Exome Sequencing of African-American Prostate Cancer Reveals Loss-of-Function *ERF* Mutations

Franklin W. Huang[1,2,3], Juan Miguel Mosquera[4,5], Andrea Garofalo[3], Coyin Oh[3], Maria Baco[3], Ali Amin-Mansour[3], Bokang Rabasha[1,3], Samira Bahl[3], Stephanie A. Mullane[3], Brian D. Robinson[4,5], Saud Aldubayan[1,3], Francesca Khani[5], Beerinder Karir[4,5], Eejung Kim[1,3], Jeremy Chimene-Weiss[3], Matan Hofree[3], Alessandro Romanel[6], Joseph R. Osborne[7,8], Jong Wook Kim[1,3], Gissou Azabdaftari[9], Anna Woloszynska-Read[10], Karen Sfanos[11,12], Angelo M. De Marzo[11,12,13], Francesca Demichelis[4,6], Stacey Gabriel[3], Eliezer M. Van Allen[1,2,3], Jill Mesirov[3,14,15], Pablo Tamayo[3,14,15], Mark A. Rubin[4,5,16], Isaac J. Powell[17,18], and Levi A. Garraway[1,2,3]

ABSTRACT

African-American men have the highest incidence of and mortality from prostate cancer. Whether a biological basis exists for this disparity remains unclear. Exome sequencing (n = 102) and targeted validation (n = 90) of localized primary hormone-naïve prostate cancer in African-American men identified several gene mutations not previously observed in this context, including recurrent loss-of-function mutations in *ERF*, an ETS transcriptional repressor, in 5% of cases. Analysis of existing prostate cancer cohorts revealed *ERF* deletions in 3% of primary prostate cancers and mutations or deletions in *ERF* in 3% to 5% of lethal castration-resistant prostate cancers. Knockdown of *ERF* confers increased anchorage-independent growth and generates a gene expression signature associated with oncogenic ETS activation and androgen signaling. Together, these results suggest that *ERF* is a prostate cancer tumor-suppressor gene. More generally, our findings support the application of systematic cancer genomic characterization in settings of broader ancestral diversity to enhance discovery and, eventually, therapeutic applications.

SIGNIFICANCE: Systematic genomic sequencing of prostate cancer in African-American men revealed new insights into prostate cancer, including the identification of *ERF* as a prostate cancer gene; somatic copy-number alteration differences; and uncommon *PIK3CA* and *PTEN* alterations. This study highlights the importance of inclusion of underrepresented minorities in cancer sequencing studies. *Cancer Discov; 7(9); 973–83. ©2017 AACR.*

[1]Department of Medical Oncology, Dana-Farber Cancer Institute, Boston, Massachusetts. [2]Department of Medicine, Harvard Medical School, Boston, Massachusetts. [3]Cancer Program, the Broad Institute of Harvard and MIT, Cambridge, Massachusetts. [4]Caryl and Israel Englander Institute for Precision Medicine, Weill Cornell Medicine-New York Presbyterian, New York, New York. [5]Department of Pathology and Laboratory Medicine, Weill Cornell Medicine, New York, New York. [6]Centre for Integrative Biology, University of Trento, Trento, Italy. [7]Department of Radiology, Weill Cornell Medicine, New York, New York. [8]Department of Radiology, Memorial Sloan Kettering Cancer Center, New York, New York. [9]Department of Pathology, Roswell Park Cancer Institute, Roswell Park, New York. [10]Department of Pharmacology and Therapeutics, Roswell Park Cancer Institute, Roswell Park, New York. [11]Department of Pathology, Johns Hopkins University School of Medicine, Baltimore, Maryland. [12]Department of Urology, James Buchanan Brady Urological Institute, Johns Hopkins University School of Medicine, Baltimore, Maryland. [13]Sidney Kimmel Comprehensive Cancer Center at Johns Hopkins, Baltimore, Maryland. [14]Department of Medicine, University of California, San Diego, La Jolla, California. [15]Moores Cancer Center, University of California, San Diego, La Jolla, California. [16]Sandra and Edward Meyer Cancer Center at Weill Cornell Medicine, New York, New York. [17]Barbara Ann Karmanos Cancer Institute, Detroit, Michigan. [18]Department of Urology, Wayne State University School of Medicine, Detroit, Michigan.

Note: Supplementary data for this article are available at Cancer Discovery Online (http://cancerdiscovery.aacrjournals.org/).

F.W. Huang and J.M. Mosquera contributed equally to this article. M.A. Rubin, I.J. Powell, and L.A. Garraway contributed equally to this article.

Corresponding Authors: Levi A. Garraway, Dana-Farber Cancer Institute, 44 Binney Street, Boston, MA 02115. Phone: 617-632-6689; Fax: 617-582-7880; E-mail: levi_garraway@dfci.harvard.edu; Franklin W. Huang, Dana-Farber Cancer Institute, 450 Brookline Avenue, Boston, MA 02215. E-mail: franklin_huang@dfci.harvard.edu; Isaac J. Powell, Wayne State University School of Medicine, 540 East Canfield Street, Detroit, MI 48201. E-mail: ipowell@med.wayne.edu; and Mark A. Rubin, Weill Cornell Medicine, 1300 York Avenue, New York, NY 10065. E-mail: rubinma@med.cornell.edu

doi: 10.1158/2159-8290.CD-16-0960

INTRODUCTION

Extensive prior work has explored socioeconomic contributions to prostate cancer disparities; however, our knowledge of the extent to which molecular and genetic mechanisms may also contribute to prostate cancer disparities has been limited (1–6). The notion that somatic genetic factors may influence tumor biology differentially across distinct ancestral backgrounds is exemplified by high *EGFR* mutation rates in patients of Asian ancestry with non–small cell lung adenocarcinoma (up to 50% of patients), compared with patients of European ancestry (10%–15% of cases; ref. 7). Large-scale genomic characterization studies are predominated by tumor samples from patients of European ancestry (8). Although these studies are of immense value, their limited racial and ethnic diversity may preclude the detection of genomic events and patterns that are unique or enriched in underrepresented groups. For example, large-scale studies such as The Cancer Genome Atlas (TCGA) have examined the genomic landscape of primary prostate cancer and have been confined mainly to men of European ancestry (81.1%, 270/333; refs. 9–11). In particular, African-American (AA) men, who have a 1.4-fold higher incidence and 2.4-fold higher mortality rate from prostate cancer compared with non-Hispanic whites, have been underrepresented in most systematic studies of prostate cancer performed to date (10–12).

We hypothesized that differences in mutational events in AA prostate cancers may in part underlie these disparities in outcomes. We also reasoned that the power to discover novel cancer genes might increase through inclusion of diverse ancestral backgrounds in large-scale cancer genome studies. To test these hypotheses, we performed whole-exome sequencing on a discovery set of 102 localized primary prostate tumors and matched normal controls from a cohort of AA men and performed targeted sequencing on an extension set of 90 primary prostate tumors.

RESULTS

We focused on intermediate- and high-risk prostate cancers corresponding to Gleason grades 7 and higher, or pathologic stages pT2a–pT3c (Supplementary Tables S1 and S2). Exome sequencing identified 3,059 somatic mutations, corresponding to a median of 7 silent and 23 nonsilent mutations per tumor (range, 0–19 silent; 4–47 nonsilent). The median mutation rate for this cohort was approximately 0.83 mutations/Mb (range, 0.11–1.75), similar to mutation rates in exome sequencing cohorts of primary prostate cancer drawn predominantly from men of European ancestry (10, 11).

Overall, the majority of tumors from AA men do not harbor recurrent mutations in known cancer driver genes. Instead, most AA prostate cancers harbor somatic copy-number alterations (SCNA) that are characteristic of those seen in other published cohorts (Fig. 1 and Supplementary Fig. S1A; refs. 11, 13). However, the overall frequencies of SCNAs appear to be lower in this AA prostate cancer (AAPC) cohort. Comparing the frequency of SCNAs in primary prostate cancer, we found that of 19 loci that undergo recurrent copy-number changes, 16 loci were altered at a lower frequency in the AAPC cohort compared with the TCGA cohort (Supplementary Fig. S1B, Fisher exact test; $P < 0.05$). Adjusting for a lower threshold for SCNAs in the AAPC cohort (see Supplementary Methods), we found that 10 of the 19 loci were still significantly different in frequencies between the two cohorts (Supplementary Fig. S2). Given that the TCGA dataset comprises higher frequencies of Gleason 8 and higher tumors in comparison with the AAPC cohort, these copy-number differences were less pronounced when stratified by Gleason score yet persisted at certain loci (Supplementary Table S3 and Supplementary Fig. S2). For example, overall, *PTEN* deletions were more common in the TCGA cohort (32%) compared with the AAPC cohort (6%), consistent with previous reports, and this difference remained when examining Gleason 7 and Gleason 8 (or higher) tumors (Supplementary Fig. S2; ref. 2). We note that focal copy-number gains at 17q25.3, a locus containing the gene *FASN* and a significant amplification by GISTIC analysis in the AAPC cohort, occur in 12.7% of AAPC tumors, in contrast with 1% in the TCGA dataset ($P = 0.0001$, Fisher exact test; Supplementary Figs. S1B and S3; refs. 14, 15). The mean fraction of the copy number–altered genome in the AAPC cohort was lower in comparison with the TCGA cohort (7.94% vs. 15.6%; Wilcoxon rank sum test, $P = 0.0025$; Supplementary Fig. S4). When we stratified by Gleason grades, the mean fraction of copy number–altered genome did not differ significantly between Gleason 7 and lower tumors but did differ between the AAPC and TCGA cohorts in Gleason 8 and higher tumors (12.8% vs. 25.7%, Wilcoxon rank sum test, $P = 0.017$; Supplementary Fig. S4).

