

ARTICLE

A genomewide linkage analysis for prostate cancer susceptibility genes in families from Germany

Christiane Maier¹, Kathleen Herkommer², Josef Hoegel¹, Walther Vogel¹ and Thomas Paiss^{*,2}

¹Abteilung Humangenetik, Universität Ulm, Germany; ²Urologische Universitätsklinik und Poliklinik, Abteilung für Urologie und Kinderurologie, Ulm, Germany

Prostate cancer is a complex disease with a substantial genetic contribution involved in the disease risk. Several genomewide linkage studies conducted so far have demonstrated a strong heterogeneity of susceptibility. In order to assess candidate regions that are particularly relevant for the German population, we performed a genomewide linkage search on 139 prostate cancer families. A nonparametric method (Z_{lr} scores), using GENEHUNTERPLUS, was applied at 500 markers (panel P1400, deCODE), with an average spacing of 7.25 cM. In the entire family collection, linkage was most evident at 8p22 ($Z_{lr} = 2.47$, $P = 0.0068$), close to the previously identified susceptibility gene *MSR1*. Further local maxima with $Z_{lr} > 2$ ($P < 0.025$) were observed at 1q, 5q and 15q. In a subgroup of 47 families, which matched the Johns Hopkins criteria of hereditary prostate cancer, suggestive linkage was found on 1p31 ($Z_{lr} = 3.37$, $P = 0.00038$), a previously not described candidate region. The remaining 92 pedigrees, with no strong disease history, revealed a maximum $Z_{lr} = 3.15$ ($P = 0.00082$) at 8q13, possibly indicating a gene with reduced penetrance or recessive inheritance. Our results suggest pronounced locus heterogeneity of prostate cancer susceptibility in Germany. In the present study population, the *MSR1* gene could play a significant role. Other conspicuous loci, like 1p31 and 8q13, need further investigation in order to verify their relevance and to identify candidate genes.

European Journal of Human Genetics (2005) 13, 352–360. doi:10.1038/sj.ejhg.5201333

Published online 10 November 2004

Keywords: prostate cancer; genomewide linkage; hereditary; familial; *MSR1*

Introduction

Prostate cancer is a complex disease with a substantial genetic contribution involved in the disease risk.¹ Mapping of the underlying genes has been attempted by different investigators from the United States and Europe, using genomewide linkage analyses in families with a

clustering of prostate cancer (for reviews, see Easton *et al*² and Schaid³). To date, these efforts have led to the identification of three susceptibility genes, *RNASEL* at 1q25,^{4,5} *ELAC2* at 17p11⁶ and *MSR1* at 8p22;^{7,8} each has been found to be mutated in a certain proportion of the family samples with particular evidence for linkage. Aside from these genes, several other loci, including 1q42–43 (*PCAP*),⁹ 1p36 (*CAPB*),¹⁰ 20q13 (*HPC20*)¹¹ and Xq27 (*HPCX*),¹² are believed to harbour additional prostate cancer genes. Most of these loci showed significant evidence of linkage in an individual scan, but were confirmed only by single studies over many attempts. A recent series of genomewide scans, performed by the

*Correspondence: Dr T Paiss, Urologische Universitätsklinik und Poliklinik, Abteilung für Urologie und Kinderurologie, Universitätsklinikum Ulm, Prittwitzstrasse 43, D-89075 Ulm, Germany. Tel: +49 731 50 02 78 07; Fax: +49 731 50 02 78 56; E-mail: thomas.paiss@medizin.uni-ulm.de
Received 18 June 2004; revised 22 September 2004; accepted 6 October 2004

International Consortium of Prostate Cancer Genetics (ICPCG), again demonstrated a lack of consensus regions for prostate cancer susceptibility.² The presence of multiple causative genes (locus heterogeneity) has been assumed to be the major cause of inconsistent results from scans performed on different family samples. It is yet unknown to what extent locus heterogeneity is correlated with the ethnic and demographic properties of the population used for linkage analyses. If samples of different origins had different population-specific frequencies of mutated genes, the evidence of linkage at a particular susceptibility locus would be inconsistent in these different samples. However, if a large number of genes were involved, the family samples might be too small to represent the whole population. Thus, the admixture of susceptibility genes in different samples analysed would be random, in which case divergent linkage results would be due to chance and need not be explained by an origin-dependant locus heterogeneity.¹³

Assuming that heterogeneity is related to demographic factors, every population must be screened individually in order to assess the particular impact of susceptibility loci and to identify the population-specific genes. We have addressed this issue for Germany by performing a genome-wide linkage study in prostate cancer families from all over the country. One advantage of our family sample may be that it was drawn from a European population, which is believed to be less heterogeneous than those of North America. Since few European populations have been investigated to date, our study may contribute to further insight into prostate cancer genetics.

