



# Gene Expression Changes and Associated Pathways Involved in the Progression of Prostate Cancer Advanced Stages

Elena A. Pudova<sup>1</sup>, George S. Krasnov<sup>1</sup>, Anastasiya A. Kobelyatskaya<sup>1</sup>, Maria V. Savvateeva<sup>1</sup>, Maria S. Fedorova<sup>1</sup>, Vladislav S. Pavlov<sup>1</sup>, Kirill M. Nyushko<sup>2</sup>, Andrey D. Kaprin<sup>2</sup>, Boris Y. Alekseev<sup>2</sup>, Dmitry Y. Trofimov<sup>3</sup>, Gennady T. Sukhikh<sup>3</sup>, Anastasiya V. Snezhkina<sup>1</sup> and Anna V. Kudryavtseva<sup>1\*</sup>

## OPEN ACCESS

### Edited by:

Yuriy L. Orlov,  
I.M. Sechenov First Moscow State  
Medical University, Russia

### Reviewed by:

Stanislav Rybtsov,  
University of Edinburgh,  
United Kingdom  
Angelika Timofeeva,  
National Medical Research Center of  
Obstetrics, Gynecology and  
Perinatology named after  
Academician V.I. Kulakova, Russia  
Vladimir Kalinin,  
TMA-Science, Ireland

### \*Correspondence:

Anna V. Kudryavtseva  
rhizamoeba@mail.ru

### Specialty section:

This article was submitted to  
Computational Genomics,  
a section of the journal  
Frontiers in Genetics

**Received:** 01 October 2020

**Accepted:** 24 December 2020

**Published:** 21 January 2021

### Citation:

Pudova EA, Krasnov GS,  
Kobelyatskaya AA, Savvateeva MV,  
Fedorova MS, Pavlov VS,  
Nyushko KM, Kaprin AD,  
Alekseev BY, Trofimov DY,  
Sukhikh GT, Snezhkina AV and  
Kudryavtseva AV (2021)  
Gene Expression Changes and  
Associated Pathways Involved in the  
Progression of Prostate Cancer  
Advanced Stages.  
*Front. Genet.* 11:613162.  
doi: 10.3389/fgene.2020.613162

<sup>1</sup>Engelhardt Institute of Molecular Biology, Russian Academy of Sciences, Moscow, Russia, <sup>2</sup>National Medical Research Radiological Center, Ministry of Health of the Russian Federation, Moscow, Russia, <sup>3</sup>National Medical Research Center for Obstetrics, Gynecology and Perinatology named after Academician V.I. Kulakov, Ministry of Health of the Russian Federation, Moscow, Russia

Prostate cancer (PC) is one of the most common cancers among men worldwide, and advanced PCs, such as locally advanced PC (LAPC) and castration-resistant PC (CRPC), present the greatest challenges in clinical management. Current indicators have limited capacity to predict the disease course; therefore, better prognostic markers are greatly needed. In this study, we performed a bioinformatic analysis of The Cancer Genome Atlas (TCGA) datasets, including RNA-Seq data from the prostate adenocarcinoma (PRAD;  $n = 55$ ) and West Coast Dream Team – metastatic CRPC (WCMT-MCRPC;  $n = 84$ ) projects, to evaluate the transcriptome changes associated with progression-free survival (PFS) for LAPC and CRPC, respectively. We identified the genes whose expression was positively/negatively correlated with PFS. In LAPC, the genes with the most significant negative correlations were *ZC2HC1A*, *SQLE*, and *KIF11*, and the genes with the most significant positive correlations were *SOD3*, *LRRC26*, *MIR22HG*, *MEG3*, and *MIR29B2CHG*. In CRPC, the most significant positive correlations were found for *BET1*, *CTAGE5*, *IFNGR1*, and *GIMAP6*, and the most significant negative correlations were found for *CLPB*, *PRPF19*, *ZNF610*, *MPST*, and *LINC02001*. In addition, we performed a gene network interaction analysis using STRINGdb, which revealed a significant relationship between genes predominantly involved in the cell cycle and characterized by upregulated expression in early recurrence. Based on the results, we propose several genes that can be used as potential prognostic markers.

**Keywords:** prostate cancer, gene expression, PFS, TCGA, PRAD, WCMT-MCRPC

## INTRODUCTION

Prostate cancer (PC) is the second most common cancer in men worldwide (Rawla, 2019). Advanced PC presents as locally advanced and castration-resistant tumors. Locally advanced PC (LAPC) is characterized by the spread of the tumor beyond the prostate capsule and is more aggressive than localized PC. Androgen deprivation therapy aimed at reducing the level of circulating testosterone is

often used in LAPC treatment (Yap et al., 2016). However, despite rapid patient responses to this therapy, after 18–36 months, the disease frequently progresses to castration-resistant PC (CRPC; Yap et al., 2016). Metastatic CRPC is a prognostically unfavorable disease that requires regular systematic examination and monitoring and significantly impairs quality of life.

During therapy, in some patients, PC has an aggressive course, leading to metastasis, while in others, the disease has an indolent course, with a low tendency for progression. For prognostic assessment of the disease course, various clinical parameters are used, such as the level of prostate-specific antigen (PSA) and/or Gleason score; however, these parameters are not sufficiently informative (Keyes et al., 2013). To solve this problem, new molecular markers are needed that are highly associated with disease progression, which, when used in combination with existing clinical parameters, have a high predictive value. Currently, a promising area in the search for potential markers is the analysis of the most significant and consistent changes in the tumor transcriptome (Larkin et al., 2012; Cao et al., 2019; Pudova et al., 2019; Song et al., 2019).

An important criterion in the study of cancer data is progression-free survival (PFS). This parameter is the time from a random assignment in a clinical trial to disease progression or death from any cause (Forsythe et al., 2018). The study of PFS-specific molecular events will help to identify the major changes associated with the onset of disease progression.

This study aimed to analyze the transcriptome profiles of LAPC and CRPC based on the RNA-Seq data from the prostate adenocarcinoma (PRAD) and West Coast Dream Team – metastatic CRPC (WCDT-MCRPC) projects in The Cancer Genome Atlas (TCGA), respectively. We analyzed the changes in gene expression and molecular pathways related to PFS. The results of this study improve our understanding of the mechanisms underlying PC progression and identify molecules with potential as prognostic markers.

## MATERIALS AND METHODS

### Data Collection

The study included RNA-Seq data for LAPC samples from the PRAD project ( $n = 55$ ) and CRPC samples from the WCDT-MCRPC project ( $n = 84$ ), which were donated to the TCGA consortium. We focused on the study of Caucasian patients (Caucasians were identified as “white” in the databases). The trimmed mean of M-values (TMM) method was used to normalize the RNA-Seq data of each dataset. Ethical approval was not available for the study as our data were revealed from the public database.

### Relative Gene Expression and Downstream Analysis

RNA-Seq data from the TCGA project (read counts per gene evaluated by HTSeq) were downloaded from the repository of GDC Data Portal<sup>1</sup> and then analyzed in the statistical

environment R.<sup>2</sup> Data normalization and gene expression analysis with generalized linear models was performed using edgeR package (Robinson et al., 2010). The obtained results were considered statistically significant when the  $p$ -value of the quasi-likelihood F-test (QLF test) of  $<0.05$ . We did not use Benjamini-Hochberg (BH)  $p$ -value adjustment here because very few genes pass the threshold after it is applied.

