



UNIVERSITAT DE
BARCELONA

Exploring the cellular origins and clonal dynamics of preclinical triple-negative breast cancer models through lineage tracing studies

Elena Vinuesa Pitarch

ADVERTIMENT. La consulta d'aquesta tesi queda condicionada a l'acceptació de les següents condicions d'ús: La difusió d'aquesta tesi per mitjà del servei TDX (www.tdx.cat) i a través del Dipòsit Digital de la UB (diposit.ub.edu) ha estat autoritzada pels titulars dels drets de propietat intel·lectual únicament per a usos privats emmarcats en activitats d'investigació i docència. No s'autoritza la seva reproducció amb finalitats de lucre ni la seva difusió i posada a disposició des d'un lloc aliè al servei TDX ni al Dipòsit Digital de la UB. No s'autoritza la presentació del seu contingut en una finestra o marc aliè a TDX o al Dipòsit Digital de la UB (framing). Aquesta reserva de drets afecta tant al resum de presentació de la tesi com als seus continguts. En la utilització o cita de parts de la tesi és obligat indicar el nom de la persona autora.

ADVERTENCIA. La consulta de esta tesis queda condicionada a la aceptación de las siguientes condiciones de uso: La difusión de esta tesis por medio del servicio TDR (www.tdx.cat) y a través del Repositorio Digital de la UB (diposit.ub.edu) ha sido autorizada por los titulares de los derechos de propiedad intelectual únicamente para usos privados enmarcados en actividades de investigación y docencia. No se autoriza su reproducción con finalidades de lucro ni su difusión y puesta a disposición desde un sitio ajeno al servicio TDR o al Repositorio Digital de la UB. No se autoriza la presentación de su contenido en una ventana o marco ajeno a TDR o al Repositorio Digital de la UB (framing). Esta reserva de derechos afecta tanto al resumen de presentación de la tesis como a sus contenidos. En la utilización o cita de partes de la tesis es obligado indicar el nombre de la persona autora.

WARNING. On having consulted this thesis you're accepting the following use conditions: Spreading this thesis by the TDX (www.tdx.cat) service and by the UB Digital Repository (diposit.ub.edu) has been authorized by the titular of the intellectual property rights only for private uses placed in investigation and teaching activities. Reproduction with lucrative aims is not authorized nor its spreading and availability from a site foreign to the TDX service or to the UB Digital Repository. Introducing its content in a window or frame foreign to the TDX service or to the UB Digital Repository is not authorized (framing). Those rights affect to the presentation summary of the thesis as well as to its contents. In the using or citation of parts of the thesis it's obliged to indicate the name of the author.



UNIVERSITAT DE
BARCELONA

PhD Program in Biomedicine

Exploring the cellular origins and clonal
dynamics of preclinical triple-negative breast
cancer models through lineage tracing studies

*Explorant els orígens cel·lulars i la dinàmica
clonal dels models preclínics de càncer de
mama triple negatiu mitjançant estudis de
rastreig de llinatges*

Elena Vinuesa Pitarch

Josep Carreras Leukaemia Research Institute

Director: Dr. Verónica Rodilla Benito

Tutor: Dr. Meritxell Rovira Clusellas

Badalona, February 2025

“Success is the sum of small efforts,
repeated day in and day out.”

— Robert Collier

ACKNOWLEDGMENTS

Este trabajo no habría sido posible sin el apoyo de todas las personas que me han acompañado a lo largo de estos años. A cada uno de vosotros, gracias de todo corazón.

En primer lugar, quiero expresar mi más profunda gratitud a la persona que ha hecho posible que hoy esté aquí. **Vero**, desde el primer momento confiaste en mí y me brindaste la oportunidad de crecer profesionalmente a tu lado. Gracias por cada oportunidad que me has dado, por tu dedicación incansable y por enseñarme a superar cualquier desafío. Tu apoyo incondicional ha sido el pilar que me ha sostenido en los momentos más difíciles, cuando mis inseguridades y dramatismos ponían a prueba tu infinita paciencia. Tu guía ha hecho que este camino, que en ocasiones parecía imposible, se sintiera mucho más fácil de recorrer. Gracias por mantener siempre mi motivación en alto, por tu empatía y por la manera tan admirable en que logras combinar cercanía y liderazgo. Siempre te estaré profundamente agradecida por la confianza que has depositado en mí, por ser una fuente constante de inspiración y por estar a mi lado en cada paso de este viaje.

También quiero agradecer a mis compañeros del **Rodilla Lab**, quienes han sido mucho más que colegas: han sido mi familia científica. **Cris**, mi mentora y ejemplo a seguir. Siempre has estado dispuesta a ayudarme, a brindarme nuevas ideas y a acompañarme en cada desafío. Más allá de la ciencia, has sido mi compañera de cafés y largas conversaciones, en las que inevitablemente terminábamos hablando de investigación. Gracias por preocuparte siempre por mí, recordarme comer y ser mi segunda madre en el laboratorio. **David**, nuestro lab manager y bioinformático, compañero desde el inicio de esta aventura. Fuimos los primeros miembros del grupo y juntos afrontamos los desafíos de montar el laboratorio desde cero. Gracias por tu paciencia, tu dedicación y por ayudar a dar forma a nuestro espacio de trabajo. **Dani**, comenzaste como mi estudiante de prácticas y en nada te convertiste en el más entusiasta del grupo. Tu pasión por la investigación es contagiosa. Gracias por todo lo que has aportado a mi proyecto, por tu sinceridad y por siempre tener ideas geniales. Tu apoyo ha sido invaluable, y, sobre todo, has sido un compañero y amigo excepcional. **Joan**, el cantante del laboratorio y eterno despistado en mis lab meetings. Espero que después de leer esta tesis entiendas mejor mi proyecto... o quizá no, ¡jajaja! Gracias por tu buen humor y por hacer del laboratorio un lugar más ameno. **Elena**, el motor que ha impulsado mi proyecto. Siempre dispuesta a ayudar, incluso antes de que te lo pidiera. Gracias por tu inagotable energía y compromiso, tu dedicación ha sido clave para avanzar en cada paso de este camino. Y, por supuesto, por ser mi nueva compañera de escritorio, haciendo que los días en el laboratorio sean mucho más llevaderos. **Mariia**, siempre te recordaré por tu nobleza y espíritu incansable, cuyo esfuerzo y entrega inspiraban a todos. También quiero agradecer a los estudiantes que han pasado por el laboratorio, en especial a **Laura**, un ejemplo de dedicación y compromiso, y a **Lidia**, cuya dulzura y pasión por la ciencia la han traído de vuelta para quedarse.

También han sido compañeros imprescindibles en mi camino los miembros del **Belver Lab**, con quienes compartí el laboratorio durante casi todo mi doctorado. En especial, a Sandra y Claudia, quienes han sido fundamentales en mi camino. No sé qué habría hecho sin vuestro sentido del humor y vuestras risas. **Sandra**, siempre serás mi primera compañera de laboratorio. Recuerdo cuando llegaste, cómo nos sentábamos en extremos opuestos de un laboratorio vacío, comenzando a descubrir este mundo juntas. Nunca olvidaré nuestras excursiones al animalario: juntas compartimos miedos, pero también muchos buenos momentos. **Clau**, mi andorrana favorita y fuente de calma. Gracias por tu apoyo y por todos los momentos compartidos que siempre recordaré con mucho cariño. **Eli**, una de mis primeras compañeras en este viaje. Gracias por compartir tu energía positiva y por hacer más agradables los días en el laboratorio. **Noel**, mi compañero de máster con quien el destino volvió a juntarme en el doctorado. Gracias por tu nobleza y generosidad. **Anna**, la Campa, mi italiana favorita. Aunque llegaste más tarde, rápidamente te ganaste un lugar en mi corazón. Gracias por tu cariño, por estar siempre pendiente de mí y por esas conversaciones que siempre me hacían sentir mejor. **Alice y Yahia**, gracias por aportar siempre un buen ambiente y compañerismo durante el tiempo que compartimos despacho. Y, por último, a **Laura**, por tus consejos sabios y tu escucha atenta cuando las inseguridades me invadían. Has sido una gran mentora y te agradezco enormemente tu apoyo.

Agradezco al equipo del animalario, especialmente a **Sergi, Yaiza y Ricard**, por acompañarme todos estos años y hacerme sentir como parte de vuestra familia.

Agradecer también a mis **amigas**, que han sido mi refugio fuera del laboratorio. **Esther**, mi compañera de toda la carrera, gracias por estar siempre ahí, por escucharme y dejarme compartir mis miedos, por nuestras risas y por todos los consejos que me has dado en nuestras interminables charlas. **Anna**, mi compañera de conservatorio y talentosa artista, gracias por tu apoyo siempre presente desde Viena. Aunque la distancia nos separe, nuestra amistad sigue tan fuerte como siempre. **Silvia, Sara y Judith**, mis amigas de toda la vida, las de siempre. Gracias por estar a mi lado durante todo este largo viaje. Os prometo que, cuando finalmente acabe, reemprenderemos las cenitas y los viajes que tenemos pendientes.

Finalmente, quiero agradecer a mi **familia**, mi mayor apoyo en todo momento. A mis padres, **Maria Teresa y Juan Ignacio**, por haberme permitido perseguir mis sueños y por estar siempre a mi lado en este camino, compartiendo conmigo cada esfuerzo y cada logro. Su cariño y confianza han sido fundamentales para que pudiera llegar hasta aquí. A mi hermana, **Blanca**, un pilar fundamental en mi vida. Cada momento juntas es único e irremplazable. Gracias por transmitirme tu vitalidad, que siempre me ha dado la fuerza y la determinación necesarias para enfrentar cualquier desafío. A mi abuela, **Paquita**, por su apoyo constante y por sus sabias palabras que siempre han iluminado mi camino. A mi tío, **Martín**, por tu buen sentido del humor y por transmitirme siempre tanta fuerza y optimismo. Y, por último, a **Jorge**, mi compañero de vida. Gracias por tu paciencia infinita, por acompañarme en cada reto y por hacer más llevadero este camino. Tu apoyo incondicional y tu fe en mí, incluso cuando yo dudaba, han sido mi mayor impulso.

FINANCIAL SUPPORT

This study has received funding from the following sources:

- **PER003:** The research leading to these results has received funding from the Generalitat de Catalunya, with support of Departament de Salut (PERIS PIF-Salut), grant SLT017/20/000140.
- **PID013:** This study is part of the project PID2020-114647RA-I00 funded by MICIU/AEI /10.13039/501100011033.



**Generalitat
de Catalunya**



INDEX

FIGURES AND TABLES INDEX	13
ABSTRACT	17
RESUM	19
ABBREVIATIONS AND ACRONYMS	21
INTRODUCTION	25
1. LINEAGE TRACING APPROACHES: TOOLS AND APPLICATIONS	25
1.1. <i>Cre/lox</i> systems for lineage tracing	26
1.1.1. Tamoxifen-inducible model.....	26
1.1.2. Doxycycline-dependent models.....	27
1.2. <i>Dre/rox</i> systems.....	27
1.3. <i>Flippase/Frt</i> system	29
1.4. <i>Pulse-chase</i> experiments.....	30
2. PHYSIOLOGY OF THE MAMMARY GLAND	31
2.1. <i>Cellular architecture of the mammary gland</i>	31
2.2. <i>Development and homeostasis of the mammary gland</i>	32
2.2.1. Embryonic development	33
2.2.2. Postnatal development during puberty	33
2.2.3. Development during pregnancy.....	34
2.2.4. Lactation	34
2.2.5. Involution.....	35
2.3. <i>Hierarchical organization of the healthy mammary gland</i>	35
2.4. <i>Mammary epithelial cell plasticity</i>	37
2.4.1. Plasticity in transplantation and cellular ablation	38
2.4.2. Hormonal influence on lineage plasticity	38
2.4.3. Ectopic expression of fate-determining factors.....	38
2.4.4. Plasticity induced by genotoxic stress and chemotherapy	39
2.4.5. Lineage infidelity induced by oncogene activation	39
3. BREAST CANCER	40
3.1. <i>Breast cancer classification and therapeutics</i>	40
3.1.1. Breast cancer morphological classification.....	40
3.1.1.1. Localized breast cancer types.....	40
3.1.1.2. Invasive breast cancer types.....	40
3.1.2. Breast cancer subtypes and associated therapies	42
3.1.2.1. Hormone receptor (HR)-positive subtypes.....	43
3.1.2.2. HER2-positive subtypes.....	43
3.1.2.3. Triple-negative breast cancer (TNBC)	43
3.1.3. Advances in molecular subtyping	44
3.1.4. Molecular heterogeneity in breast cancer.....	46
3.1.4.1. Luminal breast cancer heterogeneity.....	46
3.1.4.2. HER2-positive breast cancer heterogeneity	47

3.1.4.3. TNBC heterogeneity.....	47
3.2. <i>In vivo breast cancer models</i>	49
3.2.1. Cell line-derived xenografts	49
3.2.2. Patient-derived xenografts.....	50
3.2.3. Genetically engineered mouse models	50
3.2.4. Inducible breast cancer models	53
3.3. <i>Lineage tracing as the gold-standard approach for exploring cellular hierarchies and heterogeneity in breast cancer</i>	54
3.3.1. Unraveling the cell of origin in different breast cancer subtypes	54
3.3.2. Exploring clonal dynamics during tumor progression and dissemination using combinatorial approaches	56
3.3.2.1. 3D-whole mount imaging to explore tumor dynamics at high resolution	56
3.3.2.2. Intravital imaging for real-time monitoring of tumor cell dynamics... ..	57
3.3.2.1. Single-cell lineage tracing through cellular barcoding for mapping tumor subclonal dynamics.....	58
3.3.3. EMT heterogeneity and its role in tumor evolution	59
3.3.4. Metastatic dynamics revealed by multicolor lineage tracing	60
HYPOTHESIS AND OBJECTIVES	61
MATERIALS AND METHODS	63
1. MICE.....	63
1.1. <i>Maintenance and welfare of the animal colony</i>	63
1.2. <i>Murine models</i>	63
1.2.1. Breast cancer models	63
1.2.1.1. MMTV-PyMT	63
1.2.1.2. C3(1)-TAg.....	64
1.2.1.1. DMBA-induced breast cancer model	65
1.2.2. Murine transgenic lines for lineage tracing experiments in breast tumors	66
1.2.2.1. Inducible Cre transgenic lines	66
1.2.2.2. Rosa26 ^{mTmG} reporter line	66
1.2.2.3. Establishment of murine transgenic lines for lineage tracing.....	66
2. POLYMERASE CHAIN REACTION (PCR)-BASED GENOTYPING OF TRANSGENIC MICE.....	67
2.1. <i>Weaning and ear punching</i>	67
2.2. <i>DNA extraction from ear biopsies</i>	67
2.3. <i>PCR genotyping</i>	68
3. BULK RNA SEQUENCING.....	69
4. LINEAGE TRACING EXPERIMENTS	70
5. FLOW CYTOMETRY (FC).....	70
6. HISTOLOGICAL IMMUNOSTAININGS	73
6.1. <i>Immunohistochemistry (IHC)</i>	74
6.2. <i>Immunofluorescence (IF)</i>	76
7. STATISTICAL ANALYSIS.....	77

8. TABLES OF REAGENTS, MATERIALS, AND EQUIPMENT	78
RESULTS	81
OBJECTIVE 1: TO HISTOLOGICALLY AND MOLECULARLY CHARACTERIZE DIFFERENT PRECLINICAL MODELS OF TRIPLE-NEGATIVE BREAST CANCER.	81
1. DISTINCT MURINE TNBC PRECLINICAL MODELS REFLECT DIVERSE HUMAN BREAST CANCER SUBTYPES	81
1.1. <i>Histological characterization of TNBC preclinical models</i>	81
1.2. <i>Transcriptomic characterization of TNBC preclinical models</i>	83
1.3. <i>Claudin-low phenotype in basal and hybrid tumors</i>	84
OBJECTIVE 2: TO ELUCIDATE THE CELLULAR ORIGINS IN MULTIPLE SUBTYPES OF TRIPLE-NEGATIVE BREAST CANCER.	86
2. LINEAGE TRACING STUDIES UNRAVEL DIVERGENT ORIGINS IN TNBC SUBTYPES	86
2.1. <i>TNBC tumors with luminal B and basal-like gene expression profiles share a common ERα-negative luminal origin</i>	88
2.2. <i>TNBC tumors with a normal-like gene expression profile have a basal origin</i>	90
2.3. <i>Investigating the cellular origins of TNBC tumors with hybrid features</i>	90
2.4. <i>Distinct lineage origins and variable latencies across DMBA-induced tumor subtypes</i>	91
2.5. <i>Oncogenic plasticity drives heterogeneous TNBC subtypes in the DMBA model</i>	92
OBJECTIVE 3: TO INVESTIGATE THE CLONAL DYNAMICS OF DISTINCT MAMMARY EPITHELIAL CELL TYPES ACROSS TRANSCRIPTOMICALLY DIVERSE TRIPLE-NEGATIVE BREAST CANCER SUBTYPES.	95
3. LUMINAL PROGENITORS DRIVE TUMOR EVOLUTION IN THE LUMINAL TNBC MODELS	95
3.1. <i>Distribution of MEC populations across tumor stages in the MMTV-PyMT model</i>	95
3.2. <i>Notch1-expressing cells drive clonal expansion in early tumor stages in the MMTV-PyMT model</i>	97
3.3. <i>Differential clonal dynamics of luminal progenitors during tumor progression</i>	100
3.4. <i>Cell fate and transformation potential of mammary epithelial populations in the C3(1)-TAg model</i>	102
3.5. <i>Luminal progenitors drive basal-like tumor development in the C3(1)-TAg model</i>	104
OBJECTIVE 4: TO DETERMINE THE METASTATIC POTENTIAL OF DIFFERENT MAMMARY EPITHELIAL CELL POPULATIONS IN BREAST TUMORS.	107
4. NOTCH1-POSITIVE CELLS EXHIBIT CELLULAR PLASTICITY DURING METASTATIC SPREAD	107
4.1. <i>Lung metastases reflect primary tumor heterogeneity in MEC composition</i>	107
4.2. <i>Notch1-positive cells are primary drivers of lung metastasis in the MMTV-PyMT model</i>	109
4.3. <i>Phenotypic plasticity of Notch1-derived cells in lung metastases</i>	110

DISCUSSION 113

1. PRECLINICAL MODELS RECAPITULATE TNBC SUBTYPE DIVERSITY OBSERVED IN CLINICS 113

 1.1. *Histological and molecular variability across models* 113

 1.2. *Model-specific strengths and applications* 113

2. INTERPLAY BETWEEN CELLULAR ORIGINS AND ONCOGENIC DRIVERS IN TNBC HETEROGENEITY 114

 2.1. *Unveiling the origins of luminal tumors: insights and controversies* 114

 2.2. *Basal origins of TNBC: insights into normal-like tumorigenesis* 115

 2.3. *Hybrid tumors in TNBC: bridging luminal and basal characteristics* 115

 2.4. *DMBA-induced oncogenesis: a driver of phenotypic plasticity* 115

3. LUMINAL PROGENITORS: A PROMISING FOCUS FOR TREATING LUMINAL B AND BASAL-LIKE SUBTYPES 116

 3.1. *The evolving clonal landscape of luminal progenitors in luminal B tumors*. 116

 3.2. *Notch1-positive luminal progenitors in the C3(1)-TAg model: key drivers of basal-like tumorigenesis* 117

4. INSIGHTS INTO THE METASTATIC BEHAVIOR AND PHENOTYPIC SHIFTS OF *NOTCH1*-DERIVED CELLS 118

 4.1. *Lung metastases reflect the cellular composition of the primary tumor* 118

 4.2. *Notch1-positive cells enhance their aggressiveness in advanced tumor stages* 119

 4.3. *Phenotypic plasticity of Notch1-derived cells during metastatic dissemination* 119

5. CONCLUDING REMARKS AND FUTURE DIRECTIONS 120

 5.1. *Key contributions* 120

 5.2. *Future directions* 121

CONCLUSIONS 122

BIBLIOGRAPHY 125

FIGURES AND TABLES INDEX

INTRODUCTION

Figure I1. Two inducible Cre systems for gene activation in mouse models..... 26

Figure I2. Schematic representation of pulse-chase experiments in lineage tracing . 30

Figure I3. Schematic representation of the normal mammary gland architecture..... 32

Figure I4. Schematic representation of mammary gland development across life stages 33

Figure I5. Hierarchical structure of the mammary gland throughout development..... 36

Figure I6. Immunohistochemical classification of breast cancer based on receptor status: ER α , PR, and HER2. 42

Figure I7. TNBC subtype distribution by PAM50 (combined with claudin-low classifier) and Lehmann classifications 48

Table I1. Overview of genetically engineered and chemically induced mouse models of mammary tumors, specifying oncogene or tumor suppressor gene alterations, targeted MECs, and corresponding human breast cancer subtypes..... 52

Figure I8. Tumor subtype specification in BaCs and LCs using distinct oncogenic drivers 55

Figure I9. Schematic model of clonal dynamics in MMTV-PyMT tumors, derived from multi-color lineage tracing studies using real-time intravital imaging 57

MATERIALS AND METHODS

Table M1. Summary of the breast cancer models used in this study 63

Figure M1. Schematic representation of the breeding strategy used to generate transgenic murine lines for lineage tracing 67

Table M2. Primer sequences used for PCR-based genotyping of transgenic murine lines..... 68

Table M3. Reagents, volumes, and PCR conditions for genotyping transgenic murine lines..... 69

Table M4. Flow cytometric antibodies used to identify mouse cell surface markers .. 71

Table M5. Isotype controls used in flow cytometry 71

Table M6. Summary of fluorophores used in flow cytometry analysis. 72

Figure M2. Flow cytometry gating strategy illustrating recombination efficiency and lineage specificity during pulse experiments 73

Table M7. Primary antibodies used in IHC, with corresponding references, companies, species, and dilutions 75

Table M8. Secondary antibodies used in IHC, with corresponding references, companies, species, and dilutions 75

Table M9. Primary antibodies used in IF, with corresponding references, companies, species, and dilutions 77

Table M10. Secondary antibodies used in IF, with corresponding references, companies, species, and dilutions 77

Table M11. Reagents used in this thesis, with corresponding references and suppliers	78
Table M12. Materials used in this thesis, with corresponding references and suppliers	79
Table M13. Equipment used in this thesis, with corresponding manufacturers.....	80

RESULTS

OBJECTIVE 1

Figure R1.1. Histological characterization of different preclinical breast cancer models	82
Figure R1.2. Distribution of tumor subtypes across preclinical models	83
Figure R1.3. Molecular characterization of preclinical breast cancer models	83
Figure R1.4. Immunohistochemical staining of claudins in MMTV-PyMT, C3(1)-TAg, and DMBA tumors	84

OBJECTIVE 2

Figure R2.1. Lineage tracing of mammary epithelial cell populations using lineage-specific promoters.....	86
Figure R2.2. Recombination efficiency in total and specific MECs	87
Figure R2.3. Summary of the experimental design for the MMTV-PyMT, C3(1)-TAg, and DMBA-induced breast cancer models to investigate their cell of origin	87
Figure R2.4. ER α^{neg} LCs as the origin of TNBC tumors with luminal traits	89
Figure R2.5. TNBC tumors with luminal traits in the DMBA-induced model originate from ER α^{neg} LCs.....	89
Figure R2.6. Triple-negative tumors with histological basal traits in the DMBA-induced model originate from BaCs	90
Figure R2.7. Investigating the cellular origins of hybrid triple-negative tumors in the DMBA-induced model.....	91
Figure R2.8. Summary of cellular origins and tumor latency in DMBA-induced tumors	92
Figure R2.9. Basal-to-luminal transitions contribute to heterogeneity in DMBA-induced tumors.....	93
Figure R2.10. ER α -negative luminal progenitors largely preserve unipotency, yet may occasionally transition to an EMT-like state under oncogenic stress	94

OBJECTIVE 3

Figure R3.1. Distribution of mammary epithelial populations across tumor stages in the MMTV-PyMT model.....	95
Figure R3.2. Flow cytometry-based cell fate analysis of mammary epithelial populations across tumor stages in the MMTV-PyMT model	96
Figure R3.3. Immunofluorescence-based cell fate analysis of mammary epithelial populations in MMTV-PyMT hyperplastic lesions 24 hours post-tamoxifen injection.	97

Figure R.3.4. Early-stage targeting of MEC types reveals sustained contribution of <i>Notch1</i> -positive ER α -negative luminal progenitors to MMTV-PyMT tumor progression	98
Figure R.3.5. <i>Notch1</i> -positive cells targeted at hyperplasia stage exhibit clonal expansion during MMTV-PyMT tumor progression.	99
Figure R3.6. <i>Notch1</i> -derived cells retain luminal identity during clonal expansion in the MMTV-PyMT model	100
Figure R3.7. Clonal expansion of <i>Notch1</i> - and <i>Prom1</i> -positive luminal populations varies with tumorigenic stage in the MMTV-PyMT model	100
Figure R3.8. Tumor stage-dependent changes in the size of GFP-positive tumor clones in <i>Notch1</i> ^{mTmG} and <i>Prom1</i> ^{mTmG} MMTV-PyMT tumors	101
Figure R3.9. Differential susceptibility of mammary epithelial cells to SV40-TAg-driven transformation	103
Figure R3.10. Immunofluorescence-based cell fate analysis of mammary epithelial populations in C3(1)-TAg hyperplastic lesions 24 hours post-tamoxifen injection	103
Figure R3.11. Early-stage targeting of MEC types reveals sustained contribution of <i>Notch1</i> -positive ER α -negative luminal progenitors to C3(1)-TAg tumor progression	104
Figure R3.12. A specific subset of <i>Notch1</i> -positive basal cells undergoes clonal expansion during basal-like tumorigenesis in the C3(1)-TAg model	105
Figure R3.13. Late-stage targeting of <i>Notch1</i> -positive cells reveals continuous clonal growth.	106
OBJECTIVE 4	
Figure R4.1. Lung metastases recapitulate the cellular architecture of the primary tumor	107
Figure R.4.2. Mammary tumor cells retain their primary tumor lineage characteristics in metastatic	108
Figure R.4.3. <i>Notch1</i> -positive cells are highly metastatic in advanced tumor stages in the MMTV-PyMT model	109
Figure R4.4. <i>Notch1</i> -derived cells exhibit luminal-to-basal transition during metastatic spread	111

ABSTRACT

Triple-negative breast cancer (TNBC) is a highly aggressive and heterogeneous subtype that lacks targeted therapies, highlighting the need for a deeper understanding of its molecular and cellular mechanisms. Lineage tracing has emerged as a powerful tool to decipher the cellular hierarchies governing both normal and malignant mammary tissues. Despite significant advancements, knowledge gaps persist regarding how distinct mammary epithelial cell populations contribute to TNBC heterogeneity and progression. By integrating histological, molecular, and lineage-tracing analyses, this thesis provides critical insights into the cellular origins, clonal dynamics, and metastatic potential of different mammary epithelial lineages in preclinical TNBC models.

Lineage-tracing experiments reveal that distinct TNBC subtypes arise from specific mammary epithelial cells, with lineage identity and oncogenic plasticity shaping tumor heterogeneity. ER α -negative *Notch1*-positive luminal cells serve as the cell of origin for TNBCs with a luminal histology resembling luminal B and basal-like human subtypes, while *Acta2*-positive basal cells generate tumors with basal features mirroring normal-like human tumors. Notably, basal cells exhibit phenotypic plasticity, contributing to the emergence of hybrid tumors that blur conventional subtype boundaries. Clonal expansion studies underscore that the pivotal role of luminal progenitors in luminal B and basal-like TNBC progression. In luminal B tumors, *Notch1*-positive luminal cells undergo early, robust clonal expansion, followed by a shift toward a more invasive and plastic phenotype in advanced stages, enhancing their metastatic potential. During dissemination, these cells undergo luminal-to-basal transition, enabling them to reconstruct the primary tumor's cellular architecture at distant sites. In advanced stages, *Notch1*-positive cells are outcompeted by *Prom1*-positive luminal cells, which exhibit delayed but significant proliferative response, while *Acta2*-positive basal cells do not contribute to clonal expansion. A similar pattern emerges in basal-like tumors, where *Notch1*-positive cells, comprising ER α -negative luminal cells and a small subset of basal cells, clonally expand and drive intratumor heterogeneity.

These findings highlight the pivotal role of lineage-restricted cellular hierarchies and oncogenic plasticity in shaping TNBC heterogeneity and progression. By identifying distinct cellular origins, their evolving clonal dynamics, and metastatic trajectories, this work offers a mechanistic framework for understanding TNBC evolution. These insights may guide the development of personalized therapeutic strategies tailored to subtype-specific tumor and metastatic origins, while accounting for phenotypic plasticity to prevent tumor initiation and halt dissemination in TNBC patients.

Key words: Triple-negative breast cancer, lineage tracing, mammary epithelial cells, cellular origins, phenotypic plasticity, tumor heterogeneity, clonal dynamics, metastatic potential.

RESUM

El càncer de mama triple negatiu (TNBC) és un subtipus altament agressiu i heterogeni que manca de teràpies dirigides, fet que ressalta la necessitat de comprendre millor els seus mecanismes moleculars i cel·lulars. El rastreig de llinatge (*lineage tracing*) s'ha consolidat com una eina poderosa per desxifrar les jerarquies cel·lulars que governen tant els teixits mamaris normals com els malignes. Malgrat avenços significatius, persisteixen llacunes sobre com les diferents poblacions de cèl·lules epitelials mamàries contribueixen a la heterogeneïtat i progressió del TNBC. Integrant anàlisis histològiques, moleculars i de *lineage tracing*, aquesta tesi proporciona coneixements clau sobre els orígens cel·lulars, la dinàmica clonal i el potencial metastàtic de diferents llinatges epitelials mamaris en models preclínics de TNBC.

Els experiments de *lineage tracing* revelen que diferents subtipus de TNBC sorgeixen de cèl·lules epitelials mamàries específiques, amb la identitat de llinatge i la plasticitat oncogènica modelant la heterogeneïtat tumoral. Les cèl·lules luminals ER α -negatives *Notch1*-positives serveixen com a cèl·lules d'origen per als TNBC amb una histologia luminal que s'assemblen als subtipus luminal B i basal-like humans, mentre que les cèl·lules basals *Acta2*-positives generen tumors amb característiques basals que imiten els tumors normal-like humans. Cal destacar que les cèl·lules basals mostren plasticitat fenotípica, contribuint a l'emergència de tumors híbrids que desdibuixen les fronteres convencionals entre subtipus. Els estudis d'expansió clonal destaquen el paper fonamental dels progenitors luminals en la progressió dels TNBC luminal B i basal-like. En els tumors luminal B, les cèl·lules luminals *Notch1*-positives experimenten una expansió clonal primerenca i robusta, seguida d'un canvi cap a un fenotip més invasiu i plàstic en etapes avançades, augmentant així el seu potencial metastàtic. Durant la disseminació, aquestes cèl·lules experimenten una transició luminal-a-basal, la qual els permet reconstruir l'arquitectura cel·lular del tumor primari en llocs distants. En estadis avançats, les cèl·lules *Notch1*-positives són desplaçades per les cèl·lules luminals *Prom1*-positives, que mostren una resposta proliferativa retardada però significativa, mentre que les cèl·lules basals *Acta2*-positives no contribueixen a l'expansió clonal. Un patró similar s'observa en els tumors basal-like, on les cèl·lules *Notch1*-positives, que inclouen cèl·lules luminals ER α -negatives i un petit subconjunt de cèl·lules basals, s'expandeixen clonalment i impulsen la heterogeneïtat intratumoral.

Aquests descobriments ressalten el paper fonamental de les jerarquies cel·lulars restringides al llinatge i la plasticitat oncogènica en el modelatge de la heterogeneïtat i progressió del TNBC. En identificar orígens cel·lulars específics, la seva dinàmica clonal evolutiva i les trajectòries metastàtiques, aquest treball ofereix un marc mecanístic per comprendre l'evolució del TNBC. Aquests coneixements podrien guiar el desenvolupament d'estratègies terapèutiques personalitzades, adaptades a l'origen tumoral i metastàtic específic de cada subtipus, tenint en compte la plasticitat fenotípica per prevenir la iniciació tumoral i aturar la disseminació en pacients amb TNBC.

Paraules clau: Càncer de mama triple negatiu, *lineage tracing*, cèl·lules epitelials mamàries, orígens cel·lulars, plasticitat fenotípica, heterogeneïtat tumoral, dinàmica clonal, potencial metastàtic.

ABBREVIATIONS AND ACRONYMS

<i>Acta2</i>	Actin alpha 2, smooth muscle, aorta
<i>Aib1</i>	Nuclear receptor coactivator 3
<i>Apc</i>	Adenomatous polyposis coli
APC	Allophycocyanin
AR	Androgen receptor
BaCs	Basal cells
<i>Brca1</i>	Breast cancer 1, early onset
<i>BRCA1</i>	BRCA1 DNA repair associated
<i>Brg1</i>	SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily a, member 4
BSA	Bovine serum albumin
BV510	Brilliant Violet 510™
C3(1)	Rat prostatic steroid binding protein gene
C57BL/6	C57 Black 6 mouse strain
CDX	Cell line-derived xenografts
DAB	3,3'-diaminobenzidine
DAPI	4',6-diamidino-2-phenylindole
dH₂O	Distilled water
DMBA	7,12-dimethylbenz[a]anthracene
Dox	Doxycycline
EDTA	Ethylenediaminetetraacetic acid
EGFR	Epidermal growth factor receptor
EMT	Epithelial-to-mesenchymal transition
ER	Estrogen receptor
ERα	Estrogen receptor alpha
<i>ErbB2</i>	erb-b2 receptor tyrosine kinase 2
<i>Esr1</i>	Estrogen receptor 1, alpha
EtOH	Ethanol
<i>Etv6-Ntrk3</i>	Ets variant 6 - neurotrophic tyrosine kinase, receptor, type 3
FBS	Fetal bovine serum
FC	Flow cytometry
<i>Fgf3</i>	Fibroblast growth factor 3
FITC	Fluorescein isothiocyanate
Flp	Flippase
FVB	Friend Virus B-type
GEMMs	Genetically engineered mouse models
GFP	Green fluorescent protein
HER2	Human epidermal growth factor receptor 2
HR	Hormone receptor

<i>Hras</i>	Harvey rat sarcoma virus oncogene
HRP	Horseradish peroxidase
IF	Immunofluorescence
IHC	Immunohistochemistry
<i>Int3</i>	Notch 4
Ki-67	Proliferation marker protein Ki-67
<i>Kras</i>	Kirsten rat sarcoma viral oncogene homolog
<i>Krt5</i> / K5	Keratin 5
<i>Krt8</i> / K8	Keratin 8
<i>Krt14</i> / K14	Keratin 14
LAR	Luminal androgen receptor
LCs	Luminal cells
<i>Lgr5</i>	Leucine rich repeat containing G protein coupled receptor 5
Lin	Lineage
<i>Lpa</i>	Lysophosphatidic acid
MECs	Mammary epithelial cells
MIN	Mammary intraepithelial neoplasia
MMTV	Mouse mammary tumor virus
MPA	Medroxyprogesterone acetate
<i>Myc</i>	Myelocytomatosis oncogene
p18	Cyclin-dependent kinase 4 inhibitor C (Cdkn2c)
P63	Tumor protein 63
PAM50	Prediction analysis of microarray 50
PBS	Phosphate-buffered saline
pCR	Pathological complete response
PCR	Polymerase chain reaction
PDX	Patient-derived xenografts
PE	Phycoerythrin
PerCP	Peridinin-chlorophyll-protein
PI3K	Phosphatidylinositol 3-kinase
<i>Pik3ca</i>	Phosphatidylinositol-4,5-bisphosphate 3-kinase catalytic subunit alpha
PR	Progesterone receptor
<i>Prom1</i>	Prominin 1
PyMT	Polyoma middle T antigen
Rb	Retinoblastoma-associated protein
RCAS-TVA	Replication-competent avian sarcoma-leukosis virus long-terminal repeat with splice acceptor - tumor virus A
ROI	Region of interest
rtTA	Reverse tetracycline-controlled transactivator

scRNA-seq	Single-cell RNA sequencing
SMA	Smooth muscle actin
Src	Proto-oncogene tyrosine-protein kinase Src
SV40-TAg	Simian virus 40 large T antigen
TDLU	Terminal duct lobular unit
TEB	Terminal end bud
TNBC	Triple-negative breast cancer
TP53	Tumor protein p53
Trp53	Transformation related protein 53
Ubc	Ubiquitin C
WAP	Whey acidic protein

INTRODUCTION

INTRODUCTION

1. Lineage tracing approaches: tools and applications

Lineage tracing is a powerful and versatile methodology in biological research, enabling the tracking of the origin, development, and fate of cell populations within living organisms [1]. By genetically labeling cells and observing their behavior over time, this technique provides critical insights into fundamental biological processes, including tissue development, regeneration, disease progression, and cancer formation [2-11]. The ability to trace cellular dynamics at both single-cell and population levels has made lineage tracing indispensable for advancing our understanding of developmental biology and pathology.

Lineage tracing is based on the premise that cells can be permanently labeled in a way that the marker is inherited by all their progeny. This allows researchers to map clonal expansion and differentiation pathways over time. The technique relies on genetic constructs that are specifically designed to label and visualize the targeted cells and their descendants, often through a two-component system.

Several lineage tracing strategies exist, each with distinct advantages and applications. This system uses two genetic constructs, typically delivered via two separate genetic lines:

1. The first genetic construct utilizes a lineage-specific promoter to drive the expression of an inducible recombinase, such as Cre recombinase or Flippase (Flp). The promoter ensures that recombinase activity is restricted to a specific cell type or lineage of interest.
2. The second genetic construct contains a reporter gene that enables the visualization of recombined cells and their progeny. This reporter is often designed to produce a detectable signal, such as fluorescence or enzymatic activity, which can be observed using microscopy, histology, or sequencing techniques. Common reporter genes include:
 - a. Fluorescent proteins: Green fluorescent protein (GFP) and its derivatives (e.g., YFP, CFP, RFP) are widely used for live-cell imaging and spatial tracking of labeled cells [12,13].
 - b. *LacZ*: The *LacZ* gene encodes β -galactosidase, an enzyme detectable through substrates like X-gal that generate a blue color in labeled cells [14].
 - c. Barcoding sequences: Unique DNA sequences can be introduced as labels and subsequently analyzed by high-throughput sequencing, providing quantitative insights into clonal dynamics and lineage relationships [15,16].

Lineage tracing remains a cornerstone of modern biological research, providing unparalleled insights into the behavior and fate of cells within complex tissues. Its applications in understanding cellular hierarchies, tissue regeneration, and disease mechanisms continue to expand, driven by ongoing innovations in genetic engineering and imaging technologies.

INTRODUCTION

1.1. Cre/lox systems for lineage tracing

The Cre/loxP system is one of the most widely used genetic tools for lineage tracing, owing to its precision, flexibility, and ease of application in various biological contexts. This system relies on the Cre recombinase enzyme, which catalyzes site-specific recombination between two loxP sites (**Figure I1**), leading to targeted modifications such as gene activation, inactivation, or permanent labeling of specific cell populations.

1.1.1. Tamoxifen-inducible model

The tamoxifen-inducible Cre system introduces temporal control into lineage tracing experiments. This system employs a modified form of the Cre recombinase, fused to a mutant ligand-binding domain of the human estrogen receptor (ER) (CreER, CreER^T, or CreER^{T2}) [17-19]. This ER domain ensures that Cre remains inactive in the cytoplasm until activated by synthetic ligands such as tamoxifen or 4-hydroxytamoxifen (4-OHT).

Once activated, CreER^{T2} translocates into the nucleus, where it excises or inverts a loxP-flanked STOP cassette, thereby inducing permanent expression of a reporter gene in the targeted cells (**Figure I1A**). The recombined locus is stably inherited by the progeny of these cells, enabling long-term tracking of cellular lineages. This feature has proven invaluable in studies of tissue homeostasis, regeneration, and cancer.

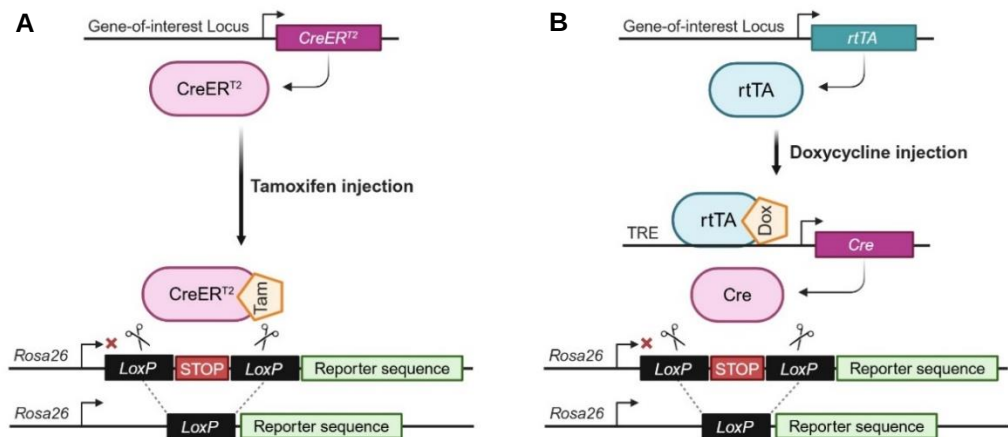


Figure I1. Two inducible Cre systems for gene activation in mouse models. **A.** In the tamoxifen-inducible CreER^{T2} system, CreER^{T2} is expressed at the locus of interest and activated by tamoxifen (Tam) injection, excising the LoxP-flanked STOP sequence to initiate reporter expression. **B.** In the doxycycline-inducible rtTA system, rtTA expression is triggered by doxycycline (Dox), leading to Cre activation and reporter expression. Figure created with BioRender.

Advantages of the tamoxifen-inducible model include activation at specific time points, enabling dynamic lineage tracking; minimal off-target recombination when appropriate promoters are used; and broad compatibility with various tissues and experimental setups. Practical considerations, such as optimizing tamoxifen concentrations, experimental timelines, and induction ages, have been largely resolved, making this approach highly versatile and effective for a variety of research applications beyond lineage tracing, such as studies involving mutation and gene deletion targeting.

INTRODUCTION

Despite its advantages, the use of tamoxifen presents some challenges. By inhibiting ER signaling, tamoxifen can delay the growth and expansion of the mammary gland during puberty. While low concentrations (up to 1.5 mg/kg) have minimal impact on mammary development [20], studies involving embryonic mammary development must account for tamoxifen's potential interference with natural pup delivery. To address this, researchers may opt for alternative methods such as C-sections or progesterone administration, though these alternatives carry additional risks to hormonal regulation and natural mammary growth [21].

1.1.2. Doxycycline-dependent models

The tetracycline-inducible system offers an alternative to tamoxifen-based approaches, providing robust temporal control of gene expression using doxycycline (Dox) or tetracycline [22,23]. This system utilizes a reverse tetracycline-controlled transactivator (rtTA), which, upon Dox administration, binds to tetracycline response elements (TRE), such as transcriptional activation elements (TetO), to activate the expression of Cre recombinase. Cre then excises the *loxP*-flanked STOP cassette, enabling lineage tracing (**Figure 11B**).

Key advantages of Dox-dependent systems include efficient labeling at saturation, as long-term Dox administration ensures that nearly all targeted cells within a lineage are labeled, overcoming the low recombination rates sometimes observed with tamoxifen-inducible systems [24]; and non-interference with endogenous ER, making it particularly useful in studies where tamoxifen's partial agonistic effects on ER could pose a challenge.

Moreover, Dox-inducible systems offer tighter control over gene expression than tamoxifen-inducible models. This is because tamoxifen has lower binding efficiency to CreER^{T2}, leading to variable and less efficient labeling. Consequently, rtTA-based systems are especially well-suited for applications requiring uniform labeling across a population or lineage, such as clonal dynamics studies or tissue regeneration experiments.

However, the main challenge of this approach lies in the requirement to generate triple-transgenic mice, which involves additional animals and an extra generation of breeding crosses to achieve the desired genotype, making it considerably more expensive and time-consuming.

1.2. Dre/rox systems

The Dre/rox recombination system is a complementary site-specific recombination system similar to Cre/*loxP* but uses a different recombinase (Dre) and recognition sites (*rox*) [25].

This system offers advantages in experiments requiring independent genetic manipulations in the same organism. For instance, the Dre/rox system can be paired with Cre/*loxP* to achieve dual recombination events, allowing researchers to selectively label and modify distinct populations, perform sequential recombination by introducing

INTRODUCTION

genetic changes stepwise to study their cumulative effects, and minimize crosstalk by ensuring that recombination events driven by Dre do not interfere with those mediated by Cre, thereby enhancing experimental flexibility [26-28].

By using dual recombinase-activated lineage tracing, researchers can enhance the precision of genetic lineage tracing using different strategies [29]:

- The **Interleaved Reporter (IR)** strategy arranges *loxP* and *rox* recombination sites in an interleaved manner, such as *loxP-rox-STOP-loxP-ZsGreen-STOP-rox-tdTomato*, ensuring sequential and cell-specific recombination. In the first cell type, Dre/*rox* recombination removes the *loxP* site between *rox* sites, activating tdTomato expression and preventing non-specific Cre/*loxP* recombination in these cells. Similarly, in the second cell type, Cre/*loxP* recombination removes the *rox* site between *loxP* sites, activating ZsGreen expression while blocking non-specific Dre/*rox* recombination. By labeling two distinct cell types with different fluorescent proteins, this strategy avoids ectopic Cre/*loxP* or Dre/*rox* recombination in non-specific cell types and ensures precise lineage tracing of cells where Dre or Cre recombination has occurred. This system was used to investigate the contributions of specific stem cells to different cell types. For instance, it clarified the roles of c-Kit^{pos} cardiac stem cells and Sox9^{pos} hepatic progenitor cells in tissue regeneration, enabling detailed insights into lineage fidelity and the occurrence of cell lineage conversion events [30].
- The **Nested Reporter (NR)** strategy utilizes nested *loxP* and *rox* recombination sites, such as *rox-loxP-STOP-loxP-ZsGreen-STOP-rox-tdTomato*, enabling simultaneous Cre and Dre recombination within the same cell. Similar to the IR system, this approach labels different cell types with distinct fluorescent proteins, with Dre recombination excluding Cre activity in non-targeted populations, ensuring precise lineage tracing. A key distinction of the NR system is its ability to switch fluorescent protein expression within the same cell. For example, a cell initially labeled with ZsGreen by Cre recombination can later switch to tdTomato upon Dre-mediated recombination. This capability makes the NR strategy particularly suited for tracking cells that change marker expression over time. An application of this strategy was demonstrated by Li *et al.*, who used the NR system to lineage trace mammary tumor cells undergoing epithelial-to-mesenchymal transition (EMT). Even when EMT gene activity was transient, Dre/*rox* recombination enabled permanent tdTomato labelling, allowing precise tracking of these cells throughout their progression [31].
- The **Intersectional** strategy facilitates double recombination, expressing a single reporter only in cells that undergo both Dre/*rox* and Cre/*loxP* recombination events. This is useful for studying double-positive cell populations defined by two genetic markers. For example, in the Ai66 reporter system, the construct is organized as *rox-STOP-rox-loxP-STOP-loxP-tdTomato*. In this setup, only double-positive cells will be labeled with tdTomato when both Dre and Cre recombinase activities occur. This system was employed by Liu *et al.* to specifically track contributions of CC10^{pos}SPC^{pos} bronchioalveolar stem cells to lung regeneration [32].

INTRODUCTION

In summary, the *Dre/rox* system has advanced multicolor lineage tracing by enabling, precise, cell-specific, and dynamic labeling of distinct or overlapping cell populations. These innovations have become invaluable tools in developmental biology, from tissue regeneration to cellular transitions and interactions, with unprecedented specificity.

However, the dual recombinase system has its limitations. One significant drawback is the limited availability of *Dre*-driver mouse lines and *rox*-flanked reporter lines. This constraint has restricted its application in certain tissues, including mammary gland studies [21]. Additionally, while effective, the complexity of managing two recombinase systems may introduce challenges in experimental design and data interpretation, making it less accessible for certain research applications.

1.3. Flippase/Frt system

The Flippase/*Frt* (Flp/*Frt*) system is a powerful genetic recombination tool that enables site-specific modifications within the genome. Originating from yeast, the Flp recombinase recognizes engineered *Frt* sequences to mediate recombination, resulting in excision, inversion, or integration of genetic material depending on the orientation of the *Frt* sites [33]. This precise control facilitates targeted manipulation of genomic loci in a cell-type or tissue-specific manner, making it a vital tool for lineage tracing, conditional gene expression, and functional studies across various biological contexts [34,35].

One of the key advantages of the Flp/*Frt* system is its independence from the *Cre/loxP* system, allowing the two to be combined for dual or sequential recombination events. This approach has enabled sophisticated genetic modifications, including targeted introduction of oncogenes in mammary epithelial cells to drive breast tumor formation [36,37], and in pancreatic cells to induce pancreatic cancer [38].

Additionally, the system offers high specificity due to its unique recognition sites, which minimize crosstalk with other recombinase systems. Temporal regulation can also be achieved using inducible Flp variants, such as FlpO-ER, which is activated through the administration of specific ligands [39].

Despite these strengths, the Flp/*Frt* system has some limitations. Flp recombinase exhibits optimal activity at lower temperatures, with recombination efficiency decreasing significantly at temperatures above 30°C, which can be a constraint in mammalian systems that operate at 37°C [40]. This temperature sensitivity, combined with variability in recombination efficiency and potential background activity, has positioned the *Cre/loxP* system as a more widely adopted due to its higher efficiency and broader range of available tools. Notably, there is compelling evidence indicating that the Flp/*Frt* system possesses considerably lower sensitivity compared to the *Cre/loxP* system [41].

Ongoing advancements, including the development of more efficient Flp variants and novel reporter constructs, continue to expand the utility of the Flp/*Frt* system. When integrated with emerging technologies such as CRISPR/Cas9, the Flp/*Frt* system holds promise for increasingly intricate genetic and cellular studies [42]. This flexibility and

INTRODUCTION

precision make the Flp/Frt system a valuable addition to the genetic toolbox, applicable to a wide range of experimental systems and model organisms.

1.4. Pulse-chase experiments

Lineage tracing approaches often utilize pulse-chase experiments to study cellular dynamics over time (**Figure 12**). In a pulse experiment, a short-term recombination event (16–24 hours) labels targeted cells, establishing a baseline population for subsequent analysis. In contrast, chase experiments track the labeled cells and their progeny over an extended period (days to weeks), allowing researchers to study their proliferation, differentiation, and long-term fate.

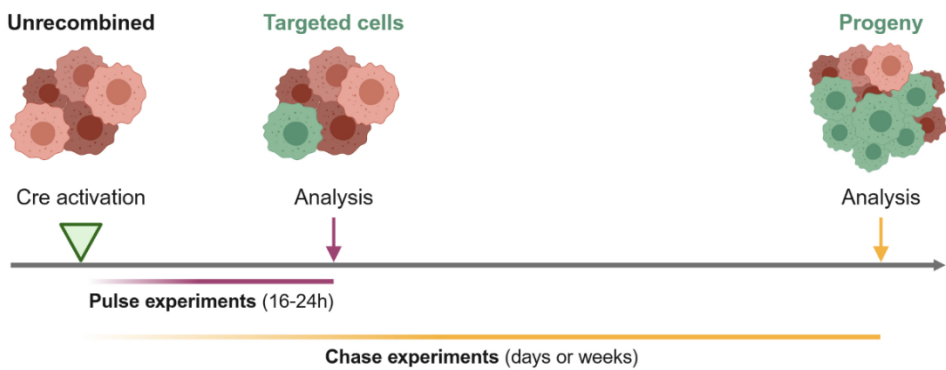


Figure 12. Schematic representation of pulse-chase experiments in lineage tracing. In pulse experiments (purple), Cre recombinase is transiently activated (16-24 hours) to label the targeted cells, establishing a baseline for subsequent tracking. In chase experiments (yellow), the initial Cre activation is followed by an extended observation period (days or weeks), during which the labeled cells and their progeny proliferate and differentiate. This method allows for detailed analysis of cellular dynamics over time, offering insights into cell lineage, fate, and differentiation processes. Figure created with *BioRender*.

Pulse-chase experiments offer invaluable insights into stem cell behavior by assessing the self-renewal and differentiation potential of progenitors. They also enhance our understanding of how specific cell populations contribute to the maintenance of tissue integrity. Furthermore, and particularly relevant to this thesis, these experiments play a crucial role in unraveling tumor heterogeneity by exploring the clonal dynamics and plasticity of cancer cells, shedding light on the mechanisms underlying their diverse behaviors and treatment responses.

Modern applications often integrate pulse-chase experiments with advanced imaging and sequencing technologies, enabling high-resolution tracking of cellular lineages and uncovering intricate details of tissue architecture.

INTRODUCTION

2. Physiology of the mammary gland

2.1. Cellular architecture of the mammary gland

The mammary gland is a highly specialized, branched epithelial exocrine organ primarily responsible for the production and secretion of milk, a process essential for the nourishment and survival of newborn mammals. Its remarkable structural complexity and dynamic adaptability enable it to meet the physiological demands of different reproductive stages, including puberty, pregnancy, lactation, and involution [43].

The gland consists of multiple cell types that work together to maintain its functional and structural integrity. The epithelial cells form the intricate ductal and alveolar network critical for milk transport and secretion. Meanwhile, the **stromal compartment** provides a supportive microenvironment [44]. The stroma is composed of:

- Adipocytes, which contribute to energy storage and secrete adipokines that are involved in gland function.
- Fibroblasts, which produce extracellular matrix components and contribute to structural support.
- Immune cells, including macrophages, neutrophils, and lymphocytes, which play roles in tissue remodeling, immune defense, and involution.
- Vascular endothelial cells, which form blood vessels that supply oxygen, nutrients, and hormones to the gland.

