2q36.3 is associated with prognosis for oestrogen receptor-negative breast cancer patients treated with chemotherapy

Jingmei Li1,2,*, Linda S. Lindström3,4,*, et al.#

Large population-based registry studies have shown that breast cancer prognosis is inherited. Here we analyse single-nucleotide polymorphisms (SNPs) of genes implicated in human immunology and inflammation as candidates for prognostic markers of breast cancer survival involving 1,804 oestrogen receptor (ER)-negative patients treated with chemotherapy (279 events) from 14 European studies in a prior large-scale genotyping experiment, which is part of the Collaborative Oncological Gene-environment Study (COGS) initiative. We carry out replication using Asian COGS samples (n = 522, 53 events) and the Prospective Study of Outcomes in Sporadic versus Hereditary breast cancer (POSH) study (n = 315, 108 events). Rs4458204_A near CCL20 (2q36.3) is found to be associated with breast cancer-specific death at a genome-wide significant level (n = 2,641, 440 events, combined allelic hazard ratio (HR) = 1.81 (1.49–2.19); P for trend = 1.90 × 10⁻⁹). Such survival-associated variants can represent ideal targets for tailored therapeutics, and may also enhance our current prognostic prediction capabilities.
We have previously shown, through large population-based registry studies, that survival from breast cancer is correlated among relatives, consistent with an inherited cancer prognosis. A potential explanation for the heritability of survival would be that family members are predisposed to developing a breast cancer tumour of predefined aetiology and predetermined tumour characteristics. This is plausible given the observation that carriers of high- and moderate-risk germline mutations in genes such as BRCA1, BRCA2, CHEK2 and PALB2, are predisposed to specific subtypes of breast cancer, and that many common variants identified through genome-wide association studies (GWAS) tend to be associated with specific subtypes, with some variants more strongly associated with oestrogen receptor (ER)-negative or triple-negative breast cancer, while others more strongly associated with ER-positive breast cancer.

It is also possible that the inherited predeterminants of survival lie not in the biology of the tumour but rather the milieu in which the tumour arises. The tumour microenvironment is composed of tumour cells, fibroblasts, endothelial cells and infiltrating immune cells, which may either inhibit or promote tumour growth and progression. There is empirical support for the concept that a host immune response might enhance the effects of conventional chemotherapy, conceivably having an influence on breast cancer outcome. For example, the presence of tumour-associated lymphocytes in a breast tumour has been suggested to be an independent predictor of neoadjuvant chemotherapy response. Other studies have shown the host immune system to be involved in the elimination of tumour cells to control cancer growth.

In this candidate pathway study, we investigate the pre-specified hypothesis that the germline common variants of genes involved in immune response and inflammation can predict the response to breast cancer survival for ER-negative, chemotherapy-treated patients. We identify a single-nucleotide polymorphism (SNP) near the CCL20 gene (2q36.3), which is associated with a difference in the clinical outcome of ER-negative breast cancer treated with chemotherapy independent of known tumour prognostic features.

**Results**

Individual patient-level genetic and phenotypic data were extracted from European studies in a prior large-scale genotyping experiment conducted in the Breast Cancer Association Consortium (BCAC), part of the Collaborative Oncological Gene-environment Study (COGS) initiative. For this study, we selected women of European descent inferred from genetic ancestry with invasive breast cancer, who have had no previous diagnosis of the disease. Subjects missing follow-up information on vital status, time to vital status, date of study entry and cause of death data were excluded.

The selection of only ER-negative patients in this study was strongly motivated by prior insight. A Swedish study of the breast cancer prognosis of 834 sister pairs in which both were affected showed that younger sisters with poor older sister survival had worse survival than younger sisters with good older sister survival (number of breast cancer deaths within 5 years from diagnosis in younger sisters, 65, P = 0.02) in a multivariate proportional hazard (Cox) analysis. When stratified by ER subtypes, the increased risk of death from ER-negative breast cancer for younger sisters with poor older sister survival compared with younger sisters with good older sister survival was found to be almost sevenfold (n = 139 sister pairs, 28, hazard ratio (HR) = 6.69 (1.36–32.91), P = 0.02) in contrast to sister pairs with the ER-positive disease (n = 584 sister pairs, 28, HR = 1.54 (0.48–4.98), P = 0.50) (unpublished data).

In addition, in a recent Breast International Group phase III trial, increasing lymphocytic infiltration was found to be associated with excellent prognosis only for patients with node-positive, ER-negative/HER2-negative disease. Twenty studies with ER-negative cases and at least one event (breast cancer-specific death) were eligible for the combined analysis (Supplementary Table 1). As we were primarily interested in response to chemotherapy, patients missing information on chemotherapy were not considered in our analyses. The 14 studies (n = 1,804) included in the combined analysis for the chemotherapy-treated subgroup are summarized in Supplementary Table 2. A total of 279 breast cancer-specific deaths were recorded in a 15-year follow-up.

