Published in final edited form as:
Nat Genet. 2008 September ; 40(9): . doi:10.1038/ng0908-1032.

Variation in *KLK* Genes, Prostate Specific Antigen, and Risk of Prostate Cancer

Jiyoung Ahn[1], **Sonja I. Berndt**[1], **Sholom Wacholder**[1], **Peter Kraft**[2], **Adam S. Kibel**[3], **Meredith Yeager**[1], **Demetrius Albanes**[1], **Edward Giovannucci**[4], **Meir J. Stampfer**[4], **Jarmo Virtamo**[5], **Michael J. Thun**[6], **Heather Spencer Feigelson**[6], **Geraldine Cancel-Tassin**[7], **Olivier Cussenot**[7], **Gilles Thomas**[1], **David J. Hunter**[1,2], **Joseph F. Fraumeni Jr**[1], **Robert Hoover**[1], **Stephen J. Chanock**[1], and **Richard B. Hayes**[1]

[1] Division of Cancer Epidemiology and Genetics, National Cancer Institute, National Institutes of Health, Department of Health and Human Services, Bethesda, Maryland 20892, USA [2] Program in Molecular and Genetic Epidemiology, Department of Epidemiology, Harvard School of Public Health, Boston, MA, USA [3] Division of Urologic Surgery, Washington University School of Medicine, St. Louis, MO [4] Department of Nutrition, Harvard School of Public Health, Boston, MA, USA [5] Department of Health Promotion and Chronic Disease Prevention, National Public Health Institute, Helsinki FIN-0030, Finland [6] Department of Epidemiology and Surveillance Research, American Cancer Society, Atlanta, Georgia 30329, USA [7] CeRePP Hopital Tenon, Assistance Publique-Hôpitaux de Paris, Paris 75020, France

Genome-wide association studies have identified SNPs associated with prostate cancer. Recently, SNPs in *KLK3* were related to prostate cancer risk in a genome-wide association study (GWAS; 1,854 cases)[1] and two candidate gene investigations (596 and 209 cases, respectively)[2,3], raising the possibility that this gene and its encoded protein PSA (prostate specific antigen; kallikrein-related peptidase), which is widely-used as a biomarker for prostate cancer detection, are etiologically related to this disease.

Understanding the contribution of common genetic variation in *KLK3* to prostate cancer risk is daunting because SNPs in *KLK3* are determinants of serum PSA concentrations[1,2,4], and thus, observed risk relationships could be due to differential identification of cases, in settings where PSA is used in screening or clinical diagnosis of this disease. Therefore, we evaluated the association of 24 tagSNPs in the *KLK* region (including *KLK3* and nearby genes *KLK1*, *KLK2*, and *KLK15*, chromosome 19: 56022744–56073589) and prostate cancer risk in 1,172 prostate cancer cases and 1,157 PSA screened controls in men of European ancestry from the Prostate, Lung, Colorectal and Ovarian (PLCO) Cancer Screening Trial[5,6], in which serum PSA was used in screening for this disease. We also examined the prostate cancer risk relationship for 12 of these tagSNPs in the *KLK3* region (chromosome 19: 56031744–56063231) in four independent studies including 4,020 prostate cancer cases and 4,028 controls (American Cancer Society Cancer Prevention Study II, the Health Professionals Follow-up Study, the CeRePP French Prostate Case-Control Study, and the Alpha-Tocopherol, Beta-Carotene Cancer Prevention Study)[6], as a component of the Cancer Genetic Markers of Susceptibility (CGEMS)-GWAS[6] (http://cgems.cancer.gov; Supplementary Methods).

Correspondence to: Richard B. Hayes, DDS, PhD, Division of Cancer Epidemiology and Genetics, National Cancer Institute, 6120 Executive Blvd., Bethesda, MD 20892; Hayesr@mail.nih.gov.

