Article:

ABCTB Investigators, , EMBRACE, , GEMO Study Collaborators, et al. (3 more authors) (2017) Identification of ten variants associated with risk of estrogen-receptor-negative breast cancer. Nature Genetics, 49 (12). pp. 1767-1778. ISSN 1061-4036

https://doi.org/10.1038/ng.3785

Identification of ten variants associated with risk of estrogen receptor negative breast cancer

Roger L. Milne[†,1,2,*], Karoline B. Kuchenbaecker[†,3,4], Kyriaki Michailidou[†,3,5], Jonathan Beesley[6], Siddhartha Kar[7], Sara Lindström[8,9], Shirley Hui[10], Audrey Lemaçon[11], Penny Soucy[11], Joe Dennis[3], Xia Jiang[9], Asha Rostamianfar[10], Hilary Finucane[9,12], Manjeet K. Bolla[3], Lesley McGuffog[3], Qin Wang[3], Cora M. Aalfs[13], ABCTB Investigators[14], Marcia Adams[15], Julian Adlard[16], Simona Agata[17], Shahana Ahmed[7], Kristiina Aittomäki[18], Fares Al-Ejeh[19], Jamie Allen[3], Christine B. Ambrosone[20], Christopher I. Amos[21], Irene L. Andrulis[22,23], Hoda Anton-Culver[24], Natalia N. Antonenkova[25], Volker Arndt[26], Norbert Arnold[27], Kristan J. Aronson[28], Bernd Auber[29], Paul L. Auer[30,31], Margreet G.E.M. Ausems[32], Jacopo Azzollini[33], François Bacot[34], Judith Balmaña[35], Monica Barile[36], Laure Barjhoux[37], Rosa B. Barkardottir[38,39], Myrto Barrdahl[40], Daniel Barnes[3], Daniel Barrowdale[3], Caroline Baynes[7], Matthias W. Beckmann[41], Javier Benitez[42-44], Marina Bermisheva[45], Leslie Bernstein[46], Yves-Jean Bignon[47], Kathleen R. Blazer[48], Marinus J. Blok[49], Carl Blomqvist[50], William Blot[51,52], Kristie Bobolis[53], Bram Boeckx[54,55], Natalia V. Bogdanova[25,56,57], Anders Bojesen[58], Stig E. Bojesen[59-61], Bernardo Bonanni[36], Anne-Lise Børresen-Dale[62], Aniko Bozsik[63], Angela R. Bradbury[64], Judith S. Brand[65], Hiltrud Brauch[66-68], Hermann Brenner[26,68,69], Brigitte Bressac-de Paillerets[70], Carole Brewer[71], Louise Brinton[72], Per Broberg[73], Angela Brooks-Wilson[74,75], Joan Brunet[76], Thomas Brüning[77], Barbara Burwinkel[78,79], Saundra S. Buys[80], Jinyoung Byun[21], Qiuyin Cai[51], Trinidad Caldés[81], Maria A. Caligo[82], Ian Campbell[83,84], Federico Canzian[85], Olivier Caron[70], Angel Carracedo[86,87], Brian D. Carter[88], J. Esteban Castelao[89], Laurent Castera[90], Virginie Caux-Moncoutier[91], Salina B. Chan[92], Jenny Chang-Claude[40,93], Stephen J. Chanock[72], Xiaoqing Chen[6], Ting-Yuan David Cheng[94], Jocelyne Chiquette[95], Hans Christiansen[56], Kathleen B.M. Claes[96], Christine L. Clarke[97], Thomas Conner[98], Don M. Conroy[7], Jackie Cook[99], Emilie Cordina-Duverger[100], Sten Cornelissen[101], Isabelle Coupier[102], Angela Cox[103], David Cox[104,105], Simon S. Cross[106], Katarina Cuk[26], Julie M. Cunningham[107], Kamila Czene[65], Mary B. Daly[108], Francesca Damiola[37], Hatef Darabi[65], Rosemarie Davidson[109], Kim De Leeneer[96], Peter Devilee[110,111], Ed Dicks[7], Orland Diez[112], Yuan Chun Ding[46], Nina Ditsch[113], Kimberly F. Doheny[15], Susan M. Domchek[64], Cecilia M. Dorfling[114], Thilo Dörk[57], Isabel dos-Santos-Silva[115], Stéphane Dubois[11], Pierre-Antoine Dugué[1,2], Martine Dumont[11], Alison M. Dunning[7], Lorraine Durcan[116,117], Miriam Dwek[118], Bernd Dworniczak[119], Diana Eccles[117], Ros Eeles[120], Hans Ehrencrona[121], Ursula Eilber[40], Bent Ejlertsen[122], Arif B. Ekici[123], A. Heather Eliassen [124,125], EMBRACE[14], Christoph Engel[126,127], Mikael Eriksson[65], Laura Fachal[7], Laurence Faivre[128,129], Peter A. Fasching[41,130], Ulrike Faust[131], Jonine Figueroa[72,132], Dieter Flesch-Janys[133,134], Olivia Fletcher[135], Henrik Flyger[136], William D. Foulkes[137], Eitan Friedman[138,139], Lin Fritschi[140], Debra Frost[3], Marike Gabrielson[65], Pragna Gaddam[141], Patricia A. Ganz[142], Susan M. Gapstur[88], Judy Garber[143], Vanesa Garcia-Barberan[81], José A. García-Sáenz[81], Mia M. Gaudet[88], Marion Gauthier-Villars[91], Andrea Gehrig[144], GEMO Study Collaborators[14], Vassilios

Georgoulias[145], Anne-Marie Gerdes[146], Graham G. Giles[1,2], Gord Glendon[22], Andrew K Godwin[147], Mark S. Goldberg[148,149], David E. Goldgar[150], Anna González-Neira[42], Paul Goodfellow[151], Mark H. Greene[152], Grethe I. Grenaker Alnæs[62], Mervi Grip[153], Jacek Gronwald[154], Anne Grundy[155], Daphne Gschwantler-Kaulich[156], Pascal Guénel[100], Qi Guo[7], Lothar Haeberle[41], Eric Hahnen[157-159], Christopher A. Haiman[160], Niclas Håkansson[161], Emily Hallberg[162], Ute Hamann[163], Natalie Hammell[34], Susan Hankinson[164], Thomas V.O. Hansen[165], Patricia Harrington[7], Steven N. Hart[162], Jaana M. Hartikainen[166-168], Catherine S. Healey[7], HEBON[14], Alexander Hein[41], Sonja Helbig[57], Alex Henderson[169], Jane Heyworth[170], Belynda Hicks[171], Peter Hillemanns[57], Shirley Hodgson[172], Frans B. Hogervorst[173], Antoinette Hollestelle[174], Maartje J. Hooning[174], Bob Hoover[72], John L. Hopper[2], Chunling Hu[107], Guanmengqian Huang[163], Peter J. Hulick[175,176], Keith Humphreys[65], David J. Hunter[9,125], Evgeny N. Imyanitov[177], Claudine Isaacs[178], Motoki Iwasaki[179], Louise Izatt[180], Anna Jakubowska[154], Paul James[84,181], Ramunas Janavicius[181,182], Wolfgang Janni[183], Uffe Birk Jensen[184], Esther M. John[185,186], Nichola Johnson[135], Kristine Jones[171], Michael Jones[187], Arja Jukkola-Vuorinen[188], Rudolf Kaaks[40], Maria Kabisch[163], Katarzyna Kaczmarek[154], Daehee Kang[189-191], Karin Kast[192], kConFab/AOCS Investigators[14], Renske Keeman[101], Michael J. Kerin[193], Carolien M. Kets[194], Machteld Keupers[195], Sofia Khan[196], Elza Khusnutdinova[45,197], Johanna I. Kiiski[196], Sung-Won Kim[156], Julia A. Knight[198,199], Irene Konstantopoulou[200], Veli-Matti Kosma[166-168], Vessela N. Kristensen[62,201,202], Torben A. Kruse[203], Ava Kwong[204-206], Anne-Vibeke Lænkholm[207], Yael Laitman[138], Fiona Lalloo[208], Diether Lambrechts[54,55], Keren Landsman[209], Christine Lasset[210], Conxi Lazaro[211], Loic Le Marchand[212], Julie Lecarpentier[3], Andrew Lee[3], Eunjung Lee[160], Jong Won Lee[213], Min Hyuk Lee[214], Flavio Lejbkowicz[209], Fabienne Lesueur[215-218], Jingmei Li[65], Jenna Lilyquist[219], Anne Lincoln[220], Annika Lindblom[221], Jolanta Lissowska[222], Wing-Yee Lo[66,67], Sibylle Loibl[223], Jirong Long[51], Jennifer T. Loud[152], Jan Lubinski[154], Craig Luccarini[7], Michael Lush[3], Robert J. MacInnis[1,2], Tom Maishman[116,117], Enes Makalic[2], Ivana Maleva Kostovska[224], Siranoush Manoukian[33], JoAnn E. Manson[225], Sara Margolin[226], John W.M. Martens[174], Maria Elena Martinez[227,228], Keitaro Matsuo[229,230], Dimitrios Mavroudis[145], Sylvie Mazoyer[231], Catriona McLean[232], Hanne Meijers-Heijboer[233], Primitiva Menéndez[234], Jeffery Meyer[107], Hui Miao[235], Austin Miller[236], Nicola Miller[193], Gillian Mitchell[84,181], Marco Montagna[17], Kenneth Muir[237,238], Anna Marie Mulligan[239,240], Claire Mulot[241], Sue Nadesan[53], Katherine L. Nathanson[64], NBSC Collaborators[14], Susan L. Neuhausen[46], Heli Nevanlinna[196], Ines Nevelsteen[195], Dieter Niederacher[242], Sune F. Nielsen[59,60], Børge G. Nordestgaard[59-61], Aaron Norman[162], Robert L. Nussbaum[243], Edith Olah[63], Olufunmilayo I. Olopade[244], Janet E. Olson[162], Curtis Olswold[162], Kai-ren Ong[245], Jan C. Oosterwijk[246], Nick Orr[135], Ana Osorio[43,44], V. Shane Pankratz[247], Laura Papi[248], Tjoung-Won Park-Simon[57], Ylva Paulsson-Karlsson[249], Rachel Peake[250], Inge Søkilde Pedersen[251], Bernard Peissel[33], Ana Peixoto[252], Jose I.A. Perez[253], Paolo Peterlongo[254], Julian Peto[115], Georg Pfeiler[156], Catherine M. Phelan[255], Mila Pinchev[209], Dijana Plaseska-Karanfilska[224], Bruce Poppe[96], Mary E Porteous[256], Ross Prentice[30], Nadege Presneau[118], Darya Prokofieva[197], Elizabth Pugh[15], Miquel

Angel Pujana[257], Katri Pylkäs[258,259], Brigitte Rack[113], Paolo Radice[260], Nazneen Rahman[261], Johanna Rantala[262], Christine Rappaport-Fuerhauser[156], Gad Rennert[209,263], Hedy S. Rennert[209], Valerie Rhenius[7], Kerstin Rhiem[157-159], Andrea Richardson[264], Gustavo C. Rodriguez[265], Atocha Romero[81,266], Jane Romm[15], Matti A. Rookus[267], Anja Rudolph[40], Thomas Ruediger[268], Emmanouil Saloustros[269], Joyce Sanders[270], Dale P. Sandler[271], Suleeporn Sangrajrang[272], Elinor J. Sawyer[273], Daniel F. Schmidt[2], Minouk J. Schoemaker[187], Fredrick Schumacher[160], Peter Schürmann[57], Lukas Schwentner[183], Christopher Scott[162], Rodney J. Scott[274,275], Sheila Seal[261], Leigha Senter[276], Caroline Seynaeve[174], Mitul Shah[7], Priyanka Sharma[277], Chen-Yang Shen[278,279], Xin Sheng[160], Hermela Shimelis[107], Martha J. Shrubsole[51], Xiao-Ou Shu[51], Lucy E Side[280], Christian F. Singer[156], Christof Sohn[281], Melissa C. Southey[282], John J. Spinelli[283,284], Amanda B. Spurdle[6], Christa Stegmaier[285], Dominique Stoppa-Lyonnet[91], Grzegorz Sukiennicki[154], Harald Surowy[78,79], Christian Sutter[286], Anthony Swerdlow[187,287], Csilla I. Szabo[288], Rulla M. Tamimi[9,124,125], Yen Y. Tan[289], Jack A. Taylor[271,290], Maria-Isabel Tejada[291], Maria Tengström[166,292,293], Soo H. Teo[294,295], Mary B. Terry[296], Daniel C. Tessier[34], Alex Teulé[297], Kathrin Thöne[134], Darcy L. Thull[298], Maria Grazia Tibiletti[299], Laima Tihomirova[300], Marc Tischkowitz[137,301], Amanda E. Toland[302], Rob A.E.M. Tollenaar[303], Ian Tomlinson[304], Diana Torres[163,305], Martine Tranchant[11], Thérèse Truong[100], Jonathan Tryer[7], Kathy Tucker[306], Nadine Tung[307], Hans-Ulrich Ulmer[308], Celine Vachon[162], Christi J. van Asperen[309], David Van Den Berg[160], Ans M.W. van den Ouweland[310], Elizabeth J. van Rensburg[114], Liliana Varesco[311], Raymonda Varon-Mateeva[312], Ana Vega[313,314], Alessandra Viel[315], Joseph Vijai[220], Daniel Vincent[34], Jason Vollenweider[107], Lisa Walker[316], Zhaoming Wang[72,317], Shan Wang-Gohrke[183], Barbara Wappenschmidt[157-159], Clarice R. Weinberg[318], Jeffrey N. Weitzel[48], Camilla Wendt[226], Jelle Wesseling[101,270], Alice S. Whittemore[186], Juul T. Wijnen[111,309], Walter Willett [125,319], Robert Winqvist[258,259], Alicja Wolk[161], Anna H. Wu[160], Lucy Xia[160], Xiaohong R. Yang[72], Drakoulis Yannoukakos[200], Daniela Zaffaroni[33], Wei Zheng[51], Bin Zhu[171], Argyrios Ziogas[24], Elad Ziv[320], Kristin K. Zorn[298], Manuela Gago-Dominguez[86,227], Arto Mannermaa[166-168], Håkan Olsson[73], Manuel R. Teixeira[252,321], Jennifer Stone[250,322], Kenneth Offit[323,324], Laura Ottini[325], Sue K. Park[189-191], Mads Thomassen[203], Per Hall[65,326], Alfons Meindl[327], Rita K. Schmutzler[157-159], Arnaud Droit[11], Gary D. Bader[#,10], Paul D.P. Pharoah[#,3,7], Fergus J. Couch[#,107], Douglas F. Easton[#,3,7], Peter Kraft[#,9,125], Georgia Chenevix-Trench[#,6], Montserrat García-Closas[#,72], Marjanka K. Schmidt[#,101,328], Antonis C. Antoniou[#,3], Jacques Simard[#,11]

1. Cancer Epidemiology Centre, Cancer Council Victoria, Melbourne, Australia.
2. Centre for Epidemiology and Biostatistics, Melbourne School of Population and Global health, The University of Melbourne, Melbourne, Australia.
3. Centre for Cancer Genetic Epidemiology, Department of Public Health and Primary Care, University of Cambridge, Cambridge, UK.
4. The Wellcome Trust Sanger Institute, Wellcome Trust Genome Campus, Hinxton, Cambridge, UK.

