Author/s:
Shi, J;Zhang, Y;Zheng, W;Michailidou, K;Ghoussaini, M;Bolla, MK;Wang, Q;Dennis, J;Lush, M;Milne, RL;Shu, XO;Beesley, J;Kar, S;Andrulis, IL;Anton-Culver, H;Arndt, V;Beckmann, MW;Zhao, Z;Guo, X;Benitez, J;Beeghly-Fadiel, A;Blot, W;Bogdanova, NV;Bojesen, SE;Brauch, H;Brenner, H;Brinton, L;Broeks, A;Brüning, T;Burwinkel, B;Cai, H;Canisius, S;Chang-Claude, J;Choi, JY;Couch, FJ;Cox, A;Cross, SS;Czene, K;Darabi, H;Devilee, P;Droit, A;Dork, T;Fasching, PA;Fletcher, O;Flyger, H;Fostira, F;Gaborieau, V;García-Closas, M;Giles, GG;Grip, M;Guenel, P;Haiman, CA;Hamann, U;Hartman, M;Miao, H;Hollestelle, A;Hopper, JL;Hsiung, CN;Ito, H;Jakubowska, A;Johnson, N;Torres, D;Kabisch, M;Kang, D;Khan, S;Knight, JA;Kosma, VM;Lambrechts, D;Li, J;Lindblom, A;Lophatananon, A;Lubinski, J;Mannermaa, A;Manoukian, S;Le Marchand, L;Margolin, S;Marme, F;Matsuo, K;McLean, C;Meindl, A;Muir, K;Neuhausen, SL;Nevanlinna, H;Nord, S;Børresen-Dale, AL;Olson, JE;Orr, N;van den Ouweland, AMW;Peterlongo, P;Choudary Putti, T;Rudolph, A;Sangrajrang, S;Sawyer, EJ;Schmidt, MK;Schmutzler, RK;Shen, CY;Hou, MF;Shrubsole, MJ;Southey, MC

Title:
Fine-scale mapping of 8q24 locus identifies multiple independent risk variants for breast cancer

Date:
2016-09-15

Citation:
Shi, J., Zhang, Y., Zheng, W., Michailidou, K., Ghoussaini, M., Bolla, M. K., Wang, Q., Dennis, J., Lush, M., Milne, R. L., Shu, X. O., Beesley, J., Kar, S., Andrulis, I. L., Anton-Culver, H., Arndt, V., Beckmann, M. W., Zhao, Z., Guo, X. ,... Southey, M. C. (2016). Fine-scale mapping of 8q24 locus identifies multiple independent risk variants for breast cancer. International Journal of Cancer, 139 (6), pp.1303-1317. https://doi.org/10.1002/ijc.30150.

Persistent Link:
https://hdl.handle.net/11343/291423

	Brenner, Hermann; German Cancer Research Center, Division of Clinical Epidemiology and Aging Research; German Cancer Research Center(DKFZ), German Cancer Consortium; German Cancer Research Center (DKFZ), Division of Preventive Oncology Brinton, Louise; National Cancer Institute, Division of Cancer Epidemiology and Genetics Broeks, Annegien; Antoni van Leeuwenhoek hospital, Netherlands Cancer Institute Brüning, Thomas; Institute for Prevention and Occupational Medicine of the German Social Accident Insurance, Burwinkel, Barbara; German Cancer Research Center, Division of Molecular Genetic Epidemiology; German Cancer Research Center, Molecular Epidemiology Group Cai, Hui; Vanderbilt University Medical Center, Medicine Canisius, Sander; Antoni van Leeuwenhoek Hospital, Netherlands Cancer Institute Chang-Claude, Jenny; German Cancer Research Center, Division of Cancer Epidemiology and Genetics Choi, Ji-Yeob; Seoul National University College of Medicine, Biomedical Sciences; Seoul National University College of Medicine, Cancer Research Institute Couch, Fergus; Mayo Clinic, Laboratory Medicine and Pathology Cox, Angela; University of Sheffield, Oncology Cross, Simon; University of Sheffield, Neuroscience Czene, Kamila; Karolinska Institutet, Medical Epidemiology and Biostatistics Darabi, Hatef; Karolinska Institutet, Medical Epidemiology and Biostatistics Devilee, Peter; Leiden University Medical Center, Pathology; Leiden University Medical Center, Human Genetics Droit, Arnaud; Laval University, Centre Hospitalier Universitaire de Québec Research Center Dork, Thilo; Hannover Medical School, Gynaecology Research Unit Fasching, Peter; Friedrich-Alexander University Erlangen-Nuremberg, Gynaecology and Obstetrics; University of California at Los Angeles, Medicine Fletcher, Olivia; Institute of Cancer Research, Breakthrough Breast Cancer Research Centre Flyger, Henrik; Copenhagen University Hospital, Breast Surgery Fostira, Florentia; National Centre for Scientific Research "Demokritos", Gaborieau, Valerie; International Agency for Research on Cancer, García-Closas, Montserrat; Institute of Cancer Research, Breakthrough Breast Cancer Research Centre; Institute of Cancer Research, Genetics and Epidemiology Giles, Graham; The Cancer Council Victoria, Cancer Epidemiology Centre; The University of Melbourne, School of Population and Global health Grip, Mervi; Oulu University Hospital and University of Oulu, Surgery Guenel, Pascal; Center for Research in Epidemiology and Population Health, a. Environmental Epidemiology of Cancer; University Paris-Sud, Haiman, Christopher; University of Southern California, Preventive Medicine Hamann, Ute; German Cancer Research Center (DKFZ), Molecular Epidemiology Group Hartman, Mikael; National University of Singapore, Saw Swee Hock School of Public Health; National University Health System, Surgery Miao, Hui; National University of Singapore, Saw Swee Hock School of Public Health Hollestelle, Antoinette; Erasmus University Medical Center, Medical Oncology Hopper, John; Technische Universität München, Gynaecology and Obstetrics Hsiung, Chia-Ni; Academia Sinica, Institute of Biomedical Sciences

	Inverstigators, kConFab; The University of Melbourne, Ito, Hidemi; Aichi Cancer Center Research Institute, Jakubowska, Anna; Pomeranian Medical University, Genetics and Pathology Johnson, Nicola; Institute of Cancer Research, Breakthrough Breast Cancer Research Centre Torres, Diana; German Cancer Research Center (DKFZ), Molecular Epidemiology Group Kabisch, Maria; German Cancer Research Center (DKFZ), Molecular Epidemiology Group Kang, Daehee; Seoul National University College of Medicine, Biomedical Sciences; Seoul National University College of Medicine, Cancer Research Institute ; Seoul National University College of Medicine, Preventive Medicine Khan, Sofia; Helsinki University Central Hospital, Obstetrics and Gynecology Knight, Julia; Lunenfeld-Tanenbaum Research Institute of Mount Sinai Hospital, Prosserman Centre for Health Research; University of Toronto, Dalla Lana School of Public Health Kosma, Veli-Matti; University of Eastern Finland, Clinical Pathology Lambrechts, Diether; Vesalius Research Center, ; University of Leuven, Oncology Li, Jingmei; Karolinska Institutet, Medical Epidemiology and Biostatistics Lindblom, Annika; Karolinska Institutet, Molecular Medicine and Surgery Lophatananon, Artitaya; Warwick University, Warwick Medical School Lubinski, Jan; Pomeranian Medical University, Genetics and Pathology Mannermaa, Arto; University of Eastern Finland, Clinical Pathology Manoukian, Siranoush; Fondazione IRCCS Istituto Nazionale dei Tumori (INT), Preventive and Predictive Medicine Le Marchand, Loic; University of Hawaii, Margolin, Sara; Karolinska Institutet, Stockholm Marme, Frederik; University of Heidelberg, National Center for Tumor Diseases; University of Heidelberg, Obstetrics and Gynecology Matsuo, Keitaro; Kyushu University Faculty of Medical Sciences, Preventive Medicine McLean, Catriona; The Alfred Hospital, Anatomical Pathology Meindl, Alfons; Technische Universität München, Gynaecology and Obstetrics Muir, Kenneth; Warwick University, Warwick Medical School; University of Manchester, b. Institute of Population Health Neuhausen, Susan; Beckman Research Institute of City of Hope, Nevanlinna, Heli; University of Helsinki, Obstetrics and Gynecology Nord, Silje; Oslo University Hospital, Genetics; University of Oslo, Børresen-Dale, Anne-Lise; Oslo University Hospital, Genetics; University of Oslo, Olson, Janet; Health Sciences Research, Orr, Nick; Institute of Cancer Research, van den Ouweland, Ans; Erasmus University Medical Center, Clinical Genetics Peterlongo, Paolo; FIRC Institute of Molecular Oncology, Putti, Thomas; National University Health System, Pathology Rudolph, Anja; German Cancer Research Center, Sangrajrang, Suleeporn; National Cancer Institute, Sawyer, Elinor; King's College London, Research Oncology Schmidt, Marjanka; Antoni van Leeuwenhoek hospital, Schmutzler, Rita; University Hospital of Cologne, Gynaecology and Obstetrics; University Hospital of Cologne, Center for Integrated Oncology; University Hospital of Cologne, Center for Molecular Medicine; University Hospital of Cologne, Center of Familial Breast and Ovarian Cancer Shen, Chen-Yang; China Medical University, School of Public Health; Academia Sinica, Institute of Biomedical Sciences Hou, Ming-Feng; Kaohsiung Medical,

	Cai, Qiuyin; Vanderbilt University Medical Center, Medicine Long, Jirong; Vanderbilt University Medical Center, Medicine
Key Words:	breast cancer, genetic susceptibility, 8q24, fine-mapping, single nucleotide polymorphism

Fine-scale mapping of 8q24 locus identifies multiple independent risk variants for breast cancer

Jiajun Shi[1†], Yanfeng Zhang[1†], Wei Zheng[1], Kyriaki Michailidou[2], Maya Ghoussaini[3], Manjeet K. Bolla[2], Qin Wang[2], Joe Dennis[3], Michael Lush[3], Roger L. Milne[4,5], Xiao-Ou Shu[1], Jonathan Beesley[6], Siddhartha Kar[3], Irene L. Andrulis[7,8], Hoda Anton-Culver[9], Volker Arndt[10], Matthias W. Beckmann[11], Zhiguo Zhao[1], Xingyi Guo[1], Javier Benitez[12,13], Alicia Beeghly-Fadiel[1], William Blot[1,14], Natalia V. Bogdanova[15], Stig E. Bojesen[16,17,18], Hiltrud Brauch[19,20,21], Hermann Brenner[10,21,22], Louise Brinton[23], Annegien Broeks[24], Thomas Brüning[25], Barbara Burwinkel[26,27], Hui Cai[1], Sander Canisius[28], Jenny Chang-Claude[29], Ji-Yeob Choi[30,31], Fergus J. Couch[32], Angela Cox[33], Simon S. Cross[34], Kamila Czene[35], Hatef Darabi[35], Peter Devilee[36,37], Arnaud Droit[38], Thilo Dork[39], Peter A. Fasching[11,40], Olivia Fletcher[41], Henrik Flyger[42], Florentia Fostira[43], Valerie Gaborieau[44], Montserrat García-Closas[41,45], Graham G. Giles[4,5], Mervi Grip[46], Pascal Guenel[47,48], Christopher A. Haiman[49], Ute Hamann[50], Mikael Hartman[51,52], Hui Miao[51], Antoinette Hollestelle[53], John L. Hopper[54], Chia-Ni Hsiung[55], kConFab Investigators[56], Hidemi Ito[57], Anna Jakubowska[58], Nichola Johnson[41],Diana Torres[50,59], Maria Kabisch[50], Daehee Kang[30,31,60], Sofia Khan[61], Julia A. Knight[62,63], Veli-Matti Kosma[64], Diether Lambrechts[65,66], Jingmei Li[35], Annika Lindblom[67], Artitaya Lophatananon[68], Jan Lubinski[58], Arto Mannermaa[64], Siranoush Manoukian[69], Loic Le Marchand[70], Sara Margolin[71], Frederik Marme[72,73], Keitaro Matsuo[74], Catriona McLean[75], Alfons Meindl[54], Kenneth Muir[68,76], Susan L. Neuhausen[77], Heli Nevanlinna[61], Silje Nord[78,79], Anne-Lise Børresen-Dale[78,79], Janet E. Olson[80], Nick Orr[81], Ans M.W. van den Ouweland[82], Paolo Peterlongo[83], Thomas Choudary Putti[84], Anja Rudolph[29], Suleeporn Sangrajrang[85], Elinor J. Sawyer[86], Marjanka K. Schmidt[24], Rita K. Schmutzler[87,88,89,90], Chen-Yang Shen[91,92], Ming-Feng Hou[93], Matha J Shrubsole[1], Melissa C. Southey[94], Anthony Swerdlow[95], Soo Hwang Teo[96,97], Bernard Thienpont[65,66], Amanda E. Toland[98], Robert A.E.M. Tollenaar[99], Ian Tomlinson[100], Therese Truong[47,48], Chiu-chen Tseng[49], Wanqing Wen[1], Robert Winqvist[101,102], Anna H. Wu[49], Cheng Har Yip[97], Pilar M. Zamora[103], Ying Zheng[104], Giuseppe Floris[105], Ching-Yu Cheng[106], Maartje J. Hooning[53], John W.M. Martens[53], Caroline Seynaeve[53], Vessela N. Kristensen[78,79,107], Per Hall[35], Paul D.P. Pharoah[2,3], Jacques Simard[38], Georgia Chenevix-Trench[6,56], Alison M. Dunning[3], Antonis C. Antoniou[2], Douglas F. Easton[2,3], Qiuyin Cai[1*], and Jirong Long[1*]

[1]Division of Epidemiology, Department of Medicine, Vanderbilt Epidemiology Center, Vanderbilt-Ingram Cancer Center, Vanderbilt University School of Medicine, Nashville, TN 37203, USA.
[2]Centre for Cancer Genetic Epidemiology, Department of Public Health and Primary Care, University of Cambridge, Cambridge CB1 8RN, UK.