In the AAPC cohort, a focused germline analysis for pathogenic mutations in genes in the DNA repair pathway revealed 4 patients with germline mutations in *BRCA1* (with a concomitant hemizygous *BRCA1* deletion), *CHEK2*, and *ATM* for an overall prevalence of 3.9% (Supplementary Table S4; ref. 16).

In the analysis of somatic mutations, we identified three genes (*SPOP, ERF,* and *FOXA1*) wherein recurrent base mutations reached statistical significance in the discovery cohort (FDR $q < 0.1$; Fig. 1; Supplementary Table S5; ref. 17). Of the significantly mutated genes, *SPOP* and *FOXA1* have previously been identified as drivers in primary prostate cancer; however, *ERF* has not been implicated in this setting (10, 11). *ERF* is a member of the ETS transcription factor family and therefore was of interest given the prominent role that ETS transcription factor rearrangements play in prostate cancer (18). Of the five nonsynonymous *ERF* mutations present in our discovery cohort, three were loss-of-function events (R183*, K91fs, and R218*) and another, which occurred within the ETS DNA-binding domain (Y89C), was predicted to be a damaging event by Polyphen-2 analysis (19). In order to determine whether *ERF* mutations led to decreased ERF expression, we tested whether *ERF* mutants in prostate tumors from the AAPC cohort showed loss of *ERF* mRNA expression. We used RNA *in situ* hybridization (RNAish) to show that *ERF* mutants from the AAPC cohort were associated with a significant loss of *ERF* mRNA expression (Fig. 2A–C and Supplementary Fig. S5).

Using GISTIC analysis to examine significant SCNAs, we also noted that a focal deletion occurred at chr19q13.2 harboring a number of genes including *ERF* and a known

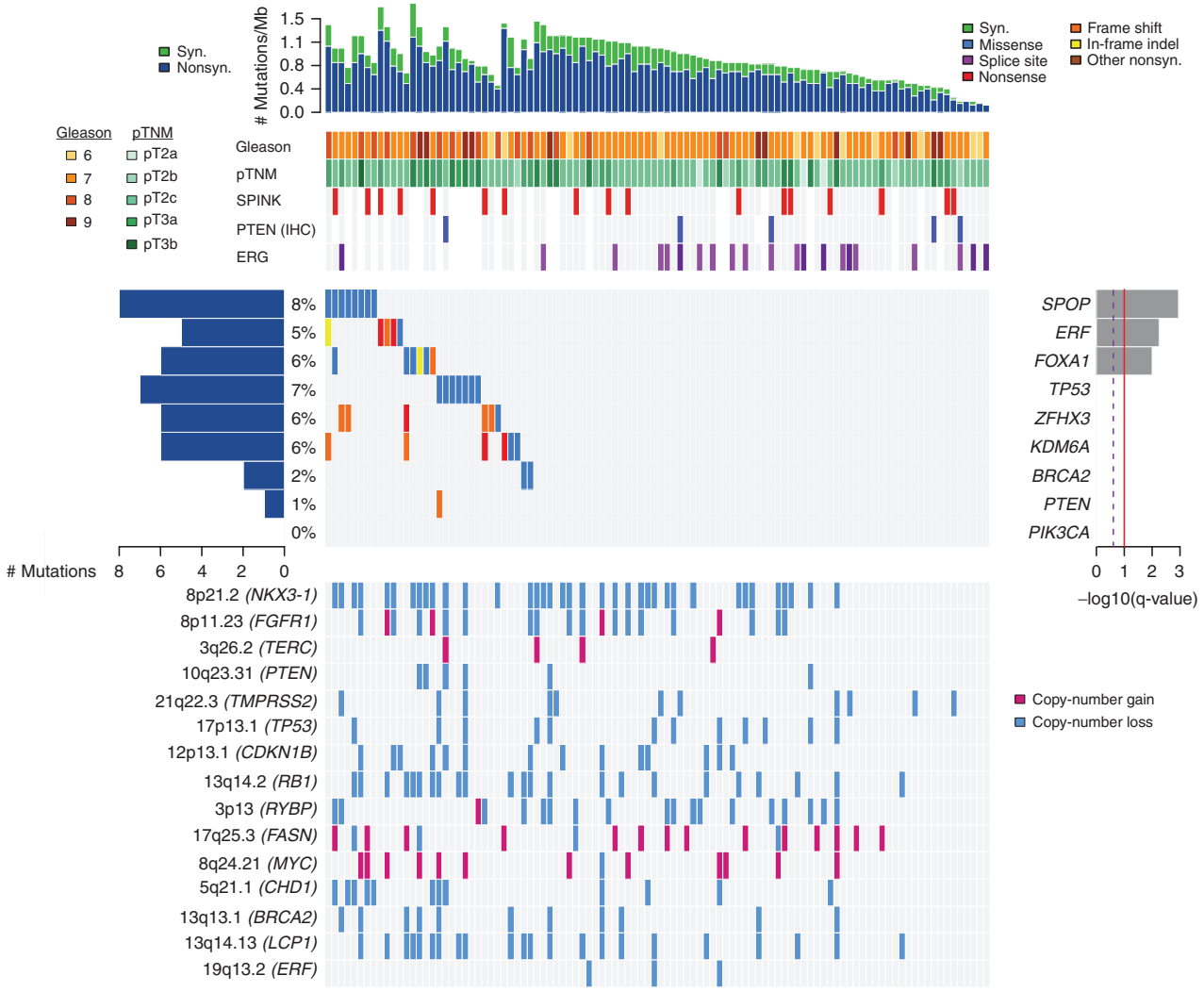

Figure 1. Exome mutation plot of 102 primary prostate cancers from AA men. Significantly mutated genes (*SPOP, ERF, FOXA1*) determined by MutSigCV ($q < 0.1$) and other selected prostate cancer genes are shown with mutation frequencies and types of mutations. Mean mutation rate per sample is shown in the top plot. Pathologic features of the prostate tumors including Gleason score, pathologic stage, and ERG by FISH, PTEN deletion by IHC, and SPINK1 overexpression are shown in the second plot for available samples. Somatic copy-number events across the samples at recurrently affected loci with noted prostate cancer genes in parentheses are shown in the bottom plot.

tumor suppressor, *CIC* (Supplementary Table S6; ref. 14). This peak represented three hemizygous copy-number losses (~3%) of *ERF* in our exome discovery cohort (Fig. 2D). These 3 patients had tumors with higher risk features: Gleason 8, pT3b; Gleason 8, pT3b with PSA of 42.3; and Gleason 7, pT2c with PSA of 12.8; 2 of these patients had biochemical recurrences. We verified copy-number loss of *ERF* in an AAPC tumor using FISH (Fig. 2E and F). To extend the finding of *ERF* copy-number loss, we assessed three prostate cancer cohorts for *ERF* deletion by FISH analysis. Only 1 of 105 cases of localized prostate cancer in AAs in a previously published cohort demonstrated hemizygous *ERF* deletion (2). None of 33 cases of localized prostate cancer in the predominantly white Early Detection Research Network cohort showed deletion in *ERF*. This evaluation included all tumor nodules in the same prostate gland, when multiple foci were present. We also interrogated 82 cases of advanced castration-resistant prostate cancer (CRPC), four of which harbored *ERF* hemizygous deletions (~5%; ref. 20). These cases were part of the Precision Medicine Clinical Trial at Weill Cornell, and an updated analysis of the whole-exome sequencing data from this cohort shows 8 of 175 cases (~5%) harbored genomic alterations of *ERF*, including 6 cases with deletion (3 hemizygous and 3 homozygous), 1 case with hemizygous deletion and concomitant H31P missense mutation, and 1 additional case with a G299 frameshift deletion (Supplementary Table S7; ref. 21).

