Genetic landscape of breast cancer

RNDr. Miroslav Tomka, PhD.
MEDIREX GROUP ACADEMY, n. o., Bratislava

Breast cancer encompasses heterogeneous group of tumors with different histological, biological and clinical behavior. Intra- and inter-tumor heterogeneity as a result of genetic and non-genetic factors has considerable impact on different responses to anticancer therapy and prognosis. Introduction of new methods such as next-generation sequencing or comparative genome hybridization uncovered complexity of genetic nature of breast cancer. This article summarizes some of the new findings on this field.

Key words: breast cancer, comparative genome hybridization, next-generation sequencing, copy number variations

Introduction

Mammary gland is a unique organ that undergoes remarkable changes during the different stages of lifetime. During each menstrual cycle, the mammary gland passes through the waves of proliferation and apoptosis. Proliferation of the mammary epithelial cells is mainly driven by cyclic fluctuations of the hormonal factors such as estrogen and progesterone. Apoptosis controlled form of cell suicide is regulated by both hormonal and non-hormonal factors, but this process is still not well elucidated. Pregnancy also has great impact on breast morphology and function since it leads to extensive ductal branching and alveogenesis (1).

Breast tissue gladly respondes to any hormonal or structural changes which makes it favorable target for tumorigenesis. Cancer can originate from any cell in the breast that has undergone tumorigenic transformation, mainly from epithelial tissue. This process can be partly explained as a progression from premalignant disease (e.g., hyperplasia, ductal carcinoma in situ) through invasive carcinoma to metastases. The whole process is accompanied and managed by accumulation of distinct genetic abnormalities. The common targets for breast cancer metastases are bone, lung, liver and brain. The prominent target organ is bone (2). The reason for this might lie in the fact that this tissue expresses higher level of hyaluronan and osteopontin. Hyaluronan interacts with osteopontin and serves as a specific ligand for CD44. This attachment complex is involved in breast cancer adhesion, migration and invasion (3). In addition, bone is an estrogen rich organ which might create a specific microenvironment for breast cancer cells (4).

Nowadays, two models are trying to explain the origin of cancer in general. First is based on the initial transformed cell that represents the cell-of-origin for the tumor. Accumulation of mutations in the regulatory genes, tumor suppressors or/and oncogenes gives such a cell growth advantage over the others what results to growth acceleration of some cell population and tumor formation. Probably, the best example illustrating the multistage neoplastic process „from one cell to tumor“ is the well-known model of colorectal tumorigenesis (5).

The second theory considers the possibility of existence of self-renewing cancer cells called cancer stem cells (CSC). CSCs are subpopulation of cancer cells with high proliferative and metastatic potential. Breast cancer CSCs were isolated and identified as a CD44+/CD24- phenotype cells (6). Some introduced other markers of CSCs including ALDH1, CD49f and CD61 (7, 8, 9). This theory is trying to explain resistance to conventional therapies by the CSCs’ ability of enhanced transmembrane transport outward by ABC transporters, specific mechanisms of DNA repair, ability to maintain specific signaling pathways involving key transcription factors, level of tumor suppressors and collaborative interactions among cancer stem cells to maintain specific microenvironment (10). Of course, each mentioned model has its strengths and weaknesses which are discussed elsewhere (11).

Classification of breast cancer

Breast cancer is classified into different subtypes that are associated with different patient survival outcomes. For many years, breast cancer has been classified based on clinicopathological features such as tumor type, tumor size, lymph node status and histological grade. One of the classification system is the Nottingham histologic score system. It is a grading system which utilizes features including tubule formation (similarity with normal breast duct structures), nuclear features (size and shape) and mitotic activity. According to this classification, low grade tumors (grade 1) tend to be less aggressive then the high grade tumors (grade 3). Despite not optimal reproducibility, grading system is the key point in the diagnosis and management of the disease (12).

Another classification of breast cancer is based on the expression of Estrogen Receptor (ER), Progesterone Receptor (PR), and Human Epidermal growth factor Receptor (HER2). Best survival rate was observed in the...
case of ER+/PR+/HER2- subtype, the difference in survival was less than 1 % between ER+/PR+/HER- and ER+/PR+/HER2+ subtypes. All of the ER+ subtypes have better survival and adjusted mortality than all of the ER- subtypes. This suggests that ER may be a more important factor in survival than HER2 (13). ER, PR and HER2 status are predictive factors of tumor response to ER, PR, and HER2 targeted drugs.

In 2000, a classification system for the breast tumors based on variations in gene expression patterns was described. It was later revised and redivided into four subtypes: luminal A, luminal B, HER2-enriched, and basal-like (14).

Luminal A type of cancers have a low histological grade and are predominantly ER-positive. Luminal B group cancers are predominantly ER-positive with intermediate to high histological grade. Prognosis is good in the case of luminal A tumors, whereas luminal B type has intermediate prognosis with high risk of relapse. Both types of cancer are suitable for hormonal therapy.

