Molecular Classification of Breast Cancer

Julia Y.S. Tsang, PhD and Gary M. Tse, FRCPC

Abstract: Cancer classification aims to provide an accurate diagnosis of the disease and prediction of tumor behavior to facilitate oncologic decision making. Traditional breast cancer classification, mainly based on clinicopathologic features and assessment of routine biomarkers, may not capture the varied clinical courses of individual breast cancers. The underlying biology in cancer development and progression is complicated. Recent findings from high-throughput technologies added important information with regard to the underlying genetic alterations and the biological events in breast cancer. The information provides insights into new treatment strategies and patient stratifications that impact on the management of breast cancer patients. This review provides an overview of recent data on high throughput analysis of breast cancers, and it analyzes the relationship of these findings with traditional breast cancer classification and their clinical potentials.

Key Words: breast cancer, molecular classification, review

Breast cancers are heterogenous, showing variable morphologic and biological features and thus different clinical behavior and response to treatment. Cancer classification aims to provide an accurate diagnosis of the disease and prediction of tumor behavior to facilitate oncologic decision making.

The mainstay of breast cancer assessment is to evaluate the following (and they are): (1) How bad the tumor is (typing and grading)? (2) How extensive is the tumor (staging)? The typing and grading of breast cancers is based on the histologic subtypes and grade, which are detailed in the WHO tumor classification (WHO). The staging of breast cancers is based on tumor size, nodal status, and distant metastasis (TNM staging). The routine assessment of breast cancer also includes estrogen receptor (ER α), progesterone receptor (PR), and human epidermal growth factor receptor 2 (HER2) expression.

For the histologic type, most of the breast cancers (> 70%) are classified as infiltrating duct carcinomas, no special type (IDC-NST), indicating that histologic typing represents a broad categorization rather than a detailed classification. Cancers of the same histologic type may show vastly different biological behavior. Thus, assessment of these parameters may not capture the varied clinical courses of individual breast cancers. In the current era of personalized medicine, a better understanding and classification is called for.

The underlying genetic alterations and the biological events involved in cancer development and progression are complex. Recent studies have focused on refining breast cancer classification using information from high-throughput technologies. Detailed biological characterization of genomic alterations may aid prognostication and risk stratification and thus better tailoring of treatment to patients. The molecular information also provides an opportunity for novel targeted treatment directed at the underlying molecular aberrations driving individual tumor growth. This review covers an overview of recent data on high throughput analysis of breast cancers and the relationship of these findings with traditional breast cancer classification and their clinical potentials.

TRADITIONAL CLASSIFICATIONS

Histologic Type

Histologic classification of breast cancers is based on the pathologic growth pattern. There are over 20 different histologic types of invasive breast cancers. The most common is IDC-NST, which accounts for 70% to 80% of all invasive cancers, followed by invasive lobular carcinomas (ILC) (around 10% of all invasive cancers). The remainder are the less common histologic types, such as mucinous, cribriform, micropapillary, papillary, tubular, medullary, metaplastic, and apocrine carcinomas.[1] Classification into histologic types is based on a wide range of criteria, including tumor cell type (eg, carcinoma with apocrine features), extracellular secretion (eg, mucinous carcinoma), architectural features (eg, papillary carcinoma), and immunohistochemical profile (eg, carcinoma with neuroendocrine differentiation).[1,2] IDC-NST do not exhibit specific morphologic characteristics of any other more specific histologic types; thus, most breast cancers fall into a single category (ie, IDC-NST). This classification cannot fully reflect the biological heterogeneity of breast cancers.

Grading

Grade encompasses microscopic assessment of histologic differentiation in the form of tubule formation, nuclear pleomorphism, and proliferation as indicated by mitotic index. The currently widely accepted Nottingham modification of the Scarff-Bloom Richardson grading evaluates each parameter with a numerical scoring system of 1 to 3 and produces a summation score for grade assignment. Grading is a powerful prognostic factor and serves as an integral component in a number of clinical decision tools such as the Nottingham prognostic index and Adjuvant online.[3] Interestingly, specific genetic and transcriptomic features of breast cancers were associated with specific and different tumor grades.[4–7]

Immunophenotype (Estrogen Receptor, Progesterone Receptor, Human Epidermal Growth Factor Receptor 2)

ER, PR, and HER2 assessment is routine in breast cancer management. They are prognostic markers and important predictive factors for hormonal and anti-HER2-targeted therapy. ER and PR are nuclear sex steroid receptors that stimulate the growth of normal and neoplastic breast

epithelium. They are expressed in ~75% of all breast cancers. When present, they are bona fide indicators for responsiveness to hormonal therapy. ER and PR are assessed immunohistochemically with a 1% nuclear expression cutoff.[8] ER/PR-positive cancers are usually low grade and less aggressive. The majority of ER-positive cancers are also PR+. However, a small percentage of breast cancers show single hormone receptor positivity. These tumors seem to be more aggressive and less responsive to hormonal therapy compared with ER/PR-positive cancers.[9,10] Approximately 15% of breast cancers overexpress HER2 with amplification of the corresponding gene at 17q12.[11] HER2 status is tested using a combination of immunohistochemical (IHC) and DNA in situ hybridization techniques. There are detailed guidelines for its assessment, which are regularly updated.[12] Currently, HER2-positive tumors are defined by >10% of cells with strong circumferential staining or HER2:CEP17 ratio ≥2. HER2 overexpression is associated with aggressive clinical course and poor prognosis, but is also predictive of response to anti-HER2 targeted treatments.[13] The remaining 10% to 15% of breast cancers that express none of these 3 markers are termed triple-negative breast cancers (TNBC). TNBC are in general high grade and associated with a poor prognosis. Patients with TNBC do not benefit from the current targeted therapies.[14]

Tumor Size, Nodal Status, and Distant Metastasis Staging

TNM staging, published by the American Joint Committee on Cancer, uses both clinical and pathologic information of tumor size (T), the status of regional lymph nodes (N), and distant metastases (M). The staging combines these factors and stratifies the disease into one of 5 stages (0, I, II, III, and IV). In the latest edition (AJCC-TNM8), the information on grade and ER, PR, and HER2 has also been incorporated to form the prognostic staging. This prognostic staging overcomes the limitation of evaluation of the anatomical disease extent alone and takes into account biological parameters that have predictive and prognostic value, and it provides more accurate prognostic information than the former staging systems.[15]

MOLECULAR CLASSIFICATION

Intrinsic Subtypes

High-throughput technologies have provided direct evidence of breast cancers' heterogeneity at the molecular level and led to changes in the paradigms in breast cancer biology. Global gene expression profiling studies classified breast cancers into 5 intrinsic subtypes by hierarchical clustering,[16] namely luminal A, luminal B, HER2-overexpressing, basal-like breast cancers (BLBC), and normal-like tumors. These studies also demonstrated the relevance of the immunophenotypic classification by hormone receptors and HER2 status. ER expression stratifies breast cancers into 2 distinct clusters: ER+ and ER−. Luminal A and B subtypes were enriched with ER-positive cancers, whereas HER2-overexpressing, BLBC and normal-like tumors were ER−. Apart from the expression of ER and ER-related genes, different subtypes showed differential expression of genes related to proliferation, HER2 amplicon, and myoepithelial cells. These different intrinsic subtypes revealed differences in incidence, clinical and pathological features, and to a large extent, overlapped with the established clinical and histopathologic classifications.

