MicroRNA Related Polymorphisms and Breast Cancer Risk

Sofia Khan¹, Dario Greco¹,2, Kyraki Michailidou³, Roger L. Milne⁴,⁵, Taru A. Muranen¹,
Tuomas Heikkinen¹, Kirsimari Aaltonen¹,⁶,⁷, Joe Dennis⁸, Manjeet K. Bolla², Jianjun Liu⁹, Per Hall⁹,
Astrid Irwanto⁴, Keith Humphreys⁹, Jingmei Li³, Kamila Czene⁶, Jenny Chang-Claude¹⁰,
Rebecca Hein¹⁰,¹¹, Anja Rudolph¹⁰, Petra Seibold¹⁰, Dieter Flesch-Janys¹², Olivia Fletcher¹³,
Julian Peto¹⁴, Isabel dos Santos Silva¹⁴, Nichola Johnson¹³, Lorna Gibson¹⁴, Zoe Aitken¹⁴,
John L. Hopper¹⁵, Helen Tsimiklis¹⁶, Minh Bui¹⁵, Enes Makalic¹⁵, Daniel F. Schmidt¹⁵,
Melissa C. Southey¹⁶, Carmel Apicella¹⁵, Jennifer Stone¹⁵, Quinten Waisfisz¹⁷, Hanne Meijers-Heijboer¹⁷,
Muriel A. Adank¹⁷, Rob B. van der Luijt¹⁸, Alfons Meindl¹⁹, Rita K. Schmutzler²⁰,²¹,²²,²³,
Bertram Müller-Myhsok²⁴, Peter Lichtner²⁵, Clare Turnbull²⁶, Nazneen Rahman²⁶, Stephen J. Chanock²⁷,
David J. Hunter²⁸,²⁹, Angela Cox³⁰, Simon S. Cross³¹, Malcolm W. R. Reed³⁰, Marjanka K. Schmida³²,
Annegien Broeks³², Laura J. Van’t Veer³², Frans B. Hogervorst³², Peter A. Fasching³³,³⁴,
Michael G. Schrauder³³, Arif B. Ekici³⁵, Matthias W. Beckmann³³, Stig E. Bojesen³⁶,³⁷,
Børge G. Nordestgaard³⁶,³⁷, Sune F. Nielsen³⁶,³⁷, Henrik Flyger³⁸, Javier Benitez³⁹,⁴⁰, Pilar M. Zamora⁴¹,
Jose I. Perez⁴², Christopher A. Haiman⁴³, Brian E. Henderson⁴⁶, Fredrick Schumacher⁴³,
Loïc Le Marchand⁴⁴, Paul D. Pharoah⁴⁵, Alison M. Dunning⁴⁵, Mitul Shah⁴⁵, Robert Luben⁴⁶,
Judith Brown⁴⁶, Fergus J. Couch⁴⁷, Xianshu Wang⁴⁷, Celine Vachon⁴⁸, Janet E. Olson⁴⁸,
Diether Lambrechts⁴⁹,⁵⁰, Matthieu Moïsse⁴⁹,⁵⁰, Robert Paridaens⁵¹, Marie-Rose Christiaens⁵¹,
Pascal Guéné⁵²,⁵³, Thérèse Truong⁵²,⁵³, Pierre Laurent-Puig⁵⁴, Claire Mulot⁵⁴, Frederick Marme⁵⁵,⁵⁶,
Barbara Burwinkel⁵⁵,⁵⁷, Andreas Schneeweiss⁵⁵,⁵⁶, Christof Sohn⁵⁵, Elinor J. Sawyer⁵⁸, Ian Tomlinson⁵⁹,
Michael J. Kerin⁶⁰, Nicola Miller⁶⁰, Irene L. Andrilis⁶¹,⁶², Julia A. Knight⁶³,⁶⁴, Sandrine Tchatchou⁶¹,
Anna Marie Mulligan⁶⁵,⁶⁶, Thilo Dörk⁶⁷, Natalia V. Bogdanova⁶⁸, Natalia N. Antonenkova⁶⁹,
Hoda Anton-Culver⁷⁰, Hatief Darabi⁷¹, Mikael Eriksson⁷¹, Montserrat Garcia-Closas⁷¹,⁷², José Filgueira⁷¹,
Jolanta Lissowska⁷³, Louise Brinton⁷³, Peter Devilee⁷⁵, Robert A. E. M. Tollenaar⁷⁵, Caroline Seynaeve⁷⁶,
Christi J. van Asperen⁷⁷, Vessela N. Kristensen⁷⁸,⁷⁹,⁸⁰, KConFab Investigators⁸¹,*
Australian Ovarian Cancer Study Group⁸¹,⁸²,*, Susan Slager⁸⁶, Amanda E. Toland⁸³,
Christine B. Ambrosone⁸⁴,†, Drakoulis Yannoukakos⁸⁵, Annika Lindblom⁸⁶, Sara Margolin⁸⁷,
Paolo Radice⁸⁸, Paolo Peterlongo⁸⁹, Monica Barile⁹⁰, Paolo Mariani⁹¹,⁹², Martje J. Hooning⁹²,
John W. M. Martens⁹², José M. Gregoire Colle⁹³, Agnes J. L. Bass⁹³, Anna Jakubowska⁹⁴, Jan Lubinski⁹⁴,
Katarzyna Jaworska-Bieniek⁹⁵, Katarzyna Durda⁹⁵, Graham G. Giles⁴,⁵, Catriona McLean⁹⁶,
Hiltrud Brauch⁹⁷,⁹⁸, Thomas Brüning⁹⁷,⁹⁸, On-Dschun Ko¹⁰⁰,
The GENICA Network⁹⁷,⁹⁸,⁹⁹,¹⁰¹,¹⁰²,¹⁰³,*, Hermann Brenner¹⁰⁴,¹⁰⁵, Aida Karina Dieffenbach¹⁰⁴,¹⁰⁵,
Volker Arndt¹⁰⁴, Christa Stegmaier¹⁰⁶, Anthony Swerdlow¹⁰⁷, Alan Ashworth¹³, Nick Orr¹³,
Michael Jones⁷¹, Jacques Simard¹⁰⁸, Mark S. Goldberg¹⁰⁹,¹¹⁰, France Labrèche¹¹¹, Martine Dumont¹⁰⁸,
Robert Winqvist¹¹², Katri Pylkäs¹¹², Arja Jukkola-Vuornen¹¹³, Mervi Grip¹¹⁴, Vesa Kataja¹¹⁵,¹¹⁶,
Veli-Matti Kosma¹¹⁷,¹¹⁸,¹¹⁹, Jaana M. Hartikainen¹¹⁷,¹¹⁸,¹¹⁹, Arto Mannermaa¹¹⁷,¹¹⁸,¹¹⁹, Ute Hamann¹⁰¹,
Georgia Chenevix-Trench¹²⁰, Carl Blomqvist¹², Kristina Aittomäki¹², Douglas F. Easton³,³⁴,
Heli Nevanlinna¹⁶.

