Evidence for a prostate cancer susceptibility locus on the X chromosome

Jianfeng Xu[1*], Deborah Meyers[1], Diha Freije[2], Sarah Isaacs[2], Kathy Wiley[2], Deborah Nusskern[2], Charles Ewing[2], Eric Wilkens[2], Piroska Bujnovszky[2], G. Steven Bova[2,3], Patrick Walsh[2] & William Isaacs[2,4] (*Group 1*)
Johanna Schleutker[5*], Mika Matikainen[5*], Teuvo Tammela[5], Tapio Visakorpi[5] & Olli-P. Kallioniemi[12] (*Group 2*)
Rebecca Berry[6*], Daniel Schaid[7*], Amy French[6], Shannon McDonnell[7], Jennifer Schroeder[6], Michael Blute[8] & Stephen Thibodeau[6] (*Group 3*)
Henrik Grönberg[9*], Monika Emanuelsson[9], Jan-Erik Damber[10], Anders Bergh[11] & Björn-Anders Jonsson[11] (*Group 4*)
Jeffrey Smith[12*], Joan Bailey-Wilson[12], John Carpten[12], Dietrich Stephan[12], Elizabeth Gillanders[12], Isaac Amundson[12], Tommi Kainu[12], Diana Freas-Lutz[12], Agnes Baffoe-Bonnie[13], Anne Van Aucken[12], Raman Sood[12], Francis Collins[12], Michael Brownstein[12] & Jeffrey Trent[12] (*Group 5*)
**These authors contributed equally to this work.*

Over 200,000 new prostate cancer cases are diagnosed in the United States each year, accounting for more than 35% of all cancer cases affecting men, and resulting in 40,000 deaths annually[1]. Attempts to characterize genes predisposing to prostate cancer have been hampered by a high phenocopy rate, the late age of onset of the disease and, in the absence of distinguishing clinical features, the inability to stratify patients into subgroups relative to suspected genetic locus heterogeneity. We previously performed a genome-wide search for hereditary prostate cancer (HPC) genes, finding evidence of a prostate cancer susceptibility locus on chromosome 1 (termed *HPC1*; ref. 2). Here we present evidence for the location of a second prostate cancer susceptibility gene, which by heterogeneity estimates accounts for approximately 16% of HPC cases. This HPC locus resides on the X chromosome (Xq27–28), a finding consistent with results of previous population-based studies suggesting an X-linked mode of HPC inheritance. Linkage to Xq27–28 was observed in a combined study population of 360 prostate cancer families collected at four independent sites in North America, Finland and Sweden. A maximum two-point lod score of 4.60 was observed at *DXS1113*, θ=0.26, in the combined data set. Parametric multipoint and non-parametric analyses provided results consistent with the two-point analysis. Significant evidence for genetic locus heterogeneity was observed, with similar estimates of the proportion of linked families in each separate family collection. Genetic mapping of the locus represents an important initial step in the identification of an X-linked gene implicated in the aetiology of HPC.

Despite the medical significance of prostate cancer in terms of morbidity, mortality and health-care costs, our understanding of the molecular determinants of prostate cancer susceptibility remains rudimentary. Epidemiological studies supporting the existence of hereditary forms of prostate cancer have led to the initiation of genome-wide searches for loci contributing to hereditary prostate cancer. A previous scan for linkage resulted in suggestive evidence (lod>1.0) for prostate cancer susceptibility loci on several chromosomes, including 1q, 4q, 5p, 7p, 13q and Xq (ref. 2). Statistically significant evidence was achieved only for the locus 1q24–25 (*HPC1*). Subsequent stratification of pedigrees showed that families linked to *HPC1* tended to have an early mean age of diagnosis (under 65 years) and a large number of affected members (>4). Even in this subset, this locus accounts for only approximately one-half of the families[3]. Further, although two confirmatory studies have corroborated linkage to *HPC1* (refs 4,5), three additional studies found no clear evidence for *HPC1*-predisposed disease in their study populations[6–8]. The disparity in these studies emphasizes the common set of obstacles for linkage detection in hereditary prostate cancer, most prominently, a high phenocopy rate and genetic locus heterogeneity.

