Published in final edited form as:

Cancer Epidemiol Biomarkers Prev. 2024 April 03; 33(4): 557–566.

doi:10.1158/1055-9965.EPI-23-0849.

The impact of inherited genetic variation on DNA methylation in prostate cancer and benign tissues of African American and European American men

Dayana Delgado[1], **Marc Gillard**[1], **Lin Tong**[1], **Kathryn Demanelis**[2,3], **Meritxell Oliva**[1,4], **Kevin J. Gleason**[5], **Meytal Chernoff**[1,6,7], **Lin Chen**[1], **Gladell P. Paner**[8], **Donald Vander Griend**[9,10], **Brandon L. Pierce**[1,11,12,*]

[1]Department of Public Health Sciences, University of Chicago, Chicago, IL 60637

[2]Department of Medicine, University of Pittsburgh, Pittsburgh, PA 15261

[3]UPMC Hillman Cancer Center, Pittsburgh, PA 15232

[4]Genomics Research Center, AbbVie, North Chicago, IL 60064

[5]Data and Statistical Sciences, AbbVie, North Chicago, IL 60064

[6]Interdisciplinary Scientist Training Program, University of Chicago, Chicago, IL, USA

[7]University of Chicago Pritzker School of Medicine, Chicago, IL, USA

[8]Department of Pathology, University of Chicago, Chicago, IL 60637

[9]Department of Pathology, University of Illinois at Chicago, Chicago, IL 60607

[10]The University of Illinois Cancer Center, Chicago, IL

[11]Department of Human Genetics, University of Chicago, Chicago, IL 60615

[12]Comprehensive Cancer Center, University of Chicago, Chicago, IL 60637.

Abstract

Background: American men of African ancestry (AA) have higher prostate cancer (PCa) incidence and mortality rates compared to American men of European ancestry (EA). Differences in genetic susceptibility mechanisms may contribute to this disparity.

Methods: To gain insights into the regulatory mechanisms of PCa susceptibility variants, we tested the association between SNPs and DNA methylation (DNAm) at nearby CpG sites across the genome in benign and cancer prostate tissue from 74 AA and 74 EA men. Genome-wide SNP data (from benign tissue) and DNAm were generated using Illumina arrays.

Results: Among AA men, we identified 6,298 and 2,641 cis-meQTLs (FDR of 0.05) in benign and tumor tissue, respectively, with 6,960 and 1,700 detected in EA men. We leveraged GWAS summary statistics to identify previously reported PCa GWAS signals likely to share a common

*Correspondence: Dr. Brandon Pierce, The University of Chicago, 5841 South Maryland Avenue, MC2000, Chicago, IL 60637; 773-702-1917; brandonpierce@uchicago.edu;.

Conflict of Interest Statement: The author declare no potential conflicts of interest.

causal variant with a detected meQTL. We identified 9 GWAS-meQTL pairs with strong evidence of co-localization (4 in EA benign, 3 in EA tumor, 2 in AA benign, and 3 in AA tumor). Among these co-localized GWAS-meQTL pairs, we identified co-localizing eQTLs impacting four eGenes with known roles in tumorigenesis.

Conclusions: These findings highlight epigenetic regulatory mechanisms by which PCa-risk SNPs can modify local DNAm and/or gene expression in prostate tissue.

Impact: Overall, our findings showed general consistency in the meQTL landscape of AA and EA men, but meQTLs often differ by tissue type (normal vs. cancer). Ancestry-based LD differences and lack of AA representation in GWAS decrease statistical power to detect co-localization for some regions.

INTRODUCTION

Prostate cancer (PCa) is the second most common cancer and cause of cancer death among men in the U.S (1). African American (AA) men are disproportionately affected by PCa, with an incidence rate that is 1.7 times higher (2–4) and a mortality rate that is 2–4 times higher than EA men (5,6). The causes of these disparities are likely complex, with social, environmental, and genetic factors contributing (3).

Genome-wide association studies (GWAS) have identified 269 common risk alleles (single nucleotide polymorphisms, SNPs) associated with PCa susceptibility, and these account for ~43% of the familial relative risk for PCa (7,8). More than 90% of PCa-risk SNPs are located in non-coding regions, suggesting that the causal variants underlying these associations impact gene regulation.

Despite the progress in identifying susceptibility loci, the biological mechanisms by which SNPs impact PCa risk are largely unknown. A common approach for understanding the regulatory mechanisms of disease-associated variants is to assess their association with local gene expression and/or epigenetic features (9,10). Regions where SNPs affect gene expression and/or DNA methylation (DNAm) are known as expression Quantitative Trait Loci (eQTLs) and methylation QTLs (meQTLs), respectively. Few studies have identified meQTLs in prostate tissue, and those studies lack adequate representation of individuals of African ancestry. One study reported 7,590 genome-wide cis-meQTLs in PCa tumor samples (11); another focused on 147 PCa-risk SNPs and identified 93 PCa-risk SNPs that were associated with DNAm at nearby CpG sites in PCa tumor tissue (12).

In this study, we attempt to improve our understanding of the mechanisms by which PCa susceptibility SNPs influence PCa biology, by examining these SNPs' association with DNAm at nearby CpG sites across the genome, in both benign and cancerous prostate tissue. To identify mechanisms that may be relevant to disparities, we analyze data from both AA and EA PCa patients and conduct analyses stratified by ancestry. We leverage existing GWAS summary statistics to determine whether GWAS and cis-meQTL association signals (identified in ancestry-specific analyses) are likely to share a common causal variant (using co-localization methods). The identification of co-localizing meQTLs and GWAS loci can provide insights on the epigenetic mechanisms by which SNPs influence PCa risk.

MATERIALS AND METHODS

PCa Patients

Subjects included in this work were male PCa patients who underwent robotic-assisted laparoscopic radical prostatectomy at University of Chicago Medical Center (UCMC) between 2011 and 2017. Trained interviewers from the Epidemiology Research Recruitment Core consented 74 AA and 74 EA eligible men for the collection of questionnaire data, prostate tissue, and access to medical records. All eligible participants were diagnosed with Gleason scores of at least seven. All enrolled patients provided written informed consent. This study was conducted in accordance with recognized ethical guidelines (U.S. Common Rule) and was approved by the Institutional Review Board of The University of Chicago.

Bio-specimen collection

After surgery, prostate specimens were sent to the Human Tissue Resource Center (HTRC) at the University of Chicago. Each prostate specimen underwent histological examination and Gleason grading by a genitourinary pathologist (G.P.P.) at the University of Chicago. The presence of adenocarcinoma was confirmed by the overexpression of alpha-methylacyl-coenzyme-A racemase (AMACR), and areas to sample for DNA extraction were marked. Benign tissue and tumor tissue were collected.

The following criteria were used for the selection of benign tissue from FFPE tissue of the resected prostate: 1) samples were selected from blocks that were free of tumor; 2) tissue from the peripheral zone was prioritized; 3) if no tumor-free areas were available in the peripheral zone, tissue was collected from the central zone. In 10 cases, neither the peripheral nor the central zone were suitable for sampling, so BPH tissue was collected. Cancer tissue was selected from the index lesion in multifocal tumors. Tissue collection was performed using either a 1mm biopunch or by laser capture microdissection of ~100 μm^2 of tissue (8μm thick sections using a Leica LMD 6500 system). In cases where two consecutive diagnostic blocks showed acceptable areas to sample (continuous benign or tumor tissue running through two blocks from base to apex), tissue was collected by punching through the base-most diagnostic block.

DNA extraction

The Gentra Puregene Tissue Kit (QIAgen) was used to extract 1–2 μg of DNA from benign and cancerous prostate samples. We assessed DNA concentration and quality with the NanoDrop and Agilent BioAnalyser. We excluded DNA samples with a concentration <40ng/μL or 260/280 ratio outside the range of <1.6 to ≥2.1 and/or fragmented DNA <2 Kb. The Illumina Infinium HD FFPE restoration kit was used to restore FFPE DNA (according the manufacturer's protocol, including qPCR-based QC of the DNA samples prior to restoration).

