Carcinogenesis, 2015, Vol. 36, No. 2, 256–271

doi:10.1093/carcin/bgu326
Advance Access publication January 13, 2015
Original Manuscript

ORIGINAL MANUSCRIPT

Inherited variants in the inner centromere protein (INCENP) gene of the chromosomal passenger complex contribute to the susceptibility of ER-negative breast cancer

Maria Kabisch[1,†,‡], Justo Lorenzo Bermejo[2,†,‡], Thomas Dünnebier[1,‡], Shibo Ying[1], Kyriaki Michailidou[3], Manjeet K.Bolla[3], Qin Wang[3], Joe Dennis[3], Mitul Shah[4], Barbara J.Perkins[4], Kamila Czene[5], Hatef Darabi[5], Mikael Eriksson[5], Stig E.Bojesen[6,7], Børge G.Nordestgaard[6,7], Sune F.Nielsen[6,7], Henrik Flyger[8], Diether Lambrechts[9,10], Patrick Neven[11], Stephanie Peeters[11], Caroline Weltens[11], Fergus J.Couch[12], Janet E.Olson[13], Xianshu Wang[12], Kristen Purrington[14], Jenny Chang-Claude[15], Anja Rudolph[15], Petra Seibold[15], Dieter Flesch-Janys[16], Julian Peto[17], Isabel dos-Santos-Silva[17], Nichola Johnson[18], Olivia Fletcher[18], Heli Nevanlinna[19], Taru A.Muranen[19], Kristiina Aittomäki[20], Carl Blomqvist[21], Marjanka K.Schmidt[22], Annegien Broeks[22], Sten Cornelissen[22], Frans B.L.Hogervorst[22], Jingmei Li[23], Judith S.Brand[5], Keith Humphreys[5], Pascal Guénel[24,25], Thérèse Truong[24,25], Florence Menegaux[24,25], Marie Sanchez[24,25], Barbara Burwinkel[26,27], Frederik Marmé[26,28], Rongxi Yang[26,27], Peter Bugert[29], Anna González-Neira[30], Javier Benitez[30–32], M.Pilar Zamora[33], Jose I.Arias Perez[34], Angela Cox[35], Simon S.Cross[36], Malcolm W.R.Reed[35], Irene L.Andrulis[37,38], Julia A.Knight[39,40], Gord Glendon[41], Sandrine Tchatchou[37], Elinor J.Sawyer[42], Ian Tomlinson[43], Michael J.Kerin[44], Nicola Miller[44], kConFab Investigators[45], Australian Ovarian Cancer Study Group[45,46], Christopher A.Haiman[47], Fredrick Schumacher[47], Brian E.Henderson[47], Loic Le Marchand[48], Annika Lindblom[49], Sara Margolin[50], Maartje J.Hooning[51], Antoinette Hollestelle[51], Mieke Kriege[51], Linetta B.Koppert[52], John L.Hopper[53], Melissa C.Southey[54], Helen Tsimiklis[54], Carmel Apicella[53], Seth Slettedahl[13], Amanda E.Toland[55], Celine Vachon[13], Drakoulis Yannoukakos[56], Graham G.Giles[53,57], Roger L.Milne[53,57], Catriona McLean[58], Peter A.Fasching[59–61], Matthias Ruebner[59,60], Arif B.Ekici[60,62], Matthias W.Beckmann[59,60], Hermann Brenner[63,64], Aida K. Dieffenbach[63,64], Volker Arndt[63], Christa Stegmaier[65], Alan Ashworth[18], Nicholas Orr[18], Minouk J.Schoemaker[66], Anthony Swerdlow[66,67], Montserrat García-Closas[18,66], Jonine Figueroa[68], Stephen J.Chanock[68], Jolanta Lissowska[69], Mark S.Goldberg[70,71], France Labrèche[72], Martine Dumont[73], Robert Winqvist[74],

Received: August 20, 2014; **Revised:** December 5, 2014; **Accepted:** December 25, 2014

Katri Pylkäs[74], Arja Jukkola-Vuorinen[75], Mervi Grip[76], Hiltrud Brauch[77–79], Thomas Brüning[80], Yon-Dschun Ko[81], The GENICA Network[1,77–83], Paolo Radice[84], Paolo Peterlongo[85], Giulietta Scuvera[86], Stefano Fortuzzi[85,87], Natalia Bogdanova[88], Thilo Dörk[89], Arto Mannermaa[90–92], Vesa Kataja[93–95], Veli-Matti Kosma[90–92], Jaana M.Hartikainen[90–92], Peter Devilee[96], Robert A.E.M.Tollenaar[97], Caroline Seynaeve[51], Christi J.Van Asperen[98], Anna Jakubowska[99], Jan Lubinski[99], Katarzyna Jaworska-Bieniek[99], Katarzyna Durda[99], Wei Zheng[100], Martha J.Shrubsole[100], Qiuyin Cai[100], Diana Torres[1,101], Hoda Anton-Culver[102], Vessela Kristensen[103–105], François Bacot[106], Daniel C.Tessier[106], Daniel Vincent[106], Craig Luccarini[4], Caroline Baynes[4], Shahana Ahmed[4], Mel Maranian[4], Jacques Simard[73], Georgia Chenevix-Trench[107], Per Hall[5], Paul D.P.Pharoah[34], Alison M.Dunning[4], Douglas F.Easton[34] and Ute Hamann[1,*,‡]

[1]Molecular Genetics of Breast Cancer, German Cancer Research Center (DKFZ), 69120 Heidelberg, Germany, [2]Institute of Medical Biometry and Informatics, University Hospital Heidelberg, 69120 Heidelberg, Germany, [3]Department of Public Health and Primary Care and [4]Department of Oncology, Centre for Cancer Genetic Epidemiology, University of Cambridge, Cambridge, CB1 8RN, UK, [5]Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Stockholm, SE-17177, Sweden, [6]Copenhagen General Population Study, [7]Department of Clinical Biochemistry, and [8]Department of Breast Surgery, Herlev Hospital, Copenhagen University Hospital, 2730 Herlev, Denmark, [9]Vesalius Research Center, VIB, 3000 Leuven, Belgium, [10] Department of Oncology, Laboratory for Translational Genetics, University of Leuven, 3000 Leuven, Belgium, [11] Department of Oncology, KU Leuven (University of Leuven), Multidisciplinary Breast Center, University Hospitals Leuven, 3000 Leuven, Belgium, [12]Department of Laboratory Medicine and Pathology and [13]Department of Health Sciences Research, Mayo Clinic, Rochester, MN 55905, USA, [14]Karmanos Cancer Institute, Detroit, MI 48201, USA, [15]Division of Cancer Epidemiology, German Cancer Research Center (DKFZ), 69120 Heidelberg, Germany, [16]Department of Cancer Epidemiology/ Clinical Cancer Registry and Institute for Medical Biometrics and Epidemiology, University Clinic Hamburg-Eppendorf, 20246 Hamburg, Germany, [17]Department of Non-Communicable Disease Epidemiology, London School of Hygiene and Tropical Medicine, London, WC1E 7HT, UK, [18]Breakthrough Breast Cancer Research Centre, The Institute of Cancer Research, London, SW3 6JB, UK, [19]Department of Obstetrics and Gynecology, [20]Department of Clinical Genetics and [21]Department of Oncology, University of Helsinki and Helsinki University Central Hospital, FI-00029 Helsinki, Finland, [22]Netherlands Cancer Institute, Antoni van Leeuwenhoek Hospital, 1066 CX Amsterdam, The Netherlands, [23]Human Genetics Division, Genome Institute of Singapore, Singapore 138672, Singapore, [24]National Institute of Health and Medical Research, Center for Research in Epidemiology and Population Health, Environmental Epidemiology of Cancer, 94807 Villejuif, France, [25]University Paris-Sud, 94807 Villejuif, France, [26]Department of Obstetrics and Gynecology, University of Heidelberg, 69120 Heidelberg, Germany, [27]Molecular Epidemiology Group, German Cancer Research Center (DKFZ), 69120 Heidelberg, Germany, [28]National Center for Tumor Diseases, University of Heidelberg, 69120 Heidelberg, Germany, [29]Institute of Transfusion Medicine and Immunology, Medical Faculty Mannheim, Heidelberg University, 68167 Mannheim, Germany, [30]Human Genotyping-CEGEN Unit and [31]Human Genetics Group, Human Cancer Genetics Program, Spanish National Cancer Research Centre (CNIO), 28029 Madrid, Spain, [32]Centro de Investigación en Red de Enfermedades Raras, 46010 Valencia, Spain, [33]Servicio de Oncología Médica, Hospital Universitario La Paz, 28046 Madrid, Spain, [34]Servicio de Cirugía General y Especialidades, Hospital Monte Naranco, 33012 Oviedo, Spain, [35]Department of Oncology, University of Sheffield, Sheffield, S10 2RX, UK, [36]Academic Unit of Pathology, Department of Neuroscience, University of Sheffield, Sheffield S10 2HQ, UK, [37]Lunenfeld-Tanenbaum Research Institute of Mount Sinai Hospital, Toronto, ON M5G 1X5, Canada, [38]Department of Molecular Genetics, University of Toronto, Toronto, ON M5S 1A8, Canada, [39]Prosserman Centre for Health Research, Lunenfeld-Tanenbaum Research Institute of Mount Sinai Hospital, Toronto, ON M5G 1X5, Canada, [40]Division of Epidemiology, Dalla Lana School of Public Health, University of Toronto, Toronto, ON M5T 3M7, Canada, [41]ON Cancer Genetics Network, Lunenfeld-Tanenbaum Research Institute of Mount Sinai Hospital, Toronto, ON, M5G 1X5, Canada, [42]Research Oncology, Division of Cancer Studies, King's College London, Guy's Hospital, London SE1 9RT, UK, [43]Wellcome Trust Centre for Human Genetics and Oxford Biomedical Research Centre, University of Oxford, Oxford OX3 7BN, UK, [44]Clinical Science Institute, University Hospital Galway, Galway, Ireland, [45]Peter MacCallum Cancer Center, Melbourne, Victoria 3002, Australia, [46]QIMR Berghofer Medical Research Institute, Brisbane, QLD 4006, Australia, [47]Department of Preventive Medicine, Keck School of Medicine, University of Southern California, Los Angeles, CA 90033, USA, [48]Epidemiology Program, University of Hawaii Cancer Center, Honolulu, HI 96813, USA, [49]Department of Molecular Medicine and Surgery and [50]Department of Oncology - Pathology, Karolinska Institutet, Stockholm SE-17177, Sweden, [51]Department of Medical Oncology and [52]Department of Surgical Oncology, Erasmus MC Cancer Institute, 3008 AE Rotterdam, The Netherlands, [53]Centre for Epidemiology and Biostatistics, Melbourne School of Population and Global Health and [54]Department of Pathology, The University of Melbourne, Melbourne, Victoria 3010, Australia, [55]Department of Molecular Virology, Immunology and Medical Genetics, Comprehensive Cancer Center, The Ohio State University, Columbus, OH 43210, USA, [56]Molecular Diagnostics Laboratory, IRRP, National Centre for Scientific Research