To determine whether *ERF* might represent a recurrently mutated gene in primary AAPC, we expanded the dataset by including an additional 90 prostate cancer samples (AAPC extension cohort; Supplementary Table S8) and performed targeted hybrid capture sequencing for 41 known or putative prostate cancer genes (Supplementary Table S9). Tumor–normal pairs from this additional AA prostate cancer cohort

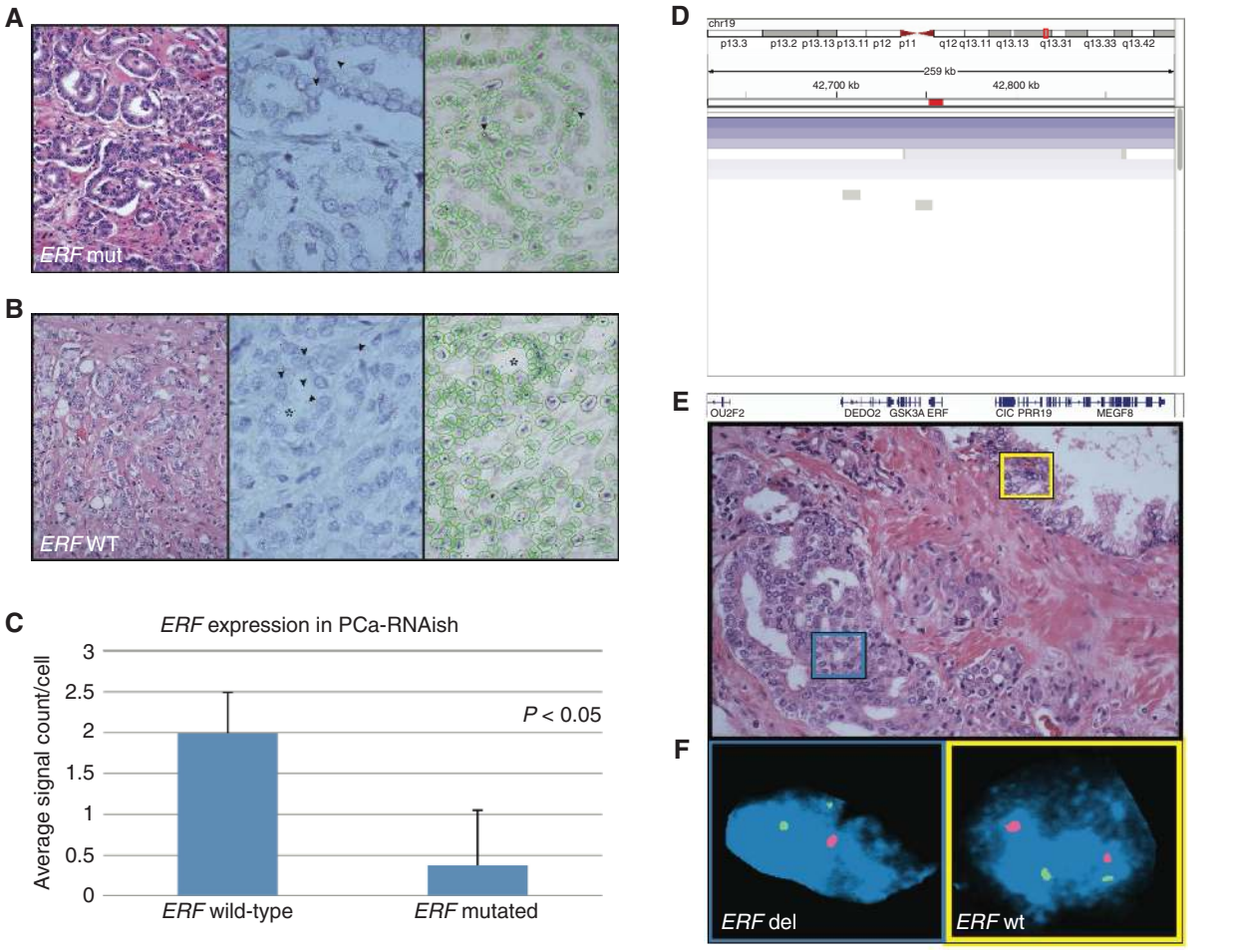

Figure 2. *ERF* RNA expression is significantly lower in *ERF*-mutated prostate cancer. **A,** Histology of prostatic adenocarcinoma, Gleason score 3 + 4 = 7 (case STID21598, left). This representative *ERF*-mutated tumor demonstrates only few *ERF* RNA signals as highlighted by arrowheads (center). RNA signals are seen as scattered yellow dots during image fragmentation for automated analysis (right). **B,** Histology of prostatic adenocarcinoma, Gleason score 4 + 3 = 7 (case STID21603, left). In contrast to *ERF*-mutated cases, this representative *ERF* wild-type (WT) prostate cancer demonstrates numerous signals by RNAish (asterisk, center). These are highlighted as yellow dots for automated image analysis (right). **C,** Automated image analysis of three *ERF* mutant cases vs. six *ERF* WT prostate cancer cases demonstrates a significant difference ($P < 0.05$) in RNA expression by this *in situ* assay. **D,** Integrative Genomics Viewer screenshot of deletions spanning the *ERF* locus in the AAPC exome cohort. **E,** Validation of *ERF*-deleted prostate cancer by FISH. Top plot shows histology of prostatic adenocarcinoma, Gleason score 4 + 3 = 7 (case STID12329). **F,** In the bottom plot, a FISH assay shows hemizygous *ERF* deletion in the tumor (left image; blue square in panel **F**), as demonstrated by only one target (red) signal and two centromeric (green) reference signals. In contrast, *ERF* wild-type status (right image; yellow square in **F**) is seen in nuclei of adjacent benign glands (original magnification: hematoxylin and eosin at ×40; FISH images at ×100 under oil immersion).

were sequenced at high coverage (mean target coverage: tumor 335x, normal 348x; Supplementary Methods). In the extension cohort, we identified five additional nonsynonymous *ERF* mutations, three of which were predicted loss-of-function frameshift mutations (Fig. 3A and Supplementary Table S10). In total, the prevalence of *ERF* mutations in the discovery and extension AA primary prostate cancer cohorts was 5.2% (10/192). We validated by Fluidigm array 7 of 7 of the *ERF* mutations that we were able to evaluate (Supplementary Fig. S6). Thus, *ERF* is recurrently mutated in primary prostate cancer in AA men.

Although *ERF* was found to be a significantly mutated gene in the AAPC cohort, it did not reach statistical significance in the TCGA cohort ($n = 333$; Fig. 3B; ref. 11). However, taking into account deletions as well as mutations in *ERF*, the frequency of somatic alterations in *ERF* is comparable between the AAPC cohort and other primary prostate cancer cohorts (5% vs. 3%), suggesting that loss-of-function by mutation or deletion may be mechanisms to dysregulate *ERF*. Overall, combining publicly available exome or whole-genome sequencing datasets from primary prostate cancer revealed that *ERF* was altered in approximately 3% of primary prostate cancer cases either by mutation (0.76%; 5/661) or by homozygous deletion (2.2%; 15/661; Supplementary Figs. S7 and S8A; Supplementary Table S11; refs. 9–11). Therefore, *ERF* is recurrently mutated or deleted in primary prostate cancer.

Although *ERF* was not previously recognized as a recurrently mutated gene in primary prostate cancer, lethal CRPC cohorts showed missense or loss-of-function *ERF* mutations in approximately 3% of CRPC tumors (Supplementary Figs. S7 and S8B–S8C; refs. 22, 23). An updated analysis of genomic

