

REVIEW

Common breast cancer risk variants in the post-COGS era: a comprehensive review

Kara N Maxwell¹ and Katherine L Nathanson^{2,3*}

Abstract

Breast cancer has a strong heritable component, with approximately 15% of cases exhibiting a family history of the disease. Mutations in genes such as *BRCA1*, *BRCA2* and *TP53* lead to autosomal dominant inherited cancer susceptibility and confer a high lifetime risk of breast cancers. Identification of mutations in these genes through clinical genetic testing enables patients to undergo screening and prevention strategies, some of which provide overall survival benefit. In addition, a number of mutant alleles have been identified in genes such as *CHEK2*, *PALB2*, *ATM* and *BRIP1*, which often display incomplete penetrance and confer moderate lifetime risks of breast cancer. Studies are underway to determine how to use the identification of mutations in these genes to guide clinical practice. Altogether, however, mutations in high and moderate penetrance genes probably account for approximately 25% of familial breast cancer risk; the remainder may be due to mutations in as yet unidentified genes or lower penetrance variants. Common low penetrance alleles, which have been mainly identified through genome-wide association studies (GWAS), are generally present at 10 to 50% population frequencies and confer less than 1.5-fold increases in breast cancer risk. A number of single nucleotide polymorphisms (SNPs) have been identified and risk associations extensively replicated in populations of European ancestry, the number of which has substantially increased as a result of GWAS performed by the Collaborative Oncological Gene-environment Study consortium. It is now estimated that 28% of familial breast cancer risk is explained by common breast cancer susceptibility loci. In some cases, SNP associations may be specific to different subsets of women with breast cancer, as defined by ethnicity or estrogen receptor status. Although not yet clinically established, it is hoped that identification of common risk variants may eventually allow identification of women at higher risk of breast cancer and enable implementation of breast cancer screening, prevention or treatment strategies that provide clinical benefit.

Introduction

Breast cancer is the most common malignancy in women, excluding nonmelanoma skin cancer, and the second leading cause of cancer death in women following lung cancer [1]. A family history of breast cancer is a significant risk factor for the development of the disease; at least 10 to 15% of all breast cancer cases may be due to inheritance of a single gene mutation or multiple genetic variants [2-4]. The identification of the hereditary breast and ovarian cancer genes *BRCA1* and *BRCA2*, which are involved in DNA damage response, specifically double-strand break

repair by homologous recombination [5,6], has greatly increased our understanding of the molecular pathways important in breast cancer susceptibility. Additionally, clinical studies have led to the development of effective screening and prevention strategies in women who carry *BRCA1* and *BRCA2* mutations [7], some of which, such as risk-reducing oophorectomy, confer a survival benefit [8]. However, ultimately only 20 to 30% of familial breast cancers are due to *BRCA1* or *BRCA2* mutations [9] and additional genetic variants, such as common low penetrance alleles, must contribute to breast cancer susceptibility. Early identification of individuals at increased risk of breast cancer due to these low-risk variants may lead to enhanced screening and prevention strategies and potentially improved overall survival for this group of patients. This review will summarize the current understanding of common low penetrance susceptibility alleles that contribute to breast cancer risk.

* Correspondence: knathans@exchange.upenn.edu

²Division of Translational Medicine and Human Genetics, Department of Medicine, Perelman School of Medicine at the University of Pennsylvania, 351 BRB 2/3, 421 Curie Blvd, Philadelphia, PA 19104, USA

³Abramson Cancer Center, Perelman School of Medicine at the University of Pennsylvania, Philadelphia, PA 19104, USA

Full list of author information is available at the end of the article

Classification of breast cancer susceptibility genes

With some exceptions, there is an inverse relationship between the risk conferred by a variant in a breast cancer susceptibility gene and the frequency of the variant in the population [4]. Variants generally can be classified as rare high-risk mutations, rare moderate-risk mutations, and common low-risk variants. The majority of variants within a given gene generally fall into one of these categories, although exceptions are being increasingly recognized; for example, the moderate penetrance mutation *BRCA1* p.Arg1699Gln [10]. In general, deleterious mutations in the *BRCA1* and *BRCA2* genes occur in approximately 1:400 to 1:800 women in an unselected population [11] and confer a high (40 to 80%) lifetime risk of breast cancer [12]. Deleterious mutations in genes such as *TP53* and *PTEN* also confer high lifetime risks of breast cancer, but occur even more rarely in the population [13-16]. A number of moderate penetrance mutations have been identified in genes such as *ATM*, *BRIP1*, *CHEK2* and *PALB2* among others, which confer a twofold to fourfold increased risk for breast cancer, although the risk appears higher in the context of a family history [17-20]. Altogether, high and moderate penetrance breast cancer susceptibility mutations in these genes probably account for just over 30% of familial breast cancer cases.

A proportion of the remaining familial breast cancer risk is explained by common low penetrance alleles, of which the majority that have been studied are single nucleotide polymorphisms (SNPs). Carriage of the risk allele at these loci typically confers between a 1.04-fold to 1.40-fold increase or a 0.75-fold to 0.95-fold decrease in breast cancer risk. These alleles are often found between 10 and 50% risk allele frequencies in the population and may act in a combinatorial polygenic manner in an individual to affect breast cancer risk. These interactions are complex; considering alleles with a relative risk for breast cancer of 1.5 and found at a population frequency of 30%, it has been estimated that an individual would need 33 to 40 such risk alleles to explain a threefold increased risk of breast cancer [21]. Whereas genes with mainly high penetrance mutations were identified by linkage and positional cloning and genes with mostly moderate penetrance mutations by candidate gene approaches, identifying the large number of important common low-risk breast cancer variants has required large-scale unbiased discovery approaches. Genome-wide association studies (GWAS) have been critical in the identification of these common low-risk variants.

Evolution in methods to identify common low-risk breast cancer susceptibility alleles

Candidate gene approaches to identify common breast cancer risk alleles

Beginning in the 1990s, multiple case-control studies were conducted to identify common low penetrance

variants associated with a risk of breast cancer. These initial studies by necessity employed a candidate gene approach. The most commonly studied loci show biallelic inheritance, giving rise to three genotypes: the common allele homozygote, the heterozygote and the minor allele homozygote [22]. In many candidate gene studies, the frequencies of the three genotypes at a given candidate locus are assayed in population-based breast cancer cases and matched controls, and, assuming a dominant genetic model, the relative risk of breast cancer for individuals with the heterozygote or minor allele homozygote genotypes are calculated using the common allele homozygotes as the baseline. Unfortunately, many of these studies yielded either false positive or nonsignificant associations with breast cancer risk, probably because the majority of the early studies were underpowered [22]. Dunning and colleagues [22] identified 46 early candidate gene association studies in breast cancer. The median number of cases and controls was only 319, and only 10 of 46 studies had >90% power to detect a 2.5-fold increase in risk for a SNP with a minor allele frequency of 0.2. Furthermore, the three loci identified as potentially remaining significant in the meta-analysis were considered to be false positive because the associations with breast cancer risk showed *P* values in the range of 0.02 to 0.002, which increased to over 0.05 considering multiple comparisons.

Therefore, to identify novel low penetrance breast cancer risk loci, large consortia were formed with the goal of conducting adequately powered case-control genetic association studies. The Breast Cancer Association Consortium (BCAC) was established with this goal and initially included over 20 collaborative groups from European, Australian, American and Asian centers. The BCAC initially employed a candidate gene approach and pooled data from 18 mostly European studies to evaluate 16 SNPs previously reported to be significantly associated with breast cancer risk [23]. By increasing the number of cases and controls genotyped by approximately threefold over any one study of a SNP, the BCAC found that only five of the 16 SNPs (31%) had borderline statistical significance in the larger pooled population [23]. Genotyping of additional individuals within the BCAC study population showed that association of only two SNPs (12%) retained statistical significance: rs1045485 (2q33.1, *CASP8* p.Asp302His) and rs1982073 (19q13.2, *TGFB1* p.Pro10Leu) (Table 1) [24]. These studies confirmed that large numbers of cases and controls are needed to identify replicable associations of SNPs and breast cancer risk.

The multistage genome-wide association study approach

The development of platforms to assay the genotypes of hundreds of thousands of SNPs simultaneously allowed identification of risk alleles without *a priori* knowledge of the location of the risk locus. Using linkage disequilibrium,

Table 1 Major genome-wide association and candidate gene studies to identify SNPs associated with breast cancer risk in populations of European and Asian ancestry

Reference	Study design	Population ^a	SNPs stratified by ER status	Number of SNPs	Number of cases / controls	Novel SNPs identified in study and later confirmed in COGS (with exceptions noted)
BCAC, 2006 [23]	Candidate gene	European and Asian (BCAC)	No	16	Variable per SNP	All with borderline statistical significance
Cox and colleagues, 2007 [24]	Candidate gene	European and Asian (BCAC)	No	9	11,391 to 18,290 / 14,753 to 22,670	rs1045485 ^b , rs1982073 ^c
Easton and colleagues, 2007 [25]	Multistage GWAS	I/II: English, III: European and Asian (BCAC)	No	I: 227,876, II: 12,711, III: 30	I: 390 / 364, II: 3,990 / 3,916, III: 21,860 / 22,578	rs889312, rs13281615, rs2981582 ^d , rs3817198
Hunter and colleagues, 2007 [26]	Multistage GWAS	I/II: European American (CGEMS)	No	I: 528,173, II: 8	I: 1,145 / 1,142, II: 1,776 / 2,072	rs1219648 ^d , rs2420946 ^d
Stacey and colleagues, 2007 [27], Stacey and colleagues, 2008 [28]	Multistage GWAS -> candidate gene	I: Icelandic, II: multi-ethnic American (MEC), Swedish, Dutch, Spanish	Yes	I: 311,524, IIa: 9, IIb: 10	I: 2,183 / 12,877, IIa: 3,898 / 6,921, IIb: 5,028 / 32,090	rs13387042 (ER+), rs10941679 (ER+), rs3803662 ^e (ER+)
Zheng and colleagues, 2009 [29]	Multistage GWAS and replication set (R)	I/II/III: Chinese, R: European American	Yes	I: 607,728, II: 29, III: 4, R: 1	I: 1,505 / 1,522, II: 1,554 / 1,576, III: 3,472 / 900, R: 1,590 / 1,466	rs2046210 (ER+/ER-)
Thomas and colleagues, 2009 [30]	Multistage GWAS	I/II: European and Asian (BCAC) and European American (CGEMS)	Yes	I: 528,173, II: 30,278, III: 24	I: 1,145 / 1,142, II: 4,547 / 4,434, III: 4,078 / 5,223	rs999737, rs11249433 (ER+)
Ahmed and colleagues, 2009 [31]	Candidate gene in two stages	I/II: European and Asian (BCAC) and European American (CGEMS)	No	I: 814, II: 3	I: 3,878 / 3,928, II: 33,134 / 36,141	rs4973768, rs6504950
Turnbull and colleagues, 2010 [32]	Multistage GWAS	I: English, II: European	No	I: 582,886, II: 15	I: 3,659 / 4,897, II: 11,431 / 11,081	rs3757318, rs1011970, rs2380205 ^d , rs10995190, rs704010, rs614367
Fletcher and colleagues, 2011 [33]	Multistage GWAS	I/II/III: English	No	I: 296,114, II: 1,223, III: 14	I: 1,136 / 1,142, II: 3,080 / 3,936, III: 4,237 / 5,044	rs865686
Ghossaini and colleagues, 2012 [34]	Candidate gene	European and Asian (BCAC)	No	72	54,588 / 58,098	rs10771399, rs1292011, rs2823093
Michailidou and colleagues, 2013 [35]	Large-scale candidate gene study	52 BCAC studies (41 European, nine Asian, two African ancestry)	Yes in associated studies	199,961 (iCOGS array)	52,675 / 49,436	41 SNPs ^f

COGS, Collaborative Oncological Gene-environment Study; ER, estrogen receptor; GWAS, genome-wide association study; SNP, single nucleotide polymorphism. ^aThe Breast Cancer Association Consortium (BCAC) population includes individuals of the following descent: Australian, European, Asian, American. The multiethnic cohort (MEC) population includes American individuals of the following descent: European, Latin, Hawaiian, Japanese, African. The Cancer Genetic Markers of Susceptibility (CGEMS) population is a nested case-control study within the Nurses' Health study in the United States. ^bSNPs showing borderline association with breast cancer risk in main COGS breast cancer GWAS. ^cNot tested in COGS and not considered an established locus by COGS. ^dThese SNPs are in the *FGFR2* genomic region; rs2981579 from this region was tested in COGS. ^ers3803662 was newly identified by Easton and colleagues as well as by Stacey and colleagues [25,27]. ^frs616488, rs11552449, rs4849887, rs2016394, rs1550623, rs16857609, rs6762644, rs12493607, rs9790517, rs6828523, rs10472076, rs1353747, rs1432679, rs11242675, rs204247, rs720475, rs9693444, rs6472903, rs2943559, rs11780156, rs10759243, rs7072776, rs11814448, rs7904519, rs11199914, rs3903072, rs11820646, rs12422552, rs17356907, rs11571833, rs2236007, rs2588809, rs941764, rs17817449, rs13329835, rs527616, rs1436904, rs4808801, rs3760982, rs132390, and rs6001930.

several hundred thousand SNPs can be used to represent the several million SNPs in the human genome (tag SNPs), allowing for GWAS associating SNP genotypes with disease risk instead of a candidate gene approach. Given the impracticality and cost of genotyping a large enough number of cases and controls to give sufficient power using a genome-wide array, the initial GWAS utilized multistage approaches [25-27] (Table 1). Easton and colleagues

initially performed genome-wide genotyping of 227,876 SNPs in a case set of 390 family-history-positive breast cancer cases and 364 controls from the United Kingdom, followed by genotyping of the 12,711 most significant SNPs in 3,990 unselected cases and 3,916 controls also from the United Kingdom [25]. From this set, 30 SNPs were selected and genotyped in 21,860 cases and 22,578 controls from the BCAC. Hunter and colleagues

performed genome-wide genotyping of 528,173 SNPs in 1,145 cases and 1,142 controls, followed by genotyping of eight SNPs in an additional 1,776 cases and 2,072 controls from the Cancer Genetic Markers of Susceptibility population from the United States [26]. Cases were non-Hispanic Caucasian women with mostly postmenopausal breast cancer, unselected for family history. Finally, Stacey and colleagues genotyped 311,524 SNPs in 2,183 cases and 12,877 controls from Iceland, again unselected for a family history of breast cancer [27]. The nine top SNPs from this study were then genotyped in an independent replication set of 3,898 cases and 6,921 controls from Iceland, Spain, Sweden and the Netherlands. In addition, Stacey and colleagues genotyped 10 candidate SNPs at 5p12-11 in 5,028 cases and 32,090 controls of European descent [28].