'Demokritos', Aghia Paraskevi Attikis, 153 10 Athens, Greece, [57]Cancer Epidemiology Centre, Cancer Council Victoria, Melbourne, Victoria 3053, Australia, [58]Anatomical Pathology, The Alfred Hospital, Melbourne, Victoria 3004, Australia, [59] Department of Gynecology and Obstetrics, University Breast Center Franconia, University Hospital Erlangen, Friedrich-Alexander University Erlangen-Nuremberg, 91054 Erlangen, Germany, [60]Comprehensive Cancer Center Erlangen-EMN, 91054 Erlangen, Germany, [61]David Geffen School of Medicine, Department of Medicine, Division of Hematology and Oncology, University of California at Los Angeles, Los Angeles, CA 90095, USA, [62]Institute of Human Genetics, University Hospital Erlangen, Friedrich-Alexander University Erlangen-Nuremberg, 91054 Erlangen, Germany, [63]Division of Clinical Epidemiology and Aging Research, German Cancer Research Center (DKFZ), 69120 Heidelberg, Germany, [64]German Cancer Consortium (DKTK), 69120 Heidelberg, Germany, [65]Saarland Cancer Registry, 66119 Saarbrücken, Germany, [66]Division of Genetics and Epidemiology and [67]Division of Breast Cancer Research, Institute of Cancer Research, London, SM2 5NG, UK, [68]Division of Cancer Epidemiology and Genetics, National Cancer Institute, Rockville, MD 20850, USA, [69]Department of Cancer Epidemiology and Prevention, M. Sklodowska-Curie Memorial Cancer Center and Institute of Oncology, 02-781 Warsaw, Poland, [70]Department of Medicine, McGill University, Montreal, QC, H3G 2M1, Canada, [71]Division of Clinical Epidemiology, McGill University Health Centre, Royal Victoria Hospital, Montreal, QC H3G 2M1, Canada, [72]Département de santé environnementale et santé au travail, Département de médecine sociale et preventive, École de santé publique, Université de Montréal, Montreal, QC, H3T 1A8, Canada, [73]Centre Hospitalier Universitaire de Québec Research Center and Laval University, QC, G1V 4G2, Canada, [74] Department of Clinical Chemistry and Biocenter Oulu, Laboratory of Cancer Genetics and Tumor Biology, University of Oulu, NordLab Oulu/Oulu University Hospital, FI-90220 Oulu, Finland, [75]Department of Oncology and [76]Department of Surgery, Oulu University Hospital, University of Oulu, FI-90220 Oulu, Finland, [77]Dr Margarete Fischer-Bosch-Institute of Clinical Pharmacology, 70376 Stuttgart, Germany, [78]University of Tübingen, 72074 Tübingen, Germany, [79]German Cancer Consortium (DKTK) and German Cancer Research Center (DKFZ), 69120 Heidelberg, Germany, [80]Institute for Prevention and Occupational Medicine of the German Social Accident Insurance, Institute of the Ruhr University Bochum (IPA), 44789 Bochum, Germany, [81]Department of Internal Medicine, Evangelische Kliniken Bonn gGmbH, Johanniter Krankenhaus, 53113 Bonn, Germany, [82]Institute of Pathology, Medical Faculty of the University of Bonn, 53127 Bonn, Germany, [83]Institute of Occupational Medicine and Maritime Medicine, University Medical Center Hamburg-Eppendorf, 20246 Hamburg, Germany, [84]Unit of Molecular Bases of Genetic Risk and Genetic Testing, Department of Preventive and Predictive Medicine, Fondazione IRCCS Istituto Nazionale dei Tumori (INT), 20133 Milan, Italy, [85]IFOM, Fondazione Istituto FIRC di Oncologia Molecolare, 20139 Milan, Italy, [86]Unit of Medical Genetics, Department of Preventive and Predictive Medicine, Fondazione IRCCS Istituto Nazionale dei Tumori (INT), 20133 Milan, Italy, [87]Cogentech Cancer Genetic Test Laboratory, 20139 Milan, Italy, [88]Department of Radiation Oncology and [89]Department of Obstetrics and Gynaecology, Hannover Medical School, 30625 Hannover, Germany, [90]School of Medicine, Institute of Clinical Medicine, Pathology and Forensic Medicine and [91]Cancer Center of Eastern Finland, University of Eastern Finland, FI-70211 Kuopio, Finland, [92]Imaging Center, Department of Clinical Pathology, Kuopio University Hospital, FI-70211 Kuopio, Finland, [93]School of Medicine, Institute of Clinical Medicine, Oncology, University of Eastern Finland, FI-70211 Kuopio, Finland, [94]Cancer Center, Kuopio University Hospital, FI-70210 Kuopio, Finland, [95]Jyväskylä Central Hospital, FI-40620 Jyväskylä, Finland, [96]Department of Human Genetics and Department of Pathology, [97]Department of Surgical Oncology and [98]Department of Clinical Genetics, Leiden University Medical Center, 2333 ZC Leiden, The Netherlands, [99]Department of Genetics and Pathology, Pomeranian Medical University, 70–115 Szczecin, Poland, [100]Division of Epidemiology, Department of Medicine, Vanderbilt Epidemiology Center, Vanderbilt-Ingram Cancer Center, Vanderbilt University School of Medicine, Nashville, TN 37203, USA, [101]Institute of Human Genetics, Pontificia Universidad Javeriana, 11001000 Bogotá, Colombia, [102]Department of Epidemiology, University of California Irvine, Irvine, CA 92697, USA, [103]Department of Genetics, Institute for Cancer Research, Oslo University Hospital, Radiumhospitalet, N-0310 Oslo, Norway, [104]Institute of Clinical Medicine, Faculty of Medicine and [105]Department of Clinical Molecular Biology (EpiGen), University of Oslo, N-0450 Oslo, Norway, [106]McGill University and Génome Québec Innovation Centre, Montréal, QC H3A 0G1, Canada and [107]Department of Genetics, QIMR Berghofer Medical Research Institute, Brisbane, QLD 4006, Australia

[†]These authors contributed equally to this work.

‡These authors comprise the manuscript writing group.

[*]To whom correspondence should be addressed. German Cancer Research Center (DKFZ), Molecular Genetics of Breast Cancer (B072), Im Neuenheimer Feld 580, 69120 Heidelberg, Germany. Tel: +0049 6221 42/2344; Fax: +0049 6221 42/4721; Email: u.hamann@dkfz-heidelberg.de

Abstract

The chromosomal passenger complex (CPC) plays a pivotal role in the regulation of cell division. Therefore, inherited CPC variability could influence tumor development. The present candidate gene approach investigates the relationship between single nucleotide polymorphisms (SNPs) in genes encoding key CPC components and breast cancer risk. Fifteen SNPs in four CPC genes (*INCENP*, *AURKB*, *BIRC5* and *CDCA8*) were genotyped in 88 911 European women from 39 case-control studies of the Breast Cancer Association Consortium. Possible associations were investigated in fixed-effects meta-analyses. The synonymous SNP rs1675126 in exon 7 of *INCENP* was associated with overall breast cancer risk [per A allele odds ratio (OR) 0.95, 95% confidence interval (CI) 0.92–0.98, $P = 0.007$] and particularly with estrogen receptor (ER)-negative breast tumors (per A allele OR 0.89, 95% CI 0.83–0.95, $P = 0.0005$). SNPs not directly genotyped were imputed based on 1000 Genomes. The SNPs rs1047739 in the 3′ untranslated region

and rs144045115 downstream of *INCENP* showed the strongest association signals for overall (per T allele OR 1.03, 95% CI 1.00–1.06, *P* = 0.0009) and ER-negative breast cancer risk (per A allele OR 1.06, 95% CI 1.02–1.10, *P* = 0.0002). Two genotyped SNPs in *BIRC5* were associated with familial breast cancer risk (top SNP rs2071214: per G allele OR 1.12, 95% CI 1.04–1.21, *P* = 0.002). The data suggest that *INCENP* in the CPC pathway contributes to ER-negative breast cancer susceptibility in the European population. In spite of a modest contribution of CPC-inherited variants to the total burden of sporadic and familial breast cancer, their potential as novel targets for breast cancer treatment should be further investigated.

Abbreviations

AIC	Akaike's information criterion
AURKB	aurora kinase B
BCAC	Breast Cancer Association Consortium
BIRC5	baculoviral IAP repeat containing 5
CDCA8	cell division cycle associated 8
COGS	Collaborative Oncological Gene-environment Study
CPC	chromosomal passenger complex
ER	estrogen receptor
FRR	familial relative risk
HER2	human epidermal growth factor receptor 2
LD	linkage disequilibrium
MAF	minor allele frequencies
OS	overall survival
PAF	population attributable fraction
RFS	relapse-free survival
SNP	single nucleotide polymorphism
TCGA	The Cancer Genome Atlas.

Introduction

Breast cancer is the most commonly occurring epithelial malignancy among women, with an estimated 1.4 million new cases and >450 000 deaths worldwide (1). Familial aggregation and twin studies have shown the substantial contribution of inherited susceptibility to breast cancer (2,3). Many genetic loci have been identified that contribute to this familial risk (4), including genes with high-penetrance mutations, notably *BRCA1* and *BRCA2*, moderate penetrance genes including *ATM*, *BRIP1*, *CHEK2* and *PALB2* and common lower penetrance alleles, of which >80 have been identified so far. In total, these loci explain ~35% of the familial risk of breast cancer (5) leaving a large portion of the observed familial clustering of the disease unexplained (6).

The chromosomal passenger complex (CPC) is a key regulator of mitosis and is essential for maintenance of genomic stability through its control of multiple processes during both nuclear and cytoplasmic division (cytokinesis) (7,8).

The core CPC is composed of the three non-enzymatic subunits, the microtubule-binding inner centromere protein (INCENP), survivin (baculoviral IAP repeat containing 5, BIRC5) and borealin (cell division cycle associated 8, CDCA8), which regulate the activity, localization and stability of the CPC's catalytic subunit, aurora kinase B (AURKB) (9). INCENP is the platform on which the CPC assembles. The INCENP N-terminus forms a triple-helix bundle with the C-terminus of survivin and N-terminus of borealin (9) that is required for CPC localization to the centromere, anaphase spindle midzone and telophase midbody (9–12). AURKB binds to a conserved region (IN box) at the INCENP C-terminus (13). Strict localization of AURKB by CPC ensures that the kinase, which has >50 substrates (8), phosphorylates the correct targets at the proper steps in cell cycle progression.

Loss of CPC function results in lagging chromosomes during metaphase, leading to segregation errors, and in addition cleavage furrows fail to maintain ingression, resulting in cytokinesis failures (14–18). Moreover, lagging chromosomes can secondarily cause cytokinesis failures during telophase. Furthermore, analyses of point mutations in CPC proteins reveal independent roles of these proteins in the initiation of cytokinesis (19–21). Disturbed CPC function may also be caused by overexpression of CPC subunits and by deregulation of its regulatory kinases and phosphatases. Indeed, high expression levels of INCENP were observed in colorectal cancer cell lines (22), whereas high expression levels of survivin (23) and AURKB (24,25) have been found in various cancers including breast cancer and shown to be associated with poor prognosis (26,27).

Given the key role of the CPC in maintaining genomic stability and the facts that chromosome segregation error (28,29) and overexpression of CPC components are frequently seen in human cancers, we hypothesize that genetic variants in the core CPC genes *INCENP*, *AURKB*, *BIRC5* and *CDCA8* affect breast cancer susceptibility. Thus, the primary aim of this investigation was to assess possible associations between selected tag single nucleotide polymorphisms (SNPs) and potentially functional SNPs in four CPC genes and breast cancer risk. Subsequently, *in silico* analyses of SNP function and gene expression were carried out to provide supportive evidence of the identified risk variants. The secondary aim was to explore genetic associations with the survival of breast cancer patients.

Materials and methods

Study participants

Study subjects were 88 911 women of European ancestry from 39 case-control studies participating in the Breast Cancer Association Consortium (BCAC). All BCAC studies had local ethical approvals and all included individuals gave informed consent (5). Seventeen SNPs were selected for genotyping including 12 tag SNPs for *INCENP*, three potentially functional SNPs in *AURKB*, *BIRC5* and *CDCA8*, one SNP in *AURKB* previously reported to be associated with familial breast cancer risk (30) and one SNP in *BIRC5* previously reported to be correlated with survivin expression (31). SNP genotyping in the BCAC samples was conducted using a custom Illumina Infinium array (iCOGS) in four centers, as part of a multiconsortia collaboration [the Collaborative Oncological Gene-environment Study (COGS)] (5). Genotypes were called using Illumina's proprietary GenCall algorithm. Quality control included checks on call rate, heterozygosity and Hardy–Weinberg equilibrium. Details on BCAC studies and the SNP selection approach can be found in Supplementary Materials and Methods and in Supplementary Tables S1 and S2, available at *Carcinogenesis* Online.

