Research Article

Cancer Epidemiology, Biomarkers & Prevention

Rare Variation in *TET2* Is Associated with Clinically Relevant Prostate Carcinoma in African Americans

Daniel C. Koboldt[1], Krishna L. Kanchi[1], Bin Gui[2], David E. Larson[1], Robert S. Fulton[1], William B. Isaacs[3], Aldi Kraja[4], Ingrid B. Borecki[4], Li Jia[2], Richard K. Wilson[1], Elaine R. Mardis[1], and Adam S. Kibel[2]

Abstract

Background: Common variants have been associated with prostate cancer risk. Unfortunately, few are reproducibly linked to aggressive disease, the phenotype of greatest clinical relevance. One possible explanation is that rare genetic variants underlie a significant proportion of the risk for aggressive disease.

Method: To identify such variants, we performed a two-stage approach using whole-exome sequencing followed by targeted sequencing of 800 genes in 652 aggressive prostate cancer patients and 752 disease-free controls in both African and European Americans. In each population, we tested rare variants for association using two gene-based aggregation tests. We established a study-wide significance threshold of 3.125×10^{-5} to correct for multiple testing.

Results: *TET2* in African Americans was associated with aggressive disease, with 24.4% of cases harboring a rare deleterious variant compared with 9.6% of controls (FET $P = 1.84 \times 10^{-5}$, OR $= 3.0$; SKAT-O $P = 2.74 \times 10^{-5}$). We report 8 additional genes with suggestive evidence of association, including the DNA repair genes *PARP2* and *MSH6*. Finally, we observed an excess of rare truncation variants in 5 genes, including the DNA repair genes *MSH6*, *BRCA1*, and *BRCA2*. This adds to the growing body of evidence that DNA repair pathway defects may influence susceptibility to aggressive prostate cancer.

Conclusions: Our findings suggest that rare variants influence risk of clinically relevant prostate cancer and, if validated, could serve to identify men for screening, prophylaxis, and treatment.

Impact: This study provides evidence that rare variants in *TET2* may help identify African American men at increased risk for clinically relevant prostate cancer. *Cancer Epidemiol Biomarkers Prev; 25(11); 1456–63. ©2016 AACR.*

Introduction

Prostate cancer is a large medical problem in the developed world. More than 180,000 new cases are diagnosed, and 26,000 men die from the disease each year in the United States alone (1). The discrepancy between the number of men diagnosed with the disease and the number of men who die of the disease is due in part to diagnosis of indolent and clinically unimportant prostate cancer in many men. Most men will develop histologic prostate cancer at some point in their lifetime (2), but only a subset will be diagnosed with clinically relevant disease. Many of the tumors detected by modern PSA-based screening prove to be clinically indolent. Yet prostate cancer still accounts for 9% of all cancer deaths in men in the United States. Although improvements in screening and treatment have led to a 3.6% drop in mortality per year since 1994 (1), many men still develop metastatic disease, which is uniformly fatal.

Identifying men who are at increased risk of clinically relevant disease is therefore one of the most important challenges in prostate cancer management. Unfortunately, whether the most common screening tool (PSA) effectively makes this discrimination is a matter of debate. Two large randomized trials have only added to the PSA screening controversy. While the ERSPC trial demonstrated a benefit to PSA screening, the PLCO trial did not (3, 4). Even in the ERSPC trial, over 1,000 patients needed to be screened, biopsied, and treated to save one life. It has been estimated that 50% or more of patients diagnosed by PSA screening have clinically indolent tumors (5, 6) and therefore do not benefit from screening.

A growing body of evidence suggests that the risk for aggressive prostate cancer also has a strong heritable component. Prostate cancer in general appears to have a particularly strong genetic component, with upward of 58% of disease risk attributed to genetic factors (7). Family history and ethnicity are core components of risk stratification for prostate cancer screening. The male offspring of patients who die of prostate cancer have 4 times the risk of presenting with aggressive pathology and 2 times the risk of prostate cancer death than the general population (8–10). Analysis of germline DNA may allow improved screening by identifying individuals predisposed to clinically relevant disease.

[1]The McDonnell Genome Institute at Washington University, St. Louis, Missouri. [2]Division of Urology, Department of Surgery, Brigham and Women's Hospital, Harvard Medical School, Boston, Massachusetts. [3]Department of Urology, The James Buchanan Brady Urological Institute, Johns Hopkins University, Baltimore, Maryland. [4]Division of Statistical Genomics, Washington University School of Medicine, St. Louis, Missouri.

Note: Supplementary data for this article are available at Cancer Epidemiology, Biomarkers & Prevention Online (http://cebp.aacrjournals.org/).

Corresponding Author: Adam S. Kibel, Division of Urology, Brigham and Women's Hospital, 45 Francis Street, Boston, MA 02115. Phone: 617-732-6665; Fax: 617-525-6348; E-mail: akibel@partners.org

doi: 10.1158/1055-9965.EPI-16-0373

Genome wide association studies (GWAS) have identified and replicated multiple loci that influence the risk of prostate cancer (11–13). Importantly, however, most of the prostate cancer cases in these studies were diagnosed with less aggressive tumors. While retrospective analyses of these datasets (e.g., meta-analysis) have reported some loci associated with disease aggressiveness (14–18), the associations have not been independently replicated, nor do they account for a large fraction of the heritability of aggressive disease.

A fundamental weakness of GWAS efforts is that they typically do not interrogate rare or private variants (19). Surveys of well-known cancer genes in other malignancies have demonstrated the important contributions of rare variants (20–23) and current estimates suggest that rare variants account for as much as 42% of prostate cancer risk (24). Importantly, such rare variants tend to have larger effect sizes than common variants identified by GWAS, and therefore could potentially guide screening, prophylaxis, and treatment (25).