In total, these three initial GWAS identified significant associations for a risk of breast cancer with nine SNPs in seven genomic regions. Three SNPs within the *FGFR2* gene were identified – rs2981582 (10q26.13, *FGFR2* c.109 + 906 T > C), rs2420946 (10q26.13, *FGFR2* c.109 + 1899A > G), and rs1219648 (10q26.13, *FGFR2* c.109 + 7033 T > C) [25,26] – and the association of this locus with breast cancer risk has been replicated in a number of other studies and retained statistical significance in meta-analyses (see references in later Table 3). The other SNPs identified were rs13387042 (2q35, intergenic), rs10941679 (5p12, intergenic), rs889312 (5q11.2, intergenic, 5' to *MAP3K1*), rs13281615 (8q24.21, intergenic), rs3817198 (11p15.5, *LSP1* c.*13 + 200 T > C), and rs3803662 (16q12.1, intergenic, 5' to *TOX3*) [25,27,28] (Table 1), and the associations with breast cancer risk have also been well replicated (see references in later Table 3). Of note, the SNPs identified by Stacey and colleagues only showed association in estrogen receptor (ER)-positive breast cancer; however, the numbers of ER-negative cases included was small and the other studies did not stratify by ER status [27,28].

Subsequent genome-wide association studies and candidate gene studies in European and Asian populations

Several other multistage GWAS and candidate gene studies were carried out between 2008 and 2012 in populations of individuals of mainly European or Asian ancestry [29-34]. These studies identified 18 SNPs showing significant associations with a risk of breast cancer. Zheng and colleagues identified rs2046210 (6q25, intergenic, 6 kb 3' to *CCDC170*) to be associated with a risk of both ER-positive and ER-negative breast cancer in Chinese individuals; this association was replicated in a population of American women of European ancestry [29]. Three GWAS and two candidate gene studies were performed to continue the analysis of populations of European ancestry [30-34]. These studies identified associations with breast cancer risk for 14 SNPs: rs11249433 (1p11.2, intergenic),

rs4973768 (3p24.1, *SLC4A7* c.*2242G > A), rs3757318 (6q25, *CCDC170* c.1294-129G > A), rs1011970 (9p21.3, *CDK2NB* antisense RNA), rs865686 (9q31.2, intergenic), rs2380205 (10p15.1, intergenic, 2.6 kb 5' to *GDI2*), rs10995190 (10q21.2, *ZNF365* c.981 + 59126G > A), rs704010 (10q22.3, *ZMIZ1* c.-337 + 12121 T > C), rs614367 (11q13.3, intergenic), rs10771399 (12p11.22, intergenic, 29 kb 5' to *PTHLH*), rs1292011 (12q24.21, intergenic), rs999737 (14q24.1, *RAD51B* c.1037-43041C > T), rs6504950 (17q22, *STXBP4* c.-156-6504G > A), and rs2823093 (21q21, intergenic, 5' to *NRIP1*) (Table 1).

The iCOGS array and the advent of the large-scale candidate gene association study

The most recently identified set of SNPs associated with breast cancer risk in Europeans and Asians come from an unprecedented genotyping effort by the Collaborative Oncological Gene-environment Study (COGS) [35]. A total of 211,155 SNPs were rationally chosen for a custom array (iCOGS) and tested in 52,675 unselected breast cancer cases and 49,436 controls from the BCAC. The SNPs on the iCOGS array were nominated by consortium members, and were: SNPs identified by meta-analysis of prior GWAS; SNPs for fine mapping of known susceptibility loci; candidate functional variants and moderate penetrance variants; and SNPs related to other traits, including other cancers and medical conditions. In all, data were obtained for 199,961 SNPs in 52,675 cases and 49,436 controls from the BCAC study groups, which now include 41 European populations, nine Asian populations, and two African-American populations.

In the first phase of the main COGS breast cancer study, associations with breast cancer risk for previously established breast cancer loci were studied [35]. For the *FGFR2* genomic region, SNP rs2981579 (10q26.13, *FGFR2* c.110-12117 T > C) was chosen (this SNP is in linkage disequilibrium at $r^2 < 0.6$ with the three SNPs described above) and showed highly significant associations with a risk of breast cancer. For the remaining 26 SNPs presented in Table 1, strong association with breast cancer risk was shown for 22 of the SNPs. The SNPs rs2380205 (10p15.1, intergenic, 2.6 kb 5' to *GDI2*) and rs1045485 (2q33.1, *CASP8* p.Asp302His) showed a trend towards association, and the SNPs rs2284378 (20q11, *RALY* c.-93 + 6158 T > C) [63] and rs1982073 (19q13.2, *TGFBI* p.Pro10Leu) [23,24,44] were not tested on the iCOGS array. Finally, three SNPs initially found to be associated with ER-negative breast cancer – rs10069690 (5p15.33, *TERT* c.1951-205G > A), rs17530068 (6q14.1, intergenic) and rs8170 (19p13.1, *BABAMI* c.837G > A) (Table 2) [57,63,92] – were tested as established loci in this phase of the study. SNPs rs10069690 and rs17530068 were found to have a significant association and rs8170 showed a trend towards association with breast cancer risk. Michailidou and colleagues

Table 2 Major genome-wide association and candidate gene studies to identify SNPs associated with breast cancer risk in other populations

Reference	Study design	Population ^a	Ethnicity	Number of SNPs	Number of cases / controls ^b	Novel SNPs identified in study
Gaudet and colleagues, 2010 [78]	Multistage GWAS	<i>BRCA2</i> carriers (CIMBA)	European and European American (majority)	I: 529,163, II: 113	I: 899 / 804, II: 1,263 / 1,222	rs3114994 ^c
Antoniou 2010 [51]	Multistage GWAS	I/II: <i>BRCA1</i> carriers, III: unselected, IV: TNBC	I/II: European and European American, III: English, IV: Caucasian (TNBC)	I: 555,616, II: 89, III: 2, IV: 2	I: 1,190 / 1,193, II: 2,974 / 3,012, III: 6,800 / 6,613, IV: 2,301 / 3,949	rs8170 ^d (<i>BRCA1+</i> , ER-, TNBC)
Haiman and colleagues, 2011 [57]	Multistage GWAS	ER- and TNBC	I: African and European American, II: European American	I: 3,154,485, II: 1	I: 3,787 / 4,345, II: 2,222 / 16,363	rs10069690 ^d (ER-, TNBC)
Siddiq and colleagues, 2012 [63]	Meta-analysis in two stages	I: ER- and TNBC, II: ER+ and ER-	I: African and European American, II: European and Asian (BCAC), Hispanic, Japanese	I: variable, II: 86	I: 4,754 / 31,663, II: 11,209 / 8,404	rs17530068 ^d , rs2284378 ^e (ER-)
Garcia-Closas and colleagues, 2013 [39]	Large-scale candidate gene study	ER- and TNBC in 40 BCAC studies	European	13,276 (iCOGS array)	6,514 / 41,455	rs4245739, rs6678914, rs12710696, rs11075995
Zheng and colleagues, 2013 [37]	Large-scale candidate gene study	I/II: 16 ABCC studies, III: 9 COGS studies	Asian (Chinese, Korean, Japanese, Thai)	I: 906,602, II: 44, III: 67 on iCOGS array	23,367 / 25,579	Confirmed associations for 31/67 of the established loci
Long and colleagues, 2013 [42]	Large-scale candidate gene study	SCCS and NBHS studies	African-American	67 on iCOGS array	1,231 / 2,069	Confirmed associations for 10/67 of the established loci
Couch and colleagues, 2013 [40]	Large-scale candidate gene study	<i>BRCA1</i> carriers (CIMBA)	European and European American (majority)	I:2,568,349, II: 201,242 on iCOGS, III: 17	I: 1,103 / 1,239, II: 4,637 / 4,681 + 2,387 from stage I, III: 1,394 / 1,252	rs2290854, rs6682208, rs11196174, rs11196175
Gaudet and colleagues, 2013 [41]	Large-scale candidate gene study	<i>BRCA2</i> carriers (CIMBA)	European and European American (majority)	200,908 on iCOGS array	10,048	rs9348512, rs619373, rs184577

BCAC, Breast Cancer Association Consortium; CIMBA, Consortium of Investigators of Modifiers of *BRCA1/2*; ER, estrogen receptor; GWAS, genome-wide association study; NBHS, Nashville Breast Health Study; SCCS, Southern Community Cohort Study; SNP, single nucleotide polymorphism; TNBC, triple-negative breast cancer. ^aAsian Breast Cancer Consortium (ABCC) includes patients of Chinese, Japanese, and Korean descent; the Collaborative Oncological Gene-environment Study (COGS) Asian studies include individuals of Chinese, Japanese, Korean, and Thai descent. ^bIn the *BRCA1/2* GWAS, cases refer to carriers with breast cancer and controls refer to carriers without breast cancer. ^cSNPs not tested in COGS and not considered an established locus by COGS. ^dSNPs showing borderline association with breast cancer risk in main COGS breast cancer GWAS. ^eHighly significant SNPs in main COGS breast cancer GWAS.

therefore considered there to be 27 established low penetrance breast cancer risk loci prior to the discovery phase of the COGS study (the 26 bold SNPs in Table 3 studied in COGS and rs2284378) [35].

To identify novel SNPs associated with breast cancer risk, 29,807 SNPs identified by analysis of nine prior GWAS but not found in prior breast cancer loci were chosen and data collected for 45,290 cases and 41,880 controls [35]. From this analysis, 41 SNPs were newly identified as significantly associated with breast cancer risk in unselected cases of mainly European and Asian descent [35]. In all, the authors estimated that ~14% of familial breast cancer risk in people of European descent is explained by these 67 established loci, with an additional ~14% of the familial risk explained by SNPs that showed

associations in the COGS study but did not reach statistical significance.

Some of the SNPs newly identified in the COGS study as associated with breast cancer risk are found in genes with a known or plausible role in cancer susceptibility. Seven of the SNPs are single nucleotide variations found within the gene boundaries of DNA repair or cell cycle genes, namely: rs11571833 (13q13.1, *BRCA2* p.Lys3326Ter), rs11552449 (1p13.2, *DCLRE1B* p.His61Tyr), rs9790517 (4q24, *TET2* c.-193 + 17535C > T), rs2046210 (6q25, intergenic, 6 kb 3' to *CCDC170*), rs2236007 (14q13.3, *PAX9* c.631 + 41G > A), rs2588809 (14q24.1, *RAD51B* c.757-98173 T > C), and rs941764 (14q32.11, *CCDC88C* c.271-15014 T > C). In addition, three SNPs are in the immediate vicinity of genes that have shown prior associations with a risk of other

Table 3 Extended information on SNPs identified in the major GWAS for breast cancer risk or SNPs used to generate published polygenic risk scores

SNP ^a	Cyto-band	Gene/HGVS designation if applicable ^b	Breast cancer risk in first reported publication	RAF (EA)	Additional populations studied (other than EA) ^c	PRS ^d				References
						A	B	C	D	
rs11249433	1p11.2	Intergenic	1.16 (1.09 to 1.24)	0.39	A, ER+	*	*	*		[30,32,35-38]
rs11552449	1p13.2	<i>DCLRE1B</i> p.His61Tyr	1.08 (1.02 to 1.14)	0.17						[35]
rs616488	1p36.22	<i>PEX14</i> c.84 + 10837A > G	0.94 (0.90 to 0.98)	0.33	ER-					[35,39]
rs6678914	1q32.1	<i>LGR6</i> c.213-7375G > A	1.10 (1.06 to 1.13)	0.59	ER- only					[39]
rs2290854	1q32.1	<i>MDM4</i> c.903 + 20A > G	1.13 (1.08 to 1.18)	0.33	B1+ only					[40]
rs4245739	1q32.1	<i>MDM4</i> c.*32C > A	1.14 (1.10 to 1.18)	0.26	ER- only					[39]
rs6682208	1q32.1	Intergenic	1.12 (1.07 to 1.17)	0.34	B1+ only					[40]
rs184577	2p22.2	CYP1B1 antisense	0.85 (0.79 to 0.91)		B2+ only					[41]
rs12710696	2p24.1	Intergenic	1.10 (1.06 to 1.13)	0.36	ER- only					[39]
rs4849887	2q14.2	Intergenic	0.90 (0.84 to 0.96)	0.10	AA					[35,42]
rs2016394	2q31.1	Intergenic	0.95 (0.92 to 0.99)	0.48						[35]
rs1550623	2q31.1	Intergenic	0.91 (0.86 to 0.96)	0.16						[35]
rs1045485^e	2q33.1	CASP8 p.Asp302His	0.89 (0.85 to 0.93)	0.13	AA^e		*	*	*	[23,24,32,35,42,43]
rs17468277 ^f	2q33.1	<i>ALS2CR12</i> c.1191G > A	0.83 (0.70 to 0.98)	0.09		*				[44]
rs13387042	2q35	Intergenic	1.20 (1.14 to 1.26)	0.50	H, A^g, AA, B1 + ^g	*	*	*	*	[27,30,32,33,35,40,42,45-50]
rs16857609	2q35	<i>DIRC3</i> noncoding transcript	1.09 (1.05 to 1.14)	0.26	A, ER-					[35,37,39]
rs4973768	3p24.1	SLC4A7 c.*2242G > A	1.11 (1.08 to 1.13)	0.46	A, B2+	*	*			[31-33,35,37,41,47,48,51,52]
rs12493607	3p24.1	<i>TGFB2</i> c.95-3300G > C	1.04 (1.00 to 1.09)	0.35						[35]
rs6762644	3p26.1	<i>ITPR1</i> c.4509 + 651A > G	1.06 (1.02 to 1.11)	0.4						[35]
rs9790517	4q24	<i>TET2</i> c.-193 + 17535C > T	1.09 (1.04 to 1.14)	0.23						[35]
rs6828523	4q34.1	<i>ADAM29</i> c.-450-5711C > A	0.89 (0.83 to 0.94)	0.13						[35,37]
rs4415084 ^f	5p12	Intergenic	1.16 (1.10 to 1.21)	0.40	B1 + ^f	*	*			[28,53]
rs10941679	5p12	Intergenic	1.19 (1.13 to 1.26)	0.24	AA, A, B2 + ^e	*	*			[30,32,33,35,37,47,50,54,55]
rs7716600 ^f	5p12	Intergenic, 54 kb 3'-MRPS30	1.24 (1.14 to 1.34)	0.23	A			*		[56]
rs10069690	5p15.33	TERT c.1951-205G > A	1.18 (1.13 to 1.25)	0.27	AA, ER+, ER-, B1+		*			[35,42,57-59]
rs16886113	5q11.2	Intergenic, 5' to <i>MAP3K1</i>	1.24 (1.11 to 1.38)	0.06	B2+ only					[41]
rs889312	5q11.2	Intergenic, 5' to <i>MAP3K1</i>	1.13 (1.10 to 1.16)	0.28	A	*	*	*	*	[25,32,35,37,48,56,60-62]
rs10472076	5q11.2	Intergenic, 5' to <i>MAP3K1</i>	1.06 (1.02 to 1.11)	0.38						[35]
rs1353747	5q11.2	<i>PDE4D</i> c.809-2683A > C	0.90 (0.84 to 0.96)	0.10						[35]
rs1432679	5q33.3	<i>EBF1</i> c.778 + 6101G > A	1.06 (1.02 to 1.10)	0.43	A, ER-					[35,37,39]
rs204247	6p23	Intergenic (11 kb 5'- <i>RANBP9</i>)	1.06 (1.02 to 1.10)	0.43						[35]
rs9348512	6p24.3	Intergenic	0.85 (0.80 to 0.90)		B2+ only					[41]
rs11242675	6p25.3	Intergenic (3.9 kb 3'- <i>FOXQ1</i>)	0.97 (0.93 to 1.01)	0.39	A					[35,37]
rs17530068	6q14.1	Intergenic	1.12 (1.08 to 1.16)	0.22	AA, ER-					[35,39,63]
rs6569479 ^f	6q22.33	<i>RNF146</i> c.3-1173 T > C	1.39 (1.23 to 1.57)	0.21	AJ		*			[64,65]
rs2253407	6q25.1	<i>SYNE1</i> c.23020-466C > A	0.92 (0.86 to 0.98)	0.47	B2+ only					[41]
rs9485372	6q25.1	Intergenic, 5' tp <i>TAB2</i>	0.89 (0.84 to 0.94)		A only					[37]

