Review Article

Cytogenetic and Genome Research

Cytogenet Genome Res 2014;144:77–91
DOI: 10.1159/000369045

Accepted: August 6, 2014
by M. Schmid
Published online: November 15, 2014

Germline DNA Variations in Breast Cancer Predisposition and Prognosis: A Systematic Review of the Literature

Yadav Sapkota

The Neurogenetics Laboratory, Department of Genetics and Computational Biology, QIMR Berghofer Medical Research Institute, Brisbane, Qld., Australia

Key Words

Breast cancer · Breast cancer recurrence · CN-LOH · CNVs · Genetic interactions · Genome-wide association study · Heritability · Linkage disequilibrium · SNPs · Susceptibility

Abstract

Breast cancer is the most common cancer and the second leading cause of death in women worldwide. The disease is caused by a combination of genetic, environmental, lifestyle, and reproductive risk factors. Linkage and family-based studies have identified many pathological germline mutations, which account for around 20% of the genetic risk of familial breast cancer. In recent years, single nucleotide polymorphism-based genetic association studies, especially genome-wide association studies (GWASs), have been very successful in uncovering low-penetrance common variants associated with breast cancer risk. These common variants alone may explain up to an additional 30% of the familial risk of breast cancer. With the advent of available genetic resources and growing collaborations among researchers across the globe, the much needed large sample size to capture variants with small effect sizes and low population frequencies is being addressed, and hence many more common variants are expected to be discovered in the coming days. Here, major GWASs conducted for breast cancer predisposition and prognosis until 2013 are summarized. Few studies investigating other forms of genetic variations contributing to breast cancer predisposition and disease outcomes are also discussed. Finally, the potential utility of the GWAS-identified variants in disease risk models and some future perspectives are presented.

Breast cancer is the most frequently occurring cancer in women worldwide. The disease has a complex etiology and is believed to result from a combined interplay of genetic and non-genetic risk factors. Linkage and family-based studies have shown that breast cancer tends to cluster in families with a more than 2-fold increased risk among the first-degree relatives of affected individuals [Lichtenstein et al., 2000; Collaborative Group on Hormonal Factors in Breast Cancer, 2001]. Some of the familial clustering has been attributed to germline mutations in high- and moderate-penetrance genes such as *BRCA1/2, PTEN, ATM, CHEK2, TP53, PALB2*, and *BRIP1* [Hall et al., 1990; Malkin et al., 1990; Wooster et al., 1995; Liaw et al., 1997; CHEK2 Breast Cancer Case-Control Consortium, 2004; Renwick et al., 2006; Seal et al., 2006; Rahman

KARGER

E-Mail karger@karger.com
www.karger.com/cgr

1424–8581/14/1442–0077$39.50/0

Dr. Yadav Sapkota
The Neurogenetics Laboratory, Department of Genetics and Computational Biology
QIMR Berghofer Medical Research Institute
300 Herston Road, Brisbane, QLD 4006 (Australia)
E-Mail Yadav.Sapkota@qimrberghofer.edu.au

et al., 2007]. Together, these genes account for ~20–25% of the genetic risk of breast cancer. Extensive research efforts to identify additional high- or moderate-penetrance genes were largely unsuccessful, leading to the proposition of a polygenic model of disease inheritance to explain the remaining genetic risk of breast cancer [Pharoah et al., 2002, 2008; Piccolo et al., 2008; Smith et al., 2008]. Under the polygenic model, a combination of multiple low-penetrance genes/alleles, each with small effect sizes, contribute to the genetic risk of breast cancer.

Genetic association studies, especially case-control association studies, have been conducted to identify low-penetrance common variants (predominantly single nucleotide polymorphisms, SNPs) associated with many complex diseases, including breast cancer. Initial genetic association studies for breast cancer were focused on interrogating a handful of SNPs in candidate genes or pathways such as DNA repair, apoptosis and cell-cycle checkpoints [Allen-Brady et al., 2006; Bewick et al., 2006; Mitra et al., 2008; Sehl et al., 2009; Economopoulos and Sergentanis, 2010; Lin et al., 2011; Mangoni et al., 2011]. Earlier this decade, completion of the Human Genome Project and the International HapMap Project, coupled with advances in genotyping technologies, made it possible to analyze the whole genome for SNPs associated with complex traits. Such studies are termed as genome-wide association studies (GWASs). With increased collaborations among researchers across the globe, many international consortia have now formed, providing better expertise and resources for conducting GWASs with several thousands of cases and controls. Such large studies have enriched the statistical power needed to capture common variants with much lower effect sizes and population frequencies. To date, GWASs conducted for breast cancer have identified more than 80 breast cancer susceptibility loci [Easton et al., 2007b; Hunter et al., 2007; Stacey et al., 2007, 2008; Gold et al., 2008; Ahmed et al., 2009; Thomas et al., 2009; Zheng et al., 2009; Gaudet et al., 2010; Long et al., 2010, 2012; Turnbull et al., 2010; Cai et al., 2011; Fletcher et al., 2011; Haiman et al., 2011; Sehrawat et al., 2011; Ghoussaini et al., 2012; Kim et al., 2012; Bojesen et al., 2013; Couch et al., 2013; French et al., 2013; Garcia-Closas et al., 2013; Michailidou et al., 2013; Sapkota et al., 2013c], and many more studies are underway. Motivated by the huge success of breast cancer GWAS, few studies have also been conducted to identify SNPs associated with breast cancer outcomes and other associated sub-phenotypes [Garcia-Closas et al., 2008; Azzato et al., 2010a, b; Shu et al., 2012; Rafiq et al., 2013].

As an alternative source of genetic heritability for breast cancer, structural variations such as copy number variations (CNVs) and copy-neutral loss of heterozygosity (CN-LOH) are being interrogated for their associations with breast cancer predisposition as well as for disease recurrence [Wellcome Trust Case Control Consortium, 2010; Sapkota et al., 2013a]. GWAS literature has also witnessed few reports providing evidence for genetic interactions as another source to explain the remaining genetic risk of breast cancer [Milne et al., 2010; Campa et al., 2011; Nickels et al., 2013; Sapkota et al., 2013b]. As these alternative approaches are more complex and sometimes computationally not feasible, future studies may focus on these once current hurdles are overcome.

The present comprehensive study is aimed to provide a systematic review of the literature on breast cancer GWASs conducted until 2013, in addition to a general description of breast cancer etiology and its genetic basis. A brief outline of alternative sources of genetic heritability of breast cancer, such as genetic interactions among common variants and structural variations, is also provided. Lastly, a general discussion on the potential application of GWAS-identified common variants in breast cancer risk/prediction models as well as future perspectives on breast cancer GWAS are outlined.

Etiology and Genetic Basis of Breast Cancer

Breast cancer is a complex multifactorial disease, which results from an interplay of environmental, reproductive, lifestyle, and genetic risk factors. It has been estimated that approximately one-third of variations in breast cancer susceptibility is accounted for by inherited genetic risk factors, while environmental and lifestyle risk factors contribute to the remaining two-thirds [Lichtenstein et al., 2000; Collaborative Group on Hormonal Factors in Breast Cancer, 2001; Key et al., 2001]. Linkage studies have been successful in identifying predisposition factors for many diseases, including breast cancer [Easton et al., 1993]. The first breast cancer predisposition gene to be identified was *BRCA1*, located on chromosome 17q21. It was found in a linkage study in 1990, with a LOD score of 2.35 with a microsatellite marker (D17S74) [Hall et al., 1990; Solomon and Ledbetter, 1990]. The linkage was stronger in families with early onset of disease (<46 years) with LOD score of 5.98, while linkage vanished in families with late onset of disease, indicating that this gene may not contribute to predisposition to breast cancers that are sporadic in nature. A linkage study conducted in 1994

identified another breast cancer predisposition gene, *BRCA2*, located on chromosome 13q12–q13 [Wooster et al., 1995]. Both *BRCA1* and *BRCA2* genes play crucial roles in maintaining genomic stability by their involvement in repair of DNA double-strand breaks. Multiple germline mutations in *BRCA1* and *BRCA2* genes have been detected; however, they occur in a small fraction of total breast cancer cases [Easton et al., 2007a]. Studies have shown that most breast cancer-associated germline mutations in *BRCA1* and *BRCA2* genes result in premature truncation of encoded proteins, translational frame shifts and defective splice sites [Ford et al., 1994; Easton et al., 1995]. Both *BRCA1* and *BRCA2* are categorized as high penetrant breast cancer genes that confer a more than 10-fold increase in disease risk [Ford et al., 1994, 1998; Easton et al., 1995; Thompson et al., 2002]. Evidence from epidemiological studies suggests that breast cancer risk by age 70 may increase up to 87% in carriers of *BRCA1* mutations and up to 84% in carriers of *BRCA2* mutations [Ford et al., 1994, 1998; Thompson et al., 2002]. Since germline mutations in *BRCA1* and *BRCA2* genes are very rare, these 2 predisposition genes could only explain 15–20% of the genetic risk of breast cancer in the overall population [Easton, 1999; Peto et al., 1999; Antoniou et al., 2000, 2002].

Continued research efforts to characterize additional breast cancer predisposition genes resulted in the identification of multiple genes conferring moderate risk for breast cancer. Germline mutational screening of cancer-related pathway genes identified 2 cancer predisposition syndromes: Li-Fraumeni and Cowden syndrome [Malkin et al., 1990; Liaw et al., 1997]. Both syndromes are characterized by a variety of different individual germline mutations in their causative tumor suppressor genes, *TP53* [Malkin et al., 1990] and *PTEN* [Liaw et al., 1997], respectively. These syndromes were also found in familial breast cancers, conferring increased risks of breast cancer [Malkin et al., 1990; Borresen et al., 1992]. Even though the exact associated risks of breast cancer due to germline mutations in *TP53* and *PTEN* genes are not certain, these are believed to exhibit moderate penetrance for breast cancer predisposition [Malkin et al., 1990; Liaw et al., 1997]. Subsequent germline mutational screening of cancer-related candidate genes identified 4 additional moderate-penetrance breast cancer predisposition genes, *CHEK2* [CHEK2 Breast Cancer Case-Control Consortium, 2004], *PALB2* [Rahman et al., 2007], *BRIP1* [Seal et al., 2006], and *ATM* [Renwick et al., 2006]. Germline mutations in these genes are also very rare in the general population and confer moderate risk for breast cancer predisposition. Together, these mutations were suggested to account for an additional 2.3% of genetic risk of breast cancer [Rahman et al., 2007].

