The Genomics of Prostate Cancer: A Historic Perspective

Mark A. Rubin1,2,3,4 and Francesca Demichelis1,5

1Englander Institute for Precision Medicine, Weill Cornell Medical College-New York Presbyterian Hospital, New York, New York 10065
2Sandra and Edward Meyer Cancer Center at Weill Cornell Medical College, New York, New York 10021
3Department of Pathology and Laboratory Medicine, Weill Cornell Medical College, New York, New York 10021
4Department of BioMedical Research, University of Bern, 3012 Bern, Switzerland
5Centre for Integrative Biology, University of Trento, 38123 Trento, Italy

Correspondence: f.demichelis@unitn.it

The genomics of prostate cancer (PCA) has been difficult to study compared with some other cancer types for a multitude of reasons, despite significant efforts since the early 1980s. Overcoming some of these obstacles has paved the way for greater insight into the genomics of PCA. The advent of high-throughput technologies coming from the initial use of microsatellite and oligonucleotide probes gave rise to techniques like comparative genomic hybridization (CGH). With the introduction of massively parallel genomic sequencing, referred to as next-generation sequencing (NGS), a deeper understanding of cancer genomics in general has occurred. Along with these technologic advances, there has been the development of computational biology and statistical approaches to address novel large data sets characterized by single base resolution. This review will provide a historic perspective of PCA genomics with an emphasis on the cardinal mutations and alterations observed to be consistently seen in PCA for both hormone-naïve localized PCA and castration-resistant prostate cancer (CRPC). There will be a focus on alterations that have the greatest potential to play a role in disease progression and therapy management.

The genomics of prostate cancer (PCA) has been difficult to study compared with some other cancer types for a multitude of reasons, despite significant efforts since the early 1980s (Fig. 1). First, the anatomic location of the prostate gland initially made it challenging to obtain samples for research. This was largely overcome with the widespread acceptance of the nerve-sparing radical prostatectomy, introduced by Dr. Walsh, as a safe procedure (Mullins et al. 2012). Perhaps an unanticipated consequence of the increase in radical prostatectomy surgery for clinically localized cancers was new availability of tumor tissue for research study.

Second, unlike melanoma, lung, and breast cancers, PCA does not lend itself to simple growth in culture. From the earliest attempts to grow prostate cancer in vitro to develop short-term cultures or to develop karyotypes, it was clear that cells would not readily grow,
and most often undergo senescence or are overtaken by benign fibromuscular stromal cells. The effects have severely hampered PCA research, limiting model systems for understanding the basic biology and genomics of the disease. Recent work in the development of PCA organoids has reignited interest in developing novel model systems for biologic and genomic studies that more closely reflect the PCA observed in patients (Gao et al. 2014).

Third, it has been traditionally difficult to obtain samples from advanced disease. However, now improved care for men with advanced castration-resistant prostate cancer (CRPC) has led to longer survival. As a consequence, there is more acceptance for metastatic bone and soft tissue biopsies, thereby making CRPC samples available for genomic studies (Robinson et al. 2015). These samples are now collected in the context of defined standard of care therapy making for more uniform comparison between samples and cohorts.

Overcoming some of these obstacles has paved the way for greater insight into the genomics of PCA. The advent of high-throughput technologies coming from the initial use of microsatellite and oligonucleotide probes gave rise to arrayed technologies. For example, comparative genomic hybridization (CGH) led to fluorescence in situ hybridization (FISH) and array of oligonucleotide probes spotted on chips—called array CGH (aCGH). Gene expression studies, or transcriptomics, first started as spotted arrays of cDNA clones that could give great insight into transcript levels producing the first heatmap data for cancer. These technically challenging spotted cDNA arrays were soon overtaken by easier to use oligonucleotide arrays. With the introduction of massively parallel genomic sequencing referred to as next-generation sequencing (NGS), a deeper understanding of cancer genomics in general has occurred. As NGS technology improved and cost decreased, base-pair resolution genomics and transcriptomics has created more detailed understanding of cancer by facilitating the in-depth study of larger and larger cohorts. Along with these technologic advances, there has been the development of computational biology and statistical approaches to address novel raw data peculiarities and large data sets.

As is often the case in science, there tends to be an emphasis on recent events. In particular, the recent advances providing the first comprehensive picture of PCA tend to overshadow the genomic knowledge that dates from the 1980s. This review will provide a historic perspective of...
PCA genomics with an emphasis on the cardinal mutations and alterations observed to be consistently seen in PCA (Fig. 1), for both hormone-naïve localized PCA and CRPC. There will be a focus on alterations that have the greatest potential to play a role in disease progression and therapy management.

THE FIRST GENOMIC LANDSCAPE OF PROSTATE CANCER (1980s–1990s)

Cancer is a genetic disease. There are two major classes of significant mutations: inactivating mutations (in tumor suppressors) and activating mutations (in oncogenes). Inactivation often comes from structural rearrangements involving loss of genomic DNA resulting in deletions (large or focal) or rearrangements. In both cases, a gene or groups of genes are disrupted. These events can be either mono- (heterozygous) or bi-allelic (homozygous). Activation can occur through amplification, point mutation, or structural rearrangements leading to gene fusions. As noted by Felix Mitelman (Mitelman et al. 2007), the cytogenetics of leukemia and lymphoma are far simpler than any of the solid tumors and the number of patient samples studies to date far greater; this has resulted in significant advances in the understanding of hematopoietic disease, whereas solid tumors are often characterized by more complex karyotypes and therefore it is difficult to untangle to point to the critical mutation. This complexity combined with the difficulty of growing PCA cells in culture made the elucidation of PCA cytogenetics difficult. Yet, despite these challenges, several cardinal chromosomal alterations were gleaned originally in the 1980s from cytogenetic studies. Details and greater level of resolution for the cardinal recurrent alterations in PCA including 8p, 8q (C-MYC locus), 10p (PTEN locus), 17q (TP53 locus), and androgen receptor (AR) first came with the advent of microsatellite polymorphic fragment length markers, CHG, aCHG, and FISH.

Cytogenetics: Critical Revelations from a Small Sampling of Prostate Cancer Samples

In 1984, a letter in the New England Journal of Medicine reported on chromosome 10 deletion at 10q24 in two men with moderate grade PCA and two with poorly differentiated PCA (Atkin and Baker 1985a). Although an earlier study in the same year on the PCA cell line LNCaP had reported 10q24 loss (Gibas et al. 1984), the work from Atkin and Baker was the first published cytogenetic study in human PCA samples. They wrote in this three-paragraph letter to the editor, “We propose … that a deletion of a chromosome 10 represents a specific change in PCA; if this is true, it would follow that genes of importance in the development of the tumor are located on this chromosome.” Later reports confirmed this 10q24 locus as a site for recurrent deletion in PCA (Atkin and Baker 1985b; Lundgren et al. 1992; Arps et al. 1993). In 1992, Mitelman’s group was able to perform cytogenetics on 57 primary PCA samples (Lundgren et al. 1992). The study also gave insight into the challenges faced with short-term culture; out of 82 cases attempted, 57 grew in media and of these 57 cases, 24 showed normal karyotypes. Therefore, from the original 82 cases, only 40% (33/82) grew tumor cells. Given that these were probably bulky tumors compared with tumors harvested in the post-PSA era, current success rates for growing tumor samples would likely be significantly worse. Despite these challenges, they were able to show that 10q loss (10q24) occurred in 5/33 (15%) PCA cases studied. Other studies reported recurrent alterations in 8p and genomic loci harboring TP53 and RB (König et al. 1989; Rubin et al. 1991; Lundgren et al. 1992). The cytogenetic work led to cardinal observations about PCA genomics but represented a highly challenging approach and one that has been abandoned by most research groups in subsequent years because of the low success rate.

