LNCaP Atlas: Gene expression associated with \textit{in vivo} progression to castration-recurrent prostate cancer

Tammy L Romanuik, Gang Wang, Olena Morozova, Allen Delaney, Marco A Marra, Marianne D Sadar

Abstract

\textbf{Background:} There is no cure for castration-recurrent prostate cancer (CRPC) and the mechanisms underlying this stage of the disease are unknown.

\textbf{Methods:} We analyzed the transcriptome of human LNCaP prostate cancer cells as they progress to CRPC \textit{in vivo} using replicate LongSAGE libraries. We refer to these libraries as the LNCaP atlas and compared these gene expression profiles with current suggested models of CRPC.

\textbf{Results:} Three million tags were sequenced using \textit{in vivo} samples at various stages of hormonal progression to reveal 96 novel genes differentially expressed in CRPC. Thirty-one genes encode proteins that are either secreted or are located at the plasma membrane, 21 genes changed levels of expression in response to androgen, and 8 genes have enriched expression in the prostate. Expression of 26, 6, 12, and 15 genes have previously been linked to prostate cancer, Gleason grade, progression, and metastasis, respectively. Expression profiles of genes in CRPC support a role for the transcriptional activity of the androgen receptor (CCNH, CUEDC2, FLNA, PSMA7), steroid synthesis and metabolism (DHCR24, DHRS7, ELOVL5, HSD17B4, OPRK1), neuroendocrine (ENO2, MAOA, OPRK1, S100A10, TRPM8), and proliferation (GAS5, GNB2L1, MT-ND3, NKX3-1, PCGEM1, PTGFR, STEAP1, TMEM30A), but neither supported nor discounted a role for cell survival genes.

\textbf{Conclusions:} The \textit{in vivo} gene expression atlas for LNCaP was sequenced and support a role for the androgen receptor in CRPC.

Background

Systemic androgen-deprivation therapy by orchietomy or agonists of gonadotropic releasing hormone are routinely used to treat men with metastatic prostate cancer to reduce tumor burden and pain. This therapy is based on the dependency of prostate cells for androgens to grow and survive. The inability of androgen-deprivation therapy to completely and effectively eliminate all metastatic prostate cancer cell populations is manifested by a predictable and inevitable relapse, referred to as castration-recurrent prostate cancer (CRPC). CRPC is the end stage of the disease and fatal to the patient within 16-18 months of onset.

The mechanisms underlying progression to CRPC are unknown. However, there are several models to explain its development. One such model indicates the involvement of the androgen signaling pathway [1-4]. Key to this pathway is the androgen receptor (AR) which is a steroid hormone receptor and transcription factor. Mechanisms of progression to CRPC that involve or utilize the androgen signaling pathway include: hypersensitivity due to \textit{AR} gene amplification [5,6]; changes in AR co-regulators such as nuclear receptor coactivators (NCOA1 and NCOA2) [7,8]; intraprostatic \textit{de novo} synthesis of androgen [9] or metabolism of AR ligands from residual adrenal androgens [10,11]; AR promiscuity of ligand specificity due to mutations [12]; and ligand-independent activation of AR by growth factors [protein kinase A (PKA), interleukin 6 (IL6), and epidermal growth factor (EGF)] [13-15]. Activation of the AR can be determined by assaying for the expression of target...
genes such as prostate-specific antigen (PSA) [16]. Other models of CRPC include the neuroendocrine differentiation [17], the stem cell model [18] and the imbalance between cell growth and cell death [3]. It is conceivable that these models may not mutual exclusive. For example altered AR activity may impact cell survival and proliferation.

Here, we describe long serial analysis of gene expression (LongSAGE) libraries [19,20] made from RNA sampled from biological replicates of the in vivo LNCaP Hollow Fiber model of prostate cancer as it progresses to the castration-recurrent stage. Gene expression signatures that were consistent among the replicate libraries were applied to the current models of CRPC.

Methods

In vivo LNCaP Hollow Fiber model

The LNCaP Hollow Fiber model of prostate cancer was performed as described previously [21-23]. All animal experiments were performed according to a protocol approved by the Committee on Animal Care of the University of British Columbia. Serum PSA levels were determined by enzymatic immunoassay kit (Abbott Laboratories, Abbott Park, IL, USA). Fibers were removed on three separate occasions representing different stages of hormonal progression that were androgen-sensitive (AS), responsive to androgen-deprivation (RAD), and castration-recurrent (CR). Samples were retrieved immediately prior to castration (AS), as well as 10 (RAD) and 72 days (CR) post-surgical castration.

RNA sample generation, processing, and quality control

Total RNA was isolated immediately from cells harvested from the in vivo Hollow Fiber model using TRIZOL Reagent (Invitrogen) following the manufacturer’s instructions. Genomic DNA was removed from RNA samples with DNasel (Invitrogen). RNA quality and quantity were assessed by the Agilent 2100 Bioanalyzer (Agilent Technologies, Mississauga, ON, Canada) and RNA 6000 Nano LabChip kit (Caliper Technologies, Hopkinton, MA, USA).

Quantitative real-time polymerase chain reaction

Oligo-d(T)-primed total RNAs (0.5 μg per sample) were reverse-transcribed with SuperScript III (Invitrogen Life Technologies, Carlsbad, CA, USA). An appropriate dilution of cDNA and gene-specific primers were combined with SYBR Green Supermix (Invitrogen) and amplified in ABI 7900 real-time PCR machine (Applied Biosystems, Foster City, CA, USA). All qPCR reactions were performed in triplicate. The threshold cycle number (Ct) and expression values with standard deviations were calculated in Excel. Primer sequences for real-time PCRs are: KLK3, F’: 5’-CCAAGTTCTAGCTGTGCT-3’ and R’: 5’-CCCAGTCGTGATACTTTGA-3’; glyceraldehyde-3-phosphate (GAPDH), F’: 5’-CTGACTTCACAGCGA-CACC-3’ and R’: 5’-TGGTGTAGCCAAATTCGTGTT-3’.

Real-time amplification was performed with initial denaturation at 95°C for 2 min, followed by 40 cycles of two-step amplification (95°C for 15 sec, 55°C for 30 sec).

LongSAGE library production and sequencing

RNA from the hollow fibers of three mice (biological replicates) representing different stages of prostate cancer progression (AS, RAD, and CR) were used to make a total of nine LongSAGE libraries. LongSAGE libraries were constructed and sequenced at the Genome Sciences Centre, British Columbia Cancer Agency. Five micrograms of starting total RNA was used in conjunction with the Invitrogen I-SAGE Long kit and protocol with alterations [24]. Raw LongSAGE data are available at Gene Expression Omnibus [25] as series accession number GSE18402. Individual sample accession numbers are as follows: S1885, GSM458902; S1886, GSM458903; S1887, GSM458904; S1888, GSM458905; S1889, GSM458906; S1890, GSM458907; S1891, GSM458908; S1892, GSM458909; and S1893, GSM458910.

Gene expression analysis

LongSAGE expression data was analyzed with DiscoverySpace 4.01 software [26]. Sequence data were filtered for bad tags (tags with one N-base call) and linker-derived tags (artifact tags). Only LongSAGE tags with a sequence quality factor (QF) greater than 95% were included in analysis. The phylogenetic tree was constructed with a distance metric of 1-r (where “r” equals the Pearson correlation coefficient). Correlations were computed (including tag counts of zero) using the Regress program of the Stat package written by Ron Perlman, and the tree was optimized using the Fitch program [27] in the Phylip package [28]. Graphics were produced from the tree files using the program TreeView [29]. Tag clustering analysis was performed using the Poisson distribution-based K-means clustering algorithm. The K-means algorithm clusters tags based on count into K partitions, with the minimum intracluster variance. PoissonC was developed specifically for the analysis of SAGE data [30]. The java implementation of the algorithm was kindly provided by Dr. Li Cai (Rutgers University, NJ, USA). An optimal value for K (K = 10) was determined [31].

Principle component analysis

Principle component analysis was performed using GeneSpring™ software version 7.2 (Silicon Genetics, CA). Affymetrix datasets of clinical prostate cancer and normal tissue were downloaded from Gene Expression Omnibus [25] (accession numbers: GDS1439 and
and analyzed in GeneSpring™. Of the 96 novel CR-associated genes, 76 genes had corresponding Affymetrix probe sets. These probe sets were applied as the gene signature in this analysis. Principle component (PC) scores were calculated according to the standard correlation between each condition vector and each principle component vector.

**Results**

**LongSAGE library and tag clustering**

RNA isolated from the LNCaP Hollow Fiber model was obtained from at least three different mice (13N, 15N, and 13R; biological replicates) at three stages of cancer progression that were androgen-sensitive (AS), responsive to androgen-deprivation (RAD), and castration-recurrent (CR). To confirm that the samples represented unique disease-states, we determined the levels of $KK3$ mRNA, a biomarker that correlates with progression, using quantitative real-time polymerase chain reaction (qRT-PCR). As expected, $KLK3$ mRNA levels dropped in the stage of cancer progression that was RAD versus AS (58%, 49%, and 37%), and rose in the stage of cancer progression that was CR versus RAD (229%, 349%, and 264%) for mice 13R, 15N, and 13N, respectively (Additional file 1). Therefore, we constructed nine LongSAGE libraries, one for each stage and replicate.