In contrast to most of the family collections previously analysed, our study sample is not limited to extended pedigrees with a strong clustering of prostate cancer. Thereby, we avoided a pre-selection for genes according to a specific mode of inheritance. It is well proven that a significant portion of causal variants are consistent with a dominant and highly penetrant model.^{14–16} Due to their strong disease effect, such high-risk susceptibility genes may be represented best by pedigrees with a strong clustering of prostate cancer. The Johns Hopkins criteria of hereditary prostate cancer¹⁷ provided a standard definition of families with a strong disease history and has been applied in the majority of genomewide linkage analyses. However, recent segregation analyses have also yielded evidence for recessive, X-chromosomal and interacting prostate cancer genes,^{18,19} which may not prevail in families consistent with the definition for 'hereditary prostate cancer'. In consequence, pedigrees exhibiting a moderate disease history, usually designated 'familial prostate cancer', may be genetically determined as well. We have therefore investigated hereditary and familial prostate cancer families in order to access the full range of potentially causal genes.

Materials and methods

Subjects

All individuals described in this report are participants in the Prostate Cancer Genetics Project of the University of Ulm. Urologists from all over Germany were asked to inform all patients with prostate cancer about our project and to motivate them to contact our institution. Although we did not apply any selection criteria to the primary sampling of these prostate cancer patients, most of them were contributed by Urological Surgical Departments from Southern Germany, and the majority were treated by radical prostatectomy. The probands' self-reported family history of prostate cancer has been used to identify prostate cancer families. A detailed family questionnaire, sent out additionally, was used as a guide for the recruitment of affected and relevant unaffected relatives. In all cases, the diagnosis of prostate cancer was confirmed by a histopathological report or by any other suitable medical record. The families used in this study were required to have at least two relatives with confirmed prostate cancer from which blood samples could be made available. All of the 146 families initially selected were Caucasian. In approximately one-third of patients, prostate cancer was diagnosed from their symptoms. In two-thirds of patients, the disease was detected by PSA screening. Informed consent, according to the Institutional Review Board at the University of Ulm, was mandatory.

Description and stratification of families

Seven of the 146 families genotyped had to be excluded due to a questionable pedigree structure. In the remaining 139 families, the number of affected men per pedigree ranged from 2 to 7 (average 2.9). On average, 2.2 (range 2–6) affected and 0.8 (0–7) unaffected individuals per family were genotyped. Table 1 shows the clinical characteristics of the affected family members. The families that contribute to subgroup analyses are described in Table 2. The selection criteria for these subsets were based on data from previously published linkage analyses.

We determined whether or not our families met the Johns Hopkins criteria for hereditary prostate cancer¹⁷ as follows: (i) prostate cancer in at least three first-degree relatives, or (ii) prostate cancer in three consecutive generations, or (iii) prostate cancer in two first-degree relatives with an age of onset <55 years. A total of 47 families meeting these criteria were classified as families with hereditary prostate cancer (HPC). The other 92 families were classified as families with familial prostate cancer (FPC).

The families were stratified according to the mean age at prostate cancer diagnosis of all affected individuals, and classified as either 'early-onset family' with a mean age of onset <65 years ($n=80$) or 'late-onset family' with a mean age of onset ≥ 65 years ($n=59$). Families were also divided according to whether or not there was evidence of

male-to-male disease transmission. Referring to Xu *et al*,¹² male-to-male disease transmission (mtm) was assumed in 46 families, where (i) a combination of an affected father and affected son(s) was present, or where (ii) prostate cancer on the paternal side with no affected relatives on the maternal side of the family was present. The 93 families not meeting these criteria were classified as families without evidence of male-to-male transmission (nm).

Genotyping and quality check

Genomic DNA was extracted from peripheral blood by standard procedures. PCR, allele calling and a first-quality check were performed by in-house techniques at deCODE (Reykjavik, Iceland), for 501 microsatellite sites from the marker panel P1400. The mean distance between markers was 7.25 cM (SD = 4.2 cM). Overall, 95.4% of the genotypes were successfully determined. We re-analysed the genotype data concerning Mendelian inheritance and relationship patterns of our families using the GRR and RELPAIRS programmes. As a consequence, we discarded seven out of the initial 146 families (4.8%) due to unclear pedigree

structures. Furthermore, one marker (D6S274) was excluded because of inconsistent genotype data in several, otherwise inconspicuous pedigrees.

Statistical analysis

The principal linkage analyses were based upon model-free multipoint approaches. Kong and Cox statistics²⁰ (Z_{lr} scores) implemented in GENEHUNTERPLUS²¹ were computed under the 'all' scoring function and the exponential model option of the ASM extension. Families were given equal weights in all analyses. To assess statistical significance, one-sided P -values were calculated using the standard normal approximation to the exact distribution. To be quoted a Z_{lr} score must exceed 2.0, which corresponds to a one-sided local level of significance slightly less than $P < 0.025$.

Comparisons of linkage between subgroups of families (eg, families that meet and do not meet the criteria for hereditary prostate cancer) were performed by applying an equal-variance T statistic, based on NPL scores which is given in Mirea *et al*.²² Approximate P -values were calculated using the adequate t -distribution.