The multiple high- and moderate-penetrance breast cancer predisposition genes identified thus far are rare in the general population and explain <25% of variations in the familial component of disease susceptibility. Further linkage studies did not yield additional *BRCA*-like genes conferring higher penetrance risk for breast cancer predisposition [Smith et al., 2006]. Search for moderate-penetrance genes through germline mutational screenings of cancer-related pathway genes was also not successful. The residual or missing heritability for breast cancer was hypothesized to be accounted for by the common disease-common variant (CDCV) hypothesis, which states that common diseases are caused by common variants [Chen et al., 2007; Hemminki et al., 2008]. According to the CDCV hypothesis, multiple common low-penetrance genes or alleles, either singly or in combination, confer breast cancer risk, conforming to a polygenic model of genetic inheritance. Under the polygenic model, each of the participating genes or alleles, also known as polygenes, has a small additive effect for breast cancer predisposition while linkage among loci and possible influence of environmental factors are ignored [Pharoah et al., 2002, 2008; Piccolo et al., 2008; Smith et al., 2008]. The CDCV hypothesis is the basis for genetic association studies conducted during the last 10 years, with an objective of characterizing additional predisposition risk factors for many complex diseases or traits, including breast cancer [Hindorff et al., 2009].

Genetic Association Studies

Genetic association studies are conducted to determine contributions of genetic variants to certain diseases or traits under study. The most commonly used strategies to evaluate genetic contributions to breast cancer in populations are case-control association studies, wherein frequencies of genetic variants in breast cancer cases are compared with those of healthy controls, and statistical significance of frequency differences is calculated; controls are free from breast cancer at the time of enrollment in such studies (fig. 1). As governed by the CDCV hypothesis, association studies largely rely on common genetic variants with population frequencies >5%, as statistically significant frequency differences between cases and controls at this cut-off are more easily demonstrable. SNPs are the most commonly used human genomic

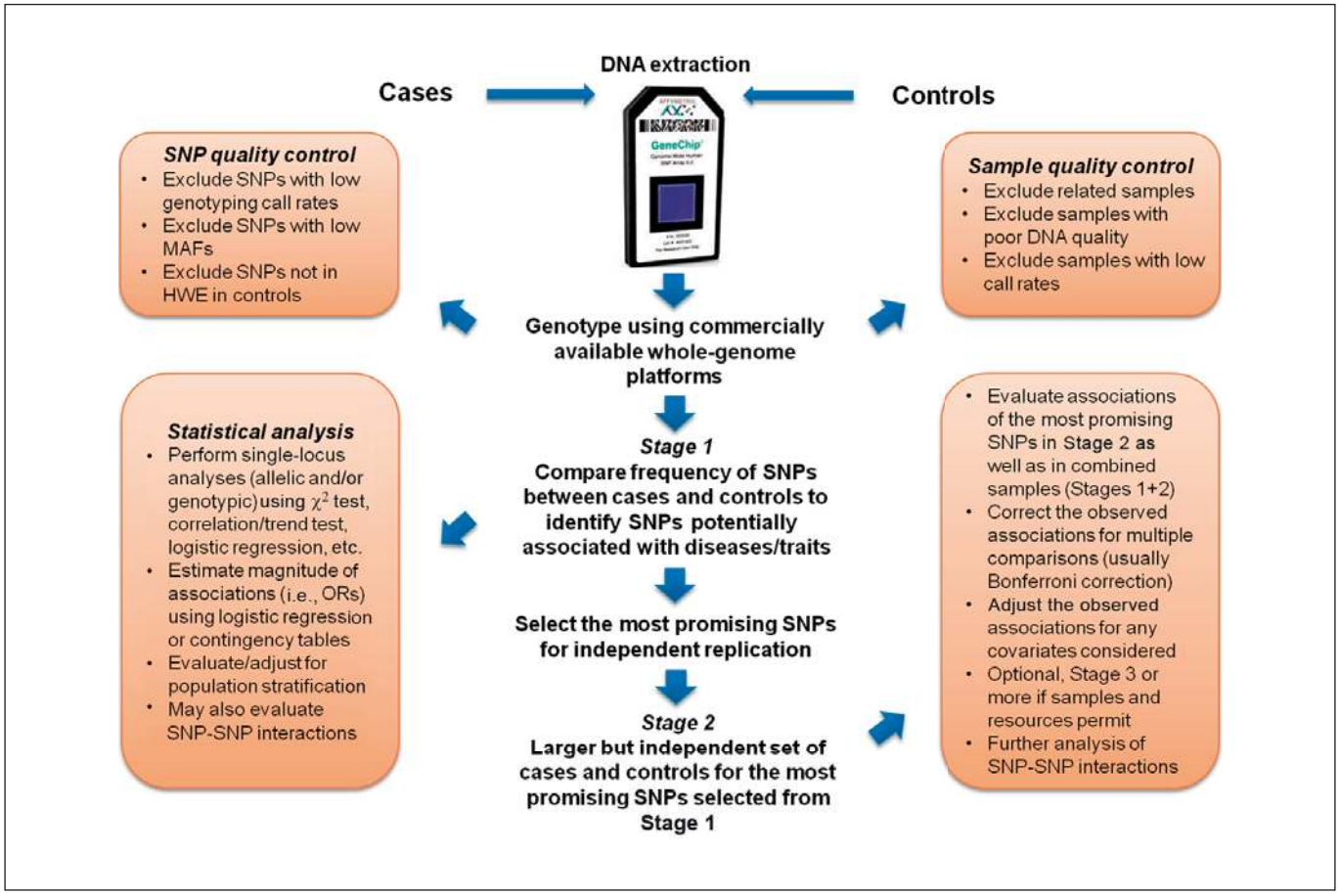

Fig. 1. A typical genetic case-control association study design and standard quality control and analysis strategies. HWE = Hardy-Weinberg equilibrium.

markers in genetic association studies to uncover associations of genetic loci in complex diseases or traits [Sachidanandam et al., 2001]. As of December 2013, the National Center for Biotechnology Information dbSNP Build 138 reported 44,278,189 validated SNPs in humans. Using SNPs as markers, 2 primary types of genetic association studies have been used to further characterize residual heritability for breast cancer predisposition: candidate gene association studies and GWASs.

Candidate Gene Association Studies

Initial genetic association studies aimed to uncover common variants for breast cancer predisposition were largely focused on genes involved in DNA repair processes, apoptosis, estrogen metabolism, inflammation, angiogenesis, and cell-cycle regulation, since these had plausible roles in breast cancer [Lipponen et al., 1994; Parshad et al., 1996; Bau et al., 2004; Kastan and Bartek, 2004; Pharoah et al., 2007; Stevens et al., 2007; Justenhoven et al., 2008; Schneider et al., 2008; Wang et al., 2008; Schonfeld et al., 2010]; such genes are also known as candidate genes. In candidate gene association studies, SNPs in or close to candidate genes are examined for their roles in breast cancer predisposition, governed by a hypothesis-driven approach. Several candidate gene association studies for breast cancer have been conducted during the last 10 years [Allen-Brady et al., 2006; Bewick et al., 2006; Haiman et al., 2008; Pooley et al., 2008; Sehl et al., 2009; Lin et al., 2011; Mangoni et al., 2011]; however, many of these studies were underpowered due to small sample size, resulting in inconsistent and irreproducible findings [Dunning et al., 1999; Pharoah et al., 2002]. To date, only 1 SNP (rs1045485) located in the coding region of *CASP8* identified by a candidate gene association study has shown promise as a breast cancer predisposition risk factor [Cox et al., 2007]. The minor allele of the coding SNP

conferred 12% less risk for breast cancer than the major allele [odds ratio (OR) = 0.88, 95% confidence interval (CI) = 0.84–0.92 and $p = 1.1 \times 10^{-7}$]. An independent case-control study also reported breast cancer risk reduction due to rs1045485 in *BRCA1* and *BRCA2* carriers [Palanca Suela et al., 2010]. However, the potential causal role of rs1045485 is still not clear.

GWASs

Advances in genotyping technologies, completion of the Human Genome Project, the International HapMap Project and the 1000 Genomes Project have led to the paradigm shift of genetic association studies from a limited candidate gene approach to a genome-wide approach, resulting in more detailed investigation of the CDCV hypothesis. GWAS premise relies on linkage disequilibrium (LD) among SNPs, which states that many neighboring SNPs are correlated and hence inherited together in a LD block [Reich et al., 2001]. Such correlations (LD) among nearby SNPs enable selection of fewer SNPs (tag SNPs) that essentially capture the information inherent to the block [Stram, 2004]. As such, by just genotyping 500,000–1,000,000 common tag SNPs (with population frequencies >5%), we could effectively capture the information content of >80% of all common SNPs in the human genome [International HapMap Consortium, 2003; Montpetit et al., 2006]. As of December 2012, >1,100 GWASs have identified >4,500 low-penetrance common SNPs associated with >700 different diseases or traits [Hindorff et al., 2009].

GWASs for Breast Cancer Predisposition

Multiple low-penetrance common breast cancer susceptibility SNPs have been identified by GWASs conducted by independent investigators and consortia. The first 3 breast cancer GWASs [Easton et al., 2007b; Hunter et al., 2007; Stacey et al., 2007] were conducted in early 2007 and identified multiple SNPs associated with breast cancer susceptibility. Easton et al. [2007b] interrogated a total of 227,876 SNPs arrayed in a custom-designed Perlegen platform in familial breast cancer cases and healthy controls from the United Kingdom and identified 5 breast cancer susceptibility loci. Of these, 4 SNPs are located in gene regions of *FGFR2, TNRC9, MAP3K1*, and *LSP1* while the fifth SNP is located in an intergenic region on chromosome 8q. Two additional novel breast cancer susceptibility loci on chromosomes 3p24 and 17q23.2 were also identified by further data mining and follow-up of the study reported by Easton et al. [2007b; Ahmed et al., 2009]. Subsequently, Stacey et al. [2007] genotyped ~300,000 SNPs arrayed in IlluminaHap300 platform in familial breast cancer cases and healthy controls from Iceland and identified 2 breast cancer susceptibility loci on chromosomes 2q35 and 16q12. In another breast cancer GWAS, Stacey et al. [2008] identified a novel breast cancer susceptibility locus on chromosome 5p12 using familial breast cancer cases and controls from Iceland. The association was stronger for breast cancer cases with estrogen receptor (ER)-positive than ER-negative tumors. *FGFR2* SNPs recently reported by Easton et al. [2007b] also showed stronger associations for ER-positive than ER-negative breast tumors [Stacey et al., 2008].

