HOXB13 G84E Mutation in Finland: Population-Based Analysis of Prostate, Breast, and Colorectal Cancer Risk

Virpi H. Laitinen1, Tiina Wahlfors1, Leena Saaristo1, Tommi Rantaperä1, Liisa M. Pelttari6, Outi Kilpivaara7, Satu-Leena Laasanen2,3, Anne Kallioniemi1, Heli Nevanlinna6, Lauri Aaltonen7, Robert L. Vessella9, Anssi Auvinen4, Tapio Visakorpi1, Teuvo L.J. Tammela5, and Johanna Schleutker1,8

Introduction

In 2010, more than 4,700 Finnish men were diagnosed with prostate cancer and 847 died of it. These figures make the disease the most commonly diagnosed cancer in Finland and the second most common cause of cancer-related death (1). Despite its high incidence and mortality rates, the exact molecular mechanisms underlying the initiation and progression of prostate cancer still remain largely unknown.

Worldwide, compelling evidence has accumulated in favor of a significant but heterogeneous genetic component in prostate cancer susceptibility. On the basis of twin studies, heritability has been estimated as high as 16% to 45% (2, 3). However, the genetics of prostate cancer has proven hard to dissect. So far, only a few risk genes have been identified, although approximately 40 loci have been associated to genetic susceptibility (4, 5). Rare Mendelian genes with high penetrance, such as ribonuclease L [RNASEL (MIM 180435); ref. 6], explain perhaps 5% of prostate cancer susceptibility, whereas the more common genetic variants found in genome-wide association studies (GWAS) explain only approximately 25% of familial risk (7). Although GWAS have discovered many loci associated with prostate cancer risk, single-nucleotide polymorphisms (SNP) related to clinical outcome, that is, disease aggressiveness, have not been found. Consequently, there is renewed interest in family studies...

Note: Supplementary data for this article are available at Cancer Epidemiology, Biomarkers & Prevention Online (http://cebp.aacrjournals.org/).

V.H. Laitinen and T. Wahlfors contributed equally to this work.

Corresponding Author: Johanna Schleutker, Medical Biochemistry and Genetics, Institute of Biomedicine, Kliinamyllynkatu 10, FI-20014 University of Turku, Turku, Finland; Phone: 358-2-3373453; Fax: 358-2-2301280; E-mail: Johanna.Schleutker@utu.fi

doi: 10.1158/1055-9965.EPI-12-1000-T

©2012 American Association for Cancer Research.

Research Article
because of the type of information they offer, especially when trying to isolate rare high-impact variants.

Linkage analyses of hereditary prostate cancer (HPC) families have detected a significant signal at the chromosomal region of 17q21-22 in both North American and Finnish populations (8–10). Recently, Ewing and colleagues (11) used targeted next generation sequencing of this region to identify a rare but recurrent germline missense mutation c.251G→A (p.G84E, rs138213197) in the first exon of the homeobox B13 [HOXB13 (MIM 604607)] gene. This mutation was associated with a significantly increased risk of early-onset, familial prostate cancer.

The HOXB13 gene belongs to a group of highly conserved homeobox genes that are essential for vertebrate embryogenesis. In humans, there are 4 HOX gene clusters (A–D) in separate chromosomes, and the HOXB cluster is localized in the 17q21-22 region (12). HOXB13 is highly expressed in both normal and cancerous prostate. The HOXB13 protein is a sequence-specific, 284-amino acid transcription factor that interacts with androgen receptor and has an important role in prostate development (13). It has been shown to regulate cellular responses to androgens, such as promotion of androgen-independent growth in prostate cancer cell lines (14) by activating or repressing the expression of most androgen receptor-responsive genes (15). In addition to prostate cancer, HOXB13 has also been shown to have a role as a tumor suppressor in primary colorectal cancers (CRC; ref. 16), and it predicts breast cancer recurrence (17) and tamoxifen response (18).

Given the linkage evidence to the 17q21-22 locus in Finnish prostate cancer families (10), and the exceptionally high proportion of Finnish families with the G84E mutation, as shown in a recent International Consortium for Prostate Cancer Genetics (ICPCG) study (19), we genotyped the G84E mutation in 4,571 prostate cancer cases and 5,467 controls, together with 516 benign prostate hyperplasia samples, 10 prostatic cell lines, and 19 LuCaP xenografts. We also investigated its role in prostate cancer case-control studies, and in 2 separate lines, 22Rv1 and CWR22R2Pc, derived from CWR22, and 19 LuCaP xenografts. DU145, PC-3, VCaP, and 2 separate lines, 22Rv1 and CWR22R2Pc, were obtained from the American Type Culture Collection. CWR22R2Pc was provided by Dr. Marja Nevalainen (Thomas Jefferson University, Philadelphia, PA). LACPC was obtained from Dr. Charles Sawyers (University of California at Los Angeles, Los Angeles, CA). VCaP and DuCaP were obtained from Dr. Jack Schalken (Radboud University Nijmegen Medical Center, Nijmegen, the Netherlands). PrEC was obtained from Lonza (Lonza Walkersville). EP156T was kindly provided by Dr. Varda Rotter (Weizmann Institute of Science, Rehovot, Israel).

