



A Pilot Study on the Whole Exome Sequencing of Prostate Cancer in the Indian Phenotype Reveals Distinct Polymorphisms

Ayam Gupta^{1,2†}, Nidhi Shukla^{1,3†}, Mamta Nehra¹, Sonal Gupta¹, Babita Malik³, Ashwani Kumar Mishra⁴, Maneesh Vijay⁵, Jyotsna Batra⁶, Nirmal Kumar Lohiya⁷, Devendra Sharma⁵ and Prashanth Suravajhala^{1*}

¹ Department of Biotechnology and Bioinformatics, Birla Institute of Scientific Research, Jaipur, India, ² Vignans Foundation for Science, Technology & Research (Deemed to be University), Guntur, India, ³ Department of Chemistry, School of Basic Sciences, Manipal University Jaipur, Jaipur, India, ⁴ DNA Xperts Private Limited, Noida, India, ⁵ Rukmani Birla Hospitals, Jaipur, India, ⁶ Australian Prostate Cancer Research Centre, Queensland Institute of Health and Biomedical Innovation and School of Biomedical Science, Queensland University of Technology, Brisbane, QLD, Australia, ⁷ Department of Zoology, University of Rajasthan, Jaipur, India

OPEN ACCESS

Edited by:

Marika Kaakinen,
University of Surrey, United Kingdom

Reviewed by:

Jyoti Sharma,
Institute of Bioinformatics (IOB), India
Pritish Kumar Varadwaj,
Indian Institute of Information
Technology Allahabad, India

*Correspondence:

Prashanth Suravajhala
prash@bisr.res.in

† These authors have contributed
equally to this work and share first
authorship

Specialty section:

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

Received: 20 March 2020

Accepted: 16 July 2020

Published: 25 August 2020

Citation:

Gupta A, Shukla N, Nehra M,
Gupta S, Malik B, Mishra AK, Vijay M,
Batra J, Lohiya NK, Sharma D and
Suravajhala P (2020) A Pilot Study on
the Whole Exome Sequencing
of Prostate Cancer in the Indian
Phenotype Reveals Distinct
Polymorphisms.
Front. Genet. 11:874.
doi: 10.3389/fgene.2020.00874

Prostate cancer (PCa) is the third most common cancer among men in India, and no next-generation sequencing (NGS) studies have been attempted earlier. Recent advances in NGS have heralded the discovery of biomarkers from Caucasian/European and Chinese ancestry, but not much is known about the Indian phenotype/variant of PCa. In a pilot study using the whole exome sequencing of benign/PCa patients, we identified characteristic mutations specific to the Indian sub-population. We observed a large number of mutations in DNA repair genes, viz. helicases, TP53, and BRCA besides the variants of unknown significance with a possibly damaging rare variant (rs730881069/chr19:55154172C/TR136Q) in the TNNI3 gene that has been previously reported as a semi-conservative amino acid substitution. Our pilot study attempts to bring an understanding of PCa prognosis and recurrence for the Indian phenotype.

Keywords: prostate cancer, genomics, exome sequencing, prognosis, biomarkers

INTRODUCTION

Prostate cancer (PCa) is the second most prevalent cancer worldwide and the third most prevalent cancer in India (Jain et al., 2014). Over 1 million PCa cases are diagnosed per year globally, and the mortality rate has grown to more than 300,000 deaths per year (Farhat et al., 2000). The Population Based Cancer Registries (PBCRs) of different cities for the time period shows that PCa has ranked among the top leading sites of cancer in many cities in India (NCRP, 2013). Prostate cancer leads in large Indian cities like Delhi, Kolkata, Pune, and Thiruvananthapuram after oral cancer, and is the third leading site of cancer in cities like Bangalore and Mumbai besides being among the top 10 leading sites of cancers in the rest of the PBCRs of India (Jain et al., 2014). The data show that almost all regions of India are equally affected, with the incidences of PCa relatively low in some states like Gujarat (Ahmedabad and Wardha PBCRs) and Madhya Pradesh (Bhopal PBCR), the lowest being the northeast regions of India (Jain et al., 2014). This could be because of lack of PBCRs in addition to PCa not actively reported in states where awareness about it is lacking (Kimura, 2012).

Although studies on PCa have dealt with the genetics, genomics, and environmental influence in the causality of PCa, no association of genotype and phenotype employing the whole exome sequencing (WES) of PCa has been done to date for the Indian population. As many Indians are vegetarian, a polygenic risk with other diseases such as diabetes could be associated with PCa, which would be interesting to deliberate (Agrawal et al., 2014). On the other hand, studies at the sequence level in India have been limited where the genetic variants have been shown to be associated with PCa but not linked to pathogenesis (Abate-Shen and Shen, 2000; Shen and Abate-Shen, 2010). In the recent past, the next-generation sequencing (NGS) approaches have provided an efficient, rapid, economical, and global biological means to understand millions of sequenced DNA reads besides recognizing ample variants from the samples (Mardis, 2008). It has an increasing potential to carry a wide spectrum of purposes including many research fields, for example, in the molecular diagnostics of genetic diseases, infectious diseases, cancer, and pharmacogenomics (Ten Bosch and Grody, 2008; Guchelaar et al., 2014). On the other hand, the integration of systems biology with genomic data can lead to the profound knowledge of the disease pathway besides allowing us to identify prognostic biomarkers for drug target identification.

