A Genetic Epidemiological Study of Hereditary Prostate Cancer (HPC) in Finland: Frequent HPCX Linkage in Families with Late-onset Disease

Johanna Schleutker, Mika Matikainen, Jeffrey Smith, Pasi Koivisto, Agnes Baffoe-Bonnie, Tommi Kainu, Elizabeth Gillanders, Risto Sankila, Eero Pukkala, John Carpten, Dietrich Stephan, Teuvo Tammela, Michael Brownstein, Elizabeth Gillanders, Risto Sankila, Olli-P. Kallioniemi

ABSTRACT

Several predisposition loci for hereditary prostate cancer (HPC) (OMIM #601518) and HPCX at Xq27–q28 (OMIM #300147) have been suggested, including HPC1 at 1q24-q25 and HPCX at Xq27-q28. Genetic heterogeneity in families with HPC and HPCX loci in a series of Finnish prostate cancer families was studied, especially in subgroups of families defined by age, number of affected cases, and the mode of disease transmission. DNA samples were collected from 57 Finnish HPC families with at least two living prostate cancer patients. Linkage analysis was carried out with 39 microsatellite markers for the HPC1 region and 22 markers for the HPCX region. The maximum two-point LOD score for the HPC1 region was 2.05 (marker DXS1205, at θ = 0.14), whereas HPCX LOD scores were all negative. In HOMOG3R analyses, significant evidence of heterogeneity was observed. Subgroup analyses performed to explore the nature of this heterogeneity indicated that families with no male-to-male (NMM) transmission and a late age of diagnosis (≥65 years) accounted for most of the HPCX-linked cases. The maximum HPCX LOD score in this subgroup was 3.12 (θ = 0.001). Nonparametric sibling pair analyses gave a peak LOD score of 3.04 (P < 0.000003) for the NMM transmission subgroup. No subgroup showed any positivity for HPC1. This study suggests that the HPCX-linked prostate cancer families represent a distinct subgroup characterized by NMM transmission of disease and late age of diagnosis.

INTRODUCTION

Inherited genetic defects have been suggested to contribute strongly to 5–10% of all prostate cancers and up to 45% of early-diagnosis cases. Segregation analyses have predicted the existence of at least one rare autosomal dominant susceptibility locus (1–3), although autosomal recessive (4) and X-linked modes of inheritance have also been suggested (5). The first prostate cancer susceptibility locus, HPC1 (OMIM #601518), was assigned to the 1q24-q25 chromosomal region by linkage analysis of high-risk families from the United States and Sweden (6). Subsequent studies have yielded conflicting evidence, some confirming linkage (7, 8), and others finding no evidence of linkage (9, 10). A meta-analysis of 772 families has confirmed linkage to HPC1 but indicated that only about 6% of all cases are linked to HPC1 (11). Recently, two other prostate cancer susceptibility loci have been mapped to chromosome 1, one on the distal q arm on 1q42-q45 (12), and another one, a putative prostate-brain tumor locus, on 1p36 (13). Further evidence of genetic heterogeneity of HPC1 was provided by our recent analysis of a combined study population of 360 HPC families from North America, Finland, and Sweden, which identified evidence for a prostate cancer susceptibility locus on the X chromosome, at Xq27–q28 (termed HPCX, OMIM #300147; Ref. 14). Since then, another new autosomal locus was identified on chromosome 20 (15), and several new, model-dependent loci from genome-wide searches have also been suggested (16, 17).

Therefore, in the past 3 years, evidence implicating several

1 The project in Finland has received federal funds from the national Human Genome Research Institute (NIH Grant NO1-HG-55389). This work was also supported in part by grants from the Medical Research Fund of Tampere University Hospital, the Reino Lahtikari Foundation, the Finnish Cancer Society, the Sigrid Juselius Foundation, and the Academy of Finland.

2 To whom requests for reprints should be addressed, at Laboratory of Cancer Genetics, Institute of Medical Technology, University of Tampere and Tampere University Hospital, FIN-33521 Tampere, Finland. Phone: (358)-3-247; Fax: (358)-3-2474168; E-mail: lojosc@uta.fi.