LongSAGE libraries were sequenced to 310,072 - 339,864 tags each, with a combined total of 2,931,124 tags, and filtered to leave only useful tags for analysis (Table 1). First, bad tags were removed because they contain at least one N-base call in the LongSAGE tag sequence. The sequencing of the LongSAGE libraries was base called using PHRED software. Tag sequence-quality factor (QF) and probability was calculated to ascertain which tags contain erroneous base-calls. The second line of filtering removed LongSAGE tags with probabilities less than 0.95 (QF < 95%). Linkers were introduced into SAGE libraries as known sequences utilized to amplify ditags prior to concatenation. At a low frequency, linkers ligate to themselves creating linker-derived tags (LDTs). These LDTs do not represent transcripts and were removed from the LongSAGE libraries. A total of 2,305,589 useful tags represented by 263,197 tag types remained after filtering. Data analysis was carried out on this filtered data.

The LongSAGE libraries were hierarchically clustered and displayed as a phylogenetic tree. In most cases, LongSAGE libraries made from the same disease stage (AS, RAD, or CR) clustered together more closely than LongSAGE libraries made from the same biological replicate (mice 13N, 15N, or 13R; Figure 1). This suggests the captured transcriptomes were representative of disease stage with minimal influence from biological variation.
decreased in the stage that was RAD, but increased again in the stage that was CR. Interestingly, steroid hormone receptor activity continues to increase throughout progression. Both of these expression trends were observed for genes with GO terms for transcription factor activity or secretion. The GO categories for genes with kinase activity and signal transduction displayed expression trends with peaks and valleys at the stage that was RAD. The levels of expression of genes involved in cell adhesion rose in the stage that was RAD, but dropped again in the stage that was CR.

Altogether, genes with functional categories that were enriched in expression trends may be consistent with the AR signaling pathway playing a role in progression of prostate cancer progression. Table 1 Composition of LongSAGE libraries

| Library | S1885 | S1886 | S1887 | S1888 | S1889 | S1890 | S1891 | S1892 | S1893 |
|---------|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| Mouse-Condition | 13N-AS* | 13N-RAD† | 13N-CR‡ | 15N-AS | 15N-RAD | 15N-CR | 13R-AS | 13R-RAD | 13R-CR |
| Unfiltered Total Tags | 310,516 | 318,102 | 339,864 | 338,210 | 310,072 | 326,870 | 337,546 | 314,440 | 335,504 |
| No. of Bad Tags | 955 | 1,010 | 1,083 | 1,097 | 983 | 737 | 900 | 744 | 832 |
| Minus Bad Tags | | | | | | | | | |
| Total Tags | 309,561 | 317,092 | 338,781 | 337,113 | 309,089 | 326,133 | 336,646 | 313,696 | 334,672 |
| Tag Types | 79,201 | 96,973 | 99,730 | 81,850 | 84,499 | 88,249 | 79,859 | 91,438 | 90,675 |
| No. of Duplicate Ditags | 19,761 | 12,220 | 12,678 | 21,973 | 17,471 | 12,836 | 24,552 | 12,786 | 13,127 |
| % of Duplicate Ditags | 6.38 | 3.85 | 3.74 | 6.52 | 5.65 | 3.94 | 7.29 | 4.08 | 3.92 |
| Average QF§ of Tags | 0.85 | 0.88 | 0.87 | 0.86 | 0.89 | 0.88 | 0.88 | 0.88 | 0.87 |
| No. of Tags QF < 0.95 | 63,057 | 62,872 | 71,576 | 68,993 | 54,627 | 54,470 | 68,981 | 101,215 | 69,647 |
| Total Tags Combined | 246,504 | 254,220 | 267,205 | 268,120 | 254,462 | 271,663 | 267,665 | 212,481 | 265,025 |
| Tag Types Combined | 52,033 | 67,542 | 66,748 | 52,606 | 59,374 | 64,985 | 53,715 | 54,682 | 64,837 |
| Total Tags Combined | 2,307,345 | 263,199 |
| Tag Types Combined | | |
| No. of LDTs Type I | 124 | 72 | 174 | 179 | 84 | 186 | 164 | 118 | 301 |
| No. of LDTs Type II | 19 | 9 | 54 | 56 | 33 | 40 | 60 | 24 | 59 |
| Minus LDTs | | | | | | | | | |
| Total Tags | 246,361 | 254,139 | 266,977 | 267,885 | 254,345 | 271,437 | 267,441 | 212,339 | 264,665 |
| Tag Types | 52,031 | 67,540 | 66,746 | 52,604 | 59,372 | 64,983 | 53,713 | 54,680 | 64,835 |
| Total Tags Combined | 2,305,589 | 263,197 |
| Tag Types Combined | | |

* AS, Androgen-sensitive/
† RAD, Responsive to androgen-deprivation.
‡ CR, Castration-recurrent.
§ QF, Quality Factor.
II LDTs, Linker-derived tags.

Figure 1 Clustering of the nine LongSAGE libraries in a hierarchical tree. The tree was generated using a Pearson correlation-based hierarchical clustering method and visualized with TreeView. LongSAGE libraries constructed from similar stages of prostate cancer progression (AS, androgen-sensitive; RAD, responsive to androgen-deprivation; and CR, castration-recurrent) cluster together. 13N, 15N, and 13R indicate the identity of each animal.
Consistent differential gene expression associated with progression of prostate cancer

Pair-wise comparisons were made between LongSAGE libraries representing the transcriptomes of different stages (AS, RAD, and CR) of prostate cancer progression from the same biological replicate (3 mice: 13N, 15N, or 13R). Among all three biological replicates, the number of consistent statistically significant differentially expressed tag types were determined using the Audic and Claverie test statistic [36] at $p \leq 0.05$, $p \leq 0.01$, and $p \leq 0.001$ (Table 2). The tags represented in Table 2 were included only if the associated expression trend was common among all three biological replicates. The Audic and Claverie statistical method is well-suited for LongSAGE data, because the method takes into account the deduplication of the libraries and tag counts. Tag types were counted multiple times if they were over, or under-represented in more than one comparison. The number of tag types differentially expressed decreased by 57% as the stringency of the p-value increased from $p \leq 0.05$ to 0.001.

Tag types consistently differentially expressed in pair-wise comparisons were mapped to RefSeq (March 4th, 2008). Tags that mapped anti-sense to genes, or mapped ambiguously to more than one gene were not included in the functional analysis. GO, Kyoto Encyclopedia of Genes and Genomes (KEGG; v45.0) [37] pathway, and SwissProt (v13.0) [38] keyword annotation enrichment analyses were conducted using EASE (v1.21; March 11th, 2008) and FatiGO (v3; March 11th, 2008) [39] (Table 3). This functional analysis revealed that the expression of genes...
involved in signaling increased during progression, but the expression of genes involved in protein synthesis decreased during progression. Cell communication increased in the stage that was RAD but leveled off in the stage that was CR. Carbohydrate, lipid and amino acid synthesis was steady in the RAD stage but increased in the CR stage. Lastly, glycolysis decreased in the RAD stage, but was re-expressed in the CR stage (Table 3).

Table 2 Number of tag types consistently and significantly differentially expressed among all three biological replicates and between conditions*

| Comparison       | Change   | p ≤ 0.001 | p ≤ 0.01 | p ≤ 0.05 |
|------------------|----------|-----------|----------|----------|
| **AS vs. RAD**   | Up in RAD| 21        | 44       | 83       |
|                  | Down in RAD| 68       | 105      | 149      |
|                  | Total     | 89        | 149      | 232      |
| **RAD vs. CR**   | Up in CR  | 24        | 45       | 89       |
|                  | Down in CR| 46       | 59       | 104      |
|                  | Total     | 70        | 104      | 193      |
| **AS vs. CR**    | Up in CR  | 111       | 167      | 294      |
|                  | Down in CR| 127       | 168      | 256      |
|                  | Total     | 238       | 335      | 550      |

* Statistics according to the Audic and Claverie test statistic.
† AS, Androgen-sensitive.
‡ RAD, Responsive to androgen-deprivation.
§ CR, Castration-recurrent.
Table 3 Top five enrichments of functional categories of tags consistently and significantly differentially expressed among all three biological replicates and between stages of prostate cancer*