Some of the greatest disparities in cancer are in prostate cancer. African American men have the highest risk of disease, with incidence rates of 208.7 per 100,000 compared with 123.0 per 100,000 in European Americans. Mortality rates demonstrate similar trends when comparing men with African to European ancestry; prostate cancer mortality rates in these populations are 47.2 and 19.9 per 100,000, respectively (1). African American men tend to present with more advanced prostate cancer and have a worse prognosis than European American men, particularly in younger age groups (26). This disparity remains even after adjusting for access to health care, tumor grade, tumor stage, and other important clinical variables (26, 27).

Observations at known prostate cancer risk loci suggest that some of the risk and mortality disparity may arise from differences in the underlying genetic architecture. For example, while 8q24 has been identified as a risk factor for prostate cancer in men of both European and African ancestry, the risk alleles differ between ancestries (28–30). Other studies have demonstrated that variants in close proximity to risk variants in European American populations are strongly associated with risk in African American populations (31). A comprehensive study of genetic variation underlying aggressive prostate cancer risk must therefore include individuals from both ancestry groups.

Here, we describe a targeted sequencing study in aggressive prostate cancer patients and disease-free controls in African Americans and European Americans. The first stage examined the whole exome and then the second stage examined 800 promising candidates.

Materials and Methods

Subjects and controls

The samples in this study were collected at two institutions: Washington University in St. Louis, Missouri (WU), and Johns Hopkins Medical Institutions in Baltimore, Maryland (JHMI). WU samples comprised 272 cases (150 European Americans and 122 African Americans) and 300 controls (150 European Americans and 150 African Americans) recruited from urology and oncology clinics in the St. Louis metropolitan area. JHMI samples comprised 384 cases (305 European Americans and 79 African Americans) and 463 controls (305 European Americans and 158 African Americans) recruited from urology and oncology clinics in the Baltimore metropolitan area.

We included only cases with aggressive prostate cancer, which we defined as follows: evidence of metastatic disease (either pathologic or radiologic), a PSA of greater than 50 ng/mL, or Gleason grade of 8 to 10. No patients fit the criteria for familial prostate cancer (defined as 3 or more first-degree relatives with prostate cancer or 2 two relatives affected before age 55). None had a history of a familial tumor syndrome.

Ancestry-matched controls were recruited from both sites from individuals deemed to have a minimal risk of developing clinically relevant prostate cancer. Men had no personal or family history of prostate cancer, PSA < 4.0 ng/mL, and a benign digital rectal examination. WU controls were over the age of 75 years or 65 years for the European Americans and African Americans, respectively. The goal was to select controls with low risk of ever developing clinically relevant prostate cancer. JHMI controls had identical inclusion criteria except that they were age-matched to the cases.

All study subjects in all two cohorts provided informed consent under a protocol approved by Human Research Protection Offices. Genomic DNA was prepared from peripheral blood samples.

Gene selection

We used two strategies to select genes for targeted sequencing. First, we selected genes already implicated in prostate cancer from several sources (Fig. 1A), including genetic studies of prostate cancer susceptibility (GWAS loci), sequencing studies of prostate tumors (COSMIC mutation database), curated databases of cancer genes (Ingenuity and the Sanger Cancer Gene Consensus), and two pathway analysis tools (GeneGo and Acumenta). In total, these sources nominated 392 unique candidate genes. Although there was some redundancy among gene sets, most genes (318, or 79%) were nominated by a single source. Even the two largest sets of previously implicated genes—from GeneGo and Acumenta pathway analysis of prostate cancer—showed surprisingly little overlap: of 230 genes, only 36 (15.7%) were nominated by both pathway analyses (Fig. 1B).

To include potential novel prostate cancer susceptibility genes, we performed discovery whole-exome sequencing on a subset of individuals in our study cohort. Our goal was to identify genes with a differential burden of rare coding variants between cases and controls, independent of prior implication in prostate cancer. We therefore generated exome sequencing data for the WU population: 572 individuals: 272 cases (150 European Americans and 122 African Americans) with aggressive disease and 300 race-matched controls (150 European Americans, 150 African Americans).

We analyzed populations separately and computed the burden of rare (MAF < 0.05), deleterious variants among cases and controls on a gene-by-gene basis (Supplementary Fig. S2). We selected 474 genes that were enriched or depleted for rare variants in cases relative to controls at a nominal level of significance ($P < 0.05$). Of these, 66 were among the genes already compiled from previous sources (Fig. 1C). The full set thus comprised 800 genes (392 known, 408 potential novel; see Supplementary Table S1) which are the focus of our targeted sequencing and analysis.

Exome and targeted sequencing

Exome capture was performed on genomic DNA from 272 cases and 300 controls (compromising the entire WU population) using the Nimblegen SeqCap EZ v2.0 exome kit according to the

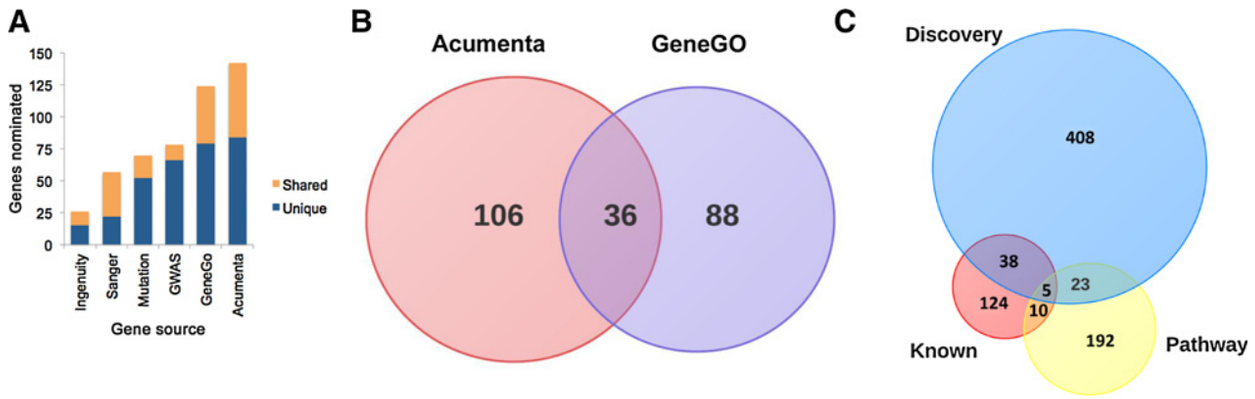

Figure 1.

