

## **A common nonsense mutation in *EphB2* is associated with prostate cancer risk in African American men with a positive family history**

Rick A. Kittles<sup>1,4</sup>, Agnes B. Baffoe-Bonnie<sup>2</sup>, Tracy Y. Moses<sup>3</sup>, Christiane M. Robbins<sup>3</sup>, Chiledum Ahaghotu<sup>4</sup>, Pia Huusko<sup>3</sup>, Curtis Pettaway<sup>5</sup>, Srinivasan Vijayakumar<sup>6</sup>, James Bennett<sup>7</sup>, Gerald Hoke<sup>8</sup>, Terry Mason<sup>9</sup>, Sally Weinrich<sup>10</sup>, Jeffrey M. Trent<sup>3</sup>, Francis S. Collins<sup>11</sup>, Spyro Mousses<sup>3</sup>, Joan Bailey-Wilson<sup>12</sup>, Paulette Furbert-Harris<sup>4</sup>, Georgia Dunston<sup>4</sup>, Isaac J. Powell<sup>13</sup>, John D. Carpten<sup>3</sup>

1 Department of Molecular Virology, Immunology & Medical Genetics, Arthur G. James Cancer Hospital and Richard J. Solove Research Institute, The Ohio State University, 494 Medical Research Facility 420 West 12<sup>th</sup> Avenue, Columbus, OH 43210

2 Population Science Division, Human Genetics Program, Fox Chase Cancer Center, 333 Cottman Avenue, Philadelphia, PA 19111-2497

3 Genetic Basis of Human Disease Research Division, Translational Genomics Research Institute, 445 N. Fifth Street, Phoenix, AZ 85004

4 National Genome Center at Howard University, 520 W. Street, NW, Washington, DC 20059

5 Department of Urology, The University of Texas MD Anderson Cancer Center, 1515 Holcombe Boulevard Houston, TX 77030

6 Department of Radiation Oncology, UC Davis Cancer Center, University of California, Davis, 4501 X Street, Sacramento, CA 95817

7 Midtown Urology Surgical Center, 128 North Avenue, NE Suite 100, Atlanta, GA 30308

8 Department of Urology, Columbia University Medical Center, Columbia University, 180 Fort Washington Avenue, New York, NY 10032

9 Department of Urology, University of Illinois at Chicago, Chicago, IL 60612

10 Medical College of Georgia, School of Nursing, EB-210, Augusta, GA 30912

11 National Human Genome Research Institute, 31 Center Drive, Bethesda, MD 20892-2152

12 Center for Inherited Disease Research, National Institutes of Health, 333 Cassell Drive, Suite 2000 Baltimore, MD 21224

13 Karmanos Cancer Institute, Wayne State University, 4100 John R, Detroit, MI 48201

Correspondence should be addressed to:

John D. Carpten, Ph.D.

Genetic Basis of Human Disease Research Division

Translational Genomics Research Institute

445 N. Fifth Street

Phoenix, AZ 85004

602-343-8819

jcarpten@tgen.org

**ABSTRACT**

**Background:** The *EphB2* gene was recently implicated as a prostate cancer (PC) tumor suppressor gene, with somatic inactivating mutations occurring in ~10% of sporadic tumors<sup>1</sup>. We evaluated the contribution of *EphB2* to inherited PC susceptibility in African Americans (AA) by screening the *EphB2* gene for germ-line polymorphisms.

**Methods:** Direct sequencing of the coding region of *EphB2* was performed in 72 probands from the African American Hereditary Prostate Cancer Study (AAHPC). A case-control association analysis was then carried out using the AAHPC probands and an additional 183 sporadic PC patients compared to 329 healthy AA male controls. Also, we performed an ancestry-adjusted association study where we adjusted for individual ancestry among all subjects, in order to rule out a spurious association due to population stratification.

**Results:** Ten coding sequence variants were identified, including the K1019X (3055A>T) nonsense mutation which was present in 15.3% of the AAHPC probands but only 1.7% of 231 European American control samples. We observed that the 3055A>T mutation significantly increased risk for prostate cancer over two-fold (Fisher's 2-sided P=0.003). The T allele was significantly more common among AAHPC probands (15.3%) than among healthy AA male controls (5.2%) (OR=3.31; 95%CI=1.5-7.4; P=0.008). The ancestry-adjusted analyses confirmed the association.

