Molecular classification of breast carcinoma

Michelle Alizart
Jodi Saunus
Margaret Cummings
Sunil R Lakhani

Abstract

Breast cancer is a complex disease encompassing multiple tumour entities each with a characteristic morphology and behaviour. Current clinical practice relies on the recognition of various pathology prognostic factors to guide patient management, including histological type and grade, stage and biomarker receptor status. However, there is increasing concern that these parameters are of limited value for the accurate prediction of individual patient outcome. The introduction of genome-wide microarray-based expression profiling studies has allowed better understanding of the molecular underpinning of several characteristics of breast cancer, including histological grade and metastatic potential. Expression profiling has also facilitated the identification of prognostic and predictive gene expression signatures and novel therapeutic targets. Here we review the evolution of molecular classification of breast cancer, including special types, the implications for clinical management, limitations of findings thus far and predictions for the future.

Keywords Breast cancer; breast cancer special types; gene expression profiling; molecular classification; prognostic signatures; therapeutic targets

Introduction

Breast cancer is not one disease, rather it comprises at least 18 morphologically distinct tumour types recognized by the World Health Organisation. It is well recognized that even breast cancers of the same histological type can display different clinical behaviour, including metastatic potential and response to therapy. More information is needed to guide management and this has motivated a huge international research effort to stratify breast tumours into clinically relevant subgroups according to their genomic and transcriptomic features. The ultimate goal of this endeavour is to contribute to a personalized model of breast cancer management, where therapies can be tailored to individual patients.

Current histopathological classification

Currently, breast cancer classification is based on the histopathological appearance of the primary tumour. Clinicopathological factors used in conjunction to assess prognosis, guide therapy and predict treatment response include tumour grade (differentiation), tumour stage (tumour size and the number of lymph nodes involved) and biomarker receptor status, together with patient age and menopausal status.

Prognostic factors: histological type and grade

Breast carcinomas are divided into *in-situ* (ductal and lobular) and invasive carcinomas, which are in turn grouped into ductal carcinoma of no special type (IDC-NST) representing 60—75% of all breast cancers or special types, of which the most common is lobular followed by tubular, papillary and mucinous types. Histological grade refers to the degree of tumour differentiation, i.e. how closely the tumour resembles its tissue of origin. The Nottingham modification of the Bloom and Richardson grading system scores three components of the tumour: the proportion showing gland/tubule formation, the degree of nuclear pleomorphism and the mitotic count. The overall grade is derived from the sum of these parameters, with Grade 1 tumours being the most differentiated and Grade 3 the least. Ki67 may also be evaluated as a proliferative marker giving information that is complementary to that given by the mitotic count or grade overall.

Predictive factors: HER2 and hormone receptor status

Each breast cancer is routinely scored for oestrogen receptor (ER), progesterone receptor (PR) and HER2 protein expression using immunohistochemistry on tissue sections, with *in-situ* hybridization (CISH/SISH or FISH) used in addition to assess *HER2* gene amplification, which is present in approximately 18% of breast cancers. These biomarkers provide predictive information with respect to patient management; for example, patients with ER-positive cancers may respond to hormonal therapy (e.g. tamoxifen, aromatase inhibitors) and patients with HER2-positive cancers may respond to trastuzumab (Herceptin®, a HER2-specific monoclonal antibody), while patients with ER-negative tumours or HER2 negative tumours are unlikely to respond.

Biomarkers can also provide important prognostic information. Patients with ER-positive tumours have longer disease-free survival times than those with ER-negative tumours, while patients with triple-negative (ER, PR and HER2 negative) tumours have a relatively short time to relapse, high relapse-associated mortality and a propensity for brain and lung metastases (reviewed elsewhere[1]). While stratification of breast tumours according to the levels of simple biomarkers (ER, PR, HER2) has been very useful clinically, combining results of multiple biomarkers may give additional prognostic information.

Michelle Alizart *BSc (Hons) MBBS is at Pathology Queensland, The Royal Brisbane and Women's Hospital, Queensland, Australia. Conflicts of interest: none.*

Jodi Saunus *BSc (Hons) PhD is at the University of Queensland, UQ Centre for Clinical Research, Royal Brisbane and Women's Hospital, Queensland, Australia. Conflicts of interest: none.*

Margaret Cummings *MBBS FRCPA PhD is at UQ Centre for Clinical Research, Royal Brisbane and Women's Hospital, Queensland, Australia. Conflicts of interest: none.*

Sunil R Lakhani *MBBS MD FRCPath FRCPA is at the UQ Centre for Clinical Research, Royal Brisbane and Women's Hospital, Queensland, Australia. Conflicts of interest: none.*

For example, assessment of basal markers such as cytokeratin 14 and cytokeratin 5/6 together with epidermal growth factor receptor (EGFR) can identify basal-like triple negative cases with a worse prognosis.

