Published in final edited form as:

Eur Urol. 2018 December ; 74(6): 710–719. doi:10.1016/j.eururo.2018.06.021.

Genome-Wide Scan Identifies Role for AOX1 In Prostate Cancer Survival

Weiqliang Li¹, Mridu Middha¹, Mesude Bicak¹, Daniel D. Sjöberg², Emily Vertosick², Anders Dahlin³, Christel Häggström⁴, Göran Hallmans⁵, Ann-Charlotte Rönn⁶, Pär Stattin⁷, Olle Melander³, David Ulmert⁸, Hans Lilja⁹, and Robert J. Klein¹

¹Icahn Institute for Genomics and Multiscale Biology and Department of Genetics and Genomic Sciences, Icahn School of Medicine at Mount Sinai, New York, NY USA

²Department of Epidemiology and Biostatistics, Memorial Sloan Kettering Cancer Center, New York, NY USA

³Department of Clinical Sciences, Malmö, Lund University, Malmö, Sweden

⁴Department of Biobank Research, Umeå University, Umeå, Sweden

⁵Department of Public Health and Clinical Medicine, Nutritional Research, Umeå University, Umeå, Sweden

⁶Clinical Research Center, Karolinska University Hospital, Huddinge, Sweden

⁷Department of Surgical Sciences, Uppsala University, Uppsala, Sweden

⁸Molecular Pharmacology Program, Sloan Kettering Institute, New York, NY USA

⁹Departments of Laboratory Medicine, Surgery, and Medicine, Memorial Sloan Kettering Cancer Center, New York, NY USA; Nuffield Department of Surgical Sciences, University of Oxford, Oxford, UK; Department of Translational Medicine, Lund University, Malmö, Sweden

Abstract

Background: Most men diagnosed with prostate cancer have low risk cancers. How to predict prostate cancer progression at the time of diagnosis remains challenging.

Objective: To identify single nucleotide polymorphisms (SNPs) associated with death from prostate cancer.

Design, Setting, and Participants: Blood samples from 11,506 men in Sweden were collected during 1991–1996. Of these, 1053 men were diagnosed with prostate cancer and 245 died from the disease. Stage and grade at diagnosis and outcome information was obtained and DNA from all cases was genotyped.

Corresponding authors: Hans Lilja, M.D., Ph.D., Memorial Sloan Kettering Cancer Center, 1275 York Avenue, New York, NY 10065, United States, Robert J. Klein, 1 Gustave L Levy Place, Box 1498, New York, NY, 10029-6574, United States, Tel: (212) 824-8949; robert.klein@mssm.edu.

Disclosures:

Hans Lilja holds patents for free PSA, hK2, and intact PSA assays, and is named on a patent for a statistical method to detect prostate cancer. The marker assay patents and the patent for the statistical model has been licensed and commercialized as the 4Kscore by OPKO Diagnostics. Dr. Lilja receives royalties from sales of this test, and owns stock in OPKO.

Outcome Measurements and Statistical Analysis: 6,126,633 SNPs were tested for association with prostate cancer specific survival time using a Cox proportional hazards model, adjusted for age at diagnosis and stage and grade at diagnosis. 1×10^{-6} was used as suggestive significance threshold. Positive candidate SNPs were tested for association with gene expression using expression quantitative trait loci (eQTLs) analysis.

Results and Limitations: We found 12 SNPs at seven independent loci associated with prostate cancer specific survival time. One of 6,126,633 SNPs tested reached genome-wide significance ($p < 5 \times 10^{-8}$) and replicated in an independent cohort: rs73055188 ($p = 5.27 \times 10^{-9}$, per-allele HR = 2.27, 95% CI 1.72–2.98) in the AOX1 gene. A second SNP reached a suggestive level of significance ($p < 1 \times 10^{-6}$) and replicated in an independent cohort: rs2702185 ($p = 7.1 \times 10^{-7}$, per-allele HR = 2.55, 95% CI = 1.76–3.69) in the SMG7 gene. rs73055188 is correlated with AOX1 expression levels, which is associated with biochemical recurrence of prostate cancer in independent cohorts. This association is yet to be validated in other ethnic groups.

Conclusions: rs73055188 at the AOX1 locus is associated with prostate cancer specific survival time and AOX1 gene expression level is correlated with biochemical recurrence of prostate cancer.

Patient summary: We identify two genetic markers that are associated with prostate cancer specific survival time.

Keywords

AOX1; eQTL; GWAS; Prostate cancer; Survival analysis

Introduction:

Prostate cancer is one of the most common cause of cancer mortality among men in the U.S.¹. While a large European study found that systematic prostate-specific antigen (PSA) screening significantly reduces death from prostate cancer², widespread use of PSA-testing has also increased the rate of diagnosis of prostate cancer. A recent joint analysis of two large trials of prostate cancer screening found consistent evidence that screening reduces mortality³. Nearly 90% of prostate cancers are clinically localized at the time of their diagnosis. While some of these patients will develop aggressive prostate cancer and eventually die from the disease, the majority of men will remain to have low risk cancers with no adverse clinical outcome. The challenging clinical question is how to predict risk of prostate cancer progression at the time of diagnosis. Several clinical parameters (age at diagnosis, Gleason score, PSA levels and cancer stage) are established predictors, albeit predictive accuracy may vary. Although radical prostatectomy is widely used to prevent unfavorable prostate cancer outcome, a recent study among men with localized prostate cancer showed that radical prostatectomy or radiation significantly reduced clinical disease progression but did not reduce prostate-cancer mortality at 10 years as compared with surveillance⁴.

