Genetic modifiers of CHEK2*1100delC associated breast cancer risk

Taru A. Muranen, M.Sc.[1], Dario Greco, PhD[4], Carl Blomqvist, M.D., PhD[2], Kristiina Aittomäki, M.D., PhD[3], Sofia Khan, PhD[1], Frans Hogervorst, PhD[5], Senno Verhoef, M.D.[5], Paul D.P. Pharoah, MB, BCh[6,7], Alison M. Dunning, PhD[6], Mitul Shah, M.Sc.[6], Robert Luben, BS[8], Stig E. Bojesen, M.D., PhD[9,10,11], Børge G. Nordestgaard, M.D., DMSc[9,10,11], Minouk Schoemaker, PhD[12], Anthony Swerdlow, DM, DSc.[12,13], Montserrat García-Closas, PhD[12,14], Jonine Figueroa, PhD[14], Thilo Dörk, PhD[15], Natalia V. Bogdanova, PhD[16], Per Hall, M.D.[17], Jingmei Li, PhD[17], Elza Khusnutdinova, M.D.[20,21], Marina Bermisheva, PhD[15,21], Vessela Kristensen, PhD[22,26,27], Anne-Lise Borresen-Dale, PhD[22,27], NBCS Investigators[22,23,24,25,26,27,28,29,30,31,32,33,34,35,36], Julian Peto, PhD[37], Isabel dos Santos Silva, PhD[37], Fergus J. Couch, PhD[38], Janet E. Olson, PhD[39], Peter Hillemans, PhD[15], Tjoung-Won Park-Simon, M.D.[15], Hiltrud Brauch, PhD[40,46,47], Ute Hamann, PhD[41], Barbara Burwinkel, PhD[42,48], Frederik Marme, M.D.[48,49], Alfons Meindl, PhD[50], Rita K. Schmutzler, M.D.[51,52,53], Angela Cox, PhD[54], Simon S. Cross, M.D.[55], Elinor J. Sawyer, PhD[56], Ian Tomlinson, PhD[57], Diether Lambrechts, PhD[58,59], Matthieu Moisse, PhD[58], Annika Lindblom, M.D.[18], Sara Margolin, M.D.[19], Antoinette Hollestelle, PhD[60], John W.M. Martens, PhD[60], Peter A. Fasching, M.D.[61,62], Matthias W. Beckmann, M.D.[61], Irene L. Andrulis, PhD[63,65], Julia A. Knight, PhD[64,66], kConFab/AOCS Investigators[67], Hoda Anton-Culver, PhD[70], Argyrios Ziogas, PhD[70], Graham G. Giles, PhD[68,71], Roger L. Milne, PhD[68,71], Hermann Brenner, M.D., M.P.H.[40,43,44], Volker Arndt, M.D., M.P.H. [44], Arto Mannermaa, PhD[72,73,74], Veli-Matti Kosma, M.D.[72,73,74], Jenny Chang-Claude, PhD[45], Anja Rudolph, PhD[45], Peter Devilee, PhD[75,76], Caroline Seynaeve, PhD[60], John L. Hopper, PhD[68], Melissa C. Southey, PhD[69], Esther M. John, PhD[77,78,79], Alice S. Whittemore, PhD[78,79], Manjeet K. Bolla, M.Sc.[7], Qin Wang, M.Sc.[7], Kyriaki Michailidou, PhD[7,80], Joe Dennis, M.SC.[7], Douglas F. Easton, PhD[6,7], Marjanka K. Schmidt, PhD[5]*, Heli Nevanlinna, PhD[1]*

*These authors contributed equally

Corresponding author: Heli Nevanlinna, PhD, post address P.O.Box 700, 00029 HUS, Finland, phone +358 9 471 71750, fax +358 9 4717 1751, email heli.nevanlinna@hus.fi

AUTHOR AFFILIATIONS

[1]Department of Obstetrics and Gynecology, [2]Department of Oncology, [3]Department of Clinical Genetics, Helsinki University Hospital, University of Helsinki, Helsinki, Finland;

[4]Unit of Systems Toxicology, Finnish Institute of Occupational Health, Helsinki, Finland;

[5]Netherlands Cancer Institute, Antoni van Leeuwenhoek hospital, Amsterdam, The Netherlands;

[6]Centre for Cancer Genetic Epidemiology, Department of Oncology, [7]Centre for Cancer Genetic Epidemiology, Department of Public Health and Primary Care, [8]Clinical Gerontology, Department of Public Health and Primary Care, University of Cambridge, Cambridge, UK;

