



Universiteit
Leiden
The Netherlands

Search for new breast cancer susceptibility genes

Oldenburg, R.A.

Citation

Oldenburg, R. A. (2008, May 29). *Search for new breast cancer susceptibility genes*. Retrieved from <https://hdl.handle.net/1887/12871>

Version: Corrected Publisher's Version

License: [Licence agreement concerning inclusion of doctoral thesis in the Institutional Repository of the University of Leiden](#)

Downloaded from: <https://hdl.handle.net/1887/12871>

Note: To cite this publication please use the final published version (if applicable).

LINKAGE SEARCHES**5.1. A GENOME WIDE LINKAGE SEARCH FOR BREAST CANCER SUSCEPTIBILITY GENES**

Cambridge Group, Paula Smith,¹ Lesley McGuffog,¹ Douglas F. Easton,^{1*} Australia Group, Graham J. Mann,² Gulietta M. Pupo,² Beth Newman,³ Georgia Chenevix-Trench,⁴ kConFab Investigators, IARC Group, Csilla Szabo,^{5,6} Melissa Southey,⁵ H el ene Renard,⁵ Fabrice Odefrey,⁵ Henry Lynch,⁷ Dominique Stoppa-Lyonnet,⁸ Fergus Couch,⁶ John L. Hopper,⁹ Graham G. Giles,¹⁰ Margaret R. E. McCredie,¹¹ Sandra Buys,¹² Irene Andrulis,¹³ Ruby Senie,¹⁴ BCFS, BRCAX Collaborators Group, David E. Goldgar,⁵ Netherlands Group, Rogier Oldenburg,¹⁵ Karin Kroeze-Jansema,¹⁶ Jaennelle Kraan,¹⁶ Hanne Meijers-Heijboer,¹⁵ Jan G. M. Klijn,¹⁵ Christi van Asperen,¹⁶ Inge van Leeuwen,¹⁶ Hans F. A. Vasen,¹⁶ Cees J. Cornelisse,¹⁶ Peter Devilee,¹⁶ United Kingdom Group, Linda Baskcomb,¹⁷ Sheila Seal,¹⁷ Rita Barfoot,¹⁷ Jon Mangion,¹⁷ Anita Hall,¹⁷ Sarah Edkins,¹⁷ Elizabeth Rapley,¹⁷ Richard Wooster,¹⁷ Jenny Chang-Claude,¹⁸ Diana Eccles,¹⁹ D. Gareth Evans,²⁰ P. Andrew Futreal,²¹ Katherine L. Nathanson,²² Barbara L. Weber,²² the Breast Cancer Susceptibility Collaboration (UK), Nazneen Rahman,¹⁷ Michael R. Stratton^{17,21}

Genes Chromosomes Cancer. 2006 Jul;45(7):646-55.

¹ CR-UK Genetic Epidemiology Unit, Department of Public Health and Primary Care, University of Cambridge, Cambridge, UK

² Westmead Institute for Cancer Research, University of Sydney, Sydney, Australia

³ Queensland University of Technology, Brisbane, Australia

⁴ Queensland Institute for Medical Research, Brisbane, Australia

⁵ Unit of Genetic Epidemiology, International Agency for Research, Lyon, France

⁶ Department of Laboratory Medicine and Pathology, Mayo Clinic, Rochester, MN

⁷ Department of Preventive Medicine, Creighton University, Omaha, NE

⁸ Unit e de G en etique Oncologique, Institut Curie and Unit e INSERM U.434, Paris, France

⁹ Centre for Genetic Epidemiology, University of Melbourne, Melbourne, Australia

¹⁰ Cancer Epidemiology Centre, Cancer Council of Victoria, Melbourne, Australia

- ¹¹ *Department of Preventative and Social Medicine, University of Otago, Dunedin, New Zealand*
- ¹² *Department of Hematology-Oncology, University of Utah, Salt Lake City, UT*
- ¹³ *Department of Laboratory Medicine, Mount Sinai Hospital, University of Toronto, Toronto, Canada*
- ¹⁴ *Department of Epidemiology, Mailman School of Public Health, Columbia University, New York, NY*
- ¹⁵ *Erasmus University Medical Centre, Rotterdam, The Netherlands*
- ¹⁶ *Leiden University Medical Centre, Leiden, The Netherlands*
- ¹⁷ *Section of Cancer Genetics, Institute of Cancer Research, Sutton, Surrey, UK*
- ¹⁸ *Division of Epidemiology, German Cancer Research Center DFKZ, Heidelberg, Germany*
- ¹⁹ *Wessex Clinical Genetics Service, Princess Ann Hospital, Southampton, UK*
- ²⁰ *Regional Genetic Service, St. Mary's Hospital, Manchester, UK*
- ²¹ *The Wellcome Trust Sanger Institute, Wellcome Trust Genome Campus, Hinxton, UK*
- ²² *University of Pennsylvania Cancer Centre, Philadelphia, PA*

ABSTRACT

Mutations in known breast cancer susceptibility genes account for a minority of the familial aggregation of the disease. To search for further breast cancer susceptibility genes, we performed a combined analysis of four genome-wide linkage screens, which included a total of 149 multiple case breast cancer families. All families included at least three cases of breast cancer diagnosed below age 60 years, at least one of whom had been tested and found not to carry a *BRCA1* or *BRCA2* mutation. Evidence for linkage was assessed using parametric linkage analysis, assuming both a dominant and a recessive mode of inheritance, and using nonparametric methods. The highest LOD score obtained in any analysis of the combined data was 1.80 under the dominant model, in a region on chromosome 4 close to marker D4S392. Three further LOD scores over 1 were identified in the parametric analyses and two in the nonparametric analyses. A maximum LOD score of 2.40 was found on chromosome arm 2p in families with four or more cases of breast cancer diagnosed below age 50 years. The number of linkage peaks did not differ from the number expected by chance. These results suggest regions that may harbor novel breast cancer susceptibility genes. They also indicate that no single gene is likely to account for a large fraction of the familial aggregation of breast cancer that is not due to mutations in *BRCA1* or *BRCA2*.

