

Article

Towards Understanding the Key Signature Pathways Associated from Differentially Expressed Gene Analysis in an Indian Prostate Cancer Cohort

Nidhi Shukla^{1,2,†}, Bhumandeep Kour^{3,†}, Devendra Sharma⁴, Maneesh Vijayvargiya⁵, TC Sadasukhi⁶, Krishna Mohan Medicherla¹, Babita Malik², Bhawana Bissa⁸, Sugunakar Vuree^{3,9}, Nirmal Kumar Lohiya⁷ and Prashanth Suravajhala^{9,10*}

¹ Department of Biotechnology and Bioinformatics, Birla Institute of Scientific Research (BISR), Statue Circle, Jaipur, India

² Department of Chemistry, School of basic sciences, Manipal University Jaipur, Jaipur

³ Department of Biotechnology, School of Bioengineering and Biosciences, Lovely Professional University, Punjab, India

⁴ Department of Urology, Rukmani Birla Hospitals, Jaipur, India

⁵ Department of Pathology, Mahatma Gandhi University of Medical Sciences and Technology, Jaipur, India

⁶ Department of Urology, Mahatma Gandhi University of Medical Sciences and Technology, Jaipur, India

⁷ Department of Zoology, Center for advanced studies, University of Rajasthan, Jaipur, India

⁸ Department of Biochemistry, Central University of Rajasthan, Ajmer, India

⁹ Bioclues.org, India

¹⁰ Amrita School of Biotechnology, Amrita Vishwavidyapeetham, Clappana 690525, Kerala, India

[†] Equal contributing authors.

* Correspondence: prash@am.amrita.edu

Abstract: Prostate cancer (PCa) is one of the most prevalent cancers among men in India. Although studies on PCa have dealt with the genetics, genomics, and the environmental influence in causality of PCa, not many studies employing the next generation sequencing (NGS) approaches of PCa have been carried out. In our previous study, we have identified some causal genes and mutations specific to Indian PCa using Whole-Exome Sequencing (WES). In the recent past, with the help of different cancer consortiums such as The Cancer Genome Atlas (TCGA), International Cancer Genome Consortium (ICGC), along with differentially expressed genes (DEGs), many cancer-associated novel non-coding RNAs have been identified as biomarkers. In this work, we attempt to identify DEGs as well as long non-coding RNAs (lncRNAs) associated with signature pathways from an Indian PCa cohort using RNA-sequencing (RNA-seq) approach. From a cohort of 60, we screened 6 patients who underwent prostatectomy; we performed a whole transcriptome shotgun sequencing (WTSS)/RNA-sequencing to decipher the DEGs. We further normalized the read counts using fragments per kilobase of transcript per million mapped reads (FPKM) and analyzed the DEGs using a cohort of downstream regulatory tools, viz. GeneMANIA Stringdb, Cytoscape-Cytohubba, cbiportal to map the inherent signatures associated with PCa. By comparing the RNA-seq data obtained from the pairs of normal and PCa tissue samples using our benchmarked in-house cuffdiff pipeline, we observed some important genes specific to PCa such as STEAP2, APP, PMEPA1, PABPC1, NFE2L2, HN1L and some other important genes known to be involved in different cancer pathways such as, COL6A1, DOK5, STX6, BCAS1, BACE1, BACE2, LMOD1, SNX9, CTNND1 etc. We also identified a few novel lncRNAs such as LINC01440, SOX2OT, ENSG00000232855, ENSG00000287903 and ENSG00000647843.1 that need to be characterized further. Deregulation of SOX2OT is observed in various tumors, including lung cancer, gastric cancer, esophageal cancer, breast cancer, hepatocellular carcinoma, ovarian cancer, pancreatic, laryngeal squamous cell carcinoma, osteosarcoma, nasopharyngeal carcinoma, and glioblastoma. It would be interesting to characterize its function in PCa as well. In comparison with publicly available datasets, we have identified characteristic DEGs and novel lncRNAs implicated in signature PCa pathways in an Indian PCa cohort which have perhaps not been reported. As a pilot study, this has set a precedent for us to validate further experimentally, and we firmly believe this will pave a way towards discovery of biomarkers and development of novel therapies.

Keywords: Prostate Cancer; RNA-Sequencing; differentially expressed genes; long non-coding RNAs,

1. Introduction

In the recent past, there is an increased occurrence of Prostate Cancer (PCa) varying among different races and countries, particularly between western (Caucasian) and eastern (Asian) populations [1] [2]. Over the last few years, there has been an increase in PCa cases across all the different parts and regions in India [3] [4]. Next Generation Sequencing (NGS) has advanced understanding of the diseased phenotypes, particularly cancers with tumor heterogeneity and capture of large amounts of genetic architecture leading to isoforms, splice junctions, post transcriptional modifications with a great accuracy and sensitivity [5] [6]. Although studies have dealt with the genetics, genomics, and the environmental influence in causal of PCa, no association of genotype and phenotype employing the NGS approaches has been discussed in the Indian population. Recent studies from our lab using the whole-exome sequencing (WES) approach of PCa in India have yielded very promising results to identify causal genes specific to PCa with 30 causal genes to the Indian cohort and over 8 genes specific to Indian phenotype with variable degree of genetic disposition [7]. Whole transcriptome shotgun sequencing (WTSS) or RNA-sequencing (RNA-Seq) on the other hand is used to study regulation, developmental biology, clinical and health sector [8]. The expected outcome of RNA-seq is identification of differentially expressed genes (DEGs) in different conditions (for eg. wild type vs mutant or control vs tumor) which could provide a detailed mechanism of an underlying disease [9]. The DEGs are selected based on the cut-off and log2 fold change that depends on p-values obtained by statistical modeling [10]. RNA-Seq generates millions of short reads that are aligned to a reference genome using different alignment softwares and based on this, the characteristic of a particular dataset is calculated [11]. Although RNAs are labile and unstable in alkaline conditions, they can be easily detected and quantified at very low abundance for varying gene expression partners, which makes them suitable to use as biomarkers [12]. Compared with DNA and protein biomarkers, RNA biomarkers have more sensitivity and specificity and are very cost-effective [13]. Similarly, RNA biomarkers have the advantage of providing dynamic insights into cellular states and regulatory processes than DNA biomarkers [14]. With the advent of high-throughput sequencing technologies, different types of non-coding RNAs (e.g., small nuclear RNA, micro RNA, small nucleolar RNA and long non-coding RNA etc.) and protein-coding RNAs (i.e., mRNAs) have been detected [15]. Interestingly, there are lots of novel non-coding RNAs discovered recently out of which mRNAs, piwiRNAs, siRNAs, ceRNAs and miRNAs have been well documented as diagnostic and prognostic markers in different types of cancers (ovary, lung, breast, colorectal) [16]. In addition to small RNAs, long noncoding RNAs (lncRNAs) previously considered as 'transcriptional noises' have been known to have a diverse and significant role in various diseases, primarily cancers [17] [18]. Although well-known lncRNAs in the form of MALAT, HOTAIR, XIST have been reported in genitourinary cancers, the increasing number of cases of PCa worldwide and as well Indian population warrants the need for identifying lncRNA based biomarkers [19] [20]. Further with the advent of RNA-seq, identification of lncRNAs will hopefully facilitate the translational research to the bench side [21]. Therefore, with considerable data available on the Indian population, transcriptome analysis on Indian patients supplementing strategies in diseased phenotypes of PCa will be of great interest [22]. In the current study, we analyzed the RNA-seq data from a cohort of 6 patients and identified DEGs as well as lncRNAs which could lay emphasis as deterministic markers. We discuss the impending results obtained from this study which we attempt to further compare and validate from our downstream analysis.

2. Materials and Methods

2.1. Patients, clinical samples and criteria

As a pilot study, from a cohort of 60, we have screened 6 patients who underwent prostatectomy (4 cases and 2 controls) from Rukmani Birla Hospital (RBH) (Table 1). The inclusion criteria include age > 55 years, non-diabetic or any other metabolic diseases, PSA > 4ng/ml, malignant while the exclusion criteria included the persons with smoking or familial history of benign prostate cases. The control inclusion on the other hand included age above 55, Gleason score <6, PSA <4ng/ml, normal/Benign Prostatic Hyperplasia (BPH) and normal BPH as exclusion criterion. The study carried out through our CA Prostate consortium of India (CAPCI; <https://bioclues.org/capci> last accessed on December 30, 2022) received approval from Institutional Ethics Committee (IEC) of Rukmani Birla Hospitals, Jaipur, India, and informed consent was judiciously taken.

Table 1. Samples used for the WTSS along with their Gleason scores constituted two high grade tumor samples, with two intermediate samples while 2 others with less than 6 from benign cases.

Sample ID	Condition	Gleason Score (primary+secondary)
69/19	Adenocarcinoma	3+4
1631/H19	Adenocarcinoma	4+4
4226/H19	Adenocarcinoma	3+4
5110/H20	Adenocarcinoma	4+4
1374/19	Benign Nodular Prostatic Hyperplasia	<6
1266/19	Benign Nodular Prostatic Hyperplasia	<6

2.2. Tissue preparation and RNA sequencing

RNA was isolated using the RNeasy FFPE kit (Qiagen, Catalog No-73504) from BPH and malignant Formalin-Fixed Paraffin Embedded (FFPE) blocks and sent for sequencing (outsourced). From an approximate 0.5 mg of cross section, the RNA was prepared wherein the NEBNext® Ultra™ II Directional RNA Library Prep Kit was used for preparing the libraries following the manufacturer’s protocol. While 100ng of FFPE RNA was used as input, it was then subjected to end repair and Illumina specific adaptors were ligated. The adaptor ligated product was then barcoded and subjected to 15 cycles of PCR. The samples after PCR were cleaned up using AMPure XP beads with the final libraries checked for quality using Qubit Fluorometer and Agilent Tapestation. The obtained libraries were pooled and diluted to final optimal loading concentration before cluster amplification on the Illumina flow cell. Once the cluster generation is completed, the cluster flow cell is loaded on Illumina HiSeq X instrument to generate 60M, 150bp paired end reads.

2.3. Bioinformatics and downstream RNA-Sequencing analysis

Cufflinks-Cuffdiff pipeline was employed to yield significant changes at the level of transcript expression [23], as we used our benchmarked pipeline from our lab to run through the workflow [24]. The sequences were aligned to human genome reference (build hg38) using HISAT2 to produce the alignment results output in SAM (sequence alignment map/file). HISAT2 aligns a set of unpaired reads (in fastq or .fq format) to the reference genome using the Ferragina and Manzini (FM)-index [25]. Cufflinks uses this map (SAM) and assembles the reads into transcripts, estimates their abundances, and finally examines DEGs from the samples. This was followed by Cuffdiff to check the DEGs which compared the aligned reads from RNA-seq samples from two or more conditions and identified transcripts that are differentially expressed using a rigorous FPKM normalization/statistical analysis [26]. These tools have gained worldwide attention and have been used in a number of transcriptomics studies. In the current study, Cuffdiff was used to perform differential analysis between the control samples and the other 4 malignant samples respectively [27]. A p-value cutoff of 0.05 and less was used to identify the significantly expressed transcripts (Figure 1). We performed a real-time PCR validation for some selected DEGs. The primers were ordered accordingly and RT-PCR was performed. We saw a significant difference in the expression of a few genes in malignant samples compared to control ones. But, some of the genes did not show any difference which could be due to small sample size as well as poor quality of FFPE blocks.

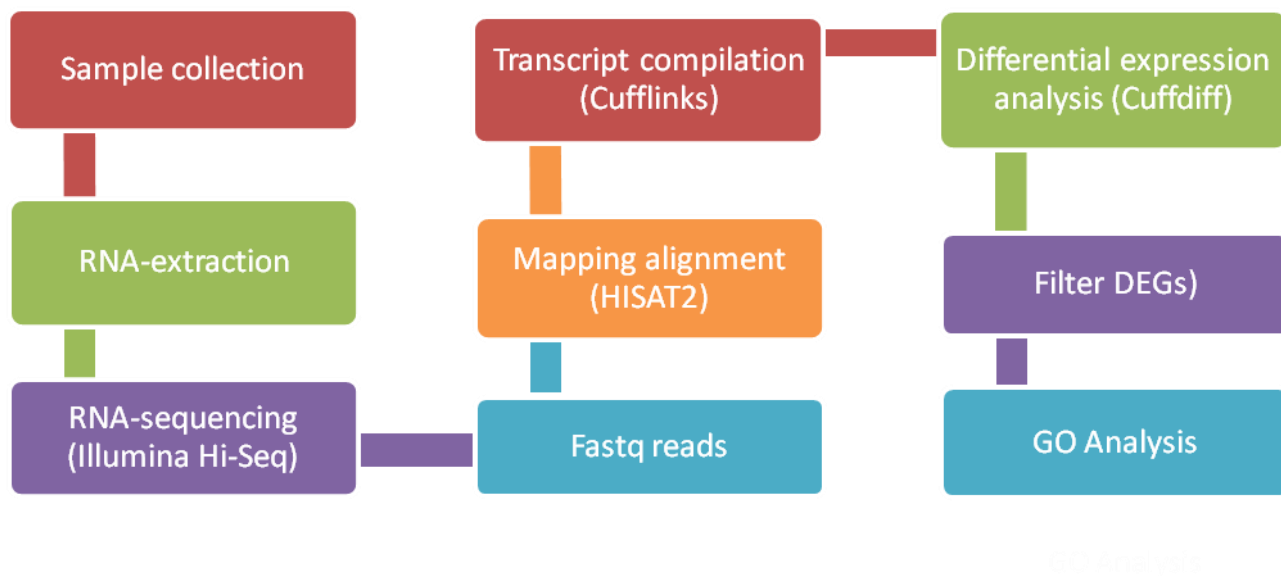


Figure 1. Pictorial representation of RNA-seq workflow depicting different steps of analysis. The input consists of FASTQ files of the sample (control and malignant), human reference genome sequence file (hg38) and gene annotations from gtf files.