For the replication phase, four iCOGS Asian studies with ER-negative breast cancer cases treated with chemotherapy and at least one death due to breast cancer in a 15-year follow-up were analysed (n = 522, 53 events, Supplementary Table 3). Early-onset breast cancer patients from the independent Prospective Study of Outcomes in Sporadic versus Hereditary breast cancer (POSH) study were used as a second replication data set. In particular, we performed our replication using ER-negative breast cancer patients treated with chemotherapy in the POSH study’s Stage 1 discovery data set samples (n = 315, 108 events) selected to facilitate studies on breast cancer prognosis. The breast-cancer-specific death rate is thus particularly high and there were few cases that drop out due to lack of phenotype information.

All women in participating studies had provided written consent for the research and approval for each study was obtained from their local ethical review board (Supplementary Tables 1 and 3). Collection of blood samples and clinical data from subjects was performed in accordance with local guidelines and regulations.
Genotyping was conducted using a custom Illumina iSelect genotyping array (iCOGS), comprising 211,115 SNPs. Details of quality control of the iCOGS data are described in detail elsewhere. Briefly, individuals were excluded for any of the following reasons: genotypically not female XX (XY, XXY or XO), overall call rate <95%, low or high heterozygosity (P<1×10^{-6}, determined separately for individuals of European and East Asian ancestry), genotypes discordant with those determined in previous genotyping such that the individual appeared to be different, genotypes for the duplicate sample that seemed to be from a different individual and cryptic duplicates. SNPs with call rates of <95%, SNPs that deviated from Hardy–Weinberg equilibrium in controls at P<1×10^{-7} and SNPs for which the genotypes were discrepant in 4.2% of duplicate samples across all COGS consortia were excluded. The final analyses in the parent COGS study were based on data from 199,961 SNPs.

Key genes related to human immunology and inflammation were identified from two comprehensive and highly curated gene panels (nCounter GX Human Immunology Kit and nCounter GX Human Inflammation Kit, NanoString Technologies, Seattle, WA, USA), which are commercially available (Supplementary Data 1). We identified all SNPs on the iCOGS within a 50-kb window of any gene on the panel. Out of 8,237 unique SNPs extracted from COGS, we further removed SNPs with low minor allele frequency.

**Table 1 | Summary of results for association of rs4458204_A with risk of dying from breast cancer.**

| Patients                                      | n  | Breast cancer-specific deaths | Per-allele HR (95% CI)* | P-value |
|-----------------------------------------------|----|-------------------------------|-------------------------|---------|
| **Discovery**                                 |    |                               |                         |         |
| ER-negative                                   | 2,218 | 332                          | 1.83 (1.47–2.27)        | 4.68×10^{-8} |
| ER-negative not treated with chemotherapy     | 411  | 53                            | 1.39 (0.69–2.81)        | 0.36    |
| ER-negative and treated with chemotherapy     | 1,804 | 279                          | 1.96 (1.55–2.47)        | 1.60×10^{-8} |

**Replication**

| ER-negative and treated with chemotherapy     | 522  | 53                            | 1.97 (0.94–4.17)        | 0.07    |
| ER-negative and treated with chemotherapy     | 315  | 108                           | 1.41 (0.95–2.09)        | 0.08    |

**Combined replication**

| ER-negative and treated with chemotherapy     | 837  | 161                          | 1.52 (1.07–2.15)        | 0.02    |

**Combined overall**

| ER-negative and treated with chemotherapy     | 2641 | 440                          | 1.81 (1.49–2.19)        | 1.90×10^{-9} |

---

**Figure 2 | Manhattan plot for association in the discovery stage.** Manhattan plot showing directly genotyped SNPs plotted according to chromosomal location (x-axis), with –log10 P-values (y-axis) derived from trend tests of association of 7,020 human immunology and inflammation SNPs with the risk of dying from breast cancer for all ER-negative patients (above) and ER-negative patients treated with chemotherapy (below) in the discovery phase. Blue and red lines indicate the Bonferroni threshold for multiple-testing correction for 2,184 (r^2>0.2) and genome-wide significance level (5×10^{-8}), respectively. SNPs with FDRs of <10% are additionally encircled and denoted in green. Chromosomal positions are based on NCBI build 36.

Genotyping was conducted using a custom Illumina iSelect genotyping array (iCOGS), comprising 211,115 SNPs. Details of quality control of the iCOGS data are described in detail elsewhere. Briefly, individuals were excluded for any of the following reasons: genotypically not female XX (XY, XXY or XO), overall call rate <95%, low or high heterozygosity (P<1×10^{-6}, determined separately for individuals of European and East Asian ancestry), genotypes discordant with those determined in previous genotyping such that the individual appeared to be different, genotypes for the duplicate sample that seemed to be from a different individual and cryptic duplicates. SNPs with call rates of <95%, SNPs that deviated from Hardy–Weinberg equilibrium in controls at P<1×10^{-7} and SNPs for which the genotypes were discrepant in >2% of duplicate samples across all COGS consortia were excluded. The final analyses in the parent COGS study were based on data from 199,961 SNPs.