INTRODUCTION

Breast cancer aggregates in families, with the disease being approximately twice as common in the first-degree relatives of cases as in the general population (Collaborative Group on Hormonal Factors in Breast Cancer, [2001]). The higher risk to monozygotic twins of breast cancer cases than to dizygotic twins of cases suggests that most of this familial clustering is likely to have a genetic basis (Peto and Mack, [2000]). However, although several important breast cancer susceptibility genes have now been identified, most of the familial aggregation of breast cancer remains unexplained.

In the 1990s, two important breast cancer susceptibility genes, *BRCA1* (MIM 113705) and *BRCA2* (MIM 600185), were identified by linkage studies in multiple case families (Miki et al. [1994]; Wooster et al. [1995]). Germline mutations in these genes confer high lifetime risks of breast cancer and ovarian cancer, together with smaller risks of some other cancer types (Antoniou et al. [2003]; Thompson and Easton, [2004]). Mutations in these genes are common in families with multiple cases of breast or ovarian cancer, and are present in most families with at least six or more cases (Ford et al. [1998]). Population-based studies have estimated that *BRCA1* and *BRCA2* mutations account for 15% of the excess familial risk of breast cancer (Peto et al. [1999]; Anglian Breast Study, [2000]; Dite et al. [2003]). Mutations in two other genes, *TP53* and *PTEN*, also confer high risks of breast cancer, but only in the context of rare syndromes. Mutations in the *ATM* and *CHEK2* genes confer more moderate (approximately twofold) risks of breast cancer (*CHEK2* Case-Control Consortium, [2004]; Thompson et al. [2005]), although some mutations in *ATM* may confer higher risks. In total, the known susceptibility genes have been estimated to account for no more than 25% of the familial aggregation of breast cancer (Easton, [1999]), suggesting strongly that other susceptibility genes remain to be identified.

BRCA1 or *BRCA2* mutations are found in the majority of families with six or more cases of breast cancer cases consistent with dominant inheritance (Ford et al. [1998]). This suggests strongly that further susceptibility genes are likely to confer smaller risks than *BRCA1* and *BRCA2* mutations, but the number and characteristics of such genes remains unknown. One model, suggested by a recent segregation analysis (Antoniou et al. [2004]), proposes that there are a large number of such genes, each conferring only small risks of the disease. If true, such loci could not be identified through linkage studies. However, it is also possible that there are further loci conferring more substantial risks that could be detected by linkage. To evaluate this

possibility, we have conducted a genome-wide linkage analysis in multiple case breast cancer families that are unlikely to be segregating *BRCA1* or *BRCA2* mutations.

MATERIALS AND METHODS

As a basis for this linkage study, we sought to identify informative families with a low probability that they contained mutations in *BRCA1* or *BRCA2*. Families were collected independently by four groups, principally through family cancer clinics or epidemiological studies of breast cancer. All families were of Caucasian ancestry. The recruitment of the families used in the study took place over the last 15 years, but all families were regularly updated with regard to their cancer status. All groups obtained appropriate Institutional Review Board approvals. Specific sources of recruitment were as follows:

Australia: Families were identified through the Kathleen Cuninghame Foundation Consortium for Research into Familial Breast Cancer (kConFab), which is a national multidisciplinary consortium for research on familial breast cancer (GJ Mann, unpublished). Several families were initially ascertained through the Australian Breast Cancer Family Registry (ABCFS); these kindreds were recruited as part of a population-based case-control-family study and all were recruited via a diagnosis of breast cancer in the proband under the age of 40 years (Hopper et al. [1999]).

IARC: Families were ascertained by a collaborative group of investigators from the USA, Canada, Australia, and France. Netherlands: The Dutch families were ascertained through the Clinical Genetic Centers in Leiden and Rotterdam, and through the Netherlands Foundation for the Detection of Hereditary Tumors (STOET).

United Kingdom: All but 17 of the families were ascertained through clinical genetics centers in the United Kingdom. Two families were initially ascertained in the Netherlands, six from centers in the USA, and nine from Heidelberg, Germany.

Initially, all families had to satisfy the following criteria: (1) at least three women diagnosed with breast cancer below age 60 years, all of whom were related such that they could share a single allele identically by descent, (2) no case of ovarian cancer or male breast cancer in a blood relative (since these phenotypes are strongly predictive of the presence of *BRCA1* or *BRCA2* mutation), and (3) DNA samples available for genotyping from at least three women affected with breast cancer, or from children of affected women such that the genotypes of at least three affected women might be inferred (in the latter case, at least two children of an affected women needed to be

TABLE 1**Summary of families by group**

Group	Number of families							Number of genotyped individuals
	Total	Number of cases of breast cancer				Cases of breast cancer diagnosed below age 50 years		
		3	4	5	6+	<4	4+	
Australia	21	4	6	6	5	17	4	127
IARC	26	7	5	8	6	23	3	122
Netherlands	22	3	6	5	8	15	7	79
U.K.	80	25	24	18	13	70	10	395
Total	149	39	41	37	32	125	24	723

available). In addition, to minimize the probability that the family segregated a *BRCA1* or *BRCA2* mutation, DNA from at least one affected individual was screened for mutations across both genes, by a method that examined the entire coding sequence and splice junctions. Whenever possible, for families with five or more cases of breast cancer, a second affected individual was screened. Subsequently, we collected detailed information on the method of mutation screening for each family, as well as genotype data on at least three microsatellite markers flanking the *BRCA1* and *BRCA2* loci. Families with insufficient mutation screening (14 families) or linkage data (a further 6 families) were not included in further analyses. Finally, we estimated the residual probability that the index-affected individual carried a *BRCA1* and or *BRCA2* mutation, based on the assumed mutation detection sensitivity, the family history and linkage data at *BRCA1* and *BRCA2* (see statistical methods). Thirteen families, where this probability exceeded 15%, were excluded from all analyses presented here. Characteristics of the 149 families included in the analysis are summarized in Table 1.

