Overview of molecular classification of breast cancer

Breast cancer is a heterogeneous disease clinically but can be stratified into meaningful groups based on 3 simple molecular tests: immunohistochemical detection of estrogen (ER) and progesterone receptor (PgR) as well as amplification/overexpression of HER2 as determined by either fluorescence in situ hybridization (FISH) or immunohistochemistry, respectively. These markers are used to predict response to specific therapies, and are weak prognostic markers. In combination with axillary node status, tumor size, histologic grade and histological subtype, they are still the most important factors guiding treatment.

Approximately two-thirds of invasive breast cancers express ER, representing a diverse group of cancers. Several genomic tests have been developed in an attempt to refine outcome prediction for ER positive tumors, (for example, the 21-gene recurrence score and 70-gene signature); however, the predictive power of these assays is far from ideal, and the assays have not been fully tested in patient populations treated with aggressive adjuvant or neoadjuvant chemotherapy. The Her2-amplified (ERBB2) group represents 10-15% of breast cancers and was associated with poor outcome prior to the introduction of anti-HER2 agents. Approximately 15-20% of breast cancers lack all three of these targets (triple negative breast cancer, TNBC) and chemotherapy is the mainstay of treatment.

Genomic Expression Profiling

Advances in gene expression microarray technology have allowed simultaneous analysis of thousands of genes in breast cancer enabling identification of ‘signatures’ which subclassify breast cancers as well as predict prognosis and response to therapy. More than 10 years ago, Perou and Sorle published the first comprehensive analyses of gene expression profiles of breast cancer. The studies caused a radical paradigm shift in breast cancer research from the “one gene-at-a-time” approach to the study of thousands of genes and their products simultaneously. These analyses showed that: 1) individual tumors vary significantly in their gene expression patterns, 2) there is multidimensionality to the variation (patterns of variation...
among numerous genes are independent) and 3) these patterns have an overriding order reflecting relationships among the genes, among the tumors and between specific genes and specific tumor types\(^8\).

Hierarchical clustering of their microarray-derived expression data identified intrinsically variable gene sets that distinguished five breast cancer subtypes- basal-like (basal epithelial/myoepithelial), luminal A (strongly ER-expressing with lower proliferative rate, luminal B (lower ER-expressing and greater expression of proliferation genes), HER2-enriched and “normal breast-like” which showed significant differences in overall and relapse-free survival\(^7-10\). All but the latter are now commonly referred to as the “intrinsic” subtypes. The majority of tumors classified as “basal-like” correspond to the TNBC group identified by standard immunohistochemical markers (IHC), while the HER2-enriched tumors correspond largely to the Her2-overexpressed/amplified group\(^7\). The luminal A and luminal B subtypes generally serve to divide ER-positive tumors based on their levels of ER expression and expression of proliferation genes\(^11,12\). However, the luminal B subtype also includes some tumors which are both ER- and HER2-positive or express PgR only\(^11,12\). In addition, many investigators have shown that the “intrinsic” subtypes show varying degrees of reproducibility across platforms and analysis methods. To address these issues, the PAM50, a test based on a minimal gene set (50 genes), was developed which classifies tumors into luminal A, luminal B, HER2-enriched, and basal-like subtypes based on expression of these assessed by qRT-PCR\(^11,13\). PAM50 can be used on paraffin-embedded tissue samples. It should be noted that the results of PAM50 and the standard IHC for ER, PR and HER2 are not entirely equivalent.

Additional molecular subtypes have been proposed for tumors that are ER-negative by immunohistochemistry. These are the “molecular apocrine” tumors which are characterized by a combination of androgen receptor (AR) expression and HER2 overexpression\(^14,15\), luminal androgen receptor (LAR) characterized by AR expression in the absence of HER2 overexpression\(^16\) and claudin-low tumors\(^17\). While the potential use of androgen receptor blockade is being tested in clinical trials, the true biologic and clinical relevance of these subtypes remains to be determined.

**Comparative genomic hybridization studies**

Comparative genomic hybridization studies (CGH) studies have shown recurrent alterations in breast cancer. These include gain of chromosomal regions 1q, 8q, 16p, 17q, and 20q; the loss of 16q and 17p; and DNA amplification in 8q12-24, 11q11, 17q12 (HER2), 17q22-24 and 20q13-
These studies, most performed more than a decade ago, showed low grade carcinomas to be diploid or near-diploid and harbor recurrent deletions of 16q, gains of 1q and gains of 16p. In contrast, high grade NST carcinomas showed complex genotypes frequently harboring loss of 11q, 14q, 8p, 13q; gain of 17q, 8q, 5p; and amplifications on 17q12, 17q22-24, 6q22, 8q22, 11q13 and 20q13.