In families with prostate cancer history, death from prostate cancer among offspring are related to prostate cancer specific survival in parents^{5,6}. This suggests the existence of genetic factors that influence prostate cancer prognosis. Genome-wide association studies have successfully identified numerous loci associated with prostate cancer risk⁷. However,

no clear direct association between these variants and prostate cancer mortality has been observed^{8–11}. Several genetic variants have been found to be associated with prostate cancer aggressiveness^{12,13} or Gleason score¹⁴, though these findings have yet to be replicated. Due in part to the requirement of long-term follow-up, scanning for genetic variants related to prostate specific survival remains unsuccessful¹⁵.

Identifying genetic variants associated with prostate cancer specific survival could inform clinical decision-making and reduce unnecessary intervention and overtreatment of indolent prostate cancer. Here we performed GWAS using multivariate analysis in the context of a population-based cohort of 11,506 men providing blood in 1991–96 for the Malmö Diet and Cancer when rates of PSA-testing in Sweden was low and with 20-year followup data available on prostate cancer specific death. We identified and replicated several SNPs associated with prostate cancer specific survival.

Materials and Methods

The MDC cohort and genotyping

The Malmö Diet and Cancer (MDC) cohort study has been described previously¹⁶. 11,506 men (European ancestry, white) aged 45–73 in Malmö, Sweden were recruited during 1991–1996. Blood specimens were collected and stored under protocols that maintained high sample quality after several years¹⁷. 1476 men were diagnosed with prostate cancer and 317 died from the disease by December 31, 2014. Stage and grade at diagnosis and outcome information was obtained from the Swedish Cancer Registry, the National Prostate Cancer Registry of the Southern Region, and cause of death obtained from the Causes of Death Registry at the National Board of Health and Welfare in Sweden¹⁸. Of these 1476 cases, we had complete clinical information on 1053 of them. These samples were genotyped on the Illumina OmniExpress Exome BeadChip. Genotype calls were made in the Illumina BeadStudio software package and processed with PLINK¹⁹. For quality control, SNPs with call rates less than 99% or minor allele frequency (MAF) less than 1% were removed. We excluded SNPs that deviated from Hardy–Weinberg equilibrium (HWE) in the genotyped cases ($p < 1 \times 10^{-5}$), leaving 586,007 SNPs were available for imputation. Genotype imputation was performed using SHAPEIT²⁰ and IMPUTE2²¹. Imputed SNPs were filtered by MAF ≥ 0.05 and imputation quality score > 0.4 resulting in 6,126,633 SNPs. To insure that we did not include potential first degree relatives, we used PLINK's --genome command to identify individuals with an IBD estimate (PI_HAT) > 0.35 . After removing one of each such pair, we repeated the survival association analysis (below) for each of the top SNPs.

GWAS of prostate cancer survival

Each SNP was tested for association with prostate cancer specific survival using a Cox proportional hazards model in the survival package for R²². Survival time was measured from the date of prostate cancer diagnosis until the date of prostate cancer specific death or last follow-up date (up to 22 years)^{23–25}. Genotypes were coded with an additive model. Aggressive prostate cancer was defined dichotomously as Gleason score ≥ 8 or WHO grade of 3; advanced prostate cancer was dichotomously defined as clinical stage T ≥ 3 or

clinical stage $M = 1$ (metastases at diagnosis)²⁶. A log-additive effect (HR per additional minor allele) was estimated for each SNP adjusting for age at diagnosis, aggressiveness, and advanced stage. SNPs with $p < 5 \times 10^{-8}$ using a 1-df Wald test were considered significant after multiple testing, and a secondary threshold suggestive of association ($p < 1 \times 10^{-6}$) was also used.

Replication in the Västerbotten Intervention Project

To test if the SNPs associations with prostate cancer survival replicates in a second independent cohort, we utilized DNA isolated from blood collected from participants in the Västerbotten Intervention Project (VIP). VIP is an on-going population-based intervention study initiated in 1985 that invites all subjects residing in the county of Västerbotten (population 260,000) to participate in a health survey when they reach the ages of 40, 50, and 60 years.

Participants are asked to complete a self-administered questionnaire and a blood sample from each subject is drawn at study entry. All participants signed an IRB-approved informed consent form. In 2012, the VIP cohort was linked to the Northern Sweden Regional Cancer Registry, part of the Swedish Cancer Registry, using the Swedish personal identity number. We identified 1,596 incident prostate cancer cases in the VIP cohort, 1,552 of which had cryopreserved blood available for analysis. Clinical data were retrieved from the National Prostate Cancer Register. We reviewed hospital medical charts of men diagnosed with cancer to identify men who later had documented evidence of metastatic disease (ie, a positive bone scan) during the follow-up period. There were 145 patients with metastatic PCa diagnosed during follow-up who subsequently died from PCa, according to the Cause of Death Registry. Cause of death was assessed by medical chart review or, when charts were not available, the Swedish Cause of Death Registry. An additional 27 men who died of PCa but who had not had metastases diagnosed prior to death were considered to have had metastatic disease at the date of death. To perform the replication analysis, DNA samples were genotyped using the MassArray mass spectrometry system manufactured by Agena Bioscience (formerly Sequenom). Statistical analyses were performed as for the MDC cohort.

