

OPEN

# Integrated RNA and metabolite profiling of urine liquid biopsies for prostate cancer biomarker discovery

Bongyong Lee<sup>1</sup>, Iqbal Mahmud<sup>5</sup>, John Marchica<sup>1,4</sup>, Paweł Dereziński<sup>6</sup>, Feng Qi<sup>4</sup>, Fubo Wang<sup>7</sup>, Piyush Joshi<sup>1</sup>, Felipe Valerio<sup>8</sup>, Inoel Rivera<sup>8</sup>, Vipul Patel<sup>9</sup>, Christian P. Pavlovich<sup>3</sup>, Timothy J. Garrett<sup>5</sup>, Gary P. Schroth<sup>10</sup>, Yinghao Sun<sup>7</sup> & Ranjan J. Perera<sup>1,2,4\*</sup>

Sensitive and specific diagnostic and prognostic biomarkers for prostate cancer (PCa) are urgently needed. Urine samples are a non-invasive means to obtain abundant and readily accessible “liquid biopsies”. Herein we used urine liquid biopsies to identify and characterize a novel group of urine-enriched RNAs and metabolites in patients with PCa and normal individuals with or without benign prostatic disease. Differentially expressed RNAs were identified in urine samples by deep sequencing and metabolites in urine were measured by mass spectrometry. mRNA and metabolite profiles were distinct in patients with benign and malignant disease. Integrated analysis of urinary gene expression and metabolite signatures unveiled an aberrant glutamate metabolism and tricarboxylic acid (TCA) cycle node in prostate cancer-derived cells. Functional validation supported a role for glutamate metabolism and glutamate oxaloacetate transaminase 1 (GOT1)-dependent redox balance in PCa, which could be exploited for novel biomarkers and therapies. In this study, we discovered cancer-specific changes in urinary RNAs and metabolites, paving the way for the development of sensitive and specific urinary PCa diagnostic biomarkers either alone or in combination. Our methodology was based on single void urine samples (i.e., without prostatic massage). The integrated analysis of metabolomic and transcriptomic data from these liquid biopsies revealed a glutamate metabolism and tricarboxylic acid cycle node that was specific to prostate-derived cancer cells and cancer-specific metabolic changes in urine.

More than 180,000 men are diagnosed with prostate cancer (PCa) in U.S. in the 2016, where 26,000 of these patients will die of the disease<sup>1</sup>. PCa is one of the second most frequently diagnosed cancer deaths among men worldwide<sup>2</sup>. The radiotherapy and surgery for localized PCa are known to be effective, however the prognosis for patients with the progressive disease is poor. A test to detect PCa with high sensitivity and specificity at an early stage is a clinical imperative. Moreover, there is a vital need for novel therapeutic tactics to manage this insidious and prevalent disease.

<sup>1</sup>Cancer and Blood Disorders Institute, Johns Hopkins All Children’s Hospital, 600 6th Avenue South, St. Petersburg, FL, 33701, USA. <sup>2</sup>Department of Oncology, Sydney Kimmel Cancer Center, Johns Hopkins University School of Medicine, 401 N. Broadway, Baltimore, MD, 21287, USA. <sup>3</sup>The James Buchanan Brady Urological Institute, Department of Urology, The Johns Hopkins University School of Medicine, 4940 Eastern Avenue, Baltimore, MD, 21224, USA. <sup>4</sup>Sanford Burnham Prebys Medical Discovery Institute, 10901 N Torrey Pines Rd, La Jolla, CA, 92037, USA. <sup>5</sup>Department Pathology, Immunology and Laboratory Medicine, University of Florida, College of Medicine, 1395 Center Drive, Gainesville, FL, 32610, USA. <sup>6</sup>Department of Inorganic and Analytical Chemistry, Poznan University of Medical Sciences, Grunwaldzka 6 Street, 60-780, Poznan, Poland. <sup>7</sup>Department of Urology, Shanghai Changhai Hospital, Second Military Medical University, 168 Changhai Road, Shanghai, China. <sup>8</sup>Florida Urology Associates, 1812 N. Mills Avenue, Orlando, FL, 32803, USA. <sup>9</sup>Global Robotics Institute, 410 Celebration Place, Suite 200, Celebration, FL, 34747, USA. <sup>10</sup>Illumina, Inc., 5200 Illumina Way, San Diego, CA, 92122, USA. \*email: [jperera2@jhmi.edu](mailto:jperera2@jhmi.edu)

Serum prostate specific antigen (PSA) levels are been used for PCa diagnosis and screening for over thirty years, and digital rectal examination (DRE) for even longer<sup>3</sup>. However, PSA has modest sensitivity and specificity and does not discriminate indolent from aggressive cancers<sup>3,4</sup>. Prostate cancer antigen 3 (PCA3), a prostate-specific non-coding RNA, was approved by the FDA in 2012 as the first PCa molecular diagnostic test for a particular clinical indication (need for recurrence prostate biopsies in men aged >50 years with assumed PSA levels and/or DRE and/or one or more earlier negative biopsies)<sup>5</sup>. Nevertheless, the importance of the PCA3 test is limited by substantial individual variability, better performance in the replication biopsy setting, and contradictory data on the relationship between score and cancer grade using the most common edge of 35<sup>6</sup>. Hence, there is a dismal need for novel molecular diagnostic tools to more precisely detect and predict the performance of localized PCa.

The kidneys produce urine to eliminate soluble waste from the bloodstream. Urine is an abundant biofluid for molecular or cellular analyses and is useful in the diagnosis and management of bladder, ovarian, and kidney diseases<sup>7–9</sup>. Urine contains over 2500 metabolites<sup>10</sup> and provides a window through which to view cellular biochemical reactions and intermediary metabolism. The metabolite signature in urine will reflect the impact of gene regulation, enzyme activities, and alterations in metabolic reactions occurring in the different cell types found along the urogenital tract.

Cancer cells exhibit perturbed metabolism that enables proliferation and survival<sup>11</sup>. Therefore, metabolomic profiling has been a fruitful approach for the identification of early cancer biomarkers<sup>12,13</sup>. Furthermore, certain metabolic states are associated with prognosis in advanced cancers<sup>13</sup>. Several metabolomics studies have revealed PCa-specific metabolic phenotypes in serum, tissue, and urine<sup>14,15</sup>. Indeed, an intermediate metabolite of glycine synthesis and degradation, sarcosine, has been described as a putative PCa biomarker in urine<sup>16</sup>. Cancer cells exhibit perturbed metabolism that enables proliferation and survival<sup>11</sup>. Therefore, metabolomic profiling has been a fruitful approach for the identification of early cancer biomarkers<sup>12,13</sup>. Furthermore, certain metabolic states are associated with prognosis in advanced cancers<sup>13</sup>. Several metabolomics studies have revealed PCa-specific metabolic phenotypes in serum, tissue, and urine<sup>14,15</sup>. Indeed, an intermediate metabolite of glycine synthesis and degradation, sarcosine, has been described as a putative PCa biomarker in urine<sup>16</sup>. However, the utility of sarcosine as a biomarker is provocative and clinical validation has been indefinable<sup>17</sup>.

Metabolomics data have been integrated with comprehensive gene expression analyses to better interrogate complex gene and metabolic networks. Integrating multiple aspects of biological complexity using different unsupervised approaches can help to pinpoint the most important and reproducible pathways driving biological processes and hence reveal robust biomarkers or promising drug targets<sup>18,19</sup>. Here we performed metabolite profiling and high-throughput RNA sequencing obtained urine from benign prostatic hyperplasia, prostatitis, and PCa patients. Our aims were to (a) discover cancer-specific changes in the urine with utility as sensitive and specific PCa biomarkers either alone or in combination and (b) identify novel drug targets for PCa. Importantly, our approach used single void urine samples (i.e., without prostatic massage) as proof-of-principle of how a simple urine specimen can be used for biomarker and target discovery. Integrated analysis of metabolomic and transcriptomic data from these liquid biopsies revealed a glutamate metabolism and tricarboxylic acid cycle node that was specific to prostate-derived cancer cells and cancer-specific metabolic changes in urine. Functional validation *in vitro* provided mechanistic support for a pivotal role for GOT1-dependent glutamate metabolism in redox balance and cancer progression.

## Results

**Deep sequencing of urine-secreted mRNAs.** Normal voided urine from men contains small numbers of exfoliated cells from different parts of the urinary tract including urothelial cells, squamous cells, renal tubular cells, and glandular cells including prostate epithelial cells<sup>20</sup>. PCa cells are shed into urine and can be successfully isolated, processed, and analyzed by various molecular techniques<sup>3</sup>, thereby providing a rich substrate for biomarker detection. We sought to exploit this readily accessible and copious substrate from PCa patients for biomarker discovery and, in turn, elucidate novel mechanistic aspects of PCa.

Quality output from current next-generation sequencing (NGS) technology depends on the availability of high-quality RNA. An initial challenge was that the quality and quantity of RNA extracted from the very small number of exfoliated cells in urine was poor<sup>21</sup> (Supplementary Fig. 1). To overcome the problem, we performed sequence-specific capture (Illumina TruSeq RNA Access) with the urine samples to reduce ribosomal RNA and enrich for exonic RNA sequences. With this approach, we successfully sequenced 11 PCa (for clinical details, see Supplementary Table 1), 12 normal, and one pooled set of three normal samples (combined due to individually low RNA yields). The 3825 RNA transcripts that were detected in 20 samples readily but not perfectly segregated into normal and PCa groups (Supplementary Fig. 2). We concluded that RNA expression analysis of urine liquid biopsies by itself was unlikely to reveal sensitive and specific PCa biomarkers.

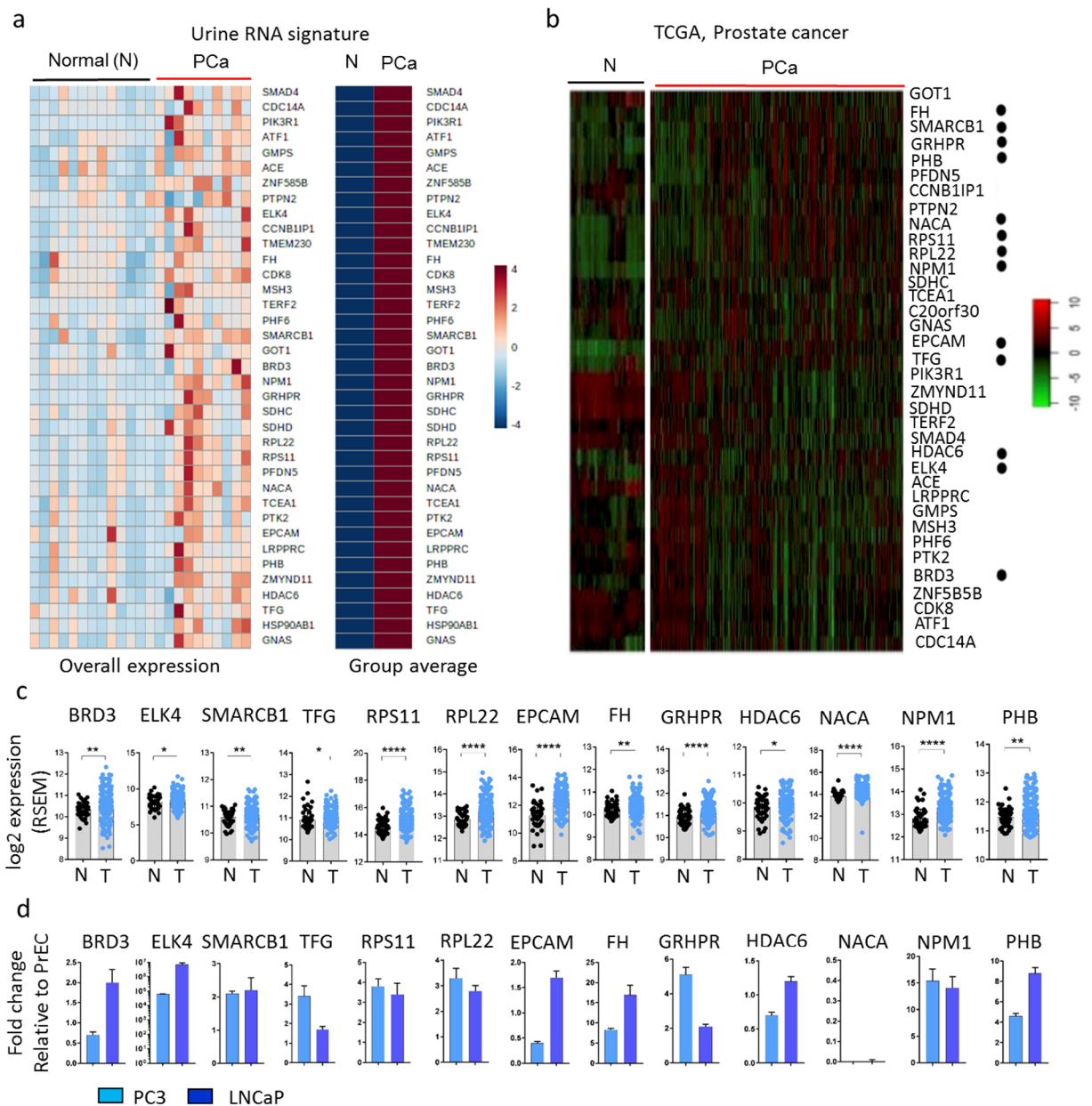
We next identified cancer-specific gene signatures. Among 5510 differentially expressed transcripts, 4662 had RPKM values greater than one, and 116 transcripts (110 genes) were significantly up- or downregulated in PCa (Table 1). Known PCa markers were upregulated in PCa urine (Supplementary Table 2), and differentially expressed genes were enriched for a number of important cancer pathways including PCa signaling, molecular mechanisms of cancer, PI3K/AKT signaling, and NF- $\kappa$ B signaling (Supplementary Table 3). To our knowledge, this is the first time that RNA-seq has been successfully applied to urine samples to profile coding genes.