**Conclusions:** Our data show that the K1019X mutation in the *EphB2* gene differs in frequency between African Americans and European Americans, is associated with increased risk for PC in AA men with a positive family history, and, may be an important genetic risk factor for prostate cancer in African Americans.

**Keywords:** EphB2, tumor suppressor, African Americans, hereditary prostate cancer, single nucleotide polymorphism

## INTRODUCTION

Prostate cancer is the most common male specific malignancy in the US and disproportionately affects African American (AA) men, who have higher incidence (~40% of all cancer cases) and mortality rates as compared to other ethnic groups<sup>2,3</sup>. The underlying reasons for these disparities are not well understood, although existing evidence implicates an important genetic component<sup>4-9</sup>. Many studies of hereditary prostate cancer (HPC) have been reported; however, few, if any, genes have been identified which are reproducibly associated with increased risk for prostate cancer across different study populations, emphasizing the heterogeneous nature of this disease<sup>10</sup>. Despite AA men having the highest incidence and mortality rates of prostate cancer in the US, very little data are available on the genetics of familial prostate cancer in this ethnic group. Studying the genetic contributions for prostate cancer in this high-risk population will have important implications for addressing the disparity of prostate cancer in African Americans.

Germ-line mutations have been found within three candidate genes for hereditary prostate cancer. These genes include ELAC2 at 17p11<sup>11</sup>, RNASEL at 1q25<sup>12,13</sup>, and MSR1 at 8p22<sup>14</sup>. The frequency spectrum of rare nonsense and missense mutations within these candidate genes vary significantly across ethnicities. These mutations in addition to common polymorphisms within the three candidate genes may also contribute to sporadic disease in European descent populations<sup>7,13,15,16</sup>. To date there has not been a clear candidate identified as contributing to hereditary prostate cancer in African Americans.

Recently, the gene *EphB2* gene, which encodes the EPHB2 receptor tyrosine kinase was discovered as being completely inactivated in the DU145 prostate cancer cell line using nonsense mediated RNA decay (NMD) inhibition in combination with array CGH, to enrich for genes likely to harbor mutations<sup>1</sup>. The exact function of the gene is unknown however evidence suggest *EphB2* may be a tumor suppressor gene since wild type *EphB2* significantly reduced clonogenic growth of DU145 prostate cancer cells (which have biallelic inactivation of *EphB2*)<sup>1</sup>. *EphB2* maps to 1p36, a region previously shown to be linked with hereditary prostate cancer among ethnically diverse sets of families<sup>17,18</sup>, including African Americans<sup>19</sup>. The strong genomic and functional characteristics of *EphB2* along with its map position near a putative HPC locus make it a strong candidate prostate cancer susceptibility gene. Therefore we set out to screen the *EphB2* gene by direct sequencing for the presence of germline mutations in 72 unrelated African American hereditary prostate cancer cases in order to determine if this gene is associated with prostate cancer predisposition in this high-risk population. A common nonsense mutation, K1019X, was genotyped in an additional 183 African American sporadic prostate cancer cases and 329 healthy age-matched controls. We provide evidence of an association of the *EphB2* nonsense mutation with the risk of prostate cancer in African Americans.

## METHODS

### *African American hereditary prostate cancer cases*

Hereditary prostate cancer cases consisted of 72 African Americans from unrelated HPC multiplex families recruited as part of the African American Hereditary Prostate Cancer (AAHPC) Study Network<sup>20-22</sup>. Ascertainment of multiplex prostate cancer families and the clinical description of the AAHPC cases have been previously described<sup>20,21</sup>. Due to

barriers in recruiting African American men into hereditary prostate cancer studies, the AAHPC Study Network developed a nation-wide effort to establish Collaborative Recruitment Centers (CRCs) in regions of the US with large numbers of African Americans including Atlanta, GA; Chicago, IL; Detroit, MI; Harlem, NY; Houston, TX; rural South Carolina, and Washington, DC<sup>20</sup>. Inclusion criteria were, 1) four or more prostate cancer cases, preferably first degree relatives, 2) at least three cases available for sampling, and 3) an average age at diagnosis of <65 years of age for the family. These families all self identified as African American and were verified by the recruitment staff. We performed mutational analysis for the *EphB2* gene using DNA samples from 72 probands from unrelated AAHPC families. The average age at diagnosis for these probands was 64.9 years of age. Clinical characteristics of the AAHPC dataset have been previously described<sup>22</sup>. All participants gave informed consent and recruitment was approved by the appropriate Institutional Review Boards (IRB).

