

Published in final edited form as:

*Sci Transl Med.* 2010 December 15; 2(62): 62ra92. doi:10.1126/scitranslmed.3001513.

## Genetic correction of PSA values using sequence variants associated with PSA levels

Julius Gudmundsson<sup>1,\*</sup>, Soren Besenbacher<sup>1,\*</sup>, Patrick Sulem<sup>1,\*</sup>, Daniel F. Gudbjartsson<sup>1</sup>, Isleifur Olafsson<sup>2</sup>, Sturla Arinbjarnarson<sup>3</sup>, Bjarni A. Agnarsson<sup>2,4</sup>, Kristrun R. Benediktsdottir<sup>2,4</sup>, Helgi J. Isaksson<sup>2</sup>, Jelena P. Kostic<sup>1</sup>, Sigurjon A. Gudjonsson<sup>1</sup>, Simon N. Stacey<sup>1</sup>, Arnaldur Gylfason<sup>1</sup>, Asgeir Sigurdsson<sup>1</sup>, Hilma Holm<sup>1</sup>, Unnur S. Bjornsdottir<sup>4</sup>, Gudmundur I. Eyjolfsson<sup>5</sup>, Sebastian Navarrete<sup>6</sup>, Fernando Fuertes<sup>6</sup>, Maria D. Garcia-Prats<sup>7</sup>, Eduardo Polo<sup>8</sup>, Ionel A. Checherita<sup>9</sup>, Mariana Jinga<sup>9</sup>, Paula Badea<sup>10</sup>, Katja K. Aben<sup>11,12</sup>, Jack A. Schalken<sup>13</sup>, Inge M. van Oort<sup>13</sup>, Fred C. Sweep<sup>14</sup>, Brian T. Helfand<sup>15</sup>, Michael Davis<sup>16</sup>, Jenny L. Donovan<sup>16</sup>, Freddie C. Hamdy<sup>17</sup>, Kristleifur Kristjansson<sup>1</sup>, Jeffrey R. Gulcher<sup>1</sup>, Gisli Masson<sup>1</sup>, Augustine Kong<sup>1</sup>, William J. Catalona<sup>15</sup>, Jose I. Mayordomo<sup>18</sup>, Gudmundur Geirsson<sup>2</sup>, Gudmundur V. Einarsson<sup>2</sup>, Rosa B. Barkardottir<sup>2,4</sup>, Eirikur Jonsson<sup>2</sup>, Viorel Jinga<sup>9</sup>, Dana Mates<sup>10</sup>, Lambertus A. Kiemeny<sup>11,12,13</sup>, David E. Neal<sup>19</sup>, Unnur Thorsteinsdottir<sup>1,4</sup>, Thorunn Rafnar<sup>1</sup>, and Kari Stefansson<sup>1,4</sup>

<sup>1</sup>deCODE genetics, IS-101 Reykjavik, Iceland. <sup>2</sup>Landspítali-University Hospital, IS-101 Reykjavik, Iceland. <sup>3</sup>Glaesibaer, Clinical Research Laboratory, IS-108 Reykjavik, Iceland. <sup>4</sup>Faculty of Medicine, University of Iceland, IS-101 Reykjavik, Iceland. <sup>5</sup>The Laboratory in Mjodd, Reykjavik, Iceland. <sup>6</sup>Division of Radiation Oncology, University Hospital, 50009, Zaragoza, Spain. <sup>7</sup>Division of Pathology, San Jorge General Hospital, 22004 Huesca, Spain <sup>8</sup>Royo Villanova Hospital, 500014 Zaragoza, Spain. <sup>9</sup>University of Medicine and Pharmacy Carol Davila, Bucharest, Romania. <sup>10</sup>National Institute of Public Health, Bucharest, Romania. <sup>11</sup>Comprehensive Cancer Center East, 6501 BG Nijmegen, The Netherlands. <sup>12</sup>Department of Epidemiology, Biostatistics and HTA, Radboud University Nijmegen Medical Centre, 6500 HB Nijmegen, The Netherlands. <sup>13</sup>Department of Urology, Radboud University Nijmegen Medical Centre, 6500 HB Nijmegen, The Netherlands. <sup>14</sup>Department of Laboratory Medicine, Radboud University Nijmegen Medical Centre, 6500 HB Nijmegen, The Netherlands. <sup>15</sup>Department of Urology, Northwestern University Feinberg School of Medicine, Chicago, IL 60611, USA. <sup>16</sup>School of Social and Community Medicine, University of Bristol, Bristol BS8 1TH, UK. <sup>17</sup>Nuffield Department of Surgical Science, University of Oxford, John Radcliffe Hospital, Oxford OX3 9DU, UK <sup>18</sup>Division of Medical Oncology, University Hospital, 50009 Zaragoza, Spain. <sup>19</sup>Oncology Centre, Addenbrooke's Hospital, University of Cambridge, Cambridge CB2 0QQ, UK.

### Abstract

Correspondence and request for materials should be addressed to: K. S., (kstefans@decode.is) or J. G., (julius@decode.is) .

\*These authors contributed equally to this work.

Author contribution: The study was designed and results were interpreted by J.G., S.B., P.S., D.F.G., A.K., U.T., T.R., and K.S. Statistical analysis was carried out by S.B., P.S., D.F.G., J.G., and A.K. Subject recruitment, biological material collection and handling along with genotyping was supervised and carried out by J.G., I.O., S.A., B.A., K.R.B., H.H., G.O., S.N.S., A.S., T.W., T.T., J.P.B., K.M.Mc., K.M.B., B.S., J.Godino, S.N., F.F., L.M., E.P., K.K.A., I.M.vO., B.K.S., B.T.H., D.K., C.Z., K.K., J.R.G., G.V.E., E.J., W.J.C., J.I.M., L.A.K., D.E.N., J.S., R.B.B., U.T. and T.R. Authors J.G., S.B., P.S., D.F.G., T.R. and K.S. drafted the manuscript. All authors contributed to the final version of the paper. Principal investigators and corresponding authors for the respective replication study populations are: The Netherlands, L.A.K.; Spain, J.I.M.; Chicago, W.J.C.; Romania, D.M.; the UK, F.C.H / D.E.N / J.L.D.

**Competing interests:** The authors from deCODE are employees at deCODE Genetics Inc. K.S. is CEO and U.T. is VP of Research at deCODE Genetics, Inc. DeCODE Genetics Inc. has a patent pending for the genetic correction of individual PSA levels.

Measuring serum levels of the prostate specific antigen (PSA) is the most common screening method for prostate cancer. However, PSA levels are affected by a number of factors apart from neoplasia. Notably, around 40% of the variability of PSA levels in the general population is accounted for by inherited factors, suggesting that it may be possible to improve both sensitivity and specificity by adjusting test results for genetic effects. In order to search for sequence variants that associate with PSA levels, we performed a genome-wide association study and follow-up analysis using PSA information from 15,757 Icelandic and 454 British men not diagnosed with prostate cancer. Overall, we detected a genome-wide significant association between PSA levels and SNPs at six loci: 5p15.33 (rs2736098), 10q11 (rs10993994), 10q26 (rs10788160), 12q24 (rs11067228), 17q12 (rs4430796), and 19q13.33 (rs17632542 (*KLK3*: I179T), each with  $P_{\text{combined}} < 3 \times 10^{-10}$ . Among 3,834 men who underwent a biopsy of the prostate, the 10q26, 12q24, and 19q13.33 alleles that associate with high PSA levels are associated with higher probability of a negative biopsy (OR between 1.15 and 1.27). Assessment of association between the 6 loci and prostate cancer risk in 5,325 cases and 41,417 controls from Iceland, the Netherlands, Spain, Romania, and the US showed that the SNPs at 10q26 and 12q24 were exclusively associated with PSA levels, whereas the other 4 loci also were associated with prostate cancer risk. We propose that a personalized PSA cutoff value, based on genotype, should be used when deciding to perform a prostate biopsy.

