Published in final edited form as:

J Proteome Res. 2014 February 7; 13(2): 1088–1100. doi:10.1021/pr401106h.

Metabolomic profiling identifies biochemical pathways associated with castrate resistant prostate cancer

Akash K Kaushik^{1,2,3,#}, Shaiju K Vareed^{2,3,#}, Sumanta Basu⁵, Vasanta Putluri^{2,3}, Nagireddy Putluri^{2,3}, Katrin Panzitt^{2,3}, Christine A Brennan⁷, Arul M Chinnaiyan⁶, Ismael A. Vergara⁹, Nicholas Erho⁹, Nancy L Weigel³, Nicholas Mitsiades³, Ali Shojaie⁸, Ganesh Palapattu⁶, George Michailidis^{5,@}, and Arun Sreekumar^{1,2,3,@,*}

¹Verna and Marrs McLean Department of Biochemistry and Molecular Biology, Baylor College of Medicine

²Alkek Center for Molecular Discovery, Baylor College of Medicine

³Molecular and Cellular Biology, Baylor College of Medicine

⁴Department of Surgery, Baylor College of Medicine

⁵Department of Statistics, University of Michigan Ann Arbor

⁶Department of Urology, University of Michigan Ann Arbor

⁷Center for Translational Pathology, University of Michigan Ann Arbor

⁸Department of Biostatistics, University of Washington Seattle

⁹GenomeDx Biosciences, Inc., Vancouver, BC, Canada

Abstract

Despite recent developments in treatment strategies, castrate resistant prostate cancer (CRPC) is still the second leading cause of cancer associated mortality among American men, the biological underpinnings of which are not well understood. To this end, we measured levels of 150 metabolites and examined the rate of utilization of 184 metabolites in metastatic androgen dependent prostate cancer (AD) and CRPC cell lines using a combination of targeted mass spectrometry and metabolic phenotyping. Metabolic data were used to derive biochemical pathways that were enriched in CRPC, using Oncomine Concept Maps (OCM). The enriched pathways were then examined in-silico for their association with treatment failure (i.e., prostate specific antigen (PSA) recurrence or biochemical recurrence) using published clinically annotated gene expression data sets. Our results indicate that a total of 19 metabolites were altered in CRPC compared to AD cell lines. These altered metabolites mapped to a highly interconnected network of biochemical pathways that describe UDP glucuronosyltransferase (UGT) activity. We observed an association with time to treatment failure in an analysis employing genes restricted to this pathway in three independent gene expression data sets. In summary, our studies highlight the

*Corresponding author: Arun Sreekumar, Ph.D, Associate Professor, Department of Molecular and Cell Biology, Alkek Center for Molecular Discovery, Baylor College of Medicine, One Baylor Plaza, Houston, TX-77479, sreekuma@bcm.edu, Phone: 713-798-3393, Fax: 713-790-1275.

#These two authors contributed equally to this work.

@Equal senior authors.

Financial Interest Disclosure: None of the authors except Ismael A. Vergara and Nicholas Erho, have any financial disclosure. Both Ismael A. Vergara and Nicholas Erho are employees of GenomeDx Biosciences, Inc., Vancouver, BC, Canada and were not involved in the study design or data interpretation.

Supporting Information **Available:** This material is available free of charge via the Internet at <http://pubs.acs.org>.

value of employing metabolomic strategies in cell lines to derive potentially clinically useful predictive tools.

Keywords

Prostate Cancer; Castrate Resistant Prostate Cancer; Metabolomics; Metabolic Phenotyping; Liquid-chromatography Mass Spectrometry; Oncomine Concept Map and Biochemical Recurrence

Introduction

Prostate Cancer (PCa) is the second leading cause of cancer-related death among men in the United States¹. For over 50 years, PCa has been known to be an androgen dependent disease^{2, 3}. Clinically localized disease can be effectively treated with either surgery or radiation. Metastatic disease, however, while initially responsive to androgen ablation therapy, often evolves into a castrate resistant state that is incurable and leads to cancer associated death. Androgen ablation therapy leads to substantial apoptosis of androgen-dependent cells and tumor regression^{4, 5}. In ~ 20% to 35% cases, however, the tumor recurs within five years^{6, 7}. There is compelling evidence that these recurrent tumors, termed castration resistant PCa (CRPC), continue to require androgen receptor (AR) signaling, despite the reduction in levels of circulating androgens^{8, 9}. Like many advanced tumors, CRPC develops in the background of diverse genetic and environmental factors. Distinct sets of genes, proteins and metabolites are thought to dictate progression from an androgen dependent phenotype to a castrate resistant state. Since AR is a transcription factor, many investigators have sought to identify androgen-regulated genes within PCa cells and tissues^{10, 11}. Remarkably, little is known about the genes regulated by AR in the CRPC state⁹. Because of multiple levels of regulation in gene activity, including translation and post-translational modifications to modify protein activity, changes in mRNA often do not correlate directly with changes in protein expression and function. Another challenge in examining gene function is pathway redundancy. Nevertheless, despite pathway compensation, subtle alterations in enzymatic activity can lead to meaningful changes in levels of metabolites required for tumor growth. Multiple groups have interrogated androgen-regulated changes at the transcriptome and proteome levels in PCa cell lines, using gene expression arrays and mass spectrometry^{12, 13}. One such seminal study using Affymetrix oligonucleotide arrays, highlighted the association of androgen signaling in PCa cells with metabolic processes associated with stress responses¹⁴. Furthermore, androgen-driven proliferation of PCa cells has been shown to involve activation of mammalian target of rapamycin (m-TOR), that itself is sensitive to metabolic perturbations in the tumor¹⁵⁻¹⁷. Until recently, means to perform large scale profiling of metabolites, which are critical effectors of cellular functions, have not been available. Using an integrative analysis of matched gene expression and proteomic data, in earlier work we predicted the activation of amino acid metabolism in androgen-treated LNCaP (androgen sensitive) prostate cancer cells¹⁸. This expectation was further strengthened by metabolomic profiling of PCa tissues that revealed amino acid metabolism as being one of the hallmarks of early tumor development¹⁹. In this tissue-based profiling study a comparative analysis of the levels of 1,126 metabolites using high-throughput liquid and gaschromatography-based mass spectrometry was able to distinguish between benign prostate, organ-confined PCa and metastatic disease. This study delineated metabolic profiles associated with metastasis and CRPC. Due to the lack of same patient pre- and post-treatment samples, there have been no studies thus far that have addressed the metabolic changes that occur during the transition of an androgen-dependent metastatic cells to a castration resistant stage, wherein a re-

programmed androgen receptor signaling is hypothesized to be a major driver of cellular growth and survival.

Here, using well characterized cell line models for metastatic androgen-dependent and castrate resistant disease, we have characterized metabolic profiles using mass spectrometry and nominated predictive pathways associated with PSA recurrence after radical prostatectomy for localized prostate cancer (i.e., biochemical recurrence (BCR)). Our results suggest a strong association of metabolic processes associated with UDP-glucuronosyltransferases activity as well as starch, sucrose, and histidine metabolism with BCR. Importantly, a distinguishing feature of our approach is the combination of mass spectrometry, bioinformatics and metabolic phenotyping with a novel prognosis-associated data mining strategy. In this way, we were able to define a CRPC cell line-derived, pathway-centric metabolic signature, with verified activity that is associated with patient-related clinical outcome.

Materials and Methods

Cell lines

Prostate-derived cell lines (androgen receptor positive – LNCaP and MDA-PCa-2a and MDA-PCa-2b; and, castrate resistant (CRPC)- C4-2, 22Rv1 and LNCaP-abl) were used for this study. Of these, MDA-PCa-2a and MDA-PCa-2b were gifts from Dr. Nora Navone, LNCaP-abl from Dr. Nicholas Mitsiades and C4-2 were originally purchased from UroCor by Dr. Nancy Weigel; the remaining cell lines were purchased from American Type Culture Collection (ATCC, Manassas, VA). LNCaP, and 22Rv1 were grown in RPMI-1640 media (Invitrogen Corp., Carlsbad, CA) supplemented with 10% fetal bovine serum (FBS; Hyclone Labs, Thermo Scientific, Rockford, IL) and 1% penicillin-streptomycin (Hyclone Labs, Thermo Scientific, Rockford, IL). C4-2 cells were grown in T-media (Gibco) supplemented with 5% FBS and 1% penicillin-streptomycin (Hyclone Labs, Thermo Scientific, Rockford, IL). LNCaP-abl cells were grown in RPMI-1640 media (Invitrogen Corp., Carlsbad, CA) lacking phenol red supplemented with 10% charcoal stripped FBS and 1% penicillin-streptomycin (Hyclone Labs, Thermo Scientific, Rockford, IL). MDA-PCa-2a and MDA-PCa-2b were grown in BRFF-HPC1 medium (Athena ES company cat#0403) with 20% FBS, Gentamicin (GIBCO cat#15750-060, final C 50ug/ml) and FNC coating mix solution (Athena ES cat#0407). All cells were maintained at 37°C and 5% CO₂ and regularly monitored for mycoplasma using MycoAlert™ (Lonza, Anaheim, CA). Prior to their analyses, cells were trypsinized and the pellet washed thrice with ice cold PBS, separated into aliquots of 5 million cells and stored at -140°C.