Statistical analyses

Single SNP association analysis

The BCAC provided genotype data along with the first 7 genetic principal components to allow adjustment for population stratification (the first 6 components based on ~37 000 uncorrelated polymorphisms plus a seventh specifically derived for the Leuven Multidisciplinary Breast Centre study). Available phenotype data included disease status, hormone receptor status, family history and survival information. The association between genotypes and overall breast cancer risk was investigated in an individual-based fixed-effects model meta-analysis comprising 88 911 study subjects. Per allele odds ratios (ORs) and corresponding 95% confidence intervals (CIs) were estimated by logistic regression in a model that incorporated study and the first 7 principal genetic components as fixed effects. Study heterogeneity

was assessed by the I^2 index. Forest plots were generated and the nearest neighbor method based on the medians of the seven first genetic principal components was used to cluster studies according to genetic similarity.

To assess the familial breast cancer risk, cases with a family history of breast cancer in a first-degree relative were compared with all controls using an additive logistic regression model adjusted for study and the first 7 principal genetic components.

A case-only analysis was carried out to explore whether SNPs in CPC genes are associated with the hormone receptor status of the tumor [estrogen receptor (ER)-positive/negative, progesterone receptor (PR)-positive/negative and human epidermal growth factor receptor 2 (HER2)-positive/negative].

Survival information was available for only ~65% of all cases. The relationship between genotype and overall survival (OS), breast-cancer-specific survival and relapse-free survival (RFS) was investigated in cases, which did not represent with distant metastases at diagnosis. OS was defined as the time between breast cancer diagnosis and death or last follow-up, whichever occurred first. For breast-cancer-specific survival, only deaths from breast cancer according to International Classification of Diseases, 10th Revision code counted as events, whereas deaths from any other cause were censored. RFS was defined as the time between breast cancer diagnosis and locoregional relapse or relapse of distant metastasis after a period of remission, whichever occurred first. Survival times were censored after 15 years (for OS and breast-cancer-specific survival) and 10 years (for RFS). If cases were diagnosed before study entry, survival times were left truncated. Survival analyses were performed by Cox regression models incorporating study and the seven genetic principal components as fixed effects. Per allele hazard ratios and corresponding 95% CIs were reported. In addition Kaplan–Meier estimates of survival were plotted stratified by genotype, and genotype-specific estimated 10 year (for OS and BCCS) and 5 year (for RFS) survival rates were reported.

If an association with the hormone receptor status of the tumor was detected, the subtype-specific disease risk and survival was investigated as well.

Haplotype analysis

Pairwise linkage disequilibrium (LD) between *INCENP* SNPs was measured by r^2 and a LD heat map was generated. Based on different combinations of SNPs and taking the LD block structure into account, we inferred haplotypes using the expectation-maximization algorithm. Haplotype frequencies were calculated. Subsequently, the association between most frequent haplotypes and overall breast cancer risk was analyzed by a logistic regression model adjusted for study and the seven principal components. The model fit was evaluated by Akaike's information criterion (AIC) in order to identify the optimal SNP combination, where the smallest AIC value represents the best model. Haplotype-specific ORs and corresponding 95% CIs were also estimated.

Interaction and pathway analysis

In order to assess the interaction effects of SNPs in different CPC genes on the risk of breast cancer, we carried out a SNP–SNP interaction analysis. The multiplicative interaction index and the interaction contrast ratio were calculated, and deviation from multiplicativity and additivity was tested based on Wilcoxon signed-rank tests. The 95% CIs were computed by bootstrapping with 10 000 simulations. To investigate whether SNPs in genes of the CPC pathway are jointly associated with overall breast cancer risk, *P*-values from single SNP analyses were summarized into one combined *P*-value using Fisher's method for independent tests.

Imputation of genotypes

Multiple imputation of genotypes was performed based on all genotyped *INCENP* SNPs, to detect associations with not directly genotyped but potentially causal SNPs. The European subpopulations from the 1000 Genomes Project phase 1 were used as reference panel (32). Genotypes of only SNPs were imputed in a region centered on the *INCENP* gene. The extent of this region was identified by visual inspection of recombination rates and, spanned at least ±150 kb starting from the first and last reference SNP. A logistic regression model adjusted for study and the seven genetic principal components was applied in subsequent association analyses for imputed genotypes summarized by minus the logarithm of the *P*-value. A gene map of the investigated region was created together with a LD map relying on pairwise r^2 values for imputed and reference SNPs.

Population attributable fraction and familial risk

The population attributable fraction (PAF) was calculated for the SNP showing the strongest association with overall breast cancer risk in order to quantify the proportion of sporadic cases related to the risk variant. Also the familial relative risk (FRR), which reflects the attributable proportion of familial cases, was estimated. The calculation of PAF and FRR relied on the estimated ORs and minor allele frequencies (MAF), together with an assumed prevalence of breast cancer in the general population equal to 7.8% until the age of 74 (33,34). Subsequently, the obtained PAF and FRR of *INCENP* were compared with the PAFs and FRRs of previously identified breast cancer susceptibility variants. Considered susceptibility variants included high-penetrance mutations in *BRCA1* and *BRCA2* (6), moderate penetrance variants in *ATM, BRIP1*, *CHEK2* and *PALB2* (6,35) as well as 80 low-penetrance variants throughout the genome (4).

Expression quantitative trait loci analysis

We examined whether identified risk variants influence gene expression. Information from HapMap, NCBI's Gene Expression Omnibus and The Cancer Genome Atlas (TCGA) was exploited (36,37). For 60 unrelated Utah residents with northern and western European ancestry from the CEPH collection (CEU population), genotype data were available from the International HapMap project. Expression data derived from Epstein-Barr virus-transformed lymphoblastoid cell lines of the same individuals have been made public through Gene Expression Omnibus. The breast cancer study (BRCA) from TCGA provides germline DNA genotypes as well as expression data for tumor and matched normal breast tissue samples. All eQTL analyses involved a two-sided Kruskal–Wallis test. Differences in expression levels between normal and tumor breast tissue samples were analyzed with a two-sided Wilcoxon–Mann–Whitney test.

Functional SNP analysis

In order to explore the functional significance of identified risk variants, the R package *FunciSNP* was used to examine *in silico* annotations with chromatin features available in ENCODE (38,39). The list of examined functional characteristics included five built-in biofeatures [CTCF binding sites, DNaseI hypersensitivity sites, formaldehyde-assisted isolation of regulatory elements signals and known promoter regions] across several cell lines as well as 57 biofeatures (DNaseI hypersensitivity sites, formaldehyde-assisted isolation of regulatory elements signals, transcription factor binding sites, methylation sites, Chromatin State Segmentation by HMM (ChromHMM) and histone modifications by Chip-seq) specifically downloaded for HMEC normal mammary epithelial cells and breast cancer cell lines MCF7 and T47D. Functional SNP analyses were carried out for variants in a 300 kb window centered on the SNP with the strongest association.

Results

Eight tag SNPs (including three surrogates) out of 12 originally selected tag SNPs for *INCENP* were genotyped by the BCAC. Additional genotype data were provided for one upstream and one downstream SNP of *INCENP*. Five SNPs (including one surrogate) out of five originally selected SNPs for *AURKB*, *BIRC5* and *CDCA8* were genotyped. A description of all 15 SNPs genotyped for *INCENP*, *AURKB*, *BIRC5* and *CDCA8* can be found in Supplementary Table S2, available at *Carcinogenesis* Online.

Associations of SNPs in CPC genes with breast cancer risk and survival

A total of 88 911 European women (46 450 cases and 42 461 controls) from 39 BCAC studies were included in association analyses. Reported probability values were not adjusted for multiplicity, they should be interpreted considering that 4 genes and 15 partially linked variants were simultaneously investigated.

Five *INCENP* SNPs (top SNP rs1675126) were associated with a decreased overall breast cancer risk. Women with variant rs1675126 showed the largest breast cancer risk reduction (per minor A allele OR 0.95, 95% CI 0.92–0.98, $P = 0.007$). The two SNPs rs4963459 and rs4963471, located, respectively, upstream and downstream of *INCENP* were associated with overall breast cancer risk as well (rs4963459: per minor A allele OR 1.02, 95% CI 1.00–1.04, $P = 0.021$; rs4963471: per minor G allele OR 1.03, 95% CI 1.01–1.05, $P = 0.003$). Association results for overall breast cancer risk of all SNPs in *INCENP* are shown in Table 1. There was also a weak association of rs2306625 in *CDCA8* with overall breast cancer risk (per minor A allele OR 0.97, 95% CI 0.95–0.99, $P = 0.040$; Table 2).

Figure 1A and B represents the clustering of studies by genetic similarity based on the first genetic principal components. The forest plot on the association with overall breast cancer risk for the top SNP rs1675126 is shown in Figure 1B. The reflection of the geographical study distribution was evident. However, a regional clustering of OR estimates was not obvious. Study heterogeneity was not apparent ($I^2 = 0\%$).

The familial breast cancer risk was increased for women with variants rs2071214 and rs3764384 in *BIRC5* (rs2071214: per minor G allele OR 1.12, 95% CI 1.04–1.21, $P = 0.002$; rs3764384: per minor A allele OR 1.04, 95% CI 1.00–1.08, $P = 0.043$; Supplementary Table S3, available at *Carcinogenesis* Online).

Case-only analysis revealed that four *INCENP* SNPs, which were associated with overall breast cancer risk, showed differential association according to ER (top SNP rs1675126: per minor A allele OR 1.09, 95% CI 1.01–1.16, $P = 0.012$), but not to PR or HER2 tumor status (Table 3). Subsequent analysis of subtype-specific disease risk revealed that five *INCENP* SNPs (top SNP rs1675126) showed stronger associations with risk of ER-negative breast tumors than with overall breast cancer risk. Women with variant rs1675126 showed the largest reduction in risk of developing ER-negative tumors (per minor A allele OR 0.89, 95% CI 0.83–0.95, $P = 0.0005$). This observed association was the strongest among all breast cancer risk analyses and remained statistically significant after correction for multiple testing. The Bonferroni-adjusted *P*-value was $P = 0.04$ (0.0005*75 – considering 15 tests on overall breast cancer risk, 15 tests on familial breast cancer and 15 tests for each of the three hormone receptors). The large sample size of the present association study provided sufficient statistical power to detect small differences between cases and controls in allele frequencies. Table 1 displays association results for all *INCENP* SNPs stratified by ER status. The forest plot on the association with ER-negative breast cancer risk for the top SNP rs1675126 is shown in Figure 1C. The *CDCA8* SNP rs2306625 was associated with HER2 (per minor A OR 0.95, 95% CI 0.91–0.99, $P = 0.033$), but not with ER or PR tumor status (Table 3). No association of rs2306625 with risk of HER2-positive or negative breast tumors was observed (Supplementary Table S4, available at *Carcinogenesis* Online).

No survival association—either with overall, breast cancer specific or relapse-free survival—was observed for the SNPs in *INCENP*, *AURKB* and *BIRC5*. The investigated SNP in *CDCA8* was associated with relapse-free survival. Patients with variant rs2306625 showed an increased risk of relapse (per minor A allele hazard ratio 1.17, 95% CI 1.05–1.31, $P = 0.004$, 89% of the survival times were censored). The 5 year RFS rate was 0.90 (95% CI 0.89–0.91) for patients homozygous for the common allele, 0.89 (95% CI 0.88–0.91) for heterozygotes and 0.88 (95% CI 0.83–0.91) for patients homozygous for the minor allele. The association of rs2306625 with relapse-free survival was stronger when cases with a HER2-positive tumor were compared with all controls (per minor A allele hazard ratio 1.56, 95% CI 1.12–2.17, $P = 0.008$, 84% of the survival times were censored). The results from survival analysis of all SNPs in CPC genes are displayed in Supplementary Tables S5–S8, available at *Carcinogenesis* Online. The relapse-free survival stratified by *CDCA8* rs2306625 genotype is shown in Supplementary Figure S1, available at *Carcinogenesis* Online.