Thirty-seven genes were significantly upregulated in PCa urine samples (Table 1 and Fig. 1a) compared to normal urine samples. To bolster confidence that the RNA originated from the patient's cancer rather than contaminating cells, we examined their expression in The Cancer Genome Atlas (TCGA) data (Fig. 1b,c and Supplementary Table 4). Of these 37 genes, 35% (13/37) were significantly upregulated in primary tumors compared to normal (Fig. 1b,c). Three of these genes were transcription factors (*ELK4*, *SMARCB1*, *BRD3*) and six were known oncogenes (*TFG*, *NACA*, *BRD3*, *ELK4*, *NPM1*, *RPL22*)<sup>22</sup>. When quantified in two representative PCa cell lines (LNCaP and PC3), most transcripts were upregulated in both cell lines compared to normal prostate

Gene	Wilcoxon p-value	log <sub>2</sub> fold-change	Gene	Wilcoxon p-value	log <sub>2</sub> fold-change	Gene	Wilcoxon p-value	log <sub>2</sub> fold-change
	Upregulated			Downregulated			Downregulated	
<i>ACE</i>	0.04	1.0	<i>ACSL3</i>	0.02	-1.3	<i>LMO7</i>	0.00	-3.0
<i>ATF1</i>	0.01	0.6	<i>ADM</i>	0.04	-2.5	<i>LRMP</i>	0.02	-3.2
<i>BRD3</i>	0.03	0.2	<i>BAZ2A</i>	0.01	-0.8	<i>MALAT1</i>	0.00	-3.2
<i>CCNB1IP1</i>	0.01	0.4	<i>BCL10</i>	0.00	-2.3	<i>MCL1</i>	0.03	-4.0
<i>CDC14A</i>	0.03	0.3	<i>BIRC3</i>	0.04	-4.1	<i>MDM2</i>	0.03	-1.2
<i>CDK8</i>	0.01	0.2	<i>BRWD3</i>	0.00	-1.5	<i>MYD88</i>	0.02	-4.2
<i>ELK4</i>	0.02	0.9	<i>CDKL5</i>	0.01	-1.1	<i>NCSTN</i>	0.02	-1.7
<i>EPCAM</i>	0.04	0.4	<i>CDKN1B</i>	0.02	-1.4	<i>NDE1</i>	0.02	-2.0
<i>FH</i>	0.04	0.0	<i>CHIC2</i>	0.01	-1.8	<i>NIPBL</i>	0.02	-1.3
<i>GMPS</i>	0.01	0.2	<i>CIC</i>	0.02	-3.9	<i>NOTCH1</i>	0.01	-2.9
<i>GNAS</i>	0.02	0.8	<i>CREBBP</i>	0.03	-1.2	<i>NT5C2</i>	0.02	-2.0
<i>GOT1</i>	0.02	0.5	<i>DDB2</i>	0.03	-1.3	<i>NUP98</i>	0.01	-2.3
<i>GRHPR</i>	0.00	1.3	<i>DOT1L</i>	0.01	-1.2	<i>PER1</i>	0.00	-1.9
<i>HDAC6</i>	0.04	0.6	<i>ELF4</i>	0.02	-1.8	<i>PICALM</i>	0.01	-3.0
<i>HSP90AB1</i>	0.04	0.6	<i>EP300</i>	0.01	-0.8	<i>PLCG2</i>	0.03	-1.6
<i>LRPPRC</i>	0.02	0.7	<i>EPHA2</i>	0.00	-3.2	<i>PRDM1</i>	0.00	-2.6
<i>MSH3</i>	0.00	0.1	<i>ERBB3</i>	0.04	-0.6	<i>PRSS8</i>	0.00	-3.3
<i>NACA</i>	0.03	1.1	<i>ERCC5</i>	0.03	-0.9	<i>PTK2B</i>	0.00	-2.1
<i>NPM1</i>	0.01	2.8	<i>ETV6</i>	0.04	-1.2	<i>RAF1</i>	0.04	-1.6
<i>PFDN5</i>	0.02	0.7	<i>FLCN</i>	0.00	-1.0	<i>RCOR1</i>	0.00	-1.3
<i>PHB</i>	0.01	0.3	<i>FOSB</i>	0.03	-2.1	<i>RNF213</i>	0.03	-1.5
<i>PHF6</i>	0.04	0.2	<i>GIT2</i>	0.04	-5.1	<i>SH3BP1</i>	0.01	-1.6
<i>PIK3R1</i>	0.00	1.4	<i>GNAQ</i>	0.02	-1.0	<i>SMAD3</i>	0.02	-1.4
<i>PTK2</i>	0.03	1.1	<i>HDAC4</i>	0.02	-1.3	<i>STRN</i>	0.02	-0.3
<i>PTPN2</i>	0.02	2.8	<i>HIPK1</i>	0.03	-1.6	<i>TMEM127</i>	0.02	-1.9
<i>RPL22</i>	0.01	0.9	<i>HIST1H2BO</i>	0.04	-0.9	<i>TMPRSS2</i>	0.02	-2.0
<i>RPS11</i>	0.01	0.9	<i>ICAM1</i>	0.03	-3.1	<i>TNFAIP3</i>	0.03	-2.2
<i>SDHC</i>	0.02	2.9	<i>IGF1R</i>	0.04	-1.1	<i>TOP1</i>	0.02	-1.3
<i>SDHD</i>	0.03	1.0	<i>IRF1</i>	0.03	-2.1	<i>TPM4</i>	0.02	-2.2
<i>SMAD4</i>	0.02	0.0	<i>JAK2</i>	0.04	-0.8	<i>TYK2</i>	0.02	-1.9
<i>SMARCB1</i>	0.00	0.3	<i>KDM5A</i>	0.04	-0.1	<i>VHL</i>	0.04	-1.8
<i>TCEA1</i>	0.03	0.1	<i>KDM5C</i>	0.01	-1.3	<i>WDFY3</i>	0.04	-1.0
<i>TERF2</i>	0.02	0.3	<i>KIAA0232</i>	0.01	-1.0	<i>WDR1</i>	0.04	-1.7
<i>TFG</i>	0.03	0.1	<i>KMT2B</i>	0.03	-0.7	<i>WSB1</i>	0.03	-1.7
<i>TMEM230</i>	0.02	0.2	<i>KMT2C</i>	0.03	-0.9	<i>XIAP</i>	0.00	-1.4
<i>ZMYND11</i>	0.00	3.1	<i>KMT2D</i>	0.02	-0.9	<i>ZNF217</i>	0.00	-1.1
<i>ZNF585B</i>	0.02	0.4	<i>LAMP2</i>	0.02	-2.9			

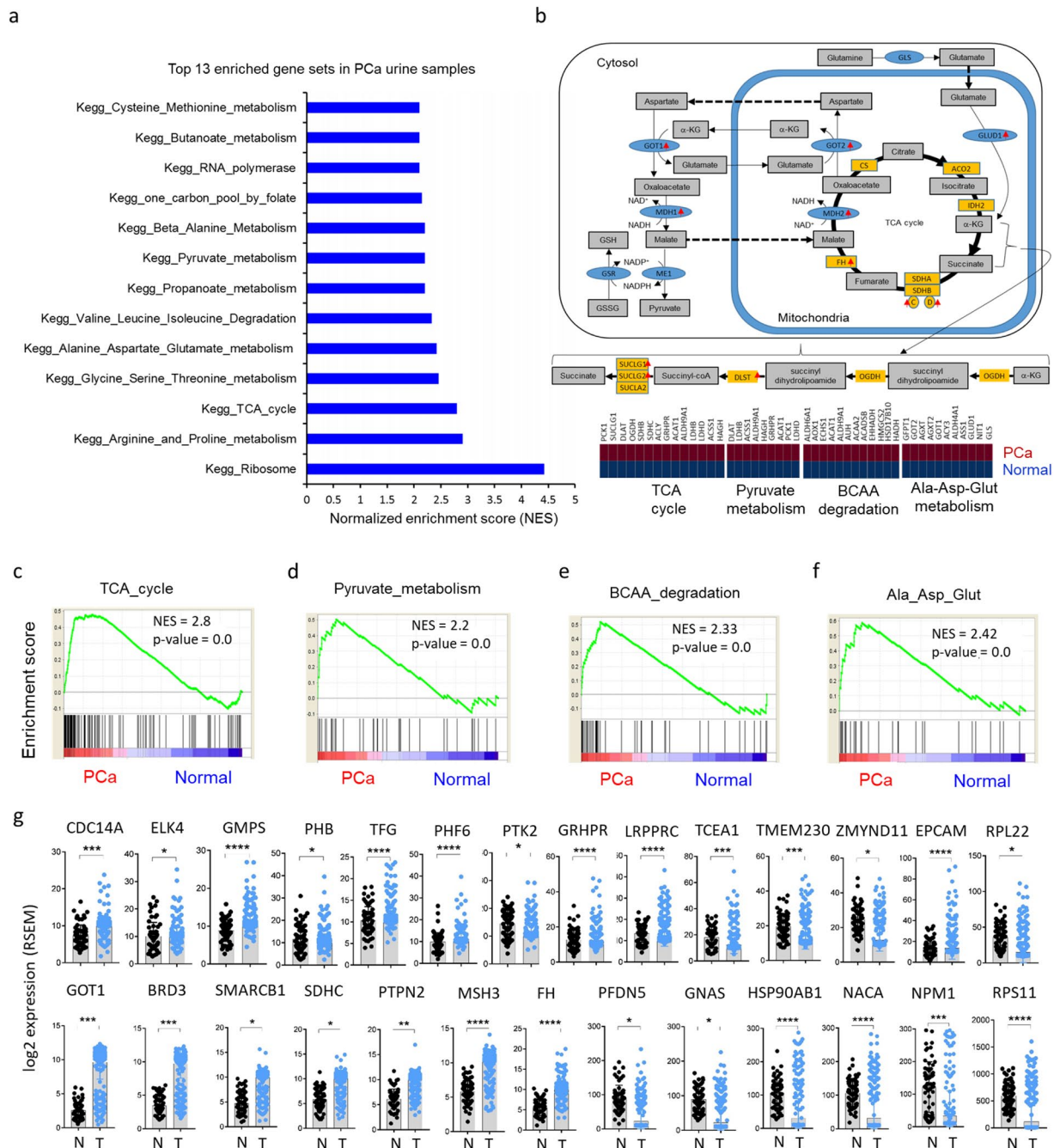
**Table 1.** Differentially expressed genes in prostate cancer urine samples.

epithelial cells (PrEC) except for *NACA* (downregulated in both cell lines), *BRD3* and *EPCAM* (decreased in PC3 cells), and *HDAC6* (downregulated in LNCaP cells) (Fig. 1d). Gene set enrichment analysis (GSEA) revealed thirteen overrepresented pathways in PCa urine compared to normal urine samples based on normalized gene set enrichment scores (NES) (Fig. 2a). Among them, pyruvate metabolism, TCA cycle, branched chain amino acid (BCAA) degradation pathway, and alanine, aspartate, and glutamate metabolism were significantly enriched (Fig. 2b–f). It is well established that tumors predominantly reprogram metabolism in demand to satisfy three major functions: bioenergetics (central carbon metabolism), biosynthesis (the production of biomaterials such as nucleotides, amino acids, and lipids), and redox balance (chemistry of oxidation-reduction homeostasis status in metabolism)<sup>23,24</sup>. Together, these processes support cancer cell maintenance, proliferation, tumor initiation, growth, and progression and are considered central cancer metabolism dogma that allows cancer cells to adapt to changing cellular and physiological conditions in the context of growth-related signaling and stress. Key carbon metabolism pathways such as pyruvate metabolism and the TCA cycle, which are the source of energy; different amino acids pathways critical for the production of macromolecules such as protein, nucleotides, and lipids; and alanine, aspartate, and glutamate metabolism, which are essential for redox balance, were significantly upregulated in urine secreted PCa cells (Fig. 2b). These urinary cell transcriptional profiles are highly consistent with the principle of altered cancer metabolism and current knowledge of prostate cancer disease progression as determined in tissues and other biofluids<sup>16,25</sup>. Taken together, our data suggest that the transcriptional profiles generated from cells residing in urine from PCa patients are likely to originate from cancerous prostate epithelial cells rather than other urinary tract contaminants.