Table 3 Extended information on SNPs identified in the major GWAS for breast cancer risk or SNPs used to generate published polygenic risk scores (Continued)

rs3757318	6q25.1	CCDC170 c.1294-129G > A	1.30 (1.17 to 1.46)	0.07	ER-	* *	[29,32,33,35,37-40,48,66-69]
rs2046210	6q25.1	Intergenic, 6 kb 3'-CCDC170	1.29 (1.21 to 1.37)	0.35	A, ER-, B1+	* *	
rs9383938 ^f	6q25.1	Intergenic, 5' to ESR1	1.18 (1.10 to 1.27)	0.08		*	[33,70]
rs720475	7q35	ARHGEF5 c.4531 + 646G > A	0.93 (0.89 to 0.98)	0.25			[35]
rs9693444	8p21	Intergenic	1.07 (1.03 to 1.12)	0.32	A		[35,37]
rs6472903	8q21.11	Intergenic	0.88 (0.84 to 0.93)	0.18	A		[35,37]
rs2943559	8q21.11	HNF4G c.-24 + 15494A > G	1.17 (1.09 to 1.26)	0.07	ER-		[35,39]
rs4733664	8q24.21	Intergenic	1.10 (1.04 to 1.17)	0.41	B2+ only		[41]
rs13281615	8q24.21	Intergenic	1.08 (1.05 to 1.11)	0.40		* * * *	[25,33,35,47,71-73]
rs1562430	8q24.21	Intergenic	0.84 (0.78 to 0.90)	0.40		*	[32,56]
rs11780156	8q24.21	Intergenic (32 kb 3'-MIR1208)	1.13 (1.07 to 1.19)	0.16			[35]
rs10965163	9p21.3	MTAP c.561C > T	0.84 (0.77 to 0.93)	0.10	B2+ only		[41]
rs1011970	9p21.3	CDK2NB antisense RNA	1.09 (1.04 to 1.14)	0.17	ER-	* *	[32,35,39,41,74]
rs865686	9q31.2	Intergenic	0.90 (0.86 to 0.96)	0.39	AA, ER+	* *	[33,35,41,74-76]
rs10759243	9q31.2	Intergenic (53 kb 5'-KLF4)	1.07 (1.02 to 1.12)	0.39	A		[35]
rs7072776	10p12.31	Intergenic (382 bp 3'-MLLT10)	1.11 (1.07 to 1.16)	0.29			[35]
rs11814448	10p12.31	Intergenic (23 kb 5'-DNAJC1)	1.35 (1.17 to 1.56)	0.02			[35]
rs2380205^e	10p15.1	Intergenic, 2.6 kb 5'-GDI2	0.94 (0.91 to 0.98)	0.43		* *	[32,35]
rs16917302	10q21.2	ZNF365 c.981 + 41642A > C	0.75 (0.66 to 0.86)	0.11	B2 + ^e	*	[77,78]
rs10995190	10q21.2	ZNF365 c.981 + 59126G > A	0.76 (0.82 to 0.91)	0.15	A, ER-, B2 + ^f	* *	[32,35,37,39,41,74,79]
rs704010	10q22.3	ZMIZ1 c.-337 + 12121 T > C	1.07 (1.03 to 1.11)	0.39	A, AA	* *	[32,35,37,41,58,80]
rs7904519	10q25.2	TCF7L2 c.382-25857A > G	1.06 (1.02 to 1.10)	0.46	ER-		[35,39]
rs11196174	10q25.2	TCF7L2 c.381 + 22730A > G	1.13 (1.07 to 1.18)	0.31	B1+ only		[40]
rs11199914	10q26.12	Intergenic (77 kb 5'-7SK)	0.94 (0.89 to 0.98)	0.32			[35]
rs2981582 ^f	10q26.13	FGFR2 c.109 + 906 T > C	1.26 (1.23 to 1.28)	0.39	AA, H, B2+	* * * *	[25,26,30,32,46,48,60,62]
rs2420946 ^f	10q26.13	FGFR2 c.109 + 1899A > G	1.24 (1.03 to 1.50)	0.38	B2+		[26,41]
rs1219648	10q26.13	FGFR2 c.109 + 7033 T > C	1.23 (1.02 to 1.48)	0.39	AA, A, H		[26,27,37,49,62,81,82]
rs2981579	10q26.13	FGFR2 c.110-12117 T > C	1.27 (1.24 to 1.29)	0.40			[35]
rs3817198	11p15.5	LSP1 c.*13 + 200 T > C	1.07 (1.04 to 1.11)	0.30	A, B1 + ^d, B2+	* * * *	[25,32,35,37,40,41,45,48,83]
rs909116 ^f	11p15.5	TNNT3 c.-19 + 954 T > C	1.17 (1.10 to 1.24)	0.53		*	[32]
rs3903072	11q13.1	Intergenic (7.4 kb 3'-CFL1)	0.92 (0.89 to 0.96)	0.47			[35]
rs614367	11q13.3	Intergenic	1.15 (1.10 to 1.20)	0.15	A, ER+, B2 + ^e	* *	[32,35,37,41,80,84,85]
rs11820646	11q24.3	Intergenic (139 kb 3'-BARX2)	0.93 (0.90 to 0.97)	0.41	A, ER-		[35,37,39]
rs27633	12p11.22	PTHLH c.-266 + 555A > C	1.14 (1.07 to 1.21)	0.39	B2+ only		[41]

Table 3 Extended information on SNPs identified in the major GWAS for breast cancer risk or SNPs used to generate published polygenic risk scores (Continued)

rs10771399	12p11.22	Intergenic, 29 kb 5'-PTHLH	0.85 (0.83 to 0.88)	0.12	A, B1+, B2 +^e, ER-	*	[34,35,37,39,40,74]
rs12422552	12p13.1	Intergenic (11 kb 5'-U6)	1.11 (1.05 to 1.16)	0.26	A		[35,37]
rs17356907	12q22	Intergenic (16 kb 3'-Y_RNA)	0.89 (0.85 to 0.93)	0.3	A, ER-		[35,37,39]
rs1292011	12q24.21	Intergenic	0.92 (0.91 to 0.94)	0.41	A, B2 +^e		[34,35,37,80]
rs11571833	13q13.1	<i>BRCA2</i> p.Lys3326Ter	1.39 (1.13 to 1.71)	0.01	ER-		[35,39]
rs2236007	14q13.3	<i>PAX9</i> c.631 + 41G > A	0.88 (0.83 to 0.93)	0.21	A		[35,37]
rs2588809	14q24.1	<i>RAD51B</i> c.757-98173 T > C	1.07 (1.01 to 1.13)	0.16			[35]
rs999737	14q24.1	<i>RAD51B</i> c.1037-43041C > T	0.94 (0.88 to 0.99)	0.24	AA, ER-, B1 +^d	* * *	[30,32,35,36,40,42,47,66,70]
rs941764	14q32.11	<i>CCDC88C</i> c.271-15014 T > C	1.05 (1.00 to 1.09)	0.34			[35]
rs3803662	16q12.1	Intergenic, 5' to <i>TOX3</i>	1.20 (1.16 to 1.24)	0.25	A, AA, B1 +^e, B2+, MBC, ER-	* * * *	[25,27,32,35,37,39,41,46,55,60,66,70,73,78]
rs4784227	16q12.1	Intergenic, 5' to <i>TOX3</i>	1.27 (1.22 to 1.32)	0.25	A		[37]
rs17817449	16q12.2	<i>FTO</i> c.-215-30685 T > G	0.95 (0.91 to 0.99)	0.4	ER-		[35,39]
rs13329835	16q23.2	<i>CDYL2</i> c.1007 + 3855 T > C	1.14 (1.09 to 1.19)	0.22	AA		[35,42]
rs2075555 ^f	17q21.33	<i>COL1A1</i> c.804 + 80A > C	No association	0.14		*	[86]
rs6504950	17q22	<i>STXBPA</i> c.-156-6504G > A	0.95 (0.92 to 0.97)	0.27	ER-	* *	[31,39,41,87]
rs527616	18q11.2	Intergenic	0.91 (0.87 to 0.95)	0.38			[35]
rs1436904	18q11.2	<i>CHST9</i> c.202 + 33413A > C	0.93 (0.9 to 0.97)	0.4			[35]
rs8170^e	19p13.1	<i>BABAM1</i> c.837G > A	1.26 (1.17 to 1.35)	0.20	AA, B1+, ER-	*	[35,39-42,58,66,77]
rs4808801	19p13.11	<i>ELL</i> c.744 + 1247 T > C	0.94 (0.90 to 0.98)	0.35	AA, ER-		[35,39,42]
rs1982073 ^f	19q13.2	<i>TGFB1</i> p.Pro10Leu	1.08 (1.04 to 1.11)	0.38			[23,24,44]
rs3760982	19q13.31	Intergenic (1.1 kb 5'- <i>KCNN4</i>)	1.06 (1.02 to 1.10)	0.46			[35]
rs2284378 ^f	20q11	<i>RALY</i> c.-93 + 6158 T > C	1.08 (1.05 to 1.12)	0.35			[63]
rs13039229	20q13.33	Intergenic (2.1 kb 3' of <i>PTK6</i>)	0.90 (0.84 to 0.97)	0.21	B2+ only		[41]
rs311499 ^f	20q13.33	Intergenic (1.4 kb 3'- <i>GMEB2</i>)	0.72 (0.61 to 0.85)	0.07	B2+ only		[78]
rs2823093	21q21	Intergenic (5' to <i>NRIP1</i>)	0.94 (0.92 to 0.96)	0.27			[34,35]
rs132390	22q12.2	<i>EMID1</i> c.403 + 272C > T	1.36 (1.19 to 1.54)	0.04			[35]
rs6001930	22q13.1	<i>MKL1</i> c.-59-16944A > G	1.17 (1.11 to 1.25)	0.11	ER-		[39]
rs619373	Xq27.1	<i>FGF13</i> c.50-73946C > T	1.30 (1.17 to 1.45)	0.03	B2+ only		[41]

COGS, Collaborative Oncological Gene-environment Study; EA, European descent; ER, estrogen receptor; GWAS, genome-wide association studies; PRS, polygenic risk score; RAF, risk allele frequency; SNP, single nucleotide polymorphism. ^aBold SNPs are considered the established breast cancer loci pre-COGS. All SNPs listed are in independent linkage disequilibrium blocks as determined using HaploReg with $r^2 > 0.8$, with the exception of the 2q33.1 SNPs rs1045485 and rs17468277. Two SNPs in linkage disequilibrium with listed SNP were tested in other populations; that is, rs76972116 at 4q34, and rs7107217 at 11q24. ^bHuman Genome Variation Society (HGVS) names are given for variants found with the coding region of a gene using the designation for transcript variant #1 as per the Ensembl/UCSC GRCh37/hg19 human genome assembly. ^cAll SNPs have been identified in populations of European descent (EA) and unselected for ER status; if other populations have been studied, this is noted. If a SNP has been identified only in a subpopulation, it is denoted with 'only'. A, Asian; AA, African-American; AJ, Ashkenazi Jewish; B1+, BRCA1-positive; B2+, BRCA2-positive; H, Hispanic; ER+, estrogen receptor-positive breast cancer; ER-, estrogen receptor-negative breast cancer; MBC, male breast cancer. SNP association listed for population subtype if reported $P < 10^{-3}$.

^dPolygenic risk scores are from: model A, Sawyer and colleagues [88]; model B, Husing and colleagues [89]; model C, Wacholder and colleagues [90]; model D, Gail [91]. ^eSNPs showing borderline association with breast cancer risk in COGS; if population specific, this is indicated in the Additional populations studied column.

^fSNPs that were not studied in any COGS study [35,37,39-42]; if population specific, this is indicated in the Additional populations studied column. ^gSNPs showing no association with breast cancer risk in COGS; if population specific, this is indicated in the Additional populations studied column.

malignancies: rs12493607 (3p24.1, *TGFBR2* c.95-3300G > C), rs7072776 (10p12.31, intergenic, 382 bp 3' of *MLLT10*), and rs6001930 (22q13.1, *MKLI* c.-59-16944A > G). Further studies will need to be done to determine whether these SNPs or other variants in linkage disequilibrium are functionally causative, and whether they influence expression or function of the gene proximate to which they are located or other surrounding genes.

From lumping to splitting – genome-wide association studies and candidate gene studies identify common low penetrance risk alleles in subgroups of patients

Common low penetrance loci in estrogen receptor-positive versus estrogen receptor-negative breast cancer

Patients with breast cancers that are negative for ERs (ER-negative) with or without negativity for Her2, the former referred to as triple-negative breast cancers (TNBC), have a significantly poorer prognosis and occur in patients with different clinical characteristics than ER-positive breast cancers [93,94]. The set of common risk variants contributing to breast cancer risk is thus likely to be different among ER-positive patients, ER-negative patients and TNBC patients. Delineation of ER-negative/TNBC versus ER-positive breast cancer-specific loci could refine risk models to assist in identification of women at higher risk for TNBC; that is, the poorer prognosis type of breast cancer. In the GWAS by Stacey and colleagues, the cases were stratified by ER positivity and a stronger association was seen for the three identified SNPs with ER-positive than ER-negative breast cancer [27,28].

