## *Review Article*
# Breast cancer intrinsic subtype classification, clinical use and future trends

Xiaofeng Dai[1], Ting Li[1], Zhonghu Bai[1], Yankun Yang[1], Xiuxia Liu[1], Jinling Zhan[1], Bozhi Shi[2]

[1]National Engineering Laboratory for Cereal Fermentation Technology, School of Biotechnology, Jiangnan University, Wuxi 214122, Jiangsu, China; [2]Bioengineering College, Chongqing University, Chongqing, 404100, China

**Abstract:** Breast cancer is composed of multiple subtypes with distinct morphologies and clinical implications. The advent of microarrays has led to a new paradigm in deciphering breast cancer heterogeneity, based on which the intrinsic subtyping system using prognostic multigene classifiers was developed. Subtypes identified using different gene panels, though overlap to a great extent, do not completely converge, and the avail of new information and perspectives has led to the emergence of novel subtypes, which complicate our understanding towards breast tumor heterogeneity. This review explores and summarizes the existing intrinsic subtypes, patient clinical features and management, commercial signature panels, as well as various information used for tumor classification. Two trends are pointed out in the end on breast cancer subtyping, i.e., either diverging to more refined groups or converging to the major subtypes. This review improves our understandings towards breast cancer intrinsic classification, current status on clinical application, and future trends.

## Introduction

Breast carcinoma is the leading cause among women in most developed countries [1]. It is not a single disease, which comprises of many biologically different entities with distinct pathological features and clinical implications [1-7]. Accumulating evidence has suggested that breast cancers with different histopathological and biological features exhibit distinct behaviors that lead to different treatment responses and should be given different therapeutic strategies [8]. Thus, accurate grouping of breast cancers into clinically relevant subtypes is of particular importance for therapeutic decision making and thus urgently called for.

Classical immunohistochemistry (IHC) markers such as ER, PR and HER2, together with traditional clinicopathological variables including, e.g., tumor size, tumor grade and nodal involvement, are conventionally used for patient prognosis and management [9, 10]. The advent of high-throughput platforms for gene expression analysis such as microarrays has shown that tumor cell response to treatment is not determined by anatomical prognostic factors but rather intrinsic molecular characteristics that can be probed using molecular methods [4-7]. This conceptual change has led to a new paradigm on how breast cancer patients are stratified and treated, which provides an incremental increase on the reproducibility and accuracy of disease prognosis and therapeutic decision making [11]. Integrating information from multiple levels or dissecting this problem from the pathway point of view, has led to an expanding spectrum on breast cancer subtypes or the other way around. With our incremental knowledge on this complex tumorigenesis progress, novel molecules with emerging roles are gaining their importance which, though contribute to deciphering breast cancer heterogeneity, complicate our understanding towards subtyping of this complex disease.

This review clarifies breast tumor intrinsic classification, clinical features and therapeutic strategies of each major intrinsic subtype, and the current status of the commercial classification signatures. Emerging information and novel perspectives on breast cancer subtyping

**Table 1.** Summary of the breast tumor molecular subtypes

| Intrinsic subtype | IHC status | Grade | Outcome | Prevalence△ |
|---|---|---|---|---|
| Luminal A* | [ER+\|PR+] HER2-KI67- | 1\|2 | Good | 23.7% [p1] [10] |
| Luminal B* | [ER+\|PR+] HER2-KI67+ | 2\|3 | Intermediate | 38.8% [p1] [10] |
| | [ER+\|PR+] HER2+KI67+ | | \|Poor | 14% [p1] [10] |
| HER2 over-expression* | [ER-PR-] HER2+ | 2\|3 | Poor | 11.2% [p1] [10] |
| Basal* | [ER-PR-] HER2-, basal marker+ | 3 | Poor | 12.3% [p1] [10] |
| Normal-like | [ER+\|PR+] HER2-KI67- | 1\|2\|3 | Intermediate | 7.8% [p2] [15] |

*Subtypes with detailed expression patterns and clinical implications discussed in the text, which take the majority of the breast tumor cases and are most commonly referred to. △The prevalence of each subtype is taken from the publication indicated in the square bracket.

are provided, with the future trends forecasted to conclude this paper.

## Gene expression profiling and intrinsic subtypes

With the development of microarrays, gene expression profiling (GEP) has been used for breast cancer prognosis, specifically aiming at identifying patients with sufficiently good prognosis to allow the safe omission of adjuvant chemotherapy [6, 7]. The pioneer studies conducted by Sørlie et al. reported a distinctive 'molecular portrait' of breast cancer using 456 cDNA clones, according to which tumors were classified into five intrinsic subtypes with distinct clinical outcomes, i.e., luminal A, luminal B, HER2 over-expression, basal and normal-like tumors [12, 13]. The rational underlying such classification is that the differences underlying the gene expression patterns among cancer subtypes reflect the fundamental differences of the tumors at the molecular level [14]. Each of the five intrinsic subtypes is nicely mapped to an IHC-defined subtype (**Table 1**) except for the normal-like tumors which account for 7.8% of all breast cancer cases in a lymph-node negative cohort [15], shares a similar IHC status with the luminal A subtype and are characterized by a normal breast tissue profiling [16].

These five intrinsic subtypes have been repeated by several other studies with varying numbers of genes included in the signature. For instance, Hu et al. found a signature containing 306 genes that can distinguish these subtypes with significant differences observed on relapse-free and overall survival [17]. Parker et al. reported a 50-gene classifier (PAM50, which contains mostly hormone receptor and proliferation related genes, and genes exhibiting myoepithelial and basal features), which has significant prognostic and predictive values on

breast tumors [18-20] and can be widely applied in the clinical setting [21]. It is worth noting that although different in gene composition, the signatures identified by different studies should converge to the pathways they imply, based on which the same sample should not be classified into different categories. However, this is not the case due to, e.g., lack of stringent standardization of the methodology and breast cancer intrinsic subtype definition [22].

The development of tissue microarray (TMA) technology has enabled the validation of gene signatures at the translational level. A study exploring the combined protein expression profiles of a large panel of well-characterized commercially available biomarkers revealed 5 major groups [23]. In particular, they identified two large subtypes exhibiting luminal epithelial cell phenotypic characteristics, hormone receptors positivity, absence of basal epithelial phenotypic features and HER2 protein over-expression; two subgroups characterized by high HER2 positivity and negative hormone receptor expression, and differ from each other by MUC1 and E-cadherin expression; and one group characteristic of strong basal epithelial marker expression, TP53 positivity, absent hormone receptor expression and weak luminal epithelial and cytokeratin expression [23]. These subtypes accord well with those intrinsic subtypes based on GEP, which confirms the biological heterogeneity of breast cancer and demonstrates the clinical relevance of the intrinsic subtypes identified using high-throughput technologies [23].

Though Sørlie's subtyping has set the standard for intrinsic breast tumor categorization, other classifications also exist. For instance, Sotiriou et al. identified 6 groups among breast carcinomas using a signature containing 706 cDNA probe elements, which include 3 luminal-like, 1

| Study | Data type | Breast tumor subtype | | | | | |
| --- | --- | --- | --- | --- | --- | --- | --- |
| | | Luminal | | HER2 | TNP | | Normal-like |
| | | Luminal A | Luminal B | over-expression | Basal | Non-basal | |
| Sørlie [12] | GEP | 1 | 1 | 1 | 1 | | 1 |
| Hu [17] | GEP | 1 | 1 | 1 | 1 | | 1 |
| PAM50 [18] | GEP | 1 | 1 | 1 | 1 | | 1 |
| Abd El-Rehim [23] | TMA | 1 | 1 | 1 | 1 | | 1 |
| Sotiriou [24] | GEP | 3 | | 1 | 2 | | |
| Fan [25] | GEP | 1 | 1 | 1 | 1 | | |
| Lehmann [26] | GEP | | | | 2 | 4 | |
| Gatza [87] | PATH | 2 | 5 | 2 | 3 | | |
| | | 4 | | | | | |
| | | 1 | | | | | |
| METABRIC [88] | INT | 3 | 3 | 1 | 1 | | |
| | | 1 | | | | | |
| | | 1 | | | | | |
| TCGA [89] | INT | 1 | 1 | 1 | 1 | | |
| Dai [33] | INT | 1 | 1 | 1 | 1 | | |

Figure 1. Intrinsic breast tumor subtypes identified according to the example studies listed in the text. Gray blocks show subtypes identified from the study, white blocks are out of the scope of the corresponding study or 0 subtype is found. The value in each block shows the number of subtypes identified for each subtype. The corresponding blocks are unified for the identified group comprising of multiple tumor subtypes. GEP: gene expression profiling. Subtyping using gene expression data. TMA: tissue microarray. Subtyping using protein expression data. PATH: pathway. Subtyping based on pathway. INT: integrative view. Subyping using data integrated from multiple levels.

