rs2735383, located at a microRNA binding site in the 3'UTR of *NBS1*, is not associated with breast cancer risk

Jingjing Liu[1], Ivona Lončar[1], J. Margriet Collée[2], Manjeet K. Bolla[3], Joe Dennis[3], Kyriaki Michailidou[3,4], Qin Wang[3], Irene L. Andrulis[5,6], Monica Barile[7], Matthias W. Beckmann[8], Sabine Behrens[9,10], Javier Benitez[11,12], Carl Blomqvist[13], Bram Boeckx[14,15], Natalia V. Bogdanova[16-18], Stig E. Bojesen[19-21], Hiltrud Brauch[22-24], Paul Brennan[25], Hermann Brenner[24,26,27], Annegien Broeks[28], Barbara Burwinkel[29,30], Jenny Chang-Claude[9,10], Shou-Tung Chen[31], Georgia Chenevix-Trench[32], Ching Y. Cheng[33-35], Ji-Yeob Choi[36,37], Fergus J. Couch[38], Angela Cox[39], Simon S. Cross[40], Katarina Cuk[26], Kamila Czene[41], Thilo Dörk[17], Isabel dos-Santos-Silva[42], Peter A. Fasching[8,43], Jonine Figueroa[44,45], Henrik Flyger[46], Montserrat García-Closas[45], Graham G. Giles[47,48], Gord Glendon[5], Mark S. Goldberg[49,50], Anna González-Neira[11], Pascal Guénel[51], Christopher A. Haiman[52], Ute Hamann[53], Steven N. Hart[54], Mikael Hartman[55,56], Sigrid Hatse[57], John L. Hopper[48], Hidemi Ito[58,59], Anna Jakubowska[60], Maria Kabisch[53], Daehee Kang[36,37,61], Veli-Matti Kosma[62-64], Vessela N. Kristensen[65,66], Loic Le Marchand[67], Eunjung Lee[52], Jingmei Li[41], Artitaya Lophatananon[68,69], Jan Lubinski[60], Arto Mannermaa[62-64], Keitaro Matsuo[59,70], Roger L. Milne[47,48], NBCS Collaborators[65,71-82], Susan L. Neuhausen[83], Heli Nevanlinna[84], Nick Orr[85], Jose I.A. Perez[86], Julian Peto[42], Thomas C. Putti[87], Katri Pylkäs[88,89], Paolo Radice[90], Suleeporn Sangrajrang[91], Elinor J. Sawyer[92], Marjanka K. Schmidt[28,93], Andreas Schneeweiss[29,94], Chen-Yang Shen[95,96], Martha J. Shrubsole[97], Xiao-Ou Shu[97], Jacques Simard[98], Melissa C. Southey[99], Anthony Swerdlow[85,100], Soo H. Teo[101,102], Daniel C. Tessier[103], Somchai Thanasitthichai[104], Ian Tomlinson[105], Diana Torres[53,106], Thérèse Truong[51], Chiu-Chen Tseng[52], Celine Vachon[54], Robert Winqvist[88,89], Anna H. Wu[52], Drakoulis

Yannoukakos[107], Wei Zheng[97], Per Hall[41], Alison M. Dunning[108], Douglas F. Easton[3,108], Maartje J. Hooning[1], Ans M.W. van den Ouweland[2], John W.M. Martens[1,109], Antoinette Hollestelle[1]*.

1 Department of Medical Oncology, Family Cancer Clinic, Erasmus MC Cancer Institute, Rotterdam, The Netherlands.

2 Department of Clinical Genetics, Erasmus University Medical Center, Rotterdam, The Netherlands.

3 Centre for Cancer Genetic Epidemiology, Department of Public Health and Primary Care, University of Cambridge, Cambridge, UK.

4 Department of Electron Microscopy/Molecular Pathology, The Cyprus Institute of Neurology and Genetics, Nicosia, Cyprus.

5 Fred A. Litwin Center for Cancer Genetics, Lunenfeld-Tanenbaum Research Institute of Mount Sinai Hospital, Toronto, Canada.

6 Department of Molecular Genetics, University of Toronto, Toronto, Canada.

7 Division of Cancer Prevention and Genetics, Istituto Europeo di Oncologia, Milan, Italy.

8 Department of Gynaecology and Obstetrics, University Hospital Erlangen, Friedrich-Alexander University Erlangen-Nuremberg, Comprehensive Cancer Center Erlangen-EMN, Erlangen, Germany.

9 Division of Cancer Epidemiology, German Cancer Research Center (DKFZ), Heidelberg, Germany.

10 University Cancer Center Hamburg (UCCH), University Medical Center Hamburg-Eppendorf, Hamburg, Germany.

11 Human Cancer Genetics Program, Spanish National Cancer Research Centre, Madrid, Spain.

12 Centro de Investigación en Red de Enfermedades Raras (CIBERER), Valencia, Spain.

13 Department of Oncology, Helsinki University Hospital, University of Helsinki, Helsinki, Finland.

14 Vesalius Research Center, VIB, Leuven, Belgium.

15 Laboratory for Translational Genetics, Department of Oncology, University of Leuven, Leuven, Belgium.

16 Department of Radiation Oncology, Hannover Medical School, Hannover, Germany.

17 Gynaecology Research Unit, Hannover Medical School, Hannover, Germany.

18 N.N. Alexandrov Research Institute of Oncology and Medical Radiology, Minsk, Belarus.

19 Copenhagen General Population Study, Herlev and Gentofte Hospital, Copenhagen University Hospital, Herlev, Denmark.

20 Department of Clinical Biochemistry, Herlev and Gentofte Hospital, Copenhagen University Hospital, Herlev, Denmark.

21 Faculty of Health and Medical Sciences, University of Copenhagen, Copenhagen, Denmark.

22 Dr. Margarete Fischer-Bosch-Institute of Clinical Pharmacology, Stuttgart, Germany.

23 University of Tübingen, Tübingen, Germany.

24 German Cancer Consortium (DKTK), German Cancer Research Center (DKFZ), Heidelberg, Germany.

25 International Agency for Research on Cancer, Lyon, France.

26 Division of Clinical Epidemiology and Aging Research, German Cancer Research Center (DKFZ), Heidelberg, Germany.

27 Division of Preventive Oncology, German Cancer Research Center (DKFZ) and National Center for Tumor Diseases (NCT), Heidelberg, Germany.

28 Division of Molecular Pathology, The Netherlands Cancer Institute - Antoni van Leeuwenhoek Hospital, Amsterdam, The Netherlands.

29 Department of Obstetrics and Gynecology, University of Heidelberg, Heidelberg, Germany.

30 Molecular Epidemiology Group, C080, German Cancer Research Center (DKFZ), Heidelberg, Germany.

31 Department of Surgery, Changhua Christian Hospital, Changhua, Taiwan.

32 Department of Genetics and Computational Biology, QIMR Berghofer Medical Research Institute, Brisbane, Australia.

33 Singapore Eye Research Institute and Singapore National Eye Center, Singapore, Singapore.

34 Department of Ophthalmology, Yong Loo Lin School of Medicine, National University of Singapore and National University Health System, Singapore, Singapore.

35 Duke-NUS Graduate Medical School, Singapore, Singapore.

36 Department of Biomedical Sciences, Seoul National University College of Medicine, Seoul, Korea.

37 Cancer Research Institute, Seoul National University, Seoul, Korea.

38 Department of Laboratory Medicine and Pathology, Mayo Clinic, Rochester, MN, USA.

39 Academic Unit of Molecular Oncology, Department of Oncology and Metabolism, University of Sheffield, Sheffield, UK.

40 Academic Unit of Pathology, Department of Neuroscience, University of Sheffield, Sheffield, UK.

41 Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Stockholm, Sweden.

42 Department of Non-Communicable Disease Epidemiology, London School of Hygiene and Tropical Medicine, London, UK.

43 David Geffen School of Medicine, Department of Medicine Division of Hematology and Oncology, University of California at Los Angeles, Los Angeles, CA, USA.

44 Usher Institute of Population Health Sciences and Informatics, The University of Edinburgh Medical School, Edinburgh, UK.

45 Division of Cancer Epidemiology and Genetics, National Cancer Institute, Rockville, MD, USA.

46 Department of Breast Surgery, Herlev and Gentofte Hospital, Copenhagen University Hospital, Herlev, Denmark.

47 Cancer Epidemiology Centre, Cancer Council Victoria, Melbourne, Australia.

48 Centre for Epidemiology and Biostatistics, Melbourne School of Population and Global health, The University of Melbourne, Melbourne, Australia.

