

Commentary

Single nucleotide polymorphisms and breast cancer: not yet a success story

Rulla M Tamimi

Channing Laboratory, Department of Medicine, Brigham and Women's Hospital and Harvard Medical School, and Department of Epidemiology, Harvard School of Public Health, Boston, Massachusetts, USA

Corresponding author: Rulla Tamimi, rulla.tamimi@channing.harvard.edu

Published: 27 July 2006

This article is online at <http://breast-cancer-research.com/content/8/4/108>

© 2006 BioMed Central Ltd

Breast Cancer Research 2006, **8**:108 (doi:10.1186/bcr1529)

Abstract

Numerous studies have examined low penetrance susceptibility polymorphisms in candidate genes, with some reporting significant findings. However, for the most part these associations could not be replicated in subsequent studies, suggesting that the original observations were due to chance. The failure to identify meaningful common genetic variation in relation to breast cancer should give us pause for thought and make us reconsider our current research strategies. The most recent directions of pooling samples to increase statistical power and pursuing whole genome screens may overcome some obstacles while also creating new challenges. Future studies should perhaps also consider alternative designs such as using surrogate (preferably continuous) markers of breast cancer, focusing on high-risk populations, and defining pathologically distinct outcomes.

Introduction

There is considerable evidence that genetics plays a role in breast cancer. Women with a family history of breast cancer are at a near twofold increased risk for breast cancer. Identification of highly penetrant genes has supported the notion that breast cancer is a genetic disease. However, these genes (e.g. *BRCA1* and *BRCA2*) account for a very small percentage of breast cancers in the population [1]. The common disease-common variant hypothesis has been one of the overarching suppositions driving many of the breast cancer association studies [2,3]. This hypothesis suggests that the risk attributable to genetics for common diseases will come from alleles that are not under severe negative selection and that are in the population at a relatively high frequency. This hypothesis is appealing from a public health perspective because common variants will have the greatest impact at the population level.

The sequencing of the human genome and discovery of millions of single nucleotide polymorphisms (SNPs) [4] provided the opportunity to characterize human genetic

variation and its impact on breast cancer systematically. Numerous studies have examined low penetrance susceptibility polymorphisms in candidate genes, with some reporting significant findings. However, for the most part these associations could not be replicated in subsequent studies, suggesting that the original observations were due to chance [5]. In a pooled analysis of 46 association studies examining polymorphisms in 18 different genes, only three polymorphisms were significantly associated with breast cancer [5]. Despite considerable efforts made during the past few years to study genetic variation and breast cancer, there has been little success in identifying important contributions to breast cancer. The lack of robust and consistent findings from single SNP association studies led to the explosion of haplotype studies, which promised to examine more comprehensively the association between common genetic variation in candidate genes and risk for disease. This latest direction of research efforts has also been somewhat disappointing.

Models of genetics and breast cancer risk

It is becoming increasingly clear that our working hypotheses and approaches may have been too simplistic for a complex and multifactorial disease such as breast cancer. A more recent model of risk suggests that a large fraction of the burden of breast cancer risk may be concentrated in a relatively small proportion of the female population [6,7]. If this theory proves correct, then it could have profound implications for screening, risk prediction and, ultimately, chemoprevention. However, additional evidence to support this model is needed. Family history of breast cancer is a fairly common and robust risk factor for breast cancer, challenging the supposition that the contribution of genetics to risk is limited to only a small fraction of the population.

The prevailing polygenic model of breast cancer risk suggests that a moderate number of genes, each conferring a

SNP = single nucleotide polymorphism.

small amount of risk alone (relative risk 1.3-1.5), together would combine multiplicatively, resulting in modest susceptibility to breast cancer [7]. According to this model, more than 100 genes may contribute to breast cancer susceptibility. These genes could have either common or rare variants. Women carrying more variant alleles would be at greater risk than those carrying fewer. Recently, a rare variant (about 1.1% in healthy populations) identified in the *Chk2* gene has been reported to confer a 1.5-fold to 2.0-fold increase in breast cancer risk in individuals that carry this allele [8,9]. The CHEK2*1100delC allele, which abolishes kinase activity of the protein, lends support to the polygenic model [8,9]. To date, this is the first polymorphism to be successfully confirmed with strong statistical power. Additional confirmation of these results and identification of additional genes conferring modest risk will lend further support to this model.