**Materials and Methods**

**Study subjects**

All cancer cases and controls genotyped in this study were of Finnish origin. Written informed consent was obtained from each study subject. The cancer diagnosis was confirmed from medical records. The study protocol was approved by the research ethics committee at Pirkanmaa Hospital District (Tampere, Finland) and by the National Supervisory Authority for Welfare and Health. Different sample types included in the analyses are presented in the Supplementary Table S1.

**Prostate cancer.** A total of 4,571 Finnish prostate cancer cases were genotyped. Of these, 3,197 unselected cases were collected in the Pirkanmaa Hospital District. Another unselected set of subjects consisted of 1,184 Finnish cancer cases recruited by the Finnish arm of The European Randomized Study of Screening for Prostate Cancer. This study was initiated in the early 1990s to evaluate the effect of prostate-specific antigen (PSA) screening on death rates from prostate cancer (20). In addition to the unselected cases, genotype data for 190 index cases derived from Finnish prostate cancer families were included. The collection of the Finnish familial prostate cancer families has been described previously (21, 22). All of the 190 families used in this study had at least 2 members affected by prostate cancer, with the majority of families (n = 151) having at least 3 confirmed cases. All affected persons were either first- or second-degree relatives of the index cases. Only an index case was originally genotyped, and additional individuals were studied only to confirm segregation of the mutation. Seventy-six index individuals overlapped with those genotyped in the large multinational ICPCG study (19). To investigate the cosegregation of the G84E mutation in nonoverlapping, mutation-positive families, additional healthy and affected family members were genotyped. The most representative clinical features for each of the 3 prostate cancer patient groups are summarized in Table 1.

Germline DNA was also available from 516 clinically and pathologically defined cases of benign prostate hyperplasia from the Urology Outpatient Clinic in Tampere University Hospital (Tampere, Finland; BPH; samples collected in 1998–2004): 254 of these cases were later diagnosed with prostate cancer. In addition to germline DNA samples, the G84E status was analyzed in 2 normal cell lines (PrEC and EP156T), 8 prostate cancer cell lines (LAPC4, LNCaP, DuCaP, DU145, PC-3, VCaP, and 2 separate lines, 22Rv1 and CWR22R2Pc, derived from CWR22), and 19 LuCaP xenografts. DU145, PC-3, 22Rv1, and LNCaP were obtained from the American Type Culture Collection. CWR22R2Pc was provided by Dr. Marja Nevalainen (Thomas Jefferson University, Philadelphia, PA). LACPC was obtained from Dr. Charles Sawyers (University of California at Los Angeles, Los Angeles, CA). VCaP and DuCaP were obtained from Dr. Jack Schalken (Radboud University Nijmegen Medical Center, Nijmegen, the Netherlands). PrEC was obtained from Lonza (Lonza Walkersville). EP156T was kindly provided by Dr. Varda Rotter (Weizmann Institute of Science, Rehovot, Israel).

**Breast cancer.** Tampere subgroup: 86 index cases from well-characterized high-risk breast cancer families were genotyped. In these families, patients with breast cancer were diagnosed at an early age or at least 3 first-degree relatives had breast or ovarian cancer. The sample set is described in more detail elsewhere (23). In addition, 410 unselected Finnish breast cancer cases, described previously by Syrjäkoski and colleagues...
were analyzed in this study. Helsinki subgroup: genotyping was conducted for 237 familial and 253 patients with sporadic breast cancer. The patients with familial breast cancer were collected at the Helsinki University Central Hospital Departments of Oncology and Clinical Genetics (Helsinki, Finland) as previously described (25). They had a strong familial background of breast cancer with 3 or more breast or ovarian cancers among first- or second-degree relatives, including the proband. The patients with sporadic breast cancer were part of an unselected series collected at the Helsinki University Central Hospital Department of Surgery in 2001 to 2004 (26). In both the Tampere and Helsinki subgroups, all of the patients with familial breast cancer tested negative for \textit{BRCA1} (MIM 113705) and \textit{BRCA2} (MIM 600185) founder mutations.

Colorectal cancer. The sample set consisted of 442 CRC cases belonging to a Finnish population-based series of 1,042 patients with CRC. Fifty-seven CRC cases were classified as familial, having at least 1 first-degree relative with CRC. The data were collected prospectively at 9 Finnish central hospitals between 1994 and 1998 as described by Aaltonen and colleagues (27) and Salovaara and colleagues (28).

\textbf{Controls.} All control subjects for breast cancer and CRC, as well as the population control group for prostate cancer, consisted of population-matched healthy individuals of ages between 18 and 65 years. The blood DNA samples were obtained from the Finnish Red Cross Blood Transfusion Service. Population control subjects for prostate cancer included 923 anonymous male blood donors. Breast cancer controls for the Tampere and Helsinki subgroups comprised 900 and 549 anonymous, healthy female blood donors, respectively. Blood-derived DNA samples from an additional 459 healthy individuals were used as CRC controls. Prostate cancer control subjects (\(n = 4,544\)) belonging to the screening trial control group were derived from the Finnish arm of the European Randomized Study of Screening for Prostate Cancer (20). All members of this control group were age-standardized (from 59 to 79 years) healthy men who had undergone PSA screening. The

