In vitro Coculture Models to Study Heterotypic Interactions in Breast Cancer Microenvironments

Patricia Casbas-Hernandez

A dissertation submitted to the faculty of The University of North Carolina at Chapel Hill in partial fulfillment of the requirements for the degree of Doctor in Philosophy in the Department of Pathology and Laboratory Medicine.

Chapel Hill 2013

Approved by:

Melissa A. Troester, Ph.D., M.P.H.

Keith D. Amos, M.D.

William B. Coleman, Ph.D.

David G. Kaufman M.D., PhD

Liza Makowski, Ph.D.

Charles M. Perou Ph.D.

ABSTRACT

PATRICIA CASBAS-HERNANDEZ: *In vitro* Coculture Models to Study Heterotypic Interactions in Breast Cancer Microenvironments
(Under the direction of Melissa A. Troester, Ph.D. M.P.H.)

Epithelial-stromal interactions are fundamental to tissue homeostasis and may alter breast cancer (BC) initiation and progression. Co-evolution of the neoplastic epithelium and the stroma implies that these compartments maintain an active dialogue with functional consequences for both parts. To study epithelial-stromal interactions and elucidate the role of stromal variation in tumor phenotypes, *in vitro* cocultures can be used. Gene expression data and cell-based assays from cocultures can identify cellular phenotypes and biomarkers with possible relevance in human studies. This research hypothesizes that stromal-epithelial interactions are altered from early in carcinogenesis (*Chapter2*) and that the different subtypes of invasive BC have distinct heterotypic interactions (*Chapter3*).

Epithelial-stroma interactions change during progression from benign disease to ductal-carcinoma in-situ (DCIS) in invasive Basal-like BC (BLBC). We cocultured fibroblasts with an isogenic series of cell lines (the MCF10 series) that represent premalignant stages of BC. The MCF10DCIS: fibroblasts cocultures have similar gene expression responses to those observed when invasive BLBC lines are cocultured with fibroblasts. Compared to MCF10A (benign) and MCF10AT1 (atypical-hyperplasia), MCF10DCIS cocultures showed high HGF secretion/activation and complimentary MET upregulation. The morphogenic development of MCF10DCIS in 3D cocultures was slowed

when the HGF-pathway was blocked. A novel HGF genomic signature was identified and was highly expressed in 86% of BLBCs. High expression of this HGF signature predicted worse overall survival among BLBC patients. These results show that HGF secretion and complementary MET overexpression occurs early in BLBC carcinogenesis with important consequences in invasive tumors.

Heterotypic interactions in cancer-adjacent tissue also differ by BC subtypes. We studied gene expression data of cancer-adjacent tissue from 158 BC patients and performed *in-vitro* cocultures. Gene expression analysis shows triple-negative BCs are associated with upregulated immune response and cytokine gene expression and Luminal BCs are associated with estrogen-response in cancer-adjacent tissues. Intrinsic tumor subtypes are reflected in the histologically normal cancer-adjacent tissue.

This research furthers our understanding on how BC interacts with surrounding tissues and how epithelial-stromal interactions play a role in BC progression.. Biomarkers derived from these studies may be helpful in earlier detection of aggressive lesions, in defining personalized surgical strategies or in predicting recurrence risk.

DEDICATION

Throughout our lives we cross paths with many people, I would like to thank all of those who have crossed mine, they have made me who I am and have made this work a reality. Thank you for everything. A special mention to breast cancer patients, their eagerness to live and overcome adversity is truly inspirational.

TABLE OF CONTENT

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

3D: three dimensional

BC: Breast Cancer

BCT: Breast Conserving Therapy

BLBC: Basal-like Breast Cancer

CI: Confidence Interval

DCIS: Ductal Carcinoma In situ

ER: Estrogen Receptor

EReS: Estrogen Response Signature

H&E: Hematoxilin and Eosin

HER2: HER2-like Breast Cancer

HGF: Hepatocyte Growth Factor

IHC: Immunohistochemistry

IPA: Ingenuity Pathway Analysis

LumA: Luminal A Breast Cancer

LumB: Luminal B Breast Cancer

N: Number

NBS: Normal Breast Study

OCT: Optical Coherence Tomography

OR: Odds Ratio

PAM: Predictive Analysis of Microarray

PR: Progesterone Receptor

PWBCS: Polish Women's Breast Cancer Study

RMF: Reduction Mammary Fibroblasts

SAM: Significant analysis of Microarray

TDLU: Terminal Ductal Lobular Units

TN: Triple-Negative

TNBC: Triple-Negative Breast Cancer

Chapter 1

EPITHELIAL-STOMAL INTERACTIONS IN BREAST CANCER

1.1. Overview

The interactions between breast epithelium and stroma are fundamental to normal tissue homeostasis and for tumor initiation and progression. Gene expression studies of in vitro coculture models demonstrate that in vitro models have relevance for tumor progression in vivo. For example, stromal gene expression has been shown to vary in association with tumor subtype in vivo, and analogous in vitro cocultures recapitulate subtype-specific biological interactions. Cocultures can be used to study cancer cell interactions with specific stromal components (e.g., immune cells, fibroblasts, endothelium) and different representative cell lines (e.g., cancer-associated versus normal-associated fibroblasts versus established, immortalized fibroblasts) can help elucidate the role of stromal variation in tumor phenotypes. Gene expression data can also be combined with cell-based assays to identify cellular phenotypes associated with gene expression changes. Coculture systems are manipulable systems that can yield important insights about cell-cell interactions and the cellular phenotypes that occur as tumor and stroma co-evolve.

1.2. The Tumor microenvironment:

The value of studying heterotypic interactions in breast cancer biology.

Breast cancer is a heterogeneous disease. Using gene expression patterns of tumors, researchers have identified six subtypes of breast cancer: Normal-like, Luminal A, Luminal B, Her2-enriched, Basal-like and Claudin-low. Each subtype is characterized by the expression of a subset of genes and has specific trends in response to treatment and overall survival [1-4]. However, in recent years it has become clear that the intrinsic characteristics of the tumor is not just a function of the epithelial cells, but also reflects interactions of epithelial cells with their microenvironment [5].

Cancer research has traditionally focused on the study of the neoplastic cells and how their intrinsic characteristics can be altered and modified to produce cell death or better respond to therapy [6]. While mutations in oncogenes and tumor suppressors cause neoplastic epithelial cells to lose many of their growth constraints, neoplastic cells do not lose their interactions with the surrounding non-malignant cells or with the extracellular architecture [7]. Instead, the interactions with cells in the microenvironment change during cancer progression and can promote or repress the tumorigenic process [8, 9]. Growth factors, cytokines and proteolytic enzymes are upregulated and secreted [10, 11], giving a histological appearance of granulation tissue similar to tissue morphology during physiological wound healing processes. The observation of histological changes in tumor adjacent tissue led Dr. Hal Dvorak to propose that tumors are "…wounds that do not heal…" [12]. More recent experimental and observational studies have expanded on these observations to further suggest that an activated stroma may be dominant in cancer progression [13, 14].

Some key evidence for the dominance of stroma comes from work identifying windows of susceptibility for breast cancer initiation and progression (for example during

pregnancy and postlactational involution). Extracellular matrix (ECM) function and composition are remodeled during pregnancy and lactation [15, 16], and these changes along with other changes in tissue cellular composition appear to contribute to increased breast cancer progression [15]. Conversely, progression can be reversed by stromal changes. Tamoxifen, a drug that primarily targets ER-positive epithelium, induces changes in mammary stroma leading to suppression of transformed phenotypes [17] and pre-malignant breast cancer cells placed on a reconstituted physiological basement membrane undergo cell growth arrest and form polarized alveolar structures as normal epithelial cells would [18]. These observations illustrate the important role of stromal response in breast cancer.

The stromal responses to a tumor can be collectively referred to as 'the tumor microenvironment'. It includes all the structures and cells that support the tumor: extracellular matrix, blood vasculature, inflammatory cells, adipocytes, myoepithelial cells and fibroblasts, all of which have been shown to contribute to cancer development [19]. However, it is important to distinguish two types of microenvironment based on location: intratumoral microenvironment and extratumoral microenvironment. Figure 1.1 shows a schematic depicting the wide variety of cells that make up intra and extratumoral microenvironments. The *intratumoral microenvironment* is what has classically been referred to as "microenvironment of the tumor" or "tumor stroma". It is physically located within the tumor mass or very directly adjacent [20]. The majority of *in vivo* studies of microenvironment have emphasized this intratumoral microenvironment as shown in Table 1.1. However, other studies have examined the *extratumoral microenvironment* or 'cancer-adjacent tissue', which extends further around the perimeter of the tumor (from millimeters to centimeters, depending on the study) and includes all the histologically benign-looking tissue that surrounds the tumor. This extratumoral tissue also provides support for and

influences tumor progression, reflecting either a tissue level response to the tumor or the baseline biological behavior of the tissue in which the tumor developed [21].

Both intratumoral and extratumoral microenvironments are related to the concept of 'field cancerization', initially defined as changes to the epithelium which are found in histologically normal tissue near the site of tumorigenesis and that could account for local recurrences [22]. In recent years the concept of field cancerization has been broadened to include stromal changes. A review of epithelial-specific field effects have been presented elsewhere [23], but in the current review, we are focusing on intratumoral and extratumoral stromal changes. We prefer the term 'microenvironment' to address these stromal-epithelial interactions and reserve the term 'field effect' for epithelium-specific changes.

Studies of intratumoral and extratumoral stroma in patient specimens have identified interesting biological associations, but it is difficult to evaluate the specific contributions of distinct cellular populations in these complex tissues. Wiseman and Werb [11] concluded a review article in 2002 with an important idea: "…if our aim is to find cures for diseases that rely on epithelial and stromal crosstalk we must increase our understanding of how these different cell types communicate with each other…". *In vitro* cell-cell communication studies can be integrated with studies from human tissue and with animal studies to better understand how heterotypic communication alters disease.

1.3. *In vitro* cocultures as models to study the microenvironment.

Monoculture studies of breast cancer cells have been the foundation for much of what we understand about molecular mechanisms and molecular signaling in cancer. Early studies showed that Basal-like and Luminal breast cancers have distinct responses to chemotherapeutics [24] and more recent studies have comprehensively profiled many established breast cancer cell lines to identify genomic models for each breast cancer

subtype [25]. Pathway focused studies in monocultures are also common. For example, Hoadley *et al*. showed that Basal-like breast cancer cell lines are more sensitive to the combination of carboplatin and cetuximab *in vitro* when compared to luminal cancer cell lines, and that EGFR-signatures have prognostic value when projected onto tumor datasets [26]. Other studies have identified p53-loss or p53-mutation associated signatures that can predict mutation status and survival *in vivo* [27, 28]. Studies of individual cell lines in monoculture have contributed to the development of new targeted therapies and are proving to have relevance *in vivo*. However, gene expression studies of monoculture experiments are not informative for microenvironment influences on progression. Coculture systems have become important in studying stromal factors.

Drastic changes occur when coculturing epithelium with different cell types. As a clear example, endometrial epithelial cells proliferate in response to estrogen only when cocultured with stromal cells, but not when they are in a monoculture [29]. Other studies have demonstrated that breast cancer cell lines in the presence of benign mammary epithelial cells have a more transformed phenotype than when grown in monoculture [9]. Thus, coculture systems can be used to better model key biological behaviors of epithelial and tumor cells advancing the complexity of the system by increments, focusing on one or a small number of particular characteristics of the tissue (e.g. fibroblast-cancer cell interactions or mechanical characteristics).

Cocultures grown in 2-dimensions on plastic are so far the most common type of culture studied by gene expression analysis. For some characteristics, 3D cocultures may be preferable, as they allow cells to organize themselves in space and to mimic tissue structures *in vitro* [30-33]. The gene expression profiles of 3D and 2D cultures of the same cell lines do show differences [34]. However, 2D cultures are easier to work with and can provide valuable genomic information. A large number of gene expression studies on 2D cocultures (Table 1.2) have been reported, demonstrating that these cocultures can

generate important insights. They can also preserve important physical characteristics. For example, fibroblast-to-myofibroblast transdifferentiation can be more easily studied on plastic (2D) than in Matrigel due to physical properties of the culture surface [35]. In both 2D and 3D cultures, there are a number of variables that play a role in determining what phenotypes are observed, including: the ratios of different cell types (using a convenient pre-specified ratio such as 1:1 vs. identifying multiple different biologically relevant ratios) [5], the number of cell types (e.g. choosing to coculture epithelial cells only with one stromal cell type or combining multiple cell types), mechanical factors (culture of cells with certain matrices or polymers to stimulate stiffness or other biophysical properties) [36] or the degree of cell contact (growing cells in direct physical contact or separating cell types on transwell cultures).

In addition to the variables that can be explicitly controlled, there are some experimental decisions that are less easily manipulated but important to consider in designing a study. For example, every cell line is unique and shows individual characteristics. To make experiments generalizable, it may be necessary to use multiple cell lines (at least 3 or more) to begin to establish reproducible trends for a given cell type (e.g. three cell lines or more may show consistency in luminal breast cancer behaviors, whereas one cell line alone cannot establish behavior of the class of luminal cell lines). Also, changes in the stromal cells over time should be considered. If using primary rather than established cell lines for stromal populations, it is important to consider that primary cells, such as fibroblasts, ultimately undergo senescence. Senescent fibroblasts create very different signaling milieus [37], so intraindividual variation in a given fibroblast line in culture (e.g. due to *in vitro* aging or passage number) should be considered when interpreting results. If a primary cell line will be used, variation in patient characteristics should be considered (i.e., due to collection of cells from different patients with different tumor subtypes, ages, and genetic or environmental exposure history) [38]. It is known that

cell lines can persistently harbor changes due to the exposure history of their donors [39]. Some studies have used hTERT immortalized cells to create a renewable source of isogenic cell lines for coculture studies [5, 40], and this has some advantages for reproducibility. On the other hand, variation may be of interest itself, such as variation that suggest differences between African American and Caucasian fibroblast lines [41]. Aligning the strengths and weaknesses of a given model system with the research question is most important given that it is typically impossible to perfectly recapitulate the complexities of the tissue.

1.3.1. Cellular phenotypes of epithelial cells in coculture:

Changes in gene expression.

Epithelial gene expression has been examined in relation to exposure to different cell types such as fibroblasts, immune cells, and even adipocytes (Table 1.2). Fibroblasts are abundant in the extratumoral and intratumoral microenvironment and play an essential role in the maintenance of normal tissue. Activation of fibroblasts to myofibroblasts, creates a sustained fibrosis and wound healing response leading to the desmoplastic reaction in advanced breast carcinomas [6, 21]. Fibroblasts also deposit the ECM necessary for cells to adhere, and their activation changes ECM and signaling to alter tumor initiation and progression [19].

Rozenchan *et al*. exposed MCF10A, a benign mammary epithelial cell line and the transformed MDA-MB-231 cell line to cancer associated fibroblasts (CAFs) and normal-tissue associated fibroblasts (NAFs) from the same patient. Through this indirect coculture they found many changes in the gene expression: MDA-MB-231 cells upregulated genes involved in the β-catenin/TCF pathway probably related to regulating cell polarity (DDX21, DICER) while MCF10A cells induced stress response (S100A9, HSP90B1, SPRR3) and

pro-survival genes when cultured with CAFs. Meanwhile, in culture with NAFs, MDA-MB-231 responded by down-regulating genes associated with glycolipid and fatty acid biosynthesis (ACSL5, AGTPAT4), potentially affecting membrane biogenesis, and MCF10A down-regulated genes critical for growth control and adhesion (DDIT4, CTNND1, PCDH1) [42]. The influence of the fibroblast on breast cancer cell gene expression has also been observed in cocultures comparing responses to (1) fibroblasts from negative and positive lymph nodes [43] and (2) fibroblasts from different anatomical sites and patients [44]. Each of these studies showed that fibroblasts from different sites and patients had distinct effects on the cancer cells with which they were cultured in both cell-based assays and gene expression. The fibroblasts from different anatomical sites (skin and lung) induce distinct proliferation effects on breast cancer cell lines and the proliferation responses could be used to segregate these cell lines on the basis of their tissue of origin [44]. The transcriptional changes induced in breast cancer cell lines when cocultured with fibroblasts from positive and negative lymph nodes had some common features. However, the fibroblasts were distinct for each breast cancer cell line, suggesting a response that is intrinsic to breast cancer subtype [43]. Likewise, in a different model system, soluble interactions between Basal and luminal cancer cells had distinct effects on fibroblast gene expression. When in a transwell coculture system Basal-like breast cancer cells induce the upregulation of genes such as IL-6, IL-8, CXCL3, TWIST and SOD2 in fibroblasts while luminal breast cancers do not [5]. These studies echo one another in demonstrating that both the fibroblasts and the cancer cells influence the character of the interaction.

The studies discussed above were conducted using transwells where cells are not in direct contact, but several gene expression studies have incorporated direct cell-cell contact with some additional technical or analytical steps. Direct cell-cell contact can create gene expression profiles that are distinct from those produced through soluble factors alone. However, methods for separating the cells may aid interpretation of the

resulting gene expression profiles. For example, cells can be transfected with a GFP reporter and grown in coculture with fibroblasts. The GFP-producing cells can then be isolated using flow cytometry, and subsequently analyzed [45]. Similarly, magnetic beads have been used to separate cells and demonstrate that tumor fibroblasts support neoplastic progression by altering the epigenome of mammary epithelial cells [46], specifically increasing hypermethylation of the CST6 gene. The authors of that study speculated that the direct cell-to-cell contact is involved in the epigenetic cascade that produces long term silencing of this gene. Others have performed a variety of cell sorting methods, ranging from use of surface markers to labeling of cells with short-lived cell tracking dyes [5]. These cell sorting methods are proving to be an important tool for deconvoluting cocultures.

Cocultures can also be deconvoluted using computational methods rather than physical cell sorting. Buess *et al.* [47]described a deconvolution method that computationally controls for cellular composition of cocultures. Using this approach, it was demonstrated that the interaction between some breast cancer cells and stromal fibroblasts induces an interferon-response signature which is correlated with survival [48]. Using the same deconvolution method, Camp *et al.* [5] recently showed that luminal and Basal-like breast cancer cells respond differently to the coculture with fibroblasts, but both show substantially altered expression relative to the monocultures. Furthermore, the direct coculture of Basal-like breast cancer cells and fibroblasts induced the expression of interleukins and chemokines such as IL-6, IL-8, CXCL3, TGF-β, TWIST and SOD1, while luminal breast cancer cell line cocultures with fibroblasts upregulated genes involved in stress response, such as the S100AB and S100A9 genes, as well as certain transcription factors (FOXP1, FOXA2). These cocultures studies raised the hypothesis that heterotypic interactions are intrinsic to breast cancer subtypes, and better understanding of cell-cell

interactions will yield important insights relative to treating and clinical course of these cancer subtypes.

Other stromal cell types (beyond fibroblasts) have been less well studied, but certainly play a critical role in tumor microenvironment. The most widely studied are endothelial, inflammatory and mesenchymal stem cells. For example, Buess *et al*. [49] have documented that endothelial cells cocultured with epithelial cells induce M-phase genes in the CD44+/CD24- epithelial cell population. This 'M-phase cell cycle gene set' consists of 70 genes such as HMGN2, CDC2, CDKN3, DICER, etc and can predicted metastasis *in vivo*. But perhaps more importantly, endothelial cocultures mirrored results with fibroblasts; gene expression studies showed complex patterns reflecting substantial variation in the abilities of normal and malignant cells to send and respond to extrinsic signals [5].

Macrophages have been evaluated for their role in tumor progression using coculture models. For example, Hagemann *et al.* [50] showed that coculture with macrophages increased tumor cell invasiveness through TNF-α dependent upregulation of matrix metalloproteinases (MMP-2,-3,-7,-9). Hou *et al.* [51] recently demonstrated that macrophages induce COX-2 expression in breast cancer cells through IL-1beta signaling. These observations gain greater importance when they are designed to confirm *in vivo*, biology, such as work following on recent studies [51] showing that tumor associated macrophages may enhance metastasis through activation of epidermal growth factor receptor signaling in neoplastic mammary epithelial cells. Continued work in cocultures with macrophages can elucidate whether these macrophage-cancer cell associations are subtype specific, as many of the markers induced in cancer cells (e.g. EGFR, COX2) are strongly associated with breast cancer subtype [52].

A common theme across stromal cell-breast cancer cocultures has been an increase in cytokines and inflammatory gene expression patterns. These results have been

observed for a variety of mesenchymal and immune cell types, therefore, it may not be surprising that similar responses have been observed in cocultures with mesenchymal stem cells (MSCs). MSC are important players in the tumor microenvironment [26, 53], as they migrate and engraft into the primary tumor site. This was compellingly demonstrated in a humanized mouse model; tibial injections of human MSCs induced increased proliferation and progression of tumor xenografts [7]. These results also demonstrate that species differences are important because the mouse mesenchymal cells in the control animals (no tibial injection) were not capable of promoting progression as strongly. Complementary *in vitro* cocultures used in this study clearly demonstrated a role of CXCL7 and IL-6 signaling in the aggressive, invasive phenotypes induced by MSCs. Other recent results also support the role of MSCs in promoting a more aggressive phenotype, showing that after direct coculture of MDA-MB-231, T47D and SK-Br3 with MSCs, the cancer cells up-regulate genes such as SNAIL, TWIST, vimentin, N-cadherin, and others. [54]. Similar observations were detected in transwell assays with SUM149 and HMEC cells [55], suggesting that many of these signals may be communicated via soluble factors, potentially including those factors identified by Liu *et al.* [54].

Cocultures have also been used to study how tumors metastasize to specific site and what molecular mechanisms are involved. A recent study by Rajski *et al*. demonstrated that cocultures of malignant breast epithelial cell lines with osteoblasts from the bone marrow increase IL-6 expression profiles; these profiles were associated with increased rates of bone metastasis *in vivo* [56]. Others have demonstrated that specific molecules increase the affinity to specific cell types in the metastatic sites, Claudin-2 increases the affinity of breast cancer cells to hepatocytes, increasing the possibility of liver metastasis [57].

An emerging area that will require additional investigation is how microRNAs modify the stromal-epithelial gene expression patterns. In a recent study, neoplastic

epithelial cells were directly cocultured with bone marrow stromal cells and microRNAs were shown to be transported via gap junctions between cancer cells and MSCs. These microRNAs led to reduced CXCL12 expression and a decreased proliferation [54]. Thus, future studies of gene expression changes in cocultured cells may find that microRNAs play an important role in controlling some of the observed gene expression profiles. The direct cell-cell transport of critical mRNA regulators suggests that the complexity of cell-cell interactions far exceeds what we have begun to understand. However, a growing database of gene expression data from coculture studies will help to advance our understanding of the unique cell-cell interactions that influence cancer progression.

While *in vitro* cocultures allow controlled investigation of signaling pathways and help to reconstruct step-by-step the complexity of cancer biology, human tumors *in vivo* are the ultimate system of interest. Thus, most studies have tested their signatures in coculture by linking the gene expression patterns with published microarray findings in patients. For example, the aforementioned 'M-phase cell cycle gene set', obtained through the coculture of endothelial cells with the stem cell portion of cancer cell lines was projected onto tumor data to demonstrate that this gene set can predict metastasis *in vivo* and patient survival [58]. Luciani *et al.* also used their *in vitro* signature of seven independent primary tumor cell line cocultures with primary fibroblasts to define two groups of patients with distinct overall survival rates [49]. Other approaches include showing that coculture derived signatures recapitulate established gene expression classes. For example, a subtype-specific fibroblast-coculture signature predicts breast cancer subtype in tumors, demonstrating that the *in vitro* signature is relevant *in vivo* [59]. These *in vivo* comparisons can also be combined with experimental data that demonstrate function, either in cell-based assays *in vitro* or in mouse models.

1.3.2. Confirming changes in cellular phenotypes:

Using cell-based assays to corroborate gene expression data with cocultures.

Because much of the research on cocultures has focused on how the stroma modulates invasive potential, cell-based assays demonstrating changes in migration and anchorage independent growth can help to establish biologic plausibility. Epithelial cells, in normal physiological conditions, are immobile, attached to a basement membrane, and bound to neighboring cells through several types of cell-junctions. They present apical versus basal polarity and these characteristics are essential for them to carry out their function *in vivo* [5]. One of the most drastic and visible changes that epithelial cells can acquire during carcinogenesis is the capacity to migrate, a hallmark of cancer [6]. Thus, migration assays *in vitro* can help assess how gene expression changes alter the capacity to migrate and invade. Common migration assays are transwell/Boyden chamber assays and scratch/wound healing assay. The *Boyden chamber assay* allows for paracrine and autocrine communication because it uses a transwell coculture system in which the cells share the same medium but cannot physically interact. The chemotactic cells are placed on the bottom well and the migrating cells are placed on the top insert. This insert has pores big enough (usually 8.0μm) to allow cells to migrate through. The transwell migration assays may or may not include an extracellular matrix (ECM) layer. If this ECM layer is present, the assay models the capacity of cells to break down ECM and invade, whereas in the absence of ECM, the migratory capacity alone is investigated. In either case, the cells migrating to the opposite side of the transwell insert are fixed, stained, and counted. The *Wound/Scratch assay* allows for paracrine, autocrine and cell-to-cell communication. Cells are seeded on the same surface in direct contact, a scratch is made when cells are nearly confluent and cells migrating into the scratch are measured overtime. By labeling one cell type with a fluorescent label, it is possible to identify which of the two cell types are closing the wound.

Many of the stromal cocultures discussed above have been evaluated for their effects on migration of cancer cells. For example, focusing on fibroblasts, Potter *et al.* [60] showed that tumor stromal cells (compared to normal stromal cells) caused greater chemotaxis of MDA-MB-231 and that this effect could be blocked by the addition of a monoclonal antibody to CCL2. MCF7s also become more migratory when cocultured with fibroblasts [60]. Fibroblast populations isolated from different distances relative to a breast tumor had distinct effects on the migratory capacity of MCF7 cells in scratch assays [61]. Similar findings have been observed for MSCs [48] and macrophages [62]. Breast adipocytes are abundant, comprising a major percentage of the extratumoral microenvironment, and have also been cocultured with breast cancer cells. Adipocytes are challenging to culture and coculture because they terminally differentiate and cannot be propagated to achieve a reproducible culture system; however, they are proving to have important implications for cancer progression. Dirat *et al*. [63] showed that the estrogen-receptor positive breast cancer cell line ZR75.1, and the estrogen-receptor negative line SUM159PT, both increased their invasive capacity after 72 hours in coculture with mature primary adipocytes.

As an epithelial cell becomes more aggressive, it becomes less dependent on extracellular matrix and basement membrane interactions for survival. After coculture with certain types of cells, benign or malignant epithelial cells can acquire or enhance their anchorage independent growth properties. There are two main types of assays to address anchorage independent growth, mammosphere and soft-agar colony formation assays. In both cases, coculture studies can be designed to evaluate paracrine, autocrine and/or cell-to-cell contacts. In the *mammosphere formation assay,* cells are cultured in suspension in a defined growth media [63]. Colonies are allowed to grow for 7-10 days and then analyzed. Only cells with anchorage-independent growth capacity will grow, so the number and size of colonies reflect acquisition of this phenotype. In *soft-agar colony formation*

assays, cells are grown in a gel-like matrix that provides more structure than a suspension culture, but the same phenotypes (colony number and size) are assessed after a period of growth, typically at least two weeks. These assays have been used to confirm anchorage-independent growth changes in coculture. For example, breast cancer fibroblasts decreased time required for MCF7s to form mammospheres, and increased the overall number of spheres relative to cocultures with normal fibroblasts. Additionally, when MDA-MB-468, a Basal-like breast cancer cell line, was cocultured with CAFs, the number of soft agar colonies were higher than when cocultured with NAFs [64]. MSCs have also been shown to induce mammosphere formation in human mammary epithelial cells (HMEC), and SUM149 but not in primary inflammatory breast cancer cells (MDA-IBC-3). These effects occurred though paracrine factors, as conditioned media from the MSCs had the same effects [65]. In 3D cultures, we can perform morphogenesis assays in which we study how proliferation, apoptosis and migratory phenotypes occur simultaneously, these assays are useful to analyze the malignant behavior of epithelial cells [32, 33, 66].