#### *African American sporadic prostate cancer cases and controls*

Unrelated men self-described as African American were enrolled for case-control studies of risk factors for prostate cancer. The subjects consisted of 512 African Americans (183 prostate cancer patients and 329 healthy male controls) recruited from the Howard University Hospital (HUH) in Washington, DC<sup>5, 23</sup>. Incident cases were identified through the Division of Urology at HUH and confirmed by review of medical records. Healthy control subjects unrelated to the cases and matched for age ( $\pm 5$  years) were ascertained from the Division of Urology at HUH and also from men participating in screening programs for prostate cancer at the Howard University Hospital. Control subjects were excluded from the study if they were ever diagnosed with benign prostatic hyperplasia (BPH), have ever had an elevated prostate specific antigen test ( $>4.0$ ng/ml), or have had an abnormal digital rectal examination (DRE). The demographic characteristics of participants in the screening program were similar to the patient population seen in the Division of Urology clinics. Recruitment of sporadic prostate cancer cases and healthy controls occurred concurrently, and they each donated a blood sample for research purposes. The participation response rates for cases and controls were 92% and 90%, respectively. All prostate cancer cases were between 40 to 85 years of age and were diagnosed with the disease within the last 4 years. Clinical characteristics including Gleason grade, prostatic specific antigen (PSA), tumor-node-metastasis (TNM) stage, age at diagnosis and family history were obtained for all cases from medical records. Disease aggressiveness was defined as “Low” (T category  $<T2c$  and/or Gleason grade  $<7$ ) or “High” (T category  $>T2c$  and/or Gleason grade  $>7$ ). The Howard University IRB approved the study and written consent was obtained from all subjects. In addition, we used previously published data from 231 European American (EA) population control genomic DNA samples commercially available from the Coriell Institute for Medical Research in order to compare the *EphB2* K1019X allele and genotype frequencies<sup>1</sup>.

#### *Mutation detection and genotyping*

DNA specimens were amplified using standard PCR protocol and intronic primer pairs with M13 tails (sequences available on request). The PCR products were purified using the QiaQuick PCR purification kit on the BioRobot 8000 Automated Nucleic Acid Purification and Liquid Handling system (Qiagen). Quarter or eighth volume cycle sequencing

reactions were prepared in 96 well format using standard M13 forward or reverse primers with the Big Dye Terminator Chemistry (PE/Applied Biosystems, Piscataway, NJ). Following Sephadex purification, sequence products were separated on an ABI 3700 or ABI 3730 Capillary DNA Analyzer (PE/Applied Biosystems, Piscataway, NJ) using manufacturer's protocols. Sequence chromatograms were aligned and analyzed using Sequencher version 4.1 (Gene Codes). Due to the complexity of the mutations (lying within a A(9) mononucleotide stretch) we used DNA sequencing for all samples. Of the 72 samples, we achieved a 100% success rate. For all 584 samples analyzed we achieved an overall success rate of >95%. Reproducibility was >99% based upon comparison of data from duplicate experiments for a subset of samples.

#### *Controlling for population stratification*

To control for possible confounding by population stratification in this study, a panel of 34 ancestry informative markers (AIMs) was also genotyped in the African American samples. These markers show large differences in frequency between the parental populations (West Africans and Europeans), and were used to control for the presence of population stratification (PS) due to admixture<sup>24, 25</sup>. Information regarding primer sequences, polymorphic sites and other relevant information on the AIMs can be found at the dbSNP NCBI database site, under the submitter handle PSU-ANTH.

#### *Statistical analysis*

Odds ratios and P values were determined by logistic regression analyses from comparison of genotypes between subjects with prostate cancer and healthy controls using SAS version 6.91 (SAS Institute, Inc, Cary, NC.). Further analyses were performed on the combined dataset consisting of all PC subjects, and for the hereditary and sporadic cases separately. For all analyses, genetic effects were adjusted for age (at time of diagnosis for case subjects and at time of ascertainment for controls). Statistical control of PS was achieved by introducing individual ancestry (IA) as a covariate in the analyses. Individual ancestry was estimated by a Bayesian method implemented in the STRUCTURE 2.0 program<sup>26</sup> and the estimate was then used as a covariate in the regression analyses. Given the large number of potential false-positive marker associations (Type II error) throughout the genome it would be appropriate to adjust the traditional thresholds of significance for P-values based on some correction for multiple testing, however given our modest sample size and since we explored the effect of only one SNP on prostate cancer risk, the only corrections on P-values we made were due to ancestry.