SNP genotyping and imputation

Genome-wide SNP data was generated (from benign tissue DNA) using the Illumina Infinium Multi-Ethnic Global-8 v1.0 array at the University of Chicago Genomics Core Facility. Genotypes were called using a GenCall Threshold 0.15. Genotype data consisted

of 148 individuals (74 AA and 74 EA) with 1,707,345 autosomal SNPs measured. We excluded 224,774 SNPs with low call rates (< 99%). Minor allele frequency (MAF) and Hardy-Weinberg equilibrium (HWE) thresholds were applied separately for AA and EA samples. For AA samples, we excluded 855,033 SNPs with MAF <0.05 and 53 SNPs with HWE p-values $<10^{-5}$, resulting in 627,485 high-quality SNPs. For EA samples, we excluded 932,143 SNPs with MAF <0.05 and 23 SNPs with HWE p-value $< 10^{-5}$, resulting in 550,405 high-quality SNPs.

We performed imputation using the Haplotype Reference Consortium (HRC, Version r1.1 2016) panel, which includes all samples from the 1,000 Genomes Project (Phase 3), using the Michigan Imputation Server. Of the 627,485 and 550,405 post-QC SNPs in AA and EA, respectively, 563,473 SNP in AA and 498,015 SNPs in EA matched the HRC panel and met the QC thresholds of the Michigan Imputation Server. A total of ~39 million SNPs were imputed in each group. In AA, we excluded ~23M SNPs with imputation accuracy (r^2) ≤0.3 and ~9.5M SNPs with MAF ≤0.07, resulting in 6,463,658 SNPs. In EA, we removed ~29.3M SNPs with r^2 ≤0.3 and ~4.9M SNPs with MAF ≤0.07, resulting in 4,900,500 SNPs. We conducted downstream analyses on the resulting 6.4M SNPs for AAs and 4.9M SNPs for EAs.

DNA methylation

The Illumina Infinium MethylationEPIC array was used to interrogate >850,000 CpG sites at the University of Chicago Genomics Core Facility. This array provides dense coverage across CpG islands (>95%), shores (>80%), and shelves (>90%). Methylation data was normalized using the BMIQ function in the ChAMP software, and methylation at each CpG was expressed as β values ranging from 0 (completed unmethylated) to 1 (completed methylated). We processed the methylation data for EA and AA groups combined, but stratified by tissue type. We removed probes with a detection p-value >0.01 (95,481), beadcount <3 in at least 5% of samples (106), non-CpG probes (2,228), non-specific probes that align to multiple locations (47), and underperforming probes (e.g., low mappability to hg38, unrecognized color channel switch for Type I probes, and contain SNPs close to the 3’ end of probe) (69,244) (13). This QC resulted in 698,812 CpG sites for benign tissue and 682,694 in tumor tissue, and these CpGs were included in all downstream analyses.

We characterized differences in DNAm between tumor and benign samples by applying principal components analysis (PCA) to all CpGs for all 296 samples. PCA demonstrated clear separation of most tumor samples from most benign samples (Supplementary Figure S1). Within tumor samples, PC1 was also associated with Gleason score (Supplementary Figure S1). These results demonstrate that tumor/benign status is the largest source of variation in our DNAm data, and our tumor samples vary with respect to the composition/abundance of tumor vs. benign cells.

Cis-meQTL Analyses

We used FastQTL to conduct genome-wide *cis*-meQTL analyses. All analyses were conducted separately for AA benign, EA benign, AA tumor, and EA tumor tissue. We tested the local (*cis*) association of SNPs and CpG sites <500 kb apart. Methylation beta

values were rank normalized to satisfy linear model assumptions. To identify CpGs affected by a meQTL, we computed CpG-level empirical p-values (i.e., the smallest P-value for each CpG) using an approximation of the beta-distribution and adaptive permutations as implemented in FastQTL (--permute 1000 10000) (14). To account for multiple testing, we used the Storey/Tibshirani false discovery rate (FDR) procedure applied at the CpG-level to identify mCpGs (implemented in the R/qvalue package). A FDR of 0.05 was applied in each of the four analyses.

All regression models were adjusted for age, 5 genotyping principal components (PCs), and 10 methylation surrogate variables (SVs), in order to capture variability in cell type composition (including tumor purity) as well as potential technical variation.

Methylation SVs were estimated using the SVA package (15). To determine the number of SVs that maximized power for meQTL detection, we conducted cis-meQTL analysis of chromosome 1 for each ancestry and tissue type using 5, 10, 15, and 20 SVs.

PC analyses were conducted (separately for AA and EA) using a genome-wide set of independent SNPs in PLINK (--indep-pairwise 50 5 0.2). An additional PC analysis was conducted for AA and EA patients combined to demonstrate that and AA and EA individuals clustered separately, as expected (Supplementary Figure S2).

Identifying GWAS signals and meQTLs likely to share a causal variant

An overview of our workflow (and results) is shown in Figure 1. We first identified meQTLs for benign (6,298 for AA and 6,960 for EA) and tumor tissue (2,641 for AA and 1,700 for EA) and identified the lead SNP for each meQTL. Next, we restricted to the 20,646 SNPs from PCa GWAS that met a threshold of $p < 5 \times 10^{-8}$ (from Schumacher et al.) and identified SNPs that were also a lead meQTL SNP, resulting in 37 AA and 40 EA co-localization candidates for benign tissue and 38 AA and 33 EA candidates for tumor tissue. We conducted co-localization tests (as described in the sections below) at loci where the lead GWAS SNP was in linkage disequilibrium (LD, $r^2 > 0.5$) with the lead meQTL SNP in at least one ancestry (8 for AA and 17 for EA in benign, 3 for AA and 14 for EA in tumor), according to LDlink (https://ldlink.nci.nih.gov/). For AA men we used the Americans of African Ancestry in the USA (ASW) reference population, and for EA men we used the European reference populations (EUR). We conducted all GWAS-meQTL co-localization analyses stratified by ancestry and tissue type, and identified 1 (AA) and 7 (EA) co-localizations for benign meQTLs and 3 (AA) and 5 (EA) co-localization for tumor meQTLs (as described in the results section). For the meQTLs identified, mCpG enrichment in genomic features was assessed using X^2 tests.

GWAS-meQTL co-localization

We used summary statistics from a prior PCa GWAS (7) and our *cis*-meQTL results to conduct Bayesian co-localization analysis (using *coloc* (16)). This approach restricts to SNPs that are present in both sets of summary statistics within a given start and end positions (basepairs) for the region under analysis.

Coloc requires three prior probabilities: the probability a variant is causal for PCa only (p_1), a QTL only (p_2), and both PCa and a QTL (p_{12}). The software uses these priors to calculate the posterior probability of a common causal variant (P(CCV)). To select a prior for PCa, we used a recent estimate of the number of independent common PCa susceptibility variants (4,530) (17). Given the ~20M SNPs tested in recent GWAS, the probability a SNP is causal for PCa is 4,530/20M, approximately 10^{-4}. This probability is equal to $p_1 + p_{12}$. To set meQTL priors, we used the previously detected 7,590 *cis*-meQTLs among 4,894,225 SNPs tested (11), indicating the probability a SNP is a causal meSNP in prostate tissue is approximately 5×10^{-3}. This probability corresponds to $p_2 + p_{12}$ (for GWAS-meQTL co-localization). Selection of these priors was informed by the literature but not intended to be exact. Because the true value of p_{12} is unknown, we varied the value of p_{12} to correspond to probabilities of a causal PCa SNP being a causal meSNP of 10%, 25%, 50%, and 75%, similar to prior meQTL studies (18,19). The resulting values for p_1, p_2, and p_{12} are shown in Supplementary Table S1.