## INTRODUCTION

In Western countries, prostate cancer is the most frequent cancer in men as well as one of the leading causes of cancer related death in men. Early diagnosis and treatment are key factors in determining the survival and prognosis of prostate cancer patients, prompting intensive searches for biomarkers for screening. Currently, measuring serum levels of prostate specific antigen (PSA) is the only accepted screening method for the disease. PSA levels in the range of 2.5 ng/mL to 4 ng/mL are commonly used as a threshold when deciding whether to perform a biopsy of the prostate. However, despite the widespread use of the PSA screening test, it is limited both in specificity and sensitivity and substantial controversy exists over its benefits for patients (1). This is mainly due to the fact that PSA is not a specific marker of prostate cancer as its levels in serum increase in benign prostatic hyperplasia and are affected by many other factors such as medications, urological manipulations and inflammation. Notably, it has been shown that 33% of men with PSA levels above 10 ng/ml had no evidence of prostate cancer at biopsy(2) and in a recent study the fraction of men with PSA levels above 10 ng/ml and no evidence of prostate cancer at biopsy was 47%(3). Furthermore, not all individuals with prostate cancer have raised levels of PSA. By performing a biopsy on a set of individuals not selected for PSA levels, Thompson and colleagues showed that 63% of prostate cancer patients (even prostate cancer with Gleason score 7) have PSA levels lower than 3 ng/ml (4). Hence, current data suggest that there is no PSA threshold below which prostate cancer can be excluded.

The utility of PSA in prostate cancer screening is currently being evaluated in two large clinical trials, the prostate arm of the Prostate, Lung, Colon and Ovary Cancer Screening Trial (PLCO) in the USA and the European Randomized Study of Screening for Prostate Cancer (ERSPC). These trials are expected to establish whether screening has an effect on prostate cancer mortality. The first results from these studies which were published last year showed either no or a small reduction in prostate cancer-specific mortality in the PSA-screened groups.(5, 6). Since then, it has been suggested that the initial reports from these trials may have underestimated the beneficial effect of PSA screening as recent publications report that PSA testing leads to a relative reduction of prostate cancer metastasis of 53% and to a relative reduction of prostate cancer mortality close to 50%. However, over-diagnosis is still reported to be substantial(7-9)

One approach to increase the specificity and sensitivity of the PSA test is to work out a model that defines the “normal” PSA value for a given man. Such a model would have to take into account age, as the prevalence of both benign and malignant conditions that can cause serum PSA elevation increases with age. Sequence variants are also important when determining personal “normal” PSA levels, as it has been shown that expression of a large percentage of genes is under genetic control of both *cis* and *trans* components (10). Importantly, it has been shown that between 40% and 45% of the variability in PSA levels among men in the general population can be accounted for by inherited factors (11, 12). However, because these factors are largely unknown, it has not been possible to adjust for them when interpreting the PSA results for a given individual. Recently, a variant (rs2735839-G) was discovered near the *KLK3* gene (which encodes PSA) that was shown to predispose to prostate cancer (odds ratio OR= 1.2) (13). However, it was subsequently shown that this variant associates with increased levels of PSA in men not diagnosed with prostate cancer (14, 15). Hence, the reported risk of prostate cancer conferred by the variant may be mediated principally through its effect on PSA levels. By correcting for this variant, as well as additional sequence variants that affect PSA levels, both the sensitivity and the specificity of the PSA screening test could be improved. This prompted us to perform a genome-wide association study (GWAS) of sequence variants affecting population variation in PSA levels. We then considered the effects of the PSA variants on subsequent prostate cancer diagnoses.

## RESULTS

### Sequence variants associated with PSA levels

We performed a GWAS on serum PSA levels (adjusted for age and laboratory center) in Icelandic men not diagnosed with prostate cancer according to data from the nationwide Icelandic Cancer Registry (ICR). In total, we had access to PSA measurements from 4,620 individuals genotyped on Illumina chips, containing either the 317K or the 370K HumanHap SNP panel. The analysis was augmented with data from 9,218 Icelanders with PSA measurements whose genetic information could be partially inferred from genotyped relatives (*in-silico* genotyping), using a previously described method (16). With respect to statistical power, this augmentation is equivalent to an additional 2,918 individuals on average (for details about the populations see Table S1). After quality control, 304,070 SNPs were available for the GWAS (for quantile-quantile plot of the results see Fig. S1). Since the mean of the  $\chi^2$  values was below 1 ( $\chi^2 = 0.91$ ) we did not apply any genomic control correction.

We selected all association signals with  $P < 1 \times 10^{-5}$  for further analysis. This represented 12 SNPs at 6 different loci, of which four loci reached genome-wide significance after accounting for the number of tests performed ( $P < 1.64 \times 10^{-7} = 0.05/304,070$ ) (Table S2a). The genome-wide significant association signals were in or near genes at the following loci: *KLK3* on 19q13.33; *HNF1B* on 17q12; *FGFR2* on 10q26.12; and *TBX3* on 12q24.21. The two suggestive association signals were at 10q11.23 near the *MSMB* gene and at 5p15.33 near the *TERT* gene (Table S2a).

To further investigate each of the six loci, we imputed genotypes based on data for 2.5 million SNPs from the Utah CEPH (Utah residents with ancestry from northern and western Europe) (CEU) HapMap individuals for all SNPs present within a window of 500 Kb centered on the most significant SNP. Based on this analysis, we identified three additional SNPs; rs2736098-A at 5p15.33, rs4430796-A at 17q12 and rs17632542-T at 19q13.33, that had a greater effect on PSA levels than any SNP present on the 317K chip (Table S2b).

In an attempt to follow-up the observed associations with PSA levels in the Icelandic discovery group, we genotyped the most significant SNP at each of the six loci in an additional 1,919 Icelandic men with PSA level measurements and not diagnosed with prostate cancer, and in 454 men from the ProtecT trial in the UK with PSA levels below 3 ng/ml and not diagnosed with prostate cancer. After combining significance levels from Iceland and the UK, at least one SNP at each locus reached genome-wide significance (Table 1).

For the strongest variant at each locus, the allele frequency was comparable in the Icelandic and UK populations with frequencies ranging from 24% to 93% (Table 1) and their observed effect on the PSA level ranges from 7% to 39% per allele in the Icelandic samples and from 5% to 102% per allele in the UK samples (see Table 1 and Table S3 for genotype effect of the variants). Variants in or near genes at three of these PSA loci, *KLK3* on 19q13.33, *HNFB* on 17q12, and *MSMB* 10q11, have previously been associated with serum PSA levels in men not diagnosed with prostate cancer (13-15, 17-20).