Associations of INCENP haplotypes with overall breast cancer risk

The five *INCENP* SNPs that were singly associated with a decreased overall breast cancer risk were in high LD ($r^2 > 0.8$) and located in two LD blocks comprising a region of ~12 kb (Supplementary Figure S2, available at *Carcinogenesis* Online). Haplotypes were estimated for these SNPs in order to assess their synergistic effect on breast cancer risk. First, the SNPs were ordered according to their *P*-values obtained from overall breast cancer risk analysis. Haplotypes were then inferred for (i) the top two SNPs, (ii) the top three SNPs and (iii) all five SNPs. Among all assessed SNP combinations, the model fit was best for the combination of the top three SNPs (AIC = 238222.1), but did not improve the model fit for rs1675126 alone (AIC = 119150.0). The haplotype frequencies and haplotype-specific estimates for all assessed SNP combinations are displayed in Supplementary Table S9, available at *Carcinogenesis* Online.

Results from interaction and pathway analyses are presented in the Supplementary Results, available at *Carcinogenesis* Online.

Genotype imputation of untyped SNPs in the INCENP region

Since several genotyped *INCENP* SNPs were associated with overall and ER-negative breast cancer risk, association mapping was refined by imputing additional variants. The reference panel for imputation comprised 379 individuals of the European subpopulations CEU, TSI (Toscani in Italia), GBR (British from England and Scotland), FIN (Finnish from Finland) and IBS (Iberian populations from Spain) from the 1000 Genomes Project. After visual inspection of recombination rates, an ~465 kb region (from 61 735 132 to 62 201 016 of chromosome 11, NCBI build 37) centered on *INCENP* and comprising 6282 SNPs was selected for genotype imputation. A total of 5078 SNPs fulfilled genotype heterozygosity and were imputed with high accuracy (99.1% median average certainty of best-guess genotypes). Subsequent association analysis revealed that the strongest signal for the association with overall breast cancer risk was obtained for rs1047739 (per minor T allele OR 1.03, 95% CI 1.00–1.06, $P = 0.0009$; Figure 2A). A marginal differential association according to ER tumor status was detected for rs1047739 (per minor T allele OR 1.04, 95% CI 1.00–1.08, $P = 0.005$), but rs144045115 showed the strongest association signal for the association with ER-negative breast cancer risk (per minor A allele OR 1.06, 95% CI 1.02–1.10, $P = 0.0002$; Figure 2B). The two variants are located in the 3′ untranslated region and downstream of *INCENP*. A gene map, recombination rates and LD in the investigated *INCENP* region are represented in Figure 2C, D and E, respectively.

PAF and FRR related to the top INCENP SNP associated with overall breast cancer risk

PAFs and FRRs for *INCENP* SNP rs1047739 compared with previously identified susceptibility variants are displayed in Figure 3. Rs1047739 showed a per allele OR of 1.03 for the association with

Table 1. Association between SNPs in *INCENP* and breast cancer risk

			Overall breast cancer risk			ER-negative breast cancer risk			ER-positive breast cancer risk		
SNP	Genotype	Controls, *N* (%)	Cases, *N* (%)	Per allele OR[a] (95% CI)	*P*-value[b]	Cases, *N* (%)	Per allele OR[a] (95% CI)	*P*-value[b]	Cases, *N* (%)	Per allele OR[a] (95% CI)	*P*-value[b]
rs4963459	GG	13 479 (31.8)	14 552 (31.3)	1.02 (1.00–1.04)	**0.021**	2346 (31.7)	1.02 (0.98–1.06)	0.175	8557 (31.6)	1.01 (0.99–1.03)	0.242
	GA	20 934 (49.3)	22 926 (49.4)			3639 (49.1)			13 322 (49.2)		
	AA	8043 (18.9)	8957 (19.3)			1426 (19.2)			5185 (19.2)		
rs17707648	GG	38 500 (90.7)	41 992 (90.4)	1.00 (0.96–1.05)	0.752	6727 (90.8)	0.98 (0.90–1.07)	0.674	24 464 (90.4)	1.01 (0.96–1.06)	0.566
	GA	3861 (9.1)	4340 (9.3)			666 (9.0)			2541 (9.4)		
	AA	100 (0.2)	118 (0.3)			20 (0.3)			69 (0.3)		
rs1628349	GG	37 684 (88.8)	41 313 (88.9)	0.95 (0.91–0.99)	**0.037**	6646 (89.7)	0.88 (0.81–0.96)	**0.004**	23 951 (88.5)	0.98 (0.93–1.02)	0.401
	GA	4609 (10.9)	4960 (10.7)			743 (10.0)			3008 (11.1)		
	AA	167 (0.4)	176 (0.4)			24 (0.3)			114 (0.4)		
rs1792949	CC	18 911 (44.9)	20 932 (45.4)	0.97 (0.95–0.99)	**0.038**	3401 (46.3)	0.94 (0.90–0.97)	**0.002**	12 024 (44.7)	0.99 (0.97–1.02)	0.799
	CA	18 550 (44.0)	20 081 (43.6)			3190 (43.4)			11 756 (43.7)		
	AA	4699 (11.2)	5100 (11.1)			761 (10.4)			3097 (11.5)		
rs1675063	AA	19 396 (45.7)	21 512 (46.3)	0.97 (0.95–0.99)	**0.029**	3473 (46.9)	0.94 (0.90–0.98)	**0.003**	12 390 (45.8)	0.99 (0.97–1.01)	0.574
	AG	18 594 (43.8)	20 074 (43.2)			3208 (43.3)			11 769 (43.5)		
	GG	4464 (10.5)	4851 (10.5)			729 (9.8)			2908 (10.7)		
rs1675126	GG	34 763 (81.9)	38 102 (82.0)	0.95 (0.92–0.98)	**0.007**	6161 (83.1)	0.89 (0.83–0.95)	**5×10^{-4}**	21 979 (81.2)	0.97 (0.94–1.01)	0.210
	GA	7213 (17.0)	7856 (16.9)			1184 (16.0)			4772 (17.6)		
	AA	482 (1.1)	488 (1.1)			68 (0.9)			319 (1.2)		
rs7129085	AA	16 264 (38.3)	18 062 (38.9)	0.97 (0.95–0.99)	**0.017**	2952 (39.8)	0.93 (0.90–0.97)	**9×10^{-4}**	10 411 (38.5)	0.98 (0.96–1.01)	0.340
	AC	19 943 (47.0)	21 601 (46.5)			3440 (46.4)			12 588 (64.5)		
	CC	6240 (14.7)	6771 (14.6)			1017 (13.7)			4066 (15.0)		
rs3781969	AA	26 025 (61.3)	28 660 (61.7)	0.99 (0.97–1.01)	0.641	4589 (61.9)	0.97 (0.93–1.02)	0.313	16 723 (61.8)	1.00 (0.97–1.02)	0.994
	AC	14 401 (33.9)	15 611 (33.6)			2499 (33.7)			9076 (33.5)		
	CC	2026 (4.8)	2168 (4.7)			324 (4.4)			1266 (4.7)		
rs11230934	AA	22 687 (53.4)	24 937 (53.7)	0.99 (0.97–1.01)	0.514	4031 (54.4)	0.96 (0.92–1.00)	0.074	14 572 (53.8)	1.00 (0.97–1.02)	0.987
	AC	16 638 (39.2)	18 148 (39.1)			2878 (38.8)			10 498 (38.8)		
	CC	3129 (7.4)	3359 (7.2)			503 (6.8)			2000 (7.4)		
rs4963471	AA	23 064 (54.3)	24 865 (53.5)	1.03 (1.01–1.05)	**0.003**	3949 53.3)	1.05 (1.01–1.09)	**0.012**	14 609 (54.0)	1.02 (1.00–1.05)	**0.028**
	AG	16 365 (38.5)	18 265 (39.3)			2915 (39.3)			10 519 (38.9)		
	GG	3032 (7.1)	3314 (7.1)			547 (7.4)			1942 (7.2)		

Probability values <5% are shown in bold.

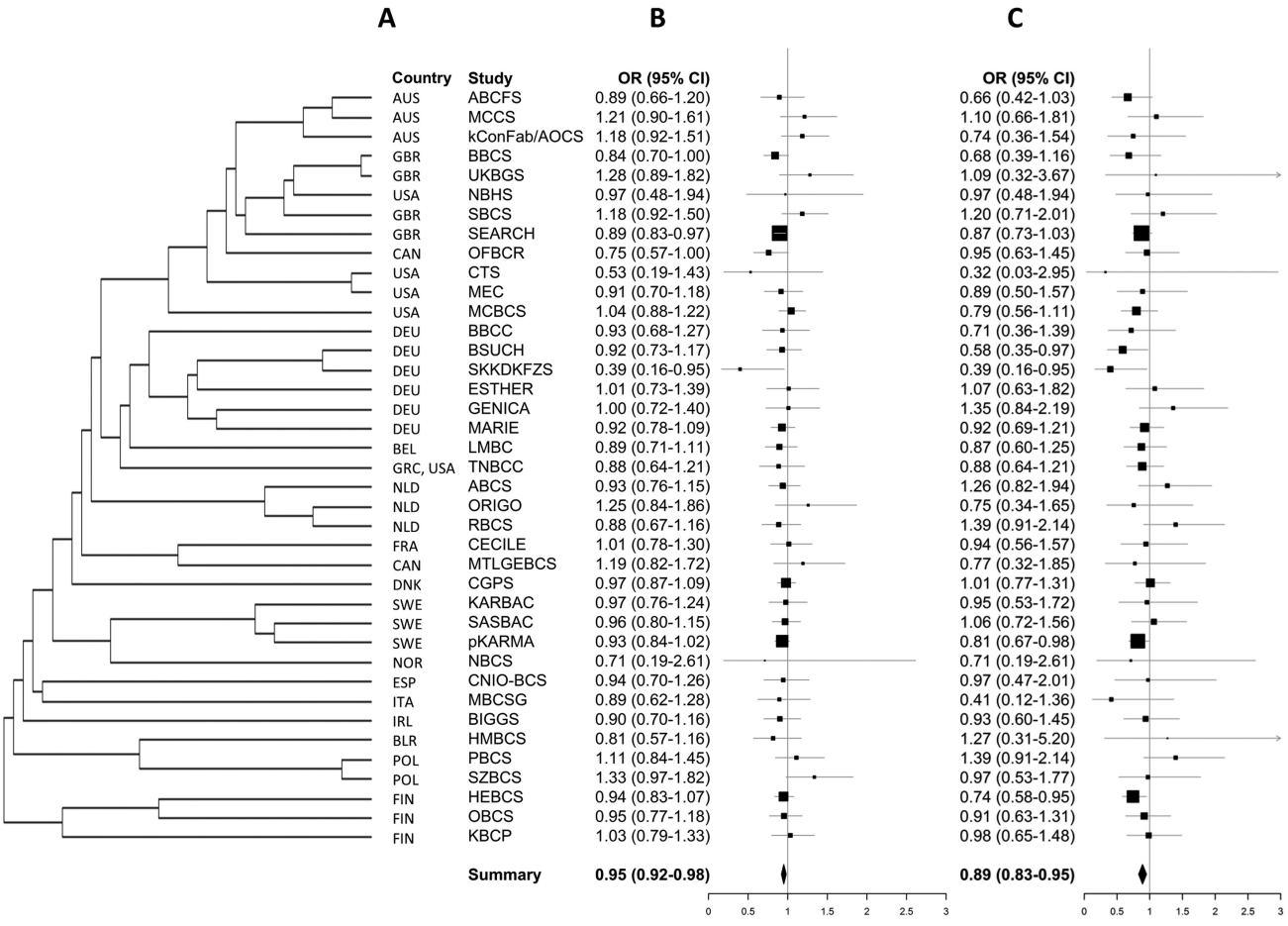
[a]OR adjusted for a fixed study effect and the first 7 principal components.

[b]Probability value based on logistic regression and an additive model.

