Chapter 11
Gene Expression Analysis

Michael Ittmann

Contents

Introduction

Alterations in gene expression in cancer are fundamental to the aberrant biology of cancer cells. Traditionally, genes are thought of as heritable units that encode proteins, and indeed changes in expression of protein-coding genes are critical in the pathogenesis of cancer. A modern definition for a gene has been proposed to be "a union of genomic sequences encoding a coherent set of potentially overlapping functional products" [1]. This definition includes many types of transcripts that do not encode proteins and yet have important cellular functions that determine phenotype. These include transcripts such as miRNAs and long noncoding RNAs (lncRNAs), which are important in the biology of prostate cancer. There is growing evidence that such lncRNAs may be extremely useful for diagnosis and prediction of prognosis and as therapeutic targets in prostate cancer. These types of transcripts can be and have been studied using gene expression arrays [2], but this chapter will focus on protein-coding transcripts. See Chap. 16 for an in-depth discussion of miRNAs and lncRNAs.

M. Ittmann, M.D., Ph.D.
Department of Pathology and Immunology, Baylor College of Medicine, Houston, TX, USA
e-mail: mittmann@bcm.edu

© Springer International Publishing AG 2018

B.D. Robinson et al. (eds.), *Precision Molecular Pathology of Prostate Cancer*,
Molecular Pathology Library 15, https://doi.org/10.1007/978-3-319-64096-9_11

Recent studies have concluded that about 80% of the variation in protein levels can be linked to differences in mRNA levels after correcting for methodological bias [3]. The remainder of protein variation can be accounted for by protein degradation and translational control. Thus, much of the variation in the proteome is controlled by mRNA abundance. In addition to the multiple factors controlling gene transcription in normal cells, mRNA levels in cancer cells can be linked to stable genomic alterations such as gains or losses of genes (copy number variation), the presence of gene fusions, promoter mutations, and alterations in DNA methylation. Such changes at the genome level are fundamental to the pathogenesis of cancer and often achieve phenotypic expression primarily by changes in mRNA levels. See Chaps. 9 and 13 for an in-depth discussion of copy number and epigenetic alterations in prostate cancer.

With the advent of the molecular biology revolution in the 1970s, mRNA levels could be measured, albeit laboriously, with techniques such as Northern blotting. The discovery of polymerase chain reaction (PCR) and later quantitative reverse transcription PCR (Q-RT-PCR) made analysis of gene expression much easier and more quantitative. Indeed, using Q-RT-PCR arrays, hundreds of genes can be assayed in multiplexed formats, and such arrays are currently widely available for focused analysis of specific gene sets. However, high-throughput technologies such as expression microarray analysis have revolutionized our ability to carry out unbiased analysis of gene expression alterations in cancer. In expression microarray analysis, thousands of cDNAs or oligonucleotides corresponding to individual genes are spotted on solid surface such as a glass slide or silicon chip. Fluorescently labeled cDNAs, generated from cellular RNA, are then hybridized to the chip, and after washing to remove nonspecific signals, the chip is scanned. The fluorescent signal is proportional to the abundance of the corresponding RNA used for labeling. The use of cDNAs on arrays has been largely supplanted by the use of oligonucleotides designed to hybridize to specific gene segments. In addition, some formats use beads rather than a solid surface. In any format, expression levels of tens of thousands of genes can be assayed simultaneously, dramatically increasing the ability to interrogate gene expression in a biological sample in a comprehensive manner.

Biological Considerations

Gene expression array technology has been used to study prostate cancer in two major ways. One major type of analysis is studying changes in gene expression in vitro using prostate and prostate cancer cell lines. In this approach, large-scale changes in gene expression can be assessed in different cell lines or more commonly a single cell line under different conditions. This technique can assess the impact of gene knockdown or overexpression, drug treatments, alterations of environmental conditions, etc. on gene expression versus appropriate controls [4–6]. In recent years, such approaches have been used extensively to understand the underlying biology of prostate cancer. In general, biological duplicates or triplicates are needed

for each cell line for each condition to allow more reliable comparisons of gene expression in the two biological states.

The other major use of gene expression arrays has been the study of human and to a lesser extent mouse tumor tissues. The study of human prostate cancer tissues was certainly one of the first major applications of this technology and has yielded many important insights into prostate cancer biology and has identified novel diagnostic and prognostic biomarkers [7–10]. Most published studies have used cancer tissues (and corresponding benign tissues) from radical prostatectomy specimens. Such tissues are potentially readily available given that tens of thousands of radical prostatectomies are performed each year in the United States alone. However, specific precautions are needed to minimize "warm ischemia," i.e., the interval between devascularization and rapid freezing of tissue for later analysis. This is a particularly important consideration in robotic prostatectomies, since warm ischemia can be prolonged in such specimens since they may remain in the patient after devascularization [11]. However, with collaboration between surgeons, pathologists, and other health-care personnel, tissue can be collected in a manner which minimizes warm ischemia and maximizes RNA integrity [11–13]. However, it should be noted that it is impossible to completely eliminate warm ischemia in a surgical specimen. Generally, 15–30 min of warm ischemia is excellent, and up to 60 min is probably acceptable [14]. Of course, tissues with warm ischemia can be used for some studies, particularly DNA analysis, but mRNAs tend to be labile. Thus mRNA integrity can be a surrogate marker of overall tissue quality assuming that RNAs are not degraded during extraction.

One important consideration in the analysis of human prostate cancer specimens is that there is a significant amount of non-cancer cells in such tissues including tumor stroma, benign prostate tissue, and inflammatory cells (both in benign prostate and tumor tissue). It should also be noted that the benign tissues can undergo atrophy and a variety of epithelial metaplasia and hyperplasia. In addition, high-grade prostatic intraepithelial neoplasia (HGPIN) can be present in the non-tumor tissues. In general, it is best to use highly enriched tumor samples for microarray analysis to minimize the contribution from benign tissue components. While there is no official standard, this author believes 50% tumor is a reasonable minimum, although usable data can be obtained with less enriched tissues. However, it should be noted that prostate cancer foci can be difficult to identify by gross examination, and inevitably there will be selection bias for larger and more cellular tumors (often with higher Gleason score) if one uses only more cellular tumor samples. That said, it can be argued that smaller tumors with lower Gleason scores are not of great interest since they rarely result in patient mortality.