The majority of the other early GWAS, however, did not stratify by ER status because the number of patients with ER-negative breast cancer was small in these studies. Focused studies with adequate power were therefore needed to identify SNPs associated with risk in this breast cancer subtype. A number of candidate SNP studies have investigated some of the known common risk variants specifically in patients with either ER-positive breast cancer, ER-negative breast cancer or TNBC [36,44,66,75,84,92,95,96]. In some cases for the SNPs at established breast cancer loci, the associations were stronger for or specific to TNBC – that is, rs13387042 (2q35, intergenic), rs889312 (5q11.2, intergenic, 5' to *MAP3KI*), rs3817198 (11p15.5, *LSP1* c.*13 + 200 T > C), rs999737 (14q24.1, *RAD51B* c.1037-43041C > T), rs3803662 (16q12.1, intergenic, 5' to *TOX3*), and rs8170 (19p13.1, *BABAMI* c.837G > A) [44,66,92] – or ER-negative breast cancer – that is, rs11249433 (1p11.2, intergenic) [36]. In other cases, SNP associations were stronger for or specific for a risk for ER-positive breast cancer – that is, rs865686 (9q31.2, intergenic) [75] and rs614367 (11q13.3, intergenic) [84].

Other studies have used the fact that certain patient subgroups have a higher incidence of TNBC, for example

BRCA1 mutation carriers and African-Americans, to study common risk variants in this subgroup. Antoniou and colleagues performed a GWAS in *BRCA1* carriers; in this study, rs8170 (19p13.1, *BABAMI* c.837G > A) was then genotyped in ER-negative and TNBC cases and found to be significantly associated with the risk of breast cancer in these subgroups (Table 2) [51]. In the study by Haiman and colleagues, a GWAS was first performed in patients of both African and European ancestry unselected for hormone receptor status; the identified SNP rs10069690 (5p15.33, *TERT* c.1951-205G > A) was then genotyped in a set of TNBC patients and found to be significantly associated with this subgroup (Table 2) [57]. Finally, Siddiq and colleagues performed a meta-analysis of prior GWAS and found that rs17530068 (6q14.1, intergenic) and rs2284378 (20q11, *RALY* c.-93 + 6158 T > C) were associated with a risk of ER-negative breast cancer (Table 2) [63]. The COGS group has also used the large-scale candidate SNP approach specifically in patients with ER-negative breast cancer and TNBC [39]. In this study, 13,276 SNPs on the iCOGS array were assayed in 6,514 ER-negative cases and 41,455 controls. The study identified four new SNPs – rs4245739 (1q32.1, *MDM4* c.*32C > A), rs6678914 (1q32.1, *LGR6* c.213-7375G > A), rs12710696 (2p24.1, intergenic), and rs11075995 (16q12.2, *FTO* c.-138 + 11162A > T) – that show specific associations with ER-negative breast cancer (Table 2) [39]. In addition, this study group performed a meta-analysis of 10,707 ER-negative cases and 76,649 controls and found that 18 of the 26 established breast cancer loci had associations with $P < 0.05$ with ER-negative breast cancer, as did 25 of the 41 new loci from the main COGS breast cancer study (Table 3) [39].

Common low penetrance loci in ethnic subgroups

There is much interest in using SNP genotyping to stratify breast cancer risk and inform clinical management (as discussed below). However, one of the recognized limitations is the extent to which the major GWAS performed in European populations can be generalized across ethnicities. A number of candidate gene studies and GWAS have now been carried out in patients of African and East Asian descent; some of the most well-studied common SNPs have shown odds ratios (ORs) consistent with those seen in European studies (Table 3) [56,58,81,97-99]. However, many of the other SNPs identified as associated with risk in Europeans have not been studied in African, East Asian or other populations. Recently, however, Zheng and colleagues performed a GWAS in women with breast cancer of Asian ancestry and combined the data with genotyping of Asian individuals using the iCOGS array. They found that 31 of the 67 established breast cancer loci also show associations with a risk of breast cancer in East Asians (Tables 2 and 3) [37]. In addition, Long and colleagues

genotyped the 67 loci in African-American cases and controls, and found that seven of the 67 SNPs showed significant associations and three SNPs had borderline associations with breast cancer risk in African-Americans with on average 83% African ancestry (Tables 2 and 3) [42].

Several SNPs have been found to be associated with breast cancer solely in studies of women of either African or Asian descent, but many need to be replicated in independent populations of the discovery ethnicity. Associations with one Asian-specific SNP that is not in linkage disequilibrium with the established SNP at that genomic region, rs9485372 (6q25.1, *TAB2* c.6 + 68962G > A), was replicated in the COGS study above (Table 3) [37]; however, other SNP associations have not yet been replicated. The 6q25.1 chromosomal region demonstrates an important caveat when attempting to generalize association of specific loci across ethnicities. Associations with breast cancer risk were found independently within 6q25.1 at rs2046210 (intergenic, 6 kb 3' to *CCDC170*) in an East Asian population with an OR of 1.29 (95% confidence interval (CI) = 1.21 to 1.37) [29], and at rs3757318 (*CCDC170* c.1294-129G > A) in European populations with an OR of 1.30 (95% CI = 1.17 to 1.46) [32]. The association at rs2046210 has been replicated in European populations [67] but not in African populations, and other SNPs have not yet been studied. However, subsequent studies have found associations for different SNPs within this region with breast cancer risk in African populations. SNP rs9397435 (intergenic, 8.9 kb 3' of *CCDC170*) has an OR of 1.34 (95% CI = 1.10 to 1.63) in Africans [99], and rs2046211 (intergenic, 6 kb 3' of *CCDC170*) has a protective effect with an OR of 0.80 (95% CI = 0.67 to 0.95) (note that none of these SNPs are in linkage disequilibrium but are within a 35 kb region near *CCDC170*) [68].

Common risk alleles in other populations

A small number of studies have begun to look for common breast cancer risk variants in other populations. Orr and colleagues performed a GWAS in European male breast cancer patients and found that SNPs at rs1314913 (14q24.1, *RAD51B* c.757-59007C > T) and rs3803662 (16q12.1, intergenic) were significantly associated with cancer risk [70]. Rafiq and colleagues performed a GWAS in women with breast cancer under the age of 40 and found an association of the SNP rs421379 (5q14, intergenic) with breast cancer prognosis [100]. However, both of these studies were carried out on small numbers of cases (~500) and therefore require independent validation. Gold and colleagues studied women of Ashkenazi Jewish descent and found an association with breast cancer risk and the SNP rs6569479 (6q22.33, *RNF146* c.3-1173 T > C) [64]. However, SNP rs2180341 (*RNF146* c.-108-746G > A), which is in linkage disequilibrium with rs6569479, was not significantly associated with breast cancer risk in the

BCAC population [57]. A weak association was seen in *BRCA1/2* carriers; however, this genomic region was not included in the recent COGS studies [40,41,65]. This example demonstrates the difficulty of studying SNPs in populations with limited numbers of individuals and demonstrates the need for large collaborative consortia to pool resources to study common variants in such populations.

Common low penetrance risk alleles modify breast cancer risk in *BRCA1* and *BRCA2* patients

Individuals with *BRCA1* or *BRCA2* germline mutations have a 60 to 80% and 40 to 60% risk of developing breast cancer in their lifetime, respectively [12,101]. Multiple environmental or genetic factors may be responsible for the variable penetrance seen within individuals and families; one possible modifier of risk is the presence of common low-risk breast cancer susceptibility variants. Study of the genetic variation influencing the risk of breast cancer and ovarian cancer in *BRCA1/2* mutation carriers has been mainly carried out by the Consortium of Investigators of Modifiers of *BRCA1/2* (CIMBA), an international group of over 40 institutions [102]. Identification of genetic loci associated with a decreased or increased risk of breast cancer in *BRCA1/2* carriers may allow the development of personalized risk prediction models. A number of candidate gene studies of the established breast cancer loci have been carried out in female *BRCA1/2* mutation carriers as part of the CIMBA effort [38,45,54,60,77,80,103].

In addition, two multistage GWAS had been carried out in *BRCA1/2* carriers prior to COGS (Table 2). Gaudet and colleagues genotyped 529,163 SNPs first in 809 *BRCA2* carriers with breast cancer (cases) and 804 *BRCA2* carriers without breast cancer (controls), followed by genotyping of 113 SNPs in 1,263 cases and 1,222 controls [78]. From this study, significant associations were found for three established breast cancer loci: rs2981582 (10q26.13, *FGFR2* c.109 + 906 T > C), rs3803662 (16q12.1, intergenic, 5' to *TOX3*), and rs16917302 (10q21.2 *ZNF365* c.981 + 41642A > C). In addition, the novel SNP rs311499 (intergenic, 1.4 kb 3' - *GMEB2*) was identified as significantly associated with breast cancer in *BRCA2* carriers. In *BRCA1* mutation carriers, Antoniou and colleagues first genotyped 1,190 *BRCA1* mutation carriers with breast cancer and 1,193 carriers without breast cancer and then studied identified SNPs in unselected cases from both the United Kingdom and the TNBC consortium, identifying rs8170 (19p13.1, *BABAMI* c.837G > A) as modifying the risk of breast cancer in these populations [51].

The COGS studies have now clarified the associations of the established breast cancer SNPs with breast cancer risk in *BRCA1/2* carriers and identified some novel associations (Table 2). Couch and colleagues [40] and Gaudet and colleagues [41] genotyped 11,705 *BRCA1* carriers and 8,211 *BRCA2* carriers, respectively, from CIMBA.

Comparing *BRCA1* mutation carriers with breast cancer with those carriers without, associations with risk of breast cancer were confirmed for three of the established breast cancer loci that have been prior associated with breast cancer risk in *BRCA1* carriers, namely: rs2046210 (6q25, intergenic, 6 kb 3' to *CCDC170*), rs10771399 (12p11.22, intergenic, 29 kb 5' to *PTHLH*) and rs8170 (19p13.1, *BABAMI* c.837G > A) [40]. Associations were borderline for rs3803662 (16q12.1, intergenic, 5' to *TOX3*) and non-significant for rs13387042 (2q35, intergenic). For *BRCA2* mutation carriers, associations with risk of breast cancer were confirmed for five of the established risk loci, namely: rs4973768 (3p24.1, *SLC4A7* c.*2242G > A), rs2420946 (10q26.13, *FGFR2* c.109 + 1899A > G), rs16917302 (10q21.2 *ZNF365* c.981 + 41642A > C), rs3817198 (11p15.5, *LSP1* c.*13 + 200 T > C), and rs3806332 (16q12.1, intergenic, 5' to *TOX3*) [41]. Six chromosomal regions that had not previously been significantly associated with breast cancer risk in *BRCA2* mutation carriers, probably due to inadequate power, were found to contain SNPs with significant associations, namely: rs27633 (12p11.22, *PTHLH* c.-266 + 555A > C), rs1688611 (5q11.2, intergenic, 5' to *MAP3K1*), rs10965163 (9p21.3, *MTAP* c.561C > T), rs4733664 (8q24.21, intergenic), rs13039229 (20q13.33, intergenic, 2.1 kb 3' of *PTK6*), and rs2253407 (6q25.1, *SYNE1* c.23020-466C > A). SNP associations were nonsignificant with breast cancer risk for nine of the established breast cancer SNPs. Importantly, four SNPs were newly found to be associated with breast cancer risk in *BRCA1* mutation carriers: rs2290854 (1q32.1, *MDM4* c.903 + 20A > G), rs6682208 (1q32.1, intergenic, 3' to *MDM4*), rs11196174 (10q25.3, *TCF7L2* c.381 + 22730A > G), and rs11196175 (10q25.3, *TCF7L2* c.381 + 25248 T > C). In *BRCA2* mutation carriers, three novel SNP associations with breast cancer risk were identified: rs184577 (2p22.2, intergenic), rs9348512 (6p24.3, intergenic), and rs619373 (Xq27.1, *FGF13* c.50-73946C > T).

Fine mapping of genomic regions – towards assigning functionality to common risk variants

The COGS efforts to fine map a number of identified loci is one of the next crucial steps in the study of common breast cancer risk variants, namely the functional characterization of identified SNPs [59,85]. A number of SNPs within or near the *TERT* genomic region have been associated with a risk of a variety of cancers [104], including breast cancer and ovarian cancer. To study the *TERT* region in breast cancer risk, 110 SNPs on the iCOGS array were genotyped in 46,451 women with breast cancer and 42,599 controls from the BCAC and in 11,709 *BRCA1* mutation carriers from CIMBA; 7,435 cases were ER-negative and 27,074 cases were ER-positive [59]. The authors identified three regions within the *TERT* locus on 5p15.33 and found that the SNP most strongly associated with risk depended on the type of breast cancer. The six

SNPs in linkage disequilibrium at $r^2 > 0.6$ in the first region corresponding to the *TERT* promoter (that is, rs2736107) were associated with breast cancer risk overall. In regions two and three, which correspond to introns 2 to 4, the five SNPs in linkage disequilibrium at $r^2 > 0.6$ in region two (for example, rs7734992), and the three SNPs in region three (rs72709458, rs2242652, and rs10069690) were associated with a risk of overall breast cancer, ER-negative breast cancer and cancer in *BRCA1* carriers, but showed only a borderline association with ER-positive breast cancer risk. The group further characterized potential functionality of the SNPs by investigating whether they were found in open chromatin regions using *in silico* ENCODE data, and whether they had effects on *TERT* promoter activity or splicing by *in vitro* assays.

French and colleagues performed fine-mapping and functional studies at the 11q13.3 region using 731 SNPs on the iCOGS array [85]. This study found three independent regions in the 11q13.3 chromosomal region, which were strongly associated only with ER-positive breast cancer and not with ER-negative breast cancer risk. Through functional studies, French and colleagues also provided evidence that the SNPs within this region influence the risk of ER-positive breast cancer via reduction of transcriptional activation of *CCND1*. This study provides a model for the necessary future studies of the functional consequences of the extensive list of common low-risk variants that have now been identified from GWAS. It is hoped that functional studies of these SNPs may lead to a better understanding of breast cancer pathogenesis and eventual development of novel prevention strategies or treatments for breast cancer.