Mass-Spectrometry

Reagents and internal standards—High-performance liquid chromatography (HPLC) grade acetonitrile, methanol and water were purchased from Burdick & Jackson (Morristown, NJ). Mass spectrometry grade formic acid and internal standards namely N-acetyl Aspartic acid-d3, [15N]2 Tryptophan, Sarcosine-d3, Glutamic acid-d5, Thymine-d4, Gibberellic acid, Trans-Zeatin, Jasmonic acid, [15N] Anthranilic acid, and Testosterone-d3 were purchased from Sigma- Aldrich (St.Louis, MO). ESI-L Low Concentration Tuning Mix (Part No. G1969-85000) from Agilent Technologies (Santa Clara) was used to calibrate the mass spectrometer.

Liquid Chromatography-Mass spectrometry—HPLC analysis was performed using an Agilent 1260 series HPLC system equipped with a degasser, binary pump, thermostatted autosampler and column oven (all from Agilent Technologies, Santa Clara, CA). The Multiple Reaction Monitoring (MRM)-based measurement of levels of 150 metabolites was

performed using 12 different methods that used either a reverse phase (RP) or aqueous normal phase (ANP) chromatographic separation, prior to mass spectrometry (Supporting information Table S1). Details of the methods, associated chromatographic conditions, metabolites measured and their corresponding MRM transitions are given in Supporting information Table S1. For reverse phase separation, a Zorbax Eclipse XDB-C18 column (50×4.6 mm i.d.; 1.8 μm particle size, Agilent Technologies, Santa Clara, CA) maintained at 37°C, was used. The mobile phase used for the RP separation contained 0.1% formic acid/water (v/v; A), and 0.1% formic acid/acetonitrile (v/v; B). The gradient conditions were 0-6 min-2%B; 6.5 min-30%B, 7 min-90%B, 12 min- 95%B, followed by re-equilibration to the initial starting condition. The flow rate throughout the separation was maintained at 0.2 mL/min. For ANP separation, a Diamond-Hydride column (100×2.1mm i.d.; 4.2μm particle size, Microsolve Technology Corporation, Eatontown, NJ) maintained at 37°C, was used. The mobile phase used for this chromatographic separation contained acetonitrile ACN (solvent A): water (solvent B), with both solvents modified by the addition of 0.2% Acetic acid and 0.1% formic acid, while the gradient conditions were 0-2 min-95%B; 3 min-90%B, 5 min-80%B, 6-7 min- 75%B, 8 min-55%B, 10 min-40%B, 12 min- 30%B, 14 min-20%B, followed by re-equilibration to the initial starting condition. The flow rate was gradually increased during the separation from 0.5 mL/min (0-8 mins) to 0.7 mL/min (10-14 min) and finally set at 0.8 mL/min (15-17 min). The total run time was set to 20 min. The samples were kept at 4°C and a constant volume of 5 pL was injected for analysis. The mass spectrometric analysis was performed on 6430 QQQ-LC/MS (Agilent Technologies, Santa Clara, CA). The mass spectrometer was operated in both electrospray (ESI) positive and negative ionization modes, with a capillary voltage of 3000 V, a collision gas flow rate of 10 L/min and a nebulizer gas flow rate of 35 L/min. The temperature of nebulizer gas was maintained at 350°C. Nitrogen was used as the collision gas at a collision cell pressure of 2.39×10^{-5} Torr. For maximum sensitivity, the fragmentor voltage and collision energy for each metabolite were optimized separately using the optimizer® software (Agilent Technologies, Santa Clara, CA) and the optimized values were used for the Multiple Reaction Monitoring (MRM) assays. A list of MRM transitions and HPLC methods used for each metabolite measured in this study, is given in Supporting information Table S1.

Internal Standard Solution and Quality controls—Aliquots (200 μL) of 10 mM solutions of N-acetyl Aspartic acid-d₃, Tryptophan-¹⁵N₂, Sarcosine-d₃, Glutamic acid-d₅, Thymine-d₄, Gibberellic acid, Trans-Zeatin, Jasmonic acid, ¹⁵N Anthranilic acid, and Testosterone-d₃ were mixed and diluted up to 8000 μL (Final concentration 0.25 mM) and aliquoted into a 20 μL portions. The aliquots were dried and stored at -80°C. Two kinds of controls were used to monitor the sample preparation and mass spectrometry. To monitor instrument performance, 20 μL of a matrix-free mixture of the internal standards described above, reconstituted in 100 μL of methanol: water (50:50) was analyzed by MRM. In addition, the process of metabolite extraction from the cell lines were monitored using pooled liver samples and spiked internal standards. In the case of the former, the 100 mg of pooled liver was extracted in tandem with the cell lines using buffer containing spiked internal standards as described above. In addition, internal standards were also spiked into cell lines during the extraction process. The matrix-free internal standards and liver samples were analyzed twice daily.

Sample preparation for mass spectrometric analysis—The metabolome extraction method described earlier²⁰ was used for the cell lines and pooled liver controls in this study. Briefly, cells were thawed at 4 °C and subjected to freeze-thaw cycles in liquid nitrogen and over ice three times to rupture the cell membrane. Following this, 750 μL of ice cold methanol:water (4:1) containing 20 μL of spiked internal standards were added to each cell line. The cells were homogenized for 1 min (30 sec pulse twice) and mixed with 450 μL of

ice cold chloroform and vortex mixed in a Multi-Tube Vortexer for 10 min. The resulting homogenate was mixed with 150 μ l of ice cold water and vortexed again for 2 min. The homogenate was incubated at -20 °C for 20 min and centrifuged at 4 °C for 10 min to partition the aqueous and organic layers. The aqueous and organic layers were separated and dried at 37 °C for 45 min in an Automatic Environmental Speed Vac® system (Thermo Fisher Scientific, Rockford, IL). The aqueous extract was reconstituted in 500 μ l of ice cold methanol:water (50:50) and filtered through 3 KDa molecular filter (Amicon Ultracel -3K Membrane, Millipore Corporation, Billerica, MA) at 4 °C for 90 min to remove proteins. The filtrate was dried at 37 °C for 45 min in speed vac and stored at -80 °C until mass spectrometry analysis. Prior to mass spectrometry analysis, the dried extract was resuspended in 100 μ L of methanol: water (50:50) containing 0.1% formic acid and analyzed using MRM.

Metabolic Phenotyping Microarray—Global carbon utilization profiles were evaluated by using Biolog Phenotyping Microarray (PM) MicroPlate (Biolog, Inc., Hayward, CA). PM technology is a high-throughput technology for testing the metabolic status of a cell line. The assay is performed in a 96-well format where each well contains a different metabolite. Utilization of the specific metabolite within the cell is detected colorimetrically by a tetrazolium-based redox dye. Reduction of this dye by reduced nicotinamide adenine dinucleotide (NADH) produced due to the utilization of the metabolite by the cells, results in formation of specific color. The assay was performed as per manufacturer's instructions. Briefly, cells were counted, pelleted and resuspended in inoculation medium (M1-medium). 50 μ l of the cell suspension containing 2×10^4 cells / well were transferred into the plates and incubated at 37 °C in a humidified atmosphere with 5% CO₂ for 24 hours. The next day, 10 μ l of redox dye (dye A) was added per well and plates were kept in the Biolog reader at 37 °C. Absorbance was measured automatically every 15 min for 3.5h using an OmniLog incubator reader.

Data-Analysis of Mass-Spectrometry and Metabolic Phenotyping microarray

The MRM-based mass spectrometry data were log2 transformed and four normalization methods were used: 1) scaling by spiked internal standard (15N2-Tryptrophan), 2) centering and scaling using the median and the inter-quartile range of all the spiked internal standards, 3) centering and scaling using the median and the inter-quartile range for the entire dataset and 4) centering and scaling of the metabolites using the median and inter-quartile range corresponding to each of the 12 methods used. To visualize results of each of these normalization methods, heat maps were drawn using the “gplots” package in the R-software language. To compute the differentials, for each of the above normalization methods, data obtained from different platforms were z-transformed and compared across samples. A two-sided t-test was used to assess the association of each metabolite with AD or CRPC status of the samples. Significance was determined by calculating the null distribution based on 10,000 permutations of the group labels, and the resulting p-values were then adjusted for multiple testing using FDR methodology, computed using the R-package fdrtool. Results were visualized using a complete linkage hierarchical clustering function, which uses the Euclidean distances in the data.