[9]Faculty of Health and Medical Sciences, University of Copenhagen, Copenhagen, Denmark;

[10]Copenhagen General Population Study, [11]Department of Clinical Biochemistry, Herlev Hospital, Copenhagen University Hospital, Herlev, Denmark;

[12]Division of Genetics and Epidemiology, [13]Division of Breast Cancer Research, The Institute of Cancer Research, London, UK;

[14]Division of Cancer Epidemiology and Genetics, National Cancer Institute, Rockville, MD, USA;

[15]Gynaecology Research Unit, [16]Department of Radiation Oncology, Hannover Medical School, Hannover, Germany;

[17]Department of Medical Epidemiology and Biostatistics, [18]Department of Molecular Medicine and Surgery, [19]Department of Oncology - Pathology, Karolinska Institutet, Stockholm, Sweden;

[20]Department of Genetics and Fundamental Medicine, Bashkir State University, Ufa, Russia;

[21]Institute of Biochemistry and Genetics, Ufa Scientific Center of Russian Academy of Sciences, Ufa, Russia;

[22]Department of Genetics, Institute for Cancer Research, [23]Department of Oncology, [24]Department of Radiology, [25]National Resource Centre for Long-term Studies after Cancer, Cancer Clinic, Radiumhospitalet, Oslo University Hospital, University of Oslo, Oslo, Norway;

[26]Department of Clinical Molecular Biology, Oslo University Hospital, [27]K.G. Jebsen Center for Breast Cancer Research, Institute of Clinical Medicine, Faculty of Medicine, [28]Department of Breast and Endocrine Surgery, Institute for Clinical Medicine, Ullevaal University Hospital, [29]Department of Clinical Molecular Biology, Institute of Clinical Medicine, Akershus University Hospital, [30]Department of Oncology, Ullevaal University Hospital, University of Oslo, Oslo, Norway;

[31]Department of Pathology, [32]Department of Surgery, Akershus University Hospital, Lørenskog, Norway;

[33]Department of Oncology, Haukeland University Hospital, Bergen, Norway;

[34]Section of Oncology, Institute of Medicine, University of Bergen, Bergen, Norway;

[35]Norwegian Centre for Integrated Care and Telemedicine, University Hospital of North Norway, Tromsø, Norway;

[36]Department of Community Medicine, Faculty of Health Sciences, University of Tromsø - The Arctic University of Norway, Tromsø, Norway;

[37]Department of Non-Communicable Disease Epidemiology, London School of Hygiene and Tropical Medicine, London, UK;

[38]Department of Laboratory Medicine and Pathology, [39]Department of Health Sciences Research, Mayo Clinic, Rochester, MN, USA;

[40]German Cancer Consortium (DKTK), [41]Molecular Genetics of Breast Cancer, [42]Molecular Epidemiology Group, [43]Division of Preventive Oncology, [44]Division of Clinical Epidemiology and Aging Research, [45]Division of Cancer Epidemiology, German Cancer Research Center (DKFZ), Heidelberg, Germany;

[46]Dr. Margarete Fischer-Bosch-Institute of Clinical Pharmacology, Stuttgart, Germany;

[47]University of Tübingen, Tübingen, Germany;

[48]Department of Obstetrics and Gynecology, [49]National Center for Tumor Diseases, University of Heidelberg, Heidelberg, Germany;

[50]Division of Gynaecology and Obstetrics, Technische Universität München, Munich, Germany;

[51]Center for Molecular Medicine Cologne (CMMC), University of Cologne, Cologne, Germany;

[52]Center for Hereditary Breast and Ovarian Cancer, [53]Center for Integrated Oncology (CIO), University Hospital of Cologne, Cologne, Germany;

[54]Sheffield Cancer Research, Department of Oncology, [55]Academic Unit of Pathology, Department of Neuroscience, University of Sheffield, Sheffield, UK;

[56]Research Oncology, Guy's Hospital, King's College London, London, UK;

[57]Wellcome Trust Centre for Human Genetics and Oxford NIHR Biomedical Research Centre, University of Oxford, Oxford, UK;

[58]Laboratory for Translational Genetics, Department of Oncology, University of Leuven, Leuven, Belgium;

[59]Vesalius Research Center, VIB, Leuven, Belgium;

[60]Department of Medical Oncology, Family Cancer Clinic, Erasmus MC Cancer Institute, Rotterdam, The Netherlands;

