Published in final edited form as:
*Prostate*. 2010 June 15; 70(9): 1012–1019. doi:10.1002/pros.21135.

# *Hepsin* Gene Variants Association with Prostate Cancer Risk and Prognosis

**Sarah K. Holt**[1], **Erika M. Kwon**[2,3], **Daniel W. Lin**[1,4], **Elaine A. Ostrander**[2], and **Janet L. Stanford**[1,5]

[1] Division of Public Health Sciences, Fred Hutchinson Cancer Research Center, Seattle, WA

[2] Cancer Genetics Branch, National Human Genome Research Institute, NIH, Bethesda, MD

[3] Program in Human Genetics and Molecular Biology, Johns Hopkins University School of Medicine, Baltimore, MD

[4] Department of Urology, Veterans Affairs Puget Sound Health Care System and University of Washington, Seattle, WA

[5] Department of Epidemiology, University of Washington, Seattle, WA

## Abstract

**Background—***Hepsin* (*HPN*) is one of the most consistently overexpressed genes in prostate cancer and there is some evidence supporting an association between *HPN* gene variants and prostate cancer risk. We report results from a population-based case-control genetic association study for six tagging single nucleotide polymorphisms (tagSNPs) in the *HPN* gene.

**Methods—**Prostate cancer risk was estimated using adjusted unconditional logistic regression in 1,401 incident prostate cancer cases diagnosed in 1993 through 1996 or 2002 through 2005 and 1,351 age-matched controls. Risks of disease recurrence/progression and prostate cancer-specific mortality were estimated using Cox proportional hazards (PH) regression in 437 cases with long term follow-up.

**Results—**There were 135 recurrence/progression events and 57 cases who died of prostate cancer. Contrary to some earlier studies, we found no evidence of altered risk of developing prostate cancer overall or when clinical measures of tumor aggressiveness were considered for any of the tagSNPs, assessed either individually or by haplotypes. There was no evidence of altered risks of tumor recurrence/progression or prostate cancer death associated with variants in the *HPN* gene.

**Conclusions—**Germline genetic variation of *HPN* does not seem to contribute to risk of prostate cancer or prognosis.

### Keywords

prostate neoplasm; single nucleotide polymorphism; *Hepsin*; case-control

REQUESTS FOR REPRINTS: Sarah K. Holt, Fred Hutchinson Cancer Research Center, Mailstop: M4-A402, PO Box 19024, Seattle, WA 98109. Phone: 206-667-7880. Fax: 206-667-4787. skholt@fhcrc.org.

**Disclosure Statement**

The authors of this manuscript do not have any affiliations that are relevant or important with any organization that to our knowledge has a direct interest, financial or otherwise, in the subject matter discussed.

## Introduction

Hepsin is a type II integral membrane serine protease that has been found to be one of the most upregulated genes in prostate cancer [1]. The physiological function of hepsin remains unknown, but within the carcinogenesis pathway of the prostate it appears to play a role in cancer cell migration/invasion rather than cell proliferation. In probasin promoter-driven hepsin (PB-hepsin) transgenic mice it has been shown that overexpression of hepsin in the mice prostate epithelium results in disorganization of the basement membrane and weakening of the epithelial-stromal adhesion [2,3]. To further explore potential function of hepsin in prostate cancer progression, PB-hepsin transgenic mice were crossed with mice expressing SV40 large T antigen in the prostate (LPB-Tag mice, line 12T-7F). LPB-Tag mice are characterized by a high level of HG-PIN (high grade prostatic intraepithelial neoplasia) lesions and preserved prostate structure. Over 55% of the double transgenic LPB-Tag/PB-hepsin mice developed prostate cancer that metastasized to the bone, liver and lung while single transgenic LPB-Tag or PB-hepsin mice remained free of metastasis [2,3]. Furthermore, antibodies that neutralize hepsin protease activity have been show to inhibit cell invasion, but do not affect cell proliferation in prostate cancer cell culture models [4].

*Hepsin* (*HPN*) germline variants could potentially change gene expression profiles and alter protein function to affect either risk of developing prostate cancer or the likelihood of progressing to comparatively more aggressive disease that results in tumor metastasis or prostate cancer death. The window of prostate cancer progression in which hepsin is most likely operating is early, “promoting” PIN lesions to localized prostate cancer. Human cDNA microarray profiling studies show *HPN* expression is strongest in HG-PIN lesions, moderate to strong in localized prostate cancer, and weak in metastasized and hormone-refractory lesions [5]. A linkage study of affected brothers identified a locus on chromosome segments 19q12–q13.11, where the *HPN* gene is located, as linked to disease aggressiveness [6,7]. Further work by Pal *et al* identified five *HPN* SNPs with significant differences in allelic frequencies in the prostate cancer cases and healthy controls [8]. To investigate the potential associations between *HPN* genetic variation and disease risk overall, risk by clinical parameters of tumor aggressiveness, prostate cancer recurrence/progression, and prostate cancer-specific mortality, we conducted a population-based case-control study with on-going patient follow-up. To capture the underlying genetic variation of *HPN* we selected 6 tagging single nucleotide polymorphisms (tagSNPs) spanning the whole gene.