The strongest overall association effect observed in the present study is with two SNPs, rs2735839 and rs17632542, located near or in the PSA coding gene *KLK3* (Table 1). Of these, rs2735839-G (and highly correlated markers) has been reported previously to associate with PSA levels (13-15, 21). The two SNPs are moderately correlated with each other ( $D'=1$  and  $r^2=0.48$  in UK;  $r^2=0.56$  in Iceland;  $r^2=0.56$  in HapMap CEU phase 3). When we adjusted the results for each SNP, using the other SNP as a covariate and only including individuals genotyped for both markers, results for rs17632542 remain significant after adjusting for rs2735839 ( $P_{\text{combined}}=5.51\times 10^{-8}$ ) whereas rs2735839 was marginally significant after adjusting for rs17632542 ( $P_{\text{combined}}=0.043$ ). This suggests that the signal from rs2735839 is subsumed by rs17632542. The SNP rs17632542 is a missense mutation (an amino acid change denoted as I179T) in *KLK3*. This amino acid alteration is defined as either neutral or deleterious by online protein structure algorithms (see Table S4). A deleterious mutation could conceivably destabilize the protein, affecting circulating PSA levels. Alternatively, the mutation may affect the antigenicity of the protein and thereby influence its detectability in PSA tests. The alleles at the 10q11 (*MSMB*) and 17q12 (*HNFB*) PSA loci are rs10993994-T and rs4430796-A or the same as those previously reported to associate with PSA levels (17-19) as well as with prostate cancer risk (17, 22).

At the new PSA locus on 10q26, two variants, rs10788160-A and rs12413088-T, have genome-wide significance and have similar effects on PSA levels. The two variants are located within a region of linkage disequilibrium (LD) not known to contain any genes 324 and 305 Kb centromeric to the start of the *FGFR2* gene, respectively. The two variants are highly correlated ( $r^2=0.85$  in Iceland and  $r^2=0.83$  in the UK) and neither remains significant after adjusting for the other. Given that the effects of the two variants cannot be distinguished from each other, we elected to focus on rs10788160-A in subsequent investigations. The PSA variant, rs10788160, is in very low linkage disequilibrium with the variant rs1219648 (23) that confers risk of breast cancer ( $D'=0.15$ ,  $r^2=0.01$  between rs1219648 and rs10788160 in Iceland).

The most significant variant on 12q24, the second new PSA locus, is rs11067228-A. This SNP is located in an LD-block that contains the gene *TBX3*, mutations in which have been found to cause the ulnar-mammary syndrome (OMIM #181450) but have not previously been shown to affect PSA levels.

At the third new PSA locus, on 5p15 near the *TERT* gene, two sequence variants, rs401681-C and rs2736098-A, have comparable effects on PSA levels. They are moderately correlated ( $D'=0.93$  and  $r^2=0.39$  between rs401681 and rs2736098 according to HapMap CEU Phase

2), and because the effects of the variants cannot be distinguished from each other, we elected to focus on rs2736098-A in subsequent analyses.

We estimated the fraction of the total variance in the levels of PSA explained by combining the effect from the best marker at each of the six loci (rs2736098, rs10993994, rs10788160, rs11067228, rs4430796 and rs17632542). The fraction accounted for is estimated to be 4.2% in Iceland and 11.8% in the UK (see Table S5). In both populations, the missense mutation in the *KLK3* gene, rs17632542, accounts for half of the fraction of variance explained.

### PSA variants and predisposition to prostate cancer

Variants at four of the six loci discussed above (*KLK3*, *TERT*, *MSMB* and *HNF1B*) have previously been reported to be associated with risk of prostate cancer, although at different degrees of significance (13, 17, 21, 22, 24) and some even with conflicting evidence (14). Due to the potential confounding effects of PSA levels and prostate cancer, we examined whether the PSA SNPs identified in this study also are associated with prostate cancer. Based on a combined analysis of over 5,325 prostate cancer cases and 41,417 controls from Iceland, the Netherlands, Spain, Romania and the US, we replicated the association of the variants at the four loci with prostate cancer, each with a similar effect as described before (odds ratios ranging from 1.10 to 1.21; see Table S6). Interestingly, in our data the missense variant in *KLK3*, rs17632542, shows a stronger association with prostate cancer than the strongest previously reported variant at this locus, rs2735839 (OR=1.39 and 1.19 for rs17632542-T and rs2735839-G, respectively; see Table S6). In contrast, we found that the variants at two of the three new PSA loci (*FGFR2* and *TBX3*) do not associate significantly with prostate cancer ( $P_{\text{combined}} = 0.27$  and  $0.54$ ;  $OR_{\text{combined}} = 0.97$  and  $1.01$ , for rs10788160-A and rs11067228-A, respectively). The effects of these six variants on both PSA levels and prostate cancer risk in Icelandic men is shown graphically in Fig. S2.

We next examined if any of the six loci associated with PSA levels have an effect on age at diagnosis or aggressiveness of prostate cancer among patients in the 6 study groups from Iceland, the Netherlands, Spain, Romania, the US and the UK. Only the missense mutation in *KLK3*, rs17632542, is significantly associated with age at diagnosis; for each allele of rs17632542-T that associates with higher PSA levels, the age at diagnosis was estimated to decrease by nearly 9 months (0.71 year decrease,  $P = 0.016$ ; see Table S7). When performing a case-only analysis, we observed that for the missense mutation in *KLK3*, rs17632542-T, the allele conferring risk of prostate cancer is significantly less frequent ( $OR=0.78$ ,  $P = 0.0099$ ; see Table S8) among cases with more aggressive prostate cancer (Gleason score  $> 6$ , and/or T3 (tumor extends through prostate capsule) or higher, and/or node positive, and/or with metastatic disease) than among those cases with less aggressive prostate cancer (Gleason score  $< 7$ , and T2 or lower). This is in agreement with findings previously reported for the correlated variant at this locus, rs2735839 (25, 26). For none of the five variants was there a significant effect on the aggressiveness of prostate cancer.

There has been controversy in the literature about the interpretation of the association of the previously reported *KLK3* variant (rs2735839) to prostate cancer, namely, whether the observed association is due to a predisposition to the disease or mediated through the strong effect of the *KLK3* locus on PSA levels and therefore driven by the increasing frequency of PSA testing (14, 15). To examine this, we stratified our Icelandic study group into cases diagnosed before 1992, a time when the majority of patients were diagnosed without undergoing PSA testing, and cases diagnosed from 1992 to 2008, a period in which PSA testing became more frequent. We used *in-silico* genotyping based on familial imputation to augment the effective sample size of the group of cases, and we used 34,124 Icelanders not known to have prostate cancer as controls. Our results for rs2735839-G show that the association effect observed for the total case study group ( $OR = 1.15$  (95% CI 1.04-1.27),  $P$



= 0.007) is confined to the group of cases diagnosed in 1992 or later (OR = 1.17 (95% CI. 1.06-1.29),  $P = 0.002$ ), whereas cases diagnosed before 1992 have no increased risk (OR = 0.97 (95% CI. 0.83-1.13),  $P = 0.7$ ; see Table S9). These results support the notion that the prostate cancer risk reported for the *KLK3* locus is driven by the increasing frequency of PSA testing and subsequent biopsies over the last few decades. In contrast, the results for the other five PSA loci are not substantially different for the two case subgroups (Table S9).