Identification of eQTLs among co-localized meSNPs

We searched for eQTLs among the GWAS-meQTLs (meSNPs) that reached the co-localization threshold of P(CCV) > 80% using the Genotype-Tissue Expression Project (GTEx v8) (20) prostate tissue eQTL results (n=221). To prepare the eQTL and GWAS summary statistics for coloc, we restricted to the common SNPs and the start/end genomic positions defined by the meQTL analysis. Genomic coordinates for GTEx SNPs were converted from GRCh38/hg38 to GRCh37/h19 using the NCBI Genome Remapping Service.

GWAS-eQTL co-localization

GWAS-eQTL co-localization was performed only for loci with strong evidence of GWAS-meQTL co-localization (P(CCV) of >80%). We conducted analyses for GWAS-eQTL pairs using *coloc*. GTEx identified 7,356 eQTLs in prostate tissue among 11.5M SNPs, suggesting the probability that a SNP is a causal eSNP in prostate tissue is ~6×10^{-4}. This probability corresponds to $p_2 + p_{12}$ (for GWAS-eQTL co-localization). We also varied the value of p_{12} to correspond to probabilities of a causal GWAS SNP being a causal eSNP of 10%, 25%, 50%, and 75% (Supplementary Table S1).

Data Availability

The raw data for this study were generated at University of Chicago, and these data are publicly available through dbGaP (dbGaP Study Accession: phs003516.v1.p1). All meQTL summary statistics are available for download from Zenodo (DOI: 10.5281/zenodo.10304061 and 10.5281/zenodo.10358156)

RESULTS

Overview of samples

Characteristics of the 74 AA and 74 EA patients included in our analyses are described in Supplementary Table S2. AA and EA patients were on average 65 years-old at diagnosis. Gleason score and tumor volume (defined as the percent of prostate that is tumor) were

slightly higher in EAs compared to AAs, but PSA was similar. The slightly less favorable clinical characteristics of EA compared to AA patients may be due to EA men with more advanced PCa residing outside the UCMC catchment area seeking care at the University of Chicago, with our AA patients being more likely to reside in the catchment area.

Cis-meQTLs in benign tissue

We identified a cis-meQTL for 6,298 and 6,960 CpGs (FDR 0.05) in benign prostate tissue of AA and EA men, respectively (Table 1). The meQTLs detected were represented by 5,855 unique lead SNPs in AA and 6,496 unique lead SNPs in EA, as some mCpGs had the same lead SNP. The CpGs impacted by meQTLs (i.e., mCpGs) were enriched in non-CpG islands (open sea) and depleted in islands compared to all measured CpGs (Supplementary Figure S3A). 64% (4,060) and 63% (4,409) of mCpGs in AA and EA were assigned to a gene (based on Illumina annotations) as compared to 73% of all measured CpGs ($p<0.001$) (Supplementary Figure S4A). mCpGs were enriched in enhancer regions ($p <0.001$) for both AAs (4%) and EAs (4%) compared to all measured CpGs (3%) and depleted in promoters (Supplementary Figure S4A). In AAs, mCpGs were enriched in DNase hypersensitivity sites (DHS) regions (67%) compared to all CpGs measured (60%) ($p <0.001$) (Supplementary Figure S4A).

Of the 6,298 cis-meQTLs detected in AA benign tissue, 4,269 (68%) had $p <0.01$ and 3,865 (61%) had $p <0.001$ in EA benign tissue. The correlation (r) between the beta coefficients from AA and EA for the 4,269 cis-meQTLs was 0.97, and 4,258 (68%) were directionally consistent and considered replicated (Supplementary Figure S5A). Similarly, among the 6,960 cis-meQTLs detected in EA benign tissue, 4,372 (63%) had $p <0.01$ and 3,576 (51%) had $p <0.001$ in AA benign tissue. The correlation between the beta coefficients for the 4,372 cis-meQTLs was 0.95, and 4,343 (62%) were directionally consistent and considered replicated (Supplementary Figure S5B).

Cis-meQTLs in tumor tissue

In tumor tissue we identified 2,641 (AA) and 1,700 (EA) *cis*-meQTLs (FDR < 0.05) (Table 1). Tumor mCpGs showed somewhat weaker depletion in islands compared to benign (Supplementary Figure S3B). In tumor, 69% (AA) and 67% (EA) of mCpGs were assigned to genes compared 73% of all measured CpGs ($p <0.001$) (Supplementary Figure S4B). Tumor mCpGs were depleted in promoters (13%) and enriched in DHS regions (67%) compared to all measured CpGs (20% and 61%, respectively, both $p <0.001$).

Among the 2,641 cis-meQTLs (FDR <0.05) detected in AA tumor tissue, 1,667 (63%) had $p <0.01$ and 1,457 (55%) had $p <0.001$ in EA tumor tissue. The correlation among the beta coefficients from AA and for the 1,667 cis-meQTLs was 0.96, and 1,659 (63%) were directionally consistent and considered replicated (Supplementary Figure S5C). Among the 1,700 cis-meQTLs (FDR <0.05) detected in EA tumor tissue, we found 1,131 (67%) at $p <0.01$ and 996 (59%) at $p <0.001$ in AA tumor tissue. The correlation between the beta coefficients in EA vs AA was 0.96 among the 1,131 cis-meQTLs and 1,126 (43%) were replicated with directional consistency (Supplementary Figure S5D).

Tissue specificity of cis-meQTLs in AA and EA men

Among the 6,298 benign-tissue cis-meQTLs detected in AAs, we replicated 62% ($p < 0.01$) in AA tumor tissue (Supplementary Table S3). Similarly, for EA, we replicated 55% benign-tissue cis-meQTLs in tumor tissue ($p < 0.01$). We observed a larger percent of replication of tumor-tissue cis-meQTLs in benign tissue (79% at $p < 0.01$ for both AA and EA) (Supplementary Table S3).

Replication of meQTLs from prior studies

We attempted replication of the 7,590 cis-meQTLs identified previously among 589 localized prostate tumors of EA men (11) (Supplementary Table S4). In AA benign and tumor tissue, 48% and 44% of the 5,231 meQTLs for which we had data were replicated ($p<0.01$) with directionally consistency. In EA benign and tumor tissue, we observed more replication of specific CpG-SNP pairs compared to AAs: 59% and 49% of the 5,590 previously reported meQTLs were replicated ($p < 0.01$) with directional consistency. (Supplementary Table S4).

In addition to genome-wide meQTLs, Houlahan et al. (11) report 75 PCa-risk cis-meQTLs (52 validated in an independent cohort), which describe the association between 27 unique PCa-risk loci and 73 CpG sites. We attempted replication for 41 and 45 of the 52 validated cis-meQTLs in AA and EA tissues, respectively (Supplementary Table S5). In AA benign and tumor tissue samples, we replicated 24% and 41% of the 41 meQTLs ($P < 0.01$ with directional consistency), respectively. In EA benign and tumor tissue samples, we replicated 29% and 53% of the 45 meQTLs ($p < 0.01$ with directional consistency), respectively.

We also sought to replicate 110 *cis*-meQTLs identified in the prostate tumor tissues of 355 EA men (12) (Supplementary Table S6). Of the 110 *cis*-meQTLs, we had summary results for 60 *cis*-meQTLs in AAs and 67 *cis*-meQTLs in EAs. We replicated 15% and 42% of the meQTLs ($P < 0.01$ with directional consistency) in AA benign and tumor tissues, and 28% and 49% of the meQTLs in EA benign and tumor tissues.