Using analysis of common breast cancer risk alleles in clinical practice

The identification of mutations in high penetrance breast cancer susceptibility genes in women allows for the delineation of individual risk and assists in guiding clinical recommendations, such as enhanced breast cancer screening (for example, magnetic resonance imaging screening starting at age 25) or chemoprevention with tamoxifen [2]. However, the majority of familial breast cancer patients are *BRCA1/2* mutation-negative. In these women, single or multiple low penetrance and/or moderate penetrance genetic variants may be responsible for an increased risk of breast cancer. Currently, genotyping of common breast cancer risk variants are not incorporated into breast cancer risk assessment models.

In clinical practice, breast cancer risk is predicted using models such as the National Cancer Institute's Breast Cancer Risk Assessment Tool based on the Gail model [105]. This model determines a woman's risk of breast cancer based on ethnicity, personal and family history of breast cancer, age of first menstruation, age of first live

birth and number of breast biopsies. As the model has been validated to accurately predict breast cancer risk in Caucasian, African and Asian American women [106-108], any risk assessment incorporating SNP genotyping must improve upon this well-validated model.

Using computer simulation, Gail demonstrated that adding genotype information from seven SNPs (Table 3, polygenic risk score (PRS) model D) slightly improved the discriminatory accuracy to predict breast cancer risk, as measured by the area under the curve, but less so than adding mammographic density [109]. Gail predicted that risk categories for women would minimally change with the inclusion of SNP data [91]. In order to study the addition of SNP genotyping to breast cancer risk models in human subjects, Wacholder and colleagues performed SNP genotyping of 10 SNPs (eight established SNPs, two others; Table 3, PRS model C) in 5,590 cases and 5,998 controls [90]. The cases were women with breast cancer between the ages of 50 and 79, and 98.5% of patients were of European ancestry. Adding genotype information only minimally improved the area under the curve of the Gail model from 58.0% to 61.8%. Interestingly, however, 32.5% of patients were reclassified into the highest risk quintile and 20.4% were reclassified into the lowest risk quintile. To test the hypothesis that additional SNPs may improve the performance of SNP genotyping, Husing and colleagues analyzed 32 SNPs (24 established SNPs, eight others; Table 3, PRS model B) in 6,009 European postmenopausal breast cancer cases and 7,827 controls [89]. Adding additional SNPs modestly improved the discriminatory accuracy of classic risk models, but again probably would not be of sufficient clinical benefit to justify genotyping costs.

It is possible that SNP genotyping did not improve breast cancer risk prediction in the above studies because they were performed in unselected patient populations. Sawyer and colleagues thus recently genotyped 22 breast cancer risk variants (Table 3, PRS model A) in 1,143 high risk women from a familial breast cancer clinic [88]. These women had on average 1.82 first-degree to third-degree relatives with breast cancer and a median age of 45 at diagnosis. Sawyer and colleagues calculated a log-additive score that weighted each risk allele by its odds ratio, called the PRS, and found this score significantly increased from 0 (95% CI = -0.03 to 0.03) in controls to 0.30 (95% CI = 0.26 to 0.33, $P = 2.4 \times 10^{-29}$) in *BRCA1/2* mutation-negative familial breast cancer cases. The PRS was significantly associated with breast cancer diagnosis under age 35 and contralateral breast cancer risk. The PRS could therefore possibly reclassify young women with a family history of breast cancer into a risk category for which intensive breast cancer screening or chemoprevention with tamoxifen might be recommended. In addition, the PRS may help predict contralateral breast cancer risk in

women with a breast cancer diagnosis, therefore assisting in the decision for or against bilateral mastectomy.

In summary, genotyping of women with familial breast cancer for common low penetrance variants may have a significant impact on risk assessment, particularly in women for whom a high penetrance mutation is not found. However, prospective trials using a PRS that includes the recent SNPs contributed by COGS will need to be conducted before this strategy can be employed routinely. The outcomes of the trials must not only include an indication of the ability of SNP genotyping to improve risk stratification, but also determine whether reassigning patients to different risk strata affects endpoints such as breast cancer detection rates, breast cancer-related morbidity and mortality, patient quality of life, or overall healthcare costs. It is important to note that these studies have so far been performed on individuals of European descent using SNPs associated with risk and therefore one must be extremely cautious as any model generated may not accurately predict risk in other ethnicities.

Conclusions

Variants in multiple genes and intergenic regions are associated with a risk of developing breast cancer. These variants range from rare mutations that confer a high risk of breast cancer to common SNPs that confer individually minimally increased or decreased risks of breast cancer and probably act in a polygenic manner. Multistage GWAS in addition to smaller association, linkage and candidate gene studies have led to the publication of thousands of variants, but only a relatively small set are likely to be truly related to breast cancer risk. Associations found in one ethnicity cannot be generalized across all ethnicities, meaning that findings in one population need to be specifically validated in another. In addition, certain common risk variants are likely to be important only in certain breast cancer subtypes, such as ER-positive or ER-negative breast cancer, *BRCA1* or *BRCA2* carriers, or male breast cancer patients. Although the identification of common breast cancer risk variants has so far had little clinical impact, recent studies pre-COGS demonstrate that it may be possible to use carefully designed SNP genotyping panels to augment breast cancer risk models for high-risk populations such as those with familial breast cancer or to predict contralateral breast cancer risk. The recent COGS studies show that approximately 28% of familial breast cancer risk is accounted for by common variants. A thorough understanding of the involved SNPs may allow them to be used in risk stratification. Given the continual expansion of technology and decreases in cost in SNP genotyping, it may now be possible to develop a low-cost SNP genotyping tool that can be used in conjunction with standard risk assessment models. These tools must be tested prospectively and across ethnicities to demonstrate clinical utility.

Abbreviations

BCAC: Breast cancer association consortium; CI: Confidence interval; CIMBA: Consortium of investigators of modifiers of *BRCA1/2*; COGS: Collaborative oncological gene-environment Study; ER: Estrogen receptor; GWAS: Genome-wide association studies; OR: Odds ratio; PRS: Polygenic risk score; SNP: Single nucleotide polymorphism; TNBC: Triple-negative breast cancer.

Competing interests

The authors declare that they have no competing interests.

Acknowledgements

KNM is supported by the National Institutes of Health T32 training grant (5T32GM008638-15), the 2013 American Association for Cancer Research Fellowship in Translational Research, and the Department of Defense Breast Cancer Research Program (DOD BC123126). KLN is supported by grants from the Breast Cancer Research Foundation (BCRF), Rooney Family Foundation, the National Institutes of Health (R01 CA114478, R01 CA164947), and the Bassler Center for BRCA Research at the University of Pennsylvania.

Author details

¹Division of Hematology/Oncology, Department of Medicine, Perelman School of Medicine at the University of Pennsylvania, Philadelphia, PA 19104, USA. ²Division of Translational Medicine and Human Genetics, Department of Medicine, Perelman School of Medicine at the University of Pennsylvania, 351 BRB 2/3, 421 Curie Blvd, Philadelphia, PA 19104, USA. ³Abramson Cancer Center, Perelman School of Medicine at the University of Pennsylvania, Philadelphia, PA 19104, USA.

Published: 20 Dec 2013

References

- Howlander N, Noone AM, Krapcho M, Neyman N, Aminou R, Altekruse SF, Kosary CL, Ruhl J, Tatalovich Z, Cho H, Mariotto A, Eisner MP, Lewis DR, Chen HS, Feuer EJ, Cronin KA: (Eds): *SEER Cancer Statistics Review, 1975–2009*. Bethesda, MD: National Cancer Institute; 2011.
- Daly MB, Axilbund JE, Buys S, Crawford B, Farrell CD, Friedman S, Garber JE, Goorha S, Gruber SB, Hampel H, Kaklamani V, Kohlmann W, Kurian A, Litton J, Marcom PK, Nussbaum R, Offit K, Pal T, Pasche B, Pilarski R, Reiser G, Shannon KM, Smith JR, Swisher E, Weitzel JN: **Genetic/familial high-risk assessment: breast and ovarian**. *J Natl Compr Canc Network* 2010, **8**:562–594.
- Foulkes WD: **Inherited susceptibility to common cancers**. *N Engl J Med* 2008, **359**:2143–2153.
- Pharoah PD, Antoniou A, Bobrow M, Zimmern RL, Easton DF, Ponder BA: **Polygenic susceptibility to breast cancer and implications for prevention**. *Nat Genet* 2002, **31**:33–36.
- Gudmundsdottir K, Ashworth A: **The roles of BRCA1 and BRCA2 and associated proteins in the maintenance of genomic stability**. *Oncogene* 2006, **25**:5864–5874.
- O'Donovan PJ, Livingston DM: **BRCA1 and BRCA2: breast/ovarian cancer susceptibility gene products and participants in DNA double-strand break repair**. *Carcinogenesis* 2010, **31**:961–967.
- Maxwell KN, Domchek SM: **Cancer treatment according to BRCA1 and BRCA2 mutations**. *Nat Rev Clin Oncol* 2012, **9**:520–528.
- Domchek SM, Friebel TM, Singer CF, Evans DG, Lynch HT, Isaacs C, Garber JE, Neuhausen SL, Matloff E, Eeles R, Pichert G, Vant'Veer L, Tung N, Weitzel JN, Couch FJ, Rubinstein WS, Ganz PA, Daly MB, Olopade OI, Tomlinson G, Schildkraut J, Blum JL, Rebbeck TR: **Association of risk-reducing surgery in BRCA1 or BRCA2 mutation carriers with cancer risk and mortality**. *JAMA* 2010, **304**:967–975.
- Peto J, Collins N, Barfoot R, Seal S, Warren W, Rahman N, Easton DF, Evans C, Deacon J, Stratton MR: **Prevalence of BRCA1 and BRCA2 gene mutations in patients with early-onset breast cancer [see comments]**. *J Natl Cancer Inst* 1999, **91**:943–949.
- Spurdle AB, Whaley PJ, Thompson B, Feng B, Healey S, Brown MA, Pettigrew C, ConFab K, Van Asperen CJ, Ausems MG, Kattentidt-Mouravieva AA, van den Ouweland AM, Dutch Belgium UV C, Lindblom A, Pigg MH, Schmutzler RK, Engel C, Meindl A, German Consortium of Hereditary Breast and Ovarian Cancer, Caputo S, Sinilnikova OM, Lidereau R, French COVAR group collaborators, Couch FJ, Guidugli L, Hansen T, Thomassen M, Eccles DM, Tucker K, et al: **BRCA1 R1699Q variant displaying ambiguous functional abrogation confers intermediate breast and ovarian cancer risk**. *J Med Genet* 2012, **49**:525–532.
- Prevalence and penetrance of BRCA1 and BRCA2 mutations in a population-based series of breast cancer cases: **Anglian Breast Cancer Study Group**. *Br J Cancer* 2000, **83**:1301–1308.
- Antoniou A, Pharoah PD, Narod S, Risch HA, Eyfjord JE, Hopper JL, Loman N, Olsson H, Johannsson O, Borg A, Pasini B, Radice P, Manoukian S, Eccles DM, Tang N, Olah E, Anton-Culver H, Warner E, Lubinski J, Gronwald J, Gorski B, Tulinius H, Thorlacius S, Eerola H, Nevanlinna H, Syrjäsoski K, Kallioniemi OP, Thompson D, Evans C, Peto J, et al: **Average risks of breast and ovarian cancer associated with BRCA1 or BRCA2 mutations detected in case series unselected for family history: a combined analysis of 22 studies**. *Am J Hum Genet* 2003, **72**:1117–1130.
- Gonzalez KD, Buzin CH, Noltner KA, Gu D, Li W, Malkin D, Sommer SS: **High frequency of de novo mutations in Li-Fraumeni syndrome**. *J Med Genet* 2009, **46**:689–693.
- Hwang SJ, Lozano G, Amos CI, Strong LC: **Germline p53 mutations in a cohort with childhood sarcoma: sex differences in cancer risk**. *Am J Hum Genet* 2003, **72**:975–983.
- Nelen MR, Kremer H, Konings IB, Schoute F, van Essen AJ, Koch R, Woods CG, Frys JP, Hamel B, Hoefsloot LH, Peeters EA, Padberg GW: **Novel PTEN mutations in patients with Cowden disease: absence of clear genotype-phenotype correlations**. *Eur J Hum Genet* 1999, **7**:267–273.
- Schrager CA, Schneider D, Gruener AC, Tsou HC, Peacocke M: **Clinical and pathological features of breast disease in Cowden's syndrome: an underrecognized syndrome with an increased risk of breast cancer**. *Hum Pathol* 1998, **29**:47–53.
- Izatt L, Greenman J, Hodgson S, Ellis D, Watts S, Scott G, Jacobs C, Liebmann R, Zvelebil MJ, Mathew C, Solomon E: **Identification of germline missense mutations and rare allelic variants in the ATM gene in early-onset breast cancer**. *Genes Chromosomes Cancer* 1999, **26**:286–294.
- 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, Elstrodt F, van Duijn C, Bartels C, Meijers C, Schutte M, McGuffog L, Thompson D, Easton D, Sodha N, Seal S, Barfoot R, Mangion J, Chang-Claude J, Eccles D, Eeles R, Evans DG, Houlston R, Murdav V, Narod S, Peretz T, 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.
- Rahman N, Seal S, Thompson D, Kelly P, Renwick A, Elliott A, Reid S, Spanova K, Barfoot R, Chagtai T, Jayatilake H, McGuffog L, Hanks S, Evans DG, Eccles D, Easton DF, Stratton MR: **PALB2, which encodes a BRCA2-interacting protein, is a breast cancer susceptibility gene**. *Nat Genet* 2007, **39**:165–167.
- Seal S, Thompson D, Renwick A, Elliott A, Kelly P, Barfoot R, Chagtai T, Jayatilake H, Ahmed M, Spanova K, North B, McGuffog L, Evans DG, Eccles D, Easton DF, Stratton MR, Rahman N: **Truncating mutations in the Fanconi anemia J gene BRIP1 are low-penetrance breast cancer susceptibility alleles**. *Nat Genet* 2006, **38**:1239–1241.
- Houlston RS, Peto J: **The search for low-penetrance cancer susceptibility alleles**. *Oncogene* 2004, **23**:6471–6476.
- 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* 1999, **8**:843–854.
- Breast Cancer Association Consortium: **Commonly studied single-nucleotide polymorphisms and breast cancer: results from the Breast Cancer Association Consortium**. *J Natl Cancer Inst* 2006, **98**:1382–1396.
- Cox A, Dunning AM, Garcia-Closas M, Balasubramanian S, Reed MW, Pooley KA, Scollen S, Baynes C, Ponder BA, Chanock S, Lissowska J, Brinton L, Peplonska B, Southey MC, Hopper JL, McCredie MR, Giles GG, Fletcher O, Johnson N, dos Santos SI, Gibson L, Bojesen SE, Nordestgaard BG, Axelsson CK, Torres D, Hamann U, Justenhoven C, Brauch H, Chang-Claude J, Kropp S, et al: **A common coding variant in CASP8 is associated with breast cancer risk**. *Nat Genet* 2007, **39**:352–358.
- Easton DF, Pooley KA, Dunning AM, Pharoah PD, Thompson D, Ballinger DG, Struwing JP, Morrison J, Field H, Luben R, Wareham N, Ahmed S, Healey CS, Bowman R, Meyer KB, Haiman CA, Kolonel LK, Henderson BE, Le Marchand L, Brennan P, Sangrajrang S, Gaborieau V, Odey F, Shen CY, Wu PE, Wang HC, Eccles D, Evans DG, Peto J, Fletcher O, et al: **Genome-wide association study identifies novel breast cancer susceptibility loci**. *Nature* 2007, **447**:1087–1093.
- Hunter DJ, Kraft P, Jacobs KB, Cox DG, Yeager M, Hankinson SE, Wacholder S, Wang Z, Welch R, Hutchinson A, Wang J, Yu K, Chatterjee N, Orr N, Willett