Restriction Fragment Length Polymorphisms: Pinpointing Genomic Alterations with Microsatellite Markers

The next wave of genomic studies took advantage of polymorphic microsatellite markers to more precisely define alterations in genomic loci. Using the restriction fragment length polymorphisms (RFLP) assay, DNA could be enzymatically sheared (e.g., using restriction en-
zymes such as RsaI or BamH1) and by using radiolabeled intragenic DNA probes, polymorphic bands could be used to determine the presence or absence of these queried fragments (Feinberg and Vogelstein 1984). Experiments could provide information regarding heterozygous or homozygous loss of genomic loci at the probed site when compared to normal DNA from the same individual (referred to as LOH). By using multiple probes and enzymes, recurrent areas of loss could be narrowed down. One of the major incentives was to narrow down genomic loss to a locus that might harbor a tumor suppressor gene.

TP53 (Cellular Tumor Antigen p53)

Studies based on the RFLP approach helped discover and define the role of p53 as a tumor suppressor gene altered in multiple cancer types, including lung, colon, breast, bladder, and brain cancers (Baker et al. 1989; Nigro et al. 1989; Takahashi et al. 1989). Using this approach, Carter et al. (1990) defined recurrent LOH on a number of loci in PCA, including the 17p locus harboring p53. In a follow-up study, Isaacs et al. (1994) showed for the first time the presence of p53 mutations in cell lines (e.g., PC3 and DU145) and two primary human PCA samples. They also showed the functional role that wild-type (WT) p53 plays as a tumor suppressor, helping to define p53 as a tumor suppressor in PCA. Subsequent studies showed that in larger numbers of patient samples, p53 mutations accumulated with disease progression (Bookstein et al. 1990a; Phillips et al. 1994; Brooks et al. 1995). In the CRPC500 cohort, RB1 is reported lost in 21% of cases (Robinson et al. 2015).

8p Loss

The first report of 8p loss in PCA was by Konig et al. (1989) using the LNCaP cell line. Numerous groups confirmed this observation in primary and advanced PCA (Bergerheim et al. 1991; Bova et al. 1993; Cher et al. 1994; MacGrogan et al. 1994; Macoska et al. 1994; Massenkeil et al. 1994; Matsuyama et al. 1994; Sakr et al. 1994; Trapman et al. 1994; Brooks et al. 1995; Emmert-Buck et al. 1995; Suzuki et al. 1995; Visakorpi et al. 1995b). Although candidate tumor suppressor genes had been proposed in this region (Bova et al. 1996), it was not until 1997 that he and colleagues identified a prostate-specific gene, NKX3.1, which is homologous to the Drosophila NK homeobox gene family (He et al. 1997). NKX3.1 is expressed at high levels in normal prostate and is androgen sensitive as they determined by androgen stimulation of LNCaP cell lines. They mapped NKX3.1 to chromosome band 8p21, a region that was previously noted to undergo loss (Fig. 1). He and colleagues proposed a potential tumor suppressor role for NKX3.1.

Androgen Receptor

Mutations have been long known to exist in the androgen receptor (AR). AR mutations occur and result in a germline disorder called androgen insensitivity syndrome (AIS), an X chromosome-linked inherited disorder (reviewed in Hughes et al. 2012; Shukla et al. 2016). Mutations in the ligand-binding domain of the AR receptor were first observed in an androgen responsive PCA cell line, LNCaP (Veldscholte et al. 1990). Newmark et al. (1992) reported the first AR mutations associated with primary
PCA. Frequent AR mutations were observed in CRPC (50%) demonstrating for the first time that AR resistance via mutation occurs with AR targeted therapy (Taplin et al. 1995). Taplin and colleagues stated, “Our results suggest that mutant androgen-receptor genes in androgen-independent PCA could be useful targets of new drugs for the treatment of PCA.” Another mechanism for AR resistance can be explained by AR gene amplification. Array CGH and fluorescence in situ hybridization (FISH) technology help define fourfold to more than 20-fold AR amplifications in hormone treated PCA patients but not in untreated hormone-naïve PCA (Fig. 2).

**Figure 2.** Androgen receptor gene copy number and expression in a patient with a recurrent prostate carcinoma and androgen deprivation therapy (ADT). Results from comparative genomic hybridization (CGH) (A–C), fluorescence in situ hybridization (FISH) (D–F), and AR immunohistochemistry (G, H) are shown. (A) Ideogram of the X chromosome. (B) Green to red fluorescence intensity ratio profile (tumor to normal DNA sequence copy number) along the length of chromosome X obtained by CGH analysis of the recurrent hormone-refractory tumor. Increased copy number is seen at the Xq12 region. (C) A digital image of the CGH experiment showing increased green fluorescence intensity at Xq12 suggesting amplification of DNA sequences from this region. (D) FISH to normal metaphase chromosomes with probes for the AR gene (in green) with DXZ1 (in red) illustrating the location of the AR gene at Xq12. (E) Interphase FISH analysis of the primary tumor with probes for AR and DXZ1. Two copies of both AR and DXZ1 are seen indicating duplication of the entire X chromosome, but no AR gene amplification. (F) Interphase FISH analysis of the recurrent tumor from the same patient. High-level amplification of the AR gene is seen as a tight cluster of approximately 12 green signals as compared with only two DXZ1 signals. Immunohistochemical demonstration of high-level AR gene expression in sections from both the (G) primary and (H) recurrent tumor. The staining intensity was slightly higher in the recurrent cancer tissue. (Panel and legend reprinted from Visakorpi et al. 1995a, with permission, from *Nature Genetics* 1995.)
Visakorpi et al. 1995a). With the development of tissue microarray technology, larger numbers of clinical samples could be detected on a single slide. Using tissue microarrays (TMAs), Bubendorf et al. (1999) queried the AR status of 371 PCA samples by FISH. AR was determined to be amplified in 23% of the 47 CRPC cases in contrast to 2 of 205 (1%) of the primary hormone-naïve PCA cases. In more recent studies using NGS, the AR aberration frequencies show the same patterns. In the TCGA study of 333 hormone-naïve PCA, no AR mutations were detected (Cancer Genome Atlas Research Network 2015). In studies in which tumors were evaluated after androgen deprivation therapy (ADT), AR mutations and amplification frequencies were in the range of the initial reports (Grasso et al. 2012b; Beltran et al. 2013, 2016; Robinson et al. 2015; Kumar et al. 2016). Other mechanism of AR resistance have been proposed including v7 AR splice variants (Sun et al. 2010) and lineage plasticity to AR indifferent CRPC (Beltran et al. 2011, 2016).

10q23 (PTEN)

The distal region of 10q is lost in a number of common cancers such as glioblastoma and breast cancer. Early studies using RFLP assays located the loss at 10q24, but a series of papers in the 1990s targeted 10q23.1 as a potential site for a tumor suppressor gene. In 1990, Carter and colleagues reported 10q loss in \(\sim 30\%\) of localized PCA (Carter et al. 1990). In 1995, Gray and colleagues suggested the critical area for a potential tumor suppressor was in 10q23-24, which was lost in 62% of the 37 PCA cases they examined (Gray et al. 1995). In 1996, Ittmann was the first to propose that 10q23.1 shows increased loss in advanced PCA. He also suggested that prior studies using aCGH approaches may have missed the 10q23.1 region, as deletions in some cases were small.

Mapping multiple cancers including brain, breast and prostate, Ramon Parson’s group pinpointed a minimal area of genomic deletion at 10q23.1, leading to the cloning of the candidate tumor suppressor gene PTEN (phosphatase and tensin homolog deleted on chromosome ten) (Li et al. 1997). Mutations were detected in brain and breast cancer. All PCA cell lines tested showed either mutations (i.e., LNCaP, DU145) or homozygous deletions (i.e., NCIH660, PC-3). Since the initial work, few inactivating PTEN mutations have been detected in PCA (Rubin et al. 2000). PTEN plays a critical role in regulating the PI3K-AKT pathway such that loss leads to downstream activation. More recent studies support nonclonal loss of PTEN in hormone-naïve tumors (Prandi et al. 2014) with enrichment with PCA progression (Armenia et al. 2018).

