ARTICLE

Genome-wide scanning for linkage in Finnish breast cancer families

Pia Huusko^{1,2,9}, Suh-Hang Hank Juo^{3,4,9}, Elizabeth Gillanders^{1,9}, Laura Sarantaus^{5,9}, Tommi Kainu¹, Pia Vahteristo⁵, Minna Allinen², MaryPat Jones¹, Katrin Rapakko², Hannaleena Eerola^{4,5}, Carol Markey¹, Paula Vehmanen⁵, Derek Gildea¹, Diane Freas-Lutz¹, Carl Blomqvist⁶, Jaakko Leisti², Guillermo Blanco⁷, Ulla Puistola⁸, Jeffrey Trent¹, Joan Bailey-Wilson³, Robert Winqvist², Heli Nevanlinna⁵ and Olli-P Kallioniemi^{*,1}

¹Cancer Genetics Branch, NHGRI, NIH, Helsinki University, Central Hospital, Helsinki, Finland; ²Department of Clinical Genetics, Oulu University Hospital/University of Oulu, Oulu, Finland; ³Inherited Disease Research Branch, NHGRI, NIH, Helsinki University Central Hospital, Helsinki, Finland; ⁴Columbia Genome Center, New York, NY, USA; ⁵Department of Obstetrics and Gynecology, Helsinki University Central Hospital, Helsinki, Finland; ⁶Department of Oncology, Helsinki University Central Hospital, Helsinki, Finland; ⁷Department of Oncology, Oulu University Hospital/University of Oulu, Oulu, Finland; ⁸Department of Obstetrics and Gynecology, Oulu University Hospital/University of Oulu, Oulu, Finland

Only a proportion of breast cancer families has germline mutations in the BRCA1 or BRCA2 genes, suggesting the presence of additional susceptibility genes. Finding such genes by linkage analysis has turned out to be difficult due to the genetic heterogeneity of the disease, phenocopies and incomplete penetrance of the mutations. Isolated populations may be helpful in reducing the level of genetic heterogeneity and in providing useful starting points for further genetic analyses. Here, we report results from a genome-wide linkage analysis of 14 high-risk breast cancer families from Finland. These families tested negative for BRCA1 and BRCA2 germline mutations and showed no linkage to the 13q21 region, recently proposed as an additional susceptibility locus. Suggestive linkage was seen at marker D2S364 (2q32) with a parametric two-point LOD score of 1.61 ($\theta = 0$), and an LOD score of 2.49 in nonparametric analyses. Additional genotyping of a 40 cM chromosomal region surrounding the region of interest yielded a maximum parametric two-point LOD score of 1.80 ($\theta = 0$) at D2S2262 and a nonparametric LOD score of 3.11 at an adjacent novel marker 11291M1 in BAC RP11-67G7. A nonparametric multipoint LOD score of 3.20 was seen at 11291M1 under the assumption of dominant inheritance. While not providing proof of linkage considering the small number of families and large number of laboratory and statistical analyses performed, these results warrant further studies of the 2q32 chromosomal region as a candidate breast cancer susceptibility locus. Both linkage and association studies are likely to be useful, particularly in other isolated populations.

European Journal of Human Genetics (2004) 12, 98–104. doi:10.1038/sj.ejhg.5201091

Published online 15 October 2003

Keywords: genome-wide linkage; breast cancer; Finland; chromosome 2q

*Correspondence: Dr Olli-P Kallioniemi, Medical Biotechnology Group, VTT Technical Research Centre of Finland, PO Box 106, FIN-20521 VTT, Finland. Tel: +358 2 478 8602; Fax: +358 2 478 8601; E-mail: olli.kallioniemi@vtt.fi