Genotyping

To evaluate linkage to *BRCA1* and *BRCA2*, the following markers were used in various combinations in the four family sets: D17S800, D17S855, D17S951, D17S1322, D17S250 (for *BRCA1*); D13S260, D13S171, D13S1700, D13S267 (for *BRCA2*). At least three markers were analyzed at each locus in each family.

The entire coding sequences of *BRCA1* and *BRCA2* in each family were screened for mutations using several methods at the different centers. These include conformation sensitive gel electrophoresis, single strand conformational analysis, protein truncation test, DNA sequencing, and denaturing gradient gel electrophoresis. All of the Netherlands and United Kingdom and three of the IARC families were additionally screened for large deletions and insertions using deletion junction-PCR, multiplex ligation probe amplification (MLPA), or Southern analysis.

For the genome-wide linkage search, the Applied Biosystems Linkage Mapping Set MD10 was analyzed on ABI 3700 DNA sequencers, either on contract at the Australian Genome Research Facility (Australian families) or at the Wellcome Trust Sanger Institute (IARC), Netherlands and United Kingdom families). Genotypes were called automatically using Genotyper or Genemapper software and were then checked manually by at least one individual. Additional markers were used to investigate potential regions of interest in subgroups of the family set.

Statistical Analysis

To compute the residual probability that the index case carried a *BRCA1* and *BRCA2* mutation, we first used the Breast and Ovarian Analysis of Disease Incidence and Carrier Estimation Algorithm (BOADICEA) model (Antoniou et al. [2004]) to calculate carrier probabilities based on the pedigree and the mutation testing that had been performed. This model allows for the effects of *BRCA1* and *BRCA2* and the combined effects of other genes in a polygenic component, and is implemented in MENDEL (Lange et al. [1988]). For this purpose, the sensitivity of mutation screening was assumed to be 70% for *BRCA1* and 80% for *BRCA2* (mutation sensitivities estimated from linked families in the Breast Cancer Linkage Consortium dataset; D.Easton, unpublished data). For samples that had been fully screened for large-scale rearrangements by MLPA, the *BRCA1* sensitivity was assumed to be 80%. The carrier probabilities were then adjusted to allow for linkage data at the *BRCA1* and *BRCA2* loci. Multipoint LOD scores were computed using Fastlink (Cottingham et al. [1993]), based on at least three markers tightly linked to each locus. The residual *BRCA1* carrier probability was then given by:

$$\frac{p_1 10^{\text{LOD}_1}}{(p_1 10^{\text{LOD}_1} + p_2 10^{\text{LOD}_2} + 1 - p_1 - p_2)}$$

and similarly for *BRCA2*, where p_1 and p_2 are the *BRCA1* and *BRCA2* probabilities generated by BOADICEA and LOD1 and LOD2 are the multipoint LOD scores at the *BRCA1* and *BRCA2* loci.

To conduct a combined linkage analysis, we first constructed a single linkage map incorporating all markers typed at any center. This map was based on the sex-averaged linkage map generated by deCODE (Kong et al. [2002]). For markers that were not present on the deCODE map, we interpolated their position between flanking markers, either using estimates from other linkage maps or based on their physical position in the human genome sequence relative to flanking markers. Allele frequencies for each marker were estimated by averaging over all typed individuals, separately for each center.

Evidence for linkage was assessed using both parametric and nonparametric (allele sharing approach) analyses. For the parametric analysis, we first assumed a model in which susceptibility to breast cancer is conferred by a dominant susceptibility allele with population frequency 0.003 that confers a cumulative breast cancer risk of 80% by age 80, when compared with 8% in noncarriers. This model is based on that derived from the segregation analysis of (Claus et al. [1981]) and has been used in most previous breast cancer linkage analyses. As in previous analyses, risks were modeled in seven age-categories (<30, 30-39, 40-49, 50-59, 60-69, 70-79, 80+) and implemented by using 14 liability classes, with separate classes for affected and unaffected individuals (Easton et al. [1993]). Since this model would have reduced power to detect a recessive susceptibility allele, we also analyzed the data under a recessive model. Under this model, the risks to carriers and noncarriers were identical to those under the dominant model, but the allele frequency was assumed to be 0.08. All analyses were carried out in the program GENEHUNTER (Kruglyak et al. [1996]), except for one large family (EUR60) that could not be run because the number of individuals exceeded the limits of the program, and where pruning the family would have lost a significant amount of information. For this family, analyses were run in VITESSE (O'Connell and Weeks, [1995]) for autosomes and FASTLINK (Cottingham et al. [1993]) for the X chromosome. This family separates into two distantly related branches, and these were treated as two distinct families (EUR60a and EUR60b) in the analysis.

For the GENEHUNTER analyses, multipoint LOD scores were calculated for locations at 1 cM intervals along each chromosome, using all markers for that chromosome. For the VITESSE and FASTLINK analyses, multipoint LOD scores based on every pair of adjacent markers and the disease locus were calculated. The LOD score for each

TABLE 2**Maximum LOD scores by chromosome under the dominant, recessive, and nonparametric analyses**

Chromosome	Position	Model	HLOD or NPL LOD score*	
All families				
2	17	Dominant	1.21	0.18
2	16	NPL	1.10	
4	79	Dominant	1.80	0.18
5	196	Recessive	1.04	0.41
14	44	NPL	1.56	
22	41	Dominant	1.15	0.06
4+ cases dx < 50				
2	17	Dominant	2.38	0.50
4	66	Dominant	1.57	0.28
10	89	Dominant	1.12	0.35
22	41	Dominant	1.43	0.12

* HLOD, heterogeneity LOD score under dominant or recessive model;
 NPL, nonparametric LOD score (see Materials and Methods).

family at each position was based on an average of the LOD scores from all analyses relevant to that position. The multipoint LOD scores for each family at each position were then used to generate heterogeneity LOD scores (HLODs) based on the standard admixture model under which a certain proportion of families are assumed to be segregating a susceptibility allele at that locus (Ott, [1983]).

The nonparametric (allele sharing) analyses were conducted using the program GENEHUNTER-plus (Kong and Cox,[1997]), using the all scoring function (Whittemore and Halpern,[1994]). Analyses were conducted separately for each of the four centers and the results files combined. Nonparametric LOD scores were then generated using the program ASM, using the exponential scoring option and equal weighting of families.

