increases over time after the loss of RUNX1. CD44, a hyaluronic acid receptor, is critical in regulating cell proliferation, survival, adhesion, and migration. CD44+ cells are considered to be tumorigenic and display

stem cell properties. CD24 is involved in cell adhesion and metastasis. Al-Hajj et al. determined that breast cells expressing $CD44^{+/high}/CD24^{-/low}$ markers gave rise to visible tumors within 12 weeks of injection into mammary pads of mice[66]. In our study, CD44 is highly expressed in both DMSO-control and dTAG1 treated cells. There is an increase in the $CD24^{-/low}$ subpopulation of cells after loss of RUNX1 up to 72 hours of dTAG1 treatment. However, by 6 days of treatment there is a switch in the expression pattern resulting in a nearly two-fold increase in the ratio of dTAG1/DMSO $CD24^{+/high}$ expression. This confirms the existence of a heterogenous population of cells that are expressing both $CD44^{+}/ CD24^{-/low}/ALDH^{High}$ and $CD44^{+}/ CD24^{+}/ALDH^{High}$ phenotype that are associated with basal mesenchymal, triple negative breast cancer subtype and basal epithelial type of breast tumors, respectively[65]. $CD44^{+}/ CD24^{+}$ cells can interconvert between CD24-positive and CD24-negative expression patterns in a context specific manner[67]. The $CD44^{+}/ CD24^{-/low}/ALDH^{High}$ are seen in the intermediate undifferentiated phase of RUNX1 loss. In an experiment conducted by Sokol et al, RUNX1 loss in primary human breast epithelial cells increased the stem cell population by two-fold and prevented cell differentiation, thus establishing RUNX1 as a mammary stem cell regulator[100].

Loss of RUNX1 alters the methylation and acetylation states of histones. During Early G1 phase of the cell cycle there is an increase in transcription activation. This is possibly mediated through 2 factors, PHF8 and EZH2. PHF8 removes the transcriptional repressive mark H3K9me2 and coactivates H3K4me3 and RNA polymerase II [86]. Loss of RUNX1 downregulates transcriptional repressor EZH2, thus inhibiting H3K27me3 methylation. This indicates that RUNX1 normally inhibits the transcription of several genes and corroborates its roles as a tumor suppressor.

In later stages, loss of RUNX1 equally activates and suppresses transcription. Downregulation of PRMT1 inhibits methylation of H4R3me2a. Methylation of Histone H4 is usually associated with an open chromatin state. Downregulation of Suv39H1 inhibits H3K9me3 methylation, thus activating transcription. Upregulation of HDAC5, HDAC9, and HDAC11, with concurrent downregulation of HAT1, results in chromatin condensation and decreased accessibility to transcription factors. TET2 and TET3 are also upregulated at 6 days, suggesting that loss of RUNX1 results in hypermethylation of DNA.

In summary, we found that normal-like human mammary epithelial cells undergo a morphological transformation, exhibit altered cellular proliferation, population heterogeneity and neoplastic properties following the loss of RUNX1. The transformed cells showed stem cell properties and expressed both epithelial and mesenchymal markers, in what is termed as a partial EMT phenotype. Pathway enrichment analysis showed that changes in gene expression seen after disruption of RUNX1 mitotic bookmarking are retained even at later stages bolstering the hypothesis that RUNX1 regulated gene expression and cell fate through epigenetic mechanisms.

REFERENCES

1. Martienssen, R. A., Riggs, A. D., & Russo, V. E. A. (1996). *Epigenetic Mechanisms of Gene Regulation*. Cold Spring Harbor laboratory.
2. Bird, A. (2007). Perceptions of epigenetics. *Nature*, *447*(7143), 396–398. https://doi.org/10.1038/nature05913
3. Wong, A. H. C., Gottesman, I. I., & Petronis, A. (2005). Phenotypic differences in genetically identical organisms: The epigenetic perspective. *Human Molecular Genetics*, *14*(suppl_1), R11–R18. https://doi.org/10.1093/hmg/ddi116
4. Shi, Y., Zhang, H., Huang, S., Yin, L., Wang, F., Luo, P., & Huang, H. (2022). Epigenetic regulation in cardiovascular disease: Mechanisms and advances in clinical trials. *Signal Transduction and Targeted Therapy*, *7*(1), 200. https://doi.org/10.1038/s41392-022-01055-2
5. Landgrave-GÃ³mez, J., Mercado-GÃ³mez, O., & Guevara-GuzmÃ¡n, R. (2015). Epigenetic mechanisms in neurological and neurodegenerative diseases. *Frontiers in Cellular Neuroscience*, *9*. https://doi.org/10.3389/fncel.2015.00058
6. You, J. S., & Jones, P. A. (2012). Cancer genetics and epigenetics: Two sides of the same coin? *Cancer Cell*, *22*(1), 9–20. https://doi.org/10.1016/j.ccr.2012.06.008
7. Fraser, P., & Bickmore, W. (2007). Nuclear organization of the genome and the potential for gene regulation. *Nature*, *447*(7143), 413–417. https://doi.org/10.1038/nature05916
8. Singh, H., Sekinger, E. A., & Gross, D. S. (2000). Chromatin and cancer: Causes and consequences. *Journal of Cellular Biochemistry. Supplement*, *Suppl 35*, 61–68.