5. Department of Electron Microscopy/Molecular Pathology, The Cyprus Institute of Neurology and Genetics, Nicosia, Cyprus.
6. Cancer Division, QIMR Berghofer Medical Research Institute, Brisbane, Australia.
7. Centre for Cancer Genetic Epidemiology, Department of Oncology, University of Cambridge, Cambridge, UK.
8. Department of Epidemiology, University of Washington School of Public Health, Seattle, WA, USA.
9. Program in Genetic Epidemiology and Statistical Genetics, Harvard T.H. Chan School of Public Health, Boston, MA, USA.
10. The Donnelly Centre, University of Toronto, Toronto, Canada.
11. Genomics Center, Centre Hospitalier Universitaire de Québec Research Center, Laval University, Québec City, Canada.
12. Department of Mathematics, Massachusetts Institute of Technology, Cambridge, MA, USA.
13. Department of Clinical Genetics, Academic Medical Center, Amsterdam, The Netherlands.
14. A list of members and affiliations appears in the Supplementary Note.
15. Center for Inherited Disease Research (CIDR), Institute ot Genetic Medicine, Johns Hopkins University School of Medicine, Baltimore, MD, USA.
16. Yorkshire Regional Genetics Service, Chapel Allerton Hospital, Leeds, UK.
17. Istituto Oncologico Veneto IOV - IRCCS, Immunology and Molecular Oncology Unit, Padua, Italy.
18. Department of Clinical Genetics, Helsinki University Hospital, University of Helsinki, Helsinki, Finland.
19. Personalised Medicine Team, QIMR Berghofer Medical Research Institute, Brisbane, Australia.
20. Roswell Park Cancer Institute, Buffalo, NY, USA.
21. Center for Genomic Medicine, Department of Biomedical Data Science, Geisel School of Medicine, Dartmouth College, Lebanon, NH, USA.
22. Fred A. Litwin Center for Cancer Genetics, Lunenfeld-Tanenbaum Research Institute of Mount Sinai Hospital, Toronto, Canada.
23. Department of Molecular Genetics, University of Toronto, Toronto, Canada.
24. Department of Epidemiology, University of California Irvine, Irvine, CA, USA.
25. N.N. Alexandrov Research Institute of Oncology and Medical Radiology, Minsk, Belarus.
26. Division of Clinical Epidemiology and Aging Research, German Cancer Research Center (DKFZ), Heidelberg, Germany.
27. Department of Gynaecology and Obstetrics, University Hospital of Schleswig-Holstein, Campus Kiel, Christian-Albrechts University Kiel, Kiel, Germany.
28. Department of Public Health Sciences, and Cancer Research Institute, Queen’s University, Kingston, ON, Canada.
29. Institute of Human Genetics, Hannover Medical School, Hannover, Germany.
30. Cancer Prevention Program, Fred Hutchinson Cancer Research Center, Seattle, WA, USA.
31. Zilber School of Public Health, University of Wisconsin-Milwaukee, Milwaukee, WI, USA.
32. Department of Medical Genetics, University Medical Center Utrecht, Utrecht, The Netherlands.

33. Unit of Medical Genetics, Department of Preventive and Predictive Medicine, Fondazione IRCCS (Istituto Di Ricovero e Cura a Carattere Scientifico) Istituto Nazionale dei Tumori (INT), Milan, Italy.
34. McGill University and Génome Québec Innovation Centre, Montréal, Canada.
35. Department of Medical Oncology. University Hospital, Vall d'Hebron, Barcelona, Spain.
36. Division of Cancer Prevention and Genetics, Istituto Europeo di Oncologia, Milan, Italy.
37. Bâtiment Cheney D, Centre Léon Bérard, Lyon, France.
38. Laboratory of Cell Biology, Department of Pathology, Landspitali-LSH v/Hringbraut, Reykjavik, Iceland.
39. BMC (Biomedical Centre), Faculty of Medicine, University of Iceland, Reykjavik, Iceland.
40. Division of Cancer Epidemiology, German Cancer Research Center (DKFZ), Heidelberg, Germany.
41. Department of Gynaecology and Obstetrics, University Hospital Erlangen, Friedrich-Alexander University Erlangen-Nuremberg, Comprehensive Cancer Center Erlangen-EMN, Erlangen, Germany.
42. Genotyping Unit, Human Cancer Genetics Programme, Spanish National Cancer Research Centre, Madrid, Spain.
43. Human Genetics Group, Human Cancer Genetics Programme, Spanish National Cancer Centre (CNIO), Madrid, Spain.
44. Spanish Network on Rare Diseases (CIBERER), Madrid, Spain.
45. Institute of Biochemistry and Genetics, Ufa Scientific Center of Russian Academy of Sciences, Ufa, Russia.
46. Department of Population Sciences, Beckman Research Institute of City of Hope, Duarte, CA, USA.
47. Department of Oncogenetics, Centre Jean Perrin, BP 392, Clermont-Ferrand, France.
48. Clinical Cancer Genetics, City of Hope, Duarte, California, USA.
49. Department of Clinical Genetics, Maastricht University Medical Center, Maastricht, The Netherlands.
50. Department of Oncology, Helsinki University Hospital, University of Helsinki, Helsinki, Finland.
51. Division of Epidemiology, Department of Medicine, Vanderbilt Epidemiology Center, Vanderbilt-Ingram Cancer Center, Vanderbilt University School of Medicine, Nashville, TN, USA.
52. International Epidemiology Institute, Rockville, MD, USA.
53. City of Hope Clinical Cancer Genomics Community Research Network, Duarte, CA, USA.
54. Vesalius Research Center, VIB, Leuven, Belgium.
55. Laboratory for Translational Genetics, Department of Oncology, University of Leuven, Leuven, Belgium.
56. Department of Radiation Oncology, Hannover Medical School, Hannover, Germany.
57. Gynaecology Research Unit, Hannover Medical School, Hannover, Germany.
58. Department of Clinical Genetics, Vejle Hospital, Vejle, Denmark.
59. Copenhagen General Population Study, Herlev and Gentofte Hospital, Copenhagen University Hospital, Herlev, Denmark.

60. Department of Clinical Biochemistry, Herlev and Gentofte Hospital, Copenhagen University Hospital, Herlev, Denmark.
61. Faculty of Health and Medical Sciences, University of Copenhagen, Copenhagen, Denmark.
62. Department of Cancer Genetics, Institute for Cancer Research, Oslo University Hospital Radiumhospitalet, Oslo, Norway.
63. Department of Molecular Genetics, National Institute of Oncology, Budapest, Hungary.
64. Department of Medicine, Abramson Cancer Center, Perelman School of Medicine at the University of Pennsylvania, hiladelphia, PA, USA.
65. Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Stockholm, Sweden.
66. Dr. Margarete Fischer-Bosch-Institute of Clinical Pharmacology, Stuttgart, Germany.
67. University of Tübingen, Tübingen, Germany.
68. German Cancer Consortium (DKTK), German Cancer Research Center (DKFZ), Heidelberg, Germany.
69. Division of Preventive Oncology, German Cancer Research Center (DKFZ) and National Center for Tumor Diseases (NCT), Heidelberg, Germany.
70. Gustave Roussy, Biopathology Department, Villejuif, F-94805, France.
71. Department of Clinical Genetics, Royal Devon & Exeter Hospital, Exeter, UK.
72. Division of Cancer Epidemiology and Genetics, National Cancer Institute, Rockville, MD, USA.
73. Department of Cancer Epidemiology, Clinical Sciences, Lund University, Lund, Sweden.
74. Genome Sciences Centre, BC Cancer Agency, Vancouver, BC, Canada.
75. Department of Biomedical Physiology and Kinesiology, Simon Fraser University, Burnaby, BC, Canada.
76. Genetic Counseling Unit, Hereditary Cancer Program, IDIBGI (Institut d'Investigació Biomèdica de Girona), Catalan Institute of Oncology, Girona, Spain.
77. Institute for Prevention and Occupational Medicine of the German Social Accident Insurance, Institute of the Ruhr University Bochum, Bochum, Germany.
78. Department of Obstetrics and Gynecology, University of Heidelberg, Heidelberg, Germany.
79. Molecular Epidemiology Group, C080, German Cancer Research Center (DKFZ), Heidelberg, Germany.
80. Department of Medicine, Huntsman Cancer Institute, 2000 Circle of Hope, Salt Lake City, UT, USA.
81. Medical Oncology Department, Hospital Clínico San Carlos, Madrid, Spain.
82. Section of Genetic Oncology, Dept. of Laboratory Medicine, University and University Hospital of Pisa, Pisa, Italy.
83. Research Department, Peter MacCallum Cancer Centre, East Melbourne, Australia.
84. Sir Peter MacCallum, Department of Oncology, The University of Melbourne, Melbourne, Australia.
85. Genomic Epidemiology Group, German Cancer Research Center (DKFZ), Heidelberg, Germany.

86. Genomic Medicine Group, Galician Foundation of Genomic Medicine, Instituto de Investigación Sanitaria de Santiago de Compostela (IDIS), Complejo Hospitalario Universitario de Santiago, Servicio Galego de Saúde SERGAS, Santiago De Compostela, Spain.
87. Centro de Investigación en Red de Enfermedades Raras (CIBERER) y Centro Nacional de Genotipado (CEGEN-PRB2), Universidad de Santiago de Compostela, Spain.
88. Epidemiology Research Program, American Cancer Society, Atlanta, GA, USA.
89. Oncology and Genetics Unit, Instituto de Investigacion Biomedica (IBI) Orense-Pontevedra-Vigo, Xerencia de Xestion Integrada de Vigo-SERGAS, Vigo, Spain.
90. Centre François Baclesse, 3 avenue Général Harris, Caen, France.
91. Service de Génétique Oncologique, Institut Curie, Paris, France.
92. 1600 Divisadero Street, C415, San Francisco, CA, USA.
93. University Cancer Center Hamburg (UCCH), University Medical Center Hamburg-Eppendorf, Hamburg, Germany.
94. Division of Cancer Prevention and Population Sciences, Roswell Park Cancer Institute, Buffalo, NY, USA.
95. Unité de recherche en santé des populations, Centre des maladies du sein Deschênes-Fabia, Hôpital du Saint-Sacrement, Québec, Canada.
96. Center for Medical Genetics, Ghent University, Gent, Belgium.
97. Westmead Institute for Medical Research, University of Sydney, Sydney, Australia.
98. Huntsman Cancer Institute, 2000 Circle of Hope, Salt Lake City, UT, USA.
99. Sheffield Clinical Genetics Service, Sheffield Children's Hospital, Sheffield, UK.
100. Cancer & Environment Group, Center for Research in Epidemiology and Population Health (CESP), INSERM, University Paris-Sud, University Paris-Saclay, Villejuif, France.
101. Division of Molecular Pathology, The Netherlands Cancer Institute - Antoni van Leeuwenhoek Hospital, Amsterdam, The Netherlands.
102. Unité d'Oncogénétique, CHU Arnaud de Villeneuve, Montpellier, France.
103. Academic Unit of Molecular Oncology, Department of Oncology and Metabolism, University of Sheffield, Sheffield, UK.
104. Department of Epidemiology and Biostatistics, School of Public Health, Imperial College London, London, UK.
105. INSERM U1052, Cancer Research Center of Lyon, Lyon, France.
106. Academic Unit of Pathology, Department of Neuroscience, University of Sheffield, Sheffield, UK.
107. Department of Laboratory Medicine and Pathology, Mayo Clinic, Rochester, MN, USA.
108. Department of Clinical Genetics, Timothy R. Talbot Jr. Chair for Cancer Research, Fox Chase Cancer Center, Philadelphia, PA, USA.
109. Department of Clinical Genetics, South Glasgow University Hospitals, Glasgow, UK.
110. Department of Pathology, Leiden University Medical Center, Leiden, The Netherlands.
111. Department of Human Genetics, Leiden University Medical Center, Leiden, The Netherlands.

112. Oncogenetics Group, Vall d’Hebron Institute of Oncology (VHIO), Clinical and Molecular Genetics Area, Vall d’Hebron University Hospital, Barcelona, Spain.
113. Department of Gynecology and Obstetrics, Ludwig-Maximilians University of Munich, Munich, Germany.
114. Cancer Genetics Laboratory, Department of Genetics, University of Pretoria, Arcadia, South Africa.
115. Department of Non-Communicable Disease Epidemiology, London School of Hygiene and Tropical Medicine, London, UK.
116. Southampton Clinical Trials Unit, Faculty of Medicine, University of Southampton, Southampton, UK.
117. Cancer Sciences Academic Unit, Faculty of Medicine, University of Southampton, Southampton, UK.
118. Department of Biomedical Sciences, Faculty of Science and Technology, University of Westminster, London, UK.
119. Institute of Human Genetics, University of Münster, Münster, Germany.
120. Oncogenetics Team, The Institute of Cancer Research and Royal Marsden NHS Foundation Trust, London, UK.
121. Department of Clinical Genetics, Lund University Hospital, Lund, Sweden.
122. Department of Oncology, Rigshospitalet, Copenhagen University Hospital, Copenhagen, Denmark.
123. Institute of Human Genetics, University Hospital Erlangen, Friedrich-Alexander University Erlangen-Nuremberg, Comprehensive Cancer Center Erlangen-EMN, Erlangen, Germany.
124. Channing Division of Network Medicine, Department of Medicine, Brigham and Women's Hospital, Harvard Medical School, Boston, MA, USA.
125. Department of Epidemiology, Harvard T.H. Chan School of Public Health, Boston, MA, USA.
126. Institute for Medical Informatics, Statistics and Epidemiology, University of Leipzig, Leipzig, Germany.
127. LIFE - Leipzig Research Centre for Civilization Diseases, University of Leipzig, Leipzig, Germany
128. Genetics Department, Dijon University Hospital, Dijon, France
129. Oncogenetics, Centre Georges-François Leclerc, Dijon, France.
130. David Geffen School of Medicine, Department of Medicine Division of Hematology and Oncology, University of California at Los Angeles, Los Angeles, CA, USA.
131. Institute of Medical Genetics and Applied Genomics, University of Tuebingen, Germany.
132. Usher Institute of Population Health Sciences and Informatics, The University of Edinburgh Medical School, Edinburgh, UK.
133. Institute for Medical Biometrics and Epidemiology, University Medical Center Hamburg-Eppendorf, Hamburg, Germany.
134. Department of Cancer Epidemiology, Clinical Cancer Registry, University Medical Center Hamburg-Eppendorf, Hamburg, Germany.
135. Breast Cancer Now Toby Robins Research Centre, The Institute of Cancer Research, London, UK.
136. Department of Breast Surgery, Herlev and Gentofte Hospital, Copenhagen University Hospital, Herlev, Denmark.
137. Program in Cancer Genetics, Departments of Human Genetics and Oncology, McGill University, Montreal, Quebec, Canada.

138. The Susanne Levy Gertner Oncogenetics Unit, Institute of Human Genetics, Chaim Sheba Medical Center, Ramat Gan, Israel.
139. Sackler Faculty of Medicine, Tel Aviv University, Ramat Aviv, Israel.
140. School of Public Health, Curtin University, Perth, Australia.
141. Clinical Cancer Genetics Laboratory, Memorial Sloane Kettering Cancer Center, New York, NY, USA.
142. UCLA Schools of Medicine and Public Health, Division of Cancer Prevention & Control Research, Jonsson Comprehensive Cancer Center, Los Angeles, CA, USA.
143. Cancer Risk and Prevention Clinic, Dana-Farber Cancer Institute, Boston, MA, USA.
144. Centre of Familial Breast and Ovarian Cancer, Department of Medical Genetics, Institute of Human Genetics, University Würzburg, Würzburg, Germany.
145. Department of Medical Oncology, University Hospital of Heraklion, Heraklion, Greece.
146. Department of Clincial Genetics, Rigshospitalet 4062, København Ø, Denmark.
147. Department of Pathology and Laboratory Medicine, University of Kansas Medical Center, Kansas City, KS, USA.
148. Department of Medicine, McGill University, Montreal, Canada.
149. Division of Clinical Epidemiology, Royal Victoria Hospital, McGill University, Montreal, Canada.
150. Department of Dermatology, Huntsman Cancer Institute, University of Utah School of Medicine, Salt Lake City, UT, USA.
151. Department of Obstetrics and Gynecology, The Ohio State University James Comprehensive Cancer Center, Columbus, OH, USA.
152. Clinical Genetics Branch, DCEG, NCI, NIH, Bethesda, MD, USA.
153. Department of Surgery, Oulu University Hospital, University of Oulu, Oulu, Finland.
154. Department of Genetics and Pathology, Pomeranian Medical University, Szczecin, Poland.
155. Centre de Recherche du Centre Hospitalier de Université de Montréal (CHUM), Montréal, Québec, Canada.
156. Department of OB/GYN and Comprehensive Cancer Centre, Medical University of Vienna, Vienna, Austria.
157. Center for Familial Breast and Ovarian Cancer, University Hospital of Cologne, Cologne, Germany.
158. Center for Integrated Oncology (CIO), University Hospital of Cologne, Cologne, Germany.
159. Center for Molecular Medicine Cologne (CMMC), University of Cologne, Cologne, Germany.
160. Department of Preventive Medicine, Keck School of Medicine, University of Southern California, Los Angeles, CA, USA.
161. Institute of Environmental Medicine, Karolinska Institutet, Stockholm, Sweden.
162. Department of Health Sciences Research, Mayo Clinic, Rochester, MN, USA.
163. Molecular Genetics of Breast Cancer, German Cancer Research Center (DKFZ), Heidelberg, Germany.
164. Department of Biostatistics & Epidemiology, University of Massachusetts, Amherst, Amherst, MA, USA.