Table 2. Association between SNPs in *AURKB*, *BIRC5* and *CDCA8* and breast cancer risk

				Overall breast cancer risk		
Gene	SNP	Genotype	Controls, *N* (%)	Cases, *N* (%)	Per allele OR[a] (95% CI)	*P*-value[b]
AURKB	rs1059476	GG	33 953 (80.0)	37 051 (79.8)	0.94 (0.97–1.00)	0.154
		AG	7966 (18.7)	8818 (19.0)		
		AA	523 (1.2)	568 (1.2)		
	rs2241909	AA	18 882 (44.5)	20 753 (44.7)	0.98 (0.96–1.00)	0.196
		AG	18 844 (44.4)	20 508 (44.2)		
		GG	4687 (11.1)	5135 (11.1)		
BIRC5	rs2071214	AA	37 965 (89.4)	41 512 (89.4)	1.02 (0.98–1.06)	0.293
		AG	4356 (10.3)	4792 (10.3)		
		GG	139 (0.3)	145 (0.3)		
	rs3764384	GG	19 315 (45.5)	20 958 (45.1)	1.00 (0.98–1.02)	0.602
		AG	18 690 (44.0)	20 492 (44.1)		
		AA	4449 (10.5)	4991 (10.8)		
CDCA8	rs2306625	GG	28 220 (66.5)	31 003 (66.8)	0.97 (0.95–0.99)	**0.040**
		AG	12 698 (29.9)	13 821 (29.8)		
		AA	1526 (3.6)	1590 (3.4)		

Probability values <5% are shown in bold.
aOR adjusted for a fixed study effect and the first 7 principal components.
bProbability value based on logistic regression and an additive model.



AUS, Australia; BEL, Belgium; BLR, Belarus; CAN, Canada; DEU, Germany; DNK, Denmark; ESP, Spain; FRA, France; FIN, Finland; GBR, United Kingdom; GRC, Greece; IRL, Ireland; ITA, Italy; NLD, Netherlands; NOR, Norway; POL, Poland; SWE, Sweden; USA, United States

Figure 1. (**A**) Clustering of studies based on the first genetic principal components. The distance between merged studies reflects their genetic similarity. (**B**) Forest plot for the association between rs1675126 and overall breast cancer risk. (**C**) Forest plot for the association between rs1675126 and ER-negative breast cancer risk.

Table 3. Association between SNPs in CPC genes and hormone receptor status

Gene	SNP	Genotype	Cases, *N* (%)	ER status		PR status		HER2 status	
				Per allele OR[a] (95% CI)	*P*-value[b]	Per allele OR[a] (95% CI)	*P*-value[b]	Per allele OR[a] (95% CI)	*P*-value[b]
INCENP	rs4963459	GG	14552 (31.3)	0.99 (0.95–1.03)	0.656	1.00 (0.96–1.04)	0.831	0.95 (0.89–1.01)	0.127
		GA	22926 (49.4)						
		AA	8957 (19.3)						
	rs17707648	GG	41992 (90.4)	1.05 (0.96–1.15)	0.234	1.02 (0.94–1.11)	0.558	0.98 (0.85–1.12)	0.797
		GA	4340 (9.3)						
		AA	118 (0.3)						
	rs1628349	GG	41313 (88.9)	1.09 (1.00–1.18)	**0.046**	1.04 (0.97–1.13)	0.215	1.07 (0.94–1.21)	0.281
		GA	4960 (10.7)						
		AA	176 (0.4)						
	rs1792949	CC	20932 (45.4)	1.04 (1.00–1.09)	**0.025**	1.03 (0.99–1.07)	0.111	1.01 (0.95–1.08)	0.607
		CA	20081 (43.6)						
		AA	5100 (11.1)						
	rs1675063	AA	21512 (46.3)	1.03 (0.99–1.08)	0.074	1.02 (0.98–1.06)	0.229	1.03 (0.96–1.09)	0.362
		AG	20074 (43.2)						
		GG	4851 (10.5)						
	rs1675126	GG	38102 (82.0)	1.09 (1.01–1.16)	**0.012**	1.05 (0.99–1.12)	0.064	1.03 (0.93–1.14)	0.569
		GA	7856 (16.9)						
		AA	488 (1.1)						
	rs7129085	AA	18062 (38.9)	1.04 (1.00–1.08)	**0.028**	1.02 (0.98–1.06)	0.243	1.01 (0.95–1.08)	0.594
		AC	21601 (46.5)						
		CC	6771 (14.6)						
	rs3781969	AA	28660 (61.7)	1.00 (0.96–1.05)	0.706	1.00 (0.96–1.05)	0.706	1.00 (0.93–1.08)	0.827
		AC	15611 (33.6)						
		CC	2168 (4.7)						
	rs11230934	AA	24937 (53.7)	1.02 (0.98–1.07)	0.218	1.01 (0.97–1.05)	0.469	1.06 (0.81–1.38)	0.654
		AC	18148 (39.1)						
		CC	3359 (7.2)						
	rs4963471	AA	24865 (53.5)	0.97 (0.93–1.02)	0.273	0.98 (0.94–1.02)	0.528	0.99 (0.92–1.06)	0.775
		AG	18265 (39.3)						
		GG	3314 (7.1)						
AURKB	rs1059476	GG	34056 (80.0)	1.04 (1.00–1.09)	0.052	1.03 (0.99–1.08)	0.071	0.99 (0.94–1.05)	0.873
		AG	7999 (18.8)						
		AA	526 (1.2)						
	rs2241909	AA	18934 (44.5)	1.00 (0.97–1.03)	0.721	0.98 (0.96–1.01)	0.423	0.98 (0.94–1.01)	0.306
		AG	18907 (44.4)						
		GG	4711 (11.1)						
BIRC5	rs2071214	AA	38090 (89.4)	1.03 (0.97–1.09)	0.316	1.04 (0.98–1.10)	0.115	0.98 (0.91–1.05)	0.593
		AG	4370 (10.3)						
		GG	139 (0.3)						
	rs3764384	GG	19384 (45.5)	0.97 (0.95–1.00)	0.110	0.97 (0.94–1.00)	0.067	0.98 (0.95–1.01)	0.374
		AG	18748 (44.0)						
		AA	4461 (10.5)						
CDCA8	rs2306625	GG	28308 (66.5)	1.00 (0.96–1.03)	0.845	0.98 (0.95–1.02)	0.508	0.95 (0.91–0.99)	**0.033**
		AG	12740 (29.9)						
		AA	1535 (3.6)						

Probability values <5% are shown in bold.

[a]OR adjusted for a fixed study effect and the first 7 principal components.

bProbability value based on logistic regression and an additive model.

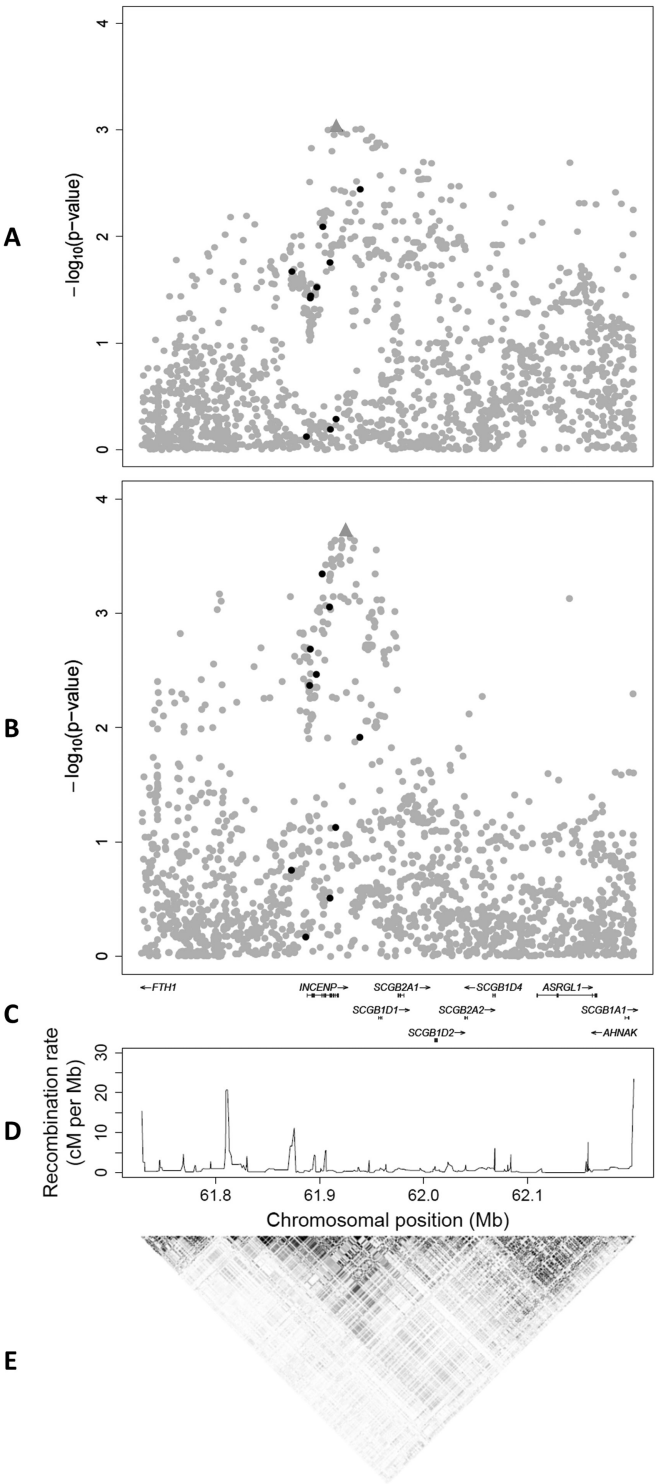


Figure 2. (**A**) Association between overall breast cancer risk and genotyped (black) and imputed (gray) SNPs in the greater *INCENP* region. (**B**) Association between ER-negative breast cancer risk and genotyped (black) and imputed (gray) SNPs in the greater *INCENP* region. Both plots show the $-\log_{10}$ *P*-values based on logistic regression adjusted for study and seven principal components. Only imputed SNPs with MAF > 0.01 are depicted. The imputed SNPs with the smallest *P*-value (rs1047739 and rs144045115) are shown as gray triangles. (**C**) Gene map including all genes in the investigated region. (**D**) Recombination rates in the investigated region. Chromosomal positions refer to NCBI build 37. (**E**) LD heatmap based on genotype data retrieved from the European subpopulations from HapMap phase 3 showing pairwise r^2 values [from 0 (white) to 1 (black)].

overall breast cancer risk and a MAF of 0.24. Assuming a cumulative risk of breast cancer in the European Union of 7.8% until the age of 74, rs1047739 results in a PAF of 1.4% and a FRR of 1.0. In comparison with other susceptibility variants, the *INCENP* SNP rs1047739 contributed to a higher PAF than any rare variant in *BRCA1*, *BRCA2*, *ATM*, *BRIP1*, *CHEK2* or *PALB2*. Most recently identified common susceptibility variants showed larger PAFs and FRRs than rs1047739, where FGFR2 rs2981579 showed the second highest PAF

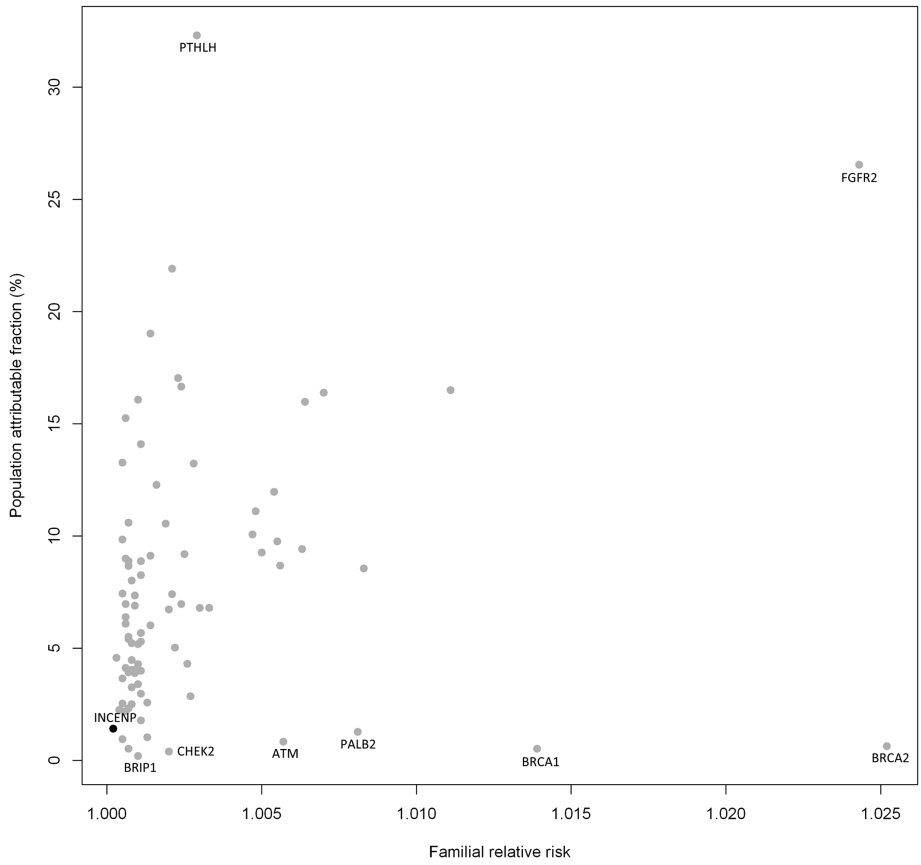


Figure 3. PAFs versus FRRs for all breast cancer susceptibility variants of low, moderate and high penetrance (gray dots). The top imputed *INCENP* SNP rs1047739 is shown as a black dot.

after PTHLH rs10771399 and the highest FRR among all common variants. The updated list of breast cancer susceptibility variants along with the corresponding ORs, MAFs, PAFs and FRRs are listed in Supplementary Table S10, available at *Carcinogenesis* Online.