MYC Amplification (8q24)

C-Myc is a transcription factor with a wide range of functions, including modulation of protein synthesis, cell cycle, and metabolism. C-myc, the protein encoded by the MYC oncogene on 8q24, was observed in 1986 as overexpressed in human primary prostate cancer. Fleming et al. (1986) reported c-myc overexpression at the transcript level using Northern blot technology. In 1997, Jenkins and colleagues conducted the first extensive study using FISH at 8q24 to show gene amplification of MYC (Jenkins et al. 1997). Amplification of MYC was observed in 25% of the clinically localized PCA tumors, but in 46% of the advanced PCA samples examined, suggesting that MYC amplification corresponds to disease progression. Interestingly, they also observed that in the localized samples, MYC amplification was often only amplified in a subset of the tumor cells in the lesion, consistent with genomic heterogeneity. MYC is one of the genes that appears to be significantly altered in CRPC versus primary PCA (Fig. 3).

The importance of co-occurring molecular alterations is well illustrated by the amplification of MYC together with activation of the PI3K-pathway. Clegg et al. (2011) observed that there is a statistically significant association between PI3K-pathway alterations (i.e., PTEN, PI3CA, AKT1, AKT2, and AKT3) and MYC amplification, with 27% and 70% co-occurrence in localized and metastatic PCA, respectively. To determine the potential impact of these co-occurring genomic alterations, they developed a series of
genetically engineered mouse models (GEMs) to explore the relationships of the individual and co-occurring alterations. Using mice with either PTEN loss or AKT overexpression and crossing them respectively with high MYC overexpressing mice (Ellwood-Yen et al. 2003) in a prostate conditional context, they showed that the addition of MYC leads to an acceleration of PIN and adenocarcinoma. Interestingly, although RAD001, a rapamycin analog, can inhibit the formation of PIN in prostate conditional AKT activated GEMs, RAD001 did not abrogate the development of PIN in mice expressing both AKT and MYC. This suggests that MYC acts in a manner that is independent from mammalian target of rapamycin (mTORC1) activation. These important studies begin to reveal the complexity of co-occurring genomic alterations in cancer, the additional challenges to therapeutic strategies, and the need to better understand them through model systems.

**SOMATIC COPY NUMBER ALTERATIONS DEFINE PROSTATE CANCER RISK PROGRESSION**

Comparative genomic hybridization (CGH) technology led to some of the first genome-wide observations regarding copy number alterations (CNAs) and disease state (Cher et al. 1994; Visakorpi et al. 1995b). CGH identified somatic CNAs in a high percentage (~75%) of localized PCA. Losses were found to be five times more common than gains and most often involved 8p (32%), 13q (32%), 6q (22%), 16q (19%), 18q (19%), and 9p (16%). These early genome-wide studies also suggested that the pattern of CNAs changes with disease progression. For example, gains of 7, 8q, and X were more often observed in the CRPC state (Visakorpi et al. 1995b). A next advance in array-based technology was the use of single-nucleotide polymorphism (SNP) arrays to access somatic copy number state (Dumur et al. 2003; Lin et al. 2004). These SNP arrays were initially developed for germline population-based studies such as the Human HapMap study. However, investigators using paired tumor and normal samples from the same individual could determine not only copy number state (loss and gain), but also LOH. As the density of the arrays increased, smaller alterations could be detected across the entire genome with increasing precision in the definition of the lesion boundaries. This platform was most recently used by the TCGA study to perform a comprehensive copy number analysis for localized PCA (Cancer Genome Atlas Research Network 2015).

---

**Figure 3.** Genes altered in castration-resistant prostate cancer (CRPC) versus primary prostate cancer (PCA). Enrichment of somatic genomic alterations in CRPC. Alteration percentages in CRPC ($n = 335$) are shown on the $x$-axis, primary PCA ($n = 583$) on the $y$-axis. The significance of enrichment (Fisher’s test $q$-value) is shown by the size of the dots. (From Robinson et al. 2015; reprinted, with permission, from Elsevier 2015.)
Both array CGH and SNP arrays were frequently used to perform genome-wide studies and identify with greater precision potential candidate genes and loci associated with disease progression. For example, Lapointe and colleagues proposed a subclassification of PCA risk based on aCGH data with three potential risk groups (Lapointe et al. 2007). Demichelis and colleagues reported on SCNA-based PCA subclasses using SNP and tiling arrays (Demichelis et al. 2009). More recently, refined maps of somatic CNAs have shown significant association with Gleason grade prognostic groups (Fig. 4) (Rubin et al. 2016) and overall SCNA as a measure of PCA relapse (Hieronymus et al. 2014). In recent studies of populations of CRPC, distinct SCNA in AR and MYC have been confirmed from prior work (Robinson et al. 2015). One important technical observation is that, although SCNA can be detected from NGS whole exome data, the quality of the calls are potentially inferior to the SNP or CGH platforms. This is because of the smaller genomic coverage and often modest depth of sequencing of large-scale sequencing experiments, together with the lack of sufficient intergenic SNPs to make high-quality copy number assessment.

ETS REARRANGEMENTS IN PROSTATE CANCER

There is an emerging view of PCA suggesting a high degree of molecular complexity involving common recurrent gene fusions (Tomlins et al. 2005), large complex genomic rearrangements (Berger et al. 2011), and common recurrent mutations (Berger et al. 2011; Barbieri et al. 2012; Baca et al. 2013; Cancer Genome Atlas Research Network 2015). This heterogeneity might be best viewed as a collection of more homogenous subgroups defined by somatic CNAs, mutations, and rearrangements. This section focuses on the recurrent gene fusions involving most commonly ETS transcription factors. The most common gene fusion in PCA is the TMPRSS2:ERG fusion (Tomlins et al. 2005). Since its initial discovery in 2005, a great deal has been learned about this common molecular event. One of the most important concepts regarding the TMPRSS2:ERG fusion is to what extent the mutation plays a role as a biologically driving event.

Hematological malignancies, such as acute myeloid leukemia (AML), are often subtyped based on the recurrent cytogenetic and molecular aberrations identified. Therefore, the discovery that 50% of PCA harbor recurrent gene rearrangements resulting in the fusion of genes (Tomlins et al. 2005), greatly facilitates a molecular subtyping of PCA, similar to what has been established for leukemias and lymphomas. In PCA, these fusions juxtapose a hormone-sensitive promoter and an ETS transcription factor that is believed to act in an oncogenic role, conferring a distinct biology to this tumor. These fusions, unlike the bcr-abl fusion in chronic myeloid leukemia (CML), do not create a chimeric protein but instead over express an ETS transcription factor in a normal, albeit truncated form. Although other molecular events play a role in PCA development and progression, defining PCA based on the presence or absence of the different gene fusions that drive cancer development provides novel insight into disease heterogeneity.

Computational Biology and the Discovery of Gene Fusion Prostate Cancer

The discovery of TMPRSS2:ETS gene fusions from a simple, statistical approach, termed “Cancer Outlier Profile Analysis” (COPA) to identify oncogene profiles in a subset of samples within publicly available cancer profiling data sets, characteristic of genes commonly associated with known genomic rearrangements (reviewed by Rubin and Chinnaiyan 2006; Hanauer et al. 2007). The application of COPA in PCA microarray experiments revealed two consistently high-scoring, and mutually exclusive candidates across 50% of PCA samples that were members of the ETS family of transcription factors, ERG and ETV1. Further experiments revealed fusions of the 5’-untranslated region of TMPRSS2 (21q22.3) with ETS transcription factor family members, either ERG (21q22.2), ETV1 (7p21.2) (Tomlins et al. 2005), or ETV4.
Figure 4. Landscape of somatic copy number alterations (CNAs) across grading groups. Landscape of somatic CNAs from 426 localized prostate cancers (PCAs) stratified by grading group and ordered from 1 (low) to 5 (high). Blue denotes deletions; red denotes amplifications. (From Rubin et al. 2016; reprinted, with permission, from Elsevier 2016.)
(Tomlins et al. 2006), suggesting a novel mechanism for overexpression of the ETS genes in PCA.