3 J. Sc., M. M., and J. Sm. contributed equally to this work.

4 The abbreviations used are: HPC, hereditary prostate cancer; PSA, prostate-specific antigen; NMM, no male-to-male; OMIM, Online Mendelian Inheritance of Man.
different loci in genetic predisposition to prostate cancer has emerged. Exploration of the genetic basis of prostate cancer susceptibility is likely to be challenging due to the large number of candidate loci, incomplete penetrance, and high frequency of phenocopies. Heterogeneity could be reduced by analyzing genetically homogeneous populations or by identifying subgroups of patients who are more likely to be linked to a particular locus.

This study was designed to evaluate the importance of the HPCX and HPC1 loci in the causation of prostate cancer in the Finnish population. The HPC1 locus was recently found to be involved mostly in families with an early age of onset and five or more affected cases (11, 18, 19). We therefore explored the role of HPCX and HPC1 in subgroups of the Finnish families defined by phenotypic features of the patients and disease transmission. The five million Finns represent a unique target population to study genetic susceptibility to cancer and other complex diseases (20, 21). Using nation-wide cancer as well as parish and population registries, one can identify cancer families in an unbiased manner. Furthermore, because of the homogeneous genetic background of the Finnish population, the heterogeneity of disease-causing loci and the heterogeneity of mutations at a given locus may be less pronounced than in ethnically more diverse populations.

In this study, we genotyped DNA samples from 57 Finnish prostate cancer families with at least two living affected cases. Linkage analysis was carried out with 39 microsatellite markers for the HPC1 region and 22 markers for the HPCX region. Statistical analyses were carried out in subgroups defined by family size, mode of transmission, and the age of prostate cancer diagnosis.

**MATERIALS AND METHODS**

**Family Identification and Sample Collection.** A total of 292 Finnish HPC families meeting the criteria of Carter et al. (1, 22) and Walsh and Partin (23) were identified in our nation-wide scan. In short, these criteria were: (a) prostate cancer present in three different generations; (b) three first-degree relatives with prostate cancer in a family; or (c) two affected family members diagnosed at the age of 55 years or earlier. HPC families in Finland were ascertained by a number of methods with appropriate approval from the Ministry of Health and Social Affairs and the local ethics committees of regional hospitals. The study was also approved by the National Human Genome Research Institute Institutional Review Board. First, a questionnaire-based approach was undertaken at the Tampere University Hospital area on 355 living prostate cancer patients diagnosed during 1988–1993. Twenty-six prostate cancer patients with a positive family history were ascertained. Second, in a nation-wide cancer registry-based search, 1547 prostate cancer patients diagnosed between 1988 and 1993 were identified from the Finnish Cancer Registry (24). Altogether, 9877 first-degree relatives (5202 males and 4675 females) were identified from church parish registries. Linking the information on the relatives back to the Cancer Registry data (1967–1995) resulted in the identification of 93 families with two or more affected cases. Third, articles and advertisements in major Finnish newspapers, television, and radio resulted in 500 contacts. Of these, 151 families met our HPC criteria. Finally, letters were sent to all 120 practicing urologists in Finland. Fifty-two families were identified in this manner. The degree of overlap (30 families) between the different ascertainment methods suggests that the 292 individual families identified represent a significant fraction of all HPC families in Finland.

Diagnoses were confirmed using the Finnish Cancer Registry or individual patient records from regional hospitals. A unique personal identification code was assigned to each Finn in 1967. All cancer diagnoses made after this time were confirmed using the disease registries, whereas those made earlier rely primarily on family history information.

If the family met the Carter et al. (1, 22) criteria, all living affected cases, as well as the spouse and adult-aged offspring of deceased patients, were contacted to obtain informed consent and to request a blood sample for linkage analyses. All male individuals of ≥45 years who participated the study were tested for total serum PSA (25) in accordance with the informed consents that the patients had given. If abnormal age-adjusted values were obtained, the men were referred to a local urology department for subsequent urological examination to exclude the presence of subclinical prostate cancer. Seven histologically confirmed prostate cancer cases were identified by PSA screening (25).

**Genotyping.** Altogether, 869 blood specimens were obtained for DNA isolation. Linkage studies of the HPC1 and HPCX regions were performed on 57 informative families. These are the same families first analyzed by Xu et al. (14). No bilinear families were included. A large number of unaffected cases were also collected to infer phase and to construct the haplotypes of deceased patients. The mean number of affected individuals was 3.2 individuals/family (range, 2–9 individuals/family), and the mean number of genotyped individuals affected was 2.4 individuals/family (range, 2–9 individuals/family). In addition, four specimens from formalin-fixed tumor specimens from selected families were studied. For population controls, anonymous, whole blood specimens from 160 normal healthy blood donors were obtained from the Blood Center of the Finnish Red Cross (Helsinki, Finland).