Associations of INCENP SNPs with gene expression

All variants located between the upstream SNP rs4963459 and the downstream SNP rs4963471 of *INCENP*, including 103 SNPs available for 60 HapMap individuals (expression in lymphoblastoid cells) and 34 SNPs available for 447 TCGA individuals (expression in 60 normal breast tissue samples and 387 tumor breast tissue samples), were examined regarding their impact on gene expression. The mean expression level was −2.84 (±0.54) in the complete set of normal breast tissue samples and −1.88 (±0.80) in the complete set of tumor breast tissue samples ($P < 0.0001$). Two SNPs [expected five (103 × 0.05)] were associated with gene expression in lymphoblastoid cells and one SNP [expected two (34 × 0.05)] was associated with gene expression in normal breast tissue. All of these three SNPs were also associated with overall and ER-negative breast cancer risk. Nine SNPs [expected two (34 × 0.05)] were associated with gene expression in tumor breast tissue. However, none of these were associated with risk of breast cancer. Distribution of INCENP expression levels per SNP genotypes is displayed in Supplementary Figure S3, available at *Carcinogenesis* Online.

Potential functional *INCENP* SNPs

The *INCENP* SNP rs1047739 was annotated with three histone modifications by H3K27me3, H3K36me3 and H4K20me1, indicating an actively transcribed and accessible chromatin region that marks RNA polymerase II elongation and a silenced promoter. Moreover, rs1047739 overlapped with the chromatin state of transcriptional elongation. Altogether, the 15 variants tightly linked ($r^2 \geq 0.8$) to SNP rs1047739 showed features consistent to open chromatin, promoter silencing, blocked enhancer activity and repressed gene expression. Detailed information is presented in Supplementary Table S11, available at *Carcinogenesis* Online.

Discussion

This is the first study that investigates whether genetic variability in genes of the core CPC including *INCENP*, *AURKB*, *BIRC5* and *CDCA8* may affect primarily the overall, familial and subtype-specific breast cancer risk and secondarily the survival.

The INCENP protein of the CPC is a scaffold protein that comprises two functional subunits: The N-terminus binds to BIRC5 and CDCA8, which is required for the localization of the complex to the centromeres of chromosomes, whereas the conserved C-terminus binds AURKB partly activating the kinase. This allows AURKB to phosphorylate a C-terminal Thr-Ser-Ser motif in INCENP and a Thr in the T-loop of its kinase domain, resulting in full AURKB activation (40,41). INCENP is phosphorylated not only by AURKB but also by Cdk1, which is involved in Polo-like kinase 1 recruitment to the kinetochores and also in the progression from metaphase to anaphase (42). Yet, the molecular mechanisms by which SNPs in *INCENP* and other CPC genes influence breast cancer risk are unknown.

We found that several genotyped and imputed SNPs within and downstream of *INCENP* were associated with overall

and particularly with ER-negative breast cancer risk. The SNP rs1675126 showed the strongest association signal with overall and ER-negative breast cancer risk among all genotyped *INCENP* SNPs. The association with ER-negative breast cancer risk is of particular interest. Only 20–25% of all breast tumors are ER-negative. ER-negative breast cancer is often diagnosed at an earlier age and has a worse prognosis than ER-positive breast cancer. So far, seven loci specifically associated to ER-negative breast cancer susceptibility have been identified (43). The imputed SNP that showed the strongest association signal in the overall breast cancer risk analysis was rs1047739 located in the 3′ untranslated region. Even though rs1047739 showed also a differentiated association regarding ER tumor status, imputed rs144045115 downstream of *INCENP* showed the strongest association with ER-negative breast cancer risk. *In silico* analyses indicated that rs1047739 is located in an accessible chromatin region actively transcribed and that three miRNAs (has-miR-346, has-miR-632 and has-miR-654-3P predicted by Targetscan and MicroCosm Targets 5) bind to the rs1047739 containing region, suggesting that it may be the causal variant. The exact molecular mechanisms of how rs1047739 influences *INCENP* transcription should be further investigated *in vitro*. It has been previously reported that expression levels of INCENP are increased in tumor cells (22). This is in line with our finding that INCENP expression was increased in tumor breast tissue samples compared with normal breast tissue samples based on data from TCGA.

In a previous publication, rs2241909 in *AURKB* was associated with familial breast cancer risk in a German study population (30). We could not replicate this association in our present large data set from several European study populations. Instead we observed that the two SNPs rs3764384 and rs2071214 in *BIRC5* were associated with familial breast cancer risk. It was also observed that rs2306625 in *CDCA8* was particularly associated with relapse-free survival. Therefore, rs2306625 may eventually influence both the risk of disease onset and in case of tumor development the pathological characteristics of the tumor and/or its response to treatment. The per G allele increased risk and better-prognosis finding would contrast with *BRCA1/2* mutations in breast cancer (44), but would mimic on the other hand mutations of mismatch repair genes in colorectal cancer (45).

Low recombination rates downstream of *INCENP* allowed for genotype imputation in the region of six secretoglobin genes (*SCGBs*) taking 10 genotyped *INCENP* SNPs as reference. *SCGBs* are members of a supergene family and most of them are localized in a dense cluster on chromosome 11q13 including *SCGB1D1*, *SCGB1D2*, *SCGB2A1*, *SCGB2A2*, *SCGB1D4* and *SCGB1A1* (46). The *SCGBs* encode small secretory proteins and seem to play a role in the modulation of inflammation, tissue repair and tumorigenesis. Some *SCBGs* are overexpressed in breast cancer (47–49) and are more frequently associated with ER-positive tumors (50,51). Interestingly, imputation results indicated variants in the region of *SCGB1D1* throughout *SCGB1A1*, which were associated with overall breast cancer risk [top SNP rs3781965 (located in intron 2 of *SCGB1D1*): per minor T allele OR 1.02, 95% CI 1.00–1.04, $P = 0.001$], whereas associations with ER-negative tumors were detected only for the upstream and gene region of *SCGB1D1* [top SNP rs2232935 (located upstream of *SCGB1D1*): per minor T allele OR 1.05, 95% CI 1.02–1.09, $P = 0.0003$]. Even though the biological functions of *SCGB* products are still poorly understood and variants in the 3′ region of *INCENP* showed stronger association signals in breast cancer risk analysis, initial results on the possible association between inherited variation in *SCGBs* and breast cancer susceptibility should be explored based on directly typed variants in future consortial work.

In conclusion, taking advantage of BCAC and COGS efforts that translated into a homogeneous, high-quality genotyping of 88 911 women from 39 European studies, we were able to identify potential novel variants in the *INCENP* gene, which associated with a 3% per allele increased risk of breast cancer, and with a 6% per allele increased risk of ER-negative breast tumors. This study demonstrates the benefit of scientific collaborations leading to large sample collections in order to identify low-penetrance variants, in particular for disease subtypes. It is likely that next generation sequencing in combination with the integration of information on additional layers of genetic variability will refine marker association signals and unravel increasing proportions of sporadic and familial cases of disease. In parallel, the identification of new susceptibility variants may point to novel drug targets. Due to the established involvement in the regulation of cell division, this is probably the most relevant aspect of the identified associations between CPC variants and breast cancer.

Supplementary material

Supplementary Materials and methods, Results, Tables S1–S11 and Figures S1–S3 can be found at http://carcin.oxfordjournals.org/

Acknowledgements

The results shown in this article are in part based upon data generated by the TCGA Research Network: http://cancergenome.nih.gov; the ENCODE Consortium and the ENCODE production laboratory(s): http://genome.ucsc.edu/ENCODE; the Gene Expression Omnibus contributors: http://www.ncbi.nlm.nih.gov/geo and the 1000 Genomes Project Consortium: http://www.1000genomes.org/home. BCAC: We thank all the individuals who took part in these studies and all the researchers, clinicians, technicians and administrative staff who have enabled this work to be carried out. COGS: This study would not have been possible without the contributions of the following: PH (COGS); DFE, PP, KM, MKB, QW (BCAC), Andrew Berchuck (OCAC), Rosalind A. Eeles, DFE, Ali Amin Al Olama, Zsofia Kote-Jarai, Sara Benlloch (PRACTICAL), GCT, Antonis Antoniou, Lesley McGuffog, FJC and Ken Offit (CIMBA), JD, AMD, Andrew Lee, and Ed Dicks, CL and the staff of the Centre for Genetic Epidemiology Laboratory, JB, AGN and the staff of the CNIO genotyping unit, JS and DCT, FB, DV, Sylvie LaBoissière and Frederic Robidoux and the staff of the McGill University and Génome Québec Innovation Centre, SEB, SFN, BGN, and the staff of the Copenhagen DNA laboratory, and Julie M. Cunningham, Sharon A. Windebank, Christopher A. Hilker, Jeffrey Meyer and the staff of Mayo Clinic Genotyping Core Facility. SEARCH: The SEARCH and EPIC teams. pKARMA: The Swedish Medical Research Counsel. CGPS: Staff and participants of the Copenhagen General Population Study. For the excellent technical assistance: Dorthe Uldall Andersen, Maria Birna Arnadottir, Anne Bank, Dorthe Kjeldgård Hansen. The Danish Breast Cancer Group (DBCG) is acknowledged for the tumor information. LMBC: Gilian Peuteman, Dominiek Smeets, Thomas Van Brussel and Kathleen Corthouts. MARIE: Judith Heinz, Nadia Obi, Alina Vrieling, Sabine Behrens, Ursula Eilber, Muhabbet Celik, Til Olchers, Stefan Nickels. BBCS: Eileen Williams, Elaine Ryder-Mills, Kara Sargus. HEBCS: Kirsimari Aaltonen, Karl von Smitten, Sofia Khan, Tuomas Heikkinen, Irja Erkkilä. ABCS: Sten Cornelissen, Richard van Hien, Linde Braaf,