In addition, multipoint parametric LOD scores were determined under a dominant and recessive model of inheritance. For both models, we adopted an age-specific scheme of six liability classes established by the International Consortium of Prostate Cancer Genetics (ICPCG; for details see Xu *et al*.²³). Penetrances were as follows, for the unsusceptible/susceptible genotype, respectively: 0.00038/0.0018 for affected men with age diagnosis ≤ 49 years; 0.00061/0.0084 for age 50–59; 0.0032/0.03 for age 60–69; 0.0082/0.04 for age 70–79 and 0.0086/0.015 for age ≥ 80 . The sixth liability class included unaffected men and women, which were coded as unknown. The disease allele frequency was specified as 0.003 under the dominant model or as 0.15 under the recessive model, according to estimations of previous segregation analyses.^{14,16} Marker allele frequencies were calculated once from the pool of all genotyped family members, regardless of their relationship. These frequencies were considered in all subgroup analyses as surrogates of the corresponding unknown

Table 1 Clinical description of prostate cancer patients

	All cases	Genotyped cases
Age at diagnosis, mean (range)	64.6 (47–89)	63.4 (47–83)
PSA value (ng/ml)	$n = 248^a$	$n = 235^b$
Median (range)	10.0 (0.2–2750)	10.0 (0.2–2750)
Tumour grade	$n = 310^a$	$n = 284^b$
GI	13.5%	14.4%
GI–II/GII	64.5%	64.4%
GII–III/GIII	21.9%	21.1%
Tumour stage	$n = 318^a$	$n = 291^b$
cT2/pT2	59.4%	59.5%
cT3/pT3	35.5%	35.4%
cT4/pT4	5.0%	5.2%

^aPatients with data available; out of a total number of 405 cases.

^bPatients with data available; out of a total number of 309 genotyped cases.

Table 2 Number of families and individuals contributing to subgroup analyses

Family subgroup	No. of families	Total	Number of individuals (per family)	
			Cases	Unaffected Genotyped
All families	139	405 (2.9)	309 (2.2)	116 (0.8)
Hereditary prostate cancer (HPC)	47	174 (3.7)	119 (2.5)	40 (0.9)
Familial prostate cancer (FPC)	92	231 (2.5)	190 (2.1)	76 (0.8)
Early onset, mean < 65 years	80	219 (2.7)	176 (2.2)	71 (0.9)
Late onset, mean ≥ 65 years	59	186 (3.2)	133 (2.3)	45 (0.8)
Male to male transmission (mtm)	46	172 (3.7)	105 (2.3)	53 (1.2)
No male to male transmission (nm)	93	233 (2.5)	204 (2.2)	63 (0.7)

population frequencies. To cope with the phenomenon of linkage heterogeneity between different families, heterogeneity LOD scores along with the proportion of linked families (α) were computed using the GENEHUNTERPLUS software.

All analyses were carried out according to the 'affected-only' principle, that is, all unaffected family members were given unknown affection status. *P*-values reported were not adjusted for multiple testing. As a threshold for 'suggestive linkage', the *P*-value 0.00074 was used,²⁴ which corresponds to a Z_{lr} score of about 3.18. However, this relates to a single genomewide analysis and does not take into account all the subgroup analyses carried out in addition.

The power to detect true linkage was evaluated by simulation analyses using the program SLINK.²⁵ Under the dominant model described above, two-point LOD scores were calculated for a representative six-allelic marker in 200 replicates. With respect to the properties of our family collection, we examined the influence of the distance between marker and disease gene (recombination fraction, θ), the number of families under study and the percentage of linked pedigrees under the assumption of heterogeneity.

Results

The characteristics of all 139 families and the subsets considered are described in Table 2. Figure 1 summarizes the nonparametric Z_{lr} scores at 500 markers across the genome of the total family set without stratification, and of the HPC and FPC subgroups. A Z_{lr} score of 4.1 ($P=0.000022$), which is required for significant linkage at a genomewide level, was not observed at any marker position.

Regions exhibiting $Z_{lr}>2$ in all families or in the HPC subset are summarized in Table 3. In the total set of families, linkage was most evident to 8p22 ($Z_{lr}=2.47$ at D8S1827), to 5q34–35 ($Z_{lr}=2.44$ at D5S2030) and to a new locus on 1q ($Z_{lr}=2.43$ at D1S189). Additional hints on linkage were found on 1p31, 5q14, 6q and on chromosome 15. Most of these loci were supported by both HPC and FPC pedigrees (Table 3). The linkage peak on 1p31 ($Z_{lr}=2.01$ at D1S438), however, was mainly contributed by families with hereditary disease, showing a Z_{lr} score of 3.37 at marker D1S438 on 1p31. This finding exceeded the threshold of $P=0.00074$ ($Z_{lr}=3.18$), which is accepted as a suggestive linkage result in a genomewide search. Further loci that became apparent in the HPC subgroup included 12q and 13q, as well as the telomeric region of 19p. Assuming a rare dominant susceptibility gene, multipoint parametric analysis showed positive LOD scores at all loci that had been considered as regions of interest after the primary analysis. LOD scores >1 were

seen for 1p31 (LOD=1.51 at D1S439 in HPC), 12q (LOD=1.29 at D12S351 in FPC) and to the telomeric region of 5q in the entire family set. On 5q, an interval of approximately 32 cM showed conspicuous signals, starting at marker D5S2040 (LOD=1.03) to our most telomeric marker D5S2030 (LOD=2.00). The maximum LOD of 2.09 was found between the marker sites D5S498 and D5S2030.