Table 1 • Characteristics of prostate cancer families

	JHU	Mayo	Tampere	Umeå	All
Number of families	139	123	57	41	360
Number of individuals typed	766	407	548	268	1989
Number of affected individuals typed	452	314	137	117	1020
Avg. number of affected/family (range)	5.1 (3–17)	4.0 (3–11)	3.2 (2–9)	4.5 (3–10)	4.3 (2–17)
Avg. number of affected individuals typed/family (range)	3.2 (2–11)	2.6 (2–6)	2.4 (2–9)	2.8 (2–8)	2.7 (2–11)
Avg. age at diagnosis (range)	64.1 (39–85)	67.1 (41–93)	68.2 (45–90)	68.0 (46–86)	66.3 (39–93)

[1]Center for the Genetics of Asthma and Complex Diseases, University of Maryland, Baltimore, Maryland 21201, USA. Departments of [2]Urology, [3]Pathology and [4]Oncology, Johns Hopkins Medical Institutions, Baltimore, Maryland 21287, USA. [5]Laboratory of Cancer Genetics, Institute of Medical Technology, University of Tampere and Tampere University Hospital, Tampere, Finland. Departments of [6]Laboratory Medicine and Pathology, [7]Health Sciences Research and [8]Urology, Mayo Clinic/Foundation, Rochester, Minnesota 55902, USA. Departments of [9]Oncology, [10]Urology & Andrology and [11]Pathology, Umeå University, Umeå, Sweden. [12]Prostate Cancer Investigation Group, National Human Genome Research Institute, National Institute of Health, Bethesda, Maryland 20892, USA. [13]Population Science Division, Fox Chase Cancer Center, Philadelphia, Pennsylvania 19012, USA. Correspondence should be addressed to W.I. (e-mail: wisaacs@jhmi.edu).

Table 2 • Two-point parametric lod scores

Marker	Heterozygosity	cM[b]	lod (θ)[a] JHU (139)	Mayo (123)[c]	Tampere (57)	Umeä (41)[d]	All (360)
DXS984	0.74	140.0	0.40 (0.36)	0.31 (0.34)	0.87 (0.22)	0.03 (0.44)	1.00 (0.34)
DXS1232	0.66	140.9	0.28 (0.36)	0.00 (0.50)	0.66 (0.22)		0.24 (0.40)
DXS1205	0.66	142.3	0.19 (0.38)	0.00 (0.50)	2.05 (0.14)		0.33 (0.36)
DXS6751	0.74	143.6	0.49 (0.36)	0.52 (0.32)	1.56 (0.18)		1.63 (0.32)
DXS6798	0.83	144.8	0.51 (0.36)		0.78 (0.22)		0.87 (0.32)
DXS8106	0.70	146.1	0.82 (0.34)	0.80 (0.30)	0.89 (0.16)		1.93 (0.30)
DXS6806	0.81	147.3	0.45 (0.36)	0.78 (0.30)	0.14 (0.28)	0.03 (0.44)	1.07 (0.34)
DXS8043	0.83	148.8	0.97 (0.32)	0.02 (0.40)	0.00 (0.50)	0.08 (0.38)	0.74 (0.36)
AFMA113zf5	0.68	149.3	0.11 (0.36)	1.24 (0.28)	1.22 (0.18)		2.01 (0.28)
DXS1200	0.60	150.4	1.98 (0.28)	0.86 (0.26)	0.17 (0.32)	0.00 (0.50)	2.80 (0.30)
DXS297	0.70	151.0	0.64 (0.34)	0.18 (0.36)	0.13 (0.00)		0.84 (0.34)
AFM136yb10	0.68	152.5	1.00 (0.30)	0.40 (0.30)	0.05 (0.38)		1.38 (0.32)
DXS8091	0.80	152.5	1.52 (0.30)	0.28 (0.34)	0.00 (0.50)		1.43 (0.32)
DXS1113	0.80	153.0	1.73 (0.28)	1.89 (0.26)	0.49 (0.22)	0.60 (0.26)	4.60 (0.26)
DXS1193	0.78	153.3	0.96 (0.32)		0.58 (0.26)	0.34 (0.32)	1.80 (0.30)
DXS8069	0.67	154.5	0.44 (0.36)	0.84 (0.30)	0.01 (0.40)	0.12 (0.38)	1.20 (0.34)
DXS8011	0.87	154.6	0.32 (0.36)		0.58 (0.26)		0.72 (0.34)
DXS8103	0.77	155.2	0.10 (0.42)	0.38 (0.34)	0.92 (0.24)	0.29 (0.32)	1.10 (0.36)
AFMA225xh9	0.74	156.3	0.31 (0.36)	0.98 (0.30)	0.00 (0.50)		0.68 (0.36)
AFMA08xa5	0.51	157.1	0.02 (0.44)	0.02 (0.40)	0.09 (0.00)		0.03 (0.42)
DXS1108	0.70	158.8	0.12 (0.42)	0.57 (0.32)	0.00 (0.50)		0.42 (0.38)

[a]Maximum lod score under homogeneity with the maximum likelihood estimate of the recombination fraction (θ), calculated using FASTLINK. [b]Distance in cM from Xpter. [c]Three markers were not genotyped in this group. [d]Thirteen markers were not genotyped in this group.