FBLH, Senno Verhoef, Laura van 't Veer, Emiel Rutgers, Ellen van der Schoot, Femke Atsma. SASBAC: The Swedish Medical Research Counsel. BSUCH: Medical Faculty Mannheim, Germany (PB). CNIO-BCS: Guillermo Pita, Charo Alonso, Daniel Herrero, Nuria Álvarez, MPZ, Primitiva Menendez, the Human Genotyping-CEGEN Unit (CNIO). SBCS: Sue Higham, Helen Cramp, Ian Brock, Sabapathy Balasubramanian and Dan Connley. OFBCR: Teresa Selander, Nayana Weerasooriya. BIGGS: Niall McInerney, Gabrielle Colleran, Andrew Rowan, Angela Jones. kConFab/AOCS: We wish to thank Heather Thorne, Eveline Niedermayr, all the kConFab research nurses and staff, the heads and staff of the Family Cancer Clinics, and the Clinical Follow Up Study (which has received funding from the National Health and Medical Research Council, the National Breast Cancer Foundation, Cancer Australia and the National Institute of Health (USA) for their contributions to this resource, and the many families who contribute to kConFab. RBCS: Petra Bos, Jannet Blom, Ellen Crepin, Anja Nieuwlaat, Annette Heemskerk, the Erasmus MC Family Cancer Clinic. ABCFS: Maggie Angelakos, Judi Maskiell, Gillian Dite. TNBCC: Robert Pilarski and Charles Shapiro were instrumental in the formation of the OSU Breast Cancer Tissue Bank. We thank the Human Genetics Sample Bank for processing of samples and providing OSU Columbus area control samples. BBCC: Silke Landrith, Alexander Hein, Sonja Oeser, Michael Schneider. ESTHER: Hartwig Ziegler, Sonja Wolf, Volker Hermann. UKBGS: We thank Breakthrough Breast Cancer and the Institute of Cancer Research for support and funding of the Breakthrough Generations Study, and the study participants, study staff, and the doctors, nurses and other health care providers and health information sources who have contributed to the study. We acknowledge National Health Service (NHS) funding to the Royal Marsden/ICR NIHR Biomedical Research Centre. PBCS: Louise Brinton, Mark Sherman, Neonila Szeszenia-Dabrowska, Beata Peplonska, Witold Zatonski, Pei Chao, Michael Stagner. MTLGEBCS: We would like to thank Martine Tranchant (CHU de Québec Research Center), Marie-France Valois, Annie Turgeon and Lea Heguy (McGill University Health Centre, Royal Victoria Hospital; McGill University) for DNA extraction, sample management and skillful technical assistance. J.S. is Chairholder of the Canada Research Chair in Oncogenetics. OBCS: Meeri Otsukka, Kari Mononen. GENICA: The GENICA Network: Dr Margarete Fischer-Bosch-Institute of Clinical Pharmacology, Stuttgart, and University of Tübingen, Germany; (HBra, Wing-Yee Lo, Christina Justenhoven), Department of Internal Medicine, Evangelische Kliniken Bonn gGmbH, Johanniter Krankenhaus, Bonn, Germany (YDK, Christian Baisch), Institute of Pathology, University of Bonn, Germany (Hans-Peter Fischer), Molecular Genetics of Breast Cancer, Deutsches Krebsforschungszentrum (DKFZ), Heidelberg, Germany (UH) and Institute for Prevention and Occupational Medicine of the German Social Accident Insurance, Institute of the Ruhr University Bochum (IPA), Bochum, Germany (TB, Beate Pesch, Sylvia Rabstein, Anne Lotz); Institute of Occupational Medicine and Maritime Medicine, University Medical Center Hamburg-Eppendorf, Germany (Volker Harth). MBCSG: Siranoush Manoukian, Bernard Peissel and Daniela Zaffaroni of the Fondazione IRCCS Istituto Nazionale dei Tumori (INT); Bernardo Bonanni, Monica Barile and Irene Feroce of the Istituto Europeo di Oncologia (IEO) and Loris Bernard the personnel of the Cogentech Cancer Genetic Test Laboratory. HMBCS: Peter Hillemanns, Hans Christiansen and Johann H. Karstens. KBCP: Eija Myöhänen, Helena Kemiläinen. ORIGO: We thank E. Krol-Warmerdam and J. Blom for patient accrual, administering questionnaires and managing clinical information. The LUMC survival data were retrieved from the Leiden hospital-based cancer registry system (ONCDOC) with the help of Dr J. Molenaar. NBHS: We thank study participants and research staff for their contributions and commitment to this study. SKKDKFZS: We thank all study participants, clinicians, family doctors, researchers and technicians for their contributions and commitment to this study. CTS: The CTS Steering Committee includes Leslie Bernstein, Susan Neuhausen, James Lacey, Sophia Wang, Huiyan Ma, Yani Lu and Jessica Clague DeHart at the Beckman Research Institute of City of Hope, Dennis Deapen, Rich Pinder, Eunjung Lee and Fred Schumacher at the University of Southern California, Pam Horn-Ross, Peggy Reynolds, Christina Clarke Dur and David Nelson at the Cancer Prevention Institute of California, and HAC, Argyrios Ziogas, and Hannah Park at the University of California Irvine. NBCS: NBCS includes the following clinical collaborators: Prof. Per Eystein Lønning, MD (Section of Oncology, Institute of Medicine, University of Bergen and Department of Oncology, Haukeland University Hospital, Bergen, Norway), Prof. Em. Sophie D. Fosså, MD (National Resource Centre for Long-term Studies after Cancer, Rikshospitalet-Radiumhospitalet Cancer Clinic Montebello, Oslo, Norway), Head physician Tone Ikdahl, MD (Department of Oncology, Oslo University Hospital, Oslo, Norway), Dr Lars Ottestad, MD (Department of Genetics and Department of Oncology, Oslo University Hospital Radiumhospitalet), Dr Marit Muri Holmen, MD (Department of Radiology, Oslo University Hospital Radiumhospitalet, Oslo, Norway), Dr Vilde Haakensen, MD (Department of Genetics and Department of Oncology, Oslo University Hospital Radiumhospitalet and Institute for Clinical Medicine, Faculty of Medicine, University of Oslo, Oslo, Norway), Prof. Bjørn Naume, MD (Division of Cancer Medicine and Radiotherapy, Department of Oncology, Oslo University Hospital Radiumhospitalet, Oslo, Norway), Associate Prof. Åslaug Helland, MD (Department of Genetics, Institute for Cancer Research and Department of Oncology, Oslo University Hospital Radiumhospitalet, Oslo, Norway and Institute of Clinical Medicine, Faculty of Medicine, University of Oslo, Oslo, Norway), Prof. Inger Torhild Gram, MD (Department of Community Medicine, Faculty of Health Sciences, University of Tromsø and Norwegian Centre for Integrated Care and Telemedicine, University Hospital of North Norway, Tromsø, Norway), Prof. Em. Rolf Kåresen, MD (Department of Breast and Endocrine Surgery, Institute for Clinical Medicine, Ullevaal Hospital, Oslo University Hospital and Institute of Clinical Medicine, Faculty of Medicine, University of Oslo, Oslo, Norway), Dr Ellen Schlichting, MD (Department for Breast and Endocrine Surgery, Oslo University Hospital Ullevaal, Oslo, Norway), Prof. Toril Sauer, MD (Department of Pathology at Akershus University Hospital, Lørenskog, Norway), Dr Olav Engebråten, MD (Institute for Clinical Medicine, Faculty of Medicine, University of Oslo and Department of Oncology, Oslo University Hospital, Oslo, Norway), Dr Margit Riis, MD (Department of Surgery, Akershus University Hospital and Department of Clinical Molecular Biology (EpiGen), Institute of Clinical Medicine, Akershus University Hospital, University of Oslo, Lørenskog, Norway).

This work was supported by the following: BCAC: Cancer Research UK [C1287/A10118, C1287/A12014] and by the European Community´s Seventh Framework Programme under grant agreement number 223175 (grant number HEALTH-F2-2009-223175) (COGS). iCOGS: the European Community's Seventh Framework Programme under grant agreement n° 223175 (HEALTH-F2-2009–223175) (COGS), Cancer Research UK (C1287/A10118, C1287/A 10710, C12292/A11174, C1281/A12014, C5047/A8384, C5047/A15007, C5047/A10692), the National Institutes of

Health (CA128978) and Post-Cancer GWAS initiative (1U19 CA148537, 1U19 CA148065 and 1U19 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. SEARCH: Cancer Research UK [C490/A10124] and UK National Institute for Health Research Biomedical Research Centre at the University of Cambridge. pKARMA: Märit and Hans Rausings Initiative Against Breast Cancer. CGPS: Chief Physician Johan Boserup and Lise Boserup Fund, the Danish Medical Research Council and Herlev Hospital. LMBC: 'Stichting tegen Kanker' (232–2008 and 196–2010). D.L. is supported by the FWO and the KULPFV/10/016-SymBioSysII. MCBCS: NIH grants CA128978, CA116167, CA176785 an NIH Specialized Program of Research Excellence (SPORE) in Breast Cancer [CA116201], and the Breast Cancer Research Foundation and a generous gift from the David F. and Margaret T. Grohne Family Foundation and the Ting Tsung and Wei Fong Chao Foundation. MARIE: Deutsche Krebshilfe e.V. [70-2892-BR I], 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 [01KH0402]. BBCS: Cancer Research UK and Breakthrough Breast Cancer and acknowledges NHS funding to the NIHR Biomedical Research Centre, and the National Cancer Research Network (NCRN). HEBCS: Helsinki University Central Hospital Research Fund, Academy of Finland (266528), the Finnish Cancer Society, The Nordic Cancer Union and the Sigrid Juselius Foundation. ABCS: Dutch Cancer Society [grants NKI 2007–3839 and 2009 4363]; BBMRI-NL, which is a Research Infrastructure financed by the Dutch government (NWO 184.021.007); and the Dutch National Genomics Initiative. SASBAC: Agency for Science, Technology and Research of Singapore (A*STAR), the United States National Institute of Health (NIH) and the Susan G. Komen Breast Cancer Foundation. CECILE: Fondation de France, Institut National du Cancer (INCa), Ligue Nationale contre le Cancer, Ligue contre le Cancer Grand Ouest, Agence Nationale de Sécurité Sanitaire (ANSES), Agence Nationale de la Recherche (ANR). BSUCH: Dietmar-Hopp Foundation, the Helmholtz Society and the German Cancer Research Center (DKFZ). CNIO-BCS: Genome Spain Foundation, the Red Temática de Investigación Cooperativa en Cáncer and grants from the Asociación Española Contra el Cáncer and the Fondo de Investigación Sanitario (PI11/00923 and PI081120). The Human Genotyping-CEGEN Unit (CNIO) is supported by the Instituto de Salud Carlos III. SBCS: Yorkshire Cancer Research S295, S299, S305PA and Sheffield Experimental Cancer Medicine Centre. Ontario Familial Breast Cancer Registry (OFBCR): UM1 CA164920 from the National Cancer Institute (USA). 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 USA Government or the BCFR. BIGGS ES: NIHR Comprehensive Biomedical Research Centre, Guy's and St. Thomas' NHS Foundation Trust in partnership with King's College London, UK. I.T. is supported by the Oxford Biomedical Research Centre. kConFab: National Breast Cancer Foundation, and previously by the National Health and Medical Research Council (NHMRC), the Queensland Cancer Fund, the Cancer Councils of New South Wales, Victoria, Tasmania and South Australia, and the Cancer Foundation of Western Australia. MEC: NIH grants CA63464, CA54281, CA098758 and CA132839. Financial support for KARBAC 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. RBCS: Dutch Cancer Society (DDHK 2004–3124, DDHK 2009–4318). The Australian Breast Cancer Family Study (ABCFS) was supported by grant UM1 CA164920 from the National Cancer Institute (USA). 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 USA Government or the BCFR. The ABCFS was also supported by the NHMRC of Australia, the New South Wales Cancer Council, the Victorian Health Promotion Foundation (Australia) and the Victorian Breast Cancer Research Consortium. J.L.H. is a NHMRC Australia Fellow and a Victorian Breast Cancer Research Consortium Group Leader. M.C.S. is a NHMRC senior research fellow and a Victorian Breast Cancer Research Consortium Group Leader. TNBCC: Specialized Program of Research Excellence (SPORE) in Breast Cancer (CA116201), a grant from the Breast Cancer Research Foundation, a generous gift from the David F. and Margaret T. Grohne Family Foundation and the Ting Tsung and Wei Fong Chao Foundation, the Stefanie Spielman Breast Cancer fund and the OSU Comprehensive Cancer Center, DBBR (a CCSG Share Resource by National Institutes of Health Grant P30 CA016056), the Hellenic Cooperative Oncology Group research grant (HR R_BG/04) and the Greek General Secretary for Research and Technology (GSRT) Program, Research Excellence II, the European Union (European Social Fund – ESF) and Greek national funds through the Operational Program 'Education and Lifelong Learning' of the National Strategic Reference Framework (NSRF)—ARISTEIA. 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. BBCC: ELAN-Fond of the University Hospital of Erlangen. ESTHER: Baden Württemberg Ministry of Science, Research and Arts. Additional cases were recruited in the context of the VERDI study, which was supported by a grant from the German Cancer Aid (Deutsche Krebshilfe). UKBGS: Breakthrough Breast Cancer and the Institute of Cancer Research (ICR), London. ICR acknowledges NHS funding to the NIHR Biomedical Research Centre. PBCS: Intramural Research Funds of the National Cancer Institute, Department of Health and Human Services, USA. MTLGEBCS: Quebec Breast Cancer Foundation, the Canadian Institutes of Health Research for the 'CIHR Team in Familial Risks of Breast Cancer' program—grant # CRN-87521 and the Ministry of Economic Development, Innovation and Export Trade—grant # PSR-SIIRI-701. OBCS: Finnish Cancer Foundation, the Academy of Finland (grant number 250083, 122715 and Center of Excellence grant number 251314), the Finnish Cancer Foundation, the Sigrid Juselius Foundation, the University of Oulu, the University of Oulu Support Foundation and the special Governmental EVO funds for Oulu University Hospital-based research activities. GENICA: Federal Ministry of Education and Research (BMBF) Germany grants 01KW9975/5, 01KW9976/8, 01KW9977/0 and 01KW0114, the Robert Bosch Foundation, Stuttgart, Deutsches Krebsforschungszentrum (DKFZ), Heidelberg, the Institute for Prevention and Occupational Medicine of the German Social Accident Insurance, Institute of the Ruhr University Bochum (IPA), Bochum, as well as the Department of Internal Medicine, Evangelische Kliniken Bonn gGmbH, Johanniter Krankenhaus, Bonn, Germany. MBCSG: Italian Association for Cancer Research (AIRC) and by funds from the Italian citizens who allocated the 5/1000 share of their tax payment in support of the Fondazione IRCCS Istituto Nazionale Tumori, according to Italian laws

