

RESEARCH ARTICLE

# Mutational Landscapes of Sequential Prostate Metastases and Matched Patient Derived Xenografts during Enzalutamide Therapy

Manish Kohli<sup>1\*</sup>, Ligu Wang<sup>2\*</sup>, Fang Xie<sup>8</sup>, Hugues Sicotte<sup>2</sup>, Ping Yin<sup>1</sup>, Scott M. Dehm<sup>3</sup>, Steven N. Hart<sup>2</sup>, Peter T. Vedell<sup>2</sup>, Poulami Barman<sup>2</sup>, Rui Qin<sup>2</sup>, Douglas W. Mahoney<sup>2</sup>, Rachel E. Carlson<sup>2</sup>, Jeanette E. Eckel-Passow<sup>2</sup>, Thomas D. Atwell<sup>4</sup>, Patrick W. Eiken<sup>4</sup>, Brendan P. McMenemy<sup>4</sup>, Eric D. Wieben<sup>5</sup>, Gautam Jha<sup>6</sup>, Rafael E. Jimenez<sup>7</sup>, Richard Weinshilboum<sup>8</sup>, Liewei Wang<sup>8</sup>



**OPEN ACCESS**

**Citation:** Kohli M, Wang L, Xie F, Sicotte H, Yin P, Dehm SM, et al. (2015) Mutational Landscapes of Sequential Prostate Metastases and Matched Patient Derived Xenografts during Enzalutamide Therapy. PLoS ONE 10(12): e0145176. doi:10.1371/journal.pone.0145176

**Editor:** Natasha Kyprianou, University of Kentucky College of Medicine, UNITED STATES

**Received:** August 10, 2015

**Accepted:** November 30, 2015

**Published:** December 22, 2015

**Copyright:** © 2015 Kohli et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](http://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Data Availability Statement:** RNA-seq data generated from metastatic prostate cancer tissues and derived mouse xenografts were submitted to NCBI Gene Expression Omnibus (<http://www.ncbi.nlm.nih.gov/geo/>) with accession number: GSE70380. Matched genome sequencing data were submitted to NCBI Short Read Archive (<http://www.ncbi.nlm.nih.gov/sra/>) with the accession number: SRP060313.

**Funding:** Funding provided by National Institutes of Health grant R01CA174777 (to S. M. D.), [http://projectreporter.nih.gov/project\\_info\\_details.cfm?aid=](http://projectreporter.nih.gov/project_info_details.cfm?aid=)

**1** Department of Oncology, Mayo Clinic, Rochester, Minnesota, United States of America, **2** Division of Biomedical Statistics and Informatics, Department of Health Sciences, Mayo Clinic, Rochester, Minnesota, United States of America, **3** Masonic Cancer Center and Department of Laboratory Medicine and Pathology, University of Minnesota, Minneapolis, Minnesota, United States of America, **4** Department of Radiology, Mayo Clinic, Rochester, Minnesota, United States of America, **5** Department of Biochemistry and Molecular Biology, Mayo Clinic, Rochester, Minnesota, United States of America, **6** Division of Hematology-Oncology, University of Minnesota, Minneapolis, Minnesota, United States of America, **7** Department of Pathology and Lab Medicine, Mayo Clinic, Rochester, Minnesota, United States of America, **8** Department of Molecular Pharmacology and Experimental Therapeutics, Mayo Clinic, Rochester, Minnesota, United States of America

\* [kohli.manish@mayo.edu](mailto:kohli.manish@mayo.edu) (MK); [wang.ligu@mayo.edu](mailto:wang.ligu@mayo.edu) (Ligu Wang)

## Abstract

Developing patient derived models from individual tumors that capture the biological heterogeneity and mutation landscape in advanced prostate cancer is challenging, but essential for understanding tumor progression and delivery of personalized therapy in metastatic castrate resistant prostate cancer stage. To demonstrate the feasibility of developing patient derived xenograft models in this stage, we present a case study wherein xenografts were derived from cancer metastases in a patient progressing on androgen deprivation therapy and prior to initiating pre-chemotherapy enzalutamide treatment. Tissue biopsies from a metastatic rib lesion were obtained for sequencing before and after initiating enzalutamide treatment over a twelve-week period and also implanted subcutaneously as well as under the renal capsule in immuno-deficient mice. The genome and transcriptome landscapes of xenografts and the original patient tumor tissues were compared by performing whole exome and transcriptome sequencing of the metastatic tumor tissues and the xenografts at both time points. After comparing the somatic mutations, copy number variations, gene fusions and gene expression we found that the patient's genomic and transcriptomic alterations were preserved in the patient derived xenografts with high fidelity. These xenograft models provide an opportunity for predicting efficacy of existing and potentially novel drugs that is based on individual metastatic tumor expression signature and molecular pharmacology for delivery of precision medicine.

8826081&icde=25365178&ddparam=&ddvalue=&ddsub=&cr=1&csb=default&cs=ASC, Minnesota Partnership for Biotechnology and Medical Genomics grant #14.37 (to M. K. and S. M. D.), <http://minnesotapartnership.info/2015/04/minnesota-partnership-awards-six-new-research-grants/> and Private and philanthropy funding sources: i) The study is funded in part by the Mayo Clinic Centre for Individualized Medicine; ii) A. T. Suharya and Ghan D. H.; iii) Joseph and Gail Gassner; iv) Mayo Clinic Schulze Cancer for Novel Therapeutics in Cancer Research. v) Pharmacogenomics Research Network. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

**Competing Interests:** The authors have declared that no competing interests exist.

## Introduction

Intra-tumoral heterogeneity may underlie hitherto unknown adaptive mechanisms of resistance to drug therapy [1]. Recent advances in massively parallel next-generation sequencing have made it possible to study cancer heterogeneity and clonal evolution within cancer [2]. This has led to in-depth elucidation of tumor specific genetic profiles of several cancers including aggressive and lethal prostate cancer [3,4,5]. However, identification of genes that are somatically mutated or have an altered expression profile has yet to yield clinical benefit for patients with advanced prostate cancer. One reason for this could be that functional and therapeutic understanding of these genomic and transcriptomic observations is lacking. Patient derived Xenografts (PDXs) can transform genomic profiles into pharmacogenetic platforms for drug testing [6,7] and offer the potential for enhancing gains from sequencing studies. However, PDX models of prostate cancer have historically been challenging to establish, with a high degree of variability in take-rates reported [8]. Perhaps not surprisingly, the take-rate with metastatic prostate cancer appears to be generally higher than the take-rate with primary prostate cancer [9]. This signifies that PDX models developed from biopsies of advanced prostate cancer may represent feasible platforms for elucidating mechanisms of response to the newest generations of therapies [10,11].

In this study, in a single patient we initially describe and then compare the mutational landscapes of metastatic tissue with PDXs from bone tissue metastasis of a castrate-resistant prostate cancer patient before and after initiating treatment with enzalutamide, a second-generation AR antagonist. We also determine the recapitulation of genomic alterations from serially obtained cancer tissue in animal models.