[61]Department of Gynaecology and Obstetrics, University Hospital Erlangen, Friedrich-Alexander University Erlangen-Nuremberg, Comprehensive Cancer Center Erlangen-EMN, Erlangen, Germany;

[62]David Geffen School of Medicine, Department of Medicine Division of Hematology and Oncology, University of California at Los Angeles, Los Angeles, CA, USA;

[63]Department of Molecular Genetics, [64]Division of Epidemiology, Dalla Lana School of Public Health, University of Toronto, Toronto, Canada;

[66]Prosserman Centre for Health Research, [65]Lunenfeld-Tanenbaum Research Institute of Mount Sinai Hospital, Toronto, Canada;

[67]Peter MacCallum Cancer Center, [68]Centre for Epidemiology and Biostatistics, Melbourne School of Population and Global health, [69]Department of Pathology, The University of Melbourne, Melbourne, Australia;

[70]Department of Epidemiology, University of California Irvine, Irvine, CA, USA;

[71]Cancer Epidemiology Centre, Cancer Council Victoria, Melbourne, Australia;

[72]Institute of Clinical Medicine, Pathology and Forensic Medicine, University of Eastern Finland, Kuopio, Finland;

[73]Cancer Center, [74]Imaging Center, Department of Clinical Pathology, Kuopio University Hospital, Kuopio, Finland;

[75]Department of Human Genetics, [76]Department of Pathology, Leiden University Medical Center, Leiden, The Netherlands;

[77]Department of Epidemiology, Cancer Prevention Institute of California, Fremont, CA, USA;

[78]Department of Health Research and Policy - Epidemiology, [79]Stanford Cancer Institute, Stanford University School of Medicine, Stanford, CA, USA;

[80]Department of Electron Microscopy/Molecular Pathology, The Cyprus Institute of Neurology and Genetics, Nicosia, Cyprus

CONFLICT OF INTEREST NOTIFICATION PAGE

CONFLICT OF INTEREST: The authors declare no conflict of interest. The funders had no role in conception and design of the study, nor in interpretation of the final results.

FUNDING

The Breast Cancer Association Consortium (BCAC) is funded by 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).

Funding for the iCOGS infrastructure came from: 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, C8197/A16565), 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.

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 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.

The Amsterdam Breast Cancer Study (ABCS) was supported by the Dutch Cancer Society [grants NKI 2007-3839; 2009 4363]; BBMRI-NL, which is a Research Infrastructure financed by the Dutch government (NWO 184.021.007); and the Dutch National Genomics Initiative.

The work of the Bavarian Breast Cancer Cases and Controls (BBCC) was partly funded by ELAN-Fond of the University Hospital of Erlangen.

The British Breast Cancer Study (BBCS) 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).

EJS is supported by NIHR 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 Breast Cancer Study of the University of Heidelberg (BSUCH) was supported by the Dietmar-Hopp Foundation, the Helmholtz Society and the German Cancer Research Center (DKFZ).

The Copenhagen General Population Study (CGPS) was supported by the Chief Physician Johan Boserup and Lise Boserup Fund, the Danish Medical Research Council and Herlev Hospital.

The ESTHER Breast Cancer 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 VERDI study, which was supported by a grant from the German Cancer Aid (Deutsche Krebshilfe).

The German Consortium of Hereditary Breast and Ovarian Cancer (GC-HBOC) is supported by the German Cancer Aid (grant no 110837, coordinator: RKS).

The Gene Environment Interaction and Breast Cancer in Germany (GENICA) was funded by the 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.

The Genetic Epidemiology Study of Breast Cancer by Age 50 (GESBC) was supported by the Deutsche Krebshilfe e. V. [70492] and the German Cancer Research Center (DKFZ).

The Hannover Breast Cancer Study (HABCS) study was supported by the Rudolf Bartling Foundation.

The Helsinki Breast Cancer Study (HEBCS) 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 work of TAM has been supported by Ida Montin Foundation, Cancer Society of Finland and Finnish Cultural Foundation.

The Hannover-Minsk Breast Cancer Study (HMBCS) was supported by a grant from the Friends of Hannover Medical School and by the Rudolf Bartling Foundation.

The Hannover-Ufa Breast Cancer Study (HUBCS) was supported by a grant from the German Federal Ministry of Research and Education (RUS08/017).

"Financial support for the Karolinska Breast Cancer Study (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 and Bert von Kantzows foundation.