Luminal A and B subgroups are characterized by gene expression profiles resembling normal luminal epithelial cells of the breast and other genes associated with ER

activation.[16] Luminal A is the most common molecular subtype, representing 40% to 50% of invasive breast cancers.[17] Typically, luminal A cancers are low grade, with the best prognosis among all intrinsic subtypes. Luminal B cancers tend to be higher grade and have a worse prognosis than luminal A. They show lower expression of ER-related genes, but higher expression in proliferation-related genes and variable expression of HER2-related genes than luminal A cancers. Clinically, the luminal A group is likely to benefit from hormonal therapy alone, whereas luminal B tumors may be candidates for additional chemotherapy.

The HER2-overexpressing subtype, comprising ~15% of all invasive breast cancers, is characterized by the overexpression of HER2/HER2 signaling-associated genes and genes located in HER2 amplicon on chromosome 17q12.[16] HER2-overexpressing tumors are likely to be high grade, ER and PR−, and run an aggressive clinical course. Nonetheless, they are highly responsive to anti-HER2-targeted therapy, resulting in a greatly improved outcome. Not all cancers of HER2-overexpressing subtype are clinically HER2+ and vice versa. A minority of HER2-positive cancers coexpress ER and are classified as luminal B.[18]

The BLBC are associated with the expression of genes in normal mammary basal/myoepithelial cells, including basal cytokeratins.[16] They also show overexpression of proliferation-related genes but lack ER, PR, and HER2-related gene expression. Histologically, BLBC are usually high grade, with high proliferation index and show triple-negative phenotype. BLBC patients have poor prognosis, and relapses may occur within 5 years after diagnosis.[19]

The normal-like cluster identified in the initial study was characterized by expression of genes similar to normal breast epithelium. However, it is a controversial subgroup and has been later considered to be an artifact due to true normal epithelial cell contamination of a low malignant cell content tumor.[20] Table 1 displays the key features of different intrinsic subtypes.

These intrinsic subtypes have been reproduced by other studies using varying numbers of genes in the signature.[21–23] However, the assignment of the individual tumor to any subtype showed only moderate reproducibility depending on the array platform used, the composition of the entire tumor population, and setting of gene expression threshold.[24,25] BLBC were the most reproducible, whereas luminal B and HER2 subtypes were the least.[25] PAM50 (Prediction Analysis of Microarray using 50 classifier genes plus 5 reference genes) has been developed as a standardized method that categorizes breast cancers into luminal A, luminal B, HER2-enriched, and BLBC. On the basis of a modified version of PAM50, Prosigna, a test approved by US Food and Drug Administration, was developed for prognostication of postmenopausal hormonal receptor–positive breast cancer patients. Prosigna shows predictive value for distant recurrence,[26] benefit of adjuvant chemotherapy,[18] and in predicting response to neoadjuvant therapy[27] and occurrence of late recurrence.[28]

Subsequent studies' gene expression profiling identified additional rare subtypes, including claudin low,[29] molecular apocrine,[30] and interferon rich.[22] However, these subtypes were only recognized in hierarchy clustering, and they lack representative signatures for identification as separate entities.[31]

The cost and technical complexities limited the application of gene expression profiling in daily clinical practice. IHC-based surrogate molecular classification has been advocated. IHC analysis of ER, PR, HER2, and Ki-67 is used for the identification of different subtypes:

TABLE 1. Overview of Different Molecular Subtypes

Intrinsic Subtype	Gene Profile	Molecular Findings	IHC Phenotype	Histologic Subtypes	Integrative Cluster	DNA Architecture	Survival
Luminal A	High expression of luminal epithelial genes and ER-related genes	Mutations in PI3KCA, MAPK3K1, and GATA3; CCDN1 amplification; no corresponding activation of PI3K pathways	ER+, PR ≥ 20%, HER2−, Ki67low	Tubular Carcinoma, low-grade IDC-NST, classic ILC	IntClust 2	11q13/14 amplification; firestorm pattern of high-level copy number gains	Poor
					IntClust 3	Low genomic instability	Good
					IntClust 4	CNA devoid	Good
					IntClust 6	High genomic instability; 8p12 amplification	Intermediate
					IntClust 7	16p gain, 16q loss, 8q amplification	Good
					IntClust 8	1q gain, 16q loss	Good
Luminal B	Lower expression of luminal epithelium and ER-related genes, but higher level of proliferation and HER2-related genes than luminal A	Similar to luminal A but with a higher prevalence of TP53 and RB pathways inactivation as well as Myc-related and FOXM1 related transcription	ER+, PR < 20%/ or HER2+/or Ki67high	IDC-NST, micropapillary carcinoma, pleomorphic ILC	IntClust 1	High genomic instability; 17q23 amplification; GATA3 mutation	Intermediate
					IntClust 2	See above	
					IntClust 5	HER2 amplification	Poor
					IntClust 6	See above	
					IntClust 9	8q gain, 20q amplification	Intermediate
HER2-OE	High expression of HER2-related genes; low expression of ER-related genes	HER2 amplicon and EGFR/HER2 signal protein signature	ER−, PR−, HER2+	High-grade IDC-NST, pleomorphic ILC	IntClust 5	See above	
Basal like	High expression of basal epithelial and proliferation genes; low expression of HER2-related and ER-related genes	Mutations in TP53; losses in RB1 and BRCA1; amplification of MYC; high PI3K/AKT pathway activation	ER−, PR−, HER2-	High-grade IDC-NST, metaplastic carcinoma, medullary carcinoma, adenoid cystic carcinoma	IntClust 10	5q loss, 8q gain, 10p gain, 12p gain; high genomic alterations with sawtooth pattern	Poor
					IntClust 4	See above	

ER indicates estrogen receptor; HER2, human epidermal growth factor receptor 2; IDC-NST, infiltrating duct carcinomas, no special type; IHC, immunohistochemical; ILC, invasive lobular carcinomas; PR, progesterone receptor.