## Material and Methods

### Clinical Case Presentation

The case presented was part of a prospectively conducted feasibility study performed in February 2013 approved by the Mayo Clinic Institutional Review Board. The goal of the study was to develop PDX models in immune-deficient mouse models from metastases of castration resistant prostate cancer (CRPC) stage patients undergoing standard of care stage specific therapy. PDX models were generated only in one of nine enrolled patients which is presented in this case report. Written informed consent was given by all participants for their clinical records to be used in this study including the patient with successful generation of the PDX models before and after enzalutamide therapy. The patient initially presented with recurrent urinary obstructive symptoms in June 2010. An investigative workup revealed a hard and enlarged prostate and transurethral resection demonstrated an adenocarcinoma with an initial Gleason score (GS) of 9 and a presenting (June 2010) prostate specific antigen (PSA) level of 0.95 ng/dL. He underwent nerve-sparing radical resection of the prostate on October, 2010, with final pathology revealing GS 9; pT3b (bilateral seminal vesicles involvement); N0 (0 of 20 lymph nodes positive for metastases) stage and several positive margins. He received adjuvant radiation combined with luteinizing hormone releasing hormone (LHRH) analogues which were continued for a year. Post therapy his PSA remained undetectable until August 2012 when it increased to 2.9 ng/dL. Staging bone and computed tomography (CT) scan of the abdomen and pelvis were non-revealing for metastasis, but an acetyl choline positron-emission tomography (PET)-CT demonstrated multiple sites of bony metastasis. He was re-started on leuprolide acetate in October 2012, with bicalutamide for hormone sensitive stage and after an initial response progressed to CRPC stage in February, 2013. Prior to initiating therapy for CRPC stage disease, he enrolled on the feasibility study and underwent a research study biopsy of a

metastatic rib lesion for sequencing and developing PDX models. He was treated with enzalutamide in combination with sipuleucel-T and indoximod on a clinical trial and underwent a second research biopsy after 12 weeks of continuous treatment with the above combination. At 12 weeks his serum PSA was observed to be 29.3 ng/ml (serially increased from pre enzalutamide PSA of 17.7 ng/ml after two measurements six weeks apart). His repeat imaging scan compared to baseline demonstrated progressive disease in the bone metastasis on a bone scan resulting in change of therapy to systemic docetaxel chemotherapy for 4.5 months and subsequently to carboplatin based chemotherapy after further progression on docetaxel.

## Patient Derived Xenograft Methods

PDXs were developed from biopsies (10mm height x 1.5mm diameter) of castrate-resistant prostate cancer (CRPC) tissue metastases collected from the patient before and after initiating stage specific standard of care treatment with enzalutamide. Renal capsule xenografting and subcutaneous 25mg testosterone pellet (Jungle Jim, Fairfield, OH) implantation in NOD-SCID mice (Jackson laboratories, Bar Harbor, Maine; Charles River laboratories, Raleigh, North Carolina) were performed as described previously for the evaluation of prostate carcinogenesis and benign prostatic hyperplasia [12]. Expansion of sub-renal xenografts was observed 4 and 6 months after implantation of baseline (visit 1) and visit 2 (after 12-weeks of continuous enzalutamide therapy) biopsy tissue, respectively. PDX tumors from each visit were further expanded in second-generation mice for molecular characterization. (Details of mice anesthesia and euthanasia are included in [S1 Animal Methods](#).) The initial xenograft mice were not biopsied again. The mice were sacrificed when their xenograft tumor diameter achieved 1.0~1.5cm, at which point the whole tumors were harvested. A portion of tumors were re-injected into the second generation mice. The rest of tumors were collected for molecular analyses, pathology confirmation or frozen for future use. Whole exome (WES) and transcriptome sequencing (RNA-seq) were performed on patient tumor tissues and on 10 PDXs (four models from the first generation and six models from the second generation) ([S1 Fig](#)).

## Whole Exome Sequencing (WES) of tumor

Deep sequencing of exome captured DNA was performed on the Illumina HiSeq 2500. Exome capture was performed using a customized version of Agilent's SureSelect Human All Exon V4+UTR capture kit with additional baits added to cover exons and introns of the androgen receptor (AR) gene [13]. DNA samples were captured following the manufacturer's protocol and sequenced as 101 nt pair-end reads in a single lane of the HiSeq 2500.

We used Genome\_GPS, a comprehensive secondary analysis pipeline for high-throughput genome sequencing data developed by Mayo Clinic, to perform sequence alignment, variants calling including single nucleotide variations (SNVs), small insertions and deletions (INDELs), structural variations, and annotation. Unless otherwise specified, all tools built into Genome\_GPS were run under the default configurations. Specifically, raw reads were aligned to the human reference genome (hg19/GRCh37) using Novoalign (VN:V2.07.13) with the following options:—hdrhd off -v 120 -c 4 -i PE 425,80 -x 5 -r Random (<http://www.novocraft.com/>). Realignment and recalibration were performed using GATK [14](VN: 2.7-4-g6f46d11) Best Practices version 3. Germline variations were called with GATK's UnifiedGenotyper and SNVMix2[15]. Variant quality score was recalibrated using the following command-line optimizations: for SNVs, -an QD -an MQRankSum -an ReadPosRankSum -an FS -an DP; and for INDELs, -an DP -an FS -an ReadPosRankSum -an MQRankSum—maxGaussians 4. Somatic SNVs were called using SomaticSniper (VN: 1.0.0.1-19) [16], JointSNVMix2 (0.8b2)[17] with the “-prob 0.1” parameter, or MuTect (1.1.4), and Somatic INDELs were called using

SomaticIndelDetector (GATK) and annotated with SnpEFF (VN: 3.0c). Reported SNVs and INDELS were filtered to exclude any variant found in the normal blood sample, variants outside the exons, variants of low or modifier impact, and any variant observed in the 1000 Genomes project [18] or the Exome Sequencing Project [19].

All somatic copy number variations (CNVs) were called using an updated version of PatternCNV [20], which uses reference samples to learn the pattern and variance of the WES coverage to better enable somatic CNV calling. The median reads per kilobase per million mapped reads (RPKM) coverage and variance of a set of germline samples was computed for each 120bp interval overlapping with the location of the exome capture probes. Probes GC content bias was corrected by grouping the probes according to their GC content and by the number of copies of each probe in the capture pool. The peak density within each group was then adjusted on a per sample basis. The base-2 logarithm ratio ( $\text{Log}_2R$ ) of the coverage between a tumor sample and the reference set was computed for each probe. The  $\text{Log}_2R$  values were segmented using CBS [21] in the DNACopy package, taking the predicted variance into account (adjusted for the coverage of the tumor sample). The  $\text{log}_2R$  was then normalized so that the median of all segmented probe values = 0. After manual review and cross-sample comparison, a few samples were re-centered so that chromosome 2 had the same copy number. A threshold of  $|\text{Log}_2R| > 0.5$  was used to call a region as amplified or deleted. Tumor purity was estimated using PurBayes, with some modifications (1). Specifically, we modified PurBayes to simultaneously identify subclones and estimate purity for each subclone across multiple sequential time points. The criteria to estimate purity included limitation to variants that had at least 50X coverage, did not overlap a called CNV ( $|\text{Log}_2R| \geq 0.5$ ) and a dbSNP variant, was not on chromosome M or chromosome X, and had at most one read carrying the mutant allele in the corresponding germline sample. Lastly, we eliminated variants that could have arisen from reads mapping to multiple locations by taking a sequence region of 101bp region centered on the variant and aligning against the human genome (hg19) using BLAT [22]. Any variant with multiple genome mappings by this approach was not used in the purity analysis.