Selection of genes for targeted resequencing. **A,** genes nominated for sequencing by source. **B,** overlap of prostate cancer pathway genes from Acumenta and GeneGo. **C,** overlap of genes from the candidate set (known, pathway) and the discovery set (rare variant aggregation tests).

manufacturer's protocols (Roche). The remaining samples (compromising the entire JHMI population) underwent targeted capture with a custom-designed Nimblegen reagent for the exons of the 800 selected genes (total space: 6.76 Mbp; design available upon request).

Exome and custom capture libraries underwent paired-end sequencing (2 × 101 bp) on Illumina HiSeq2000 instruments according to the manufacturer's protocols (Illumina). On average, 17.2 Gbp and 0.48 Gbp of raw sequence data were generated for each exome and custom capture library, respectively. On average, 87% of target bases were covered at ≥20× and 93% were covered at ≥10× in the exome data. Fifteen samples (3 European American cases, one African American case, and 11 African American controls) failed to meet our threshold for minimum coverage (>70% of bases covered ≥20×) and were removed prior to analysis.

SNP array data and MDS analysis

All samples that underwent exome sequencing were genotyped using Affymetrix 600K SNP arrays according to the manufacturer's protocols. We performed multidimensional scaling (MDS) clustering of autosomal SNPs using PLINK v1.07 to obtain the first 10 dimensions (–mds-cluster 10). MDS plots of the first two dimensions readily distinguished African American from European American samples (Supplementary Fig. S1).

Sequence alignment and processing

Data were aligned with BWA-MEM v0.7.10 with quality trimming (−q 8) to remove low-quality bases at the ends of reads to the GRCh37-lite reference sequence. Data from individual runs were merged (if necessary) with Picard (http://picard.sourceforge.net) v1.113. Duplicates were marked using Picard MarkDuplicates. Reads marked as duplicates were retained in the alignment file (BAM file) but not used for subsequent analysis.

Variant calling

Single-nucleotide polymorphisms were called using Samtools v1.16 (samtools pileup –cv) with default settings and VarScan v2.3.6. Variant calls from both tools were filtered to remove false positives as previously described, and then merged (unique-union) using joinx v1.3 (https://github.com/genome/joinx). Conflicts where both callers predicted a variant at a position, but with differing genotypes, were resolved by taking the Samtools prediction. Small insertions and deletions (indels) were called using VarScan v2.3.6 and filtered using a different set of parameters.

Statistical analysis

MDS clustering of SNP array data was performed using PLINK v1.07 and specifying 10 dimensions (−k 10). In the discovery phase, single variant tests and SKAT aggregation tests were performed using a pipeline of Perl scripts and R libraries. Following the replication phase, MDS clustering was performed for all samples using the common (MAF > 1%) variant calls in the shared target region set. For the combined analysis (discovery + replication samples, 800 genes), the SKAT-O aggregation tests were performed using the EPACTS framework. Fisher exact tests (FET) were conducted using a customized Perl script and the fet() function of the R package.

All aggregation tests utilized only variants that were rare (defined as MAF < 5% in the population set) and either truncating (frameshift, essential splice site, and nonsense) or missense and predicted to be deleterious (by at least one of Polyphen, SIFT, or Condel) as annotated by Variant Effect Predictor (VEP) release 74. The analysis of rare truncating mutations, however, only included variants annotated as nonsense (SNVs only), essential splice site (SNVs/indels), or frameshift (indels only).

Results

The goal of this study was to identify rare variation associated with clinically relevant prostate cancer in two ethnic groups: African Americans and European Americans. We recruited 652 patients with aggressive prostate cancer and 752 healthy age-matched controls from the urology and oncology clinics at two institutions (Table 1). Because whole-exome sequencing of 1,400 samples was not feasible, we instead sought to identify a subset of genes in which rare variants might affect risk of aggressive prostate cancer.

Rare variant aggregation testing

We combined both exome and targeted sequencing data for the 800 selected genes, which yielded a total sample size of 497

Table 1. Study participants recruited at Washington University in St. Louis, Missouri (WU), and Johns Hopkins Medical Institutions in Baltimore, Maryland (JHMI)

Study	African Americans		European Americans	
Site	Cases	Controls	Cases	Controls
WU	122	150	150	150
JHMI	78	147	302	305
Total	200	297	452	455

African Americans (200 cases and 297 controls) and 907 European Americans (452 cases and 455 controls). To identify rare variants associated with aggressive prostate cancer, we performed two gene-level aggregation tests, FET and SKAT-O. Only rare (MAF < 0.05) deleterious variants were included, and populations were analyzed separately and combined (Fig. 2).

To correct for multiple testing, we established a threshold of 3.125×10^{-5} (0.05/1,600 tests) for study-wide significance. The only gene to meet that threshold was *TET2* in African Americans, in which 24.35% of cases harbored a rare coding variant compared with 9.61% of controls (FET $P = 1.84 \times 10^{-5}$, OR = 3.0; SKAT-O $P = 2.74 \times 10^{-5}$). Rare variants unique to cases clustered in at least two regions of the gene (Fig. 3A), suggesting that rare variation in specific protein domains drives this association.

In addition, we identified an additional 56 genes for whom there was enrichment of rare variants among either cases or controls in one of our gene-based tests which was suggestive of an association ($P < 0.01$). Eight demonstrated suggestive evidence of enrichment in both statistical tests (Table 2).