None of the 24 *KLK*-region tagSNPs showed strong evidence for association with prostate cancer in PLCO (Table 1: Cases vs Controls in PLCO) and none of the 12 *KLK3*-region SNPs was significant in the CGEMS combined analyses (Table 1: Cases vs Controls in CGEMS combined). No substantial differences in risk were noted by disease aggressiveness (Supplement Table 1). Two tagSNPs (rs1058205 and rs2735839), located in the 3′ untranslated region of KLK3, in high linkage disequilibrium (LD; r2=0.8; Figure 1), along with four variants 8–23 kb upstream of the gene (rs2659056, rs2569729, rs266849, and rs266870) and several variants about 7 kb downstream (rs1506684 and rs2569739, in high LD, r^2=1.0), were associated with strong differentials in serum PSA concentrations (e.g., rs2735839, p trend=4.1×10^{-9}; Table 1, Adjusted Mean PSA Level, and Figure 1-Panel A). However, when these SNPs were simultaneously included in the multivariate model, associations remained for rs2735839, but others were no longer significantly associated with PSA concentrations.

Eeles *et al* [1] reported strong associations between three *KLK3* SNPs (p-values for stage 1 were 1.2×10^{-7}, 1.0×10^{-16} and 2.4×10^{-20} for rs2659056, rs266849, and rs2735839) and prostate cancer, in a GWAS, where controls were selected by design for low PSA (<0.5 ng/ml) and no limitations were placed on case-group PSA values. The replications in non-PSA selected groups showed attenuated associations (p-values for stage 2 were 0.424, 0.228 and 0.0002, respectively). Pal *et al* [2] also reported a strong association of prostate cancer with *KLK3* variants, where control subjects were also preferentially selected by PSA levels (< 2.5 ng/ml). When we preferentially limited the men in the control group in the PLCO Trial GWAS to those with PSA <0.5 ng/ml, as an exercise, we show prostate cancer risk associations (Figure 1-Panel B [▲] and Supplement Table 2) greater than 5-fold for some genotypes (rs1058205, OR=0.19, 95%CI=0.10–0.36, p=0.000015), whereas the same SNPs showed no clear association when the full control group was employed (Figure 1-Panel B [●] and Table 1). Furthermore, modest *KLK3* associations with prostate cancer risk observed in our full case-control series and in other studies where controls are not selected by design for PSA level[3] may still be related to PSA-based case ascertainment, because PSA is widely used as a clinical aid in selecting men for biopsy for suspect prostate cancer. In PLCO, *KLK* SNPs were also not associated with prostate cancer when only cases and controls with high PSA levels or only cases and controls with low PSA levels were compared (Supplement Table 3). Thus, the observed associations between the *KLK* SNPs and prostate cancer risk in earlier studies[4,5,3] and in our exercise in PLCO may be due to selection for PSA differentials and not causally related to underlying disease risk.

PSA is a serine protease, functioning in the liquefaction of seminal coagulum[7]; consistent with our data, Pal *et al*[2] showed serum PSA associations with rs266849 and rs1058205, and Cramer *et al*[4] showed that two SNPs in the 5′ region of *KLK3* are related to increased *KLK3* promoter activity and serum PSA level. In our study, the association with PSA was strongest with rs2735839, which is highly correlated with rs1058205 (r^2=0.8); associations observed with the other SNPs in our study appeared to be largely driven by their linkage disequilibrium with rs2735839. Sliding window haplotype analysis (Supplement Figure 1) and close inspection of the most strongly associated haplotype window (Supplementary Table 4) suggested that additional unmeasured risk variant(s) may influence PSA concentrations.

PSA is an early marker predictive for the presence of prostate cancer and its use has contributed to the near doubling of diagnosed cases over the past two decades in countries where use is common[8]. While trials are underway to determine whether screening with PSA reduces prostate cancer mortality[5,9], it is apparent that many men are diagnosed with clinically indolent cancer on the basis of PSA screening; the over-detection rate is estimated to range from 18 to 56%[9,10]. In addition, the majority of men with an elevated PSA have a

negative biopsy[11]. In our study, *KLK3* variant status was related to PSA test-positivity in control men (Supplementary Table 5). While we show that *KLK3* SNPs are not risk factors highly correlated with prostate cancer *per se*, a combination of *KLK3* SNP and serum PSA monitoring, possibly including a broader spectrum of protein and DNA tests[12], is a potential tool for advancement in prostate cancer early detection.