9. Nurk, S., Koren, S., Rhie, A., Rautiainen, M., Bzikadze, A. V., Mikheenko, A., Vollger, M. R., Altemose, N., Uralsky, L., Gershman, A., Aganezov, S., Hoyt, S. J., Diekhans, M., Logsdon, G. A., Alonge, M., Antonarakis, S. E., Borchers, M., Bouffard, G. G., Brooks, S. Y., … Phillippy, A. M. (2022). The complete sequence of a human genome. *Science (New York, N.Y.)*, *376*(6588), 44–53. https://doi.org/10.1126/science.abj6987
10. Cooper, G. M. (2000). *The Cell: A Molecular Approach.* (2nd Edition). Sinauer Associates.
11. Zaidi, S. K., Fritz, A. J., Tracy, K. M., Gordon, J. A., Tye, C. E., Boyd, J., Van Wijnen, A. J., Nickerson, J. A., Imbalzano, A. N., Lian, J. B., Stein, J. L., & Stein, G. S. (2018). Nuclear organization mediates cancer-compromised genetic and epigenetic control. *Advances in Biological Regulation*, *69*, 1–10. https://doi.org/10.1016/j.jbior.2018.05.001
12. Fritz, A. J., Gillis, N. E., Gerrard, D. L., Rodriguez, P. D., Hong, D., Rose, J. T., Ghule, P. N., Bolf, E. L., Gordon, J. A., Tye, C. E., Boyd, J. R., Tracy, K. M., Nickerson, J. A., van Wijnen, A. J., Imbalzano, A. N., Heath, J. L., Frietze, S. E., Zaidi, S. K., Carr, F. E., … Stein, G. S. (2019). Higher order genomic organization and epigenetic control maintain cellular identity and prevent breast cancer. *Genes, Chromosomes & Cancer*, *58*(7), 484–499. https://doi.org/10.1002/gcc.22731

13. Davey, C. A., Sargent, D. F., Luger, K., Maeder, A. W., & Richmond, T. J. (2002). Solvent mediated interactions in the structure of the nucleosome core particle at 1.9 a resolution. *Journal of Molecular Biology*, *319*(5), 1097–1113. https://doi.org/10.1016/S0022-2836(02)00386-8
14. Luger, K. (2006). Dynamic nucleosomes. *Chromosome Research*, *14*(1), 5–16. https://doi.org/10.1007/s10577-005-1026-1
15. Freitas, M. A., Sklenar, A. R., & Parthun, M. R. (2004). Application of mass spectrometry to the identification and quantification of histone post-translational modifications. *Journal of Cellular Biochemistry*, *92*(4), 691–700. https://doi.org/10.1002/jcb.20106
16. Hergeth, S. P., & Schneider, R. (2015). The H1 linker histones: Multifunctional proteins beyond the nucleosomal core particle. *EMBO Reports*, *16*(11), 1439–1453. https://doi.org/10.15252/embr.201540749
17. Allis, C. D., & Jenuwein, T. (2016). The molecular hallmarks of epigenetic control. *Nature Reviews. Genetics*, *17*(8), 487–500. https://doi.org/10.1038/nrg.2016.59
18. Okosun, J., Bödör, C., Wang, J., Araf, S., Yang, C.-Y., Pan, C., Boller, S., Cittaro, D., Bozek, M., Iqbal, S., Matthews, J., Wrench, D., Marzec, J., Tawana, K., Popov, N., O'Riain, C., O'Shea, D., Carlotti, E., Davies, A., … Fitzgibbon, J. (2014). Integrated genomic analysis identifies recurrent mutations and evolution patterns driving the initiation and progression of follicular lymphoma. *Nature Genetics*, *46*(2), 176–181. https://doi.org/10.1038/ng.2856
19. Sherr, C. J., & Roberts, J. M. (1995). Inhibitors of mammalian G1 cyclin-dependent kinases. *Genes & Development*, *9*(10), 1149–1163. https://doi.org/10.1101/gad.9.10.1149
20. Alabert, C., & Groth, A. (2012). Chromatin replication and epigenome maintenance. *Nature Reviews Molecular Cell Biology*, *13*(3), 153–167. https://doi.org/10.1038/nrm3288
21. Egli, D., Birkhoff, G., & Eggan, K. (2008). Mediators of reprogramming: Transcription factors and transitions through mitosis. *Nature Reviews. Molecular Cell Biology*, *9*(7), 505–516. https://doi.org/10.1038/nrm2439
22. Martínez-Balbás, M. A., Dey, A., Rabindran, S. K., Ozato, K., & Wu, C. (1995). Displacement of sequence-specific transcription factors from mitotic chromatin. *Cell*, *83*(1), 29–38. https://doi.org/10.1016/0092-8674(95)90231-7
23. Littau, V. C., Allfrey, V. G., Frenster, J. H., & Mirsky, A. E. (1964). ACTIVE AND INACTIVE REGIONS OF NUCLEAR CHROMATIN AS REVEALED BY ELECTRON MICROSCOPE AUTORADIOGRAPHY. *Proceedings of the National Academy of Sciences of the United States of America*, *52*(1), 93–100. https://doi.org/10.1073/pnas.52.1.93