49 Department of Medicine, McGill University, Montreal, Canada.

50 Division of Clinical Epidemiology, Royal Victoria Hospital, McGill University, Montreal, Canada.

51 Cancer & Environment Group, Center for Research in Epidemiology and Population Health (CESP), INSERM, University Paris-Sud, University Paris-Saclay, Villejuif, France.

52 Department of Preventive Medicine, Keck School of Medicine, University of Southern California, Los Angeles, CA, USA.

53 Molecular Genetics of Breast Cancer, German Cancer Research Center (DKFZ), Heidelberg, Germany.

54 Department of Health Sciences Research, Mayo Clinic, Rochester, MN, USA.

55 Saw Swee Hock School of Public Health, National University of Singapore, Singapore, Singapore.

56 Department of Surgery, National University Health System, Singapore, Singapore.

57 Leuven Multidisciplinary Breast Center, Department of Oncology, Leuven Cancer Institute, University Hospitals Leuven, Leuven, Belgium.

58 Division of Epidemiology and Prevention, Aichi Cancer Center Research Institute, Nagoya, Japan.

59 Department of Epidemiology, Nagoya University Graduate School of Medicine, Nagoya, Japan.

60 Department of Genetics and Pathology, Pomeranian Medical University, Szczecin, Poland.

61 Department of Preventive Medicine, Seoul National University College of Medicine, Seoul, Korea.

62 Translational Cancer Research Area, University of Eastern Finland, Kuopio, Finland.

63 Institute of Clinical Medicine, Pathology and Forensic Medicine, University of Eastern Finland, Kuopio, Finland.

64 Imaging Center, Department of Clinical Pathology, Kuopio University Hospital, Kuopio, Finland.
65 Department of Cancer Genetics, Institute for Cancer Research, Oslo University Hospital Radiumhospitalet, Oslo, Norway.
66 Department of Clinical Molecular Biology, Oslo University Hospital, University of Oslo, Oslo, Norway.
67 Epidemiology Program, University of Hawaii Cancer Center, Honolulu, HI, USA.
68 Division of Health Sciences, Warwick Medical School, Warwick University, Coventry, UK.
69 Institute of Population Health, University of Manchester, Manchester, UK.
70 Division of Molecular Medicine, Aichi Cancer Center Research Institute, Nagoya, Japan.
71 Department of Research, Vestre Viken, Drammen, Norway.
72 Institute of Clinical Medicine, Faculty of Medicine, University of Oslo, Oslo, Norway.
73 Department of Breast and Endocrine Surgery, Oslo University Hospital, Ullevål, Oslo, Norway.
74 Department of Radiology and Nuclear Medicine, Oslo University Hospital Radiumhospitalet, Oslo, Norway.
75 Department of Pathology, Akershus University Hospital, Lørenskog, Norway.
76 Department of Tumor Biology, Institute for Cancer Research, Oslo University Hospital Radiumhospitalet, Oslo, Norway.
77 Department of Oncology, Division of Surgery and Cancer and Transplantation Medicine, Oslo University Hospital Radiumhospitalet, Oslo, Norway.
78 National Advisory Unit on Late Effects after Cancer Treatment, Division of Surgery and Cancer and Transplantation Medicine, Oslo University Hospital, Oslo, Norway.
79 Department of Oncology, Oslo University Hospital Ullevål, Oslo, Norway.
80 Department of Breast-Endocrine Surgery, Akershus University Hospital, Oslo, Norway.
81 Section of Oncology, Institute of Medicine, University of Bergen, Bergen, Norway.
82 Department of Oncology, Haukeland University Hospital, Bergen, Norway.

83 Department of Population Sciences, Beckman Research Institute of City of Hope, Duarte, CA, USA.

84 Department of Obstetrics and Gynecology, Helsinki University Hospital, University of Helsinki, Helsinki, Finland.

85 Division of Breast Cancer Research, The Institute of Cancer Research, London, UK.

86 Servicio de Cirugía General y Especialidades, Hospital Monte Naranco, Oviedo, Spain.

87 Department of Pathology, Yong Loo Lin School of Medicine, National University of Singapore, Singapore, Singapore.

88 Laboratory of Cancer Genetics and Tumor Biology, Cancer and Translational Medicine Research Unit, Biocenter Oulu, University of Oulu, Oulu, Finland.

89 Laboratory of Cancer Genetics and Tumor Biology, Northern Finland Laboratory Centre NordLab, Oulu, Finland.

90 Unit of Molecular Bases of Genetic Risk and Genetic Testing, Department of Preventive and Predictive Medicine, Fondazione IRCCS (Istituto Di Ricovero e Cura a Carattere Scientifico) Istituto Nazionale dei Tumori (INT), Milan, Italy.

91 National Cancer Institute, Bangkok, Thailand.

92 Research Oncology, Guy's Hospital, King's College London, London, UK.

93 Division of Psychosocial Research and Epidemiology, The Netherlands Cancer Institute - Antoni van Leeuwenhoek hospital, Amsterdam, The Netherlands.

94 National Center for Tumor Diseases, University of Heidelberg, Heidelberg, Germany.

95 School of Public Health, China Medical University, Taichung, Taiwan.

96 Taiwan Biobank, Institute of Biomedical Sciences, Academia Sinica, Taipei, Taiwan.

97 Division of Epidemiology, Department of Medicine, Vanderbilt Epidemiology Center, Vanderbilt-Ingram Cancer Center, Vanderbilt University School of Medicine, Nashville, TN, USA.

98 Genomics Center, Centre Hospitalier Universitaire de Québec Research Center, Laval University, Québec City, Canada.

99 Genetic Epidemiology Laboratory, Department of Pathology, The University of Melbourne, Melbourne, Australia.

100 Division of Genetics and Epidemiology, The Institute of Cancer Research, London, UK.

101 Cancer Research Malaysia, Subang Jaya, Selangor, Malaysia.

102 Breast Cancer Research Unit, Cancer Research Institute, University Malaya Medical Centre, Kuala Lumpur, Malaysia.

103 McGill University and Génome Québec Innovation Centre, Montréal, Canada.

104 Research and Technology Assessment Department, National Cancer Institute, Thailand.

105 Wellcome Trust Centre for Human Genetics and Oxford NIHR Biomedical Research Centre, University of Oxford, Oxford, UK.

106 Institute of Human Genetics, Pontificia Universidad Javeriana, Bogota, Colombia.

107 Molecular Diagnostics Laboratory, INRASTES, National Centre for Scientific Research "Demokritos", Athens, Greece.

108 Centre for Cancer Genetic Epidemiology, Department of Oncology, University of Cambridge, Cambridge, UK.

109 Cancer Genomics Netherlands.

*Correspondence to a.hollestelle@erasmusmc.nl.

Abstract

NBS1, also known as NBN, plays an important role in maintaining genomic stability. Interestingly, rs2735383 G>C, located in a microRNA binding site in the 3'-untranslated region (UTR) of *NBS1*, was shown to be associated with increased susceptibility to lung and colorectal cancer. However, the relation between rs2735383 and susceptibility to breast cancer is not yet clear. Therefore, we genotyped rs2735383 in 1,170 familial non-*BRCA1/2* breast cancer cases and 1,077 controls using PCR-based restriction fragment length polymorphism (RFLP-PCR) analysis, but found no association between rs2735383CC and breast cancer risk (OR=1.214, 95% CI=0.936-1.574, *P*=0.144). Because we could not exclude a small effect size due to a limited sample size, we further analyzed imputed rs2735383 genotypes (r^2>0.999) of 47,640 breast cancer cases and 46,656 controls from the Breast Cancer Association Consortium (BCAC). However, rs2735383CC was not associated with overall breast cancer risk in European (OR=1.014, 95% CI=0.969-1.060, *P*=0.556) nor in Asian women (OR=0.998, 95% CI=0.905-1.100, *P*=0.961). Subgroup analyses by age, age at menarche, age at menopause, menopausal status, number of pregnancies, breast feeding, family history and receptor status also did not reveal a significant association. This study therefore does not support the involvement of the genotype at *NBS1* rs2735383 in breast cancer susceptibility.