¹ Department of Obstetrics and Gynecology, University of Helsinki and Helsinki University Central Hospital, Helsinki, Finland, ² Finnish Institute of Occupational Health, Helsinki, Finland, ³ Centre for Cancer Genetic Epidemiology, Department of Public Health and Primary Care, University of Cambridge, Cambridge, United Kingdom, ⁴ Cancer Epidemiology Centre, Cancer Council Victoria, Melbourne, Australia, ⁵ Centre for Epidemiology and Biostatistics, Melbourne School of Population and Global Health, The University of Melbourne, Melbourne, Australia, ⁶ Department of Clinical Genetics, University of Helsinki and Helsinki University Central Hospital, Helsinki, Finland, ⁷ Department of Oncology, University of Helsinki and Helsinki University Central Hospital, Helsinki, Finland, ⁸ Department of Epidemiology and Biostatistics, Karolinska Institutet, Stockholm, Sweden, ⁹ Division of Cancer Epidemiology, German Cancer Research Center (DKFZ), Heidelberg, Germany, ¹⁰ PMV Research Group at the Department of Child and Adolescent Psychiatry and Psychotherapy, University of Cologne, Cologne, Germany, ¹¹ Department of Non-Communicable Disease Epidemiology Department, London School of Hygiene and Tropical Medicine, London, United Kingdom, ¹² Centre for Cancer Epidemiology/Genetic Cancer Registry and Institute for Medical Biometrics and Epidemiology, University Clinic Hamburg-Eppendorf, Hamburg, Germany, ¹³ Breakthrough Breast Cancer Research Centre, The Institute of Cancer Research, London, United Kingdom, ¹⁴ Department of Obstetrics and Gynecology, University of Helsinki and Helsinki University Central Hospital, Helsinki, Finland, ¹⁵ Centre for Cancer Research, Department of Non-Communicable Disease Epidemiology, Department, London School of Hygiene and Tropical Medicine, London, United Kingdom, ¹⁶ Department of Obstetrics and Gynecology, University of Helsinki and Helsinki University Central Hospital, Helsinki, Finland.

* Corresponding author: Heli Nevanlinna, helen@husari.oulu.fi
Abstract

Genetic variations, such as single nucleotide polymorphisms (SNPs) in microRNAs (miRNA) or in the miRNA binding sites may affect the miRNA dependent gene expression regulation, which has been implicated in various cancers, including breast cancer, and may alter individual susceptibility to cancer. We investigated associations between miRNA related SNPs and breast cancer risk. First we evaluated 2,196 SNPs in a case-control study combining data from twelve genome wide association studies (GWAS). Second, we further investigated 42 SNPs with suggestive evidence for association using 41,785 cases and 41,880 controls from 41 studies included in the Breast Cancer Association Consortium (BCAC). Combining the GWAS and BCAC data within a meta-analysis, we estimated main effects on breast cancer risk as well as risks for estrogen receptor (ER) and age defined subgroups. Five miRNA binding site SNPs associated significantly with breast cancer risk: rs1045494 (odds ratio (OR) 0.92; 95% confidence interval (CI): 0.88–0.96), rs1052532 (OR 0.97; 95% CI: 0.95–0.99), rs10719 (OR 0.97; 95% CI: 0.94–0.99), rs4687554 (OR 0.97; 95% CI: 0.95–0.99), and rs3134615 (OR 1.03; 95% CI: 1.01–1.05) located in the 3’ UTR of CASP8, HDCC3, DROSHA, MUSTN1, and MYCL1, respectively. DROSHA belongs to miRNA machinery genes and has a central role in initial miRNA processing. The remaining genes are involved in different molecular functions, including apoptosis and gene expression regulation. Further studies are warranted to elucidate whether the miRNA binding site SNPs are the causative variants for the observed risk effects.


Data Availability: Authors confirm that, for approved reasons, some access restrictions apply to the data underlying the findings. Data are available via the Breast Cancer Association Consortium (BCAC) Data Access Coordination Committee (DACC) (http://ccgc.medschl.cam.ac.uk/consortia/bcac/). To request the data, readers may contact Manjeet Humphreys (mkh39@medschl.cam.ac.uk) or Douglas Easton (df20@medschl.cam.ac.uk).

