Sawyer, E; Roylance, R; Petridis, C; Brook, MN; Nowinski, S; Papoulis, E; Fletcher, O; Pinder, S; Hanby, A; Kohut, K; Gorman, P; Caneppele, M; Peto, J; Dos Santos Silva, I; Johnson, N; Swann, R; Dwek, M; Perkins, KA; Gillett, C; Houlston, R; Ross, G; De Ieso, P; Southey, MC; Hopper, J; Provenzano, E; Apicella, C; Wesseling, J; Cornelissen, S; Keeman, R; Fasching, PA; Jud, SM; Ekici, AB; Beckmann, MW; Kerin, MJ; Marme, F; Schneeweiss, A; Sohn, C; Burwinkel, B; Gunel, P; Truong, T; Laurent-Puig, P; Kerbrat, P; Bojesen, SE; Nordestgaard, BG; Nielsen, SF; Flyger, H; Milne, RL; Perez, JI; Meundez, P; Benitez, J; Brenner, H; Dieffenbach, AK; Arndt, V; Stegmaier, C; Meindl, A; Lichtner, P; Schmutzler, RK; Lochmann, M; Brauch, H; Fischer, HP; Ko, YD; GENICA Network; Nevanlinna, H; Muranen, TA; Aittomaki, K; Blomqvist, C; Bogdanova, NV; Drk, T; Lindblom, A; Margolin, S; Mannermaa, A; Kataja, V; Kosma, VM; Hartikainen, JM; Chenevix-Trench, G; Investigators, K; Lambrechts, D; Weltens, C; Van Limbergen, E; Hatse, S; Chang-Claude, J; Rudolph, A; Seibold, P; Flesch-Janys, D; Radice, P; Peterlongo, P; Bonanni, B; Volorio, S; Giles, GG; Severi, G; Baglietto, L; McLean, CA; Haiman, CA; Henderson, BE; Schumacher, F; Le Marchand, L; Simard, J; Goldberg, MS; Labrche, F; Dumont, M; Kristensen, V; Winqvist, R; Pylls, K; Jukkola-Vuorinen, A; Kauppila, S; Andrlis, IL; Knight, JA; Glendon, G; Mulligan, AM; Devilee, P; Tollenaar, RA; Seynaeve, CM; Kriege, M; Figueroa, J; Chanock, SJ; Sherman, ME; Hooning, MJ; Hollestelle, A; van den Ouweland, AM; van Deurzen, CH; Li, J; Czene, K; Humphreys, K; Cox, A; Cross, SS; Reed, MW; Shah, M; Jakubowska, A; Lubinski, J; Jaworska-Bieniek, K; Durda, K; Swerdlow, A; Ashworth, A; Orr, N; Schoemaker, M; Couch, FJ; Hallberg, E; Gonzalez-Neira, A; Pita, G; Alonso, MR; Tessier, DC; Vincent, D; Bacot, F; Bolla, MK; Wang, Q; Dennis, J; Michailidou, K; Dunning, AM; Hall, P; Easton, D; Pharoah, P; Schmidt, MK; Tomlinson, I; Garcia-Closas, M (2014) Genetic predisposition to in situ and invasive lobular carcinoma of the breast. PLoS genetics, 10 (4). e1004285. ISSN 1553-7390 DOI: 10.1371/journal.pgen.1004285

Downloaded from: http://researchonline.lshtm.ac.uk/1673603/

DOI: 10.1371/journal.pgen.1004285
Usage Guidelines

Please refer to usage guidelines at http://researchonline.lshtm.ac.uk/policies.html or alternatively contact researchonline@lshtm.ac.uk.

Available under license: http://creativecommons.org/licenses/by/2.5/
Genetic Predisposition to In Situ and Invasive Lobular Carcinoma of the Breast


1 Research Oncology, Division of Cancer Studies, King’s College London, Guy’s Hospital, London, United Kingdom, 2 Centre for Molecular Oncology, Barts Cancer Institute, Queen Mary University of London, London, United Kingdom, 3 Division of Genetics and Epidemiology, Institute of Cancer Research, London, United Kingdom, 4 Biomedical Research Centre, King’s College London, Guy’s Hospital, London, United Kingdom, 5 Breakthrough Breast Cancer Research Centre, The Institute of Cancer Research, London, United Kingdom, 6 Leeds Institute of Molecular Medicine, St James’s University Hospital, Leeds, United Kingdom, 7 London School of Hygiene and Tropical Medicine, London, United Kingdom, 8 Department of Molecular and Applied Biosciences, University of Westminster, London, United Kingdom, 9 The Royal Marsden NHS Foundation Trust, London, United Kingdom, 10 Department of Pathology, The University of Melbourne, Melbourne, Australia, 11 Centre for Molecular, Environmental, Genetic and Analytic Epidemiology, The University of Melbourne, Melbourne, Australia, 12 NHRI Cambridge Biomedical Research Centre, Addenbrookes Hospital, Cambridge, United Kingdom, 13 Netherlands Cancer Institute, Antoni van Leeuwenhoek Hospital, Amsterdam, The Netherlands, 14 David Geffen School of Medicine, Department of Medicine Division of Hematology and Oncology, University of California at Los Angeles, California, United States of America, 15 Department of Gynecology and Obstetrics, University Hospital Erlangen, Friedrich-Alexander-Universität Erlangen-Nürnberg, Comprehensive Cancer Center Erlangen-EMN, Erlangen, Germany, 16 Institute of Human Genetics, Friedrich-Alexander-Universität Erlangen-Nürnberg, Erlangen, Germany, 17 Surgery, Clinical Science Institute, National University of Ireland, Galway, Ireland, 18 Department of Obstetrics and Gynecology, University of Heidelberg, Heidelberg, Germany, 19 National Center for Tumor Diseases, University of Heidelberg, Heidelberg, Germany, 20 Molecular Epidemiology Group, German Cancer Research Center (DKFZ), Heidelberg, Germany, 21 InsERM (National Institute of Health and Medical Research), CESP (Center for Research in Epidemiology and Population Health), U1018, Environmental Epidemiology of Cancer, Villejuif, France, 22 University Paris-Sud, UMRIS 1018, Villejuif, France, 23 Université Paris Sorbonne Cité, UMR-5775 Inserm, Paris, France, 24 Centre Eugène Marquis, Department of Medical Oncology, Rennes, France, 25 Copenhagen General Population Study and Department of Clinical Biochemistry, Herlev Hospital, Copenhagen University Hospital, University of Copenhagen, Copenhagen, Denmark, 26 Department of Breast Surgery, Herlev Hospital, Copenhagen University Hospital, Copenhagen, Denmark, 27 Genetic & Molecular Epidemiology Group, Human Cancer Genetics Program, Spanish National Cancer Research Centre (CNIO), Madrid, Spain.
Invasive lobular breast cancer (ILC) accounts for 10–15% of all invasive breast carcinomas. It is generally ER positive (ER+) and often associated with lobular carcinoma in situ (LCIS). Genome-wide association studies have identified more than 70 common polymorphisms that predispose to breast cancer, but these studies included predominantly ductal (IDC) carcinomas. To identify novel common polymorphisms that predispose to ILC and LCIS, we pooled data from 6,023 cases (5,622 ILC, 401 pure LCIS, 140 pure IDC, and 83 ILC+IDC tumors) and compared them with 7,163 controls. Using the ICOS chip, six novel SNPs most strongly associated with ILC were identified. These six SNPs were also associated with ILC in a subsequent independent cohort (rs2981579/10q26) and the pooled analysis (rs2981579/10q26, rs10995190/10q21, rs1550623/2q31, rs11752916/1q4, rs15288809/1q4, and rs15288810/1q4) were genotyped in 2,767 ILC, 1,718 IDC, and 2,571 controls. These analyses identified a lobular-specific SNP at rs2981579 (rs2981579, P = 6.6 × 10^-13). Of the 75 known breast cancer polymorphisms that were genotyped, 56 were associated with ILC and 15 with IDC at P < 0.05. Two SNPs showed significantly stronger associations for ILC than LCIS (rs2981579/10q26, P = 6.6 × 10^-13); and four showed stronger associations for LCIS than ILC (rs6678914/1q32, P = 6.6 × 10^-13). In addition, seven of the 75 known loci showed significant differences between ER+ tumors with IDC and ILC histology. The Netherlands, 83 Division of Cancer Prevention, National Cancer Institute, Bethesda, Maryland, United States of America, 84 Department of Medical Oncology, Family Cancer Clinic, Erasmus MC Cancer Institute, Rotterdam, The Netherlands, 85 Department of Clinical Genetics, Erasmus University Medical Center, Rotterdam, The Netherlands, 87 Human Genetics Division, Genome Institute of Singapore, Singapore, 88 Medical Epidemiology and Biostatistics, Karolinska Institutet, Stockholm, Sweden, 89 CRUK/YCR Sheffield Cancer Research Centre, Department of Oncology, University of Sheffield, Sheffield, United Kingdom, 90 Academic Unit of Pathology, Department of Neuroscience, University of Sheffield, Sheffield, United Kingdom, 91 Centre for Cancer Genetic Epidemiology, Department of Oncology, University of Cambridge, Cambridge, United Kingdom, 92 Department of Genetics and Pathology, Pomeranian Medical University, Szczecin, Poland, 93 Postgraduate School of Molecular Medicine, Warsaw Medical University, Warsaw, Poland, 94 Division of Genetics and Epidemiology and Division of Breast Cancer Research, The Institute of Cancer Research, Sutton, Surrey, United Kingdom, 95 Breakthrough Breast Cancer Research Centre, Division of Breast Cancer Research, The Institute of Cancer Research, London, United Kingdom, 96 Department of Laboratory Medicine and Pathology, Mayo Clinic, Rochester, Minnesota, United States of America, 97 Department of Health Sciences Research, Mayo Clinic, Rochester, Minnesota, United States of America, 98 Human Genotyping-CEGEN Unit, Human Cancer Genetics Program, Spanish National Cancer Research Centre (CNIO), Madrid, Spain, 99 McGill University and Genome Québec Innovation Centre, Montréal, Québec, Canada, 100 Centre for Cancer Genetic Epidemiology, Department of Public Health and Primary Care, University of Cambridge, Cambridge, United Kingdom, 101 Welcome Trust Centre for Human Genetics and Oxford Biomedical Research Centre, University of Oxford, Oxford, United Kingdom