Metabolic phenotyping microarray data were first corrected for background signal, obtained from readings of empty wells. The data for each metabolite was then adjusted by subtracting average values of the three negative controls, and subsequently z-scored. The metabolic flux through amino acid and sugar pathways in androgen dependent vs. CRPC cell lines were compared using heatmaps and boxplots. In addition, gene set analysis (R-package GSA²¹) was used to assess the overall enrichment of amino acids, and sugar using KEGG-derived pathways, in the respective plates.

Core CRPC Metabolic Signature and bioprocess mapping

As described above, the metabolic data for the AD and CRPC cell lines was normalized using four different methods. Following this, the list of metabolites that were commonly altered between AD and CRPC lines, across all four normalization methods, at two FDR thresholds, 6 and 20 %, were derived. The latter threshold was used to describe metabolite-level changes between AD and CRPC, while the former, termed Core CRPC-Associated Metabolic Signature (CCAMS), was used for Oncomine Concept Map-based (OCM) bioprocess enrichment, as described below. In parallel to determine metabolic alterations and pathways between AD and CRPC cells of the same parental origin, we compared the metabolome of LNCaP (AD) with C42 and LNCaP-abl (both CRPC). As described earlier, the data was normalized using four different methods and the commonly altered metabolites at FDR 20 % were used to enrich for bioprocesses using OCM.

For the OCM bioprocess enrichment, metabolites in the CCAMS or those that were altered in LNCaP vs C42 and LNCaP-abl comparison, were converted to enzyme-coding genes using matched KEGG datasets (KEGG API version v6.2, containing a total of 4435 enzyme coding genes and 1405 corresponding metabolites). In parallel, to generate a null set, all the metabolites identified in this study (n=150), were also converted to their respective enzyme-coding genes. As the first step towards this, in the CCAMS signature, all 17/19, mapped to their respective KEGG IDs while only 124/150 (82.66%) metabolites in the entire data set mapped to unique KEGG IDs. Further, out of the 17 metabolites with unique KEGG ID's, 12 metabolites mapped to 78 unique enzyme coding genes (CRPC gene signature). Similarly, the 124 metabolites from the entire dataset mapped to 769 enzyme coding genes. On the other hand, in the metabolic signature altered in LNCaP vs C4-2 and LNCaP-abl comparison, 31 metabolites were altered at FDR corrected P value 0.2 out of which 28 mapped to unique KEGG IDs and mapped to 236 unique enzyme coding genes. This entire process of converting metabolites to their corresponding enzyme coding genes was done using in-house Python language scripts.

Molecular Concept Enrichment Analysis

We used the Oncomine Concept Map (OCM) (private.molecularconcepts.org)²² to derive enriched bioprocesses or molecular concepts associated with CRPC in a network format. A molecular concept also termed molecular signature is a list of genes or proteins that are derived from various external databases and computationally-derived regulatory networks. OCM is composed of more than 14,000 molecular concepts and provides a platform to systematically analyze lists of differentially expressed gene sets. The compendia of molecular concepts for OCM analysis consists of data from 13 databases and 335 high-throughput datasets. The method calculates parameters of association for each individual molecular concept with a gene set (e.g. CRPC gene signature) using Fisher's exact test. The analysis is carried out using a null set to control for background enrichments. In this study, the CCAMS was used to perform bioprocess enrichment, in the background of 769 genes associated with the set of 150 metabolites examined in this study, as the null set.

Further, as additional controls, we also looked for enriched molecular concepts using five randomly generated gene sets, each containing 78 enzyme coding genes. To generate a random gene signature set, we first subtracted the list of genes associated with CCAMS, from the total set of KEGG-derived genes (n=4435). This resulted in a set of 4357 genes, which were used to generate five independent signatures, each containing 78 genes, after multiple rounds of randomization (n=10,000). The resulting five random gene sets were used in parallel with CCAMS, to perform OCM-based enrichment analysis.

Gene Expression Datasets

We used publicly available gene-expression data obtained from the studies of Glinsky²³, Sun²⁴ and Taylor²⁵, to assess the predictive value of the CCAMS-derived molecular concepts, with regards to biochemical recurrence. Raw data for the Taylor and Sun studies were downloaded from GEO (<http://www.ncbi.nlm.nih.gov/geo>, with GEO ID GSE21035 and GSE25136 respectively), while the Glinsky data was provided by Dr. Setlur (Department of Pathology, Harvard Medical School, Massachusetts General Hospital). While the Glinsky and Sun studies contained gene expression data generated on the Affymetrix U95Av2 and Affymetrix Human Genome U133A Array platforms respectively, the Taylor data contained expression values generated on the Agilent-014693 Human Genome platform. All the data sets were generated using prostatectomy tissues and contained follow up information on time to treatment failure or biochemical recurrence (BCR), that is to say PSA failure following radical prostatectomy (i.e., PSA recurrence, a measure of treatment failure).

The Glinsky data contained expression measurements for 12,625 transcripts across 79 discrete prostate tumors (we used gene expression values of 78 samples) of which 37 patients experienced BCR, Sun data contained 14021 transcripts across 79 samples of which 37 patients experienced BCR, and the Taylor data measured the expression of 22,329 transcripts, across 140 individual prostate cancer tissues, of which 36 patients experienced BCR. Data from patients with BCR and no-BCR, from the above three datasets were used for the current analysis. Prior to any downstream analysis, all data sets were log2 transformed and quantile normalized using the R-language “limma” package.

Survival Analysis

To evaluate the predictive potential of pathways enriched by CCAMS, we selected the top 10 OCM concepts that were highly interconnected. These 10 molecular concepts described starch & sucrose metabolism, androgen & estrogen metabolism, pentose and glucuronate conversion and importantly shared a common set of 18 highly redundant genes belonging to the UGT family of transferases^{26, 27}. Thus, to examine the predictive value of CCAMS for biochemical recurrence (BCR), genes belonging to the UGT cluster were selected.

These selected genes were combined into a single eigengene, using Principal Components Analysis (PCA). The resulting PCA scores were partitioned into positive and negative values, thus creating a factorial score to signify high or low expression. Then, a Cox Proportional Hazards Model was built, using this factorial score as the predictor and time to BCR as the outcome variable, with censoring corresponding to time-to-recurrence. To select the combination of genes with the best predictive ability for BCR in all the three data sets (Glinsky, Sun and Taylor), a best subset selection procedure was employed. Thus to begin with, for our PCA analysis, genes were weighted on the basis of their expression profiles across different metastatic samples available in Oncomine²⁸. Following this PCA was applied to subsets of genes, of varying sizes, ranging from 2 to 18 for the UGT family. The PCA-based selection was repeated 10000 times for the UGT gene family. The results of the Cox Proportional Hazard Model that was finally selected, describing the high and low groups, were visualized using Kaplan-Meier (KM) plots. The entire analysis described above was performed using the core package and the “survival” package in the R-language.

RNA seq Data-analysis

The RNA-Seq data contained gene expression information on 267 prostate cancer tissue samples analyzed at the MCTP (Michigan center for translational pathology), at University of Michigan. The sequencing data was aligned using TopHat 2.0.8 against the Ensembl 69 GRCh37 human genome build and FPKM (fragments per kilobase of exon per million

fragments mapped) was calculated across genes in Ensemble 69 using Cufflinks 2.0.2. Graphs were created using a variety of custom R and python scripts. The data was further log2 transformed and the expression of UGT2B15, UGT2B28, UGT2A1 as well as their combined median expression was plotted using package “lattice” in R-language.

Results

In this study (refer to Figure 1A), we used liquid chromatography-coupled-mass-spectrometry (LC-MS) to profile metabolic signatures of both androgen responsive (AD; LNCaP, MDA-PCa-2a and MDA-PCa-2b) and castrate resistant prostate cancer (CRPC; LNCaP-abl, C4-2 and 22Rv1) cell lines. Although the AD cells (LNCaP, MDA-PCa-2a and MDA-PCa-2b) were derived from metastatic sites^{29, 30}, they respond to androgen for their growth and sustenance. In contrast, both LNCaP-abl 31 and C4-2 32 were originally derived by castration of LNCaP cells and can grow independent of androgen levels in the culture medium. 22Rv1, on the other hand is a unique CRPC cell line that harbors an alternately spliced androgen receptor (AR) termed V7-variant that has been reported in at least 23% of CRPC patients³³. Furthermore, these cells also over express SPINK1 that is seen in ~10 % of the PCa patients³⁴. Overall, these prostate cancer cell lines represent distinct features of metastatic androgen dependent (AD) tumors that develop castration resistance after androgen ablation therapy.

The mass spectrometry data were examined using different data-preprocessing methods and a commonly altered CRPC signature, termed “Core CRPC-Associated Metabolic Signature or CCAMS”, was determined. This was subsequently used to identify enriched bioprocesses or metabolic pathways, a subset of which was examined *in silico*, for their predictive power (Figure 1B). Furthermore, enriched pathways were independently verified using metabolic phenotype microarrays (Figure 1A).