Introduction

The DNA damage response (DDR) pathway maintains the stability of the human genome via a complex network of pathways integrating signal transduction, regulation of the cell cycle and repair of DNA. Double-strand breaks (DSBs), a particularly severe form of DNA damage, arise as a consequence of cell replication, programmed DNA rearrangements (*i.e.* meiosis and VDJ recombination) and exposure to carcinogens. When left unrepaired, DSBs may cause genomic instability, cell death and cancer[1,2]. In fact, mutations in genes involved in DSB repair, but also in the DDR pathway in general, are involved in the etiology of many human cancers. The two major repair pathways that mediate the repair of DSBs are the template-mediated homologous recombination repair pathway and the more error-prone non-homologous end-joining pathway[3,4]. The MRE11/RAD50/NBS1 complex is an important regulator of DSB repair through these pathways as this complex not only acts as a sensor of DSBs, but also recruits and activates the ATM protein to the break and activates it[5]. Activation of ATM, the central mediator of response to DSBs, initiates a cascade of signaling pathways involved in cell cycle checkpoint control, DNA repair and, when necessary, apoptosis by phosphorylation of p53, CHEK2, BRCA1, FANCD2 and NBS1 amongst others[6].

The DDR plays an important role in susceptibility to breast cancer. In fact, all of the currently identified high- and moderate-risk breast cancer genes (*i.e. BRCA1, BRCA2, CHEK2, ATM, NBS1* and *PALB2)* are involved in DNA repair[7,8]. As the majority of familial breast cancer risk is not yet attributable to known risk genes, this makes other genes encoding proteins involved in the DDR pathway attractive candidates for breast cancer susceptibility genes. The recent identification of the early DNA damage response gene *MCPH1*as a novel breast cancer susceptibility gene illustrates that this hypothesis still holds[9].

In this respect, the *NBS1* gene is located at chromosome 8q21 and bi-allelic germline mutations in *NBS1* cause the chromosomal instability syndrome Nijmegen breakage syndrome[10]. In addition, heterozygous carriers of *NBS1* mutations are at an increased risk to develop several types of cancer[11]. The *NBS1* c.657del5 founder mutation is the most prevalent mutation implicated in Nijmegen breakage syndrome (*i.e.* 90%) and has its origin in the Slavic population[12]. The mutation confers an overall 2.5- to 3-fold increased cancer risk and is associated with increased risk for breast cancer, prostate cancer and lymphoma specifically[13]. Two other *NBS1* mutations implicated in Nijmegen breakage syndrome are p.I171V and p.R215W. Although both mutations associate with an overall cancer risk of 4-fold and 2-fold, respectively, there does not seem to be an increased risk to develop breast cancer specifically[13].

Besides the rare Nijmegen breakage syndrome-associated mutations, two common variants in *NBS1* (*i.e.* p.E185Q; rs1805794 and c.2265+541G>C; rs2735383) have also been reported to be associated with risks for several cancer types. Recent meta-analyses for rs1805794 have, however, shown that this variant does not associate with breast cancer risk[13-16], while associations with lung cancer and urinary system cancer are still inconclusive[13,16-18]. The functional variant rs2735383, localized in in the 3' UTR of *NBS1*, has been shown to modulate the binding ability of microRNA-629 in lung cancer cells and microRNA-509-5p in colorectal cancer cells, affect *NBS1* transcriptional activity and decrease *NBS1* mRNA and NBS1 protein levels[19,20]. Although rs2735383 has been associated with an increased risk of lung cancer and colorectal cancer[13,20], its association with breast cancer risk is yet unclear. For this reason, we assessed whether *NBS1* rs2735383 is associated with breast cancer risk in the Rotterdam Breast Cancer Study (RBCS) by RFLP-PCR and in 45 studies of BCAC through imputation of the iCOGS array[24].

Results

To evaluate the association between *NBS1* rs2735383 and breast cancer risk, we analyzed *NBS1* rs2735383 by RFLP-PCR in 1,269 non-*BRCA1/2* familial breast cancer patients and 1,159 controls from RBCS. Since genetic risk factors are usually enriched in familial/early-onset breast cancer cases, specifically selecting these breast cancer patients improves statistical power. Among the cases, 516 had the GG genotype, 507 had the GC genotype and 147 had the CC genotype at rs2735383 (minor allele frequency (MAF)=0.342). Among the controls, 462 had the GG genotype, 501 had the GC genotype and 114 had the CC genotype (MAF=0.338). For both cases and controls, the genotypes of rs2735383 were in Hardy-Weinberg equilibrium (HWE). Because rs2735383 CC was associated with an increased risk of lung cancer and colorectal cancer under a recessive genetic model[13,20], we analyzed the association of rs2735383 with breast cancer in a similar way. However, rs2735383 was not significantly associated with the risk of breast cancer (OR=1.214, 95% CI=0.936-1.574, P=0.144; Table 1). In this respect, the lung cancer risk conferred by the rs2735383 CC genotype had been associated with an OR of 1.28 (95% CI=1.21-1.46, P<0.001), whereas the colorectal cancer risk had been associated with an OR of 1.55 (95% CI=1.27-1.94, $P<10^{-4}$)[13,20]. Here, we do not observe a similar effect size for breast cancer as for lung and colorectal cancer. However, RBCS is underpowered to detect effect sizes smaller than those observed for lung cancer *(i.e.* OR<1.28). Therefore, we cannot exclude rs2735383 CC is associated with breast cancer, but confers a smaller risk.

For this reason we analyzed *NBS1* rs2735383 in BCAC studies through imputation. Since we had data available for RBCS on rs2735383 from both the PCR-based RFLP and from imputation, we first evaluated the concordance between the two methods. In total, from 1,313

samples (*i.e.* 646 cases and 667 controls) we had genotypes for rs2738353 available from both RFLP-PCR and imputation. Importantly, the agreement between the two methods was 97.1% (*i.e.* concordance in 1,275 of 1,313 samples, r^2=0.933) and was similar among cases and controls (*i.e.* 98.1% versus 96.1%). Moreover, case-control ORs for imputed data were comparable to ORs obtained by RFLP-PCR (OR=1.14, 95% CI=0.80-1.62 versus OR=1.17, 95% CI=0.83-1.66). Therefore, we used the imputed data on rs2735383 to evaluate further its association with breast cancer risk.

For the overall analysis in Europeans we had 41,915 cases and 40,042 controls available from 36 case-control studies. However, rs2735383 was not associated with breast cancer risk in Europeans, neither under a recessive genetic model (OR=1.014, 95% CI=0.969-1.060, *P*=0.556; Table 2 and Fig. 1), nor under a dominant (OR=1.006, 95% CI=0.978-1.035, *P*=0.684; Table 2) or additive model (per allele OR=1.000, 95% CI=0.979-1.021, *P*=0.984; Table 2). Because the association with increased lung and colorectal cancer risk was observed in the Asian population[13,20], we also performed the same analysis in the nine Asian BCAC studies. In total, we had 5,725 cases and 6,614 controls available for this analysis from nine case-control studies. Also in Asians we did not find any association between rs2735383 and breast cancer risk for either the recessive (OR=0.998, 95% CI=0.905-1.100, *P*=0.961; Table 2 and Fig. 1), dominant (OR=0.995, 95% CI=0.922-1.074, *P*=0.900; Table 2) or additive genetic model (per allele OR=0.997, 95% CI=0.946-1.050, *P*=0.911; Table 2). These results imply that *NBS1* rs2735383 is not associated with an increased risk to develop invasive breast cancer.

A previous study had shown that rs2735383 may be associated with breast cancer risk in women > 50 years, women with age at menarche >13 years, women with premenopausal status, women with number of abortions ≤ 2 and women who have breast fed, but not by age at

menopause, number of pregnancies and family history[21]. Therefore, to exclude that an association of rs2735383 with breast cancer risk exists in a particular subgroup of individuals or breast cancer patients, we performed subgroup analysis according to age, age at menarche, age at menopause, menopausal status, number of full-term pregnancies, breast feeding, family history and receptor status. We did, however, not find any association between the genotype at rs2735383 and the risk of breast cancer in any of these subgroups for the recessive genetic model (Table 3). Also for the dominant and additive genetic models we found no association between the genotype at rs2735383 and breast cancer risk that would withstand multiple testing correction (Supplementary Tables S1 and S2). *NBS1* rs2735383 is thus not associated with the risk for breast cancer, either in the overall analyses or in specific subgroups.