## Materials and Methods

### Study Population

Study subjects were enrolled in one of two population-based prostate cancer case-control studies that have been described previously [9,10]. Cases were Caucasian or African American and were newly diagnosed with histologically confirmed prostate cancer in the two study periods, either January 1, 1993 to December 31, 1996 (Study I, age range 40–64 years) or January 1, 2002 to December 31, 2005 (Study II, age range 35–74 years). Prostate cancer cases were identified from the metropolitan Seattle-Puget Sound population-based tumor registry that is operated as part of the National Cancer Institute’s Surveillance, Epidemiology, and End Results (SEER) program. Of the 1,754 eligible, interviewed cases we obtained peripheral blood leukocyte samples for genotyping from 1,458 men (Study I: n=631; Study II: n=827). We excluded 57 of these subjects because they did not have sufficient DNA for this specific study for a total of 1,401 cases. African American and Caucasian controls were recruited evenly throughout the ascertainment periods for cases using random digit telephone dialing (RDD) and frequency matched to cases by 5-year age groups. Of the 1,645 eligible interviewed controls we obtained peripheral blood leukocyte samples from 1,351 men (Study I: n=565; Study II: n=786). We excluded 30 of these

subjects because they did not have sufficient DNA for this specific study for a total of 1,321 controls.

The SEER registry provided information on tumor characteristics, primary therapy, vital status, and underlying cause of death. Death certificates were obtained to confirm fatal prostate cancer. The most recent registry linkage update for mortality was December 1, 2008. Cases included in survival analysis were limited to Caucasians (n=1,258). In January 2004 a self-administered follow-up questionnaire collecting information on use of secondary therapies, follow-up PSA results, and evidence for prostate cancer recurrence/progression was sent to cases in Study I. Out of the 631 cases who were alive and had consented to future contact, 520 completed the survey for a response level of 82%. An additional 51 cases who were diagnosed with local/regional disease and who were deceased at the time of the follow-up survey had recurrence/progression data available: 1) for 34 of these cases the next-of-kin provided consent for medical record review, which was used to determine recurrence/progression status; and, 2) 17 of these cases died of metastatic prostate cancer and were coded as having recurred. Of the 571 cases with data, subjects were excluded from recurrence/progression analyses if they had no DNA for genotyping (n=101), were African American (n=17), or had an initial diagnosis of metastatic disease (n=6), for a total of 437 eligible cases in the recurrence/progression analysis. This study was approved by Fred Hutchinson Cancer Research Center's Institutional Review Board and genotyping was approved by the Internal Review Board of the National Human Genome Research Institute.

## SNP Selection and Genotyping

SNPs that captured genetic variability in the *HPN* gene used publicly available data from the HapMap consortium[1] and the tagSNP algorithm LDselect available from SeattleSNPs[2]. Using the parameters of $r^2 \geq 0.8$ and minor allele frequency ≥ 5% [11], a total of 6 tagging SNPs for *HPN* (chromosome19q11–q13.2, length 26 kb) were chosen [12]. The Applied Biosystems (ABI) SNPlex™ Genotyping System was used for genotyping, and proprietary GeneMapper® software was used for calling alleles[3]. Discrimination of the specific SNP allele was determined by an ABI 3730*xl* DNA Analyzer and is based on the presence of a unique sequence assigned to the original allele-specific oligonucleotide. Quality control included genotyping of 76 blind duplicate samples, which revealed 99% agreement on genotyping calls across all SNPs assayed. In addition, each batch of DNA aliquots genotyped incorporated similar numbers of case and control samples, and laboratory personnel were blinded to the case-control status of samples. The call rate was ≥98% for all but one SNP (rs1688029, 97%).

