Prostate Cancer Susceptibility Locus on Chromosome 1q: a Confirmatory Study

Kathleen A. Cooney, Jennifer D. McCarthy, Ethan Lange, Li Huang, Susan Miesfeldt, James E. Montie, Joseph E. Oesterling, Howard M. Sandler, Kenneth Lange*

**Background:** Recent recognition that a predisposition to prostate cancer can be inherited has led to a search for specific genes associated with the disease. Through a study of families with three or more affected first-degree relatives, a region on the long arm of chromosome 1 (i.e., 1q24-25) has been tentatively identified as containing a gene, HPC1, involved in the development of hereditary prostate cancer. Confirmation of this finding is needed, however, before attempts are made to isolate and characterize the putative HPC1 gene.

**Purpose:** To confirm that chromosome 1q24-25 contains a gene relevant to hereditary prostate cancer, we analyzed an independent set of families, each with two or more affected individuals.

**Methods:** Fifty-nine unrelated families were selected for analysis on the sole criterion that more than one living family member was affected by prostate cancer. DNA samples were subsequently isolated from 130 individuals with the disease. These samples were genotyped at six polymorphic marker sequences (D1S215, D1S2883, D1S466, D1S158, D1S518, and D1S2757) covering the chromosomal region proposed to contain HPC1. The resulting data were analyzed by nonparametric multipoint linkage (NPL) methods, yielding NPL Z scores and corresponding one-sided P values.

**Results:** When the entire set of 59 families was considered, the occurrence of prostate cancer (and, presumably, the HPC1 gene) was most tightly linked to marker D1S466 (NPL Z score = 1.58; P = .0574). Analysis of the 20 families (51 affected individuals) fulfilling one or more of the proposed clinical criteria for hereditary prostate cancer (i.e., three or more affected individuals within one nuclear family; affected individuals in three successive generations [maternal or paternal lineage]; and/or clustering of two or more individuals affected before the age of 55 years) revealed more convincing evidence of disease linkage to chromosome 1q24-25 (maximum NPL Z score [at marker D1S466] = 1.72; P = .0451). The 39 families (79 affected individuals) that did not meet the clinical criteria for hereditary prostate cancer exhibited no significant evidence of disease linkage to DNA sequences at chromosome 1q24-25 (maximum NPL Z score [at marker D1S466] = 0.809; P = .208). The six African-American families in our study contributed disproportionately to the observation of linkage, with a maximum NPL Z score at marker D1S158 of 1.39 (P = .0848) for these families. **Conclusions and Implications:** Our data confirm that chromosome 1q24-25 is likely to contain a prostate cancer susceptibility gene. Future efforts at positional cloning of the HPC1 gene should focus on families who meet the proposed clinical criteria for hereditary prostate cancer. [J Natl Cancer Inst 1997;89:955-9]
inherited and sporadic forms of the disease (4). Finally, prostate cancer is exceedingly common in the general population, with a one in five lifetime risk to American men (3). This high phenocopy rate further complicates genetic linkage analysis in prostate cancer families.

Despite these obstacles, Smith et al. (5) have tentatively identified a genetic locus, HPC1, on chromosome 1 that is associated with prostate cancer predisposition. These investigators conducted a genome screen on 66 high-risk prostate cancer families, each with three or more affected first-degree family members. In the initial stage of their screen with 341 polymorphic markers, Smith et al. found evidence that marker DIS218 is linked to a locus associated with the development of prostate cancer (LOD score \( \sigma = 2.75 \)). Subsequent typing of additional markers mapping to chromosome 1q24-25 and analysis by nonparametric as well as parametric statistical methods strengthened the case for linkage (5). Furthermore, admixture analysis implicated HPC1 in approximately one third of the families. Unfortunately, the most likely candidate region extends over a broad 20 centimorgan (cM) interval (i.e., approximately 20 million base pairs).

Before the HPC1 gene is isolated and characterized, its existence should be statistically confirmed in an independent set of families. In our study, we employed nonparametric linkage methods (6) and six polymorphic markers spanning the HPC1 candidate interval to analyze 59 families, each with two or more individuals affected by prostate cancer. Nonparametric methods are particularly appropriate in view of recent suggestions of genetic heterogeneity in this disease (7,8). We obtained further evidence for HPC1 linkage to prostate cancer, in particular in the families who meet one or more of the proposed clinical criteria for HPC. These criteria are as follows: 1) three or more affected individuals within one nuclear family; 2) affected individuals occurring in three successive generations (maternal or paternal lineage); or 3) a cluster of two or more relatives each affected before the age of 55 years (2).