Cis-eQTL analysis

To perform eQTL analysis, we downloaded SNP and gene expression data from a cohort of 470 prostate samples from dbGaP (phs000985.v1.p1) that were described before²⁷. Candidate SNPs with $P < 1 \times 10^{-6}$ from the survival GWAS were selected for the analysis. For each SNP, cis-genes within 10^6 base pairs were taken for analysis. Each SNP was tested for association with corresponding genes using linear regression. Positive association was reported when $P < 0.05$.

AOX1 gene expression analysis using TCGA data

To examine the relationship between SNP, DNA methylation, and AOX1 gene expression in prostate tumor samples, we used data from the Cancer Genome Atlas (TCGA)²⁸. Raw CEL files for blood or adjacent normal tissue was downloaded from the TCGA portal in November 2014. SNPs were called using the Affymetrix birdseed algorithm (apt-probeset-

genotype). SNPs were removed if they had more than 10% missing data; individuals with more than 10% missing data were then removed. Phasing and imputation was performed on these individuals using SHAPEIT and IMPUTE2, above. Transcriptome data and DNA methylation data from 498 prostate cancer case were downloaded from the Genomic Data Commons (<https://portal.gdc.cancer.gov/>). Association between SNP, DNA methylation, and gene expression was tested using multivariate linear regression in R. All controlled access TCGA data was obtained through dbGaP authorization under phs000178.

Biochemical recurrence analysis

AOX1 gene expression level was tested for association with biochemical recurrence of prostate cancer using the “camcApp” app²⁹. Transcriptome data and biochemical recurrence data from 111 prostate cancer cases in the Cambridge cohort³⁰ and 140 prostate cancer cases in the MSKCC cohort³¹ were used. For each cohort, prostate cancer cases were allocated into two different groups based on the expression level of AOX1 gene. A Kaplan-Meier plot was then displayed to compare the probability of freedom from biochemical recurrence between two groups. Statistical significance was determined by P value from Cox proportional hazards test.

Results

Demographics of prostate cancer cases in the MDC cohort

Of 1476 men who had incident diagnosis of prostate cancer in the Malmö Diet and Cancer (MDC) cohort (Fig. 1), 317 of them died from it (Supplementary Table S1). The median age at prostate cancer diagnosis is 70.0. The five-year cumulative incidence of death from prostate cancer was 10.6%. The median length of observation time from time of diagnosis was 8.6 years.

Genome-wide scan for novel SNPs associated with prostate cancer specific survival

DNA available from each of the 1053/1476 cases was genotyped for 586,000 SNPs. From this, genotypes of 6,126,633 SNPs were imputed and tested for association with prostate cancer specific survival. After adjusting for age, stage, and grade at diagnosis, two of the SNPs tested in this analysis reached genome-wide significance: rs73055188 ($p=5.27 \times 10^{-9}$, per-allele HR = 2.27, 95% CI 1.72–2.98) and rs35105661 ($p=1.10 \times 10^{-8}$, per-allele HR = 2.03, 95% CI 1.59–2.59). Additionally, 10 candidate SNPs for which the evidence of association was suggestive ($P < 1 \times 10^{-6}$) were found (Fig. 2A). No evidence for test statistic inflation was observed in a quantile-quantile plot (Fig. 2B) by plotting observed P values against theoretical P values based on uniform distribution. These 12 candidate SNPs are found at 7 unique loci (Table 1) that include five cancer related genes (AOX1, HERC4, TLE4, CASC5, SMG7). We note that numerous SNPs recently reported to be associated with prostate cancer survival^{32,33} do not appear to be significant in our data at all.

To test the robustness of these predictions, we separately added PSA at diagnosis (either on the linear or log scale) and initial treatment (categorized as conservative, curative, and non-curative) to the model. In all cases, the 12 candidate SNPs remain associated with prostate cancer specific survival time (Supplementary Table 2). We also removed 15 individuals who

were predicted to be first degree relatives of others in the cohort by PLINK, recognizing that some of these could be false positives. Removing these individuals did not substantially change the results (data not shown).

To validate our results, four candidate SNPs from this analysis were genotyped in VIP cohort. Using the same Cox proportional hazard model, two SNPs are validated in VIP cohort: rs73055188 ($p=0.019$, per-allele HR = 1.76, 95% CI 1.30–2.25) located in the AOX1 gene and rs2702185 ($p=0.042$, per-allele HR = 1.67, 95% CI 1.02–2.73) located in the SMG7 gene (Supplementary Table 3). We also tested eight SNPs that had been previously reported to be associated with prostate cancer survival; as in the MDC cohort, no associations were observed for any of these SNPs (Supplementary Table 3).