In this pilot study, while we report mutations from northwestern Indian patients, we aimed to understand the marked impact of the risk of PCa specific genes on affected individuals versus controls. We claim that this is perhaps the first study of its kind, where WES was employed to screen mutations and infer the genetic evaluation of PCa in India. With the new Genome India Project initiative (GenomeIndia Initiative, 2017), the genetic makeup for such diseases could allow us to understand the pathways associated with PCa native to India. We also deliberate on mutations associated with polygenic diseases and vitamin D deficiency, besides those that are unique to human diet, socioeconomic imbalances, and PCa manifestation when compared to the Caucasian population (Layne et al., 2019).

MATERIALS AND METHODS

Subjects

The samples were collected from Rukmani Birla Hospital (RBH), Jaipur, and all patients were identified as native to north or northwest India. All methods were performed in accordance with relevant guidelines and regulations. Archival pathological FFPE blocks containing specimens were obtained retrospectively after clearance from the institutional ethics committee of RBH, and informed consent wherever applicable was duly taken. Annotated FFPE blocks of five benign and five tumors (malignant) were taken and sent to Xcelris Labs, Ahmedabad, for WES (Table 1). DNA from sample tissues was extracted using the FFPE tissue gDNA isolation kit (Qiagen/56404) for DNA isolation, and quality control (QC) for contamination was done. The baits for exon capture with a mean target coverage depth of 110× per sample were achieved with as many as several targets covered for more than 20×.

TABLE 1 | Samples used for the WES along with their Gleason scores which constituted one high-grade tumor sample, with three intermediate samples and one just on par above grade 6, while five others are with less than six from benign cases.

Sample ID	Condition	Gleason Score (primary + secondary)
Z785	Adenocarcinoma	3 + 4 = 7
Z786	Acinar Adenocarcinoma	3 + 4 = 7
Z789	Acinar Adenocarcinoma	3 + 3 = 6
Z791	Acinar Adenocarcinoma	3 + 4 = 7
Z794	Cribriform Adenocarcinoma	4 + 4 = 8
Z787	Benign Nodular Prostatic Hyperplasia	<6
Z788	Benign Prostatic Hyperplasia	<6
Z790	Benign Nodular Prostatic Hyperplasia	<6
Z792	Benign Nodular Prostatic Hyperplasia	<6
Z793	Benign Nodular Prostatic Hyperplasia	<6

Sample Preparation, Exome Capture, and Sequencing

DNA was isolated from samples using the Qiagen FFPE DNA extraction kit. The quality of genomic DNA was checked on 0.8% agarose gel (loaded 3 μl) for the single intact band. The gel was run at 110 V for 30 min; 1 μl of each sample was used for determining the concentration using the Qubit 2.0 Fluorometer. All the libraries were prepared using the Agilent SureSelect XT target enrichment system by following the manufacturer's instruction. Briefly, 200 ng of each DNA sample was used for fragmentation using the Covaris S2 system. While the fragmented DNA was subjected to end repair, A-tailing, it was followed by adapter ligation which was hybridized to RNA baits that are designed to physically capture specific DNA sequences. The captured DNA is then eluted from the baits, purified using a biotin-based precipitation, and then amplified by PCR to yield an exome library for sequencing. All prepared libraries were checked on the Agilent high-sensitivity (HS) chip on Bioanalyzer 2100 and quantified on fluorometer by the Qubit dsDNA HS Assay kit (Life Technologies). The average size and concentration of each library was calculated from the HS chip and Qubit, respectively (see Table 2).

The amplified library was analyzed on Bioanalyzer 2100 (Agilent Technologies) using a HS DNA chip as per the manufacturer's instructions. After obtaining the Qubit concentration for the library and mean peak size from Bioanalyzer, the profile library was loaded onto the Illumina platform for cluster generation and sequencing. Paired-end (PE) sequencing allowing the template fragments to be sequenced in both forward and reverse directions was ensured with the library molecules bound to complementary adapter oligos on PE flow cells. The adapters were designed to allow the selective cleavage of the forward strands after the re-synthesis of the reverse strand during sequencing. Finally, the copied reverse strand was used to sequence from the opposite end of the fragment, and all the libraries were sequenced with 150 × 2 PE reads, which resulted in 8–10 GB of data per sample.

TABLE 2 | Data statistics of the sequenced samples.

File Name	Total Reads	Total Bases	Size in GB
Z785	34,589,627	4,524,820,258	9,035,044,752
Z785	34,589,627	4,510,224,494	
Z786	36,192,378	4,782,261,247	9,552,530,147
Z786	36,192,378	4,770,268,900	
Z787	29,342,454	3,504,276,014	7,023,683,778
Z787	29,342,454	3,519,407,764	
Z788	32,007,925	4,590,702,434	9,142,339,304
Z788	32,007,925	4,551,636,870	
Z789	32,912,966	4,688,597,419	9,332,868,751
Z789	32,912,966	4,644,271,332	
Z790	44,212,484	6,398,843,741	12,730,582,043
Z790	44,212,484	6,331,738,302	
Z791	33,311,709	4,822,305,019	9,600,461,768
Z791	33,311,709	4,778,156,749	
Z792	36,245,502	5,367,958,348	10,658,814,541
Z792	36,245,502	5,290,856,193	
Z793	37,261,051	5,344,729,573	10,640,702,373
Z793	37,261,051	5,295,972,800	
Z794	33,200,697	4,584,192,242	9,148,425,756
Z794	33,200,697	4,564,233,514	