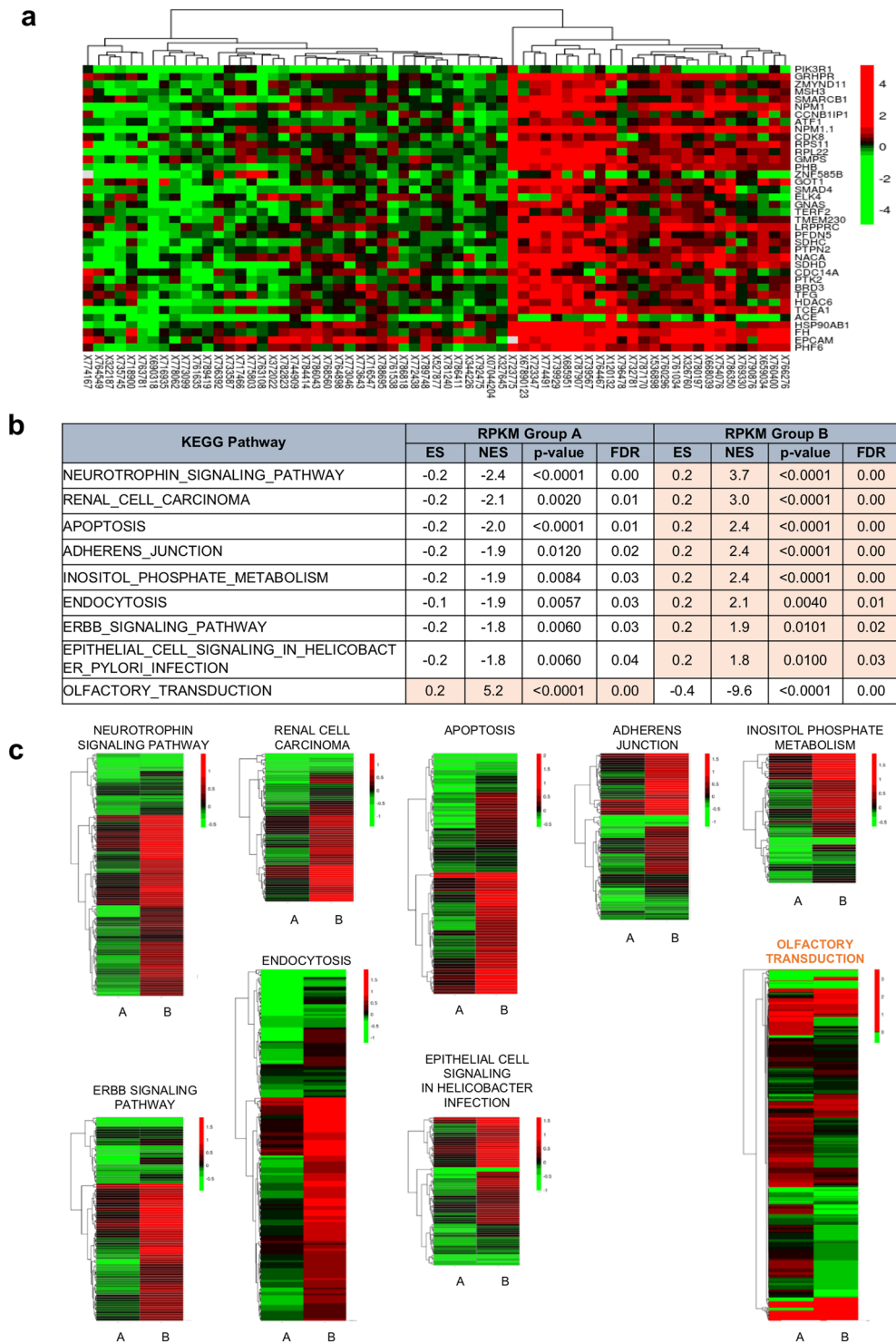
**Figure 1.** RNA-seq of cells extracted from the urine of patients with ( $n = 8$ ) and without ( $n = 12$ ) prostate cancer and validation of their differential expression in The Cancer Genome Atlas (TCGA) data. **(a)** A heatmap showing expression of 37 significantly upregulated genes in PCa cancer liquid biopsies. **(b,c)** Among 37 upregulated genes, 13 genes (*FH*, *SMARCB1*, *GRHPR*, *PHB*, *NACA*, *RPS11*, *RPL22*, *NPM1*, *EPCAM*, *TFG*, *HDAC6*, *ELK4*, *BRD3*) were significantly upregulated in primary tumors ( $n = 497$ ) compared to normal ( $n = 52$ ) in TCGA data. In the heatmap, black dots next to the gene name mark the genes upregulated in primary tumors compared to normal in TCGA data. The TCGA project for PCa data is publicly available for download at <https://portal.gdc.cancer.gov/projects/TCGA-PRAD>. **(d)** These 13 genes were also tested in two prostate cancer cell lines (LNCaP and PC3), and most were overexpressed apart from *NACA*, which was downregulated in both cell lines; *BRD3* and *EPCAM*, which were decreased in PC3 cells; and *HDAC6*, which was downregulated in LNCaP cells.

**Validation of urine gene signatures in tumor tissue.** Since exfoliated cells in urine represent a mixture of cell types, we next established that the gene expression profiles of urine-exfoliated cells represented expression in prostate tissue using RNA-seq data derived from 65 PCa and matched normal prostate tissues<sup>25</sup>. Most of the 110 up- or downregulated genes in urine exfoliated cells (Table 1) agreed with tissue gene expression (Fig. 2g and Supplementary Table 5). Of 37 upregulated genes, 34 genes were upregulated in PCa tissue, and 27 of 34 genes (79%) were significantly upregulated (Fig. 2g). For the downregulated genes, 46 out of 73 genes (60%) were downregulated in PCa tissue compared to normal tissue (Supplementary Table 5). These results suggest that the PCa gene signature detected in urine exfoliated cells represents a *bone fide* PCa signature, especially with regard to upregulated genes, which may therefore represent more robust biomarkers.



**Figure 2.** Gene set enrichment analysis (GSEA) of urine transcriptomes and validation of the expression of 37 genes in prostate tumor tissues. **(a)** Normalized Enrichment Score plot of the top 13 pathways in PCa urine. **(b)** Schematic view of TCA cycle and glutamate metabolism. Red arrows indicate upregulated genes in PCa urine compared to normal urine samples. **(c–f)** GSEA Plot for the TCA cycle, pyruvate metabolism, BCAA degradation, and alanine, aspartate, and glutamate metabolism, respectively. GSEA was conducted using GSEA software from the Broad Institute (<http://software.broadinstitute.org/gsea/index.jsp>). **(g)** The 37 gene signature from PCa patient urine RNA was confirmed in 65 pairs of tumors and adjacent normal tissue by RNA-seq data. Thirty four out of 37 genes were upregulated in PCa tumor tissue. Among them, 27 genes were significantly upregulated (Student *t*-test, \**p* < 0.05, \*\**p* < 0.01, \*\*\**p* < 0.001, \*\*\*\**p* < 0.0001).

Thus, we performed principal component analysis (PCA) of the 37 upregulated genes in the 65 patient tissue RNA-seq data. This 37-gene signature divided the tumor samples into two distinct groups, A and B (Fig. 3a), which did not differ with respect to Gleason score, tumor stage, or metastasis status (Supplementary Fig. 3). However, the two groups did show significant differences in *PCA3* and *KLK3* expression (Supplementary Fig. 4), which were both significantly higher in group B than in group A patients (Supplementary Fig. 4).



**Figure 3.** The urine 37 gene signature distinguishes PCa patients into two groups. **(a)** Unsupervised clustering of log-fold change (logFC) values shows that patients can be grouped into two groups, Group A and Group B, exhibiting lower and higher expression of the 37 genes compared to the matched normal, respectively. The logFC was obtained from the log<sub>2</sub> ratio of RPKM values of tumor and matched normal of the listed genes. **(b,c)** KEGG pathway analysis suggests that the two molecularly identified groups are physiologically distinct, as shown by opposite enrichment of the depicted pathways. The logFC values were obtained by comparing tumor versus matched normal samples from each group via paired limma analysis.

With respect to pathway differences between groups, nine pathways were significantly different: eight pathways were significantly upregulated (FDR > 0.05) in group B, and one pathway, olfactory transduction, was upregulated in group A (Fig. 3b,c). Therefore, the 37-gene signature in urine samples represents prostate tissue

gene expression and might be useful to distinguish advanced PCa (higher *PCA3* and *KLK3* levels in cancer) as well as to detect PCa itself.

**Metabolomic profiling of urine from normal subjects and patients with diseased prostates.** With this in mind, we performed global metabolite profiling of urine from patients with normal prostates, BPH, PTT, and PCa to discover cancer-specific metabolic changes. In global metabolite profiling, the metabolic profiles of urine specimens from normal subjects and patients with cancer were distinct and separate by PCA (Fig. 4a,b), whilst there was significant overlap between the profiles obtained from patients with BPH and PTT (Fig. 4c). Positive and negative ion data were first normalized to the specific gravity and then normalized to the total ion signal for all subsequent statistical analyses (Supplementary Fig. 5a). Positive and negative ion data sets were treated separately, and initial analysis was performed with PCA. The negative ion data set separation by PCA was very distinct between PCa and control groups, with BPH and PTT clustering together but as a separate cluster from PCa and control. Separation was primarily observed along PC1 (Fig. 4a). However, no correlation was observed between PSA scores and metabolic profiling. This result was reproducible with a second set of urine samples (Supplementary Fig. 5b). Further, subsequent hierarchical clustering-based heatmap analysis revealed distinctly higher abundance of global metabolites in PCa urine samples compared to normal samples (Supplementary Fig. 5c); over 180 metabolites in positive ion mode and 140 metabolites in negative ion mode were detected from the extracted urine samples, respectively.