## **RESULTS**

The clinical characteristics of the 72 AAHPC probands, the 183 sporadic PC cases and the 329 healthy AA male controls are presented in Table 1. The mean age of 69 years for the sporadic PC cases was higher than that for the controls (66.1 years) and the HPC probands (64.9 years). The Wilcoxon Sign Rank test showed that the mean PSA for both the HPC probands and the sporadic PC cases compared with AA male controls was significantly different (P value <0.01). For 52 HPC probands on whom disease aggressiveness categorization was available, only 17% had a high index compared with 47% of those with sporadic disease.

Mutational analysis in 72 AAHPC probands resulted in the discovery of ten unique coding sequence variants within the *EphB2* gene. All sequence variants are summarized in Table 2. Six silent exonic variants were observed with frequencies ranging from 1.4% to 30% among the probands. Only four of the ten variants actually resulted in amino acid changes and are considered mutations (Table 2).

**Table 1. Clinical Characteristics of African American PC patients and population-based control subjects.**

Characteristic	Hereditary PC (N = 72)	Sporadic PC (N = 183)	Controls (N = 329)
Mean age in years (SD)	64.9 (20.8)	69.0 (8.9)	66.1 (12.6)
Serum PSA in ng/ml <sup>a</sup>			
Mean (SD)	52.3 (90.1) <sup>b</sup>	71.3 (195.1) <sup>b</sup>	2.8 (1.1)
Median (range)	7.6 (0.1-230)	8.6 (0.1-150)	1.0 (0.1-3.6)
Disease Aggressiveness: <sup>c</sup>			
Low (%)	43 (83)	54 (53)	-
High (%)	9 (17)	48 (47)	-
Unknown	20	81	-

<sup>a</sup> Serum PSA measured at time of diagnosis for cases and at most recent clinical visit for controls.

<sup>b</sup> P-value < 0.01 from Wilcoxon Sign Rank test comparison with control population.

<sup>c</sup> Low aggressiveness: Gleason <7 and T category <T2c; High aggressiveness: Gleason ≥7 or T category ≥ T2c.

**Table 2. *EphB2* sequence variants discovered in 72 AAHPC probands**

Nucleotide position	Amino Acid change	dbSNP	Frequency in probands
510 C/T	none	-	5.6%
624 G/A	none	-	1.4%
657 G/A	none	rs1371869	1.4%
<b>835 G/T</b>	<b>A279S</b>	-	<b>2.8%</b>
930 C/T	none	-	5.6%
1377 G/A	none	rs2229872	30%
<b>1949 T/C</b>	<b>V650A</b>	-	<b>2.8%</b>
2640 G/A	none	-	1.4%
<b>2647 A/G</b>	<b>M883V</b>	-	<b>2.8%</b>
<b>3055 A/T</b>	<b>K1019X</b>	-	<b>15.3%</b>

<sup>a</sup> Nucleotide and amino acid positions are based upon coding sequence for Genbank accession file NM\_017449.

Two of the amino acid substitutions, V650A and M883V were novel in contrast to the previously observed A279S variant <sup>1</sup> but all three were observed at a frequency of 2.8%. Both V650A and M883V appear to be located in the mutation rich kinase domain of the EphB2 protein. The kinase domain is essential for receptor signaling <sup>1</sup>. Also observed

among the coding mutations was the previously reported K1019X nonsense mutation (3055 A>T) in exon 15 of the *EphB2* gene. The K1019X mutation was present in 15.3% (11 of 72) of the AAHPC probands, though it was previously shown to be present in 1.7% (4 of 231) of control DNA samples from the Coriell Institute for Medical Research (two-sided Fisher's exact test P-value of 0.000043). We observed that this nonsense mutation was three times more common in the AA controls (5.17%, 17 of 329) than in the EA controls.

An association analysis of the K1019X variant (3055 A>T) was performed combining all AAHPC (N=72) and sporadic cases (N=183) and compared them to African American male controls (N=329) controlling for age at diagnosis. We did not examine the other coding variants since they were less frequent (< 3.0%) and thus would require an extremely large number of cases and controls in order to perform a reliable association study. Table 3 reveals that the presence of the (T) allele significantly increased risk for prostate cancer (OR = 2.44; 95% CI = 1.4-4.3; Fisher's 2 sided P=0.003) in the combined dataset. Subset analyses revealed that the frequency of K1019X was significantly higher for the AAHPC probands (15.3%) as compared to AA healthy male controls (5.2%) (OR = 3.31; CI 1.48-7.41; Fisher's 2-sided P=0.008). We compared the 6.6% (12 of 183) frequency of the mutation among the 183 sporadic prostate cancer cases with the AA healthy male controls, and found no significant difference between the two groups (OR, 1.3; 95% CI, 0.60-2.76; Fisher's 2-sided P=0.55) (Table 3). Stratifying the analysis by prostate cancer disease aggressiveness revealed no significant associations (data not shown).