It is known that a significant fraction of the alterations in gene expression in prostate cancer versus benign tissue is derived from the cancer stroma [15]. One method of insuring gene expression profiles which are derived from the cancer cells, cancer stroma, benign epithelium, or benign stroma is to carry out laser-capture microdissection [16, 17]. This technique is quite labor intensive but yields more specific results. Generally, amplification of the RNA obtained is required to yield sufficient quantities for gene expression analysis, and controls should always be

carried out to insure that no bias is occurring during amplification [16, 17]. A relatively small number of studies have been carried out using this technique, but they have revealed novel paracrine interactions of tumor stroma with the cancer cells that biologically relevant to tumor progression [18].

If one is comparing cancer tissues to benign tissues from the same prostate, there are also caveats. Since most cancers arise in the peripheral zone, it is presumably better to use this tissue for comparison. Benign prostatic hyperplasia, which arises in the transition zone, is known to induce extensive changes in gene expression and is extremely common in radical prostatectomy specimens [19]. It should also be noted that the same issues of tissue heterogeneity of benign tissues apply to the control benign samples. In particular it is important to exclude high-grade prostatic intraepithelial neoplasia (HGPIN) in the benign samples. However, it would also important to exclude severe acute and/or granulomatous prostatitis, which can be seen in radical prostatectomy specimens. Unfortunately, it can be difficult at times to discern from published reports what are the criteria used, if any, to exclude benign tissues for analysis. Finally, it should be noted that benign adjacent tissues from radical prostatectomies may not be truly "normal" since there is evidence for premalignant field effects in such tissues [20–22]. In practice, age-matched normal prostate tissue is very hard to obtain. Prostate can be harvested from organ donors, but such men are almost always considerably younger than the typical prostate cancer patient. Prostate tissue can also be obtained from radical cystoprostatectomies for treatment of bladder cancer, but one must be cautious about potential gene expression changes induced by bladder cancer treatments, i.e., BCG, etc. Thus, no perfect control tissue is readily available, but carefully characterized benign peripheral zone tissue is probably the best tissue to use in practice.

Metastatic tissues have also been used for gene expression arrays [23, 24]. Such tissues are not commonly obtained in routine clinical practice except for palliative procedures (channel TURPs, relief of spinal cord compression, etc.), but several leading institutions have developed rapid autopsy programs to obtain tissue from lethal prostate cancer. Of course, it should be noted that there is some contamination of normal tissues from various sites as well as tumor stroma, so not all of the RNA is derived from cancer cells in these tissues.

Given that most gene expression arrays have been done on clinically localized disease treated by radical prostatectomy or heavily treated end-stage disease, there is little known about specific gene expression patterns in other clinical prostate cancer states such as treatment-naïve advanced local disease or treatment-naïve metastatic disease to the lymph nodes, bone, or other sites. With emphasis on targeted therapy, biopsy and molecular analysis of such disease states is likely to become more common in the research setting, allowing for further studies of such tissues.

While all of the above caveats may seem to detract from the utility of gene expression analysis of cancer tissues, in reality major advances have arisen from such studies as will be outlined below. However, it is best to be aware of these issues and try and minimize their impact on future studies. In addition, one must keep them in mind when interpreting data from various studies, particularly being aware that not all genes altered in prostate cancer tissues are actually altered in cancer cell themselves.

Expression microarrays can also be used to analyze mouse tissues or human tissues in mice. Analysis of xenografts of human tissues in immunocompromised mice is straightforward and can be used to examine treatment effects, impact of genetic alterations, etc. similar to studies in tissue culture [25–27]. Gene expression arrays can also be used to analyze gene expression in genetically engineered mouse models of prostate cancer [28, 29]. This can give unique insights into the biology of prostate cancer with genetically defined lesions. The mouse prostate consists of distinct lobes [30] (ventral, dorsal, lateral, and anterior) with their own distinct gene expression profiles [31]. Early small lesions can be examined in each lobe (with appropriate controls), while larger cancers usually invade adjacent tissues and cannot be divided into lobes. A number of laboratories have recently used a combined approach in which both human and mouse expression data are mined to define key regulatory pathways in prostate cancer [32–34].

Analytical Considerations

A key element in obtaining accurate gene expression array data is the quality of the input RNA. This is commonly expressed as an RNA integrity number (RIN) or RIN value [35]. The maximum value of RIN is ten, which is quite hard to achieve in practice. Generally, a RIN number of >7 is considered necessary for acceptable quality for RNA from fresh tissues. RNA can be partially degraded due to prolonged ischemia, slow freezing, thawing of tissue samples, and at many points during RNA extraction due to inadequate technique. The use of formalin-fixed paraffin-embedded (FFPE) samples for large-scale mRNA microarray analysis has remained challenging. RNA quality and degradation can be variable, and to date mRNA from frozen tissues remains the gold standard [36]. However, more focused mRNA analysis is certainly possible using specifically designed analytical platforms. DNA analysis in FFPE is more easily performed due to its higher stability but again is more difficult than analysis of DNAs from frozen tissues.

It goes without saying that meticulous attention to detail for reverse transcription, labeling, hybridization, washing, and scanning is critical to obtain accurate results. It is beyond the scope of this review to examine these factors in details, particularly since multiple platforms and approaches are in use, each with distinct technical requirements. That said, several studies have shown generally high concordance rates between different platforms and techniques of labeling [8].

There are two basic approaches to array hybridization. Two-color hybridization uses dyes of two different colors labeled to RNAs representing two conditions, i.e., Cy3 for cancer and Cy5 for matched benign. This gives a direct readout of the relative expression of any probe by the ratio of fluorescence at the two appropriate wavelengths. One-color arrays use only a single dye, and intensity is measured directly and compared to intensity in other arrays directly. The assumption is that labeling and hybridization are relatively similar for all arrays and that any differences between samples in different batches can be accounted for by mathemati-

cal normalization and correction of batch effects. The two methods yield similar results; one-color approach is the dominant mode of analysis and in practice gives robust data [37].