2.4. Interaction networks, statistical analysis, gene ontology and cbiportal analyses

We generated an interaction network considering a flexible and intuitive approach for generating gene function hypotheses, evaluating gene lists, and choosing genes for functional studies [28]. To check this, GeneMANIA was employed which overcomes the limitations of earlier methods for predicting gene function on yeast and animal benchmarks using GeneMANIA (<https://genemania.org/> last accessed on December 26, 2022) [36]. In addition, we also used the search tool for the retrieval of interacting genes (STRING) database (<https://string-db.org/>), and integrated DEGs to STRING to evaluate the interaction. Experimentally valid interactions with a score of minimum 0.4 (40% or more) was chosen to be an ideal one with the resulting file saved as a tab-separated values (TSV) file [29]. The raw data files from the STRING database were then imported into Cytoscape 3.5.1 and the cytoHubba (<http://hub.iis.sinica.edu.tw/cytohubba/> last accessed on December 22, 2022) plugin was employed with clustering coefficient, betweenness and closeness centralities to calculate the significant modules in the PPI network [30]. Different Cytoscape plugins can score and rank the nodes using different algorithms [31]. CytoHubba is one such plugin which uses a simple interface to analyze the different networks. CytoHubba implements eleven nodes such as degree, betweenness, closeness, clustering coefficient, stress etc. to rank any network [32]. PANTHER (<http://www.pantherdb.org/> last accessed on December 22, 2022) is a gene ontology based functional annotation tool which takes a variety of inputs such as Gene IDs, UniProtKB IDs, Ensembl IDs etc. and results in either functional analysis or statistical enrichment analysis [33]. Furthermore, we deemed to check the expression and mutational profile of some of the genes from our study in cbiportal (<https://www.cbiportal.org> last accessed on December 22, 2022) for Cancer Genomics provides visualization, analysis and download of large-scale cancer genomics data sets [34]. We used data from TCGA, Pan cancer Atlas of Prostate Adenocarcinoma where 489 samples/patients were screened [35]. The statistics for the aforementioned analyses were based on network analyses with every approach having a function F attached to it that gives each node v a numerical value (v). If a node's score — $F(u)$ — is higher than another node's score — $F(v)$ we say that node u has a higher ranking than node v [32] [38].

3. Results

3.1. Distinct DEGs were obtained

Although there has been a difficulty in extracting RNA from FFPE blocks which is well known [37], from the RNA-seq and downstream analyses, all the samples yielded good quality reads from FastQC with ca. 40M transcript reads with no exposure in tiles. With ensuing cufflinks-cuffdiff pipeline, we obtained approximately 70 DEGs among which 65 were up-regulated and 5 were down-regulated with an inherent p-value heuristics ≤ 0.05 and $\leq -2 \text{ Log}_2\text{FC} \geq \text{Log}_2\text{FC}$. (Table 2).

Table 2. List of top up-regulated DEGs after cuffdiff pipeline. Some of the identified DEGs were PCa specific such as APP, STEAP2, PABPC1, RPS24 etc.

gene id	locus	log2 fold	p-value	gene name
CUFF.13238	chr10:78033882-78040677	6.18617	0.0252	RPS24
CUFF.101498	chr8:100713839-100714204	3.23209	0.03765	PABPC1
CUFF.17462	chr11:57782675-57782947	4.93401	0.0003	CTNND1
CUFF.134059	chr5:55937823-55938153	4.64641	0.0334	IL6ST
CUFF.157342	chr7:90232992-90235254	3.248	0.03545	STEAP2
CUFF.102981	chr21:46003384-46005044	7.4652	0.0004	COL6A1
CUFF.119448	chr3:181441034-181441358	3.356307	0.0131	SOX2OT
CUFF.101277	chr21:25880415-25881777	5.6321	0.03945	APP
CUFF.41953	chr17:48060789-48061102	4.18946	0.04285	NFE2L2
CUFF.110085	chrX:74232840-74233182	3.23919	0.03765	FTX

Among the top niche specific DEGs, collagen type VI $\alpha 1$ chain (COL6A1), a gene which is located on chromosome 21, encoding the $\alpha 1$ (VI) chain of type VI collagen (which is a primary extracellular matrix protein) was found which maintains the integrity of various tissues. Catenin delta-1 (CTNND1) functions as an oncogene and is known to be the driver of metastatic cancer progression [40]. A very important gene which we identified is six-transmembrane epithelial antigen of Prostate-2 (STEAP2), known to be over-expressed in aggressive PCa, which corroborates our study as it was identified in a high grade tumor [44]. Furthermore, we also observed that docking protein 5 (*DOK5*), a member of a subgroup of the DOK family, is known to be expressed using c-Ret in several neuronal tissues [49]. Amongst the lncRNAs, we identified SOX2OT mapped to the chromosome locus 3q26.3 and is highly expressed in embryonic stem cells [55]. The lncRNA FTX (five prime to XIST) possesses an X-inactive specific transcript and is involved in X-chromosome inactivation [57].

3.2. Protein interactions yielded innate pathways responsible for PCa

From the input genes shown with cross-hatched circles of uniform size, GeneMANIA added relevant genes which are shown with solid circles and their size is proportional to the number of interactions they have (Figure 2). In addition, we have identified several interacting partners that are coexpressed such as DZIP1, COL6A2, TAGLN, ZBTB33, LAMP1, LAMP2, IKBKB, DNAJB11 which acts as oncogenes in different cancers such as gastric cancer, renal cell carcinoma, colorectal cancer, to name a few.

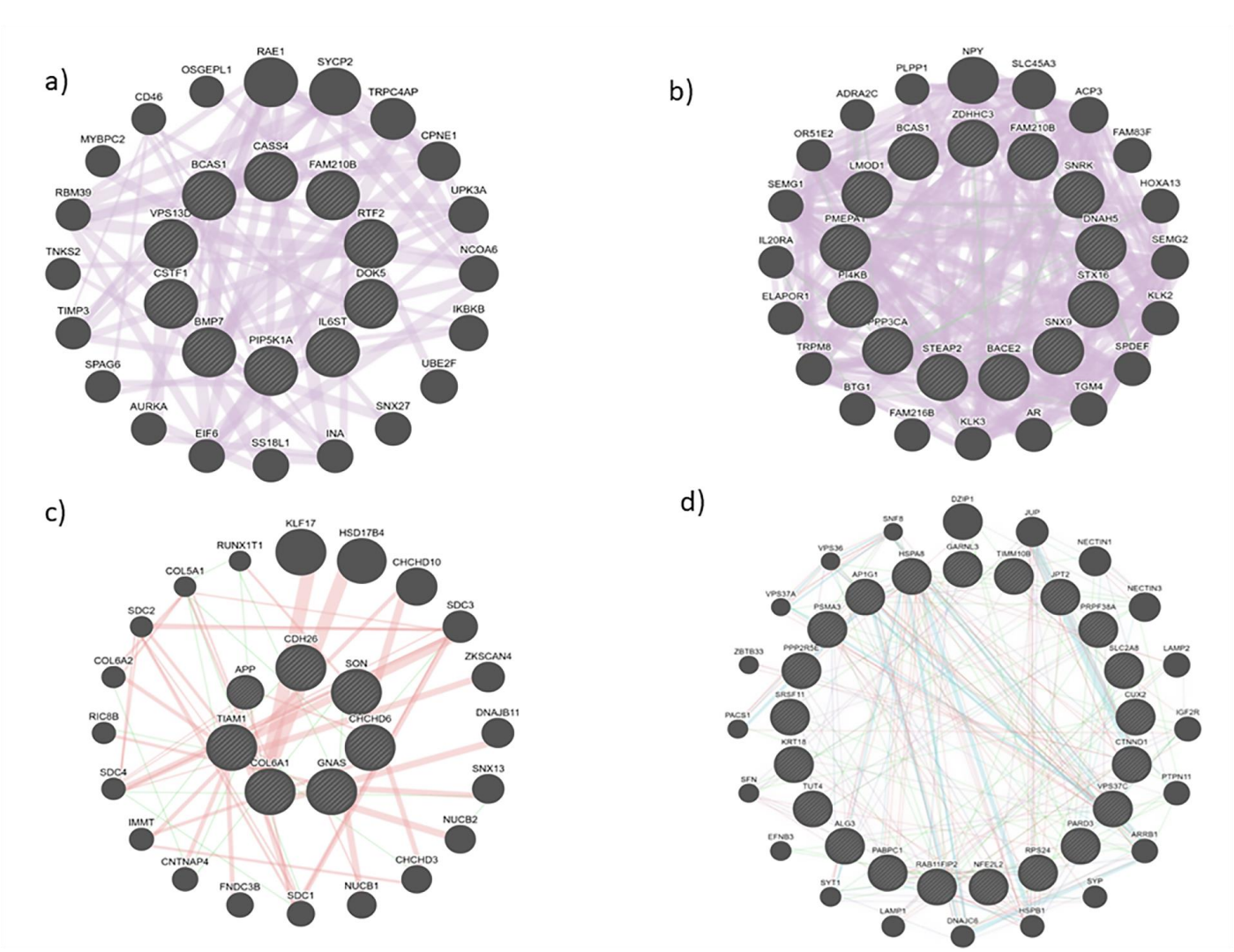


Figure 2. Protein-protein interaction of different malignant samples using GeneMANIA with purple edges representing co-expression, green -genetic interaction, red - physical interaction and blue representing co-localization. (a-d represents different malignant samples, the circles with edges are the input genes whereas solid circles are the interacting partners).

When we used the Cytoscape-cytoHubba plugin, top networks were ranked by clustering coefficient with the top 20 hub genes in our different malignant samples. Some of the prominent genes we identified through the networks are DOK5, APP, CTNND1, STX6, STX10, STX16, BACE1 and BACE2 (Figure 3 A-D)) which are in agreement with regulation of various cancers based on their coexpression patterns in GeneMANIA.

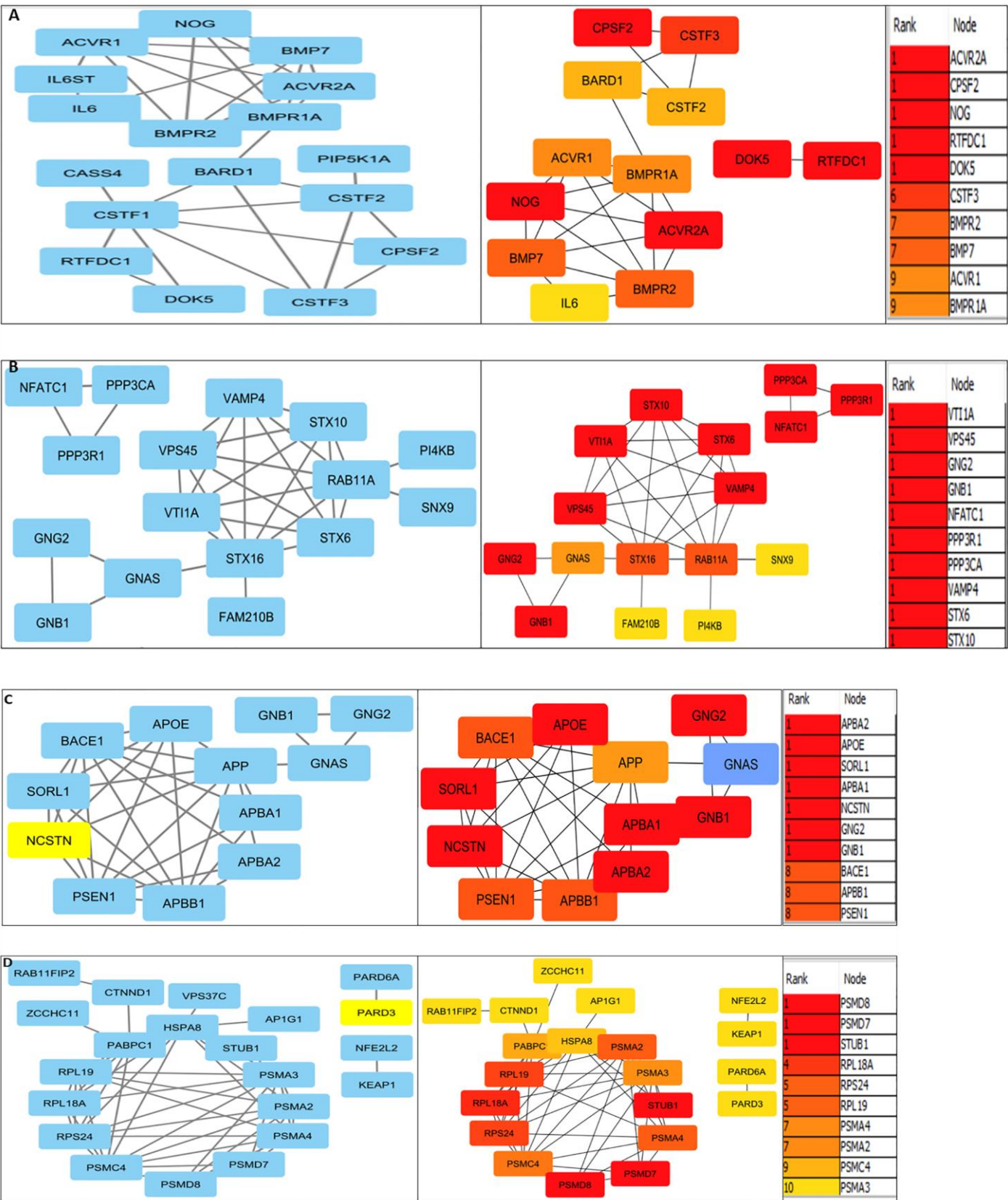


Figure 3. A compendium of top 20 genes from the PPI network that are regulated and expressed in malignant samples were considered from Cytoscape-cytoHubba plugin. The network was ranked based on clustering coefficients which yielded hierarchically high confidence interactions. These genes with highest clustering coefficients are indicated in red, while high to moderate clustering coefficients are in orange and those with low clustering coefficients are shown in yellow. Some of the important genes we identified through all 4 networks (A, B, C and D) are, *DOK5*, *APP*, *CTNND1*, *STX6*, *STX10*, *STX16*, *BACE1* and *BACE2* which are amongst the top ranking genes in the network.

