The G84E mutation of HOXB13 is associated with increased risk for prostate cancer: results from the REDUCE trial

Zhuo Chen1*, Celia Greenwood2,3, William B.Isaacs4, William D.Foulkes5, Jielin Sun6,7, Sigun L.Zheng1,6, Lynn D.Con dreay7 and Jianfeng Xu1,6,8,9,*

1Center for Cancer Genomics, Wake Forest University School of Medicine, Winston-Salem, NC 27157, USA; 2Lady Davis Institute for Medical Research, Jewish General Hospital, Montreal, Quebec H3T 1E2, Canada; 3Departments of Oncology and Epidemiology, Biostatistics, and Occupational Health, McGill University, Montreal, Quebec H3A 1A2, Canada; 4Department of Urology, John Hopkins Medical Institutions, Baltimore, MD 21205, USA; 5Program in Cancer Genetics, McGill University, Montreal, Quebec H2W 1S6, Canada; 6Department of Genetics and Personalized Medicine Research, Wake Forest University School of Medicine, Winston-Salem, NC 27157, USA; 7Genetics, GlaxoSmithKline, Research Triangle Park, NC 27709, USA; 8Department of Urology and 9Department of Cancer Biology, Wake Forest University School of Medicine, Winston-Salem, NC 27157, USA

*To whom correspondence should be addressed. Center for Cancer Genomics, Wake Forest University School of Medicine, Winston-Salem, NC 27157, USA. Tel: +1 336 713 7500; Fax: +1 336 713 7566; Email: jxu@wakehealth.edu

A novel rare mutation, homeobox B13 (HOXB13) G84E, was reported to co-segregate with prostate B3 (PCa) in hereditary PCa families and associate with PCa risk in unrelated cases and controls. In this study, we aim to compare the G84E mutation frequency among subjects of different races/ethnicities from various geographic regions in the world and to assess its risk for developing PCa, in the Reduction by Dutasteride of Prostate Cancer Events (REDUCE) trial. All the 3508 subjects had initial negative prostate biopsy and were biopsied at Year 2 and 4 for detection of PCa. The G84E mutation was detected only in Caucasians, with the highest carrier frequency in Northern Europe (1.06%), followed by Western Europe (0.60%) and North America (0.31%). No mutation carrier was observed in Southern Europe, Eastern Europe, Latin America, Australia and South Africa. In Caucasians, the G84E mutation frequency was 0.99% and 0.24% in positive and negative biopsy subjects, respectively (P = 0.01). In positive biopsy subjects, the frequency was significantly higher in subjects with a positive family history than those without (4.31% versus 0.34%, P = 0.002). In the 4 year follow-up, the PCa detection rate was 53.8% among the 13 mutation carriers and 22.0% among 3186 non-carriers, relative risk = 2.45 (95% confidence interval: 1.48–4.07). All mutation carriers shared a common haplotype, suggesting a founder effect. In Finland, the G84E mutation was estimated to occur in the year 1792 (95% credible interval: 1735–1831). In conclusion, the G84E mutation of HOXB13, a relatively recent mutation that likely occurred in Northern Europe, significantly increases risk for PCa.

Introduction

Prostate cancer (PCa) is one of the most prevalent cancers worldwide. The incidence of PCa is higher in African Americans and in the European descent than in Hispanics and Asians (1,2). It is widely hypothesized that genetic factors contribute to racial/ethnic differences in incidence of PCa. However, specific genes underlying the variation remain largely unknown.