- WC, Colditz GA, Ziegler RG, Berg CD, Buys SS, McCarty CA, Feigelson HS, Calle EE, Thun MJ, Hayes RB, Tucker M, Gerhard DS, Fraumeni Jr JF, Hoover RN, Thomas G, Chanock SJ: **A genome-wide association study identifies alleles in FGFR2 associated with risk of sporadic postmenopausal breast cancer.** *Nat Genet* 2007, **39**:870–874.
27. Stacey SN, Manolescu A, Sulem P, Rafnar T, Gudmundsson J, Gudjonsson SA, Masson G, Jakobsdottir M, Thorlacius S, Helgason A, Aben KK, Strobbe LJ, Albers-Akkers MT, Swinkels DW, Henderson BE, Kolonel LN, Le Marchand L, Millastre E, Andres R, Saez B, Lambea J, Godino J, Polo E, Tres A, Mouy M, Saemundsdottir J, Backman VM, Gudmundsson L, Kristjansson K, Bergthorsson JT, Kostic J, et al: **Common variants on chromosomes 2q35 and 16q12 confer susceptibility to estrogen receptor-positive breast cancer.** *Nat Genet* 2007, **39**:865–869.
28. Stacey SN, Manolescu A, Sulem P, Thorlacius S, Gudjonsson SA, Jonsson GF, Jakobsdottir M, Bergthorsson JT, Gudmundsson J, Aben KK, Strobbe LJ, Swinkels DW, van Engelenburg KC, Henderson BE, Kolonel LN, Le Marchand L, Millastre E, Andres R, Saez B, Lambea J, Godino J, Polo E, Tres A, Picelli S, Rantala J, Margolin S, Jonsson T, Sigurdsson H, Jonsdottir T, Hrafnkelsson J, et al: **Common variants on chromosome 5p12 confer susceptibility to estrogen receptor-positive breast cancer.** *Nat Genet* 2008, **40**:703–706.
29. Zheng W, Long J, Gao YT, Li C, Zheng Y, Xiang YB, Wen W, Levy S, Deming SL, Haines JL, Gu K, Fair AM, Cai Q, Lu W, Shu XO: **Genome-wide association study identifies a new breast cancer susceptibility locus at 6q25.1.** *Nat Genet* 2009, **41**:324–328.
30. Thomas G, Jacobs KB, Kraft P, Yeager M, Wacholder S, Cox DG, Hankinson SE, Hutchinson A, Wang Z, Yu K, Chatterjee N, Garcia-Closas M, Gonzalez-Bosquet J, Prokunina-Olsson L, Orr N, Willett WC, Colditz GA, Ziegler RG, Berg CD, Buys SS, McCarty CA, Feigelson HS, Calle EE, Thun MJ, Diver R, Prentice R, Jackson R, Kooperberg C, Chlebowski R, Lissowska J, 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* 2009, **41**:579–584.
31. Ahmed S, Thomas G, Ghousaini M, Healey CS, Humphreys MK, Platte R, Morrison J, Maranian M, Pooley KA, Luben R, Eccles D, Evans DG, Fletcher O, Johnson N, Dos Santos Silva I, Peto J, Stratton MR, Rahman N, Jacobs K, Prentice R, Anderson GL, Rajkovic A, Curb JD, Ziegler RG, Berg CD, Buys SS, McCarty CA, Feigelson HS, Calle EE, Thun MJ, et al: **Newly discovered breast cancer susceptibility loci on 3p24 and 17q23.2.** *Nat Genet* 2009, **41**:585–590.
32. Turnbull C, Ahmed S, Morrison J, Pernet D, Renwick A, Maranian M, Seal S, Ghousaini M, Hines S, Healey CS, Hughes D, Warren-Perry M, Tapper W, Eccles D, Evans DG, Hooning M, Schutte M, van den Ouweland A, Houlston R, Ross G, Langford C, Pharoah PD, Stratton MR, Dunning AM, Rahman N, Easton DF: **Genome-wide association study identifies five new breast cancer susceptibility loci.** *Nat Genet* 2010, **42**:504–507.
33. Fletcher O, Johnson N, Orr N, Hosking FJ, Gibson LJ, Walker K, Zelenika D, Gut I, Heath S, Palles C, Coupland B, Broderick P, Schoemaker M, Jones M, Williamson J, Chilcott-Burns S, Tomczyk K, Simpson G, Jacobs KB, Chanock SJ, Hunter DJ, Tomlinson IP, Swerdlow A, Ashworth A, Ross G, Dos Santos Silva I, Lathrop M, Houlston RS, Peto J: **Novel breast cancer susceptibility locus at 9q31.2: results of a genome-wide association study.** *J Natl Cancer Inst* 2011, **103**:425–435.
34. Ghousaini M, Fletcher O, Michailidou K, Turnbull C, Schmidt MK, Dicks E, Dennis J, Wang Q, Humphreys MK, Luccarini C, Baynes C, Conroy D, Maranian M, Ahmed S, Driver K, Johnson N, Orr N, dos Santos SI, Waisfisz Q, Meijers-Heijboer H, Uitterlinden AG, Rivadeneira F, Hall P, Czene K, Irwanto A, Liu J, Nevanlinna H, Aittomaki K, Blomqvist C, Meindl A, et al: **Genome-wide association analysis identifies three new breast cancer susceptibility loci.** *Nat Genet* 2012, **44**:312–318.
35. Michailidou K, Hall P, Gonzalez-Neira A, Ghousaini M, Dennis J, Milne RL, Schmidt MK, Chang-Claude J, Bojesen SE, Bolla MK, Wang Q, Dicks E, Lee A, Turnbull C, Rahman N, Fletcher O, Peto J, Gibson L, Dos Santos SI, Nevanlinna H, Muranen TA, Aittomaki K, Blomqvist C, Czene K, Irwanto A, Liu J, Waisfisz Q, Meijers-Heijboer H, Adank M, van der Luijt RB, et al: **Large-scale genotyping identifies 41 new loci associated with breast cancer risk.** *Nat Genet* 2013, **45**:353–361.
36. Figueroa JD, Garcia-Closas M, Humphreys M, Platte R, Hopper JL, Southey MC, Apicella C, Hammet F, Schmidt MK, Broeks A, Tollenaar RA, Van't Veer LJ, Fasching PA, Beckmann MW, Ekici AB, Strick R, Peto J, Dos Santos Silva I, Fletcher O, Johnson N, Sawyer E, Tomlinson I, Kerin M, Burwinkel B, Marme F, Schneeweiss A, Sohn C, Bojesen S, Flyger H, Nordestgaard BG, et al: **Associations of common variants at 1p11.2 and 14q24.1 (RAD51L1) with breast cancer risk and heterogeneity by tumor subtype: findings from the Breast Cancer Association Consortium.** *Hum Mol Genet* 2011, **20**:4693–4706.
37. Zheng W, Zhang B, Cai Q, Sung H, Michailidou K, Shi J, Choi JY, Long J, Dennis J, Humphreys MK, Wang Q, Lu W, Gao YT, Li C, Cai H, Park SK, Yoo KY, Noh DY, Han W, Dunning AM, Benitez J, Vincent D, Bacot F, Tessier D, Kim SW, Lee MH, Lee JW, Lee JY, Xiang YB, Zheng Y, et al: **Common genetic determinants of breast-cancer risk in East Asian women: a collaborative study of 23 637 breast cancer cases and 25 579 controls.** *Hum Mol Genet* 2013, **22**:2539–2550.
38. Antoniou AC, Kartsonaki C, Sinilnikova OM, Soucy P, McGuffog L, Healey S, Lee A, Peterlongo P, Manoukian S, Peissel B, Zaffaroni D, Cattaneo E, Barile M, Pensotti V, Pasini B, Dolcetti R, Giannini G, Putignano AL, Varesco L, Radice P, Mai PL, Greene MH, Andrulis IL, Glendon G, Ozelik H, Thomassen M, Gerdes AM, Kruse TA, Birk Jensen U, Cruger DG, et al: **Common alleles at 6q25.1 and 1p11.2 are associated with breast cancer risk for BRCA1 and BRCA2 mutation carriers.** *Hum Mol Genet* 2011, **20**:3304–3321.
39. Garcia-Closas M, Couch FJ, Lindstrom S, Michailidou K, Schmidt MK, Brook MN, Orr N, Rhie SK, Riboli E, Feigelson HS, Le Marchand L, Buring JE, Eccles D, Miron P, Fasching PA, Brauch H, Chang-Claude J, Carpenter J, Godwin AK, Nevanlinna H, Giles GG, Cox A, Hopper JL, Bolla MK, Wang Q, Dennis J, Dicks E, Howat WJ, Schoof N, Bojesen SE, et al: **Genome-wide association studies identify four ER negative-specific breast cancer risk loci.** *Nat Genet* 2013, **45**:392–398.
40. Couch FJ, Wang X, McGuffog L, Lee A, Olsword C, Kuchenbaecker KB, Soucy P, Fredericksen Z, Barrowdale D, Dennis J, Gaudet MM, Dicks E, Kosel M, Healey S, Sinilnikova OM, Bacot F, Vincent D, Hogervorst FB, Peock S, Stoppa-Lyonnet D, Jakubowska A, K ConFab Investigators, Radice P, Schmutzler RK, Domchek SM, Piedmonte M, Singer CF, Friedman E, Thomassen M, Hansen TV, et al: **Genome-wide association study in BRCA1 mutation carriers identifies novel loci associated with breast and ovarian cancer risk.** *PLoS Genet* 2013, **9**:e1003212.
41. Gaudet MM, Kuchenbaecker KB, Vijai J, Klein RJ, Kirchoff T, McGuffog L, Barrowdale D, Dunning AM, Lee A, Dennis J, Healey S, Dicks E, Soucy P, Sinilnikova OM, Pankratz VS, Wang X, Eldridge RC, Tessier DC, Vincent D, Bacot F, Hogervorst FB, Peock S, Stoppa-Lyonnet D, Peterlongo P, Schmutzler RK, Nathanson KL, Piedmonte M, Singer CF, Thomassen M, Hansen TV, et al: **Identification of a BRCA2-specific modifier locus at 6p24 related to breast cancer risk.** *PLoS Genet* 2013, **9**:e1003173.
42. Long J, Zhang B, Signorello LB, Cai Q, Deming-Halverson S, Shrubsole MJ, Sanderson M, Dennis J, Michailiou K, Easton DF, Shu XO, Blot WJ, Zheng W: **Evaluating genome-wide association study-identified breast cancer risk variants in African-American women.** *PLoS One* 2013, **8**:e58350.
43. Sargentanis TN, Economopoulos KP: **Association of two CASP8 polymorphisms with breast cancer risk: a meta-analysis.** *Breast Cancer Res Treat* 2010, **120**:229–234.
44. Broeks A, Schmidt MK, Sherman ME, Couch FJ, Hopper JL, Dite GS, Apicella C, Smith LD, Hammet F, Southey MC, Veer LJ V t, de Groot R, Smit VT, Fasching PA, Beckmann MW, Jud S, Ekici AB, Hartmann A, Hein A, Schulz-Wendtland R, Burwinkel B, Marme F, Schneeweiss A, Sinn HP, Sohn C, Tchatchou S, Bojesen SE, Nordestgaard BG, Flyger H, Orsted DD, et al: **Low penetrance breast cancer susceptibility loci are associated with specific breast tumor subtypes: findings from the Breast Cancer Association Consortium.** *Hum Mol Genet* 2011, **20**:3289–3303.
45. Antoniou AC, Sinilnikova OM, McGuffog L, Healey S, Nevanlinna H, Heikkinen T, Simard J, Spurdle AB, Beesley J, Chen X, Neuhausen SL, Ding YC, Couch FJ, Wang X, Fredericksen Z, Peterlongo P, Peissel B, Bonanni B, Viel A, Bernard L, Radice P, Szabo CI, Foretova L, Zikan M, Claes K, Greene MH, Mai PL, Rennert G, Lejbkowitz F, Andrulis IL, et al: **Common variants in LSP1, 2q35 and 8q24 and breast cancer risk for BRCA1 and BRCA2 mutation carriers.** *Hum Mol Genet* 2009, **18**:4442–4456.
46. Reeves GK, Travis RC, Green J, Bull D, Tipper S, Baker K, Beral V, Peto R, Bell J, Zelenika D, Lathrop M: **Incidence of breast cancer and its subtypes in relation to individual and multiple low-penetrance genetic susceptibility loci.** *JAMA* 2010, **304**:426–434.
47. Higginbotham KS, Breyer JP, Bradley KM, Schuyler PA, Plummer WD Jr, Freudenthal ME, Trentham-Dietz A, Newcomb PA, Sanders ME, Page DL, Parl FF, Egan KM, Dupont WD, Smith JR: **A multistage association study identifies a breast cancer genetic locus at NCOA7.** *Cancer Res* 2011, **71**:3881–3888.
48. Long J, Shu XO, Cai Q, Gao YT, Zheng Y, Li G, Li C, Gu K, Wen W, Xiang YB, Lu W, Zheng W: **Evaluation of breast cancer susceptibility loci in Chinese women.** *Cancer Epidemiol Biomarkers Prev* 2010, **19**:2357–2365.
49. Slattery ML, Baumgartner KB, Giuliano AR, Byers T, Herrick JS, Wolff RK: **Replication of five GWAS-identified loci and breast cancer risk among**