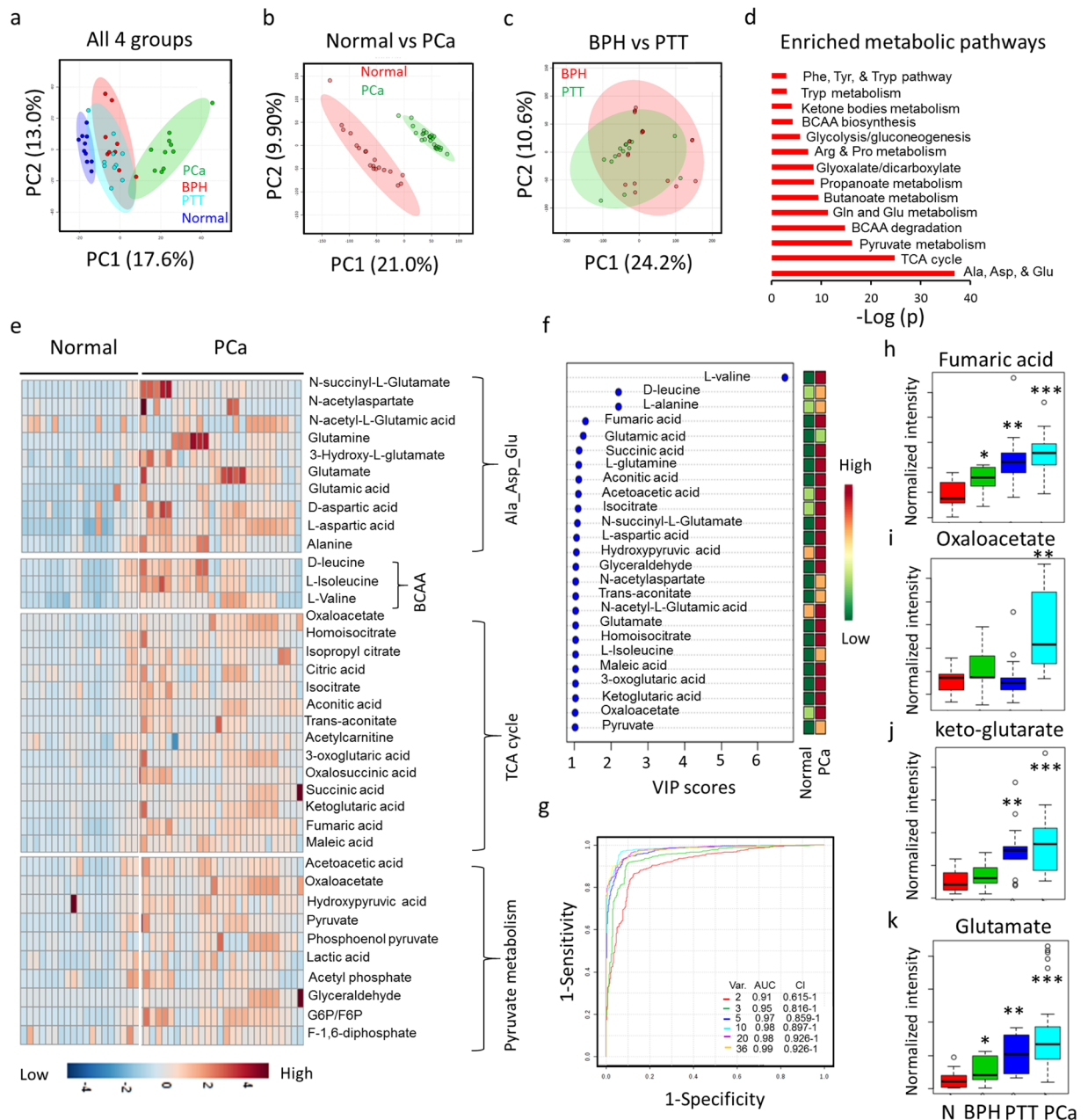
Comprehensive metabolic networks in PCa urine samples have not been well studied, and understanding differential metabolic pathway utilization in PCa might contribute towards the development of robust biomarkers. Therefore, we performed metabolite-based pathway enrichment analysis (Supplementary Fig. 5d), which revealed 14 significantly impacted metabolic pathways in the urine metabolome (Fig. 4d). Notably, alanine, aspartate, and glutamate metabolism, TCA cycle metabolism, pyruvate metabolism, and several amino acid pathway metabolites were identified at higher levels in PCa urine samples compared to normal (Fig. 4e), consistent with previous tissue-based studies<sup>16,25</sup>.

We next conducted metabolite-based urinary biomarker screening, with metabolite contribution assessed by examining the variable importance in projection (VIP) score, which is calculated from the weighted sum of the square for each partial least square design (PLS) loadings for each principal component. Of the top twenty five variables identified by VIP scores, all were metabolite variables that significantly contributed to the class separation of normal and PCa samples (Fig. 4f). We then conducted multivariate receiver operating characteristics (ROC) curve-based exploratory biomarker analysis to identify a diagnostic PCa-related metabolite signature. To better predict PCa, the top 50 discriminatory metabolites were identified via logistic regression (Fig. 4g). A combination of six metabolites showed better discrimination ( $AUC > 98\%$ ) than each metabolite individually ( $AUC < 91\%$ ) (Fig. 4g): aconitic acid ( $AUC = 0.97$ ), succinic acid ( $0.96$ ), fumaric acid ( $AUC = 0.955$ ), oxaloacetate ( $AUC = 0.952$ ),  $\alpha$ -ketoglutaric acid ( $AUC = 0.921$ ), and glutamate ( $AUC = 0.951$ ) (Supplementary Fig. 6). Subsequent data analysis identified significantly higher levels of TCA cycle metabolites (fumaric acid, oxaloacetate, and keto-glutaric acid) (Fig. 4h–j) and glutamate metabolism (glutamate) pathway (Fig. 4k) metabolites in PCa urine samples compared to urine samples from healthy and inflammatory prostates such as BPH and PTT.

Given that Gleason score (GS) status correlates with PCa tumor progression, an additional independent 11 normal, 11 GS-6, 11 GS-7, 11 GS-8, and 11 GS-9 PCa urine samples were collected and again subjected to global metabolomics analysis using mass spectrometry (Fig. 5). Metabolite data were analyzed using same statistical approaches as in Fig. 4. Furthermore, metabolome data accuracy was validated by a partial least squares-discriminant analysis (PLS-DA)-based Q2 model (Fig. 5a). Supervised multivariate statistical analysis of the global metabolome revealed a profound trend of clustering with respect to the four different GS groups and the normal urine samples (Fig. 5b). Hierarchical clustering heatmap analysis revealed the top 25 distinct metabolite signatures among normal and different GS groups (Fig. 5c). Metabolic pathway analysis between normal versus individual GS groups identified several potentially unique and common metabolic networks (Supplementary Fig. 7a–d). Differential metabolites from four GS groups and integration of their corresponding biochemical pathway analysis revealed twelve metabolic networks that were commonly present and upregulated with increased GS score (Fig. 5d). Among these pathways, TCA cycle and glutamate metabolism pathways were the pathways most upregulated with PCa disease progression (Fig. 5e). Notably, levels of TCA cycle and glutamate metabolism pathway metabolites were significantly increased with the highest GS, the most aggressive form of PCa (Supplementary Fig. 7e–j).

**Integrated gene expression and metabolite analysis.** We reasoned that integrating changes in gene expression and metabolite levels evident in the urine samples would better reveal the key pathways driving PCa and hence pinpoint the most robust biomarkers. The integrated pathway analysis module of MetaboAnalyst<sup>26</sup> was used to map both genes and metabolites to KEGG pathways to determine not just overrepresented pathways but also the relative importance of the genes and compounds based on their relative locations (topology). The top three pathways most significantly enriched for differentially expressed genes and metabolites were: aminoacyl-tRNA biosynthesis; Ala, Asp, and Glu metabolism; and the TCA cycle ( $p < 0.001$ ; Fig. 6a and Supplementary Table 6). Aminoacyl-tRNA biosynthesis probably represents an increase in global protein translation and demand for protein synthesis in cancer cells<sup>27</sup>. However, Ala, Asp, and Glu metabolism and the TCA cycle are closely related pathways that are critical for energy generation and carbon and nitrogen metabolism for biomass accumulation<sup>28</sup>, especially in rapidly dividing cells such as cancer cells.

Analysis of the top 25 significant metabolites identified by UHPLC-HRMS and their corresponding genes from RNA-seq revealed that GOT1- and GOT2-mediated metabolism was the main gene-metabolite interactive node influencing Ala, Asp, and Glu metabolism and the TCA cycle metabolism (Fig. 6b). In several PCa

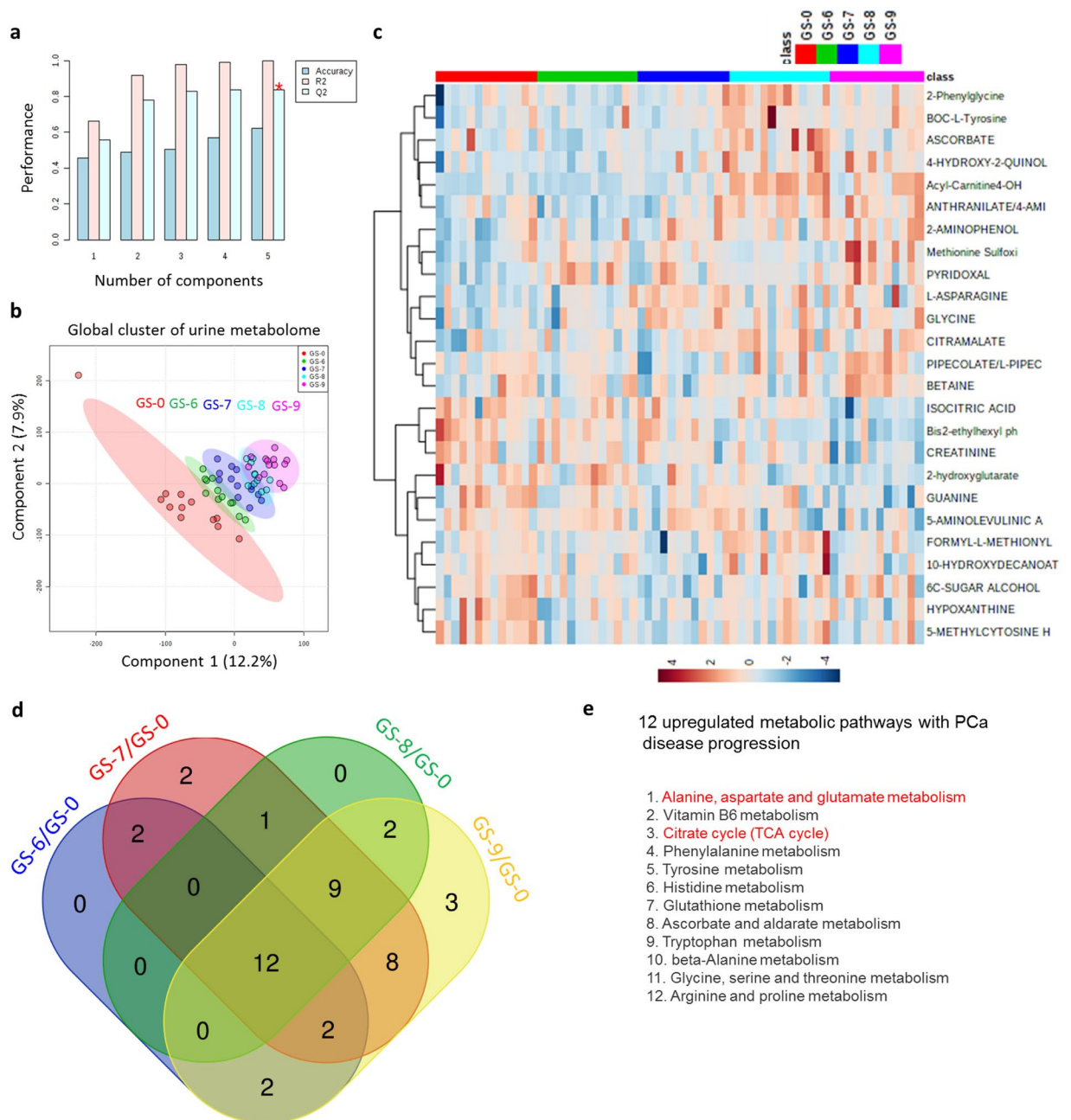


**Figure 4.** Global untargeted urine metabolomics profile data reveals distinct metabolic differences between PCa and normal. **(a)** Multivariate principal component analysis (PCA) scores plotted for normal ( $n = 20$ ), BPH ( $n = 20$ ), PTT ( $n = 11$ ), and PCa ( $n = 20$ ). **(b)** PCA score plot for BPH and PTT. **(c)** PCA score plot for normal and PCa. **(d)** List of significant metabolic pathways in PCa urine samples. The p-value is shown in negative  $\log_{10}$  scale. **(e)** Heat map of the differential metabolites in PCa tissue. Color bar indicates the relative abundance of the metabolites, with red indicating a higher concentration and blue indicating a lower concentration. Ala\_asp\_Glu, alanine, aspartate, and glutamate metabolism; BCAA, branched chain amino acid metabolism, G6P, glucose-6-phosphate; F6P, fructose-6-phosphate.  $P < 0.05$  was considered significant. **(f)** Metabolites ranked by their contributions and shown as variable importance in the projection (VIP) scores. **(g)** ROC curves for the predictive model. Shown as combination metabolite models calculated from the logistic regression analysis. **(h–k)** Significantly higher levels of metabolites in the TCA cycle (**h,i,j**) and alanine, aspartate, and glutamate metabolism (**k**) pathways were found in the PCa urine samples compare to normal, BPH, and PTT.

clinical datasets, GOT1 expression was significantly higher in malignant prostate compared to normal (Fig. 6c–e), whereas GOT2 expression showed no significant differences (data not shown).