Funding: Funding for the iCOGS infrastructure came from the European Community’s Seventh Framework Programme under grant agreement number 223175 (HEALTH-F2-2009-223175) (DCGS). iCOGS was also partly supported by the Canadian Institutes of Health Research for the “CIfHR Team in Familial Risks of Breast Cancer” program (US & DFE), and the Ministry of Economic Development, Innovation and Export Trade of Quebec – grant # PS04-08-711-01 & (US & DFE, P.Hall). HEMBCS was financially supported by the Helsinki University Central Hospital Research Fund, Academy of Finland (266528), the Finnish Cancer Society, The Nordic Cancer Union and the Sigrid Juselius Foundation. The population allele and genotype frequencies were obtained from the data source funded by the Nordic Center of Excellence in Disease Genetics based on samples regionally selected from Finland, Sweden and Denmark. The UK2 GWAS was funded by Wellcome Trust and Cancer Research UK. It included samples collected through the TRBC study which is funded by Cancer Research UK [C6020/A8372]. The WTCC2 was funded by the Wellcome Trust. The ABCCS and OFBCR studies were supported by the United States National Cancer Institute, National Institutes of Health (NIH) under RFA-CA-06-503 and through cooperative agreements with members of the Breast Cancer Family Registry (BCFR) and Principal Investigators, including Cancer Care Ontario (U01 CA69467), Northern California Cancer Center (U01 CA69417), University of Melbourne (U01 CA69638). Samples from the NC-BCFR were processed and distributed by the Coriell Institute for Medical Research. OFBCR was supported by the Canadian Institutes of Health Research for the “CIfHR Team in Familial Risks of Breast Cancer” program and grant UM1 CA164920 from the National Cancer Institute. 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. The ABCFS was also supported by the National Health and Medical Research Council of Australia, the New South Wales Cancer Council, the Victorian Health Promotion Foundation (Australia) and the Victorian Breast Cancer Research Consortium. JLH is a National Health and Medical Research Council (NHMRC) Australia Fellow and a Victorian Breast Cancer Research Consortium Group Leader. MCS is a NHMRC Senior Research Fellow and a Victorian Breast Cancer Research Consortium Group Leader. JLH and MCS are both group leaders of the Victoria Breast Cancer Research Consortium. The ABCS study was supported by the Dutch Cancer Society [grants NKI 2007-3839, 2009-4563], BMBB-NL, which is a Research Infrastructure financed by the Dutch government (NWO 184.021.007), and the Dutch National Genomics Initiative. The BRCs is funded by Cancer Research UK and Breakthrough Breast Cancer and acknowledges NHS funding to the NIHR Biomedical Research Centre, and the National Cancer Research Network (NCRN). The BCBS GWAS received funding from The Institut National de Cancer. The work of the BCBS was partly supported by ELAN-Fond of the University Hospital of Erlangen. ES (BIGGS) is supported by NHF Comprehensive Biomedical Research Centre, Guy’s & St. Thomas’ NHS Foundation Trust in partnership with King’s College London, United Kingdom. It is supported by the Oxford Biomedical Research Centre. The BSUCH study was supported by the Dietmar-Hopp Foundation, the Helmholtz Society and the German Cancer Research Center (DKFZ). The CECILE study was founded by Fondation de France (contract grant number 2004012618 and 2007005156), Institut National du Cancer (INCA) [2007-1/5P2C, 2008-1-C1-P4 and 2009-1-IH5/SP-04], Ligue Nationale contre le Cancer, Association pour la Recherche contre le Cancer (ARC) [2008-1-C1-P4], Agence Française de Sécurité Sanitaire de l’Environnement et du Travail [AFSSET - ANSES] [ST-2005-003, EST2008/1/26, and VS-2009-21], Ligue contre le Cancer Grand Ouest. The CGPS was supported by the Chief Physician Johan Bosunep and Lise Boserup Fund, the Danish Medical Research Council and Herlev Hospital. The CNIO-BCS was supported by the Genome Spain Foundation, the Red Temática de Investigación Cooperativa en Cancer and grants from the Asociación Española Contra el Cáncer and the Fondo de Investigación Sanitario (PI11/00923 and PI081120). We acknowledge the support of Alvarez Ivanov, Daniel Herrero, Primitiva Menendez and the Human Genotyping-CEGEN Unit (CNIO). The Human Genotyping-CEGEN Unit is supported by the Instituto de Salud Carlos III. The CTS was supported by the California Breast Cancer Act of 1993; National Institutes of Health (grants R01 CA77398 and the Lon V Smith Foundation [LV539420]); and the California Breast Cancer Research Fund (contract 97-10500). Collection of cancer incidence data used in this study 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. DEMORITOS is supported by a Hellenic Cooperator Ongroup Group research grant (HR R, RG/04) and the Greek General Secretary for Research and Technology (GSRT) Program, Research Excellence II, funded at 75% by the European Union. The DFBBCS GWAS was funded by The Netherlands Organisation for Scientific Research (NWO) as part of a ZonMW/VIDI grant number 91756341. The generation and management of GWAS genotype data for the Rotterdam Study is supported by the Netherlands Organisation of Scientific Research NWO Investments (nr. 175.010.2005.011, 911-03-012). This study is funded by the Research Institute for Diseases in the Elderly (014-93-015; RIDE2), the Netherlands Genomics Initiative (Netherlands Organisation for Scientific Research (NWO) project nr. 050-060-810. The Rotterdam Study is funded by Erasmus Medical Center and Erasmus University, Rotterdam, Netherlands Organization for the Health Research and Development (ZonMW), the Research Institute for Diseases in the Elderly (RIDE), the Ministry of Education, Culture, Science, the Ministry for Health, Welfare and Sports, the European Commission (DG XII), and the Municipality of Rotterdam. The ESTHER study was supported by a grant from the Baden-Württemberg Ministry of Science, Research and Arts. Additional cases were recruited in the context of the VEB0 study, which was supported by a grant from the German Cancer Aid (Deutsche Krebshilfe). The HMBCS was supported by a grant from the Friends of Hannover Medical School and by the Rudolf Bartling Foundation. The 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 and Bert von Kantzow foundation. The GC-HBOC was supported by Deutsche Krebshilfe [107054], the Dietmar-Hopp Foundation, the Helmholtz society and the German Cancer Research Centre (DKFZ). The GC-HBOC GWAS was...
supported by the Genetic Cancer Aid (grant no. 107352). The GENICA was funded by the Federal Ministry of Education and Research (BMBF) Germany grants 01KWK9975/5, 01KWK9976/8, 01KWK9977/0 and 01KWK0114, the Robert Bosch Foundation, Stuttgart, Deutsches Krebsforschungszentrum (DKFZ), Heidelberg, German Social Accident Insurance, Institute of the Ruhr University Bochum (IFP), Germany, as well as the Department of Internal Medicine, Evangelische Kliniken Bonn gGmbH, Johanniter Krankenhaus, Bonn, Germany. The KBCP was financially supported by the special 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. KConFab is supported by a grant from the 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. The KConFab Clinical Follow Up Study was funded by the NHMRC [145684, 288704, 454508]. Financial support for the AOC5 was provided by the United States Army Medical Research and Materiel Command (DAMD17-01-7-0028), the Cancer Council of Tasmania and Cancer Foundation of Western Australia and the NHMRC [196600]. GCT is supported by the NHMRC, LIMIC is supported by the 'Stichting tegen Kanker' [232-2008 and 196-2010], Dieter Lambrichts is supported by the FWO and the KULPVF10/10-016-SymbioSys. The MARIE study was supported by the Deutsche Krebshilfe e.V. [70-2892-BB 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 [01KHO402]. MBC5a is supported by grants from the Italian Association for Cancer Research (AIRC) and by funds from the Italian citizens who allocated the $1/1000 share of their tax payment in support of the Fondazione IRCCS Istituto Nazionale Tumori, according to the Italian laws (INT-Institutional strategic projects “S x 1000”). The MCBCs was supported by the NIH grants [CA122340, CA128978] and a Specialized Program of Research Excellence (SPRE) in Breast Cancer [CA116201], 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. MCCS cohort recruitment was funded by VicHealth and Cancer Council Victoria. The MCCS was further supported by Australian NHMRC grants 290077, 251553 and 504711 and by infrastructure provided by Cancer Council Victoria. The MEC was supported by NIH grants CA63464, CA54281, CA98678 and CA132839. For the MTLGEBCS study, the initial case–control study was supported by the Canadian Breast Cancer Research Initiative. Work was also supported by the Quebec Breast Cancer Foundation, the Canadian Institutes of Health Research for the “CIHR Team in Familial Risks of Breast Cancer” program – grant # FRN-7571 and the Ministry of Economic Development, Innovation and Export Trade – grant # PSR-SIRI-701. The NBCS was supported by grants from the Norwegian Research council, 155218/V40, 175240/510 to ALBD, FUGE-NFRI 186100/V1 to VNK and a Swiss Bridge Award to ALBD. The OBCS was supported by research grants from the Finnish Cancer Academy, the Academy of Finland, the University of Oulu, and the Oulu University Hospital. The ORIGO study was supported by the Dutch Society for Science (RUL 1997-1505) and the Biobanking and Biomolecular Resources Research Infrastructure (BBMRI-NL CPL16). The OSU study was funded by the Stefanie Spielman fund and the OSU Comprehensive Cancer Center. The PBCS was funded by Intramural Research Funds of the National Cancer Institute, Department of Health and Human Services, USA. The gKARMA study was supported by Marit and Hans Rausings Initiative Against Breast Cancer and Breast Cancer Risk Prediction Center, a Unneus Centre (contract 70867902) financed by the Swedish Research Council. The RBCS was funded by the Dutch Society Cancer (DDHK 2004-3124, DDHK 2009-4318). The RPCCI study was supported by RPCCI DataBank and BioRepository (DBBR), a Cancer Center Support Grant Shared Resource (P30 CA016562-32). The SASBAC study 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. The SBCS was supported by funding from Yorkshire Cancer Research S295, S299, 33050.PA. SEARCH is funded by a programme grant from Cancer Research UK [C490/A10124] and supported by the the UK National Institute for Health Research Biomedical Research Centre at the University of Cambridge. AMD has been supported by Cancer Research UK grant [C8197/AB065] and by the Joseph Mitchell Fund. SKKDF25 is supported by the DKFZ. The SzBCS was supported by Grant PBK, I22, 2705/2004; Katarzyna Jaworska is a fellow of International PhD program, Postgraduate School of Molecular Medicine, Warsaw Medical University, supported by the Polish Foundation of Science. The TNBCC was supported by the NIH grant (CA128978), the Breast Cancer Research Foundation, Komen Foundation for the Cure, the Ohio State University Comprehensive Cancer Center, the Stefanie Spielman fund for Breast Cancer Research and a generous gift from the David F. and Margaret T. Grohne Family Foundation and the Ting Tsung and Wei Fong Chao Foundation. Part of the TNBCC project has been co-financed by the Greek national funds through the Operational Program “Education and Lifelong Learning” of the National Strategic Reference Framework (NSRF) - Research Funding Program of the General Secretariat for Research & Technology: ARISTEIA. The UKBGS is funded by the Breakthrough Breast Cancer and the Institute of Cancer Research (ICR). ICR acknowledges NHS funding to the NIHR Biomedical Research Centre. CGEMS, The Nurses’ Health Studies are supported by NIH grants CA 65725, CA87969, CA49449, CA67262, CA50385 and SUOIC0A98233. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* Email: heli.nevanlinna@hus.fi

† Membership of the GENICA Network, KConFab Investigators, and AOC5 is provided in the Acknowledgements.