Third type – HER2-enriched – is characterized by high histological grade, ER- and PR-negative, HER2-positive with poor prognosis. Therapy uses drugs aimed against HER2 (e.g. trastuzumab).

Tumors that belong to the basal type or triple negative for ER, PR and HER2 group are of high histological grade with poor prognosis. Treatment of these cancer types is difficult (12).

**Mutations in familial breast cancer**

 Genome-wide scanning revealed genetic variation in over 75 loci significantly associated with familial breast cancer. Despite progress on this field, currently known risk alleles explain only about 40 % of familial bound breast cancer (15). Up to date, BRCA1 and BRCA2 remain the two most significant genes linked to familial breast and ovarian cancer and account for about 20 % of familial breast cancer (16).

Traditional view on the subject suggests the key role of BRCA1/2 in familial breast and ovarian cancer onset. Accordingly, mutation in one of these genes significantly increases cancer risk. However, things are not so easy. There are some reports out there claiming that some BRCA1/2 mutations increase breast cancer risk only moderately (e.g. BRCA1 c.1966Gln (17), or even mean low risk for the affected person (e.g. BRCA2 p.Lys3326*)) (18). Spectrum of mutations found in BRCA1 and BRCA2 is extremely wide. Some of the mutations have been found one time only, whereas some have been found recurrently, often in enclosed ethnic groups, for example Ashkenazi Jews (e.g. BRCA1 c.66_67delAG; BRCA2 c.5946delCT) (19).

Both BRCA1 and BRCA2 play critical roles in DNA repair, cell cycle checkpoint control and maintenance of genomic stability. Thus, searching for other genes involved in corresponding networks led to discovery of genes related to breast cancer, including ATM, CHEK2, BRIP1, PALB2 and NBS1. Mutations in these genes increase breast cancer risk approximately twice. Another group of genes with influence on familial breast cancer is composed of genes such as FAM175A, BARD1, RAD51C, MRE11, RAD50 (16).

Familial syndromes, such as Li-Fraumeni syndrome (TP53), Cowden syndrome (PTEN Hamartoma Tumour Syndrome) (PTEN), Peutz-Jeghers syndrome (STK11), and hereditary diffuse gastric cancer syndrome (CHD1) are known to be one of the prominent diagnostic marks of breast cancer. Families affected by breast cancer seldom carry mutations in these genes but person with mutation in one of these genes has 2- to 10-fold increased risk of getting breast cancer early during the lifetime (16).

Modern screening methods such as next generation sequencing (NGS) revealed rare alleles associated with breast cancer risk, including FANCM, BLM, FANC, XRCC2 and MCPH1/BRIT1 (20). The MCPH1 c.904_916del mutation was genotyped in 1 370 breast cancer cases (145 familial cases, 75 young cases diagnosed below the age of 40 years, and 1 150 cases unselected for a family history of cancer or age at disease onset) and 1 160 healthy geographically matched controls. The highest prevalence for MCPH1 c.904_916del was observed among the familial cases (5/145, 3.4 %), whereas only 5 of the 1 160 healthy controls (0.4 %) carried the mutation (20).

Castéra et al. used NGS 69 to detect germline deleterious alterations within BRCA1 and BRCA2 in 708 patients, 4 TP53 mutations in 468 patients and also 36 variations inducing either a premature stop codon or a splicing defect among other genes: 5/708 in CHEK2, 3/708 in RAD51C, 1/708 in RAD50, 7/708 in PALB2, 3/708 in MRE11A, 5/708 in ATM, 3/708 in NBS1, 1/708 in CDH1, 3/468 in MSH2, 2/468 in PMS2, 1/708 in BARD1, 1/468 in PMS1 and 1/468 in MLH3 (21).

Alorafi et al. screened BRCA1/2 negative families with breast cancer. As expected, they identified mutations in several well-known high-susceptibility and moderate-susceptibility genes, including ATM (~ 5 %), RAD50 (~ 3 %), CHEK2 (~ 2 %), TP53 (~ 1 %), PALB2 (~ 1 %), and MRE11A (~ 1 %). They also identified novel pathogenic variants in 30 other genes: MAP3K1, CASPB, RAD51B, ZNF217, CDKN2B-AS1 and ERBB2 including a splice site mutation which, as they predict, would generate a constitutively active HER2 protein (22).

Chrupek et al. screened 289 African American women for inherited mutations. African Americans have a disproportionate burden of aggressive young-onset breast cancer. Of patients with mutations, 80 % (52/65) carried mutations in BRCA1 and BRCA2 genes and 20 % (13/65) carried mutations in PALB2, CHEK2, BARD1, ATM, PTEN, or TP53. The mutational allelic spectrum was highly heterogeneous with 57 different mutations in 65 patients (23).
A total of 133 patients were enrolled in Li’s study. Total 30 patients (22.6%) were found to carry germline deleterious mutations: 9 in BRCA1, 11 in BRCA2, 2 in RAD50, 2 in TP53 and one each in ATM, BRIPI, FANCI, MSH2, MUTYH, and RAD51C (24).