As microarray-based CGH (aCGH) technologies have advanced over the past decade, they have demonstrated the remarkably heterogeneity of breast cancer at the genomic level, mirroring the complexity seen in gene expression profiling. When examined without prior knowledge of breast cancer subtype, investigators have shown that breast cancer can be subdivided based on the architecture of the genomic alterations reflecting different types of genomic instability. Although the “intrinsic” subtypes, as defined by gene-expression array analysis have distinct patterns of gene copy number aberrations, significant intrasubtype heterogeneity remains evident.

Despite the overall complexity, a few genomic-phenotypic correlations have been made. The most well-known is in carcinomas with a lobular phenotype, where the target gene of 16q deletions is \( CDH1 \) which encodes E-cadherin. However, in the remaining low grade ER-positive carcinomas with 16q deletions, both special type and NST, the target gene(s) remains unknown.

**Next Generation Sequencing**

Next generation sequencing (NGS) is unquestionably one of the most significant technical advances in the biological sciences in the past 30 years. NGS describes a group of technologies capable of sequencing millions of DNA templates in parallel. This technology has enabled researchers to characterize the full spectrum of mutations in breast cancer genomes including segmental duplications, amplifications and deletions, translocations, inversions, small insertion-deletions and point mutations. Three major patterns are seen: 1) tumors with a few interchromosomal translocations with copy number alterations involving large DNA fragments or whole chromosome arms, 2) tumors with complex interchromosomal translocations affecting shorter regions with high-level amplifications and 3) tumors with small intrachromosomal segmental alterations such as duplications, deletions and/or inversions. The latter has been termed the “mutator phenotype” and is most commonly seen in TNBC. These patterns are also highly consistent with those deduced from aCGH. Relationships between the structural
rearrangement patterns and the gene expression subtypes have been described\textsuperscript{23}; however, as with aCGH, significant intrasubtype heterogeneity remains evident.

The biggest surprise of comprehensive sequencing studies of breast cancer is the apparent lack of a substantial number of recurrent mutations in breast cancer genes when mutational profiles are compared across patients\textsuperscript{30-33}. Only a few mutations have been shown to be highly recurrent across all breast cancers and the recent publication by TCGA reported that somatic mutations in only 3 genes (TP53, PIK3CA and GATA3) occurred at >10% incidence (\(n = 466\) tumors)\textsuperscript{33}. This suggests that individual tumors are unique, each harboring large numbers of infrequently mutated genes that uniquely characterize its genome\textsuperscript{34,35}. Even when mutations occurred in the same cancer genes, they were in different codons or protein domains\textsuperscript{33}. TCGA also analyzed the spectrum of mutations with respect to the four "intrinsic" subtypes\textsuperscript{33}.

Significantly mutated genes were considerably more diverse and recurrent within luminal A and luminal B tumors than within the basal-like and HER2-enriched subtypes; however, the overall mutation rate was lowest in the luminal A tumors and highest in the basal-like and HER2-enriched tumors. The luminal A subtype harbored the most significantly mutated genes, with the most frequent being \textit{PIK3CA} (45\%) followed by \textit{MAP3K1}, \textit{GATA3}, \textit{TP53}, \textit{CDH1}, and \textit{MAP2K4}. Luminal B cancers exhibited a diversity of significantly mutated genes, with \textit{TP53} (29\%) and \textit{PIK3CA} (29\%) being the most frequent. In contrast, basal-like tumors displayed \textit{TP53} mutations in 80\% of cases and the majority of mutations present in the luminal tumors, except \textit{PIK3CA} (9\%), were absent/near absent. The HER2-enriched subtype, characterized by frequent Her2 amplification (80\%), had a hybrid pattern with a high frequency of \textit{TP53} (72\%) and \textit{PIK3CA} (39\%). Mutation types also differed among the "intrinsic" subtypes. \textit{TP53} mutations in basal-like tumors were mostly non-sense and frame shift, whereas missense mutations predominated in luminal A and B tumors. The recurrent \textit{PIK3CA} E545K mutation was present almost exclusively with in luminal A tumors, whereas the H1047L mutation was observed in all subtypes and was the most prevalent type in luminal A, luminal B and HER2-enriched tumors.