\begin{table}[h]
\centering
\caption{Clinical characteristics of the 3 prostate cancer patient groups analyzed in this study}
\begin{tabular}{|c|c|c|c|}
\hline
\textbf{Characteristics} & \textbf{Variables} & \textbf{All FAM}\textsuperscript{a} (n) & \textbf{UNS}\textsuperscript{b} (n) & \textbf{SCRCase}\textsuperscript{c} (n) \\
\hline
Average age at onset & Age at onset (y) & 62.8 (9) & 68.6 (234) & 67.0 (152) \\
Prostate specific antigen & \(\leq4.0\) ng/mL & 5.4 (59) & 8.0 (740) & 12.9 (211) \\
& 4.1–9.9 ng/mL & 35.5 (29) & 43.0 (1,258) & 61.1 (719) \\
& 10.0–19.9 ng/mL & 26.5 (44) & 25.3 (740) & 17.9 (211) \\
& 20.0–49.9 ng/mL & 21.1 (35) & 13.3 (389) & 6.5 (77) \\
& 50.0–99.9 ng/mL & 4.8 (8) & 4.7 (137) & 0.8 (9) \\
& \(\geq100\) ng/mL & 6.6 (11) & 5.7 (167) & 0.8 (9) \\
Missing data & 12.4 (24) & 8.5 (272) & 0.6 (7) \\
Primary treatment & Prostatectomy & 46.0 (82) & 34.7 (1,030) & 23.0 (32) \\
& Radiotherapy & 16.9 (30) & 18.4 (546) & 39.1 (55) \\
& Hormonal therapy & 30.9 (55) & 37.9 (1,124) & 9.8 (14) \\
& Active surveillance & 4.5 (8) & 5.6 (166) & 14.7 (21) \\
& Brachytherapy & 1.7 (3) & 2.9 (86) & 12.6 (18) \\
& Cystectomy & 0.5 (15) & 0.7 (1) \\
& Missing data & 6.3 (12) & 7.2 (230) & 88.0 (1,043) \\
Gleason score for biopsy & 3 & 2.7 (4) & 2.7 (72) & 2.5 (29) \\
& 4 & 11.6 (17) & 4.1 (109) & 8.8 (102) \\
& 5 & 15.7 (23) & 11.4 (304) & 12.3 (143) \\
& 6 & 32.7 (48) & 36.5 (972) & 42.8 (496) \\
& 7 & 22.4 (33) & 27.8 (740) & 25.2 (292) \\
& 8 & 8.2 (12) & 8.4 (224) & 6.0 (70) \\
& 9 & 6.0 (9) & 8.4 (224) & 2.0 (23) \\
& 10 & 0.7 (1) & 0.7 (19) & 0.4 (5) \\
& Missing data & 22.6 (43) & 16.7 (534) & 2.1 (25) \\
Progression & PSA progression & 13.7 (26) & 30.9 (988) & 8.8 (104) \\
Cause of death & Overall deaths & 42.1 (80) & 43.1 (1,378) & 8.8 (104) \\
& Prostate cancer & 35.0 (67) & 26.6 (850) & 5.7 (67) \\
\hline
\textsuperscript{a}All FAM, familial index cases from all 190 Finnish prostate cancer families. \\
\textsuperscript{b}UNS, unselected cases. \\
\textsuperscript{c}SCRCase, screening trial cases.
\end{tabular}
\end{table}
disease status is annually evaluated from the records of the Finnish Cancer Registry.

**SNP genotyping**
Prostate and breast cancer samples, as well as the cell lines and xenografts, were genotyped for the G84E mutation (rs138213197) using a Custom TaqMan SNP assay (Applied Biosystems/Life Technologies) according to the manufacturer’s instructions. Duplicate test samples and 4 negative controls were included in each 384-well plate.

BPH samples were genotyped by the Technology Centre, Institute for Molecular Medicine Finland (FIMM), University of Helsinki (Helsinki, Finland) using the MassARRAY iPLEX platform (Sequenom, Inc.).

**DNA sequencing**
The mutation was confirmed in a selected set of prostate and breast cancer samples by standard Sanger sequencing using an ABI PRISM BigDye Termination Cycle Sequencing Ready Reaction Kit (Applied Biosystems/Life Technologies). CRC cases and controls were genotyped by sequencing the coding exons of HOXB13. CRC DNA from all 7 G84E carriers was extracted from freshly frozen tissue, and the coding region of HOXB13 was sequenced for LOH analysis. Primer sequences are available upon request.

**Statistical analysis**
The statistical significance of the association between the HOXB13 G84E mutation and prostate cancer, breast cancer, or CRC was evaluated using a Fisher exact test, implemented in PLINK (29) and GraphPad Prism 5.02 (GraphPad Software, Inc.) softwares. In addition to case-control comparisons, case–case analyses evaluated the impact of the mutation to the clinical features (PSA, Gleason score, age at onset, and progression). All P values were 2-sided. The association between the mutation and overall survival was analyzed using a Cox model. Survival time (years) after diagnosis was compared between carriers and noncarriers. Statistical significance of the survival differences between the G84E carriers and noncarriers were calculated with log-rank and Gehan–Breslow–Wilcoxon tests.