24. Johnson, T. C., & Holland, J. J. (1965). Ribonucleic acid and protein synthesis in mitotic HeLa cells. *The Journal of Cell Biology*, *27*(3), 565–574. https://doi.org/10.1083/jcb.27.3.565
25. Hsiung, C. C.-S., Morrissey, C. S., Udugama, M., Frank, C. L., Keller, C. A., Baek, S., Giardine, B., Crawford, G. E., Sung, M.-H., Hardison, R. C., & Blobel, G. A. (2015). Genome accessibility is widely preserved and locally modulated during mitosis. *Genome Research*, *25*(2), 213–225. https://doi.org/10.1101/gr.180646.114
26. Zaret, K. S. (2014). Genome reactivation after the silence in mitosis: Recapitulating mechanisms of development? *Developmental Cell*, *29*(2), 132–134. https://doi.org/10.1016/j.devcel.2014.04.019

27. Grandy, R. A., Whitfield, T. W., Wu, H., Fitzgerald, M. P., VanOudenhove, J. J., Zaidi, S. K., Montecino, M. A., Lian, J. B., van Wijnen, A. J., Stein, J. L., & Stein, G. S. (2016). Genome-Wide Studies Reveal that H3K4me3 Modification in Bivalent Genes Is Dynamically Regulated during the Pluripotent Cell Cycle and Stabilized upon Differentiation. *Molecular and Cellular Biology*, *36*(4), 615–627. https://doi.org/10.1128/MCB.00877-15
28. Zhao, R., Nakamura, T., Fu, Y., Lazar, Z., & Spector, D. L. (2011). Gene bookmarking accelerates the kinetics of post-mitotic transcriptional re-activation. *Nature Cell Biology*, *13*(11), 1295–1304. https://doi.org/10.1038/ncb2341
29. Ming, X., Zhang, Z., Zou, Z., Lv, C., Dong, Q., He, Q., Yi, Y., Li, Y., Wang, H., & Zhu, B. (2020). Kinetics and mechanisms of mitotic inheritance of DNA methylation and their roles in aging-associated methylome deterioration. *Cell Research*, *30*(11), 980–996. https://doi.org/10.1038/s41422-020-0359-9
30. Kadauke, S., Udugama, M. I., Pawlicki, J. M., Achtman, J. C., Jain, D. P., Cheng, Y., Hardison, R. C., & Blobel, G. A. (2012). Tissue-specific mitotic bookmarking by hematopoietic transcription factor GATA1. *Cell*, *150*(4), 725–737. https://doi.org/10.1016/j.cell.2012.06.038
31. Cang, Y., & Prelich, G. (2002). Direct stimulation of transcription by negative cofactor 2 (NC2) through TATA-binding protein (TBP). *Proceedings of the National Academy of Sciences*, *99*(20), 12727–12732. https://doi.org/10.1073/pnas.202236699
32. Christova, R., & Oelgeschläger, T. (2002). Association of human TFIID-promoter complexes with silenced mitotic chromatin in vivo. *Nature Cell Biology*, *4*(1), 79–82. https://doi.org/10.1038/ncb733
33. Imai, Y., Kurokawa, M., Tanaka, K., Friedman, A. D., Ogawa, S., Mitani, K., Yazaki, Y., & Hirai, H. (1998). TLE, the human homolog of groucho, interacts with AML1 and acts as a repressor of AML1-induced transactivation. *Biochemical and Biophysical Research Communications*, *252*(3), 582–589. https://doi.org/10.1006/bbrc.1998.9705
34. Chuang, L. S. H., Ito, K., & Ito, Y. (2013). RUNX family: Regulation and diversification of roles through interacting proteins. *International Journal of Cancer*, *132*(6), 1260–1271. https://doi.org/10.1002/ijc.27964
35. Komori, T. (2011). Signaling networks in RUNX2-dependent bone development. *Journal of Cellular Biochemistry*, *112*(3), 750–755. https://doi.org/10.1002/jcb.22994
36. Lee, K.-S., Lee, Y.-S., Lee, J.-M., Ito, K., Cinghu, S., Kim, J.-H., Jang, J.-W., Li, Y.-H., Goh, Y.-M., Chi, X.-Z., Wee, H., Lee, H.-W., Hosoya, A., Chung, J.-H., Jang, J.-J., Kundu, J. K., Surh, Y.-J., Kim, W.-J., Ito, Y., … Bae, S.-C. (2010). Runx3 is required for the differentiation of lung epithelial cells and suppression of lung cancer. *Oncogene*, *29*(23), 3349–3361. https://doi.org/10.1038/onc.2010.79
37. Fujii, S., Ito, K., Ito, Y., & Ochiai, A. (2008). Enhancer of zeste homologue 2 (EZH2) down-regulates RUNX3 by increasing histone H3 methylation. *The Journal of Biological Chemistry*, *283*(25), 17324–17332. https://doi.org/10.1074/jbc.M800224200
38. Banerji, S., Cibulskis, K., Rangel-Escareno, C., Brown, K. K., Carter, S. L., Frederick, A. M., Lawrence, M. S., Sivachenko, A. Y., Sougnez, C., Zou, L., Cortes, M. L., Fernandez-Lopez, J. C., Peng, S., Ardlie, K. G., Auclair, D., Bautista-Piña, V., Duke, F., Francis, J., Jung, J., … Meyerson, M. (2012).