By targeted exome sequencing in a PCa linkage region at 17q21–q22 among 94 probandes of hereditary PCa families, Ewing et al. (3) recently identified a rare but recurrent non-synonymous mutation, G84E (rs138213197), in the transcription factor homebox B13 (HOXB13) gene. The G84E mutation is located in the conserved functional domains of HOXB13, which mediates the binding of HOX13 paralogs (including HOXB13) to the MEIS homeodomain family of HOX cofactor proteins. The amino acid substitution from glycine to glutamic acid is predicted to be deleterious to HOXB13 protein function. In Ewing’s study, the G84E mutation was found to co-segregate completely with PCa. In addition, they found that the mutation occurred in 1.4% of 5083 unrelated PCa cases and 0.1% of 1401 controls of Caucasian men. Carriers of the G84E mutation had 20.1 times the odds of PCa compared with the non-carriers (odds ratio [OR] = 20.1) (3). Subsequent studies by Breyer et al. (4) in 928 familial PCa probands and 930 controls of European descent and by Akbari et al. (5) in 1525 PCa sporadic cases and 1757 controls from Canada confirmed the association. However, the reported OR in the latter studies were 7.9 (95% confidence interval [CI]: 1.8–34.5) and 5.8 (95% CI: 1.3–26.5), respectively, which were considerably lower than the initial report (3).

To better understand the frequency of the G84E mutation in various races/ethnicities and geographic regions in the world and estimate its risk for PCa, we evaluated the mutation in the Reduction by Dutasteride of Prostate Cancer Events (REDUCE) trial, an international multicenter chemoprevention trial originally designed to evaluate the effect of dutasteride on PCa risk. The broad distribution of study subjects across the world and the prospective nature of the study design provide a unique opportunity to address these two questions.

Materials and methods

Study subjects

The study was performed in a total of 3508 subjects (3239 Caucasians, 45 African descendants, 199 Hispanics and 25 Asians from 31 countries) from the REDUCE trial who consented for genetic studies. The REDUCE trial is a 4 year, multicenter, randomized, double-blind, placebo-controlled, parallel group study. The design and rational of the REDUCE trial have been described in the previous reports (2,3). Among the 3508 subjects included in this study, 1794 subjects were randomized to the placebo arm, whereas 1714 subjects were in the dutasteride arm. The inclusion criteria for the REDUCE trial were as follows: (i) 50–75 years of age; (ii) prostate-specific antigen (PSA) level of 2.5–10 ng/ml for subjects <60 years old and 3–10 ng/ml for those ≥60 years; (iii) prostate volume ≤80 cm3; (iv) a 4 year, multicenter chemoprevention trial originally designed to evaluate the effect of dutasteride on PCa risk. The broad distribution of study subjects across the world and the prospective nature of the study design provide a unique opportunity to address these two questions.

Genetic markers and genotyping

We genotyped the HOXB13 G84E mutation and additional 14 variants within or flanking the HOXB13 gene region (chromosome 17, bp 46,179,399–46,827, 590, build 37). All variants were genotyped using the iPLEX MassARRAY system (Sequenom, San Diego, CA). Duplicates and negative (water) controls were included in each 96-well plates as quality control samples. All assays were performed in a blinded fashion. The overall genotyping call rate is 98.5%. Negative controls (water samples) did not give a genotype call. All of the variants were in Hardy–Weinberg Equilibrium in negative biopsy subjects (Supplementary Table 2, available at Carcinogenesis Online).

Abbreviations: HOXB13, homeobox B13; OR, odds ratio; PCa, prostate cancer; PSA, prostate-specific antigen; REDUCE, Reduction by Dutasteride of Prostate Cancer Events; RR, relative risk; SNP, single nucleotide polymorphism.

© The Author 2013. Published by Oxford University Press. All rights reserved. For Permissions, please email: journals.permissions@oup.com

Carcinogenesis vol.34 no.6 pp.1260–1264, 2013
doi:10.1093/carcin/bgt055
Advance Access publication February 7, 2013

1260
G84E frequencies were compared between positive prostate biopsy subjects and negative prostate biopsy subjects, using Fisher’s exact test implemented in SAS software (SAS Institute, Cary, NC). Due to the prospective nature of this study, relative risk (RR, rate ratio) was used to measure the magnitude of association between G84E mutation and PCa among Caucasians. In this study, RR is the ratio of the probability of PCa (positive prostate biopsy) occurring in the G84E mutation carriers versus the probability in the non-carriers. RR was estimated for all subjects and among the subsets of positive biopsy subjects versus all negative biopsy subjects. We also conducted the Breslow–Day test for homogeneity of stratified analysis according to family history, age at diagnosis and Gleason grade, respectively. RR estimation and the Breslow–Day test were conducted using SAS. Linear regression analysis was conducted using PLINK software (9) to assess the association between G84E mutation and baseline clinical variables, including total PSA, free PSA, prostate volume, testosterone, dihydrotestosterone and International Prostate Symptom Score, adjusting age as a covariate. All clinical variables were log transformed to approximate normal distributions. The interaction effect between G84E and treatment (dutasteride) was analyzed using logistic regression implemented in PLINK by adding an interaction term (9). All tests were two tailed.

