

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
<i>Only common tests should be described solely by name; describe more complex techniques in the Methods section.</i> |
| ☐ | ☒ | 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
<i>Give P values as exact values whenever suitable.</i> |
| ☒ | ☐ | 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	Data processing was performed using fastp (v0.20.0), STAR (v2.7.0), RSEM (v1.2.28), RNASEQC (v2.3.5), samtools (v1.11).
Data analysis	Data analysis was performed using R (v4.1.3) or prism (v10). Expression normalization was done using the R packages tximport (v1.22) and edgeR (v3.36). Differential expression analyses used the R packages DESeq2 (v1.34) and fgsea (v1.20). Machine learning models were performed using python (v3.13.7) and scikit-learn (v1.6.1).

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](#)

De-identified clinical and demographic data for individuals included in this study as well as model scores and other data are provided in the Supplementary Tables. Gene expression data from GTEx and TCGA were retrieved from the UCSC Xena Repository. Gene expression data from the UROMOL BLCA study were obtained

from Lindskrog et al (EGAS00001004693). Cell type marker gene lists were downloaded from the PanglaoDB resource (<https://panglaoDB.se/markers.html>). Custom gene signatures were generated using the Tabula Sapiens gene expression data. Interferon gamma and T cell gene lists were obtained from Ayers et al. and Rijnders et al, respectively. The RNA sequencing data generated in this study is available in dbGaP (phs004116). The bioinformatic codes used for various classifiers is available on <https://github.com/kevinjlustanford/uRARESeq>. Requests for additional information should be addressed to M.D.

Research involving human participants, their data, or biological material

Policy information about studies with [human participants or human data](#). See also policy information about [sex, gender \(identity/presentation\), and sexual orientation](#) and [race, ethnicity and racism](#).

Reporting on sex and gender

This study includes both male and female participants and sex was self-reported and identified from medical records. Summarized sex statistics is in Extended Data Table 1 and clinical and demographic characteristics such as age, sex, and smoking history of all participants in the study can be found in Supplemental Table 1. Our cohort consist of approximately ~74-77% male and ~23-26% female samples in the control cohort and ~82-87% male and ~13-18% female samples in the BLCA cohort. The sex ratio observed in this study is consistent with the sex prevalence of bladder cancer. Sex was considered as part of experimental design. Both males and females were eligible to participate. To minimize the risk of our machine learning-based biomarkers being impacted by sex, we removed genes with sex-specific expression differences as potential features

Reporting on race, ethnicity, or other socially relevant groupings

Social relevant grouping were not collected in this study.

Population characteristics

For the control cohort, the median age of the training cohort was 61 (range 23-96) and 67 in the validation cohort (range 20-93). For the bladder cancer cohort, the median age of the training cohort was 75 (range 29-94) and 75 (range 27-100) in the validation cohort. Our cohort consist of approximately ~72-74% male and ~26-28% female control samples and ~82-87% male and ~13-18% female samples in the cancer cohort. Summarized sex statistics is in Extended Data Table 1 and clinical and demographic characteristics such as age, sex, and smoking history of all participants in the study can be found in Supplemental Table 1.

Recruitment

Patients were recruited based on confirmed diagnosis of bladder cancer, prostate cancer, renal cell carcinoma, or based on negative pathology findings from urine cystoscopy. All participants provided written informed consent for the use of their clinical data and biospecimens for research using protocols approved by Institutional Review Boards at their respective centers. Collection centers included Stanford University and VA Palo Alto. Participants were not compensated for this study.

Ethics oversight

All participants provided written informed consent for the use of their clinical data and biospecimens for research and approved by Institutional Review Boards at their respective centers. Collection centers included Stanford University and VA Palo Alto.

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](https://www.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 cohort sizes were determined by sample availability, so no statistical calculations were performed to determine cohort size prior to patient enrollment

Data exclusions

No data were excluded in this study

Replication

Technical experiments performed in this study were validated in n=3 biological samples. In addition, the study findings were validated in an independent validation cohort

Randomization

Randomization was not applicable in this study because all samples were assigned based on confirmed diagnosis of bladder cancer, prostate cancer, kidney cancer, or non-cancer based on negative cystoscopy results.

Blinding

Blinding was not performed in this study because this was not necessary for this observational study for the purpose of method development and proof-of-concept data

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
☒	☐ Plants

Methods

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

Plants

Seed stocks

Report on the source of all seed stocks or other plant material used. If applicable, state the seed stock centre and catalogue number. If plant specimens were collected from the field, describe the collection location, date and sampling procedures.

Novel plant genotypes

Describe the methods by which all novel plant genotypes were produced. This includes those generated by transgenic approaches, gene editing, chemical/radiation-based mutagenesis and hybridization. For transgenic lines, describe the transformation method, the number of independent lines analyzed and the generation upon which experiments were performed. For gene-edited lines, describe the editor used, the endogenous sequence targeted for editing, the targeting guide RNA sequence (if applicable) and how the editor was applied.

Authentication

Describe any authentication procedures for each seed stock used or novel genotype generated. Describe any experiments used to assess the effect of a mutation and, where applicable, how potential secondary effects (e.g. second site T-DNA insertions, mosaicism, off-target gene editing) were examined.