Discussion

The CC genotype of the common variant rs2735383 in the 3'UTR of *NBS1* has been shown to be associated with an increased cancer risk, specifically for lung and colorectal cancer (lung cancer: OR=1.28, 95% CI=1.21-1.46, $P<0.001$ and colorectal cancer: OR=1.55, 95% CI=1.27-1.94, $P<10^{-4}$)[13,20]. In the current study, we evaluated the association of *NBS1* rs2735383 with breast cancer risk. We found that the CC genotype of rs2735383 did not confer an increased breast cancer risk, neither in the overall analyses nor in the subgroup analyses.

In agreement with these results, a small study by Han *et al.* consisting of 239 premenopausal breast cancer patients and 477 matched controls from the Nurses' Health Study II showed that rs2735383 did not associate with breast cancer risk under an additive genetic model (OR=0.92, 95% CI=0.72-1.16, $P=0.469$)[22]. Moreover, the study of Wu *et al*. consisting of 450 breast cancer patients and 450 cancer-free controls from the Henan Province in China also found

no association with overall breast cancer risk[21]. However, after stratification according to reproductive factors, rs2735383 CC was found to be associated with an increased breast cancer risk for women > 50 years, women with age at menarche >13 years, women with premenopausal status, women with number of abortions ≤ 2 and women who have breast fed, but not by age at menopause, number of pregnancies and family history[21]. In the current study we therefore also performed subgroup analysis by age, age at menarche, age at menopause, menopausal status, number of full-term pregnancies, breast feeding, family history and receptor status, but did not find any association between rs2735383 and risk of breast cancer in any of these subgroups that would withstand multiple testing correction. We thus could not replicate the earlier positive findings in women > 50 years, women with age at menarche >13 years, premenopausal women and women who have breast fed. A possible, but not very likely, explanation for the difference in outcome between the studies may be the European versus Asian ethnicity. In the current study we chose to perform the subgroup analysis only in the European studies and not the Asian studies as this made sure that we had sufficient power in the subgroup analysis to identify smaller effects of rs2735383 on breast cancer risk. In this respect, a more plausible explanation would be that subgroup analyses, especially in a small study population (*i.e.* 450 cases and 450 controls in the study of Wu *et al.*), could have easily given rise to false positive findings. Therefore, one should be careful when reporting positive findings from multiple small subgroup comparisons and always use appropriate levels of statistical significance[23]. Unfortunately, in the study from Wu *et al.* there is no mention of multiple testing correction.

It was found that the rs2735383CC genotype significantly decreased the expression of the *NBS1* gene through either binding of microRNA-629 to the 3'-UTR of *NBS1* gene in lung cancer cells or the binding of microRNA-509-5p to the 3'-UTR of *NBS1* gene in colorectal cancer

cells[19,20]. Since low expression of *NBS1* may reduce the efficiency of DSB repair, this way the rs2735383CC genotype likely confers an increased lung and colorectal cancer risk. According to our study, however, the rs2735383CC genotype does not confer an increased breast cancer risk. Considering that in lung and colorectal cancer cells different microRNAs appear to be downregulating *NBS1* expression, tissue specific expression of these microRNAs may likely play a role. Besides microRNA-509 and microRNA-629, the C allele at rs2735383 has also been predicted to enhance the binding of microRNA-499 and microRNA-508 to the 3'UTR of *NBS1*[21]. However, if these microRNAs are not expressed in normal breast tissue, the CC genotype of rs2735383 will not associate with breast cancer risk as *NBS1* cannot be downregulated by any of these microRNAs. At least in breast cancer cells, none of these microRNAs, except microRNA-629, are expressed at substantial levels (source: TCGA Research Network; http://cancergenome.nih.gov/). Since microRNAs are often deregulated between normal tissue and cancer tissue, this does not necessarily represent the situation in normal breast cells. Thus, further evaluation of the miRNA expression levels in normal (breast) tissue, but also their correlation with the genotype at rs2735383 should provide more insight for the tissue specificity of rs2735383 and cancer risk.

Importantly, in contrast to lung and colorectal cancer susceptibility, the results of this study do not support the presence of an association (*i.e.* OR>1.04 for Europeans and OR>1.11 for Asians) between the genotype at rs2735383 in the 3'UTR of *NBS1* and breast cancer susceptibility.

Materials and Methods

Study population

RBCS cases (N=1,269) came from the database of the Clinical Genetics Department at Erasmus University Medical Centre in Rotterdam, representing the Southwestern part of the Netherlands. First, we selected families that presented with at least two cases of female breast cancer or at least one case of female breast cancer and one case of ovarian cancer in first- or second-degree relatives. In addition, at least one of these two cases needed to be diagnosed before the age of 60. For each selected family, the youngest breast cancer patient who had been tested for *BRCA1* and *BRCA2* was then assigned to be the index case and included in RBCS. Furthermore, breast cancer cases were also included if they were diagnosed either before the age of 40 years with unilateral breast cancer or before 50 years of age with bilateral breast cancer and did not report a family history of either breast or ovarian cancer in a first or second degree relative. All index cases and their tested relatives did not carry a *BRCA1* or *BRCA2* mutation. The median age of the RBCS cases was 44 years (range 18-92 years). RBCS controls (N=1,159) came from the same database and geographic location as the RBCS cases and included women from cystic fibrosis families who were either spouses of individuals at risk of being carrier of a *CFTR* mutation or individuals who were tested negative for a *CFTR* mutation. The median age of the RBCS controls was 41 years (range 10-97 years).

BCAC consists of case-control studies of unrelated women[24]. For the purpose of the current analyses, only studies with participants of European and Asian ancestry were included, resulting in a total of 45 case-control studies (Supplementary Table S3). Studies with participants of African ancestry (*i.e.* two studies) were not included because power in the analyses would be low due to a relatively low MAF (*i.e.* 0.123) and small amount of cases (*i.e.* 1,046). Each study was approved by its relevant governing research ethics committee and all study participants provided written informed consent. The experimental protocol was approved by the Medical

Ethical Committee of the Erasmus Medical Center Rotterdam and the study was carried out in accordance with the Code of Conduct of the Federation of Medical Scientific Societies in the Netherlands (http://www.fmwv.nl). Following genotyping on the iCOGS array[24], quality control exclusions (described below), and analysis-specific exclusions, data from the following women were available for analysis: 47,640 patients with invasive breast cancer and 46,656 controls, totaling 94,296 BCAC participants.

PCR-based RFLP analysis

A 324 bp fragment of the 3'UTR of *NBS1* including rs2735383 was amplified in a duplex PCR reaction together with a 713 bp fragment of the *LRRC4* gene. Primers for *NBS1* and *LRRC4* were present in the PCR reaction at a final concentration of 0.25 and 1 µM, respectively, and sequences are available in Supplementary Table S4. The amplified *LRRC4* fragment served as an internal digestion control and generated two fragments of 549 and 164 bp upon complete digestion with ScfI (New England Biolabs, Frankfurt am Main, Germany). The 324 bp amplified *NBS1* fragment was only digested when the major G allele was present at rs2735383, thereby generating two fragments of 233 and 91 bp. Thus upon successful digestion with ScfI, samples with rs2735383 GG generated four fragments, samples with rs2735383 GC generated five fragments and samples with rs2735383 CC generated three fragments (Fig. 2).

iCOGS genotyping and imputation

Genotyping of BCAC studies was performed previously using the custom iCOGS Illumina Infinium iSelect BeadChip[24]. Briefly, DNA samples from 114,255 BCAC participants were genotyped, along with HapMap2 DNAs for European, African, and Asian populations. Raw

intensity data files underwent centralized genotype calling and quality control[24]. The HapMap2 samples were used to identify women with predicted European and Asian ancestry by performing principal component (PC) analysis using a set of over 37,000 unlinked markers[25]. Nine European PCs and two Asian PCs were found to control adequately for residual population stratification in BCAC data. Samples with a low conversion rate, extreme heterozygosity, non-female sex, or one of a first-degree relative pair were excluded. Variants were excluded if they were monomorphic or had a call rate <95% (*i.e.* when MAF >0.05) or <99% (*i.e.* when MAF <0.05), deviation from HWE (*i.e.* P <10-7), or >2% duplicate discordance.