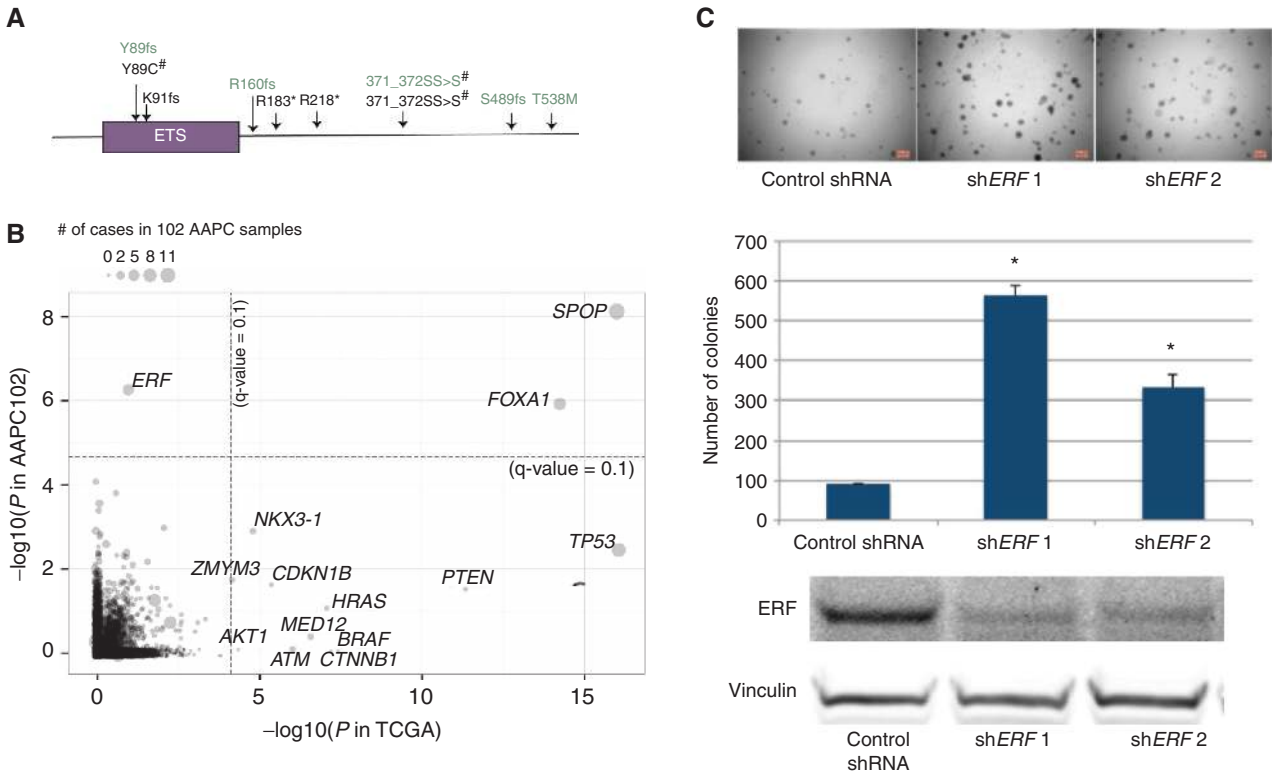

Figure 3. *ERF* is mutated in the AAPC cohort and can act as a tumor suppressor. **A,** Types and locations of mutations in *ERF* found in the discovery (black) and extension (green) AAPC cohorts. Mutations were validated by Fluidigm Access Array except those indicated by #. **B,** Mutation significance *P* values for genes from the AAPC cohort ($n = 102$) and TCGA ($n = 333$) were plotted. Dotted lines are drawn at the *q* value = 0.1 for significance within each cohort. **C,** Anchorage-independent growth assays of ERF knockdown in PC3 prostate cancer cell lines were performed in triplicate. Colonies were quantified with CellProfiler; *, $P < 0.05$, Student *t* test. Data shown are representative of two independent experiments. Western blot results are depicted for knockdown of ERF with two short hairpins and control hairpin.

data from a cohort of CRPC cases shows that *ERF* is recurrently mutated at a frequency of approximately 3% (8/269) and undergoes copy-number loss (hemizygous or homozygous) at a frequency of approximately 17.5% (47/269; Supplementary Fig. S8C; ref. 22). We also examined prostate cancer cell line data from the Cancer Cell Line Encyclopedia (CCLE) and Catalogue of Somatic Mutations in Cancer (COSMIC) databases and found that *ERF* is mutated in one of six prostate cancer cell lines (DU-145; p.A132S; refs. 24, 25).

We next asked whether alterations in *ERF* might be associated with more aggressive disease. *ERF* copy-number loss was associated with a number of aggressive pathologic features, including higher Gleason grade (Supplementary Fig. S9; $P = 0.0035$), higher pathologic T stage ($P = 0.00696$), and residual tumor ($P = 0.0435$; www.firebrowse.org). Of the 5 patients with *ERF* mutations (5/492) in the TCGA dataset, 4 had Gleason 8 or higher tumors (3 with Gleason 9), and 2 of these patients experienced biochemical recurrences. Among the 8 patients with *ERF* mutations or deletions in the AAPC exome cohort, 3 experienced biochemical recurrences (vs. 16 of 73 *ERF* wild-type) and 4 of 8 were pT3 (vs. 31 of 94 *ERF* wild-type). Moreover, a germline analysis of *ERF* in the AAPC cohort revealed an *ERF* coding variant (S295I) in a patient with Gleason 9 prostate cancer that experienced a biochemical recurrence (Supplementary Table S12). Overall, these data raise the possibility that *ERF* mutations and deletions may be linked to more aggressive forms of prostate cancer.

ERF was first characterized as an ETS and RAS tumor-suppressor protein with a transcriptional repressor function (26, 27). In addition to prostate cancer, mutations in *ERF* occur in other tumor types at similar frequencies: stomach adenocarcinoma (~4%), colorectal adenocarcinoma (~4%), and Ewing sarcoma (~3%), which is also notably driven by a common ETS rearrangement, *EWS–FLI* (28–30). We asked whether *ERF* might function as a tumor suppressor in prostate cancer cells. Using lentiviral shRNAs, we knocked down ERF in the PC3 prostate cancer cell line and demonstrated a significant increase in anchorage-independent growth of prostate cancer cells (Fig. 3C; Supplementary Fig. S10). ERF knockdown also increased invasion of PC3 cells and increased mouse tumor xenograft growth (Supplementary Fig. S11A–S11C). Overexpression of ERF reduced colony growth, whereas a mutant ERF harboring a mutation (Y89C) identified in the AAPC cohort diminished this effect (Supplementary Fig. S12A–S12C). Furthermore, ERF knockdown in another prostate cancer cell line (LNCaP) and immortalized prostate epithelial cell line (RWPE-1) showed increased growth proliferation but no significant increase in invasion (Supplementary Fig. S13A–S13H). Overexpression of wild-type ERF in the DU-145 cell line harboring a mutation in *ERF* had no effect on

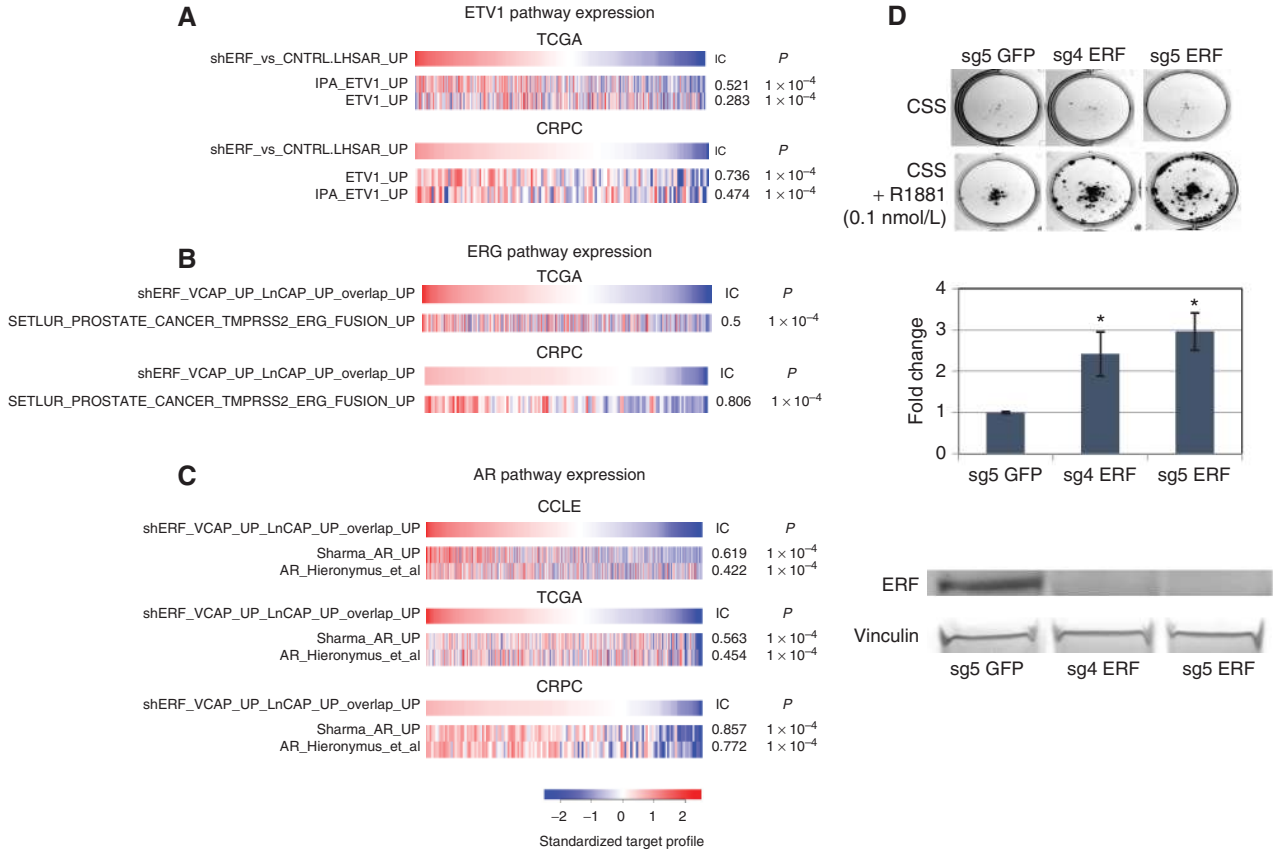

Figure 4. The association of *ERF* knockdown signatures with the single-sample GSEA enrichment profiles of ETS and AR gene sets. **A,** An *ERF* knockdown signature profile in the prostate epithelial cell line LHS-AR associates with the enrichment profiles of ETV1 gene sets in the TCGA and CRPC expression datasets. **B,** A composite *ERF* knockdown signature profile, i.e., the enrichment profile of the overlap between *ERF* knockdown signatures in the VCAP and LNCaP cell lines, associates with the enrichment profile of an ERG gene set across the TCGA and CRPC expression datasets. **C,** The same composite *ERF* knockdown signature association with enrichment profile of AR gene sets across the CCLE, TCGA, and CRPC datasets. **D,** Focus formation assay of LNCaP-sgERF cell lines in the context of charcoal-stripped serum (CSS) and androgen treatment (0.1 nmol/L R1881). Crystal violet staining was quantitated and compared with control. Mean and SE of three replicates is shown. *, $P < 0.05$ by Student *t* test, fold-change sg4 ERF versus sg5 GFP, sg5 ERF versus sg5 GFP. Data shown are representative of two independent experiments.

proliferation in a focus formation assay but led to a reduction in colony growth in a low-attachment assay and decreased invasion (Supplementary Fig. S14A–S14D). These results were consistent with a possible tumor-suppressor role for *ERF* in prostate cancer.