⁹These authors contributed equally to this work

Received 2 October 2002; revised 15 July 2003; accepted 5 August 2003

Introduction

The two major breast cancer susceptibility genes, BRCA1 and BRCA2, account for a majority of high-risk families with both early-onset breast cancer and ovarian cancer.¹ However, most of the families with less than six cases of

female breast cancer and no ovarian cancer do not carry BRCA1 or BRCA2 mutations that can be detected by conventional sequencing.^{1–5} Furthermore, the proportion of breast cancer families attributable to these genes may also vary among different populations. For example, a large fraction of breast cancer in Ashkenazi Jewish families can be explained by prevalent founder mutations in the BRCA1 and BRCA2 genes,⁶ whereas in the Finnish population mutations in these genes seem to be less frequent.^{5,7,8}

Many efforts have been made to discover additional, previously unknown susceptibility genes by linkage or association analyses. Positive linkage has been observed at chromosome region 8p12–p22 in two German families.⁹ Recently, we presented evidence for a novel putative breast cancer susceptibility locus at chromosome region 13q21.¹⁰ Neither of these findings could be confirmed by the Breast Cancer Linkage Consortium,^{11,12} suggesting that these linkages may either represent population-, phenotype- or family cohort-specific effects or simply type I or II error. Indeed, our recent reanalyses of the 13q21–q22 linkage data suggest that linkage was primarily seen in Finnish families that exhibited only breast cancer and relatively few affected cases (S Juo, personal communication).

The aim of this study was to perform a genome-wide linkage (GWL) search for additional susceptibility loci in a putatively homogeneous group of 14 Finnish breast cancer families without BRCA1 or BRCA2 involvement, and with no evidence of linkage to the 13q21 site, followed by fine mapping of interesting chromosomal regions. Using relatively homogeneous populations, such as the Finns, as the target for this study, was likely to reduce the extent of genetic heterogeneity that usually makes linkage analyses of complex diseases difficult. Since recent segregation analyses have suggested an involvement of a recessive breast cancer gene,^{13–15} we also tested both recessive and dominant models of inheritance.

Materials and methods

Families

In all, 14 Finnish breast cancer families collected at the Helsinki and Oulu University Hospitals were included in the study. All families had at least three breast cancer cases with DNA available for genotyping. Both Oulu and Helsinki families were previously tested negative for BRCA1 and BRCA2 mutations by heteroduplex analysis (ie DGGE, SSCP or CSGE), protein truncation test and linkage analysis (data not shown).^{5,7,8} In addition, the Oulu families were excluded having large rearrangements by Southern blot analysis.¹⁶ All families were screened for CHEK2 mutation 1100delC, which is present in 1.1–1.4% of the normal population and has recently been identified as a low-penetrance allele with a two-fold increased breast cancer risk.^{17,18} Some individuals in families 178 and 277 were found to carry this mutation;¹⁸ however, due to the

incomplete segregation and the low penetrance of the allele, these families were included in this study. Of the 14 families now investigated, 10 were included in our previous 13q-linkage study. None of the families showed linkage to chromosome region 13q21.¹⁰ Initially, 48 affected and 45 unaffected individuals were genotyped for GWL. After the initial scan, 32 additional family members (five affected and 27 unaffected) were genotyped for follow-up chromosomal fine mapping. Mean ages of onset in the 14 families varied from 43.3 to 63.6 years. The detailed characteristics of the families are described in Table 1. Breast cancer diagnoses in the families were confirmed through hospital records and the Finnish Cancer Registry. The study was performed under informed consent and with appropriate permission from the Ethical Committees of Helsinki University Central Hospital, Oulu University Hospital, Ministry of Social Affairs and Health in Finland as well as the National Institutes of Health (NIH).

