Associations of Genetic Polymorphisms with Breast Cancer Risk Among Chinese Han Women

XIN-XIA TIAN*, WEI-GANG FANG*
*Department of Pathology, Peking University Health Science Center, Beijing, China

Increasing evidence has demonstrated that the high-frequency, low-penetrant genetic variations play important roles in the carcinogenesis of breast cancer. Although each genetic variant confers modest effect on breast cancer risk, the multiple genetic variants can cumulatively result in considerable effects on breast cancer over decades. This review summarizes the study strategies on the most common genetic polymorphisms, single nucleotide polymorphisms, mainly including candidate gene approach and genome-wide association study (GWAS). The authors briefly introduce their research works as well as some other studies performed among Chinese Han women. The future challenge of genetic polymorphism association study is to identify the causal variants and elucidate their molecular mechanisms.

Key words: genetic polymorphism, breast cancer, Chinese Han women

Introduction

Breast cancer (BC) is the most common malignancy among women and it is caused by a complex combination of genetic, environmental and lifestyle factors. The major environmental and lifestyle risk factors include age, estrogen exposure (from endogenous and exogenous sources), smoking, radiation exposure, obesity, and lifestyle in general. The inherited predisposition to this malignancy has also been thoroughly studied. Researchers have identified several breast cancer predisposition genes, such as low-frequency, high-penetrance genes BRCA1, BRCA2, PTEN and p53, as well as low-frequency, intermediate-penetrance genes CHEK2, ATM and PALB2. However, these genes explain only a small proportion of the total genetic risk of BC. It is currently believed that the high-frequency, low-penetrant genetic variations play important roles in the carcinogenesis of breast cancer. The common genetic variations include single nucleotide polymorphisms (SNP), small insertion/deletions (Indels), small variable repeats, variable long tandem repeats and copy number variation, etc. SNPs, a single base change in a DNA sequence, are the most commonly studied ones among Chinese populations. Majority of these studies are carried out among Chinese Han population, which is the largest ethnic group and constitutes about 92% of the population of the People’s Republic of China.

Associations of SNPs with breast cancer risk

In the early stage of association studies, researchers focused on some functional SNPs of candidate...
genes. The term ‘functional variants’ refers to those polymorphisms that can be shown through laboratory experiments to alter encoded protein’s structure or expression level, and affect the protein’s function, interaction with other proteins, or half-life and stability (Figure-1). Over 100 association studies of candidate functional SNPs with breast cancer risk were performed among Chinese women via case-control study, involving genes that regulate estrogen biosynthesis and metabolism, carcinoegen metabolizing enzymes, cell signaling and proliferation pathways, repair of normal DNA damage, the mitotic cycle and apoptosis, etc. Although the effect of an individual SNP is generally small, the cumulative effect of functionally relevant SNPs may be large after several decades, and contribute to increased breast cancer risk.

Since individual SNPs may fail to capture the whole contributions of a candidate gene to a particular trait, haplotype-based association analyses are believed to provide higher resolution and potentially greater power for identifying modest etiological effects of genes. With the development of international HapMap Project and 1,000 Genomes Project, which provides a detailed catalogue of human genetic variations, more and more haplotype or gene-wide association studies have been performed. A set of closely linked SNP alleles in a region of a chromosome which tend to be inherited together (not easily separable by recombination) is called a haplotype, while a pair of haplotypes forms a diplotype (Figure-2). In the human genome, there is on average one SNP in every 500-1,000 nucleotides. Currently, it is still inconvenient to genotype all these known SNPs in a candidate gene. Fortunately, based on linkage disequilibrium (LD) theory, a much smaller subset of informative SNPs called haplotype-tagging SNPs (htSNP) or tagging SNPs (tSNP) can represent gene-wide common variations\(^4\)\(^5\), and capture the contribution of the whole SNPs in a candidate gene to a trait. So, haplotype-based analysis is a cost-effective strategy for investigating associations between candidate genes and complex traits. Owing to the frequent involvement of centrosome defects in breast cancer, we have been conducting a series of studies to determine whether common genetic variants in centrosomal genes, including *Aurora-A*, *Brcal*, *Centrobin*, *Nek2*, *CDK2*, *CCNE1*, *CDK1*, and *CCNB1*, contribute to breast cancer susceptibility, progression and patients’ survival. We have performed a comprehensive single nucleotide polymorphism (SNP) and haplotype association study in around 1,200 BC cases and 1,200 age-matched controls among Chinese Han women. Our studies indicate that genetic polymorphisms of centrosome-related genes affect breast cancer susceptibility, progression, and survival in Chinese Han women\(^6\). Apart from single gene or a few genes analyses, candidate signaling pathway analyses have become popular in recent years. Ma X, et al. evaluated 341 SNPs in 11 TGF-\(\beta\) signaling pathway genes using a multistage, case-control study among Asian women, and demonstrated that those women who carried minor allele homozygotes (GG) of TGFBR2(rs1078985 had a 24% reduced risk of breast cancer compared with major allele carriers (AG or AA; OR = 0.76; 95%CI = 0.65–0.89; \(p = 8.42 \times 10^{-4}\))\(^7\).