## A Multitude of Gene Fusions in Prostate Cancer

Since the initial discovery of ETS fusions in PCA, a number of more recent studies have identified fusion events involving additional ETS family members (i.e., ELK4) (Maher et al. 2009a; Rickman et al. 2009), novel 5-prime (upstream) partners, and a class of non-ETS based fusions (Fig. 5) (Rubin et al. 2011).

The largest category, ETS fusions, is comprised of the highly recurrent TMPRSS2:ERG fusion, which contrasts with the remaining, less common, fusion events. Interestingly, the ETS family member fusions involve a diverse set of 5-prime upstream partners, as exemplified by ETV1 having at least nine different fusion partners. In addition to TMPRSS2, three other

![Figure 5. Prostate cancer gene fusion classification. The ongoing effort to screen prostate cancer patients for gene fusions, in combination with the recent technology advances, has resulted in a comprehensive gene fusion landscape. This schematic highlights all published gene fusions categorized into ETS rearrangements, RAF kinase gene fusions, and SPINK1-positive, ETS rearrangement–negative prostate cancers. The percentages highlight the estimated frequency of each gene fusion on the basis of published screens. (Based on data in Rubin et al. 2011.)](http://perspectivesinmedicine.cshlp.org/)

---

**ETS**

<table>
<thead>
<tr>
<th>Gene</th>
<th>5′-10′ Partners</th>
</tr>
</thead>
<tbody>
<tr>
<td>TMPRSS2</td>
<td>1</td>
</tr>
<tr>
<td>SLC45A3</td>
<td>1</td>
</tr>
<tr>
<td>NDRG1</td>
<td>1</td>
</tr>
<tr>
<td>ETV1</td>
<td>4</td>
</tr>
<tr>
<td>ETV4</td>
<td>1</td>
</tr>
<tr>
<td>ETV5</td>
<td>2</td>
</tr>
<tr>
<td>ESRP1</td>
<td>1</td>
</tr>
</tbody>
</table>

**RAF**

<table>
<thead>
<tr>
<th>Gene</th>
<th>5′-10′ Partners</th>
</tr>
</thead>
<tbody>
<tr>
<td>CRAF</td>
<td>6</td>
</tr>
<tr>
<td>ESRP1</td>
<td>14-15</td>
</tr>
<tr>
<td>SLC45A3</td>
<td>1</td>
</tr>
</tbody>
</table>

**SPINK1**

<table>
<thead>
<tr>
<th>5′-10′ Partners</th>
</tr>
</thead>
<tbody>
<tr>
<td>SPINK1 8%</td>
</tr>
</tbody>
</table>

**SPOP Mutations**

<table>
<thead>
<tr>
<th>5′-10′ Partners</th>
</tr>
</thead>
<tbody>
<tr>
<td>SPOP 13%</td>
</tr>
</tbody>
</table>
androgen-responsive 5’ partners SLC45A3 (Han et al. 2008; Esgueva et al. 2010), HERPUD1 (Maher et al. 2009b), and NDRG1 (Pflueger et al. 2009) have been found to fuse with ERG. However, many of the 5’ partners appear to fuse to multiple ETS family members, such as SLC45A3 (-ERG, -ELK4, -ETV1, and -ETV5) and TMPRSS2 (-ERG, -ETV1, -ETV4, and -ETV5), both of which are androgen-responsive. The majority of these AR regulated promoters confer an organ/tissue specificity to these gene fusions. Interestingly, as these events occur as early as the precursor lesion, high-grade PIN, they suggest one of the first hormonally regulated mutations in PCA development. This may have implication in how individual men respond to endogenous hormone and/or hormone manipulation as part of systemic treatment for more aggressive PCA.

Recent advances in next-generation transcriptome sequencing have facilitated the discovery of RAF kinase gene fusions, SLC45A3-BRAF, ESRP1-RAFI1, and RAF1-ESRP1 in advanced PCA (Palanisamy et al. 2010). Although rare, detected in ~1%–2% of prostate cancers, RAF kinase fusions represent the first “driver” fusion in PCA that do not involve an ETS family member.

The TMPRSS2:ETS Fusion Is a Common, Early Gatekeeper of Prostate Cancer

Microscopic examination of PCA using an in situ FISH assay reveal that gene fusion occurs in neoplastic cells but not in adjacent benign nuclei or stromal cells (Tomlins et al. 2005; Perner et al. 2006; Mosquera et al. 2008). A larger study drawn from a wide spectrum of benign prostatic lesions and precursors of PCA (Perner et al. 2007) failed to detect TMPRSS2:ERG fusion in benign prostate tissue, BPH, or PIA (also commonly referred to as “focal prostate atrophy” or “prostate atrophy”) (for a review, see De Marzo et al. 2006). The TMPRSS2:ERG fusion was observed in ~20% of high-grade PIN lesions intermingled with prostate cancer that carried the same fusion pattern. This was the same frequency previously detected by Cerveira et al. (2006) using an RT-PCR based assay. More recently, immunohistochemistry has been used to evaluate the gene fusions in situ (Park et al. 2010). Using an antibody highly specific for ERG rearrangements, one can detect the earliest overexpression of the ERG oncogene in the morphologic area of high-grade PIN, but not in directly adjacent benign prostate tissue. A significant clinical implication for this finding is the potential utility of assessing the TMPRSS2:ERG fusion status in diagnostically problematic prostate needle core biopsies with high-grade PIN and adjacent small atypical glands.

The prevalence of TMPRSS2:ERG PCA has been reported to range from 40%–70%, depending on the clinical cohorts investigated. The first large clinical study on a German prostatectomy cohort reported ~50% of cases with a TMPRSS2:ERG fusion (Perner et al. 2006). Multiple, retrospective studies from PSA-screened prostatectomy cohorts have reported frequencies of the TMPRSS2:ERG fusion between 35% and 50% when fluorescence in situ assay (FISH) was used to detect the rearrangement (Yoshimoto et al. 2006; Attard et al. 2007; Clark et al. 2007; Mehra et al. 2007b; Rajput et al. 2007; Tu et al. 2007). This has now been confirmed in the TCGA study (Cancer Genome Atlas Research Network 2015), which also includes the TMPRSS2:FLI1 fusion that had been predicted owing to the similarity to ERG (i.e., they both belong to the same clade of ETS transcription factors). The exact frequency of more rare fusions may be below 1% but could also be more common in ethnic groups that have yet to be characterized.

The common ETS gene fusions are widely believed to have an important oncogenic role, but ERG or ETV1 over expression alone is not sufficient to lead to PCA. Some evidence suggests that ETV1 has a stronger phenotype than ERG in PCA disease progression (Baena et al. 2013). There is now mounting molecular data for an important concomitant role of TMPRSS2:ERG and Pten/PI3K/ATK pathway activation in PCA oncogenesis. Carver et al. (2009) examined data from human samples and determined that high-grade PIN shows ERG gene rearrangement in 10% of cases, whereas loss of PTEN as determined by immunohistochemistry was observed

The Genomics of Prostate Cancer

Advanced Online Article. Cite this article as Cold Spring Harb Perspect Med doi: 10.1101/cshperspect.a034942
perspectivesinmedicine.org

M.A. Rubin and F. Demichelis

in 45% of cases. Similar to previous observations on high-grade PIN with ERG rearrangement (Perner et al. 2007; Mosquera et al. 2008), the associated PCA also showed the same gene fusion status. They interpreted this data to be consistent with TMPRSS2:ERG fusion representing an early molecular event facilitating the transition from high-grade PIN to PCA. They further analyzed tumor samples demonstrating that 68% (27/40) of PCA had loss of PTEN expression by immunohistochemistry in contrast to 38% (15/40) of cases with ERG rearrangement (Carver et al. 2009). Interestingly, 14 of 15 ERG rearranged cases showed PTEN loss. King et al. (2009) presented almost identical data with 14 of 17 TMPRSS2:ERG fusion PCA showing PTEN loss. These human tumor data support the hypothesis that PTEN and ERG are concomitant molecular lesions. This data has now been supported by more recent next-generation sequencing studies.