**Table 3. Association between Prostate Cancer and *Ephb2* K1019X Polymorphism.**

Population	N	No. (%) of subjects with			OR (95% CI) <sup>a</sup>	P-value
		K/ K	K/ X	X/ X		
All cases	255	234 (91.7)	19 (7.5)	2 (0.7)	2.44 (1.4-4.3)	0.003
AAHPC	72	61 (84.7)	11 (15.3)	-	3.31 (1.5-7.4)	0.008
AA sporadic PC	183	173 (94.5)	8 (4.4)	2 (1.1)	1.28 (0.6-2.8)	0.55
AA controls	329	312 (94.8)	9 (3.2)	8 (1.9)	1.00 (reference)	-
Coriell controls	231	270 (98.5)	4 (1.5)	-		

<sup>a</sup>OR by logistic regression for K/X and X/X genotype comparisons with AA controls after adjusting for age at diagnosis.

We note that the K1019X variant was not in Hardy Weinberg Equilibrium (HWE) within our AA sporadic case and control samples ( $p < 0.05$ ). The observed departure from HWE was not unexpected given that the African American population is the product of recent admixture<sup>27</sup> and there were significant differences in K1019X allele frequency between African Americans and European Americans. Thus, in order to rule out a spurious

association of K1019X with prostate cancer in African Americans due to admixture stratification we controlled for ancestral differences between the prostate cancer cases and controls by estimating individual ancestry (IA) for each subject using 34 admixture informative markers (AIMs). We used the individual ancestry estimate for each subject as a covariate in the association analysis in order to take into account differences in ancestral proportions between cases and controls<sup>24, 25</sup>. Individual ancestry (West African) ranged from 10% to 93.5% in the cases with an average IA estimate of  $71.3 \pm 1.9$ . The estimates for West African ancestry for the controls ranged from 6.5% to 95.3% (average was  $69.0 \pm 0.8$ ). After adjusting for individual ancestry, the association of the *EphB2* K1019X mutation with prostate cancer in the AAHPC probands as compared to the AA healthy male controls was still significant ( $P=0.01$ ).

## DISCUSSION

The potential genetic basis for the high incidence and mortality rates of prostate cancer among African Americans is still under investigation. In this study, we identified ten sequence variants in the *EphB2* gene, including a common nonsense mutation K1019X among 72 African American hereditary prostate cancer patients. Further, we revealed an association of *EphB2* K1019X and prostate cancer risk in African Americans. The K1019X variant was observed in much higher frequency among AA prostate cancer patients than among healthy AA male controls ( $P=0.003$ ). The association was mainly due to men with hereditary prostate cancer ( $P=0.008$ ). In fact the risk for prostate cancer was increased 3-fold among African American men who carried at least one copy of the K1019X allele and had a family history of prostate cancer. Given its high frequency in hereditary cases, K1019X likely is associated with familial prostate cancer in African American men. Of course, further replication is necessary before these results can be accepted as a true positive. The most reliable replication study should use prostate cancer patients with strong family history and not sporadic cases. A larger case-control study could also evaluate the rare missense variants such as A279S, V650A, and M883V which may also contribute to prostate cancer risk.

The results of this study, and the recent identification of biallelic inactivation of the *EphB2* gene using the nonsense-mediated decay (NMD) microarray technique<sup>1</sup>, further implicates *EphB2* as a prostate cancer tumor suppressor gene. K1019X is an A to T transversion within a poly-A tract of the last exon of the *EphB2* gene. This polymorphism may exhibit an affect similar to the APC I1307K polymorphism, carried by 6.1% of Ashkenazi Jews, that is associated with colorectal cancer<sup>28, 29</sup>. Mononucleotide repeat sequences such as where K1019X occurs have been shown to be genetically unstable and prone to somatic mutation<sup>28, 30</sup>. Performing somatic mutational analyses surrounding the poly-A tract region is strongly warranted in order to confirm the role of *EphB2* in prostate tumorigenesis.

