UNIVERSIDAD DE GRANADA

MASTER THESIS

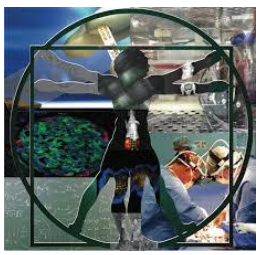
RNA-Seq Transcriptome Analysis for the Study of Prostate Cancer Development and Evolution

Esther Sauras Colón

Tutors: María Jesús Álvarez Cubero, Eduardo Andrés León

September 2019

MASTER IN TRANSLATIONAL RESEARCH AND PERSONALIZED MEDICINE



International Journal of
Molecular Sciences

MDPI

Master Thesis

RNA-Seq Transcriptome Analysis for the Study of Prostate Cancer Development and Evolution

Esther Sauras Colón [1]*, María Jesús Álvarez Cubero [2,3]* and Eduardo Andrés León [4]*

[1] Master in Translational Research and Personalized Medicine. Faculty of Medicine, University of Granada. Granada, Spain.
[2] GENYO. Centre for Genomics and Oncological Research. Pfizer / University of Granada / Andalusian Regional Government. Granada, Spain.
[3] Department of Biochemistry and Molecular Biology III. Faculty of Medicine, University of Granada. Granada, Spain.
[4] Bioinformatics Unit, Institute of Parasitology and Biomedicine "López-Neyra" (IPBLN). Spanish National Research Council (CSIC). Granada, Spain.
* Correspondence: Esther Sauras Colón: esc6@correo.ugr.es. María Jesús Álvarez Cubero: mjesusac@ugr.es. Eduardo Andrés León: eduardo.andres@csic.es.

September 2019

Abstract: Prostate cancer (PCa) is one of the most common cancers worldwide. Even though prostate specific antigen (PSA) test is the non-invasive routine blood test for the detection of asymptomatic disease, it can result in problems of diagnostic accuracy and overdiagnosis. Thus, it is necessary to continue investigating new efficient biomarkers for the prevention, diagnosis and prognosis of PCa. Here, we analyse the transcriptome of seven individuals by the use of next-generation sequencing (NGS) technique to identify differentially expressed genes, which can help to better understand PCa aggressiveness. Present analysis show that there are two upregulated genes in PCa regarding to controls: HP (Haptoglobin) and HLA-G (Human Leukocyte Antigen-G). On the other hand, there are seven downregulated genes, where TP53TG3 and their transcripts should be highlighted. Also, we make a comparison between a more aggressive PCa phenotype and the rest of PCa samples to investigate about genes implicated in aggressiveness, obtaining a high number of upregulated genes including CENPF, DLGAP5 and RRM2, among others. These genes could serve as predictive, diagnostic and prognostic biomarkers, as well as molecular targets. Nevertheless, further studies would be needed to confirm the obtained results.

Keywords: Prostate cancer; Biomarkers; Next-generation sequencing; RNA-Seq; Transcriptome; Differentially expressed genes.

1. Introduction

The prostate is a small gland located under the urinary bladder that is part of the reproductive and urinary systems in men. Changes in DNA prostate cells can lead to benign prostate hyperplasia, prostatic intraepithelial neoplasia or prostate cancer (PCa). PCa is the second most common cancer affecting men with 1.28 million incident cases worldwide in 2018, and the eighth leading cause of cancer-specific mortality in males [1]. Some of the main risk factors associated with PCa are advanced age, African-American race, geographic residence area, family history or genetic changes, among others [2].

There are three types of PCa: familial, hereditary and sporadic. Although familial and hereditary are very similar, they are not synonymous. Both of them refer to the presence of affected members within the same family, but the criterion for hereditary PCa is a family with three

generations affected, three first-degree relatives affected or two close relatives affected before 55 years old [3]. In the other hand, sporadic PCa is caused by DNA damage during the life of the individual. Familial and hereditary cancer have an incidence rate of around 15%, while in sporadic cancer the incidence rate is 80-90% [4].

The increase of incidence, but also of the overall survival, are attributed to an earlier stage at diagnosis due to the introduction of prostate specific antigen (PSA) test as a non-invasive blood test for the detection of asymptomatic disease, and advances in treatments [5]. However, PSA has low specificity. Elevated levels of PSA are a sign of PCa, but can indicate other conditions, such as prostatic hyperplasia, prostatitis or urinary tract infection, resulting in a problem of false positives that can lead to a diagnostic error [6]. Otherwise, the U.S. Preventive Services Task Force recommended not to use PSA test in asymptomatic men of all ages because a considerable percentage of them have tumors that either will not progress or will progress so slowly that it would have not presented symptoms for the rest of their life, resulting in a problem of overdiagnosis [7].

Prostate biopsy confirms the diagnosis of the tumor and its grade. For over 40 years, Gleason score has been the most accepted system for PCa grading. Gleason system scores the primary and secondary histopathologic patterns of cancer, and Gleason score is the sum of both patterns (up to 10). Nevertheless, in recent years, there have been some changes in the grading system and new histologic grade groups (from 1 to 5) have been established for PCa [8].

The American Joint Committee on Cancer published in 1977 the first edition of a staging system based on T (tumor extent), N (lymph node invasion) and M (presence or absence of metastasis) classifications. The eighth edition of the cancer staging system is currently in force. Staging is important to classify the severity of the disease, estimate the prognosis and recommend treatment. TNM staging, used in combination with tumor grade and PSA test, is a tool to stratify risk of patients with PCa, classifying it into four large groups (I to IV) with its subgroups, and is used as a basis to guide treatment decisions [8]. Stratification by these extension and progression stages with its prognostic value allows relating diagnosis and aggressiveness of PCa with potential biomarkers and their relevance in clinical practice. Due to the above problems of false positives and overdiagnosis, it is necessary to continue investigating new biomarkers for prevention, diagnosis and prognosis of PCa in order to solve these problems [9].

A biomarker is "a defined characteristic that is measured as an indicator of normal biological processes, pathogenic processes or responses to an exposure or intervention" [10]. Identification of new sensitive and specific biomarkers will lead to development of the emerging method of clinical practice known as precision medicine. In recent years, new biomarkers related to PCa have emerged, including prostate health index (PHI), four-kallikrein panel (4Kscore), prostate cancer antigen 3 (PCA3), transmembrane protease serine 2:ETS-related gene (TMPRSS2:ERG) fusion, circulating tumor cells (CTCs), microRNAs (miRNAs), phosphatase and tensin homolog (PTEN) or androgen receptor splice variant-7 (AR-V7) [11].

PHI combines the values of total PSA, free PSA and the precursor form of PSA [-2] proPSA in a specific formula for the detection of PCa. The Food and Drug Administration (FDA) approved this test in 2012 [12].

4Kscore consists of a panel that includes a combination of four kallikrein proteins (total PSA, free PSA, intact PSA and human kallikrein 2) for detection and prognosis of PCa. Opko Diagnostics commercialized this test, which also incorporates clinical information, such as age and history of prior biopsy, in an algorithm to predict the risk of aggressive PCa [13].

PCA3 is a noncoding RNA that is overexpressed in PCa compared with normal prostate tissue. The FDA approved in 2012 a test to detect PCA3 RNA expression in urine samples for the diagnosis of PCa [14,15].

TMPRSS2:ERG fusion gene was identified in urine samples of patients with PCa. Genomic rearrangements are present in hematologic malignancies and solid tumors. Specifically, the two partners of this fusion gene are TMPRSS2, an androgen regulated gene, and ERG, a member of the ETS transcription factor family. Similar to PCA3, the rearrangement TMPRSS2:ERG could improve the diagnosis of PCa [15].

CTCs are a population of cancer cells able to detach primary or metastatic tumor and enter into bloodstream through vessel walls. The study of CTCs provides a new approach to monitor tumor evolution and study development of resistance to therapy. Nowadays, the FDA only has approved the CellSearch® platform for that purpose in PCa [16].

MiRNAs are small (22 bp), non-coding, single-stranded RNA molecules, which function as regulators of gene expression through RNA silencing and post-transcriptional regulation. They play an important role in different physiological processes, such as development, differentiation, proliferation or apoptosis, but also in carcinogenesis. Some examples of miRNAs involved in PCa are let-7a, miR-145, miR-296-5p or miR-218 [17].

PTEN is a tumor suppressor gene frequently mutated or deleted in several types of cancer, including PCa. The Cancer Genome Atlas Research Network presented an analysis where PTEN was mutated or deleted in 17% of PCa cases [18].

AR-V7 is a potential biomarker for treatment selection in metastatic PCa. Patients with AR-V7, a widely studied variant of the androgen receptor (AR), are associated with resistance to certain drugs. There is a commercialized test to detect this variant known as OncotypeDX AR-V7 Nucleus Detect [19].

Androgens (testosterone and/or dihydrotestosterone) are key in the normal physiology of prostate through binding and activation of AR, but they are also implicated in the pathogenesis and progression of PCa [20]. Thus, pharmacologic treatments for metastatic or high-risk localized disease consist of androgen deprivation therapy (ADT), either alone or in combination with other therapies [21]. Despite a high initial response rate to ADT, there is a number of patients that present either a continuous rise in serum PSA levels, progression of pre-existing disease, and/or appearance of new metastases. This is known as castration resistant prostate cancer (CRPC) [22]. Many new agents have been developed for CRPC over the last few years, such as small-molecule inhibitors, vaccines or other immunomodulating agents, but it is still incurable [23].