[3]Centre for Cancer Genetic Epidemiology, Department of Oncology, University of Cambridge, Cambridge CB1 8RN, UK.
[4]Cancer Epidemiology Centre, The Cancer Council Victoria, Melbourne, Victoria 3053, Australia.
[5]Centre for Epidemiology and Biostatistics, School of Population and Global health, The University of Melbourne, Melbourne, Victoria 3053, Australia.
[6]Department of Genetics, QIMR Berghofer Medical Research Institute, Brisbane, Australia.
[7]Lunenfeld-Tanenbaum Research Institute of Mount Sinai Hospital, Toronto, ON, M5G 1X5, Canada.
[8]Department of Molecular Genetics, University of Toronto, Toronto, ON, M5S 1A8, Canada.
[9]Department of Epidemiology, University of California Irvine, Irvine, CA 92697, USA.
[10]Division of Clinical Epidemiology and Aging Research, German Cancer Research Center, Heidelberg 69120, Germany.
[11]Department of Gynaecology and Obstetrics, University Hospital Erlangen, Friedrich-Alexander University Erlangen-Nuremberg, Erlangen 91054, Germany.
[12]Human Cancer Genetics Program, Spanish National Cancer Research Centre, Madrid 28029, Spain.
[13]Centro de Investigación en Red de Enfermedades Raras, Valencia, Spain.
[14]International Epidemiology Institute, Rockville, MD 20850, USA.
[15]Department of Radiation Oncology, Hannover Medical School, Hannover 30625, Germany,
[16]Copenhagen General Population Study, Herlev Hospital, 2730 Herlev, Denmark.
[17]Department of Clinical Biochemistry, Herlev Hospital, Copenhagen University Hospital, 2730 Herlev, Denmark.
[18]Faculty of Health and Medical Sciences, University of Copenhagen, Copenhagen, Denmark.
[19]Dr. Margarete Fischer-Bosch-Institute of Clinical Pharmacology, Stuttgart 70376, Germany.
[20]University of Tübingen, Tübingen 72074, Germany.
[21]German Cancer Consortium, German Cancer Research Center(DKFZ), Heidelberg 69120, Germany.
[22]Division of Preventive Oncology, German Cancer Research Center (DKFZ), Heidelberg 69120, Germany.
[23]Division of Cancer Epidemiology and Genetics, National Cancer Institute, Rockville, MD 20850, USA.
[24]Netherlands Cancer Institute, Antoni van Leeuwenhoek hospital, Amsterdam 1066 CX, The Netherlands.
[25]Institute for Prevention and Occupational Medicine of the German Social Accident Insurance, Bochum 44789, Germany.
[26]Division of Molecular Genetic Epidemiology, German Cancer Research Center, Heidelberg 69120, Germany.
[27]Molecular Epidemiology Group, German Cancer Research Center, Heidelberg 69120, Germany.
[28]Netherlands Cancer Institute, Antoni van Leeuwenhoek Hospital, 1066 CX Amsterdam, the Netherlands.
[29]Division of Cancer Epidemiology, German Cancer Research Center, Heidelberg 69120, Germany.
[30]Department of Biomedical Sciences, Seoul National University College of Medicine, Seoul 110-799, Korea.

[31]Cancer Research Institute, Seoul National University College of Medicine, Seoul 110-799, Korea.
[32]Department of Laboratory Medicine and Pathology, Mayo Clinic, Rochester, MN 55905, USA.
[33]Sheffield Cancer Research Centre, Department of Oncology, University of Sheffield, Sheffield S10 2RX, UK.
[34]Academic Unit of Pathology, Department of Neuroscience, University of Sheffield, Sheffield S10 2HQ, UK.
[35]Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Stockholm SE-17177, Sweden.
[36]Separtment of Pathology, Leiden University Medical Center, Leiden 2333 ZC, The Netherlands.
[37]Department of Human Genetics, Leiden University Medical Center, Leiden 2333 ZC, The Netherlands.
[38]Centre Hospitalier Universitaire de Québec Research Center, Laval University, Québec City G1V 4G2, Canada.
[39]Gynaecology Research Unit, Hannover Medical School, Hannover 30625, Germany.
[40]David Geffen School of Medicine, Department of Medicine Division of Hematology and Oncology, University of California at Los Angeles, Los Angeles, CA 90095, USA.
[41]Division of Cancer Studies, Breakthrough Breast Cancer Research Centre, Institute of Cancer Research, London SW3 6JB, UK.
[42]Department of Breast Surgery, Herlev Hospital, Copenhagen University Hospital, 2730 Herlev, Denmark.
[43]Molecular Diagnostics Laboratory, IRRP, National Centre for Scientific Research "Demokritos", 153 10 Athens, Greece.
[44]International Agency for Research on Cancer, Lyon 69372, France.
[45]Division of Genetics and Epidemiology, Institute of Cancer Research, London SW7 3RP, UK.
[46]Department of Surgery, Oulu University Hospital and University of Oulu, Oulu FI-90220, Finland.
[47]Environmental Epidemiology of Cancer, Center for Research in Epidemiology and Population Health, INSERM, Villejuif 94807, France.
[48]University Paris-Sud, Villejuif 94807, France.
[49]Department of Preventive Medicine, Keck School of Medicine, University of Southern California, Los Angeles, CA 90033, USA.
[50]Molecular Genetics of Breast Cancer, German Cancer Research Center (DKFZ), Heidelberg 69120, Germany.
[51]Saw Swee Hock School of Public Health, National University of Singapore, Singapore 119077, Singapore.
[52]Department of Surgery, National University Health System, Singapore 117597.
[53]Department of Medical Oncology, Erasmus MC Cancer Institute, 3015 CN Rotterdam, The Netherlands.
[54]Division of Gynaecology and Obstetrics, Technische Universität München, Munich 81675, Germany.
[55]Institute of Biomedical Sciences, Academia Sinica, Taipei 115, Taiwan.
[56]Peter MacCallum Cancer Centre, The University of Melbourne, East Melbourne, VIC 3002, Australia

[57]Division of Epidemiology and Prevention, Aichi Cancer Center Research Institute, Aichi 464-8681, Japan.
[58]Department of Genetics and Pathology, Pomeranian Medical University, Szczecin 70-115, Poland.
[59]Institute of Human Genetics, Pontificia Universidad Javeriana, Bogota 12362, Colombia
[60]Department of Preventive Medicine, Seoul National University College of Medicine, Seoul 110-799, Korea.
[61]Department of Obstetrics and Gynecology, Helsinki University Central Hospital, University of Helsinki, Helsinki, FI-00029 HUS, Finland.
[62]Prosserman Centre for Health Research, Lunenfeld-Tanenbaum Research Institute of Mount Sinai Hospital, Toronto, ON, M5G 1X5, Canada.
[63]Division of Epidemiology, Dalla Lana School of Public Health, University of Toronto, Toronto, ON, M5S 1A8, Canada.
[64]School of Medicine, Institute of Clinical Medicine, Pathology and Forensic Medicine and Cancer Center of Eastern Finland, University of Eastern Finland, Kuopio, Finland; Imaging Center, Department of Clinical Pathology, Kuopio University Hospital, Kuopio 70210, Finland.
[65]Vesalius Research Center, Leuven 3000, Belgium.
[66]Laboratory for Translational Genetics, Department of Oncology, University of Leuven, Leuven 3000, Belgium.
[67]Department of Molecular Medicine and Surgery, Karolinska Institutet, Stockholm SE-17177, Sweden.
[68]Division of Health Sciences, Warwick Medical School, Warwick University, Coventry CV4 7AL, UK.
[69]Unit of Medical Genetics, Department of Preventive and Predictive Medicine, Fondazione IRCCS Istituto Nazionale dei Tumori (INT), Milan 20133, Italy.
[70]University of Hawaii Cancer Center, Honolulu, HI 96813, USA.
[71]Department of Oncology - Pathology, Karolinska Institutet, Stockholm SE-17177, Sweden.
[72]National Center for Tumor Diseases, University of Heidelberg, Heidelberg 69120, Germany.
[73]Department of Obstetrics and Gynecology, University of Heidelberg, Heidelberg 69120, Germany.
[74]Department of Preventive Medicine, Kyushu University Faculty of Medical Sciences, Fukuoka, Japan.
[75]Anatomical Pathology, The Alfred Hospital, Melbourne, , Victoria 3004, Australia.
[76]Institute of Population Health, University of Manchester, Manchester M13 9PL, UK.
[77]Beckman Research Institute of City of Hope, Duarte, CA 91010, USA.
[78]Department of Genetics, Institute for Cancer Research, Oslo University Hospital, Radiumhospitalet, Ullernchausseen 70, N-0310 Oslo, Norway.
[79]K.G. Jebsen Center for Breast Cancer Research, Institute for Clinical Medicine, Faculty of Medicine, University of Oslo, Kirkeveien 166, 0450 Oslo, Norway.
[80]Department of Health Sciences Research, Mayo Clinic, Rochester, MN 55905, USA.
[81]Division of Breast Cancer Research, Institute of Cancer Research, London, UK; Cancer Research, Institute of Cancer Research, London SW3 6JB, UK.
[82]Department of Clinical Genetics, Erasmus University Medical Center, 3000 CA Rotterdam, The Netherlands.
[83]IFOM, the FIRC Institute of Molecular Oncology, Milan 20139, Italy.
[84]Department of Pathology, National University Health System, Singapore 117597.

[85]National Cancer Institute, Bangkok 10400, Thailand.
[86]Research Oncology, Guy's Hospital, King's College London, London SE1 9RT, UK.
[87]Division of Molecular Gyneco-Oncology, Department of Gynaecology and Obstetrics, University Hospital of Cologne, Cologne 50931, Germany.
[88]Center for Integrated Oncology, University Hospital of Cologne, Cologne 50931, Germany.
[89]Center for Molecular Medicine, University Hospital of Cologne, Cologne 50931, Germany.
[90]Center of Familial Breast and Ovarian Cancer, University Hospital of Cologne, Cologne 50931, Germany.
[91]School of Public Health, China Medical University, Taichung 404, Taiwan.
[92]Taiwan Biobank, Institute of Biomedical Sciences, Academia Sinica, Taipei 115, Taiwan.
[93]Cancer Center and Department of Surgery, Chung-Ho Memorial Hospital, Kaohsiung Medical University, Kaohsiung 807, Taiwan.
[94]Department of Pathology, The University of Melbourne, Melbourne, Victoria 3010, Australia.
[95]Division of Genetics and Epidemiology and Division of Breast Cancer Research, Institute of Cancer Research, London SW7 3RP, UK
[96]Cancer Research Initiatives Foundation, 47500 Subang Jaya, Selangor, Malaysia.
[97]Breast Cancer Research Unit, Cancer Research Institute, University Malaya Medical Centre, 59100 KualaLumpur, Malaysia.
[98]Department of Molecular Virology, Immunology and Medical Genetics, Comprehensive Cancer Center, The Ohio State University, Columbus, OH 43210, USA.
[99]Department of Surgical Oncology, Leiden University Medical Center, 2333 ZC Leiden, The Netherlands.
[100]Wellcome Trust Centre for Human Genetics and Oxford Biomedical Research Centre, University of Oxford, Oxford OX3 7BN, UK.
[101]Laboratory of Cancer Genetics and Tumor Biology, Department of Clinical Chemistry, University of Oulu, Oulu FI-90220, Finland.
[102]Laboratory of Cancer Genetics and Tumor Biology, Northern Finland Laboratory Centre NordLab, Oulu FI-90220, Finland.
[103]Servicio de Oncología Médica, Hospital Universitario La Paz, Madrid 28046, Spain.
[104]Shanghai Municipal Center for Disease Control and Prevention, Shanghai, 200336, PR China.
[105]University Hospital Gashuisberg, Leuven, Belgium.
[106]Singapore Eye Research Institute, National University of Singapore, Singapore, Singapore.
[107]Department of Clinical Molecular Biology (EpiGen), Akershus University Hospital, University of Oslo (UiO), Oslo, Norway

8Department of Medical Oncology, Erasmus University Medical Center, Rotterdam, the Netherlands.

† These authors contributed equally to this work.