| Top 5 GO † categories | P-value ‡ | Top 5 KEGG § annotations | P-value ‡ | Top 5 SwissProt annotations | P-value ‡ |
|------------------------|-----------|---------------------------|-----------|----------------------------|-----------|
| AS vs. RAD: Up in RAD¶ |           |                           |           |                             |           |
| Cell communication     | 2E-02     | Stilbene, coumarine and lignin biosynthesis | 1E-02     | Antioxidant                 | 7E-04     |
| Extracellular          | 2E-02     | Butanoyl metabolism       | 2E-02     | Cell adhesion               | 5E-03     |
| Extracellular matrix   | 2E-02     | 2,4-Dichlorobenzoate degradation | 2E-02     | Signal                      | 6E-03     |
| Synaptic vesicle transport | 3E-02   | Cell adhesion molecules (CAMts) | 2E-02     | Fertilization               | 7E-03     |
| Synapse                | 4E-02     | Alkaloid biosynthesis II   | 5E-02     | Amyotrophic lateral sclerosis | 7E-03    |
| AS vs. RAD: Down in RAD |           |                           |           |                             |           |
| Glycolysis             | 3E-05     | Glycolysis/Gluconeogenesis | 3E-05     | Glycolysis                  | 3E-07     |
| Glucose catabolism     | 1E-04     | Ribosome                   | 2E-03     | Pyrroline carboxylic acid   | 8E-05     |
| Hexose catabolism      | 1E-04     | Carbon fixation            | 3E-03     | Pyridoxal phosphate         | 2E-04     |
| Hexose metabolism      | 2E-04     | Fructose and mannose metabolism | 2E-02     | Gluconeogenesis             | 3E-04     |
| Monosaccharide catabolism | 2E-04   | Urea cycle and metabolism of amino groups | 3E-02     | Coiled coil                 | 5E-03     |
| RAD vs. CR: Up in CR   |           |                           |           |                             |           |
| Acid phosphatase activity | 4E-02 | gamma-Hexachlorocyclohexane degradation | 5E-03 | Lyase                        | 2E-03     |
| Lyase activity**      | 7E-02     | Glycolysis/Gluconeogenesis | 3E-02     | Immune response             | 5E-03     |
| Carbohydrate metabolism** | 9E-02 | O-Glycan biosynthesis      | 5E-02     | Signal                      | 6E-03     |
| Extracellular**        | 1E-01     | Ether lipid metabolism**   | 6E-02     | Glycolysis                  | 7E-03     |
| Catabolism**           | 1E-01     | Phenylalanine, tyrosine and tryptophan biosynthesis** | 6E-02 | Progressive external ophthalmoplegia | 1E-02 |
| RAD vs. CR: Down in CR |           |                           |           |                             |           |
| Cytosolic ribosome     | 2E-09     | Ribosome                   | 2E-11     | Ribosomal protein           | 6E-10     |
| Large ribosomal subunit | 1E-07 | Urea cycle and metabolism of amino groups | 1E-02 | Ribonucleaseprotein         | 3E-08     |
| Cytosol                | 2E-07     | Arginine and proline metabolism | 4E-02 | Acetylation                 | 1E-05     |
| Cytosolic large ribosomal subunit | 2E-07 | Type II diabetes mellitus** | 1E-01 | Elongation factor           | 1E-03     |
| Protein biosynthesis   | 2E-07     | Phenylalanine metabolism** | 1E-01 | rRNA-binding                | 2E-03     |
| AS vs. CR: Up in CR    |           |                           |           |                             |           |
| Synapse                | 4E-03     | Butanoate metabolism       | 2E-03     | Glycoprotein                | 2E-03     |
| Extracellular          | 5E-03     | Ascorbate and aldarate metabolism | 2E-02 | Vitamin C                   | 7E-03     |
| Transition metal ion binding | 7E-03 | Phenylalanine metabolism  | 2E-02 | Lipoprotein                 | 1E-02     |
| Metal ion binding      | 2E-02     | Linoleic acid metabolism   | 2E-02     | Signal                      | 1E-02     |
| Extracellular matrix   | 2E-02     | gamma-Hexachlorocyclohexane degradation | 2E-02 | Heparin-binding             | 1E-02     |
| AS vs. CR: Down in CR  |           |                           |           |                             |           |
| Cytosolic ribosome     | 4E-12     | Ribosome                   | 2E-09     | Acetylation                 | 2E-07     |
| Biosynthesis           | 7E-11     | Carbon fixation            | 9E-04     | Ribosomal protein           | 1E-06     |
| Macromolecule biosynthesis | 2E-10 | Glycolysis/Gluconeogenesis | 3E-03 | Glycolysis                  | 7E-05     |
| Protein biosynthesis   | 1E-08     | Glycosphingolipid biosynthesis - lactoseries | 4E-02 | Ribonucleaseprotein         | 8E-05     |
| Eukaryotic 43 S preinitiation complex | 2E-08 | Glutamate metabolism** | 8E-02 | Protein biosynthesis        | 1E-04     |

* Statistics according to the Audic and Claverie test statistic (p ≤ 0.05).
† GO, Gene Ontology.
‡ P-value represents the raw EASE (Expression Analysis Systematic Explorer) score.
§ KEGG, Kyoto Encyclopedia of Genes and Genomes.
II Unadjusted p-value was computed using FatiGO.
¶ AS, androgen-sensitive; RAD, responsive to androgen-deprivation; CR, castration-recurrent.
** Not statistically significant (p > 0.05).
Table 4 Gene expression trends of LongSAGE tags that consistently and significantly altered expression in CR prostate cancer*

| Tag Sequence | S1885 | S1886 | S1887 | S1888 | S1889 | S1890 | S1891 | S1892 | S1893 | Trend‡ | Gene** | Accession§§ |
|--------------|-------|-------|-------|-------|-------|-------|-------|-------|-------|--------|---------|------------|
| TCTAGAAGAACACTGTGC | 121 | 79 | 382 | 7 | 67 | 136 | 7 | 52 | 200 | A | ACPP‡ | NM_001099 |
| TAATTTTTCACTTATGGTGT | 101 | 311 | 648 | 119 | 397 | 7 | 67 | 136 | 7 | 52 | 200 | A | C1ORF80 | ENSG00000186063 |
| TGAGGAGGCCAGGACAA | 8 | 39 | 150 | 4 | 39 | 144 | 7 | 33 | 95 | A | N/A | Genomic |
| CTCATAGGAAGGAGGTTA | 637 | 952 | 1680 | 653 | 1170 | 1540 | 688 | 1620 | 1930 | A | RNF208 | BC009006 |
| GATTITCTATTTGTITTT | 89 | 169 | 446 | 116 | 208 | 399 | 86 | 311 | 555 | A | SERINC5 | ENSG00000163400 |
| GTGGAAGAGGACTACCC | 426 | 571 | 742 | 273 | 417 | 741 | 262 | 363 | 495 | A | STEAP1 | NM_012449 |
| GAGGATACCTGAGGCGC | 191 | 299 | 449 | 134 | 189 | 589 | 187 | 203 | 314 | B | AQP3 | NM_004925 |
| TTGTTGATTGAAAATTT | 219 | 197 | 528 | 273 | 197 | 479 | 232 | 391 | 586 | B | AMD1‡‡ | NM_001634 |
| TGAGGAGGCC | 8 | 39 | 150 | 4 | 39 | 144 | 7 | 33 | 95 | A | N/A | Genomic |
| TCTCATAGGAAGGAGGTTA | 637 | 952 | 1680 | 653 | 1170 | 1540 | 688 | 1620 | 1930 | A | RNF208 | BC009006 |
| GATTITCTATTTGTITTT | 89 | 169 | 446 | 116 | 208 | 399 | 86 | 311 | 555 | A | SERINC5 | ENSG00000163400 |
| GTGGAAGAGGACTACCC | 426 | 571 | 742 | 273 | 417 | 741 | 262 | 363 | 495 | A | STEAP1 | NM_012449 |
| GAGGATACCTGAGGCGC | 191 | 299 | 449 | 134 | 189 | 589 | 187 | 203 | 314 | B | AQP3 | NM_004925 |
| TTGTTGATTGAAAATTT | 219 | 197 | 528 | 273 | 197 | 479 | 232 | 391 | 586 | B | AMD1‡‡ | NM_001634 |
| TGAGGAGGCC | 8 | 39 | 150 | 4 | 39 | 144 | 7 | 33 | 95 | A | N/A | Genomic |
| TCTCATAGGAAGGAGGTTA | 637 | 952 | 1680 | 653 | 1170 | 1540 | 688 | 1620 | 1930 | A | RNF208 | BC009006 |

*Romanuik et al. BMC Medical Genomics 2010, 3:43 http://www.biomedcentral.com/1755-8794/3/43*. 

Page 8 of 19
Table 4: Gene expression trends of LongSAGE tags that consistently and significantly altered expression in CR prostate cancer* (Continued)

| Gene ID | Description | Fstart | Fend | Rstart | Rend | Gstart | Gend | H | N/A | Genomic |
|---------|-------------|--------|------|--------|------|--------|------|---|-----|---------|
| CTGATCGTGTTCTCCT | 575 | 607 | 399 | 429 | 267 | 419 | 197 | 44 | 51 | 56 | NM_001015 |
| AGCGGATTTGATGTT | 1070 | 1092 | 289 | 311 | 250 | 292 | 205 | 49 | 53 | 56 | NM_001958 |
| CTGCCCTCCACAGAGAG | 1150 | 1172 | 339 | 361 | 270 | 312 | 225 | 45 | 49 | 53 | NM_002539 |
| CGGTGTTCCACAGAGAG | 1230 | 1252 | 363 | 385 | 294 | 336 | 249 | 49 | 53 | 56 | NR_002578 |
| AGCGGATTTGATGTT | 1310 | 1332 | 357 | 379 | 288 | 330 | 251 | 47 | 51 | 56 | NM_001961 |
| AGCGGATTTGATGTT | 1390 | 1412 | 375 | 397 | 306 | 348 | 260 | 50 | 54 | 58 | NM_001012334 |
| ACCACGAGGCTTCCT | 1470 | 1492 | 394 | 416 | 325 | 367 | 279 | 52 | 56 | 60 | NM_002129 |
| AGCGGATTTGATGTT | 1550 | 1572 | 413 | 435 | 344 | 386 | 297 | 53 | 57 | 61 | NM_006098 |
| CGGTGTTCCACAGAGAG | 1630 | 1652 | 431 | 453 | 362 | 404 | 316 | 55 | 59 | 63 | NM_001962 |
| CTGCCCTCCACAGAGAG | 1710 | 1732 | 449 | 471 | 380 | 422 | 333 | 57 | 61 | 65 | NR_002591 |
| AGCGGATTTGATGTT | 1790 | 1812 | 467 | 489 | 400 | 442 | 351 | 60 | 64 | 68 | NM_000991 |
| ACCACGAGGCTTCCT | 1870 | 1892 | 485 | 507 | 416 | 458 | 369 | 63 | 67 | 71 | NM_000975 |
| CTGCCCTCCACAGAGAG | 1950 | 1972 | 503 | 525 | 434 | 476 | 387 | 67 | 71 | 75 | NM_002539 |
| AGCGGATTTGATGTT | 2030 | 2052 | 521 | 543 | 452 | 494 | 395 | 71 | 75 | 79 | NM_000998 |
| CTGCCCTCCACAGAGAG | 2110 | 2132 | 539 | 561 | 469 | 511 | 412 | 75 | 79 | 83 | NM_012423 |
| AGCGGATTTGATGTT | 2190 | 2212 | 557 | 579 | 488 | 530 | 431 | 79 | 83 | 87 | NM_000917 |
| ACCACGAGGCTTCCT | 2270 | 2292 | 575 | 597 | 506 | 548 | 449 | 83 | 87 | 91 | NM_012423 |
| AGCGGATTTGATGTT | 2350 | 2372 | 593 | 615 | 524 | 566 | 467 | 91 | 95 | 99 | NM_000917 |