As part of the National Cancer Institute Cancer Genetics Markers of Susceptibility (CGEMS), Hunter et al. [2007] conducted a second breast cancer GWAS by genotyping 528,252 SNPs in postmenopausal sporadic breast cancer cases and healthy controls of European ancestry in IlluminaHapMap500 arrays. The findings from this GWAS confirmed association signals from intron 2 of the *FGFR2* gene reported by Easton et al. [2007b]. Similarly, Thomas et al. [2009] conducted a follow-up genetic association study of postmenopausal sporadic breast cancer GWAS reported by Hunter et al. [2007] that included study subjects of European ancestry and identified 2 additional breast cancer susceptibility loci on chromosomes 1p11.2 and 14q24.1. The reported loci also showed stronger associations in cases with ER-positive than ER-negative breast tumors. Results from this study also confirmed previous breast cancer susceptibility signals reported by Easton et al. [2007b], Hunter et al. [2007] and Stacey et al. [2007, 2008]. Fletcher et al. [2011] identified a novel breast cancer susceptibility locus on chromosome 9q31.2 by further data mining and follow-up of the postmenopausal sporadic breast cancer GWAS reported by Hunter et al. [2007].

Turnbull et al. [2010] conducted a breast cancer GWAS by genotyping familial breast cancer cases and healthy controls from the UK on Illumina Infinium 660K array and identified 5 additional breast cancer susceptibility loci not reported previously. This study also confirmed associations of breast cancer susceptibility loci found by Easton et al. [2007b], Hunter et al. [2007], Stacey et al. [2007, 2008], Thomas et al. [2009], Zheng et al. [2009], and Ahmed et al. [2009] in their study populations. Gold et al. [2008] interrogated 150,080 SNPs arrayed on Illumina GoldenGate platform in breast cancer cases and controls among non-*BRCA1/2* mutation carriers from a genetically isolated population, Ashkenazi Jews, and

identified a novel breast cancer susceptibility locus on chromosome 6q22.33. Antoniou et al. [2010] analyzed 620,601 SNPs on Illumina Infinium 610K array using *BRCA1* mutation carriers of European ancestry with and without breast cancer diagnoses and reported a novel breast cancer susceptibility locus on chromosome 19p13. The observed association was stronger in cases with triple-negative breast tumors. Gaudet et al. [2010] genotyped 592,163 SNPs on Affymetrix SNP 6.0 using *BRCA2* mutation carriers of European descent with and without breast cancer diagnoses and implicated a breast cancer susceptibility SNP in *FGFR2* intron 2, a locus reported and confirmed by many independent association studies. Haiman et al. [2011] evaluated 3,154,485 SNPs genotyped on Illumina 550-Duo SNP array and imputed in ER-negative breast cancer cases and healthy controls of European ancestry and reported a novel susceptibility locus for ER-negative breast cancer on chromosome 5p15. A meta-analysis of multiple breast cancer GWASs conducted by Siddiq et al. [2012] identified 2 novel breast cancer susceptibility loci on chromosomes 20q11 and 6q14 in women of European ancestry. Of these, rs2284378 on 20q11 showed stronger associations in ER-negative breast cancers as compared to ER-positive and breast cancer cases unselected for ER status.

Zheng et al. [2009] analyzed 607,728 SNPs genotyped on Affymetrix 500K and SNP 6.0 arrays using breast cancer cases and healthy controls of Chinese ancestry and established a breast cancer susceptibility locus on chromosome 6q25.1. An additional breast cancer susceptibility SNP on chromosome 6q25.1 for East-Asian women (Chinese, Japanese and Korean) was detected by Long et al. [2012] through data mining and follow-up of breast cancer GWAS reported by Zheng et al. [2009]. Further data mining and follow-up of breast cancer GWAS by Zheng et al. [2009] identified a potential causal breast cancer susceptibility SNP in Chinese women at a chromosomal locus (16q12.1) reported earlier by Stacey et al. [2007]. Cai et al. [2011] conducted breast cancer GWAS by genotyping breast cancer cases and healthy controls of Chinese ancestry on Affymetrix SNP 6.0 array and found a novel breast cancer susceptibility locus on chromosome 10q21.2 which contained a zinc finger protein encoded by the *ZNF365* gene. Kim et al. [2012] interrogated 555,525 SNPs genotyped on Affymetrix SNP 6.0 array using breast cancer cases and healthy controls of Korean ancestry and described a novel breast cancer susceptibility locus on chromosome 2q34, which contained the *ERBB4* gene. This study also successfully reproduced previously GWAS-identified breast cancer susceptibility loci reported by international consortia [Easton et al., 2007b; Hunter et al., 2007; Stacey et al., 2008; Ahmed et al., 2009; Turnbull et al., 2010].

Furthermore, Ghoussaini et al. [2012] combined data from multiple independent breast cancer GWASs and conducted a large-scale replication study that identified 3 novel breast cancer susceptibility loci on chromosomes 12p11, 12q24 and 21q21 in women with European ancestry. SNPs on 12q24 and 21q21 showed stronger associations for susceptibility to ER-positive than to ER-negative breast cancers, whereas a SNP on 12p11 conferred a similar risk for both ER-positive and ER-negative breast cancers. This was by far the largest GWAS conducted for breast cancer predisposition until early 2013. Subsequently, Orr et al. [2012] analyzed data for 447,760 SNPs genotyped in male breast cancer cases and healthy controls of European ancestry and identified a novel SNP located on chromosome 14q24.1, which contained the *RAD51B* gene, while providing supportive evidence for association of a SNP on 16q21.1 with male breast cancer risk. This was the first and only GWAS conducted to date for male breast cancer predisposition.

More recently, the Collaborative Oncological Gene-Environment Study (COGS) conducted by far the largest GWAS for breast cancer including >100,000 individuals of European ancestry and reported 41 novel breast cancer susceptibility loci located on chromosomes 1–14, 16, 18, 19, and 22 [Michailidou et al., 2013]. The study, an excellent collaborative model example of a consortium of consortia, interrogated 29,807 carefully selected candidate SNPs arrayed on a custom Illumina iSelect genotyping platform. These 41 newly identified breast cancer susceptibility loci alone explain about 5% of familial risk of breast cancer. The authors claimed that the overall contribution of SNPs to breast cancer susceptibility is ~18%, as suggested by the overall excess of significant associations of the COGS-selected SNPs [Michailidou et al., 2013]. With these estimations, high- and moderate-penetrance genes together with GWAS-identified low-penetrance common variants may explain about 50% of the familial risk of breast cancer [Michailidou et al., 2013]. Additionally, the COGS reported 4 variants located on chromosomes 1q32.1, 2p24.1 and 16q12.2 associated with ER-negative breast cancer [Garcia-Closas et al., 2013]; one of these variants on chromosome 1q32.1 was also associated with breast cancer risk in *BRCA1* mutation carriers [Couch et al., 2013]. A novel breast cancer susceptibility variant on chromosome 6p24 was also detected using *BRCA2* mutation carriers of European descent [Gaudet et al., 2013]. Besides, a comprehensive fine-

mapping study conducted by COGS led to the identification of 3 additional independent loci on chromosome 11q13 [French et al., 2013] and 2 novel breast cancer susceptibility loci on chromosome 5p15 [Bojesen et al., 2013]. Importantly, results from COGS indicated shared genetic susceptibility for breast, prostate and ovarian cancers, providing supporting evidence for a common genetic etiology in the development and progression of these hormone-related cancers. COGS reported 18 chromosomal regions containing risk variants associated with more than one of these cancers, and many of these regions harbor cancer-related genes such as *MDM4, TET2, TERT, KLF4, POU5F1B, RAD51B*, and *BABM1*. Of these, 2 genomic regions, 5p15.33 *(TERT)* and 8q24.21 *(MYC, POU5F1B)*, contained susceptibility loci for all 3 cancers. *TERT* plays an important role in making telomerase that maintains the telomere ends of chromosomes by addition of 6-nucleotide telomere repeats. *MYC* encodes a transcription factor that activates expression of many genes. It is a strong proto-oncogene and is involved in cell proliferation, cell growth and apoptosis.

A brief summary of breast cancer GWASs conducted between 2007–2013 was retrieved from the National Human Genome Research Institute catalog of published GWASs (http://www.genome.gov/gwastudies/) on July 11, 2014 [Hindorff et al., 2009] and is provided in online supplementary table 1 (see www.karger.com/doi/10.1159/000369045). Detailed description of the COGS findings would demand substantial space in this review article, and hence original research articles by the COGS [Bojesen et al., 2013; Couch et al., 2013; French et al., 2013; Garcia-Closas et al., 2013; Gaudet et al., 2013; Michailidou et al., 2013] and the GWAS catalog [Hindorff et al., 2009] are suggested for interested readers.