* **Correspondence to:**

Jirong Long, PhD
Vanderbilt Epidemiology Center and Vanderbilt-Ingram Cancer Center
Vanderbilt University School of Medicine
2525 West End Avenue, 8th Floor, Nashville, TN 37203

Phone: (615) 343-6741; Fax: (615) 936-8241
E-mail: Jirong.Long@vanderbilt.edu

Qiuyin Cai, M.D., Ph.D.
Vanderbilt Epidemiology Center and Vanderbilt-Ingram Cancer Center
Vanderbilt University School of Medicine
1161 21st Avenue South, Nashville, TN 37232
Phone: (615) 936-1351; Fax: (615)936-8291
E-mail: qiuyin.cai@vanderbilt.edu

Running title: Novel 8q24 Variants for Breast Cancer Risk

Key words: breast cancer, genetic susceptibility, 8q24, fine-mapping, single nucleotide polymorphism

Manuscript category: Cancer Genetics

Conflict of interest statement

None of the authors has conflicts of interests to declare.

Abstract word count: 181

Main text word count: 3773

Number of figures: 2 in main text and 3 as additional supporting information

Number of tables: 4 in main text and 7 as additional supporting information

What's new?

Previous genome-wide association studies identified rs13281615 and rs11780156 on 8q24 as breast cancer susceptibility loci. The authors performed a fine-mapping study including 55,540 breast cancer cases and 51,168 controls within the Breast Cancer Association Consortium and identified three additional, represented by rs35961416, rs7815245, and rs2033101, respectively. *In silico* analysis indicated two putatively functional variants rs7815245 and rs1121948.

Abstract

Previous genome-wide association studies among women of European ancestry identified two independent breast cancer susceptibility loci represented by single nucleotide polymorphisms (SNPs) rs13281615 and rs11780156 at 8q24. We conducted a fine-mapping study across 2.06 Mb (chr8:127,561,724 -129,624,067, hg19) in 55,540 breast cancer cases and 51,168 controls within the Breast Cancer Association Consortium. We found three additional independent association signals in women of European ancestry, represented by rs35961416 (OR = 0.95, 95% CI = 0.93-0.97, conditional $P = 5.8 \times 10^{-6}$), rs7815245 (OR = 0.94, 95% CI = 0.91-0.96, conditional $P = 1.1 \times 10^{-6}$), and rs2033101 (OR = 1.05, 95% CI = 1.02-1.07, conditional $P = 1.1 \times 10^{-4}$). Integrative analysis using functional genomic data from the Roadmap Epigenomics, the Encyclopedia of DNA Elements project, the Cancer Genome Atlas, and other public resources implied that SNPs rs7815245 in Signal 3, and rs1121948 in Signal 5 (in linkage disequilibrium with rs11780156, $r^2 = 0.77$), were putatively functional variants for two of the five independent association signals. Our results highlight multiple 8q24 variants associated with breast cancer susceptibility in women of European ancestry.

Introduction

Breast cancer is one of the most common malignancies among women worldwide.[1] Genome-wide association studies (GWASs) have identified approximately 100 loci associated with breast cancer.[2] Multiple independent variants on 8q24 have been shown to confer susceptibility for multiple types of cancer, including breast,[3,4] prostate, colorectal, bladder, ovarian, renal cell, glioma, chronic lymphocytic leukemia, and Hodgkin's lymphoma,[5] (also see Supplementary Figure S1). Although most of these loci are located in a "gene desert" region, several hundred kilobases (kb) telomeric to several genes including *FAM84B*, *POU5F1B*, *MYC*, and the long non-coding gene *PVT1*. A growing number of studies have shown that the 8q24 locus may harbor long-range regulatory elements involved in regulating expression of the *MYC* [6] or *PVT1* genes.[7]

In most GWAS, only the single nucleotide polymorphism (SNP) showing the strongest statistical association at each locus (hereinafter referred to as the index SNP) is reported. Those index SNPs themselves are usually not the causal variants but are in linkage disequilibrium (LD) with the functional variants. In addition to the common variants identified by GWAS, low-frequency variants in susceptible loci may also be associated with disease risk.[8] Furthermore, in each locus, there may exist allelic heterogeneity and multiple independent variants that may be associated with complex diseases.[8] Some of the missing heritability for disease may be derived from the incomplete coverage of genetic variants and poor representation of the full spectrum of causal variants on commercial genotyping arrays.[9,10] Therefore, it is necessary to conduct fine-scale mapping studies to investigate comprehensively all genetic variants in the LD blocks where a GWAS index SNP is located.

We conducted a fine-mapping study of the 8q24 region using data from 106,708 individuals within the Breast Cancer Association Consortium (BCAC). We systematically evaluated the associations with breast cancer risk of the SNPs across 2.06 Mb in this chromosome region. We aimed to identify additional independent association signals and potentially functional variants that may be responsible for the observed associations of variants in this locus with breast cancer risk.

Materials and Methods

Ethics statement

All studies were approved by the relevant institutional review committee and informed consent was obtained from all participants.

Subjects

Epidemiological and genotype data were obtained from 50 breast cancer case-control studies participating in the BCAC.[4] The sample set for the current project included 48,155 cases and 43,612 controls of European ancestry from 39 studies, 6,269 cases and 6,624 controls of Asian ancestry from 9 studies, and 1,116 cases and 932 controls of African ancestry from 2 studies. The estrogen receptor (ER) status of the primary tumor was available for 35,824 cases of European ancestry; 28,038 (78%) cases were ER+ and 7,786 (22%) were ER-.

SNP selection and genotyping

Fine-mapping SNPs were selected for inclusion on the custom Illumina iSelect array (iCOGS),[4,11–13] with the following criteria: 1) Defining the interval to include all SNPs with $r^2 >$ 0.1 with the index SNPs rs13281615 and rs11780156 based on HapMap 2 CEU, which identified a region of 2.06 Mb (base positions 127,561,724 -129,624,067; NCBI build 37 assembly); 2) Identifying all SNPs in the interval using the 1000 Genomes Project CEU (April 2010), and HapMap 3; 3) Selecting high-quality SNPs: only variants with the minor allele called at least twice in the 1000 Genomes Project and an Illumina designability score > 0.8 were included; 4) Selecting all SNPs with $r^2 > 0.1$ with the index SNPs rs13281615 and rs11780156 from the CEU data set of the 1000 Genomes Project or HapMap 3; 5) Selecting tagging SNPs at $r^2 > 0.9$ to capture the remaining SNPs that are not in LD with the index SNPs ($r^2 < 0.1$). Genotyping of the iCOGS array and details of the genotyping calling and quality control has been described elsewhere.[4,11,12] In order to improve SNP density and imputation quality, we conducted one-step imputation (without phasing) using the program IMPUTE2 (see URLs) with the March 2012 release of the 1000 Genomes Project as reference. Genotypes were successfully imputed for 10,593 variants in samples of European ancestry, 9,218 variants in samples of Asian ancestry, and 17,964 variants in samples of African ancestry, all with imputation-$r^2 > 0.3$. After excluding SNPs with minor allele frequency (MAF) < 0.02, the final genotype data in this project included 6,631 SNPs in samples of European ancestry, 6,459 SNPs in samples of Asian ancestry, and 10,830 SNPs in samples of African ancestry.

Statistical analysis

The per-allele odds ratio (OR) and 95% confidence interval (CI) for each SNP was estimated for breast cancer risk using a log-additive logistic model with covariates of study site and principal components (PC; eight PCs with one additional principal component from the LMBC study in Europeans, two PCs in Asians and two in African Americans). Per-allele ORs and CIs were estimated separately for each population for overall disease, for ER+ and for ER- breast cancer. To identify potential independent susceptibility variant(s), stepwise forward logistic regression analyses were employed with or without the index SNPs rs13281615 and rs11780156 forced in the model. SNPs with a P value $< 10^{-4}$ from the single variant analysis were included in this analysis.[11] To identify potentially functional variant candidate(s), we computed a likelihood ratio for each SNP relative to the representative SNP in each signal and excluded SNPs with a likelihood ratio < 0.01. Because no SNPs showed $P < 1\times10^{-4}$ in Asian or African ancestry data, such analyses were performed only on data from subjects of European ancestry. We used the haplo.stats package in R for haplotype analyses for the SNPs that are independently associated with breast cancer risk in women of European ancestry, with study sites and principal components as covariates. The familial relative risk (FRR) was estimated with the formula $\ln(\lambda)/\ln(\lambda_o)$, where λ is the FRR to offspring of an affected individual due to a single genetic locus or assumed multiplicatively interacting loci and λ_o is the overall FRR, which was assumed to be 1.8 for breast cancer.[14] All analyses were conducted using R version 3.0.1.

Functional annotation

We annotated a total of 245 breast cancer risk associated variants ($P<10^{-4}$ from univariate analysis) for potential functional significance using data from the Encyclopedia of DNA Elements (ENCODE), the Roadmap Epigenomics Mapping Consortium, and The Cancer

Genome Atlas (TCGA) (see URLs). For each variant, we investigated whether it is mapped to transcriptional elements primarily associated with enhancers (H3K4me1) or promoters (H3K4me3), in any of nine cell lines: normal human mammary epithelial cell line (HMEC), GM12878, H1-hESC, K562, HepG2, HSMM, HUVEC, NHEK, and NHLF. The epigenetic landscape of histone markers H3K4Me1, H3K4Me3, and H3K27Ac was also examined through layered histone tracks on seven ENCODE cell lines, including GM12878, H1-hESC, K562, HSMM, HUVEC, NHEK, and NHLF from the UCSC Genome Browser (see URLs). DNase I hypersensitive and transcription factor (TF) ChIP-Seq datasets were investigated in all available ENCODE cell lines, including HMEC and the breast cancer cell lines T-47D and MCF-7. Publicly available tools RegulomeDB[15] and HaploReg v4.1[16] were also used to evaluate potential functional variants.

For regions lacking ChIP-seq peaks data, we collected raw ChIP-seq data for the estrogen receptor-α (ESR1) and forkhead box protein A1 (FOXA1) in MCF-7, TAMR and ZR751 breast cancer cell lines (Supplementary Table S1) from the study by Hurtado *et al.* [17] The raw ChIP-seq data in .FASTQ format from different lanes in the same experiment were first merged and mapped to the human reference genome (hg19) using the Bowtie2 program[18] with the default setting. Aligned data were processed and converted into Binary Sequence Alignment/Map format (BAM) files using the SAMtools program.[19] After removing duplicated reads, we used the MACS14 (version 1.4.2) algorithm[20] to identify peaks with 50 bp resolution using the matched DNA input data as the control. The peaks were ranked by the number of uniquely aligned reads and only the top 5% of peaks were selected for motif discovery. The summits of the top 5% peaks were extended by 100 bp on either side. Similar methodological strategy has been used elsewhere.[21,22] Motifs between 5 and 30 bp in length were identified on both strands. We

employed the MEME 4.9.1 toolkit[23] to search DNA motifs and enrichment significance for ESR1 and FOXA1.

Expression quantitative trait (eQTL) analysis

eQTL analysis was performed followings the method described previously.[24] Briefly, RNA-Seq V2 data (level 3) of 1,006 breast cancer tumor tissues were downloaded from the TCGA data portal (see URLs). DNA methylation data measured by the Illumina HumanMethylation450 BeadChip and genotype data from the Affymetrix SNP 6.0 array were also retrieved from TCGA level 3 data. Genotype data of the flanking 2 Mb region of the index SNPs on 8q24 were extracted and then imputed to the 1000 Genomes Project data with Minimac (see URLs). Only common SNPs (MAF > 0.05) with high imputation quality ($r^2 > 0.3$) were included in the present study. For the interrogated 2 Mb region, copy number variation (CNV) data spanning the 8q24 genes *FAM84B*, *POU5F1B*, *MYC,* and *PVT1* from TCGA tumor tissue samples were collected from the CbioPortal (see URLs).