Genotyping

Genotyping was performed in NHGRI, NIH. For the genome-wide scan, 398 microsatellite markers from the ABI Prism Linkage Mapping Set version 2 (Applied Biosystems, Foster City, CA, USA) were used, spanning the human genome in approximately 10 cM intervals (Figure 1). Each marker was independently amplified according to standard PCR protocol. The PCR products were then pooled and run on the ABI 377 DNA Sequencer. Electrophoretic data were analyzed by two independent individuals using Genescan and Genotyper software programs (Applied Biosystems, Foster City, CA, USA). The allele sizes were determined using the CEPH family member 134702. For fine mapping, an additional set of 34 markers, spanning a 40 cM chromosomal region from 2q24 to 2q33, was used (Figure 1) (www.marshfieldclinic.org). The primer sequences for four markers in BAC RP11-67G7 are as follows: 11291M1 forward, 5'-TTTCAAGAGCAACCTTTCAAGA-3'; 11291M1 reverse, 5'-GTGTCTGGGATAACGTTGATGGGATTT-3'; 11291M2 forward, 5'-TGGGGAAGATGAGCAATGTT-5'; 11291M2 reverse, 5'-GTGTCTGTGACCAGGGTAGAGGCAAG-3'; 11291M3 forward, 5'-GGGCTGCAGTTTGTCTGT-3'; 11291M3 reverse, 5'-GTGTCTCGGGATGTTATCCCATCAG-3'; 11291M4 forward, 5'-TGAGCCTACAAAATGCCTCTG-3'; 11291M4 reverse, 5'-GTGTCTTGGCTCACATACTTGCTCA-3'. The heterozygosities for these markers are 60, 71, 70 and 64%, respectively. PCR reactions were performed in a 15 µl volume containing 20 ng of genomic DNA, 0.33 µM each primer, 0.25 mM each dNTP, 2.5 mM MgCl₂, 10 mM Tris-HCl, 50 mM KCl and 0.5 U *Taq* polymerase. PCR amplification was performed using the GeneAmp 9600 or 9700 thermocyclers (Applied Biosystems, Foster City, CA, USA). PCR cycling conditions were as follows: 95°C for 12 min; followed by 10 cycles of 94°C for 15 s, 55°C for 15 s, 72°C for 30 s; and 20 cycles of

Table 1 Families analyzed in GWL

Family ID	Hospital	No. of affecteds (breast cancer)	No. of affecteds genotyped	No. of unaffecteds genotyped	Mean ages of dx (years)	Other cancers	Max. two-point LOD at D2S364 Parametric/nonparametric
7	Helsinki	5	3	2	54.2	Melanoma, neck, ovary, skin ^a	0.05/0.00
96	Helsinki	4	4	5	44.5	Connective tissue, esophagus	−0.01/−0.01
123	Helsinki	4	3	6	60.7	Colon, gastrointestinal, skin ^a	0.00/−0.03
178	Helsinki	4	3	5	52.5	Gastrointestinal, skin ^a	0.00/0.33
261	Helsinki	5	4	7	52.4	Brain, esophagus, kidney, lung, testis, urinary system	0.14/0.66
277	Helsinki	5	3	2	55.0	Prostate	0.09/0.33
381	Helsinki	5	3	4	51.6	Esophagus, liver, lung, skin ^a	0.19/0.39
437	Helsinki	3	3	2	63.6	Ovary, skin ^a	0.00/0.12
477	Helsinki	5	3	3	52.4	Gastrointestinal, skin ^a , thyroid	0.25/0.31
574	Helsinki	5	4	4	58.2	Buccal cavity, Hodgkin's	0.67/0.91
3004	Oulu	7	3	—	43.3	Brain, gastrointestinal, pancreas, prostate, sarcoma	0.14/0.19
3008	Oulu	3	3	2	50.3	Gastrointestinal	0.00/0.82
3009	Oulu	6	6	3	50.6	Esophagus, lung, melanoma	0.65/0.82
3011	Oulu	4	3	—	56.8	Colon, liver, ovary, stomach	0.10/0.20

^aOther than melanoma.

89°C for 15 s, 55°C for 15 s, 72°C for 30 s, with a final extension of 72°C for 10 min.

Statistical analyses

Genotyping data were checked for Mendelian inconsistency using the genetic analysis system and PedCheck programs.¹⁹ Any marker violating the rules of Mendelian transmission was double checked by the genotyping laboratory. Ambiguous marker genotypes were deleted. The biological relationships were also examined using the RelCheck computer program.²⁰ The program uses the genotypic information of autosomal markers and calculates the likelihoods to infer the true relationship of a putative sibling pair. Five relationships were considered: monozygotic twins, parental/offspring, full sibs, half-sibs and unrelated individuals.