Although the cause of PCa is still unknown, there are many genetic changes involved in its etiology. Defects in DNA damage repair play an important role in the development of malignant tumors; in fact, inherited mutations in DNA-repair genes are associated with increased risk of lethal PCa. Some of the DNA-repair implicated genes are BRCA2, ATM, CHEK2, BRCA1, RAD51D and PALB2 [24]. Moreover, several polymorphisms in the main genes involved in the synthesis and metabolism of sex hormones have been related with PCa progression and aggressiveness. Robles-Fernandez et al. found that polymorphisms in genes CYP17A1, LHCGR and ESR2 are associated with a more aggressive PCa according to clinical variables [25]. Also, Glutathione S-Transferases (GSTs) enzymes have been linked to PCa. GSTs are phase II enzymes whose function is the detoxification of potential carcinogens, steroid hormones and xenobiotics and there are evidences that GSTM1, a GST isoform, has a decreased expression in PCa patients exposed to environmental chemicals [26]. Other genes with a relevant role in cancer are oncogenes and tumor

suppressor genes. Somatic mutations in this type of genes, such as KIT, KRAS and TP53, seem to have a high presence in patients with PCa. Specifically, mutations in KIT are associated with PCa aggressiveness [27]. Furthermore, germline polymorphisms located in RNASEL, ELAC2 and MSR1 genes are implicated in PCa progression and aggressiveness [28]. One of the main problems in the search of specific biomarkers in PCa is the high heterogeneity this type of tumor presents, which makes difficult to establish a classification. A more extended review of genetic markers in PCa can be found in Cozar et al., 2017 [29].

Nowadays, one step further in the comprehension of PCa involves the use of high-throughput techniques to identify genetic markers. There have been exponential advances in genomic biology associated with new technologies, reduced costs and bioinformatics approaches to process the amount of resulting data. Thus, now it is possible to study different omics levels (genome, epigenome, transcriptome, proteome or metabolome) for research and clinical applications. Genetic markers could estimate the predisposition to cancer within affected families and the response to specific therapeutics. Mateo et al. demonstrated that patients with metastatic CRPC who have defects in DNA-repair genes present a better response to pharmacologic inhibitors of PARP1 and these type of tumors also appear to be responsive to platinum-based chemotherapy [30]. Another example that reflects the need to understand the genetics of PCa is that patients with the same tumor grade can present different outcomes [31]. Hence, one of the main challenges in the study of PCa is to find specific biomarkers.

The aim of the present work is the integration of omics data for the study of biomarkers with relevance in PCa development and progression. Here, we analyse the transcriptome of seven individuals of the same family by the use of next-generation sequencing (NGS) in order to identify differentially expressed genes, which can help to better understand the aggressiveness of the disease.

2. Results

2.1. Filtering and normalization

Firstly, data analysis requires initial filtering and normalization. Approximately around 38% of genes have zero counts throughout the seven samples; therefore, the aim of filtering is to remove these genes that are unexpressed or lowly expressed in all the samples. Instead of considering gene expression at the level of raw counts, we transform raw counts into counts per million (CPM). For a specific gene, a CPM value of 1 means having 16 counts in the sample with the lowest sequencing depth (sample 3, library size ≈16 million) or 18 counts in the sample with the greatest sequencing depth (sample 4, library size ≈18 million). The filtering criterion keeps genes whose CPM value is greater than 1 in at least one of the samples, because one is the lowest number of replicates corresponding to breast cancer (BCa) group (Table A1). Using this criterion, the number of genes is reduced from 59205 to 19833 (Figure 1).

Gene counts across different samples cannot be compared directly, as there are some external factors during sample processing that can affect to their expression. Normalization is necessary to equalize the expression distributions between all the samples (Figure 2). Normalization process is performed by the trimmed mean of M-values (TMM) method [32], which calculates a normalization factor for each sample to scale the library size (Table A1).

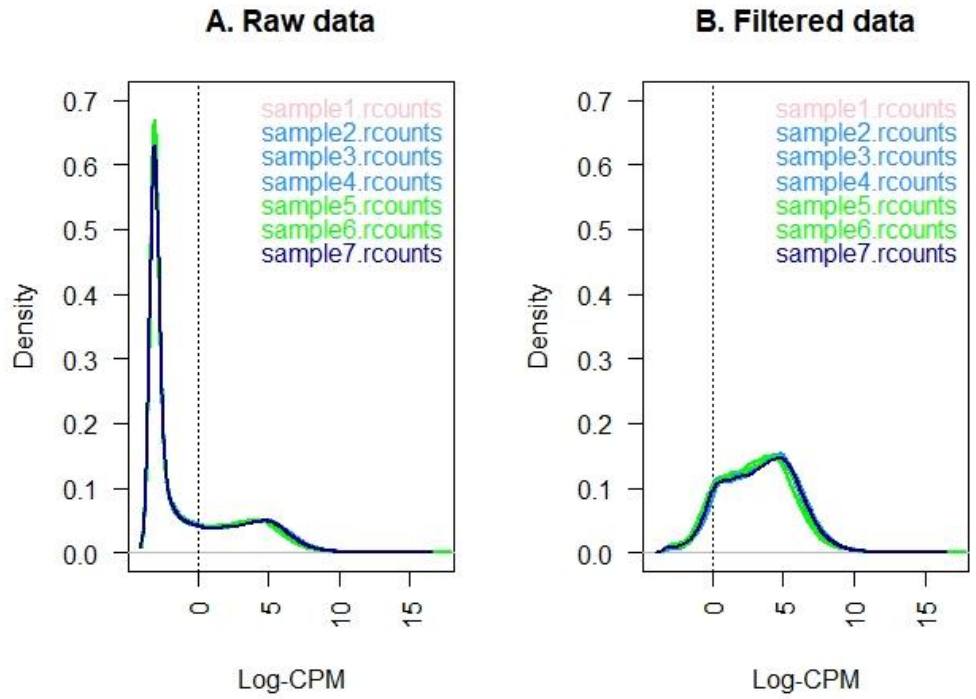


Figure 1. Density plot of log-CPM values for (A) raw data and (B) filtered data. Filtering removes genes that are unexpressed or lowly expressed in all the samples. Samples are represented in different colors corresponding to their groups: group 1 – control (green), group 2 – BCa (pink), and group 3 – PCa (blue, intensity of blue color represents the aggressiveness of PCa).

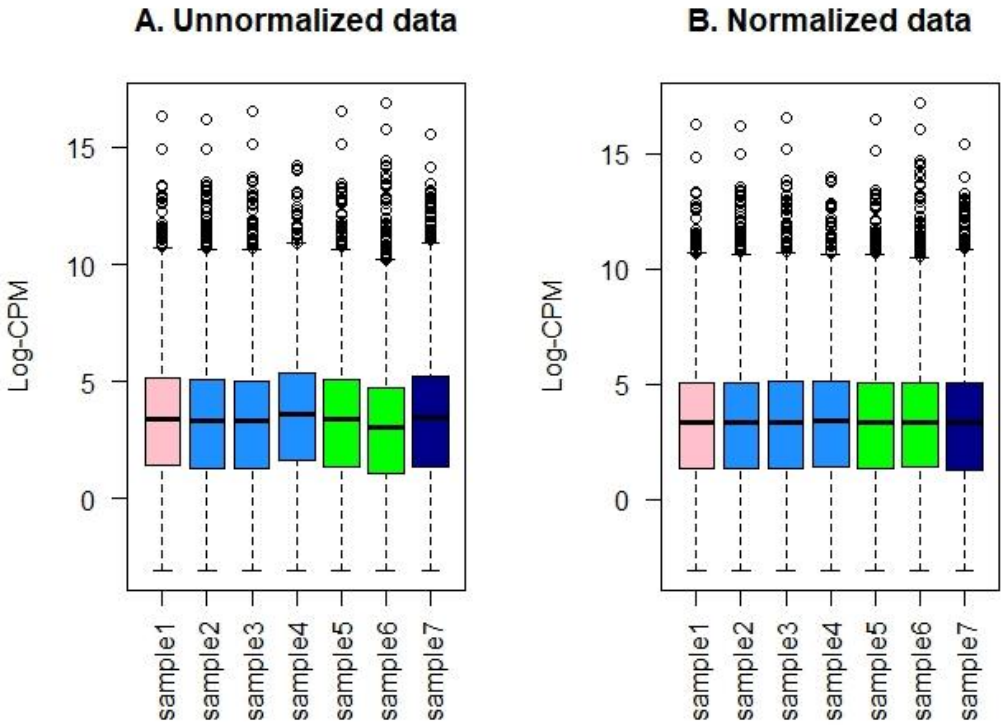


Figure 2. Boxplots of log-CPM values that reflect expression distributions of each sample for (A) unnormalized data and (B) normalized data. Normalization ensures that expression distributions of each sample are similar to make them comparable. Samples are represented in different colors corresponding to their groups: group 1 - control (green), group 2 – BCa (pink), and group 3 – PCa (blue, intensity of blue color represents the aggressiveness of PCa).

2.2. Identification of differentially expressed genes

Once filtering and normalization are performed, our goal is to identify differentially expressed genes in PCa samples. For that purpose, we compare gene expression between control group and PCa group using the exact test, which calculates differences in the means between both groups with a negative binomial distribution. The result shows a total of seven downregulated genes and two upregulated genes (Figure 3).