Genomic DNA was prepared from a 10-ml whole blood sample or from a paraffin-embedded tissue sample using Puregene kit (Gentra Systems, Inc., Minneapolis, MN). A total of 20–60 ng of genomic DNA was used per PCR reaction, and fluorescence-labeled PCR primers for 39 different microsatellite markers at 1q (6) and 22 different microsatellite markers at Xq were used. High throughput, semiautomated genotyping was accomplished by means of ABI 377 sequencers. Archived tissue specimens were analyzed using an ABI 310 DNA sequencer. GENESCAN and GENOTYPER programs (Applied Biosystems, Foster City, CA) were used in data analysis as described by Smith et al. (6).

**Linkage Analyses.** Standard two-point and multipoint parametric likelihood analysis was performed using the computer program FASTLINK (26). The parameters of the trait model used in the linkage analyses were identical to those used by Smith et al. (6). In brief, only individuals with verified diagnoses of prostate cancer were considered to be affected. Males under the age of 75 years with normal (age-adjusted) PSA values were treated as having an unknown status. Males over 75 years with normal PSA values were considered to be unaffected. Age-dependent penetrance values with three liability classes
were used. The frequency of the \textit{HPC1} and \textit{HPCX} genes was set to 0.003. In sliding four-point FASTLINK analysis for \textit{HPCX}, markers DXS1232, DXS1205, and DXS6571 were used with recombination fractions of 0.02 and 0.006, respectively. Additional parametric and nonparametric multipoint analyses were performed with GENEHUNTER (27). The X chromosome version of GENEHUNTER was used in X chromosome analyses. For the subgroup analyses, 33 families were classified as having NMM transmission of the disease (no affected fathers or affected uncles on the paternal side of the family), and 24 families were classified as having male-to-male transmission of the disease (affected father or paternal uncle).

Finnish allele frequencies for each marker were estimated from the founders of the 57 linkage families and from 160 anonymous control samples obtained from blood donors. In the subgroup linkage analyses, families with a mean age of onset of 65 years or less were considered as early-diagnosis families, and all others were considered as late-diagnosis families. For subgroup analyses, one representative, positive marker for each HPC candidate area was selected [DXS158 for \textit{HPC1} (6) and DXS1205 for \textit{HPCX} (14)].

\textbf{Sibling Pair and Heterogeneity Tests.} Nonparametric sibling pair tests of linkage were performed with the program package \textsc{analyze}, using \textsc{polylocus} (28). HOMOG3R was used to calculate the log likelihoods of heterogeneity under the assumptions that \((a)\) in proportion \(a_1\) of families, the trait is linked to marker 1 (DXS1205), \((b)\) in proportion \(a_2\) of families, it is linked to marker 2 (DXS1240); \((c)\) the two markers (or maps of markers) are located in different regions of the genome; \((d)\) the trait is never truly linked to both markers; and \((e)\) there may be a third proportion, \(a_3\) \([a_3 = (1 - a_1 - a_2)]\), of families without linkage to markers 1 and 2. Family-specific LOD scores at both markers (DXS1205 and DXS1240) are provided to HOMOG3R, which then estimates \(a_1\) and \(a_2\) and calculates the difference in Ln(likelihood) of the models under heterogeneity versus homogeneity. Twice this difference is distributed asymptotically as a \(\chi^2\) with 3 degrees of freedom and tests for evidence of heterogeneity. To test for evidence of two linked loci as opposed to only one locus, HOMOG3R calculates the difference in Ln(likelihood) of the model when \(a_1\) and \(a_2\), and the recombination fractions at both loci are estimated compared to the best-fitting model when it is assumed that there are no families linked to one of the loci. Twice this difference is asymptotically distributed as a \(\chi^2\) with 2 degrees of freedom.

The predivided sample test was used to test for heterogeneity between the various subgroups described in Table 2 (29, 30). A Bonferroni correction was performed to correct for multiple testing.