*Continued from previous page.
Additional file 1, Figure S3. Mapping information was provided where available.

We cross-referenced these 114 candidate genes with 28 papers that report global gene expression analyses on tissue samples from men with ‘castration-recurrent’, ‘androgen independent’, ‘hormone refractory’, ‘androgen-ablation resistant’, ‘relapsed’, or ‘recurrent’ prostate cancer, or animal models of castration-recurrence [42-69].

Table 4: Gene expression trends of LongSAGE tags that consistently and significantly altered expression in CR prostate cancer* (Continued)

| Tag Sequence                  | Observed Count | Fold Change | p-value | Gene Symbol | Gene Symbol | Annotation |
|-------------------------------|----------------|-------------|---------|-------------|-------------|------------|
| AGGCTGATGTCCTGGTGGTGGTGGTG    | 2453           | 67          | 0.006   | ACAT        | ACAT        | RCS1       |
| GTGTGATGTCCTGGTGGTGGTGGTG     | 1235           | 31          | 0.012   | ACAT        | ACAT        | RCS1       |
| GCTGATGTCCTGGTGGTGGTGGTG     | 421            | 11          | 0.026   | ACAT        | ACAT        | RCS1       |
| TCTGATGTCCTGGTGGTGGTGGTG     | 345            | 9           | 0.043   | ACAT        | ACAT        | RCS1       |
| TGTGATGTCCTGGTGGTGGTGGTG     | 217            | 5           | 0.058   | ACAT        | ACAT        | RCS1       |
| GCTGATGTCCTGGTGGTGGTGGTG     | 183            | 4           | 0.086   | ACAT        | ACAT        | RCS1       |
| TCTGATGTCCTGGTGGTGGTGGTG     | 159            | 4           | 0.111   | ACAT        | ACAT        | RCS1       |
| TGTGATGTCCTGGTGGTGGTGGTG     | 135            | 3           | 0.137   | ACAT        | ACAT        | RCS1       |
| GCTGATGTCCTGGTGGTGGTGGTG     | 111            | 2           | 0.163   | ACAT        | ACAT        | RCS1       |
| TCTGATGTCCTGGTGGTGGTGGTG     | 98             | 2           | 0.189   | ACAT        | ACAT        | RCS1       |
| TGTGATGTCCTGGTGGTGGTGGTG     | 75             | 2           | 0.215   | ACAT        | ACAT        | RCS1       |
| GCTGATGTCCTGGTGGTGGTGGTG     | 56             | 1           | 0.241   | ACAT        | ACAT        | RCS1       |
| TCTGATGTCCTGGTGGTGGTGGTG     | 44             | 1           | 0.267   | ACAT        | ACAT        | RCS1       |
| TGTGATGTCCTGGTGGTGGTGGTG     | 35             | 1           | 0.293   | ACAT        | ACAT        | RCS1       |
| GCTGATGTCCTGGTGGTGGTGGTG     | 28             | 1           | 0.319   | ACAT        | ACAT        | RCS1       |
| TCTGATGTCCTGGTGGTGGTGGTG     | 21             | 1           | 0.345   | ACAT        | ACAT        | RCS1       |
| TGTGATGTCCTGGTGGTGGTGGTG     | 17             | 1           | 0.371   | ACAT        | ACAT        | RCS1       |
| GCTGATGTCCTGGTGGTGGTGGTG     | 14             | 1           | 0.397   | ACAT        | ACAT        | RCS1       |
| TCTGATGTCCTGGTGGTGGTGGTG     | 11             | 1           | 0.423   | ACAT        | ACAT        | RCS1       |
| TGTGATGTCCTGGTGGTGGTGGTG     | 8               | 1           | 0.449   | ACAT        | ACAT        | RCS1       |
| GCTGATGTCCTGGTGGTGGTGGTG     | 6               | 1           | 0.475   | ACAT        | ACAT        | RCS1       |

* Statistics according to the Audic and Claverie test statistic (p ≤ 0.05).
† Tag count per 1 million = (observed tag count/total tags in the library) × 1,000,000.
‡ Trends are visually represented from A to P in Additional file 1, Figure S3. In addition to p-value considerations, significantly different trends were also required to display uniform directions of change in each biological replicate.
§ AS, Androgen-sensitive.
¶ RAD, Responsive to androgen-deprivation.
** CR, Castration-recurrent.
†† Human Genome Nomenclature Committee (HGNC)-approved gene names were used when possible. Non-HGNC-approved gene names were not italicized.
‡‡ Tag maps antisense to gene.
§§ Accession numbers were displayed following the priority (where available): RefSeq > Mammalian Gene Collection > Ensembl Gene. If the tag mapped to more than one transcript variant of the same gene, the accession number of the lowest numerical transcript variant was displayed.
The candidate genes were identified with HUGO Gene Nomenclature Committee (HGNC) approved gene names, aliases, descriptions, and accession numbers. The gene expression trends of 18 genes of 114 genes were previously associated with CRPC. These genes were: ACPP, ADAM2, AMACR, AMD1, ASAHI1, DHCRC24, FLNA, KLK3, KPNB1, PLAG2G2A, RPL13A, RPL35A, RPL37A, RPL39, RPLP2, RPS20, STEAP2, and TACC (Table 4). To our knowledge, the gene expression trends of the remaining 96 genes have never before been associated with CRPC (Tables 4 & 5).

A literature search helped to gauge the potential of these 96 genes to be novel biomarkers or therapeutic targets of CRPC. The results of this literature search are presented in Table 5. We found 31 genes that encode for protein products that are known, or predicted, to be plasma membrane bound or secreted extracellularly (Bioinformatic Harvester). These genes were: ABHD2, AQP3, B2 M, C19orf48, CD151, CXC7, DHR57, ELOVL5, ENDO1, ENO2, FGFR1L, GNB2L1, GRB10, HLA-B, MARCKSL1, MDK, NAT14, NELF, OPRK1, OR51E2, PLCB4, PTGFR, RAMP1, S100A10, SPON2, STEAP1, TFPI, TMEM30A, TMEM66, TRPM8, and VPS13B. Secretion of a protein could facilitate detection of the putative biomarkers in blood, urine, or biopsy sample. Twenty-one of the candidate genes are known to alter their levels of expression in response to androgen. These genes were: ABHD2, B2 M, BTG1, C19orf48, CAMK2N1, CXC7, EEF1A2, ELOVL5, ENDO1, HSD17B4, MAOA, MDK, NKX3-1, ODC1, P4HA1, PCGEM1, PGK1, SELENBP1, TMEM66, TPD52, and TRPM8 [9,22,70-81]. Genes regulated by androgen may be helpful in determining the activation status of AR in CRPC. Enriched expression of a protein in prostate tissue could be indicative of whether a tumor is of prostatic origin. Eight of these 96 genes are known to be over-represented in prostate tissue [75,82-85]. These genes were: ELOVL5, NKX3-1, PCGEM1, PCOTH, RAMP1, SPON2, STEAP1, and TPD52. Twenty-six genes (ABHD2, BNIP3, EEF1A2, ELOVL5, GALT3, GLO1, HSD17B4, MARCKSL1, MDK, NGFRAP1, ODC1, OR51E2, PCGEM1, PCOTH, PGK1, PP2CB, PSMA7, RAMP1, RPS18, SELENBP1, SLC25A4, SLC25A6, SPON2, STEAP1, TPD52, and TRPM8) have known associations to prostate cancer [59-81].

### Novel CR-associated genes identify both clinical samples of CRPC and clinical metastasis of prostate cancer

The expression of novel CR-associated genes were validated in publically available, independent sample sets representing different stages of prostate cancer progression (Gene Expression Omnibus accession numbers: GDS1390 and GDS1439). Dataset GDS1390 includes expression data of ten AS prostate tissues, and ten CRPC tissues from Affymetrix U133A arrays [47]. Dataset GDS1439 includes expression data of six benign prostate tissues, seven localized prostate cancer tissues, and seven metastatic prostate cancer tissues from Affymetrix U133 2.0 arrays [97].