Co-localized GWAS-meQTL pairs in benign tissue

Among our 6,298 (AA) and 6,960 (EA) meQTLs identified in benign tissue, we identified *cis*-meQTLs residing in the same location as PCa-risk loci (Figure 1 and Supplementary Tables S7–S8). Restricting to lead SNPs with $p < 5\times10^{-8}$ in the Schumacher et al. PCa GWAS, we searched for meQTL lead SNPs with an LD ($r^2 > 0.5$) with a GWAS lead SNP in AA and/or EA. We identified 20 such SNP-CpG pairs (Supplementary Table S9).

We found evidence of co-localization (P(CCV) >80%) (based on 50% prior probability that a GWAS SNP is an meSNP, see Supplementary Table S1) for 2 GWAS-meQTL pairs in AA men and 4 pairs in EA men (Table 2 and Supplementary Table S10 for full results). The co-localized signals detected in AA men were in an intergenic region near *MLPH* and in a promoter of *HAUS6* (Figure 2a-b). The signal in *HAUS6* was also suggestive in the EA cohort. The four co-localized signals detected in EA men were located in the gene body of *TNS3,* promoter near *MSMB*, promoter near *MRPL52/MMP14,* and intergenic region of *COPRS/UTP6* (Figure 2c-f). The signals in *TNS3* and *COPRS* are also

suggestive in AA and highlight the potential impact of cross-population LD differences on co-localization results. The posterior probabilities that the GWAS and meQTL signal are caused by independent causal variants are shown in Supplementary Table S11.

Co-localized GWAS-meQTL pairs in tumor tissue

In tumor tissue, we identified *cis*-meQTLs located at previously reported PCa-risk loci ($p < 5\times10^{-8}$) in 38 regions (for AA) and 33 regions (for EA), respectively (Figure 1 and Supplementary Table S12–S13). Restricting to loci with LD (r^2) > 0.5 between the lead GWAS and lead meQTL SNP in at least one ancestry, we identified 16 loci to test for co-localization (Supplementary Table S14).

There was evidence of co-localization (based on 50% prior probability that a GWAS SNP is a meSNP, Supplementary Table S1) for 3 GWAS-meQTL pairs in both AAs and EAs (Table 3 and Supplementary Table S15 for full results). Two co-localization signals were shared between AAs and EAs; one near the gene body of *IRX4* and the other in the 5'UTR of *MMP7* (Figure 3a-b). In EA, PCa-risk SNP rs12653946 was associated with 9 CpGs in the *IRX4* region (Supplementary Figure S6), most showing strong co-localization evidence, including several promoter CpGs. In AA men, evidence of co-localization was observed for a CpG near the gene body of *MYO9B* (Figure 3c). This region showed suggestive evidence of co-localization in EA (Table 3). One additional co-localized signal identified in EA was located in the gene body of *TNS3* (Figure 3d), a signal also observed in EA benign tissue and suggestive AA benign and tumor tissue. The posterior probabilities that the GWAS and meQTL signals are caused by independent causal variants are shown in Supplementary Table S16.

Co-localized GWAS-eQTL pairs

Among the 9 co-localized GWAS-meQTL pairs across both tissue types and ancestries, we identified eQTLs for four PCa-risk SNPs (rs12653946, rs11568818, rs11666569, and rs10993994) in GTEx normal prostate tissue. We conducted co-localization analyses for the four PCa-risk SNPs and the corresponding eGenes (*IRX4, MMP7, MYO9B,* and *MSMB,* respectively). Using the prior sets listed in Supplementary Table S1, we found strong evidence of shared common causal variants affecting both GWAS and eQTL traits for all four SNP-eGenes tested (Supplementary Table S17 and Figure S7), suggesting these PCa risk variants regulate both local methylation and expression.

DISCUSSION

We performed a genome-wide search for *cis*-meQTLs in both benign and cancerous prostate tissue of AA and EA men. We identified 6,298 and 6,960 *cis*-meQTLs in the benign tissue of AA and EA men, respectively, and 2,641 and 1,700 *cis*-meQTLs in the tumor tissue of AA and EA men, respectively. To determine if known PCa susceptibility loci (7) impact local DNA methylation in the prostate, we used Bayesian co-localization methods to identify 9 regions in which PCa risk loci likely share a causal variant with a meQTL in either cancer and/or benign tissue. Four of the nine regions showed evidence of a co-localizing eQTL (*IRX4, MMP7, MYO9B,* and *MSMB*), all of which have roles in tumorigenesis (7,12,21–

23). Our findings highlight potential regulatory mechanisms by which PCa-risk variants impact gene regulation in prostate tissue.

We identified 6 regions with strong evidence of GWAS-meQTL co-localization in benign tissue (2 in AA,4 in EA) and 4 in tumor tissue (2 in both AA and EA, 1 in AA, and 1 in EA). According to QTLbase (24) and GTEx, only three of our 9 co-localized meSNPs are associated with methylation and/or expression in whole blood (rs2292884-cg14458575: eGene = *MLPH*, rs10993994-cg17030820: eGene = *MSMB*, rs11666569- cg19418318: eGene = MYO9B) (Supplementary Table S18), indicating that some PCa susceptibility SNPs may have prostate-specific regulatory effects.

Interestingly, we observed co-localization for the *MLPH* region (for AA benign tissue) despite distinct differences in LD between AA (meQTL) and EA (GWAS) in this region (Figure 2A). The lead meQTL SNP for AA (rs72620822) was among the top GWAS SNPs (Supplementary Figure S8), reflecting LD between these SNPs in populations of European ancestry (r^2= 0.65 in EUR). However, the lead GWAS SNP (rs2292884) is not among the top AA meQTL SNPs (Figure 2A), reflecting lack of LD between these SNPs in populations of African ancestry (r^2= 0.002 in AFR). Evidence for co-localization for the EA meQTL in the *MLPH* region is much weaker. This this locus requires further investigation in future studies.

Our results suggest that the causal PCa risk variant represented by rs12653946 affects methylation in tumor tissue at (at least) 9 CpGs near *IRX4* in EA men (Supplementary Table S14, Supplementary Figure S6). Six of these CpGs are in a CpG island near the IRX4 start site, and one is in an enhancer (Supplementary Table S14). The rs12653946 risk allele (T) was associated with increased methylation at all 7 CpGs within islands (Supplementary Table S14). The meQTL at *IRX4* appears substantially weaker (or absent) in benign tissue ($p > 0.001$ for all 9 CpGs) compared to tumor (Supplementary Table S19 and Supplementary Figure S9). In GTEx normal prostate tissue, the rs12653946 risk allele (T) is associated with decreased gene expression of IRX4 (Supplementary Table S17), a previously reported tumor suppressor gene (25). Additional PCa risk variants showing clear association with both DNAm and gene expression (in GTEx prostate tissue) include rs11568818 (MMP7), rs11666569 (MYO9B), and rs10993994 (MSMB). For MMP7 and MYO9B, the risk allele decreases DNAm (in 5` UTR) and increases gene expression. For MSMB, the risk allele increases DNAm (near the TSS) and decreases MSMB expression.

A primary challenge of conducting co-localization using meQTL results from AA men is the LD mismatch with the largely European ancestry GWAS of PCa, which decreases power to detect co-localization using meQTL results from AAs. For example, in AA benign and tumor tissue, we identified a meQTL (cg23694490- rs834603) near *TNS3* but the lead meQTL SNP was in low LD (r^2 = 0.15) with the lead GWAS SNP (rs56232506), while in EA men these SNPs were in strong LD (r^2 = 0.86) and co-localization was observed. Thus, evidence of co-localization was lower for AA likely due LD differences in this region (Figure 2C). This LD discordance also likely affects the generalizability of associations reported in prior PCa GWAS, as less than half of PCa risk SNPs (identified in studies of largely European ancestry men) have been replicated in men of African ancestry (26). These

examples highlight the need for large PCa GWAS focusing on AAs. Such results would improve our ability to describe the functional impacts of PCa risk alleles using AA QTL results.