### Effect of prostate cancer risk variants on PSA levels

We assessed the effect on PSA serum levels of the 47 sequence variants (23) conferring risk of prostate cancer reported to date (see Table S10 and Fig. S2). Some loci have more than one reported SNP. According to our results, there is a clear tendency for the allele associated with prostate cancer risk also to be associated with high levels of PSA (see Table S10). This is comparable to results previously reported by Wiklund et al. (15). For the vast majority of the loci ( $N=41$ ), their effect on PSA levels is weak (well below 0.1 standard units on the scale of the distribution of the PSA measurements (see Methods in Supplementary Material) and likely reflects undiagnosed prostate cancer cases in the PSA study group [also suggested by Wiklund et al. (15)]. Exceptions are the variants at the *KLK3* (rs2735839 and rs17632542), *HNF1B* (rs4430769), *MSMB* (rs10993994) and *TERT* loci (rs2736098), which are the loci of genome-wide significance in our PSA genome-wide association study. Variants at two other loci 11q13 (rs11228565) and 8q24 (rs16901979) also have greater effects on PSA levels but did not reach genome-wide significance (see Table S10). These six loci can roughly be divided into two groups: those with a moderate effect on PSA levels compared to their effect on prostate cancer risk (8q24, 11q13, 10q11 and 17q12) and those comprising variants that have a relatively strong PSA effect compared to their effect on prostate cancer risk (i.e. variants at *KLK3* on 19q13.33, and *TERT* on 5p15; see Fig. S2).

### Sequence variants and benign prostatic hyperplasia

Benign prostatic hyperplasia (BPH), a common non-malignant overgrowth of prostate tissue, can affect serum PSA levels. To determine if any of the PSA variants discussed above are associated with BPH, we compared a set of 2,312 Icelandic men with symptomatic BPH with 33,779 healthy Icelandic men. Men with symptomatic BPH were defined as individuals either diagnosed after undergoing transurethral resection of the prostate (TURP) or men over the age of 50 repeatedly using drugs in the G04C group of the ATC classification (e.g. Tamsulosin, Finasteride and Dutasteride) for treating BPH between the years 2003 and 2009. Except for rs2736098-T on 5p15 that showed a nominally significant association ( $P = 0.048$ , OR=1.08; Table S11), no association was observed between BPH and any of the remaining five PSA variants. Hence, BPH is unlikely to account for a significant fraction of the observed association between PSA levels and the variants discussed here.

### PSA sequence variants and prostate biopsies

When screening for prostate cancer, a PSA level above a certain cutoff value is considered an indicator that a needle biopsy is required. We wanted to assess if the variants that associate with increased PSA levels also make men more prone to undergo a prostate tissue biopsy. In our study group of 2,300 Icelandic men who underwent a prostate biopsy between 1998 and 2008, we observed a higher frequency of the allele increasing PSA levels in those undergoing a biopsy compared with population controls for all six variants (1.04 OR 1.46; all SNPs have  $P < 0.05$  except rs11067228 on 12q24 which has  $P = 0.25$ , see Table S12). Among the 2,300 individuals who had undergone a biopsy, cancer was diagnosed in ~50% (a positive biopsy). When restricting the analysis to individuals with a biopsy but no detectable prostate cancer (negative biopsy) and comparing them to population controls, similar or even stronger results were observed (1.03 OR 1.82; all SNPs have  $P < 0.05$

except rs10993994 near *MSMB* which has  $P = 0.48$ , see Table S13). These results suggest that unnecessary biopsies are more likely to be performed on individuals with elevated PSA levels due to inherited factors.

From the UK study group, we had access to a group of approximately 1,400 men who had undergone a biopsy. Of those, about one third were diagnosed with prostate cancer. Using the Icelandic and the UK study groups of men who had been biopsied, we compared the frequency of the PSA variants in positive and negative biopsies. Of the six loci, we found that for the three PSA variants not primarily associated with prostate cancer risk (*KLK3*, *FGFR2* and *TBX3*), the alleles associated with increased PSA levels are significantly less frequent among men with a positive biopsy than in men with a negative biopsy (rs10788160-A near *FGFR2* has  $OR_{combined} = 0.79$  and  $P_{combined} = 5.4 \times 10^{-6}$ , rs11067228-A near *TBX3* has  $OR_{combined} = 0.87$  and  $P_{combined} = 0.0034$ , rs17632542-T in *KLK3* has  $OR_{combined} = 0.77$  and  $P_{combined} = 0.013$ ; see Table 2 and Table S14 for full details). The results for these three variants demonstrate that the alleles associated with increased PSA levels increase the probability that a normal prostate is biopsied.

### Combined effect of PSA sequence variants

To summarize the overall effect on PSA levels, we combined the effect of the PSA variants independently for the Icelandic and UK study populations, assuming a multiplicative model. We chose to include in the analysis only the four sequence variants located near *TERT*, *FGFR2*, *TBX3* and *KLK3* (rs2736098, rs10788160, rs11067228, and rs17632542, respectively) that are primarily associated with PSA levels. The variants at the *MSMB* and *HNF1B* loci were not included, as we consider them not to be primarily associated with PSA levels. We realize that disentangling the effect of prostate cancer on PSA levels and the impact of PSA levels on prostate cancer diagnosis is challenging. However, we based our selection on the fact that out of the six PSA variants reported here, the variants at the *MSMB* and *HNF1B* loci show the most significant association with prostate cancer (each SNP with  $P < 4 \times 10^{-13}$ ; Table S6).

Results from Iceland for the combined effect of the PSA variants show that for the top 5% of the genetic PSA level distribution, the measured PSA levels are estimated to be 23% to 47% higher than the population average. Similarly, for the bottom 5% of genetic PSA level distribution, the measured PSA levels are estimated to be 30% to 56% lower than the population average. In the UK study population, the estimated relative effect on PSA levels is even greater; the range of increase is 40% to 92% for the top 5% of the distribution with the greatest genotypic effect compared to the population average, whereas for the bottom 5% of the distribution, the range of the decrease is 53% to 80% compared with the population average.

To demonstrate how the genetic effect of the four PSA sequence variants influences individual PSA levels, we calculated a personalized PSA cutoff value corresponding to the commonly used cutoff of 4 ng/ml. This was done by multiplying the value of 4 ng/ml with the estimated relative genetic effect for the PSA SNPs. For individuals with the highest (top 5% of the distribution) genotypic effect, the personalized PSA cutoff value increased from 4 ng/ml to cutoff values between 4.9 and 5.9 ng/ml based on the estimates from Iceland, and to cutoff values between 5.6 and 7.7 ng/ml based on the UK estimates. For the bottom 5% of the relative genetic effect distribution, the personalized PSA cutoff values move from 4 ng/ml to cutoff values between 1.7 and 2.8 ng/ml according to the Icelandic estimates, and to cutoff values between 0.8 and 1.9 ng/ml according to the UK estimates (see Fig. 1). These results demonstrate that for a substantial fraction of men undergoing PSA-based prostate cancer screening, the personalized PSA cutoff value is shifted following correction for the effect of the PSA sequence variants.

