Molecular Classification of Breast Cancer: An Overview with Emphasis on Ethnic Variations and Future Perspectives

Mohamed A. Shawarby, Dalal M. Al-Tamimi, Ayesha Ahmed
Department of Pathology, College of Medicine, University of Dammam and King Fahd Hospital of the University, Al-Khobar, Kingdom of Saudi Arabia

Address for correspondence: Dr. Mohamed A. Shawarby, P. O. Box 2208, Al-Khobar 31952, Kingdom of Saudi Arabia.
E-mail: melshawarby46@hotmail.com

ABSTRACT
Morphologically identical breast cancers can display divergent clinical outcomes and responses to therapy. This can predominantly be attributed to molecular class differences that exist amongst histologically similar cancer types. Consequently, molecular classification can be more powerful than histopathology as a predictive factor for the different treatments. This article reviews the molecular classification of breast cancer and emphasizes that ethnic variations may exist in molecular class prevalence patterns. It also highlights key insights into the currently defined molecular classes as provided by ongoing research on primary breast cancers using recent state-of-the-art technology. Such research is revealing that significant molecular heterogeneity may exist within the molecular classes themselves. More diverse ethnic variations may also be unraveled. The results of ongoing and upcoming research may provide more precise prognostic and predictive information about breast cancer and perhaps a breakthrough step toward “personalization” of breast cancer treatment. Forty-one relevant articles (2000-2012) extracted through PubMed and Google advanced searches and at our institute’s library were utilized to prepare the article, along with results of published and ongoing research by the authors.

Key words: Breast cancer, classifications, molecular, prognosis, treatment

INTRODUCTION
Although significant progress in breast cancer diagnosis and therapy has been made, breast cancer remains a heterogeneous disease with marked clinical and morphological diversities as well as variability in prognosis and response to various therapies. Heterogeneity also exists between breast cancers originating from distinct geographic locations. Around the world, distinct ethnic groups have been found to share similarly higher breast cancer incidence and grade relative to wider populations. Such observations can predominantly be attributed to molecular class differences that exist amongst histologically similar breast cancer types.\(^\text{1}\) Gene expression analysis has demonstrated distinct molecular classes of breast cancers based on the degree of expression of a select number of genes, which can be translated into more prognostically and therapeutically useful information than can be provided by existing histological classification systems. In this article, we present an overview of the molecular classification of breast cancer with emphasis on ethnic variations and future perspectives.

<table>
<thead>
<tr>
<th>Access this article online</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Quick Response Code:</strong></td>
</tr>
<tr>
<td>Website: <a href="http://www.sjmms.net">www.sjmms.net</a></td>
</tr>
<tr>
<td>DOI: 10.4103/1658-631X.112908</td>
</tr>
</tbody>
</table>
MATERIALS AND METHODS

Forty-one relevant articles (2000-2012) extracted through PubMed and Google advanced searches and at our institute’s library were utilized to prepare the article, along with results of published and ongoing research by the authors.

Molecular classification of breast cancer

Breast cancer is a heterogeneous disease showing marked clinical and morphological diversities as well as variability in prognosis and response to different therapeutic modalities. The existing histological classification systems for breast cancer are far from being accurate in predicting the prognosis or selecting the appropriate treatment of a given patient. Morphologically identical tumors can display divergent clinical outcomes and responses to therapy. This can predominantly be attributed to molecular class differences that exist amongst histologically similar cancer types. Consequently, molecular classification can be more powerful than histopathology as a predictive factor for the different treatments. This would result in less frequent use of chemotherapy with considerable advantages in reducing toxicity and costs.

The development of molecular analytical methods dates back to a quarter of a century, when immunohistochemistry (IHC) first allowed us to segregate breast cancers into two main classes: estrogen receptor positive (ER+) and estrogen receptor negative (ER−). A decade later, the next step forward was the emergence of nucleic acid in situ hybridization. This led to the identification of two new categories, dependent on whether human epidermal growth factor receptor-2 (HER2) was amplified or not. Low-grade cancers have positive ER and PR, but the high-grade cancers are found to be negative for ER and PR and also have an overexpression or amplification of HER2 with complex karyotypes. Many other single gene molecular markers were assessed, but failed to prove clinically relevant. Further progress in this area was the development of gene expression profiling, providing the simultaneous assessment of multiple genes and thus offering a more reliable prognostic and predictive indicator.