We used the TCGA breast cancer data described above to perform *cis*-eQTL analyses in tumor tissues. Several steps were taken to reduce the batch and other technical effects on gene expressions following the approach described by Pickrell et al.[25] First, the RNA-Seq by Expectation-Maximization value of each gene was log2 transformed and genes with a median expression level of 0 across tissues were removed. We then performed the principal component correction on gene expression to remove potential batch effects. A linear regression of expression values on the first five principal components was constructed and the residuals were used to replace the expression values of each gene among tissues. To make the data more closely conform to the linear model for the eQTL analysis, we further transformed the gene expression

levels to fit quantiles of N(0,1) distribution based on the ranks of the expression values to their respective quantiles. Finally, to further adjust for the potential effects of methylation and CNV on the expression of each gene in tumor tissues, we constructed residual linear regression models to detect eQTLs according to the approach used by Li et al.[26]

Results

Associations with breast cancer risk among women of European ancestry

We first conducted univariate analysis for 2,391 genotyped and 4,240 well-imputed SNPs in samples from women of European ancestry. A total of 359 SNPs were associated with breast cancer risk with a statistical significance of $P < 10^{-4}$ (Figure 1 and Supplementary Table S2). Confirming previous GWAS results, the index SNPs rs13281615 (Signal 2 in Table 1) and rs11780156 (Signal 5 in Table 1) showed significant associations with ORs of 1.11 (95% CI = 1.08-1.13, $P = 2.0 \times 10^{-24}$) and 1.07 (95% CI = 1.05-1.10, $P = 4.1 \times 10^{-8}$), respectively (Table 1, univariate analysis). We then conducted forward stepwise regression analysis for each of the 359 SNPs to identify potential independent association signals. When two index SNPs rs13281615 and rs11780156 were forced into the model, we found two additional independent association signals at statistical significance of $P < 1 \times 10^{-4}$ and a third one with suggestive evidence ($P = 1.1 \times 10^{-4}$) (Table 1 and Supplementary Figure S2). The first independent signal (Signal 1 in Table 1), represented by rs35961416 (chr8:128213561:I) with an insertion of base A, showed a P value of 5.8×10^{-6} after adjustment for other four signals (conditional OR = 0.95, 95% CI = 0.93-0.97). The second independent signal (Signal 3) represented by rs7815245 (conditional OR = 0.94, 95% CI = 0.91-0.96, $P = 1.1 \times 10^{-6}$, Table 1), was in moderate LD with the index SNP

rs13281615 in Signal 2 ($r^2 = 0.48$) but not with the second index SNP rs11780156 in Signal 5 ($r^2 < 0.01$). The third suggestive independent variant (Signal 4, rs2033101), not in LD with either of the two index SNPs, showed a conditional *P* value of 1.1×10^{-4} (OR = 1.05, 95% CI = 1.02-1.07, Table 1). We also performed forward stepwise regression analysis without the two index SNPs forced into the model. Five similar independent breast cancer risk associated SNPs were selected: Signal 1 (rs35961416) (conditional $P = 3.2 \times 10^{-6}$) and the suggestive Signal 4 (rs2033101) (conditional $P = 1.4 \times 10^{-4}$) remained the same; Signal 2 (rs13281615) was tagged by the highly correlated SNP rs10110330 ($r^2 = 0.97$, conditional OR = 1.06, 95% CI = 1.04-1.09, $P = 9.5 \times 10^{-6}$), Signal 3 (rs7815245) tagged by the most significant SNP rs17465052 ($r^2 = 0.93$, conditional OR = 0.94, 95% CI = 0.91-0.97, $P = 3.2 \times 10^{-6}$), and Signal 5 (rs11780156) tagged by c8_pos129263191 (rs67397162, $r^2 = 1$, conditional OR = 1.07, 95% CI = 1.05-1.10, $P = 1.1 \times 10^{-7}$). These results consistently showed four independent risk association signals and another suggestive one (Supplementary Figure S2). No significant evidence of between-study heterogeneity was observed for any of these independently risk-associated SNPs (data not shown).

Stratified by ER status, all five independent signals showed significant associations for ER+ breast cancer; however, with the exception of rs11780156 showing a *P* value of 0.012 (signal 5), no significant associations were observed for ER- breast cancer (Table 2).

Haplotype analyses were performed using data from the five independent risk signals (Table 3). In women of European ancestry, a total of 16 haplotypes with frequency of > 1% were observed. Compared to the reference haplotype, which carries the alleles associated with a reduced risk in all five SNPs, most haplotypes were associated with increased breast cancer risk. Haplotype 5, which carries the risk-associated alleles of the signals 1-3, showed the most

significant association ($P = 3.4 \times 10^{-11}$ for overall breast cancer) while Haplotype 1, which carries the risk-associated alleles of all signals except for Signal 4, was associated with the highest estimated OR (OR = 1.27, 95% CI = 1.15-1.39 for overall breast cancer) (Table 3). As shown in Table 3, similar haplotype associations were observed for ER+ but not for ER- breast cancer.

Association with breast cancer risk in women of Asian or African ancestry

Of the five independently risk-associated variants identified in women of European ancestry, only rs35961416 (Signal 1) showed a nominal association in African-American women at $P < 0.05$ ($P = 0.04$, Table 4). Based on univariate analyses of all SNPs on 8q24 that passed QC, SNP rs76382129 showed a P value of 8.3×10^{-4} in women of Asian ancestry and five SNPs showed P values of between 9.6×10^{-4} and 1.6×10^{-4} in women of African ancestry (Supplementary Table S3). Another 16 SNPs showed breast cancer risk association with P values between 0.01 and 0.001 in either population and in the same direction across the two populations (Table 4).

Functional annotation

For each of the five independent signals identified among women of European ancestry, we excluded SNPs with $r^2 \leq 0.2$ with the representative SNP in each signal region and then calculated the likelihood of all risk-associated variants to select potentially functional variant candidates. Setting a likelihood ratio threshold of > 0.01 relative to the representative/index SNP in each signal region, we did not identify any functional variant candidates for rs35961416

(Signal 1) and rs2033101 (Signal 4) whereas we identified 154 functional variant candidates for Signal 2, 170 variants for Signal 3 (143 variants overlap with those for Signal 2), and 62 variants for Signal 5, respectively (Supplementary Tables S4-S6). Thus, a total of 245 unique SNPs including the five representative SNPs in five signal regions were further evaluated for their potentially functional significance.

Our integrative functional annotation from ENCODE, Roadmap Epigenomics, the RegulomeDB[15], the HaploReg databases[16], and other public data identified the representative SNP rs7815245 in Signal 3 and SNP rs1121948 (in LD with rs11780156) in Signal 5 as most likely functional variant candidates underlying respective independent association signals (Figure 2). Based on the Roadmap Epigenomics data, SNP rs7815245 in Signal 3 is mapped to a conserved enhancer region with a genomic evolutionary rate profiling (GERP) score of 5.04 among eight tissues including breast variant human mammary epithelial cells (vHMEC) and breast myoepithelial primary cells. It is in a DNase I hypersensitive region in eight tissues including vHMEC. It is also predicted to change the transcription factor TCF12 binding motif. This SNP is also located in the binding regions of two critical nuclear hormone responsible receptors, estrogen receptor-α (ESR1), and forkhead box protein A1 (FOXA1) (Supplementary Figure S3). ChIP-seq data from different breast cancer cell lines and technical replicates showed consistent results (Supplementary Figure S3A). DNA binding motif analysis further confirmed that SNP rs7815245 is located in the ESR1 DNA binding motif ($P = 1.5 \times 10^{-3}$) and is very close to the FOXA1 DNA binding motif ($P = 5.2 \times 10^{-3}$) (Supplementary Figure S3B). In addition, the breast cancer risk-associated T allele was correlated with decreased expression of the *POU5F1B* gene ($P = 0.04$, Supplementary Table S7).

SNP rs1121948, which is in strong LD with the index SNP rs11780156 (r^2=0.77) in Signal 5, resides in the binding motifs of the TFs GATA-binding protein 3 (GATA3) and MYC-associated factor X (MAX) in the breast cancer cell line MCF-7. HaploReg data shows that this SNP resides in promoter regions of lung and muscle tissues, in strong enhancer regions of 14 tissues including HSMM and NHLF cells, and in DNase I hypersensitive sites of four tissues. Two active epigenetic markers (H3K4Me1 and H3K27Ac) were enriched in the interval containing rs1121948 in seven ENCODE cell lines (Figure 2C). We examined the effect of the associations of the 62 SNPs from the likelihood analysis for Signal 5 and expression of genes within 1 Mb of the index SNP rs11780156. We found that the risk-associated G allele of rs1121948 was weakly associated with decreased expression of the *PVT1* gene (*P*=0.037, Supplementary Table S7).

Discussion

In this study, we conducted a fine-mapping investigation at the breast cancer susceptibility locus on 8q24. Among women of European ancestry, we identified four independent association signals represented by rs35961416, rs13281615, rs7815245 and rs11780156, respectively, and another suggestive one tagged by rs2033101. This discovery increases the proportion of familial risk of breast cancer explained by variation on 8q24 from 0.25% (due to the GWAS index SNPs rs13281615 in Signal 2 and rs11780156 in Signal 5) to 0.55%.

SNP rs7815245 (Signal 3) showed a more significant association than the previously GWAS-identified index SNPs rs13281615 (Signal 2) and rs11780156 (Signal 5). This SNP is located in an enhancer region among eight tissues including breast variant HMEC and

myoepithelial cells. TF occupancy data showed that SNP rs7815245 falls within the DNA binding motifs for ESR1 and FOXA1, two critical DNA binding proteins for the development of several hormone-dependent cancers including breast cancer. Breast cancer susceptibility variants rs4784227 on 16q12.1 and rs2981578 on 10q26 have also been reported to modulate the affinity for these two transcription factors.[11,27] In addition, FOXA1 has shown a critical role in estrogen-ESR1 activity and endocrine response in breast cancer cells.[17,28,29] These results imply that the association between SNP rs7815245 at Signal 3 and breast cancer risk might be mediated by their functional effects through these two transcription factors. The risk allele T of rs7815245 down-regulated expression of its downstream gene *POU5F1B*, which encodes a weak transcriptional activator highly similar to the POU class 5 homeobox 1 transcription factor and is overexpressed in prostate cancer.[30] However, further functional studies are needed to clarify the biological mechanism of this SNP in breast cancer susceptibility.

The most attractive candidate gene for cancer risk variants in the gene-desert 8q24 region is the proto-oncogene *MYC*, because it plays a vital role in tumorigenesis and metastasis of several types of cancer including breast cancer.[31–33] As a key transcription factor, MYC forms heterodimers with MAX, and then regulates transcription of genes involved in cell growth, and proliferation.[31] Aberrant MYC signaling can promote cell transformation and tumor progression.[32,33] Although most of the GWAS-identified SNPs on 8q24 for multiple types of cancers[5,34] are not mapped to the *MYC* genic region, they may cis-regulate nearby genes including *MYC* and its 53-kb downstream non-coding gene *PVT1*.[5] For example, cancer risk-associated variants may regulate *MYC* expression by forming a large chromatin loop with the *MYC* locus.[6,26,35] This hypothesis has been partially supported by the fact that trait-associated loci are frequently found to be *cis*-eQTL.[26,36,37] Our e-QTL analysis of the TCGA breast cancer tumor

tissues showed a trend that rs1121948 in Signal 5 might affect *MYC* or *PVT1* expression levels (Supplementary Table 7), consistent with a co-expression pattern.[38] However, the risk-increased alleles are associated with down-regulated gene expression of *MYC* or *PVT1*. This is inconsistent with overexpression of these two genes commonly observed in breast cancer tumors,[39–42] leaving challenges to link the risk alleles and the possible candidate gene(s) in malignancy.

Of the five independent association signals observed among women of European ancestry, only rs35961416 showed a significant association in women of African ancestry. This could be due to small effect size, different allele frequency, or allelic heterogeneity by race. Differences in GWAS findings across populations have commonly been observed for breast cancer and many other complex traits.[43–45] Taking the GWAS index SNP rs13281615 as an example, the risk allele frequency in women of European ancestry was 0.40.[46] Under an additive inheritance mode to detect the same per-allele effect (OR = 1.08) at $P = 0.05$, our Asian sample with 6,269 breast cancer cases and 6,624 controls (risk allele frequency of 0.53) and African American sample with 1,116 breast cancer cases and 932 controls (risk allele frequency of 0.44) has a power of 0.09% and 0.1%, respectively.

In addition to a smaller sample size for women of Asian or African ancestry, there are several other limitations in this study. First, no functional laboratory experiments were conducted for any of the putative functional SNPs implicated in our study, preventing us from drawing a more definitive conclusion regarding the functionality of these variants. For example, our *in silico* analyses suggest that rs1121948 is located in binding sites of GATA3 and MAX, which may regulate *MYC* expression, but such potential interaction needs to be demonstrated experimentally.[47–50] Second, we limited our investigation to variants with a MAF > 0.02, and

thus it is possible that some rare variants in these loci may also contribute to the risk of breast cancer.

In conclusion, our fine mapping study identified two additional and another suggestive independent association signals on 8q24 among women of European ancestry, which together with two previous reported GWAS index signals plain approximately 0.55% of excess familial risk of breast cancer. In addition, our functional analyses revealed two putatively functional variants that can be further investigated experimentally. Our study provides additional evidence of the importance of common independent variants on 8q24 in breast cancer susceptibility.

URLs. 1000 Genomes, http://browser.1000genomes.org; BCAC, http://apps.ccge.medschl.cam.ac.uk/consortia/bcac/; CbioPortal, http://www.cbioportal.org/public-portal/; ENCODE, http://genome.ucsc.edu/ENCODE/; HaploReg v4.1, http://www.broadinstitute.org/mammals/haploreg/haploreg.php; HapMap project, http://hapmap.ncbi.nlm.nih.gov/; iCOGs, http://ccge.medschl.cam.ac.uk/research/consortia/icogs/; IMPUTE v.2.2, https://mathgen.stats.ox.ac.uk/impute/impute_v2.html; LocusZoom, http://csg.sph.umich.edu/locuszoom/; Minimac, http://genome.sph.umich.edu/wiki/Minimac; R version 3.0.1, http://www.r-project.org/; RegulomeDB, http://regulome.stanford.edu/; UCSC Genome Browser, http://genome.ucsc.edu; TCGA, http://cancergenome.nih.gov/.