**In silico pathogenicity prediction**
The pathogenicity of G84E was evaluated by using a machine learning-based method PON-P (Pathogenic- or-Not Pipeline; ref. 30) that includes 6 independent tolerance predictors (SIFT, PolyPhen-2, SNAP, PHD-SNP, PANTHER, and SNP&GO) and the pipeline’s own meta-predictor, which integrates the output of 5 predictors (SIFT, SNAP, PolyPhen-2, PHD-SNP, and I-Mutant-3) as the input to make the pathogenicity prediction. Two additional programs, NetSurfP (31) and SABLE 2 (32), were used to investigate the sequence environment of G84. These programs predict features such as the secondary structure, transmembrane regions, and the relative solvent accessibilities of the amino acids based on the amino acid sequence of the given protein. Protein stability was examined using the I-Mutant-3 (33) and MuPro (34) programs, also implemented in PON-P, and an additional program called iPTREE-STAB (35).

**Results**

**Prostate cancer**
The overall call rate of the mutation site among prostate cancer samples was 99.8%, and the average concordance of duplicated samples was 99.9%. The G84E mutation was in Hardy–Weinberg equilibrium in both cases and controls. The overall minor allele frequency in the entire sample set was 1.9%. The G84E mutation was detected in 188 subjects, of which 160 were patients with prostate cancer (carrier frequency 3.5%) and 28 were healthy controls (0.5%). Of the cases carrying G84E, 3.4% (155 of 4,571) were heterozygous, and 0.1% (5 of 4,571) were homozygous for the mutation. The observed G84E carrier frequency for the unselected cases from the Pirkanmaa Hospital District was 3.6% (114 of 3,197), but the frequency was only 2.2% (26 of 1,184) for the screening trial patients. The highest carrier frequency of 8.4% (16 of 190) was observed among index patients with a positive family history of prostate cancer. In this group, the case subjects were significantly more likely to carry the mutation compared with population controls [carrier frequency 1.0%; P = 2.318 e-18; OR, 8.8; 95% confidence interval (CI), 4.9–15.7]. In addition, statistically significantly higher carrier frequencies were detected among cases with a positive family history of prostate cancer compared with unselected cases (P = 1.982e-06; OR, 2.5; 95% CI, 1.7–3.6). Table 2 summarizes the results of the association analyses.

Case–case analysis of the G84E mutation in relation to clinical features of prostate cancer revealed a significant association with younger age (<55 years) at diagnosis (P = 0.0008; OR, 2.0; 95% CI, 1.3–3.0). Likewise, carrier frequency was significantly higher among men with serum PSA concentrations 20 ng/mL or more at diagnosis (P = 0.006; OR, 1.4; 95% CI, 1.1–1.9). However, no evidence for an association with tumor grade (Gleason score >3 vs. ≤3) or prostate cancer progression based on elevated PSA (present vs. absent) was observed (Table 3). Gleason 7 was left out of the analysis to decrease the heterogeneity of the compared groups because it was not possible to differentiate Gleason scores of 7 as either “3+4” or “4+3.” A slightly but not significantly poorer overall survival (HR, 1.16; 95% CI, 0.9–1.5) was observed in mutation carriers relative to noncarriers. A significantly elevated risk of prostate cancer was found to be associated with the G84E mutation in a group of patients with prior BPH diagnosis (P = 0.01084; OR, 4.6; 95% CI, 1.3–16.2). Interestingly, none of the prostate cell lines or LuCaP xenografts carried the A allele of the mutation.

Of the 190 Finnish prostate cancer families included in this study, 32 indexes (17%) were found to be carriers of the G84E mutation. Fifteen of these 32 families
overlapped with the ICPCG dataset (19). Cosegregation of G84E with prostate cancer in the remaining 17 families was assessed by genotyping an additional 28 healthy and 37 affected family members, for whom DNA samples were available. In 11 of 17 families, the G84E mutation cosegregated with the disease in 20 genotyped cases, representing 53% of the total cancer cases in these families. Segregation of the mutation with the disease was incomplete in 6 families, as both unaffected mutation carriers (n = 5) and mutation-negative

### Table 2. Summary of results obtained from the case–control and case–case association analyses of the G84E mutation and prostate cancer risk