In conclusion, SNPs in the *KLK3* region were not associated strongly with prostate cancer in large series of cases and controls, where controls are not selected by design for low PSA and even the modest *KLK3*-prostate cancer risk associations observed in non-selected populations may be due to PSA-directed differential identification of prostate cancer cases with particular *KLK3*-PSA profiles.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

The Prostate Lung Colorectal Ovarian Cancer Screening Trial (PLCO) is supported by the Intramural Research Program of the Division of Cancer Epidemiology and Genetics and by contracts from the Division of Cancer Prevention, National Cancer Institute, US National Institutes of Health (NIH), Department of Health and Human Services (DHHS). The Health Professionals Follow-up Study (HPFS) study is supported by NIH grants CA55075 and U01CA098233. The American Cancer Society (ACS) study is supported by U01 CA098710. The Alpha Tocopherol Beta-Carotene Cancer Prevention study (ATBC) Study is supported by the Intramural Research Program of the National Cancer Institute, NIH, and by US Public Health Service contracts N01-CN-45165, N01-RC-45035, and N01-RC-37004 from the National Cancer Institute, NIH, DHHS. A.S.K is supported by RO1CA112028. Centre de Recherche pour les Pathologies Prostatiques (CeRePP) is grateful to J.P.B. and Generali for their charitable donations to this project. The genome wide scans have been supported by the National Cancer Institute, National Institutes of Health, under contract N01-CO-12400. The content of this publication does not necessarily reflect the views or policies of the Department of Health and Human Services, nor does mention of trade names, commercial products, or organizations imply endorsement by the US Government.

References

1. Eeles RA, et al. Nat Genet. 2008; 40:316–21. [PubMed: 18264097]
2. Pal P, et al. Hum Genet. 2007; 122:251–259. [PubMed: 17593395]
3. Lai J, et al. Carcinogenesis. 2007; 28:1032–1039. [PubMed: 17151093]
4. Cramer SD, et al. J Natl Cancer Inst. 2003; 95:1044–1053. [PubMed: 12865450]
5. Prorok PC, et al. Control Clin Trials. 2000; 21:273S–309S. [PubMed: 11189684]
6. Thomas G, et al. 2008; 40:310–5.
7. Borgono CA, et al. Mol Cancer Res. 2004; 2:257–280. [PubMed: 15192120]
8. Woolf SH. N Engl J Med. 1995; 333:1401–1405. [PubMed: 7477122]
9. Draisma G, et al. J Natl Cancer Inst. 2003; 95:868–878. [PubMed: 12813170]
10. Etzioni R, et al. J Natl Cancer Inst. 2002; 94:981–990. [PubMed: 12096083]
11. Presti JC Jr, et al. J Urol. 2003; 169:125–129. [PubMed: 12478119]
12. Nam RK, et al. J Clin Oncol. 2003; 21:2312–2319. [PubMed: 12805332]