Given its known role as a repressor of the ETS transcription factor family, we hypothesized that loss of *ERF* might activate a transcriptional program that resembles the output of ETS transcriptional activators such as ERG or ETV1 (31, 32). To test this hypothesis, we performed lentiviral shRNA knockdown of *ERF* in an immortalized prostate epithelial cell line (LHS-AR) and two prostate cancer cell lines (VCaP and LNCaP) that each harbor oncogenic *ERG* (VCaP) or *ETV1* (LNCaP) rearrangements (Supplementary Fig. S10; ref. 33). We then generated transcriptome data (RNA sequencing) to derive a gene signature of ERF knockdown from the top 100 genes upregulated, according to the difference of means, when ERF is knocked down in comparison with control (see Methods). The ERF knockdown (KD)_UP signature from the LHS-AR cells was correlated with ETV1 and ERG signatures across the CCLE and was correlated with ETV1's target expression across the TCGA and CRPC tumor datasets (Fig. 4A; Supplementary Fig. S15). We then generated a combined ERF KD_UP signature from VCaP and LNCaP cell lines, hypothesizing that in the context of ETS activation that ERF loss might augment an ETS oncogenic transcriptional program. We found that this signature correlated with ERG pathway expression across TCGA and CRPC tumor datasets (Fig. 4B). Supporting the idea that *ERF* loss or dysfunction may function similarly to ERG activation, *ERF* deletions and mutations were mutually exclusive of *ERG* rearrangements in the AAPC exome cohort (Fig. 1). Furthermore, mutations and homozygous deletions of *ERF* were mutually exclusive of *ERG* fusion events in the published TCGA dataset (Supplementary Fig. S16; Fisher exact test: $P < 0.05$). In addition, overexpression of wild-type ERF in DU-145, an *ERF* mutant cell line, diminished the ERF KD signatures and ERG signature (Supplementary Fig. S17A).

We next asked whether the ERF KD signature could be associated with a more aggressive phenotype in prostate cancer. We tested whether the ERF KD gene signature was also correlated with features of aggressive prostate cancer in the TCGA dataset and found that higher Gleason scores

correlated with the ERF KD gene signature (Supplementary Fig. S17B).

In an unbiased pathway analysis of cell lines from the CCLE, we found an androgen receptor signaling (AR) signature (NELSON_RESPONSE_TO_ANDROGEN_UP) as the top correlated signature with respect to the ERF KD signature_UP (Supplementary Fig. S18; ref. 34). We projected this ERF KD signature into the RNA-sequencing datasets generated from tumor samples in cohorts of primary prostate cancer (TCGA) and CRPC (11, 22). We found that AR signatures were correlated with the ERF KD signature_UP in the TCGA and CRPC datasets (Fig. 4C; ref. 35). These data suggest that AR signatures are correlated with the transcriptional program of *ERF* knockdown and that loss of *ERF* is associated with a transcriptional program that can mimic ETS activation and may impinge on androgen signaling. To test the hypothesis that *ERF* loss may promote androgen signaling, we used the CRISPR/Cas9 system to target the *ERF* coding sequence in prostate cancer cells (36). We showed that loss of *ERF* is associated with an increase in androgen-dependent growth (Fig. 4D).

We also investigated other genes that were recurrently mutated but did not reach statistical significance in the AAPC discovery cohort. *YBX1* (K81T, R244*, N76S) was mutated 3 times and has been implicated as an oncogene in prostate cancer but had not been previously identified as recurrently mutated in any prostate cancer cohorts (37). We identified missense mutations in the steroid hydroxylase *CYP11B1*, as well as other cytochrome P450 family members that occurred in a total of approximately 14% of samples in the discovery cohort (Supplementary Table S13).

Among known cancer genes, we observed that *PIK3CA*, which is recurrently mutated at a frequency of approximately 3% (20/667) in primary prostate cancers, was not mutated in the AAPC discovery or extension cohorts (Fisher exact test, $P = 0.0115$; Fig. 1; refs. 10, 11, 30). Previous analyses also suggest that *PTEN* is less commonly deleted in cohorts of men of African ancestry (2, 38). Our data suggest that alteration of the PI3K signaling pathway through either *PTEN* deletion or *PIK3CA* mutation is a less common event in AAPC. Interestingly, we identified missense mutations in *FOXA1* (F254V and H247L) in the AAPC cohort that occurred at the same residues only in men of African ancestry in the TCGA dataset (F254V and H247Y), raising the possibility of somatic mutations in prostate cancer that may be associated with ancestry (39).

Finally, to test the addition of tumors from AA men to a large cohort of primary prostate cancer largely from men of European ancestry, we performed a combined analysis of the AAPC discovery ($n = 102$) and TCGA ($n = 457$) cohorts which nominated several new significantly mutated genes, including *SMARCA1* and *ZFHX3* (Supplementary Fig. S19; ref. 17).

DISCUSSION

Prostate cancer sequencing studies have been comprised primarily of men of European ancestry, and there have been few large studies focused on men of African ancestry. Here, we use exome sequencing to identify genomic features of primary prostate cancer in AA men. We report the identification of recurrent mutations in *ERF*, a prostate cancer gene that had not been previously appreciated. Although analysis of this cohort identified *ERF* as a significantly mutated gene in prostate cancer, we note that *ERF* was mutated and located within a focal deletion in the TCGA cohort, composed primarily of men of European ancestry. Focal deletions at chr19q13.2 encompass *ERF* and *CIC*, a tumor suppressor that has been shown to regulate the ETS factors *ETV1, ETV4*, and *ETV5* (40). Our exome sequencing study implicates *ERF* as a potential target of these deletions. Still, we cannot exclude the possibility that deletion of both *ERF* and *CIC* contributes to prostate cancer. Therefore, our data in conjunction with TCGA data suggest that *ERF* can be altered through mutation or deletion.

Dysregulation of ETS transcription factors plays a major role in prostate carcinogenesis, and several studies have reported a lower prevalence of *TMPRSS2–ERG* rearrangements in prostate tumors of AA men (2, 41). Here, we present evidence of a mechanism of affecting ETS transcriptional output through an ETS transcriptional repressor. Furthermore, alteration of *ERF* may also be associated with more aggressive prostate cancers.

We observed a number of recurrent SCNAs that differed between the primary AAPC and TCGA cohorts. The limitations of this comparison with TCGA include differences in technologies (exome sequencing vs. SNP arrays) and stringencies for copy-number detection. Still, our results, in conjunction with other studies that have examined specific copy-number alterations, suggest that overall the SCNA landscape may be distinct in primary prostate cancer in AA men. In addition, we find that alterations of the PI3K signaling pathway through deletion of *PTEN* or mutation of *PIK3CA* are uncommon in primary AAPC, suggesting that distinct patterns of genomic alterations may occur in this cohort with implications for precision medicine.

Our results suggest that increasing the ancestral diversity of study populations for cancer genomic characterization may help increase the discovery potential of these studies, which we believe to date have not included sufficiently large numbers of men of African ancestry. Given the relatively lower mutation rate of prostate cancer, larger cohorts of AA patients with prostate cancer may be required to identify recurrently mutated genes that may contribute to prostate carcinogenesis or to aggressive prostate cancer features in this population. These studies will inform whether alterations in these genes may be enriched in certain ancestral groups. Recent studies have implicated prostate tumor location, differential gene expression, and somatic genomic events such as *LSAMP* deletions in prostate cancers in AA men (2–4, 42–44). Our study suggests that there are still unexplained reasons for the aggressive nature of prostate cancer in AA men, which is only partially explained by the genomic studies to date. Additional studies focused on metastatic CRPC samples from AA men, and development of methodologies to integrate analyses of somatic and germline data may improve our understanding of the nature of aggressive prostate cancer in these patients (23, 45–47). Our results suggest that inclusion of sufficient numbers of patients of African ancestry in cancer genomic studies may enable the discovery of new cancer genes and inform the inclusion of diverse populations toward precision cancer medicine (48).

METHODS

Cohort Description and Pathology Evaluation

The discovery cohort comprised specimens from Weill Cornell Medicine (WCM) and Karmanos Cancer Center (KCC; Supplementary Table S2). The extension cohort comprised samples from KCC, Johns Hopkins University/Prostate Cancer Biorepository Network (PCBN), and Roswell Park Cancer Institute (RPCI; Supplementary Table S8). All cohort samples were from patients self-identified as African-American. Archival pathology specimens were obtained retrospectively from four different institutions under Institutional Review Board (IRB) protocols: WCM (IRB #1007011157), RPCI (IRB #BDR-035413), Johns Hopkins University (JHU; IRB # NA_00048544), and KCC (IRB # 044812MP4E). Hematoxylin and eosin–stained slides were reviewed by study pathologists at WCM (J.M. Mosquera, B.D. Robinson, F. Khani, and M.A. Rubin), RPCI (G. Azabdaftari), JHU (A.M. De Marzo), and KCC. Annotated slides containing tumor and benign tissue were used for somatic and germline DNA. Clinical and pathologic data are summarized in Supplementary Tables S2 and S8. Data on *ERG* rearrangement, *PTEN* deletion, *SPOP* mutation, and SPINK1 expression were available for the WCM cohort, as previously published (2).