Abstract

Invasive lobular breast cancer (ILC) accounts for 10–15% of all invasive breast carcinomas. It is generally ER positive (ER+) and often associated with lobular carcinoma in situ (LCIS). Genome-wide association studies have identified more than 70 common polymorphisms that predispose to breast cancer, but these studies included predominantly ductal (IDC) carcinomas. To identify novel common polymorphisms that predispose to ILC and LCIS, we pooled data from 6,023 cases (5,622 ILC, 401 pure ILC) and 34,271 controls from 36 studies genotyped using the ICOS chip. Six novel SNPs most strongly associated with ILC/LCIS in the pooled analysis were genotyped in a further 516 lobular cases (482 ILC, 36 LCIS) and 1,467 controls. These analyses identified a lobular-specific SNP at rs2981579 (rs2981579, OR = 1.13 (1.09–1.18), P = 6.6 × 10^-13). Of the 75 known breast cancer polymorphisms that were genotyped, 56 were associated with ILC and 15 with IDC at P < 0.05. Two SNPs showed significantly stronger associations for ILC than LCIS (rs2981579/10q26, P = 6.6 × 10^-13), and two showed stronger associations for LCIS than ILC (rs6678914/1q32, P = 6.6 × 10^-13). In addition, seven of the 75 known loci showed significant differences between ER+ tumors with IDC and ILC histology. Three of these showed stronger associations for ILC (rs11249433/1p11, rs2981579/10q26, and rs10995190/10q21) and four associated only with IDC (5p12/rs10941679, rs2588809/1q4, rs15288809/1q4, and rs15288810/1q4). In conclusion, we have identified one novel lobular breast cancer specific predisposition polymorphism at rs2981579, and shown for the first time that common breast cancer polymorphisms predispose to LCIS. We have shown that many of the ER+ breast cancer predisposition loci also predispose to ILC, although there is some heterogeneity between ER+ lobular and ER+ IDC tumors. These data provide evidence for overlapping, but distinct etiological pathways within ER+ breast cancer between morphological subtypes.
The authors have declared that no competing interests exist.
Author Summary

Invasive lobular breast cancer (ILC) accounts for 10–15% of invasive breast cancer and is generally ER positive (ER+). To date, none of the genome-wide association studies that have identified loci that predispose to breast cancer in general or to ER+ or ER-negative breast cancer have focused on lobular breast cancer. In this lobular breast cancer study we identified a new variant that appears to be specific to this morphological subtype. We also ascertained which of the known variants predispose specifically to lobular breast cancer and show for the first time that some of these loci are also associated with lobular carcinoma in situ, a non-obligate precursor of breast cancer and also a risk factor for contralateral breast cancer. Our study shows that the genetic pathways of invasive lobular cancer and ER+ ductal carcinoma mostly overlap, but there are important differences that are likely to provide insights into the biology of lobular breast tumors.

Introduction

Invasive lobular breast cancer (ILC) accounts for 10–15% of all invasive breast carcinomas and it has distinct etiological, clinical and biological characteristics compared with the more common invasive ductal/no special type carcinoma (IDC) [1]. Lobular cancers show stronger associations with the use of hormone replacement therapy (HRT) than IDC, [2] and its incidence follows a similar temporal pattern as the use of combined HRT [3]. ILC is characterized by E-cadherin loss and the malignant cells therefore infiltrate the breast stroma in single files with a typical associated stromal reaction. This makes it difficult to detect these tumors by palpation or mammography, and they are often larger at presentation than IDCs [4]. ILCs are generally of histological grade 2 and estrogen receptor positive (ER+), with the exception of the pleomorphic subgroup. They typically have a different pattern of metastatic spread to IDCs, tending to infiltrate the peritoneum, ovary and gastrointestinal system. There is some evidence that they are less chemo-sensitive than IDC and that the 10-year survival rate of women with ILC is lower than that of ER+ IDCs [5,6].

ILC is often associated with lobular carcinoma in situ (LCIS), a form of non-invasive breast cancer that is difficult to detect clinically and typically found incidentally on biopsy. The increased breast biopsy rate associated with screening mammography has led to an increase in the diagnosis of LCIS. LCIS shares many of the same genetic aberrations as ILC, suggesting that it is a precursor lesion in an analogous manner to ductal carcinoma in situ (DCIS) and IDC [7]. Women who have had LCIS are 2.4 times more likely to develop invasive breast cancer compared to the general population, with an excess of ILC (23–80% of cases) [8,9]. However only 50–70% of invasive cancers associated with LCIS have lobular morphology [10, unpublished data from GLACIER study]. The remaining cancers have an IDC or mixed ductal-lobular appearance, but again are generally ER+ (95% of IDC and mixed ductal-lobular cancers associated with LCIS in the GLACIER study were ER+). Unlike DCIS, LCIS is also a risk factor for developing invasive cancer in the contralateral breast [8].