To help make the results of this study more comparable with those of previous studies, we also estimated the OR. Using PLINK software to measure the association of each single nucleotide polymorphism (SNP) with PCa risk under an additive genetic model in all subjects (9). For OR estimation, subjects with positive prostate biopsy were treated as cases, whereas subjects with negative prostate biopsy were treated as controls.

**Phylogenetic tree generation**

The phylogenetic tree of haplotypes, which appeared over 10 times in all subjects, was constructed using MEGA version 5.05 (10). The neighbor-joining method (11), a simplified version of the minimum evolution method (12), was used to assess the relationships among major haplotypes.

**Mutation age estimation**

DMLE (Disease Mapping using Linkage disEquilibrium) software was used to estimate G84E mutation age in Finland, based on the observed linkage disequilibrium at multiple genetic markers (14). DMLE uses Markov Chain Monte Carlo methods to allow Bayesian estimation of the posterior probability density of the position of a disease mutation relative to a set of markers as well as the age of the mutation (14). The key parameters necessary for estimating the mutation age are the population growth rate and the proportion of disease chromosomes sampled. The population growth rate was calculated using the following formula: \( \rho_1 = \rho_0 \cdot e^\beta \) (15), where \( \rho_0 \) is the population of Finland in 1750 AD, \( \rho_1 \) is the Finnish population in 2003. In the REDUCE trial started, \( g \) is number of generations between these two population estimates and \( r \) is the growth rate per generation. According to Statistics Finland (http://www.stat.fi/meta/svt/index_en.html), the population was 421,500 in the year 1750 and 5,219,732 in 2003, respectively. Because the average age of primiparous women was 24.9 in 1975 in Finland (http://www.stat.fi/meta/svt/index_en.html) and that women gave birth at a relatively younger age in history than in recent years, we assumed 22 years per generation in the history in Finland. We assumed a constant growth rate over these 250 years as well. For the proportion of disease chromosomes sampled, we used the formula: \( \rho = N \rho' \), where \( N \) is the number of cases in the sample. \( N \) is the population of Finland in 2003. \( N \) is the lifetime risk of PCa and \( F \) is the estimated frequency of the G84E mutation. The distribution for G84E carriers is presented in Table I. Briefly, the G84E mutation of HOXB13 was observed in 13 of 3239 Caucasian men (0.4%) but was not detected in other races/ethnicities, including 45 African descendants, 199 Hispanics and 25 Asians. Among Caucasians, carrier frequency of the G84E mutation varied from different geographic regions; it was highest in Northern Europe (1.06%), followed by Western Europe (0.60%) and North America (0.31%). No mutation carrier was observed in Southern Europe (\( N = 419 \)), Latin America (\( N = 233 \)), Australia (\( N = 37 \)) and South Africa (\( N = 31 \)).

**Results**

### G84E carrier frequency distributions

The distribution for G84E carriers is presented in Table I. Briefly, the G84E mutation of HOXB13 was observed in 13 of 3239 Caucasian men (0.4%) but was not detected in other races/ethnicities, including 45 African descendants, 199 Hispanics and 25 Asians. Among Caucasians, carrier frequency of the G84E mutation varied from different geographic regions; it was highest in Northern Europe (1.06%), followed by Western Europe (0.60%) and North America (0.31%). No mutation carrier was observed in Southern Europe (\( N = 419 \)), Latin America (\( N = 233 \)), Australia (\( N = 37 \)) and South Africa (\( N = 31 \)).