RESULTS

Table 2 summarizes all linkage peaks with LOD scores greater than 1 in the combined dataset, for the whole family set and for analyses restricted to families with four or

TABLE 3**Maximum HLOD scores by chromosome and group, under the dominant model**

Chrom	Australia	IARC	Netherlands	UK	TOTAL
1	0.78 (134)	0.54 (7)	0.12 (108)	0.41 (20)	0.22 (9)
2	1.07 (17)	0.57 (7)	0.28 (253)	0.68 (27)	1.21 (17)
3	0.62 (38)	0.72 (101)	0.08 (30)	0.68 (107)	0.93 (102)
4	1.02 (35)	2.02 (73)	0.08 (189)	0.92 (79)	1.80 (79)
5	0.91 (170)	0.43 (42)	0.15 (5)	0.15 (104)	0.13 (170)
6	0.65 (89)	0.72 (171)	0	0.04 (189)	0.04 (94)
7	0.37 (103)	0.19 (46)	0.03 (57)	0.86 (170)	0.20 (171)
8	0.62 (39)	0.12 (50)	0	0.35 (13)	0.07 (42)
9	0.48 (121)	0	0.70 (81)	0.07 (64)	0.01 (113)
10	0.53 (89)	0	0.16 (138)	0.04 (20)	0.12 (102)
11	0.50 (0)	0.89 (26)	0.47 (95)	0.01 (22)	0.42 (0)
12	0.57 (161)	0.15 (119)	0.43 (150)	1.04 (3)	0.43 (150)
13	0.86 (67)	0.15 (99)	0	0.19 (108)	0.0 (99)
14	0.44 (44)	0.93 (37)	0.20 (44)	0.49 (116)	0.79 (44)
15	0.0	0.11 (26)	0	1.30 (9)	0.41 (21)
16	0.13 (0)	0.14 (14)	0	0.27 (25)	0.0 (3)
17	1.17 (93)	0.09 (103)	0.12 (41)	0.64 (115)	0.60 (103)
18	0.20 (49)	0.01 (89)	0.07 (16)	0.79 (18)	0.49 (17)
19	0.0	0.2 (54)	0.0 (32)	0.21 (86)	0
20	0.34 (58)	1.40 (73)	0	2.06 (8)	0.70 (2)
21	0.84 (51)	0.54 (45)	0	0.22 (15)	0.94 (51)
22	0.03 (0)	0.01 (11)	0.28 (15)	1.52 (40)	1.15 (41)
X	0.06 (188)	0.25 (18)	0.65 (135)	0.11 (123)	0.06 (122)

* Numbers in parentheses indicate position (cM).

more breast cancer cases diagnosed below age 50 years. Tables 3-5 give the highest LOD scores for each chromosome by group, for each of the three analyses. Figure 1 gives the maximum scores for all chromosomal locations for the three analyses in the combined dataset.

In the parametric analysis under the dominant model, the highest HLOD was 1.80 on chromosome arm 4q, close to D4S392. Positive scores were obtained at this location

TABLE 4**Maximum HLOD scores by chromosome and group, under the recessive model**

Chrom	Australia	IARC	Netherlands	UK	Total
1	0.18 (4)	0.70 (68)	0.29 (16)	0.64 (228)	0.37 (4)
2	0.67 (0)	0.62 (171)	0.13 (54)	0.66 (124)	0.64 (172)
3	0.17 (150)	0.0 (95)	0.09 (30)	0.51 (5)	0.07 (30)
4	0.57 (80)	0.31 (51)	0.12 (132)	0.30 (108)	0.15 (115)
5	0.70 (170)	0.26 (47)	0.44 (11)	0.88 (200)	1.04 (196)
6	0.10 (35)	0.63 (120)	0.02 (120)	0.27 (138)	0.37 (121)
7	0.25 (181)	0.14 (23)	0.95 (18)	1.03 (166)	0.25 (171)
8	0.40 (23)	0.61 (97)	0.02 (97)	0.96 (27)	0.95 (27)
9	0.37 (18)	0.02 (113)	1.36 (80)	0.19 (159)	0.06 (33)
10	0.34 (102)	0.22 (1)	0.02 (111)	0.71 (102)	0.76 (102)
11	0.27 (116)	0.19 (104)	0.03 (89)	0.27 (133)	0.23 (116)
12	0.86 (24)	0.31 (115)	1.29 (170)	0.56 (36)	0.14 (61)
13	0.16 (69)	0.16 (121)	0.0	0.12 (22)	0.0 (113)
14	0.33 (44)	0.90 (36)	0.0	0.10 (36)	0.46 (36)
15	0.02 (108)	0.29 (0)	0.28 (23)	0.21 (0)	0.08 (0)
16	0.41 (122)	1.16 (92)	0.0	0.37 (46)	0.66 (50)
17	0.38 (86)	0.20 (103)	0.15 (94)	0.74 (30)	0.55 (103)
18	0.63 (115)	0.0 (108)	0.0	0.94 (1)	0.19 (18)
19	0.0	0.0	0.65 (33)	0.69 (97)	0.03 (37)
20	0.0	0.0	0.06 (62)	0.98 (7)	0.0
21	0.01 (33)	0.12 (43)	0.03 (55)	0.08 (56)	0.08 (55)
22	0.28 (8)	0.04 (4)	0.36 (15)	0.13 (56)	0.33 (21)
X	0.32 (188)	0.88 (117)	0.37 (45)	1.27 (188)	0.63 (120)

* Numbers in parentheses indicate position (cM).

in the Australian, IARC, and United Kingdom series, but not in the Dutch dataset. Two other HLODs over 1 were found, on 2p (1.20, close to marker D2S2211) and on chromosome 22 (1.15, between D22S278 and D22S283). The latter result is predominantly due to a single family, EUR60, which includes 18 breast cancer cases and is the most informative family in the dataset. One branch of this family (EUR60b) generates a LOD score of 2.62. Seven women with breast cancer in this family, all belong-