\textbf{RESULTS} \textit{HPC1} and \textit{HPCX} Linkage Analyses. \textit{HPC1} and \textit{HPCX} linkage studies were carried out using the 57 informative families as described in the study by Xu et al. (14). \textit{HPC1} linkage studies of families demonstrated two-point parametric LOD scores ranging from \(-20.74\) (DXS1581) to \(-3.55\) (DXS230) \((\theta = 0)\) for the 39 markers used. Two-point parametric LOD scores for nine markers are shown in Table 1. Multipoint parametric LOD scores and nonparametric NPL scores from the GENEHUNTER program were also negative. Two-point nonparametric affected sibling pair tests using \textsc{analyze} were not significant.

Two-point parametric LOD scores for \textit{HPCX} are presented in the lower portion of Table 1. The maximum LOD of 2.05 occurred at \(\theta = 0.14\) with marker DXS1205. Five other markers on both sides of DXS1205 also showed positive LOD scores. Sliding four-point linkage using FASTLINK gave a maximum parametric multipoint LOD of 1.39 near DXS1205. The maximum NPL score obtained with GENEHUNTER for this region was 1.97 at position 2.62, corresponding to the location of marker DXS8043 \((P = 0.024)\). The two-point nonparametric affected sibling pair tests were most significant at DXS1205 \((P = 0.006)\).

\textbf{Stratified Analyses.} Using HOMOG3R analyses, significant evidence of heterogeneity was observed \((P < 0.05)\), with an estimated \(a_1\) of 0.5 for \textit{HPCX} and an \(a_2\) of 0.3 for \textit{HPC1}, assuming that there are two known loci (\textit{HPCX} and \textit{HPC1}) and \(a_3 = 1 - a_1 - a_2 = 0.20\) families unlinked to either loci. The test for two loci versus at most one locus was not significant, yielding most parsimonious estimates of \(a_1 = 0.45\), \(a_2 = 0\), and \(a_3 = 0.55\). We then performed subgroup analysis of these families by the age of diagnosis and by the number of affected individuals (Table 2).

For \textit{HPC1}, all subgroups defined in this manner had only strongly negative two-point LOD scores. In contrast, \textit{HPCX} LOD scores were positive in several subgroups. The 33 families classified as having NMM transmission of disease accounted for most of the positive LOD scores for the \textit{HPCX} region, with a maximum two-point LOD score of 2.16 \((\theta = 0.079)\) for DXS1205. In nonparametric two-point sibling pair analysis of the NMM transmission group, the LOD score for this marker was even higher (3.04; \(P < 0.00093\)). In contrast, the remaining families with male-to-male transmission had a peak two-point LOD score of 0.17 \((\theta = 0.49)\) at DXS1205 and a nonsignificant affected sibling pair test.

Further stratification of the data indicated that most of the \textit{HPCX} positivity came from the subgroup of families having NMM transmission and a late age of diagnosis (>65 years; Table 2). These families were also relatively small, having two to three affected cases. Late-onset NMM transmission families gave an overall two-point maximum LOD score of 3.12 \((\theta = 0.001)\) for DXS1205 and a nonparametric affected sibling pair LOD of 2.23 \((P < 0.00068)\).

To further evaluate whether the subgroups described above were the source of the observed heterogeneity, Morton's predivided sample test was used. After adjusting for multiple testing, only the division based on the presence of male-to-male transmission and age of diagnosis was of statistical significance (adjusted \(P < 0.05)\).

\textbf{DISCUSSION} The genetic epidemiological study of prostate cancer in the Finnish population resulted in the following major conclusions: \((a)\) there is significant evidence of genetic heterogeneity, suggesting that multiple loci are likely to contribute to prostate cancer even in the homogeneous Finnish population; \((b)\) most evidence for linkage to the \textit{HPCX} region on Xq27–q28 came...
from families with NMM transmission and late age (≥65 years) of diagnosis of prostate cancer. This suggests that HPCX-linked families are very different from the HPC1 families that are characterized by ≥5 affected cases/family, male-to-male transmission, and early age of diagnosis (11, 18, 19); and (c) no evidence of HPC1 linkage was seen in the entire material or in any subgroup of the Finnish prostate cancer patients. Therefore, most of the large, early age of diagnosis Finnish HPC families still remain unaccounted for with these two loci.