The number of oligonucleotide probes in an array for any given gene is variable. If multiple probes for a single gene show similar alterations of intensity in a given analysis, it adds confidence that the observed change reflects biology and is not an artifact. Of course differences between probes for the same gene may also reflect alternative RNA species arising by alternative splicing, alternative promoters, etc. Alternative splicing plays an important role in prostate cancer biology [38]. For example, changes in alternative splicing of FGF receptors [39] and androgen receptor [40] are well documented in prostate cancer. One can use microarrays to detect alternative splicing using either custom-built arrays or arrays with all known exons [41]. The former is used for more focused analysis since it targets known slice junctions. The latter reveals differences in exon usage but cannot really document the full repertoire of alternatively spliced transcripts. As a hypothetical example, if one detects decreased intensity of exon 2 of a given gene and exon 6 shows increased intensity, it can be hard to determine if these two alterations are occurring in the same or in independent transcripts or both. That said, exon arrays can provide important clues to potentially biologically important alternative RNA species.

Data Analysis

After hybridization, arrays are scanned, and quality control performed. These files are then used to generate gene expression values using programs such as Bioconductor. Data is normalized to adjust the overall chip brightness of the arrays to a similar level [42]. This is needed since differences in labeling efficiency, hybridization, and wash conditions result in differences in signal between arrays. Techniques such as loess normalization, total intensity normalization, quantile normalization, or invariant set normalization can be used. Batch effects also need to be assessed and corrected for in larger experiments.

Of course, data analysis performed depends on the design of the specific experiment or study. Quite often, the goal of the microarray experiment is to define genes that are differentially expressed between two biological states, i.e., cancer and normal, and treated and untreated. Statistical analysis using t-tests can be used to define statistically significant differences between the two sets of data. However, such tests have significant issues in very large datasets. Simplistically, with 60,000 features, using a cutoff of $p < 0.01$ will yield ~600 false positive signals; if 6000 genes are differentially expressed, 10% of these are likely to be false positives. This has led to the use of the false discovery rate or permutation testing to estimate rates of false positivity [43]. While in-depth discussion of statistical approaches to the analysis of microarray data is beyond the scope of this chapter, users need to be aware that such false positive is always an issue. Of course, the actual fold changes are strong indicators of biological significance, i.e., a threefold increase is more likely to be bio-

logically significant than a 10% increase. It should be noted that fold increases tend to be underestimated at the higher end in expression microarrays due to technical factors with array hybridization such as saturation of probe on the array, so that two- or threefold changes in expression are usually highly significant. Outlier analysis has also emerged as important method of identifying differentially expressed genes that are biologically significant [44, 45]. Of course, correlation of gene expression analysis data with copy number analysis, mutations, and other genomic analysis can provide strong support for the importance of a given alteration. For example, if a gene with loss of expression also undergoes frequent homozygous deletion, it is a strong indication that the gene may be a critical tumor suppressor.

In addition to identifying specific genes that are differentially expressed in two datasets, more complex patterns of linked gene expression can be sought. Cluster analysis can be carried out to identify natural groupings of genes that may reflect biological subtypes or other natural groupings. A variety of approaches can be used such as hierarchical clustering, K-means clustering, and principle components analysis. The data is commonly visualized using heat maps [46]. Another useful approach is gene set enrichment analysis to compare a given gene expression signature to a gene set indicative of a specific function, chromosomal location, or regulation [47]. It should be noted that of the thousands of genes that are differentially expressed in prostate cancer and benign tissues, we only understand the biological significance of a fraction of these genes. Thus, we are only beginning to understand the "big data" that has been unleashed over the last 15 years, and novel approaches almost certainly allow new insights into prostate cancer.

An important aspect of gene expression data is that much of it is publicly available. Most published data is deposited in publicly available websites. Of course, analysis of such data requires significant skill. Other sites, such as Oncomine [48] and cBioPortal [49], are usable by general cancer scientists and clinicians and are a very useful avenue for hypothesis testing and generation.

Microarray-Based Discoveries in Prostate Cancer

Prostate Cancer Biology and Pathogenesis

The discovery of recurrent fusion of the androgen-regulated TMPRSS2 gene to the ETS transcription factors, particularly the ERG gene, in the majority of prostate cancer lesions, has led to a paradigm shift in the study of prostate [44]. This discovery will be discussed in greater detail in the next chapter but was based on outlier analysis of gene expression data by the Chinnaiyan group [44]. The TMPRSS2/ERG (T/E) fusion gene occurs in approximately 50% of prostate cancers [50–58]. Experiments in prostate cancer cells containing the T/E fusion [44] indicate that the TMPRSS2 promoter, which contains androgen receptor (AR)-responsive promoter elements [59], increases ERG expression in response to androgens. The ubiquitous activity of AR in prostate cancer cells thus results in high expression of ERG fusion

transcripts. Immunohistochemical studies have shown that ERG overexpression is almost never seen in benign prostate epithelial cells. The high frequency of this genetic alteration argues that it plays a key role in the pathogenesis of those prostate cancers bearing the fusion gene. Indeed, it has been shown that downregulation of the T/E fusion gene by stable shRNA or liposomal siRNA targeting the fusion gene results in markedly reduced tumor growth in vivo [60, 61]. Thus, like the BCR-ABL gene in chronic myelogenous leukemia, it is an attractive therapeutic target as well as a potential diagnostic marker.

Another major finding growing out of gene expression microarray studies is the finding that serine protease inhibitor Kazal type 1 (SPINK1) is overexpressed in approximately 5–15% of prostate cancers [45, 62–64]. SPINK1 overexpression is essentially mutually exclusive with ERG overexpression and thus constitutes a distinct subtype of prostate cancer. Many but not all studies have found SPINK1 overexpression to be associated with adverse outcome [45, 62, 65]. SPINK1 has been shown potentiate EGFR signaling and as an extracellular protein is potentially therapeutically targetable [63].