Acknowledgments

We thank all the individuals who took part in these studies and all the researchers, study staff, clinicians and other healthcare providers, technicians and administrative staff who have enabled this work to be carried out. In particular, we thank: Andrew Berchuck (OCAC), Rosalind A. Eeles, Ali Amin Al Olama, Zsofia Kote-Jarai, Sara Benlloch (PRACTICAL), Antonis Antoniou, Lesley McGuffog, Ken Offit (CIMBA), Andrew Lee, and Ed Dicks, Craig Luccarini and the staff of the Centre for Genetic Epidemiology Laboratory, Daniel C. Tessier, Francois Bacot, Daniel Vincent, Sylvie LaBoissière, Frederic Robidoux and the staff of the McGill University and Génome Québec Innovation Centre, Sune F. Nielsen and the staff of the Copenhagen DNA laboratory, Julie M. Cunningham, Sharon A. Windebank, Christopher A. Hilker, Jeffrey Meyer and the staff of Mayo Clinic Genotyping Core Facility, Maggie Angelakos, Judi Maskiell, Ellen van der Schoot (Sanquin Research), Emiel Rutgers, Senno Verhoef, Frans Hogervorst, the Thai Ministry of Public Health (MOPH), Dr Prat Boonyawongviroj (former Permanent Secretary of MOPH), Dr Pornthep Siriwanarungsan (Department Director-General of Disease Control), Michael Schrauder, Matthias Rübner, Sonja Oeser, Silke Landrith, Eileen Williams, Elaine Ryder-Mills, Kara Sargus, Niall McInerney, Gabrielle Colleran, Andrew Rowan, Angela Jones, Christof Sohn, Andeas Schneeweiß, Peter Bugert, the Danish Breast Cancer Group, Núria Álvarez, the CTS Steering Committee (including Leslie Bernstein, James Lacey, Sophia Wang, Huiyan Ma, Yani Lu and Jessica Clague DeHart at the Beckman Research Institute of the City of Hope; Dennis Deapen, Rich Pinder, Eunjung Lee and Fred Schumacher at the University of Southern California; Pam Horn-Ross, Peggy Reynolds and David Nelson at the Cancer Prevention Institute of California; and Hannah Park at the University of California Irvine), Hartwig Ziegler, Sonja Wolf, Volker Hermann, The GENICA network: Dr Margarete Fischer-Bosch-Institute of Clinical Pharmacology, Stuttgart, and University of Tübingen, Germany (HB, Wing-Yee Lo, Christina Justenhoven); Department of Internal Medicine, Evangelische Kliniken Bonn gGmbH, Johanniter Krankenhaus, Bonn, Germany (YDK, Christian Baisch), Institute of Pathology, University of Bonn, Germany (Hans-Peter Fischer), Molecular Genetics of Breast Cancer, Deutsches Krebsforschungszentrum (DKFZ) Heidelberg, Germany (UH), Institute for Prevention and Occupational Medicine of the German Social Accident Insurance, Institute of the Ruhr University Bochum (IPA), Germany (Thomas Brüning, Beate Pesch, Sylvia Rabstein, Anne Lotz), Institute of Occupational Medicine and Maritime Medicine, University Medical Center Hamburg-Eppendorf, Germany (Volker Harth)], Tuomas Heikkinen, Irja Erkkilä, Kirsimari Aaltonen, Karl von Smitten, Natalia Antonenkova, Peter Hillemanns, Hans Christiansen, Eija Myöhänen, Helena Kemiläinen, Heather Thorne, Eveline Niedermayr, the AOCS Management Group (D Bowtell, G Chenevix-Trench, A deFazio, D Gertig, A Green, P Webb), the ACS Management Group (A. Green, P. Parsons, N. Hayward, P. Webb, D. Whiteman), the LAABC data collection team, especially Annie Fung and June Yashiki, Gilian Peuteman, Dominiek Smeets, Thomas Van Brussel, Kathleen Corthouts, Nadia Obi, Judith Heinz, Sabine Behrens, Ursula Eilber, Muhabbet Celik, Til Olchers, Siranoush Manoukian,

Bernard Peissel, Giulietta Scuvera, Daniela Zaffaroni, Bernardo Bonanni, Monica Barile, Irene Feroce, Angela Maniscalco, Alessandra Rossi, Loris Bernard, the personnel of the Cogentech Cancer Genetic Test Laboratory; The Mayo Clinic Breast Cancer Patient Registry, Martine Tranchant, Marie-France Valois, Annie Turgeon, Lea Heguy, Phuah Sze Yee, Peter Kang, Kang In Nee, Shivaani Mariapun, Yoon Sook-Yee, Daphne Lee, Teh Yew Ching, Nur Aishah Mohd Taib, Meeri Otsukka, Kari Mononen, Teresa Selander, Nayana Weerasooriya, OFBCR staff, E. Krol-Warmerdam, J. Molenaar, J. Blom, Louise Brinton, Neonila Szeszenia-Dabrowska, Beata Peplonska, Witold Zatonski, Pei Chao, Michael Stagner, Petra Bos, Jannet Blom, Ellen Crepin, Anja Nieuwlaat, Annette Heemskerk, the Erasmus MC Family Cancer Clinic, Sue Higham, Simon Cross, Helen Cramp, Dan Connley, Sabapathy Balasubramanian, Ian Brock, The Eastern Cancer Registration and Information Centre, the SEARCH and EPIC teams, Craig Luccarini, Don Conroy, Caroline Baynes, Kimberley Chua, the Ohio State University Human Genetics Sample Bank and Robert Pilarski. Data on SCCS cancer cases used in this publication were provided by the: Alabama Statewide Cancer Registry; Kentucky Cancer Registry, Lexington, KY; Tennessee Department of Health, Office of Cancer Surveillance; Florida Cancer Data System; North Carolina Central Cancer Registry, North Carolina Division of Public Health; Georgia Comprehensive Cancer Registry; Louisiana Tumor Registry; Mississippi Cancer Registry; South Carolina Central Cancer Registry; Virginia Department of Health, Virginia Cancer Registry; Arkansas Department of Health, Cancer Registry; Dr. Kristine Kleivi, PhD (K.G. Jebsen Centre for Breast Cancer Research, Institute of Clinical Medicine, University of Oslo, Oslo, Norway and Department of Research, Vestre Viken, Drammen, Norway), Dr. Lars Ottestad, MD (Department of Genetics, Institute for Cancer Research, Oslo University Hospital Radiumhospitalet, Oslo, Norway), Prof. Em. Rolf Kåresen, MD (Department of Oncology, Oslo University Hospital and Faculty of Medicine, University of Oslo, Oslo, Norway), Dr. Anita Langerød, PhD (Department of Genetics, Institute for Cancer Research, Oslo University Hospital Radiumhospitalet, Oslo, Norway), Dr. Ellen Schlichting, MD (Department for Breast and Endocrine Surgery, Oslo University Hospital Ullevaal, Oslo, Norway), Dr. Marit Muri Holmen, MD (Department of Radiology and Nuclear Medicine, Oslo University Hospital, Oslo, Norway), Prof. Toril Sauer, MD (Department of Pathology at Akershus University hospital, Lørenskog, Norway), Dr. Vilde Haakensen, MD (Department of Genetics, Institute for Cancer Research, Oslo University Hospital Radiumhospitalet, Oslo, Norway), Dr. Olav Engebråten, MD (Institute for Clinical Medicine, Faculty of Medicine, University of Oslo and Department of Oncology, Oslo University Hospital, Oslo, Norway), Prof. Bjørn Naume, MD (Division of Cancer Medicine and Radiotherapy, Department of Oncology, Oslo University Hospital Radiumhospitalet, Oslo, Norway), Dr. Cecile E. Kiserud, MD (National Advisory Unit on Late Effects after Cancer Treatment, Department of Oncology, Oslo University Hospital, Oslo, Norway and Department of Oncology, Oslo University Hospital, Oslo, Norway), Dr. Kristin V. Reinertsen, MD (National Advisory Unit on Late Effects after Cancer Treatment, Department of Oncology, Oslo University Hospital, Oslo, Norway and Department of Oncology, Oslo University Hospital, Oslo, Norway), Assoc. Prof. Åslaug Helland, MD (Department of Genetics, Institute for Cancer Research and

Department of Oncology, Oslo University Hospital Radiumhospitalet, Oslo, Norway), Dr. Margit Riis, MD (Dept of Breast- and Endocrine Surgery, Oslo University Hospital, Ullevål, Oslo, Norway), Dr. Ida Bukholm, MD (Department of Breast-Endocrine Surgery, Akershus University Hospital, Oslo, Norway and Department of Oncology, Division of Cancer Medicine, Surgery and Transplantation, Oslo University Hospital, Oslo, Norway), Prof. Per Eystein Lønning, MD (Section of Oncology, Institute of Medicine, University of Bergen and Department of Oncology, Haukeland University Hospital, Bergen, Norway) and Grethe I. Grenaker Alnæs, M.Sc. (Department of Genetics, Institute for Cancer Research, Oslo University Hospital Radiumhospitalet, Oslo, Norway); Paolo Radice, Bernard Peissel and Daniela Zaffaroni (Fondazione IRCCS Istituto Nazionale dei Tumori - INT, Milan, Italy); Bernardo Bonanni, Monica Barile and Irene Feroce (Istituto Europeo di Oncologia - IEO, Milan, Italy); and the personnel of the Cogentech Cancer Genetic Test Laboratory, Milan, Italy. We thank Jing He for data process and analyses and Chenjie Zeng for help with plotting (Vanderbilt Epidemiology Center, Nashville, TN, USA).

Financial support

The work conducted for this project at Vanderbilt Epidemiology Center is supported in part by NIH grant R37CA070867 and endowment funds for the Ingram Professorship and Anne Potter Wilson Chair. BCAC is funded by Cancer Research UK (C1287/A10118, C1287/A12014) and by the European Community's Seventh Framework Programme under grant agreement n° 223175 (HEALTH-F2–2009-223175) (COGS). Meetings of the BCAC have been funded by the European Union COST programme (BM0606). Genotyping of the iCOGS array was funded by the European Union (HEALTH-F2-2009-223175), Cancer Research UK (C1287/A10710), the Canadian Institutes of Health Research for the 'CIHR Team in Familial Risks of Breast Cancer' program and the Ministry of Economic Development, Innovation and Export Trade of Quebec (PSR-SIIRI-701). Additional support for the iCOGS infrastructure was provided by 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), Komen Foundation for the Cure, the Breast Cancer Research Foundation, and the Ovarian Cancer Research Fund. The ABCFS and OFBCR work 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 US Government or the BCFR. The ABCFS was also supported by the National Health and Medical Research Council of Australia, the New South Wales Cancer Council, the Victorian Health Promotion Foundation (Australia) and the Victorian Breast Cancer Research Consortium. J.L.H.

is a National Health and Medical Research Council (NHMRC) Senior Principal Research Fellow and M.C.S. is a NHMRC Senior Research Fellow. The OFBCR work was also supported by the Canadian Institutes of Health Research 'CIHR Team in Familial Risks of Breast Cancer' program. The ABCS was funded by the Dutch Cancer Society Grant no. NKI2007-3839 and NKI2009-4363. The ACP study is funded by the Breast Cancer Research Trust, UK. The work of the BBCC was partly funded by ELAN-Programme of the University Hospital of Erlangen. The 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). E.S. is supported by NIHR Comprehensive Biomedical Research Centre, Guy's & St. Thomas' NHS Foundation Trust in partnership with King's College London, UK. Core funding to the Wellcome Trust Centre for Human Genetics was provided by the Wellcome Trust (090532/Z/09/Z). I.T. 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 funded by the Fondation de France, the French National Institute of Cancer (INCa), The National League against Cancer, the National Agency for Environmental and Occupational Health and Food Safety (ANSES), the National Agency for Research (ANR), and the Association for Research against Cancer (ARC). The CGPS was supported by the Chief Physician Johan Boserup and Lise Boserup Fund, the Danish Medical Research Council and Herlev Hospital. The CNIO-BCS was supported by the Genome Spain Foundation, the Red Temática de Investigación Cooperativa en Cáncer and grants from the Asociación Española Contra el Cáncer and the Fondo de Investigación Sanitario (PI11/00923 and PI081120). The Human Genotyping-CEGEN Unit, CNIO is supported by the Instituto de Salud Carlos III. D.A. was supported by a Fellowship from the Michael Manzella Foundation (MMF) and was a participant in the CNIO Summer Training Program. 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). 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, Institute for Prevention and Occupational Medicine of the German Social Accident Insurance, Institute of the Ruhr University Bochum (IPA), as well as the Department of Internal Medicine, Evangelische Kliniken Bonn gGmbH, Johanniter Krankenhaus Bonn, Germany. The HEBCS was supported by the Helsinki University Central Hospital Research Fund, Academy of Finland (132473), the Finnish Cancer Society, The Nordic Cancer Union and the Sigrid Juselius