We detected nearly triple the number of cis-meQTLs in benign tissue compared to tumor tissue (in both AA and EA), a finding likely attributable to differences in cell type composition. Benign tissue samples may be more homogeneous, with patient samples consisting largely of epithelial and stromal cells. In contrast, tumor tissue samples will contain cancer cells, which can differ in various cellular phenotypes across individuals (and within individuals), as well as adjacent normal (epithelial and stromal cells). Assuming meQTL patterns differ to some extent across cell types, power to detect cell type-specific QTLs will be higher in more homogeneous cell type mixtures. Despite these differences in cell type, we found that 79% of tumor-tissue meQTLs were likely present in benign tissue ($p < 0.01$) but observed less replication of benign-tissue meQTLs in tumor tissue (62% in AAs, 55% in EAs).

Several of the genes/regions identified in this work have established roles in PCa. For example, studies in MMP7 knockout mice show decreased proliferation, increased apoptosis, and inhibition of angiogenesis and epithelial-to-mesenchymal transition, leading to decreased tumor burden (22). In addition, increased levels of serum MMP7 were recently reported to associate with reduced overall survival in castration-resistant PCa. (27) IRX4 is a tumor suppressor gene (25), and has been identified as a target of an eQTL associated with PCa risk (21), with alternatively spliced transcripts potentially playing an important role in PCa prognosis (28). MSMB has been previously identified as a target of a PCa risk eQTL, with downstream/trans effects on SNGH11 (23). MSMB expression is reduced in tumors and adjacent benign prostate tissue, potentially explaining why serum MSMB are decreased in PCa patients (29).

Our study has several strengths compared to prior work. First, we employed co-localization tests, which are critical for determining if causal variants for PCa risk also affect DNAm (and are not simply associated with DNAm due to LD with a nearby meSNP). While our work appears to suggest that the number of PCa risk loci potentially explained by meQTLs is smaller than previously reported, we are underpowered to detect weaker meQTL effects, and sample size likely prevented detection of co-localization for some meQTLs, such as in the EA tumor tissue meQTL (cg02493740 - rs10187424) near GGCX (Supplementary Table S15, Supplementary Figure S10). Second, we analyzed equal numbers of AA and EA samples, emphasizing detection of meQTLs in patients of African ancestry. Third, we conducted meQTL analyses using paired tumor and benign samples; however, it is possible that these benign samples may contain cancer-related molecular characteristics and are not representative of prostate tissue from PCa-free men. Lastly, while we did not generate gene expression data, we leveraged eQTL results from GTEx normal prostate tissue.

Our study utilized FFPE tissues as a DNA source, and we used a qPCR-based QC procedure to ensure the quality of DNA samples were appropriate for restoration using the Illumina Infinium HD FFPE restoration kit. Prior studies show high concordance of EPIC array DNA methylation measures obtained from FFPE samples (repaired using the Illumina kit)

compared to those from fresh frozen tissue from the same individuals, reporting correlations >90% (30–32). However, the potential for decreased quality of FFPE samples, and increased noise in the DNAm data generated, may reduce power for meQTL detection compared to frozen tissue.

We adjusted for global ancestry in our meQTL analyses, but not local ancestry. It has recently been shown that local ancestry adjustment can improve power for QTL detection in admixed samples (33), but power gains are likely to be modest (34).

In summary, we conducted a comprehensive search for meQTLs in tumor and paired benign prostate tissues of AA and EA men. We used summary statistics from prior GWAS to identify PCa susceptibility loci whose biological mechanism likely involves alteration of local DNAm. These results are a resource to explore differences in prostate meQTL profiles of AA and EA men. To better understand the biological mechanisms by which PCa susceptibility SNPs influence PCa biology, larger and more diverse studies and meta-analyses of prostate meQTLs and eQTLs are needed, as are larger PCa GWAS of AA men.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

ACKNOWLEDGMENTS

Funding:

U.S Department of Defense CDMRP Health Disparity Research Award (W81XWH-14-1-0529 to B.L.P.), National Institute of Environmental Health Sciences award (R35ES028379 to B.L.P. and F30ES031858 to M.C), National Institute of General Medicine (T32GM150375 and T32GM007281 to Marcus Ramsey Clark), National Cancer Institute (P30CA14599 to Kunle Odunsi), Susan G. Komen Research Training Grant (GTDR16376189 to Eileen Dolan), and the National Institute of Aging (T32AG51146 to David O. Meltzer).

This work was supported by the Epidemiology Research and Recruitment Core, the Human Tissue Research Core, and the Genomics Core Facility at the University of Chicago. This publication was neither originated nor managed by AbbVie, and it does not communicate results of AbbVie-sponsored Scientific Research. Thus, it is not in scope of the AbbVie Publication Procedure (PUB-100).

References

1. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2020. CA: a cancer journal for clinicians 2020;70(1):7–30 doi 10.3322/caac.21590. [PubMed: 31912902]
2. Torre LA, Bray F, Siegel RL, Ferlay J, Lortet-Tieulent J, Jemal A. Global cancer statistics, 2012. CA: a cancer journal for clinicians 2015;65(2):87–108 doi 10.3322/caac.21262. [PubMed: 25651787]
3. Rawla P Epidemiology of Prostate Cancer. World journal of oncology 2019;10(2):63–89 doi 10.14740/wjon1191. [PubMed: 31068988]
4. Pietro GD, Chornokur G, Kumar NB, Davis C, Park JY. Racial Differences in the Diagnosis and Treatment of Prostate Cancer. International neurourology journal 2016;20(Suppl 2):S112–9 doi 10.5213/inj.1632722.361. [PubMed: 27915474]
5. DeSantis CE, Siegel RL, Sauer AG, Miller KD, Fedewa SA, Alcaraz KI, et al. Cancer statistics for African Americans, 2016: Progress and opportunities in reducing racial disparities. CA: a cancer journal for clinicians 2016;66(4):290–308 doi 10.3322/caac.21340. [PubMed: 26910411]