### Effects of G84E on PCa

In the end of the 4 year follow-up of the REDUCE trial, we found that 7 out of 13 (53.8%) G84E mutation carriers developed PCa, whereas the number in non-carriers was 701 of 3186 (22.0%). The G84E mutation carriers were at 2.45-fold increased risk of developing PCa, compared with non-carriers (RR = 2.45, 95% CI: 1.48–4.07) (Table II). The G84E carrier frequency was 0.99% (7/708) in positive biopsy subjects and 0.24% (6/2491) in negative biopsy subjects (Table II).

In addition, we compared carrier frequency of G84E mutation in the positive biopsy subjects stratified by family history, age at diagnosis and Gleason grade (Table II). Among 708 positive biopsy subjects, we found 5 G84E mutation carriers in 116 subjects with a family history of PCa and 2 G84E mutation carriers in 592 subjects without a family history, which was significantly different (4.31% versus 0.34%, \( P = 0.002 \). With regard to age at diagnosis, we did not detect significant difference of G84E carrier frequency between positive biopsy subjects diagnosed before or after 65 years old (1.05% [3/285] versus 0.95% [4/423], \( P = 0.99 \)). However, we noticed that all of the three positive biopsy subjects carrying G84E mutation diagnosed before 65 years old had a family history of PCa. Although carrier frequency of the G84E mutation was not significantly different in positive biopsy subjects with high (\( >7 \)) or low (\( <7 \)) Gleason grade (1.71% [4/234] versus 0.63% [3/471], \( P = 0.23 \)), G84E mutation carriers were at 4.73-fold increased risk of developing high Gleason grade PCa (RR = 4.72, 95% CI: 2.19–10.19) compared with all negative biopsy subjects.

No significant association was detected between G84E mutation and PCAn-related baseline clinical variables, including total PSA, free PSA, prostate volume, testosterone, dihydrotestosterone and International Prostate Symptom Score (Table III). In addition, no significant interaction between G84E mutation and treatment (dutasteride) was found (\( P = 0.57 \)).

To make our results more comparable with previous studies in which OR was used as a measurement for association, we also estimated OR to measure the association of each SNP with PCa risk under an additive genetic model among all subjects. Carriers of the G84E mutation had 4.14 times the odds of PCa compared with the non-carriers (OR = 4.14, 95% CI: 1.38–12.28, \( P = 0.006 \)). None of the other SNPs were significantly associated with PCa risk (Supplementary Table 2, available at Carcinogenesis Online).

### Haplotype inference

A total of 91 haplotypes were inferred. All 13 G84E mutation carriers shared a unique haplotype G-T-A-G-A-C-C-A-T with the G84E mutation had 4.14 times the odds of PCa compared with the non-carriers (OR = 4.14, 95% CI: 1.38–12.28, \( P = 0.006 \)). None of the other SNPs were significantly associated with PCa risk (Supplementary Table 2, available at Carcinogenesis Online).
Mutation age estimation

We estimated age of mutation for G84E in subjects from Finland, where the highest G84E carrier frequency (1.96%) was observed, except in Belgium where only 29 subjects were included and may not be representative. Assuming 22 years per generation (population growth rate 1.24 per generation), the G84E mutation was estimated to originate approximately 9.6 (95% credible interval: 7.8–12.2) generations ago in Finland. In term of year, the G84E mutation was estimated to originate around the year 1792 (95% credible interval: 1735–1831). When assuming 20 or 25 years per generation in Finland, the G84E mutation may originate in the year 1807 (95% credible interval: 1753–1843) or 1863 (95% credible interval: 1803–1890), respectively. Overall, the HOXB13 G84E is most likely to originate around the turn of the 19th century.

Discussion

In this study, we confirmed the rare mutation G84E in HOXB13 gene as susceptibility loci for PCa in Caucasians from the REDUCE trial in a prospective manner. In addition, in positive biopsy subjects, the G84E frequency was significantly higher in subjects with a positive family history than those without. Moreover, we found the G84E mutation was likely to be a founder mutation, suggesting this mutation affects risk of hereditary PCa. Given the G84E frequency in Northern Europe (1.06%), Western Europe (0.60%), and North America (0.31%), we speculated that G84E might spread from Finland to other geographic regions during historic population migration, as nearly 0.7 million Finnish emigrated to Sweden, the USA and Canada during 1866–1970 (16).