Genome-wide association studies (GWAS) in breast cancer have identified loci that predispose to invasive breast cancer in general, or specifically to ER+ or ER-negative disease [11–25]. However, no previous study has focused specifically on lobular carcinomas. Only one common single nucleotide polymorphism (SNP; rs11249433 at 1p11.2) has been shown to be more strongly associated with lobular than ductal histology [26]. For the remaining SNPs predisposing to ER+ tumors, it is unclear whether the studies have lacked statistical power to identify differential associations by histology, or whether associations tend to be non-differential by morphology after accounting for ER status.

The aim of this study was to identify new breast cancer susceptibility loci specific to lobular carcinoma, and to evaluate the heterogeneity of associations of known loci by morphology. This involved pooling genotyping data from over 6,000 cases of lobular carcinoma (ILC and/or LCIS) and over 34,000 controls genotyped using the iCOGS chip, a custom SNP array that comprises 211,155 SNPs enriched at predisposition loci for breast and other cancers [24].

Results

In a phase I analysis, we evaluated risk associations between SNPs on the iCOGS chip and risk of ILC and LCIS using 1,782 lobular cases (1,470 ILC with or without LCIS, 312 pure LCIS) from GLACIER, a UK study of lobular breast cancer, and 4,755 UK controls from the Breast Cancer Association Consortium, BCAC (Figure 1). There was little evidence for systematic inflation of the test statistics, based on 37,544 uncorrelated SNPs that had not been selected on the basis of breast cancer risk (k = 1.04; Figure S1). Data were combined by meta-analysis with a further 4,241 cases (4,152 ILC, 89 LCIS) and 29,519 controls of European ancestry, derived from 34 studies in BCAC, and previously typed on the iCOGS chip (Tables S1 and S2). This resulted in a total of 6,023 cases (5,622 ILC, 401 LCIS) and 34,271 controls with data on 199,961 iCOGS SNPs (after quality control exclusions and with minor allele frequency (MAF) >0.01) included in the meta-analysis.

Search for new lobular breast cancer predisposition loci

All SNPs reaching genome-wide significance (P<5×10−8) in the meta-analysis were correlated with one of the known breast cancer predisposition loci. In order to identify new loci that predispose to lobular carcinoma, we selected six uncorrelated SNPs (rs11977670, rs2121783, rs2747652, rs3909680, rs9948182, rs7034265) that were only weakly correlated (r2 = 0.25) with known loci and that showed the best evidence of association (P between 5×10−6 and 5×10−8) in the overall lobular case-control analysis (ILC and LCIS). These SNPs were genotyped in a Phase II including 516 cases (481 ILC, 35 LCIS) and 1,467 controls, all from white European donors (Figure 1).

One of the six SNPs, rs11977670 at 7q34, reached genome-wide significance in a pooled analysis of phase I and II ILC cases and controls (OR = 1.13, 95%CI = 1.09–1.18, P = 6.0×10−10; Table 1, Figure 2); rs11977670 showed a similar association with LCIS (P-het for ILC vs LCIS = 0.198), and a very weak or no association with IDC (OR = 1.02, 95%CI = 1.00–1.05, P = 0.070; P-het for ILC vs IDC = 1.3×10−5), indicating that this is a lobular specific predisposition locus (Table 2). The risk allele appeared to act in a dominant rather than additive manner: ORAG = 1.21, 95%CI = 1.14–1.30; ORAG = 1.27, 95%CI = 1.17–1.38; P for departure from log-additivity = 0.009; Table S3. rs11977670 was not significantly associated with age at onset of ILC (P = 0.16) and risk alleles were not significantly over-represented in cases with a positive family history (FH) (P = 0.90, FH+ vs FH−). None of the other 5 SNPs genotyped...
were associated with lobular breast cancer at a genome-wide significance level, with the strongest association being for rs2121783 at 3p13 (OR = 1.11, 95%CI = 1.07–1.15, \(P = 4.5 \times 10^{-7}\); Table S4).

rs11977670 at 7q34 (position:139942304, GRCh Build 37) is intergenic, 65 kb from the nearest gene, \(JHDM1D\), a histone demethylase and 500 kb from \(BRAF\), a gene frequently mutated in melanoma. It is also in close proximity to a predicted novel U1 spliceosomal RNA that contains two U1 specific promoter motifs (Figure S2). ENCODE data on normal human mammary epithelial cells (HMEC), and breast carcinoma (MCF-7), were used to establish chromatin states in the region and showed that rs11977670 lies in region marked by H3K27 acetylation, Figure S3.

Using expression data from the Cancer Genome Atlas Network (TCGA database) [27], we assessed expression of the nine genes within 0.5 Mb of rs11977670 by breast cancer subtype (ER+ ILC, 40 cases; ER+ IDC, 341 cases; and ER-negative IDC, 108 cases; Figure S4). Three genes showed differential expression in ER+ ILC compared to ER+ IDC (\(BRAF\), \(P = 0.006\); \(NDUFB2\), \(P = 0.02\); \(SLC37A3\), \(P = 0.05\)), however none reached statistical significance when correcting for multiple testing. Another two genes, \(JHDM1D\) and \(ADCK2\), showed a difference in expression between ER-negative and ER+ cancers, but this was not lobular-specific. To further investigate which genes may be influenced by SNPs tagged by rs11977670, germline genotype data for rs13225058 (A/G), a surrogate for rs11977670 (G/A) (\(r^2 = 0.79\)) was taken from the TCGA database (SNP6.0 Affymetrix array) and compared to expression of these genes, correcting for copy number variation, in 335 ER+ primary breast cancers where both genotype and expression data was available. A significant difference, after correcting for multiple testing, was found in expression between the AA and GG genotype for two genes \(JHDM1D\) (\(P = 0.0005\)) and \(SLC37A3\) (\(P = 0.004\)), Figure S5a. Confining the analysis to the 36 ILC cases with data in TCGA showed no significant genotype specific expression due small numbers although there was the suggestion of a trend towards overexpression with the GG genotype (2 cases), Figure S5b. 48 of the cases also had expression data on adjacent normal breast tissue, but due to the small numbers no significant genotype specific expression changes were detected, Figure S6. There was no evidence of copy number variation around rs11977670 and no evidence of an excess of somatic mutations in \(JHDM1D\), \(SLC37A3\) or \(BRAF\) in ILC.
Assessment of the 75 known breast cancer susceptibility loci for association with ILC and LCIS

Most (56 of 75) known common breast cancer susceptibility loci were associated with ILC at P<0.05 with the effect in the same direction as previously reported (Table S5), and 13 of these reached genome-wide significance (P<5×10⁻⁸). The strongest associations were with SNPs close to FGFR2 (rs2981579, OR = 1.38, P = 5.1×10⁻³²), TOX3 (rs3803662, OR = 1.33, P = 1.1×10⁻³⁵) at 1p11.2, rs11249433, OR = 1.25, P = 2.7×10⁻²⁵ and 11q13.3 (rs554219, OR = 1.33, P = 1.6×10⁻²⁵). All 13 loci had previously been shown to be associated with ER+ breast cancer and one locus, rs1249433 (lp11.2), with lobular histology in subgroup analysis. Of the remaining 19 SNPs with P<0.05, 18 had ORs in the same direction as previously reported for overall breast cancer (Sign test P<0.0001), suggesting that these SNPs are also likely to predispose to LCIS. Only one of the seven ER-negative specific loci on the iCOGS array showed a significant association with ILC (rs12710693, OR = 1.17, 95% CI = 1.00–1.36, P = 0.05; Table S5). Forty-seven of the remaining 60 SNPs at P>0.05 had ORs in the same direction as for ILC. This is greater than one would expect by chance (Sign Test P = 1.2×10⁻⁵) suggesting many of these SNPs predispose to LCIS, but the study did not have enough power to detect these associations with the small sample size.

