

# Divergent Evolution in Bilateral Prostate Cancer: a Case Study

Roni Haas<sup>1,2,3,4,†</sup>, Yash Patel<sup>1,2,3,4</sup>, Lydia Y. Liu<sup>1,2,3,4</sup>, Rong Rong Huang<sup>5</sup>, Adam Weiner<sup>2,3,4</sup>, Takafumi N. Yamaguchi<sup>1,2,3,4</sup>, Raag Agrawal<sup>1,2,3</sup>, Paul C. Boutros<sup>1,2,3,4,†</sup>, Robert E. Reiter<sup>2,3,4,†</sup>

<sup>1</sup> Department of Human Genetics, University of California, Los Angeles, USA

<sup>2</sup> Department of Urology, University of California, Los Angeles, USA

<sup>3</sup> Jonsson Comprehensive Cancer Center, University of California, Los Angeles, USA

<sup>4</sup> Institute for Precision Health, University of California, Los Angeles, USA

<sup>5</sup> Department of Pathology, University of California, Los Angeles, USA

† Corresponding authors

## Abstract

Multifocal prostate cancer is a prevalent phenomenon, with most cases remaining uncharacterized from a genomic perspective. A patient presented with bilateral prostate cancer. On systematic biopsy, two indistinguishable clinicopathologic lesions were detected. Whole-genome sequencing displayed somatically unrelated tumours with distinct driver CNA regions, suggesting independent origins of the two tumors. We demonstrated that similar clinicopathologic multifocal tumours, which might be interpreted as clonal disease, can in fact represent independent cancers. Genetic prognostics can prevent mischaracterization of multifocal disease to enable optimal patient management.

## Results and Discussion

In this report, we investigated the bilateral prostate cancer of a patient in his 60s, who is genetically similar to the European reference population (**Supplementary Table 1**). The patient was diagnosed with bilateral Gleason Grade (GG) 1 tumours on systematic biopsy and was upgraded to GG3 in both on pathology from surgical specimens. Both tumours harboured 60% pattern 4 and 4% tertiary pattern 5. There were no cribriform or intraductal components and both tumors were confined to the prostate. The uncanny clinical and histologic similarities between the two tumors (**Figure 1A**) inspired molecular evaluations to test if nearly indistinguishable clinicopathologic features reflect similar tumor biology.

We sampled the separate tumours (A and B) from freshly frozen radical prostatectomy tissue. A benign prostate tissue, procured simultaneously from the same patient, was used as a reference control. DNA whole genome sequencing (WGS) was carried out to interrogate the genomic landscape and evolutionary relationship of the tumors (**Figure 1A; Supplementary Figure 1**).

We first evaluated the extent of genomic instability in each tumour as reflected by mutation counts. Larger numbers of somatic single nucleotide variants (SNVs) and structural variants (SV) were detected in tumour A compared to B, and none were shared between the two (**Table 1; Supplementary Tables 2-3; Supplementary Figure 2**). Despite this, SNVs in both tumours were largely the consequences of the same mutagenic processes (**Figure 1B**). The tumors experienced similar profiles of base change mutations (**Figure 1C**) and contained multiple SNVs that are predicted to have a significant protein functional effect. No genes shared functional SNV in both tumours

**(Figure 1D).** Among driver genes, *BCL6* and *NUP214* were functionally affected in Tumour A, and *SPOP* in Tumour B.

Next, we compared the copy number aberration (CNA) profiles of tumours A and B to identify possible shared CNA-driver regions, which could imply a shared evolutionary origin (**Table 1; Figure 1E; Supplementary Table 4**). Three CNA regions overlapped (**Supplementary Figure 3**), but only one, on chromosome 16, was a CNA-driver region spanning *FBXO3* (**Supplementary Figure 3C**). However, the breakpoints of these events differ, indicating convergent evolution rather than a shared origin.

Consistent with SNV and SV results, slightly higher genomic instability originating from CNAs was observed in Tumour A (**Table 1**). This involved CNA-driver regions, including *MYC* and *NCOA2* gains on chromosome 8q and *NKX3-1* deletion on chromosome 8p, typically seen in prostate cancer. By contrast Tumour B was characterized by *PTEN* deletion on chromosome 10, and *RB1* deletion on chromosome 13. Overall, the CNA landscape displayed distinctive clonal drivers.

Taken together, our results support distinct genetic origin of the two tumours, indicating that they arose independently and represent two separate cancers. The remarkably similar clinicopathologic features could have been mistakenly considered clonal without further genetic characterization. This study, and previous works<sup>1-3</sup> emphasize that the genetic prognostic should be carefully considered in situations of multifocal prostate tumors, which are relatively common<sup>4,5</sup>. Mischaracterizing multifocal lesions as a single prostate cancer may prevent optimal patient management.