In order to identify the genes whose expression change is most strongly associated with PFS, we evaluated Spearman's rank correlation coefficient between the normalized gene expression level and PFS. Here, the obtained results were also considered statistically significant when both Spearman's  $p < 0.05$  and QLF test  $p < 0.05$ . For this analysis, we pre-filtered genes by the following parameters: read counts per million (CPM)  $> 8$  and gene expression fold change is two- or (LogFC)  $> 1$ .

Next, we performed KEGG (Kyoto Encyclopedia of Genes and Genomes database) pathway enrichment analysis using the clusterProfile package (Yu et al., 2012). Additionally, we performed interaction network reconstruction and Gene Ontology (GO) pathway enrichment analysis using the STRINGdb (Szklarczyk et al., 2017). For enrichment analyses, BH adjustment to calculate the false discovery rate (FDR) was applied. The obtained results of this analysis were considered statistically significant when the FDR-value of  $<0.05$ . When constructing the networks, only data on direct protein-protein interactions were used (other associations such as co-expression, co-occurrence, gene fusions, and neighborhood, which are set by default, were excluded).

## RESULTS

Using PRAD and WCDT-MCRPC datasets, we analyzed changes of the relative gene expression associated with PFS (the number of days to recurrence) and found 889 genes for the LAPC and 1,889 genes for CRPC with the  $p < 0.05$  according to the QLF test (Supplementary Tables S1 and S2).

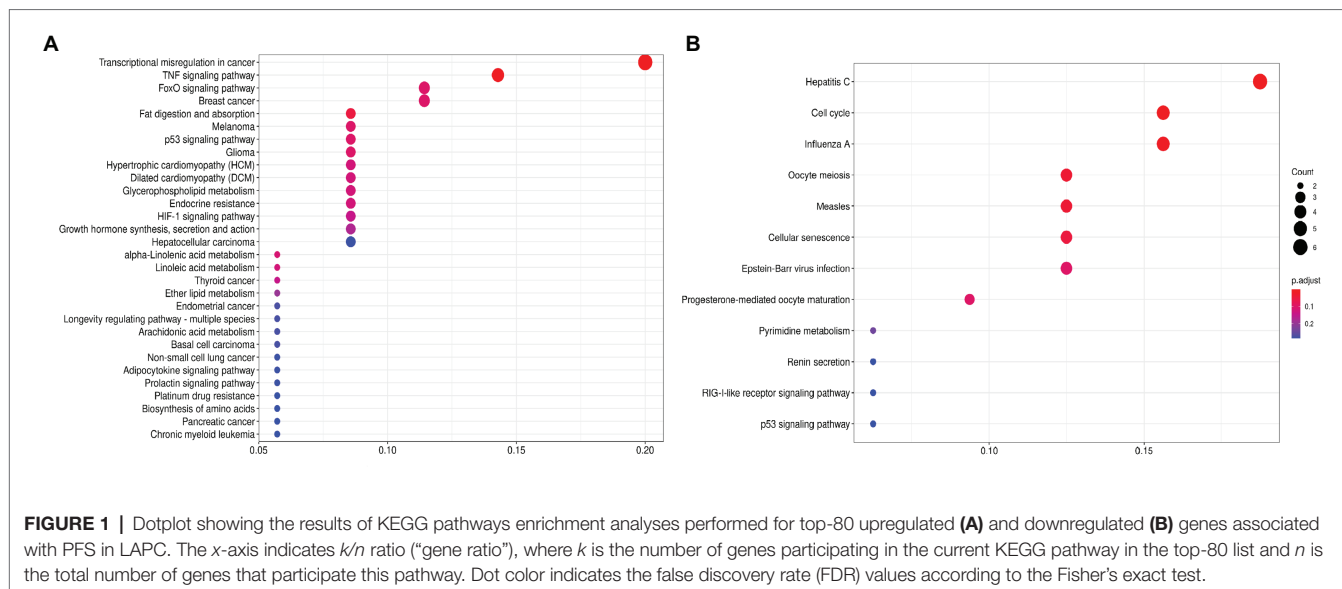
### Pathways Enrichment Analysis

First, we focused on the pathways enrichment analysis (KEGG database) to identify the major pathways, the expression of whose members may be positively or negatively associated with PFS (Figure 1). For this analysis, we selected top-80 genes, increased expression of which was positively associated with PFS (further – “upregulated genes”) and top-80 genes with negatively associated expression (further – “downregulated genes”). In LAPC, the most statistically significant enrichment with upregulated genes was noted for “Transcriptional misregulation in cancer” pathway (hsa05202). The analysis of downregulated genes showed a significant enrichment for the “Cell cycle” pathway (hsa04110; Table 1).

For CRPC, a pathway enrichment analysis (top-250 up- and downregulated genes) revealed two pathways, “Complement and coagulation cascades” (hsa04610) and “Drug metabolism-cytochrome

<sup>1</sup><https://portal.gdc.cancer.gov/>

<sup>2</sup><https://www.r-project.org/>



**TABLE 1 |** Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways that are significantly enriched with genes associated with progression-free survival (PFS) in locally advanced prostate cancer (LAPC).

KEGG ID	Pathway name	Enrichment p-value	Enrichment FDR	Genes
<b>Enriched with upregulated genes</b>				
hsa05202	Transcriptional misregulation in cancer	1,53E-05	1,98E-03	<i>CEBPB, RARA, GADD45B, CDKN1A, NR4A3, IGF1, SUPT3H</i>
hsa04668	TNF signaling pathway	1,12E-04	7,29E-03	<i>BCL3, CEBPB, JUNB, SOCS3, IRF1</i>
hsa04975	Fat digestion and absorption	8,31E-04	3,60E-02	<i>PLA2G5, CD36, PLA2G2A</i>
<b>Enriched with downregulated genes</b>				
hsa05160	Hepatitis C	2,81E-05	2,84E-03	<i>RSAD2, IFIT1, MX1, DDX58, OAS3, EIF2AK2</i>
hsa04110	Cell cycle	1,17E-04	5,89E-03	<i>E2F5, CCNB1, RAD21, CCNB2, BUB1</i>
hsa05164	Influenza A	5,08E-04	1,71E-02	<i>RSAD2, MX1, DDX58, OAS3, EIF2AK2</i>
hsa04114	Oocyte meiosis	1,57E-03	3,96E-02	<i>CALML3, CCNB1, CCNB2, BUB1</i>
hsa05162	Measles	2,07E-03	4,17E-02	<i>MX1, DDX58, OAS3, EIF2AK2</i>

P-450” (hsa00982), which were enriched with genes that have expression positively associated with PFS. When considering pathways negatively associated with PFS, we have seen an overrepresentation of genes participating in such cancer-associated pathways as the “TGF-beta signaling pathway” (hsa04350), “Hippo signaling pathway” (hsa04390), and others (Figure 2; Table 2).

### Correlation Analysis Between Gene Expression and PFS

Further, a correlation analysis between the gene expression and PFS was performed in LAPC and CRPC using Spearman’s rank correlation calculation.

For the LAPC, 34 genes were found to be correlated with PFS ( $p < 0.05$ ); an equal number of genes had positive (17/34) and negative (17/34) correlation with PFS (Figure 3).

For the CRPC, 118 genes were found, including 81 genes with positive correlation and 37 genes with negative correlation with PFS (Figure 4).

