nature portfolio

Corresponding author(s): Alexander W. Wyatt

Last updated by author(s): Aug 10, 2023

Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our Editorial Policies and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒	A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☐	☒	A description of all covariates tested
☐	☒	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒	For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☐	☒	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	No custom software was used to collect the data. Raw sequencing data was generated in-house. Sample meta-data was assembled in-house from sample/data quality metrics and clinical charts.
Data analysis	All statistical tests and data analyses were conducted in Python 3.7 (using pandas 0.25.0, numpy 1.16.4, scipy 1.6.2, lifelines v0.26.4, and statsmodels 0.12.2), Julia 1.5.1 (using HypothesisTests 0.10.9, and Distributions 0.25.53), and R 3.6.1 (using dplyr 1.0.7) languages. Visualizations were generated using matplotlib version 3.3.4 (Python) and ggplot 3.1.1 (R). The following bioinformatics/genomic analysis software was used: cutadapt-1.11, seqkit-0.8, Bowtie-2.3.0, 2.3.4.3, samblaster-0.1.24, bedtools-2.25.0, samtools 1.12 (htslib 1.12), Picard 2.25.6, Mutato version 0.7, ANNOVAR (version 20191024), Picard 2.25.6, SCHISM 1.1.2, and PyClone 0.13.1. All custom code is available at https://github.com/amurth/de_novo_cspc_multifocal_sequencing. Comprehensive methodology including all software (and versions) used is annotated within the manuscript and our Supplementary Information document.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our policy

The human reference genome hg38 was downloaded from UCSC. We used exon and TSS coordinates from the RefSeq Matched Annotation from NCBI and EMBL-EBI (MANE) database. All de-identified sequencing data have been deposited in the European Genome-Phenome Archive (EGA) database under the accession code (EGAS00001006466) and is available under standard EGA controlled release. All other data supporting the findings of this study are available within the article (including its supplementary information and data files). Source data for Figures 1, 2, 3, 4, 5, and 7 and Extended Data Figures 1, 2, 3, 5, 6, 9, and 10 are provided as source data files.

The following databases were used for mutation annotation: COSMICv77, ExAC v0.3, Kaviar (2016-02-04 release), ClinVar (2019-03-06 release)

Mutation and copy number data from the following studies (referenced in the main text) were used in Figure 4a,c and Extended Data Figure 3d-e:
https://www.cbioportal.org/study/summary?id=prad_su2c_2015
https://www.cbioportal.org/study/summary?id=prad_mcspc_mskcc_2020
https://www.cbioportal.org/study/summary?id=prad_tcga_pub

Tumor fraction data from the following studies were used in Figure 2d (reference in the main text and on the figure):
Vandekerkhove, G. et al. Circulating Tumor DNA Abundance and Potential Utility in De Novo Metastatic Prostate Cancer. Eur. Urol. 75, 667–675 (2019)
Mayrhofer, M. et al. Cell-free DNA profiling of metastatic prostate cancer reveals microsatellite instability, structural rearrangements and clonal hematopoiesis. Genome Med. 10, 85 (2018).
Annala, M. et al. Circulating Tumor DNA Genomics Correlate with Resistance to Abiraterone and Enzalutamide in Prostate Cancer. Cancer Discov. 8, 444–457 (2018).
Annala, M. et al. Cabazitaxel versus abiraterone or enzalutamide in poor prognosis metastatic castration-resistant prostate cancer: a multicentre, randomised, open-label, phase 2 trial. Ann. Oncol. 2021/04/10, (2021).

Human research participants

Policy information about studies involving human research participants and Sex and Gender in Research.