We noted previously that the prevalence of K1019X was significantly higher among African American controls than among European American controls ( $P<0.001$ ), suggesting that it could be in admixture disequilibrium in the AA population. These findings were extremely interesting and inspired us to further investigate the role of this mutation as a prostate cancer genetic risk factor; however, several questions remained. Ethnic differences in allele frequency and disease risk can create false-positive results in case-control studies, especially when using recently admixed populations such as African



Americans. Thus, in order to control for possible confounding we introduced individual ancestry as a covariate in our analyses. This approach has been used to limit spurious associations that are the result of differences in ancestral proportions (admixture). Our ancestry-adjusted analyses provided additional support for a strong association of the K1019X and prostate cancer in African Americans.

The frequencies of sequence variants in a number of candidate genes for prostate cancer differ significantly between African Americans and European Americans. Among the examples are the CAG repeat tract within the androgen receptor gene<sup>31, 32</sup>, a TA-repeat tract within the SRD5AR gene<sup>33</sup>, the CYP3A4 promoter variant<sup>6, 25</sup> and frequent variants within MSR1<sup>7</sup>. The *EphB2* K1019X mutation represents a novel addition to this group of allelic variants. Population differences in allele frequencies for many of the candidate genes are real but whether the differences contribute to differences in disease susceptibility is difficult to assess with the current available data and traditional approaches. One potential confounding issue is population stratification (PS) created by admixture<sup>34, 35</sup>. Here we typed unlinked genetic markers informative for ancestry in order to detect, quantify and correct for PS<sup>26, 36-39</sup>. Even though studies have consistently shown evidence of PS in African Americans<sup>25, 27, 40, 41</sup> there continues to be debate on whether PS exists<sup>42, 43</sup> or whether its impact is significant<sup>44</sup>. A recent study has shown that even in highly inbred populations, significant PS exists, lending support that this phenomenon must exist in highly outbred populations such as African Americans<sup>45</sup>.

Our examination of sequence variants in the *EphB2* gene and subsequent case-control study among African American men suggest that *EphB2* may play an important role in familial prostate cancer. This finding is potentially significant given the higher frequency of the K1019X nonsense mutation and the higher prevalence of prostate cancer among African American men compared to their US counterparts. The functional significance of the common K1019X nonsense mutation is not known, however our results in addition to previous functional analyses of wildtype EphB2 using the DU145 prostate cancer cell line<sup>1</sup>, suggests a pathogenic role for EphB2 in prostate cancer and further study is warranted.

**URL.**

The NCBI dbSNP website is available at [www.ncbi.nlm.nih.gov/SNP/](http://www.ncbi.nlm.nih.gov/SNP/).

**ACKNOWLEDGEMENTS**

The authors would like to first thank the families and study participants for their continued involvement in this research. We would also like to thank the study coordinators of the AAHPC study network, including M. Franklin, P. Roberson, E. Johnson, L. Faison-Smith, C. Meegan, M. Johnson, L. Kososki, C. Jones, R. Mejia. In addition we thank F. Akereyeni and C. Bonilla for technical assistance. This research was funded in part by the NIH Center for Minority Health and Health Disparities (1-HG-75418), the NCI (1U54CA91431-01) and the Department of Defense (DAMD17-00-1-0025 and DAMD 17-02-1-0067). A.B.B-B. also received support from USPHS grant CA-06927 and an appropriation from the Commonwealth of Pennsylvania.

**The authors declare that they have no competing interests.**

**License for publication statement**

"The Corresponding Author has the right to grant on behalf of all authors and does grant on behalf of all authors, an exclusive license (or non exclusive for government employees) on a worldwide basis to the BMJ Publishing Group Ltd to permit this article (if accepted) to be published in Journal of Medical Genetics and any other BMJPGL products and sublicenses such use and exploit all subsidiary rights, as set out in our license (<http://JMG.bmjournals.com/misc/ifora/licenceform.shtml>)."