Introduction

Breast cancer is the most common women’s cancer and is a leading cause of cancer mortality [1]. Inherited genetic variation has been associated with the initiation, development and progression of breast cancer. Studies on twins have suggested that hereditary predisposing factors are involved in up to one third of all breast cancers [2]. Many genetic loci have been associated with breast cancer risk and collectively explain approximately 35% of the familial risk [3-4]. The largest genetic association study of breast cancer to date identified 41 novel low penetrance susceptibility loci [4] by selecting nearly 30,000 SNPs from a meta-analysis of nine genome-wide association (GWA) studies and genotyping them using 41,785 cases and 41,880 controls of European ancestry from studies in the Breast Cancer Association Consortium (BCAC). These 41 susceptibility loci probably represent the tip of the iceberg, and additional SNPs from the combined GWAS might explain a similar fraction of familial risk to that attributed to the already identified loci [4].

Mature miRNAs are 20–23 nucleotide, single-stranded RNA molecules that play a crucial role in gene expression regulation for many cellular processes including differentiation potential and development pattern. miRNAs undergo a stepwise maturation process involving an array of RNA machinery components. Drosha and DGCGR8 mediate the cleavage of long primary miRNA transcripts (pri-miRNAs) into shorter pre-miRNAs in the nucleus [5,6]. The pre-miRNAs are then transported to the cytoplasm where they are further cleaved by Dicer to produce mature miRNAs [7]. MiRNAs interact by pairing with the 3’ untranslated region (UTR), and also within the coding region and 5’ UTR of the corresponding mRNAs leading to mRNA destabilization, cleavage or translation repression. More effective mRNA destabilization is achieved when miRNA targets the 5’ UTR rather than other mRNA regions [8–10]. An individual miRNA may regulate approximately 100 distinct mRNAs, and together more than 1000 human miRNAs are believed to modulate more than half of the mRNA species encoded in the genome [11,12]. Additionally, most miRNAs possess binding sites for miRNAs [13]. MiRNAs are involved in tumorigenesis in that they can be either oncogenic when tumor suppressor genes are targeted, or genomic guardians (tumor suppressor miRNAs) when oncogenes are targeted [14]. Additionally it has been suggested that they may modulate both metastasis [15] and chemotherapy resistance [16]. MiRNAs have also been shown to have altered expression levels in tumours compared to normal tissue and between tumor subtypes in breast cancer among other carcinoma types [17–19]. SNPs may affect miRNA binding sites in their binding regions. Polymorphisms in miRNA binding sites have been studied in regard to the risk of several cancers [20], including breast cancer [21–23]. These studies have found evidence for association of miRNA related SNPs and cancer risk, but the study sample sizes have been relatively small.

In this study, we investigate associations between miRNA-related polymorphisms and breast cancer risk by using a meta-
Materials and Methods

SNP selection and genotyping

SNPs in mature or pre-miRNAs, in genes of the miRNA machinery and in 3’UTR regions of protein coding genes with a potential effect on miRNA binding were systematically searched from Ensembl (hg18/build36) and Patrocles databases [24]. Additionally, tagging SNPs for each with $r^2 \geq 0.8$ were also identified utilizing the public HapMap SNP database. By this in silico approach we identified altogether 147,801 candidate SNPs and 12,550 tagging SNPs. These SNPs were then overlayed with those from the combined GWAS from the BCAC [4] and altogether 2196 SNPs were present (either genotyped or imputed) in the combined GWAS. These SNPs were genotyped with Illumina or Affymetrix arrays, as described previously [25–32]. The combined GWAS data were imputed for all scans using HapMap version 2 CEU as a reference in similar fashion to that presented by Michailidou and colleagues [4].

The genotyped DROSHA SNP tags the 3’ UTR miRNA binding site variant in the DROSHA gene. The remaining 38 candidate or tag SNPs were located in or tagged to a predicted miRNA binding site in the 3’ UTR of protein coding genes. All 42 SNPs are described in Table 1. The workflow of the SNP selection in different stages is illustrated in Figure 1.

Study sample

The combined GWAS included nine breast cancer studies totalling 10,052 cases and 12,575 controls of European ethnic background. Details and study-specific subject numbers are presented in Table S1. Since the GWAS were limited to patients of European ethnic background we further utilized 41,785 cases ascertained for their first primary, invasive breast cancer and 41,890 controls of European ancestry from 41 BCAC studies genotyped using the iCOGS array (Table S2). For a subgroup analysis of ER negative and ER positive cases, as well as cases aged less than 50 years at diagnosis, we included all the cases for which the respective data were available. The ER subgroup analysis was based on 702 ER negative cases and 2,019 ER positive cases from five GWAS studies and 7,200 ER negative cases from 40 BCAC studies and 26,302 ER positive cases from 34 BCAC studies.

The analysis of cases aged less than 50 years at diagnosis was based on 3,470 cases from three GWAS studies and 9,483 cases from 35 BCAC studies. All participating studies conform to the Declaration of Helsinki and were approved by the respective ethical review boards and ethics committees (Tables S1 and S2), and all participants in these studies had provided written consent for the research.

Statistical methods

We used logistic regression to estimate per-allele log-odds ratios and standard errors including the study as a covariate. We also included principal components as covariates in order to correct for potential hidden population structure. In the GWAS, for two studies (UK2 and HEBCS) the estimates were adjusted for the first three principal components and in the iCOGS analysis we used the first six principal components and an additional component to reduce inflation for the LMBC study, as described previously [4].

For meta-analysis, we combined the estimates from the combined GWAS and iCOGS with a fixed effects model using the inverse variance weighted method. In the meta-analysis, the subjects involved in both combined GWAS and iCOGS (1880) were only taken into account once. In order to adjust for $P$-values against multiple testing, we used Benjamini Hochberg correction. The adjusted $P$-values are shown in Table 2 along with the nominal $P$-values. In the text we report the nominal $P$-values. The statistical analyses were conducted using the R 2.14.0 statistical computing environment (http://www.r-project.org/).

Results

For the 42 SNPs we successfully genotyped, estimates of association from the combined GWAS and from iCOGS analysis are shown in Table S3. Twenty-one SNPs showed consistent
Table 1. The 42 studied SNPs in miRNAs, miRNA machinery genes and miRNA target genes.