3.3. Validation of RNA-seq result using TCGA dataset by cbiportal

We sought to ask whether any DEGs were relatively expressed in publicly available datasets from various studies. To check this, we used TCGA datasets and checked for alteration of frequencies and expression in the cohorts. We used the PanCancer Atlas [39] dataset for Prostate Adenocarcinoma, where 489 samples were screened for their functional roles and molecular aberrations (Figures 4 and 5). Mutations and CNV analysis for some of the DEGs were done and the results are summarized below:

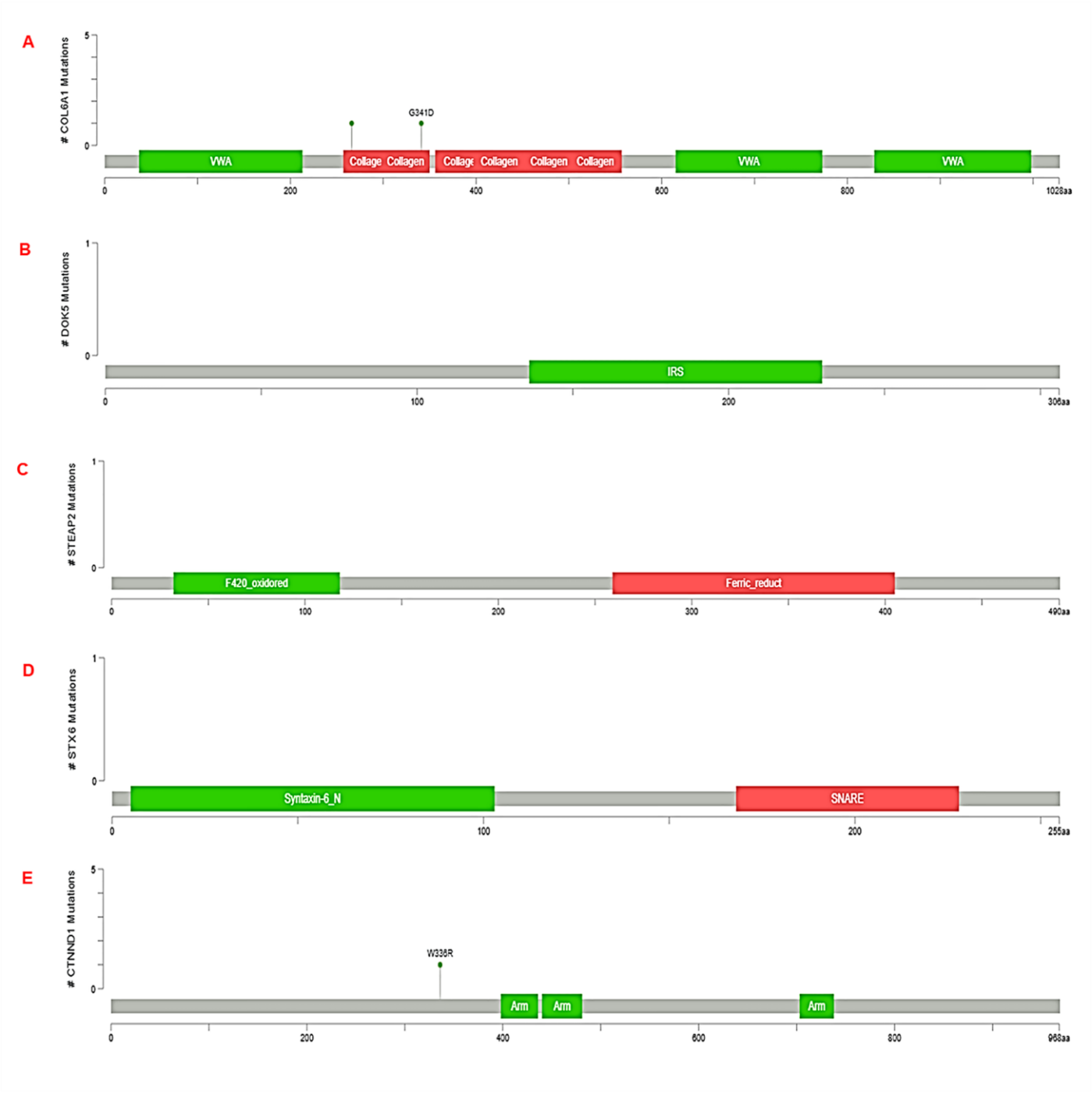


Figure 4. Overall percent and type of genetic mutations in *COL6A1*, *DOK5*, *STEAP2*, *STX6* and *CTNND1* related to PCa. For STX, it is fused with SNARE proteins which we also have identified in our study as one of the causal genes for PCa.

Furthermore, an attempt was made to analyze genomic alterations such as gene amplifications, deep deletions (that is equivalent to homozygous deletions), shallow deletions (heterozygous loss), and truncating mutations, inframe mutations, or missense mutations. Among the DEGs, amplification was the most prominent one in case of DOK5 and

STEAP2 whereas deep deletions and mutations were observed in COL6A1, STX6 and CTNND1. We argue that similar analysis could check performance of all the DEGs which will highlight the alteration frequencies across the cohort (Figure 5).

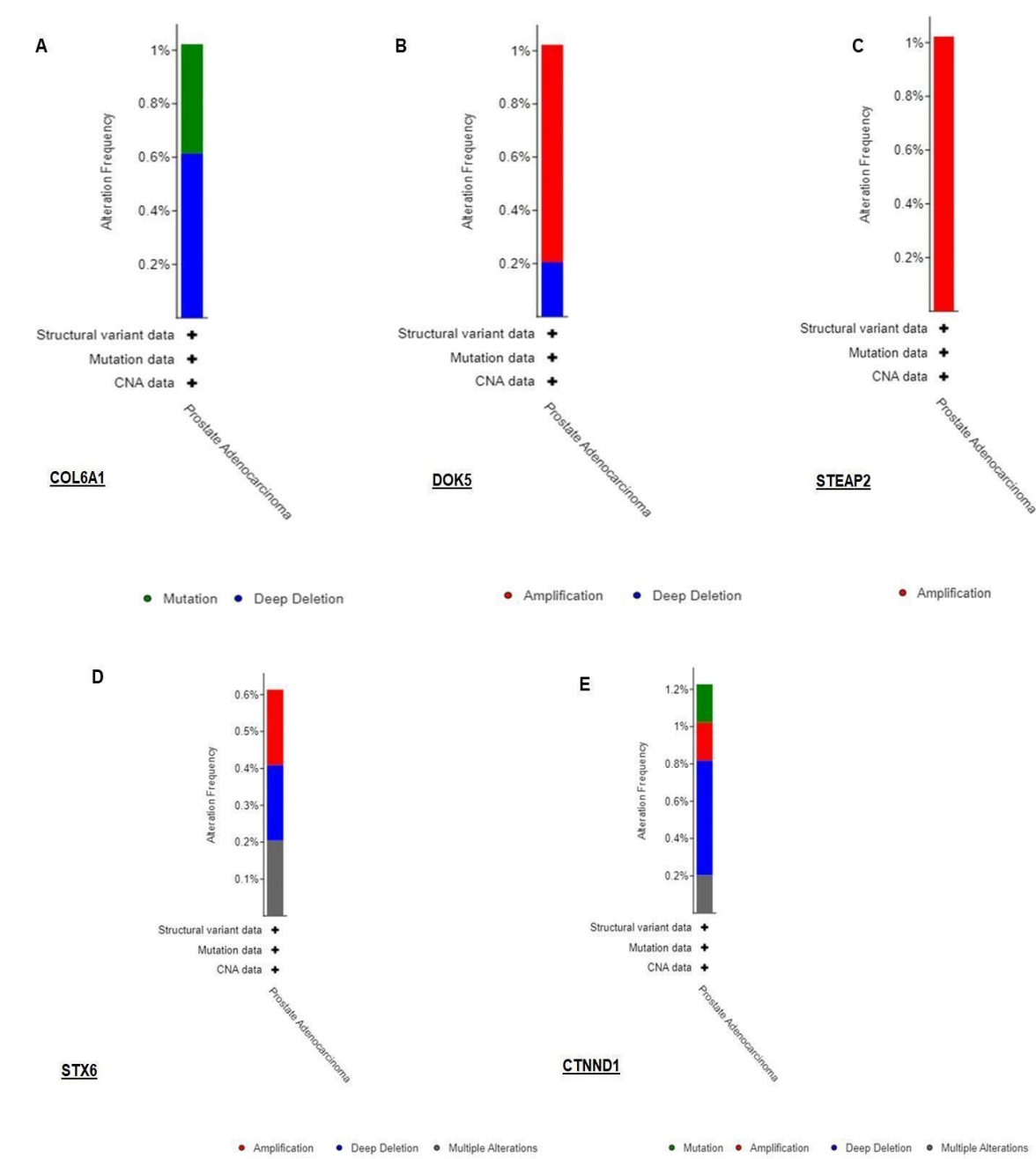


Figure 5. Alteration frequencies in (A)COL6A1 (B)DOK5 (C)STEAP2 (D)STX6 and (E)CTNND1. where the queried gene is altered either in 1% or less than 1% of queried patients/samples (Total number of samples-489).

3.4. Gene ontology yielded distinct pathways regulating biogenesis

All the DEGs obtained from our cuffdiff pipeline were subjected to GO analysis by Pantherdb which shows the role of DEGs in biological adhesions, biological regulations, biogenesis, cellular, metabolic and developmental processes, localization, locomotion, multicellular organismal process (Figures 6 and 7).

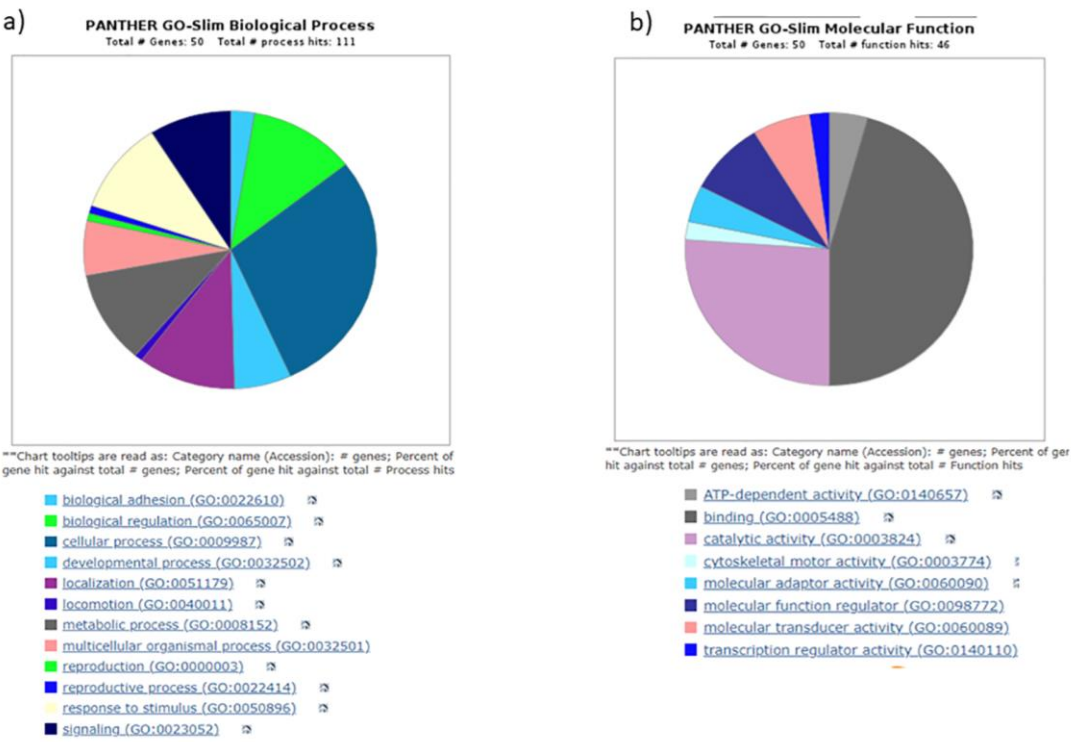


Figure 6. Summary of the (a) Biological functions (BP) and Molecular function (MP) terms in up-regulated DEGs.