(1) luminal A-like (ER+, PR ≥ 20%, HER2−, Ki67 < 20%),
(2) luminal B-like (ER+, PR < 20% and/or HER2+ and/or Ki67 ≥ 20%),
(3) HER2-overexpression (ER−, PR−, HER2+), and
(4) basal-like (triple negative: ER−, PR−, HER2−).

The classification has been endorsed by the St Gallen Consensus Conference in planning individual patient treatment.[32] It should be noted that the oncologic guideline in molecular classification for patient management is actually based on IHC surrogate rather than gene profiling. In case of discrepancies, patients should be managed according to IHC results. Although IHC classification shows much overlapping with gene expression profiling classification, some discrepancies exist. Only ~80% of TNBC belonged to intrinsic BLBC subtype, and 65% of HER2-positive tumors belong to intrinsic HER2-overexpressing subtype.[33] Moreover, the cutoff for Ki67 is still a matter of debate. The most recent St Gallen consensus adopted 20%; however, interlaboratory variation in measurement and cutoff exists. The other contentious issue is the low ER expression (1% to 9%). These tumors are rare and are usually classified as luminal cancers on the basis of the current criteria. However, it has been suggested that these tumors were more similar to BLBC both molecularly and biologically.[34] Their clinical outcome is controversial, with some reporting similarity to ER-negative cancers, and others showing an intermediate outcome between ER-negative and ER high tumors.[35] Pragmatically, they may be considered luminal cancers for the potential benefit from empirical adjuvant endocrine therapy.

Integrative Clusters

Inherited genetic variation and acquired genomic aberrations can contribute to breast cancer carcinogenesis by inducing abnormal gene expression. To capture the genomic aberration–driven gene expression changes in molecular classification, an integrated analysis of 2000 breast cancers on gene expression profiling and genomic alterations for class discovery was reported by the Molecular Taxonomy of Breast Cancer International Consortium (METABRIC).[36] Ten integrative clusters (IntClust 1 to 10) have been assigned, and each was associated with distinct copy number aberrations (CNAs) and gene expression changes. These 10 IntClust subtypes partly captured the intrinsic subtyping and also showed the genomic heterogeneity within the individual subtype (Table 1). Findings from IntClust subtypes have been validated in a large cohort of breast cancers.[37] IntClust 1, 2, 3, 6, 7, and 8 are mainly ER+ and luminal intrinsic subtypes. IntClust 1, comprising mainly luminal B cancers, showed that intermediate prognosis is characterized by amplification of 17q23 and high genomic instability. GATA3 mutations are frequently found in IntClust 1, a feature that separates this subtype from other luminal B tumors. IntClust 2 includes both luminal A and luminal B cancers and shows 11q13/14 amplification with characteristic firestorm pattern of high-level copy number gains. Several known and putative driver genes reside in this region, including CCND1 (11q13.3), EMSY (11q13.5), PAK1 (11q14.1), and RSF1 (11q14.1), and these genes have been previously linked to breast or ovarian cancers. Remarkably, this group is associated with the worst prognosis of all luminal cancers. IntClusts 3, 7, and 8 are comprised primarily of luminal A cancers and have a good prognosis. IntClust 3 has very few genomic alterations and is one of the most prevalent IntClust subgroups. IntClusts 7 and 8 have an intermediate level of genomic alterations. IntClust 7 shows specific 16p gain/16q loss with higher frequencies of 8q amplification but lacks the 1q alteration, while IntClust 8 shows

the classical 1q gain/16q loss and corresponds to a common translocation event. Intclust 6 encompasses both luminal A and luminal B cancers with intermediate prognosis and is characterized by the specific amplification of the 8p12 locus and high level of genomic instability.

IntClusts 4, 5, and 9 can be both luminal and non-luminal tumors. IntClust 4, similar to IntClust 3, is a prevalent IntClust subtype, showing a low level of genomic alterations and good prognosis. It has an essentially flat copy number landscape; this is termed the "CNA devoid" subtype. Histologically, most of the IntClust 4 tumors exhibit extensive lymphocytic infiltration. IntClust 5 is dominated by a high level of amplification on 17q12 encompassing HER2 gene, including mainly luminal B and HER2-enriched subtypes. IntClust 9 comprises a mixture of intrinsic subtypes and shows an intermediate prognosis. It is characterized by 8q cis-acting alterations and 20q amplification.

IntClust 10 includes the majority of BLBC, mostly with high genomic instability, and shows a sawtooth pattern of alterations. These tumors rarely show high-level amplifications, but copy number alterations, including 5q loss and gains in 8q, 10p, and 12p, are common. Numerous signaling molecules, transcription factors, and cell division genes were associated in trans with a 5q loss. These cancers represent a high-risk group in the first 5 years of diagnosis but have a relatively good long-term outcome.

Next-Generation Sequencing

Next generation sequencing (NGS) technologies allow a comprehensive characterization of the mutational landscape of breast cancers, including base substitutions, small insertion/deletion, and structural rearrangements. NGS studies have established a repertoire of driver genes' mutations and copy number alterations in breast cancers and furnished further information on the molecular classification.[38–42] At least 40 driver genes in breast cancers have been identified. The most frequently mutated genes were TP53, PIK3CA, GATA3, MYC, CCND1, PTEN, FGFR1, RB1, ERBB2, and MAP3K1.[40,41] Generally, mutations are rare in breast cancers, and only the 3 most common mutations (TP53, PIK3CA, and GATA3) have incidences over 10%.[39] Many of these mutations were associated with distinct clinical and pathologic features. Variations in mutation frequency and mutation type were found across different intrinsic or IntClust subtypes. For instance, there was a subtype-dependent distribution of PIK3CA mutations, being more prevalent in luminal A and B cancers as well as IntClusts 3, 7, and 8. By contrast, TP53 mutations were dominant in basal-like cancers and IntClust 10.[39,41] The overall rate of mutation was lowest in luminal A subtype, but this subtype harbors the most significantly mutated genes. The highest overall mutation rate was found in basal-like and HER2-enriched breast cancer.[39]