The Kuopio Breast Cancer Project (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.

The Kathleen Cuningham Foundation Consortium for research into Familial Breast Cancer (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.

Financial support for the Australian Ovarian Cancer Study (AOCS) was provided by the United States Army Medical Research and Materiel Command [DAMD17-01-1-0729], Cancer Council Victoria, Queensland Cancer Fund, Cancer Council New South Wales, Cancer Council South Australia, The Cancer Foundation of Western Australia, Cancer Council Tasmania and the National Health and Medical Research Council of Australia (NHMRC; 400413, 400281, 199600).

The Leuven Multidisciplinary Breast Centre (LMBC) is supported by the 'Stichting tegen Kanker' (232-2008 and 196-2010). DL is supported by the FWO and the KULPFV/10/016-SymBioSysII.

The Mayo Clinic Breast Cancer Study (MCBCS) was supported by the 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.

The Melbourne Collaborative Cohort Study (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. Cases and their vital status were ascertained through the Victorian Cancer Registry (VCR) and the Australian Institute of Health and Welfare (AIHW), including the National Death Index.

The Norwegian Breast Cancer Study (NBCS) has received funding from the K.G. Jebsen Centre for Breast Cancer Research; the Research Council of Norway grant 193387/V50 (to ALBD and VKr) and grant 193387/H10 (to ALBD and VKr), South Eastern Norway Health Authority (grant 39346 to to ALBD and VKr) and the Norwegian Cancer Society (to to ALBD and VKr).

The Northern California Breast Cancer Family Registry (NC-BCFR) 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 Ontario Familial Breast Cancer Registry (OFBCR) 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 Leiden University Medical Centre Breast Cancer Study (ORIGO) was supported by the Dutch Cancer Society (RUL 1997-1505) and the Biobanking and Biomolecular Resources Research Infrastructure (BBMRI-NL CP16).

The NCI Polish Breast Cancer Study (PBCS) was funded by Intramural Research Funds of the National Cancer Institute, Department of Health and Human Services, USA.

The Rotterdam Breast Cancer Study (RBCS) was funded by the Dutch Cancer Society (DDHK 2004-3124, DDHK 2009-4318).

The Singapore and Sweden Breast Cancer Study (SASBAC) 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 Sheffield Breast Cancer Study (SBCS) was supported by Yorkshire Cancer Research S295, S299, S305PA and Sheffield Experimental Cancer Medicine Centre.

The Study of Epidemiology and Risk factors in Cancer Heredity (SEARCH) is funded by a programme grant from Cancer Research UK [C490/A10124] and supported by the UK National Institute for Health Research Biomedical Research Centre at the University of Cambridge.

The UCI Breast Cancer Study (UCIBCS) component of this research was supported by the NIH [CA58860, CA92044] and the Lon V Smith Foundation [LVS39420].

The UK Breakthrough Generations Study (UKBGS) is funded by Breast Cancer Now and the Institute of Cancer Research (ICR), London. ICR acknowledges NHS funding to the NIHR Biomedical Research Centre.

ABSTRACT

Purpose

*CHEK2**1100delC is a founder variant in European populations conferring a 2-3 fold increased risk of breast cancer (BC). Epidemiologic and family studies have suggested that the risk associated with *CHEK2**1100delC is modified by other genetic factors in a multiplicative fashion. We have investigated this empirically using data from the Breast Cancer Association Consortium (BCAC).

Methods

With genotype data of 39,139 (624 1100delC carriers) BC patients and 40,063 (224) healthy controls from 32 BCAC studies, we analyzed the combined risk effects of *CHEK2**1100delC and 77 common variants in terms of a polygenic risk score (PRS) and pairwise interaction.

Results

The PRS conferred an odds ratio (OR) of 1.59 [95% CI 1.21-2.09] per standard deviation for BC for *CHEK2**1100delC carriers and 1.58 [1.55-1.62] for non-carriers. No evidence for deviation from the multiplicative model was found. The OR for the highest quintile of the PRS was 2.03 [0.86-4.78] and for the lowest quintile 0.52 [0.16-1.74] for *CHEK2**1100delC carriers, corresponding to over 34.0% and less than 15.0% life-time risk, respectively.

Conclusion

Our results confirm the multiplicative nature of risk effects conferred by *CHEK2**1100delC and the common susceptibility variants. Furthermore, the PRS could identify the carriers at a high life-time risk for clinical actions.