The promise that recurrent somatic mutations represent potential drug targets for breast cancer therapy has fueled this enterprise. Recent data indicate that activation of PI3K signaling is associated with the poor outcome luminal B subtype of breast cancer and accompanied by the development of endocrine therapy resistance\textsuperscript{36}. Importantly, inhibition of PI3K pathway signaling in endocrine resistant breast cancer cell lines reduces cell survival and improves treatment
response to endocrine agents\textsuperscript{36}. In fact, inhibitors of PI3K and well as other components of the PI3K pathway (AKT and mTOR) are being evaluated in clinic trials in combination with endocrine therapy and/or chemotherapy in ER-positive and HER2-positive breast cancers\textsuperscript{36,37}. Currently, there are no distinctive histologic features associated with the presence of PIK3CA mutations, however it has been suggested that among ER-positive cancers, lobular features are highly predictive\textsuperscript{38}, and that among TNBC the presence of AR-positivity and mesenchymal/metaplastic features are predictive\textsuperscript{16}.

A few other reported genomic-phenotypic correlations include t(12;15) in secretory carcinomas leading to the chimeric fusion gene ETV6-NTRK3\textsuperscript{39} and the t(6;9) in adenoid cystic carcinomas leading to the chimeric fusion gene MYB-NFIB\textsuperscript{40}.

What are the practical implications of this work and how can it guide therapy for individual patients?

The pace of acquisition of genomic data in both the research and clinical setting in cancer patients far outstrips the utility of that information to oncologists in choosing specific therapies for their patients. How should it be integrated practically with current clinical and pathological parameters? Are there instances where histology alone or in combination with immunohistochemistry is in fact currently still more informative than complex genomic analysis?

Special types of carcinomas with important clinical implications

There are special type carcinomas that are associated with such good prognosis that overall survival is equivalent to patients without breast cancer. These carcinomas are defined by a specific constellation of morphologically-defined features, and fall into the ‘luminal A” subgroup of carcinomas. Invariably estrogen receptor positive with a low proliferative rate, these carcinomas would be identified as not requiring chemotherapy, however, stronger statements regarding long term survival can be made based on morphologic grounds alone. These tumors include tubular carcinomas, invasive cribriform carcinoma, and pure mucinous carcinoma.

Pure Tubular Carcinoma
Tubular carcinomas are usually incidental findings on screening mammography and not associated with a palpable mass on physical exam. The frequency of cases detected by mammographic screening is 9-19\%\textsuperscript{41,42}.

Pure tubular carcinoma is diagnosed when more than 90% of the lesion is characterized by a haphazard infiltration of distinctly tubular structures lined by a single layer of bland epithelial cells\textsuperscript{43,44}. Any component that is not tubular should have grade 1 nuclei and a low proliferative rate, similar to the tubular component, and should compose 10% or less of the tumor. The tubules may be round, oval, or “bent teardrop” shaped. The epithelial cells lining the tubules are small and round with minimal pleomorphism. Mitotic figures are absent. Tumors containing between 70-89\% tubules are diagnosed as tubular variant carcinomas and tumors containing less than 70\% of the classic tubules we designate as ‘no special type tumors” with tubular features\textsuperscript{43}. Several studies have indicated that tubular histology is of prognostic value, and that the greater the purity of pattern the more favorable the prognosis. Several studies have shown that when >90\% purity of pattern is observed an excellent prognosis can be expected, even in the presence of a positive lymph node\textsuperscript{45-47}. Patients with pure tubular carcinoma can expect survival rates similar to the general population\textsuperscript{46}.

Tubular carcinoma always expresses estrogen and progesterone receptors, has a low proliferation rate, and is Her2 and EGFR negative, and is classified as a luminal A subtype of breast cancer\textsuperscript{43,44}. For pure tubular carcinomas, the excellent prognosis indicates that conservative but complete surgical excision is adequate therapy for the overwhelming majority of cases\textsuperscript{48}.

**Invasive Cribriform Carcinoma**

Invasive cribriform carcinoma (ICC) is uncommon, representing approximately 0.08\% to 3.5\% of invasive breast carcinomas and is biologically analogous to tubular carcinoma\textsuperscript{49,50}.

Pure ICC is composed of infiltrating islands of tumor cells with a cribriform architecture which resemble cribriform patterns of ductal carcinoma in situ. The tumor cells have low to intermediate grade nuclei and scant to moderate pale cytoplasm. Individual tumor islands contain bars and arches of cells creating well-defined spaces. To make a diagnosis of pure invasive cribriform carcinoma, more than 90\% of the tumor must have cribriform architecture, unless the non-cribriform areas represent less than 50\% of the tumor and have tubular
morphology\textsuperscript{43,44}. Mixed tubular and cribriform carcinomas in any proportion have the same excellent prognosis as pure tubular carcinomas.