HER2-like and 2 basal-like subtypes [24]. Fan et al. have suggested a 70-gene signature to classify tumors into 4 groups where the normal-like subtype was not identified according to Sørlie's subtyping [25]. Lehmann et al. have subdivided triple negative tumors into 6 stable groups, i.e., 2 basal-like (BL1 and BL2), 1 immunomodulatory (IM), 1 mesenchymal (M), 1 mesenchymal stem-like (MSL) and 1 luminal androgen receptor (LAR) subtype using GEP. The BL1 subtype is heavily enriched in cell cycle and cell division components, suggesting the potential response to anti-mitotic agents such as taxanes (paclitaxel or docetaxel) of tumors belonging to this subtype. The BL2 subtype displays unique gene ontologies involving growth factor signaling and has features suggestive of basal/myoepithelial origin. The IM subtype is enriched for genes involved in immune cell processes. The M subtype displays a variety of unique gene ontologies that are heavily enriched in cell motility, ECM receptor interaction, and cell differentiation pathways. The MSL subtype shares genes for similar biological processes with the M subtype. Besides, this subtype contains genes involved in growth factor signaling, and displays low expression of claudin 3, 4, 7 (similar to the claudin-low subtype). The LAR subtype is ER negative but displays luminal gene expression patterns [26].

Despite these inconsistent naming and number of categories grouped by different studies (**Figure 1**), breast tumors fall primarily into three major classes, i.e., luminal, HER2 over-expression and triple negative phenotypic tumors (TNP), where triple negative tumors are the most heterogeneous and comprise largely of the basal subtype. The expression pattern, treatment response and clinical outcome of the major breast tumor subgroups, i.e., luminal, HER2 over-expression, basal, are discussed below [27]. All basic breast tumor intrinsic subtypes, their intrinsic nomenclature, featured IHC status, prevalence, as well as the association with clinical variables, i.e., tumor grade and patient outcome, are summarized in **Table 1**.

## Expression and clinical features of basic intrinsic subtypes

### Luminal tumors

The luminal-like tumors express hormone receptors, with expression profiles reminiscent of the luminal epithelial component of the breast [16]. These patterns include the expression of luminal cytokeratins 8/18, ER and genes associated with ER activation such as *LIV1* and *CCND1* [16, 24]. At least two subtypes exist within luminal-like tumors, i.e., luminal A and luminal B. Approximately speaking, the luminal A and luminal B tumors each represents the [ER+|PR+]HER2- (tumors with ER or PR positivity and HER2 negativity) and [ER+|PR+]HER2+ (tumors with ER or PR positivity and HER2 positivity) subtype, respectively, using the IHC nomenclature introduced in the previous section [28]. However, this equivalence does not always hold as, e.g., only part of luminal B tumors are HER2+ [10]. Luminal A tumors have higher expression of ER-related genes and lower expression of proliferative genes than luminal B cancers [12, 14]. Luminal B tumors tend to be of higher grade than luminal A tumors.

Luminal tumors are the most common subtypes among breast cancer, with luminal A being the majority. In the Carolina Breast Cancer Study [29], luminal breast tumors represent 64.3% of all patients, where luminal A cancers account for 54.3% (i.e., 57%, 67%, 40% and 55% of premenopausal white, postmenopausal white, premenopausal African American and postmenopausal African American women, respectively). In general, the luminal subtypes carry a good prognosis, and luminal B tumors have a significantly worse prognosis than the luminal A subtype [14]. Luminal tumors response well to hormone therapy but poorly to conventional chemotherapy [27]. Treatment response differs between luminal subtypes. According to the Recurrence Score, which is resulted from a RT-PCR based 16-gene predictor (half of them are ER and proliferation-related genes), tumors with low Recurrence Scores are luminal A while those with high Recurrence Scores are luminal B [30]. Thus, luminal A tumors could be adequately treated with endocrine therapy, while luminal B tumors which are more proliferative may benefit more from the combined therapeutic strategy of chemotherapy and hormonal treatment. Other targeted approaches such as anti-angiogenic strategies were suggested to be effective for luminal tumors as well. For example, the anti-VEGF antibody, bevacizumab, was shown to improve progression free survival in metastatic breast cancer when combined with paclitaxel, among which 60% of the patients carrying luminal tumors [27]. In 2012, the mTOR inhibitor everolimus (Afinitor) was approved in combination with exemestane for treating ER-positive, HER2-negative advanced breast cancer that recurs on standard therapies [31]. In addition, Palbociclib (under development by Pfizer), a cyclin-dependent kinase (CDK) 4/6 inhibitor, is approaching approval for treating such patients on the basis of data from a phase II study [32].

### HER2 over-expression tumors

The intrinsic HER2 over-expression tumors refer to those identified using gene expression array, which is similar to the ER-PR-HER2+ (ER negative, PR negative, HER2 positive) subgroup by immunostaining or fluorescence *in situ* hybridization (FISH) [28]. However, tumors classified by these two systems do not perfectly match, as not all clinically HER2-positive tumors show changes at the transcriptional level. The HER2 over-expression tumors are characterized by over-expressing other genes in the HER2 amplicon such as GRB7 [16, 33] and PGAP3 [33]. 40% to 80% of these tumors harbor TP53 mutation. HER2 over-expression tumors are more likely to be of grade 3.

No association with age or race was found for HER2 over-expression tumors [29], as well as known risk factors [27, 34]. Though HER2 over-expression breast tumors carry a poor prognosis [12, 14, 24], they are sensitive to anthracycline and taxane-based neoadjuvant chemotherapy, with significantly higher pathological complete response than luminal breast tumors [27]. The poor prognosis of this subtype as well as the basal tumors seem to derive from a higher risk of early relapse among those without complete eradication of tumor cells, and cancers of these two classes are suggested to derive the most benefit from improvements in chemotherapy [27]. Unlike the basal tumors, molecularly targeted agents such as the anti-HER2 monoclonal antibody, trastuzumab, are available for HER2 over-expression cancers. Not all HER2 over-expression tumors respond
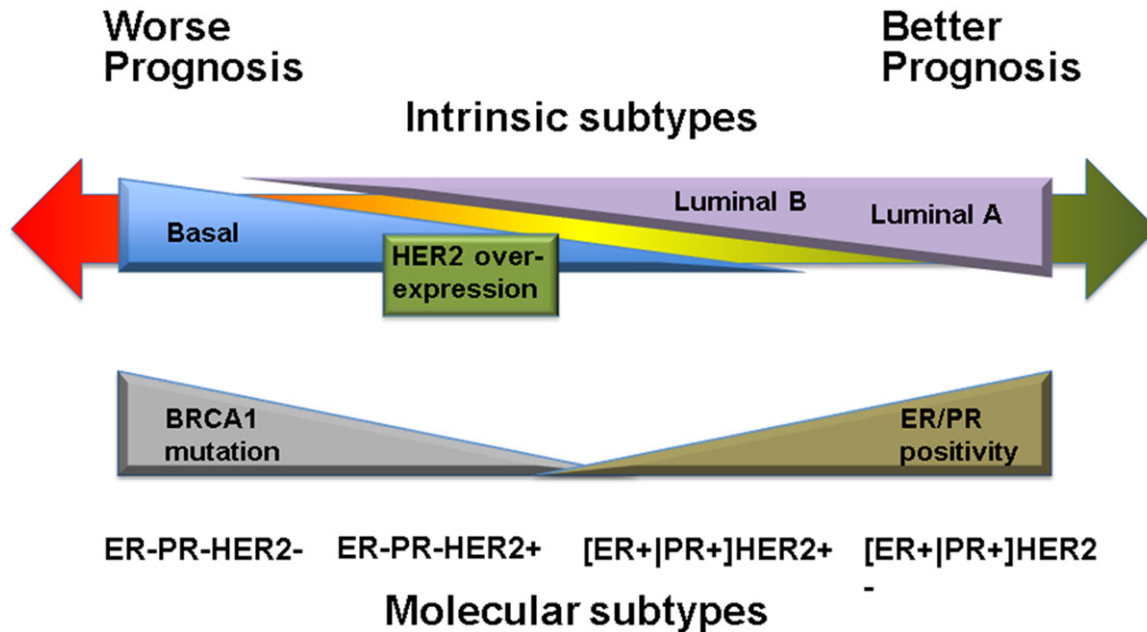
**Figure 2.** Patient outcome based on breast tumor intrinsic subtypes.

to trastuzumab. PTEN loss [35] and CXCR4 up-regulation [27] have been implicated in trastuzumab resistance which provide targets in the combined strategies to improve clinical outcome in the future.