Metabolic profiling of prostate-derived cell lines

We used MRM liquid chromatography-coupled-mass-spectrometry (LC-MS) to define the metabolic signatures of both AD and CRPC cell lines (outlined in Figure 1A). The targeted mass spectrometry analysis used in this study, was monitored for its robustness using both matrix-free standards and pooled matrix controls. The assay platform had a median CV <2.9% for both matrix-free internal standards and liver pool respectively (Supporting information Figure S1). A total of 150 metabolites (refer to Supporting information Table S1 for list of MRM transitions) were measured for their relative steady state levels (Figure 2A and Supporting information Table S2). This included mostly polar and mid-polar metabolites that were selected from prior prostate cancer publications^{19, 20} and included carbohydrate, amino acid, nucleotide, amino sugar, polyamine, organic acid, fatty acid and carnitine metabolism. Importantly, each metabolite was identified using two specific transitions that included a qualifier and a quantifier (Supporting information Table S1). The mass spectrometry data was pre-processed using Mass Hunter software, followed by use of in-house statistical tools (see above), to derive the CRPC-associated metabolic signature. Based on the knowledge that the compendia of differential compounds identified by analysis of high throughput data is sensitive to the data pre-processing steps³⁵, we examined the AD and CRPC cell line-derived mass spectrometry-derived data, using four different normalization methods, to derive a core CRPC associated metabolic signature (CCAMS) (Figure 1A), (refer to Supporting information Table S1 and Materials and Methods for details). A differential list of metabolites was generated using each of these pre-processing methods (FDR corrected *P* value<0.2, Supporting information Table S3 for the list of differential metabolites for each normalization method). A total of between 40-50 metabolites (FDR corrected *P* value<0.2) were found to be differentially altered in each of

these normalization processes, of which 38 compounds were commonly altered in all 4 preprocessing methods. Commonly altered metabolites between AD and CRPC cell lines (FDR corrected P value < 0.2), termed core CRPC-associated metabolic signature (CCAMS) are shown in Figure 2B, while Supporting information Figures S2-S5, show the distinct metabolic signatures obtained for each of the normalization methods. Out of 38 commonly altered metabolites, levels of 12 compounds were elevated in CRPC while the remaining 26 metabolites were higher in AD cells. Metabolites elevated in CRPC included sugars like fructose, and galactose, as well as energy / signaling intermediates like NAD, GMP, AMP and ADP. Metabolites elevated in AD included carnitines as well as metabolites in the amino sugar or hexose monophosphate pathway. Also, AD cells contained higher levels of methylated metabolites like methylalanine, and dimethylglycine, as well as elevated levels of amino acids, supporting our earlier observation of androgen signaling regulating the methylation axis and amino acid metabolism, in AD cells²⁰. We also compared the metabolic signature of LNCaP (AD) with LNCaP-abl and C4-2 (CRPC) cells which provided further insights into metabolic alterations in AD vs CRPC cells derived from the same parental cell line. There were 31 differential metabolites (FDR corrected P value 0.2) in this comparison, of which 9/31 metabolites were also common to CCAMS (Supporting information Figure S6 A). The common metabolites included galactose, glucuronic acid, N-acetyl galactosamine, asparagine, guanosine monophosphate (GMP), aminobutyric acid, 2-oxo-3-methyl valerate, palmitoylcarnitine and hexonycarnitine. Similar to CCAMS, galactose, asparagine and GMP were elevated in the CRPC cells whereas glucuronic acid, N-acetyl galactosamine, hexonycarnitines and palmitoylcarnitine were elevated in AD cells. Overall, based on the above findings, it appears that CRPC cells, tend to generate higher levels of acetyl CoA, NAD and second messengers, while being inefficient in mitochondrial transport of fatty acid break down intermediates. In contrast, AD cells seem to be more efficient in utilizing amino acids for their survival.

OCM-based enrichment analysis of altered metabolites

To obtain better insight into pathways represented in our metabolic data (see Figure 1B), a subset of metabolites associated with the CCAMS (FDR corrected P value < 0.06 , $n=19$), were first mapped to their KEGG IDs, which were then sequentially mapped to their corresponding enzymes and associated gene IDs. Thus, as described in Figure 1B, the 19 differential metabolites were mapped to their KEGG IDs, using KEGG database (KEGG API version v6.2). All except two metabolites mapped to a unique KEGG identifier. The two metabolites that did not map to a KEGG ID, hexanoylcarnitine and ethyl-3-indoleacetate, were not used for further analysis. The KEGG ID's for the metabolites were then mapped to their corresponding enzyme IDs that were in turn matched up with their gene IDs. Overall, of the 17 metabolites with unique identifiers, 12 compounds mapped to a total of 78 enzyme coding genes (termed CRPC-gene-signature, Supporting information Table S4). The CCAMS, thus generated, was used to carry out the OCM-based enrichment analysis. As controls for the enrichment analysis, five random metabolic gene signatures (Supporting information Table S5), each 78 gene long were also enriched using OCM. These random gene sets were selected from a pool of 4435 KEGG-associated metabolic genes, after 10,000 independent iterations (see Methods). Notably, when used in OCM, only the CCAMS, enriched for significant biochemical processes, while the random metabolic gene signatures did not.

The CCAMS enriched a total of 61 concepts at FDR Q value < 0.05 (Supporting information Table S6). The top 10 pathways enriched by the CCAMS (P value $< 2E-05$, FDR Q value < 0.002 , see Supporting information Table S7 for the list) were then visualized as a network (Figure 3). From these data we observed a tightly interconnected network of molecular concepts (brown bridges) that described glucuronosyltransferase activity, pentose and

glucuronate interconversions, porphyrin and chlorophyll metabolism, microsome, starch and sucrose metabolism, UDP-glucuronosyl/UDP-glucosyltransferase, androgen and estrogen metabolism, and chromosomal arms 4q13, and 2q37. Interestingly, all these concepts share a common set of genes belonging to the UDP-Glucuronosyltransferase (UGT) family. This includes UGT1 (genes n=9) and UGT2 (genes n=9). We further investigated the OCM based pathway enrichment using metabolites that were significantly altered (FDR corrected P value<0.2) between LNCaP and LNCaP-abl/C4-2 cells. The enriched molecular concepts (FDR corrected P value<1E-05) were clustered into three groups of which one of them described UDP-glucuronosyltransferases (Supporting information Figure S6 B). The other two clusters described enrichment of receptor interaction/activity and metal dependent phosphodiesterases activity, respectively (Supporting information Figure S6 B). Furthermore, the 9 metabolites that were common between CCAMS and our LNCaP vs C42 and LNCaP-abl comparison also enriched for UDP-glucuronosyltransferases along with concepts describing 3-5-cyclic nucleotide and metal dependent phosphodiesterases activity (Supporting information Figure S6 C).

Importantly, UDP-glucuronosyltransferases and its components have been previously identified in independent studies to play an essential role in prostate cancer progression^{36, 37}, and therefore, reinforce the power of our approach in being able to obtain biologically relevant pathways using bioinformatics analysis of cell line-derived metabolic datasets. We also examined the list of differential metabolites (FDR Q value<0.05) obtained in 3 out of 4 normalization methods, for enriched bioprocesses (Supporting information Figure S7). Here, in addition to the bioprocesses identified by the CCAMS, we identified a cluster of pathways associated with amino acid metabolism describing those associated with glutamate, alanine & aspartate metabolism (Supporting information Figure S7, red bridges), and, an independent set of concepts describing galactose and aminosugar metabolism (Supporting information Figure S7, green bridges).

Next, we examined for any overlap between the cell line-derived CRPC concepts and those obtained from the metastatic tissue-associated metabolic data that we had published earlier¹⁹. As shown in Supporting information Figure S8, a total of 7 bioprocesses overlapped between these two datasets, supporting the clinical relevance of our cell line-based findings.

Metabolic Phenotyping Microarray

To ascertain the validity of the pathways identified by OCM analysis, metabolic phenotyping microarrays (MPM) were carried out on AD and CRPC cell lines. The underlying principle of MPM is that the rate of utilization of various metabolites that serve as the sole source of carbon, or nitrogen can be determined by measuring the amount of NADH produced over time. The latter is quantified by measuring absorbance of reduced tetrazolium dye at 450 nm.

Thus, in MPM, the rate of NADH production is used as a proxy of activity through biochemical pathways that utilize the given metabolite. MPM measured the rate of utilization of 184 (92 amino acids and 92 sugar intermediates) metabolites (Supporting information Table S8) distributed in two 96-well plates. These metabolites were part of the PMM1 and PMM2 plates marketed by BIOLOG Inc, and were chosen based on our mass spectrometry findings that show alteration in metabolites associated with the sugar and amino acid pathways. Data from these plates were examined using the GSA enrichment-based approach²¹, to obtain pathways that were significantly (FDR corrected P value<0.05) altered between CRPC and AD cell lines. A total of 23 metabolic pathways were found to be altered at FDR level of 5% (refer Table 1 for list of altered pathways). Included among the top 9 in this list were fructose and mannose metabolism, galactose metabolism,

phenylalanine metabolism, glycoxylate and dicarboxylate metabolism, ABC transporters, amino sugar metabolism, starch and sucrose metabolism and TCA cycle, among others.

