Breast Cancer Biomarker Discovery: Proteomics and Genomics Approaches

Akram Safaei¹, Mostafa Rezaei-Tavirani¹, Sara Sobhi², Mohammad Esmaeil Akbari²

Abstract
Breast cancer is one of the major health problems of the Eastern world. Regardless of the survival rate improvement with progression in screening and adjuvant systemic therapies, still one – third of the patients with primary breast cancer have recurrence of micrometastasis after 10 years. It is important to discover a reliable biomarker for detection of breast cancer. The underlying molecular mechanism of the disease needs to be better understood. Allied to genomics, proteomics technologies promise to be valuable for identifying new markers that improve screening, early diagnosis, prognosis and prediction of therapeutic response or toxicity, as well as the identification of new therapeutic targets. In this review, we present proteomic and genomic sciences have been used for differential analysis of breast cancer to find molecular changes of cancer for detection candidate biomarkers.

Keywords: Breast cancer; Biological Markers; Proteomics; Genomics

Introduction
Breast Cancer (BC) is a heterogeneous group of different tumor subtypes that vary in prognosis and response to therapy [1]. At present, the best available tool for the early detection of breast cancer is mammography. This imaging is the most effective approach for diagnosing BC in women older than 50 years of age. Although new improvements are being made in the resolution of these imaging techniques, tumors smaller than 5 mm usually go undiagnosed. Moreover, as dense breast tissue decreases the mammographic sensitivity in young women, the effectiveness of mammography has not been established [2].

Finally, high-grade tumors cannot be diagnosed with 1 to 2 years of regular mammography imaging. For these reasons, new approaches should be developed to improve diagnosis of breast cancer and to increase the overall and disease free survival rates of patients who were diagnosed with this disease.

Therefore, there is an urgent need for breast cancer early detection biomarkers given that none are currently available and given the considerable public health importance of breast cancer [3]. A biomarker is a substance that is objectively measured, indicates the presence of an abnormal condition within a patient and allows disease progression and/or therapeutic response to be monitored [4]. Biomarkers provide a powerful and dynamic approach to understanding the spectrum of malignancies with applications in observational and analytic epidemiology, randomized clinical trials, screening, diagnosis and prognosis [5].

Breast cancer has been among the earliest and most intensely-studied diseases using gene expression profiling and protein profiling technologies. The resulting molecular signatures help reveal the biological spectrum of breast cancers, providing diagnostic tools as well as prognostic and predictive gene signatures [6-8].

Recent advances in genomics and proteomics have contributed to our understanding of the natural history of cancers. Genomic techniques, such as DNA microarray analysis and proteomic methods, for example, 2-dimensional electrophoresis and mass spectrometry, are now commonly used to evaluate the expression profiles of genes and proteins in cancer cells, their surrounding tissues, and body fluids [9]. Biomarker discoveries for breast cancer detection have focused on blood and/or tissue, using proteomic [10-13], transcriptomic[14-17], and genomic approaches [18,19].

Genomics:

Initially, the field of genomics was focused on mapping the human genome and understanding the effects of single genes in a system. The sequencing of the human genome, along with the development of high throughput technologies, have shifted the
focus of the field of genomics to gene expression profiling and the quest to identify genetic signatures of disease [20]. Many different gene expression profiling technologies are currently in use, including cDNA microarrays, oligonucleotide arrays, and Serial Analysis of Gene Expression (SAGE) [21]. These technologies allow the simultaneous study and comparison of the expression of thousands of genes in varying conditions.

Inherent instability in genes that maintain genomic integrity, have been implicated in breast-cancer development. Although molecular mechanisms of tumor genesis are unclear at present, carcinogenic agents could contribute to fields of genomic instability localized to specific areas of the breast. Understanding the functional importance of genomic instability in breast cancer is important for improvement of diagnostic and treatment strategies [22].

1. Genetic mutations

To address the association between variants and breast cancer, an increasing number of articles on genetic association studies, Genome-Wide Association studies (GWASs), and related meta-analyzed pooled analyses have been published [23-25]. Genetic mutation in BRCA1 (Breast Cancer1) and BRCA2 (Breast Cancer 2) [26, 27], BRIPl [28], CHEK2 [29], ATM [30] and TP53 [31] result in increased risk of breast cancer. However, these are estimated to account for only 5% to 10% of breast cancer cases. In breast cancer, mutations in the BRCA1 and BRCA2 genes have been well-characterized to carry a high risk of the disease during a woman’s lifespan. These high risk genes contribute to only a small proportion of the familial cases of breast cancer. Further efforts reported the contribution of genetic mutations in other genes, including the estrogen receptor gene, TP53, CYP19, and mismatch repair genes to further investigate the genetic component of breast cancer [32]. Recent large-scale sequencing analysis of over 13,000 genes in a small collection of breast tumors identified 122 genes with somatic mutation frequencies higher than the background frequency.