Allele frequencies were calculated from all genotyped family members in the genome-wide analysis by the Gconvert program (<http://www2.qimr.edu.au/davidD>). For the fine-mapping markers on 2q, we also included additional 270 unrelated Finnish controls to obtain better allele frequency estimations. Both parametric and non-parametric analyses were performed under both dominant and recessive inheritance models. For the dominant parametric analysis, the CASH model²¹ as modified by Easton *et al*²² was used. In brief, this model specifies dominant inheritance of a susceptibility allele and 14 age-specific liabilities (seven age-specific cumulative risks for unaffecteds, and another seven age-specific densities for affecteds). The population frequency of a disease allele is determined to be 0.0033. For the recessive model, the parameters were adopted from Cui *et al*.¹⁴ The recessive model specifies a disease allele frequency of 0.063, and the

Figure 1 GWL plot of two-point parametric (a) and nonparametric pseudomarker (b) analysis using dominant mode of inheritance. Vertical lines indicate the chromosome boundaries and horizontal lines the LOD score values.

risk of a homozygote is 50% by age 40 years and near certainty by age 60 years. Similarly, a total of 14 age-specific liability classes were used in the recessive model. Nonparametric analyses were performed using a newly developed pseudomarker approach,^{23,24} which approximates a 'model-free' affected relative pair analysis, but maintains an important property of LOD score analysis: pedigree correlations between all relatives are considered jointly, and the pedigree is not broken into sets of all possible relative pairs. The nonparametric dominant inheritance analysis assumed that all affecteds are gene carriers, and the disease allele is infinitesimally rare, according to the pseudomarker strategy.²⁴ Similarly, the nonparametric recessive inheritance analysis assumed a rare disease allele and no phenocopies. Both parametric and nonparametric analyses (ie pseudomarker approach) were carried out using the FASTLINK program.^{25–27} Sliding

multipoint analysis was conducted in the region of most significant two-point LOD scores by using the FASTLINK program. In all analyses, the breast cancer cases were coded as affected and all other cancer cases as unaffected.

Results

In the GWL analyses, assuming dominant inheritance, marker D2S364 at chromosome 2q32 gave the highest parametric two-point LOD score of 1.61 ($\theta=0$). The second highest LOD score of 1.12 ($\theta=0$) was seen for marker D9S283 at 9q21. Marker D2S364 also showed the highest LOD score (2.49) in the nonparametric pseudomarker analysis, using a dominant mode of inheritance (Figure 1). Under the assumption of recessive inheritance, seven out of 398 markers showed LOD scores above one in parametric analysis. These markers were D2S367 (2p22), D2S364 (2q32), D3S1304 (3p26), D11S4177 (11p15), D19S921 (19q13), DXS1001 (Xq24) and DXS1227 (Xq27). The LOD score at D2S364 was 1.77 ($\theta=0.05$), while the highest LOD score of 1.96 ($\theta=0.05$) was seen at marker DXS1001 (Xq24). In the nonparametric recessive analysis, the following five markers had LOD scores above one: D2S364 (2q32), D3S1304 (3p26), D9S1690 (9q31), D11S902 (11p15) and DXS1073 (Xq28). The LOD score at D2S364 was 1.43, while the highest LOD score of 1.72 was seen at D11S902 (11p15) (data not shown).

Since the 2q32 region was most consistently positive in the statistical analyses, we chose to focus on additional genotyping at this region, and performed linkage analyses assuming a dominant mode of inheritance. We genotyped 34 additional markers covering approximately 40 cM around the peak marker D2S364. A maximum parametric two-point LOD score, 1.80 ($\theta=0$), was observed for marker D2S2262 at 2q32.2. The highest LOD score of 3.11 was seen at marker 11291M1 in nonparametric analysis. This marker was developed using a polymorphic nucleotide repeat in BAC RP11-67G7, which maps to chromosome region 2q32.1 in between markers D2S364 and D2S2273. The highest multipoint LOD score was 3.20, seen at the same marker (Figure 2). The HOMOG program was used to test for evidence of genetic heterogeneity for markers D2S364 and 11291M1 which yielded the highest LOD scores. The results from HOMOG gave no significant evidence of genetic heterogeneity. However, it is well known that the A-test used in HOMOG has low power to detect heterogeneity in complex diseases. Thus, the negative results from HOMOG do not necessarily mean that no heterogeneity exists in these families.