6. Kelly SP, Rosenberg PS, Anderson WF, Andreotti G, Younes N, Cleary SD, et al. Trends in the Incidence of Fatal Prostate Cancer in the United States by Race. European urology 2017;71(2):195–201 doi 10.1016/j.eururo.2016.05.011. [PubMed: 27476048]
7. Schumacher FR, Al Olama AA, Berndt SI, Benlloch S, Ahmed M, Saunders EJ, et al. Association analyses of more than 140,000 men identify 63 new prostate cancer susceptibility loci. Nature Genetics 2018;50(7):928–36 doi 10.1038/s41588-018-0142-8. [PubMed: 29892016]
8. Conti DV, Darst BF, Moss LC, Saunders EJ, Sheng X, Chou A, et al. Trans-ancestry genome-wide association meta-analysis of prostate cancer identifies new susceptibility loci and informs genetic risk prediction. Nature Genetics 2021;53(1):65–75 doi 10.1038/s41588-020-00748-0. [PubMed: 33398198]
9. Do C, Shearer A, Suzuki M, Terry MB, Gelernter J, Greally JM, et al. Genetic–epigenetic interactions in cis: a major focus in the post-GWAS era. Genome Biology 2017;18(1):120 doi 10.1186/s13059-017-1250-y. [PubMed: 28629478]
10. Albert FW, Kruglyak L. The role of regulatory variation in complex traits and disease. Nature Reviews Genetics 2015;16(4):197–212 doi 10.1038/nrg3891.
11. Houlahan KE, Shiah Y-J, Gusev A, Yuan J, Ahmed M, Shetty A, et al. Genome-wide germline correlates of the epigenetic landscape of prostate cancer. Nature Medicine 2019;25(10):1615–26 doi 10.1038/s41591-019-0579-z.
12. Dai JY, Wang X, Wang B, Sun W, Jordahl KM, Kolb S, et al. DNA methylation and cis-regulation of gene expression by prostate cancer risk SNPs. PLOS Genetics 2020;16(3):e1008667 doi 10.1371/journal.pgen.1008667. [PubMed: 32226005]
13. Zhou W, Laird PW, Shen H. Comprehensive characterization, annotation and innovative use of Infinium DNA methylation BeadChip probes. Nucleic Acids Research 2016;45(4):e22–e doi 10.1093/nar/gkw967.
14. Aguet F, Brown AA, Castel SE, Davis JR, He Y, Jo B, et al. Genetic effects on gene expression across human tissues. Nature 2017;550(7675):204–13 doi 10.1038/nature24277. [PubMed: 29022597]
15. Leek JT, Johnson WE, Parker HS, Jaffe AE, Storey JD. The sva package for removing batch effects and other unwanted variation in high-throughput experiments. Bioinformatics 2012;28(6):882–3 doi 10.1093/bioinformatics/bts034. [PubMed: 22257669]
16. Giambartolomei C, Vukcevic D, Schadt EE, Franke L, Hingorani AD, Wallace C, et al. Bayesian Test for Colocalisation between Pairs of Genetic Association Studies Using Summary Statistics. PLOS Genetics 2014;10(5):e1004383 doi 10.1371/journal.pgen.1004383. [PubMed: 24830394]
17. Zhang YD, Hurson AN, Zhang H, Choudhury PP, Easton DF, Milne RL, et al. Assessment of polygenic architecture and risk prediction based on common variants across fourteen cancers. Nature Communications 2020;11(1):3353 doi 10.1038/s41467-020-16483-3.
18. Pierce BL, Tong L, Argos M, Demanelis K, Jasmine F, Rakibuz-Zaman M, et al. Co-occurring expression and methylation QTLs allow detection of common causal variants and shared biological mechanisms. Nature Communications 2018;9(1):804 doi 10.1038/s41467-018-03209-9.
19. Wallace C Eliciting priors and relaxing the single causal variant assumption in colocalisation analyses. PLOS Genetics 2020;16(4):e1008720 doi 10.1371/journal.pgen.1008720. [PubMed: 32310995]
20. Consortium GTEx. The GTEx Consortium atlas of genetic regulatory effects across human tissues. Science 2020;369(6509):1318–30 doi 10.1126/science.aaz1776. [PubMed: 32913098]
21. Xu X, Hussain WM, Vijai J, Offit K, Rubin MA, Demichelis F, et al. Variants at IRX4 as prostate cancer expression quantitative trait loci. European journal of human genetics : EJHG 2014;22(4):558–63 doi 10.1038/ejhg.2013.195. [PubMed: 24022300]
22. Zhang Q, Liu S, Parajuli KR, Zhang W, Zhang K, Mo Z, et al. Interleukin-17 promotes prostate cancer via MMP7-induced epithelial-to-mesenchymal transition. Oncogene 2017;36(5):687–99 doi 10.1038/onc.2016.240. [PubMed: 27375020]
23. Bicak M, Wang X, Gao X, Xu X, Väänänen R-M, Taimen P, et al. Prostate cancer risk SNP rs10993994 is a trans-eQTL for SNHG11 mediated through MSMB. Human Molecular Genetics 2020;29(10):1581–91 doi 10.1093/hmg/ddaa026. [PubMed: 32065238]

24. Zheng Z, Huang D, Wang J, Zhao K, Zhou Y, Guo Z, et al. QTLbase: an integrative resource for quantitative trait loci across multiple human molecular phenotypes. Nucleic Acids Res 2020;48(D1):D983–D91 doi 10.1093/nar/gkz888. [PubMed: 31598699]
25. Nguyen HH, Takata R, Akamatsu S, Shigemizu D, Tsunoda T, Furihata M, et al. IRX4 at 5p15 suppresses prostate cancer growth through the interaction with vitamin D receptor, conferring prostate cancer susceptibility. Hum Mol Genet 2012;21(9):2076–85 doi 10.1093/hmg/dds025. [PubMed: 22323358]
26. Lachance J, Berens AJ, Hansen MEB, Teng AK, Tishkoff SA, Rebbeck TR. Genetic Hitchhiking and Population Bottlenecks Contribute to Prostate Cancer Disparities in Men of African Descent. Cancer research 2018;78(9):2432–43 doi 10.1158/0008-5472.Can-17-1550. [PubMed: 29438991]
27. Szarvas T, Csizmarik A, Varadi M, Fazekas T, Huttl A, Nyirady P, et al. The prognostic value of serum MMP-7 levels in prostate cancer patients who received docetaxel, abiraterone, or enzalutamide therapy. Urol Oncol 2021;39(5):296 e11–e19 doi 10.1016/j.urolonc.2020.09.005.
28. Fernando A, Liyanage C, Moradi A, Janaththani P, Batra J. Identification and Characterization of Alternatively Spliced Transcript Isoforms of IRX4 in Prostate Cancer. Genes (Basel) 2021;12(5) doi 10.3390/genes12050615.
29. Bergstrom SH, Jaremo H, Nilsson M, Adamo HH, Bergh A. Prostate tumors downregulate microseminoprotein-beta (MSMB) in the surrounding benign prostate epithelium and this response is associated with tumor aggressiveness. Prostate 2018;78(4):257–65 doi 10.1002/pros.23466. [PubMed: 29250809]
30. Oliveira D, Hentze J, O'Rourke CJ, Andersen JB, Hogdall C, Hogdall EV. DNA Methylation in Ovarian Tumors-a Comparison Between Fresh Tissue and FFPE Samples. Reprod Sci 2021;28(11):3212–8 doi 10.1007/s43032-021-00589-0. [PubMed: 33891290]
31. Kling T, Wenger A, Beck S, Caren H. Validation of the MethylationEPIC BeadChip for fresh-frozen and formalin-fixed paraffin-embedded tumours. Clin Epigenetics 2017;9:33 doi 10.1186/s13148-017-0333-7. [PubMed: 28392843]
32. de Ruijter TC, de Hoon JP, Slaats J, de Vries B, Janssen MJ, van Wezel T, et al. Formalin-fixed, paraffin-embedded (FFPE) tissue epigenomics using Infinium HumanMethylation450 BeadChip assays. Lab Invest 2015;95(7):833–42 doi 10.1038/labinvest.2015.53. [PubMed: 25867767]
33. Zhong Y, Perera MA, Gamazon ER. On Using Local Ancestry to Characterize the Genetic Architecture of Human Traits: Genetic Regulation of Gene Expression in Multiethnic or Admixed Populations. Am J Hum Genet 2019;104(6):1097–115 doi 10.1016/j.ajhg.2019.04.009. [PubMed: 31104770]
34. Gay NR, Gloudemans M, Antonio ML, Abell NS, Balliu B, Park Y, et al. Impact of admixture and ancestry on eQTL analysis and GWAS colocalization in GTEx. Genome Biol 2020;21(1):233 doi 10.1186/s13059-020-02113-0. [PubMed: 32912333]