*Basal tumors*

As discussed before, the basal subtype is composed of ER-PR-HER2- (triple negative) tumors with expression profiles mimicking that of the basal epithelial cells of other parts of the body and normal breast myoepithelial cells [16]. Such expression patterns include lacking or low expression of hormone receptors and HER2, and high expression of basal markers (such as keratins 5, 6, 14, 17, EGFR) and proliferation related genes [16, 24]. Tumors characterized by basal cytokeratin expression are more probable to have low BRCA1 expression [23] and harbor TP53 mutation [12, 29]. Similar with HER2 over-expression tumors, basal cancers are likely to be of grade 3 tumors [12, 29].

Basal tumors account for 60% to 90% triple negative cases [25, 36]. These tumors are of particular interest because they follow aggressive clinical course and currently lack any form of standard targeted systemic therapy. Compared with the other subtypes, these tumors are associated with younger patient age, more common to develop in African-American women and especially among pre-menopausal individ-

uals [37]. Risk factors for this subtype include earlier menarche, high waist-to-hip ratio, and a lack of breast-feeding together with high parity [38]. Unlike the luminal A subtype, where having multiple children and a younger age at the first full-term pregnancy are protective, these factors increase the hazard for basal tumors [39]. These tumors are associated with a lower disease-specific survival and a higher risk of local and regional relapse. Follow-up studies have revealed a time-dependent survival profile for basal breast tumors, with a very poor early outlook diminishing after around 5 years. The metastasis pattern also separates basal tumors from the other breast cancers, with a tendency towards visceral organs (excluding bone) and less likely to involve lymph nodes [39]. The size of basal tumors is, in general, larger than the other subtypes, with a median size of 2 cm in one series [40]. Also, tumors of this class tend to show rapid growth [39]. Given the triple negative receptor status, basal tumors are not amenable to conventional targeted breast cancer therapies, leaving chemotherapy the only option in the therapeutic armamentarium. Two independent studies examining the chemo-resistance of basal cancers have shed light on patients experiencing tumors of this subtype. Their studies converge to the view that these aggressive tumors are sensitive to conventional chemotherapies such as anthracycline and taxane, and their poor prognosis is

**Table 2.** Commercially available prognostic multi-gene signatures for breast cancer patients [6, 54]

| | MammaPrint [43, 44] | Veridex 76-gene [44] | MapQuant Dx [45] | MapQuant Dx simplified [46] | Oncotype DX [30, 47] | Theros [48, 49] |
|---|---|---|---|---|---|---|
| Technique | DNA microarray | DNA microarray | DNA microarray | qRT-PCR | qRT-PCR | qRT-PCR |
| Provider | Agendia | Currently not commercially available | Ipsogen | Ipsogen | Genomic Health | bioTheranostics |
| Assay type | 70-gene signiture | 76-gene signiture | 97-gene signiture | 8-gene signature | 21-gene recurrence score | 2-gene ratio of HOXB13 to IL17R (H/I)/molecular grade index |
| Tissue type | Fresh or frozen | Fresh or Frozen | Fresh or Frozen | FFPE | FFPE | FFPE |
| Discovery set | 78 ER±, N0, < 5 cm diameter cancers, age < 55 years | 115 ER±, N0 cancers | 64 ER+ cancers | 64 ER+ cancers | 447 ER+ samples, including samples from the tamoxifen only group of the NSABP B-20 trial | 60 ER+ tumors, tamoxifen only treated patients 20 microdissected FFPE samples |
| Initial validation set | 295 ER±, N±, < 5 cm diameter cancer, age < 52 years | 171 ER±, N0 cancers | 597 ER± cancers, of which 125 profiled in-house | 597 ER± cancers, of which 125 profiled in-house | 668 ER+ samples from NSABP B-14trial[90] (tamoxifen-treated) | 20 ER+ FFPE samples |
| Outcome | Distant metastasis at 5 years | Distant metastasis at 5 years | Good (GG II) or poor (GG I III) prognosis | Good (GG II) or poor (GG I III) prognosis | Disease-free relapse at 10 years | Relapse-free and overall survival |
| Clinical application | To aid in prognostic prediction in patients < 61 year of age with stage I or II, N0 disease with a tumor size of ≤ 5 cm | To prognose N0 patients | To restratify grade 2 tumors into low-risk grade 1 or high-risk grade 3 tumors, specifically for invasive, primary, ER+ grade 2 tumors | To restratify grade 2 tumors into low-risk grade 1 or high-risk grade 3 tumors, specifically for invasive, primary, ER+ grade 2 tumors | To predict the risk of recurrence in patients with ER+, N0 disease treated with tamoxifen; to identify patients with a low risk of recurrence who may not need adjuvant chemotherapy | To stratify ER+ patients into groups with a predicted low-risk or high-risk of recurrence and a predicted good or poor response to endocrine therapy |
| Resulets presentation | Dichotomous; good or poor prognosis | Dichotomous; good or poor prognosis | Dichotomous, GGII or GG I III | Dichotomous, GGII or GG I III | Continuous variable; recurrence score | Continuous variable; risk of recurrence score |
| Additional information provided | mRNA levels of ER, PR, and HER2 (Targetprint); Intrinsic subtypes (Blueprint) | - | - | - | mRNA levels of ER, PR, and HER2 | Molecular grade index |
| Prognostic value in other populations | Up to 3 positive nodes, and HER2+ disease | ER+, N0 patients treated with tamoxifen | ER+, receiving aromatase inhibitors | ER+, receiving aromatase inhibitors | ER+ and 1-3 N+, ER+ postmenopausal receiving aromatase inhibitors | - |
| Predictive value | Chemotherapy response (poor prognosis group) | Chemotherapy response (poor prognosis group) | Chemotherapy response (GG I III) | Chemotherapy response (GG I III) | Chemotherapy response (high recurrence score) | Chemotherapy response (high risk of recurrence score)* |
| Level of evidence | II | III | III | III | I | III |
| FDA approval | Yes | No | No | No | No | No |
| Availability | Europe and USA | | Europe | Europe | Europe and USA | USA |

FDA: US Food and Drug Administration; MGI: molecular grade index; GGI: genomic grade index; FFPE: formalin-fixed paraffin-embedded; ER: oestrogen receptor; PR: progesterone receptor; N: lymph nodes; ±: positive and negative; +: positive; *: based on indirect evidence.

not driven by the initial chemoresistance but rather due to the relatively few treatment options available for triple negative tumors [27]. Besides these conventional therapeutic strategies, many studies have kept suggesting novel targets for basal tumors. It has been suggested that basal tumors may be EGFR-driven [41]. The 'wound response' signature, encompassing genes involved in matrix remodelling and angiogenesis, has been shown to be associated with basal tumors, suggesting other potential avenues of targeting [42].

Patient outcome of these basic intrinsic subtypes are compared in **Figure 2**.

**Commercial multi-gene signatures for breast cancer prognosis**

Commercially, six genomic assays, i.e., MammaPrint [43, 44], Veridex 76-gene signature [44], MapQuant Dx [45] and its simplified version [46], Oncotype DX [30, 47] and Theros [48, 49] were developed for the prediction of clinical outcome among breast cancer patients, which are summarized in **Table 2**.

MammaPrint (Agendia, Amsterdam, Netherlands), a feature set containing 70 genes, is the first successfully developed prognostic signature [43, 44]. It has been approved by the US Food and Drug Administration (FDA), and used for the prognostication of patient with stage 1 or 2, node negative, invasive breast tumors of size less than 5 cm. Level II evidence suggests that in cases where the clinicopathological measures disagree with MammaPrint, the latter predicts outcome with higher accuracy [50, 51]. Following MammaPrint, Veridex 76-gene signature was developed [44]. By contrast with MammaPrint, this signature was identified by conducting the analysis, separately, within ER-positive and ER-negative cancers, leading to 60 genes diagnostic of distant metastasis within 5 years in ER-positive patients and 16 genes prognostic of distant metastasis among ER-negative patients [44]. A test of this signature using 171 node negative patients revealed that this 76-gene signature was a strong prognostic factor for 5 years distant metastasis, which outperforms the St Gallen's and NCI guidelines in identifying patients with good prognosis and could forgo chemotherapy [44]. MapQuant Dx (Ipsogen SA, Marseille, France) is a hypothesis-driven prognostic signature based on the premise that histological grade is a strong prognostic factor in ER positive tumors [30]. This signature could stratify grade II cancers into grade I-like (with a low frequency of distant relapses) and grade III-like (having a clinical behavior similar to that of grade III) cancers [45, 52]. Akin to MammaPrint, MapQuant Dx correlates with the benefit from chemotherapy. However, the prognostic value provided by MapQuant Dx is only applicable to ER positive tumors [53]. MammaPrint, the Veridex 76-gene signature, and MapQuant Dx are all DNA microarrays based and require fresh or frozen samples for the assays. This poses challenges for the prospective validation, limiting their prognostic power supported by the evidence level. The prognostic information of these three assays stem almost exclusively from the expression of the proliferation-related genes and time dependent, setting further limitations for their application [54].