The **epithelial compartment** of the mammary gland is organized into a bilayer tubular structure [45]. This bilayer comprises (**Figure I3**):

- An outer **basal layer**, which includes basal progenitor cells, which serve as precursors to myoepithelial cells and play a role in epithelial maintenance; and myoepithelial cells, which are contractile cells responsible for milk ejection during lactation. These cells form a continuous layer adjacent to the basement membrane and provide structural support to the luminal cells.
- An inner **luminal layer**, composed of two distinct populations of luminal cells:
 - Hormone-sensing luminal cells, which express estrogen receptor alpha (ER α) and/or progesterone receptor (PR). These cells respond to hormonal signals and regulate the behavior of surrounding cells through paracrine signaling.
 - Hormone-responding luminal cells, which lack ER α /PR expression and include luminal progenitors and differentiated alveolar cells involved in milk production during lactation.

Approximately 10-15% of the luminal epithelial cells express ER α (ER α -positive or ER α^{pos}), and these cells are essential for glandular function. Through paracrine signaling, ER α^{pos} cells influence the proliferation and differentiation of nearby ER α -negative (ER α^{neg}) luminal cells, particularly during pregnancy and lactation [46,47].

INTRODUCTION

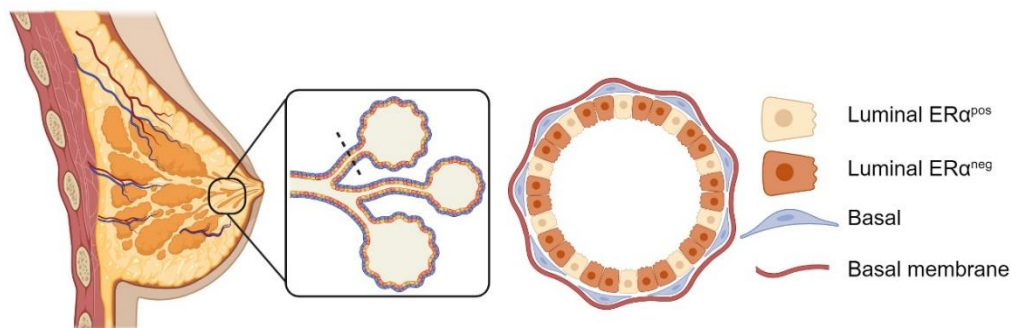


Figure 13. Schematic representation of the normal mammary gland architecture. The tissue is organized into ducts consisting of basal cells adjacent to the basal membrane, ERα^{pos} luminal cells, and ERα^{neg} luminal cells. The dotted black line delineates the cross-section of the mammary duct illustrated in the magnified inset on the right. Illustration adapted from *Vinuesa-Pitarch et al., 2021* [45].

The epithelial and stromal compartments are separated by a specialized extracellular matrix structure known as the basement membrane. This thin but robust barrier consists of proteins such as laminins, collagen IV, nidogens, and perlecan, providing both physical separation and biochemical signaling interfaces [48]. The basement membrane plays a critical role in maintaining the structural integrity and spatial organization of the gland, as well as in regulating signaling by mediating interactions between epithelial and stromal cells that influence gland development, function, and repair [49-51].

During lactation, the coordinated activity of the epithelial layers becomes critical. Milk synthesis occurs in the alveolar luminal cells, which are specialized for this function. Once synthesized, milk is stored in the alveolar lumen and transported through the ductal network. The outer myoepithelial cells contract in response to oxytocin, a hormone released during nursing, to expel milk from the alveoli into the ductal system for delivery to the offspring [52].

This highly orchestrated process underscores the importance of cellular coordination and structural integrity in maintaining the functionality of the gland during lactation.

2.2. Development and homeostasis of the mammary gland

The mammary gland is a highly dynamic organ that begins its development during the embryonic stage and completes its morphological and functional maturation postnatally. The development of the mammary gland occurs across distinct life stages, including puberty, pregnancy, lactation, and involution. Each stage demonstrates the capacity of the gland to adapt its structure and function to meet the reproductive requirements of the individual [53] (**Figure 14**).

INTRODUCTION

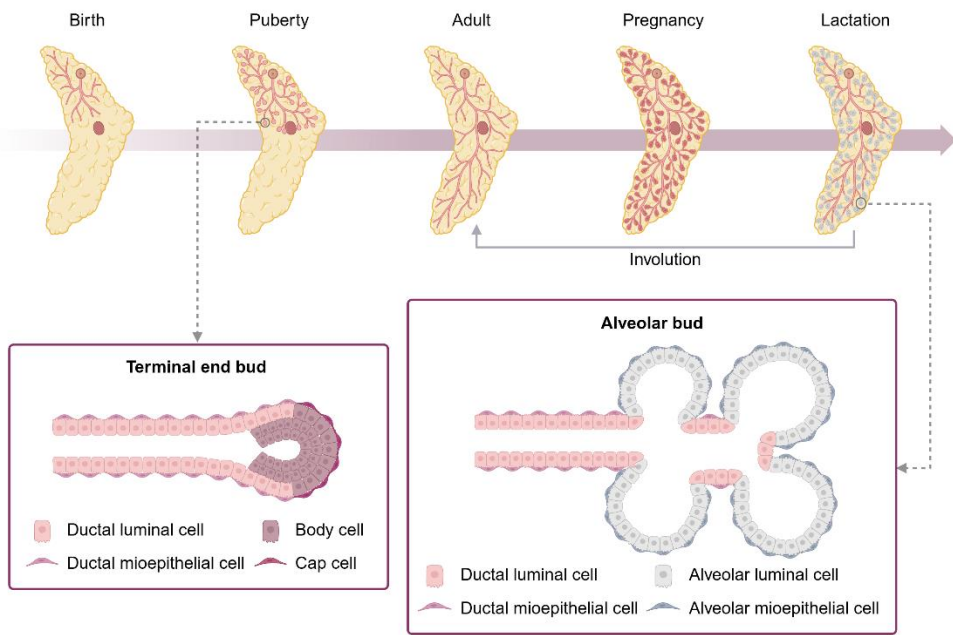


Figure 14. Schematic representation of mammary gland development across life stages. The mammary gland undergoes dynamic remodeling from birth through puberty, adulthood, pregnancy, lactation, and involution. Below, the structures involved in mammary gland morphogenesis are highlighted: the terminal end bud (TEB), responsible for ductal elongation and branching during puberty, and the alveolar bud, which clusters to form the milk-secreting alveoli during lactation. Cell types include ductal luminal and myoepithelial cells, body cells, cap cells, and alveolar luminal and myoepithelial cells. Illustration adapted from *Wu et al., 2022* [54].

2.2.1. Embryonic development

The embryonic development of the mammary gland involves a series of tightly regulated stages. It begins with the formation of mammary placodes, localized thickenings of the surface ectoderm along the mammary line, a region extending along the flank of the embryo. These placodes invaginate into the underlying mesenchyme to form mammary buds, which subsequently undergo branching morphogenesis, leading to the formation of a primitive ductal system [53,55].

This process is orchestrated by a complex interplay of signaling pathways, including Wnt signaling, which regulates epithelial-mesenchymal interactions and mammary line specification [56,57]; FGF (fibroblast growth factor), which drives placode formation and early ductal outgrowth [58,59]; and BMP (bone morphogenetic protein), which modulates epithelial differentiation and positioning of the mammary buds [60]. Together, these pathways ensure proper spatial and temporal organization of the nascent mammary gland and establish the framework for postnatal development.

2.2.2. Postnatal development during puberty

Puberty marks a period of rapid and extensive development of the mammary gland, driven by hormonal and growth factor cues. This stage is characterized by the formation

INTRODUCTION

of terminal end buds (TEBs) at the tips of elongating ducts. These highly dynamic, bulbous structures are essential for ductal elongation and branching into the mammary fat pad [61,62]. TEBs consist of two main cell populations (**Figure I4**):

- **Cap cells**, which form the outer layer of TEBs and act as myoepithelial progenitors, guiding ductal elongation.
- **Body cells**, located beneath the cap cells, include rapidly dividing luminal progenitor cells responsible for ductal growth.

Under the influence of growth hormone and estrogen, TEBs invade the mammary fat pad, resulting in the development of a complex, tree-like ductal architecture embedded within the stromal matrix [63]. This phase establishes the primary ductal network, which will later support functional differentiation during pregnancy.

2.2.3. Development during pregnancy

The transition to pregnancy triggers extensive remodeling of the mammary gland to prepare it for lactation. Repeated estrous cycles prime the ductal system, which undergoes side branching to increase its complexity. By mid-to-late pregnancy, alveologenesis occurs, characterized by the formation of alveolar buds, which give rise to alveoli [64] (**Figure I4**). These alveoli emerge and aggregate into clusters, forming functional units known as terminal duct lobular units (TDLUs) in humans, while lobuloalveolar structures are present in mice [65]. Each functional unit consists of terminal ducts, which serve as conduct for milk, and associated lobules, which house clusters of alveoli that differentiate into milk-secreting units during lactation [66].

Hormones such as prolactin, progesterone, and placental lactogen play critical roles during this stage, stimulating the growth and differentiation of the alveolar epithelium [53]. The result is a mammary gland primed for lactation, with a highly specialized architecture optimized for milk production and delivery.

2.2.4. Lactation

Lactation is a finely regulated process involving the synthesis, secretion, and ejection of milk. This process begins with the preparation of the mammary gland during pregnancy, when alveoli, which are expanded during this stage, differentiate into specialized milk-producing cells. This differentiation is primarily driven by prolactin, which stimulates the maturation of alveolar luminal cells to produce and store milk components in readiness for lactation [63].

Following parturition, lactation is initiated and maintained through a combination of hormonal and mechanical stimuli [67]:

- **Prolactin** ensures the continued production of milk by stimulating mammary epithelial cells (MECs) to synthesize essential milk components, including proteins, lipids, and carbohydrates.
- **Oxytocin**, released in response to suckling, induces contraction of myoepithelial cells surrounding the alveoli. This action facilitates the ejection of milk from the alveoli into the ductal system and out through the nipple.

INTRODUCTION

This coordination between milk synthesis and ejection ensures a steady and reliable supply of milk to meet the nutritional needs of offspring. Additionally, the mechanical stimulation of suckling reinforces milk production through a feedback mechanism, sustaining lactation over the nursing period.

2.2.5. Involution

After weaning, the mammary gland undergoes involution, a tightly regulated process aimed at restoring the gland to a near pre-pregnant state [68]. This process involves:

1. **Apoptosis:** Programmed cell death of milk-secreting epithelial cells, leading to the regression of glandular tissue.
2. **Tissue remodeling:** Reorganization of the extracellular matrix and glandular architecture, resulting in the resorption of expanded alveolar structures and ductal branches.

An immune response is integral to this phase, with macrophages clearing apoptotic cells and debris to facilitate remodeling. Following these changes, a period of regeneration often occurs, during which epithelial cells and stromal components reorganize to restore the baseline architecture of the gland, preparing it for future reproductive cycles.

The cyclical processes of development, functional differentiation, and regression underscore the remarkable plasticity of the mammary gland. Its ability to undergo extensive structural and functional changes in response to hormonal and physiological cues ensures its adaptability to the reproductive status of the individual. These processes highlight the mammary gland as a model of tissue plasticity and regeneration, offering insights into both normal physiology and pathological conditions, such as cancer.

2.3. Hierarchical organization of the healthy mammary gland

Numerous transplantation studies have sought to identify mammary stem cell populations by isolating MECs and injecting them into cleared mammary fat pads at limiting dilutions [4,69-72]. Pioneering research demonstrated that different subsets of basal cells (BaCs) efficiently reconstitute the mammary gland, suggesting the presence of multipotent mammary stem cells responsible for tissue maintenance [2,70-72]. However, despite the higher transplantation efficiency of BaCs, subsequent studies revealed that specific subsets of luminal cells (LCs) also contribute to mammary reconstitution [4,69]. These experiments setups often involve non-physiological conditions that may confer MECs with stem cell-like properties not typically observed in normal development, suggesting that they measure cell plasticity rather than inherent multipotency [73].

Over the past decade, lineage tracing has become the preferred method for defining and understanding the cellular hierarchies within the mammary gland. This strategy has been instrumental in assessing clonal expansion, cell fate, and stem cell potency across multiple tissues, such as intestine, prostate, skin, mammary gland, lung, muscle, brain, and blood [2,74-80].

INTRODUCTION

In the context of the mammary gland, several genetic promoters have been employed to address the question about cellular hierarchies (**Figure 15**):

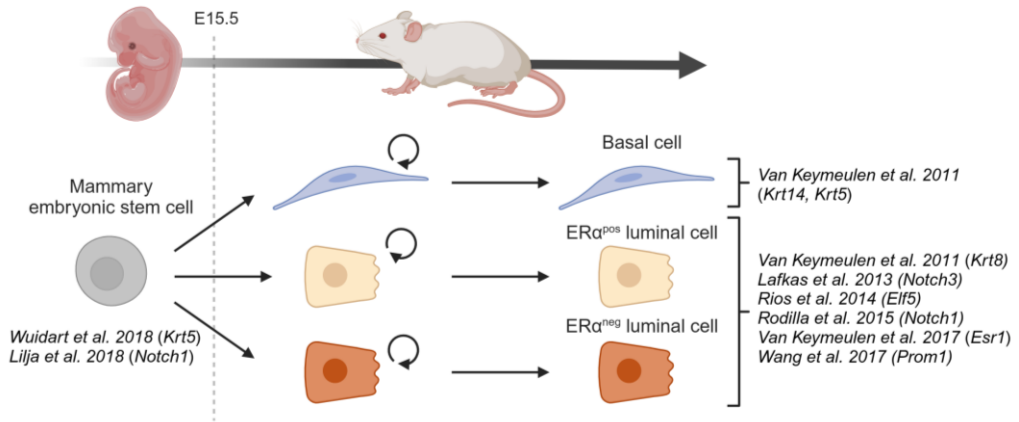


Figure 15. Hierarchical structure of the mammary gland throughout development. Lineage tracing studies using different lineage-specific promoters have provided evidence for the adult mammary gland being maintained by different pools of unipotent mammary epithelial progenitors that self-renew within their respective lineages, whereas multipotent stem cells, with the ability of giving rise to both luminal and myoepithelial cells, have been found exclusively during embryonic development. By embryonic day 15.5 (E15.5), there is no longer evidence of multipotent stem cells in the mammary gland. Figure created with BioRender.

- **Ubiquitous promoters:** *Rosa26-CreER^{T2}* mice, using a ubiquitous promoter, have been applied for unbiased labeling of single proliferating cells, enabling broader insights into the cellular behavior across all compartments [81,82].
- **Basal cell markers:** Promoters such as *Acta2* (actin alpha 2, smooth muscle, aorta), *Krt5* (keratin 5), and *Krt14* (keratin 14), predominantly label postnatal BaCs that contribute to the myoepithelial layer across mammary gland life stages, confirming the existence of long-lived unipotent basal progenitors in the virgin state and throughout pregnancy [2,83]. Nevertheless, clonal analysis using *Dll1* (delta like canonical Notch ligand 1)- *CreER^{T2}*, *Lgr5*- and *Lgr6-CreER^{T2}* lines (leucine rich repeat containing G protein coupled receptors 5 and 6, respectively), failed to provide a clear consensus on the existence of unipotent or multipotent mammary stem cells, as these genes are predominantly expressed in BaCs, but are also found in some LCs [7,84,85].
- **Luminal cell markers:** Lineage tracing studies by Van Keymeulen *et al.* using the *Krt8* (keratin 8) promoter identified a luminal population of long-lived unipotent stem cells capable of clonal expansion across multiple pregnancy cycles, consistently maintaining their luminal identity. Conversely, LCs targeted by *Krt18* (keratin 18) represent a more committed subset of luminal progenitors, as they fail to persist through pregnancy and lactation [2]. Similarly, Rios *et al.* showed that *Elf5* (E74-like factor 5)-labeled luminal progenitors give rise to alveolar LCs during the first pregnancy, although these cells die off during involution and must be replaced with a new pool of progenitors for subsequent pregnancies [20]. Moreover, promoters for *Notch1*, *Notch3*, *Sox9* (SRY (sex determining region Y)-box 9), *Prom1* (prominin

INTRODUCTION

1), and *Esr1* (estrogen receptor 1, alpha) have provided valuable insights into lineage-specific maintenance. Specifically, *Notch3* labels both ER α^{pos} and ER α^{neg} LCs, capturing the entirety of the luminal compartment [6]. *Esr1* and *Prom1* selectively mark ER α^{pos} LCs, highlighting their unique lineage specificity [86,87]. In contrast, *Notch1* exclusively targets ER α^{neg} LCs in postnatal glands, emphasizing the compartmentalized nature of luminal hierarchies [4]. Of note, *Notch1*-labeled cells represent long-lived unipotent stem cells that could survive multiple successive involutions, demonstrating extensive self-renewal capacity. However, *Sox9*-expressing cells contribute to the maintenance of ER α^{neg} LCs even during several pregnancies, while also sustaining a subset of BaCs to a lesser extent, thus compromising the lineage specificity of the *Sox9* promoter [86].

Despite the extensive research supporting the unipotency of the three mammary epithelial compartments of the mammary gland – BaCs, ER α^{neg} LCs, and ER α^{pos} LCs – a study revealed rare BaCs capable of generating both basal and luminal progeny in the healthy mammary gland, suggesting a potential for multipotency [20]. These findings, though intriguing, were subject to discrepancies largely due to variations in the regions of the *Krt5* promoter used, which altered its specificity for different cell types [73]. Moreover, the study faced criticism for its lack of short time periods demonstrating clear targeting of BaCs at the initial timepoint. This limitation hindered the ability to draw definitive conclusions regarding the dynamic behavior and potential multipotent capabilities of BaCs.

Recent studies clarified these conflicting results using clonal analysis at saturation with *Krt14^{rtTA}-Cre^{TetO}* and *Krt8^{rtTA}-Cre^{TetO}* systems, showing no evidence of multipotent stem cells in the postnatal mouse mammary gland [24]. In contrast, multipotent stem cells, capable of generating both LCs and BaCs, are found exclusively during embryonic development [5,88]. This transition from multipotency to unipotency begins as early as embryonic day 12.5 (E12.5) and is nearly complete by E15.5, with no significant evidence of bipotency persisting beyond this stage [5] (**Figure 15**).

Collectively, these studies demonstrate that both BaCs and LCs are self-sustained by distinct unipotent progenitors, with limited evidence of interconversion between these lineages in the adult homeostatic mouse mammary gland. Within the luminal compartment, ER α^{pos} and ER α^{neg} luminal subsets represent two independent lineages, consistently maintaining their distinct identities during physiological mammary gland development. Importantly, all these populations continue to uphold their respective lineages into adulthood, even after hormonal fluctuations and multiple pregnancies, demonstrating their capacity for long-term self-renewal.

2.4. Mammary epithelial cell plasticity

MECs, despite their lineage commitment under normal physiological conditions, exhibit remarkable plasticity in response to various stressors. This plasticity allows MECs to adapt and reprogram their fate, challenging the traditional view of strict unipotency within adult epithelial lineages. Evidence of this adaptability has been observed under

INTRODUCTION

diverse conditions, including transplantation [2,4], cellular ablation [3], hormone stimulation [89], ectopic expression of fate-determining factors [5,88], genotoxicity [90], and oncogene activation [91-94].

2.4.1. Plasticity in transplantation and cellular ablation

Transplantation assays have been pivotal in demonstrating the plastic potential of mammary epithelial cells. *Van Keymeulen et al.* demonstrated that isolated *Krt14*-positive myoepithelial cells, but not *Krt8*-positive LCs, can regenerate a fully functional mammary gland comprising both basal and luminal compartments [2]. This finding indicates that under transplantation conditions, BaCs can exhibit bipotent behavior. Moreover, subsequent research by *Rodilla et al.* reported that *Notch1*-positive LCs exhibit low regenerative capacity in virgin females, although their repopulating capacity significantly increases in parous females under hormonal stimulation [4]. Conversely, co-transplantation of BaCs and LCs preserved their unipotency, indicating that MECs retain their lineage-restricted differentiation when both unipotent stem cell types coexist, as observed in unperturbed mammary glands [2,4].

Similarly, in genetic ablation experiments, the removal of LCs promoted a shift in BaC fate, enabling them to adopt multipotent behavior. This plasticity was shown to be constrained in the presence of LCs, driven by their secretion of TNF (tumor necrosis factor), which restricted BaC lineage reprogramming [3]. Notably, during this transition, BaCs adopted an intermediate hybrid basal-luminal state before fully differentiating into LCs. This process reactivated a genetic program reminiscent of embryonic multipotent stem cells, underscoring the ability of BaCs to revert to a more primitive state under specific conditions [88].

2.4.2. Hormonal influence on lineage plasticity

Hormonal stimulation during pregnancy provides another example of mammary epithelial cell plasticity. Endogenous ovarian hormones induced basal-like cell formation from *Krt8*-expressing LCs. Remarkably, these basal-like cells persisted through involution and expanded during subsequent pregnancies. This demonstrates the profound influence of hormonal signals on lineage reprogramming and highlights the dynamic nature of MECs across reproductive cycles [89].

2.4.3. Ectopic expression of fate-determining factors

Studies exploring the role of specific transcription factors have further elucidated the mechanisms underlying MEC plasticity. Functional studies have identified distinct fate-specification factors that regulate lineage commitment within the mammary gland:

- **Basal specification factors.** Overexpression of P63 (tumor protein 63) reprograms adult LCs into BaCs by inducing a multipotent embryonic-like hybrid state before differentiation into fully functional BaCs [88].
- **Luminal specification factors.** *Gata3* (GATA binding protein 3) deletion impairs the expression of luminal differentiation markers in LCs, disrupting TEB formation during puberty and causing improper development of milk-secreting alveoli during gestation, ultimately resulting in lactational defects [95]. Loss of C/EBP β

INTRODUCTION

(CCAAT/enhancer binding protein beta) in the adult mammary gland disrupts luminal progenitor identity and results in aberrant basal marker expression in the luminal compartment [96]. Activation of *Notch1* dictates the transition of BaCs to fully committed LCs during pubertal development [5].

These findings underscore the ability of specific molecular cues to override established lineage commitments, demonstrating that ectopic expression or suppression of key transcription factors can reprogram MECs into alternative cell fates.

2.4.4. Plasticity induced by genotoxic stress and chemotherapy

Genotoxic stress and exposure to chemotherapeutic agents have also been shown to drive MEC plasticity. Chemotherapy has been associated with the proliferation of myoepithelial cells, accompanied by transitions from basal to luminal fates [90]. This finding suggests that therapeutic interventions can inadvertently influence cellular hierarchies, potentially contributing to altered tissue dynamics or even tumorigenesis.

The plasticity of mammary epithelial cells challenges traditional hierarchical models of unipotent progenitors, highlighting their dynamic nature and ability to adapt under various physiological and pathological conditions.

2.4.5. Lineage infidelity induced by oncogene activation

In cancer research, two independent groups showed the generation of luminal ER α^{pos} tumors upon expression of the oncogenic form of phosphatidylinositol-4,5-bisphosphate 3-kinase catalytic subunit alpha, *Pik3ca*^{H1047R}, in both BaCs and LCs [92,93]. In the same direction, directing *Brca1* (breast cancer 1, early onset) and *Trp53* (transformation related protein 53) deficiency to luminal progenitors enabled these cells to adopt a basal-like phenotype, phenocopying human *BRCA1* (BRCA1 DNA repair associated) loss-of-function breast cancer [91]. Moreover, LCs gave rise to BaCs during tumorigenesis when either PyMT (polyoma middle T antigen) or *ErbB2* (receptor tyrosine-protein kinase erbB-2) signaling was activated [94]. Additional studies demonstrated the generation of tumors with hybrid and/or basal phenotypes when different oncogenes were overexpressed in LCs, *Etv6-Ntrk3* (Ets variant 6 - neurotrophic tyrosine kinase, receptor, type 3), *Notch1* intracellular domain (N1ICD), and *Kras*^{G12D}, the mutant human form of Kirsten rat sarcoma viral oncogene homolog [97-99].

Those studies suggested that adult cells can undergo reprogramming to a multipotent stem cell state during tumorigenesis, supporting the concept that cancers may arise from a reactivation of embryonic developmental programs in postnatal tissues.

INTRODUCTION

3. Breast cancer

3.1. Breast cancer classification and therapeutics

3.1.1. Breast cancer morphological classification

Breast cancer stands as the most common diagnosed cancer among women and accounts for approximately 6.9% of all cancer-related deaths worldwide [100]. Most breast carcinomas originate within the TDLU [101], and are classified morphologically by their degree of invasion and cohesiveness [102,103]:

3.1.1.1. Localized breast cancer types

Breast cancer includes non-invasive forms confined to the ducts or lobules, such as ductal carcinoma in situ and lobular carcinoma in situ:

- **Ductal carcinoma in situ (DCIS):** DCIS represents the most prevalent localized breast cancer, comprising approximately 10% of all breast cancers diagnosed [104]. This condition involves abnormal cell growth confined to the milk ducts without breaching the basement membrane or infiltrating surrounding stromal tissues. Frequently identified through mammographic screenings, DCIS typically presents as clusters of microcalcifications, which are small clusters of calcium deposits serving as early signs of atypical ductal activity, rather than symptomatic manifestations such as lumps or nipple discharge [105]. Although DCIS itself is non-invasive, it is widely regarded as a precursor to invasive ductal carcinoma if left untreated [106].
- **Lobular carcinoma in situ (LCIS):** LCIS is a less common localized lesion originating in the breast lobules, accounting for 1-2% of breast cancer cases. Unlike DCIS, LCIS often eludes detection through routine mammographic imaging due to its lack of calcifications or distinct masses, making it an incidental finding frequently identified during breast biopsies conducted for other reasons. The presence of LCIS significantly raises the lifetime risk of developing both invasive ductal and invasive lobular carcinomas, necessitating vigilant monitoring and consideration of risk-reduction strategies [107].

3.1.1.2. Invasive breast cancer types

Invasive breast cancer types are characterized by the ability of tumor cells to invade surrounding tissues, comprising a range of distinct histological subtypes with unique clinical and pathological features:

- **Invasive ductal carcinoma of no special type (IDC-NST):** The most common type of invasive breast cancer, IDC-NST, constitutes approximately 60% of breast cancer cases [108]. The term 'no special type' indicates that it is characterized by a diverse range of cell morphology and the presence of tubular or glandular structures, with an absence of growth patterns or cytological features typical of special histological subtypes [109]. It arises in ductal tissues, invades surrounding breast tissue, and typically forms a solid mass, often detectable as a palpable lump. IDC-NST is known for its potential to metastasize to regional lymph nodes and

INTRODUCTION

distant organs, with common sites of spread including the bones, lungs, liver, and brain [110].

- **Special types of invasive breast cancer:** Collectively sum up to around 25% of cases and include distinct subtypes with varying histopathology [109]:
 - Invasive lobular carcinoma (ILC): The most common special type of breast cancer, ILC, represents up to 15% of cases [109]. It originates in the lobules, the milk-producing glands, and invades surrounding tissues in a less cohesive, more dispersed pattern compared to IDC-NST. This diffuse growth pattern makes ILC less visible on imaging and more challenging to detect early [111]. ILC tends to metastasize to unusual sites like the gastrointestinal tract and ovaries, alongside common sites like bones, while showing reduced tendency to metastasize to the liver and lungs [112]. ILC is associated with poorer prognosis and a less favorable response to chemotherapy compared to IDC-NST [113,114].
 - Rare special subtypes: These subtypes represent a smaller subset of breast cancer cases, showcasing unique histological and molecular characteristics that contribute to the heterogeneity of breast cancer, each requiring specific diagnostic and treatment approaches [109]:
 - Mucinous carcinoma: Features nests of cells in a mucinous matrix, typically associated with older women and low incidence of lymph node involvement [115].
 - Mucinous cystadenocarcinoma: Characterized by cystic structures with papillae and abundant extracellular mucin. It commonly presents as a palpable mass, often detected in post-menopausal women. It typically has well-defined margins on imaging and a favorable prognosis, with uncommon nodal involvement [116].
 - Cribriform carcinoma: Features a distinctive cribriform growth pattern with well-defined cell nests in a glandular structure. It shows rare axillary metastases and has a 10-year survival rate of about 90% [117].
 - Micropapillary carcinoma: Composed of tumor cells growing in a papillary-like structure without a central fibrovascular core. It is associated with a higher risk of lymph node metastases, despite a relatively favorable prognosis [118].
 - Papillary carcinoma: Features well-defined, finger-like projections into cystic spaces with a fibrovascular core. It is often low-grade with a good prognosis, rarely metastasizing, and commonly affecting post-menopausal women [119].
 - Tubular carcinoma: Distinguished by its well-formed, tube-like structures with excellent prognosis and rare lymph node metastases [120].
 - Medullary carcinoma: Characterized by a syncytial growth pattern, a dense infiltrate of immune cells, and clear tissue boundaries. It often presents as a palpable lump and is associated with a favorable prognosis, with a 5-year overall survival rate of approximately 89% [121].

INTRODUCTION

- Metaplastic carcinoma: Exhibits mixed differentiation with areas of squamous or mesenchymal differentiation, often associated with a more aggressive clinical course. Imaging shows large, oval-shaped masses with high density and indistinct margins. Prognosis varies, with a 5-year overall survival between 64% and 83% [122].
- Apocrine carcinoma: Contains cells with distinct apocrine features, including abundant eosinophilic cytoplasm and apical snouts. Imaging shows irregular masses with poorly defined margins and associated microcalcifications. It has a varied prognosis depending on grade [123].
- Neuroendocrine carcinoma: Composed of cells with neuroendocrine differentiation, often showing a more aggressive behavior, particularly in advanced stages. It appears as irregular, poorly defined masses and may be associated with skin thickening and swollen lymph nodes [124].
- Mixed invasive ductal and lobular carcinomas: These tumors combine features of both IDC and ILC and generally have a better prognosis than pure ILC tumors, accounting for 3-5% of all breast cancers [125,126].

3.1.2. Breast cancer subtypes and associated therapies

Therapeutic decision-making for breast cancer relies on the immunophenotypic classification, which evaluates the expression of key molecular markers such as ERα, PR, and the human epidermal growth factor receptor 2 (HER2) [127,128]. This classification categorizes tumors into hormone receptor-positive, HER2-positive and triple-negative subtypes, as further detailed below (Figure I6).

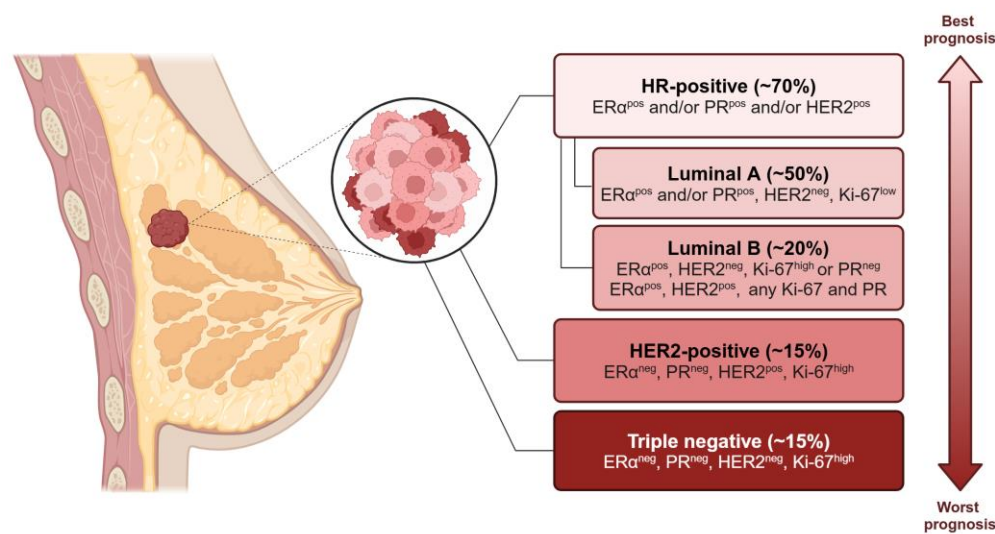


Figure I6. Immunohistochemical classification of breast cancer based on receptor status: ERα, PR, and HER2. Hormone receptor (HR)-positive subtypes include Luminal A and Luminal B, distinguished by differing levels of HER2 and the proliferation marker Ki-67. The HER2-enriched subtype is characterized by HER2 overexpression, while triple-negative breast cancer lacks expression of ERα, PR, and HER2. Illustration adapted from *Vinuesa-Pitarch et al., 2021* [45].

INTRODUCTION

3.1.2.1. Hormone receptor (HR)-positive subtypes

Luminal A is the most common breast cancer subtype, making up about half of newly diagnosed breast cancer cases [103]. It is defined by being ER α ^{pos} ($\geq 1\%$), PR^{pos} ($\geq 20\%$), HER2^{neg} ($\leq 10\%$), and having low levels of proliferation marker protein Ki-67 ($<14\%$) [129,130]. These tumors are typically low-grade and resemble luminal epithelial cells of the normal breast, showing high expression of K7/8/18/19 luminal keratins [103]. They generally have a favorable prognosis, a slow progression, and typically they are of low-grade with limited lymph node involvement [131]. Patients benefit from hormone therapies either with estrogen receptor modulators (tamoxifen) or with aromatase inhibitors (anastrozole) [132].

The **luminal B** subtype comprises 20% of invasive breast cases [103]. This subtype expresses keratins from the luminal compartment and can be further categorized into luminal B HER2-negative: ER α ^{pos} ($\geq 1\%$), HER2^{neg} ($\leq 10\%$) and at least one of the following: high Ki-67 levels ($\geq 20\%$), PR^{neg} or $< 20\%$, or high recurrence risk based on multi-gene expression array, if available; or luminal B HER2-positive; ER α ^{pos} ($\geq 1\%$), HER2^{pos} ($> 10\%$) or amplified, and any level of PR and Ki-67 [129]. Luminal B subtype is considered the most aggressive hormone-dependent breast cancer, requiring additional chemotherapy to the hormonal treatment for HER^{pos/neg} cases and additional HER2-targeted therapy for HER2^{pos} patients [133-136]. Luminal B tumors present higher histological grades and an increased proliferation index rate expressed by Ki-67 compared to luminal A tumors [131]. They exhibit higher expression of proliferation-related genes, such as *CCND1* (cyclin D1) and *CCNE2* (cyclin E2), compared to luminal A tumors, which is associated with poorer clinical outcomes in tamoxifen-treated women [137,138]. Additionally, these tumors show activation of PI3K (phosphatidylinositol 3-kinase) [139], a mechanism known to contribute to acquired endocrine resistance [140].

3.1.2.2. HER2-positive subtypes

The HER2-positive subtype represents 15% of newly diagnosed breast cancer cases [103]. This subtype is defined by HER2 overexpression, ER α and PR negativity, and high expression of Ki-67 ($>20\%$) [129], by immunophenotypic characterization, and HER2 amplification by fluorescence in situ hybridization (FISH) [141]. These tumors are more likely to be high-grade and associated with lymph node metastases [103]. HER2^{pos} patients benefit from HER2-targeted therapies that block HER2 activity, such as Herceptin (trastuzumab), Perjeta (pertuzumab), Tykerb (lapatinib), and Kadcyła (T-DM1 or ado-trastuzumab emtansine) [142,143].

3.1.2.3. Triple-negative breast cancer (TNBC)

TNBC, accounting for 15% of all breast cancers, is characterized by the lack of ER α , PR, and HER2 expression, being more prevalent among younger women and patients with *BRCA1* mutations [144-146]. This subtype is considered the most aggressive subtype, with a high proliferation index rate of Ki-67, a higher grade, and an elevated metastatic incidence to the brain, liver and lungs [146-148]. The aggressive nature of

INTRODUCTION

TNBC, combined with the lack of targeted therapies, results in a poorer prognosis and a higher rate of distant recurrence compared to other breast cancer subtypes [146].

Non-surgical treatment of TNBC has been extensively limited to chemotherapy, which benefits only 30-40% of patients with early-stage TNBC [149]. Common chemotherapy combinations include the AC-T regimen, composed of an anthracycline (doxorubicin or epirubicin) plus cyclophosphamide, followed by a taxane (paclitaxel or docetaxel), both in neoadjuvant and adjuvant settings [150-152]. For patients who are not suitable for standard chemotherapy regimens involving anthracyclines and taxanes, or those who refuse treatment involving alopecia, a non-anthracycline regimen, including paclitaxel plus carboplatin, may offer a viable alternative [153].

The evolving understanding of TNBC biology has spurred the development of innovative therapies, such as poly (ADP-ribose) polymerase (PARP) inhibitors, antibody-drug conjugates (ADCs), and immunotherapy agents. These advancements are transforming the treatment landscape, offering new possibilities for both early- and late-stage patients with TNBC. Specifically, PARP inhibitors, such as olaparib, have been approved for patients with *BRCA1* and *BRCA2* (*BRCA2* DNA repair associated) mutations, either as monotherapy or in combination with a carboplatin/paclitaxel regimen [154,155]. Additionally, immune-checkpoint inhibitors, such as atezolizumab and pembrolizumab have been authorized for use in combination with chemotherapy [156,157]. Another important addition is sacituzumab govitecan, an ADC targeting human Trop-2 (trophoblast cell-surface antigen 2), granted for patients with metastatic TNBC who have received at least two lines of therapy [132]. Trop-2 is overexpressed in TNBC, and its linkage to SN-38 enhances the drug's ability to deliver chemotherapy specifically to cancer cells.

While this classification provides a practical framework for treatment, it does not fully capture the molecular heterogeneity of breast cancer. Variations in tumor biology, even within the same subtype, contribute to diverse clinical outcomes, underscoring the need to integrate additional molecular and genomic insights to refine treatment strategies and improve prognostic accuracy.

3.1.3. Advances in molecular subtyping

Recent advances in gene expression profiling have significantly refined breast cancer classification. Using complementary DNA (cDNA) microarrays, *Perou et al.* identified five intrinsic molecular subtypes of breast tumors: luminal A, luminal B, HER2-enriched, basal-like, and normal. These molecular subtypes exhibit distinct gene expression profiles and correlate with varying clinical outcomes, with basal-like and HER2-enriched subtypes generally associated with poorer prognoses [158,159]:

- **Luminal tumors:** This category is characterized by gene clusters associated with ER α and luminal-specific genes [158]. Luminal tumors exhibit considerable heterogeneity and are further divided into two subgroups:
 - **Luminal A tumors:** Distinguished by high ER α protein expression and low proliferation rates, these tumors are associated with favorable clinical outcomes

INTRODUCTION

and strong responses to hormone therapy [130,160]. They generally exhibit low *ErbB2* (erb-b2 receptor tyrosine kinase 2) expression and more homogeneous gene expression profiles compared to other subtypes [159].

- **Luminal B tumors:** These tumors display lower ER α expression and higher proliferation rates compared to luminal A tumors, contributing to worse clinical outcomes and increased endocrine resistance [130,161,162]. Their gene expression profiles overlap with HER2-positive and basal-like subtypes, reflecting a more aggressive biological phenotype [159].
- **HER2-enriched tumors:** Defined by low ER α expression and overexpression of *HER2/ERBB2*, HER2-enriched tumors are known for their aggressive nature and high proliferation rates, often associated with a poor prognosis [159]. Notably, this subtype encompasses a diverse range of molecular characteristics, including varying levels of hormone receptor expression (ER α and PR) and distinct genetic alterations [163]. This heterogeneity results in a spectrum of biological behaviors, making them complex in terms of clinical management. While HER2-targeted therapies are central to treatment, additional molecular features often influence therapeutic strategies and outcomes [164-168].
- **Basal-like tumors:** Characterized by the expression of basal-specific genes, including keratin 5, keratin 17, integrin- β 4, and laminin [158], basal-like tumors are associated with high-grade features and an aggressive clinical course [169]. These tumors are frequently triple-negative, lacking ER α , PR, and HER2 expression. The absence of effective targeted therapies contributes to their poor prognosis and elevated risk of metastasis [159,170].
- **Normal-like tumors:** This rare subtype is characterized by high expression of basal epithelial and adipose cell genes but lower expression of luminal epithelial markers. Normal-like tumors generally present a more favorable prognosis compared to basal-like or HER2-enriched subtypes [158,159].

Building on these intrinsic molecular subtypes, *Parker et al.* developed *Prosigna*, a powerful tool for breast cancer classification based on the prediction analysis of microarray 50 (PAM50) classifier [171]. This gene expression-based assay leverages the expression levels of 50 key genes to categorize breast cancers into four intrinsic subtypes and calculate a risk of recurrence score, a valuable prognostic indicator estimating the likelihood of distant recurrence. By integrating subtype-specific gene expression and recurrence risk, PAM50 profiling tool has become invaluable in guiding treatment decisions. Notably, it is particularly effective in identifying low-risk luminal A patients with node-negative disease, for whom adjuvant endocrine therapy alone is recommended, sparing these patients the need for chemotherapy [172,173]. This underscores the utility of PAM50 algorithm in guiding personalized treatment strategies tailored to molecular subtype and individual risk profiles.

Further advancing molecular characterization, *Prat et al.* identified the **claudin-low subtype**, characterized by low expression of cell adhesion genes, including claudins, and enrichment in EMT and cancer stem cell features [174]. This subtype is associated with poor clinical outcomes and displays intermediate sensitivity to conventional

INTRODUCTION

chemotherapy, bridging the therapeutic responses of basal-like and luminal tumors [174-176]. Most claudin-low tumors are triple-negative (61 to 71%) and frequently exhibit metaplastic and medullary differentiation, contributing to their aggressive biological behavior. Unique molecular pathways, such as Wnt/ β -catenin, Notch, and TGF- β (transforming growth factor beta), play central roles in driving the invasive and metastatic potential of these tumors [177]. These insights emphasize the importance of understanding subtype-specific features to develop more effective treatment strategies tailored to this challenging subtype. Claudin-low tumors are identified using a specific claudin-low predictor and not explicitly by the PAM50 classifier, which often misclassifies them as basal-like or normal-like, with a smaller proportion falling into luminal A, HER2-enriched, and luminal B categories [176,178].

3.1.4. Molecular heterogeneity in breast cancer

In clinical practice, pathological markers such as ER α , PR, and HER2 fail to comprehensively capture the complexity of breast cancer. To address this limitation, gene expression profiling tests based on microarrays have become valuable tools in breast cancer management, providing critical insights into tumor biology and guiding treatment decisions. Notable examples include *Oncotype DX* [179], which predicts the risk of recurrence in HR-positive breast cancer, and *MammaPrint* [180], which assesses a 70-gene signature to determine the need for adjuvant chemotherapy in early-stage breast cancer. Additionally, *Prosigna* [171], described in **Section 3.1.3**, integrates the PAM50 gene expression signature to classify tumors into intrinsic subtypes while providing a risk or recurrence score.

Despite their utility, the high cost of these tests, along with technical limitations and the complexity of data management, may restrict their widespread use. Moreover, the inherent heterogeneity of each breast cancer subtype poses considerable challenges for diagnostic tools, hindering their ability to fully encapsulate the diverse and multifaceted nature of the disease. The following sections will explore how the heterogeneity across different breast cancer subtypes impacts the effectiveness of these tools in accurately diagnosing and personalizing treatment strategies.

3.1.4.1. Luminal breast cancer heterogeneity

Luminal subtypes of breast cancer exhibit significant heterogeneity, primarily categorized into luminal A and luminal B, both of which are HR-positive [159]. The distinction between these subtypes is crucial for clinical decision-making, as luminal B tumors show elevated expression of proliferation-related genes such as *MKI67* (marker of proliferation Ki-67) and *CCNB1* (cyclin B1), key components in genomic predictors like *Oncotype DX* [179]. The use of Ki-67 as an IHC marker further supports the classification of luminal subtypes, with high Ki-67 expression indicating luminal B and low expression suggesting luminal A [130].

Despite these advancements, challenges remain, as luminal B tumors are not easily categorized based on HR and HER2 statuses, with up to 72% classified as ER^{pos}/HER2^{neg}, 20% as ER^{pos}/HER2^{pos}, 7% as ER^{neg}/HER2^{neg}, and 1% as

INTRODUCTION

ER^{neg}/HER2^{pos} [181]. Additionally, luminal B breast cancers also demonstrate significant heterogeneity in relation to PR expression. Specifically, PR-positive luminal B tumors are associated with a better prognosis and increased tamoxifen sensitivity compared to PR-negative counterparts [182-184]. Moreover, researchers have identified distinct subgroups within PR-negative luminal B tumors based on *PIK3CA* and *TP53* (tumor protein p53) status, each exhibiting differing prognostic outcomes. This highlights the need for refined classification methods to differentiate these subtypes and tailor therapeutic approaches accordingly [185].

3.1.4.2. HER2-positive breast cancer heterogeneity

Importantly, the clinical HER2-positive subtype is not synonym of the molecular HER2-enriched subtype in the PAM50 classification. Approximately 50% of the clinical HER2-positive breast cancers are HER2-enriched, while the remaining 50% predominantly consist of luminal B tumors, followed by basal-like and luminal A subtypes, with the HER2-enriched subtype benefiting the most from HER2-targeted agents [186,187].

Notably, HER2-positive tumors exhibit varying sensitivity to standard treatments based on HR status, with HR^{pos}HER2^{pos} tumors showing lower pathological complete response (pCR) rates compared to HR^{neg}HER2^{pos} tumors after neoadjuvant anti-HER2 therapy combined with chemotherapy [188].

Furthermore, about 30% of HER2-enriched tumors are HER2-negative in the clinical setting [181]. Remarkably, integrating HER2-enriched and *ERBB2* mRNA into a single assay can accurately identify tumors with strong responsiveness to HER2-targeted therapies. This approach could potentially reduce chemotherapy use in around 40% of HER2-positive patients who harbored HER2-enriched/*ERBB2*^{high} tumors, optimizing treatment strategies while minimizing unnecessary toxicity [167].

3.1.4.3. TNBC heterogeneity

TNBCs exhibit significant heterogeneity, encompassing various molecular subtypes at the transcriptomic level [189]. Using the PAM50 molecular assay, combined with a claudin-low classifier, these tumors can be categorized into several subtypes, including basal-like (49%), claudin-low (30%), HER2-enriched (9%), luminal B (6%), luminal A (5%), and normal-breast like (1%) [174,181,190] (**Figure I7A**).

In the clinical setting, a significant portion of basal-like tumors are mistakenly classified as triple-negative, leading to inaccuracies in identifying true basal-like tumors. Specifically, 6-29% and 9-13% of basal-like tumors exhibit ER α or HER2 positivity, respectively [159,191], demonstrating that triple-negative tumors can inaccurately include cases that are not truly basal-like. However, these tumors, while not strictly triple-negative, may respond to targeted therapies traditionally used for HR-positive or HER2-positive tumors, emphasizing the need for refined classification methods that account for ER α ^{pos} and HER2^{pos} statuses.

INTRODUCTION

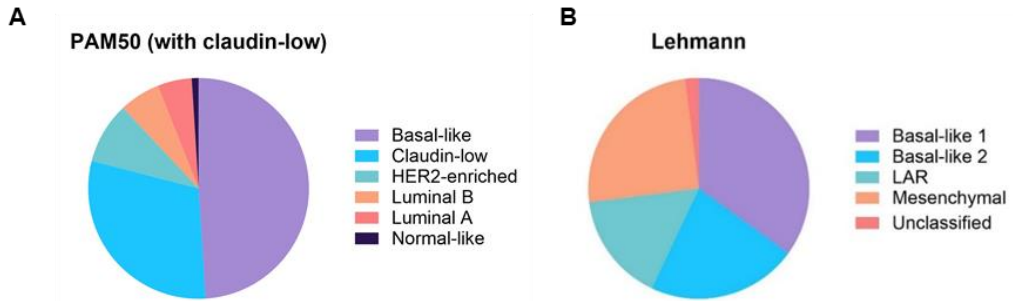


Figure 17. TNBC subtype distribution by PAM50 (combined with claudin-low classifier) and Lehmann classifications. **A.** Classification by PAM50 and the claudin-low centroid-based predictor highlights TNBC heterogeneity, with the majority of tumors classified as basal-like (49%), followed by claudin-low (30%), HER2-enriched (9%), luminal B (6%), luminal A (5%), and normal-like (1%). **B.** Lehmann's classification further subdivides TNBC into basal-like 1 (35%), basal-like 2 (22%), LAR (16%), mesenchymal (25%), and an unclassified cluster (2%). References: *Prat et al., 2011*; *Lehmann et al., 2016*. Graphs were created with GraphPad Prism.

Importantly, the presence of basal markers is associated with significantly worse prognosis [192-195]. Not all TNBC tumors, however, express these markers in clinical practice. To address this issue, *Nielsen et al.* proposed an immunohistochemical surrogate for basal-like cancers, defining these tumors by ER α , PR and HER2 negativity, along with expression of K5/6 basal keratins and EGFR (epidermal growth factor receptor) [191]. This highlights the need for nuanced classification methods that better capture the molecular heterogeneity of TNBC.

Further gene expression cluster analysis by *Lehmann et al.* identified several TNBC subtypes, each differing in their response to standard treatments [196,197] (**Figure 17B**):

- **Basal-like 1 (BL1)**, representing approximately 35% of TNBCs, is characterized by high expression of genes involved in cell cycle regulation and DNA damage response pathways. BL1 tumors are particularly sensitive to DNA-damaging agents, such as cisplatin, highlighting their vulnerability to therapies targeting genomic instability.
- **Basal-like 2 (BL2)**, accounting for 22% of TNBCs, exhibits enriched expression of genes associated with growth factor signaling and metabolic pathways. While BL2 tumors share some similarities with BL1 in terms of aggressiveness, their chemosensitivity is lower. Their unique molecular profile may necessitate tailored therapeutic strategies to target the specific pathways involved in their pathogenesis.
- **Mesenchymal (M)**, comprising 25% of TNBCs, is enriched for genes involved in EMT and growth factor signaling pathways, such as those involving PI3K/mTOR (mammalian target of rapamycin) and Abl (abelson tyrosine-protein kinase 1)/Src (proto-oncogene tyrosine-protein kinase Src). These tumors are characterized by their invasive nature and show sensitivity to inhibitors targeting these signaling pathways, such as the Src inhibitor dasatinib and the dual PI3K/mTOR inhibitor NVP-BE2235.

INTRODUCTION

- **Luminal androgen receptor (LAR)**, representing 16% of TNBCs, is driven by androgen receptor (AR) signaling and highly expresses luminal epithelial markers. Their dependency on AR signaling renders them responsive to AR antagonists, such as bicalutamide, offering a targeted treatment approach for this unique TNBC subset.

Based on PAM50 classification, the distribution of these TNBC subtypes reveals that most BL1, BL2, and M tumors are basal-like, while LAR tumors exhibit higher frequencies of HER2 and luminal subtypes [196].

Comparisons among BL1, BL2, M, and LAR subtypes reveal significant differences across various clinical variables [196]. BL1 tumors demonstrate the strongest response to standard neoadjuvant chemotherapy, with 41% achieving pCR, compared to 18% for BL2 and 29% for LAR. BL1 patients exhibit better relapse-free survival, with nearly 60% surviving beyond 10 years. Despite being of histologically higher grade, BL1 tumors present at a lower clinical stage (6% stage 3) compared to BL2 (30% stage 3) and LAR (22% stage 3). LAR shows the highest rate of bone metastasis (46%), while BL1 and BL2 show lower incidences (16% and 30%, respectively), consistent with the tropism of hormone-sensitive cancers for bone metastasis. The LAR subtype also exhibits a substantial enrichment of regional lymph node metastases (47%), compared to BL2 (30%). In contrast, the M subtype demonstrates the lowest lymph node involvement (21%) and the highest lung metastasis frequency (46%).

These findings collectively highlight the substantial molecular heterogeneity of TNBC, with distinct subtypes such as BL1, BL2, M, and LAR demonstrating varying responses to treatment [196,197]. Each subtype requires a personalized therapeutic approach, with BL1 showing strong sensitivity to DNA-damaging agents, while BL2, M, and LAR necessitate targeted strategies to address their specific biological characteristics.

In summary, the molecular classification of breast cancer continues to evolve, offering critical insights into tumor heterogeneity and guiding the development of more targeted and effective therapies.

3.2. *In vivo* breast cancer models

There are several *in vivo* models available for studying human breast tumors, including cell line-derived xenografts (CDX), patient-derived xenografts (PDX), and genetically engineered mouse models (GEMMs) [198]. These models are essential tools in cancer research, facilitating the study of tumor biology, progression, metastasis, and therapeutic responses within a living organism.

3.2.1. Cell line-derived xenografts

CDX models involve culturing tumor cell lines *in vitro*, which can introduce significant biases, influencing which cell types ultimately thrive. There are well-characterized cell lines representing the major clinical subtypes, such as MCF7 and T47D, resembling the luminal A subtype, BT474 and MDA-MB-361, sharing the gene expression profile of the luminal B subtype, SKB43 and HCC202, recapitulating the features of HER2-positive

INTRODUCTION

tumors, and BT20, MDA-MB-231, MDA-MB-468, resembling the triple-negative human tumors [198,199].

However, not all these cell lines can be successfully established *in vivo* and can result in aberrant clonal selection, where only a subset of cells from the original tumor dominate, leading to the loss of important molecular features that define the heterogeneity and complexity of the original tumor. The clonal nature of CDX models often fails to capture the full spectrum of genetic and phenotypic diversity seen in human tumors, reducing their ability to accurately mimic clinical responses to therapy. Furthermore, CDX models are typically established by subcutaneous implantation of these cell lines into the flank of mice, which can hinder natural tumor-stromal interactions critical for tumor progression, metastasis, and treatment responses [199]. Additionally, the use of immunodeficient mice eliminates the host immune system, posing challenges for immunotherapy research.

3.2.2. Patient-derived xenografts

PDX models provide a more physiologically relevant approach compared to CDX models, as they involve direct transplantation of patient-derived tumor cells or tumor pieces into host mice, bypassing *in vitro* culture [200]. This preserves the original tumor's diversity, including its genetic and phenotypic heterogeneity, which is often lost in cell line-based models. Consequently, PDX models can more accurately reflect the complexity of human cancers, including the diversity in treatment responses [201-204]. They offer a valuable platform for studying drug efficacy, resistance mechanisms, and disease progression.

