Published in final edited form as:

Hum Mutat. 2013 September ; 34(9): 1231–1241. doi:10.1002/humu.22346.

Somatic alterations contributing to metastasis of a castration resistant prostate cancer

Michael L. Nickerson^{1,*}, Kate M. Im¹, Kevin J. Misner¹, Wei Tan², Hong Lou², Bert Gold¹, David W. Wells³, Hector C. Bravo⁴, Karin M. Fredrikson⁵, Timothy T. Harkins⁵, Patrice Milos⁶, Berton Zbar¹, W. Marston Linehan⁷, Meredith Yeager⁸, Thorkell Andresson⁹, Michael Dean^{1,*}, and G. Steven Bova^{10,11}

¹Cancer and Inflammation Program, National Cancer Institute, National Institutes of Health, Frederick, Maryland ²Basic Science Program, SAIC-Frederick, National Cancer Institute, National Institutes of Health, Frederick, Maryland ³Genetics Core, Center for Cancer Research, National Cancer Institute, National Institutes of Health, Frederick, Maryland ⁴Center for Bioinformatics and Computational Biology, Department of Computer Science, University of Maryland, College Park, Maryland ⁵Roche Diagnostics Corporation, Indianapolis, Indiana ⁶Helicos BioSciences Corporation, Cambridge, MA ⁷Urologic Oncology Branch, National Cancer Institute, National Institutes of Health, Bethesda, Maryland ⁸Division of Cancer Epidemiology and Genetics, National Cancer Institute, National Institutes of Health, Bethesda, Maryland ⁹Laboratory of Proteomics and Analytical Technologies, Advanced Technology Program, SAIC-Frederick, National Cancer Institute, National Institutes of Health, Frederick, Maryland ¹⁰Department of Pathology, School of Medicine, Johns Hopkins University, Baltimore, Maryland ¹¹Institute of Biomedical Technology, University of Tampere, Tampere, Finland

Abstract

Metastatic castration resistant prostate cancer (mCRPC) is a lethal disease and molecular markers that differentiate indolent from aggressive subtypes are needed. We sequenced the exomes of five metastatic tumors and healthy kidney tissue from an index case with mCRPC to identify lesions associated with disease progression and metastasis. An Ashkenazi Jewish (AJ) germline founder mutation, del185AG in *BRCA1*, was observed and AJ ancestry was confirmed. Sixty-two somatic variants altered proteins in tumors, including cancer-associated genes, *TMPRSS2-ERG*, *PBRM1*, and *TET2*. The majority (n=53) of somatic variants were present in all metastases and only a subset (n=31) was observed in the primary tumor. Integrating tumor next generation sequencing (NGS) and DNA copy number showed somatic loss of *BRCA1* and *TMPRSS2-ERG*. We sequenced 19 genes with deleterious mutations in the index case in additional mCRPC samples and detected a frameshift, two somatic missense alterations, tumor loss of heterozygosity (LOH), and combinations of germline missense SNPs in *TET2*. In summary, genetic analysis of metastases from an index case permitted us to infer a chronology for the clonal spread of disease based on sequential accrual of somatic lesions. The role of *TET2* in mCRPC deserves additional analysis and may define a subset of metastatic disease.

*Correspondence to: Michael L. Nickerson, PhD, Cancer and Inflammation Program, National Cancer Institute, Building 560, Rm. 21-21, Frederick, MD 21702. Phone: 301-846-6973; Fax: 301-846-7042; nickersonml@mail.nih.gov; and Michael Dean, PhD, Cancer and Inflammation Program, National Cancer Institute, Building 560, Rm. 21-89B, Frederick, MD 21702. Phone: 301-846-5931; Fax: 301-846-7042; deanm@mail.nih.gov.

Supporting Information for this preprint is available from the *Human Mutation* editorial office upon request (humu@wiley.com)

Conflicts of Interest: The authors have no conflicts to disclose.

Keywords

tumor heterogeneity; somatic mutation; metastasis; epigenetic modifiers; *BRCA1*; *TMPRSS2*; *ERG*; *PBRM1*; *TET2*

Introduction

Metastatic prostate cancer is a lethal disease poorly controlled using existing therapies and responsible for over 258,000 deaths worldwide each year [Garcia et al., 2007; Jemal et al., 2011]. The precise molecular chronology from prostate cancer (PC) to castration resistant PC (CRPC) to metastatic CRPC (mCRPC) remains unknown. This is in part due to confounding effects from the multifocal etiology and tumor heterogeneity of primary prostate adenocarcinomas and also a lack of genetic data characterizing somatic alterations in metastatic tumors. Several groups recently analyzed PC by genomic copy number and targeted re-sequencing of candidate genes [Taylor et al., 2010; Robbins et al., 2011] and by exome and whole genome sequencing of CRPC [Berger et al., 2011; Kumar et al., 2011; Barbieri et al., 2012; Grasso et al., 2012]. Most of these NGS studies examined primary tumors where the presence of stroma admixed with tumor may confound identification of somatic alterations. In the present study, we sought to examine the chronology of genetic changes occurring in a cancer genome during progression from PC to mCRPC in an index patient who died of the disease. For this purpose, we used exome capture followed by NGS on DNA from five mCRPC tumors and healthy kidney tissue. We validated selected germline and all NGS predicted nonsynonymous somatic variants and integrated the sequence changes with copy number analysis of the same tumors. We used computational inference methods to generate a chronologic and anatomic map of the spread of disease in the index case. This 'high-resolution' study of disease by collection and analysis of 13 samples from a single case was necessary to identify alterations in cancer genes that were acquired sequentially during disease progression.

Materials and Methods

Human Subjects and DNA

mCRPC subjects consented to participate in the Project to Eliminate Lethal Prostate Cancer, an integrated clinical-genomic autopsy study of PC under Johns Hopkins Medicine Institutional Review Board approval (Supp. Table S1). Patient age at diagnosis has been altered $\pm$ 1 to 5 years in accordance with patient consent. Tissue samples obtained during rapid autopsy were frozen and cryostat-microdissected before DNA was isolated using proteinase K digestion followed by phenol/chloroform extraction. DNA samples were obtained from eleven metastatic tumors from the index patient including the five tumors (abdominal paraaortic lymph node (16472), right supraclavicular lymph node (16473), right femur marrow (16474), subdural fossa C (16475), and left axillary lymph node (16476) metastases) that were subjected to exome sequencing. DNA was also extracted from a cross-section of the non-microdissected, paraffin-embedded primary adenocarcinoma, but was not subjected to NGS.

Exome Capture and Sequencing

Exome capture libraries were prepared using 5 μ g DNA and the Sequence Capture 2.1M Human Exome Array and standard protocols (NimbleGen, Madison, WI) [Albert et al., 2007; Hodges et al., 2007; Okou et al., 2007]. Probes for capture of $\sim$ 180,000 protein-coding exons were designed based upon curated genes in the Consensus Coding Sequence

(CCDS) database [Pruitt et al., 2009]. Libraries were sequenced using a 454 Genome Sequencer FLX system and Titanium reagent chemistry (Roche/454; Branford, CT).

Single Molecule Genome Sequencing

100 ng DNA from a formalin-fixed, paraffin embedded section of a metastasis was tailed with terminal transferase, hybridized for sequencing, and loaded onto three channels of a HeliScope (Helicos Biosciences, Cambridge, MA). The yield was 10.9 million unique and 7.7 million aligned reads which were placed in 20 kb bins to determine copy number.