Perou et al. were the first to provide a classification system based on gene expression analysis, and this consisted of four major molecular classes of breast cancer: luminal-like, basal-like, normal-like, and HER-2 positive. Subsequent studies suggested the existence of more molecular classes and this ultimately led to addition of a fifth category, with the molecular spectrum now expanding to luminal A (LUMA), luminal B (LUMB), HER2 overexpressing, basal-like, and normal-like.

A further advancement in the field was the use of IHC as a surrogate for DNA microarray classification. Studies confirmed that it could reliably identify the major molecular classes of invasive breast carcinoma. This method represents a feasible alternative because many of the cases of breast cancer occur in places where analysis of prognostic factors needs to be economical, easy, and reproducible.

Recently published studies have used five surrogate IHC markers (ER, PR, HER2, CK5/6, and EGFR) for molecular class distinction with luminal tumors being categorized by hormone receptor (HR) positivity, HER2 expression, a feature of HER 2 tumors, and CK5/6 and/or EGFR (HER1) indicative of basal-like tumors.

At least five main molecular classes of breast cancer are currently recognized:
- Luminal A
- Luminal B
- HER2
- Basal
- Unclassified

Luminal classes express hormone receptors and have a pattern that agrees with the luminal epithelial component of the mammary gland. They express luminal CK8/18, ER and genes associated with its activation, such as LIV1 and CCND1.

Luminal A tumors are ER positive, PR positive or negative, HER2 negative, and CK5/6 and EGFR negative. Luminal A is the most frequent subtype. It shows a good prognosis and responds well to hormone therapy. Various studies have reported that ER+ tumors have little response to conventional chemotherapy. It has been demonstrated that patients with ER- tumors have more complete pathological responses to neoadjuvant chemotherapy than ER+ tumors.

Identification of luminal B tumors at the protein level is a point of controversy. Some authors have used the co-expression of HR and HER2 to define this group, based on the fact that the HER2-associated genes (i.e., ERBB2 and GRB7) are expressed in 30-50% of luminal tumors. However, these tumors have a poorer prognosis than LUMA tumors and are endocrine (tamoxifen) resistant and require estrogen deprivation in addition to blockage of HER2 pathways. Therefore, including
them as an integral component of endocrine-sensitive luminal tumors may not be justified. Due to these complexities, the HER2+ tumors need to be considered separately from pure luminal tumors, which should be further categorized as luminal A and luminal B, with those showing co-positivity of HER2 grouped into a separate hybrid category termed “luminal–HER2 hybrids.”\cite{16,17}

Bhargava et al. defined LUMA and LUMB as pure hormone receptor positive, the differentiating feature between them being the strong intensity of ER positivity in LUMA tumors.\cite{17}

Cheang et al. added that a Ki67 proliferation index of more than 13.25% is a hallmark of LUMB tumors. Ki67 is a nuclear marker of cell proliferation, and its expression correlates proportionally to poorer clinical outcomes.\cite{20}

HER2+ tumors are HER2 positive, ER and PR negative, and CK5/6 and EGFR negative.\cite{16,17} Overexpression of HER2 in tumor cells implies a poor prognosis.\cite{21,22} It also demonstrates the highest sensitivity to neoadjuvant chemotherapy based on anthracyclines and taxanes.\cite{21}

The poor prognosis of HER2 originates in its high risk of early relapse.\cite{24}

Basal-like tumors are CK5/6 and/or EGFR positive, ER and PR negative, and HER2 negative.\cite{16,17} The basal class is so named due to its pattern of expression that is similar to basal epithelial cells and normal myoepithelial cells of mammary tissue.\cite{6} This similarity is a product of the lack of ER expression and related genes, low expression of HER2, intense expression of CKs 5, 6, and 17, and the gene expression related with cellular proliferation.\cite{6} Using IHC, this class has also been called “triple negative” for not expressing ER, PR, or HER2.\cite{3} It has been associated with the \textit{BRCA1} mutation.\cite{25-27} Ribeiro et al. demonstrated that normal luminal cells that express CK5/6, which act as stem cells. These cells undergoing malignant transformation originate the basal phenotype of breast cancer. Under normal circumstances, \textit{BRCA1} could regularize proliferation of these cells; however, low regulation of \textit{BRCA1} could stimulate the expression of p53, leading to an abnormal growth of these CK5/6 positive cells.\cite{27} These tumors tend to be very aggressive, high grade, and with p53 mutation.\cite{24} Various studies have demonstrated the poor prognosis of this class.\cite{24,27,28} It is still not clear if this prognosis is due to a lack of therapeutic options or to an inherent aggressiveness.\cite{29} For being triple negative (ER, PR, and HER2 negative), it is not susceptible to conventional target treatments. However, these tumors present high sensitivity to chemotherapy.\cite{23}