ERF *FISH*

To assess *ERF* deletion in tissues, we developed a dual-color FISH assay consisting of a locus-specific probe (W12-2967N22) plus reference probe spanning a stable region of the chromosome (RP11-46I12). All clones were tested on normal metaphase spreads of comparative genomic hybridization target slides as previously described (9, 21). *ERF* deletion was defined as the presence of 0 or 1 copy on average per nucleus compared with two reference signals. At least 100 nuclei were evaluated per tissue section using a fluorescence microscope (Olympus BX51; Olympus Optical).

ERF *RNAish*

This single-color chromogenic detection assay uses pairs of specially designed oligonucleotide probes that, through sequence-specific hybridization, recognize both the specific target *ERF* RNA sequence and the signal amplification system (see *ERF* oligonucleotide list, Supplementary Table S14; Affymetrix, Inc.). Based on unique coordinates on chromosome 19, unique target probe oligonucleotides were designed (see *ERF* map sequence, Supplementary Fig. S20). The latter is designed to hybridize in tandem to the target RNA. Cross-hybridization to other sequences is minimized by screening against the entire human RNA sequence database. The signal amplification system consists of the preamplifier, amplifier, and enzyme-conjugated label probe, which assemble into a tree-like complex through sequential hybridization. Signal amplification occurs at target sites bound by probe pairs only. Nonspecific off-target binding by single probes does not result in signal amplification. All steps of *ERF* RNAish staining of the slides were performed manually (49). Briefly, formalin-fixed, paraffin-embedded (FFPE) unstained tissue sections (5 μm) were mounted on positively charged microscopic glass slides, deparaffinized in xylene, and dehydrated through a series of alcohols. The dehydrated sections were then treated, and sequential hybridization of probe and amplifiers was performed according to Affymetrix protocols. The rehydrated sections were treated with 3% hydrogen peroxide at room temperature for 10 minutes to block endogenous peroxidase. Sections were then boiled in 1 × citric buffer (10 nmol/L Nacitrate, pH 6.0) for 15 minutes and incubated with protease (2.5 mg/mL; Sigma Aldrich) at 40°C for 30 minutes. The slides were hybridized sequentially with target probes (20 nmol/L) in hybridization buffer A [6 × saline sodium citrate (SSC) buffer (1 × SSC is 0.15 mol/L NaCl and 0.015 mol/L Na-citrate), 25% formamide, 0.2% lithium dodecyl sulfate (LDS), and blocking reagents] at 40°C for 2 hours, signal preamplifier in hybridization buffer B (20% formamide, 5 × SSC, 0.3% LDS, 10% dextran sulfate, and blocking reagents) at 40°C for 30 minutes, amplifier in hybridization buffer B at 40°C for 30 minutes, and horseradish peroxidase– or alkaline phosphatase–labeled probes in hybridization buffer C (5 × SSC, 0.3% LDS, and blocking reagents) at 40°C for 15 minutes. Hybridization signals were detected under brightfield microscope as red colorimetric staining followed by counterstaining with hematoxylin. Signals were granular and discrete red dots corresponding to individual RNA targets.

Quantitative ERF *RNAish Analysis*

RNAscope SpotStudio Software from Definiens, Inc. was utilized for image analysis of *ERF* RNA *in situ* expression at a single-cell resolution.

Exome Sequencing and Analysis

For our whole-exome sequencing discovery cohort, we collected treatment-naïve radical prostatectomy specimens collected from two primary sites: New York City (WCM) and Detroit (Karmanos Cancer Institute/Wayne State University). All patients were self-reported AA. All tissue DNA was extracted from FFPE tissues. We sequenced 128 matched tumor–normal pairs, and after quality control for contamination and low tumor purity (Supplementary Table S15), we analyzed 102 matched pairs (see Supplementary Methods). The baits for exon capture targeted 98.2% of genes in the Consensus CDS (CCDS) database. A mean target coverage depth of 100× per sample was achieved, with 80% of targets covered at a depth of ≥20×.

Somatic Alterations, Filters, and Germline Analysis

We used Mutect and the Indelocator (http://www.broadinstitute.org/cancer/cga/indelocator) for calling single-nucleotide changes and insertion/deletions (Supplementary Methods). We used an FFPE filter to remove mutations likely from FFPE artifacts. We also used the GATK HaplotypeCaller to find germline variants within a specific gene.

Germline Variant Interpretation

The analysis of germline variants focused on variants identified among 20 genes that are associated with autosomal-dominant cancer-predisposition syndromes. These genes were chosen for their crucial role in the maintenance of DNA integrity. Pathogenicity of germline variants was determined according to the most recent guidelines published jointly by the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. Germline variants were evaluated against published literature and publicly available databases such as ClinVar and variant-specific databases. In addition, population-based frequency databases including 1000 Genomes and the Exome Aggregation Consortium were examined. Only pathogenic and likely pathogenic variants with high or moderate penetrance were reported in this study. Low penetrance variants were excluded.

DNA Sequencing Validation

For validation, we used the Fluidigm Access Array microfluidic device. PCR products were barcoded, pooled, and subjected to Illumina sequencing on a MiSeq instrument.

Cell Culture/Lentiviral Transduction

LNCaP and PC3 cells were cultured in RPMI (Gibco) with 10% FBS. VCAP cells were cultured in DMEM with 10% FBS. Cells stably expressing shRNAS were generated by seeding cells at 3×10^5 cells per well of a 6-well followed by transduction with lentivirus expressing given shRNAs. shRNAs in pLKO vectors were obtained from the Genetic

Perturbation Platform at the Broad Institute: sh*ERF*.1—shRNA TRCN0000349615: CCTGGTGTCTTCCGAGTCTAT; sh*ERF*.2—shRNA TRCN0000273970: CCACACCCAAAGCGTCTACAA; shRNA TRCN 0000013908: CCTGTCTCTGTGGGTTTCTAA; shRNA TRCN00000 13911: GAGGTGACTGACATCAGTGAT. Selection was performed subsequently with RPMI media with 10% FBS containing puromycin (2 μg/mL). Cell lines were kindly provided by Dr. William C. Hahn during 2014 to 2016 and were authenticated annually by DDC Medical or by the Molecular Diagnostics Laboratory at the Dana-Farber Cancer Institute using short tandem repeat profiling. The LHS-AR cell line was provided by Dr. William C. Hahn (33).

Western Blot Analysis

Cell lysates were prepared by lysing cells in 1% NP-40 with protease inhibitors (Roche) and phosphatase inhibitors (Calbiochem). Lysates were fractionated by SDS-polyacrylamide gel electrophoresis and transferred to nitrocellulose membranes using the iBlot system (Life Technologies). Immunoblotting was performed using LI-COR reagents (Odyssey Blocking Buffer and IRDye 800CW and IRDye 680RD secondary antibodies) according to the manufacturer's instructions (LI-COR Biosciences). Fluorescence detection was performed using an Odyssey CLx Infrared Imaging System, and quantitation was performed using Image Studio software (LI-COR). Anti-ERF antibody was purchased from Abcam (ab61108). Antibody for vinculin was purchased from Sigma (#V9131).

qRT-PCR

Total RNA was isolated using RNAeasy (Qiagen). cDNA was prepared using approximately 1 μg of RNA and the Superscript III Kit (Life Technologies). ERF transcript levels were quantified using SYBR green (Applied Biosystems) and measured using Quantstudio 6. The following primers were used for ERF:

ERF_1_FP 5′ GCA AGC CCC AGA TGA ATT ACG 3′
ERF_1_RP 5′ CCC CTT GGT CTT GTG CAG AA 3′

RNA Sequencing

Total RNA was isolated using RNAeasy (Qiagen), then processed with the NEBNext PolyA mRNA Magnetic Isolation Module (NEB, E7490), and then further processed with the NEBNext Ultra Directional RNA Library Prep Kit (NEB, E7420S). Libraries for RNA sequencing were then sequenced on a NextSeq (Illumina).

ERF Signature

To generate the ERF signatures, we analyzed the RPKM RNA-sequencing profiles of *ERF* shRNA knockouts in LNCaP and VCaP cell lines (see *ERF* shRNA experiments above). We independently ranked genes according to the difference of means between the sh*ERF*-infected LNCaP and control samples. The same procedure was performed for the VCaP cell line. To obtain a consensus sh*ERF* signature, we computed the overlap between the top 400 differentially expressed (upregulated and downregulated) genes in LNCaP and VCaP cell lines. This overlap resulted in 65 upregulated genes and 61 downregulated genes. These two gene sets are provided as part of the Supplementary Materials and will also be made publicly available as part of the C6 subcollection of the Molecular Signatures Database (MSigDB) in a future release (Supplementary Table S16; ref. 50).