Sequence analysis of mutations and translocations across breast cancer subtypes. *Nature*, *486*(7403), 405–409. https://doi.org/10.1038/nature11154
39. van Bragt, M. P. A., Hu, X., Xie, Y., & 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
40. 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., Zaidi, S. K., & Stein, G. S. (2020). Inhibition of the RUNX1-CBFβ transcription factor complex compromises mammary epithelial cell identity: A phenotype potentially stabilized by mitotic gene bookmarking. *Oncotarget*, *11*(26), 2512–2530. https://doi.org/10.18632/oncotarget.27637
41. Singh, P., Saxena, R., Srinivas, G., Pande, G., & Chattopadhyay, A. (2013). Cholesterol biosynthesis and homeostasis in regulation of the cell cycle. *PloS One*, *8*(3), e58833. https://doi.org/10.1371/journal.pone.0058833
42. Snaebjornsson, M. T., Janaki-Raman, S., & Schulze, A. (2020). Greasing the Wheels of the Cancer Machine: The Role of Lipid Metabolism in Cancer. *Cell Metabolism*, *31*(1), 62–76. https://doi.org/10.1016/j.cmet.2019.11.010
43. Arun, G., Aggarwal, D., & Spector, D. L. (2020). MALAT1 Long Non-Coding RNA: Functional Implications. *Non-Coding RNA*, *6*(2), 22. https://doi.org/10.3390/ncrna6020022
44. Klec, C., Prinz, F., & Pichler, M. (2019). Involvement of the long noncoding RNA NEAT1 in carcinogenesis. *Molecular Oncology*, *13*(1), 46–60. https://doi.org/10.1002/1878-0261.12404
45. Jadaliha, M., Zong, X., Malakar, P., Ray, T., Singh, D. K., Freier, S. M., Jensen, T., Prasanth, S. G., Karni, R., Ray, P. S., & Prasanth, K. V. (2016). Functional and prognostic significance of long non-coding RNA MALAT1 as a metastasis driver in ER negative lymph node negative breast cancer. *Oncotarget*, *7*(26), 40418–40436. https://doi.org/10.18632/oncotarget.9622
46. Chang, J. W., Seo, S. T., Im, M. A., Won, H.-R., Liu, L., Oh, C., Jin, Y. L., Piao, Y., Kim, H. J., Kim, J. T., Jung, S.-N., & Koo, B. S. (2022). Claudin-1 mediates progression by regulating EMT through AMPK/TGF-β signaling in head and neck squamous cell carcinoma. *Translational Research: The Journal of Laboratory and Clinical Medicine*, *247*, 58–78. https://doi.org/10.1016/j.trsl.2022.04.003
47. Ma, X., Miao, H., Jing, B., Pan, Q., Zhang, H., Chen, Y., Zhang, D., Liang, Z., Wen, Z., & Li, M. (2015). Claudin-4 controls the proliferation, apoptosis, migration and in vivo growth of MCF-7 breast cancer cells. *Oncology Reports*, *34*(2), 681–690. https://doi.org/10.3892/or.2015.4037
48. Khan, S. F., Damerell, V., Omar, R., Du Toit, M., Khan, M., Maranyane, H. M., Mlaza, M., Bleloch, J., Bellis, C., Sahm, B. D. B., Peres, J., ArulJothi, K. N., & Prince, S. (2020). The roles and regulation of TBX3 in development and disease. *Gene*, *726*, 144223. https://doi.org/10.1016/j.gene.2019.144223
49. Wansleben, S., Peres, J., Hare, S., Goding, C. R., & Prince, S. (2014). T-box transcription factors in cancer biology. *Biochimica et Biophysica Acta (BBA) - Reviews on Cancer*, *1846*(2), 380–391. https://doi.org/10.1016/j.bbcan.2014.08.004
50. Peres, J., Damerell, V., Chauhan, J., Popovic, A., Desprez, P.-Y., Galibert, M.-D., Goding, C. R., & Prince, S. (2021). TBX3 Promotes Melanoma Migration by Transcriptional Activation of ID1, which Prevents Activation of E-Cadherin

by MITF. *Journal of Investigative Dermatology*, *141*(9), 2250-2260.e2. https://doi.org/10.1016/j.jid.2021.02.740
51. Roschger, C., & Cabrele, C. (2017). The Id-protein family in developmental and cancer-associated pathways. *Cell Communication and Signaling: CCS*, *15*(1), 7. https://doi.org/10.1186/s12964-016-0161-y
52. Lee, J., Kim, K., Kim, J. H., Jin, H. M., Choi, H. K., Lee, S.-H., Kook, H., Kim, K. K., Yokota, Y., Lee, S. Y., Choi, Y., & Kim, N. (2006). Id helix-loop-helix proteins negatively regulate TRANCE-mediated osteoclast differentiation. *Blood*, *107*(7), 2686–2693. https://doi.org/10.1182/blood-2005-07-2798