## RNA sequencing (RNAseq) of tumor specimens

mRNA libraries were prepared using Illumina's TruSeq RNA prep kit and standard protocol. The RNA libraries were sequenced as 101 nt pair-end reads at 2 samples per lane of an Illumina HiSeq 2500, generating an average of 140 million reads per sample.

Mayo Analysis Pipeline for RNA-Seq (MAP-RSeq) is the comprehensive secondary analysis pipeline for RNAseq data at Mayo Clinic [23]. The MAP-RSeq workflow performs a thorough processing of paired-end RNAseq data in three steps which included alignment and parameter optimization; quality control of samples and summary of genomic features per sample or run. These three steps were run using the following options:

- i. Alignment and parameter optimization: Reads were aligned by TopHat 2.0.6 [24] using the hg19 genome build and bowtie1 aligner option [25]. TopHat requires mean and standard deviation of fragment size to perform alignment. Hence, fragment size parameters were estimated by sampling the first 100K read pairs that were uniquely mapped. Most of the tools in MAP-RSeq use the default configurations; however, we used the following settings for TopHat: `--keep-fastq-order, --keep-tmp, --no-coverage-search, --bowtie1, --library-type fr-unstranded, --max-multihits 20, --solexa1.3-quals, --fusion-search, --fusion-ignore-chromosomes chrM, --fusion-min-dist 50000`.
- ii. Quality control of samples: MAP-RSeq also provides detailed quality-control plots per sample to estimate the distance between paired-end reads, evaluate the sequencing depth

for alternate splicing analysis, determine the rate of duplicate reads, and evaluate coverage of reads across genes using RSeQC software (<http://rseqc.sourceforge.net/>) [26].

- iii. Summary of genomic features per sample or run: Gene expression counts were generated using HTseq software (<http://www-huber.embl.de/users/anders/HTSeq/doc/overview.html>) from Illumina gene annotation files ([http://support.illumina.com/sequencing/sequencing\\_software/igenome.html](http://support.illumina.com/sequencing/sequencing_software/igenome.html)). The MAP-RSeq also provides a list of expressed fusion transcripts using the TopHat-Fusion algorithm [27].

## Xenograft sequencing Methods

Sequencing data from patient-derived xenograft (PDX) samples contain a mixture of reads originating from the host (mouse) and from the graft (human). Since the proportion of reads from the different sources can vary significantly across PDXs, classifying reads by source is useful to improve the precision of analysis. For both DNA and RNA, we employed the classification scheme proposed by Conway *et al.* [28], which classifies reads as “graft”, “host”, “both”, “neither” or “ambiguous”. We used the Non-Obese Diabetic (NOD) genome sequences as the host reference genome (<http://cgd.jax.org/tools/Seqnature.shtml>), where NOD is the background strain of the SCID strain, *NOD.CB17-Prkdcscid/J*, used in the engrafting. Using this host reference genome rather than the mouse reference genome (mm10) allowed for less reads to be categorized as “ambiguous” and more being categorized as “host”. For both RNA and DNA, we chose to exclude only those classified as “host” to maximize the amount of graft-originating reads for downstream analysis. Then, in each case, we proceeded with the relevant workflow (Genome\_GPS for WES and MAPR-seq for RNAseq) as described previously.

## Recall and Precision

We use recall (i.e. sensitivity) and precision metrics to measure the relevance between tumor and xenograft genomes:

$$recall = \frac{P \cap X}{P}$$

$$precision = \frac{P \cap X}{X}$$

Where  $P$  denotes the “mutations called from patient tissues” and  $X$  denotes the “mutations called from xenograft”.

Somatic mutations selected for this analysis had to be called in either the patient tumors or xenografts, had to have high or moderate impact under SNPEff, had to not be found in the 1000 genome European or African samples, had to have 0 reads in the germline, and the 100-bp region surrounding the variant was required to map uniquely to the human reference genome ( $\geq 90\%$  identity).

## AR variant 7 and AR full length expression

AR variant 7 (*AR-V7*) and AR full length (*AR-FL*) are the two commonly expressed isoforms in castrate resistant stage. Since *AR-V7* and *AR-FL* share the first 3 exons, expression level of *AR-V7* was therefore measured by the number of splice reads that span exon-3 (chrX:66905852–66905968) and the downstream cryptic exon (chrX:66914515–66915580) located in intron-3. The expression levels of *AR-FL* were measured by the number of splice reads that span exon-4 (chrX:66931244–66931531) and exon-5 (chrX:66937320–66937464).

The number of splice reads were then normalized by total splice reads to correct sequencing depth (i.e. SRPM, splice read per million).

## Quantitative real time RT-PCR validation

RNA samples were extracted from snap-frozen xenograft tumor tissues by miRNeasy mini kit (Qiagen, Redwood city, CA). QRT-PCR primers were designed and synthesized to cover the possible regions of gene fusion breakpoints. The GAPDH primer was purchased from Qiagen. The QRT-PCR was performed using the one-step QRT-PCR kit (Qiagen). The amplification Ct value was normalized to that of GAPDH. The PCR amplified products were visualized on agarose gels.