TABLE 5**Maximum nonparametric scores by chromosome and group**

Chrom	Australia	IARC	Netherlands	UK	Total
1	0.37 (89)	0.49 (75)	0.06 (171)	0.04 (271)	0.23 (83)
2	1.22 (17)	0.86 (4)	0.10 (129)	0.83 (28)	1.10 (16)
3	0.54 (112)	0.59 (100)	0.04 (30)	0.72 (106)	0.71 (102)
4	0.24 (35)	2.15 (73)	0.04 (131)	0.03 (170)	0.32 (73)
5	0.68 (169)	1.31 (47)	0.15 (9)	0.20 (205)	0.28 (205)
6	0.29 (90)	0.56 (172)	0.0	0.0	0.0
7	0.72 (104)	0.72 (46)	0.24 (22)	0.58 (122)	0.42 (122)
8	0.61 (39)	0.26 (43)	0.0	0.70 (0)	0.36 (21)
9	0.02 (45)	0.00 (34)	0.74 (81)	0.06 (66)	0.02 (31)
10	0.30 (138)	0.06 (111)	0.0	0.06 (103)	0.04 (102)
11	1.00 (117)	0.14 (152)	0.98 (95)	0.16 (130)	0.38 (133)
12	0.20 (161)	0.38 (117)	0.41 (150)	1.65 (31)	0.76 (150)
13	0.17 (76)	0.43 (106)	0.00 (99)	0.23 (109)	0.12 (109)
14	0.68 (98)	0.69 (44)	0.71 (37)	0.21 (37)	1.56 (44)
15	0.0	0.02 (15)	0.02 (23)	0.99 0	0.0
16	0.24 (50)	0.13 (51)	0.0	0.33 (101)	0.09 (85)
17	0.75 (93)	0.57 (55)	0.03 (86)	0.62 (29)	0.19 (5)
18	0.23 (121)	0.74 (18)	0.11 (99)	1.31 (14)	0.85 (15)
19	0.08 (24)	0.04 (63)	0.17 (32)	0.40 (86)	0.01 (16)
20	0.10 (60)	1.34 (74)	0.0	2.22 (8)	0.46 (0)
21	0.21 (34)	0.02 (43)	0.00 (55)	0.0	0.0
22	0.26 (3)	0.15 (3)	0.65 (15)	0.00 (50)	0.00 (0)
X	0.32 (141)	0.67 (116)	0.51 (28)	0.01 (120)	0.35 (137)

* Numbers in parentheses indicate position (cM).

ging to branch EUR60b, carry the *CHEK2* 1100delC variant (Meijers-Heijboer et al. [2002]). When both branches of this family were removed, the maximum HLOD on chromosome 22 reduced to 0.06.

When analyses were restricted to families with at least four cases of breast cancer diagnosed below age 50, the maximum HLOD on 2p rose to 2.38. HLODs over 1 in this

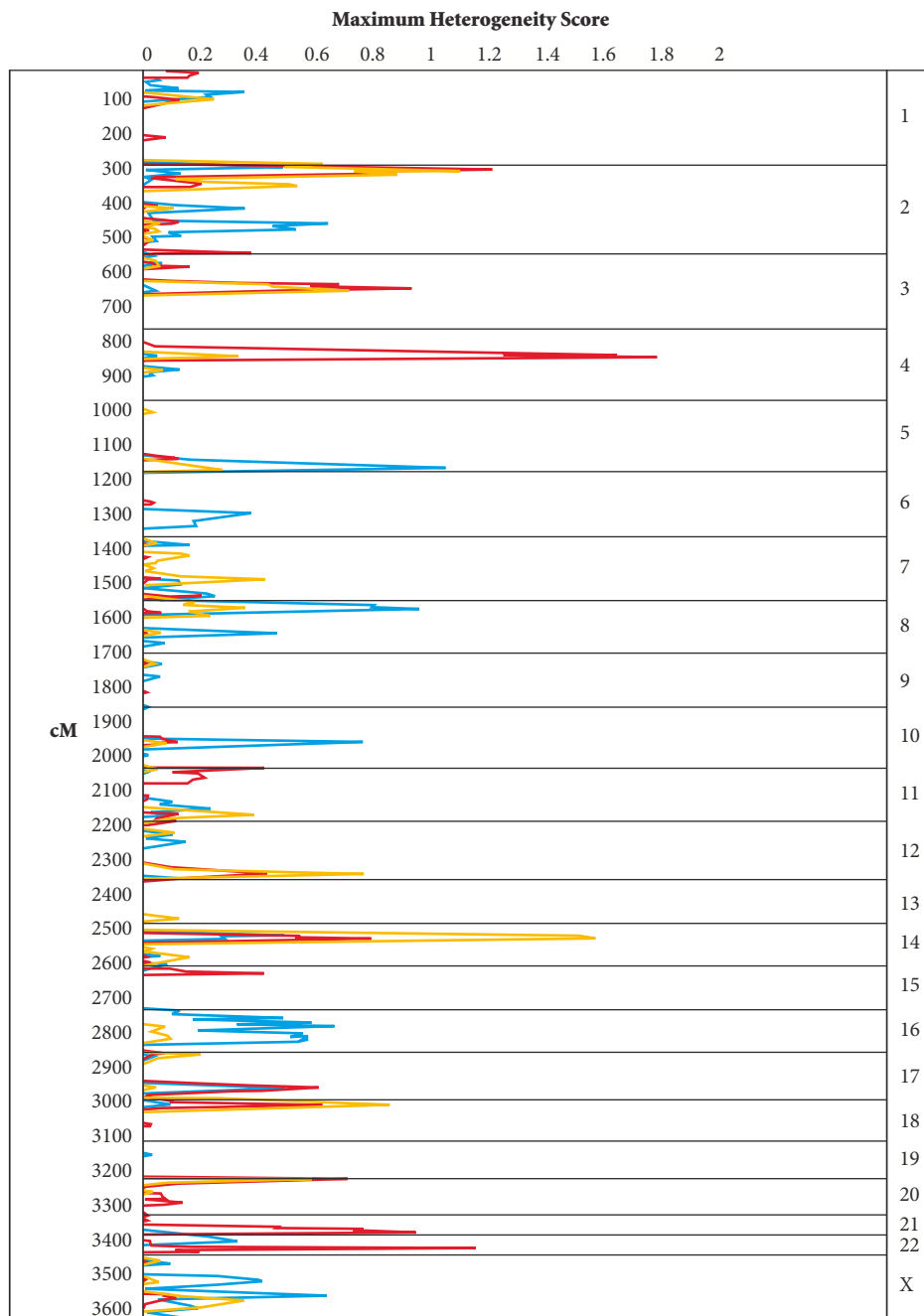


Figure 1. Maximum H_{LOD} 's by location for the dominant model (red line), the recessive model (blue line), and maximum nonparametric LOD scores (yellow line).