Reporting on sex and gender	All participants in this study were biologically male with de novo metastatic castration-sensitive prostate cancer.
Population characteristics	For patients sourced from LoMP I/II studies, participants must be 18-85 years old with no symptoms clearly-related to metastatic disease or other previous or current malignancies that may interfere with treatment or assessment. Full inclusion/exclusion criteria is available at clinicaltrials.gov (NCT02138721, NCT03655886). Key clinical characteristics are included in Figure 1 (cohort averages) and Supplementary Table 1 (per-patient). All patients voluntarily consented to provide samples for future research.
Recruitment	Patients were primarily selected from the prospective LoMP1/2 clinical trials (NCT02138721, NCT03655886) and voluntarily consented to providing samples for future research. Clinical eligibility for the clinical trials (from which patients for this present study were sourced) may influence the baseline clinical demography characteristics of our study cohort.
Ethics oversight	Approval for collection and analysis of patient tumor samples was granted by the Committee for Ethics of Ghent University Hospital and the University of British Columbia Research Ethics Board. This study was conducted in accordance with the Declaration of Helsinki and all participants provided written informed consent prior to enrollment.

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	Sample size was not predetermined for this analysis. Patients who underwent surgery via the clinical trials (NCT02138721, NCT03655886) were available for this study, subject to limits in sample acquisition and patient consent. The exploratory and descriptive nature of the study plus lac of predefined analyses means that target sample size were not possible.

Data exclusions	Pre-specified exclusion criteria is listed in the clinical trial descriptions (NCT02138721, NCT03655886). These clinical protocols are previously published and publicly available. The rationale and denominators for sub-analyses with sample-exclusions are clearly stated in the manuscript and/or figure captions.
Replication	All analyses are descriptive in nature and can be reproduced using the available code, public EGA data, structured data with Supplementary Tables, and Source Data files. No in vitro experiments requiring technical or biological replicates were performed. As described in the manuscript, select plasma or tissue DNA libraries were subject to both targeted and whole exome sequencing, serving as technical replicates for variables identified through targeted sequencing. All in silico analyses/experiments/models were either deterministic (not requiring replication) or had high convergence (i.e. negligibly small probability of inter-replicate variance that was empirically tested). Random seeds were used for all relevant analyses for reproducibility (documented in our online code repository).
Randomization	This exploratory study did not involve any analyses requiring patient/sample randomization. All analyses were descriptive in nature, and all relevant samples were analyzed together (unless they were categorically irrelevant for the analysis at hand).
Blinding	This is an exploratory biomarker study (i.e. not a clinical study) and therefore blinding was not performed.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☐	☒ Antibodies
☒	☐ Eukaryotic cell lines
☒	☐ Palaeontology and archaeology
☒	☐ Animals and other organisms
☐	☒ Clinical data
☒	☐ Dual use research of concern

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging

Antibodies

Antibodies used	All antibody information is described in Supplementary Table 10 (and summarized below); Methodology used is as described in:Verbeke, S. L. J. et al. A reappraisal of hemangiopericytoma of bone; analysis of cases reclassified as synovial sarcoma and solitary fibrous tumor of bone. Am. J. Surg. Pathol. 34, 777–783 (2010). p53 (Clone: DO-7; Source: Roche; Dilution: RTU; Antigen Retrieval: EDTA; Positive Control: Tonsil) PTEN (Clone: D4.3; Source: Cell Signaling Tech.; Dilution: 1:50; Antigen Retrieval: EDTA; Positive Control: Tonsil) RB1 (Clone: G3-245; Source: BD; Dilution: 1:25 Antigen Retrieval: EDTA; Positive Control: Tonsil)
Validation	All antibodies were tested on tonsil tissue as a positive control.

Clinical data

Policy information about clinical studies

All manuscripts should comply with the ICMJE guidelines for publication of clinical research and a completed CONSORT checklist must be included with all submissions.

Clinical trial registration	40/43 patients were enrolled in NCT02138721 or NCT03655886 (while the remaining 3 were sourced from a hospital-based registry).
Study protocol	Patient inclusion is specified in the original clinical trial protocol available at clinicaltrials.gov or in the previously published manuscript (https://doi.org/10.1016/j.euros.2021.05.006).
Data collection	Patient enrollment and sample collection occurred between May 2014 - April 2021 in the context of the clinical trials (NCT02138721 or NCT03655886).
Outcomes	For several exploratory analyses, we utilize castration-resistance free survival (which is a previously reported endpoint in the LoMP I study), defined as the date of diagnostic biopsy to date of castration resistance (European Association of Urology guidelines)