53. Mansky, K. C., Marfatia, K., Purdom, G. H., Luchin, A., Hume, D. A., & Ostrowski, M. C. (2002). The *microphthalmia* transcription factor (MITF) contains two N-terminal domains required for transactivation of osteoclast target promoters and rescue of *mi* mutant osteoclasts. *Journal of Leukocyte Biology*, *71*(2), 295–303. https://doi.org/10.1189/jlb.71.2.295
54. Zhong, P., Shu, R., Wu, H., Liu, Z., Shen, X., & Hu, Y. (2021). Low KRT15 expression is associated with poor prognosis in patients with breast invasive carcinoma. *Experimental and Therapeutic Medicine*, *21*(4), 305. https://doi.org/10.3892/etm.2021.9736
55. Bianchini, G., Balko, J. M., Mayer, I. A., Sanders, M. E., & Gianni, L. (2016). Triple-negative breast cancer: Challenges and opportunities of a heterogeneous disease. *Nature Reviews. Clinical Oncology*, *13*(11), 674–690. https://doi.org/10.1038/nrclinonc.2016.66
56. Merkin, R. D., Vanner, E. A., Romeiser, J. L., Shroyer, A. L. W., Escobar-Hoyos, L. F., Li, J., Powers, R. S., Burke, S., & Shroyer, K. R. (2017). Keratin 17 is overexpressed and predicts poor survival in estrogen receptor–negative/human epidermal growth factor receptor-2–negative breast cancer. *Human Pathology*, *62*, 23–32. https://doi.org/10.1016/j.humpath.2016.10.006
57. Tang, S., Liu, W., Yong, L., Liu, D., Lin, X., Huang, Y., Wang, H., & Cai, F. (2022). Reduced Expression of KRT17 Predicts Poor Prognosis in HER2high Breast Cancer. *Biomolecules*, *12*(9), 1183. https://doi.org/10.3390/biom12091183
58. Chen, G., Sun, J., Xie, M., Yu, S., Tang, Q., & Chen, L. (2021). PLAU Promotes Cell Proliferation and Epithelial-Mesenchymal Transition in Head and Neck Squamous Cell Carcinoma. *Frontiers in Genetics*, *12*, 651882. https://doi.org/10.3389/fgene.2021.651882
59. Urban, P., Vuaroqueaux, V., Labuhn, M., Delorenzi, M., Wirapati, P., Wight, E., Senn, H.-J., Benz, C., Eppenberger, U., & Eppenberger-Castori, S. (2006). Increased Expression of Urokinase-Type Plasminogen Activator mRNA Determines Adverse Prognosis in ErbB2-Positive Primary Breast Cancer. *Journal of Clinical Oncology*, *24*(26), 4245–4253. https://doi.org/10.1200/JCO.2005.05.1912
60. Vinchure, O. S., Sharma, V., Tabasum, S., Ghosh, S., Singh, R. P., Sarkar, C., & Kulshreshtha, R. (2019). Polycomb complex mediated epigenetic reprogramming alters TGF-β signaling via a novel EZH2/miR-490/TGIF2 axis thereby inducing migration and EMT potential in glioblastomas. *International Journal of Cancer*, *145*(5), 1254–1269. https://doi.org/10.1002/ijc.32360
61. Nuytten, M., Beke, L., Van Eynde, A., Ceulemans, H., Beullens, M., Van Hummelen, P., Fuks, F., & Bollen, M. (2008). The transcriptional repressor NIPP1 is an essential player in EZH2-mediated gene silencing. *Oncogene*, *27*(10), 1449–1460. https://doi.org/10.1038/sj.onc.1210774

62. Takayama, K., Suzuki, T., Tsutsumi, S., Fujimura, T., Urano, T., Takahashi, S., Homma, Y., Aburatani, H., & Inoue, S. (2015). RUNX1, an androgen- and EZH2-regulated gene, has differential roles in AR-dependent and -independent prostate cancer. *Oncotarget*, *6*(4), 2263–2276. https://doi.org/10.18632/oncotarget.2949
63. Cave, D. D., Buonaiuto, S., Sainz, B., Fantuz, M., Mangini, M., Carrer, A., Di Domenico, A., Iavazzo, T. T., Andolfi, G., Cortina, C., Sevillano, M., Heeschen, C., Colonna, V., Corona, M., Cucciardi, A., Di Guida, M., Batlle, E., De Luca, A., & Lonardo, E. (2022). LAMC2 marks a tumor-initiating cell population with an aggressive signature in pancreatic cancer. *Journal of Experimental & Clinical Cancer Research*, *41*(1), 315. https://doi.org/10.1186/s13046-022-02516-w