## Statistical Analysis

SNP genotype frequencies were examined for Hardy-Weinberg equilibrium (HWE) using the $\chi^2$ statistic and all were found to be consistent ($P > 0.05$) with HWE among Caucasian controls. Data were analyzed using unconditional logistic regression to calculate an odds ratio (OR) as an estimate of relative risk of prostate cancer associated with SNP genotypes. We included age and stratified by race in all regression models. We assessed possible confounding effects of variables listed in Table 1 and found none appreciably altered risk estimates, thus did not include them as covariates. We used both dominant and co-dominant models, excepting SNPs for which no or too few individuals were homozygous for the variant genotype. Trend tests, which used a single indicator variable coded as the number of variant alleles for each SNP, were used to assess gene dosage. Global tests of association,

[1] www.hapmap.org
[2] pga.gs.washington.edu
[3] www.appliedbiosystems.com

which were estimated by comparing an adjusted model that included all SNPs to the null model that only included age, automatically adjusted for multiple testing based on degrees of freedom of the corresponding $\chi^2$ test [13]. Multiple comparisons were also accounted for by using permutations to calculate exact p-values for each significant individual SNP ($\alpha$= 0.05). Haploview software version 4.1 was used to generate pairwise LD estimates and define haplotype blocks [14]. Haplotype risk was assessed within the each block using HPlus version 3.1, which employs an empirical estimating equation (EE) technique [15]. Associations with individual htSNPs according to Gleason score [2−7(3+4) vs. 7(4+3)−10], tumor stage (local vs. regional/distant) and a composite prostate cancer aggressiveness score were examined using polytomous regression. Classification parameters of the composite score for the "high" aggressive phenotype included death from prostate cancer *or* initial diagnosis of distant metastases. All other cases were classified as "low/moderate".

To examine associations between individual SNPs with prostate cancer recurrence/progression and prostate cancer-specific mortality we used Cox proportional hazards (PH) regression models adjusting for age, Gleason score [2−7(3+4) and 7(4+3)−10], stage at diagnosis (local, regional, and distant), diagnostic PSA (prostate-specific antigen) level (0–9.9 vs. ≥10.0 ng/mL) and primary treatment [radical prostatectomy, radiation with or without androgen deprivation therapy (ADT), ADT only, other treatment, and active surveillance]. Analysis of outcomes was restricted to Caucasians because sample size within African Americans was too limited. Evidence of recurrence included: biochemical (PSA ≥ 0.2 ng/mL after radical prostatectomy, PSA ≥ nadir PSA level plus 2.0 ng/mL for men treated with radiation therapy, or rising PSA while on primary ADT); receipt of secondary treatment [ADT (medications or orchiectomy), radiation or chemotherapy]; a positive bone scan, biopsy, or MRI showing evidence of prostate cancer after primary treatment; and/or, a self-reported physician's diagnosis of prostate cancer recurrence/progression. For estimating recurrence/progression risk, the time-dependant variable was defined as time from diagnosis to the first reported evidence of recurrence. The censoring date was the date that the follow-up questionnaire was returned. For those cases that died of prostate cancer prior to the follow-up survey, time to recurrence was estimated using multiple imputations from patients who died from prostate cancer and who had a known time to recurrence [16]. This method used a linear regression model that included attempted curative treatment (radical prostatectomy or radiation therapy vs. neither) and Gleason score (<7 vs. ≥ 7). In cases where the imputed time to recurrence was longer than the time to prostate cancer death, a uniformly distributed error from 0 to 1 year was subtracted from time to prostate cancer death and used as imputed time to. For calculation of prostate cancer-specific mortality, the time-dependant variable was defined as the time from diagnosis to death or censoring for men who remained alive. Living cases were censored on the date of the most recent linkage with the cancer registry. Cases that died from other causes were censored at the time of death. All analyses, with the exception of the haplotype analyses, were done using the STATA statistical package (version 10.1, STATA Corp., College Station, TX).

## Results

Cases and controls were similar in age (mean in cases, 59.8 years; in controls, 59.3 years). Cases had a higher percentage of African Americans and subjects reporting a family history of prostate cancer (Table 1). The majority of prostate cancers were local stage tumors with low/moderate Gleason scores. There was no strong evidence of altered risk of developing prostate cancer for any tagSNPs evaluated (Table 2). TagSNPs rs2451996, rs1688043, rs2305746 and rs2305747 formed a block for haplotype analysis. The most common haplotype (TAGT) at a frequency of 58% served as the reference group. The three haplotypes with frequencies greater than 5% (CAGC at 20%, CAGT at 16%, and CGAC at 6%) did not show any significant associations with prostate cancer risk (data not shown).

Stratification by measures of tumor aggressiveness did not reveal significant associations of risk with genotypes (data not shown).