In the subset of FPC pedigrees, Z_{lr} scores greater than 2 were found on 8q, 9p, 9q and 11q. The results for each region of interest are shown in Table 4, including parametric LOD scores calculated for the dominant and the recessive model. Overall, the highest peak ($Z_{lr}=3.15$, $P=0.00082$) was seen at marker D8S543 on the long arm of chromosome 8. Linkage to this site was not supported by the subset of HPC families. Using the equal-variance *T* statistic,²² the difference between the Z_{lr} scores of the FPC and the HPC was highly significant ($P=0.0007$). Parametric analyses at D8S543 revealed a slightly higher evidence of linkage under the recessive (LOD=2.37) than under the dominant model (LOD=1.73).

The families were also evaluated for linkage after stratification according to the average age of diagnosis. The most prominent finding was a $Z_{lr}=3.39$ ($P=0.00035$) on the long arm of chromosome 5 (D5S2501) in late-onset families. This linkage to late-onset disease was evident over a 43 cM large region on 5q, ranging from 5q14 ($Z_{lr}=2.89$ at D5S428) to 5q31 ($Z_{lr}=2.91$ at D5S2017). Marker D5S428 already peaked up in the entire family set ($Z_{lr}=2.04$, Table 3), and was more strongly linked to late-onset disease. In a similar way, evidence of linkage to 8p22 was more obvious in the late-onset group ($Z_{lr}=2.99$ at D8S1827) than in the total sample ($Z_{lr}=2.47$, Table 3). Chromosome 6q25 (D6S305) was a further region of interest in the late-onset group ($Z_{lr}=2.72$). This site is 14 cM telomeric to marker D6S503, which peaked up ($Z_{lr}=2.07$) in the primary analysis. Peaks in early-onset families were seen on 1q12 ($Z_{lr}=2.59$ at D1S189) and 5q34 ($Z_{lr}=2.52$ at D5S2030), which were also regions of interest in the primary analysis.

Table 5 summarizes our findings at regions previously implicated in hereditary prostate cancer. The analysis of the whole family set without stratification and the best result from the subset analyses are presented. The most pronounced locus was 8p22, where the presumptive susceptibility gene *MSR1* had previously been identified. This peak, as described above, was seen at marker D8S1827 ($Z_{lr}=2.47$ in all families; $Z_{lr}=2.99$ in the late-onset group) and is located approximately 3 Mb telomeric to *MSR1*. Parametric linkage analysis, under the assumption of heterogeneity, indicated that 56% of our families in general, and almost all (99%) of the pedigrees with late-onset disease, might be linked to *MSR1*. We also noticed some evidence for linkage to 17p11 (*ELAC2*), particularly in HPC families ($Z_{lr}=1.94$ at marker D17S969). In contrast,