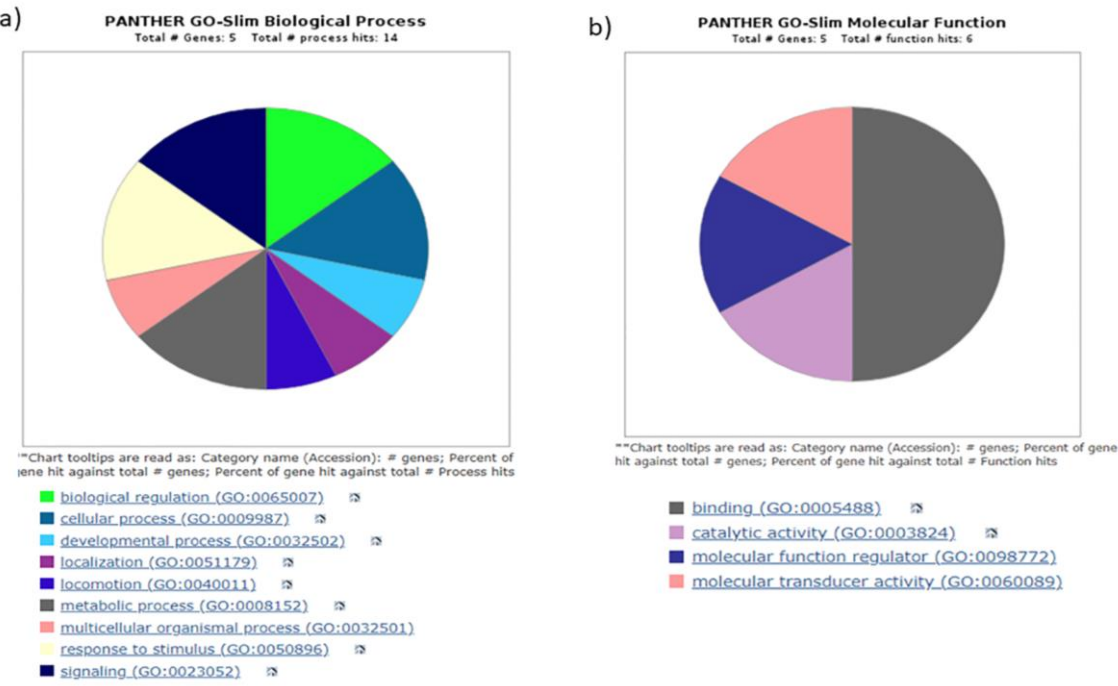


Figure 7. Summary of the (a) Biological functions (BP) and Molecular function (MP) terms in down-regulated DEGs.

For the molecular function category, the terms are binding factors, catalytic, adapter, transducer, transcription regulator activities which are associated with differential expressed genes. This indicates the role of DEGs in different important processes like transcription, cell migration, differentiation etc.

Comparative analysis of RNA-seq data with other publicly available datasets

To screen the potential DEGs across the datasets, we compared and analyzed the DEGs from our current study to that of the DEGs from cbiportal and methylation/array specific datasets that are publicly available from NCBI gene expression omnibus (GEO; <https://www.ncbi.nlm.nih.gov/geo/> last accessed December 22, 2022). As no concrete list of RNA-Seq datasets were associated in sequence read archive in lieu of PCa phenotype, we compared our DEGs to GEO datasets, viz. GSE6919 and GSE45016 in addition to previously benchmarked RNA-Seq dataset of PCa in the Chinese population (Ren et al, 2012). From this, we obtained only 135 hits in SRA for RNA-seq of PCa and most datasets are either from PCa cell-lines or RIP-seq and so we deemed them not useful keeping in view of diffident phenotypes, experiments and perhaps correlation studies they may be heralded with. Nevertheless, we identified a few common genes, viz. STEAP2, DOK5, IL6ST, LMOD1, CTNND1 etc., which could be the key candidate genes in our study. Likewise, when our data was compared with GEO datasets, we observed IL6ST, BACE2, SOX2-OT, STEAP2, APP, SNX9, STX16, CTNND1 among the other genes that are expressed. This, we believe, strengthens our finding that the DEGs which we have mentioned in the current study could be a valid signature for PCa diagnosis. On the other hand, when we compared Prostate Adenocarcinoma (TGCA, PanCancer Atlas) dataset to that of our list of DEGs, we found IL6ST, ZBTB20. These key DEGs are common between Indian population and western population which can be validated further as it is beyond the scope of this current analysis.

Phenolyzer highlights important DEGs

A cross-sectional comparison of phenotypes and DEGs using phenolyzer would provide us indicators for extent of expression across diseases. Therefore, we asked how many DEGs among all including the samples, cbiportal and comparative analyses are distinctly associated with disease/phenotype terms. Our Phenolyzer results with disease names, viz. Prostate Cancer, PCa neoplasia which are clinical phenotype terms have largely been connected with each other (seen in pink edges in Figure 8). On the other hand, in an attempt to reach consensus with identification of DEGs, we also employed the DESeq/EdgeR normalization method which resulted in the identification of approximately 1,230 genes as DEG: 490 were up-regulated and 1,215 were down-regulated. We identified distinct genes including KLK4, FN1, PBOV1, TPM2, and FLNA which are known to be involved in PCa pathways along with IGF1, TPD52, and SRSF1 that are involved in different cancer pathways. While KLK4 is a very important gene which is known to be involved in the progression of prostate cancer by promoting proliferation, migration and epithelial to mesenchymal transition, we have checked its expression in our sample by qRT-PCR and we did significant change in benign vs malignant samples. We also found a few novel lncRNAs such as LINC00940, FLJ16779 that have not been reported earlier besides SNHG19, NPBWR1 lncRNAs that are known to be involved in cancer and other diseases as well. Further attempt was made to understand the regulatory mechanisms underpinning PCa signature pathways by mapping lncRNAs with protein encoding genes.

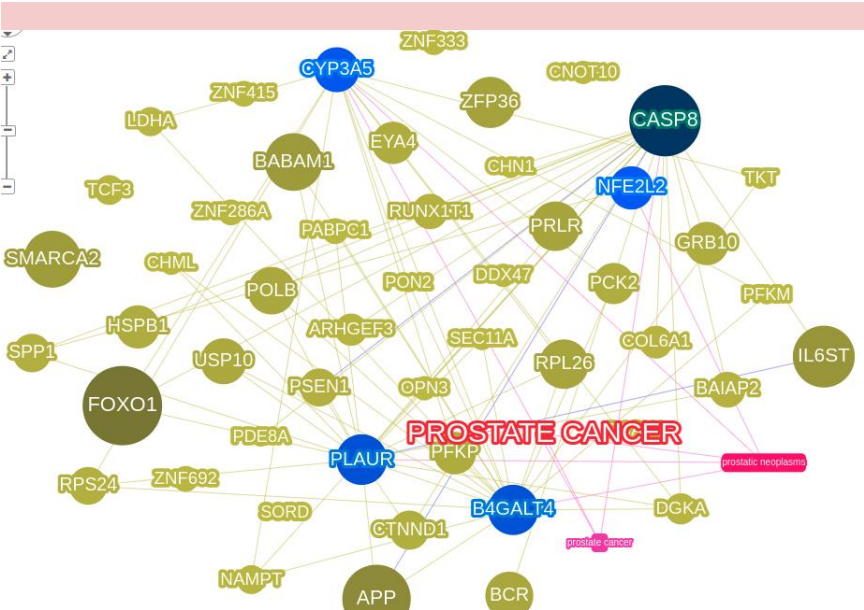


Figure 8. Phenolyzer hub of DEGs associated with distinct clinical terms, viz. PCa and PCa neoplasia.

4. Discussion

In the current study, we have identified many DEGs which are known to be involved in different cancer pathways including PCa. While some of the DEGs are specifically known to be associated with PCa, we also discovered a few novel lncRNAs which need further investigation. Expression of COL6A1 is significantly elevated in different tumors such as lung, prostate, cervical and pancreatic cancer compared to normal tissues. Interestingly, our previous WES studies in PCa have identified COL6A1 as one of the causal genes [7], whereas, we also identified this through our RNA-seq analysis. COL6A1 was shown to be physically interacting with DNAJB11, APP along with some other genes. DNAJB11 is involved in aberrant signaling pathways associated with different cancers. Similarly, APP is known to be associated with androgen-responsive genes and regulates proliferation and migration of PCa cells. Therefore, we argue that COL6A1 might act as a prognostic marker for PCa in the Indian population. Transcriptional factor Kaiso/ZBTB33 was identified as a CTNND1-specific binding partner and this complex is a modulator of the canonical Wnt/ β -catenin signaling pathway [41]. There is a large amount of research focusing on the role of CTNND1 in cancer development and progression, but in PCa, it is still not well elucidated [42]. In our study, we have identified CTNND1 and ZBTB33, through interaction studies but what is more interesting is that we have earlier identified CTNND1 as a co-localization partner with one of the lncRNAs which is known to be highly expressed in PCa. Furthermore, CTNND1 is one of the main interacting partners of ACE2/TMPRSS2, the main receptors which are responsible for SARS-CoV-2 entry into the cell. We had hypothesized that CTNND1 interacts with EGFR and this interaction could uphold SARS-CoV-2 infection independent of its endocytosis and associate with cell viability [43]. The interacting partners of another important gene STEAP2 are KLK3, KLK2 and AR all of which are hallmarks of PCa. KLK3 is a protein coding gene and its protein product Prostate specific antigen (PSA) is a well-established biomarker of PCa [45] [46]. Similarly, human kallikrein 2 (KLK2), interacts with AR and drives PCa progression [47] [48]. Earlier studies have shown that DOK5 are expressed in T-cells and their expression is regulated upon T-cell activation. DOK5 is shown to be involved in invasion and metastasis of cancer specifically in Gastric cancer but it has not been well studied in PCa [50]. Since, we got a strong interaction of DOK5 in our clustering coefficient studies using cytoHubba plugin, we argue that it would be worth analyzing this gene further. Interestingly, through our cytoscape-cytoHubba analysis, we identified many soluble N-ethylmaleimide-sensitive factor (NSF) attachment protein receptor (SNARE) proteins that are key mediators of membrane fusion [51]. All the SNARE proteins share a common sequence of 60–70-residue called the SNARE motifs, that helps in mediating the interaction between vesicle SNAREs (v-SNAREs) and target membrane SNAREs (t-SNAREs) [52]. One of the t-SNARE proteins, syntaxin 6 (STX6) is particularly important in vesicle fusion. STX6 is up-regulated in a variety of cancers including breast, colon, liver, pancreatic, prostate, bladder, skin, testicular, tongue, cervical, lung and gastric cancers and it has been identified as a common transcriptional target of p53 family members (p53, p63 and p73). Along with STX6, we have identified STX10, VTI1A and STX16 which can be further studied. Overall, we also found other important DEGs, viz. BACE1 and BACE2 which belong to a class of proteases called β -secretases that are extensively studied in Alzheimer's disease [53]. Not many studies have been done with respect to their role in cancer, but there are some recent studies which have shown their involvement in pancreatic and skin cancers [54].

A major chunk of lncRNAs are novel and regulated in distinct pathways

lncRNAs do not code for proteins but they are involved in almost all biological processes such as gene expression, epigenetic regulation, cell cycle regulation etc in different cancers, PCa being one of them. Previous studies have highlighted the oncogenic role of lncRNAs in metastasis, proliferation and development of PCa whereas they act as tumor suppressants but still there are several lncRNAs whose functions are still unknown. With the advancement in NGS, bioinformatics analysis has enabled identification of many lncRNAs which show dysregulated expression in PCa [62]. Their diverse role has made them a target for all stages of PCa development which includes screening, diagnosis, prognosis and treatment further establishing their role as biomarkers in PCa. For example, MALAT-1 (Metastasis associated lung adenocarcinoma transcript 1), an lncRNA which is used to predict metastasis and survival in non-small cell lung cancer [63], has also been correlated with PCa development and progression [64]. It has also been reported that MALAT-1 expression closely correlates with PSA levels, Gleason scores, and tumor sizes [64]. Similarly, PCAT-18 (prostate cancer-associated non-coding RNA transcript 18 and SchLAP1 (second chromosome locus associated with prostate-1) has also been used as diagnostic and prognostic biomarkers in PCa [65] [66]. Some of the lncRNAs we identified from our current study includes LINC01440, SOX2OT, ENSG00000232855, ENST00000647843.1 and FTX. FTX has been reported to be involved in the tumorigenesis of multiple cancer types [58]. Long intergenic non-protein coding RNA 1440 (LINC01440) is novel lncRNA which needs to be explored for its role in cancers. A recent study has implicated

its role in a spinal disorder known as Ossification of ligamentum flavum (OLF) where it was found to be up-regulated in the diseased patients compared to the healthy population [59]. Different studies have shown that SOX2OT acts as an oncogene and is elevated in different tumor types [56]. But its role and significance is still not explored in PCa, which makes it a very promising candidate for further analysis. Another lncRNA, ENSG00000287903 (NONHSAT106693) was earlier screened in our Vitamin K deficiency cohort from our recent study [61] wherein qRT-PCR showed a significant up-regulation in malignant samples compared to control samples (supplementary information) which provides us a *raison d'être* to further validate the remaining lncRNAs as well. Given the role of lncRNAs in PCa, it would be interesting to see if any of these lncRNAs can serve as biomarkers, albeit several downstream experiments and well-designed clinical trials are to be employed. Taken together, there are a few limitations for our study. (i) The qRT-PCR validation needs more tumor and adjacent normal tissue samples as our sample size is very small. (ii) Some additional experiments, such as immunohistochemistry and Western blot, when performed could confirm the protein levels in PCa. (iii) Given the scarcity of PCa data in India, the survival analysis across the entire transcriptome could not be drawn as it is largely towards identification of biomarkers indicating the prognostic power but given the sample size conundrum we have, we are limited to check this power. Nevertheless, In our previous study, we have performed WES on PCa samples which we have cited but again the sample size is limited. (iv) Due to lack of fresh-frozen prostate/radical prostatectomy tissues, only FFPE blocks were used and isolating RNA from them is an arduous and challenging task [37], because of which we could not perform qRT-PCR for all the selected DEGs.