In parallel with microarray-based prognostic signatures, the technique of qRT-PCR has also been used for developing the prognostic assay. These methods extract RNA from formalin-fixed paraffin-embedded (FFPE) tissue samples, thus do not have difficulties in specimen procurement. The simplified version of MapQuant Dx [46], Oncotype DX [30, 47], and Theros [48, 49] belong to this category. Simplified MapQuant Dx contains 8 genes and has the same predictive value as the microarray-based version [46]. Oncotype DX measures the expression of 21 genes, including 16 cancer-related and 5 reference genes [30, 47]. The recurrence score is used to prognose the risk of distant relapse at 10 years for ER positive, node negative breast cancer patients, and an independent predictive factor for these patients receiving adjuvant tamoxifen [30]. Also, evidences show that the recurrence score has predictive value among ER positive patients treated with aromatase inhibitors [55], and is prognostic of patients with ER positive tumors harboring up to three positive lymph nodes [56]. The prognostic use of Oncotype DX is supported by level I evidence, and this test has been incorporated in the National Comprehensive Cancer Network as a predictor of recurrence and a guide when making therapeutic decisions among early ER positive node negative breast tumors [54]. Oncotype DX assay has also been included in American Society of Clinical Oncology guidelines as a tumor marker of recurrence [54]. Theros measures the ratio between two genes

HOXB13 and IL17R (H/I ratio) to provide the prognostic value on the relapse-free and overall survival among ER positive patients [48, 49]. It is also predictive of the treatment response to endocrine among these patients, with a high H/I ratio being associated with a high recurrence risk [48, 49]. The accuracy of the H/I assay was improved by including the molecular-grade index which contains 5 genes and mainly measures cell proliferation [48].

Different gene signatures share very few genes in common due to the complexity of gene expression data containing large numbers of highly correlated variables [6]. However, studies comparing the performance of different signatures have revealed a concordant risk assignment despite the few common genes shared [25], e.g., only one gene (SCUBE2) is in common between MammaPrint and Oncotype DX.

### Emerging information for intrinsic subtyping

*MiRNA expression profiling*

MicroRNA (miRNA) is a category of small (approximately 22-nucleotide) non-coding RNAs with regulatory activities, which functions mainly by inhibiting protein synthesis via binding the complementary sequences on target mRNA genes [57]. It is estimated that miRNAs could regulate the expression of 30% to 60% of all human protein-coding genes [58, 59]. Also, miRNA expression profiles are unique for a wide range of human diseases including different stages of tumor progression and metastasis [60], and circulating miRNAs are extremely stable in blood and serum [61, 62]. These features, altogether, empower miRNAs a new class of promising diagnostic and prognostic markers in the clinic. The first comprehensive report on miRNA signatures for breast tumors was published in 2005 [63], where a number of miRNAs were shown deregulated in such tumors. MiR-125b, miR-145, miR-21 and miR-155 were reported to be the most dysregulated miRNAs in their study, and some of these malfunctional miRNAs were correlated with specific clinical features of breast tumors including the expression of hormone receptors, tumor stage, vascular invasion and proliferation [63, 64]. They later reported that miR-205 is down-modulated in breast carcinomas as compared with normal breast tissues [65]. Another study revealed that miR-221 and miR-222 are involved in a regulatory loop with ER [66]. The

expression of miR-145 was reported severely decreased in tumor specimens [67]. MiRNA145 was shown to collaborate with TP53 in a death-promoting regulatory loop and target ER in breast tumor cells, suggesting a miR-145 re-expression therapy for patients with ER+ and/or TP53 wild-type tumors [67]. MiR-9 was shown to participate in a breast tumor metastasis-promoting network involving E-cadherin and β–catenin [68].

MiRNA profiling brings an improvement for breast cancer subtyping and prediction of treatment response. Our recent study on breast tumor subtyping integrating mRNA and miRNA profilings revealed 69 miRNAs differentially expressed among subgroups classified using ER, PR and HER2 status, with the majority comes from the triple negative group and especially the basal subtype [33]. Among the miRNAs differentiating breast tumor subtypes in this signature, miR-135a, miR-135b, miR-365 and miR-7 were found to distinguish breast tumors by ER status [33]. Blenkiron et al. studied the miRNA expression profiling among breast malignancies classified using intrinsic subtypes, and found miR-155 differentially expressed in ER+ versus ER- tumors [57]. A miRNA panel consisting of largely under-expressed miRNAs was identified to characterize male from female breast carcinomas [69, 70].

Taken together, miRNAs are potential excellent biomarkers of breast carcinomas, and could be employed for innovative therapies for targeted patients. MiRNA expression profiling could avail as a critical means for breast cancer subtyping as well as the associated prognosis and therapeutic prediction.

*lncRNA*

Long non-coding RNA (lncRNA) is originally defined as RNA molecules longer than 200 nucleotides that do not encode a protein. Our understandings toward the roles and functions of lncRNAs are rapidly advancing nowadays [71-73]. A recent review distilled the myriad functions of lncRNA into 4 mutually unexclusive archetypes, i.e., signals; decoys (acting as molecular sponges pulling away RNA binding proteins that play regulatory roles such as transcription factors and chromatin modifiers); guides (recruiting chromatin-modifying enzymes to target genes); and ropes (keep multiple proteins together to form ribonucleoprotein

complexes) [73, 74]. Its disease subtyping role has also been explored by several studies. A recent study has reported that lncRNAs, but not mRNAs or miRNAs, could discriminate failing hearts of different pathologies, i.e., identifying the ischemic and nonischemic failing cardiomyopathy from nonfailing cases [75]. Using lncRNA expression profiles, glioma has been classified into three mole cular subtypes [76]. The prognostic and predictive roles of lncRNAs have been suggested in several cancers including breast, prostate, bladder and kidney tumors [77, 78]. In breast cancer, several lncRNAs including, e.g., HOTAIR, MALAT1, GAS5, BC200, SRA-1 and LSINCT5, have been reported differentially expressed in tumor cells [77] and the lncRNA such as Zfas1 has been reported as a potential marker of such carcinomas [79]. Given the prominent roles played by lncRNA in breast cancer and its success in disease subtyping, it is promising to better decipher the high heterogeneity of breast cancer with the avail of lncRNA.

*Epigenetics*

Epigenetics mainly refers to the modification of DNA and histone as well as their roles in regulating the transcriptional program. Similar to aberrant gene expression, epigenetic alterations contribute to the pathogenesis and molecular heterogeneity of cancers. DNA methylation signatures were reported to segregate patients with CEBPA aberrations from the other subtypes of leukemia and define four epigenetically distinct forms of acute myeloid leukemia (AML) with NPM1 mutations [80]. A 15-gene methylation classifier was reported to be predictive of overall survival in AML [80]. Abnormal methylation of CpG island in gene promoters has been identified as a common mechanism for suppressing gene expression in cancer cells [81]. CpG Island Methylator Phenotype (CIMP) has been reported as novel subtype in colorectal cancers [82]. Though the hypermethylation pattern was reported less distinctive across subtypes in breast cancer than among tumors of different tissue origins [83], the importance of hypermethylation in breast cancer classification has been recognized lately. A recent study revealed that CIMP in breast cancer is associated with the lobular subtype [84]. The histone chaperone HJURP (Holliday Junction Recognition Protein), an epigenetic regulator, was identified as a new independent prognostic marker

for luminal A breast carcinoma [85]. Epigenetics is an indispensible and promising tool in assessing breast cancer heterogeneity, with its power being unfolded.