By combining the expression data suggesting a certain phenotypic trait and with cell based assays, new treatment-relevant advances are possible [67]. Genome expression data along with cell based assays can be used to test targeted perturbations (e.g. blocking cytokines or treating with growth factors such as in [5, 59]) to study how these phenotypes are regulated. Given that these studies can be done with human cells, and with careful control of cell ratio, cell physical environment, polarity, etc., these systems can provide interesting and important insights about how cancers become more invasive and aggressive through interactions with their environments.

1.4. Mouse models for comparative biology of tumor microenvironment

Given identification of novel hypotheses from *in vitro* cocultures and confirmation of the cellular phenotypes *in vitro*, a complete picture of stromal-epithelial interactions requires linkages with studies *in vivo.* As described above, public genomic data can be useful for this purpose, but mouse models have contributed to our fundamental understanding of the reciprocal signaling between stroma and epithelial compartments. Noel *et al.* [68] performed the first inoculation of cocultured fibroblasts and breast cancer cell lines with matrigel in an athymic mice model. The inoculation of these cocultures decreased the latency time and enhanced tumor growth. Both tumor growth and latency time were dependent on the number of inoculated fibroblasts in the coculture. In another classic example, it was demonstrated that when non-tumorgenic cell lines are introduced into irradiated cleared fat pads, they form tumors. Conversely, when introduced into cleared fat pads that have not been irradiated the same tumorogenic cell lines do not form tumors. This indicates that radiation induces changes in the stromal microenvironment that contribute to neoplastic progression *in vivo* [68]. More recently, Hu *et al.* [69] have shown that myoepithelial cells suppress, while fibroblasts enhance, tumor progression from DCIS to invasive cancer in a mouse xenograph model. Novel models for combining and humanizing the microenvironment have also been proposed, including a humanized mouse xenograft model into cleared fat pads [69], and an intraductal xenografts, where human cell lines can be injected alone or with stromal components [70]. An advantage of these models is that some of the innate immune responses are preserved, as is the systemic circulation and the three dimensional structure of the tissue.

Recently, it has also been established that different mouse models can be used to represent the heterogeneity of human breast cancers [71]. For example, the C3Tag mice over-express the SV40Tag transgene in distal mammary ductal epithelium and terminal ductal lobular units. This over-expression allows for a targeted inactivation of two tumor

suppressor genes: p53 and Rb, giving rise to a very predictable onset of tumors [72]. They most commonly develop tumors with features of Basal-like breast cancer. Thus, these models may be useful for studying Basal-like microenvironments. Future studies should examine how microenvironment characteristics, such as obesity or immune cell ablation influence the progression of tumors in some of these model systems, to gain a perspective on the role of microenvironment in different breast cancer subtypes. These models, when combined with coculture-based mechanistic studies, can be a powerful combination.

1.5. *In vivo* studies of breast cancer microenvironments.

In recent years, tissue-level wound and stromal responses have been more thoroughly characterized using molecular data [73, 74]. A growing body of gene microarray data support a role for stromal gene expression in breast cancer progression (Table 1.1). Finak *et al*. [75] analyzed biopsies of cancer tissue and non-affected tissue from breast cancer patients. By laser capture microdissection they separated the tumor compartment from the stromal compartment and performed microarrays to identify a prognostic gene set from tumor stroma that predicted patient surviva. Ma *et al*. [14] compared gene expression of ductal carcinoma *in situ* (DCIS)-associated stroma to stroma from individuals with invasive disease and showed that the majority of stromal alterations occur at the DCIS stage. These authors argued that invasiveness is dependent on the signals the epithelial cells receive from myoepithelial cells, fibroblasts and myofibroblasts. Allinen *et al.* isolated pure stromal cell populations from reduction mammoplasties, DCIS, and invasive breast cancer patients. Analysis of gene expression of these purified cell populations revealed widespread molecular changes in all cell types of the breast cancer stroma [76]. We and others have shown an activated wound response in the tumor microenvironment of breast cancer [73, 77]. Signatures of wound response from *in vitro* [77] or *in vivo* [73, 74] predict

breast cancer survival and relapse in independent datasets. Finally, Beck *et al.* have studied both macrophage infiltration-associated gene expression [78] and fibromatosis-associated gene expression [79] as predictors of outcome. These studies cumulatively suggest that tumor progression occurs due and is the results of the concerted action of a variety of stromal responses.

1.6. Significance

The tissue stroma is crucial for normal organ homeostasis as well as for tumor initiation and progression. Additionally, both intra and extratumoral microenvironments play essential roles in tumor biology. Thus, improved understanding of the interactions that take place between epithelial cells and stromal compartments is critical to advancing our knowledge of human cancer. *In vitro* coculture systems are controllable systems that can be used to study gene expression changes and corresponding cellular phenotypes that occur as tumor and stroma co-evolve. These systems can be used to define critical factors mediating the communication between the cell types. Although cocultures have limitations, the growing body of gene expression coculture data, when combined with observational studies, demonstrates that these models are generating important insights in the biology of breast cancer.

1.7. Figures and Tables

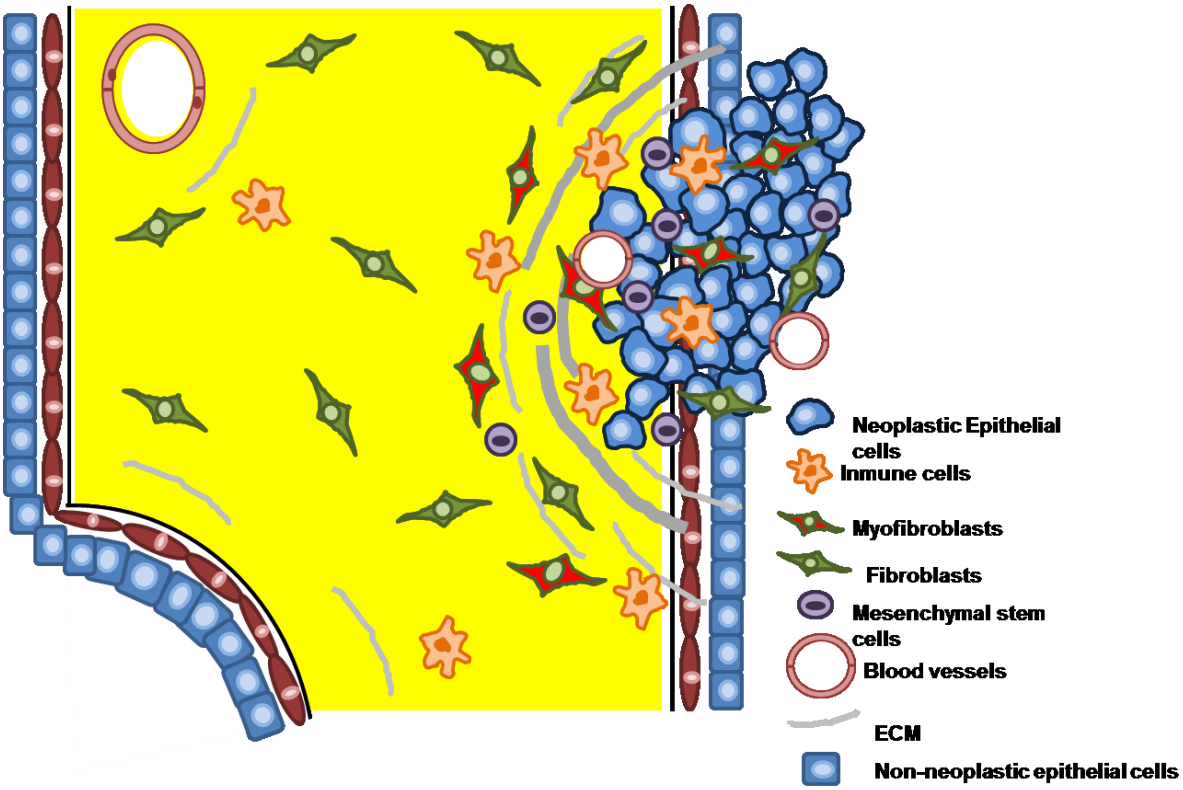


Figure 1.1. **Intra and extratumoral microenvironments and cellular components of these compartments.** Cell types present in the extratumoral and intratumoral microenvironment are similar and include fibroblasts, immune cells, endothelial cells, and mesenchymal stem cells. Abundance and signaling of these cells vary widely between and within individuals with cancer.

Table 1.1: **Whole Genome Microarrays studies to investigate breast cancer microenvironments in human tissues.**

Authors [citation]	Type of specimen studied	Processing of specimen	Type of micro-environment	Major Findings
Finak *et. al*.(2008) [75]	Fresh, frozen tissue from primary cancers (53) and adjacent non-affected tissue (31) from breast cancer patients	Laser capture microdisection of tumor stroma	Intratumoral vs Extratumoral	Stromal Derived Prognostic Predictor (SDPP),a gene set that stratifies patients by disease outcome. Genes are involved in immune response, angiogenesis and hypoxic response
Ma *et al*. (2009) [14]	Fresh frozen biopsies from disease free tissue, DCIS and invasive breast cancer (14).	Laser capture microdisection	Intratumoral	Tumor microenvironment participates in tumorogensis before tumor cells invade. Invasiveness is dependent on the signals from myoepithelial cells, fibroblasts and myofirboblasts.
Allinen *et al.* (2004) [76]	Snap frozen biopsies from reduction mammoplasties, DCIS and invasive breast cancer.	Isolation of pure cell populations by differential centrifugation	Intratumoral	Widespread genome changes in all stromal cell types. Genetic alterations only occur in epithelial cancer cells.
Troester *et al*. (2009)[73]	Snap frozen tissue from histologically normal tissue adjacent to breast cancer (47) and reduction mammoplasties (60).	Whole genome profiles of the tissue	Extratumoral	A wound response is activated in the tumor microenvironment. The wound response signature predicts cancer progression.
Chang *et al*. (2004) [77] , Chang *et al*. (2005) [74]	Isolated fibroblasts from 10 different anatomical sites & Tissue from early breast cancer patients (295)	*In vitro* response of the fibroblast populations to serum	Intratumoral Normal tissue	Identification of an *in vitro* wound response, enriched in early stage tumors High expression of this signature correlates with worse overall survival and increased distant metastasis.
Beck *et al* (2008) [79]	Desmoid fibromatosis and solitary fibrous tumors.		Intratumoral	DTF core gene set (derived mainly from fibroblasts) is a robust descriptor of stromal response that is associated with improved clinical outcome in public genomic data from breast cancer patients.
Beck *et al*. (2009) [78]	Tenosynovial giant cell tumors and pigmented villonodular synovitus		Intratumoral	The CSF1 gene expression signature (derived mainly from macrophages) is present in more aggressive cancers.
Luciani *et al*. (2011) [59]	Tissue from primary breast tumors and reduction mammoplasties	Isolation of epithelial and fibroblast cells.	Intratumoral	A 'fibroblast triggered gene expression' gene set generated by coculture of primary breast tumor cell lines and fibroblasts is enriched for inflammatory signaling, cell death and cell proliferation genes. Predicts survival in independent datasets.
Roman-Perez et al. (2012) [80]	Cancer-adjacent tissue from invasive tumors (72)		Extratumoral	There are 2 subtypes of cancer-adjacent tissue independent of tumor subtype with distinct survival patterns.

Table 1.1: **Whole genome microarray studies to investigate breast cancer microenvironments in human tissues**.

Table 1.2: **Whole genome microarray studies to investigate breast cancer tumor microenvironment *in vitro***

Authors [citation]	Cancer cell lines used	Stromal cell lines used	Type of coculture	Special separation techniques	Linked to human *in vivo* data	Major Findings
Rozenchan *et al*. (2009) [42]	MCF10A MDA-MB-231	Primary CAFs and NAFs	Transwell	No	No	Epithelial cell lines upregulate different pathways when cocultured with the two types of fibroblasts. MDA-MB-231-CAF cocultures CAFs upregulate β-catenin/TCF pathway genes; MDA-MB-231-NAF cocultures .down regulate glycolipid and fatty acid biosynthesis. MCF10A-CAF cocultures upregulate stress response genes, while MCF10A-NAF cocultures downregulate growthcontrol and adhesion genes.
Santos *et al*. (2011) [43]	MDA-MB-231, MDA-MB-435, MCF7	Primary fibroblasts from positive and negative LN	Transwell	No	No	Gene expression changes induced by coculture with fibroblasts from positive and negative nodes are distinct and intrinsic to each tumor subtype.
Camp *et al*. (2010) [5]	MCF7, T47D, ZR75, Sum102, Sum149, HCC1537	Immortalized reduction mammary fibroblasts	Direct physical contact & transwell	Yes	Computational deconvolution	The response to fibroblast coculture differs between Basal like and luminal cancer cell lines. The genes that distinguish Basal-like vs. luminal cultures also distinguishes human tumors. Basal-likes upregulate interleukins and chemokines (IL-6, IL-8, CXCL1, CXCL3, TGF-β) also TWIST and SOD1. Luminal cells increase stress response genes.
Buess *et al*. (2009)[49]	Hs578T, BT549, MDA-MB-436, MDA-MB-231, HMEC, SKBR3, MCF7, T47D, HMECs	Stromal fibroblasts: human dermal fibroblasts, embryonic lung fibroblasts, breasts stromal fibroblasts	Transwell & direct physical contact	Yes	Computational deconvolution	Interaction between some breast cancer cells and stromal fibroblasts induced interferon response. The presence of this response is associated with higher risk of tumor progression
Buess *et al*. (2009)[49]**]**	HMECs, MCF7, T47D, MDA-MB-231, SKBR-3, Hs578T, BT549	HuVECs & Human dermal microvascular endothelial cells	Direct physical coculture & transwell	Yes	Computational deconvolution	Induction of an 'M-phase cell cycle genes' in breast cancer cell lines but not in normal epithelium. Tumors with this gene signature have increased metastasis and worse overall survival. Endothelial cells induce proliferation in CD44+/CD24- cancer cells.

Liu *et al*. (2011) [54]	Sum159, Sum149, MCF7	Human bone marrow derived mesenchymal cells	Direct physical coculture & transwell	No	No	MSCs regulate cancer cell behavior through their effects on cancer stem cells. Networks of cytokines (IL-6, Il-8, CXCl1, CXCl5, CXCl6 are associated with migration of cancer cells).
Wadlow *et al*. (2009) [44]	Many commercially available cancer cell lines	Many commercially available normal skin and lung fibroblasts		GFP expression in epithelial cells		Cancer cell proliferation is modulated both by the cancer cell and the fibroblasts. Two functionally distinct pathways associated with altered proliferation were identified, one of which showed features of activated mesenchyme.
Stewart et al (2012) [81]	HCC1937, MDA-MB-468, SUM149 MCF-7, T47D, ZR-75-1	THP-1 monocytes	Transwells			Basal-like BC mediates a specific stromal immune response, implicating specific cytokines that are differentially expressed in Basal-like microenvironments

Table 1.2: **Whole genome microarray studies to investigate breast cancer tumor microenvironment *in vitro*.**

Chapter 2

ROLE OF HGF IN EPITHELIAL-STROMAL CELL INTERACTIONS DURING PROGRESSION FROM BENIGN BREAST DISEASE TO DUCTAL CARCINOMA IN SITU.

2.1. Overview

Introduction: Basal-like and Luminal breast cancers have distinct stromal-epithelial interactions, which play a role in progression to invasive cancer. However, little is known about how stromal-epithelial interactions evolve in benign and pre-invasive lesions.

Methods: To study epithelial-stroma interactions in Basal-like breast cancer progression, we cocultured reduction mammoplasty fibroblasts (RMFs) with the isogenic MCF10 series of cell lines (representing benign/normal, atypical hyperplasia, and ductal carcinoma *in situ*). We used gene expression microarrays to identify pathways induced by coculture in premalignant cells (MCF10DCIS) compared to normal and benign (MCF10A and MCF10AT1). Relevant pathways were then (1) evaluated *in vivo* for associations with Basal-like subtype and (2) targeted *in vitro* and effects on morphogenesis were evaluated.

Results: Our results show that premalignant MCF10DCIS cells express characteristic gene expression patterns of invasive Basal-like microenvironments. Furthermore, while HGF secretion is upregulated (relative to normal, MCF10A levels) when fibroblasts are cocultured with either atypical (MCF10AT cells) or premalignant (MCF10DCIS) cells, only MCF10DCIS cells upregulate the HGF receptor, MET. In 3-dimensional cultures,

upregulation of HGF/MET in MCF10DCIS cells induced morphological changes suggestive of malignant potential, and these changes were reversed by antibody-based blocking of HGF signaling. These results are relevant to *in vivo* progression because high expression of a novel MCF10DCIS-derived HGF signature was correlated with Basal-like subtype and with worse overall survival, with approximately 86% of Basal-like cancers highly expressing the HGF signature.

Conclusions: In this study we document coordinated and complementary changes in HGF and MET expression in epithelium and stroma in pre-invasive lesions. These results suggest that targeting stroma-derived HGF signaling in early carcinogenesis may block progression of Basal-like precursor lesions.

2.2. Introduction

Normal development and homeostasis requires stromal-epithelial interactions. Cancers must evolve and adapt in stromal context and therefore, cancer progression depends upon an initiated cell's ability to utilize permissive signals and circumvent repressive signals [8]. Under evolutionary theories of cancer, tumors that progress have characteristics that are advantageous given their microenvironments [82]. Cancer cells may also modify their environments to induce growth-promoting signals. Recent data suggest that host and/or stromal factors affect tumor subtype. For example, aging stroma may influence which tumor subtypes develop or may promote more aggressive disease [37, 83]. Conversely, tumor characteristics may define epithelium-stromal interactions. Basal-like breast cancers have a distinct microenvironment interaction pattern relative to other breast cancer subtypes [5] and appear to be associated with distinct immune microenvironments [6, 81, 84]. These and many other data suggest that complementary epithelial-stromal coevolution is influential in cancer development. However, since most of these studies have examined epithelial-stroma interactions after tumors have acquired

invasive characteristics, it is not well known how host-tumor interactions are maintained earlier in disease progression.

We hypothesized that Basal-like breast cancers may have unique interactions with their microenvironments beginning in the early stages of progression. In epidemiologic studies, there is evidence that Basal-like breast cancers progress very rapidly through the ductal carcinoma *in situ* stage (DCIS) compared to other cancers [85]. However, many of the DCIS-adjacent stromal tissue studies have been from patients who also have invasive cancers in the same breast [14], and given the cross-sectional nature of these studies (with data at only a single time point in the progression of disease), it is difficult to identify epithelial-stromal interactions that are induced during progression. In addition, stroma from DCIS lesions and invasive tumors are very similar, suggesting that stromal changes may occur prior to invasion [14, 76]. It is important to identify pathways that are altered in the stroma prior to invasion as these pathways may be targetable.

To study epithelial-stromal interactions in the pre-invasive phases of Basal-like breast cancer development, we employed the MCF10 cell line series in cocultures. The MCF10 cell lines represent an isogenic background (being derived from a single patient), but express pathologic characteristics in xenografts, ranging from non-neoplastic benign morphology (MCF10A) to atypical hyperplasia (MCF10AT1) to DCIS (MCF10DCIS). These lines were cocultured with fibroblasts (both two dimensional (2D) on plastic and three dimensional (3D) in Matrigel®/collagen). Cell-based assays and gene expression profiling were conducted to track the evolution of cell-cell interactions with progression. The resulting experimental data, together with patient data, suggests an important role for HGF signaling in premalignant to invasive Basal-like breast cancer.

2.3. Methods

2.3.1. Cell lines and treatments

MCF10A, MCF10AT1 and MCF10DCIS.com (referred to as MCF10DCIS) were purchased from Karmanos Cancer Institute (Detroit, MI) and Asterand (Detroit, MI). These cell lines were maintained in DMEM supplemented with 5% horse serum, 50 units/mL penicillin, and 50 units/mL streptomycin, 5 µg/mL insulin (GIBCO, Life technologies, Carlsbad, CA), 1 µg/mL hydrocortisone (Sigma-Aldrich), cholera toxin (EMD, Millipore, Darmstadt, Germany) and EGF (Invitrogen, Life technologies, Carlsbad). Cocultures were also performed in this media after ascertaining that reduction mammary fibroblasts (RMFs) maintained their RPMI 1640 doubling times in this DMEM/F12. MCF7 (Luminal cell line) and SUM149 (Basal-like cell line) were purchased from ATCC. RMFs (htert-immortalized fibroblasts from reduction mammoplasty [70]) were provided by Dr. Charlotte Kupperwasser (Tufts University). We selected an htert immortalized cell line for our experiments over primary cell lines (as other studies have done [38]), for several reasons. Primary cells have a limited life; after 9 passages they senesce allowing insufficient time to perform many assays (such as the 3D assays described below). Even prior to senescence, the aging process and patient-to-patient variation affects gene expression as shown by quantitative reverse transcriptase (RT) PCR for HGF. Figure 2.1 shows HGF RNA levels vary by up to 546-fold for a panel of 14 primary CNAF (Cancer-normal associated fibroblasts, obtained from histologically normal tissue adjacent to a tumor) and CAFs (Cancer-associated fibroblasts) at different passages. These cell lines were maintained at 37˚C and 5% CO_2 in RPMI 1640 or DMEM/F12 with L-glutamine (GIBCO) supplemented with 10% FBS (Sigma-Aldrich) and 50 units/mL penicillin and 50 units/mL streptomycin (GIBCO) as described in [5, 38]. All cell lines were tested for mycoplasma prior to use by the University of North Carolina at Chapel Hill Tissue Culture Facility, NC.

2.3.2. Primary fibroblasts

Primary fibroblasts were obtained from breast tissue of patients undergoing breast surgery for primary invasive breast carcinoma at UNC Hospital. Tissue specimens were procured under an IRB-approved protocol by the Lineberger Cancer Center Tissue Procurement Facility. Isolation protocol of these CNAFs and CAFs was previously described in [38].

2.3.3. Coculture conditions and treatments

Two types of cocultures were performed using media and culture conditions described above and previously in [5]. *Direct cocultures* are defined as a coculture where the two cell types are grown in direct physical contact, in the same well. The following epithelial:RMF ratios were plated for these cocultures: 1:4, 1:2, 1:1, 2:1 as well as the monocultures of each cell line and cells were maintained for 48 hours before RNA isolation. *Interaction cocultures* are defined as a coculture where the fibroblasts and cancer cells are separated by a porous membrane that allows cell-cell communication via soluble factors. For interaction cultures, fibroblasts were seeded in inserts on Corning Transwell plates with 0.4 μm pore polycarbonate membranes; epithelial cells were grown in the bottom well. Interaction cultures were plated at a 1:1 ratio and maintained for 48 hours before RNA isolation.

2.3.4. RNA and expression microarrays

Cells were harvested by scraping in RNA lysis buffer. Total RNA was isolated using the RNeasy mini kit (Qiagen, Valencia, CA) and RNA quality was analyzed on an Agilent 2100 Bioanalyzer using an RNA6000 nano chip. Quantification was performed on a ND-1000

Nanodrop spectrophotometer. Microarrays were performed according to Agilent protocol using two-color Agilent 4×44K V2 (Agilent G4845A). We used the Agilent Quick Amp labeling kit and protocol to synthesize Cy3-labeled reference from Stratagene Universal Human Reference spiked at 1:1,000 with MCF7 RNA and 1:1,000 with ME16C RNA to increase expression of breast cancer genes. The identical protocol was applied to total RNA from cocultured or monocultured cell lines to label these samples with Cy5. Labeled cDNAs were hybridized to arrays overnight and washed before scanning on an Agilent G2505C microarray scanner.

2.3.5. Coculture data normalization and analysis

Data from 58 microarrays (representing monocultures, direct cocultures, and indirect cocultures from 6 different cell lines) were included in this study. Microarray data are available in the Gene Expression Omnibus (GSE43467). Only those genes where more than 70% of microarrays had signal in both channels greater than 10 dpi were included. Data were Lowess normalized and missing data were imputed using *k*-nearest neighbors' imputation. For the *direct coculture* analyses, we excluded genes that did not have at least two-fold deviation from the mean in at least one sample and the method of Buess *et al*. [47] was used to normalize cocultures to appropriate monocultures performed in the same media and under identical conditions as previously described in [5]. Briefly, the Buess method is an example of an expression deconvolution approach applied to coculture data; this method estimates the percent of fibroblasts and cancer cells in each coculture, and normalizes the data for composition differences prior to estimating the effect of epithelial-stromal interaction on gene expression. The Buess interaction coefficient “I” was calculated as the ratio of observed to expected gene expression and an “I-matrix” representing the epithelial-stromal interaction coefficient for each gene in each coculture

was generated. The estimated "I" for each gene and coculture can be thought of as an indicator of the ratio of that gene's expression level relative to the expected level based on the cellular composition and the monoculture expression values. For coculture studies, I-matrices were analyzed using multiclass Significance Analysis of Microarrays (SAM [86]), comparing MCF10A to MCF10AT to MCF10DCIS cocultures (three classes). Microarray analysis was done using R.1.14. Heatmap generation and visualization were done using Cluster 3.0 and Java treeview, respectively. Functional and pathway analyses were done using Ingenuity Pathway Analysis (IPA), with Benjamini–Hochberg multiple testing correction to identify significant functions and pathways with *P*-values less than 0.05. Pathways and functions with less than 2 genes were excluded from our analysis.

2.3.6. Calculation of Basal-like interaction score

We utilized gene sets identified in Camp *et al.* [5] to score each coculture for the degree to which it expressed Basal-like microenvironment genes. In Camp *et al*. a 30 gene signature was identified that predicted Basal-like vs. Luminal interactions in cocultures and that also distinguished Basal-like vs. Luminal tumors. Using I-values (as described above) for each of these 30 genes across all cocultures, we computed a Basal-like interaction score. Briefly, using the method of Creighton *et al*. [87], vectors corresponding to the thirty genes in the Basal-like signature were constructed, with 1 assigned to genes up-regulated in Basal-like cocultures/cancers and -1 assigned to down-regulated genes. A Pearson correlation coefficient was calculated for this standard vector versus the vector of I-values for each coculture experiment. The Pearson correlation coefficient from this analysis is defined as the 'Basal-like interaction score'.

2.3.7. Analysis of cytokine expression in conditioned media

To identify soluble mediators of Basal-like microenvironments in the MCF10DCIS cells, conditioned media samples from direct 1:1 cocultures (48 hours, standard coculture media conditions as described above) were analyzed according to manufacturer protocol on a RayBio Human Cytokine Antibody Array 5 (80) (Raybiotech, Norcross GA) designed to detect 80 cytokines and chemokines. Briefly, slides were blocked by incubation with blocking buffer at room temperature for 30 min and incubated with 100 µL of conditioned media at room temperature for 90 minutes. Slides were washed and incubated with biotin-conjugated antibodies overnight at 4˚C. Finally, the slides were washed and incubated with fluorescent dye-conjugated streptavidin at room temperature for 2 hours. After final washing, slides were dried by centrifugation at 0.2 RCF (Centrifuge 5702R Eppendorf, Hauppauge, NY) for 3 minutes. Fluorescent signal was detected on a laser scanner (Axon scanner) using a Cy3 (green) channel (excitation frequency 532 nm). Data for each cytokine was normalized to positive controls on the same slide to estimate relative protein expression. Each monoculture or direct coculture was analyzed in duplicate.

2.3.8. Western Blot

Cells were harvested from interaction cocultures and protein was isolated and quantified as described previously [27]. Lysates were denatured by boiling with β-mercaptoethanol and 15-30 µg of protein were electrophoresed on a 4–20% Tris-HCl Criterion precast gel (Bio-Rad,Hercules, CA) and transferred to a Hybond-P membrane (Amersham-GE Healthscience) by electroblotting. The blots were probed with antibodies against the receptor MET (Cell signaling #8198S), HGF–α chain (Santa Cruz sc-166724) and β-actin (Cell signaling #4967). Blots were washed three times with Tris-buffered saline supplemented with 0.1% TWEEN and then were probed with ECL anti-mouse IgG horseradish peroxidase-linked whole antibody from sheep (Amersham-GE Healthscience).