A global test in case-only analysis (ILC vs LCIS) indicated no significant differences in associations between the 75 SNPs associated with LCIS and ILC (likelihood ratio test (75 df) P = 5.9×10⁻⁶). The strongest associations were for rs2981579 at chr 1p11.2 (P=het = 2.7×10⁻²⁶; Table 4), a SNP previously associated with lobular histology. At P<0.05, a further two loci were associated more strongly with ILC than IDC: rs2981579, FGFR2 (P=het = 5.3×10⁻¹⁵) and rs10995190, 10q21.2 (P=het = 0.002). This analysis also identified four IDC-specific SNPs at P<0.05: rs10941679, 5p12 (P=het = 1.5×10⁻¹⁵); rs2588809, RAD51L1 (P=het = 0.001); rs6472903, 8q21 (P=het = 0.004); rs1550623, CDC47 (P=het = 0.031) Table S6.

Assessment of the 75 known susceptibility SNPs for differential effects on ILC and IDC

In order to identify lobular specific SNPs, we performed a case-only analysis of 3,201 ER+ ILC cases and 15,024 ER+ IDC cases from BCAC. Analysis was confined to ER+ cases since 94% of IDC cases were ER+ (compared to 78% of IDC in BCAC). A global test indicated significant differences in SNP associations between ILC and IDC (likelihood ratio test (75 df) P = 5.9×10⁻⁶). The SNP showing the largest difference between ILC and IDC was rs11249433 at chr 1p11.2 (P=het = 2.7×10⁻²⁶; Table 4), a SNP previously associated with lobular histology. At P<0.05, a further two loci were associated more strongly with ILC than IDC: rs2981579, FGFR2 (P=het = 5.3×10⁻¹⁵) and rs10995190, 10q21.2 (P=het = 0.002). This analysis also identified four IDC-specific SNPs at P<0.05: rs10941679, 5p12 (P=het = 1.5×10⁻¹⁵); rs2588809, RAD51L1 (P=het = 0.001); rs6472903, 8q21 (P=het = 0.004); rs1550623, CDC47 (P=het = 0.031) Table S6.

Assessment of the 75 known susceptibility SNPs for effects on mixed ILC-IDC cancer predisposition

Case-control analysis of 690 mixed ductal-lobular carcinomas revealed 25 loci that showed an association with these mixed cancers at P<0.05. The top hits were at FGFR2 (rs2981579, OR = 1.37, P = 1.6×10⁻³⁷), rs941764 (CCDC88C, OR = 1.25, P = 3.6×10⁻³⁶) and rs10995190 (ZNZF365, OR = 0.74, P = 3.9×10⁻¹⁵). The case-only analysis above showed that two of these SNPs are more strongly associated with ILC than IDC (rs2981579, FGFR2 (P=het = 5.3×10⁻¹⁵) and rs10995190, 10q21.2 (P=het = 0.002)). This analysis also identified four IDC-specific SNPs at P<0.05: rs10941679, 5p12 (P=het = 1.5×10⁻¹⁵); rs2588809, RAD51L1 (P=het = 0.001); rs6472903, 8q21 (P=het = 0.004); rs1550623, CDC47 (P=het = 0.031) Table S6.

Discussion

Our analyses of a total of 6,539 lobular cancers (including 436 cases of pure LCIS) and 35,710 controls has identified for the first time a lobular-specific SNP, rs11977670 (JHDM1D; OR = 1.13 P = 4.2×10⁻¹⁰), that showed little evidence of association with IDC (P = 0.07) or DCIS (P = 0.23). Identification of the target of this association will require fine mapping of the region, followed by functional assays to determine which gene(s) the key SNPs regulate. The preliminary genome-wide significance (P=5×10⁻⁸) Table S1. The remaining SNPs showed no significant heterogeneity between ILC and LCIS.

Table 1. rs11977670, chromosome 7:139942304 G>A, and association ILC in populations of European ancestry.

<table>
<thead>
<tr>
<th>Study/Consortia</th>
<th>N Studies</th>
<th>Cases</th>
<th>Controls</th>
<th>MAF</th>
<th>OR (95% CI)*</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Phase I</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GLACER</td>
<td>1</td>
<td>1,470</td>
<td>4,755</td>
<td>0.437</td>
<td>1.16 (1.07, 1.26)</td>
<td>6.1×10⁻⁴</td>
</tr>
<tr>
<td>BCAC</td>
<td>34</td>
<td>4,150</td>
<td>29,488</td>
<td>0.429</td>
<td>1.10 (1.05, 1.16)</td>
<td>4.0×10⁻⁵</td>
</tr>
<tr>
<td>Combined</td>
<td>35</td>
<td>5,620</td>
<td>34,243</td>
<td>0.430</td>
<td>1.12 (1.07, 1.16)</td>
<td>1.4×10⁻⁷</td>
</tr>
<tr>
<td>Phase II</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>UK PHASE II</td>
<td>1</td>
<td>479</td>
<td>1,452</td>
<td>0.426</td>
<td>1.38 (1.19, 1.60)</td>
<td>2.9×10⁻⁵</td>
</tr>
<tr>
<td>Phase III</td>
<td>36</td>
<td>6,099</td>
<td>5,695</td>
<td>0.430</td>
<td>1.13 (1.09, 1.18)</td>
<td>6.0×10⁻¹⁰</td>
</tr>
</tbody>
</table>

* per allele.
doi:10.1371/journal.pgen.1004285.t001

Discussion

Our analyses of a total of 6,539 lobular cancers (including 436 cases of pure LCIS) and 35,710 controls has identified for the first time a lobular-specific SNP, rs11977670 (JHDM1D; OR = 1.13 P = 4.2×10⁻¹⁰), that showed little evidence of association with IDC (P = 0.07) or DCIS (P = 0.23). Identification of the target of this association will require fine mapping of the region, followed by functional assays to determine which gene(s) the key SNPs regulate. The preliminary in silico functional analysis suggests that SNPs in this region may be influencing expression of JHDM1D (a histone demethylase) and SLCA37A3 (a sugar-phosphate exchanger). For JHDM1D this appears to be a recessive effect, in contrast to the susceptibility data, which suggests a dominant effect. There are little data on the role of these genes in cancer. There is some evidence that increased expression of JHDM1D can...
suppress tumor growth by regulating angiogenesis [28] and decreased expression promotes invasiveness, which is contrary to what one would expect from the risk data [29]. This inconsistency does shed some doubt on these results and further analysis of the region is required before any firm conclusion can be made. Studies of syndecan-1-deficient breast cancer cells, which show increased cell motility and invasiveness, demonstrate decreased expression of both JHDM1D and E-cadherin [29], suggesting the two genes may interact.

Somatic mutations in $CDH1$ (E-Cadherin) are frequent in ILC and rare germline frameshift mutations in $CDH1$ have been described in ILC, particularly in families with hereditary diffuse gastric cancer (HDGC), but also in cases of familial ILC with no HDGC [30,31]. However, none of the 56 SNPs in $CDH1$ that were typed on the iCOGS chip showed any association with lobular cancer at $P<0.05$.