A further confounding issue in prostate cancer linkage studies is the lack of a clear delineation of the mode(s) of inheritance. Segregation analyses of familial prostate cancer have supported an autosomal dominant mode of inheritance for prostate cancer susceptibility alleles[9–11], although formal testing of possible X chromosome segregation has not been performed. On the basis of studies of prostate cancer risk in relatives of affected men, it has been suggested that an HPC susceptibility locus may reside on the X chromosome. Several population-based studies have reported a statistically significant excess risk of prostate cancer in men with affected brothers, as compared with those with affected fathers, consistent with the hypothesis of an X-linked, or recessive, model of inheritance[12–16]. In our initial genome-wide search for prostate cancer linkage, there was suggestive evidence of linkage to the X chromosome[2]. These indications have prompted a more detailed analysis of potential X-linkage in HPC families.

To carry out this analysis, we have assembled 360 prostate cancer pedigrees consisting of families collected at sites in the US (Johns Hopkins University (JHU) in Baltimore, Maryland and the Mayo Clinic in Rochester, Minnesota), Finland (University of Tampere, Tampere) and Sweden (Umeå University, Umeå). Characteristics of the various family collections are given (Table 1). Overall, these 360 families contained 1,568 affected members. DNA samples, either from blood or archival tissue samples, were available from 1,020 affected individuals, and from an additional 969 individuals who were either female or unaffected. Over one-half of the families had at least one case of apparent male-to-male disease transmission. As it is possible that some of these occurrences result from a high phenocopy rate, the entire data set was analysed for possible evidence of X linkage.

The results from our previous 10-cM genome-wide screen using 66 North American prostate cancer families implicated a 40-cM interval from *DXS1001* to *DXS1108*, reaching a maximum two-point lod score of 1.08 at marker *DXS1193* at Xq27–28 (ref. 2). To more rigorously test the hypothesis of linkage to this region, an additional 28 markers were selected to augment the five original survey markers across the X chromosome interval. These markers were genotyped to create density map intervals of 1.2-cM for the 139 North American HPC families collected at JHU. A subset of 26 of these markers, spanning 19 cM from *DXS984* to *DXS1108* (140–159 cM from Xpter), were genotyped for the 123 Mayo Clinic and the 57 Finnish HPC families, and a less dense, 4-cM map of eight markers in this interval was completed for the 41 Swedish families. Allele frequencies were estimated from independent individuals in the complete data set. Two-point parametric lod scores are listed (Table 2). Twelve of the markers tested had lod scores greater than 1 in the combined data set, with a maximum score of 4.6 at marker *DXS1113*, θ=0.26. These results were supported by non-parametric affected sibpair analysis (Table 3). Fourteen consecutive markers had an excess mean identical-by-descent (IBD) sharing (0.55), with the lowest *P*-value of 0.00006 at *DXS1113*. The lod score, on the basis of sibpair IBD sharing,

Table 3 • Two-point affected sibpair analysis

	cM[a]	Mean IBD[b]	*P*-value[c]	lod
DXS984	140.0	0.54	0.08	0.42
DXS1232	140.9	0.51	0.33	0.04
DXS1205	142.3	0.53	0.15	0.24
DXS6751	143.6	0.56	0.005	1.41
DXS6798	144.8	0.55	0.047	0.60
DXS8106	146.1	0.57	0.005	1.43
DXS6806	147.3	0.55	0.039	0.67
DXS8043	148.8	0.55	0.023	0.86
AFMA113zf5	149.3	0.58	0.013	1.08
DXS1200	150.4	0.60	0.00008	3.11
DXS297	151.0	0.56	0.025	0.83
AFM136yb10	152.5	0.57	0.007	1.28
DXS8091	152.5	0.57	0.003	1.63
DXS1113	153.0	0.60	0.00006	3.20
DXS1193	153.3	0.57	0.006	1.37
DXS8069	154.5	0.55	0.048	0.60
DXS8011	154.6	0.55	0.04	0.65
DXS8103	155.2	0.52	0.16	0.20
AFMA225xh9	156.3	0.54	0.06	0.50
AFMA08xa5	157.1	0.52	0.32	0.05
DXS1108	158.8	0.52	0.21	0.14

[a]Distance in cM from Xpter. [b]Affected sibpair analyses were performed using ANALYZE. [c]All possible sibpairs were used in the analysis, however, a weight of (n–1) was given to the sibship of multiple sibs, where n is the number of sibs.