Studies that have overexpressed ETV1, ETV5, and ERG have shown an increase in cell migration capability, not an increase in proliferation or the ability to transform these cells into tumor cells (Tomlins et al. 2007a, 2008; Helgeson et al. 2008). Klezovitch et al. (2008) also showed that the overexpression of ERG is associated with tumor cell migration through a proteolytic molecular program. Carver et al. (2009) performed a series of studies to explore the concomitant relationship between TMPRSS2:ERG and the Pten/PI3K/Akt pathways. Benign prostate cells (BPH-1) show increased proliferation with constitutively overexpressed Akt (AKT-1) as compared with controls. ERG overexpression did not lead to an increase in cell proliferation. In contrast, cell migration studies showed an additive effect with ERG and Akt. These results suggest that the oncogenic role of TMPRSS2:ERG fusion is in tumor cell migration that is enhanced by the proliferative effects of Pten/PI3K/Akt pathway activation.

The first mouse models used a probasin promoter to overexpress ERG (Klezovitch et al. 2008; Tomlins et al. 2008) and ETV1 (Tomlins et al. 2007a). These models showed the ability of the transgene (i.e., ERG or ETV1) to develop early molecular changes referred to as mouse PIN (Shappell et al. 2004). These subtle changes did not reach the level of invasive cancer.

Based on observations described above in human samples, more recent work explored the in vivo significance of TMPRSS2:ERG fusion in the context of Pten/PI3K/Akt pathway activation. The Pandolfi group had previously showed that a transgenic mouse model with loss of one Pten allele (Pten\(^{−/−}\)) develops mouse high-grade PIN after a long latency period with incomplete penetrance (Di Cristofano et al. 1998). By developing a Pten\(^{+/−}\) transgenic mouse that overexpresses ERG using a modified probasin promoter (ARR\(_{PB}\)), they observed a florid high-grade PIN with a shorter latency period and more complete penetrance as compared with the Pten\(^{−/−}\) mouse. Taken together, Carver and colleagues concluded that alterations of Pten/PI3K/Akt pathway play an early role in tumor cell proliferation and the TMPRSS2:ERG gene fusion contributes through cell migration. They posited that the combination lead to an aggressive phenotype as suggested by Yoshimoto et al. (2008). King et al. (2009) reported supportive data using a different mouse model that overexpresses the most common TMPRSS2:ERG isoform under control of a probasin promoter. This TMPRSS2:ERG model does not show an apparent phenotype. However, bigenic mice with either TMPRSS2:ERG; Pten\(^{+/−}\) or TMPRSS2:ERG and probasin driven Akt showed high-grade PIN. They also tested the potential concomitant effect of TMPRSS2:ERG with MYC overexpressing mice (Hi-Myc) but did not see any additional effect; the Hi-Myc mice alone develop PCA. These data are also consistent with clinical data that failed to show associations with MYC amplification and ERG rearrangement. Both of these studies suggest that Pten/PI3K/Akt and TMPRSS2:ERG alterations need to be explored as important events in PCA disease progression.

**Gene Fusion Is a Clonal Event and Helps Understand Prostate Cancer Heterogeneity**

TMPRSS2:ERG fusions, when present, are distributed evenly amongst all tumor nuclei within a discrete tumor lesion. We reported that 243...
out of 246 prostate cancer cases showed homogeneity within a discrete tumor nodule (Perner et al. 2006). This observation was extended when multiple microdissected foci of cancer from individual patients were examined by RT-PCR for gene fusions and showed either all or no foci overexpressed ERG and its family members ETV1 and ETV4 (Tomlins et al. 2007b). Thus, within a discrete nodule, the fusion rearrangement must occur early because, if present, all of the tumor nuclei harbor the fusion. However, when we undertook studies to evaluate rearrangements between the multiple nodules within a single prostate gland from one individual, each discrete lesion may occur independently from one another. This has been observed in three independently conducted studies (Barry et al. 2007; Mehra et al. 2007a; Clark et al. 2008). For example, in the study by Barry and colleagues, 32 prostatectomy samples with clear-cut discrete tumors showed fusion by balanced translocation and fusion by interstitial deletion occurring as distinct events, suggesting that these are clonal mechanisms for achieving TMPRSS2:ERG fusion. Interestingly, that study found a high rate (41%) of interfocal heterogeneity for fusion status. These observations have both biological and clinical implications. Biologically, the presence of multiple clonally distinct lesions suggests that within a single gland, complex molecular events such as the gene rearrangement can occur in some but not all lesions. This makes classifying PCA more challenging. From a clinical perspective, how does one determine the most aggressive nodule to target? It has been long assumed that the dominant nodule harbors the most aggressive tumor and therefore dictates the clinical course.

Emerging Understanding of Prostate Cancer Genomic Complexity

Genomic rearrangements appear to be nonrandom, locus-specific and depend, in part, on the proximity of chromosomal regions in the nucleus (Mani and Chinnaiyan 2010). Moreover, there is mounting evidence suggesting that transcription factors are associated with DNA double-strand breaks, thus predisposing transcribed regions to genomic rearrangements. For example, both androgen and estrogen signaling recruit the enzyme topoisomerase 2B (TOP2B) to target gene promoters, which creates DNA double-strand breaks and facilitates translocation (Ju et al. 2006; Haffner et al. 2010). The AR and TOP2B are coexpressed in human PCA precursor lesions in which TMPRSS2:ERG rearrangements are known to occur, suggesting a critical role of TOP2B in the recurrent ETS rearrangements. Three studies have also shown that androgen signaling promotes TMPRSS2:ERG fusion formation (Lin et al. 2009; Mani et al. 2009; Bastus et al. 2010), in part by recruiting DNA break-inducing enzymes (e.g., activation of induced cytidine deaminase (AID) to translocation breakpoint sites (Lin et al. 2009). More recently, Berger et al. (2011) showed that rearrangement breakpoints were enriched near open chromatin, AR and ERG DNA-binding sites in the setting of the ETS gene fusion TMPRSS2:ERG, but inversely correlated with these regions in tumors lacking ETS fusions. Hence, transcription factors can contribute to the formation of genomic rearrangements by facilitating the juxtaposition of chromosomal loci and recruiting enzymatic machinery involved in DNA breaks to these target loci.

More recent NGS studies also provided important insights into DNA genomic complexity (see below). It probably is important to note that recurrent gene fusions such as TMPRSS2:ERG are a relatively rare event when compared with all of the rearrangements and structural variants now observed by NGS studies. It is interesting to note that there are probably many rare, nonrecurrent gene fusions that occur and do provide some growth advantage to tumor cells. Pfueger et al. (2009), reported non-ETS fusions involving CDKN1A (p21), CD9, and IKBKB (IKK-β) that were present at both the transcript and DNA levels. However, this RNA-seq study foreshadowed an observation that would be more definitively made in whole-genome sequencing studies, namely that ETS fusion PCA are more prone to interchromosomal rearrangements as a class (Berger et al. 2011; Baca et al. 2013), and hence the majority of gene fusions found in the...
RNAseq studies tend to occur in ETS fusion cancers.

In summary, ETS fusions are the most common recurrent genetic mutation identified in PCA. Although a number of ETS and non-ETS family members have been observed to be fused with TMPRSS2 or other 5’ partners, the vast majority of fusions involve TMPRSS2:ERG. This fusion can be studied in large numbers, as it was identified in ~45% of all PSA screened PCA.