Diagnostic Biomarkers

Surgical pathologists frequently face the problem of trying to determine if a small cluster of glands in a needle biopsy is malignant. A very early discovery using expression microarrays was that alpha methyl-acyl CoA racemase (AMACR) [9, 66, 67] is markedly increased in prostate cancer. This has led to the widespread use of immunohistochemistry to detect this marker as an adjunct to diagnosis of difficult lesions [64, 65]. Typically, this is combined with basal cell markers such as p63, CK5/6, and/or high molecular weight cytokeratin (34β[beta]E12), which are absent in cancer lesions. While these are caveats about this approach, using a combination of a positive marker such as AMACR along with negative basal cell-specific markers has proven to be a powerful diagnostic tool [68]. However, AMACR is not prostate cancer specific and can be expressed in benign prostatic glands and benign mimics of prostate cancer and high-grade PIN [68, 69]. Similarly, basal cells may be absent in some benign mimics of prostate cancer, particularly in a small clusters of glands. ERG is more specific to prostate cancer and thus is also useful as an immunohistochemical adjunct to diagnosis [70, 71]. Its utility is somewhat limited by the fact that it is present in only ~50% of prostate cancers in European-Americans [72] and ~20% of prostate cancers in African-Americans [73, 74]. See Chap. 27 for further discussion.

Detection Biomarkers

Given the limitations of prostate-specific antigen (PSA) testing, it is clear that better biomarkers for the early detection are needed of clinically significant prostate cancer. A number of protein-coding genes initially detected as elevated in prostate cancer by gene expression arrays have been included in potential multiplexed panels to detect

prostate cancer using urine including T/E and SPINK1 (see above) as well as GOLPH2 [75] and GOLM1 [76]. Recently increasing interest has focused on non-coding RNAs including lncRNAs and microRNAs. Novel blood-/serum-based tests are also under active development, although none are as yet in widespread clinical use. See Chap. 27 for additional discussion of this rapidly evolving area.

Prognostic Biomarkers

Given the highly heterogeneous clinical behavior of prostate cancer, it is critical to define those men requiring treatment and, hopefully, the optimal treatment needed. A number of genes identified by gene expression analysis have been proposed as potential biomarkers of disease aggressiveness such as EZH2 [23, 77] and SPINK1 [45]. One can also seek to define patterns of gene expression that are predictive of clinical outcome and response to therapy, etc. A number of such signatures have been proposed [10, 78–88] including signatures based on stromal markers [89]. One signature based on gene expression data is now commercially available.

(Oncotype Dx for prostate cancer). Whether such signatures will achieve widespread usage is not yet clear. In-depth discussion of such approaches is beyond the scope of this chapter but is discussed in detail in Chap. 29.

Prostate Cancer Classification

Expression array analysis has proven to be extremely powerful in defining biological and clinically linked subclasses of cancer in a number of malignancies, notably in lymphoma [90] and breast cancer [91]. Expression microarray analysis suggests a classification of localized prostate cancer that certainly has biological significance and perhaps clinical significance as well. As described above, approximately 50% of localized prostate cancers have ERG overexpression due to the presence of the T/E fusion gene. Another 10–15% of prostate cancers have overexpression of ERG via alternative fusions or overexpression of other ETS factors. SPINK1 is overexpressed by approximately 10% of prostate cancers. A final category is defined by mutation of the SPOP gene (see Chap. 12 for details). Of note, these categories are almost mutually exclusive and account for approximately 80% of all prostate cancers. See Chap. 30 for further discussion.

RNA-Seq and Future of Expression Microarrays

Gene expression arrays have had a tremendous impact on our understanding of the pathobiology of prostate cancer over the last 15 years. However, the use of gene expression microarrays is now being supplemented by the use of RNA-Seq, i.e., the

use of next-generation sequencing to sequence RNAs. Gene expression measured by this technology is highly concordant with measurements using one-channel microarrays [92]. RNA-Seq also has significant advantages over gene expression arrays. First, it can discover novel genes and is not limited by what is arrayed on the chip. Second, it has an almost limitless linear measurement range, while gene expression microarrays have limited sensitivity for low-expression genes and show signal saturation for highly expressed genes. Third, exon usage and alternative splicing can be more accurately evaluated. Finally, fusion genes, point mutations, and small deletions can be identified. While RNA-Seq is still more expensive than gene arrays, its cost is decreasing and is approaching the cost of gene arrays. This higher cost and more complex data analysis is still limiting penetration of RNA-Seq into prostate cancer analysis of gene expression, but it is clear that RNA-Seq is likely to substantially replace gene arrays in the future.

References

1. Gerstein MB, Bruce C, Rozowsky JS, Zheng D, Du J, Korbel JO, et al. What is a gene, post-ENCODE? History and updated definition. [Historical Article Research Support, N.I.H., Extramural Review]. Genome Res. 2007;17(6):669–81.
2. Ozen M, Creighton CJ, Ozdemir M, Ittmann M. Widespread deregulation of microRNA expression in human prostate cancer. Oncogene. 2008;27(12):1788–93.
3. Li JJ, Biggin MD. Gene expression. Statistics requantitates the central dogma. [Comment Research Support, Non-U.S. Gov't Research Support, U.S. Gov't, Non-P.H.S.]. Science. 2015;347(6226):1066–7.
4. Yu W, Feng S, Dakhova O, Creighton CJ, Cai Y, Wang J, et al. FGFR-4 Arg(3)(8)(8) enhances prostate cancer progression via extracellular signal-related kinase and serum response factor signaling. Clin Cancer Res. 2011;17(13):4355–66.
5. Yang F, Chen Y, Shen T, Guo D, Dakhova O, Ittmann MM, et al. Stromal TGF-beta signaling induces AR activation in prostate cancer. Oncotarget. 2014;5(21):10854–69.
6. Yan J, Yu CT, Ozen M, Ittmann M, Tsai SY, Tsai MJ. Steroid receptor coactivator-3 and activator protein-1 coordinately regulate the transcription of components of the insulin-like growth factor/AKT signaling pathway. Cancer Res. [Research Support, N.I.H., Extramural.]. 2006;66(22):11039–46.
7. Chaib H, Cockrell EK, Rubin MA, Macoska JA. Profiling and verification of gene expression patterns in normal and malignant human prostate tissues by cDNA microarray analysis. Neoplasia. [Comparative Study Research Support, U.S. Gov't, P.H.S.]. 2001;3(1):43–52.
8. Rhodes DR, Barrette TR, Rubin MA, Ghosh D, Chinnaiyan AM. Meta-analysis of microarrays: interstudy validation of gene expression profiles reveals pathway dysregulation in prostate cancer. Cancer Res. 2002;62(15):4427–33.
9. Rubin MA, Zhou M, Dhanasekaran SM, Varambally S, Barrette TR, Sanda MG, et al. Alpha-Methylacyl coenzyme A racemase as a tissue biomarker for prostate cancer. JAMA. [Research Support, Non-US Gov't Research Support, US Gov't, PHS]. 2002;287(13):1662–70.
10. Lapointe J, Li C, Higgins JP, van de Rijn M, Bair E, Montgomery K, et al. Gene expression profiling identifies clinically relevant subtypes of prostate cancer. Proc Natl Acad Sci U S A. 2004;101(3):811–6.
11. Voss BL, Santiano K, Milano M, Mangold KA, Kaul KL. Integrity and amplification of nucleic acids from snap-frozen prostate tissues from robotic-assisted laparoscopic and open prostatectomies. Arch Pathol Lab Med. [Research Support, N.I.H., Extramural Research Support, Non-U.S. Gov't]. 2013;137(4):525–30.