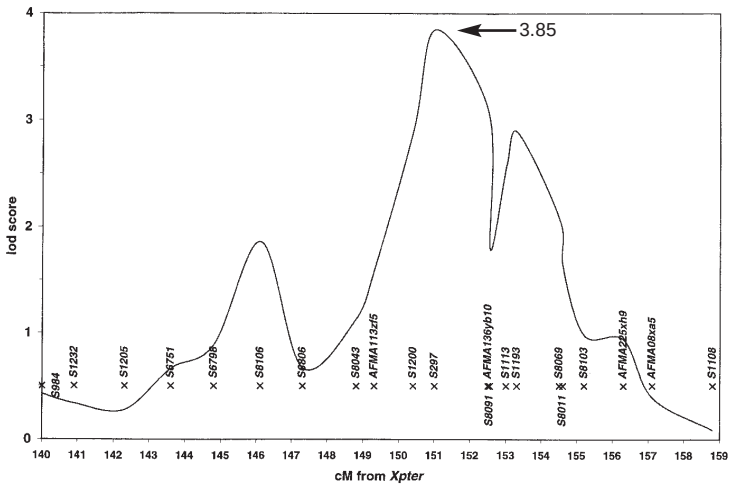

Fig. 1 Graph of multipoint lod scores assuming heterogeneity. The peak multipoint lod score of 3.85 is located between *DXS1200* and *DXS297*.

was 3.2 for this marker. When population-specific allele frequencies were used, similar results were obtained.

Simulation studies were performed to estimate the probability of obtaining a two-point parametric lod score of 4.6 or greater, or a *P*-value less than 0.00006 for non-parametric affected sibpair analysis (mean test), at a single marker on the X chromosome in the absence of linkage (false positive rate). Among 10,000 replicates in the simulation, there were no two-point parametric lod scores greater than 4.0, nor were there any *P*-values less than 0.00006 for affected sibpair analysis. There were three lod scores greater than 3, and only once was there a *P*-value less than 0.0001 among the 10,000 replicates.

Results from parametric multipoint linkage analyses were consistent with the two-point analyses. Data from the Swedish families were not included in the multipoint analysis, because only eight markers were genotyped in this dataset. Analysis was carried out using a sliding multipoint approach[17–19], and heterogeneity analysis was then performed using HOMOG (ref. 20). The maximum lod score assuming heterogeneity was 3.85, occurring 151 cM from Xpter, between loci *DXS1200* and *DXS297* (Fig. 1). Significant evidence for locus heterogeneity was obtained, with the proportion (α) of families linked estimated at 16% (χ^2=17.73, df=1, *P*=0.00002; Table 4).

Each study population had positive two-point and multipoint lod scores for multiple markers in the Xq27–28 region (Tables 2,4). Estimates of the proportions of linked families in each collection ranged from 15% (JHU) to 41% (Tampere), although the differences among groups are not statistically significant (χ^2=0.53, *P*=0.77).

As a possible source of genetic heterogeneity, we stratified families into two subsets on the basis of consistency with an X-linked mode of inheritance, using the apparent presence or absence of male-to-male transmission as a single, surrogate, stratification criterion. Following this stratification, 129 families without male-to-male transmission contribute disproportionately to the evidence of linkage to this region (maximum multipoint lod score assuming heterogeneity=2.46 at 151 cM from Xpter, estimated proportion linked=19%). In contrast, for families with male-to-male transmission (n=190), the maximum lod score assuming heterogeneity was 1.47, also at 151 cM, with a lower estimated proportion linked (13%). Although this difference is not statistically significant, the observed trend is consistent with the hypothesis of X chromosome linkage in this data set. The observation of positive lod scores in families with apparent male-to-male disease transmission may result from the presence of phenocopies as affected fathers or other relatives.

As there was evidence for linkage of HPC susceptibility loci to both 1q24–25 (ref. 2) and Xq27–28 in families collected at JHU, we tested the hypothesis (H_1) that there are three types of prostate cancer families in this cohort: (i) a proportion of the families linked to Xq27–28; (ii) a proportion of the remaining families linked to 1q24–25; and (iii) the rest linked to neither region. Using the admixture test[20] (HOMOG3R) with multipoint lod score data for the 139 families in this group, significant evidence of locus heterogeneity was observed (Table 5). The data were made at least 360-fold more probable given the hypothesis (H_1) that subsets of HPC families are linked to Xq27–28 or to *HPC1*, and the remainder unlinked, than the hypotheses of either as a sole locus (H_2 or H_3). Multipoint data suggested that 15% of the families in this group were linked to the X chromosome locus, and that 30% were linked to *HPC1*. Similarly, in the 59 families in this collection that are not linked to *HPC1* (lod<–0.1), the multipoint lod score under heterogeneity is 1.96 for Xq27–28, whereas the lod score is 0.48 in the remaining 80 families.