165. Center for Genomic Medicine, Rigshospitalet, Copenhagen University Hospital, Copenhagen, Denmark.
166. Translational Cancer Research Area, University of Eastern Finland, Kuopio, Finland.
167. Institute of Clinical Medicine, Pathology and Forensic Medicine, University of Eastern Finland, Kuopio, Finland.
168. Imaging Center, Department of Clinical Pathology, Kuopio University Hospital, Kuopio, Finland.
169. Institute of Genetic Medicine, Centre for Life, Newcastle Upon Tyne Hospitals NHS Trust, Newcastle upon Tyne, UK.
170. School of Population Health, University of Western Australia, Perth, Australia.
171. Cancer Genomics Research Laboratory (CGR), Division of Cancer Epidemiology and Genetics, National Cancer Institute, Bethesda, MD, USA.
172. Medical Genetics Unit, St George's, University of London, UK.
173. Family Cancer Clinic, The Netherlands Cancer Institute - Antoni van Leeuwenhoek hospital, Amsterdam, The Netherlands.
174. Department of Medical Oncology, Family Cancer Clinic, Erasmus MC Cancer Institute, Rotterdam, The Netherlands.
175. Center for Medical Genetics, NorthShore University HealthSystem, Evanston, IL, USA.
176. Pritzker School of Medicine, University of Chicago, Evanston, IL, USA.
177. N.N. Petrov Institute of Oncology, St.-Petersburg, Russia.
178. Lombardi Comprehensive Cancer Center, Georgetown University, Washington, DC, USA.
179. Division of Epidemiology, Center for Public Health Sciences, National Cancer Center, Tokyo, Japan.
180. Clinical Genetics, Guy’s and St. Thomas’ NHS Foundation Trust, London, UK.
181. Familial Cancer Centre, Peter MacCallum Cancer Centre, Melbourne, Australia.
182. State Research Institute Centre for Innovative medicine, Vilnius, Lithuania.
183. Department of Gynaecology and Obstetrics, University of Ulm, Ulm, Germany.
184. Department of Clinical Genetics, Aarhus University Hospital, Aarhus N, Denmark.
185. Department of Epidemiology, Cancer Prevention Institute of California, Fremont, CA, USA.
186. Departments of Health Research and Policy and Biomedical Data Sciences, Stanford University School of Medicine, Stanford, CA, USA.
187. Division of Genetics and Epidemiology, The Institute of Cancer Research, London, UK.
188. Department of Oncology, Oulu University Hospital, University of Oulu, Oulu, Finland.
189. Department of Preventive Medicine, Seoul National University College of Medicine, Seoul, Korea.
190. Department of Biomedical Sciences, Seoul National University College of Medicine, Seoul, Korea.
191. Cancer Research Institute, Seoul National University, Seoul, Korea.
192. Department of Gynecology and Obstetrics, University Hospital Carl Gustav Carus, TU Dresden, Dresden, Germany.
193. School of Medicine, National University of Ireland, Galway, Ireland.

194. Department of Human Genetics, Radboud University Nijmegen Medical Centre, Nijmegen, The Netherlands.
195. Leuven Multidisciplinary Breast Center, Department of Oncology, Leuven Cancer Institute, University Hospitals Leuven, Leuven, Belgium.
196. Department of Obstetrics and Gynecology, Helsinki University Hospital, University of Helsinki, Helsinki, Finland.
197. Department of Genetics and Fundamental Medicine, Bashkir State University, Ufa, Russia.
198. Prosserman Centre for Health Research, Lunenfeld-Tanenbaum Research Institute of Mount Sinai Hospital, Toronto, Canada.
199. Division of Epidemiology, Dalla Lana School of Public Health, University of Toronto, Toronto, Canada.
200. Molecular Diagnostics Laboratory, INRASTES, National Centre for Scientific Research "Demokritos", Athens, Greece.
201. Institute of Clinical Medicine, Faculty of Medicine, University of Oslo, Oslo, Norway.
202. Department of Clinical Molecular Biology, Oslo University Hospital, University of Oslo, Oslo, Norway.
203. Department of Clinical Genetics, Odense University Hospital, Odense C, Denmark.
204. Hong Kong Hereditary Breast Cancer Family Registry, Hong Kong.
205. Department of Surgery, The University of Hong Kong, Hong Kong.
206. Department of Surgery, Hong Kong Sanatorium and Hospital, Hong Kong.
207. Department of Pathology, University Hospital of Region Zealand, division Slagelse, Slagelse, Denmark.
208. Genetic Medicine, Manchester Academic Health Sciences Centre, Central Manchester University Hospitals NHS Foundation Trust, Manchester, UK.
209. Clalit National Cancer Control Center, Haifa, Israel.
210. Unité de Prévention et d'Epidémiologie Génétique, Centre Léon Bérard, Lyon, France.
211. Molecular Diagnostic Unit, Hereditary Cancer Program, IDIBELL (Bellvitge Biomedical Research Institute),Catalan Institute of Oncology, Barcelona, Spain.
212. University of Hawaii Cancer Center, Honolulu, HI, USA.
213. Department of Surgery, University of Ulsan College of Medicine and Asan Medical Center, Seoul, Korea.
214. Department of Surgery, Soonchunhyang University and Hospital, Seoul, Korea.
215. Institut Curie, Paris, France.
216. PSL Research Unviersity, Paris, France.
217. Inserm, U900, Paris, France.
218. Mines Paris Tech, Fontainebleau, France.
219. Department of Health Sciences Research, Mayo Clinic, Scottsdale, AZ, USA.
220. Clinical Genetics Research Laboratory, Dept. of Medicine, Memorial Sloan Kettering Cancer Center, New York, NY, USA.
221. Department of Molecular Medicine and Surgery, Karolinska Institutet, Stockholm, Sweden.
222. Department of Cancer Epidemiology and Prevention, M. Sklodowska-Curie Memorial Cancer Center & Institute of Oncology, Warsaw, Poland.
223. German Breast Group, Neu Isenburg, Germany.

224. Research Centre for Genetic Engineering and Biotechnology "Georgi D. Efremov" , Macedonian Academy of Sciences and Arts, Skopje, Republic of Macedonia.
225. Department of Medicine, Brigham and Women's Hospital, Harvard Medical School, Boston, MA, USA.
226. Department of Oncology - Pathology, Karolinska Institutet, Stockholm, Sweden.
227. Moores Cancer Center, University of California San Diego, La Jolla, CA, USA.
228. Department of Family Medicine and Public Health, University of California San Diego, La Jolla, CA, USA.
229. Division of Molecular Medicine, Aichi Cancer Center Research Institute, Nagoya, Japan.
230. Department of Epidemiology, Nagoya University Graduate School of Medicine, Nagoya, Japan.
231. Lyon Neuroscience Research Center - CRNL, INSERM U1028, CNRS UMR5292, University of Lyon, Lyon, France.
232. Anatomical Pathology, The Alfred Hospital, Melbourne, Australia.
233. Department of Clinical Genetics, VU University Medical Centre, Amsterdam, the Netherlands.
234. Servicio de Anatomía Patológica, Hospital Monte Naranco, Oviedo, Spain.
235. Saw Swee Hock School of Public Health, National University of Singapore, Singapore, Singapore.
236. NRG Oncology, Statistics and Data Management Center, Roswell Park Cancer Institute, Buffalo, NY, USA.
237. Institute of Population Health, University of Manchester, Manchester, UK.
238. Division of Health Sciences, Warwick Medical School, Warwick University, Coventry, UK.
239. Department of Laboratory Medicine and Pathobiology, University of Toronto, Toronto, Canada.
240. Laboratory Medicine Program, University Health Network, Toronto, Canada.
241. Université Paris Sorbonne Cité, INSERM UMR-S1147, Paris, France.
242. Department of Gynecology and Obstetrics, University Hospital Düsseldorf, Heinrich-Heine University Düsseldorf, Düsseldorf, Germany.
243. 513 Parnassus Ave., HSE 901E, San Francisco, CA, USA.
244. Center for Clinical Cancer Genetics and Global Health, The University of Chicago, Chicago, IL, USA.
245. West Midlands Regional Genetics Service, Birmingham Women’s Hospital Healthcare NHS Trust, Edgbaston, Birmingham, UK.
246. Department of Genetics, University Medical Center, Groningen University, Groningen, The Netherlands.
247. University of New Mexico Health Sciences Center, University of New Mexico, Albuquerque, NM, USA.
248. Unit of Medical Genetics, Department of Biomedical, Experimental and Clinical Sciences, University of Florence, Florence, Italy.
249. Department of Immunology, Genetics and Pathology, Uppsala University,Uppsala, Sweden.
250. The Curtin UWA Centre for Genetic Origins of Health and Disease, Curtin University and University of Western Australia, Perth, Australia.
251. Section of Molecular Diagnostics, Clinical Biochemistry, Aalborg University Hospital, Aalborg, Denmark.

252. Department of Genetics, Portuguese Oncology Institute, Porto, Portugal.
253. Servicio de Cirugía General y Especialidades, Hospital Monte Naranco, Oviedo, Spain.
254. IFOM, The FIRC (Italian Foundation for Cancer Research) Institute of Molecular Oncology, Milan, Italy.
255. Department of Cancer Epidemiology, Moffitt Cancer Center, Tampa, FL, USA.
256. South East of Scotland Regional Genetics Service, Western General Hospital, Edinburgh, UK.
257. ProCURE, Catalan Institute of Oncology, IDIBELL (Bellvitge Biomedical Research Institute), Barcelona, Spain.
258. Laboratory of Cancer Genetics and Tumor Biology, Cancer and Translational Medicine Research Unit, Biocenter Oulu, University of Oulu, Oulu, Finland.
259. Laboratory of Cancer Genetics and Tumor Biology, Northern Finland Laboratory Centre Oulu, Oulu, Finland.
260. Unit of Molecular Bases of Genetic Risk and Genetic Testing, Department of Preventive and Predictive Medicine, Fondazione IRCCS (Istituto Di Ricovero e Cura a Carattere Scientifico) Istituto Nazionale dei Tumori (INT), Milan, Italy.
261. Section of Cancer Genetics, The Institute of Cancer Research, London, UK.
262. Department of Clinical Genetics, Karolinska University Hospital, Stockholm, Sweden.
263. Carmel Medical Center and B. Rappaport Faculty Of Medicine-Technion, Haifa, Israel.
264. Brigham and Women's Hospital, Dana-Farber Cancer Institute, Boston, MA USA.
265. Division of Gynecologic Oncology, NorthShore University HealthSystem, University of Chicago, Evanston, IL, USA.
266. Medical Oncology Department, Hospital Universitario Puerta de Hierro, Madrid, Spain.
267. Department of Epidemiology. Netherlands Cancer Institute, Amsterdam, The Netherlands.
268. Institute of Pathology, Staedtisches Klinikum Karlsruhe, Karlsruhe, Germany.
269. Hereditary Cancer Clinic, University Hospital of Heraklion, Heraklion, Greece.
270. Department of Pathology, The Netherlands Cancer Institute - Antoni van Leeuwenhoek hospital, Amsterdam, The Netherlands.
271. Epidemiology Branch, National Institute of Environmental Health Sciences, NIH, Research Triangle Park, NC, USA.
272. National Cancer Institute, Bangkok, Thailand.
273. Research Oncology, Guy’s Hospital, King's College London, London, UK.
274. Division of Molecular Medicine, Pathology North, John Hunter Hospital, Newcastle, Australia.
275. Discipline of Medical Genetics, School of Biomedical Sciences and Pharmacy, Faculty of Health, University of Newcastle, Callaghan, Australia.
276. Clinical Cancer Genetics Program, Division of Human Genetics, Department of Internal Medicine, The Comprehensive Cancer Center, The Ohio State University, Columbus, USA.
277. Department of Medicine, Kansas Medicial Center, Kansas City, KS, USA.
278. School of Public Health, China Medical University, Taichung, Taiwan.
279. Taiwan Biobank, Institute of Biomedical Sciences, Academia Sinica, Taipei, Taiwan.

280. North East Thames Regional Genetics Service, Great Ormond Street Hospital for Children NHS Trust, London, UK.
281. National Center for Tumor Diseases, University of Heidelberg, Heidelberg, Germany.
282. Department of Pathology, The University of Melbourne, Melbourne, Australia.
283. Cancer Control Research, BC Cancer Agency, Vancouver, BC, Canada.
284. School of Population and Public Health, University of British Columbia, Vancouver, BC, Canada.
285. Saarland Cancer Registry, Saarbrücken, Germany.
286. Institute of Human Genetics, University Hospital Heidelberg, Heidelberg, Germany.
287. Division of Breast Cancer Research, The Institute of Cancer Research, London, UK.
288. National Human Genome Research Institute ,National Institutes of Health, Bethesda, MD, USA.
289. Dept of OB/GYN and Comprehensive Cancer Center, Medical University of Vienna, Vienna, Austria.
290. Epigenetic and Stem Cell Biology Laboratory, National Institute of Environmental Health Sciences, NIH, Research Triangle Park, NC, USA.
291. Molecular Genetics Laboratory, Clinical Genetics Service, Cruces University Hospital. BioCruces Health Research Institute, Barakaldo, Spain.
292. Cancer Center, Kuopio University Hospital, Kuopio, Finland.
293. Institute of Clinical Medicine, Oncology, University of Eastern Finland, Kuopio, Finland.
294. Cancer Research Malaysia, Subang Jaya, Selangor, Malaysia.
295. Breast Cancer Research Unit, Cancer Research Institute, University Malaya Medical Centre, Kuala Lumpur, Malaysia.
296. Department of Epidemiology, Mailman School of Public Health, Columbia University, New York, NY, USA.
297. Genetic Counseling Unit, Hereditary Cancer Program, IDIBELL (Bellvitge Biomedical Research Institute), Catalan Institute of Oncology, Barcelona, Spain.
298. Magee-Womens Hospital, University of Pittsburgh School of Medicine, Pittsburgh, PA, USA.
299. Ospedale di Circolo ASST Settelaghi, Varese, Italy.
300. Latvian Biomedical Research and Study Centre, Riga, Latvia.
301. Department of Medical Genetics, Addenbrooke's Treatment Centre, Addenbrooke's Hosptital, Cambridge, UK.
302. Department of Molecular Virology, Immunology and Medical Genetics, Comprehensive Cancer Center, The Ohio State University, Columbus, OH, USA.
303. Department of Surgery, Leiden University Medical Center, Leiden, The Netherlands.
304. Wellcome Trust Centre for Human Genetics and Oxford NIHR Biomedical Research Centre, University of Oxford, Oxford, UK.
305. Institute of Human Genetics, Pontificia Universidad Javeriana, Bogota, Colombia.
306. Hereditary Cancer Clinic, Department of Medical Oncology, Prince of Wales Hospital, Randwick, Australia.