MOLECULAR LANDSCAPE OF PROSTATE CANCER IN THE NEXT-GENERATION SEQUENCING ERA

Advances in high-throughput technologies, such as NGS, have provided extraordinary insight into the PCA genome/transcriptome (Tomlins et al. 2007b; Taylor et al. 2010; Berger et al. 2011; Rubin et al. 2011; Barbieri et al. 2012; Grasso et al. 2012b; Baca et al. 2013; Beltran and Rubin 2013; Lindberg et al. 2013; Tomlins 2013; Weischenfeldt et al. 2013). There have been numerous recent NGS genomic studies that largely confirm many of the known genomic events in PCA such as PTEN loss, TP53 mutation/loss, and ETS gene fusions. One important observation is that some events such as ETS gene fusions and SPOP mutation (the most common point mutation in primary PCA, see below) are mutually exclusive, leading to the view that PCA represent a collection of potentially definable molecular subclasses (Rubin et al. 2011; Beltran and Rubin 2013; Tomlins 2013). Such subtyping is largely based on the presence or absence of recurrent gene fusions between the androgen regulated gene TMPRSS2, and ETS family members (most commonly ERG). Comprehensive copy number profiling, whole-exome sequencing (WES) and whole-genome sequencing (WGS) studies characterizing the PCA genomic landscape have identified a few highly recurrent somatically mutated genes (including SPOP, TP53, PTEN, and FOXA1, all <15%), with recurrent broad CNAs (i.e., 8p loss and 8q gain), but relatively few focal and/or high-level CNAs (most commonly focal PTEN, TP53, and RB1 losses). The recent TCGA publication of 333 prostate cancer genomes, transcriptomes, and methylomes solidify the idea of PCA molecular subclasses (Fig. 6) (Cancer Genome Atlas Research Network 2015). One major advantage of NGS generated tumor data sets is the possibility to computationally assess tumor cellularity (fraction of tumor cells in cells constellation whose DNA was analyzed, also referred to as tumor purity) by exploiting single base resolution data and to infer tumor aneuploidy, as accomplished by various tools including ABSOLUTE (Carter et al. 2012) and CLONET (Prandi et al. 2014). In addition, NGS data make it possible to assess the clonality of each single somatic lesion (fraction of tumor cells that harbor a specific somatic lesion) allowing for the definition of the first primary PCA evolution chart in 2013 (Fig. 7) (Baca et al. 2013), and of phylogenic trees to trace tumor history in one patient when for instance multiple metastatic biopsies are sequenced (Gundem et al. 2015; Kumar et al. 2016). Together, these computational assessments from large collection of prostatectomies samples characterized through WES or WGS deepened our recognition of earlier versus later events in prostate carcinogenesis.

Insights into Genomic Complexity

With the advent of WGS, insights into global genomic rearrangements are possible. Berger et al. (2011) published the first seven whole genomes (Fig. 8), revealing structurally complex rearrangements. The associations were significantly enriched for disruption of tumor suppressor genes or their orthologs. There was also an observation the rearrangements in the ETS rearrangement cancers tended to occur in the proximity of transcriptionally open chromatin marks, suggesting that subclasses of prostate cancer may act differently based on early genomic alterations. In a significantly larger WGS study, Baca et al. (2013) extended these observations based on 57 prostate cancers. Using statistical arguments and simulations, they found that the many complex rearrangements likely originated within a single cell cycle, providing a proliferative advantage to a cancerous cell causing multiple oncogenic events. This was termed “chromoplexy” to refer to these distinc-
Figure 6. Recurrent alterations in primary prostate cancer. The spectrum and type of recurrent alterations and genes (mutations, fusions, deletions, and overexpression) in the cohort are shown (left to right) grouped by main molecular subtypes. On the right, the statistical significance of individual mutant genes is shown. Mutations in IDH1, PIK3CA, RB1, KMT2D, CHD1, BRCA2, and CDK12 are also shown, despite their not being statistically significant. SPINK1 overexpression is shown for reference. (From Cancer Genome Atlas Research Network 2015; reprinted, with permission from Elsevier 2015.)
tive genomic restructuring events. They showed that this process is distinct from the phenomenon of chromothripsis (Stephens et al. 2011). Chromoplexy is active across the majority of prostate cancers and in several nonprostate tumor lineages. Similarly, in a recent publication, Fraser and colleagues (2017) describe results from a large WGS effort on localized, intermediate-risk PCA that includes specific single-nucleotide variants and rearrangement signatures with DNA footprints associated with aging and clustered-mutation phenomenon called kataegis. This large WGS study shows a high burden of complex rearrangements, including samples with evidence of chromothripsis (chromosome shattering). Although not reported, their study also showed chromoplexy (chains of balanced rearrangements), suggesting highly altered 3D structures of PCA cell genomes (personal communication).

Another observation from WGS analysis revealed that, although the ETS fusion cases showed a high frequency of intergenic gene rearrangements, a second group of ~10% of PCA show high frequencies of intragenic rearrangements. This later class is best defined as harboring SPOP mutations and CHD1 loss (Boysen et al. 2015). This recent work showed that SPOP mutations are associated with a homology repair defect phenotype.

**SPOP Mutations Define a Distinct Molecular Subclass of Prostate Cancer**

Recurrent missense mutations in SPOP are the most common point mutations in primary PCA, occurring in ~10% of both clinically localized and metastatic CRPC (Berger et al. 2011; Barbieri et al. 2012; Blattner et al. 2014; Cancer Genome Atlas Research Network 2015; Robin-
Figure 8. Graphical representation of seven prostate cancer genomes. Each Circos plot depicts the genomic location in the outer ring and chromosomal copy number in the inner ring (red, copy gain; blue, copy loss). Interchromosomal translocations and intrachromosomal rearrangements are shown in purple and green, respectively. Genomes are organized according to the presence (top row) or absence (bottom row) of the TMPRSS2-ERG gene fusion. (From Berger et al. 2011; reprinted, with permission, from the authors.)
The most common hot spot SPOP mutations occur at F133, Y87, F102, and W131 and often co-occur with specific genomic features including deletions at 5q21, 6q15, and 2q21 (Barbieri et al. 2012; Boysen et al. 2015). Molecularly, human PCA can be classified into those harboring rearrangements in ETS transcription factors (e.g., TMPRSS2-ERG) and those lacking ETS rearrangements. SPOP mutant PCA also defines characteristic genomic rearrangements, gene expression profiles, and methylation patterns (Barbieri et al. 2012; Blattner et al. 2014; Boysen et al. 2015; Cancer Genome Atlas Research Network 2015). SPOP mutations occur early in the natural history of prostate cancer solely as heterozygous missense mutations with dominant-negative, selective loss-of-function toward the remaining wild-type allele (Baca et al. 2013; Prandi et al. 2014; Theurillat et al. 2014; Boysen et al. 2015).

SPOP encodes the substrate recognition component of a CUL3-based E3 ubiquitin ligase, and PCA derived SPOP mutants appear to act as dominant-negative with selective loss-of-function (Boysen et al. 2015). Known substrates of SPOP are numerous, and the specific substrates that are deregulated by SPOP mutations are starting to be defined. These include the chromatin-associated oncopogene DEK (Theurillat et al. 2014), the oncogenic coactivator TRIM24 (Theurillat et al. 2014; Groner et al. 2016), and the AR itself (An et al. 2014; Geng et al. 2014).

Initial models have established the role of SPOP mutation as a driver of prostate neoplasia in vivo, and studies exploring the downstream effects of SPOP mutations have largely relied on overexpression of mutant SPOP protein in cell lines with alterations outside the genetic context of SPOP mutant prostate cancer (Geng et al. 2013, 2014; González-Billalabeitia et al. 2014; Theurillat et al. 2014; Pierce et al. 2015).

More recently, Blattner et al. (2017) reported the development of the first conditional mouse model showing that SPOP mutation drives prostate tumorigenesis in vivo. Mice conditionally expressing mutant SPOP in the prostate have minimal histologic phenotype but show focal areas of cytologic atypia. In contrast, mutant SPOP results in early high-grade PIN with striking nuclear atypia in the setting of heterozygous Pten loss, and invasive, poorly differentiated carcinoma with homozygous Pten loss. PTEN deletions and mutations, although rare in the early phases of SPOP mutant human prostate cancer (Barbieri et al. 2012; Cancer Genome Atlas Research Network 2015), become more frequent in CRPC (Grasso et al. 2012a; Robinson et al. 2015) suggesting that PTEN deletion may contribute to progression of SPOP mutant prostate cancer (Haffner et al. 2013; Barbieri et al. 2014). Using in vitro models derived from these mice, they showed that mutant SPOP activated both PI3K/mTOR signaling and AR signaling, effectively uncoupling the normal negative feedback between these two pathways. Together, these findings show that SPOP mutation drives prostate neoplasia in vivo through deregulation of the PI3K/mTOR and AR pathways, and underscore the critical role of these two signaling pathways across molecular subtypes of human prostate cancer (Fig. 9).