Blots were rewashed, and detection was by enhanced chemiluminescence western blotting detection system (Amersham-GE Healthscience). Relative MET and HGF protein concentration was quantified using Image J software, with pixel intensity of the MET or HGF protein band divided by pixel intensity of the β-actin band. Fold-change expression was calculated by dividing the coculture expression by the monoculture expression at that same time point.

2.3.9. qPCR for MET and HGF

The relative abundances of HGF (Hs00300159_m1 Cat. # 4331182) and MET (Hs01565584_m1 Cat. # 4331182) mRNA were quantified by qPCR using an ABI 7900HT machine (Life Technologies, Carlsbad, CA). mRNA was isolated from cells using Qiagen's RNeasy mini kit and protocols (Qiagen, Valencia, CA). 1 µg of total RNA was reverse transcribed into cDNA using the iScript cDNA synthesis kit and protocol from Bio-Rad. The cDNA was then diluted five-fold by the addition of 80 µl of water. Subsequently, 2 µl of cDNA and 18 µl of master mix 10 µl SsoFast 2X Probes Supermix (Bio-Rad), 0.5 µl 18S-VIC and 0.5 µl gene specific Assay-On-Demand-FAM (ABI), 7 µl water were used in each well of the qPCR 96-well plate. Amplification conditions were as follows: 1 cycle of 95°C for 1 minute; 40 cycles of 95°C for 5 seconds, 60°C for 20 seconds.

2.3.10. Generation of coculture-derived HGF signature.

To identify HGF-regulated genes, monocultures of MCF10DCIS cells were grown in serum-free media and treated with recombinant human HGF (rhHGF, 294-HG/CF, R&D Systems) for six hours with addition of HGF every hour at a 100 ng/ml concentration (half-life of HGF is approximately four minutes). Total RNA was isolated after six hours of treatment and analogous cocultures with fibroblasts were performed. Microarray data from

both HGF-treated and fibroblast-cocultured cell lines were normalized to sham monocultures (monocultures in serum free media with no rhHGF and monocultures in regular media) by subtracting the log2(R/G) values of the monoculture. The resulting log2(R/G) ratio represents the response to coculture or treatment relative to that of sham. To identify genes that were differentially regulated by both coculture and HGF treatment, a one class SAM analysis was performed with all HGF-treated and cocultured arrays first normalized to sham monocultures. Functional and pathway analyses of the resulting gene signatures were performed using Ingenuity Pathway Analysis (IPA), with significant functions and pathways defined as those with *P* values less than 0.05 after Benjamini–Hochberg multiple testing correction.

2.3.11. Correlation with HGF signature in human tumors

We evaluated the behavior of our HGF Signature (described above) in 707 breast cancer samples from three publically available datasets: 1) NKI295 (N=295) [88], 2) Naderi *et al.* (N=135) [89] and 3) UNC337 samples (N=277) [2]. Intrinsic subtype classification was performed using the PAM50 predictor Parker *et al.* [4]. By mapping the 280 gene signature to all three datasets, a common probe set representing 109 unique genes was identified. These probes were median centered across samples, and then duplicate genes were collapsed to a unique entrez ID by statistical mean. To classify tumors according to HGF expression (positively or negatively correlated), we applied methods described in Creighton *et al.* (as described above) [87]. Correlations that were > 0 were classified as positive for the HGF signature (HGF-positive) and correlations that were <= 0 were classified as negative for the HGF signature (HGF-negative). We then obtained a Chi square statistic (four degrees of freedom) to test the association between tumor subtype (Basal-like, Luminal A, Luminal B, Her2, Normal-like) and HGF score (HGF-positive vs.

HGF-negative). All statistical analyses were performed using R 1.14 and Bioconductor packages.

2.3.12. Migration/Wound assays

Migration/wound assays were performed in six-well plates. 1:1 ratios of epithelial and RMFs were seeded in direct cocultures for 48 hours prior to performing the scratch. Epithelial cells were stained for better visualization using 5 µmol/L of Invitrogen (Life technologies) Cell Tracker (Green CMFDA) following the manufacturers protocol and as previously described in [5]. A scratch was made with a pipette tip and allowed to close for 6 hours. Phase and green fluorescent pictures were taken at 0 and 6 hours with an Olympus IX70 at 4X magnification. Image J was used to quantify percent wound closure by measuring the area free of cells at 0 hours and at 6 hours.

2.3.13. 3D Morphogenesis assay

3D cocultures were performed as previously described in [90, 91]. A model system similar to that of Jedeszko *et al*. was used; however our RMF lines are not engineered to overexpress HGF [92]. Briefly, a 1:3 ratio of epithelial:RMF cells were cocultured in a 3D extracellular scaffold composed of a 1:1 mixture of biologically derived collagen I and Matrigel® (BD Biosciences). The final concentration of Collagen I was 1 mg/mL. One well from a 24 well plate was prepped for coculture by coating with 500 µL of Matrigel-collagen mix. Then, 1 mL of cell suspension in Martrigel-collagen mix was plated. Cultures were maintained in a humidified, 37ºC, 5% CO_2 incubator for two weeks, with media change every 2 days. To test the role of HGF in morphogenesis of these cocultures, a set of plates were cultured with neutralizing, anti-HGF antibody (0.5 mg/ml, Abcam10678) was added every day for 2 weeks.

The Matrigel-collagen embedded 3D structures were fixed after 2 weeks using 4% paraformaldehyde (USB) overnight. Fixed cultures were then cryopreserved in 20% sucrose in 0.1 M phosphate buffer at 4°C and washed before embedding and freezing in optimal cutting temperature compound (Tissue-tek 4538). Frozen sections (6 µm) were cut for immunohistochemistry using a Leica 1950 cryostat. For immunostaining, slides were brought to room temperature, hydrated and placed in Citrate buffer pH 6.0 (Thermo-Fisher TA135-HBH). Heat-induced epitope retrieval (HIER) was performed using a decloaking chamber (Biocare Medical) at 95°C for 5 minutes followed by 90°C for 10 seconds. Slides were cooled for 20 minutes, washed in Tris-buffer (0.05M pH 7.6) and blocked in 10% Normal Goat Serum (NGS) in Tris for 1 hour at room temperature. Samples were incubated overnight at 4°C in Mouse monoclonal Smooth Muscle Actin (Dako M085; 1:100) and Rabbit polyclonal Cytokeratin (Dako Z0622; 1:100). After rinsing, slides were incubated at room temperature for 3 hours in a mixture containing Goat anti-Mouse (Alexafluor 568, Invitrogen A21134; 1:400) and Goat anti-Rabbit (Alexafluor 488, Invitrogen A11008; 1:400) antibody. Slides were washed and coverslipped with Fluorogel II containing DAPI (EM Sciences). For hematoxylin and eosin stain (H&E), frozen sections were stained 1 minute in acidified Harris Hematoxylin (Thermo Scientific 6765003), rinsed in running tap water for 4 minutes, stained 3 minutes in alcoholic Eosin Y (Thermo Scientific 6766007), dehydrated in 95% alcohol for 2 minutes and 5 minutes in absolute alcohol, cleared in Xylene for 6 minutes and coverslipped with DPX mountant. Phase and fluorescent images were obtained using an Olympus IX-81 microscope at 10X magnification.

Acinar morphology and apoptosis were assessed using H&E slides of 3D sections. Apoptosis was present if apoptotic bodies (small, sealed membrane vesicles) were present [93]. Immunofluorescence (IF) staining of pan-cytokeratin was used to score each organoid

for lumen (present or absent) and to confirm epithelial cell identity of acinar cells. Each structure was visualized and classified as “with lumen” if there was a clear open space within the center of the structure, “no lumen” if the lumen was filled by cells or cellular debris. For lumen and apoptosis, 30-35 acinar structures were analyzed per condition.

Lumen size and total acinar size were also measured by a method based on optical coherence tomography (OCT). Imaging of the 3D cultures was performed using a custom, ultrahigh-resolution, spectral-domain optical coherence tomography (SD-OCT) system as previously described in [90]. The OCT image stacks were resampled into an isotropic pixel resolution of 1.55 μm after correcting for the refractive index of the aqueous gels, and displayed in a “hot” color map using MATLAB® (2011a, MathWorks). From color mapped OCT images, cell clusters resembling acini were selected based on their ‘spherical’ shape. The OCT image containing the central position of each acinus was determined by sifting through the OCT image-stack to find the image with the largest acinus size. The overall acinus area and acinus lumen area were each characterized from these central OCT image slices using ImageJ. The mean acini area and mean lumen area were calculated for each gel. A total of 50-60 acinar structures were analyzed per condition.

2.3.14. Statistical analysis

Statistical analyses for experimental results were performed using SAS 9.2 (32). ANOVA was performed to compare Basal-like interaction scores across MCF10A, MCF10AT1 and MCF10DCIS cocultures. Two-tailed t-tests were performed to determine statistical differences between lumen size (by OCT measurement in 3D morphogenesis assay) of MCF10A and MCF10DCIS, with and without anti-HGF. A Chi-square analysis was performed comparing the presence of apoptosis (Yes/No) and lumen (Yes/No) in the 3D morphogenesis in MCF10A, MCF10DCIS and MCF10DCIS + anti-HGF cocultures.

2.4. Results

2.4.1. The MCF10A series acquire "Basal-like microenvironment" characteristics at DCIS stage

Each cell line in the MCF10 isogenic panel had a distinct response to coculture with reduction mammoplasty fibroblasts (RMF). By multiclass SAM (significant analysis of microarrays) we identified approximately 700 genes as differentially expressed across these three cell lines *(*Figure 2.2 A*).* One set of genes was particularly upregulated in MCF10DCIS cells, and not in MCF10A or MCF10AT cells (grey bar in Figure 2.2 A). This cluster of genes was analyzed by Ingenuity Pathway Analysis (IPA) (Tables 2.1 and 2.2) and results suggest immune response processes and connective tissue disorders, such as immune cell trafficking (*p-value*=0.021), cell mediated immune response (*p-value*=0.005) acute phase response signaling (*p-value*<0.001) or cell mediated immune response (*p-value*=0.005). Many of these processes were also upregulated in invasive Basal-like breast cancers in direct cocultures [5].

We utilized a gene set identified by Camp *et al*. [5] to directly test whether the cocultures upregulated Basal-like microenvironment characteristics. This signature can distinguish Basal-like from Luminal cocultures *in vitro* and can distinguish Basal-like from Luminal tumors *in vivo*. We calculated a score for each sample based on its correlation with the signature, and this score was termed the 'Basal-like interaction score'. MCF10DCIS cells had a high Basal-like interaction score *(*Figure 2.2. B*),* similar to that of invasive Basal-like breast cancer cell line, SUM149. In contrast, the cocultures of premalignant MCF10AT1 and MCF10A cells showed weakly positive Basal-like interaction scores and the Luminal MCF7 cell line showed a negative score *(*Figure 2.2. B*).* Pearson correlations were also performed for the monocultures and interaction cocultures of these same cell lines (Figure 2.3). The correlation observed in the indirect cocultures followed

the same trend as observed in direct cocultures, although the strength of the Basal-like score was attenuated. This attenuation was expected based on dilution and diffusion requirements in indirect coculture; signals from the fibroblasts are diluted in large volumes of media, whereas direct coculture reduces the dependence on diffusion kinetics and protein stability. While neither direct nor interaction cocultures simulate the inhibitory effects of basement membrane on cellular signaling, direct cocultures offer important advantages to address cell-cell signaling, particularly when considering effectors with short half-lives. Thus, all coculture data are from direct cocultures unless otherwise stated.

2.4.2. Upregulation of secreted cytokines in MCF10DCIS-fibroblast cocultures.

Having established that Basal-like microenvironments are induced by soluble factors, we sought to identify the secreted mediators. Eighty cytokines and chemokines were measured in the conditioned media of the direct cocultures. A striking increase in the number of cytokines expressed occurred in the MCF10DCIS cocultures, with a total of 62 cytokines upregulated by more than 1.5-fold. In contrast, MCF10A and MCF10AT cocultures each upregulated only a small number of cytokines (Table 2.3; A full list of cytokines and their fold-change relative to monoculture is provided in Table 2.4). The most highly upregulated cytokine in DCIS cocultures was hepatocyte growth factor (HGF), which increased monotonically from MCF10A to MCF10AT1 to MCF10DCIS and was upregulated more than 80-fold in MCF10DCIS and 70-fold in MCF10AT1 direct cocultures (Figure 2.4. A).

When HGF secretion is measured in the conditioned media from cocultures, the source of the HGF is not discernible. To identify which component of the system is producing the HGF, intracellular HGF RNA and protein levels were measured in each cell line after 48 hours of coculture. Figure 2.5 shows that HGF is exclusively being produced

by the fibroblasts. The epithelial cells had no detectable levels of transcript or protein in the monoculture; however, in coculture some HGF protein was observed in the epithelial cells, presumably due to internalization of the receptor-ligand complex. HGF secretion and activation is part of a complex cascade that regulates the actions of HGF. Consistent with previously published results [94], we found HAI-1 differentially expressed in our heatmap from Figure 2.2. A. HAI-1 inhibits the activity of HGF-activator thereby inhibiting the activation and subsequent activity of HGF. HAI-1 was found in higher levels in the MCF10A cocultures, accounting for different levels of HGF observed between MCF10A and MCF10DCIS in our cytokine arrays (which are presumed to measure active HGF).

HGF is the major ligand for the MET tyrosine kinase receptor and is also a negative regulator of MET transcription [95]. To evaluate whether MET is present in these cell lines, and whether MET levels are differentially regulated in coculture conditions, we assayed expression of this receptor in the MCF10 series, alone and in coculture over the course of 48 hours. Both at the RNA (Figure 2.5. B) and the protein level (Figure 2.5. C), we observed that contact with fibroblasts (6 hours) induced the MCF10DCIS cells to markedly upregulate MET RNA. Peak RNA induction at 6 hours is followed by peak protein expression at 12 hours. This effect was not observed (in MCF10A) or markedly diminished (in MCF10AT1) in the other two cell lines of the series. Thus, the interaction of MCF10DCIS cells with RMF in coculture stimulates an increase in HGF secretion and a concomitant increase in epithelial HGF receptor, MET, expression.

2.4.3. An HGF gene signature correlates with Basal-like tumors

Our coculture results established a fold-change increase in HGF signaling in premalignant, Basal-like microenvironments; however, if these changes are essential for Basal-like breast carcinogenesis, then they should also be present in invasive Basal-like

breast cancers. To assess this hypothesis, we generated an *in vitro* HGF signature. We identified gene expression changes that occurred in both (1) MCF10DCIS monocultures treated with rhHGF and (2) cocultures of MCF10DCIS with RMFs. These HGF-regulated genes are most likely to be effecting the action of HGF on Met in coculture. IPA analysis of the HGF signature suggested Sonic Hedgehog Signaling (*p-value*=0.007), Basal Cell Carcinoma Signaling (*p-value*=0.017), Tight Junction Signaling (*p-value*=0.020) among other signaling pathways were upregulated by HGF signaling in cocultures. Using this signature we scored 707 invasive tumors from three independent data sets as having a high or low correlation with this HGF signature. Figure 2.6.A shows that 86% of the aggressive Basal-like tumors in these data sets are positively correlated with the HGF signature, whereas only 23.6% of the Luminal A tumors present a positive association. Additionally, 40% of Her2-like and 36% of Luminal B tumors, both of which are more aggressive than Luminal A tumors, were positively associated with the HGF signature (*p-value*<2.2e-16). Furthermore, among Basal-like patients that are positive for the HGF signature, patients had worse overall survival (Figure 2.6.C). These results emphasize the importance of HGF signaling in aggressive breast cancer. While our coculture results show that HGF signaling is already present at the DCIS stage, the importance of this pathway in survival illustrates that the dysregulation of the HGF pathway persists in invasive Basal-like tumors and contributes to their progression.

2.4.4. Blocking of HGF inhibits three dimensional phenotypes

Our previous studies of Basal-like vs. Luminal cocultures indicated 'hepatic fibrosis' signaling was upregulated in Basal-like cocultures [5], and in light of our current data illustrating (1) that MCF10DCIS:RMF cocultures have high Basal-like interaction scores, (2) that HGF was secreted/active in MCF10DCIS cocultures, and (3) that HGF signaling is

over-represented among invasive Basal-like tumors and is predictive of overall survival, we used HGF-targeted antibodies to study the role of HGF in Basal-like interaction score and functional morphogenic coculture assays. First, MCF10DCIS:RMF cocultures were incubated for 36 hours to allow for epithelial-stromal interactions to occur, and for the last 12 hours anti-HGF antibody was added every 2 hours. In Figure 2.7.A we observe a decrease in the Basal-like interaction score due to anti-HGF treatment. While there is variation (due to random or other unexplained causes) in the time course data such that 4, 8 or 12 h after exposure have Basal-like interaction scores ranging from negative to slightly positive, it is clear that at all three time points the score is dramatically reduced.

Morphogenic assays have been shown to track normal, physiological acinar development as well as pathological malignant potential of epithelial cells [96]. The Brugge and Bissell labs have characterized the development of 3D structures over time in MCF10A cells, identifying important morphogenic mechanisms [97-99]. Namely, morphogenesis represents a balance between a variety of physiologic and pathologically-relevant processes including cell proliferation, apoptosis, and cellular migration [97] To first confirm that blocking HGF signaling with anti-HGF antibody was affecting these phenotypes, we performed wound healing assays. Figure 2.8 shows that the MCF10DCIS:RMF cocultures are greatly affected by inhibition of HGF, migration of the epithelial cells was reduced in the absence of HGF signaling. Therefore, we aimed to develop an integrated picture of how key cellular phenotypes change by studying morphogenesis in 3D with and without HGF signaling. Briefly, the expected development of MCF10A cells in 3D dictates that by 6-8 days they have reached their final size and will then start acquiring a lumen through apoptosis of the centrally located cells [31, 33, 100] Using this as a metric, we performed 3D cocultures of MCF10A and MCF10DCIS with RMFs in a matrigel-collagen mix for two weeks. Cells were treated in the presence or absence of an anti-HGF antibody and the resulting acinar structures were analyzed using

two techniques 1) longitudinal Optical Coherence Tomography (OCT) imaging at two time points (one week and two weeks) and 2) traditional H&E and IF staining at two weeks.

To track the morphogenesis of these cell lines over time, we used OCT. as we have previously found that OCT imaging allows for longitudinal 3D *in vitro* imaging without disruption of the acinar structures [101]. In addition, with these images we can estimate the size of the acinar structures and their lumens. After counting 50 structures per condition, we observed no overall difference in the size of MCF10DCIS cocultures with and without anti-HGF antibody treatment. However, blocking HGF resulted in a statistically significant decrease in lumen size (*p-value*=0.017, Figure 2.7.B). The MCF10DCIS cocultures treated with the anti-HGF antibody bore a greater resemblance to the non-malignant MCF10A (Figure 2.9). As shown in Figure 2.7.B, the number of structures without a lumen is high in MCF10A cocultures, and similar in the MCF10DCIS cocultures treated with anti-HGF, whereas the untreated MCF10DCIS cocultures have progressed to form a lumen (Table 2.5; p-value=0.0007). These differences cannot be attributable to differences in proliferation rates of the cell lines, because the population doubling times (PDT) of MCF10A and MCF10DCIS with RMF are very similar (29.7 vs. 30.8 hours, respectively).

Finally, H&E stains were used to measure apoptosis. By counting the presence of apoptotic bodies in the lumens, structures were classified as having apoptotic bodies or not [93]. Since apoptosis is the mechanism by which the lumens are formed [31, 33, 100], we would expect that in cells where lumen formation is not complete (MCF10A), the levels of apoptosis would be higher at a given time point. Figure 2.7.C shows that as expected, apoptosis was greater in the cocultures of MCF10A compared to MCF10DCIS cocultures and that treatment with anti-HGF antibody restores MCF10 apoptosis levels to those more similar to MCF10A (Table 2.5; p-value=0.0450053). Taking these results collectively, treatment with anti-HGF reverts the MCF10DCIS cells to morphogenic phenotypes that

resemble the less malignant cells (MCF10A), thereby blocking a microenvironment-mediated increase in malignant potential.

2.5. Discussion

Breast cancer progression requires that epithelial cells acquire capabilities that enhance their growth and survival. While biological models of cancer have traditionally emphasized cell-autonomous characteristics, it is clear that changes in the microenvironment are also necessary [6]. Castro *et al*. demonstrated that the stroma of DCIS lesions already possess alterations found in full invasive tumors [102], Ma *et al*. and Allinen *et al*. demonstrated that genomic changes occur in many cell populations of the microenvironment [14, 76]. These models explicitly hypothesize that epithelial cells must undergo an evolutionary adaptation to their microenvironments [8, 13, 103]. Other models propose that the microenvironment is the driving force of the benign lesion evolution into full invasive tumors; Gatenby and Gillies even speculate that the origin of cancer may lie not in mutations within epithelial cells but within acquired or somatic mutations changes in the mesenchymal cells that control tissue structure [104]. While our data do not support these latter models, our data do suggest that progression is not isolated to a single compartment (the epithelium), but rather reflective of epithelium-stroma co-evolution. Heterotypic interactions between epithelium and stroma foster this co-evolution by selecting for complementary phenotypes in stroma and epithelium.

Evidence of progressive complementarity between stroma and epithelium is critical to documenting co-evolution. For example, considering transition from pre-invasive to invasive tumors, Hu *et al.* recently demonstrated that NFKB and COX-2 signaling contribute to epithelial-stromal mediated invasive potential [69]. In the current study we document co-evolution in the earlier transition between benign and pre-invasive lesions. A

striking upregulation of HGF was evident early in progression (at the atypical hyperplasia stage), but was not sufficient to induce the characteristic Basal-like stromal-interaction phenotype. MET must also be expressed, and at the DCIS stage MET was upregulated with drastic consequences on the behavior of the epithelial cells: MCF10DCIS cells expressed high Basal-like interaction scores and an associated ability to progress in morphogenesis assays. Similarly increased malignant potential has been observed in xenograft models with MCF10DCIS, which also preserve stromal interactions [105, 106]. The importance of the HGF pathway was further documented in our system by blocking HGF signaling with an antibody, resulting in reversal of expression and morphologic phenotypes. In fact, this latter experimental work underscores an advantage of our *in vitro* coculture system: in studying cancer evolution using this cell line panel, pathways can be manipulated and direct causal effects can be examined. We were also able to separately assay both epithelial and stromal characteristics and show complementary changes. Epithelial cells upregulate the MET receptor only at the DCIS stage, despite HGF expression by all cocultured fibroblasts. While other studies have reported HGF expression by the epithelial cells [107, 108], these studies utilize different cell line models of progression and we did not observe a similar autocrine pathway of HGF-MET interaction in our studies. The previous studies focused on cell line models of metastatic progression, so perhaps our focus on pre-invasive stages of disease also accounts for some differences between findings. It is interesting to consider that by later stages of progression, this pathway may develop autocrine capability. By studying the interactions and the relative contributions of each component at different stages of progression, we can better understand how stroma and epithelium are co-evolving.

Our results emphasize that HGF/MET signaling is important in Basal-like breast cancer progression from early in the disease, but HGF has long been studied in invasive cancer biology and in normal development [109, 110]. Previous studies have linked

HGF/MET signaling with poor outcome in invasive breast cancers [111]. Several recent publications have demonstrated the importance of HGF/MET signaling and the microenvironment in melanoma treatment-resistance [112, 113]. Our data, using a novel HGF signature and three independent tumor datasets, indicate that HGF/MET signaling is highly correlated with Basal-like breast cancer subtype and worse overall survival in patients. Mouse models with over expression of MET induce Basal-like tumors with signatures of WNT and epithelial to mesenchymal transition (EMT) [114], suggesting that this pathway's importance in tumor biology is conserved across species. In normal tissue, HGF is produced by stromal fibroblasts and acts as a mitogen, motogen and morphogen on MET-expressing epithelial cells [115]. MET is a tyrosine kinase receptor that, when activated by its ligand (HGF), auto-phosphorylates and initiates an intracellular signaling cascade that involves many targets. In the developing mammary duct, deletion of epithelial MET inhibits ductal branching [116] and in adult glands, HGF is critical for tubulogenesis [117]. Thus, the HGF/MET pathway is an essential player in normal development and wound healing [118, 119]. Given the high expression of wound response genes in the tissue adjacent to cancer [73] and the important role of HGF in normal ductal morphogenesis and invasive breast cancer, a better understanding of HGF/MET in progression of Basal-like breast cancers is important.

Future work should focus on studying how molecules that block MET signaling affect stromal-epithelial interactions and should study HGF/MET in tissue from pre-invasive Basal-like lesions. One case series reported by Lindemann *et al*. [120] attempted to link HGF and MET signaling in earlier lesions by immunohistochemistry studies of HGF and MET in DCIS. That study concluded that an imbalance in MET expression between the tumor and the surrounding normal tissue is associated with aggressive DCIS phenotypes. However, uncertainty remains about how the imbalance can best be characterized in human tissue. Studying cell-cell communication is difficult in tissue, so the availability of

complementary *in vitro* models is helpful. These coculture models advance our understanding of the reciprocal molecular changes in the pre-invasive stages of breast cancer and can guide research on tissue.

In conclusion, heterotypic interactions are crucial for disease progression both compartments, the stroma and the epithelium must coevolve to produce a successful tumor. Understanding the reciprocal epithelial and stromal changes that occur in early lesions will help to identify strategies to treat patients and/or prevent invasive breast cancers. HGF/MET signaling is a strong candidate pathway for treating premalignant Basal-like lesions and the application of MET inhibitors should be considered in preclinical models to advance this plausible strategy.

2.6. Figure and Tables

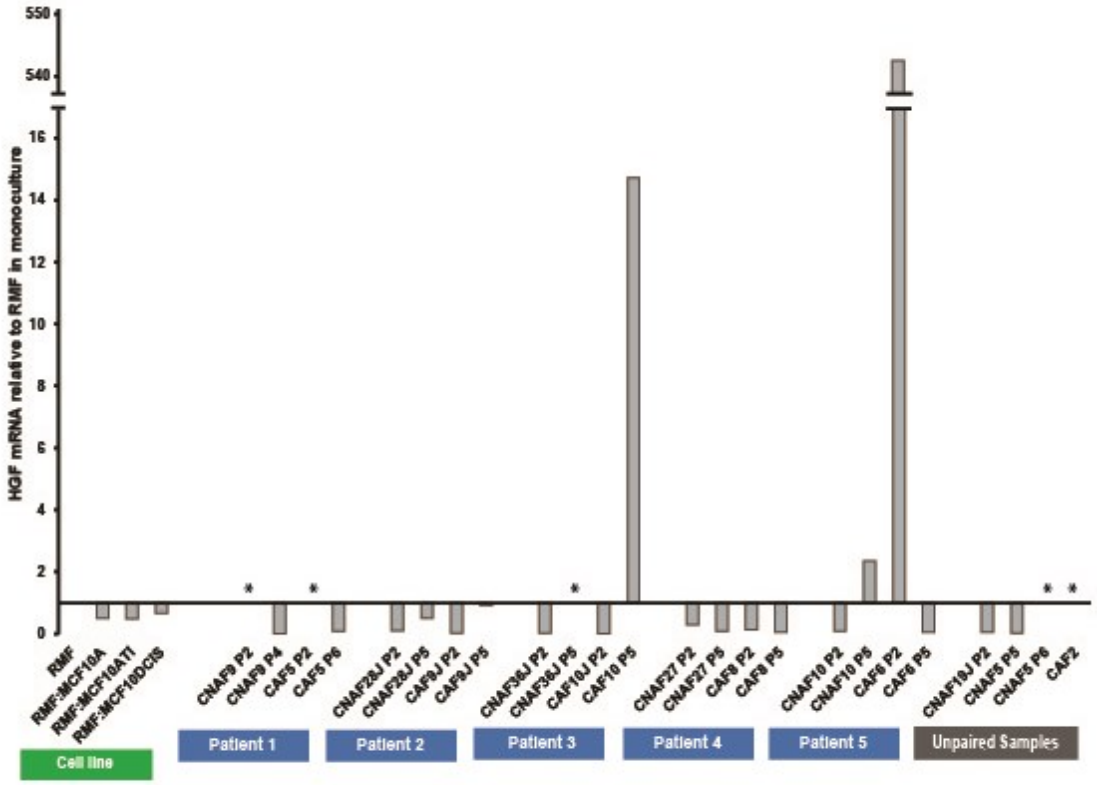


Figure 2.1. **Primary fibroblasts have temporal and intra-individual instability.** mRNA quantification of HGF across a panel of 14 primary fibroblasts lines and the RMF cell line in coculture with the MCF10A progression series. Primary fibroblasts were isolated from 5 patients both from the cancer-adjacent tissue (CNAF: cancer-normal associated fibroblasts) and the tumor itself (CAF: cancer-associated-fibroblasts) at different passages (p=2 and p=5). HGF levels vary at the transcriptional level between patients and between passages. Samples with a (*) had no detectable levels of transcript by qPCR.