<table>
<thead>
<tr>
<th>Functional SNP (Tag SNP, R-squared)</th>
<th>Chr</th>
<th>Position</th>
<th>Coding</th>
<th>Gene</th>
<th>miRNA</th>
<th>SNP effect*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Located within miRNA</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>rs17301408</td>
<td>10</td>
<td>115923895</td>
<td>GA</td>
<td>hsa-miR-2110</td>
<td></td>
<td></td>
</tr>
<tr>
<td>rs13447640 (rs1805360, r^2 = 1)</td>
<td>11</td>
<td>93866677</td>
<td>GA</td>
<td>hsa-miR-548l</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Located in miRNA biogenesis machinery genes</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>rs3764941</td>
<td>5</td>
<td>135497426</td>
<td>AC</td>
<td>SMADS</td>
<td></td>
<td></td>
</tr>
<tr>
<td>rs17151639</td>
<td>7</td>
<td>127425052</td>
<td>AG</td>
<td>SND1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>rs174060616</td>
<td>7</td>
<td>134773600</td>
<td>CG</td>
<td>CNOT4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>rs10719</td>
<td>5</td>
<td>31437204</td>
<td>GA</td>
<td>DROSHA</td>
<td>hsa-miR-1298</td>
<td>AC</td>
</tr>
<tr>
<td>Located in miRNA target genes</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>rs2550303</td>
<td>16</td>
<td>54953111</td>
<td>AG</td>
<td>AMFR</td>
<td>hsa-miR-577</td>
<td>AC</td>
</tr>
<tr>
<td>rs7513934</td>
<td>1</td>
<td>52590776</td>
<td>GA</td>
<td>CC2D1B</td>
<td>hsa-miR-384/hsa-miR-577</td>
<td>CNC</td>
</tr>
<tr>
<td>rs1128226</td>
<td>7</td>
<td>21908194</td>
<td>AC</td>
<td>CDA7L</td>
<td>hsa-miR-548g</td>
<td>AC</td>
</tr>
<tr>
<td>rs3796133</td>
<td>3</td>
<td>10000533</td>
<td>GA</td>
<td>DCBLD2</td>
<td>hsa-miR-624*</td>
<td>AC</td>
</tr>
<tr>
<td>rs7441</td>
<td>12</td>
<td>90063806</td>
<td>GA</td>
<td>DCN</td>
<td>hsa-miR-135b*</td>
<td>AC</td>
</tr>
<tr>
<td>rs1803439</td>
<td>21</td>
<td>37807312</td>
<td>AG</td>
<td>Dyrk1A</td>
<td>hsa-miR-550</td>
<td>AC</td>
</tr>
<tr>
<td>rs3797</td>
<td>15</td>
<td>2719858</td>
<td>AG</td>
<td>FAM189A1</td>
<td>hsa-miR-570</td>
<td>AC</td>
</tr>
<tr>
<td>rs7130622</td>
<td>11</td>
<td>128186721</td>
<td>AC</td>
<td>FLI1</td>
<td>hsa-miR-138-2*</td>
<td>AC</td>
</tr>
<tr>
<td>rs1052532</td>
<td>15</td>
<td>89275240</td>
<td>AG</td>
<td>HSCDC3</td>
<td>hsa-miR-1224-3p/hsa-miR-1260/hsa-miR-1280</td>
<td>AC</td>
</tr>
<tr>
<td>rs7040123</td>
<td>9</td>
<td>7160742</td>
<td>AG</td>
<td>KDM4C</td>
<td>hsa-miR-154*/hsa-miR-487a</td>
<td>AC</td>
</tr>
<tr>
<td>rs1062225</td>
<td>10</td>
<td>49513232</td>
<td>AG</td>
<td>MAPK8</td>
<td>hsa-miR-203</td>
<td>AC</td>
</tr>
<tr>
<td>rs41739</td>
<td>7</td>
<td>116224740</td>
<td>AG</td>
<td>MET</td>
<td>hsa-miR-578-5p</td>
<td>AC</td>
</tr>
<tr>
<td>rs702681</td>
<td>5</td>
<td>56253786</td>
<td>AG</td>
<td>MIER3</td>
<td>hsa-miR-196a*</td>
<td>AC</td>
</tr>
<tr>
<td>rs3134615</td>
<td>1</td>
<td>40134653</td>
<td>CA</td>
<td>MYCL1</td>
<td>hsa-miR-1827</td>
<td>AC</td>
</tr>
<tr>
<td>rs2304669</td>
<td>2</td>
<td>238830402</td>
<td>AG</td>
<td>PER2</td>
<td>hsa-miR-485-3p</td>
<td>AC</td>
</tr>
<tr>
<td>rs13422</td>
<td>17</td>
<td>15074900</td>
<td>AC</td>
<td>PMP22</td>
<td>hsa-miR-129b-1*</td>
<td>AC</td>
</tr>
<tr>
<td>rs7562391</td>
<td>2</td>
<td>20144441</td>
<td>AC</td>
<td>PPL3</td>
<td>hsa-miR-4937/hsa-miR-499-3p</td>
<td>AC</td>
</tr>
<tr>
<td>rs7520333</td>
<td>1</td>
<td>40862837</td>
<td>AG</td>
<td>RIMS3</td>
<td>hsa-let-7d/hsa-let-7e</td>
<td>CNC</td>
</tr>
<tr>
<td>rs739962</td>
<td>18</td>
<td>53717824</td>
<td>GA</td>
<td>ST8A3</td>
<td>hsa-miR-96/hsa-miR-1271/hsa-miR-182</td>
<td>AC</td>
</tr>
<tr>
<td>rs1058450</td>
<td>4</td>
<td>120200088</td>
<td>GA</td>
<td>SYNPO2</td>
<td>hsa-miR-183</td>
<td>AC</td>
</tr>
<tr>
<td>rs4351800</td>
<td>11</td>
<td>7446395</td>
<td>CA</td>
<td>SYT9</td>
<td>hsa-miR-544</td>
<td>AC</td>
</tr>
<tr>
<td>rs12438324</td>
<td>15</td>
<td>55366808</td>
<td>AG</td>
<td>TCF12</td>
<td>hsa-miR-591</td>
<td>AC</td>
</tr>
<tr>
<td>rs1269870</td>
<td>13</td>
<td>99415306</td>
<td>AG</td>
<td>ZIC5</td>
<td>hsa-miR-34a/hsa-miR-34c-5p/hsa-miR-449a/hsa-miR-449b</td>
<td>AC</td>
</tr>
<tr>
<td>rs9990 (rs1444448, r^2 = 1)</td>
<td>10</td>
<td>64230476</td>
<td>AG</td>
<td>ADO</td>
<td>hsa-miR-512-5p/hsa-miR-510</td>
<td>AC</td>
</tr>
<tr>
<td>rs757537 (rs4705870, r^2 = 1)</td>
<td>5</td>
<td>13218703</td>
<td>GA</td>
<td>ANKRD43</td>
<td>hsa-miR-320a/hsa-miR-320b/hsa-miR-320c/hsa-miR-320d</td>
<td>AC</td>
</tr>
<tr>
<td>rs3774729 (rs2037119, r^2 = 0.943)</td>
<td>3</td>
<td>63969919</td>
<td>GA</td>
<td>ATXN7</td>
<td>hsa-miR-1206</td>
<td>AC</td>
</tr>
</tbody>
</table>
Table 1. Cont.

<table>
<thead>
<tr>
<th>SNP (Tag SNP, R² = 1)</th>
<th>Chr (Position)</th>
<th>Coding</th>
<th>Gene</th>
<th>miRNA (miRNA binding site SNP, R² ≥ 0.96)</th>
</tr>
</thead>
</table>
| rs702681 (rs1045494, 1.0) | 6 (15647231) | AG | MYCL1 | hsa-miR-938 binding site SNP rs1045494, (OR 0.92 [95%CI 0.88–0.96]; $P = 5.90 \times 10^{-4}$); rs1052532, (OR 0.97 [95%CI 0.95–0.99]; $P = 7.78 \times 10^{-6}$); rs10719, (OR 0.97 [95%CI 0.94–0.99]; $P = 1.35 \times 10^{-5}$); rs3134615 (OR 0.97 [95%CI 0.95–0.99]; $P = 1.71 \times 10^{-5}$); rs3134615 (OR 1.03 [95%CI 1.01–1.05]; $P = 5.07 \times 10^{-5}$) located in 3' UTR of Caspase-8 (CASP8), HD Domain Containing 3 (HDDC3), DROSHA, Musculoskeletal, Embryonic Nuclear Protein 1 (MUSTN1) and V-Myelocytomatosis Viral Oncogene Homolog 1 (MYCL1), respectively. (Table 2). SNP rs1045494 is tagging the hsa-miR-938 binding site SNP rs1045494 (r² = 1.0) of CASP8 and the SNP rs1052532 in HDDC3 is predicted to abolish the binding site for hsa-miR-1224. The SNP rs10719 is predicted to abolish the hsa-miR-1290 binding site in the 3' UTR of DROSHA. SNP rs4607554 tags the hsa-miR-99a binding site SNP rs645330 (r² = 1.0) of MUSTN1 and rs3134615 is located at the binding site of hsa-miR-1027 of MYCL1. There was no evidence for heterogeneity in the per-allele OR for any SNP. The per study per allele ORs for these five miRNA binding site SNPs from the combined GWAS along with per-SNP heterogeneity variance $P$-values are shown in Figure S1 and from the iCOGS in Figure S2. Next we analysed the SNPs by ER status-defined subtype, and for cases aged less than 50 years at diagnosis, for risk associations in the meta-analysis of combined GWAS and iCOGS (Tables S4, S5 and S6). These analyses did not reveal any additional significant results. For rs1045494 in CASP8, a second SNP, rs4607554 in MUSTN1 and rs3134615 in MYCL1 (OR 1.03 [95%CI 1.01–1.05]; $P = 7.75 \times 10^{-4}$), more significant association with breast cancer risk was found for the ER positive subgroup than in the main analysis, but the result from the test for heterogeneity by ER status was not significant (data not shown). All associations were estimated using an additive inheritance model. Dominant and recessive models did not improve the estimates (data not shown).