Interestingly, *GOT1* expression was significantly elevated in high GS tumors compared to controls (Fig. 6f). *GOT1* expression was significantly elevated in GS 7 tumors [both (3 + 4) and (4 + 3)] compared to normal. Furthermore, we identified aberrant levels of glutamate, the product of GOT1, in primary and metastatic prostate cancers compared to normal (Fig. 6g,h). We also evaluated the ROC curves for glutamate levels in PCa urine

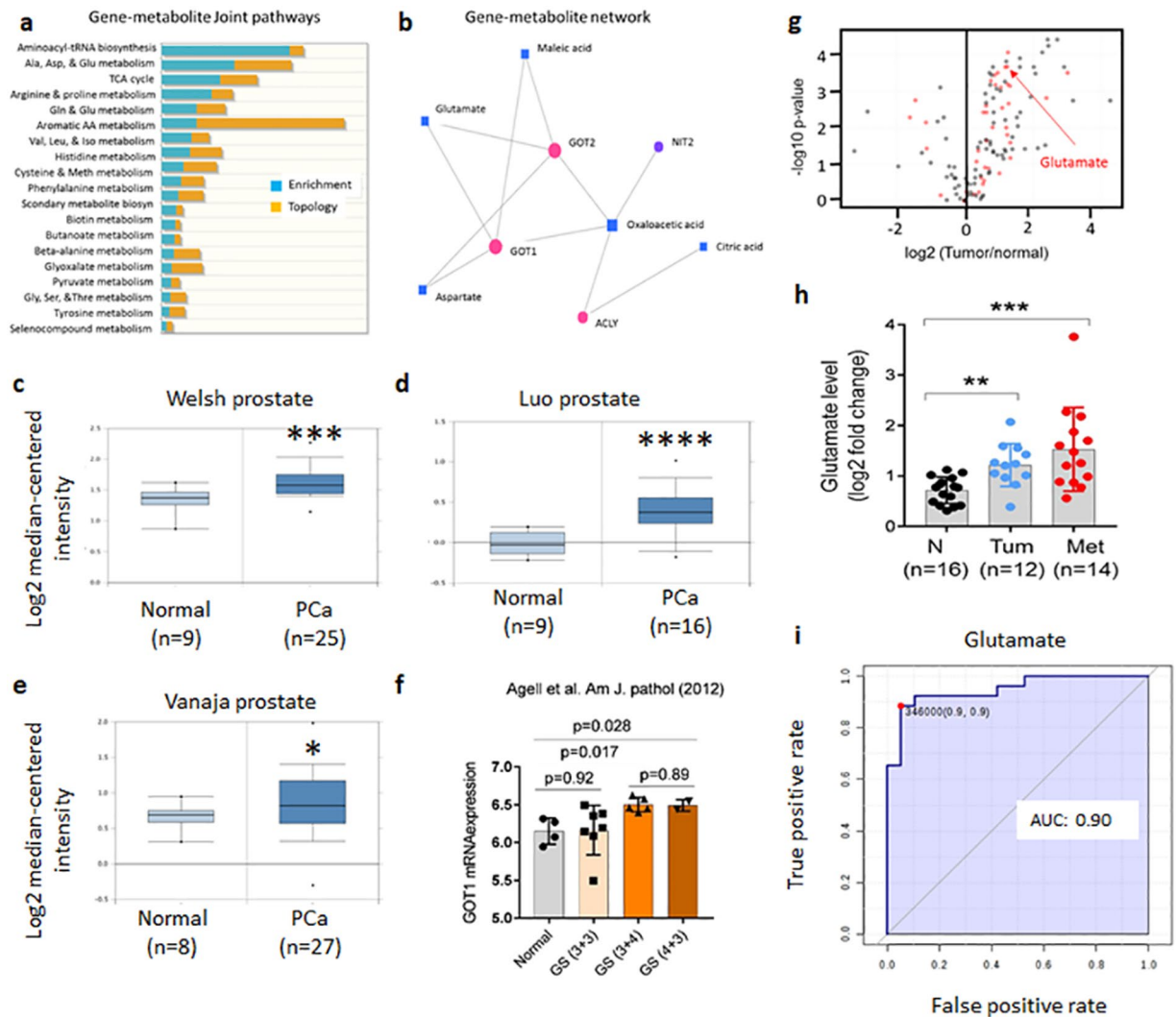




**Figure 5.** Correlation of TCA cycle and glutamate metabolism pathway metabolite levels with PCa Gleason score (GS). **(a)** Validation of PLS-DA (partial least square-differential analysis) model. Q2 indicates the accuracy of the model is over 80% when it includes the top five components. **(b)** Plot of PLS-DA component 1 and component 2 for log<sub>2</sub> autoscaled metabolite abundance data for normal versus different PCa GS. Circles indicate the 95% confidence intervals of each sample group. **(c)** Heatmap showing the autoscaled abundances of the top 25 untargeted metabolites in normal versus different PCa GS-based urine samples. Data were analyzed using the Euclidean distance calculation and ward clustering algorithm. **(d,e)** Venn diagram represents common and unique metabolic pathways that were significantly upregulated in PCa disease progression. **(f)** List of the top 12 significant metabolic pathways that were commonly screened as upregulated in different GS of PCa.

specimens normalized to BPH/PTT, and glutamate had an AUC of 90% (Fig. 6i). Therefore, GOT1-mediated glutamate metabolism might be critical for PCa disease progression, and a better understanding of GOT1-driven metabolism could reveal potential drug targets and biomarkers for PCa.

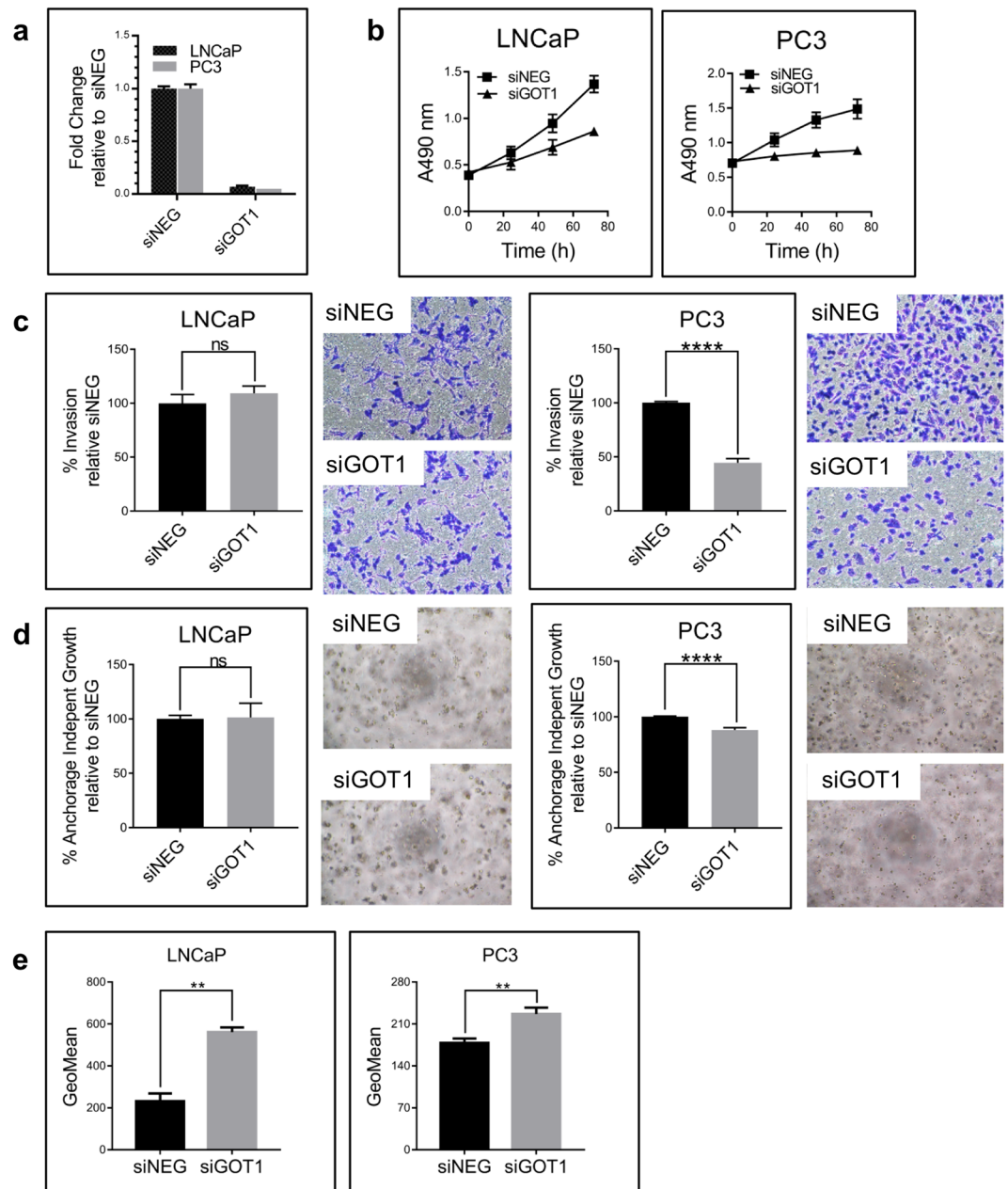
**Glutamate metabolism contributes to the cancerous phenotype via GOT1-mediated redox balance.** GOT1, a cytosolic transaminase that converts Asp to Glu, and other genes involved in glutamine metabolism such as *GLUD1*, *GOT1*, *GOT2*, and *MDH1* were significantly upregulated in PCa urine samples (Fig. 2b and



**Figure 6.** Integrated metabolic pathway enrichment analysis. **(a)** Gene-metabolite joint pathways were identified using the MetaboAnalyst integrated pathway analysis module. **(b)** Gene-metabolite network exploration identified a GOT1- and GOT2-mediated interactive network that might influence the TCA cycle and other metabolic pathways. **(c–e)** Box plots of GOT1 expression in malignant prostate tumor compared to normal (data sources <https://www.oncomine.org/resource/login.html>). **(f)** A bar graph of GOT1 expression across the prostate cancer Gleason score (GS) (data sources GSE30521). **(g–i)** A volcano plots of the differential metabolites **(g)**, bar graph of glutamate level among normal, tumor, and metastatic clinical specimens of Pan-Cancer Metabolism Data Explorer **(h)**, and ROC curve for glutamate level prediction in PCa urine samples compared to normal **(i)**. Statistical significance was analyzed by Student's t-test.

Supplementary Table 7). To investigate GOT1's role as a regulatory metabolic node in prostate cancer, we knocked down *GOT1* in the prostate cancer cell lines LNCaP and PC3 using siRNA (Fig. 7a). As expected, *GOT1* knockdown upregulated the upstream metabolites (Supplementary Fig. 8) Glu [1.2-fold (LNCaP;  $p = 0.01$ ) and 1.4-fold (PC3;  $p = 0.03$ )] and Asp [1.5-fold (LNCaP;  $p = 0.0004$ ) and 2.6-fold (PC3;  $p = 0.0006$ )] in both cell lines. *GOT1* knockdown significantly decreased the viability of both LNCaP and PC3 cells (Fig. 7b), consistent with previous reports that GOT1 repression suppresses tumor growth<sup>29,30</sup>, and the invasiveness and colony-forming ability of PC3 cells (Fig. 7c,d). We therefore examined the mechanism by which *GOT1* regulated prostate cancer cell viability.

Maintaining NAD/NADH balance supports *de novo* aspartic acid (Asp) biosynthesis and is required for proliferation<sup>31,32</sup>. Since GOT1 is part of the malate-Asp shuttle<sup>31</sup>, we checked whether *GOT1* knockdown affected the NAD/NADH ratio; the NAD/NADH ratio was indeed decreased (Supplementary Fig. 9), suggesting that this reduction may have influenced the cell proliferation inhibition in both LNCaP and PC3 cells. As depicted in Fig. 2b, GOT1 is necessary to convert Asp derived from the Gln TCA cycle into oxaloacetate and malate to produce NADPH<sup>33</sup>, which is essential for maintaining intracellular redox balance via detoxification of damaging reactive oxygen species (ROS). Both LNCaP and PC3 cells showed increased ROS levels upon *GOT1* knockdown (Fig. 7e), suggesting that GOT1 plays a role in cellular redox balance and can be manipulated to reduce the viability of prostate cancer cells.