Of the additional 4 genes identified in African American cases and controls, 3 demonstrated enrichment of rare variants in the cases (*CEP63*, *TRBV7-7*, and *NUBP2*) and one demonstrated enrichment of rare variants in the controls (*SPRR3*). No gene reached our study-wide significance threshold in European Americans. However, 2 genes demonstrated suggestive evidence of association using both statistical tests. Both *ZSWIM2* and *PARP2* variants were more common in controls than in cases suggesting a protective effect.

We also analyzed European and African American samples together, using individual ancestry as a covariate (SKAT-O only). *TET2* and indeed virtually all of the promising genes showed weaker association in the combined set, suggesting that the associations (if true) are specific to only one ancestry. However, an exception was *PARP2*. The combined analysis yielded stronger evidence of association for *PARP2* by both tests (FET $P = 7.06 \times 10^{-4}$, OR = 0.3; SKAT-O $P = 1.88 \times 10^{-3}$), though it still did not reach our threshold for study-wide significance.

A key advantage of sequencing studies over GWAS is the ability to identify rare truncating mutations (nonsense, splice site, and frameshift) that are poorly interrogated by most SNP arrays. Only 226 (~28%) of the genes surveyed here harbored a rare truncating

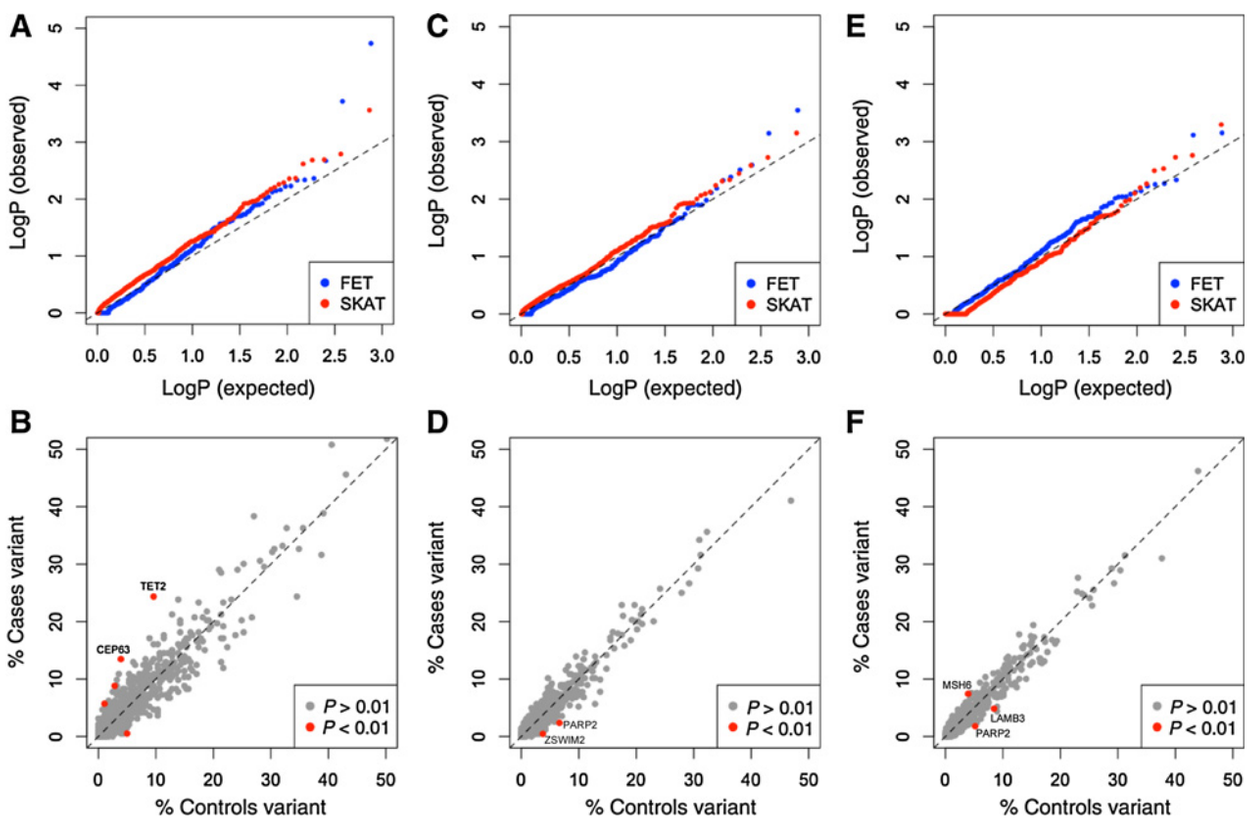

Figure 2.
Rare variant aggregation test results for 800 genes. Top, quantile–quantile plots for the African American (**A**), European American (**C**), and combined (**E**) sample sets. Bottom, fraction of samples with at least one rare deleterious variant in each gene in African American (**B**), European American (**D**), and combined (**F**) sample sets. Genes with FET $P < 0.01$ are marked in red.