Discussion

We investigated associations between genetic variation in miRNAs, in the genes of the miRNA machinery and in the miRNA binding sites and the risk of breast cancer. We identified several SNPs that are predicted to abolish an miRNA binding site and that are significantly associated with breast cancer risk. Previous studies investigating miRNA related SNPs, especially in miRNA binding sites have included predefined sets of genes. Nicoloso and colleagues investigated 38 previously identified breast cancer risk SNPs and found two to modify miRNA binding sites in TGFB1 and XRCC1 in vitro [23]. Neither of these were included in our data set. Liang and colleagues investigated 134 potential miRNA binding sites in cancer-related genes and found six miRNA binding site SNPs that were associated with ovarian cancer risk [34].
### Table 2. Associations of SNPs in the GWAS and iCOGS separately and combined GWAS + iCOGS and breast cancer risk.

<table>
<thead>
<tr>
<th>SNP</th>
<th>Chr</th>
<th>Position</th>
<th>coding</th>
<th>GWAS OR (95% CI)</th>
<th>iCOGS OR (95% CI)</th>
<th>Combined GWAS + iCOGS OR (95% CI)</th>
<th>Combined GWAS + iCOGS P^2(BH corrected)</th>
<th>Gene</th>
</tr>
</thead>
<tbody>
<tr>
<td>rs702681</td>
<td>5</td>
<td>56253786</td>
<td>AG</td>
<td>1.07 (1.02–1.11)</td>
<td>3.92×10⁻⁵</td>
<td>1.06 (1.04–1.09)</td>
<td>2.76×10⁻⁹</td>
<td>MIER3</td>
</tr>
<tr>
<td>rs1045494</td>
<td>2</td>
<td>201860026</td>
<td>AG</td>
<td>0.90 (0.81–1.00)</td>
<td>4.74×10⁻²</td>
<td>0.92 (0.88–0.96)</td>
<td>4.47×10⁻⁹</td>
<td>CASP8</td>
</tr>
<tr>
<td>rs1052532</td>
<td>15</td>
<td>89275240</td>
<td>AG</td>
<td>0.94 (0.90–0.98)</td>
<td>7.94×10⁻³</td>
<td>0.97 (0.95–0.99)</td>
<td>1.47×10⁻²</td>
<td>HDDC3</td>
</tr>
<tr>
<td>rs10719</td>
<td>5</td>
<td>31437204</td>
<td>GA</td>
<td>0.92 (0.88–0.97)</td>
<td>8.79×10⁻⁴</td>
<td>0.98 (0.95–1.00)</td>
<td>5.32×10⁻²</td>
<td>DROSHA</td>
</tr>
<tr>
<td>rs4687554</td>
<td>3</td>
<td>52839175</td>
<td>AG</td>
<td>0.94 (0.90–0.99)</td>
<td>1.23×10⁻²</td>
<td>0.97 (0.95–1.00)</td>
<td>2.39×10⁻²</td>
<td>MLI1</td>
</tr>
<tr>
<td>rs1334615</td>
<td>1</td>
<td>40134653</td>
<td>CA</td>
<td>1.04 (0.99–1.09)</td>
<td>9.97×10⁻²</td>
<td>1.03 (1.00–1.05)</td>
<td>2.09×10⁻²</td>
<td>MYC1</td>
</tr>
<tr>
<td>rs7635553</td>
<td>3</td>
<td>168646064</td>
<td>GA</td>
<td>0.89 (0.83–0.95)</td>
<td>9.73×10⁻⁴</td>
<td>0.98 (0.95–1.01)</td>
<td>1.98×10⁻¹</td>
<td>SERPIN2</td>
</tr>
<tr>
<td>rs3796133</td>
<td>3</td>
<td>10000533</td>
<td>GA</td>
<td>1.18 (1.08–1.29)</td>
<td>4.18×10⁻⁴</td>
<td>1.01 (0.97–1.06)</td>
<td>5.74×10⁻⁵</td>
<td>DCBLD2</td>
</tr>
<tr>
<td>rs4351800</td>
<td>11</td>
<td>7446395</td>
<td>CA</td>
<td>1.06 (1.00–1.08)</td>
<td>4.48×10⁻²</td>
<td>1.01 (0.99–1.03)</td>
<td>1.98×10⁻¹</td>
<td>SYT9</td>
</tr>
<tr>
<td>rs17512204</td>
<td>2</td>
<td>118449301</td>
<td>GA</td>
<td>1.06 (0.98–1.14)</td>
<td>1.20×10⁻¹</td>
<td>1.03 (0.99–1.06)</td>
<td>1.63×10⁻¹</td>
<td>CCDC93</td>
</tr>
<tr>
<td>rs3809828</td>
<td>17</td>
<td>7187575</td>
<td>GA</td>
<td>1.17 (1.06–1.28)</td>
<td>1.97×10⁻³</td>
<td>1.01 (0.97–1.05)</td>
<td>5.22×10⁻¹</td>
<td>KCTD1</td>
</tr>
<tr>
<td>rs7441</td>
<td>12</td>
<td>90063806</td>
<td>GA</td>
<td>1.11 (1.03–1.20)</td>
<td>8.70×10⁻³</td>
<td>1.01 (0.97–1.05)</td>
<td>5.98×10⁻¹</td>
<td>DCRN</td>
</tr>
<tr>
<td>rs7086917</td>
<td>10</td>
<td>49667441</td>
<td>AC</td>
<td>0.96 (0.93–1.00)</td>
<td>6.35×10⁻²</td>
<td>0.99 (0.97–1.01)</td>
<td>4.38×10⁻¹</td>
<td>WDFY4</td>
</tr>
<tr>
<td>rs7040123</td>
<td>9</td>
<td>7160742</td>
<td>AG</td>
<td>1.11 (0.99–1.23)</td>
<td>7.59×10⁻²</td>
<td>1.02 (0.97–1.07)</td>
<td>5.14×10⁻⁵</td>
<td>KDM4C</td>
</tr>
<tr>
<td>rs7674744</td>
<td>4</td>
<td>78874296</td>
<td>GA</td>
<td>0.94 (0.89–0.99)</td>
<td>2.83×10⁻²</td>
<td>0.99 (0.97–1.02)</td>
<td>6.91×10⁻¹</td>
<td>CNOT6L</td>
</tr>
<tr>
<td>rs12438324</td>
<td>15</td>
<td>55366808</td>
<td>AG</td>
<td>0.87 (0.79–0.97)</td>
<td>1.01×10⁻²</td>
<td>1.00 (0.94–1.05)</td>
<td>8.69×10⁻⁹</td>
<td>TCF12</td>
</tr>
<tr>
<td>rs17151639</td>
<td>7</td>
<td>127425052</td>
<td>AG</td>
<td>0.96 (0.92–1.01)</td>
<td>1.09×10⁻¹</td>
<td>0.99 (0.97–1.02)</td>
<td>5.66×10⁻⁵</td>
<td>SND1</td>
</tr>
<tr>
<td>rs17406169</td>
<td>7</td>
<td>134736000</td>
<td>CG</td>
<td>0.87 (0.72–1.04)</td>
<td>1.27×10⁻¹</td>
<td>0.99 (0.93–1.04)</td>
<td>6.39×10⁻¹</td>
<td>CNOT4</td>
</tr>
<tr>
<td>rs7513934</td>
<td>1</td>
<td>52590776</td>
<td>GA</td>
<td>1.04 (1.00–1.08)</td>
<td>7.98×10⁻²</td>
<td>1.00 (0.98–1.02)</td>
<td>9.99×10⁻¹</td>
<td>CCDC18</td>
</tr>
<tr>
<td>rs2304669</td>
<td>2</td>
<td>238830420</td>
<td>GA</td>
<td>0.96 (0.91–1.02)</td>
<td>1.86×10⁻¹</td>
<td>1.00 (0.97–1.02)</td>
<td>8.17×10⁻¹</td>
<td>PER2</td>
</tr>
<tr>
<td>rs1058450</td>
<td>4</td>
<td>120200888</td>
<td>GA</td>
<td>0.96 (0.91–1.01)</td>
<td>1.33×10⁻¹</td>
<td>1.00 (0.97–1.02)</td>
<td>9.28×10⁻¹</td>
<td>SYNPO2</td>
</tr>
</tbody>
</table>