We then assessed what fraction of 12,779 PSA measurements from 4,569 Icelandic men would be reclassified, with respect to certain PSA cut-off values, after correcting them for the four PSA sequence variants located near *TERT*, *FGFR2*, *TBX3* and *KLK3* (rs2736098, rs10788160, rs11067228, and rs17632542, respectively). For a PSA cut-off value of 4ng/ml, 6.0% of the men had at least one PSA measurement reclassified; 3.0% moved from below to above the cut-off value and 3.0% moved in the opposite direction. The results for a cut-off value of 3ng/ml were similar, 6.9% of the men had at least one PSA measurement reclassified; 3.1% moved from below to above the cut-off value and 3.8% moved in the opposite direction (Table S15). If applied clinically, these men would be reclassified with respect to whether or not they should undergo a biopsy.

### Discriminatory power of biopsy outcome models

We calculated the area under the receiver-operating-characteristic curve (AUC) to assess the discriminatory power of four models on the outcome of performing a biopsy of the prostate. The four models included the following data: Model 1, PSA levels; model 2, the combined prostate cancer risk estimates of 23 established sequence variants; model 3, genetic correction of PSA values based on the sequence variants at the four PSA loci (5p15, 10q26, 12q24 and 19q33.3); model 4, PSA levels corrected for sequence variants and the combined risk estimates of the 23 prostate cancer risk variants. In the analyses of the models, we used 415 Icelandic and 1,291 British men with information on biopsy outcome (i.e. biopsy positive or biopsy negative) and PSA levels, as well as genotypes for the 23 established prostate cancer variants and the PSA variants reported above. The model with genetic correction of PSA levels (model 3) has an AUC of 70.9% and 58.5% for the Icelandic and UK men, respectively (Fig. S3). When compared to model 1, which has an AUC of 70.4% and 57.1% for the Icelandic and UK men, respectively, the inclusion of PSA levels corrected for sequence variants (model 3) increases the discriminatory power by 0.5 and 1.4 percentage points in the Icelandic and UK groups, respectively. However, of the four models assessed, model 4 has the greatest discriminatory power, with an AUC of 73.2% and 63.6% for the Icelandic and UK men, respectively. Compared to model 1, the increased AUC of model 4 is 2.8 and 6.5 percentage points for Icelandic and UK men, respectively. Hence, the most gain in discriminatory power is achieved by including both the 23 prostate cancer risk variants and the genetic correction of PSA levels. However, in order to better assess the effect of the PSA and prostate cancer risk variants on PSA-based biopsies, this type of modeling would have to be done in a population where biopsies are performed systematically irrespective of individual PSA levels, similar to the approach taken in the Prostate Cancer Prevention Trial (4).

## DISCUSSION

In this study, we identified 6 loci that associate with PSA levels with genome-wide significance. Variants at three of these loci had previously been shown to associate with PSA levels, whereas three of the loci---at 10q26, 5p15 and 12q24---are new. Unlike the variants previously reported to associate with PSA levels, two of the new loci, 12q24 and 10q26, do not associate with prostate cancer risk and the third locus, at 5p15, has only a moderate effect on prostate cancer. Furthermore, we have shown that two of these variants (rs10788160-A on 10q26 and rs11067228-A on 12q24), together with the *KLK3* variant, are associated with a greater probability of having a normal prostate biopsied. Hence, these new markers primarily predict the outcome of PSA-based prostate cancer screening---that is, the decision to perform a biopsy and the outcome of the biopsy--rather than predisposition to prostate cancer.

In our study, we showed that a missense mutation, rs17632542-T, in the *KLK3* gene on 19q33.33 is associated with higher PSA levels. This variant has a stronger effect on PSA



levels than the variant rs2735839 previously reported at this locus. The *KLK3* variant was also found to predispose to prostate cancer but the effect was confined to the group of cases diagnosed after the introduction of the PSA test. Furthermore, the association with prostate cancer at the *KLK3* locus was shown to be primarily with the less aggressive form of the disease. For individuals undergoing a biopsy of the prostate, we have shown that the variant rs17632542-T is associated with a greater probability of not being diagnosed with cancer. Together, these results suggest that the reported association with prostate cancer at the *KLK3* locus is mainly driven by its effect on PSA levels and the increased frequency of PSA testing in men.

Our results from estimating the combined relative effect of the 4 variants primarily associated with PSA levels demonstrate a considerable variation in PSA levels between individuals based on the genotypes of these 4 variants. By applying the combined genetic effect on commonly used PSA cutoff values, a personalized PSA cutoff value can be obtained. Furthermore, our results show after applying a genetic correction that 6-7% of Icelandic men undergoing PSA-based prostate cancer screening have at least one PSA measurement shifted above or below a commonly applied PSA cutoff value for deciding whether to biopsy. Hence, these men would be reclassified with respect to whether they should undergo a biopsy. The results of the AUC analysis show that when genetic correction is applied to the PSA data from Icelandic and UK men an improvement in the prediction accuracy is observed. The greatest improvement in prediction accuracy is seen when both the genetic correction of PSA levels and the combined effect of prostate cancer risk variants are included. For a screening test as important and widely used as the PSA test, having a better way to interpret the measured serum PSA levels is likely to improve substantially the clinical utility of the test.

We note that the estimates provided here are based solely on our Icelandic and UK study populations and more accurate estimates could be obtained from large prospective studies. Also, given the fast pace of discoveries in the current era of GWAS, additional variants associated with PSA levels could be discovered, suggesting the need for updating of such multivariate models.

In combination with information about age, ethnicity, and family history of the disease, estimates of the effect of genetic variation on prostate cancer risk and PSA levels could lay a foundation for the development of individual prostate cancer screening strategies that would have the ultimate goal of reducing cost and improving quality of life.

## Material and Methods

Results from PSA testing were collected from the three clinical laboratories performing the great majority of all PSA measurements in Iceland. The series of data spanned a period of 15 years (from 1994 to 2009). In total we had information about PSA values from 15,757 men (for further details see Table S1). The men have not been diagnosed with prostate cancer according to the nation-wide Icelandic Cancer Registry (ICR), and had not undergone TURP between 1983 and 2008, based on a list from the Landspítali-University Hospital where 90% of all TURP procedures in the country are performed.

Recruitment of men aged 50-69 in the 'Prostate Testing for Cancer and Treatment' trial (ProtecT) took place at nine sites in the UK. From the ProtecT trial study group, the following number of samples were selected for the present study: 524 men with PSA values >3 ng/ml and diagnosed with prostate cancer after undergoing a needle biopsy, 960 men with PSA values between 3 ng/ml and 10 ng/ml but not diagnosed with prostate cancer after undergoing a 10-core needle biopsy, and 454 men with PSA values < 3 ng/ml who had not

undergone prostate biopsies. For a detailed description of study subjects, genotyping and statistical methods see Supplementary Material.