Cis-eQTL analysis identifies candidate genes whose expression level is influenced by prostate cancer survival-associated SNPs

As none of these significant SNPs alter the coding sequence of the genes, we hypothesized that they could influence gene expression. To address this hypothesis, we performed cis-eQTL analysis to test each SNP for association with the expression levels of nearby genes. Of the 54 tests performed, 14 were nominally significant ($p<0.05$; Table 2). rs35105661 is correlated with the expression level of three genes (SIRT1, DNAJC12, HERC4). Five SNPs (rs16834021, rs16834027, rs16834034, rs73055188, rs73055201) at the same locus correlate with AOX1 expression. rs16970844 is correlated with the expression level of four genes (C15orf57, CASC5, GCHFR, IVD).

rs73055188 is a lead SNP that is associated with prostate cancer specific survival and AOX1 gene expression

The most significant association with prostate cancer specific survival was observed for rs73055188, which is associated with prostate cancer specific survival regardless of cancer grade or cancer stage (Fig. 3). This SNP appears to fit a dominant model, as similar hazard ratios were observed for heterozygotes and minor homozygotes (Heterozygote HR=2.27, 95% CI 1.67–3.11, $p=2.26\times 10^{-7}$; homozygote HR=2.28, 95% CI 1.37–3.8, $p=0.0014$). Four other SNPs at this locus also met our threshold of suggestive significance (Table 1). These SNPs all appear to represent one association signal (Fig. 4A).

DNA methylation changes at AOX1 in prostate tumors^{34–36} could be associated with prostate cancer progression^{37,38}. To examine the relationship between rs73055188, DNA methylation and gene expression, we utilized data from The Cancer Genome Atlas (TCGA)²⁸. Of 18 CpG sites near AOX1, 14 show a difference in methylation between tumor and normal tissue (Fig. 4B). In a multivariate model, both methylation levels at some of the 18 CpG sites and genotype of rs73055188 are associated with AOX1 expression levels (Fig. 4C). To ask if rs73055188 influences AOX1 expression in non-cancerous prostate tissue, we used a cohort of matched gene expression and genotype data available in dbGaP ($n=470$). Minor allele A at rs73055188 was associated with lower AOX1 expression (Fig. 4D; $P = 0.0015$).

AOX1 expression is correlated with biochemical recurrence of prostate cancer

Based on these results, we hypothesized that lower AOX1 expression would correlate with worse outcomes in prostate cancer. To address this hypothesis, we made use of two prostate cancer cohorts whose data is available through camcAPP (<http://bioinformatics.cruk.cam.ac.uk/apps/camcAPP/>). In both studies from Cambridge (n = 111) and MSKCC (n = 140), lower expression level of AOX1 is associated with shorter time to biochemical recurrence (Fig. 5).

Discussion

We performed a genome-wide scan for genetic variants associated with prostate cancer specific mortality and identified several SNPs associated with unfavorable prognosis. Notably, by adjusting for known predictors of prostate cancer outcome, the SNPs we identified have direct effects on survival time after initial cancer diagnosis rather than risk of developing disease or stage or grade at diagnosis.

Twelve SNPs at seven independent loci were associated with prostate cancer specific survival time ($P < 1 \times 10^{-6}$). Two independent loci are genome-wide significant ($P < 5 \times 10^{-8}$). Two of four SNPs tested in the replication cohort, including the top SNP from the discovery cohort, replicated ($p < 0.05$). Multiple candidate genes were identified and cis-eQTL analysis further found 14 positive associations between SNPs and gene expression level. Several of these putative target genes have been previously implicated in cancer. Importantly, the AOX1 gene was shown to be hypermethylated in prostate tumor cells and its methylation signature was suggested as a candidate biomarker for prostate cancer outcome^{37,38}. The mouse ortholog of AOX1 is regulated by testosterone^{39,40}. TLE4 has been shown to play roles in multiple cancer types such as acute myeloid leukemia⁴¹ and colorectal cancer⁴². SIRT1 has been suggested to play roles in many cancer types including prostate cancer⁴³.

As rs73055188 at AOX1 was the most significant finding, we focused on this locus. We show that rs73055188 is associated with AOX1 gene expression. Evidence from two additional cohorts suggests that lower expression of AOX1 is correlated with earlier biochemical recurrence of prostate cancer.

The role of the second replicated SNP, rs2702185, is less clear. This SNP is located in the intron of SMG7. SMG7 binds both the tumor suppressor p53 and its regulator MDM2, and has been implicated in ATM-mediated phosphorylation of MDM2⁴⁴. However, the eQTL analysis points to two other genes at this locus: NCF2 and ARPC5. NCF2 encodes a subunit of NADPH oxidase that is a direct target of p53⁴⁵. ARPC5 plays a role in actin polymerization and is regulated by miR-141, which has tumor and metastasis suppressing properties in prostate cancer⁴⁶. Further work will be needed to understand the relative contribution of these genes to prostate cancer survival.

A previous meta-analysis GWAS of prostate cancer survival did not find any significant result despite a larger number of observed events¹⁵. One potential reason is that the meta-analysis used large cohorts of cases from various ethnic groups and various countries. Thus, both heterogeneity in the effect of SNPs on survival among different ancestry groups as well

as differences in standard of care for prostate cancer among different healthcare systems could dilute the effect observed in such a study. In contrast, our study focuses on one cohort of northern European ancestry, where rates of PSA-testing until recently have remained low and who were treated within the Swedish healthcare system. However, the gene expression evidence from Cambridge and MSKCC supports a role for AOX1 beyond the context of Sweden.