Common Variants for Breast Cancer Prognosis

Successes from GWASs in identifying low-penetrance common variants for breast cancer predisposition led to investigations examining potential roles of GWAS-identified breast cancer susceptibility SNPs for breast cancer prognosis. None of the breast cancer susceptibility loci reported by GWASs showed significant associations with breast cancer prognosis, except for a SNP (rs13281615) on chromosome 8q24 reported by Easton et al. [2007b], which showed statistically significant association with overall survival ($p = 0.009$) in an independent study comprising 13,527 invasive breast cancer cases [Garcia-Closas et al., 2008]. In 2010, Azzato et al. [2010a] conducted the first GWAS for breast cancer survival after diagnosis using the follow-up and genotype data of 528,252 SNPs for 1,145 postmenopausal sporadic breast cancer cases from the CGEMS initiative [Hunter et al., 2007]. However, the results did not find any SNPs statistically significantly associated with breast cancer prognosis. The authors concluded that a different set of low-penetrance common alleles, rather than susceptibility alleles, may be responsible for variations in breast cancer prognosis. In the same year, Azzato et al. [2010b] conducted a second GWAS for breast cancer prognosis by using existing stage 1 GWAS data from Easton et al. [2007b] that consisted of 3,761 invasive breast cancer cases genotyped for 10,621 SNPs on a custom-based Perlegen platform. The authors reported a SNP (rs4778137) on chromosome 15q13.1 as statistically significantly associated with breast cancer survival for triple-negative breast cancer cases ($p = 5.0 \times 10^{-5}$), and the association was successfully replicated in an independent set of 14,096 invasive breast cancer cases. Subsequently, a third GWAS for breast cancer prognosis was conducted by Shu et al. [2012] by interrogating 613,031 SNPs genotyped on Affymetrix SNP 6.0 array for 6,110 invasive breast cancer cases of Chinese ancestry. The results indicated 2 SNPs, rs3784099 and rs9934948, located on chromosomes 14 and 16, respectively, as significantly associated with breast cancer survival ($p < 5.0 \times 10^{-6}$). Recently, a multi-stage association study conducted by us using a combined sample size of 7,219 breast cancer cases and healthy controls of Caucasian origin from Alberta, Canada, confirmed a potential prognostic value of SNP rs13280615 on chromosome 8q24.21 for breast cancer [Sehrawat et al., 2011; Sapkota et al., 2013c]. The SNP was for the first time shown to be of prognostic value with breast cancer outcomes (recurrence-free and overall survival) by independent investigators [Garcia-Closas et al., 2008]. In GWAS literature, corroborative evidence of this kind is highly recommended in diverse ethnic groups. Even though the prognostic SNP replicated here in this study is also from Caucasian subjects, its replication in other ethnic groups is awaited. More recently, Rafiq et al. [2013] conducted a 2-stage association study using breast cancer cases of European ancestry and reported suggestive association of a SNP near the *ARRDC3* gene with breast cancer prognosis ($p = 9.5 \times 10^{-7}$). Overall, these studies have shown that inherited germline genetic variations may contribute to the observed variations in breast cancer prognosis, and hence, larger GWASs in future are warranted to identify additional common variants for breast cancer outcomes.

Summary of Genetic Risk Accounted to Date and the Search for Missing Heritability

Over the last few years, several GWASs and a candidate gene association study conducted for breast cancer led to the identification of multiple low-penetrance common variants conferring single-locus effects for breast cancer risk, lending credence to the CDCV hypothesis and the polygenic model of risk for complex diseases [Pharoah et al., 2002; Cox et al., 2007; Hemminki et al., 2008]. The majority of breast cancer-associated SNPs identified thus far are common in the study population with minor allele frequency >10% [Cox et al., 2007; Easton et al., 2007b; Hunter et al., 2007; Stacey et al., 2007, 2008; Gold et al., 2008; Ahmed et al., 2009; Thomas et al., 2009; Zheng et al., 2009; Antoniou et al., 2010; Gaudet et al., 2010, 2013; Long et al., 2010, 2012; Turnbull et al., 2010; Cai et al., 2011; Fletcher et al., 2011; Haiman et al., 2011; Ghoussaini et al., 2012; Siddiq et al., 2012; Bojesen et al., 2013; Couch et al., 2013; French et al., 2013; Garcia-Closas et al., 2013; Michailidou et al., 2013]. However, the effect sizes of these associations are modest and explain ~30% of additional genetic risk for breast cancer predisposition [Ghoussaini et al., 2012; Michailidou et al., 2013]. Taken together, known high- and moderate-penetrance genes identified through linkage studies and mutational screenings of candidate genes, in addition to recently identified low-penetrance common SNPs by genetic association studies, only account for <50% of variations in familial breast cancer predisposition [Michailidou et al., 2013], suggesting that more variants exist.

One of the major challenges to uncover the remainder of breast cancer heritability is the sample size needed for sufficient statistical power since the (yet unidentified) common variants are expected to confer much smaller effect sizes (ORs <1.5). International consortia such as the Breast Cancer Association Consortium (BCAC), the CGEMS, the Breast and Prostate Cancer Cohort Consortia, and the COGS have already made an effort to increase the number of studied individuals to >100,000 by including breast cancer cases and healthy controls from several individual research centers and consortia [Ghoussaini et al., 2012; Michailidou et al., 2013]. However, results from such giant consortia are also limited to SNPs with very small effect sizes, indicating that future GWASs are unlikely to identify common variants with very large individual effect sizes (ORs >1.5), regardless of sufficiently large sample sizes. Consequently, there is a clear need to explore other forms of genetic variations contributing to breast cancer predisposition. Current debates suggest that one of the possible sources for 'residual or missing heritability' for breast cancer are contributions of structural variations, such as CNVs, and genetic interactions (gene-gene and gene-environment interactions), in addition to the contributions from strong single-locus effects through continued efforts for sufficiently powered systematic GWASs [Moore, 2003, 2005; Manolio et al., 2009; Anonymous, 2010; Bodmer and Tomlinson, 2010; Eichler et al., 2010; Gibson, 2010; Thomas, 2010; Zuk et al., 2012]. Considering recent advances in next-generation sequencing technologies, possible involvement of yet 'unknown' high-penetrance mutations/genes contributing to breast cancer susceptibility may also be investigated through whole-genome and whole-exome studies. At present, extreme cases with a strong family history may be ascertained for these studies, as whole genome/exome sequencing of a large number of breast cancer cases and controls is still a cost-limiting factor. Such a comprehensive approach may identify a larger proportion of breast cancer heritability, leading to possibilities for population level screening and prophylactic interventions in the near future.

CNVs

One possible source of residual heritability for breast cancer is the contribution of CNVs. CNVs are the most common type of structural variations in the human genome. These are DNA segments >1 kb in size that vary in their copy numbers due to gains or losses (fig. 2) [Redon et al., 2006; Henrichsen et al., 2009; Kuiper et al., 2010]. Throughout this review, we refer to results from germline CNVs (and not CNVs from tumor cells/somatic origins) as is our focus with SNPs for their potential value in disease susceptibility or prognosis. As of January 2014, there were 2,304,349 CNVs reported in the Database of Genomic Variants, Toronto, Canada (http://dgv.tcag.ca/dgv/app/home), a curated catalog of human genomic structural variation, and more CNVs may be identified in the coming years. This catalog is by no means a complete database, but is continually evolving. CNVs are believed to affect expression of many genes, either through gene dosage (gains or losses) or by *cis*-acting regulatory activities [Locke et al., 2006; Stranger et al., 2007; Henrichsen et al., 2009]. Studies have shown that germline CNVs may predispose to many complex diseases, and SNPs are generally underrepresented in genomic regions harboring CNVs and, therefore, GWASs utilizing CNVs are slowly emerging [Tuzun et al., 2005; McCarroll and Altshuler,

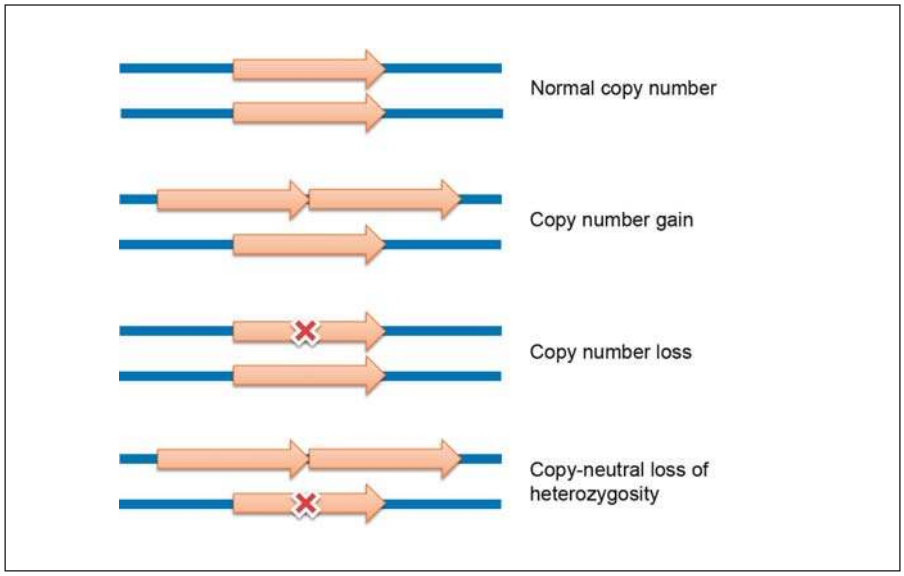

Fig. 2. Illustration of copy number variations and copy-neutral loss of heterozygosity.

2007; Wong et al., 2007; McCarroll et al., 2008; Shlien and Malkin, 2010].

The first large-scale GWAS investigating potential roles of CNVs for genetic susceptibility to complex diseases was conducted by the Wellcome Trust Case Control Consortium [2010]. The study investigated 3,432 common CNVs for their roles in 7 diseases and found no strong evidence for CNVs as better (than SNPs) sources of residual heritability for complex diseases. Even though the study reported lack of associations of common CNVs with complex diseases, perhaps due to a limited number of common CNVs considered in the analysis, the study certainly opened possibilities for a more systematic and comprehensive analysis (larger coverage) of germline CNVs as predisposition factors for complex diseases. Another study that evaluated germline CNV profiles between *BRCA1*-associated and sporadic ovarian cancer patients reported substantial differences in copy number gains and losses between these 2 groups of cancer patients [Yoshihara et al., 2011]. Germline CNVs were also reported to be associated with susceptibility to familial pancreatic and breast cancers [Al-Sukhni et al., 2012].