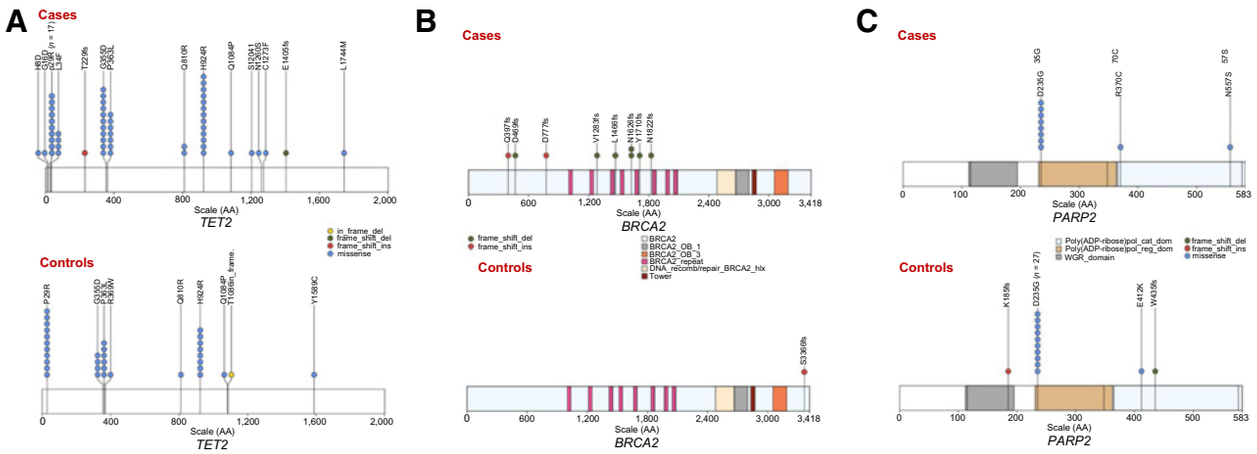

Figure 3.

Patterns of rare deleterious variants in selected genes. **A,** rare deleterious variants observed in *TET2* in African American cases (top) and controls (bottom). **B,** rare truncating variants observed in *BRCA2* in European American cases (top) and controls (bottom). **C,** rare deleterious variants in *PARP2* observed in European American cases (top) and controls (bottom).

variant in at least one sample. Of those, five exhibited an excess of rare truncating variants in cases relative to controls (Table 3) FET $P < 0.05$. Although none of these achieved the threshold for study-wide significance, the presence of three well-studied cancer predisposition genes (*BRCA1*, *BRCA2*, and *MSH6*) lends support to the growing body of evidence that rare loss-of-function variants in the DNA repair pathway may contribute to inherited prostate cancer risk.

Discussion

In this study, we performed targeted sequencing of 800 genes in men from two ancestry groups (African Americans and European Americans) to identify rare variation associated with clinically relevant prostate cancer. Genes were chosen for analysis on the basis of previously implicated gene (known set), implicated pathway (pathway set) or whole-exome sequencing (discovery set). Most of the prior genetic studies either included benign as well as clinically relevant tumors (11–13) or were retrospective analyses (14–18) whose reported associations have not been independently replicated. Importantly, all of these studies relied on GWAS approaches, which generally do not interrogate rare genetic variation.

Table 2. Genes with suggestive evidence of association ($P < 0.01$) by both FET and SKAT-O aggregation tests

Gene	Source	Controls	Cases	FET	OR	SKAT
African Americans (200 cases and 297 controls)						
TET2	Known	9.61%	24.35%	1.84×10^{-5}	3.0	0.00027
CEP63	Discovery	3.91%	13.47%	0.00019	3.8	0.00241
TRBV7-7	Discovery	1.07%	5.70%	0.00468	5.6	0.00621
NUBP2	Pathway	2.85%	8.81%	0.00588	3.3	0.00206
SPRR3	Discovery	4.98%	0.52%	0.00598	0.1	0.00515
European Americans (452 cases and 455 controls)						
ZSWIM2	Discovery	3.76%	0.47%	0.00072	0.1	0.00189
PARP2	Discovery	6.64%	2.36%	0.00310	0.3	0.00867
Both Ancestries (652 cases and 752 controls)						
PARP2	Discovery	5.18%	1.78%	0.00071	0.3	0.00188
MSH6	Pathway	3.96%	7.43%	0.00599	1.9	0.00051
LAMB3	Pathway	8.46%	4.85%	0.00915	0.6	0.00296

NOTE: The table shows the proportions of controls and cases with at least one rare (MAF < 5%) deleterious (truncating or damaging missense) variant. SKAT = SKAT-O.

The expectation that rare variants could predispose individuals to prostate cancer is consistent with well-established models of susceptibility for other cancer types. In breast cancer, for example, up to 24% of patients carry germline loss-of-function variants BRCA1/2 or one of ten other genes involved in genome stability and DNA repair (32). While the variants are independently rare, they confer an increased risk that can be as high as 10-fold, and the cumulative effect from multiple genes is even larger (33, 34).

Several studies have demonstrated that rare variants in specific genes play a role in the hereditary predisposition to prostate cancer. Ewing and colleagues demonstrated that the G84E variant of *HOXB13* was associated with a more than 20-fold increased risk of prostate carcinoma (35). Subsequent studies have demonstrated a more modest 3- to 4-fold increased risk; however, this demonstrates that a single rare variant found in less than 1% of the general population can account for an appreciable increased risk of disease (20). Mononen and colleagues examined a single rare SNP in the hormone-binding region of *AR* (R726L) which altered the transactivation specificity of the receptor. In a Finnish population, the rare leucine allele confers a 6-fold increase in risk for prostate cancer (36). A recent multi-ancestry study estimates that 42% of prostate cancer genetic risk is due to variants with frequencies between 0.1% and 1% (24).

We demonstrated that rare variation in *TET2* is associated with aggressive prostate cancer in African Americans. *TET2* variants appear to cluster in at least two domains (Fig. 3A) and only demonstrated evidence of association in African Americans. *TET2* is expressed in many tissues including prostate (37). It is thought to be a tumor suppressor on the basis of ~15% of myeloid cancers harboring somatic mutations (38). Prior work has identified variants in *TET2* in GWAS studies (39). Eeles and colleagues examined 1,854 prostate cancer cases and in 1,894 controls and identified 7 novel variants associated with prostate cancer risk, one of which was in close proximity to *TET2*. Eeles and colleagues found that the *TET2* variant was protective in European American, which is not consistent with it being a tumor suppressor gene. Interestingly, in contrast, the same variant was associated with

Table 3. Genes with suggestive evidence of enrichment (FET $P < 0.05$) for rare truncating variants, defined here as nonsense, splice site, or frameshift variants with MAF < 5%

Gene	Source	Controls	Cases	FET	OR
African Americans (200 cases and 297 controls)					
ELK4	Known	1.42%	5.18%	0.0250	3.8
European Americans (452 cases and 455 controls)					
BRCA2	Pathway	0.22%	2.12%	0.0092	9.8
DLEC1	Discovery	0.00%	1.42%	0.0126	14.1
Both Ancestries (652 cases and 752 controls)					
MSH6	Pathway	2.05%	4.68%	0.0083	2.4
BRCA1	Pathway	0.00%	0.81%	0.0199	13.1

nonstatistically increased risk in the relatively smaller African American subgroup. It is plausible that *TET2* could play a role in both ethnicities but through different mechanisms of action or the genetic architecture differs between African Americans and European Americans. The fact that variants increase the risk of aggressive prostate cancer is supported by findings that somatic mutations in *TET2* may contribute to metastatic potential in prostate cancer (40). At least some of the *TET2* variants identified in cases would be predicted to be function altering, e.g., C1273F affecting coordination of Zn++ (41).