Patients and Methods

Patient Selection

Patients with prostate cancer evaluated at the University of Michigan were interviewed to obtain information about a family history of the disease. Individuals with at least one living affected relative were asked to participate in the next phase of the study, which included providing a blood sample, information on extended family history, and access to medical records. The majority of the families (50 [85%] of 59) were identified at one of the investigators’ institutions. The other nine families either contacted our research group directly or were referred by local physicians because of a family history of prostate cancer. All research protocols and consent forms were approved by the Institutional Review Board at the University of Michigan. Blood was collected from affected individuals and was used for DNA preparation and cell immortalization (i.e., establishment of permanent cell lines). DNA was isolated from nucleated blood cells by use of the Puregene kit (Gentra Systems, Inc., Plymouth, MN).

Genotype Analysis

All individuals were genotyped by use of polymorphic markers spanning the HPC1 candidate region of chromosome 1q24-25 that was reported by Smith et al. (5) (Table 1). Primers for the polymerase chain reaction (PCR) were purchased from Research Genetics (Huntsville, AL), and PCR was performed as described by Nichols et al. (9) with minor modifications. Briefly, one primer from each primer pair was labeled at its 5’ end with \( \text{[32P]} \)-adenosine triphosphate (New England Nuclear Life Sciences Products, Boston, MA) and T4 polynucleotide kinase (New England Biolabs, Inc., Beverly, MA).

Table 1. Polymorphic markers on chromosome 1q used for linkage analysis

<table>
<thead>
<tr>
<th>Locus</th>
<th>Polymorphism</th>
<th>Heterozygote frequency</th>
<th>cM</th>
</tr>
</thead>
<tbody>
<tr>
<td>DIS215</td>
<td>Dinucleotide</td>
<td>0.73</td>
<td>—</td>
</tr>
<tr>
<td>DIS2883</td>
<td>Dinucleotide</td>
<td>0.75</td>
<td>0</td>
</tr>
<tr>
<td>DIS466</td>
<td>Dinucleotide</td>
<td>0.77</td>
<td>4.0</td>
</tr>
<tr>
<td>DIS158</td>
<td>Dinucleotide</td>
<td>0.89</td>
<td>2.6</td>
</tr>
<tr>
<td>DIS518</td>
<td>Tetranucleotide</td>
<td>0.70</td>
<td>5.0</td>
</tr>
<tr>
<td>DIS2757</td>
<td>Dinucleotide</td>
<td>0.85</td>
<td>6.0</td>
</tr>
</tbody>
</table>

*Marker information was obtained from Human Genome Database (GDB [TM], Baltimore, MD): The Johns Hopkins University, 1990-; available from Internet at URL http://gdbwww.gdb.org. Five of the six markers were developed by Genethon (23); DIS158 was originally described by Overbeck et al. (13). The relative distances between markers were assigned by use of the computer program MULTIMAP (19). tCm = centimorgan (see “Notes” section for additional information).

All reaction mixtures contained 1 \( \mu \text{L} \) 10x PCR buffer (Promega Corp., Madison, WI), 1 \( \mu \text{L} \) 2 mM standard deoxynucleotide triphosphates, 5 ng each primer, 0.07 \( \mu \text{L} \) Taq polymerase (Promega Corp.), 20 ng genomic DNA, and water in total volume of 10 \( \mu \text{L} \). PCR was performed by use of an MJ Research (Watertown, MA) 96-well thermocycler, with annealing temperatures optimized for each primer set. PCR products were analyzed in 8%-polyacrylamide DNA-sequencing gels, with end-labeled MspI-digested pBR322 fragments as size markers. The gels were run at 65 W for approximately 2.5 hours and then exposed to x-ray film for 2-10 hours at –80°C (in the presence of an intensifying screen). The genotypes were scored visually by two observers.