More recently, with the application of high-throughput SNP genotyping arrays, large chromosomal lesions characterized by loss of heterozygosity (LOH), but with diploid copy numbers, were observed (as were also in many malignancies), possibly resulting from non-homologous recombination during meiosis, trisomic rescue or mitotic recombination [Mohamedali et al., 2007; O'Keefe et al., 2010; Kryh et al., 2011; Lapunzina and Monk, 2011; Melcher et al., 2011; Saeki et al., 2011]. These chromosomal defects, also known as CN-LOHs or uniparental disomies, are characterized by loss of 1 allele with simultaneous replacement by an exact copy of another allele, resulting in retention of diploid copy number but loss of polymorphic differences (both alleles are from the same parent) (fig. 2). CN-LOHs have been reported to be associated with gain of oncogenic alleles and inactivation of tumor suppressors and may be an important mechanism in cancer development [Mohamedali et al., 2007; O'Keefe et al., 2010; Melcher et al., 2011; Saeki et al., 2011]. With the advent of SNP genotyping platforms that can measure both CNVs and CN-LOHs, it is now possible to investigate potential roles of these large chromosomal defects as genetic determinants for complex diseases using germline DNA. A recent study conducted by us that aimed to evaluate the role of germline CNVs and CN-LOHs in breast cancer recurrence identified multiple copy number aberrations (2 copy number gains and 5 CN-LOHs) showing statistically significant differences between recurrence-free survival probabilities with and without these aberrations [Sapkota et al., 2013a]. Of these, 3 CN-LOHs were successfully validated by qPCR. Even though these results need to be further investigated in independent breast cancer cases to confirm robustness of their associations, findings from this first study of its kind clearly indicate the importance of germline DNA variations as potential prognostic markers for breast cancer.

Genetic Interactions

GWASs representing common SNPs and emerging studies of CNVs or CN-LOHs primarily focus on single-locus effects (also known as main genetic effects). However, the risk for complex diseases, including breast cancer, is also attributed to 2 types of genetic interactions: gene-gene and gene-environment interactions. At present, GWASs including these genetic interactions are limited because of the need for large sample sizes to achieve the statistical power, as well as the exposure data (health, lifestyle and reproductive) that are difficult to obtain and, where available, may not have banked DNA in most cohorts.

Gene-Gene Interactions

The etiology of complex diseases includes a substantial proportion of gene-gene interactions, commonly referred to as epistasis. Epistasis is a ubiquitous phenomenon that describes how genes or loci interact to affect phenotypes [Moore, 2005]. Such interactions are believed to explain a large proportion of genetic heritability of complex diseases. At present, epistatic interactions involving 2 loci or SNPs can be evaluated using logistic regressions. However, it has been limited to candidate gene studies with a small number of SNPs [Onay et al., 2006]. Testing every combination of pair-wise interactions or even extending to multi-SNPs (multi-way) interactions in a GWAS is computationally intensive. Several strategies can be adopted to reduce the number of tests that may arise from a brute-force approach, which tests all pairs of interactions. One approach that we employed in a recent study would be to test for interactions among SNPs that show weak but consistent associations in at least 2 independent stages [Sapkota et al., 2012, 2013b]. This would dramatically reduce the number of SNPs to include in epistatic interaction analyses (i.e. multiple testing burden) while excluding thousands of SNPs that are not obvious candidates for the interaction analyses. More recently, logic regressions have been proposed for testing multi-way interactions among SNPs and have been successfully applied to GWAS of Crohn's disease and a candidate gene association study of cervical cancer [Feng et al., 2005; Dinu et al., 2012]. Future studies that focus on potential epistatic interactions among SNPs, in addition to single-locus effects, in a GWAS or a candidate gene association study may identify additional heritability for complex diseases, including breast cancer.

Gene-Environment Interactions

Complex diseases such as breast cancer also result from combined effects of both genetic and environmental risk factors. Even though these forms of genetic interactions are believed to explain a large proportion of heritability for breast cancer, especially for sporadic breast cancer, investigations of such interactions are limited due to difficulty in obtaining the environmental, lifestyle and reproductive data.

The Breast and Prostate Cancer Cohort Consortium conducted a comprehensive study to evaluate possible interactions between the common breast cancer susceptibility loci and the established breast cancer risk factors, such as age at menarche, parity, age at menopause, use of hormone replacement therapy, body mass index, smoking habit, alcohol consumption, family history, and height, using data from 8,576 cases and 11,892 controls [Campa et al., 2011]. The study findings indicated that the common breast cancer susceptibility loci do not affect the associations of the examined established risk factors with breast cancer. These findings were also supported by another study conducted by the BCAC that evaluated possible interactions among common breast cancer susceptibility loci and known breast cancer risk factors using genotype and questionnaire data from 26,349 cases and 32,208 controls from 21 case-control studies [Milne et al., 2010]. More recently, COGS assessed 23 known breast cancer susceptibility loci and their potential interactions with 10 established environmental risk factors and provided first strong evidence for an important role of gene-environmental interactions in breast cancer susceptibility [Nickels et al., 2013]. Statistically significant associations were observed for interactions between *LSP1* variant (rs3817198) and parity and between *CASP8* variant (rs17468277) and alcohol consumption. Once sufficient exposure data becomes readily accessible to incorporate in GWASs, future studies may also focus on this form of genetic interaction to address the residual heritability of breast cancer.

Risk Prediction Models for Breast Cancer

In general, risk assessment of breast cancer has been broadly grouped into 2 categories – empirical models and genetic risk models [Amir et al., 2010]. Empirical models estimate the probability of carrying a mutation (such as *BRCA1/2*) without making any explicit assumptions about the underlying genetic architecture. Myriad II [Frank et al., 2002] and the Manchester model [Evans et

al., 2004] are examples of empirical models. On the contrary, genetic risk models make explicit assumptions about the genetic risks, such as the number of susceptibility variants, their allele frequencies in the general population and the risks conferred by these alleles. Examples of genetic risk models include BRCAPRO [Berry et al., 2002], The Breast and Ovarian Analysis of Disease Incidence and Carrier Estimation Algorithm (BOADICEA) [Antoniou et al., 2008a] and the International Breast Cancer Intervention Study (IBIS) [Tyrer et al., 2004]. Both empirical and genetic risk models have good discriminatory accuracy, as measured by the area under the curve (AUC), for predicting *BRCA1/2* mutation status. The AUC values for BOADICEA, BRCAPRO, IBIS, Myriad II, and the Manchester model were 0.77, 0.76, 0.74, 0.75, and 0.72, respectively [Antoniou et al., 2008b]. For a more accurate assessment of breast cancer risks over time, additional 'known' risk factors such as age, obesity, mammography density, family history of breast cancer, hormonal and reproductive factors have been incorporated into the risk assessment models. Even though there are many risk prediction models for breast cancer, the Gail model is the mainstream model currently in use [Gail et al., 1989] and is the only model validated in 3 large population-based databases [Bondy et al., 1994; Spiegelman et al., 1994; Costantino et al., 1999]. The Gail model uses a number of breast cancer risk factors, namely age, age at menarche, age at first live birth (or nulliparity), number of first-degree relatives with breast cancer, number of prior breast biopsies, and presence of atypical hyperplasia [Gail et al., 1989]. In its present form, the Gail model offers the discriminatory accuracy at ~0.5–0.6 [Rockhill et al., 2001; Tice et al., 2005; Bondy and Newman, 2006], which is slightly better than the baseline, indicating that additional risk factors need to be identified and incorporated into the Gail model to improve its precision.

In 2010, Wacholder et al. for the first time attempted to assess the disease risk conferred by common breast cancer susceptibility variants [Wacholder et al., 2010]. The authors demonstrated that inclusion of 10 common breast cancer susceptibility variants identified through GWASs into the widely used Gail model moderately improved the performance of risk models for breast cancer from 58.0 to 61.8%, as measured by the AUC. While this scant improvement in risk assessment may not be sufficient for inclusion of common variants to identify women who might benefit from prophylactic intervention, it is possible that many more variants remain to be identified that could eventually improve clinical risk assessment for breast cancer. More recently, Sawyer et al. [2012] evaluated the associated familial breast cancer risk conferred by 22 common breast cancer susceptibility variants identified through multiple GWASs using a polygenic risk score (PRS), calculated as the sum of the log OR for each allele. Using PRS for risk assessment, the 22 common variants could explain 18.5% of the genetic risk for breast cancer, while the predictive power of PRS in non-*BRCA1/2* familial breast cancer cases was 65.4%, as measured by AUC. These results also indicated that PRS was significantly higher among individuals with familial breast cancer than in healthy controls ($p = 1.0 \times 10^{-16}$). Moreover, the PRS was significantly higher among familial cases without *BRCA1/2* mutations than among cases with mutation carriers ($p = 2.3 \times 10^{-6}$). Women who tested negative for *BRCA1/2* mutations but had higher PRS were more likely to have an early onset of disease before 30 years of age (OR = 3.37) and a higher chance of second breast cancer (OR = 1.96) as compared to women with low PRS. Presently, the current model of genetic testing for familial breast cancer only identifies *BRCA1/2* mutations in ~1 in 5 women. The test is uninformative for familial cases that test negative for *BRCA1/2* mutations. However, after addition of common variants (i.e. PRS) into the current model of genetic testing, it may be now possible to subdivide non-*BRCA1/2* familial breast cancer cases into high-, intermediate- and low-risk groups, as described by these investigators. A similar model of genetic testing, also taking into account newly identified variants by COGS and genetic interactions, can be attempted for risk assessment of sporadic breast cancer when sufficient common variants will be identified through more GWASs and candidate gene association studies aided by large international consortia.

Concluding Remarks

Identification of a large number of breast cancer predisposition factors (usually common SNPs), with either single-locus or epistatic effects, could be of use for breast cancer risk assessment. The combined effect of informative genetic risk factors into the Gail model may increase the risk prediction accuracy and eventually may allow for development of population-based risk screening and stratification programs.

Furthermore, few studies conducted to date have indicated that germline DNA variations (SNPs and copy number aberrations) may serve as potential prognostic markers for breast cancer. Clearly, more research is needed to identify additional germline variants of potential

prognostic value. The most promising markers that show consistent statistically significant associations with breast cancer prognosis can be further evaluated in prospective clinical trials. If successful, these germline markers, in addition to currently utilized tumor-based prognostic and predictive factors, may help us realize the practical value of breast cancer prevention and control through applications of genetically stratified populations to benefit from emerging genomics medicine.

Acknowledgements

I would like to thank Dr. Sambasivarao Damaraju and Professors Carol E. Cass, Raymond Lai, John R. Mackey, and Yutaka Yasui for their useful suggestions and feedback. Funding to pay publication charges for this article was provided by QIMR Berghofer Medical Research Institute, Brisbane, Qld., Australia.