Quality Control and Variant Calling

The datasets were run through our in-house pipeline described earlier (Meena et al., 2018). The quality assessment was done for all the samples using FastQC (Andrews, 2010) with raw reads checked for quality, GC bias, K-mer quality, and duplication levels. Calling and filtering of variants and indels were performed using Vt (Tan et al., 2015) and Annovar (Wang et al., 2010) before we established the sensitivity and specificity of called variants. Mutations were counted as heterozygous (“het”) using awk/bash one-liners and the two sets of average depth (≥ 5 DP ≤ 20 and DP ≥ 20) were parsed for further downstream annotation (Supplementary Information). Population stratification at a gross-level was deemed not necessary given the small size of the population on a small scale; however, contamination crosscheck both at sequence (fastq and bam levels) were ensured to assess the heterogeneity of each given site of pooled dataset. Variant analysis using Vt predicted the variant types and were filtered by setting criteria for a false discovery rate (FDR) that yielded the genomic variants in affected individuals versus controls. The samples were analyzed for mutations predicted as causal, and the corresponding genes were filtered against GeneMania (Warde-Farley et al., 2010) and checked for the significant association of pathogenesis with the risk of genitourinary/PCa-specific genes. A brief overview of the methodology is described in Figure 1.

Downstream Analyses

Downstream annotation was done using various open source tools. The reported PCa variants were confirmed from NCBI's ClinVar (NCBI ClinVar Database, 2019) using the search term “prostate” with filters “pathogenic” and “single nucleotide” and checked against the generated VCF files. The bcftools was

employed to retrieve the heterozygous mutations with the varied depth as mentioned earlier. A final list of variants was validated using Sanger sequencing. For the latter, we have maintained an average of 400–600 bp of product size with the 200–250 bp flanking from both sides of the location of the SNP mutation. Primers were designed in such a way that secondary structures in the genome were avoided. We adjusted the GC content as 50% to maintain a balanced distribution of GC-rich in CDS regions. The PCR was set with initial denaturation at 95°C for 5 min, followed by 35 cycles of denaturation at 95°C for 30 s, annealing at 60°C for 30 s and an extension at 72°C for 45 s and final extension of 5 min at 72°C.

RESULTS AND DISCUSSION

Characteristics of PCa Relevant to the Indian Phenotype

Ten subjects, including five cases, those including one high Gleason grade sample (8), and four between 6 and 7 grades were subjected to WES. The sequences were aligned to human genome reference (build hg38) using bowtie2 to produce the sequence alignment file. We observed 944,509 annotated variants and upon further normalization, we discovered 897,926 SNPs, indels and a few copy number variants (CNVs) among them. Post downstream analyses, the generated VCF files were filtered for the variants which were heterozygous and 42 SNPs were confirmed that exhibited significant association with PCa in affected subjects (Supplementary Table 1). Among them, 17 SNPs were subjected further to Sanger validation in an independent platform wherein primers were designed with a length of 18–25 nucleotides (Supplementary Table 2). We found that all the SNPs/samples except one unaffected (rs73598374) passed the sequencing quality thresholds from quality checks with a mean read depth of 14.4. The false positives were carefully checked between benign prostate hyperplasia (BPH) and PCa samples even as a number of exome sites, low-coverage sites, and large deletions could not be validated. In total, we observed 671,609 heterozygous variants confirming the above thresholds that were filtered across all samples.

Association of Variants With PCa

The variants segregating from cases were retained, while those from controls were excluded to rule out the false positives. The minor allele frequency cut-off 0.05 intersecting the allele frequency were judiciously checked for low penetrated variants that are associated with a large number of disease susceptibility genes, primarily cancers. Therefore, we setup the variant identification to as low as DP = 5 and further, segregated the variants those that fall in ≥ 5 DP ≤ 20 and DP ≥ 20 (Figure 2A).

High-Impact Variants in High-Grade Tumors

The identification of tumor-specific somatic mutations was possible with our pipeline. Among the inferred mutations, the prominent were BRCA2 mutations (rs276174889, rs80358600,