Soft-Agar Assays

PC3 cells were transduced with lentivirus generated with pLKO-puro-shRNA plasmids. Twenty-four hours after infection, cells were selected with 2 μg/mL puromycin. After 48 hours of selection, cells were split for qRT-PCR assays and for passaging in normal serum-containing media for soft-agar assays. A total of 1×10^4 cells were suspended in 1 mL of 0.33% select agar in RPMI/FBS and plated on a bottom layer of 0.5% select agar in 6-well plates. Each cell line was analyzed in triplicate. Colonies were photographed after 11 days and quantified using CellProfiler.

CRISPR/Cas9 Experiments

Single-guide RNA guide sequences targeting the ERF coding region were cloned into the plentiCRISPR vector (pxpr001; ref. 36). sg4 ERF: CAC CGG GGT ACA TCG GGC TCA GCG T sg5 ERF: CAC CGG ATC CCC GCG CCC GAC CAC C; control guide sg5 GFP: GAAGTTC GAGGGCGACACCC. Lentivirus was produced in 293T cells and then LNCaP cells were lentivirally transduced followed by puromycin (2 μg/mL) selection. Western blot analysis was used to confirm *ERF* knockdown. LNCaP cells were seeded at a density of 1×10^4 cells in triplicate in 12-well plates in RPMI media supplemented with 10% charcoal-stripped serum (CSS; Gibco #12676). The following day cells were treated with R1881 synthetic androgen in RPMI/CSS media, and media were changed every 3 to 4 days. After 14 days, cells were fixed with 4% formaldehyde and stained with 0.5% crystal violet solution. Cells were photographed using a Leica microscope and imaging software. Quantification of crystal violet uptake for each sample was performed by destaining cells with 10% acetic acid and measurement of absorbance at 595 nm using a SpectraMax 190 instrument.

Animal Experiments

All animal experiments were approved by the Dana-Farber Cancer Institute Institutional Animal Care and Use Committee and were performed in accordance with institutional and national guidelines.

Accession Numbers

The accession number for sequencing files is dbGAP: phs000945.

Disclosure of Potential Conflicts of Interest

L.A. Garraway reports receiving commercial research grants from Astellas, BMS, Merck, and Novartis, has ownership interest (including patents) in Foundation Medicine, and is a consultant/advisory board member for Boehringer Ingelheim, Foundation Medicine, Novartis, Third Rock, and Warp Drive. No potential conflicts of interest were disclosed by the other authors.

Authors' Contributions

Conception and design: F.W. Huang, J.M. Mosquera, J.R. Osborne, E.M. Van Allen, M.A. Rubin, L.A. Garraway

Development of methodology: F.W. Huang, J.M. Mosquera, J. Chimene-Weiss, J.R. Osborne, J.W. Kim, M.A. Rubin, L.A. Garraway

Acquisition of data (provided animals, acquired and managed patients, provided facilities, etc.): F.W. Huang, J.M. Mosquera, M. Baco, B. Rabasha, S. Bahl, B.D. Robinson, F. Khani, J. Chimene-Weiss, G. Azabdaftari, A. Woloszynska-Read, A.M. De Marzo, S. Gabriel, M.A. Rubin, I.J. Powell, L.A. Garraway

Analysis and interpretation of data (e.g., statistical analysis, biostatistics, computational analysis): F.W. Huang, J.M. Mosquera, A. Garofalo, C. Oh, M. Baco, A. Amin-Mansour, S.A. Mullane, S. Aldubayan, E. Kim, M. Hofree, A. Romanel, F. Demichelis, E.M. Van Allen, J. Mesirov, P. Tamayo, M.A. Rubin, L.A. Garraway

Writing, review, and/or revision of the manuscript: F.W. Huang, J.M. Mosquera, A. Garofalo, M. Baco, A. Amin-Mansour, S.A. Mullane, B.D. Robinson, E. Kim, J.R. Osborne, A. Woloszynska-Read, E.M. Van Allen, J. Mesirov, M.A. Rubin, I.J. Powell, L.A. Garraway

Administrative, technical, or material support (i.e., reporting or organizing data, constructing databases): F.W. Huang, S. Bahl, B.D. Robinson, B. Karir, K. Sfanos, I.J. Powell, L.A. Garraway

Study supervision: F.W. Huang, J.M. Mosquera, S. Gabriel, M.A. Rubin, L.A. Garraway

Acknowledgments

We thank David Takeda, Rick Wilson, Ginevra Botta, Cory Johannessen, Paz Polak, and Manaswi Gupta for helpful discussions. We thank the Broad Genomics Platform. We thank Susan Bolton from Wayne State University; Himisha Beltran from WCM Division of Hematology and Medical Oncology and Englander Institute for Precision Medicine; Andrea Sboner from WCM Pathology and Laboratory Medicine, Institute for Computational Biomedicine, and Englander Institute for Precision Medicine; and Kyung Park, Peyman Tavassoli, Zohal Noorzad, and Jaclyn Croyle from WCM Pathology and Laboratory Medicine. The authors also thank Elena Pop and Elizabeth Brese from Roswell Park, and Helen Fedor from Johns Hopkins University.

Grant Support

This work was supported by the NCI U01 CA162148 (L.A. Garraway), the Department of Defense Prostate Cancer Research Program Physician Research Training Award W81XWH-14-1-0514 (F.W. Huang), Prostate Cancer Foundation Young Investigator Award (F.W. Huang), ASCO Young Investigator Award (F.W. Huang), the US NIH R01 CA116337 (M.A. Rubin), 5U01 CA111275-09 (J.M. Mosquera and M.A. Rubin), U54 HG003067 (S.Gabriel), R01CA154480 (J. Mesirov and P. Tamayo), R01CA121941 (J. Mesirov and P. Tamayo), U01CA176058 (J. Mesirov and P. Tamayo), R01CA109467 (J. Mesirov and P. Tamayo), U54 CA137788 (J.R. Osborne), the Starr Cancer Consortium (M.A. Rubin), and by a Stand Up To Cancer–Prostate Cancer Foundation Prostate Dream Team Translational Research Grant (Grant Number: SU2C-AACR-DT0712; M.A. Rubin and L.A. Garraway). Stand Up To Cancer is a program of the Entertainment Industry Foundation. Research grants are administered by the American Association for Cancer Research, the Scientific Partner of SU2C. This work was also supported in part by the Translational Research Program at WCM Pathology and Laboratory Medicine. This work was in part supported by the Department of Defense Prostate Cancer Research Program, DOD Award No. W81XWH-10-2-0056 and W81XWH-10-2-0046 PCRP Prostate Cancer Biorepository Network (PCBN). This work was supported by NCI grant P30CA016056 and the Pathology Network and Clinical Data Network Shared Resources at Roswell Park Cancer Institute.

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 August 29, 2016; revised February 22, 2017; accepted May 4, 2017; published OnlineFirst May 17, 2017.