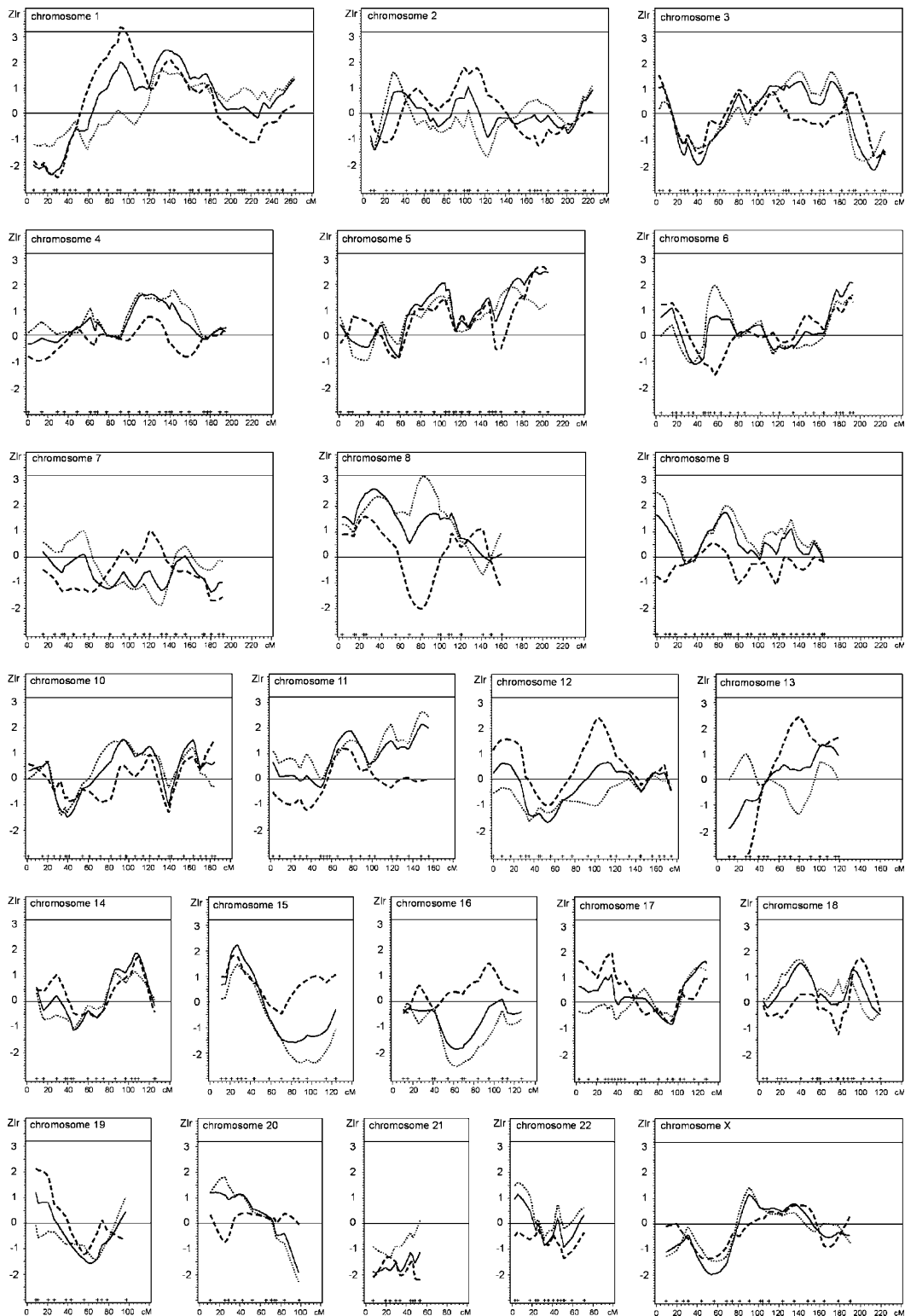


Figure 1 Multipoint nonparametric Z_{Ir} scores plotted for each chromosome. Graphs are shown for all 139 families (solid line), 42 hereditary prostate cancer families (HPC, dashed line) and 92 familial prostate cancer families (FPC, dotted line). A horizontal line at $Z_{Ir} = 3.18$ indicates the threshold of suggestive linkage ($P = 0.00074$). Marker positions are indicated by cross signs on the genetic length scaled in centimorgan (cM).

Table 3 Loci exhibiting nonparametric linkage of $Z_{lr} > 2$ ($P < 0.025$) in families with hereditary prostate cancer or in the total family set

Chr.	Locus	Position (cM)	Marker	All families (n = 139) (Z_{lr})	HPC (n = 47) (Z_{lr})	FPC (n = 92) (Z_{lr})
1	1p31.1	91	D1S438	2.01	3.37	0.06
	1q12	140	D1S189	2.43	2.10	1.51
5	5q14.3	104	D5S428	2.04	1.42	1.52
	5q34	174	D5S2040	2.19	1.19	1.84
	5q35.3	204	D5S2030	2.44	2.42	1.25
6	6q27	190	D6S503	2.07	1.52	1.45
8	8p22	27	D8S1827	2.47	1.61	1.94
12	12q21.3	103	D12S351	0.53	2.42	-1.00
13	13q31.1	80	D13S271	0.32	2.46	-1.38
15	15q13.1	27	D15S231	2.23	1.75	1.47
19	19p13.3	9	D19S565	1.19	2.13	-0.08

Table 4 Evidence of linkage (Z_{lr} scores > 2) especially in families with familial prostate cancer (FPC)

chr.	Locus	pos.	Marker	FPC (Z_{lr})	Model-free analyses		Parametric multipoint analysis of FPC pedigrees	
					HPC (Z_{lr})	Difference P-value	Dominant LOD	Recessive LOD
8	8q13.1	82	D8S543	3.15	-2.00	0.0007	1.73	2.37
	8q21.2	98	D8S273	2.30	-0.33	0.119	0.73	0.32
9	9p24.3	0	D9S1858	2.52	-0.73	0.046	0.68	1.06
	9q12	67	D9S1777	2.00	0.23	0.312	0.39	0.24
11	11q23	119	D11S908	2.13	-0.33	0.118	0.05	0.20
	11q25	148	D11S2367	2.62	-0.05	0.115	0.74	1.08

Table 5 Linkage results at previously described prostate cancer loci

Locus/gene, marker	All families		Subgroup of highest evidence		
	Z_{lr}	$hLOD$ (α) ^a	Families	Z_{lr}	$hLOD$ (α)
HPC1/ <i>RNASEL</i> , D1S218	1.28	0.50 (0.50)	All	—	—
PCAP, D1S304	1.53	0.07 (0.33)	All	—	—
CAPB, D1S507	-2.31	0.00 (0.00)	FPC, n = 92	-1.23	0.00 (0.00)
8p22/ <i>MSR1</i> , D8S1827	2.47	0.73 (0.56)	Late onset, n = 59	2.99	0.93 (0.99)
HPC2/ <i>ELAC2</i> , D17S969	1.08	0.61 (0.57)	HPC, n = 47	1.94	0.99 (0.86)
HPC20, D20S839	-0.39	0.00 (0.06)	HPC, n = 47	0.40	0.00 (0.12)
HPCX, DXS8073	-0.47	0.00 (0.01)	Early onset, n = 80	0.62	0.00 (0.07)

^aParametric analyses represent the dominant model. The LOD score under the assumption of heterogeneity ($hLOD$) is given along with the fraction of contributing pedigrees (α).

the number of families linked to the *RNASEL* gene at 1q25 (*HPC1*) and the *PCAP* region at 1q42–43 was small. No evidence for linkage was seen at the loci 1p36 (*CAPB*), *HPC20* and Xq27 (*HPCX*). Stratification according to the presence or absence of male-to-male transmission did not yield a stronger evidence of linkage to *HPCX* in either subset ($Z_{lr} = -0.58$ in n-mtm families, $Z_{lr} = 0.02$ in mtm families).