2. Microsatellite instability

Understanding the functional importance of genomic instability in early carcinogenesis has important implications for improvement of diagnostic and treatment strategies [33]. Microsatellites are tandem repeats of simple polymorphic sequences randomly distributed in non-coding regions of DNA [34]. They can be used in cancer genetics and indirect cancer diagnosis and can help unraveling the genetic basis of tumor formation and progression of cancer [35]. Breast cancer is a complex disease in which numerous genetic alterations occur. The knowledge of specific genetic changes and their biological consequences is critical to understand breast cancer tumorigenesis, screening and treatment of patients. Microsatellite instability reflects replication errors induced by defective function of mismatch repair genes and is demonstrated with the appearance of novel, no inherited alleles in tumor cells and represents a specific pathway of tumor development. Both events serve as prognostic markers, which can be correlated with clinicopathological features and can help exploring breast cancer formation [36, 37].

3. Loss of chromosome

Similar to other solid tumors, chromosome loss is a common molecular defect in breast cancer [38]. Callahan et al. have found that Loss of Heterozygosity (LOH) on chromosome 17p in tumor DNAs is associated with breast tumors having a high proliferative index and that LOH on chromosome 7 is associated with patients having a poor prognosis [39]. In sporadic breast cancer Loss of Heterozygosity (LOH) at BRCA1 and BRCA2 loci is common [38, 40]. In fact, LOH at BRCA1 locus could be an independent prognostic biomarker useful in identifying a subgroup of patients with poor prognosis [38, 41]. Loss of heterozygosity at chromosome 17p are frequent in a variety of human malignancies such as sporadic breast carcinomas. Some study suggested the presence of tumor suppressor genes, independent of TP53, on 17p13.3 region [42] and loss of heterozygosity at the long arm of chromosome 16, E-cadherin deletion, in breast cancer cell [43]. Small interstitial deletions of the proximal-central region of 3p, with band 3p14 as a minimal common deleted segment, have recently been shown to occur in as many as 10% of carcinomas of the breast, often as the only chromosomal change [44]. High incidence and frequency of LOH in HER2, which indicate increased genetic instability, were found to be associated with the aggressive features of breast cancer [45]. Nowacka et al. demonstrated the association of Loss of Heterozygosity (LOH) and Microsatellite Instability (MSI) in the 12p13.3 and 1p32 chromosomal regions where RAD52 and RAD54 genes (Role in both DNA repair and DNA recombination) are localized with breast cancer [46].

4. Telomere shortening

Telomeres stabilize and protect chromosomal termini, but shorten due to cell division and oxidative damage [47, 48]. Critically short telomeres, in the setting of abrogated DNA damage checkpoints, cause chromosomal instability due to end-to-end chromosomal fusions, subsequent breakage, and rearrangement, resulting in an increased cancer incidence in animal models. Alterations in telomere maintenance mechanisms leading to short telomeres underlie different genetic disorders of ageing and cancer predisposition syndromes. It is known that short telomeres and subsequent genomic instability contribute to malignant transformation, and it is therefore likely that people with shorter telomeres are at higher risk for different types of cancer. Recently, the authors...
demonstrated that the genes BRCA1 and BRCA2 are modifiers of Telomere Length (TL) in familial breast cancer [49]. Reduced telomere DNA content is correlated with genomic instability and metastasis in invasive human breast carcinoma [50]. Zheng et al. showed telomere deficiencies on chromosomes 9p, 15p, 15q and Xp were associated with breast cancer risk in pre-menopausal women and it has been suggested a useful panel of blood-based biomarkers for breast cancer risk assessment [51]. Telomere shortening is a strong candidate for the cause of structural chromosome defects that contribute to breast cancer development [52-54].

5. Epigenetic biomarkers

Epigenetic changes are critical for development and progression of cancers, including breast cancer [55]. Significant progress has been made in the basic understanding of how various epigenetic changes such as DNA methylation, histone modification, miRNA expression, and higher order chromatin structure affect gene expression [56]. In addition to genetic alterations, epigenetic abnormalities are associated with all cancer types. The epigenetic mechanisms currently believed to play a role in cancer include:

1) DNA methylation of cytosine bases in CG rich sequences, called CpG Islands; 2) post-translational modifications of histones, which are proteins that form the nucleosomes, which regulate the packaging structure of the DNA (called chromatin); 3) microRNAs and noncoding RNAs; and 4) nucleosome positioning [57]. In cancer, epigenetic changes such as covalent addition of methyl groups to the genomic DNA itself are more prominent than genetic changes. DNA promoter methylation frequently occurs during breast carcinogenesis and is an early event in this process [58]. Some of epigenetic alterations including promoter hypermethylation of genes like P16INK4a (a prototypic inhibitor kinase protein), BRCA1, BRCA2, ERα (Estrogen Receptor alpha) and RARβ2 (Retinoic Acid Receptor β2), APC, and RASSF1A (RAS association domain Family 1A) have been associated with early stages of mammary gland tumorigenesis and have been suggested to be included in the models that evaluate individual breast cancer risk [59]. Hypermethylation can silence tumor suppressor genes and activate oncogenes, respectively, and thereby alter functional genes to promote abnormal cell growth [60]. In breast cancer, DNA methylation has shown promise as a potential biomarker for early detection, therapy monitoring, assessment of prognosis or prediction of therapy response [61]. The results of Tomassi et al's research suggested a critical role of homeobox gene such as the TLX1, HOXB13 (homeobox13), and HNF1B (Hepatocyte Nuclear Factor 18) genes methylation in the insurgence and/or progression of breast cancer [62]. Many of the changes affect gene expression and genome stability through inappropriate regulation of local chromatin structure [63]. Mapping of nucleosome positioning on several long gene regions subject to DNA methylation has identified instances of nucleosome repositioning [64]. In breast cancer, abnormal histone modification in combination with DNA hypermethylation is frequently associated with epigenetic silencing of tumor suppressor genes and genomic instability [65, 66]. Some recent studies show histone modification (another type of epigenetic alteration) is associated with increased risk of breast cancer [67, 68].

Figure 1. Steps of breast cancer biology study for biomarker discovery
Proteomics

The foundation for any biomarker discovery effort is based on identification of proteins that show differential expression between disease and control samples [69]. In general, there are two approaches to proteomic biomarker discovery: target specific and global/non directed. Target-specific approaches often use antibodies to screen specific proteins through western blot analysis, enzyme-linked immunosorbent assays, and antibody arrays. While these techniques are clinically applicable, they are generally low-throughput with regard to the number of proteins that can be surveyed at any one time. Thus, they may not be ideal for biomarker discovery [70]. In contrast, global/nondirected approaches may be better suited for biomarker discovery because they are relatively unbiased, high-throughput screens. Nondirected approaches can also be divided into two groups: those studies that rely on profiling of unidentified proteins and those that generate profiles of identified proteins [71].

1. Candidate biomarker for breast cancer

Proteomic studies show HER2 is the most prominent and commonly used biomarker for breast cancer detection [72], as well as MMP-2 (Matrix Metallo Proteinase-2 immuno-reactive protein), absence of estrogen and progesterone receptors and high expression of Ki-67 (Mib-1) antigen, Osteopontin (OPN), urinase type Plasmonogen Activator and its Inhibitors (PAI-1 and 2) and cathepsins (B and L) have also been indicated as prognostic biomarkers for breast cancer [73]. In a proteomics study of breast cancer serum, HSP27 (up-regulated) and 14-3-3 sigma (down regulated) were identified using 2D-PAGE coupled with Matrix Assisted Laser Desorption Ionization- Time of Flight- Mass Spectrometry (MALDI-TOF-MS) [74]. These proteins are involved in the regulation of the cell cycle machinery at several key points. Additionally, they appear to be associated, directly or indirectly, with signalling proteins including IGF-1 receptor, Raf, MEK kinases and PI3-kinase [76]. Several studies have reported sub regulation of the 14-3-3 protein in breast cancer, suggesting its role as a tumor suppressor [77]. Protein 14-3-3 was an early detection marker of breast cancer [78]. In another study 2D-DIGE analysis of serum samples and controls revealed that proapoptoprotein A-I, transferrin, and hemoglobin were up-regulated and three proteins, apolipoprotein A-I, apolipoprotein C-III, and haptoglobin a2 were downregulated in cancer patients [79]. Sanders et al. identified ubiquitin and S100-A8 to be decreased in tumor tissue (n=122) compared to normal tissue (n=167) [80]. In pathological characteristics of breast cancer, 97 biomarker proteins were found, including ER, PR, HER-2, p53, CK5/6 (cytokeratin), CK8/18, cyclin E, Ki-67, BCL2, BRCA1, cyclin D1 and E-cadherin [81]. Hudelist et al.'s results show a total of 32 epithelial proteins were differentially expressed and identified as cytokines, structural proteins, tumor suppressor genes, signal-transducers or cell-cycle regulators [32]. These results confirm results of previous studies that most biomarkers role in key points DNA- repair systems, cell proliferation and cell-adhesion [82-84]. Some suggested candidate biomarkers for breast cancer have been shown in Table 1.