Unsupervised principal component analysis based on the largest three principal components revealed separate clustering of tumor samples representing AS and CR stages of cancer progression, with the exception of two CR samples and one AS sample (Figure 4a).

Metastatic prostate cancer is expected to have a more progressive phenotype and is associated with hormonal progression. Therefore, the gene expression signature obtained from the study of hormonal progression may be common to that observed in clinical metastases. Unsupervised principal component analysis based on the largest three principal components revealed separate clustering of not only benign and malignant, but also localized and metastatic tissue samples (Figure 4b).

### Discussion

Genes that change levels of expression during hormonal progression may be indicative of the mechanisms involved in CRPC. Here we provide the most comprehensive gene expression analysis to date of prostate cancer with approximately 3 million long tags sequenced using in vivo samples of biological replicates at various stages of hormonal progression to improve over the previous libraries that are approximately 70,000 short tags or less. Previous large-scale gene expression analyses have been performed with tissue samples from men with advanced prostate cancer [42-58], and animal or xenograft models of CRPC [59-69]. Most of these previous studies compared differential expression between CRPC samples with the primary samples obtained before androgen ablation. This experimental design cannot distinguish changes in gene expression that are a direct response to androgen ablation, or from changes in proliferation/survival that have been obtained as the prostate cancer cells progress to more a more advanced phenotype. Here we are the first to apply an in vivo model of hormonal progression to compare gene
expression between serial samples of prostate cancer before (AS), and after androgen ablation therapy (RAD) as well as when the cells become CR. This model is the LNCaP Hollow Fiber model [21] which has genomic similarity with clinical prostate cancer [23] and mimics the hormonal progression observed clinically in response to host castration as measured by levels of expression of PSA and cell proliferation. Immediately prior to castration, when the cells are AS, PSA levels are elevated and the LNCaP cells proliferate. A few days following castration, when the cells are RAD, PSA levels drop and the LNCaP cells cease to proliferate, but do not apoptose in this model. Approximately 10 weeks following castration, when the cells are CR, PSA levels rise and the LNCaP cells proliferate in the absence of androgen. This model overcomes some limitations in other studies using xenografts that include host contamination of prostate cancer cells. The hollow fibers prevent infiltration of host cells into the fiber thereby allowing retrieval of pure populations of prostate cells from within the fiber. The other important benefit of the fiber model is the ability to examine progression of cells to CRPC at various stages within the same host mouse over time, because the retrieval of a subset of fibers entails only minor surgery. The power to evaluate progression using serial samples from the same mouse minimizes biological variation to enhance the gene expression analyses. However, limitations of this model include the lack of cell-cell contact with stroma cells, and lack of heterogeneity in tumors. Typically, these features would allow paracrine interactions as expected in clinical situations. Consistent with the reported clinical relevance of this model [23], here principal component analysis based on the expression of these novel genes identified by LongSAGE, clustered the clinical samples of CRPC separately from the androgen-dependent

Table 5: Characteristics of genes with novel association to castration-recurrence in vivo (Continued)

| Gene      | PM | S | A | P | Y | I | G | R |
|-----------|----|---|---|---|---|---|---|---|
| NELF      | -  | - | - | - | - | - | - | - |
| PM -      | -  | - | - | - | - | - | - | - |
| WDR45L    | -  | - | - | - | - | - | - | - |
| YWHAQ     | -  | - | - | - | - | - | - | - |

* Human Genome Nomenclature Committee (HGNC)-approved gene names were used when possible. Non-HGNC-approved gene names were not italicized.
† S or PM, gene product is thought to be secreted (S) or localize to the plasma membrane (PM).
‡ Reg. by A, gene expression changes in response to androgen in prostate cells.
§ Spec. to P, gene expression is specific to- or enriched in- prostate tissue compared to other tissues.
II CaP, gene is differentially expressed in prostate cancer compared to normal, benign prostatic hyperplasia, or prostatic intraepithelial neoplasia.
* GG, gene is differentially expressed in higher Gleason grade tissue versus lower Gleason grade tissue.
** Prog, gene expression correlates with late-stage prostate cancer or is a risk factor that predicts progression.
†† Mets, gene expression is associated with prostate cancer metastasis in human samples or in vivo models.
‡‡ CR, gene is associated with castration-recurrent prostate cancer in human tissue or in vivo models, but exhibits an opposite trend of this report
§§ Y, yes; ↑, high gene expression; ↓, low gene expression.

Figure 4 Principle component analyses of clinical samples. A, Principle component analysis based on the expression of novel CR-associated genes in the downloaded dataset GDS1390 clustered the AS and CR clinical samples into two groups. B, Principle component analysis based on the expression of novel CR-associated genes in the downloaded dataset GDS1439 clustered the clinical samples (benign prostate tissue, benign; localized prostate cancer, Loc CaP; and metastatic prostate cancer, Met CaP) into three groups.
samples. Principal component analysis based on the expression of these genes also revealed separate clustering of the different stages of tumor samples and also showed separate clustering of the benign samples from the prostate cancer samples. Therefore, some common changes in gene expression profile may lead to the survival and proliferation of prostate cancer and contribute to both distant metastasis and hormonal progression. We used this LNCaP atlas to identify changes in gene expression that may provide clues of underlying mechanisms resulting in CRPC. Suggested models of CRPC involve: the AR; steroid synthesis and metabolism; neuroendocrine prostate cancer cells; and/or an imbalance of cell growth and cell death.

**Androgen receptor (AR)**

**Transcriptional activity of AR**

The AR is suspected to continue to play an important role in the hormonal progression of prostate cancer. The AR is a ligand-activated transcription factor with its activity altered by changes in its level of expression or by interactions with other proteins. Here, we identified changes in expression of some known or suspected modifier of transcriptional activity of the AR in CRPC versus RAD such as Cyclin H (CCNH) [107], proteasome macropain subunit alpha type 7 (PSMA7) [108], CUE-domain-containing-2 (CUEDC2) [109], filamin A (FLNA) [110], and high mobility group box 2 (HMGB2) [111]. CCNH and PSMA7 displayed increased levels of expression, while CUEDC2, FLNA, and HMGB2 displayed decreased levels of expression in CR. The expression trends of CCNH, CUEDC2, FLNA, and HMGB2 in CRPC may result in increased AR signaling through mechanisms involving protein-protein interactions or altering levels of expression of AR. CCNH protein is a component of the cyclin-dependent activating kinase (CAK). CAK interacts with the AR and increases its transcriptional activity [107]. Over-expression of the proteosome subunit PSMA7 promotes AR transactivation of a PSA-luciferase reporter [108]. A fragment of the protein product of FLNA negatively regulates transcription by AR through a physical interaction with the hinge region [110]. CUEDC2 protein promotes the degradation of progesterone and estrogen receptors [109]. These steroid receptors are highly related to the AR, indicating a possible role for CUEDC2 in AR degradation. Thus decreased expression of FLNA or CUEDC2 could result in increased activity of the AR. Decreased expression of HMGB2 in CRPC is predicted to decrease expression of at least a subset of androgen-regulated genes that contain palindromic AREs [111]. Here, genes known to be regulated by androgen were enriched in expression trend categories with a peak or valley at the RAD stage of prostate cancer progression. Specifically, of the 13 tags (62%) exhibiting these expression trends ‘E’, ‘F’, ‘J’, ‘K’, or ‘L’ represented known androgen-regulated genes, in contrast to only 22 of the remaining 122 tags (18%; Tables 4 & 5). Overall, this data supports increased AR activity in CRPC, which is consistent with re-expression of androgen-regulated genes as previously reported [68] and similarity of expression of androgen regulated genes between CRPC and prostate cancer before androgen ablation [23].

**Steroid synthesis and metabolism**

In addition to changes in expression of AR or interacting proteins altering the transcriptional activity of the AR, recent suggestion of sufficient levels of residual androgen in CRPC provides support for an active ligand-bound receptor [112]. The AR may become re-activated in CRPC due to the presence of androgen that may be synthesized by the prostate de novo [4] or through the conversion of adrenal androgens. Here, the expression of 5 genes known to function in steroid synthesis or metabolism were significantly differentially expressed in CRPC versus RAD. They are 24-dehydrocholesterol reductase (DHCR24) [113], dehydrogenase/reductase SDR-family member 7 (DHRS7) [114], elongation of long chain fatty acids family member 5 (ELOVL5) [115,116], hydroxysteroid (17-beta) dehydrogenase 4 (HSD17B4) [117], and opioid receptor kappa 1 (OPRK1) [118]. Increased levels of expression of these genes may be indicative of the influence of adrenal androgens, or the local synthesis of androgen, to reactivate the AR to promote the progression of prostate cancer in the absence of testicular androgens.

**Neuroendocrine**

Androgen-deprivation induces neuroendocrine differentiation of prostate cancer. Here, the expression of 8 genes that are associated with neuroendocrine cells were significantly differentially expressed in CRPC versus RAD. They either responded to androgen ablation such as hairy and enhancer of split 6 (HES6) [119], karyopherin/importin beta 1 (KPNB1) [120], monoamine oxidase A (MAOA) [121], and receptor (calcitonin) activity modifying protein 1 (RAMP1) [122], or were increased expressed in CRPC such as ENO2 [122], OPRK1 [118], S100 calcium binding protein A10 (S100A10) [123], and transient receptor potential cation channel subfamily M member 8 (TRPM8) [124].