REFERENCES

1. Taksler GB, Keating NL, Cutler DM. Explaining racial differences in prostate cancer mortality. Cancer 2012;118:4280–9.
2. Khani F, Mosquera JM, Park K, Blattner M, O'Reilly C, MacDonald TY, et al. Evidence for molecular differences in prostate cancer between African American and Caucasian men. Clin Cancer Res 2014;20:4925–34.
3. Petrovics G, Li H, Stumpel T, Tan SH, Young D, Katta S, et al. A novel genomic alteration of LSAMP associates with aggressive prostate cancer in African American men. EBioMedicine 2015;2:1957–64.
4. Lindquist KJ, Paris PL, Hoffmann TJ, Cardin NJ, Kazma R, Mefford JA, et al. Mutational landscape of aggressive prostate tumors in African American men. Cancer Res 2016;76:1860–8.
5. Cher ML, Lewis PE, Banerjee M, Hurley PM, Sakr W, Grignon DJ, et al. A Similar pattern of chromosomal alterations in prostate cancers from African-Americans and Caucasian Americans. Clin Cancer Res 1998;4:1273–8.
6. Rose AE, Satagopan JM, Oddoux C, Zhou Q, Xu R, Olshen AB, et al. Copy number and gene expression differences between African American and Caucasian American prostate cancer. J Transl Med 2010;8:70.
7. Shi Y, Li J, Zhang S, Wang M, Yang S, Li N, et al. Molecular epidemiology of EGFR mutations in Asian patients with advanced non-small-cell lung cancer of adenocarcinoma histology - Mainland China Subset Analysis of the PIONEER study. PLoS One 2015;10:e0143515.
8. Spratt DE, Chan T, Waldron L, Speers C, Feng FY, Ogunwobi OO, et al. Racial/ethnic disparities in genomic sequencing. JAMA Oncol 2016;2:1070–4.
9. Baca SC, Prandi D, Lawrence MS, Mosquera JM, Romanel A, Drier Y, et al. Punctuated evolution of prostate cancer genomes. Cell 2013;153:666–77.
10. Barbieri CE, Baca SC, Lawrence MS, Demichelis F, Blattner M, Theurillat JP, et al. Exome sequencing identifies recurrent SPOP, FOXA1 and MED12 mutations in prostate cancer. Nat Genet 2012;44:685–9.
11. Cancer Genome Atlas Research Network. The molecular taxonomy of primary prostate cancer. Cell 2015;163:1011–25.
12. 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 Cancer J Clin 2016;66:290–308.
13. Taylor BS, Schultz N, Hieronymus H, Gopalan A, Xiao Y, Carver BS, et al. Integrative genomic profiling of human prostate cancer. Cancer Cell 2010;18:11–22.
14. Mermel CH, Schumacher SE, Hill B, Meyerson ML, Beroukhim R, Getz G. GISTIC2.0 facilitates sensitive and confident localization of the targets of focal somatic copy-number alteration in human cancers. Genome Biol 2011;12:R41.
15. Migita T, Ruiz S, Fornari A, Fiorentino M, Priolo C, Zadra G, et al. Fatty acid synthase: a metabolic enzyme and candidate oncogene in prostate cancer. J Natl Cancer Inst 2009;101:519–32.
16. Mateo J, Carreira S, Sandhu S, Miranda S, Mossop H, Perez-Lopez R, et al. DNA-repair defects and olaparib in metastatic prostate cancer. 2015;373:1697–708.
17. Lawrence MS, Stojanov P, Polak P, Kryukov GV, Cibulskis K, Sivachenko A, et al. Mutational heterogeneity in cancer and the search for new cancer-associated genes. Nature 2013;499:214–8.
18. Tomlins SA, Rhodes DR, Perner S, Dhanasekaran SM, Mehra R, Sun X, et al. Recurrent fusion of TMPRSS2 and ETS transcription factor genes in prostate cancer. Science 2005;310:644–8.
19. Adzhubei IA, Schmidt S, Peshkin L, Ramensky VE, Gerasimova A, Bork P, et al. A method and server for predicting damaging missense mutations. Nat Methods 2010;7:248–9.
20. Beltran H, Prandi D, Mosquera JM, Benelli M, Puca L, Cyrta J, et al. Divergent clonal evolution of castration-resistant neuroendocrine prostate cancer. Nat Med 2016;22:298–305.
21. Beltran H, Eng K, Mosquera JM, Sigaras A, Romanel A, Rennert H, et al. Whole-exome sequencing of metastatic cancer and biomarkers of treatment response. JAMA Oncol 2015;1:466–74.
22. 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.
23. Kumar A, Coleman I, Morrissey C, Zhang X, True LD, Gulati R, et al. Substantial interindividual and limited intraindividual genomic diversity among tumors from men with metastatic prostate cancer. Nat Med 2016;22:369–78.
24. Barretina J, Caponigro G, Stransky N, Venkatesan K, Margolin AA, Kim S, et al. The Cancer Cell Line Encyclopedia enables predictive modelling of anticancer drug sensitivity. Nature 2012;483:603–7.
25. Forbes SA, Bindal N, Bamford S, Cole C, Kok CY, Beare D, et al. COSMIC: Mining complete cancer genomes in the catalogue of somatic mutations in cancer. Nucleic Acids Res 2011;39(Database issue):D945–50.
26. Sgouras DN, Athanasiou MA, Beal GJ, Fisher RJ, Blair DG, Mavrothalassitis GJ. ERF: an ETS domain protein with strong transcriptional repressor activity, can suppress ets-associated tumorigenesis and is regulated by phosphorylation during cell cycle and mitogenic stimulation. EMBO J 1995;14:4781–93.
27. Le Gallic L, Sgouras D, Beal G Jr, Mavrothalassitis G. Transcriptional repressor ERF is a Ras/mitogen-activated protein kinase target that regulates cellular proliferation. Mol Cell Biol 1999;19:4121–33.

28. Giannakis M, Mu XJ, Shukla SA, Qian ZR, Cohen O, Nishihara R, et al. Genomic correlates of immune-cell infiltrates in colorectal carcinoma. Cell Rep 2016;17:1206.
29. Crompton BD, Stewart C, Taylor-Weiner A, Alexe G, Kurek KC, Calicchio ML, et al. The genomic landscape of pediatric Ewing sarcoma. Cancer Discov 2014;4:1326–41.
30. Cerami E, Gao J, Dogrusoz U, Gross BE, Sumer SO, Aksoy BA, et al. The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data. Cancer Discov 2012;2:401–4.
31. Chen Y, Chi P, Rockowitz S, Iaquinta PJ, Shamu T, Shukla S, et al. ETS factors reprogram the androgen receptor cistrome and prime prostate tumorigenesis in response to PTEN loss. Nat Med 2013;19:1023–9.
32. Baena E, Shao Z, Linn DE, Glass K, Hamblen MJ, Fujiwara Y, et al. ETV1 directs androgen metabolism and confers aggressive prostate cancer in targeted mice and patients. Genes Dev 2013;27:683–98.
33. Berger R, Febbo PG, Majumder PK, Zhao JJ, Mukherjee S, Signoretti S, et al. Androgen-induced differentiation and tumorigenicity of human prostate epithelial cells. Cancer Res 2004;64:8867–75.
34. Nelson PS, Clegg N, Arnold H, Ferguson C, Bonham M, White J, et al. The program of androgen-responsive genes in neoplastic prostate epithelium. Proc Natl Acad Sci U S A 2002;99:11890–5.
35. Sharma NL, Massie CE, Ramos-Montoya A, Zecchini V, Scott HE, Lamb AD, et al. The androgen receptor induces a distinct transcriptional program in castration-resistant prostate cancer in man. Cancer Cell 2013;23:35–47.
36. Shalem O, Sanjana NE, Hartenian E, Shi X, Scott DA, Mikkelsen TS, et al. Genome-scale CRISPR-Cas9 knockout screening in human cells. Science 2014;343:84–7.
37. Sheridan CM, Grogan TR, Nguyen HG, Galet C, Rettig MB, Hsieh AC, et al. YB-1 and MTA1 protein levels and not DNA or mRNA alterations predict for prostate cancer recurrence. Oncotarget 2015;6:7470–80.
38. Tosoian JJ, Almutairi F, Morais CL, Glavaris S, Hicks J, Sundi D, et al. Prevalence and prognostic significance of PTEN loss in African-American and European-American men undergoing radical prostatectomy. Eur Urol 2017;71:697–700.
39. Romanel A, Zhang T, Elemento O, Demichelis F. EthSEQ: ethnicity annotation from whole exome sequencing data. Bioinformatics. 2017 Mar 27. [Epub ahead of print].
40. Dissanayake K, Toth R, Blakey J, Olsson O, Campbell DG, Prescott AR, et al. ERK/p90(RSK)/14-3-3 signalling has an impact on expression of PEA3 Ets transcription factors via the transcriptional repressor capicua. Biochem J 2011;433:515–25.
41. Magi-Galluzzi C, Tsusuki T, Elson P, Simmerman K, LaFargue C, Esgueva R, et al. TMPRSS2-ERG gene fusion prevalence and class are significantly different in prostate cancer of Caucasian, African-American and Japanese patients. Prostate 2011;71:489–97.
42. Blattner M, Lee DJ, O'Reilly C, Park K, MacDonald TY, Khani F, et al. SPOP mutations in prostate cancer across demographically diverse patient cohorts. Neoplasia 2014;16:14–20.
43. Faisal FA, Sundi D, Tosoian JJ, Choeurng V, Alshalalfa M, Ross AE, et al. Racial variations in prostate cancer molecular subtypes and androgen receptor signaling reflect anatomic tumor location. Eur Urol 2016;70:14–7.
44. Yamoah K, Johnson MH, Choeurng V, Faisal FA, Yousefi K, Haddad Z, et al. Novel biomarker signature that may predict aggressive disease in African American men with prostate cancer. J Clin Oncol 2015;33:2789–96.
45. Grasso CS, Wu YM, Robinson DR, Cao X, Dhanasekaran SM, Khan AP, et al. The mutational landscape of lethal castration-resistant prostate cancer. Nature 2012;487:239–43.
46. Duggan D, Zheng SL, Knowlton M, Benitez D, Dimitrov L, Wiklund F, et al. Two genome-wide association studies of aggressive prostate cancer implicate putative prostate tumor suppressor gene DAB2IP. J Natl Cancer Inst 2007;99:1836–44.
47. Rand KA, Rohland N, Tandon A, Stram A, Sheng X, Do R, et al. Whole-exome sequencing of over 4100 men of African ancestry and prostate cancer risk. Hum Mol Genet 2016;25:371–81.
48. Manrai AK, Funke BH, Rehm HL, Olesen MS, Maron BA, Szolovits P, et al. Genetic misdiagnoses and the potential for health disparities. N Engl J Med 2016;375:655–65.
49. Chakravarty D, Sboner A, Nair SS, Giannopoulou E, Li R, Hennig S, et al. The oestrogen receptor alpha-regulated lncRNA NEAT1 is a critical modulator of prostate cancer. Nat Commun 2014;5:5383.
50. Liberzon A, Subramanian A, Pinchback R, Thorvaldsdottir H, Tamayo P, Mesirov JP. Molecular signatures database (MSigDB) 3.0. Bioinformatics 2011;27:1739–40.