We compared the effect of the G84E mutation on PCa risk in our study with previous reports (3–5). Only OR estimates were available from the previous reports due to the case-control study design adopted. Therefore, we also estimated OR in our study. The OR estimate in our study (4.14) was smaller than that in the previous reports by Ewing et al. (3) (OR = 20.1), Breyer et al. (4) (OR = 7.9) and Akbari et al. (5) (OR = 5.8). Two possible reasons may account for the variations of the point estimates. The first possible reason may be due to the different sampling strategy used in the studies. For case selection, PCa patients in the original study by Ewing et al. (3) were either hereditary PCa patients or diagnosed at a young age (<55 years). Similarly, study by Breyer et al. (4) was also conducted in familial cases, whereas only 15.8% patients in our study and 17.7% patients in Akbari’s study had a family history of PCa (5). In addition to case selection, sampling strategy of control may also influence the point estimation of OR. For example, controls in Ewing’s study were selected based on low PSA levels and digital rectal examination, whereas controls in our study had a PSA ranging from 2.5–10ng/ml with negative prostate biopsy. Therefore, the variations in OR estimates may due to different sampling strategies used among studies. Second, potential population stratification may also influence OR estimation. Although

Discussion

In this study, we confirmed the rare mutation G84E in HOXB13 gene as susceptibility loci for PCa in Caucasians from the REDUCE trial in a prospective manner. In addition, in positive biopsy subjects, the G84E frequency was significantly higher in subjects with a positive family history than those without. Moreover, we found the G84E mutation was likely to be a founder mutation, suggesting this mutation affects risk of hereditary PCa. Given the G84E frequency in Northern Europe (1.06%), Western Europe (0.60%), and North America (0.31%), we speculated that G84E might spread from Finland to other geographic regions during historic population migration, as nearly 0.7 million Finnish emigrated to Sweden, the USA and Canada during 1866–1970 (16).

We compared the effect of the G84E mutation on PCa risk in our study with previous reports (3–5). Only OR estimates were available from the previous reports due to the case-control study design adopted. Therefore, we also estimated OR in our study. The OR estimate in our study (4.14) was smaller than that in the previous reports by Ewing et al. (3) (OR = 20.1), Breyer et al. (4) (OR = 7.9) and Akbari et al. (5) (OR = 5.8). Two possible reasons may account for the variations of the point estimates. The first possible reason may be due to the different sampling strategy used in the studies. For case selection, PCa patients in the original study by Ewing et al. (3) were either hereditary PCa patients or diagnosed at a young age (<55 years). Similarly, study by Breyer et al. (4) was also conducted in familial cases, whereas only 15.8% patients in our study and 17.7% patients in Akbari’s study had a family history of PCa (5). In addition to case selection, sampling strategy of control may also influence the point estimation of OR. For example, controls in Ewing’s study were selected based on low PSA levels and digital rectal examination, whereas controls in our study had a PSA ranging from 2.5–10ng/ml with negative prostate biopsy. Therefore, the variations in OR estimates may due to different sampling strategies used among studies. Second, potential population stratification may also influence OR estimation. Although
Table II. *HOXB13* G84E carrier frequency in positive and negative biopsy subjects