307. Department of Medical Oncology, Beth Israel Deaconess Medical Center, MA, USA.
308. Frauenklinik der Stadtklinik Baden-Baden, Baden-Baden, Germany.
309. Department of Clinical Genetics, Leiden University Medical Center, Leiden, The Netherlands.
310. Department of Clinical Genetics, Erasmus University Medical Center, Rotterdam, The Netherlands.
311. Unit of Hereditary Cancer, Department of Epidemiology, Prevention and Special Functions, IRCCS (Istituto Di Ricovero e Cura a Carattere Scientifico) AOU San Martino - IST Istituto Nazionale per la Ricerca sul Cancro, largo Rosanna Benzi 10, 16132 Genoa, Italy.
312. Institute of Human Genetics, Campus Virchov Klinikum, Charite Berlin, Germany.
313. Fundación Pública Galega de Medicina Xenómica, Servizo Galego de Saúde (SERGAS), Instituto de Investigaciones Sanitarias (IDIS), Santiago de Compostela, Spain.
314. Grupo de Medicina Xenómica, Centro de Investigación Biomédica en Red de Enfermedades Raras (CIBERER), Universidade de Santiago de Compostela (USC), Santiago de Compostela, Spain.
315. Unit of Functional onco-genomics and genetics, CRO Aviano, National Cancer Institute, Via Franco Gallini 2, 33081 Aviano (PN), Italy.
316. Oxford Regional Genetics Service, Churchill Hospital, Oxford, UK.
317. Department of Computational Biology, St. Jude Children’s Research Hospital, Memphis, TN, USA.
318. Biostatistics and Computational Biology Branch, National Institute of Environmental Health Sciences, NIH, Research Triangle Park, NC, USA.
319. Department of Nutrition, Harvard T.H. Chan School of Public Health, Boston, MA, USA.
320. Department of Medicine, Institute for Human Genetics, UCSF Helen Diller Family Comprehensive Cancer Center, University of California San Francisco, San Francisco, CA, USA.
321. Biomedical Sciences Institute (ICBAS), University of Porto, Porto, Portugal.
322. Department of Obstetrics and Gynaecology, University of Melbourne and the Royal Women's Hospital, Melbourne, Australia.
323. Clinical Genetics Research Lab, Cancer Biology and Genetics Program, Sloan Kettering Institute, New York, NY, USA
324. Clinical Genetics Service, Department of Medicine, Memorial Sloan-Kettering Cancer Center, New York, NY, USA.
325. Department of Molecular Medicine, University La Sapienza, c/oPoliclinico Umberto I, Rome, Italy.
326. Department of Oncology, South General Hospital, Stockholm, Sweden
327. Division of Gynaecology and Obstetrics, Technische Universität München, Munich, Germany.
328. Division of Psychosocial Research and Epidemiology, The Netherlands Cancer Institute - Antoni van Leeuwenhoek hospital, Amsterdam, The Netherlands.

[ǂ]Co-first authorship
[#]Co-senior authorship

*Correspondence to: Roger L. Milne, Cancer Epidemiology Centre, Cancer Council Victoria, Melbourne, Australia; E-mail: roger.milne@cancervic.org.au.

Most common breast cancer susceptibility variants have been identified through genome-wide association studies (GWASs) of predominantly estrogen receptor (ER)-positive disease[1]. We conducted a GWAS using 21,468 ER-negative cases and 100,594 controls combined with 18,908 *BRCA1* mutation carriers (9,414 with breast cancer), all of European origin. We identified independent associations at $P<5x10^{-8}$ with 10 variants at nine novel loci. At P<0.05, we replicated associations with 10 of 11 variants previously reported in ER-negative or *BRCA1* mutation carrier GWASs, and observed consistent associations with ER-negative disease for 105 susceptibility variants identified by other breast cancer GWASs. These 125 variants explain approximately 16% of the familial risk of this breast cancer subtype. There was high genetic correlation (0.72) between risk of ER-negative breast cancer and breast cancer risk for *BRCA1* carriers. These findings will likely lead to improved risk prediction and inform further fine-mapping and functional work to better understand the biological basis of ER-negative breast cancer.

GWASs have identified 107 single nucleotide polymorphisms (SNPs) that are independently associated with breast cancer risk[2-32]. Association studies focused on ER-negative disease, or *BRCA1* mutation carriers, who are more likely to develop ER-negative disease (70-80% of cases)[33], have identified 11 of these SNPs[3,9,12,19,29,30]. We aimed to discover additional ER-negative breast cancer susceptibility variants by performing a GWAS in women of European origin.

New genotyping data were generated for 9,655 ER-negative cases and 45,494 controls from 68 Breast Cancer Association Consortium (BCAC) studies and 15,566 *BRCA1* mutation carriers (7,784 with breast cancer) from 58 Consortium of Investigators of Modifiers of BRCA1/2 (CIMBA) studies (Supplementary Tables 1 and 2) using the Illumina OncoArray beadchip, a 570K SNP custom array with genome-wide coverage[34]. Imputation was used to derive estimated genotypes for ~21M SNPs, using the 1000 Genomes Project (Phase 3) as reference; ~11.5M of those with imputation $r^2>0.3$ and minor allele frequency (MAF)>0.005 were included in further analyses. For BCAC data, we estimated per-allele odds ratios (ORs) using logistic regression, adjusting for country and principal components. For CIMBA data, we estimated per-allele hazard ratios (HR) using a retrospective cohort analysis framework, modelling time to breast cancer and stratifying on country, Ashkenazi Jewish origin and birth cohort[35,36] (see Online Methods). These analyses were also applied to an independent set of previously generated data from other genome-wide genotyping of additional European participants in 44 BCAC studies (11,813 ER-negative cases and 55,100 controls)[9,12,16,20,37,38] and 54 CIMBA studies (3,342 *BRCA1* mutation carriers, 1,630 with breast cancer) (Supplementary Tables 1 and 2). Fixed-effects meta-analysis was used to combine results across genotyping initiatives within consortia and, assuming that the OR and HR estimates approximate the same underlying relative risk, across consortia[39].

Results from the combined meta-analysis are summarised in Supplementary Figures 1 and 2. There was minimal inflation of test statistics (lambda1000=1.004;

Supplementary Figure 3). We identified 10 variants at nine novel loci that were independently associated with risk of ER-negative breast cancer at $P<5x10^{-8}$ (Table 1; Supplementary Table 3; Supplementary Figures 4-11). Two independent signals were observed within 12kb at 11q22.3, for rs74911261 (MAF=0.02) and rs11374964 (MAF=0.42); OR estimates and statistical significance were largely unchanged when each variant was adjusted for the other (Supplementary Table 4). The association with 8p23.3-rs66823261 was not observed for *BRCA1* mutation carriers (P=0.32, P-heterogeneity=0.030).

For each of these 10 novel signals, we identified candidate causal SNPs analytically[40,41] (see Online Methods) and combined multiple sources of *in silico* functional annotation from public databases[42-52] to identify likely functional variants and target genes. Results are summarised in Supplementary Table 5 (including UCSC Genome Browser links; see also Supplementary Note), Figure 1 and Supplementary Figures 4-11 (data sources in Supplementary Table 6). Many candidate causal SNPs lie in predicted regulatory regions and are associated with expression of nearby genes in blood or other tissues. At 2p23, the predicted target genes include *ADCY3* and *NCOA1* (Supplementary Figure 4). At 6q23.1 (Supplementary Figure 5), the most plausible target gene is *L3MBTL3*[53]. A predicted target at 8q24.13 is *FBXO32*, which is expressed in ER-negative HMECs but not ER-positive MCF7 breast cancer cells (Supplementary Figure 7) and has a known role in cancer cachexia[54]. At 11q22.3 (Figure 1), a predicted target gene of common risk-associated variants is *NPAT*[55]. The rarer SNPs underlying the other 11q22.3 signal are predicted to target *ATM*, a known breast cancer susceptibility gene[56]. Three rare coding variants (MAF≤0.03) in *ATM, NPAT* and *KDELC2*, are also among the candidate causal SNPs at this locus. At 16p13, predicted target genes include *ADCY9* and *CREBBP* (Supplementary Figure 8). At 19q12 (Supplementary Figure 11), a potential target gene encodes cyclin E1 which is involved in cell cycle control and phosphorylation of NPAT[57].

Expression QTL associations were assessed between each candidate causal variant and genes within 1Mb using 79 ER-negative breast tumours from TCGA and 135 normal breast tissue samples from METABRIC[58-60]. The strongest associations identified were 6q23.1-rs6569648-*L3MBTL3* ($P=4.3x10^{-6}$) and 18q12.1-rs12965632-*CDH2* ($P=1.0x10^{-4}$), both in METABRIC (Supplementary Table 5). SNP rs6569648 was the top *cis*-eQTL (of all imputed variants within 1 Mb) for *L3MBTL3* while the p-value for the rs12965632-*CDH2* eQTL was within two orders of magnitude of the top *cis*-eQTLs for this gene (Supplementary Figures 12-13).

For 10 of the 11 variants previously identified through GWASs of ER-negative disease or overall disease in *BRCA1* mutation carriers[3,9,12,18,19,30,31], or reported as more strongly associated with ER-negative breast cancer[29], associations with ER-negative disease were replicated (P<0.05) using OncoArray data from BCAC, which does not overlap with any of the discovery studies (Table 2). Effect sizes were generally similar to those originally reported. Using all available CIMBA data, six of these 11 variants were associated with breast cancer risk (P<0.05) for *BRCA1* mutation carriers (Table 2). No evidence of association was observed for 20q11-rs2284378[12] in either BCAC or CIMBA (P≥0.46).

Based on estimated ORs using BCAC data for all cases with known ER status (16,988 ER-negative; 65,275 ER-positive), all 10 new and 10 previously reported and replicated ER-negative disease susceptibility SNPs were more strongly associated with risk of ER-negative than ER-positive subtype (P-heterogeneity<0.05, except for novel hit 19p13.2-rs322144; Supplementary Table 7). Two variants (1q32.1-rs4245739 and 19p13.11-rs67397200) were not associated with ER-positive disease. For four variants (11q22.3- rs11374964, 11q22.3-rs74911261, 1q32.1-rs6678914 and 2p23.2-rs4577244), the risk-associated allele for ER-negative disease was associated with reduced risk of ER-positive disease (P<0.05).

For these 20 ER-negative breast cancer susceptibility SNPs, we also assessed associations by triple-negative (TN) status (negative for ER, progesterone receptor and HER2; Table 3), tumour grade (Table 4) and age at diagnosis (Supplementary Table 8) using BCAC data only. Five, including the novel susceptibility variants 11q22.3-rs11374964 and 11q22.3-rs74911261, were more strongly associated with risk of both TN and higher-grade disease (P<0.05), although after adjustment for TN status, heterogeneity by grade was observed only for 11q22.3-rs74911261 and 1q32.1-rs4245739 (P<0.05). For 2p23.3-rs4577244, heterogeneity was observed for grade only, while 6q25.2-rs2747652 was more strongly associated with risk of other (non-TN) ER-negative breast cancer subtypes (P<0.05). At younger ages, associations appeared to be stronger for two variants (5p15.33-rs10069690 and 19p13.11-rs67397200), and weaker for one (6q25.2-rs2747652) (P<0.05).

Elsewhere we report 65 novel susceptibility loci for overall breast cancer[1]. Three of these overlap within 500kb with the novel ER-negative disease-associated loci reported here (variants 2p23.3-rs200648189, 6q23.1-rs6569648 and 8q24.13-rs17350191). We assessed associations with risk of ER-negative disease, and with risk of overall breast cancer for *BRCA1* mutation carriers, for SNPs at the remaining 62 loci, as well as for the 96 previously reported breast cancer susceptibility variants that were not ER-negative specific. Of these 158 SNPs, 105 were associated (P<0.05) with risk of ER-negative breast cancer, and 24 with risk for *BRCA1* mutation carriers (Supplementary Tables 9-10). Results for *BRCA2* mutation carriers are presented in Supplementary Table 11.

Pathway analysis based on mapping each SNP to the nearest gene was performed using summary association statistics from the meta-analysis of BCAC and CIMBA data combined[61-64] (see Online Methods). This identified several pathways implicated in ER-negative disease (enrichment score [ES]≥0.41; Supplementary Figure 14; Supplementary Tables 12-13), including a subset that was not enriched in susceptibility to ER-positive disease (ES<0; Supplementary Table 14). One of the latter subsets was the adenylate cyclase (AC) activating pathway (ES=0.62; Supplementary Figure 15). Two of the predicted target genes for the 10 novel ER-negative breast cancer susceptibility variants, based on the eQTL analysis (Supplementary Table 5), *ADCY3* (P[TCGA]=6.7×10^{-3}] and *ADCY9* (P[METABRIC]=1.3×10^{-4}), are part of this pathway, and their association signals were critical to the elevated ES observed (Supplementary Figure 14). ADCY9 is stimulated by β2 adrenergic receptor (β2AR) signalling[65] in ER-negative breast cancer[66], which in turn drives AC-cAMP signalling, including for example mitogenic signalling through β-arrestin-Src-ERK[67].

To further explore the functional properties of the genome that contribute to ER-negative breast cancer heritability, we conducted a partitioned heritability analysis using linkage disequilibrium (LD) score regression[68]. Considering 52 “baseline” genomic features, we observed the greatest enrichment for super-enhancers (2.5-fold, $p=2x10^{-7}$) and the H3K4me3 histone mark (2.4-fold, p=0.0005), with 33% depletion (p=0.0002) observed for repressed regions (Supplementary Table 15). No differences in enrichment for these features were observed between susceptibility to ER-negative and ER-positive breast cancer, but baseline genomic features are not specific to cell type[68]. The estimated correlation between ER-negative and ER-positive breast cancer based on ~1M common genetic variants[69,70] was 0.60 (standard error [SE], 0.03) indicating that, although these two breast cancer subtypes have a shared genetic component, a substantial proportion is distinct. The estimated correlation between ER-negative disease in the general population and overall breast cancer for *BRCA1* mutation carriers was 0.72 (SE, 0.11).

In summary, in this study of women of European origin, we have identified 10 novel susceptibility variants for ER-negative breast cancer and replicated associations with ER-negative disease for 10 SNPs identified by previous GWASs. Most of these were not associated, or more weakly associated, with ER-positive disease, consistent with the findings from pathway and partitioned heritability analyses showing that ER-negative breast cancer has a partly distinct genetic aetiology. We also observed consistent associations with ER-negative disease for a further 105 overall breast cancer susceptibility SNPs. Together, these 125 variants explain ~14% of an assumed 2-fold increased risk of developing ER-negative disease for the first degree female relatives of women affected with this subtype (the newly identified SNPs explain ~1.5%); Supplementary Table 16) and ~40% of the estimated familial risk that is attributable to all variants imputable from the Oncoarray (see Online Methods). We have also identified nine novel breast cancer susceptibility variants for *BRCA1* mutation carriers and confirmed associations for a further 30 previously reported SNPs; these 39 variants explain ~8% of the variance in polygenic risk for carriers of these mutations (Supplementary Table 17). However, the lower number of *BRCA1* risk-associated variants may merely be a consequence of the smaller sample size, since the genetic correlation with ER-negative breast cancer is high. These findings will likely inform improved risk prediction, both for the general population and for *BRCA1* mutation carriers[30,71,72]. Further investigation is required for other populations of non-European origin. Fine-mapping and functional studies should lead to a better understanding of the biological basis of ER-negative breast cancer, and perhaps inform the design of more effective preventive interventions, early detection and treatments for this disease.

Data availability

A subset of the data that support the findings of this study is publically available via dbGaP (see URLs section; accession number phs001265.v1.p1). The complete dataset will not be made publically available due to restraints imposed by the ethics committees of individual studies; requests for data can be made to the corresponding author or the Data Access Coordination Committees (DACCs) of BCAC (see URLs section) and CIMBA (see URLs section). BCAC DACC approval is required to access data from studies ABCFS, ABCS, ABCTB, BBCC, BBCS, BCEES, BCFR-NY, BCFR-PA, BCFR-UT, BCINIS, BSUCH, CBCS, CECILE, CGPS, CTS, DIETCOMPLYF, ESTHER, GC-HBOC, GENICA, GEPARSIXTO, GESBC, HABCS, HCSC, HEBCS, HMBCS, HUBCS, KARBAC, KBCP, LMBC, MABCS,

MARIE, MBCSG, MCBCS, MISS, MMHS, MTLGEBCS, NC-BCFR, OFBCR, ORIGO, pKARMA, POSH, PREFACE, RBCS, SKKDKFZS, SUCCESSB, SUCCESSC, SZBCS, TNBCC, UCIBCS, UKBGS and UKOPS (see Supplementary Table 1). CIMBA DACC approval is required to access data from studies BCFR-ON, CONSIT TEAM, DKFZ, EMBRACE, FPGMX, GC-HBOC, GEMO, G-FAST, HEBCS, HEBON, IHCC, INHERIT, IOVHBOCS, IPOBCS, MCGILL, MODSQUAD, NAROD, OCGN, OUH and UKGRFOCR (see Supplementary Table 2).