**Proliferation and Cell survival**

The gene expression trends of GAS5 [125], GNB2L1 [126], MT-ND3, NKX3-1 [127], PCGEM1 [128], PTGFR [129], STEAP1 [130], and TMEM30A [131] were in agreement with the presence of proliferating cells in CRPC. Of particular interest is that we observed a transcript anti-sense to NKX3-1, a tumor suppressor,
highly expressed in the stages of cancer progression that were AS and CR, but not RAD. Anti-sense transcription may hinder gene expression from the opposing strand, and therefore, represents a novel mechanism by which NXX3-1 expression may be silenced. There were also some inconsistencies including the expression trends of BTG1 [132], FGFR1 [133], and PCDH [134] and that may be associated with non-cycling cells. Overall, there was more support at the transcriptome level for proliferation than not, which was consistent with increased proliferation observed in the LNCaP Hollow Fiber model [21].

Gene expression trends of GLO1 [135], S100A10 [136], TRPM8 [137], and P13KCD [138] suggest cell survival pathways are active following androgen-deprivation and/or in CRPC, while gene expression trends of CAMK2NI [139], CCT2 [140], MDK [141,142], TMEM66 [143], and YWHAQ [136] may oppose such suggestion. Taken together, these data neither agree nor disagree with the activation of survival pathways in CRPC. In contrast to earlier reports in which MDK gene and protein expression was determined to be higher in late stage cancer [63,142], we observed a drop in the levels of MDK mRNA in CRPC versus RAD. MDK expression is negatively regulated by androgen [65]. Therefore, the decreased levels of MDK mRNA in CRPC may suggest that the AR is reactivated in CRPC.

Other

The significance of the gene expression trends of AMD1, BNP3, GRB10, MARCKSL1, NGRAP1, ODC1, PPP2CB, PPP2R1A, SLC25A4, SLC25A6, and WDR45L that function in cell growth or cell death/survival were not straightforward. For example, BNP3 and WDR45L, both relatively highly expressed in CRPC versus RAD, may be associated with autophagy. BNP3 promotes autophagy in response to hypoxia [144], and the WDR45L-related protein, WIPI-49, co-localizes with the autophagic marker LC3 following amino acid depletion in autophagosomes [145]. It is not known if BNP3 or putative WDR45L-associated autophagy results in cell survival or death. Levels of expression of NGFRAP1 were increased in CRPC versus RAD. The protein product of NGFRAP1 interacts with p75 (NTR). Together they process caspase 2 and caspase 3 to active forms, and promote apoptosis in 293T cells [146]. NGFRAP1 requires p75 (NTR) to induce apoptosis. However, LNCaP cells do not express p75 (NTR), and so it is not clear if apoptosis would occur in this cell line [147].

Overall, genes involved in cell growth and cell death pathways were altered in CRPC. Increased tumor burden may develop from a small tip in the balance when cell growth outweighs cell death. Unfortunately, the contributing weight of each gene is not known, making predictions difficult based on gene expression alone of whether proliferation and survival were represented more than cell death in this model of CRPC. It should be noted that LNCaP cells are androgen-sensitive and do not undergo apoptosis in the absence of androgens. The proliferation of these cells tends to decrease in androgen-deprived conditions, but eventually with progression begins to grow again mimicking clinical CRPC.

Conclusion

Here, we describe the LNCaP atlas, a compilation of LongSAGE libraries that catalogue the transcriptome of human prostate cancer cells as they progress to CRPC in vivo. Using the LNCaP atlas, we identified differential expression of 96 genes that were associated with castration-recurrence in vivo. These changes in gene expression were consistent with the suggested model for a role of the AR, steroid synthesis and metabolism, neuroendocrine cells, and increased proliferation in CRPC.

Additional material

Additional file 1: Supplementary Figures. Figure S1: qRT-PCR analysis of AKU gene expression during hormonal progression of prostate cancer to castration-recurrence. RNA samples were retrieved from the in vivo LNCaP Hollow Fiber model at different stages of cancer progression that were: AS, androgen-sensitive, day zero (just prior to surgical castration and 7 days post-fiber implantation); RAD, responsive to androgen-deprivation, 10 days post-surgical castration; and CR, castration-recurrent, 72 days post-surgical castration. MNE, mean normalized expression, calculated by normalization to glyceraldehyde-3-phosphate (GAPDH). Error bars represent ± standard deviation of technical triplicates. Each mouse represents one biological replicate. Figure S2: Ten K-means clusters are optimal to describe the expression trends present during progression to castration-recurrence. K-means clustering was conducted over a range of K (number of clusters) from K = 2 to K = 20 and the within-cluster dispersion was computed for each clustering run and plotted against K. The within-cluster dispersion declined with the addition of clusters and this decline was most pronounced at K = 10. The graph of within cluster dispersion versus K shown here is for mouse 13N, but the results were similar for mice 15N and 13R. Figure S3: Trend legend for Table 4. Gene expression trends of LongSAGE tags that consistently and significantly altered expression in CR prostate cancer are represented graphically with trends labeled A-P. * Statistics according to the Audie and Claviner test statistic (p < 0.05).

Abbreviations

ACPP: prostate acid phosphatase; ACTH: adrenocorticotropic hormone; AR: androgen receptor; AREs: androgen response elements; AS: androgen-sensitive; BAX: BCL2-associated X protein; BCL-2: B-cell CLL/Lymphoma 2; BCL2L1: BCL2-like 1; CAK: cyclin-dependent activating kinase; CCND1: cyclin D1; CCNH: Cyclin H; CDK11A: cyclin-dependent kinase inhibitor 1A; CDKN1B: cyclin-dependent kinase inhibitor 1B; CHG: chromogranin; CR: castration-recurrent; CRPC: castration-recurrent prostate cancer; CUEDC2: CUE-domain-containing-2; DHCR24: 24-dehydrocholesterol reductase; DHRS7: dehydrogenase/reductase SDR-family member 7; EASE: Expression Analysis Systematic Explorer; ELOVL5: elongation of long chain fatty acids family member 5; ENU2: neuronal enolase 2; FLNA: filamin A; GO: Gene Ontology; HE56: hairy and enhancer of split 6; HGG: HUGO Gene Nomenclature Committee; HMGB2: high mobility group box 2; HMGCS1: 3-hydroxy-3-methylglutaryl-Coenzyme A synthase 1; HP3A: hypothalamus-pituitary-adrenal;
HSD17B3: hydroxysteroid (17-beta) dehydrogenase 3, HSD17B4: hydroxysteroid (17-beta) dehydrogenase 4, HSD17B5: hydroxysteroid (17-beta) dehydrogenase 5, IL6: interleukin 6; KEGG: Kyoto Encyclopedia of Genes and Genomes; KLK3: kallikrein/kidney 3; KPNB1: karyopherin/importin beta 1; LH/H: LH/H: Luteinizing hormone releasing hormone; LongSAGE: long serial analysis of gene expression; MAOA: monoamine oxidase A; NCOA: nuclear receptor coactivator; NKX3-1: NK homeobox 1; NT: neurotensin; OPRK1: opioid receptor kappa 1; PCA: protein kinase A; PSA: prostate-specific antigen also known as KLK3, PSMA7: proteasome macropain subunit alpha type 7; PTH/P: parathyroid hormone-related protein; qRT-PCR: quantitative real time-polymerase chain reaction; RAD: responsive to androgen-deprivation; RAMP1: receptor (caliclitinin) activity modifying protein 1; RBP: retinoblastoma 1; S100A10: S100 calcium binding protein A10, SQUE: squelene exoposide; TRPM8: transient receptor potential cation channel subfamily M member 8.

Acknowledgements

The authors would like to thank Jean Wang for her excellent technical assistance and Dr. Simon Haile for helpful discussions. This work was supported by funding from NIH, Grant CA105304 (M.D.S.).

Authors’ contributions

TLR and MDS conceived, designed, conducted, and analyzed all experiments described in this manuscript. TLR and MDS wrote the manuscript. GW performed the principle component analysis. MAM was responsible for SAGE library construction and sequencing. OM (tag clustering) and AD (library clustering) aided in bioinformatic analysis. All authors read and approved the final manuscript.

Author’s information

M.D.S. and M.A.M. are Terry Fox Young Investigators. M.A.M. is a Senior Scholar of the Michael Smith Foundation for Health Research.

Competing interests

The authors declare that they have no competing interests.