5. Conclusions

Prostate cancer cases are increasing in India even as NGS studies are just beginning to be explored. In our current RNA-seq study and subsequent bioinformatics analyses, we sought to characterize many DEGs including lncRNAs that are specific to PCa of Indian sub-population. While we identified some of the important genes, *viz.* *DOK5*, *COL6A1*, *CTNND1*, *STEAP2*, *APP*, their role in PCa is still not clear. We envisage that characterizing their functional aspects would help us understand PCa progression. Since search for non-invasive and more sensitive biomarkers is on the anvil across all solid tumors, we firmly hope that these lncRNAs amongst the DEGs would serve as a precedent in development of NGS panels for PCa detection in Indian phenotype.

Supplementary Materials: The following supporting information can be downloaded at: www.mdpi.com/xxx/s1, Table S1: P values for GO analysis. Table S2: LncPro table for lncRNA-protein interaction. Supplementary Information: qRT-PCR validation of important DEGs, Primer sequences used in the qRT-PCR, and PCR conditions for the experiment.

Author Contributions: NS and BK contributed equally. All the other authors chipped in with lateral sections. NS wrote the first draft. PS conceived the project and proofread the manuscript before all the authors agreed to the manuscript.

Funding: Nidhi Shukla gratefully acknowledges the Department of Science and Technology (DST), Government of India for women scientist fellowship (WOS-A).

Institutional Review Board Statement: This study was authorized by the Office of the Institutional Ethics committee at the Rukmani Birla Hospital, Jaipur, India. Reference number: RBH/IEC/21/008.

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study. Participants in this study were recruited from the Urology department at Rukmani Birla Hospitals, Jaipur, India. All the participants were well informed about the use of their samples.

Data Availability Statement: The data that support the findings of this study are available at National Center for Biotechnology Information (NCBI) Sequence Read Archive (SRA) database; accession number-**PRJNA616165**.

Acknowledgments: NS acknowledges the Department of Science and Technology, GOI, for the Women Scientist fellowship (SR/WOS-A/LS-404/2018) award.

Conflicts of Interest: The author(s) declared no potential conflicts of interest with respect to the research, authorship and/or publication of this article.

References

- [1] Jemal, A.; Bray, F.; Center, M.M.; Ferlay, J.; Ward, E.; Forman, D. Global Cancer Statistics. *CA. Cancer J. Clin.*, **2011**.
- [2] Sim, H.G.; Cheng, C.W.S. Changing Demography of Prostate Cancer in Asia. *Eur. J. Cancer*, **2005**.
- [3] Jain, S.; Saxena, S.; Kumar, A. Epidemiology of Prostate Cancer in India. *Meta Gene*, **2014**, 2, 596–605.
- [4] NCRP. *Three-Year Report of the Population Based Cancer Registries- 2009-2011*; **2013**.
- [5] Trapnell C, Pachter L, Salzberg SL. TopHat: discovering splice junctions with RNA-Seq. *Bioinformatics*. **2009**, 1;25(9):1105-11.
- [6] Guan YF, Li GR, Wang RJ, Yi YT, Yang L, Jiang D, Zhang XP, Peng Y. Application of next-generation sequencing in clinical oncology to advance personalized treatment of cancer. *Chinese journal of cancer*. 2012, 31(10):463.
- [7] Gupta, A.; Shukla, N.; Nehra, M.; Gupta, S.; Malik, B.; Mishra, A.K.; Vijay, M.; Batra, J.; Lohiya, N.K.; Sharma, D.; Suravajhala, P. A Pilot Study on the Whole Exome Sequencing of Prostate Cancer in the Indian Phenotype Reveals Distinct Polymorphisms. *Front. Genet.*, **2020**, 11.
- [8] Stark, R.; Grzelak, M.; Hadfield, J. RNA Sequencing: The Teenage Years. *Nature Reviews Genetics*, **2019**.
- [9] Marco-Puche, G.; Lois, S.; Benítez, J.; Trivino, J.C. RNA-Seq Perspectives to Improve Clinical Diagnosis. *Frontiers in Genetics*, **2019**.
- [10] Stupnikov A, McInerney CE, Savage KI, McIntosh SA, Emmert-Streib F, Kennedy R, Salto-Tellez M, Prise KM, McArt DG. Robustness of differential gene expression analysis of RNA-seq. *Computational and structural biotechnology journal*. **2021**, 1;19:3470-81.
- [11] Kukurba, K.R.; Montgomery, S.B. RNA Sequencing and Analysis. *Cold Spring Harb. Protoc.*, **2015**.
- [12] Xi, X.; Li, T.; Huang, Y.; Sun, J.; Zhu, Y.; Yang, Y.; Lu, Z.J. RNA Biomarkers: Frontier of Precision Medicine for Cancer. *Non-coding RNA*, **2017**.
- [13] Chatterjee, S.K.; Zetter, B.R. Cancer Biomarkers: Knowing the Present and Predicting the Future. *Future Oncology*, **2005**.
- [14] Chen, L.L. The Biogenesis and Emerging Roles of Circular RNAs. *Nature Reviews Molecular Cell Biology*, **2016**.
- [15] Lopez, J.P.; Diallo, A.; Cruceanu, C.; Fiori, L.M.; Laboissiere, S.; Guillet, I.; Fontaine, J.; Ragoussis, J.; Benes, V.; Turecki, G.; Ernst, C. Biomarker Discovery: Quantification of MicroRNAs and Other Small Non-Coding RNAs Using next Generation Sequencing. *BMC Med. Genomics*, **2015**.
- [16] Krishnan P, Damaraju S. The challenges and opportunities in the clinical application of noncoding RNAs: the road map for miRNAs and piRNAs in cancer diagnostics and prognostics. *International journal of genomics*. **2018**, 30;2018..
- [17] Grixti, J.M.; Ayers, D. Long Noncoding RNAs and Their Link to Cancer. *Non-coding RNA Research*, **2020**.
- [18] Evans, J.R.; Feng, F.Y.; Chinnaiyan, A.M. The Bright Side of Dark Matter: LncRNAs in Cancer. *Journal of Clinical Investigation*, **2016**.
- [19] Arun, G.; Diermeier, S.D.; Spector, D.L. Therapeutic Targeting of Long Non-Coding RNAs in Cancer. *Trends in Molecular Medicine*, **2018**.
- [20] Gao, N.; Li, Y.; Li, J.; Gao, Z.; Yang, Z.; Li, Y.; Liu, H.; Fan, T. Long Non-Coding RNAs: The Regulatory Mechanisms, Research Strategies, and Future Directions in Cancers. *Frontiers in Oncology*, **2020**.
- [21] Winkle, M.; El-Daly, S.M.; Fabbri, M.; Calin, G.A. Noncoding RNA Therapeutics — Challenges and Potential Solutions. *Nature Reviews Drug Discovery*, **2021**.
- [22] Kumar P, Chakraborty J, Sukumar G, Dalgard C, Chatterjee R, Biswas R. Comparative RNA-seq analysis reveals dys-regulation of major canonical pathways in ERG-inducible LNCaP cell progression model of prostate cancer. *Oncotarget*. 2019 Jul 2;10(42):4290-4306. doi: 10.18632/oncotarget.27019. PMID: 31303963; PMCID: PMC6611515..
- [23] Trapnell, C.; Roberts, A.; Goff, L.; Pertea, G.; Kim, D.; Kelley, D.R.; Pimentel, H.; Salzberg, S.L.; Rinn, J.L.; Pachter, L. Differential Gene and Transcript Expression Analysis of RNA-Seq Experiments with TopHat and Cufflinks. *Nat. Protoc.*, **2012**.
- [24] Gupta, A.; Gupta, S.; Jatawa, S.K.; Kumar, A.; Suravajhala, P. A Simplest Bioinformatics Pipeline for Whole Transcriptome Sequencing: Overview of the Processing and Steps from Raw Data to Downstream Analysis. *bioRxiv*, **2019**.
- [25] Ferragina P, Manzini G. *Proc. 41st Annual Symposium on Foundations of Computer Science*. IEEE Comput. Soc.; 2000. pp. 390–398
- [26] Trapnell, C.; Hendrickson, D.G.; Sauvageau, M.; Goff, L.; Rinn, J.L.; Pachter, L. Differential Analysis of Gene Regulation at Transcript Resolution with RNA-Seq. *Nat. Biotechnol.*, **2013**.
- [27] Liu, X.; Zhao, J.; Xue, L.; Zhao, T.; Ding, W.; Han, Y.; Ye, H. A Comparison of Transcriptome Analysis Methods

with Reference Genome. *BMC Genomics*, **2022**.

- [28] Mostafavi, S.; Morris, Q. Combining Many Interaction Networks to Predict Gene Function and Analyze Gene Lists. *Proteomics*, **2012**.
- [29] Szklarczyk, D.; Gable, A.L.; Nastou, K.C.; Lyon, D.; Kirsch, R.; Pyysalo, S.; Doncheva, N.T.; Legeay, M.; Fang, T.; Bork, P.; Jensen, L.J.; von Mering, C. The STRING Database in 2021: Customizable Protein-Protein Networks, and Functional Characterization of User-Uploaded Gene/Measurement Sets. *Nucleic Acids Res.*, **2021**.
- [30] Shannon, P.; Markiel, A.; Ozier, O.; Baliga, N.S.; Wang, J.T.; Ramage, D.; Amin, N.; Schwikowski, B.; Ideker, T. Cytoscape: A Software Environment for Integrated Models of Biomolecular Interaction Networks. *Genome Res.*, **2003**.
- [31] Saito R, Smoot ME, Ono K, Ruscheinski J, Wang PL, Lotia S, Pico AR, Bader GD, Ideker T. A travel guide to Cytoscape plugins. *Nat Methods*. 2012 Nov;9(11):1069-76. doi: 10.1038/nmeth.2212. Epub 2012 Nov 6. PMID: 23132118; PMCID: PMC3649846.
- [32] Chin, C.H.; Chen, S.H.; Wu, H.H.; Ho, C.W.; Ko, M.T.; Lin, C.Y. CytoHubba: Identifying Hub Objects and Sub-Networks from Complex Interactome. *BMC Syst. Biol.*, **2014**.
- [33] Mi, H.; Muruganujan, A.; Casagrande, J.T.; Thomas, P.D. Large-Scale Gene Function Analysis with the Panther Classification System. *Nat. Protoc.*, **2013**.
- [34] Cerami E, Gao J, Dogrusoz U, Gross BE, Sumer SO, Aksoy BA, Jacobsen A, Byrne CJ, Heuer ML, Larsson E, Antipin Y, Reva B, Goldberg AP, Sander C, Schultz N. The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data. *Cancer Discov*. 2012 May;2(5):401-4. doi: 10.1158/2159-8290.CD-12-0095. Erratum in: *Cancer Discov*. 2012 Oct;2(10):960. PMID: 22588877; PMCID: PMC3956037.
- [35] Liu, J.; Lichtenberg, T.; Hoadley, K.A.; Poisson, L.M.; Lazar, A.J.; Cherniack, A.D.; Kovatich, A.J.; Benz, C.C.; Levine, D.A.; Lee, A. V.; Omberg, L.; Wolf, D.M.; Shriver, C.D.; Thorsson, V.; Caesar-Johnson, S.J.; Demchok, J.A.; Felau, I.; Kasapi, M.; Ferguson, M.L.; Hutter, C.M.; Sofia, H.J.; Tarnuzzer, R.; Wang, Z.; Yang, L.; Zenklusen, J.C.; Zhang, J. (Julia); Chudamani, S.; Liu, J.; Lolla, L.; Naresh, R.; Pihl, T.; Sun, Q.; Wan, Y.; Wu, Y.; Cho, J.; DeFreitas, T.; Frazer, S.; Gehlenborg, N.; Getz, G.; Heiman, D.I.; Kim, J.; Lawrence, M.S.; Lin, P.; Meier, S.; Noble, M.S.; Saksena, G.; Voet, D.; Zhang, H.; Bernard, B.; Chambwe, N.; Dhankani, V.; Knijnenburg, T.; Kramer, R.; Leinonen, K.; Liu, Y.; Miller, M.; Reynolds, S.; Shmulevich, I.; Thorsson, V.; Zhang, W.; Akbani, R.; Broom, B.M.; Hegde, A.M.; Ju, Z.; Kanchi, R.S.; Korkut, A.; Li, J.; Liang, H.; Ling, S.; Liu, W.; Lu, Y.; Mills, G.B.; Ng, K.S.; Rao, A.; Ryan, M.; Wang, J.; Weinstein, J.N.; Zhang, J.; Abeshouse, A.; Armenia, J.; Chakravarty, D.; Chatila, W.K.; de Bruijn, I.; Gao, J.; Gross, B.E.; Heins, Z.J.; Kundra, R.; La, K.; Ladanyi, M.; Luna, A.; Nissan, M.G.; Ochoa, A.; Phillips, S.M.; Reznik, E.; Sanchez-Vega, F.; Sander, C.; Schultz, N.; Sheridan, R.; Sumer, S.O.; Sun, Y.; Taylor, B.S.; Wang, J.; Zhang, H.; Anur, P.; Peto, M.; Spellman, P.; Benz, C.; Stuart, J.M.; Wong, C.K.; Yau, C.; Hayes, D.N.; Parker, J.S.; Wilkerson, M.D.; Ally, A.; Balasundaram, M.; Bowlby, R.; Brooks, D.; Carlsen, R.; Chuah, E.; Dhalla, N.; Holt, R.; Jones, S.J.M.; Kasaian, K.; Lee, D.; Ma, Y.; Marra, M.A.; Mayo, M.; Moore, R.A.; Mungall, A.J.; Mungall, K.; Robertson, A.G.; Sadeghi, S.; Schein, J.E.; Sipahimalani, P.; Tam, A.; Thiessen, N.; Tse, K.; Wong, T.; Berger, A.C.; Beroukhi, R.; Cherniack, A.D.; Cibulskis, C.; Gabriel, S.B.; Gao, G.F.; Ha, G.; Meyerson, M.; Schumacher, S.E.; Shih, J.; Kucherlapati, M.H.; Kucherlapati, R.S.; Baylin, S.; Cope, L.; Danilova, L.; Bootwalla, M.S.; Lai, P.H.; Maglinte, D.T.; Van Den Berg, D.J.; Weisenberger, D.J.; Auman, J.T.; Balu, S.; Bodenheimer, T.; Fan, C.; Hoadley, K.A.; Hoyle, A.P.; Jefferys, S.R.; Jones, C.D.; Meng, S.; Mieczkowski, P.A.; Mose, L.E.; Perou, A.H.; Perou, C.M.; Roach, J.; Shi, Y.; Simons, J. V.; Skelly, T.; Soloway, M.G.; Tan, D.; Veluvolu, U.; Fan, H.; Hinoue, T.; Laird, P.W.; Shen, H.; Zhou, W.; Bellair, M.; Chang, K.; Covington, K.; Creighton, C.J.; Dinh, H.; Doddapaneni, H.V.; Donehower, L.A.; Drummond, J.; Gibbs, R.A.; Glenn, R.; Hale, W.; Han, Y.; Hu, J.; Korchina, V.; Lee, S.; Lewis, L.; Li, W.; Liu, X.; Morgan, M.; Morton, D.; Muzny, D.; Santibanez, J.; Sheth, M.; Shinbro, E.; Wang, L.; Wang, M.; Wheeler, D.A.; Xi, L.; Zhao, F.; Hess, J.; Appelbaum, E.L.; Bailey, M.; Cordes, M.G.; Ding, L.; Fronick, C.C.; Fulton, L.A.; Fulton, R.S.; Kandoth, C.; Mardis, E.R.; McLellan, M.D.; Miller, C.A.; Schmidt, H.K.; Wilson, R.K.; Crain, D.; Curley, E.; Gardner, J.; Lau, K.; Mallery, D.; Morris, S.; Paulauskis, J.; Penny, R.; Shelton, C.; Shelton, T.; Sherman, M.; Thompson, E.; Yena, P.; Bowen, J.; Gastier-Foster, J.M.; Gerken, M.; Leraas, K.M.; Lichtenberg, T.M.; Ramirez, N.C.; Wise, L.; Zmuda, E.; Corcoran, N.; Costello, T.; Hovens, C.; Carvalho, A.L.; de Carvalho, A.C.; Fregnani, J.H.; Longatto-Filho, A.; Reis, R.M.; Scapulatempo-Neto, C.; Silveira, H.C.S.; Vidal, D.O.; Burnette, A.; Eschbacher, J.; Hermes, B.; Noss, A.; Singh, R.; Anderson, M.L.; Castro, P.D.; Ittmann, M.; Huntsman, D.; Kohl, B.; Le, X.; Thorp, R.; Andry, C.; Duffy, E.R.; Lyadov, V.; Paklina, O.; Setdikova, G.; Shabunin, A.; Tavobilov, M.; McPherson, C.; Warnick, R.; Berkowitz, R.; Cramer, D.; Feltmate, C.; Horowitz, N.; Kibel, A.; Muto, M.; Raut, C.P.; Malykh, A.; Barnholtz-Sloan, J.S.; Barrett, W.; Devine, K.; Fulop, J.; Ostrom, Q.T.; Shimmel, K.; Wolinsky, Y.; Sloan, A.E.; De Rose, A.; Giulianti, F.; Goodman, M.; Karlan, B.Y.; Hagedorn, C.H.; Eckman, J.; Harr, J.; Myers, J.; Tucker, K.; Zach, L.A.; Deyarmin, B.; Hu, H.; Kvecher, L.; Larson, C.; Mural, R.J.; Somiari, S.; Vicha, A.; Zelinka, T.; Bennett, J.; Iacocca, M.; Rabeno, B.; Swanson, P.; Latour, M.; Lacombe, L.; Têtu, B.; Bergeron, A.; McGraw, M.