Analysis of NGS Data

Sequences from multiple (2.5–3.5) picotiterplates were combined and aligned to the NCBI build 36/hg18 human genome assembly using the Genome Sequencer Reference Mapper program (GS Mapper, Roche) and default alignment parameters. This yielded 1.96–3.6 million uniquely mapped reads and 1.4–2.6 million reads that mapped to exome target regions, including splice junctions. Mean depth was estimated using read bases aligning “In Region” divided by the total target region bases. Default GS Mapper settings and the following criteria were used for variant identification: reads starting at the same genomic coordinate were considered duplicates and were removed, and a variant must be observed in both forward and reverse reads, a minimum of three non-duplicate reads, and in >20% of total reads across a given nucleotide position. Annotated variants were sorted based on known variants (dbSNP130 and other personal genomic variants), synonymous variants, non-coding variants found outside of splice junctions (>10bp from an intron–exon boundary), and novel variants that alter proteins or splice junctions (>10 bp from an intron–exon boundary).

NGS read alignments were viewed using the Integrated Genome Browser (IGB; Affymetrix, Santa Clara, CA) [Nicol et al., 2009], the Integrated Genomics Viewer (IGV; Broad Institute, Cambridge, MA) [Robinson et al., 2011], and other programs, including the Basic Local Alignment Search Tool (BLAST, the National Center for Biotechnology Information, Bethesda, MD), the BLAST-Like Alignment Tool (BLAT) [Kent, 2002], and the In-Silico PCR program in the UCSC Genome Browser [Rhead et al., 2010].

Integrated NGS and Copy Number

Affymetrix SNP6 single nucleotide polymorphism (SNP) and array Comparative Genomic Hybridization (aCGH) copy number data was available from index case metastases as previously described [Liu et al., 2009]. This data was not available for the primary adenocarcinoma. Copy number status at each gene altered by a somatic mutation was analyzed by Fisher exact test. The data was visualized using a Circos plot [Krzywinski et al., 2009].

Analyses of Variants

Variants were analyzed by manual annotation of proteins and review of curated websites such as UniProt (The UniProt Consortium, 2010) and the published literature, and by assigning arbitrary points to analysis program outputs to rank candidate genes for additional analysis (details available upon request). SIFT [Kumar et al., 2009], POLYPHEN [Adzhubei et al., 2010], Mutation Taster [Schwarz et al., 2010], and CHASM [Carter et al., 2009] were run using standard settings. Genes were examined in Gene2MeSH (<http://gene2mesh.ncibi.org/>), COSMIC [Forbes et al., 2008], and for expression in prostate tissue using the Unigene DB. Interacting proteins were characterized using the Michigan Molecular Interactions (MiMI) database (<http://mimi.ncibi.org/MimiWeb/main-page.jsp>) [Jayapandian et al., 2007; Tarcea et al., 2009]. Ingenuity Variant Analysis

(www.ingenuity.com/variants) was leveraged for biological interpretation of somatic variants, and an interactive online supplement of all somatic variants summarized in Table 1 is available at: https://variants.ingenuity.com/Nickerson_et al.

Variants described in this study have been deposited in COSMIC (accession # COSP31410) and dbSNP.

PCR and Sanger Sequencing

Primers designed using Primer 3 [Rozen and Skaletsky, 2000] were synthesized by Invitrogen (Carlsbad, CA). Primers for specific variants are available upon request. FastStart PCR reagents (Roche; South San Francisco, CA) and GeneAmp 9700 PCR System (ABI; Foster City, CA) thermal cyclers were used for PCR and sequencing. Whole genome amplification (WGA, Qiagen, Valencia, CA) required 100 ng tumor DNA and standard PCR conditions at an annealing temperature of 58°C. Big Dye v.3.1 chemistry (ABI) sequencing reactions were subjected to chromatography on a 3730 Genetic Analyzer (ABI) and chromatograms were examined using Sequencher, v.4.8 (GeneCodes, Ann Arbor, MI) and Mutation Surveyor (Softgenetics, State College, PA). Somatic alterations confirmed in metastatic tumor exomes were examined by duplicate (or triplicate if sufficient DNA was available), independent PCRs and double stranded Sanger sequencing of DNA extracted from the primary adenocarcinoma of the index patient.

Reverse Transcriptase-PCR (RT-PCR): Approximately 1 µg of total RNA was reverse-transcribed using SuperScript III (Invitrogen, Carlsbad, CA) and random hexamers in a total volume of 20 µl. cDNA was diluted 1:100 and 5 µl was PCR-amplified in a 50 µl final reaction volume using Platinum High Fidelity Taq DNA polymerase (Invitrogen, Carlsbad, CA). PCR conditions were: denaturation for 30 seconds at 94°C, annealing for 1 min at 60°C, and extension for 2 min at 68°C, for 35 cycles with a final 10 minute extension at 68°C. PCR products were examined in 1% agarose in 0.5x TBE, stained with ethidium bromide, and visualized under ultraviolet light. Primers for the analysis of *TMPRSS2* and *ERG* are available upon request.

Oncogenetic Tree

We constructed several oncogenetic trees [Desper et al., 1999] using nine variants detected in six metastases and the most parsimonious example is shown.

Results

Patient History and Samples

The index case, subject A17 (Supp. Table S1), was a Caucasian, male nonsmoker diagnosed at age 58 with an elevated serum prostate-specific antigen (PSA) of 5.1 ng/mL. Despite treatment with continuous luteinizing hormone-releasing hormone agonist therapy, his PSA level increased over three years to 9,940 ng/mL ten weeks prior to death. His primary tumor was removed one year after diagnosis and the composite radical retropubic prostatectomy (RRP) Gleason score was 4+4=8. DNA was isolated from sections from one RRP block, which was estimated to contain 70% noncancerous cells, 15% Gleason grade 4 prostate cancer, 10% Gleason grade 3 prostate cancer, and 5% Gleason grade 5 prostate cancer based on analysis of adjacent H&E stained sections. Metastatic tumors were isolated approximately 2.5 years later, after death, as part of an integrated clinical-molecular autopsy study and were 60–98% estimated cancer purity. No DNA-damaging chemotherapy was administered during treatment. Thus, metastatic tissues likely represent mCRPC.

Exome Sequencing and Variant Validation

Exome capture from five metastatic tumors and control kidney tissue targeted ~180,000 protein-coding exons (NimbleGen, Madison, WI). An average of over three million reads and 1.13 Gigabases per sample were sequenced on a FLX Genome Sequencer (Roche/454, Branford, CT) to a mean depth of 29x coverage over 97% of the target region (Supp. Table S2). Between 6,800 (metastatic tumor 1) and 12,000 (control kidney) variants were identified, including 646 novel variants observed in at least one metastasis and not in control kidney tissue (range 67–225, average 129). Two hundred twenty three novel variants altering proteins or splice junctions not present in dbSNP130 or as a personal genome variant (UCSC genome browser) in genome-build HG18 were identified for validation. Manual review of NGS alignments in IGV prompted removal of 27 variants due to alignment errors (data not shown). The remaining 196 putative somatic variants were examined by PCR and Sanger re-sequencing in control kidney and metastatic tissue DNA resulting in 57 (29%) false positives, 77 (39%) novel germline nonsynonymous polymorphisms, and 62 (32%) somatic nonsynonymous alterations (Supp. Table S3, S4). Somatic variants included four insertions or deletions (indels) causing frameshifts, one nonsense, 29 nonconservative missense, and four intronic splice junctions alterations within 10 bp of exons. Exome sequencing was integrated with copy number datasets (SNP array and aCGH) [Liu et al., 2009] from the same metastases (Supp. Figure S1). We found no significant association between the presence of a somatic sequence alteration and a copy number alteration within a given gene (p-value = 0.45, Fisher's exact test).

We compared the 62 genes with somatic alterations in the index patient to results from recently published studies characterizing sequence variation in PC [Taylor et al., 2010; Berger et al., 2011; Robbins et al., 2011; Barbieri et al., 2012]. In total, 43% of genes with somatic alterations in the index patient were mutated in at least one other study (available upon request).