Key genes related to human immunology and inflammation were identified from two comprehensive and highly curated gene panels (nCounter GX Human Immunology Kit and nCounter GX Human Inflammation Kit, NanoString Technologies, Seattle, WA, USA), which are commercially available (Supplementary Data 1). We identified all SNPs on the iCOGS within a 50-kb window of any gene on the panel. Out of 8,237 unique SNPs extracted from COGS, we further removed SNPs with low minor allele frequency.
and low call rate (<0.95). After quality-control exclusions, we analysed 7,020 non-overlapping SNPs in 557 unique gene regions (from 597 genes on the original nCounter panels).

In the POSH study, rs4458204 was genotyped on the Illumina 660 W-Quad SNP array. Details can be found in the parent POSH article. Briefly, genotyping for the samples was conducted in

| Study   | n  | Events | maf | Hazard ratio (95% CI) |
|---------|----|--------|-----|-----------------------|
| ABCFS   | 149| 32     | 0.14| 2.07 (1.07, 4.01)     |
| KBCP    | 28 | 11     | 0.04| 1.43 (0.07, 29.07)    |
| MARIE   | 279| 53     | 0.12| 1.30 (0.70, 2.43)     |
| OFBCR   | 101| 10     | 0.14| 0.52 (0.06, 4.67)     |
| RBCS    | 76 | 18     | 0.08| 2.75 (0.49, 15.43)    |
| SEARCH  | 505| 107    | 0.14| 1.97 (1.41, 2.76)     |
| SKKDKFZS| 78 | 14     | 0.14| 4.69 (1.12, 19.64)    |

Random effects: 1.88 (1.45, 2.43)

Fixed effects: 1.88 (1.45, 2.43)

Test for heterogeneity: I² = 0%; P = 0.6
NATURE COMMUNICATIONS | DOI: 10.1038/ncomms5051

ARTICLE

SNPs were thinned using the ‘—indep-pairwise’ option in PLINK such that all SNPs within a window size of 50 SNPs (step size of 10) were required to have \( r^2 < 0.2 \). This procedure resulted in a set of 2,184 independent SNPs pruned by linkage disequilibrium. The Bonferroni-adjusted threshold for 2,184 independent tests is \( 2.29 \times 10^{-5} \). In addition to standard Bonferroni adjustment, a 10% false discovery rate (FDR) threshold was applied to try to identify more candidate SNPs associated with breast cancer outcome. An FDR-adjusted \( P \)-value of 0.10 implies that 10% of significant tests will result in false positives.

The results for tests of association between 7,020 human immunology and inflammation SNPs and risk of death from ER-negative breast cancer are summarized in Supplementary Data 2 and 3. The deviation of the smaller observed \( P \)-values from those expected (\( \lambda = 1.16 \)) is consistent with multiple weak associations between these SNPs and survival for ER-negative breast cancer patients (Fig. 1). In particular, for a single SNP rs4458204_A located on chromosome 2:228637113 (minor allele frequency = 0.12), the \( \chi^2 \) (1 df) association test statistic was much higher than for the other SNPs and was close to surpassing the threshold for experiment-wide significance after Bonferroni adjustment (\( P < 2.29 \times 10^{-5} \)) in the partially adjusted analysis stratified by study and adjusted for population stratification and age (\( n = 2,218,332 \) events, per-allele HR = 1.54 (1.26–1.90), \( P \) for trend = 3.62 \( \times 10^{-5} \); Supplementary Data 3). However, after further adjusting for appropriate patient tumour and treatment characteristics, the SNP association surpassed the threshold for genome-wide significance (\( P < 5 \times 10^{-8} \)) (per-allele HR = 1.83 (1.47–2.27), \( P \) for trend = 4.68 \( \times 10^{-8} \); Table 1 and Fig. 2), a conservative threshold which is likely to be overly stringent\(^{27}\). The lack of an association signal tower could be because the iCOGS was designed to have minimum linkage disequilibrium across SNPs. No SNP within a 100-kb window is correlated to rs4458204 with \( r^2 > 0.2 \) (Fig. 3). The association was stronger for a subset of ER-negative patients who had been treated with chemotherapy (\( n = 1,804,279 \) events, per-allele HR = 1.96 (1.55–2.47), \( P \) for trend = 1.60 \( \times 10^{-8} \)). We found no evidence of heterogeneity in the per-allele HR across 14 studies (\( P^2 = 0\%\), \( P \) for heterogeneity = 0.84; forest plot in Fig. 4).