The SNPs with consistent odds ratios in combined GWAS and iCOGS analysis are shown. (Results for all 42 SNPs are presented in Table S3.)

1. Build 36 position.
2. Per allele odds ratio for the minor allele relative to the major allele.
3. Adjusted p-trend.
4. Adjusted p-trend adjusted against multiple testing by Benjamini-Hochberg correction method.

doi:10.1371/journal.pone.0109973.t002
In the meta-analysis of combined GWAS and iCOGS for main effects, for four of the five most significant miRNA binding site SNPs, the minor allele was associated with a decreased breast cancer risk. The minor allele of SNP rs3134615 in 3' UTR of MYC/L was associated with an increased breast cancer risk. All the five most significant miRNA binding site SNPs locate in 3' UTR and have been predicted to abolish the miRNA binding site. The defect in miRNA-mediated regulation would be expected to lead to an increase in the translation of the corresponding encoded protein. The five genes, whose regulation may be affected by the miRNA-associated SNPs, include the pre-apoptotic gene CASP8, HDEC3, miRNA biogenesis master regulator DROSHA, MYC-family member MYC/L and MUSTN1. CASP8 is involved in apoptosis in breast cancer cells [35], and many studies have reported polymorphisms in this gene to be associated with risks for several cancers [36,37] including breast cancer [38,39], indicating the importance of CASP8 in tumor development. SNP rs1045494 studied here is located close to the coding region SNP rs1045485 that has been previously shown to have a stronger protective effect [38,40,41]. Interestingly, Michalidou and colleagues reported this SNP as having only weak evidence for an association (P 0.0013 in combined GWAS and iCOGS) [4], but these two SNPs (rs1045485 and rs1045494) are not correlated (r² = 0.01 in Caucasian population). Neither is rs1045494 correlated with the more strongly associated rs1830298 SNP, identified through fine-mapping of the region (r² = 0.02) [42]. Rs1045494 tags SNP rs1045487 (r² = 1.0) which is predicted to abolish the hsa-miR-938 binding site and thus may affect CASP8 expression. There is very little reported evidence on the involvement of HDDC3 or the hsa-miR-1224-3p in cancer, indicating a novel association with risk. HDDC3 has been suggested to be involved in the starvation response [43]. The HDDC3 gene is expressed at higher levels by several different tumor types, including breast tumors, than by normal tissue [44]. DROSHA is a miRNA master regulator. It is a member of the RNase III enzyme family, belongs to the miRNA biogenesis pathway and is the core nuclease that processes pri-miRNAs into pre-miRNAs in the nucleus [5,6]. The SNP rs10719 in the 3' UTR of DROSHA is predicted to abolish the hsa-miR-1298 binding site. Hsa-miR-1298 is predicted to target DROSHA by the Patrocles prediction as well as by TargetScan [45] and PITA [46] prediction algorithms. Recently a small Korean study reported another SNP rs644236, tagging the SNP rs10719 (r² = 0.955 in CEU population and r² = 0.876 in Asian population (combined CHB and JPT)) to be associated with elevated breast cancer risk [47]. When taking into account the opposite major and minors alleles in the Asian and European populations for SNPs rs644236 and rs10719, this result is in concordance with our results where both the combined GWAS as well as the iCOGS analysis consistently indicated an association of the minor allele of SNP rs10719 with reduced breast cancer risk. We also found the minor allele of SNP rs3134615 in the 3' UTR of MYC/L to be associated with an increased risk. MYC/L (L-MYC) belongs to the same family of transcription factors as the known proto-oncogene MYC (C-MYC) and they share a high degree of structural similarity [48]. The MYC/L gene has previously been reported to be amplified and overexpressed in ovarian cancer [49]. A case-control study by Xiong and colleagues reported SNP rs3134615 to be significantly associated with increased risk of small cell lung cancer [50]. SNP rs3134615 was predicted by Patrocles to abolish the hsa-miR-1827 binding site. This has also been suggested by functional studies where MYC/L was found as the target of hsa-miR-1827 and the SNP rs3134615 was also found to increase MYC/L expression [50]. The evidence from functional studies is consistent with our finding that SNP rs3134615 might increase breast cancer risk. MUSTN1 has been shown to be involved in the development and regeneration of the musculoskeletal system [51]. Thus far no evidence of association between MUSTN1 and breast cancer has been reported, but the MUSTN1 gene is expressed in the mammary glands [52].

Since only a small fraction of miRNA binding sites has been experimentally validated, we selected SNPs that had been computationally predicted to affect miRNA binding sites. For our original SNP selection we used the Patrocles database that contains predicted miRNA binding sites and also compiles perturbation prediction of SNP effects. There are a multitude of prediction programs and their performance has been evaluated [53]. Witkos and colleagues find target prediction algorithms that utilize orthologous sequence alignment, like Patrocles, to be the most reliable.

The followup of the 42 miRNA related SNPs identified five significant associations with breast cancer risk. Although the individual risk effects were subtle, considering that we could only investigate a small proportion of our initial in silico data set of miRNA related SNPs (over 140,000 SNPs) this may suggest that genetic polymorphisms affecting the miRNA regulation could have a considerable combined effect on breast cancer risk.

It should be noted that, until fine mapping studies are carried out for these loci, it is not clear whether these miRNA-related SNPs are the variants responsible for the observed associations.

This comprehensive analysis of miRNA related polymorphisms using a large two stage study of women with European ancestry provides evidence for miRNA related SNPs being potential modulators of breast cancer risk.

Supporting Information

Figure S1 Forest plots for the five most significant miRNA binding site SNPs from the combined GWAS. Squares indicate the estimated per-allele OR for the minor allele in Europeans. The horizontal lines indicate 95% confidence limits. The vertical blue dashed lines indicate clipping of the confidence intervals for presentation purpose. The area of the square is inversely proportional to the variance of the estimate. The diamond indicates the estimated per-allele OR from the combined analysis.

(PDF)

Figure S2 Forest plots for the five most significant miRNA binding site SNPs from the iCOGS. Squares indicate the estimated per-allele OR for the minor allele in Europeans. The horizontal lines indicate 95% confidence limits. The vertical blue dashed lines indicate clipping of the confidence intervals for presentation purpose. The area of the square is inversely proportional to the variance of the estimate. The diamond indicates the estimated per-allele OR from the combined analysis.