## REFERENCES

1. Huusko P, Ponciano-Jackson D, Wolf M, *et al.* Nonsense-mediated decay microarray analysis identifies mutations of EPHB2 in human prostate cancer. *Nat Genet* 2004;**36**(9):979-83.
2. Ghafoor A, Jemal A, Cokkinides V, *et al.* Cancer statistics for African Americans. *CA Cancer J Clin* 2002;**52**(6):326-41.
3. Jemal A, Tiwari RC, Murray T, *et al.* Cancer statistics, 2004. *CA Cancer J Clin* 2004;**54**(1):8-29.
4. Makridakis NM, Ross RK, Pike MC, *et al.* Association of mis-sense substitution in SRD5A2 gene with prostate cancer in African-American and Hispanic men in Los Angeles, USA. *Lancet* 1999;**354**(9183):975-8.
5. Kittles RA, Panguluri RK, Chen W, *et al.* Cyp17 promoter variant associated with prostate cancer aggressiveness in African Americans. *Cancer Epidemiol Biomarkers Prev* 2001;**10**(9):943-7.
6. Zeigler-Johnson CM, Walker AH, Mancke B, *et al.* Ethnic differences in the frequency of prostate cancer susceptibility alleles at SRD5A2 and CYP3A4. *Hum Hered* 2002;**54**(1):13-21.
7. Miller DC, Zheng SL, Dunn RL, *et al.* Germ-line mutations of the macrophage scavenger receptor 1 gene: association with prostate cancer risk in African-American men. *Cancer Res* 2003;**63**(13):3486-9.
8. Cunningham GR, Ashton CM, Annegers JF, *et al.* Familial aggregation of prostate cancer in African-Americans and white Americans. *Prostate* 2003;**56**(4):256-62.
9. Oakley-Girvan I, Feldman D, Eccleshall TR, *et al.* Risk of early-onset prostate cancer in relation to germ line polymorphisms of the vitamin D receptor. *Cancer Epidemiol Biomarkers Prev* 2004;**13**(8):1325-30.
10. Easton DF, Schaid DJ, Whittemore AS, *et al.* Where are the prostate cancer genes? - A summary of eight genome wide searches. *Prostate* 2003;**57**(4):261-69.
11. Tavtigian SV, Simard J, Teng DH, *et al.* A candidate prostate cancer susceptibility gene at chromosome 17p. *Nat Genet* 2001;**27**(2):172-80.
12. Carpten J, Nupponen N, Isaacs S, *et al.* Germline mutations in the ribonuclease L gene in families showing linkage with HPC1. *Nat Genet* 2002;**30**(2):181-4.
13. Casey G, Neville PJ, Plummer SJ, *et al.* RNASEL Arg462Gln variant is implicated in up to 13% of prostate cancer cases. *Nat Genet* 2002;**32**(4):581-3.
14. Xu J, Zheng SL, Komiya A, *et al.* Germline mutations and sequence variants of the macrophage scavenger receptor 1 gene are associated with prostate cancer risk. *Nat Genet* 2002;**32**(2):321-5.
15. Rebbeck TR, Walker AH, Zeigler-Johnson C, *et al.* Association of HPC2/ELAC2 genotypes and prostate cancer. *Am J Hum Genet* 2000;**67**(4):1014-9.
16. Stanford JL, Sabacan LP, Noonan EA, *et al.* Association of HPC2/ELAC2 polymorphisms with risk of prostate cancer in a population-based study. *Cancer Epidemiol Biomarkers Prev* 2003;**12**(9):876-81.
17. Gibbs M, Stanford JL, McIndoe RA, *et al.* Evidence for a rare prostate cancer-susceptibility locus at chromosome 1p36. *Am J Hum Genet* 1999;**64**(3):776-87.
18. Matsui H, Suzuki K, Ohtake N, *et al.* Genomewide linkage analysis of familial prostate cancer in the Japanese population. *J Hum Genet* 2004;**49**(1):9-15.