Intriguingly, 9/23 altered pathways (shown in bold red fonts in Table 1) were also nominated by the OCM analysis to be associated with CRPC. Thus, OCM and MPM revealed alterations in starch/sucrose metabolism, aminosugar metabolism, pentose phosphate pathway, alanine and aspartate pathway, pentose and glucuronate interconversion, glutamate metabolism, glycolysis and pyruvate metabolism. This study verifies the robustness of the metabolic pathways defined by mass spectrometry-coupled bioinformatics analysis for CRPC.

Evaluation of the clinical relevance of CRPC-associated metabolic pathways

To examine the clinical relevance of the CCAMS-associated metabolic pathways, we carried out an in-silico analysis, looking for any association between the component genes in our identified metabolic pathways enriched in CRPC with BCR. For this association-based analysis, Glinsky, Sun and Taylor prostate cancer gene expression data sets were used. A summary of the clinical information associated with each of these datasets is provided in Supporting information Tables S9-S11. These data sets contained approximately 70-140 prostate cancer patients with a median follow up time of 54-60 months post-prostatectomy. For the analysis, initially the gene expression values for the metabolic genes associated with CRPC were collected from Glinsky, Sun and Taylor datasets. Following this, the combination of genes associated with a specific metabolic pathway were examined using Principal Component Analysis (PCA) to derive a threshold value for their average expression that was then used to divide the samples into two groups, namely those having low or high expression. The association of the low or high group with BCR and hence the ability of the PCA-derived threshold to make this distinction, was then examined using a Cox-proportional hazard model and the results were visualized using Kaplan Meier (KM) plots (See method section for complete description).

For the UDP-glucuronosyltransferase pathway, higher combined expression of three genes, namely UGT2B15, UGT2A1 and UGT2B28, was found to be highly associated (P value < 0.05 in Glinsky, Sun and Taylor) with BCR (Figures 4 A-C). Furthermore, consistent with this observation, the combined median expression of UGT2B15, UGT2A1 and UGT2B28 was significantly elevated (P value < 0.05) in metastatic prostate cancer compared to localized tumors in the Grasso³⁸, Taylor²⁵ and Varambally³⁹ datasets whereas the expression was not significant due to small sample numbers in the Vanaja⁴⁰ dataset (Figures 4D-G). The clinical relevance of combined expression of UGT2A1, UGT2B15 and UGT2B28 was further validated using a RNA-seq dataset containing expression values for 267 prostate cancer tissues, sequenced at the Michigan Center for Translational Pathology, University of Michigan. The 267 tissues included 26 benign adjacent, 93 organ-confined tumors and 148 metastatic tumors. The metastatic tissues were collected during warm autopsy. Importantly, combined expression of UGT2A1, UGT2B15 and UGT2B28 was significantly higher in metastatic prostate cancer tissues compared to organ-confined disease ($P = 1.2E-05$, Figure 5D). Further, when examined independently, both UGT2B28 and UGT2B15 were found to be significantly elevated in metastatic disease compared to localized PCa ($P < 1E-05$, Figure 5A-B). Importantly however, genes associated with pathways enriched using randomly generated metabolic gene signatures, did not show a significant correlation with BCR-based clinical outcome.

Discussion

Cancer is a complex disease governed by multiple alterations at the level of transcriptomics, proteomics and metabolomics. Recent developments in high-throughput technologies have

strengthened the methods used for the integration of reproducible data from different platforms to determine a global picture of the pathogenesis of cancer. One such study in breast cancer, integrated high-throughput gene-expression and proteomics data with biological network information and determined defining molecular components of breast tumorigenesis⁴¹. This and other similar studies, highlight the power of an integrative pathway-centric approach to studying cancer. In the present study, we have employed a similar approach but with a distinct method to combine (steady state levels) data from metabolomics with those from gene-expression in a pathway centric manner and have further validated the (flux) activity of identified pathways using metabolic phenotyping microarrays. Using such an approach, we have determined clinically relevant biochemical pathways that are associated with the development of CRPC, as well as predict primary treatment failure or biochemical recurrence.

AR signaling plays a critical role in the development of PCa, by regulating both oncogenic and metabolic genes^{42, 43}. In our earlier study, we described an AR-regulated metabolic signature in AD PCa, wherein alterations in amino acid metabolism were a key component²⁰. Clinically, the dependence of PCa on androgen signaling is exploited by the administration of androgen ablation as the first line of treatment for metastatic tumors and those that recur and progress following failed primary treatment (radiation or surgery). For men who receive androgen ablation for metastatic prostate cancer, the median time to development of CRPC is approximately 16-20 months^{44, 45}. In these patients, AR signaling appears to be active and independent of circulating levels of androgen. Studies from many groups have revealed activating changes in AR such as mutations⁴⁶, amplification⁴⁷, as well as alterations in the receptor that permit promiscuous activity via binding to other steroid hormones like estrogen⁴⁸, glucocorticoids⁴⁹ etc. Recent experimental evidence also suggests expression of AR splice variants in androgen insensitive and bone metastasis associated with poor survival outcome in patients with CRPC^{50, 51}. Changes in AR co-repressors and co-activators are also thought to be important in CRPC⁵². Irrespective of its mode of activation, AR in the castrate resistant state has been shown to activate distinct gene expression signatures^{53,54} as well as specific metabolic programs in metastatic prostate cancer¹⁹ that together support PCa progression. In a previous study using metastatic PCa tissues, we have shown the existence of a distinct metabolome compared to localized tumors¹⁹. Metastatic tissues however, represent the distal end of the progression cascade wherein the tumor has colonized into distal sites. From the clinical perspective, it is important to capture the changes that occur at the transition of AD to CRPC, something that has been difficult to study given the scarcity of such samples. Recently, different groups have addressed this issue by developing AD cell lines from metastatic tissue sites that exhibit CRPC-like properties⁹. Here, we adopted a cell line-based metabolic profiling approach to identify metabolic alterations in AD and CRPC and further examined our results in the context of patient data.

In this study, we compared metabolic profiles of LNCaP, MDA-PCa-2a and MDA-PCa-2b, all AD; with LNCaP-abl, C4-2 and 22Rv1, all CRPC like lines. Notably, both LNCaP-abl and C4-2 were generated by androgen deprivation of LNCaP AD cells, while 22Rv1, is known to harbor a V7 variant of AR and express high levels of SPINK1, both of which have been found in advanced prostate cancer patients^{34, 55}. Importantly, we used two orthogonal methods to interrogate the metabolome, namely mass spectrometry, and metabolic phenotyping microarray. The former, measures the steady state levels of the metabolites, while the latter is an NADH-dependent colorimetric method that evaluates pathway activity.

Using a mass spectrometry-based approach, our observed CRPC signature appears to contain elevated levels of sugars as well as intermediates associated with energy metabolism and signaling. In contrast, AD cells seem to possess elevated levels of amino acids and their

methyated derivatives. These data support our earlier work describing amino acids to be sensitive to AR signaling and also important in localized PCa development^{19, 20}. In addition, this finding also suggests a metabolic re-programming in CRPC, shifting cellular dependence to sugar metabolism.

Importantly, OCM analysis of the isolated metabolic profiles corroborates this theme wherein the major metabolic perturbations when comparing AD with CRPC appear to occur in sugar-associated pathways. Intriguingly, our OCM network contains highly interconnected nodes describing UDP-glucuronosyltransferase activity, sucrose metabolism and pentose glucuronate interconversions. Biochemically, these are known to be highly interconnected and were found to be co-enriched by a set of UGT-associated genes in our OCM network. Importance of UGTs in CRPC was further highlighted when we compared metabolic profile of AD (LNCaP) and CRPC (LNCaP-abl and C4-2) cells derived from sample parental line.

UGT-dependent metabolism has gained major attention in CRPC with the finding that it plays an important role in the metabolism of steroid hormones including androgens²⁶. Our mass spectrometry derived metabolic data show lower levels of glucuronic acid, a key component of the UGT pathway in CRPC cells, suggestive of its increased utilization for glucuronidation reactions and hence potentially regulating the levels of androstane-3 α ,17 β -diol-G glucuronides in these cell²⁶. Supporting this idea, previous studies have reported higher expression of UGT2B15 and UGT2B17 in CRPC cells like LNCaP-abl compared to AD LNCaP cells, and AR negative DU145 and PC3 cells, implying interaction of androgen signaling with this pathway^{36, 56}. These findings motivated us to then examine the predictive capacity of the UGT-associated pathway for treatment failure or biochemical recurrence (BCR).