Univariate Kaplan–Meier survival curves of breast cancer-specific survival for ER-negative patients treated with chemotherapy by rs4458204 genotypes are presented in Fig. 5 (log-rank \( P = 3.18 \times 10^{-6} \)). The median survival time for the AA genotype at rs4458204 was 11.5 years. SNPs in three other loci corresponding to regions around the transforming growth factor beta receptor II (\( TGFB2 \)), interleukin 12B (\( IL12B \)) and interferon induced with helicase C domain 1 (\( IFIH1 \)) genes were found to be associated with breast cancer-specific death with FDR-adjusted \( P < 0.10 \) (Fig. 2).

From our replication study of rs4458204_A using multi-ethnic iCOGS Asian samples (522 ER-negative patients treated with chemotherapy, 53 events; see Supplementary Table 3), the per-allele HR after controlling for tumour characteristics and treatment was 1.97 (0.94–4.17; \( P \) for trend = 0.07, Table 1). Together with multivariable-adjusted results from a second replication of the SNP using early-onset breast cancer patients POSH study, significant evidence of replication was observed (combined per-allele HR = 1.52 (1.07–2.15), \( P \) for trend = 0.02, Table 1). From a meta-analysis of both discovery and replication stages, the association of the SNP with risk of dying from breast cancer was found to be 1.81 (1.49–2.19; \( P \) for trend = 1.90 \( \times 10^{-9} \)) with no observed heterogeneity (\( I^2 = 1.4\%\), \( P \) for heterogeneity = 0.36; Table 1).

The cluster plots for the most significant SNP in our analysis, rs4458204 (\( CCL20 \)), and three other index SNPs of loci for which two separate batches in two locations (Mayo Clinic and the Genome Institute of Singapore). To ensure harmonization of the genotype calling, the intensity data were combined and used to generate genotypes based on the algorithm available in the genotyping module of Illumina’s Genome Studio software.
representative SNPs indicate good discrimination of the three genotypes. homozygote genotypes for the SNP (AA/aa), green heterozygote (Aa) and black undetermined. Three distinct, tight clusters exhibited by all four on cluster formation in scatter plots of normalized allele intensities X and Y. Each circle represents one individual’s genotype. Blue and red clouds indicate homozgyote genotypes for the SNP (AA/aa), green heterozygote (Aa) and black undetermined. Three distinct, tight clusters exhibited by all four representative SNPs indicate good discrimination of the three genotypes.

the associated test statistic passed FDR<0.1, namely rs1367610 (TGFBR2), rs2569254 (IL12B) and rs13422767 (IFIH1), were examined. All SNPs showed good discrimination of the three genotypes in cluster plots for the BCAC samples that passed quality control in the parent COGS study (Fig. 6).

Discussion

rs4458204 is located ~41.5 kb upstream of the chemokine (C–C motif) ligand 20 (CCL20) gene. Chemokines are important mediators of immune response, and CCL20 has previously been shown to induce migration and proliferation of breast epithelial cells. CCL20 has also been reported to be strongly chemotactic for lymphocytes and weakly attracts neutrophils. However, rs4458204 was not found to be a significant (P for trend > 0.05) expression trait quantitative locus in any of the tissues (that is, adipose subcutaneous, artery tibial, blood, heart, lung, muscle skeletal, nerve tibial, skin and thyroid) reported on the publicly available Genotype-Tissue Expression Portal.

It is of note that the association of rs4458204_A with the survival of ER-negative breast cancer patients treated with chemotherapy increased and the strength of the association became stronger after adjustment for tumour characteristics and type of treatment (per-allele HR (95% confidence interval) from 1.64 (1.31–2.05) to 1.96 (1.55–2.47), P for trend from 1.27 × 10⁻⁵ to 1.60 × 10⁻⁸). This suggests that tumour characteristics and treatment covariates are likely to be confounders and thus it is desirable to include them in the fully adjusted model to obtain a more accurate effect size of the genetic factor. Moreover, it has also been shown that adjustment for prognostic factors will lead to a gain in power for statistical analyses. Genes in other regions identified by the less stringent FDR threshold (TGFBR2, IL12B and IFIH1) have been implicated to play a role in breast cancer disease progression, suggesting that there are potentially more variants in immune response and inflammation genes that are associated with breast cancer prognosis. Although TGFBR2 is a breast cancer susceptibility locus, none of the SNPs annotated to this gene was significantly associated with breast cancer risk (P>0.05) in the parent COGS study.