Similar to tubular carcinoma, the prognosis is excellent despite the presence of lymph node metastasis (usually single node, rarely up to 3 nodes involved) in approximately 14\% of patients\textsuperscript{51}. In Page’s original series none of the patients died from ICC during 14.5 years average follow-up, despite an average size of 3 cm\textsuperscript{51}. A 100\% five year survival rate was also confirmed in a subsequent study\textsuperscript{52}.

Similar to tubular carcinoma, invasive cribriform carcinomas is classified as a luminal A type breast cancer, expressing estrogen receptor with a low proliferation index; HER-2 is not over-expressed\textsuperscript{53}.

**Pure Mucinous Carcinoma**

Pure mucinous carcinomas represent about 2\% if invasive mammary carcinomas and is usually detected as a well-circumscribed lesion on mammography. The median size of mucinous carcinoma is 2.0 cm is grossly circumscribed with a gelatinous appearance on cross section.

The hallmark histologic appearance of mucinous carcinoma is small nests, trabeculae, or sheets of epithelial cells, with smooth borders and usually with some glandular lumen formation, entirely surrounded by pools of extra-cellular mucin\textsuperscript{49}. The designation of pure mucinous carcinoma should be used only when mucinous morphology is present in over 90\% of the tumor which is of low Nottingham grade\textsuperscript{49,54,55}. Tumors with 50\% to 90\% mucinous morphology should be designated as invasive carcinomas of no special type with mucinous features.

Patients with pure mucinous carcinomas have 10 year survival rates similar to age-matched controls without carcinoma\textsuperscript{46}.

Estrogen receptor positivity is seen in approximately 95\% of pure mucinous carcinoma and progesterone receptor positivity in about 80\% of cases. Pure mucinous carcinomas are negative for Her-2.

**Triple-Negative Breast Cancers with Excellent Prognosis**
Fibromatosis-like metaplastic carcinoma

Metaplastic carcinomas of the breast are a heterogeneous group of cancers which collectively represent less than 1% of all breast cancers. Fibromatosis-like metaplastic carcinoma is a special variant of metaplastic carcinoma with relatively indolent growth which is capable of local recurrence but not distant spread. Fibromatosis-like metaplastic carcinomas have an infiltrative growth pattern with cellular finger-like projections extending into adjacent mammary structures and fatty tissue reminiscent of fibromatosis. These tumors are composed of bland appearing, plump spindle cells arranged in interlacing short fascicles, often with a storiform pattern and usually without obvious epithelial differentiation on H&E. In some cases the spindle cells merge with clusters of spindle cells which have fusiform or polygonal shapes with epithelioid nuclei which may cluster together. Rarely, foci of glandular or squamous elements may be seen, usually comprising less than 5% of the tumor. Cellularity is low, and consists of cells with low grade nuclei and indistinct nucleoli. The cytoplasm is pale, eosinophilic, and the neoplastic nuclei are tapered, and have finely dispersed chromatin.

Mitotic activity is low (range none to 3/10 HPF). The stroma is mildly myxoid to edematous appearing with a minimal to mild lymphoplasmacytic infiltrate which may focally resemble nodular fasciitis. Some cases also show varying degrees of fine stromal collagenization. Rarely spindle cell metaplastic carcinomas may be seen arising from a number of fibrosclerotic breast lesions including papillomas, complex sclerosing lesions, and nipple duct adenomas. Performance of immunohistochemistry using antibodies to p63 and high molecular weight cytokeratins is necessary to establish the epithelial nature of the process, and may also help delineate the extent of the tumor, especially helpful in assessing margins. Low grade fibromatosis-like metaplastic carcinomas are negative for ER, PgR and HER2 expression.

Adenoid Cystic Carcinoma

Adenoid cystic carcinoma (ACC) of the breast is a rare, but distinctive type of carcinoma, accounting for about 0.1% of all invasive breast malignancies. Strict criteria were identified by the 1970’s, aided by the use of electron microscopic features and histochemistry. ACC of the breast is so named because of its microscopic similarity to tumors of the major and minor
salivary glands. Among the several carefully defined special types of breast cancer, ACC is a prime example of the importance of recognizing a cluster of unique pathologic features that guarantee a usually excellent prognosis\textsuperscript{49,58,59}. This is exemplified by the greatly increased representation of women with ACC among twenty-year survivors of invasive breast cancer\textsuperscript{60}.