Statistical Methods

A genetic map of markers DIS215, DIS2883, DIS466, DIS158, DIS518, and DIS2757 on chromosome 1q24-25 was constructed by use of the computer program MULTIMAP (10) (Table 1) and a subset of the CEPH families (11). All markers could be ordered unambiguously with the exception of the adjacent markers DIS215 and DIS2883. Because no recombination events were observed between these two markers, they were assigned a genetic distance of 0 cM. The order of markers DIS2883, DIS466, and DIS158 agrees with that reported by Smith et al. (5).

Nonparametric multipoint linkage (NPL) analysis was carried out by use of the computer program GENEHUNTER (6), which implicitly assumes Hardy–Weinberg and linkage equilibrium. A version of the computer program MENDEL (12) was used to estimate allele frequencies from data obtained in the analysis of our 59 families. The NPL Z scores and corresponding one-sided \( P \) values were calculated twice—once using the allele frequencies calculated by MENDEL and once using allele frequencies derived from CEPH family data (for markers DIS215, DIS2883, DIS466, and DIS2757) (11) or data from a Causpan population (for marker DIS158) (13). Allele frequencies for marker DIS518 are unavailable from CEPH, so MENDEL-derived frequencies were invoked in both cases. With either set of allele frequencies, the assumption of genetic equilibrium in calculations of NPL Z scores is likely to be slightly violated because the data from our patient families were pooled across ethnic lines. However, the six markers that we have chosen appear to be highly polymorphic in both African-American and Caucasian families.

Results

Patient Demographic Data

DNA samples were obtained from 130 men with prostate cancer who were from 59 unrelated families. The diagnosis of the disease was confirmed by a review of medical records for 118 (91%) of the 130 men. Two affected individuals were studied from each of 48 families, and three affected individuals were examined from each of 10 families. Four affected individuals were typed in the remaining family, including the proband and the proband’s father, paternal uncle, and paternal cousin (son of a paternal aunt). Thirty-eight (79%) of the 48 paired relatives...
were brother pairs, and the remaining pairs were either uncle/nephew or cousin pairs. Five of the families with three affected individuals contained three affected brothers. The other families with three affected individuals contained varying combinations of affected first- and second-degree relatives. Six (15 affected individuals) of the 59 families were of African-American descent, 51 families (109 affected individuals) were of Caucasian descent, and two families (six affected individuals) were of unknown ethnic origin.

Further examination of the 59 families revealed the following characteristics. The average number of affected individuals per family was 3.1 (range, two to 15 individuals). The average number of affected men available for genotyping was 2.2 per family. Among the 130 men who were genotyped, the average age at the diagnosis of prostate cancer was 65.2 years. The average age at the onset of prostate cancer in a family could be determined for 52 (88%) of the 59 families; accurate information could not be obtained for seven families. Using the available data, 20 (34%) of the 59 families had an average age at disease onset under 65 years. Twenty (34%) of the 59 families met at least one of the clinical diagnostic criteria for HPC (2). Data describing the 20 HPC families (51 affected individuals) and the 39 non-HPC families (79 affected individuals) are summarized in Table 2. Notably, the average age at the onset of prostate cancer was lower in the HPC families than in the non-HPC families (63.4 years versus 66.3 years; two-sided t test, $P = .064$).

### Linkage Analysis

All of the 130 men with prostate cancer were genotyped using the six chromosome 1q24-25 markers listed in Table 1. In addition, seven unaffected individuals from each of five families were genotyped; the data from these unaffected individuals were used only to infer parental genotypes and to improve the likelihood of determining whether an affected relative pair was identical by descent rather than by state at the various marker loci. NPL analysis of all 59 families, using allele frequencies determined primarily from an analysis of CEPH families, yielded a maximum NPL Z score of 1.58 ($P = .0574$) at marker D1S466 (Table 3). When only the 20 families that fulfilled one or more of the clinical diagnostic criteria for HPC were analyzed, the NPL Z score at marker D1S466 increased to 1.72 ($P = .0451$). The maximum NPL Z score for the remaining 39 families who did not meet the clinical criteria for HPC was 0.809 (at marker D1S466; $P = .208$). Importantly, analysis of just the six African-American families yielded a maximum NPL Z score at marker D1S158 of 1.39 ($P = .0848$).