Foundation. The HERPACC was supported by a Grant-in-Aid for Scientific Research on Priority Areas from the Ministry of Education, Science, Sports, Culture and Technology of Japan, by a Grant-in-Aid for the Third Term Comprehensive 10-Year Strategy for Cancer Control from Ministry Health, Labour and Welfare of Japan, by a research grant from Takeda Science Foundation, by Health and Labour Sciences Research Grants for Research on Applying Health Technology from Ministry Health, Labour and Welfare of Japan and by National Cancer Center Research and Development Fund. The HMBCS was supported by short-term fellowships from the German Academic Exchange Program (to N.B), and the Friends of Hannover Medical School (to N.B.). Financial support for KARBAC was provided through the regional agreement on medical training and clinical research (ALF) between Stockholm County Council and Karolinska Institutet, the Stockholm Cancer Foundation and the Swedish Cancer Society. 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, the Academy of Finland and by the strategic funding of the University of Eastern Finland. kConFab is supported by grants from the National Breast Cancer Foundation, the NHMRC, the Queensland Cancer Fund, the Cancer Councils of New South Wales, Victoria, Tasmania and South Australia and the Cancer Foundation of Western Australia. The kConFab Clinical Follow Up Study was funded by the NHMRC (145684, 288704, 454508). Financial support for the AOCS was provided by the United States Army Medical Research and Materiel Command (DAMD17-01-1-0729), the Cancer Council of Tasmania and Cancer Foundation of Western Australia and the NHMRC (199600). G.C.T. and P.W. are supported by the NHMRC. LAABC is supported by grants (1RB-0287, 3PB-0102, 5PB-0018 and 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' (232-2008 and 196-2010). The MARIE study was supported by the Deutsche Krebshilfe e.V. (70-2892-BR I), the Federal Ministry of Education and Research (BMBF) Germany (01KH0402), the Hamburg Cancer Society and the German Cancer Research Center (DKFZ). MBCSG is supported by grants from the Italian Association for Cancer Research (AIRC) and by funds from the Italian citizens who allocated a 5/1000 share of their tax payment in support of the Fondazione IRCCS Istituto Nazionale Tumori, according to Italian laws (INT-Institutional strategic projects '5 × 1000'). The MCBCS was supported by the NIH grants (CA122340, CA128978) and a Specialized Program of Research Excellence (SPORE) in Breast Cancer (CA116201), the Breast Cancer Research Foundation and a generous gift from the David F. and Margaret T. Grohne Family Foundation and the Ting Tsung and Wei Fong Chao Foundation. MCCS cohort recruitment was funded by VicHealth and Cancer Council Victoria. The MCCS was further supported by Australian NHMRC grants 209057, 251553 and 504711 and by infrastructure provided by Cancer Council Victoria. The MEC was supported by NIH grants CA63464, CA54281, CA098758 and

CA132839. The work of MTLGEBCS was supported by the Quebec Breast Cancer Foundation, the Canadian Institutes of Health Research (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 Science, Technology and Innovation (MOSTI), Malaysian Ministry of Higher Education (UM.C/HlR/MOHE/06) and Cancer Research Initiatives Foundation (CARIF). Additional controls were recruited by the Singapore Eye Research Institute, which was supported by a grant from the Biomedical Research Council (BMRC08/1/35/19<tel:08/1/35/19>/550), Singapore and the National medical Research Council, Singapore (NMRC/CG/SERI/2010). The NBCS was supported by grants from the Norwegian Research council (155218/V50 to A.L.B.D. and V.N. K., 193387/H10 to A.L.B.D., and V.N. K., 175240/S10 to A.L.B.D., FUGE-NFR 181600/V11 to V.N.K. and a Swizz Bridge Award to A.L.B.D.), South Eastern Norway Health Authority (grant 39346 to A.L.B.D.), the K.G. Jebsen Centre for Breast Cancer Research and the Norwegian Cancer Society (to A.L.B.D. and V.N. K.). 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 Sigrid Juselius Foundation, the Academy of Finland, the University of Oulu, and the Oulu University Hospital. The ORIGO study was supported by the Dutch Cancer Society (RUL 1997-1505) and the Biobanking and Biomolecular Resources Research Infrastructure (BBMRI-NL CP16). The PBCS was funded by Intramural Research Funds of the National Cancer Institute, Department of Health and Human Services, USA. pKARMA is a combination of the KARMA and LIBRO-1 studies. KARMA was supported by Märit and Hans Rausings Initiative Against Breast Cancer. KARMA and LIBRO-1 were supported the Cancer Risk Prediction Center (CRisP; www.crispcenter.org), a Linnaeus Centre (Contract ID 70867902) financed by the Swedish Research Council. The RBCS was funded by the Dutch Cancer Society (DDHK 2004-3124, DDHK 2009-4318). 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. KC was financed by the Swedish Cancer Society (5128-B07-01PAF). The SBCGS was supported primarily by NIH grants R01CA64277, R01CA148667, and R37CA70867. The SBCS was supported by Yorkshire Cancer Research S305PA, S299 and S295. Funding for the SCCS was provided by NIH grant R01 CA092447. The Arkansas Central Cancer Registry is fully funded by a grant from National Program of Cancer Registries, Centers for Disease Control and Prevention (CDC). Data on SCCS cancer cases from Mississippi were collected by the Mississippi Cancer Registry which participates in the National Program of Cancer Registries (NPCR) of the Centers for Disease Control and Prevention (CDC). The contents of this publication are solely the responsibility of the authors and do not necessarily represent the official views of the CDC or the Mississippi Cancer Registry. 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 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 National Medical Research Council Start-up Grant and Centre Grant (NMRC/CG/NCIS /2010). The recruitment of controls by the Singapore Consortium of Cohort Studies-Multi-ethnic cohort (SCCS-MEC) 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. K. J. is a fellow of International PhD program, Postgraduate School of Molecular Medicine, Warsaw Medical University, supported by the Polish Foundation of Science. The TNBCC was supported by the NIH grant (CA128978), the Breast Cancer Research Foundation, Komen Foundation for the Cure, the Ohio State University Comprehensive Cancer Center, the Stefanie Spielman Fund for Breast Cancer Research and a generous gift from the David F. and Margaret T. Grohne Family Foundation and the Ting Tsung and Wei Fong Chao Foundation. Part of the TNBCC (DEMOKRITOS) has been co-financed by 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)—Research Funding Program of the General Secretariat for Research & Technology: ARISTEIA. The TWBCS is supported by the Institute of Biomedical Sciences, Academia Sinica and the National Science Council, Taiwan. The UKBGS is funded by Breakthrough Breast Cancer and the Institute of Cancer Research (ICR). ICR acknowledges NHS funding to the NIHR Biomedical Research Centre.

References

1. Mavaddat N, Antoniou AC, Easton DF, Garcia-Closas M. Genetic susceptibility to breast cancer. *Mol Oncol* 2010;4:174–91.

2. Welter D, MacArthur J, Morales J, Burdett T, Hall P, Junkins H, Klemm A, Flicek P, Manolio T, Hindorff L, Parkinson H. The NHGRI GWAS Catalog, a curated resource of SNP-trait associations. *Nucleic Acids Res* 2014;42:D1001–1006.

3. Easton DF, Pooley KA, Dunning AM, Pharoah PDP, Thompson D, Ballinger DG, Struewing JP, Morrison J, Field H, Luben R, Wareham N, Ahmed S, et al. Genome-wide association study identifies novel breast cancer susceptibility loci. *Nature* 2007;447:1087–93.

4. Michailidou K, Hall P, Gonzalez-Neira A, Ghoussaini M, Dennis J, Milne RL, Schmidt MK, Chang-Claude J, Bojesen SE, Bolla MK, Wang Q, Dicks E, et al. Large-scale genotyping identifies 41 new loci associated with breast cancer risk. *Nat Genet* 2013;45:353–361, 361e1–2.

5. Huppi K, Pitt JJ, Wahlberg BM, Caplen NJ. The 8q24 gene desert: an oasis of non-coding transcriptional activity. *Front Genet* 2012;3:69.

6. Ahmadiyeh N, Pomerantz MM, Grisanzio C, Herman P, Jia L, Almendro V, He HH, Brown M, Liu XS, Davis M, Caswell JL, Beckwith CA, et al. 8q24 prostate, breast, and colon cancer risk loci show tissue-specific long-range interaction with MYC. *Proc Natl Acad Sci U S A* 2010;107:9742–6.

7. Meyer KB, Maia A-T, O'Reilly M, Ghoussaini M, Prathalingam R, Porter-Gill P, Ambs S, Prokunina-Olsson L, Carroll J, Ponder BAJ. A Functional Variant at a Prostate Cancer Predisposition Locus at 8q24 Is Associated with PVT1 Expression. *PLoS Genet* 2011;7:e1002165.

8. Manolio TA, Collins FS, Cox NJ, Goldstein DB, Hindorff LA, Hunter DJ, McCarthy MI, Ramos EM, Cardon LR, Chakravarti A, Cho JH, Guttmacher AE, et al. Finding the missing heritability of complex diseases. *Nature* 2009;461:747–53.

9. McCarthy MI, Abecasis GR, Cardon LR, Goldstein DB, Little J, Ioannidis JPA, Hirschhorn JN. Genome-wide association studies for complex traits: consensus, uncertainty and challenges. *Nat Rev Genet* 2008;9:356–69.

10. Sanna S, Li B, Mulas A, Sidore C, Kang HM, Jackson AU, Piras MG, Usala G, Maninchedda G, Sassu A, Serra F, Palmas MA, et al. Fine mapping of five loci associated with low-density lipoprotein cholesterol detects variants that double the explained heritability. *PLoS Genet* 2011;7:e1002198.

11. Meyer KB, O'Reilly M, Michailidou K, Carlebur S, Edwards SL, French JD, Prathalingham R, Dennis J, Bolla MK, Wang Q, de Santiago I, Hopper JL, et al. Fine-scale mapping of the FGFR2 breast cancer risk locus: putative functional variants differentially bind FOXA1 and E2F1. *Am J Hum Genet* 2013;93:1046–60.

12. Agarwal D, Pineda S, Michailidou K, Herranz J, Pita G, Moreno LT, Alonso MR, Dennis J, Wang Q, Bolla MK, Meyer KB, Menéndez-Rodríguez P, et al. FGF receptor genes and breast cancer susceptibility: results from the Breast Cancer Association Consortium. *Br J Cancer* 2014;110:1088–100.

3. Glubb DM, Maranian MJ, Michailidou K, Pooley KA, Meyer KB, Kar S, Carlebur S, O'Reilly M, Betts JA, Hillman KM, Kaufmann S, Beesley J, et al. Fine-scale mapping of the 5q11.2 breast cancer locus reveals at least three independent risk variants regulating MAP3K1. *Am J Hum Genet* 2015;96:5–20.

4. Collaborative Group on Hormonal Factors in Breast Cancer. Familial breast cancer: collaborative reanalysis of individual data from 52 epidemiological studies including 58,209 women with breast cancer and 101,986 women without the disease. *Lancet* 2001;358:1389–99.

5. Boyle AP, Hong EL, Hariharan M, Cheng Y, Schaub MA, Kasowski M, Karczewski KJ, Park J, Hitz BC, Weng S, Cherry JM, Snyder M. Annotation of functional variation in personal genomes using RegulomeDB. *Genome Res* 2012;22:1790–7.

6. Ward LD, Kellis M. HaploReg: a resource for exploring chromatin states, conservation, and regulatory motif alterations within sets of genetically linked variants. *Nucleic Acids Res* 2012;40:D930–934.

7. Hurtado A, Holmes KA, Ross-Innes CS, Schmidt D, Carroll JS. FOXA1 is a key determinant of estrogen receptor function and endocrine response. *Nat Genet* 2011;43:27–33.

8. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. *Nat Methods* 2012;9:357–9.

9. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R, 1000 Genome Project Data Processing Subgroup. The Sequence Alignment/Map format and SAMtools. *Bioinforma Oxf Engl* 2009;25:2078–9.

0. Zhang Y, Liu T, Meyer CA, Eeckhoute J, Johnson DS, Bernstein BE, Nusbaum C, Myers RM, Brown M, Li W, Liu XS. Model-based analysis of ChIP-Seq (MACS). *Genome Biol* 2008;9:R137.

1. Jothi R, Cuddapah S, Barski A, Cui K, Zhao K. Genome-wide identification of in vivo protein-DNA binding sites from ChIP-Seq data. *Nucleic Acids Res* 2008;36:5221–31.

2. Ramagopalan SV, Heger A, Berlanga AJ, Maugeri NJ, Lincoln MR, Burrell A, Handunnetthi L, Handel AE, Disanto G, Orton S-M, Watson CT, Morahan JM, et al. A ChIP-seq defined genome-wide map of vitamin D receptor binding: associations with disease and evolution. *Genome Res* 2010;20:1352–60.

3. Bailey TL, Boden M, Buske FA, Frith M, Grant CE, Clementi L, Ren J, Li WW, Noble WS. MEME SUITE: tools for motif discovery and searching. *Nucleic Acids Res* 2009;37:W202–208.

4. Cai Q, Zhang B, Sung H, Low S-K, Kweon S-S, Lu W, Shi J, Long J, Wen W, Choi J-Y, Noh D-Y, Shen C-Y, et al. Genome-wide association analysis in East Asians identifies breast cancer susceptibility loci at 1q32.1, 5q14.3 and 15q26.1. *Nat Genet* 2014;46:886–90.

5. Pickrell JK, Marioni JC, Pai AA, Degner JF, Engelhardt BE, Nkadori E, Veyrieras J-B, Stephens M, Gilad Y, Pritchard JK. Understanding mechanisms underlying human gene expression variation with RNA sequencing. *Nature* 2010;464:768–72.