Linkage analysis is valuable for identification of genetic loci predisposing to prostate cancer. The presence of genetic heterogeneity both in and across populations necessitates large-scale studies to provide significant statistical power to identify major loci. Among the JHU study population, loci at 1q24–25 and Xq27–28 are estimated to account for approximately 30% and 15% of the prostate cancer families, respectively. In contrast, of these two loci, only the X-chromosome locus appears to have a prominent role in prostate cancer predisposition in the Finnish study population, in which a larger fraction of families (over 40%) are estimated to be X-linked, and *HPC1* shows only a marginal role (J. Schleutker *et al.*, in preparation). A similar situation exists in the Mayo Clinic data set, although the proportion of families linked to the X chromosome is the same as in the JHU study population. From these results, we anticipate that confirmatory studies will also encounter genetic heterogeneity. Indeed, a recently described factor contributing to the lack of linkage to *HPC1* in several family collections may be the presence of an increased proportion of X-linked pedigrees in these cohorts. Similarly, linkage to the X chromosome may be most readily apparent upon stratification of pedigrees by male-to-male disease transmission in these populations, although, as we have seen, evidence for this linkage is not restricted to particular subsets of this stratification. Further, as the major proportion of the families examined in this study are not linked to either *HPC1* or the X-chromosome locus, and as collection of additional study

Table 4 • Heterogeneity test using multipoint lod score for each family collection

		MLE estimates (3-unit support interval)		
Group	lod[a]	α	map position[b]	*P*-value
JHU	2.34	0.15 (0.03–0.30)	152.5 (140.0–154.6)	0.001
Mayo	1.03	0.16 (0.01–0.34)	154.5 (140.0–158.8)	0.029
Tampere	2.03	0.41 (0.08–0.71)	143.6 (140.0–151.0)	0.002
All	3.85	0.16 (0.06–0.26)	151.0 (140.0–153.3)	0.00002

[a]Heterogeneity test was based on sliding multipoint lod scores, using the admixture test (HOMOG), where lod is calculated assuming heterogeneity. [b]Distance in cM from Xpter.

Table 5 • Admixture test using multipoint lod scores for Xq27–28 and 1q24–25 (139 JHU families)

Hypothesis[a]	% linked to 1q24–25[b] α1	% linked to Xq27–28[c] α2	ln L	χ^2 (df)[d]	*P*-value
H1	0.30	0.15	16.43		
H2	0.29	[0]	10.52	11.82 (2)	0.0027
H3	[0]	0.16	5.42	22.02 (2)	1.6×10^{-5}
H4	[0]	[0]	0	32.86 (5)	4.0×10^{-7}

[a]Heterogeneity test was performed using the admixture test (HOMOG3R). [b]Multipoint lod scores at 1q24–25 were based on markers *D1S158* and *D1S422*. [c]Multipoint lod score at Xq27–28 were based on markers AFMA113zf5, *DXS1200* and *DXS297*. [d]χ^2 is –2ln likelihood difference between H1 and each alternative hypotheses.

populations increases the statistical power, additional loci may be proven to account for a portion of prostate cancer predisposition. In this regard, a recent study of 47 French and German families had a multipoint lod score, assuming heterogeneity, of 2.2 (α=50%) and two-point score of 2.7 at 1q42.2–43 (ref. 8).

Significance of the results achieved here is supported by several different lines of evidence. Most importantly, the linkage data derived from each of four independent family collections provides evidence of linkage to Xq27–28. When combined, this data set yields a maximum two-point lod score of 4.6, meeting the proposed criteria for significant linkage[21]. Second, non-parametric methods supported this result and provided a model-independent significance level of *P*=0.00006 for linkage. Third, simulations performed to provide an empirical nominal significance level for the observed linkage results never yielded a two-point lod score greater than 4.0, nor any *P*-value less than 0.00006 in 10,000 replicates. The data support the newly identified locus as predisposing to hereditary prostate cancer at Xq27–28.