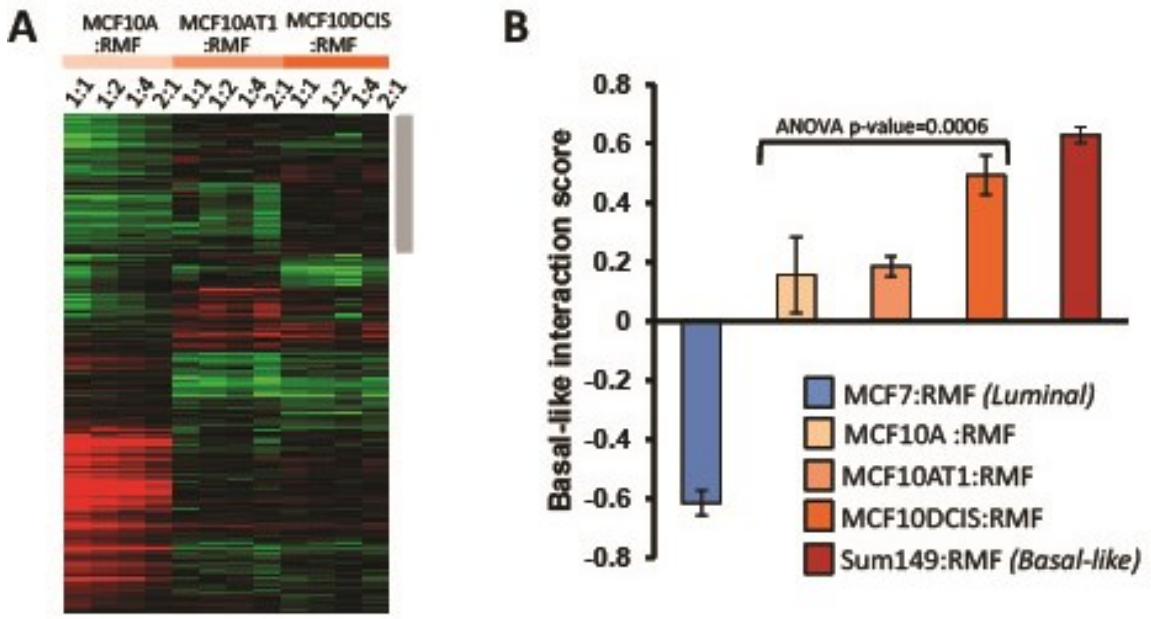


Figure 2.2. ***The MCF10A series progressively acquires "Basal-like microenvironment" characteristics***. A) One-dimensional (genes only) hierarchical clustering showing interaction scores (I-values) for genes that significantly change due to coculture. Distinct clusters of genes are upregulated in each cell line. A grey bar adjacent to the heatmap shows a cluster of genes that are uniquely upregulated in the MCF10DCIS:RMF cocultures. B) The 'Basal-like interaction score' was developed by Camp et al [5] previously, the coculture data shows a strong relationship between progression and expression of Basal-like microenvironment characteristics. Luminal cocultures had a negative Basal-like interaction score, while the score increased with progression from A/AT1 to DCIS, with DCIS having interaction scores similar to those in the Basal-like invasive cancer cell line, SUM149.

Ingenuity Canonical Pathways	p-value	Ratio	Molecules
FXR/RXR Activation	5.37032E-06	1.43E-01	BAAT,ABCG5,MLXIPL,APOB,CYP27A1,NR1H4,FBP1,APOC2,PLTP,FOXA3,HNF1A,MTTP
Primary Immunodeficiency Signaling	5.7544E-06	1.64E-01	PTPRC,BTK,LCK,CD19,IL2RG,ZAP70,IGHA1,CD8A,JAK3
Acute Phase Response Signaling	2.95121E-05	9.3E-02	HAMP,C4BPB,ORM1/ORM2,APOH,HNF1A,SERPINF2,F2,CRABP1,SERPIND1,ALB,TF,C4BPA,ITIH4,PIK3CD,IL1RAP,C2
CTLA4 Signaling in Cytotoxic T Lymphocytes	0.000281838	1.05E-01	CD3G,LCK,PPP2R5C,ZAP70,HLA-DRB1,PIK3CD,HLA-DQB1,HLA-DRB5,TRA@,LCP2
Leukocyte Extravasation Signaling	0.000354813	7.94E-02	MMP15,ARHGAP4,CLDN6,RHOH,BTK,CYBA,CYBB,PIK3CD,RASSF5,ACTG2,ACTC1,CLDN3,MMP9,ACTA1,PRKCB
Human Embryonic Stem Cell Pluripotency	0.00040738	8.28E-02	FGFR3,BMP4,LEFTY1,TDGF1,SMO,FGFR2,PIK3CD,LEFTY2,HNF1A,BMP5,TCF7L2,POU5F1
CD28 Signaling in T Helper Cells	0.000446684	9.02E-02	PTPRC,CD3G,LCK,ZAP70,HLA-DRB1,PIK3CD,HLA-DQB1,CARD11,HLA-DRB5,TRA@,LCP2
T Cell Receptor Signaling	0.000467735	9.8E-02	PTPRC,BTK,CD3G,LCK,ZAP70,PIK3CD,CD8A,CARD11,TRA@,LCP2
Intrinsic Prothrombin Activation Pathway	0.000977237	1.72E-01	F11,KNG1,F5,COL2A1 (includes EG:1280),F2
B Cell Development	0.001348963	1.67E-01	PTPRC,CD19,HLA-DRB1,HLA-DQB1,HLA-DRB5

Table 2.1. **IPA pathway analysis of genes upregulated in MCF10DCIS:RMFs direct cocultures.**

Category	Functions Annotation	p-value	# Mol	Molecules
Cell-To-Cell Signaling and Interaction	activation of cells	1.42E-03	25	APP, CD1A, CD1B, CD1D, CD1E, CD2, CD200, CD8A, CTSG, F2, F5, IL2RG, KNG1, LCK, LGALS3, LTB, MAGEA3/MAGEA6, MLANA, PTPRC, TCL1A, TLR2, TNFRSF8, TYR, VTN, ZAP70
Cell-mediated Immune Response	contact growth inhibition of T lymphocytes	5.22E-03	3	LCK, PTPRC, ZAP70
Cell-To-Cell Signaling and Interaction	activation of blood cells	1.12E-02	17	CD1A, CD1B, CD1D, CD1E, CD2, CD200, CTSG, F2, F5, KNG1, LCK, LTB, MAGEA3/MAGEA6, MLANA, PTPRC, TCL1A, TYR
Cellular Growth and Proliferation	proliferation of leukocytes	1.52E-02	17	BATF, BMP4, CARD11, CD19, CD2, CD6, CD8A, HLA-DRB1, IL10RA, IL28A, IL2RG, LAIR1, PDCD1, PTPRC, TCL1A, TYR, ZAP70
Cellular Development	developmental process of leukocytes	1.72E-02	21	APP, BMP4, CBFA2T3, CD1D, CD2, CD6, GFRA1, IFI16, IL10RA, IL2RG, LAIR1, LCK, LIF, MLANA, MMP9, PDCD1, STAT5A, TCL1A, TLR2, TYR, ZAP70
Cell-To-Cell Signaling and Interaction	response of peripheral blood leukocytes	1.91E-02	3	IL10RA, IL1RAP, TLR2
Cell-To-Cell Signaling and Interaction	activation of leukocytes	2.07E-02	15	CD1A, CD1B, CD1D, CD1E, CD2, CD200, F5, KNG1, LCK, LTB, MAGEA3/MAGEA6, MLANA, PTPRC, TCL1A, TYR
Inflammatory Response	activation of leukocytes	2.07E-02	15	CD1A, CD1B, CD1D, CD1E, CD2, CD200, F5, KNG1, LCK, LTB, MAGEA3/MAGEA6, MLANA, PTPRC, TCL1A, TYR
Immune Cell Trafficking	activation of leukocytes	2.07E-02	15	CD1A, CD1B, CD1D, CD1E, CD2, CD200, F5, KNG1, LCK, LTB, MAGEA3/MAGEA6, MLANA, PTPRC, TCL1A, TYR
Inflammatory Response	activation of lymphocytes	2.58E-02	11	CD1A, CD1B, CD1D, CD1E, CD2, LCK, MAGEA3/MAGEA6, MLANA, PTPRC, TCL1A, TYR
Cell-To-Cell Signaling and Interaction	binding of leukocytes	2.60E-02	10	APOH, CD2, CD6, CTSG, CX3CL1, F2, IL28A, NPY, ORM1/ORM2, VTN
Inflammatory Response	shape change of phagocytes	2.83E-02	3	ALB, CORO1A, VTN
Cellular Growth and Proliferation	proliferation of lymphocytes	3.00E-02	15	BATF, BMP4, CARD11, CD19, CD2, CD6, CD8A, HLA-DRB1, IL2RG, LAIR1, PDCD1, PTPRC, TCL1A, TYR, ZAP70
Cellular Development	developmental process of T lymphocytes	3.00E-02	13	BMP4, CD1D, CD2, CD6, IL2RG, LCK, MLANA, PDCD1, STAT5A, TCL1A, TLR2, TYR, ZAP70
Cellular Movement	homing of cell lines	3.00E-02	12	APP, CX3CL1, L1CAM, LCK, LIF, PIK3CD, SNAI2, TFF2, TFF3, TLR2, TNFRSF8, VTN
Hematological System Development and Function	survival of lymphocytes	5.95E-02	5	BMP4, BTK, CD8A, JAK3, NFIL3
Cell-To-Cell Signaling and Interaction	response of leukocytes	6.76E-02	6	CORO1A, IGHA1, IL10RA, IL1RAP, TLR2, VTN
Cell-To-Cell Signaling and Interaction	activation of T lymphocytes	7.06E-02	8	CD1D, CD2, LCK, MAGEA3/MAGEA6, MLANA, PTPRC, TCL1A, TYR
Inflammatory Response	activation of T lymphocytes	7.06E-02	8	CD1D, CD2, LCK, MAGEA3/MAGEA6, MLANA, PTPRC, TCL1A, TYR
Cellular Growth and Proliferation	proliferation of T lymphocytes	8.27E-02	11	BATF, BMP4, CARD11, CD2, CD6, CD8A, HLA-DRB1, IL2RG, LAIR1, PDCD1, ZAP70

Table 2.2.. **IPA biological function analysis of genes upregulated in MCF10DCIS:RMFs direct cocultures**

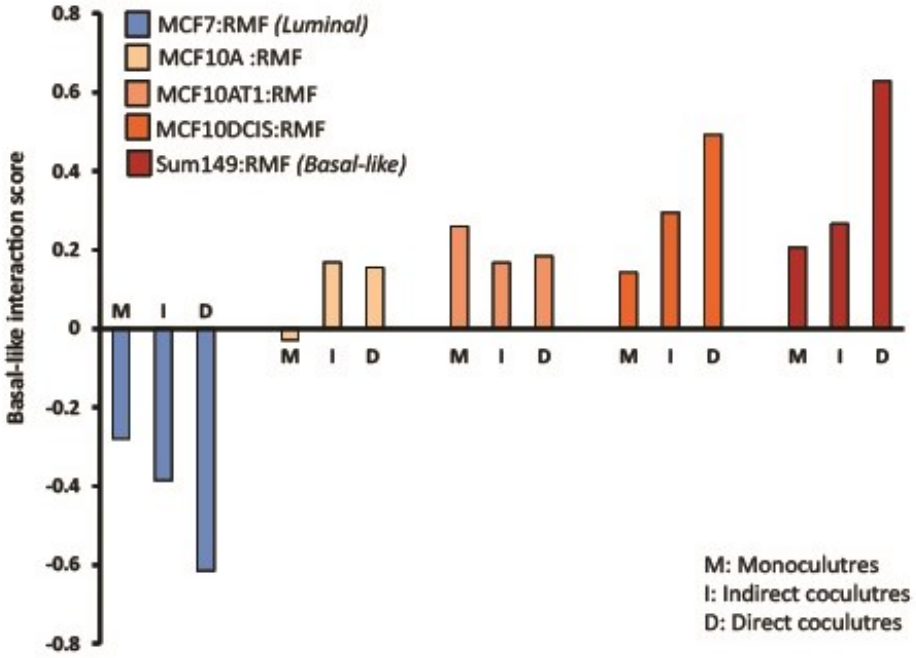


Figure 2.3. **The MCF10A series progressively acquires "Basal-like microenvironment" characteristics in interaction cocultures**. Basal-like interaction score of monoculutres (M), Indirect cocultures (I) and direct cocultures (D). Indirect cocultures maintain the trends of the 'Basal-like interaction score" present in the direct cocultures (from Figure 2.2.B).

	Cytokine Fold Change			
	1.5-3	3-10	> 10	Total #
MCF10A:RMF	-	1	-	1
MCF10AT1:RMF	5	2	1	8
MCF10DCIS:RMF	50	11	1	62

Table 2.3.. **Number of cytokine expressed in the cocultures of the MCF10A series**.

	MCF10A:RMFs	MCF10AT1:RMFs	MCF10DCIS:RMFs
CXbeta6-1	-	-	1.57
MIP-3alpha	-	-	1.63
I-I-4	-	-	1.65
Eotaxin	-	-	1.66
TARC	-	-	1.68
I-I-3	-	-	1.7
FGF-4	-	-	1.71
RANTES	3.13	2.32	1.8
BDNF	-	-	1.83
IGFBP-1	-	-	1.85
FGF-9	-	-	1.86
NAP-2	-	-	1.86
IL-13	-	-	1.95
Thrombopoetin	-	-	1.98
GDNF	-	-	2
leptin	-	-	2.01
BLC	-	-	2.02
Osteopontin	-	-	2.03
MDC	-	-	2.04
Eotaxin-3	-	-	2.05
MIP-1delta	-	-	2.06
PIGF	-	-	2.08
Fractalkine	-	-	2.09
SCF	-	-	2.11
Eotaxin-2	-	-	2.16
SDF-1	-	-	2.17
FGF-7	-	-	2.19
IGFBP-4	-	-	2.25
I-I-1B	-	-	2.26
TNF-alpha	-	-	2.28
GM-CSF	-	-	2.3
LIF	-	-	2.32
Flt-3 ligand	-	-	2.33
MCSF	-	-	2.33
MIF	-	-	2.35
NT-3	-	-	2.35
IGF-1	-	-	2.35
IL-16	-	-	2.37
IGFBP-3	-	-	2.39
PARC	-	-	2.39
I-I-10	-	-	2.4
TGF-beta3	-	-	2.42
NT-4	-	-	2.48
LIGHT	-	-	2.5
I-I-5	-	-	2.51
MCP-4	-	-	2.55
I-I-7	-	-	2.66
I-309	-	-	2.66
TNF-beta	-	-	2.82
MIG	-	-	3.21
GCSF	-	-	3.59
I-I-6	-	6.02	3.67
MCP-2	-	-	3.72
I-I-2	-	-	3.8
FGF-6	-	1.56	5.04
GCP-2	-	3.16	5.56
MCP-3	-	2.5	6.09
Osteoprotegrin	-	-	6.21
ENA-78	-	2.37	6.24
MCP-1	-	7.96	7.57
Oncostatin M	-	-	11.14
HGF	-	70.12	80.65

Table 2.4. **Fold change relative to monoculture of all cytokine that change in coculture of the MCF10A series. List of the cytokines detected on the antibody based array with the values for each coculture.**

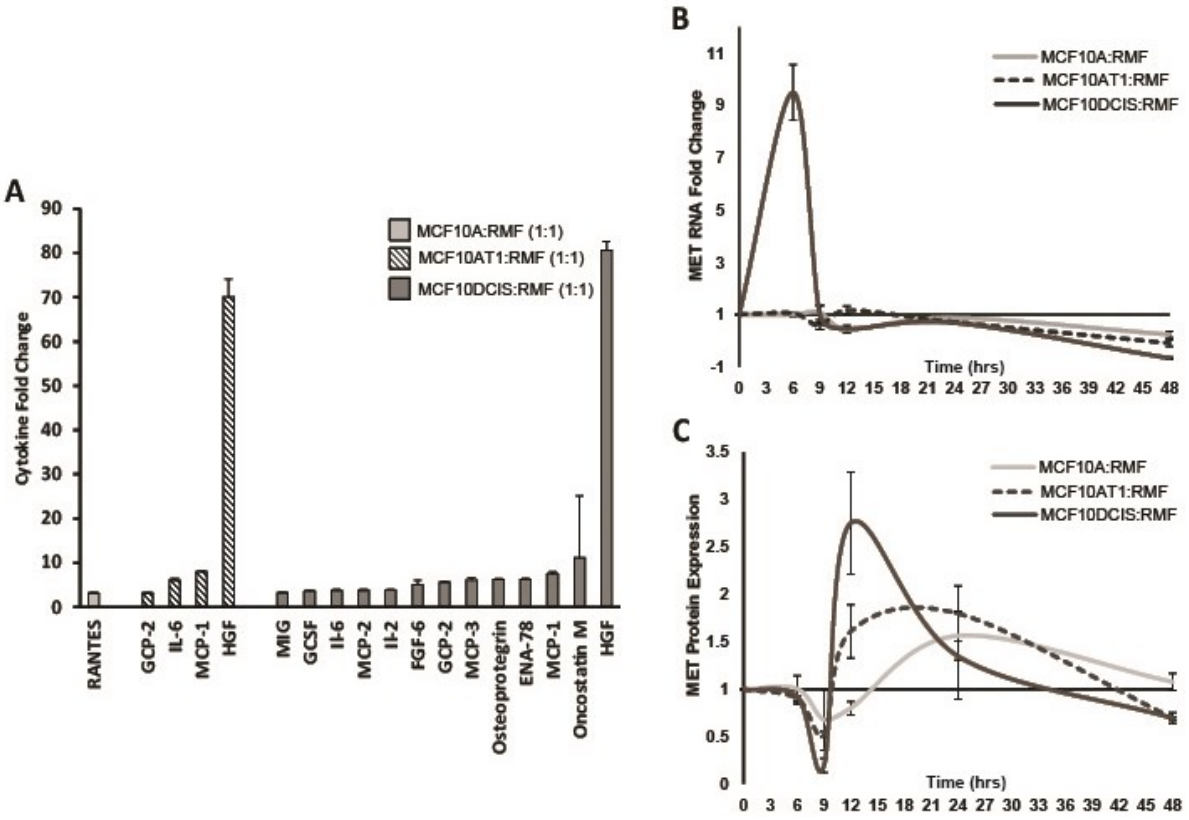


Figure 2.4. ***The MCF10A series cocultures present differentially secreted cytokines and MET receptor status.*** A) A variety of cytokine are overexpressed in the direct cocultures of all three cell lines. The number of cytokines overexpressed increases with progression. The cytokine most strongly overexpressed was HGF. B) MET RNA fold change expression relative to the corresponding cell line in monoculture over a 48hour period. We observe a sharp increase of transcript in the MCF10DCIS cells when in coculture after 6 hours. C) MET protein fold change expression relative to the corresponding cell line in monoculture over a 48hour period. The protein increase is delayed 6 hours when compared with RNA expression due to the delay in translation.

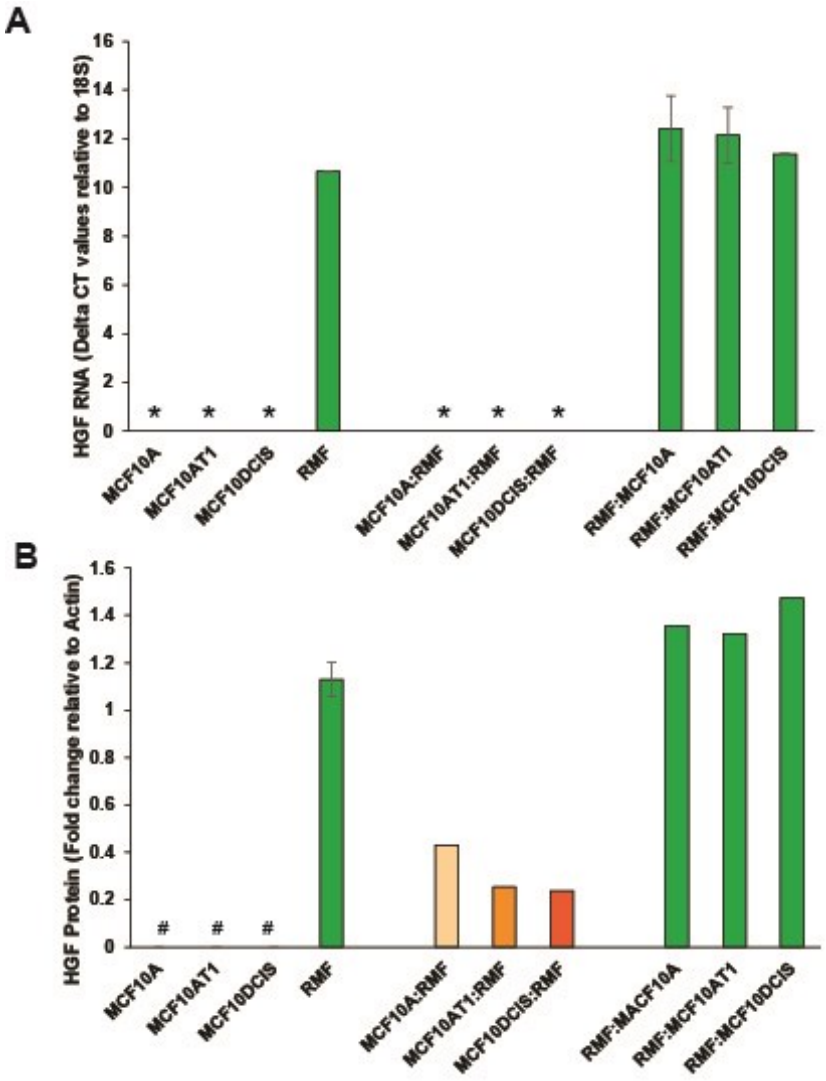


Figure 2.5. **HGF is produced by the stromal component of the cocultures, the RMFs**. A) RNA levels of HGF (delta CT values relative to 18s gene) in the monocultures of the MCF10A series and the RMFs, as well as these cell lines in coculture. The epithelial cells had no detectable levels of HGF transcript (*) even in coculture with RMFs; only RMFs had high levels of the transcript. B) Protein levels of HGF, again of cells in monoculture or coculture. Epithelial cells in monoculture had no detectable levels of HGF (#) however in coculture they appear to have some HGF protein, we argue this is due to the internalization of the receptor-ligand complex. RMFs, had high levels of HGF expression both in monoculture and in coculture. Both graphs demonstrate the same trend, the stromal cells are responsible for HGF secretion in this coculture system.

A

Tumor subtype	Correlation with HGF signature	
	Negative	Positive
Basal-like	18	117
Her2 - like	60	40
Luminal A	177	55
Luminal B	103	59
Normal - like	29	45

X-squared = 149.6211 df=4 p-value < 2.2e-16

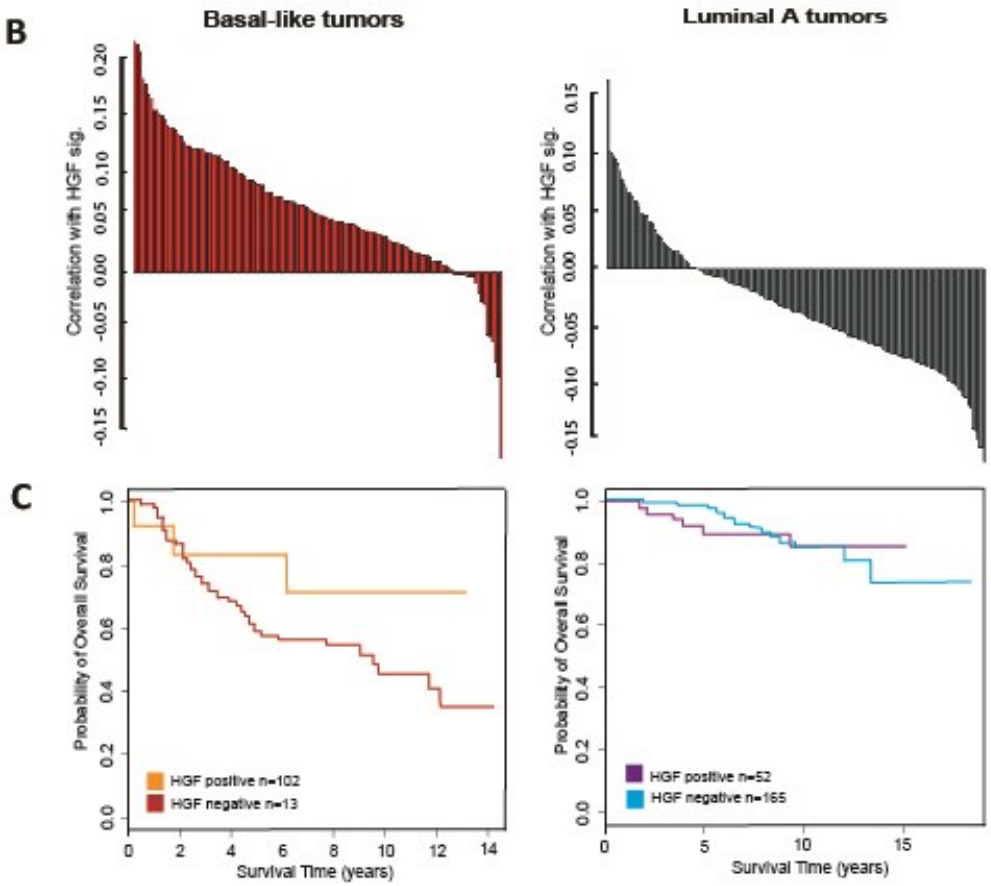


Figure 2.6. **HGF signaling is present *in vivo* in Basal-like tumors.** An *in vitro* generated signature that captures HGF signaling is highly correlated with invasive Basal-like tumors. A) Table with the number of tumors that are positively and negatively correlated with the HGF signature by subtype. B) Bar graph that shows the Creighton correlation of each Basal-like and Luminal tumors with the HGF signature, the degree of positivity among Basal-like tumors is higher when compared to the Luminal tumors (scale bar) C) Kaplan-Meier survival curves for overall survival among patients that were positive or negative for the HGF signature. Patients with Basal-like tumors with positive HGF signatures had worse overall survival over a period of 14 years.