**Figure 7.** GOT1 supports proliferation, invasion, and colony formation in prostate cancer cell lines. **(a)** GOT1 knockdown in LNCaP and PC3 prostate cancer cell lines. **(b)** GOT1 knockdown significantly inhibits cell viability in PC3 and LNCaP cells. **(c)** Invasion and **(d)** anchorage-independent growth in prostate cancer cell lines upon GOT1 knockdown. **(e)** GOT1 knockdown significantly increases ROS production in PC3 and LNCaP cells. The data from three independent experiments were expressed as mean  $\pm$  SD.

## Discussion

Recent advances in transcriptomics and metabolomics have led to the identification of various candidate biomarkers for cancer diagnosis and prognosis<sup>16,34</sup>. However, biomarkers derived from one dataset may not be reliable, and reproducibility in independent cohorts is challenging<sup>35</sup>.

Prostate carcinogenesis is known to involve metabolic reprogramming to provide sufficient energy for rapid cellular proliferation<sup>28,36</sup>. Many cancer cells exhibit augmented aerobic glycolysis, that previously reported as the Warburg effect, even in high-oxygen environments<sup>37</sup>. This metabolic adaptation helps to provide essential cellular components such as lipids and nucleotides to support the anabolic needs of rapidly proliferating tumor cells. Beyond the Warburg effect, the TCA cycle and oxidative phosphorylation are known to play important roles in PCa<sup>36,38</sup>. Prostate epithelial cells normally produce certain components of prostatic fluid such as citrate, PSA, and polyamines<sup>38</sup>. Increased citrate production by prostate cells means that they favor citrate synthesis over citrate utilization. However, PCa cells degrade citrate and accumulate oxidized citrate, resulting in more efficient energy production<sup>36,39</sup>.

There are significant advantages to using biofluids including blood, urine, saliva, and seminal plasma as sources of biomarkers<sup>40,41</sup>. Among them, urine is a promising liquid biopsy as it is noninvasive, replenishable, and convenient to collect. Urine has been at the center of clinical proteomics and provided biomarkers for renal disease<sup>42</sup>, renal cell carcinoma<sup>43</sup>, bladder cancer<sup>44</sup>, and prostate cancer<sup>45</sup>. In addition to proteins and peptides, urine contains various nucleic acids, metabolites, and lipids. Recently, the long noncoding RNA *PCA3* and the fusion gene *TMPRSS2:ERG* have been proposed as urinary PCa biomarkers<sup>46</sup>. Here we report for the first time a global transcriptomic profile of PCa in urine. We applied capture-based enrichment (RNA Access protocol), in which probes target exonic regions, and were able to separate PCa samples from normal healthy individual samples by unsupervised methods (Supplementary Fig. 2). Targeted or global strategies have been used to profile metabolites in urine samples and identify PCa biomarkers<sup>14,16,47,48</sup>, but results have been highly variable<sup>49</sup>. In the first unbiased metabolomics study measuring 1126 metabolites in 262 clinical samples including 110 urine samples, the glycine derivative sarcosine was elevated in PCa tissue and urine from PCa patients, and functional validation of the oncogenic role of sarcosine was provided *in vitro*<sup>16</sup>. However, sarcosine was not a reproducible prognostic marker in independent cohorts<sup>15,17</sup>, a common finding in single-biomarker studies that possess neither the specificity nor sensitivity for clinical development<sup>50</sup>.

We have shown that urine global metabolite profiling could distinguish PCa patient samples from normal prostates, BPH, and PTT (Fig. 4). Furthermore, global metabolomics analysis was able to distinguish the four different GS groups and the normal urine samples (Fig. 5). Recently, due to the limitations of Gleason grading, a new PCa grading system proposed by Epstein and colleagues was introduced, which assigns the tumor histologic patterns 1 through 5<sup>51</sup>. Notably, the new system distinguishes Gleason score 3 + 4 = 7 (Group 2) and 4 + 3 = 7 (Group 3), which are prognostically very different. Due to the limited patient information available, we were not able to determine the metabolic profile differences between Groups 2 and 3; it would be interesting to analyze the metabolic profiles of these two groups in the future to establish whether they are distinguishable. The recently proposed prostate cancer grading system also divides GS 8–10 disease into GS 8 and GS 9–10 on the foundation of biochemical repetition (BCR) succeeding radical prostatectomy (RP) as an outcome. However, BCR does not inevitably portend worse survival outcomes, so this is a less important distinction to make biochemically.

In pancreatic ductal adenocarcinoma (PDAC), GOT1 is required to sustain cell growth by enabling the production of NADPH to compensate for internal ROS<sup>33</sup>. We showed that GOT1 is essential for PCa cell line (PC3 and LNCaP) growth. *GOT1* knockdown increased ROS levels, suggesting that GOT1 may be involved in NADPH generation. Consistent with this, chronic exposure to cycling anoxia/re-oxygenation induced GOT1 in DU145 prostate cancer cells, that decreased cellular ROS levels with increased glutathione<sup>52</sup>. It would be interesting to establish if *GOT1* knocked-down PC3 and LNCaP cells are more sensitive to ROS-inducing treatments such as ionizing radiation. In PDAC, GOT1 plays a role in anaplerotic metabolism countering ROS production in low pH conditions<sup>53</sup>. However, this is unlikely to be the case in PCa. Pancreatic cancer cell growth was reduced in low pH conditions, but acidic pH stimulated both proliferation and colony formation of PC3 cells<sup>54</sup>, while LNCaP cells underwent apoptosis<sup>55</sup>. Although the effect of low pH in PCa cells was not similar to PDAC cells, it remains to be determined whether low pH increases GOT1 levels or whether anaplerotic glutamine metabolism is affected by pH in PCa cells.

As reported previously, GOT1 also functions as a member of the malate-aspartate shuttle<sup>32</sup>, in which two pairs of enzymes, glutamate oxaloacetate transaminases (GOT) and malate dehydrogenase (MDH), serve to transfer reducing equivalents across the mitochondrial membrane (Fig. 2b). Our transcriptomic analysis revealed the upregulation of all members of the shuttle including *GOT1*, *GOT2*, *MDH1*, and *MDH2* (Supplementary Table 7). These results suggest that the malate-aspartate shuttle may play an important role in cell growth in PCa. This hypothesis is supported by the reduction in NAD/NADH ratio upon *GOT1* knockdown in both cell lines. In agreement with our observation, Zhou *et al.* showed that GOT1 inhibition in osteosarcoma and lung cancer cell lines decreased proliferation and NADH accumulation. Higher GOT1 expression was linked to poor survival in thyroid carcinoma, breast carcinoma, and lung adenocarcinoma, suggesting that GOT1 also plays a pivotal role in other cancer types<sup>56</sup>.

## Conclusions

In this study we identified a group of RNA and metabolite signatures in urine as putative PCa biomarkers and therapeutic targets. To improve the accuracy of disease classification, we carried out metabolic and transcriptomic profiling of urine obtained from BPH, PTT, and PCa patients (without prostatic massage). Urine from normal healthy individuals was used as the control. Through an integrated analysis of metabolomic and transcriptomic data, we identified GOT1 as a key regulator of metabolic changes in PCa patients. Prostate cancers appear to undergo GOT1-dependent metabolic adaptation to promote a malignant phenotype and resist oxidative stress. The glutamate phenotype represented by the gene expression and metabolic changes in urine reflect this GOT1-dependent pathway in PCa cells. In addition to focusing on these pathway components as urine biomarkers of PCa, enzymes involved in this pathway might be excellent targets for PCa therapy. Indeed, small molecule inhibitors of GLS1 (mitochondrial glutaminase), which converts glutamine to glutamate, already exist<sup>57,58</sup>. Targeting this pathway is worthy of further investigation either with or without concurrent ROS-induced cellular stress<sup>57</sup>, this latter approach a particularly appealing strategy in patients with prostate cancers treated with ionizing radiotherapy. Finally, we postulate that liquid biopsies are an extremely useful tool for non-invasive biomarker and target discovery not only in PCa but also in other cancers.

## Methods

**Sample collection and preparation.** The Sanford Burnham Prebys Medical Discovery Institute, Adventishealth (formerly named as the Florida Hospital) and the John Hopkins University School of Medicine IRB protocols approved the study to consent participants and collect samples (patient and healthy individual urine and tissue samples). We have obtained the consent from all subjects, and the samples were de-identified prior distributing to the lab members. Our studies agreed to the principles set out in the WMA Declaration of Helsinki and the Department of Health Human Services Belmont Report.

This was a retrospective study of individuals attending the Global Robotics Institute (Celebration, FL, USA) and Florida Urology Associates (Orlando, FL, USA) of Florida Hospital and the Johns Hopkins Hospital between 2008 and 2014. Urine samples were collected from 20 benign prostatic hyperplasia (BPH), 11 prostatitis (PTT), and 20 prostate cancer (PCa) patients and 20 normal healthy individuals with no history of cancer attending the Global Robotics Institute and Florida Urology Associates of Florida Hospital. An additional 55 urine samples (11 samples from each different Gleason score) were obtained from patients attending the Johns Hopkins Hospital. Urine samples were collected using urine preservation tubes (Norgen Bioteck, Thorold, ON, Canada) and kept at room temperature until centrifugation to separate the exfoliated cells in the urine samples. Cell-free urine was then stored at  $-80^{\circ}\text{C}$  until further use for metabolite analysis. The exfoliated cells from normal and PCa urine samples were used for total RNA purification using the Urine (Exfoliated Cell) RNA purification kit (Norgen Bioteck). Total RNA was subjected to RNA-seq to identify gene signatures.

**Global untargeted metabolomics.** We conducted global metabolomics using ultra-high performance liquid chromatography coupled with high-resolution mass spectrometry (UHPLC-HRMS) on a Thermo Q Exactive with Dionex UHPLC (Thermo Fisher Scientific, Waltham, MA). To 50  $\mu\text{L}$  of urine, 20  $\mu\text{L}$  of internal standard was added (40  $\mu\text{g}/\text{mL}$  tryptophan-d3, 4  $\mu\text{g}/\text{mL}$  leucine-d10, 4  $\mu\text{g}/\text{mL}$  creatine-d3, and 4  $\mu\text{g}/\text{mL}$  caffeine-d3) followed by 400  $\mu\text{L}$  of 98:2 acetonitrile:water with 0.1% sodium azide. The solution was vortexed and spun down at 20,000  $\times$  g ( $8^{\circ}\text{C}$ ) for 10 min. Transferred supernatant was dried under a gentle stream of nitrogen and the dried sample restored in 50  $\mu\text{L}$  of 0.1% formic acid in water was transferred to a LC vial with a fused glass insert for analysis. Liquid chromatography–high resolution mass spectrometry (LC-HRMS)<sup>59,60</sup> analysis was done in positive and negative ion modes as separate injections, injecting 2  $\mu\text{L}$  for positive and 4  $\mu\text{L}$  for negative ions. Separation was achieved on a C18-pfp column (ACE Excel 100 $\times$ 2.1 mm, 2  $\mu\text{m}$ , Advanced Chromatography Technologies, Aberdeen, Scotland) with 0.1% formic acid in water as A and acetonitrile as B. An in-house retention library of 600 metabolites was used to identify metabolites. All the data normalization, multivariate analyses, pathway analysis, and biomarker discovery were carried out using Metaboanalyst 4.0 (<http://www.metaboanalyst.ca>). Integrated gene-metabolite network analysis was conducted using the Metascape plugin (<http://metscape.ncibi.org/>) in the Cytoscape (<http://www.cytoscape.org/>).

**Cell culture.** PCa cell lines, PC3 (ATCC<sup>®</sup> CRL-7934<sup>™</sup>) and LNCaP (ATCC<sup>®</sup> CRL-1740<sup>™</sup>) were cultured in Dulbecco's Modified Eagle Medium and RPMI 1640 medium (Thermo Fisher Scientific), respectively, with 10% FBS and penicillin/streptomycin.