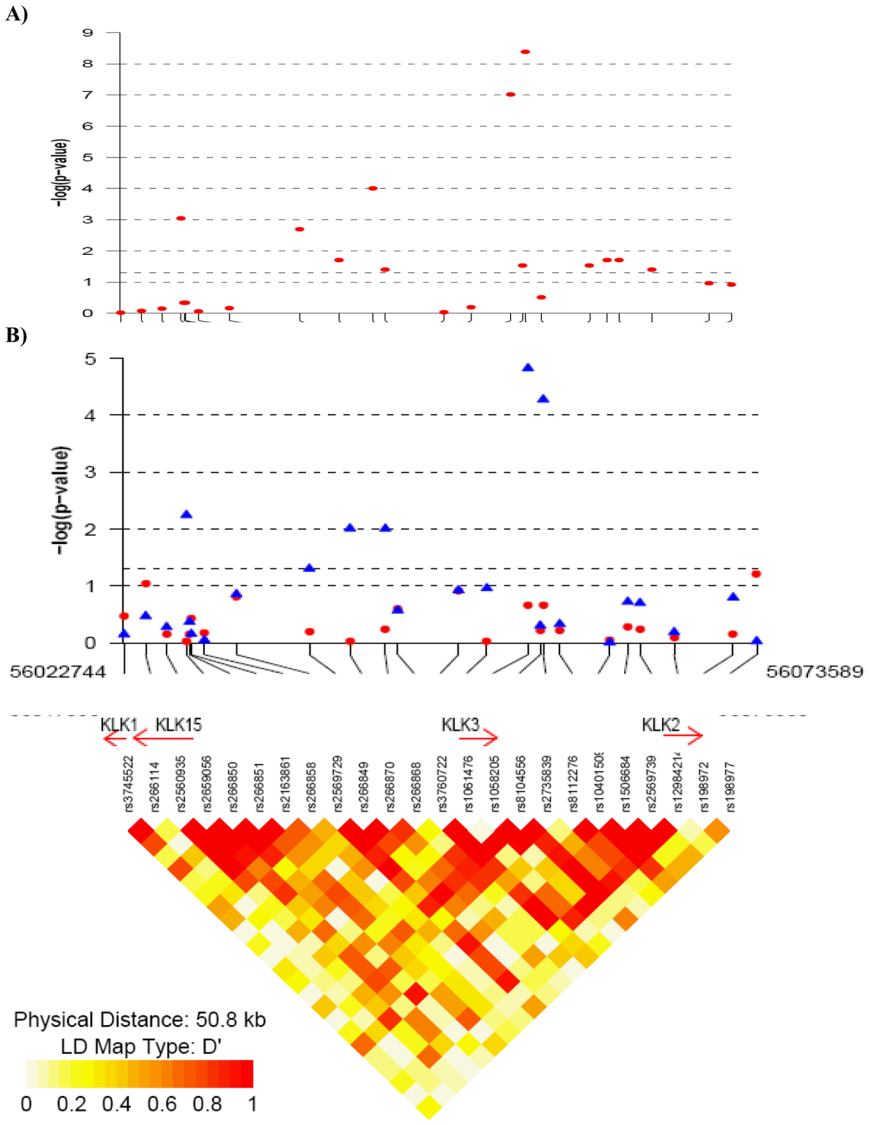

Figure 1. Associations of tagSNPs in the *KLK* region with PSA concentrations and prostate cancer risk in the PLCO Trial

A) P values for association of *KLK3* with serum PSA concentrations in the control group (2df test).

B) P values for association of *KLK3* with prostate cancer risk (2df test); the blue triangles represent p values for the association when we restricted the control group to men with low PSA (<0.5 ng/ml). The red circles represent p values for the association when we did not restrict the control group to men with low PSA.

Table 1

Associations of tagSNPs in the *KLK* region with prostate cancer risk and PSA concentrations in the PLCO Cancer Screening Trial and with prostate cancer in the combined CGEMS follow-up study.