TABLE 6**LOD scores greater than 1.5 in individual families, under the dominant model**

Study center	Family	Chromosome	Position	LOD score
Australia	699003	2	67	1.67
IARC	2191	4	61	1.84
IARC	2191	20	70	1.80
IARC	MAYO151	3	95	1.52
IARC	MAYO151	11	43	1.59
Netherlands	RUL153	11	88	1.67
UK	EUR60a	15	24	1.50
UK	EUR60b	4	79	1.91
UK	EUR60b	22	41	2.62

subset were also found on chromosomes 4 and 22 close to the peaks in the overall analysis, and a further peak on chromosome 10 (HLOD 1.12) was also identified.

In addition to the aforementioned loci, an HLOD of 2.06 was found on 20p (at 8 cM) in the United Kingdom family set. There was, however, no evidence of linkage in the families from the other groups. LOD scores greater than 1.5 in individual families are summarized in Table 6. Of the eight scores, three contribute to the linkage peaks on chromosomes 2 and 4 found in the overall dataset. In addition two families showed linkage on chromosome 11. These peaks were however separated by over 40 cM and there was no evidence of linkage to this region in the overall analysis.

In analyses under a recessive model, only one locus reached a HLOD of greater than 1 (1.04 on 5q). In the nonparametric analysis, the highest peak was on chromosome 14 (LOD 1.56 at position 43). The only other LOD over 1 was on chromosome 2 (LOD 1.10, position 16), almost coincident with the peak in the analysis under the dominant model.

DISCUSSION

The analyses of 149 families reported here represented by far the largest genome-wide linkage screen for breast cancer susceptibility loci. The only other report since the identification of *BRCA1* and *BRCA2* was that by (Huusko et al. [2004]), who studied 14 *BRCA1/2* negative breast cancer families from Finland. Other reports have

examined specific loci on chromosome arms 6q, 8p, and 13q (Zuppan et al. [1991]; Kerangueven et al. [1995]; Seitz et al. [1997]; Kainu et al. [2000]; Rahman et al. [2000]; Thompson et al. [2002]).

The rationale for the genome-wide linkage searches is that there exist further breast cancer genes in which alleles confer high risks. The pattern of familial risks indicates that such alleles are likely to be dominant, and we therefore considered the parametric analysis assuming a dominant model to be the primary analysis. To provide some protection against model misspecification, we also conducted analyses under a recessive model and using an allele sharing approach. These approaches, however, identified no further strong linkage signals.

Under the dominant model, we found three regions with HLODs in excess of 1, but none with HLODs over 2. Of these linkage peaks, one on chromosome 22 is explained entirely by a single family (EUR60). This family is the most informative in the study, containing 18 breast cancer cases. Seven cases of breast cancer have been shown to carry the *CHEK2* variant 1100delC (Mieijers-Heijboer et al. [2002]). Since *CHEK2* is located on chromosome 22, one might hypothesize that the linkage signal is a reflection of the segregation of this variant. However, the breast cancer risk conferred by *CHEK2* 1100delC is only twofold, and this would not be expected to generate strong linkage evidence. Furthermore, the LOD score in the larger branch of EUR60 at *CHEK2* itself is only 0.3. Thus, it remains unclear whether the linkage signal on chromosome 22 reflects the effect of *CHEK2* 1100delC together with chance segregation, or whether there is an additional susceptibility allele segregating in this family. If the latter is true, given the lack of any linkage evidence from other families, susceptibility alleles at this other locus must be rare.

The strongest linkage signal in our set was found on the short arm of chromosome 4. This score was also, in part, due to EUR60 (LOD score 1.91 in the larger branch), although some evidence of linkage remained when EUR60 was excluded. The third linkage peak was on 2p (HLOD 1.2). This evidence increased (HLOD 2.4) when analyses were restricted to families with at least four cases of breast cancer diagnosed below age 50 years.

Huusko et al. ([2004]) reported evidence for linkage to markers on 2q32 in 14 Finnish breast cancer families, with a maximum LOD score of 3.20 close to D2S2262. We found no evidence of linkage in this region (maximum HLOD under the dominant model 0.0, = 0.0; nonparametric LOD = 0.05). Huusko et al. ([2004]) found one other LOD score over 1 under a dominant model, at D9S283 (1.12). Again, we found no evidence for linkage in this region. Similarly, Zuppan et al. ([1991]) found

evidence of linkage to the estrogen receptor gene on 6q in two families. In our study, we found no evidence of linkage to this region ($H_{LOD} = 0$ for both the dominant and recessive models).

Theoretical calculations indicate that, for a fully informative marker map, the expected number of regions with LOD scores of greater than 1 and 1.5 will be 5 and 2, respectively (Lander and Kruglyak, [1995]). These predictions are not strictly comparable to our analyses, since our marker sets are not fully informative. Nevertheless, they indicate that the number of linkage peaks is not clearly in excess of the number that might be expected by chance and, therefore, that the observed peaks may reflect the play of chance rather than true susceptibility loci.