## Results

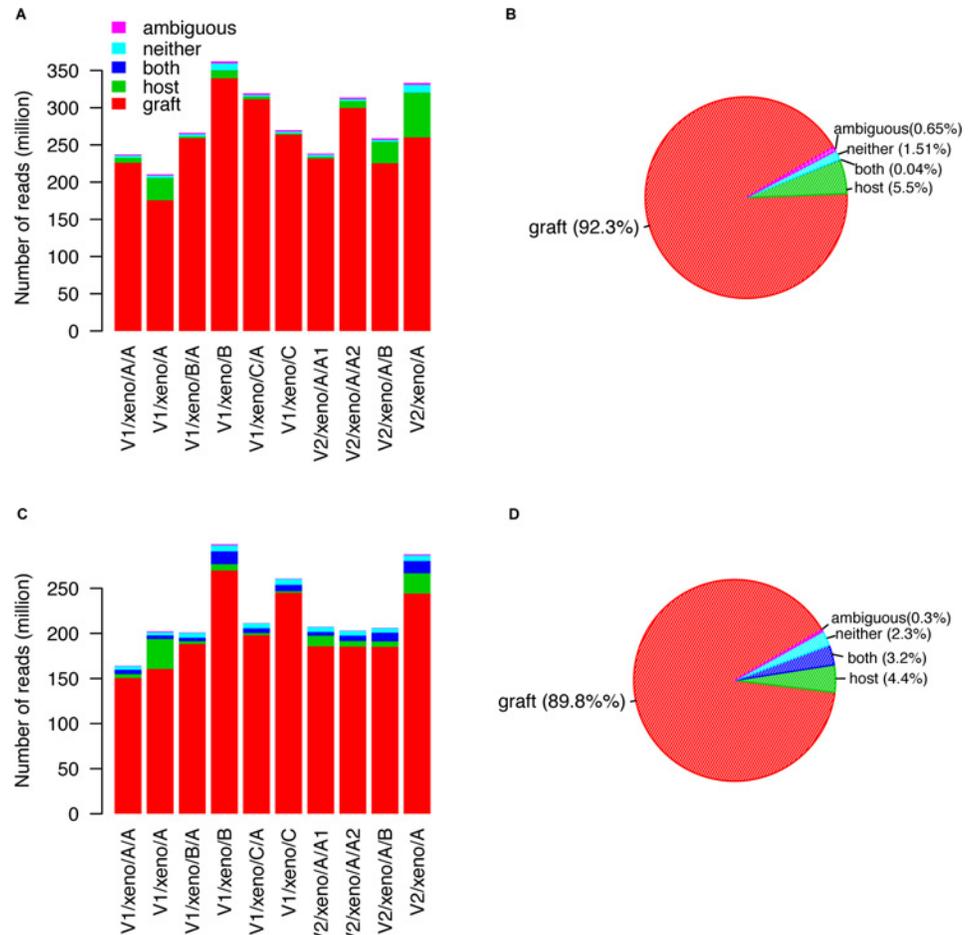
### Sequencing statistics

A total of 13 metastatic tissue and PDX tumor specimens were obtained for whole exome sequencing (WES) which included two patient tumor tissue specimens (visit 1 and visit 2), 10 PDXs, and the germline DNA from the patient. These generated a total of 3.58 billion reads ( $275.6 \pm 4.9$  million reads per sample) with 3.49 billion (97.49%) mapped to human reference genome hg19/GRCh37 (S1 Table). Additionally RNA-seq was performed on all (except the germline) specimen including 10 PDXs and two patient tumor tissue (visit 1 and visit 2) and these produced a total of 2.67 billion reads ( $221.8 \pm 53.3$  million reads per sample) with 2.48 billion (92.88%) mapped to the reference genome (S2 Table). The deep sequencing depth and high mappability promised comprehensive exploration of the mutational genome and transcriptome landscape.

We used the classification methods developed by Conway et al [28] to divide all WES and RNA-seq reads into 5 categories including “graft”, “host”, “both”, “neither” and “ambiguous”. Overall, 92.3% and 89.8% of reads were assigned as graft specific for WES and RNA-seq, respectively, demonstrating high species specificity of all xenografts (Fig 1; S3 & S4 Tables).

### Somatic mutations

To evaluate genome-wide homogeneity of biopsy tumor tissue and corresponding PDXs, we calculated the recall and precision using all detected somatic mutations. Fig 2A shows that most somatic mutations identified from biopsy tissues were rediscovered in PDXs with high recall rate (mean = 0.953, median = 0.958), indicating PDX models faithfully preserved genomic somatic mutations of the original tumors. However, the precision was relatively lower (mean = 0.656, median = 0.657) compared to recall, suggesting there were considerable amount of PDX specific variant calls. This low precision is likely not due to mapping artifacts as the sequence surrounding those variants (101bp centered on the variant) has poor mapping to the mouse genome using BLAT [22]. Presumably, PDX specific mutations mainly originated from two sources: host (mouse) reads contamination and subclone differences. Although we used Xenome to remove host-specific reads, this process is not completely effective given the relatively high similarity between mouse and human genome. Predictably, aligning mouse-derived sequences to the human genome would generate a considerable amount of false positive somatic mutations. In addition, the implantation of tumor cells in NOD-SCID mice could potentially change the frequency of tumor cell populations because of clonal selection, and such subclone drift could also produce many PDX specific mutations. In addition, the relative allele frequencies of somatic mutations were concordant between biopsy tissue and PDXs (Fig 2B). For example, on comparing patient tumors and germline DNA a somatic *TP53* mutation (I255F) was detected, which is reported to destabilize the *TP53* protein [29]. This variant was also detected in all PDXs (S5 Table). This mutation in *TP53* appeared to be heterozygous in



**Fig 1. Classification of xenograft whole exome sequencing (panels A and B) and RNA sequencing reads (panels C and D).** Reads generated from xenograft samples were divided into five groups including “graft”, “host”, “both”, “neither” and “ambiguous” using tool developed by Conway et al. (A) Reads assignments for 10 xenograft whole exome sequencing data. (B) Average proportion of whole exome sequencing reads assigned to the 5 groups mentioned above. (C) Reads assignments for 10 xenograft RNA-seq data. (D) Average proportion of RNA-seq reads assigned to the 5 groups mentioned above.

doi:10.1371/journal.pone.0145176.g001

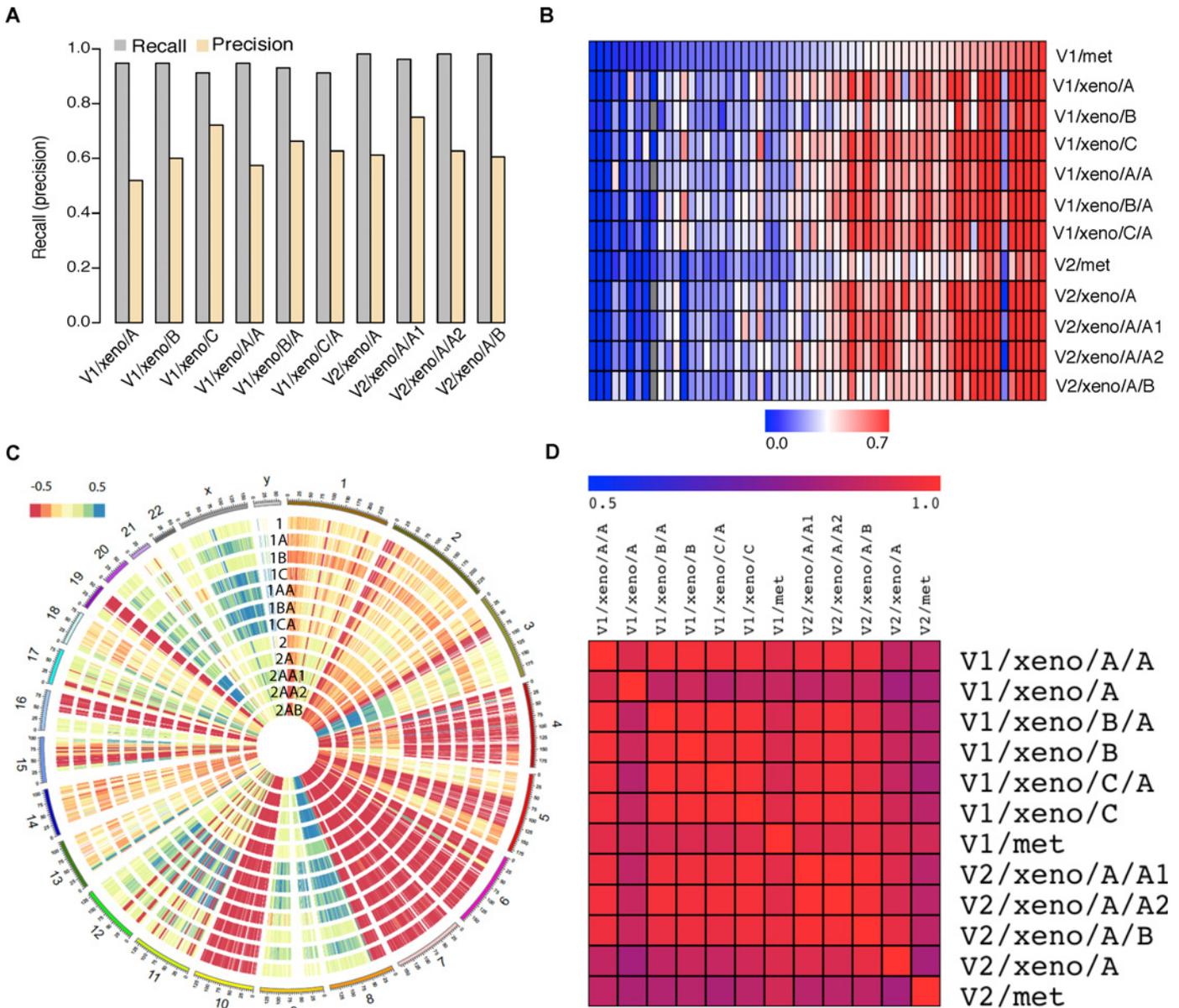
the DNA of the clinical tumor samples and of most PDX’s, but the mutant allele dominated the RNASeq reads in all samples, suggesting the mutant allele specific expression or genome deletion of another copy. After evaluating the CNV data, we found another copy of TP53 was indeed deleted in PDXs with  $\text{Log}_2R = -1.007 \pm 0.15$ . In patient tumor samples, the  $\text{Log}_2R$  was -0.6 and -0.65 in metastatic tissue from visit 1 and visit 2, respectively. The lower  $\text{log}_2R$  value in patient tissues might due to the lower tumor purity compared to that of PDXs.