There were 135 events with self-reported physician-diagnosed recurrence/progression, with an average 8.1 years of follow-up (range 0.6–12.8 years) after diagnosis. Risk of recurrence/progression was associated with diagnosis under the age of 50 years, regional stage of disease (cases with distant stage of disease were excluded from risk of recurrence/progression analyses), higher Gleason score, and a diagnostic PSA of more than 10 ng/mL. There were 57 cases who died of prostate cancer in the average 8.4 years of follow-up (range 0.8–15.9 years). Survival was associated with localized stage of disease, lower Gleason score, and a diagnostic PSA less than 10 ng/mL. None of the tagSNPs showed altered risks of disease recurrence/progression or death from prostate cancer (Table 3).

## Discussion

We did not find any evidence that germline genetic variation in the *HPN* gene altered risk of prostate cancer overall, risk of more aggressive phenotypes of prostate cancer, risk of prostate cancer recurrence/progression or risk of prostate cancer death. This study was prompted by hepsin expression patterns observed in malignant prostate cells and genetic associations found by Pal *et al* [8]. These authors genotyped for 11 SNPs spanning ~26 kb in the *HPN* gene that were selected from the NCBI database (www.ncbi.nih.gov) based on SNP location, validity status and heterozygosity. We took a tagSNP approach in a 36 kb region that included 5 kb upstream and downstream of the *HPN* gene. Although we covered a larger region as compared to Pal *et al*, we only genotyped for 6 tagSNPs and included only two SNPs, rs2305747 and rs2305746, from the five SNPs previously identified to be associated with prostate cancer risk. This was because we used only SNPs in the HapMap dataset with allelic frequencies greater than 5%. Two of the 5 SNPs (rs10410046 and rs1350290) shown to be associated with risk of prostate cancer in the Pal study had allelic frequencies less than 5% in the HapMap Caucasian population and were in high linkage disequilibrium (LD) with rs2305746, thus were not included. Another SNP, rs2305745, was in perfect LD with rs2305747, thus we only genotyped for the latter to avoid redundancy.

Pal *et al*. reported five SNPs, all located within a single haplotype block, had significant allele frequency differences between prostate cancer cases and controls [8]. We did not observe an association with prostate cancer risk for either of rs2305747 and rs2305746, which covers this haplotype block in our dataset. The role of hepsin in prostate carcinogenesis seems to be in aiding progression from localized lesions to metastasis via disorganization of the cellular adhesion complex, thus germline polymorphisms may have an impact on disease aggressiveness rather than disease initiation. Pal *et al*. reported SNP rs1688043 had a reported association between MAF (minor allele frequency) and tumor aggressiveness, as measured by Gleason scores of 2–6 versus 7–9 [8]. We did not replicate this finding when we stratified by Gleason score, by tumor stage, or by more stringent composite prostate cancer aggressiveness score. Furthermore, none of the tagSNPs included in this study were associated with tumor recurrence/progression or prostate cancer death.

The discrepancy between these findings could be due to the different populations of the studies. This study identified 1,258 sporadic Caucasian cases from the SEER registry which had only 21.4% reporting a first-degree relative with prostate cancer, while the 590 familial cases in the Pal study were from 304 prostate cancer families that were part of a genome screen of multiplex sibships with prostate cancer. Not only could risk estimates not be comparable in familial versus sporadic prostate cancer cases, but cases in the Pal study were slightly older (mean of 65.5 years, range 42–91, versus 59.8, range 35–74). We did attempt to make our subjects more comparable to the previous study by limiting our cases to those

with a reported family history of prostate cancer and controls with no reported family history, but this did not appreciable alter our results. Family history of prostate cancer in a first degree relative was not found to be an effect modifier or confounder in our dataset. While the Pal *et al.* statistical analysis accounted for relatedness of the subjects it did not adjust for age or any other factors and only compared allele frequencies using a likelihood ratio test rather than presenting risk estimates. We could, however, directly compare the MAF for SNPs that were genotyped in both studies. The MAF in cases and controls were similar with the exception of rs1688043, the SNP identified by Pal *et al*. to be associated with Gleason. Pal *et al.* reported a control MAF of 15% whereas our control MAF was 6% which matches the HapMap Caucasion reported MAF of 6%.

Controls also differed between the two studies. Our controls were age-matched (range 35–74 years) and population-based while the Pal controls were from a screening population that was over 65, had no registered PSA level above 2.5 ng/ml, never had a DRE suspicious for prostate cancer, and had no known history of prostate cancer. In contrast to these eligibility requirements, in our controls only 13.2% were 65 or over, 11.2% reported a first-degree relative with prostate cancer, and 13.5% had no reported screening test (either PSA or DRE) in the 5 years prior to enrollment in the study. We did account for this potential misclassification of disease in the control population by measuring PSA levels in plasma that was collected at time of interview. Only 4.4% of controls with no reported screening history had PSA values greater than or equal to 4.0 ng/ml. Furthermore, exclusion of the 18 controls had measured PSA greater than 10 ng/mL had no effect on the results reported here.