**Quantitative real-time PCR (qPCR).** Total RNAs from cell lines were purified using the Direct-zol RNA Miniprep kit (Zymo Research, Irvine, CA). Normal prostate epithelial cell RNA was purchased from BioChain Institute Inc. (Catalog # R1234201-50, Newark, CA). RNA (0.5  $\mu\text{g}$ ) was then used for cDNA synthesis using a high capacity cDNA reverse transcription kit (Applied Biosystems, Foster City, CA). qPCR was performed using a Power SYBR Green PCR master mix (Applied Biosystems) in the 7500 Real-Time PCR system (Applied Biosystems). A final reaction volume of 10  $\mu\text{L}$  was used containing 1  $\mu\text{L}$  (corresponding to 10 ng) of cDNA template, 5  $\mu\text{L}$  of 2X Power SYBR Green PCR master mix (Applied Biosystems), and 0.2  $\mu\text{M}$  of each primer. The reaction was subjected to denaturation at  $95^{\circ}\text{C}$  for 10 min followed by 40 cycles of denaturation at  $95^{\circ}\text{C}$  for 15 sec and annealing at  $58^{\circ}\text{C}$  for 1 min. The comparative Ct analysis was performed using SDS1.2.3 software (Applied Biosystems) with TATA-box binding protein (TBP),  $\beta$ -actin (ACTB), and hypoxanthine phosphoribosyltransferase 1 (HPRT1) serving as endogenous controls (average Ct value of three control genes was used for calculating  $\Delta\text{Ct}$  value). The primer sequences for the genes are listed in Supplementary Table 8.

**RNA access.** The quantity and integrity of the RNA was measured using both the Qubit RNA HS Assay Kit (Thermo Fisher Scientific) and the Agilent 2100 Bioanalyzer RNA Pico kit (Agilent Technologies, Santa Clara, CA). Following the Illumina DV<sub>200</sub> metric (percentage of RNA fragments greater than 200 nucleotides), 100 ng of RNA with DV<sub>200</sub> > 30% was used to prepare sequencing libraries in accordance with the TruSeq RNA Access protocol (Illumina, Inc., San Diego, CA). The random primers were used for the first strand cDNA synthesis followed by second strand synthesis. The cDNA then underwent 3' adenylation followed by adapter ligation and PCR amplification (15 cycles). Library quality was measured using both the Qubit dsDNA HS Assay Kit and Agilent Bioanalyzer DNA kit. A 4-plex pool of libraries was then made (200 ng of each sample) followed by two rounds of hybridization/capture and a final amplification (10 cycles). The quality and quantity of the final libraries were determined using the Agilent 2100 Bioanalyzer DNA HS kit and Kapa Biosystems qPCR (Kapa Biosystems, Inc., Wilmington, MA). Multiplexed libraries were pooled and normalized to 17.5 pM. The libraries were sequenced using a 75 bp paired-end run on the Illumina MiSeq instrument. Reads were mapped to the hg19 human genome using tophat2.0.1; mapped reads were filtered based on the mapping quality. The overall mapping rates were about 93%. mRNA quantification was conducted in Partek Genomics Suite 6.6. The differential expression of mRNAs was analyzed using a R package edgeR.

**Clustering and principal component analysis (PCA).** The resulting mRNA expression profile included 46,459 transcripts, and the non-parametric Mann-Whitney U-test was used to identify significantly regulated transcripts. 5510 transcripts were identified ( $p \leq 0.05$ ) as significantly differentially expressed between normal and PCa groups. Within those 5,510 transcripts, 1,118 transcripts had Reads Per Kilobase of transcript per Million mapped reads (RPKM) values greater than 1.0 for all samples. We pre-compiled a gene panel that listed all prostate cancer-related genes (with the help of Illumina). By comparing with this panel, we obtained 542 transcripts with RPKMs greater than 1.0 for all samples. Within these 542 transcripts, 116 were significantly regulated (Table 1). All overlapping transcripts between the expression profile and prostate cancer panel, a total of 3825 transcripts, were used to run the unsupervised clustering analysis and PCA. Hierarchical cluster analysis was performed in R using the correlation between samples to characterize similarity. Each sample was assigned to its own group and cluster algorithm was repeated until there was a single cluster. Correlations between samples were calculated using the expression values of the 3825 transcripts. We also used PCA to visualize sample-to-sample distance. The transformation was defined that the first principal component accounted for the largest variance (as much of the variability in the dataset as possible). In the results, each sample was projected onto the 3D space in which the three axes were the first three highest principle components (see Supplementary Fig. 2b).

**Gene set enrichment analysis (GSEA) of tissue RNA-seq data.** Gene expression was analyzed in 65 patients using RNA-seq data<sup>25</sup>. Briefly, RPKM values from tumor and matched normal samples from 65 patients were analyzed for differential gene expression between tumor and matched normal using the Bioconductor limma package<sup>61</sup>. To compare between groups for tumor or matched normal tissues, linear model analysis using the limma package was used on RPKM values for all genes. The analysis was performed in R (version 3.4.4, [www.R-project.org](http://www.R-project.org)).

GSEA for Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways was conducted using the GSEA desktop application ([borad.mit.edu/gsea](http://borad.mit.edu/gsea)). All genes were ranked using scores based on fold-change direction and p-value, and enrichment analysis was conducted using GSEA pre-ranked with the 'classic' enrichment statistic. Significantly enriched genes sets were identified using q-value  $\leq 0.05$  as a cutoff.

**Transient transfection and cell proliferation assay.** Cells ( $0.3 \times 10^6$  cells) were mixed with siRNA (Thermo Fisher Scientific; final concentration 20 nM) and lipofectamine RNAiMAX (Thermo Fisher Scientific) mixture in 2 ml medium containing 10% FBS. In a 96-well plate, cells were seeded in duplicate at 7500 cells per wells and, after 24 h, the medium was replaced. Cell proliferation assay was performed using the CellTiter96 Aqueous One Solution Cell Proliferation Assay (MTS) kit (Promega, Madison, WI).

**NAD and NADH quantification.** NAD/NADH ratio was determined using the NAD/NADH-Glo Assay kit (Promega). In a 96-well plate, cells were seeded in duplicate at 7500 cells per wells. Cells were lysed with base solution (100 mM sodium carbonate, 20 mM sodium bicarbonate, 10 mM nicotinamide, and 0.05% Triton X-100) with 1% dodecyltrimethylammonium bromide (Sigma Aldrich, St Louis, MO; D8638). Lysed cells were heated at 60 °C for 20 min in the presence/absence of acid, then subjected to the luciferase assay according to the manufacturer's protocol.

**Measurement of reactive oxygen species (ROS).** One million cells were suspended in 1 ml medium having 20  $\mu$ M 2',7'-dichlorofluorescein diacetate (H<sub>2</sub>DCF-DA, Sigma Aldrich, D6883) and incubated for 30 min at 37 °C in 5% CO<sub>2</sub>. DCF fluorescence were detected using a FACSCalibur flow cytometer (Becton Dickinson, Anaheim, CA). Flowing Software v. 2.5.1 ([www.flowingsoftware.com](http://www.flowingsoftware.com)) was used for data analysis.

**Invasion assay.** The cell invasion assay was performed using Corning BioCoat Matrigel Invasion Chambers (Discovery Labware, Bedford, MA). Cells were incubated in serum-free medium for 24 h and added to the upper chambers in serum-free medium ( $0.2 \times 10^6$  cells). The lower chamber was filled with medium containing 10% fetal bovine serum. After 48 hr incubation at 37 °C, cells at the bottom side of upper chamber were stained with 0.5% crystal violet dye. The dye from invaded cells was extracted using methanol and absorbance at 570 nm was measured.

**Soft agar colony formation assay.** Anchorage-independent growth was assessed using a CytoSelect 96-well cell transformation assay kit (Cell Biolabs Inc., San Diego, CA). Briefly, cells were seeded in soft agar at a density of 10,000 cells per well and incubated at 37 °C in 5% CO<sub>2</sub> for 7 days. Colony formation was quantified with the 3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide (MTT) assay provided with the kit according to the assay protocol.

**Ethics approval and consent to participate.** The institutional review board (IRB) of the Florida Hospital and The Johns Hopkins Hospital approved sample use. All participants were required to sign and provide written consent. All animal studies was conducted with the approval of the appropriate ethics committee in conformity with the ethical standards.

### Data availability

The data are available from the corresponding author upon request.

Received: 6 October 2019; Accepted: 4 February 2020;