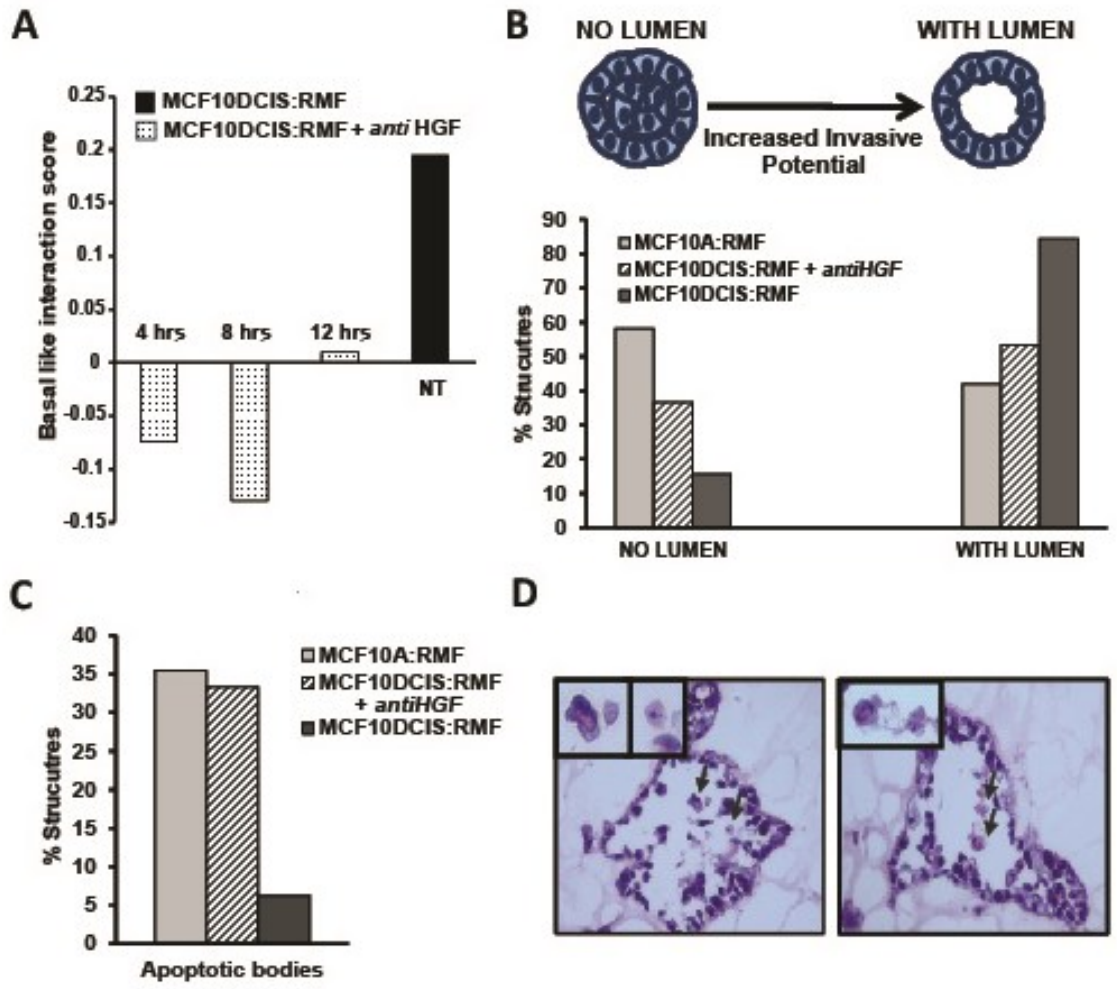


Figure 2.7. **Blocking HGF signaling reverts Basal-like microenvironments and slows down morphogenesis in a 3D coculture assay.** A) Basal-like interaction score of MCF10DCIS:RMF direct cocultures when treated with anti-HGF antibody (as described in the methods section), despite the variability Basal-like score is reversed from positive association to negative association. B) Morphogenesis represents a balance between cell proliferation, apoptosis, and cellular migration. Morphogenesis assays track these processes *in vitro*. Bar graph shows the quantification of structures with and without lumens at 2 weeks in coculture. When HGF signaling is blocked in the DCIS cocultures, the morphogenesis process is slowed down causing these acinar structures to display an intermediate phenotype between the MCF10DCIS and the MCF10A cocultures. C) Apoptosis was lowest for DCIS at 2 weeks, these cocultures had already undergone cavitation at that time point. Treatment with anti-HGF increases the apoptosis levels at this time-point because of delayed the cavitation process, D) Representative H&E images of cross-sections of the 3D structures and with apoptotic bodies.

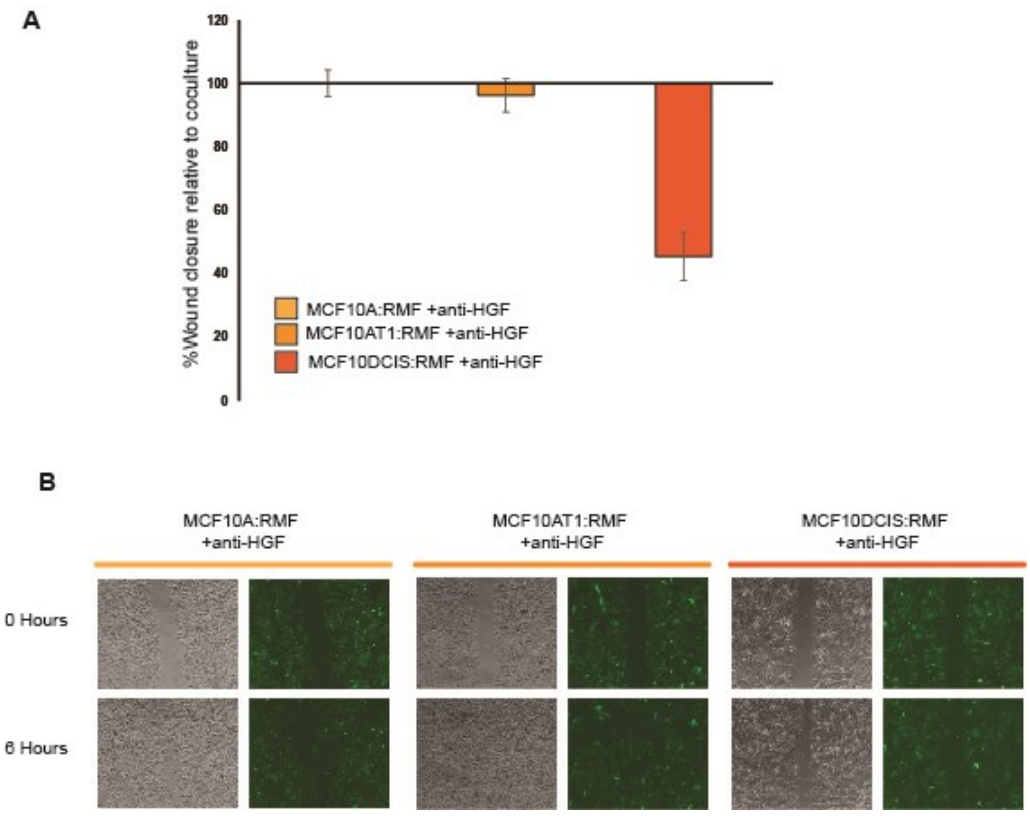


Figure 2.8. **Blocking HGF signaling in coculture conditions inhibits migratory phenotypes in the MCF10DCIS cells.** A) Quantification of % wound closure after 6 hours in a scratch assay of direct cocultures with anti-HGF relative to the cocultures without anti-HGF. B) Representative phase and fluorescent pictures of the direct cocultures with anti-HGF at 0 and 6 hours. The epithelial cells were dyed green for better visualization.

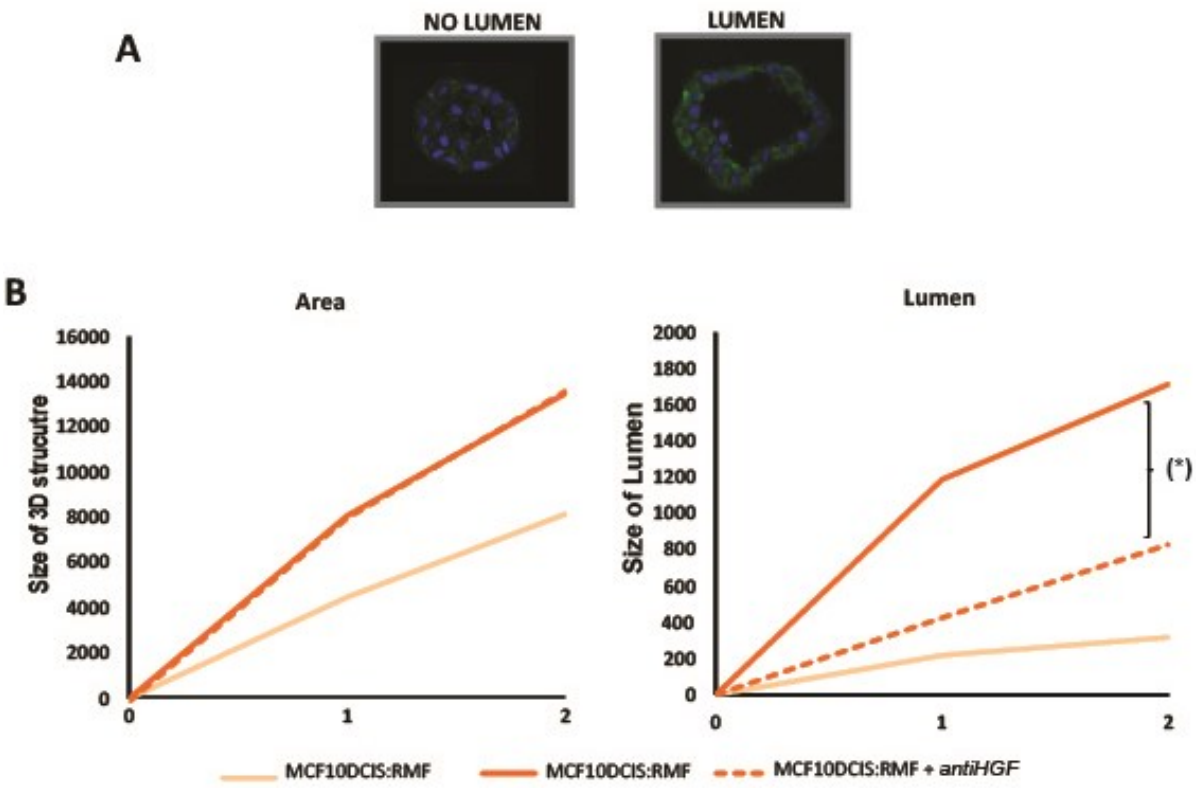


Figure 2.9. **Quantification of OCT measurements of the acini structures.** A) Representative fluorescent pictures of acinar structures stained with pan-cytokeratin (green) and DAPI (nucleus), the left picture shows a structure without a lumen and the right picture represent a structure with a very well defined lumen. B) Graphs representing the evolution of the overall size (Area) of the acini and the size of the lumen (Lumen). Anti-HGF treatment does not affect the overall size of the 3D structures; however, it has a big influence on the area of the lumen (*p-value=0.017). Acini with anti-HGF treatment present smaller lumens resembling the more benign cell line MCF10A.

	MCF10A:RMF	MCF10DCIS:RMF	MCF10DCIS:RMF *+antiHGF*	P-value
Lumen				
Yes	13	27	16	0.00074575
No	18	5	11	
Undetermined	-	-	3	
Apoptotic bodies				
Yes	8	2	8	0.0450053
No	19	30	20	
Undetermined	3	-	2	
Total (n)	31	32	30	

Table 2.5. **Chi square analysis of 3D quantification of morphological assay. Lumen and apoptosis quantification.**

Chapter 3

TUMOR INTRINSIC SUBTYPE IS REFLECTED IN CANCER ADJACENT TISSUE

3.1. Overview

Introduction: Overall survival of early-stage breast cancer patients is similar for those who undergo breast conserving therapy and mastectomy, however, approximately 10-15% of women undergoing BCT suffer an ipsilateral breast tumor recurrence. The risk of recurrence may vary according to age or breast cancer subtype. Understanding the gene expression of the cancer-adjacent tissue and/or stromal response to specific tumor subtypes is important for developing clinical strategies to reduce recurrence risk.

Methods: We studied gene expression data in cancer-adjacent tissue from 158 patients with breast cancer. Complementary *in vitro* cocultures were used to study cell-cell communication between fibroblasts and specific breast cancer subtypes.

Results: Our results suggest that intrinsic tumor subtypes are reflected in the histologically normal cancer-adjacent tissue. Gene expression analysis of disease-free cancer-adjacent tissues shows that triple-negative (Claudin-low or Basal-like tumors) exhibit increased expression of genes involved in inflammation and immune response. While such changes could reflect distinct immune populations present in cancer-adjacent tissue of the different breast cancer subtypes, altered immune response gene expression is also observed in cocultures in the absence of immune cell infiltrates, emphasizing that these inflammatory mediators are secreted by breast-specific cell types. In addition, while Basal-like breast

cancers are associated with upregulated immune response gene expression, Luminal breast cancers are more commonly associated with estrogen-response in adjacent tissues. **Conclusions**: Specific characteristics of breast cancers are reflected in the normal tissue that surrounds them. This commonality between tumor and surrounding normal may underlie second primaries and local recurrences. Biomarkers derived from normal tissue may be helpful in defining personalized surgical strategies or in predicting recurrence risk.

3.2. Introduction

Breast conservation therapy (BCT) with lumpectomy and radiotherapy has become an established alternative to mastectomy for treatment of early breast cancer. While overall survival of early-stage breast cancer patients is similar for those who were treated with combinatorial breast-conserving surgery and radiotherapy and those who underwent radical mastectomy [121, 122], approximately 10-15% of women undergoing BCT suffer an ipsilateral breast tumor recurrence [123]. Younger age has been associated with higher rates of recurrence [124], but in recent years it has become apparent that other characteristics, such as tumor subtype, may underline these associations. Aggressive breast cancers tend to be diagnosed in younger women [83] and have higher local recurrence rates when compared to less aggressive breast cancers [125, 126]. Furthermore, while at least some studies among all patients (BCT and Mastectomy combined) have found increased risk of recurrence among younger women with Basal-like or triple-negative breast cancers, the association may vary by surgery type [127], indicating complex interactions between treatment and biology of recurrence. The merit of aggressive surgical strategies may depend on the degree to which definitive biological changes can be detected in the tissue left behind following BCT.

In this study, we hypothesized that the stromal responses or other genomic features of histologically normal, cancer-adjacent tissue differ by intrinsic subtype. Stromal responses to breast tumors have previously been associated with clinical outcomes of breast cancer [75, 80], however, few studies have identified subtype-specific changes in surrounding stroma. A recent study reported histological differences such as differences in epithelial microarchitecture in the cancer-adjacent tissue of triple-negative breast cancers versus Luminal breast cancers [128]. Another recent study suggested that ER-positive tumors may more commonly be associated with high expression of estrogen receptor RNA in histologically normal, cancer-adjacent tissue [129]. These findings have not been confirmed in subsequent studies or in larger populations. To address whether adjacent tissue reflects the biology of the tumor itself, we investigated gene expression profiles of cancer-adjacent tissue from 158 patients. To identify the contributing cell types (cancer cell versus stromal cell contributions) we investigated whether these gene expression associations could be recapitulated *in vitro* using a coculture model system. Our results suggest that estrogen responsiveness may be more common in the normal tissue of patients developing Luminal breast cancers and distinct stromal responses are more common in Basal-like and Claudin-low tumors. The triple-negative stromal response is dependent upon fibroblast interactions and which may have important consequences for the immune microenvironment.

3.3. Methods

3.3.1. Patient samples:

Polish Women's Breast Cancer Study and The Normal Breast Study

The study population included 139 women from the PWBCS with available snap frozen extratumoral breast tissues and gene expression data. The PWBCS is a

population-based case-control study conducted in two major cities in Poland (Warsaw and Łódź) during 2000-2003 [130]. PWBCS cases were women aged 20-74 years with newly-diagnosed, pathologically-confirmed *in situ* or invasive breast carcinoma identified through a rapid identification system organized at five participating hospitals and via cancer registries. Fresh tissues from invasive tumors, non-neoplastic cancer-adjacent breast tissue and mammary fat tissue were collected at the time of breast surgery and snap frozen in liquid nitrogen. Tumor adjacent breast tissues used in this study were <2 cm from the tumor margin. Based on *in vitro* evidence of their distinctive microenvironments, Basal-like and Luminal tumors were oversampled in this study. Information on clinicopathological, demographic, and anthropometric factors was collected from medical records and in-person interviews as described previously. All participants provided written informed consent under a protocol approved by the U.S. National Cancer Institute and local (Polish) institutional review boards.

The Normal Breast Study (NBS) is a hospital-based cross-sectional study currently being conducted in UNC Hospitals (Chapel Hill, NC, USA) since 2009. A small subset of 19 patients with 45 cancer-adjacent samples was used from this study. All patients had a newly diagnosed invasive breast carcinoma. Fresh tissues were collected at the time of breast surgery and snap frozen in liquid nitrogen. Tumor adjacent breast tissues used in this study were classified as peritumoral (<2 cm from the tumor margin) and remote (>2 cm from the tumor margin). Information on clinicopathological, demographic, and anthropometric factors was collected from medical records and in-person interviews. All of the participants provided written informed consent under approved by the IRB (Protocol # LCCC-0913).

3.3.2. RNA, expression microarrays and subtyping: Tumor samples

Fresh tumor tissue was collected at the time of surgery and snap frozen. Illumina Ref-8 Beadchip Version2 microarray platform was used and normalization of the data was done using Lumi software package. To classify tumors according to intrinsic subtype, genes were median-centered and samples were standardized to zero mean and unit variance. Then the PAM50 predictor was performed as described in Parker *et al.* [4] to categorize the tumors into the five subtypes (Luminal A, Luminal B, Her2-enriched, Basal-like, and normal-like). The Claudin-low predictor was applied as described in [2].

3.3.3. RNA and expression microarrays: Cancer-adjacent samples

Fresh cancer-adjacent tissue was collected at the time of surgery and snap frozen. All microarrays on cancer-adjacent breast tissue were performed at the University of North Carolina at Chapel Hill. The medial section of fresh frozen cancer-adjacent tissue was homogenized using a MagnaLyser homogenizer (Roche), and RNA was isolated by Qiazol extraction followed by purification on an RNeasy column as described in Troester et al. [73]. RNA quality and quantity were analyzed on an Agilent 2100 Bioanalyzer and a ND-1000 Nanodrop spectrophotometer, respectively, before running two-color 4X44K Agilent whole genome arrays. Cy3-labeled reference was produced from total RNA from Stratagene Universal Human Reference (spiked 1:1,000 with MCF-7 RNA and 1:1,000 with ME16C RNA to increase expression of breast cancer genes) following amplification with Agilent low RNA input amplification kit. The same protocol was applied to total RNA from breast tissues, with all patient samples labeled with Cy5. Data were lowess-normalized, and probes that had a signal of <10 dpi in either channel were excluded as missing. Probes that had more than 20% missing data across all samples were excluded from further analysis. In expression data preprocessing, we 1) eliminated the probes without

corresponding ENTREZ ID, 2) collapsed the duplicate probes by averaging, 3) imputed missing data using k-nearest neighbors (KNN) method with k=10, and 4) median-centered each gene.

3.3.4. Supervised analysis of cancer-adjacent tissue

Four-class significance analysis of microarrays (SAM) was used to identify differentially expressed genes associated with distinct intrinsic breast cancer subtypes (Luminal A, Luminal B, Basal-like and Claudin-low) [86]. False discovery rate (FDR) was used to address multiple comparisons, with significance defined as FDR < 5%. Considering that tumors classified as "normal-like" may result from extensive normal tissue or stromal content in the tumor specimen [131], we excluded normal-like tumors from the supervised analysis. HER2-like tumors were also excluded due to their low number (n=9). From of the remaining 130 tumor-adjacent, histologically normal tissues, a gene list of 126 genes was identified as differentially expressed across these 4 tumor subtypes. Functional and pathway analyses were done using Ingenuity Pathway Analysis (IPA), with Benjamini–Hochberg multiple testing correction to identify significant functions and pathways with *P*-values less than 0.05. Pathways and functions with less than 2 genes were excluded from our analysis.

3.3.5. Composition analysis of cancer-adjacent tissues from PWBCS

Frozen sections were obtained adjacent to each specimen used for microarray analysis, and 127 samples from the PWBCS had hemotoxylin and eosin (H&E) stained sections that were of sufficient quality to be analyzed for tissue composition. Slides that were out of focus or had folded tissues were not included. Briefly, 20 μm slices obtained

from the histologically normal-appearing tissue adjacent to tumors were H&E stained and scanned using Aperio ScanScope CS V11.0.2.725, a digital pathology platform for image analysis. A set of 15 H&E slides were used to train an analysis macro in Genie (Aperio Technologies, Vista, CA, USA) to segment adipose tissue, epithelium, non-fatty stroma, and glass on each slide. Methods for analysis of composition are detailed the Appendix 1, briefly, each slide was manually annotated for area of each tissue component and this was used as a 'gold standard' to compare with the automated Genie-based method. For both manual and Genie annotation, the area of each component was computed and divided by the sum of the areas of adipose, epithelium and non-fatty stroma. All image analyses were performed by UNC Translational Pathology Laboratory. The final training accuracy of the Genie macro was 99.37%. Slides were also analyzed by pathologist in terms of percentages of adipose tissue, epithelium, and non-fatty stroma in each slide for comparison.

3.3.6. Cell lines and Coculture conditions

Cell lines were purchased and maintained as previously explained in [5]. All cell lines were tested for mycoplasma by the University of North Carolina at Chapel Hill Tissue Culture Facility, NC. Cells were maintained at 37°C and 5% CO_2. Cultures were 2-dimensional, grown on plastic. Cancer cell lines and fibroblasts cells were grown in the appropriate cancer cell media (e.g., MCF 7 in RPMI) for 48-hours, after performing growth curves to ascertain that the reduction mammary fibroblasts (RMFs) maintained doubling times in each media similar to the doubling times observed in RPMI 1640. Direct cocultures were performed as previously explained [5]. The following RMF:cancer cell ratios were plated for most direct cocultures: 0:1, 1:4, 1:2, 1:1, 2:1, 1:0, and cells were maintained in coculture for 48 hours before harvesting cells for RNA isolation.

3.3.7. RNA and expression microarrays: Cell lines

Cells were harvested by scraping in RNA lysis buffer. Total RNA was isolated using the RNeasy mini kit (Qiagen, Valencia, CA) and RNA quality was analyzed on an Agilent 2100 Bioanalyzer using an RNA6000 nano chip. Quantification was performed on a ND-1000 Nanodrop spectrophotometer. Microarrays were performed according to Agilent protocol using 2-color Agilent 4×44K (Agilent G4112F) human arrays and 244K (Agilent G4502A) custom human arrays. Only probes present on the 4×44K array were utilized and all 4×44K probes were present on the 244K custom array. We used the Agilent Quick Amp labeling kit and protocol to synthesize Cy3-labeled reference from Stratagene Universal Human Reference spiked at 1:1,000 with MCF7 RNA and 1:1,000 with ME16C RNA to increase expression of breast cancer genes. The identical protocol was applied to total RNA from cocultured or monocultured cell lines to label these samples with Cy5. Labeled cDNAs were hybridized to arrays overnight and washed before scanning on an Agilent G2505C microarray scanner. All array data are available through the Gene Expression Omnibus (GSE26411).

3.3.8. Coculture data normalization and analysis

Data from 122 microarrays (representing monocultures and direct cocultures from 14 different cell lines) were included in this study. Only those genes where more than 70% of microarrays had signal in both channels greater than 10 dpi were included. Data were Lowess normalized and missing data were imputed using *k*-nearest neighbors' imputation. For the *direct coculture* analyses, we excluded genes that did not have at least 2-fold deviation from the mean in at least 1 sample and the method of Buess *et al*. [47] was used to normalize cocultures to appropriate monocultures performed in the same media and

under identical conditions as previously described in [5]. Briefly, the Buess method is an example of an expression deconvolution approach applied to coculture data; this method estimates the percent of fibroblasts and cancer cells in each coculture, and normalizes the data for composition differences prior to estimating the effect of epithelial-stromal interaction on gene expression. The Buess interaction coefficient "I" was calculated as the ratio of observed to expected gene expression and an "I-matrix" representing the epithelial-stromal interaction coefficient for each gene in each coculture was generated. The estimated "I" for each gene and coculture can be thought of as an indicator of the ratio of that gene's expression level relative to the expected level based on the cellular composition and the monoculture expression values. For coculture studies, I-matrices were analyzed using multiclass Significance Analysis of Microarrays (SAM [86]), comparing Basal-like and Claudin-low cocultures to the rest (Her2-like and Luminal cocultures). Microarray analysis was done using R.1.14. Heatmap generation and visualization were done using Cluster 3.0 and Java treeview, respectively.

3.3.9. Statistical analysis

R 1.14 was used to generate box plots and chi-square test to generate p-values associated with them. SAS 9.2 (32) was used for the remaining analysis and table generation. Odds ratios and confidence intervals were calculated for comparison of estrogen response signature and the triple-negative stromal response signature. Fisher's Exact test was calculated for association of clinical features among both groups of patients ('Triple-negative' and 'Others') as well as for subtype specific tissue composition due to the low 'n' in some of the cells. Chi-square analysis was performed to analysis composition and estrogen response.

3.4. Results

3.4.1. Tumor intrinsic subtype is reflected in Cancer-adjacent tissue.

We used samples from the Polish Women's Breast Cancer Study (PWBCS) to identify subtype-associated changes in normal tissue (Population characteristics of the PWBCS are shown in Table 3.1). The triple-negative-adjacent tissues had a unique stromal response, with Figure 3.1 showing a heatmap of genes whose expression differed significantly between tissues adjacent to Basal-like, Claudin-low, Luminal B and Luminal A tumors (in multiclass supervised analysis). Because there was a common signature associated with both triple-negative tumor subtypes, we collapsed tumor subtypes to conduct a 2-class comparison of Basal-like and Claudin-low to the rest of the subtypes. Gene ontology analysis (using Ingenuity Pathway Analysis) revealed that genes associated with Basal-like and Claudin-low tumors are involved in functions and pathways such as activation of leukocytes (p-value<0.001), proliferation of mononuclear leukocytes (p-value<0.001), cell movement of leukocytes (p-value<0.001), interferon signaling (p-value< 0.001), Hepatic Fibrosis (p-value < 0.0001), T-helper cell differentiation (p-value< 0.004) or antigen presentation pathway (p-value< 0.02) (a full list can be found in Tables 3.2 and 3.3).

Our results also suggest that the cancer-adjacent tissue shares biology of the tumors themselves. Four genes (NAT1, FOXA1, MLPH, ESR1) used to identify Luminal breast cancers by the PAM50 subtyping [4] have elevated expression levels in the tissue adjacent to Luminal breast cancers. Having observed high expression of ESR1 adjacent to Luminal breast cancers, and in light of previous reports suggesting similarities between ER positive tumors and their adjacent tissue [129], we utilized an estrogen response signature (EReS) generated by Oh *et al.* [132] (Figure 3.1A) to calculate Pearson correlation coefficients for each cancer-adjacent specimen. Based on these Pearson correlation

coefficients, samples were classified as being positive or negative for the EReS. The Odds Ratios (using Basal-like cancer-adjacent tissue as the reference) clearly shows an increased proportion of estrogen responsive tissues among Luminal A (OR=2.36) and HER2 (OR=1.79) (non-triple-negative) patients (Table 3.4).

To further investigate if distance affects the expression of EReS, we utilized a small cohort of 19 patients with 45 cancer-adjacent samples (from The Normal Breast Study, explained in the Method section) with carefully annotated distances from the tumor. We tested the EReS among peri-tumoral samples (less than 2cm from tumor) and ER positives were more likely to express this signature (62.5%) than ER negatives (16.6%). In addition, the EReS was tested in remote samples (more than 2cm from tumor), of ER+ and ER- tumors, we observe that it follows the previously identified association where an increased proportion of ER+ tumors have estrogen responsive cancer-adjacent tumors in samples as far as 4cm from the tumor (62.5%) (Table 3.5). This shows that estrogen receptor positive tumors arise from normal tissues that are more likely to be estrogen responsive and that the estrogen responsiveness is not localized to just the area adjacent to tumor.

Conversely, the aggressive signature appears to be highly localized. To estimate the strengths of the association between triple-negative subtype and the newly identified triple-negative gene expression signature, we also computed Pearson correlation coefficients for each tissue's association with this signature (from Figure 3.1). Again samples were classified as positive or negative for this signature using Pearson correlation (as described in methods) (Figure 3.1.B). Odds ratios (using Luminal cancer-adjacent tissues as a reference) show that Basal-likes and Claudin-lows (OR=8.2 and OR=8.4) are much more likely to express the signature associated with triple-negative cancer adjacent tissue. Luminal B (OR=3.4) were more positively associated with this signature than Luminal A breast cancers, but were more weakly associated than the triple-negatives breast cancers (Table 3.6).