Hepsin has been shown to be over-expressed in 90% of prostate tumors, however the mechanism by which this up-regulation occurs is not clear [17,18]. Germline polymorphisms can potentially impact gene expression at any level of gene regulation: transcription (e.g., modification of a regulatory binding site), post-transcription (e.g., variations affecting alternate splicing or mRNA stability) or translation (e.g. differences in mRNA sequence affecting translational efficiency). There are not any known polymorphisms in the coding region of *HPN*, non-synonymous or synonymous, nor have there been any polymorphisms reported with demonstrated functional consequences in factors affecting gene expression. None of our *HPN* tagSNPs are evolutionarily conserved (conservation score < 0.001) and are all intronic, with the exception of rs1688029, which is located in the 3′ UTR. The tagSNPs included in this study were selected to cover genetic variability of the gene with the hope that any significant findings would lead to subsequent functional studies. These findings do not provide any leads to better understand role of the *HPN* gene in prostate cancer etiology.

The mechanisms by which hepsin affects prostate cancer progression are not fully understood, thus a pathway approach to discovering a combination of *HPN* related gene variants contributing to a phenotype such as aggressive prostate cancer may be challenging. Disorganization of the basement membrane via enzymatic activity of hepsin may be through activation of urokinase-type plasminogen activator (uPA) and tissue-type plasminogen activator and the subsequent initiation of the plasminogen/plasmin proteolytic pathway [19]. Impairment of the basement membrane by plasmin arises partly through the degradation of proteins such as laminin-332, an extracellular matrix (ECM) molecule responsible for cellular adhesion [20]. Not only is hepsin a substrate for pro-uPA, but it is also a substrate for laminin-332. Cleaving this ECM molecule leads to enhanced prostate cancer cell migration *in vitro* [21]. Interestingly, transgenic mice models where overexpression of hepsin in prostate tissue showed disorganization of the basement membrane also showed weaker staining of laminin-332 [2].

This analysis was unique in that it examined *HPN* genetic variation with both prostate cancer risk and prostate cancer outcomes in a dataset that allowed for long-term follow-up of cases in a population-based study. To test the assumption that HPN germline variants could be associated with aggressive disease Pal *et al*. measured the association between Gleason scores and allelic frequencies while we were able to examine multiple clinical parameters as well as outcomes such as tumor recurrence and prostate cancer death. We were able to control for potential selection bias by accounting for screening practices in the entire study sample, examine potential confounders such as family history, and stratify the analysis by various definitions of disease aggressiveness. If response to the follow-up survey was correlated with clinical features then results of the recurrence/progression analyses could be biased. Comparing the 520 responders to the 111 non-responders revealed that responders were younger (p=0.002), but no correlation was found with clinical features such as Gleason score, stage, or PSA at diagnosis. Although we took the tagSNP approach to allow for more comprehensive gene coverage, we were limited to known genetic variation of *HPN* within publicly available data from the HapMap consortium. There may be unknown variants, especially in regulatory regions, which are not in strong linkage disequilibrium with any of our genotyped SNPs. Another weakness of our study was the sample size with respect to the analysis of prostate cancer recurrence/progression and prostate cancer death. We were also underpowered to examine risk within African Americans; moreover, since tagSNP selection was based on a Caucasian population, gene coverage for the African American population is not complete.

## Conclusion

We did not find associations with prostate cancer risk overall, aggressive prostate cancer, prostate cancer recurrence/progression or prostate cancer-specific mortality with any of the *HPN* tagSNPs included in this study. However, we recognize that the limited number of prostate cancer deaths and cases with recurrence/progression data increases the plausibility of false-negative findings for the outcome analyses. *HPN* expression in prostate tissue may be a superior phenotypic measurement as compared to examining prostate cancer risk or outcomes for capturing germline variants that impact gene expression or protein function, thus future studies could potentially correlate genotypic variation with *HPN* expression in prostate tissue. Lack of replication of the previous study done by Pal *et al.* may be due to differences in the populations studied, previous significant findings could be attributed to chance, or, at least in the instance of the SNP rs1688043, genotyping error.

## Acknowledgments

**FINANCIAL SUPPORT:** This work was supported by grants RO1-CA56678, RO1-CA092579, RO1-CA082554 and P50-CA097186 from the National Cancer Institute; additional support was provided by the Fred Hutchinson Cancer Research Center and the Intramural Program of the National Human Genome Research Institute.