12. Esgueva R, Park K, Kim R, Kitabayashi N, Barbieri CE, Dorsey PJ Jr, et al. Next-generation prostate cancer biobanking: toward a processing protocol amenable for the International Cancer Genome Consortium. Diagn Mol Pathol. [Research Support, Non-U.S. Gov't Research Support, U.S. Gov't, Non-P.H.S.]. 2012;21(2):61–8.
13. Best S, Sawers Y, Fu VX, Almassi N, Huang W, Jarrard DF. Integrity of prostatic tissue for molecular analysis after robotic-assisted laparoscopic and open prostatectomy. Urology. [Research Support, N.I.H., Extramural Research Support, Non-U.S. Gov't.]. 2007;70(2):328–32.
14. Dash A, Maine IP, Varambally S, Shen R, Chinnaiyan AM, Rubin MA. Changes in differential gene expression because of warm ischemia time of radical prostatectomy specimens. Am J Pathol. [Research Support, U.S. Gov't, P.H.S.]. 2002;161(5):1743–8.
15. Stuart RO, Wachsman W, Berry CC, Wang-Rodriguez J, Wasserman L, Klacansky I, et al. In silico dissection of cell-type-associated patterns of gene expression in prostate cancer. Proc Natl Acad Sci U S A. [Research Support, U.S. Gov't, PHS]. 2004;101(2):615–20.
16. Dakhova O, Ozen M, Creighton CJ, Li R, Ayala G, Rowley D, et al. Global gene expression analysis of reactive stroma in prostate cancer. Clin Cancer Res. 2009;15(12):3979–89.
17. Tomlins SA, Mehra R, Rhodes DR, Cao X, Wang L, Dhanasekaran SM, et al. Integrative molecular concept modeling of prostate cancer progression. Nat Genet. [Research Support, N.I.H., Extramural Research Support, Non-U.S. Gov't Research Support, U.S. Gov't, Non--P.H.S.]. 2007;39(1):41–51.
18. Dakhova O, Rowley D, Ittmann M. Genes upregulated in prostate cancer reactive stroma promote prostate cancer progression in vivo. Clin Cancer Res. [Research Support, N.I.H., Extramural Research Support, Non-U.S. Gov't]. 2014;20(1):100–9.
19. Luo J, Dunn T, Ewing C, Sauvageot J, Chen Y, Trent J, et al. Gene expression signature of benign prostatic hyperplasia revealed by cDNA microarray analysis. Prostate. [Research Support, Non-U.S. Gov't Research Support, U.S. Gov't, P.H.S.]. 2002;51(3):189–200.
20. Kosari F, Cheville JC, Ida CM, Karnes RJ, Leontovich AA, Sebo TJ, et al. Shared gene expression alterations in prostate cancer and histologically benign prostate from patients with prostate cancer. Am J Pathol. [Research Support, Non-U.S. Gov't]. 2012;181(1):34–42.
21. Risk MC, Knudsen BS, Coleman I, Dumpit RF, Kristal AR, LeMeur N, et al. Differential gene expression in benign prostate epithelium of men with and without prostate cancer: evidence for a prostate cancer field effect. Clin Cancer Res. [Research Support, N.I.H., Extramural]. 2010;16(22):5414–23.
22. Chandran UR, Dhir R, Ma C, Michalopoulos G, Becich M, Gilbertson J. Differences in gene expression in prostate cancer, normal appearing prostate tissue adjacent to cancer and prostate tissue from cancer free organ donors. BMC Cancer. [Research Support, N.I.H., Extramural Research Support, U.S. Gov't, Non-P.H.S. Research Support, U.S. Gov't, P.H.S.]. 2005;5:45.
23. Varambally S, Dhanasekaran SM, Zhou M, Barrette TR, Kumar-Sinha C, Sanda MG, et al. The polycomb group protein EZH2 is involved in progression of prostate cancer. Nature. 2002;419(6907):624–9.
24. Taylor BS, Schultz N, Hieronymus H, Gopalan A, Xiao Y, Carver BS, et al. Integrative genomic profiling of human prostate cancer. Cancer Cell. [Research Support, N.I.H., Extramural Research Support, Non-U.S. Gov't.]. 2010;18(1):11–22.
25. Glinsky GV, Krones-Herzig A, Glinskii AB, Gebauer G. Microarray analysis of xenograft-derived cancer cell lines representing multiple experimental models of human prostate cancer. Mol Carcinog. [Research Support, Non-U.S. Gov't Research Support, U.S. Gov't, P.H.S.]. 2003;37(4):209–21.
26. Sirotnak FM, She Y, Khokhar NZ, Hayes P, Gerald W, Scher HI. Microarray analysis of prostate cancer progression to reduced androgen dependence: studies in unique models contrasts early and late molecular events. Mol Carcinog. [Comparative Study Research Support, Non--U.S. Gov't Research Support, U.S. Gov't, P.H.S.]. 2004;41(3):150–63.
27. Narita S, Tsuchiya N, Saito M, Inoue T, Kumazawa T, Yuasa T, et al. Candidate genes involved in enhanced growth of human prostate cancer under high fat feeding identified by microarray analysis. Prostate. [Research Support, Non-U.S. Gov't]. 2008;68(3):321–35.