*Pathways*

Breast cancers are comprised of molecularly distinct subtypes that respond differently to pathway-targeted therapies. Transcription factors and signaling proteins associated with transcriptional programs were found specific to intrinsic subtypes [86], suggesting the link between signaling and subtyping. Thus, data from microarrays have also been used for pathway analysis, with the aim of identifying subtypes sharing common disease-causing functions. Gatza et al. have presented a pathway-based method to classify breast cancer subtypes based on oncogenic and tumor suppressor pathway deregulation [87]. Seventeen groups were identified in [87] where the widely accepted intrinsic subtypes were mixed in each pathway-defined subgroup. Particularly, subgroups 11 and 17 are luminal A tumors; subgroups 3, 4, 6, 9 and 16 belong to the luminal B subtype; subgroups 7 and 10 contain HER2 positive tumors, and subgroups 2, 5, 8 are basal tumors. Subgroups 1, 12, 13, 15 are comprised of luminal tumors, and the subgroup 14 is a mixture of all intrinsic subtypes [87]. These pathway-based subtypes were shown to differentiate tumors exhibiting similar clinical and biological properties, including distinct patterns of chromosomal alterations that were not evident using, e.g., intrinsic subtyping [87].

## Trends in breast tumor instrinsic subtyping

With the advent of high-throughput technology, overwhelming information on various levels, such as genomic, transcriptional, translational and epigenetic, has become available for cancer research. An increasing effort has been devoted to integrating information at multiple levels, with the aim of understanding the core functional differences driving breast tumor heterogeneity and seeking the effective treatment. Two emerging trends exist in this area, i.e., expanding the subtypes with refined features, and converging the subtypes identified using various methods.

The METABRIC (Molecular Taxonomy of Breast Cancer International Consortium) paved the way in the first direction. This study revealed a

refined breast cancer molecular taxonomy, i.e., 10 integrative clusters which are named IntClust 1 to 10, by integrating copy number and GEP of 2000 breast tumors [88]. Among these subtypes, IntClusts 3, 7, 8 are primarily composed of luminal A tumors, where IntClust 3 is marked by a paucity of copy number and cis-acting alterations, IntClust7 lacks the 1q alteration, harbors the 16p gain/16q loss and has a higher frequency of 8q amplification, and IntClust 8 is characterized by 1q gain/16q loss; IntClusts 1, 6, 9 are enriched for luminal B cancers, which are characterized by 17q23/20q cis-acting aberrations, 8p12 cis-acting aberrations, and 8q cis-acting aberrations/20q amplifications, respectively; IntClust 5 consists of HER2-amplified cancers regardless of ER status; IntClust 10 contains the majority of basal tumors which is the most instable at the genomic level; IntClust 2 is enriched in luminal cancers but exhibits a high mortality rate, bearing cis-acting aberrations in the 11q13/14 region (where several driver genes reside); IntClust 4 is composed of both ER+ and ER- tumors and varied intrinsic subtypes, and shares similar genomic features with IntClust 3, i.e., lacking copy number and cis-acting alterations [88].

The Cancer Genome Atlas Network (TCGA) pioneers in the second direction. It investigated breast cancer subtypes by incorporating information from multiple platforms, i.e., genomic DNA copy number arrays, DNA methylation, exome sequencing, mRNA arrays, miRNA sequencing and reverse-phase protein arrays. By classifying tumors using each individual platform and comparing results at different levels, they conclude that diverse genetic and epigenetic alterations converge phenotypically into four major breast tumor subgroups (*i.e.,* luminal A, luminal B, HER2 positive, triple negative) that are previously identified using mRNA profiling [89]. Our previous endeavor on unveiling breast tumor heterogeneity using mRNA and miRNA expression profiling has also demonstrated the power of an integrative view on breast tumor subtyping and contributed in this domain [33]. A feature set containing 1015 mRNA and 69 miRNAs was found differentially expressed among breast tumor subtypes defined by ER, PR and HER2. It could well characterize breast tumors into [ER+|PR+]HER2-, [ER+|PR+]HER2+, ER-PR-HER2+, ER-PR-HER2-, as validated us-

ing multiple independent datasets, converging breast tumor subtypes obtained using different methodologies [33]. We further reduced this gene panel to 119 mRNAs using hierarchical clustering and nearest-to-center principle for the convenience of clinic use (results to be published).

## Discussion

Among the four intrinsic subtypes, i.e., luminal A, luminal B, HER2 over-expression, which are commonly referred to, basal tumors are of particular interest due to the aggressive clinical course they follow and the lack of standard targeted systemic therapy. Normal-like and luminal A tumors share the same status on the basic IHC markers, i.e. [ER+|PR+]HER2-KI67- (ER or PR positive, HER2 negative, KI67 negative) but differ on expression pattern, with the normal-like tumors resembling the normal breast profiling and having poor outcome.

Information such as lncRNA and epigenetic data plays critical roles in tumor progression and classification, providing novel perspectives on breast tumor subtyping. With the available information accumulating, taking an integrative view on breast tumor subtyping has been gaining increasing interest on deciphering the heterogeneity of such tumors. Current studies incorporating multiple types of data, though few, cast a relatively comprehensive view and reflect two trends on breast tumor subtyping, i.e., expansion of and convergence to the current four major subtypes.

Despite the growing number of clinically relevant molecular subtypes being identified, current breast cancer patient management still depends on traditional pathology assessment supplemented with biomarker testing using validated commercial assays (i.e., MammaPrint, MapQuant Dx and its simplified version, Oncotype DX, Theros). The clinical relevance of molecular subtypes identified using either molecular markers or signature patterns, i.e., guiding in individualized therapy, is evident. However, it is important to standardize the methodology used for molecular subtyping, whose reproducibility needs to be extensively tested. Further effort is called for to make these theoretical advances technically convenient and available for clinical use.

## Concluding remarks

Breast cancer is comprised of different entities, each being associated with distinct outcome and therapeutic approaches. Intrinsic subtyping, born with the advent of high-throughput technologies, has gained its favour in recent years assuming that tumors sharing similar expression profiling follow the same pathologic pathway and should be given the same treatment. Armed with this concept and methodology, availed by various emerging information and novel perspectives, and equipped with our incremental knowledge on the carcinogenesis of such a complex disease, we are ready to decipher breast tumor heterogeneity and make it beneficial to clinical patients.

## Acknowledgements

## Disclosure of conflict of interest

None.

Address correspondence to: Dr. Xiaofeng Dai, National Engineering Laboratory for Cereal Fermentation Technology, School of Biotechnology, No. 1800, Lihu Avenue, Wuxi 214122, Jiangsu, China. Tel: 86-186-1147-9958; Fax: 86-0510-8532-9306; E-mail: Xiaofeng.dai@me.com

## References

[1] Spitale A, Mazzola P, Soldini D, Mazzucchelli L and Bordoni A. Breast cancer classification according to immunohistochemical markers: clinicopathologic features and short-term survival analysis in a population-based study from the South of Switzerland. Ann Oncol 2009; 20: 628-635.

[2] Tang P, Wang J and Bourne P. Molecular classifications of breast carcinoma with similar terminology and different definitions: are they the same? Hum Pathol 2008; 39: 506-513.

[3] Desmedt C, Sotiriou C and Piccart-Gebhart MJ. Development and validation of gene expression profile signatures in early-stage breast cancer. Cancer Invest 2009; 27: 1-10.

[4] Iwamoto T and Pusztai L. Predicting prognosis of breast cancer with gene signatures: are we lost in a sea of data? Genome Med 2010; 2: 81.

[5] Reis-Filho JS, Weigelt B, Fumagalli D and Sotiriou C. Molecular profiling: moving away from tumor philately. Sci Transl Med 2010; 2: 47ps43.

[6] Sotiriou C and Pusztai L. Gene-expression signatures in breast cancer. N Engl J Med 2009; 360: 790-800.

[7] Weigelt B, Baehner FL and 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] Blows FM, Driver KE, Schmidt MK, Broeks A, van Leeuwen FE, Wesseling J, Cheang MC, Gelmon K, Nielsen TO, Blomqvist C, Heikkila P, Heikkinen T, Nevanlinna H, Akslen LA, Begin LR, Foulkes WD, Couch FJ, Wang X, Cafourek V, Olson JE, Baglietto L, Giles GG, Severi G, McLean CA, Southey MC, Rakha E, Green AR, Ellis IO, Sherman ME, Lissowska J, Anderson WF, Cox A, Cross SS, Reed MW, Provenzano E, Dawson SJ, Dunning AM, Humphreys M, Easton DF, Garcia-Closas M, Caldas C, Pharoah PD and Huntsman D. Subtyping of breast cancer by immunohistochemistry to investigate a relationship between subtype and short and long term survival: a collaborative analysis of data for 10,159 cases from 12 studies. PLoS Med 2010; 7: e1000279.

[9] Vallejos CS, Gomez HL, Cruz WR, Pinto JA, Dyer RR, Velarde R, Suazo JF, Neciosup SP, Leon M, de la Cruz MA and Vigil CE. Breast Cancer Classification According to Immunohistochemistry Markers: Subtypes and Association With Clinicopathologic Variables in a Peruvian Hospital Database. Clinical Breast Cancer 2010; 10: 294-300.