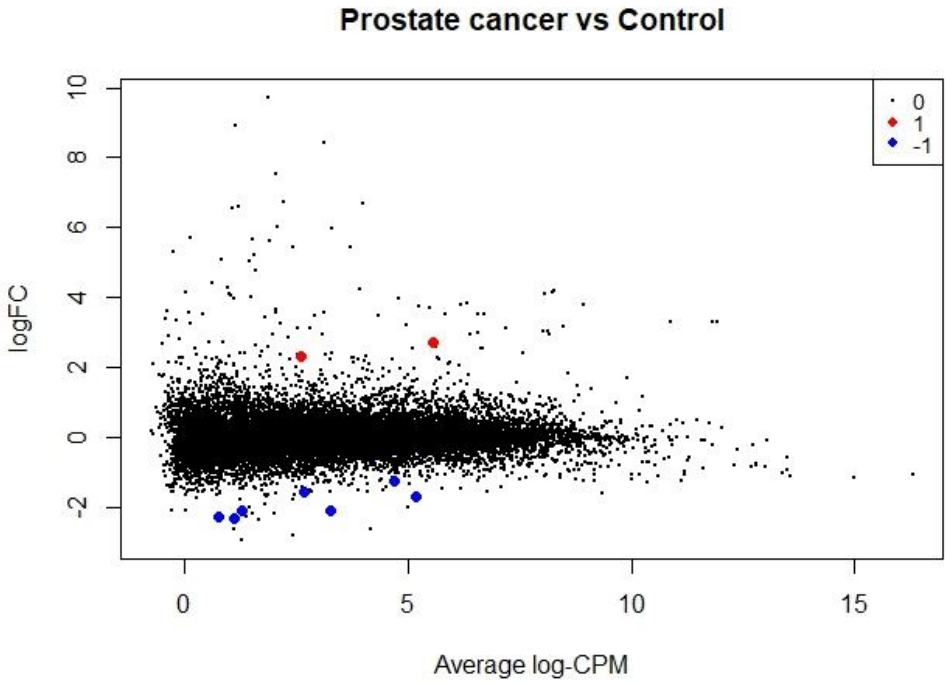


Figure 3. MA plot of the identified genes with differential expression between control group and PCa group. X axis represents the average log-CPM, while Y axis represents logFC. There are 7 downregulated genes (blue) and 2 upregulated genes (red) that are statistically significant.

Upregulated genes are HP and HLA-G, whereas downregulated genes are TP53TG3B, ZDHHC11B, TP53TG3, IGHV1-3, TP53TG3D, ZNF542P and LEPR (Table 1). A false discovery rate (FDR) value of 0.05 was established as threshold for the selection of differentially expressed genes, maintaining in the study those genes whose FDR value was lower.

Table 1. Differentially expressed genes in PCa. Upregulated genes are those with a positive logFC value, while downregulated genes have a negative logFC value. LogFC represents logarithm to the base two of the difference between expression means of a specific gene in both comparison groups. On the other hand, FDR are adjusted P-values through Benjamini-Hochberg method to control the false discovery rate in multiple testing.

Gene name	logFC	FDR
HP	2.302753	0.001844
HLA-G	2.716145	0.002491
TP53TG3B	-2.293182	0.026770
ZDHHC11B	-1.547803	0.026770
TP53TG3	-2.111520	0.026770
IGHV1-3	-2.111649	0.026770
TP53TG3D	-2.286385	0.035143
ZNF542P	-1.238904	0.035143
LEPR	-1.679310	0.047721

Later, we create a multidimensional scaling (MDS) plot to observe if samples are properly separated by their corresponding group (Figure 4). MDS plot is a method to visualize the level of similarity of the samples, so we would expect to be able to differentiate three groups: BCa, PCa and control group. Looking at the plot, control group can be differentiated clearly (samples 5 and 6), however, BCa group (sample 1) is among PCa group (samples 2, 3, 4 and 7) and cannot be separated as well as control group.

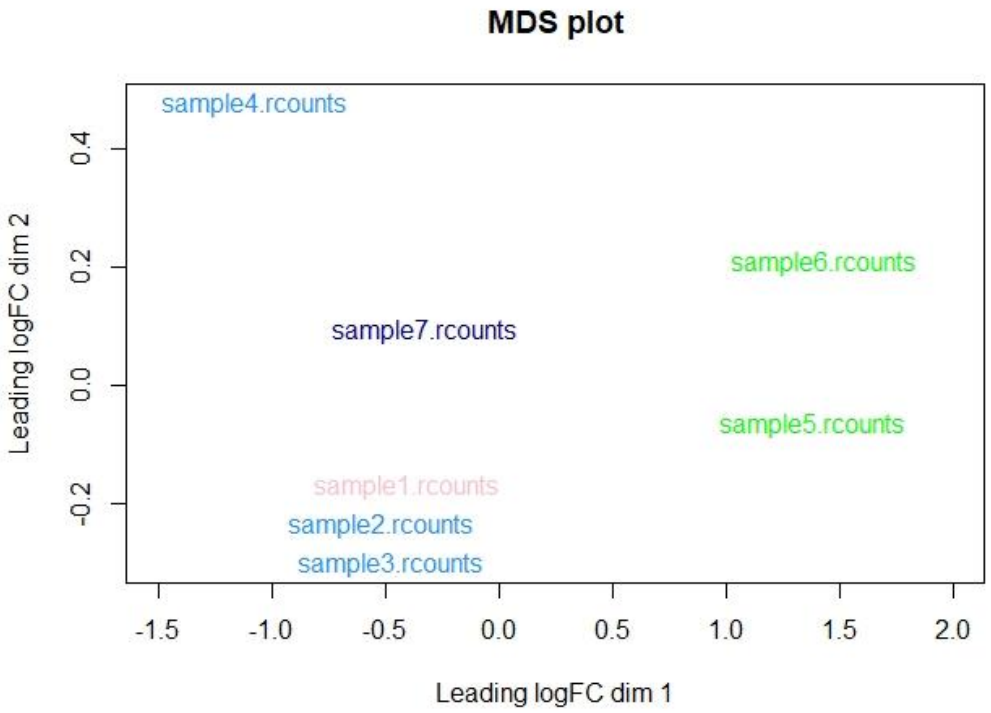


Figure 4. MDS plot. Samples are represented in different colors corresponding to their groups: group 1 - control (green), group 2 – BCa (pink), and group 3 – PCa (blue, intensity of blue color represents the aggressiveness of PCa). Control group can be differentiated clearly from the other two groups; nevertheless, sample corresponding to BCa group is among samples of PCa group and cannot be separated as well.

We continue comparing gene expression between control group and BCa group using the exact test, as previously, to find out if there are some common differentially expressed genes between BCa and PCa groups regarding to controls. There are 25 differential expressed genes in BCa, including 16 downregulated genes and 9 upregulated genes, and 2 of them are also significant in PCa (Figure 5). Common genes are HP and HLA-G, which are included in the list of differentially expressed genes in BCa (Table A2). These genes are upregulated, as it happened in PCa, so maybe they are the reason why samples of BCa and PCa groups are located so close in the previous MDS plot.

Finally, one of the patients presents a more aggressive phenotype of PCa. Thus, it is interesting also to make a comparison between this aggressive phenotype of PCa (sample 7) and the rest of PCa samples to investigate about genes implicated in aggressiveness (Figure 6).

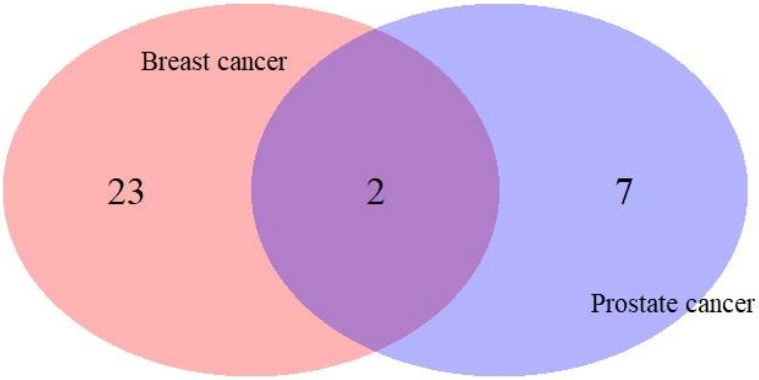


Figure 5. Venn diagram. Pink represents BCa, whereas blue represents PCa. The diagram shows that there are 25 differential expressed genes in breast cancer and 9 differential expressed genes in prostate cancer, 2 of which are common in both cases. Area of the circles is not proportional to the group size.

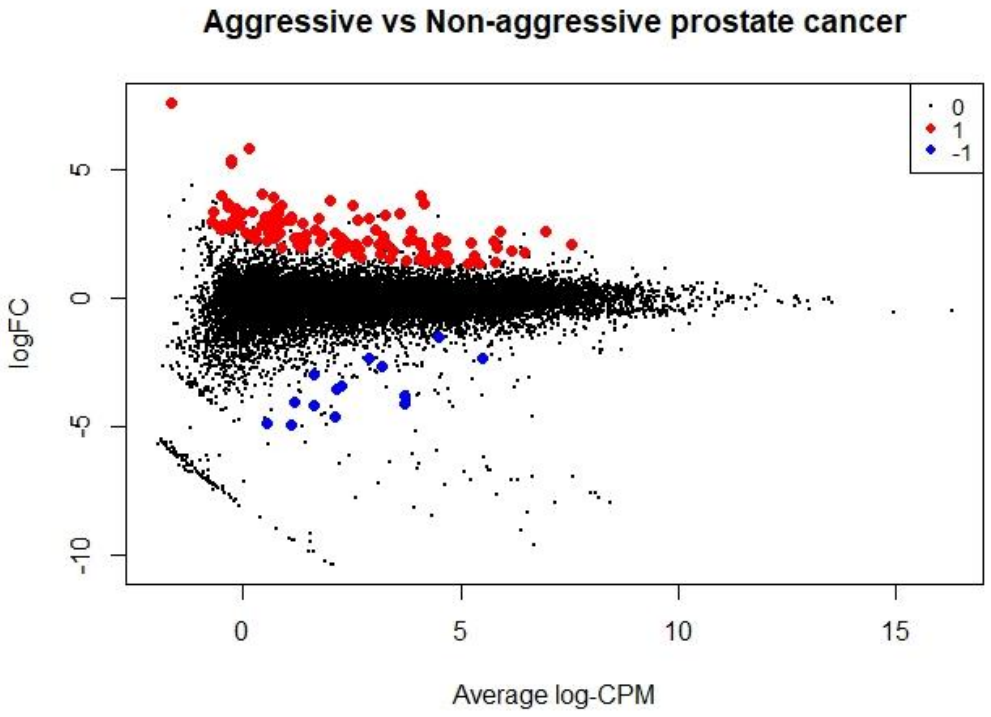


Figure 6. MA plot of the identified genes with differential expression between aggressive PCa sample and the rest of PCa samples. X axis represents the average log-CPM, while Y axis represents logFC. There are 14 downregulated genes (blue) and 118 upregulated genes (red) that are statistically significant.