Our study also highlighted some genes with suggestive evidence of association that did not reach study-wide significance, most notably those in the DNA repair pathway. *BRCA1*, *BRCA2*, *MSH6*, and *PARP2* were all implicated, predominantly in the European American, raising the possibility that these are important in this population.

Our observation that ~3% of aggressive prostate cancer cases harbored rare truncating mutations in *BRCA1* or *BRCA2* adds to the growing body of literature implicating them as *bona fide* prostate cancer predisposition genes. Prior work has identified single base pair variants (as opposed to frameshift indels) in *BRCA1* and *BRCA2* as risk alleles for prostate cancer in patients in familial cohorts. Our results implicate truncating variants. The only frameshift variant identified in the control population was at the 3′ end of the gene and therefore less likely to alter function (Fig. 3B). Comparison of known carriers of *BRCA1* and *BRCA2* mutations has demonstrated a 4- and 5-fold increased risk of prostate cancer, respectively (21, 22) and in particular to an increased risk of aggressive disease (23, 42–44). Of greatest interest, men with *BRCA* variants undergoing curative therapy for localized prostate cancer demonstrated decreased metastasis-free survival and overall survival compared with men without germline variants (23). A recent sequencing study of 150 metastatic prostate tumors reported *BRCA1/BRCA2* somatic alterations in 13.3% of cases (45). The observation that 90% of patients with germline alteration in *BRCA2* exhibited biallelic loss in metastatic deposits (45) supports a causal role for loss-of-function mutations in this gene. Our study represents an important addition to this prior work, in that we examined patients without a family history of cancer syndromes.

Less is known about the role of other candidate DNA repair genes in prostate cancer. Analysis of Lynch families revealed that they had a 5-fold increased risk of prostate cancer; however, only 2 of the 11 men diagnosed had *MSH6* mutations and the syndrome was not associated with aggressive disease. *PARP2* variants clustered in a known regulatory domain (Fig. 3C). A recent familial study of patients with prostate cancer also identified *PARP2* as a candidate susceptibility gene and reported a variant (A283G) in the same regulatory domain (46). That study, along with the report that pharmacologic PARP inhibition is therapeutic for advanced prostate cancer (47), lends further support to the idea that disruption of the PARP-associated DNA repair pathway may protect against aggressive disease.

A possible concern about our study is that we examined aggressive prostate cancer patients and disease-free controls, but not patients with indolent or low-risk disease. Although it would be interesting to examine rare genetic variation across the full spectrum of patients, we excluded those with indolent disease for two reasons. First, patients with indolent prostate cancer are clinically similar to patients without disease because both groups have excellent survival rates when given close observation (48). Second, we excluded patients with indolent tumors because the clinical course remains unknown: without intervention, they might very well have progressed to aggressive disease (case status). However, it is important to recognize that because we did not examine indolent disease, the variation identified could be associated with general risk of prostate cancer, not specific to aggressive disease. A study examining these variants in patients with indolent disease will be needed to prove the association is specific for aggressive prostate cancer.

Sample size is often the main limitation of genetic studies, due to either cost or the availability of samples. In our target sequencing cohort, we decided to focus on 800 genes (rather than the entire exome) so that we could sequence all available samples while controlling costs. Yet even our entire cohort (652 cases and 752 controls) is modest in size. Unlike indolent prostate cancer, aggressive, potentially lethal prostate cancer has a relatively low incidence; ascertaining thousands of samples remains a significant challenge. Even so, to our knowledge, our study of 800 genes in ~1,400 samples represents the largest germline sequencing study of aggressive prostate cancer to date. As the cost of sequencing continues to fall, more ambitious studies that use exome or whole-genome sequencing in larger populations may become feasible.

Conclusion

We applied targeted sequencing of 800 genes to in over 1,400 individuals and found that rare variation in *TET2* is associated with aggressive prostate cancer in African Americans. We also identified several genes with promising (if not statistically significant) evidence of association, including several members of DNA repair pathways. If confirmed in clinically relevant populations, these targets have the potential to identify a population for targeted screening, prophylaxis, and treatment.

Disclosure of Potential Conflicts of Interest

I.B. Borecki is Head Analytical Genetics at Regeneron. No potential conflicts of interest were disclosed by the other authors.