Published online: 28 February 2020

## References

- Siegel, R. L., Miller, K. D. & Jemal, A. Cancer statistics, 2016. *CA Cancer J. Clin.* **66**, 7–30, <https://doi.org/10.3322/caac.21332> (2016).
- Torre, L. A. *et al.* Global cancer statistics, 2012. *CA Cancer J. Clin.* **65**, 87–108, <https://doi.org/10.3322/caac.21262> (2015).
- Truong, M., Yang, B. & Jarrard, D. F. Toward the detection of prostate cancer in urine: a critical analysis. *J. Urol.* **189**, 422–429, <https://doi.org/10.1016/j.juro.2012.04.143> (2013).
- Prensner, J. R., Rubin, M. A., Wei, J. T. & Chinnaiyan, A. M. Beyond PSA: the next generation of prostate cancer biomarkers. *Sci. Transl. Med.* **4**, 127rv123, <https://doi.org/10.1126/scitranslmed.3003180> (2012).
- Hessels, D. & Schalken, J. A. The use of PCA3 in the diagnosis of prostate cancer. *Nat. Rev. Urol.* **6**, 255–261, <https://doi.org/10.1038/nrurol.2009.40> (2009).
- Evaluation of Genomic Applications in, P. & Prevention Working, G. Recommendations from the EGAPP Working Group: does PCA3 testing for the diagnosis and management of prostate cancer improve patient health outcomes? *Genet. Med.* **16**, 338–346, <https://doi.org/10.1038/gim.2013.141> (2014).
- Issaq, H. J. *et al.* Detection of bladder cancer in human urine by metabolomic profiling using high performance liquid chromatography/mass spectrometry. *J. Urol.* **179**, 2422–2426, <https://doi.org/10.1016/j.juro.2008.01.084> (2008).
- Kim, K. *et al.* Urine metabolomics analysis for kidney cancer detection and biomarker discovery. *Mol. Cell. Proteomics* **8**, 558–570, <https://doi.org/10.1074/mcp.M800165-MCP200> (2009).
- Pasikanti, K. K. *et al.* Noninvasive urinary metabolomic diagnosis of human bladder cancer. *J. Proteome Res.* **9**, 2988–2995, <https://doi.org/10.1021/pr901173v> (2010).
- Bouatra, S. *et al.* The human urine metabolome. *Plos One* **8**, e73076, <https://doi.org/10.1371/journal.pone.0073076> (2013).
- Hanahan, D. & Weinberg, R. A. Hallmarks of cancer: the next generation. *Cell* **144**, 646–674, <https://doi.org/10.1016/j.cell.2011.02.013> (2011).
- Caro, P. *et al.* Metabolic signatures uncover distinct targets in molecular subsets of diffuse large B cell lymphoma. *Cancer Cell* **22**, 547–560, <https://doi.org/10.1016/j.ccr.2012.08.014> (2012).
- Chinnaiyan, P. *et al.* The metabolomic signature of malignant glioma reflects accelerated anabolic metabolism. *Cancer Res.* **72**, 5878–5888, <https://doi.org/10.1158/0008-5472.CAN-12-1572-T> (2012).
- Jiang, Y., Cheng, X., Wang, C. & Ma, Y. Quantitative determination of sarcosine and related compounds in urinary samples by liquid chromatography with tandem mass spectrometry. *Anal. Chem.* **82**, 9022–9027, <https://doi.org/10.1021/ac1019914> (2010).
- Cao, D. L. *et al.* A multiplex model of combining gene-based, protein-based, and metabolite-based with positive and negative markers in urine for the early diagnosis of prostate cancer. *Prostate* **71**, 700–710, <https://doi.org/10.1002/pros.21286> (2011).
- Sreekumar, A. *et al.* Metabolomic profiles delineate potential role for sarcosine in prostate cancer progression. *Nature* **457**, 910–914, <https://doi.org/10.1038/nature07762> (2009).
- Jentzmik, F. *et al.* Sarcosine in urine after digital rectal examination fails as a marker in prostate cancer detection and identification of aggressive tumours. *Eur. Urol.* **58**, 12–18; discussion 20–11, <https://doi.org/10.1016/j.eururo.2010.01.035> (2010).
- Ferrara, C. T. *et al.* Genetic networks of liver metabolism revealed by integration of metabolic and transcriptional profiling. *PLoS Genet.* **4**, e1000034, <https://doi.org/10.1371/journal.pgen.1000034> (2008).
- Zhang, G. *et al.* Integration of metabolomics and transcriptomics revealed a fatty acid network exerting growth inhibitory effects in human pancreatic cancer. *Clin. Cancer Res.* **19**, 4983–4993, <https://doi.org/10.1158/1078-0432.CCR-13-0209> (2013).
- Sullivan, P. S., Chan, J. B., Levin, M. R. & Rao, J. Urine cytology and adjunct markers for detection and surveillance of bladder cancer. *Am. J. Transl. Res.* **2**, 412–440 (2010).
- Lee, B. *et al.* Long noncoding RNAs as putative biomarkers for prostate cancer detection. *J. Mol. Diagn.* **16**, 615–626, <https://doi.org/10.1016/j.jmoldx.2014.06.009> (2014).
- Subramanian, A. *et al.* Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl. Acad. Sci. USA* **102**, 15545–15550, <https://doi.org/10.1073/pnas.0506580102> (2005).
- Cantor, J. R. & Sabatini, D. M. Cancer cell metabolism: one hallmark, many faces. *Cancer Discov.* **2**, 881–898, <https://doi.org/10.1158/2159-8290.CD-12-0345> (2012).
- DeBerardinis, R. J. & Chandel, N. S. Fundamentals of cancer metabolism. *Sci. Adv.* **2**, e1600200, <https://doi.org/10.1126/sciadv.1600200> (2016).
- Shao, Y. *et al.* Metabolomics and transcriptomics profiles reveal the dysregulation of the tricarboxylic acid cycle and related mechanisms in prostate cancer. *Int. J. Cancer* **143**, 396–407, <https://doi.org/10.1002/ijc.31313> (2018).
- Xia, J. & Wishart, D. S. Using MetaboAnalyst 3.0 for Comprehensive Metabolomics Data Analysis. *Curr. Protoc. Bioinformatics* **55**(14), 10.11–14.10.91, <https://doi.org/10.1002/cpbi.11> (2016).
- Kim, S., You, S. & Hwang, D. Aminoacyl-tRNA synthetases and tumorigenesis: more than housekeeping. *Nat. Rev. Cancer* **11**, 708–718, <https://doi.org/10.1038/nrc3124> (2011).
- Altman, B. J., Stine, Z. E. & Dang, C. V. From Krebs to clinic: glutamine metabolism to cancer therapy. *Nat. Rev. Cancer* **16**, 619–634, <https://doi.org/10.1038/nrc.2016.71> (2016).
- Son, J. *et al.* Glutamine supports pancreatic cancer growth through a KRAS-regulated metabolic pathway. *Nature* **496**, 101–105, <https://doi.org/10.1038/nature12040> (2013).
- Ying, H. *et al.* Oncogenic Kras maintains pancreatic tumors through regulation of anabolic glucose metabolism. *Cell* **149**, 656–670, <https://doi.org/10.1016/j.cell.2012.01.058> (2012).
- Sullivan, L. B. *et al.* Supporting Aspartate Biosynthesis Is an Essential Function of Respiration in Proliferating Cells. *Cell* **162**, 552–563, <https://doi.org/10.1016/j.cell.2015.07.017> (2015).
- Birsoy, K. *et al.* An Essential Role of the Mitochondrial Electron Transport Chain in Cell Proliferation Is to Enable Aspartate Synthesis. *Cell* **162**, 540–551, <https://doi.org/10.1016/j.cell.2015.07.016> (2015).
- Son, J. *et al.* Glutamine supports pancreatic cancer growth through a KRAS-regulated metabolic pathway (vol 496, pg 101, 2013). *Nature* **499**, <https://doi.org/10.1038/nature12317> (2013).
- van 't Veer, L. J. *et al.* Gene expression profiling predicts clinical outcome of breast cancer. *Nature* **415**, 530–536, <https://doi.org/10.1038/415530a> (2002).
- Dupuy, A. & Simon, R. M. Critical review of published microarray studies for cancer outcome and guidelines on statistical analysis and reporting. *J. Natl. Cancer Inst.* **99**, 147–157, <https://doi.org/10.1093/jnci/djk018> (2007).
- Dakubo, G. D., Parr, R. L., Costello, L. C., Franklin, R. B. & Thayer, R. E. Altered metabolism and mitochondrial genome in prostate cancer. *J. Clin. Pathol.* **59**, 10–16, <https://doi.org/10.1136/jcp.2005.027664> (2006).
- DeBerardinis, R. J., Lum, J. J., Hatzivassiliou, G. & Thompson, C. B. The biology of cancer: metabolic reprogramming fuels cell growth and proliferation. *Cell Metab.* **7**, 11–20, <https://doi.org/10.1016/j.cmet.2007.10.002> (2008).
- Costello, L. C., Franklin, R. B. & Feng, P. Mitochondrial function, zinc, and intermediary metabolism relationships in normal prostate and prostate cancer. *Mitochondrion* **5**, 143–153, <https://doi.org/10.1016/j.mito.2005.02.001> (2005).
- Costello, L. C. & Franklin, R. B. Concepts of citrate production and secretion by prostate. 1. *Metabolic relationships. Prostate* **18**, 25–46 (1991).
- Fendler, A., Stephan, C., Yousef, G. M., Kristiansen, G. & Jung, K. The translational potential of microRNAs as biofluid markers of urological tumours. *Nature Reviews Urology* **13**, 734–752, <https://doi.org/10.1038/nrurol.2016.193> (2016).
- Kulbe, J. R. & Geddes, J. W. Current status of fluid biomarkers in mild traumatic brain injury. *Exp. Neurol.* **275**(Pt 3), 334–352, <https://doi.org/10.1016/j.expneurol.2015.05.004> (2016).

42. Nguyen, M. T., Ross, G. F., Dent, C. L. & Devarajan, P. Early prediction of acute renal injury using urinary proteomics. *Am. J. Nephrol.* **25**, 318–326, <https://doi.org/10.1159/000086476> (2005).
43. Morrissey, J. J. *et al.* Evaluation of Urine Aquaporin-1 and Perilipin-2 Concentrations as Biomarkers to Screen for Renal Cell Carcinoma A Prospective Cohort Study. *Jama Oncology* **1**, 204–212, <https://doi.org/10.1001/jamaoncol.2015.0213> (2015).
44. Chen, C. L. *et al.* Identification of potential bladder cancer markers in urine by abundant-protein depletion coupled with quantitative proteomics. *J. Proteomics* **85**, 28–43, <https://doi.org/10.1016/j.jprot.2013.04.024> (2013).
45. Flatley, B., Wilmott, K. G., Malone, P. & Cramer, R. MALDI MS Profiling of Post-DRE Urine Samples Highlights the Potential of beta-Microseminoprotein as a Marker for Prostatic Diseases. *Prostate* **74**, 103–111, <https://doi.org/10.1002/pros.22736> (2014).
46. Sanda, M. G. *et al.* Association between combined tmprss2:eRg and pca3 rna urinary testing and detection of aggressive prostate cancer. *JAMA Oncology*, <https://doi.org/10.1001/jamaoncol.2017.0177> (2017).
47. Heger, Z. *et al.* Determination of common urine substances as an assay for improving prostate carcinoma diagnostics. *Oncol. Rep.* **31**, 1846–1854, <https://doi.org/10.3892/or.2014.3054> (2014).
48. Struck-Lewicka, W. *et al.* Urine metabolic fingerprinting using LC-MS and GC-MS reveals metabolite changes in prostate cancer: A pilot study. *J. Pharm. Biomed. Anal.* **111**, 351–361, <https://doi.org/10.1016/j.jpba.2014.12.026> (2015).
49. Lima, A. R., Bastos Mde, L., Carvalho, M. & Guedes de Pinho, P. Biomarker Discovery in Human Prostate Cancer: an Update in Metabolomics Studies. *Transl. Oncol.* **9**, 357–370, <https://doi.org/10.1016/j.tranon.2016.05.004> (2016).
50. Sharma, S. Tumor markers in clinical practice: General principles and guidelines. *Indian J. Med. Paediatr. Oncol.* **30**, 1–8, <https://doi.org/10.4103/0971-5851.56328> (2009).
51. Pierorazio, P. M., Walsh, P. C., Partin, A. W. & Epstein, J. I. Prognostic Gleason grade grouping: data based on the modified Gleason scoring system. *BJU Int.* **111**, 753–760, <https://doi.org/10.1111/j.1464-410X.2012.11611.x> (2013).
52. Matschke, J. *et al.* Targeted Inhibition of Glutamine-Dependent Glutathione Metabolism Overcomes Death Resistance Induced by Chronic Cycling Hypoxia. *Antioxid Redox Signal* **25**, 89–107, <https://doi.org/10.1089/ars.2015.6589> (2016).
53. Abrego, J. *et al.* GOT1-mediated anaplerotic glutamine metabolism regulates chronic acidosis stress in pancreatic cancer cells. *Cancer Lett.* **400**, 37–46, <https://doi.org/10.1016/j.canlet.2017.04.029> (2017).
54. Huang, S. *et al.* Acidic extracellular pH promotes prostate cancer bone metastasis by enhancing PC-3 stem cell characteristics, cell invasiveness and VEGF-induced vasculogenesis of BM-EPCs. *Oncol. Rep.* **36**, 2025–2032, <https://doi.org/10.3892/or.2016.4997> (2016).
55. Iguchi, K. *et al.* Characterization of the low pH/low nutrient-resistant LNCaP cell subline LNCaP-F10. *Oncol. Rep.* **28**, 2009–2015, <https://doi.org/10.3892/or.2012.2039> (2012).
56. Zhou, X., Curbo, S., Li, F., Krishnan, S. & Karlsson, A. Inhibition of glutamate oxaloacetate transaminase 1 in cancer cell lines results in altered metabolism with increased dependency of glucose. *BMC Cancer* **18**, 559, <https://doi.org/10.1186/s12885-018-4443-1> (2018).
57. Chakrabarti, G. *et al.* Targeting glutamine metabolism sensitizes pancreatic cancer to PARP-driven metabolic catastrophe induced by ss-lapachone. *Cancer Metab* **3**, 12, <https://doi.org/10.1186/s40170-015-0137-1> (2015).
58. Shukla, K. *et al.* Design, synthesis, and pharmacological evaluation of bis-2-(5-phenylacetamido-1,2,4-thiadiazol-2-yl)ethyl sulfide 3 (BPTES) analogs as glutaminase inhibitors. *J. Med. Chem.* **55**, 10551–10563, <https://doi.org/10.1021/jm301191p> (2012).
59. Liu, H., Garrett, T. J., Tayyari, F. & Gu, L. Profiling the metabolome changes caused by cranberry procyanidins in plasma of female rats using (1) H NMR and UHPLC-Q-Orbitrap-HRMS global metabolomics approaches. *Mol. Nutr. Food Res.* **59**, 2107–2118, <https://doi.org/10.1002/mnfr.201500236> (2015).
60. Mahmud, I., Sternberg, S., Williams, M. & Garrett, T. J. Comparison of global metabolite extraction strategies for soybeans using UHPLC-HRMS. *Anal. Bioanal. Chem.* **409**, 6173–6180, <https://doi.org/10.1007/s00216-017-0557-6> (2017).
61. Ritchie, M. E. *et al.* limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res.* **43**, e47, <https://doi.org/10.1093/nar/gkv007> (2015).

## Acknowledgements

We thank Sanford Burnham Prebys Medical Discovery Institute (SBP) Analytical Genomics core facility for deep sequencing, Bioinformatics core for data analysis support, and Ms. Debbie McFadden for formatting the manuscript. This manuscript was deposited bioRxiv in draft form (<https://www.biorxiv.org/content/10.1101/599514v1>). This work was supported by National Institutes of Health grant NCI 5P30CA030199 (SBP), P30 CA006973 (JHU SKCCC), Florida Department of Health, Bankhead-Coley Cancer Research Program 5BC08 and International Prostate Cancer Foundation (IPCF) to R.J.P. and Southeast Center for Integrated Metabolomics (SECIM) – NIH Grant #U24 DK097209 to T.G.

## Author contributions

R.J.P. incept the idea and, planned the experiment with B.L., T.J.G., I.M., J.M., F.Q. and analyzed data. R.J.P., B.L., I.M. and T.J.G., wrote the manuscript. T.J.G., P.D. and I.M. process samples and analyzed metabolomics data. J.M. performed RNA access-seq. F.Q. and P.J. analyzed RNA access-seq results. F.W. and Y.S. provided prostate cancer tissue RNA-seq data and analysis. F.V., I.R., V.P. and C.P.P. provided urine samples. G.P.S. supported RNA access-seq. All authors read and approved the final manuscript.

## Competing interests

The authors declare no competing interests.

## Additional information

**Supplementary information** is available for this paper at <https://doi.org/10.1038/s41598-020-60616-z>.

**Correspondence** and requests for materials should be addressed to R.J.P.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.





**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2020