According to the recombinations seen in one of the most informative families, 3009, the critical region spans from marker D2S2177 to D2S117 (Figure 3). No common haplotypes between the linked families were seen (data

Figure 2 Results from the fine mapping of the chromosome 2q candidate region with 38 markers. The columns beginning from the left show the markers, physical distance, nonparametric LOD scores (two-point and multipoint) and parametric LOD scores. The GWL markers are shown in red. The peak LOD scores are underlined and values above one are in italics.

not shown), suggesting multiple different mutations in these families.

Discussion

We report in this study suggestive evidence of a novel breast cancer susceptibility locus at chromosome region 2q32 in Finnish families without involvement of BRCA1, BRCA2 or chromosome region 13q21. The existence of a novel breast cancer susceptibility gene was initially supported by both parametric (LOD score 1.61) and nonparametric (LOD score 2.49) analyses of the GWL data. Further evidence for linkage was obtained from additional genotyping with more markers, providing the final non-parametric multipoint LOD score of 3.20.

The entire area of mild linkage positivity spans about 40 cM from D2S156 (2q24) to D2S325 (2q33). The number of known genes in this region is 194. The 5 cM core region from D2S384 to D2S2262, where the LOD scores are above one in both parametric and nonparametric analyses, includes 28 known genes. Several candidate genes of potential biological interest are located here, ITGAV a member of integrin gene family, as well as PMS1, one of the HNPCC (hereditary nonpolyposis colorectal cancer) genes.

Figure 3 Haplotypes of family 3009. The cancer types and ages at diagnosis, as well as ages at last breast exam for unaffected individuals, when known, are below each affected individual. The numbers for the affected individuals have been encircled. Red haplotype represents the disease-linked haplotype. Gray area indicates the critical region between the recombinations. X indicates marker tests that failed.

The sequence in this region is not complete and additional genes are likely to be found within the gaps.

Regarding the involvement of recessive genes in breast cancer,^{13–15} chromosome region Xq24 may be interesting. Previously, candidate loci for prostate (HPC-X) and testicular cancer have been linked to this chromosomal arm, but substantially more distally, at the Xq27 region.^{28,29}

Based on the prior experience in replicating linkage results of complex diseases from one population to another, it is possible that 2q32 will not turn out to be a major predisposition locus in more heterogeneous sample sets available to most investigators. In fact, based on the paucity of published linkage results in breast cancer, it is increasingly likely that no region in the human genome will display overwhelmingly positive LOD scores in the general Western European or US populations. Obviously, after the BRCA1 and BRCA2 genes are excluded, many factors may contribute to the familial clustering of breast cancer. It is possible that some of the remaining familial breast cancers are due to chance clustering of apparently sporadic cases.³⁰ Shared lifestyle, dietary and environmental effects within a family may also have a significant contribution to disease. Mutations in TP53,³¹ PTEN³² and ATM³³ genes may each explain a small fraction of hereditary breast cancer. Recent studies have also implied a polygenic model, where the susceptibility to breast

cancer is conferred by a large number of alleles.³⁴ The contribution of low-penetrance susceptibility genes, such as CHEK2^{17,18} to breast cancer causation appears very important. Considering these multiple genetic and non-genetic causes that may explain familial clustering of breast cancer, further efforts to investigate the reported 2q32 linkage are needed before any conclusions on the significance of the reported data at this locus are drawn. These analyses should incorporate both linkage- and nonlinkage-based methods, including association analyses and studies of candidate genes in the region. Studies in other isolated populations would be particularly informative.