Alterations in Genes Associated with Cancer

NGS of control kidney tissue identified a heterozygous AJ germline founder mutation (c. 185_186delAG, rs80357713, Fig. 1A) in *BRCA1* previously linked to breast, ovarian, and prostate cancer [Tonin et al., 1996; Struwing et al., 1997; King et al., 2003]. A review of patient records revealed self-reported Ashkenazi Jewish (AJ) ancestry in both parents and diagnoses of breast, colon, pancreatic, and melanoma cancers, but not PC, in immediate family members. Thus, the patient belongs to a likely new cancer family but samples from additional family members were unavailable for genetic analysis. NGS read counts, Sanger re-sequencing, and aCGH show a homozygous mutant *BRCA1* in all 11 metastases, indicating somatic LOH of the wild type (wt) allele (Fig. 1A, B). These results suggest both germline and somatic loss of *BRCA1* are associated with this case of CRPC and *BRCA1* loss continued to be selected for as the disease progressed to mCRPC.

PBRM1 was altered by a somatic variant (Fig. 1C) introducing a nonconservative substitution (p.M672R) of a highly conserved methionine in the primary adenocarcinoma and all metastases (Fig. 1D). *PBRM1* has previously been shown to be altered by somatic mutations and/or LOH in breast, kidney, and prostate cancers [Xia et al., 2008; Varela et al., 2011; Berger et al., 2011]. The variant alters a bromodomain (BRD), a domain that binds acetylated lysine on histones and transcription factors to modulate gene transcription (Fig. 1E). BRD5 preferentially binds acetylated lysines 36 and 122 on histone 3 as part of the polybromo, BRG1-associated factor (PBAF) subgroup of the mating type switch/sucrose nonfermenting-B (SWI/SNF-B) complex [Thompson, 2009; Charlop-Powers et al., 2010]. The substitution occurs in the acetylated lysine binding pocket near catalytic di-tyrosine residues next to an essential proline between alpha helices 3 and 4 found in all BRDs (Fig.

1E). Altered binding to acetylated lysines on histones may disrupt chromatin remodeling, transcription, or cell cycle progression [Roberts and Orkin, 2004]. Somatic missense and frameshift alterations of *PBRM1* were observed in three other studies suggesting a low but significant frequency of somatic alteration in PCa [Berger et al., 2011; Barbieri et al., 2012; Grasso et al., 2012].

TMPRSS2 was altered by a somatic 10 bp heterozygous insertion in exon 10 (Fig. 2A, Supp. Table S5) in all metastases, yet the alteration was not detected in the primary tumor. The insertion was a duplication predicted to cause a frameshift and truncation of the protein, which deleted the majority of the catalytic domain. *TMPRSS2* is known to form fusion transcripts in ~50% of primary prostate tumors, most commonly with *ERG* [Tomlins et al., 2005]. Examination of copy number across *TMPRSS2* and *ERG* loci in metastatic tumors (Fig. 2B) revealed a ~3 Mb somatic deletion of chromosome 21q in all five metastases, consistent with previous reports that this deletion may “drive” the malignancy once it occurs. We sought to use NGS reads to demarcate the deletion borders in greater detail by subjecting tumor DNA to low pass (3–4x coverage) single-molecule genome sequencing using a HeliScope (Helicos BioSciences, Cambridge, MA, details available upon request). Although not sufficient for high confidence variant calling, reads aligning to the genome in 25 kb increments were binned and visualized to show an ~50% reduction in NGS reads in the region of loss. This permitted us to deduce borders for the LOH in *TMPRSS2* intervening sequence (IVS)-11 and *ERG* IVS-1 (Fig. 2B, middle and bottom). RT-PCR of metastasis RNA and sequencing revealed the junction of the fusion transcript between *TMPRSS2* exon 1 and *ERG* exon 4. This implies that splicing to create the fusion transcript skipped intact exons of each gene lying outside the region of loss (Fig. 2B). RT-PCR of metastatic tumor RNA (Fig. 2C) using *TMPRSS2* primers located 5′ and 3′ of the somatic 10 bp insertion revealed expression of the transcript containing the insertion; however, no product was obtained using primers in the *TMPRSS2* insertion and primers 3′ of the fusion transcript junction in *ERG*. Taken together, these results indicate the fusion transcript and insertion are present on separate alleles (Fig. 2D). Thus, both copies of *TMPRSS2* were altered by somatic genetic events, including LOH favouring formation of a fusion transcript that likely occurred in the primary tumor followed by an insertion that truncated the remaining protein in all metastases.

TET2 was altered in all metastases but not in the primary tumor by a somatic C>G encoding a nonconservative substitution of an alanine for proline 562 (p.P562A), which is phylogenetically conserved (Fig. 3A–C). Somatic alteration of *TET2* has been implicated in the pathogenesis of 15–20% of myeloproliferative, myelodysplastic, and myeloid leukemia [Delhommeau et al., 2009; Nibourel et al., 2010; Schaub et al., 2010].

Somatic Alterations and Disease Progression

We examined the 62 nonsynonymous, somatic variants observed in exome sequences of five metastases in six additional metastases from the index case by duplicate, independent PCRs and Sanger re-sequencing (Supp. Table S5). Most variants (n=53, 85%) were present in all metastases (‘common metastatic’ variants, Table 1), indicating they arose in a primary tumor cell that expanded in a clonal fashion to give rise to 11 metastases. However, nine variants (15%) were observed in a subset of metastases (‘late metastatic’ variants, Table 1). Thus, somatic sequence alterations continued to accumulate in individual metastasis and/or tumor lineages after metastatic expansion was initiated, as stochastic passenger variants or in response to new or continued selective pressure as tumor cells dispersed to distant anatomic locations (Fig. 4A).

We mapped the anatomic location of metastases on the basis of the observed distribution of somatic variants. Anatomic locations imputed as occupied early in metastatic progression,

such as the abdominal paraaortic lymph node and femur bone marrow (tumors 1 and 3, respectively, Fig. 4B), possessed metastases with only common metastatic variants. Anatomic sites imputed as colonized late in metastatic progression, such as supraclavicular and axillary lymph nodes, and brain, contained lesions with additional late metastatic variants. These genetic results recapitulate a chronologically defined order of metastatic spread that is in agreement with bone and computerized tomography scans, and physical examination of the patient.

Somatic alterations observed in metastasis exomes were examined by duplicate, independent PCRs and double stranded Sanger sequencing of DNA extracted from a cross-section of non-microdissected, paraffin-embedded primary adenocarcinoma. Of the 53 somatic, common metastatic variants, only 22 (42%) were detected (Table 1, Supp. Table S5) despite PCRs in triplicate for negative variants, and none of the nine somatic, late metastatic variants were detected. The location of cells containing the remainder of the common metastatic variants is of great interest as these gave rise to all metastatic tumors, and these could reside in a different region of the primary tumor (not analyzed), or in a physically distinct metastatic niche hypothetically colonized early in the spread of disease.

Somatic variants that were detected in the primary tumor displayed large differences in signal intensity relative to wt bases in Sanger sequencing chromatograms (Fig. 4C). Relative Signal Intensity (RSI) values are visual (qualitative) estimates of somatic mutation versus total signal (mutation peak height / mutation + wt nucleotide peak heights) in Sanger chromatograms, and were previously shown to be a reliable indicator of mutant allele frequency, tumor cell number, and tumor:normal tissue ratios in heterogeneous cancer samples [Nickerson et al., 2008]. Sequencing showed a somatic *EIF1AX* gene variant with a RSI = 0.7, indicating ~70% of cells in the primary tumor section possessed the variant. Conversely, PCR and sequencing in triplicate confirmed a low RSI = 0.1 for the somatic variant in *PBRM1*, suggesting few primary tumor cells (~10%) contained this alteration. Thus, within a single, 'typical' section of primary tumor, signal from somatic mutations that were detected were highly variable.

Evaluation of Altered Genes in Additional mCRPC Cases

Somatic variants and altered genes were evaluated as potential cancer drivers using multiple programs as described in the Methods (Supp. Table S4). A risk score was developed to prioritize genes for further analysis and protein coding regions of eighteen genes (gray, Supp. Table S4) were examined by PCR of WGA DNA and re-sequencing in a validation panel of 29 mCRPC metastases, each from an independent case (Supp. Table S1). Additional somatic nonsynonymous variants were observed only in *TET2* as described below.