It should also be noted that this study is not a true genome wide association study for lobular breast cancer as the SNPs on the iCOGS chips were chosen on the basis of some prior evidence of association with breast cancer as a whole. Although ILC would have been a small proportion of the samples in the discovery sets for these SNPs it is possible that other lobular specific loci exist that have not been included on the iCOGS chip. This is particularly true for LCIS, which would only have been included in the discovery set as a parallel phenotype when associated with invasive disease.
<table>
<thead>
<tr>
<th>Tumour type</th>
<th>N Studies</th>
<th>Cases</th>
<th>Controls</th>
<th>OR (95% CI)</th>
<th>P</th>
<th>Phet</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>All tumours combined</strong></td>
<td>43</td>
<td>48,286</td>
<td>43,776</td>
<td>1.05 (1.03, 1.07)</td>
<td>5.1 × 10⁻⁷</td>
<td></td>
</tr>
<tr>
<td>Ductal invasive</td>
<td>34</td>
<td>23,549</td>
<td>29,488</td>
<td>1.02 (1.00, 1.05)</td>
<td>0.070</td>
<td></td>
</tr>
<tr>
<td>ER pos</td>
<td>32</td>
<td>15,010</td>
<td>29,250</td>
<td>1.03 (1.00, 1.06)</td>
<td>0.047</td>
<td></td>
</tr>
<tr>
<td>ER neg</td>
<td>33</td>
<td>4,266</td>
<td>29,068</td>
<td>0.97 (0.93, 1.02)</td>
<td>0.208</td>
<td></td>
</tr>
<tr>
<td>DCIS</td>
<td>22</td>
<td>847</td>
<td>23,372</td>
<td>1.06 (0.96, 1.17)</td>
<td>0.231</td>
<td>0.900</td>
</tr>
<tr>
<td>Lobular invasive</td>
<td>36</td>
<td>5,099</td>
<td>35,695</td>
<td>1.13 (1.09, 1.18)</td>
<td>6.0 × 10⁻¹⁰</td>
<td>1.3 × 10⁻⁸</td>
</tr>
<tr>
<td>ER pos</td>
<td>34</td>
<td>4,135</td>
<td>35,457</td>
<td>1.13 (1.08, 1.19)</td>
<td>2.1 × 10⁻⁷</td>
<td>1.8 × 10⁻⁴</td>
</tr>
<tr>
<td>ER neg</td>
<td>32</td>
<td>268</td>
<td>33,791</td>
<td>1.22 (1.03, 1.45)</td>
<td>0.022</td>
<td>0.009</td>
</tr>
<tr>
<td>LCIS</td>
<td>16</td>
<td>437</td>
<td>25,488</td>
<td>1.10 (0.96, 1.26)</td>
<td>0.179</td>
<td>0.198</td>
</tr>
<tr>
<td>Lobular (invasive+LCIS)</td>
<td>36</td>
<td>6,536</td>
<td>35,695</td>
<td>1.13 (1.09, 1.18)</td>
<td>4.2 × 10⁻¹⁰</td>
<td></td>
</tr>
<tr>
<td>ER pos</td>
<td>34</td>
<td>4,202</td>
<td>35,457</td>
<td>1.13 (1.08, 1.18)</td>
<td>3.1 × 10⁻⁷</td>
<td></td>
</tr>
<tr>
<td>ER neg</td>
<td>32</td>
<td>278</td>
<td>33,791</td>
<td>1.20 (1.01, 1.42)</td>
<td>0.035</td>
<td></td>
</tr>
</tbody>
</table>

- aDuctal (invasive) vs Ductal (in situ) in subjects where ER status exists, adjusted for ER status.
- bDuctal (invasive) vs Lobular (invasive) in subjects where ER status exists, UK population-based studies grouped together.
- biDuctal (invasive) vs Lobular (invasive), ER positive only.
- biiDuctal (invasive) vs Lobular (invasive), ER negative only.
- cLobular (invasive) vs Lobular (in situ) in subjects where ER status exists, adjusted for ER status.
Table 3. SNPs associated with ILC ($P < 5 \times 10^{-8}$) or LCIS ($P < 0.05$) in a pooled lobular analysis (GLACIER and BCAQ).

<table>
<thead>
<tr>
<th>Cytoband</th>
<th>Gene</th>
<th>RS number</th>
<th>MAF ILC vs Controls*</th>
<th>OR (95% CI)</th>
<th>P</th>
<th>OR (95% CI)</th>
<th>P</th>
<th>Phet ILC vs LCIS</th>
</tr>
</thead>
<tbody>
<tr>
<td>10q26.13</td>
<td>FGFR2</td>
<td>rs2981579</td>
<td>0.40</td>
<td>1.38 (1.32, 1.44)</td>
<td>$5.1 \times 10^{-24}$</td>
<td>1.19 (1.03, 1.37)</td>
<td>0.019</td>
<td>0.04</td>
</tr>
<tr>
<td>16q12.1</td>
<td>TOX3</td>
<td>rs38034862</td>
<td>0.26</td>
<td>1.33 (1.27, 1.39)</td>
<td>$1.1 \times 10^{-10}$</td>
<td>1.35 (1.16, 1.57)</td>
<td>1.2 $\times 10^{-4}$</td>
<td>0.99</td>
</tr>
<tr>
<td>1p11.2</td>
<td></td>
<td>rs11249433</td>
<td>0.40</td>
<td>1.25 (1.20, 1.30)</td>
<td>$2.7 \times 10^{-23}$</td>
<td>1.15 (1.00, 1.33)</td>
<td>0.050</td>
<td>0.94</td>
</tr>
<tr>
<td>11q13.3</td>
<td></td>
<td>rs554219</td>
<td>0.12</td>
<td>1.33 (1.26, 1.41)</td>
<td>$1.6 \times 10^{-22}$</td>
<td>1.31 (1.08, 1.60)</td>
<td>0.007</td>
<td>0.80</td>
</tr>
<tr>
<td>9q31.2</td>
<td></td>
<td>rs865686</td>
<td>0.38</td>
<td>0.83 (0.79, 0.86)</td>
<td>$1.0 \times 10^{-17}$</td>
<td>0.72 (0.61, 0.84)</td>
<td>0.2 $\times 10^{-5}$</td>
<td>0.12</td>
</tr>
<tr>
<td>2q35</td>
<td></td>
<td>rs13387042</td>
<td>0.49</td>
<td>0.84 (0.80, 0.87)</td>
<td>$5.7 \times 10^{-17}$</td>
<td>0.90 (0.78, 1.04)</td>
<td>0.145</td>
<td>0.21</td>
</tr>
<tr>
<td>11q13.3</td>
<td></td>
<td>rs7591566</td>
<td>0.06</td>
<td>1.40 (1.29, 1.51)</td>
<td>$1.2 \times 10^{-16}$</td>
<td>1.55 (1.20, 2.01)</td>
<td>7.8 $\times 10^{-4}$</td>
<td>0.54</td>
</tr>
<tr>
<td>11q13.3</td>
<td></td>
<td>rs641367</td>
<td>0.14</td>
<td>1.24 (1.18, 1.31)</td>
<td>$7.2 \times 10^{-15}$</td>
<td>1.32 (1.10, 1.58)</td>
<td>0.003</td>
<td>0.46</td>
</tr>
<tr>
<td>10q21.2</td>
<td>ZNF365</td>
<td>rs10993190</td>
<td>0.16</td>
<td>0.80 (0.75, 0.85)</td>
<td>$1.7 \times 10^{-13}$</td>
<td>0.69 (0.55, 0.87)</td>
<td>0.002</td>
<td>0.10</td>
</tr>
<tr>
<td>5q11.2</td>
<td>MAP3K1</td>
<td>rs889312</td>
<td>0.28</td>
<td>1.18 (1.13, 1.23)</td>
<td>$9.1 \times 10^{-13}$</td>
<td>1.05 (0.89, 1.22)</td>
<td>0.576</td>
<td>0.03</td>
</tr>
<tr>
<td>10q22.3</td>
<td>ZMIZ1</td>
<td>rs704010</td>
<td>0.38</td>
<td>1.14 (1.10, 1.19)</td>
<td>$3.7 \times 10^{-10}$</td>
<td>1.15 (0.99, 1.33)</td>
<td>0.063</td>
<td>0.49</td>
</tr>
<tr>
<td>10p12.31</td>
<td>MLLT10</td>
<td>rs1243182</td>
<td>0.32</td>
<td>1.14 (1.09, 1.19)</td>
<td>$6.1 \times 10^{-9}$</td>
<td>1.29 (1.11, 1.49)</td>
<td>7.8 $\times 10^{-4}$</td>
<td>0.49</td>
</tr>
<tr>
<td>4q34.1</td>
<td>ADAM29</td>
<td>rs6828253</td>
<td>0.12</td>
<td>0.82 (0.77, 0.88)</td>
<td>$1.6 \times 10^{-8}$</td>
<td>0.96 (0.77, 1.21)</td>
<td>0.751</td>
<td>0.21</td>
</tr>
<tr>
<td>8q24.21</td>
<td></td>
<td>rs13281615</td>
<td>0.41</td>
<td>1.13 (1.08, 1.18)</td>
<td>$2.1 \times 10^{-8}$</td>
<td>1.13 (0.97, 1.30)</td>
<td>0.116</td>
<td>0.86</td>
</tr>
<tr>
<td>2q35</td>
<td>DIRC3</td>
<td>rs16857609</td>
<td>0.26</td>
<td>1.10 (1.05, 1.15)</td>
<td>$3.5 \times 10^{-5}$</td>
<td>1.25 (1.07, 1.46)</td>
<td>0.006</td>
<td>0.625</td>
</tr>
<tr>
<td>2q14.2</td>
<td></td>
<td>rs4849887</td>
<td>0.10</td>
<td>0.91 (0.84, 0.97)</td>
<td>0.007</td>
<td>0.71 (0.54, 0.93)</td>
<td>0.012</td>
<td>0.11</td>
</tr>
<tr>
<td>6q14.1</td>
<td></td>
<td>rs17529111</td>
<td>0.22</td>
<td>1.06 (1.01, 1.11)</td>
<td>0.020</td>
<td>1.25 (1.06, 1.48)</td>
<td>0.009</td>
<td>0.04</td>
</tr>
<tr>
<td>2p24.1</td>
<td></td>
<td>rs17270696</td>
<td>0.36</td>
<td>1.05 (1.00, 1.09)</td>
<td>0.037</td>
<td>1.17 (1.01, 1.35)</td>
<td>0.034</td>
<td>0.10</td>
</tr>
<tr>
<td>5p15.33</td>
<td>TERT-JH TERT</td>
<td>rs7726159</td>
<td>0.34</td>
<td>1.03 (0.99, 1.07)</td>
<td>0.195</td>
<td>1.22 (1.05, 1.42)</td>
<td>0.008</td>
<td>0.254</td>
</tr>
<tr>
<td>1q32.1</td>
<td>LGR6</td>
<td>rs6678914</td>
<td>0.41</td>
<td>1.02 (0.98, 1.06)</td>
<td>0.415</td>
<td>0.77 (0.67, 0.90)</td>
<td>8.0 $\times 10^{-4}$</td>
<td>0.0007</td>
</tr>
<tr>
<td>8q21.11</td>
<td></td>
<td>rs6472903</td>
<td>0.18</td>
<td>0.98 (0.93, 1.03)</td>
<td>0.459</td>
<td>0.81 (0.66, 0.99)</td>
<td>0.036</td>
<td>0.11</td>
</tr>
<tr>
<td>5p15.33</td>
<td>TERT</td>
<td>rs10069690</td>
<td>0.26</td>
<td>1.01 (0.97, 1.06)</td>
<td>0.615</td>
<td>1.18 (1.01, 1.38)</td>
<td>0.040</td>
<td>0.19</td>
</tr>
</tbody>
</table>