Prognostic algorithms

The St Gallen breast cancer guidelines provide clinicians with a consensus set of recommendations for determining the primary management of early breast cancer. Importantly, criteria are re-evaluated biannually by a panel of international experts using the most up to date clinical evidence, and are endorsed by many international oncology societies. An algorithm was developed at the 2009 consensus meeting to determine the most suitable adjuvant chemotherapy regimen. It incorporates tumour size, histological grade, vascular invasion and lymph node status together with standardized cut-offs for ER, PR, HER2 and Ki67.

Other algorithms have also been developed, with the view to incorporating multiple parameters known to influence patient outcome into a single score. The Nottingham Prognostic Index, calculated using tumour size, grade and lymph node status, assigns patients to one of six prognostic groups to predict 5-year-survival post surgery. Adjuvant! Online, is used to predict the benefit of adjuvant treatment given in early breast cancer and is calculated using patient age, tumour size and grade, nodal involvement and margin status. But evaluation of Adjuvant! Online has shown that its use does not significantly alter patient management when compared with decisions made by a multi-disciplinary clinical team.[2]

Current algorithms for breast cancer management are generally used to predict who will not respond to treatment; tests that provide a positive predictive value are more limited. Also, results of these positive predictive tests are not absolute. For example, some patients with HER2-positive disease show either *de novo* resistance or will develop resistance to trastuzumab over time. Additionally, approximately 15% of patients with breast carcinomas categorized as low grade with a low risk of aggressive behaviour will develop recurrent disease. There is a need to provide better and more detailed prognostic information in the clinical setting and a model that incorporates conventional histopathological parameters with molecular data derived from the patient's own tumour is very attractive from a management point of view.

Molecular classification in breast cancer

The molecular basis of breast cancer phenotypes was initially investigated using loss of heterozygosity analysis (LOH), which can indicate tumour suppressor gene inactivation. This was followed by comparative genomic hybridization (CGH),[3] which identified key genomic losses, gains and amplified loci in breast cancer, suggesting the early framework for its molecular classification, including distinct low grade and high grade arms.[4]

Since then, high-resolution, microarray-based gene expression profiling has allowed the molecular basis of breast cancer to be further elucidated. These techniques provide a platform with which to study of thousands of genes in a single experiment. A gene expression array is a grid comprising tens of thousands of short DNA fragments (probes) that have been spotted in picomolar amounts via covalent attachment to a solid surface, usually a glass chip. Probe sequences map to genes of interest such as protein-coding or regulatory RNA genes, with at least two different probes per gene usually incorporated for specificity. Complementary DNA is prepared by reverse transcription of RNA from clinical samples, fluorescently labelled, and then hybridized to the chip. Expression of each RNA molecule can be quantified from the fluorescence intensity relative to a control sample (labelled with a different dye — two colour experiment), or relative to the level in other samples processed in the same experiment (one colour approach).

Gene expression signatures from large cohorts of tumour can be grouped and classified using bioinformatics tools. In an unsupervised classification, tumours are clustered according to their similarities and differences, without any *a priory* assumption (class discovery). In contrast, supervised analysis is done to look for similarities and differences in specified groups e.g. grade 1 versus grade 3, responders versus non-responders (class comparison). In order to translate this type of molecular classification to clinical practice, algorithms known as single sample predictors (SSPs) have also been developed. These are used to assign individual patients to molecular groups for the purpose of clinical trials and to decide on therapy. It is worth noting that currently, these SSPs are not consistent in their abilities to assign individual tumours to the same molecular groups and hence further work will be needed before they are ready for clinical use.

Despite these limitations (including that mRNA levels are often not directly correlated with levels of their functional, and more clinically relevant protein products), global snapshots of tumour transcriptional activity have undoubtedly helped to explain the molecular basis of breast cancer heterogeneity. Furthermore, correlating molecular data with clinical follow-up has highlighted prognostic and predictive signatures and novel therapeutic targets. Molecular classification of breast cancer is undoubtedly a work in progress that continues to evolve in line with advances in genomic analysis and bioinformatics. The following sections will review the developing molecular taxonomy of breast cancer, its clinical impact and limitations.

Breast cancer intrinsic subtypes

In seminal studies by the Stanford group, unsupervised gene expression array analysis of two series of invasive breast cancers revealed that the most critical discriminator was ER status. Additionally, they characterized the five main intrinsic subtypes of breast cancer (Stanford taxonomy): ER-positive luminal A and luminal B subgroups, the ER-negative basal-like, HER2, and normal-like subgroups,[5,6] Significantly, these clusters are reproducible across tumour cohorts and array platforms, and correlate with incidence, treatment response and survival.

Luminal A tumours show high expression of ER and related gene networks, have lower proliferation rates, tend to be of low histological grade and have the best prognosis (examples shown in Figure 1). Luminal B tumours show lower expression of ER networks, are more often of higher histological grade, have higher proliferation rates and a worse prognosis than luminal A tumours. HER2 and basal-like tumours are typically associated with aggressive clinical behaviour and the former show over-expression and amplification of HER2 (17q11). However, a significant proportion of HER2-positive tumours is ER positive, and clusters with the luminal B subgroup. Basal-like tumours show gene expression signatures similar to that of normal basal/myoepithelial cells of the breast. They express cytokeratin 5/6, cytokeratin 14,