Breast cancer — one term, many entities?

Nicholas R. Bertos1,2 and Morag Park1,2,3,4

1Rosalind and Morris Goodman Cancer Research Centre, 2Department of Medicine, and 3Departments of Oncology and Biochemistry, McGill University, Montreal, Quebec, Canada. 4Department of Medicine, McGill University Health Centre, Montreal, Quebec, Canada.

Breast cancer, rather than constituting a monolithic entity, comprises heterogeneous tumors with different clinical characteristics, disease courses, and responses to specific treatments. Tumor-intrinsic features, including classical histological and immunopathological classifications as well as more recently described molecular subtypes, separate breast tumors into multiple groups. Tumor-extrinsic features, including microenvironmental configuration, also have prognostic significance and further expand the list of tumor-defining variables. A better understanding of the features underlying heterogeneity, as well as of the mechanisms and consequences of their interactions, is essential to improve targeting of existing therapies and to develop novel agents addressing specific combinations of features.

Normal architecture of the breast

To understand how such variation arises, it is instructive to examine the normal architecture of the breast environment (Figure 1A). Each lobe arises from multiple lobules, which connect to a common terminal interlobular duct. These ducts then continue to their outlet at the nipple. Histologically, lobules and ducts are lined by a single layer of luminal epithelial cells, surrounded by transversely oriented myoepithelial cells. These structures are separated from the surrounding tissue, or stroma, by a basement membrane, the breach of which distinguishes invasive carcinoma from carcinoma in situ (1). The surrounding stroma comprises ECM, discrete cells (e.g., fibroblasts, immune cells, and adipocytes), and organized structures (e.g., blood vessels), each of which contributes to the overall configuration of the local microenvironment.

Structural heterogeneity

Classical pathology has segregated breast tumors into multiple categories, based on their overall morphology and structural organization. The most common type observed and reported is invasive ductal carcinoma, not otherwise specified (IDC NOS; about 75% of cases), while invasive lobular carcinoma (ILC) represents the next most frequent histologic type of breast tumor (about 10% of cases) (2). Together, these two categories and combinations thereof make up the vast majority (about 90%) of breast cancers, while the remainder are categorized as medullary, apocrine, metaplastic, mucinous (A and B), inflammatory, comedo, adenoid cystic, and micropapillary types (2, 3). Interestingly, histologic type is linked to prognosis. While IDC NOS, ILC, apocrine, and medullary carcinomas have similar 10-year survival rates, adenoid cystic, medullary, mucinous, and tubular carcinomas exhibit relatively better overall outcomes (4–6). However, the rarity of specific non-ILC/non-IDC tumors (known as “special types”) has resulted in a relative paucity of in-depth characterization of larger cohorts of such cases, and thus the details of how membership in these groups interacts with other factors contributing to tumor heterogeneity are not well understood.

Immunopathological classifications

The presence of specific markers in breast cancer has long been recognized to both define subtypes with differential overall prognosis and to identify tumors susceptible to targeted treatments. The chief markers assessed are estrogen receptor (ER), progesterone receptor (PR), and human epidermal receptor 2 (HER2). Expression of the first two is assayed almost exclusively by immunohistochemistry (IHC)-based methods, which report levels of the corresponding proteins, while HER2 assays combine IHC and FISH approaches. Ambiguous IHC-derived HER2 results are subjected to FISH testing for genomic amplification of HER2; in cases in which the overall ratio of copies of the HER2 gene to those of its chromosome is greater than 2.2 are reported as HER2+ (7). ER status is utilized to identify tumors that may respond to anti-estrogen (endocrine) therapy, including ER antagonists or aromatase inhibitors, which target ER-dependent signaling (8, 9). PR status, generally correlated with ER status, has less clinical significance. PR status does not appear to predict relative benefit from specific types of endocrine therapy (10, 11), and overall, ER+/PR− cases may not receive additional benefit from endocrine therapy compared with ER+/PR+ cases (12–14). HER2+ cases are treated with targeted therapies such as the monoclonal antibody trastuzumab, which binds to HER2, mediates antibody-dependent cytotoxicity, and disrupts HER2-dependent signaling (15, 16). There is currently no standard targeted therapy for cases assessed as ER− and HER2− by IHC, although this represents an intensive area of research.