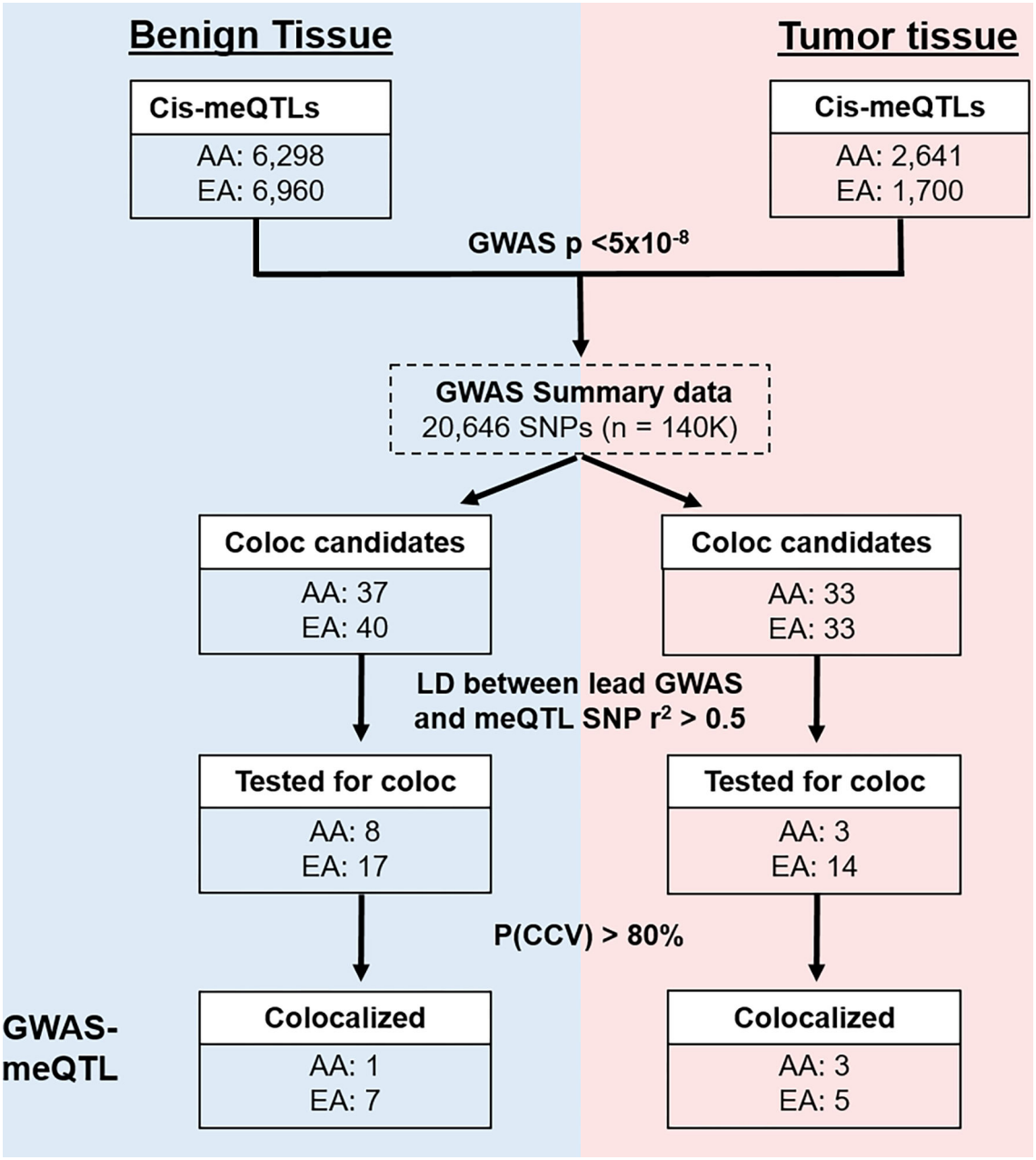

Figure 1. Summary of cis-meQTLs identified and tested for co-localization prostate cancer GWAS signals.

Results are reported for benign prostate tissue (left, blue) and prostate tumor tissue (right, red) for patients of both African Ancestry (AA) and European Ancestry (EA).

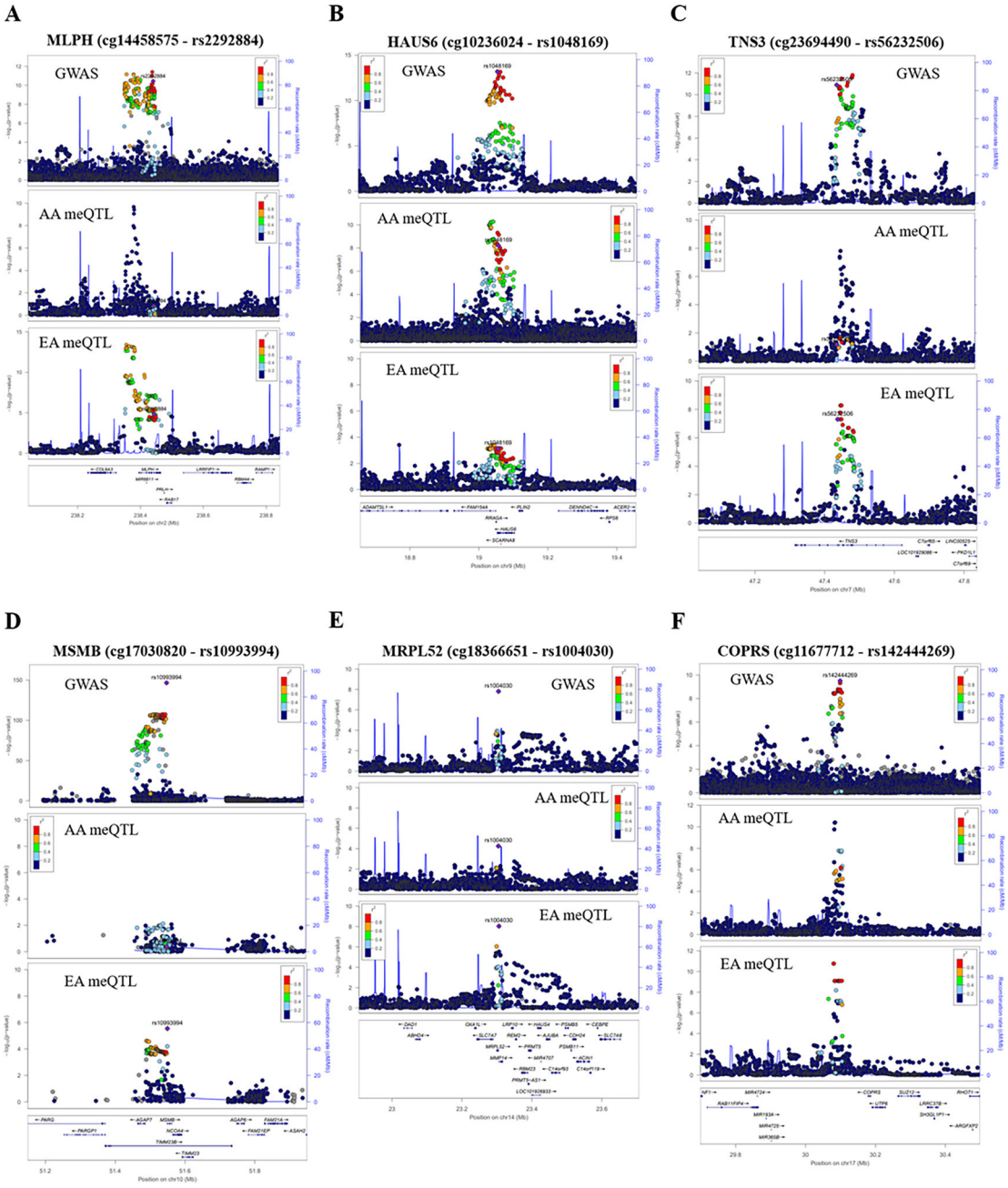

Figure 2. Examples of co-localization between meQTLs in benign prostate tissue and prostate cancer GWAS signals.

Co-localized signals detected in AA men were near MLPH (A) and HAUS6 (B). Co-localized signals detected in EA men were in the proximity of TNS3 (C), MSMB (D), MRPL52/MMP14 (E), and COPRS/UTP6 (F). LD information for EA meQTLs and GWAS is from EUR, while LD for AA meQTLs is from AFR.