Single base mutations in DNA do not occur randomly but in a distinct context related to the underlying etiology of the tumor.[43] Different mutational processes, including carcinogen exposures, aberrant DNA editing, replication errors, and defective DNA repair, imprint particular patterns of mutations on cancer genomes, that is, mutational signatures.[44,45] Whole-genome sequencing on 560 breast cancers revealed 12 base substitution and 6 rearrangement signatures.[40] Three of the rearrangement signatures, characterized by tandem duplications or deletions, seem to be associated with defective homologous recombination-based DNA repair. Each tumor may harbor > 1 mutational signature. A characteristic set of mutational signatures (HRDetect) has been identified to predict BRCA1/2

deficiency, even for those patients without detectable aberrations in the genes,[46] thus potentially for selecting breast cancers that are sensitive PARP inhibitors. A similar approach has been developed for the identification of mismatch repair deficiency in breast cancers.[47] Mismatch repair–deficient breast cancers, although rare, have been discovered to be highly responsive to immune therapies such as PD-1 checkpoint blockade, making their identification of paramount importance.

Multiomics Approach

High-throughput technologies have amassed multiple layers of information on the molecular characteristics of breast cancers, including gene expression, genomic alterations, epigenetic changes, and protein expression. While each layer already provides a wealth of information, combining these provides a powerful repertoire of biological information. Classification of breast cancers from these different analyses showed a high degree of overlap, reflecting a common biological trait in different classifications. An integrated analysis may provide a more holistic picture of the genetic complexities of breast cancers. The Cancer Genome Atlas Network (TCGA) has utilized data from 6 different platforms, including mRNA expression microarray, DNA methylation chips, single-nucleotide polymorphism arrays, microRNA (miRNA) sequencing, whole-exome sequencing, and reverse-phase protein array to examine specific genetic, epigenetic, and proteomic alterations in breast cancers, providing further insights into the genetic landscape of breast cancers.[39]

The integrated information across platforms showed the existence of 4 main breast cancer classes that correlated well with the mRNA intrinsic subtypes (Table 1). Luminal cancers were the most heterogenous in terms of gene expression, DNA alterations, and patient outcomes. A luminal expression signature can be detected in these cancers at both mRNA and protein levels. They have a high prevalence of PIK3CA mutation but without activation of the corresponding PI3K pathways. MAP3K1 and MAP2K4 mutations, frequent CCND1 amplification, and high MYB protein were also featured in luminal cancers. The pathway differences in luminal A and B relate mainly to hyperactivation of transcriptional activity associated with MYC and FOXM1 proliferation. TP53 and RB1 pathways were also differentially inactivated in the 2 luminal subtypes, and these pathways are expressed more commonly in luminal B cancers. The clinically HER2+ cancers can be divided into 2 different subgroups: HER2-enriched group, which is associated with HER2 amplicon and EGFR/HER2 signal protein signature, and HER2+ luminal cancers with a higher expression luminal cluster of genes. These 2 subgroups showed differential gene mutations. The signature of the HER2-enriched group may allow response prediction to anti-HER2-targeted therapy, whereas the HER2+ luminal cancers show mutational profile more similar to luminal cancers. BLBC proved to be distinct on every platform, highlighting TP53 mutations, which were most prevalent, and losses of RB1 and BRCA1, amplification of MYC, as well as high PI3K/AKT pathway activation. These molecular features overlap with changes in high-grade serous ovarian cancers. Two additional novel groups, named reactive I and II, which relate to stroma/microenvironment elements were defined on the basis of protein expression. The 2 groups differed only in mRNA and protein expression, but not in their DNA alterations, miRNA expression, and methylation profile.[39]

Results of transcriptional profiling and RRPA platforms correlated with the consensus clustering using information content from CNAs, miRNA expression, and DNA methylation, indicating that the heterogeneity of different molecular subtypes of breast cancers was largely captured at the level of gene transcription and protein function. Thus, this provides sound evidence that diverse genetic and epigenetic alterations may converge phenotypically into the same gene/protein expression classes.

MOLECULAR BASIS OF TRADITIONAL CLASSIFICATION

Grade is one of the most important traditional prognostic features in breast cancer classification. There is evidence that breast cancers with different grades show genomic and transcriptomic characteristics. The 3 morphologic features for grading, namely tubule formation, mitosis, and nuclear pleomorphism, showed different molecular traits, with tubule formation being the most dissimilar. Tumors with medium/high mitotic counts and marked nuclear pleomorphism shared TP53 mutation, high PAM50 proliferative score, and basal-like features on the basis of methylation and miRNA profile. Tumors with poor epithelial tubule formation showed more frequent CDH1 mutation, association with luminal A subtype, and inflammation gene set. The transcriptomic signature of the tubule-poor tumors showed a mixture of molecular traits and was more prognostic than grade in ER-positive breast cancer.[48]

Unlike IDC-NST that may actually belong to any intrinsic subtypes, special histologic subtypes were associated with distinct histologic appearance, clinical behavior, and a specific intrinsic subtype. These special subtypes can be separated on the basis of ER status. The ER-positive group includes ILC, tubular cancers, micropapillary cancers, mucinous cancers, and neuroendocrine cancers. The ER-negative group includes apocrine carcinomas, pleomorphic ILC, adenoid cystic carcinomas, metaplastic carcinomas, and carcinomas with medullary features. Accordingly, they fell into the relevant intrinsic subtypes, and, indeed, these special subtypes were more genetically homogenous than IDC-NST, with most tumors belonging to one intrinsic subtype. Tubular, mucinous, and neuroendocrine tumors were classified into the luminal subtype, while adenoid cystic, medullary, and metaplastic carcinomas were classified into the BLBC subtype.[49] ILC was found mainly in the luminal subtypes but infrequently in HER2-enriched and BLBC subtypes. It is an interesting paradox that some special subtypes with good prognosis, like adenoid cystic carcinomas and secretory carcinomas, fell within the BLBC group, which is generally considered to have a poor outcome, thus underscoring the heterogeneity existing within the intrinsic subtyping. Moreover, within the same intrinsic subtype, there were genetic and transcriptomic differences between special subtypes and IDC-NST. Metaplastic cancers, although mostly classified as BLBC/TNBC, were clinicopathologically and genetically distinct from their IDC-NST counterparts. Metaplastic breast cancers showed more downregulation of DNA repair pathways, frequent genetic activation of Wnt signaling, and higher expression of genes in epithelial mesenchymal transition compared with IDC-NST of basal-like subtype.[50] They were also enriched for PIK3CA/PIK3R1 and RAS-MAP kinase aberrations compared with other TNBC.[51] Within the group of metaplastic carcinomas, there exist morphologic variants including spindle cell and squamous or heterologous chondroid element, and these morphologic variants showed different transcriptome and mutation landscape, but showed little variations in copy number changes. PI3KCA mutations were inversely related to chondroid morphology,[51–53] whereas those with spindle cell differentiation differed from cancers with chondroid or squamous cell

features in showing the expression of epithelial-to-mesenchymal transition-related genes.[53]