However, their development and maintenance are time-consuming and require significant infrastructure and financial resources, limiting their scalability in routine research settings. Additionally, modelling luminal ER α -positive subtypes, which represent over 70% of diagnosed breast cancers, has proven especially difficult in part because xenotransplantation tends to favor the most aggressive cancer forms, leading to a bias in both CDX and PDX models toward the triple-negative subtype [203,205]. Moreover, PDX models still rely on immunodeficient mice, limiting their applicability for immunotherapy studies.

3.2.3. Genetically engineered mouse models

GEMMs provide a versatile platform to investigate breast cancer biology, progression, and therapeutic responses [206]. These models closely replicate the spontaneous development of breast cancer, capturing the stepwise progression observed in human disease, enabling the study of tumor initiation, progression, and metastasis in a controlled environment. GEMMs are specifically engineered with conditional and/or constitutively active mutant alleles to mimic the genetic alterations found in human breast cancers. Unlike cell culture or xenograft models, GEMMs develop tumors within an intact immune system and a native tumor microenvironment.

Importantly, many GEEMs are engineered to model specific molecular subtypes of breast cancer, allowing the investigation of subtype-specific mechanisms of

INTRODUCTION

tumorigenesis and treatment response. Researchers typically use different promoters to drive oncogene expression in a broad range of cells or in a subset of cells, depending on the purpose of the study:

- **Tissue-specific promoters:** Examples include the MMTV (mouse mammary tumor virus) and C3(1) (rat prostatic steroid binding protein gene) promoters, which are specifically activated in the mammary epithelium. These promoters induce widespread expression across all MECs, providing robust models for studying breast tumorigenesis [207-215].
- **Lineage-specific promoters:** Oncogene expression is selectively driven in specific cell subsets. In the adult mammary gland, promoters such as *Lgr5*, *Krt5*, and *Krt14* are activated in BaCs, while *Krt8* and *Blg* (beta-lactoglobulin) drive expression in LCs [91-93]. The WAP (whey acidic protein) promoter specifically drives expression in alveolar ERα^{neg} luminal progenitors during pregnancy and lactation [99,216-220], and is also transiently activated in a subset of alveolar luminal cells during estrus in virgin female mice [221]. This approach provides a more nuanced understanding of breast cancer initiation and progression in defined cellular contexts.

These promoters often use the Cre/loxP recombination system to enable precise spatial control over oncogene expression, targeting specific cell lineages while minimizing off-target expression in unintended cell types [91-93,99,210,221].

Comparative genomic analyses between various mouse models and human cancer datasets have enabled researchers to identify the molecular subtypes in humans that align closely with different mouse models [91-93,99,222,223] (**Table 11**):

- **Basal-like tumors:** Captured by models such as MMTV-Wnt1 [207], WAP-TAg [217], *Blg*-Cre/*Brca1*^{ff}/p53^{+/-} [91], *Krt14*-Cre/*Brca1*^{ff}/p53^{+/-} [91], MMTV-Myc [208], WAP-Myc [218], C3(1)-TAg [215], MMTV-Cre/*Brca1*^{Co/Co}/p53^{+/-} [210], *Krt8*-Cre/*Pik3ca*^{H1047R} [92,93], and MMTV-Aib1 [209].
- **Luminal tumors:** Exemplified by MMTV-Myc, WAP-Myc, *Krt8*-Cre/*Pik3ca*^{H1047R}, WAP-Int3 [219], MMTV-Hras [208], *Lgr5*-Cre/*Pik3ca*^{H1047R} [92], MMTV-Neu [212], *Krt5*-Cre/*Pik3ca*^{H1047R} [93], MMTV-PyMT [211], *Rb*^{-/-} [224], and MMTV-Aib1.
- **HER2-enriched tumors:** Modeled by *Krt8*-Cre/*Pik3ca*^{H1047R}, WAP-Cre/*Etv6* [221], *Rb*^{-/-}, MMTV-Fgf3 [213], WAP-T₁₂₁ [220], MMTV-Aib1, p18^{-/-} [225], and *Brg1*^{+/-} [226].
- **Normal-like tumors:** Reflected in models such as MMTV-Neu, MMTV-Fgf3, p18^{-/-}, and MMTV-Lpa [214].
- **Claudin-low tumors:** Defined by mesenchymal traits, these tumors are represented by a subset of tumors within the C3(1)-TAg model, along with those from MMTV-Cre/*Brca1*^{Co/Co}/p53^{+/-}, WAP-T₁₂₁, p18^{-/-}, MMTV-Lpa, WAP-Cre/*Kras*^{G12D} [99], and *Brg1*^{+/-}.

INTRODUCTION

Mouse model	Oncogene status and targeted cells	Human breast cancer subtype
MMTV-Wnt1	<i>Wnt1</i> overexpression in MECs	Basal-like
WAP-TAg	<i>TAg</i> overexpression in LCs	Basal-like
<i>Blg-Cre/Brca1^{flf}/p53^{+/-}</i>	<i>Trp53</i> heterozygous and <i>Brca1</i> loss in BaCs	Basal-like
<i>Krt14-Cre/Brca1^{flf}/p53^{+/-}</i>	<i>Trp53</i> heterozygous and <i>Brca1</i> loss in BaCs	Basal-like
MMTV-Myc	<i>Myc</i> overexpression in MECs	Basal-like, Luminal B
WAP-Myc	<i>Myc</i> overexpression in LCs	Basal-like, Luminal B
C3(1)-TAg	<i>TAg</i> overexpression in MECs	Basal-like, Claudin-low
MMTV-Cre/ <i>Brca1^{ColCo}/p53^{+/-}</i>	<i>Trp53</i> heterozygous and <i>Brca1</i> loss in MECs	Basal-like, Claudin-low
<i>p53^{+/-}</i> irradiated	<i>Trp53</i> heterozygous and irradiated	Basal-like, Claudin-low
<i>Krt8-Cre/Pik3ca^{H1047R}</i>	<i>Pik3ca^{H1047R}</i> overexpression in LCs	Basal-like, HER2-enriched, Luminal B
WAP-Int3	<i>Notch4</i> overexpression in LCs	Luminal-like
DMBA-induced	Random mutations in MECs	Luminal-like, Basal-like
MMTV-Hras	<i>Hras</i> overexpression in MECs	Luminal A
<i>Lgr5-Cre/Pik3ca^{H1047R}</i>	<i>Pik3ca^{H1047R}</i> overexpression in BaCs	Luminal A, Luminal B
MMTV-Neu	<i>Neu/ErbB2</i> overexpression in MECs	Luminal A, Normal-like
<i>Krt5-Cre/Pik3ca^{H1047R}</i>	<i>Pik3ca^{H1047R}</i> overexpression in BaCs	Luminal B
MMTV-PyMT	<i>PyMT</i> overexpression in MECs	Luminal B
WAP-Cre/ <i>Etv6</i>	<i>Etv6-Ntrk3</i> overexpression in LCs	HER2-enriched
<i>Rb^{-/-}</i>	<i>Rb</i> homozygous null	HER2-enriched, Luminal A
MMTV-Fgf3	<i>Fgf3</i> overexpression in MECs	HER2-enriched, Normal-like
WAP-T ₁₂₁	<i>T₁₂₁</i> overexpression in LCs	HER2-enriched, Claudin-low
MMTV-Aib1	<i>Aib</i> overexpression in MECs	HER2-enriched, Basal-like, Luminal B
<i>p18^{-/-}</i>	<i>Cdkn2c</i> homozygous null	HER2-enriched, Normal-like, Claudin-low
MMTV-Lpa	<i>Lpa</i> overexpression in MECs	Normal-like, Claudin-low
WAP-Cre/ <i>Kras^{G12D}</i>	<i>Kras^{G12D}</i> overexpression in LCs	Claudin-low
<i>Brg1^{+/-}</i>	<i>Brg1</i> heterozygous deletion	Claudin-low, HER2-enriched
MPA+DMBA-induced	Random mutations in MECs	Claudin-low, Normal-like, Basal-like, HER2-enriched

Table I1. Overview of genetically engineered and chemically induced mouse models of mammary tumors, specifying oncogene or tumor suppressor gene alterations, targeted MECs, and corresponding human breast cancer subtypes. Subtypes are based on PAM50 gene profiling by *Parker et al., 2009* [171] and the intrinsic gene list by *Hu et al., 2006* [227], including basal-like, luminal A, luminal B, HER2-enriched, normal-like, and claudin-low subtypes. References: *Herschkowitz et al., 2007*; *Molyneux et al., 2010*; *Pfefferle et al., 2013*; *Koren et al., 2015*; *Van Keymeulen et al. 2015*; *Abba et al., 2016*; *Radler et al., 2021*. Table adapted from *Vinuesa-Pitarch et al., 2021* [45].

INTRODUCTION

3.2.4. Inducible breast cancer models

Inducible breast cancer models utilize chemical, physical, and biological approaches to induce carcinogenesis, each offering unique advantages for studying specific aspects of tumor biology.

- **Chemically induced breast cancer models:** Relying on chemical treatments to induce DNA damage and mutations, these models are valuable for studying the mutagenic processes involved in breast cancer initiation and progression.
 - DMBA (7,12-dimethylbenz[a]anthracene) model: Exposure to DMBA, either alone or in combination with medroxyprogesterone acetate (MPA), induces a diverse array of tumor subtypes [222,223,228,229] (**Table I1**). This model, used in this thesis, will be further described in the Methods section.
 - NMU (N-nitroso-N-methylurea) model: Exposure to NMU predominantly leads to mammary tumors with a gene expression profile similar to HR-positive human breast tumors [230,231].
- **Radiation-induced breast cancer models:** Physical agents, such as ionizing radiation, are widely used in breast cancer models to induce DNA damage and genomic instability, leading to breast cancer development [232]. Bagg Albino (BALB/c) mice are commonly used in radiation carcinogenesis studies due to a unique polymorphism in the *Prkdc* gene. This gene encodes the DNA-dependent protein kinase catalytic subunit, which plays a key role in DNA repair processes. The polymorphism in *Prkdc* increases susceptibility to cancer, particularly following radiation exposure, making BALB/c mice an ideal model for exploring how deficiencies in DNA repair mechanisms contribute to breast cancer development [233]. Mice harboring mutations in *Trp53* show a significantly higher incidence of mammary tumors after irradiation. These tumors molecularly resemble human basal-like and claudin-low tumors (**Table I1**), making these models particularly useful for studying aggressive breast cancer subtypes [234]. Combined mutations further enhance radiation-induced carcinogenesis in murine models. *Trp53*^{+/-} irradiated mice exhibit heightened breast cancer susceptibility, especially when combined with mutations in *Atm* (ataxia telangiectasia mutated) or *Brca1*, both of which are critical for maintaining genomic integrity [235,236]. Similarly, mutations in *Apc* (adenomatous polyposis coli) significantly increase breast cancer risk following radiation exposure. These combined genetic and environmental perturbations closely mimic the complex interactions that contribute to human breast cancer [237]. Radiation-induced breast cancer models are particularly valuable for understanding how environmental factors, such as radiation from medical treatments or environmental sources, contribute to the initiation and progression of breast cancer. By replicating these processes *in vivo*, researchers can investigate the mechanisms of radiation-induced malignancy and identify potential strategies for prevention and intervention.
- **Biologically-induced breast cancer models:** Biological approaches, such as lentiviral infections, provide precise methods for investigating gene-specific contributions to breast cancer [238]. Lentiviral vectors are engineered to introduce

INTRODUCTION

oncogenes or silence tumor suppressor genes, enabling targeted modifications in specific mammary epithelial cells. This approach offers researchers the ability to activate or suppress gene expression after birth, providing greater flexibility compared to the permanent genetic alterations used in GEMMs. The RCAS-TVA (replication-competent avian sarcoma-leukosis virus long-terminal repeat with splice acceptor - tumor virus A) system is a notable example of a biologically induced breast cancer model. In this system, the TVA receptor is expressed on target cells, enabling them to be selectively infected by RCAS viral vectors carrying genes of interest [239]. For instance, *Ding et al.* used this system to infect MMTV-TVA mice with RCAS-*ErbB2*, inducing *ErbB2* expression in both ER α^{pos} and ER α^{neg} LCs, leading to the transformation of these cells and the formation of HER2-positive tumors [9].

3.3. Lineage tracing as the gold-standard approach for exploring cellular hierarchies and heterogeneity in breast cancer

3.3.1. Unraveling the cell of origin in different breast cancer subtypes

The cell of origin for breast cancer remains one of the most debated topics in cancer biology, with ongoing discussions regarding the factors that influence tumor initiation. Over the years, comparative expression profiling has been widely used to identify shared expression patterns among normal human MECs and various breast cancer subtypes, providing clues about a plausible cell of origin. In this direction, several studies have pointed to luminal progenitors (ER α^{neg} LCs) as a potential target population for the initiation of basal-like breast cancer, given their similar gene expression profiles [240-242]. Conversely, the BaC-associated signature showed the greatest similarity to the claudin-low subtype. Meanwhile, the ER α^{pos} LC-associated (mature luminal) signature was predominantly upregulated in luminal A and B breast cancer subtypes. These findings highlight the utility of comparative expression profiling in pinpointing the cell of origin across diverse breast cancer subtypes and suggest that both progenitor and differentiated MECs may serve as potential targets for oncogenic events.

Compelling *in vivo* evidence has shown that distinct breast cancer subtypes can emerge from oncogenic expression or tumor suppressor depletion in any mammary cell lineage. Both luminal and basal progenitors are susceptible to transformation, with basal-like breast cancers more likely to originate from LCs rather than BaCs [91-93] (**Figure 18**). For instance, *Pik3ca*^{H1047R} expression in *Lgr5*-expressing BaCs generated luminal A and luminal B tumors, while its expression in *Krt8*-expressing LCs resulted in luminal B, HER2-enriched, and basal-like tumors [92]. Similarly, *Pik3ca*^{H1047R} expression in *Krt5*-expressing BaCs induced luminal B tumors, whereas expression in *Krt8*-positive LCs generated luminal B and basal-like tumors [93]. In a separate study using *Trp53* heterozygous mice, *Brca1* deficiency was directed to either *Krt14*-expressing BaCs or *Blg*-expressing LCs, primarily targeting ER α^{neg} progenitors. Both models produced mammary tumors resembling the genetic profile of human basal-like tumors [91], although only luminal-derived tumors shared histologic similarities with human *BRCA1* loss-of-function breast cancers. Furthermore, oncogenic activation in ER α^{neg} LCs

INTRODUCTION

through the WAP promoter yielded a broad spectrum of tumor subtypes [99,222,223] (**Table 11**). Specifically, *Myc* (myelocytomatosis oncogene) overexpression led to both luminal B and basal-like tumors, while large T antigen (*TAg*) overexpression predominantly produced basal-like tumors. Introduction of the *Etv6-Ntrk3* fusion gene or the truncated form of *TAg* (*T₁₂₁*) that inactivates Rb (retinoblastoma-associated protein), resulted in HER2-enriched tumors. Lastly, expression of *Kras^{G12D}* or *T₁₂₁* overexpression induced claudin-low tumors (**Figure 18**).

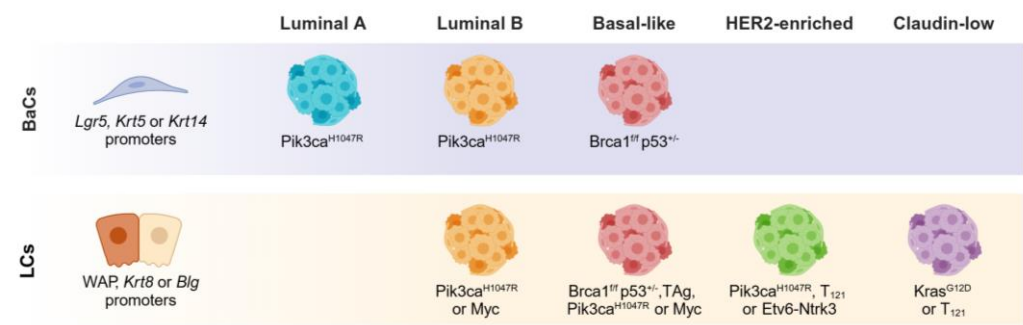


Figure 18. Tumor subtype specification in BaCs and LCs using distinct oncogenic drivers. This figure illustrates different breast tumors resulting from oncogenic activation or tumor suppressor deletion in BaCs using *Lgr5*, *Krt5*, and *Krt14* promoters; in ERα^{neg} LCs using WAP promoter; and in both ERα^{pos} and ERα^{neg} LCs using *Krt8* and *Blg* promoters. Each oncogenic murine model aligns with specific human breast cancer subtypes: luminal A, luminal B, basal-like, HER2-enriched, and claudin-low. References: Herschkowitz et al., 2007; Molyneux et al., 2010; Pfefferle et al., 2013; Koren et al., 2015; Van Keymeulen et al. 2015; Radler et al., 2021. Figure created with BioRender.

Collectively, these findings suggest that the oncogene itself can initiate the tumorigenic process in any cell type, regardless of its lineage. However, the cell of origin appears to play a critical role in determining the aggressiveness of the resulting tumor. The ongoing debate centers on whether tumor development is predominantly driven by the cell of origin, the specific oncogenic mutations, or an interplay between these two factors. It is important to note that traditional lineage-specific oncogenic models may hinder the accurate identification of the true tumor cell of origin. These models often artificially induce tumorigenesis through the forced expression of oncogenes, which can misrepresent the natural course of disease development. Moreover, breast cancer cells exhibit substantial plasticity upon oncogenic activation [91-94], adding another layer of complexity to identifying a specific cell of origin. To address these challenges, combining lineage tracing tools with genetic and inducible breast cancer models provides a more authentic representation of cancer development. This approach enables researchers to trace the natural origins and progression of the disease, offering valuable insights into tumor biology. **This methodology constitutes the primary focus of this thesis.**

Lineage tracing has emerged as a powerful tool for investigating the cell of origin in breast cancer, although relatively few studies have utilized this approach. One notable study used lineage tracing to track a subset of BaCs expressing *Lgr5* in C3(1)-TAg mice, a murine model of TNBC that shares a basal-like genetic profile with its human

INTRODUCTION

counterpart [215,222] (**Table I1**). The study demonstrated that hyperplastic lesions predominantly arose from *Lgr5*-derived progeny, identifying *Lgr5*-positive BaCs as the cell of origin of basal-like breast tumors [8]. In contrast, lineage tracing studies in the MMTV-PyMT breast cancer model, which transcriptomically resembles human luminal B tumors [211,223] (**Table I1**), revealed that these luminal tumors originated from *Krt8*-positive LCs rather than *Krt5*-positive BaCs [243]. Conversely, *Rios et al.*, using the MPA+DMBA-induced breast cancer model, showed that *Krt5*-positive BaCs give rise to luminal-enriched clones that expand during tumor progression [244], supporting a basal origin for luminal tumors.

An innovative study combined lineage tracing with the RCAS-TVA biological tool (see **Section 3.2.4**) to investigate the role of ER α in the development and progression of HER2-positive tumors [9]. Researchers utilized *Esr1*-Cre/MMTV-TVA/*Rosa26*-tdRFP mice infected with an RCAS-*ErbB2* viral vector, enabling the transformation of both ER α^{pos} and ER α^{neg} LCs. This approach allowed effective lineage tracing of both cell populations, with ER α^{pos} -derived cells labeled in red, while ER α^{neg} -derived progeny remained unlabeled. The study revealed that HER2-positive tumors originating from ER α^{pos} LCs are more aggressive compared to those arising from ER α^{neg} LCs. Remarkably, ER α -positive cells were required to lose *Esr1* expression in order to expand and metastasize, underscoring the critical influence of the cell of origin in shaping tumor behavior and determining breast cancer outcomes.

3.3.2. Exploring clonal dynamics during tumor progression and dissemination using combinatorial approaches

Recent advancements in microscopy techniques, such as intravital imaging and 3D whole-mount imaging, combined with multicolor lineage tracing and single-cell lineage tracing via cellular barcoding, have transformed our understanding of mammary tumor progression and tumor cell population dynamics [10,244]. These cutting-edge approaches provide unprecedented resolution in tracking the behavior and interactions of individual tumor clones throughout tumorigenesis.

3.3.2.1. 3D-whole mount imaging to explore tumor dynamics at high resolution

Rios et al. advanced multicolor lineage tracing by the developing a novel tissue clearing technique called FUnGI (Fructose, Urea, and Glycerol for Imaging). This method enabled the efficient clearing of mammary tissue within two hours at room temperature, while preserving structural integrity and achieving high transparency for high-resolution imaging. FUnGI was integrated with a large-scale, single-cell resolution 3D imaging protocol [20], creating a powerful platform for analyzing clonal expansion dynamics in mammary tumors. This innovative approach allowed whole-mammary gland visualization with minimal light scattering and no fluorescence loss, facilitating the effective visualization of fluorescent reporter genes across various tissues. As a result, it permitted precise tracking of clonal populations, yielding valuable insights into tumor progression and cellular plasticity within the mammary gland. By allowing high-resolution monitoring of multiple clonal populations within the same tissue, this imaging

INTRODUCTION

technology offered an unprecedented view of clonal interactions, evolution, and their contribution to breast cancer development.

3.3.2.2. Intravital imaging for real-time monitoring of tumor cell dynamics

Intravital imaging, which involves the real-time visualization of living tissues within an organism, is a powerful tool for studying cellular processes in their native microenvironment [245]. This technique enables dynamic monitoring of tumor cell behaviors, including proliferation, migration, and interaction with surrounding cells, in live animals. A key feature of intravital imaging is the use of surgically implanted imaging windows, which provide repeated, non-invasive access to tissues over extended periods. These windows facilitate longitudinal studies, offering valuable insights into the temporal progression of tumor development and metastatic spread. By capturing these processes *in vivo*, intravital imaging reveals a more nuanced and complex understanding of tumor cell behavior than was previously possible.

The *Rosa26*-Confetti reporter system, which randomly activates one of four distinct fluorescent proteins (GFP, YFP, RFP, and CFP) in the targeted compartment, enables precise tracking of individual clones within a tumor using multiple color markers [246] (**Figure I9**). Unlike traditional single-color lineage tracing, this multicolor system allows researchers to differentiate and monitor multiple clonal populations within the same tissue.

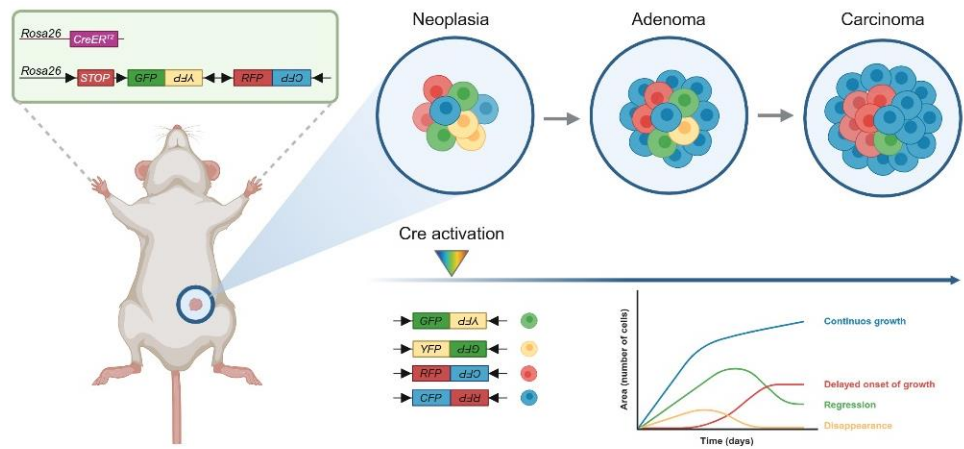


Figure I9. Schematic model of clonal dynamics in MMTV-PyMT tumors, derived from multi-color lineage tracing studies using real-time intravital imaging. The CreERT² system allows temporal control of recombination in neoplastic cells, leading to the random expression of different fluorescent proteins (GFP, YFP, RFP, and CFP). The model depicts the progression of tumor growth from neoplasia to adenoma and carcinoma, with colors indicating different clonal populations. Growth patterns of these clones are intravitaly imaged over time through an imaging window, showing continuous growth, delayed onset, regression, and disappearance, each represented by a specific color and growth curve in the graph on the right. Reference: Zomer *et al.* 2013. Figure created with BioRender.

In a pioneering study by Zomer *et al.*, intravital imaging was combined with random multi-labeling using *Rosa26*-Cre^{ERT2}/*Rosa26*-Confetti mice, allowing real-time tracing of individual tumor cells within the mammary gland [10] (**Figure I9**). The study uncovered

INTRODUCTION

surprising findings that challenged conventional assumptions about tumor cell proliferation. Contrary to the prevailing notion that most tumor cells proliferate during cancer progression, only a small subset of cells underwent clonal expansion. This highly proliferative subset plays a critical role in driving tumor growth and progression, while most cells within the primary tumor exhibit limited proliferative capacity. Many tumor cells either disappear, grow at a slower rate, or partially expand before eventually regressing. These observations suggest that the aggressive behavior of a tumor may be driven by a minority of highly proliferative cells, whereas most tumor cells remain relatively dormant or non-contributive.

In a related study, *Scheele et al.* also demonstrated that protective mechanisms within mammary tissues, such as tissue remodeling and the stochastic loss of self-renewing cells, limit the spread of mutations and ensure that only a subset of mutant clones expand, providing further insights into the selective clonal dynamics underlying tumorigenesis [247]. Together, these studies underscore the importance of focusing on specific clonal populations to understand cancer progression and metastasis. Such insights underscore the importance of focusing on specific clonal populations to better understand the mechanisms underlying cancer progression and metastasis.

3.3.2.1. Single-cell lineage tracing through cellular barcoding for mapping tumor subclonal dynamics

Single-cell lineage tracing, which combines genetic barcoding with single-cell RNA sequencing (scRNA-seq), has emerged as a transformative technique for studying cellular hierarchies and tumor heterogeneity [248]. Genetic barcoding involves labeling individual cells or cell populations with unique DNA sequences or fluorescent tags, enabling researchers to track these cells over time. By linking these barcodes to gene expression profiles at single-cell resolution, this approach provides a powerful framework for mapping cellular trajectories and identifying the functional roles of distinct subpopulations in various biological processes, including tumor progression and therapy resistance.

In breast cancer research, single-cell lineage tracing has proven invaluable for unraveling the clonal dynamics of tumor cells. For example, one study utilized cellular barcoding combined with single-cell transcriptomics to reconstruct the subclonal dynamics of EGFR-amplified TNBC cells during treatment with afatinib, a tyrosine kinase inhibitor that irreversibly blocks EGFR [249]. This method revealed how distinct subclones within a tumor respond differently to targeted therapy, highlighting the complexity of tumor evolution and resistance mechanisms. In another example, *Ginzel et al.* applied this approach to investigate the tumorigenic potential of distinct HER2 isoforms, including wild-type HER2 (^{WT}HER2), ^{d16}HER2, and ^{p95}HER2, within the same mammary gland environment [250]. These researchers used MMTV-Cre mice crossed with HER2-Crainbow mice, a model engineered to express three HER2 variants, each uniquely fluorescently barcoded and flanked by *LoxP* sites. By combining fluorescence imaging with scRNA-seq, they were able to delineate tumor phenotypes associated with each isoform. Their findings demonstrated that ^{WT}HER2 infrequently induced indolent

INTRODUCTION

tumors, while ^{d16}HER2 led to the formation of luminal-like proliferative in situ lesions that eventually progressed. In contrast, ^{p95}HER2 was linked to the early onset of invasive cancers characterized by the co-expression of luminal and basal epithelial markers. These results underscore the complexity of HER2-driven tumorigenesis and emphasize the importance of subclassifying HER2-positive breast cancer patients based on isoform expression; such subclassification could guide personalized treatment strategies, optimize clinical outcomes, and improve therapeutic precision.

By enabling the reconstruction of subclonal dynamics and uncovering mechanisms of tumor progression, single-cell lineage tracing through cellular barcoding has significantly advanced our understanding of breast cancer biology. This methodology continues to offer unparalleled insights into the molecular and cellular processes driving tumor heterogeneity and therapeutic resistance.

3.3.3. EMT heterogeneity and its role in tumor evolution

Lineage tracing has become an invaluable tool for studying the role of EMT in cancer progression and metastasis, shedding light on the intricate dynamics between epithelial and mesenchymal states. EMT is a biological process in which epithelial cells lose their polarity and adhesion properties, adopting mesenchymal traits such as enhanced motility, invasiveness, and resistance to apoptosis. These changes enable tumor cells to detach from the primary tumor and migrate to distant organs [251,252].

Studies using breast cancer models, such as MMTV-PyMT and MMTV-Neu, have monitored mesenchymal markers like FSP1 (fibroblast-specific protein 1) and vimentin to explore the role of EMT in metastasis. Surprisingly, these studies suggested that EMT does not directly contribute to lung metastasis, as the metastases formed were not derived from *Fsp1*- or *Vim*-positive cells [253,254]. However, these findings were limited by the small fraction of cells undergoing EMT that were studied. For example, mesenchymal cells marked by low E-Cadherin levels constituted only 5% of the population, whereas *Fsp1*-positive cells accounted for just 0.3% [254]. Notably, E-Cadherin-low cells demonstrate metastatic potential when injected into circulation [254], yet whether this transition naturally occurs *in vivo* during disease progression remains unclear. Despite the unclear link between EMT and metastasis, EMT significantly enhances cancer cell survival under chemotherapy, rendering these cells more resilient and difficult to eliminate. This resilience contributes to the development of recurrent lung metastases, underscoring EMT's role in therapy resistance [253].

Interestingly, *Li et al.* developed a sophisticated lineage tracing tool called *EMTracer*, which utilizes the Cre and Dre combinatorial systems to precisely trace *Kit*-positive LCs transitioning through EMT states during tumor progression [31]. *EMTracer* is a triple transgenic mouse model comprising *Kit*-CreER, EMT gene-LSL-Dre, and an NR1-reporter (**see Section 1.2**). By integrating *EMTracer* with the MMTV-PyMT breast cancer model, the authors investigated EMT dynamics throughout tumor development. In this system, tamoxifen administration induced Cre/*LoxP* recombination, labeling luminal *Kit*-positive cells with zsGreen. Simultaneously, cells expressing EMT markers like vimentin or N-cadherin activated Dre expression, triggering Dre/*rox* recombination,

INTRODUCTION

which switched the fluorescent label from zsGreen to tdTomato. This dual-reporter approach enabled precise tracking of cells transitioning through EMT states during tumor progression. Functional studies revealed that vimentin was dispensable for metastasis, whereas N-cadherin activation was critical for successful lung colonization. The findings also demonstrated that breast cancer cells initiate the EMT program during the early stages of primary tumor growth, rather than later during dissemination or metastatic colonization. This discovery challenges the widely held notion that EMT predominantly occurs in advanced tumor stages and underscores the importance of investigating EMT across all stages of tumor progression.

Overall, these findings highlight EMT as a heterogeneous and dynamic process, where tumor cells often undergo partial transitions rather than a complete EMT, resulting in hybrid states that exhibit both epithelial and mesenchymal traits. Such hybrid states are thought to contribute to tumor plasticity, invasiveness, and therapy resistance. Future research could benefit from employing combinatorial lineage tracing systems to label and track cells expressing various mesenchymal markers. Additionally, integrating lineage tracing with scRNA-seq at different tumor stages could provide deeper insights into intermediate EMT states, helping to unravel their functional roles in tumor progression and metastasis. These advancements emphasize the need to investigate EMT as a continuum of states, rather than a binary process, to fully understand its implications for cancer biology and therapeutic intervention.

3.3.4. Metastatic dynamics revealed by multicolor lineage tracing

Multicolor fluorescent lineage tracing in the MMTV-PyMT model provided critical insights into the mechanisms of metastasis, revealing that cancer cells spread as cohesive clusters rather than through the sequential seeding of individual cells [11]. Following the orthotopic implantation of mammary tumors composed of cells labeled with distinct fluorescent proteins, lung metastases were found to contain cells of multiple colors, indicating that metastases originated from multicellular clusters rather than single cells. This study outlined metastatic progression through five distinct stages, each characterized by the presence of multicolored tumor cell clusters. These stages include collective invasion, local dissemination into the adjacent stroma, extravasation into tumor emboli, circulation of tumor cell clusters, and the formation of distant micro- and macro-metastases. By identifying these stages, the research underscores the critical role of collective cell behavior in metastatic progression, offering valuable insights into potential therapeutic targets aimed at disrupting cluster-based dissemination.

HYPOTHESIS AND OBJECTIVES

HYPOTHESIS AND OBJECTIVES

Lineage tracing has emerged as a powerful tool for unraveling the intricate cellular hierarchies in both healthy and malignant tissues. Despite significant advancements in breast cancer research, critical gaps persist in literature. Most studies have narrowly focused on specific mammary epithelial cell subpopulations, failing to provide a comprehensive overview of the diverse cellular hierarchies within the mammary gland. Moreover, there is limited understanding of how clonal dynamics evolve across different tumor stages and breast cancer subtypes, particularly in aggressive forms such as triple-negative breast cancer. These gaps underscore the need for an integrated and thorough exploration of cellular behaviors across diverse contexts in breast cancer biology.

This study hypothesizes that the application of lineage tracing technology will provide critical insights into the cellular origins, clonal dynamics, and metastatic behavior of distinct mammary epithelial cell types in preclinical models that closely mimic clinical TNBC. By deciphering cellular hierarchies across different tumor stages and subtypes, this research aims to deepen our understanding of tumor heterogeneity and identify potential biomarkers for diagnosis and prognosis. These insights may ultimately guide the development of personalized therapeutic strategies, improving treatment efficacy and outcomes for women with TNBC.

This thesis is structured around four primary objectives:

- **Objective 1:** To histologically and transcriptomically characterize different preclinical models of triple-negative breast cancer.
- **Objective 2:** To elucidate the cellular origins in multiple subtypes of triple-negative breast cancer.
- **Objective 3:** To investigate the clonal dynamics of distinct mammary epithelial cell types across transcriptomically diverse triple-negative breast cancer subtypes.
- **Objective 4:** To determine the metastatic potential of different mammary epithelial cell populations in triple-negative breast tumors.

This thesis aims to address fundamental questions in breast cancer biology, with a focus on advancing our understanding of tumor heterogeneity.

HYPOTHESIS AND OBJECTIVES

MATERIALS AND METHODS

1. Mice

1.1. Maintenance and welfare of the animal colony

The animals used in this project were housed in individual, ventilated, and autoclavable cages within specific pathogen-free facilities, ensuring a controlled environment free from known infectious agents. To support animal welfare and maintain the reproducibility of experimental results, the colony had ad libitum access to water and was kept at a stable temperature of 20 to 22 °C under a 12-hour light/dark cycle.

The experimental protocols used in this study were carefully designed and approved in accordance with the regulations and legal requirements established by the Department of Agriculture, Livestock, and Fisheries of the Catalonia Government (Approval num. 11195). Additionally, all protocols adhered to the guidelines of the Animal Experimentation Ethics Committee of the Germans Trias i Pujol Research Institute and the Comparative Medicine and Bioimage Centre of Catalonia (FUE-2020-01747689; ID-9HSW349C9). These procedures complied with the international standards set by the Association for Assessment and Accreditation of Laboratory Animal Care (AAALAC), ensuring the highest levels of ethical and scientific rigor in animal research.

1.2. Murine models

1.2.1. Breast cancer models

This study utilized three distinct commercially available murine models of breast cancer: two genetically engineered models (MMTV-PyMT and C3(1)-TAg) and one chemically induced model (DMBA). These mouse models induce tumorigenesis through different mechanisms, effectively representing the diverse range of molecular subtypes observed in clinical settings (**Table M1**).

Model	Induction	Oncogenic events	Subtype
MMTV-PyMT	Genetically Engineered	PyMT oncogene drives ShcA, PI3K, PLCγ, and STAT3 activation	Luminal B [222,223]
C3(1)-TAg	Genetically Engineered	TAg oncogene induces p53 and Rb inactivation, and Kras amplification	Basal-like, Claudin-low [222,223]
DMBA-induced	Chemical treatment	DMBA-induced random mutagenesis promotes PI3K, COX-2, and PD-L1 activation, Pten and Rb inactivation, AhR, Ccnd1 and Myc overexpression	Claudin-low, Basal-like, Normal-like, HER2-enriched, Luminal-like [222,223,228,229]

Table M1. Summary of the breast cancer models used in this study. The table outlines the induction methods, key oncogenic events, and corresponding human breast cancer subtypes represented by the three murine models utilized.

1.2.1.1. MMTV-PyMT

The Tg(MMTV-PyVT)634Mul/J transgenic strain (Jackson Laboratory, #002374) is among the most widely used genetically engineered mouse models for cancer research.

MATERIALS AND METHODS

Introduced in 1992 [211], this model carries the PyMT antigen under the control of the MMTV promoter, resulting in the spontaneous development of highly fibrotic, multifocal tumors in 100% of female mice. Tumors progress through distinct stages: hyperplasia (~4 weeks), mammary intraepithelial neoplasia (MIN) or adenoma (~9 weeks), and carcinoma (~14 weeks), with lung metastases occurring in 80–90% of tumor-bearing females [211,255].

PyMT tumors typically exhibit a microacinar histology, featured by small, less organized glandular-like structures, but they can also progress to more advanced forms, such as papillary formations, which display more organized and well-differentiated glandular-like projections [256]. Gene expression profiling clustered PyMT tumors with luminal B human breast cancers [222,223]. Although molecularly classified as luminal-like, these tumors are negative for ER α and PR, limiting their utility for studying hormone-dependent breast cancers [257]. Nevertheless, they express abundant levels of AR, with a genetic profile resembling that of the LAR TNBC subtype, which is characterized by high AR expression [257].

Although the PyMT oncogene is not found in human breast cancers, it is highly potent and activates multiple signaling pathways that are critical to human breast cancer progression. PyMT oncogene activation occurs through a stepwise process. PyMT binds PP2A (protein phosphatase 2A), forming a complex that recruits and activates c-Src. This interaction leads to the phosphorylation of PyMT at specific tyrosine residues: Tyr250, binding ShcA (SHC-transforming protein 1); Tyr315, binding PI3K; and Tyr322, binding PLC γ 1 (phospholipase C gamma 1). These interactions activate critical signaling pathways, including Ras/Raf/MEK/MAPK, PI3K/PDK1/Akt, and PLC γ /PKC, which drive tumor transformation [258]. Additionally, the activation of STAT3 (signal transducer and activator of transcription 3) via Tyr705 phosphorylation, mediated by c-Src or JAK (Janus kinase), plays a crucial role in PyMT-driven progression [259-264].

The genetic background of MMTV-PyMT mice affects tumor onset and metastasis rates. Tumors develop faster in Friend Virus B-type (FVB) mice compared to the C57 Black 6 (C57BL/6) strain [265]. For this study, the FVB background was selected to ensure accelerated tumor development and facilitate research objectives.

1.2.1.2. C3(1)-TAg

The Tg(C3-1-TAg)cJeg/JegJ mouse model (Jackson Laboratory, #013591) expresses SV40-TAg (simian virus 40 large T antigen) under the control of the C3(1) promoter, resulting in prostate cancer in males and mammary tumors in females [215]. Female hemizygous mice develop multifocal mammary tumors, progressing from low-grade MIN (8 weeks) to high-grade MIN (12 weeks), and adenocarcinomas (21 weeks), with penetrance exceeding 90%, and an average latency of palpable tumors of 19 weeks [206,266]. The lung is the primary site for metastasis, although they have also been detected in the liver, adrenal glands, and heart [266]. FVB females present a 15-30% incidence of pulmonary metastases and typically require euthanasia by 7 months of age due to the development of large mammary adenocarcinomas [266].

MATERIALS AND METHODS

Histologically, C3(1)-TAg tumors are characterized by irregular, poorly differentiated glandular structures [215]. These tumors are triple-negative (ER α , PR, and HER2-negative), resembling aggressive human basal-like cancers, often with *TP53* and *RB1* (RB transcriptional corepressor 1) loss of function, and a minor proportion resembling claudin-low tumors [222,223]. C3(1)-TAg-driven tumorigenesis results from p53 (cellular tumor antigen p53) and Rb inactivation through TAg, mirroring gene disruptions in human cancers [267,268] (**Table I1**). Additional oncogenic drivers include *Kras* amplification and MAPK (mitogen-activated protein kinase) activation, accelerating tumor development [269].

Similar to the PyMT model, tumor progression is faster in FVB mice than in C57BL/6 mice, making FVB the preferred strain for this study [266].

1.2.1.1. DMBA-induced breast cancer model

The DMBA-induced model involves weekly oral administration of 1 mg/mouse DMBA (10 mg/ml in sunflower oil) for six consecutive weeks. In FVB mice, this protocol achieves a 75% tumor penetrance by 34 weeks of age [270]. Importantly, DMBA can also lead to the development of tumors in other organs, including the lungs, bone, lymph nodes, and skin, all of which can adversely affect survival.

DMBA tumors are highly heterogeneous, varying in location, histology, and molecular subtype. Long-latency tumors generated by DMBA alone predominantly align with luminal-like subtype and, less frequently, basal-like subtype [228] (**Table I1**). When combined with hormone treatments such as MPA, DMBA shortens tumor latency and primarily promotes the development of normal-like and claudin-low subtypes [222,223,229], although basal-like and HER2-enriched subtypes can also emerge under these conditions [228].

DMBA is a polycyclic aromatic hydrocarbon that activates the aryl hydrocarbon receptor (AhR), triggering transcription of AhR-responsive genes. Metabolism of DMBA by cytochrome P450 enzymes generates mutagenic epoxide intermediates that form DNA adducts, driving tumorigenesis [271]. Long-latency tumors often harbor *Pik3ca* and/or *Pten* (phosphatase and tensin homolog) mutations, with 82% carrying the *Pik3ca* H1047R/L mutation, the most frequent mutation in human breast cancers [228]. In contrast, short-latency tumors exhibit mutations in drivers like *Hras* (Harvey rat sarcoma virus oncogene) and *Apc*, rarely observed in human cancers. DMBA tumors display *AhR*, *Ccnd1* (cyclin D1), and *Myc* overexpression, Rb inactivation, and upregulation of NF- κ B and Wnt signaling pathways, contributing to tumorigenesis [270]. MPA+DMBA-induced mutations also activate immunosuppressive genes like *Ptgs2* (COX-2) and the immune checkpoint gene *Cd274* (PD-L1), fostering a pro-tumorigenic microenvironment [229]. These features are characteristic of claudin-low tumors in the human METABRIC (Molecular Taxonomy of Breast Cancer International Consortium) dataset [272].

MATERIALS AND METHODS

1.2.2. Murine transgenic lines for lineage tracing experiments in breast tumors

1.2.2.1. Inducible Cre transgenic lines

The different preclinical breast cancer models were mated with various tamoxifen-inducible Cre lines, where Cre recombinase activity is driven by either the ubiquitin C (*Ubc*) promoter, which indiscriminately tags all cell types, or by lineage-specific promoters, each labeling a distinct mammary epithelial cell population. These Cre lines include:

- *Ubc*-Cre^{ERT2} (B6.Cg-Ndor1^{Tg(UBC-cre/ERT2)1Ejb/1J}) transgenic mice [273], generously provided by Dr M. Graupera (IJC), for random labeling of all cell types.
- *Acta2*-Cre^{ERT2} (Tg(*Acta2*-cre/ERT2)12Pcn) transgenic mice [274], kindly provided by Prof P. Chambon (IGBMC), for labeling and tracing BaCs [83].
- *Notch1*-Cre^{ERT2} (*Notch1*^{tm1.1(cre/ERT2)Sat}) knock-in mice [275], gifted by Dr S. Fre (Institut Curie), for labeling and tracing ER α ^{neg} LCs [4].
- *Prom1*-Cre^{ERT2} (129S-Prom1^{tm1(cre/ERT2)Gilb/J}) knock-in mice [276], obtained from Jackson Laboratory (#017743), for labeling and tracing ER α ^{pos} LCs [86].

1.2.2.2. Rosa26^{mTmG} reporter line

The Cre lines were crossed with the Rosa26^{mTmG} (Gt(ROSA)26Sortm4(ACTB-tdTomato,-EGFP)Luo/J) double fluorescent reporter mouse strain (Jackson Laboratory, #007676) [277]. This reporter line enables the fluorescent visualization and long-term tracing of targeted cellular populations. The Rosa26 gene in this line is engineered to express two fluorescent proteins: mTomato (membrane-localized tdTomato) and mGFP (membrane-localized EGFP). Initially, *mTomato* is constitutively expressed in all cells, providing a baseline fluorescence. Cre-mediated recombination switches the expression from mTomato (red) to mGFP (green), specifically in cells where Cre is active. This transition allows for precise tracking of clonal expansion and differentiation in mGFP-positive cells, facilitating comprehensive investigations of cellular dynamics within each experimental framework.

1.2.2.3. Establishment of murine transgenic lines for lineage tracing

The three breast cancer models were crossed with the inducible CreER^{T2} lines and the Rosa26^{mTmG} reporter line to conduct lineage tracing experiments. **Figure M1** outlines the breeding strategy used to generate transgenic murine lines. These models enabled the specific labeling and tracing of the three mammary epithelial cell compartments throughout tumor progression.

MATERIALS AND METHODS

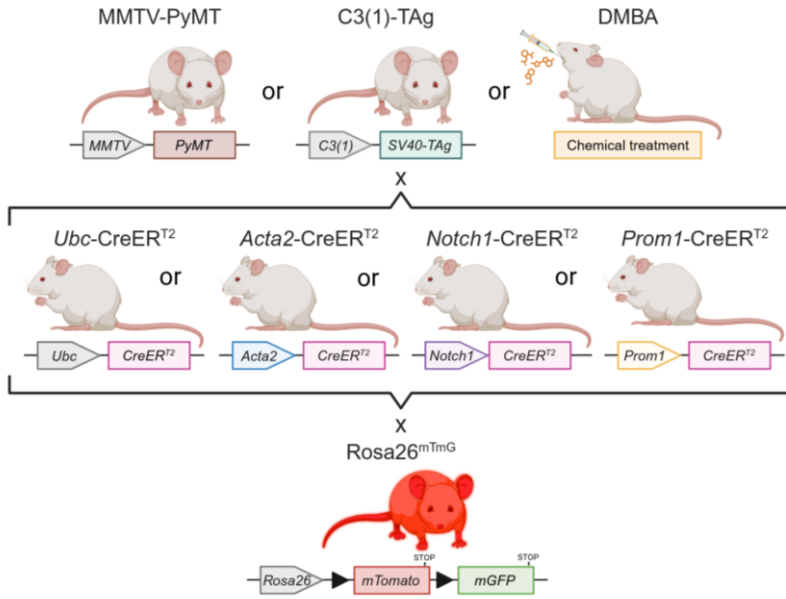


Figure M1. Schematic representation of the breeding strategy used to generate transgenic murine lines for lineage tracing. The genetically engineered (MMTV-PyMT and C3(1)-TAG) and chemically induced (DMBA) breast cancer models were crossed with *Ubc*-, *Acta2*-, *Notch1*-, and *Prom1*-CreER^{T2} lines and the Rosa26^{mTmG} reporter line. This approach enabled the specific labeling and tracking of distinct mammary epithelial cell populations throughout tumor progression. Figure created with BioRender.

2. Polymerase chain reaction (PCR)-based genotyping of transgenic mice

2.1. Weaning and ear punching

Mouse offspring were housed with their parents for three weeks after birth. At this point, they were sexed, labeled, and biopsied for genotyping. Ear punches were collected to identify mice carrying the Rosa26^{mTmG} reporter construct through fluorescence microscopy. Positive identification was based on the emission of red fluorescence. Ear biopsies displaying red fluorescence were selected for further DNA extraction to detect the non-fluorescent transgenes by PCR.

2.2. DNA extraction from ear biopsies

DNA extractions from ear punches were performed using the Hot Sodium Hydroxide and Tris (HotSHOT) method described by *Truett et al.* [278], a rapid and cost-effective protocol. Ear tissue samples were placed into individual Eppendorf tubes containing 75 μ L of alkaline lysis buffer (Solution A), composed of 25 mM sodium hydroxide (NaOH) and 0.2 mM ethylenediaminetetraacetic acid (EDTA). The samples were incubated in a thermocycler at 98°C for 1 hour to facilitate tissue lysis, then cooled to 15°C. Next, 75 μ L of neutralizing reagent (Solution B), containing 40 mM Tris-HCl at pH 5.5, was added to neutralize the reaction. Samples were centrifuged at 4000 rpm for 3 minutes to pellet debris, and 2 μ L of the supernatant was used as a template for PCR amplification. This efficient DNA extraction method was well-suited for routine genotyping.

MATERIALS AND METHODS

2.3. PCR genotyping

Genotyping of transgenic mice was performed using PCR with specific primers designed to amplify the targeted DNA sequences. Two separate PCR assays were conducted to identify the mutant (mut) and wild-type (wt) versions of the *PyMT* and *SV40 TAg* transgenes in the MMTV-*PyMT* and C3(1)-*TAg* models, respectively. For the four CreER^{T2} lines (*Ubc*, *Acta2*, *Notch1*, and *Prom1*), a PCR was conducted to detect the *Cre* transgene, using the *Rosa26* gene as an internal positive control.

PCR reactions were prepared using the 2x PCRBIO HS Taq Mix Red, containing PCRBIO HS Taq DNA polymerase, 6 mM MgCl₂, 2 mM dNTPs, enhancers, stabilizers, and a tracking dye for gel electrophoresis. Each 50 µL reaction included 2 µL of DNA template and forward and reverse primers designed by Integrated DNA Technologies (IDT) for each target sequence (**Table M2**). PCR-grade distilled water (dH₂O) was used to adjust the final volume. Negative controls were included in each PCR run to monitor contamination or non-specific amplification.

Target gene	Forward primer sequence 5' → 3'	Reverse primer sequence 5' → 3'	Size product
<i>PyMT</i> wt	CAAATGTTGCTTGTCTGGTG	GTCAGTCGAGTGCACAGTTT	205 bp
<i>PyMT</i> mut	CGCACATACTGCTGGAAGAA	TGCCGGGAACGTTTTATTAG	453 bp
<i>SV40 TAg</i> wt	CTCCCAACCCAGAGGTAGT	AGACCCAGATCCAGAAAGG	320 bp
<i>SV40 TAg</i> mut	CAGAGCAGAATTGTGGAGTGG	GGACAAACCACAAGTGAATGCAGTG	500 bp
<i>Rosa26</i>	CTTTAAGCCTGCCAGAAGA	AGGGAGCTGCAGTGGAGTAG	211 bp
<i>Cre</i>	GTAGTTATTCGGATCATCAGCT	GCTGCCACGACCAAGTGACAGC	402 bp

Table M2. Primer sequences used for PCR-based genotyping of transgenic murine lines. The table includes the forward and reverse primer sequences (5' → 3'), and the size of the resulting PCR product in base pairs (bp).

The specific reagents, volumes, and conditions for each PCR are outlined in **Table M3**.

Agarose gel electrophoresis was performed to separate and visualize PCR products by size. A 2% agarose gel was prepared by dissolving 1 g of agarose in 50 mL of 1x Tris-Acetate-EDTA (TAE) buffer, followed by heating until fully dissolved. After cooling slightly, GelRed dye was added, and the solution was poured into a gel tray. Once solidified, the gel was placed in an electrophoresis chamber filled with 1x TAE buffer. PCR samples (15 µL) and DNA ladder (2 µL) were loaded into the wells. Electrophoresis was conducted at 130 volts for 20-30 minutes, and DNA bands were visualized using UV light.

MATERIALS AND METHODS

PCR	Reagents	Volume	Concentration	PCR conditions
PyMT wt / PyMT mut	2x PCRBIO HS Taq Mix Red	25 µL	1x	94°C 2 min
	5 µM PyMT wt primers	1.5 µL	150 nM	94°C 20 sec
	5 µM PyMT mut primers	5 µL	500 nM	62°C 30 sec
	Template DNA	2 µL	5-500 ng	72°C 30 sec
	PCR grade dH ₂ O	16.5 µL		72°C 7 min
SV40 TAg wt / SV40 TAg mut	2x PCRBIO HS Taq Mix Red	25 µL	1x	94°C 2 min
	5 µM SV40 TAg wt primers	5 µL	500 nM	94°C 20 sec
	5 µM SV40 TAg mut primers	5 µL	500 nM	60.5°C 25 sec
	Template DNA	2 µL	5-500 ng	72°C 30 sec
	PCR grade dH ₂ O	13 µL		72°C 7 min
Rosa 26 / Cre	2x PCRBIO HS Taq Mix Red	25 µL	1x	95°C 2 min
	5 µM Rosa 26 primers	1 µL	100 nM	95°C 20 sec
	5 µM Cre primers	2 µL	200 nM	60°C 30 sec
	Template DNA	2 µL	5-500 ng	72°C 30 sec
	PCR grade dH ₂ O	20 µL		72°C 7 min

Table M3. Reagents, volumes, and PCR conditions for genotyping transgenic murine lines. This table summarizes the reagents, their respective volumes, and final concentrations used in each 50 µL PCR reaction to genotype transgenic mice for the *PyMT* (wild-type and mutant), *SV40 TAg* (wild-type and mutant), *Rosa26*, and *Cre* genes. It also outlines the specific PCR cycling conditions, including denaturation, annealing, and elongation steps, for each target gene.

3. Bulk RNA sequencing

Bulk RNA sequencing was conducted to characterize the transcriptomic profiles of the preclinical breast cancer models used in this thesis. This approach provided a comprehensive analysis of gene expression patterns, allowing for the classification of molecular subtypes and alignment with established breast cancer classifications.

Total RNA was extracted from tumor samples using the NZY miRNA Isolation & RNA Clean-up Kit (NZYtech), a method optimized to isolate high-quality RNA while removing impurities such as genomic DNA and contaminants. RNA purity and concentration were assessed using the Agilent TapeStation with the RNA ScreenTape Analysis Kit (Agilent Technologies). This quality control step ensured that only RNA of sufficient integrity and purity was used for subsequent mRNA enrichment and sequencing.