URLs

dbGaP: https://www.ncbi.nlm.nih.gov/gap
BCAC: http://bcac.ccge.medschl.cam.ac.uk/
CIMBA: http://cimba.ccge.medschl.cam.ac.uk/
PCcalc software: http://ccge.medschl.cam.ac.uk/software/pccalc/
SNPTEST: https://mathgen.stats.ox.ac.uk/genetics_software/snptest/snptest.html
GeneSets: http://baderlab.org/GeneSets
GenGen package: http://gengen.openbioinformatics.org/en/latest/

Acknowledgements

Genotyping of the OncoArray was funded by the Government of Canada through Genome Canada and the Canadian Institutes of Health Research (GPH-129344), the *Ministère de l'Économie, de la Science et de l'Innovation du Québec* through Genome Québec, the Quebec Breast Cancer Foundation for the PERSPECTIVE project, the US National Institutes of Health (NIH) [1 U19 CA 148065 for the Discovery, Biology and Risk of Inherited Variants in Breast Cancer (DRIVE) project and X01HG007492 to the Center for Inherited Disease Research (CIDR) under contract number HHSN268201200008I], Cancer Research UK [C1287/A16563], Odense University Hospital Research Foundation (Denmark), the National R&D Program for Cancer Control - Ministry of Health & Welfare (Republic of Korea) [1420190], the Italian Association for Cancer Research [AIRC, IG16933], the Breast Cancer Research Foundation, the National Health and Medical Research Council (Australia) and German Cancer Aid [110837].

Genotyping of the iCOGS array was funded by the European Union [HEALTH-F2-2009-223175], Cancer Research UK [C1287/A10710, C1287/A10118, C12292/A11174], NIH grants (CA128978, CA116167, CA176785) and Post-Cancer GWAS initiative (1U19 CA148537, 1U19 CA148065 and 1U19 CA148112 - the GAME-ON initiative), an NCI Specialized Program of Research Excellence (SPORE) in Breast Cancer (CA116201) the Canadian Institutes of Health Research (CIHR) for the CIHR Team in Familial Risks of Breast Cancer, the *Ministère de l'Économie, Innovation et Exportation du Québec* (#PSR-SIIRI-701), Komen Foundation for the Cure, the Breast Cancer Research Foundation and the Ovarian Cancer Research Fund.

Combining the GWAS data was supported in part by NIH Cancer Post-Cancer GWAS initiative [1 U19 CA 148065] (DRIVE, part of the GAME-ON initiative). LD score regression analysis was supported by CA194393.

BCAC is funded by Cancer Research UK [C1287/A16563] and by the European Union via its Seventh Framework Programme [HEALTH-F2-2009-223175, (COGS)] and Horizon 2020 Research & Innovation Programme [633784 (B-CAST); 634935 (BRIDGES)]. CIMBA is funded by Cancer Research UK [C12292/A20861 and C12292/A11174].

We thank all the individuals who took part in these studies and all the researchers, clinicians, technicians and administrative staff who have enabled this work to be carried out.

For a full description of funding and acknowledgments, see the Supplementary Note.

Author Contributions

Writing group: R.L.M., K.B.K, K.Michailidou, J.Beesley, S.Kar, S.Lindström, S.Hui., G.D.B., P.D.P.P., F.J.C., D.F.E., P.K., G.CT., M.GC., M.K.S., A.C.A., J.Simard.

Conceived and coordinated the synthesis of the Oncoarray: D.F.E., A.C.A., J. Simard, C.I.A., J.Byun, S.J.C., E.D., D.J.H., A.Lee, P.D.P.P., J.T., Z.W.

OncoArray genotyping: M.A., A.C.A., S.E.B., M.K.B., F.B., G.CT., J.M.C., K.F.D., D.F.E., N.Hammell, B.Hicks, K.J., C.Luccarini, L.M, J.M., E.P., J.Romm, M.K.S., X.S., J.Simard., P.Soucy, D.C.T., D.V., J.Vollenweider, L.X., B.Z.

Oncoarray genotype calling and quality control: X.C., J.D., E.D., D.F.E., K.B.K, J.Lecarpentier, A.Lee, M.Lush.

Database management: D.Barrowdale., M.K.B., M.L., L.McG., Q.W., R.Keeman, M.K.S.

Statistical analysis: K.B.K, K.Michailidou, S.Hui, S.Kar, X.J., A.Rostamianfar, H.Finucane, S.Lindström, D.Barnes, P.K., P.D.P.P., G.D.B., R.L.M., A.C.A., D.F.E.

Bioinformatic analysis: J.Beesley, P.Soucy, A.Lemaçon, D.Barnes, F.AE. A.D., J. Simard, G.CT.

Provided DNA samples and/or phenotypic data: ABCTB.I., C.M.A., J.Adlard, S.Agata, S.Ahmed, J.Allen, K.A., C.B.A., I.L.A., H.AC., N.N.A., A.C.A., V.A., N.A., K.J.A., B.A., P.L.A., M.G.E.M.A., J.Azzollini, J.Balmaña, M.Barile, L.Barjhoux, R.B.B., M.Barrdahl, D.Barnes, D. Barrowdale, C.Baynes, M.W.B., J.Beesley, J.Benitez, M.Bermisheva, L.Bernstein, YJ.B., K.R.B., M.J.B., C.Blomqvist, W.B., K.B., B.Boeckx, N.V.B., A.Bojesen, S.E.B., M.K.B., B.Bonanni, A.Bozsik, A.R.B., J.S.B., H.Brauch, H.Brenner, B.BdP., C.Brewer, L.Brinton, P.B., A.BW., J.Brunet, T.B., B.Burwinkel, S.S.B., AL.BW., Q.C., T.Caldés, M.A.C., I.Campbell, F.C., O.C., A.Carracedo, B.D.C., J.E.C., L.C., V.CM., S.B.C., J.CC., S.J.C., X.C., G.CT., TYD.C., J.Chiquette, H.C., K.B.M.C., C.L.C., NBCS.C., T.Conner, D.M.C., J.Cook, E.CD., S.C., F.J.C., I.Coupier, D.C., A.Cox, S.S.C., K.Cuk, K.Czene, M.B.D., F.D., H.D., R.D., K.D., J.D., P.D., O.D., YC.D., N.D., S.M.D., C.M.D., S.D., PA.D., M.Dumont, A.M.D., L.D., M.Dwek, B.D., T.D., EMBRACE, D.F.E., D.E., R.E., H.Ehrencrona, U.E., B.E., A.B.E., A.H.E., C.E., M.E., L.Fachal, L.Faivre, P.A.F., U.F., J.F., D.FJ., O.F., H.Flyger, W.D.F., E.F., L.Fritschi, D.F., GEMO.S.C., M.Gabrielson, P.Gaddam, M.GD., P.A.G., S.M.G., J.Garber, V.GB., M.GC., J.A.GS., M.M.G., M.GV., A.Gehrig, V.G., AM.G., G.G.G., G.G., A.KG., M.S.G., D.E.G., A.GN., P.Goodfellow, M.H.G., G.I.GA., M.Grip, J.Gronwald, A.Grundy, D.GK., Q.G., P.Guénel, HEBON, L.H., E.Hahnen, C.A.H., P.Hall, E.Hallberg, U.H., S.Hankinson, T.V.O.H., P.Harrington, S.N.H., J.M.H., C.S.H., A.Hein, S.Helbig, A.Henderson, J.H., P.Hillemanns, S.Hodgson, F.B.H., A.Hollestelle, M.J.H., B.Hoover, J.L.H., C.H., G.H., P.J.H., K.H., D.J.H., N.Håkansson, E.N.I., C.I., M.I., L.I., A.J., P.J., R.J., W.J., UB.J., E.M.J., N.J., M.J., A.JV., R.Kaaks, M.Kabisch, K.Kaczmarek, D.K., K.Kast, R.Keeman, M.J.K., C.M.K., M.Keupers, S.Khan, E.K., J.I.K., J.A.K., I.K., V.K., P.K., V.N.K., T.A.K., K.B.K., A.K., Y.L., F.Lalloo, K.L., D.L., C.Lasset, C.Lazaro, L.IM., J.Lecarpentier, M.Lee, A.Lee, E.L., J.Lee, F.Lejbkowicz, F.Lesueur, J.Li, J.Lilyquist, A.Lincoln, A.Lindblom, S.Lindström, J.Lissowska, WY.L., S.Loibl, J.Long, J.T.L., J.Lubinski, C.Luccarini, M.Lush, AV.L., R.J.M., T.M., E.M., I.MK., A.Mannermaa, S.Manoukian, J.E.M., S.Margolin, J.W.M.M., ME.M., K.Matsuo, D.M., S.Mazoyer, L.M., C.McLean, H.MH., A.Meindl, P.M., H.M., K.Michailidou, A.Miller, N.M., R.L.M., G.M., M.M., K.Muir, A.M.M., C.Mulot, S.N., K.L.N., S.L.N., H.N., I.N., D.N., S.F.N., B.G.N., A.N., R.L.N., K.Offit, E.O., O.I.O., J.E.O., H.O., C.O., K.Ong, J.C.O., N.O., A.O., L.O., VS.P., L.P., S.K.P., TW.PS., Y.PK., R.Peake, IS.P., B.Peissel, A.P., J.I.A.P., P.P., J.P., G.P., P.D.P.P., C.M.P., M.P., D.PK., B.Poppe, M.EP., R.Prentice, N.P., D.P., MA.P., K.P., B.R., P.R., N.R., J.Rantala, C.RF., H.S.R., G.R., V.R., K.R., A.Richardson, G.C.R., A.Romero, M.A.R., A.Rudolph, T.R., E.S., J.Sanders, D.P.S., S.Sangrajrang, E.J.S., D.F.S., M.K.S., R.K.S., M.J.Schoemaker, F.S., L.Schwentner, P.Schürmann, C.Scott, R.J.S., S.Seal, L.Senter, C.Seynaeve, M.S., P.Sharma, CY.S., H.Shimelis, M.J.Shrubsole, XO.S., L.E.S., J.Simard, C.F.S., C.Sohn, P.Soucy, M.C.S., J.J.S., A.B.S., C.Stegmaier, J.Stone, D.SL., G.S., H.Surowy, C.Sutter, A.S., C.I.S., R.M.T., Y.Y.T., J.A.T., M.R.T., MI.T., M.Tengström, S.H.T., M.B.T., A.T., M.Thomassen, D.L.T., K.Thöne, MG.T., L.T., M.Tischkowitz, A.E.T., R.A.E.M.T., I.T., D.T., M.Tranchant, T.T., K.Tucker, N.T., HU.U., C.V., D.vdB., L.V., R.VM., A.Vega, A.Viel, J.Vijai, L.W., Q.W., S.WG., B.W., C.R.W., J.N.W., C.W., J.W., A.S.W., J.T.W., W.W., R.W., A.W., A.H.W., X.R.Y., D.Y., D.Z., W.Z., A.Z., E.Z., K.K.Z., I.dSS., kConFab.AOCS.I., C.J.v.A., E.vR., A.M.W.vdO.

All authors read and approved the final version of the manuscript.

Competing Financial Interests

The authors confirm that they have no competing financial interests

References

1. Michailidou, K. *et al.* Identification of more than 70 new breast cancer susceptibility loci for breast cancer and definition of risk-associated genomic features *Nature* (under review).
2. Ahmed, S. *et al.* Newly discovered breast cancer susceptibility loci on 3p24 and 17q23.2. *Nat Genet* **41**, 585-90 (2009).
3. Antoniou, A.C. *et al.* A locus on 19p13 modifies risk of breast cancer in BRCA1 mutation carriers and is associated with hormone receptor-negative breast cancer in the general population. *Nat Genet* **42**, 885-92 (2010).
4. Cai, Q. *et al.* Genome-wide association study identifies breast cancer risk variant at 10q21.2: results from the Asia Breast Cancer Consortium. *Hum Mol Genet* **20**, 4991-9 (2011).
5. Cox, A. *et al.* A common coding variant in CASP8 is associated with breast cancer risk. *Nat Genet* **39**, 352-8 (2007).
6. Easton, D.F. *et al.* Genome-wide association study identifies novel breast cancer susceptibility loci. *Nature* **447**, 1087-93 (2007).
7. Fletcher, O. *et al.* Novel breast cancer susceptibility locus at 9q31.2: results of a genome-wide association study. *J Natl Cancer Inst* **103**, 425-35 (2011).
8. Ghoussaini, M. *et al.* Genome-wide association analysis identifies three new breast cancer susceptibility loci. *Nat Genet* **44**, 312-8 (2012).
9. Haiman, C.A. *et al.* A common variant at the TERT-CLPTM1L locus is associated with estrogen receptor-negative breast cancer. *Nat Genet* **43**, 1210-4 (2011).
10. Hein, R. *et al.* Comparison of 6q25 breast cancer hits from Asian and European Genome Wide Association Studies in the Breast Cancer Association Consortium (BCAC). *PLoS One* **7**, e42380 (2012).
11. Hunter, D.J. *et al.* A genome-wide association study identifies alleles in FGFR2 associated with risk of sporadic postmenopausal breast cancer. *Nat Genet* **39**, 870-4 (2007).
12. Siddiq, A. *et al.* A meta-analysis of genome-wide association studies of breast cancer identifies two novel susceptibility loci at 6q14 and 20q11. *Hum Mol Genet* **21**, 5373-84 (2012).
13. Stacey, S.N. *et al.* Common variants on chromosomes 2q35 and 16q12 confer susceptibility to estrogen receptor-positive breast cancer. *Nat Genet* **39**, 865-9 (2007).
14. Stacey, S.N. *et al.* Common variants on chromosome 5p12 confer susceptibility to estrogen receptor-positive breast cancer. *Nat Genet* **40**, 703-6 (2008).
15. Thomas, G. *et al.* A multistage genome-wide association study in breast cancer identifies two new risk alleles at 1p11.2 and 14q24.1 (RAD51L1). *Nat Genet* **41**, 579-84 (2009).
16. Turnbull, C. *et al.* Genome-wide association study identifies five new breast cancer susceptibility loci. *Nat Genet* **42**, 504-7 (2010).
17. Zheng, W. *et al.* Genome-wide association study identifies a new breast cancer susceptibility locus at 6q25.1. *Nat Genet* **41**, 324-8 (2009).
18. Bojesen, S.E. *et al.* Multiple independent variants at the TERT locus are associated with telomere length and risks of breast and ovarian cancer. *Nat Genet* **45**, 371-84 (2013).
19. Garcia-Closas, M. *et al.* Genome-wide association studies identify four ER negative-specific breast cancer risk loci. *Nat Genet* **45**, 392-8 (2013).