For ILC, recent molecular studies identified numerous molecular features discriminating it from IDC-NST. Apart from the well-known E-cadherin loss, ILC showed a lower incidence of GATA3 mutations and expression, accompanied by a differential modulation of ER signaling, compared with matched IDC-NST.[54] In addition, ILC also showed more frequent mutations targeting PTEN, TBX3, PYGM, and FOXA1, and AKT activation.[54–56] The findings suggested a different pathogenetic pathway for ILC. AKT activation was consistently increased in ILC versus IDC, making selective inhibition of this pathway in ILC a particularly attractive treatment strategy.[54] Other recurrent therapeutically relevant targets, including HER2 and HER3 were also altered in ILC, particularly the recurrent tumors.[55] In addition, ILC in luminal A subtype can be further separated into 3 groups by mRNA expression: reactive like, immune related, and proliferative classes. These subtypes showed no distinguishing somatic mutations or DNA copy number alterations, but different clinical outcome.[54]

REFINING SUBTYPES

Among each of the 4 intrinsic subtypes, there is still significant heterogeneity, particularly in BLBC/TNBC and luminal A cancers. Luminal A tumors, representing the most frequent subtype, showed great heterogeneity, as shown by its widely spread pattern in integrative cluster analysis (IntClusts 2, 3,4, 7, and 8).[36] On the basis of somatic mutations and copy number variations, luminal A cancers can be further classified into 5 different subtypes: CNA quiet, 1q/16q, chromosome 8 associated, CNA-high and mixed subtypes. Each subtype shows specific mutational landscapes, with PIK3CA mutations enriched in the 1q/16q subtype, and MAP3K1 mutations in the Chr8-associated subtype. The CNA-high subtype showed the worst prognosis and overexpressed regulators of mitosis,[57] and showed TP53 mutations. Those genes were associated with the 5q loss in BLBC.[36] Another study that performed cluster-of-clusters' analysis on the basis of gene/protein/miRNA expression, CNAs, and metabolic profiles found that luminal A cancers mainly fell into 2 of their identified clusters. The differences between the 2 groups were mainly due to 71 differentially expressed miRNAs. The 2 subgroups seemed to show prognostic differences across several data sets. One of the clusters corresponded to the reverse-phase protein array defined reactive I and II subgroups in the TCGA analysis.[58]

Despite the apparent homogeneity of BLBC/TNBC in the integrative cluster and TCGA analysis,[36,39] TNBC could be heterogenous in both morphologic and molecular levels (Fig. 1). Six subtypes of TNBC, including 2 basal-like (BL1 and BL2), an immunomodulatory (IM), a mesenchymal (M), a mesenchymal stem-like (MSL), and luminal androgen receptor (LAR) subtypes were identified by gene expression profiling.[59] The BL1 cancers were enriched with genes involved in cell proliferation and DNA damage
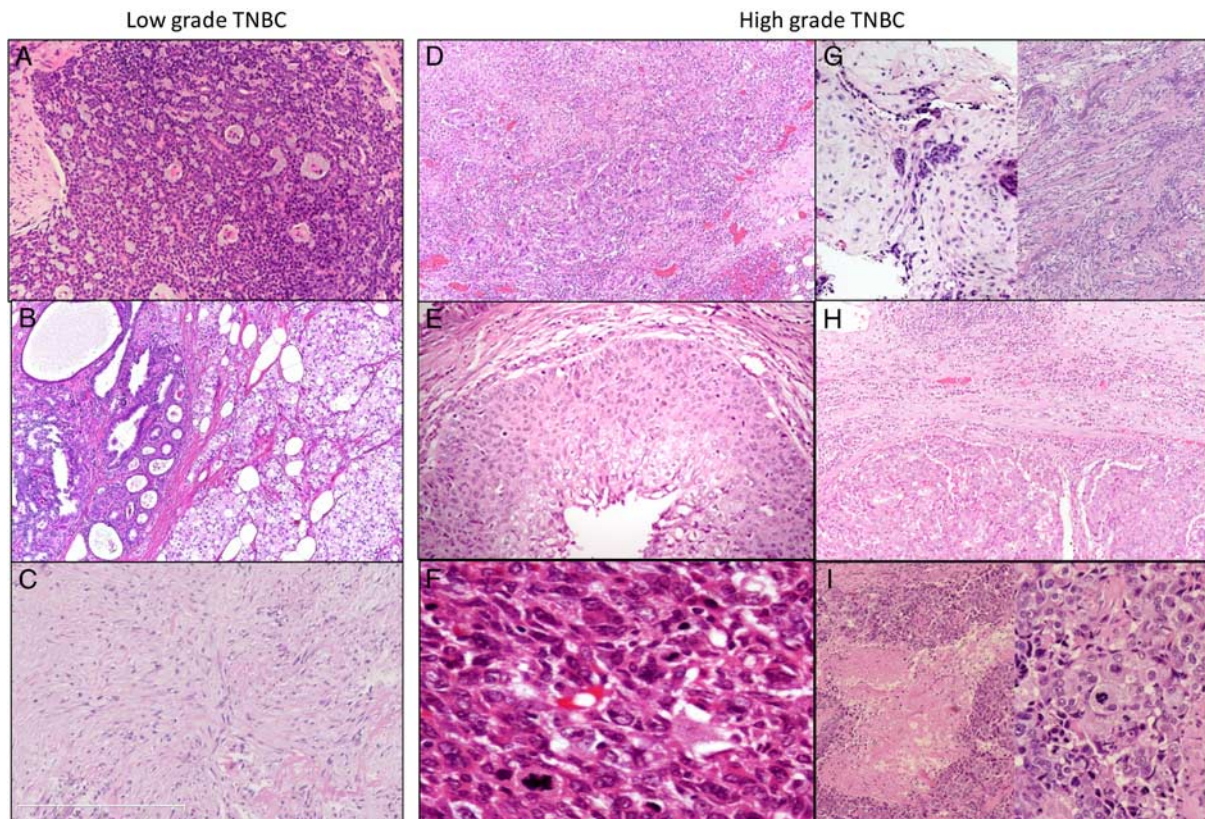