64. Marcato, P., Dean, C. A., Pan, D., Araslanova, R., Gillis, M., Joshi, M., Helyer, L., Pan, L., Leidal, A., Gujar, S., Giacomantonio, C. A., & Lee, P. W. K. (2011). Aldehyde Dehydrogenase Activity of Breast Cancer Stem Cells Is Primarily Due To Isoform ALDH1A3 and Its Expression Is Predictive of Metastasis. *Stem Cells*, *29*(1), 32–45. https://doi.org/10.1002/stem.563
65. Li, W., Ma, H., Zhang, J., Zhu, L., Wang, C., & Yang, Y. (2017). Unraveling the roles of CD44/CD24 and ALDH1 as cancer stem cell markers in tumorigenesis and metastasis. *Scientific Reports*, *7*(1), 13856. https://doi.org/10.1038/s41598-017-14364-2
66. Al-Hajj, M., Wicha, M. S., Benito-Hernandez, A., Morrison, S. J., & Clarke, M. F. (2003). Prospective identification of tumorigenic breast cancer cells. *Proceedings of the National Academy of Sciences of the United States of America*, *100*(7), 3983–3988. https://doi.org/10.1073/pnas.0530291100
67. Jaggupilli, A., & Elkord, E. (2012). Significance of CD44 and CD24 as cancer stem cell markers: An enduring ambiguity. *Clinical & Developmental Immunology*, *2012*, 708036. https://doi.org/10.1155/2012/708036
68. Phillips, J. L., Taberlay, P. C., Woodworth, A. M., Hardy, K., Brettingham-Moore, K. H., Dickinson, J. L., & Holloway, A. F. (2018). Distinct mechanisms of regulation of the ITGA6 and ITGB4 genes by RUNX1 in myeloid cells. *Journal of Cellular Physiology*, *233*(4), 3439–3453. https://doi.org/10.1002/jcp.26197
69. Mercurio, A. M., & Rabinovitz, I. (2001). Towards a mechanistic understanding of tumor invasion—Lessons from theα 6 β 4integrin. *Seminars in Cancer Biology*, *11*(2), 129–141. https://doi.org/10.1006/scbi.2000.0364
70. Diaz, L. K., Cristofanilli, M., Zhou, X., Welch, K. L., Smith, T. L., Yang, Y., Sneige, N., Sahin, A. A., & Gilcrease, M. Z. (2005). B4 integrin subunit gene expression correlates with tumor size and nuclear grade in early breast cancer. *Modern Pathology*, *18*(9), 1165–1175. https://doi.org/10.1038/modpathol.3800411
71. Kitajiri, S., Hosaka, N., Hiraumi, H., Hirose, T., & Ikehara, S. (2002). Increased expression of integrin β-4 in papillary thyroid carcinoma with gross lymph node metastasis. *Pathology International*, *52*(7), 438–441. https://doi.org/10.1046/j.1440-1827.2002.01379.x
72. Bierie, B., Pierce, S. E., Kroeger, C., Stover, D. G., Pattabiraman, D. R., Thiru, P., Liu Donaher, J., Reinhardt, F., Chaffer, C. L., Keckesova, Z., & Weinberg, R. A. (2017). Integrin-β4 identifies cancer stem cell-enriched populations of partially mesenchymal carcinoma cells. *Proceedings of the National Academy of Sciences*, *114*(12). https://doi.org/10.1073/pnas.1618298114

73. Masugi, Y., Yamazaki, K., Emoto, K., Effendi, K., Tsujikawa, H., Kitago, M., Itano, O., Kitagawa, Y., & Sakamoto, M. (2015). Upregulation of integrin β4 promotes epithelial–mesenchymal transition and is a novel prognostic marker in pancreatic ductal adenocarcinoma. *Laboratory Investigation*, *95*(3), 308–319. https://doi.org/10.1038/labinvest.2014.166
74. Pei, Y.-F., Liu, J., Cheng, J., Wu, W.-D., & Liu, X.-Q. (2019). Silencing of LAMC2 Reverses Epithelial-Mesenchymal Transition and Inhibits Angiogenesis in Cholangiocarcinoma via Inactivation of the Epidermal Growth Factor Receptor Signaling Pathway. *The American Journal of Pathology*, *189*(8), 1637–1653. https://doi.org/10.1016/j.ajpath.2019.03.012
75. Moon, Y. W., Rao, G., Kim, J. J., Shim, H.-S., Park, K.-S., An, S. S., Kim, B., Steeg, P. S., Sarfaraz, S., Changwoo Lee, L., Voeller, D., Choi, E. Y., Luo, J., Palmieri, D., Chung, H. C., Kim, J.-H., Wang, Y., & Giaccone, G. (2015). LAMC2 enhances the metastatic potential of lung adenocarcinoma. *Cell Death & Differentiation*, *22*(8), 1341–1352. https://doi.org/10.1038/cdd.2014.228