References

Ahmed S, Thomas G, Ghoussaini M, Healey CS, Humphreys MK, et al: Newly discovered breast cancer susceptibility loci on 3p24 and 17q23.2. Nat Genet 41:585–590 (2009).

Allen-Brady K, Cannon-Albright LA, Neuhausen SL, Camp NJ: A role for *XRCC4* in age at diagnosis and breast cancer risk. Cancer Epidemiol Biomarkers Prev 15:1306–1310 (2006).

Al-Sukhni W, Joe S, Lionel AC, Zwingerman N, Zogopoulos G, et al: Identification of germline genomic copy number variation in familial pancreatic cancer. Hum Genet 131:1481–1494 (2012).

Amir E, Freedman OC, Seruga B, Evans DG: Assessing women at high risk of breast cancer: a review of risk assessment models. J Natl Cancer Inst 102:680–691 (2010).

Anonymous: On beyond GWAS (editorial). Nat Genet 42:551 (2010).

Antoniou AC, Gayther SA, Stratton JF, Ponder BA, Easton DF: Risk models for familial ovarian and breast cancer. Genet Epidemiol 18: 173–190 (2000).

Antoniou AC, Pharoah PD, McMullan G, Day NE, Stratton MR, et al: A comprehensive model for familial breast cancer incorporating *BRCA1, BRCA2* and other genes. Br J Cancer 86:76–83 (2002).

Antoniou AC, Cunningham AP, Peto J, Evans DG, Lalloo F, et al: The BOADICEA model of genetic susceptibility to breast and ovarian cancers: updates and extensions. Br J Cancer 98:1457–1466 (2008a).

Antoniou AC, Hardy R, Walker L, Evans DG, Shenton A, et al: Predicting the likelihood of carrying a *BRCA1* or *BRCA2* mutation: validation of BOADICEA, BRCAPRO, IBIS, Myriad and the Manchester scoring system using data from UK genetics clinics. J Med Genet 45:425–431 (2008b).

Antoniou AC, Wang X, Fredericksen ZS, McGuffog L, Tarrell R, et al: A locus on 19p13 modifies risk of breast cancer in *BRCA1* mutation carriers and is associated with hormone receptor-negative breast cancer in the general population. Nat Genet 42:885–892 (2010).

Azzato EM, Pharoah PD, Harrington P, Easton DF, Greenberg D, et al: A genome-wide association study of prognosis in breast cancer. Cancer Epidemiol Biomarkers Prev 19:1140–1143 (2010a).

Azzato EM, Tyrer J, Fasching PA, Beckmann MW, Ekici AB, et al: Association between a germline *OCA2* polymorphism at chromosome 15q13.1 and estrogen receptor-negative breast cancer survival. J Natl Cancer Inst 102: 650–662 (2010b).

Bau DT, Fu YP, Chen ST, Cheng TC, Yu JC, et al: Breast cancer risk and the DNA double-strand break end-joining capacity of nonhomologous end-joining genes are affected by *BRCA1*. Cancer Res 64:5013–5019 (2004).

Berry DA, Iversen ES Jr, Gudbjartsson DF, Hiller EH, Garber JE, et al: BRCAPRO validation, sensitivity of genetic testing of *BRCA1/BRCA2*, and prevalence of other breast cancer susceptibility genes. J Clin Oncol 20:2701–2712 (2002).

Bewick MA, Conlon MS, Lafrenie RM: Polymorphisms in *XRCC1, XRCC3*, and *CCND1* and survival after treatment for metastatic breast cancer. J Clin Oncol 24:5645–5651 (2006).

Bodmer W, Tomlinson I: Rare genetic variants and the risk of cancer. Curr Opin Genet Dev 20:262–267 (2010).

Bojesen SE, Pooley KA, Johnatty SE, Beesley J, Michailidou K, et al: Multiple independent variants at the *TERT* locus are associated with telomere length and risks of breast and ovarian cancer. Nat Genet 45:371e2–384e2 (2013).

Bondy ML, Newman LA: Assessing breast cancer risk: evolution of the Gail model. J Natl Cancer Inst 98:1172–1173 (2006).

Bondy ML, Lustbader ED, Halabi S, Ross E, Vogel VG: Validation of a breast cancer risk assessment model in women with a positive family history. J Natl Cancer Inst 86:620–625 (1994).

Borresen AL, Andersen TI, Garber J, Barbier-Piraux N, Thorlacius S, et al: Screening for germ line *TP53* mutations in breast cancer patients. Cancer Res 52:3234–3236 (1992).

Cai Q, Long J, Lu W, Qu S, Wen W, et al: Genome-wide association study identifies breast cancer risk variant at 10q21.2: results from the Asia Breast Cancer Consortium. Hum Mol Genet 20:4991–4999 (2011).

Campa D, Kaaks R, Le Marchand L, Haiman CA, Travis RC, et al: Interactions between genetic variants and breast cancer risk factors in the breast and prostate cancer cohort consortium. J Natl Cancer Inst 103:1252–1263 (2011).

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 74: 1175–1182 (2004).

Chen GK, Jorgenson E, Witte JS: An empirical evaluation of the common disease-common variant hypothesis. BMC Proc 1, Suppl 1:S5 (2007).

Collaborative Group on Hormonal Factors in Breast Cancer: Familial breast cancer: collaborative reanalysis of individual data from 52 epidemiological studies including 58,209 women with breast cancer and 101,986 women without the disease. Lancet 358:1389–1399 (2001).

Costantino JP, Gail MH, Pee D, Anderson S, Redmond CK, et al: Validation studies for models projecting the risk of invasive and total breast cancer incidence. J Natl Cancer Inst 91:1541–1548 (1999).

Couch FJ, Wang X, McGuffog L, Lee A, Olswold C, et al: Genome-wide association study in *BRCA1* mutation carriers identifies novel loci associated with breast and ovarian cancer risk. PLoS Genet 9:e1003212 (2013).

Cox A, Dunning AM, Garcia-Closas M, Balasubramanian S, Reed MW, et al: A common coding variant in *CASP8* is associated with breast cancer risk. Nat Genet 39:352–358 (2007).

Dinu I, Mahasirimongkol S, Liu Q, Yanai H, Sharaf Eldin N, et al: SNP-SNP interactions discovered by logic regression explain Crohn's disease genetics. PLoS One 7:e43035 (2012).

Dunning AM, Healey CS, Pharoah PD, Teare MD, Ponder BA, Easton DF: A systematic review of genetic polymorphisms and breast cancer risk. Cancer Epidemiol Biomarkers Prev 8:843–854 (1999).

Easton DF: How many more breast cancer predisposition genes are there? Breast Cancer Res 1: 14–17 (1999).

Easton DF, Bishop DT, Ford D, Crockford GP: Genetic linkage analysis in familial breast and ovarian cancer: results from 214 families. The Breast Cancer Linkage Consortium. Am J Hum Genet 52:678–701 (1993).

Easton DF, Ford D, Bishop DT: Breast and ovarian cancer incidence in BRCA1-mutation carriers. Breast Cancer Linkage Consortium. Am J Hum Genet 56:265–271 (1995).

Easton DF, Deffenbaugh AM, Pruss D, Frye C, Wenstrup RJ, et al: A systematic genetic assessment of 1,433 sequence variants of unknown clinical significance in the *BRCA1* and *BRCA2* breast cancer-predisposition genes. Am J Hum Genet 81:873–883 (2007a).

Easton DF, Pooley KA, Dunning AM, Pharoah PD, Thompson D, et al: Genome-wide association study identifies novel breast cancer susceptibility loci. Nature 447:1087–1093 (2007b).

Economopoulos KP, Sergentanis TN: XRCC3 Thr241Met polymorphism and breast cancer risk: a meta-analysis. Breast Cancer Res Treat 121:439–443 (2010).

Eichler EE, Flint J, Gibson G, Kong A, Leal SM, et al: Missing heritability and strategies for finding the underlying causes of complex disease. Nat Rev Genet 11:446–450 (2010).

Evans DG, Eccles DM, Rahman N, Young K, Bulman M, et al: A new scoring system for the chances of identifying a *BRCA1/2* mutation outperforms existing models including BRCAPRO. J Med Genet 41:474–480 (2004).

Feng Q, Balasubramanian A, Hawes SE, Toure P, Sow PS, et al: Detection of hypermethylated genes in women with and without cervical neoplasia. J Natl Cancer Inst 97:273–282 (2005).

Fletcher O, Johnson N, Orr N, Hosking FJ, Gibson LJ, et al: Novel breast cancer susceptibility locus at 9q31.2: results of a genome-wide association study. J Natl Cancer Inst 103:425–435 (2011).

Ford D, Easton DF, Bishop DT, Narod SA, Goldgar DE: Risks of cancer in *BRCA1*-mutation carriers. Breast Cancer Linkage Consortium. Lancet 343:692–695 (1994).

Ford D, Easton DF, Stratton M, Narod S, Goldgar D, et al: Genetic heterogeneity and penetrance analysis of the *BRCA1* and *BRCA2* genes in breast cancer families. The Breast Cancer Linkage Consortium. Am J Hum Genet 62:676–689 (1998).

Frank TS, Deffenbaugh AM, Reid JE, Hulick M, Ward BE, et al: Clinical characteristics of individuals with germline mutations in *BRCA1* and *BRCA2*: analysis of 10,000 individuals. J Clin Oncol 20:1480–1490 (2002).

French JD, Ghoussaini M, Edwards SL, Meyer KB, Michailidou K, et al: Functional variants at the 11q13 risk locus for breast cancer regulate cyclin D1 expression through long-range enhancers. Am J Hum Genet 92:489–503 (2013).

Gail MH, Brinton LA, Byar DP, Corle DK, Green SB, et al: Projecting individualized probabilities of developing breast cancer for white females who are being examined annually. J Natl Cancer Inst 81:1879–1886 (1989).

Garcia-Closas M, Hall P, Nevanlinna H, Pooley K, Morrison J, et al: Heterogeneity of breast cancer associations with five susceptibility loci by clinical and pathological characteristics. PLoS Genet 4:e1000054 (2008).

Garcia-Closas M, Couch FJ, Lindstrom S, Michailidou K, Schmidt MK, et al: Genome-wide association studies identify four ER negative-specific breast cancer risk loci. Nat Genet 45:392–398, 398e1–2 (2013).