FIGURE 1. The diverse morphologic spectrum of triple-negative breast cancers (TNBCs). Some TNBCs are considered low grade, including adenoid cystic carcinoma (A), secretory carcinoma (B), and fibromatosis-like metaplastic carcinoma (C). Others are high grade, including infiltrating duct carcinomas, no special type (D), squamous cell carcinoma (variant of metaplastic carcinoma) (E), spindle cell carcinoma (variant of metaplastic carcinoma) (F), carcinoma of heterologous element [chondroid (left panel) and osteoid (right panel)] (G), and carcinoma with medullary features (H) including necrosis (left panel) and high-grade pleomorphic morphology (right panel) (I).

response, whereas the BL2 subtype was associated with genes in growth factor signaling, glycolysis, and gluconeogenesis. The IM subtype involved genes in immune cell processes. The M and MSL subtypes expressed genes in epithelial mesenchymal transition and stem cell properties. LAR subtype was enriched with genes in androgen receptor signaling.[59] All BL1, BL2, and IM subtypes were classified into the basal-like intrinsic subtype, and all LARs were non–basal like by gene expression profiling.[60] However, using histopathologic quantification and laser capture microdissection to examine transcripts in the IM and MSL groups, the genes defining the 2 groups were later shown to be from tumor-infiltrating lymphocytes and peritumoral stromal cells, respectively; thus, the classification had been revised into 4 subtypes (BL1, BL2, M, and LAR).[61] On the basis of both RNA and DNA profiles, others have divided TNBC into 4 subtypes: basal-like immune suppressed (BLIS), basal-like immune activated (BLIA), mesenchymal, and LAR subtypes.[62] BLIS subtype exhibited downregulation of immune-regulating pathways and cytokine pathways and vice versa for the BLIA subtype. This classification overlaps with the earlier gene expression profiling classification, in particular, the LAR and mesenchymal groups. BL1 and BL2 subtypes were spread among the BLIA and BLIS. Importantly, both classifications showed significant correlation with clinical outcome.[60,62] These subtypes also showed differential response to therapies. In cell line studies, BL1 and BL2 preferentially responded to cisplatin, M and MSL subtypes responded to PI3K/mTOR as well as Abl/Src inhibition, and LAR was sensitive to AR antagonist.[59] In TNBC patients, those with BL1 showed a higher response rate to neoadjuvant therapy than the other subtypes.[60,61] A number of clinical trials using androgen receptor–targeting therapy for treatment of AR-positive TNBC, currently, are underway.[63]

Integrative multiomics analysis showed specific differences in mutational and copy number profiles characterizing each TNBC molecular subtype, and these have direct bearings in treatment selection.[64] BL1 cancers were characterized by high genomic instability, high copy number losses for TP53, BRCA1/2 and RB1 genes, and high copy number gains for PPAR1 gene, suggesting they may respond to PARP inhibitors. In addition, BL1 cancers may also be a candidate for MEK1/2 inhibitors or PI3K/AKT inhibitors, and they also display copy number gains for KRAS, NRAS, and BRAF as well as PIK3CA, with significant mRNA overexpression of the corresponding gene. LAR and MSL tumors retained RB1 while showing significantly lower CDK4 and CDK6 mRNA expression level. In cell line models, LAR and MSL tumors showed higher sensitivity to CDK4/6 inhibition, especially for the former.[65] Of note, CDK4/6 inhibitors were synergistic with PI3K inhibitors in PIK3CA-mutant TNBC cell lines.[65] With 75% of LAR having somatic mutations in PI3K signaling pathway, those patients may benefit from a combined treatment of the 2 inhibitors.

PERSPECTIVE

The traditional classifications of breast cancers based on pathologic features and IHC evaluation of hormonal receptors and HER2 have been well established for their clinical application and validity. They are inexpensive and can be easily applicable in routine practice. The advent of high-throughput technologies revealed an unprecedented amount of data on transcriptomic, epigenetic, genomic, and proteomic alterations,

providing a more complete description of the pathogenic changes in breast cancers. The vast amount of information now available has undoubtedly enhanced our understanding of breast cancer biology, provided insights into breast cancer stratification, and, more importantly, aid in identifying novel drivers and biomarkers in breast cancer.

Despite the usefulness of the information in formulating the molecular classification of breast cancers, there is still a long road ahead before practical implementation of these findings. The clinical translation to improve treatment strategies and management for individual patients remains in its infancy. Efforts are needed to unify the different findings. To identify a classification on the basis of molecular changes but with clinical relevance is only the first step. There are still hurdles that need to overcome. The identification of intrinsic subtypes to the clinical translation into an approved diagnostic assay has taken over 10 years of development. To date, many technical and interpretive complex high-throughput technologies are still mainly for research purposes. Although NGS has become more affordable and NGS panels are regularly used for cancer diagnostics, an NGS panel has yet to be designed for molecular classification of breast cancers. Many other techniques will take time to evolve into a robust economical diagnostic tool. The findings from high-throughput analysis are often affected by the various bioinformatics pipelines used. The molecular subgroups identified by the test may not be stable. Empirical data are required to support whether the classification represents a robust recurrent and meaningful subtype. As a natural extension of development in molecular classification, the list of potentially targetable cancer driver genes is growing in each subtype. Yet, multiple genetic variants might alter a response to a given therapy. Hence, the presence of an alteration does not guarantee a therapeutic response. Many targeted inhibitors have yet to be proven clinically effective when matched with their specific mutation. Further well-designed trials and studies are definitely required.

As of now, the molecular classification should be used as a complement to the histopathologic assessment and not their replacement. The traditional histologic variables' assessment, when adequately carried out, provides a simple, inexpensive, and highly accurate method for assessing tumor biological characteristics and patient prognosis. It is indispensable in many parts of the world with lower resources. Evidence from molecular analysis provides convincing proof supporting the relevance of histopathologic features in the underlying tumor biology. Although molecular classification, in general, may provide additional prognostic information, it is not without limitations, for instance, in some special subtypes. Therefore, information derived from molecular classification should be provided to complement histopathologic classification.

It is conceivable that, in the future, more genetic information of breast cancers will be available and provide data as basis for more subtypes' classification or finer division between the subtypes. These, however, will need to be put into a clinical perspective, and the ultimate aim of any classification is for accurate prognostication and treatment of breast cancer patients.