The graphic shows a high number of upregulated genes in aggressive PCa. There is a total of 118 upregulated genes, including CENPF, DLGAP5 or RRM2, among others (Table A3). An enrichment analysis of the upregulated genes indicates that they participate in biological processes, such as mitotic cell cycle, cell cycle phase transition, chromosome segregation or nuclear division. On the other hand, there are 14 downregulated genes (Table A3). A FDR value of 0.05 was established as threshold for the selection of differentially expressed genes, maintaining in the study those genes whose FDR value was lower.

Taking into account the differentially expressed genes in aggressive PCa with a FDR value lower than 0.05, we create a heat map (Figure 7). It represents the expression level of the selected genes in each individual sample and generates a dendrogram using the Euclidean distance as a similarity measure. Heat map makes it easier to visualize if there are some genes with different expression in any of the samples and dendrogram groups these genes into clusters. Otherwise, the result shows that samples are clustered by their corresponding groups as follows: PCa group (samples 2, 3 and 4), BCa group (sample 1), control group (samples 5 and 6) and aggressive PCa group (sample 7).

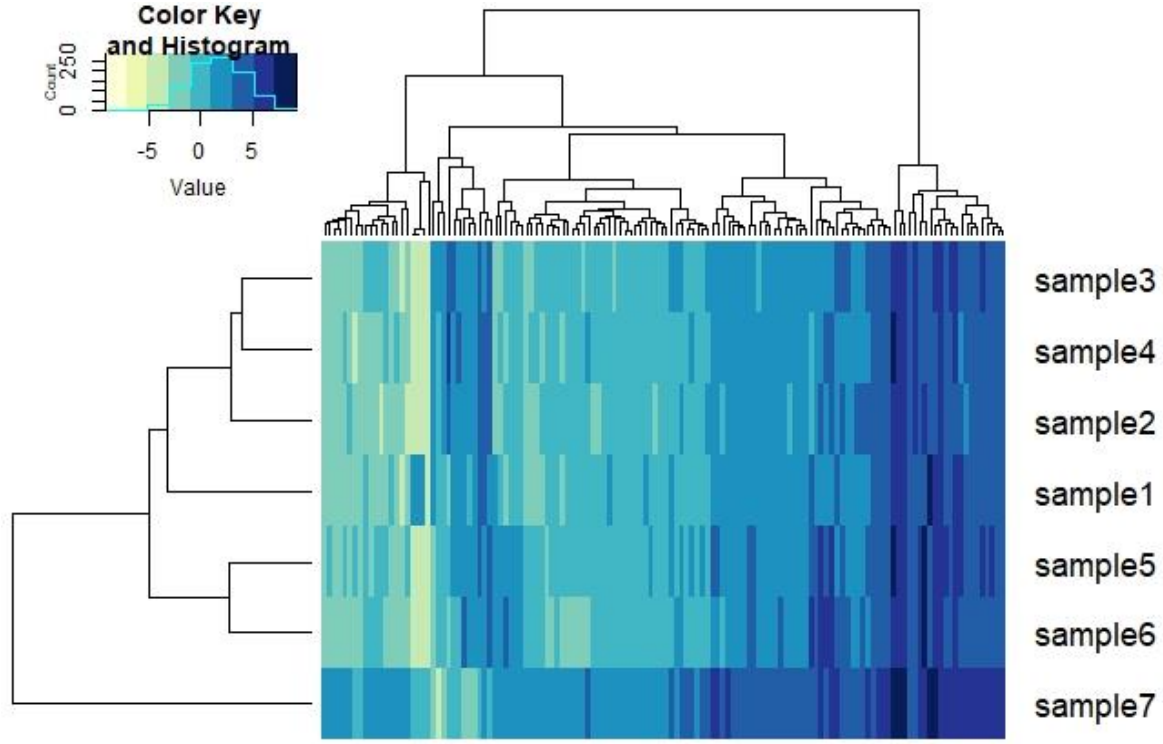


Figure 7. Heat map for statistically significant differentially expressed genes in aggressive PCa. A FDR value of 0.05 was established as threshold. Rows represent each sample and columns represent different genes. The cells in the matrix represent the expression level of a gene in an individual sample (blue and yellow in cells reflect high and low expression levels, respectively).

3. Discussion

Hereditary breast and ovarian cancer (HBOC) syndrome is characterized by the presence of mutations in breast cancer type 1 (BRCA1) and breast cancer type 2 (BRCA2) genes with a higher risk of developing breast and ovarian cancer, as well as other solid tumors as PCa, pancreatic cancer and melanoma [33]. Specifically, PCa risk is increased in BRCA2 compared to BRCA1 mutation carriers [34]. These genes are a link between BCa and PCa, and could be the reason of several members affected by familial cancer within the same family. Nevertheless, a genetic test showed that the woman with BCa does not present mutations in BRCA1 or BRCA2. This suggests that it is not a case of HBOC syndrome.

In this study, we find other genes, such as haptoglobin (HP) and human leukocyte antigen-G (HLA-G), which are differentially expressed in both PCa and BCa groups, where they are upregulated regarding to controls.

HP gene encodes a protein known as haptoglobin. The function of haptoglobin is to capture and bind free plasma hemoglobin, so that prevents kidney damage and loss of iron. It also acts as an antioxidant, has antibacterial activity, and plays a role in modulating many aspects of the acute phase response [35]. In addition, it is involved in the pathogenesis of tumors and infections. Some studies have demonstrated that haptoglobin levels in serum increase significantly when there are inflammation processes and cancer development [36,37].

On the other hand, HLA-G encodes a non-classical human leukocyte antigen (HLA) class I molecule. This molecule is expressed in physiological conditions, but it has also been detected in different types of tumors and its role in cancer has been broadly investigated in the recent years. There are seven isoforms of HLA-G as a result of alternative splicing, including four membrane-bound (HLA-G1, HLA-G2, HLA-G3 and HLA-G4) and three soluble (HLA-G5, HLA-G6 and HLA-G7) isoforms. Heterogeneity of HLA-G in cancer implies a great variation in the levels of expression and the isoform profiles between tumor types and patients with the same type of tumor. However, HLA-G expression in cancer is related to immune suppressive microenvironments, advanced tumor stage, and poor therapeutic responses and prognosis [38].

Tumors develop mechanisms for immune system evasion, such as the inhibition of tumor antigen presentation through aberrant expression of the antigen processing machinery [39]. HLA-G expression was first described in melanoma [40]. Through binding to its receptors ILT2, ILT4 and KIR2DL4, HLA-G could inhibit the function of immune effector cells and antigen presenting cells (APCs) [41]. It allows tumor cells to achieve a higher invasive, metastatic potential promoting immune surveillance escape and metastasis during the progression of disease. Thus, HLA-G could be an interesting molecular biomarker and a therapeutic target. Some strategies for cancer immunotherapy based on HLA-G are the use of a delivery system based on nanoparticles carrying anticancer drugs to the tumor, the downregulation of HLA-G by interfering RNA, the use of antibodies blocking HLA-G or its receptors and a HLA-G derived peptide that could induce cytotoxic activity against cells expressing HLA-G [42]. Anti-HLA-G blocking antibodies in animal models seem to restore antitumor immunity against tumor cells expressing HLA-G *in vivo* [43]. However, this constitutes a novel immune escape mechanism and therapeutic antibodies are currently in preclinical development [44].

About downregulated differentially expressed genes in PCa, we mainly find different transcripts of TP53 target gene 3 (TP53TG3). TP53TG3 is a TP53-inducible gene that plays a significant role in TP53-mediated signaling pathway, forming complexes with other proteins to be transferred into the nucleus, where it participates in functions such as cell cycle arrest, apoptosis, DNA repair, chromosomal stability, and/or inhibition of angiogenesis [45]. TP53 is known as the guardian of the genome and many studies have demonstrated its importance in several types of cancer [46,47]. Nevertheless, there is not so much investigation about TP53TG3. Mutations causing loss of function in TP53 tumor suppressor protein are frequent in cancer, what consequently affects the decreased expression of its targeted genes. It would explain that TP53TG3 and its transcripts appear as downregulated differentially expressed genes in PCa.

Finally, we find a high number of genes with differential expression when we make a comparison between an aggressive phenotype of PCa and the rest of PCa samples. The major part of the differentially expressed genes are upregulated genes, indicating that upregulation and aggressiveness may be associated. This is reflected in the previous heat map, where higher levels of expression appear in the aggressive PCa sample and the dendrogram establishes a greater distance between this sample and the rest of them (Figure 7). Moreover, these genes could act as biomarkers of aggressiveness and progression of disease. From the resulting 132 differentially expressed genes, we look in previous published articles for those with a high logFC value and a low FDR value, obtaining names such as centromere protein F (CENPF), DLG associated protein 5 (DLGAP5) and

ribonucleotide reductase small subunit M2 (RRM2). However, there is a large list of genes that can be studied in the future to identify their role in PCa aggressiveness.

CENPF encodes a protein associated with the centromere–kinetochore complex and chromosomal segregation during mitosis. There are studies that report the clinical significance of CENPF as a potential biomarker of PCa malignancy [48,49]. In fact, Shahid et al. reported that higher clinical grades of PCa showed increased expression of CENPF [49]. Branched-chain amino acid (BCAA) metabolism is one of the pathways regulated by CENPF that is related with cancer progression.