Alteration of *TET2* in mCRPC

We sequenced *TET2* in 29 additional metastases and seven PC cell lines and observed three novel missense alterations and a 1-bp insertion in DU-145, a cell line derived from a mPC brain tumor [Stone et al., 1978; Mickey et al., 1980], which introduced a frameshift resulting in truncation of 89% of the protein (Table 2, Fig. 3, Supp. Table S6). Two of the variants were confirmed as somatic using matched control kidney or blood DNA. The somatic variants were nonconservative substitutions that altered highly conserved residues (Fig. 3C). All three patients with a somatic *TET2* alteration were Caucasian men diagnosed in their fifth decade (average 56 years of age) who survived an average of only 3.8 years. DU-145 was derived from a 69 year old Caucasian male with a history of lymphocytic leukemia who expired six months after diagnosis [Mickey et al., 1980].

The analysis above revealed that *TET2* was frequently altered by germline missense SNPs as well. A third novel alteration encoded an alanine for aspartic acid 551 (p.D551A) in the germline of an African-American man, diagnosed at age 56, who survived 6 years, from a family in which his father and brother were diagnosed with PC and his mother died of leukemia. This variant was not observed in DNA from over 250 healthy donors recruited at the National Cancer Institute, 629 individuals from the 1000 Genomes Project [2010], or in healthy controls from published studies of *TET2* in leukemia [Delhommeau et al., 2009; Nibourel et al., 2010]. We observed a surprising number of additional germline SNPs, including 13 encoding missense alterations and an intronic, splice variant (Fig. 3B; Supp. Tables S7, S8). These were observed in 21 different combinations in PC subjects and may be significant to PC risk as SNP rs7679673, located ~5 kb upstream of the *TET2* transcription start site, is associated with PC risk by genome-wide association study (GWAS) [Eeles et al., 2009], as are additional intronic variants in the 5' end of the gene (data not shown). Lastly, we re-analyzed aCGH data from tumors from all 30 mCRPC patients [Liu et al., 2009] and observed LOH encompassing *TET2* in a surprising 19/30 (63%), including tumors with focal losses specific for *TET2* (Fig. 3D).

Discussion

We examined the genetic basis of a mCRPC using exome sequencing of five anatomically distinct metastases and matched noncancerous kidney tissue integrated with genomic copy number from a single index case. Metastatic tumors exhibited 62 somatic nonsynonymous alterations in addition to a common AJ founder mutation in *BRCA1*. Genes identified as altered in cancer in the literature manifested sequence changes and LOH at different stages of this patient's disease. Sequentially accrued somatic alterations included *BRCA1* LOH, a *TMPRSS2-ERG* double-hit, and missense changes to genes encoding epigenetic-associated proteins, *PBRM1* and *TET2*. These data provide convincing evidence supporting multistage models of carcinogenesis whereby alteration of multiple cancer genes have been speculated to underlie cancer [Vogelstein and Kinzler, 2004; Hornsby et al., 2007; Hanahan and Weinberg, 2011; Stratton, 2011].

Our data shows the somatic alteration of *PBRM1* (and other genes, Fig. 4) in the primary tumor from the index case possessed a low signal in sequence chromatograms (RSI=0.1) compared to signal of the same variant in metastases (RSI=0.5) suggesting that a *PBRM1* mutation might be missed in a routine analysis of a primary prostate tumor. The results suggest mutation analysis of a single primary tumor section may not provide a complete picture of somatic mosaicism within a sample such as was analyzed here.

We identified 62 somatic nonsynonymous alterations in metastases that accrued during disease progression. Twenty-two variants (35%) detected in the primary adenocarcinoma showed reproducible, different RSI estimates [Nickerson et al., 2008] providing experimental evidence for sequential accumulation of variants in the adenocarcinoma. Thirty-one variants (50%) not observed in the primary adenocarcinoma but detected in all metastases were attributed to a metastatic progenitor cell that gave rise to all subsequent metastases. The physical location of these genetically-defined metastatic progenitor cells is of great interest and more detailed somatic genetic studies might determine whether these arose in distal regions of the primary adenocarcinoma or in a distinct metastatic niche. Nine variants (15%) were detected in individual metastases indicating few additional somatic changes were required after metastatic dispersal had initiated. However, variants in individual metastases were useful to define a chronological order for the metastatic spread of disease (Fig. 4). Thus, variants in the index case were acquired at each stage of the disease, in agreement with a previous study that sequenced candidate genes from three metastases from a single patient [Robbins et al., 2011].

Our results indicate that somatic alteration of *TET2* occurs at a significant frequency in mCRPC (Table 2, Fig. 3). The gene encodes a protein belonging to a family of three related proteins, TET1–3, that are regulators of DNA CpG demethylation through a methylcytosine dioxygenase function producing hydroxymethylcytosine (hmC). Additionally, TET2 has recently been shown to promote histone O-GlcNAcylation [Chen et al., 2012] and thus can be considered an epigenetic modifier of both chromatin and DNA. We characterized multiple molecular mechanisms that appear to act on *TET2*. Novel somatic missense alterations and LOH were observed in 3/30 (10%) and 19/30 (63%), respectively, of mCRPC cases, and a novel frameshift leading to a truncation of the protein that is likely to be of somatic origin was detected in DU-145, a cell line derived from an mPC brain tumor. Combinations of 14 germline SNPs encoding missense amino acids (Fig. 3; Supp. Table S7, Table S8), may confer susceptibility to PC in accordance with GWAS-identified PC risk linked to the *TET2* locus [Eeles et al., 2009]. Germline SNPs may act in a combinatorial fashion with somatic sequence and copy number alterations to confer risk, or rare germline variants may be highly penetrant, as indicated by the novel germline SNP, D551A, in a putative PC family.

TET2 mutations have been previously observed in 15–20% of myeloid cancers [Delhommeau et al., 2009; Nibourel et al., 2010]. Review of the COSMIC database (March, 2013) shows a lower frequency of mutation in cancers from other tissues producing solid tumors, such as endometrial (2.8%) and urinary (1.8%); and colon, lung, nervous system, and skin (all ~1%). Our findings of *TET2* alterations in tumors from the metastatic stage of mPC add complexity to existing data on epithelial-derived cancers, where predominantly primary tumors have been investigated. Mutations observed across cancers, predominated by analyses of hematopoietic and lymphoid cancers, include a substantial fraction of nonsense substitutions (339/1041, 32.6%) and indels causing frameshifts (390/1041, 37.5%) expected to substantially alter protein structure in 70.1% of cases, suggesting *TET2* loss of function drives these cancers. The mutations identified in this study support a similar role for *TET2* loss of function as a driver in mCRPC as evidenced by frequent somatic LOH and sequence changes, including a confirmed somatic, nonconservative missense alteration (N1890S) in the oxoglutarate-iron-binding catalytic domain and the novel frameshift-truncation in the DU-145 cell line. Significant association of *TET2* locus germline variants to disease risk by GWAS independently confirms a role for *TET2* in prostate cancer. We show complex haplotypes of missense polymorphisms may interact with a previously identified regulatory SNP [Eeles et al., 2009] to influence disease. We are actively investigating mechanisms whereby individual variants and, more generally, *TET2* loss of function contributes to mCRPC. We hypothesize *TET2* alterations may provide survival benefits to circulating cancer cells, alter epithelial cell identity to avoid immune detection, or facilitate androgen receptor signalling, a known driver of mCRPC.