GENOMIC ALTERATIONS AND RESISTANCE

It has been long observed that AR mutations are only seen in the setting of treated prostate cancer. Recent genomics analysis has provided a deeper understanding as to the extent and repertoire of alterations associated with resistance.

Multiple studies have now assessed the molecular landscape of lethal CRPC (predominantly from rapid autopsy samples) (Tomlins et al. 2007b; Liu et al. 2009; Grasso et al. 2012b; Aryee et al. 2013; Gundem et al. 2015; Hong et al. 2015). WES and targeted NGS studies have shown that AR is the most recurrent alteration (~50%) in CPRC but not in untreated PCA (from unrelated patients or previous studies), consistent with the known role of AR in mediating CRPC progression (Grossmann et al. 2001; Chen et al. 2004, 2008; Richter et al. 2007; Bluemn and Nelson 2012; Bastos et al. 2014; Hovelson et al. 2015). Nevertheless, numerous genes are altered more frequently in CRPC versus untreated PCA (i.e., TP53 and PTEN), and additional potential low-frequency driving alterations are observed more frequent-
ly in lethal CRPC than untreated PCA (i.e., deleterious APC alterations in 20% of CRPC and 3%–5% of untreated PCA from TCGA and our previous untreated PCA studies [Barbieri et al. 2012; Grasso et al. 2012b; Baca et al. 2013; Hovelson et al. 2015]). Likewise, single-gene, functional, and profiling studies have identified pathways that may mediate progression to CRPC along with AR pathway deregulation, such as the WNT/APC/CTNNB1 or PI3K pathways (Carver et al. 2011; Kumar et al. 2011; Grasso et al. 2012b; Kypta and Waxman 2012; Karantanos et al. 2013; Rajan et al. 2014; Hovelson et al. 2015); however, it is unclear if these alterations are present before ADT, mediate progression to CRPC, or are acquired after CRPC progression.

In recent work from Armenia et al. (2018), more than 1000 prostate cancer genomes were compared. The samples were all run on a similar computational pipeline facilitating a proper comparison. One of the important observations is the alterations that are statistically increased or decreased with disease progression. As noted above, these are AR, PTEN, and MYC, as well as some less common alterations (Robinson et al. 2015).

**AR Indifferent Prostate Cancer: Another Mechanism for Resistance**

Treatment-related neuroendocrine prostate cancers (NEPC) are clonally derived from a castration-resistant adenocarcinoma precursor. Although they retain prostate cancer genomic alterations, NEPC tend to lose expression of common prostate markers and are often treated with regimens similar to small cell lung cancer. There is a spectrum of disease within CRPC, with small cell carcinoma being the extreme; mixed and intermediate atypical phenotypes are also observed, which can retain AR expression and harbor overlapping molecular features. Clinical features associated with platinum sensitivity (termed “anaplastic” or “aggressive variant” prostate cancer) have been defined for patient selection in Phase 2 chemotherapy studies and share pathologic and molecular features of NEPC. The cell plasticity associated with NEPC is associated with decreased or absent expres-
sion of the AR and downstream androgen-regulated genes such as prostate-specific antigen (PSA). Moreover, data from preclinical models also favor a transdifferentiation model of treatment-related NEPC. For example, adenocarcinoma cells (e.g., LNCaP) display a neuroendocrine phenotype similar to the few established NEPC cell lines that have been described (e.g., NCI-H660) (Lai et al. 1995; Mertz et al. 2007; Beltran et al. 2011) following exposure to a variety of stimuli (e.g., androgen deprivation [Burchardt et al. 1999] or treatment with cAMP [Bang et al. 1994], IL-6 [Qiu et al. 1998] or fractionated ionizing radiation [Deng et al. 2008]). Similar results have been observed in vivo. For example, a well characterized patient-derived prostate adenocarcinoma xenografts (LTL331) implanted into the subrenal capsule of mice develops small cell NEPC following castration (LTL331R) (Lin et al. 2014) and which phenotypically resembled other lines that were generated from bona fide NEPC tumor tissue (e.g., LTL352, LTL370 [Lin et al. 2014], LuCAP-49 [True et al. 2002], UCRU-PR-2 [van Haaften-Day et al. 1987], and WISH-PC2 [Pinthus et al. 2000]).

Activation of Myc family oncogenes commonly occur across poorly differentiated neuroendocrine carcinomas and in small cell lung cancer, up to 20% harbor genomic amplification of N-myc, C-Myc, and L-Myc, or a recurrent fusion involving MYCL1 (9%) (Wistuba et al. 2001; Huijbers et al. 2014; Rudin and Poirier 2014; Teicher 2014). N-Myc amplification occurs in high-risk neuroblastoma and a subset of NEPC and SCLC. N-Myc has also been shown to be amplified and overexpressed in NEPC and can drive the NEPC phenotype (Beltran et al. 2011; Mosquera et al. 2013; Lee et al. 2016). Recently, Dardenne et al. (2016) used a variety of preclinical models including isogenic cell lines, xenografts, GEMMs, and mouse tumor organoids to show that N-Myc overexpression is associated with highly proliferative, invasive prostate cancer with pathologic features of NEPC (Dardenne et al. 2016). Dardenne and colleagues further showed that N-Myc interacts and cooperates with Enhancer of Zeste Homolog 2 (EZH2) to drive NEPC transcriptional programs including an abrogation of AR signaling, enhanced AKT and EMT signaling, repression of Polycomb Repressive Complex 2 (PRC2) target genes, and expression of neuroendocrine markers. Lee et al. (2016) used forward engineering of benign human prostate cells that overexpress N-Myc in the context of myristoylated AKT and provided compelling data showing that N-Myc overexpression resulted in castrate resistant tumors with NEPC or adenocarcinoma histology foci. These complementary results solidify N-Myc as a driver of the NEPC phenotype. In both studies, N-Myc was shown to form a complex with Aurora kinase A (Aurora-A) that stabilized N-Myc and that was sensitive to allosteric inhibitors to Aurora kinase A. These results showed the feasibility of exploiting the mutual dependence of N-Myc and Aurora-A to revert their oncogenic functions.

Although N-Myc is overexpressed in the majority of NEPC cases, there is a spectrum of N-Myc expression in CRPC samples with 20% of CRPC tumors demonstrating transcript levels in the range of NEPC (Beltran et al. 2016; Dardenne et al. 2016). N-Myc overexpression in prostate adenocarcinoma was sufficient to drive the NE phenotype (Dardenne et al. 2016). Specifically, RNA-seq analyses from multiple stable LNCaP or 22Rv1 cell populations overexpressing N-Myc showed that N-Myc overexpression resulted in repression of AR signaling and activation of PRC2 signaling and gene expression signatures associated with RB1 loss (Dardenne et al. 2016), all of which are molecular features associated with NEPC (Beltran et al. 2011, 2016; Lee et al. 2016). These results suggest that N-Myc is at least one driver of NE plasticity in prostate cancer. Recent preclinical data also suggest that lineage plasticity can be observed in the context of RB1/TP53 loss (Ku et al. 2017; Mu et al. 2017). EZH2 inhibitors can also sensitize NEPC tumor cells to enzalutamide. This suggests that reversing or delaying lineage plasticity with this type of epigenetic approach may provide a clinical benefit to a larger number of patients than previously appreciated and would extend clinical responses to antiandrogen therapy in the case of prostate cancer.
HETEROGENEITY

The Impact of Multifocality and Heterogeneity on Tracking Lethal CRPC

At radical prostatectomy, ~80% of patients harbor multiclonal (also referred to as multifocal) PCA, where spatially distinct tumor foci, which may show similar morphology and/or grade (Gleason score), are present in the same prostate (Arora et al. 2004; Meiers et al. 2007). Multifocal PCA represent clones of independent origin, as supported by numerous approaches, including ERG rearrangement status (by FISH or IHC) (Sakr et al. 1994; Qian et al. 1995; Cheng et al. 1998; Barry et al. 2007; Mehra et al. 2007a; Attard et al. 2008; Furusato et al. 2008; Kobayashi et al. 2008; Perner et al. 2010; Young et al. 2012; Lindberg et al. 2013; Kunju et al. 2014; Smith et al. 2014). In contrast, lethal, metastatic CRPC appears uniformly ERG rearrangement positive or negative in all sites in a given patient, consistent with clonal origin, although extensive subclonal structure is present (Mehra et al. 2008; Liu et al. 2009; Grasso et al. 2012b; Aryee et al. 2013; Nickerson et al. 2013; Gundem et al. 2015; Hong et al. 2015).