6. Li Q, Seo J-H, Stranger B, McKenna A, Pe'er I, Laframboise T, Brown M, Tyekucheva S, Freedman ML. Integrative eQTL-based analyses reveal the biology of breast cancer risk loci. *Cell* 2013;152:633–41.

7. Cowper-Sal lari R, Zhang X, Wright JB, Bailey SD, Cole MD, Eeckhoute J, Moore JH, Lupien M. Breast cancer risk-associated SNPs modulate the affinity of chromatin for FOXA1 and alter gene expression. *Nat Genet* 2012;44:1191–8.

8. Carroll JS, Liu XS, Brodsky AS, Li W, Meyer CA, Szary AJ, Eeckhoute J, Shao W, Hestermann EV, Geistlinger TR, Fox EA, Silver PA, et al. Chromosome-wide mapping of estrogen receptor binding reveals long-range regulation requiring the forkhead protein FoxA1. *Cell* 2005;122:33–43.

9. Bernardo GM, Lozada KL, Miedler JD, Harburg G, Hewitt SC, Mosley JD, Godwin AK, Korach KS, Visvader JE, Kaestner KH, Abdul-Karim FW, Montano MM, et al. FOXA1 is an essential determinant of ERalpha expression and mammary ductal morphogenesis. *Dev Camb Engl* 2010;137:2045–54.

0. Kastler S, Honold L, Luedeke M, Kuefer R, Möller P, Hoegel J, Vogel W, Maier C, Assum G. POU5F1P1, a putative cancer susceptibility gene, is overexpressed in prostatic carcinoma. *The Prostate* 2010;70:666–74.

1. Chen Y, Olopade OI. MYC in breast tumor progression. *Expert Rev Anticancer Ther* 2008;8:1689–98.

2. Hynes NE, Stoelzle T. Key signalling nodes in mammary gland development and cancer: Myc. *Breast Cancer Res* 2009;11:210.

3. Wolfer A, Ramaswamy S. MYC and metastasis. *Cancer Res* 2011;71:2034–7.

4. Grisanzio C, Freedman ML. Chromosome 8q24-Associated Cancers and MYC. *Genes Cancer* 2010;1:555–9.

5. Wright JB, Brown SJ, Cole MD. Upregulation of c-MYC in cis through a Large Chromatin Loop Linked to a Cancer Risk-Associated Single-Nucleotide Polymorphism in Colorectal Cancer Cells. *Mol Cell Biol* 2010;30:1411–20.

6. Nicolae DL, Gamazon E, Zhang W, Duan S, Dolan ME, Cox NJ. Trait-associated SNPs are more likely to be eQTLs: annotation to enhance discovery from GWAS. *PLoS Genet* 2010;6:e1000888.

7. Li Q, Stram A, Chen C, Kar S, Gayther S, Pharoah P, Haiman C, Stranger B, Kraft P, Freedman ML. Expression QTL-based analyses reveal candidate causal genes and loci across five tumor types. *Hum Mol Genet* 2014;23:5294–302.

8. Colombo T, Farina L, Macino G, Paci P. PVT1: a rising star among oncogenic long noncoding RNAs. *BioMed Res Int* 2015;2015:304208.

9. Guan Y, Kuo W-L, Stilwell JL, Takano H, Lapuk AV, Fridlyand J, Mao J-H, Yu M, Miller MA, Santos JL, Kalloger SE, Carlson JW, et al. Amplification of PVT1 contributes to the pathophysiology of ovarian and breast cancer. *Clin Cancer Res Off J Am Assoc Cancer Res* 2007;13:5745–55.

0. Sørlie T, Perou CM, Tibshirani R, Aas T, Geisler S, Johnsen H, Hastie T, Eisen MB, van de Rijn M, Jeffrey SS, Thorsen T, Quist H, et al. Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. *Proc Natl Acad Sci U S A* 2001;98:10869–74.

1. Sotiriou C, Neo S-Y, McShane LM, Korn EL, Long PM, Jazaeri A, Martiat P, Fox SB, Harris AL, Liu ET. Breast cancer classification and prognosis based on gene expression profiles from a population-based study. *Proc Natl Acad Sci U S A* 2003;100:10393–8.

2. Xu J, Chen Y, Olopade OI. MYC and Breast Cancer. *Genes Cancer* 2010;1:629–40.

3. Zheng W, Zhang B, Cai Q, Sung H, Michailidou K, Shi J, Choi J-Y, Long J, Dennis J, Humphreys MK, Wang Q, Lu W, et al. Common genetic determinants of breast-cancer risk in East Asian women: a collaborative study of 23 637 breast cancer cases and 25 579 controls. *Hum Mol Genet* 2013;22:2539–50.

4. Bojesen SE, Pooley KA, Johnatty SE, Beesley J, Michailidou K, Tyrer JP, Edwards SL, Pickett HA, Shen HC, Smart CE, Hillman KM, Mai PL, et al. Multiple independent variants at the TERT locus are associated with telomere length and risks of breast and ovarian cancer. *Nat Genet* 2013;45:371–384, 384e1–2.

5. Wu Y, Waite LL, Jackson AU, Sheu WH-H, Buyske S, Absher D, Arnett DK, Boerwinkle E, Bonnycastle LL, Carty CL, Cheng I, Cochran B, et al. Trans-ethnic fine-mapping of lipid loci identifies population-specific signals and allelic heterogeneity that increases the trait variance explained. *PLoS Genet* 2013;9:e1003379.

6. Easton DF, Pooley KA, Dunning AM, Pharoah PDP, Thompson D, Ballinger DG, Struewing JP, Morrison J, Field H, Luben R, Wareham N, Ahmed S, et al. Genome-wide association study identifies novel breast cancer susceptibility loci. *Nature* 2007;447:1087–93.

7. Tomlinson I, Webb E, Carvajal-Carmona L, Broderick P, Kemp Z, Spain S, Penegar S, Chandler I, Gorman M, Wood W, Barclay E, Lubbe S, et al. A genome-wide association scan of tag SNPs identifies a susceptibility variant for colorectal cancer at 8q24.21. *Nat Genet* 2007;39:984–8.

8. Junttila MR, Westermarck J. Mechanisms of MYC stabilization in human malignancies. *Cell Cycle Georget Tex* 2008;7:592–6.

9. Ghoussaini M, Song H, Koessler T, Al Olama AA, Kote-Jarai Z, Driver KE, Pooley KA, Ramus SJ, Kjaer SK, Hogdall E, DiCioccio RA, Whittemore AS, et al. Multiple Loci With Different Cancer Specificities Within the 8q24 Gene Desert. *JNCI J Natl Cancer Inst* 2008;100:962–6.

0. Lin CY, Lovén J, Rahl PB, Paranal RM, Burge CB, Bradner JE, Lee TI, Young RA. Transcriptional amplification in tumor cells with elevated c-Myc. *Cell* 2012;151:56–67.

FIGURE LEGENDS

Figure 1. Manhattan plot of overall breast cancer risk association in Europeans at the 8q24 locus. SNPs are plotted based on their chromosomal positions (hg19) and -$\log_{10}$ *P*-values for univariate association. The regions spanning five independent association signals (representative SNP for each signal are shown in Table 1) and their highly correlated SNPs are indicated by dashed rectangles.

Figure 2. Functional annotation of the independent signal regions. (**A**) Chromatin states across the 45.6 kb region harboring two associated signals rs13281615 and rs7815245. The top 3 tracks show enrichment of transcription regulatory histone markers H3K4me1, H3K4me3 and H3K27ac from seven cell lines in ENCODE. The next 9 tracks are the chromatin state annotation by ChromHMM derived from 9 cell types. ChromHMM color coding is as follows: orange, strong enhancer; yellow, weak enhancer; light green, weak transcribed; light gray, low signal. The next 5 tracks show the designated histone modifications in the HMEC cell line. The last two tracks show the open chromatin enrichment from DNase clusters and evolutionary conservation measurement by PhastCons from 100 vertebrates. (B) LD structure of the region harboring the index SNP rs13281615 and the independent signal rs7815245 in European samples. (C) Chromatin states across the 58.9 kb region harboring the second index SNP rs11780156 and the potential underlying functional SNP rs1121948. The contents of the tracks are the same as described in (A). (D) LD structure of the 58.9 kb region wherein SNPs rs11780156 and rs1121948 lie, marked with red arrows.

Table 1. Identification of five independent association signals for overall breast cancer risk among women of European ancestry: a collaborative study of 48,155 breast cases and 43,612 controls.

Signal	SNP	Base position (hg 19)	Alleles[a]	MAF	Imputation r^2	LD[b] (r^2)	Univariate Analysis		Conditional Analysis	
							per-allele OR (95% CI)[c]	*P*[c]	Adjusted per-allele OR (95% CI)[d]	Adjusted *P*[d]
Signal 1	rs35961416	128213561	**A**/-	0.44	0.75	-	0.95(0.93-0.97)	$1.3x10^{-5}$	0.95(0.93-0.97)	$5.8x10^{-6}$
Signal 2	rs13281615[e]	128355618	**G**/A	0.40	0.97	-	1.11(1.08-1.13)	$2.0x10^{-24}$	1.06(1.03-1.09)	$6.6x10^{-6}$
Signal 3	rs7815245	128383597	**T**/C	0.42	1	0.48	0.90(0.88-0.92)	$1.3x10^{-27}$	0.94(0.91-0.96)	$1.1x10^{-6}$
Signal 4	rs2033101	128964222	**T**/C	0.19	0.95	-	1.05(1.03-1.08)	$5.2x10^{-5}$	1.05(1.02-1.07)	$1.1x10^{-4}$
Signal 5	rs11780156[e]	129194641	**T**/C	0.16	1	-	1.07(1.05-1.10)	$4.1x10^{-8}$	1.07(1.04-1.10)	$1.7x10^{-7}$
	rs1121948	129165056	**G**/A	0.20	1	0.77	1.07(1.04-1.09)	$2.0x10^{-7}$	-	-

Abbreviations: LD, linkage disequilibrium; OR, odds ratio; CI, confidence interval; MAF, minor (effect) allele frequency.

[a] Minor/major alleles; effect alleles are shown in bold.

[b] r^2 for linkage disequilibrium with index SNP rs13281615 (0.48) and rs11780156 (0.77),respectively; "-" indicates $r^2<0.01$ with any index SNPs.

[c] Adjusted for age, study site and principal components.

[d] Adjusted for the other four signals, age, study site and principle components.

[e] Independent index SNPs rs13281615 and rs11780156 from previously reported genome-wide association studies.

Table 2. Association of the five independent signals with breast cancer risk by ER status among women of European ancestry.

SNP	ER+ cases (n=28,038)		ER- cases (n=7,786)		*P* for heterogeneity test[b]
	Adjusted per-allele OR (95% CI)[a]	Adjusted *P*[a]	Adjusted per-allele OR (95% CI)[a]	Adjusted *P*[a]	
rs35961416	0.93(0.91-0.96)	$4.8x10^{-7}$	0.97(0.93-1.01)	0.138	0.217
rs13281615	1.08(1.04-1.11)	$3.1x10^{-6}$	1.00(0.95-1.05)	0.971	0.032
rs7815245	0.94(0.91-0.97)	$1.6x10^{-4}$	0.96(0.91-1.01)	0.104	0.276
rs2033101	1.06(1.03-1.09)	$1.3x10^{-4}$	1.04(0.99-1.08)	0.140	0.307
rs11780156	1.08(1.05-1.11)	$6.6x10^{-7}$	1.06(1.01-1.12)	0.012	0.118

Abbreviations: ER, estrogen receptor; OR, odds ratio; CI, confidence interval.

[a] Adjusted for other four independent signal, age, study site and principle components.

[b] Heterogeneity test between ER-positive and ER-negative disease.

Table 3. Haplotype analyses of the two independent signals in relation to breast cancer risk among women of European ancestry.