28. Acevedo VD, Gangula RD, Freeman KW, Li R, Zhang Y, Wang F, et al. Inducible FGFR-1 activation leads to irreversible prostate adenocarcinoma and an epithelial-to-mesenchymal transition. Cancer Cell. 2007;12(6):559–71.
29. Carstens JL, Shahi P, Van Tsang S, Smith B, Creighton CJ, Zhang Y, et al. FGFR1-WNT-TGF-beta signaling in prostate cancer mouse models recapitulates human reactive stroma. Cancer Res. [Research Support, N.I.H., Extramural.]. 2014;74(2):609–20.
30. Ittmann M, Huang J, Radaelli E, Martin P, Signoretti S, Sullivan R, et al. Animal models of human prostate cancer: the consensus report of the New York meeting of the Mouse Models of Human Cancers Consortium Prostate Pathology Committee. Cancer Res. [Research Support, N.I.H., Extramural Research Support, Non-U.S. Gov't Research Support, U.S. Gov't, Non--P.H.S.]. 2013;73(9):2718–36.
31. Berquin IM, Min Y, Wu R, Wu H, Chen YQ. Expression signature of the mouse prostate. J Biol Chem. [Research Support, N.I.H., Extramural Research Support, U.S. Gov't, P.H.S.]. 2005;280(43):36442–51.
32. Aytes A, Mitrofanova A, Lefebvre C, Alvarez MJ, Castillo-Martin M, Zheng T, et al. Cross-species regulatory network analysis identifies a synergistic interaction between FOXM1 and CENPF that drives prostate cancer malignancy. Cancer Cell. [Research Support, N.I.H., Extramural Research Support, Non-U.S. Gov't.]. 2014;25(5):638–51.
33. Irshad S, Bansal M, Castillo-Martin M, Zheng T, Aytes A, Wenske S, et al. A molecular signature predictive of indolent prostate cancer. Sci Transl Med. [Research Support, N.I.H., Extramural Research Support, Non-U.S. Gov't]. 2013;5(202):202ra122.
34. Thompson VC, Day TK, Bianco-Miotto T, Selth LA, Han G, Thomas M, et al. A gene signature identified using a mouse model of androgen receptor-dependent prostate cancer predicts biochemical relapse in human disease. Int J Cancer. [Research Support, N.I.H., Extramural.]. 2012;131(3):662–72.
35. Imbeaud S, Graudens E, Boulanger V, Barlet X, Zaborski P, Eveno E, et al. Towards standardization of RNA quality assessment using user-independent classifiers of microcapillary electrophoresis traces. Nucleic Acids Res. [Evaluation Studies Research Support, Non-U.S. Gov't.]. 2005;33(6):e56.
36. Greytak SR, Engel KB, Bass BP, Moore HM. Accuracy of molecular data generated with FFPE biospecimens: lessons from the literature. Cancer Res. [Review]. 2015;75(8):1541–7.
37. Oberthuer A, Juraeva D, Li L, Kahlert Y, Westermann F, Eils R, et al. Comparison of performance of one-color and two-color gene-expression analyses in predicting clinical endpoints of neuroblastoma patients. Pharmacogenomics J. [Comparative Study Research Support, Non--U.S. Gov't.]. 2010;10(4):258–66.
38. Lapuk AV, Volik SV, Wang Y, Collins CC. The role of mRNA splicing in prostate cancer. Asian J Androl. [Research Support, Non-U.S. Gov't]. 2014;16(4):515–21.
39. Kwabi-Addo B, Ropiquet F, Giri D, Ittmann M. Alternative splicing of fibroblast growth factor receptors in human prostate cancer. Prostate. 2001;46(2):163–72.
40. Dehm SM, Schmidt LJ, Heemers HV, Vessella RL, Tindall DJ. Splicing of a novel androgen receptor exon generates a constitutively active androgen receptor that mediates prostate cancer therapy resistance. Cancer Res. 2008;68(13):5469–77.
41. Cuperlovic-Culf M, Belacel N, Culf AS, Ouellette RJ. Microarray analysis of alternative splicing. OMICS. [Research Support, Non-U.S. Gov't Review]. 2006;10(3):344–57.
42. Steinhoff C, Vingron M. Normalization and quantification of differential expression in gene expression microarrays. Brief Bioinform. [Research Support, Non-U.S. Gov't Review]. 2006;7(2):166–77.
43. Gusnanto A, Calza S, Pawitan Y. Identification of differentially expressed genes and false discovery rate in microarray studies. Curr Opin Lipidol. [Review]. 2007;18(2):187–93.
44. Tomlins SA, Rhodes DR, Perner S, Dhanasekaran SM, Mehra R, Sun XW, et al. Recurrent fusion of TMPRSS2 and ETS transcription factor genes in prostate cancer. Science. [Research Support, N.I.H., Extramural Research Support, Non-U.S. Gov't Research Support, U.S. Gov't, Non-P.H.S. Research Support, U.S. Gov't, P.H.S.]. 2005;310(5748):644–8.
45. Tomlins SA, Rhodes DR, Yu J, Varambally S, Mehra R, Perner S, et al. The role of SPINK1 in ETS rearrangement-negative prostate cancers. Cancer Cell. [Meta-Analysis Multicenter Study