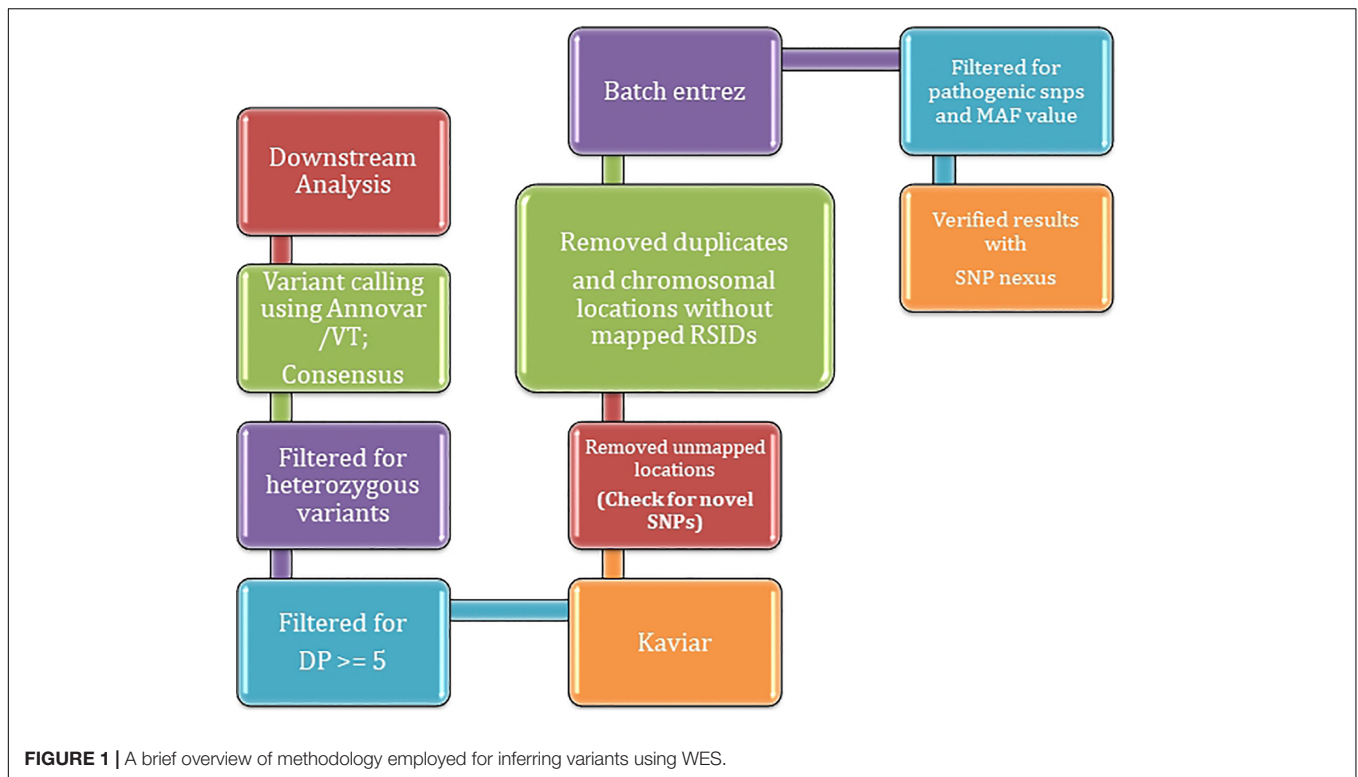


FIGURE 1 | A brief overview of methodology employed for inferring variants using WES.

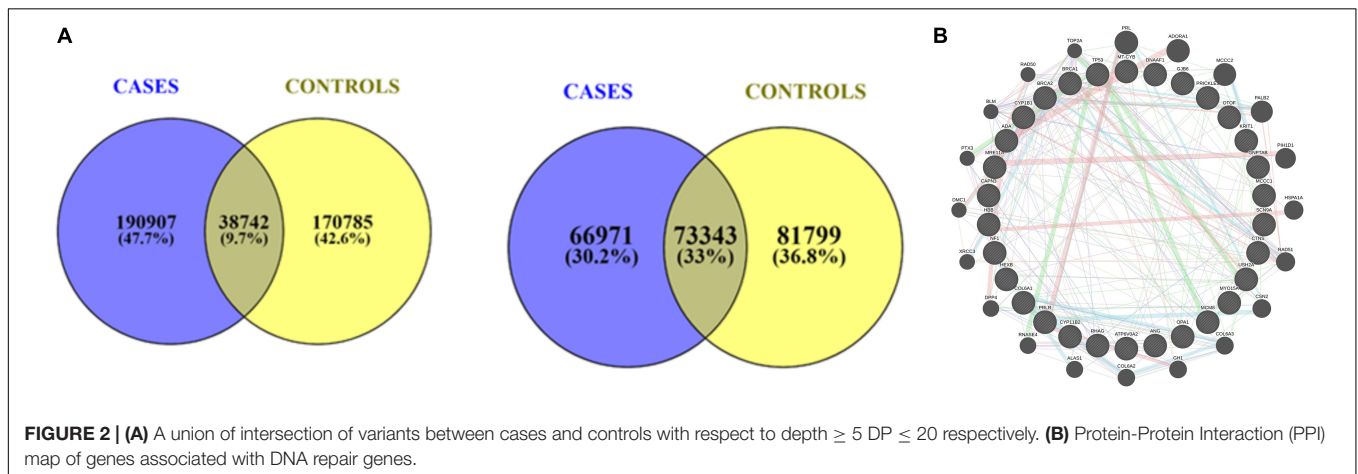


FIGURE 2 | (A) A union of intersection of variants between cases and controls with respect to depth ≥ 5 DP ≤ 20 respectively. (B) Protein-Protein Interaction (PPI) map of genes associated with DNA repair genes.