## Methods

### Sample preparation and WGS

Two tumor foci, A and B were punched out from fresh OCT embedded tissue blocks after central pathology assessment. Tumor DNA was extracted from the same tumor tissue cores using Qiagen All Prep DNA mini kit (Qiagen, cat. 80204). The sample from Tumor A was in the anterior right, and from Tumor B in the posterior left, close to the dominant tumor (0.8 cm). A normal prostate core was sampled from a distal OCT block, from which the normal DNA was extracted using Qiagen QIAamp DNA mini kit (Qiagen, Cat 51304).

AnaPrep Tissue DNA Extraction kit (BioChain, cat. Z1322004) was used for DNA isolation. 500 to 1000 ng of input DNA was used for the library construction. Libraries were prepared using KAPA HyperPrep kit (Cat. KR0961, Roche). High throughput sequencing was performed by Illumina NovaSeq 6000 sequencer S4 platform with 150bp paired-end reads in 1.5 lanes. Coverage depths of 120x for the tumor samples and 80x for the normal tissue were achieved. A data quality check was done on Illumina SAV. Demultiplexing was performed with Illumina Bcl2fastq v2.19.1.403 software.

### WGS processing

DNA reads were tested with FastQC v0.11.8<sup>6</sup> for quality assurance before alignment. The reads were then mapped to the human GRCh38 reference genome using BWA-MEM2 v2.2.1<sup>7</sup>, and aligned SAM files were converted to BAM files using SAMtools v1.12<sup>8</sup>. Next,

using Picard Tool's v2.26.10<sup>9</sup>, the resulting BAM files were sorted in coordinate order, duplicates were marked, and the BAM files were indexed. The indexed BAM files went through indel-realignment using the Genome Analysis Toolkit (GATK) v3.7.0<sup>10</sup> and base quality score recalibration was done using GATK v4.1.9.0<sup>10</sup> in tumour-normal pairs. Separate BAM files were generated for the tumour and normal samples, and their headers were modified using SAMtools v1.12<sup>8</sup>. Qualimap version v.2.2.2-dev<sup>11</sup> was used to verify high coverage across the genome (**Supplementary Figure 1**). The mean coverage was 105X for T1, and 95X for T2.

### **Mutation calling**

Processed BAM files from the previous step were taken for mutation calling in tumour-normal pairs. SNVs were called separately using three algorithms, Mutect2 of GATK v4.2.0.0<sup>10</sup>, Strelka2 v2.9.10<sup>12</sup> and SomaticSniper v1.0.5.0<sup>13</sup>. To increase reliability, SNVs were declared only for sites that were detected using at least two algorithms for each tumour site. Strict germline variant filtering was performed, excluding any variant in the normal control sample and in gnomAD v3.1.2<sup>14</sup>. The called SNVs were annotated using SnpEff v5.0e<sup>15</sup>. Mutational signatures were identified using SigProfilerExtractor v1.14<sup>16</sup>.

SVs were identified using DELLY v0.8.7<sup>17</sup>. For reliability, detected sites were filtered to exclude germline structural variants. SVs that were identified with DELLY, despite very low coverage in their area in the normal sample, were excluded. Annotations were obtained using SnpEff v5.0e<sup>15</sup>. CNAs were detected using Battenberg v2.2.9<sup>18</sup>, with a purity of 0.41 and 0.21 for tumours A and B respectively, and ploidy of ~2 for both. Each call was reviewed manually to filter out false findings.

### **Genetic ancestry**

Genetic ancestry was predicted using Peddy v0.4.8, with WGS of the normal sample as input<sup>19</sup>. The prediction results calculated by Peddy are presented in **Supplementary Table 1**.

### **Visualization**

Visualizations were generated in the R environment v4.2.0 using the BoutrosLab.plotting.general package v7.0.3<sup>20</sup>.

## Competing Interests

P.C.B. sits on the Scientific Advisory Boards of BioSymetrics Inc and Intersect Diagnostics Inc., and formerly sat on that of Sage Bionetworks. All other authors declare they have no conflicts of interest.

## Author Contributions

Sample preparation: R.R.H

Data processing analyses and visualization: R.H

Generated tools: Y.P, L.Y.L., T.N.Y, R.A

Clinical annotation: A.W., R.E.R.

Supervised research: R.E.R, P.C.B

Wrote the first draft of the manuscript: R.H

Approved the manuscript: all authors.

## Ethical Statement

The Institutional Review Board of the University of California Los Angeles gave ethical approval for this work. The patient signed a universal consent for analyses of biological samples (#15-001395).