76. Zhang, H., Pan, Y., Cheung, M., Cao, M., Yu, C., Chen, L., Zhan, L., He, Z., & Sun, C. (2019). LAMB3 mediates apoptotic, proliferative, invasive, and metastatic behaviors in pancreatic cancer by regulating the PI3K/Akt signaling pathway. *Cell Death & Disease*, *10*(3), 230. https://doi.org/10.1038/s41419-019-1320-z
77. Miyazaki, K., Miyazaki, M., Guo, Y., Yamasaki, N., Kanno, M., Honda, Z., Oda, H., Kawamoto, H., & Honda, H. (2010). The Role of the Basic Helix-Loop-Helix Transcription Factor *Dec1* in the Regulatory T Cells. *The Journal of Immunology*, *185*(12), 7330–7339. https://doi.org/10.4049/jimmunol.1001381
78. Ehata, S., Hanyu, A., Hayashi, M., Aburatani, H., Kato, Y., Fujime, M., Saitoh, M., Miyazawa, K., Imamura, T., & Miyazono, K. (2007). Transforming growth factor-beta promotes survival of mammary carcinoma cells through induction of antiapoptotic transcription factor DEC1. *Cancer Research*, *67*(20), 9694–9703. https://doi.org/10.1158/0008-5472.CAN-07-1522
79. Sethuraman, A., Brown, M., Krutilina, R., Wu, Z.-H., Seagroves, T. N., Pfeffer, L. M., & Fan, M. (2018). BHLHE40 confers a pro-survival and pro-metastatic phenotype to breast cancer cells by modulating HBEGF secretion. *Breast Cancer Research*, *20*(1), 117. https://doi.org/10.1186/s13058-018-1046-3
80. Logarajah, S., Hunter, P., Kraman, M., Steele, D., Lakhani, S., Bobrow, L., Venkitaraman, A., & Wagner, S. (2003). BCL-6 is expressed in breast cancer and prevents mammary epithelial differentiation. *Oncogene*, *22*(36), 5572–5578. https://doi.org/10.1038/sj.onc.1206689
81. Gurel, B., Ali, T. Z., Montgomery, E. A., Begum, S., Hicks, J., Goggins, M., Eberhart, C. G., Clark, D. P., Bieberich, C. J., Epstein, J. I., & De Marzo, A. M. (2010). NKX3.1 as a marker of prostatic origin in metastatic tumors. *The American Journal of Surgical Pathology*, *34*(8), 1097–1105. https://doi.org/10.1097/PAS.0b013e3181e6cbf3
82. Abdulkadir, S. A., Magee, J. A., Peters, T. J., Kaleem, Z., Naughton, C. K., Humphrey, P. A., & Milbrandt, J. (2002). Conditional loss of Nkx3.1 in adult mice induces prostatic intraepithelial neoplasia. *Molecular and Cellular Biology*, *22*(5), 1495–1503. https://doi.org/10.1128/MCB.22.5.1495-1503.2002
83. Atlas, E., Cardillo, M., Mehmi, I., Zahedkargaran, H., Tang, C., & Lupu, R. (2003). Heregulin is sufficient for the promotion of tumorigenicity and

metastasis of breast cancer cells in vivo. *Molecular Cancer Research: MCR*, *1*(3), 165–175.
84. Chung, Y., Sheu, M., Yang, S., Lin, C., & Yen, S. (2002). Resistance to tamoxifen-induced apoptosis is associated with direct interaction between Her2/neu and cell membrane estrogen receptor in breast cancer. *International Journal of Cancer*, *97*(3), 306–312. https://doi.org/10.1002/ijc.1614
85. Wolf-Yadlin, A., Kumar, N., Zhang, Y., Hautaniemi, S., Zaman, M., Kim, H., Grantcharova, V., Lauffenburger, D. A., & White, F. M. (2006). Effects of HER2 overexpression on cell signaling networks governing proliferation and migration. *Molecular Systems Biology*, *2*(1), 54. https://doi.org/10.1038/msb4100094
86. Fortschegger, K., de Graaf, P., Outchkourov, N. S., van Schaik, F. M. A., Timmers, H. T. M., & Shiekhattar, R. (2010). PHF8 targets histone methylation and RNA polymerase II to activate transcription. *Molecular and Cellular Biology*, *30*(13), 3286–3298. https://doi.org/10.1128/MCB.01520-09
87. Li, Z., Wang, D., Lu, J., Huang, B., Wang, Y., Dong, M., Fan, D., Li, H., Gao, Y., Hou, P., Li, M., Liu, H., Pan, Z.-Q., Zheng, J., & Bai, J. (2020). Methylation of EZH2 by PRMT1 regulates its stability and promotes breast cancer metastasis. *Cell Death and Differentiation*, *27*(12), 3226–3242. https://doi.org/10.1038/s41418-020-00615-9