The cDNA libraries and RNA sequencing were outsourced to BGI Genomics, which utilizes next-generation sequencing technology based on DNA nanoballs for high-throughput sequencing. Following sequencing, the raw data were preprocessed using *SOAPnuke* software, also developed by BGI Genomics, to ensure high-quality datasets for downstream analysis. RNA sequencing data were analyzed using the Galaxy workbench platform [279], following standard bioinformatic protocols [280]. Quality control and read trimming were performed using *MultiQC* to summarize quality metrics [281], and *Cutadapt* to remove adapter sequences [282]. The cleaned reads were aligned to the Grcm38 (mm10) mouse reference genome using *STAR* [283], which efficiently handles splicing events. Gene-level read counts were obtained with *featureCounts* [284], which quantifies the number of reads aligned to each gene, providing gene expression levels. Finally, differential gene expression analysis was performed using *DESeq2* [285], which normalized the read counts and identified differentially expressed genes.

MATERIALS AND METHODS

4. Lineage tracing experiments

Postnatal recombination of the inducible Cre recombinase was achieved in each breast cancer model through intraperitoneal injection of tamoxifen (10 mg/mL). Tamoxifen was dissolved in a 1:9 mixture of absolute ethanol and sunflower oil and administered at a dose of 0.1 mg/g of body weight using 26G needles. Pulse and chase experiments were designed with tailored injection timings to address the specific objectives of each study.

At each experimental endpoint, mammary glands and/or lungs were harvested and processed for flow cytometry and/or immunohistochemical analysis, depending on the specific objectives of the study. This comprehensive approach enabled the characterization of recombination efficiency, clonal expansion, and metastatic behavior across the different models and experimental conditions.

5. Flow cytometry (FC)

Dissected mammary glands were finely chopped into 2 mm³ tissue fragments using a scalpel. The tissue was then placed into GentleMACS C Tubes containing Gibco™ Dulbecco's Modified Eagle Medium F12 (DMEM/F12) without red phenol, supplemented with L-glutamine, 15 mM HEPES, and an enzymatic cocktail (600 U/mL collagenase and 200 U/mL hyaluronidase). The tubes were processed in a GentleMACS Octo Dissociator with Heaters, running the 37C_m_TDK_2 program for 40 minutes at 37°C to achieve mechanical and enzymatic digestion. Following dissociation, red blood cells were lysed using eBioscience™ 1X RBC Lysis Buffer to enrich the epithelial population. The resulting cell suspension was filtered through a 70 µm cell strainer to obtain a single-cell preparation for flow cytometry.

Single cells were isolated and stained with fluorochrome-conjugated antibodies targeting established markers to identify distinct mammary epithelial populations by FC [4]. Specifically, CD24 (cluster of differentiation 24), CD29 (integrin β1), EpCAM (epithelial cell adhesion molecule), CD49f (integrin α6), and Sca-1 (stem cell antigen-1). BaCs were defined as CD24^{low}CD29^{high} and EpCAM^{low}CD49f^{high}, while LCs were defined as CD24^{high}CD29^{low} and EpCAM^{high}CD49f^{low}. ERα^{pos} and ERα^{neg} LCs were further identified by Sca-1 expression, being Sca-1^{high} and Sca-1^{neg} [242], respectively.

Additionally, a biotinylated lineage (Lin) antibody panel was used to exclude hematopoietic cells (Lin^{pos}). Cells were further stained with Brilliant Violet 510™ (BV510) Streptavidin as a secondary antibody. Dead cells were excluded using DAPI (4',6-diamidino-2-phenylindole), a DNA-binding dye, and the samples were transferred to flow cytometry tubes with 35 µm strainer caps for analysis.

Cells were incubated with primary antibodies for 20 minutes at 4°C in flow buffer, composed of DMEM/F12 without red phenol, 1% bovine serum albumin (BSA), 1% fetal bovine serum (FBS), and 5 mM EDTA. The specific antibodies and their properties are listed in **Table M4**.

MATERIALS AND METHODS

Antibody	Reference	Supplier	Specie	Dilution
Anti-mouse CD24 PerCP/Cyanine5.5	101824	BioLegend	Rat	1:300
Anti-mouse CD326 (Ep-CAM) PerCP/Cyanine5.5	118220	BioLegend	Rat	1:300
Anti-mouse/rat CD29 APC	102216	BioLegend	Armenian Hamster	1:300
Anti-human/mouse CD49f APC	313616	BioLegend	Rat	1:300
Anti-mouse Ly-6A/E (Sca-1) APC/Cyanine7	108126	BioLegend	Rat	1:300
Anti-mouse TER-119/Erythroid cells Biotin	133307	BioLegend	Armenian Hamster	1:300
Anti-mouse Ly-6G/Ly-6C (Gr-1) Biotin	133307	BioLegend	Armenian Hamster	1:300
Anti-mouse CD3ε Biotin	133307	BioLegend	Armenian Hamster	1:300
Anti-mouse CD45R/B220 Biotin	133307	BioLegend	Armenian Hamster	1:300
Brilliant Violet 510 Streptavidin	405233	BioLegend	N/A	1:300

Table M4. Flow cytometric antibodies used to identify mouse cell surface markers. This table includes fluorochrome-conjugated and biotinylated primary antibodies, along with the corresponding streptavidin secondary antibody used for detection. The table provides antibody references, species of origin, and dilution factors for each reagent.

Compensation controls were established to correct for spectral overlap between fluorochromes. Bead-based controls were prepared using compensation beads stained with isotype control antibodies conjugated to each fluorochrome, while cell-based controls accounted for autofluorescence. **Table M5** details the isotype controls used.

Antibody	Reference	Supplier	Specie	Dilution
κ Isotype Ctrl PerCP/Cyanine5.5	400426	BioLegend	Rat	1:60
κ Isotype Ctrl APC	400412	BioLegend	Rat	1:60
κ Isotype Ctrl APC/Cyanine7	400422	BioLegend	Rat	1:60
κ Isotype Ctrl FITC	400634	BioLegend	Rat	1:60
κ Isotype Ctrl BV510	400435	BioLegend	Rat	1:60

Table M5. Isotype controls used in flow cytometry. This table lists the isotype control antibodies utilized for flow cytometry experiments, including their references, suppliers, species of origin, and dilution factors.

All fluorophores used in this study, including fluorescent dyes, conjugated antibodies, and endogenous reporter proteins, along with their excitation and emission peaks, labeled markers, targeted cell populations, and corresponding compensation controls, are summarized in **Table M6**.

MATERIALS AND METHODS

Fluorophore	Excitation peak (nm)	Emission peak (nm)	Labeled marker(s)	Targeted cell population(s)	Compensation control
DAPI dye	405	455	DNA	Dead cells	DAPI-stained cells
BV510-conjugated antibody	405	510	CD3ε, Ly-6G/Ly-6C, CD45R/B220, TER-119	Hematopoietic (Lin ^{pos}) cells	Isotype control beads conjugated to BV510
GFP endogenous protein	488	525	-	GFP ^{pos} cells	Isotype control beads conjugated to FITC
PerCP/Cyanine5.5-conjugated antibody	488	695	CD24, CD326 (EpCAM)	BaCs (low expression) and LCs (high expression)	Isotype control beads conjugated to PerCP/Cy5.5
tdTomato endogenous protein	488	575-595	-	tdTomato ^{pos} cells	tdTomato ^{pos} cells
APC-conjugated antibody	633	660	CD29, CD49f	BaCs (high expression) and LCs (low expression)	Isotype control beads conjugated to APC
APC/Cyanine7-conjugated antibody	633	780	Ly-6A/E (Sca-1)	ERα ^{pos} LCs	Isotype control beads conjugated to APC/Cy7

Table M6. Summary of fluorophores used in flow cytometry analysis. This table includes their excitation and emission wavelengths (nm), associated labeled markers, targeted cell populations, and corresponding compensation controls.

Samples were analyzed using a BD FACSCanto II cytometer. Data was processed in FlowJo for gating and population identification and visualized in GraphPad Prism. Key gating steps included (**Figure M2**):

- **SSC-A vs FSC-A (Side Scatter Area vs Forward Scatter Area):** This gating step distinguished larger cells from smaller debris or cellular fragments. FSC-A correlates with cell size, while SSC-A reflects cell complexity or internal granularity.
- **SSC-W vs SSC-A (Side Scatter Width vs Side Scatter Area):** This gate was used to isolate single cells from clumps or doublets. SSC-W measures the width of the scatter signal, and SSC-A measures the overall intensity of side scatter. By comparing these parameters, doublets (two cells stuck together) were effectively excluded from the analysis.
- **Pacific blue (DAPI) vs FSC-A:** Dead cells were excluded by gating on DAPI-negative cells, ensuring that only live cells were included in the subsequent analysis.
- **FSC-A vs BV510 (Lin):** The BV510 signal, labeling the lineage panel, was used to exclude cells positive for hematopoietic markers, thereby selecting for the non-lineage population (Lin^{neg}), which is enriched for MECs.
- **PerCP-Cyanine5.5 (CD24 or EpCAM) vs APC (CD29 or CD49f):** This gating strategy differentiated stromal cells from MECs within the Lin^{neg} population. Among

MATERIALS AND METHODS

MECs, BaCs were identified as $CD24^{low}CD29^{high}$ or $EpCAM^{low}CD49f^{high}$, while LCs were classified as $CD24^{high}CD29^{low}$ or $EpCAM^{high}CD49f^{low}$.

- *PE (tdTomato) vs FITC (GFP)*: Endogenous tdTomato (TOM) and GFP fluorescence were analyzed within the MEC population to determine recombination efficiency for each lineage-specific promoter either across the total epithelial fraction or within specific MEC types (BaCs, Sca-1^{pos} LCs, and Sca-1^{neg} LCs). Through this gating, TOM^{pos} GFP^{pos} cells were identified in pulse experiments, as exemplified in **Figure M2**, while TOM^{neg} GFP^{pos} cells were detected in chase experiments. For lineage specificity assessments, BaCs and LCs were further plotted within the GFP^{pos} cells.
- *FITC (GFP) vs APC-Cyanine7 (Ly-6A/E)*: Ly-6A/E (Sca-1) expression was used to identify ER α ^{pos} cells, while the absence of Sca-1 identified ER α ^{neg} cells within the luminal compartment. These populations were analyzed alongside GFP expression to determine the proportions of Sca-1^{pos} and Sca-1^{neg} LCs that were recombined (GFP^{pos}) or non-recombined (GFP^{neg}).

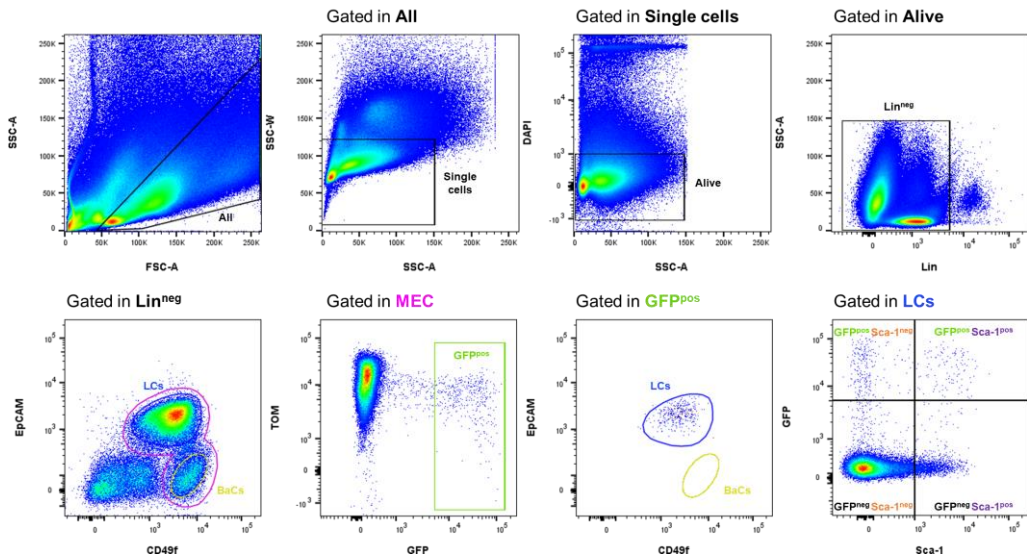


Figure M2. Flow cytometry gating strategy illustrating recombination efficiency and lineage specificity during pulse experiments. The gating steps used to identify specific cell populations are as follows: (1) SSC-A vs FSC-A to select larger (all) cells; (2) SSC-W vs SSC-A within all cells to isolate single cells; (3) DAPI vs SSC-A within single cells to select alive cells (4) SSC-A vs Lin within alive cells to identify Lin^{neg} cells (5) EpCAM vs CD49f within Lin^{neg} cells to separate stromal cells from total MECs, including BaCs and LCs; (6) TOM vs GFP within MECs to detect TOM^{pos} GFP^{pos} cells, (7) EpCAM vs CD49f within GFP^{pos} MECs to identify recombined BaCs and LCs; and (8) GFP vs Sca-1 within LCs to distinguish non-recombined and recombined ER α ^{pos} (Sca-1^{pos}) and ER α ^{neg} (Sca-1^{neg}) subpopulations.

6. Histological immunostainings

Mammary glands and/or lungs were dissected and fixed in 4% formaldehyde for 2–4 hours, depending on tumor size. Tissues were washed with PBS (phosphate-buffered saline) and processed in an automated tissue processor, employing an ethanol gradient for dehydration over 11 hours. Clearing was performed with xylene, followed by paraffin

MATERIALS AND METHODS

infiltration and embedding. Paraffin-embedded tissues were sectioned into 4 μm slices using a microtome, mounted onto glass slides, and dried at 37°C. These slides were subsequently used for immunostaining procedures, including immunohistochemistry and immunofluorescence.

6.1. Immunohistochemistry (IHC)

IHC visualizes antigens in formaldehyde-fixed, paraffin-embedded tissue using antigen-specific primary antibodies and enzyme-conjugated secondary antibodies. The secondary antibody is typically conjugated to horseradish peroxidase (HRP), which reacts with DAB (3,3'-diaminobenzidine) to produce a brown precipitate indicating antigen localization. The process includes:

- *Dewaxing and hydration*: tissue sections were incubated at 65°C for 1 hour to soften the paraffin wax. Dewaxing was performed using xylol (a hydrophobic solvent), which effectively dissolved the paraffin to expose the tissue. Following xylol treatment, the samples were rehydrated through a graded ethanol (EtOH) series to gradually replace the xylol with water, necessary for subsequent aqueous staining steps. The rehydration sequence was as follows: Xylol (x4) → EtOH-100% (x3) → EtOH-96% (x3) → EtOH-70% → dH₂O.
- *Antigen retrieval*: Tissue samples underwent heat-induced epitope retrieval in 1x citrate buffer at pH 6. They were heated to boiling for 20 minutes to unmask epitopes by breaking formaldehyde cross-links formed during fixation, enhancing antibody binding. After boiling, the samples were cooled for 1 hour at room temperature (RT).
- *Endogenous peroxidase activity blockage*: To prevent non-specific staining from endogenous peroxidase activity, sections were incubated with 3% hydrogen peroxide in methanol for 10 minutes. This step deactivates endogenous peroxidases, which could otherwise cause background staining upon DAB incubation.
- *Permeabilization*: Tissues were permeabilized with 0.3% Triton X-100, diluted in blocking buffer (5% FBS and 2% BSA in PBS), to enhance antibody penetration and block non-specific binding sites.
- *Primary antibody incubation*: Samples were incubated overnight at 4°C with primary antibodies, diluted in blocking buffer, specifically binding the proteins indicated in **Table M7**.
- *Secondary antibody incubation*: After washing the primary antibody, samples were incubated with HRP-conjugated secondary antibodies detailed in **Table M8**. This enzyme-linked antibody allows for visualization of the GFP through the subsequent DAB reaction.
- *DAB precipitation*: Using the SignalStain® DAB Substrate Kit, 300 μL of DAB working solution (30 μL DAB Chromogen Concentrate in 1 mL DAB Diluent) was applied to each slide for 1 to 5 minutes. The HRP enzyme catalyzed a colorimetric reaction with the DAB substrate, producing a brown precipitate at the site of the protein expression, facilitating accurate localization of target cells within the tissue

MATERIALS AND METHODS

- sections. The incubation time for the DAB reaction was optimized for each protein to ensure optimal staining intensity and specificity.
- *Hematoxylin counterstaining*: Harris hematoxylin was used to counterstain cell nuclei in purple, distinguishing nuclear structures from the DAB-stained areas.
 - *Dehydration*: To prepare for permanent mounting, samples were dehydrated through a graded ethanol series, followed by xylol incubation to completely remove hydrous compounds from the slides, following this sequence: EtOH-25 → EtOH-50 → EtOH-70 → EtOH-96 (×2) → EtOH-Abs (×2) → Xylol (×3) (2' each step).
 - *Mounting*: Tumor sections were mounted with DPX mounting medium, covered with coverslips, and left to dry overnight at RT in a fume hood. This process permanently preserves the sections for imaging and analysis.
 - *Image acquisition*: IHC-stained slides were scanned using the Panoramic SCAN II and ZEISS Axioscan 7 scanners. Images were visualized and navigated in SlideViewer and ZEISS ZEN Lite software, respectively.

Antibody	Reference	Supplier	Specie	Dilution
GFP	ab13970	Abcam	Chicken	1:1000
Claudin 3	34-1700	Thermo Fisher Scientific	Rabbit	1:250
Claudin 7	PA5-32356	Thermo Fisher Scientific	Rabbit	1:250

Table M7. Primary antibodies used in IHC, with corresponding references, companies, species, and dilutions.

Antibody	Reference	Supplier	Specie	Dilution
Anti-chicken HRP	31401	Thermo Fisher Scientific	Rabbit	1:500
Anti-rabbit HRP	7074	Cell Signaling Technology	Goat	1:500

Table M8. Secondary antibodies used in IHC, with corresponding references, companies, species, and dilutions.

Quantification of DAB-positive staining, corresponding to GFP-expressing cells, was performed using *QuPath* software, and different approaches were followed depending on the purpose of each experiment:

- *Quantification of GFP-positive tumor cells per mammary gland*: Regions of interest (ROIs) within the mammary glands were defined by selecting the tumor areas, and the positive cell detection algorithm was applied to identify DAB-positive and total cells within each area. The percentage of GFP-positive cells was calculated by dividing the number of DAB-positive cells by the total number of cells in the ROI. Subsequently, the total GFP value for each mammary gland was calculated by merging the different ROIs within each gland.
- *Quantification of GFP-positive tumor clone sizes within mammary glands*: ROIs corresponding to individual clones were defined in each mammary gland, and the positive cell detection algorithm was applied to quantify the percentage of GFP-positive cells within each clone. Clones were then categorized based on size: single

MATERIALS AND METHODS

cells, 2-10 cells, and ≥ 11 cells. The data were exported to excel for analysis of the distribution of clone sizes across the mammary glands.

- *Quantification of GFP-positive tumors cells per metastasis:* ROIs were defined within the lungs by selecting metastatic nodules, and the positive cell detection algorithm was applied to quantify the percentage of GFP-positive cells within each metastasis.
- *Quantification of metastases containing GFP-positive tumor cells:* ROIs corresponding to individual metastatic nodules in the lungs were defined and categorized as either positive or negative for GFP-positive cells. The data were then exported to Excel to calculate the percentage of GFP-positive metastases.

Results were analyzed and visualized using *GraphPad Prism*.

6.2. Immunofluorescence (IF)

IF is a technique used to detect specific proteins in tissue samples using antibodies conjugated to fluorescent dyes. Multiple primary antibodies can be combined, each with a different fluorochrome, allowing the simultaneous detection of several targets on the same slide. Fluorescence emitted by the labels is captured using a fluorescence microscope, enabling the analysis of protein distribution and cellular localization. The following steps outline the procedure:

- *Dewaxing and hydration:* Sections were de-waxed with xylene and rehydrated through a decreasing ethanol gradient to prepare for staining: Xylol (x4) $\rightarrow$ EtOH-100% (x3) $\rightarrow$ EtOH-96% (x3) $\rightarrow$ EtOH-70% $\rightarrow$ dH₂O.
- *Antigen retrieval:* Sections were subjected to antigen retrieval in boiling citrate buffer for 20 minutes to unmask antigen epitopes.
- *Permeabilization:* Tissue was permeabilized with 0.3% Triton X-100, diluted in blocking buffer (5% FBS and 2% BSA in PBS), to allow antibody penetration and block non-specific binding sites.
- *Primary antibody incubation:* Samples were incubated overnight at 4°C with the chicken anti-GFP antibody, in combination with different lineage markers, which are summarized in **Table M9**.
- *Secondary antibody incubation:* After washing the primary antibody, sections were incubated with Alexa Fluor 488-conjugated anti-chicken secondary antibody to label GFP in green, combined with anti-mouse and anti-rabbit Cyanine 3 (Cy3) and/or Cyanine 5 (Cy5)-conjugated secondary antibodies recognizing the different lineage markers, also being detailed in **Table M10**.
- *Mounting:* Sections were mounted using Fluoroshield™ mounting medium with DAPI to counterstain nuclei.
- *Imaging and Processing:* Fluorescent images were acquired with a Leica STELLARIS 8 confocal fluorescence microscope, and images were processed with *ImageJ* software for analysis.

MATERIALS AND METHODS

Antibody	Reference	Supplier	Specie	Dilution
GFP	ab13970	Abcam	Chicken	1:1000
ERα	sc-8005	Santa Cruz Biotechnology	Mouse	1:100
ERα	HPA000449	Sigma-Aldrich	Rabbit	1:100
K8	ab53280	Abcam	Rabbit	1:300
K14	HPA023040	Sigma-Aldrich	Mouse	1:500
K5	ab13970	Abcam	Rabbit	1:250
P63	ab735	Abcam	Mouse	1:250
αSMA-Cy3	C6198	Sigma-Aldrich	Mouse	1:250
SV40 T-Antigen	MA1-90661	Thermo Fisher Scientific	Mouse	1:250

Table M9. Primary antibodies used in IF, with corresponding references, companies, species, and dilutions.

Antibody	Reference	Supplier	Specie	Dilution
Anti-chicken Alexa Fluor 488	ab150169	Abcam	Goat	1:2000
Anti-mouse Cyanine CyTM3	715-165-150	Jackson ImmunoResearch	Donkey	1:500
Anti-rabbit Cyanine CyTM3	711-167-003	Jackson ImmunoResearch	Donkey	1:500
Anti-mouse Cyanine CyTM5	715-175-151	Jackson ImmunoResearch	Donkey	1:500
Anti-rabbit Cyanine CyTM5	711-175-152	Jackson ImmunoResearch	Donkey	1:500

Table M10. Secondary antibodies used in IF, with corresponding references, companies, species, and dilutions.

7. Statistical analysis

All graphical representations and statistical analyses were conducted using GraphPad Prism software. Normality of the data was assessed using the Shapiro-Wilk test. For non-parametric data, the Kruskal-Wallis test was used for comparisons involving three or more groups, while the Mann-Whitney U test was applied for pairwise comparisons between two groups. One-way ANOVA was used when the data met parametric assumptions, comparing the means of multiple groups within a single factor. Two-way ANOVA was employed for analyses incorporating both categorical and continuous variables, accounting for interactions between factors. Statistical significance was interpreted using the following p-value thresholds: ns (not significant) for $p > 0.05$, * for $p \leq 0.05$, ** for $p \leq 0.01$, *** for $p \leq 0.001$, and **** for $p \leq 0.0001$.

MATERIALS AND METHODS

8. Tables of reagents, materials, and equipment

Reagent	Reference	Supplier
2x PCRBIO HS Taq Mix Red	PB10.23-02	PCR Biosystems
100 bp DNA Ladder	LR100-5	AttendBio Research
AgaPure™ Agarose LE (Low EEO)	AG006	SG Servicios Hospitalarios
BSA	821006	Sigma-Aldrich
Citrate Buffer, pH 6.0, 10×	C9999	Sigma-Aldrich
Collagenase A from Clostridium histolyticum	10103586001	Sigma-Aldrich
DAPI	D9542	Sigma-Aldrich
DMBA	408181000	Thermo Fisher Scientific
DMEM/F-12, HEPES, without red phenol	11039047	Thermo Fisher Scientific
DPX mounting medium	06522	Sigma-Aldrich
eBioscience™ 1X RBC Lysis Buffer	00-4333-57	Thermo Fisher Scientific
EDTA	E5134	Sigma-Aldrich
Ethyl alcohol, pure	51976	Sigma-Aldrich
FBS	91S1810-500	Cultek
Fluoroshield™ with DAPI	F6057	Sigma-Aldrich
Formaldehyde 3.4-4 %	2.529.311.315	Suministros Generales para Laboratorio
GelRed dye	BT-41003	LabNet Biotechnica
Harris Hematoxylin Solution, Modified	HHS32	Sigma-Aldrich
Hyaluronidase from bovine testes	H3506	Sigma-Aldrich
Hydrogen peroxide 30%	108597	Sigma-Aldrich
NaOH	10396240	Thermo Fisher Scientific
SignalStain® DAB Substrate Kit	8059	Cell Signaling Technology
TAE buffer 50X	1610743	Bio-Rad
Tamoxifen	T5648	Sigma-Aldrich
Tris-HCl	108315	Sigma-Aldrich
Triton™ X-100	108603	Sigma-Aldrich
Xylenes histological grade	534056	Sigma-Aldrich

Table M11. Reagents used in this thesis, with corresponding references and suppliers.

MATERIALS AND METHODS

Materials	Reference	Supplier
ClearLine® RNase/DNase-free microtube 1.5 mL	390690	dDBioLab
15 mL centrifuge tubes	352095	Corning
50 mL centrifuge tubes	352070	Corning
70 µm strainers	141379C	dDBioLab
2000 mL Beaker, low form	213-0469	VWR International
Adhesive Slides, SuperFrost® Plus 75 x 25 mm	631-0108	VWR International
Bacteriological Petri Dishes	P101	Bellés Diagnòstic
BD Microlance 26G Hypodermic Needles, 13 mm, Sterile	303800	Becton Dickinson
BD Plastipak™ 1 mL Syringe	303172	Becton Dickinson
Cover Slips 24 x 50 mm	631-1574	VWR International
Disposable Cover Plates™	97091	Electron Microscopy Sciences
Flow cytometry tubes	352235	Cultek
GentleMACS C Tubes	130-096-334	Miltenyi Biotec
Holder for 20 slides	391058	dDBioLab
Pasteur Pipettes High-Performance 7 mL	612-1681	VWR International
PCR Tubes & Caps, RNase-free, 0.2 mL	AM12230	Thermo Fisher Scientific
Student Anatomical Standard Pattern Forceps	91100-12	Fine Science Tools
Student Fine Scissors	91460-11	Fine Science Tools
Swann-Morton™ Stainless Steel Surgical Scalpels	11748353	Thermo Fisher Scientific

Table M12. Materials used in this thesis, with corresponding references and suppliers.

MATERIALS AND METHODS

Equipment	Manufacturer
BD FACSCanto II flow cytometer	BD Biosciences
Cover Plate™ Slide Rack	Electron Microscopy Sciences
GentleMACS Octo Dissociator with Heaters	Miltenyi Biotec
Centrifuge 5427 R	Eppendorf
Confocal Microscope Stellaris 8	Leica Microsystems
Electronic Balance BJ610C	Precisa
Histo bath HIER-3D	Kunz Instruments
HistoStar™ Embedding Workstation	Epredia
HM 340E Rotary Microtome	Epredia
Hot plate PC800	Bio Optica
Incubat bacteriological and culture ovens	J.P. Selecta
LifeECO Thermal Cycler TC-96/G/H(b)C	Bioer Technology
Manual slide staining set, 300 ml capacity	Bio Optica
Pannoramic SCAN II	3DHISTECH
Refrigerated centrifuge 5810R	Eppendorf
Cole-Parmer™ Stuart™ Hot Plate Stirrer	Thermo Fisher Scientific
Syngene™ Gbox Chemi XRQ	Syngene
STP 120 Centrifugal Tissue Processor	Epredia
ZEISS Axioscan 7	ZEISS

Table M13. Equipment used in this thesis, with corresponding manufacturers.

RESULTS

Objective 1: To histologically and molecularly characterize different preclinical models of triple-negative breast cancer.

1. Distinct murine TNBC preclinical models reflect diverse human breast cancer subtypes

This study employed three distinct preclinical breast cancer models (MMTV-PyMT, C3(1)-TAg, and DMBA-induced tumors), each characterized by unique biological traits and tumor progression dynamics (**Table M1**). Comprehensive histological and transcriptomic analyses were conducted to classify murine tumor profiles and assess their alignment with human breast cancer subtypes.

1.1. Histological characterization of TNBC preclinical models

Histological analysis used IF staining of established cell identity markers [4], including P63, K5 (keratin 5), SMA (smooth muscle actin) and K14 (keratin 14) for BaCs, and K8 (keratin 8) and ER α for different luminal populations, to classify tumors into distinct subtypes (**Figure R1.1**).

The MMTV-PyMT and C3(1)-TAg models were found to develop ER α^{neg} luminal breast tumors characterized by K8 expression (**Figure R1.2**). Typically restricted to the basal compartment in the healthy mammary gland, K14 was aberrantly expressed in luminal lineages in these tumor models, consistent with previously published data [243,286]. These observations underscore the promiscuous nature of K14 expression across different tumorigenic contexts.

In contrast, the DMBA-induced model displayed a high degree of heterogeneity, resulting in the formation of three distinct histological subtypes (**Figure R1.2**). The first, referred to as Type I, constituted approximately 50% of all DMBA tumors and was defined as ER α^{neg} luminal. These tumors exclusively express K8, with basal markers limited to the adjacent healthy basal layer. The second subtype, Type II, accounted for 25% of tumors and exhibited a basal phenotype characterized by exclusive expression of P63, K5, and K14, with SMA confined to stromal fibroblasts. The third subtype, Type III, made up 25% of tumors and displayed a hybrid basal-luminal phenotype with co-expression of K8 and P63. These patterns were particularly evident in merged IF images (**Figure R1.1**).

RESULTS: OBJECTIVE 1

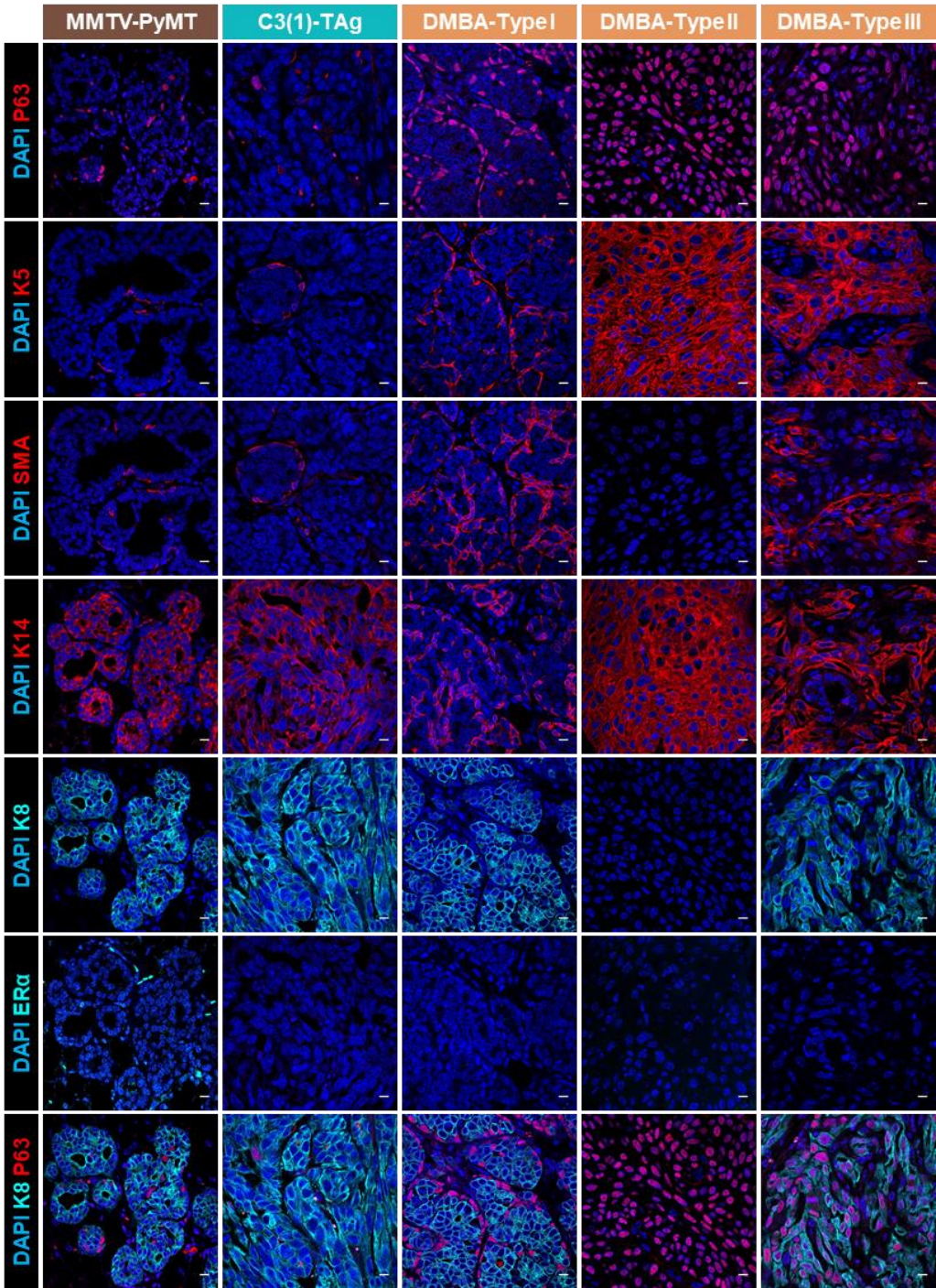


Figure R1.1. Histological characterization of different preclinical breast cancer models. Representative IF images of tumor sections from MMTV-PyMT, C3(1)-TAg and DMBA-induced models, all classified as TNBC. Basal markers (P63, K5, SMA, and K14) are displayed in red, while luminal markers (K8 and ERα) are shown in cyan. The final row presents merged images of K8 and P63 staining. DNA is counterstained with DAPI in blue. Scale bars represent 10 μm.

RESULTS: OBJECTIVE 1

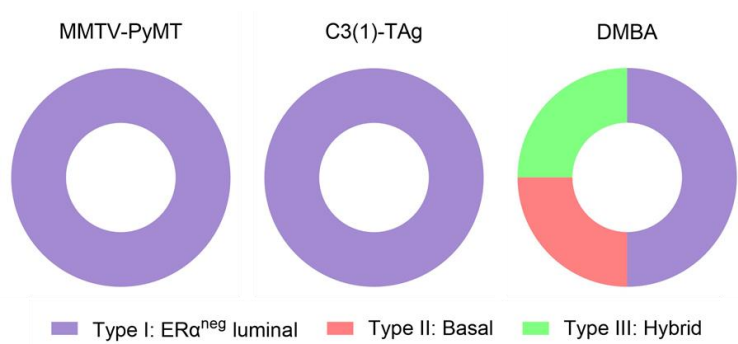


Figure R1.2. Distribution of tumor subtypes across preclinical models. The MMTV-PyMT and C3(1)-TAg models exclusively generated homogeneous ERα^{neg} luminal tumors (n=24 and 27, respectively). In contrast, the DMBA-induced model exhibited a more diverse distribution of tumor subtypes: 50% ERα^{neg} luminal (n=22), 25% basal (n=11), and 25% hybrid (n=11). Tumor subtypes were classified based on the expression of specific cell identity markers, as described in the text.

1.2. Transcriptomic characterization of TNBC preclinical models

To further characterize the models, bulk RNA sequencing was performed on tumors from the MMTV-PyMT, C3(1)-TAg, DMBA-Type I, and DMBA-Type II subtypes. The PAM50 gene signature was used to classify tumors into intrinsic molecular subtypes commonly observed in human breast cancers, including luminal A, luminal B, basal-like, HER2-enriched, and normal-like.

The transcriptomic profiles revealed that MMTV-PyMT tumors were closely aligned with the luminal B subtype, while C3(1)-TAg tumors exhibited a basal-like profile (**Figure R1.3**), consistent with prior studies [222,223]. On the other hand, DMBA-induced tumors exhibited notable molecular heterogeneity reflective of their diverse histological features. Specifically, DMBA-Type I tumors displayed gene expression patterns associated with the basal-like subtype, whereas DMBA-Type II tumors showed profiles characteristic of the normal-like subtype. These findings are consistent with previously reported molecular subtypes for the DMBA model [222,228] (**Figure R1.3**).

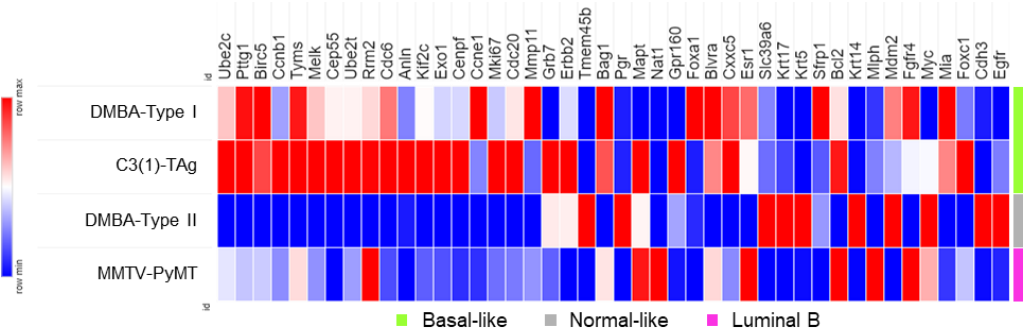


Figure R1.3. Molecular characterization of preclinical breast cancer models. Heatmap showing the transcriptomic profiles of MMTV-PyMT, C3(1)-TAg, DMBA-Type I, and DMBA-Type II tumors, classified according to the PAM50-associated genes. Gene expression levels are represented on a red-to-blue scale, with red indicating the highest expression and blue indicating the lowest. The intrinsic breast cancer subtypes associated with each model are indicated in the legend.

RESULTS: OBJECTIVE 1

Overall, **these models successfully mimic the diversity of molecular subtypes observed in clinical settings**, providing valuable platforms for studying subtype-specific mechanisms of tumorigenesis and progression.

1.3. Claudin-low phenotype in basal and hybrid tumors

About 15% of basal-like and normal-like tumors classified by PAM50 have been previously linked to the claudin-low subtype [178]. To investigate this connection further, we analyzed claudin protein expression in the different models, aiming to explore the molecular overlap between intrinsic subtypes and claudin-low tumors and provide a deeper understanding of their shared phenotypic characteristics.

Claudins are key tight junction proteins involved in maintaining cell-cell adhesion and tissue integrity. Dysregulation of claudins has been implicated in increased tumor invasiveness, EMT, and poor clinical prognosis [99,174,177,287]. To assess their expression in our tumor models, we performed immunohistochemical staining of Claudin-3 and Claudin-7.

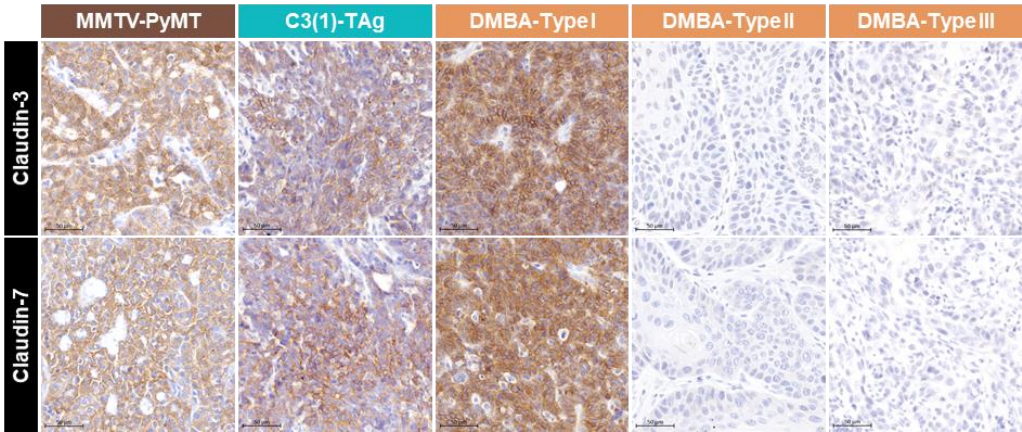


Figure R1.4. Immunohistochemical staining of claudins in MMTV-PyMT, C3(1)-TAg, and DMBA tumors. Strong expression of both Claudin-3 (top row) and Claudin-7 (bottom row) is evident in MMTV-PyMT, C3(1)-TAg, and DMBA-Type I tumors (ER α -negative luminal), whereas no detectable staining is observed in DMBA-Type II (true basal) and DMBA-Type III (hybrid) tumors. Scale bars: 50 μ m.

The analysis demonstrated that histologically luminal tumors across all three models exhibited high claudin expression (claudin-high), while histologically basal (DMBA-Type II) and hybrid (DMBA-Type III) tumors showed reduced or absent claudin expression (claudin-low) (**Figure R1.4**). Remarkably, the claudin-low profile of hybrid tumors closely resembled that of basal tumors, suggesting a potential shared molecular signature. These findings further reinforce the correlation between claudin expression and tumor subtype, emphasizing its potential as a biomarker of tumor aggressiveness.

Through comprehensive histological and transcriptomic analyses, this study revealed that the MMTV-PyMT, C3(1)-TAg, and DMBA-induced models generate triple-negative breast cancers with distinct and diverse profiles. Notably, tumors transcriptomically classified as “basal-like” demonstrated luminal characteristics at the histological level.

RESULTS: OBJECTIVE 1

Conversely, tumors exhibiting basal and claudin-low histological features were found to transcriptomically align with the normal-like subtype. To clarify this distinction, tumors with histological basal features will henceforth be referred to as “true basal” tumors, emphasizing their unique characteristics within the broader TNBC landscape.

Objective 2: To elucidate the cellular origins in multiple subtypes of triple-negative breast cancer.

2. Lineage tracing studies unravel divergent origins in TNBC subtypes

This objective aimed to determine whether distinct TNBC subtypes originate from specific MEC populations. Using tamoxifen-inducible Cre lines (**Figure R2.1**), lineage tracing experiments were conducted in MMTV-PyMT, C3(1)-TAg, and DMBA-induced TNBC models to label *Acta2*-expressing BaCs, *Notch1*-positive ER α^{neg} LCs, *Prom1*-positive ER α^{pos} LCs, and *Ubc*-positive cells for random labeling.

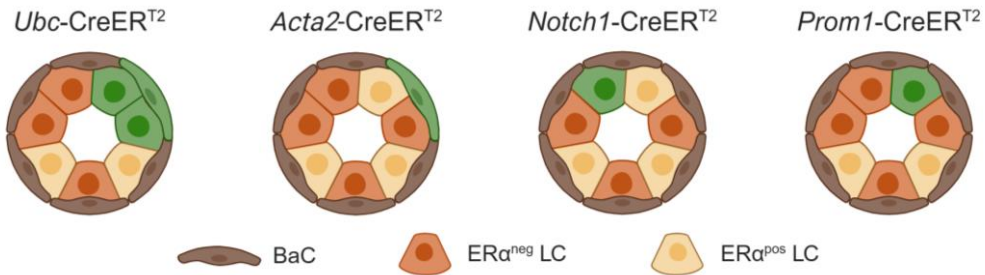


Figure R2.1. Lineage tracing of mammary epithelial cell populations using lineage-specific promoters. Representative cross-sections showing GFP-labeled epithelial cell subtypes (depicted in green) across different CreER^{T2} lines following tamoxifen administration. *Ubc*-CreER^{T2} randomly labels all cell types, *Acta2*-CreER^{T2} labels BaCs, *Notch1*-CreER^{T2} targets ER α^{neg} LCs, and *Prom1*-CreER^{T2} marks ER α^{pos} LCs. Figure created with BioRender.

First, to evaluate the baseline recombination efficiency of each Cre line, non-transformed mammary glands subjected to three consecutive tamoxifen injections were analyzed 24 hours after the final administration. MEC populations were identified through FC using well-established markers (**Figure M2**). The analysis revealed recombination rates of $46.65 \pm 1.55\%$ for *Ubc*^{mTmG}, $16.10 \pm 1.30\%$ for *Acta2*^{mTmG}, $59.10 \pm 5.15\%$ for *Notch1*^{mTmG}, and $18.95 \pm 4.35\%$, for *Prom1*^{mTmG} mammary glands (**Figure R2.2A**). However, when assessing the efficiency of the selected promoters in labeling specific MEC populations, these differences became less pronounced: in *Acta2*^{mTmG} mice, $67.25 \pm 1.65\%$ of BaCs were successfully labeled, while *Notch1*^{mTmG} exhibited labeling of $76.53 \pm 4.12\%$ of Sca-1^{neg} LCs, and *Prom1*^{mTmG} mice marked $47.25 \pm 1.25\%$ of Sca-1^{pos} LCs (**Figure R2.2B**). Using the *Ubc*^{mTmG} line, all three MEC lineages were labeled in similar proportions: $30.00 \pm 0.00\%$ of BaCs, $49.95 \pm 1.15\%$ of Sca-1^{neg} LCs, and $41.05 \pm 0.15\%$ of Sca-1^{pos} LCs (**Figure R2.2B**).

These results highlight the promoter-specific recombination efficiencies, demonstrating that each Cre line targets its specific MEC subpopulation with comparable precision. This precise labeling lays the foundation for exploring how these distinct MEC subpopulations contribute to TNBC initiation, progression, and heterogeneity, shedding light on the cellular dynamics driving tumor evolution.

RESULTS: OBJECTIVE 2

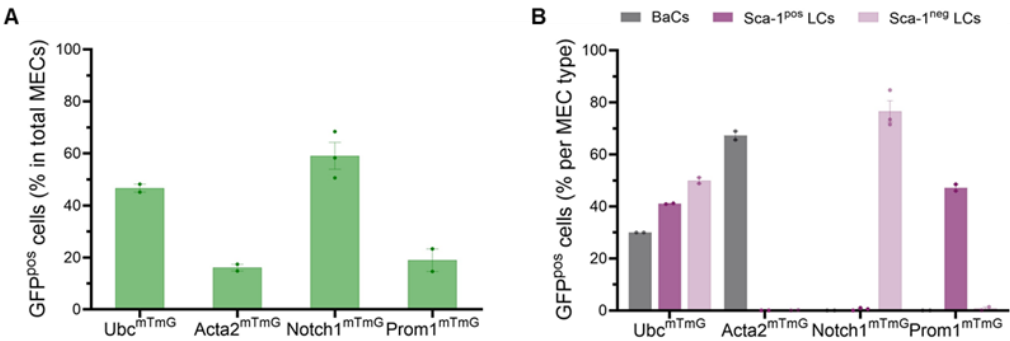


Figure R2.2. Recombination efficiency in total and specific MECs. Females were injected with three consecutive daily doses of tamoxifen at 6 weeks of age and analyzed 72 hours later using FC. **A.** Recombination rates were calculated by determining the percentage of GFP^{pos} cells within the MEC population for each Cre line. **B.** The efficiency of specific MEC type labeling (BaCs, Sca-1^{pos} LCs, and Sca-1^{neg} LCs) was assessed for each Cre line. At least two animals per genotype were analyzed. Data are presented as mean $\pm$ standard error of the mean (S.E.M).

Thus, we aimed to investigate whether distinct breast cancer subtypes arise from specific MECs. To ensure precise recombination prior to tumor onset, the timing and dosage of tamoxifen administration were meticulously adjusted to align with the tumor latency of each model (**Figure R2.3**).

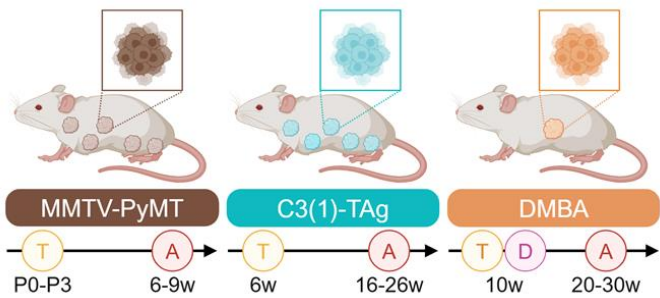


Figure R2.3. Summary of the experimental design for the MMTV-PyMT, C3(1)-Tag, and DMBA-induced breast cancer models to investigate their cell of origin. (T) indicates the time point of tamoxifen-induced recombination prior to tumor onset, (D) marks the six consecutive weekly DMBA injections, and (A) represents when tumors became palpable and were analyzed (at specific weeks of age). Figure created with *BioRender*.

In the MMTV-PyMT model, where malignant transformation occurs early (around 4 weeks of age), recombination was induced in neonates by administering a single dose of tamoxifen to lactating mothers shortly after delivery. This approach facilitated tamoxifen transfer through maternal milk to neonates during postnatal days P0–P3. This widely adopted strategy for Cre/*loxP* recombination in neonates ensures uniform and efficient recombination across the litter while minimizing stress and potential harm to the pups. Recombination was thus initiated at birth, a developmental stage when the three distinct mammary epithelial lineages had already segregated, with each lineage independently self-renewing throughout postnatal mammary gland development [4,83,86].

RESULTS: OBJECTIVE 2

In the C3(1)-TAg model, which is characterized by a longer tumor latency with hyperplasia appearing at around 8 weeks of age, we administered three consecutive daily doses of tamoxifen at 6 weeks of age, prior to malignant transformation. This triple-dose regimen ensured efficient labeling without significantly altering tumor latency or the natural progression of tumorigenesis.

In the DMBA model, a single tamoxifen dose was administered at 10 weeks of age, followed by weekly oral injections of the DMBA carcinogen for six consecutive weeks [270], beginning three days after tamoxifen exposure. Attempts to use a triple-dose tamoxifen regimen resulted in significantly delayed tumor development and increased off-target tumorigenesis in tissues such as the skin, lungs, and digestive tract, often leading to animal mortality before mammary tumors could develop.

2.1. TNBC tumors with luminal B and basal-like gene expression profiles share a common ER α -negative luminal origin

We start by investigating tumors with a luminal ER α -negative phenotype generated by the three preclinical models: MMTV-PyMT tumors, resembling the luminal B subtype, and C3(1)-Tag and DMBA-Type I tumors, which mimic human basal-like breast cancer. The experimental design outlined in **Figure R2.3** was followed for each model, and at the specified endpoints, tumor sections were stained for GFP protein using IHC to quantify the percentage of mammary glands with GFP^{pos} tumor clones.

Immunohistochemical analysis revealed a high prevalence of GFP^{pos} tumor clones in Notch1^{mTmG} mammary glands from both MMTV-PyMT and C3(1)-TAg mice, with detection rates of $72.22 \pm 14.70\%$ and $93.33 \pm 3.69\%$, respectively. These results were consistent with the findings in Ubc^{mTmG} line, where $88.57 \pm 8.57\%$ and $72.67 \pm 15.36\%$ of mammary glands contained GFP^{pos} clones, respectively (**Figure R2.4**).

Approximately 2% of the Acta2^{mTmG}-analyzed mammary glands exhibited GFP^{pos} clones in both models. In the MMTV-PyMT model, this phenomenon can be attributed to the presence of a small population of undifferentiated stem cells that co-express both luminal and basal lineage markers during the early postnatal period (P0 to P3). Conversely, in the C3(1)-TAg model, this event was observed only once and was restricted to a dysplastic lesion, suggesting that Acta2-expressing cells rarely contribute to neoplastic transformation. These findings indicate that while Acta2-positive cells may play a role in early-stage cellular processes, their contribution to tumor development appears limited.

DMBA-Type I breast tumors, which constituted 50% of DMBA-induced tumors, were characterized by K8 positivity and exclusively originated from Notch1-positive cells, with GFP-positive clones detected in 100% of analyzed Notch1^{mTmG} tumors (n=4) (**Figure R2.5**). In contrast, no GFP^{pos} clones were ever detected in six DMBA-Type I tumors analyzed in the Acta2^{mTmG} line, reinforcing the conclusion that luminal ER α -negative tumors arise exclusively from ER α ^{neg} LCs.

RESULTS: OBJECTIVE 2

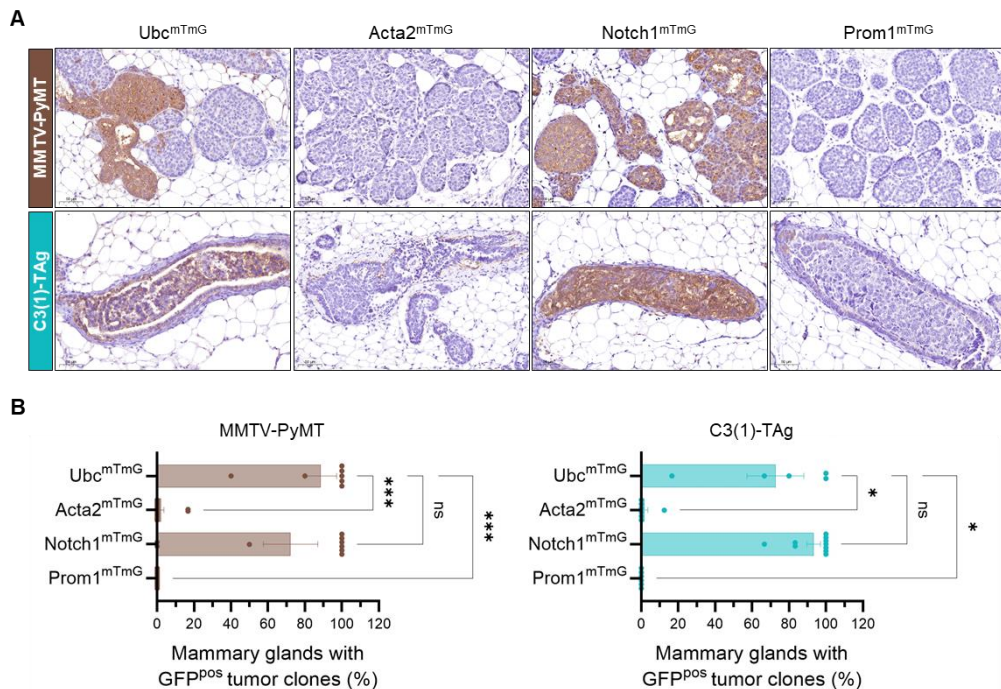


Figure R2.4. ERα^{neg} LCs as the origin of TNBC tumors with luminal traits. **A.** IHC analysis of GFP expression in tumor sections from MMTV-PyMT and C3(1)-TAG models several weeks after tamoxifen injection prior tumor onset. Transformed clones derived from the initially labeled cells in the indicated Cre lines are visualized. GFP staining appears brown, while nuclei are counterstained in blue. Scale bars represent 50 μm. **B.** Quantification of the percentage of mammary glands containing GFP^{pos} tumor clones in the indicated Cre/MMTV-PyMT models (n=36, 59, 28, 55) and the Cre/C3(1)-TAG models (n=29, 51, 52, 38). Each dot represents an individual animal, with at least three animals analyzed per time point and genotype. “n” refers to the total number of mammary glands analyzed. Data are presented as mean ± S.E.M. Statistical significance was determined using the Kruskal-Wallis test (p < 0.05).