Staugaitis, S.M.; Chabot, J.; Hibshoosh, H.; Sepulveda, A.; Su, T.; Wang, T.; Potapova, O.; Voronina, O.; Desjardins, L.; Mariani, O.; Roman-Roman, S.; Sastre, X.; Stern, M.H.; Cheng, F.; Signoretti, S.; Berchuck, A.; Bigner, D.; Lipp, E.; Marks, J.; McCall, S.; McLendon, R.; Secord, A.; Sharp, A.; Behera, M.; Brat, D.J.; Chen, A.; Delman, K.; Force, S.; Khuri, F.; Magliocca, K.; Maithel, S.; Olson, J.J.; Owonikoko, T.; Pickens, A.; Ramalingam, S.; Shin, D.M.; Sica, G.; Van Meir, E.G.; Zhang, H.; Eijckenboom, W.; Gillis, A.; Korpershoek, E.; Looijenga, L.; Oosterhuis, W.; Stoop, H.; van Kessel, K.E.; Zwarthoff, E.C.; Calatozzolo, C.; Cuppini, L.; Cuzzubbo, S.; DiMeco, F.; Finocchiaro, G.; Mattei, L.; Perin, A.; Pollo, B.; Chen, C.; Houck, J.; Lohavanichbutr, P.; Hartmann, A.; Stoeher, C.; Stoeher, R.; Taubert, H.; Wach, S.; Wullich, B.; Kyder, W.; Murawa, D.; Wiznerowicz, M.; Chung, K.; Edenfield, W.J.; Martin, J.; Baudin, E.; Bubley, G.; Bueno, R.; De Rienzo, A.; Richards, W.G.; Kalkanis, S.; Mikkelsen, T.; Noushmehr, H.; Scarpaccia, L.; Girard, N.; Aymerich, M.; Campo, E.; Giné, E.; Guillermo, A.L.; Van Bang, N.; Hanh, P.T.; Phu, B.D.; Tang, Y.; Colman, H.; Evason, K.; Dottino, P.R.; Martignetti, J.A.; Gabra, H.; Juhl, H.; Akeredolu, T.; Stepa, S.; Hoon, D.; Ahn, K.; Kang, K.J.; Beuschlein, F.; Breggia, A.; Birrer, M.; Bell, D.; Borad, M.; Bryce, A.H.; Castle, E.; Chandan, V.; Cheville, J.; Copland, J.A.; Farnell, M.; Flotte, T.; Giana, N.; Ho, T.; Kendrick, M.; Kocher, J.P.; Kopp, K.; Moser, C.; Nagorney, D.; O'Brien, D.; O'Neill, B.P.; Patel, T.; Petersen, G.; Que, F.; Rivera, M.; Roberts, L.; Smallridge, R.; Smyrk, T.; Stanton, M.; Thompson, R.H.; Torbenson, M.; Yang, J.D.; Zhang, L.; Brimo, F.; Ajani, J.A.; Angulo Gonzalez, A.M.; Behrens, C.; Bondaruk, J.; Broaddus, R.; Czerniak, B.; Esmaeli, B.; Fujimoto, J.; Gershenwald, J.; Guo, C.; Logothetis, C.; Meric-Bernstam, F.; Moran, C.; Ramondetta, L.; Rice, D.; Sood, A.; Tamboli, P.; Thompson, T.; Troncoso, P.; Tsao, A.; Wistuba, I.; Carter, C.; Haydu, L.; Hersey, P.; Jakrot, V.; Kakavand, H.; Kefford, R.; Lee, K.; Long, G.; Mann, G.; Quinn, M.; Saw, R.; Scolyer, R.; Shannon, K.; Spillane, A.; Stretch, J.; Synnott, M.; Thompson, J.; Wilmott, J.; Al-Ahmadie, H.; Chan, T.A.; Ghossein, R.; Gopalan, A.; Levine, D.A.; Reuter, V.; Singer, S.; Singh, B.; Tien, N.V.; Broudy, T.; Mirsaidi, C.; Nair, P.; Drwiega, P.; Miller, J.; Smith, J.; Zaren, H.; Park, J.W.; Hung, N.P.; Kebebew, E.; Linehan, W.M.; Metwalli, A.R.; Pacak, K.; Pinto, P.A.; Schiffman, M.; Schmidt, L.S.; Vocke, C.D.; Wentzensen, N.; Worrell, R.; Yang, H.; Moncrieff, M.; Goparaju, C.; Melamed, J.; Pass, H.; Botnariuc, N.; Caraman, I.; Cernat, M.; Chemencedji, I.; Clipca, A.; Doruc, S.; Gorincioi, G.; Mura, S.; Pirtac, M.; Stancul, I.; Tcaciuc, D.; Albert, M.; Alexopoulou, I.; Arnaout, A.; Bartlett, J.; Engel, J.; Gilbert, S.; Parfitt, J.; Sekhon, H.; Thomas, G.; Rassl, D.M.; Rintoul, R.C.; Bifulco, C.; Tamakawa, R.; Urba, W.; Hayward, N.; Timmers, H.; Antenucci, A.; Facciolo, F.; Grazi, G.; Marino, M.; Merola, R.; de Krijger, R.; Gimenez-Roqueplo, A.P.; Piché, A.; Chevalier, S.; McKercher, G.; Birsoy, K.; Barnett, G.; Brewer, C.; Farver, C.; Naska, T.; Pennell, N.A.; Raymond, D.; Schilero, C.; Smolenski, K.; Williams, F.; Morrison, C.; Borgia, J.A.; Liptay, M.J.; Pool, M.; Seder, C.W.; Junker, K.; Omberg, L.; Dinkin, M.; Manikhas, G.; Alvaro, D.; Bragazzi, M.C.; Cardinale, V.; Carpino, G.; Gaudio, E.; Chesla, D.; Cottingham, S.; Dubina, M.; Moiseenko, F.; Dhanasekaran, R.; Becker, K.F.; Janssen, K.P.; Slotta-Huspenina, J.; Abdel-Rahman, M.H.; Aziz, D.; Bell, S.; Cebulla, C.M.; Davis, A.; Duell, R.; Elder, J.B.; Hilty, J.; Kumar, B.; Lang, J.; Lehman, N.L.; Mandt, R.; Nguyen, P.; Pilarski, R.; Rai, K.; Schoenfield, L.; Senecal, K.; Wakely, P.; Hansen, P.; Lechan, R.; Powers, J.; Tischler, A.; Grizzle, W.E.; Sexton, K.C.; Kastl, A.; Henderson, J.; Porten, S.; Waldmann, J.; Fassnacht, M.; Asa, S.L.; Schadendorf, D.; Couce, M.; Graefen, M.; Huland, H.; Sauter, G.; Schlomm, T.; Simon, R.; Tennstedt, P.; Olabode, O.; Nelson, M.; Bathe, O.; Carroll, P.R.; Chan, J.M.; Disaia, P.; Glenn, P.; Kelley, R.K.; Landen, C.N.; Phillips, J.; Prados, M.; Simko, J.; Smith-McCune, K.; VandenBerg, S.; Roggin, K.; Fehrenbach, A.; Kendler, A.; Sifri, S.; Steele, R.; Jimeno, A.; Carey, F.; Forgie, I.; Mannelli, M.; Carney, M.; Hernandez, B.; Campos, B.; Herold-Mende, C.; Jungk, C.; Unterberg, A.; von Deimling, A.; Bossler, A.; Galbraith, J.; Jacobus, L.; Knudson, M.; Knutson, T.; Ma, D.; Milhem, M.; Sigmund, R.; Godwin, A.K.; Madan, R.; Rosenthal, H.G.; Adebamowo, C.; Adebamowo, S.N.; Boussioutas, A.; Beer, D.; Giordano, T.; Mes-Masson, A.M.; Saad, F.; Bocklage, T.; Landrum, L.; Mannel, R.; Moore, K.; Moxley, K.; Postier, R.; Walker, J.; Zuna, R.; Feldman, M.; Valdivieso, F.; Dhir, R.; Luketich, J.; Mora Pinero, E.M.; Quintero-Aguilo, M.; Carlotti, C.G.; Dos Santos, J.S.; Kemp, R.; Sankarankuty, A.; Tirapelli, D.; Catto, J.; Agnew, K.; Swisher, E.; Creaney, J.; Robinson, B.; Shelley, C.S.; Godwin, E.M.; Kendall, S.; Shipman, C.; Bradford, C.; Carey, T.; Haddad, A.; Moyer, J.; Peterson, L.; Prince, M.; Rozek, L.; Wolf, G.; Bowman, R.; Fong, K.M.; Yang, I.; Korst, R.; Rathmell, W.K.; Fantacone-Campbell, J.L.; Hooke, J.A.; Kovatich, A.J.; Shriver, C.D.; DiPersio, J.; Drake, B.; Govindan, R.; Heath, S.; Ley, T.; Van Tine, B.; Westervelt, P.; Rubin, M.A.; Lee, J. II; Aredes, N.D.; Mariamidze, A.; Hu, H. An Integrated TCGA Pan-Cancer Clinical Data Resource to Drive High-Quality Survival Outcome Analytics. *Cell*, **2018**.

[36] Warde-Farley, D.; Donaldson, S.L.; Comes, O.; Zuberi, K.; Badrawi, R.; Chao, P.; Franz, M.; Grouios, C.; Kazi, F.; Lopes, C.T.; Maitland, A. The GeneMANIA prediction server: biological network integration for gene prioritization and predicting gene function. *Nucleic acids research*, **2010**.