Authors' Contributions

Conception and design: D.C. Koboldt, D.E. Larson, R.S. Fulton, I.B. Borecki, R.K. Wilson, E.R. Mardis, A.S. Kibel

Development of methodology: D.C. Koboldt, K.L. Kanchi, R.S. Fulton, A. Kraja, R.K. Wilson, A.S. Kibel

Acquisition of data (provided animals, acquired and managed patients, provided facilities, etc.): R.S. Fulton, W.B. Isaacs, R.K. Wilson, E.R. Mardis, A.S. Kibel

Analysis and interpretation of data (e.g., statistical analysis, biostatistics, computational analysis): D.C. Koboldt, K.L. Kanchi, D.E. Larson, W.B. Isaacs, I.B. Borecki, L. Jia, R.K. Wilson, E.R. Mardis, A.S. Kibel

Writing, review, and/or revision of the manuscript: D.C. Koboldt, K.L. Kanchi, W.B. Isaacs, A. Kraja, R.K. Wilson, A.S. Kibel
Administrative, technical, or material support (i.e., reporting or organizing data, constructing databases): B. Gui, D.E. Larson, R.K. Wilson, A.S. Kibel
Study supervision: R.K. Wilson, E.R. Mardis

Grant Support

A.S. Kibel was supported by the Anthony DeNovi Fund, the Donald C. McGraw Foundation, and the St. Louis Men's Group Against Cancer. W.B. Isaacs was supported by William Gerrard, Mario Duhon, John and Jennifer Chalsty and the Patrick C. Walsh Cancer Research Fund. The remaining authors were supported by the National Human Genome Research Institute (grant #5U54HG003079, P.I. R.K. Wilson).

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked *advertisement* in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

Received May 5, 2016; revised June 26, 2016; accepted July 20, 2016; published OnlineFirst August 2, 2016.

References

1. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2016. CA Cancer J Clin 2016;66:7–30.
2. Scardino PT, Weaver R, Hudson MA. Early detection of prostate cancer. Hum Pathol 1992;23:211–22.
3. Andriole GL, Crawford ED, Grubb RL III, Buys SS, Chia D, Church TR, et al. Mortality results from a randomized prostate-cancer screening trial. N Engl J Med 2009;360:1310–9.
4. Schroder FH, Hugosson J, Roobol MJ, Tammela TL, Ciatto S, Nelen V, et al. Screening and prostate-cancer mortality in a randomized European study. N Engl J Med 2009;360:1320–8.
5. Etzioni R, Penson DF, Legler JM, di Tommaso D, Boer R, Gann PH, et al. Overdiagnosis due to prostate-specific antigen screening: lessons from U.S. prostate cancer incidence trends. J Natl Cancer Inst 2002;94:981–90.
6. Draisma G, Etzioni R, Tsodikov A, Mariotto A, Wever E, Gulati R, et al. Lead time and overdiagnosis in prostate-specific antigen screening: importance of methods and context. J Natl Cancer Inst 2009;101:374–83.
7. Hjelmborg JB, Scheike T, Holst K, Skytthe A, Penney KL, Graff RE, et al. The heritability of prostate cancer in the Nordic twin study of cancer. Cancer Epidemiol Biomarkers Prev 2014;23:2303–10.
8. 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;8:1001–6.
9. Hemminki K, Ji J, Forsti A, Sundquist J, Lenner P. Concordance of survival in family members with prostate cancer. J Clin Oncol 2008;26:1705–9.
10. Jansson KF, Akre O, Garmo H, Bill-Axelson A, Adolfsson J, Stattin P, et al. Concordance of tumor differentiation among brothers with prostate cancer. Eur Urol 2012;62:656–61.
11. Al Olama AA, Kote-Jarai Z, Berndt SI, Conti DV, Schumacher F, Han Y, et al. A meta-analysis of 87,040 individuals identifies 23 new susceptibility loci for prostate cancer. Nat Genet 2014;46:1103–9.
12. Eeles R, Goh C, Castro E, Bancroft E, Guy M, Al Olama AA, et al. The genetic epidemiology of prostate cancer and its clinical implications. Nat Rev Urol 2014;11:18–31.
13. Eeles RA, Olama AA, Benlloch S, Saunders EJ, Leongamornlert DA, Tymrakiewicz M, et al. Identification of 23 new prostate cancer susceptibility loci using the iCOGS custom genotyping array. Nat Genet 2013;45:385–91, 91e1–2.
14. Amin Al Olama A, Kote-Jarai Z, Schumacher FR, Wiklund F, Berndt SI, Benlloch S, et al. A meta-analysis of genome-wide association studies to identify prostate cancer susceptibility loci associated with aggressive and non-aggressive disease. Hum Mol Genet 2013;22:408–15.
15. Berndt SI, Wang Z, Yeager M, Alavanja MC, Albanes D, Amundadottir L, et al. Two susceptibility loci identified for prostate cancer aggressiveness. Nat Commun 2015;6:6889.
16. Cheng I, Plummer SJ, Neslund-Dudas C, Klein EA, Casey G, Rybicki BA, et al. Prostate cancer susceptibility variants confer increased risk of disease progression. Cancer Epidemiol Biomarkers Prev 2010;19:2124–32.
17. Helfand BT, Roehl KA, Cooper PR, McGuire BB, Fitzgerald LM, Cancel-Tassin G, et al. Associations of prostate cancer risk variants with disease aggressiveness: results of the NCI-SPORE Genetics Working Group analysis of 18,343 cases. Hum Genet 2015;134:439–50.
18. Machiela MJ, Chen C, Liang L, Diver WR, Stevens VL, Tsilidis KK, et al. One thousand genomes imputation in the National Cancer Institute Breast and Prostate Cancer Cohort Consortium aggressive prostate cancer genome-wide association study. Prostate 2013;73:677–89.
19. Kraft P, Hunter DJ. Genetic risk prediction–are we there yet? N Engl J Med 2009;360:1701–3.
20. Karlsson R, Aly M, Clements M, Zheng L, Adolfsson J, Xu J, et al. A population-based assessment of germline HOXB13 G84E mutation and prostate cancer risk. Eur Urol 2014;65:169–76.
21. Leongamornlert D, Mahmud N, Tymrakiewicz M, Saunders E, Dadaev T, Castro E, et al. Germline BRCA1 mutations increase prostate cancer risk. Br J Cancer 2012;106:1697–701.
22. Breast Cancer Linkage C. Cancer risks in BRCA2 mutation carriers. J Natl Cancer Inst 1999;91:1310–6.
23. Castro E, Goh C, Olmos D, Saunders E, Leongamornlert D, Tymrakiewicz M, et al. Germline BRCA mutations are associated with higher risk of nodal involvement, distant metastasis, and poor survival outcomes in prostate cancer. J Clin Oncol 2013;31:1748–57.
24. Mancuso N, Rohland N, Rand KA, Tandon A, Allen A, Quinque D, et al. The contribution of rare variation to prostate cancer heritability. Nat Genet. 2015.
25. Manolio TA, Collins FS, Cox NJ, Goldstein DB, Hindorff LA, Hunter DJ, et al. Finding the missing heritability of complex diseases. Nature 2009;461:747–53.
26. Powell IJ, Schwartz K, Hussain M. Removal of the financial barrier to health care: does it impact on prostate cancer at presentation and survival? A comparative study between black and white men in a Veterans Affairs system. Urology 1995;46:825–30.
27. Moul JW, Douglas TH, McCarthy WF, McLeod DG. Black race is an adverse prognostic factor for prostate cancer recurrence following radical prostatectomy in an equal access health care setting. J Urol 1996;155:1667–73.
28. Gudmundsson J, Sulem P, Manolescu A, Amundadottir LT, Gudbjartsson D, Helgason A, et al. Genome-wide association study identifies a second prostate cancer susceptibility variant at 8q24. Nat Genet 2007;39:631–7.
29. Haiman CA, Patterson N, Freedman ML, Myers SR, Pike MC, Waliszewska A, et al. Multiple regions within 8q24 independently affect risk for prostate cancer. Nat Genet 2007;39:638–44.
30. Freedman ML, Haiman CA, Patterson N, McDonald GJ, Tandon A, Waliszewska A, et al. Admixture mapping identifies 8q24 as a prostate cancer risk locus in African-American men. Proc Natl Acad Sci U S A 2006;103:14068–73.
31. Xu J, Kibel AS, Hu JJ, Turner AR, Pruett K, Zheng SL, et al. Prostate cancer risk associated loci in African Americans. Cancer Epidemiol Biomarkers Prev 2009;18:2145–9.
32. Walsh T, Casadei S, Lee MK, Pennil CC, Nord AS, Thornton AM, et al. Mutations in 12 genes for inherited ovarian, fallopian tube, and peritoneal carcinoma identified by massively parallel sequencing. Proc Natl Acad Sci U S A 2011;108:18032–7.
33. McClellan J, King MC. Genetic heterogeneity in human disease. Cell. 141:210–7.
34. Meindl A, Hellebrand H, Wiek C, Erven V, Wappenschmidt B, Niederacher D, et al. Germline mutations in breast and ovarian cancer pedigrees establish RAD51C as a human cancer susceptibility gene. Nat Genet 2010;42:410–4.
35. Ewing CM, Ray AM, Lange EM, Zuhlke KA, Robbins CM, Tembe WD, et al. Germline mutations in HOXB13 and prostate-cancer risk. N Engl J Med 2012;366:141–9.
36. Mononen N, Syrjakoski K, Matikainen M, Tammela TL, Schleutker J, Kallioniemi OP, et al. Two percent of Finnish prostate cancer patients have a germ-line mutation in the hormone-binding domain of the androgen receptor gene. Cancer Res 2000;60:6479–81.