To address this, we examined three independent gene expression datasets. Importantly 47.4, 46.8 and 25.7 % of patients in the Glinsky, Sun and Taylor studies, respectively, experienced BCR within the follow up period of the study. Using a PCA-coupled Cox-proportional hazard model, on these datasets, we asked whether genes associated with the UGT-associated pathways had any association with time to BCR (i.e., treatment failure). Interestingly, our results suggested that an aggregate of expression values for UGT2B15, UGT2A1 and UGT2B28, (all belonging to the UDP-glucuronosyltransferase) in PCa tissues collected at surgery prior to any other treatment, could predict time to BCR. When visualized using Kaplan-Meier plots, tumors having higher combined expression for these three genes showed BCR at an earlier time point (i.e., worse outcome) compared to those having lower combined expression levels for these genes. This was validated in 3/4 independent gene expression datasets wherein higher combined expression of these UGT-associated genes (UGT2B15, UGT2A1 and UGT2B28) was observed in metastatic PCa compared to organ-confined tumors^{25, 38-40}, a finding that was independently confirmed using a RNA-seq dataset. From a larger perspective, these results indicate that monitoring the expression levels of a set of UGT genes could be used as a means to refine prediction tools for men with localized PCa who undergo radical prostatectomy. Our data need to be validated; however do demonstrate the potential utility of a cell line derived metabolomic approach to studying human cancer.

In summary, this study, for the first time, defined a cell line-based metabolic signature for CRPC. This signature was then translated into pathways using Oncomine concept map-based enrichment and validated using metabolic phenotyping microarray. Importantly, expression levels of the delineated metabolic pathways in localized PCa, appear to correlate with the risk of treatment failure or biochemical recurrence.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We would like to acknowledge Praveenkumar Akula Sureshkumar for his assistance in metabolic phenotyping microarray (MPM). This research was supported by the following grant support NIH 1R01CA133458-01 (ASK), Susan Komen Foundation Grant KG110818 (ASK), NIH U01 CA167234 (ASK, GP and GM), DOD W81XWH-12-1-0130 (ASK, GP and GM), NSF DMS-1161759 (ASK), NSF DMS-12-28164 (GM) and NSF DMS-11617838 (GM), RP120092 (ASK) and Funds from the Alkek Center for Molecular Discovery (ASK).

References

1. Siegel R, Naishadham D, Jemal A. Cancer statistics, 2012. *CA Cancer J Clin.* 2012; 62(1):10–29. [PubMed: 22237781]
2. Huggins C, Hodges CV. Studies on prostatic cancer. I. The effect of castration, of estrogen and androgen injection on serum phosphatases in metastatic carcinoma of the prostate. *CA Cancer J Clin.* 1972; 22(4):232–40. [PubMed: 4625049]
3. Wang X, Yin L, Rao P, Stein R, Harsch KM, Lee Z, Heston WD. Targeted treatment of prostate cancer. *J Cell Biochem.* 2007; 102(3):571–9. [PubMed: 17685433]
4. Ohlson N, Wikstrom P, Stattin P, Bergh A. Cell proliferation and apoptosis in prostate tumors and adjacent non-malignant prostate tissue in patients at different time-points after castration treatment. *Prostate.* 2005; 62(4):307–15. [PubMed: 15389788]
5. Westin P, Bergh A. Apoptosis and other mechanisms in androgen ablation treatment and androgen independent progression of prostate cancer: a review. *Cancer Detect Prey.* 1998; 22(5):476–84.
6. Uchio EM, Asian M, Wells CK, Calderone J, Concato J. Impact of biochemical recurrence in prostate cancer among US veterans. *Arch Intern Med.* 2010; 170(15):1390–5. [PubMed: 20696967]
7. Ward JF, Blute ML, Slezak J, Bergstralh EJ, Zincke H. The long-term clinical impact of biochemical recurrence of prostate cancer 5 or more years after radical prostatectomy. *J Urol.* 2003; 170(5):1872–6. [PubMed: 14532796]
8. Massie CE, Lynch A, Ramos-Montoya A, Boren J, Stark R, Fazli L, Warren A, Scott H, Madhu B, Sharma N, Bon H, Zecchini V, Smith DM, Denicola GM, Mathews N, Osborne M, Hadfield J, Macarthur S, Adryan B, Lyons SK, Brindle KM, Griffiths J, Gleave ME, Rennie PS, Neal DE, Mills IG. The androgen receptor fuels prostate cancer by regulating central metabolism and biosynthesis. *EMBO J.* 2011; 30(13):2719–33. [PubMed: 21602788]
9. Wang Q, Li W, Zhang Y, Yuan X, Xu K, Yu J, Chen Z, Beroukhim R, Wang H, Lupien M, Wu T, Regan MM, Meyer CA, Carroll JS, Manrai AK, Janne OA, Balk SP, Mehra R, Han B, Chinnaiyan AM, Rubin MA, True L, Fiorentino M, Fiore C, Loda M, Kantoff PW, Liu XS, Brown M. Androgen receptor regulates a distinct transcription program in androgen-independent prostate cancer. *Cell.* 2009; 138(2):245–56. [PubMed: 19632176]
10. Ngan S, Stronach EA, Photiou A, Waxman J, Ali S, Buluwela L. Microarray coupled to quantitative RT-PCR analysis of androgen-regulated genes in human LNCaP prostate cancer cells. *Oncogene.* 2009; 28(19):2051–63. [PubMed: 19363526]
11. Velasco AM, Gillis KA, Li Y, Brown EL, Sadler TM, Achilleos M, Greenberger LM, Frost P, Bai W, Zhang Y. Identification and validation of novel androgen-regulated genes in prostate cancer. *Endocrinology.* 2004; 145(8):3913–24. [PubMed: 15131019]
12. Khan AP, Poisson LM, Bhat VB, Fermin D, Zhao R, Kalyana-Sundaram S, Michailidis G, Nesvizhskii AI, Omenn GS, Chinnaiyan AM, Sreekumar A. Quantitative proteomic profiling of prostate cancer reveals a role for miR-128 in prostate cancer. *Mol Cell Proteomics.* 2010; 9(2): 298–312. [PubMed: 19955085]
13. Xu LL, Su YP, Labiche R, Segawa T, Shanmugam N, McLeod DG, Moul JW, Srivastava S. Quantitative expression profile of androgen-regulated genes in prostate cancer cells and identification of prostate-specific genes. *Int J Cancer.* 2001; 92(3):322–8. [PubMed: 11291065]