3.4.2. Cancer-adjacent expression is not associated with tissue composition.

One source of variation in cancer-adjacent signatures is the heterogeneous composition of normal breast tissue. To evaluate the role of tissue composition in estrogen response and triple-negative signatures, we quantified the proportion of different components (epithelium, non-fatty stroma and fatty/adipose) present in each individual sample. Sections adjacent to the portion of tissue used for RNA extraction were used to perform analyses. We used a computational approach to automate analysis of the percentage of epithelium, non-fatty stroma and adipose present in H&E stained tissue sections. Results of this analysis show that the gene expression patterns detected in Figure 3.1 are not due to differences in composition. The average percentage of epithelium in these samples was 9.8% while average percentage non-fatty stroma was 26.8%. Table 3.7 dichotomizes the samples as below or above average for each component. There is no statistically significant association between the subtypes and epithelial content (Fisher's Exact test p-value=0.1652) or stromal content (Fisher's Exact Test p-value=0.2346). The EReS was also not associated with epithelial content (p-value=0.8570), but was significantly associated with non-fatty stromal content (p-value<0.001).

3.4.3. Cancer-adjacent biology can be recapitulated in vitro.

To further understand the unique biology of the triple-negative tumors, it is important to understand which stromal cell types are contributing to the cytokine and interleukin gene expression observed in triple-negative adjacent tissues. Previous studies examined cytokines in cancer-adjacent stroma and hypothesize that they were due to inflammatory infiltrates [75], however the origin of these signatures are not well documented. *In vitro* cocultures are useful systems to model specific heterotypic

interactions and we have previously demonstrated that they recapitulate tumor biology [5]. However, in the current study we wished to address whether these changes extend to cancer-adjacent tissue and to expand the definition of triple-negative to include the newly identified Claudin-low subtype [2]. Thus, we cocultured breast cancer subtype models with reduction mammary fibroblasts (RMFs) to study gene expression changes due to heterotypic interactions. By analyzing these cocultures, we identified a unique set of genes upregulated in triple-negative cocultures (Figure 3.2.A, grey bar). To address whether this *in vitro* gene signature is associated with triple-negative-adjacent tissue *in vitro* we took two approaches: (1) we compared pathways and biological functions *in vitro* with those identified in tissue (as shown in Figure 3.1) and (2) we studied the expression of the *in vitro* genes in breast tumors and the cancer adjacent tissue.

Using the first approach, the genes identified through the *in vitro* cocultures (Tables 3.8 and 3.9) and *in vivo* are involved in similar pathways (Tables 3.2 and 3.3). Statistically significant biological functions and pathways common in both settings were activation of cells (p-value<0.001), proliferation of mononuclear leukocytes (p-value<0.001), cell movement of leukocytes (p-value<0.001), inflammatory response (p-value<0.001), hepatic fibrosis (p-value<0.001) and those involved in immune cell such activation role of cytokines in mediating communication between immune cells (p-value=0.002) or IL-6 signaling (p-value=0.002). Using the second approach, the *in vitro,* coculture-derived triple-negative gene signatures were associated with subtype *in viv*o. Basal-like and Claudin-low tumors had high expression of coculture genes both in the intratumoral microenvironment (Figure 3.2.B), p-value=8.44e-15 and in the cancer-adjacent tissue (Figure 3.2.C), p-value=0.196. The magnitude of the differences were lower among cancer-adjacent tissues and the sample size was smaller, so the resulting p-value was not statistically significant, but the patterns of expression mirror the statistically significant patterns observed intratumorally.

3.5. Discussion

Stromal responses to tumor are well documented in many gene expression and pathology studies [75, 76, 79, 128], occurring in invasive disease and early during disease progression (Chapter 2). Specific stromal responses, such as wound response signatures in the tissue adjacent to cancers, can predict cancer survival and relapse [73, 74]. However, while most tumors appear to be associated with some stromal reaction, there is substantial heterogeneity in the gene expression and other characteristics of cancer-adjacent tissue. Roman-Perez *et al.*, identified two distinct subtypes of cancer-adjacent tissues that were independent of tumor subtypes but that presented distinct survival patterns [80]. Graham *et al*. also showed that ER positive tumors more commonly express estrogen receptor in adjacent tissue [129]. Yang *et al.* conclude that compared to Luminal tumors, tumors with basal phenotypes (by immunohistochemistry markers) are associated with less involution in cancer-adjacent terminal-ductal lobular units (TDLU) [128]. Given that there appear to be subtype-specific stromal responses and tissue characteristics, this research may have important translational implications. Of note, aggressive forms of breast cancer (i.e.: triple-negative breast cancers) have higher rates of local recurrence [125, 126].

Our results support biological differences in the cancer-adjacent tissue of Basal-like vs. Luminal breast cancers and suggest that biologically distinct tissue characteristics exist in the histologically normal tissue left behind after tumor resection. Tissue adjacent to Basal-like and Claudin-low tumor subtypes (triple-negative breast cancers) have highly expressed genes involved in immune and inflammatory processes, consistent with the proinflammatory milieu reported previously for aggressive tumors [2, 81, 133]. This response of the stroma or ‘reactive stroma’ [23] in these triple-negative breast cancers, may be important for outcome and may reflects a response to tumor instead of a host predisposition. In less aggressive Luminal cancers, genes used to identify Luminal tumors

[4] as well as their characteristic estrogen response were found upregulated in the cancer-adjacent tissue.

The current study is cross-sectional in nature and therefore cannot identify whether the cancer-adjacent signatures are a response to tumor or pre-exist tumor formation. However, if these changes are host factors that precede the tumors, they may reflect subtype-specific risk. Some authors have proposed that host factors or widespread field effects underlie phenotypic concordance between first and second primaries, however molecular evidence of this has not been identified [134, 135]. Our results may reflect molecular evidence of etiologic distinctiveness: we observed that ER positive tumors are more likely to occur in EReS normal tissues. An important question is whether this difference is present only peritumorally, or across a broader geographic range within the tissue.

Distance from tumor has been hypothesized to be important for determining whether molecular changes surrounding tumors relate to recurrence or second primary risk. Yang et al. hypothesized that changes observed close to the tumor are due to cancerization effects, especially mutations in epithelium, and those observed further away are a result of the environmental risk factors [136]. We found that ER positive tumors tend to have estrogen-responsive cancer-adjacent tissues both peritumoral and remote. Because of this association we suspect it is a host or etiologic pre-disposition factor. The triple-negative signature could not be assessed in samples from >2cm from the tumor because intrinsic tumor subtype information was not available for these remote samples. However, based on previous observation we speculate it to be localized response, and related to stromal reaction to the tumor. If this is true, it seems less likely to reflect risk of second primaries. Thus, the different factors that influence the phenotype of the cancer-adjacent tissue may have distinct clinical implications.

Cancer-adjacent tissue is composed of a high percentage of stroma (both fatty and non-fatty) and gene expression analyses of these tissues are enriched for stromal pathways. Because of this, many studies have approached the study of stromal response and field cancerization effects by microdissecting and studying the individual cellular components [14, 75, 76]. However, cellular heterogeneity is an important factor, and while our results do not allow definitive attribution of signatures to specific cell types, we are able to see that stromal composition plays a distinct role in the estrogen responsiveness signature. While mechanisms by which the stromal content mediates the estrogen response remain to be determined, some evidence suggests that stroma may modulate hormone receptor levels in breast epithelium [137]. By studying whole tissue we are able to capture important epithelial-stromal interactions among the different components. However, tissue studies, even with microdissection, cannot perfectly isolate the immune infiltrates from stroma and epithelium, thus other approaches are needed to identify the cell type of origin for the immune mediators that are upregulated in the triple-negative microenvironment.

Our use of cocultures allowed us to identify fibroblasts as important contributors to cytokine/chemokine expression in the tissue adjacent to triple-negative cancers. Our data suggest that triple-negative stromal response is dependent upon fibroblasts interactions because epithelial-fibroblast cocultures, even in the absence of immune cells, produce molecules that are important in immune cell recruitment and activation. These molecules (i.e.: cytokines, chemokines, etc.) can also play important roles in regulating epithelial cells. For example, IL-6 appears to regulate epithelial cellular differentiation [138] . Many others have previously shown that *in vitro* epithelial-cell generated signatures have predictive accuracy when using tumor gene expression data [5, 47, 59, 77], however we expand this to show that even stroma-derived signatures can be predictive and that some of the signatures are not only highly expressed in tumors but also in cancer adjacent

tissues. The robustness of the biological phenotype across settings strengthens our observation that triple-negative BCs have unique microenvironments.

Our study had some limitations. Our analysis is limited to 139 patients from Poland, and 19 from a North Carolina-based study. While the U.S. study appears to validate the findings from Poland, complete tumor subtype data was not available and so medical record abstraction of receptor status was used instead. Future work should verify our observations in other populations with accurate tumor subtype information. In addition, while neither EReS nor triple-negative subtype were associated with epithelial content, it is possible that epithelium specific signals were under-represented in our tissue. Epithelial content is low in normal breast tissue (median 9.8% in our data), and therefore microdissection studies may be necessary to enrich for these changes. Simultaneously, consideration of other demographic variables (such as BMI, race or age) in tandem with tumor subtype may help to further understand heterogeneity in cancer-adjacent tissue expression among patients with particular tumor subtypes. Future work could also evaluate specific immune cell populations and their associations with the cytokine profiles that are upregulated in triple-negative cancers.

In conclusion, we found distinct biological characteristics of the cancer adjacent tissue depending upon the intrinsic characteristics of the tumor. This commonality between tumor and surrounding normal may underlie second primaries and local recurrences and may provide plausible explanations as to why aggressive tumors recur more often. These results also suggest that it may be feasible to develop normal tissue biomarkers that can help to define appropriate, personalized surgical strategies.

3.6. Figures and Tables

	Other tumors Luminal A (n=53) Luminal B (n=49) Her2 (n=9)	%	Triple Negative Basal-like (n=17) Claudin-low (n=13)	%	TOTAL	%	Fisher's Exact Test p-value
ER status							
Missing	10	9.17	2	6.67	12	8.63	
(-)	17	15.60	20	66.67	37	26.62	
(+)	82	75.23	8	26.67	90	64.75	<0.0001
Tumor size							
>2cm	59	54.13	11	36.67	70	50.36	
2-5cm	45	41.28	19	63.33	64	46.04	
>5cm	5	4.59	0	0	5	3.6	0.0654
Grade							
Missing	19	17.43	7	23.33	26	18.71	
Well differentiated	17	15.60	2	6.67	19	13.67	
Moderately diff	52	47.71	5	16.67	57	41.01	
Poorly diff	21	19.27	16	53.33	37	26.62	<0.0001
Node status							
0	53	48.62	20	66.67	73	52.52	
1-3 positive nodes	33	30.28	8	26.67	41	29.50	
>4 positive nodes	23	21.10	2	6.67	25	17.99	0.1075
Age							
>=29	1	0.92	0	0			
30-39	3	2.75	3	10	1	0.72	
40-49	22	20.18	6	20	6	4.32	
50-59	39	35.78	11	36.67	50	35.97	
60-69	31	28.44	7	23.33	39	27.34	
<=70	13	11.93	3	10	16	11.51	0.6530
TOTAL	109		30		139		

Table 3.1. **Characteristics of Polish Women's Breast Cancer Study patients with gene expression profiles**

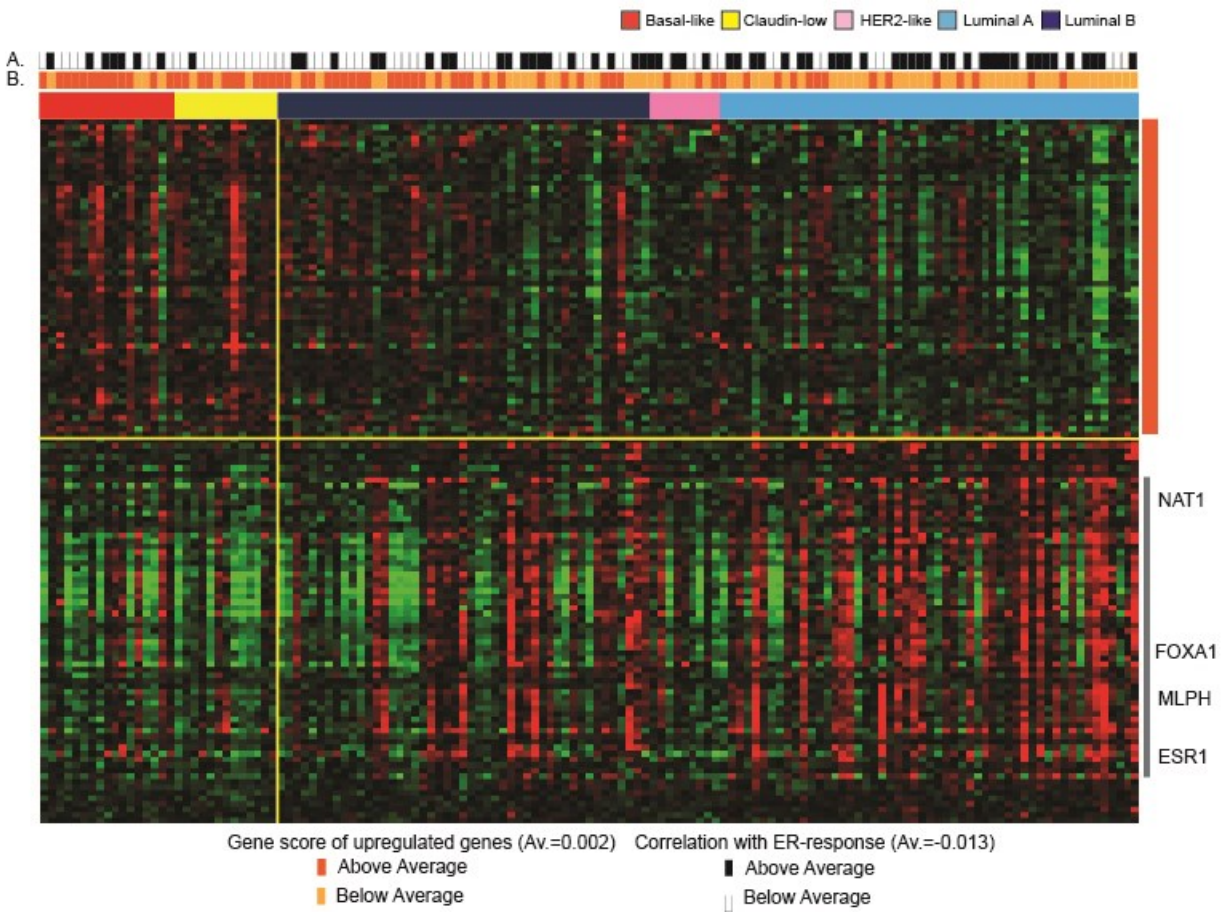


Figure 3.1. **Tumor Intrinsic Subtype is reflected in Cancer-Adjacent Tissue.** Heat map representing 126 differentially expressed genes across the cancer-adjacent tissue of Claudin-low (triple-negative), Basal-like (triple-negative), Her-2-like, Luminal B and Luminal A. Distinct clusters of up and down regulated genes show a trend in the gene expression of these cancer-adjacent tissues. Among those 126 genes, 4 genes that are used in the PAM50 classification system to identify Luminal tumors are aso highly expressed in their adjacent tissue. A) Samples were dichotomized as having a high (black) or a low (white) expression of an Estrogen response signature (EReS) developed by Oh *et al*. Luminal cancers had the highest proportion of EReS positive cancer-adjacent tissues. B) Samples were dichotomized as having high (dark orange) or low (light orange) expression of the newly identified triple-negative signature (upregulated genes in cancer-adjacent tissue of Basal-like and Claudin-low tumors, represented here by the orange bar on the side). The triple-negative cancers (Claudin-low and Basal-like) had the highest proportion of samples with high expression of this signature.

Category	Functions Annotation	p-Value	# Mol
Cell Death and Survival	apoptosis	1.44E-03	44
Cell-To-Cell Signaling and Interaction	activation of cells	1.52E-04	16
	binding of tumor cell lines	6.05E-04	10
	activation of leukocytes	8.71E-04	11
Cellular Development	proliferation of blood cells	1.46E-05	17
	proliferation of mononuclear leukocytes	2.29E-05	15
	proliferation of lymphocytes	6.24E-05	14
	differentiation of tumor cell lines	1.52E-03	10
Cellular Growth and Proliferation	proliferation of blood cells	1.46E-05	17
	proliferation of mononuclear leukocytes	2.29E-05	15
Cellular Movement	cell movement of leukocytes	3.96E-04	15
Dermatological Diseases and Conditions	Psoriasis	2.46E-05	27
Hematological System Development and Function	proliferation of mononuclear leukocytes	2.29E-05	15
Immunological Disease	systemic autoimmune syndrome	5.36E-04	31
Inflammatory Disease	relapsing-remitting multiple sclerosis	7.88E-06	6
	arthritis	4.29E-05	34
	chronic granulomatous disease	9.95E-05	3
	rheumatoid arthritis	3.49E-04	27

Table 3.2. **Biological functions associated with genes differentially expressed in cancer adjacent tissue of Claudin-low, Basal-like, Luminal B and Luminal A tumors**.

Ingenuity Canonical Pathways	p-value
Interferon Signaling	0.000107
Hepatic Fibrosis / Hepatic Stellate Cell Activation	0.00049
Role of Pattern Recognition Receptors in Recognition of Bacteria and Viruses	0.000589
Macropinocytosis Signaling	0.001148
Autoimmune Thyroid Disease Signaling	0.001698
Allograft Rejection Signaling	0.002455
T Helper Cell Differentiation	0.003802
Production of Nitric Oxide and Reactive Oxygen Species in Macrophages	0.008511
Sphingosine-1-phosphate Signaling	0.00912
Antigen Presentation Pathway	0.016218
Complement System	0.016218
Histamine Biosynthesis	0.017378
Graft-versus-Host Disease Signaling	0.020893
Role of NFAT in Cardiac Hypertrophy	0.025704
Cytotoxic T Lymphocyte-mediated Apoptosis of Target Cells	0.026303
Leukocyte Extravasation Signaling	0.030903
Corticotropin Releasing Hormone Signaling	0.032359
iNOS Signaling	0.036308

Table 3.3. **Pathways associated with genes differentially expressed in cancer adjacent tissue of Claudin-low, Basal-like, Luminal B and Luminal A tumors.**

Breast Cancer Subtype	Odds Ratios (95% C.I.)
Basal-like	Ref.
Claudin-low	0.12 (0.01-1.14)
Luminal B	0.97 (0.31-3)
Her2-like	1.79 (0.35-9.13)
Luminal A	2.36 (0.77-7.18)

.
Table 3.4. **Odds Ratio of estrogen response signaling in the cancer-adjacent tissue (PWBCS).**

	Peritumoral <2cm from tumor n=23		Remote >2cm from tumor n=22	
	EReS (+)	**EReS (-)**	**EReS (+)**	**EReS (-)**
ER+	10 (62.5%)	6 (37.5%)	10 (62.5%)	6 (37.5%)
ER-	3 (42.9%)	4 (57.1%)	1 (16.6%)	5 (83.34%)

Table 3.5. **Cancer-adjacent tissue classified by ER status as well as distance from tumor (NBS).** Association between estrogen response gene signature and distance is shown.

Breast Cancer Subtype	Odds Ratios (95% C.I.)
Basal-like	8.23 (2.31-29.32)
Claudin-low	8.44 (2.04-35)
Luminal B	3.42 (1.5-7.9)
Her2-like	0.97 (0.31-3)
Luminal A	Ref.

Table 3.6. **Odds Ratio of triple-negative microenvironment signature in cancer-adjacent tissue (PWBCS).**

	Basal-like	Claudin-low	LumB	LumA	*Fisher's Exact Test p-value*	Estrogen Response		*Chi square test P-value*
	n=17 (%)	n=11 (%)	n=47 (%)	n=52 (%)		(+) n=59 (%)	(-) n=68 (%)	
% Epithelium								
Above AV n=51	4 (23.5)	7 (63.6)	21 (44.7)	19 (36.5)	0.1652	23 (38.98)	28 (41.18)	0.857
Below AV n=76	13 (76.5)	4 (36.3)	26 (55.3)	33 (63.5)		36 (61.02)	40 (58.82)	
% Stroma								
Above AV n=54	8 (47)	2 (18.2)	18 (38.3)	26 (50)	0.2346	42 (71.19)	12 (17.65)	<0.001
Below AV n=73	9 (53)	9 (81.2)	29 (61.7)	26 (50)		17 (28.81)	56 (82.35)	

Effective Sample Size = 127
Frequency Missing = 3

Table 3.7. **Cancer-adjacent tissue classified by subtype and by %epithelium** (average % epithelium=9.8%) **as well as by %non-fatty stroma** (average % non-fatty stroma=26.8%). **Association between estrogen response gene signature (EReS) and epithelial content**.

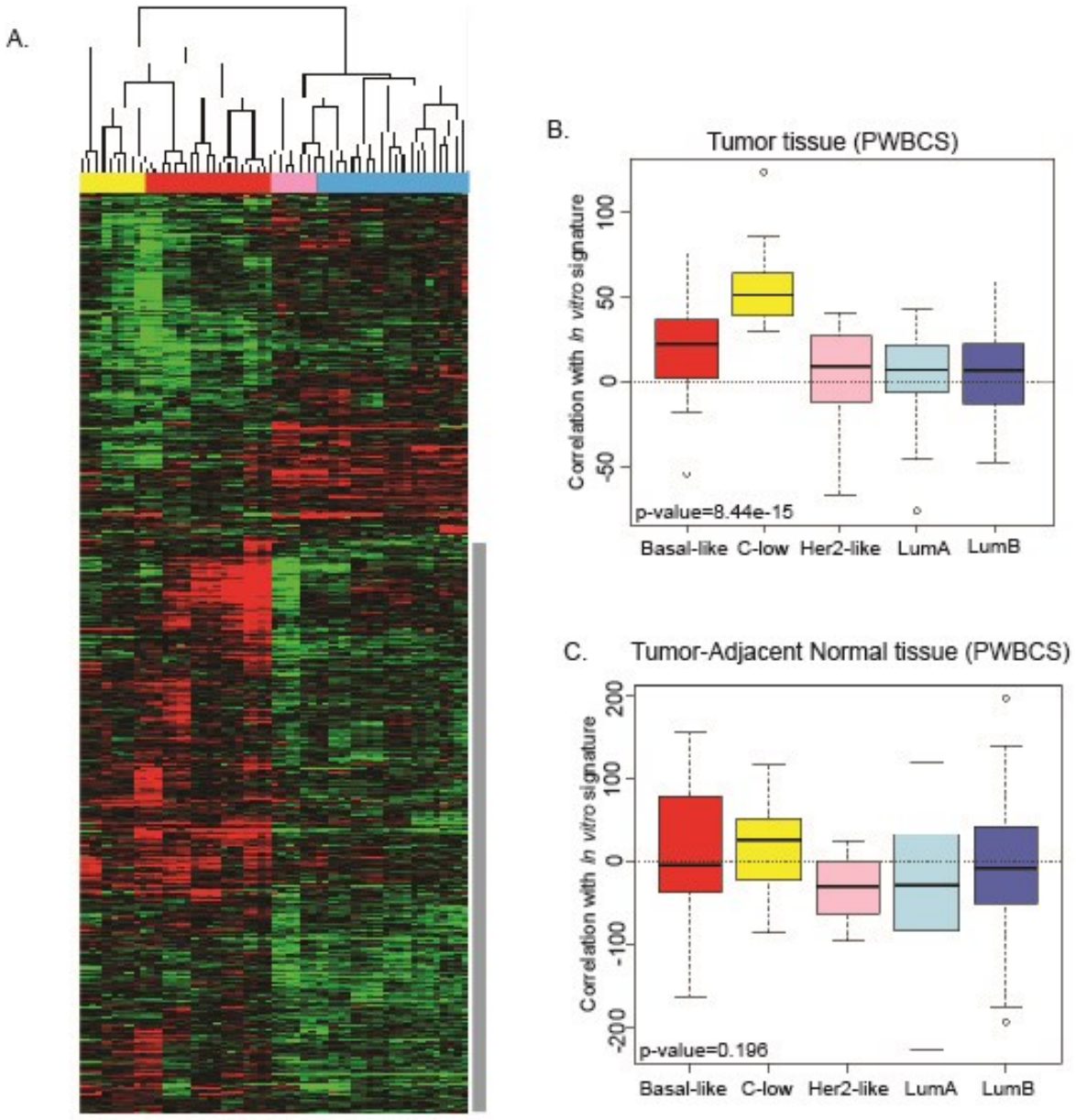


Figure 3.2. **Triple-negative microenvironments can be recapitulated *in vitro* using coculture systems.** (A) Direct *in vitro* cocultures were performed by culturing fibroblasts with cell lines models of different subtypes of breast cancer. Heat map represents interaction values (I) for genes differentially expressed in triple-negatives versus all the other cocultures. The grey bar represent genes upregulated in triple-negative tumors. B) The *in vitro* generated list of genes (grey bar in A), is highly expressed both in the intratumoral microenvironment (B) and in the cancer-adjacent tissue (C).

Category	Functions Annotation	p-Value	#Mol
Cell-To-Cell Signaling and Interaction	activation of cells	4.62E-06	34
Cellular Growth and Proliferation	proliferation of cells	1.38E-07	146
	proliferation of mononuclear leukocytes	3.09E-05	25
Cellular Movement	migration of cells	3.29E-08	75
	migration of granulocytes	2.43E-06	13
	cell movement of leukocytes	6.48E-06	32
Immune Cell Trafficking	migration of granulocytes	2.43E-06	13
	leukocyte migration	4.75E-06	34
	cell movement of leukocytes	6.48E-06	32
	transmigration of phagocytes	1.24E-05	11
	transmigration of leukocytes	1.79E-05	12
Immunological Disease	autoimmune disease	2.17E-07	84
Inflammatory Response	inflammatory response	6.37E-06	33

Table 3.8. **Biological functions associated with genes differentially expressed in Basal-like and Claudin-low cocultures.**

Ingenuity Canonical Pathways	p-value
Hepatic Fibrosis / Hepatic Stellate Cell Activation	5.9E-08
Communication between Innate and Adaptive Immune Cells	3E-06
Dendritic Cell Maturation	3.5E-06
Role of Macrophages, Fibroblasts and Endothelial Cells in Rheumatoid Arthritis	1.9E-05
Differential Regulation of Cytokine Production in Macrophages and T Helper Cells by IL-17A and IL-17F	4.4E-05
TREM1 Signaling	6.8E-05
Atherosclerosis Signaling	0.00014
p38 MAPK Signaling	0.00015
Altered T Cell and B Cell Signaling in Rheumatoid Arthritis	0.00071
Hematopoiesis from Pluripotent Stem Cells	0.00135
Role of Cytokines in Mediating Communication between Immune Cells	0.00151
NF-κB Signaling	0.00195
Glucocorticoid Receptor Signaling	0.0024
IL-6 Signaling	0.00257

Table 3.9. **Pathways associated with genes differentially expressed in Basal-like and Claudin-low cocultures**.

Chapter 4

DISCUSSION

4.1. Summary of main findings

The findings in Chapters 2 and 3 stress the importance of heterotypic interactions in breast cancer progression. Utilizing *in vitro* coculture, in tandem with *in vivo* observational studies, it is possible to increase our understanding of complementary epithelial and stromal changes in breast cancer progression. Figure 4.1 puts our findings in context of the natural history of a tumor. We showed how epithelial-stromal interactions were altered in progression of Basal-like breast cancer from benign disease to DCIS (*Chapter 2*) by combining coculture of fibroblasts with the MCF10 progression series and tumor data. We show that HGF secretion and the complementary MET overexpression occurs early in Basal-like breast cancer progression and is maintained in invasive tumors. We then showed how epithelial-stromal interactions vary with different tumor subtypes using cocultures to identify gene expression changes induced by heterotypic interactions and then evaluating these signatures *in vivo* in tumor and cancer-adjacent normal tissue (*Chapter 3*). These findings in premalignant and invasive breast lesions have possible consequences for theories of carcinogenesis and for clinical prevention of tumor recurrence.

4.2. Implications for evolutionary theories of carcinogenesis.

Biological models of cancer have traditionally emphasized epithelial characteristics, however it is clear that changes in the microenvironment are also important [6].