Combinations of these markers allow for the assignment of individual cases to specific categories, namely ER−/HER2−, HER2+ (ER−/HER2+), triple negative (TN; ER−/PR−/HER2−), and triple positive (ER+/PR+/HER2+). From a prognostic viewpoint, ER+ tumors exhibit the best overall outcome. Following the advent of HER2-targeted therapies, HER2+ tumors, previously associated with poor outcome, now exhibit an improved overall outcome when treated with such therapies (17, 18). TN tumors, on the other hand, are linked to the worst prognosis among the subtypes (19), while triple-positive cases appeared to have a prognosis intermediate between those of ER+ and HER2+ cases prior to the introduction of HER2-targeted treatments (20).

Intratumoral heterogeneity

It is important to note that heterogeneity within individual tumors is significantly more prevalent than the status assignments by ER-, PR-, and HER2-directed assays described above would indicate. For example, ER− status is currently reported when the proportion of ER− tumor cells within the tumor exceeds a 1% threshold (21), while the IHC threshold for HER2 positivity is met when more than 30%
of cells display strong membrane-associated HER2 expression. These low thresholds necessarily imply that in many cases, the majority of the tumor cells present display features inconsistent with the overall status assigned. In addition, the clinical course of disease may be governed by alterations occurring within relatively small subsets of primary tumor cells. This is reflected by differences between the primary tumor and metastatic lesions at the ER/HER2 status (22) and genomic (23) levels. In this context, it is important to note that reporting thresholds vary over time and between laboratories (e.g., the threshold for ER positivity has changed from 10% to 1% of tumor cells exhibiting ER positivity in IHC analysis; ref. 24). Thus, conclusions and correlations based on reported receptor status from different studies may not be directly comparable (25).

Significant intratumoral heterogeneity also exists at the genomic level. Investigation of discrete tumor regions by comparative genomic hybridization established that significant genomic heterogeneity was present in more than half (5 of 9) of all tumors examined (26). Studies of genomic alterations in multiple single cells from individual tumors have revealed that over half of the samples examined are polygenic, containing multiple clonal subpopulations that may be spatially separated or intermixed (27, 28). Future studies will be required to address the implication of these observations.

**Molecular heterogeneity at the gene expression level**

Over the past decade, the advent of high-throughput/high-content microarray-based gene expression profiling technologies has facilitated relatively large-scale studies of breast cancer cohorts, leading to the identification of multiple molecular subtypes of breast cancer. Studies by Perou and colleagues (29–32) described molecular subtypes defined by distinct transcriptional signatures that partially recapitulate the original immunopathological classes, while adding an additional level of detail. Two luminal subtypes (A and B) contain principally ER+ cases and are distinguished by the presence of genes regulated by the ER signaling pathway. The luminal A subtype is associated with higher levels of ESR1, ER, and ER-regulated genes (32), decreased proliferation (29, 33, 34), and improved overall outcome (29, 30, 32). Luminal B tumors appear to exhibit decreased levels of ESR1, ER, and ER-regulated genes as well as increased proliferation and relatively worse prognosis.

Other classification schemes have also been proposed, including one that divides ER+ disease into four subtypes that are chiefly distinguished by differential expression of proliferation-related genes (35). Recent investigations suggest that the distribution of luminal cases, rather than forming two distinct categories, may be better modeled as a continuum along which ER-regulated elements and proliferation are inversely related.
(36, 37). This model is further supported by the finding that gene expression signatures designed to classify novel samples by molecular subtype (29, 38, 39) exhibit substantial disagreement for luminal cases (40, 41), implying that the luminal subtypes are not as distinct as initially believed.