With regard to the targeted directed therapeutic options, some early trials suggest that this class can be managed with manipulation of the epidermal growth factor.\cite{3}

Unclassified (penta −ve) tumors are ER and PR negative, HER2 negative, and CK5/6 and EGFR negative. They correspond to those triple-negative tumors not exhibiting basal markers.\cite{16} Unclassified cases were initially considered to be synonymous with “normal-like” breast cancers. These tumors cluster with non-tumoral breast cells and exhibit overexpression of \textit{PIK3R1} and \textit{AKR1C1}, in addition to other genomic alterations.\cite{30} The current concept states that the “normal-like” subtype is absolutely different from the unclassified (penta −ve) “ER−, PR−, HER2−, and CK5/6 and EGFR−” group, as absent or decreased expression of basal markers is not a feature compatible with the “normal-like” molecular class.\cite{31} They are very good prognostically\cite{30} and are grouped with the luminals, both of which exhibit low pathologic complete remission rates of 6%.\cite{23}

According to Huo et al., the unclassified category comprises two contrasting branches, a bad prognostic branch, characterized by the expression of vascular endothelial growth factor, B-cell lymphoma extra-large protein, and cyclin E, and a good prognosis branch, characterized by expression of B-cell lymphoma protein 2 and cyclin D1 as the distinguishing features.\cite{32} The unclassified and “normal-like” are completely separate entities and IHC surrogates for these categories have not yet been developed. Associating these with a particular set of negative or absent markers may lead to misinterpretations of their intrinsic biological characteristics.\cite{32}

**Ethnic variations**

Heterogeneity exists between breast cancers originating from distinct geographic locations. Around the world, distinct ethnic groups have been found to share similarly higher breast cancer incidence and grade relative to wider populations. In a recent study performed in the Eastern Province of Saudi Arabia, Al Tamimi et al. analyzed the spectrum of the molecular classes of breast cancer present in a cohort of Saudi patients. ER, PR, HER2, EGFR, and CK5/6 were used as surrogate markers for gene expression profiling to classify 231 breast cancer specimens. The prevalence patterns obtained were compared and contrasted with Western patterns and other regionally based studies.\cite{16}

Out of the 231 cases, 3.9% were classified as luminal A (strong ER +ve, PR +ve or −ve), 16% as luminal...
B (weak to moderate ER +ve and/or PR +ve), 17.3% as HER2+ (strong or moderately positive HER2 with confirmation by silver-enhanced in situ hybridization), and 10% as basal (CK5/6 or EGFR +ve). Co-positivity of different markers in varied patterns was seen in 10% of cases which were grouped into a hybrid category comprising luminal B–HER2, HER2–basal, and luminal–basal hybrids. 42.8% of the tumors were negative for all five immunohistochemical markers and were labeled as unclassified (penta−ve).

Comparing the prevalence patterns in that study with that of Western and other regionally based studies revealed striking differences. The main differences were in the luminal and the unclassified groups. Luminal (HR +ve) tumors as a group had a low prevalence in the Saudi luminal and the unclassified groups. Luminal (HR +ve) tumors as a group had a low prevalence in the Saudi cases (19.9%), in contrast to its high prevalence as reported in the Western studies (70.28-78.6%) and in regionally based studies from North Korea (44.5%) and Nigeria (80.2%). In addition, LUMB (16%) was more prevalent than LUMA (3.9%) in the Saudi study, while in the other studies LUMA was the more prevalent (ranging from 39.9 to 77.6%). Realizing that discrepancies could arise because of the different methods used in the differentiation between LUMA and LUMB, the authors grouped them together as luminal tumors for comparison studies. The unclassified tumors represented a small group in the studies from other regions, whereas they constituted a large proportion of the Saudi cases. The prevalence of HER2 (17.31%) in the Saudi study was also higher than that in the compared studies, with a range of 4-6.6%.