Under the admixture model, the estimated proportion of families linked to the loci are 0.18, 0.18, and 0.06 for chromosomes 2, 4, and 22, respectively. Such estimates can be misleading, since they are highly dependent on the genetic model that is assumed, and the true model is unknown. However, they indicate that, even if one or more of these linkage peaks is ultimately shown to harbor a true susceptibility locus, its contribution to the familial aggregation of breast cancer is likely to be modest. Moreover, under the assumed parametric dominant model, 87% of the genome achieved an H_{LOD} of -1 or lower if the proportion of linked families was assumed to be 0.3, and 66% of the genome achieved an $H_{LOD} < -2$, indicating that such a locus was unlikely to have been missed elsewhere in the genome.

The failure to detect strong linkage signals might reflect extensive locus heterogeneity, whereby the disease is only linked to a particular locus in a small proportion of families. Under this scenario, greater power might be achievable by considering subsets of families from more homogeneous populations where genetic heterogeneity might be reduced. We were able to examine this to a limited extent by performing separate analyses of the families in each of the four study sets. Since the Australian families were largely of British and Irish origin, these two groups might be considered comparable. The Dutch population exhibits distinct founder mutations for many diseases and this group is, to an extent, genetically distinct, while the IARC families originated from many sources and are genetically heterogeneous. In the event, no strong linkage signals were observed either in the Dutch set or in the combined United Kingdom/Australian set. In particular, the linkage peaks identified in family EUR60 were not supported by linkage evidence in other Dutch families. The linkage peak on chromosome 2 did, however, become somewhat stronger when the Dutch families were excluded.

The failure to detect strong evidence for linkage may also reflect disease heterogene-

ity. Recent studies have demonstrated that breast tumors can be categorized into groups on the basis of CGH profiles and expression patterns, and that these patterns differ between *BRCA1*, *BRCA2*, and non-*BRCA1/2* familial breast cancer (Hedenfalk et al. [2001],[2003]; Gronwald et al. [2005]; Macguire et al. [2005]). These observations raise the possibility that mutations in other breast cancer susceptibility genes are associated with distinct tumor profiles. If so, incorporating tumor characteristics into the analyses could identify linkage signals that are not evident using breast cancer as a whole as the disease end point.

The positive signals found in this study indicate the most promising locations for further high-risk susceptibility genes, and would be worth following up in further families. Our results also indicate, however, that many genes are likely to be involved in breast cancer predisposition, with no gene accounting for a large fraction of the familial aggregation, and that alternative strategies will probably be necessary to identify them.

ACKNOWLEDGEMENTS

The authors would like to thank all of the families for their participation in this study. For the Australian study, the authors would like to thank the kConFab research nurses and staff for data collection, Heather Thorne, Lynda Williams, and Dani Surace for DNA preparation, Jan Groves for establishment of the LCL's, Eveline Niedermayr and Sandra Picken for supplying data, the staff of the Familial Cancer Clinics for their support of kConFab. B.A. Oostra helped out with the second phase genotyping of the Dutch study. The Breast Cancer Susceptibility Collaboration (United Kingdom) consists of the following contributors: A. Ardern-Jones, J. Berg, A. Brady, C. Brewer, G. Brice, B. Bullman, R. Cetnarsryj, C. Chapman, C. Chu, N. Coates, T. Cole, R. Davidson, A. Donaldson, H. Dorkins, F. Douglas, D. Eccles, R. Eeles, F. Elmslie, D.G. Evans, S. Goff, D. Goudie, J. Gray, L. Greenhalgh, H. Gregory, N. Haites, S.V. Hodgson, T. Homfray, R.S. Houlston, L. Izatt, L. Jeffers, V. Johnson-Roffey, F. Laloo, M. Longmuir, J. Mackay, A. Magee, S. Mansour, Dr. Zosia Miedzybrodzka, J. Miller, P. Morrison, V. Murday, J. Paterson, M. Porteous, N. Rahman, K. Redman, M. Rogers, S. Rowe, A. Sagar, A. Schofield, L. Side, M. Steel. Anita Hall and Elizabeth Mackie supported the genotyping of the United Kingdom set. Families from the IARC group were contributed in part through the Breast Cancer Family Registry (BCFR), which is supported through collaborative agreements with The University of Melbourne, Cancer Care Ontario, the Huntsman Cancer Institute and Columbia University. DFE is a Principal Research Fellow of Cancer Research UK.