Keywords: Breast cancer; *CHEK2**1100delC; Polygenic risk score (PRS); common variants; Breast Cancer Association Consortium (BCAC)

INTRODUCTION

The protein truncating mutation *CHEK2**1100delC (checkpoint kinase 2) is a moderate penetrance breast cancer risk variant with relative risk estimate of 2-3 fold.[1, 2] However, several studies have shown that the absolute risk of breast cancer in *CHEK2**1100delC carriers is markedly higher in women with a family history than without,[3-5] and that *CHEK2**1100delC carriers have a higher probability of developing bilateral breast cancer.[6] These observations are quantitatively consistent with a simple polygenic model suggesting that *CHEK2**1100delC combines multiplicatively with other genetic loci. However, this has not yet been established empirically.

Genome wide association studies have identified common genetic variants that are associated with increased risk of breast cancer. A polygenic risk score (PRS), based on 77 low penetrance variants has been estimated to explain approximately 12-14% of the excess familial risk and shown to identify individuals at highest genetic risk at the population level.[7, 8] Some of these variants predominantly predispose to either estrogen receptor positive (ER+) or estrogen receptor negative (ER-) disease, which represent the two main etiological subclasses of breast cancer.[9] *CHEK2**1100delC carriers are more strongly predisposed to ER+ disease: about 90% of carrier tumors are ER+ in comparison to 77-78% of non-carrier tumours.[10]

Here, we investigate the synergistic risk effects attributable to *CHEK2**1100delC and the common breast cancer susceptibility variants both individually and summarized in terms of the PRS.[7, 8]

PATIENTS AND METHODS

Study participants

Female invasive breast cancer patients and healthy controls of European ancestry were included from studies participating in the Breast Cancer Association Consortium (BCAC)(Table S1). Data from a study were included if the study provided genotype data of the common variants from at least one breast cancer patient carrying the 1100delC variant. This selection yielded data from 32 studies and

a total of 79 202 study subjects, including 848 *CHEK2**1100delC carriers (Table S2) for pairwise interaction analyses. Complete quality controlled[7, 10] genotype data for all 77 common variants and *CHEK2**1100delC were available from 33 624 study subjects (369 *CHEK2**1100delC carriers, Table S2). This data were used in the analyses involving the PRS.

All participating studies were approved by their institutional review committees. Each study followed national guidelines for participant inclusion and for informed consent procedures.

Genotyping

All variants except *CHEK2**1100delC were genotyped centrally using a custom Illumina iSelect genotyping array (iCOGS, Illumina, Inc. San Diego, CA, USA) as part of the COGS consortium studies as described earlier.[7, 8] *CHEK2**1100delC was primarily genotyped using a custom made TaqMan assay (Applied Biosystems, Foster City, CA, USA), with a small minority being genotyped using iPLEX.[10] In addition to the 38,549 study subjects genotyped using the iCOGS array, 40,653 BCAC study subjects were genotyped for up to 25 of the common risk variants and these data were used in the pairwise interaction analysis (Table S2, Table S3). These samples were genotyped by independent studies following BCAC genotyping standards as described previously.[11, 12]

Statistical analyses

Statistical analyses were performed using Stata SE 10 (StataCorp, Texas, USA) and R version 2.15.2.[13] For the common variants a log-additive model was assumed; i.e. the risk was analyzed in terms of the number of disease-associated alleles [0, 1, 2] carried. For practical reasons, *CHEK2**1100delC was assumed to follow a dominant inheritance model (i.e. rare homozygotes (n=19) were combined with heterozygotes). All analyses were adjusted for study and seven principal components defined on the basis of the genome-wide data from the iCOGS project as described earlier.[7]

Polygenic risk score

In order to investigate the combined effects of the 77 common variants and *CHEK2**1100delC, a polygenic risk score (PRS) based on the main effects of the common variants was calculated using the formula: $\sum_{i=1}^{n} a_i \log_2 OR_i$

where n is the number of loci included in the model, a is the number of susceptibility alleles in locus i and OR is the per allele odds ratio for breast cancer, estimated separately for each variant in the whole data set (Table S4, column "All"). The PRS was standardized by mean and standard deviation of the PRS distribution among the healthy individuals.[8] Noteworthy, of pairs of linked variants with r2>0.75, we included in the PRS only the lead variant (rs2981579, not rs2981582; rs12662670, not rs3757318; rs554219, not rs614367). Furthermore, we did not include in the analyses any imputed data, nor rs17879961, the *CHEK2* missense variant I157T, because the number of study subjects carrying both 1100delC and I157T was very low (n=5). The interaction between PRS and *CHEK2**1100delC was assessed by comparing nested logistic regression models: a model including the PRS and 1100delC genotype and a model supplemented with an interaction term, coded as the product of the PRS and 1100delC. In analyses of the PRS and positive family history of breast cancer, positive family history was defined as at least one first degree relative with breast cancer.