ERG rearrangement status (ERG+ or ERG−) is a useful clonal marker to show spatially distant multifocal tumors (Barry et al. 2007; Mehra et al. 2007a; Furusato et al. 2008, 2010; Perner et al. 2010; Young et al. 2012; Smith et al. 2014). Several anecdotal NGS studies (n’s ≤ 5–10) have added complexity to tracking the eventual CRPC clone through identifying intrafocal heterogeneity at RP (Aryee et al. 2013; Brooks et al. 2014; Lindberg et al. 2014; VanderWeele et al. 2014; Cooper et al. 2015; Gundem et al. 2015; Hong et al. 2015). These series of locally advanced PCA vary from reporting little divergence to complete lack of shared alterations between the index focus and lymph node metastases and/or CRPC. Haffner and colleagues tracked the lethal clone in a single patient (Aryee et al. 2013). Remarkably, they found that at RP, a small organ-confined low-grade (Gleason score 6) area of a large, high-grade primary tumor was the only area that harbored all alterations present in distant CRPC and lethal metastases. Critically, these alterations were absent from the vast majority of the primary tumor and lymph node metastasis at RP. Hence, in this patient, the lethal CRPC clone arose from a small, low-grade area of a histologically defined single index focus, rather than the higher-grade area or concurrent lymph node metastasis. Whether this “n of 1” case represents the exception, rather than the rule, can only be assessed in a large cohort of paired RP and CRPC specimens, rather than locally advanced PCA.

Gundem et al. (2015) recently explored PCA clonal evolution in 10 men with heavily treated CRPC at rapid autopsy. Like other published rapid autopsy series, this cohort did not represent a clinical trial and did not include patients treated with current second-line agents targeting AR signaling (enzalutamide and abiraterone). With these caveats, their study presents a key snapshot of heavily treated lethal CRPC. In their study, primary prostate tumors (retained during treatment of advanced disease) showed the presence of a large “trunk” of mutations seen subclonally. Among the mutations found in the trunk, a subset of potential driver mutations was observed in a more pure, clonal form in the metastatic lesions. They show the feasibility of tracking clonal mutations in metastases back to initiating lesions. Current ongoing studies from the SU2C-CRPC500 cohort are actively addressing this in a more formal manner.

In summary, the advent of the radical prostatectomy coupled with advances in sequencing technology led for steady increase in our understanding of prostate cancer genomics. More recently, the examination of patients treated for advanced CRPC either at time of autopsy or through metastatic biopsies is leading to important insights into tumor plasticity and resistance. Future advances in cancer mouse models and organoids derived from patients will enable the development of new treatment strategies.

ACKNOWLEDGMENTS

This review has intentionally focused on more recent work while trying to adopt a historic perspective for all the initial groundwork that set the foundation for our current understanding. We apologize sincerely in advance to authors.
who feel we have not included their studies in this article on genomics.

REFERENCES


Bookstein R, Shew JY, Chen PL, Scully P, Lee WH. 1990b. Suppression of tumorigenicity of human prostate carci-


The Genomics of Prostate Cancer

Advanced Online Article. Cite this article as Cold Spring Harb Perspect Med doi: 10.1101/cshperspect.a034942


The Genomics of Prostate Cancer


---

**Cite this article as:** Cold Spring Harb Perspect Med doi: 10.1101/cshperspect.a034942

---

Downloaded from http://perspectivesinmedicine.cshlp.org/ at Cold Spring Harbor Laboratory Library on May 11, 2018 - Published by Cold Spring Harbor Laboratory Press
M.A. Rubin and F. Demichelis


The Genomics of Prostate Cancer: A Historic Perspective

Mark A. Rubin and Francesca Demichelis

*Cold Spring Harb Perspect Med* published online April 30, 2018

<table>
<thead>
<tr>
<th>Subject Collection</th>
<th>Prostate Cancer</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Anatomy and Histology of the Human and Murine Prostate</strong></td>
<td>The Genomics of Prostate Cancer: A Historic Perspective</td>
</tr>
<tr>
<td>Michael Ittmann</td>
<td>Mark A. Rubin and Francesca Demichelis</td>
</tr>
<tr>
<td><strong>Metastases in Prostate Cancer</strong></td>
<td>Genetically Engineered Mouse Models of Prostate Cancer in the Postgenomic Era</td>
</tr>
<tr>
<td>Federico La Manna, Sofia Karkampouna, Eugenio Zoni, et al.</td>
<td>Juan M. Arriaga and Cory Abate-Shen</td>
</tr>
<tr>
<td><strong>Preclinical and Coclinical Studies in Prostate Cancer</strong></td>
<td>New Opportunities for Targeting the Androgen Receptor in Prostate Cancer</td>
</tr>
<tr>
<td>Ming Chen and Pier Paolo Pandolfi</td>
<td>Margaret M. Centenera, Luke A. Selth, Esmaeil Ebrahimie, et al.</td>
</tr>
<tr>
<td><strong>Transcriptional Regulation in Prostate Cancer</strong></td>
<td>DNA Damage Response in Prostate Cancer</td>
</tr>
<tr>
<td>David P. Labbé and Myles Brown</td>
<td>Matthew J. Schiewer and Karen E. Knudsen</td>
</tr>
<tr>
<td><strong>Anatomic and Molecular Imaging in Prostate Cancer</strong></td>
<td>Molecular Biomarkers in the Clinical Management of Prostate Cancer</td>
</tr>
<tr>
<td>Eric T. Miller, Amirali Salmasi and Robert E. Reiter</td>
<td>Aaron M. Udager and Scott A. Tomlins</td>
</tr>
<tr>
<td><strong>Preclinical Models of Prostate Cancer: Patient-Derived Xenografts, Organoids, and Other Explant Models</strong></td>
<td>Chemoprevention in Prostate Cancer: Current Perspective and Future Directions</td>
</tr>
<tr>
<td>Gail P. Risbridger, Roxanne Toivanen and Renea A. Taylor</td>
<td>J. Ricardo Rivero, Jr., Ian M. Thompson, Jr., Michael A. Liss, et al.</td>
</tr>
<tr>
<td><strong>The Epidemiology of Prostate Cancer</strong></td>
<td>Metabolic Vulnerabilities of Prostate Cancer: Diagnostic and Therapeutic Opportunities</td>
</tr>
<tr>
<td>Claire H. Pernar, Ericka M. Ebot, Kathryn M. Wilson, et al.</td>
<td>Giorgia Zadra and Massimo Loda</td>
</tr>
<tr>
<td><strong>Prostate Cancer Disparities by Race and Ethnicity: From Nucleotide to Neighborhood</strong></td>
<td>Prostate Organogenesis</td>
</tr>
<tr>
<td>Timothy R. Rebeck</td>
<td>Jeffrey C. Francis and Amanda Swain</td>
</tr>
</tbody>
</table>

For additional articles in this collection, see [http://perspectivesinmedicine.cshlp.org/cgi/collection/](http://perspectivesinmedicine.cshlp.org/cgi/collection/)

Copyright © 2018 Cold Spring Harbor Laboratory Press; all rights reserved