Imputation of genotypes was performed using 1000 Genomes Project data (v3 April 2012 release) as the reference panel[26,27]. To improve computation efficiency we used a two-step procedure which involved pre-phasing by chromosome and by chunk using SHAPEIT software in the first step[28] and imputation of the phased data using IMPUTE version 2 software in the second[29]. *NBS1* rs2735383 was imputed with an imputation $r^2 > 0.999$ in both Europeans and Asians.

Statistical analyses

The association between *NBS1* rs2735383 and invasive breast cancer risk was evaluated by logistic regression analysis providing ORs and 95% CIs. In the analyses of BCAC studies, ORs were adjusted for study, age, and PCs. In the analyses of RBCS, ethnicity was not a confounding factor thus reported ORs were unadjusted for PCs. For the European and Asian BCAC studies, we additionally performed study-specific logistic regression analysis adjusting for age and PCs, and pooled the log ORs in a fixed-effects meta-analysis. Subgroup analyses within the European BCAC studies were based on age (*i.e.* ≤ 50 years and > 50 years), age at menarche (*i.e.* ≤ 13

years and > 13 years), age at menopause (*i.e.* ≤ 50 years and > 50 years), menopausal status (*i.e.* premenopausal and postmenopausal), number of full-term pregnancies (*i.e.* ≤ 2 and > 2), breast feeding (*i.e.* no and yes), first-degree family history of breast cancer and receptor status (*i.e.* ER positive, ER negative and triple negative). Clinical and demographic characteristics of the BCAC cases are presented in Supplementary Table S5. Association between *NBS1* rs2735383 and the clinical and demographic characteristics were evaluated using a χ^2 test. All *P*-values were two-sided and $P<0.05$ was considered to be statistically significant after correction for multiple testing by the Bonferroni procedure. Logistic regression analyses were performed using SPSS statistics version 23 (IBM Corporation, Armonk, NY) and fixed-effects meta-analyses using Stata version 13 (StataCorp, College Station, TX).

References

1 Khanna, K. K. & Jackson, S. P. DNA double-strand breaks: signaling, repair and the cancer connection. *Nat Genet* **27**, 247-254 (2001).

2 Stracker, T. H. & Petrini, J. H. The MRE11 complex: starting from the ends. *Nat Rev Mol Cell Biol* **12**, 90-103 (2011).

3 Sung, P. & Klein, H. Mechanism of homologous recombination: mediators and helicases take on regulatory functions. *Nat Rev Mol Cell Biol* **7**, 739-750 (2006).

4 van Gent, D. C. & van der Burg, M. Non-homologous end-joining, a sticky affair. *Oncogene* **26**, 7731-7740 (2007).

5 Lavin, M. F. ATM and the Mre11 complex combine to recognize and signal DNA double-strand breaks. *Oncogene* **26**, 7749-7758 (2007).

6 Lee, J. H. & Paull, T. T. Activation and regulation of ATM kinase activity in response to DNA double-strand breaks. *Oncogene* **26**, 7741-7748 (2007).

7 Hollestelle, A., Wasielewski, M., Martens, J. W. & Schutte, M. Discovering moderate-risk breast cancer susceptibility genes. *Curr Opin Genet Dev* **20**, 268-276 (2010).

8 Easton, D. F. *et al.* Gene-panel sequencing and the prediction of breast-cancer risk. *N Engl J Med* **372**, 2243-2257 (2015).

9 Mantere, T. *et al.* Targeted Next-Generation Sequencing Identifies a Recurrent Mutation in MCPH1 Associating with Hereditary Breast Cancer Susceptibility. *PLoS Genet* **12**, e1005816 (2016).

10 Varon, R. *et al.* Nibrin, a novel DNA double-strand break repair protein, is mutated in Nijmegen breakage syndrome. *Cell* **93**, 467-476 (1998).

11 di Masi, A. & Antoccia, A. NBS1 Heterozygosity and Cancer Risk. *Curr Genomics* **9**, 275-281 (2008).

12 Varon, R. *et al.* Clinical ascertainment of Nijmegen breakage syndrome (NBS) and prevalence of the major mutation, 657del5, in three Slav populations. *Eur J Hum Genet* **8**, 900-902 (2000).

13 Gao, P., Ma, N., Li, M., Tian, Q. B. & Liu, D. W. Functional variants in NBS1 and cancer risk: evidence from a meta-analysis of 60 publications with 111 individual studies. *Mutagenesis* **28**, 683-697 (2013).

14 Wang, F., Zhao, S. & Qin, H. NBS1 8360G > C polymorphism is associated with breast cancer risk was not credible: appraisal of a recent meta-analysis. *Breast Cancer Res Treat* **128**, 291-292 (2011).

15 Yao, F., Fang, Y., Chen, B., Jin, F. & Wang, S. Association between the NBS1 Glu185Gln polymorphism and breast cancer risk: a meta-analysis. *Tumour Biol* **34**, 1255-1262 (2013).

16 He, Y. Z. *et al.* NBS1 Glu185Gln polymorphism and cancer risk: update on current evidence. *Tumour Biol* **35**, 675-687 (2014).

17 Fang, W. *et al.* The functional polymorphism of NBS1 p.Glu185Gln is associated with an increased risk of lung cancer in Chinese populations: case-control and a meta-analysis. *Mutat Res* **770**, 61-68 (2014).

18 Zhang, Y. *et al.* NBS1 Glu185Gln polymorphism and susceptibility to urinary system cancer: a meta-analysis. *Tumour Biol* **35**, 10723-10729 (2014).

19 Yang, L. *et al.* A functional polymorphism at microRNA-629-binding site in the 3'-untranslated region of NBS1 gene confers an increased risk of lung cancer in Southern and Eastern Chinese population. *Carcinogenesis* **33**, 338-347 (2012).

20 Li, J. T. *et al.* Associations between NBS1 Polymorphisms and Colorectal Cancer in Chinese Population. *PLoS One* **10**, e0132332 (2015).

21 Wu, Z. *et al.* Evaluation of miRNA-binding-site SNPs of MRE11A, NBS1, RAD51 and RAD52 involved in HRR pathway genes and risk of breast cancer in China. *Mol Genet Genomics* **290**, 1141-1153 (2015).

22 Han, J. *et al.* Genetic variation in DNA repair pathway genes and premenopausal breast cancer risk. *Breast Cancer Res Treat* **115**, 613-622 (2009).

23 Kivimaki, M. *et al.* Don't let the truth get in the way of a good story: an illustration of citation bias in epidemiologic research. *Am J Epidemiol* **180**, 446-448 (2014).

24 Michailidou, K. *et al.* Large-scale genotyping identifies 41 new loci associated with breast cancer risk. *Nat Genet* **45**, 353-361 (2013).

25 Price, A. L. *et al.* Principal components analysis corrects for stratification in genome-wide association studies. *Nat Genet* **38**, 904-909 (2006).

26 Michailidou, K. *et al.* Genome-wide association analysis of more than 120,000 individuals identifies 15 new susceptibility loci for breast cancer. *Nat Genet* **47**, 373-380 (2015).

27 Abecasis, G. R. *et al.* An integrated map of genetic variation from 1,092 human genomes. *Nature* **491**, 56-65 (2012).

28 Delaneau, O., Marchini, J. & Zagury, J. F. A linear complexity phasing method for thousands of genomes. *Nat Methods* **9**, 179-181 (2012).

29 Howie, B., Fuchsberger, C., Stephens, M., Marchini, J. & Abecasis, G. R. Fast and accurate genotype imputation in genome-wide association studies through pre-phasing. *Nat Genet* **44**, 955-959 (2012).

NBCS Collaborators are Kristine K. Sahlberg[66,72], Lars Ottestad[66], Rolf Kåresen[73,74], Anita Langerød[66], Ellen Schlichting[74], Marit Muri Holmen[75], Toril Sauer[73,76], Vilde Haakensen[66], Olav Engebråten[73,77,78], Bjørn Naume[78], Cecile E. Kiserud[79,80], Kristin V. Reinertsen[79,80], Åslaug Helland[66,78], Margit Riis[74], Ida Bukholm[80,81], Per Eystein Lønning[82,83], Anne-Lise Børresen-Dale[66] and Grethe I. Grenaker Alnæs[66].