- Hispanic and non-Hispanic white women living in the Southwestern United States. *Breast Cancer Res Treat* 2011, **129**:531–539.
50. Zheng W, Cai Q, Signorello LB, Long J, Hargreaves MK, Deming SL, Li G, Li C, Cui Y, Blot WJ: **Evaluation of 11 breast cancer susceptibility loci in African-American women.** *Cancer Epidemiol Biomarkers Prev* 2009, **18**:2761–2764.
51. Antoniou AC, Wang X, Fredericksen ZS, McGuffog L, Tarrell R, Sinilnikova OM, Healey S, Morrison J, Kartsonaki C, Lesnick T, Ghousaini M, Barrowdale D, Peock S, Cook M, Oliver C, Frost D, Eccles D, Evans DG, Eeles R, Izatt L, Chu C, Douglas F, Paterson J, Stoppa-Lyonnet D, Houdayer C, Mazoyer S, Giraud S, Lasset C, Remenieras A, Caron O, *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* 2010, **42**:885–892.
52. Chen W, Zhong R, Ming J, Zou L, Zhu B, Lu X, Ke J, Zhang Y, Liu L, Miao X, Huang T: **The SLC4A7 variant rs4973768 is associated with breast cancer risk: evidence from a case-control study and a meta-analysis.** *Breast Cancer Res Treat* 2012, **136**:847–857.
53. Ruiz-Narvaez EA, Rosenberg L, Rotimi CN, Cupples LA, Boggs DA, Adeyemo A, Cozier YC, Adams-Campbell LL, Palmer JR: **Genetic variants on chromosome 5p12 are associated with risk of breast cancer in African American women: the Black Women's Health Study.** *Breast Cancer Res Treat* 2010, **123**:525–530.
54. Wang X, Pankratz VS, Fredericksen Z, Tarrell R, Karaus M, McGuffog L, Pharaoh PD, Stoppa-Lyonnet D, Dunning AM, Peock S, Cook M, Oliver C, Frost D, Sinilnikova OM, Stoppa-Lyonnet D, Mazoyer S, Houdayer C, Hogervorst FB, Hooning MJ, Ligtenberg MJ, Spurdle A, Chenevix-Trench G, Schmutzler RK, Wappenschmidt B, Engel C, Meindl A, Domchek SM, Nathanson KL, Rebbeck TR, Singer CF, *et al*: **Common variants associated with breast cancer in genome-wide association studies are modifiers of breast cancer risk in BRCA1 and BRCA2 mutation carriers.** *Hum Mol Genet* 2010, **19**:2886–2897.
55. Hutter CM, Young AM, Ochs-Balcom HM, Carty CL, Wang T, Chen CT, Rohan TE, Kooperberg C, Peters U: **Replication of breast cancer GWAS susceptibility loci in the Women's Health Initiative African American SHARE Study.** *Cancer Epidemiol Biomarkers Prev* 2011, **20**:1950–1959.
56. Kim HC, Lee JY, Sung H, Choi JY, Park SK, Lee KM, Kim YJ, Go MJ, Li L, Cho YS, Park M, Kim DJ, Oh JH, Kim JW, Jeon JP, Jeon SY, Min H, Kim HM, Park J, Yoo KY, Noh DY, Ahn SH, Lee MH, Kim SW, Lee JW, Park BW, Park WY, Kim EH, Kim MK, Han W, *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* 2012, **14**:R56.
57. Haiman CA, Chen GK, Vachon CM, Canzian F, Dunning A, Millikan RC, Wang X, Ademuyiwa F, Ahmed S, Ambrosone CB, Baglietto L, Balleine R, Bandera EV, Beckmann MW, Berg CD, Bernstein L, Blomqvist C, Blot WJ, Brauch H, Buring JE, Carey LA, Carpenter JE, Chang-Claude J, Chanock SJ, Chasman DI, Clarke CL, Cox A, Cross SS, Deming SL, Diasio RB, *et al*: **A common variant at the TERT-CLPTM1L locus is associated with estrogen receptor-negative breast cancer.** *Nat Genet* 2011, **43**:1210–1214.
58. Palmer JR, Ruiz-Narvaez EA, Rotimi CN, Cupples LA, Cozier YC, Adams-Campbell LL, Rosenberg L: **Genetic susceptibility loci for subtypes of breast cancer in an african american population.** *Cancer Epidemiol Biomarkers Prev* 2013, **22**:127–134.
59. Bojesen SE, Pooley KA, Johnatty SE, Beesley J, Michailidou K, Tyrer JP, Edwards SL, Pickett HA, Shen HC, Smart CE, Hillman KM, Mai PL, Lawrenson K, Stutz MD, Lu Y, Karevan R, Woods N, Johnston RL, French JD, Chen X, Weischer M, Nielsen SF, Maranian MJ, Ghousaini M, Ahmed S, Baynes C, Bolla MK, Wang Q, Dennis J, McGuffog L, *et al*: **Multiple independent variants at the TERT locus are associated with telomere length and risks of breast and ovarian cancer.** *Nat Genet* 2013, **45**:371–384.
60. Antoniou AC, Spurdle AB, Sinilnikova OM, Healey S, Pooley KA, Schmutzler RK, Vermold B, Engel C, Meindl A, Arnold N, Hofmann W, Sutter C, Niederacher D, Deissler H, Caldes T, Kampjarvi K, Nevanlinna H, Simard J, Beesley J, Chen X, Neuhausen SL, Rebbeck TR, Wagner T, Lynch HT, Isaacs C, Weitzel J, Ganz PA, Daly MB, Tomlinson G, Olopade OL, *et al*: **Common breast cancer-predisposition alleles are associated with breast cancer risk in BRCA1 and BRCA2 mutation carriers.** *Am J Hum Genet* 2008, **82**:937–948.
61. Lu PH, Yang J, Li C, Wei MX, Shen W, Shi LP, Jiang ZY, Zhou N, Tao GQ: **Association between mitogen-activated protein kinase kinase 1 rs889312 polymorphism and breast cancer risk: evidence from 59,977 subjects.** *Breast Cancer Res Treat* 2011, **126**:663–670.
62. Shan J, Mahfoudh W, Dsouza SP, Hassen E, Bouaouina N, Abdelhak S, Benhadjayed A, Memmi H, Mathew RA, Aigha II, Gabbouj S, Remadi Y, Chouchane L: **Genome-wide association studies (GWAS) breast cancer susceptibility loci in Arabs: susceptibility and prognostic implications in Tunisians.** *Breast Cancer Res Treat* 2012, **135**:715–724.
63. Siddiq A, Couch FJ, Chen GK, Lindstrom S, Eccles D, Millikan RC, Michailidou K, Stram DO, Beckmann L, Rhie SK, Ambrosone CB, Aittomaki K, Amiano P, Apicella C, Baglietto L, Bandera EV, Beckmann MW, Berg CD, Bernstein L, Blomqvist C, Brauch H, Brinton L, Bui QM, Buring JE, Buys SS, Campa D, Carpenter JE, Chasman DI, Chang-Claude J, Chen C, *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* 2012, **21**:5373–5384.
64. Gold B, Kirchoff T, Stefanov S, Lautenberger J, Viale A, Garber J, Friedman E, Narod S, Olshen AB, Gregersen P, Kosarin K, Olsh A, Bergeron J, Ellis NA, Klein RJ, Clark AG, Norton L, Dean M, Boyd J, Offit K: **Genome-wide association study provides evidence for a breast cancer risk locus at 6q22.33.** *Proc Natl Acad Sci U S A* 2008, **105**:4340–4345.
65. Kirchoff T, Gaudet MM, Antoniou AC, McGuffog L, Humphreys MK, Dunning AM, Bojesen SE, Nordestgaard BG, Flyger H, Kang D, Yoo KY, Noh DY, Ahn SH, Dork T, Schurmann P, Karstens JH, Hillemanns P, Couch FJ, Olson J, Vachon C, Wang X, Cox A, Brock I, Elliott G, Reed MW, Burwinkel B, Meindl A, Brauch H, Hamann U, Ko YD, *et al*: **Breast cancer risk and 6q22.33: combined results from Breast Cancer Association Consortium and Consortium of Investigators on Modifiers of BRCA1/2.** *PLoS One* 2012, **7**:e35706.
66. Stevens KN, Vachon CM, Lee AM, Slager S, Lesnick T, Olswold C, Fasching PA, Miron P, Eccles D, Carpenter JE, Godwin AK, Ambrosone C, Winqvist R, Brauch H, Schmidt MK, Cox A, Cross SS, Sawyer E, Hartmann A, Beckmann MW, Schulz-Wendtland R, Ekici AB, Tapper WJ, Gerty SM, Durcan L, Graham N, Hein R, Nickels S, Flesch-Janys D, Heinz J, *et al*: **Common breast cancer susceptibility loci are associated with triple-negative breast cancer.** *Cancer Res* 2011, **71**:6240–6249.
67. Hein R, Maranian M, Hopper JL, Kapuscinski MK, Southey MC, Park DJ, Schmidt MK, Broeks A, Hogervorst FB, Bueno-de-Mesquit HB, Muir KR, Lophatananon A, Rattanamongkongul S, Puttawibul P, Fasching PA, Hein A, Ekici AB, Beckmann MW, Fletcher O, Johnson N, dos Santos SI, Peto J, Sawyer E, Tomlinson I, Kerin M, Miller N, Marmee F, Schneeweiss A, Sohn C, Burwinkel B, *et al*: **Comparison of 6q25 breast cancer hits from Asian and European Genome Wide Association Studies in the Breast Cancer Association Consortium (BCAC).** *PLoS One* 2012, **7**:e42380.
68. Ruiz-Narvaez EA, Rosenberg L, Yao S, Rotimi CN, Cupples AL, Bandera EV, Ambrosone CB, Adams-Campbell LL, Palmer JR: **Fine-mapping of the 6q25 locus identifies a novel SNP associated with breast cancer risk in African-American women.** *Carcinogenesis* 2012, **34**:287–291.
69. Guo H, Ming J, Liu C, Li Z, Zhang N, Cheng H, Wang W, Shi W, Shen N, Zhao Q, Li D, Yi P, Wang L, Wang R, Xin Y, Zhao X, Nie X, Huang T: **A common polymorphism near the ESR1 gene is associated with risk of breast cancer: evidence from a case-control study and a meta-analysis.** *PLoS One* 2012, **7**:e52445.
70. Orr N, Lemnrau A, Cooke R, Fletcher O, Tomczyk K, Jones M, Johnson N, Lord CJ, Mitsopoulos C, Zvelebil M, McDade SS, Buck G, Blancher C, Trainer AH, James PA, Bojesen SE, Bokmand S, Nevanlinna H, Mattsson J, Friedman E, Laitman Y, Palli D, Masala G, Zanna I, Ottini L, Giannini G, Hollestelle A, Ouweland AM, Novakovic S, Krajac M, *et al*: **Genome-wide association study identifies a common variant in RAD51B associated with male breast cancer risk.** *Nat Genet* 2012, **44**:1182–1184.
71. McInerney N, Collieran G, Rowan A, Walther A, Barclay E, Spain S, Jones AM, Tuohy S, Curran C, Miller N, Kerin M, Tomlinson I, Sawyer E: **Low penetrance breast cancer predisposition SNPs are site specific.** *Breast Cancer Res Treat* 2009, **117**:151–159.
72. Antoniou AC, Wang X, Fredericksen ZS, McGuffog L, Tarrell R, Sinilnikova OM, Healey S, Morrison J, Kartsonaki C, Lesnick T, Ghousaini M, Barrowdale D, Embrace, Peock S, Cook M, Oliver C, Frost D, Eccles D, Evans DG, Eeles R, Izatt L, Chu C, Douglas F, Paterson J, Stoppa-Lyonnet D, Houdayer C, Mazoyer S, Giraud S, Lasset C, Remenieras A, *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* 2010, **42**:885–892.
73. Easton DF, Deffenbaugh AM, Pruss D, Frye C, Wenstrup RJ, Allen-Brady K, Tavtigian SV, Monteiro AN, Iversen ES, Couch FJ, Goldgar DE: **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* 2007, **81**:873–883.
74. Antoniou AC, Kuchenbaecker KB, Soucy P, Beesley J, Chen X, McGuffog L, Lee A, Barrowdale D, Healey S, Sinilnikova OM, Caligo MA, Loman N, Harbst K, Lindblom A, Arver B, Rosenquist R, Karlsson P, Nathanson K, Domchek S, Rebbeck T, Jakubowska A, Lubinski J, Jaworska K, Durda K, Zlowowcka-Perlowska E, Osorio A, Duran M, Andres R, Benitez J, Hamann U, et al: **Common variants at 12p11, 12q24, 9p21, 9q31.2 and in ZNF365 are associated with breast cancer risk for BRCA1 and/or BRCA2 mutation carriers.** *Breast Cancer Res* 2012, **14**:R33.
75. Warren H, Dudbridge F, Fletcher O, Orr N, Johnson N, Hopper JL, Apicella C, Southey MC, Mahmoodi M, Schmidt MK, Broeks A, Cornelissen S, Braaf LM, Muir KR, Lophatananon A, Chaiwerawattana A, Wiangnon S, Fasching PA, Beckmann MW, Ekici AB, Schulz-Wendtland R, Sawyer EJ, Tomlinson I, Kerin M, Burwinkel B, Marme F, Schneeweiss A, Sohn C, Guenel P, Truong T, et al: **9q31.2-rs865686 as a susceptibility locus for estrogen receptor-positive breast cancer: evidence from the Breast Cancer Association Consortium.** *Cancer Epidemiol Biomarkers Prev* 2012, **21**:1783–1791.
76. Chen F, Chen GK, Millikan RC, John EM, Ambrosone CB, Bernstein L, Zheng W, Hu JJ, Ziegler RG, Deming SL, Bandera EV, Nyante S, Palmer JR, Rebbeck TR, Ingles SA, Press MF, Rodriguez-Gil JL, Chanock SJ, Le Marchand L, Kolonel LN, Henderson BE, Stram DO, Haiman CA: **Fine-mapping of breast cancer susceptibility loci characterizes genetic risk in African Americans.** *Hum Mol Genet* 2011, **20**:4491–4503.
77. Couch FJ, Gaudet MM, Antoniou AC, Ramus SJ, Kuchenbaecker KB, Soucy P, Beesley J, Chen X, Wang X, Kirchoff T, McGuffog L, Barrowdale D, Lee A, Healey S, Sinilnikova OM, Andrulis IL, Ozcelik H, Mulligan AM, Thomassen M, Gerdes AM, Jensen UB, Skytte AB, Kruse TA, Caligo MA, von Wachenfeldt A, Barbany-Bustinza G, Loman N, Soller M, Ehrencrona H, Karlsson P, et al: **Common variants at the 19p13.1 and ZNF365 loci are associated with ER subtypes of breast cancer and ovarian cancer risk in BRCA1 and BRCA2 mutation carriers.** *Cancer Epidemiol Biomarkers Prev* 2012, **21**:645–657.
78. Gaudet MM, Kirchoff T, Green T, Vijai J, Korn JM, Guiducci C, Segre AV, McGee K, McGuffog L, Kartsonaki C, Morrison J, Healey S, Sinilnikova OM, Stoppa-Lyonnet D, Mazoyer S, Gauthier-Villars M, Sobol H, Longy M, Frenay M, GEMO Study Collaborators, Hogervorst FB, Rookus MA, Collee JM, Hoogerbrugge N, van Roozendaal KE, Piedmonte M, Rubinstein W, Nerenstone S, Van Le L, Blank SV, et al: **Common genetic variants and modification of penetrance of BRCA2-associated breast cancer.** *PLoS Genet* 2010, **6**:e1001183.
79. Cai Q, Wen W, Qu S, Li G, Egan KM, Chen K, Deming SL, Shen H, Shen CY, Gammon MD, Blot WJ, Matsuo K, Haiman CA, Khoo US, Iwasaki M, Santella RM, Zhang L, Fair AM, Hu Z, Wu PE, Signorello LB, Titus-Ernstoff L, Tajima K, Henderson BE, Chan KY, Kasuga Y, Newcomb PA, Zheng H, Cui Y, Wang F, et al: **Replication and functional genomic analyses of the breast cancer susceptibility locus at 6q25.1 generalize its importance in women of chinese, Japanese, and European ancestry.** *Cancer Res* 2011, **71**:1344–1355.
80. Antoniou AC, Kuchenbaecker KB, Soucy P, Beesley J, Chen X, McGuffog L, Lee A, Barrowdale D, Healey S, Sinilnikova OM, Caligo MA, Loman N, Harbst K, Lindblom A, Arver B, Rosenquist R, Karlsson P, Nathanson K, Domchek S, Rebbeck T, Jakubowska A, Lubinski J, Jaworska K, Durda K, Zlowowcka-Perlowska E, Osorio A, Duran M, Andres R, Benitez J, Hamann U, et al: **Common variants at 12p11, 12q24, 9p21, 9q31.2 and in ZNF365 are associated with breast cancer risk for BRCA1 and/or BRCA2 mutation carriers.** *Breast Cancer Res* 2012, **14**:R33.
81. Long J, Cai Q, Sung H, Shi J, Zhang B, Choi JY, Wen W, Delahanty RJ, Lu W, Gao YT, Shen H, Park SK, Chen K, Shen CY, Ren Z, Haiman CA, Matsuo K, Kim MK, Khoo US, Iwasaki M, Zheng Y, Xiang YB, Gu K, Rothman N, Wang W, Hu Z, Liu Y, Yoo KY, Noh DY, Han BG, et al: **Genome-wide association study in east Asians identifies novel susceptibility loci for breast cancer.** *PLoS Genet* 2012, **8**:e1002532.
82. Higginbotham KS, Breyer JP, McReynolds KM, Bradley KM, Schuyler PA, Plummer WD, Freudenthal ME, Trentham-Dietz A, Newcomb PA, Parl FF, Sanders ME, Page DL, Egan KM, Dupont WD, Smith JR: **A multistage genetic association study identifies breast cancer risk loci at 10q25 and 16q24.** *Cancer Epidemiol Biomarkers Prev* 2012, **21**:1565–1573.
83. Chen MB, Li C, Shen WX, Guo YJ, Shen W, Lu PH: **Association of a LSP1 gene rs3817198T > C polymorphism with breast cancer risk: evidence from 33,920 cases and 35,671 controls.** *Mol Biol Rep* 2011, **38**:4687–4695.
84. Lambrechts D, Truong T, Justenhoven C, Humphreys MK, Wang J, Hopper JL, Dite GS, Apicella C, Southey MC, Schmidt MK, Broeks A, Cornelissen S, van Hien R, Sawyer E, Tomlinson I, Kerin M, Miller N, Milne RL, Zamora MP, Perez JJ, Benitez J, Hamann U, Ko YD, Bruning T, Chang-Claude J, Eilber U, Hein R, Nickels S, Flesch-Janys D, Wang-Gohrke S, et al: **11q13 is a susceptibility locus for hormone receptor positive breast cancer.** *Hum Mutat* 2012, **33**:1123–1132.
85. French JD, Ghoussaini M, Edwards SL, Meyer KB, Michailidou K, Ahmed S, Khan S, Maranian MJ, O'Reilly M, Hillman KM, Betts JA, Carroll T, Bailey PJ, Dicks E, Beesley J, Tyrer J, Maia AT, Beck A, Knoblach NW, Chen C, Kraft P, Barnes D, Gonzalez-Neira A, Alonso MR, Herrero D, Tessier DC, Vincent D, Bacot F, Luccarini C, Baynes C, et al: **Functional variants at the 11q13 Risk locus for breast cancer regulate cyclin D1 expression through long-range enhancers.** *Am J Hum Genet* 2013, **92**:489–503.
86. Murabito JM, Rosenberg CL, Finger D, Kreger BE, Levy D, Splansky GL, Antman K, Hwang SJ: **A genome-wide association study of breast and prostate cancer in the NHLBI's Framingham Heart Study.** *BMC Med Genet* 2007, **8**(Suppl 1):S6.
87. Tang L, Xu J, Wei F, Wang L, Nie WW, Chen LB, Guan XX: **Association of STXBP4/COX11 rs6504950 (G > A) polymorphism with breast cancer risk: evidence from 17,960 cases and 22,713 controls.** *Arch Med Res* 2012, **43**:383–388.
88. Sawyer S, Mitchell G, McKinley J, Chenevix-Trench G, Beesley J, Chen XQ, Bowtell D, Trainer AH, Harris M, Lindeman GJ, James PA: **A role for common genomic variants in the assessment of familial breast cancer.** *J Clin Oncol* 2012, **30**:4330–4336.
89. Husing A, Canzian F, Beckmann L, Garcia-Closas M, Diver WR, Thun MJ, Berg CD, Hoover RN, Ziegler RG, Figueroa JD, Isaacs C, Olsen A, Viallon V, Boeing H, Masala G, Trichopoulos D, Peeters PH, Lund E, Ardanaz E, Khaw KT, Lenner P, Kolonel LN, Stram DO, Le Marchand L, McCarty CA, Buring JE, Lee IM, Zhang S, Lindstrom S, Hankinson SE, et al: **Prediction of breast cancer risk by genetic risk factors, overall and by hormone receptor status.** *J Med Genet* 2012, **49**:601–608.
90. Wacholder S, Hartge P, Prentice R, Garcia-Closas M, Feigelson HS, Diver WR, Thun MJ, Cox DG, Hankinson SE, Kraft P, Rosner B, Berg CD, Brinton LA, Lissowska J, Sherman ME, Chlebowski R, Kooperberg C, Jackson RD, Buckman DW, Hui P, Pfeiffer R, Jacobs KB, Thomas GD, Hoover RN, Gail MH, Chanock SJ, Hunter DJ: **Performance of common genetic variants in breast-cancer risk models.** *N Engl J Med* 2010, **362**:986–993.
91. Gail MH: **Value of adding single-nucleotide polymorphism genotypes to a breast cancer risk model.** *J Natl Cancer Inst* 2009, **101**:959–963.
92. Stevens KN, Fredericksen Z, Vachon CM, Wang X, Margolin S, Lindblom A, Nevanlinna H, Greco D, Aittomaki K, Blomqvist C, Chang-Claude J, Vrieling A, Flesch-Janys D, Sinn HP, Wang-Gohrke S, Nickels S, Brauch H, Ko YD, Fischer HP, Schmutzler RK, Meindl A, Bartram CR, Schott S, Engel C, Godwin AK, Weaver J, Pathak HB, Sharma P, Brenner H, Muller H, et al: **19p13.1 is a triple-negative-specific breast cancer susceptibility locus.** *Cancer Res* 2012, **72**:1795–1803.
93. Anders CK, Johnson R, Litton J, Phillips M, Bleyer A: **Breast cancer before age 40 years.** *Semin Oncol* 2009, **36**:237–249.
94. Kaplan HG, Malmgren JA: **Impact of triple negative phenotype on breast cancer prognosis.** *Breast J* 2008, **14**:456–463.
95. Garcia-Closas M, Hall P, Nevanlinna H, Pooley K, Morrison J, Richesson DA, Bojesen SE, Nordestgaard BG, Axelsson CK, Arias JI, Milne RL, Ribas G, Gonzalez-Neira A, Benitez J, Zamora P, Brauch H, Justenhoven C, Hamann U, Ko YD, Bruening T, Haas S, Dork T, Schurmann P, Hillemanns P, Bogdanova N, Bremer M, Karstens JH, Fagerholm R, Aaltonen K, Aittomaki K, et al: **Heterogeneity of breast cancer associations with five susceptibility loci by clinical and pathological characteristics.** *PLoS Genet* 2008, **4**:e1000054.
96. Milne RL, Benitez J, Nevanlinna H, Heikkinen T, Aittomaki K, Blomqvist C, Arias JI, Zamora MP, Burwinkel B, Bartram CR, Meindl A, Schmutzler RK, Cox A, Brock I, Elliott G, Reed MW, Southey MC, Smith L, Spurdle AB, Hopper JL, Couch FJ, Olson JE, Wang X, Fredericksen Z, Schurmann P, Bremer M, Hillemanns P, Dork T, Devilee P, van Asperen CJ, et al: **Risk of estrogen receptor-positive and -negative breast cancer and single-nucleotide polymorphism 2q35-rs13387042.** *J Natl Cancer Inst* 2009, **101**:1012–1018.
97. Cai Q, Long J, Lu W, Qu S, Wen W, Kang D, Lee JY, Chen K, Shen H, Shen CY, Sung H, Matsuo K, Haiman CA, Khoo US, Ren Z, Iwasaki M, Gu K, Xiang YB, Choi JY, Park SK, Zhang L, Hu Z, Wu PE, Noh DY, Tajima K, Henderson BE, Chan KY, Su F, Kasuga Y, Wang 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* 2011, **20**:4991–4889.
98. Chen F, Chen GK, Stram DO, Millikan RC, Ambrosone CB, John EM, Bernstein L, Zheng W, Palmer JR, Hu JJ, Rebbeck TR, Ziegler RG, Nyante S, Bandera EV,