The cumulative life-time breast cancer risk of *CHEK2**1100delC carriers in different PRS-percentiles was derived assuming an average life-time risk of 23%[14] and previously published relative risk estimates associated with the PRS.[8]

Pairwise interaction analyses

We tested for pairwise interaction between each common variant and *CHEK2**1100delC as described above for the interaction between the PRS and 1100delC. P-values were corrected for 77 parallel tests using the Benjamini-Hochberg method.[15] Furthermore, the OR for breast cancer was estimated separately for each of the common variants for the whole dataset and for the subgroup of 1100delC carriers. These analyses were also performed separately on the subgroup of breast cancer

patients with ER+ disease, because 1100delC is associated with ER+ breast cancer.[10] Statistical power was estimated as previously suggested for risk interaction analyses.[16]

RESULTS

We analyzed the combined effects of *CHEK2**1100delC and 77 common low penetrance breast cancer risk variants using data from the international Breast Cancer Association Consortium (Table S2). The PRS summarizing the individual effects of the common variants was strongly associated with breast cancer risk among *CHEK2**1100delC carriers (OR per unit standard deviation 1.59 [1.21 - 2.09], P=0.0008) and the OR was similar to that in non-carriers (1.58 [1.55 - 1.62], $P_{interaction}$ 0.93). ORs for the highest and lowest quintiles of the PRS distribution were 2.03 [0.86 - 4.78] and 0.52 [0.16 - 1.74] for *CHEK2**1100delC carriers, respectively, when compared to the middle quintile (Table 1). Both estimates were well in line with those made among non-carriers.

The OR associated with *CHEK2**1100delC in the analysis data set 2.99 [2.32 – 3.85] was attenuated, when the model was adjusted for positive family history of breast cancer. Also, the OR associated with the PRS was slightly attenuated (Table 2). Any significant interaction between risk effects associated with 1100delC, PRS and positive family history was not found. However, in a case-only analysis there was a significant association between the PRS and family history of breast cancer, among both *CHEK2**1100delC carriers (OR 1.29 [1.01 - 1.65], P=0.04) and non-carriers (OR 1.17 [1.12 - 1.21], P=4E-16) (Figure S1).

When the common variants were considered individually, we found nominally significant interactions between five variants and *CHEK2**1100delC for overall breast cancer (rs11249433, rs11780156, rs204247, rs2981582 and rs704010; Table S4a). Two of these represented synergistic (more than multiplicative) and three antagonistic interactions (the estimated effect in 1100delC carriers being in the opposite direction to that in non-carriers). However, none of the interactions were significant after correction for multiple testing. Nine variants showed a nominally significant interaction for ER-positive breast cancer (Table S4b).

DISCUSSION

Our analyses on the synergistic effects of *CHEK2**1100delC and 77 common low penetrance variants on breast cancer risk give strong support to the predicted multiplicative polygenic model.[8, 17, 18] While this has previously been shown for combinations of low penetrance variants,[8] and for variants in combination with BRCA1 and BRCA2 mutations,[19] this is the first direct demonstration for a "moderate" risk gene and has important implications for risk prediction. The PRS was a significant risk factor for *CHEK2**1100delC carriers, and the estimated OR per unit standard deviation was very similar in *CHEK2**1100delC carriers and in non-carriers, consistent with the hypothesis that the common susceptibility variants combine with the rare *CHEK2**1100delC variant in an approximately multiplicative fashion. Similarly, the PRS risk estimates for the highest and lowest quintiles did not differ between the *CHEK2**1100delC carriers and non-carriers. These two estimates made for the *CHEK2**1100delC carriers alone did not reach statistical significance (Table 1), possibly reflecting limited statistical power due to the relatively low number of healthy variant carriers (Table S2). However, this is the largest cohort genotyped for *CHEK2**1100delC and these common variants, and even though some of the point estimates are not significant as such, they are consistent with the previous reports. Most importantly, we did not find evidence for deviation from the multiplicative model, suggesting that the PRS could be used in risk stratification of 1100delC carriers as it can be used for non-carriers.