In the scheme proposed by Perou and colleagues, three subtypes contain predominantly ER- cases (31). The molecular ERBB2- subtype generally (but imperfectly) overlaps with IHC-defined HER2+ tumors (31), while the basal or basal-like subtype broadly corresponds to the TN (ER-/PR-/HER2-) cohort (42, 43). A normal-like molecular subtype resembles normal epithelial tissue (44) and may comprise cases in which samples contain large amounts of non-tumor tissue (29, 39). In an alternate scheme, two ER- subtypes are defined by different expression levels of the androgen receptor, while HER2+ cases do not form a separate group, but rather cluster according to their ER status (35). Other molecular subtypes within the ER- cohort have also been reported. These include a claudin-low subtype (45) linked to metastatic breast cancer and poor outcome, which shows similarity to stem cell–linked and epithelial-to-mesenchymal transition–linked (EMT-linked) signatures, and the molecular apocrine class, enriched in ER-/HER2+ tumors and displaying positivity for androgen receptor and downstream signaling (46).

TN breast cancers are necessarily defined by the absence of specific markers, suggesting that significant heterogeneity may exist within this group. An investigation of expression profiles from S87 TN breast tumors further subdivided these into 6 groups (47). These exhibited preferential responses to specific chemotherapeutic regimens, as well as differential expression of basal-specific, immunomodulatory, mesenchymal, mesenchymal stem-like, and androgen receptor-related genes, likely reflecting several of the classes discussed above.

Clustering by its nature leads to definitive assignments of tumors to specific subtypes. However, individual tumors likely display combinations of features of different subtypes, whether due to intratumoral heterogeneity or to a mixed phenotype of individual cells. Thus, the specific nature of a tumor may be better defined by its location within a multidimensional continuum, in which the canonical subtype-defining signatures represent vertices, than by a single label. It is important to note that gene expression profiles derived from bulk tumor by definition must reflect an average value of the cells surveyed. The fact that these profiles nevertheless significantly correlate with disease course suggests that they may reflect the propensity of tumors exhibiting distinct processes, e.g., proliferation, to give rise to subpopulations capable of metastasis.

**Genomic heterogeneity**

The principal molecular subtypes were each found to be associated with specific patterns of genomic alterations (48–53). Six genomic subtypes have been identified in breast cancer. Four of these overlap with gene expression–defined groups (ERBB2-, basal-like, luminal); however, the overall luminal cohort was imperfectly segregated between luminal A and luminal B cases, while two genomic classes (amplifier and mixed) contained tumors from multiple gene expression–defined subtypes (54). These results further suggest that gene expression–defined classifications may not fully define the spectrum of diversity present across breast tumors. Certain genomic changes, e.g., amplification of the HER2-containing amplicon, are clearly linked to the molecular features of the corresponding tumor subtype. However, the extent to which genomic alterations determine features of tumor subtype, or whether they also partially reflect variations in mechanisms of genomic instabil-ity between subtypes, remains an open question.

A new area for studying the behavior of breast cancers has arisen with the discovery of microRNAs (miRNAs), short RNA molecules (approximately 22 nucleotides) that play roles in transcriptional and posttranscriptional regulation of gene expression (55, 56). Analyses of miRNA profiles in breast cancer have determined that many miRNAs display expression patterns linked to molecular subtype (57–60) as well as ER status, tumor grade (58), and other tumor-related processes (61–63). Experimental evidence has confirmed that miRNA levels can play a role in determining disease course; for example, re-expression of miRNA-193b, down-regulated in highly metastatic derivatives of the MDA-MB-231 cell line, significantly inhibited tumor growth and dissemination in a mouse xenograft model (64). Thus, integrated analysis of miRNA and mRNA profiles in breast cancer constitutes an important new frontier (60, 65).

While some changes in miRNA expression are correlated with genomic changes or local alterations in primary transcription rates, changes in overall miRNA biogenesis may also underlie alterations of miRNA levels (58). Transcriptional levels of the miRNA-processing elements AGO2 and DICER1 correlate with ER status, proliferation status, tumor molecular subtype, and disease outcome (58, 66–68), while decreased levels of the miRNA processor Drosha are linked to HER2 positivity (67).