Even in the best case scenario, in which researchers are able to identify a set of 100 genes that each modestly contribute to breast cancer risk, it is unclear what the public health utility of testing in the population would be. Perhaps, the greatest advancement would come from the potential identification of new pathways that contribute to breast cancer that could ultimately be targets for screening, chemoprevention, or treatment.

Limitations of current studies

There are several potential reasons for a lack of consistent findings between SNPs and breast cancer. First, breast cancer is a heterogeneous disease, but most analyses treat it as one. Just as nongenetic breast cancer risk factors may be differentially associated with specific tumor types [10], so might genetic factors. Second, most association studies to date have utilized a candidate gene approach, in which we are limited by our biologic knowledge. The genome contains an estimated 20,000-25,000 genes, many of which play redundant roles or are components of pathways. Thus, a functional polymorphism in one gene may be compensated for by a functionally similar or downstream gene. Third, by definition, low penetrance genes do not cause disease - they make individuals more susceptible. The same landmark study that provided statistical evidence that breast cancer was heritable also concluded that environmental factors and somatic changes are the major contributors to breast cancers occurring in the population [11]. Without considering other factors that influence this susceptibility, whether they be dietary factors, reproductive factors, or somatic alterations, it will be very difficult to detect an association in our traditional population studies.

One solution to overcome some of these challenges is to increase statistical power. Comprehensive efforts, such as the HapMap project to identify variation and determine linkage disequilibrium patterns, and advances in genotyping technology are providing the necessary reagents to perform adequately powered large-scale association studies. The

Breast and Prostate Cancer Cohort Consortium (BPC3) is pooling resources from nine cohort studies to include more than 5000 cases of breast cancer to address the role of common variants with respect to breast cancer [12]. That study should have sufficient power to detect small to moderate effects and will examine gene-gene and gene-environment interactions. The limited success of the candidate gene approach, together with technologic advances in the development of whole genome scans, has opened the door for less hypothesis driven research. Because many of the identified millions of SNPs are perfectly correlated with one another, one tagging SNP can serve as a proxy for other variants. It is believed that a few hundred thousand well selected SNPs can capture the majority of common genetic variation in the genome [13]. Whole genome scans can now be performed, and will allow researchers to conduct genome-wide association studies with no *a priori* assumptions about loci [14]. The most evident concern regarding the genome scans is that they will lead to false-positive results. Therefore, these studies also have foreseeable challenges in determining appropriate sample sizes, statistical issues of data analysis and accounting for multiple testing, and reproducibility in other data sets.

Conclusion

The failure to identify meaningful common genetic variation in relation to breast cancer should give us pause for thought and make us reconsider our current research strategies. The most recent directions of increasing sample size through pooling and pursuing whole genome screens may overcome some of the obstacles. However, both are still directed at detecting common genetic variation and may miss meaningful risk alleles occurring at less than 5%. Use of a highly heritable, continuous outcome measure such as mammographic density may be an alternative and powerful tool for identifying genes associated with breast cancer. Based on twin studies, the heritability of breast cancer has been estimated to be 27% [11]. In contrast, mammographic density - considered by some to be a surrogate marker of breast cancer risk - is a highly heritable trait, more so than breast cancer. The correlation between mammographic density in monozygotic twins ranges from 0.61 to 0.67 [15], suggesting that the majority of variation in mammographic density is explained by genetic factors. Future studies should perhaps also consider alternative designs such as using surrogate (preferably continuous) markers of breast cancer, focusing on high-risk populations, and defining pathologically distinct outcomes.

Competing interests

The author declares that they have no competing interest.