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## Acknowledgments

We thank the individuals who participated in the study and whose contribution made this work possible. Funding: This project was funded in part by contract number 202059 (PROMARK) from the 7th Framework Program of the European Union, in part by FP7-MC-IAPP Grant agreement no.: 218071 (CancerGene), in part by the Urological Research Foundation, Prostate SPORE Grant (P50 CA90386-05S2), and the Robert H. Lurie Comprehensive Cancer Center Grant (p30 CA60553) to W.J.C. The ProtecT study is ongoing and is funded by the Health Technology Assessment Programme (projects 96/20/06, 96/20/99). The ProtecT trial is supported by Department of Health, England; Cancer Research UK grant number C522/A8649, Medical Research Council of England grant number G0500966, ID 75466, the NCRI, UK, and the Southwest National Health Service Research and Development. The bio-repository from ProtecT is supported by the NCRI (ProMPT) study and the Cambridge and Oxford BMRC grants from NIHR.

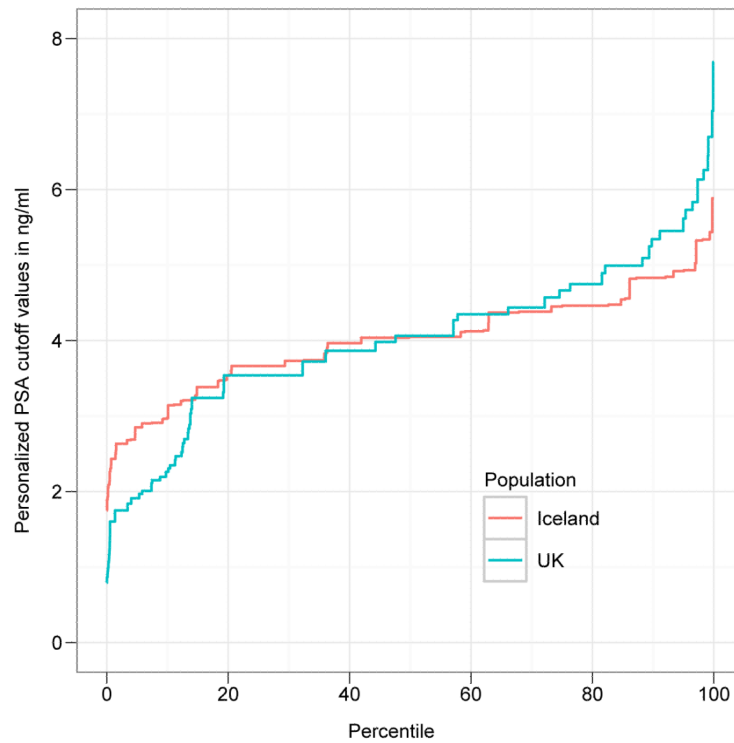
## REFERENCES

1. Barry MJ. Screening for prostate cancer--the controversy that refuses to die. *N Engl J Med.* 2009; 360:1351–1354. [PubMed: 19297564]
2. Aus G, Damber JE, Khatami A, Lilja H, Stranne J, Hugosson J. Individualized screening interval for prostate cancer based on prostate-specific antigen level: results of a prospective, randomized, population-based study. *Arch Intern Med.* 2005; 165:1857–1861. [PubMed: 16157829]
3. Nam RK, Zhang WW, Trachtenberg J, Seth A, Klotz LH, Stanimirovic A, Punnen S, Venkateswaran V, Toi A, Loblaw DA, Sugar L, Siminovitch KA, Narod SA. Utility of incorporating genetic variants for the early detection of prostate cancer. *Clin Cancer Res.* 2009; 15:1787–1793. [PubMed: 19223501]
4. Thompson IM, Ankerst DP, Chi C, Goodman PJ, Tangen CM, Lucia MS, Feng Z, Parnes HL, Coltman CA Jr. Assessing prostate cancer risk: results from the Prostate Cancer Prevention Trial. *J Natl Cancer Inst.* 2006; 98:529–534. [PubMed: 16622122]
5. Schroder FH, Hugosson J, Roobol MJ, Tammela TL, Ciatto S, Nelen V, Kwiatkowski M, Lujan M, Lilja H, Zappa M, Denis LJ, Recker F, Berenguer A, Maattanen L, Bangma CH, Aus G, Villers A, Rebillard X, van der Kwast T, Blijenberg BG, Moss SM, de Koning HJ, Auvinen A. Screening and prostate-cancer mortality in a randomized European study. *N Engl J Med.* 2009; 360:1320–1328. [PubMed: 19297566]
6. Andriole GL, Crawford ED, Grubb RL 3rd, Buys SS, Chia D, Church TR, Fouad MN, Gelmann EP, Kvale PA, Reding DJ, Weissfeld JL, Yokochi LA, O'Brien B, Clapp JD, Rathmell JM, Riley TL, Hayes RB, Kramer BS, Izmirlian G, Miller AB, Pinsky PF, Prorok PC, Gohagan JK, Berg CD. Mortality results from a randomized prostate-cancer screening trial. *N Engl J Med.* 2009; 360:1310–1319. [PubMed: 19297565]
7. van Leeuwen PJ, Connolly D, Gavin A, Roobol MJ, Black A, Bangma CH, Schroder FH. Prostate cancer mortality in screen and clinically detected prostate cancer: estimating the screening benefit. *Eur J Cancer.* 46:377–383. [PubMed: 19804966]
8. Hugosson J, Carlsson S, Aus G, Bergdahl S, Khatami A, Lodding P, Pihl CG, Stranne J, Holmberg E, Lilja H. Mortality results from the Goteborg randomised population-based prostate-cancer screening trial. *Lancet Oncol.* 2010
9. Crawford ED, Grubb R 3rd, Black A, Andriole GL Jr, Chen MH, Izmirlian G, Berg CD, D'Amico AV. Comorbidity and Mortality Results From a Randomized Prostate Cancer Screening Trial. *J Clin Oncol.*
10. Emilsson V, Thorleifsson G, Zhang B, Leonardson AS, Zink F, Zhu J, Carlson S, Helgason A, Walters GB, Gunnarsdottir S, Mouy M, Steinthorsdottir V, Eiriksdottir GH, Bjornsdottir G, Reynisdottir I, Gudbjartsson D, Helgadottir A, Jonasdottir A, Jonasdottir A, Styrkarsdottir U,

Gretarsdottir S, Magnusson KP, Stefansson H, Fossdal R, Kristjansson K, Gislason HG, Stefansson T, Leifsson BG, Thorsteinsdottir U, Lamb JR, Gulcher JR, Reitman ML, Kong A, Schadt EE, Stefansson K. Genetics of gene expression and its effect on disease. *Nature*. 2008; 452:423–428. [PubMed: 18344981]