Although several GWAS have aimed to find genetic markers associated with breast cancer survival to date, few credible variants have been robustly identified. The threefold greater breast cancer mortality for affected sisters is comparable in magnitude to the familial relative risk for breast cancer incidence, for which close to 100 independent susceptibility loci based on common variants (SNPs) have been identified, and these explain only a small proportion of familial aggregation of risk. The failure to identify a similar number of survival-associated loci influencing survival may reflect the much lower statistical power for survival analyses to date, but may also reflect the substantial heterogeneity in tumour characteristics and treatment. As such, it...
has been suggested that sufficiently powered studies investigating specific cancer subtypes or treatment subgroups would need to be much larger to discover more regions in the genome associated with breast cancer prognosis. In agreement, the association between rs4458204 and breast cancer survival for this study was found to be more pronounced (larger HR) for women with ER-negative disease treated with chemotherapy (Table 1). However, as we did not study the association for women with ER-positive disease, the impact of this SNP on survival for those women remains unclear. One of the strengths of our study is that we have based our gene selection on commercially pre-designed panels of genes known to be differentially expressed in immunology and inflammation, which covers a comprehensive and validated list of relevant genes. The use of the iCOGS array in the BCAC consortium allowed us to investigate genetic variation across >500 immune response genes and provided an unprecedented large sample size with detailed clinical information to examine their associations with breast cancer survival. The results were also replicated by the POSH study, which is not part of the COGS consortium. However, SNPs related to immune response and inflammation were not specifically selected to be put on the iCOGS panel to give comprehensive coverage of these genes; only 557 of the 597 genes (~93%) were represented. The proportion of total phenotypic variance (Nagelkerke pseudo-R-squared) explained by this SNP alone was also small, at ~1.3%, suggesting that many more variants will need to be discovered for such genetic data to be useful in a clinical setting.

Our findings suggest that host factors affecting the ability to respond to systemic treatment or to mount an effective immunologic response contribute to the heritability of prognosis. Such survival-associated variants can represent ideal targets for tailored therapeutics and may also enhance our current prognostic prediction capabilities.

References

1. Lindstrom, L. S. et al. Familial concordance in cancer survival: a Swedish population-based study. Lancet Oncol. 8, 1001–1006 (2007).
2. Hartman, M. et al. Breast cancer prognosis inherited? Breast Cancer Res. 9, R39 (2007).
3. Lindström, L. et al. Association of breast cancer survival among sisters is not explained by shared primary tumour characteristics. Submitted (2013).
4. Verkooijen, H. M. et al. Breast cancer prognosis is inherited independently of patient, tumor and treatment characteristics. Int. J. Cancer 130, 2103–2110 (2012).
5. Roy, R., Chun, J. & Powell, S. N. BRCA1 and BRCA2: different roles in a common pathway of genome protection. Nat. Rev. Cancer 12, 68–78 (2012).
6. Honrado, E., Osorio, A., Palacios, J. & Benitez, J. Pathology and gene expression in inflammation were not specifically selected to be put on the iCOGS panel to give comprehensive coverage of these genes; only 557 of the 597 genes (~93%) were represented. The proportion of total phenotypic variance (Nagelkerke pseudo-R-squared) explained by this SNP alone was also small, at ~1.3%, suggesting that many more variants will need to be discovered for such genetic data to be useful in a clinical setting.

Our findings suggest that host factors affecting the ability to respond to systemic treatment or to mount an effective immunologic response contribute to the heritability of prognosis. Such survival-associated variants can represent ideal targets for tailored therapeutics and may also enhance our current prognostic prediction capabilities.

Acknowledgements

We thank all the individuals who took part in these studies and all the researchers, study staff, clinicians and other health care providers, technicians and administrative staff who have enabled this work to be carried out. This work was financed by the Swedish Research Council (grant number: 521-2001-3187 and 524-2011-6857 to LL; Gösta Miltons Donationsfond to LL, Swedish Cancer Society (grant number: CAN 2010/807 and NKI2009-4363. The work of the BBCC was partly funded by ELAN-Programme of the European Community’s Seventh Framework Programme under grant agreement number 223175 (HEALTH-F2-2009-232175 (COGS), Cancer Research UK (C1287/A10118, C1287/A10710, C1287/A112174, C1281/A12014, C5047/A8384, C5047/A15007, C5047/A10692), the National Institutes of Health (CA128978) and Post-Cancer GWAS initiative (1U19 CA148537, U19 CA148065 and U19 CA148112—the GAME-ON initiative), the Department of Defence (W81XWH-10-1-0341), the Canadian Institutes of Health Research (CIHR) for the CIHR Team in Familial Risks of Breast Cancer, Komen Foundation for the Cure, the Breast Cancer Research Foundation and the Ovarian Cancer Research Fund. The ABCS and OFBCR work was supported by the Canadian Institutes of Health Research (CIHR) in Familial Risks of Breast Cancer, Komen Foundation for the Cure, the Breast Cancer Research Foundation and the Ovarian Cancer Research Fund. The ABCS and OFBCR work was supported by grant UM1 CA165492 from the National Cancer Institute (USA). The OFBCR work was also supported by the Canadian Institutes of Health Research (CIHR) in Familial Risks of Breast Cancer’ programme. The ABCS was funded by Dutch Cancer Society grant number NKI2007-3839, M.K.S. was funded by Dutch Cancer Society grant number NKI2009-4363. The work of the BBCC was partly funded by ELAN-Programme of the University Hospital of Erlangen. The BCAC is funded by CR-UK (C1287/A10118 and C1287/A12014). Meetings of the BCAC have been funded by the European Union COST programme (BM0606). E.S. is supported by NIH Comprehensive Biomedical Research Centre, Guy’s & St Thomas’ NHS Foundation Trust in partnership with King’s College London, United Kingdom. I.T. is supported by the Oxford Biomedical Research Centre.