REFERENCE LIST

- **Anglian Breast Cancer Study Group.** Prevalence and penetrance of *BRCA1* and *BRCA2* mutations in a population-based series of breast cancer cases. *Br J Cancer* 83: 1301-1308 (2000)
- **Antoniou A, Pharoah PD, Narod S, et al.** Average risks of breast and ovarian cancer associated with *BRCA1* or *BRCA2* mutations detected in case series unselected for family history: a combined analysis of 22 studies. *Am J Hum Genet* 72: 1117-1130 (2003)
- **Antoniou AC, Pharoah PDP, Smith P, Easton DF.** The BOADICEA model of genetic susceptibility to breast and ovarian cancer. *Br J Cancer* 91: 1580-1590. (2004)
- **The CHEK2 Breast Cancer Case-Control Consortium.** CHEK2*1100delC and susceptibility to breast cancer: A collaborative analysis involving 10,860 breast cancer cases and 9,065 controls from ten studies. *Am J Hum Genet* 74: 1175-1182 (2004)
- **Claus EB, Risch N, Thompson WD.** Genetic analysis of breast cancer in the cancer and steroid hormone study. *Am J Hum Genet* 48: 232-242. (1991)
- **Collaborative Group on Hormonal Factors in Breast Cancer.** Familial breast cancer: Collaborative reanalysis of individual data from 52 epidemiological studies including 58,209 women with breast cancer and 101,986 women without the disease. *Lancet* 358: 1389-1399. (2001)
- **Cottingham RW, Jr, Idury RM, Schaffer AA.** Faster sequential genetic linkage computations. *Am J Hum Genet* 53: 252-263. (1993)
- **Dite GS, Jenkins MA, Southey MC, et al.** Familial risks, early-onset breast cancer, and *BRCA1* and *BRCA2* germline mutations. *J Natl Cancer Inst* 95: 448-457 (2003)
- **Easton DF.** How many more breast cancer predisposition genes are there? *Breast Cancer Res* 1: 14-17 (1999)
- **Easton DF, Bishop DT, Ford D, Crookford GP.** Genetic linkage analysis in familial breast and ovarian cancer: results from 214 families. The Breast Cancer Linkage Consortium. *Am J Hum Genet.* 52: 678-701 (1993)
- **Ford D, Easton DF, Stratton M, et al. the Breast Cancer Linkage Consortium.** Genetic heterogeneity and Penetrance analysis of the *BRCA1* and *BRCA2* genes in breast cancer families. *Am J Hum Genet* 62: 334-345 (1998)
- **Gronwald J, Jauch A, Cybulski C, et al.** Comparison of genomic abnormalities between BRCAx and sporadic breast cancers studied by comparative genomic hybridization. *Int J Cancer* 20: 230-236 (2005)
- **Hedenfalk I, Duggan D, Chen Y, et al.** Gene-expression profiles in hereditary breast cancer. *N Engl J Med* 22: 539-548 (2001)
- **Hedenfalk I, Ringner M, Ben-Dor A, et al.** Molecular classification of familial non-*BRCA1/BRCA2* breast cancer. *Proc Natl Acad Sci USA* 100: 2532-2537 (2003)
- **Hopper JL, Chenevix-Trench G, Jolley DJ, et al.** Design and analysis issues in a population-based, case-control-family study of the genetic epidemiology of breast cancer and the Co-operative Family Registry for Breast Cancer Studies (CFRBCS). *J Natl Cancer Inst Monogr* 26: 95-100 (1999)
- **Huusko P, Juo SH, Gillanders E, et al.** Genome-wide scanning for linkage in Finnish breast cancer families. *Eur J Hum Genet* 12: 98-104 (2004)
- **Kainu T, Juo SH, Desper R, et al.** Somatic deletions in hereditary breast cancers implicate 13q21 as a putative novel breast cancer susceptibility locus. *Proc Natl Acad Sci USA* 97: 9603-9608 (2000)
- **Kerangueven F, Essioux L, Dib A, et al.** Loss of heterozygosity and linkage analysis in breast carcinoma: indication for a putative third susceptibility gene on the short arm of chromosome 8. *Oncogene* 10: 1023-1026 (1995)

- **Kong A, Cox NJ.** Allele-sharing models: LOD scores and accurate linkage tests. *Am J Hum Genet* 61: 1179-1188 (1997)
- **Kong A, Gudbjartsson D, Sainz J, et al.** A high-resolution recombination map of the human genome. *Nat Genet* 31: 241-247 (2002)
- **Kruglyak L, Daly MJ, Reeve-Daly MP, Lander ES.** Parametric and nonparametric linkage analysis: A unified multipoint approach. *Am J Hum Genet* 58: 1347-1363 (1996)
- **Lander E, Kruglyak L.** Genetic dissection of complex traits: Guidelines for interpreting and reporting linkage results. *Nat Genet* 11: 241-247 (1995)
- **Lange K, Weeks D, Boehnke M.** Programs for Pedigree Analysis: MENDEL, FISHER, and dGENE. *Genet Epidemiol* 5: 471-472 (1988)
- **Maguire P, Holmberg K, Kost-Alimova M, et al.** CGH analysis of familial non-*BRCA1/BRCA2* breast tumors and mutation screening of a candidate locus on chromosome 17q11.2-12. *Int J Mol Med* 16: 135-141 (2005)
- **Meijers-Heijboer H, van den Ouweland A, Klijn J, et al.** Low-penetrance susceptibility to breast cancer due to *CHK2* 110delC in non-carriers of *BRCA1* or *BRCA2* mutations. *Nat Genet* 31: 55-59 (2002)
- **Miki Y, Swensen J, Shattuck-Eidens D, et al.** A strong candidate for the breast and ovarian cancer susceptibility gene *BRCA1*. *Science* 266: 66-71 (1994)
- **O'Connell JR, Weeks DE.** The VITESSE algorithm for rapid exact multilocus linkage analysis via genotype set-recoding and fuzzy inheritance. *Nat Genet* 11: 402-408 (1995)
- **Ott J.** Linkage analysis and family classification under heterogeneity. *Ann Hum Genet* 47: 311-320 (1983)
- **Peto J, Mack TM.** High constant incidence in twins and other relatives of women with breast cancer. *Nat Genet* 26: 411-414 (2000)
- **Peto J, Collins N, Barfoot R, et al.** Prevalence of *BRCA1* and *BRCA2* gene mutations in patients with early-onset breast cancer. *J Natl Cancer Inst* 91: 943-949 (1999)
- **Rahman N, Teare MD, Seal S, et al.** Absence of evidence for a familial breast cancer susceptibility gene at chromosome 8p12-p22. *Oncogene* 19: 4170-4173 (2000)
- **Seitz S, Rohde K, Bender E, et al.** Strong indication for a breast cancer susceptibility gene on chromosome 8p12-p22: Linkage analysis in German breast cancer families. *Oncogene* 14: 741-743 (1997)
- **Thompson D, Easton D.** The genetic epidemiology of breast cancer genes. *J Mammary Gland Biol Neoplasia* 9: 221-236 (2004)
- **Thompson D, Duedal S, Kirner J, et al.** Cancer risks and mortality in heterozygous *ATM* mutation carriers. *J Natl Cancer Inst* 97: 813-822 (2005)
- **Thompson D, Szabo CI, Mangion J, et al.** Evaluation of linkage of breast cancer to the putative *BRCA3* locus on chromosome 13q21 in 128 multiple case families from the Breast Cancer Linkage Consortium. *Proc Natl Acad Sci USA* 99: 827-831 (2002)
- **Whittemore AS, Halpern J.** A class of tests for linkage using affected pedigree members. *Biometrics* 50: 118-127 (1994)
- **Wooster R, Bignell G, Swift S, et al.** Identification of the breast cancer susceptibility gene *BRCA2*. *Nature* 378: 789-792 (1995)
- **Zuppan P, Hall JM, Lee MK, et al.** Possible linkage of the estrogen receptor gene to breast cancer in a family with late-onset disease. *Am J Hum Genet* 48: 1065-1068 (1991)