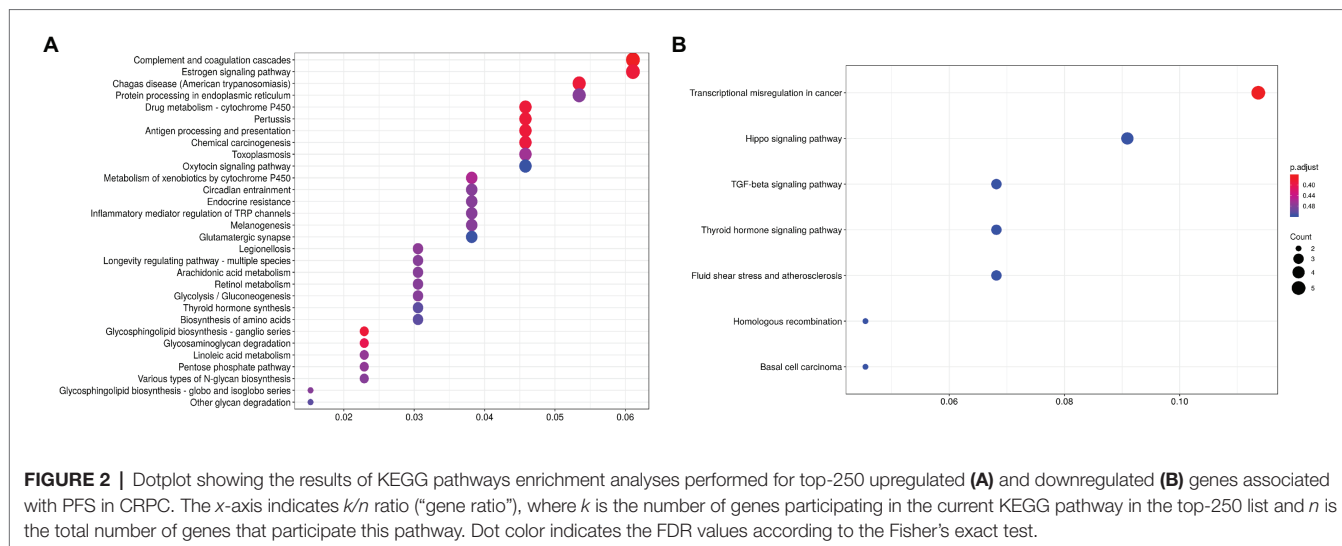
### Interaction Network Analysis Using STRINGdb

To analyze the LAPC, we examined 190 genes, expression of which is negatively, and 115 genes – positively associated with PFS.

These genes passed a threshold of  $p < 0.05$  (both QLF test and Spearman’s rank correlation analysis), irrespectively of CPM and abs(LogFC) values. A statistically significant result was obtained for genes with a negative correlation. For these genes, the network has a strong enrichment of protein-protein interactions (PPI) with  $p < 1.0e-16$  (Figure 5).

The genes shown in Figure 5 encode for proteins participating in such processes as the “Cell cycle” (GO: 0007049; enrichment FDR = 2.24e-15), “Organelle organization” (GO: 0006996; FDR = 8.53e-06), “Response to stress” (GO: 0006950; FDR = 0.0490) terms. We have also noted the cluster with the highest number of interactions, including *RAD21, SGOL1, DYNC1L1, CKAP5, CENPI, PPP2R5C, SKA2, SPC25, MAD2L1, CENPL, BUB1, ERCC6L, SGOL2,* and *CDCA8* genes. Most of them participate cell cycle process.

Next, we performed a protein-protein interaction network analysis using 196 genes with negative expression correlation with PFS for CRPC. The analysis (PPI enrichment  $p = 0.000112$ ) revealed involved enrichment in genes participating the “Cell cycle” (GO: 0007049; enrichment FDR = 0.0207), “Regulation of gene expression” (GO: 0010468; FDR = 0.0215), and “Metabolic process” (GO: 0008152; FDR = 0.0345) terms (Figure 6).



**TABLE 2 |** KEGG pathways that are significantly enriched with genes associated with PFS in castration-resistant prostate cancer (CRPC).

KEGG ID	Pathway name	Enrichment p-value	Enrichment FDR	Genes
<b>Enriched with upregulated genes</b>				
hsa04610	Complement and coagulation cascades	8,14E-05	1,88E-02	<i>CPB2, SERPIND1, CFI, VTN, KNG1, SERPING1, SERPINA1, C3</i>
hsa00982	Drug metabolism – cytochrome P450	1,45E-04	1,88E-02	<i>ADH6, ADH1B, CYP3A4, CYP2E1, CYP2C8, GSTP1</i>
<b>Enriched with downregulated genes</b>				
hsa05202	Transcriptional misregulation in cancer	3,28E-03	3,65E-01	<i>HIST1H3C, TAF15, ASPSCR1, NCOR1, EYA1</i>
hsa04390	Hippo signaling pathway	1,04E-02	5,17E-01	<i>BMP4, TGFB2, BMP5, NKD1</i>
hsa04350	TGF-beta signaling pathway	1,45E-02	5,17E-01	<i>BMP4, TGFB2, BMP5</i>

The interaction network formed with genes positively correlated with PFS did not pass the PPI enrichment *p*-value threshold.

## DISCUSSION

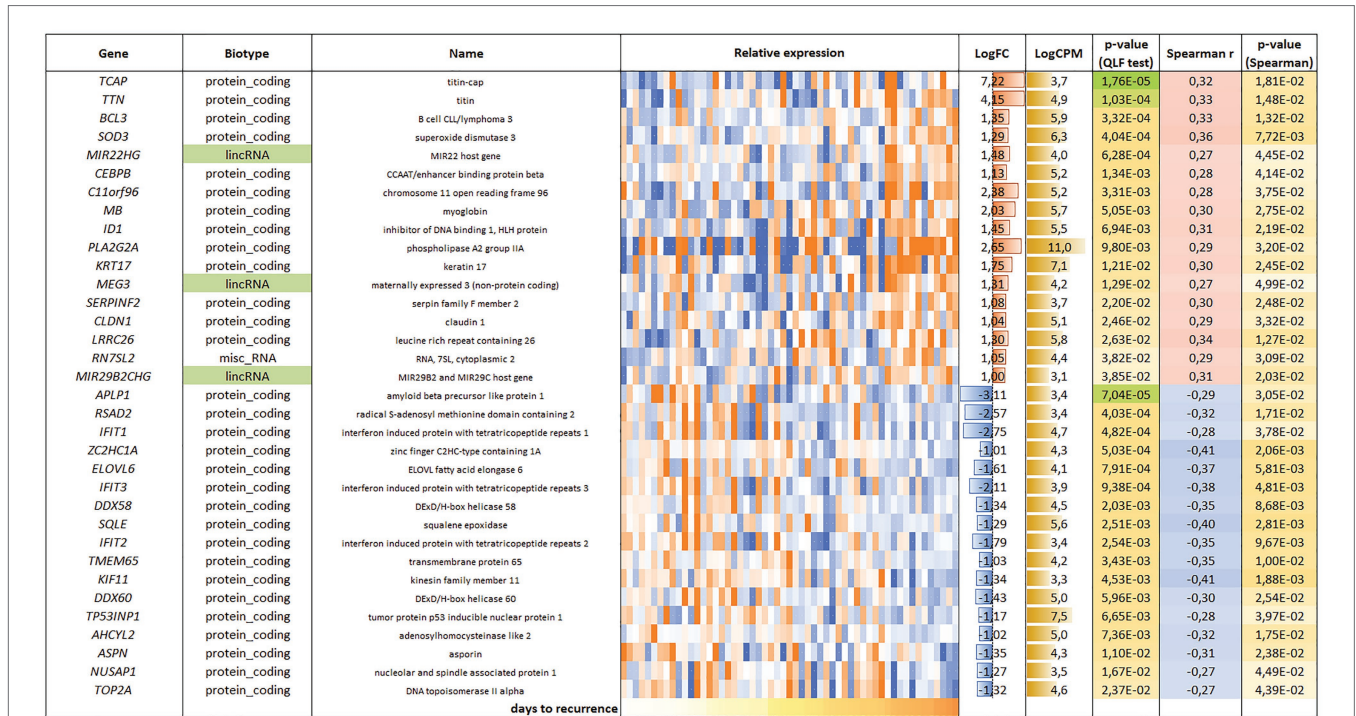
In the present study, we examined RNA-Seq data from the two most advanced stages of PC, LAPC and CRPC, with an aim to reveal the major changes in the transcriptome associated with disease progression within each analyzed category. The pathological manifestation of LAPC is invasion of the prostatic capsule as well as invasion beyond it, whereas CRPC, which is the next progressive stage, is characterized by the presence of distant metastases.