Received: 20 April 2010 Accepted: 24 September 2010
Published: 24 September 2010

References

1. Feldman BJ, Feldman D: The development of androgen-independent prostate cancer. Nat Rev Cancer 2001, 1(1):34-45.
2. Scher HI, Sawyers CL: Biology of progressive, castration-resistant prostate cancer: directed therapies targeting the androgen-receptor signaling axes. J Clin Oncol 2005, 23(32):8253-8261.
3. So A, Geave M, Hurtado-Col A, Nelson C: Mechanisms of the development of androgen independence in prostate cancer. World J Urol 2005, 23(1):1-9.
4. McPhaul MJ: Mechanisms of prostate cancer progression to androgen independence. Best Pract Res Clin Endocrinol Metab 2008, 22(2):373-388.
5. Visakorpi T, Hytynen E, Kvoitov P, Tanner M, Keinanen R, Paliberg C, Palote A, Tammela T, Isona J, Kallioniemi OP: In vivo amplification of the androgen receptor gene and progression of human prostate cancer. Nat Genet 1995, 9(4):401-406.
6. Ford OH, Gregory CW, Kim D, Smitherman AB, Mohler JL: Androgen receptor gene amplification and protein expression in recurrent prostate cancer. J Urol 2003, 170(5):1817-1821.
7. Gregory CW, He B, Johnson RT, Ford OH, Mohler JL, French FS, Wilson EM: A mechanism for androgen receptor-mediated prostate cancer recurrence after androgen deprivation therapy. Cancer Res 2001, 61(11):4315-4319.
8. Chmela R, Buchanan G, Need EF, Tilley W, Greenberg NM: Androgen receptor coregulators and their involvement in the development and progression of prostate cancer. Int J Cancer 2007, 120(4):719-733.
9. Holzbieletz J, Lai P, LaTulippe E, Smith A, Satagopan J, Zhang L, Ryan C, Smith S, Scher H, Scardino P: et al Gene expression analysis of human prostate carcinoma during hormonal therapy identifies androgen-responsive genes and mechanisms of therapy resistance. Am J Pathol 2004, 164(1):217-227.
10. Mostaghel EA, Nelson PS: Intracrine androgen metabolism in prostate cancer: mechanisms of castration resistance and therapeutic implications. Best Pract Res Clin Endocrinol Metab 2008, 22(2):243-258.
11. Labrie F: Adrenal androgens and intracrinology. Semin Reprod Med 2004, 22(4):299-309.
12. Veldscholte J, Berrevoets CA, Ro-Stalpers C, Kuiper G, Jenster G, Trapman J, Binkmann AO, Mulder E: The androgen receptor in LNCaP cells contains a mutation in the ligand binding domain which affects steroid binding characteristics and response to antiandrogens. J Steroid Biochem Mol Biol 1992, 41(3-8):665-669.
13. Culig Z, Hobisch A, Crnacuer MV, Radmayr C, Trapman J, Hittmair A, Bartsch G, Klocker H: Androgen receptor activation in prostate tumor cell lines by insulin-like growth factor-I, keratinocyte growth factor, and epidermal growth factor. Cancer Res 1994, 54(20):5474-5478.
14. Hobisch A, Eder E, Putz T, Horninger W, Bartsch G, Klocker H, Culig Z: Interleukin-6 regulates prostate-specific protein expression in prostate carcinoma cells by activation of the androgen receptor. Cancer Res 1998, 58(20):4640-4645.
15. Nazareth LV, Weigel NL: Activation of the human androgen receptor through a protein kinase A signaling pathway. J Biol Chem 1996, 271(33):19900-19907.
16. Riegman PH, Vliestra RJ, van der Korppt JA, Binkmann AO, Trapman J: The promoter of the prostate-specific antigen gene contains a functional androgen responsive element. Mol Endocrinol 1991, 5(12):1921-1930.
17. Yuan TC, Veeramani S, Lin MF: Neuroendocrine-like prostate cancer cells: neuroendocrine transdifferentiation of prostate adenocarcinoma cells. Endocr Relat Cancer 2007, 14(3):531-547.
18. Isaac JT: Biology of hormone refractory prostate cancer. Why does it develop? Urol Clin North Am 1999, 26(2):263-273.
19. Velculescu VE, Zhang L, Vogelstein B, Kinzler KW: Serial analysis of gene expression. Science 1995, 270(5235):484-487.
20. Saha S, Sparks AB, Page C, Aikawa V, Wang CJ, Vogelstein B, Kinzler KW, Velculescu VE: Using the transcriptome to annotate the genome. Nat Biotechnol 2002, 20(5):508-512.
21. Sadar MD, Akopian VA, Berardi E: Characterization of a new in vivo hollow fiber model for the study of progression of prostate cancer to androgen independence. Mol Cancer Ther 2002, 1(8):629-637.
22. Wang G, Jones SJ, Marra MA, Sadar MD: Identification of genes targeted by the androgen and P4A signaling pathways in prostate cancer cells. Oncogene 2006, 25(55):7311-7323.
23. Wang G, Wang J, Sadar MD: Crossstalk between the androgen receptor and beta-catenin in castrate-resistant prostate cancer. Cancer Res 2008, 68(23):9918-9927.
24. Siddiqui AS, Khattra J, Delaney AD, Zhao Y, Astell C, Asano J, Babakaiff R, Barber S, Beland J, Bohacek S, et al: A mouse atlas of gene expression: large-scale digital gene-expression profiles from precisely defined developing C57BL/6J mouse tissues and cells. Proc Natl Acad Sci USA 2005, 102(15):18485-18490.
25. Gene Expression Omnibus. [http://www.ncbi.nlm.nih.gov/geo/].
26. DiscoverySpace. [http://www.bcgsc.ca/biosinfo/software/discoveryspace/].
27. Fitch WM, Margoliash E: Construction of phylogenetic trees. Science 1967, 155(760):279-284.
28. Felsenstein J: Numerical methods for inferring evolutionary trees. J Biol Chem 2001, 276(52):47870-47878.
29. Page RD: TreeView: an application to display phylogenetic trees on personal computer. Comput Appl Biosci 1996, 12(4):357-358.
30. Cai L, Huang H, Blackshaw S, Liu JS, Cepko C, Wong WH: Clustering analysis of SAGE data using a Poisson approach. Genome Biol 2004, 5(7):R51.
31. Blackshaw S, Harpavat S, Trimarchi J, Cai L, Huang H, Kuo WP, Weber G, Lee K, Fraioli RE, Cho SH, et al: Genomic analysis of mouse retinal development. PLoS Biol 2004, 2(9):E247.
32. Pruitt KD, Tatusova T, Maglott DR: NCBI reference sequences (RefSeq): a curated nonredundant sequence database of genomes, transcripts and proteins. Nucleic Acids Res 2007, 35 Database: D61-65.
33. Robertson N, Oveisi-Fordorei M, Zuyderduyn SD, Varhol RJ, Fjell C, Marra MA, A mouse atlas of gene expression: large-scale digital gene-expression profiles from precisely defined developing C57BL/6J mouse tissues and cells. Proc Natl Acad Sci USA 2005, 102(15):18485-18490.
34. Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, Davis AP, Dolinski K, Dwight SS, Eppig JT, et al: Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. Nat Genet 2000, 25(1):25-29.
35. Hosack DA, Dennis G Jr, Sherman BT, Lane HC, Lempicki RA: Identifying biological themes within lists of genes with EASE. Genome Biol 2003, 4(10):R70.
36. Audic S, Claverie JM: The significance of digital gene expression profiles. Genome Res 1997, 7(10):986-995.

37. Kanehisa M, Akata M, Goto S, Hattori M, Hirakawa M, Itaguchi M, Katayama T, Kawashima S, Okuda S, Tokimatsu T, et al: KEGG for linking genomes to life and the environment. Nucleic Acids Res 2008, 36(Database):D480-484.

38. Baroch A, Apweiler R: The SWISS-PROT protein sequence database and its supplement TrEMBL in 2000. Nucleic Acids Res 2000, 28(1):45-48.

39. Al-Shahour F, Minguiez P, Tarraga J, Medina I, Alloza E, Montaner D, Papandreou CN: Alterations in primary prostate cancer after androgen ablation therapy. Clin Cancer Res 2004, 10(46):1722-1728.

40. Shiekhattar R, Jain M, Ramakrishnan AS, Kaur S, Nguyen C, Sweeney C, Viale A, Petrovich Z, Walia N, et al: Identification of androgen-regulated genes in prostate cancer. Proc Natl Acad Sci USA 2003, 101(1):290-297.

41. Le Gros AM, Rajan BV, Monda AC, Choo H, Chodosh LA, Daugherty LA, Moinfar F, Weinberger G, Kooistra T, et al: Androgen-regulated gene expression predicts clinical outcome of prostate cancer. J Clin Invest 2004, 113(10):1299-1308.

42. Bismar TA, Demichelis F, Riva A, Kim R, Varambally S, He L, Kutok J, Aster JC, Tangrea MA, Duray PH, et al: Identification of novel androgen-regulated genes in prostate cancer by sequencing of LongSAGE libraries. J Urol 2005, 173(10):3219-3227.

43. Mohler JL, Morris TL, Ford OH, Alvey RF, Sakamoto C, Gregory CW: Identification of differentially expressed genes associated with prostate cancer progression: a rapid expression profiling assay. J Clin Endocrinol Metab 2000, 85(11):3913-3924.

44. Romanuik et al: BMC Medical Genomics 2010, 3:43 http://www.biomedcentral.com/1755-8794/3/43

45. Assikis VJ, Do KA, Wen S, Wang X, Cho-Vega JH, Brisbay S, Lopez R, Bubley GJ, Ross K, Golub TR, Rubin MA, Penning TM, Logothetis CJ, Troncoso P, Papandreou CN: Failure of hormone therapy in hormone-refractory prostate cancer involves systematic restoration of androgen responsive and nonresponsive genes and activation of rapamycin sensitive signaling. Oncogene 2004, 23(21):4242-4252.

46. Stanbrough M, Bubley GJ, Cho-Vega JH, Christodoulou K, Kallioniemi OP, Bubley GJ: Gene expression profiling identifies clinically relevant subtypes of prostate cancer. Proc Natl Acad Sci USA 2003, 101(3):811-816.