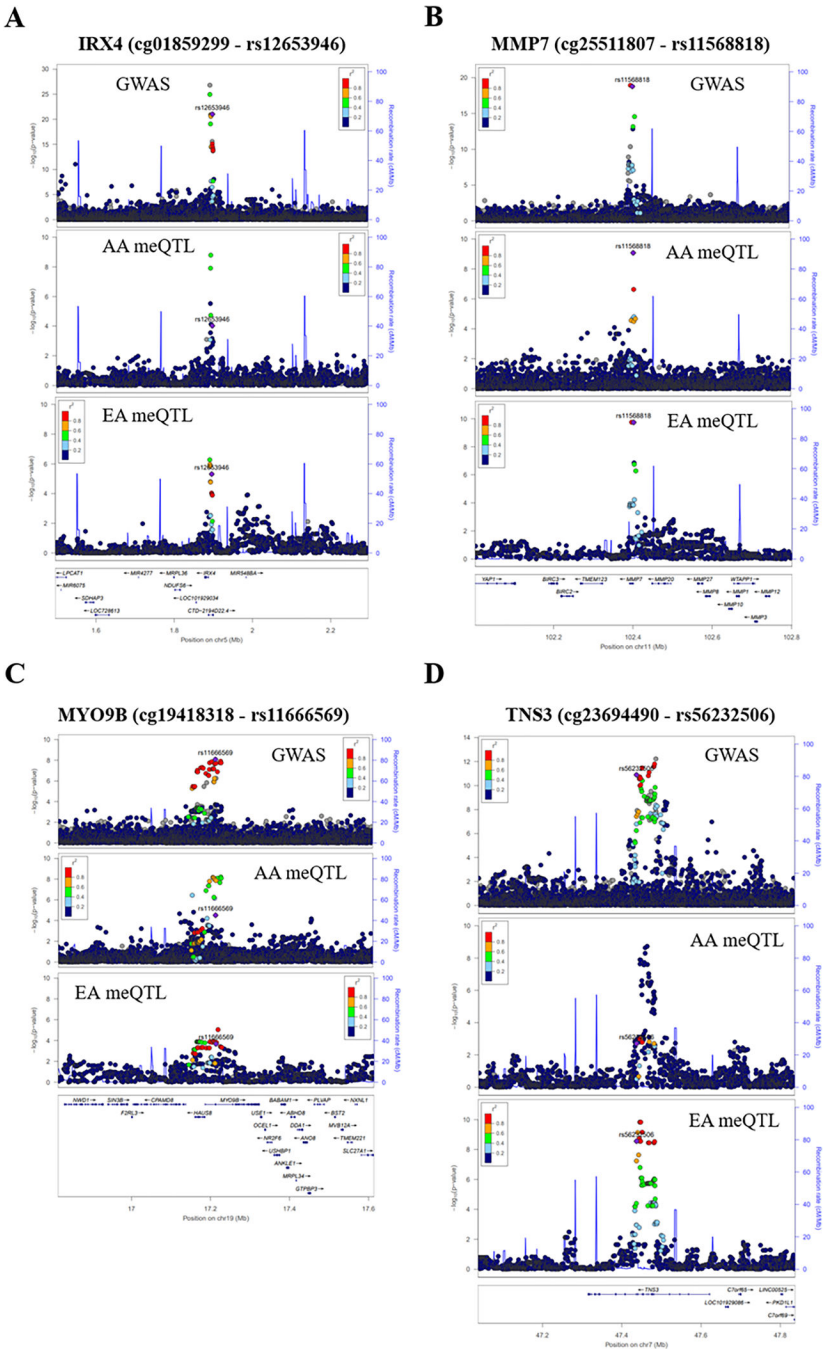

Figure 3. Examples of co-localization between meQTLs in tumor tissue and prostate cancer GWAS signals.

Co-localization signals shared between AAs and EAs were near IRX4 (A) and MMP7 (B). In AA men, co-localized signals were observed at MYO9B (C). In EA, a co-localized meQTL was located at TNS3 (D). LD information for EA meQTLs and GWAS is from EUR, while LD for AA meQTLs is from AFR.

Table 1.

Summary of genome-wide cis-meQTLs identified in analyses stratified by ancestry and tissue type

	AA *cis*-meQTL analysis (n = 74)		EA *cis*-meQTL analysis (n = 74)	
	Benign	**Cancer**	**Benign**	**Cancer**
SNPs analyzed	n = 6,463,658		n = 4,900,500	
CpG sites analyzed	698,812	682,694	698,812	682,694
mCpGs detected[a]	6,298	2,641	6,960	1,700
Unique lead SNPs	5,855	2,434	6,496	1,586
Average distance between CpG and lead SNP (bp)	23,741	25,647	30,487	28,085

[a]CpG sites affected by a meQTL, detected at a false discovery rate (FDR) of 0.05

Table 2.

Prostate cancer risk SNPs showing posterior probabilities >80% that GWAS and meQTL signals (in benign prostate tissue) share the same causal variant.

Chr	PCa-risk SNP [a]	CpG	Nearest gene	Discovery cohort [b]	P(CCV) in AA [c]		P(CCV) in EA [c]	
					PP 25%[d]	PP 50%[e]	PP 25% [d]	PP 50% [e]
2	rs2292884	cg14458575	*MLPH*	EA	69.4%	**87.2%**	42.4%	69%
7	rs56232506	cg23694490	*TNS3*	EA	27.5%	53.4%	58.7%	**81.1%**
9	rs1048169	cg10236024	*HAUS6*	AA	60.1%	**82%**	18.3%	40.3%
10	rs10993994	cg17030820	*MSMB*	EA	6.8%	18.1%	**81.2%**	**92.9%**
14	rs1004030	cg18366651	MRPL52	EA	40.5%	66.8%	**98.1%**	**99.4%**
17	rs142444269	cg11677712	*COPRS*	EA	0.5%	1.4%	77.9%	**91.4%**

[a] The PCa-risk SNP is the lead GWAS SNP reported in Schumacher et al.

[b] The discovery cohort is where the GWAS SNP and the meQTL lead SNP are in LD (r^2>0.5)

[c] Posterior probabilities >80% are shown in bold text

[d] Prior probability that 25% of GWAS SNPs are also meSNPs in prostate benign tissue

[e] Prior probability that 50% of GWAS SNPs are also meSNPs in prostate benign tissue

Table 3.

Prostate cancer risk SNPs showing posterior probabilities >80% that GWAS and meQTL signals (in prostate tumor tissue) share the same causal variant.

Chr	PCa-risk SNP[a]	CpG	Nearest gene	Discovery cohort[b]	P(CCV) in AA [c]		P(CCV) in EA [c]	
					PP 25%[d]	PP 50%[e]	PP 25% [d]	PP 50% [e]
5	rs12653946	cg01859299	*IRX4*	AA	**93.1%**	**97.6%**	77.2%	**91.1%**
5	rs12653946	cg14051264	*IRX4*	EA	66.5%	**85.7%**	**94.8%**	**98.2%**
7	rs56232506	cg23694490	*TNS3*	EA	5.4%	14.7%	69.8%	**87.5%**
11	rs11568818	cg25511807	*MMP7*	AA & EA	**96%**	**98.6%**	**98.3%**	**99.4%**
19	rs11666569	cg19418318	*MYO9B*	AA	70.5%	**87.8%**	27.7%	53.5%

[a] The PCa-risk SNP is the lead GWAS SNP reported in Schumacher et al.

[b] The discovery cohort is where the GWAS SNP and the meQTL lead SNP are in LD (r^2>0.5)

[c] Posterior probabilities >80% are shown in bold text

[d] Prior probability that 25% of GWAS SNPs are also meSNPs in prostate benign tissue

[e] Prior probability that 50% of GWAS SNPs are also meSNPs in prostate benign tissue

[f] The IRX4 region has 7 additional mCpGs showing strong evidence of co-localization, and these are shown in Supplementary Table 15.