There are several limitations to this study. As both cohorts were from Sweden, it is not clear if these results will apply in other populations with non-European ancestry. As has been found for other traits, it may be feasible to identify additional loci associated with prostate cancer survival through studies of non-European populations. We were unable to perform a multivariable analysis using the Cambridge and MSKCC cohorts because of the lack of clinical annotations and limited access to the data. For future study, GWAS of prostate cancer survival may be further improved by stratifying based on initial clinical management. Increasing sample size and follow up time will also help to identify SNPs with lower minor allele frequencies or smaller effect size. Because of limited sample size, we are not able to confidently test whether rs73055188 may change AOX1 methylation level. Genetic and epigenetic factors may function independently or cooperatively to regulate AOX1 expression.

Conclusions:

We performed a genome-wide association study and found 12 SNPs at seven independent loci associated with prostate cancer specific survival time. rs73055188 is associated with prostate cancer specific survival regardless of cancer grade or stage and may regulate AOX1 gene expression. Lower expression of AOX1 is correlated with earlier biochemical recurrence of prostate cancer.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

The work was supported in part by grants from National Cancer Institute [R01 CA175491, P30 CA196521, P50 CA92629, and P30 CA008748], the Sidney Kimmel Center for Prostate and Urologic Cancers, and David H. Koch through the Prostate Cancer Foundation, the National Institute for Health Research (NIHR) Oxford Biomedical Research Centre Program in UK, the Swedish Cancer Society (CAN 2017/559), and the Swedish Research Council (VR-MH project no. 2016–02974). This work was supported in part through the computational resources and staff expertise provided by Scientific Computing at the Icahn School of Medicine at Mount Sinai, including infrastructure supported by the Office of Research Infrastructure of the National Institutes of Health under award number S10OD018522.

References

1. Siegel RL, Miller KD, Jemal A. Cancer Statistics, 2017. *CA Cancer J Clin.* 2017;67(1):7–30. [PubMed: 28055103]
2. Schroder FH, Hugosson J, Roobol MJ, et al. Screening and prostate cancer mortality: results of the European Randomised Study of Screening for Prostate Cancer (ERSPC) at 13 years of follow-up. *Lancet.* 2014;384(9959):2027–2035. [PubMed: 25108889]

3. Tsodikov A, Gulati R, Heijnsdijk EAM, et al. Reconciling the Effects of Screening on Prostate Cancer Mortality in the ERSPC and PLCO Trials. *Ann Intern Med*. 2017;167(7):449–455. [PubMed: 28869989]
4. Hamdy FC, Donovan JL, Lane JA, et al. 10-Year Outcomes after Monitoring, Surgery, or Radiotherapy for Localized Prostate Cancer. *N Engl J Med*. 2016;375(15):1415–1424. [PubMed: 27626136]
5. Lindstrom LS, Hall P, Hartman M, Wiklund F, Gronberg H, Czene K. Familial concordance in cancer survival: a Swedish population-based study. *Lancet Oncol*. 2007;8(11):1001–1006. [PubMed: 17921068]
6. Hemminki K Familial risk and familial survival in prostate cancer. *World J Urol*. 2012;30(2):143–148. [PubMed: 22116601]
7. Al Olama AA, Kote-Jarai Z, Berndt SI, et al. A meta-analysis of 87,040 individuals identifies 23 new susceptibility loci for prostate cancer. *Nature genetics*. 2014;46(10):1103–1109. [PubMed: 25217961]
8. Gallagher DJ, Vijai J, Cronin AM, et al. Susceptibility loci associated with prostate cancer progression and mortality. *Clinical cancer research : an official journal of the American Association for Cancer Research*. 2010;16(10):2819–2832.
9. Sullivan J, Kopp R, Stratton K, et al. An analysis of the association between prostate cancer risk loci, PSA levels, disease aggressiveness and disease-specific mortality. *British journal of cancer*. 2015;113(1):166–172. [PubMed: 26068399]
10. Shui IM, Lindstrom S, Kibel AS, et al. Prostate cancer (PCa) risk variants and risk of fatal PCa in the National Cancer Institute Breast and Prostate Cancer Cohort Consortium. *European urology*. 2014;65(6):1069–1075. [PubMed: 24411283]
11. Helfand BT, Roehl KA, Cooper PR, et al. Associations of prostate cancer risk variants with disease aggressiveness: results of the NCI-SPORE Genetics Working Group analysis of 18,343 cases. *Human genetics*. 2015;134(4):439–450. [PubMed: 25715684]
12. Xu J, Zheng SL, Isaacs SD, et al. Inherited genetic variant predisposes to aggressive but not indolent prostate cancer. *Proc Natl Acad Sci U S A*. 2010;107(5):2136–2140. [PubMed: 20080650]
13. Amin Al Olama A, Kote-Jarai Z, Schumacher FR, et al. A meta-analysis of genome-wide association studies to identify prostate cancer susceptibility loci associated with aggressive and non-aggressive disease. *Human molecular genetics*. 2013;22(2):408–415. [PubMed: 23065704]
14. Berndt SI, Wang Z, Yeager M, et al. Two susceptibility loci identified for prostate cancer aggressiveness. *Nat Commun*. 2015;6:6889. [PubMed: 25939597]
15. Szulkin R, Karlsson R, Whittington T, et al. Genome-Wide Association Study of Prostate Cancer-Specific Survival. *Cancer Epidemiol Biomarkers Prev*. 2015;24(11):1796–1800. [PubMed: 26307654]
16. Manjer J, Carlsson S, Elmstahl S, et al. The Malmo Diet and Cancer Study: representativity, cancer incidence and mortality in participants and non-participants. *Eur J Cancer Prev*. 2001;10(6):489–499. [PubMed: 11916347]
17. Pero RW, Olsson A, Bryngelsson C, et al. Quality control program for storage of biologically banked blood specimens in the Malmo Diet and Cancer Study. *Cancer Epidemiol Biomarkers Prev*. 1998;7(9):803–808. [PubMed: 9752989]
18. Sjoberg DD, Vickers AJ, Assel M, et al. Twenty-year Risk of Prostate Cancer Death by Midlife Prostate-specific Antigen and a Panel of Four Kallikrein Markers in a Large Population-based Cohort of Healthy Men. *European urology*. 2018.
19. Purcell S, Neale B, Todd-Brown K, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet*. 2007;81(3):559–575. [PubMed: 17701901]
20. Delaneau O, Marchini J, Zagury JF. A linear complexity phasing method for thousands of genomes. *Nat Methods*. 2011;9(2):179–181. [PubMed: 22138821]
21. Howie BN, Donnelly P, Marchini J. A flexible and accurate genotype imputation method for the next generation of genome-wide association studies. *PLoS genetics*. 2009;5(6):e1000529. [PubMed: 19543373]