The results of the Saudi study clearly demonstrated that ethnic variations may exist in molecular class prevalence patterns, a view supported by earlier studies from Saudi Arabia that have similarly shown variability in molecular breast cancer class distribution between Saudi and Caucasian populations, as well as a predominance of the high-grade pathway in breast cancer development in Middle East women. However, in a more recent study performed on a large cohort of Saudi breast cancer patients from the Western Province (852 cases), Satti reported an incidence of 64% for ER/PR+, 23% for HER2+, and 24% for triple-negative classes. ER/PR and HER2 status did not differ from that reported previously, showing a direct correlation to tumor type and grade of ductal carcinoma. However, a difference existed in the relatively lower ER positivity in patients aged >50 years and a higher percentage of triple-negative cases. The discrepancy between those results and those from some other parts of the Kingdom may be attributed to variation in the number of cases included in the different studies, to differences in the classification criteria employed or to variable heterogeneity amongst populations from different regions of the country. More national studies as well as comparative population-based studies using advanced molecular techniques to classify breast cancer (rather than IHC) may help unfold whether ethnic variations do really exist in molecular class prevalence.

Future perspectives

It would be fair to say that currently, molecular profiling of breast cancer does provide additional prognostic and predictive information to clinicopathological features and routinely used immunohistochemical markers. However, this information benefits a limited number of patients and is largely restricted to patients with ER-positive cancers. Moreover, assignment of molecular classes of breast cancer based on hierarchical cluster analysis is subjective and shows modest inter-observer reproducibility. In spite of this, ongoing research on primary breast cancers using recent state-of-the-art technology platforms including DNA methylation arrays, miRNA sequencing, Affimexier SNP arrays, exome sequencing, and reverse phase protein arrays is offering further opportunities to characterize more completely the molecular architecture of breast cancer and is revealing that significant molecular heterogeneity may exist within its currently defined molecular classes. More diverse ethnic variations may also be unraveled. In a recent pilot study performed on Canadian and Saudi breast cancer patient populations, Amemiya et al., using Next Generation SOLI dances sequencing and IonTorrent exome-targeted sequencing technologies, found a high prevalence for an SNV in FAM175A gene predicted to be deleterious in the Canadian as compared to the Saudi patients. Moreover, a high prevalence of MSH6 gene deletions was seen in the Saudi patients, resulting in a frameshift in the Saudi population compared to the Canadian population. The results of ongoing and upcoming research may provide more precise prognostic and predictive information about breast cancer and perhaps a breakthrough step toward “personalization” of breast cancer treatment.

CONCLUSION

Breast cancer remains a heterogeneous disease showing marked diversities spanning its morphology, clinical presentation, prognosis, and response to various therapies. Heterogeneity also exists between breast cancers affecting distinct ethnic backgrounds. Such heterogeneity is predominantly attributed to molecular differences that exist amongst histologically similar
cancer types. Gene expression analysis has demonstrated distinct molecular classes of breast cancer that can be translated into more prognostically and therapeutically useful information than can be provided by existing histological classification. Ongoing research on primary breast cancers using recent state-of-the-art technology is revealing that significant molecular heterogeneity may exist within its currently defined molecular classes. More diverse ethnic variations may also be unraveled. This may provide more precise prognostic and predictive information about breast cancer and perhaps a breakthrough step towards “personalization” of breast cancer treatment.

REFERENCES

28. Ribeiro-Silva A, RAMALHO LN, GARCIA SB, BRANDAO DF, CHAHUD S, ZUCOLOTO S. p65 correlates with both BRCA1 and cytokeratin 5 in invasive breast carcinomas: Further evidence for the pathogenesis
38. Wiegelt B Reis-Filho JS. Molecular profiling currently offers no more than tumour morphology and basic immunohistochemistry. Breast Cancer Res 2010;12 Suppl 4:S5.


Source of Support: Nil, Conflict of Interest: None declared.