Evolutionary theories of carcinogenesis argue that tumors are initiated and progress with pressure from the surrounding tissue, and their survival is dependent upon acquisition of mutations that improve 'fitness' (Figure 4.2.A). In other words, these models explicitly hypothesize that epithelial cells must undergo an evolutionary adaptation to their microenvironments [8, 13, 103]. Many models have been proposed to explain how this process occurs, some propose the microenvironment is the driving force of progression, others even speculate that mutations within the stromal cells are necessary [104]. Based on our data we suggest that tumor progression is not isolated to a single compartment, but rather reflective of epithelium-stroma co-evolution. Heterotypic interactions between epithelium and stroma foster this co-evolution by selecting for complementary phenotypes therefore having long lasting consequences for the phenotypes of both compartments.

These evolutionary theories of carcinogenesis imply that cancers must overcome many hurdles the microenvironment puts on them. Physiological tissue architecture and composition or 'normal tissue contexts' are known to suppress tumor growth [8, 139]. Studies have shown that neoplastic cells introduced into normal tissue contexts will revert back to benign phenotypes [140]. Conversely, we know that during the invasive stages of cancer, the microenvironment has a promoting effects on the tumor, helping it grow and expand. Bissell *et al.* argue that the microenvironment undergoes two stages during carcinogenesis, one in which it still maintains its physiological organization and therefore suppresses tumor growth, another in which it becomes promoting of tumor growth [8] (Figure 4.3.B). In Chapter 2 and 3 we investigated the crucial dialogue between the microenvironment and the tumor during the early stages of tumorigenesis which are essential for the tumor, as well as the effects of this dialogue on the tissue directly adjacent to the tumor which may have consequences later, when the microenvironment has become permissive and may have implications for recurrence and reflect etiologic heterogeneity of the tumors.

We now know that the distinct tumor subtypes have different risk factors and present different natural histories. Development of Basal-like tumors is associated with premenopausal obesity, having children as well as having a BRCA1 mutation [141-144]; Basal-like breast cancers progress much more rapidly through the pre-invasive phases of the disease and tend to present with recurrences more often [85, 125-127]. This etiologic heterogeneity is intimately tied with the evolution and progression of tumors. The final outcome of a given tumor will be the interplay of environmental, host and tumor specific factors; in other words how the epithelial-stromal interactions evolve given the context in which they develop. Many authors argue that we can use cancer microenvironments as windows into the tumors past and potentially its future. Our findings show that (1) investigating the early events in the natural history a tumor will shed light on how and why certain factors allow a tumor to overcome the microenvironmental hurdles and (2) the cancer-adjacent tissue may provide insight into further understanding the origins of certain tumors as well as understand why and how different subtypes recur more than others.

Future work should focus on further understanding these heterotypic interactions and how they play a role in tumor progression. Specifically, tumor-stroma interactions should be studied to understand clonal expansion of neoplastic epithelial cells within tumors [145], how and why certain clones progress and others do not. It would be interesting to study if this clonal expansion also occurs among stromal cells; are certain subpopulations of fibroblasts or immune cells selected in individual tumors? Do certain subtypes of breast cancer select for stromal cells with similar alterations? Another concept that should be further investigated is the fact that at the time of diagnosis cancer patients present many micrometastasis but not all of them progress to form macrometastasis [8]. The microenvironment of the metastatic organ is clearly playing a role in this process, understanding how the heterotypic interactions allow for some metastasis to occur but not others will be essential for improving clinical management. There are many remaining

questions about evolutionary microenvironments that should be addressed. Future studies should also be designed to address the limitations of the current work. In particular, there are limitations and opportunities for future research using both coculture model systems and human tissue-based genomic methods.

4.3. Limitations and future directions

4.3.1. Coculture models of gene expression

In vitro cocultures have led to significant advances in our understanding of heterotypic interactions among different cell types, complementing mouse studies and human *in vivo* studies. However, making inferences about the relevance of coculture findings to human patients *in vivo* has limitations and requires some assumptions. Some limitations of these systems are as follows: representativeness of stromal cell lines, variation introduced by other cell types such as macrophages or mesenchymal cells, understanding the differences between 2D and 3D cultures or the limitation of only being able to studying one or few signaling pathways.

We utilized a single fibroblast line through much of our work: an immortalized reduction mammoplasty fibroblast [70]. It is clear that this cell line does not represent the full variability present in human breast tissue fibroblasts, but this system had advantages for allowing reproducibility of our experiments. It also enabled us to compare findings across cancer cell lines. While we hypothesized that stromal interactions were important, our work focused more on understanding how the variation in cancer cells modulated stromal responses. Future work could expand the number and type (for example: primary cell lines) of fibroblasts and study a single breast cancer cell line's responses to a range of stromal cells. In fact, some studies [38, 146] have already demonstrated that stromal characteristics may vary considerably from patient to patient. We focused on isolating

fibroblast-epithelial interactions in coculture and it was beyond the scope of this dissertation to study how these interactions are affected by other cell types. While inflammatory responses and cytokine milieus emerge as important biological determinants of Basal-like vs. Luminal cancer microenvironments, it will be interesting to investigate the effect of other immune cell populations on epithelial cells in the context of a coculture. Macrophages have been studied in coculture with breast cancer cell lines [50, 81, 133], but there is more to be learned in this area. The plasticity of immune cells makes this a particularly challenging area of study. For example, macrophages can react to coculture by differentiating and polarizing to a wide range of phenotypes. Understanding these changes *in vitro* is a first step to interpreting immune cell infiltration and tissue level gene expression in tumors. It will also be useful to consider specific inflammatory signals in monoculture, as previously addressed with IL-1B, [51] but adding other cells to the culture system will improve the *in vivo* relevance of these findings. Also, a variety of mesenchymal cells appear to induce similar responses in cancer cells [47, 54, 67]. It is also possible that the responses can even be induced by de-differentiating epithelial cells or by interactions between Luminal and myoepithelium. These transient phenotypes are challenging or impossible to study *in vivo*, but could be more readily manipulated in coculture systems.

It is also important to consider how the physics of the cell culture model influence cell behavior. For example, we performed 3D cocultures in a mix of matrigel and collagen and compared these findings to findings from 2D. Previous papers have shown that cells have distinct behavior in 3D environments compared to 2D environments [99]. Each system also has unique advantages for studying specific characteristics. For example, we analyzed morphogenesis in 3D, allowing us to study the integration of several cellular phenotypes in a single assay [32, 33, 97]. This experimental model also leaves gaps and results likely depend upon the type of extracellular matrix used, the stiffness present in the

matrix and the amount of time that the cell are allowed to grow [147, 148]. Therefore, studies focusing on ECM in heterotypic contexts [149], the role of mechanical forces [150] and the overall 3D architecture of the tissue [151] will add new biological insights. Biomechanical features are of established importance in cancer progression, and nanotechnology tools for dynamically altering physical environments may help address this [11, 101, 152] .

We must acknowledge that a common limitation to all cocultures is that the unit of study is often a single pathway or a small number of pathways in isolation, and typically limited to cancer cells and one other stromal cell type. Each coculture system recapitulates some aspect of the whole tissue and the interactions that are occurring *in vivo*, but given the complexity of the whole tissue, cocultures cannot fully recapitulate all dynamics and dimensions of the tissue. Assays and bioinformatics methods are available for studying interactions between two or three cell types; but incorporating more cell types has not yet been accomplished. In any case, models with increasing complexity will be needed to advance our understanding of heterotypic interactions in breast cancer.

Beyond these limitations, there is much to be learned about epithelial-stroma interactions and cocultures will play an important role in further understanding key processes in carcinogenesis. Future studies using coculture systems should address several issues, including how the microenvironment of a tumor responds to hypoxic conditions and how this affects disease progression. It is established that hypoxia occurs during tumor progression but the epithelial-stromal interactions affected by these conditions are unknown. These cocultures also could address other data types beyond gene expression, such as metabolomics. The Warburg effect, defined as energy overproduction primarily using the glycolytic pathway, is not easily studied in solid tumors and could be studied in detail in coculture models [38]. As metabolism has gained

recognition as an important driver of cancer progression [6], novel methods for studying metabolic microenvironments are needed.

Finally, there may be important translational implications of extending of cocultures to study responsiveness to cytotoxic chemotherapeutics. If signatures associated with increased toxicity were able to predict pathologic complete response *in vivo*, cocultures could help identify pathways that are promising as biomarkers or targets in neoadjuvant therapies. Given the wealth of public data from tumors and their adjacent microenvironment that has accrued in recent decades, it is becoming increasingly possible to test the relevance of *in vitro* gene expression results in patient populations at little additional cost. These types of analyses could help refine lists of candidate pathways involved in interactions to those that are most likely to play an important role in disease progression.

4.3.2. Tissue-based genomic studies.

There is an urgent need for reliable, predictive biomarkers,and these biomarkers may develop out of a better understanding of how tumors arise and progress. The majority of the studies to understand tumor biology have been done on tumor samples collected at the clinical or expression stage (Figure 4.2), which is the final step in a very long process that takes many years [8]. These studies have been extremely helpful in understanding the phenotypes of invasive tumors, but they can provide little information about the early stages of disease. Chapter 2 illustrates that some changes observed in invasive Basal-like tumors *in vivo* (e.g. HGF signaling) occur early. By examining only invasive tumors, we would have missed that this signaling pathway becomes important much earlier in the disease progression. Additional tissue samples, collected from earlier stages of progression such as atypical hyperplasia or DCIS, could help to further ascertain the

relevance of these pathways in early lesions. In an ideal study, researchers collect tissue samples in a longitudinal design, beginning with benign or histologically normal tissues (during the induction period,) and study how they are altered over time, ultimately leading to tumor formation. This design is impractical for ethical and other reasons, however we argue that a good surrogate is cross-sectional study of many people at different stages, including analysis of the tissue adjacent to cancer. These cross-sectional and cancer-adjacent tissues are easily obtained and can reflect the causal action period providing a window into the natural history of the cancer [73, 153] (Figure 4.2.). The 'genomic fingerprints' of lifetime exposures and risks may be detectable within these normal and benign tissues, prior to onset of genomic instability [80]. Genomic instability may mask the etiologically-relevant and early progression-relevant changes. The results presented in Chapter 3 represent how we could utilize cancer-adjacent gene expression patterns to further understand breast cancer etiology. Future studies and analysis should devote more efforts to the collection of these samples, with a focus on understanding the etiology of this heterogeneous disease, and understanding how risk factors and host response to the cancer work together to alter the normal tissue

4.4. Translational implications of findings

In these studies we have analyzed epithelial and stromal changes that occur in early lesions and in tissue adjacent to tumors. Specifically, our results suggest that 1) HGF/MET signaling is a strong candidate pathway for treating premalignant Basal-like lesions and the application of MET inhibitors should be considered in preclinical models to advance this plausible strategy, 2) understanding the heterogeneous etiology of breast cancer as well as the responses to the tumors will help design better treatments. For example, triple-negative breast cancers may warrant wider margins during surgery given

the expression of aggressive gene signatures present in the cancer-adjacent tissue. Together these advances contribute to our larger goal: to better understand cancer incidence, identify new ways of preventing the disease, and improving clinical management of patients with breast cancer.

4.5. Figures

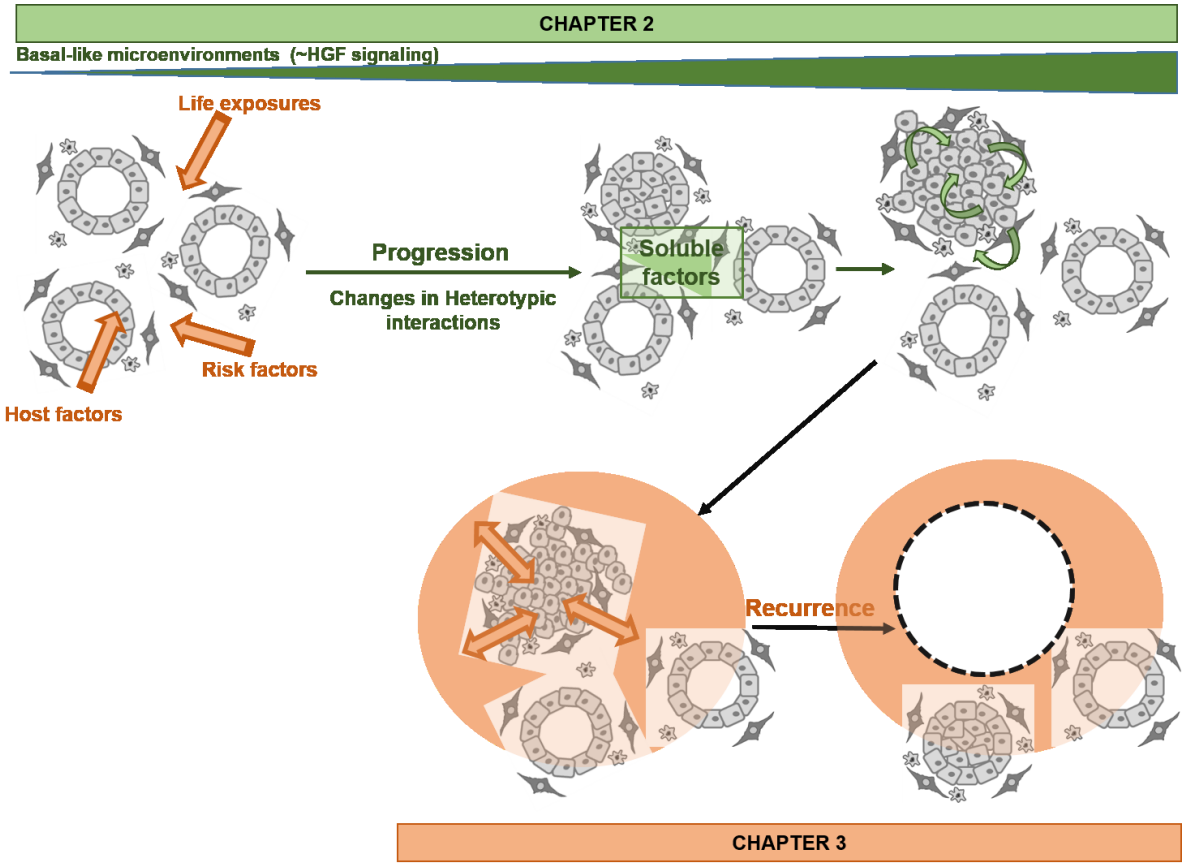


Figure 4.1. **Summary of findings**. Schematic representing how the studies from this document are integrated into the natural history of a tumor. Highlighted in green are the major finding of Chapter 2 (*Role of HGF in epithelial-stromal cell interactions during the progression from benign disease to DCIS*). We show how epithelial-stromal interactions are altered in progression of Basal-like breast cancer from benign disease to DCIS, that HGF secretion and the complementary MET overexpression occurs early Basal-like breast cancer progression and that it is maintained in invasive tumors. Highlighted in orange are the major finding of Chapter 3 (*Tumor intrinsic subtype is reflected in the cancer-adjacent tissue*), we show how epithelial-stromal interactions vary with different tumor subtypes using cocultures to identify gene expression changes induced by heterotypic interactions and then evaluating these signatures *in vivo* in tumor and cancer-adjacent normal tissue. These alterations may have consequences for breast cancer recurrence.

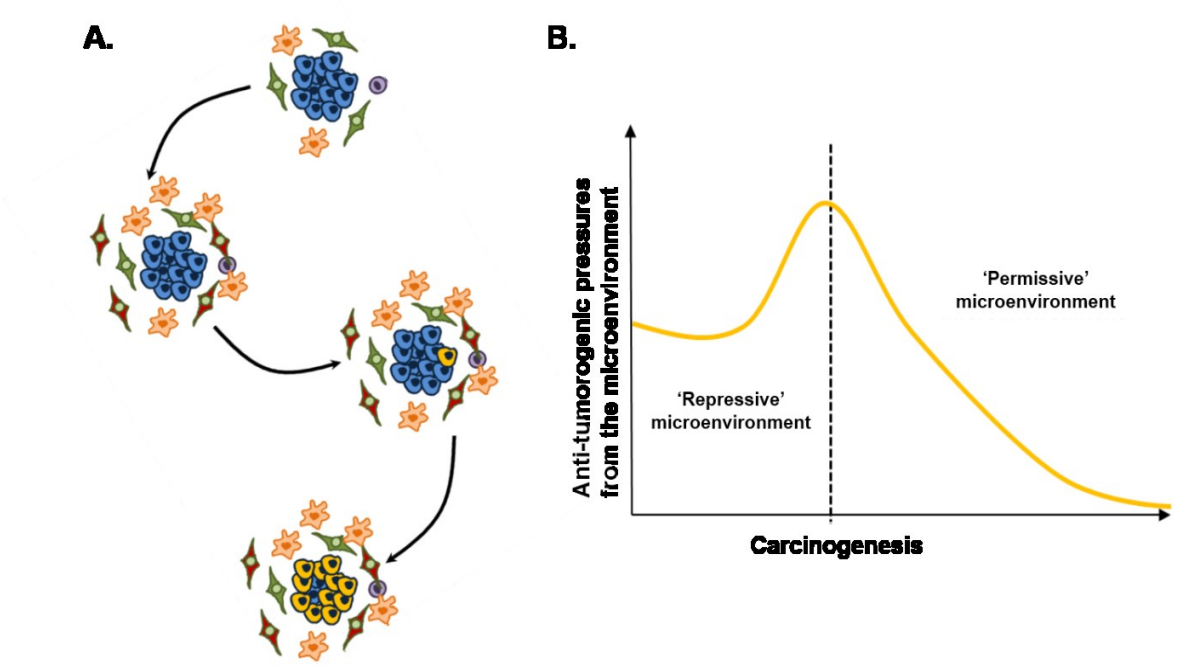


Figure 4.2. **Evolutionary theories of carcinogenesis**. A) Schematic representing the clonal evolution of a small tumor in response to a change in the microenvironment. B) Diagram adapted from Bissel et al. (cite), it shows the anti-tumorigenic pressures the carcinogenic process encounters because of the microenvironment.

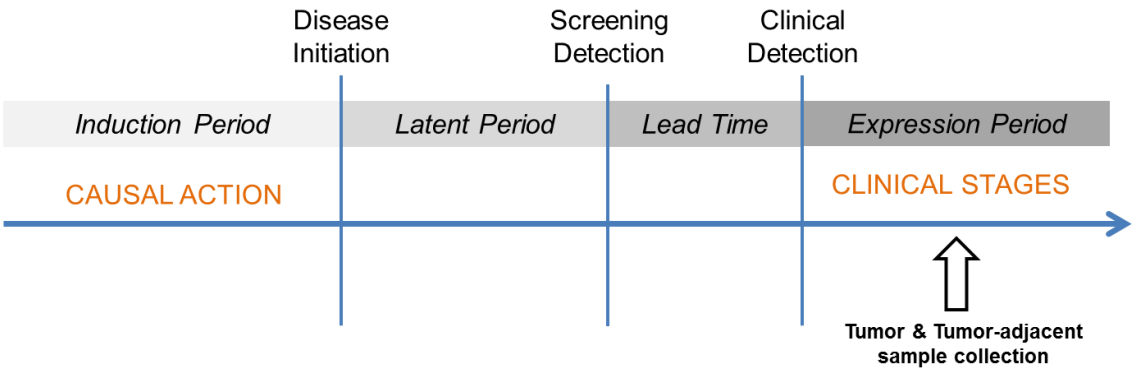


Figure 4.3. **Natural history of breast cancer**. A tumor undergoes many stages from a clinical perspective. There is an initial stage in which risk factors and genetic susceptibility are interacting with the tissue until a given cell undergoes the first driver mutation that will lead to cancer (disease initiation). The Latent period is when, the disease is progressing but we cannot detect it clinically. Lead time is very dependent on tumor site and the available technologies, it represent the time between when we can detect the tumor (in the case of breast cancer through mammography) and the time in which it become clinically detectable. After a tumor is clinically detectable is when tumor samples are usually collected; this is also the time in which the tumor-adjacent sample should be collected. Collection of this tissue will increase our understanding of the preceeding stages.

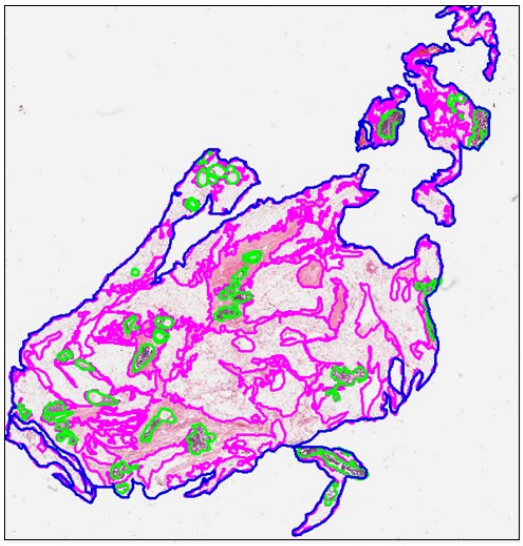
Appendix 1**. BREAST TISSUE COMPOSITION MEASURMENTS**.

Frozen non-neoplastic breast specimens of approximately 100 mg were cut over dry ice and then used to cut frozen sections. Sections were collected at both ends of the specimen and then constructed into 20 µm slides. The central portion was used for RNA extraction. After H&E staining, the slides were scanned into high-resolution digital images using the Aperio Scan-Scope XT Slide Scanner (Aperio Technologies, Vista, CA, USA) in the UNC Translational Pathology Laboratory. After excluding slides with poor resolution or having folded tissues, 118 (97.5%) slides were subjected to breast tissue composition analysis. To train the composition estimator in Aperio's Genie software, 15 representative H&E slides were selected and evaluated by eye by a pathologist who provided semi-quantitative estimates of the percentage of adipose tissue (10% bin width), epithelium (1% bin width), and non-fatty stroma (10% bin width). Each pathologist-scored slide was also manually annotated for epithelial area, stromal area, and total area (mm^2) using Aperio ImageScope software. These digital area-based, quantitative estimates were used to train Aperio's Genie Classifier to partition epithelium, adipose tissue, non-fatty stroma, and glass into percentages. Examples of annotated digital images are presented in Figure 1 of this Appendix. The trained classifier was strongly correlated with pathologist review, particularly for stroma and adipose (Table 1 of this Appendix, Pearson correlation coefficient ranged 0.95-0.96). The trained classifier was also positively and strongly correlated with manually scored area, for all three tissue compartments. Compared with digital assessment, visual assessment (by human eye) of small percentage differences is weaker, and thus the digital image analysis data were used in analyses of epithelial tissue, which is sparse (<10%) in benign breast. The trained Genie Classifier was then applied to the remaining slides.

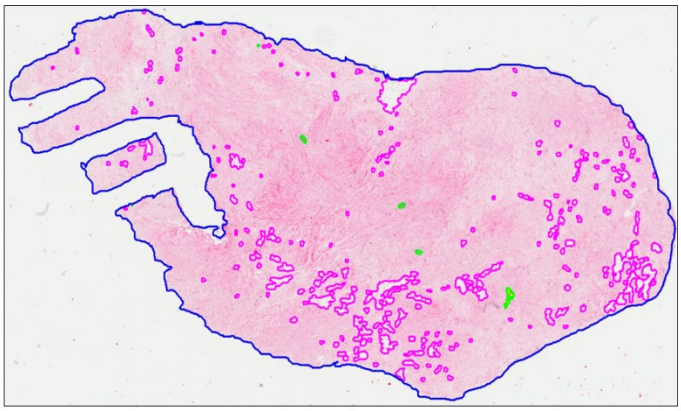
	Epithelium	Adipose	Stroma
Manual &Genie	0.98	0.96	0.96
Manual & Pathologist	0.71	0.96	0.95
Pathologist & Genie	0.68	0.96	0.96

Appendix 1 Table 1. **Pearson correlation coefficient between evaluation approaches by breast composition** (n=118).

A.



B.



Appendix 1 Figure 1. **Representative pictures of breast composition by digital pathological image analysis**. (Pink=adipose tissue, Green=epithelial tissue and Blue= outline of specimen) A) Adipose=62.93% Epithelium=6.68% Stroma=30.38%, B) Adipose=5.14% Epithelium=0.09% Stroma=94.80%

REFERENCES

1. Perou, C.M., *Molecular Stratification of Triple-Negative Breast Cancers.* The Oncologist, 2011. **16**(suppl 1): p. 61-70.

2. Prat, A., et al., *Phenotypic and molecular characterization of the claudin-low intrinsic subtype of breast cancer.* Breast Cancer Research, 2010. **12**(5): p. R68.

3. Perou, C.M., et al., *Molecular portraits of human breast tumours.* Nature, 2000. **406**(6797): p. 747-752.

4. Parker, J.S., et al., *Supervised Risk Predictor of Breast Cancer Based on Intrinsic Subtypes.* Journal of Clinical Oncology, 2009. **27**(8): p. 1160-1167.

5. Camp, J.T., et al., *Interactions with Fibroblasts Are Distinct in Basal-Like and Luminal Breast Cancers.* Molecular Cancer Research, 2011. **9**(1): p. 3-13.

6. Hanahan, D. and Robert A. Weinberg, *Hallmarks of Cancer: The Next Generation.* Cell, 2011. **144**(5): p. 646-674.

7. Joyce, J.A. and J.W. Pollard, *Microenvironmental regulation of metastasis.* Nat Rev Cancer, 2009. **9**(4): p. 239-252.

8. Bissell, M.J. and W.C. Hines, *Why don't we get more cancer? A proposed role of the microenvironment in restraining cancer progression.* Nat Med, 2011. **17**(3): p. 320-329.

9. Poczobutt, J., et al., *Benign mammary epithelial cells enhance the transformed phenotype of human breast cancer cells.* BMC Cancer, 2010. **10**(1): p. 373.

10. Mueller, M.M. and N.E. Fusenig, *Friends or foes [mdash] bipolar effects of the tumour stroma in cancer.* Nat Rev Cancer, 2004. **4**(11): p. 839-849.

11. Wiseman, B.S. and Z. Werb, *Stromal Effects on Mammary Gland Development and Breast Cancer.* Science, 2002. **296**(5570): p. 1046-1049.

12. Dvorak, H.F., *Tumors: Wounds That Do Not Heal.* New England Journal of Medicine, 1986. **315**(26): p. 1650-1659.

13. Gatenby, R.A., Gillies, Robert J., *A microenvironmental model of carcinogenesis.* Nat Rev Cancer, 2008. **8**(1).

14. Ma, X.-J., et al., *Gene expression profiling of the tumor microenvironment during breast cancer progression.* Breast Cancer Research, 2009. **11**(1): p. R7.

15. Schedin, P. and P.J. Keely, *Mammary Gland ECM Remodeling, Stiffness, and Mechanosignaling in Normal Development and Tumor Progression.* Cold Spring Harbor Perspectives in Biology, 2011. **3**(1).

16. Schedin, P., et al., *Mammary ECM composition and function are altered by reproductive state.* Molecular Carcinogenesis, 2004. **41**(4): p. 207-220.

17. Hattar, R., et al., *Tamoxifen induces pleiotrophic changes in mammary stroma resulting in extracellular matrix that suppresses transformed phenotypes.* Breast Cancer Research, 2009. **11**(1): p. R5.

18. Roskelley, C.D. and M.J. Bissell, *The dominance of the microenvironment in breast and ovarian cancer.* Seminars in Cancer Biology, 2002. **12**(2): p. 97-104.

19. Cichon, M., et al., *Microenvironmental Influences that Drive Progression from Benign Breast Disease to Invasive Breast Cancer.* Journal of Mammary Gland Biology and Neoplasia, 2010. **15**(4): p. 389-397.

20. Albini, A. and M.B. Sporn, *The tumour microenvironment as a target for chemoprevention.* Nat Rev Cancer, 2007. **7**(2): p. 139-147.

21. Trujillo, K.A., et al., *Markers of fibrosis and epithelial to mesenchymal transition demonstrate field cancerization in histologically normal tissue adjacent to breast tumors.* International Journal of Cancer, 2011: p. n/a-n/a.

22. DP, S., S. HW, and S. W, *Field cancerization in oral stratified squamous epithelium; clinical implications of multicentric origin.* Cancer, 1953. **5**: p. 963-968.