DNA tumor purity was estimated at 75% for visit 1 and 76% for visit 2 metastatic samples using the frequency of the most frequent subclone, as computed by PurBayes. The clonality analyses indicated that all three major clones were maintained at relative stable frequencies between patient tissue and PDXs, as well as among different passages of PDXs (S2 Fig).

### Copy number variation

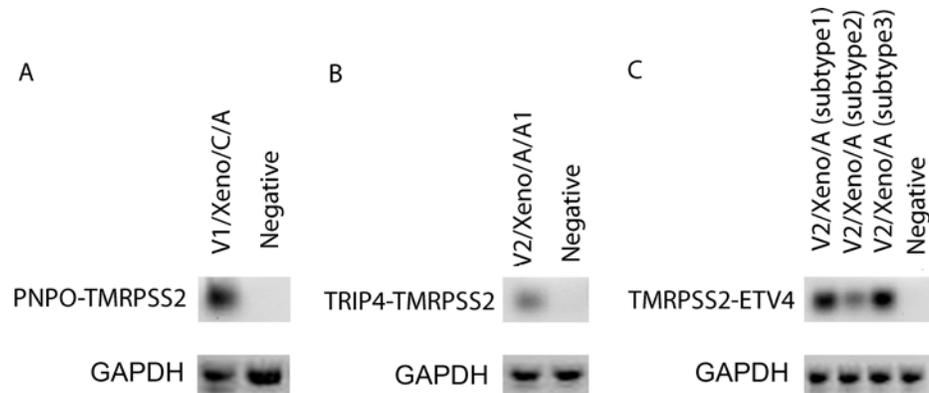
Genome wide copy number variation (CNV) analyses revealed high consistence among all samples (Fig 2C), with limited lineage-specific CNV differences (e.g. chr3q). We detected

somatic copy number gains for *AR* ( $\text{Log}_2R = 3.91 \pm 0.32$ ) and *FGFR1* ( $\log_2R = 1.49 \pm 0.13$ ) genes in tissue biopsies of both visits. As expected, similar levels of copy number gains of *AR* and *FGFR1* were also detected from all PDX samples (S6 Table). The *FGFR1* amplification was part of a multi-gene amplified region on chromosome 8 (S3 Fig). The *AR* amplification also included a number of other genes on the high level amplification (3 distal and 1 proximal), as



**Fig 2. Comparison of genome and transcriptome landscapes between patient tumor tissue and patient derived xenograft models (PDXs).** (A) Recall (grey) and precision (tan) of detected somatic mutations. Recall = number of somatic mutations called from both patient tissue and xenograft divided by total somatic mutations called from patient tissue; Precision = number of somatic mutations called from both patient tissue and xenograft divided by total mutations called from xenograft. (B) Heatmap showing the concordance of relative allele frequency of somatic mutations between patient tissues and xenografts. Rows correspond to patients and xenograft samples and columns correspond to 60 selected somatic mutations. (C) Circos plot showing profiles of copy number variation for patient tumor tissues and xenografts. From outside to inside, tracks correspond to 1 = V1/met, 1A = V1/xeno/A, 1B = V1/xeno/B, 1C = V1/xeno/C, 1AA = V1/xeno/A/A, 1BA = V1/xeno/B/A, V1/xeno/C/A, V2/met, V2/xeno/A, V2/xeno/A/A1, V2/xeno/A/A2 and V2/xeno/A/B. (D) Pair-wise gene expression correlation between patient tissues and xenografts. Correlation was measured by Pearson correlation coefficient. Gene expression was measured by  $\log_{10}$  (RPKM, Reads Per Kilobase exon per Million mapped reads).

doi:10.1371/journal.pone.0145176.g002



**Fig 3. Validation of gene fusion products identified by RNA seq.** The upper gel images showed the targeted amplification products, and the lower gel images showed the loading control amplification (GAPDH). (A) The amplification products for PNPO-TMRPSS2 fusion was detected in the xenograft sample V1/xenon/C/A. (B) TRIP4-TMRPSS2 fusion was detected in V2/xeno/A/A1. (C) Three different types of TMRPSS2-ETV4 fusions were detected in V2/Xeno/A. The sample V1/Xeno/A1 was used as negative control.

doi:10.1371/journal.pone.0145176.g003

well as two contiguous region of Copy Number gain, with successively diminishing amplitude away from the *AR* gene, potentially indicating that the *AR* amplification occurred as at least three successive events, each of which contributed to an accumulative number of *AR* gene copies over time. This three level structure was also observed in all xenograft DNA (S4 Fig). The *MLL3 (KMT2C)* gene was also deleted as part of a 4 gene deleted segment on chromosome 7 from 151253201 to 151884347, which also included *GALNTL5*, *GALNT11*, and *PRKAG2* (S5 Fig). There was also a significant 2.5 Mb deletion on chromosome 15 including *CYP1A1*, *CYP1A2*, *CSK*, *UBE2Q2*, *FBXO22*, *TSPAN3* etc. (S6 Fig). The same two regions were also deleted in the second visit sample. Overall, there were 385 and 451 genes deleted at the  $\text{Log}_2R \leq -0.5$  level and 97 and 84 genes amplified at the  $\text{Log}_2R \geq 0.5$  in the first and second visit tumor specimens, respectively.