Research Support, N.I.H., Extramural Research Support, Non-U.S. Gov't Research Support, U.S. Gov't, Non-PHS]. 2008;13(6):519–28.
46. Saldanha AJ. Java Treeview – extensible visualization of microarray data. Bioinformatics. 2004;20(17):3246–8.
47. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci U S A. 2005;102(43):15545–50.
48. Rhodes DR, Kalyana-Sundaram S, Mahavisno V, Varambally R, Yu J, Briggs BB, et al. Oncomine 3.0: genes, pathways, and networks in a collection of 18,000 cancer gene expression profiles. Neoplasia. 2007;9(2):166–80.
49. Cerami E, Gao J, Dogrusoz U, Gross BE, Sumer SO, Aksoy BA, et al. The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data. Cancer Discov. [Research Support, N.I.H., Extramural Research Support, Non-U.S. Gov't.]. 2012;2(5):401–4.
50. Wang J, Cai Y, Ren C, Ittmann M. Expression of variant TMPRSS2/ERG fusion messenger RNAs is associated with aggressive prostate cancer. Cancer Res. 2006;66(17):8347–51.
51. Clark J, Merson S, Jhavar S, Flohr P, Edwards S, Foster CS, et al. Diversity of TMPRSS2-ERG fusion transcripts in the human prostate. Oncogene. [Research Support, Non-U.S. Gov't.]. 2007;26(18):2667–73.
52. Yoshimoto M, Joshua AM, Chilton-Macneill S, Bayani J, Selvarajah S, Evans AJ, et al. Three-color FISH analysis of TMPRSS2/ERG fusions in prostate cancer indicates that genomic microdeletion of chromosome 21 is associated with rearrangement. Neoplasia. [Research Support, Non-U.S. Gov't.]. 2006;8(6):465–9.
53. Rajput AB, Miller MA, De Luca A, Boyd N, Leung S, Hurtado-Coll A, et al. Frequency of the TMPRSS2:ERG gene fusion is increased in moderate to poorly differentiated prostate cancers. J Clin Pathol. 2007;60(11):1238–43.
54. Mehra R, Tomlins SA, Shen R, Nadeem O, Wang L, Wei JT, et al. Comprehensive assessment of TMPRSS2 and ETS family gene aberrations in clinically localized prostate cancer. Mod Pathol. 2007;20(5):538–44.
55. Nam RK, Sugar L, Wang Z, Yang W, Kitching R, Klotz LH, et al. Expression of TMPRSS2:ERG gene fusion in prostate cancer cells is an important prognostic factor for cancer progression. Cancer Biol Ther. 2007;6(1):40–5.
56. Demichelis F, Fall K, Perner S, Andren O, Schmidt F, Setlur SR, et al. TMPRSS2:ERG gene fusion associated with lethal prostate cancer in a watchful waiting cohort. Oncogene. [Research Support, N.I.H., Extramural Research Support, Non-U.S. Gov't.]. 2007;26(31):4596–9.
57. Soller MJ, Isaksson M, Elfving P, Soller W, Lundgren R, Panagopoulos I. Confirmation of the high frequency of the TMPRSS2/ERG fusion gene in prostate cancer. Genes Chromosomes Cancer. 2006;45(7):717–9.
58. Attard G, Clark J, Ambroisine L, Fisher G, Kovacs G, Flohr P, et al. Duplication of the fusion of TMPRSS2 to ERG sequences identifies fatal human prostate cancer. Oncogene. [Research Support, Non-U.S. Gov't.]. 2008;27(3):253–63.
59. Lin B, Ferguson C, White JT, Wang S, Vessella R, True LD, et al. Prostate-localized and androgen-regulated expression of the membrane-bound serine protease TMPRSS2. Cancer Res. 1999;59(17):4180–4.
60. Wang J, Cai Y, Yu W, Ren C, Spencer DM, Ittmann M. Pleiotropic biological activities of alternatively spliced TMPRSS2/ERG fusion gene transcripts. Cancer Res. [Research Support, N.I.H., Extramural Research Support, U.S. Gov't, Non-P.H.S.]. 2008;68(20):8516–24.
61. Shao L, Tekedereli I, Wang J, Yuca E, Tsang S, Sood A, et al. Highly specific targeting of the TMPRSS2/ERG fusion gene using liposomal nanovectors. Clin Cancer Res. [Research Support, Non-U.S. Gov't Research Support, U.S. Gov't, Non-P.H.S.]. 2012;18(24):6648–57.
62. Leinonen KA, Tolonen TT, Bracken H, Stenman UH, Tammela TL, Saramaki OR, et al. Association of SPINK1 expression and TMPRSS2:ERG fusion with prognosis in endocrine-treated prostate cancer. Clin Cancer Res. 2010;16(10):2845–51.

63. Ateeq B, Tomlins SA, Laxman B, Asangani IA, Cao Q, Cao X, et al. Therapeutic targeting of SPINK1-positive prostate cancer. Sci Transl Med. 2011;3(72):72ra17.
64. Bismar TA, Yoshimoto M, Duan Q, Liu S, Sircar K, Squire JA. Interactions and relationships of PTEN, ERG, SPINK1 and AR in castration-resistant prostate cancer. Histopathology. 2012;60(4):645–52.
65. Grupp K, Diebel F, Sirma H, Simon R, Breitmeyer K, Steurer S, et al. SPINK1 expression is tightly linked to 6q15- and 5q21-deleted ERG-fusion negative prostate cancers but unrelated to PSA recurrence. Prostate. 2013;73(15):1690–8.
66. Jiang Z, Woda BA. Diagnostic utility of alpha-methylacyl CoA racemase (P504S) on prostate needle biopsy. Adv Anat Pathol. [Review]. 2004;11(6):316–21.
67. Jiang Z, Woda BA, Wu CL, Yang XJ. Discovery and clinical application of a novel prostate cancer marker: alpha-methylacyl CoA racemase (P504S). Am J Clin Pathol. [Review]. 2004;122(2):275–89.
68. Kumaresan K, Kakkar N, Verma A, Mandal AK, Singh SK, Joshi K. Diagnostic utility of alpha-methylacyl CoA racemase (P504S) & HMWCK in morphologically difficult prostate cancer. Diagn Pathol. [Evaluation Studies]. 2010;5:83.
69. Worschech A, Meirelles L, Billis A. Expression of alpha-methylacyl coenzyme A racemase in partial and complete focal atrophy on prostate needle biopsies. Anal Quant Cytol Histol. [Research Support, Non-U.S. Gov't]. 2009;31(6):424–31.
70. Shah RB, Tadros Y, Brummell B, Zhou M. The diagnostic use of ERG in resolving an "atypical glands suspicious for cancer" diagnosis in prostate biopsies beyond that provided by basal cell and alpha-methylacyl-CoA-racemase markers. Hum Pathol. 2013;44(5):786–94.
71. Dabir PD, Ottosen P, Hoyer S, Hamilton-Dutoit S. Comparative analysis of three- and two-antibody cocktails to AMACR and basal cell markers for the immunohistochemical diagnosis of prostate carcinoma. Diagn Pathol. [Comparative Study Research Support, Non-U.S. Gov't.]. 2012;7:81.
72. Tomlins SA, Palanisamy N, Siddiqui J, Chinnaiyan AM, Kunju LP. Antibody-based detection of ERG rearrangements in prostate core biopsies, including diagnostically challenging cases: ERG staining in prostate core biopsies. Arch Pathol Lab Med. [Evaluation Studies Research Support, N.I.H., Extramural.]. 2012;136(8):935–46.
73. Magi-Galluzzi C, Tsusuki T, Elson P, Simmerman K, LaFargue C, Esgueva R, et al. TMPRSS2-ERG gene fusion prevalence and class are significantly different in prostate cancer of Caucasian, African-American and Japanese patients. Prostate. 2011;71(5):489–97.
74. Rosen P, Pfister D, Young D, Petrovics G, Chen Y, Cullen J, et al. Differences in frequency of ERG oncoprotein expression between index tumors of Caucasian and African American patients with prostate cancer. Urology. [Research Support, N.I.H., Extramural]. 2012;80(4):749–53.
75. Laxman B, Morris DS, Yu J, Siddiqui J, Cao J, Mehra R, et al. A first-generation multiplex biomarker analysis of urine for the early detection of prostate cancer. Cancer Res. 2008;68(3):645–9.
76. Varambally S, Laxman B, Mehra R, Cao Q, Dhanasekaran SM, Tomlins SA, et al. Golgi protein GOLM1 is a tissue and urine biomarker of prostate cancer. Neoplasia. [Meta-Analysis Research Support, N.I.H., Extramural Research Support, U.S. Gov't, Non-P.H.S.]. 2008;10(11):1285–94.
77. van Leenders GJ, Dukers D, Hessels D, van den Kieboom SW, Hulsbergen CA, Witjes JA, et al. Polycomb-group oncogenes EZH2, BMI1, and RING1 are overexpressed in prostate cancer with adverse pathologic and clinical features. Eur Urol. 2007;52(2):455–63.
78. Mucci LA, Pawitan Y, Demichelis F, Fall K, Stark JR, Adami HO, et al. Testing a multigene signature of prostate cancer death in the Swedish Watchful Waiting Cohort. Cancer Epidemiol Biomark Prev. 2008;17(7):1682–8.
79. Mucci LA, Pawitan Y, Demichelis F, Fall K, Stark JR, Adami HO, et al. Nine-gene molecular signature is not associated with prostate cancer death in a watchful waiting cohort. Cancer Epidemiol Biomarkers Prev. [Research Support, N.I.H., Extramural Research Support, Non--U.S. Gov't.]. 2008;17(1):249–51.