A review of signalling pathways potentially disrupted by alteration of cancer-associated genes (*BRCA1*, *TMPRSS2-ERG*, *PBRM1*, and *TET2*) in the index patient suggests alterations targeted the SWI/SNF/RSC chromatin remodelling pathway, similar to observations in other urogenital cancers [Jones et al., 2010; Gui et al., 2011; Varela et al., 2011]. SWI/SNF is essential for androgen receptor modulation of gene expression in prostate tissue [Marshall et al., 2003; Link et al., 2005] and has a role in ERG-driven transcriptional control [Lin et al., 2009; Cai et al., 2010; Haffner et al., 2010]. PBRM1 is a selective component of a subset of SWI/SNF complexes containing BRG1/SMARCA4. BRCA1, which has a role in homologous recombination DNA repair, directly interacts with BRG1 during transcriptional modulation by androgen [Bochar et al., 2000; Ye et al., 2001; Sun et al., 2007]. TET2 is a CpG demethylase [Tahiliani et al., 2009; Figueroa et al., 2010; Ito et al., 2010] and cooperation between SWI/SNF-mediated chromatin modification and DNA CpG methylation has been recently observed [Cheng and Blumenthal, 2010; Tang et

al., 2010; Chen et al., 2012]. Additional analyses of pathways including proteins encoded by genes altered at each stage of disease (Table 1) are shown in Supp. Figure S2 and in an interactive online supplement (https://variants.ingenuity.com/Nickerson_etal).

This study identified at least two cancer drivers that are potential therapeutic targets. *BRCA1* mutations have been associated with an increased risk for breast, ovarian, and prostate cancer [Tonin et al., 1996; Struwing et al., 1997; King et al., 2003] and a recent study showed that *BRCA1* mutation carriers had a higher risk of recurrence and death due to PC than non-*BRCA* carriers [Gallagher et al., 2010]. The *BRCA1* 185_186delAG allele in the index patient has implications for a man who might present to his physician today. Identification of a *BRCA1* mutation prior to a diagnosis of PC could have enabled effective preventive measures to be taken [Gallagher et al., 2010]. Identification of a *BRCA1* mutation after diagnosis might indicate salutary outcomes through the prescription of PARP inhibitors [Fong et al., 2009], one of which is now being tested for safety and efficacy in the TOPARP trial (NCT01682772).

Secondly, *TET2* status was recently shown to be prognostic in chronic myelomonocytic leukemia [Grossmann et al., 2011] and we theorize that hypomethylating agents such as 5-azacytidine, currently approved by the FDA for treatment of myelodysplastic syndromes, may have a therapeutic effect in mCRPC with *TET2* alterations. This has been the logic rationalizing a recently completed phase II US oncology trial (NCT00384839) and is the basis for a clinical trial now recruiting (NCT00503984). We therefore suggest that additional therapy may have been considered had this patient received molecular characterization of his disease as shown here.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank Drs. William B. Isaacs (Johns Hopkins University), Mary Carrington and Karobi Moitra (National Cancer Institute), and Richard O. Snyder (University of Florida) for comments and assistance; Rachel Karchin and Hannah Carter (Johns Hopkins University) for access to CHASM; Peter Pinto and Cathy Vocke (National Cancer Institute), Mario Eisenberger, Michael Carducci, and Vicki Sinibaldi for oncology clinical support; and Claudia Stewart, Michael Smith, Michael Malasky, Dorothy Jay, Xiande Yi, Praveen Kondreddi, Roger Paul, Sokhom Pin, Heather Martis, Doug Keen, Jesse Himmelstein, Marc Rohrer, Jiro Wada, and Allen Kane for technical assistance and reagents. The content of this publication does not necessarily reflect the views or policies of the Department of Health and Human Services, nor does mention of trade names, commercial products, or organizations imply endorsement by the U.S. Government.

Funding

This work was supported by the Intramural Research Program of the NIH, the National Cancer Institute Center for Cancer Research, and National Institutes of Health funding (CA92234 to G.S.B., HHSN26120080001E to W.T.); John and Kathie Dyson; The American Cancer Society; and the Association for the Cure of Cancer of the Prostate. The content of this publication does not necessarily reflect the views or policies of the Department of Health and Human Services, nor does mention of trade names, commercial products, or organizations imply endorsement by the U.S. Government. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

References

1000 Genomes Project Consortium. A map of human genome variation from population-scale sequencing. *Nature*. 2010; 467:1061–1073. [PubMed: 20981092]

- Adzhubei IA, Schmidt S, Peshkin L, Ramensky VE, Gerasimova A, Bork P, Kondrashov AS, Sunyaev SR. A method and server for predicting damaging missense mutations. *Nat Methods*. 2010; 7:248–249. [PubMed: 20354512]
- Albert TJ, Molla MN, Muzny DM, Nazareth L, Wheeler D, Song X, Richmond TA, Middle CM, Rodesch MJ, Packard CJ, Weinstock GM, Gibbs RA. Direct selection of human genomic loci by microarray hybridization. *Nat Methods*. 2007; 4:903–905. [PubMed: 17934467]
- Barbieri CE, Baca SC, Lawrence MS, Demichelis F, Blattner M, Theurillat JP, White TA, Stojanov P, Van Allen E, Stransky N, Nickerson E, Chae SS, et al. Exome sequencing identifies recurrent *SPOP*, *FOXA1* and *MED12* mutations in prostate cancer. *Nat Genet*. 2012; 44:685–689. [PubMed: 22610119]
- Berger M, Lawrence M, Demichelis F, Drier Y, Cibulskis K, Sivachenko A, Sboner A, Esgueva R, Pflueger D, Sougnez C, Onofrio R, Carter SL, et al. The genomic complexity of primary human prostate cancer. *Nature*. 2011; 470:214–220. [PubMed: 21307934]
- Bochar D, Wang L, Beniya H, Kinev A, Xue Y, Lane W, Wang W, Kashanchi F, Shiekhhattar R. BRCA1 is associated with a human SWI/SNF-related complex: linking chromatin remodeling to breast cancer. *Cell*. 2000; 102:257–265. [PubMed: 10943845]
- Cai J, Kandagatla P, Singareddy R, Kropinski A, Sheng S, Cher M, Chinni S. Androgens induce functional CXCR4 through ERG factor expression in *TMPRSS2-ERG* fusion-positive prostate cancer cells. *Translational Oncology*. 2010; 3:195–203. [PubMed: 20563261]
- Carter H, Chen S, Isik L, Tyekucheva S, Velculescu VE, Kinzler KW, Vogelstein B, Karchin R. Cancer-specific high-throughput annotation of somatic mutations: computational prediction of driver missense mutations. *Cancer Res*. 2009; 69:6660–6667. [PubMed: 19654296]
- Charlop-Powers Z, Zeng L, Zhang Q, Zhou MM. Structural insights into selective histone H3 recognition by the human Polybromo bromodomain 2. *Cell Res*. 2010; 20:529–538. [PubMed: 20368734]
- Chen Q, Chen Y, Bian C, Yu X. TET2 promotes histone O-GlcNAcylation during gene transcription. *Nature*. 2012
- Cheng X, Blumenthal R. Coordinated chromatin control: structural and functional linkage of DNA and histone methylation. *Biochemistry*. 2010; 49:2999–3008. [PubMed: 20210320]
- Delhommeau F, Dupont S, Della VV, James C, Trannoy S, Masse A, Kosmider O, Le Couedic JP, Robert F, Alberdi A, Lécluse Y, Plo I, et al. Mutation in *TET2* in myeloid cancers. *New Engl J Med*. 2009; 360:2289–2301. [PubMed: 19474426]
- Desper R, Jiang F, Kallioniemi OP, Moch H, Papadimitriou CH, Schaffer AA. Inferring tree models for oncogenesis from comparative genome hybridization data. *J Comput Biol*. 1999; 6:37–51. [PubMed: 10223663]
- Eeles RA, Kote-Jarai Z, Al Olama AA, Giles GG, Guy M, Severi G, Muir K, Hopper JL, Henderson BE, Haiman CA, Schleutker J, Hamdy FC, et al. Identification of seven new prostate cancer susceptibility loci through a genome-wide association study. *Nat Genet*. 2009; 41:1116–1121. [PubMed: 19767753]
- Figueroa ME, Abdel-Wahab O, Lu C, Ward PS, Patel J, Shih A, Li Y, Bhagwat N, Vasanthakumar A, Fernandez HF, Tallman MS, Sun Z, et al. Leukemic *IDH1* and *IDH2* mutations result in a hypermethylation phenotype, disrupt TET2 function, and impair hematopoietic differentiation. *Cancer Cell*. 2010; 18:553–567. [PubMed: 21130701]
- Fong PC, Boss DS, Yap TA, Tutt A, Wu P, Mergui-Roelvink M, Mortimer P, Swaisland H, Lau A, O'Connor MJ, Ashworth A, Carmichael J, et al. Inhibition of poly(ADP-ribose) polymerase in tumors from *BRCA* mutation carriers. *New Engl J Med*. 2009; 361:123–134. [PubMed: 19553641]
- Forbes SA, Bhamra G, Bamford S, Dawson E, Kok C, Clements J, Menzies A, Teague JW, Futreal PA, Stratton MR. The Catalogue of Somatic Mutations in Cancer (COSMIC). *Current Protocols in Human Genetics*. Apr.2008 Unit 10.11.
- Gallagher DJ, Gaudet MM, Pal P, Kirchhoff T, Balistreri L, Vora K, Bhatia J, Stadler Z, Fine SW, Reuter V, Zelefsky M, Morris MJ, et al. Germline *BRCA* mutations denote a clinicopathologic subset of prostate cancer. *Clin Cancer Res*. 2010; 16:2115–2121. [PubMed: 20215531]
- Garcia M, Jemal A, Ward EM, Center MM, Hao Y, Siegel RL, Thun MJ. Global Cancer Facts & Figures 2007. American Cancer Society Website. 2007:1–44.