### Gene fusion

A total of forty candidate fusions were detected, 8 of which were inter-chromosomal events (S7 Fig). We did not detect *TMRPSS2-ERG* gene fusion in this patient, however, we found that *TMRPSS2* gene was fused with other partners including *ETV4*, *TRIP4*, and *PNPO* in the visit 1 patient tumor tissue, and all these *TMRPSS2* involved fusion events were rediscovered in PDXs (S7 Table). We were able to confirm *TMRPSS2* and *ETV4*, *TRIP4*, and *PNPO* fusion events with QRT-PCR (Fig 3). These events were not observed in the visit 2 specimen and the failure to detect these fusions in visit 2 may have been due to RNA degradation and lower tumor content in the biopsy or due to the possibility that clones harboring these fusions represented therapy-sensitive disease.

### Gene expression

Gene expression profiles measured by RNA-seq exhibited an overall high degree of correlation between patient tumor tissue and derived PDX models (Pearson correlation coefficient = 0.933; min: 0.825; max: 0.996) (Fig 2D). However, we observed higher expression of *AR* splice variant 7 (*AR-V7*; fold change of 6.25) and *AR* full length (*AR-FL*; fold change of 4.71) in the first visit metastatic tissue relative to the second visit. The corresponding fold change in the PDXs for *ARV7* and *AR* full length between the first and second visits were 1.60 and 1.29, respectively (S8 Fig). We also compared the expression values (measured by RPKM) between visit 1 and

visit 2 tumors. When using the threshold of 2 fold change, we detected 461 and 311 genes which were up- and down-regulated in the metastatic tissue from the second visit tumor, respectively (S8 Table). We further performed the pathway analysis using Ingenuity IPA platform, we found that the top enriched pathways for up-regulated genes included *Oxidative Phosphorylation* ( $P = 2.2E-47$ ), *Mitochondrial Dysfunction* ( $P = 1.9E-42$ ) and *Protein Ubiquitination* ( $P = 6.9E-8$ ). The top enriched pathways for down-regulated genes included *Interferon Signaling* ( $P = 6.2E-7$ ), *Clathrin-mediated endocytosis signaling* ( $P = 6.8E-5$ ) and *Activation of IRF by cytosolic pattern recognition receptors* ( $P = 2.6E-4$ ).

## Discussion

The AR inhibitory action of enzalutamide includes competitive blocking of the ligand binding domain of the receptor, inhibition of the translocation of AR to the cell nucleus and inhibiting binding of AR to the DNA [30]. Pre-chemotherapy PSA response rate to enzalutamide (decrease in PSA by 50% of baseline after 12 weeks of treatment) is 78% and a complete or partial soft tissue response rate is 56% [31]. Recently AR-V7 expression has been reported as a molecular predictive factor of enzalutamide efficacy with AR-V7 expression being associated with low PSA response and shorter duration of progression free survival [32]. Our patient did not have a biochemical or a substantive clinical response after 12-weeks of enzalutamide drug exposure and both his metastases as well as the PDX specimens were detected to have AR-V7 expressions. Interestingly minimal decreases in expression levels were observed in the pre to post treatment specimens at 12-weeks. While these observations on AR-V7 expression and resistance to therapy are in agreement with previous reports, the decrease in expression levels of this splice variant post therapy suggests a role for other resistance factors as well including a role for other AR splice variants (AR23, AR45) as resistance factors of drug efficacy. Alternately a role for *TMPRSS2-ETV4* fusion protein or the expression of other androgen regulated gene fusions with *TMPRSS* interference in determining enzalutamide activity is also unclear. Finally a role for non-AR pathway genes in drug resistance could also be involved. We observed deletion of genes such as *PRKAG2* in tumor tissue and PDXs, a key gene which inactivates enzymes of the fatty acid and cholesterol biosynthesis pathways (a precursor of sex steroids) and which may affect intra tumor sex steroid levels. Increased *FGFR1* copy number was observed in tumor tissue and PDXs. Increase in copy number of this gene has been postulated to play a role in the development of initial hormone resistance to androgen deprivation therapy [33]. Its effect on causing resistance to secondary hormonal therapies such as enzalutamide remains unexplored.

We successfully developed PDXs from mCRPC sites which were able to faithfully represent patient tumor heterogeneity. The concordance in mutation detection and gene expression levels based on exome and RNA-seq data between human tumor and corresponding xenograft tumors was very high. Some of the PDX specific mutations have low accuracy which might be due to mixture of mouse genome. These observations indicate that these PDXs were able to recapitulate the genomic and transcriptomic characteristics of the patient's underlying cancer. These PDXs now allow the opportunity for evaluating several hypotheses for mechanisms of drug resistance and ultimately develop novel therapeutics that can guide patient intervention. One of the main challenges of developing prostate cancer PDXs for individualized testing of therapies in real-time (i.e. tumor "avatars") has been a highly variable take-rate, which may be due to several reasons including the source of tissue in prostate cancer (primary vs. metastatic); tissue procurement methods; mouse host strain and engraftment site (subcutaneous or renal sub-capsular). For example, take-rates have been reported to range from 5%–95% for primary prostate cancer, and 25%–95% for metastatic disease [11]. An additional challenge to widespread use of PDX models for lab to bedside translation is the requirement for an immune-

deficient host. As a result, PDX models are unable to faithfully recapitulate the relationships between different cell types and cytokines in the tumor microenvironment. This is of particular relevance when evaluating cancer immunotherapies or the role of the immune system in disease progression and/or response to AR-targeted therapy or chemotherapy. Another challenge for the successful adoption of using PDXs to enhance precision medicine is the feasibility of translating results from the lab to the bedside in real time after obtaining a functional validation of targetable genes. Finally, the availability of drugs that can be used against the driver mutations identified in this and other studies [4,34] is another practical consideration

## Conclusions

PDX models remain an attractive option as functional tumor models if they are able to reflect the tumor heterogeneity observed in the original patient tumor through initial stages of passaging [35]. We successfully elucidated the establishment of sequential PDX models from a patient on enzalutamide therapy in castrate resistant stage. Attempts to evaluate the underlying drug resistance mechanisms using this model system for the delivery of precision medicine in future are at present on-going.