Using KEGG pathway enrichment analysis, we found several pathways, activation of which is possibly associated with PFS period, either positively or negatively. In LAPC, the most prominent associations were the increased expression of the participants of “TNF signaling pathway” and “Transcriptional misregulation in cancer.” Transcriptional misregulation involves an extensive network of processes that play dual roles in cancers, and probably in LAPC.

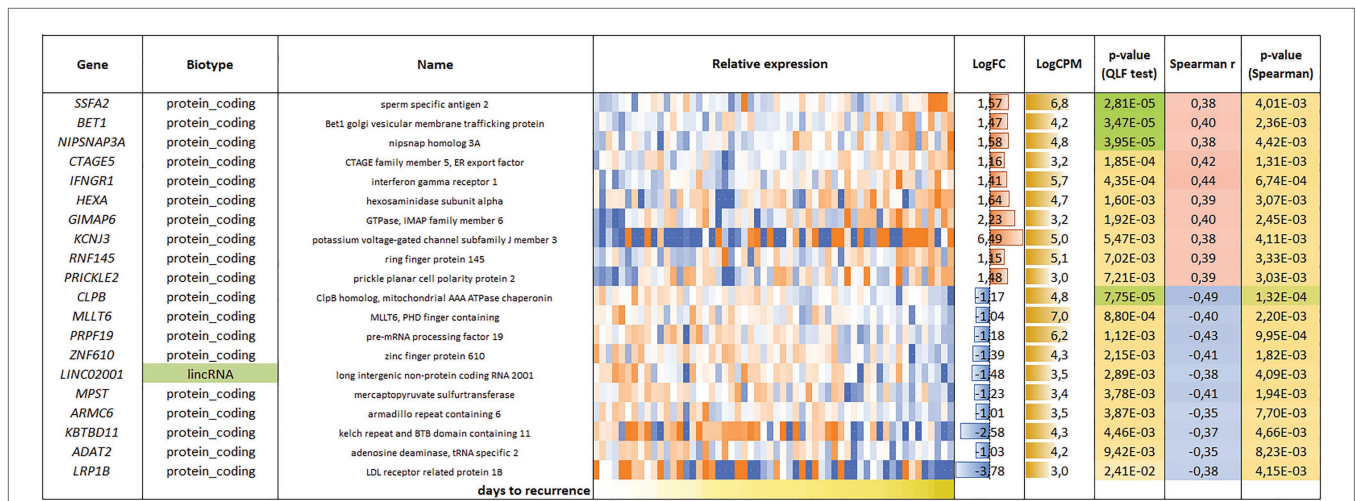
In CRPC, we identified putative activation of “Complement and coagulation cascades” and “Drug metabolism-cytochrome

P-450” KEGG pathways to be associated with increased PFS time. According to various studies, the expression levels of complement system genes vary in different cancers. In PC, upregulated expression of complement system genes in both the classical and alternative pathways is associated with a good prognosis and long-term patient survival (Roumenina et al., 2019). Based on our results, putative activation of the complement pathway increased the relapse time in CRPC. Therefore, in patients with activation of the complement pathway genes, a longer response to therapy is expected, which may lead to the elimination of tumor-induced immunosuppression and increased antitumor immunity. Among the identified pathways enriched with genes negatively associated with PFS in CRPC, we noticed several cancer-associated pathways such as the “Hippo signaling pathway” and “TGF-beta signaling pathway” as well as “Transcriptional misregulation in cancer.”

Next, we performed PPI network and GO enrichment analysis using the STRINGdb. We examined the interaction networks for sets of genes that were positively and negatively correlated with PFS. In both datasets, statistically significant results were obtained only for genes with negative correlations. In LAPC, an extensive network of protein-protein interactions was identified for downregulated genes, which suggests close biological relationships among the genes and proteins under study.



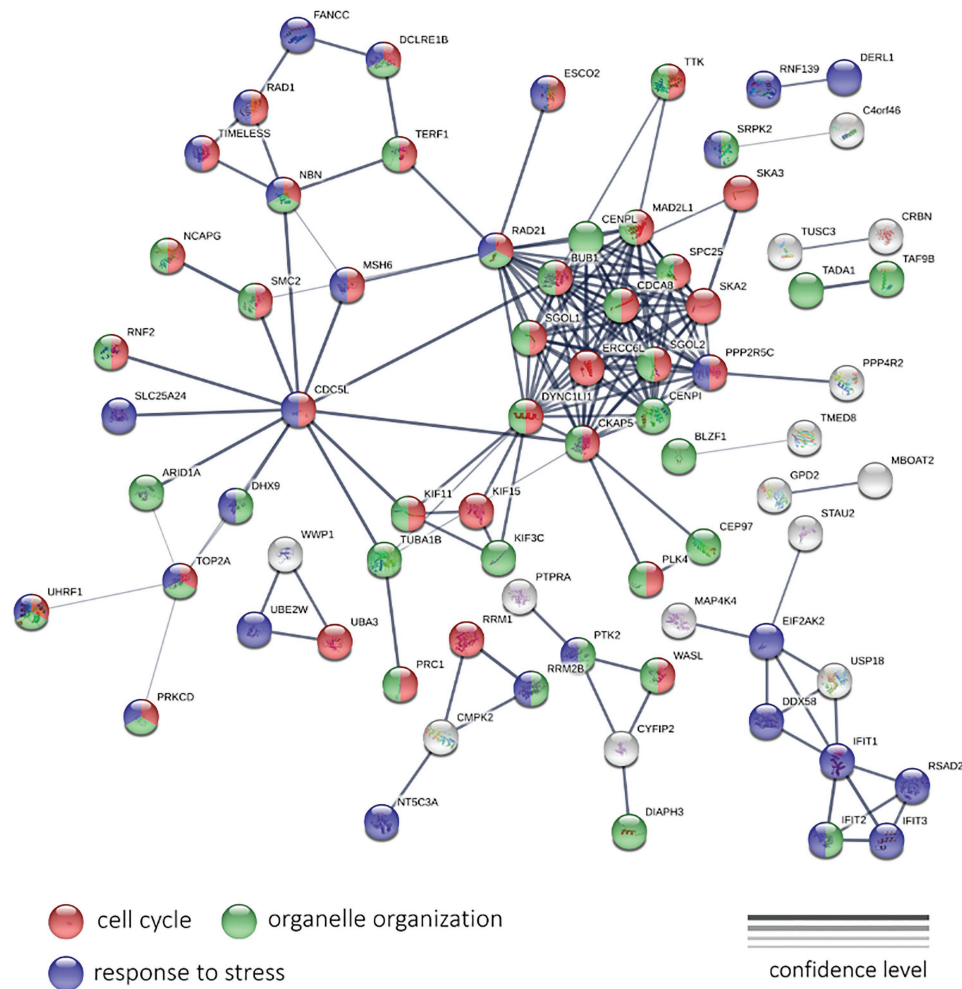
**FIGURE 3 |** List and heatmap demonstrating log relative expression level of genes with significant Spearman’s rank correlation coefficient relative to PFS for LAPC. Cell colors (blue-white-orange gradient) correspond to the binary logarithm of the ratio of the expression level in a current sample to the average level across all the samples (per each gene). Blue – expression level is below the average, orange – above the average. LogFC – binary logarithm of the ratio of GLM-approximated expression values between samples with maximal and minimal PFS; LogCPM – binary logarithm of read counts per million (CPM); p-value (QLF test) – p-value according to quasi-likelihood F-test (edgeR); Spearman r – Spearman’s rank correlation coefficient between gene expression level and PFS; p-value (Spearman) – p-value according to Spearman’s rank correlation test.