With the introduction of array technologies for the simultaneous examination of huge numbers of polymorphisms, genome-wide association study (GWAS) was brought to this association study. GWAS is a hypothesis-free approach to comprehensively measure polymorphisms in genomes. By using GWAS approach, researchers have identified many novel genetic loci that were not anticipated by the candidate gene strategy. The first breast cancer GWAS was published in 2007, discovering
FGFR2 loci. Zheng W, et al. published a GWAS study on breast cancer among Chinese women in 2009, showing that SNP rs2046210 at 6q25.1, located upstream of the gene encoding estrogen receptor alpha (ESR1), were strongly associated with breast cancer risk. Although GWAS can help the researchers better understand diseases, Chang CQ, et al. revealed that cancer GWAS and candidate gene meta-analyses reveals limited overlap but similar effect sizes, suggesting that well-conducted candidate gene studies may continue to contribute to our understanding of the genetic associations in the post-GWAS era.

Candidate gene and GWAS association studies have been criticized for non-replication of results, false positives, and small sample sizes. These concerns have prompted the use of meta-analyses or pooled analyses of multiple studies to minimize false-positive associations and assess the credibility of findings. Quite a few studies have suggested that the genetic variants identified in European studies may not have the same effect on breast cancer risk in a Chinese population, which could be due to the differences in allele frequencies and the extent of linkage disequilibrium across populations. For example, the Arg25Pro (rs1800471) variant in TGFBI, which has been reported to be associated with breast cancer in European populations, is rare in Asian populations. Therefore, it is important to evaluate the genetic risk of breast cancer among different populations.

**Identification of causal genetic polymorphisms**

The resulting associated SNPs, even those with high statistical significance, are at best proxies for the truly causal loci, most of which have not been clarified. The exact causal variants can only be obtained through further fine-mapping of the associated loci and well-designed appropriate laboratory experiments. In order to identify the causal variants at 6q15.1, in which region the SNP rs2046210 has been identified to be strongly associated with breast cancer risk by GWAS analysis, Wang Y, et al. assessed six potentially functional SNPs within the CCDC170 and ESR1 gene regions at 6q25.1. They found that variant rs9383935 was in high linkage disequilibrium (LD) with rs2046210, and rs9383935 showed a strong independent effect on breast cancer risk. Further functional analyses revealed that the rs9383935 (G>A) risk allele A could decrease the activity of the reporter gene in both MCF-7 and BT-474 breast cancer cell lines, probably due to an altered binding capacity of miR-27a to the 3' untranslated region sequence of CCDC170. The greatest challenge in the ‘post-GWAS’ era is to understand the functional consequences of these loci. Biological insights can clarify the mechanisms of carcinogenesis, and then be translated to clinical benefits, such as reliable biomarkers and effective strategies for BC screening and prevention.
Associations of other genetic polymorphisms with breast cancer

Apart from SNPs, there are also many other heritable genetic factors including insertion/deletion polymorphisms, copy number variations (CNVs), and chromosomal changes, etc. By using a haplotype-tagging SNP approach, Sun T, et al. identified a six-nucleotide deletion (-652 6N del) variant in the CASP8 promoter, and demonstrated that this genetic variant was associated with reduced susceptibility to multiple cancers, including breast cancers\textsuperscript{13}. Functional analyses reveal that this variant reduces the expression or activity of caspase-8 and activation-induced cell death (AICD) of T lymphocytes by destroying a binding element for stimulatory protein 1 (Sp1). Recently, Li XC, et al. showed that ErbB2 copy number variation was the frequent early event of human breast cancer\textsuperscript{14}. We believe that augmenting the SNP data with insertion/deletion and copy number variations, etc., could improve BC risk prediction.

Conclusion

Many studies confirm that genetic polymorphisms do affect breast cancer susceptibility among Chinese women. Despite conferring minimal increase in breast cancer risk, individual genetic variants can cumulatively result in considerable effects on breast cancer over decades. In order to fully use these genetic variants in the clinical practice, future research needs to identify causal variants and determine how these low-penetrance alleles interact with each other and with environmental factors.

References