Similar results were obtained with allele frequencies calculated from the families in this study by use of the computer program MENDEL (12). The maximum NPL Z score for the entire group of 59 families was 1.44 ($P = .0756$) at marker D1S466. For the 20 HPC families, the maximum NPL Z score, also at this marker locus, was 1.61 ($P = .0558$). The lowering of the maximum NPL Z scores and the inflation of the corresponding $P$ values when using allele frequencies estimated from our prostate cancer families can best be explained by the major impact of rare marker alleles on sharing statistics. If allele frequencies are estimated from affected families, then the frequencies of rare alleles that actually appear in the families are overestimated and their contributions to the sharing statistics are correspondingly underestimated.

### Discussion

One of the difficulties in studying the genetics of prostate cancer is the establishment of an accurate clinical definition of an inherited syndrome. The criteria for diagnosis should be sufficiently broad to encompass most families with the inherited syndrome but specific enough to exclude clusters of sporadic cases. The proposed definition of HPC (2) is based on the results of a segregation analysis of families at apparently increased risk of prostate cancer, which were reported by Carter et al. (1). This analysis suggested that a rare autosomal dominant gene could be causing prostate cancer in some families an average of 5-10 years earlier than is observed for sporadic cases (2). Since evidence from a segregation analysis of this complexity is necessarily indirect, we chose to enroll all patients with prostate cancer who had at least one living affected relative, but then to partition the families into subgroups on the basis of whether or not they met one or more of the proposed diagnostic criteria for HPC. The 20 families in our study fulfilling the diagnostic criteria for HPC provided compelling evidence of disease linkage to the putative HPC1 gene.

Mutations in hereditary cancer genes often predispose individuals to more than one histologic type of cancer. Cataloging these associated cancers in patients and their close relatives may sharpen the clinical diagnosis of a cancer predisposition syndrome. For example, mutations in BRCA2 predispose carriers to male breast cancer (14,15) and to pancreatic cancer (15) in addition to female breast cancer. Recognition of a homozygous deletion in a pancreatic cancer narrowed the BRCA2 candidate interval and facilitated the eventual cloning of the gene (16,17). On the basis of an analysis of a large number of high-risk prostate cancer families, Isaacs et al. (18) suggested that HPC may be relatively site-specific, possibly with the exception of an excess number of primary cancers of the central nervous system in affected families. Although limited data are available on cancers affecting other family members in our study, 10 of the 130 individuals studied had at least one additional primary cancer. These cancers included three cases of melanoma, four cases of nonmelanoma skin cancer, one medullary thyroid cancer, one neuroendocrine cancer, one probable leiomyosarcoma, one superficial bladder cancer, and one renal cancer. There were no cases of primary cancer of the central nervous system. A continued search for secondary cancers in HPC families is warranted.

The age-adjusted incidence of prostate cancer in African-
American men is approximately 50% greater than that in Caucasian-American men (19), but the relative risk associated with a positive family history of prostate cancer is similar in African Americans and Caucasian Americans (20,21), emphasizing the importance of predisposing germline mutations in both ethnic groups. In general, however, African-American families have been underrepresented in studies of HPC. Smith et al. (5) studied two African-American families (of 79 North-American pedigrees) that had a combined LOD score of 1.4 using markers at chromosome 1q24-25. Six of the families in our study were of African-American descent, and four of the six families met at least one of the clinical diagnostic criteria for HPC. Analysis of the six African-American families alone yielded a maximum NPL Z score of 1.39 at marker D1S158 ($P = .0848$). However, one needs to exercise caution in interpreting these results because the distribution of the NPL Z statistic is only asymptotically correct, and marker frequencies were estimated from Caucasian families. Despite these reservations, our data strongly suggest that the African-American families contributed disproportionately to the evidence for disease linkage to chromosome 1q24-25 in our prostate cancer families. This observation merits confirmation both for its clinical importance as well as for its implications in the positional cloning of the HPC1 gene.

Nonparametric linkage methods were chosen for the analysis of our dataset rather than standard parametric methods (LOD scores) for two reasons. First, a good genetic model for the transmission of the putative HPC1 gene is simply lacking. Such a model would, at a minimum, require a disease-allele frequency and separate age-at-onset curves for each disease-locus genotype. Our dataset is too small to sustain a model of this complexity. Second, parametric tests are more sensitive than nonparametric tests to genetic and environmental heterogeneity. The marker allele clustering statistics suggested by Kruglyak et al. (6) and Sobel and Lange (22) are relatively insensitive to departures from strict patterns of Mendelian inheritance. This robustness property makes the nonparametric statistics ideal for linkage studies in common diseases such as prostate cancer.