<table>
<thead>
<tr>
<th>Prostate cancer datasets</th>
<th>F_A%</th>
<th>F_U%</th>
<th>P value</th>
<th>OR (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>All cases and controls</td>
<td>3.5</td>
<td>0.5</td>
<td>1.1 x 10^-62</td>
<td>7.1 (5.5–9.3)</td>
</tr>
<tr>
<td>UNS vs. Pco a</td>
<td>3.6</td>
<td>1.0</td>
<td>1.8 x 10^-8</td>
<td>3.6 (2.2–5.7)</td>
</tr>
<tr>
<td>UNS vs. SCRco b</td>
<td>3.6</td>
<td>0.3</td>
<td>6.2 x 10^-57</td>
<td>13.4 (8.9–20.3)</td>
</tr>
<tr>
<td>SCRcase c vs. SCRco</td>
<td>2.2</td>
<td>0.3</td>
<td>1.1 x 10^-23</td>
<td>8.0 (4.9–12.9)</td>
</tr>
<tr>
<td>SCRcase vs. Pco</td>
<td>2.2</td>
<td>1.0</td>
<td>0.004603</td>
<td>2.1 (1.2–3.6)</td>
</tr>
<tr>
<td>All FAM d vs. Pco</td>
<td>8.4</td>
<td>1.0</td>
<td>2.3 x 10^-18</td>
<td>8.8 (4.9–15.7)</td>
</tr>
<tr>
<td>All FAM vs. SCRco</td>
<td>8.4</td>
<td>0.3</td>
<td>1.8 x 10^-89</td>
<td>33.1 (19.4–56.5)</td>
</tr>
<tr>
<td>All FAM vs. UNS</td>
<td>8.4</td>
<td>3.6</td>
<td>2.0 x 10^-6</td>
<td>2.5 (1.7–3.6)</td>
</tr>
<tr>
<td>All FAM vs. SCRcase</td>
<td>8.4</td>
<td>2.2</td>
<td>4.2 x 10^-11</td>
<td>4.2 (2.6–6.6)</td>
</tr>
<tr>
<td>FAM e vs. Pco</td>
<td>7.9</td>
<td>1.0</td>
<td>1.5 x 10^-13</td>
<td>8.2 (4.3–16.0)</td>
</tr>
<tr>
<td>FAM vs. SCRco</td>
<td>7.9</td>
<td>0.3</td>
<td>4.4 x 10^-63</td>
<td>31.1 (16.7–57.8)</td>
</tr>
<tr>
<td>FAM vs. UNS</td>
<td>7.9</td>
<td>3.6</td>
<td>0.0006835</td>
<td>2.3 (1.4–3.8)</td>
</tr>
<tr>
<td>FAM vs. SCRcase</td>
<td>7.9</td>
<td>2.2</td>
<td>2.6 x 10^-7</td>
<td>3.9 (2.2–6.8)</td>
</tr>
<tr>
<td>BPHcase g vs. BPHco h</td>
<td>2.6</td>
<td>0.6</td>
<td>0.011</td>
<td>4.6 (1.3–16.2)</td>
</tr>
</tbody>
</table>

NOTE: F_A and F_U represent the frequencies of G84E carriers among affected and unaffected subjects, respectively. All P values are statistically significant.

aUNS, unselected cases. bPco, population controls. cSCRco, screening trial controls. dSCRcase, screening trial cases.

All FAM, familial index cases from all 190 Finnish prostate cancer families. eFAM, familial index cases from the 114 Finnish prostate cancer families analyzed in this study (the 76 familial cases overlapping with the ICPCG dataset are omitted).

fBPHcase, patients with BPH with a later diagnosis of prostate cancer. gBPHco, patients with BPH with no diagnosis of prostate cancer.

### Table 3. Summary of results obtained from the case–case association analysis of the G84E mutation and selected clinical features

<table>
<thead>
<tr>
<th>Age at diagnosis</th>
<th>G84E carriers% (n)</th>
<th>G84E noncarriers% (n)</th>
<th>P value</th>
<th>OR (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>≤55 y</td>
<td>6.25 (13)</td>
<td>93.75 (195)</td>
<td>0.0007959</td>
<td>2.0 (1.3–3.0)</td>
</tr>
<tr>
<td>&gt;55 y</td>
<td>3.40 (148)</td>
<td>96.60 (4,209)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PSA at diagnosis</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>≥20 ng/mL</td>
<td>4.56 (39)</td>
<td>95.44 (816)</td>
<td>0.006187</td>
<td>1.4 (1.1–1.9)</td>
</tr>
<tr>
<td>&lt;20 ng/mL</td>
<td>3.19 (110)</td>
<td>96.81 (3,336)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PSA progression</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Present</td>
<td>3.76 (39)</td>
<td>96.24 (997)</td>
<td>0.5034</td>
<td>1.1 (0.8–1.4)</td>
</tr>
<tr>
<td>Absent</td>
<td>3.51 (124)</td>
<td>96.49 (3,406)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gleason score</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>≥8</td>
<td>4.04 (22)</td>
<td>95.96 (523)</td>
<td>0.09918</td>
<td>1.3 (1.0–1.9)</td>
</tr>
<tr>
<td>≤6</td>
<td>3.11 (70)</td>
<td>96.89 (2,182)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

NOTE: The statistically significant P values are italicized.
patients ($n = 7$) were observed. (segregation presented in the Supplementary Table S2).

**Breast cancer**

The G84E mutation was identified in 6 of 323 (1.9%) of the familial breast cancer cases, 10 of 663 (1.5%) of the sporadic breast cancer cases and 16 of 1,449 (1.1%) of the population controls. Case-control association analyses were conducted for the entire dataset and separately for both subgroups (familial and sporadic), but no statistically significant differences in carrier frequencies between cases and controls were observed (data not shown). However, in the high-risk, familial Tampere subgroup, the frequency of G84E carriers was 3.5%, a figure similar to the number of mutation carriers among the Finnish patients with prostate cancer. The OR of 3.2 (95% CI, 0.9–11.9) is suggestive of an association between G84E and increased breast cancer risk.