Finally, we examined our power to detect true linkage at a single marker depending on variables given in Table 6. No major effect was seen for an increasing marker-to-gene

distance from $\theta = 0.01$ to 0.05, which may represent our relatively high marker density (average spacing = 7.25 cM). Also, a reduction in family number due to sample splitting, as shown for the HPC subgroup, may not severely influence the study power. In contrast, the probability to observe linkage mainly depends on the degree of heterogeneity, a parameter that cannot reliably be determined. Although we chose a conservative option for simulation (parametric two-point analysis), our study might be underpowered with respect to identify a true prostate cancer locus.

Table 6 Estimation of power depending on distance, decreasing sample size and percentage of linked families

Varied (nuisance) parameter	θ	Number of families	Linked families (%)	Probability of reaching LOD ≥ 3.0 (%)
Optimum	0.01	139	100	100
Marker-to-gene distance	0.05	139	100	98
Sample splitting ^a	0.01	47	100	94
Heterogeneity	0.01	139	75	63

^aThe given example refers to the subgroup of 47 families fulfilling the criteria of HPC. Values that were modified with respect to the optimal parameter set are bold.

Discussion

In our prostate cancer family sample from Germany, we obtained neither significant linkage to a new locus nor unequivocal confirmation for any of the accepted loci. Below the level of formal significance, however, there were a number of findings which coincided with observations reported earlier, that is, accepted and suspected loci, plus some indications of new regions.

In this study, we included families with two affected relatives. Selecting families with a strong clustering of prostate cancer is the common strategy for patient recruitment when disease risk is inherited by a dominant and highly penetrant mode of transmission. Using family samples with a strong history of prostate cancer has so far led to the identification of three known high-risk genes, *ELAC2*, *RNASEL* and *MSR1*. There is accumulating evidence, however, that only a small portion of hereditary prostate cancer is based on mutations in these high-risk genes. Prostate cancer is believed to be a genetically heterogeneous and complex disease rather than a classical Mendelian-like disorder, and the involvement of multiple low-risk genes, high-risk genes and gene interaction is assumed. The purpose of our study was to evaluate the entire range of genes predisposing to prostate cancer. Affected relative pairs represent the vast majority of familial prostate cancer in the German population²⁶ and were consequently included in the analysis. To cope with the deliberate excess of heterogeneity in our families, we applied sample-splitting definitions in order to increase our chance to detect linkage at particular susceptibility loci. As demonstrated by simulation analysis, a power gain achieved from increasing homogeneity outweighs the power loss due to the reduced size of the family subgroup. We used age of onset, paternal transmittance and family history as the commonly accepted stratification criteria to increase homogeneity within subsets. The most relevant stratification was the classification of families as 'hereditary' (HPC group) and 'familial' (FPC group), using the Johns Hopkins definition of hereditary prostate cancer, in order to separate dominant high-risk genes from recessive genes and those with reduced penetrance. Due to a limited number of descendants within a family and the sex specificity of prostate cancer, high-risk genes may be present in both families with a high and with a low

number of affected relatives. For example, in the case of the high-risk breast cancer susceptibility genes *BRCA1* and *BRCA2*, it was shown that only 50% of mutation carriers have additional affected relatives.²⁷ On the other hand, dominant genes with reduced penetrance and recessive loci are unlikely to be present in families with a strong clustering of the disease. We therefore considered the entire family set and the HPC group in order to screen for high-risk variants (Table 3). The scan of the FPC group, however, may have increased the chance to detect genes with reduced penetrance or recessive inheritance (Table 4).

According to this concept, we found several loci that seemed to be specific for a moderate family history. All the five marker positions at 8q13/8q21, 9p24 and 11q23/11q25 that showed a $Z_{lr} > 2$ in the FPC group had negative linkage signals in hereditary disease pedigrees. For two loci (8q13 and 9p24), the corresponding Z_{lr} scores showed nominal significant differences. Only a few studies have included prostate cancer families with only two affected members in their linkage analysis. Witte *et al*²⁸ recently analysed a set of 114 newly collected relative pairs and found evidence of linkage to prostate cancer at two loci, 16q23 and 11q24. For the latter locus, our data may provide some confirmation, as it lies within a region of 30 cM within 11q23–25, where we observed Z_{lr} scores greater than 2 in familial prostate cancer.

The highest linkage signal within our FPC group was seen at locus 8q13. A nonparametric Z_{lr} of 3.15 was supported by a maximum parametric LOD score of 2.37 using a recessive model. Gibbs *et al*²⁹ previously described linkage to a site on 8q that was located within 5 cM of our local maximum. These investigators only included hereditary prostate cancer families, but their evidence of linkage was stronger using a recessive model (two-point LOD = 2.17) rather than assuming dominant inheritance. Hence, both studies support the presence of a recessive susceptibility gene on 8q.

In the HPC group, we detected a region at 1p31 reaching the level of suggestive evidence of linkage according to Kruglyak and Lander²⁴ ($Z_{lr} = 3.37$). The marker with the highest score (D1S438) is more than 60 cM centromeric to 1p36 and cannot be attributed to the presumptive susceptibility gene at the *CAPB* locus, which was previously detected in brain-prostate cancer families.¹⁰ Support for a

region of interest at 1p, that is different from *CAPB*, was also seen in a recent genomewide linkage search of Swedish HPC families.³⁰ In the late-onset group of their sample, Wiklund *et al* observed a maximum hLOD score of 1.35 on 1p, 15 cM telomeric from the critical region in our scan. The coincidence of the two independent findings at 1p31 may indicate a new candidate region for hereditary prostate cancer.