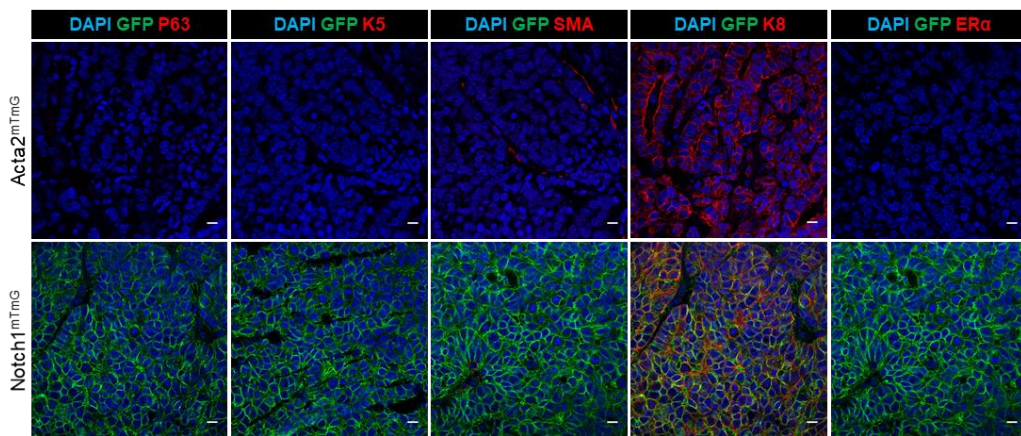


Figure R2.5. TNBC tumors with luminal traits in the DMBA-induced model originate from ER α ^{neg} LCs. Representative IF images of DMBA-Type I tumors from Acta2^{mTmG} and Notch1^{mTmG} mice. GFP is depicted in green, while basal markers (P63, K5, and SMA), and luminal markers (K8 and ER α) are shown in red. DNA is counterstained with DAPI in blue. Scale bars represent 10 μ m.

RESULTS: OBJECTIVE 2

These findings confirm that TNBC tumors with histological luminal traits arise from ER α ^{neg} LCs, regardless of their transcriptomic classification. These results highlight the intricate relationship between cellular origin and tumor subtype, emphasizing the complexity of breast cancer pathogenesis and the limitations of traditional molecular classifications in fully capturing tumor heterogeneity.

2.2. TNBC tumors with a normal-like gene expression profile have a basal origin

“True basal” tumors (DMBA-Type II), accounting for 25% of DMBA-induced tumors, were transcriptomically classified as normal-like. Lineage tracing revealed that these tumors exclusively originated from Acta2-expressing BaCs, with the only true basal tumor observed in the Acta2^{mTmG} line largely composed of GFP-positive clones (**Figure R2.6**). Conversely, no GFP-positive clones were detected in DMBA-Type II tumors from Notch1^{mTmG} and Prom1^{mTmG} mice (n=3 and 1, respectively), excluding a luminal origin.

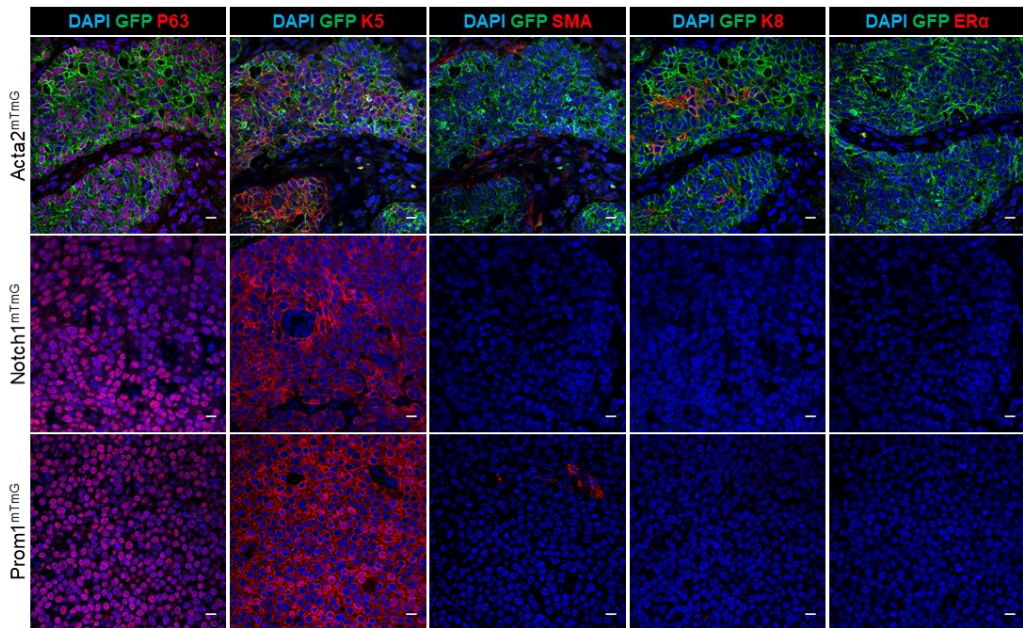


Figure R2.6. Triple-negative tumors with histological basal traits in the DMBA-induced model originate from BaCs. Representative IF images of DMBA-Type II tumors from Acta2^{mTmG}, Notch1^{mTmG}, and Prom1^{mTmG} mice. GFP is depicted in green, while basal markers (P63, K5, and SMA), and luminal markers (K8 and ER α) are shown in red. DNA is counterstained with DAPI in blue. Scale bars represent 10 μ m.

2.3. Investigating the cellular origins of TNBC tumors with hybrid features

DMBA-Type III tumors, characterized by the co-expression of luminal and basal markers, represented 25% of all DMBA-induced tumors. This unique phenotype poses critical questions regarding their cellular origins: do these tumors arise from BaCs acquiring luminal traits, or from LCs adopting basal characteristics?

Lineage tracing experiments in Acta2^{mTmG}, Notch1^{mTmG} and Prom1^{mTmG} mice (n=2, 3, and 1, respectively) did not reveal GFP^{pos} clones in any of the Cre lines (**Figure R2.7**),

RESULTS: OBJECTIVE 2

leaving their cellular origin unresolved. However, these findings do not conclusively exclude the involvement of these lineages, as recombination efficiencies for BaCs and ERα^{neg} LCs were not 100% at the early time points analyzed (Figure R2.2).

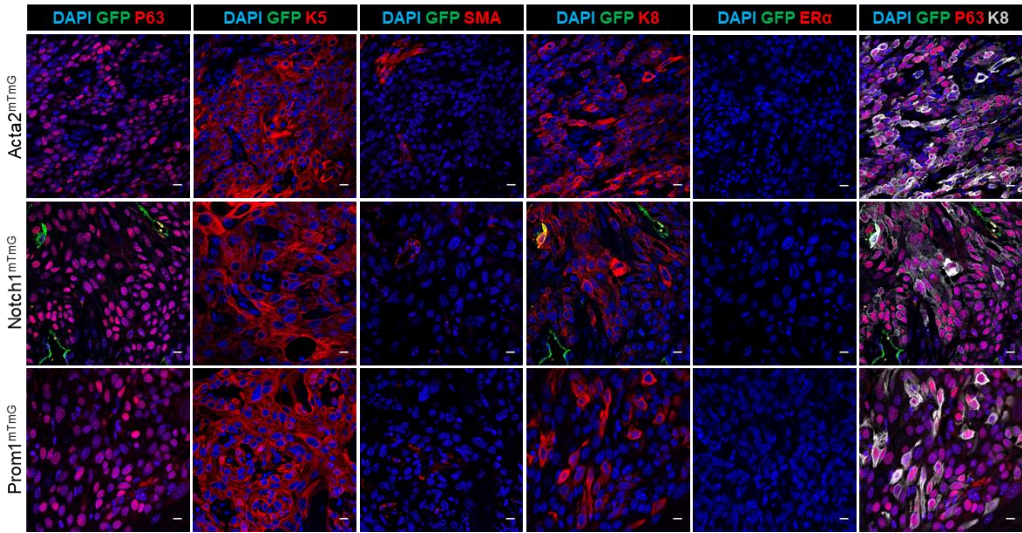


Figure R2.7. Investigating the cellular origins of hybrid triple-negative tumors in the DMBA-induced model. Representative IF images of DMBA-Type III tumors from Acta2^{mTmG}, Notch1^{mTmG}, and Prom1^{mTmG} mice. GFP is depicted in green, while basal markers (P63, K5, and SMA), and luminal markers (K8 and ERα) are shown in red. The final row presents merged images of P63 (red) and K8 (gray) staining. DNA is counterstained with DAPI in blue. Scale bars represent 10 μm.

Although inconclusive, these observations underscore the phenotypic plasticity observed in DMBA models and suggest that hybrid tumors may arise from untraced progenitors or through transdifferentiation events. Further research is essential to unravel the precise mechanisms driving the development of this intriguing tumor subtype.

2.4. Distinct lineage origins and variable latencies across DMBA-induced tumor subtypes

To gain deeper insights into how cellular origins shape the development of diverse breast cancer subtypes, we combined lineage tracing with latency analyses to examine the developmental trajectories and temporal dynamics of tumors arising from distinct progenitor cells. Lineage tracing of different MEC types revealed that ERα^{neg} luminal DMBA-Type I tumors consistently arose from *Notch1*-positive cells (ERα^{neg} LCs), while basal DMBA-Type II tumors exclusively derived from *Acta2*-positive BaCs, strengthening the evidence for their lineage-specific origins (Figure R2.8A).

RESULTS: OBJECTIVE 2

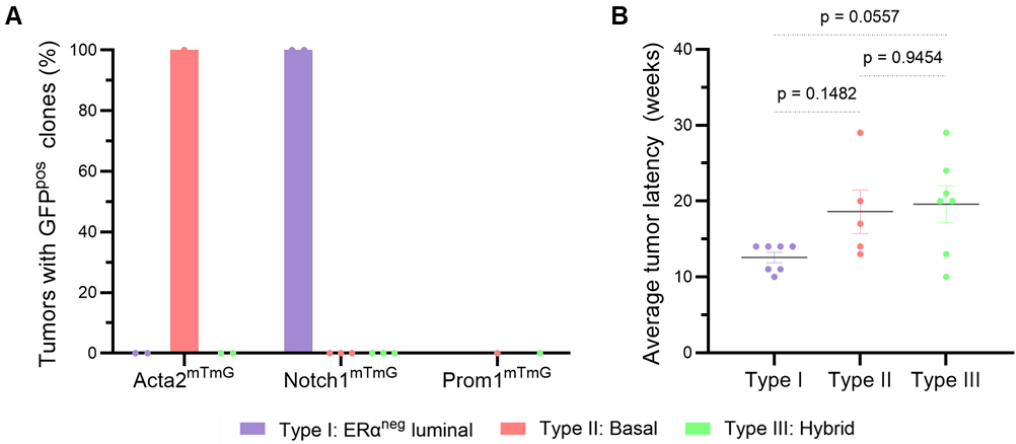


Figure R2.8. Summary of cellular origins and tumor latency in DMBA-induced tumors. **A.** Quantification of the percentage of DMBA-induced tumors containing GFP^{pos} tumor clones in Acta2^{mTmG}, Notch1^{mTmG} and Prom1^{mTmG} mice across different types: DMBA-Type I (luminal) (n= 6, 4, 0), DMBA-Type II (basal) (n= 1, 3, 1), and DMBA-Type III (hybrid) (n= 2, 3, 1). Each dot represents the average percentage of tumors with GFP^{pos} clones per animal. **B.** Distribution of latencies across DMBA-Type I, DMBA-Type II, and DMBA-Type III tumors (n=7, 5, and 7). “n” refers to the total number of tumors analyzed. All data are presented as mean ± S.E.M. Significance was determined using the one-way ANOVA test with p < 0.05 considered significant.

Latencies, reflecting the time required for tumors to emerge after carcinogen exposure, provided additional insights into subtype-specific differences. Luminal DMBA-Type I tumors exhibited shorter latencies (12.57 ± 0.69 weeks), consistent with their higher proliferative potential and activation of pathways that drive rapid tumor initiation. In contrast, basal (Type II) and hybrid (Type III) tumors developed more slowly and with greater variability (18.60 ± 2.87 weeks and 19.57 ± 2.42 weeks, respectively), suggesting distinct molecular or biological constraints shaping their progression. Notably, the similarity in latency between hybrid and basal tumors hints at shared mechanisms, possibly implicating a common cellular BaC origin (**Figure R2.8B**).

These findings underscore the divergent trajectories of TNBC subtypes, from their cellular beginnings to the dynamics of tumor growth. By linking tumor latency with cellular origin, this work illuminates how intrinsic properties of progenitor cells and their responses to carcinogenic insults contribute to the heterogeneity of breast cancer. Further exploration of hybrid tumor origins and the molecular drivers of subtype-specific latencies is essential to unravel the complexity of TNBC pathogenesis.

2.5. Oncogenic plasticity drives heterogeneous TNBC subtypes in the DMBA model

MECs exhibit remarkable plasticity under oncogenic stress, which can blur their lineage identities and contribute to intra-tumoral heterogeneity [91-94]. This section explores how cellular plasticity influences tumor outcomes, challenging the conventional paradigm of a fixed cell of origin. Indeed, the DMBA model provided compelling evidence of this plasticity.

RESULTS: OBJECTIVE 2

Tumors induced in Acta2^{mTmG} mice, which were expected to primarily reflect basal origins, exhibited unexpected heterogeneity. Some clones consisted of K8-positive LCs that had lost P63 expression (marked with an asterisk), indicating a complete basal-to-luminal transition. Others co-expressed both basal (P63) and luminal (K8) markers (marked with an arrow), suggesting an intermediate hybrid state during this transition (**Figure R2.9**). These findings suggest that DMBA-Type III tumors, characterized by a hybrid phenotype, may arise from BaCs undergoing lineage plasticity. However, further investigation is required to confirm this hypothesis and to determine the molecular drivers underlying this phenotypic shift.

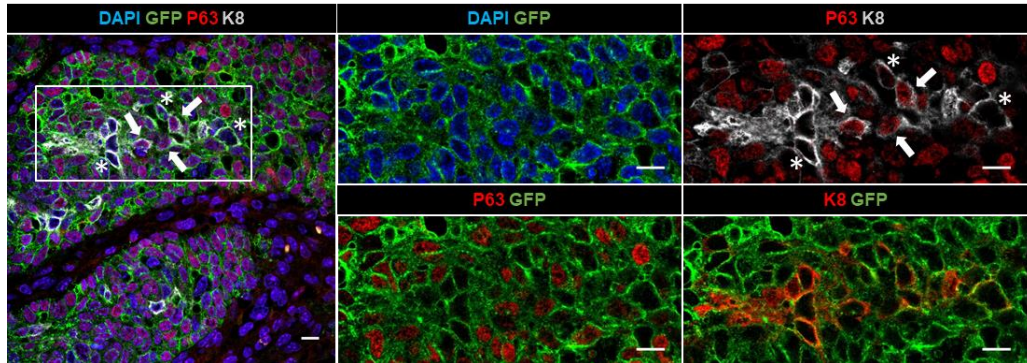


Figure R2.9. Basal-to-luminal transitions contribute to heterogeneity in DMBA-induced tumors. Representative IF images of GFP^{POS} clones in a DMBA Type I (basal) tumor from Acta2^{mTmG} mice. Merged images highlight the basal marker P63 in red and the luminal marker K8 in gray, with insets providing magnified views of framed luminal clones. GFP is shown in green, and DNA is counterstained with DAPI in blue. White arrows indicate cells co-expressing luminal and basal markers, while white asterisks highlight cells with a fully luminal identity. Scale bars: 10 μ m for both the main images and insets.

Notably, LCs exhibited their own intriguing behaviors under oncogenic pressure. In the DMBA model, *Notch1*-positive LCs consistently gave rise to K8-positive luminal clones, even within mixed tumors that appeared to result from the fusion of basal (Type II) and luminal (Type I) tumor clones (**Figure R2.10A**). This observation, aligned with our previous revelations in the MMTV-PyMT and C3(1)-TAg models, reinforces the idea that luminal progenitors are predominantly unipotent, retaining their identity even in the face of significant oncogenic challenges.

Nevertheless, exceptions to the presumed unipotency of *Notch1*-positive cells were observed. In one of the eleven tumors analyzed, *Notch1*-positive luminal tumors exhibited a scattered SMA expression pattern (**Figure R2.10B**). The presence of SMA, in the absence of other basal markers, suggests that this transition is more indicative of EMT rather than a direct shift to a basal identity. Indeed, SMA has been previously recognized as an EMT marker associated with increased invasiveness and metastatic potential [288]. This observation raises the possibility that, under certain oncogenic pressures, *Notch1*-positive LCs may undergo partial EMT, enhancing their capacity for tumor progression and dissemination. Further research is required to elucidate the molecular mechanisms underlying this phenotypic shift and its broader implications for TNBC progression.

RESULTS: OBJECTIVE 2

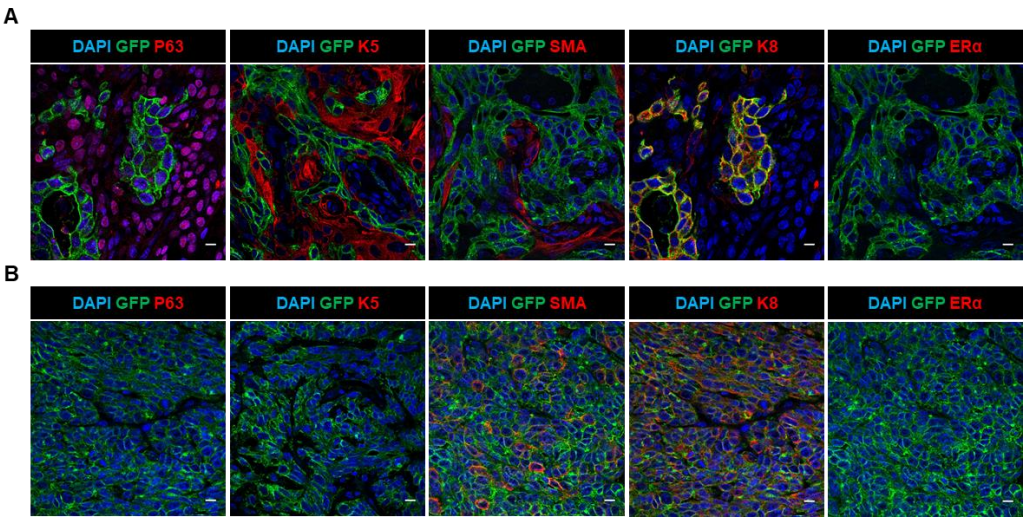


Figure R2.10. ERα-negative luminal progenitors largely preserve unipotency, yet may occasionally transition to an EMT-like state under oncogenic stress. Representative IF images of Notch1^{mTmG} tumors in the DMBA model are shown. **A.** Mixed tumor consisting of fused luminal (DMBA Type I) and basal (DMBA Type II) clones, where only the luminal portion consists of GFP^{pos} Notch1-derived cells. **B.** Tumor composed entirely of GFP^{pos} clones, positive for both K8 and SMA. GFP is depicted in green, while basal markers (P63, K5, and SMA), and luminal markers (K8 and ERα) are shown in red. DNA is counterstained with DAPI in blue. Scale bars represent 10 μm.

Together, these findings highlight the plastic nature of *Acta2*- and *Notch1*-positive cells during tumorigenesis. **While the initial cellular origins of distinct TNBC subtypes can be traced to specific MEC populations, the final tumor identity is shaped by the interplay of oncogenic events and cellular plasticity.** This dynamic process not only contributes to the heterogeneity observed within tumors but also underscores the complexity of breast cancer pathogenesis. Understanding these plasticity-driven shifts is crucial for developing targeted therapies capable of addressing the multifaceted nature of TNBC.

Objective 3: To investigate the clonal dynamics of distinct mammary epithelial cell types across transcriptomically diverse triple-negative breast cancer subtypes.

3. Luminal progenitors drive tumor evolution in the luminal TNBC models

The development of TNBC involves dynamic changes in MEC populations as tumors progress. To investigate how different MEC populations contribute to clonal expansion in histological luminal TNBC subtypes, we used the MMTV-PyMT and C3(1)-TAG models. By leveraging tamoxifen-inducible Cre lines, we aimed to trace the fate of specific MEC populations and uncover their roles at distinct tumor stages, from early hyperplasia to advanced carcinoma.

3.1. Distribution of MEC populations across tumor stages in the MMTV-PyMT model

To assess the persistence of MEC populations throughout tumor progression, we analyzed tumors at three stages: hyperplasia (4 weeks), adenoma (9 weeks), and carcinoma (14 weeks) (**Figure R3.1A**).

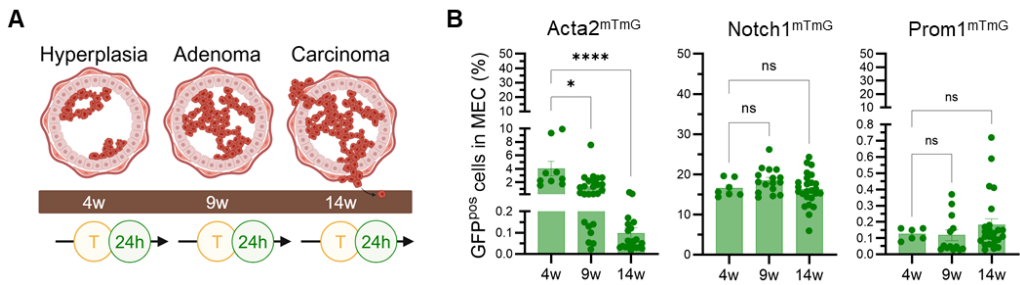


Figure R3.1. Distribution of mammary epithelial populations across tumor stages in the MMTV-PyMT model. **A.** Experimental design followed in 24-hour experiments in the MMTV-PyMT TNBC model. Tumors were analyzed 24 hours after tamoxifen injection at the hyperplasia (4w), adenoma (9w), or carcinoma (14w) stages. **B.** Percentage of GFP^{pos} mammary epithelial cells in MMTV-PyMT tumors from Acta2^{mTmG}, Notch1^{mTmG} and Prom1^{mTmG} mice following a 24-hour pulse at hyperplasia (n=9, 7, 6), adenoma (n=25, 16, 13), and carcinoma (n=21, 25, 24), obtained by FC. At least three animals per time point and genotype were analyzed. "n" refers to the total number of mammary glands analyzed. Data are presented as mean ± S.E.M. Significance was determined using the Kruskal-Wallis test with $p < 0.05$ considered significant.

We quantified GFP^{pos} cells within MECs 24 hours post-recombination using FC. Results revealed that the proportion of GFP^{pos} LCs labeled by *Notch1* and *Prom1* remained consistent across all stages, indicating a sustained contribution of luminal lineages. In contrast, GFP^{pos} BaCs labeled by *Acta2* significantly declined in advanced tumors, consistent with the gradual loss of the basal layer during tumor progression (**Figure R3.1B**). Specifically, at the hyperplasia stage, $4.04 \pm 1.07\%$ of cells in Acta2^{mTmG} tumors were GFP^{pos}, but this proportion decreased to $1.33 \pm 0.33\%$ at the adenoma stage and nearly disappeared ($0.10 \pm 0.03\%$) by the carcinoma stage. Meanwhile, the percentage of GFP^{pos} cells in Notch1^{mTmG} tumors remained stable at $16.56 \pm 0.81\%$ (hyperplasia), $18.45 \pm 0.79\%$ (adenoma), and $16.30 \pm 0.82\%$ (carcinoma). Similarly, Prom1^{mTmG}

RESULTS: OBJECTIVE 3

tumors maintained a low but consistent contribution of GFP^{pos} cells: $0.12 \pm 0.01\%$ (hyperplasia), $0.12 \pm 0.03\%$ (adenoma), and $0.18 \pm 0.04\%$ (carcinoma) (**Figure R3.1B**).

To further explore the lineage-specific contributions, we examined the distribution of distinct MEC subtypes (BaCs and LCs) within the GFP-positive population at each tumor stage. FC analysis showed that *Acta2*-positive cells in *Acta2*^{mTmG} tumors were exclusively basal, with $99.75 \pm 0.12\%$ and $99.89 \pm 0.07\%$ of GFP^{pos} cells identified as BaCs at the hyperplasia and adenoma stages, respectively. *Notch1*-positive cells in *Notch1*^{mTmG} tumors were confined to the luminal lineage, comprising $99.89 \pm 0.05\%$ of GFP^{pos} cells at hyperplasia and $99.13 \pm 0.32\%$ at adenoma. Similarly, *Prom1*-positive cells in *Prom1*^{mTmG} tumors were almost exclusively luminal, with $99.92 \pm 0.04\%$ of GFP^{pos} cells at hyperplasia and $99.89 \pm 0.11\%$ at adenoma (**Figure R3.2A**).

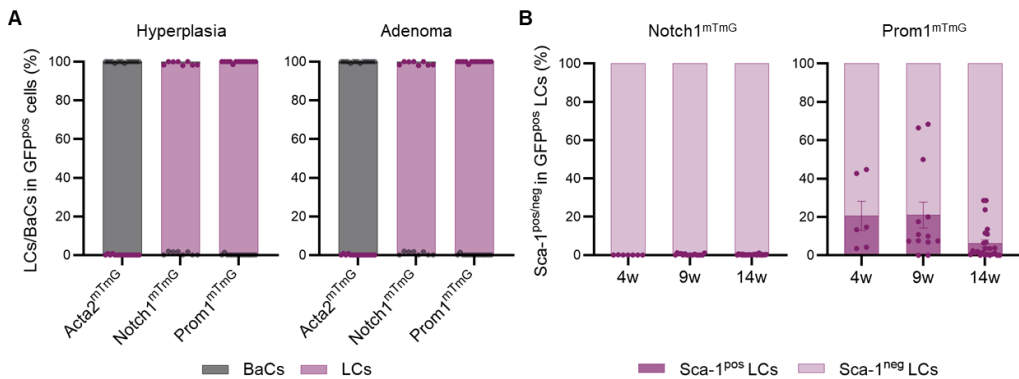


Figure R3.2. Flow cytometry-based cell fate analysis of mammary epithelial populations across tumor stages in the MMTV-PyMT model. **A.** Distribution of LCs and BaCs within GFP^{pos} MECs in *Acta2*^{mTmG}, *Notch1*^{mTmG}, and *Prom1*^{mTmG} after 24-hour pulses at hyperplasia (n=4, 7, 6) and adenoma (n=15, 8, 13) stages. **B.** Sca-1 expression in GFP^{pos} LCs after 24-hour pulses at hyperplasia (4w), adenoma (9w), and carcinoma (14w) in *Notch1*^{mTmG} (n=7, 16, 25) and *Prom1*^{mTmG} (n=6, 13, 24) tumors. At least three animals per time point and genotype were analyzed. "n" refers to the total number of mammary glands analyzed. Data are presented as mean $\pm$ S.E.M.

We also analyzed Sca-1 expression in GFP^{pos} LCs in *Notch1*^{mTmG} and *Prom1*^{mTmG} tumors to distinguish ER α -positive and ER α -negative LC subsets. In *Notch1*^{mTmG}, recombination exclusively targeted Sca-1^{neg} ER α -negative LCs, which persisted across all tumor stages ($99.97 \pm 0.02\%$ at hyperplasia, $99.68 \pm 0.11\%$ at adenoma, and $99.75 \pm 0.05\%$ at carcinoma). In contrast, in *Prom1*^{mTmG}, recombination initially labeled a subset of Sca-1^{pos} ER α -positive LCs ($20.57 \pm 7.57\%$ at hyperplasia and $21.00 \pm 6.71\%$ at adenoma), but this population declined to $6.31 \pm 1.82\%$ by the carcinoma stage, leaving most GFP^{pos} LCs in *Prom1*^{mTmG} tumors as Sca-1^{neg} LCs (**Figure R3.2B**). These findings suggest that, unlike in the healthy mammary gland, *Prom1* expression in tumors is not restricted to ER α -positive LCs but also marks an ER α -negative luminal subpopulation with stem-like traits, as previously reported [289-293].

To validate these findings, we performed IF analysis on hyperplastic tumor sections, staining for lineage-specific markers. *Acta2*-positive cells corresponded to P63-positive BaCs, while *Notch1*-positive cells were K8-positive and confined to the ER α -negative

RESULTS: OBJECTIVE 3

luminal lineage. *Prom1* expression labeled K8-positive LCs, with a small proportion co-expressing ER α (**Figure R3.3**), confirming the results obtained by FC.

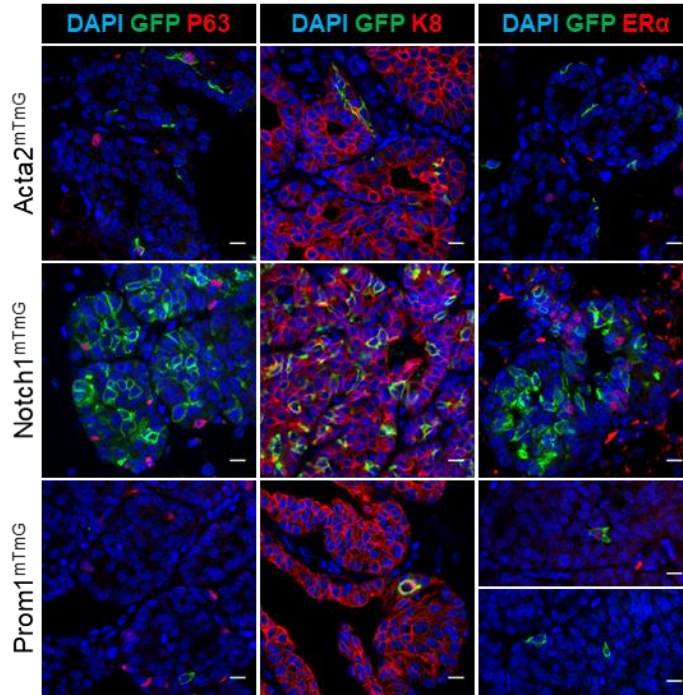


Figure R3.3. Immunofluorescence-based cell fate analysis of mammary epithelial populations in MMTV-PyMT hyperplastic lesions 24 hours post-tamoxifen injection. GFP^{pos} cells label *Acta2*-expressing BaCs (K8^{neg}, ER α ^{neg}, P63^{pos}) in the *Acta2*^{mTmG} line, *Notch1*-expressing ER α ^{neg} LCs (K8^{pos}, ER α ^{neg}, P63^{neg}) in the *Notch1*^{mTmG} line, and *Prom1*-expressing ER α ^{pos} (upper panel) or ER α ^{neg} (bottom panel) LCs (K8^{pos}, ER α ^{pos/neg}, P63^{neg}) in the *Prom1*^{mTmG} line. GFP is depicted in green, while the basal marker P63, and the luminal markers K8 and ER α are shown in red. DNA is counterstained with DAPI in blue. Scale bars represent 10 μ m.

Together, these results highlight the differential contributions of MEC populations to the luminal B subtype of TNBC in the MMTV-PyMT model. Luminal progenitors, marked by *Notch1* and *Prom1* expression, persist throughout tumor progression and likely play a central role in driving tumor growth. In contrast, the basal population, labeled by *Acta2*, diminishes as tumors progress, reflecting its limited role in advanced stages. These findings underscore the importance of LCs in tumor evolution and provide insight into the cellular dynamics underlying TNBC heterogeneity.

3.2. *Notch1*-expressing cells drive clonal expansion in early tumor stages in the MMTV-PyMT model

To dissect how distinct MEC lineages contribute to clonal expansion during breast tumor progression, we targeted specific MEC populations at the hyperplasia stage in the MMTV-PyMT model. Tamoxifen was administered to triple transgenic females, and their mammary glands were analyzed after either 5- or 10-week chase periods, corresponding to the adenoma and carcinoma stages, respectively (**Figure R3.4A**).

RESULTS: OBJECTIVE 3

This approach enabled us to trace the clonal dynamics of *Acta2*-, *Notch1*-, and *Prom1*-expressing populations, revealing lineage-specific contributions to tumor development.

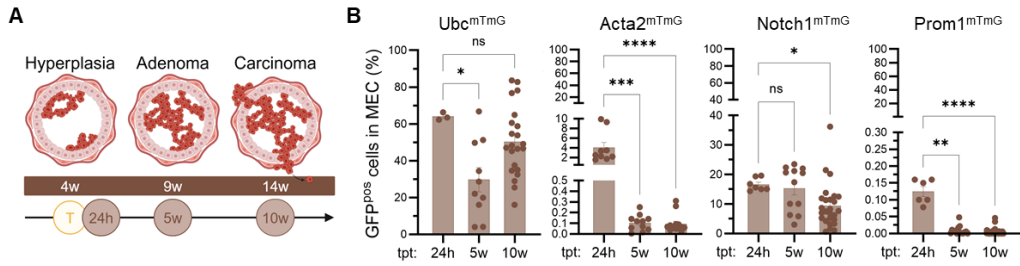


Figure R3.4. Early-stage targeting of MEC types reveals sustained contribution of *Notch1*-positive ER α -negative luminal progenitors to MMTV-PyMT tumor progression. **A.** Experimental design for clonal expansion analysis in the MMTV-PyMT model from hyperplasia. Tumors were analyzed 24 hours after tamoxifen injection at the hyperplasia stage and following 5- and 10-week chase periods. **B.** Percentage of GFP^{pos} mammary epithelial cells in MMTV-PyMT tumors from Ubc^{mtmG}, Acta2^{mtmG}, Notch1^{mtmG}, and Prom1^{mtmG} mice after a 24-hour pulse (24h) at hyperplasia (n=3, 9, 7, 6), following a 5-week chase (5w) (n=10, 11, 12, 15), or a 10-week chase (10w) (n=22, 13, 26, 41), obtained by FC. At least three animals per time point and genotype were analyzed. “n” refers to the total number of mammary glands analyzed. Data are presented as mean $\pm$ S.E.M. Significance was determined using the Kruskal-Wallis test with $p < 0.05$ considered significant.

Initial FC analysis uncovered striking differences in the clonal expansion potential of the different MEC populations, providing quantitative insights into these lineage dynamics. After 24 hours of tamoxifen administration at hyperplasia, GFP^{pos} cells accounted for $16.56 \pm 0.81\%$ of the MEC population in Notch1^{mtmG} tumors. This proportion remained stable after a 5-week chase ($15.26 \pm 2.19\%$) and declined only slightly after 10 weeks ($9.37 \pm 1.49\%$) (**Figure R3.4B**). These results indicate that *Notch1*-derived progeny persists and expands throughout tumor progression.

In contrast, the Acta2^{mtmG} and Prom1^{mtmG} lines exhibited sharp declines in GFP^{pos} cells following the chase periods. At 24 hours, GFP^{pos} cells made up $4.04 \pm 1.07\%$ and $0.12 \pm 0.01\%$ of the MEC population in Acta2^{mtmG} and Prom1^{mtmG} tumors, respectively. However, by the 5-week chase, these proportions fell to $0.10 \pm 0.02\%$ (Acta2^{mtmG}) and $0.01 \pm 0.00\%$ (Prom1^{mtmG}). By 10 weeks, no detectable GFP^{pos} cells remained in Prom1^{mtmG} tumors ($0.00 \pm 0.00\%$), and only $0.09 \pm 0.02\%$ persisted in Acta2^{mtmG} tumors (**Figure R3.4B**). These findings suggest that *Prom1*- and *Acta2*-expressing populations play a minimal role in early tumor development.

Interestingly, the Ubc^{mtmG} line, which targets all MEC subpopulations, exhibited a marked decline in GFP^{pos} cells at the adenoma stage, dropping from $64.03 \pm 1.19\%$ at 24 hours to $29.72 \pm 6.60\%$ after 5 weeks. This reduction likely reflects the lower proliferative potential of *Notch1*-negative luminal cells. However, by the carcinoma stage, the proportion of GFP^{pos} cells in the Ubc^{mtmG} line rebounded to $50.34 \pm 3.99\%$ (**Figure R3.4B**), suggesting selective survival and expansion of highly proliferative subpopulations.

RESULTS: OBJECTIVE 3

IF analysis further showcased the striking differences in the clonal expansion potential of the different MEC populations. Large GFP^{pos} clones were observed exclusively in tumors derived from *Notch1*^{mTmG} mice. These clones were comparable in size to those seen in the control *Ubc*^{mTmG} line, suggesting that *Notch1*-positive cells play a prominent role in driving clonal expansion during early tumorigenesis. In contrast, *Prom1*^{mTmG} tumors exhibited rare and small GFP^{pos} clones, while *Acta2*-derived cells remained confined to the basal layer without evidence of clonal expansion. These findings underscore the limited proliferative potential of *Prom1*- and *Acta2*-expressing populations during early tumor progression compared to *Notch1*-positive cells (**Figure R3.5**).

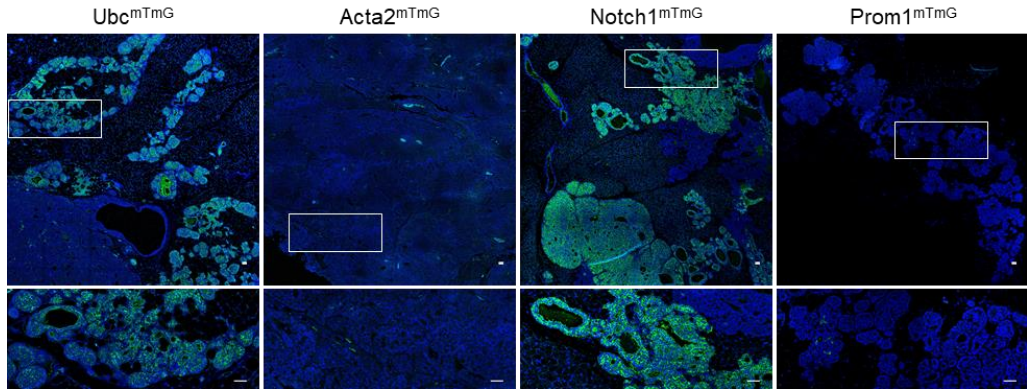


Figure R3.5. *Notch1*-positive cells targeted at hyperplasia stage exhibit clonal expansion during MMTV-PyMT tumor progression. Representative IF images from MMTV-PyMT tumor sections 10 weeks post-tamoxifen injection at the hyperplasia stage in the indicated Cre lines. GFP is depicted in green, and DNA is counterstained with DAPI in blue. Scale bars represent 40 μ m for both the main images and insets.

Moreover, cell fate characterization of GFP-positive clones in the *Notch1*^{mTmG} line revealed that *Notch1*-derived cells preserved their luminal identity during clonal expansion in the MMTV-PyMT model, exhibiting K8 positivity and ER negativity, while lacking expression of the basal marker P63, thereby highlighting their unipotent potential in luminal B tumorigenesis (**Figure R3.6**).

These results unequivocally identify *Notch1*-positive cells as the primary drivers of early breast tumorigenesis in the MMTV-PyMT model. Their sustained contribution to clonal expansion highlights their dominant role in tumor progression. In contrast, *Prom1*- and *Acta2*-expressing populations exhibit limited clonal expansion potential, suggesting that their involvement in tumor progression is restricted. The observed dynamics underscore the hierarchical contributions of distinct MEC lineages to tumor evolution, with *Notch1*-positive cells emerging as a critical target population in luminal B-like TNBC tumors.

RESULTS: OBJECTIVE 3

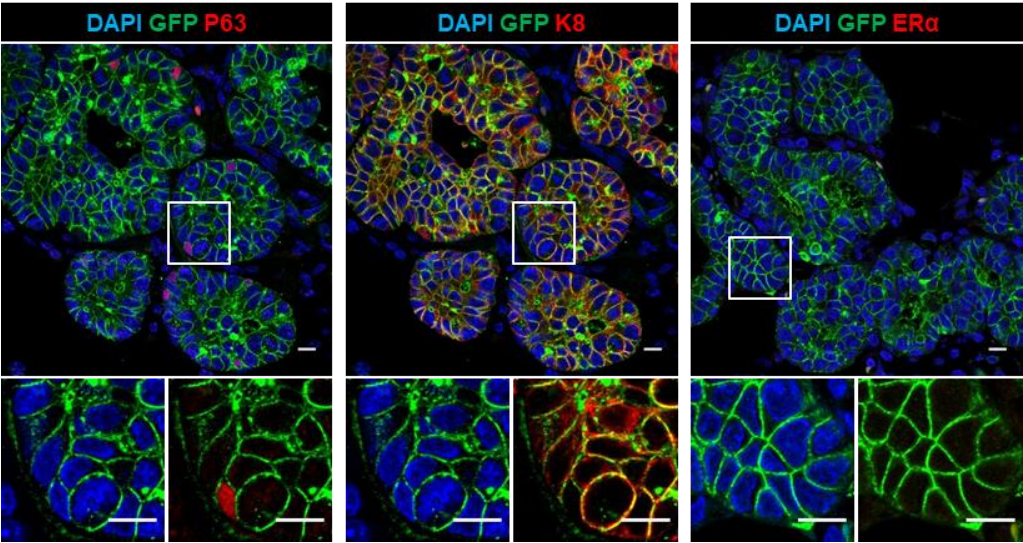


Figure R3.6. *Notch1*-derived cells retain luminal identity during clonal expansion in the MMTV-PyMT model. Representative IF images from MMTV-PyMT tumor sections 10 weeks post-tamoxifen injection at the hyperplasia stage in the *Notch1*^{mTmG} line. GFP is depicted in green, while the basal marker P63, and the luminal markers K8 and ERα are shown in red. Insets provide magnified views of the highlighted regions. DNA is counterstained with DAPI in blue. Scale bars represent 10 μm for both the main images and insets.

3.3. Differential clonal dynamics of luminal progenitors during tumor progression

To investigate how clonal dynamics of distinct MEC populations vary during advanced tumor progression, we analyzed their proliferative potential when targeted at later stages of tumorigenesis. MECs were labeled at the adenoma stage (9 weeks), and their clonal expansion was assessed after a 5-week chase at the carcinoma stage (**Figure R3.7A**).

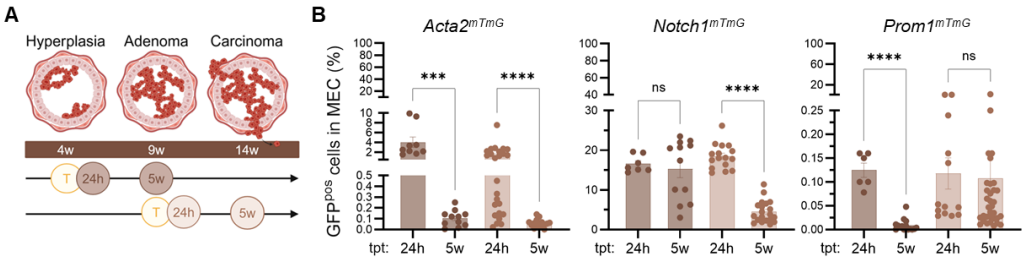


Figure R3.7. Clonal expansion of *Notch1*- and *Prom1*-positive luminal populations varies with tumorigenic stage in the MMTV-PyMT model. **A.** Experimental design for clonal expansion analysis in the MMTV-PyMT model, with tumors analyzed 24 hours post-tamoxifen at hyperplasia (dark brown) or adenoma (light brown) stages, followed by a 5-week chase period. **B.** Percentage of GFP^{pos} MECs in MMTV-PyMT tumors from *Acta2*^{mTmG}, *Notch1*^{mTmG}, and *Prom1*^{mTmG} mice after a 24-hour pulse (24h) at hyperplasia (n=9, 7, 6), following a 5-week chase (5w) from hyperplasia (n=11, 12, 15), after a 24-hour pulse at adenoma (n=25, 16, 13), and following a 5-week chase from adenoma (n=16, 20, 33), obtained by FC. At least three animals per time point and genotype were analyzed. “n” refers to the total number of mammary glands analyzed. Data are presented as mean ± S.E.M. Significance was determined using the Kruskal-Wallis test with p < 0.05 considered significant.

RESULTS: OBJECTIVE 3

Our analyses revealed that *Notch1*- and *Prom1*-expressing cells exhibited distinct behaviors influenced by tumor stage, reflecting their unique roles in tumor progression. *Notch1*-positive cells demonstrated robust clonal expansion at earlier stages but exhibited restricted behavior as tumors advanced. When targeted at the adenoma stage, *Notch1*-positive cells initially accounted for $18.45 \pm 0.79\%$ of GFP^{pos} cells; however, this proportion significantly declined to $4.50 \pm 0.61\%$ after the 5-week chase. This reduction suggests that while some *Notch1*-positive cells continued to expand, a substantial fraction experienced cell death, reduced growth, or entered a non-dividing state (**Figure R3.7B**).

In contrast, *Prom1*-positive cells showed a delayed but consistent clonal response. Following a 24-hour pulse at the adenoma stage, GFP^{pos} cells comprised $0.12 \pm 0.03\%$ of the total population. This proportion remained stable at $0.11 \pm 0.05\%$ after the 5-week chase, suggesting a steady contribution to tumor growth despite their relatively late activation compared to *Notch1*-positive cells (**Figure R3.7B**).

Acta2-positive cells, however, maintained a stable, non-proliferative phenotype throughout tumor progression. At the adenoma stage, GFP^{pos} cells accounted for $1.33 \pm 0.33\%$ after a 24-hour pulse, dropping further to $0.06 \pm 0.01\%$ after the 5-week chase (**Figure R3.7B**). These findings underscore the limited contribution of *Acta2*-positive BaCs to tumor progression, as they consistently remained confined to the healthy basal layer without evidence of clonal expansion.

To further characterize clonal growth patterns, we examined GFP^{pos} clone sizes following 5-week chase periods in tumors from *Notch1*^{mTmG} and *Prom1*^{mTmG} mice. Clones were categorized as single cells (1 cell), small clones (2–10 cells), or large clones (≥ 11 cells) (**Figure R3.8A**).

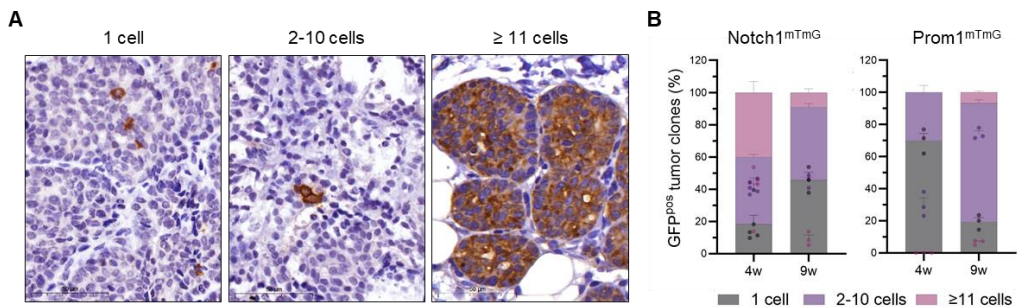


Figure R3.8. Tumor stage-dependent changes in the size of GFP-positive tumor clones in *Notch1*^{mTmG} and *Prom1*^{mTmG} MMTV-PyMT tumors. **A.** Representative IHC images of GFP^{pos} tumor clones categorized as single cells (1 cell), small clones (2–10 cells), or large clones (≥ 11 cells). Scale bars: 50 μ m. **B.** Quantification of GFP^{pos} tumor clones categorized by size as single cells (gray), small clones (purple), and large clones (pink) in *Notch1*^{mTmG} and *Prom1*^{mTmG} lines after a 5-week chase from hyperplasia (4w) ($n=5$, 2109 clones and $n=3$, 48 clones) and adenoma (9w) ($n=3$, 3708 clones and $n=3$, 434 clones). At least three animals per time point and genotype were analyzed. “n” refers to the total number of mammary glands analyzed. Data are presented as mean $\pm$ S.E.M.

Notch1-positive cells displayed a shift in clonal dynamics as tumors progressed, with large clones diminishing in prevalence at later stages. At 4 weeks, GFP^{pos} clones from

RESULTS: OBJECTIVE 3

Notch1^{mTmG} tumors included $18.41 \pm 5.47\%$ single cells, $41.46 \pm 1.81\%$ small clones, and $40.13 \pm 6.82\%$ large clones. By 9 weeks, the percentage of large clones sharply declined to $9.15 \pm 2.40\%$, while single cells increased to $45.85 \pm 4.62\%$ and small clones remained relatively stable at $45.01 \pm 2.28\%$ (**Figure R3.8B**). This pattern suggests a shift toward reduced expansion potential and increased cell attrition at advanced stages.

Conversely, *Prom1*-positive cells demonstrated an opposite trend, reflecting a delayed activation of clonal expansion. At 4 weeks, clones were predominantly single cells ($70.01 \pm 4.39\%$), with small clones comprising $29.92 \pm 4.39\%$ and no large clones observed. By 9 weeks, single cells decreased significantly to $19.32 \pm 2.65\%$, small clones rose to $74.08 \pm 1.96\%$, and large clones emerged at $6.61 \pm 0.86\%$ (**Figure R3.8B**), indicating a late-stage proliferative response.

These observations highlight the interplay of biological processes shaping clonal dynamics, including expansion, delayed growth, cell cycle arrest, and cell death, as previously proposed by other research groups (**Figure I9**) [10]. *Notch1*-positive cells dominate early-stage clonal expansion, while *Prom1*-positive cells activate their proliferative potential later. *Acta2*-positive cells, by contrast, remain confined to a basal, non-proliferative role throughout tumor progression.

Collectively, **these results emphasize the divergent contributions of MEC types to tumorigenesis, with distinct temporal and clonal behaviors that shape the heterogeneity observed in TNBC.**

3.4. Cell fate and transformation potential of mammary epithelial populations in the C3(1)-TAg model

To investigate how different MEC lineages contribute to basal-like tumors, we began by evaluating the fate and susceptibility of *Acta2*-, *Notch1*-, and *Prom1*-expressing cells to transformation in the C3(1)-TAg model. Through 24-hour pulse experiments at the hyperplasia stage (16 weeks), we used immunofluorescence to analyze cell fates and lineage-specific transformation.

When evaluating the expression of SV40-TAg oncogene, we found that all three epithelial lineages were susceptible to transformation, though with varying efficiencies (**Figure R3.9**). *Notch1*-positive cells exhibited the highest transformation rate, with 99.09% of 692 cells analyzed across 34 fields undergoing transformation. In contrast, transformation was significantly less frequent in *Acta2*-positive (0.89% of 445 cells analyzed across 30 fields) and *Prom1*-positive (7.22% of 376 cells analyzed across 35 fields) populations. This discrepancy suggests that *Acta2*- and *Prom1*-expressing populations may exhibit inherent resistance to transformation or undergo increased cell death following SV40-TAg expression.

Cell fate analysis confirmed that *Acta2* expression remained exclusively confined to the P63-positive BaCs, while *Notch1* and *Prom1* labeled distinct luminal populations (**Figure R3.10**). *Notch1* was predominantly expressed in ER α ^{neg} LCs (K8-positive), but also labeled a small subset of BaCs, with 2.15% of *Notch1*-positive cells (810 cells analyzed across 32 fields) expressing the P63 basal marker. In contrast, *Prom1*

RESULTS: OBJECTIVE 3

primarily targeted ERα^{neg} LCs, with only 23.92% of *Prom1*-positive cells expressing ERα (291 cells analyzed across 26 fields).

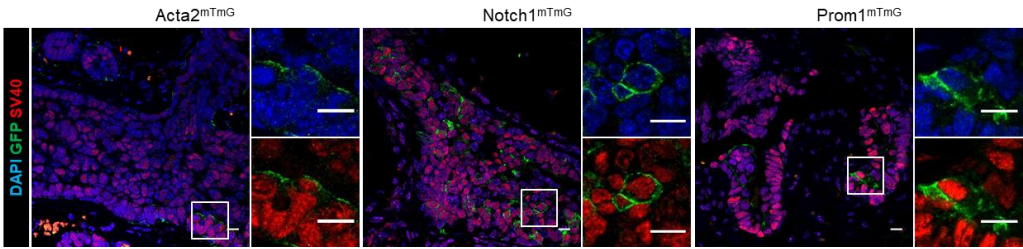


Figure R3.9. Differential susceptibility of mammary epithelial cells to SV40-TAg-driven transformation. IF images of C3(1)-TAG tumor sections 24 hours post-tamoxifen injection at hyperplasia stage (16 weeks). GFP is depicted in green and SV40 in red. Insets provide magnified views of the highlighted regions. DNA is counterstained with DAPI in blue. Scale bars represent 10 μm for both the main images and insets.