## References

1. Wu Q, Parry G. Hepsin and prostate cancer. Front Biosci. 2007; 12:5052–9. [PubMed: 17569629]
2. Klezovitch O, Chevillet J, Mirosevich J, Roberts RL, Matusik RJ, Vadioukhin V. Hepsin promotes prostate cancer progression and metastasis. Cancer Cell. 2004; 6:185–95. [PubMed: 15324701]
3. Vasioukhin V. Hepsin paradox reveals unexpected complexity of metastatic process. Cell Cycle. 2004; 3:1394–7. [PubMed: 15539945]
4. Xuan JA, Schneider D, Toy P, Lin R, Newton A, Zhu Y, Finster S, Vogel D, Mintzer B, Dinter H, Light D, Parry R, Polokoff M, Whitlow M, Wu Q, Parry G. Antibodies neutralizing hepsin protease activity do not impact cell growth but inhibit invasion of prostate and ovarian tumor cells in culture. Cancer Res. 2006; 66:3611–9. [PubMed: 16585186]

5. Dhanasekaran SM, Barrette TR, Ghosh D, Shah R, Varambally S, Kurachi K, Pienta KJ, Rubin MA, Chinnaiyan AM. Delineation of prognostic biomarkers in prostate cancer. Nature. 2001; 412:822–6. [PubMed: 11518967]
6. Witte JS, Elston RC, Lin J, Suarez BK, Broman KW, Burmester JK, Weber JL, Catalona WJ. Genome-wide scan for prostate cancer-aggressiveness loci. Am J Hum Genet. 2000; 67:92–9. [PubMed: 10825281]
7. Witte JS, Suarez GK, Thiel B, Lin J, Yu A, Banerjee TK, Burmester JK, Casey G, Catalona WJ. Genome-wide scan of brothers: Replication and fine mapping of prostate cancer susceptibility and aggressiveness loci. Prostate. 2003; 57:298–308. [PubMed: 14601026]
8. Pal P, Xi H, Kaushal R, Sun G, Jin CH, Jin L, Suarez BK, Catalona WJ, Deka R. Variants in the HEPSIN gene are associated with prostate cancer in men of European origin. Hum Genet. 2006; 120:187–92. [PubMed: 16783571]
9. Stanford JL, Wicklund KG, McKnight B, Daling JR, Brawer MK. Vasectomy and risk of prostate cancer. Cancer Epidemiol Biomarkers Prev. 1999; 8:881–6. [PubMed: 10548316]
10. Agalliu I, Salinas CA, Hansten PD, Ostrander EA, Stanford JL. Statin use and risk of prostate cancer: results from a population-based epidemiologic study. Am J of Epidemiol. 2008; 168:250–60. [PubMed: 18556686]
11. Clayton D, Leung HT. An R package for analysis of whole-genome association studies. Hum Hered. 2007; 64:45–51. [PubMed: 17483596]
12. Carlson CS, Eberle MA, Rieder MJ, Yi Q, Kruglyak L, Nickerson DA. Selecting a maximally informative set of single-nucleotide polymorphisms for association analyses using linkage disequilibrium. Am J Hum Genet. 2004; 74:106–20. [PubMed: 14681826]
13. Chapman JM, Cooper JD, Todd JA, Clayton DG. Detecting disease associations due to linkage disequilibrium using haplotype tags: a class of tests and the determinants of statistical power. Human Heredity. 2003; 56:18–31. [PubMed: 14614235]
14. Barrett JC, Fry B, Maller J, Daly MJ. Haploview: analysis and visualization of LD and haplotype maps. Bioinformatics. 2005; 21:263–5. [PubMed: 15297300]
15. Li SS, Khalid N, Carlson C, Zhao LP. Estimating haplotype frequencies and standard errors for multiple single nucleotide polymorphisms. Biostatistics. 2003; 4:513–22. [PubMed: 14557108]
16. Harrell, FE. Regression modeling strategies: with applications to linear models, logistic regression, and survival analysis. New York: Springer; 2001. p. 147-156.
17. Magee JA, Araki T, Patil S, Ehrig T, True L, Humphrey PA, Catalona WJ, Watson MA, Milbrandt J. Expression profiling reveals hepsin overexpression in prostate cancer. Cancer Research. 2001; 61:5692–6. [PubMed: 11479199]
18. Stephan C, Yousef GM, Scorilas A, Jung K, Jung M, Kristiansen G, Hauptmann S, Kishi T, Nakamura T, Loening SA, Diamandis EP. Hepsin is highly over expressed in and a new candidate for a prognostic indicator in prostate cancer. J Urol. 2004; 171:187–91. [PubMed: 14665873]
19. Moran P, Li W, Fan B, Vij R, Eigenbrot C, Kirchholfer D. Pro-urokinase-type plasminogen activator is a substrate for hepsin. J Biol Chem. 2006; 281:30439–46. [PubMed: 16908524]
20. Ogura Y, Matsunaga Y, Nishiyama T, Amano S. Plasmin induces degradation and dysfunction of laminin 332 (laminin 5) and impaired assembly of basement membrane at the dermal-epidermal junction. Br J Dermatol. 2008; 159:49–60. [PubMed: 18460030]
21. Tripathi M, Nandana S, Yamashita H, Ganesan R, Kirchhofer D, Quaranta V. Laminin-332 is a substrate for hepsin, a protease associated with prostate cancer progression. J Biol Chem. 2008; 283:30576–84. [PubMed: 18784072]