- Grasso CS, Wu YM, Robinson DR, Cao X, Dhanasekaran SM, Khan AP, Quist MJ, Jing X, Lonigro RJ, Brenner JC, Asangani IA, Ateeq B, et al. The mutational landscape of lethal castration-resistant prostate cancer. *Nature*. 2012; 487:239–243. [PubMed: 22722839]
- Grossmann V, Kohlmann A, Eder C, Haferlach C, Kern W, Cross N, Haferlach T, Schnittger S. Molecular profiling of chronic myelomonocytic leukemia reveals diverse mutations in >80% of patients with *TET2* and *EZH2* being of high prognostic relevance. *Leukemia*. 2011; 25:877–879. [PubMed: 21339759]
- Gui Y, Guo G, Huang Y, Hu X, Tang A, Gao S, Wu R, Chen C, Li X, Zhou L, He M, Li Z, et al. Frequent mutations of chromatin remodeling genes in transitional cell carcinoma of the bladder. *Nat Genet*. 2011; 43:875–878. [PubMed: 21822268]
- Haffner MC, Aryee MJ, Toubaji A, Esopi DM, Albadine R, Gurel B, Isaacs WB, Bova GS, Liu W, Xu J, Meeker AK, Netto G, et al. Androgen-induced TOP2B-mediated double-strand breaks and prostate cancer gene rearrangements. *Nat Genet*. 2010; 42:668–675. [PubMed: 20601956]
- Hanahan D, Weinberg RA. Hallmarks of cancer: the next generation. *Cell*. 2011; 144:646–674. [PubMed: 21376230]
- Hodges E, Xuan Z, Balija V, Kramer M, Molla MN, Smith SW, Middle CM, Rodesch MJ, Albert TJ, Hannon GJ, McCombie WR. Genome-wide in situ exon capture for selective resequencing. *Nat Genet*. 2007; 39:1522–1527. [PubMed: 17982454]
- Hornsby C, Page KM, Tomlinson IP. What can we learn from the population incidence of cancer? Armitage and Doll revisited. *Lancet Oncol*. 2007; 8:1030–1038. [PubMed: 17976613]
- Ito S, D'Alessio AC, Taranova OV, Hong K, Sowers LC, Zhang Y. Role of Tet proteins in 5mC to 5hmC conversion, ES-cell self-renewal and inner cell mass specification. *Nature*. 2010; 466:1129–1133. [PubMed: 20639862]
- Jayapandian M, Chapman A, Tarcea VG, Yu C, Elkiss A, Ianni A, Liu B, Nandi A, Santos C, Andrews P, Athey B, States D, et al. Michigan Molecular Interactions (MiMI): putting the jigsaw puzzle together. *Nucleic Acids Res*. 2007; 35:D566–D571. [PubMed: 17130145]
- Jemal A, Bray F, Center MM, Ferlay J, Ward E, Forman D. Global cancer statistics. *CA Cancer J Clin*. 2011; 61:69–90. [PubMed: 21296855]
- Jones S, Chen WD, Parmigiani G, Diehl F, Beerewinkel N, Antal T, Traulsen A, Nowak MA, Siegel C, Velculescu VE, Kinzler KW, Vogelstein B, et al. Comparative lesion sequencing provides insights into tumor evolution. *Proc Natl Acad Sci U S A*. 2008; 105:4283–4288. [PubMed: 18337506]
- Jones S, Wang T, Shih I, Mao T, Nakayama K, Roden R, Glas R, Slamon D, Diaz L Jr. Vogelstein B, Kinzler KW, Velculescu VE, et al. Frequent mutations of chromatin remodeling gene *ARID1A* in ovarian clear cell carcinoma. *Science*. 2010; 330:228–231. [PubMed: 20826764]
- Kent WJ. BLAT-the BLAST-like alignment tool. *Genome Res*. 2002; 12:656–664. [PubMed: 11932250]
- King MC, Marks JH, Mandell JB. Breast and ovarian cancer risks due to inherited mutations in *BRCA1* and *BRCA2*. *Science*. 2003; 302:643–646. [PubMed: 14576434]
- Krzywinski M, Schein J, Birol I, Connors J, Gascoyne R, Horsman D, Jones SJ, Marra MA. Circos: an information aesthetic for comparative genomics. *Genome Res*. 2009; 19:1639–1645. [PubMed: 19541911]
- Kumar A, White TA, MacKenzie AP, Clegg N, Lee C, Dumpit RF, Coleman I, Ng SB, Salipante SJ, Rieder MJ, Nickerson DA, Corey E, et al. Exome sequencing identifies a spectrum of mutation frequencies in advanced and lethal prostate cancers. *Proc Natl Acad Sci U S A*. 2011; 108:17087–17092. [PubMed: 21949389]
- Kumar P, Henikoff S, Ng PC. Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm. *Nat Protoc*. 2009; 4:1073–1081. [PubMed: 19561590]
- Lin C, Yang L, Tanasa B, Hutt K, Ju B, Ohgi K, Zhang J, Rose D, Fu X, Glass C, Rosenfeld MG. Nuclear receptor-induced chromosomal proximity and DNA breaks underlie specific translocations in cancer. *Cell*. 2009; 139:1069–1083. [PubMed: 19962179]
- Link K, Burd C, Williams E, Marshall T, Rosson G, Henry E, Weissman B, Knudsen K. BAF57 governs androgen receptor action and androgen-dependent proliferation through SWI/SNF. *Molec Cell Biol*. 2005; 25:2200–2215. [PubMed: 15743818]