<table>
<thead>
<tr>
<th></th>
<th>G84E carriers</th>
<th>G84E non-carriers</th>
<th>G84E carrier frequency</th>
<th>RR (95% CI)*</th>
<th>P</th>
<th>Homogeneity P*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Negative biopsy subjects</td>
<td>6</td>
<td>2485</td>
<td>0.24%</td>
<td>1.00</td>
<td>NA</td>
<td></td>
</tr>
<tr>
<td>Positive biopsy subjects</td>
<td>7</td>
<td>701</td>
<td>0.99%</td>
<td>2.45 (1.48–4.07)</td>
<td>0.01</td>
<td></td>
</tr>
<tr>
<td>Family history</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Positive family history</td>
<td>5</td>
<td>111</td>
<td>4.31%</td>
<td>10.63 (5.43–20.83)</td>
<td>0.04</td>
<td>0.004</td>
</tr>
<tr>
<td>Negative family history</td>
<td>2</td>
<td>590</td>
<td>0.34%</td>
<td>1.31 (0.39–4.34)</td>
<td>0.65</td>
<td></td>
</tr>
<tr>
<td>Age at diagnosis</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;65 y</td>
<td>3</td>
<td>282</td>
<td>1.05%</td>
<td>3.28 (1.29–8.31)</td>
<td>0.06</td>
<td>0.91</td>
</tr>
<tr>
<td>≥65 y</td>
<td>4</td>
<td>419</td>
<td>0.95%</td>
<td>2.78 (1.29–5.97)</td>
<td>0.04</td>
<td></td>
</tr>
<tr>
<td>Gleason grade</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>High grade (≥7)</td>
<td>4</td>
<td>230</td>
<td>1.71%</td>
<td>4.73 (2.19–25.71)</td>
<td>0.01</td>
<td>0.29</td>
</tr>
<tr>
<td>Low grade (&lt;7)</td>
<td>3</td>
<td>471</td>
<td>0.63%</td>
<td>2.10 (0.83–5.30)</td>
<td>0.16</td>
<td></td>
</tr>
</tbody>
</table>

* RR was estimated for overall or subsets of positive biopsy subjects versus all negative biopsy subjects.
* Homogeneity P of Breslow–Day test was estimated for homogeneity of stratified analysis according to family history, age at diagnosis and Gleason grade, respectively.

Table III. Associations between *HOXB13* G84E mutation and baseline clinical variables

<table>
<thead>
<tr>
<th>Clinical variables</th>
<th>Mean*</th>
<th>P*</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>G84E carriers</td>
<td>G84E non-carriers</td>
</tr>
<tr>
<td>PSA (ng/ml)</td>
<td>5.43</td>
<td>5.58</td>
</tr>
<tr>
<td>Free PSA (ng/ml)</td>
<td>0.70</td>
<td>0.87</td>
</tr>
<tr>
<td>Prostate volume (cm³)</td>
<td>37.26</td>
<td>43.38</td>
</tr>
<tr>
<td>Testosterone (nmol/l)</td>
<td>16.98</td>
<td>14.38</td>
</tr>
<tr>
<td>Dihydrotestosterone (nmol/l)</td>
<td>1.35</td>
<td>1.20</td>
</tr>
<tr>
<td>International Prostate Symptom Score</td>
<td>6.14</td>
<td>7.14</td>
</tr>
</tbody>
</table>

*Tests were based on variables by linear regression analysis, adjusted for age; mean were back-transformed to original scale.

Funding

National Cancer Institute RC2 grant (CA148463 to J.X.); and a research contract by GlaxoSmithKline to J.X.

Acknowledgements

We thank the patients enrolled in the REDUCE trial (sponsored by GlaxoSmithKline) who provided consent and genetic samples, which enabled this study, and the clinicians who contributed their expertise in recruiting study patients for the REDUCE clinical study. Dave Pulford, Jennifer Aponte, Jon Charnecki and Mary Eilyn Volk participated in consent reconciliation and sample management to enable genetic sample selection for inclusion and genotype determination. Karen King provided data management support for this project. We appreciate the assistance of Lauren Marmor in coordinating the project. We thank the patients enrolled in the REDUCE trial and hold stock in GlaxoSmithKline. J.X. certifies that all conflicts of interest, including specific financial interests and relationships and affiliations relevant to the subject matter or materials discussed in the manuscript (e.g. employment/affiliation, grants or funding, consultations, honoraria, stock ownership or options, expert testimony, royalties, or patents filed, received or pending) have been disclosed.

References


Received September 13, 2012; revised January 30, 2013; accepted February 3, 2013