(INT-Institutional strategic projects '5 × 1000'). HMBCS: Friends of Hannover Medical School and by the Rudolf Bartling Foundation. KBCP: Government Funding (EVO) of Kuopio University Hospital grants, Cancer Fund of North Savo, the Finnish Cancer Organizations and by the strategic funding of the University of Eastern Finland. ORIGO: Dutch Cancer Society (RUL 1997-1505) and the Biobanking and Biomolecular Resources Research Infrastructure (BBMRI-NL CP16). The SZBCS was supported by grant PBZ_KBN_122/P05/2004. NBHS: NIH grant R01CA100374. Biological sample preparation was conducted the Survey and Biospecimen Shared Resource, which is supported by P30 CA68485. SKKDKFZS: DKFZ. CTS: California Breast Cancer Act of 1993 and the California Breast Cancer Research Fund (contract 97-10500) and is currently funded through the National Institutes of Health (R01 CA77398). Collection of cancer incidence data was supported by the California Department of Public Health as part of the statewide cancer reporting program mandated by California Health and Safety Code Section 103885. HAC receives support from the Lon V Smith Foundation (LVS39420). NBCS: Norwegian Research council, 155218/V40, 175240/S10 to ALBD, FUGE-NFR 181600/V11 to VNK and a Swizz Bridge Award to ALBD.
Conflict of Interest Statement: None declared.

References

1. Ferlay, J. et al. (2010) Estimates of worldwide burden of cancer in 2008: GLOBOCAN 2008. Int. J. Cancer, 127, 2893–2917.
2. Lichtenstein, P. et al. (2000) Environmental and heritable factors in the causation of cancer–analyses of cohorts of twins from Sweden, Denmark, and Finland. N. Engl. J. Med., 343, 78–85.
3. Peto, J. et al. (2000) High constant incidence in twins and other relatives of women with breast cancer. Nat. Genet., 26, 411–414.
4. Ghoussaini, M. et al. (2013) Inherited genetic susceptibility to breast cancer: the beginning of the end or the end of the beginning? Am. J. Pathol., 183, 1038–1051.
5. Michailidou, K. et al. (2013) Large-scale genotyping identifies 41 new loci associated with breast cancer risk. Nat. Genet., 45, 353–361.
6. Hemminki, K. et al. (2008) Etiologic impact of known cancer susceptibility genes. Mutat. Res., 658, 42–54.
7. Carmena, M. et al. (2012) The chromosomal passenger complex (CPC): from easy rider to the godfather of mitosis. Nat. Rev. Mol. Cell Biol., 13, 789–803.
8. van der Waal, M.S. et al. (2012) Cell division control by the Chromosomal Passenger Complex. Exp. Cell Res., 318, 1407–1420.
9. Jeyaprakash, A.A. et al. (2007) Structure of a Survivin-Borealin-INCENP core complex reveals how chromosomal passengers travel together. Cell, 131, 271–285.
10. Ainsztein, A.M. et al. (1998) INCENP centromere and spindle targeting: identification of essential conserved motifs and involvement of heterochromatin protein HP1. J. Cell Biol., 143, 1763–1774.
11. Klein, U.R. et al. (2006) Centromere targeting of the chromosomal passenger complex requires a ternary subcomplex of Borealin, Survivin, and the N-terminal domain of INCENP. Mol. Biol. Cell, 17, 2547–2558.
12. Vader, G. et al. (2006) Survivin mediates targeting of the chromosomal passenger complex to the centromere and midbody. EMBO Rep., 7, 85–92.
13. Adams, R.R. et al. (2000) INCENP binds the Aurora-related kinase AIRK2 and is required to target it to chromosomes, the central spindle and cleavage furrow. Curr. Biol., 10, 1075–1078.
14. Adams, R.R. et al. (2001) Essential roles of Drosophila inner centromere protein (INCENP) and aurora B in histone H3 phosphorylation, metaphase chromosome alignment, kinetochore disjunction, and chromosome segregation. J. Cell Biol., 153, 865–880.
15. Earnshaw, W.C. et al. (1991) Analysis of the distribution of the INCENPs throughout mitosis reveals the existence of a pathway of structural changes in the chromosomes during metaphase and early events in cleavage furrow formation. J. Cell Sci., 98 (Pt 4), 443–461.
16. Fraser, A.G. et al. (1999) Caenorhabditis elegans inhibitor of apoptosis protein (IAP) homologue BIR-1 plays a conserved role in cytokinesis. Curr. Biol., 9, 292–301.
17. Honda, R. et al. (2003) Exploring the functional interactions between Aurora B, INCENP, and survivin in mitosis. Mol. Biol. Cell, 14, 3325–3341.
18. Kaitna, S. et al. (2000) Incenp and an aurora-like kinase form a complex essential for chromosome segregation and efficient completion of cytokinesis. Curr. Biol., 10, 1172–1181.
19. Hümmer, S. et al. (2009) Cdk1 negatively regulates midzone localization of the mitotic kinesin Mklp2 and the chromosomal passenger complex. Curr. Biol., 19, 607–612.
20. Szafer-Glusman, E. et al. (2011) Role of Survivin in cytokinesis revealed by a separation-of-function allele. Mol. Biol. Cell, 22, 3779–3790.
21. Terada, Y. et al. (1998) AIM-1: a mammalian midbody-associated protein required for cytokinesis. EMBO J., 17, 667–676.
22. Adams, R.R. et al. (2001) Human INCENP colocalizes with the Aurora-B/AIRK2 kinase on chromosomes and is overexpressed in tumour cells. Chromosoma, 110, 65–74.
23. Ambrosini, G. et al. (1997) A novel anti-apoptosis gene, survivin, expressed in cancer and lymphoma. Nat. Med., 3, 917–921.
24. Tanaka, T. et al. (1999) Centrosomal kinase AIK1 is overexpressed in invasive ductal carcinoma of the breast. Cancer Res., 59, 2041–2044.
25. Tatsuka, M. et al. (1998) Multinuclearity and increased ploidy caused by overexpression of the aurora- and Ipl1-like midbody-associated protein mitotic kinase in human cancer cells. Cancer Res., 58, 4811–4816.
26. Lens, S.M. et al. (2010) Shared and separate functions of polo-like kinases and aurora kinases in cancer. Nat. Rev. Cancer, 10, 825–841.
27. Jha, K. et al. (2012) Survivin expression and targeting in breast cancer. Surg. Oncol., 21, 125–131.
28. Janssen, A. et al. (2011) Chromosome segregation errors as a cause of DNA damage and structural chromosome aberrations. Science, 333, 1895–1898.
29. Ricke, R.M. et al. (2008) Whole chromosome instability and cancer: a complex relationship. Trends Genet., 24, 457–466.
30. Tchatchou, S. et al. (2007) Aurora kinases A and B and familial breast cancer risk. Cancer Lett., 247, 266–272.
31. Xu, Y. et al. (2004) A mutation found in the promoter region of the human survivin gene is correlated to overexpression of survivin in cancer cells. DNA Cell Biol., 100, 527–537.
32. Abecasis, G.R. et al. (2012) An integrated map of genetic variation from 1,092 human genomes. Nature, 491, 56–65.
33. Boyle, P. et al. (2005) Cancer incidence and mortality in Europe, 2004. Ann. Oncol., 16, 481–488.
34. Hemminki, K. et al. (2007) Constraints for genetic association studies imposed by attributable fraction and familial risk. Carcinogenesis, 28, 648–656.
35. Mavaddat, N. et al. (2010) Genetic susceptibility to breast cancer. Mol. Oncol., 4, 174–191.
36. The Cancer Genome Atlas Network (2012) Comprehensive molecular portraits of human breast tumours. Nature, 490, 61–70.
37. Stranger, B.E. et al. (2007) Relative impact of nucleotide and copy number variation on gene expression phenotypes. Science, 315, 848–853.
38. Coetzee, S.G. et al. (2012) FunciSNP: an R/bioconductor tool integrating functional non-coding data sets with genetic association studies to identify candidate regulatory SNPs. Nucleic Acids Res., 40, e139.
39. The ENCODE Project Consortium (2012) An integrated encyclopedia of DNA elements in the human genome. Nature, 489, 57–74.
40. Bishop, J.D. et al. (2002) Phosphorylation of the carboxyl terminus of inner centromere protein (INCENP) by the Aurora B Kinase stimulates Aurora B kinase activity. J. Biol. Chem., 277, 27577–27580.
41. Sessa, F. et al. (2005) Mechanism of Aurora B activation by INCENP and inhibition by hesperadin. Mol. Cell, 18, 379–391.
42. Goto, H. et al. (2006) Complex formation of Plk1 and INCENP required for metaphase-anaphase transition. Nat. Cell Biol., 8, 180–187.
43. Campa, D. et al. (2014) A genome-wide "pleiotropy scan" does not identify new susceptibility loci for estrogen receptor negative breast cancer. PLoS One, 9, e85955.
44. Huzarski, T. et al. (2013) Ten-year survival in patients with BRCA1-negative and BRCA1-positive breast cancer. J. Clin. Oncol., 31, 3191–3196.
45. Clark, A.J. et al. (2004) Prognosis in DNA mismatch repair deficient colorectal cancer: are all MSI tumours equivalent? Fam. Cancer, 3, 85–91.

46. Jackson, B.C. et al. (2011) Update of the human secretoglobin (SCGB) gene superfamily and an example of 'evolutionary bloom' of androgen-binding protein genes within the mouse Scgb gene superfamily. Hum. Genomics, 5, 691–702.
47. Watson, M.A. et al. (1996) Mammaglobin, a mammary-specific member of the uteroglobin gene family, is overexpressed in human breast cancer. Cancer Res., 56, 860–865.
48. Watson, M.A. et al. (1999) Mammaglobin expression in primary, metastatic, and occult breast cancer. Cancer Res., 59, 3028–3031.
49. Culleton, J. et al. (2007) Lipophilin B: a gene preferentially expressed in breast tissue and upregulated in breast cancer. Int. J. Cancer, 120, 1087–1092.
50. O'Brien, N. et al. (2002) Mammaglobin a: a promising marker for breast cancer. Clin. Chem., 48, 1362–1364.
51. Span, P.N. et al. (2004) Mammaglobin is associated with low-grade, steroid receptor-positive breast tumors from postmenopausal patients, and has independent prognostic value for relapse-free survival time. J. Clin. Oncol., 22, 691–698.