(PDF)

Table S1 A description of each GWAS study, number of subjects and genotyping platform used in combined GWAS.

(DOC)

Table S2 A description of each BCAC study with subjects of European origin in iCOGS.

(DOC)

Table S3 Frequencies and effect sizes of the 42 SNPs in the main analysis; combined GWAS and iCOGS.

(DOC)
Table S4 Results for SNPs in the GWAS and iCOGS separately and combined GWAS+iCOGS analysis for ER negative subgroup.

(DOC)

Table S5 Results for SNPs in the GWAS and iCOGS separately and combined GWAS+iCOGS analysis for ER positive subgroup.

(DOC)

Table S6 Results for SNPs in the GWAS and iCOGS separately and combined GWAS+iCOGS analysis for cases less than 50 years at diagnosis.

(DOC)

Acknowledgments

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. The **HEBCS** thanks Dr. Karl von Smitten and KN Irja Erkal for their help with the HEBCS data and samples. The **ABCFS** thanks Maggie Angelakos, Judi Maskell and Gillian Dite. The **OFBCR** thanks Teresa Selander, Nayana Weerasooriya and Gerd Glendon. The **ABCFS** would like to acknowledge Ellen van der Schoot for DNA of controls. The **BBCC** thanks Silke Landrath, Sonja Oesser, Matthias Rubiner. The **BBCC** thanks Eileen Williams, Elaine Ryder-Mills and Kara Sargus. The **BIGGS** thanks Niall McNerney, Gabrielle Collemer, Andrew Rowan and Angela Jones. The **BSUCH** thanks Peter Bugert and the Medical Faculty, Mainzheim. The **BIGGS**, **HEBCS**, **BBCC**, **BCS**, **OFBCR** and **ABCFS** would like to thank the following persons for providing additional information and samples: Prof. Dr. Norbert Arnold, Dr. Sabine Preissler-Adams, Dr. Monika Mareeva-Varon, Dr. Dieter Niederacher, Prof. Dr. Brigitte Schlegelberger, Dr. Clemens Muller, Heide Hellebrand, and Stefanie Peissel and Daniela Zaffaroni of the Fondazione IRCCS Istituto Nazionale dei Tumori (I.N.T.); Bernardo Bonanni, Irene Feroco and Angela Peissel; Doctor Stephane Nieuwlaat, Annette Heemskerk and the Erasmus MC Family Cancer Clinic. The **SBCS** thanks Sue Higham, Ian Brock, Sabapathy Balasubramanian, Helen Cramp and Dan Connelly. The **SEARCH** thanks the SEARCH and EPI-Cancer Nordzelt research groups. The **ABCFS** study would not have been possible without the contributions of the following: Qin Wang (BCAC), Andrew Berchuck (OCAC), Rosalind A. Eles, Ali Amin Al Olama, Zofia Kote-Jarai, Sara Benlloch (PRACTICAL), Antonia Antoniou, Lesley McGuigof and Ken Offit (CIAMBA), Andrew Lee, and Ed Dicks, Craig Lazzarini and the staff of the Centre for Genetic Epidemiology Laboratory, Anna Gonzalez-Neira and the staff of the CNIO genotyping unit, Daniel C. Tessler, Francois Bacot, Daniel Vincent, Sylvie LaBoissiere and Frederic Robidoux and the staff of the McGill University and Genome Quebec Innovation Centre, and the staff of the Copenhagen DNA laboratory, and Julie M. Cunningham, Sharon A. Windelband, Christopher A. Hiler, Kerry Meyer and the staff of Mayo Clinic Genotyping Core Facility.

**Consortia members**

**GENICA Network.** Hilruit Brauch, Wing-Yee Lo, Christina Justenhoven: Dr. Margaret Fischer-Bosch-Institute of Clinical Pharmacology, Stuttgart, and University of Tubingen, Germany. Yon-Dochno Ko, Christian Baitsch: Department of Internal Medicine, Evangelische Kliniken Bonn, Bonn, Germany. Bonnier Hansen: Hvidovre University Hospital and Genome Quebec Innovation Centre, and the staff of the Copenhagen DNA laboratory, and Julie M. Cunningham, Sharon A. Windelband, Christopher A. Hiler, Kerry Meyer and the staff of Mayo Clinic Genotyping Core Facility.

**kConFab Investigators.** See http://www.konfab.org/Organisation/Members.aspx

**AOCS.** See: http://www.aocstudy.org/olg_coll.asp

**Author Contributions**

Conceived and designed the experiments: HG DQ GCT AC RL DPE. Performed the experiments: SK KG DM DFE. Analyzed the data: SK KG DM HL DFE. Contributed reagents/materials/analysis tools: SK KG DM DFE.

**MOBCS** thanks Ilja Myohanen and Helena Kemilainen. **kConFab/AOCS** 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 for their contributions to this resource, and the many families who contribute to kConFab. The **LMBC** thanks Gilian Peuteman, Dominick Smeets, Thomas Van Brussel and Kathleen Corbouts. The **MARIE** would like to thank Aleina Vrieling, Katharina Buck, Ursula Eilber, Muhlabett Celik, and Sabine Behrens. The **MBCSG** thanks Sranounesh Manoukian, Bernard Peters and the patients and families of the Fondazione IRCCS Istituto Nazionale dei Tumori (INT); Bernardo Bonanni, Irene Feroco and Angela Manichelico of the Istituto Europeo di Oncologia (IEO) and the personnel of the KConFab Cancer Genetic Test Laboratory. The **MTLGBCS** gratefully acknowledge the assistance of Lesley Richardson and Marie-Claire Goulet in conducting the study. We would like to thank Martine Tranchant (Cancer Genomics Laboratory, CHU de Quebec Research Center), Madeleine Maillard, Valerie Baldacchino, Sarah Freund and Len Heguy (McGill University Health Center, Royal Victoria Hospital; McGill University) for DNA extraction, sample management and skilled technical assistance. J.S. is Chairholder of the Canada Research Chair in Oncogenetics. The **OBGS** thanks Meeri Otsukka and Kari Mononen. The **ORIGO** thanks 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 (ONCODC) with the help of Dr. J. Molaen. The **OSU** thanks Robert Pilarski and Charles Shapiro, who were instrumental in the formation of the OSU Breast Cancer Tissue Bank. We thank the Human Genetics Sample Bank for processing of samples. OSU Columbus area control specimens were provided by the Ohio State University’s Human Genetics Sample Bank. The **PBCC** thanks Mark Sherman, Neolina Szeszenia-Dabrowska, Beata Peklonska, Witold Zatorowski, Fei Chiao and Michael Shipp. The **PBCS** thanks Dr. Richard Shpruengel, Bonnier Hansen, Lisabett Grone, Helena Cramp and Charles Shapiro, who were instrumental in the formation of the OSU Breast Cancer Tissue Bank. We thank the Human Genetics Sample Bank for processing of samples. OSU Columbus area control specimens were provided by the Ohio State University’s Human Genetics Sample Bank. The **PBCS** thanks Mark Sherman, Neolina Szeszenia-Dabrowska, Beata Peklonska, Witold Zatorowski, Fei Chiao and Michael Shipp. The **PBCS** thanks Dr. Richard Shpruengel, Bonnier Hansen, Lisabett Grone, Helena Cramp and Charles Shapiro, who were instrumental in the formation of the OSU Breast Cancer Tissue Bank. We thank the Human Genetics Sample Bank for processing of samples. OSU Columbus area control specimens were provided by the Ohio State University’s Human Genetics Sample Bank.