rs80359171, rs771203198, and rs145988146) associated with high-grade and intermediate tumors (Z794, Z785, and Z786) (Couch et al., 2007). While, BRCA gene mutations lead to development of breast and ovarian cancer, they have been linked to PCa in addition to pancreatic cancer and other myelomas (Paul and Paul, 2014; Cavanagh and Rogers, 2015). The BRCA2 mutations in particular have been associated with an 8.6-fold increased risk for the PCa manifestation (Castro and Eeles, 2012) even as BRCA1 was also seen characteristically in one sample (Z785): rs28897696. Common genetic alterations are usually associated with heterozygous BRCA1 or BRCA2 mutations, and these include loss of the wild-type BRCA1 or BRCA2 allele (LOH), loss of TP53 (which encodes p53),

and ATM or CHK2 function (Roy et al., 2012). These additional alterations may allow cells to bypass checkpoint controls and evade apoptosis, and thereby initiate tumorigenesis. Additionally, tumor complexity makes the detection of cancer-specific CNVs even more difficult (Zare et al., 2017), and the variants observed are frameshift, loss, or gain of stop codons which have significant insertions/deletions at a number of bases in different reading frames with altered protein sequences. While we found them to be significantly associated, we observed that mutations in TP53 and COL6A1 among the other 102 frameshift variants were shown to be affected with spliced regions. Sanger validation (orthogonal) was used to disprove a false-positive variant wherein we performed a

systematic validation of variants of these genes and found no discrepancies among them.

The majority of mutations that are detected in the current study (*viz.*, *GJB6*, *KRIT1*, *GNPTAB*, *ANG*, *MCM8*, and *NF1*) are only significant to high-grade PCa. The SNPs have been widely studied in context to cancer genomes and its predisposition, and so, we believe these SNPs cannot be associated with the familial aspect of the cancer. On the other hand, we observed that a few pathogenic SNPs were absent in high- and median-grade PCa samples and present in the low-grade sample, *viz.*, *BRAF*, *RAD51C*, *RHAG*, *CYP11B2*, *PRICKLE1*, *CAPN3*. With the CNVs considered one of the most important somatic aberrations in cancer since oncogene activation, it is often attributed to chromosomal copy number amplification and tumor suppressor gene inactivation leading to largely heterozygous deletion. With the identification of somatic CNVs known to have an important role in cancer prognosis and treatment, we asked if any of them were significantly associated. On the other hand, we also found 76 CNVs associated with the gain/loss of function that we mapped using SNP nexus (Dayem Ullah et al., 2018) (**Supplementary Table 3**).

As expected, a large number of DNA repair genes (e.g., helicase, TP53, BRCA2) harbor characteristic mutations, which is in agreement with the WES studies. We also found possibly damaging rare variants of unknown significance (VUS) (rs730881069/chr19:55154172: C/T: R136Q) in the *TNNI3* gene that has been previously reported as a semi-conservative amino acid substitution. The mutation, however, is known to impact the PCa cases and seen across multiple hypertrophic cardiomyopathy case reports and, hence, classified as likely pathogenic. From the pathway and interaction network analyses, we inferred that *BRCA2* and *BRCA1* have a majority of interacting partners that are largely associated with either DNA repair genes or helicases, *viz.*, *KRIT1*, *CTNS*, *USH2A*, *HEXB*, *MCCC1*, *MYO15A*, *CAPN3*, *GNPTAB*, *RHAG*, *PRLR*, *OTOE*, *GJB6*, *MCM8*, *CYP11B1*, *PRICKLE1*, *TP53*, *COL6A1*, *DNAAF1*, *SCN9A*, *CYP11B2*, *ATP6V0A2*, *NF1*, *ADA*, *MT-CYB*, *OPA1*, *ANG*, *HBB*, *MRE11* (**Figure 2B**). Notable among them is RecQ/BLM helicase, which has been recently reported to regulate cancer cell proliferation (Qian et al., 2017). In the recent past, recommendations demonstrated the role of prostate serum antigen (PSA) levels correlating the patient's race, age, and prostate volume (Wu et al., 2017). Since the use of PSA is limited and controversial, the search for novel PCa-specific biomarkers, especially from non-invasive bio-fluids, is an important task (Wu et al., 2017). Furthermore, while it is assumed that the impacts of the discovered SNPs on tumor initiation and progression cannot be established due to small sample size, we found that certain genes have a vivid association with the expression or tumor type in general agreement. For example, *CYP11B1* is known for increases in high-grade PCa and correlates with Gleason grades (Chang et al., 2017), P53 and PRLR are shown to correlate with Gleason grades (Jacobson et al., 2011; Martinez-Gonzalez et al., 2018), *MRE11* is known to have an elevated expression associated with progression and poor outcome in PCa (Wang et al., 2019), and *ANG* is up-regulated in the prostate and involved in prostate intraepithelial neoplasia (PIN) formation (Li et al., 2011; Vanli and Guo-Fu, 2015). As no single biomarker is

available for diagnostic and prognostic use (Velonas et al., 2013), the protein-protein interactions complementing our analyses show characteristic associations with a large number of pathways enriched between them (**Figure 2B**). Although only 5.42% of the protein interaction map density constitutes the pathways, *COL6A1*, which is known to be upregulated in castration-resistant prostate cancer (CRPC), remains to be the most critical oncogene regulating the androgen signaling pathways (Zhu et al., 2015). This is also in agreement with the pathway association of its sibling genes *COL6A2* and *COL6A3* (as indicated in light blue edges). Taken together, our results suggest the importance of these pathways and support the potential use of *COL6A* family members in PCa.