**Colorectal cancer**

Of the 442 patients with CRC, 7 (1.6%) were identified as carriers of the HOXB13 G84E mutation, and none of these were familial. No evidence of allelic imbalance was observed in the LOH analysis of the G84E-positive tumors. In a case-control association analysis, the difference in carrier frequencies was nonsignificant between cases (1.6%) and population controls (0.9%).

**In silico analysis**

To further explore the mechanistic function of G84E in prostate cancer risk, in silico analyses were conducted. In a pathogenicity prediction analysis, the G84E mutation was predicted to be deleterious in 5 of 6 of the tolerance predictors included in PON-P. However, the pipeline's own meta-predictor indicated the mutation to be tolerated. NetSurfP and SABLE 2 estimated glycine 84 to be located in a region buried inside the protein structure. Moreover, the sequence surrounding glycine 84 was found to be relatively hydrophobic, suggesting that G84 is located in the hydrophobic core of HOXB13. However, the applied programs gave conflicting results in protein stability tests. (all results presented in the Supplementary Tables S3–S5 and Supplementary Fig. S1).

**Discussion**

The present results validate an important role for HOXB13 G84E in prostate cancer predisposition. In the Finnish population, the mutation was detected in 3.5% of all cases and 8.4% of familial prostate cancer, which suggests that G84E may be the strongest genetic marker of prostate cancer reported to date. In the original article by Ewing and colleagues (11), the highest HOXB13 G84E carrier frequency of 3.1% was observed among men with a positive family history of prostate cancer and an early age (â‰¤ 55 years) at diagnosis. In Finland, the carrier frequency of the mutation among familial prostate cancer cases was almost 3-fold higher (8.4%), a frequency that is strikingly similar to the carrier frequency observed in Swedish prostate cancer families (8.2%). Moreover, both the Finnish and Swedish mutation carrier families share a common rare haplotype indicating a likely founder effect for the mutation (19). Founder mutations are typical for isolated populations, such as the Finnish population, and they may explain a major fraction of all mutations in specific genes (36, 37). In the Finnish population, strong founder mutations have been detected in breast cancers and CRCs (38, 39). In Finland, founder mutations are often present in geographic clusters when the birthplaces of ancestors are known (40). Here, however, the birthplaces of the grandparents of the G84E-positive patients did not show such a pattern, which may indicate a very old origin of the mutation.

Ewing and colleagues (11) reported control subject carrier frequencies to vary between 0.1% and 0.2%. In our study, the frequency distribution of the G84E mutation in different prostate cancer control groups ranged from 0.3% to 1.0%. The lowest frequency (0.3%) was detected in the age-matched, PSA-screened control group. The variations in carrier frequencies are explained by differences in age distributions between control groups, with the oldest subjects belonging to screening trial controls and the youngest to population controls.

When compared with patients with prostate cancer, G84E carrier frequencies were substantially lower in patients with breast cancer and CRC. On the basis of these results, the HOXB13 G84E mutation seems to be prostate cancer–specific, although this needs to be verified in larger breast and CRC datasets. However, it is noteworthy that all patients with G84E-positive familial breast cancer tested negative for the Finnish BRCA1/2 founder mutations, and the highest carrier frequency of 3.5% occurred within the high-risk Tampere subgroup among all studied patients with breast cancer.

Previously, similar prostate cancer–associated mutation carrier frequencies in the Finnish population have been obtained only for mutations 1100delC (3.3%) and 1157T (10.8%) in the checkpoint kinase 2 [CHEK2 (MIM 604373)] gene (41). Analogous to the current HOXB13 mutation, CHEK2 mutation frequencies in prostate cancer were significantly higher in populations from Northern and Eastern European countries as compared with North American populations, reflecting population-specific differences (42–44). CHEK2 is also a known breast cancer risk gene, and the frequency of 1100delC among patients with breast cancer varies similarly between European and North American populations (45, 46).

In previous linkage studies of the 69 Finnish HPC families, a strong signal was observed for the 17q21-22 region (10), and before the present study, G84E was thought to explain this finding. However, of the 32 G84E-positive families analyzed in this study, only 2 families showed linkage (LOD score >0.6) to chromosome 17, suggesting that the G84E-positive and linkage-contributing families are not overlapping. Moreover,
cosegregation with prostate cancer was not complete in many of the G84E-positive families, and incomplete penetrance and genetic heterogeneity were observed in 35.3% (6 of 17) of the families, which is consistent with the results of the ICPCG study (19). Of the 5 unaffected mutation carriers observed in this study, 3 were in their sixties and are therefore still at risk for the disease, but the 2 oldest carriers were already 80 and 87 years of age. Contrary to the results reported by Ewing and colleagues (11), we found 5 of the analyzed patients with prostate cancer to be homozygous for the rs138213197 A allele. Two of them represented familial prostate cancer (1 initially reported in the above-mentioned ICPCG study), whereas the other 3 were unselected cases. The 5 homozygous patients did not share any distinctive clinical features relating to disease aggressiveness.