Many of the mRNA- and miRNA-level alterations observed in breast cancer cannot be ascribed to genomic changes. These are likely linked to epigenetic factors, including DNA methylation and histone modifications. DNA methylation within the ESR1-containing CpG island was increased in ER- breast cancer samples compared with ER+ samples (69–71), suggesting that methylation-induced ESR1 silencing may govern ER expression in some cases. Recent studies including samples from different molecular subtypes demonstrate that characteristic differential DNA methylation profiles subdivide samples into multiple luminal A–enriched and basal-like/ERBB2-–enriched clusters (72, 73). Interestingly, methylation cluster–specific survival differences were observed within luminal A samples (73), suggesting that additional information beyond molecular subtype alone is captured by methylation profiling.

**Differences in cell of origin**

The expression patterns of specific cytokeratins (CKs) in the luminal and basal molecular breast cancer subtypes resemble those of normal luminal epithelial (CK8/18-expressing) and basal stem (CK5/6-expressing) cells of the breast, respectively. Therefore, it has been proposed that these tumors arise from different cells of origin, and that luminal tumors arise from luminal progenitor cells, while basal tumors originate from the basal or stem cell compartment (38, 74, 75). However, comparisons of the transcriptional signatures of the molecular breast cancer subtypes with gene expression profiles of cell subsets isolated from mammary tissue suggest that different molecular subtypes of breast cancer may arise from cells at various stages of differentiation along the mammary epithelial hierarchy (76).

Claudin-low tumors display gene expression profiles similar to those of mammary stem cells, while basal tumors resemble bipotent or early luminal progenitors (77–79). ERBB2+ tumors are...
most similar to late luminal progenitors, while luminal tumors are closest to the differentiated luminal cell compartment (76). This scheme is somewhat complicated by the possibility that tumor cells may drift from their original configuration, losing specific markers of the cell of origin and taking on others (epithelial plasticity). A prime example of this is EMT, in which transformed epithelial cells switch to a mesenchymal phenotype whose expression profile resembles that of stem cells (80). EMT is associated with increased motility and enhanced invasion. However, once they reach a distant site, these cells can then revert to an epithelial phenotype following interactions with the local microenvironment (81–84).

Tumors of special types

The principal studies reporting the diversity of molecular breast cancer subtypes were conducted using samples from the most common histological subtypes, i.e., IDC and occasionally ILC tumors. Tumors of special types, i.e., non-IDC NOS cases, constitute 25% of breast tumors, generally display high uniformity with respect to ER and HER2 status within each type (85), and are commonly associated with good clinical outcome (86). Microarray-based gene expression profiling of small cohorts of breast tumors of special types category established that special types tend to cluster within single subtypes (87), while IDC NOS tumors display a variety of molecular subtypes. For example, tubular carcinomas and standard ILC samples display similar gene expression patterns and fall into the luminal category, while pleomorphic ILC samples group with apocrine tumors (119) and enhances angiogenesis through recruitment of endothelial progenitor cells (120), as well as matrix metalloproteinases, which can mediate degradation of the ECM (121). Interestingly, manipulation of stromal fibroblasts via specific ablation of the tumor suppressor gene PTEN, which negatively regulates the Akt kinase and cell survival and proliferation pathways, supports enhanced tumor development of a Neu/ErbB2-driven tumor xenograft (122). In contrast, fibroblast-specific ablation of the ETS2 transcription factor reduces Neu/ErbB2 tumor growth, demonstrating the functional importance of fibroblast-dependent interactions in disease outcome (122).

Angiogenesis is a necessary step for tumor growth, and microvessel density has weak prognostic value in breast cancer (123, 124), demonstrating that this element is also heterogeneously distributed. Thus, a range of potential microenvironmental states and responses further enhances the heterogeneity of breast cancer. The full spectrum of these combinations, and the nature and extent of their individual interactions with tumor-intrinsic determinants of disease course, remains to be investigated.