**FIGURE 4 |** List and heatmap demonstrating log relative expression level of genes with significant Spearman’s rank correlation coefficient relative to PFS for CRPC. For comments, see the Figure 3 legend.

We observed significant enrichment with participants of several biological processes such as “Cell cycle,” “Organelle organization,” and “Response to stress.” In CRPC, we observed

strong enrichment with genes involved “Cell cycle,” “Regulation of gene expression,” and “Metabolic process.” This suggests that the activation of the cell cycle and an increase in the

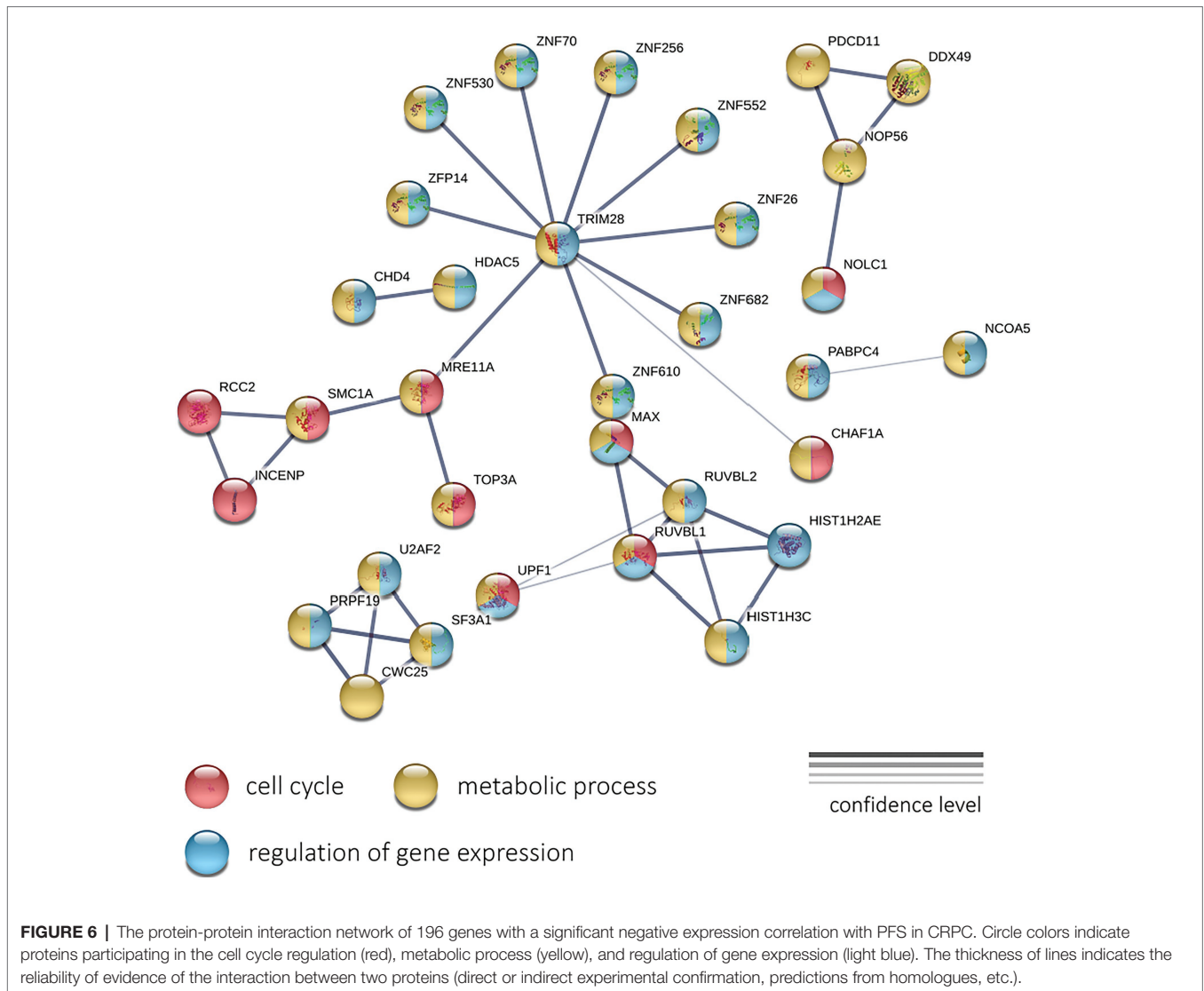


**FIGURE 5 |** The protein-protein interaction network of 190 genes with a significant negative expression correlation with PFS in LAPC. Circle colors indicate proteins participating in the cell cycle regulation (red), organelle organization (green), and response to stress (dark blue). If a gene does not belong to these categories, it is marked with gray. The thickness of lines indicates the reliability of evidence of the interaction between two proteins (direct or indirect experimental confirmation, predictions from homologues, etc.).

expression of cell cycle-related genes may be behind the formation of more aggressive tumor phenotype and shortened PFS period, for both LAPC and CRPC. This is to be expected. Needless to say once again, the cell cycle, in particular cellular mitosis, and cancer are closely related, as cancer cells undergo abnormal, uncontrolled mitosis, which supports tumor growth and metastasis, two processes that are integral to disease progression (Weaver and Cleveland, 2005).

In our study, we also focused on identifying potential prognostic markers among the genes that have the strongest correlations with PFS. In LAPC, *KIF11* (Spearman's rank correlation coefficient,  $r_s = 0.41$ ), *ZC2HC1A* ( $r_s = -0.41$ ), and *SQLE* ( $r_s = -0.40$ ) had the greatest negative correlations with PFS, while *SOD3* ( $r_s = 0.36$ ) and *LRRC26* ( $r_s = 0.34$ ) had the highest positive correlations. We reviewed the literature regarding the role of these genes in cancer, especially PC.