DLGAP5 is a cell cycle related gene that has been reported to be overexpressed in a subpopulation of CRPC cells, suggesting that it is associated with an aggressive stage [50]. Also, the loss of DLGAP5 sensitizes androgen-dependent cells to docetaxel, a chemotherapeutic agent used in PCa treatment [51].

RRM2 is associated with poor outcomes in multiple tumors because it promotes epithelial-mesenchymal transition and angiogenesis [52]. RRM2 is responsible of dNTP production, leading to increased genomic instability, cancer progression and resistance to treatments. However, it can be silenced by using a small interfering RNA (siRRM2), a specific micro RNA (miR-193b) [53] or a ribonucleotide reductase inhibitor (COH29) [54].

The translational relevance of this study relapse in the need of finding new biomarkers for PCa due to the current problems of diagnostic accuracy and overdiagnosis. Even though PCa is one of the most prevalent tumors in our society, tools for early disease diagnosis or monitoring of the treatment are still not effective enough. Here, we propose some genes that could serve as predictive, diagnostic and prognostic biomarkers, as well as molecular targets. Nevertheless, further studies would be needed to confirm the obtained results. The main limitations this study presents are sample size and available information at time of the study. We have samples from seven individuals within the same family, but we would need a higher number of samples to extrapolate results to the rest of population. In addition, we take as controls individuals that at time of the study did not have the disease. However, we do not know if they will develop cancer throughout their life.

Future challenges for this study include the analysis of single-nucleotide polymorphisms (SNPs) on both RNA and DNA and expression quantitative trait loci (eQTLs). SNPs are substitutions of a single nucleotide that occurs in at least 1% of population. There is a great interest in their study because they are involved in differences in development, aggressiveness, and/or treatment response of certain diseases [55]. On the other hand, eQTLs are genetic variants that explain variation in gene expression levels [56].

Data from this study are relevant because integration of omics data can help to better understand the aggressiveness of PCa; mainly due to its high incidence, relevance and the lack of current biomarkers.

4. Materials and Methods

4.1. Patients

Seven individuals within the same family were analysed in the current study. In this family, there are five individuals affected of PCa, and there is a woman with BCa who does not present mutations in BRCA1 or BRCA2 genes. The available information suggests that it is a case of familial PCa (Figure 8).

Blood samples in EDTA tubes were taken from individuals 1 to 7. All subjects of the study provided a written informed consent to be enrolled in the project, which was previously approved by the Research Ethics Committee of Granada Center (CEI-Granada) code 0166-N-19 following Helsinki ethical declaration.

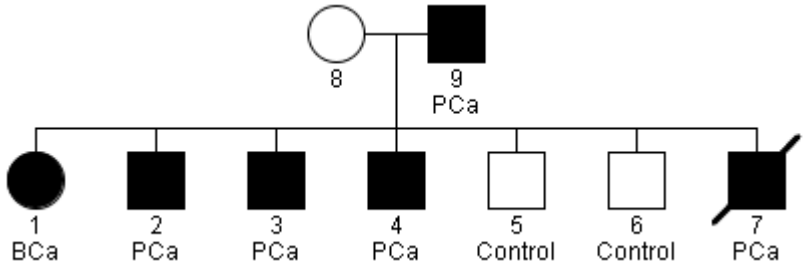


Figure 8. Pedigree of the studied family. Circles indicate women whereas squares indicate men. Solid symbols represent affected people and open symbols represent unaffected people. Individual 7 died during the study. PCa denotes prostate cancer while BCa denotes breast cancer.

4.2. Sample processing

After collecting blood samples, RNA extraction was performed. The process began with the separation of plasma by centrifugation and later RNA was extracted according to QIAamp RNA miRNeasy Serum/Plasma (QIAGEN/QIACUBE) protocol [57].

Once RNA was obtained, next step was the preparation of libraries following the TruSeq® Stranded mRNA Sample Preparation Guide from Illumina (Figure 9). The workflow comprises seven processes:

- Purify and fragment mRNA: it allows to purify the polyA containing mRNA molecules through oligo-dT attached magnetic beads. Then, the mRNA purified is fragmented and primed for cDNA synthesis.

- Synthesize first strand cDNA: this process reverse transcribes the RNA fragments obtained previously primed with random hexamers into first strand cDNA.

- Synthesize second strand cDNA: the second strand cDNA is synthesized using first strand cDNA as a template and incorporating dUTP instead of dTTP. Consequently, RNA is removed.

- Adenylate 3′ ends: one adenine is added to the 3' end of the blunt fragments to prevent them from ligating to one another during next step. Also, one thymine is added to the 3' end of the adapter for a complementary ligation between the adapter and the fragment. This strategy prevent chimera formation.

- Ligate adapters: this process ligates multiple indexing adapters to the ends of the double strand cDNA fragments, preparing them for hybridization onto a flow cell.

- Enrich DNA fragments: there is a selectively enrichment of those DNA fragments that have adapter molecules on both ends by PCR and an amplification of the amount of DNA in the library.

- Normalize and pool libraries: this process prepares DNA templates for cluster generation. Afterwards, the sequencing is performed on a NextSeq 500 System (Illumina).

For further details, see the Preparation Guide from Illumina [58].

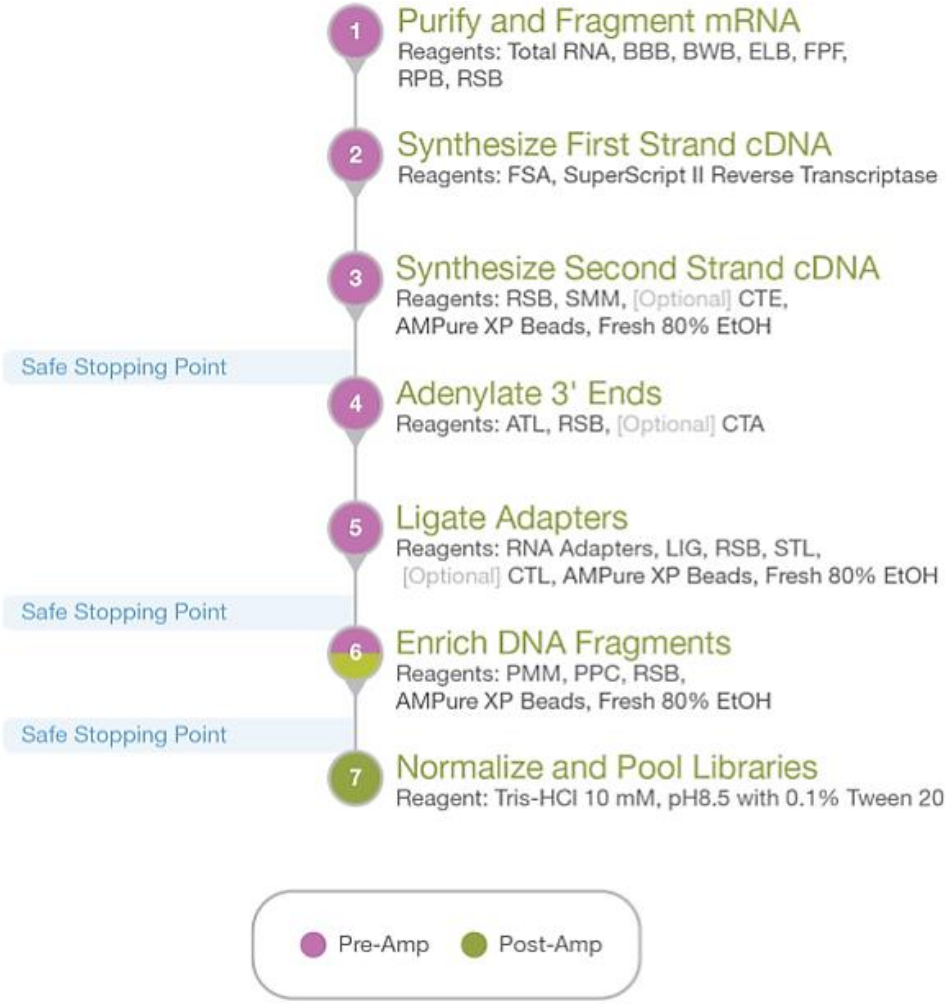


Figure 9. Workflow of the library preparation following TruSeq® Stranded mRNA Sample Preparation Guide from Illumina.

4.3. Data analysis

The bioinformatics analysis starts with a study of the sequencing quality. This is important to avoid possible mistakes, such as adapters accumulation or poor quality, among others. For that purpose, FastQC software was used [59]. In order to have a complete coverage of the sequence, it is necessary to add one nucleotide more in the sequencing process. So that, this nucleotide needs to be removed later using the SeqTk toolkit [60]. Hisat2 [61] allows the alignment of the sequences to human genome (GRCh38/hg38). The output format of the files after alignment is SAM, therefore, we convert them to BAM format using Samtools [62]. Also, with Samtools it is possible to obtain the percentage of properly paired reads for each one of the samples (Table 2).

Once the sequences are aligned to the reference genome, quantification is performed using featureCounts [63]. This function has different parameters we can specify to optimize the percentage of successfully assigned fragments of each sample (Table 3). Some of the parameters we take into account are those for avoiding the count of chimeric fragments, counting multi-mapping reads and assigning a read to the feature that has the largest number of overlapping bases.

Table 2. Percentage of properly paired reads of the seven samples.

Number of sample	Properly paired reads
1	88.41%
2	93.13%
3	93.91%
4	94.05%
5	94.41%
6	94.86%
7	94.59%

Table 3. Percentage of successfully assigned fragments of the seven samples.