REFERENCES

1. Lakhani SR, Ellis IO, Schnitee SJ, et al. *World Health Organisation Classification of Tumours of the Breast*. Lyon: IARC Press; 2012.

2. Rosai J. *Rosai and Ackerman's Surgical Pathology*, 10 ed. Lyon, France: Elsevier; 2011.
3. Mook S, Schmidt MK, Rutgers EJ, et al. Calibration and discriminatory accuracy of prognosis calculation for breast cancer with the online adjuvant! Program: a hospital-based retrospective cohort study. *Lancet Oncol.* 2009;10:1070–1076.
4. Ivshina AV, George J, Senko O, et al. Genetic reclassification of histologic grade delineates new clinical subtypes of breast cancer. *Cancer Res.* 2006;66:10292–10301.
5. Sotiriou C, Pusztai L. Gene-expression signatures in breast cancer. *N Engl J Med.* 2009;360:790–800.
6. Sotiriou C, Wirapati P, Loi S, et al. Gene expression profiling in breast cancer: understanding the molecular basis of histologic grade to improve prognosis. *J Natl Cancer Inst.* 2006;98:262–272.
7. Weigelt B, Baehner FL, Reis-Filho JS. The contribution of gene expression profiling to breast cancer classification, prognostication and prediction: a retrospective of the last decade. *J Pathol.* 2010;220:263–280.
8. Hammond ME, Hayes DF, Dowsett M, et al. American society of clinical oncology/college of american pathologists guideline recommendations for immunohistochemical testing of estrogen and progesterone receptors in breast cancer. *J Clin Oncol.* 2010;28:2784–2795.
9. Cui X, Schiff R, Arpino G, et al. Biology of progesterone receptor loss in breast cancer and its implications for endocrine therapy. *J Clin Oncol.* 2005;23:7721–7735.
10. Ethier JL, Ocana A, Rodriguez Lescure A, et al. Outcomes of single versus double hormone receptor-positive breast cancer. A geicam/9906 sub-study. *Eur J Cancer.* 2018;94:199–205.
11. Slamon DJ, Clark GM, Wong SG, et al. Human breast cancer: correlation of relapse and survival with amplification of the her-2/neu oncogene. *Science.* 1987;235:177–182.
12. Wolff AC, Hammond MEH, Allison KH, et al. Human epidermal growth factor receptor 2 testing in breast cancer: American Society Of Clinical Oncology/College Of American Pathologists clinical practice guideline focused update. *J Clin Oncol.* 2018;36:2105–2122.
13. Bedard PL, Cardoso F, Piccart-Gebhart MJ. Stemming resistance to her-2 targeted therapy. *J Mammary Gland Biol Neoplasia.* 2009;14:55–66.
14. Yam C, Mani SA, Moulder SL. Targeting the molecular subtypes of triple negative breast cancer: understanding the diversity to progress the field. *Oncologist.* 2017;22:1086–1093.
15. Weiss A, Chavez-MacGregor M, Lichtensztajn DY, et al. Validation study of the american joint committee on cancer eighth edition prognostic stage compared with the anatomic stage in breast cancer. *JAMA Oncol.* 2018;4:203–209.
16. Perou CM, Sorlie T, Eisen MB, et al. Molecular portraits of human breast tumours. *Nature.* 2000;406:747–752.
17. Voduc KD, Cheang MC, Tyldesley S, et al. Breast cancer subtypes and the risk of local and regional relapse. *J Clin Oncol.* 2010;28:1684–1691.
18. Parker JS, Mullins M, Cheang MC, et al. Supervised risk predictor of breast cancer based on intrinsic subtypes. *J Clin Oncol.* 2009;27:1160–1167.
19. Badve S, Dabbs DJ, Schnitt SJ, et al. Basal-like and triple-negative breast cancers: a critical review with an emphasis on the implications for pathologists and oncologists. *Mod Pathol.* 2011;24:157–167.
20. Peppercorn J, Perou CM, Carey LA. Molecular subtypes in breast cancer evaluation and management: divide and conquer. *Cancer Invest.* 2008;26:1–10.
21. Calza S, Hall P, Auer G, et al. Intrinsic molecular signature of breast cancer in a population-based cohort of 412 patients. *Breast Cancer Res.* 2006;8:R34.
22. Hu Z, Fan C, Oh DS, et al. The molecular portraits of breast tumors are conserved across microarray platforms. *BMC Genomics.* 2006;7:96.
23. Sorlie T, Tibshirani R, Parker J, et al. Repeated observation of breast tumor subtypes in independent gene expression data sets. *Proc Natl Acad Sci U S A.* 2003;100:8418–8423.
24. Mackay A, Weigelt B, Grigoriadis A, et al. Microarray-based class discovery for molecular classification of breast cancer: analysis of interobserver agreement. *J Natl Cancer Inst.* 2011;103:662–673.
25. Weigelt B, Mackay A, A'Hern R, et al. Breast cancer molecular profiling with single sample predictors: a retrospective analysis. *Lancet Oncol.* 2010;11:339–349.
26. Gnant M, Filipits M, Greil R, et al. Predicting distant recurrence in receptor-positive breast cancer patients with limited clinicopathological risk: using the pam50 risk of recurrence score in 1478 postmenopausal patients of the abcsg-8 trial treated with adjuvant endocrine therapy alone. *Ann Oncol.* 2014;25:339–345.
27. Prat A, Galvan P, Jimenez B, et al. Prediction of response to neoadjuvant chemotherapy using core needle biopsy samples with the prosigna assay. *Clin Cancer Res.* 2016;22:560–566.
28. Filipits M, Nielsen TO, Rudas M, et al. The pam50 risk-of-recurrence score predicts risk for late distant recurrence after endocrine therapy in postmenopausal women with endocrine-responsive early breast cancer. *Clin Cancer Res.* 2014;20:1298–1305.
29. Prat A, Parker JS, Karginova O, et al. Phenotypic and molecular characterization of the claudin-low intrinsic subtype of breast cancer. *Breast Cancer Res.* 2010;12:R68.
30. Farmer P, Bonnefoi H, Becette V, et al. Identification of molecular apocrine breast tumours by microarray analysis. *Oncogene.* 2005;24:4660–4671.
31. Russnes HG, Lingjaerde OC, Borresen-Dale AL, et al. Breast cancer molecular stratification: from intrinsic subtypes to integrative clusters. *Am J Pathol.* 2017;187:2152–2162.
32. Goldhirsch A, Winer EP, Coates AS, et al. Personalizing the treatment of women with early breast cancer: highlights of the St Gallen International expert consensus on the primary therapy of early breast cancer 2013. *Ann Oncol.* 2013;24:2206–2223.
33. Prat A, Carey LA, Adamo B, et al. Molecular features and survival outcomes of the intrinsic subtypes within her2-positive breast cancer. *J Natl Cancer Inst.* 2014;106:dju152.
34. Cheang MC, Martin M, Nielsen TO, et al. Defining breast cancer intrinsic subtypes by quantitative receptor expression. *Oncologist.* 2015;20:474–482.
35. Morigi C. Highlights from the 15th St Gallen International Breast Cancer Conference 15-18 March, 2017, Vienna: tailored treatments for patients with early breast cancer. *Ecancermedicalscience.* 2017;11:732.
36. Curtis C, Shah SP, Chin SF, et al. The genomic and transcriptomic architecture of 2,000 breast tumours reveals novel subgroups. *Nature.* 2012;486:346–352.
37. Ali HR, Rueda OM, Chin SF, et al. Genome-driven integrated classification of breast cancer validated in over 7,500 samples. *Genome Biol.* 2014;15:431.
38. Banerji S, Cibulskis K, Rangel-Escareno C, et al. Sequence analysis of mutations and translocations across breast cancer subtypes. *Nature.* 2012;486:405–409.
39. Cancer Genome Atlas Network. Comprehensive molecular portraits of human breast tumours. *Nature.* 2012;490:61–70.
40. Nik-Zainal S, Davies H, Staaf J, et al. Landscape of somatic mutations in 560 breast cancer whole-genome sequences. *Nature.* 2016;534:47–54.
41. Pereira B, Chin SF, Rueda OM, et al. The somatic mutation profiles of 2,433 breast cancers refines their genomic and transcriptomic landscapes. *Nat Commun.* 2016;7:11479.
42. Stephens PJ, Tarpey PS, Davies H, et al. The landscape of cancer genes and mutational processes in breast cancer. *Nature.* 2012;486:400–404.
43. Forbes SA, Beare D, Boutselakis H, et al. Cosmic: somatic cancer genetics at high-resolution. *Nucleic Acids Res.* 2017;45:D777–D783.
44. Alexandrov LB, Nik-Zainal S, Wedge DC, et al. Deciphering signatures of mutational processes operative in human cancer. *Cell Rep.* 2013;3:246–259.
45. Helleday T, Eshtad S, Nik-Zainal S. Mechanisms underlying mutational signatures in human cancers. *Nat Rev Genet.* 2014;15:585–598.