37. Lorsbach RB, Moore J, Mathew S, Raimondi SC, Mukatira ST, Downing JR. TET1, a member of a novel protein family, is fused to MLL in acute myeloid leukemia containing the t(10;11)(q22;q23). Leukemia 2003;17:637–41.
38. Delhommeau F, Dupont S, Della Valle V, James C, Trannoy S, Masse A, et al. Mutation in TET2 in myeloid cancers. N Engl J Med 2009;360:2289–301.
39. Eeles RA, Kote-Jarai Z, Al Olama AA, Giles GG, Guy M, Severi G, et al. Identification of seven new prostate cancer susceptibility loci through a genome-wide association study. Nat Genet 2009;41:1116–21.
40. Nickerson ML, Im KM, Misner KJ, Tan W, Lou H, Gold B, et al. Somatic alterations contributing to metastasis of a castration-resistant prostate cancer. Hum Mut 2013;34:1231–41.
41. Hu L, Li Z, Cheng J, Rao Q, Gong W, Liu M, et al. Crystal structure of TET2–DNA complex: insight into TET-mediated 5mC oxidation. Cell 2013;155: 1545–55.
42. Tryggvadottir L, Vidarsdottir L, Thorgeirsson T, Jonasson JG, Olafsdottir EJ, Olafsdottir GH, et al. Prostate cancer progression and survival in BRCA2 mutation carriers. J Natl Cancer Inst 2007;99:929–35.
43. Mitra A, Fisher C, Foster CS, Jameson C, Barbachanno Y, Bartlett J, et al. Prostate cancer in male BRCA1 and BRCA2 mutation carriers has a more aggressive phenotype. Br J Cancer 2008;98:502–7.
44. Gallagher DJ, Gaudet MM, Pal P, Kirchhoff T, Balistreri L, Vora K, et al. Germline BRCA mutations denote a clinicopathologic subset of prostate cancer. Clin Cancer Res 2010;16:2115–21.
45. Robinson D, Van Allen EM, Wu YM, Schultz N, Lonigro RJ, Mosquera JM, et al. Integrative clinical genomics of advanced prostate cancer. Cell 2015;161:1215–28.
46. Johnson AM, Zuhlke KA, Plotts C, McDonnell SK, Middha S, Riska SM, et al. Mutational landscape of candidate genes in familial prostate cancer. Prostate 2014;74:1371–8.
47. Mateo J, Carreira S, Sandhu S, Miranda S, Mossop H, Perez-Lopez R, et al. DNA-repair defects and olaparib in metastatic prostate cancer. N Engl J Med 2015;373:1697–708.
48. Klotz L, Vesprini D, Sethukavalan P, Jethava V, Zhang L, Jain S, et al. Long-term follow-up of a large active surveillance cohort of patients with prostate cancer. J Clin Oncol 2015;33:272–7.