A candidate prostate cancer susceptibility gene residing on the X chromosome is the androgen receptor gene[22–25] (*AR*). *AR*, however, is located at Xq12, over 50 cM from the region implicated in this study. Furthermore, direct assessment of linkage to *AR* in the North American families studied here provides no evidence of linkage (unpublished observations). Several genes at Xq27–28 have been mapped (http://www.ncbi.nlm.nih.gov/genemap), and these and other novel genes in the Xq27–28 region will need to be evaluated as candidate prostate cancer susceptibility genes.

We have presented evidence for linkage of a significant subset of prostate cancer families to a locus on Xq27–28. Contingent upon confirmation, we suggest the designation *HPCX* for this locus.

Methods

North American families. Johns Hopkins family collection: The 79 North American families that were described in the report of linkage to *HPC1* (ref. 2) are included in this study, as are an additional 60 pedigrees collected at the Brady Urologic Institute at Johns Hopkins. A majority of these families were ascertained through referrals from physicians; some families were recruited from earlier epidemiological studies[9] and through news articles. Age of diagnosis of prostate cancer was confirmed either through medical records or from two other independent sources. All individuals in this study gave full informed consent.

Mayo Clinic family collection: The 123 North American families in this collection were ascertained by a cancer family-history survey, sent to over 5,000 men who underwent a radical prostatectomy for clinically localized prostate cancer in the Department of Urology at the Mayo Clinic during 1966–1995 (ref. 11). Prostate cancer diagnosis and the age of onset was confirmed through medical records at the Mayo Clinic and elsewhere. All participants in this study gave full informed consent.

Finnish families. In Finland, 302 prostate cancer families with two or more affected cases were identified through referrals from physicians, family questionnaires sent to patients, a nationwide registry-based search and advertisements in newspapers, radio and television. Of this group, 57 families that were informative for linkage analyses were included in this study. Diagnosis of all prostate cancer patients was confirmed through hospital records or from the Finnish cancer registry. All individuals participating in this study gave full informed consent.

Swedish families. Since 1995, families with three or more relatives affected with prostate cancer have been collected at the Department of Oncology of Umeå University, mainly from referrals from urologists throughout Sweden. From approximately 300 referrals, 41 families informative for linkage analysis have been selected. Twelve of these families were included in an earlier report[2]. When blood samples were unavailable, tissue samples were collected from affected men whenever possible. Tissue samples were reviewed by an experienced pathologist and microdissection was performed to separate normal and tumour tissue. For genotyping, only normal tissue was used. All prostate cancer diagnoses in the families were confirmed by the National cancer registry and medical records.

Genotyping methods. Techniques of preparing DNA and genotyping were as described[2]. Markers were derived from the Genome Database (Johns Hopkins University School of Medicine). Marker data was obtained for 33 polymorphic loci available in the GDB, spanning the approximately 40-cM interval between *DXS1001* and *DXS1108*. Order and distance for these markers was estimated from the entire genotype data set using CRIMAP (ref. 26). The most likely order thus determined agrees with the published order[27]. Allele frequencies were estimated from genotypes of independent individuals in the 360 families.

Statistical methods. Both parametric and non-parametric linkage approaches were used in this study. The parametric analysis used a previous model[2,3] with regard to disease allele frequency (0.003) and age-specific penetrances, although affected men were assumed to be carriers of an X-linked, sex-limited, dominant gene. A fixed 15% phenocopy rate, that is, P (non-predisposing genotype/disease), was assumed, whereas all unaffected men under 75, and all women, were assumed to be of unknown phenotype. In men over age 75, the lifetime penetrance of gene carriers was estimated to be 63%, and the lifetime risk of prostate cancer for a non-carrier was 16% in this age class. FASTLINK (refs 18,19) and ANALYZE (ftp://linkage.cpmc.columbia.edu/software/analyze) were used for the parametric two-point analysis. For the non-parametric analysis, affected sibpairs were used for the two-point analysis as implemented by ANALYZE, using the mean test and likelihood based test. The mean test compares the number of alleles shared IBD with the number of alleles not shared IBD among affected sibpairs. When there are multiple sibs in a sibship, a weight of (n–1) is given to the sibship, where n is the number of sibs. When parents are not genotyped, the program computes the likelihood of each possible genotype for the parents, and computes the number of alleles shared IBD in a sibpair as the average over all possible parental genotype combinations, weighted by their conditional probabilities given the known data.