20. Michailidou, K. *et al.* Large-scale genotyping identifies 41 new loci associated with breast cancer risk. *Nat Genet* **45**, 353-61 (2013).
21. Cai, Q. *et al.* Genome-wide association analysis in East Asians identifies breast cancer susceptibility loci at 1q32.1, 5q14.3 and 15q26.1. *Nat Genet* **46**, 886-90 (2014).
22. Long, J. *et al.* Genome-wide association study in east Asians identifies novel susceptibility loci for breast cancer. *PLoS Genet* **8**, e1002532 (2012).
23. Michailidou, K. *et al.* Genome-wide association analysis of more than 120,000 individuals identifies 15 new susceptibility loci for breast cancer. *Nat Genet* **47**, 373-80 (2015).
24. Milne, R.L. *et al.* Common non-synonymous SNPs associated with breast cancer susceptibility: findings from the Breast Cancer Association Consortium. *Hum Mol Genet* **23**, 6096-111 (2014).
25. Gaudet, M.M. *et al.* Identification of a BRCA2-specific modifier locus at 6p24 related to breast cancer risk. *PLoS Genet* **9**, e1003173 (2013).
26. Meyer, K.B. *et al.* Fine-scale mapping of the FGFR2 breast cancer risk locus: putative functional variants differentially bind FOXA1 and E2F1. *Am J Hum Genet* **93**, 1046-60 (2013).
27. Orr, N. *et al.* Fine-mapping identifies two additional breast cancer susceptibility loci at 9q31.2. *Hum Mol Genet* **24**, 2966-84 (2015).
28. French, J.D. *et al.* Functional variants at the 11q13 risk locus for breast cancer regulate cyclin D1 expression through long-range enhancers. *Am J Hum Genet* **92**, 489-503 (2013).
29. Dunning, A.M. *et al.* Breast cancer risk variants at 6q25 display different phenotype associations and regulate ESR1, RMND1 and CCDC170. *Nat Genet* **48**, 374-86 (2016).
30. Couch, F.J. *et al.* Identification of four novel susceptibility loci for oestrogen receptor negative breast cancer. *Nat Commun* **7**, 11375 (2016).
31. Lawrenson, K. *et al.* Functional mechanisms underlying pleiotropic risk alleles at the 19p13.1 breast-ovarian cancer susceptibility locus. *Nat Commun* **7**, 12675 (2016).
32. Wyszynski, A. *et al.* An intergenic risk locus containing an enhancer deletion in 2q35 modulates breast cancer risk by deregulating IGFBP5 expression. *Hum Mol Genet* (2016).
33. Mavaddat, N. *et al.* Pathology of breast and ovarian cancers among BRCA1 and BRCA2 mutation carriers: results from the Consortium of Investigators of Modifiers of BRCA1/2 (CIMBA). *Cancer Epidemiol Biomarkers Prev* **21**, 134-47 (2012).
34. Amos, C.I. *et al.* The OncoArray Consortium: a Network for Understanding the Genetic Architecture of Common Cancers. *Cancer Epidemiol Biomarkers Prev* (2016).
35. Antoniou, A.C. *et al.* A weighted cohort approach for analysing factors modifying disease risks in carriers of high-risk susceptibility genes. *Genet Epidemiol* **29**, 1-11 (2005).
36. Barnes, D.R., Lee, A., Easton, D.F. & Antoniou, A.C. Evaluation of association methods for analysing modifiers of disease risk in carriers of high-risk mutations. *Genet Epidemiol* **36**, 274-91 (2012).
37. Ahsan, H. *et al.* A genome-wide association study of early-onset breast cancer identifies PFKM as a novel breast cancer gene and supports a

common genetic spectrum for breast cancer at any age. *Cancer Epidemiol Biomarkers Prev* **23**, 658-69 (2014).

38. Stevens, K.N. *et al.* 19p13.1 is a triple-negative-specific breast cancer susceptibility locus. *Cancer Res* **72**, 1795-803 (2012).
39. Willer, C.J., Li, Y. & Abecasis, G.R. METAL: fast and efficient meta-analysis of genomewide association scans. *Bioinformatics* **26**, 2190-1 (2010).
40. Maller, J.B. *et al.* Bayesian refinement of association signals for 14 loci in 3 common diseases. *Nat Genet* **44**, 1294-301 (2012).
41. Udler, M.S., Tyrer, J. & Easton, D.F. Evaluating the power to discriminate between highly correlated SNPs in genetic association studies. *Genet Epidemiol* **34**, 463-8 (2010).
42. ENCODE Project Consortium. A user's guide to the encyclopedia of DNA elements (ENCODE). *PLoS Biol* **9**, e1001046 (2011).
43. Kheradpour, P. & Kellis, M. Systematic discovery and characterization of regulatory motifs in ENCODE TF binding experiments. *Nucleic Acids Res* **42**, 2976-87 (2014).
44. Kundaje, A. *et al.* Integrative analysis of 111 reference human epigenomes. *Nature* **518**, 317-30 (2015).
45. Boyle, A.P. *et al.* Annotation of functional variation in personal genomes using RegulomeDB. *Genome Res* **22**, 1790-7 (2012).
46. He, B., Chen, C., Teng, L. & Tan, K. Global view of enhancer-promoter interactome in human cells. *Proc Natl Acad Sci U S A* **111**, E2191-9 (2014).
47. Rao, S.S. *et al.* A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping. *Cell* **159**, 1665-80 (2014).
48. Corradin, O. *et al.* Combinatorial effects of multiple enhancer variants in linkage disequilibrium dictate levels of gene expression to confer susceptibility to common traits. *Genome Res* **24**, 1-13 (2014).
49. Forrest, A.R. *et al.* A promoter-level mammalian expression atlas. *Nature* **507**, 462-70 (2014).
50. GTEx Consortium. Human genomics. The Genotype-Tissue Expression (GTEx) pilot analysis: multitissue gene regulation in humans. *Science* **348**, 648-60 (2015).
51. Hnisz, D. *et al.* Super-enhancers in the control of cell identity and disease. *Cell* **155**, 934-47 (2013).
52. Westra, H.J. *et al.* Systematic identification of trans eQTLs as putative drivers of known disease associations. *Nat Genet* **45**, 1238-43 (2013).
53. James, L.I. *et al.* Small-molecule ligands of methyl-lysine binding proteins: optimization of selectivity for L3MBTL3. *J Med Chem* **56**, 7358-71 (2013).
54. Sukari, A., Muqbil, I., Mohammad, R.M., Philip, P.A. & Azmi, A.S. F-BOX proteins in cancer cachexia and muscle wasting: Emerging regulators and therapeutic opportunities. *Semin Cancer Biol* **36**, 95-104 (2016).
55. Ling Zheng, L. *et al.* Interaction of Heat Shock Protein Cpn10 with the Cyclin E/Cdk2 Substrate Nuclear Protein Ataxia-Telangiectasia (NPAT) Is Involved in Regulating Histone Transcription. *J Biol Chem* **290**, 29290-300 (2015).
56. Easton, D.F. *et al.* Gene-panel sequencing and the prediction of breast-cancer risk. *N Engl J Med* **372**, 2243-57 (2015).
57. Rogers, S. *et al.* Cyclin E2 is the predominant E-cyclin associated with NPAT in breast cancer cells. *Cell Div* **10**, 1 (2015).
58. Li, Q. *et al.* Integrative eQTL-based analyses reveal the biology of breast cancer risk loci. *Cell* **152**, 633-41 (2013).

59. Cancer Genome Atlas Network. Comprehensive molecular portraits of human breast tumours. *Nature* **490**, 61-70 (2012).
60. Curtis, C. *et al.* The genomic and transcriptomic architecture of 2,000 breast tumours reveals novel subgroups. *Nature* **486**, 346-52 (2012).
61. Merico, D., Isserlin, R., Stueker, O., Emili, A. & Bader, G.D. Enrichment map: a network-based method for gene-set enrichment visualization and interpretation. *PLoS One* **5**, e13984 (2010).
62. Wang, K., Li, M. & Bucan, M. Pathway-based approaches for analysis of genomewide association studies. *Am J Hum Genet* **81**, 1278-83 (2007).
63. Wang, K., Li, M. & Hakonarson, H. Analysing biological pathways in genome-wide association studies. *Nat Rev Genet* **11**, 843-54 (2010).
64. Wang, L., Jia, P., Wolfinger, R.D., Chen, X. & Zhao, Z. Gene set analysis of genome-wide association studies: methodological issues and perspectives. *Genomics* **98**, 1-8 (2011).
65. Hacker, B.M. *et al.* Cloning, chromosomal mapping, and regulatory properties of the human type 9 adenylyl cyclase (ADCY9). *Genomics* **50**, 97-104 (1998).
66. Melhem-Bertrandt, A. *et al.* Beta-blocker use is associated with improved relapse-free survival in patients with triple-negative breast cancer. *J Clin Oncol* **29**, 2645-52 (2011).
67. Pon, C.K., Lane, J.R., Sloan, E.K. & Halls, M.L. The beta2-adrenoceptor activates a positive cAMP-calcium feedforward loop to drive breast cancer cell invasion. *FASEB J* **30**, 1144-54 (2016).
68. Finucane, H.K. *et al.* Partitioning heritability by functional annotation using genome-wide association summary statistics. *Nat Genet* **47**, 1228-35 (2015).
69. Bulik-Sullivan, B. *et al.* An atlas of genetic correlations across human diseases and traits. *Nat Genet* **47**, 1236-41 (2015).
70. Bulik-Sullivan, B.K. *et al.* LD Score regression distinguishes confounding from polygenicity in genome-wide association studies. *Nat Genet* **47**, 291-5 (2015).
71. Milne, R.L. & Antoniou, A.C. Genetic modifiers of cancer risk for BRCA1 and BRCA2 mutation carriers. *Ann Oncol* **22 Suppl 1**, i11-7 (2011).
72. Mavaddat, N. *et al.* Prediction of breast cancer risk based on profiling with common genetic variants. *J Natl Cancer Inst* **107**(2015).

Figure legends

Figure 1. Genomic region around independent ER negative risk associated variants, 11_108345515_G_A (rs11374964) and 11_108357137_G_A (rs74911261). One Mb region showing statistical significance of all genotyped and imputed SNPs and positions of candidate causal variants for two independent signals (shown below as red or blue ticks) in relation to RefSeq genes. Missense variants are labelled with asterisks. Breast cell enhancers overlapping candidate SNPs predicted to target nearby genes by IM-PET[46] are depicted as black bars. Chromatin interactions from ENCODE ChIA-PET in MCF7 cells overlapping candidate variants are shaded to reflect interaction confidence scores. Epigenomic features (derived from publicly available ChIP-seq and DNase-seq) that overlap candidate variants are shown as red or blue segments, depending on the intersected signal. Density tracks show the summed occurrence of ChIP-seq and DNase-seq peak signals at each position. Roadmap Epigenomics Project chromatin state models for HMEC and myoepithelial cells grouped into enhancer, promoter or transcribed annotations are shown as yellow, red or green segments, respectively. Transcript levels in MCF7 and HMEC cells are represented by histograms depicting mean normalised RNA-seq expression. All MCF7 ChIA-PET (ENCODE) and HMEC Hi-C[47] chromatin interactions are represented by black and blue arcs, respectively. NHGRI catalog GWAS SNPs are shown as green ticks. All Oncoarray SNPs (genotyped or imputed) are shown as black ticks and uninterrogated, common SNPs (dbSNP138, EUR MAF > 1%) as red ticks. Features may be examined in detail via exploration of a custom UCSC Genome Browser session accessible via hyperlinks within Supplementary Table 5.

Table 1: Ten novel loci associated with risk of estrogen receptor (ER)-negative breast cancer using meta-analysis of BCAC and CIMBA data

Location	SNP	Chr	Position	Nearest gene	Alleles[#]	BCAC ER-negative[†]			CIMBA *BRCA1* mutation carriers[‡]			Meta-analysis	Heterogeneity
						MAF	OR (95%CI)	P-value	MAF	HR (95%CI)	P-value	P-value	P-value[¥]
2p23.3	rs200648189	2	24739694	*NCOA1*	CT/C	0.19	0.94 (0.91-0.97)	$4.7x10^{-4}$	0.20	0.88 (0.84-0.92)	$3.3x10^{-7}$	$9.7x10^{-9}$	$2.0x10^{-2}$
6q23.1	rs6569648	6	130349119	*L3MBTL3*	T/C	0.23	0.93 (0.90-0.95)	$4.3x10^{-8}$	0.22	0.94 (0.90-0.98)	$5.4x10^{-3}$	$8.3x10^{-10}$	0.64
8p23.3	rs66823261	8	170692	*RPL23AP53*	T/C	0.23	1.09 (1.06-1.12)	$5.6x10^{-9}$	0.22	1.02 (0.98-1.07)	0.32	$3.3x10^{-8}$	$3.0x10^{-2}$
8q24.13	rs17350191	8	124757661	*ANXA13*	C/T	0.34	1.07 (1.04-1.09)	$2.0x10^{-8}$	0.34	1.08 (1.04-1.12)	$1.9x10^{-4}$	$1.7x10^{-11}$	0.81
11q22.3	rs11374964	11	108345515	*KDELC2*	G/GA	0.42	0.94 (0.92-0.96)	$3.6x10^{-8}$	0.43	0.91 (0.88-0.95)	$1.3x10^{-6}$	$4.1x10^{-13}$	0.26
11q22.3	rs74911261	11	108357137	*KDELC2*	G/A	0.02	0.82 (0.75-0.89)	$2.3x10^{-6}$	0.02	0.74 (0.65-0.84)	$2.0x10^{-6}$	$5.4x10^{-11}$	0.17
16p13.3	rs11076805	16	4106788	*ADCY9*	C/A	0.25	0.92 (0.90-0.95)	$2.2x10^{-8}$	0.25	0.96 (0.92-1.00)	0.073	$1.4x10^{-8}$	0.14
18q12.1	rs36194942	18	25401204	*CDH2*	A/AT	0.30	0.94 (0.91-0.96)	$2.5x10^{-7}$	0.31	0.95 (0.91-0.99)	$1.4x10^{-2}$	$1.4x10^{-8}$	0.50
19p13.2	rs322144	19	11423703	*TSPAN16*	C/G	0.47	0.95 (0.93-0.97)	$2.4x10^{-5}$	0.46	0.92 (0.89-0.96)	$3.7x10^{-5}$	$7.4x10^{-9}$	0.23
19q12	rs113701136	19	30277729	*CCNE1*	C/T	0.32	1.07 (1.04-1.09)	$1.7x10^{-7}$	0.32	1.05 (1.01-1.09)	$1.2x10^{-2}$	$6.8x10^{-9}$	0.57

[#] More common allele listed first, minor allele second; [†] Combined data from 21,468 ER-negative cases and 100,594 controls of European ancestry from the Breast Cancer Association Consortium (BCAC); [‡] Combined data from 18,908 *BRCA1* mutation carriers from the Consortium of Investigators of Modifiers of BRCA1/2 (CIMBA), 9,414 of whom had developed breast cancer; [¥] Test for heterogeneity in effect size for ER-negative disease and overall disease for *BRCA1* mutation carriers

Chr, chromosome; MAF, minor allele frequency; OR, odds ratio per copy of the minor allele; CI, confidence interval; HR, hazard ratio per copy of the minor allele

Table 2: Previously reported estrogen receptor (ER)-negative hits: replication using independent data from BCAC and combined results using all BCAC and CIMBA data

							INDEPENDENT REPLICATION			ALL AVAILABLE DATA COMBINED			
Location	SNP	Chr	Position	Ref	Nearest gene	Alleles[#]	BCAC ER-negative (OncoArray)*			BCAC ER-negative[†]		CIMBA *BRCA1*[‡]	
							MAF	OR (95%CI)	P-value	OR (95%CI)	P-value	HR (95%CI)	P-value
1q32.1	rs6678914	1	202187176	19	*LGR6*	G/A	0.41	0.94 (0.91-0.97)	$1.1x10^{-4}$	0.92 (0.90-0.94)	$2.6x10^{-12}$	0.98 (0.95-1.02)	0.31
1q32.1	rs4245739	1	204518842	19	*MDM4*	A/C	0.26	1.12 (1.09-1.17)	$9.2x10^{-11}$	1.14 (1.11-1.16)	$3.1x10^{-23}$	1.09 (1.04-1.13)	$7.3x10^{-5}$
2p24.1	rs12710696	2	19320803	19	*MIR4757*	C/T	0.37	1.04 (1.00-1.07)	$2.5x10^{-2}$	1.06 (1.04-1.09)	$6.5x10^{-8}$	1.01 (0.98-1.05)	0.49
2p23.2	rs4577244[‡]	2	29120733	30	*WDR43*	C/T	0.34	0.93 (0.89-0.96)	$9.6x10^{-5}$	0.92 (0.90-0.95)	$1.5x10^{-9}$	0.92 (0.88-0.96)	$1.3x10^{-4}$
5p15.33	rs10069690	5	1279790	9,18	*TERT*	C/T	0.26	1.19 (1.14-1.23)	$3.8x10^{-21}$	1.18 (1.15-1.21)	$1.5x10^{-35}$	1.18 (1.14-1.23)	$3.7x10^{-16}$
6q25.1	rs3757322[‡]	6	151942194	29	*ESR1*	T/G	0.32	1.14 (1.10-1.18)	$5.5x10^{-14}$	1.15 (1.12-1.18)	$2.8x10^{-31}$	1.14 (1.10-1.19)	$2.9x10^{-12}$
6q25.2	rs2747652[‡]	6	152437016	29	*ESR1*	C/T	0.48	0.92 (0.89-0.95)	$1.1x10^{-7}$	0.91 (0.89-0.93)	$1.9x10^{-18}$	1.00 (0.97-1.04)	0.96
13q22.1	rs6562760[‡]	13	73957681	30	*KLF5*	G/A	0.24	0.92 (0.88-0.95)	$5.0x10^{-6}$	0.92 (0.90-0.95)	$8.7x10^{-10}$	0.89 (0.86-0.93)	$3.5x10^{-7}$
16q12.2	rs11075995	16	53855291	19	*FTO*	T/A	0.30	1.07 (1.03-1.11)	$3.3x10^{-4}$	1.09 (1.06-1.12)	$1.0x10^{-10}$	1.01 (0.97-1.06)	0.49
19p13.11	rs67397200	19	17401404	3,31	*ANKLE1*	C/G	0.32	1.17 (1.13-1.21)	$7.0x10^{-20}$	1.17 (1.14-1.19)	$2.7x10^{-37}$	1.18 (1.14-1.23)	$2.7x10^{-17}$
20q11.21	rs2284378	20	32588095	12	*RALY*	C/T	0.32	0.99 (0.95-1.02)	0.46	1.03 (1.01-1.06)	$1.7x10^{-2}$	1.00 (0.97-1.04)	0.81