In our PCa cohort, TP53 is the highest prevalent mutation which is in agreement with the previous studies reported (Ecke et al., 2010). Apart from well-known and promising cancer genes, our study uncovered several genes with poorly characterized functional roles in cancer which needs further experimental investigation. For example, *KRIT1* is an intracellular protein with ankyrin repeats and a FERM domain that interacts with the Ras-family GTPase Krev1/Rap1a inferring a role in GTPase signaling cascades. Most of the studies linked were its involvement in cardiovascular development but it is known to be ubiquitously expressed in many cells and tissues (Laberge-le Couteux et al., 1999). *USH2A*, a gene responsible for Usher syndrome, has been found to be mutated in our cohort demonstrating its role as tumor suppressor (Van Wijk et al., 2004). Although, some other genes such as *MCC1*, *COL6A1*, and *CYP11B2* are associated with different cancers such as pancreatic, colorectal, and hepatocellular carcinoma, their involvement in PCa is not yet explored (Fukuyama et al., 2008; Fan et al., 2016; Owusu-Ansah et al., 2019). Previous studies, however, have reported the role of diet in PCa (Labbé et al., 2015), so it will be enticing to explore if the genes mentioned above have any role in diet induced PCa. One of the genes, *viz.*, *OPA1*, is mutated in our cohort that is important for maintaining normal mitochondrial morphology and function, with deficiency of *OPA1* checking diet-induced obesity and insulin resistance (Pereira et al., 2018). Likewise, *CYP11B1* prevents diet-induced obesity and glucose intolerance through AMPK activation (Liu et al., 2015). We also observed that *CTNS* (17p13), a gene that causes cystinosis, is associated with multisystem genetic disorder, wherein cysteine is accumulated in different tissues including the adipose and prostate. Strikingly, it is known that vitamin D repletion could be associated with *CTNS* knockout or mutated mice indicating vitamin D as a key-player in the pathogenesis of PCa (Cheung et al., 2019).

CONCLUSION

Prostate cancer is burgeoning in India. Our comprehensive bioinformatics analysis confirmed some characteristic known or unknown mutations from a WES study native to India. Although our study shows characteristic mutations in certain genes, an assay comprising multiple biomarkers that are differentially expressed could be attempted in the future. If this is successful, the number of biomarkers developed will depend on their

validation in a large cohort of patients and the translation of these findings to clinical practice. This pilot study, we believe, is decisive to understand the inherent genes and mutations responsible for PCa in India. Furthermore, an attempt was made to develop a conceptual framework for research particularly in propagating information on the causal genes and mutations responsible for PCa. Although the work was limited to a small number of samples studied, we deem this pilot work would have an impending role in understanding mutations that are of particular interest to Indian genealogy.

DATA AVAILABILITY STATEMENT

The raw reads of WES are deposited in NCBI's sequence read archive (SRA) with accession PRJNA616165.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by CK Birla/Rukmani Birla Hospital. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

AG, MN, and PS analyzed the PCA samples using our NGS pipeline. NS, SG, and BM cross-checked the downstream

analysis. MV and DS provided the clinical samples. AM, JB, and NL helped in the Sequenom analysis. AG wrote the first draft with PS. PS proofread the manuscript before uploading and all others agreed to the changes in the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

The authors declare that this study received seed funding from Brainpan Innovations. The funder was not involved in the study design, collection, analysis, and interpretation of data, the writing of this article or the decision to submit it for publication.

ACKNOWLEDGMENTS

The authors gratefully acknowledge the critical comments of Jayendra Nath Shukla, Praveen Mathur, and Krishna Mohan Medicherla for their kind inputs and suggestions.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fgene.2020.00874/full#supplementary-material>