*5622 cases, 34272 controls.
**401 cases, 24045 controls.
doi:10.1371/journal.pgen.1004285.t003
75 of the known common breast cancer susceptibility loci were assessed for association with ILC and LCIS. As cases of ILC were included in the discovery sets that generated these susceptibility loci and lobular breast cancer is generally ER+ (94% of the ILC cases in this study were ER+) with the majority of ILCs classified as luminal tumors [32], it is not surprising that the majority of SNPs that we found to be associated with ILC were known to also predispose to ER+ breast cancer. However, some loci were only associated with ER+ IDC and not with ILC; particularly rs10941679 at 5p12, previously shown to predispose more strongly to ER-positive, lower-grade cancers [33]. None of these showed a much stronger association with ILC than IDC, particularly rs12149433 at 1p11.2, as previously described [26]. These data suggest specific etiological pathways for the development of different histological subtypes of breast cancer, in addition to the common pathways that predispose to multiple tumor subtypes.

Despite the small number of pure LCIS cases without invasive disease, our analyses have shown for the first time that many of the SNPs that predispose to ILC also predispose to LCIS. Although only 15 of the known breast cancer SNPs were associated with LCIS at P<0.05, 47 of the remaining 60 SNPs at P>0.05 had ORs in the same direction as for ILC (Sign Test P = 1.2×10^{-5}) suggesting that many more SNPs are likely to be associated with pure LCIS but did not reach statistical significance individually because of the relatively few LCIS cases without associated ILC in our sample set. This is not unexpected if LCIS is an intermediate phenotype for ILC. However, a small number of SNPs had differential effects on LCIS or ILC risk. Specifically, rs667914 at 1q32.1 (LGR6), known to be an ER-negative specific SNP [25], that appeared to be associated with LCIS but not ILC (P = 0.0007), and rs17529111 at 6q14 preferentially associated with ER-negative tumors [23] that had a stronger association with LCIS than IDC (P = 0.04). We also identified SNPs in FGFR2 and at 5q11.2 (MAP5K1) that appear only to predispose to ILC, but have little effect on LCIS suggesting that SNPs affect different parts of the lobular carcinoma pathway. These findings are surprising and as based on small numbers need confirmation in future studies.

Some of the SNPs associated with both ILC and LCIS showed a stronger effect size in LCIS compared to ILC for example SNPs at 7q35.2. 3q21.1, 7q31.3, 14q24.1, and 7q31.2. These SNPs are not unexpected because of the relatively few LCIS cases without associated ILC in this study. It is also strongly correlated with rs4572903 (r^2 = 0.69), with rs7072776, a recently identified ER+ breast cancer predisposition locus, which showed no association with LCIS in this study. It is also strongly correlated with rs1243180 (r^2 = 0.80), an ovarian cancer predisposition variant [34] and rs11012732 (r = 0.57), which predisposes to meningioma [35]. The ovarian SNP, rs1243180, also showed a strong association with lobular cancer (LCIS: P = 5.54×10^{-5}; OR = 1.11). Conditional analysis confirmed that this was not independent of rs1243180. SNPs that predispose strongly to LCIS were also associated with ER+ IDCs but again with stronger effect sizes in LCIS, consistent with the fact that 30–40% of invasive tumors associated with LCIS will not be ILC but will be IDC, mixed ductal-lobular or other morphology.

One SNP, rs1243182 (MLLT10), that showed a strong association with LCIS (LCIS: P = 7.8×10^{-5}, OR = 1.29; IDC: P = 6.1×10^{-3}, OR = 1.14), with LCIS: P = 5×10^{-3}, OR = 1.15. IDC: P = 1.4×10^{-3}, OR = 1.07, is partially correlated (r^2 = 0.69) with rs7072776, a recently identified ER+ breast cancer predisposition locus, which showed no association with LCIS in this study. The effect size is smaller in ILC as not all cases of LCIS will become invasive cancer. SNPs that predispose strongly to LCIS were also not associated with ER+ IDCs but again with stronger effect sizes in LCIS, consistent with the fact that 30–40% of invasive tumors associated with LCIS will not be ILC but will be IDC, mixed ductal-lobular or other morphology.