References

1. Lindstrom, L. S. et al. Familial concordance in cancer survival: a Swedish population-based study. Lancet Oncol. 8, 1001–1006 (2007).
2. Hartman, M. et al. Breast cancer prognosis inherited? Breast Cancer Res. 9, R39 (2007).
3. Lindström, L. et al. Association of breast cancer survival among sisters is not explained by shared primary tumour characteristics. Submitted (2013).
4. Verkooijen, H. M. et al. Breast cancer prognosis is inherited independently of patient, tumor and treatment characteristics. Int. J. Cancer 130, 2103–2110 (2012).
5. Roy, R., Chun, J. & Powell, S. N. BRCA1 and BRCA2: different roles in a common pathway of genome protection. Nat. Rev. Cancer 12, 68–78 (2012).
6. Honrado, E., Osorio, A., Palacios, J. & Benitez, J. Pathology and gene expression in inflammation were not specifically selected to be put on the iCOGS panel to give comprehensive coverage of these genes; only 557 of the 597 genes (~93%) were represented. The proportion of total phenotypic variance (Nagelkerke pseudo-R-squared) explained by this SNP alone was also small, at ~1.3%, suggesting that many more variants will need to be discovered for such genetic data to be useful in a clinical setting.

Our findings suggest that host factors affecting the ability to respond to systemic treatment or to mount an effective immunologic response contribute to the heritability of prognosis. Such survival-associated variants can represent ideal targets for tailored therapeutics and may also enhance our current prognostic prediction capabilities.