Acknowledgements

We thank all the individuals who took part in these studies and all the researchers, clinicians, technicians and administrative staff who have enabled this work to be carried out. This study would not have been possible without the contributions of the following: Paul Pharoah (BCAC), Andrew Berchuck (OCAC), Rosalind A. Eeles, Ali Amin Al Olama, Zsofia Kote-Jarai and Sara Benlloch (PRACTICAL), Antonis Antoniou, Lesley McGuffog and Ken Offit (CIMBA), Andrew Lee, Ed Dicks, Craig Luccarini and the staff of the Centre for Genetic Epidemiology Laboratory, the staff of the CNIO genotyping unit, Francois Bacot, Daniel Vincent, Sylvie La Boissière and Frederic Robidoux and the staff of the McGill University and Génome Québec Innovation Centre, Sune F. Nielsen, Borge G. Nordestgaard, and the staff of the Copenhagen DNA laboratory, and Julie M. Cunningham, Sharon A. Windebank, Christopher A. Hilker, Jeffrey Meyer and the staff of Mayo Clinic Genotyping Core Facility; **ABCFS:** Maggie Angelakos, Judi Maskiell, Gillian Dite; **ABCS:** Blood bank Sanquin (The Netherlands); **ACP:** The ACP study wishes to thank the participants in the Thai Breast Cancer study. Special Thanks also go to the Thai Ministry of Public Health (MOPH), doctors and nurses who helped with the data collection process. Finally, the study would like to thank Dr Prat Boonyawongviroj, the former Permanent Secretary of MOPH and Dr Pornthep Siriwanarungsan, the Department Director-General of Disease Control who have supported the study throughout; **BBCS:** Eileen Williams, Elaine Ryder-Mills, Kara Sargus; **BIGGS:** Niall McInerney, Gabrielle Colleran, Andrew Rowan, Angela Jones; **BSUCH:** Peter Bugert, Medical Faculty Mannheim; **CGPS:** Staff and participants of the Copenhagen General Population Study. For the excellent technical assistance: Dorthe Uldall Andersen, Maria Birna Arnadottir, Anne Bank, Dorthe Kjeldgård Hansen. The Danish Cancer Biobank is acknowledged for providing infrastructure for the collection of blood samples for the cases; **CNIO-BCS:** Guillermo Pita, Charo Alonso, Nuria

Álvarez, Pilar Zamora, Primitiva Menendez, the Human Genotyping-CEGEN Unit (CNIO); **CTS:** The CTS Steering Committee includes Leslie Bernstein, Susan Neuhausen, James Lacey, Sophia Wang, Huiyan Ma, and Jessica Clague DeHart at the Beckman Research Institute of City of Hope, Dennis Deapen, Rich Pinder, and Eunjung Lee at the University of Southern California, Pam Horn-Ross, Peggy Reynolds, Christina Clarke Dur and David Nelson at the Cancer Prevention Institute of California, Hoda Anton-Culver, Argyrios Ziogas, and Hannah Park at the University of California Irvine, and Fred Schumacher at Case Western University. Hartwig Ziegler, Sonja Wolf, Volker Hermann, Christa Stegmaier, Katja Butterbach; **GENICA:** The GENICA Network: Dr. Margarete Fischer-Bosch-Institute of Clinical Pharmacology, Stuttgart, and University of Tübingen, Germany [HB, Wing-Yee Lo, Christina Justenhoven], German Cancer Consortium (DKTK) and German Cancer Research Center (DKFZ) [HB], Department of Internal Medicine, Evangelische Kliniken Bonn gGmbH, Johanniter Krankenhaus, Bonn, Germany [Yon-Dschun Ko, Christian Baisch], Institute of Pathology, University of Bonn, Germany [Hans-Peter Fischer], Molecular Genetics of Breast Cancer, Deutsches Krebsforschungszentrum (DKFZ), Heidelberg, Germany, Institute for Prevention and Occupational Medicine of the German Social Accident Insurance, Institute of the Ruhr University Bochum (IPA), Bochum, Germany [Thomas Brüning, Beate Pesch, Sylvia Rabstein, Anne Lotz]; and Institute of Occupational Medicine and Maritime Medicine, University Medical Center Hamburg-Eppendorf, Germany [Volker Harth]; **HEBCS:** Kristiina Aittomäki, Sofia Khan, Taru A. Muranen, Kirsimari Aaltonen, Karl von Smitten, Irja Erkkilä; **HMBCS:** Peter Hillemanns, Hans Christiansen and Johann H. Karstens; **pKARMA:** The Swedish Medical Research Counsel; **KBCP:** Eija Myöhänen, Helena Kemiläinen; **kConFab/AOCS:** We wish to thank Heather Thorne, Eveline Niedermayr, all the kConFab research nurses and staff, the heads

and staff of the Family Cancer Clinics, and the Clinical Follow Up Study (which has received funding from the NHMRC, the National Breast Cancer Foundation, Cancer Australia, and the National Institute of Health (USA)) for their contributions to this resource, and the many families who contribute to kConFab; **LAABC:** We thank all the study participants and the entire data collection team, especially Annie Fung and June Yashiki; **LMBC:** Gilian Peuteman, Thomas Van Brussel, EvyVanderheyden and Kathleen Corthouts; **MARIE:** Anja Rudolph, Petra Seibold, Dieter Flesch-Janys, Judith Heinz, Nadia Obi, Ursula Eilber, Muhabbet Celik; **MBCSG:** Paolo Peterlongo of IFOM, the FIRC Institute of Molecular Oncology; Siranoush Manoukian, Bernard Peissel, Jacopo Azzollini, Daniela Zaffaroni and Lidia Pezzani of the Fondazione IRCCS Istituto Nazionale dei Tumori (INT); Bernardo Bonanni, and Irene Feroce of the Istituto Europeo di Oncologia (IEO) and the personnel of the Cogentech Cancer Genetic Test Laboratory; **MTLGEBCS:** We would like to thank Martine Tranchant (CHU de Québec Research Center), Marie-France Valois, Annie Turgeon and Lea Heguy (McGill University Health Center, Royal Victoria Hospital; McGill University) for DNA extraction, sample management and skillful technical assistance. J.S. is Chairholder of the Canada Research Chair in Oncogenetics. We thank study partcipants and research staff (particularly Patsy Ng, Nurhidayu Hassan, Yoon Sook-Yee, Daphne Lee, Lee Sheau Yee, Phuah Sze Yee and Norhashimah Hassan) for their contributions and commitment to this study; **NBCS:** OSBREAC Investigators: Prof. Solveig Hofvind, PhD (Cancer Registry of Norway, Oslo, Norway, Oslo and Akershus University College of Applied Sciences, Faculty of Health Science, Oslo, Norway), Prof. Tone F. Bathen, PhD (Department of Circulation and Medical Imaging, Norwegian University of Science and Technology NTNU), Trondheim, Norway), Dr. Elin Borgen, MD (Department of Pathology, Division of Diagnostics and Intervention, Oslo University Hospital, Oslo, Norway), Prof. Em.

Øystein Fodstad, MD (Department of Tumor Biology, Institute for Cancer Research, Oslo University Hospital, Oslo, Norway and Institute for Clinical Medicine, Faculty of Medicine, University of Oslo, Oslo, Norway), Dr. Øystein Garred, MD (Department of Pathology, Oslo University Hospital, Oslo, Norway), Gry Aarum Geitvik, Unit leader (Department of Cancer Genetics, Institute for Cancer Research, Oslo University Hospital-Radiumhospitalet, Oslo, Norway), Prof. Gunhild Mari Mælandsmo, PhD (Department of Tumor Biology, Institute for Cancer Research, Oslo University Hospital, Oslo, Norway, Department of Pharmacy, Faculty of Health Sciences, University of Tromsø, Tromsø, Norway), Dr. Hege G Russnes, MD (Department of Cancer Genetics, Institute for Cancer Research, Oslo University Hospital-Radiumhospitalet, Oslo, Norway and Department of Pathology, Oslo University Hospital, Oslo, Norway), Dr. Therese Sørlie, PhD (Department of Cancer Genetics, Institute for Cancer Research, Oslo University Hospital-Radiumhospitalet, Oslo, Norway), Prof. Ole Christian Lingjærde, PhD (Centre for Cancer Biomedicine, University of Oslo, Oslo, Norway and Department of Computer Science, University of Oslo, Oslo, Norway), Dr. Helle Kristine Skjerven, MD (Breast and Endocrine Surgery, Department of Breast and Endocrine Surgery, Vestre Viken Hospital, Drammen, Norway), Dr. Britt Fritzman, MD (Østfold Hospital, Østfold, Norway); **NBHS:** We thank study participants and research staff for their contributions and commitment to this study; **OBCS:** We thank Arja Jukkola-Vuorinen, Mervi Grip, Saila Kauppila, Meeri Otsukka, Leena Keskitalo and Kari Mononen for their contributions to this study; **OFBCR:** Teresa Selander, Nayana Weerasooriya; **PBCS:** Louise Brinton, Mark Sherman, Neonila Szeszenia-Dabrowska, Beata Peplonska, Witold Zatonski, Pei Chao, Michael Stagner; **RBCS:** Petra Bos, Jannet Blom, Ellen Crepin, Elisabeth Huijskens, Anja Kromwijk-Nieuwlaat, Annette Heemskerk, the Erasmus MC Family Cancer Clinic; **SASBAC:** The Swedish Medical