---

### Table 4. SNPs showing differential lobular and ductal associations with breast cancer risk in BCAC subjects (ER+ tumours only).

<table>
<thead>
<tr>
<th>IDC vs Controls*</th>
<th>ILC vs Controls**</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>P</strong></td>
<td><strong>P</strong></td>
</tr>
<tr>
<td>Cytoband</td>
<td>Genes</td>
</tr>
<tr>
<td>1p11.2</td>
<td>Cytoband</td>
</tr>
<tr>
<td>6p21.13</td>
<td>FGFR2</td>
</tr>
<tr>
<td>10q26.13</td>
<td>ZNF365</td>
</tr>
<tr>
<td>8q21.11</td>
<td>TOX3</td>
</tr>
<tr>
<td>2q31.1</td>
<td>CDCA7</td>
</tr>
</tbody>
</table>

*15023 cases, 29273 controls. **3201 cases, 29273 controls.
suggesting an association between breast cancer and meningioma [37].

In conclusion, we have identified a novel lobular-specific predisposition SNP at 7q34 close to JHDM1D that does not appear to be associated with IDC. Most known breast cancer predisposition SNPs also predispose to ILC, with some differential effects between ILC and IDC. In addition, many SNPs predisposing to invasive cancer are also likely to increase the risk for LCIS. Overall, our analyses show that genetic predisposition to IDC and lobular lesions (both ILC and LCIS) overlap to a large extent, but there are important differences that are likely to provide insights into the biology of lobular breast tumors.

Methods

Ethics statement

All studies were performed with ethical committee approval, Table S7, and subjects participated in the studies after providing informed consent.

Study populations

Phase I. Cases and controls came from 34 studies forming part of the Breast Cancer Association Consortium (BCAC) included in the COGS Project [13] (Table S1), and GLACIER (A study to investigate the Genetics of Lobular Ar Carcinoma In situ in Eu rope MREC 06/Q1702/64), a UK case-only study of lobular breast cancer. BCAC studies recruited all types of breast cancer. Pathological information in BCAC was collected by the studies individually but combined and checked through standardized data control in a central database. A total of 4,152 ILC and 89 LCIS cases were identified by the central BCAC pathology database (see Table S2 for number of cases by study).

The GLACIER study recruited patients from participating centers throughout the UK with the aim of identifying predisposition genes for LCIS and/or ILC. Any patients aged 60 or less at the time of diagnosis, with a current or past history of LCIS (with or without invasive disease of any histological subtype) were eligible. A total of 2,539 cases were recruited: 2,167 were identified from local pathology reports in 97 UK hospitals, 346 cases were identified through the British Breast Cancer Study (BBCS) using UK Cancer Registry data and 26 cases from the Royal Marsden Breast Tissue Bank. Cryptic relatedness analysis showed no evidence of overlap between these samples and the BCAC samples. All these cases were genotyped with the iCOGS chip and compared to 5,000 UK controls selected from four UK studies participating in BCAC and already typed on the iCOGS chip. Controls were randomly selected prior to analysis so that each of these UK studies, including GLACIER, had a case:control ratio of at least 1:2 (Table S8). These controls were excluded from case-control comparisons with BCAC cases from the originating study. This report includes only cases of pure LCIS or ILC with or without LCIS. Cases of LCIS with IDC or mixed lobular and ductal carcinoma in GLACIER were excluded in order to perform meta-analyses with the BCAC studies which do not have information on the presence or absence of LCIS associated with an invasive cancer. After excluding individuals based on genotyping quality (see Genotyping and Analysis) and non-European ancestry, data for the GLACIER study available for analyses included 1,782 cases (1,470 ILC (with or without LCIS), 312 pure LCIS) and 4,755 controls.

Phase II. A further 516 cases (481 ILC, 35 LCIS) and 1,465 controls were analyzed as part of Phase II. Controls were recruited through the GLACIER study, but were not genotyped in Phase I on the iCOGS chip to reduce costs, and were all white West European. Cases came from the following studies: 232 cases from GLACIER, 176 from BBCS, 71 from DietCompLyf [38], 39 from King’s Health Partners Cancer Biobank (KHP-CB). All cases were white West European, apart from the 39 samples from the KHP-CB where there were no associated ethnicity data. For studies that had also participated in Phase I, we selected samples so there was no overlap with the samples in Phase I.

Genotyping and analysis

Phase I. After DNA extraction from peripheral blood, GLACIER samples were genotyped on the iCOGS custom Illumina iSelect, which contains 211,155 SNPs, at King’s College, London. The remaining cases and controls were genotyped as part of the COGS project described in detail elsewhere [13]. The GLACIER cases were analyzed using the same QC criteria as the COGS project. Briefly, genotypes were called using Illumina’s proprietary GenCall algorithm and 10,000 SNPs were manually inspected to verify the algorithm calling. Individuals were excluded if genotyped not female, had overall call rate <95% or were ethnic outliers (248 cases) as identified by multi-dimensional scaling, combining the genotyping data with the three Hapmap2 populations. SNPs with a Gencall rate of <0.25, call rate <95% (call rate <99% if MAF <0.1) and HWE 10−7 or evidence of poor clustering on inspection of cluster plots were excluded. All SNPs with MAF <0.01 were excluded for this analysis. A cryptic relatedness analysis of the whole dataset was performed using 46,918 uncorrelated SNPs and there was no evidence of any duplicates.

For GLACIER cases and controls, principal component analysis (PCA) was carried out on a subset of 46,918 uncorrelated SNPs and used to exclude individuals or groups distinct from the main cluster using the first five principal components (PCs), Figure S7. Following removal of outliers (166 cases and 245 controls), the PCA was repeated and the first five PCs included as covariates in the analysis. The adequacy of the case-control matching was evaluated using quantile-quantile plots of test statistics and the inflation factor (λ) calculated using only uncorrelated SNPs that were not selected by BCAC and were not within one of the four common fine-mapping regions, to minimize selection for SNPs associated with breast cancer, Figure S1. As the majority of the SNPs on the iCOGS array were selected from GWAS of breast, ovarian and prostate cancer the SNPs selected for this analysis were taken from the set of SNPs selected by the prostate consortium, with the assumption that these SNPs were more likely to be representative of common SNPs in terms of population structure in our study than those selected by the breast or ovarian consortia.

For each SNP, we estimated a per-allele log-odds ratio (OR) and standard error by logistic regression, including the 5 PCs as covariates, using PLINK v1.07 (http://pngu.mgh.harvard.edu/purcell/plink/).

Genotyping and analysis of BCAC studies is described in detail elsewhere [24], in brief data were analyzed using the Genotype Library and Utilities (GLU) package to estimate per-allele ORs and standard errors for each SNP using unconditional logistic regression. All analyses were performed in subjects of European ancestry (determined by PC analyses) and adjusted for study and seven principal components.

Case-control odds ratio (OR) for ILC or LCIS cases vs controls from BCAC and GLACIER were combined using inverse variance-weighted fixed-effects meta-analysis, as implemented in
METAL. [39]. Case-only analyses were also carried out to compare genotype frequencies for ILC vs LCIS (GLACIER and BCAC) and ILC vs IDC (BCAC studies only), and were used as a test for heterogeneity of ORs by tumor subtype. Any study without data on both histological subtypes was dropped from the case-only analysis.

Phase II. SNPs showing the strongest evidence for association with lobular tumors (P<3×10⁻⁵) in the meta-analysis (after excluding previously reported loci) were genotyped at LGC Genomics (formerly KBiosciences) in Phase II samples. Duplicate samples genotyped on the iCOGS chip were included to assess the concordance of the two genotyping methods. Cluster plots for rs11977670 are shown in Figure S8.