Acknowledgements

We thank all the individuals who took part in these studies and all the researchers, study staff, clinicians and other health care providers, technicians and administrative staff who have enabled this work to be carried out. This work was financed by the Swedish Research Council (grant number: 521-2001-3187 and 524-2011-6857 to LL; Gösta Miltons Donationsfond to LL, Swedish Cancer Society (grant number: CAN 2010/807 and NKI2009-4363. The work of the BBCC was partly funded by ELAN-Programme of the European Community’s Seventh Framework Programme under grant agreement number 223175 (HEALTH-F2-2009-232175 (COGS), Cancer Research UK (C1287/A10118, C1287/A10710, C1282/A112174, C1281/A12014, C5047/A8384, C5047/A15007, C5047/A10692), the National Institutes of Health (CA128978) and Post-Cancer GWAS initiative (1U19 CA148537, U19 CA148065 and U19 CA148112—the GAME-ON initiative), the Department of Defence (W81XWH-10-1-0341), the Canadian Institutes of Health Research (CIHR) for the CIHR Team in Familial Risks of Breast Cancer, Komen Foundation for the Cure, the Breast Cancer Research Foundation and the Ovarian Cancer Research Fund. The ABCS and OFBCR work was supported by grant UM1 CA165492 from the National Cancer Institute (USA). The OFBCR work was also supported by the Canadian Institutes of Health Research (CIHR) Team in Familial Risks of Breast Cancer’ programme. The ABCS was funded by Dutch Cancer Society grant number NKI2007-3839, M.K.S. was funded by Dutch Cancer Society grant number NKI2009-4363. The work of the BBCC was partly funded by ELAN-Programme of the University Hospital of Erlangen. The BCAC is funded by CR-UK (C1287/A10118 and C1287/A12014). Meetings of the BCAC have been funded by the European Union COST programme (BM0606). E.S. is supported by NIH Comprehensive Biomedical Research Centre, Guy’s & St Thomas’ NHS Foundation Trust in partnership with King’s College London, United Kingdom. I.T. is supported by the Oxford Biomedical Research Centre.
The HEBCS was supported by the Helsinki University Central Hospital Research Foundation, Academy of Finland (132473), the Finnish Cancer Society, The Nordic Cancer Union and the Sigrid Juselius Foundation. Financial support for KARBC was provided through the regional agreement on medical training and clinical research (ALF) between Stockholm County Council and Karolinska Institutet, the Swedish Cancer Society, The Gustav V Jubilee foundation and Bert von Kantzows foundation. The KBPC was financially supported by the special Government Funding (EVO) of Kuopio University Hospital grants, Cancer Fund of North Savo, the Finnish Cancer Organizations, the Academy of Finland and by the strategic funding of the University of Eastern Finland. kConFab is supported by grants from the National Breast Cancer Foundation, the NHMRC, the Queensland Cancer Fund, the Cancer Councils of New South Wales, Victoria, Tasmania and South Australia, and the Cancer Foundation of Western Australia. The KConFab Clinical Follow-Up Study was funded by the NHMRC, Cancer Australia and the National Breast Cancer Foundation. K.A.P. is a National Breast Cancer Foundation Practitioner Fellow. L.M.B.C. is supported by the ‘Stichting tegen Kanker’ (232-2008 and 196-2010). The MARIE study was supported by the Deutsche Krebshilfe e.V. (70-2892-BR II), the Hamburg Cancer Society, the German Cancer Research Center and the genotype work in part by the Federal Ministry of Education and Research (BMBF) Germany (01KHO10420). MCCS cohort recruitment was funded by VicHealth and Cancer Council Victoria. The MCCS was further supported by Australian NHMRC grants 209057, 251553 and 504711, and by infrastructure provided by Cancer Council Victoria. The MEC was supported by NIH grants CA66346, CA54281, CA09758 and CA128393. The NBCS was supported by grants from the Norwegian Research Council (155218/V40, 157340/S10 to A.L.R.D, FUGE-NER-181600/V11 to Y.N.K. and a Swiss Bridge Award to A.L.R.D.). The ORIGO study was supported by the Dutch Cancer Society (RUL 1997-1505) and the Biobanking and Biomolecular Resources Research Infrastructure (BBMRI-NL CPI6). pKARMA is a combination of the KARMA and LIBRO-1 studies. KARMA was supported by Märit and Hans Rausings Initiative Against Breast Cancer. KARMA and LIBRO-1 were supported the Cancer Risk Prediction Center (CRiP; www.criscenter.org), a Lineaus Centre (Contract ID 70867902) financed by the Swedish Research Council. The RBCS was funded by the Dutch Cancer Society (DDHK 2004-3124, DDHK 2009-4318). SASBC was supported by funding from the Agency for Science, Technology and Research of Singapore (A*STAR), the US National Institute of Health (NIH) and the Susan G. Komen Breast Cancer Foundation. K.C. was funded by the Swedish Cancer Society (5128-B07-01PAF). SEARCH is funded by a programme grant from Cancer Research UK (C490/ A10124) and supported by the UK National Institute for Health Research Biomedical Research Centre at the University of Cambridge. SKDKRFSZ is supported by the DKFZ. The TWBCS is supported by the Institute of Biomedical Sciences, Academia Sinica and the National Science Council, Taiwan. MYBRCA is funded by research grants from the Malaysian Ministry of Science, Technology and Innovation (MOSTI), Malaysian Ministry of Higher Education (UM/C/HIR/NOH/06) and Cancer Research Initiatives Foundation (CARIF). The HERPACC was supported by a Grant-in-Aid for Scientific Research on Priority Areas from the Ministry of Education, Science, Sports, Culture and Technology of Japan. MCCS was supported by a Grant-in-Aid for the Third Term Comprehensive 10-Year Strategy for Cancer Control from Ministry Health, Labour and Welfare of Japan, by a research grant from Takeda Science Foundation, by the Health and Labour Sciences Research Grants for Research on Applying Health Technology from Ministry Health, Labour and Welfare of Japan and by the National Cancer Center Research and Development Fund. The SEBCS was supported by the Korea Health 21 R&D Project (A030801), Ministry of Health and Welfare, Republic of Korea. POSH was supported by Funding Breast Cancer Campaign (NOV210P862) and Cancer Research UK (C1275/A896). The content of this manuscript does not necessarily reflect the views or policies of the National Cancer Institute or any of the collaborating centers in the Breast Cancer Family Registry (BCFR), nor does mention of trade names, commercial products, or organizations imply endorsement by the US Government or the BCFR. J.L.H. is a National Health and Medical Research Council (NHMRC) Senior Principal Research Fellow and M.C.S. is a NHMRC Senior Research Fellow. D.F.E. is a Principal Fellow of CR-UK.The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Author contributions
J.Li, K.C., L.L. M.L. and J.N.F. performed the statistical analysis and drafted the manuscript. D.E. conceived and coordinated the synthesis of the iCOGS array and led BCAC. P.H. coordinated COGS, M.K.S., P.D.P.P., K.M., J.D., M.K.B. and Q.W. provided data management support for BCAC. M.K.K., L.I.V.V., S.C. and E.R. coordinated ABCS. M.C.S., C.A., G.S.D. and J.L.H coordinated ARCBS, P.A.F., L.H., A.B.E. and M.W.B. coordinated BBCC. H.N., C.B., T.A.M. and K.A. coordinated HEBCS. A.L. and S.M. coordinated KARBC. A.M., V.-M., J.M.H. and V. Kataja coordinated KRCB. G.C.-T., K.A.P. and S.-A.M. coordinated kConFab. D.L., R.T., A.S. and H.W. coordinated LIBRC. J.C.-C., D.F.-J., P.S. and A.R. coordinated MARIE. G.G.G., L.B. and G.S. coordinated MCCS. C.A.H., B.E.H., F.S. and L.L.M. coordinated MEC. V. Kristensen, G.I.G.A., A.-L.B. and S.N. coordinated NBCS. R.W., K.P., A.J.-V. and M.G. coordinated OBCCS. F.D., J.A.K., G.G. and S.T. coordinated OFBCR. P.D., R.T. and C.S. coordinated ORIGO. P.H., K.H., K.C. and J.Li coordinated pKARMA. H.M., M.K., H.A. and A.v.d.O. coordinated RBCS: U.H., D.T., H.U.U. and T.R. coordinated SKDKRFSZ. J.Liu, I.N.F., C.C.K. and Y.L. coordinated SASBAC. A.D., M.S., R.L. coordinated SEARCH. C.-Y.S., C.-N.H., P.-E.W. and S.-T.C coordinated TBWCS. S.H.T., N.A.M.T., C.H.Y. and G.F.H coordinated MYBRCA. K.M., H. Ito, H. Iwata and K.T. coordinated HERPACC. D.K., J.-Y.C., S.K.P. and K.-Y.Y. coordinated SEBCS. S.R., T.M., J.W.T. and D.M.E. coordinated POSH. All authors provided critical review of the manuscript.