11. Bansal A, Murray DK, Wu JT, Stephenson RA, Middleton RG, Meikle AW. Heritability of prostate-specific antigen and relationship with zonal prostate volumes in aging twins. *J Clin Endocrinol Metab*. 2000; 85:1272–1276. [PubMed: 10720075]
12. Pilia G, Chen WM, Scuteri A, Orru M, Albai G, Dei M, Lai S, Usala G, Lai M, Loi P, Mameli C, Vacca L, Deiana M, Olla N, Masala M, Cao A, Najjar SS, Terracciano A, Nedorezov T, Sharov A, Zonderman AB, Abecasis GR, Costa P, Lakatta E, Schlessinger D. Heritability of cardiovascular and personality traits in 6,148 Sardinians. *PLoS Genet*. 2006; 2:e132. [PubMed: 16934002]
13. Eeles RA, Kote-Jarai Z, Giles GG, Olama AA, Guy M, Jugurnauth SK, Mulholland S, Leongamornlert DA, Edwards SM, Morrison J, Field HI, Southey MC, Severi G, Donovan JL, Hamdy FC, Dearnaley DP, Muir KR, Smith C, Bagnato M, Arden-Jones AT, Hall AL, O'Brien LT, Gehr-Swain BN, Wilkinson RA, Cox A, Lewis S, Brown PM, Jhavar SG, Tymrakiewicz M, Lophatananon A, Bryant SL, Horwich A, Huddart RA, Khoo VS, Parker CC, Woodhouse CJ, Thompson A, Christmas T, Ogden C, Fisher C, Jamieson C, Cooper CS, English DR, Hopper JL, Neal DE, Easton DF. Multiple newly identified loci associated with prostate cancer susceptibility. *Nat Genet*. 2008; 40:316–321. [PubMed: 18264097]
14. Ahn J, Berndt SI, Wacholder S, Kraft P, Kibel AS, Yeager M, Albanes D, Giovannucci E, Stampfer MJ, Virtamo J, Thun MJ, Feigelson HS, Cancel-Tassin G, Cussenot O, Thomas G, Hunter DJ, Fraumeni JF Jr, Hoover RN, Chanock SJ, Hayes RB. Variation in *KLK* genes, prostate-specific antigen and risk of prostate cancer. *Nat Genet*. 2008; 40:1032–1034. author reply 1035-1036. [PubMed: 19165914]
15. Wiklund F, Zheng SL, Sun J, Adami HO, Lilja H, Hsu FC, Stattin P, Adolfsson J, Cramer SD, Duggan D, Carpten JD, Chang BL, Isaacs WB, Gronberg H, Xu J. Association of reported prostate cancer risk alleles with PSA levels among men without a diagnosis of prostate cancer. *Prostate*. 2009; 69:419–427. [PubMed: 19116992]
16. Gudbjartsson DF, Walters GB, Thorleifsson G, Stefansson H, Halldorsson BV, Zusmanovich P, Sulem P, Thorlacius S, Gylfason A, Steinberg S, Helgadóttir A, Ingason A, Steinthorsdóttir V, Ólafsdóttir EJ, Ólafsdóttir GH, Jonsson T, Borch-Johnsen K, Hansen T, Andersen G, Jørgensen T, Pedersen O, Aben KK, Witjes JA, Swinkels DW, den Heijer M, Franke B, Verbeek AL, Becker DM, Yanek LR, Becker LC, Tryggvadóttir L, Rafnar T, Gulcher J, Kiemeny LA, Kong A, Thorsteinsdóttir U, Stefansson K. Many sequence variants affecting diversity of adult human height. *Nat Genet*. 2008; 40:609–615. [PubMed: 18391951]
17. Thomas G, Jacobs KB, Yeager M, Kraft P, Wacholder S, Orr N, Yu K, Chatterjee N, Welch R, Hutchinson A, Crenshaw A, Cancel-Tassin G, Staats BJ, Wang Z, Gonzalez-Bosquet J, Fang J, Deng X, Berndt SI, Calle EE, Feigelson HS, Thun MJ, Rodriguez C, Albanes D, Virtamo J, Weinstein S, Schumacher FR, Giovannucci E, Willett WC, Cussenot O, Valeri A, Andriole GL, Crawford ED, Tucker M, Gerhard DS, Fraumeni JF Jr, Hoover R, Hayes RB, Hunter DJ, Chanock SJ. Multiple loci identified in a genome-wide association study of prostate cancer. *Nat Genet*. 2008; 40:310–315. [PubMed: 18264096]
18. Waters KM, Stram DO, Le Marchand L, Klein RJ, Valtonen-Andre C, Peltola MT, Kolonel LN, Henderson BE, Lilja H, Haiman CA. A Common Prostate Cancer Risk Variant 5' of Microseminoprotein- $\beta$  (MSMB) Is a Strong Predictor of Circulating  $\beta$ -Microseminoprotein (MSP) Levels in Multiple Populations. *Cancer Epidemiol Biomarkers Prev*.
19. Xu X, Valtonen-Andre C, Savblom C, Hallden C, Lilja H, Klein RJ. Polymorphisms at the Microseminoprotein- $\beta$  locus associated with physiologic variation in  $\beta$ -microseminoprotein and prostate-specific antigen levels. *Cancer Epidemiol Biomarkers Prev*. 2010; 19:2035–2042. [PubMed: 20696662]
20. Klein RJ, Hallden C, Cronin AM, Ploner A, Wiklund F, Bjartell AS, Stattin P, Xu J, Scardino PT, Offit K, Vickers AJ, Gronberg H, Lilja H. Blood biomarker levels to aid discovery of cancer-related single-nucleotide polymorphisms: kallikreins and prostate cancer. *Cancer Prev Res (Phila)*. 3:611–619. [PubMed: 20424135]