A further goal of our study was to determine the particular impact of accepted susceptibility regions in the German population. The prostate cancer loci *HPC1*, *PCAP*, *CAPB*, 8p22, *HPC2*, *HPC20* and *HPCX* were analysed in detail by nonparametric, parametric and sample-splitting methods. We confirmed 8p22 as a high-risk locus initially detected in North American HPC families.⁷ An Z_{lr} of 2.47 was seen in our sample without stratification. Moreover, we found higher evidence in pedigrees with a mean age of onset of over 65 years ($Z_{lr}=2.99$), resembling the observation of two earlier studies^{7,29} that detected linkage particularly in the late-onset groups. In contrast to a recent Swedish replicatory approach, reporting linkage of 8p22 in early onset families,³¹ our data strengthen the hypothesis that 8p22 may account for late-onset disease. In fact almost all (99%) of our late onset families may be attributed to 8p22 (Table 5). Recently, the gene encoding the macrophage scavenger receptor (*MSR1*) has been proposed as a strong candidate gene at 8p22.⁸ Nonsense and missense germline mutation were found at moderate frequencies in North American and Scandinavian prostate cancer families.^{8,32–34} Since our linkage results at 8p22 suggest a high prevalence of disease alleles at least within our study sample, mutation analyses of *MSR1* in German and possibly in further European prostate cancer families seem promising. Aside from the *MSR1* locus, we retrieved some evidence for the *HPC2/ELAC* locus at 17p11, mainly in the HPC group, whereas other accepted susceptibility regions appear to be under-represented in our study sample.

Overall, we detected several regions of potential interest in our study sample by applying a threshold of Z_{lr} greater than 2 ($P=0.025$). However, the high number of linkage signals below the level of significant evidence demonstrates the heterogeneity of prostate cancer susceptibility in the German population. Locus heterogeneity was partially resolved by subgroup analyses, as we detected suggestive linkage once in HPC ($Z_{lr}=3.37$ at 1p31) and another peak close to suggestive linkage in FPC ($Z_{lr}=3.15$ at 8q). However, the splitting into two subsamples must consequently be considered as two independent analyses, of which we would expect suggestive linkage peaks arising by chance exactly two times. Some of our loci were supported by previous studies, although the number of genomewide scans has been inflated and may have decreased the value of 'replicated' findings. Remarkable, however, is a pattern of three of our loci on 19p, 15q and 1p, which were

recently reported from a genomewide scan on Swedish HPC families.³⁰ This similarity may possibly support the hypothesis of a population based genetic heterogeneity.

Conclusion and perspectives

The strong genetic heterogeneity, that is present even in a rather uniform population from Central Europe, has hampered the mapping of relevant prostate cancer genes to date. Thus, novel concepts to increase discriminatory power at individual loci are warranted. Meanwhile, several groups in this field have made use of sample-splitting methods in order to increase homogeneity within subsamples. The benefit of this conventional principle is indicated by power simulation and by our results, since each of the subgroup scans highlighted one particular locus of interest. In addition to the stratifications based on epidemiological criteria, the selection of families based on clinical characteristics could help to define phenotypes of specific genes.

Our results at 1p31 and 8q13 reached suggestive significance only. Given prior data from previous genomewide scans, however, these loci may be preferred as candidate regions. Fine-mapping attempts are under way in order to identify the genes potentially involved in the development of prostate cancer.

Acknowledgements

We want to thank all patients, their families and attending urologists for contributing to this study. We gratefully acknowledge Petra Reutter and Margot Brugger for technical assistance. This work was supported by the Deutsche Krebshilfe, grant number 70-3111-V03 and in part by NCI Grant #U01 CA89600-02.

References

- 1 Lichtenstein P, Holm NV, Verkasalo PK *et al*: Environmental and heritable factors in the causation of cancer—analyses of cohorts of twins from Sweden, Denmark, and Finland. *N Engl J Med* 2000; **343**: 78–85.
- 2 Easton DF, Schaid DJ, Whittemore AS *et al*: Where are the prostate cancer genes? A summary of eight genome wide searches. *Prostate* 2003; **57**: 261–269.
- 3 Schaid DJ: The complex genetic epidemiology of prostate cancer. *Hum Mol Genet* 2004; **13** (Spec No 1): R103–R121.
- 4 Smith JR, Freije D, Carpten JD *et al*: Major susceptibility locus for prostate cancer on chromosome 1 suggested by a genome-wide search. *Science* 1996; **274**: 1371–1374.
- 5 Carpten J, Nupponen N, Isaacs S *et al*: Germline mutations in the ribonuclease L gene in families showing linkage with HPC1. *Nat Genet* 2002; **30**: 181–184.
- 6 Tavtigian SV, Simard J, Teng DH *et al*: A candidate prostate cancer susceptibility gene at chromosome 17p. *Nat Genet* 2001; **27**: 172–180.
- 7 Xu J, Zheng SL, Hawkins GA *et al*: Linkage and association studies of prostate cancer susceptibility: evidence for linkage at 8p22–23. *Am J Hum Genet* 2001; **69**: 341–350.
- 8 Xu J, Zheng SL, Komiya A *et al*: Germline mutations and sequence variants of the macrophage scavenger receptor 1 gene are associated with prostate cancer risk. *Nat Genet* 2002; **32**: 321–325.