## Supporting Information

**S1 Fig. Schematic representation of relationship between patient tissue (visit1 and visit2) and derived xenograft models (1st and 2nd generation).**

(TIFF)

**S2 Fig. The output of PurBayes is visualized by plotting the mutant allele counts versus the total reads at each site for V1/Met and derived PDXs (A-G), V2/Met and derived PDXs (H-L). Three clonal populations were indicated as black, blue and green lines, respectively. The corresponding dotted lines indicate the 95% confidence interval.**

(TIFF)

**S3 Fig. Screenshot of UCSC genome browser showing multi-gene amplified region on chromosome 8 harboring gene *FGFR1*. Y-axis indicates the log<sub>2</sub> ratio of reads coverage between tumor and germline DNA. V1/Met and derived PDXs were indicated as red tracks, V2/Met and derived PDXs were indicated as green tracks.**

(TIFF)

**S4 Fig. Screenshot of UCSC genome browser showing multiple levels of copy number gain around the AR gene (highlighted) on chromosome X. Y-axis indicates the log<sub>2</sub> ratio of reads coverage between tumor and germline DNA. V1/Met and derived PDXs were indicated using red tracks, V2/Met and derived PDXs were indicated using green tracks.**

(TIF)

**S5 Fig. Screenshot of UCSC genome browser showing 4 gene deleted segment on chromosome 7. Y-axis indicates the log<sub>2</sub> ratio of reads coverage between tumor and germline DNA. V1/Met and derived PDXs were indicated as red tracks, V2/Met and derived PDXs were indicated as green tracks.**

(TIFF)

**S6 Fig. Screenshot of UCSC genome browser showing 2.5 Mb segment deletion on chromosome 15. Y-axis indicates the log<sub>2</sub> ratio of reads coverage between tumor and germline DNA. V1/Met and derived PDXs were indicated as red tracks, V2/Met and derived PDXs were indicated as green tracks.**

(TIFF)

**S7 Fig. Circos plot showing 40 candidate gene fusions detected from patient tissue and xenografts.** Tmprss2 encompassing fusions were indicated by red links.

(TIFF)

**S8 Fig. Isoform specific expression for AR variant 7 (ARV7, left) and AR full length (ARF, right).** SRPM (splice reads per million) is number of splice reads that specifically support ARV7 (or ARF) normalized by total splice reads.

(TIFF)

**S1 Table. Whole exome sequencing (WES) statistics for 13 samples including 10 patient derived xenografts, patient tumor (Visit1 and Visit2) and blood tissue (Visit1).**

(XLSX)

**S2 Table. RNA sequencing (RNAseq) statistics for 12 samples including 10 patient derived xenografts, patient tumor (Visit1 and Visit2).**

(XLSX)

**S3 Table. Whole exome sequencing reads classification of 10 patient derived xenografts.**

(XLSX)

**S4 Table. RNA-seq reads classification of 10 patient derived xenografts.**

(XLSX)

**S5 Table. Clinically relevant DNA somatic variants.** Reads supporting reference allele and mutant allele are shown for both the WES and RNAseq data. Gene name (codon change) are provided.

(XLSX)

**S6 Table. Copy number variation (CNV) in clinically relevant genes.**

(XLSX)

**S7 Table. Tmprss2 encompassed fusions.** Number of paired-end reads supporting the fusion.

(XLSX)

**S8 Table. Differentially expressed genes (DEG) between patient tumor tissue before and after treatment (i.e. Visit 1 vs Visit 2).** DEGs were defined as those genes with RPKM fold change  $> 2$  or  $< 0.5$ .

(XLSX)

**S1 Animal Methods. Description of mice anesthesia and Euthanasia methods.**

(DOCX)

## Acknowledgments

The study is supported by the Mayo Clinic Centre for Individualized Medicine and the Pharmacogenomics Research Network (PGRN). This work was also supported by NIH grant R01CA174777 (to S.M.D.)

## Author Contributions

Conceived and designed the experiments: MK Ligu Wang FX PY RQ RW Liewei Wang. Performed the experiments: MK FX PY TDA PWE BPM REJ Liewei Wang. Analyzed the data: MK Ligu Wang HS SMD SNH PTV PB RQ DWM REC JEEP EDW Liewei Wang.

Contributed reagents/materials/analysis tools: Ligu Wang HS EDW PTV PB RQ DWM JEEP REC FX PY Liewei Wang. Wrote the paper: MK Ligu Wang HS SMD RQ GJ REJ Liewei Wang.

## References

1. Gerlinger M, Rowan AJ, Horswell S, Larkin J, Endesfelder D, et al. (2012) Intratumor heterogeneity and branched evolution revealed by multiregion sequencing. *The New England journal of medicine* 366: 883–892. doi: [10.1056/NEJMoa1113205](https://doi.org/10.1056/NEJMoa1113205) PMID: [22397650](https://pubmed.ncbi.nlm.nih.gov/22397650/)
2. Yates LR, Campbell PJ (2012) Evolution of the cancer genome. *Nature reviews Genetics* 13: 795–806. doi: [10.1038/nrg3317](https://doi.org/10.1038/nrg3317) PMID: [23044827](https://pubmed.ncbi.nlm.nih.gov/23044827/)
3. Barbieri CE, Bangma CH, Bjartell A, Catto JW, Culig Z, et al. (2013) The mutational landscape of prostate cancer. *European urology* 64: 567–576. doi: [10.1016/j.eururo.2013.05.029](https://doi.org/10.1016/j.eururo.2013.05.029) PMID: [23759327](https://pubmed.ncbi.nlm.nih.gov/23759327/)
4. Grasso CS, Wu YM, Robinson DR, Cao X, Dhanasekaran SM, et al. (2012) The mutational landscape of lethal castration-resistant prostate cancer. *Nature* 487: 239–243. doi: [10.1038/nature11125](https://doi.org/10.1038/nature11125) PMID: [22722839](https://pubmed.ncbi.nlm.nih.gov/22722839/)
5. Beltran H, Tomlins S, Aparicio A, Arora V, Rickman D, et al. (2014) Aggressive variants of castration-resistant prostate cancer. *Clinical cancer research: an official journal of the American Association for Cancer Research* 20: 2846–2850.
6. Spiotto MT, Pytynia M, Liu GF, Ranck MC, Widau R (2013) Animal models to study the mutational landscape for oral cavity and oropharyngeal cancers. *Journal of oral & maxillofacial research* 4: e1.
7. Siolas D, Hannon GJ (2013) Patient-derived tumor xenografts: transforming clinical samples into mouse models. *Cancer research* 73: 5315–5319. doi: [10.1158/0008-5472.CAN-13-1069](https://doi.org/10.1158/0008-5472.CAN-13-1069) PMID: [23733750](https://pubmed.ncbi.nlm.nih.gov/23733750/)
8. Lawrence MG, Taylor RA, Toivanen R, Pedersen J, Norden S, et al. (2013) A preclinical xenograft model of prostate cancer using human tumors. *Nature protocols* 8: 836–848. doi: [10.1038/nprot.2013.043](https://doi.org/10.1038/nprot.2013.043) PMID: [23558784](https://pubmed.ncbi.nlm.nih.gov/23558784/)
9. Corey E, Quinn JE, Vessella RL (2003) A novel method of generating prostate cancer metastases from orthotopic implants. *The Prostate* 56: 110–114. PMID: [12746835](https://pubmed.ncbi.nlm.nih.gov/12746835/)
10. Van Allen EM, Foye A, Wagle N, Kim W, Carter SL, et al. (2014) Successful whole-exome sequencing from a prostate cancer bone metastasis biopsy. *Prostate cancer and prostatic diseases* 17: 23–27. doi: [10.1038/pcan.2013.37](https://doi.org/10.1038/pcan.2013.37) PMID: [24366412](https://pubmed.ncbi.nlm.nih.gov/24366412/)
11. Lin D, Xue H, Wang Y, Wu R, Watahiki A, et al. (2014) Next generation patient-derived prostate cancer xenograft models. *Asian journal of andrology* 16: 407–412. doi: [10.4103/1008-682X.125394](https://doi.org/10.4103/1008-682X.125394) PMID: [24589467](https://pubmed.ncbi.nlm.nih.gov/24589467/)
12. Nicholson TM, Uchtmann KS, Valdez CD, Theberge AB, Miralem T, et al. (2013) Renal capsule xenografting and subcutaneous pellet implantation for the evaluation of prostate carcinogenesis and benign prostatic hyperplasia. *Journal of visualized experiments: JoVE*.
13. Li Y, Hwang TH, Oseth LA, Hauge A, Vessella RL, et al. (2012) AR intragenic deletions linked to androgen receptor splice variant expression and activity in models of prostate cancer progression. *Oncogene* 31: 4759–4767. doi: [10.1038/onc.2011.637](https://doi.org/10.1038/onc.2011.637) PMID: [22266865](https://pubmed.ncbi.nlm.nih.gov/22266865/)
14. McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, et al. (2010) The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome research* 20: 1297–1303. doi: [10.1101/gr.107524.110](https://doi.org/10.1101/gr.107524.110) PMID: [20644199](https://pubmed.ncbi.nlm.nih.gov/20644199/)
15. Shah SP, Morin RD, Khattra J, Prentice L, Pugh T, et al. (2009) Mutational evolution in a lobular breast tumour profiled at single nucleotide resolution. *Nature* 461: 809–813. doi: [10.1038/nature08489](https://doi.org/10.1038/nature08489) PMID: [19812674](https://pubmed.ncbi.nlm.nih.gov/19812674/)
16. Larson DE, Harris CC, Chen K, Koboldt DC, Abbott TE, et al. (2012) SomaticSniper: identification of somatic point mutations in whole genome sequencing data. *Bioinformatics* 28: 311–317. doi: [10.1093/bioinformatics/btr665](https://doi.org/10.1093/bioinformatics/btr665) PMID: [22155872](https://pubmed.ncbi.nlm.nih.gov/22155872/)
17. Roth A, Morin R, Ding J, Crisan A, Ha G, et al. (2012) JointSNVMix: A Probabilistic Model For Accurate Detection Of Somatic Mutations In Normal/Tumour Paired Next Generation Sequencing Data. *Bioinformatics*.
18. Abecasis GR, Auton A, Brooks LD, DePristo MA, Durbin RM, et al. (2012) An integrated map of genetic variation from 1,092 human genomes. *Nature* 491: 56–65. doi: [10.1038/nature11632](https://doi.org/10.1038/nature11632) PMID: [23128226](https://pubmed.ncbi.nlm.nih.gov/23128226/)
19. Pritchyk KM, Schiff BA, Newkirk KA, Krowiak E, Deeb ZE (2002) Metastatic renal cell carcinoma to the head and neck. *Laryngoscope* 112: 1598–1602. PMID: [12352670](https://pubmed.ncbi.nlm.nih.gov/12352670/)