[10] Cheang MC, Chia SK, Voduc D, Gao D, Leung S, Snider J, Watson M, Davies S, Bernard PS, Parker JS, Perou CM, Ellis MJ, Nielsen TO. Ki67 index, HER2 status, and prognosis of patients with luminal B breast cancer. J Natl Cancer Inst 2009; 101: 736-50.

[11] Pusztai L, Broglio K, Andre F, Symmans WF, Hess KR and Hortobagyi GN. Effect of molecular disease subsets on disease-free survival in randomized adjuvant chemotherapy trials for estrogen receptor-positive breast cancer. J Clin Oncol 2008; 26: 4679-4683.

[12] Sørlie T, Perou CM, Tibshirani R, Aas T, Geisler S, Johnsen H, Hastie T, Eisen MB, van de Rijn M, Jeffrey SS, Thorsen T, Quist H, Matese JC, Brown PO, Botstein D, Lonning PE and Borresen-Dale AL. Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. Proc Natl Acad Sci U S A 2001; 98: 10869-10874.

[13] Perou CM, Sørlie T, Eisen MB, van de Rijn M, Jeffrey SS, Rees CA, Pollack JR, Ross DT, John-

sen H, Akslen LA, Fluge O, Pergamenschikov A, Williams C, Zhu SX, E LP, Børresen-Dale AL, Brown PO and Botstein D. Molecular portraits of human breast tumors. Nature 2000; 406: 747-752.

[14] Sørlie T, Tibshirani R, Parker J, Hastie T, Marron JS, Nobel A, Deng S, Johnsen H, Pesich R, Geisler S, Demeter J, Perou CM, Lønning PE, Brown PO, Børresen-Dale A and Botstein D. Repeated observation of breast tumor subtypes in independent gene expression data sets. PNAS 2003; 100: 8418-8423.

[15] Smid M, Wang Y, Zhang Y, Sieuwerts AM, Yu J, Klijn JG, Foekens JA and Martens JW. Subtypes of breast cancer show preferential site of relapse. Cancer Res 2008; 68: 3108-3114.

[16] Perou CM, Sorlie T, Eisen MB, van de Rijn M, Jeffrey SS, Rees CA, Pollack JR, Ross DT, Johnsen H, Akslen LA, Fluge O, Pergamenschikov A, Williams C, Zhu SX, Lonning PE, Borresen-Dale AL, Brown PO and Botstein D. Molecular portraits of human breast tumours. Nature 2000; 406: 747-752.

[17] Hu Z, Fan C, Oh DS, Marron JS, He X, Qaqish BF, Livasy C, Carey LA, Reynolds E, Dressler L, Nobel A, Parker J, Ewend MG, Sawyer LR, Wu J, Liu Y, Nanda R, Tretiakova M, Ruiz Orrico A, Dreher D, Palazzo JP, Perreard L, Nelson E, Mone M, Hansen H, Mullins M, Quackenbush JF, Ellis MJ, Olopade OI, Bernard PS and Perou CM. The molecular portraits of breast tumors are conserved across microarray platforms. BMC Genomics 2006; 7: 96.

[18] Gnant M, Filipits M, Greil R, Stoeger H, Rudas M, Bago-Horvath Z, Mlineritsch B, Kwasny W, Knauer M, Singer C, Jakesz R, Dubsky P, Fitzal F, Bartsch R, Steger G, Balic M, Ressler S, Cowens JW, Storhoff J, Ferree S, Schaper C, Liu S, Fesl C, Nielsen TO, Austrian B and Colorectal Cancer Study G. 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.

[19] Dowsett M, Sestak I, Lopez-Knowles E, Sidhu K, Dunbier AK, Cowens JW, Ferree S, Storhoff J, Schaper C and Cuzick J. Comparison of PAM50 risk of recurrence score with oncotype DX and IHC4 for predicting risk of distant recurrence after endocrine therapy. J Clin Oncol 2013; 31: 2783-2790.

[20] Parker JS, Mullins M, Cheang MC, Leung S, Voduc D, Vickery T, Davies S, Fauron C, He X, Hu Z, Quackenbush JF, Stijleman IJ, Palazzo J, Marron JS, Nobel AB, Mardis E, Nielsen TO, Ellis MJ, Perou CM and Bernard PS. Supervised risk predictor of breast cancer based on intrinsic subtypes. J Clin Oncol 2009; 27: 1160-1167.

[21] Ades F, Zardavas D, Bozovic-Spasojevic I, Pugliano L, Fumagalli D, de Azambuja E, Viale G, Sotiriou C and Piccart M. Luminal B breast cancer: molecular characterization, clinical management, and future perspectives. J Clin Oncol 2014; 32: 2794-2803.

[22] Weigelt B, Mackay A, A'hern R, Natrajan R, Tan DSP, Dowsett M, Ashworth A and Reis-Filho JS. Breast cancer molecular profiling with single sample predictors: a retrospective analysis. Lancet Oncol 2010; 11: 339-349.

[23] Abd El-Rehim DM, Ball G, Pinder SE, Rakha E, Paish C, Robertson JF, Macmillan D, Blamey RW and Ellis IO. High-throughput protein expression analysis using tissue microarray technology of a large well-characterised series identifies biologically distinct classes of breast cancer confirming recent cDNA expression analyses. Int J Cancer 2005; 116: 340-350.

[24] Sotiriou C, Neo SY, McShane LM, Korn EL, Long PM, Jazaeri A, Martiat P, Fox SB, Harris AL and Liu ET. Breast cancer classification and prognosis based on gene expression profiles from a population-based study. Proc Natl Acad Sci U S A 2003; 100: 10393-10398.

[25] Fan C, Oh DS, Wessels L, Weigelt B, Nuyten DS, Nobel AB, van't Veer LJ and Perou CM. Concordance among gene-expression-based predictors for breast cancer. N Engl J Med 2006; 355: 560-569.

[26] Lehmann BD, Bauer JA, Chen X, Sanders ME, Chakravarthy AB, Shyr Y and Pietenpol JA. Identification of human triple-negative breast cancer subtypes and preclinical models for selection of targeted therapies. J Clin Invest 2011; 121: 2750-2767.

[27] Brenton JD, Carey LA, Ahmed AA and Caldas C. Molecular classification and molecular forecasting of breast cancer: ready for clinical application? J Clin Oncol 2005; 23: 7350-7360.

[28] Vallejos CS, Gómez HL, Gruz WR, Pinto JA, Dyer RR, Velarde R, Suazo JF, Neciosup SP, M L, de la Cruz MA and Vigil CE. Breast Cancer Classification According to Immunohistochemistry Markers: Subtypes and Association With Clinicopathologic Variables in a Peruvian Hospital Database. Clinical Breast Cancer 2010; 10: 294-300.

[29] O'Brien KM, Cole SR, Tse CK, Perou CM, Carey LA, Foulkes WD, Dressler LG, Geradts J and Millikan RC. Intrinsic breast tumor subtypes, race, and long-term survival in the Carolina Breast Cancer Study. Clin Cancer Res 2010; 16: 6100-6110.

[30] Paik S, Shak S, Tang G, Kim C, Baker J, Cronin M, Baehner FL, Walker MG, Watson D, Park T, Hiller W, Fisher ER, Wickerham DL, Bryant J

and Wolmark N. A multigene assay to predict recurrence of tamoxifen-treated, node-negative breast cancer. N Engl J Med 2004; 351: 2817-2826.

[31] Abbas S, Linseisen J, Slanger T, Kropp S, Mutschelknauss EJ, Flesch-Janys D and Chang-Claude J. Serum 25-hydroxyvitamin D and risk of post-menopausal breast cancer-results of a large case-control study. Carcinogenesis 2008; 29: 93-99.

[32] Morikawa A and Henry NL. Palbociclib for the Treatment of Estrogen Receptor-Positive, HER2-Negative Metastatic Breast Cancer. Clin Cancer Res 2015;

[33] Dai X, Chen A and Bai Z. Integrative investigation on breast cancer in ER, PR and HER2-defined subgroups using mRNA and miRNA expression profiling. Sci Rep 2014; 4: 6566.

[34] Colditz GA, Rosner BA, Chen WY, Holmes MD and Hankinson SE. Risk factors for breast cancer according to estrogen and progesterone receptor status. J Natl Cancer Inst 2004; 96: 218-228.

[35] Nagata Y, Lan KH, Zhou X, Tan M, Esteva FJ, Sahin AA, Klos KS, Li P, Monia BP, Nguyen NT, Hortobagyi GN, Hung MC and Yu D. PTEN activation contributes to tumor inhibition by trastuzumab, and loss of PTEN predicts trastuzumab resistance in patients. Cancer Cell 2004; 6: 117-127.