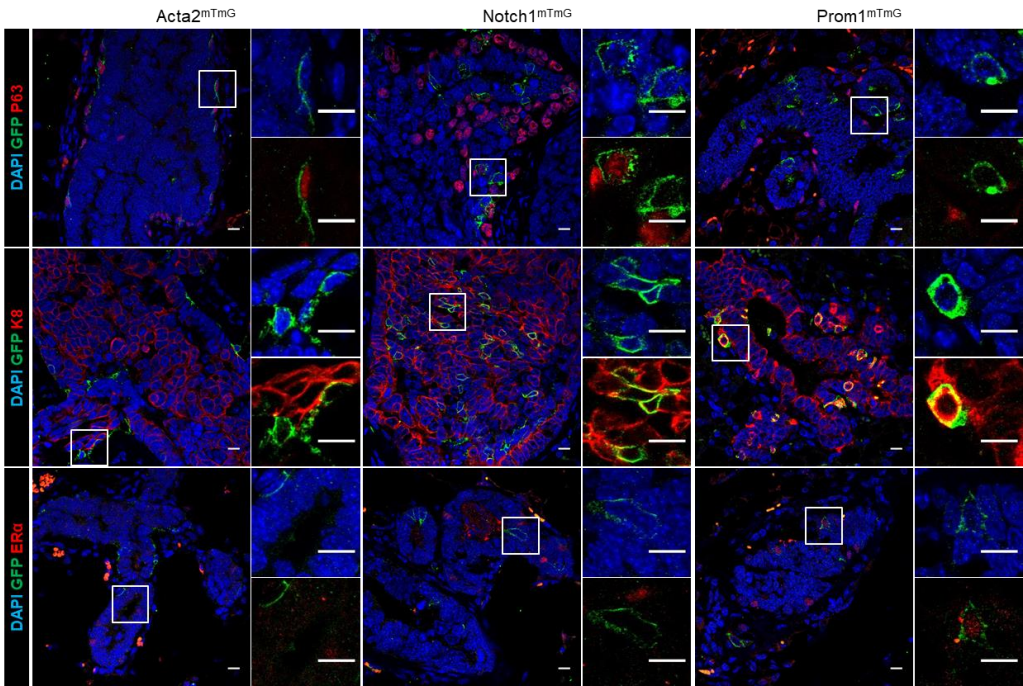


Figure R3.10. Immunofluorescence-based cell fate analysis of mammary epithelial populations in C3(1)-TAG hyperplastic lesions 24 hours post-tamoxifen injection. GFP is depicted in green, while the basal marker P63, and the luminal markers K8 and ERα are shown in red. Insets provide magnified views of the highlighted regions. DNA is counterstained with DAPI in blue. Scale bars represent 10 μm for both the main images and insets.

In summary, our findings reveal that distinct MECs exhibit differential transformation potential in the C3(1)-TAG model, with *Notch1*-positive luminal cells being the most susceptible to SV40-TAg-driven oncogenesis. Interestingly, a subset of *Notch1*-expressing cells was identified as BaCs, suggesting that *Notch1* marks a broader epithelial population under oncogenic pressure.

RESULTS: OBJECTIVE 3

3.5. Luminal progenitors drive basal-like tumor development in the C3(1)-TAg model

To assess how distinct MEC populations contribute to tumor growth, we performed clonal expansion experiments targeting *Acta2*-, *Notch1*-, and *Prom1*-positive cells at the hyperplasia stage, followed by 5- and 10-week chase periods (**Figure R3.10A**). This allowed us to evaluate clonal dynamics as tumors progressed from hyperplasia to carcinoma.

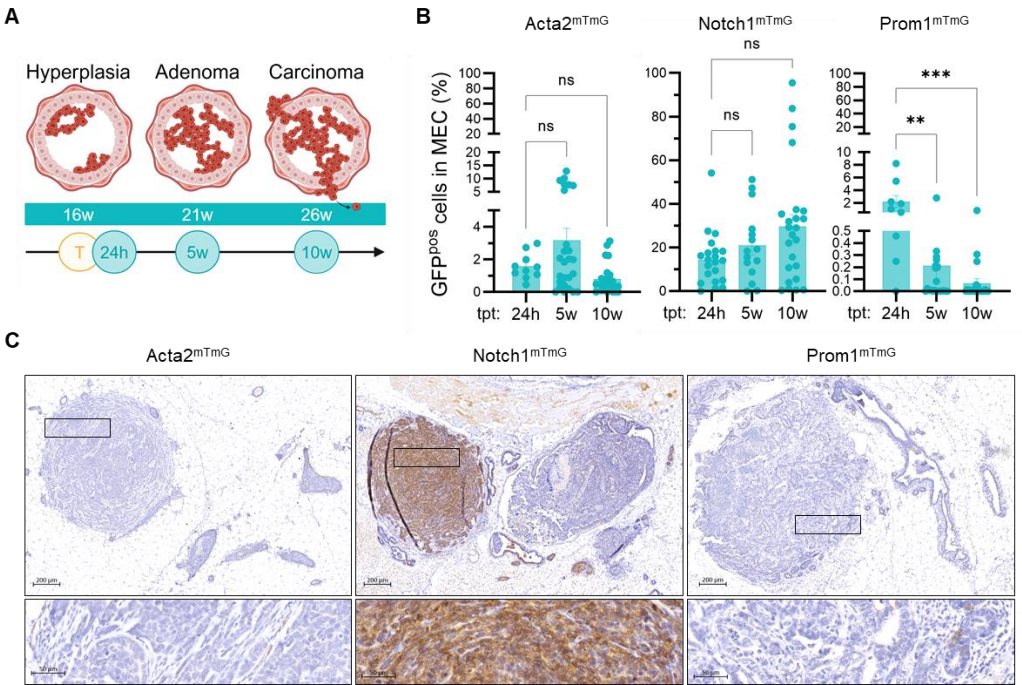


Figure R3.11. Early-stage targeting of MEC types reveals sustained contribution of *Notch1*-positive ER α -negative luminal progenitors to C3(1)-TAg tumor progression. **A.** Experimental design for clonal expansion analysis in the C3(1)-TAg model from hyperplasia. Tumors were analyzed 24 hours after tamoxifen injection at the hyperplasia stage and following 5- and 10-week chase periods. **B.** Percentage of GFP^{pos} MECs in C3(1)-TAg tumors from Acta2^{mTmG}, Notch1^{mTmG}, and Prom1^{mTmG} mice after a 24-hour pulse at the hyperplasia stage (n=10, 21, 9), after a 5-week chase (n=26, 15, 22), or a 10-week chase (n=24, 24, 23), based on IHC quantification. At least two animals per time point and genotype were analyzed. "n" refers to the total number of mammary glands analyzed. Data are presented as mean $\pm$ S.E.M. Significance was determined using the Kruskal-Wallis test with $p < 0.05$ considered significant. **C.** Representative IHC images of C3(1)-TAg tumor sections 10 weeks post-tamoxifen injection at the hyperplasia stage in the indicated Cre lines. Brown indicates positive GFP staining, and blue represents nuclear counterstaining. Scale bars represent 200 μ m in the main images and 50 μ m in the insets.

Immunohistochemical analysis revealed that *Notch1*-positive cells exhibited robust clonal expansion throughout tumor development, forming large GFP-positive clones. In contrast, *Prom1*-positive cells contributed minimally to tumor growth, with GFP^{pos} cells declining significantly over time. *Acta2*-positive cells remained restricted to the basal layer, showing no evidence of clonal expansion during tumorigenesis (**Figures R3.11B-C**).

RESULTS: OBJECTIVE 3

Quantitative analyses of GFP^{pos} tumor cells confirmed these observations. *Notch1*-derived progeny steadily increased their contribution, from $14.43 \pm 2.64\%$ at the 24-hour pluses to $20.91 \pm 4.22\%$ after 5 weeks and $29.62 \pm 5.49\%$ after 10-weeks. *Prom1*-derived progeny decline sharply, from $2.24 \pm 0.93\%$ at the 24-hour pulse to $0.21 \pm 0.13\%$ after 5 weeks and $0.07 \pm 0.04\%$ after 10 weeks, reflecting a reduced contribution to the tumor bulk as the tumor advances. *Acta2*-derived progeny remained stable over time, representing $1.57 \pm 0.25\%$ at the 24-hour pulse, $3.18 \pm 0.73\%$ after 5 weeks, and $0.79 \pm 0.19\%$ after 10 weeks (**Figure R3.11B**). These data indicate a more gradual loss of the healthy basal layer in this model, aligning with its significantly slower tumor progression compared to the MMTV-PyMT model.

Cell fate analysis of GFP-positive clones in the *Notch1*^{mTmG} line revealed that, while most tumors contained GFP-positive clones with a K8-positive luminal identity, 24.32% of the 37 tumors analyzed exhibited GFP-positive basal clones characterized by P63 expression (**Figure R3.12**). This finding suggests that a subset of *Notch1*-positive BaCs, targeted early in tumorigenesis, can clonally expand and contribute to tumor progression. However, these clones remained small and sparsely distributed within larger tumors, indicating that while they may play a role in early tumor development, their influence diminishes as the tumor advances.

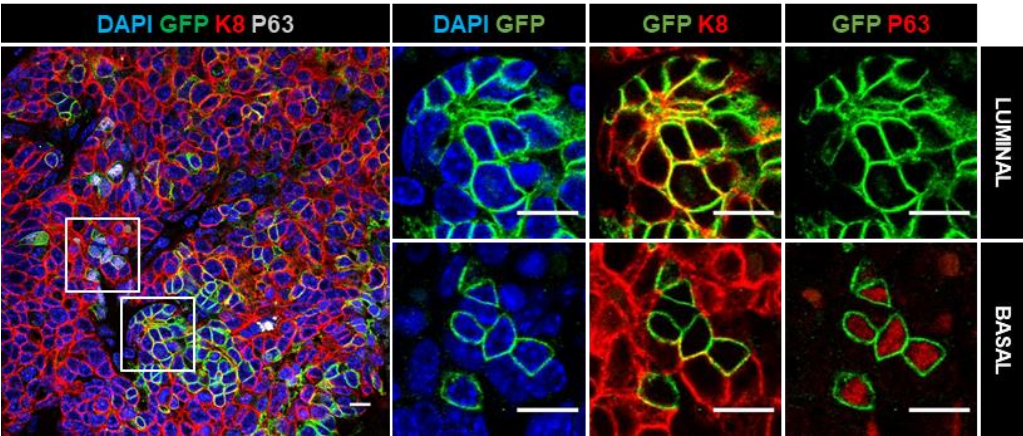


Figure R3.12. A specific subset of *Notch1*-positive basal cells undergoes clonal expansion during basal-like tumorigenesis in the C3(1)-TAg model. Representative IF images from C3(1)-TAg tumor sections 10 weeks post-tamoxifen injection at the hyperplasia stage in the *Notch1*^{mTmG} line. Merged images show GFP in green, the luminal marker K8 in red, and the basal marker P63 in gray. Insets provide magnified views of the highlighted regions. DNA is counterstained with DAPI in blue. Scale bars represent 10 μm for both the main images and insets.

Lastly, we investigated whether the clonal potential of *Notch1*-positive cells varies with tumor stage, by targeting them at the adenoma stage and analyzing them after a 5-week chase (**Figure R3.13A**). The quantification of GFP^{pos} cells showed stable proportions, with $11.13 \pm 2.29\%$ at the 24-hour pulse and $9.50 \pm 1.46\%$ after 5 weeks (**Figure R3.13B**).

RESULTS: OBJECTIVE 3

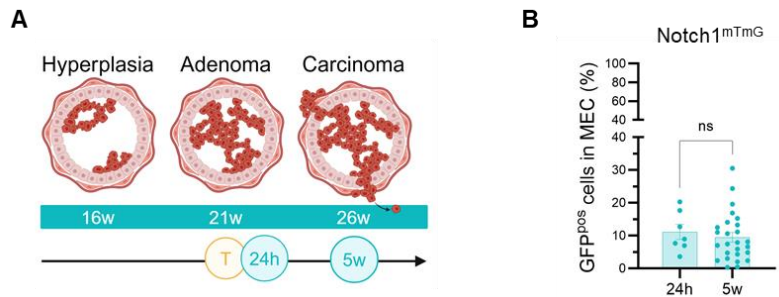


Figure R3.13. Late-stage targeting of *Notch1*-positive cells reveals continuous clonal growth. **A.** Experimental design for clonal expansion analysis in the C3(1)-TAg model from adenoma, with tumors analyzed 24 hours after tamoxifen injection at the adenoma stage and following a 5-week chase period. **B.** Percentage of GFP^{pos} MECs in C3(1)-TAg tumors from Notch1^{mTmG} mice after a 24-hour pulse (24h) at the adenoma stage (n=7), or after a 5-week chase (5w) (n=26), based on IHC quantification. At least two animals per time point and genotype were analyzed. “n” refers to the total number of mammary glands analyzed. Data are presented as mean ± S.E.M. Significance was determined using the Mann-Whitney test with p < 0.05 considered significant.

These findings underscore the dominant role of *Notch1*-positive LCs in driving basal-like TNBC, while *Prom1*- and *Acta2*-positive populations contribute minimally, mirroring observations in the MMTV-PyMT model. Further clonal expansion studies using additional Cre lines and extended time points could help determine whether other cell populations have the potential to outcompete *Notch1*-positive cells over time.

Objective 4: To determine the metastatic potential of different mammary epithelial cell populations in breast tumors.

4. *Notch1*-positive cells exhibit cellular plasticity during metastatic spread

4.1. Lung metastases reflect primary tumor heterogeneity in MEC composition

To investigate the metastatic potential of distinct MEC populations, we utilized the MMTV-PyMT model, renowned for its high incidence of lung metastases (80–90%) at the carcinoma stage. This model provided an ideal platform to trace the contribution of *Acta2*-, *Notch1*-, and *Prom1*-expressing MECs to tumor spread and metastasis.

As an initial step, we compared the cellular composition of metastatic lung nodules with that of primary tumors at the hyperplasia stage after similar pulses (**Figure R4.1A**).

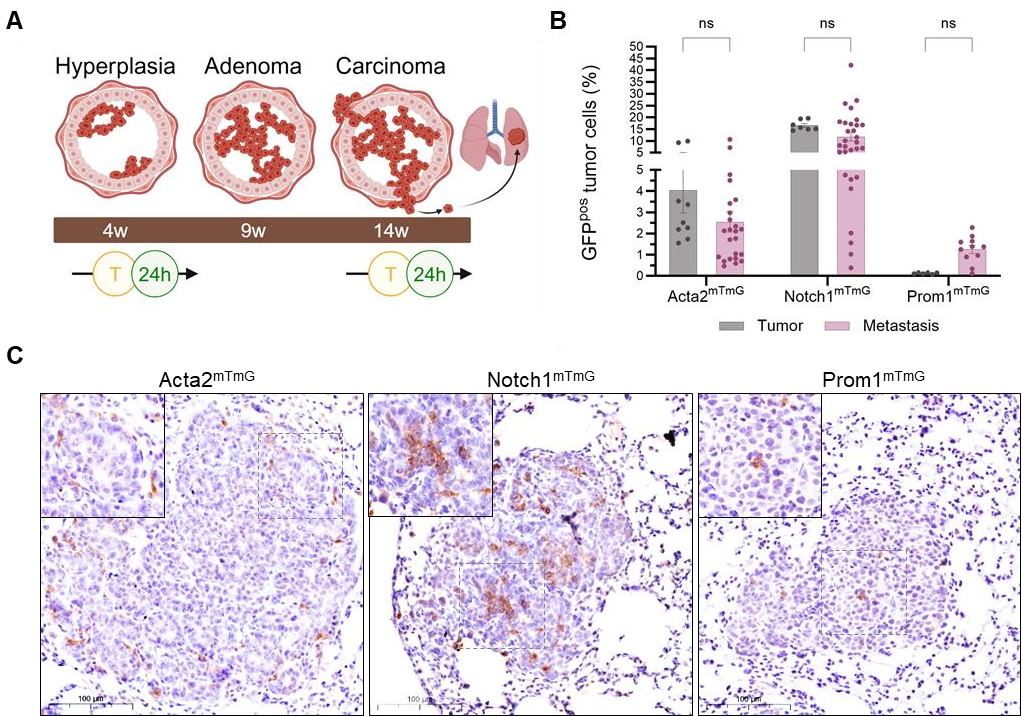


Figure R4.1. Lung metastases recapitulate the cellular architecture of the primary tumor. **A.** Experimental design in the MMTV-PyMT model to analyze MEC distribution in hyperplastic tumors and lung metastases following 24-hour tamoxifen pulses at the hyperplasia and carcinoma stages in *Acta2*^{mTmG}, *Notch1*^{mTmG}, and *Prom1*^{mTmG} mice. **B.** Percentage of GFP^{pos} tumor cells in lung metastases (pink) (n=24, 30, 11) and primary tumors (gray) (n= 9, 7, 6) after a 24-hour pulse for each Cre line. Each dot represents an individual tumor or metastasis. At least three animals per genotype and timepoint were analyzed. Data are presented as mean ± S.E.M. Significance was determined using the two-way ANOVA test, with p < 0.05 considered significant. **C.** Representative IHC images illustrate GFP^{pos} cells in lung metastases from indicated Cre lines after a 24-hour tamoxifen pulse at carcinoma stage. Brown indicates positive GFP staining, and blue represents nuclear counterstaining. Scales bars correspond to 100 µm. Enlarged views of the highlighted regions are displayed in the magnified panels.

RESULTS: OBJECTIVE 4

Remarkably, lung metastases retained three distinct MEC lineages present in primary tumors at the hyperplasia stage. GFP^{pos} cells derived from *Acta2*-, *Notch1*-, and *Prom1*-expressing MECs were identified in metastases, suggesting that metastatic cells attempt to replicate the primary tumor's cellular architecture within the lung environment (**Figures R4.1B-C**).

Quantitative analysis revealed consistent proportions of GFP^{pos} cells in both the primary tumors and metastases (**Figure R4.1B**). *Acta2*-positive cells accounted for $4.04 \pm 1.07\%$ of GFP-positive cells in primary tumors and $2.55 \pm 0.48\%$ in lung metastases. Similarly, *Notch1*-positive cells represented $16.56 \pm 0.81\%$ in primary tumors and $11.76 \pm 1.72\%$ in metastases. Notably, *Prom1*-positive cells showed a relative increase in metastases, rising from $0.13 \pm 0.02\%$ in primary tumors to $1.25 \pm 0.19\%$ in the lungs.

Further immunofluorescence analysis showed that *Acta2*-positive cells within metastatic nodules retained basal lineage characteristics, expressing markers such as P63, K14, SMA, and K5. In contrast, *Notch1*- and *Prom1*-positive cells exhibited K8 expression and were ER α -negative (**Figure R4.2**). These findings suggest that metastatic initiation and survival depend on preserving the basal-to-luminal ratio observed in primary tumors, highlighting the role of cellular heterogeneity in sustaining metastatic outgrowth.

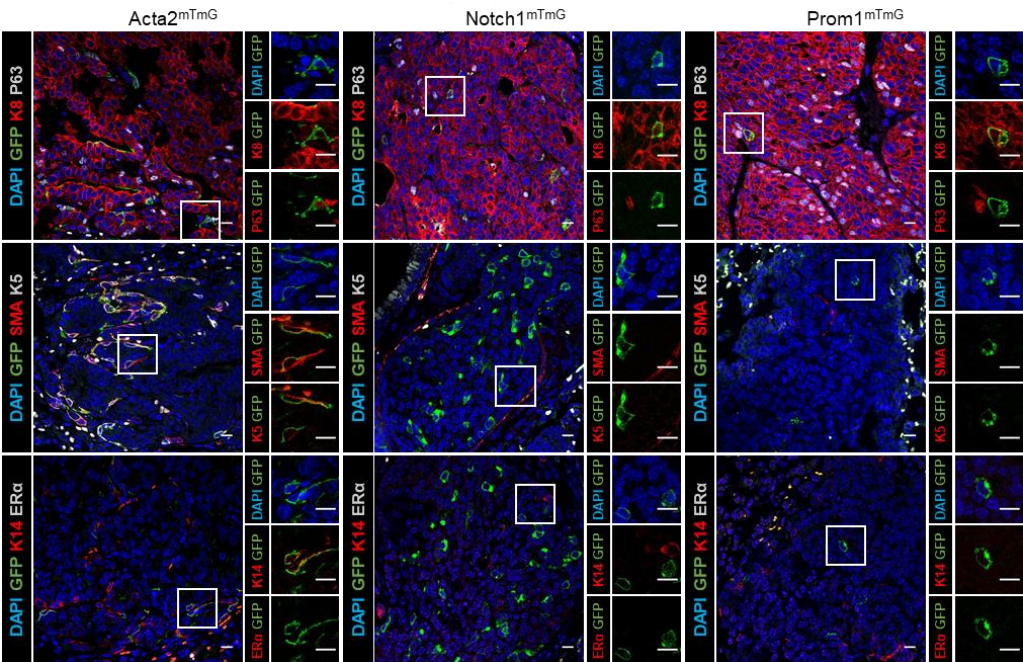


Figure R4.2. Mammary tumor cells retain their primary tumor lineage characteristics in metastatic sites. Representative IF images of lung metastases from *Acta2*^{mTmG}, *Notch1*^{mTmG}, and *Prom1*^{mTmG} mice after a 24-hour pulse at carcinoma stage. Merged images show GFP in green, with K8, SMA, or K14 in red, and P63, K5, or ER α in gray. Insets provide magnified views of the highlighted regions. DNA is counterstained with DAPI in blue. Enlarged views of the highlighted cells are shown in the magnified panels. Scale bars represent 10 μ m for both the main images and insets.

RESULTS: OBJECTIVE 4

Together, these observations emphasize the ability of tumor cells to adapt to distant tissue environments while preserving the lineage characteristics of their primary tumor. This cellular diversity highlights the complex biological mechanisms driving metastatic progression and the survival of heterogeneous tumor populations at secondary sites.

4.2. *Notch1*-positive cells are primary drivers of lung metastasis in the MMTV-PyMT model

To examine the metastatic potential of distinct MEC populations, we administered a single intraperitoneal injection of tamoxifen at either hyperplasia or adenoma stage in the different Cre mice and dissected the lungs at carcinoma stage, 10 or 5 weeks later, respectively (**Figure R4.3A**). This experimental approach provided insights into how tumor progression influences the metastatic capacity of distinct MEC populations.

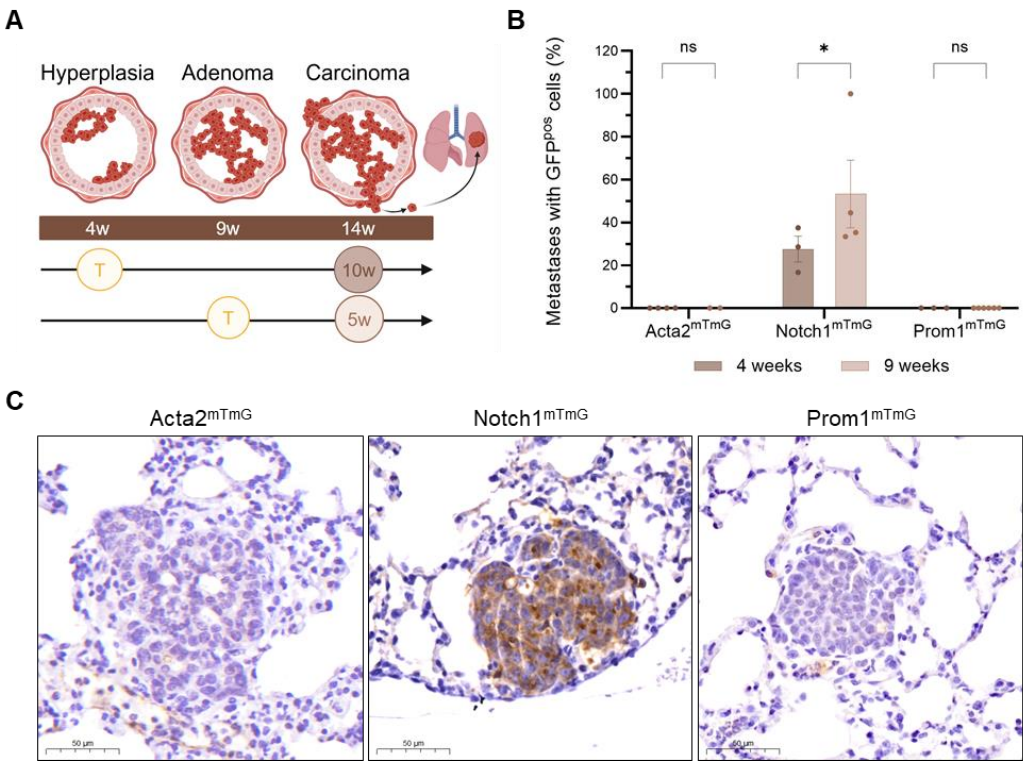


Figure R4.3. *Notch1*-positive cells are highly metastatic in advanced tumor stages in the MMTV-PyMT model. **A.** Experimental design to investigate early and late mammary epithelial cell contributions to lung metastasis in the MMTV-PyMT model. Cre recombination was initiated at the hyperplasia (4 weeks) or adenoma (9 weeks) stages, with lung tissues analyzed 10 or 5 weeks later, respectively. **B.** Bar plot representing the percentage of lung metastases containing GFP^{pos} cells per animal in lung sections from Acta2^{mTmG}, Notch1^{mTmG}, and Prom1^{mTmG} mice after a 10-week chase from hyperplasia (n= 49, 27, 47 metastases), and after a 5-week chase from adenoma (n=39, 52, 39 metastases). At least two animals per time point and genotype were analyzed. Data are presented as mean $\pm$ S.E.M. Significance was determined using the two-way ANOVA test, with $p < 0.05$ considered significant. **C.** Representative IHC images of lung metastases showcasing GFP-labeled progeny exclusively in the Notch1^{mTmG} mice. GFP staining appears brown, while nuclei are counterstained in blue. Scales bars represent 50 μ m.

RESULTS: OBJECTIVE 4

Immunohistochemical analysis revealed that GFP-labeled progeny was predominantly found within the metastatic nodules of *Notch1*^{mTmG} mice (**Figures R4.3B-C**). Quantitative analyses showed that *Notch1*-positive cells targeted at the hyperplasia stage contributed to $27.58 \pm 6.03\%$ of metastatic nodules (**Figure R4.3B**). Notably, their contribution increased significantly when labeled at later tumor stages, reaching $53.27 \pm 15.76\%$ following a 5-week chase from the adenoma stage. This nearly twofold increase highlights the progressively enhanced metastatic potential of *Notch1*-positive cells, likely influenced by tumor microenvironmental changes and dynamic oncogenic signaling as the disease advances.

In contrast, *Prom1*-positive cells showed no evidence of metastatic ability when targeted at different tumor stages (**Figure R4.3B**). However, the recombination efficiency in the *Prom1*^{mTmG} line was extremely low (0.12%), limiting definitive conclusions. Similarly, *Acta2*-positive BaCs did not contribute to lung metastases at either early or advanced tumor stages (**Figure R4.3B**). Despite the absence of GFP-positive BaC-derived clones in metastatic sites, *Acta2*-positive cells were present within metastatic nodules, suggesting that LCs may acquire basal-like features through cellular plasticity during metastasis.

Overall, these findings establish *Notch1*-positive LCs as the primary contributors to lung metastases in the MMTV-PyMT model. Their metastatic potential increases significantly as tumors progress, positioning them as a promising therapeutic target for intervention.

4.3. Phenotypic plasticity of *Notch1*-derived cells in lung metastases

To further explore the metastatic behavior of *Notch1*-expressing cells, we characterized their cellular progeny within lung metastases. Surprisingly, GFP^{pos} cells predominantly formed small clusters within metastatic nodules, rather than constituting the entire metastasis (**Figure R4.4**). This observation aligns with prior studies suggesting that mammary epithelial tumor cells invade the lungs as cohesive clusters rather than as single cells [11]. These clusters, integrating cooperatively within the lung microenvironment, contribute to the formation of metastatic lesions, emphasizing the importance of collective tumor cell migration as a key mechanism in the metastatic cascade.

Immunofluorescence analysis of cell identity markers revealed distinct populations of *Notch1*-derived cells within lung metastases, providing valuable insights into their plasticity during metastatic progression. Notably, most of these cells retained their luminal identity, exclusively expressing the K8 luminal marker (**Figure R4.4A**, upper panel). However, a subset of *Notch1*-derived cells co-expressed the P63 basal marker, indicative of a hybrid luminal-basal state (**Figure R4.4A**, bottom panel). This transitional phenotype suggests that *Notch1*-derived cells may undergo a stepwise process, acquiring hybrid characteristics before fully transitioning to a basal-like state.

Further evidence of this luminal-to-basal lineage shift was provided by the detection of *Notch1*-derived cells expressing K5 and SMA basal markers (**Figure R4.4B**). These

RESULTS: OBJECTIVE 4

markers, typically associated with basal identity, underscore the potential for *Notch1*-derived cells to adopt a more invasive and migratory phenotype during metastatic outgrowth.

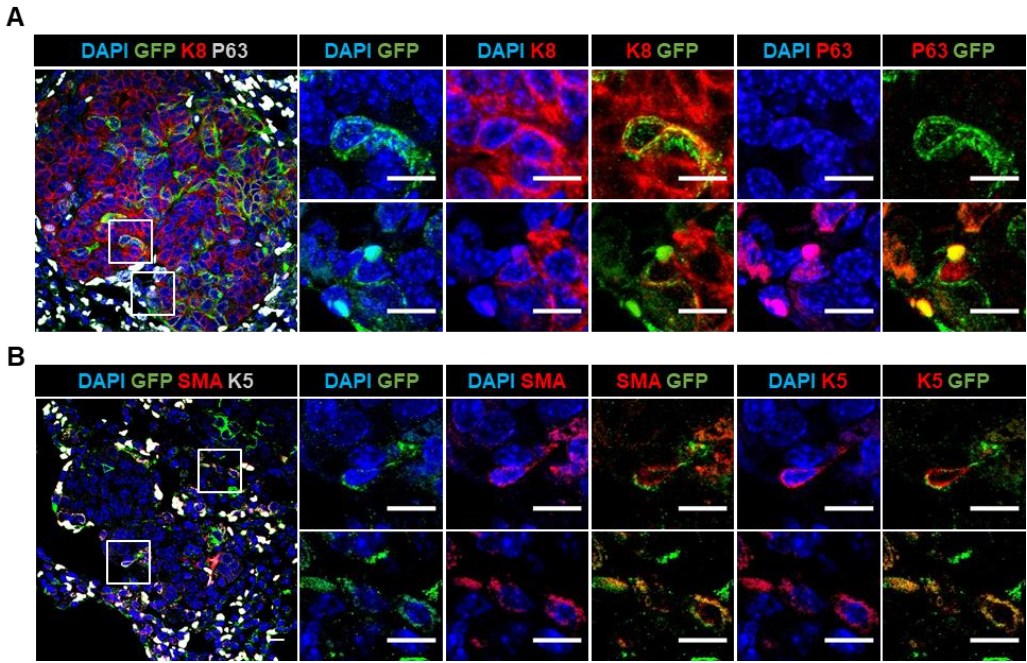


Figure R4.4. *Notch1*-derived cells exhibit luminal-to-basal transition during metastatic spread. IF images of lung metastases with GFP^{pos} cells after a 5-week chase from adenoma in the *Notch1*^{mtmG} line. **A.** Lung metastasis containing GFP-labeled *Notch1*-derived cells with a K8^{pos} luminal phenotype (upper panel) or a K8^{pos} P63^{pos} hybrid phenotype (bottom panel). **B.** Lung metastasis with GFP^{pos} *Notch1*-derived cells co-expressing SMA and K5 basal markers, indicative of a luminal-to-basal lineage shift. Merged images show GFP in green, K8 and SMA markers in red, and P63 and K5 markers in gray. Insets provide magnified views of the highlighted regions. DNA is counterstained with DAPI in blue. Scale bars represent 10 μ m for both the main images and insets.

These findings demonstrate that *Notch1*-derived cells undergo luminal-to-basal transition during metastatic dissemination, enabling the re-establishment of primary tumor architecture at distant sites. The luminal-to-basal transition in *Notch1*-derived cells underscores their remarkable phenotypic plasticity, equipping them with enhanced migratory and invasive capabilities, facilitating metastatic dissemination and colonization. This adaptability enables them to leverage diverse biological programs for survival and growth in distant microenvironments. These insights open new avenues for therapeutic strategies aimed at disrupting the cellular processes that facilitate the migration of metastatic breast cancer cells.

RESULTS: OBJECTIVE 4

DISCUSSION

1. Preclinical models recapitulate TNBC subtype diversity observed in clinics

1.1. Histological and molecular variability across models

Our study provides an in-depth characterization of preclinical TNBC models, highlighting their distinct phenotypic and transcriptomic profiles. Collectively, these models mirror the diversity of breast cancer subtypes seen in clinical settings, offering invaluable tools for investigating TNBC.

While the MMTV-PyMT and C3(1)-TAg models predominantly develop homogeneous ER α -negative luminal tumors, the DMBA model produces a broader spectrum of tumor types, including ER α -negative luminal, basal, and hybrid phenotypes. These findings reveal a critical observation: significant mismatches often exist between histological and transcriptomic classifications. For example, basal-like tumors in the C3(1)-TAg and DMBA models exhibit luminal histological features akin to MMTV-PyMT tumors, which have a luminal B transcriptomic profile. Conversely, tumors with basal histological traits in the DMBA model (designated as “true-basal”) align with normal-like molecular characteristics, mirroring similar discrepancies reported in clinical TNBC cases. Such inconsistencies underscore the importance of integrating both histological and molecular data in preclinical and clinical studies.

The dual-level analysis provides a more accurate reflection of tumor biology and potential therapeutic responses, addressing the heterogeneity that complicates TNBC management. Notably, the molecular classification of TNBC has practical implications for therapy, with luminal B subtypes potentially benefiting from endocrine-like or CDK4/6 inhibitor treatments typically reserved for hormone receptor-positive cancers. These findings reinforce the critical role of transcriptomic profiling in guiding personalized treatment strategies, as immunohistochemical classification alone may fail to capture actionable insights.

The basal phenotype, often associated with poorer prognosis but increased chemotherapy sensitivity [192-195], was faithfully reproduced in the DMBA model. This model generated basal and hybrid tumor types (DMBA-Type II and III, respectively), mimicking the heterogeneity seen in clinical basal-like TNBC. However, some TNBCs observed in the clinic lack basal marker expression, a feature recapitulated by MMTV-PyMT, C3(1)-TAg, and DMBA-Type I tumors.

Thus, the preclinical models used in this study effectively capture the molecular and phenotypic variability observed in human TNBC, making them powerful platforms for investigating its biology and developing subtype-specific therapeutic strategies.

1.2. Model-specific strengths and applications

Preclinical models vary in their utility and translational relevance based on their tumor induction mechanisms. Genetically engineered models, such as MMTV-PyMT and C3(1)-TAg, generate consistent tumor phenotypes, making them ideal for studying

DISCUSSION

tumor evolution and specific pathway dynamics. For instance, MMTV-PyMT tumors progress rapidly and metastasize frequently, providing a robust model for investigating luminal B progression and metastatic dissemination. Conversely, the slower tumor progression and lower metastatic incidence in the C3(1)-TAG model make it well-suited for studying basal-like tumor biology.

In comparison, the DMBA model, which relies on chemically induced random mutagenesis, generates diverse tumor subtypes, even within the same animal. This variability closely mirrors the heterogeneity observed in human breast cancers, making it a powerful model for studying tumor plasticity, intra-tumoral heterogeneity, and the effects of random oncogenic events under varying selective pressures.

Together, these models complement each other, offering a robust preclinical framework for TNBC research. The DMBA model captures the complexity and adaptability of TNBC, while the MMTV-PyMT and C3(1)-TAG models allow for controlled investigations of specific pathways, clonal dynamics, and metastatic behavior. Leveraging the unique strengths of each model is essential for advancing our understanding of TNBC and developing effective therapeutic strategies tailored to the distinct features of this heterogeneous disease.

2. Interplay between cellular origins and oncogenic drivers in TNBC heterogeneity

2.1. Unveiling the origins of luminal tumors: insights and controversies

Our lineage tracing experiments reveal that luminal progenitors, specifically ER α -negative LCs, serve as the primary source of aggressive luminal B and basal-like tumors in the MMTV-PyMT, C3(1)-TAG, and DMBA models. These findings align with prior studies that identified luminal progenitors as critical drivers of TNBC upon activation of specific oncogenes [92,93,217,218]. However, our results diverge from earlier reports suggesting a basal origin for certain tumor types. For instance, *Hagerling et al.* attributed C3(1)-TAG tumors to a basal origin using the *Lgr5* promoter [8]. Yet, as *Lgr5* predominantly labels myoepithelial cells and rare LCs, questions remain about the specificity of this marker for basal cells. Similarly, *Rios et al.* reported luminal tumors arising from *Krt5*-positive BaCs in the MPA+DMBA model [244]. However, concerns over *Krt5* promoter specificity, including potential luminal leakage, complicate the interpretation of these findings. Such discrepancies underscore the importance of highly specific lineage tracing approaches to delineate the distinct roles of MEC subpopulations in tumor initiation and progression.

Our findings emphasize the correlation between cellular origin and tumor subtype. ER α -negative luminal progenitors predominantly drive tumors classified as ER α -negative luminal, suggesting that the cellular origin plays a critical role in shaping tumor identity. These insights highlight the complexity of TNBC biology, where cellular origin interacts with oncogenic drivers to determine the histological and molecular characteristics of tumors.

DISCUSSION

2.2. Basal origins of TNBC: insights into normal-like tumorigenesis

Our findings demonstrate that true basal tumors with normal-like gene expression profiles originate exclusively from *Acta2*-positive BaCs. Lineage tracing experiments identified GFP-positive clones only in *Acta2*^{mTmG} tumors with a basal histology, while *Notch1*^{mTmG} tumors lacked GFP-positive clones, ruling out luminal progenitors as the source of these tumors.

Remarkably, true basal tumors exhibited low levels of claudin protein, which may serve as a potential marker of aggressive tumors [99,174,177,287]. Computational analyses further supported this finding, showing strong similarities between basal cell signatures and claudin-low tumors [240-242]. The reduced proliferative capacity of true basal, claudin-low tumors is consistent with previous studies linking the claudin-low subtype with diminished expression of luminal and proliferation-related genes [174].

In the DMBA model, random mutagenesis appears to favor BaCs for transformation into claudin-low tumors. Conversely, specific oncogenic drivers, such as KRAS, can override lineage constraints, reprogramming luminal cells into basal-like or stem-like states [99]. Notably, in our study, *Notch1*-positive luminal progenitors resisted transformation into claudin-low phenotypes, likely due to their limited plasticity or the absence of pathways necessary for dedifferentiation. This contrast underscores the interplay between cell type, oncogenic driver, and tumor microenvironment in shaping tumor phenotypes.

2.3. Hybrid tumors in TNBC: bridging luminal and basal characteristics

Hybrid tumors, characterized by co-expression of luminal and basal markers, present unresolved questions about their cellular origins. In our models, lineage tracing failed to detect GFP-positive clones in any Cre line, suggesting that these tumors may arise from progenitor populations not efficiently targeted by our experimental approaches. Despite this ambiguity, hybrid tumors shared several features with true basal tumors, including claudin-low protein profiles and slower growth rates, hinting at a possible common basal origin. These findings highlight the biological and clinical complexity of hybrid tumors, emphasizing the need for further research to clarify their origins and therapeutic implications.

2.4. DMBA-induced oncogenesis: a driver of phenotypic plasticity

Our lineage tracing experiments revealed that DMBA-induced mutations drive phenotypic plasticity in both BaCs and LCs, contributing to the formation of tumors with diverse histological and molecular characteristics. While ER α -negative luminal tumors (DMBA-Type I) originate from LCs, true basal tumors (DMBA-Type II) arise from BaCs. However, the final tumor phenotype is shaped by the interplay of oncogenic events and cellular plasticity, rather than being solely determined by the cell of origin.

DMBA-induced plasticity enables MECs to undergo dedifferentiation, reprogramming, or transdifferentiation, resulting in intra-tumoral heterogeneity [92,93]. This phenomenon was exemplified by luminal clones arising from *Acta2*-positive BaCs in true basal tumors, highlighting basal-to-luminal transition as a key feature of DMBA-

DISCUSSION

driven tumorigenesis. Additionally, hybrid tumors (DMBA-Type III) may originate from BaCs undergoing partial lineage transitions, as evidenced by the presence of intermediate hybrid states within these tumors.

Moreover, LCs targeted by the *Notch1* promoter generated tumors expressing SMA, a marker associated with EMT, which is linked to enhanced invasiveness and metastatic potential [288]. While SMA expression is rare in human tumors, its association with basal-like breast cancer suggests that DMBA-induced tumors may provide valuable insights into rare clinical subtypes [294].

In summary, the DMBA model captures the phenotypic plasticity and heterogeneity inherent to TNBC, making it an invaluable tool for studying the interplay between oncogenic drivers and cellular plasticity. Future research should focus on identifying the mutations driving specific TNBC subtypes to inform the development of targeted therapies and advance personalized medicine.

3. Luminal progenitors: a promising focus for treating luminal B and basal-like subtypes

3.1. The evolving clonal landscape of luminal progenitors in luminal B tumors

In the physiological mammary gland, *Prom1* is exclusively associated with ER α -positive luminal cells, marking hormonally regulated luminal subsets [86]. However, in the MMTV-PyMT and C3(1)-TAg models, *Prom1* expression extends beyond its conventional lineage restriction, labeling a subset of highly proliferative ER-negative luminal progenitors. Notably, *Prom1* (CD133) is widely recognized as a surface marker of cells with stem-like properties across various tumor contexts [290]. Isolated *Prom1*-positive cells have demonstrated tumor-initiating capacity, along with enhanced tumorigenicity and metastatic potential both *in vitro* and *in vivo* [291-293]. The identification of this *Prom1*-positive population in our TNBC models supports previous findings linking *Prom1* to the most aggressive breast cancer subtypes, including triple-negative and ER-negative/HER2-positive tumors [293].

Clonal expansion experiments in the MMTV-PyMT model, which molecularly resembles luminal B tumors, revealed the contribution of ER α -negative luminal progenitors across tumor stages, contrasting with the progressive loss of basal cells over time. Notably, the clonal dynamics of *Notch1*-positive and *Prom1*-positive luminal progenitors was closely tied to tumor stage, with their relative contributions shifting over time. Interestingly, these clonal patterns mirrored the differential growth trajectories observed during intravital imaging of multicolored PyMT tumor cells, encompassing continuous growth, delayed proliferation, cell cycle arrest, and eventual attrition [10].

Notch1-positive cells emerged as dominant drivers of early clonal expansion, forming large clones during hyperplasia and establishing the tumor's architectural framework. This early dominance highlights their intrinsic self-renewal capacity and heightened susceptibility to tumor-promoting cues. However, as tumors progress, the clonal expansion of *Notch1*-positive cells declines due to increasing cell attrition and competition within the evolving tumor microenvironment.

DISCUSSION

In contrast, *Prom1*-positive cells exhibited a delayed but significant proliferative response, particularly during adenoma and carcinoma stages. Initially quiescent, these cells appear to acquire a more proliferative phenotype under tumorigenic pressures, highlighting their plasticity and adaptive potential. This compensatory expansion may enable *Prom1*-positive cells to sustain tumor growth as the contribution of *Notch1*-positive cells declines. A plausible hypothesis is that the *Prom1*-positive cells may originate from *Notch1*-positive progenitors that expand during tumor progression, suggesting that some late-stage clones observed in the *Notch1*^{mTmG} line could include a subset of *Prom1*-positive cells. Further investigation is needed to determine whether *Prom1*-positive cells arise from *Notch1*-positive progenitors or constitute a distinct subset of ER α -negative luminal progenitors.

Acta2-positive BaCs, by comparison, demonstrated a consistent lack of clonal expansion throughout tumor progression. Their confinement to the basal layer and absence of proliferative activity highlights their limited oncogenic potential in triple-negative tumors with a luminal phenotype, playing a primarily structural or supportive role within the tumor microenvironment.

These findings underscore the temporal heterogeneity in the contributions of luminal progenitors to luminal B tumorigenesis. While *Notch1*-positive cells drive early tumor progression, *Prom1*-positive cells assume a more prominent role at later stages. This dynamic interplay suggests that effective therapeutic strategies should target *Notch1*-positive cells to intercept early tumorigenesis while addressing the adaptive responses of *Prom1*-positive cells in advanced-stage tumors.

3.2. *Notch1*-positive luminal progenitors in the C3(1)-TAg model: key drivers of basal-like tumorigenesis

The C3(1)-TAg model, which molecularly mirrors human basal-like breast tumors, highlights the central role of *Notch1*-positive luminal progenitors in driving basal-like tumorigenesis. In this model, the three mammary epithelial populations exhibited varying susceptibilities to malignant transformation by the SV40-TAg oncogene. *Notch1*-positive cells displayed the highest transformation efficiency, whereas only a minor subset of *Acta2*- and *Prom1*-expressing cells underwent transformation. The absence of transformation in these populations raises two possibilities: either they possess intrinsic resistance to SV40-TAg, or they undergo senescence or apoptosis in response to oncogenic stress, as previously reported in other tumorigenic contexts [295-297]. These findings contrast with studies using alternative oncogenes, where both ER α -positive and ER α -negative populations exhibited a proliferative advantage upon HER2 expression [9], highlighting the distinct lineage-specific responses to oncogenic drivers.

This variation in oncogene susceptibility across MECs may explain the differential contributions observed in clonal expansion experiments. Notably, only the *Notch1*-positive population exhibited sustained clonal expansion throughout tumor progression, reinforcing their dominance and adaptability within the tumor microenvironment.

DISCUSSION

However, these findings do not exclude the potential contributions of untraced populations, which may gain a proliferative advantage and outcompete the *Notch1*-positive lineage over time. Future studies using alternative Cre lines will be crucial to elucidate these dynamics. Additionally, given the slower tumor progression rate in the C3(1)-TAg model, extended chase experiments may better capture the long-term contributions of traced progenitors.

A particularly striking feature of this model is the emergence of a *Notch1*-positive basal subset. While these cells have a more limited clonogenic potential than luminal clones, they may still contribute to early tumor development. This phenomenon is likely a consequence of oncogenic stress induced by SV40-TAg, as a comparable basal subset was absent in the MMTV-PyMT model. The early presence of *Notch1*-positive tumor BaCs during hyperplasia raises questions about their origin. One possibility is that a small group of BaCs might start expressing *Notch1* in response to oncogenic stress, which could boost their proliferative potential and promote EMT [298,299]. Another possibility is that these basal cells could come from *Notch1*-positive LCs that undergo phenotypic changes. Previous studies in ER α -positive breast cancer have shown that HR-positive luminal cells can adopt a more basal-like phenotype due to *Esr1* mutations or long-term estrogen deprivation, with ER α antagonists further exacerbating this transition [300]. Future lineage tracing studies targeting SV40-TAg-naïve *Notch1*-expressing cells at early stages could help clarify the source of this *Notch1*-positive basal subset in the C3(1)-TAg model.

Importantly, *Notch1*-positive BaCs represent a distinct subset from those characterized by *Acta2* (SMA) expression, as *Acta2*-traced cells did not undergo clonal expansion. Instead, this *Notch1*-positive basal subset may represent a unique progenitor pool, particularly sensitive to oncogenic pressure, whose expansion could fuel intratumor heterogeneity.

In summary, the sustained dominance and diverse contributions of *Notch1*-positive subsets underscore their pivotal role in shaping tumor heterogeneity and evolution in basal-like TNBC. These findings offer critical insights for the development of therapeutic strategies targeting luminal progenitors in basal-like breast cancer.

4. Insights into the metastatic behavior and phenotypic shifts of *Notch1*-derived cells

4.1. Lung metastases reflect the cellular composition of the primary tumor

This study provides novel insights into how different MEC populations contribute to lung metastases in the MMTV-PyMT model, a widely used preclinical model for studying metastatic breast cancer. Lineage tracing experiments revealed that lung metastases mirrored the cellular diversity of the primary tumor, comprising the same luminal and basal populations observed in hyperplastic lesions. The proportional distribution of *Acta2*-, *Notch1*-, and *Prom1*-positive cells between primary tumors and metastatic sites remained remarkably consistent, suggesting that the metastatic process involves a preservation of the primary tumor's cellular architecture.

DISCUSSION

These findings emphasize the inherent heterogeneity of breast tumors, not only at the primary site but also in their metastatic outgrowth. This reflects the complex interplay of epithelial subtypes during tumor progression and spread, highlighting the importance of considering cellular diversity when studying metastatic breast cancer.

4.2. *Notch1*-positive cells enhance their aggressiveness in advanced tumor stages

Despite the presence of all three MEC populations in lung metastases, lineage tracing studies revealed that only *Notch1*-positive LCs exhibit significant metastatic capacity. In contrast, *Prom1*-positive luminal and *Acta2*-positive basal progenitors lack this potential, even at advanced tumor stages. However, their presence within metastatic nodules suggests that *Notch1*-positive cells may undergo lineage plasticity during the metastatic process, adopting traits that facilitate their survival and adaptation in the distant organ environment.

Notch1-positive cells contributed to metastases as early as the hyperplastic stage, with their representation within metastatic nodules increasing in later tumor stages. This progressive enrichment suggests that selective pressures within the evolving tumor microenvironment favor *Notch1*-positive cells enhancing their metastatic potential over time. Interestingly, this increase in metastatic activity coincided with a decline in the clonal expansion capacity of *Notch1*-positive cells at the primary site. This shift suggests a reallocation of resources, with cells prioritizing metastatic dissemination over local proliferation at advanced stages of tumor progression.

These findings establish *Notch1*-positive cells as key players in the metastatic cascade, offering a potential target for therapeutic intervention in advanced breast cancer. However, our analyses do not conclusively rule out the potential metastatic behavior of *Prom1*-positive cells. Due to the low recombination efficiency in the *Prom1*^{mTmG} line, a larger number of metastases must be analyzed to draw firm conclusions. Future studies combining lineage tracing studies with single-cell RNA sequencing will definitely uncover their molecular profiles and clarify their metastatic potential.

4.3. Phenotypic plasticity of *Notch1*-derived cells during metastatic dissemination

A key discovery of this study is the phenotypic plasticity of *Notch1*-derived cells during metastatic progression. While previous research identified K8-positive luminal tumor cells as the primary drivers of distant metastases [243], it likely overlooked luminal-to-basal transitions due to reliance on the K14 marker, which can be expressed by both luminal and basal lineages depending on the tumor context. In contrast, this study detected a subset of *Notch1*-derived cells transitioning from an initial luminal identity to a basal-like phenotype, marked by the expression of robust basal markers such as P63, K5, and SMA.

This plasticity likely enhances the migratory and invasive capacities of *Notch1*-derived cells, equipping them to adapt to the challenging microenvironments encountered during metastatic colonization. The luminal-to-basal shift may also confer additional advantages, such as increased resistance to apoptosis and enhanced survival in

DISCUSSION

secondary tissues. Similar plasticity has been observed in *ErbB2*-positive breast cancer models, where the loss of ER α correlates with greater metastatic potential [9].

The presence of cohesive clusters of *Notch1*-derived cells within metastatic nodules further reinforces the notion that collective cell migration is a critical mechanism for breast cancer metastasis. These clusters, rather than individual cells, likely cooperate to successfully establish secondary lesions, as shown in previous multicolored lineage tracing studies [11]. This study adds to the growing body of literature by demonstrating that *Notch1*-positive cells exhibit dynamic phenotypic plasticity, transitioning from luminal to basal-like phenotypes to better survive and thrive in metastatic environments.

5. Concluding remarks and future directions

This thesis has provided critical insights into the cellular origins, clonal dynamics, and metastatic behavior of distinct MEC populations across preclinical TNBC models. By integrating lineage tracing, transcriptomic analysis, and clonal expansion studies, we have advanced our understanding of how distinct oncogenic drivers and cellular origins shape TNBC heterogeneity, progression, and metastasis. These findings not only reinforce the value of preclinical models in recapitulating the diversity of human TNBC but also open new avenues for therapeutic interventions targeting specific cellular populations.

5.1. Key contributions

Preclinical models reflect TNBC subtype diversity: Our comparative analysis of the MMTV-PyMT, C3(1)-TAg, and DMBA-induced models has demonstrated their ability to capture the molecular and histological heterogeneity of TNBC subtypes observed in clinical settings. The alignment between model-specific tumor phenotypes and human TNBC subtypes validates their translational relevance, with each model offering unique advantages for investigating specific aspects of TNBC biology.

Cellular origins drive TNBC heterogeneity: We confirmed that luminal ER α -negative progenitors are the primary cells of origin for luminal B and basal-like tumors across multiple models, both of which are histologically classified as luminal. In contrast, true basal tumors, which molecularly resemble the normal-like subtype, arise exclusively from basal progenitors. These findings underscore the critical role of cellular origin in shaping tumor identity and provide a framework for understanding how lineage-specific traits influence tumor behavior and therapeutic response.

Dynamic clonal expansion of luminal progenitors: Clonal expansion studies revealed that *Notch1*-positive luminal progenitors dominate tumor initiation and early progression, while *Prom1*-positive cells contribute to later stages. The temporal and spatial dynamics of these populations highlight the plasticity and adaptability of luminal progenitors under tumorigenic pressure, providing potential targets for stage-specific therapies.

DISCUSSION

Metastatic potential and phenotypic plasticity: *Notch1*-positive cells emerged as key drivers of lung metastasis in the MMTV-PyMT model, exhibiting phenotypic plasticity that enabled transitions between luminal and basal states during metastatic dissemination. This plasticity likely enhances their migratory and invasive capabilities, providing a new perspective on the cellular mechanisms underpinning TNBC metastasis.

5.2. Future directions

Building on these findings, several key areas warrant further exploration:

Unraveling oncogenic pathway interactions: Future studies should investigate how specific oncogenic drivers interact with cellular origin to influence tumor heterogeneity and progression. Leveraging single-cell RNA sequencing and proteomics could provide a deeper understanding of how tumor microenvironments and mutational landscapes shape these interactions.

Therapeutic targeting of luminal progenitors: Given their dominant role in tumor initiation and progression, targeting luminal progenitors, particularly *Notch1*-positive cells, offers a promising therapeutic strategy. Efforts should focus on identifying vulnerabilities within these populations, such as signaling pathways that drive their clonal expansion or promote metastatic dissemination.

Exploring basal progenitor contributions: While basal progenitors showed limited direct contribution to metastasis, their involvement in the formation of normal-like tumors highlights their relevance in TNBC biology. Future studies could explore how environmental cues, or specific oncogenic signals enhance their tumorigenic potential and plasticity.