						All cases (n=48,155)			ER+ cases (n=28,038)			ER- cases (n=7,786)		
	SNP 1[a]	SNP 2[a]	SNP 3[a]	SNP 4[a]	SNP 5[a]	Frequency[b]	per-allele OR (95% CI)[c]	*P*[c]	Frequency[b]	per-allele OR (95% CI)[c]	*P*[c]	Frequency[b]	per-allele OR (95% CI)[c]	*P*[c]
Haplotypes	**A**/-	**G**/A	**T**/C	**T**/C	**T**/C									
Baseline	A	A	T	C	C	0.125	Reference (1.00)	-	0.127	Reference (1.00)	-	0.130	Reference (1.00)	-
Haplo.1	-	G	C	C	T	0.029	1.27(1.15-1.39)	1.1×10^{-6}	0.028	1.30(1.17-1.46)	2.9×10^{-6}	0.028	1.22(1.03-1.46)	0.024
Haplo.2	-	A	C	C	T	0.017	1.24(1.11-1.40)	2.1×10^{-4}	0.016	1.29(1.12-1.48)	2.9×10^{-4}	0.016	1.19(0.96-1.48)	0.115
Haplo.3	-	G	C	T	C	0.033	1.23(1.13-1.35)	3.2×10^{-6}	0.033	1.28(1.15-1.42)	3.0×10^{-6}	0.032	1.10(0.93-1.30)	0.276
Haplo.4	A	G	C	T	C	0.031	1.22(1.12-1.33)	1.2×10^{-5}	0.031	1.21(1.09-1.35)	3.6×10^{-4}	0.030	1.08(0.91-1.28)	0.370
Haplo.5	-	G	C	C	C	0.140	1.17(1.12-1.22)	3.4×10^{-11}	0.139	1.20(1.13-1.26)	1.1×10^{-10}	0.136	1.06(0.97-1.15)	0.220
Haplo.6	A	G	C	C	C	0.131	1.13(1.07-1.19)	2.2×10^{-5}	0.131	1.13(1.06-1.21)	2.1×10^{-4}	0.130	1.04(0.94-1.16)	0.470
Haplo.7	-	A	T	C	T	0.029	1.13(1.03-1.23)	0.117	0.028	1.12(1.01-1.25)	0.038	0.029	1.04(0.87-1.24)	0.654
Haplo.8	A	G	C	C	T	0.024	1.11(1.01-1.23)	0.036	0.024	1.06(0.94-1.20)	0.339	0.024	1.03(0.85-1.25)	0.758
Haplo.9	-	A	C	C	C	0.089	1.11(1.05-1.17)	2.3×10^{-4}	0.090	1.12(1.05-1.19)	5.7×10^{-4}	0.089	1.02(0.92-1.13)	0.688
Haplo.10	-	A	T	T	C	0.037	1.11(1.02-1.20)	0.011	0.037	1.13(1.03-1.24)	9.5×10^{-3}	0.037	1.05(0.90-1.22)	0.529
Haplo.11	-	A	C	T	C	0.020	1.09(0.98-1.21)	0.121	0.020	1.04(0.92-1.18)	0.510	0.021	1.27(1.05-1.53)	0.013
Haplo.12	A	A	C	C	C	0.034	1.05(0.97-1.15)	0.229	0.033	1.01(0.91-1.12)	0.877	0.035	1.09(0.93-1.28)	0.300
Haplo.13	A	A	T	C	T	0.023	1.04(0.92-1.16)	0.548	0.023	1.05(0.91-1.20)	0.494	0.023	1.11(0.90-1.37)	0.318
Haplo.14	-	A	T	C	C	0.148	1.02(0.97-1.08)	0.445	0.149	1.03(0.96-1.09)	0.447	0.153	1.03(0.92-1.14)	0.618
Haplo.15	A	A	T	T	C	0.029	0.97(0.88-1.08)	0.596	0.030	0.97(0.86-1.10)	0.654	0.030	0.96(0.80-1.16)	0.697
Haplo.rare	*	*	*	*	*	0.060	1.20(1.13-1.28)	7.0×10^{-9}	0.060	1.26(1.17-1.35)	6.1×10^{-10}	0.058	1.09(0.96-1.22)	0.183

Abbreviations: ER, estrogen receptor; OR, odds ratio; CI, confidence interval.

[a] SNPs1-5 represent signals rs35961416, rs13281615, rs7815245, rs2033101, and rs11780156, respectively; effect alleles are shown in bold.

[b] Haplotype frequency.

[c] Adjusted for age, study site, and principal components.

Table 4. Association of top SNPs identified in women of European and non-European ancestry with breast cancer risk among women of Asian (6,269 cases and 6,624 controls) and African ancestry (1,116 cases and 932 controls).

Top SNPs	Alleles[a]	Univariate Analysis (Asian)			Univariate Analysis (African)		
		EAF	OR(95% CI)[b]	*P*[b]	EAF	OR(95% CI)[b]	*P*[b]
Identified in women of European ancestry							
rs35961416	**A**/-	0.10	1.01(0.91-1.13)	0.804	0.39	0.85(0.73-0.99)	0.040
rs13281615	**G**/A	0.53	1.02(0.97-1.08)	0.357	0.44	1.02(0.90-1.16)	0.721
rs7815245[c]	**T**/C	0.17	0.96(0.89-1.02)	0.202	0.36	0.99(0.87-1.13)	0.908
rs2033101	**T**/C	0.33	1.00(0.95-1.05)	0.957	0.09	0.85(0.67-1.07)	0.174
rs11780156	**T**/C	0.20	0.99(0.93-1.06)	0.842	0.04	0.97(0.70-1.34)	0.838
Identified in women of non-European ancestry							
rs16901629	**G**/A	0.14	1.12(1.03-1.22)	6.6×10^{-3}	0.38	1.07(0.92-1.24)	0.407
rs974451	**A**/G	0.23	0.91(0.85-0.96)	1.4×10^{-3}	0.56	0.99(0.88-1.13)	0.915
rs7014860	**C**/A	0.24	1.14(1.03-1.25)	9.8×10^{-3}	0.25	1.01(0.83-1.24)	0.896
rs979200	**C**/T	0.45	1.07(1.02-1.13)	8.2×10^{-3}	0.67	1.09(0.95-1.24)	0.230
rs16901857	**G**/A	0.24	1.08(1.02-1.15)	6.6×10^{-3}	0.08	1.02(0.82-1.28)	0.841
rs75127456	A/C	0.10	0.89(0.81-0.97)	8.9×10^{-3}	0.07	0.92(0.71-1.19)	0.524
rs56005245	**C**/T	0.26	1.09(1.03-1.16)	2.5×10^{-3}	0.37	1.02(0.89-1.16)	0.780
chr8:128272219:I	**A**/AG	0.04	1.32(1.10-1.59)	2.9×10^{-3}	0.12	1.04(0.83-1.32)	0.715
rs28392817	**T**/G	0.17	1.10(1.03-1.18)	7.1×10^{-3}	0.78	1.1(0.94-1.28)	0.237
rs4733807	**A**/G	0.12	0.89(0.82-0.96)	3.6×10^{-3}	0.16	0.99(0.83-1.18)	0.907
rs55971392	**G**/A	0.10	0.88(0.8-0.96)	4.5×10^{-3}	0.04	0.88(0.61-1.27)	0.501
rs35686742	**C**/T	0.09	0.87(0.79-0.96)	4.4×10^{-3}	0.03	0.91(0.62-1.32)	0.619
rs6988558	**G**/C	0.44	0.98(0.93-1.04)	0.476	0.49	0.83(0.73-0.96)	9.6×10^{-3}
rs73356177	**A**/G	0.05	1.05(0.93-1.19)	0.426	0.10	1.36(1.11-1.67)	2.8×10^{-3}
rs1516964	**C**/T	0.04	0.88(0.76-1.01)	0.072	0.09	0.67(0.51-0.89)	6.1×10^{-3}
rs56142222	**G**/A	0.03	1.08(0.92-1.25)	0.351	0.22	1.22(1.05-1.42)	9.8×10^{-3}

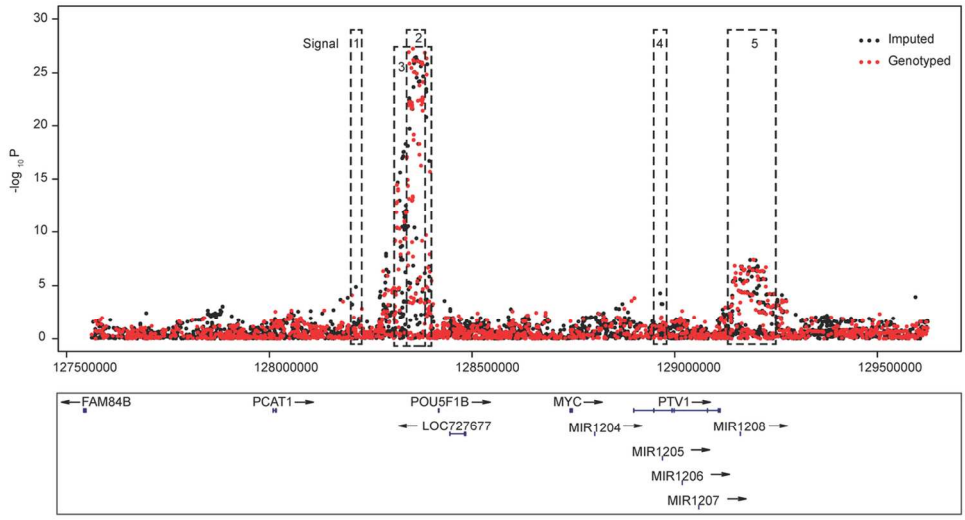
Abbreviations: EAF, effect allele frequency; OR, odds ratio; CI, confidence interval.

[a] Effect/reference allele; effect alleles are shown in bold.

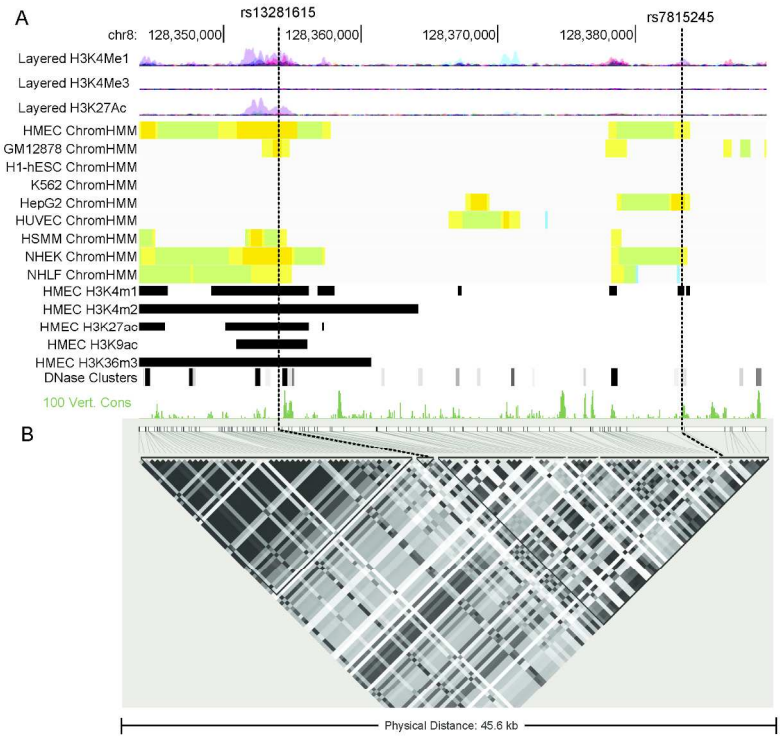
[b] Adjusted for study site and two principal components for each population.

[c] Except for r^2 of 0.23 and 0.43 for linkage disequilibrium between rs7815245 and the index SNP rs13281615 in Asians and African American, all other SNPs are not in LD with either of the index SNPs rs13281615 or rs11780156 in non-European populations ($r^2 < 0.02$).

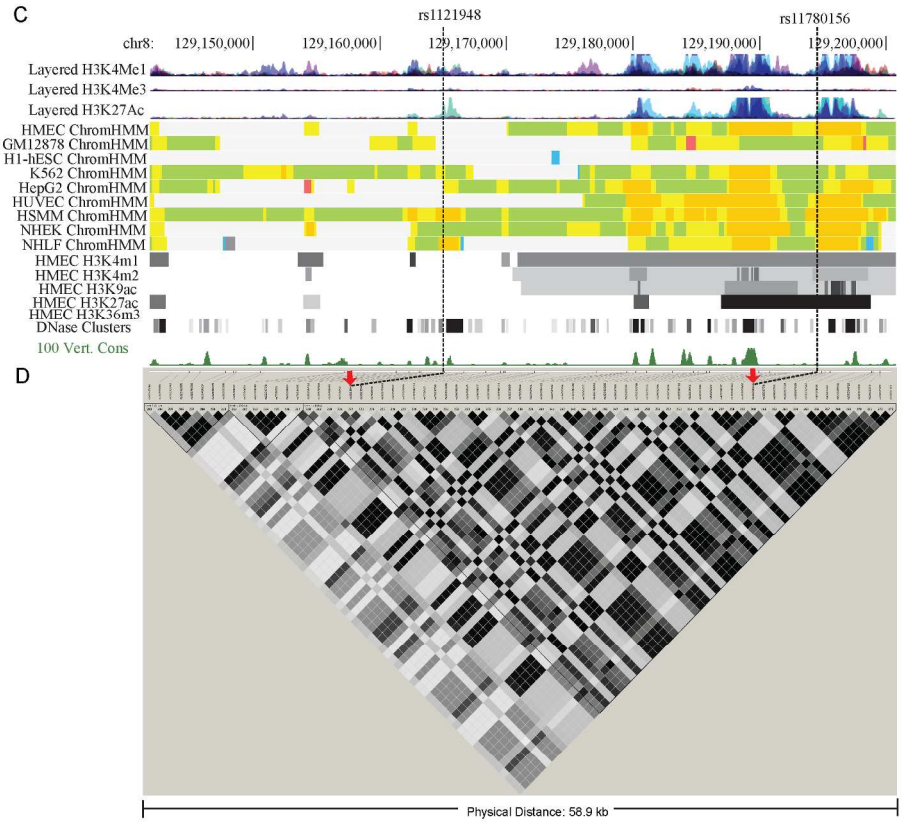
John Wiley & Sons, Inc.



118x63mm (300 x 300 DPI)



210x297mm (300 x 300 DPI)



215x279mm (300 x 300 DPI)