The simulation study was performed using FASTSLINK (ftp://watson.hgen.pitt.edu/pub). A 10-allele marker, which represents the marker *DXS1113*, was simulated unlinked to the disease locus using the exact pedigree structure and availability of genotype information for the 360 families analysed. The marker *DXS1113* has 15 alleles, six of which have frequencies of approximately 1% or less. To make the simulation of a large number of replicates (10,000) more practical, we collapsed the six less frequent alleles into one allele.

The multipoint approach is critical in linkage analysis of a late age-of-onset disease such as prostate cancer, because parental genotypic data are often missing, making inference of IBD ambiguous. Additionally, multipoint analysis is more robust to misspecification of allele frequencies and statistical fluctuations at individual loci. When more markers are used simultaneously in the analysis (multipoint analysis), the probability distribution is concentrated on certain inheritance vectors, thus the determination of IBD is less dependent on the marker allele frequencies[28]. However, multipoint analyses of X-chromosome marker data are hampered by the lack of fully functional X-chromosome versions of the most appropriate multipoint analysis computer programs (for example, GENEHUNTER).

In this study, the parametric multipoint analysis was performed using FASTLINK (LINKMAP; refs 18,19). Due to computer memory constraint, only 4-point analyses (disease locus against three marker loci) were performed. A sliding multipoint approach was used as described[17]. Briefly, this approach consists of sliding a group of three loci down the map and analysing the disease locus only in the interval between the second and third marker. Heterogeneity analysis was then performed using HOMOG (ref. 20).

The admixture model was used to test several hypotheses for genetic locus heterogeneity (HOMOG3R; ref. 20). α_1 is the proportion of families linked to the first disease locus (that is, 1q24–25), and α_2 is the proportion linked to the second disease locus (that is, Xq27–28). Hypothesis 1 (H_1) assumes that there are three types of families in the sample, (α_1, α_2 and $1-(\alpha_1+\alpha_2)$). Hypothesis 2 (H_2) assumes that there are two types of families, α_1 and $1-\alpha_1$. Hypothesis 3 (H_3) assumes that there are two types of families, α_2 and $1-\alpha_2$. Hypothesis 4 (H_4) assumes no linkage to either disease locus ($\alpha_1=\alpha_2=0$). Maximum likelihood for each of these hypotheses was calculated from the data. Chi-square (χ^2) tests were performed by calculating twice the difference of the natural log likelihood between two hypotheses, with the degrees of freedom (df) equal to the difference in the number of parameters estimated for the two hypotheses. The asymptotic null distribution of the test statistic has not been well investigated, but this approach is conservative[20].

Stratification of families. The criteria used to categorize a family as having evidence of male-to-male disease transmission were as follows: (i) presence of affected father and affected son(s) combinations, or (ii) prostate cancer case(s) on the paternal side of the family, with no evidence of affected relatives on maternal side. Families that did not meet these criteria were classified as families without evidence of male-to-male transmission.

Acknowledgements

The authors thank all the family members who participated in this study. In addition, the assistance of physicians who referred patients to this study is gratefully acknowledged. The authors wish to thank K. Vaherto, M. Kuorilehto, R. Sankila, E. Pukkala and J. Viitanen for assistance in family collection in Finland. This work was supported by the National Institutes of Health (SPORE CA58236, CA72818 and NO1-HG-55389); The Fund for Research and Progress in Urology, The Johns Hopkins University; Academy of Finland; Reino Lahtikari and Sigrid Juselius Foundations; Finnish Cancer Society; Medical Research Fund of Tampere University Hospital; Swedish Cancer Society (Cancerfonden); Lion's Cancer Foundation, Department of Oncology, Umeå University; and CaPCURE. S.I. is supported by the GE Fund. D.F. is supported by the American Foundation for Urologic Disease.

Received 3 June; accepted 1 September, 1998.