Number of sample	Properly paired reads
1	83.30%
2	81.90%
3	81.00%
4	81.20%
5	83.20%
6	83.10%
7	82.00%

Finally, statistical analysis was performed using the edgeR package from Bioconductor [64]. EdgeR enables to examine differential expression of replicated count data using exact tests based on the negative binomial distribution. The result consists of a list with differentially expressed genes and their corresponding logFC (logarithm to the base two of the difference between expression means of a specific gene in both comparison groups) and FDR (adjusted P-values through Benjamini-Hochberg method to control the false discovery rate in multiple testing) values. Level of significance was established at 0.05. Additionally, this report includes density plots of the raw and filtered data, boxplots of the normalized and unnormalized data, MA plots, a MDS plot, a Venn diagram, a dendrogram and a heat map. Workflow of the entire data analysis is shown in Figure 10.

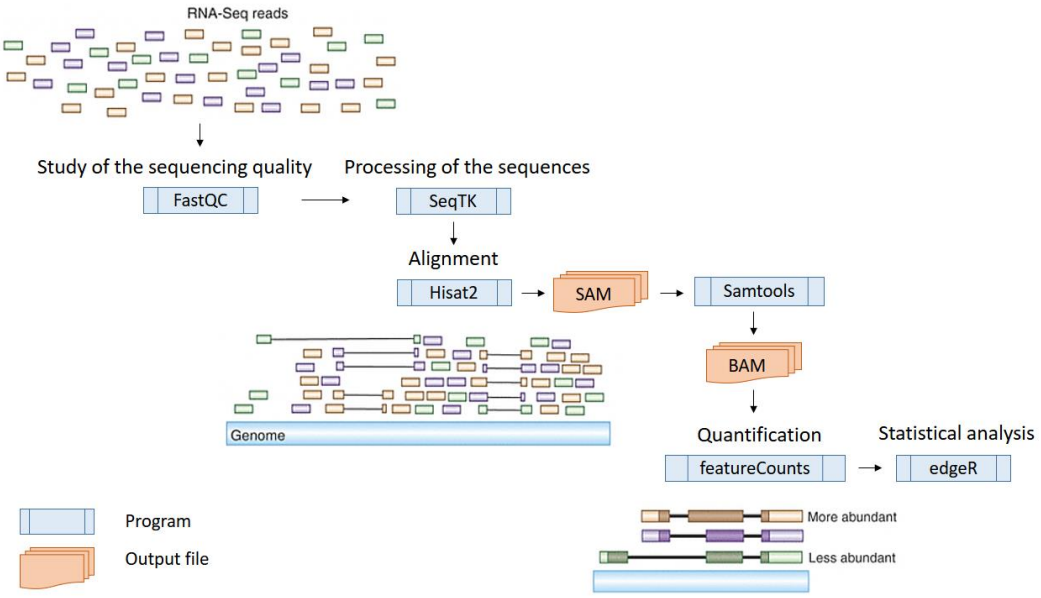


Figure 10. Diagram representing the workflow of the data analysis. Blue represents programs used for the analysis and orange represents output files.

Funding: This research received no external funding.

Acknowledgments: We would like to extend our gratitude to all the patients who participated in the study. Also, we would like to thank the Genomic Unit of GENYO for making the sample processing and RNA-Seq analysis possible. Special thanks to María Jesús and Eduardo for their support, help and patience.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

4Kscore	Four-Kallikrein Panel
ADT	Androgen Deprivation Therapy
APCs	Antigen Presenting Cells
AR	Androgen Receptor
AR-V7	Androgen Receptor Splice Variant-7
BCa	Breast Cancer
BCAA	Branched-Chain Amino Acid
BRCA1	Breast Cancer Type 1
BRCA2	Breast Cancer Type 2
cDNA	Complementary DNA
CENPF	Centromere Protein F
CPM	Counts Per Million
CRPC	Castration Resistant Prostate Cancer
CTCs	Circulating Tumor Cells
DLGAP5	DLG Associated Protein 5
dTTP	Deoxythymidine Triphosphate
dUTP	Deoxyuridine Triphosphate
EDTA	Ethylenediaminetetraacetic Acid
eQTLs	Expression Quantitative Trait Loci
FDA	Food and Drug Administration
FDR	False Discovery Rate
GSTs	Glutathione S-Transferases
HBOC	Hereditary Breast and Ovarian Cancer
HLA	Human Leukocyte Antigen
HLA-G	Human Leukocyte Antigen-G
HP	Haptoglobin
logFC	Log Fold Change
MDS	Multidimensional Scaling
miRNAs	MicroRNAs
mRNA	Messenger RNA
NGS	Next-Generation Sequencing
PCa	Prostate Cancer
PCA3	Prostate Cancer Antigen 3
PCR	Polymerase Chain Reaction
PHI	Prostate Health Index
PSA	Prostate Specific Antigen
PTEN	Phosphatase and Tensin Homolog
RNA-Seq	RNA Sequencing
RRM2	Ribonucleotide Reductase Small Subunit M2
SNPs	Single-Nucleotide Polymorphisms
TMM	Trimmed Mean of M-values
TMPRSS2:ERG	Transmembrane Protease Serine 2:ETS-Related Gene
TP53TG3	TP53 Target Gene 3

Appendix A

Table A1. Information about the seven samples included in the study. This table contains their name, as well as their group, library size and normalization factor. Group 1 correspond to controls, group 2 correspond to BCa and group 3 correspond to PCa.

Sample	File	Group	Library size	Normalization factor
Sample 1	sample1.rcounts	2	16238583	1.0334897
Sample 2	sample2.rcounts	3	17659773	0.9736855
Sample 3	sample3.rcounts	3	15669713	0.9506980
Sample 4	sample4.rcounts	3	17665865	1.1730208
Sample 5	sample5.rcounts	1	17232820	1.0233840
Sample 6	sample6.rcounts	1	16771680	0.8024918
Sample 7	sample7.rcounts	3	17260036	1.0850446

Table A2. Differentially expressed genes in BCa. Upregulated genes are those with a positive logFC value, while downregulated genes have a negative logFC value. FDR is also shown in the table. A FDR value of 0.05 was established as threshold for the selection of genes with differential expression.

Gene name	logFC	FDR
KDM5D	-14.134162	5.23E-43
TXLNGY	-13.924992	5.38E-39
USP9Y	-13.274073	1.82E-32
UTY	-12.135202	4.02E-27
DDX3Y	-11.124095	4.13E-26
RPS4Y1	-11.637882	1.51E-25
PRKY	-6.389185	1.39E-23
EIF1AY	-11.658487	1.37E-22
ZFY	-10.798070	8.47E-20
Xist_exon4	10.798606	3.18E-12
XIST_intron	10.635697	2.32E-11
BCORP1	-11.530795	8.53E-11
RP11-424G14.1	-9.577773	1.82E-06
TMSB4Y	-11.759517	5.73E-06
HP	2.987812	9.08E-06
TTTY15	-7.148742	4.72E-05
RP11-88G17.6	8.907718	0.000338
ALDH1A1	2.245514	0.007066
RNU6-941P	-5.084648	0.009597
IGHV3-30	2.113688	0.012709
AC126755.2	-4.458948	0.012709
AC244250.4	5.245980	0.017916
AC005578.3	2.937979	0.044212
KALP	-7.528456	0.044212
HLA-G	2.486229	0.044212

Table A3. Differentially expressed genes in aggressive PCa. There are 118 upregulated genes and 14 downregulated genes, obtaining 132 genes with differential expression. A FDR value of 0.05 was established as threshold for the selection of those genes.

Upregulated genes				
FAM118A	RPS26P15	ANAPC1P1	CDCA5	AC068522.4
CENPF	RP11-116P24.2	HLA-DQB1-AS1	ARHGAP11A	CCNB1
GS1-184P14.2	RP11-403B2.6	ASPM	VANGL1	FCRL6
DLGAP5	RPS26	RPS26P6	TOX2	HJURP
DTL	CDCA7	CD38	TRIM15	CD8BP
RP11-551L14.1	CDC45	TYMSOS	KIF19	TRAJ36
AC004057.1	GBP5	TK1	GINS2	ORC1
RRM2	CDC20	RP11-275E15.2	MELK	CD8A
TRIM10	RPS26P31	TRBV12-3	IDH2	ERICH3
TYMS	CA1	RP11-330L19.1	KIF4A	CXCL9
MKI67	RPS26P11	UHRF1	HEATR9	PLK4
TPX2	CCNB2	PCNA	AFAP1L2	FRMPD3
MCM4	PITX1	CLIC3	RP11-222K16.2	SKA3
KIF11	CDC6	MCM10	ETV7	RP11-22B23.1
BIRC5	KIF18B	KIF15	E2F8	CDCA2
POLQ	MCM2	RPS26P8	CH507-513H4.1	FAM225A
RPS26P3	TOP2A	NUSAP1	TICRR	SBK1
CENPE	MYBL2	GZMA	RP11-277P12.20	WDR63
RP4-620F22.2	DEPDC1B	SKA1	SH2D1A	CENPM
CLSPN	ZWINT	CDT1	GTSE1	OSBP2
LAG3	KIAA0101	HMMR	CD8B	TRBV12-4
RPS26P47	SPATA3-AS1	MCM6	SMC4	GFI1
CEP55	STMN1	KIFC1	GBP1P1	
TRAV21	EZH2	CDK1	SYNM	
Downregulated genes				
HLA-U	HIST1H1E	AKR1C1	MEG3	IL1R1
AKR1C3	C2orf27A	GPR183	CA4	CNTNAP3
LYPD2	RHD	COL18A1	CMTM2	