- Ingles SA, Press MF, Ruiz-Narvaez EA, Deming SL, Rodriguez-Gil JL, Demichele A, Chanock SJ, Blot W, Signorello L, Cai Q, Li G, Long J, Huo D, Zheng Y, Cox NJ, Olopade OI, *et al*: **A genome-wide association study of breast cancer in women of African ancestry.** *Hum Genet* 2013, **132**:39–48.
99. Huo D, Zheng Y, Ogundiran TO, Adebamowo C, Nathanson KL, Domchek SM, Rebbeck TR, Simon MS, John EM, Hennis A, Nemesure B, Wu SY, Leske MC, Ambis S, Niu Q, Zhang J, Cox NJ, Olopade OI: **Evaluation of 19 susceptibility loci of breast cancer in women of African ancestry.** *Carcinogenesis* 2012, **33**:835–840.
100. Rafiq S, Tapper W, Collins A, Khan S, Politopoulos I, Gerty S, Blomqvist C, Couch FJ, Nevanlinna H, Liu J, Eccles D: **Identification of inherited genetic variations influencing prognosis in early onset breast cancer.** *Cancer Res* 2013, **73**:1883–1891.
101. Chen S, Parmigiani G: **Meta-analysis of BRCA1 and BRCA2 penetrance.** *J Clin Oncol* 2007, **25**:1329–1333.
102. Chenevix-Trench G, Milne RL, Antoniou AC, Couch FJ, Easton DF, Goldgar DE: **An international initiative to identify genetic modifiers of cancer risk in BRCA1 and BRCA2 mutation carriers: the Consortium of Investigators of Modifiers of BRCA1 and BRCA2 (CIMBA).** *Breast Cancer Res* 2007, **9**:104.
103. Antoniou AC, Sinilnikova OM, Simard J, Leone M, Dumont M, Neuhausen SL, Struewing JP, Stoppa-Lyonnet D, Barjhoux L, Hughes DJ, Coupier I, Belotti M, Lasset C, Bonadona V, Bignon YJ, Rebbeck TR, Wagner T, Lynch HT, Domchek SM, Nathanson KL, Garber JE, Weitzel J, Narod SA, Tomlinson G, Olopade OI, Godwin A, Isaacs C, Jakubowska A, Lubinski J, Gronwald J, *et al*: **RAD51 135G->C modifies breast cancer risk among BRCA2 mutation carriers: results from a combined analysis of 19 studies.** *Am J Hum Genet* 2007, **81**:1186–1200.
104. Li C, Yin Z, Wu W, Li X, Zhou B: **Genetic variants in TERT-CLPTM1L genetic region associated with several types of cancer: a meta-analysis.** *Gene* 2013, **526**:390–399.
105. Costantino JP, Gail MH, Pee D, Anderson S, Redmond CK, Benichou J, Wieand HS: **Validation studies for models projecting the risk of invasive and total breast cancer incidence.** *J Natl Cancer Inst* 1999, **91**:1541–1548.
106. Gail MH, Brinton LA, Byar DP, Corle DK, Green SB, Schairer C, Mulvihill JJ: **Projecting individualized probabilities of developing breast cancer for white females who are being examined annually.** *J Natl Cancer Inst* 1989, **81**:1879–1886.
107. Gail MH, Costantino JP, Pee D, Bondy M, Newman L, Selvan M, Anderson GL, Malone KE, Marchbanks PA, McCaskill-Stevens W, Norman SA, Simon MS, Spirtas R, Ursin G, Bernstein L: **Projecting individualized absolute invasive breast cancer risk in African American women.** *J Natl Cancer Inst* 2007, **99**:1782–1792.
108. Matsuno RK, Costantino JP, Ziegler RG, Anderson GL, Li H, Pee D, Gail MH: **Projecting individualized absolute invasive breast cancer risk in Asian and Pacific Islander American women.** *J Natl Cancer Inst* 2011, **103**:951–961.
109. Gail MH: **Discriminatory accuracy from single-nucleotide polymorphisms in models to predict breast cancer risk.** *J Natl Cancer Inst* 2008, **100**:1037–1041.

10.1186/bcr3591

Cite this article as: Maxwell and Nathanson: Common breast cancer risk variants in the post-COGS era: a comprehensive review. *Breast Cancer Research* 2013, **15**:212