Gaudet MM, Kirchhoff T, Green T, Vijai J, Korn JM, et al: Common genetic variants and modification of penetrance of *BRCA2*-associated breast cancer. PLoS Genet 6:e1001183 (2010).

Gaudet MM, Kuchenbaecker KB, Vijai J, Klein RJ, Kirchhoff T, et al: Identification of a *BRCA2*-specific modifier locus at 6p24 related to breast cancer risk. PLoS Genet 9:e1003173 (2013).

Ghoussaini M, Fletcher O, Michailidou K, Turnbull C, Schmidt MK, et al: Genome-wide association analysis identifies three new breast cancer susceptibility loci. Nat Genet 44:312–318 (2012).

Gibson G: Hints of hidden heritability in GWAS. Nat Genet 42:558–560 (2010).

Gold B, Kirchhoff T, Stefanov S, Lautenberger J, Viale A, et al: Genome-wide association study provides evidence for a breast cancer risk locus at 6q22.33. Proc Natl Acad Sci USA 105:4340–4345 (2008).

Haiman CA, Hsu C, de Bakker PI, Frasco M, Sheng X, et al: Comprehensive association testing of common genetic variation in DNA repair pathway genes in relationship with breast cancer risk in multiple populations. Hum Mol Genet 17:825–834 (2008).

Haiman CA, Chen GK, Vachon CM, Canzian F, Dunning A, et al: A common variant at the *TERT-CLPTM1L* locus is associated with estrogen receptor-negative breast cancer. Nat Genet 43:1210–1214 (2011).

Hall JM, Lee MK, Newman B, Morrow JE, Anderson LA, et al: Linkage of early-onset familial breast cancer to chromosome 17q21. Science 250:1684–1689 (1990).

Hemminki K, Försti A, Bermejo JL: The 'common disease-common variant' hypothesis and familial risks. PLoS One 3:e2504 (2008).

Henrichsen CN, Chaignat E, Reymond A: Copy number variants, diseases and gene expression. Hum Mol Genet 18:R1–R8 (2009).

Hindorff LA, MacArthur J, Morales J, Junkins HA, Hall PN, et al: A Catalog of Published Genome-Wide Association Studies (2009). http://www.genome.gov/gwastudies (accessed July 11, 2014).

Hunter DJ, Kraft P, Jacobs KB, Cox DG, Yeager M, et al: A genome-wide association study identifies alleles in *FGFR2* associated with risk of sporadic postmenopausal breast cancer. Nat Genet 39:870–874 (2007).

International HapMap Consortium: The International HapMap Project. Nature 426:789–796 (2003).

Justenhoven C, Hamann U, Schubert F, Zapatka M, Pierl CB, et al: Breast cancer: a candidate gene approach across the estrogen metabolic pathway. Breast Cancer Res Treat 108:137–149 (2008).

Kastan MB, Bartek J: Cell-cycle checkpoints and cancer. Nature 432:316–323 (2004).

Key TJ, Verkasalo PK, Banks E: Epidemiology of breast cancer. Lancet Oncol 2:133–140 (2001).

Kim HC, Lee JY, Sung H, Choi JY, Park SK, et al: A genome-wide association study identifies a breast cancer risk variant in *ERBB4* at 2q34: results from the Seoul Breast Cancer Study. Breast Cancer Res 14:R56 (2012).

Kryh H, Caren H, Erichsen J, Sjöberg RM, Abrahamsson J, et al: Comprehensive SNP array study of frequently used neuroblastoma cell lines; copy neutral loss of heterozygosity is common in the cell lines but uncommon in primary tumors. BMC Genomics 12:443 (2011).

Kuiper RP, Ligtenberg MJ, Hoogerbrugge N, Geurts van Kessel A: Germline copy number variation and cancer risk. Curr Opin Genet Dev 20:282–289 (2010).

Lapunzina P, Monk D: The consequences of uniparental disomy and copy number neutral loss-of-heterozygosity during human development and cancer. Biol Cell 103:303–317 (2011).

Liaw D, Marsh DJ, Li J, Dahia PL, Wang SI, et al: Germline mutations of the *PTEN* gene in Cowden disease, an inherited breast and thyroid cancer syndrome. Nat Genet 16:64–67 (1997).

Lichtenstein P, Holm NV, Verkasalo PK, Iliadou A, Kaprio J, et al: Environmental and heritable factors in the causation of cancer – analyses of cohorts of twins from Sweden, Denmark, and Finland. New Engl J Med 343:78–85 (2000).

Lin WY, Camp NJ, Cannon-Albright LA, Allen-Brady K, Balasubramanian S, et al: A role for *XRCC2* gene polymorphisms in breast cancer risk and survival. J Med Genet 48:477–484 (2011).

Lipponen P, Aaltomaa S, Kosma VM, Syrjänen K: Apoptosis in breast cancer as related to histopathological characteristics and prognosis. Eur J Cancer 30A:2068–2073 (1994).

Locke DP, Sharp AJ, McCarroll SA, McGrath SD, Newman TL, et al: Linkage disequilibrium and heritability of copy-number polymorphisms within duplicated regions of the human genome. Am J Hum Genet 79:275–290 (2006).

Long J, Cai Q, Shu XO, Qu S, Li C, et al: Identification of a functional genetic variant at 16q12.1 for breast cancer risk: results from the Asia Breast Cancer Consortium. PLoS Genet 6:e1001002 (2010).

Long J, Cai Q, Sung H, Shi J, Zhang B, et al: Genome-wide association study in East Asians identifies novel susceptibility loci for breast cancer. PLoS Genet 8:e1002532 (2012).
Malkin D, Li FP, Strong LC, Fraumeni JF Jr, Nelson CE, et al: Germ line p53 mutations in a familial syndrome of breast cancer, sarcomas, and other neoplasms. Science 250:1233–1238 (1990).
Mangoni M, Bisanzi S, Carozzi F, Sani C, Biti G, et al: Association between genetic polymorphisms in the *XRCC1, XRCC3, XPD, GSTM1, GSTT1, MSH2, MLH1, MSH3*, and *MGMT* genes and radiosensitivity in breast cancer patients. Int J Radiat Oncol Biol Phys 81:52–58 (2011).
Manolio TA, Collins FS, Cox NJ, Goldstein DB, Hindorff LA, et al: Finding the missing heritability of complex diseases. Nature 461:747–753 (2009).
McCarroll SA, Altshuler DM: Copy-number variation and association studies of human disease. Nat Genet 39:S37–42 (2007).
McCarroll SA, Kuruvilla FG, Korn JM, Cawley S, Nemesh J, et al: Integrated detection and population-genetic analysis of SNPs and copy number variation. Nat Genet 40:1166–1174 (2008).
Melcher R, Hartmann E, Zopf W, Herterich S, Wilke P, et al: LOH and copy neutral LOH (cnLOH) act as alternative mechanism in sporadic colorectal cancers with chromosomal and microsatellite instability. Carcinogenesis 32:636–642 (2011).
Michailidou K, Hall P, Gonzalez-Neira A, Ghoussaini M, Dennis J, et al: Large-scale genotyping identifies 41 new loci associated with breast cancer risk. Nat Genet 45:353–361, 361e1–2 (2013).
Milne RL, Gaudet MM, Spurdle AB, Fasching PA, Couch FJ, et al: Assessing interactions between the associations of common genetic susceptibility variants, reproductive history and body mass index with breast cancer risk in the breast cancer association consortium: a combined case-control study. Breast Cancer Res 12:R110 (2010).
Mitra AK, Singh N, Singh A, Garg VK, Agarwal A, et al: Association of polymorphisms in base excision repair genes with the risk of breast cancer: a case-control study in North Indian women. Oncol Res 17:127–135 (2008).
Mohamedali A, Gaken J, Twine NA, Ingram W, Westwood N, et al: Prevalence and prognostic significance of allelic imbalance by single-nucleotide polymorphism analysis in low-risk myelodysplastic syndromes. Blood 110:3365–3373 (2007).
Montpetit A, Nelis M, Laflamme P, Magi R, Ke X, et al: An evaluation of the performance of tag SNPs derived from HapMap in a Caucasian population. PLoS Genet 2:e27 (2006).
Moore JH: The ubiquitous nature of epistasis in determining susceptibility to common human diseases. Hum Hered 56:73–82 (2003).
Moore JH: A global view of epistasis. Nat Genet 37:13–14 (2005).
Nickels S, Truong T, Hein R, Stevens K, Buck K, et al: Evidence of gene-environment interactions between common breast cancer susceptibility loci and established environmental risk factors. PLoS Genet 9:e1003284 (2013).
O'Keefe C, McDevitt MA, Maciejewski JP: Copy neutral loss of heterozygosity: a novel chromosomal lesion in myeloid malignancies. Blood 115:2731–2739 (2010).
Onay VU, Briollais L, Knight JA, Shi E, Wang Y, et al: SNP-SNP interactions in breast cancer susceptibility. BMC Cancer 6:114 (2006).
Orr N, Lemnrau A, Cooke R, Fletcher O, Tomczyk K, et al: Genome-wide association study identifies a common variant in *RAD51B* associated with male breast cancer risk. Nat Genet 44:1182–1184 (2012).
Palanca Suela S, Esteban Cardenosa E, Barragan Gonzalez E, de Juan Jiménez I, Chirivella González I, et al: *CASP8* D302H polymorphism delays the age of onset of breast cancer in *BRCA1* and *BRCA2* carriers. Breast Cancer Res Treat 119:87–93 (2010).
Parshad R, Price FM, Bohr VA, Cowans KH, Zujewski JA, Sanford KK: Deficient DNA repair capacity, a predisposing factor in breast cancer. Br J Cancer 74:1–5 (1996).
Peto J, Collins N, Barfoot R, Seal S, Warren W, et al: Prevalence of *BRCA1* and *BRCA2* gene mutations in patients with early-onset breast cancer. J Natl Cancer Inst 91:943–949 (1999).
Pharoah PD, Antoniou A, Bobrow M, Zimmern RL, Easton DF, Ponder BA: Polygenic susceptibility to breast cancer and implications for prevention. Nat Genet 31:33–36 (2002).
Pharoah PD, Tyrer J, Dunning AM, Easton DF, Ponder BA; SEARCH Investigators: Association between common variation in 120 candidate genes and breast cancer risk. PLoS Genet 3:e42 (2007).
Pharoah PD, Antoniou AC, Easton DF, Ponder BA: Polygenes, risk prediction, and targeted prevention of breast cancer. N Engl J Med 358:2796–2803 (2008).
Piccolo SR, Camp N, Frey LJ: Polygenic model for predicting breast cancer risk via genome-wide polymorphisms. AMIA Annu Symp Proc 1094 (2008).
Pooley KA, Baynes C, Driver KE, Tyrer J, Azzato EM, et al: Common single-nucleotide polymorphisms in DNA double-strand break repair genes and breast cancer risk. Cancer Epidemiol Biomarkers Prev 17:3482–3489 (2008).
Rafiq S, Tapper W, Collins A, Khan S, Politopoulos I, et al: Identification of inherited genetic variations influencing prognosis in early-onset breast cancer. Cancer Res 73:1883–1891 (2013).
Rahman N, Seal S, Thompson D, Kelly P, Renwick A, et al: *PALB2*, which encodes a BRCA2-interacting protein, is a breast cancer susceptibility gene. Nat Genet 39:165–167 (2007).
Redon R, Ishikawa S, Fitch KR, Feuk L, Perry GH, et al: Global variation in copy number in the human genome. Nature 444:444–454 (2006).
Reich DE, Cargill M, Bolk S, Ireland J, Sabeti PC, et al: Linkage disequilibrium in the human genome. Nature 411:199–204 (2001).
Renwick A, Thompson D, Seal S, Kelly P, Chagtai T, et al: *ATM* mutations that cause ataxia-telangiectasia are breast cancer susceptibility alleles. Nat Genet 38:873–875 (2006).
Rockhill B, Spiegelman D, Byrne C, Hunter DJ, Colditz GA: Validation of the Gail et al. model of breast cancer risk prediction and implications for chemoprevention. J Natl Cancer Inst 93:358–366 (2001).
Sachidanandam R, Weissman D, Schmidt SC, Kakol JM, Stein LD, et al: A map of human genome sequence variation containing 1.42 million single nucleotide polymorphisms. Nature 409:928–933 (2001).
Saeki H, Kitao H, Yoshinaga K, Nakanoko T, Kubo N, et al: Copy-neutral loss of heterozygosity at the *p53* locus in carcinogenesis of esophageal squamous cell carcinomas associated with *p53* mutations. Clin Cancer Res 17:1731–1740 (2011).
Sapkota Y, Robson P, Lai R, Cass CE, Mackey JR, Damaraju S: A two-stage association study identifies methyl-CpG-binding domain protein 2 gene polymorphisms as candidates for breast cancer susceptibility. Eur J Hum Genet 20:682–689 (2012).
Sapkota Y, Ghosh S, Lai R, Coe BP, Cass CE, et al: Germline DNA copy number aberrations identified as potential prognostic factors for breast cancer recurrence. PLoS One 8:e53850 (2013a).
Sapkota Y, Mackey JR, Lai R, Franco-Villalobos C, Lupichuk S, et al: Assessing SNP-SNP interactions among DNA repair, modification and metabolism related pathway genes in breast cancer susceptibility. PLoS One 8:e64896 (2013b).
Sapkota Y, Yasui Y, Lai R, Sridharan M, Robson PJ, et al: Identification of a breast cancer susceptibility locus at 4q31.22 using a genome-wide association study paradigm. PLoS One 8:e62550 (2013c).
Sawyer S, Mitchell G, McKinley J, Chenevix-Trench G, Beesley J, et al: A role for common genomic variants in the assessment of familial breast cancer. J Clin Oncol 30:4330–4336 (2012).
Schneider BP, Radovich M, Sledge GW, Robarge JD, Li L, et al: Association of polymorphisms of angiogenesis genes with breast cancer. Breast Cancer Res Treat 111:157–163 (2008).
Schonfeld SJ, Bhatti P, Brown EE, Linet MS, Simon SL, et al: Polymorphisms in oxidative stress and inflammation pathway genes, low-dose ionizing radiation, and the risk of breast cancer among US radiologic technologists. Cancer Causes Control 21:1857–1866 (2010).
Seal S, Thompson D, Renwick A, Elliott A, Kelly P, et al: Truncating mutations in the Fanconi anemia J gene *BRIP1* are low-penetrance breast cancer susceptibility alleles. Nat Genet 38:1239–1241 (2006).