These data clearly support the idea to implement NGS into clinical management of familial breast cancer. It is highly probable that near future will bring routine usage of panels containing all known breast cancer associated loci for mutation screening in affected families.

Copy number variations in breast cancer

During last two decades, our understanding of cancer related loci has increased. The comparative genome hybridization (CGH) is tightly connected with this progress, despite its limitations. CGH is used for mapping losses and gains of DNA in the cancer genomes. These experiments revealed some loci that are commonly affected in breast cancer. Most frequently gained regions are localized on the chromosomal arms 1q, 8q, 17q, as well as 11q and 20q. On the contrary, most frequently lost regions were observed on 8p, 11q and 16q (25,26,27).

Closer view on the frequency plots shows two most prevalent regions with high gain alterations: 1q and 8q (figure 1). Andre et al. observed the gain of the 153 Mb region at 1q in 55 % cases out of 106 breast cancer patients. The second most frequently observed region of gain was spanning throughout the 116.7 and 127.5 Mb of the 8q. These gains were shown in 58 % of all cases (27). Interestingly, long arm of chromosome 1 is almost exclusively affected by gain of genetic material in breast tumors. On the other hand, its short arm is mainly involved in loss (28).

On the contrary, loss of genetic material is more or less equally distributed throughout the genome. However, at least a 24.2 Mb region at 8p is more frequently lost in 51 % cases, followed by a 47.8 Mb region at 13q seen in 41 % cases (27). Similar findings were spotted by others as well (25). Some patients show more complex anomalies in their tumor genome (figure 2).

Several studies have shown positive correlation between copy number alteration and gene expression level. For example, Andre et al. identified 3,007 such genes by Affimetrix U133A probes, including MYC, FOXA1, FGFR3, FGFR4, CCND1, PAK1. In addition, amplification with positive gain/loss and gene expression were observed in two amplicons (8p11.1-12 and 17q11-21) where genes PROSC, GPR124, ADRB3, as well as genes for trans-membrane tyrosine kinases -ERBB2, FGFR1 are localized (27).

On the other hand, regions that were shown to be lost and highly correlated with changes in gene expression contain genes such as BRCA1, STAT3, STAT5A, STAT5B, and MAPT genes, as well as genes encoding chemokines (CCL3, CCL4, CCL5, CCL14, CCL15, CCL16, CCL18, CCL23) (27).

According to the accepted theory, cancerogenesis is a multistep process leading from primary cancer cells through tumor formation to metastasis. According to this theory, one can expect differences in genetic variations within different stages of tumor. Navin et al. dissected primary breast cancer tumors into sections and analyzed them by flow cytometry and CGH. Based on their observations, breast carcinomas divided into two groups: (1) mono-genomic and (2) polygenomic. Monogenomic tumors contained a single major clonal subpopulation with a highly stable chromosome structure. Polygenic tumors contained only a few, no more than three major chromosome subpopulations. They do not found any series of gradual intermediates which can confirm the multistage carcinogenesis (29).

In this context, it is not surprising that differences between primary tumors and metastasis are not so significant. Comparison of known driver oncogenes located within regions frequently amplified in breast cancer showed 100 % concordance for ERBB2 (17q12) and FGFR1 (8p11.23), 96 % for CCND1 (11q13.3) and PAK1 (11q14.1), and 88 % for MYC (8q24) (25). Mutational status of primary tumors and metastases by NGS showed tumor-metastases concordance of variants 92 % for recurrent mutations (AKT1, ERBB2, PIK3CA, TP53) and 73 % for non-recurrent variants (25). High level of similarity was found by others as well, for instance Vollebergh (30).

However, genetic footprint of the primary tumors and metastases is not always in perfect concordance. For example, an ATM-containing region was found to be deleted in metastases, but in primary tumors it was usually amplified (25).
Thus, it is evident that primary tumors and metastases share higher level of copy number variation, but a little less concordance is found in mutational status. This suggests that metastasis formation does not require much more genetic events as was thought. To elucidate the carcinogenesis in general, further investigation will need information from the other fields, for example proteomics (25).

Conclusion

Breast cancer is the most common cancer in women worldwide. Despite improvements in the diagnostics and treatment in the past decade, our knowledge of this disease is still limited. Differences on the gene level observed not only between two different patients, but even among the cells of the same tumor cause problems which makes treatment and prognosis more challenging. Treatment regime which is successful in one patients might not be of any value for another. Therefore, involvement of new techniques such as CGH might be a good way towards personalized medicine from which each patient can benefit the most.

Acknowledgments: The author acknowledges the funding provided by Research and Development Operational Programme ITMS 26210120026, 2013/1.02-SORO. Additionally, I would like to thank Andrej Gnip, BSc. for proofreading of the manuscript.

References