14. Segawa T, Nau ME, Xu LL, Chilukuri RN, Makarem M, Zhang W, Petrovics G, Sesterhenn IA, McLeod DG, Moul JW, Vahey M, Srivastava S. Androgen-induced expression of endoplasmic reticulum (ER) stress response genes in prostate cancer cells. *Oncogene*. 2002; 21(57):8749–58. [PubMed: 12483528]
15. Kaarbo M, Mikkelsen OL, Malerod L, Qu S, Lobert VH, Akgul G, Halvorsen T, Maelandsmo GM, Saatcioglu F. PI3K-AKT-mTOR pathway is dominant over androgen receptor signaling in prostate cancer cells. *Cell Oncol*. 2010; 32(1-2):11–27. [PubMed: 20203370]
16. Kobayashi T, Shimizu Y, Terada N, Yamasaki T, Nakamura E, Toda Y, Nishiyama H, Kamoto T, Ogawa O, Inoue T. Regulation of androgen receptor transactivity and mTOR-S6 kinase pathway by Rheb in prostate cancer cell proliferation. *Prostate*. 2010; 70(8):866–74. [PubMed: 20127734]
17. Wu Y, Chhipa RR, Cheng J, Zhang H, Mohler JL, Ip C. Androgen receptormTOR crosstalk is regulated by testosterone availability: implication for prostate cancer cell survival. *Anticancer Res*. 2010; 30(10):3895–901. [PubMed: 21036700]
18. Vellaichamy A, Sreekumar A, Strahler JR, Rajendiran T, Yu J, Varambally S, Li Y, Omenn GS, Chinnaiyan AM, Nesvizhskii AI. Proteomic interrogation of androgen action in prostate cancer cells reveals roles of aminoacyl tRNA synthetases. *PLoS One*. 2009; 4(9):e7075. [PubMed: 19763266]
19. Sreekumar A, Poisson LM, Rajendiran TM, Khan AP, Cao Q, Yu J, Laxman B, Mehra R, Lonigro RJ, Li Y, Nyati MK, Ahsan A, Kalyana-Sundaram S, Han B, Cao X, Byun J, Omenn GS, Ghosh D, Pennathur S, Alexander DC, Berger A, Shuster JR, Wei JT, Varambally S, Beecher C, Chinnaiyan AM. Metabolomic profiles delineate potential role for sarcosine in prostate cancer progression. *Nature*. 2009; 457(7231):910–4. [PubMed: 19212411]
20. Putluri N, Shojaie A, Vasu VT, Nalluri S, Vareed SK, Putluri V, Vivekanandan-Giri A, Byun J, Pennathur S, Sana TR, Fischer SM, Palapattu GS, Creighton CJ, Michailidis G, Sreekumar A. Metabolomic profiling reveals a role for androgen in activating amino acid metabolism and methylation in prostate cancer cells. *PLoS One*. 2011; 6(7):e21417. [PubMed: 21789170]
21. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, Paulovich A, Pomeroy SL, Golub TR, Lander ES, Mesirov JP. 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. [PubMed: 16199517]
22. Rhodes DR, Kalyana-Sundaram S, Mahavisno V, Varambally R, Yu J, Briggs BB, Barrette TR, Anstet MJ, Kincead-Beal C, Kulkarni P, Varambally S, Ghosh D, Chinnaiyan AM. Oncomine 3.0: genes, pathways, and networks in a collection of 18,000 cancer gene expression profiles. *Neoplasia*. 2007; 9(2):166–80. [PubMed: 17356713]
23. Glinksy GV, Glinkii AB, Stephenson AJ, Hoffman RM, Gerald WL. Gene expression profiling predicts clinical outcome of prostate cancer. *J Clin Invest*. 2004; 113(6):913–23. [PubMed: 15067324]
24. Sun Y, Goodison S. Optimizing molecular signatures for predicting prostate cancer recurrence. *Prostate*. 2009; 69(10):1119–27. [PubMed: 19343730]
25. Taylor BS, Schultz N, Hieronymus H, Gopalan A, Xiao Y, Carver BS, Arora VK, Kaushik P, Cerami E, Reva B, Antipin Y, Mitsiades N, Landers T, Dolgalev I, Major JE, Wilson M, Succi ND, Lash AE, Heguy A, Eastham JA, Scher HI, Reuter VE, Scardino PT, Sander C, Sawyers CL, Gerald WL. Integrative genomic profiling of human prostate cancer. *Cancer Cell*. 2010; 18(1):11–22. [PubMed: 20579941]
26. Belanger A, Pelletier G, Labrie F, Barbier O, Chouinard S. Inactivation of androgens by UDP-glucuronosyltransferase enzymes in humans. *Trends Endocrinol Metab*. 2003; 14(10):473–9. [PubMed: 14643063]
27. Mackenzie PI, Owens IS, Burchell B, Bock KW, Bairoch A, Belanger A, Fournel-Gigleux S, Green M, Hum DW, Iyanagi T, Lancet D, Louisot P, Magdalou J, Chowdhury JR, Ritter JK, Schachter H, Tephly TR, Tipton KF, Nebert DW. The UDP glycosyltransferase gene superfamily: recommended nomenclature update based on evolutionary divergence. *Pharmacogenetics*. 1997; 7(4):255–69. [PubMed: 9295054]
28. Rhodes DR, Yu J, Shanker K, Deshpande N, Varambally R, Ghosh D, Barrette T, Pandey A, Chinnaiyan AM. ONCOMINE: a cancer microarray database and integrated data-mining platform. *Neoplasia*. 2004; 6(1):1–6. [PubMed: 15068665]

29. Horoszewicz JS, Leong SS, Kawinski E, Karr JP, Rosenthal H, Chu TM, Mirand EA, Murphy GP. LNCaP model of human prostatic carcinoma. *Cancer Res.* 1983; 43(4):1809–18. [PubMed: 6831420]
30. Navone NM, Olive M, Ozen M, Davis R, Troncoso P, Tu SM, Johnston D, Pollack A, Pathak S, von Eschenbach AC, Logothetis CJ. Establishment of two human prostate cancer cell lines derived from a single bone metastasis. *Clin Cancer Res.* 1997; 3(12 Pt 1):2493–500. [PubMed: 9815652]
31. Culig Z, Hoffmann J, Erdel M, Eder IE, Hobisch A, Hittmair A, Bartsch G, Utermann G, Schneider MR, Parczyk K, Klocker H. Switch from antagonist to agonist of the androgen receptor bicalutamide is associated with prostate tumour progression in a new model system. *Br J Cancer.* 1999; 81(2):242–51. [PubMed: 10496349]
32. Wu HC, Hsieh JT, Gleave ME, Brown NM, Pathak S, Chung LW. Derivation of androgen-independent human LNCaP prostatic cancer cell sublines: role of bone stromal cells. *Int J Cancer.* 1994; 57(3):406–12. [PubMed: 8169003]
33. Zhang X, Morrissey C, Sun S, Ketchandji M, Nelson PS, True LD, Vakar-Lopez F, Vessella RL, Plymate SR. Androgen receptor variants occur frequently in castration resistant prostate cancer metastases. *PLoS One.* 2011; 6(11):e27970. [PubMed: 22114732]
34. Tomlins SA, Rhodes DR, Yu J, Varambally S, Mehra R, Perner S, Demichelis F, Helgeson BE, Laxman B, Morris DS, Cao Q, Cao X, Andren O, Fall K, Johnson L, Wei JT, Shah RB, Al-Ahmadie H, Eastham JA, Eggener SE, Fine SW, Hotakainen K, Stenman UH, Tsodikov A, Gerald WL, Lilja H, Reuter VE, Kantoff PW, Scardino PT, Rubin MA, Bjartell AS, Chinnaiyan AM. The role of SPINK1 in ETS rearrangement-negative prostate cancers. *Cancer Cell.* 2008; 13(6):519–28. [PubMed: 18538735]
35. Dillies MA, Rau A, Aubert J, Hennequet-Antier C, Jeanmougin M, Servant N, Keime C, Marot G, Castel D, Estelle J, Guernec G, Jagla B, Jouneau L, Laloe D, Le Gall C, Schaeffer B, Le Crom S, Guedj M, Jaffrezic F, on behalf of The French StatOmique, C. A comprehensive evaluation of normalization methods for Illumina high-throughput RNA sequencing data analysis. *Brief Bioinform.* 2012
36. Bao BY, Chuang BF, Wang Q, Sartor O, Balk SP, Brown M, Kantoff PW, Lee GS. Androgen receptor mediates the expression of UDP-glucuronosyltransferase 2 B15 and B17 genes. *Prostate.* 2008; 68(8):839–48. [PubMed: 18302198]
37. Paquet S, Fazli L, Grosse L, Verreault M, Tetu B, Rennie PS, Belanger A, Barbier O. Differential expression of the androgen-conjugating UGT2B15 and UGT2B17 enzymes in prostate tumor cells during cancer progression. *J Clin Endocrinol Metab.* 2012; 97(3):E428–32. [PubMed: 22170718]
38. Grasso CS, Wu YM, Robinson DR, Cao X, Dhanasekaran SM, Khan AP, Quist MJ, Jing X, Lonigro RJ, Brenner JC, Asangani IA, Ateeq B, Chun SY, Siddiqui J, Sam L, Anstett M, Mehra R, Prensner JR, Palanisamy N, Ryslik GA, Vandin F, Raphael BJ, Kunju LP, Rhodes DR, Pienta KJ, Chinnaiyan AM, Tomlins SA. The mutational landscape of lethal castration-resistant prostate cancer. *Nature.* 2012; 487(7406):239–43. [PubMed: 22722839]
39. Varambally S, Yu J, Laxman B, Rhodes DR, Mehra R, Tomlins SA, Shah RB, Chandran U, Monzon FA, Becich MJ, Wei JT, Pienta KJ, Ghosh D, Rubin MA, Chinnaiyan AM. Integrative genomic and proteomic analysis of prostate cancer reveals signatures of metastatic progression. *Cancer Cell.* 2005; 8(5):393–406. [PubMed: 16286247]
40. Vanaja DK, Cheville JC, Iturria SJ, Young CY. Transcriptional silencing of zinc finger protein 185 identified by expression profiling is associated with prostate cancer progression. *Cancer Res.* 2003; 63(14):3877–82. [PubMed: 12873976]
41. Imielinski M, Cha S, Rejtar T, Richardson EA, Karger BL, Sgroi DC. Integrated proteomic, transcriptomic, and biological network analysis of breast carcinoma reveals molecular features of tumorigenesis and clinical relapse. *Mol Cell Proteomics.* 2012; 11(6):M111 014910. [PubMed: 22240506]
42. Creighton CJ. Multiple oncogenic pathway signatures show coordinate expression patterns in human prostate tumors. *PLoS One.* 2008; 3(3):e1816. [PubMed: 18350153]
43. Ning QY, Wu JZ, Zang N, Liang J, Hu YL, Mo ZN. Key pathways involved in prostate cancer based on gene set enrichment analysis and meta analysis. *Genet Mol Res.* 2011; 10(4):3856–87. [PubMed: 22194210]