- 9 Berthon P, Valeri A, Cohen-Akenine A et al: Predisposing gene for early-onset prostate cancer, localized on chromosome 1q42.2–43. *Am J Hum Genet* 1998; **62**: 1416–1424.
- 10 Gibbs M, Stanford JL, McIndoe RA et al: Evidence for a rare prostate cancer-susceptibility locus at chromosome 1p36. *Am J Hum Genet* 1999; **64**: 776–787.
- 11 Berry R, Schroeder JJ, French AJ et al: Evidence for a prostate cancer-susceptibility locus on chromosome 20. *Am J Hum Genet* 2000; **67**: 82–91.
- 12 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.
- 13 Ostrander EA, Stanford JL: Genetics of prostate cancer: too many loci, too few genes. *Am J Hum Genet* 2000; **67**: 1367–1375.
- 14 Carter BS, Beaty TH, Steinberg GD et al: Mendelian inheritance of familial prostate cancer. *Proc Natl Acad Sci USA* 1992; **89**: 3367–3371.
- 15 Gronberg H, Damber L, Damber JE et al: Segregation analysis of prostate cancer in Sweden: support for dominant inheritance. *Am J Epidemiol* 1997; **146**: 552–557.
- 16 Schaid DJ, McDonnell SK, Blute ML et al: Evidence for autosomal dominant inheritance of prostate cancer. *Am J Hum Genet* 1998; **62**: 1425–1438.
- 17 Carter BS, Bova GS, Beaty TH et al: Hereditary prostate cancer: epidemiologic and clinical features. *J Urol* 1993; **150**: 797–802.
- 18 Cui J, Staples MP, Hopper JL et al: Segregation analyses of 1,476 population-based Australian families affected by prostate cancer. *Am J Hum Genet* 2001; **68**: 1207–1218.
- 19 Gong G, Oakley-Girvan I, Wu AH et al: Segregation analysis of prostate cancer in 1,719 white, African-American and Asian-American families in the United States and Canada. *Cancer Causes Control* 2002; **13**: 471–482.
- 20 Kong A, Cox NJ: Allele-sharing models: LOD scores and accurate linkage tests. *Am J Hum Genet* 1997; **61**: 1179–1188.
- 21 Kruglyak L, Daly MJ, Reeve-Daly MP et al: Parametric and nonparametric linkage analysis: a unified multipoint approach. *Am J Hum Genet* 1996; **58**: 1347–1363.
- 22 Mirea L, Briollais L, Bull S: Tests for covariate-associated heterogeneity in IBD allele sharing of affected relatives. *Genet Epidemiol* 2004; **26**: 44–60.
- 23 Xu J: Combined analysis of hereditary prostate cancer linkage to 1q24–25: results from 772 hereditary prostate cancer families from the International Consortium for Prostate Cancer Genetics. *Am J Hum Genet* 2000; **66**: 945–957.
- 24 Lander E, Kruglyak L: Genetic dissection of complex traits: guidelines for interpreting and reporting linkage results. *Nat Genet* 1995; **11**: 241–247.
- 25 Ott J: Computer-simulation methods in human linkage analysis. *Proc Natl Acad Sci USA* 1989; **86**: 4175–4178.
- 26 Paiss T, Herkommer K, Chab A et al: Familial prostate carcinoma in Germany. *Urology A* 2002; **41**: 38–43.
- 27 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.
- 28 Witte JS, Suarez BK, Thiel B et al: Genome-wide scan of brothers: replication and fine mapping of prostate cancer susceptibility and aggressiveness loci. *Prostate* 2003; **57**: 298–308.
- 29 Gibbs M, Stanford JL, Jarvik GP et al: A genomic scan of families with prostate cancer identifies multiple regions of interest. *Am J Hum Genet* 2000; **67**: 100–109.
- 30 Wiklund F, Gillanders EM, Albertus JA et al: Genome-wide scan of Swedish families with hereditary prostate cancer: suggestive evidence of linkage at 5q11.2 and 19p13.3. *Prostate* 2003; **57**: 290–297.
- 31 Wiklund F, Jonsson BA, Goransson I et al: Linkage analysis of prostate cancer susceptibility: confirmation of linkage at 8p22–23. *Hum Genet* 2003; **112**: 414–418.
- 32 Wang L, McDonnell SK, Cunningham JM et al: No association of germline alteration of MSR1 with prostate cancer risk. *Nat Genet* 2003; **35**: 128–129.
- 33 Seppala EH, Ikonen T, Autio V et al: Germ-line alterations in MSR1 gene and prostate cancer risk. *Clin Cancer Res* 2003; **9**: 5252–5256.
- 34 Lindmark F, Jonsson BA, Bergh A et al: Analysis of the macrophage scavenger receptor 1 gene in Swedish hereditary and sporadic prostate cancer. *Prostate* 2004; **59**: 132–140.