19. Brown WM, Lange EM, Chen H, *et al.* Hereditary prostate cancer in African American families: linkage analysis using markers that map to five candidate susceptibility loci. *Br J Cancer* 2004;**90**(2):510-4.
20. Royal C, Baffoe-Bonnie A, Kittles R, *et al.* Recruitment experience in the first phase of the African American Hereditary Prostate Cancer (AAHPC) study. *Ann Epidemiol* 2000;**10**(8 Suppl):S68-77.
21. Powell IJ, Carpten J, Dunston G, *et al.* African-American heredity prostate cancer study: a model for genetic research. *J Natl Med Assoc* 2001;**93**(12 Suppl):25S-28S.
22. Ahaghotu C, Baffoe-Bonnie A, Kittles R, *et al.* Clinical characteristics of African-American men with hereditary prostate cancer: the AAHPC Study. *Prostate Cancer Prostatic Dis* 2004;**7**(2):165-9.
23. Panguluri RC, Long LO, Chen W, *et al.* COX-2 gene promoter haplotypes and prostate cancer risk. *Carcinogenesis* 2004;**25**(6):961-6.
24. Hoggart CJ, Parra EJ, Shriver MD, *et al.* Control of confounding of genetic associations in stratified populations. *Am J Hum Genet* 2003;**72**(6):1492-1504.
25. Kittles RA, Chen W, Panguluri RK, *et al.* CYP3A4-V and prostate cancer in African Americans: causal or confounding association because of population stratification? *Hum Genet* 2002;**110**(6):553-60.
26. Pritchard JK, Stephens M, Donnelly P. Inference of population structure using multilocus genotype data. *Genetics* 2000;**155**(2):945-59.
27. Freedman ML, Reich D, Penney KL, *et al.* Assessing the impact of population stratification on genetic association studies. *Nat Genet* 2004;**36**(4):388-93.
28. Gryfe R, Di Nicola N, Gallinger S, *et al.* Somatic instability of the APC I1307K allele in colorectal neoplasia. *Cancer Res* 1998;**58**(18):4040-3.
29. Laken SJ, Petersen GM, Gruber SB, *et al.* Familial colorectal cancer in Ashkenazim due to a hypermutable tract in APC. *Nat Genet* 1997;**17**(1):79-83.
30. Parsons R, Myeroff LL, Liu B, *et al.* Microsatellite instability and mutations of the transforming growth factor beta type II receptor gene in colorectal cancer. *Cancer Res* 1995;**55**(23):5548-50.
31. Kittles RA, Young D, Weinrich S, *et al.* Extent of linkage disequilibrium between the androgen receptor gene CAG and GGC repeats in human populations: implications for prostate cancer risk. *Hum Genet* 2001;**109**(3):253-61.
32. Sartor O, Zheng Q, Eastham JA. Androgen receptor gene CAG repeat length varies in a race-specific fashion in men without prostate cancer. *Urology* 1999;**53**(2):378-80.
33. Reichardt JK, Makridakis N, Henderson BE, *et al.* Genetic variability of the human SRD5A2 gene: implications for prostate cancer risk. *Cancer Res* 1995;**55**(18):3973-5.
34. Chakraborty R, Weiss KM. Admixture as a tool for finding linked genes and detecting that difference from allelic association between loci. *Proc Natl Acad Sci U S A* 1988;**85**(23):9119-23.
35. Lander ES, Schork NJ. Genetic dissection of complex traits. *Science* 1994;**265**(5181):2037-48.
36. Pritchard JK, Rosenberg NA. Use of unlinked genetic markers to detect population stratification in association studies. *Am J Hum Genet* 1999;**65**(1):220-8.

37. Devlin B, Roeder K. Genomic control for association studies. *Biometrics* 1999;**55**(4):997-1004.
38. Schork NJ, Fallin D, Thiel B, *et al.* The future of genetic case-control studies. *Adv Genet* 2001;**42**:191-212.
39. Reich DE, Goldstein DB. Detecting association in a case-control study while correcting for population stratification. *Genet Epidemiol* 2001;**20**(1):4-16.
40. Parra EJ, Kittles RA, Argyropoulos G, *et al.* Ancestral proportions and admixture dynamics in geographically defined African Americans living in South Carolina. *Am J Phys Anthropol* 2001;**114**(1):18-29.
41. Pfaff CL, Parra EJ, Bonilla C, *et al.* Population structure in admixed populations: effect of admixture dynamics on the pattern of linkage disequilibrium. *Am J Hum Genet* 2001;**68**(1):198-207.
42. Wacholder S, Rothman N, Caporaso N. Counterpoint: bias from population stratification is not a major threat to the validity of conclusions from epidemiological studies of common polymorphisms and cancer. *Cancer Epidemiol Biomarkers Prev* 2002;**11**(6):513-20.
43. Thomas DC, Witte JS. Point: population stratification: a problem for case-control studies of candidate-gene associations? *Cancer Epidemiol Biomarkers Prev* 2002;**11**(6):505-12.
44. Wang Y, Localio R, Rebbeck TR. Evaluating bias due to population stratification in case-control association studies of admixed populations. *Genet Epidemiol* 2004;**27**(1):14-20.
45. Helgason A, Yngvadottir B, Hrafnkelsson B, *et al.* An Icelandic example of the impact of population structure on association studies. *Nat Genet* 2005;**37**(1):90-5.