Sehl ME, Langer LR, Papp JC, Kwan L, Seldon JL, et al: Associations between single nucleotide polymorphisms in double-stranded DNA repair pathway genes and familial breast cancer. Clin Cancer Res 15:2192–2203 (2009).
Sehrawat B, Sridharan M, Ghosh S, Robson P, Cass CE, et al: Potential novel candidate polymorphisms identified in genome-wide association study for breast cancer susceptibility. Hum Genet 130:529–537 (2011).
Shlien A, Malkin D: Copy number variations and cancer susceptibility. Curr Opin Oncol 22: 55–63 (2010).
Shu XO, Long J, Lu W, Li C, Chen WY, et al: Novel genetic markers of breast cancer survival identified by a genome-wide association study. Cancer Res 72:1182–1189 (2012).
Siddiq A, Couch FJ, Chen GK, Lindström S, Eccles D, et al: A meta-analysis of genome-wide association studies of breast cancer identifies two novel susceptibility loci at 6q14 and 20q11. Hum Mol Genet 21:5373–5384 (2012).
Smith P, McGuffog L, Easton DF, Mann GJ, Pupo GM, et al: A genome wide linkage search for breast cancer susceptibility genes. Genes Chromosomes Cancer 45:646–655 (2006).
Smith TR, Levine EA, Freimanis RI, Akman SA, Allen GO, et al: Polygenic model of DNA repair genetic polymorphisms in human breast cancer risk. Carcinogenesis 29:2132–2138 (2008).
Solomon E, Ledbetter DH: Report of the committee on the genetic constitution of chromosome 17. Cytogenet Cell Genet 55:198–215 (1990).
Spiegelman D, Colditz GA, Hunter D, Hertzmark E: Validation of the Gail et al. model for predicting individual breast cancer risk. J Natl Cancer Inst 86:600–607 (1994).
Stacey SN, Manolescu A, Sulem P, Rafnar T, Gudmundsson J, et al: Common variants on chromosomes 2q35 and 16q12 confer susceptibility to estrogen receptor-positive breast cancer. Nat Genet 39:865–869 (2007).
Stacey SN, Manolescu A, Sulem P, Thorlacius S, Gudjonsson SA, et al: Common variants on chromosome 5p12 confer susceptibility to estrogen receptor-positive breast cancer. Nat Genet 40:703–706 (2008).
Stevens VL, McCullough ML, Pavluck AL, Talbot JT, Feigelson HS, et al: Association of polymorphisms in one-carbon metabolism genes and postmenopausal breast cancer incidence. Cancer Epidemiol Biomarkers Prev 16:1140–1147 (2007).
Stram DO: Tag SNP selection for association studies. Genet Epidemiol 27:365–374 (2004).
Stranger BE, Forrest MS, Dunning M, Ingle CE, Beazley C, et al: Relative impact of nucleotide and copy number variation on gene expression phenotypes. Science 315:848–853 (2007).
Thomas D: Gene–environment-wide association studies: emerging approaches. Nat Rev Genet 11:259–272 (2010).
Thomas G, Jacobs KB, Kraft P, Yeager M, Wacholder S, et al: A multistage genome-wide association study in breast cancer identifies two new risk alleles at 1p11.2 and 14q24.1 *(RAD51L1)*. Nat Genet 41:579–584 (2009).
Thompson D, Easton DF, Breast Cancer Linkage Consortium: Cancer incidence in *BRCA1* mutation carriers. J Natl Cancer Inst 94:1358–1365 (2002).
Tice JA, Miike R, Adduci K, Petrakis NL, King E, Wrensch MR: Nipple aspirate fluid cytology and the Gail model for breast cancer risk assessment in a screening population. Cancer Epidemiol Biomarkers Prev 14: 324–328 (2005).
Turnbull C, Ahmed S, Morrison J, Pernet D, Renwick A, et al: Genome-wide association study identifies five new breast cancer susceptibility loci. Nat Genet 42:504–507 (2010).
Tuzun E, Sharp AJ, Bailey JA, Kaul R, Morrison VA, et al: Fine-scale structural variation of the human genome. Nat Genet 37: 727–732 (2005).
Tyrer J, Duffy SW, Cuzick J: A breast cancer prediction model incorporating familial and personal risk factors. Stat Med 23:1111–1130 (2004).
Wacholder S, Hartge P, Prentice R, Garcia-Closas M, Feigelson HS, et al: Performance of common genetic variants in breast-cancer risk models. New Engl J Med 362:986–993 (2010).
Wang X, Goode EL, Fredericksen ZS, Vierkant RA, Pankratz VS, et al: Association of genetic variation in genes implicated in the beta-catenin destruction complex with risk of breast cancer. Cancer Epidemiol Biomarkers Prev 17:2101–2108 (2008).
Wellcome Trust Case Control Consortium, Craddock N, Hurles ME, Cardin N, Pearson RD, et al: Genome-wide association study of CNVs in 16,000 cases of eight common diseases and 3,000 shared controls. Nature 464:713–720 (2010).
Wong KK, deLeeuw RJ, Dosanjh NS, Kimm LR, Cheng Z, et al: A comprehensive analysis of common copy-number variations in the human genome. Am J Hum Genet 80:91–104 (2007).
Wooster R, Bignell G, Lancaster J, Swift S, Seal S, et al: Identification of the breast cancer susceptibility gene *BRCA2*. Nature 378:789–792 (1995).
Yoshihara K, Tajima A, Adachi S, Quan J, Sekine M, et al: Germline copy number variations in *BRCA1*-associated ovarian cancer patients. Genes Chromosomes Cancer 50:167–177 (2011).
Zheng W, Long J, Gao YT, Li C, Zheng Y, et al: Genome-wide association study identifies a new breast cancer susceptibility locus at 6q25.1. Nat Genet 41:324–328 (2009).
Zuk O, Hechter E, Sunyaev SR, Lander ES: The mystery of missing heritability: genetic interactions create phantom heritability. Proc Natl Acad Sci USA 109:1193–1198 (2012).