The kinesin family member 11 (*KIF11*) gene encodes mitotic kinesin, which plays a central role in mitosis (Wojcik et al., 2013). In PC, assessment of tumor cell differentiation, which is recorded as the Gleason score, is an important prognostic parameter. A value of 8 or higher corresponds to a poorly differentiated tumor and is associated with an unfavorable prognosis (Gordetsky and Epstein, 2016). A value of 8 or higher corresponds to a poorly differentiated tumor and is associated with an unfavorable prognosis. It was shown that *KIF11* gene expression was higher in PC tumor samples with a Gleason score of 8 (poorly differentiated tumors) than in tumor samples with a Gleason score of 7 (moderately differentiated tumors; Piao et al., 2017). In our study, we found that *KIF11* gene expression was negatively correlated with an increase in the number of days before recurrence, which confirms the relationship of this gene with unfavorable prognosis in PC.



There are few data regarding zinc finger C2HC-type containing 1A (*ZC2HC1A*). However, Zhu et al. (2019) showed that overexpression of the *ZC2HC1A* gene was associated with unfavorable prognosis in hepatocellular carcinoma. In our study, *ZC2HC1A* expression was increased in PC samples with early onset recurrence.

The *SQLE* gene encodes squalene epoxidase, which is involved in cholesterol synthesis. Cholesterol is an essential component of cell membranes and a precursor for the synthesis of androgens (Pelton et al., 2012). According to previous data, the *SQLE* gene is expressed in aggressive PC, and its expression is correlated with the Gleason score. It has been suggested that the progression of PC depends on the *de novo* synthesis of cholesterol catalyzed by *SQLE* (Stopsack et al., 2017). In our study, we observed higher *SQLE* gene expression in samples with early onset relapse.

Prostate cancer can also progress due to oxidative stress, which produces reactive oxygen species (ROS; Shiota et al., 2011). Cells are protected against ROS by antioxidant enzymes, such

as superoxide dismutase (SOD), which functions as a first line antioxidant enzyme. The *SOD3* gene encodes an extracellular isoform of the enzyme, and data suggest that *SOD3* acts as a tumor suppressor in PC (Faraci and Didion, 2004). Thus, the effect of *SOD3* gene expression on cell proliferation, migration, and invasion of PC-3 cells was assessed, and the results showed that overexpression of *SOD3* inhibits these processes (Kim et al., 2014). Our data also demonstrate that *SOD3* gene expression was increased relative to time to relapse and was lower in tumor samples from patients with an unfavorable prognosis.

Studies on the function of *LRRC26* gene in LNCaP cell culture have shown that its overexpression leads to suppression of the NF- $\kappa$ B pathway, which is involved in cancer progression and metastasis (Liu et al., 2012). Thus, decreased expression of the *LRRC26* gene is associated with an unfavorable prognosis in PC, which is consistent with our results.

Correlation analysis of gene expression with PFS in CRPC revealed that the following genes had the strongest Spearman

correlation coefficients: *CLPB* ( $r_s = -0.49$ ), *PRPF19* ( $r_s = -0.43$ ), *MPST* ( $r_s = 0.41$ ), *IFNGR1* ( $r_s = 0.44$ ), *CTAGE5* ( $r_s = 0.42$ ), *GIMAP6* ( $r_s = 0.40$ ), and *BET1* ( $r_s = 0.40$ ).

The caseinolytic mitochondrial matrix peptidase chaperone subunit B (*CLPB*) is an ATPase associated with a variety of cellular processes. Currently, there are no data on the role of this gene in cancer. Our study showed that an increase in expression was associated with an unfavorable prognosis, as evidenced by the strong correlation and high statistical significance.

The pre-mRNA processing factor 19 (*PRPF19*) gene encodes the hPrp19 protein, which is involved in many physiological processes, such as the ubiquitin-proteasome system, DNA damage response, proliferation, and apoptosis (Yin et al., 2012). The hPrp19 protein has been reported to play a potential pro-oncogenic role due to its proliferation-promoting activity. hPrp19 is also required for the expression of p21, which has an intense cell cycle arrest-promoting effect (Chen et al., 2011). The mechanism of hPrp19 in cancer is still unclear; however, its involvement in DNA repair is expected to be correlated with tumor progression. According to our results, in CRPC, increased expression of the *PRPF19* gene is associated with early relapse.

The mercaptopyruvate sulfurtransferase (*MPST*) gene encodes an enzyme involved in the catalysis of endogenous hydrogen sulfide from L-cysteine. Various studies have shown that endogenous hydrogen sulfide can regulate the occurrence and development of tumors and can participate in cancer progression by stimulating angiogenesis and cell growth in colon and ovarian cancer (Bhattacharyya et al., 2013; Szabo et al., 2013; Hellmich and Szabo, 2015). In CRPC, we observed an increase in *MPST* expression at relapse, which suggests the involvement of this gene in PC progression.

Interferon-gamma receptor 1 (*IFNGR1*) encodes a subunit of the IFN- $\gamma$  receptor that acts in IFN- $\gamma$  pathways and regulates the immune response. Reduced expression of this gene was observed in MYC-dependent metastatic PC. Experiments that activate *IFNGR1* gene expression demonstrated strong activation of tumor-suppression signaling and sustained apoptosis (Wee et al., 2014). We observed similar results in CRPC, which showed decreased expression of the *IFNGR1* gene in tumor samples with early relapse.

The CTAGE family member 5 (*CTAGE5*) gene product (an ER export factor) is involved in the transport of collagen VII in the endoplasmic reticulum. In PC, the *CTAGE5* gene is involved in tumor-specific splicing (Ren et al., 2012). Our study showed a decrease in the expression of *CTAGE5* in samples with early relapse.

The GTPase, IMAP family member 6 (*GIMAP6*) gene encodes a protein belonging to the GIMAP family of proteins, which are mainly expressed in immune system cells and contribute to the development of thymocytes, apoptosis of peripheral lymphocytes, and T-helper differentiation (Filen et al., 2009). Dysregulation of *GIMAP6* expression is observed in non-small cell lung cancer, where the *GIMAP6* gene expression is decreased in tumor samples (Shiao et al., 2008). Decreased expression of the *GIMAP6* gene has also been

noted in hepatocellular carcinoma (Huang et al., 2016). In CRPC, we also observed that decreased expression of *GIMAP6* was associated with the onset of disease relapse.

Bet1 Golgi vesicular membrane trafficking protein (*BET1*) is a membrane protein associated with the Golgi complex that is involved in vesicular transport. There are few published studies on the function of this gene in cancer. There is a report that *BET1* was identified as part of a gene expression signature associated with a favorable prognosis in glioblastoma (Cai et al., 2020). We showed that decreased *BET1* expression was associated with early relapse in CRPC. In addition to the identified protein-coding genes, in our correlation analysis, we also found long non-coding RNAs (lncRNAs) with altered expression patterns. lncRNAs are involved in the regulation of various biological processes, including many cancer-associated pathways (Spizzo et al., 2012; Flynn and Chang, 2014; Quinn and Chang, 2016). Aberrant lncRNAs expression has been observed in a variety of cancers, and many studies have shown a link between these molecules and cancer development and progression. In LAPC, we found changed expression of the following lncRNAs that were associated with PFS: *MIR22HG* ( $r_s = 0.27$ ), *MEG3* ( $r_s = 0.27$ ), and *MIR29B2CHG* ( $r_s = 0.31$ ).

Regarding the role of MIR22 host gene (*MIR22HG*) in PC, decreased *MIR22HG* expression has been shown to be significantly associated with a higher Gleason score and shorter PFS time, highlighting its prognostic potential (Shen et al., 2019). In our analysis, we also observed an association between decreased *MIR22HG* expression and early onset of disease recurrence.