22. Willis JA, Olson SH, Orlov I, et al. A replication study and genome-wide scan of single-nucleotide polymorphisms associated with pancreatic cancer risk and overall survival. *Clinical cancer research : an official journal of the American Association for Cancer Research*. 2012;18(14):3942–3951. [PubMed: 22665904]
23. Fall K, Stromberg F, Rosell J, Andren O, Varenhorst E, South-East Region Prostate Cancer G. Reliability of death certificates in prostate cancer patients. *Scand J Urol Nephrol*. 2008;42(4):352–357. [PubMed: 18609293]
24. Sandblom G, Dufmats M, Olsson M, Varenhorst E. Validity of a population-based cancer register in Sweden--an assessment of data reproducibility in the South-East Region Prostate Cancer Register. *Scand J Urol Nephrol*. 2003;37(2):112–119. [PubMed: 12745718]
25. Van Hemelrijck M, Wigertz A, Sandin F, et al. Cohort Profile: the National Prostate Cancer Register of Sweden and Prostate Cancer data Base Sweden 2.0. *Int J Epidemiol*. 2013;42(4):956–967. [PubMed: 22561842]
26. Klein RJ, Hallden C, Gupta A, et al. Evaluation of multiple risk-associated single nucleotide polymorphisms versus prostate-specific antigen at baseline to predict prostate cancer in unscreened men. *European urology*. 2012;61(3):471–477. [PubMed: 22101116]
27. Larson NB, McDonnell S, French AJ, et al. Comprehensively evaluating cis-regulatory variation in the human prostate transcriptome by using gene-level allele-specific expression. *Am J Hum Genet*. 2015;96(6):869–882. [PubMed: 25983244]
28. Cancer Genome Atlas Research N. The Molecular Taxonomy of Primary Prostate Cancer. *Cell*. 2015;163(4):1011–1025. [PubMed: 26544944]
29. Dunning MJ, Vowler SL, Lalonde E, et al. Mining Human Prostate Cancer Datasets: The “camcAPP” Shiny App. *EBioMedicine*. 2017;17:5–6. [PubMed: 28286059]
30. Ross-Adams H, Lamb AD, Dunning MJ, et al. Integration of copy number and transcriptomics provides risk stratification in prostate cancer: A discovery and validation cohort study. *EBioMedicine*. 2015;2(9):1133–1144. [PubMed: 26501111]
31. Taylor BS, Schultz N, Hieronymus H, et al. Integrative genomic profiling of human prostate cancer. *Cancer Cell*. 2010;18(1):11–22. [PubMed: 20579941]
32. Briollais L, Ozcelik H, Xu J, et al. Germline Mutations in the Kallikrein 6 Region and Predisposition for Aggressive Prostate Cancer. *J Natl Cancer Inst*. 2017;109(4).
33. FitzGerald LM, Zhao S, Leonardson A, et al. Germline variants in IL4, MGMT and AKT1 are associated with prostate cancer-specific mortality: An analysis of 12,082 prostate cancer cases. *Prostate Cancer Prostatic Dis*. 2018.
34. Moller M, Strand SH, Mundbjerg K, et al. Heterogeneous patterns of DNA methylation-based field effects in histologically normal prostate tissue from cancer patients. *Sci Rep*. 2017;7:40636. [PubMed: 28084441]
35. Geybels MS, Zhao S, Wong CJ, et al. Epigenomic profiling of DNA methylation in paired prostate cancer versus adjacent benign tissue. *Prostate*. 2015;75(16):1941–1950. [PubMed: 26383847]
36. Kim JW, Kim ST, Turner AR, et al. Identification of new differentially methylated genes that have potential functional consequences in prostate cancer. *PLoS One*. 2012;7(10):e48455. [PubMed: 23119026]
37. Shui IM, Wong CJ, Zhao S, et al. Prostate tumor DNA methylation is associated with cigarette smoking and adverse prostate cancer outcomes. *Cancer*. 2016;122(14):2168–2177. [PubMed: 27142338]
38. Haldrup C, Mundbjerg K, Vestergaard EM, et al. DNA methylation signatures for prediction of biochemical recurrence after radical prostatectomy of clinically localized prostate cancer. *J Clin Oncol*. 2013;31(26):3250–3258. [PubMed: 23918943]
39. Vila R, Kurosaki M, Barzago MM, et al. Regulation and biochemistry of mouse molybdo-flavoenzymes. The DBA/2 mouse is selectively deficient in the expression of aldehyde oxidase homologues 1 and 2 and represents a unique source for the purification and characterization of aldehyde oxidase. *J Biol Chem*. 2004;279(10):8668–8683. [PubMed: 14665639]
40. Kurosaki M, Demontis S, Barzago MM, Garattini E, Terao M. Molecular cloning of the cDNA coding for mouse aldehyde oxidase: tissue distribution and regulation in vivo by testosterone. *Biochem J*. 1999;341 (Pt 1):71–80. [PubMed: 10377246]