Protein/gene expression does not necessarily reflect protein activity, which is often regulated via Post-Translation Modifications (PTM), of which phosphorylation is one of the most prominent. This is an important consideration because the activity of protein is a more relevant phenotype than its expression during pathogenesis. PTM is very important class of enzymes that are critical regulators of mitogenic and angiogenic signalling which resulting in a gain of function in various human cancers such as breast [85].

2. Proteomic approaches in biomarker discovery in breast cancer

LC-MS and CE-MS have been widely used analytical techniques in the biomarker discovery in proteomics [86]. His approach was used to identification of biomarkers in different cancers such as breast cancer [87].

MALDI-TOF (Matrix Assisted Laser Desorption Ionization- Time Of Flight- Mass Spectrometry) is another popular technique in proteomics for biomarker discovery. This strategy relies on detection of protein/peptide peaks that differ in their mass-to-charge ratio (m/z) peptide mass can be calculated from the time taken to reach the detector (“time of flight”). Protein biomarkers have been discovered by MALDI-TOF in breast cancer [88].

In some isotope labeled method in proteomics, extracted proteins from treatment and control samples are labeled with either light or heavy ICAT reagents by reacting with cysteinylthiols on the proteins. Peptides containing the labeled and unlabeled ICAT tags are recovered by avidin affinity chromatography and are then analyzed by LC-MS/MS. Differential protein expression is determined by the isotope peak ratio of the peptide [89]. Un-Beom Kang et al. show profiling of breast cancer plasma proteins by ICAT. They reported potential serological biomarker for the detection of breast cancer [90].

In SELDI, basically, the sample to be analyzed is spotted onto a "protein-chip array" designed for the analysis of eight samples on a single chip [91]. The surface may consist of various materials of different physico-chemical characteristics. It has...
been introduced as a strong approach to discover of breast cancer [92].

After diagnosis of biomarkers they can be introduced for validation by Western Blotting.

Table 1. Candidate protein biomarkers in breast cancer

<table>
<thead>
<tr>
<th>Candidate biomarkers</th>
<th>Function</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>HER2</td>
<td>Proto oncogene/development of cardiac and neural tissue(72)</td>
<td>71</td>
</tr>
<tr>
<td>MMP-2</td>
<td>Protease that digest ECM(73)</td>
<td>72</td>
</tr>
<tr>
<td>OSTEOPROTEIN</td>
<td>immune function</td>
<td>72</td>
</tr>
<tr>
<td>PAI-1</td>
<td>regulation of plasmin formation</td>
<td>72</td>
</tr>
<tr>
<td>Cathepsin b</td>
<td>functions in protein turnover</td>
<td>72</td>
</tr>
<tr>
<td>14-3-3 sigma</td>
<td>cell cycle progression 100</td>
<td>73,74,77</td>
</tr>
<tr>
<td>P53</td>
<td>Apoptosis</td>
<td>80</td>
</tr>
<tr>
<td>Cyclin E</td>
<td>cell proliferation</td>
<td>80</td>
</tr>
<tr>
<td>E-cadherin</td>
<td>cell adhesion</td>
<td>80</td>
</tr>
<tr>
<td>BRCA1</td>
<td>repair damage DNA</td>
<td>81</td>
</tr>
</tbody>
</table>

Conclusion

There are hundreds of identified candidate biomarkers, but these must be validated to prove their specificity and clinical relevance. The many lessons learned from the development of the genetic assays currently being used clinically are also applicable to the proteomic study design strategy. The goal of the proteomic and genomic assays should be to develop biomarkers for screening, diagnosis, prognosis, and treatment monitoring. Together with genomics, proteomics is well on the way to molecularly characterizing the different types of biomarkers in breast tumor, and thus defining new therapeutic targets for future treatment, as well as proteomics may be easily coupled with functional tests that are proximally impossible with genomics.

Acknowledgment

This research has been supported by student committee Research of Shahid Beheshti University of Medical Sciences.

Conflict of Interest

The authors have no conflict of interest in this article.

Authors’ Contribution

Mohammad Esmaeil Akbari and Mostafa Rezaei-Tavirani designed and revised the study; Akram Safaei contributed to the search and wrote the paper. Sara Sobhi helped in writing the manuscript. All the authors read and approved it finally.

References


36. Fukin K, Shen L, Patocs A, Mutter GL, Eng C. Genomic instability within tumor stroma and clinicopathological characteristics of sporadic primary invasive breast carcinoma. JAMA: the

Vol6, Suppl., Winter 2013


73. Leppä S, Saarto T, Vehmanen L, Blomqvist C, Elomaa I. A high serum matrix metalloproteinase-