23. Heaphy, C., J. Griffith, and M. Bisoffi, *Mammary field cancerization: molecular evidence and clinical importance.* Breast Cancer Research and Treatment, 2009. **118**(2): p. 229-239.

24. Troester, M.A., et al., *Cell-Type-Specific Responses to Chemotherapeutics in Breast Cancer.* Cancer Research, 2004. **64**(12): p. 4218-4226.

25. Neve, R.M., et al., *A collection of breast cancer cell lines for the study of functionally distinct cancer subtypes.* Cancer cell, 2006. **10**(6): p. 515-527.

26. Hoadley, K., et al., *EGFR associated expression profiles vary with breast tumor subtype.* BMC Genomics, 2007. **8**(1): p. 258.

27. Troester, M., et al., *Gene expression patterns associated with p53 status in breast cancer.* BMC Cancer, 2006. **6**(1): p. 276.

28. Miller, L.D., et al., *An expression signature for p53 status in human breast cancer predicts mutation status, transcriptional effects, and patient survival.* Proceedings of the National Academy of Sciences of the United States of America, 2005. **102**(38): p. 13550-13555.

29. Astrahantseff, K.N. and J.E. Morris, *Estradiol-17β Stimulates Proliferation of Uterine Epithelial Cells Cultured with Stromal Cells but Not Cultured Separately.* In Vitro Cellular & Developmental Biology. Animal, 1994. **30A**(11): p. 769-776.

30. Weigelt, B. and M.J. Bissell, *Unraveling the microenvironmental influences on the normal mammary gland and breast cancer.* Seminars in Cancer Biology, 2008. **18**(5): p. 311-321.

31. Hebner, C., V.M. Weaver, and J. Debnath, *Modeling Morphogenesis and Oncogenesis in Three-Dimensional Breast Epithelial Cultures.* Annual Review of Pathology: Mechanisms of Disease, 2008. **3**(1): p. 313-339.

32. Brugge, J., *Into the deep: Refocusing on 3D.* Nay Cell Biol, 2012. **14**(4): p. 332.

33. Debnath, J., S.K. Muthuswamy, and J.S. Brugge, *Morphogenesis and oncogenesis of MCF-10A mammary epithelial acini grown in three-dimensional basement membrane cultures.* Methods, 2003. **30**(3): p. 256-268.

34. Weigelt, B., et al., *HER2 signaling pathway activation and response of breast cancer cells to HER2-targeting agents is dependent strongly on the 3D microenvironment.* Breast Cancer Research and Treatment, 2010. **122**(1): p. 35-43.

35. Q, Y., et al., *CLIC4 mediates TGF-beta1-induced fibroblast-to-myofibroblast transdifferentiation in ovarian cancer.* Oncol. Rep., 2009. **3**: p. 541-548.

36. Paszek, M.J., et al., *Tensional homeostasis and the malignant phenotype.* Cancer cell, 2005. **8**(3): p. 241-254.

37. Krtolica A, C.J., *Integrating epithelial cancer, aging stroma and cellular senescence.* Adv. Gerontol., 2003. **11**: p. 109-16.

38. Brauer, H.A., et al., *Impact of Tumor Microenvironment and Epithelial Phenotypes on Metabolism in Breast Cancer.* Clinical Cancer Research, 2013. **19**(3): p. 571-585.

39. Bell, J.A., et al., *Lipid Partitioning, Incomplete Fatty Acid Oxidation, and Insulin Signal Transduction in Primary Human Muscle Cells: Effects of Severe Obesity, Fatty Acid Incubation, and Fatty Acid Translocase/CD36 Overexpression.* Journal of Clinical Endocrinology & Metabolism, 2010. **95**(7): p. 3400-3410.

40. Sampey, B.P., et al., *Genistein effects on stromal cells determines epithelial proliferation in endometrial co-cultures.* Experimental and Molecular Pathology, 2011. **90**(3): p. 257-263.

41. Fleming, J., et al., *The normal breast microenvironment of premenopausal women differentially influences the behavior of breast cancer cells in vitro and in vivo.* BMC Medicine, 2010. **8**(1): p. 27.

42. Rozenchan, P.B., et al., *Reciprocal changes in gene expression profiles of cocultured breast epithelial cells and primary fibroblasts.* International Journal of Cancer, 2009. **125**(12): p. 2767-2777.

43. Santos, R., et al., *Influence of the interaction between nodal fibroblast and breast cancer cells on gene expression.* Tumor Biology, 2011. **32**(1): p. 145-157.

44. Wadlow, R.C., et al., *Systems-Level Modeling of Cancer-Fibroblast Interaction.* PLoS ONE, 2009. **4**(9): p. e6888.

45. Hayashi, S.-i. and Y. Yamaguchi, *Estrogen signaling in cancer microenvironment and prediction of response to hormonal therapy.* The Journal of Steroid Biochemistry and Molecular Biology, 2008. **109**(3-5): p. 201-206.

46. Lin, H.-J.L., et al., *Breast Cancer–Associated Fibroblasts Confer AKT1-Mediated Epigenetic Silencing of Cystatin M in Epithelial Cells.* Cancer Research, 2008. **68**(24): p. 10257-10266.

47. Buess, M., et al., *Characterization of heterotypic interaction effects in vitro to deconvolute global gene expression profiles in cancer.* Genome Biology, 2007. **8**(9): p. R191.

48. Gao, M.-Q., et al., *Stromal fibroblasts from the interface zone of human breast carcinomas induce an epithelial–mesenchymal transition-like state in breast cancer cells in vitro.* Journal of Cell Science, 2010. **123**(20): p. 3507-3514.

49. Buess M, R.M., Vogel-Durrer BM, Herrmann R, Rochlitz C., *Tumor-endothelial interaction links the CD44(+)/CD24(-) phenotype with poor prognosis in early-stage breast cancer.* Neoplasia, 2009. **10**: p. 987-1002.

50. Hagemann, T., et al., *Enhanced invasiveness of breast cancer cell lines upon co-cultivation with macrophages is due to TNF-α dependent up-regulation of matrix metalloproteases.* Carcinogenesis, 2004. **25**(8): p. 1543-1549.

51. Hou, Z., et al., *Macrophages induce COX-2 expression in breast cancer cells: role of IL-1β autoamplification.* Carcinogenesis, 2011. **32**(5): p. 695-702.

52. DeNardo, D.G., et al., *CD4+ T Cells Regulate Pulmonary Metastasis of Mammary Carcinomas by Enhancing Protumor Properties of Macrophages.* Cancer cell, 2009. **16**(2): p. 91-102.

53. Gauthier, M.L., et al., *Abrogated Response to Cellular Stress Identifies DCIS Associated with Subsequent Tumor Events and Defines Basal-like Breast Tumors.* Cancer cell, 2007. **12**(5): p. 479-491.

54. Liu, S., et al., *Breast Cancer Stem Cells Are Regulated by Mesenchymal Stem Cells through Cytokine Networks.* Cancer Research, 2011. **71**(2): p. 614-624.

55. Martin, F., et al., *Potential role of mesenchymal stem cells (MSCs) in the breast tumour microenvironment: stimulation of epithelial to mesenchymal transition (EMT).* Breast Cancer Research and Treatment, 2010. **124**(2): p. 317-326.

56. Michal Rajski, et al., *Global gene expression analysis of the interaction between cancer cells and osteoblasts to predict bone metastasis in breast cancer.* PLoS One, 2012.

57. Tabaris, S., et al., *Claudin-2 Promotes Breast Cancer Liver Metastasis by Facilitating Tumor Cell Interactions with Hepatocytes.* Molecular and Cellular Biology, 2012. **32**(15): p. 2979-2991.

58. Lim, P.K., et al., *Gap Junction–Mediated Import of MicroRNA from Bone Marrow Stromal Cells Can Elicit Cell Cycle Quiescence in Breast Cancer Cells.* Cancer Research, 2011. **71**(5): p. 1550-1560.

59. Luciani, M.G., et al., *Distinctive Responsiveness to Stromal Signaling Accompanies Histologic Grade Programming of Cancer Cells.* PLoS ONE, 2011. **6**(5): p. e20016.

60. Potter, S., et al., *Influence of stromal–epithelial interactions on breast cancer in vitro and in vivo.* Breast Cancer Research and Treatment, 2011: p. 1-11.

61. Tobar, N., et al., *NOX4-dependent ROS production by stromal mammary cells modulates epithelial MCF-7 cell migration.* Br J Cancer, 2010. **103**(7): p. 1040-1047.

62. Rhodes, L., et al., *Effects of human mesenchymal stem cells on ER-positive human breast carcinoma cells mediated through ER-SDF-1/CXCR4 crosstalk.* Molecular Cancer, 2010. **9**(1): p. 295.

63. Dirat, B., et al., *Cancer-Associated Adipocytes Exhibit an Activated Phenotype and Contribute to Breast Cancer Invasion.* Cancer Research, 2011. **71**(7): p. 2455-2465.

64. Dontu, G., et al., *In vitro propagation and transcriptional profiling of human mammary stem/progenitor cells.* Genes & Development, 2003. **17**(10): p. 1253-1270.

65. Tyan, S.-W., et al., *Breast Cancer Cells Induce Cancer-Associated Fibroblasts to Secrete Hepatocyte Growth Factor to Enhance Breast Tumorigenesis.* PLoS ONE, 2011. **6**(1): p. e15313.

66. Mailleux, A.A., M. Overholtzer, and J.S. Brugge, *Lumen formation during mammary epithelial morphogenesis: insights from in vitro and in vivo models.* Cell Cycle, 2008. **7**(1): p. 57-62.

67. Klopp, A.H., et al., *Mesenchymal Stem Cells Promote Mammosphere Formation and Decrease E-Cadherin in Normal and Malignant Breast Cells.* PLoS ONE, 2010. **5**(8): p. e12180.

68. Noël A, D.P.-G.M., Purnell G, Nusgens B, Lapiere CM, Foidart JM., *Enhancement of tumorigenicity of human breast adenocarcinoma cells in nude mice by matrigel and fibroblasts.* Br J Cancer, 1993. **68(5)**: p. 909-915.

69. Hu, M., et al., *Regulation of In Situ to Invasive Breast Carcinoma Transition.* Cancer cell, 2008. **13**(5): p. 394-406.

70. Proia, D.A. and C. Kuperwasser, *Reconstruction of human mammary tissues in a mouse model.* Nat. Protocols, 2006. **1**(1): p. 206-214.

71. Behbod, F., et al., *An intraductal human-in-mouse transplantation model mimics the subtypes of ductal carcinoma in situ.* Breast Cancer Research, 2009. **11**(5): p. R66.

72. Wright, M., et al., *Molecular analysis reveals heterogeneity of mouse mammary tumors conditionally mutant for Brca1.* Molecular Cancer, 2008. **7**(1): p. 29.

73. Troester, M.A., et al., *Activation of Host Wound Responses in Breast Cancer Microenvironment.* Clinical Cancer Research, 2009. **15**(22): p. 7020-7028.

74. Chang, H.Y., et al., *Robustness, scalability, and integration of a wound-response gene expression signature in predicting breast cancer survival.* Proceedings of the National Academy of Sciences of the United States of America, 2005. **102**(10): p. 3738-3743.

75. Finak, G., et al., *Stromal gene expression predicts clinical outcome in breast cancer.* Nat Med, 2008. **14**(5): p. 518-527.

76. Allinen, M., et al., *Molecular characterization of the tumor microenvironment in breast cancer.* Cancer cell, 2004. **6**(1): p. 17-32.

77. Chang, H.Y., et al., *Gene Expression Signature of Fibroblast Serum Response Predicts Human Cancer Progression: Similarities between Tumors and Wounds.* PLoS Biol, 2004. **2**(2): p. e7.

78. Beck, A.H., et al., *The Macrophage Colony-Stimulating Factor 1 Response Signature in Breast Carcinoma.* Clinical Cancer Research, 2009. **15**(3): p. 778-787.

79. Beck, A.H., et al., *The fibromatosis signature defines a robust stromal response in breast carcinoma.* Lab Invest, 2008. **88**(6): p. 591-601.

80. Roman-Perez, E., et al., *Gene expression in extratumoral microenvironment predicts clinical outcome in breast cancer patients.* Breast Cancer Research, 2011. **14**(2): p. R51.

81. Stewart, D.A., et al., *Basal-like Breast Cancer Cells Induce Phenotypic and Genomic Changes in Macrophages.* Molecular Cancer Research, 2012. **10**(6): p. 727-738.

82. Gillies, R.A.G.R.J., *A microenvironmental model of carcinogenesis.* Nature Reviews Cancer, 2008. **8**: p. 56-61.

83. Pirone, J.R., et al., *Age-Associated Gene Expression in Normal Breast Tissue Mirrors Qualitative Age-at-Incidence Patterns for Breast Cancer.* Cancer Epidemiology Biomarkers & Prevention, 2012. **21**(10): p. 1735-1744.

84. Shiao, S.L., et al., *Immune microenvironments in solid tumors: new targets for therapy.* Genes & Development, 2011. **25**(24): p. 2559-2572.

85. Livasy, C.A., et al., *Identification of a basal-like subtype of breast ductal carcinoma in situ.* Human pathology, 2007. **38**(2): p. 197-204.

86. Tusher, V.G., R. Tibshirani, and G. Chu, *Significance analysis of microarrays applied to the ionizing radiation response.* Proceedings of the National Academy of Sciences, 2001. **98**(9): p. 5116-5121.

87. Creighton, C.J., et al., *Insulin-Like Growth Factor-I Activates Gene Transcription Programs Strongly Associated With Poor Breast Cancer Prognosis.* Journal of Clinical Oncology, 2008. **26**(25): p. 4078-4085.

88. van de Vijver, M.J., et al., *A Gene-Expression Signature as a Predictor of Survival in Breast Cancer.* New England Journal of Medicine, 2002. **347**(25): p. 1999-2009.

89. Naderi, A.T., A E Barbosa-Morais, N L Pinder, S E Green, A R Powe, D G Robertson, J F R Aparicio, S Ellis, I O Brenton, J D Caldas, C, *A gene-expression signature to predict survival in breast cancer across independent data sets.* Oncogene, 2006. **26**(10).

90. Chhetri, R.K., et al., *Longitudinal Study of Mammary Epithelial and Fibroblast Co-Cultures Using Optical Coherence Tomography Reveals Morphological Hallmarks of Pre-Malignancy.* PLoS ONE, 2012. **7**(11): p. e49148.

91. Johnson, K.R., et al., *Demystifying the Effects of a Threeâ€Dimensional Microenvironment in Tissue Morphogenesis*, in *Methods in Cell Biology*. 2007, Academic Press. p. 547-583.

92. Jedeszko, C., et al., *Fibroblast Hepatocyte Growth Factor Promotes Invasion of Human Mammary Ductal Carcinoma In situ.* Cancer Research, 2009. **69**(23): p. 9148-9155.

93. West, N.J., et al., *Apoptosis in the Colonic Crypt, Colorectal Adenomata, and Manipulation by Chemoprevention.* Cancer Epidemiology Biomarkers & Prevention, 2009. **18**(6): p. 1680-1687.

94. Kataoka, H., et al., *Hepatocyte Growth Factor Activator Inhibitor Type 1 Is a Specific Cell Surface Binding Protein of Hepatocyte Growth Factor Activator (HGFA) and Regulates HGFA Activity in the Pericellular Microenvironment.* Journal of Biological Chemistry, 2000. **275**(51): p. 40453-40462.

95. Stella, M.C., et al., *Negative Feedback Regulation of Met-Dependent Invasive Growth by Notch.* Molecular and Cellular Biology, 2005. **25**(10): p. 3982-3996.

96. AT, L., et al., *Constructing three-dimensional models to study mammary gland branching morphogenesis and functional differentiation.* Journal of Mammary Gland Biology and Neoplasia, 2012.

97. Mailleux A, O.M., Brugge J. , *Lumen formation during mammary epithelial morphogenesis: insights from in vitro and in vivo models. .* Cell Cycle, 2008. **7**.

98. Nelson, C.M. and M.J. Bissell, *Modeling dynamic reciprocity: Engineering three-dimensional culture models of breast architecture, function, and neoplastic transformation.* Seminars in Cancer Biology, 2005. **15**(5): p. 342-352.

99. Kenny, P.A., et al., *The morphologies of breast cancer cell lines in three-dimensional assays correlate with their profiles of gene expression.* Molecular Oncology, 2007. **1**(1): p. 84-96.

100. Whelan, K.A., et al., *Hypoxia Suppression of Bim and Bmf Blocks Anoikis and Luminal Clearing during Mammary Morphogenesis.* Molecular Biology of the Cell, 2010. **21**(22): p. 3829-3837.

101. Raghav K. Chhetri, Z.F.P., Melissa A. Troester, Amy L. Oldenburg, *Longitudinal Study of Mammary Epithelial and Fibroblast Co-Cultures using Optical Coherence Tomography Reveals Morphological Hallmarks of Pre-Malignancy.* Plos ONE, In press.

102. Castro, N., et al., *Evidence that molecular changes in cells occur before morphological alterations during the progression of breast ductal carcinoma.* Breast Cancer Research, 2008. **10**(5): p. R87.

103. Kareca, I., *What can Ecology teach Us About Cancer?* Translational Onocology, 2011. **4**(5): p. 266-279.

104. Gatenby, R.A.G., Robert J. Brown, Joel S., *Of cancer and cave fish.* Nature Publishing group, 2011. **11**(4): p. 237-238.

105. Valdez, K.E., et al., *Human primary ductal carcinoma in situ (DCIS) subtype-specific pathology is preserved in a mouse intraductal (MIND) xenograft model.* The Journal of Pathology, 2011. **225**(4): p. 565-573.

106. P. J. Dawson, S.R.W., L. Tait, G. H. Heppner, F. R. Miller, *MCF10AT: a model for the evolution of cancer from proliferative breast disease.* The American Journal of Pathology, 1996. **1**: p. 313-319.

107. Wright, T.G., et al., *Increased production and secretion of HGF α-chain and an antagonistic HGF fragment in a human breast cancer progression model.* International Journal of Cancer, 2009. **125**(5): p. 1004-1015.

108. Ma, J., et al., *Somatic mutation and functional polymorphism of a novel regulatory element in the HGF gene promoter causes its aberrant expression in human breast cancer.* The Journal of Clinical Investigation, 2009. **119**(3): p. 478-491.

109. Khoury, H., et al., *HGF Converts ErbB2/Neu Epithelial Morphogenesis to Cell Invasion.* Molecular Biology of the Cell, 2005. **16**(2): p. 550-561.

110. Lamorte, L., et al., *Crk Adapter Proteins Promote an Epithelial-Mesenchymal-like Transition and Are Required for HGF-mediated Cell Spreading and Breakdown of Epithelial Adherens Junctions.* Molecular Biology of the Cell, 2002. **13**(5): p. 1449-1461.

111. Ghoussoub, R.A.D., et al., *Expression of c-met is a strong independent prognostic factor in breast carcinoma.* Cancer, 1998. **82**(8): p. 1513-1520.

112. Timothy R. Wilson, et al., *Widespread potential for growth-factor-driven resistance to anticancer kinase inhibitors.* Nature, 2012. **487**: p. 505-509.

113. Straussman, R., et al., *Tumour micro-environment elicits innate resistance to RAF inhibitors through HGF secretion.* Nature, 2012. **487**: p. 500-504.

114. Ponzo, M.G., et al., *Met induces mammary tumors with diverse histologies and is associated with poor outcome and human basal breast cancer.* Proceedings of the National Academy of Sciences, 2009. **106**(31): p. 12903-12908.

115. Niranjan, B., et al., *HGF/SF: a potent cytokine for mammary growth, morphogenesis and development.* Development, 1995. **121**(9): p. 2897-2908.

116. Garner, O.B., et al., *Stage-dependent regulation of mammary ductal branching by heparan sulfate and HGF-cMet signaling.* Developmental Biology, 2011. **355**(2): p. 394-403.

117. Meyer, G., et al., *A Potential Role of Progestin-Induced Laminin-5/6-Integrin Signaling in the Formation of Side Branches in the Mammary Gland.* Endocrinology, 2012. **153**(10): p. 4990-5001.

118. Conway, K., et al., *The molecular and clinical impact of hepatocyte growth factor, its receptor, activators, and inhibitors in wound healing.* Wound Repair and Regeneration, 2006. **14**(1): p. 2-10.

119. Huh, C.-G., et al., *Hepatocyte growth factor/c-met signaling pathway is required for efficient liver regeneration and repair.* Proceedings of the National Academy of Sciences of the United States of America, 2004. **101**(13): p. 4477-4482.

120. Lindemann, K., et al., *Differential expression of c-Met, its ligand HGF/SF and HER2/neu in DCIS and adjacent normal breast tissue.* Histopathology, 2007. **51**(1): p. 54-62.

121. Veronesi, U., et al., *Twenty-Year Follow-up of a Randomized Study Comparing Breast-Conserving Surgery with Radical Mastectomy for Early Breast Cancer.* New England Journal of Medicine, 2002. **347**(16): p. 1227-1232.

122. Fisher, B., et al., *Twenty-Year Follow-up of a Randomized Trial Comparing Total Mastectomy, Lumpectomy, and Lumpectomy plus Irradiation for the Treatment of Invasive Breast Cancer.* New England Journal of Medicine, 2002. **347**(16): p. 1233-1241.

123. Silverstein, M.J., et al., *The Influence of Margin Width on Local Control of Ductal Carcinoma in Situ of the Breast.* New England Journal of Medicine, 1999. **340**(19): p. 1455-1461.

124. Arvold, N.D., et al., *Age, Breast Cancer Subtype Approximation, and Local Recurrence After Breast-Conserving Therapy.* Journal of Clinical Oncology, 2011. **29**(29): p. 3885-3891.

125. Gabos, Z., et al., *The association between biological subtype and locoregional recurrence in newly diagnosed breast cancer.* Breast Cancer Research and Treatment, 2010. **124**(1): p. 187-194.

126. Voduc, K.D., et al., *Breast Cancer Subtypes and the Risk of Local and Regional Relapse.* Journal of Clinical Oncology, 2010. **28**(10): p. 1684-1691.

127. Abdulkarim, B.S., et al., *Increased Risk of Locoregional Recurrence for Women With T1-2N0 Triple-Negative Breast Cancer Treated With Modified Radical Mastectomy Without Adjuvant Radiation Therapy Compared With Breast-Conserving Therapy.* Journal of Clinical Oncology, 2011. **29**(21): p. 2852-2858.

128. Yang, X., et al., *Analysis of terminal duct lobular unit involution in luminal A and basal breast cancers.* Breast Cancer Research, 2013. **14**(2): p. R64.

129. Graham, K., et al., *Gene Expression Profiles of Estrogen Receptor Positive and Estrogen Receptor Negative Breast Cancers Are Detectable in Histologically Normal Breast Epithelium.* Clinical Cancer Research, 2011. **17**(2): p. 236-246.

130. Garcia-Closas, M.B., L A Lissowska, J Chatterjee, N Peplonska, B Anderson, W F Szeszenia-Dabrowska, N Bardin-Mikolajczak, A Zatonski, W Blair, A Kalaylioglu, Z Rymkiewicz, G Mazepa-Sikora, D Kordek, R Lukaszek, S Sherman, M E, *Established breast cancer risk factors by clinically important tumour characteristics.* Br J Cancer, 2006. **95**(1): p. 123-129.

131. Elloumi, F., et al., *Systematic Bias in Genomic Classification Due to Contaminating Non-neoplastic Tissue in Breast Tumor Samples.* BMC Medical Genomics, 2011. **4**(1): p. 54.

132. Oh, D.S., et al., *Estrogen-Regulated Genes Predict Survival in Hormone Receptorâ€"Positive Breast Cancers.* Journal of Clinical Oncology, 2006. **24**(11): p. 1656-1664.

133. Campbell, M., et al., *Proliferating macrophages associated with high grade, hormone receptor negative breast cancer and poor clinical outcome.* Breast Cancer Research and Treatment, 2011. **128**(3): p. 703-711.

134. Largent, J.A., et al., *Reproductive History and Risk of Second Primary Breast Cancer: The WECARE Study.* Cancer Epidemiology Biomarkers & Prevention, 2007. **16**(5): p. 906-911.

135. Begg, C.B. and M. Berwick, *A note on the estimation of relative risks of rare genetic susceptibility markers.* Cancer Epidemiology Biomarkers & Prevention, 1997. **6**(2): p. 99-103.

136. Yang, X., et al., *Estrogen receptor and progesterone receptor expression in normal terminal duct lobular units surrounding invasive breast cancer.* Breast Cancer Research and Treatment, 2013: p. 1-11.

137. Haslam, S. and T. Woodward, *Host microenvironment in breast cancer development: Epithelial-cell-stromal-cell interactions and steroid hormone action in normal and cancerous mammary gland.* Breast Cancer Res, 2003. **5**(4): p. 208 - 215.

138. Sansone, P., et al., *IL-6 triggers malignant features in mammospheres from human ductal breast carcinoma and normal mammary gland.* The Journal of Clinical Investigation, 2007. **117**(12): p. 3988-4002.

139. Bissell, M.J., P.A. Kenny, and D.C. Radisky, *Microenvironmental Regulators of Tissue Structure and Function Also Regulate Tumor Induction and Progression: The Role of Extracellular Matrix and Its Degrading Enzymes.* Cold Spring Harbor Symposia on Quantitative Biology, 2005. **70**: p. 343-356.

140. Weaver, V.M., et al., *Reversion of the Malignant Phenotype of Human Breast Cells in Three-Dimensional Culture and In Vivo by Integrin Blocking Antibodies.* The Journal of Cell Biology, 1997. **137**(1): p. 231-245.

141. Carey La, P.C.M.L.C.A. and et al., *RAce, breast cancer subtypes, and survival in the carolina breast cancer study.* JAMA, 2006. **295**(21): p. 2492-2502.

142. Prat, A. and C.M. Perou, *Deconstructing the molecular portraits of breast cancer.* Molecular Oncology, 2011. **5**(1): p. 5-23.

143. O'Brien, K.M., et al., *Intrinsic Breast Tumor Subtypes, Race, and Long-Term Survival in the Carolina Breast Cancer Study.* Clinical Cancer Research, 2010. **16**(24): p. 6100-6110.

144. Millikan, R., et al., *Epidemiology of basal-like breast cancer.* Breast Cancer Research and Treatment, 2008. **109**(1): p. 123-139.

145. Wong, D.J., et al., *p16INK4a Lesions Are Common, Early Abnormalities that Undergo Clonal Expansion in Barrettâ€™s Metaplastic Epithelium.* Cancer Research, 2001. **61**(22): p. 8284-8289.

146. Heneweer, M., et al., *Co-culture of Primary Human Mammary Fibroblasts and MCF-7 Cells as an In Vitro Breast Cancer Model.* Toxicological Sciences, 2005. **83**(2): p. 257-263.

147. Krause, S., et al., *The microenvironment determines the breast cancer cells' phenotype: organization of MCF7 cells in 3D cultures.* BMC Cancer, 2010. **10**(1): p. 263.

148. Griffith, L.G. and M.A. Swartz, *Capturing complex 3D tissue physiology in vitro.* Nature publishing group, 2006. **7**(3): p. 211-224.

149. Yumei, Y., et al., *Molecular Mechanisms of Breast Cancer Progression: Lessons from Mouse Mammary Cancer Models and Gene Expression Profiling.* Breast Disease, 2004. **19**(1): p. 69-82.

150. Kim, B.G., et al., *Laminin-332-Rich Tumor Microenvironment for Tumor Invasion in the Interface Zone of Breast Cancer.* The American Journal of Pathology, 2011. **178**(1): p. 373-381.

151. Paszek, M.J. and V.M. Weaver, *The Tension Mounts: Mechanics Meets Morphogenesis and Malignancy.* Journal of Mammary Gland Biology and Neoplasia, 2004. **9**(4): p. 325-342.

152. Wang, X., et al., *A complex 3D human tissue culture system based on mammary stromal cells and silk scaffolds for modeling breast morphogenesis and function.* Biomaterials, 2010. **31**(14): p. 3920-3929.

153. Chen, D.-T., et al., *Proliferative genes dominate malignancy-risk gene signature in histologically-normal breast tissue.* Breast Cancer Research and Treatment, 2011. **119**(2): p. 335-346.