## Data and Materials Availability

All data produced in the present study are available upon reasonable request to the authors.

## Acknowledgments

R.H is supported by EMBO Postdoctoral Fellowship ALTF 1131-2021 and the Prostate Cancer Foundation Young Investigator Award 22YOUN32. R.A is supported by the NIH grant T32GM008042. ABW was supported by the UCLA Dr. Allen and Charlotte Ginsburg Fellowship in Precision Genomic Medicine and the Prostate Cancer Foundation Young Investigator Award (23YOUN21). This work was supported by the NIH through awards P30CA016042, U2CCA271894, R01CA270108 and P50CA092131. It was supported by the DOD through awards W81XWH2210247 and W81XWH2210751. This work was supported by a Prostate Cancer Foundation Special Challenge Award to PCB (Award ID #: 20CHAS01) made possible by the generosity of Mr. Larry Ruvo.

## References

1. Boutros, P. C. *et al.* Spatial genomic heterogeneity within localized, multifocal prostate cancer. *Nat. Genet.* **47**, 736–745 (2015).
2. Wilkinson, S. *et al.* A case report of multiple primary prostate tumors with differential drug sensitivity. *Nat. Commun.* **11**, 837 (2020).
3. the ICGC Prostate Group *et al.* Analysis of the genetic phylogeny of multifocal prostate cancer identifies multiple independent clonal expansions in neoplastic and morphologically normal prostate tissue. *Nat. Genet.* **47**, 367–372 (2015).
4. Djavan, B. *et al.* Predictability and significance of multifocal prostate cancer in the radical prostatectomy specimen. *Tech. Urol.* **5**, 139–142 (1999).
5. Wise, A. M., Stamey, T. A., McNeal, J. E. & Clayton, J. L. Morphologic and clinical significance of multifocal prostate cancers in radical prostatectomy specimens. *Urology* **60**, 264–269 (2002).
6. Wingett, S. W. & Andrews, S. FastQ Screen: A tool for multi-genome mapping and quality control. *F1000Research* **7**, 1338 (2018).
7. Vasimuddin, Md., Misra, S., Li, H. & Aluru, S. Efficient Architecture-Aware Acceleration of BWA-MEM for Multicore Systems. in *2019 IEEE International Parallel and Distributed Processing Symposium (IPDPS)* 314–324 (IEEE, Rio de Janeiro, Brazil, 2019). doi:10.1109/IPDPS.2019.00041.
8. Li, H. *et al.* The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**, 2078–2079 (2009).
9. Broad Institute. *Picard Toolkit*. <https://Broadinstitute.Github.io/Picard/>. (2019).