Mechanisms of phenotypic plasticity: The luminal-to-basal transitions observed in *Notch1*-positive cells underscore the importance of understanding cellular plasticity in metastatic progression. Investigating the molecular drivers of this plasticity, including epigenetic regulators and transcription factors, may reveal novel targets for disrupting metastasis.

Refining models for rare TNBC subtypes: The DMBA model's ability to recapitulate rare basal-like subtypes highlights the value of chemically induced models in studying tumor heterogeneity. Expanding these models to explore additional rare subtypes or mutational contexts could offer new insights into less-studied aspects of TNBC.

Integrating multi-omic approaches: Combining transcriptomic, proteomic, and spatial analysis in preclinical models will allow for a more holistic understanding of TNBC biology. Such integrative approaches could identify novel biomarkers for subtype-specific diagnosis and therapy.

CONCLUSIONS

CONCLUSIONS

OBJECTIVE 1

- MMTV-PyMT and C3(1)-TAg models develop homogeneous ER α -negative luminal tumors, whereas DMBA-induced tumors exhibit significant heterogeneity, forming three distinct histological subtypes: ER α -negative luminal (Type I), basal (Type II), and hybrid basal-luminal (Type III).
- Murine models effectively recapitulate the molecular diversity of clinical TNBC subtypes: MMTV-PyMT aligns with luminal B, C3(1)-TAg with basal-like, and DMBA-induced tumors exhibit heterogeneity, with Type I resembling basal-like and Type II mirroring normal-like.
- Tumors with luminal histology (luminal B and basal-like) are claudin-high, while tumors with basal (normal-like) or hybrid histology are claudin-low.

OBJECTIVE 2

- ER α -negative luminal (*Notch1*-positive) cells give rise to TNBC luminal tumors, regardless of their classification as luminal B or basal-like based on gene expression profiles.
- TNBC tumors with a histologically basal phenotype and a normal-like molecular profile originate from *Acta2*-positive basal cells.
- *Acta2*-positive basal cells exhibit phenotypic plasticity during DMBA-induced tumorigenesis, transitioning into hybrid and luminal phenotypes, supporting a basal origin for hybrid (Type III) DMBA-induced tumors.
- Luminal tumors exhibit shorter latencies compared to basal and hybrid tumors, highlighting their heightened proliferative capacity.

OBJECTIVE 3:

- In the MMTV-PyMT model (luminal B), *Notch1*- and *Prom1*-expressing ER α -negative luminal cells persist across all tumor stages, whereas *Acta2*-expressing basal cells progressively decline, reflecting the loss of the basal layer in advanced tumors.
- *Notch1*-positive luminal progenitors drive early clonal expansion in the MMTV-PyMT model, whereas *Prom1*-positive luminal cells exhibit a delayed but robust proliferative response in advanced tumor stages.
- *Acta2*-positive basal cells display limited proliferative potential throughout tumorigenesis in the MMTV-PyMT model, remaining confined to the basal layer without significant contributions to tumor progression.
- In the C3(1)-TAg model (basal-like), *Notch1*-positive cells, including ER α -negative luminal cells and a minor subset of basal cells, show high susceptibility to SV40-TAg-driven transformation, sustaining clonal expansion throughout tumor progression.
- *Prom1*-positive luminal cells and *Acta2*-positive basal cells are rarely transformed by SV40-TAg in the C3(1)-TAg model, consistent with their negligible proliferative activity in both early and advanced basal-like tumors.

CONCLUSIONS

OBJECTIVE 4:

- *Notch1*-positive cells are key contributors to lung metastases in the MMTV-PyMT model, with their metastatic potential significantly increasing in advanced tumor stages.
- *Notch1*-derived cells undergo luminal-to-basal transition during metastatic spread, facilitating the reconstruction of the primary tumor architecture at distant metastatic sites.
- *Acta2*-positive basal cells lack intrinsic metastatic capacity; their presence in metastatic nodules is solely attributed to transdifferentiation processes.

CONCLUSIONS

BIBLIOGRAPHY

BIBLIOGRAPHY

1. Kretzschmar, K.; Watt, F.M. Lineage tracing. *Cell* **2012**, *148*, 33-45, doi:10.1016/j.cell.2012.01.002.
2. Van Keymeulen, A.; Rocha, A.S.; Ousset, M.; Beck, B.; Bouvencourt, G.; Rock, J.; Sharma, N.; Dekoninck, S.; Blanpain, C. Distinct stem cells contribute to mammary gland development and maintenance. *Nature* **2011**, *479*, 189-193, doi:10.1038/nature10573.
3. Centonze, A.; Lin, S.; Tika, E.; Sifrim, A.; Fioramonti, M.; Malfait, M.; Song, Y.; Wuidart, A.; Van Herck, J.; Dannau, A.; et al. Heterotypic cell-cell communication regulates glandular stem cell multipotency. *Nature* **2020**, *584*, 608-613, doi:10.1038/s41586-020-2632-y.
4. Rodilla, V.; Dasti, A.; Huyghe, M.; Lafkas, D.; Laurent, C.; Rey, F.; Fre, S. Luminal progenitors restrict their lineage potential during mammary gland development. *PLoS Biol* **2015**, *13*, e1002069, doi:10.1371/journal.pbio.1002069.
5. Lilja, A.M.; Rodilla, V.; Huyghe, M.; Hannezo, E.; Landragin, C.; Renaud, O.; Leroy, O.; Rulands, S.; Simons, B.D.; Fre, S. Clonal analysis of Notch1-expressing cells reveals the existence of unipotent stem cells that retain long-term plasticity in the embryonic mammary gland. *Nat Cell Biol* **2018**, *20*, 677-687, doi:10.1038/s41556-018-0108-1.
6. Lafkas, D.; Rodilla, V.; Huyghe, M.; Mourao, L.; Kiaris, H.; Fre, S. Notch3 marks clonogenic mammary luminal progenitor cells in vivo. *J Cell Biol* **2013**, *203*, 47-56, doi:10.1083/jcb.201307046.
7. Blaas, L.; Pucci, F.; Messal, H.A.; Andersson, A.B.; Josue Ruiz, E.; Gerling, M.; Douagi, I.; Spencer-Dene, B.; Musch, A.; Mitter, R.; et al. Lgr6 labels a rare population of mammary gland progenitor cells that are able to originate luminal mammary tumours. *Nat Cell Biol* **2016**, *18*, 1346-1356, doi:10.1038/ncb3434.
8. Hagerling, C.; Owyong, M.; Sitarama, V.; Wang, C.Y.; Lin, C.; van den Bijgaart, R.J.E.; Koopman, C.D.; Brenot, A.; Nanjaraj, A.; Warnberg, F.; et al. LGR5 in breast cancer and ductal carcinoma in situ: a diagnostic and prognostic biomarker and a therapeutic target. *BMC Cancer* **2020**, *20*, 542, doi:10.1186/s12885-020-06986-z.
9. Ding, Y.; Liu, Y.; Lee, D.K.; Tong, Z.; Yu, X.; Li, Y.; Xu, Y.; Lanz, R.B.; O'Malley, B.W.; Xu, J. Cell lineage tracing links ERalpha loss in Erbb2-positive breast cancers to the arising of a highly aggressive breast cancer subtype. *Proc Natl Acad Sci U S A* **2021**, *118*, doi:10.1073/pnas.2100673118.
10. Zomer, A.; Ellenbroek, S.I.; Ritsma, L.; Beerling, E.; Vrisekoop, N.; Van Rheenen, J. Intravital imaging of cancer stem cell plasticity in mammary tumors. *Stem Cells* **2013**, *31*, 602-606, doi:10.1002/stem.1296.
11. Cheung, K.J.; Padmanaban, V.; Silvestri, V.; Schipper, K.; Cohen, J.D.; Fairchild, A.N.; Gorin, M.A.; Verdone, J.E.; Pienta, K.J.; Bader, J.S.; et al. Polyclonal breast cancer metastases arise from collective dissemination of keratin 14-expressing tumor cell clusters. *Proc Natl Acad Sci U S A* **2016**, *113*, E854-863, doi:10.1073/pnas.1508541113.

BIBLIOGRAPHY

12. Mao, X.; Fujiwara, Y.; Chapdelaine, A.; Yang, H.; Orkin, S.H. Activation of EGFP expression by Cre-mediated excision in a new ROSA26 reporter mouse strain. *Blood* **2001**, *97*, 324-326, doi:10.1182/blood.v97.1.324.
13. Srinivas, S.; Watanabe, T.; Lin, C.S.; William, C.M.; Tanabe, Y.; Jessell, T.M.; Costantini, F. Cre reporter strains produced by targeted insertion of EYFP and ECFP into the ROSA26 locus. *BMC Dev Biol* **2001**, *1*, 4, doi:10.1186/1471-213x-1-4.
14. Soriano, P. Generalized lacZ expression with the ROSA26 Cre reporter strain. *Nat Genet* **1999**, *21*, 70-71, doi:10.1038/5007.
15. Peikon, I.D.; Gizatullina, D.I.; Zador, A.M. In vivo generation of DNA sequence diversity for cellular barcoding. *Nucleic Acids Res* **2014**, *42*, e127, doi:10.1093/nar/gku604.
16. Pei, W.; Feyerabend, T.B.; Rossler, J.; Wang, X.; Postrach, D.; Busch, K.; Rode, I.; Klapproth, K.; Dietlein, N.; Quedenau, C.; et al. Polylox barcoding reveals haematopoietic stem cell fates realized in vivo. *Nature* **2017**, *548*, 456-460, doi:10.1038/nature23653.
17. Danielian, P.S.; White, R.; Hoare, S.A.; Fawell, S.E.; Parker, M.G. Identification of residues in the estrogen receptor that confer differential sensitivity to estrogen and hydroxytamoxifen. *Mol Endocrinol* **1993**, *7*, 232-240, doi:10.1210/mend.7.2.8469236.
18. Feil, R.; Brocard, J.; Mascrez, B.; LeMeur, M.; Metzger, D.; Chambon, P. Ligand-activated site-specific recombination in mice. *Proc Natl Acad Sci U S A* **1996**, *93*, 10887-10890, doi:10.1073/pnas.93.20.10887.
19. Metzger, D.; Clifford, J.; Chiba, H.; Chambon, P. Conditional site-specific recombination in mammalian cells using a ligand-dependent chimeric Cre recombinase. *Proc Natl Acad Sci U S A* **1995**, *92*, 6991-6995, doi:10.1073/pnas.92.15.6991.
20. Rios, A.C.; Fu, N.Y.; Lindeman, G.J.; Visvader, J.E. In situ identification of bipotent stem cells in the mammary gland. *Nature* **2014**, *506*, 322-327, doi:10.1038/nature12948.
21. Rodilla, V.; Fre, S. Lineage Tracing Methods to Study Mammary Epithelial Hierarchies In Vivo. *Methods Mol Biol* **2022**, *2471*, 141-157, doi:10.1007/978-1-0716-2193-6_7.
22. Gossen, M.; Bujard, H. Tight control of gene expression in mammalian cells by tetracycline-responsive promoters. *Proc Natl Acad Sci U S A* **1992**, *89*, 5547-5551, doi:10.1073/pnas.89.12.5547.
23. Gossen, M.; Freundlieb, S.; Bender, G.; Muller, G.; Hillen, W.; Bujard, H. Transcriptional activation by tetracyclines in mammalian cells. *Science* **1995**, *268*, 1766-1769, doi:10.1126/science.7792603.
24. Wuidart, A.; Ousset, M.; Rulands, S.; Simons, B.D.; Van Keymeulen, A.; Blanpain, C. Quantitative lineage tracing strategies to resolve multipotency in tissue-specific stem cells. *Genes Dev* **2016**, *30*, 1261-1277, doi:10.1101/gad.280057.116.

BIBLIOGRAPHY

25. Sauer, B.; McDermott, J. DNA recombination with a heterospecific Cre homolog identified from comparison of the pac-c1 regions of P1-related phages. *Nucleic Acids Res* **2004**, *32*, 6086-6095, doi:10.1093/nar/gkh941.
26. Zhang, C.; Chang, F.; Miao, H.; Fu, Y.; Tong, X.; Feng, Y.; Zheng, W.; Ma, X. Construction and application of a multifunctional CHO cell platform utilizing Cre/lox and Dre/rox site-specific recombination systems. *Front Bioeng Biotechnol* **2023**, *11*, 1320841, doi:10.3389/fbioe.2023.1320841.
27. Anastassiadis, K.; Fu, J.; Patsch, C.; Hu, S.; Weidlich, S.; Duerschke, K.; Buchholz, F.; Edenhofer, F.; Stewart, A.F. Dre recombinase, like Cre, is a highly efficient site-specific recombinase in *E. coli*, mammalian cells and mice. *Dis Model Mech* **2009**, *2*, 508-515, doi:10.1242/dmm.003087.
28. Chuang, K.; Nguyen, E.; Sergeev, Y.; Badea, T.C. Novel Heterotypic Rox Sites for Combinatorial Dre Recombination Strategies. *G3 (Bethesda)* **2015**, *6*, 559-571, doi:10.1534/g3.115.025841.
29. Jin, H.; Liu, K.; Zhou, B. Dual recombinases-based genetic lineage tracing for stem cell research with enhanced precision. *Sci China Life Sci* **2021**, *64*, 2060-2072, doi:10.1007/s11427-020-1889-9.
30. He, L.; Li, Y.; Li, Y.; Pu, W.; Huang, X.; Tian, X.; Wang, Y.; Zhang, H.; Liu, Q.; Zhang, L.; et al. Enhancing the precision of genetic lineage tracing using dual recombinases. *Nat Med* **2017**, *23*, 1488-1498, doi:10.1038/nm.4437.
31. Li, Y.; Lv, Z.; Zhang, S.; Wang, Z.; He, L.; Tang, M.; Pu, W.; Zhao, H.; Zhang, Z.; Shi, Q.; et al. Genetic Fate Mapping of Transient Cell Fate Reveals N-Cadherin Activity and Function in Tumor Metastasis. *Dev Cell* **2020**, *54*, 593-607 e595, doi:10.1016/j.devcel.2020.06.021.
32. Liu, Q.; Liu, K.; Cui, G.; Huang, X.; Yao, S.; Guo, W.; Qin, Z.; Li, Y.; Yang, R.; Pu, W.; et al. Lung regeneration by multipotent stem cells residing at the bronchioalveolar-duct junction. *Nat Genet* **2019**, *51*, 728-738, doi:10.1038/s41588-019-0346-6.
33. Gronostajski, R.M.; Sadowski, P.D. The FLP recombinase of the *Saccharomyces cerevisiae* 2 microns plasmid attaches covalently to DNA via a phosphotyrosyl linkage. *Mol Cell Biol* **1985**, *5*, 3274-3279, doi:10.1128/mcb.5.11.3274-3279.1985.
34. Park, Y.N.; Masison, D.; Eisenberg, E.; Greene, L.E. Application of the FLP/FRT system for conditional gene deletion in yeast *Saccharomyces cerevisiae*. *Yeast* **2011**, *28*, 673-681, doi:10.1002/yea.1895.
35. Voutev, R.; Hubbard, E.J. A "FLP-Out" system for controlled gene expression in *Caenorhabditis elegans*. *Genetics* **2008**, *180*, 103-119, doi:10.1534/genetics.108.090274.
36. Luond, F.; Bill, R.; Vettiger, A.; Oller, H.; Pelczar, P.; Christofori, G. A Transgenic MMTV-Flippase Mouse Line for Molecular Engineering in Mammary Gland and Breast Cancer Mouse Models. *J Mammary Gland Biol Neoplasia* **2019**, *24*, 39-45, doi:10.1007/s10911-018-9412-4.
37. Radler, P.D.; Vistisen, K.; Triplett, A.A.; Dennaoui, R.; Li, Y.; Shrestha, H.; Ferraiuolo, R.M.; Thangasamy, A.; Saur, D.; Wagner, K.U. Dual recombinase

BIBLIOGRAPHY

- action in the normal and neoplastic mammary gland epithelium. *Sci Rep* **2021**, *11*, 20775, doi:10.1038/s41598-021-00231-8.
38. Schonhuber, N.; Seidler, B.; Schuck, K.; Veltkamp, C.; Schachtler, C.; Zukowska, M.; Eser, S.; Feyerabend, T.B.; Paul, M.C.; Eser, P.; et al. A next-generation dual-recombinase system for time- and host-specific targeting of pancreatic cancer. *Nat Med* **2014**, *20*, 1340-1347, doi:10.1038/nm.3646.
39. Goodrich, M.M.; Talhouk, R.; Zhang, X.; Goodrich, D.W. An approach for controlling the timing and order of engineered mutations in mice. *Genesis* **2018**, *56*, e23243, doi:10.1002/dvg.23243.
40. Buchholz, F.; Ringrose, L.; Angrand, P.O.; Rossi, F.; Stewart, A.F. Different thermostabilities of FLP and Cre recombinases: implications for applied site-specific recombination. *Nucleic Acids Res* **1996**, *24*, 4256-4262, doi:10.1093/nar/24.21.4256.
41. Zhao, C.; Ries, C.; Du, Y.; Zhang, J.; Sakimura, K.; Itoi, K.; Deussing, J.M. Differential CRH expression level determines efficiency of Cre- and Flp-dependent recombination. *Front Neurosci* **2023**, *17*, 1163462, doi:10.3389/fnins.2023.1163462.
42. Li, J.; Li, Y.; Pawlik, K.M.; Napierala, J.S.; Napierala, M. A CRISPR-Cas9, Cre-lox, and Flp-FRT Cascade Strategy for the Precise and Efficient Integration of Exogenous DNA into Cellular Genomes. *CRISPR J* **2020**, *3*, 470-486, doi:10.1089/crispr.2020.0042.
43. Dawson, C.A.; Visvader, J.E. The Cellular Organization of the Mammary Gland: Insights From Microscopy. *J Mammary Gland Biol Neoplasia* **2021**, *26*, 71-85, doi:10.1007/s10911-021-09483-6.
44. Biswas, S.K.; Banerjee, S.; Baker, G.W.; Kuo, C.Y.; Chowdhury, I. The Mammary Gland: Basic Structure and Molecular Signaling during Development. *Int J Mol Sci* **2022**, *23*, doi:10.3390/ijms23073883.
45. Vinuesa-Pitarch, E.; Ortega-Alvarez, D.; Rodilla, V. How Lineage Tracing Studies Can Unveil Tumor Heterogeneity in Breast Cancer. *Biomedicines* **2021**, *10*, doi:10.3390/biomedicines10010003.
46. Clarke, R.B.; Howell, A.; Potten, C.S.; Anderson, E. Dissociation between steroid receptor expression and cell proliferation in the human breast. *Cancer Res* **1997**, *57*, 4987-4991.
47. Rusidze, M.; Adlanmerini, M.; Chantalat, E.; Raymond-Letron, I.; Cayre, S.; Arnal, J.F.; Deugnier, M.A.; Lenfant, F. Estrogen receptor-alpha signaling in post-natal mammary development and breast cancers. *Cell Mol Life Sci* **2021**, *78*, 5681-5705, doi:10.1007/s00018-021-03860-4.
48. Maller, O.; Martinson, H.; Schedin, P. Extracellular matrix composition reveals complex and dynamic stromal-epithelial interactions in the mammary gland. *J Mammary Gland Biol Neoplasia* **2010**, *15*, 301-318, doi:10.1007/s10911-010-9189-6.
49. Englund, J.I.; Ritchie, A.; Blaas, L.; Cojoc, H.; Penttimikko, N.; Dohla, J.; Iqbal, S.; Patarroyo, M.; Katajisto, P. Laminin alpha 5 regulates mammary gland

BIBLIOGRAPHY

- remodeling through luminal cell differentiation and Wnt4-mediated epithelial crosstalk. *Development* **2021**, 148, doi:10.1242/dev.199281.
50. Manucci, A.C.; Fiore, A.; Genesi, G.L.; Bruni-Cardoso, A. The basement membrane regulates the cellular localization and the cytoplasmic interactome of Yes-Associated Protein (YAP) in mammary epithelial cells. *J Cell Biochem* **2024**, 125, e30606, doi:10.1002/jcb.30606.
51. Fata, J.E.; Werb, Z.; Bissell, M.J. Regulation of mammary gland branching morphogenesis by the extracellular matrix and its remodeling enzymes. *Breast Cancer Res* **2004**, 6, 1-11, doi:10.1186/bcr634.
52. Haaksma, C.J.; Schwartz, R.J.; Tomasek, J.J. Myoepithelial cell contraction and milk ejection are impaired in mammary glands of mice lacking smooth muscle alpha-actin. *Biol Reprod* **2011**, 85, 13-21, doi:10.1095/biolreprod.110.090639.
53. Macias, H.; Hinck, L. Mammary gland development. *Wiley Interdiscip Rev Dev Biol* **2012**, 1, 533-557, doi:10.1002/wdev.35.
54. Wu, D.; Thompson, L.U.; Comelli, E.M. MicroRNAs: A Link between Mammary Gland Development and Breast Cancer. *Int J Mol Sci* **2022**, 23, doi:10.3390/ijms232415978.
55. Cowin, P.; Wysolmerski, J. Molecular mechanisms guiding embryonic mammary gland development. *Cold Spring Harb Perspect Biol* **2010**, 2, a003251, doi:10.1101/cshperspect.a003251.
56. Chu, E.Y.; Hens, J.; Andl, T.; Kairo, A.; Yamaguchi, T.P.; Briskin, C.; Glick, A.; Wysolmerski, J.J.; Millar, S.E. Canonical WNT signaling promotes mammary placode development and is essential for initiation of mammary gland morphogenesis. *Development* **2004**, 131, 4819-4829, doi:10.1242/dev.01347.
57. Veltmaat, J.M.; Van Veelen, W.; Thiery, J.P.; Bellusci, S. Identification of the mammary line in mouse by Wnt10b expression. *Dev Dyn* **2004**, 229, 349-356, doi:10.1002/dvdy.10441.
58. Mailleux, A.A.; Spencer-Dene, B.; Dillon, C.; Ndiaye, D.; Savona-Baron, C.; Itoh, N.; Kato, S.; Dickson, C.; Thiery, J.P.; Bellusci, S. Role of FGF10/FGFR2b signaling during mammary gland development in the mouse embryo. *Development* **2002**, 129, 53-60, doi:10.1242/dev.129.1.53.
59. Veltmaat, J.M.; Relaix, F.; Le, L.T.; Kratochwil, K.; Sala, F.G.; van Veelen, W.; Rice, R.; Spencer-Dene, B.; Mailleux, A.A.; Rice, D.P.; et al. Gli3-mediated somitic Fgf10 expression gradients are required for the induction and patterning of mammary epithelium along the embryonic axes. *Development* **2006**, 133, 2325-2335, doi:10.1242/dev.02394.
60. Hens, J.R.; Dann, P.; Zhang, J.P.; Harris, S.; Robinson, G.W.; Wysolmerski, J. BMP4 and PTHrP interact to stimulate ductal outgrowth during embryonic mammary development and to inhibit hair follicle induction. *Development* **2007**, 134, 1221-1230, doi:10.1242/dev.000182.
61. Hinck, L.; Silberstein, G.B. Key stages in mammary gland development: the mammary end bud as a motile organ. *Breast Cancer Res* **2005**, 7, 245-251, doi:10.1186/bcr1331.

BIBLIOGRAPHY

62. Messal, H.A.; van Rheenen, J.; Scheele, C. An Intravital Microscopy Toolbox to Study Mammary Gland Dynamics from Cellular Level to Organ Scale. *J Mammary Gland Biol Neoplasia* **2021**, *26*, 9-27, doi:10.1007/s10911-021-09487-2.
63. Hannan, F.M.; Elajnaf, T.; Vandenberg, L.N.; Kennedy, S.H.; Thakker, R.V. Hormonal regulation of mammary gland development and lactation. *Nat Rev Endocrinol* **2023**, *19*, 46-61, doi:10.1038/s41574-022-00742-y.
64. Briskin, C.; Rajaram, R.D. Alveolar and lactogenic differentiation. *J Mammary Gland Biol Neoplasia* **2006**, *11*, 239-248, doi:10.1007/s10911-006-9026-0.
65. Oakes, S.R.; Hilton, H.N.; Ormandy, C.J. The alveolar switch: coordinating the proliferative cues and cell fate decisions that drive the formation of lobuloalveoli from ductal epithelium. *Breast Cancer Res* **2006**, *8*, 207, doi:10.1186/bcr1411.
66. Tripathy, S.; Singh, S.; Das, S.K. Potential of breastmilk in stem cell research. *Cell Tissue Bank* **2019**, *20*, 467-488, doi:10.1007/s10561-019-09791-6.
67. Canul-Medina, G.; Fernandez-Mejia, C. Morphological, hormonal, and molecular changes in different maternal tissues during lactation and post-lactation. *J Physiol Sci* **2019**, *69*, 825-835, doi:10.1007/s12576-019-00714-4.
68. Jena, M.K.; Jaswal, S.; Kumar, S.; Mohanty, A.K. Molecular mechanism of mammary gland involution: An update. *Dev Biol* **2019**, *445*, 145-155, doi:10.1016/j.ydbio.2018.11.002.
69. Chang, T.H.; Kunasegaran, K.; Tarulli, G.A.; De Silva, D.; Voorhoeve, P.M.; Pietersen, A.M. New insights into lineage restriction of mammary gland epithelium using parity-identified mammary epithelial cells. *Breast Cancer Res* **2014**, *16*, R1, doi:10.1186/bcr3593.
70. Plaks, V.; Brenot, A.; Lawson, D.A.; Linnemann, J.R.; Van Kappel, E.C.; Wong, K.C.; de Sauvage, F.; Klein, O.D.; Werb, Z. Lgr5-expressing cells are sufficient and necessary for postnatal mammary gland organogenesis. *Cell Rep* **2013**, *3*, 70-78, doi:10.1016/j.celrep.2012.12.017.
71. Shackleton, M.; Vaillant, F.; Simpson, K.J.; Stingl, J.; Smyth, G.K.; Asselin-Labat, M.L.; Wu, L.; Lindeman, G.J.; Visvader, J.E. Generation of a functional mammary gland from a single stem cell. *Nature* **2006**, *439*, 84-88, doi:10.1038/nature04372.
72. Stingl, J.; Eirew, P.; Ricketson, I.; Shackleton, M.; Vaillant, F.; Choi, D.; Li, H.I.; Eaves, C.J. Purification and unique properties of mammary epithelial stem cells. *Nature* **2006**, *439*, 993-997, doi:10.1038/nature04496.
73. Rodilla, V.; Fre, S. Cellular Plasticity of Mammary Epithelial Cells Underlies Heterogeneity of Breast Cancer. *Biomedicines* **2018**, *6*, doi:10.3390/biomedicines6040103.
74. Barker, N.; van Es, J.H.; Kuipers, J.; Kujala, P.; van den Born, M.; Cozijnsen, M.; Haegebarth, A.; Korving, J.; Begthel, H.; Peters, P.J.; et al. Identification of stem cells in small intestine and colon by marker gene Lgr5. *Nature* **2007**, *449*, 1003-1007, doi:10.1038/nature06196.

BIBLIOGRAPHY

75. Clayton, E.; Doupe, D.P.; Klein, A.M.; Winton, D.J.; Simons, B.D.; Jones, P.H. A single type of progenitor cell maintains normal epidermis. *Nature* **2007**, *446*, 185-189, doi:10.1038/nature05574.
76. Scaramozza, A.; Park, D.; Kollu, S.; Beerman, I.; Sun, X.; Rossi, D.J.; Lin, C.P.; Scadden, D.T.; Crist, C.; Brack, A.S. Lineage Tracing Reveals a Subset of Reserve Muscle Stem Cells Capable of Clonal Expansion under Stress. *Cell Stem Cell* **2019**, *24*, 944-957 e945, doi:10.1016/j.stem.2019.03.020.
77. Wang, L.L.; Serrano, C.; Zhong, X.; Ma, S.; Zou, Y.; Zhang, C.L. Revisiting astrocyte to neuron conversion with lineage tracing in vivo. *Cell* **2021**, *184*, 5465-5481 e5416, doi:10.1016/j.cell.2021.09.005.
78. Chapple, R.H.; Tseng, Y.J.; Hu, T.; Kitano, A.; Takeichi, M.; Hoegenauer, K.A.; Nakada, D. Lineage tracing of murine adult hematopoietic stem cells reveals active contribution to steady-state hematopoiesis. *Blood Adv* **2018**, *2*, 1220-1228, doi:10.1182/bloodadvances.2018016295.
79. Ousset, M.; Van Keymeulen, A.; Bouvencourt, G.; Sharma, N.; Achouri, Y.; Simons, B.D.; Blanpain, C. Multipotent and unipotent progenitors contribute to prostate postnatal development. *Nat Cell Biol* **2012**, *14*, 1131-1138, doi:10.1038/ncb2600.
80. Yang, Y.; Riccio, P.; Schotsaert, M.; Mori, M.; Lu, J.; Lee, D.K.; Garcia-Sastre, A.; Xu, J.; Cardoso, W.V. Spatial-Temporal Lineage Restrictions of Embryonic p63(+) Progenitors Establish Distinct Stem Cell Pools in Adult Airways. *Dev Cell* **2018**, *44*, 752-761 e754, doi:10.1016/j.devcel.2018.03.001.
81. Davis, F.M.; Lloyd-Lewis, B.; Harris, O.B.; Kozar, S.; Winton, D.J.; Muresan, L.; Watson, C.J. Single-cell lineage tracing in the mammary gland reveals stochastic clonal dispersion of stem/progenitor cell progeny. *Nat Commun* **2016**, *7*, 13053, doi:10.1038/ncomms13053.
82. Scheele, C.L.; Hannezo, E.; Muraro, M.J.; Zomer, A.; Langedijk, N.S.; van Oudenaarden, A.; Simons, B.D.; van Rheenen, J. Identity and dynamics of mammary stem cells during branching morphogenesis. *Nature* **2017**, *542*, 313-317, doi:10.1038/nature21046.
83. Prater, M.D.; Petit, V.; Alasdair Russell, I.; Giraddi, R.R.; Shehata, M.; Menon, S.; Schulte, R.; Kalajzic, I.; Rath, N.; Olson, M.F.; et al. Mammary stem cells have myoepithelial cell properties. *Nat Cell Biol* **2014**, *16*, 942-950, 941-947, doi:10.1038/ncb3025.
84. Chakrabarti, R.; Celia-Terrassa, T.; Kumar, S.; Hang, X.; Wei, Y.; Choudhury, A.; Hwang, J.; Peng, J.; Nixon, B.; Grady, J.J.; et al. Notch ligand Dll1 mediates cross-talk between mammary stem cells and the macrophageal niche. *Science* **2018**, *360*, doi:10.1126/science.aan4153.
85. de Visser, K.E.; Ciampricotti, M.; Michalak, E.M.; Tan, D.W.; Speksnijder, E.N.; Hau, C.S.; Clevers, H.; Barker, N.; Jonkers, J. Developmental stage-specific contribution of LGR5(+) cells to basal and luminal epithelial lineages in the postnatal mammary gland. *J Pathol* **2012**, *228*, 300-309, doi:10.1002/path.4096.

BIBLIOGRAPHY

86. Wang, C.; Christin, J.R.; Oktay, M.H.; Guo, W. Lineage-Biased Stem Cells Maintain Estrogen-Receptor-Positive and -Negative Mouse Mammary Luminal Lineages. *Cell Rep* **2017**, *18*, 2825-2835, doi:10.1016/j.celrep.2017.02.071.
87. Van Keymeulen, A.; Fioramonti, M.; Centonze, A.; Bouvencourt, G.; Achouri, Y.; Blanpain, C. Lineage-Restricted Mammary Stem Cells Sustain the Development, Homeostasis, and Regeneration of the Estrogen Receptor Positive Lineage. *Cell Rep* **2017**, *20*, 1525-1532, doi:10.1016/j.celrep.2017.07.066.
88. Wuidart, A.; Sifrim, A.; Fioramonti, M.; Matsumura, S.; Brisebarre, A.; Brown, D.; Centonze, A.; Dannau, A.; Dubois, C.; Van Keymeulen, A.; et al. Early lineage segregation of multipotent embryonic mammary gland progenitors. *Nat Cell Biol* **2018**, *20*, 666-676, doi:10.1038/s41556-018-0095-2.
89. Song, W.; Wang, R.; Jiang, W.; Yin, Q.; Peng, G.; Yang, R.; Yu, Q.C.; Chen, J.; Li, J.; Cheung, T.H.; et al. Hormones induce the formation of luminal-derived basal cells in the mammary gland. *Cell Res* **2019**, *29*, 206-220, doi:10.1038/s41422-018-0137-0.
90. Seldin, L.; Macara, I.G. DNA Damage Promotes Epithelial Hyperplasia and Fate Mis-specification via Fibroblast Inflammasome Activation. *Dev Cell* **2020**, *55*, 558-573 e556, doi:10.1016/j.devcel.2020.09.021.
91. Molyneux, G.; Geyer, F.C.; Magnay, F.A.; McCarthy, A.; Kendrick, H.; Natrajan, R.; Mackay, A.; Grigoriadis, A.; Tutt, A.; Ashworth, A.; et al. BRCA1 basal-like breast cancers originate from luminal epithelial progenitors and not from basal stem cells. *Cell Stem Cell* **2010**, *7*, 403-417, doi:10.1016/j.stem.2010.07.010.
92. Koren, S.; Reavie, L.; Couto, J.P.; De Silva, D.; Stadler, M.B.; Roloff, T.; Britschgi, A.; Eichlisberger, T.; Kohler, H.; Aina, O.; et al. PIK3CA(H1047R) induces multipotency and multi-lineage mammary tumours. *Nature* **2015**, *525*, 114-118, doi:10.1038/nature14669.
93. Van Keymeulen, A.; Lee, M.Y.; Ousset, M.; Brohee, S.; Rorive, S.; Giraddi, R.R.; Wuidart, A.; Bouvencourt, G.; Dubois, C.; Salmon, I.; et al. Reactivation of multipotency by oncogenic PIK3CA induces breast tumour heterogeneity. *Nature* **2015**, *525*, 119-123, doi:10.1038/nature14665.
94. Hein, S.M.; Haricharan, S.; Johnston, A.N.; Toneff, M.J.; Reddy, J.P.; Dong, J.; Bu, W.; Li, Y. Luminal epithelial cells within the mammary gland can produce basal cells upon oncogenic stress. *Oncogene* **2016**, *35*, 1461-1467, doi:10.1038/onc.2015.206.
95. Kouros-Mehr, H.; Slorach, E.M.; Sternlicht, M.D.; Werb, Z. GATA-3 maintains the differentiation of the luminal cell fate in the mammary gland. *Cell* **2006**, *127*, 1041-1055, doi:10.1016/j.cell.2006.09.048.
96. LaMarca, H.L.; Visbal, A.P.; Creighton, C.J.; Liu, H.; Zhang, Y.; Behbod, F.; Rosen, J.M. CCAAT/enhancer binding protein beta regulates stem cell activity and specifies luminal cell fate in the mammary gland. *Stem Cells* **2010**, *28*, 535-544, doi:10.1002/stem.297.
97. Tao, L.; van Bragt, M.P.; Li, Z. A Long-Lived Luminal Subpopulation Enriched with Alveolar Progenitors Serves as Cellular Origin of Heterogeneous

BIBLIOGRAPHY

- Mammary Tumors. *Stem Cell Reports* **2015**, 5, 60-74, doi:10.1016/j.stemcr.2015.05.014.
98. Phoon, Y.P.; Chivukula, I.V.; Tsoi, Y.L.; Kanatani, S.; Uhlen, P.; Kuiper, R.; Lendahl, U. Notch activation in the mouse mammary luminal lineage leads to ductal hyperplasia and altered partitioning of luminal cell subtypes. *Exp Cell Res* **2020**, 395, 112156, doi:10.1016/j.yexcr.2020.112156.
99. Radler, P.D.; Wehde, B.L.; Triplett, A.A.; Shrestha, H.; Shepherd, J.H.; Pfefferle, A.D.; Rui, H.; Cardiff, R.D.; Perou, C.M.; Wagner, K.U. Highly metastatic claudin-low mammary cancers can originate from luminal epithelial cells. *Nat Commun* **2021**, 12, 3742, doi:10.1038/s41467-021-23957-5.
100. Bray, F.; Laversanne, M.; Sung, H.; Ferlay, J.; Siegel, R.L.; Soerjomataram, I.; Jemal, A. Global cancer statistics 2022: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin* **2024**, doi:10.3322/caac.21834.
101. Wellings, S.R.; Jensen, H.M.; Marcum, R.G. An atlas of subgross pathology of the human breast with special reference to possible precancerous lesions. *J Natl Cancer Inst* **1975**, 55, 231-273.
102. Cserni, G. Histological type and typing of breast carcinomas and the WHO classification changes over time. *Pathologica* **2020**, 112, 25-41, doi:10.32074/1591-951X-1-20.
103. Makki, J. Diversity of Breast Carcinoma: Histological Subtypes and Clinical Relevance. *Clin Med Insights Pathol* **2015**, 8, 23-31, doi:10.4137/CPath.S31563.
104. Reeves, G.K.; Beral, V.; Green, J.; Gathani, T.; Bull, D.; Million Women Study, C. Hormonal therapy for menopause and breast-cancer risk by histological type: a cohort study and meta-analysis. *Lancet Oncol* **2006**, 7, 910-918, doi:10.1016/S1470-2045(06)70911-1.
105. Venkatesh, S.L.; Oseni, T.O.; Bahl, M. Symptomatic ductal carcinoma in situ (DCIS): Upstaging risk and predictors. *Clin Imaging* **2021**, 73, 101-107, doi:10.1016/j.clinimag.2020.11.050.
106. Leonard, G.D.; Swain, S.M. Ductal carcinoma in situ, complexities and challenges. *J Natl Cancer Inst* **2004**, 96, 906-920, doi:10.1093/jnci/djh164.
107. Cutuli, B.; De Lafontan, B.; Kirova, Y.; Auvray, H.; Tallet, A.; Avigdor, S.; Brunaud, C.; Delva, C. Lobular carcinoma in situ (LCIS) of the breast: is long-term outcome similar to ductal carcinoma in situ (DCIS)? Analysis of 200 cases. *Radiat Oncol* **2015**, 10, 110, doi:10.1186/s13014-015-0379-7.
108. Badowska-Kozakiewicz, A.M.; Liszcz, A.; Sobol, M.; Patera, J. Retrospective evaluation of histopathological examinations in invasive ductal breast cancer of no special type: an analysis of 691 patients. *Arch Med Sci* **2017**, 13, 1408-1415, doi:10.5114/aoms.2015.53964.
109. Conti, M.; Morciano, F.; Amodeo, S.; Gori, E.; Romanucci, G.; Belli, P.; Tommasini, O.; Fornasa, F.; Rella, R. Special Types of Breast Cancer: Clinical Behavior and Radiological Appearance. *J Imaging* **2024**, 10, doi:10.3390/jimaging10080182.

BIBLIOGRAPHY

110. Chen, X.; Zhang, C.; Guo, D.; Wang, Y.; Hu, J.; Hu, J.; Wang, S.; Liu, X. Distant metastasis and prognostic factors in patients with invasive ductal carcinoma of the breast. *Eur J Clin Invest* **2022**, *52*, e13704, doi:10.1111/eci.13704.
111. McCart Reed, A.E.; Kalinowski, L.; Simpson, P.T.; Lakhani, S.R. Invasive lobular carcinoma of the breast: the increasing importance of this special subtype. *Breast Cancer Res* **2021**, *23*, 6, doi:10.1186/s13058-020-01384-6.
112. Mathew, A.; Rajagopal, P.S.; Villgran, V.; Sandhu, G.S.; Jankowitz, R.C.; Jacob, M.; Rosenzweig, M.; Oesterreich, S.; Brufsky, A. Distinct Pattern of Metastases in Patients with Invasive Lobular Carcinoma of the Breast. *Geburtshilfe Frauenheilkd* **2017**, *77*, 660-666, doi:10.1055/s-0043-109374.
113. Pestalozzi, B.C.; Zahrieh, D.; Mallon, E.; Gusterson, B.A.; Price, K.N.; Gelber, R.D.; Holmberg, S.B.; Lindtner, J.; Snyder, R.; Thurlimann, B.; et al. Distinct clinical and prognostic features of infiltrating lobular carcinoma of the breast: combined results of 15 International Breast Cancer Study Group clinical trials. *J Clin Oncol* **2008**, *26*, 3006-3014, doi:10.1200/JCO.2007.14.9336.
114. O'Connor, D.J.; Davey, M.G.; Barkley, L.R.; Kerin, M.J. Differences in sensitivity to neoadjuvant chemotherapy among invasive lobular and ductal carcinoma of the breast and implications on surgery-A systematic review and meta-analysis. *Breast* **2022**, *61*, 1-10, doi:10.1016/j.breast.2021.11.017.
115. Budzik, M.P.; Fudalej, M.M.; Badowska-Kozakiewicz, A.M. Histopathological analysis of mucinous breast cancer subtypes and comparison with invasive carcinoma of no special type. *Sci Rep* **2021**, *11*, 5770, doi:10.1038/s41598-021-85309-z.
116. Cao, X.; Luo, Y.; Shen, S.; Ren, X. Primary mucinous cystadenocarcinoma of the breast: A case report and literature review. *Oncol Lett* **2025**, *29*, 60, doi:10.3892/ol.2024.14806.
117. Demir, S.; Sezgin, G.; Sari, A.A.; Kucukzeybek, B.B.; Yigit, S.; Etit, D.; Yazici, A.; Kucukzeybek, Y. Clinicopathological analysis of invasive cribriform carcinoma of the breast, with review of the literature. *Ann Diagn Pathol* **2021**, *54*, 151794, doi:10.1016/j.anndiagnpath.2021.151794.
118. Adrada, B.; Arribas, E.; Gilcrease, M.; Yang, W.T. Invasive micropapillary carcinoma of the breast: mammographic, sonographic, and MRI features. *AJR Am J Roentgenol* **2009**, *193*, W58-63, doi:10.2214/AJR.08.1537.
119. Jagmohan, P.; Pool, F.J.; Putti, T.C.; Wong, J. Papillary lesions of the breast: imaging findings and diagnostic challenges. *Diagn Interv Radiol* **2013**, *19*, 471-478, doi:10.5152/dir.2013.13041.
120. Cabral, A.H.; Recine, M.; Paramo, J.C.; McPhee, M.M.; Poppiti, R.; Mesko, T.W. Tubular carcinoma of the breast: an institutional experience and review of the literature. *Breast J* **2003**, *9*, 298-301, doi:10.1046/j.1524-4741.2003.09409.x.
121. Pintican, R.; Duma, M.; Chiorean, A.; Fetica, B.; Badan, M.; Bura, V.; Szep, M.; Feier, D.; Dudea, S. Mucinous versus medullary breast carcinoma: mammography, ultrasound, and MRI findings. *Clin Radiol* **2020**, *75*, 483-496, doi:10.1016/j.crad.2019.12.024.

BIBLIOGRAPHY

122. Jia, Y.; He, C.; Liu, L.; Sun, L.; Chen, Y.; Luo, Y.; Yu, T. A Retrospective Study of the Imaging and Pathological Features of Metaplastic Breast Carcinoma and Review of the Literature. *Med Sci Monit* **2019**, *25*, 248-258, doi:10.12659/MSM.912107.
123. Seo, K.J.; An, Y.Y.; Whang, I.Y.; Chang, E.D.; Kang, B.J.; Kim, S.H.; Park, C.S.; Kim, J.S.; Hong, H. Sonography of Invasive Apocrine Carcinoma of the Breast in Five Cases. *Korean J Radiol* **2015**, *16*, 1006-1011, doi:10.3348/kjr.2015.16.5.1006.
124. Park, Y.M.; Wu, Y.; Wei, W.; Yang, W.T. Primary neuroendocrine carcinoma of the breast: clinical, imaging, and histologic features. *AJR Am J Roentgenol* **2014**, *203*, W221-230, doi:10.2214/AJR.13.10749.
125. Metzger-Filho, O.; Ferreira, A.R.; Jeselsohn, R.; Barry, W.T.; Dillon, D.A.; Brock, J.E.; Vaz-Luis, I.; Hughes, M.E.; Winer, E.P.; Lin, N.U. Mixed Invasive Ductal and Lobular Carcinoma of the Breast: Prognosis and the Importance of Histologic Grade. *Oncologist* **2019**, *24*, e441-e449, doi:10.1634/theoncologist.2018-0363.
126. Rakha, E.A.; Gill, M.S.; El-Sayed, M.E.; Khan, M.M.; Hodi, Z.; Blamey, R.W.; Evans, A.J.; Lee, A.H.; Ellis, I.O. The biological and clinical characteristics of breast carcinoma with mixed ductal and lobular morphology. *Breast Cancer Res Treat* **2009**, *114*, 243-250, doi:10.1007/s10549-008-0007-4.
127. Caparlar, M.A.; Dokcu, S.; Erogu, A. Significance of immunohistochemical markers in women with breast cancer. *Niger J Clin Pract* **2023**, *26*, 314-318, doi:10.4103/njcp.njcp_252_22.
128. Loibl, S.; Andre, F.; Bachelot, T.; Barrios, C.H.; Bergh, J.; Burstein, H.J.; Cardoso, M.J.; Carey, L.A.; Dawood, S.; Del Mastro, L.; et al. Early breast cancer: ESMO Clinical Practice Guideline for diagnosis, treatment and follow-up. *Ann Oncol* **2024**, *35*, 159-182, doi:10.1016/j.annonc.2023.11.016.
129. Goldhirsch, A.; Winer, E.P.; Coates, A.S.; Gelber, R.D.; Piccart-Gebhart, M.; Thurlimann, B.; Senn, H.J.; Panel, m. Personalizing the treatment of women with early breast cancer: highlights of the St Gallen International Expert Consensus on the Primary Therapy of Early Breast Cancer 2013. *Ann Oncol* **2013**, *24*, 2206-2223, doi:10.1093/annonc/mdt303.
130. Cheang, M.C.; Chia, S.K.; Voduc, D.; Gao, D.; Leung, S.; Snider, J.; Watson, M.; Davies, S.; Bernard, P.S.; Parker, J.S.; et al. Ki67 index, HER2 status, and prognosis of patients with luminal B breast cancer. *J Natl Cancer Inst* **2009**, *101*, 736-750, doi:10.1093/jnci/djp082.
131. Hashmi, A.A.; Aijaz, S.; Khan, S.M.; Mahboob, R.; Irfan, M.; Zafar, N.I.; Nisar, M.; Siddiqui, M.; Edhi, M.M.; Faridi, N.; et al. Prognostic parameters of luminal A and luminal B intrinsic breast cancer subtypes of Pakistani patients. *World J Surg Oncol* **2018**, *16*, 1, doi:10.1186/s12957-017-1299-9.
132. Spring, L.M.; Gupta, A.; Reynolds, K.L.; Gadd, M.A.; Ellisen, L.W.; Isakoff, S.J.; Moy, B.; Bardia, A. Neoadjuvant Endocrine Therapy for Estrogen Receptor-Positive Breast Cancer: A Systematic Review and Meta-analysis. *JAMA Oncol* **2016**, *2*, 1477-1486, doi:10.1001/jamaoncol.2016.1897.

BIBLIOGRAPHY

133. Schettini, F.; Buono, G.; Cardalesi, C.; Desideri, I.; De Placido, S.; Del Mastro, L. Hormone Receptor/Human Epidermal Growth Factor Receptor 2-positive breast cancer: Where we are now and where we are going. *Cancer Treat Rev* **2016**, *46*, 20-26, doi:10.1016/j.ctrv.2016.03.012.
134. Tripathy, D.; Kaufman, P.A.; Brufsky, A.M.; Mayer, M.; Yood, M.U.; Yoo, B.; Quah, C.; Yardley, D.; Rugo, H.S. First-line treatment patterns and clinical outcomes in patients with HER2-positive and hormone receptor-positive metastatic breast cancer from registHER. *Oncologist* **2013**, *18*, 501-510, doi:10.1634/theoncologist.2012-0414.
135. Criscitiello, C.; Disalvatore, D.; De Laurentiis, M.; Gelao, L.; Fumagalli, L.; Locatelli, M.; Bagnardi, V.; Rotmensz, N.; Esposito, A.; Minchella, I.; et al. High Ki-67 score is indicative of a greater benefit from adjuvant chemotherapy when added to endocrine therapy in luminal B HER2 negative and node-positive breast cancer. *Breast* **2014**, *23*, 69-75, doi:10.1016/j.breast.2013.11.007.
136. Wu, P.; Lv, W. Concurrent neoadjuvant endocrine therapy with chemotherapy in HR+HER2- breast cancer: a systematic review and meta-analysis. *Front Endocrinol (Lausanne)* **2024**, *15*, 1254213, doi:10.3389/fendo.2024.1254213.
137. Rudas, M.; Lehnert, M.; Huynh, A.; Jakesz, R.; Singer, C.; Lax, S.; Schippinger, W.; Dietze, O.; Greil, R.; Stiglbauer, W.; et al. Cyclin D1 expression in breast cancer patients receiving adjuvant tamoxifen-based therapy. *Clin Cancer Res* **2008**, *14*, 1767-1774, doi:10.1158/1078-0432.CCR-07-4122.
138. Caldon, C.E.; Sergio, C.M.; Kang, J.; Muthukaruppan, A.; Boersma, M.N.; Stone, A.; Barraclough, J.; Lee, C.S.; Black, M.A.; Miller, L.D.; et al. Cyclin E2 overexpression is associated with endocrine resistance but not insensitivity to CDK2 inhibition in human breast cancer cells. *Mol Cancer Ther* **2012**, *11*, 1488-1499, doi:10.1158/1535-7163.MCT-11-0963.
139. Loi, S.; Haibe-Kains, B.; Majaj, S.; Lallemand, F.; Durbecq, V.; Larsimont, D.; Gonzalez-Angulo, A.M.; Pusztai, L.; Symmans, W.F.; Bardelli, A.; et al. PIK3CA mutations associated with gene signature of low mTORC1 signaling and better outcomes in estrogen receptor-positive breast cancer. *Proc Natl Acad Sci U S A* **2010**, *107*, 10208-10213, doi:10.1073/pnas.0907011107.
140. Miller, T.W.; Hennessy, B.T.; Gonzalez-Angulo, A.M.; Fox, E.M.; Mills, G.B.; Chen, H.; Higham, C.; Garcia-Echeverria, C.; Shyr, Y.; Arteaga, C.L. Hyperactivation of phosphatidylinositol-3 kinase promotes escape from hormone dependence in estrogen receptor-positive human breast cancer. *J Clin Invest* **2010**, *120*, 2406-2413, doi:10.1172/JCI41680.
141. Desai, N.V.; Torous, V.; Parker, J.; Auman, J.T.; Rosson, G.B.; Cruz, C.; Perou, C.M.; Schnitt, S.J.; Tung, N. Intrinsic molecular subtypes of breast cancers categorized as HER2-positive using an alternative chromosome 17 probe assay. *Breast Cancer Res* **2018**, *20*, 75, doi:10.1186/s13058-018-1005-z.
142. Feng, Y.; Spezia, M.; Huang, S.; Yuan, C.; Zeng, Z.; Zhang, L.; Ji, X.; Liu, W.; Huang, B.; Luo, W.; et al. Breast cancer development and progression: Risk factors, cancer stem cells, signaling pathways, genomics, and molecular pathogenesis. *Genes Dis* **2018**, *5*, 77-106, doi:10.1016/j.gendis.2018.05.001.