Reference

1. Rebbeck TR: **Inherited genetic predisposition in breast cancer. A population-based perspective.** *Cancer* 1999, **Suppl**:2493-2501.
2. Lander ES, Linton LM, Birren B, Nusbaum C, Zody MC, Baldwin J, Devon K, Dewar K, Doyle M, FitzHugh W, *et al.*: **Initial sequencing and analysis of the human genome.** *Nature* 2001, **409**:860-921.

3. Reich DE, Lander ES: **On the allelic spectrum of human disease.** *Trends Genet* 2001, **17**:502-510.
4. Sachidanandam R, Weissman D, Schmidt SC, Kakol JM, Stein LD, Marth G, Sherry S, Mullikin JC, Mortimore BJ, Willey DL, *et al.*: **A map of human genome sequence variation containing 1.42 million single nucleotide polymorphisms.** *Nature* 2001, **409**: 928-933.
5. Dunning AM HC, Pharoah PD, Teare MD, Ponder BA, Easton DF: **A systematic review of genetic polymorphisms and breast cancer risk.** *Cancer Epidemiol Biomarkers Prev* 1999, **8**:843-854.
6. Peto J, Mack TM: **High constant incidence in twins and other relatives of women with breast cancer.** *Nat Genet* 2000, **26**: 411-414.
7. Antoniou AC, Pharoah PD, McMullan G, Day NE, Stratton MR, Peto J, Ponder BJ, Easton DF: **A comprehensive model for familial breast cancer incorporating BRCA1, BRCA2 and other genes.** *Br J Cancer* 2002, **86**:76-83.
8. CHEK2 Breast Cancer Case-Control Consortium: **CHEK2*1100delC and susceptibility to breast cancer: a collaborative analysis involving 10,860 breast cancer cases and 9,065 controls from 10 studies.** *Am J Hum Genet* 2004, **74**: 1175-1182.
9. Meijers-Heijboer H, van den Ouweland A, Klijn J, Wasielewski M, de Snoo A, Oldenburg R, Hollestelle A, Houben M, Crepin E, van Veghel-Plandsoen M, *et al.*: **Low-penetrance susceptibility to breast cancer due to CHEK2(*)1100delC in noncarriers of BRCA1 or BRCA2 mutations.** *Nat Genet* 2002, **31**:55-59.
10. Colditz GA, Rosner BA, Chen WY, Holmes MD, Hankinson SE: **Risk factors for breast cancer according to estrogen and progesterone receptor status.** *J Natl Cancer Inst* 2004, **96**:218-228.
11. Lichtenstein P, Holm NV, Verkasalo PK, Iliadou A, Kaprio J, Koskenvuo M, Pukkala E, Skytthe A, Hemminki K: **Environmental and heritable factors in the causation of cancer: analyses of cohorts of twins from Sweden, Denmark, and Finland.** *N Engl J Med* 2000, **343**:78-85.
12. Hunter DJ, Riboli E, Haiman CA, Albanes D, Altshuler D, Chanock SJ, Haynes RB, Henderson BE, Kaaks R, Stram DO, *et al.*: **A candidate gene approach to searching for low-penetrance breast and prostate cancer genes.** *Nat Rev Cancer* 2005, **12**:977-985.
13. Gabriel SB, Schaffner SF, Nguyen H, Moore JM, Roy J, Blumenstiel B, Higgins J, DeFelice M, Lochner A, Faggart M, *et al.*: **The structure of haplotype blocks in the human genome.** *Science* 2002, **296**:2225-2229.
14. Hirschhorn JN, Daly MJ: **Genome-wide association studies for common diseases and complex traits.** *Nat Rev Genet* 2005, **6**: 95-108.
15. Boyd NF, Dite GS, Stone J, Gunasekara A, English DR, McCredie MR, Giles GG, Tritchler D, Chiarelli A, Yaffe MJ, *et al.*: **Heritability of mammographic density, a risk factor for breast cancer.** *N Engl J Med* 2002, **347**:886-894.