Studies on maternally expressed 3 (*MEG3*) have shown that this lncRNA inhibits the proliferation and metastasis of gastric cancer through p53 signaling (Wei and Wang, 2017). In PC, *MEG3* is characterized by a downregulated expression, and increased expression has an inhibitory effect on tumor growth (Wu et al., 2019). Therefore, our results are consistent with the published data.

There are currently no data on the role of MIR29B2 and MIR29C host gene (*MIR29B2CHG*) in cancer. Therefore, we report here, for the first time, the involvement of *MIR29B2CHG* in the progression of LAPC.

In CRPC, we identified an association between an increase in the expression of *LINC02001* (small nucleolar RNA host gene 30 [*SNHG30*];  $r_s = -0.38$ ) and early onset of disease relapse. No studies on the involvement of *LINC02001* in cancer have been reported.

Summing up, in this study, genes and pathways associated with PFS in advanced-stage PC (LAPC and CRPC) were identified. Our results are consistent with the previous studies that reported the participation of *KIF11*, *SQLE*, *SOD3*, *LRRC26*, *IFNGR1*, *MIR22HG*, and *MEG3* in the carcinogenesis and progression of PC. The possible association with the progression of PC was first shown for genes *ZC2HC1A*, *CLPB*, *PRPF19*, *MPST*, *GIMAP6*, *BET1*, *MIR29B2CHG*, and *LINC02001*. All listed genes showed strong correlations with PFS and thus could be considered as potential prognostic markers.



## DATA AVAILABILITY STATEMENT

The study included RNA-Seq data for LAPC samples from the PRAD project ( $n = 55$ ) and CRPC samples from the WCDT-MCRPC project ( $n = 84$ ), which were donated to the TCGA consortium.

## AUTHOR CONTRIBUTIONS

EAP, AVK, KMN, ADK, BYA and GTS conceived and designed the work. GSK, AAK, and VSP carried out bioinformatics analysis. MVS, MSF, and DYT carried out statistical analysis. EAP, AVS, and GSK wrote the manuscript. All authors contributed to the article and approved the submitted version.

## FUNDING

This research was funded by the Russian Science Foundation (grant 18-75-10127) and RFBR (grant 17-29-06085).

## REFERENCES

- Bhattacharyya, S., Saha, S., Giri, K., Lanza, I. R., Nair, K. S., Jennings, N. B., et al. (2013). Cystathionine beta-synthase (CBS) contributes to advanced ovarian cancer progression and drug resistance. *PLoS One* 8:e79167. doi: 10.1371/journal.pone.0079167
- Cai, H. Q., Liu, A. S., Zhang, M. J., Liu, H. J., Meng, X. L., Qian, H. P., et al. (2020). Identifying predictive gene expression and signature related to temozolomide sensitivity of glioblastomas. *Front. Oncol.* 10:669. doi: 10.3389/fonc.2020.00669
- Cao, Z. X., Xiao, G. A., Zhang, W., Ji, J., Ye, C., Liu, D., et al. (2019). Comprehensive investigation of alternative splicing and development of a prognostic risk score for prostate cancer based on six-gene signatures. *J. Cancer* 10, 5585–5596. doi: 10.7150/jca.31725
- Chen, Y., Zhang, L., and Jones, K. A. (2011). SKIP counteracts p53-mediated apoptosis via selective regulation of p21Cip1 mRNA splicing. *Genes Dev.* 25, 701–716. doi: 10.1101/gad.2002611
- Faraci, F. M., and Didion, S. P. (2004). Vascular protection: superoxide dismutase isoforms in the vessel wall. *Arterioscler. Thromb. Vasc. Biol.* 24, 1367–1373. doi: 10.1161/01.ATV.0000133604.20182.cf
- Filen, J. J., Filen, S., Moulder, R., Tuomela, S., Ahlfors, H., West, A., et al. (2009). Quantitative proteomics reveals GIMAP family proteins 1 and 4 to be differentially regulated during human T helper cell differentiation. *Mol. Cell. Proteomics* 8, 32–44. doi: 10.1074/mcp.M800139-MCP200
- Flynn, R. A., and Chang, H. Y. (2014). Long noncoding RNAs in cell-fate programming and reprogramming. *Cell Stem Cell* 14, 752–761. doi: 10.1016/j.stem.2014.05.014
- Forsythe, A., Chandiwana, D., Barth, J., Thabane, M., Baeck, J., and Tremblay, G. (2018). Progression-free survival/time to progression as a potential surrogate for overall survival in HR+, HER2– metastatic breast cancer. *Breast Cancer* 10, 69–78. doi: 10.2147/BCTT.S162841
- Gordetsky, J., and Epstein, J. (2016). Grading of prostatic adenocarcinoma: current state and prognostic implications. *Diagn. Pathol.* 11:25. doi: 10.1186/s13000-016-0478-2
- Hellmich, M. R., and Szabo, C. (2015). Hydrogen sulfide and cancer. *Handb. Exp. Pharmacol.* 230, 233–241. doi: 10.1007/978-3-319-18144-8\_12
- Huang, Z., Zhang, W., Gao, C., Ji, B., Chi, X., Zheng, W., et al. (2016). Dysregulation of GTPase IMAP family members in hepatocellular cancer. *Mol. Med. Rep.* 14, 4119–4123. doi: 10.3892/mmr.2016.5764
- Keyes, M., Crook, J., Morton, G., Vigneault, E., Usmani, N., and Morris, W. J. (2013). Treatment options for localized prostate cancer. *Can. Fam. Physician* 59, 1269–1274.

## ACKNOWLEDGMENTS

We thank the National Medical Research Center for Obstetrics, Gynecology and Perinatology named after Academician V.I. Kulakov and the Center for Precision Genome Editing and Genetic Technologies for Biomedicine, EIMB RAS for providing the computing power and techniques for the data analysis. This work was performed using the equipment of EIMB RAS “Genome” center ([http://www.eimb.ru/ru1/ckp/ccu\\_genome\\_c.php](http://www.eimb.ru/ru1/ckp/ccu_genome_c.php)).

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fgene.2020.613162/full#supplementary-material>