REFERENCES

- Abate-Shen, C., and Shen, M. M. (2000). Molecular genetics of prostate cancer. *Genes Dev.* 14, 2410–2434. doi: 10.1101/gad.819500
- Agrawal, S., Millett, C. J., Dhillon, P. K., Subramanian, S., and Ebrahim, S. (2014). Type of vegetarian diet, obesity and diabetes in the adult Indian population. *Nutr. J.* 13:89. doi: 10.1186/1475-2891-13-89
- Andrews, S. (2010). FastQC: a quality control tool for high throughput sequence data. *Babraham Bioinform.*
- Castro, E., and Eeles, R. (2012). The role of BRCA1 and BRCA2 in prostate cancer. *Asian J. Androl.* 14, 409–414. doi: 10.1038/aja.2011.150
- Cavanagh, H., and Rogers, K. M. A. (2015). The role of BRCA1 and BRCA2 mutations in prostate, pancreatic and stomach cancers. *Hered. Cancer Clin. Pract.* 13:16. doi: 10.1186/s13053-015-0038-x
- Chang, I., Mitsui, Y., Kim, S. K., Sun, J. S., Jeon, H. S., Kang, J. Y., et al. (2017). Cytochrome P450 1B1 inhibition suppresses tumorigenicity of prostate cancer via caspase-1 activation. *Oncotarget* 8:34190. doi: 10.18632/oncotarget.16598
- Cheung, W. W., Hao, S., and Wang, Z. (2019). Vitamin D repletion ameliorates adipose tissue browning and muscle wasting in infantile nephropathic cystinosis-associated cachexia. *J. Cachexia Sarcopenia Muscle* 11, 120–134. doi: 10.1002/jcsm.12497
- Couch, F. J., Johnson, M. R., Rabe, K. G., Brune, K., De Andrade, M., Goggins, M., et al. (2007). The prevalence of BRCA2 mutations in familial pancreatic cancer. *Cancer Epidemiol. Biomark. Prev.* 16, 342–346. doi: 10.1158/1055-9965.EPI-06-0783
- Dayem Ullah, A. Z., Oscanoa, J., Wang, J., Nagano, A., Lemoine, N. R., and Chelala, C. (2018). SNPnexus: assessing the functional relevance of genetic variation to facilitate the promise of precision medicine. *Nucleic Acids Res.* 46, W109–W113. doi: 10.1093/nar/gky399
- Ecke, T. H., Schlechte, H. H., Schiemenz, K., Sachs, M. D., Lenk, S. V., Rudolph, B. D., et al. (2010). TP53 gene mutations in prostate cancer progression. *Anticancer Res.* 30, 1579–1586.
- Fan, Z., Wang, Z., Chen, W., Cao, Z., and Li, Y. (2016). Association between the CYP11 family and six cancer types. *Oncol. Lett.* 12, 35–40. doi: 10.3892/ol.2016.4567
- Farhat, W. A., Habbal, A. A., and Khauli, R. B. (2000). A guideline to clinical utility of prostate specific antigen. *Saudi Med. J.* 21, 223–227.
- Fukuyama, R., Nicolaita, R., Ng, K. P., Obusez, E., Sanchez, J., Kalady, M., et al. (2008). Mutated in colorectal cancer, a putative tumor suppressor for serrated colorectal cancer, selectively represses β -catenin-dependent transcription. *Oncogene* 27, 6044–6055. doi: 10.1038/nc.2008.204
- Guchelaar, H. J., Gelderblom, H., Van Der Straaten, T., Schellens, J., and Swen, J. (2014). Pharmacogenetics in the cancer clinic: from candidate gene studies to next-generation sequencing. *Clin. Pharmacol. Ther.* 95, 383–385. doi: 10.1038/clpt.2014.13
- GenomeIndia Initiative (2017). *GenomeIndia Initiative Launched to Unlock India's DNA. 1*. Available online at: <https://www.iisc.ac.in/events/genomeindia-initiative-launched-to-unlock-indias-dna/> (accessed August 27, 2019).
- Jacobson, E. M., Hugo, E. R., Borchering, D. C., and Ben-Jonathan, N. (2011). Prolactin in breast and prostate cancer: molecular and genetic perspectives. *Discov. Med.* 11, 315–324.
- Jain, S., Saxena, S., and Kumar, A. (2014). Epidemiology of prostate cancer in India. *Meta Gene* 2, 596–605. doi: 10.1016/j.mgene.2014.07.007
- Kimura, T. (2012). East meets west: ethnic differences in prostate cancer epidemiology between East Asians and Caucasians. *Chin. J. Cancer* 31, 421–429. doi: 10.5732/cjc.011.10324
- Labbé, D. P., Zadra, G., Ebot, E. M., Mucci, L. A., Kantoff, P. W., Loda, M., et al. (2015). Role of diet in prostate cancer: the epigenetic link. *Oncogene* 34, 4683–4691. doi: 10.1038/nc.2014.422
- Laberge-le Couteulx, S., Jung, H. H., Labauge, P., Houtteville, J. P., Lescoat, C., Cecillon, M., et al. (1999). Truncating mutations in CCM1, encoding KRIT1, cause hereditary cavernous angiomas. *Nat. Genet.* 23, 189–193. doi: 10.1038/13815