[37] Shukla, N., Siva, N., Sivakumar, M. B., Parveen, R., Mishra, A. K., Shah, A., Medicherla, K. M. and Suravajhala, P. (2021). Extraction of DNA and RNA from Formalin-fixed Paraffin-embedded Tissue Specimens. *Bio-101*: e4095. DOI:

10.21769/BioProtoc.4095. Shukla, N., Siva, N., Sivakumar, M. B., Parveen, R., Mishra, A. K., Shah, A., Medicherla, K. M. and Suravajhala, P. (2021). Extraction of DNA and RNA from Formalin-fixed Paraffin-embedded Tissue Specimens. *Bio-101*: e4095. DOI: 10.21769/BioProtoc.4095.

[38] Assenov, Y.; Ramírez, F.; Schelhorn, S.E.; Lengauer, T.; Albrecht, M. Computing topological parameters of biological networks. *Bioinformatics*, **2008**.

[39] Sanchez-Vega, F.; Mina, M.; Armenia, J.; Chatila, W.K.; Luna, A.; La, K.C.; Dimitriadou, S.; Liu, D.L.; Kantheti, H.S.; Saghafein, S.; Chakravarty, D.; Daian, F.; Gao, Q.; Bailey, M.H.; Liang, W.W.; Foltz, S.M.; Shmulevich, I.; Ding, L.; Heins, Z.; Ochoa, A.; Gross, B.; Gao, J.; Zhang, H.; Kundra, R.; Kandoth, C.; Bahceci, I.; Dervishi, L.; Dogrusoz, U.; Zhou, W.; Shen, H.; Laird, P.W.; Way, G.P.; Greene, C.S.; Liang, H.; Xiao, Y.; Wang, C.; Iavarone, A.; Berger, A.H.; Bivona, T.G.; Lazar, A.J.; Hammer, G.D.; Giordano, T.; Kwong, L.N.; McArthur, G.; Huang, C.; Tward, A.D.; Frederick, M.J.; McCormick, F.; Meyerson, M.; Caesar-Johnson, S.J.; Demchok, J.A.; Felau, I.; Kasapi, M.; Ferguson, M.L.; Hutter, C.M.; Sofia, H.J.; Tarnuzzer, R.; Wang, Z.; Yang, L.; Zenklusen, J.C.; Zhang, J. (Julia); Chudamani, S.; Liu, J.; Lolla, L.; Naresh, R.; Pihl, T.; Sun, Q.; Wan, Y.; Wu, Y.; Cho, J.; DeFreitas, T.; Frazer, S.; Gehlenborg, N.; Getz, G.; Heiman, D.I.; Kim, J.; Lawrence, M.S.; Lin, P.; Meier, S.; Noble, M.S.; Saksena, G.; Voet, D.; Zhang, H.; Bernard, B.; Chambwe, N.; Dhankani, V.; Knijnenburg, T.; Kramer, R.; Leinonen, K.; Liu, Y.; Miller, M.; Reynolds, S.; Shmulevich, I.; Thorsson, V.; Zhang, W.; Akbani, R.; Broom, B.M.; Hegde, A.M.; Ju, Z.; Kanchi, R.S.; Korkut, A.; Li, J.; Liang, H.; Ling, S.; Liu, W.; Lu, Y.; Mills, G.B.; Ng, K.S.; Rao, A.; Ryan, M.; Wang, J.; Weinstein, J.N.; Zhang, J.; Abeshouse, A.; Armenia, J.; Chakravarty, D.; Chatila, W.K.; de Bruijn, I.; Gross, B.E.; Heins, Z.J.; Kundra, R.; La, K.; Ladanyi, M.; Luna, A.; Nissan, M.G.; Ochoa, A.; Phillips, S.M.; Reznik, E.; Sanchez-Vega, F.; Sander, C.; Schultz, N.; Sheridan, R.; Sumer, S.O.; Sun, Y.; Taylor, B.S.; Wang, J.; Zhang, H.; Anur, P.; Peto, M.; Spellman, P.; Benz, C.; Stuart, J.M.; Wong, C.K.; Yau, C.; Hayes, D.N.; Parker, J.S.; Wilkerson, M.D.; Ally, A.; Balasundaram, M.; Bowlby, R.; Brooks, D.; Carlsen, R.; Chuah, E.; Dhalla, N.; Holt, R.; Jones, S.J.M.; Kasaian, K.; Lee, D.; Ma, Y.; Marra, M.A.; Mayo, M.; Moore, R.A.; Mungall, A.J.; Mungall, K.; Robertson, A.G.; Sadeghi, S.; Schein, J.E.; Sipahimalani, P.; Tam, A.; Thiessen, N.; Tse, K.; Wong, T.; Berger, A.C.; Beroukhi, R.; Cherniack, A.D.; Cibulskis, C.; Gabriel, S.B.; Gao, G.F.; Ha, G.; Meyerson, M.; Schumacher, S.E.; Shih, J.; Kucherlapati, M.H.; Kucherlapati, R.S.; Baylin, S.; Cope, L.; Danilova, L.; Bootwalla, M.S.; Lai, P.H.; Maglinte, D.T.; Van Den Berg, D.J.; Weisenberger, D.J.; Auman, J.T.; Balu, S.; Bodenheimer, T.; Fan, C.; Hoadley, K.A.; Hoyle, A.P.; Jefferys, S.R.; Jones, C.D.; Meng, S.; Mieczkowski, P.A.; Mose, L.E.; Perou, A.H.; Perou, C.M.; Roach, J.; Shi, Y.; Simons, J. V.; Skelly, T.; Soloway, M.G.; Tan, D.; Veluvolu, U.; Fan, H.; Hinoue, T.; Laird, P.W.; Shen, H.; Zhou, W.; Bellair, M.; Chang, K.; Covington, K.; Creighton, C.J.; Dinh, H.; Doddapaneni, H.V.; Donehower, L.A.; Drummond, J.; Gibbs, R.A.; Glenn, R.; Hale, W.; Han, Y.; Hu, J.; Korchina, V.; Lee, S.; Lewis, L.; Li, W.; Liu, X.; Morgan, M.; Morton, D.; Muzny, D.; Santibanez, J.; Sheth, M.; Shinbrot, E.; Wang, L.; Wang, M.; Wheeler, D.A.; Xi, L.; Zhao, F.; Hess, J.; Appelbaum, E.L.; Bailey, M.; Cordes, M.G.; Ding, L.; Fronick, C.C.; Fulton, L.A.; Fulton, R.S.; Kandoth, C.; Mardis, E.R.; McLellan, M.D.; Miller, C.A.; Schmidt, H.K.; Wilson, R.K.; Crain, D.; Curley, E.; Gardner, J.; Lau, K.; Mallory, D.; Morris, S.; Paulauskis, J.; Penny, R.; Shelton, C.; Shelton, T.; Sherman, M.; Thompson, E.; Yena, P.; Bowen, J.; Gastier-Foster, J.M.; Gerken, M.; Leraas, K.M.; Lichtenberg, T.M.; Ramirez, N.C.; Wise, L.; Zmuda, E.; Corcoran, N.; Costello, T.; Hovens, C.; Carvalho, A.L.; de Carvalho, A.C.; Fregnani, J.H.; Longatto-Filho, A.; Reis, R.M.; Scapulatempo-Neto, C.; Silveira, H.C.S.; Vidal, D.O.; Burnette, A.; Eschbacher, J.; Hermes, B.; Noss, A.; Singh, R.; Anderson, M.L.; Castro, P.D.; Ittmann, M.; Huntsman, D.; Kohl, B.; Le, X.; Thorp, R.; Andry, C.; Duffy, E.R.; Lyadov, V.; Paklina, O.; Setdikova, G.; Shabunin, A.; Tavobilo, M.; McPherson, C.; Warnick, R.; Berkowitz, R.; Cramer, D.; Feltmate, C.; Horowitz, N.; Kibel, A.; Muto, M.; Raut, C.P.; Malykh, A.; Barnholtz-Sloan, J.S.; Barrett, W.; Devine, K.; Fulop, J.; Ostrom, Q.T.; Shimmel, K.; Wolinsky, Y.; Sloan, A.E.; De Rose, A.; Giulianti, F.; Goodman, M.; Karlan, B.Y.; Hagedorn, C.H.; Eckman, J.; Harr, J.; Myers, J.; Tucker, K.; Zach, L.A.; Deyarmin, B.; Hu, H.; Kvecher, L.; Larson, C.; Mural, R.J.; Somiari, S.; Vicha, A.; Zelinka, T.; Bennett, J.; Iacocca, M.; Rabeno, B.; Swanson, P.; Latour, M.; Lacombe, L.; Têtu, B.; Bergeron, A.; McGraw, M.; Staugaitis, S.M.; Chabot, J.; Hibshoosh, H.; Sepulveda, A.; Su, T.; Wang, T.; Potapova, O.; Voronina, O.; Desjardins, L.; Mariani, O.; Roman-Roman, S.; Sastre, X.; Stern, M.H.; Cheng, F.; Signoretti, S.; Berchuck, A.; Bigner, D.; Lipp, E.; Marks, J.; McCall, S.; McLendon, R.; Secord, A.; Sharp, A.; Behera, M.; Brat, D.J.; Chen, A.; Delman, K.; Force, S.; Khuri, F.; Magliocca, K.; Maithel, S.; Olson, J.J.; Owonikoko, T.; Pickens, A.; Ramalingam, S.; Shin, D.M.; Sica, G.; Van Meir, E.G.; Zhang, H.; Eijckenboom, W.; Gillis, A.; Korpershoek, E.; Looijenga, L.; Oosterhuis, W.; Stoop, H.; van Kessel, K.E.; Zwarthoff, E.C.; Calatrazzolo, C.; Cuppini, L.; Cuzzubbo, S.; DiMeco, F.; Finocchiaro, G.; Mattei, L.; Perin, A.; Pollo, B.; Chen, C.; Houck, J.; Lohavanichbutr, P.; Hartmann, A.; Stoehr, C.; Stoehr, R.; Taubert, H.; Wach, S.; Wullich, B.; Kyler, W.; Murawa, D.; Wiznerowicz, M.; Chung, K.; Edenfield, W.J.; Martin, J.; Baudin, E.; Bubley, G.; Bueno, R.; De Rienzo, A.; Richards, W.G.; Kalkanis, S.; Mikkelsen, T.; Noushmehr, H.; Scarpace, L.; Girard, N.; Aymerich, M.; Campo, E.; Giné, E.; Guillermo, A.L.; Van Bang, N.; Hanh, P.T.; Phu, B.D.; Tang, Y.; Colman, H.; Evason, K.; Dottino, P.R.; Martignetti, J.A.; Gabra, H.;