[36] Swenson RR, Rizzo CJ, Brown LK, Payne N, DiClemente RJ, Salazar LF, Vanable PA, Carey MP, Valois RF, Romer D and Hennessy M. Prevalence and correlates of HIV testing among sexually active African American adolescents in 4 US cities. Sex Transm Dis 2009; 36: 584-591.

[37] Carey LA, Perou CM, Livasy CA, Dressler LG, Cowan D, Conway K, Karaca G, Troester MA, Tse CK, Edmiston S, Deming SL, Geradts J, Cheang MC, Nielsen TO, Moorman PG, Earp HS and Millikan RC. Race, breast cancer subtypes, and survival in the Carolina Breast Cancer Study. JAMA 2006; 295: 2492-2502.

[38] Ho-Yen C, Bowen RL and Jones JL. Characterization of basal-like breast cancer: an update. Mini-Symposium: the Biological Phenotype of Breast Cancer 2012; 18: 104-111.

[39] Ho-Yen C, Bowen RL and Jones J. Characterization of basal-like breast cancer: an update. Diagnostic Histopathology 2012; 18: 104-111.

[40] Rakha EA, Putti TC, Abd El-Rehim DM, Paish C, Green AR, Powe DG, Lee AH, Robertson JF and Ellis IO. Morphological and immunophenotypic analysis of breast carcinomas with basal and myoepithelial differentiation. J Pathol 2006; 208: 495-506.

[41] Nielsen TO, Hsu FD, Jensen K, Cheang M, Karaca G, Hu Z, Hernandez-Boussard T, Livasy C, Cowan D, Dressler L, Akslen LA, Ragaz J, Gown AM, Gilks CB, van de Rijn M and Perou CM. Immunohistochemical and clinical characterization of the basal-like subtype of invasive breast carcinoma. Clin Cancer Res 2004; 10: 5367-5374.

[42] Chang HY, Nuyten DS, Sneddon JB, Hastie T, Tibshirani R, Sørlie T, Dai H, He YD, van't Veer LJ, Bartelink H, van de Rijn M, Brown PO and van de Vijver MJ. Robustness, scalability, and integration of a wound-response gene expression signature in predicting breast cancer survival. Proc Natl Acad Sci U S A 2005; 102: 3738-3743.

[43] van't Veer LJ, Dai H, van de Vijver MJ, He YD, Hart AA, Mao M, Peterse HL, van der Kooy K, Marton MJ, Witteveen AT, Schreiber GJ, Kerkhoven RM, Roberts C, Linsley PS, Bernards R and Friend SH. Gene expression profiling predicts clinical outcome of breast cancer. Nature 2002; 415: 530-536.

[44] Wang Y, Klijn JG, Zhang Y, Sieuwerts AM, Look MP, Yang F, Talantov D, Timmermans M, Meijer-van Gelder ME, Yu J, Jatkoe T, Berns EM, Atkins D and Foekens JA. Gene-expression profiles to predict distant metastasis of lymph-node-negative primary breast cancer. Lancet 2005; 365: 671-679.

[45] Sotiriou C, Wirapati P, Loi S, Harris A, Fox S, Smeds J, Nordgren H, Farmer P, Praz V, Haibe-Kains B, Desmedt C, Larsimont D, Cardoso F, Peterse H, Nuyten D, Buyse M, Van de Vijver MJ, Bergh J, Piccart M and Delorenzi M. Gene expression profiling in breast cancer: understanding the molecular basis of histologic grade to improve prognosis. J Natl Cancer Inst 2006; 98: 262-272.

[46] Toussaint J, Sieuwerts AM, Haibe-Kains B, Desmedt C, Rouas G, Harris AL, Larsimont D, Piccart M, Foekens JA, Durbecq V and Sotiriou C. Improvement of the clinical applicability of the Genomic Grade Index through a qRT-PCR test performed on frozen and formalin-fixed paraffin-embedded tissues. BMC Genomics 2009; 10: 424.

[47] Paik S, Tang G, Shak S, Kim C, Baker J, Kim W, Cronin M, Baehner FL, Watson D, Bryant J, Costantino JP, Geyer CE Jr, Wickerham DL and Wolmark N. Gene expression and benefit of chemotherapy in women with node-negative, estrogen receptor-positive breast cancer. J Clin Oncol 2006; 24: 3726-3734.

[48] Ma XJ, Salunga R, Dahiya S, Wang W, Carney E, Durbecq V, Harris A, Goss P, Sotiriou C, Erlander M and Sgroi D. A five-gene molecular grade index and HOXB13:IL17BR are complementary prognostic factors in early stage breast cancer. Clin Cancer Res 2008; 14: 2601-2608.

[49] Ma XJ, Wang Z, Ryan PD, Isakoff SJ, Barmettler A, Fuller A, Muir B, Mohapatra G, Salunga R, Tuggle JT, Tran Y, Tran D, Tassin A, Amon P,

Wang W, Wang W, Enright E, Stecker K, Estepa-Sabal E, Smith B, Younger J, Balis U, Michaelson J, Bhan A, Habin K, Baer TM, Brugge J, Haber DA, Erlander MG and Sgroi DC. A two-gene expression ratio predicts clinical outcome in breast cancer patients treated with tamoxifen. Cancer Cell 2004; 5: 607-616.

[50] van de Vijver MJ, He YD, van't Veer LJ, Dai H, Hart AA, Voskuil DW, Schreiber GJ, Peterse JL, Roberts C, Marton MJ, Parrish M, Atsma D, Witteveen A, Glas A, Delahaye L, van der Velde T, Bartelink H, Rodenhuis S, Rutgers ET, Friend SH and Bernards R. A gene-expression signature as a predictor of survival in breast cancer. N Engl J Med 2002; 347: 1999-2009.

[51] Buyse M, Loi S, van't Veer L, Viale G, Delorenzi M, Glas AM, d'Assignies MS, Bergh J, Lidereau R, Ellis P, Harris A, Bogaerts J, Therasse P, Floore A, Amakrane M, Piette F, Rutgers E, Sotiriou C, Cardoso F, Piccart MJ and Consortium T. Validation and clinical utility of a 70-gene prognostic signature for women with node-negative breast cancer. J Natl Cancer Inst 2006; 98: 1183-1192.

[52] Ivshina AV, George J, Senko O, Mow B, Putti TC, Smeds J, Lindahl T, Pawitan Y, Hall P, Nordgren H, Wong JE, Liu ET, Bergh J, Kuznetsov VA and Miller LD. Genetic reclassification of histologic grade delineates new clinical subtypes of breast cancer. Cancer Res 2006; 66: 10292-10301.

[53] Wirapati P, Sotiriou C, Kunkel S, Farmer P, Pradervand S, Haibe-Kains B, Desmedt C, Ignatiadis M, Sengstag T, Schutz F, Goldstein DR, Piccart M and Delorenzi M. Meta-analysis of gene expression profiles in breast cancer: toward a unified understanding of breast cancer subtyping and prognosis signatures. Breast Cancer Res 2008; 10: R65.

[54] Reis-Filho JS and Pusztai L. Gene expression profiling in breast cancer: classification, prognostication, and prediction. Lancet 2011; 378: 1812-1823.

[55] Dowsett M, Cuzick J, Wale C, Forbes J, Mallon EA, Salter J, Quinn E, Dunbier A, Baum M, Buzdar A, Howell A, Bugarini R, Baehner FL and Shak S. Prediction of risk of distant recurrence using the 21-gene recurrence score in node-negative and node-positive postmenopausal patients with breast cancer treated with anastrozole or tamoxifen: a TransATAC study. J Clin Oncol 2010; 28: 1829-1834.

[56] Albain KS, Barlow WE, Shak S, Hortobagyi GN, Livingston RB, Yeh IT, Ravdin P, Bugarini R, Baehner FL, Davidson NE, Sledge GW, Winer EP, Hudis C, Ingle JN, Perez EA, Pritchard KI, Shepherd L, Gralow JR, Yoshizawa C, Allred DC, Osborne CK, Hayes DF and Breast Cancer Intergroup of North A. Prognostic and predictive value of the 21-gene recurrence score assay in postmenopausal women with node-positive, oestrogen-receptor-positive breast cancer on chemotherapy: a retrospective analysis of a randomised trial. Lancet Oncol 2010; 11: 55-65.

[57] Blenkiron C, Goldstein LD, Thorne NP, Spiteri I, Chin S, Dunning MJ, Barbosa-Morais NL, Teschendorff AE, Green AR, Ellis IO, Tavaré S, Caldas C and Miska EA. MicroRNA expression profiling of human breast cancer identifies new markers of tumor subtype. Genome Biol 2007; 8: R214.