dbSNP [a]	LOC [b]	Risk Allele (freq)	Gene	Cases vs. Controls in PLCO (n=1,172/1,157) [c]			Cases vs. Controls in CGEMS combined (n= 5,192/5,185) [d]			Adjusted mean PSA level for the PLCO controls (N=1,157) [e]			
				P-value (2df)	OR_{hetero} (95% CI)	OR_{homo} (95% CI)	P-value (2df)	OR_{hetero} (95% CI)	OR_{homo} (95% CI)	P-value (2df)	Homozygote	Heterozygote	Homozygote variant
rs3745522	56022744	T (0.26)	KLK1, KLK15	0.33	0.87 (0.73–1.05)	0.98 (0.72–1.34)				0.96	1.10 (0.79–1.53)	1.08 (0.77–1.51)	1.10 (0.76–1.59)
rs266114	56024482	C (0.34)	KLK1, KLK15	0.09	1.18 (0.98–1.41)	1.27 (0.97–1.66)				0.86	1.07 (0.77–1.50)	1.10 (0.79–1.54)	0.84 (0.40–1.73)
rs2560935	56026167	T (0.07)	KLK1, KLK15	0.70	0.92 (0.71–1.19)	1.34 (0.45–4.01)				0.72	1.09 (0.79–1.52)	1.09 (0.77–1.55)	0.84 (0.40–1.73)
rs2659056	56027755	G (0.26)	KLK1, KLK15	0.94	1.01 (0.85–1.20)	0.95 (0.68–1.33)				0.0009	1.00 (0.72–1.39)	1.23 (0.88–1.71)	1.07 (0.73–1.57)
rs266850	56028009	G (0.16)	KLK1, KLK15	0.68	1.00 (0.83–1.21)	1.26 (0.75–2.11)				0.46	1.06 (0.76–1.48)	1.12 (0.80–1.57)	1.23 (0.79–1.93)
rs266851	56028151	A (0.19)	KLK1, KLK15	0.37	0.92 (0.76–1.10)	1.24 (0.79–1.96)				0.47	1.12 (0.80–1.56)	1.07 (0.77–1.50)	0.96 (0.63–1.46)
rs2163861	56029204	T (0.50)	KLK1, KLK15	0.66	0.95 (0.78–1.17)	0.90 (0.71–1.13)				0.88	1.08 (0.77–1.51)	1.10 (0.79–1.53)	1.06 (0.75–1.50)
rs266858	56031774	C (0.23)	KLK3, KLK1, KLK15	0.15	1.13 (0.95–1.35)	1.38 (0.93–2.06)	0.61	1.02 (0.94–1.11)	1.09 (0.92–1.30)	0.69	1.07 (0.76–1.49)	1.11 (0.79–1.54)	1.15 (0.77–1.72)
rs2569729	56037633	A (0.18)	KLK3, KLK1, KLK15	0.63	0.96 (0.80–1.15)	1.20 (0.75–1.93)	0.37	0.94 (0.86–1.03)	0.95 (0.76–1.19)	0.002	1.13 (0.82–1.57)	0.95 (0.67–1.33)	0.85 (0.56–1.31)
rs266849	56040902	G (0.19)	KLK3, KLK15	0.91	1.01 (0.84–1.21)	0.91 (0.58–1.43)	0.49	0.96 (0.88–1.05)	1.08 (0.86–1.35)	0.02	1.14 (0.82–1.59)	1.02 (0.73–1.42)	0.85 (0.56–1.29)
rs266870	56043746	T (0.49)	KLK3, KLK15	0.56	1.09 (0.89–1.33)	1.14 (0.90–1.44)	0.13	0.92 (0.84–1.01)	0.90 (0.80–1.01)	0.0001	1.21 (0.87–1.69)	1.06 (0.76–1.48)	0.89 (0.63–1.25)
rs266868	56044749	A (0.31)	KLK3, KLK15	0.25	0.93 (0.78–1.11)	1.20 (0.89–1.60)	0.67	1.03 (0.95–1.13)	1.05 (0.91–1.20)	0.04	1.00 (0.72–1.40)	1.15 (0.82–1.60)	1.11 (0.77–1.60)
rs3760722	56049628	T (0.10)	KLK3, KLK2	0.12	1.18 (0.94–1.47)	2.01 (0.82–4.97)	0.75	1.04 (0.93–1.16)	1.06 (0.71–1.61)	0.95	1.10(0.79–1.53)	1.07 (0.76–1.51)	1.10 (0.53–2.27)
rs1061476	56051899	A (0.49)	KLK3, KLK2	0.94	0.97 (0.79–1.18)	0.97 (0.76–1.22)	0.73	1.03 (0.94–1.14)	1.04 (0.93–1.17)	0.64	1.10 (0.79–1.54)	1.10 (0.79–1.54)	1.04 (0.74–1.47)
rs1058205	56055210	C (0.18)	KLK3, KLK2	0.21	0.91 (0.75–1.09)	0.68 (0.41–1.12)	0.07	0.95 (0.87–1.04)	0.77 (0.60–0.99)	9.7×10^{-8}	1.13 (0.82–1.56)	1.04 (0.74–1.45)	0.53 (0.35–1.25)