41. Dayyani F, Wang J, Yeh JR, et al. Loss of TLE1 and TLE4 from the del(9q) commonly deleted region in AML cooperates with AML1-ETO to affect myeloid cell proliferation and survival. *Blood*. 2008;111(8):4338–4347. [PubMed: 18258796]
42. Wang SY, Gao K, Deng DL, et al. TLE4 promotes colorectal cancer progression through activation of JNK/c-Jun signaling pathway. *Oncotarget*. 2016;7(3):2878–2888. [PubMed: 26701208]
43. Yu Y, Zhang Q, Meng Q, et al. Mesenchymal stem cells overexpressing Sirt1 inhibit prostate cancer growth by recruiting natural killer cells and macrophages. *Oncotarget*. 2016;7(44):71112–71122. [PubMed: 27764779]
44. Luo H, Cowen L, Yu G, Jiang W, Tang Y. SMG7 is a critical regulator of p53 stability and function in DNA damage stress response. *Cell Discov*. 2016;2:15042. [PubMed: 27462439]
45. Italiano D, Lena AM, Melino G, Candi E. Identification of NCF2/p67phox as a novel p53 target gene. *Cell Cycle*. 2012;11(24):4589–4596. [PubMed: 23187810]
46. Liu C, Liu R, Zhang D, et al. MicroRNA-141 suppresses prostate cancer stem cells and metastasis by targeting a cohort of pro-metastasis genes. *Nat Commun*. 2017;8:14270. [PubMed: 28112170]

Figure 1.
Flowchart of analysis.

Figure 2.

Results from a GWAS of prostate cancer specific survival time in the MDC cohort. A: Manhattan plot of P values for each SNP highlighting key chromosomal regions. P values are for variant associations with prostate cancer survival time from Cox proportional hazard model, each adjusted for cancer diagnosis age, principal components and cancer stages at diagnosis. Horizontal red line indicates genome-wide statistical significance ($P < 5 \times 10^{-8}$) and blue line indicate suggestive significance ($P < 1 \times 10^{-5}$). B: Quantile-quantile (q-q) plot of GWAS results. Observed P value from association study for each SNP is plotted against expected P value from continuous uniform distribution.

Figure 3.

rs73055188 is associated with prostate cancer specific survival time regardless of cancer grade or cancer stage. The x-axis is survival years from cancer diagnosis to cancer specific death. Y-axis is percentage of survived patients. Prostate cancer cases from MDC cohort are divided into four groups based on their cancer grade and cancer stage at diagnosis. Survival data from each group is plotted: non-aggressive and non-advanced prostate cancer cases (A), aggressive and non-advanced prostate cancer cases (B), non-aggressive and advanced prostate cancer cases (C), aggressive and advanced prostate cancer cases (D). P value is calculated from Cox-proportional hazard model.

Figure 4.

Linear regression identifies that DNA methylation and rs73055188 are associated with *AOX1* gene expression. A: Regional plot from GWAS in MDC cohort of the top risk SNP (rs73055188). Y-axes represent the statistical significance ($-\log_{10}$ -transformed P values) of SNP association results (left) and the recombination rate (right). The color code for the points represents the r^2 of each SNP with rs73055188 (ranges defined in the legend). Multiple risk SNPs with low P values are located in *AOX1* gene. B: DNA hypermethylation is observed at multiple CpG sites near *AOX1* gene in prostate tumor cells comparing with

adjacent normal cells. TCGA data is used and CpG loci identification is from Illumina HumanMethylation450 BeadChip. C: Relationship between the expression of *AOXI* gene and genotype of rs73055188 and DNA methylation levels at nearby CpG sites. Data from prostate tumors in TCGA is used. Y-axis is regression coefficient for each risk factor using multivariate linear regression model. Error bar shows standard deviation of regression coefficient. Statistical significance is labeled as * ($P < 0.05$) and *** ($P < 0.0005$). D: rs73055188 is associated with *AOXI* expression in adjacent normal prostate tissue from prostate or bladder cancer patients. Box plot shows *AOXI* expression level stratified by different genotypes at rs73055188. P value is calculated using fixed-effect association test.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Figure 5.