21. Pal P, Xi H, Sun G, Kaushal R, Meeks JJ, Thaxton CS, Guha S, Jin CH, Suarez BK, Catalona WJ, Deka R. Tagging SNPs in the kallikrein genes 3 and 2 on 19q13 and their associations with prostate cancer in men of European origin. *Hum Genet.* 2007; 122:251–259. [PubMed: 17593395]
22. Gudmundsson J, Sulem P, Steinthorsdottir V, Bergthorsson JT, Thorleifsson G, Manolescu A, Rafnar T, Gudbjartsson D, Agnarsson BA, Baker A, Sigurdsson A, Benediktsdottir KR, Jakobsdottir M, Blondal T, Stacey SN, Helgason A, Gunnarsdottir S, Olafsdottir A, Kristinsson KT, Birgisdottir B, Ghosh S, Thorlacius S, Magnusdottir D, Stefansdottir G, Kristjansson K, Bagger Y, Wilensky RL, Reilly MP, Morris AD, Kimber CH, Adeyemo A, Chen Y, Zhou J, So WY, Tong PC, Ng MC, Hansen T, Andersen G, Borch-Johnsen K, Jorgensen T, Tres A, Fuertes F, Ruiz-Echarri M, Asin L, Saez B, van Boven E, Klaver S, Swinkels DW, Aben KK, Graif T, Cashy J, Suarez BK, van Vierssen Trip O, Frigge ML, Ober C, Hofker MH, Wijmenga C, Christiansen C, Rader DJ, Palmer CN, Rotimi C, Chan JC, Pedersen O, Sigurdsson G, Benediktsson R, Jonsson E, Einarsson GV, Mayordomo JI, Catalona WJ, Kiemeny LA, Barkardottir RB, Gulcher JR, Thorsteinsdottir U, Kong A, Stefansson K. Two variants on chromosome 17 confer prostate cancer risk, and the one in TCF2 protects against type 2 diabetes. *Nat Genet.* 2007; 39:977–983. [PubMed: 17603485]
23. The NIH Catalog of Published Genome-Wide Association Studies. Jun. 2010 <http://www.genome.gov/26525384#1>
24. Rafnar T, Sulem P, Stacey SN, Geller F, Gudmundsson J, Sigurdsson A, Jakobsdottir M, Helgadóttir H, Thorlacius S, Aben KK, Blondal T, Thorgeirsson TE, Thorleifsson G, Kristjansson K, Thorisdottir K, Ragnarsson R, Sigurgeirsson B, Skuladottir H, Gudbjartsson T, Isaksson HJ, Einarsson GV, Benediktsdottir KR, Agnarsson BA, Olafsson K, Salvarsdottir A, Bjarnason H, Asgeirsdottir M, Kristinsson KT, Matthiasdottir S, Sveinsdottir SG, Polidoro S, Hoiom V, Botella-Estrada R, Hemminki K, Rudnai P, Bishop DT, Campagna M, Kellen E, Zeegers MP, de Verdier P, Ferrer A, Isla D, Vidal MJ, Andres R, Saez B, Juberias P, Banzo J, Navarrete S, Tres A, Kan D, Lindblom A, Gurzau E, Koppova K, de Vegt F, Schalken JA, van der Heijden HF, Smit HJ, Termeer RA, Oosterwijk E, van Hooij O, Nagore E, Porru S, Steineck G, Hansson J, Buntinx F, Catalona WJ, Matullo G, Vineis P, Kiltie AE, Mayordomo JI, Kumar R, Kiemeny LA, Frigge ML, Jonsson T, Saemundsson H, Barkardottir RB, Jonsson E, Jonsson S, Olafsson JH, Gulcher JR, Masson G, Gudbjartsson DF, Kong A, Thorsteinsdottir U, Stefansson K. Sequence variants at the TERT-CLPTM1L locus associate with many cancer types. *Nat Genet.* 2009; 41:221–227. [PubMed: 19151717]
25. Xu J, Isaacs SD, Sun J, Li G, Wiley KE, Zhu Y, Hsu FC, Wiklund F, Turner AR, Adams TS, Liu W, Trock BJ, Partin AW, Chang B, Walsh PC, Gronberg H, Isaacs W, Zheng S. Association of prostate cancer risk variants with clinicopathologic characteristics of the disease. *Clin Cancer Res.* 2008; 14:5819–5824. [PubMed: 18794092]
26. Kader AK, Sun J, Isaacs SD, Wiley KE, Yan G, Kim ST, Fedor H, DeMarzo AM, Epstein JI, Walsh PC, Partin AW, Trock B, Zheng SL, Xu J, Isaacs W. Individual and cumulative effect of prostate cancer risk-associated variants on clinicopathologic variables in 5,895 prostate cancer patients. *Prostate.* 2009; 69:1195–1205. [PubMed: 19434657]



**Figure 1. Cumulative distribution of the personalized PSA cutoff values after genetic correction** Shown is the cumulative distribution (the x-axis represents the percentile of the population) of the personalized PSA cutoff values after applying a genetic correction to the commonly used PSA cutoff value of 4 ng/ml (shown on the y-axis), using the relative genetic estimates for the 4 PSA SNPs (rs2736098, rs10788160, rs11067228, and rs17632542), in the Icelandic (red line) and UK (blue line) study populations.



Table 1

Association results for SNPs and PSA levels, based on samples from Iceland and UK.

SNP	Allele	Chr	Position (bp)	Iceland			UK			Combined		
				P-value	Freq.	Total (n)	Increase per allele (%)	P-value	Freq.		Total (n)	Increase per allele (%)
rs401681	C	5	1,375,087	$1.88 \times 10^{-09}$	0.55	9,049	7.0	0.002	0.53	451	19.0	$1.20 \times 10^{-10}$
rs2736098*	A	5	1,347,086	$5.10 \times 10^{-10}$	0.33	6,347	10.5	0.021	0.27	450	14.8	$2.84 \times 10^{-10}$
rs10788160	A	10	123,023,539	$8.88 \times 10^{-14}$	0.31	8,686	10.2	0.0012	0.24	453	22.9	$4.50 \times 10^{-15}$
rs10993994	T	10	51,219,502	$9.25 \times 10^{-14}$	0.39	8,870	9.2	0.46	0.38	453	5.4	$6.66 \times 10^{-13}$
rs11067228	A	12	113,578,643	$1.09 \times 10^{-11}$	0.56	8,882	8.3	0.074	0.56	441	9.2	$1.93 \times 10^{-11}$
rs4430796*	A	17	33,172,153	$1.40 \times 10^{-11}$	0.52	6,222	9.4	0.21	0.50	449	6.3	$5.60 \times 10^{-11}$
rs2735839	G	19	56,056,435	$4.84 \times 10^{-43}$	0.87	8,869	25.4	$1.18 \times 10^{-06}$	0.86	445	49.7	$6.26 \times 10^{-47}$
rs17632542*	T	19	56,053,569	$9.00 \times 10^{-40}$	0.91	6,078	39.1	$2.66 \times 10^{-09}$	0.93	435	102.2	$3.05 \times 10^{-46}$

Shown are results for alleles that associate with increased (%) levels of PSA. Results for SNPs present on the Illumina chips are based on genotypes from chip (~50%), in-silico genotyping using family imputation (~30%), and single track assay genotyping (~20%)

\* These SNPs (rs273098, rs4430796, and rs17632542) are not on the Illumina chips used in the present study and results are based on genotypes from HapMap SNP imputation (~70%) and single track assay (~30%) genotyping.

**Table 2**  
Association results for PSA SNPs and outcome from a biopsy of the prostate, combined results for Iceland and UK

SNP	Allele increasing PSA-levels	Chr	Position (bp)	Persons with positive biopsy (n)	Persons with positive biopsy, freq.	Persons with negative biopsy (n)	Persons with negative biopsy, freq.	OR 95% CI	P-value	P <sub>het</sub>
rs2736098	A	5	1,347,086	1,718	0.34	1,907	0.32	1.04 (0.94,1.16)	0.47	0.082
rs10993994	T	10	51,219,502	1,696	0.41	2,082	0.40	1.05 (0.96,1.15)	0.31	0.82
rs10788160	A	10	123,023,539	1,679	0.28	2,084	0.32	0.79 (0.71,0.87)	5.40×10 <sup>-06</sup>	0.092
rs11067228	A	12	113,578,643	1,706	0.55	2,106	0.59	0.87 (0.79,0.95)	0.0034	0.51
rs4430796	A	17	33,172,153	1,858	0.55	1,919	0.53	1.03 (0.97,1.10)	0.37	0.067
rs17632542	T	19	56,053,569	1,873	0.93	1,924	0.95	0.77 (0.63,0.95)	0.013	0.56
rs2735839	G	19	56,056,435	1,743	0.88	2,091	0.89	0.85 (0.74,0.98)	0.026	0.44

Shown are the results from a combined analysis of the Icelandic and UK study groups, the number of individuals (n) that have undergone a biopsy of the prostate and have been diagnosed with cancer of the prostate (positive biopsy; maximum number of individuals with genotypes used in the analysis is 1,870, of those 1,354 are from Iceland and 516 from the UK), the number of individuals (n) that have undergone a biopsy of the prostate and have not been diagnosed with cancer of the prostate (negative biopsy; maximum number of individuals with genotypes used in the analysis is 2,124, of those 1,169 are from Iceland and 955 from the UK), the allele associated with increased PSA levels and the allelic frequency (freq.), the odds ratio (OR), and the two-sided P-value. The OR and P-values were estimated using the Mantel-Haenszel model.