## Table 1

Distribution and risk estimates for selected characteristics of cases and controls, King County, Washington, 1993–1996 and 2002–2005

| | Cases (%) N = 1,401 | Controls (%) N = 1,321 | Adjusted OR[a] (95% CI) |
|---|---|---|---|
| Age group | | | |
| 35–49 | 113 (8.1) | 124 (9.4) | |
| 50–54 | 211 (15.1) | 199 (15.1) | |
| 55–59 | 343 (24.5) | 350 (26.5) | |
| 60–64 | 408 (29.1) | 342 (25.9) | |
| 65–69 | 173 (12.3) | 164 (12.4) | |
| 70–74 | 153 (10.9) | 142 (10.7) | |
| Race | | | |
| Caucasian | 1,258 (89.8) | 1,241 (93.9) | ref |
| African American | 143 (10.2) | 80 (6.1) | 1.81 (1.36–2.41) |
| First-degree relative with prostate cancer | | | |
| No | 1,101 (78.6) | 1,173 (88.8) | ref |
| Yes | 300 (21.4) | 148 (11.2) | 2.17 (1.75–2.68) |
| Stage of PCa at diagnosis | | | |
| Local | 1,098 (78.4) | | |
| Regional | 267 (19.1) | | |
| Distant | 36 (2.6) | | |
| Gleason score at diagnosis | | | |
| 2−6, 7 (3+4) | 1,173 (83.7) | | |
| 7(4+3), 8−10 | 222 (15.8) | | |
| Unknown | 6 (0.4) | | |
| Composite Aggressiveness Score[b] | | | |
| High | 77 (5.5) | | |
| Moderate/Low | 1,324 (94.5) | | |

[a] Adjusted for age.

[b] Classification parameters of the composite score for “high” aggressive phenotype included death from prostate cancer *or* initial diagnosis of distant metastasis. “Moderate/Low” included all other cases.

**Table 2**

Genotype distribution and odds ratios (95% CI) for associations between *HPN* genotypes and prostate cancer risk by race