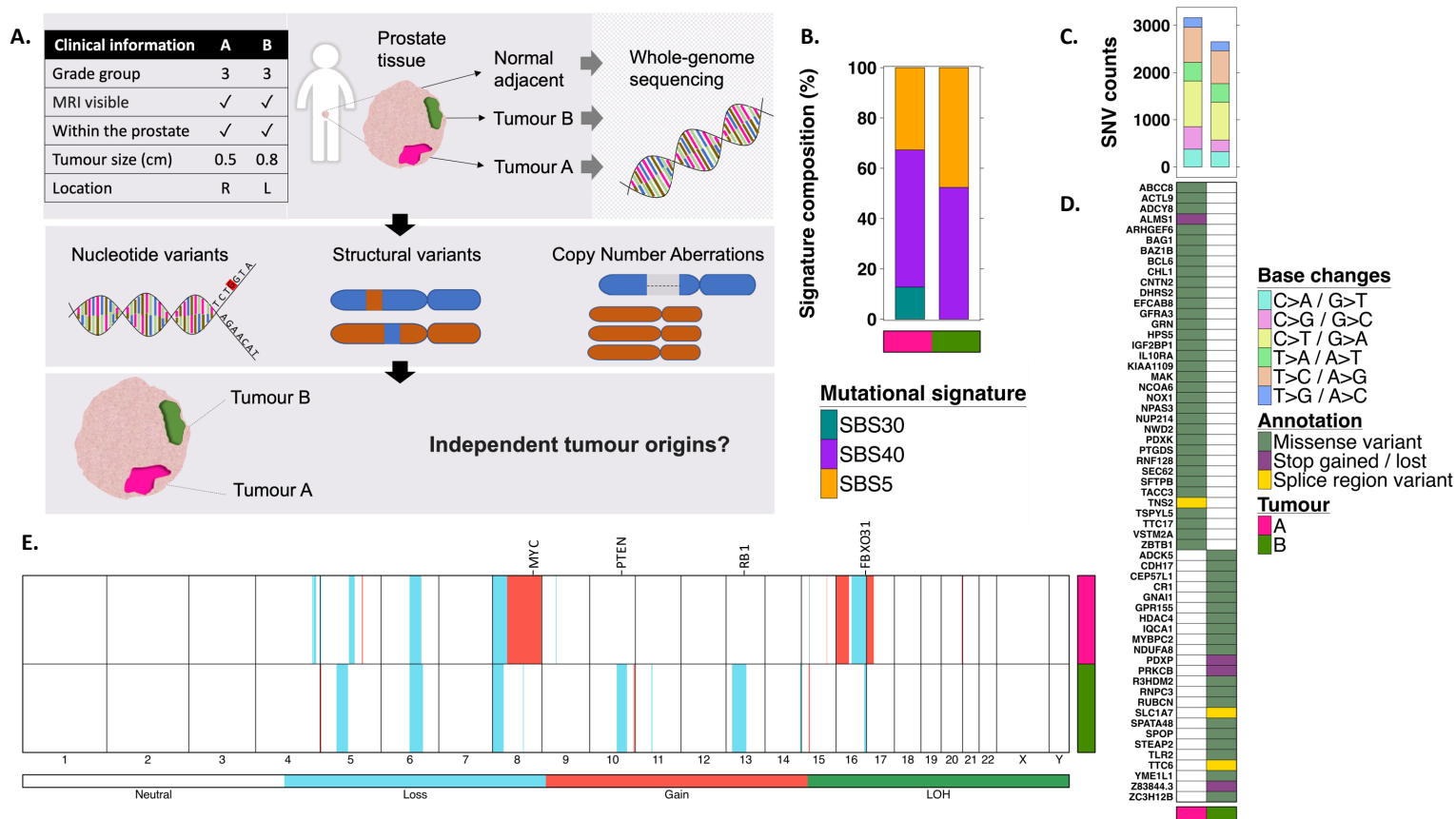
10. Auwera, G. A. *et al.* From FastQ Data to High-Confidence Variant Calls: The Genome Analysis Toolkit Best Practices Pipeline. *Curr. Protoc. Bioinforma.* **43**, (2013).
11. García-Alcalde, F. *et al.* Qualimap: evaluating next-generation sequencing alignment data. *Bioinformatics* **28**, 2678–2679 (2012).
12. Kim, S. *et al.* Strelka2: fast and accurate calling of germline and somatic variants. *Nat. Methods* **15**, 591–594 (2018).
13. Larson, D. E. *et al.* SomaticSniper: identification of somatic point mutations in whole genome sequencing data. *Bioinformatics* **28**, 311–317 (2012).
14. Chen, S. *et al.* A genomic mutational constraint map using variation in 76,156 human genomes. *Nature* **625**, 92–100 (2024).
15. Cingolani, P. *et al.* A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of *Drosophila melanogaster* strain w1118; iso-2; iso-3. *Fly (Austin)* **6**, 80–92 (2012).
16. Islam, S. M. A. *et al.* Uncovering novel mutational signatures by de novo extraction with SigProfilerExtractor. *Cell Genomics* **2**, 100179 (2022).
17. Rausch, T. *et al.* DELLY: structural variant discovery by integrated paired-end and split-read analysis. *Bioinformatics* **28**, i333–i339 (2012).
18. Nik-Zainal, S. *et al.* The Life History of 21 Breast Cancers. *Cell* **149**, 994–1007 (2012).
19. Pedersen, B. S. & Quinlan, A. R. Who's Who? Detecting and Resolving Sample Anomalies in Human DNA Sequencing Studies with Peddy. *Am. J. Hum. Genet.* **100**, 406–413 (2017).
20. P'ng, C. *et al.* BPG: Seamless, automated and interactive visualization of scientific data. *BMC Bioinformatics* **20**, 42 (2019).

## Figures and Tables

Table 1. mutation counts

	Tumour A	Tumour B	Shared
<b>SNV</b>	3,164	2,654	0
<b>SV</b>	62	46	0
<b>Gains</b>	6	3	0
<b>Loses</b>	9	10	3
<b>LOH</b>	1	0	0

<sup>1</sup>SNV, somatic single nucleotide variants (SNV), somatic structural variants (SV), loss of heterozygosity (LOH).



**Figure 1. Genomic landscape.** **A.** Schematic of study design and workflow. **B.** Mutational signatures. Signature etiologies: SBS30, defective DNA base excision repair due to NTHL1 mutation; SBS40, unknown; SBS5, unknown. **C.** SNV counts categorized by the base change type. **D.** Functionally affected genes by SNVs **E.** Genome-wide view of CNAs. Chromosome numbers are presented on the x-axis. The colored lines represent CNA classes. LOH, loss of heterozygosity.