Additional information
**Supplementary Information** accompanies this paper at http://www.nature.com/naturecommunications

**Competition of financial interests:** The authors declare no competing interests.

**Reprints and permissions** information is available online at http://npg.nature.com/reprintsandpermissions/

**How to cite this article:** Li, J. et al. 2q36.3 is associated with prognosis for oestrogen receptor-negative breast cancer patients treated with chemotherapy. Nat. Commun. 5:4051 doi: 10.1038/ncomms5051 (2014).

This work is licensed under a Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License. The images or other third party material in this article are included in the article’s Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-nd/4.0/

Jia N. Foo,1 Saajad Rafiq,5 Marjanka K. Schmidt,6,7 Paul D.P. Pharoah,8,9 Kyriaki Michailidou,8 Joe Dennis,8 Manjeet K. Bolla,8 Qin Wang,8 Laura J. Van ‘t Veer,7 Sten Cornelissen,7 Emriel Rutgers,10 Melissa C. Southey11, Carmel Apicella12, Gillian S. Dite12, John L. Hopper12,13, Peter A. Fasching14,15, Lothar Haebelerle14, Arif B. Ekiç16, Matthias W. Beckmann14, Carl Blomqvist17, Taru A. Muranen,8 Kristiina Aittomäki,9 Annika Lindblom,20 Sara Margolin21, Arto Mannermäe22,23,24, Veli-Matti Kosma22,23,24, Jaana M. Hartikainen22,23,24, Vesa Kataja22,24,25, Georgia Chenевич-Trench26, Qin Wang8, Laura J. Van ‘t Veer7, Sten Cornelissen7, Emriel Rutgers10, Melissa C. Southey11, Carmel Apicella12, Gillian S. Dite12, John L. Hopper12,13, Peter A. Fasching14,15, Lothar Haebelerle14, Arif B. Ekiç16, Matthias W. Beckmann14, Carl Blomqvist17, Taru A. Muranen,8 Kristiina Aittomäki9, Annika Lindblom20, Sara Margolin21, Arto Mannermäe22,23,24, Veli-Matti Kosma22,23,24, Jaana M. Hartikainen22,23,24, Vesa Kataja22,24,25, Georgia Chenевич-Trench26, kConFab Investigators27, Kelly-Anne Phillips12,18, Sue-Anne McLachlan29, Diether Lambrechts30,31, Bernard Thienpont30,31, Ann Smeets32, Hans Wildiers32, Jenny Chang-Claude33, Dieter Flesch-Janys33,34,35, Petra Seibold33, Anja Rudolph33, Graham G. Giles32,36, Laura Baglietto32,36, Gianluca Severi12,36, Christopher A. Haiman37, Brian E. Henderson38, Irene L. Andrulis44,45, Julia A. Knight46,47, Gord Glendon44, Sandrine Tchatchou48, Peter Devilee49,50, Robert Tolenaar51, Caroline Seynaeve52, Maartje Hooning52, Mieke Krieger52, Antoinette Hollestelle52, Ans van den Ouweland53, Yi Li1,