**SUPPLEMENTARY TABLE S1** | List statistically significant genes according to quasi-likelihood F-test for the LAPC.

**SUPPLEMENTARY TABLE S2** | List statistically significant genes according to quasi-likelihood F-test for the CRPC.

- Kim, J., Mizokami, A., Shin, M., Izumi, K., Konaka, H., Kadono, Y., et al. (2014). SOD3 acts as a tumor suppressor in PC-3 prostate cancer cells via hydrogen peroxide accumulation. *Anticancer Res.* 34, 2821–2831.
- Larkin, S. E., Holmes, S., Cree, I. A., Walker, T., Basketter, V., Bickers, B., et al. (2012). Identification of markers of prostate cancer progression using candidate gene expression. *Br. J. Cancer* 106, 157–165. doi: 10.1038/bjc.2011.490
- Liu, X. F., Xiang, L., Zhang, Y., Becker, K. G., Bera, T. K., and Pastan, I. (2012). CAPC negatively regulates NF-kappaB activation and suppresses tumor growth and metastasis. *Oncogene* 31, 1673–1682. doi: 10.1038/onc.2011.355
- Pelton, K., Freeman, M. R., and Solomon, K. R. (2012). Cholesterol and prostate cancer. *Curr. Opin. Pharmacol.* 12, 751–759. doi: 10.1016/j.coph.2012.07.006
- Piao, X. -M., Byun, Y. J., Jeong, P., Ha, Y. -S., Yoo, E. S., Yun, S. J., et al. (2017). Kinesin family member 11 mRNA expression predicts prostate cancer aggressiveness. *Clin. Genitourin. Cancer* 15, 450–454. doi: 10.1016/j.clgc.2016.10.005
- Pudova, E. A., Lukyanova, E. N., Nyushko, K. M., Mikhaylenko, D. S., Zaretsky, A. R., Snezhkina, A. V., et al. (2019). Differentially expressed genes associated with prognosis in locally advanced lymph node-negative prostate cancer. *Front. Genet.* 10:730. doi: 10.3389/fgene.2019.00730
- Quinn, J. J., and Chang, H. Y. (2016). Unique features of long non-coding RNA biogenesis and function. *Nat. Rev. Genet.* 17, 47–62. doi: 10.1038/nrg.2015.10
- Rawla, P. (2019). Epidemiology of prostate cancer. *World J. Oncol.* 10, 63–89. doi: 10.14740/wjon1191
- Ren, S., Peng, Z., Mao, J. H., Yu, Y., Yin, C., Gao, X., et al. (2012). RNA-seq analysis of prostate cancer in the Chinese population identifies recurrent gene fusions, cancer-associated long noncoding RNAs and aberrant alternative splicings. *Cell Res.* 22, 806–821. doi: 10.1038/cr.2012.30
- Robinson, M. D., McCarthy, D. J., and Smyth, G. K. (2010). edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* 26, 139–140. doi: 10.1093/bioinformatics/btp1616
- Roumenina, L. T., Daugan, M. V., Petitprez, F., Sautes-Fridman, C., and Fridman, W. H. (2019). Context-dependent roles of complement in cancer. *Nat. Rev. Cancer* 19, 698–715. doi: 10.1038/s41568-019-0210-0
- Shen, H., Weng, X. D., Yang, D., Wang, L., and Liu, X. H. (2019). Long noncoding RNA MIR22HG is down-regulated in prostate cancer. *Math. Biosci. Eng.* 17, 1776–1786. doi: 10.3934/mbe.2020093
- Shiao, Y. M., Chang, Y. H., Liu, Y. M., Li, J. C., Su, J. S., Liu, K. J., et al. (2008). Dysregulation of GIMAP genes in non-small cell lung cancer. *Lung Cancer* 62, 287–294. doi: 10.1016/j.lungcan.2008.03.021

- Shiota, M., Yokomizo, A., and Naito, S. (2011). Oxidative stress and androgen receptor signaling in the development and progression of castration-resistant prostate cancer. *Free Radic. Biol. Med.* 51, 1320–1328. doi: 10.1016/j.freeradbiomed.2011.07.011
- Song, Z. Y., Chao, F., Zhuo, Z., Ma, Z., Li, W., and Chen, G. (2019). Identification of hub genes in prostate cancer using robust rank aggregation and weighted gene co-expression network analysis. *Aging* 11, 4736–4756. doi: 10.18632/aging.102087
- Spizzo, R., Almeida, M. I., Colombatti, A., and Calin, G. A. (2012). Long non-coding RNAs and cancer: a new frontier of translational research? *Oncogene* 31, 4577–4587. doi: 10.1038/onc.2011.621
- Stopsack, K. H., Gerke, T. A., Andr en, O., Andersson, S. -O., Giovannucci, E. L., Mucci, L. A., et al. (2017). Cholesterol uptake and regulation in high-grade and lethal prostate cancers. *Carcinogenesis* 38, 806–811. doi: 10.1093/carcin/bgx058
- Szabo, C., Coletta, C., Chao, C., Modis, K., Szczesny, B., Papapetropoulos, A., et al. (2013). Tumor-derived hydrogen sulfide, produced by cystathionine-beta-synthase, stimulates bioenergetics, cell proliferation, and angiogenesis in colon cancer. *Proc. Natl. Acad. Sci. U. S. A.* 110, 12474–12479. doi: 10.1073/pnas.1306241110
- Szklarczyk, D., Morris, J. H., Cook, H., Kuhn, M., Wyder, S., Simonovic, M., et al. (2017). The STRING database in 2017: quality-controlled protein-protein association networks, made broadly accessible. *Nucleic Acids Res.* 45, D362–D368. doi: 10.1093/nar/gkw937
- Weaver, B. A., and Cleveland, D. W. (2005). Decoding the links between mitosis, cancer, and chemotherapy: the mitotic checkpoint, adaptation, and cell death. *Cancer Cell* 8, 7–12. doi: 10.1016/j.ccr.2005.06.011
- Wee, Z. N., Li, Z., Lee, P. L., Lee, S. T., Lim, Y. P., and Yu, Q. (2014). EZH2-mediated inactivation of IFN-gamma-JAK-STAT1 signaling is an effective therapeutic target in MYC-driven prostate cancer. *Cell Rep.* 8, 204–216. doi: 10.1016/j.celrep.2014.05.045
- Wei, G. H., and Wang, X. (2017). lncRNA MEG3 inhibit proliferation and metastasis of gastric cancer via p53 signaling pathway. *Eur. Rev. Med. Pharmacol. Sci.* 21, 3850–3856.
- Wojcik, E. J., Buckley, R. S., Richard, J., Liu, L., Huckaba, T. M., and Kim, S. (2013). Kinesin-5: cross-bridging mechanism to targeted clinical therapy. *Gene* 531, 133–149. doi: 10.1016/j.gene.2013.08.004
- Wu, M., Huang, Y., Chen, T., Wang, W., Yang, S., Ye, Z., et al. (2019). lncRNA MEG3 inhibits the progression of prostate cancer by modulating miR-9-5p/QKI-5 axis. *J. Cell. Mol. Med.* 23, 29–38. doi: 10.1111/jcmm.13658
- Yap, T. A., Smith, A. D., Ferraldeschi, R., Al-Lazikani, B., Workman, P., and De Bono, J. S. (2016). Drug discovery in advanced prostate cancer: translating biology into therapy. *Nat. Rev. Drug Discov.* 15, 699–718. doi: 10.1038/nrd.2016.120
- Yin, J., Zhu, J. M., and Shen, X. Z. (2012). New insights into pre-mRNA processing factor 19: a multi-faceted protein in humans. *Biol. Cell.* 104, 695–705. doi: 10.1111/boc.201200011
- Yu, G., Wang, L. G., Han, Y., and He, Q. Y. (2012). clusterProfiler: an R package for comparing biological themes among gene clusters. *OMICS* 16, 284–287. doi: 10.1089/omi.2011.0118
- Zhu, G. Q., Yang, Y., Chen, E. B., Wang, B., Xiao, K., Shi, S. M., et al. (2019). Development and validation of a new tumor-based gene signature predicting prognosis of HBV/HCV-included resected hepatocellular carcinoma patients. *J. Transl. Med.* 17:203. doi: 10.1186/s12967-019-1946-8

**Conflict of Interest:** The reviewer AT declared a shared affiliation with two of the authors, GS and DT, to the handling editor at time of review.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2021 Pudova, Krasnov, Kobelyatskaya, Savvateeva, Fedorova, Pavlov, Nyushko, Kaprin, Alekseev, Trofimov, Sukhikh, Snezhkina and Kudryavtseva. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.