A pooled analysis of ILC including Phase I (GLACIER and BCAC) and Phase II data was performed. Data were analyzed using STATA v.12 to estimate per-allele ORs and standard errors for each SNP using unconditional logistic regression. Differences in the strength of the associations with ILC, IDC and LCIS were assessed using case-only analyses. A sign test was used to test whether the number of SNPs showing associations in the same direction in two different subtypes (i.e. LCIS vs ILC, and IDC vs ILC) was significantly greater than expected by chance. A likelihood ratio test was used as a global test of the null hypothesis of no differences between subtypes for any of the ORs of the 75 known loci evaluated. Stratum-specific estimates of per-allele OR by categories of age and family history of disease were obtained from logistic regression models and differences in ORs across strata were tested using an interaction term.

Bioinformatics
In order to establish the SNP’s functional role, a window of 10 kb both up and downstream was formed around the marker and pairwise r² values calculated using 1000 genome CEU population data. Three SNPs were identified as being in LD (r²>0.5) with rs11977670 and were compared to next generation sequence technologies to elucidate the overlap between chromatin states (ENCODE Project). Two cell lines, normal human mammary epithelial (HMEC), and breast carcinoma (MCF-7), were used to establish these chromatin states, i.e. active or engaged enhancers (H3K27ac), nucleosome-depleted regions (DNase I and FAIRE), and RNA polymerase linked regions (Pol II). Expression data from the Cancer Genome Atlas Network for each gene within a 1 Mb window of rs11977670 was analyzed looking for differential expression in each breast cancer subtype (ER+ IDC; 40 cases; ER+ IDC, 341 cases; and ER-negative IDC, 108 cases). Allele data for surrogate SNP rs13225058 was obtained for all ER+ cases from TCGA. These 335 cases were used to produce genotype specific gene expression data in R. Differences in gene expression between the three genotypes were tested for using one-way-anova, verified by t-test and visually by boxplot. Linear regression was performed across all three genotypes using copy number variation as a co-variate. Level 3 copy number variation data [hg19 build] was obtained from the TCGA data portal.

Supporting Information

**Figure S1** Quantile-quantile plot for GLACIER. A: QQ plot based on the 37544 uncorrelated SNPs not selected on the basis of breast cancer risk (λ = 1.04). B: QQ plot for all SNPs in dataset (λ = 1.09). (PPTX)

**Figure S2** LD block containing rs1197790. (PPTX)

**Figure S3** rs1197790 falls in a high H3K27ac region using ENCODE data from normal human mammary epithelial (HMEC), and breast carcinoma (MCF-7) cell lines to establish chromatin states in the region. (PPTX)

**Figure S4** Gene expression data taken from TCGA for genes in a 1 Mb window of rs11977670. Three genes showed differential expression in ER+ ILC compared to ER+ IDC (BRAF, P = 0.006; NDUF22, P = 0.02; SLC37A3, P = 0.05). (PPTX)

**Figure S5 a:** Genotype specific gene expression in ER+ Breast Cancers. Gene expression and genotype data was taken from TCGA and compared using a surrogate for rs11977670. A significant difference between the AA and GG genotype was only found for 353 ER+ cancers. A significant difference between the AA and GG genotype was only found for two genes, JHDM1D and SLC37A3. **b:** Genotype specific gene expression in 36 Invasive Lobular Cancers. Gene expression and genotype data was taken from TCGA and compared using a surrogate for rs11977670, rs1225058 (r² = 0.79).

**Figure S6** Genotype specific gene expression in 48 cases of normal breast tissue associated with ER+ breast cancer. Gene expression and genotype data was taken from TCGA and compared using a surrogate for rs11977670, rs1225058 (r² = 0.79) for 48 cases with normal breast tissue. (PPTX)

**Figure S7** Results of principal components analysis (PCA) – GLACIER cohort. A: PCA with the 3 HapMap populations. B: PCA after exclusion of outliers (414 cases and 245 controls). (PPTX)

**Figure S8** Cluster plots for rs11977670 is on chromosome 7 (139942304). A: Phase I – iCOGS Array – GLACIER (Illumina). B: Phase I – iCOGS Array – BCAC (Illumina). C: Phase II KASPAR (LGC Genomics). (PPTX)

**Table S1** Participating studies from the BCAC.

**Table S2** Number of lobular breast cancer cases per study.

**Table S3** Genotype-specific odds ratios for rs11977670 and risk of lobular-specific breast cancer (based on pooled analysis of phase I and II).

**Table S4** Results for borderline SNPs not reaching GWS after Phase II.

**Table S5** Pooled lobular analysis of known SNPs (BCAC and GLACIER).

**Table S6** Lobular and ductal associations with breast cancer risk in BCAC subjects (ER pos only).

**Table S7** Details of ethical approval boards for each study.

**Table S8** Lobular cases and controls from UK BCAC studies. Controls from each of these studies were randomly selected to obtain a control group for GLACIER cases in a 1:2 case to control ratio.
Acknowledgments


References


Author Contributions

Conceived and designed the experiments: ES RR IT MG. Performed the experiments: CP EPa AGN GP MRA DCT DV FB JD AD. Analyzed the data: ES CP EPa AGN GP MRA DCT DV FB JD AD. Contributed reagents/materials/analysis tools: ES RR CP MNB SN MKB QW KM IT MG. Wrote the paper: ES CP MNB SN MKB QW KM IT MG. Provided critical review of the manuscript: ES RR IT MG. Contributed to the discussion: ES CP MNB SN MKB QW KM IT MG.

Acknowledgments


Author Contributions

Conceived and designed the experiments: ES RR IT MG. Performed the experiments: CP EPa AGN GP MRA DCT DV FB JD AD. Analyzed the data: ES CP EPa AGN GP MRA DCT DV FB JD AD. Contributed reagents/materials/analysis tools: ES RR CP MNB SN MKB QW KM IT MG. Wrote the paper: ES CP MNB SN MKB QW KM IT MG. Provided critical review of the manuscript: ES RR IT MG. Contributed to the discussion: ES CP MNB SN MKB QW KM IT MG.

Author Contributions

Conceived and designed the experiments: ES RR IT MG. Performed the experiments: CP EPa AGN GP MRA DCT DV FB JD AD. Analyzed the data: ES CP EPa AGN GP MRA DCT DV FB JD AD. Contributed reagents/materials/analysis tools: ES RR CP MNB SN MKB QW KM IT MG. Wrote the paper: ES CP MNB SN MKB QW KM IT MG. Provided critical review of the manuscript: ES RR IT MG. Contributed to the discussion: ES CP MNB SN MKB QW KM IT MG.

Author Contributions

Conceived and designed the experiments: ES RR IT MG. Performed the experiments: CP EPa AGN GP MRA DCT DV FB JD AD. Analyzed the data: ES CP EPa AGN GP MRA DCT DV FB JD AD. Contributed reagents/materials/analysis tools: ES RR CP MNB SN MKB QW KM IT MG. Wrote the paper: ES CP MNB SN MKB QW KM IT MG. Provided critical review of the manuscript: ES RR IT MG. Contributed to the discussion: ES CP MNB SN MKB QW KM IT MG.

Author Contributions

Conceived and designed the experiments: ES RR IT MG. Performed the experiments: CP EPa AGN GP MRA DCT DV FB JD AD. Analyzed the data: ES CP EPa AGN GP MRA DCT DV FB JD AD. Contributed reagents/materials/analysis tools: ES RR CP MNB SN MKB QW KM IT MG. Wrote the paper: ES CP MNB SN MKB QW KM IT MG. Provided critical review of the manuscript: ES RR IT MG. Contributed to the discussion: ES CP MNB SN MKB QW KM IT MG.