- Layne, T. M., Graubard, B. I., Ma, X., Mayne, S. T., and Albanes, D. (2019). Prostate cancer risk factors in black and white men in the NIH-AARP diet and health study. *Prostate Cancer Prostatic Dis.* 22, 91–100. doi: 10.1038/s41391-018-0070-79
- Li, S., Ibaragi, S., and Hu, G. (2011). Angiogenin as a molecular target for the treatment of prostate cancer. *Curr. Cancer Ther. Rev.* 7, 83–90. doi: 10.2174/1573394711107020083
- Liu, X., Huang, T., Li, L., Tang, Y., Tian, Y., Wang, S., et al. (2015). CYP1B1 deficiency ameliorates obesity and glucose intolerance induced by high fat diet in adult C57BL/6J mice. *Am. J. Transl. Res.* 7, 761–771.
- Mardis, E. R. (2008). The impact of next-generation sequencing technology on genetics. *Trends Genet.* 24, 133–141. doi: 10.1016/j.tig.2007.12.007
- Martinez-Gonzalez, L. J., Pascual Geler, M., Robles Fernandez, I., Cozar, J. M., Lorente, J. A., and Alvarez Cubero, M. J. (2018). Improving the genetic signature of prostate cancer, the somatic mutations. *Urol. Oncol. Semin. Orig. Investig.* 36, 312.e17–312.e23. doi: 10.1016/j.urolonc.2018.03.012
- Meena, N., Mathur, P., Medicherla, K., and Suravajhala, P. (2018). A bioinformatics pipeline for WHOLE exome sequencing: overview of the processing and steps from raw data to downstream analysis. *Biol. Protoc.* 8:2805. doi: 10.21769/BioProtoc.2805
- NCBI ClinVar Database (2019). *NCBI NCBI ClinVar Database*. Available online at: <https://www.ncbi.nlm.nih.gov/clinvar> (accessed October 25, 2019).
- NCRP (2013). *Three-Year Report of the Population Based Cancer Registries- 2009-2011*. Bethesda, MD: NCRP.
- Owusu-Ansah, K. G., Song, G., Chen, R., Edo, M. I. A., Li, J., Chen, B., et al. (2019). COL6A1 promotes metastasis and predicts poor prognosis in patients with pancreatic cancer. *Int. J. Oncol.* 55, 391–404. doi: 10.3892/ijo.2019.4825
- Paul, A., and Paul, S. (2014). The breast cancer susceptibility genes (BRCA) in breast and ovarian cancers. *Front. Biosci. Landmark* 19, 605–618. doi: 10.2741/4230
- Pereira, R., Olvera, A. C., Marti, A. A., Hewezi, R., Bui Tran, W. A., Potthoff, M. J., et al. (2018). Resistance to diet-induced obesity in mice lacking OPA1 in adipose tissue occurs independently of fat-derived FGF-21 and BAT function. *Diabetes* 67(Supplement 1) 276-LB. doi: 10.2337/db18-276-lb
- Qian, X., Feng, S., Xie, D., Feng, D., Jiang, Y., and Zhang, X. (2017). RecQ helicase BLM regulates prostate cancer cell proliferation and apoptosis. *Oncol. Lett.* 14, 4206–4212. doi: 10.3892/ol.2017.6704
- Roy, R., Chun, J., and Powell, S. N. (2012). BRCA1 and BRCA2: different roles in a common pathway of genome protection. *Nat. Rev. Cancer* 12, 68–78. doi: 10.1038/nrc3181
- Shen, M., and Abate-Shen, C. (2010). Molecular genetics of prostate cancer: new prospects for old challenges. *Genes Dev.* 24, 1967–2000. doi: 10.1101/gad.1965810.GENES
- Tan, A., Abecasis, G. R., and Kang, H. M. (2015). Unified representation of genetic variants. *Bioinformatics* 31, 2202–2204. doi: 10.1093/bioinformatics/btv112
- Ten Bosch, J. R., and Grody, W. W. (2008). Keeping up with the next generation: massively parallel sequencing in clinical diagnostics. *J. Mol. Diagnost.* 10, 484–492. doi: 10.2353/jmoldx.2008.080027
- Van Wijk, E., Pennings, R. J. E., Te Brinke, H., Claassen, A., Yntema, H. G., Hoefsloot, L. H., et al. (2004). Identification of 51 novel exons of the usher syndrome Type 2A (USH2A) gene that encode multiple conserved functional domains and that are mutated in patients with usher syndrome Type II. *Am. J. Hum. Genet.* 74, 738–744. doi: 10.1086/383096
- Vanli, N., and Guo-Fu, H. U. (2015). Mechanism and function of angiogenin in prostate cancer. *Chin. J. Biochem. Mol. Biol.* 31, 1261–1266. doi: 10.13865/j.cnki.cjmb.2015.12.06
- Velonas, V., Woo, H., Remedios, C., and Assinder, S. (2013). Current status of biomarkers for prostate cancer. *Int. J. Mol. Sci.* 14, 11034–11060. doi: 10.3390/ijms140611034
- Wang, J., Xu, W.-H., Wei, Y., Zhu, Y., Qin, X.-J., Zhang, H.-L., et al. (2019). Elevated MRE11 expression associated with progression and poor outcome in prostate cancer. *J. Cancer* 10, 4333–4340. doi: 10.7150/jca.31454
- Wang, K., Li, M., and Hakonarson, H. (2010). ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Res.* 38:e164. doi: 10.1093/nar/gkq603
- Wardle-Farley, D., Donaldson, S. L., Comes, O., Zuberi, K., Badrawi, R., Chao, P., et al. (2010). The GeneMANIA prediction server: biological network integration for gene prioritization and predicting gene function. *Nucleic Acids Res.* 38, W214–W220. doi: 10.1093/nar/gkq537
- Wu, D., Ni, J., Beretov, J., Cozzi, P., Willcox, M., Wasinger, V., et al. (2017). Urinary biomarkers in prostate cancer detection and monitoring progression. *Crit. Rev. Oncol. Hematol.* 118, 15–26. doi: 10.1016/j.critrevonc.2017.08.002
- Zare, F., Dow, M., Monteleone, N., Hosny, A., and Nabavi, S. (2017). An evaluation of copy number variation detection tools for cancer using whole exome sequencing data. *BMC Bioinform.* 18:286. doi: 10.1186/s12859-017-1705-x
- Zhu, Y. P., Wan, F. N., Shen, Y. J., Wang, H. K., Zhang, G. M., and Ye, D. W. (2015). Reactive stroma component COL6A1 is upregulated in castration-resistant prostate cancer and promotes tumor growth. *Oncotarget* 6, 14488–14496. doi: 10.18632/oncotarget.3697

Conflict of Interest: AM was employed by company DNA Xperts Pvt Ltd.

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

Copyright © 2020 Gupta, Shukla, Nehra, Gupta, Malik, Mishra, Vijay, Batra, Lohiya, Sharma and Suravajhala. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.