In the recent multicenter multinational study (14) involving HPC families from the United States (Johns Hopkins University and Mayo Clinic), Sweden, and Finland, a highly significant two-point LOD score of 4.6 was observed for the HPCX locus. In the present analyses of the Finnish families, stronger evidence for HPCX linkage was observed by sibling pair analysis than by two-point parametric analyses. The peak sibling pair LOD score was 3.04 (P < 0.000093) as compared with a maximum parametric two-point LOD of 2.16 in the NMM transmission group, suggesting that the model used for the parametric analyses may not be optimal. However, this finding was reversed for the NMM transmission/late diagnosis subgroup, with parametric LOD being more significant than the affected sibling pair LOD. This may be due to differences in the way these analyses deal with genetic heterogeneity or to the fact that the model for linkage was derived from segregation analyses (2) that suggested a rare locus with a high penetrance at an early age of diagnosis. Single locus segregation analyses cannot distinguish between different loci or their separate age-dependent penetrance functions. Furthermore, most investigators studying HPC have actively tried to collect only large, early age of diagnosis families, often with paternal transmission. Therefore, the finding of HPCX linkage in late diagnosis families would suggest that the linked fraction could be larger if prostate cancer families were ascertained regardless of family size, transmission pattern, or age of diagnosis. The HPCX gene could therefore have a significantly higher population frequency than estimated based on the original study by Xu et al. (14), where the family material was enriched with extended prostate cancer families with early age of onset.

In our recent population-based analysis of cancer risks among first-degree relatives of prostate cancer patients in Finland (31), we found a significantly elevated risk of prostate cancer in the relatives of the prostate cancer patients. As expected, increased risk of prostate cancer was seen in relatives of patients with a young age of diagnosis (<65 years). However, the risk was also elevated for relatives of patients ≥80 years but not in those diagnosed around the average age of diagnosis of prostate cancer (70–80 years). This epidemiological observa-
tion suggests the involvement of a specific genetic factor contributing to late age of prostate cancer diagnosis. Based on the results of this study, it is possible that this increased genetic risk of late diagnosis prostate cancer may be associated with the effects of the $HPCX$ gene.

Our negative results for $HPC1$ linkage are compatible with those of recent studies by McIndoe et al. (9), who analyzed 49 families from Seattle, Washington, as well as those of Eeles et al. (10), who analyzed 136 families from the United Kingdom; Quebec, Canada; and Texas. The results by Cooney et al. (7) and by Hsieh et al. (8) confirmed the presence of the $HPC1$ locus by reporting borderline significant linkage. Gronberg et al. (18, 19) reported that almost all of the evidence of linkage in $HPC1$ families came from those families with an early age of diagnosis or with five or more affected cases. However, in our study, stratification of the families by age and size did not significantly increase the $HPC1$ LOD scores.

In conclusion, there is significant evidence of heterogeneity in the loci causing prostate cancer, even in the genetically homogeneous Finnish population. The $HPCX$ locus on Xq27–q28 seems to explain a large portion of the Finnish HPC cases, especially among families with NMM transmission and late age of diagnosis. This may suggest that the $HPCX$ gene has a low penetrance and perhaps a higher prevalence than previous studies have suggested. In contrast, $HPC1$ is not likely to be the major locus contributing to HPC in Finland. Therefore, there are many Finnish prostate cancer families, especially those with many affected cases diagnosed at an early age, in which genes other than $HPC1$ or $HPCX$ are likely to be important.

ACKNOWLEDGMENTS

We thank all participating families for excellent cooperation. We thank Markku Kuorilehto, Kaisa Vaherto, Lila Hakala, Isaac Amudson, Derrick Gildea, Diana Freas-Lutz, and Anne Van Aucken for assistance.

REFERENCES


A Genetic Epidemiological Study of Hereditary Prostate Cancer (HPC) in Finland: Frequent HPCX Linkage in Families with Late-onset Disease

Johanna Schleutker, Mika Matikainen, Jeffrey Smith, et al.


Updated version
Access the most recent version of this article at:
http://clincancerres.aacrjournals.org/content/6/12/4810

Cited articles
This article cites 28 articles, 8 of which you can access for free at:
http://clincancerres.aacrjournals.org/content/6/12/4810.full.html#ref-list-1

Citing articles
This article has been cited by 12 HighWire-hosted articles. Access the articles at:
/content/6/12/4810.full.html#related-urls

E-mail alerts
Sign up to receive free email-alerts related to this article or journal.

Reprints and Subscriptions
To order reprints of this article or to subscribe to the journal, contact the AACR Publications Department at pubs@aacr.org.

Permissions
To request permission to re-use all or part of this article, contact the AACR Publications Department at permissions@aacr.org.