The unadjusted OR for the *CHEK2**110delC variants (Table 2) was higher in our analysis data set than in previous reports.[2, 14] Adjusting for positive family history markedly attenuated the *CHEK2**1100delC associated, suggestive of some oversampling of familial cases. The PRS was also slightly attenuated after the adjustment. However, *CHEK2**1100delC, PRS and family history remained significant risk factors in the combined model (Table 2) suggesting that the common variants together explain part of the excess familial risk as previously suggested,[17] but that the PRS has predictive value also in breast cancer families segregating *CHEK2**1100delC.

Recently, a large study estimating the risk associated with *CHEK2**1100delC in relation to age, tumor subtype and family history reported the cumulative life-time risk for 1100delC carriers to be about 23%.[14] Assuming that the genetic risk attributable to the common variants (the PRS) would vary around this estimate similarly as published previously for non-carriers (OR higher than 1.48 [1.39 - 1.57] or lower than 0.65 [0.60 – 0.70] for percentiles above 80% or lower than 20%, respectively),[8] 20% of the 1100delC carriers with highest PRS would have life-time risk higher than 34.0% [32.0% - 36.1%] exceeding the threshold for the high-risk category (>30%) according to the UK NICE guidelines for familial breast cancer.[20] Similarly, for the 20% of 1100delC carriers with lowest PRS, the life-time risk would be lower than 15.0% [13.8% - 16.1%], i.e. close to population risk (<17%). These observations imply that, if *CHEK2**1100delC is to be used in risk prediction, it can be made more effectively by including in the prediction also the PRS representing the risk modifying effects of the common variants.

*CHEK2**1100delC carrier cancers do not represent a phenotypically distinct subgroup of breast carcinomas. Instead, the phenotypic diversity of *CHEK2**1100delC associated cancers resembles that of breast tumors in general.[10] Thus, it was not surprising that the relative risks conferred by the common variants were similar for the *CHEK2**1100delC carriers and for non-carriers, and no significant pairwise interaction was found. We estimated that we had sufficient statistical power (80%) to exclude the possibility of such pairwise interaction between *CHEK2**1100delC and any of the common variants that would have an effect size of 2.5 (OR for interaction term higher than 2.5) but not enough power to investigate interactions comparable in magnitude to the risk effects associated with the low penetrance variants (OR 1.1-1.5). Thus, it remains possible that more modest departures from a multiplicative model may exist. If so, however, much larger case-control studies, perhaps combined with pedigree analyses, will be required to detect them.

In conclusion, our analyses confirm the predicted multiplicative relationship between *CHEK2**1100delC and the common low penetrance variants. Hence, the PRS could be similarly applied for risk prediction for the variant carriers as for the general population. Most importantly,

the PRS could help identifying the high risk group of the *CHEK2**1100delC carriers, who would best benefit from clinical intervention.

ACKNOWLEDGEMENTS

We thank all the individuals who took part in these studies and all the researchers, clinicians, technicians and administrative staff who enabled this work to be carried out (details in online Supplementary data). Especially we thank the staffs of the Centre for Genetic Epidemiology Laboratory, the CNIO genotyping unit, the McGill University and Génome Québec Innovation Center, the Copenhagen DNA laboratory and the Mayo Clinic Genotyping Core Facility.

REFERENCES

1. CHEK2 Breast Cancer Case-Control Consortium. CHEK2*1100delC and susceptibility to breast cancer: A collaborative analysis involving 10,860 breast cancer cases and 9,065 controls from 10 studies. *Am J Hum Genet* 2004;**74**:1175-1182.

2. Weischer M, Bojesen SE, Ellervik C, Tybjaerg-Hansen A, Nordestgaard BG. CHEK2*1100delC genotyping for clinical assessment of breast cancer risk: Meta-analyses of 26,000 patient cases and 27,000 controls. *J Clin Oncol* 2008;**26**:542-548.

3. Cybulski C, Wokolorczyk D, Jakubowska A, et al. Risk of breast cancer in women with a CHEK2 mutation with and without a family history of breast cancer. *J Clin Oncol* 2011;**29**:3747-3752.

4. Adank MA, Verhoef S, Oldenburg RA, et al. Excess breast cancer risk in first degree relatives of CHEK2 *1100delC positive familial breast cancer cases. *Eur J Cancer* 2013;**49**:1993-1999.