[#] More common allele listed first, minor allele second; *Includes Breast Cancer Association Consortium (BCAC) OncoArray data from 9,655 ER-negative cases and 45,494 controls cases and controls not included in previously published studies; [†] Combined data from 21,468 ER-negative cases and 100,594 controls of European ancestry from BCAC, which includes overlapping samples with previous publications for all SNPs; [‡] Combined data from 18,908 *BRCA1* mutation carriers from the Consortium of Investigators of Modifiers of BRCA1/2 (CIMBA), 9,414 of whom had developed breast cancer - includes overlapping samples with previous publications for SNPs rs4577244, rs3757322, rs2747652 and rs6562760

Chr, chromosome; Ref, publication(s) in reference list in which the association was identified; MAF, minor allele frequency; OR, odds ratio per copy of the minor allele; CI, confidence interval; HR, hazard ratio per copy of the minor allele

Table 3: Associations for 10 novel and 10 previously reported (and replicated) ER-negative breast cancer susceptibility loci, by triple-negative status (BCAC data only: ER-negative cases‡, all controls))

Location	SNP	Triple-negative		Other ER-negative		Heterogeneity P-value*
		OR (95%CI)	P-value	OR (95%CI)	P-value	
Loci identified by the present study						
2p23.3	rs200648189	0.95 (0.90-1.00)	4.8x10^{-2}	0.96 (0.91-1.03)	0.24	0.36
6q23.1	rs6569648	0.93 (0.89-0.97)	1.4x10^{-3}	0.93 (0.88-0.98)	5.6x10^{-3}	0.91
8p23.3	rs66823261	1.11 (1.05-1.16)	3.3x10^{-5}	1.12 (1.07-1.19)	2.4x10^{-5}	0.91
8q24.13	rs17350191	1.07 (1.03-1.11)	7.9x10^{-4}	1.07 (1.02-1.12)	4.0x10^{-3}	0.67
11q22.3	rs11374964	0.88 (0.85-0.91)	1.9x10^{-11}	0.99 (0.95-1.04)	0.75	1.5x10^{-5}
11q22.3	rs74911261	0.76 (0.66-0.87)	1.1x10^{-4}	0.98 (0.84-1.13)	0.76	3.0x10^{-2}
16p13.3	rs11076805	0.91 (0.87-0.96)	1.5x10^{-4}	0.95 (0.90-1.00)	4.5x10^{-2}	0.20
18q12.1	rs36194942	0.93 (0.89-0.96)	2.4x10^{-4}	0.92 (0.88-0.97)	9.9x10^{-4}	0.94
19p13.2	rs322144	0.94 (0.91-0.98)	5.9x10^{-3}	0.94 (0.90-0.98)	9.7x10^{-3}	0.68
19q12	rs113701136	1.10 (1.06-1.15)	9.1x10^{-7}	1.07 (1.02-1.12)	4.4x10^{-3}	0.12
Previously reported loci (associations replicated by the present study)						
1q32.1	rs6678914	0.94 (0.91-0.98)	2.1x10^{-3}	0.91 (0.87-0.95)	2.0x10^{-5}	0.45
1q32.1	rs4245739	1.18 (1.13-1.23)	4.3x10^{-15}	1.04 (1.00-1.10)	7.5x10^{-2}	6.5x10^{-4}
2p24.1	rs12710696	1.07 (1.03-1.11)	1.1x10^{-3}	1.04 (1.00-1.09)	6.1x10^{-2}	0.52
2p23.2	rs4577244	0.90 (0.86-0.94)	5.3x10^{-6}	0.94 (0.89-0.99)	1.9x10^{-2}	0.15
5p15.33	rs10069690	1.28 (1.23-1.33)	2.4x10^{-33}	1.07 (1.02-1.12)	5.4x10^{-3}	5.6x10^{-8}
6q25.1	rs3757322	1.15(1.10-1.19)	4.3x10^{-12}	1.14(1.10-1.20)	4.8x10^{-9}	0.35
6q25.2	rs2747652	0.93(0.89-0.96)	5.7x10^{-5}	0.87(0.83-0.91)	2.9x10^{-10}	9.6x10^{-3}
13q22.1	rs6562760	0.94 (0.90-0.98)	2.8x10^{-3}	0.92 (0.87-0.96)	8.8x10^{-4}	0.46
16q12.2	rs11075995	1.06 (1.02-1.11)	6.5x10^{-3}	1.08 (1.03-1.13)	3.1x10^{-3}	0.81
19p13.11	rs67397200	1.27 (1.22-1.32)	2.0x10^{-32}	1.05 (1.01-1.10)	2.7x10^{-2}	4.7x10^{-10}

‡Combined Breast Cancer Association Consortium (BCAC) data from 6,877 triple-negative and 4,467 other ER-negative cases and 83,700 controls; *ER-negative case-only analysis, by triple-negative status; OR, odds ratio per copy of the minor allele; CI, confidence interval

Table 4: Associations for 10 novel and 10 previously reported (and replicated) ER-negative breast cancer susceptibility loci, by grade (BCAC data only: ER-negative cases‡, all controls)

Location	SNP	Grade 1		Grade 2		Grade 3		Heterogeneity P-value*
		OR (95%CI)	P-value	OR (95%CI)	P-value	OR (95%CI)	P-value	
Loci identified by the present study								
2p23.3	rs200648189	1.11 (0.92-1.33)	0.28	0.95 (0.88-1.03)	0.23	0.96 (0.91-1.00)	6.8x10^{-2}	0.70
6q23.1	rs6569648	0.93 (0.79-1.09)	0.37	0.93 (0.87-0.99)	1.6x10^{-2}	0.94 (0.91-0.98)	3.8x10^{-3}	0.34
8p23.3	rs66823261	1.13 (0.96-1.34)	0.14	1.12 (1.04-1.19)	1.2x10^{-3}	1.10 (1.05-1.15)	1.3x10^{-5}	0.11
8q24.13	rs17350191	1.16 (1.01-1.34)	3.0x10^{-2}	1.05 (0.99-1.11)	0.10	1.09 (1.05-1.12)	4.1x10^{-6}	0.94
11q22.3	rs11374964	0.91 (0.79-1.04)	0.16	0.99 (0.94-1.05)	0.85	0.93 (0.90-0.96)	1.3x10^{-5}	3.0x10^{-2}
11q22.3	rs74911261	1.22 (0.81-1.84)	0.35	0.89 (0.73-1.07)	0.21	0.74 (0.65-0.85)	7.4x10^{-6}	6.7x10^{-4}
16p13.3	rs11076805	0.90 (0.76-1.06)	0.21	0.93 (0.87-0.99)	3.2x10^{-2}	0.92 (0.88-0.95)	4.5x10^{-5}	0.71
18q12.1	rs36194942	0.97 (0.84-1.13)	0.73	0.93 (0.88-0.99)	2.2x10^{-2}	0.96 (0.92-0.99)	2.3x10^{-2}	0.98
19p13.2	rs322144	0.94 (0.81-1.08)	0.38	0.95 (0.90-1.01)	0.11	0.96 (0.93-1.00)	6.4x10^{-2}	0.48
19q12	rs113701136	1.02 (0.89-1.18)	0.77	1.06 (1.01-1.13)	3.0x10^{-2}	1.10 (1.06-1.14)	2.5x10^{-7}	0.12
Previously reported loci (associations replicated by the present study)								
1q32.1	rs6678914	0.95 (0.83-1.09)	0.46	0.90 (0.85-0.95)	9.3x10^{-5}	0.92 (0.89-0.95)	1.2x10^{-6}	0.75
1q32.1	rs4245739	1.02 (0.88-1.19)	0.75	1.05 (0.99-1.12)	8.7x10^{-2}	1.18 (1.14-1.22)	2.5x10^{-18}	4.3x10^{-5}
2p24.1	rs12710696	1.08 (0.94-1.23)	0.28	1.10 (1.04-1.16)	9.6x10^{-4}	1.04 (1.01-1.08)	1.6x10^{-2}	0.28
2p23.2	rs4577244	1.02 (0.88-1.20)	0.77	0.95 (0.89-1.01)	9.4x10^{-2}	0.90 (0.86-0.93)	1.2x10^{-7}	4.0x10^{-2}
5p15.33	rs10069690	0.96 (0.83-1.12)	0.64	1.07 (1.01-1.14)	2.2x10^{-2}	1.21 (1.17-1.26)	1.5x10^{-24}	7.3x10^{-4}
6q25.1	rs3757322	1.16 (1.01-1.34)	0.04	1.13 (1.07-1.20)	7.5x10^{-6}	1.18 (1.14-1.22)	4.5x10^{-20}	0.16
6q25.2	rs2747652	0.86 (0.75-0.98)	0.02	0.92 (0.87-0.97)	1.9x10^{-3}	0.90 (0.87-0.93)	1.6x10^{-9}	0.61
13q22.1	rs6562760	0.98 (0.84-1.15)	0.82	0.92 (0.87-0.98)	1.4x10^{-2}	0.91 (0.88-0.95)	1.2x10^{-5}	0.52
16q12.2	rs11075995	1.16 (1.00-1.35)	4.7x10^{-2}	1.09 (1.02-1.15)	7.5x10^{-3}	1.08 (1.04-1.13)	5.2x10^{28}	0.42
19p13.11	rs67397200	1.01 (0.87-1.16)	0.91	1.08 (1.02-1.14)	9.8x10^{-3}	1.22 (1.18-1.26)	5.3x10^{-37}	1.3x10^{-3}

‡Combined Breast Cancer Association Consortium (BCAC) data from 492 grade 1, 3,243 grade 2 and 8,568 grade 3 cases and 82,347 controls; * ER-negative case-only analysis of BCAC data, by grade (trend test, 1df); OR, odds ratio per copy of the minor allele; CI, confidence interval

Online Methods

Study subjects

Supplementary Table 1 summarises the studies from the Breast Cancer Association Consortium (BCAC) that contributed data. The majority were case-control studies. Sixty-eight BCAC studies participated in the ER-negative breast cancer component of the OncoArray, contributing 9,655 cases and 45,494 controls. All studies provided core data on disease status and age at diagnosis/observation, and the majority provided information on clinico-pathological and lifestyle factors, which have been curated and incorporated into the BCAC database (version 6). Estrogen receptor status for most (~70%) cases was obtained from clinical records. After removal of overlapping participants, genotype data were also available from eight GWASs[9,12,16,37,38] (4,480 ER-negative cases and 12,632 controls) and 40 studies previously genotyped using the Illumina iCOGS custom array[20] (7,333 ER-negative cases and 42,468 controls).

A total of 21,468 ER-negative cases were included in the combined analyses. Of those 5,793 had tumours that were also negative for progesterone receptor (PR) and human epidermal growth factor receptor 2 (HER2) and were defined as triple-negative (TN). PR and HER2 status was also obtained predominantly from clinical records. A further 4,217 were positive for PR or HER and were considered non-TN. The remainder had unknown PR or HER status. All participating studies were approved by their appropriate ethics review boards and all subjects provided informed consent.

Subjects included from the Consortium of Investigators of Modifiers of BRCA1/2 (CIMBA) were women of European ancestry aged 18 years or older with a pathogenic variant in *BRCA1*. The majority of the participants were sampled through cancer genetics clinics. Multiple members of the same families were included in some instances. Fifty-eight studies from 24 countries contributed Oncoarray genotype data. After quality control (see below) and removal of overlapping participants with the BCAC OncoArray study, data were available on 15,566 *BRCA1* mutation carriers, of whom 7,784 were affected with breast cancer (Supplementary Table 2). We also obtained iCOGS genotype data on 3,342 *BRCA1* mutation carriers (1,630 with breast cancer) from 54 studies through CIMBA. All mutation carriers provided written informed consent and participated under ethically approved protocols.

OncoArray SNP selection

Approximately 50% of the SNPs for the OncoArray were selected as a “GWAS backbone” (Illumina HumanCore), which aimed to provide high coverage for the majority of common variants through imputation. The remaining allocation was selected from lists supplied by each of six disease-based consortia, together with a seventh lists of SNPs of interest to multiple disease groups. Approximately 72k SNPs were selected specifically for their relevance to breast cancer, based on prior evidence of association with overall or subtype-specific disease, with breast density or with breast tissue specific gene expression. Lists were merged, as described previously[34].

Genotype calling and quality control

Details of the genotype calling and quality control (QC) for the iCOGS and GWAS are described elsewhere[19,20,23,30], and those for OncoArray are described in the Supplementary Note.

Imputation

Genotypes for ~21M SNPs were imputed for all samples using the October 2014 (Phase 3) release of the 1000 Genomes Project data as the reference panel and Nhap=800. The iCOGS, OncoArray and six of the GWAS datasets were imputed using a two-stage imputation approach, using SHAPEIT[73] for phasing and IMPUTEv2[74] for imputation. The imputation was performed in 5Mb non-overlapping intervals. All subjects were split into subsets of ~10,000 samples, with subjects from the same grouped in the subset. The Breast and Prostate Cancer Cohort Consortium (BPC3) and Breast Cancer Family Registry (BCFR) GWAS performed the imputation separately using MACH and Minimac[75,76]. We imputed genotypes for all SNPs that were polymorphic (MAF>0.1%) in either European or Asian samples. For the BCAC GWAS, data were included in the analysis for all SNPs with MAF>0.01 and imputation $r^2>0.3$. For iCOGS and OncoArray we included data for all SNPs with imputation $r^2>0.3$ and MAF>0.005.

Statistical analyses of BCAC data

Per-allele odds ratios and standard errors were generated for the Oncoarray, iCOGS and each GWAS, adjusting for principal components using logistic regression. The Oncorray and iCOGS analyses were additionally adjusted for country and study, respectively. For the OncoArray dataset, principal components analysis was performed using data for 33,661 SNPs (which included the 2,318 markers of continental ancestry) with a MAF≥0.05 and maximum correlation of 0.1, using purpose-written software (PCcalc; see URLs section) to allow standard calculations to be performed sufficiently rapidly on a very large dataset. We used the first 10 principal components, as additional components did not further reduce inflation in the test statistics. We used nine principal components for the iCOGS and up to 10 principal components for the other GWAS, where this was found to reduce inflation.

OR estimates were derived using MACH for the BCFR GWAS, ProbABEL[77] for the BPC3 GWAS, SNPTEST (see URLs section) for the remaining GWAS and purpose written software for the iCOGS and Oncoarray datasets. OR estimates and standard errors were combined by a fixed effects inverse variance meta-analysis using METAL[39]. This was first done across the eight GWAS, applying genomic control, as described previously[20]. It was then applied (without genomic control) to combine findings from the three BCAC genotyping initiatives (GWAS, iCOGS, OncoArray).