Although G84E seems to explain a considerable fraction of Finnish familial prostate cancer, the linkage signal cannot be explained by HOXB13 alone and there must be other, yet unidentified genes and variants on chromosome 17 that are responsible for the remaining and quite substantial proportion of HPC cases in Finland. Because of the observed heterogeneity, we evaluated other cancers in the G84E-positive families. In these 32 families, 35 individuals were diagnosed with a cancer other than that of the prostate. Altogether, 17 different cancer types were detected in the patients (10 males and 25 females). No particular cancer type was over-represented. Another cancer was diagnosed in 5 of the patients with G84E-positive prostate cancer, and 5 females had a diagnosis of breast cancer.

Several studies have shown an increased risk of prostate cancer incidence among patients with BPH, although BPH is not considered a premalignant lesion (47, 48). Our collection of BPH cases, from years 1998 to 2004, has been followed-up since and almost half of these cases have been diagnosed with prostate cancer during this follow-up time. In this study, the aim was to assess whether the HOXB13 G84E mutation has a risk-associated role in prostate cancer occurrence in the BPH cohort. As shown, patients with BPH carrying the G84E mutation were at a significantly increased risk of developing prostate cancer as compared with noncarriers. Because all of these BPH cases were histologically confirmed, there is no chance for misclassification of clinical BPH. Furthermore, the relatively long follow-up time of 8 to 14 years enhances the reliability of the data. Histologic BPH is observed in 50% of men of ages 51 to 60 years and in 70% of men of ages 61 to 70 years (49). Genetic markers that can separate the patients with high-risk BPH from the considerably larger low-risk group would be desirable. Therefore, at least in Finland, G84E deserves serious attention, and genetic testing could be an option for patients with histologically confirmed BPH.

Although numerous genetic variants have been associated with prostate cancer predisposition, their roles as prognostic factors have been limited. Here, the G84E mutation was found to be associated with a high (≥20 ng/mL) PSA concentration at the time of diagnosis, providing evidence for the clinical relevance of G84E in the Finnish population. To our knowledge, this is the first time that G84E has been significantly associated with a clinical feature commonly considered a marker of aggressive disease. However, no difference in other clinical features related to disease aggressiveness, such as Gleason score or prostate cancer progression, was observed between mutation carriers and noncarriers. We also analyzed the association of G84E with overall survival, but the median survival period after prostate cancer diagnosis did not differ between carriers and noncarriers (data not shown). The association of G84E with PSA concentrations may perhaps be explained by a possible regulatory role of HOXB13 on androgen-responsive genes, which warrants further study.

Ewing and colleagues (11) analyzed tumor tissues obtained from G84E carriers and showed that these tumors maintain the expression of HOXB13, a finding consistent with the hypothesis that HOXB13 functions as an oncogene. We confirmed the observation of HOXB13 expression by analyzing tumor tissue from G84E carriers and noncarriers with immunohistochemistry (data not shown). The pathogenic role of the G84E mutation has not yet been shown by functional studies. We investigated the pathogenicity of G84E using diverse in silico predictors. On the basis of our results, it is possible that G84E affects protein stability because a small hydrophobic glycine is replaced with hydrophilic glutamate. To confirm the functionality, in vivo studies are needed.

In summary, the rare HOXB13 mutation has been shown to contribute to prostate cancer risk in Finland, confirming the high frequency of the G84E mutation in this Nordic population. The risk was highest in familial prostate cancer cases. No such effect was observed for CRC, but a suggestive risk effect was detected in a subset of familial breast cancer cases. These results indicate that the G84E mutation may have clinical implications for prostate cancer management in the Finnish population.

Disclosure of Potential Conflicts of Interest
No potential conflicts of interest were disclosed.

Authors’ Contributions
Conception and design: V.H. Laitinen, T. Wahlfors, T.L.J. Tammela, J. Schleutker
Acquisition of data (provided animals, acquired and managed patients, provided facilities, etc.): L. Saaristo, L.M. Pettari, O. Kilpivaara, S.-L. Laasanen, A. Kallioniemi, H. Nevanlinna, L. Aaltonen, R.L. Vessella, A. Auvinen, T. Visakorpi, T.L.J. Tammela, J. Schleutker
Analysis and interpretation of data (e.g., statistical analysis, biostatistics, computational analysis): V.H. Laitinen, T. Wahlfors, L. Saaristo, T. Rantapero, O. Kilpivaara, H. Nevanlinna, A. Auvinen, T. Visakorpi, T.L.J. Tammela, J. Schleutker
Writing, review, and/or revision of the manuscript: V.H. Laitinen, T. Wahlfors, L. Saaristo, T. Rantapero, L.M. Pettari, A. Kallioniemi, H. Nevanlinna, L. Aaltonen, R.L. Vessella, A. Auvinen, T. Visakorpi, T.L.J. Tammela, J. Schleutker
Published OnlineFirst January 4, 2013; DOI: 10.1158/1055-9965.EPI-12-1000-T
Ewing CM, Ray AM, Lange EM, Zuhlke KA, Robbins CM, Tembe WD,
and Sini Karjalainen for technical assistance in sequencing the CRC samples.