80. Klein EA, Cooperberg MR, Magi-Galluzzi C, Simko JP, Falzarano SM, Maddala T, et al. A 17-gene assay to predict prostate cancer aggressiveness in the context of Gleason grade heterogeneity, tumor multifocality, and biopsy undersampling. Eur Urol. [Research Support, Non--U.S. Gov't Research Support, U.S. Gov't, Non-P.H.S.]. 2014;66(3):550–60.
81. Erho N, Crisan A, Vergara IA, Mitra AP, Ghadessi M, Buerki C, et al. Discovery and validation of a prostate cancer genomic classifier that predicts early metastasis following radical prostatectomy. PLoS One. [Research Support, N.I.H., Extramural Research Support, Non--U.S. Gov't.]. 2013;8(6):e66855.
82. Sun Y, Goodison S. Optimizing molecular signatures for predicting prostate cancer recurrence. Prostate. [Comparative Study Research Support, N.I.H., Extramural Research Support, Non--U.S. Gov't.]. 2009;69(10):1119–27.
83. Glinsky GV, Glinskii AB, Stephenson AJ, Hoffman RM, Gerald WL. Gene expression profiling predicts clinical outcome of prostate cancer. J Clin Invest. 2004;113(6):913–23.
84. Yu YP, Landsittel D, Jing L, Nelson J, Ren B, Liu L, et al. Gene expression alterations in prostate cancer predicting tumor aggression and preceding development of malignancy. J Clin Oncol. [Research Support, U.S. Gov't, P.H.S.]. 2004;22(14):2790–9.
85. Singh D, Febbo PG, Ross K, Jackson DG, Manola J, Ladd C, et al. Gene expression correlates of clinical prostate cancer behavior. Cancer Cell. [Research Support, Non-U.S. Gov't Research Support, U.S. Gov't, P.H.S.]. 2002;1(2):203–9.
86. Cheville JC, Karnes RJ, Therneau TM, Kosari F, Munz JM, Tillmans L, et al. Gene panel model predictive of outcome in men at high-risk of systemic progression and death from prostate cancer after radical retropubic prostatectomy. J Clin Oncol. 2008;26(24):3930–6.
87. Chen X, Xu S, McClelland M, Rahmatpanah F, Sawyers A, Jia Z, et al. An accurate prostate cancer prognosticator using a seven-gene signature plus Gleason score and taking cell type heterogeneity into account. PLoS One. [Research Support, N.I.H., Extramural Research Support, U.S. Gov't, Non-P.H.S.]. 2012;7(9):e45178.
88. Dhanasekaran SM, Barrette TR, Ghosh D, Shah R, Varambally S, Kurachi K, et al. Delineation of prognostic biomarkers in prostate cancer. Nature. [Research Support, Non-U.S. Gov't.]. 2001;412(6849):822–6.
89. Jia Z, Rahmatpanah FB, Chen X, Lernhardt W, Wang Y, Xia XQ, et al. Expression changes in the stroma of prostate cancer predict subsequent relapse. PLoS One. [Clinical Trial Comparative Study Research Support, N.I.H., Extramural Research Support, Non-U.S. Gov't Research Support, U.S. Gov't, Non-P.H.S.]. 2012;7(8):e41371.
90. Alizadeh AA, Eisen MB, Davis RE, Ma C, Lossos IS, Rosenwald A, et al. Distinct types of diffuse large B-cell lymphoma identified by gene expression profiling. Nature. 2000;403(6769):503–11.
91. Sorlie T. Molecular portraits of breast cancer: tumour subtypes as distinct disease entities. Eur J Cancer. 2004;40(18):2667–75.
92. Guo Y, Sheng Q, Li J, Ye F, Samuels DC, Shyr Y. Large scale comparison of gene expression levels by microarrays and RNAseq using TCGA data. PLoS One. [Comparative Study]. 2013;8(8):e71462.