47. DeSantis CE Jr, Weir HK, Buchholz UI, Fleshman JW, Sullivan D, Goding S, Weir JR, Brown ML, Rowland J, et al: Androgen-independent prostate cancer: a heterogeneous group of diseases: lessons from a rapid expression profiling assay. J Urol 2006, 175(6):2131-2136.

48. Tarnawski A, Filik J, Wajs P, Chojnacka A, Skrzypczak I, Kwiatkowski P, Szymonik J, Majewska E, Krysiak M, et al: Identification of novel androgen-regulated genes in prostate cancer. Gene 2004, 325(1-2):107-116.

49. Romanuik et al: BMC Medical Genomics 2010, 3:43 http://www.biomedcentral.com/1755-8794/3/43

50. Ryan DP, Al-Shahrour F, Minguez P, Tarraga J, Medina I, Alloza E, Montaner D, Papandreou CN: Failure of hormone therapy in hormone-refractory prostate cancer involves systematic restoration of androgen responsive and nonresponsive genes and activation of rapamycin sensitive signaling. Oncogene 2004, 23(21):4242-4252.

51. Romanuik et al: BMC Medical Genomics 2010, 3:43 http://www.biomedcentral.com/1755-8794/3/43

52. Romanuik et al: BMC Medical Genomics 2010, 3:43 http://www.biomedcentral.com/1755-8794/3/43
transcriptional activity in mammalian cells. Mol Cell Biol 1998, 18(8):4471-4487.

112. Moeller JL, Gregory CW, Ford OH, Kim D, Weaver CM, Petruzzi P, Wilson EM, French FJ. The androgen axis in recurrent prostate cancer. Clin Cancer Res 2004, 10(6):440-448.

113. Waterham HR, Koster J, Romeijn GI, Hennekam RC, Verken P, Anderson HC, FitzPatrick DR, Kilner RI, Wanders RJ. Mutations in the 3beta-hydroxysteroid Delta24-reductase gene cause desmosterolism, an autosomal recessive disorder of cholesterol biosynthesis. Am J Hum Genet 2001, 69(4):685-694.

114. Haessler F, Palczewski K. Short-chain dehydrogenases/reductases in retina. Methods Enzymol 2000, 316:573-583.

115. Momozawa Y, Takeuchi Y, Kitago M, Masuda K, Kabukia Y, Hashizume C, Ichimaru T, Mogi K, Okamura H, Yonezawa T, et al. Gene expression profiles linked to the hormonal induction of male-effect phenome androgen synthesis in (Capra hircus). Biol Reprod 2007, 77(1):102-107.

116. Kitago M, Momozawa Y, Masuda K, Wakabayashi Y, Date-tto A, Hagiino-Yamagishi K, Kikusui T, Takeuchi Y, Mori Y. Localization of the candidate genes ELOVL5 and SCDF1 for 'male effect' phenome synthesis in goats (Capra hircus). J Reprod Dev 2007, 53(6):1329-1333.

117. Penning TM. Hydroxysteroid dehydrogenases and pre-receptor regulation of steroid hormone action. Hum Reprod Update 2003, 9(3):193-205.

118. Pascoe JE, Williams KL, Mukhopadhyay P, Rice KC, Woods JH, Ko MC. Effects of mu, kappa, and delta opioid receptor agonists on the function of hypothalamic-pituitary-adrenal axis in monkeys. Psychoneuroendocrinology 2008, 33(4):478-486.

119. Hu Y, Wang T, Stormo GD, Gordon JI. Functional screening for achaete-scute homolog 1 in mouse prostate neuroendocrine cells reveals its gene targets and DNA binding sites. Proc Natl Acad Sci USA 2003, 101(15):5559-5564.

120. Lam VH, Thomas RJ, Loveland KL, Schilders S, Gu M, Martin TJ, Gillespie MT, Jans DA. Nuclear transport of parathyroid hormone (PTH)-related protein (PTHrP) receptor is dependent on microtubules. Mol Endocrinol 2002, 16(2):390-401.

121. Zhao H, Nolley R, Chen Z, Reese SW, Peehl DM. Inhibition of monoamine oxidase A promotes secretory differentiation in basal prostatic epithelial cells. Differentiation 2006, 76(7):820-830.

122. Abrahamsson PA: PCGEM1, a prostate-specific gene with cell growth-promoting function, overexpressed in prostate cancers, promotes prostate cancer cell growth through phosphorylation of oncoprotein TAF-Ib/SET. Cancer Res 2005, 65(11):4578-4586.

123. Svenningsson P, Chergui K, Rachleff I, Flajolet M, Zhang X, El Yacoubi M, Romanuik S, Van Order RR, Chang K, Kuzmak A, et al. The phosphatidylinositol 3-kinase pathway is a dominant growth factor-activated cell survival pathway in LNCaP human prostate carcinoma cells. Cancer Res 1999, 59(12):2891-2897.

124. Rohklin OW, Taghyev AF, Bayer KU, Bunclot D, Koteliansk VE, Glover RA, Cohen MB. Calcium/calmodulin-dependent kinase II plays an important role in prostate cancer cell survival. Cancer Biol Ther 2007, 6(5):732-742.

125. Mourtada-Maarabouni M, Hedge VL, Kirkham L, Farzaneh F, Williams GT. Potential channel TRPM8 agonists stimulate calcium influx and intercellular communication in vitro and growth of human tumor xenografts in vivo. Cancer Res 2007, 67(12):5798-5805.

126. Atanazova Y, Nakagawa H, Furusha M, Ashida S, Tamura K, Yoshihko H, Shuin T, Fujikoa T, Kajagki T, Nakamura Y. PCOITH, a novel gene overexpressed in prostate cancers, promotes prostate cancer cell growth through phosphorylation of oncoprotein TAF-Ib/SET. Cancer Res 2005, 65(11):4578-4586.

127. Petrovics G, Zhang W, Makarem M, Street JP, Connelly R, Sun L, Nkx3.1 mutant mice: 14-3-3 isoforms and P11. Cancer Lett 2008, 275(2):93-102.

128. Proikas-Cezanne T, Waddell S, Gaugel A, Frickey T, Lupas A, Nordheim A. Identification and characterization of CDC50A, CDC50B and CDC50C genes in silico. Oncol Rep 2004, 12(4):939-943.

129. Matsuura S, Rouault J, Magaud J, Berthet C. In search of a function for the T521/PAC/8T1G1/TOB family. FEBS Lett 2001, 497(2-3):67-72.

130. Truex B, Zhuang L, Taeschler S, Wiedemann M. Characterization of FGFR1L, a novel fibroblast growth factor (FGF) receptor preferentially expressed in skeletal tissues. J Biol Chem 2003, 278(36):33857-33865.

131. Kamataki T, Hitaishi T, Saito H, Yamagishi K, Kikusui T, Takeuchi Y, Mori Y. RNA interference of achaete-scute homolog 1 in mouse prostate neuroendocrine cells reveals its gene targets and DNA binding sites. Proc Natl Acad Sci USA 2003, 101(15):5559-5564.

132. Matsuda S, Rouault J, Magaud J, Berthet C. Evidence that TRPM8 is an androgen-dependent Ca2+ channel required for the survival of prostate cancer cells. Cancer Res 2004, 64(22):6365-6373.

133. Lin J, Adam RM, Santievez E, Freeman MR. The phosphatidylinositol 3-kinase pathway is a dominant growth factor-activated cell survival pathway in LNCaP human prostate carcinoma cells. Cancer Res 1999, 59(12):2891-2897.

134. Petrovics G, Zhang W, Makarem M, Street JP, Connelly R, Sun L, Nkx3.1 mutant mice: 14-3-3 isoforms and P11. Cancer Lett 2008, 275(2):93-102.

135. Mourtada-Maarabouni M, Hedge VL, Kirkham L, Farzaneh F, Williams GT. Potential channel TRPM8 agonists stimulate calcium influx and intercellular communication in vitro and growth of human tumor xenografts in vivo. Cancer Res 2007, 67(12):5798-5805.

136. Atanazova Y, Nakagawa H, Furusha M, Ashida S, Tamura K, Yoshihko H, Shuin T, Fujikoa T, Kajagki T, Nakamura Y. PCOITH, a novel gene overexpressed in prostate cancers, promotes prostate cancer cell growth through phosphorylation of oncoprotein TAF-Ib/SET. Cancer Res 2005, 65(11):4578-4586.

137. Svenningsson P, Chergui K, Rachleff I, Flajolet M, Zhang X, El Yacoubi M, Vaugois JM, Nomikos GG, Greengard P. Nuclear transport of parathyroid hormone (PTH)-related protein (PTHrP) receptor is dependent on microtubules. Mol Endocrinol 2002, 16(2):390-401.

138. Petrovics G, Zhang W, Makarem M, Street JP, Connelly R, Sun L, Nkx3.1 mutant mice: 14-3-3 isoforms and P11. Cancer Lett 2008, 275(2):93-102.

Pre-publication history

The pre-publication history for this paper can be accessed here: http://www.biomedcentral.com/1755-8794/3/43/prepub

doi:10.1186/1755-8794-3-43

Cite this article as: Romanuik et al.: LNCaP Atlas: Gene expression associated with in vivo progression to castration-recurrent prostate cancer. BMC Medical Genomics 2010 3:43.