46. Davies H, Glodzik D, Morganella S, et al. Hrdetect is a predictor of brca1 and brca2 deficiency based on mutational signatures. *Nat Med*. 2017;23:517–525.

47. Davies H, Morganella S, Purdie CA, et al. Whole-genome sequencing reveals breast cancers with mismatch repair deficiency. *Cancer Res*. 2017;77:4755–4762.

48. Heng YJ, Lester SC, Tse GM, et al. The molecular basis of breast cancer pathological phenotypes. *J Pathol*. 2017;241:375–391.

49. Weigelt B, Horlings HM, Kreike B, et al. Refinement of breast cancer classification by molecular characterization of histological special types. *J Pathol*. 2008;216:141–150.

50. Weigelt B, Kreike B, Reis-Filho JS. Metaplastic breast carcinomas are basal-like breast cancers: a genomic profiling analysis. *Breast Cancer Res Treat*. 2009;117:273–280.

51. Krings G, Chen YY. Genomic profiling of metaplastic breast carcinomas reveals genetic heterogeneity and relationship to ductal carcinoma. *Mod Pathol*. 2018;31:1661–1674.

52. McCart Reed AE, Kalaw E, Nones K, et al. Phenotypic and molecular dissection of metaplastic breast cancer and the prognostic implications. *J Pathol*. 2019;247:214–227.

53. Ng CKY, Piscuoglio S, Geyer FC, et al. The landscape of somatic genetic alterations in metaplastic breast carcinomas. *Clin Cancer Res*. 2017;23:3859–3870.

54. Ciriello G, Gatza ML, Beck AH, et al. Comprehensive molecular portraits of invasive lobular breast cancer. *Cell*. 2015;163:506–519.

55. Desmedt C, Zoppoli G, Gundem G, et al. Genomic characterization of primary invasive lobular breast cancer. *J Clin Oncol*. 2016;34:1872–1881.

56. Dieci MV, Smutna V, Scott V, et al. Whole exome sequencing of rare aggressive breast cancer histologies. *Breast Cancer Res Treat*. 2016;156:21–32.

57. Ciriello G, Sinha R, Hoadley KA, et al. The molecular diversity of luminal a breast tumors. *Breast Cancer Res Treat*. 2013;141:409–420.

58. Aure MR, Vitelli V, Jernstrom S, et al. Integrative clustering reveals a novel split in the luminal a subtype of breast cancer with impact on outcome. *Breast Cancer Res*. 2017;19:44.

59. Lehmann BD, Bauer JA, Chen X, et al. Identification of human triple-negative breast cancer subtypes and preclinical models for selection of targeted therapies. *J Clin Invest*. 2011;121:2750–2767.

60. Masuda H, Baggerly KA, Wang Y, et al. Differential response to neoadjuvant chemotherapy among 7 triple-negative breast cancer molecular subtypes. *Clin Cancer Res*. 2013;19:5533–5540.

61. Lehmann BD, Jovanovic B, Chen X, et al. Refinement of triple-negative breast cancer molecular subtypes: implications for neo-adjuvant chemotherapy selection. *PLoS One*. 2016;11:e0157368.

62. Burstein MD, Tsimelzon A, Poage GM, et al. Comprehensive genomic analysis identifies novel subtypes and targets of triple-negative breast cancer. *Clin Cancer Res*. 2015;21:1688–1698.

63. Denkert C, Liedtke C, Tutt A, et al. Molecular alterations in triple-negative breast cancer-the road to new treatment strategies. *Lancet*. 2017;389:2430–2442.

64. Bareche Y, Venet D, Ignatiadis M, et al. Unravelling triple-negative breast cancer molecular heterogeneity using an integrative multiomic analysis. *Ann Oncol*. 2018;29:895–902.

65. Asghar US, Barr AR, Cutts R, et al. Single-cell dynamics determines response to cdk4/6 inhibition in triple-negative breast cancer. *Clin Cancer Res*. 2017;23:5561–5572.