[58] Berezikov E, Guryev V, van de Belt J, Wienholds E, Plasterk RH and Cuppen E. Phylogenetic shadowing and computational identification of human microRNA genes. Cell 2005; 120: 21-24.

[59] Griffiths-Jones S, Saini HK, van Dongen S and Enright AJ. miRBase: tools for microRNA genomics. Nucleic Acids Res 2008; 36: D154-158.

[60] Jansson MD and Lund AH. MicroRNA and cancer. Mol Oncol 2012; 6: 590-610.

[61] Farazi TA, Hoell JI, Morozov P and Tuschl T. MicroRNAs in human cancer. Adv Exp Med Biol 2013; 774: 1-20.

[62] Nana-Sinkam SP and Croce CM. Clinical applications for microRNAs in cancer. Clin Pharmacol Ther 2013; 93: 98-104.

[63] Iorio MV, Ferracin M, Liu CG, Veronese A, Spizzo R, Sabbioni S, Magri E, Pedriali M, Fabbri M, Campiglio M, Menard S, Palazzo JP, Rosenberg A, Musiani P, Volinia S, Nenci I, Calin GA, Querzoli P, Negrini M and Croce CM. MicroRNA gene expression deregulation in human breast cancer. Cancer Res 2005; 65: 7065-7070.

[64] Volinia S, Calin GA, Liu CG, Ambs S, Cimmino A, Petrocca F, Visone R, Iorio M, Roldo C, Ferracin M, Prueitt RL, Yanaihara N, Lanza G, Scarpa A, Vecchione A, Negrini M, Harris CC and Croce CM. A microRNA expression signature of human solid tumors defines cancer gene targets. Proc Natl Acad Sci U S A 2006; 103: 2257-2261.

[65] Iorio MV, Casalini P, Piovan C, Di Leva G, Merlo A, Triulzi T, Menard S, Croce CM and Tagliabue E. microRNA-205 regulates HER3 in human breast cancer. Cancer Res 2009; 69: 2195-2200.

[66] Di Leva G, Gasparini P, Piovan C, Ngankeu A, Garofalo M, Taccioli C, Iorio MV, Li M, Volinia S, Alder H, Nakamura T, Nuovo G, Liu Y, Nephew KP and Croce CM. MicroRNA cluster 221-222 and estrogen receptor alpha interactions in breast cancer. J Natl Cancer Inst 2010; 102: 706-721.

[67] Spizzo R, Nicoloso MS, Lupini L, Lu Y, Fogarty J, Rossi S, Zagatti B, Fabbri M, Veronese A, Liu X, Davuluri R, Croce CM, Mills G, Negrini M and Calin GA. miR-145 participates with TP53 in a death-promoting regulatory loop and targets

estrogen receptor-alpha in human breast cancer cells. Cell Death Differ 2010; 17: 246-254.

[68] Almeida MI, Reis RM and Calin GA. MYC-microRNA-9-metastasis connection in breast cancer. Cell Res 2010; 20: 603-604.

[69] Pinto R, De Summa S, Danza K, Popescu O, Paradiso A, Micale L, Merla G, Palumbo O, Carella M and Tommasi S. MicroRNA expression profiling in male and female familial breast cancer. Br J Cancer 2014; 111: 2361-2368.

[70] Fassan M, Baffa R, Palazzo JP, Lloyd J, Crosariol M, Liu CG, Volinia S, Alder H, Rugge M, Croce CM and Rosenberg A. MicroRNA expression profiling of male breast cancer. Breast Cancer Res 2009; 11: R58.

[71] Esteller M. Non-coding RNAs in human disease. Nat Rev Genet 2011; 12: 861-874.

[72] Mercer TR, Dinger ME and Mattick JS. Long non-coding RNAs: insights into functions. Nat Rev Genet 2009; 10: 155-159.

[73] Van Roosbroeck K, Pollet J and Calin GA. miRNAs and long noncoding RNAs as biomarkers in human diseases. Expert Rev Mol Diagn 2013; 13: 183-204.

[74] Wang KC and Chang HY. Molecular mechanisms of long noncoding RNAs. Mol Cell 2011; 43: 904-914.

[75] Yang KC, Yamada KA, Patel AY, Topkara VK, George I, Cheema FH, Ewald GA, Mann DL and Nerbonne JM. Deep RNA sequencing reveals dynamic regulation of myocardial noncoding RNAs in failing human heart and remodeling with mechanical circulatory support. Circulation 2014; 129: 1009-1021.

[76] Li R, Qian J, Wang YY, Zhang JX and You YP. Long noncoding RNA profiles reveal three molecular subtypes in glioma. CNS Neurosci Ther 2014; 20: 339-343.

[77] Qi P and Du X. The long non-coding RNAs, a new cancer diagnostic and therapeutic gold mine. Mod Pathol 2013; 26: 155-165.

[78] Martens-Uzunova ES, Bottcher R, Croce CM, Jenster G, Visakorpi T and Calin GA. Long non-coding RNA in prostate, bladder, and kidney cancer. Eur Urol 2014; 65: 1140-1151.

[79] Askarian-Amiri ME, Crawford J, French JD, Smart CE, Smith MA, Clark MB, Ru K, Mercer TR, Thompson ER, Lakhani SR, Vargas AC, Campbell IG, Brown MA, Dinger ME and Mattick JS. SNORD-host RNA Zfas1 is a regulator of mammary development and a potential marker for breast cancer. RNA 2011; 17: 878-891.

[80] Figueroa ME, Lugthart S, Li Y, Erpelinck-Verschueren C, Deng X, Christos PJ, Schifano E, Booth J, van Putten W, Skrabanek L, Campagne F, Mazumdar M, Greally JM, Valk PJ, Lowenberg B, Delwel R and Melnick A. DNA methylation signatures identify biologically distinct subtypes in acute myeloid leukemia. Cancer Cell 2010; 17: 13-27.

[81] Lao VV and Grady WM. Epigenetics and colorectal cancer. Nat Rev Gastroenterol Hepatol 2011; 8: 686-700.

[82] Toyota M, Ahuja N, Ohe-Toyota M, Herman JG, Baylin SB and Issa JP. CpG island methylator phenotype in colorectal cancer. Proc Natl Acad Sci U S A 1999; 96: 8681-8686.

[83] Bae YK, Brown A, Garrett E, Bornman D, Fackler MJ, Sukumar S, Herman JG and Gabrielson E. Hypermethylation in histologically distinct classes of breast cancer. Clin Cancer Res 2004; 10: 5998-6005.

[84] Roessler J, Ammerpohl O, Gutwein J, Steinemann D, Schlegelberger B, Weyer V, Sariyar M, Geffers R, Arnold N, Schmutzler R, Bartram CR, Heinrich T, Abbas M, Antonopoulos W, Schipper E, Hasemeier B, Kreipe H and Lehmann U. The CpG island methylator phenotype in breast cancer is associated with the lobular subtype. Epigenomics 2015; 7: 187-99.

[85] Montes de Oca R, Gurard-Levin ZA, Berger F, Rehman H, Martel E, Corpet A, de Koning L, Vassias I, Wilson LO, Meseure D, Reyal F, Savignoni A, Asselain B, Sastre-Garau X and Almouzni G. The histone chaperone HJURP is a new independent prognostic marker for luminal A breast carcinoma. Mol Oncol 2014; 9: 657-74.

[86] Osmanbeyoglu HU, Pelossof R, Bromberg JF and Leslie CS. Linking signaling pathways to transcriptional programs in breast cancer. Genome Res 2014; 24: 1869-1880.

[87] Gatza ML, Lucas JE, Barry WT, Kim JW, Wang Q, Crawford MD, Datto MB, Kelley M, Mathey-Prevot B, Potti A and Nevins JR. A pathway-based classification of human breast cancer. PNAS 2010; 107: 6994-6999.

[88] Curtis C, Shah SP, Chin SF, Turashvili G, Rueda OM, Dunning MJ, Speed D, Lynch AG, Samarajiwa S, Yuan Y, Graf S, Ha G, Haffari G, Bashashati A, Russell R, McKinney S, Langerod A, Green A, Provenzano E, Wishart G, Pinder S, Watson P, Markowetz F, Murphy L, Ellis I, Purushotham A, Borresen-Dale AL, Brenton JD, Tavare S, Caldas C and Aparicio S. The genomic and transcriptomic architecture of 2,000 breast tumours reveals novel subgroups. Nature 2012; 486: 346-352.

[89] TCGA. Comprehensive molecular portraits of human breast tumours. Nature 2012; 490: 61-70.