In summary, we present preliminary evidence of linkage in Finnish breast cancer families to a susceptibility locus at the chromosomal region 2q32. The presence of a gene predisposing to breast cancer at this locus will obviously have to be confirmed by an independent study, involving refined linkage mapping, genetic association and studies of candidate genes. Finally, analyses of this region at 2q as a target for somatic alterations in sporadic breast cancer may be informative.

Acknowledgements

We thank Leena Kukkola, Tuija Löppönen, Minna Merikivi and Outi Vierimaa for patient contacts as well as Marika Kujala and Kati Outila for technical assistance. John Carpten is acknowledged for the data on the markers in BAC RP11-67G7. This work has been partly supported by the Academy of Finland, Cancer Foundation of Northern Finland, Finnish Breast Cancer Group, Finnish Cancer Society, Finnish Cultural Foundation, Helsinki University Central Hospital Research Fund, Helsinki University Science Foundation, Oulu University Hospital, Sigrid Juselius Foundation and University of Oulu. We also thank all patients for participation.

References

- 1 Ford D, Easton DF, Stratton M *et al*: Genetic heterogeneity and penetrance analysis of the BRCA1 and BRCA2 genes in breast cancer families. The Breast Cancer Linkage Consortium. *Am J Hum Genet* 1998; **62**: 676–689.
- 2 Rebbeck TR, Couch FJ, Kant J *et al*: Genetic heterogeneity in hereditary breast cancer: role of BRCA1 and BRCA2. *Am J Hum Genet* 1996; **59**: 547–553.
- 3 Serova OM, Mazoyer S, Puget N *et al*: Mutations in BRCA1 and BRCA2 in breast cancer families: are there more breast cancer susceptibility genes? *Am J Hum Genet* 1997; **60**: 486–495.
- 4 Schubert EL, Lee MK, Mefford HC *et al*: BRCA2 in American families with four or more cases of breast or ovarian cancer: recurrent and novel mutations, variable expression, penetrance and the possibility of families whose cancer is not attributable to BRCA1 or BRCA2. *Am J Hum Genet* 1997; **60**: 1031–1040.
- 5 Vehmanen P, Friedman LS, Eerola H *et al*: Low proportion of BRCA1 and BRCA2 mutations in Finnish breast cancer families: evidence for additional susceptibility genes. *Hum Mol Genet* 1997; **13**: 2309–2315.
- 6 Roa BB, Boyd AA, Volcik K, Richards CS: Ashkenazi Jewish population frequencies for common mutations in BRCA1 and BRCA2. *Nat Genet* 1996; **4**: 185–187.
- 7 Vehmanen P, Friedman LS, Eerola H *et al*: A low proportion of BRCA2 mutations in Finnish breast cancer families. *Am J Hum Genet* 1997; **60**: 1050–1058.
- 8 Huusko P, Paakkonen K, Launonen V *et al*: Evidence of founder mutations in Finnish BRCA1 and BRCA2 families. *Am J Hum Genet* 1998; **62**: 1544–1548.
- 9 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* 1997; **14**: 741–743.
- 10 Kainu T, Juo S-H, 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* 2000; **97**: 9603–9608.
- 11 Rahman N, Teare MD, Seal S *et al*: Absence of evidence for a familial breast cancer susceptibility gene at chromosome 8p12–p22. *Oncogene* 2000; **19**: 4170–4173.
- 12 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* 2002; **99**: 827–831.
- 13 Antoniou AC, Pharoah PDP, McMullan G, Day NE, Ponder BAJ, Easton D: Evidence for further breast cancer susceptibility genes in addition to BRCA1 and BRCA2 in a population-based study. *Genet Epidemiol* 2001; **21**: 1–18.
- 14 Cui J, Antoniou AC, Dite GS *et al*: After BRCA1 and BRCA2 - What next? Multifactorial segregation analysis of three-generation, population-based Australian families affected by female breast cancer. *Am J Hum Genet* 2001; **68**: 420–431.
- 15 Antoniou AC, Pharoah PDP, McMullan G *et al*: A comprehensive model for familial breast cancer incorporating BRCA1, BRCA2 and other genes. *Br J Cancer* 2002; **86**: 76–83.
- 16 Lahti-Domenici J, Rapakko K, Paakkonen K *et al*: Exclusion of large deletions and other rearrangements in BRCA1 and BRCA2 in Finnish breast and ovarian cancer families. *Cancer Genet Cytogenet* 2001; **129**: 120–123.
- 17 Meijers-Heijboer H, van den Ouweland A, Klijn J *et al*: Low-penetrance susceptibility to breast cancer due to CHEK2 1100delC in noncarriers of BRCA1 or BRCA2 mutations. *Nat Genet* 2002; **31**: 55–59.
- 18 Vahteristo P, Bartkova J, Eerola H *et al*: A CHEK2 variant contributing to a substantial fraction of familial breast cancer. *Am J Hum Genet* 2002; **71**: 432–438.
- 19 O'Connell JR, Weeks DE: PedCheck: a program for identification of genotype incompatibilities in linkage analysis. *Am J Hum Genet* 1998; **63**: 259–266.
- 20 Broman KW, Weber JL: Estimation of pairwise relationships in the presence of genotyping errors. *Am J Hum Genet* 1998; **63**: 1563–1564.
- 21 Claus EB, Risch N, Thompson WD: Genetic analysis of breast cancer in the cancer and steroid hormone study. *Am J Hum Genet* 1991; **48**: 232–242.
- 22 Easton DF, Bishop DT, Ford D, Crockford GP: Genetic linkage analysis in familial breast and ovarian cancer: results from 214 families. The Breast Cancer Linkage Consortium. *Am J Hum Genet* 1993; **52**: 678–701.
- 23 Goring HH, Terwilliger JD: Linkage analysis in the presence of errors III: marker loci and their map as nuisance parameters. *Am J Hum Genet* 2000; **66**: 1298–1309.
- 24 Goring HH, Terwilliger JD: Linkage analysis in the presence of errors IV: joint pseudomarker analysis of linkage and/or linkage disequilibrium on a mixture of pedigrees and singletons when the mode of inheritance cannot be accurately specified. *Am J Hum Genet* 2000; **66**: 1310–1327.
- 25 Lathrop GM, Lalouel JM, Julier C, Ott J: Strategies for multilocus linkage analysis in humans. *Proc Natl Acad Sci USA* 1984; **81**: 3443–3446.
- 26 Cottingham RW, Idury RM, Schäffer AA: Faster sequential genetic linkage computations. *Am J Hum Genet* 1993; **53**: 252–263.