5. Johnson AD, Handsaker RE, Pulit SL, Nizzari MM, O'Donnell CJ, de Bakker PI. SNAP: A web-based tool for identification and annotation of proxy SNPs using HapMap. *Bioinformatics* 2008;**24**:2938-2939.

6. Fletcher O, Johnson N, Dos Santos Silva I, et al. Family history, genetic testing, and clinical risk prediction: Pooled analysis of CHEK2 1100delC in 1,828 bilateral breast cancers and 7,030 controls. *Cancer Epidemiol Biomarkers Prev* 2009;**18**:230-234.

7. Michailidou K, Hall P, Gonzalez-Neira A, et al. Large-scale genotyping identifies 41 new loci associated with breast cancer risk. *Nat Genet* 2013;**45**:353-361.

8. Mavaddat N, Pharoah PD, Michailidou K, et al. Prediction of breast cancer risk based on profiling with common genetic variants. *J Natl Cancer Inst* 2015;**107**:10.1093/jnci/djv036. Print 2015 May

9. Anderson WF, Rosenberg PS, Prat A, Perou CM, Sherman ME. How many etiological subtypes of breast cancer: Two, three, four, or more? *J Natl Cancer Inst* 2014;**106**:10.1093/jnci/dju165. Print 2014 Aug.

10. Weischer M, Nordestgaard BG, Pharoah P, et al. CHEK2*1100delC heterozygosity in women with breast cancer associated with early death, breast cancer-specific death, and increased risk of a second breast cancer. *J Clin Oncol* 2012;**30**:4308-4316.

11. Cox A, Dunning AM, Garcia-Closas M, et al. A common coding variant in CASP8 is associated with breast cancer risk. *Nat Genet* 2007;**39**:352-358.

12. Easton DF, Pooley KA, Dunning AM, et al. Genome-wide association study identifies novel breast cancer susceptibility loci. *Nature* 2007;**447**:1087-1093.

13. R Foundation for Statistical Computing, Vienna, Austria [computer program].R Core Team, 2013.

14. Schmidt MK, Hogervorst F, van Hien R, et al. Age- and tumor Subtype–Specific breast cancer risk estimates for CHEK2*1100delC carriers. *J Clin Oncol.* In press.

15. Benjamini Y, Hochberg Y. Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of royal statistical society. Serier B (Methodological)* 1995;**vol. 57**:289-300.

16. Demidenko E. Sample size and optimal design for logistic regression with binary interaction. *Stat Med* 2008;**27**:36-46.

17. Johnson N, Fletcher O, Naceur-Lombardelli C, dos Santos Silva I, Ashworth A, Peto J. Interaction between CHEK2*1100delC and other low-penetrance breast-cancer susceptibility genes: A familial study. *Lancet* 2005;**366**:1554-1557.

18. Antoniou AC, Pharoah PD, McMullan G, et al. A comprehensive model for familial breast cancer incorporating BRCA1, BRCA2 and other genes. *Br J Cancer* 2002;**86**:76-83.

19. Kuchenbaecker KB, Neuhausen SL, Robson M, et al. Associations of common breast cancer susceptibility alleles with risk of breast cancer subtypes in BRCA1 and BRCA2 mutation carriers. *Breast Cancer Res* 2014;**16**:3416-014-0492-9.

20. National Collaborating Centre for Cancer (UK). 2013.

LEGENDS TO FIGURES AND TABLES

Table 1. Breast cancer risk associated with the polygenic risk score (PRS) for non-carriers and the carriers of *CHEK2**1100delC.

Table 2. Relative breast cancer risk associated with *CHEK2**1100delC, PRS and positive family history of breast cancer in the analysis data set.

SUPPLEMENTARY INFORMATION

Figure S1. Relationship between the polygenic risk score (PRS) and positive family history of breast cancer.

Table S1. Description of study design and genotype data availability of 32 studies participating in the Breast Cancer Association Consortium (BCAC).

Table S2. *CHEK2**1100delC genotype data availability for breast cancer (BC) cases and controls.

Table S3. Description of genotype data coverage and genotyping methods for each low penetrance variant.

Table S4. Odds ratios (OR) and 95% confidence intervals (CI) estimated for the whole dataset and for the carriers of *CHEK2**1100delC, as well as for pairwise interaction between each variant and *CHEK2**1100delC for (a) breast cancer (b) estrogen receptor positive (ER+) breast cancer.

Supplementary data. Detailed acknowledgements.