rs8104556	56056192	T (0.17)	KLK3, KLK2	0.61	1.01 (0.83–1.21)	1.27 (0.79–2.04)				0.03	1.13 (0.81–1.57)	1.03 (0.74–1.44)	0.81 (0.53–1.25)
rs2735839	56056435	A (0.15)	KLK3, KLK2	0.21	0.93 (0.76–1.12)	0.60 (0.32–1.12)	0.27	0.95 (0.87–1.04)	0.82 (0.60–1.12)	4.1×10^{-9}	1.10 (0.79–1.51)	0.94 (0.67–1.31)	0.44 (0.28–1.44)
rs8112276	56057764	A (0.49)	KLK3, KLK2	0.61	0.93 (0.76–1.13)	1.02 (0.80–1.29)	0.23	0.92 (0.84–1.01)	0.94 (0.84–1.05)	0.31	1.18 (0.84–1.66)	1.07 (0.77–1.50)	1.08 (0.77–1.52)
rs10401509	56061787	A (0.005)	KLK3, KLK2	0.90	0.80 (0.30–2.12)	NA	0.14	1.16 (0.86–1.57)	NA	0.03	1.11 (0.80–1.54)	0.64 (0.36–1.14)	NA
rs1506684	56063231	T (0.46)	KLK3, KLK2	0.52	0.93 (0.76–1.13)	0.87 (0.69–1.11)	0.09	0.91 (0.83–0.99)	0.91 (0.81–1.02)	0.02	1.23 (0.88–1.72)	1.08 (0.77–1.50)	1.01 (0.72–1.41)
rs2569739	56064207	C (0.47)	KLK3, KLK2	0.56	0.93 (0.76–1.13)	0.88 (0.70–1.12)				0.02	1.21 (0.87–1.70)	1.06 (0.76–1.49)	1.00 (0.71–1.41)
rs12984214	56066957	G (0.35)	KLK2	0.81	1.03 (0.86–1.23)	0.94 (0.73–1.22)				0.04	1.02 (0.73–1.42)	1.14 (0.82–1.60)	1.20 (0.84–1.70)
rs198972	56071705	T (0.31)	KLK2	0.70	1.08 (0.90–1.28)	1.08 (0.80–1.44)				0.11	1.14 (0.82–1.59)	1.03 (0.74–1.44)	1.00 (0.69–1.44)
rs198977	56073589	T (0.24)	KLK2	0.06	1.20 (1.01–1.43)	1.33 (0.94–1.90)				0.12	1.13 (0.81–1.58)	1.02 (0.73–1.43)	0.99 (0.67–1.47)

[a] SNP identifier based on NCBI dbSNP; SNPs are included in the region of the *KLK* genes including 2 kb of transcription start sites and 10 kb of last exon.

[b] Chromosomal location based on NCBI Human Genome Build 35 coordinates.

[c] The result of 2 df test based on logistic regression in the PLCO study - adjusted for age in five year intervals, study center, three eigenvectors to control population stratification in an incident density sampling strategy.

[d] The result of 2 df test based on logistic regression of the combined genotypes generated in the initial PLCO study and the four follow-up studies (American Cancer Society Cancer Prevention Study II, 1,790/1,797; the Health Professionals Follow-up Study, 619/620; the CeRePP French Prostate Case-Control Study, 671/671; and the Alpha-Tocopherol, Beta-Carotene Cancer Prevention Study, 940/940)- adjusted for age in ten-year intervals, study/study center, and four eigenvectors to control population stratification.

[e] The result of 2 df test based on generalized linear regression model in the controls of the PLCO study, after log transformation of PSA – adjusted for age in five year intervals, study center, and three eigenvectors to control population stratification. PLCO Trial PSA concentration-closest in time prior to selection of cases and controls.