20. Wang C, Evans JM, Bhagwate AV, Prodduturi N, Sarangi V, et al. (2014) PatternCNV: a versatile tool for detecting copy number changes from exome sequencing data. *Bioinformatics*.
21. Olshen AB, Venkatraman ES, Lucito R, Wigler M (2004) Circular binary segmentation for the analysis of array-based DNA copy number data. *Biostatistics* 5: 557–572. PMID: [15475419](#)
22. Kent WJ (2002) BLAT—the BLAST-like alignment tool. *Genome research* 12: 656–664. PMID: [11932250](#)
23. Kalari KR, Nair AA, Bhavsar JD, DR OB, Davila JI, et al. (2014) MAP-RSeq: Mayo Analysis Pipeline for RNA sequencing. *BMC bioinformatics* 15: 224. doi: [10.1186/1471-2105-15-224](#) PMID: [24972667](#)
24. Trapnell C, Pachter L, Salzberg SL (2009) TopHat: discovering splice junctions with RNA-Seq. *Bioinformatics* 25: 1105–1111. doi: [10.1093/bioinformatics/btp120](#) PMID: [19289445](#)
25. Langmead B (2010) Aligning short sequencing reads with Bowtie. *Current protocols in bioinformatics / editorial board, Andreas D Baxeavanis [et al]* Chapter 11: Unit 11 17.
26. Wang L, Wang S, Li W (2012) RSeQC: quality control of RNA-seq experiments. *Bioinformatics* 28: 2184–2185. doi: [10.1093/bioinformatics/bts356](#) PMID: [22743226](#)
27. Kim D, Salzberg SL (2011) TopHat-Fusion: an algorithm for discovery of novel fusion transcripts. *Genome biology* 12: R72. doi: [10.1186/gb-2011-12-8-r72](#) PMID: [21835007](#)
28. Conway T, Wazny J, Bromage A, Tymms M, Sooraj D, et al. (2012) Xenome—a tool for classifying reads from xenograft samples. *Bioinformatics* 28: i172–178. doi: [10.1093/bioinformatics/bts236](#) PMID: [22689758](#)
29. Bullock AN, Henckel J, Fersht AR (2000) Quantitative analysis of residual folding and DNA binding in mutant p53 core domain: definition of mutant states for rescue in cancer therapy. *Oncogene* 19: 1245–1256. PMID: [10713666](#)
30. Tran C, Ouk S, Clegg NJ, Chen Y, Watson PA, et al. (2009) Development of a second-generation anti-androgen for treatment of advanced prostate cancer. *Science* 324: 787–790. doi: [10.1126/science.1168175](#) PMID: [19359544](#)
31. Beer TM, Armstrong AJ, Rathkopf DE, Loriot Y, Sternberg CN, et al. (2014) Enzalutamide in metastatic prostate cancer before chemotherapy. *The New England journal of medicine* 371: 424–433. doi: [10.1056/NEJMoa1405095](#) PMID: [24881730](#)
32. Antonarakis ES, Lu C, Wang H, Lubner B, Nakazawa M, et al. (2014) AR-V7 and resistance to enzalutamide and abiraterone in prostate cancer. *The New England journal of medicine* 371: 1028–1038. doi: [10.1056/NEJMoa1315815](#) PMID: [25184630](#)
33. Edwards J, Krishna NS, Witton CJ, Bartlett JM (2003) Gene amplifications associated with the development of hormone-resistant prostate cancer. *Clinical cancer research: an official journal of the American Association for Cancer Research* 9: 5271–5281.
34. Robinson D, Van Allen EM, Wu YM, Schultz N, Lonigro RJ, et al. (2015) Integrative clinical genomics of advanced prostate cancer. *Cell* 161: 1215–1228. doi: [10.1016/j.cell.2015.05.001](#) PMID: [26000489](#)
35. Hidalgo M, Amant F, Biankin AV, Budinska E, Byrne AT, et al. (2014) Patient-derived xenograft models: an emerging platform for translational cancer research. *Cancer discovery* 4: 998–1013. doi: [10.1158/2159-8290.CD-14-0001](#) PMID: [25185190](#)