References

1. GLOBOCAN 2018: Estimated number of incident cases and deaths worldwide, both sexes, all ages. Available online: http://gco.iarc.fr/ (accessed on Jul 10, 2019).
2. American Cancer Society - Prostate Cancer. Available online: https://www.cancer.org/cancer/prostate-cancer.html (accessed on Jul 10, 2019).
3. Potter, S.R.; Partin, A.W. Hereditary and Familial Prostate Cancer: Biologic Aggressiveness and Recurrence. *Rev Urol*. **2000**, *2*, 35–36.
4. Alvarez-Cubero, M.J.; Saiz, M.; Martinez-Gonzalez, L.J.; Alvarez, J.C.; Lorente, J.A.; Cozar, J.M. Genetic analysis of the principal genes related to prostate cancer: A review. *Urol Oncol*. **2013**, *31*, 1419–1429. DOI: 10.1016/j.urolonc.2012.07.011.
5. Siegel, R.L.; Miller, K.D.; Jemal, A. Cancer statistics, 2019. *CA Cancer J Clin*. **2019**, *69*, 7–34. DOI: 10.3322/caac.21551.
6. LeBlanc, A.G.; Demers, A.; Shaw, A. Recent trends in prostate cancer in Canada. *Health Rep*. **2019**, *30*, 12–17. DOI: 10.25318/82-003-x201900400002-eng.
7. Moyer, V.A.; U.S. Preventive Services Task Force. Screening for Prostate Cancer: U.S. Preventive Services Task Force Recommendation Statement. *Ann Intern Med*. **2012**, *157*, 120–134. DOI: 10.7326/0003-4819-157-2-201207170-00459.
8. Buyyounouski, M.K.; Choyke, P.L.; McKenney, J.K.; Sartor, O.; Sandler, H.M.; Amin, M.B.; Kattan, M.W.; Lin, D.W. Prostate Cancer - Major Changes in the American Joint Committee on Cancer Eighth Edition Cancer Staging Manual. *CA Cancer J Clin*. **2017**, *67*, 245–253. DOI: 10.3322/caac.21391.
9. Hernández, J.; Thompson, I.M. Prostate-Specific Antigen: A Review of the Validation of the Most Commonly Used Cancer Biomarker. *Cancer*. **2004**, *101*, 894–904. DOI: 10.1002/cncr.20480.
10. FDA-NIH Biomarker Working Group. *BEST (Biomarkers, EndpointS, and other Tools) Resource* [Internet]. Silver Spring (MD): Food and Drug Administration (US); Bethesda (MD): National Institutes of Health (US). 2016 [Updated 2017]. Available from: www.ncbi.nlm.nih.gov/books/NBK326791/
11. Alford, A.V.; Brito, J.M.; Yadav, K.K.; Yadav, S.S.; Tewari, A.K.; Renzulli, J. The Use of Biomarkers in Prostate Cancer Screening and Treatment. *Rev Urol*. **2017**, *19*, 221–234. DOI: 10.3909/riu0772.
12. Filella, X.; Giménez, N. Evaluation of [-2] proPSA and Prostate Health Index (phi) for the detection of prostate cancer: A systematic review and meta-analysis. *Clin Chem Lab Med*. **2013**, *51*, 729–739. DOI: 10.1515/cclm-2012-0410.
13. Gaudreau, P.O.; Stagg, J.; Soulières, D.; Saad, F. The Present and Future of Biomarkers in Prostate Cancer: Proteomics, Genomics, and Immunology Advancements. *Biomark Cancer*. **2016**, *8*, 15–33. DOI: 10.4137/BIC.S31802.
14. Sharma, P.; Zargar-Shoshtari, K.; Pow-Sang, J.M. Biomarkers for prostate cancer: present challenges and future opportunities. *Futur Sci OA*. **2016**, *2*, FSO72. DOI: 10.4155/fso.15.72.
15. Salagierski, M.; Schalken, J.A. Molecular Diagnosis of Prostate Cancer: PCA3 and TMPRSS2:ERG Gene Fusion. *J Urol*. **2012**, *187*, 795–801. DOI: 10.1016/j.juro.2011.10.133.
16. Liu, W.; Yin, B.; Wang, X.; Yu, P.; Duan, X.; Liu, C.; Wang, B.; Tao, Z. Circulating tumor cells in prostate cancer: Precision diagnosis and therapy (Review). *Oncol Lett*. **2017**, *14*, 1223–1232. DOI: 10.3892/ol.2017.6332.
17. Khanmi, K.; Ignacimuthu, S.; Paulraj, M.G. MicroRNA in prostate cancer. *Clin Chim Acta*. **2015**, *451*, 154–160. DOI: 10.1016/j.cca.2015.09.022.
18. Cancer Genome Atlas Research Network. The molecular taxonomy of primary prostate cancer. *Cell*. **2015**, *163*, 1011–1025. DOI: 10.1016/j.cell.2015.10.025.
19. Filella, X.; Fernández-Galan, E.; Bonifacio, R.F.; Foj, L. Emerging biomarkers in the diagnosis of prostate cancer. *Pharmgenomics Pers Med*. **2018**, *11*, 83–94. DOI: 10.2147/PGPM.S136026.
20. Alukal, J.P.; Lepor, H. Testosterone Deficiency and the Prostate. *Urol Clin North Am*. **2016**, *43*, 203–208. DOI: 10.1016/j.ucl.2016.01.013.
21. Harris, W.P.; Mostaghel, E.A.; Nelson, P.S.; Montgomery, B. Androgen deprivation therapy: progress in understanding mechanisms of resistance and optimizing androgen depletion. *Nat Clin Pract Urol*. **2009**, *6*, 76–85. DOI: 10.1038/ncpuro1296.
22. Gomella, L.G.; Petrylak, D.P.; Shayegan, B. Current management of advanced and castration resistant prostate cancer. *Cancer J Urol*. **2014**, *21*, 1–6.

23. Saad, F.; Shore, N.; Zhang, T.; Sharma, S.; Cho, H.K.; Jacobs, I.A. Emerging therapeutic targets for patients with advanced prostate cancer. *Cancer Treat Rev*. **2019**, *76*, 1–9. DOI: 10.1016/j.ctrv.2019.03.002.
24. Pritchard, C.C.; Mateo, J.; Walsh, M.F.; De Sarkar, N.; Abida, W.; Beltran, H.; Garofalo, A.; Gulati, R.; Carreira, S.; Eeles, R.; Elemento, O.; Rubin, M.A.; et al. Inherited DNA-Repair Gene Mutations in Men with Metastatic Prostate Cancer. *N Engl J Med*. **2016**, *375*, 443–453. DOI: 10.1056/NEJMoa1603144.
25. Robles-Fernandez, I.; Martinez-Gonzalez, L.J.; Pascual-Geler, M.; Cozar, J.M.; Puche-Sanz, I.; Serrano, M.J.; Lorente, J.A.; Alvarez-Cubero, M.J. Association between polymorphisms in sex hormones synthesis and metabolism and prostate cancer aggressiveness. *PLoS One*. **2017**, *12*, e0185447. DOI: 10.1371/journal.pone.0185447.
26. Gómez-Martín, A.; Martinez-Gonzalez, L.J.; Puche-Sanz, I.; Cozar, J.M.; Lorente, J.A.; Hernández, A.F.; Alvarez-Cubero, M.J. GSTM1 gene expression and copy number variation in prostate cancer patients—Effect of chemical exposures and physical activity. *Urol Oncol*. **2019**, *37*, 290.e9–290.e15. DOI: 10.1016/j.urolonc.2018.12.010.
27. Martinez-Gonzalez, L.J.; Pascual Geler, M.; Robles Fernandez, I.; Cozar, J.M.; Lorente, J.A.; Alvarez Cubero, M.J. Improving the genetic signature of prostate cancer, the somatic mutations. *Urol Oncol*. **2018**, *36*, 312.e17–312.e23. DOI: 10.1016/j.urolonc.2018.03.012.
28. Puche-Sanz, I.; Alvarez-Cubero, M.J.; Pascual-Geler, M.; Rodríguez-Martínez, A.; Delgado-Rodríguez, M.; García-Puche, J.L.; Expósito, J.; Robles-Fernández, I.; Entrala-Bernal, C.; Lorente, J.A.; Cózar-Olmo, J.M.; Serrano, M.J. A comprehensive study of circulating tumour cells at the moment of prostate cancer diagnosis: biological and clinical implications of EGFR, AR and SNPs. *Oncotarget*. **2017**, *8*, 70472–70480. DOI: 10.18632/oncotarget.19718.
29. Cozar, J.M.; Robles-Fernandez, I.; Martinez-Gonzalez, L.J.; Pascual-Geler, M.; Rodriguez-Martinez, A.; Serrano, M.J.; Lorente, J.A.; Alvarez-Cubero, M.J. Genetic markers a landscape in prostate cancer. *Mutat Res*. **2018**, *775*, 1–10. DOI: 10.1016/j.mrrev.2017.11.004.
30. Mateo, J.; Carreira, S.; Sandhu, S.; Miranda, S.; Mossop, H.; Perez-Lopez, R.; Nava Rodrigues, D.; Robinson, D.; Omlin, A.; Tunariu, N.; Boysen, G.; Porta, N.; et al. DNA-Repair Defects and Olaparib in Metastatic Prostate Cancer. *N Engl J Med*. **2015**, *373*, 1697–1708. DOI: 10.1056/NEJMoa1506859.
31. Wallis, C.J.D.; Nam, R.K. Prostate cancer genetics: a review. *EJIFCC*. **2015**, *26*, 79–91.
32. Robinson, M.D.; Oshlack, A. A scaling normalization method for differential expression analysis of RNA-seq data. *Genome Biol*. **2010**, 11, R25. DOI: 10.1186/gb-2010-11-3-r25.
33. Petrucelli, N.; Daly, M.B.; Pal, T. *BRCA1- and BRCA2-Associated Hereditary Breast and Ovarian Cancer* [Internet]. Seattle (WA): University of Washington. 1998 [Updated 2016]. Available from: https://www.ncbi.nlm.nih.gov/books/NBK1247/
34. Levy-Lahad, E.; Friedman, E. Cancer risks among BRCA1 and BRCA2 mutation carriers. *Br J Cancer*. **2007**, *96*, 11-15. DOI: 10.1038/sj.bjc.6603535.
35. UniProt. Available online: https://www.uniprot.org/ (accessed on Jul 22, 2019).
36. Fujimura, T.; Shinohara, Y.; Tissot, B.; Pang, P.C.; Kurogochi, M.; Saito, S.; Arai, Y.; Sadilek, M.; Murayama, K.; Dell, A.; Nishimura, S.; Hakomori, S.I. Glycosylation status of haptoglobin in sera of patients with prostate cancer vs. benign prostate disease or normal subjects. *Int J Cancer*. **2008**, *122*, 39–49. DOI: 10.1002/ijc.22958.
37. Wulaningsih, W.; Holmberg, L.; Garmo, H.; Malmstrom, H.; Lambe, M.; Hammar, N.; Walldius, G.; Jungner, I.; Van Hemelrijck, M. Prediagnostic serum inflammatory markers in relation to breast cancer risk, severity at diagnosis and survival in breast cancer patients. *Carcinogenesis*. **2015**, *36*, 1121-1128. DOI: 10.1093/carcin/bgv096.
38. Lin, A.; Yan, W.H. Heterogeneity of HLA-G Expression in Cancers: Facing the Challenges. *Front Immunol*. **2018**, *9*, 2164. DOI: 10.3389/fimmu.2018.02164.
39. Drake, C.G.; Jaffee, E.; Pardoll, D.M. Mechanisms of Immune Evasion by Tumors. *Adv Immunol*. **2006**, *90*, 51–81. DOI: 10.1016/S0065-2776(06)90002-9.
40. Paul, P.; Rouas-Freiss, N.; Khalil-Daher, I.; Moreau, P.; Riteau, B.; Le Gal, F.A.; Avril, M.F.; Dausset, J.; Guillet, J.G.; Carosella, E.D. HLA-G expression in melanoma: A way for tumor cells to escape from immunosurveillance. *Proc Natl Acad Sci USA*. **1998**, *95*, 4510–4515. DOI: 10.1073/pnas.95.8.4510.
41. Amiot, L.; Ferrone, S.; Grosse-Wilde, H.; Seliger, B. Biology of HLA-G in cancer: a candidate molecule for therapeutic intervention? *Cell Mol Life Sci*. **2011**, *68*, 417–431. DOI: 10.1007/s00018-010-0583-4.