In conclusion, this linkage study of 59 families with two or more members diagnosed with prostate cancer confirms the existence of a prostate cancer susceptibility locus at chromosome 1q24-25, as originally proposed by Smith et al. (5). The observation that marker D1S466 has the highest NPL Z score in both the current study and that of Smith et al. may assist in the genetic localization of the HPC1 gene. Our data further suggest that the criteria for HPC proposed by Carter et al. (2) accurately identify a relevant subset of families for mapping studies. Finally, our data hint that mutations in the HPC1 gene are more prevalent in African Americans than in Caucasian Americans. Oncologists and geneticists should concentrate on the identification of additional HPC families for genotyping with chromosome 1q24-25 markers. This strategy should narrow the candidate interval and facilitate the positional cloning of HPC1.

### References


### Table 3. Results of nonparametric multipoint linkage (NPL) analysis*

<table>
<thead>
<tr>
<th></th>
<th>All families (n = 59)</th>
<th>HPC families (n = 20)</th>
<th>Non-HPC families (n = 39)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>NPL Z score</td>
<td>P</td>
<td>NPL Z score</td>
</tr>
<tr>
<td>10 cM proximal</td>
<td>0.649</td>
<td>.257</td>
<td>0.599</td>
</tr>
<tr>
<td>D1S215/D1S2883</td>
<td>1.11</td>
<td>.133</td>
<td>0.965</td>
</tr>
<tr>
<td>D1S466</td>
<td>1.58</td>
<td>.0574</td>
<td>1.72</td>
</tr>
<tr>
<td>D1S158</td>
<td>1.47</td>
<td>.0714</td>
<td>1.65</td>
</tr>
<tr>
<td>D1S518</td>
<td>0.978</td>
<td>.164</td>
<td>1.17</td>
</tr>
<tr>
<td>D1S2757</td>
<td>0.868</td>
<td>.193</td>
<td>1.48</td>
</tr>
<tr>
<td>10 cM distal</td>
<td>0.522</td>
<td>.300</td>
<td>0.965</td>
</tr>
</tbody>
</table>

*The NPL Z score and corresponding one-sided $P$ value for each marker on chromosome 1q24-25 (calculated by use of allele frequencies derived from Centre d’Études du Polymorphisme Humain family data [see “Notes” section for additional information]) is reported for the analysis of all 59 families with one or more living family members affected by prostate cancer, the 20 families who met one or more of the proposed clinical criteria for hereditary prostate cancer (HPC) (2), and the remaining 39 families that did not meet these criteria. The bold values indicate the highest NPL Z score (and the smallest $P$ value) for each group. Proximal and distal refer to relative positions on the chromosome, i.e., 10 centimorgans (cM) closer to the centromere from the marker cluster or 10 cM farther away. See “Notes” section for additional information on centimorgans.


Notes

1A LOD score is a base 10 logarithm of the odds favoring the linkage of two loci. A maximum LOD score of 3.0 has been traditionally accepted as strong evidence for linkage. When Smith et al. (5) tested additional markers at chromosome 1q24-25, a LOD score of 3.65 was obtained.

2Centimorgans (cM) are units of genetic distance. Short distances measured in cM are roughly equal to recombination fractions (i.e., 1 cM = a recombination frequency of ∼1%).

3The CEPH (Centre d’Etudes du Polymorphisme Humain) families are a set of 61 large families that are standardly used for genetic mapping studies.

Supported by Public Health Service grants P50 CA69568 (National Cancer Institute), RO1 GM53275 (National Institute of General Medical Sciences), and T32 HG00040 (National Center for Human Genome Research), National Institutes of Health, Department of Health and Human Services; the Department of Surgery Research Advisory Committee, University of Michigan; and the Schmidt Family Medical School Scholarship, University of Michigan.

We thank all of the families involved in this study as well as J. Neuenschwander, K. Sparks, K. Jones, J. Purakal, P. Kielczewski, and K. King. We also acknowledge D. Siemieniak for assistance in use of MULTIMAP software; E. Fearon, D. Ginsburg, and E. Petty for critical review of the manuscript; and W. Nichols, V. Terry, W. Isaacs, and J. Trent for helpful advice.

Manuscript received April 2, 1997; revised May 27, 1997; accepted May 29, 1997.