| SNP | Genotype | Caucasians | | | | African Americans | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | Cases (%) (n=1,258)[a] | Controls (%) (n=1,241)[a] | Adjusted OR[b] (95% CI) | $P_{trend}$[c] | Cases (%) (n=143)[a] | Controls (%) (n=80)[a] | Adjusted OR[b] (95% CI) | $P_{trend}$[c] |
| *rs12461158* | GG | 692(55.4) | 651(52.4) | ref | | 71(50.0) | 41(51.3) | ref | |
| | AG | 476(38.1) | 506(40.7) | 0.89 (0.75–1.04) | | 56(39.4) | 33(41.3) | 0.97 (0.52–1.78) | |
| | AA | 81(6.5) | 85(6.8) | 0.90 (0.65–1.24) | | 15(10.6) | 6(7.5) | 1.35 (0.47–3.88) | |
| | AA+AG | | | 0.89 (0.76–1.04) | *0.18* | | | 1.03 (0.58–1.84) | *0.72* |
| *rs2451996* | TT | 425(34.2) | 409(33.1) | ref | | 14(9.9) | 7(8.8) | ref | |
| | CT | 612(49.3) | 593(48.0) | 0.99 (0.83–1.18) | | 49(34.5) | 42(52.5) | 0.46 (0.16–1.35) | |
| | CC | 204(16.4) | 233(18.9) | 0.84 (0.67–1.06) | | 79(55.6) | 31(38.8) | 1.16 (0.40–3.33) | |
| | CC+CT | | | 0.95 (0.80–1.12) | *0.20* | | | 0.76 (0.28–2.08) | *0.09* |
| *rs1688043* | AA | 1,089(88.0) | 1,084(88.0) | ref | | 95(67.9) | 57(73.1) | ref | |
| | AG | 145(11.7) | 142(11.5) | 1.01 (0.79–1.30) | | 40(28.6) | 21(26.9) | -- | |
| | GG | 4(0.3) | 6(0.5) | 0.66 (0.19–2.34) | | 5(3.6) | 0(0.0) | -- | |
| | GG+AG | | | 1.00 (0.78–1.27) | *0.90* | | | 1.23 (0.65–2.35) | |
| *rs2305746* | GG | 1,093(88.0) | 1,093(88.1) | ref | | 88(62.0) | 55(68.8) | ref | |
| | AG | 145(11.7) | 141(11.4) | 1.03 (0.80–1.31) | | 47(33.1) | 22(27.5) | 1.23 (0.65–2.34) | |
| | AA | 4(0.3) | 7(0.6) | 0.57 (0.17–1.94) | | 7(4.9) | 3(3.8) | 1.24 (0.30–5.19) | |
| | AA+AG | | | 1.00 (0.79–1.28) | *0.88* | | | 1.23 (0.67–2.27) | *0.53* |
| *rs2305747* | TT | 668(54.5) | 673(55.0) | ref | | 22(15.5) | 12(15.0) | ref | |
| | CT | 485(39.6) | 457(37.3) | 1.07 (0.91–1.26) | | 62(43.7) | 48(60.0) | 0.64 (0.27–1.49) | |
| | CC | 73(6.0) | 94(7.7) | 0.78 (0.56–1.08) | | 58(40.8) | 20(25.0) | 1.49 (0.60–3.71) | |
| | CC+CT | | | 1.02 (0.87–1.20) | *0.62* | | | 0.89 (0.40–2.00) | *0.15* |
| *rs1688029* | AA | 571(47.3) | 571(47.0) | ref | | 17(12.2) | 10(12.5) | ref | |
| | AG | 520(43.0) | 517(42.6) | 1.00 (0.85–1.19) | | 51(36.7) | 39(48.8) | 0.64 (0.25–1.65) | |
| | GG | 117(9.7) | 126(10.4) | 0.93 (0.70–1.23) | | 71(51.1) | 31(38.8) | 1.26 (0.49–3.26) | |
| | GG+AG | | | 0.99 (0.84–1.16) | *0.72* | | | 0.91 (0.37–2.23) | *0.19* |

[a] Variable numbers of cases and controls reflect failed genotyping.

[b] Adjusted for age. If there are no case or control homozygote carriers of the less common allele, then only the dominant model risk estimate is shown.

[c] Analysis for linear trend according to the number of variant alleles. If there are homozygote carriers of the less common allele this analysis is omitted.

**Table 3**

Hazard ratios (95% CI) for Prostate Cancer Recurrence/Progression and Prostate Cancer-Specific Deaths by *HPN* Genotypes Among Caucasian Men

| SNP | Genotype | Recurrence/Progression (n=437 with follow-up) | | Prostate Cancer-Specific Deaths (n=1,258 with follow-up) | |
|---|---|---|---|---|---|
| | | No. of events (n=135) [a] | HR (95% CI) [b] | No. of events (n=57) [a] | HR (95% CI) [b] |
| *rs12461158* | GG | 67 | Ref | 28 | Ref |
| | AA+AG | 65 | 1.12 (0.79–1.59) | 29 | 1.16 (0.67–2.02) |
| *rs2451996* | TT | 37 | Ref | 15 | Ref |
| | CC+CT | 94 | 1.39 (0.94–2.06) | 39 | 1.65 (0.87–3.13) |
| *rs1688043* | AA | 114 | Ref | 47 | Ref |
| | GG+AG | 14 | 1.05 (0.59–1.86) | 8 | 1.35 (0.62–2.97) |
| *rs2305746* | GG | 117 | Ref | 48 | Ref |
| | AA+AG | 14 | 1.05 (0.59–1.85) | 8 | 1.32 (0.60–2.89) |
| *rs2305747* | TT | 64 | Ref | 27 | Ref |
| | CC+CT | 65 | 1.22 (0.86–1.75) | 28 | 1.58 (0.90–2.76) |
| *rs1688029* | AA | 47 | Ref | 22 | Ref |
| | GG+AG | 72 | 1.21 (0.83–1.77) | 28 | 1.29 (0.72–2.33) |

[a] Variable numbers of events reflect failed genotyping.

[b] Hazard Ratios (HR) adjusted for age, Gleason, stage, diagnostic PSA level, and primary treatment.