42. Lin, A.; Yan, W.H. Human Leukocyte Antigen-G (HLA-G) Expression in Cancers: Roles in Immune Evasion, Metastasis and Target for Therapy. *Mol Med*. **2015**, *21*, 782–791. DOI: 10.2119/molmed.2015.00083.
43. Lin, A.; Zhang, X.; Xu, H.H.; Xu, D.P.; Ruan, Y.Y.; Yan, W.H. HLA-G expression is associated with metastasis and poor survival in the Balb/c nu/nu murine tumor model with ovarian cancer. *Int J Cancer*. **2012**, *131*, 150–157. DOI: 10.1002/ijc.26375.
44. Carosella, E.D.; Ploussard, G.; LeMaoult, J.; Desgrandchamps, F. A Systematic Review of Immunotherapy in Urologic Cancer: Evolving Roles for Targeting of CTLA-4, PD-1/PD-L1, and HLA-G. *Eur Urol*. **2015**, *68*, 267–279. DOI: 10.1016/j.eururo.2015.02.032.
45. Ng, C.C.; Koyama, K.; Okamura, S.; Kondoh, H.; Takei, Y.; Nakamura, Y. Isolation and Characterization of a Novel TP53-Inducible Gene, TP53TG3. *Genes Chromosomes Cancer*. **1999**, *26*, 329–335. DOI: 10.1002/(SICI)1098-2264(199912)26:4<329::AID-GCC7>3.0.CO;2-C.
46. Wang, Y.; Pascal, L.E.; Zhong, M.; Ai, J.; Wang, D.; Jing, Y.; Pilch, J.; Song, Q.; Rigatti, L.H.; Graham, L.E.; Nelson, J.B.; Parwani, A.V.; et al. Combined Loss of EAF2 and p53 Induces Prostate Carcinogenesis in Male Mice. *Endocrinology*. **2017**, *158*, 4189–4205. DOI: 10.1210/en.2017-00409.
47. Hientz, K.; Mohr, A.; Bhakta-Guha, D.; Efferth, T. The role of p53 in cancer drug resistance and targeted chemotherapy. *Oncotarget*. **2017**, *8*, 8921–8946. DOI: 10.18632/oncotarget.13475.
48. Göbel, C.; Özden, C.; Schroeder, C.; Hube-Magg, C.; Kluth, M.; Möller-Koop, C.; Neubauer, E.; Hinsch, A.; Jacobsen, F.; Simon, R.; Sauter, G.; Michl, U.; et al. Upregulation of centromere protein F is linked to aggressive prostate cancers. *Cancer Manag Res*. **2018**, *10*, 5491–5504. DOI: 10.2147/CMAR.S165630.
49. Shahid, M.; Kim, M.; Lee, M.Y.; Yeon, A.; You, S.; Kim, H.L.; Kim, J. Downregulation of CENPF Remodels Prostate Cancer Cells and Alters Cellular Metabolism. *Proteomics*. **2019**, *19*, e1900038. DOI: 10.1002/pmic.201900038.
50. Horning, A.M.; Wang, Y.; Lin, C.K.; Louie, A.D.; Jadhav, R.R.; Hung, C.N.; Wang, C.M.; Lin, C.L.; Kirma, N.B.; Liss, M.A.; Kumar, A.P.; Sun, L.; et al. Single-Cell RNA-seq Reveals a Subpopulation of Prostate Cancer Cells with Enhanced Cell-Cycle-Related Transcription and Attenuated Androgen Response. *Cancer Res*. **2018**, *78*, 853–864. DOI: 10.1158/0008-5472.CAN-17-1924.
51. Hewit, K.; Sandilands, E.; Martinez, R.S.; James, D.; Leung, H.Y.; Bryant, D.M.; Shanks, E.; Markert, E.K. A functional genomics screen reveals a strong synergistic effect between docetaxel and the mitotic gene DLGAP5 that is mediated by the androgen receptor. *Cell Death Dis*. **2018**, *9*, 1069. DOI: 10.1038/s41419-018-1115-7.
52. Huang, Y.; Liu, X.; Wang, Y.H.; Yeh, S.D.; Chen, C.L.; Nelson, R.A.; Chu, P.; Wilson, T.; Yen, Y. The prognostic value of ribonucleotide reductase small subunit M2 in predicting recurrence for prostate cancers. *Urol Oncol*. **2014**, *32*, 51.e9–51.e19. DOI: 10.1016/j.urolonc.2013.08.002.
53. Mazzu, Y.Z.; Yoshikawa, Y.; Nandakumar, S.; Chakraborty, G.; Armenia, J.; Jehane, L.E.; Lee, G.M.; Kantoff, P.W. Methylation-associated miR-193b silencing activates master drivers of aggressive prostate cancer. *Mol Oncol*. **2019**. DOI: 10.1002/1878-0261.12536.
54. Mazzu, Y.Z.; Armenia, J.; Chakraborty, G.; Yoshikawa, Y.; Coggins, S.A.; Nandakumar, S.; Gerke, T.A.; Pomerantz, M.M.; Qiu, X.; Zhao, H.; Atiq, M.; Khan, N.; et al. A Novel Mechanism Driving Poor-Prognosis Prostate Cancer: Overexpression of the DNA Repair Gene, Ribonucleotide Reductase Small Subunit M2 (RRM2). *Clin Cancer Res*. **2019**, *25*, 4480–4492. DOI: 10.1158/1078-0432.CCR-18-4046.
55. Asthana, S.; Sunyaev, S. Understanding the Functional Importance of Human Single Nucleotide Polymorphisms. In *Discovering Biomolecular Mechanisms with Computational Biology*; Springer US: Boston, MA, 2006; pp. 126–132. DOI: 10.1007/0-387-36747-0_11.
56. Nica, A.C.; Dermitzakis, E.T. Expression quantitative trait loci: present and future. *Philos Trans R Soc Lond B Biol Sci*. **2013**, *368*, 20120362. DOI: 10.1098/rstb.2012.0362.
57. QIAGEN miRNeasy Serum/Plasma Handbook. 2012. Available online: www.qiagen.com/HB-1002 (accessed on Jul 22, 2019).
58. Illumina TruSeq® Stranded mRNA Reference Guide. 2017. Available online: https://www.illumina.com/techniques/sequencing/rna-sequencing/mrna-seq.html (accessed on Jul 22, 2019).
59. Babraham Bioinformatics – FastQC: A Quality Control tool for High Throughput Sequence Data. Available online: http://www.bioinformatics.babraham.ac.uk/projects/fastqc/ (accessed on Jul 22, 2019).
60. Github – Seqtk. Available online: https://github.com/lh3/seqtk (accessed on Jul 22, 2019).

61. Kim, D.; Langmead, B.; Salzberg, S.L. HISAT: a fast spliced aligner with low memory requirements. *Nat Methods*. **2015**, *12*, 357–360. DOI: 10.1038/nmeth.3317.
62. Li, H.; Handsaker, B.; Wysoker, A.; Fennell, T.; Ruan, J.; Homer, N.; Marth, G.; Abecasis, G.; Durbin, R.; 1000 Genome Project Data Processing Subgroup. The Sequence Alignment/Map format and SAMtools. *Bioinformatics*. **2009**, *25*, 2078–2079. DOI: 10.1093/bioinformatics/btp352.
63. Liao, Y.; Smyth, G.K.; Shi, W. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics*. **2014**, *30*, 923–930. DOI: 10.1093/bioinformatics/btt656.
64. Robinson, M.D.; McCarthy, D.J.; Smyth, G.K. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics*. **2010**, *26*, 139–140. DOI: 10.1093/bioinformatics/btp616.