Microenvironmental heterogeneity

Recent studies have demonstrated that disease course is not solely linked to tumor-intrinsic features, but that features of the local microenvironment, or stroma, can strongly influence outcome (94–100). Patient-specific variations in the abundance and status of different stromal cell types, as well as communication between and mutual modulation of each compartment (tumor and stroma), induce additional levels of complexity and contribute to breast cancer heterogeneity and disease course (Figure 1B). Stromal elements contributing to this heterogeneity include the ECM itself, alterations of which have been proposed to contribute to tumorigenesis (101) and response to treatment (102). A mouse model with increased stromal type I collagen exhibited increased tumor formation and invasion (103), likely via a \( \beta_1 \)-integrin–dependent mechanism (104), while normal, but not tumor-associated, myoepithelial cells reversed invasion in DCIS-like cells via synthesis of laminin-1 (105). While myoepithelial cells are commonly absent in breast tumors, the interactions described above likely influence early stages of disease initiation and progression to malignancy.

Immune cells, fibroblasts, and endothelial cells also vary in number and type among different tumors. Increased numbers of tumor-associated macrophages are associated with poor disease outcome (106), reflecting their ability to enhance tumor cell invasion (107–108). This is mediated in part via a mutual paracrine loop involving growth factors produced by macrophages that influence tumor cell migration, such as epidermal growth factor, and growth factors produced by tumor cells that attract macrophages, such as colony stimulating factor 1 (109). Tumor-associated T cells can have different effects, depending on activation status. Th1-type T cells are associated with good outcome (99, 110–113), while immunosuppressive Tregs are associated with tumor progression and poor prognosis (113–116), and response to neoadjuvant therapy correlates with changes in Treg numbers (117). In addition, microenvironmental factors, including T cell status, may play a more prominent role in determining outcome in ER\(^-\) and HER2\(^-\) tumors (99, 118), while tumor proliferation is more closely correlated with outcome in ER\(^+\) disease.

Carcinoma-associated fibroblasts (CAFs) are known to secrete tumor-promoting factors, including the chemokine SDF-1/CXCL12, which promotes tumor cell migration into the stroma (119) and enhances angiogenesis through recruitment of endothelial progenitor cells (120), as well as matrix metalloproteinases, which can mediate degradation of the ECM (121). Interestingly, manipulation of stromal fibroblasts via specific ablation of the tumor suppressor gene PTEN, which negatively regulates the Akt kinase and cell survival and proliferation pathways, supports enhanced tumor development of a Neu/ErbB2-driven tumor xenograft (122). In contrast, fibroblast-specific ablation of the ETS2 transcription factor reduces Neu/ErbB2 tumor growth, demonstrating the functional importance of fibroblast-dependent interactions in disease outcome (122).

Macrogenenvironmental heterogeneity

Beyond those features of the stroma that are specific to the local tumor microenvironment, systemic factors may also contribute to the range of behaviors evinced by breast tumors. These include factors such as age, menopausal status, and variations in body mass index (the latter two affecting systemic estrogen and progesterone levels; ref. 125), as well as overall immune status. For example, obesity is associated with increased tumor recurrence and decreased survival (126–128) as well as with a distinct tumor transcriptional signature (129), potentially due to the influence of surrounding adipocytes on tumor tissue (126). How these variables interact with tumor-intrinsic and local microenvironmental features remains an area for further study.
molecules supporting tumor cell homing. Recent studies demonstrate that the expression of specific molecules, including chemokine receptors as well as claudin-2, CCN3, and tenascin-C, correlate with and/or enhance metastasis to distinct sites (119, 136–138). The chemokine CXCL12 is expressed in bone marrow, and expression of its cognate receptor CXCR4 on breast tumor cells is correlated with an increased risk of metastasis to bone (139). Interestingly, cell lines selected for enhanced metastasis to bone display a specific transcriptional signature (including elevated CXCR4) that is also present in subsets of cells from the parental tumor cell population (140), suggesting that subpopulations of cells capable of metastasis to specific sites form part of the initial spectrum of heterogeneity in the primary tumor.

The “seed-and-soil” model may also apply to the local environment of the primary tumor. Tumor self-seeding, in which cells escaping from the primary tumor are hypothesized to later return to the same anatomical site and to be concentrated at the outer surface of the lesion (141–143), also leads to increased intratumoral heterogeneity and may contribute to the polygenic nature of many tumors. This model supports a view of cancer progression as a process dependent on multiple factors beyond the local disease site. Support for this concept comes from a study demonstrating that indolent tumors are induced to grow by the presence of an actively proliferating tumor at an anatomically distinct site, and that this occurs via osteopontin-mediated mobilization and recruitment of bone marrow–derived cells into the stroma of the previously indolent tumor (144). Similarly, circulating estrogen can induce bone marrow–derived cell recruitment to the stroma of ER+ tumors in a mouse model, leading to the promotion of tumor growth (145) — the recruitment of such cells further enhances heterogeneity in the tumor microenvironment.