1. Landis, S.H., Murray, T., Bolden, S. & Wingo, P.A. Cancer statistics, 1998. *CA Cancer J. Clin*. **48**, 6–29 (1998).
2. Smith, J.R. *et al*. Major susceptibility locus for prostate cancer on chromosome 1 suggested by a genome-wide search. *Science* **274**, 1371–1374 (1996).
3. Grönberg, H. *et al*. Early age at diagnosis in families proving evidence of linkage to the hereditary prostate cancer locus *(HPC1)* on chromosome 1. *Cancer Res.* **57**, 4707–4709 (1997).
4. Cooney, K.A. *et al*. Prostate cancer susceptibility locus on chromosome 1q: a confirmatory study. *J. Natl Cancer Inst*. **89**, 955–959 (1997).
5. Hsieh, C.L. *et al*. Prostate cancer susceptibility locus on chromosome 1q: a confirmatory study. *J. Natl Cancer Inst*. **89**, 1893–1894 (1997).
6. McIndoe, R.A. *et al*. Linkage analysis of 49 high-risk families does not support a common familial prostate cancer-susceptibility gene at 1q24–25. *Am. J. Hum. Genet*. **61**, 347–353 (1997).
7. Eeles, R.A. *et al*. Linkage analysis of chromosome 1q markers in 136 prostate cancer families. *Am. J. Hum. Genet*. **62**, 653–658 (1998).
8. Berthon, P. *et al*. Predisposing gene for early-onset prostate cancer, localized on chromosome 1q42.2–43. *Am. J. Hum. Genet*. **62**,1416–1424 (1998).
9. Carter, B.S., Beaty, T.H., Steinberg, G.D., Childs, B. & Walsh P.C. Mendelian inheritance of familial prostate cancer. *Proc. Natl Acad. Sci. USA* **89**, 3367–3371 (1992).
10. Grönberg, H., Damber, L., Damber, J.E. & Iselius, L. Segregation analysis of prostate cancer in Sweden: support for dominant inheritance. *Am. J. Epidemiol*. **146**, 552–557 (1997).
11. Schaid, D.J., McDonnell, S.K., Blute, M.L. & Thibodeau, S.N. Evidence for autosomal dominant inheritance of prostate cancer. *Am. J. Hum Genet*. **62**, 1425–1438 (1998).
12. Woolf, C.M. An investigation of the familial aspects of carcinoma of the prostate. *Cancer* **13**, 361–364 (1960).
13. Narod, S.A. *et al*. The impact of family history on early detection of prostate cancer. *Nature Med.* **1**, 99–101 (1995).
14. Whittemore, A.S. *et al*. Family history and prostate cancer risk in black, white, and Asian men in the United States and Canada. *Am. J. Epidemiol*. **141**, 732–740 (1997).
15. Monroe, K.R. *et al*. Evidence of an X-linked or recessive genetic component to prostate cancer risk. *Nature Med.* **1**, 827–829 (1995).
16. Hayes, R.B. *et al*. Prostate cancer risk in U.S. blacks and whites with a family history of cancer. *Int. J. Cancer* **60**, 361–364 (1995).
17. Terwilliger, J.D. & Ott, J. *Handbook of Human Genetic Linkage*. (Johns Hopkins University Press, Baltimore, 1994).
18. Cottingham Jr, R.W., Idury, R.M. & Schaffer, A.A. Faster sequential genetic linkage computations. *Am. J. Hum. Genet*. **53**, 252–263 (1993).
19. Lathrop, G.M., Lalouel, J.M., Julier, C. & Ott, J. Strategies for multilocus linkage analysis in humans. *Proc. Natl Acad. Sci. USA* **81**, 3443–3446 (1984).
20. Ott, J. *Analysis of Human Genetic Linkage* (Johns Hopkins University Press, Baltimore, 1991).
21. Lander, E.S. & Kruglayk, L. Genetic dissection of complex traits: guideline for interpreting and reporting linkage results. *Nature Genet*. **11**, 241–247 (1995).
22. Stanford, J.L. *et al.* Polymorphic repeats in the androgen receptor gene: molecular markers of prostate cancer risk. *Cancer Res.* **57**, 1194–1198 (1997).
23. Hakimi, J.M., Rondinelli, R.H., Schoenberg, M.P. & Barrack, E.R. Androgen-receptor gene structure and function in prostate cancer. *World J. Urol.* **14**, 329–337 (1996).
24. Giovannucci, E. *et al.* The CAG repeat within the androgen receptor gene and its relationship to prostate cancer. *Proc. Natl Acad. Sci. USA* **94**, 3320–3325 (1997).
25. Irvine, R.A., Yu, M.C., Ross, R.K. & Coetzee, G.A. The CAG and GGC microsatellites of the androgen receptor gene are in linkage disequilibrium in men with prostate cancer. *Cancer Res.* **55**, 1937–1940 (1995).
26. Lander, E.S. & Green, P. Construction of multilocus genetic linkage maps in humans. *Proc. Natl Acad. Sci. USA* **84**, 2363–2367 (1987).
27. Collins, A., Teague, J., Keats, B.J. & Morton, N.E. Linkage map integration. *Genomics* **36**, 157–162 (1996).
28. Kruglayk, L., Daly, M.J., Reeve-Daly, M.P. & Lander, E.S. Parametric and nonparametric linkage analysis: A unified multipoint approach. *Am. J. Hum. Genet*. **58**, 1347–1363 (1996).