Carl Blomqvist, and Karl von Smitten, as well as Sara Vilske, for their help
authors also thank Kirsi Kuusisto for her contribution to the collection of the
T.L.J. Tammela

Acknowledgments

The authors thank the patients and families who participated in this
study. Riitta Vaalavuo and Riina Liikala are thanked for assistance. The authors also thank Kirs Kuusisto for her contribution to the collection of the
T. Wahlfors, T.L.J. Tammela, J. Schleutker

References

M, et al. Environmental and heritable factors in the causation of
cancer—analyses of cohorts of twins from Sweden, Denmark, and
prostate, breast, and colorectal cancer among Nordic twins. Bio-
4. Varghese JS, Easton DF. Genome-wide association studies in com-
mon cancers—what have we learnt? Curr Opin Genet Dev 2010;20:
201–9.
S, et al. Genome-wide association study identifies new prostate
Germline mutations in the ribonuclease L gene in families showing
et al. Seven prostate cancer susceptibility loci identified by a multi-
JE, et al. Combined genome-wide scan for prostate cancer suscepti-
A combined genomewide linkage scan of 1,233 families for prostate
cancer-susceptibility genes conducted by the international consor-
Genome-wide linkage scan for prostate cancer susceptibility in Fin-
land: evidence for a novel locus on 2q37.3 and confirmation of signal
11. Ewing CM, Ray AM, Lange EM, Zuhlke KA, Robbins CM, Tembe WD,
et al. Germline mutations in HOXB13 and prostate-cancer risk. N Engl
201.
13. Luang L, Pu Y, Heppos D, Danielpou D, Prins GS. Posterior Hox gene
expression and differential androgen regulation in the developing and
promotes androgen independent growth of LNCaP prostate cancer
The homeodomain protein HOXB13 regulates the cellular response to
HOXB13, a target of DNMT3B, is methylated at an upstream CpG
island, and functions as a tumor suppressor in primary colorectal
17. Jerevall PL, Brännmosser S, Strand C, Gruveberger-Saal S, Malmström
cancer—individual roles for HOXB13 and IL17R in prediction of
Predictive relative relevance of HOXB13 protein expression for tamoxifen
is a susceptibility gene for prostate cancer: results from the Interna-
tional Consortium for Prostate Cancer Genetics (ICPCG). Hum Genet
20. Schroder FH, Hugosson J, Roobol MJ, Tammela TL, Ciatto S, Nelen, V,
et al. Screening and prostate-cancer mortality in a randomized Euro-
21. Carter BS, Beatty TH, Steinberg GD, Childs B, Walsh PC. Mendelian
inheritance of familial prostate cancer. Proc Natl Acad Sci U S A
Kainu T, et al. A genetic epidemiological study of hereditary prostate
cancer (HPC) in Finland: frequent HPCX linkage in families with late-
for BRCA1, BRCA2, CHEK2, PALB2, BRIP1, RAD50, and CDH1
mutations in high-risk Finnish BRCA1/2-founder mutation-negative
breast and/or ovarian cancer families. Breast Cancer Res 2011;13:
R20.
mutations in 1035 unselected Finnish breast cancer patients. J Natl
25. Erola H, Blomqvist C, Pukkala E, Pyhronen S, Nevanlinna H. Familial
breast cancer in southern Finland: how prevalent are breast cancer
families and can we trust the family history reported by patients? Eur J
K, et al. NAD(P)H:Quinone oxidoreductase 1 NQO1
71–gene
–
201.
P, et al. Incidence of hereditary nonpolyposis colorectal cancer and the
inen M, et al. Population-based molecular detection of
hereditary nonpolyposis colorectal cancer. J Clin Oncol 2000;18:
2193–200.
29. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D,
et al. PLINK: a tool set for whole-genome association and population-
integrated predictor for pathogenicity of missense variants. Hum Mutat
2012;33:1166–74.
31. Petersen B, Petersen TN, Andersen P, Nielsen M, Lundegaard C. A
generic method for assignment of reliability scores applied to solution
32. Adamczak R, Porollo A, Meller J. Combining prediction of secon-
dary and solvent accessibility in proteins. Proteins 2005;59:
467–75.

Grant Support

This work was supported by the Sigrid Juselius Foundation, the
Finnish Cancer Organizations, and the Competitive Research Fund-
ing of the Pirkanmaa Hospital District (959/94 and 9%39/9) grants to
J. Schleutker; the Academy of Finland grants 116437 and 251074
to J. Schleutker, grant 132473 to H. Nevanlinna, grant 250345
Finnish Center of Excellence in Cancer Genetics Research to L.
Aaltonen; and the Helsinki University Hospital Research Funds to
H. Nevanlinna.

Received August 29, 2012; revised November 14, 2012; accepted Decem-

Published OnlineFirst January 4, 2013; DOI: 10.1158/1055-9965.EPI-12-1000-T
HOXB13 G84E Mutation in Finland: Population-Based Analysis of Prostate, Breast, and Colorectal Cancer Risk

Virpi H. Laitinen, Tiina Wahlfors, Leena Saaristo, et al.

Cancer Epidemiol Biomarkers Prev  Published OnlineFirst January 4, 2013.

Updated version  Access the most recent version of this article at:
doi:10.1158/1055-9965.EPI-12-1000-T

Supplementary Material  Access the most recent supplemental material at:
http://cebp.aacrjournals.org/content/suppl/2013/01/04/1055-9965.EPI-12-1000-T.DC1

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.