- Liu W, Laitinen S, Khan S, Vihinen M, Kowalski J, Yu G, Chen L, Ewing CM, Eisenberger MA, Carducci MA, Nelson WG, Yegnasubramanian S, et al. Copy number analysis indicates monoclonal origin of lethal metastatic prostate cancer. *Nat Med.* 2009; 15:560–565.
- Marshall T, Link K, Petre-Draviam C, Knudsen K. Differential requirement of SWI/SNF for androgen receptor activity. *J Biol Chem.* 2003; 278:30605–30613. [PubMed: 12775722]
- Mickey DD, Stone KR, Wunderli H, Mickey GH, Paulson DF. Characterization of a human prostate adenocarcinoma cell line (DU 145) as a monolayer culture and as a solid tumor in athymic mice. *Prog Clin Biol Res.* 1980; 37:67–84. [PubMed: 7384095]
- Nibourel O, Kosmider O, Cheok M, Boissel N, Renneville A, Philippe N, Dombret H, Dreyfus F, Quesnel B, Geffroy S, Quentin S, Roche-Lestienne C, et al. Incidence and prognostic value of *TET2* alterations in de novo acute myeloid leukemia achieving complete remission. *Blood.* 2010; 116:1132–1135. [PubMed: 20489055]
- Nickerson M, Jaeger E, Shi Y, Durocher J, Mahurkar S, Zaridze D, Matveev V, Janout V, Kollarova H, Bencko V, Navratilova M, Szeszenia-Dabrowska N, et al. Improved identification of *von Hippel-Lindau* gene alterations in clear cell renal tumors. *Clin Cancer Res.* 2008; 14:4726–4734. [PubMed: 18676741]
- Nicol JW, Helt GA, Blanchard SG Jr. Raja A, Loraine AE. The Integrated Genome Browser: free software for distribution and exploration of genome-scale datasets. *Bioinformatics.* 2009; 25:2730–2731. [PubMed: 19654113]
- Okou DT, Steinberg KM, Middle C, Cutler DJ, Albert TJ, Zwick ME. Microarray-based genomic selection for high-throughput resequencing. *Nat Methods.* 2007; 4:907–909. [PubMed: 17934469]
- Pruitt KD, Harrow J, Harte RA, Wallin C, Diekhans M, Maglott DR, Searle S, Farrell CM, Loveland JE, Ruef BJ, Hart E, Suner MM, et al. The consensus coding sequence (CCDS) project: Identifying a common protein-coding gene set for the human and mouse genomes. *Genome Res.* 2009; 19:1316–1323. [PubMed: 19498102]
- Rhead B, Karolchik D, Kuhn RM, Hinrichs AS, Zweig AS, Fujita PA, Diekhans M, Smith KE, Rosenbloom KR, Raney BJ, Pohl A, Pheasant M, et al. The UCSC Genome Browser database: update 2010. *Nucleic Acids Res.* 2010; 38:D613–D619. [PubMed: 19906737]
- Robbins C, Tembe W, Baker A, Sinari S, Moses T, Beckstrom-Sternberg S, Beckstrom-Sternberg J, Barrett M, Long J, Chinnaiyan A, Lowey J, Suh E, et al. Copy number and targeted mutational analysis reveals novel somatic events in metastatic prostate tumors. *Genome Res.* 2011; 21:47–55. [PubMed: 21147910]
- Roberts C, Orkin S. The SWI/SNF complex-chromatin and cancer. *Nat Rev Cancer.* 2004; 4:133–142. [PubMed: 14964309]
- Robinson JT, Thorvaldsdottir H, Winckler W, Guttman M, Lander ES, Getz G, Mesirov JP. Integrative genomics viewer. *Nat Biotechnol.* 2011; 29:24–26. [PubMed: 21221095]
- Rozen S, Skaletsky H. Primer3 on the WWW for general users and for biologist programmers. *Methods Mol Biol.* 2000; 132:365–386. [PubMed: 10547847]
- Schaub FX, Looser R, Li S, Hao-Shen H, Lehmann T, Tichelli A, Skoda RC. Clonal analysis of *TET2* and *JAK2* mutations suggests that *TET2* can be a late event in the progression of myeloproliferative neoplasms. *Blood.* 2010; 115:2003–2007. [PubMed: 20061559]
- Schwarz JM, Rodelsperger C, Schuelke M, Seelow D. MutationTaster evaluates disease-causing potential of sequence alterations. *Nat Methods.* 2010; 7:575–576. [PubMed: 20676075]
- Stone KR, Mickey DD, Wunderli H, Mickey GH, Paulson DF. Isolation of a human prostate carcinoma cell line (DU 145). *Int J Cancer.* 1978; 21:274–281. [PubMed: 631930]
- Stratton MR. Exploring the genomes of cancer cells: progress and promise. *Science.* 2011; 331:1553–1558. [PubMed: 21436442]
- Struwing JP, Hartge P, Wacholder S, Baker SM, Berlin M, McAdams M, Timmerman MM, Brody LC, Tucker MA. The risk of cancer associated with specific mutations of *BRCA1* and *BRCA2* among Ashkenazi Jews. *N Engl J Med.* 1997; 336:1401–1408. [PubMed: 9145676]
- Sun J, Blair A, Aiyar S, Li R. Cofactor of BRCA1 modulates androgen-dependent transcription and alternative splicing. *J Steroid Biochem Mol Biol.* 2007; 107:131–139. [PubMed: 17659869]

- Tahiliani M, Koh KP, Shen Y, Pastor WA, Bandukwala H, Brudno Y, Agarwal S, Iyer LM, Liu DR, Aravind L, Rao A. Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1. *Science*. 2009; 324:930–935. [PubMed: 19372391]
- Tang L, Nogales E, Ciferri C. Structure and function of SWI/SNF chromatin remodeling complexes and mechanistic implications for transcription. *Prog Biophys Mol Biol*. 2010; 102:122–128. [PubMed: 20493208]
- Tarcea VG, Weymouth T, Ade A, Bookvich A, Gao J, Mahavisno V, Wright Z, Chapman A, Jayapandian M, Ozgur A, Tian Y, Cavalcoli J, et al. Michigan molecular interactions r2: from interacting proteins to pathways. *Nucleic Acids Res*. 2009; 37:D642–D646. [PubMed: 18978014]
- Taylor B, Schultz N, Hieronymus H, Gopalan A, Xiao Y, Carver B, Arora V, Kaushik P, Cerami E, Reva B, Antipin Y, Mitsiades N, et al. Integrative genomic profiling of human prostate cancer. *Cancer Cell*. 2010; 18:11–22. [PubMed: 20579941]
- Thompson M. Polybromo-1: the chromatin targeting subunit of the PBAF complex. *Biochimie*. 2009; 91:309–319. [PubMed: 19084573]
- Tomlins SA, Rhodes DR, Perner S, Dhanasekaran SM, Mehra R, Sun XW, Varambally S, Cao X, Tchinda J, Kuefer R, Lee C, Montie JE, et al. Recurrent fusion of *TMPRSS2* and *ETS* transcription factor genes in prostate cancer. *Science*. 2005; 310:644–648. [PubMed: 16254181]
- Tonin P, Weber B, Offit K, Couch F, Rebbeck TR, Neuhausen S, Godwin AK, Daly M, Wagner-Costalos J, Berman D, Grana G, Fox E, et al. Frequency of recurrent *BRCA1* and *BRCA2* mutations in Ashkenazi Jewish breast cancer families. *Nat Med*. 1996; 2:1179–1183. [PubMed: 8898735]
- Varela I, Tarpey P, Raine K, Huang D, Ong CK, Stephens P, Davies H, Jones D, Lin ML, Teague J, Bignell G, Butler A, et al. Exome sequencing identifies frequent mutation of the SWI/SNF complex gene *PBRM1* in renal carcinoma. *Nature*. 2011; 469:539–542. [PubMed: 21248752]
- Vogelstein B, Kinzler KW. Cancer genes and the pathways they control. *Nat Med*. 2004; 10:789–799. [PubMed: 15286780]
- Xia W, Nagase S, Montia AG, Kalachikov SM, Keniry M, Su T, Memeo L, Hibshoosh H, Parsons R. BAF180 is a critical regulator of p21 induction and a tumor suppressor mutated in breast cancer. *Cancer Res*. 2008; 68:1667–1674. [PubMed: 18339845]
- Ye Q, Hu Y, Zhong H, Nye A, Belmont A, Li R. BRCA1-induced large-scale chromatin unfolding and allele-specific effects of cancer-predisposing mutations. *J Cell Biol*. 2001; 155:911–921. [PubMed: 11739404]

Figure 1.