The histologic hallmark of adenoid cystic carcinoma is the presence of a dual cell population consisting of islands of admixed epithelial and myoepithelial cells forming small, sharply defined glandular and pseudoglandular structures, with true glands filled with brightly eosinophilic mucin and pseudolumens filled with more basophilic basement membrane-like material\textsuperscript{43,49}. Another common pattern is the formation of inter-anastomosing serpentine cords of tumor; however, the dual cell population and presence of basement membrane-like material usually remains obvious. In some cases, the basement membrane-like material may predominate forming irregular islands. Nodular and diffusely infiltrating smaller, usually rounded collections of infiltrating tumor may coexist, but maintenance of the same cytology and intercellular arrangements is the defining feature.

A solid variant of ACC has been described which exhibits a >90% solid growth pattern composed of basaloïd cells with moderate to occasionally marked nuclear atypia and rare to brisk mitotic activity\textsuperscript{61}. These tumors show the same immunohistochemical profile a typical ACC and are associated with an excellent prognosis, even in the presence of a lymph node metastasis.

PAS stains the contents of the ‘true’ lumens pink as is typical of epithelial mucins which are of near neutral pH whereas the basement membrane material or ‘stromal mucin’ of the ‘pseudolumens’ is more acidic and stains pale pink or blue. The main proliferating element in ACC is a population of modified myoepithelial cells as demonstrated by positive staining for smooth muscle actins, vimentin, and p63\textsuperscript{62}. The luminal epithelial cells are positive for c-kit (CD117)\textsuperscript{62}. These cells are grouped in nests outlining the ‘pseudolumens’ or more irregular islands of basement membrane-like material. The true epithelial lumens also referred to as the ‘glandular’ or ‘adenoid’ component are lined by cytokeratin and epithelial membrane antigen positive cells which demonstrate conserved basolateral markers of normal epithelial polarity as exemplified by positivity for fodrin, E-cadherin, and beta-catenin expression\textsuperscript{63}. Strictly defined adenoid cystic carcinomas are ER, PgR, HER2 negative\textsuperscript{64-66}; however, this finding should not be used as an indicator of poor outcome.
Secretory Carcinoma

Secretory carcinoma, originally described in adolescents, may also occur in adults. Secretory carcinoma was first termed juvenile carcinoma because the original cases were identified in children; however, it is now recognized that less than half of all cases present in patients less than 20 years old. In both males and females less than 20 years old, these tumors have indolent behavior and an excellent prognosis, even with axillary nodal involvement.

Histologically, lobules of tumor are separated by prominent bands of often densely sclerotic collagen. A microcystic appearance is imparted histologically as a result of numerous intercellular lumina containing extracellular secretions. Individual tumor cells show only mild to moderate cytologic atypia. The tumor cell cytoplasm may be clear as a result of cytoplasmic vacuolization or finely granular. The secretory material both within tumor cells and within the intercellular lumina stains positively for periodic acid-Schiff (diastase resistant) and Alcian blue. Secretory carcinoma is ER, PgR, HER2 negative.

Recently, a t(12;15)(p12;q26.1) translocation was identified in 12 of 13 cases of secretory carcinoma tested. Three of these patents were younger than 20 years old. This translocation has been previously identified in congenital fibrosarcoma, mesoblastic nephroma and a case report of acute myelod leukemia. The resulting ETV6-NTRK3 fusion gene encodes a chimeric tyrosine kinase with potent transforming activity in fibroblasts. Since secretory carcinoma is a histologically distinct lesion it is unclear whether there will be a role for translocation studies in diagnosis of this lesion; however, it is remarkable in that the occurrence of such translocations as the sole cytogenetic abnormality in epithelial tumors is extremely rare.

CONCLUSIONS

There is no question that enormous insight has been gained from multilevel studies of breast cancer genomics; however, our ability to translate these findings to improved treatment strategies for individual patient’s remains in its infancy. The genomic architecture of breast cancer is highly complex consistent with clinical observation of different responses to treatment in patients with seemingly similar clinicopathologic features. Our windows into this complexity include studies of gene expression profiling, analysis of DNA copy number changes,
determination of gene mutation status, determination of gene amplification status and detection for fusion genes in breast cancer. These approaches have provided an enormous volume of both complementary and overlapping information. Despite these exciting insights, the discussed special type carcinomas are clear examples in which histology alone or in combination with immunohistochemistry is in fact currently still more informative than complex genomic analysis in guiding therapy for individual patients.

REFERENCES