Lower expression of *AOX1* gene is correlated with shorter time to biochemical recurrence of prostate cancer. Kaplan–Meier plot is used to plot time to biochemical recurrence against percentage of freedom from biochemical recurrence using data from Cambridge cohort (A) and MSKCC cohort (B). Prostate cancer cases are divided into two groups based on expression of *AOX1*. P value is calculated from the Cox proportional hazard model.

Table 1.

Genome-wide association study results for novel risk polymorphisms associated with prostate cancer specific survival time. Chromosome positions are given for genome build 37. All hazard ratios reflect the increased hazard for carrying the minor allele. The MAF is the pooled frequency across both cases and controls. Gene is closest RefSeq annotated gene within 20kb. HR, 95% CI of HR and P value are calculated for each SNP using cox proportional hazard model as described. Hazard ratios are multiple variable adjusted using age, stage and grade at diagnosis as covariates. Impute score is generated during imputation using IMPUTE2.

Polymorphism	CHR	BP	Gene	Major Allele	Minor Allele	Patients	MAF	HR	CI:low	CI:high	P	Impute score
rs73055188	2q33.1	201532698	<i>A0X1</i>	G	A	1021	0.091	2.266	1.722	2.982	5.27E-09	0.962
rs35105661	10q21.3	69796848	<i>HERC4</i>	CT	C	723	0.284	2.033	1.594	2.593	1.10E-08	0.842
rs11294346	9q21.31	82230557	<i>TLE4</i>	TA	T	1051	0.069	2.293	1.691	3.109	9.16E-08	0.997
rs16834021	2q33.1	201540988	<i>A0X1</i>	C	T	1053	0.097	2.020	1.551	2.630	1.80E-07	1
rs16834027	2q33.1	201541827	<i>A0X1</i>	G	A	1053	0.097	2.020	1.551	2.630	1.80E-07	1
rs16834034	2q33.1	201542014	<i>A0X1</i>	T	G	1053	0.097	2.020	1.551	2.630	1.80E-07	1
rs73055201	2q33.1	201538788	<i>A0X1</i>	G	A	1052	0.098	2.003	1.538	2.609	2.61 E-07	0.993
rs16970844	15q 15.1	40881268	<i>CASC5</i>	C	T	1009	0.057	2.465	1.741	3.489	3.61 E-07	0.938
rs200216754	15q 11.2	23639450	-	C	CA	207	0.251	3.881	2.297	6.559	4.07E-07	0.495
rs13440432	9q21.31	82246520	<i>TLE4</i>	T	G	1053	0.075	2.134	1.591	2.862	4.18E-07	1
rs2702185	1q25.3	183501350	<i>SMG7</i>	c	T	828	0.057	2.548	1.761	3.688	7.05E-07	0.767
rs1075725	15q25.3	88108316	<i>LINC00052</i>	T	c	1036	0.476	1.634	1.342	1.989	9.94E-07	0.991

Abbreviation: CHR: chromosome loci; BP: chromosome coordinate; MAF: minor allele frequency; HR: hazard ratio; CI:low and CI:high: 95% confidence interval of hazard ratio; P: P value.

Table 2.

Cis-eQTL analysis results for top risk SNPs associated with prostate cancer specific survival time. Chromosome positions are given in b37.

CHR	SNP	BP	Gene	Start-End	P	Ref. allele	Risk allele	β
10	rs35105661	69796848	<i>SIRT1</i>	69644427–69678147	1.30E-09	CT	C	0.053
10	rs35105661	69796848	<i>DNAJC12</i>	69556427–69597924	2.59E-09	CT	C	0.154
15	rs 16970844	40881268	<i>C15orf57</i>	40820882–40857256	6.57E-06	C	T	0.173
10	rs35105661	69796848	<i>HERC4</i>	69681665–69835105	9.06E-05	CT	c	0.020
2	rs73055188	201532698	<i>AOX1</i>	201450591–201541787	1.51 E-03	G	A	–0.143
15	rs 16970844	40881268	<i>GCHFR</i>	41056218–41059906	3.21 E-03	C	T	–0.099
2	rs 16834021	201540988	<i>AOX1</i>	201450591–201541787	3.87E-03	C	T	–0.1304
2	rs16834027	201541827	<i>AOX1</i>	201450591–201541787	3.87E-03	G	A	–0.1304
2	rs 16834034	201542014	<i>AOX1</i>	201450591–201541787	3.87E-03	T	G	–0.1304
2	rs73055201	201538788	<i>AOX1</i>	201450591–201541787	6.63E-03	G	A	–0.1197
1	rs2702185	183501350	<i>NCF2</i>	183524698–183560011	0.0153	C	T	0.102
15	rs 16970844	40881268	<i>IVD</i>	40697686–40728146	0.0159	C	T	–0.051
1	rs2702185	183501350	<i>ARPC5</i>	183592401–183604892	0.0180	C	T	–0.029
15	rs 16970844	40881268	<i>CASC5</i>	40886218–40956540	0.0203	C	T	0.138

Abbreviation: CHR: chromosome number; BP: chromosome coordinate of each SNP; Start-End: chromosome coordinate of each gene; P: P value; β : regression coefficient.