The independence of signals from two variants at 11q22.3 was by fitting the logistic regression models described above with both variants as covariates. This was done separately for iCOGS and OncoArray data and results for each variant combined by meta-analysis.

For selected SNPs we estimated per-allele ORs by ER-status using all available BCAC data for 82,263 cases with known ER status and 87,962 controls from the iCOGS and OncoArray studies. We also estimated the per-allele ORs by TN status (TN versus other ER-negative subtypes) and tumour grade, using available BCAC data for ER-negative cases and corresponding controls. Tests for heterogeneity by subtype were derived by applying logistic regression to cases only. This was done separately for the iCOGS and Oncoarray datasets, adjusted as before, and then combined in a fixed-effects meta-analysis. Multinomial regression was applied to cases only to test a linear trend for grade, with the model constrained so that the difference between grade 1 and 3 was double that for the difference between grade 2 and 3; this method was also used to test for a linear trend with age with ordinal values 1, 2, 3 and 4 representing ages <40, 40-49, 50-59 and ≥60, respectively.

Statistical analyses of CIMBA data

Associations between genotypes and breast cancer risk for *BRCA1* mutation carriers were evaluated using a 1*df* per allele trend-test (*P*-trend), based on modeling the retrospective likelihood of the observed genotypes conditional on breast cancer phenotypes[36]. This was done separately for iCOGS and OncoArray data. To allow for the non-independence among related individuals, an adjusted test statistic was used which took into account the correlation in genotypes[3]. All analyses were stratified by country of residence and, for countries where strata were sufficiently large (USA and Canada), by Ashkenazi Jewish ancestry. The results from the iCOGS and OncoArray datasets were then pooled using fixed effects meta-analysis. We repeated these analyses modelling ovarian cancer as a competing risk and observed no substantial difference in the results obtained.

The independence of signals from two variants at 11q22.3 was assessed using OncoArray data only, fitting a Cox regression model with per-allele effects for both variants, adjusting for birth cohort, stratified by country of residence and using robust standard errors and clustered observations for relatives. This approach provides valid significance tests of associations, although the HR estimates can be biased[35].

Meta-analysis of BCAC and CIMBA

A fixed effects meta-analysis of results from BCAC and CIMBA was conducted using an inverse variance approach assuming fixed effects, as implemented in METAL[39]. The effect estimates used were the logarithm of the per-allele hazard ratio (HR) estimate for the association with breast cancer risk in *BRCA1* mutation carriers from CIMBA and the logarithm of the per-allele OR estimate for the association with risk of ER-negative breast cancer based on BCAC data, both of which were assumed to approximate the same relative risk. We assessed genomic inflation using common (MAF>1%) GWAS backbone variants. As lambda is influenced by sample size, we calculated lambda1000 to be comparable with other studies.

All statistical tests conducted were two-sided.

Definition of known hits

We identified all associations previously reported from genome-wide or candidate analysis at a significance level $P<5\times10^{-8}$ for overall breast cancer, ER-negative or ER-positive breast cancer, in *BRCA1* or *BRCA2* carriers, or in meta-analyses of these categories. We included only one SNP in any 500kb interval, unless joint analysis provided genome-wide significant evidence (conditional $P<5\times10^{-8}$) of more than one independent signal. Where multiple studies reported associations in the same region, we considered the first reported association unless a later study identified a different variant in the same region that was more strongly associated with breast cancer risk. One hundred and seven previously reported hits were identified, 11 of these through GWAS of ER-negative disease or of breast cancer in *BRCA1* mutation carriers, or reported as more strongly associated with ER-negative breast cancer. These are listed in Table 2. The other 96 previously reported hits are listed in Supplementary Table 10.

Definition of new hits

To search for novel loci, we assessed all SNPs excluding those within 500kb of a known hit. This identified 206 SNPs in nine regions that were associated with disease risk at $P<5\times10^{-8}$ in the meta-analysis of BCAC ER-negative breast cancer and CIMBA *BRCA1* mutation carriers. The SNP with lowest p-value from this analysis was considered the lead SNP. No additional loci were detected from the analysis of BCAC data only. Imputation quality, as assessed by the IMPUTE2 imputation r^2 in the Oncoarray dataset, was ≥0.89 for the 10 lead SNPs reported (Supplementary Table 3).

Candidate causal SNPs

To define the set of potentially causal variants at each of the novel susceptibility loci, we selected all variants with p-values within two orders of magnitude of the most significant SNP at each of the 10 novel loci. This is approximately equivalent to selecting variants whose posterior probability of causality is within two orders of magnitude of the most significant SNP[40,41]. This approach was applied to identify potentially causal variants for the signal given by the more frequent lead SNP at 11q22.3 (rs11374964). A similar approach was applied for the rarer lead SNP at this locus (rs74911261), but based on p-values from analyses adjusted for rs11374964.

Proportion of familial risk explained

The relative risk of ER-negative breast cancer for the first degree female relative of a woman with ER-negative disease has not been estimated. We therefore assumed that the 2-fold risk observed for overall disease also applied to ER-negative disease. In order to estimate the proportion of this explained by the 125 variants associated with ER-negative disease, we used minor allele frequency and OR estimates from the OncoArray-based genotype data and applied the formula: $\sum_i p_i(1-p_i)(\beta_i^2-\tau_i^2)/\ln(\lambda))$, where p_i is the minor allele frequency for variant *i*, β_i is the log(OR) estimate for variant *i*, τ_i is the standard error of β_i and $\lambda=2$ is the assumed overall familial relative risk.

The corresponding estimate for the FRR due to all variants is the *frailty scale* heritability, defined as $h_f^2=\sum_i 2p_i(1-p_i)\gamma_i^2$, where the sum over all variants and γ_i is the true relative risk conferred by variant *i*, assuming a log-additive model. We first obtained the estimated heritability based on the full set of summary estimates using LD Score Regression[68], which derives a heritability estimate on the observed scale. We then converted this to an estimate on the frailiity scale using the formula $h_f^2 = {h_{obs}^2}/{P(1-P)}$, where *P* is the proportion of samples in the population that are cases.

Proportion of polygenic risk-modifying variance explained for *BRCA1* carriers.

The proportion of the variance in the polygenic frailty modifying risk in BRCA1 carriers explained by the set of associated SNPs was estimated by $\sum_i \ln c_i/\sigma^2$, where

c_i is the squared estimated coefficient of variation in incidences associated with SNP_i[78] and σ^2 is the total polygenic variance, estimated from segregation data[79].

In Silico Annotation of Candidate Causal variants

We combined multiple sources of *in silico* functional annotation from public databases to help identify potential functional SNPs and target genes, based on previous observations that breast cancer susceptibility alleles are enriched in *cis*-regulatory elements and alter transcriptional activity[28,80-82]. The influence of candidate causal variants on transcription factor binding sites was determined using the ENCODE-Motifs resource[43]. To investigate functional elements enriched across the region encompassing the strongest candidate causal SNPs, we analysed chromatin biofeatures data from the Encyclopedia of DNA Elements (ENCODE) Project[42], Roadmap Epigenomics Projects[44] and other data obtained through the National Center for Biotechnology Information (NCBI) Gene Expression Omnibus (GEO) namely: Chromatin State Segmentation by Hidden Markov Models (chromHMM), DNase I hypersensitive and histone modifications of epigenetic markers H3K4, H3K9, and H3K27 in Human Mammary Epithelial (HMEC) and myoepithelial (MYO) cells, T47D and MCF7 breast cancer cells and transcription factor ChIP-seq in a range of breast cell lines (Supplementary Table 6). To identify the SNPs most likely to be functional we used RegulomeDB[45], and to identify putative target genes, we examined potential functional chromatin interactions between distal and proximal regulatory transcription-factor binding sites and the promoters at the risk regions, using Hi-C data generated in HMECs[47] and Chromatin Interaction Analysis by Paired End Tag (ChiA-PET) in MCF7 cells. This detects genome-wide interactions brought about by, or associated with, CCCTC-binding factor (CTCF), DNA polymerase II (POL2), and Estrogen Receptor (ER), all involved in transcriptional regulation[47]. Annotation of putative *cis*-regulatory regions and predicted target genes used the Integrated Method for Predicting Enhancer Targets (IM-PET)[46], the “Predicting Specific Tissue Interactions of Genes and Enhancers” (PreSTIGE) algorithm[48], Hnisz[51] and FANTOM[49]. Intersections between candidate causal variants and regulatory elements were identified using Galaxy, BedTools v2.24 and HaploReg v4.1, and visualised in the UCSC Genome Browser. Publically available eQTL databases including Gene-Tissue Expression (GTEx;[50] version 6, multiple tissues) and Westra[52] (blood), were queried for candidate causal variants.

eQTL analyses

Expression quantitative trait loci (eQTL) analyses were performed using data from The Cancer Genome Atlas (TCGA) and Molecular Taxonomy of Breast Cancer International Consortium (METABRIC) projects[59,60].

The TCGA eQTL analysis was based on 79 ER-negative breast tumors that had matched gene expression, copy number, and methylation profiles together with the corresponding germline genotypes available. All 79 individuals were of European ancestry as ascertained using the genotype data and the Local Ancestry in adMixed Populations (LAMP) software package (LAMP estimate cut-off >95% European)[83]. Germline genotypes were imputed into the 1000 Genomes reference panel (October 2014 release) using IMPUTE2[75,84]. Gene expression had been measured on the Illumina HiSeq 2000 RNA-Seq platform (gene-level RSEM normalized counts[85]), copy number estimates were derived from the Affymetrix SNP 6.0 (somatic copy number alteration minus germline copy number variation called using the GISTIC2 algorithm[86]), and methylation beta values measured on the Illumina Infinium HumanMethylation450, as previously described[59]. Primary TCGA eQTL analysis focused on all potentially causal variants in the 10 new regions associated with breast cancer risk in the meta-analysis of ER-negative cases and controls from BCAC and *BRCA1* mutation carriers from CIMBA. We considered all genes located up to 1 Mb on either side of each of these variants. The effects of tumor copy number and methylation on gene expression were first removed using a method described previously[58], and eQTL analysis was performed by linear regression as implemented in the R package Matrix eQTL[87].

The METABRIC eQTL analysis was based on 135 normal breast tissue samples resected from breast cancer patients of European ancestry. Germline genotyping for the METABRIC study was also done on the Affymetrix SNP 6.0, and ancestry estimation and imputation for this data set was conducted as described for TCGA. Gene expression in the METABRIC study had been measured using the Illumina HT12 microarray platform and we used probe-level estimates. As for TCGA, we considered all genes in 10 regions using Matrix eQTL.

We also performed additional eQTL analyses using the METABRIC data set for all variants within 1 Mb of *L3MBTL3* and *CDH2* and the expression of these specific genes.

Global Genomic Enrichment Analyses

We performed stratified LD score regression analyses[68] for ER- breast cancer using the summary statistics based on the meta-analyses of OncoArray, GWAS, iCOGS and CIMBA. We used all SNPs in the 1000 Genomes Project phase 1 v3 release that had a minor allele frequency > 1% and an imputation quality score $R^2>0.3$ in the OncoArray data. LD scores were calculated using the 1000 Genomes Project Phase 1 v3 EUR panel. Further details are provided in the Supplementary Note.

We tested the differences in functional enrichment between ER-positive and ER-negative subsets for individual features through a Wald test, using the regression coefficients and standard errors for the two subsets based on the models described above. Finally, we assessed the heritability due to genotyped and imputed SNPs[70] and estimated the genetic correlation between ER-positive and ER-negative breast cancer[69]. The genetic correlation analysis was restricted to the ~1M SNPs included in HapMap 3.

Pathway Enrichment Analyses (see also the Supplementary Note)

The pathway gene set database Human_GOBP_AllPathways_no_GO_iea_January_19_2016_symbol.gmt (GeneSets; see URLs section)[61], was used for all analyses. Pathway size was determined by the total number of genes in the pathway to which SNPs in the imputed GWAS dataset could be mapped. To provide more biologically meaningful results, and reduce false positives, only pathways that contained between 10 and 200 genes were considered.

SNPs were mapped to the nearest gene within 500kb; those that were further than 500 kb away from any gene were excluded. Gene significance was calculated by assigning the lowest p-value observed across all SNPs assigned to a gene[63,64], based on the meta-analysis of BCAC and CIMBA data described above.

The gene set enrichment analysis (GSEA)[61] algorithm, as implemented in the GenGen package (see URLs section)[62,63] was used to perform pathway analysis. Briefly, the algorithm calculates an enrichment score (ES) for each pathway based on a weighted Kolmogorov-Smirnov statistic[62]. Pathways that have most of their genes at the top of the ranked list of genes obtain higher ES values.

We defined an ES threshold (ES≥0.41) to yield a true-positive rate (TPR) of 0.20 and a false-positive rate (FPR) of 0.14, with true-positive pathways defined as those observed with false discovery rate (FDR)<0.05 in a prior analysis carried out using the analytic approach defined above applied to iCOGS data for ER-negative disease.

To visualize the pathway enrichment analysis results, an enrichment map was created using the Enrichment Map (EM) v 2.1.0 app[61] in Cytoscape v3.30[88], applying an edge-weighted force directed layout. To measure the contribution of each gene to enriched pathways and annotate the map, we reran the pathway enrichment analysis multiple times, each time excluding one gene. A gene was considered to drive the enrichment if the ES dropped to zero or less (pathway enrichment driver) after it was excluded. Pathways were grouped in the map if they shared >70% of their genes or their enrichment was driven by a shared gene.

Online-only References

73. Delaneau, O., Marchini, J. & Zagury, J.F. A linear complexity phasing method for thousands of genomes. *Nat Methods* **9**, 179-81 (2012).
74. Howie, B.N., Donnelly, P. & Marchini, J. A flexible and accurate genotype imputation method for the next generation of genome-wide association studies. *PLoS Genet* **5**, e1000529 (2009).
75. Howie, B., Fuchsberger, C., Stephens, M., Marchini, J. & Abecasis, G.R. Fast and accurate genotype imputation in genome-wide association studies through pre-phasing. *Nat Genet* **44**, 955-9 (2012).
76. Li, Y., Willer, C.J., Ding, J., Scheet, P. & Abecasis, G.R. MaCH: using sequence and genotype data to estimate haplotypes and unobserved genotypes. *Genet Epidemiol* **34**, 816-34 (2010).
77. Aulchenko, Y.S., Struchalin, M.V. & van Duijn, C.M. ProbABEL package for genome-wide association analysis of imputed data. *BMC Bioinformatics* **11**, 134 (2010).
78. Antoniou, A.C. & Easton, D.F. Polygenic inheritance of breast cancer: Implications for design of association studies. *Genet Epidemiol* **25**, 190-202 (2003).
79. Antoniou, A.C. *et al.* The BOADICEA model of genetic susceptibility to breast and ovarian cancers: updates and extensions. *Br J Cancer* **98**, 1457-66 (2008).
80. Darabi, H. *et al.* Polymorphisms in a Putative Enhancer at the 10q21.2 Breast Cancer Risk Locus Regulate NRBF2 Expression. *Am J Hum Genet* **97**, 22-34 (2015).
81. Glubb, D.M. *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* **96**, 5-20 (2015).

82. Ghoussaini, M. *et al.* Evidence that breast cancer risk at the 2q35 locus is mediated through IGFBP5 regulation. *Nat Commun* **4**, 4999 (2014).
83. Baran, Y. *et al.* Fast and accurate inference of local ancestry in Latino populations. *Bioinformatics* **28**, 1359-67 (2012).
84. Abecasis, G.R. *et al.* An integrated map of genetic variation from 1,092 human genomes. *Nature* **491**, 56-65 (2012).
85. Li, B. & Dewey, C.N. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinformatics* **12**, 323 (2011).
86. Mermel, C.H. *et al.* GISTIC2.0 facilitates sensitive and confident localization of the targets of focal somatic copy-number alteration in human cancers. *Genome Biol* **12**, R41 (2011).
87. Shabalin, A.A. Matrix eQTL: ultra fast eQTL analysis via large matrix operations. *Bioinformatics* **28**, 1353-8 (2012).
88. Shannon, P. *et al.* Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res* **13**, 2498-504 (2003).