Juhl, H.; Akeredolu, T.; Stepa, S.; Hoon, D.; Ahn, K.; Kang, K.J.; Beuschlein, F.; Breggia, A.; Birrer, M.; Bell, D.; Borad, M.; Bryce, A.H.; Castle, E.; Chandan, V.; Cheville, J.; Copland, J.A.; Farnell, M.; Flotte, T.; Giana, N.; Ho, T.; Kendrick, M.; Kocher, J.P.; Kopp, K.; Moser, C.; Nagorney, D.; O'Brien, D.; O'Neill, B.P.; Patel, T.; Petersen, G.; Que, F.; Rivera, M.; Roberts, L.; Smallridge, R.; Smyrk, T.; Stanton, M.; Thompson, R.H.; Torbenson, M.; Yang, J.D.; Zhang, L.; Brimo, F.; Ajani, J.A.; Gonzalez, A.M.A.; Behrens, C.; Bondaruk, J.; Broaddus, R.; Czerniak, B.; Esmaeli, B.; Fujimoto, J.; Gershenwald, J.; Guo, C.; Logothetis, C.; Meric-Bernstam, F.; Moran, C.; Ramondetta, L.; Rice, D.; Sood, A.; Tamboli, P.; Thompson, T.; Troncso, P.; Tsao, A.; Wistuba, I.; Carter, C.; Haydu, L.; Hersey, P.; Jakrot, V.; Kakavand, H.; Kefford, R.; Lee, K.; Long, G.; Mann, G.; Quinn, M.; Saw, R.; Scolyer, R.; Shannon, K.; Spillane, A.; Stretch, J.; Synott, M.; Thompson, J.; Wilmott, J.; Al-Ahmadie, H.; Chan, T.A.; Ghossein, R.; Gopalan, A.; Levine, D.A.; Reuter, V.; Singer, S.; Singh, B.; Tien, N.V.; Broudy, T.; Mirsaidi, C.; Nair, P.; Drwiega, P.; Miller, J.; Smith, J.; Zaren, H.; Park, J.W.; Hung, N.P.; Kebebew, E.; Linehan, W.M.; Metwalli, A.R.; Pacak, K.; Pinto, P.A.; Schiffman, M.; Schmidt, L.S.; Vocke, C.D.; Wentzensen, N.; Worrell, R.; Yang, H.; Moncrieff, M.; Goparaju, C.; Melamed, J.; Pass, H.; Botnariuc, N.; Caraman, I.; Cernat, M.; Chemencedji, I.; Clipca, A.; Doruc, S.; Gorincioi, G.; Mura, S.; Pirtac, M.; Stancul, I.; Tcaciuc, D.; Albert, M.; Alexopoulou, I.; Arnaout, A.; Bartlett, J.; Engel, J.; Gilbert, S.; Parfitt, J.; Sekhon, H.; Thomas, G.; Rassl, D.M.; Rintoul, R.C.; Bifulco, C.; Tamakawa, R.; Urba, W.; Hayward, N.; Timmers, H.; Antenucci, A.; Facciolo, F.; Grazi, G.; Marino, M.; Merola, R.; de Krijger, R.; Gimenez-Roqueplo, A.P.; Piché, A.; Chevalier, S.; McKercher, G.; Birsoy, K.; Barnett, G.; Brewer, C.; Farver, C.; Naska, T.; Pennell, N.A.; Raymond, D.; Schilero, C.; Smolenski, K.; Williams, F.; Morrison, C.; Borgia, J.A.; Liptay, M.J.; Pool, M.; Seder, C.W.; Junker, K.; Omberg, L.; Dinkin, M.; Manikhas, G.; Alvaro, D.; Bragazzi, M.C.; Cardinale, V.; Carpino, G.; Gaudio, E.; Chesla, D.; Cottingham, S.; Dubina, M.; Moiseenko, F.; Dhanasekaran, R.; Becker, K.F.; Janssen, K.P.; Slotta-Huspenina, J.; Abdel-Rahman, M.H.; Aziz, D.; Bell, S.; Cebulla, C.M.; Davis, A.; Duell, R.; Elder, J.B.; Hilty, J.; Kumar, B.; Lang, J.; Lehman, N.L.; Mandt, R.; Nguyen, P.; Pilarski, R.; Rai, K.; Schoenfield, L.; Senecal, K.; Wakely, P.; Hansen, P.; Lechan, R.; Powers, J.; Tischler, A.; Grizzle, W.E.; Sexton, K.C.; Kastl, A.; Henderson, J.; Porten, S.; Waldmann, J.; Fassnacht, M.; Asa, S.L.; Schadendorf, D.; Couce, M.; Graefen, M.; Huland, H.; Sauter, G.; Schlomm, T.; Simon, R.; Tennstedt, P.; Olabode, O.; Nelson, M.; Bathe, O.; Carroll, P.R.; Chan, J.M.; Disaia, P.; Glenn, P.; Kelley, R.K.; Landen, C.N.; Phillips, J.; Prados, M.; Simko, J.; Smith-McCune, K.; VandenBerg, S.; Roggin, K.; Fehrenbach, A.; Kendler, A.; Sifri, S.; Steele, R.; Jimeno, A.; Carey, F.; Forgie, I.; Mannelli, M.; Carney, M.; Hernandez, B.; Campos, B.; Herold-Mende, C.; Jungk, C.; Unterberg, A.; von Deimling, A.; Bossler, A.; Galbraith, J.; Jacobus, L.; Knudson, M.; Knutson, T.; Ma, D.; Milhem, M.; Sigmund, R.; Godwin, A.K.; Madan, R.; Rosenthal, H.G.; Adebamowo, C.; Adebamowo, S.N.; Boussioutas, A.; Beer, D.; Giordano, T.; Mes-Masson, A.M.; Saad, F.; Bocklage, T.; Landrum, L.; Mannel, R.; Moore, K.; Moxley, K.; Postier, R.; Walker, J.; Zuna, R.; Feldman, M.; Valdivieso, F.; Dhir, R.; Luketich, J.; Pinero, E.M.M.; Quintero-Aguilo, M.; Carlotti, C.G.; Dos Santos, J.S.; Kemp, R.; Sankarankuty, A.; Tirapelli, D.; Catto, J.; Agnew, K.; Swisher, E.; Creaney, J.; Robinson, B.; Shelley, C.S.; Godwin, E.M.; Kendall, S.; Shipman, C.; Bradford, C.; Carey, T.; Haddad, A.; Moyer, J.; Peterson, L.; Prince, M.; Rozek, L.; Wolf, G.; Bowman, R.; Fong, K.M.; Yang, I.; Korst, R.; Rathmell, W.K.; Fantacone-Campbell, J.L.; Hooke, J.A.; Kovatich, A.J.; Shriver, C.D.; DiPersio, J.; Drake, B.; Govindan, R.; Heath, S.; Ley, T.; Van Tine, B.; Westervelt, P.; Rubin, M.A.; Lee, J. II; Aredes, N.D.; Mariamidze, A.; Van Allen, E.M.; Cherniack, A.D.; Ciriello, G.; Sander, C.; Schultz, N. Oncogenic Signaling Pathways in The Cancer Genome Atlas. *Cell*, **2018**.

[40] Lu, Q. δ -Catenin Dysregulation in Cancer: Interactions with E-Cadherin and Beyond. *Journal of Pathology*, **2010**.

[41] Kantidze, O.L.; Kamalyukova, I.M.; Razin, S. V. Association of the Mammalian Transcriptional Regulator Kaiso with Centrosomes and the Midbody. *Cell Cycle*, **2009**.

[42] Schackmann, R.C.J.; Tenhagen, M.; van de Ven, R.A.H.; Derksen, P.W.B. P120-Catenin in Cancer - Mechanisms, Models and Opportunities for Intervention. *J. Cell Sci.*, **2013**.

[43] Shukla, N.; Prasad, A.; Kanga, U.; Suravajhala, R.; Nigam, V.K.; Kishor, P.B.K.; Polavarapu, R.; Chaubey, G.; Singh, K.K.; Suravajhala, P. SARS-CoV-2 Transgressing LncRNAs Uncovers the Known Unknowns. *Physiol. Genomics*, **2021**.

[44] Burnell, S.E.A.; Spencer-Harty, S.; Howarth, S.; Bodger, O.; Kynaston, H.; Morgan, C.; Doak, S.H. STEAP2 Knockdown Reduces the Invasive Potential of Prostate Cancer Cells. *Sci. Rep.*, **2018**.

[45] Stamey, T.A.; Yang, N.; Hay, A.R.; McNeal, J.E.; Freiha, F.S.; Redwine, E. Prostate-Specific Antigen as a Serum Marker for Adenocarcinoma of the Prostate. *N. Engl. J. Med.*, **1987**.

[46] Ahn, J.; Berndt, S.I.; Wacholder, S.; Kraft, P.; Kibel, A.S.; Yeager, M.; Albanes, D.; Giovannucci, E.; Stampfer, M.J.; Virtamo, J.; Thun, M.J.; Feigelson, H.S.; Cancel-Tassin, G.; Cussenot, O.; Thomas, G.; Hunter, D.J.; Fraumeni, J.F.; Hoover, R.N.; Chanock, S.J.; Hayes, R.B. Variation in KLK Genes, Prostate-Specific Antigen and Risk of Prostate Cancer. *Nature Genetics*, **2008**.

- [47] Shang, Z.; Niu, Y.; Cai, Q.; Chen, J.; Tian, J.; Yeh, S.; Lai, K.P.; Chang, C. Human Kallikrein 2 (KLK2) Promotes Prostate Cancer Cell Growth via Function as a Modulator to Promote the ARA70-Enhanced Androgen Receptor Transactivation. *Tumor Biol.*, **2014**.
- [48] McDevitt MR, Thorek DL, Hashimoto T, Gondo T, Veach DR, Sharma SK, Kalidindi TM, Abou DS, Watson PA, Beattie BJ, Timmermand OV. Feed-forward alpha particle radiotherapy ablates androgen receptor-addicted prostate cancer. *Nature communications*. **2018**, 24;9(1):1-1.
- [49] Grimm, J.; Sachs, M.; Britsch, S.; Di Cesare, S.; Schwarz-Romond, T.; Alitalo, K.; Birchmeier, W. Novel P62dok Family Members, Dok-4 and Dok-5, Are Substrates of the c-Ret Receptor Tyrosine Kinase and Mediate Neuronal Differentiation. *J. Cell Biol.*, **2001**.
- [50] Luo, F.; Wang, Z.; Chen, S.; Luo, Z.; Wang, G.; Yang, H.; Tang, L. DOK5 as a Prognostic Biomarker of Gastric Cancer Immunoinvasion: A Bioinformatics Analysis. *Biomed Res. Int.*, **2022**.
- [51] Meng, J.; Wang, J. Role of SNARE Proteins in Tumourigenesis and Their Potential as Targets for Novel Anti-Cancer Therapeutics. *Biochimica et Biophysica Acta - Reviews on Cancer*, **2015**.
- [52] Han, J.; Pluhackova, K.; Böckmann, R.A. The Multifaceted Role of SNARE Proteins in Membrane Fusion. *Frontiers in Physiology*, **2017**.
- [53] Yan, R. Physiological Functions of the β -Site Amyloid Precursor Protein Cleaving Enzyme 1 and 2. *Frontiers in Molecular Neuroscience*, **2017**.
- [54] Farris, F.; Matafora, V.; Bachi, A. The Emerging Role of β -Secretases in Cancer. *Journal of Experimental and Clinical Cancer Research*, **2021**.
- [55] Amaral, P.P.; Neyt, C.; Wilkins, S.J.; Askarian-Amiri, M.E.; Sunkin, S.M.; Perkins, A.C.; Mattick, J.S. Complex Architecture and Regulated Expression of the Sox2ot Locus during Vertebrate Development. *RNA*, **2009**.
- [56] Li, Y.; Du, M.; Wang, S.; Zha, J.; Lei, P.; Wang, X.; Wu, D.; Zhang, J.; Chen, D.; Huang, D.; Lu, J.; Li, H.; Sun, M. Clinicopathological Implication of Long Non-Coding RNAs SOX2 Overlapping Transcript and Its Potential Target Gene Network in Various Cancers. *Frontiers in Genetics*, **2020**.
- [57] Li, S.; Zhang, Q.; Liu, W.; Zhao, C. Silencing of FTX Suppresses Pancreatic Cancer Cell Proliferation and Invasion by Upregulating MiR-513b-5p. *BMC Cancer*, **2021**.
- [58] Huo, X.; Wang, H.; Huo, B.; Wang, L.; Yang, K.; Wang, J.; Wang, L.; Wang, H. FTX Contributes to Cell Proliferation and Migration in Lung Adenocarcinoma via Targeting MiR-335-5p/NUCB2 Axis. *Cancer Cell Int.*, **2020**.
- [59] Kong, D.; Zhao, Q.; Liu, W.; Wang, F. Identification of Crucial MiRNAs and lncRNAs for Ossification of Ligamentum Flavum. *Mol. Med. Rep.*, **2019**.
- [60] Ren, S., Peng, Z., Mao, J.H. *et al.* 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 (2012). <https://doi.org/10.1038/cr.2012.30>
- [61] Rajagopal, S.; Sharma, A.; Simlot, A.; Mathur, P.; Mehta, S.; Mehta, S.; Naravula, J.; Medicherla, K.M.; Kumar, A.; Kanga, U.; Suravajhala, R.; Bhandari, R.K.; Nair, B.G.; Kishor, P.B.K.; Suravajhala, P. Inferring *bona fide* Differentially Expressed Genes and Their Variants Associated with Vitamin K Deficiency Using a Systems Genetics Approach. *Genes* **2022**, 13, 2078. <https://doi.org/10.3390/genes13112078>.
- [62] Guo H, Ahmed M, Zhang F, Yao CQ, Li S, Liang Y, Hua J, Soares F, Sun Y, Langstein J, Li Y, Poon C, Bailey SD, Desai K, Fei T, Li Q, Sendorek DH, Fraser M, Prensner JR, Pugh TJ, Pomerantz M, Bristow RG, Lupien M, Feng FY, Boutros PC, Freedman ML, Walsh MJ, He HH. Modulation of long noncoding RNAs by risk SNPs underlying genetic predispositions to prostate cancer. *Nat Genet.* 2016 Oct;48(10):1142-50. doi: 10.1038/ng.3637. Epub 2016 Aug 15. PMID: 27526323.
- [63] Ji P, Diederichs S, Wang W, Böing S, Metzger R, Schneider PM, Tidow N, Brandt B, Buerger H, Bulk E, Thomas M, Berdel WE, Serve H, Müller-Tidow C. MALAT-1, a novel noncoding RNA, and thymosin beta4 predict metastasis and survival in early-stage non-small cell lung cancer. *Oncogene.* 2003 Sep 11;22(39):8031-41. doi: 10.1038/sj.onc.1206928. PMID: 12970751.
- [64] Ren S, Liu Y, Xu W, Sun Y, Lu J, Wang F, Wei M, Shen J, Hou J, Gao X, Xu C, Huang J, Zhao Y, Sun Y. Long noncoding RNA MALAT-1 is a new potential therapeutic target for castration resistant prostate cancer. *J Urol.* 2013 Dec;190(6):2278-87. doi: 10.1016/j.juro.2013.07.001. Epub 2013 Jul 8. PMID: 23845456.
- [65] Crea F, Watahiki A, Quagliata L, Xue H, Pikor L, Parolia A, Wang Y, Lin D, Lam WL, Farrar WL, Isogai T, Morant R, Castori-Eppenberger S, Chi KN, Wang Y, Helgason CD. Identification of a long non-coding RNA as a novel biomarker and potential therapeutic target for metastatic prostate cancer. *Oncotarget.* 2014 Feb 15;5(3):764-74. doi: 10.18632/oncotarget.1769. PMID: 24519926; PMCID: PMC3996663.

[66] Prensner JR, Zhao S, Erho N, Schipper M, Iyer MK, Dhanasekaran SM, Magi-Galluzzi C, Mehra R, Sahu A, Siddiqui J, Davicioni E, Den RB, Dicker AP, Karnes RJ, Wei JT, Klein EA, Jenkins RB, Chinnaiyan AM, Feng FY. RNA biomarkers associated with metastatic progression in prostate cancer: a multi-institutional high-throughput analysis of SChLAP1. *Lancet Oncol.* 2014 Dec;15(13):1469-1480. doi: 10.1016/S1470-2045(14)71113-1. Epub 2014 Nov 17. PMID: 25456366; PMCID: PMC4559342.