BIBLIOGRAPHY

143. Swain, S.M.; Shastry, M.; Hamilton, E. Targeting HER2-positive breast cancer: advances and future directions. *Nat Rev Drug Discov* **2023**, *22*, 101-126, doi:10.1038/s41573-022-00579-0.
144. Malorni, L.; Shetty, P.B.; De Angelis, C.; Hilsenbeck, S.; Rimawi, M.F.; Elledge, R.; Osborne, C.K.; De Placido, S.; Arpino, G. Clinical and biologic features of triple-negative breast cancers in a large cohort of patients with long-term follow-up. *Breast Cancer Res Treat* **2012**, *136*, 795-804, doi:10.1007/s10549-012-2315-y.
145. Greenup, R.; Buchanan, A.; Lorizio, W.; Rhoads, K.; Chan, S.; Leedom, T.; King, R.; McLennan, J.; Crawford, B.; Kelly Marcom, P.; et al. Prevalence of BRCA mutations among women with triple-negative breast cancer (TNBC) in a genetic counseling cohort. *Ann Surg Oncol* **2013**, *20*, 3254-3258, doi:10.1245/s10434-013-3205-1.
146. Dent, R.; Trudeau, M.; Pritchard, K.I.; Hanna, W.M.; Kahn, H.K.; Sawka, C.A.; Lickley, L.A.; Rawlinson, E.; Sun, P.; Narod, S.A. Triple-negative breast cancer: clinical features and patterns of recurrence. *Clin Cancer Res* **2007**, *13*, 4429-4434, doi:10.1158/1078-0432.CCR-06-3045.
147. Xiao, W.; Zheng, S.; Yang, A.; Zhang, X.; Zou, Y.; Tang, H.; Xie, X. Breast cancer subtypes and the risk of distant metastasis at initial diagnosis: a population-based study. *Cancer Manag Res* **2018**, *10*, 5329-5338, doi:10.2147/CMAR.S176763.
148. Li, H.; Han, X.; Liu, Y.; Liu, G.; Dong, G. Ki67 as a predictor of poor prognosis in patients with triple-negative breast cancer. *Oncol Lett* **2015**, *9*, 149-152, doi:10.3892/ol.2014.2618.
149. von Minckwitz, G.; Untch, M.; Blohmer, J.U.; Costa, S.D.; Eidtmann, H.; Fasching, P.A.; Gerber, B.; Eiermann, W.; Hilfrich, J.; Huober, J.; et al. Definition and impact of pathologic complete response on prognosis after neoadjuvant chemotherapy in various intrinsic breast cancer subtypes. *J Clin Oncol* **2012**, *30*, 1796-1804, doi:10.1200/JCO.2011.38.8595.
150. Sparano, J.A.; Wang, M.; Martino, S.; Jones, V.; Perez, E.A.; Saphner, T.; Wolff, A.C.; Sledge, G.W., Jr.; Wood, W.C.; Davidson, N.E. Weekly paclitaxel in the adjuvant treatment of breast cancer. *N Engl J Med* **2008**, *358*, 1663-1671, doi:10.1056/NEJMoa0707056.
151. Zhang, X.; Li, H.; Wu, F.; Sun, D.; Zhang, H.; Jin, L.; Kang, X.; Wang, Z. Neoadjuvant Chemotherapy Combined with Breast-Conserving Surgery in the Treatment of Triple-Negative Breast Cancer. *J Oncol* **2022**, *2022*, 7847889, doi:10.1155/2022/7847889.
152. Martin, M.; Ruiz Simon, A.; Ruiz Borrego, M.; Ribelles, N.; Rodriguez-Lescure, A.; Munoz-Mateu, M.; Gonzalez, S.; Margeli Vila, M.; Barnadas, A.; Ramos, M.; et al. Epirubicin Plus Cyclophosphamide Followed by Docetaxel Versus Epirubicin Plus Docetaxel Followed by Capecitabine As Adjuvant Therapy for Node-Positive Early Breast Cancer: Results From the GEICAM/2003-10 Study. *J Clin Oncol* **2015**, *33*, 3788-3795, doi:10.1200/JCO.2015.61.9510.
153. Chen, X.S.; Nie, X.Q.; Chen, C.M.; Wu, J.Y.; Wu, J.; Lu, J.S.; Shao, Z.M.; Shen, Z.Z.; Shen, K.W. Weekly paclitaxel plus carboplatin is an effective

BIBLIOGRAPHY

- nonanthracycline-containing regimen as neoadjuvant chemotherapy for breast cancer. *Ann Oncol* **2010**, *21*, 961-967, doi:10.1093/annonc/mdq041.
154. Tutt, A.N.J.; Garber, J.E.; Kaufman, B.; Viale, G.; Fumagalli, D.; Rastogi, P.; Gelber, R.D.; de Azambuja, E.; Fielding, A.; Balmana, J.; et al. Adjuvant Olaparib for Patients with BRCA1- or BRCA2-Mutated Breast Cancer. *N Engl J Med* **2021**, *384*, 2394-2405, doi:10.1056/NEJMoa2105215.
155. Dieras, V.; Han, H.S.; Kaufman, B.; Wildiers, H.; Friedlander, M.; Ayoub, J.P.; Puhalla, S.L.; Bondarenko, I.; Campone, M.; Jakobsen, E.H.; et al. Veliparib with carboplatin and paclitaxel in BRCA-mutated advanced breast cancer (BROCADE3): a randomised, double-blind, placebo-controlled, phase 3 trial. *Lancet Oncol* **2020**, *21*, 1269-1282, doi:10.1016/S1470-2045(20)30447-2.
156. Schmid, P.; Chui, S.Y.; Emens, L.A. Atezolizumab and Nab-Paclitaxel in Advanced Triple-Negative Breast Cancer. Reply. *N Engl J Med* **2019**, *380*, 987-988, doi:10.1056/NEJMc1900150.
157. Cortes, J.; Rugo, H.S.; Cescon, D.W.; Im, S.A.; Yusof, M.M.; Gallardo, C.; Lipatov, O.; Barrios, C.H.; Perez-Garcia, J.; Iwata, H.; et al. Pembrolizumab plus Chemotherapy in Advanced Triple-Negative Breast Cancer. *N Engl J Med* **2022**, *387*, 217-226, doi:10.1056/NEJMoa2202809.
158. Perou, C.M.; Sorlie, T.; Eisen, M.B.; van de Rijn, M.; Jeffrey, S.S.; Rees, C.A.; Pollack, J.R.; Ross, D.T.; Johnsen, H.; Akslen, L.A.; et al. Molecular portraits of human breast tumours. *Nature* **2000**, *406*, 747-752, doi:10.1038/35021093.
159. Sorlie, T.; Perou, C.M.; Tibshirani, R.; Aas, T.; Geisler, S.; Johnsen, H.; Hastie, T.; Eisen, M.B.; van de Rijn, M.; Jeffrey, S.S.; et al. Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. *Proc Natl Acad Sci U S A* **2001**, *98*, 10869-10874, doi:10.1073/pnas.191367098.
160. Babyshkina, N.; Vtorushin, S.; Dronova, T.; Patalyak, S.; Slonimskaya, E.; Kzhyshkowska, J.; Cherdyntseva, N.; Choynzonov, E. Impact of estrogen receptor alpha on the tamoxifen response and prognosis in luminal-A-like and luminal-B-like breast cancer. *Clin Exp Med* **2019**, *19*, 547-556, doi:10.1007/s10238-019-00583-6.
161. Loi, S.; Sotiriou, C.; Haibe-Kains, B.; Lallemand, F.; Conus, N.M.; Piccart, M.J.; Speed, T.P.; McArthur, G.A. Gene expression profiling identifies activated growth factor signaling in poor prognosis (Luminal-B) estrogen receptor positive breast cancer. *BMC Med Genomics* **2009**, *2*, 37, doi:10.1186/1755-8794-2-37.
162. Loi, S.; Haibe-Kains, B.; Desmedt, C.; Lallemand, F.; Tutt, A.M.; Gillet, C.; Ellis, P.; Harris, A.; Bergh, J.; Foekens, J.A.; et al. Definition of clinically distinct molecular subtypes in estrogen receptor-positive breast carcinomas through genomic grade. *J Clin Oncol* **2007**, *25*, 1239-1246, doi:10.1200/JCO.2006.07.1522.
163. Godoy-Ortiz, A.; Sanchez-Munoz, A.; Chica Parrado, M.R.; Alvarez, M.; Ribelles, N.; Rueda Dominguez, A.; Alba, E. Deciphering HER2 Breast Cancer Disease: Biological and Clinical Implications. *Front Oncol* **2019**, *9*, 1124, doi:10.3389/fonc.2019.01124.

BIBLIOGRAPHY

164. Veeraraghavan, J.; De Angelis, C.; Mao, R.; Wang, T.; Herrera, S.; Pavlick, A.C.; Contreras, A.; Nuciforo, P.; Mayer, I.A.; Forero, A.; et al. A combinatorial biomarker predicts pathologic complete response to neoadjuvant lapatinib and trastuzumab without chemotherapy in patients with HER2+ breast cancer. *Ann Oncol* **2019**, *30*, 927-933, doi:10.1093/annonc/mdz076.
165. Schettini, F.; Pascual, T.; Conte, B.; Chic, N.; Braso-Maristany, F.; Galvan, P.; Martinez, O.; Adamo, B.; Vidal, M.; Munoz, M.; et al. HER2-enriched subtype and pathological complete response in HER2-positive breast cancer: A systematic review and meta-analysis. *Cancer Treat Rev* **2020**, *84*, 101965, doi:10.1016/j.ctrv.2020.101965.
166. Holmes, F.A.; Espina, V.; Liotta, L.A.; Nagarwala, Y.M.; Danso, M.; McIntyre, K.J.; Osborne, C.R.; Anderson, T.; Krekow, L.; Blum, J.L.; et al. Pathologic complete response after preoperative anti-HER2 therapy correlates with alterations in PTEN, FOXO, phosphorylated Stat5, and autophagy protein signaling. *BMC Res Notes* **2013**, *6*, 507, doi:10.1186/1756-0500-6-507.
167. Prat, A.; Pascual, T.; De Angelis, C.; Gutierrez, C.; Llombart-Cussac, A.; Wang, T.; Cortes, J.; Rexer, B.; Pare, L.; Forero, A.; et al. HER2-Enriched Subtype and ERBB2 Expression in HER2-Positive Breast Cancer Treated with Dual HER2 Blockade. *J Natl Cancer Inst* **2020**, *112*, 46-54, doi:10.1093/jnci/djz042.
168. Denkert, C.; Huober, J.; Loibl, S.; Prinzler, J.; Kronenwett, R.; Darb-Esfahani, S.; Brase, J.C.; Solbach, C.; Mehta, K.; Fasching, P.A.; et al. HER2 and ESR1 mRNA expression levels and response to neoadjuvant trastuzumab plus chemotherapy in patients with primary breast cancer. *Breast Cancer Res* **2013**, *15*, R11, doi:10.1186/bcr3384.
169. Cancer Genome Atlas, N. Comprehensive molecular portraits of human breast tumours. *Nature* **2012**, *490*, 61-70, doi:10.1038/nature11412.
170. Valentin, M.D.; da Silva, S.D.; Privat, M.; Alaoui-Jamali, M.; Bignon, Y.J. Molecular insights on basal-like breast cancer. *Breast Cancer Res Treat* **2012**, *134*, 21-30, doi:10.1007/s10549-011-1934-z.
171. Parker, J.S.; Mullins, M.; Cheang, M.C.; Leung, S.; Voduc, D.; Vickery, T.; Davies, S.; Fauron, C.; He, X.; Hu, Z.; et al. Supervised risk predictor of breast cancer based on intrinsic subtypes. *J Clin Oncol* **2009**, *27*, 1160-1167, doi:10.1200/JCO.2008.18.1370.
172. Prat, A.; Parker, J.S.; Fan, C.; Cheang, M.C.U.; Miller, L.D.; Bergh, J.; Chia, S.K.L.; Bernard, P.S.; Nielsen, T.O.; Ellis, M.J.; et al. Concordance among gene expression-based predictors for ER-positive breast cancer treated with adjuvant tamoxifen. *Ann Oncol* **2012**, *23*, 2866-2873, doi:10.1093/annonc/mds080.
173. Wallden, B.; Storhoff, J.; Nielsen, T.; Dowidar, N.; Schaper, C.; Ferree, S.; Liu, S.; Leung, S.; Geiss, G.; Snider, J.; et al. Development and verification of the PAM50-based Prosigna breast cancer gene signature assay. *BMC Med Genomics* **2015**, *8*, 54, doi:10.1186/s12920-015-0129-6.
174. Prat, A.; Parker, J.S.; Karginova, O.; Fan, C.; Livasy, C.; Herschkowitz, J.I.; He, X.; Perou, C.M. Phenotypic and molecular characterization of the claudin-low

BIBLIOGRAPHY

- intrinsic subtype of breast cancer. *Breast Cancer Res* **2010**, *12*, R68, doi:10.1186/bcr2635.
175. Dias, K.; Dvorkin-Gheva, A.; Hallett, R.M.; Wu, Y.; Hassell, J.; Pond, G.R.; Levine, M.; Whelan, T.; Bane, A.L. Claudin-Low Breast Cancer; Clinical & Pathological Characteristics. *PLoS One* **2017**, *12*, e0168669, doi:10.1371/journal.pone.0168669.
176. Sabatier, R.; Finetti, P.; Guille, A.; Adelaide, J.; Chaffanet, M.; Viens, P.; Birnbaum, D.; Bertucci, F. Claudin-low breast cancers: clinical, pathological, molecular and prognostic characterization. *Mol Cancer* **2014**, *13*, 228, doi:10.1186/1476-4598-13-228.
177. Pan, C.; Xu, A.; Ma, X.; Yao, Y.; Zhao, Y.; Wang, C.; Chen, C. Research progress of Claudin-low breast cancer. *Front Oncol* **2023**, *13*, 1226118, doi:10.3389/fonc.2023.1226118.
178. Fougner, C.; Bergholtz, H.; Norum, J.H.; Sorlie, T. Re-definition of claudin-low as a breast cancer phenotype. *Nat Commun* **2020**, *11*, 1787, doi:10.1038/s41467-020-15574-5.
179. Paik, S.; Shak, S.; Tang, G.; Kim, C.; Baker, J.; Cronin, M.; Baehner, F.L.; Walker, M.G.; Watson, D.; Park, T.; et al. A multigene assay to predict recurrence of tamoxifen-treated, node-negative breast cancer. *N Engl J Med* **2004**, *351*, 2817-2826, doi:10.1056/NEJMoa041588.
180. van 't Veer, L.J.; Dai, H.; van de Vijver, M.J.; He, Y.D.; Hart, A.A.; Mao, M.; Peterse, H.L.; van der Kooy, K.; Marton, M.J.; Witteveen, A.T.; et al. Gene expression profiling predicts clinical outcome of breast cancer. *Nature* **2002**, *415*, 530-536, doi:10.1038/415530a.
181. Prat, A.; Perou, C.M. Deconstructing the molecular portraits of breast cancer. *Mol Oncol* **2011**, *5*, 5-23, doi:10.1016/j.molonc.2010.11.003.
182. Arpino, G.; Weiss, H.; Lee, A.V.; Schiff, R.; De Placido, S.; Osborne, C.K.; Elledge, R.M. Estrogen receptor-positive, progesterone receptor-negative breast cancer: association with growth factor receptor expression and tamoxifen resistance. *J Natl Cancer Inst* **2005**, *97*, 1254-1261, doi:10.1093/jnci/dji249.
183. Thakkar, J.P.; Mehta, D.G. A review of an unfavorable subset of breast cancer: estrogen receptor positive progesterone receptor negative. *Oncologist* **2011**, *16*, 276-285, doi:10.1634/theoncologist.2010-0302.
184. Cui, X.; Schiff, R.; Arpino, G.; Osborne, C.K.; Lee, A.V. Biology of progesterone receptor loss in breast cancer and its implications for endocrine therapy. *J Clin Oncol* **2005**, *23*, 7721-7735, doi:10.1200/JCO.2005.09.004.
185. Lopez, G.; Costanza, J.; Colleoni, M.; Fontana, L.; Ferrero, S.; Miozzo, M.; Fusco, N. Molecular Insights into the Classification of Luminal Breast Cancers: The Genomic Heterogeneity of Progesterone-Negative Tumors. *Int J Mol Sci* **2019**, *20*, doi:10.3390/ijms20030510.
186. Prat, A.; Carey, L.A.; Adamo, B.; Vidal, M.; Tabernero, J.; Cortes, J.; Parker, J.S.; Perou, C.M.; Baselga, J. Molecular features and survival outcomes of the

BIBLIOGRAPHY

- intrinsic subtypes within HER2-positive breast cancer. *J Natl Cancer Inst* **2014**, *106*, doi:10.1093/jnci/dju152.
187. Prat, A.; Bianchini, G.; Thomas, M.; Belousov, A.; Cheang, M.C.; Koehler, A.; Gomez, P.; Semiglazov, V.; Eiermann, W.; Tjulandin, S.; et al. Research-based PAM50 subtype predictor identifies higher responses and improved survival outcomes in HER2-positive breast cancer in the NOAH study. *Clin Cancer Res* **2014**, *20*, 511-521, doi:10.1158/1078-0432.CCR-13-0239.
188. Baselga, J.; Bradbury, I.; Eidtmann, H.; Di Cosimo, S.; de Azambuja, E.; Aura, C.; Gomez, H.; Dinh, P.; Fauria, K.; Van Dooren, V.; et al. Lapatinib with trastuzumab for HER2-positive early breast cancer (NeoALTTO): a randomised, open-label, multicentre, phase 3 trial. *Lancet* **2012**, *379*, 633-640, doi:10.1016/S0140-6736(11)61847-3.
189. Yam, C.; Mani, S.A.; Moulder, S.L. Targeting the Molecular Subtypes of Triple Negative Breast Cancer: Understanding the Diversity to Progress the Field. *Oncologist* **2017**, *22*, 1086-1093, doi:10.1634/theoncologist.2017-0095.
190. Treeck, O.; Schuler-Toprak, S.; Ortmann, O. Estrogen Actions in Triple-Negative Breast Cancer. *Cells* **2020**, *9*, doi:10.3390/cells9112358.
191. Nielsen, T.O.; Hsu, F.D.; Jensen, K.; Cheang, M.; Karaca, G.; Hu, Z.; Hernandez-Boussard, T.; Livasy, C.; Cowan, D.; Dressler, L.; et al. Immunohistochemical and clinical characterization of the basal-like subtype of invasive breast carcinoma. *Clin Cancer Res* **2004**, *10*, 5367-5374, doi:10.1158/1078-0432.CCR-04-0220.
192. Rakha, E.A.; El-Sayed, M.E.; Green, A.R.; Lee, A.H.; Robertson, J.F.; Ellis, I.O. Prognostic markers in triple-negative breast cancer. *Cancer* **2007**, *109*, 25-32, doi:10.1002/cncr.22381.
193. Dogra, A.; Mehta, A.; Doval, D.C. Are Basal-Like and Non-Basal-Like Triple-Negative Breast Cancers Really Different? *J Oncol* **2020**, *2020*, 4061063, doi:10.1155/2020/4061063.
194. Thike, A.A.; Cheok, P.Y.; Jara-Lazaro, A.R.; Tan, B.; Tan, P.; Tan, P.H. Triple-negative breast cancer: clinicopathological characteristics and relationship with basal-like breast cancer. *Mod Pathol* **2010**, *23*, 123-133, doi:10.1038/modpathol.2009.145.
195. Tischkowitz, M.; Brunet, J.S.; Begin, L.R.; Huntsman, D.G.; Cheang, M.C.; Akslen, L.A.; Nielsen, T.O.; Foulkes, W.D. Use of immunohistochemical markers can refine prognosis in triple negative breast cancer. *BMC Cancer* **2007**, *7*, 134, doi:10.1186/1471-2407-7-134.
196. Lehmann, B.D.; Jovanovic, B.; Chen, X.; Estrada, M.V.; Johnson, K.N.; Shyr, Y.; Moses, H.L.; Sanders, M.E.; Pietenpol, J.A. Refinement of Triple-Negative Breast Cancer Molecular Subtypes: Implications for Neoadjuvant Chemotherapy Selection. *PLoS One* **2016**, *11*, e0157368, doi:10.1371/journal.pone.0157368.
197. Lehmann, B.D.; Bauer, J.A.; Chen, X.; Sanders, M.E.; Chakravarthy, A.B.; Shyr, Y.; Pietenpol, J.A. Identification of human triple-negative breast cancer subtypes and preclinical models for selection of targeted therapies. *J Clin Invest* **2011**, *121*, 2750-2767, doi:10.1172/JCI45014.

BIBLIOGRAPHY

198. Holen, I.; Speirs, V.; Morrissey, B.; Blyth, K. In vivo models in breast cancer research: progress, challenges and future directions. *Dis Model Mech* **2017**, *10*, 359-371, doi:10.1242/dmm.028274.
199. Holliday, D.L.; Speirs, V. Choosing the right cell line for breast cancer research. *Breast Cancer Res* **2011**, *13*, 215, doi:10.1186/bcr2889.
200. Whittle, J.R.; Lewis, M.T.; Lindeman, G.J.; Visvader, J.E. Patient-derived xenograft models of breast cancer and their predictive power. *Breast Cancer Res* **2015**, *17*, 17, doi:10.1186/s13058-015-0523-1.
201. Marangoni, E.; Vincent-Salomon, A.; Auger, N.; Degeorges, A.; Assayag, F.; de Cremoux, P.; de Plater, L.; Guyader, C.; De Pinieux, G.; Judde, J.G.; et al. A new model of patient tumor-derived breast cancer xenografts for preclinical assays. *Clin Cancer Res* **2007**, *13*, 3989-3998, doi:10.1158/1078-0432.CCR-07-0078.
202. DeRose, Y.S.; Wang, G.; Lin, Y.C.; Bernard, P.S.; Buys, S.S.; Ebbert, M.T.; Factor, R.; Matsen, C.; Milash, B.A.; Nelson, E.; et al. Tumor grafts derived from women with breast cancer authentically reflect tumor pathology, growth, metastasis and disease outcomes. *Nat Med* **2011**, *17*, 1514-1520, doi:10.1038/nm.2454.
203. Tentler, J.J.; Tan, A.C.; Weekes, C.D.; Jimeno, A.; Leong, S.; Pitts, T.M.; Arcaroli, J.J.; Messersmith, W.A.; Eckhardt, S.G. Patient-derived tumour xenografts as models for oncology drug development. *Nat Rev Clin Oncol* **2012**, *9*, 338-350, doi:10.1038/nrclinonc.2012.61.
204. Hidalgo, M.; Amant, F.; Biankin, A.V.; Budinska, E.; Byrne, A.T.; Caldas, C.; Clarke, R.B.; de Jong, S.; Jonkers, J.; Maelandsmo, G.M.; et al. Patient-derived xenograft models: an emerging platform for translational cancer research. *Cancer Discov* **2014**, *4*, 998-1013, doi:10.1158/2159-8290.CD-14-0001.
205. Cottu, P.; Marangoni, E.; Assayag, F.; de Cremoux, P.; Vincent-Salomon, A.; Guyader, C.; de Plater, L.; Elbaz, C.; Karboul, N.; Fontaine, J.J.; et al. Modeling of response to endocrine therapy in a panel of human luminal breast cancer xenografts. *Breast Cancer Res Treat* **2012**, *133*, 595-606, doi:10.1007/s10549-011-1815-5.
206. Usary, J.; Darr, D.B.; Pfefferle, A.D.; Perou, C.M. Overview of Genetically Engineered Mouse Models of Distinct Breast Cancer Subtypes. *Curr Protoc Pharmacol* **2016**, *72*, 14 38 11-14 38 11, doi:10.1002/0471141755.ph1438s72.
207. Tsukamoto, A.S.; Grosschedl, R.; Guzman, R.C.; Parslow, T.; Varmus, H.E. Expression of the int-1 gene in transgenic mice is associated with mammary gland hyperplasia and adenocarcinomas in male and female mice. *Cell* **1988**, *55*, 619-625, doi:10.1016/0092-8674(88)90220-6.
208. Sinn, E.; Muller, W.; Pattengale, P.; Tepler, I.; Wallace, R.; Leder, P. Coexpression of MMTV/v-Ha-ras and MMTV/c-myc genes in transgenic mice: synergistic action of oncogenes in vivo. *Cell* **1987**, *49*, 465-475, doi:10.1016/0092-8674(87)90449-1.
209. Torres-Arzayus, M.I.; Font de Mora, J.; Yuan, J.; Vazquez, F.; Bronson, R.; Rue, M.; Sellers, W.R.; Brown, M. High tumor incidence and activation of the

BIBLIOGRAPHY

- PI3K/AKT pathway in transgenic mice define AIB1 as an oncogene. *Cancer Cell* **2004**, 6, 263-274, doi:10.1016/j.ccr.2004.06.027.
210. Xu, X.; Wagner, K.U.; Larson, D.; Weaver, Z.; Li, C.; Ried, T.; Hennighausen, L.; Wynshaw-Boris, A.; Deng, C.X. Conditional mutation of Brca1 in mammary epithelial cells results in blunted ductal morphogenesis and tumour formation. *Nat Genet* **1999**, 22, 37-43, doi:10.1038/8743.
 211. Guy, C.T.; Cardiff, R.D.; Muller, W.J. Induction of mammary tumors by expression of polyomavirus middle T oncogene: a transgenic mouse model for metastatic disease. *Mol Cell Biol* **1992**, 12, 954-961, doi:10.1128/mcb.12.3.954-961.1992.
 212. Guy, C.T.; Webster, M.A.; Schaller, M.; Parsons, T.J.; Cardiff, R.D.; Muller, W.J. Expression of the neu protooncogene in the mammary epithelium of transgenic mice induces metastatic disease. *Proc Natl Acad Sci U S A* **1992**, 89, 10578-10582, doi:10.1073/pnas.89.22.10578.
 213. Muller, W.J.; Lee, F.S.; Dickson, C.; Peters, G.; Pattengale, P.; Leder, P. The int-2 gene product acts as an epithelial growth factor in transgenic mice. *EMBO J* **1990**, 9, 907-913.
 214. Liu, S.; Umezū-Goto, M.; Murph, M.; Lu, Y.; Liu, W.; Zhang, F.; Yu, S.; Stephens, L.C.; Cui, X.; Murrow, G.; et al. Expression of autotaxin and lysophosphatidic acid receptors increases mammary tumorigenesis, invasion, and metastases. *Cancer Cell* **2009**, 15, 539-550, doi:10.1016/j.ccr.2009.03.027.
 215. Maroulakou, I.G.; Anver, M.; Garrett, L.; Green, J.E. Prostate and mammary adenocarcinoma in transgenic mice carrying a rat C3(1) simian virus 40 large tumor antigen fusion gene. *Proc Natl Acad Sci U S A* **1994**, 91, 11236-11240, doi:10.1073/pnas.91.23.11236.
 216. Wagner, K.U.; Wall, R.J.; St-Onge, L.; Gruss, P.; Wynshaw-Boris, A.; Garrett, L.; Li, M.; Furth, P.A.; Hennighausen, L. Cre-mediated gene deletion in the mammary gland. *Nucleic Acids Res* **1997**, 25, 4323-4330, doi:10.1093/nar/25.21.4323.
 217. Husler, M.R.; Kotopoulis, K.A.; Sundberg, J.P.; Tennent, B.J.; Kunig, S.V.; Knowles, B.B. Lactation-induced WAP-SV40 Tag transgene expression in C57BL/6J mice leads to mammary carcinoma. *Transgenic Res* **1998**, 7, 253-263, doi:10.1023/a:1008865911660.
 218. Sandgren, E.P.; Schroeder, J.A.; Qui, T.H.; Palmiter, R.D.; Brinster, R.L.; Lee, D.C. Inhibition of mammary gland involution is associated with transforming growth factor alpha but not c-myc-induced tumorigenesis in transgenic mice. *Cancer Res* **1995**, 55, 3915-3927.
 219. Gallahan, D.; Jhappan, C.; Robinson, G.; Hennighausen, L.; Sharp, R.; Kordon, E.; Callahan, R.; Merlino, G.; Smith, G.H. Expression of a truncated Int3 gene in developing secretory mammary epithelium specifically retards lobular differentiation resulting in tumorigenesis. *Cancer Res* **1996**, 56, 1775-1785.
 220. Simin, K.; Wu, H.; Lu, L.; Pinkel, D.; Albertson, D.; Cardiff, R.D.; Van Dyke, T. pRb inactivation in mammary cells reveals common mechanisms for tumor

BIBLIOGRAPHY

- initiation and progression in divergent epithelia. *PLoS Biol* **2004**, 2, E22, doi:10.1371/journal.pbio.0020022.
221. Li, Z.; Tognon, C.E.; Godinho, F.J.; Yasaitis, L.; Hock, H.; Herschkowitz, J.I.; Lannon, C.L.; Cho, E.; Kim, S.J.; Bronson, R.T.; et al. ETV6-NTRK3 fusion oncogene initiates breast cancer from committed mammary progenitors via activation of AP1 complex. *Cancer Cell* **2007**, 12, 542-558, doi:10.1016/j.ccr.2007.11.012.
222. Herschkowitz, J.I.; Simin, K.; Weigman, V.J.; Mikaelian, I.; Usary, J.; Hu, Z.; Rasmussen, K.E.; Jones, L.P.; Assefnia, S.; Chandrasekharan, S.; et al. Identification of conserved gene expression features between murine mammary carcinoma models and human breast tumors. *Genome Biol* **2007**, 8, R76, doi:10.1186/gb-2007-8-5-r76.
223. Pfefferle, A.D.; Herschkowitz, J.I.; Usary, J.; Harrell, J.C.; Spike, B.T.; Adams, J.R.; Torres-Arzuayus, M.I.; Brown, M.; Egan, S.E.; Wahl, G.M.; et al. Transcriptomic classification of genetically engineered mouse models of breast cancer identifies human subtype counterparts. *Genome Biol* **2013**, 14, R125, doi:10.1186/gb-2013-14-11-r125.
224. Jiang, Z.; Deng, T.; Jones, R.; Li, H.; Herschkowitz, J.I.; Liu, J.C.; Weigman, V.J.; Tsao, M.S.; Lane, T.F.; Perou, C.M.; et al. Rb deletion in mouse mammary progenitors induces luminal-B or basal-like/EMT tumor subtypes depending on p53 status. *J Clin Invest* **2010**, 120, 3296-3309, doi:10.1172/JCI41490.
225. Pei, X.H.; Bai, F.; Smith, M.D.; Usary, J.; Fan, C.; Pai, S.Y.; Ho, I.C.; Perou, C.M.; Xiong, Y. CDK inhibitor p18(INK4c) is a downstream target of GATA3 and restrains mammary luminal progenitor cell proliferation and tumorigenesis. *Cancer Cell* **2009**, 15, 389-401, doi:10.1016/j.ccr.2009.03.004.
226. Bultman, S.J.; Herschkowitz, J.I.; Godfrey, V.; Gebuhr, T.C.; Yaniv, M.; Perou, C.M.; Magnuson, T. Characterization of mammary tumors from Brg1 heterozygous mice. *Oncogene* **2008**, 27, 460-468, doi:10.1038/sj.onc.1210664.
227. Hu, Z.; Fan, C.; Oh, D.S.; Marron, J.S.; He, X.; Qaqish, B.F.; Livasy, C.; Carey, L.A.; Reynolds, E.; Dressler, L.; et al. The molecular portraits of breast tumors are conserved across microarray platforms. *BMC Genomics* **2006**, 7, 96, doi:10.1186/1471-2164-7-96.
228. Abba, M.C.; Zhong, Y.; Lee, J.; Kil, H.; Lu, Y.; Takata, Y.; Simper, M.S.; Gaddis, S.; Shen, J.; Aldaz, C.M. DMBA induced mouse mammary tumors display high incidence of activating Pik3caH1047 and loss of function Pten mutations. *Oncotarget* **2016**, 7, 64289-64299, doi:10.18632/oncotarget.11733.
229. Fougner, C.; Bergholtz, H.; Kuiper, R.; Norum, J.H.; Sorlie, T. Claudin-low-like mouse mammary tumors show distinct transcriptomic patterns uncoupled from genomic drivers. *Breast Cancer Res* **2019**, 21, 85, doi:10.1186/s13058-019-1170-8.
230. Chan, M.M.; Lu, X.; Merchant, F.M.; Iglehart, J.D.; Miron, P.L. Gene expression profiling of NMU-induced rat mammary tumors: cross species comparison with human breast cancer. *Carcinogenesis* **2005**, 26, 1343-1353, doi:10.1093/carcin/bgi100.

BIBLIOGRAPHY

231. Gullino, P.M.; Pettigrew, H.M.; Grantham, F.H. N-nitrosomethylurea as mammary gland carcinogen in rats. *J Natl Cancer Inst* **1975**, *54*, 401-414.
232. Imaoka, T.; Nishimura, M.; Iizuka, D.; Daino, K.; Takabatake, T.; Okamoto, M.; Kakinuma, S.; Shimada, Y. Radiation-induced mammary carcinogenesis in rodent models: what's different from chemical carcinogenesis? *J Radiat Res* **2009**, *50*, 281-293, doi:10.1269/jrr.09027.
233. Yu, Y.; Okayasu, R.; Weil, M.M.; Silver, A.; McCarthy, M.; Zabriskie, R.; Long, S.; Cox, R.; Ullrich, R.L. Elevated breast cancer risk in irradiated BALB/c mice associates with unique functional polymorphism of the Prkdc (DNA-dependent protein kinase catalytic subunit) gene. *Cancer Res* **2001**, *61*, 1820-1824.
234. Backlund, M.G.; Trasti, S.L.; Backlund, D.C.; Cressman, V.L.; Godfrey, V.; Koller, B.H. Impact of ionizing radiation and genetic background on mammary tumorigenesis in p53-deficient mice. *Cancer Res* **2001**, *61*, 6577-6582.
235. Cressman, V.L.; Backlund, D.C.; Hicks, E.M.; Gowen, L.C.; Godfrey, V.; Koller, B.H. Mammary tumor formation in p53- and BRCA1-deficient mice. *Cell Growth Differ* **1999**, *10*, 1-10.
236. Umesako, S.; Fujisawa, K.; Iiga, S.; Mori, N.; Takahashi, M.; Hong, D.P.; Song, C.W.; Haga, S.; Imai, S.; Niwa, O.; et al. Atm heterozygous deficiency enhances development of mammary carcinomas in p53 heterozygous knockout mice. *Breast Cancer Res* **2005**, *7*, R164-170, doi:10.1186/bcr968.
237. van der Houven van Oordt, C.W.; Smits, R.; Schouten, T.G.; Houwing-Duistermaat, J.J.; Williamson, S.L.; Luz, A.; Meera Khan, P.; van der Eb, A.J.; Breuer, M.L.; Fodde, R. The genetic background modifies the spontaneous and X-ray-induced tumor spectrum in the Apc1638N mouse model. *Genes Chromosomes Cancer* **1999**, *24*, 191-198, doi:10.1002/(sici)1098-2264(199903)24:3<191::aid-gcc3>3.0.co;2-l.
238. Fisher, G.H.; Orsulic, S.; Holland, E.; Hively, W.P.; Li, Y.; Lewis, B.C.; Williams, B.O.; Varmus, H.E. Development of a flexible and specific gene delivery system for production of murine tumor models. *Oncogene* **1999**, *18*, 5253-5260, doi:10.1038/sj.onc.1203087.
239. Du, Z.; Podsypanina, K.; Huang, S.; McGrath, A.; Toneff, M.J.; Bogoslovskaja, E.; Zhang, X.; Moraes, R.C.; Fluck, M.; Allred, D.C.; et al. Introduction of oncogenes into mammary glands in vivo with an avian retroviral vector initiates and promotes carcinogenesis in mouse models. *Proc Natl Acad Sci U S A* **2006**, *103*, 17396-17401, doi:10.1073/pnas.0608607103.
240. Lim, E.; Vaillant, F.; Wu, D.; Forrest, N.C.; Pal, B.; Hart, A.H.; Asselin-Labat, M.L.; Gyorki, D.E.; Ward, T.; Partanen, A.; et al. Aberrant luminal progenitors as the candidate target population for basal tumor development in BRCA1 mutation carriers. *Nat Med* **2009**, *15*, 907-913, doi:10.1038/nm.2000.
241. Nguyen, Q.H.; Pervolarakis, N.; Blake, K.; Ma, D.; Davis, R.T.; James, N.; Phung, A.T.; Willey, E.; Kumar, R.; Jabart, E.; et al. Profiling human breast epithelial cells using single cell RNA sequencing identifies cell diversity. *Nat Commun* **2018**, *9*, 2028, doi:10.1038/s41467-018-04334-1.
242. Shehata, M.; Teschendorff, A.; Sharp, G.; Novcic, N.; Russell, I.A.; Avril, S.; Prater, M.; Eirew, P.; Caldas, C.; Watson, C.J.; et al. Phenotypic and functional

BIBLIOGRAPHY

- characterisation of the luminal cell hierarchy of the mammary gland. *Breast Cancer Res* **2012**, 14, R134, doi:10.1186/bcr3334.
243. Mohamed, G.A.; Mahmood, S.; Ognjenovic, N.B.; Lee, M.K.; Wilkins, O.M.; Christensen, B.C.; Muller, K.E.; Pattabiraman, D.R. Lineage plasticity enables low-ER luminal tumors to evolve and gain basal-like traits. *Breast Cancer Res* **2023**, 25, 23, doi:10.1186/s13058-023-01621-8.
244. Rios, A.C.; Capaldo, B.D.; Vaillant, F.; Pal, B.; van Ineveld, R.; Dawson, C.A.; Chen, Y.; Nolan, E.; Fu, N.Y.; Group, D.; et al. Intracolon Plasticity in Mammary Tumors Revealed through Large-Scale Single-Cell Resolution 3D Imaging. *Cancer Cell* **2019**, 35, 618-632 e616, doi:10.1016/j.ccell.2019.02.010.
245. Alieva, M.; Ritsma, L.; Giedt, R.J.; Weissleder, R.; van Rheenen, J. Imaging windows for long-term intravital imaging: General overview and technical insights. *Intravital* **2014**, 3, e29917, doi:10.4161/intv.29917.
246. Snippert, H.J.; van der Flier, L.G.; Sato, T.; van Es, J.H.; van den Born, M.; Kroon-Veenboer, C.; Barker, N.; Klein, A.M.; van Rheenen, J.; Simons, B.D.; et al. Intestinal crypt homeostasis results from neutral competition between symmetrically dividing Lgr5 stem cells. *Cell* **2010**, 143, 134-144, doi:10.1016/j.cell.2010.09.016.
247. Ciwinska, M.; Messal, H.A.; Hristova, H.R.; Lutz, C.; Bornes, L.; Chalkiadakis, T.; Harkes, R.; Langedijk, N.S.M.; Hutten, S.J.; Menezes, R.X.; et al. Mechanisms that clear mutations drive field cancerization in mammary tissue. *Nature* **2024**, 633, 198-206, doi:10.1038/s41586-024-07882-3.
248. Serrano, A.; Berthelet, J.; Naik, S.H.; Merino, D. Mastering the use of cellular barcoding to explore cancer heterogeneity. *Nat Rev Cancer* **2022**, 22, 609-624, doi:10.1038/s41568-022-00500-2.
249. Pelliccia, S.; Franchini, M.; Viscido, G.; Arnese, R.; Gambardella, G. Single cell lineage tracing reveals clonal dynamics of anti-EGFR therapy resistance in triple negative breast cancer. *Genome Med* **2024**, 16, 55, doi:10.1186/s13073-024-01327-2.
250. Ginzel, J.D.; Acharya, C.R.; Lubkov, V.; Mori, H.; Boone, P.G.; Rochelle, L.K.; Roberts, W.L.; Everitt, J.I.; Hartman, Z.C.; Crosby, E.J.; et al. HER2 Isoforms Uniquely Program Intratumor Heterogeneity and Predetermine Breast Cancer Trajectories During the Occult Tumorigenic Phase. *Mol Cancer Res* **2021**, 19, 1699-1711, doi:10.1158/1541-7786.MCR-21-0215.
251. Kalluri, R.; Weinberg, R.A. The basics of epithelial-mesenchymal transition. *J Clin Invest* **2009**, 119, 1420-1428, doi:10.1172/JCI39104.
252. Tsai, J.H.; Donaher, J.L.; Murphy, D.A.; Chau, S.; Yang, J. Spatiotemporal regulation of epithelial-mesenchymal transition is essential for squamous cell carcinoma metastasis. *Cancer Cell* **2012**, 22, 725-736, doi:10.1016/j.ccr.2012.09.022.
253. Fischer, K.R.; Durrans, A.; Lee, S.; Sheng, J.; Li, F.; Wong, S.T.; Choi, H.; El Rayes, T.; Ryu, S.; Troeger, J.; et al. Epithelial-to-mesenchymal transition is not required for lung metastasis but contributes to chemoresistance. *Nature* **2015**, 527, 472-476, doi:10.1038/nature15748.

BIBLIOGRAPHY

254. Bornes, L.; van Scheppingen, R.H.; Beerling, E.; Schelfhorst, T.; Ellenbroek, S.I.J.; Seinstra, D.; van Rheenen, J. Fsp1-Mediated Lineage Tracing Fails to Detect the Majority of Disseminating Cells Undergoing EMT. *Cell Rep* **2019**, *29*, 2565-2569 e2563, doi:10.1016/j.celrep.2019.10.107.
255. Lin, E.Y.; Jones, J.G.; Li, P.; Zhu, L.; Whitney, K.D.; Muller, W.J.; Pollard, J.W. Progression to malignancy in the polyoma middle T oncoprotein mouse breast cancer model provides a reliable model for human diseases. *Am J Pathol* **2003**, *163*, 2113-2126, doi:10.1016/S0002-9440(10)63568-7.
256. Hollern, D.P.; Swiatnicki, M.R.; Andrechek, E.R. Histological subtypes of mouse mammary tumors reveal conserved relationships to human cancers. *PLoS Genet* **2018**, *14*, e1007135, doi:10.1371/journal.pgen.1007135.
257. Christenson, J.L.; Butterfield, K.T.; Spoelstra, N.S.; Norris, J.D.; Josan, J.S.; Pollock, J.A.; McDonnell, D.P.; Katzenellenbogen, B.S.; Katzenellenbogen, J.A.; Richer, J.K. MMTV-PyMT and Derived Met-1 Mouse Mammary Tumor Cells as Models for Studying the Role of the Androgen Receptor in Triple-Negative Breast Cancer Progression. *Horm Cancer* **2017**, *8*, 69-77, doi:10.1007/s12672-017-0285-6.
258. Fluck, M.M.; Schaffhausen, B.S. Lessons in signaling and tumorigenesis from polyomavirus middle T antigen. *Microbiol Mol Biol Rev* **2009**, *73*, 542-563, Table of Contents, doi:10.1128/MMBR.00009-09.
259. Garcia, R.; Yu, C.L.; Hudnall, A.; Catlett, R.; Nelson, K.L.; Smithgall, T.; Fujita, D.J.; Ethier, S.P.; Jove, R. Constitutive activation of Stat3 in fibroblasts transformed by diverse oncoproteins and in breast carcinoma cells. *Cell Growth Differ* **1997**, *8*, 1267-1276.
260. Bromberg, J.F.; Horvath, C.M.; Besser, D.; Lathem, W.W.; Darnell, J.E., Jr. Stat3 activation is required for cellular transformation by v-src. *Mol Cell Biol* **1998**, *18*, 2553-2558, doi:10.1128/MCB.18.5.2553.
261. Gluschnaider, U.; Hertz, R.; Ohayon, S.; Smeir, E.; Smets, M.; Pikarsky, E.; Bar-Tana, J. Long-chain fatty acid analogues suppress breast tumorigenesis and progression. *Cancer Res* **2014**, *74*, 6991-7002, doi:10.1158/0008-5472.CAN-14-0385.
262. Berishaj, M.; Gao, S.P.; Ahmed, S.; Leslie, K.; Al-Ahmadie, H.; Gerald, W.L.; Bornmann, W.; Bromberg, J.F. Stat3 is tyrosine-phosphorylated through the interleukin-6/glycoprotein 130/Janus kinase pathway in breast cancer. *Breast Cancer Res* **2007**, *9*, R32, doi:10.1186/bcr1680.
263. Chang, Q.; Bournazou, E.; Sansone, P.; Berishaj, M.; Gao, S.P.; Daly, L.; Wels, J.; Theilen, T.; Granitto, S.; Zhang, X.; et al. The IL-6/JAK/Stat3 feed-forward loop drives tumorigenesis and metastasis. *Neoplasia* **2013**, *15*, 848-862, doi:10.1593/neo.13706.
264. Yu, H.; Pardoll, D.; Jove, R. STATs in cancer inflammation and immunity: a leading role for STAT3. *Nat Rev Cancer* **2009**, *9*, 798-809, doi:10.1038/nrc2734.
265. Davie, S.A.; Maglione, J.E.; Manner, C.K.; Young, D.; Cardiff, R.D.; MacLeod, C.L.; Ellies, L.G. Effects of FVB/NJ and C57Bl/6J strain backgrounds on

BIBLIOGRAPHY

- mammary tumor phenotype in inducible nitric oxide synthase deficient mice. *Transgenic Res* **2007**, *16*, 193-201, doi:10.1007/s11248-006-9056-9.
266. Green, J.E.; Shibata, M.A.; Yoshidome, K.; Liu, M.L.; Jorcyk, C.; Anver, M.R.; Wigginton, J.; Wiltout, R.; Shibata, E.; Kaczmarczyk, S.; et al. The C3(1)/SV40 T-antigen transgenic mouse model of mammary cancer: ductal epithelial cell targeting with multistage progression to carcinoma. *Oncogene* **2000**, *19*, 1020-1027, doi:10.1038/sj.onc.1203280.
267. Mietz, J.A.; Unger, T.; Huibregtse, J.M.; Howley, P.M. The transcriptional transactivation function of wild-type p53 is inhibited by SV40 large T-antigen and by HPV-16 E6 oncoprotein. *EMBO J* **1992**, *11*, 5013-5020, doi:10.1002/j.1460-2075.1992.tb05608.x.
268. Dyson, N.; Bernards, R.; Friend, S.H.; Gooding, L.R.; Hassell, J.A.; Major, E.O.; Pipas, J.M.; Vandyke, T.; Harlow, E. Large T antigens of many polyomaviruses are able to form complexes with the retinoblastoma protein. *J Virol* **1990**, *64*, 1353-1356, doi:10.1128/JVI.64.3.1353-1356.1990.
269. Liu, M.L.; Von Lintig, F.C.; Liyanage, M.; Shibata, M.A.; Jorcyk, C.L.; Ried, T.; Boss, G.R.; Green, J.E. Amplification of Ki-ras and elevation of MAP kinase activity during mammary tumor progression in C3(1)/SV40 Tag transgenic mice. *Oncogene* **1998**, *17*, 2403-2411, doi:10.1038/sj.onc.1202456.
270. Currier, N.; Solomon, S.E.; Demicco, E.G.; Chang, D.L.; Farago, M.; Ying, H.; Dominguez, I.; Sonenshein, G.E.; Cardiff, R.D.; Xiao, Z.X.; et al. Oncogenic signaling pathways activated in DMBA-induced mouse mammary tumors. *Toxicol Pathol* **2005**, *33*, 726-737, doi:10.1080/01926230500352226.
271. Denison, M.S.; Nagy, S.R. Activation of the aryl hydrocarbon receptor by structurally diverse exogenous and endogenous chemicals. *Annu Rev Pharmacol Toxicol* **2003**, *43*, 309-334, doi:10.1146/annurev.pharmtox.43.100901.135828.
272. Pereira, B.; Chin, S.F.; Rueda, O.M.; Vollan, H.K.; Provenzano, E.; Bardwell, H.A.; Pugh, M.; Jones, L.; Russell, R.; Sammut, S.J.; et al. Erratum: The somatic mutation profiles of 2,433 breast cancers refine their genomic and transcriptomic landscapes. *Nat Commun* **2016**, *7*, 11908, doi:10.1038/ncomms11908.
273. Ruzankina, Y.; Pinzon-Guzman, C.; Asare, A.; Ong, T.; Pontano, L.; Cotsarelis, G.; Zediak, V.P.; Velez, M.; Bhandoola, A.; Brown, E.J. Deletion of the developmentally essential gene ATR in adult mice leads to age-related phenotypes and stem cell loss. *Cell Stem Cell* **2007**, *1*, 113-126, doi:10.1016/j.stem.2007.03.002.
274. Wendling, O.; Bornert, J.M.; Chambon, P.; Metzger, D. Efficient temporally-controlled targeted mutagenesis in smooth muscle cells of the adult mouse. *Genesis* **2009**, *47*, 14-18, doi:10.1002/dvg.20448.
275. Fre, S.; Hannezo, E.; Sale, S.; Huyghe, M.; Lafkas, D.; Kissel, H.; Louvi, A.; Greve, J.; Louvard, D.; Artavanis-Tsakonas, S. Notch lineages and activity in intestinal stem cells determined by a new set of knock-in mice. *PLoS One* **2011**, *6*, e25785, doi:10.1371/journal.pone.0025785.

BIBLIOGRAPHY

276. Zhu, L.; Gibson, P.; Currie, D.S.; Tong, Y.; Richardson, R.J.; Bayazitov, I.T.; Poppleton, H.; Zakharenko, S.; Ellison, D.W.; Gilbertson, R.J. Prominin 1 marks intestinal stem cells that are susceptible to neoplastic transformation. *Nature* **2009**, *457*, 603-607, doi:10.1038/nature07589.
277. Muzumdar, M.D.; Tasic, B.; Miyamichi, K.; Li, L.; Luo, L. A global double-fluorescent Cre reporter mouse. *Genesis* **2007**, *45*, 593-605, doi:10.1002/dvg.20335.
278. Truett, G.E.; Heeger, P.; Mynatt, R.L.; Truett, A.A.; Walker, J.A.; Warman, M.L. Preparation of PCR-quality mouse genomic DNA with hot sodium hydroxide and tris (HotSHOT). *Biotechniques* **2000**, *29*, 52, 54, doi:10.2144/00291bm09.
279. Galaxy, C. The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2022 update. *Nucleic Acids Res* **2022**, *50*, W345-W351, doi:10.1093/nar/gkac247.
280. Batut, B.; van den Beek, M.; Doyle, M.A.; Soranzo, N. RNA-Seq Data Analysis in Galaxy. *Methods Mol Biol* **2021**, *2284*, 367-392, doi:10.1007/978-1-0716-1307-8_20.
281. Ewels, P.; Magnusson, M.; Lundin, S.; Kaller, M. MultiQC: summarize analysis results for multiple tools and samples in a single report. *Bioinformatics* **2016**, *32*, 3047-3048, doi:10.1093/bioinformatics/btw354.
282. Martin, M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.journal* **2011**, *17*, 10, doi:10.14806/ej.17.1.200.
283. Dobin, A.; Davis, C.A.; Schlesinger, F.; Drenkow, J.; Zaleski, C.; Jha, S.; Batut, P.; Chaisson, M.; Gingeras, T.R. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **2013**, *29*, 15-21, doi:10.1093/bioinformatics/bts635.
284. Liao, Y.; Smyth, G.K.; Shi, W. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* **2014**, *30*, 923-930, doi:10.1093/bioinformatics/btt656.
285. Love, M.I.; Huber, W.; Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* **2014**, *15*, 550, doi:10.1186/s13059-014-0550-8.
286. Hoenerhoff, M.J.; Shibata, M.A.; Bode, A.; Green, J.E. Pathologic progression of mammary carcinomas in a C3(1)/SV40 T/t-antigen transgenic rat model of human triple-negative and Her2-positive breast cancer. *Transgenic Res* **2011**, *20*, 247-259, doi:10.1007/s11248-010-9406-5.
287. Yadav, R.; Sharma, A.; Dahiya, D.; Bal, A.; Bhatia, A. Comparative morphology of tumour microenvironment in claudin-low and claudin-high breast cancers. *Pathol Res Pract* **2024**, *261*, 155502, doi:10.1016/j.prp.2024.155502.
288. Damonte, P.; Gregg, J.P.; Borowsky, A.D.; Keister, B.A.; Cardiff, R.D. EMT tumorigenesis in the mouse mammary gland. *Lab Invest* **2007**, *87*, 1218-1226, doi:10.1038/labinvest.3700683.
289. Moreno-Londono, A.P.; Robles-Flores, M. Functional Roles of CD133: More than Stemness Associated Factor Regulated by the Microenvironment. *Stem Cell Rev Rep* **2024**, *20*, 25-51, doi:10.1007/s12015-023-10647-6.

BIBLIOGRAPHY

290. Brugnoli, F.; Grassilli, S.; Al-Qassab, Y.; Capitani, S.; Bertagnolo, V. CD133 in Breast Cancer Cells: More than a Stem Cell Marker. *J Oncol* **2019**, 2019, 7512632, doi:10.1155/2019/7512632.
291. Wright, M.H.; Calcagno, A.M.; Salcido, C.D.; Carlson, M.D.; Ambudkar, S.V.; Varticovski, L. Brca1 breast tumors contain distinct CD44+/CD24- and CD133+ cells with cancer stem cell characteristics. *Breast Cancer Res* **2008**, 10, R10, doi:10.1186/bcr1855.
292. Croker, A.K.; Goodale, D.; Chu, J.; Postenka, C.; Hedley, B.D.; Hess, D.A.; Allan, A.L. High aldehyde dehydrogenase and expression of cancer stem cell markers selects for breast cancer cells with enhanced malignant and metastatic ability. *J Cell Mol Med* **2009**, 13, 2236-2252, doi:10.1111/j.1582-4934.2008.00455.x.
293. Meyer, M.J.; Fleming, J.M.; Lin, A.F.; Hussnain, S.A.; Ginsburg, E.; Vonderhaar, B.K. CD44posCD49fhiCD133/2hi defines xenograft-initiating cells in estrogen receptor-negative breast cancer. *Cancer Res* **2010**, 70, 4624-4633, doi:10.1158/0008-5472.CAN-09-3619.
294. Livasy, C.A.; Karaca, G.; Nanda, R.; Tretiakova, M.S.; Olopade, O.I.; Moore, D.T.; Perou, C.M. Phenotypic evaluation of the basal-like subtype of invasive breast carcinoma. *Mod Pathol* **2006**, 19, 264-271, doi:10.1038/modpathol.3800528.
295. Serrano, M.; Lin, A.W.; McCurrach, M.E.; Beach, D.; Lowe, S.W. Oncogenic ras provokes premature cell senescence associated with accumulation of p53 and p16INK4a. *Cell* **1997**, 88, 593-602, doi:10.1016/s0092-8674(00)81902-9.
296. Bazarov, A.V.; Lee, W.J.; Bazarov, I.; Bosire, M.; Hines, W.C.; Stankovich, B.; Chicas, A.; Lowe, S.W.; Yaswen, P. The specific role of pRb in p16 (INK4A) - mediated arrest of normal and malignant human breast cells. *Cell Cycle* **2012**, 11, 1008-1013, doi:10.4161/cc.11.5.19492.
297. Yaeger, R.; Mezzadra, R.; Sinopoli, J.; Bian, Y.; Marasco, M.; Kaplun, E.; Gao, Y.; Zhao, H.; Paula, A.D.C.; Zhu, Y.; et al. Molecular Characterization of Acquired Resistance to KRASG12C-EGFR Inhibition in Colorectal Cancer. *Cancer Discov* **2023**, 13, 41-55, doi:10.1158/2159-8290.CD-22-0405.
298. Bolos, V.; Mira, E.; Martinez-Poveda, B.; Luxan, G.; Canamero, M.; Martinez, A.C.; Manes, S.; de la Pompa, J.L. Notch activation stimulates migration of breast cancer cells and promotes tumor growth. *Breast Cancer Res* **2013**, 15, R54, doi:10.1186/bcr3447.
299. Shao, S.; Zhao, X.; Zhang, X.; Luo, M.; Zuo, X.; Huang, S.; Wang, Y.; Gu, S.; Zhao, X. Notch1 signaling regulates the epithelial-mesenchymal transition and invasion of breast cancer in a Slug-dependent manner. *Mol Cancer* **2015**, 14, 28, doi:10.1186/s12943-015-0295-3.
300. Liang, J.; Yao, X.; Aouad, P.; Wang, B.E.; Crocker, L.; Chaudhuri, S.; Liang, Y.; Darmanis, S.; Giltneane, J.; Moore, H.M.; et al. ERalpha dysfunction caused by ESR1 mutations and therapeutic pressure promotes lineage plasticity in ER(+) breast cancer. *Nat Cancer* **2025**, doi:10.1038/s43018-024-00898-8.