BRCA1 and *PBRM1* alterations. **A:** Sanger re-sequencing of a *BRCA1* mutation in matched normal DNA (top) which is homozygous in metastases (bottom). **B:** Copy number indicates *BRCA1* (arrow) LOH. The horizontal center line indicates signal for two copies. **C:** Sanger re-sequencing of *PBRM1* is wt in noncancerous kidney (top) and confirms a somatic A/C variant (p.M672R) in all metastases and the primary tumor (representative tumor, bottom). **D:** High conservation of p.M672 and neighboring residues. **E:** *PBRM1* protein domains and somatic missense alterations of BRD5 in tumors. Diamonds, PC (this study and Berger et al. [2011]); and dots, kidney cancer; cross, squamous cell carcinoma [Varela et al., 2011]. Triangles, acetylated lysine binding pocket catalytic residues; α3, α4, alpha helices conserved in BRDs; Pro, an essential proline. Adapted from [Charlop-Powers et al., 2010; Thompson, 2009].

Figure 2.

Alteration of *TMPRSS2* and fusion with *ERGA*: *TMPRSS2* re-sequencing in noncancerous kidney (top) confirms a somatic insertion (arrow) in reverse sequencing reads from metastases (bottom). **B**: Array CGH (top, representative metastasis) indicates a deletion between *ERG* (left dotted line) and *TMPRSS2* (right dotted line). Binned Helicos NGS reads (middle) identify LOH borders in *TMPRSS2* IVS-11 and *ERG* IVS-1 (± 25 kb, bottom). **C**: RT-PCR of metastatic tumor RNA. Top: PCR primer locations (bent arrows), exons (boxes) not to scale, // indicates exons not shown. Left gel, lane 1: amplified fusion transcript (arrow) with a junction between *TMPRSS2* exon 1 and *ERG* exon 4. Right gel:

RT-PCR using TMPR_E10_s1 and TMPR_E13_as1 (lane 1) or TMPR_E10_s1 and TMPR_E14_as1 (lane 2) showed the *TMPRSS2* exon 10 insertion. TMPR_E10_s1 and ERG_E5_as1 (lane 3) showed no product. Data indicates the insertion and fusion transcript are on separate alleles (summarized, panel **D**). Primer sequences are available upon request; M, 1-kb DNA ladder (Invitrogen).

Figure 3.

Alteration of *TET2* in mCRPC. **A:** Sanger chromatograms from matched noncancerous tissue (NL, left), and metastases (met, right). **1.** Subject A2, somatic G>T (p.R43I) in a liver metastasis; **2.** Subject A7, somatic A>G (p.N1890S) in a post-subdural metastasis; **3.** Index case A17, somatic C>G (p.P562A) in a supraclavicular lymph node metastasis; **4.** Subject A23, germline A>C (p.D551A); **5.** Cell line DU-145, insertion (arrows) in forward (top) and reverse (bottom) chromatograms. **B:** Protein diagram with novel PC variants (above) and germline missense SNPs (below). An oxoglutarate-iron-binding domain (2OGFE domain) and proline-, glutamine-, and cysteine (Cys)-rich regions are indicated. Post-translational

modifications: acetylated residue (p.K53, vertical purple line) and phosphorylation sites (p.T97, p.S99, p.S1107, and p.Y1345; yellow lines). Iron-binding catalytic residues (p.H1382, p.D1384, and p.H1881, brown lines), a 2-oxoglutarate-binding site (p.R1896, pink line), and conserved Box 1 and Box 2 [Delhommeau et al., 2009]. **C:** Somatic variants alter conserved residues. Asterisks, complete conservation; two dots, highly conserved; one dot, moderately conserved. **D:** Copy number in mCRPC show LOH (blue) and a gain (red) across *TET2* (shaded).

Figure 4.

Sequential acquisition of somatic variants. **A:** An oncogenetic tree of late metastatic variants indicates metastases with no variants (blue circles, panel B) and defines metastatic lineages. **B:** One possible model of the spread of disease (see text). Blue, metastases with no late metastatic variants; green, subdural metastasis; red, upper thorax region metastases; yellow, prostate primary tumor. **C:** RSI of somatic mutations in chromatograms from the primary tumor. A germline *BRCA1* mutation shows a RSI = 0.5, indicating no LOH.

Table 1Genes Altered During Progression of a mCRPC^a

Germline	Primary Adenocarcinoma	Common Metastatic	Late Metastatic
Cancer-associated genes			
<i>BRCA1</i>	<i>TMPRSS2-ERG</i> ^b	<i>BRCA1 LOH</i>	
	<i>PBRM1</i>	<i>TET2</i>	
		<i>TMPRSS2</i>	
Other genes	<i>ADAM19</i>	<i>ACAP2</i>	<i>ABCA12</i>
	<i>ALDH8A1</i>	<i>ARAP3</i>	<i>C6</i>
	<i>BUB1</i>	<i>C7ORF47</i>	<i>COL19A1</i>
	<i>CERK</i>	<i>CDH10</i>	<i>FAM70B</i>
	<i>CKAP5</i>	<i>CTSD</i>	<i>LAMA3</i>
	<i>COL22A1</i>	<i>CYP4F11</i>	<i>POLR2B</i>
	<i>DPY19L2</i>	<i>DHX38</i>	<i>SLC12A5</i>
	<i>EIF1AX</i>	<i>DPP10</i>	<i>TCF12 LOH</i>
	<i>MTMR12</i>	<i>FOXP1</i>	<i>TMEM143</i>
	<i>MYADM</i>	<i>FUT3</i>	<i>ZNF334</i>
	<i>MYH6</i>	<i>G3BP2</i>	
	<i>NELL1</i>	<i>GTF3C3</i>	
	<i>PIK3AP1</i>	<i>IL27RA</i>	
	<i>PLS1</i>	<i>ITGB6</i>	
	<i>PLS3</i>	<i>KCNJ3</i>	
	<i>PTAFR</i>	<i>KIF9</i>	
	<i>SDCCAG1</i>	<i>LCT</i>	
	<i>SKIL</i>	<i>LY75</i>	
	<i>TDP1</i>	<i>MAPKAPK2</i>	
	<i>TNFRSF18</i>	<i>PCDP1</i>	
	<i>TTC26</i>	<i>PLD1</i>	
		<i>PRIC285</i>	
		<i>SETD1A</i>	
		<i>SIN3A</i>	
		<i>SYCP1</i>	
		<i>TCF12</i>	
		<i>UGGT1</i>	
		<i>UIMC1</i>	
		<i>XRN2</i>	

^aSequence alteration unless indicated.^bHigh probability in the primary tumor.

Table 2

Novel *TET2* Variants in mCRPC

Patient/Cell Line	Race	Age Diagnosis/ TTD ^a	Family History of PC	Family History of Any Cancer	Tumor/Tissue Examined	Mutation Type	Nucleotide	Amino Acid Change ^b	Reference>Variant Amino Acid
A2	Caucasian	51/3.8 yrs	grandfather	unknown	Liver Met C1	Somatic	G>T	R43I	Basic, Hydrophilic<>Hydrophobic
A7	Caucasian	59/4.1 yrs	unknown	paternal aunt had cancer, unknown type	R Post Subdural Met	Somatic	A>G	N1890S	Hydrophilic<>Hydrophilic
A17	Caucasian	58/3.5 yrs	unknown	mother and sister-breast cancer	Multiple	Somatic	C>G	P562A	Hydrophobic<>Hydrophobic
A23	African American	56/6 yrs	father and brother	mother-eukemia	Blood	Germline	A>C	D551A	Acidic, Hydrophilic<>Hydrophobic
DU-145	Caucasian	69/0.5 yrs	x	x	Brain	ND	insA	T229fs24*	Truncation

^aTTD, time to death.

^bReference protein, NP_001120680.

yrs, years; x = not applicable; ND, no data.