Research Counsel; **SBCGS:** We thank study partcipants and research staff for their contributions and commitment to this study; **SBCS:** Sue Higham, Helen Cramp, Dan Connley, Ian Brock, Sabapathy Balasubramanian and Malcolm W.R. Reed; **SEARCH:** The SEARCH and EPIC teams; **SGBCC:** We thank the participants and research coordinator Ms Tan Siew Li; **SKKDKFZS:** We thank all study participants, clinicians, family doctors, researchers and technicians for their contributions and commitment to this study; **SZBCS:** Ewa Putresza; **UKBGS:** We thank Breast Cancer Now and the Institute of Cancer Research for support and funding of the Breakthrough Generations Study, and the study participants, study staff, and the doctors, nurses and other health care providers and health information sources who have contributed to the study. We acknowledge NHS funding to the Royal Marsden/ICR NIHR Biomedical Research Centre.

Funding

This study was funded by the Cancer Genomics Netherlands (CGC.nl) and a grant for the Netherlands Organization of Scientific Research (NWO). J Liu received a scholarship from the China Scholarship Council (Beijing, China).

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. J.L.H. is a National Health and Medical Research Council (NHMRC) Senior Principal Research Fellow. M.C.S. is a NHMRC Senior Research Fellow. The **ABCS** study was supported by the Dutch Cancer Society [grants NKI 2007-3839; 2009 4363]. The **ACP** study is funded by the Breast Cancer Research Trust, UK. The work of the **BBCC** was partly funded by ELAN-Fond of the University Hospital of Erlangen. The **BBCS** is funded by Cancer Research UK and Breast Cancer Now and acknowledges NHS funding to the NIHR Biomedical Research Centre, and the National Cancer Research Network (NCRN). ES 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 **BSUCH** study was supported by the Dietmar-Hopp Foundation, the Helmholtz Society and the German Cancer Research Center (DKFZ). The **CECILE** study was supported by Fondation de France,

Institut National du Cancer (INCa), Ligue Nationale contre le Cancer, Agence Nationale de Sécurité Sanitaire, de l'Alimentation, de l'Environnement et du Travail (ANSES), Agence Nationale de la Recherche (ANR). The **CGPS** was supported by the Chief Physician Johan Boserup and Lise Boserup Fund, the Danish Medical Research Council, and Herlev and Gentofte Hospital. The **CNIO-BCS** was supported by the Instituto de Salud Carlos III (CEGEN) and grants from the Fondo de Investigación Sanitario (PI11/00923 and PI12/00070) with FEDER funds and H2020 (BRIDGES project). The **CTS** was initially supported by the California Breast Cancer Act of 1993 and the California Breast Cancer Research Fund (contract 97-10500) and is currently funded through the National Institutes of Health (R01 CA77398, UM1 CA164917, and U01 CA199277). Collection of cancer incidence data was supported by the California Department of Public Health as part of the statewide cancer reporting program mandated by California Health and Safety Code Section 103885. HAC receives support from the Lon V Smith Foundation (LVS39420). 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 VERDI study, which was supported by a grant from the German Cancer Aid (Deutsche Krebshilfe). The **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 **HEBCS** was financilly 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 **HERPACC** was supported by MEXT Kakenhi (No. 170150181 and 26253041) from the Ministry of Education, Science, Sports, Culture and Technology of Japan, by a Grant-in-Aid for the Third Term Comprehensive 10-Year Strategy for Cancer Control from Ministry Health, Labour and Welfare of Japan, by Health and Labour Sciences Research Grants for Research on Applying Health Technology from Ministry Health, Labour and Welfare of Japan, by National Cancer Center Research and Development Fund, and "Practical Research for Innovative Cancer Control (15ck0106177h0001)" from Japan Agency for Medical Research and development, AMED, and Cancer Bio Bank Aichi. The **HMBCS** was supported by a grant from the Friends of Hannover Medical School and by the Rudolf Bartling Foundation. The **pKARMA** study was supported by Märit and Hans Rausings Initiative Against Breast Cancer. 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. Financial support for the **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). G.C.T. and P.W. are supported by the NHMRC. RB was a Cancer Institute NSW Clinical Research Fellow. **LAABC** is supported by grants

(1RB-0287, 3PB-0102, 5PB-0018, 10PB-0098) from the California Breast Cancer Research Program. Incident breast cancer cases were collected by the USC Cancer Surveillance Program (CSP) which is supported under subcontract by the California Department of Health. The CSP is also part of the National Cancer Institute's Division of Cancer Prevention and Control Surveillance, Epidemiology, and End Results Program, under contract number N01CN25403. **LMBC** is supported by the 'Stichting tegen Kanker'. Diether Lambrechts is supported by the FWO. The **MARIE** study was supported by the Deutsche Krebshilfe e.V. [70-2892-BR I, 106332, 108253, 108419, 110826, 110828], the Hamburg Cancer Society, the German Cancer Research Center (DKFZ) and the Federal Ministry of Education and Research (BMBF) Germany [01KH0402]. **MBCSG** is supported by grants from the Italian Association for Cancer Research (AIRC) and by funds from the Italian citizens who allocated the 5/1000 share of their tax payment in support of the Fondazione IRCCS Istituto Nazionale Tumori, according to Italian laws (INT-Institutional strategic projects "5x1000"). The **MCBCS** was supported by the NIH grants CA192393, 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. **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 and the Australian Cancer Database. The **MEC** was support by NIH grants CA63464, CA54281, CA098758, CA132839 and CA164973. The work of **MTLGEBCS** was 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 # CRN-87521 and the Ministry of Economic Development, Innovation and Export Trade – grant # PSR-SIIRI-701. **MYBRCA** is funded by research grants from the Malaysian Ministry of Higher Education (UM.C/HlR/MOHE/06) and Cancer Research Malaysia. MYMAMMO is supported by research grants from Yayasan Sime Darby LPGA Tournament and Malaysian Ministry of Higher Education (RP046B-15HTM). The **NBCS** has been supported by the Research Council of Norway grant 193387/V50 (to A-L Børresen-Dale and V.N. Kristensen) and grant 193387/H10 (to A-L Børresen-Dale and V.N. Kristensen), South Eastern Norway Health Authority (grant 39346 to A-L Børresen-Dale and 27208 to V.N.Kristensen) and the Norwegian Cancer Society (to A-L Børresen-Dale and 419616 - 71248 - PR-2006-0282 to V.N. Kristensen). It has received funding from the K.G. Jebsen Centre for Breast Cancer Research (2012-2015). The **NBHS** was supported by NIH grant R01CA100374. Biological sample preparation was conducted the Survey and Biospecimen Shared Resource, which is supported by P30 CA68485. The **OBCS** was supported by research grants from the Finnish Cancer Foundation, the Academy of Finland (grant number 250083, 122715 and Center of Excellence grant number 251314), the Finnish Cancer Foundation, the Sigrid Juselius Foundation, the University of Oulu, the University of Oulu Support Foundation and the special Governmental EVO funds for Oulu University Hospital-based research activities. 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 **PBCS** was funded by Intramural Research Funds of the National