44. Crawford ED, Eisenberger MA, McLeod DG, Spaulding JT, Benson R, Dorr FA, Blumenstein BA, Davis MA, Goodman PJ. A controlled trial of leuprolide with and without flutamide in prostatic carcinoma. *N Engl J Med*. 1989; 321(7):419–24. [PubMed: 2503724]
45. Eisenberger MA, Blumenstein BA, Crawford ED, Miller G, McLeod DG, Loehrer PJ, Wilding G, Sears K, Culkin DJ, Thompson IM Jr, Bueschen AJ, Lowe BA. Bilateral orchiectomy with or without flutamide for metastatic prostate cancer. *N Engl J Med*. 1998; 339(15):1036–42. [PubMed: 9761805]
46. Taplin ME, Bubley GJ, Shuster TD, Frantz ME, Spooner AE, Ogata GK, Keer HN, Balk SP. Mutation of the androgen-receptor gene in metastatic androgen-independent prostate cancer. *N Engl J Med*. 1995; 332(21):1393–8. [PubMed: 7723794]
47. Koivisto P, Visakorpi T, Kallioniemi OP. Androgen receptor gene amplification: a novel molecular mechanism for endocrine therapy resistance in human prostate cancer. *Scand J Clin Lab Invest Suppl*. 1996; 226:57–63. [PubMed: 8981668]
48. Veldscholte J, Voorhorst-Ogink MM, Bolt-de Vries J, van Rooij HC, Trapman J, Mulder E. Unusual specificity of the androgen receptor in the human prostate tumor cell line LNCaP: high affinity for progestagenic and estrogenic steroids. *Biochim Biophys Acta*. 1990; 1052(1):187–94. [PubMed: 2322591]
49. Zhao XY, Malloy PJ, Krishnan AV, Swami S, Navone NM, Peehl DM, Feldman D. Glucocorticoids can promote androgen-independent growth of prostate cancer cells through a mutated androgen receptor. *Nat Med*. 2000; 6(6):703–6. [PubMed: 10835690]
50. Guo Z, Yang X, Sun F, Jiang R, Linn DE, Chen H, Chen H, Kong X, Melamed J, Tepper CG, Kung HJ, Brodie AM, Edwards J, Qiu Y. A novel androgen receptor splice variant is up-regulated during prostate cancer progression and promotes androgen depletion-resistant growth. *Cancer Res*. 2009; 69(6):2305–13. [PubMed: 19244107]
51. Hornberg E, Ylitalo EB, Crnalic S, Antti H, Stattin P, Widmark A, Bergh A, Wikstrom P. Expression of androgen receptor splice variants in prostate cancer bone metastases is associated with castration-resistance and short survival. *PLoS One*. 2011; 6(4):e19059. [PubMed: 21552559]
52. Culig Z, Comuzzi B, Steiner H, Bartsch G, Hobisch A. Expression and function of androgen receptor coactivators in prostate cancer. *J Steroid Biochem Mol Biol*. 2004; 92(4):265–71. [PubMed: 15663989]
53. Cai C, Wang H, Xu Y, Chen S, Balk SP. Reactivation of androgen receptor-regulated TMPRSS2:ERG gene expression in castration-resistant prostate cancer. *Cancer Res*. 2009; 69(15):6027–32. [PubMed: 19584279]
54. Decker KF, Zheng D, He Y, Bowman T, Edwards JR, Jia L. Persistent androgen receptor-mediated transcription in castration-resistant prostate cancer under androgen-deprived conditions. *Nucleic Acids Res*. 2012; 40(21):10765–79. [PubMed: 23019221]
55. Hu R, Dunn TA, Wei S, Isharwal S, Veltri RW, Humphreys E, Han M, Partin AW, Vessella RL, Isaacs WB, Bova GS, Luo J. Ligand-independent androgen receptor variants derived from splicing of cryptic exons signify hormone-refractory prostate cancer. *Cancer Res*. 2009; 69(1):16–22. [PubMed: 19117982]
56. Chouinard S, Barbier O, Belanger A. UDP-glucuronosyltransferase 2B15 (UGT2B15) and UGT2B17 enzymes are major determinants of the androgen response in prostate cancer LNCaP cells. *J Biol Chem*. 2007; 282(46):33466–74. [PubMed: 17848572]

Figure 1.

A) Outline of the experimental approach used to define the Castrate Resistant Prostate Cancer (CRPC)-associated metabolic signature. Metabolites were extracted from Androgen dependent (AD) and CRPC cell lines (n=3 each) and 150 metabolites were measured using liquid chromatography-coupled mass spectrometry, along with analysis of standards for monitoring the profiling process. Mass spectrometry data was examined using in-house statistical tools and used for Oncomine-Concept (OCM)-based enrichment analysis. The enriched bioprocesses were validated using Metabolic Phenotyping Microarrays (MPM) and their association with prostate cancer recurrence (a.k.a. biochemical recurrence or BCR) was assessed using a combination of Principal Component Analysis (PCA) and Cox-proportional hazard model, on Glinsky, Sun and Taylor prostate cancer gene expression datasets. The prognostic power of the CRPC-associated pathways was visualized using Kaplan Meier

(KM) plots. B) Outline of the steps involved in data analysis, OCM-based pathway mapping and evaluation of prognostic potential of CRPC-associated metabolic pathways.

Figure 2.

A) Heat map showing the steady state levels of 150 metabolites examined using Multiple Reaction Monitoring (MRM)-based targeted profiling of Androgen Dependent (AD) and Castrate Resistant (CRPC) Prostate Cancer Cell lines. Columns correspond to cell lines and rows correspond to metabolites. Shades of red represent elevated levels of metabolite and shades of green represent reduced levels of the metabolites (refer to scale for details). **B)** same as in A, but showing the list of metabolites that were commonly altered between AD and CRPC cells, across four different data pre-processing methods (termed Core CRPC-Associated Metabolic Signature or CCAMS). The heat map shows 38 metabolites that were altered at an FDR corrected P value of 0.2. Color scheme is similar to A.

Figure 3.

Network representation of bioprocesses enriched by the CCAMS shown in Figure 2B. In the network, each node represents a unique bioprocess, size of the node represents number of genes associated with the bioprocess, and the color of the node represents the different databases from which the bioprocess was derived. The top 10 bioprocesses with P -value $< 2E-05$ (Q -value < 0.002) are represented in the network. Nine of the enriched bioprocesses formed a highly interconnected cluster, represented by red bridges, and contained genes belonging to the uridine diphosphate-glucuronosyltransferase (UGT) family.

Figure 4.

A-C). are the Kaplan Meier plots, showing the prognostic value of UDP-glucuronosyltransferases, using the Glinsky, Sun and Taylor gene expression datasets for prostate cancer²³⁻²⁵. Time to biochemical recurrence (BCR) was used as the prognostic endpoint. **D-E).** Boxplots showing higher combined expression of UGT2B28, UGT2B15 and UGT2A1 in metastatic PCa compared to organ confined PCa in Grasso³⁸, Taylor²⁵, Varambally³⁹ and Vanaja⁴⁰ gene expression datasets. The combined expression of these genes were significantly higher in metastatic tissues, in 3/4 datasets examined.

Figure 5.

A-C) are plots showing the relative expression of UGT2B15, UGT2B28 and UGT2A1 measured using RNAseq, across 26 benign prostate (green), 93 localized (blue) and 148 metastatic (red) prostate cancer tissues. The Y-axis describes the log 2 values of FPKM (fragments per kilobase of exon per million fragments mapped) **D)** same as in A-C, but for combined median expression of UGT2B15, UGT2B28 and UGT2A1 in the prostate-derived tissue samples.

Table 1

List of metabolic pathways that showed significantly (FDR corrected P value < 0.05) altered activity, in AD vs CRPC cells, in metabolic phenotyping microarray. Pathways that are in bold red fonts were also determined to be altered in core CRPC associated metabolic signature (CCAMS) using OCM derived bioprocess enrichment.

Pathway	FDR Corrected Pvalue
Fructose and mannose metabolism	0.0005
Galactose metabolism	0.0005
Phenylalanine metabolism	0.0005
Glyoxylate and dicarboxylate metabolism	0.0007
Aminosugar metabolism	0.0008
ABC transporters	0.0014
Starch and sucrose metabolism	0.0014
Citrate cycle (TCA cycle)	0.0019
Pentose phosphate pathway	0.0019
Reductive carboxylate cycle (CO ₂ fixation)	0.0019
Alanine and aspartate metabolism	0.0034
Tyrosine and tryptophan biosynthesis	0.0034
Pyruvate metabolism	0.0035
Pentose and glucuronate interconversion	0.005
Glutamate metabolism	0.0064
Beta-alanine and D-alanine metabolism	0.0132
Pathways in cancer	0.0152
Renal cell carcinoma	0.0152
Phenyl alanine	0.0189
Butanoate metabolism	0.0372
Taurine and hypotaurine	0.0415
Glycolysis/ Gluconeogenesis	0.0445
Inositol phosphate metabolism	0.0445