- 27 Schaffer AA, Gupta SK, Shriram K, Cottingham RW: Avoiding recomputation in linkage analysis. *Hum Hered* 1994; **44**: 225–237.
- 28 Xu J, Meyers D, Freije D *et al*: Evidence for a prostate cancer susceptibility locus on the X chromosome. *Nat Genet* 1998; **20**: 175–179.
- 29 Rapley EA, Crockford GP, Teare D *et al*: Localization to Xq27 of a susceptibility gene for germ-cell tumours. *Nat Genet* 2000; **24**: 197–200.
- 30 Cui J, Hopper JL: Why are the majority of hereditary cases of early-onset breast cancer sporadic? A simulation study. *Cancer Epidemiol Biomarkers Prev* 2000; **9**: 805–812.
- 31 Malkin D, Li FP, Strong LC *et al*: Germ line p53 mutations in a familial syndrome of breast cancer, sarcomas and other neoplasms. *Science* 1990; **250**: 1233–1238.
- 32 Nelen MR, Padberg GW, Peeters EAJ *et al*: Localization of the gene for Cowden disease to chromosome 10q22–23. *Nat Genet* 1996; **13**: 14–116.
- 33 Savitsky K, Bar-Shira A, Gilad S *et al*: A single ataxia telangiectasia gene with a product similar to PI-3 kinase. *Science* 1995; **268**: 1749–1753.
- 34 Pharoah PDP, Antoniou A, Bobrow M, Zimmern RL, Easton DF, Ponder BAJ: Polygenic susceptibility to breast cancer and implications for prevention. *Nat Genet* 2002; **31**: 33–36.