88. Scheel, C., & Weinberg, R. A. (2012). Cancer stem cells and epithelial–mesenchymal transition: Concepts and molecular links. *Seminars in Cancer Biology*, *22*(5–6), 396–403. https://doi.org/10.1016/j.semcancer.2012.04.001
89. Nieto, M. A., Huang, R. Y.-J., Jackson, R. A., & Thiery, J. P. (2016). EMT: 2016. *Cell*, *166*(1), 21–45. https://doi.org/10.1016/j.cell.2016.06.028
90. Tsirigoti, C., Ali, M. M., Maturi, V., Heldin, C.-H., & Moustakas, A. (2022). Loss of SNAI1 induces cellular plasticity in invasive triple-negative breast cancer cells. *Cell Death & Disease*, *13*(9), 832. https://doi.org/10.1038/s41419-022-05280-z
91. Bernardo, G. M., Bebek, G., Ginther, C. L., Sizemore, S. T., Lozada, K. L., Miedler, J. D., Anderson, L. A., Godwin, A. K., Abdul-Karim, F. W., Slamon, D. J., & Keri, R. A. (2013). FOXA1 represses the molecular phenotype of basal breast cancer cells. *Oncogene*, *32*(5), 554–563. https://doi.org/10.1038/onc.2012.62
92. Fu, X., Pereira, R., De Angelis, C., Veeraraghavan, J., Nanda, S., Qin, L., Cataldo, M. L., Sethunath, V., Mehravaran, S., Gutierrez, C., Chamness, G. C., Feng, Q., O'Malley, B. W., Selenica, P., Weigelt, B., Reis-Filho, J. S., Cohen, O., Wagle, N., Nardone, A., … Schiff, R. (2019). FOXA1 upregulation promotes enhancer and transcriptional reprogramming in endocrine-resistant breast cancer. *Proceedings of the National Academy of Sciences*, *116*(52), 26823–26834. https://doi.org/10.1073/pnas.1911584116
93. Lupien, M., Eeckhoute, J., Meyer, C. A., Wang, Q., Zhang, Y., Li, W., Carroll, J. S., Liu, X. S., & Brown, M. (2008). FoxA1 translates epigenetic signatures into enhancer-driven lineage-specific transcription. *Cell*, *132*(6), 958–970. https://doi.org/10.1016/j.cell.2008.01.018
94. Cruz, R. G. B., Madden, S. F., Brennan, K., & Hopkins, A. M. (2022). A Transcriptional Link between HER2, JAM-A and FOXA1 in Breast Cancer. *Cells*, *11*(4), 735. https://doi.org/10.3390/cells11040735

95. Naderi, A., Meyer, M., & Dowhan, D. H. (2012). Cross-regulation between FOXA1 and ErbB2 Signaling in Estrogen Receptor-Negative Breast Cancer. *Neoplasia*, *14*(4), 283-IN3. https://doi.org/10.1593/neo.12294
96. Mani, S. A., Guo, W., Liao, M.-J., Eaton, E. N., Ayyanan, A., Zhou, A. Y., Brooks, M., Reinhard, F., Zhang, C. C., Shipitsin, M., Campbell, L. L., Polyak, K., Brisken, C., Yang, J., & Weinberg, R. A. (2008). The epithelial-mesenchymal transition generates cells with properties of stem cells. *Cell*, *133*(4), 704–715. https://doi.org/10.1016/j.cell.2008.03.027
97. Morel, A.-P., Lièvre, M., Thomas, C., Hinkal, G., Ansieau, S., & Puisieux, A. (2008). Generation of breast cancer stem cells through epithelial-mesenchymal transition. *PloS One*, *3*(8), e2888. https://doi.org/10.1371/journal.pone.0002888
98. Noh, J. M., Choi, D. H., Baek, H., Nam, S. J., Lee, J. E., Kim, J. W., Ki, C.-S., Park, W., & Huh, S. J. (2012). Associations between BRCA Mutations in High-Risk Breast Cancer Patients and Familial Cancers Other than Breast or Ovary. *Journal of Breast Cancer*, *15*(3), 283–287. https://doi.org/10.4048/jbc.2012.15.3.283
99. Shi, R., Liu, L., Wang, F., He, Y., Niu, Y., Wang, C., Zhang, X., Zhang, X., Zhang, H., Chen, M., & Wang, Y. (2020). Downregulation of cytokeratin 18 induces cellular partial EMT and stemness through increasing EpCAM expression in breast cancer. *Cellular Signalling*, *76*, 109810. https://doi.org/10.1016/j.cellsig.2020.109810
100. Sokol, E. S., Sanduja, S., Jin, D. X., Miller, D. H., Mathis, R. A., & Gupta, P. B. (2015). Perturbation-Expression Analysis Identifies RUNX1 as a Regulator of Human Mammary Stem Cell Differentiation. *PLOS Computational Biology*, *11*(4), e1004161. https://doi.org/10.1371/journal.pcbi.1004161