**Circulating tumor cells**

In breast cancer, the majority of patients present with local disease, and the primary lesions are generally removed by surgery prior to the development of clinically detectable metastases. Therefore, prognostic information obtained through studies of the primary tumor must essentially reflect the relationship between features of this tissue and the likelihood that cells with the potential to proliferate at distant sites have already disseminated before primary tumor resection. In order to gain a better understanding of
the elements governing recurrence, an intermediate between the primary tumor and distant metastases must be sought. Circulating tumor cells (CTCs) represent cells that have already escaped the primary tumor site and thus may be an appropriate candidate. These cells have a high level of agreement (82%–89%) at the level of HER2 status with the primary tumor (146, 147); however, concordances for ER and PR status are lower (41% and 45%, respectively) (148). These observations suggest that CTCs may act as a proxy for the subset of cells within the primary tumor capable of leading to disease recurrence. Interestingly, characterization of CTCs shows that subpopulations of these cells are enriched for stem cell and EMT markers (149, 150), suggesting that they arise from specific subgroups within the tumor. Further investigations of the correlation between markers of CTCs and those of distant metastases are expected to clarify the origin of CTCs and how their features influence disease course.

Conclusions and future directions

The identification of multiple sources of tumor heterogeneity at both tumor-intrinsic and -extrinsic levels, as well as the changes seen during progression (Table 1), suggests that an enhanced understanding of the effect of this heterogeneity on the impact of specific treatments represents an urgent clinical need. Therefore, it is desirable that provision is made within large-scale clinical trials for the collection of sufficient samples to assign tumors to specific combinations of classes. While markers for specific tumor-intrinsic and -extrinsic levels, as well as the changes in intrinsic protein and transcript features either have already been included in standard clinical practice (e.g., IHC-based assessment of receptor status) or are beginning to approach standardization (i.e., the 50-gene PAM50 classifier for assignment to molecular subtypes), assays for microenvironmental features remain outside the realm of clinical use. Additional studies of large cohorts (including the histologic “special types” of tumors) at multiple concurrent levels, including integrated tumor-intrinsic, microenvironmental and macroenvironmental profiling, as well as analyses of matched CTCs and distant metastases and assessment of intratumoral heterogeneity, will be necessary. This will ultimately allow the development of a complete set of classifiers that, taken together, can define an individual tumor by its location in a multidimensional coordinate system comprising all of the variables contributing to breast cancer heterogeneity. Such schemes can then be applied to tissue and blood samples from clinical trials of specific therapies, generating an improved stratification system to predict the relative benefit of each potential intervention, or combinations thereof, in a model tailored to the specific combination of features defining an individual cancer.

Acknowledgments

Space constraints render it impossible to adequately acknowledge all of the contributions made to this field by the many key individuals and groups who have studied different aspects of this research area in detail, and in many cases have made it necessary to refer the reader to reviews. The authors apologize for any omissions. The authors acknowledge the valuable insights gained into breast tumor heterogeneity through discussions with colleagues from different fields at McGill University and the McGill University Health Centre, including Michael Hallett (Bioinformatics), Catalin Mihalceiu (Medical Oncology), Sarkis Michaelian (Surgery), and Attila Omeroglu (Pathology). M. Park holds the Diane and Sal Guerra Chair in Cancer Genetics at McGill University.

Address correspondence to: Nicholas R. Bertos, 1160 Ave. des Pins O., Montreal, Quebec, Canada H3A 1A3. Phone: 514.398.2325; Fax: 514.398.6769; E-mail: nicholas.bertos@mcgill.ca.

24. Welsh AW, et al. Standardization of estrogen receptor measurement in breast cancer suggests false-negative results are a function of threshold intensity rather than percentage of positive cells.