Cancer Institute, Department of Health and Human Services, USA. The **RBCS** was funded by the Dutch Cancer Society (DDHK 2004-3124, DDHK 2009-4318). 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 **SBCGS** was supported primarily by NIH grants R01CA64277, R01CA148667, UMCA182910, and R37CA70867. Biological sample preparation was conducted the Survey and Biospecimen Shared Resource, which is supported by P30 CA68485. The scientific development and funding of this project were, in part, supported by the Genetic Associations and Mechanisms in Oncology (GAME-ON) Network U19 CA148065. The **SBCS** was supported by Sheffield Experimental Cancer Medicine Centre and Breast Cancer Now. **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. **SEBCS** was supported by the BRL (Basic Research Laboratory) program through the National Research Foundation of Korea funded by the Ministry of Education, Science and Technology (2012-0000347). **SGBCC** is funded by the NUS start-up Grant, National University Cancer Institute Singapore (NCIS) Centre Grant and the NMRC Clinician Scientist Award. Additional controls were recruited by the Singapore Consortium of Cohort Studies-Multi-ethnic cohort (SCCS-MEC), which was funded by the Biomedical Research Council, grant number: 05/1/21/19/425. **SKKDKFZS** is supported by the DKFZ. The **SZBCS** was supported by Grant PBZ_KBN_122/P05/2004. The **TBCS** was funded by The National Cancer Institute Thailand. The **TNBCC** was supported by: a Specialized Program of Research Excellence (SPORE) in Breast Cancer (CA116201), a grant from the Breast Cancer Research Foundation, a generous gift from the David F. and Margaret T. Grohne Family Foundation; the Hellenic Cooperative

Oncology Group research grant (HR R_BG/04) and the Greek General Secretary for Research and Technology (GSRT) Program, Research Excellence II, the European Union (European Social Fund – ESF), and Greek national funds through the Operational Program "Education and Lifelong Learning" of the National Strategic Reference Framework (NSRF) - ARISTEIA. The **TWBCS** is supported by the Taiwan Biobank project of the Institute of Biomedical Sciences, Academia Sinica, Taiwan. The **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.

Author contributions

J.Liu. designed and performed the experiments, analyzed the data, wrote the paper and provided funding. I.L. designed and performed the experiments and analyzed the data. J.M.C. designed the study, contributed samples and provided clinical data. M.K.B. and Q.W. provided database management, contributed samples and provided clinical data. J.D. and K.M. analyzed the data, contributed samples and provided clinical data. I.A., M.B., M.W.B., S.B., C.B., B.B., N.V.B., S.E.B., H.Brauch, P.B., A.B., B.B., S-T.C., C.Y.C., J-Y.C., F.J.C., A.C., S.S.C., K.C., K.C., I.d-S-S., P.A.F., J.F., H.F., M.G-C., G.G.G., G.G., M.S.G., P.G., C.A.H., S.N.H., M.H., S.H., H.I., A.J., M.K., D.K., V-M.K., V.N.K., L.L.M., E.L., J.Li, A.L., J.Lubinski, A.M., K.M., R.L.M., NBCS Collaborators, S.L.N., H.N., N.O., J.I.A.P., J.P., T.C.P., K.P., P.R., S.S., E.J.S., A.S., C-Y.S., M.J.S., X-O.S., M.C.S., S.H.T., S.T., I.T., D.T., T.T., C-C.T., C.V., R.W., A.H.W., D.Y. and W.Z. contributed samples and provided clinical data. H.Brenner, J.C-C., J.L.H., G.C-T., T.D., U.H., M.K.S. and A.S. contributed samples, provided clinical data and revised the manuscript. A.G-N., J.B. and D.C.T. coordinated the study, contributed samples and provided

clinical data. A.M.D. designed and coordinated the study, contributed samples and provided clinical data. J.S designed and coordinated the study, contributed samples, provided clinical data and funding. P.H. led and designed the study, contributed samples, provided clinical data and funding. D.F.E. conceived of the study, designed the study, contributed samples, provided clinical data and funding. M.J.H. designed the study, provided clinical data and revised the manuscript. A.M.W.O. designed the study, contributed samples and provided clinical data. J.W.M.M. designed the study, revised the manuscript and provided funding. A.H. conceived of the study, designed the study and the experiments, analyzed the data and wrote the paper. All authors read and approved the final version of the manuscript.

Additional information

Supplementary information accompanies this paper at http://www.nature.com/srep.

The authors declare no competing financial interests.

Legends to Figures

Figure 1. Forest plots for the association between rs2735383 and breast cancer risk. A) for the 36 European BCAC studies and B) for the nine Asian BCAC studies. Study-specific ORs (squares) were from a recessive genetic model and adjusted by age and PCs. Overall or pooled ORs (diamonds) were from a fixed-effects meta-analysis.

Figure 2. Microchip electrophoresis of the RFLP-PCR products of 23 RBCS cases. After PCR amplification of *NBS1* and *LRRC4* fragments (*i.e.* 324 and 713 bp), digestion with SfcI generated four fragments (*i.e.* 549, 233, 164 and 91 bp) for samples with rs2735383 GG genotypes, five fragments (*i.e.* 549, 324, 233, 164 and 91 bp) for samples with rs2735383 GC genotypes and three fragments (*i.e.* 549, 324 and 164 bp) for samples with rs2735383 CC genotypes. UM, upper marker; LM, lower marker.

Table 1 Association of *NBS1* rs2735383 with breast cancer risk in the RBCS study

Genetic model	N Controls	N Cases	OR (95% CI)	*P*-value
Recessive				
GG + GC	963	1023	1	
CC	114	147	1.214 (0.936-1.574)	0.144
	1077	1170		

N, number of; MAF, minor allele frequency; OR, odds ratio; CI, confidence interval.

Table 2 Association of *NBS1* rs2735383 with overall breast cancer risk in the European and Asian BCAC studies

Ethnicity	Genetic model	N Controls	N Cases	MAF Controls	MAF Cases	OR (95% CI)*	*P*-value*
European		40,0042	41,915	33.44%	33.39%		
	Recessive					1.014 (0.969-1.060)	0.556
	Dominant					1.006 (0.978-1.035)	0.684
	Additive					1.000 (0.979-1.021)	0.984
Asian		6,614	5,725	40.71%	40.58%		
	Recessive					0.998 (0.905-1.100)	0.961
	Dominant					0.995 (0.922-1.074)	0.900
	Additive					0.997 (0.946-1.050)	0.911

N, number of; MAF, minor allele frequency; OR, odds ratio; CI, confidence interval.
*adjusted for age, study and principal components (PCs). In the European analyses nine PCs were added to the regression model and in the Asian analyses two PCs.

Table 3 Subgroup analysis of *NBS1* rs2735383 and breast cancer risk in the European BCAC studies

Subgroup	N Controls	N Cases	MAF Controls	MAF Cases	OR (95% CI)*	*P*-value*
Age						
≤ 50 years	13,055	13,362	33.76%	33.41%	0.977 (0.899-1.062)	0.581
> 50 years	26,987	28,553	33.28%	33.38%	1.026 (0.971-1.084)	0.356
Age at menarche						
≤ 13 years	14,312	13,843	33.72%	33.13%	0.984 (0.914-1.060)	0.677
> 13 years	8,964	8,095	32.65%	33.63%	1.077 (0.978-1.187)	0.131
Age at menopause						
≤ 50 years	5,571	7,288	32,79%	33.43%	1.019 (0.906-1.146)	0.755
> 50 years	3,366	4,262	33.50%	33.24%	0.993 (0.855-1.154)	0.926
Menopausal status						
Premenopausal	8,974	7,412	33.66%	33.07%	0.981 (0.887-1.085)	0.715
Postmenopausal	19,648	17,353	33.39%	33.56%	1.007 (0.943-1.075)	0.844
Number of full-term pregnancies						
≤ 2	21,008	19,722	33.53%	33.20%	1.004 (0.942-1.071)	0.893
> 2	8,258	7,327	33.26%	33.44%	1.032 (0.931-1.144)	0.549
Breast feeding						
No	6,849	6,805	33.36%	33.25%	0.988 (0.884-1.104)	0.828
Yes	11,947	12,709	33.68%	33.28%	0.978 (0.903-1.060)	0.594
Family history						
1st degree relative with BC	23,648	4,119	33.21%	32.68%	0.990 (0.884-1.108)	0.859
Receptor status						
ER positive	39,699	25,959	33.47%	33.40%	1.021 (0.970-1.075)	0.427
ER negative	39,618	6,774	33.42%	32.90%	0.991 (0.908-1.082)	0.846
Triple negative	30,696	2,712	33.10%	32.04%	0.980 (0.847-1.134)	0.788

N, number; MAF, minor allele frequency; OR, odds ratio; CI, confidence interval; ER, estrogen receptor; BC, breast cancer

*recessive genetic model adjusted for age, study and nine principal components.