

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 (<i>n</i>) 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. <i>F</i> , <i>t</i> , <i>r</i>) with confidence intervals, effect sizes, degrees of freedom and <i>P</i> 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 <i>d</i> , Pearson's <i>r</i>), 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	scWGS pipeline: https://github.com/mondrian-scwgs/mondrian v0.93.0 svtyper: https://github.com/marcjwilliams1/svtyper v0.7.7 duplex sequencing pipeline: https://github.com/msk-access/nucleo v3.1.2
Data analysis	SIGNALS: https://github.com/shahcompbio/signals v0.12.0 HMMclone: https://github.com/shahcompbio/HMMclone v0.1.0 Figure generation: https://github.com/marcjwilliams1/spectrum_cfdna_figures v0.1.0

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 MSK SPECTRUM study is registered on dbGaP under accession number phs002857.v2.p1 (https://www.ncbi.nlm.nih.gov/projects/gap/cgi-bin/study.cgi?study_id=phs002857.v2.p1).

A Synapse page for processed data is available under accession syn66399325

Summary tables include sequencing coverage, cfDNA tumor fractions, clone frequencies from SVs and SNVs, genomic features of defined clones and error rates per patient are provided as an excel spreadsheet.

Data availability:

• scWGS:

- Raw sequencing reads are available for controlled access from the NCBI Sequence Read Archive via dbGaP (https://www.ncbi.nlm.nih.gov/projects/gap/cgi-bin/study.cgi?study_id=phs002857.v3.p1).

- Processed objects will be made available via Synapse prior to publication.

• scRNA-seq:

- Raw and processed expression data are available from NCBI Gene Expression Omnibus (<https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE180661>).

- Processed objects are available from Synapse (<https://www.synapse.org/#!Synapse:syn33521743/datasets/>).

• cfDNA:

- Raw sequencing reads are available for controlled access from the NCBI Sequence Read Archive via dbGaP (https://www.ncbi.nlm.nih.gov/projects/gap/cgi-bin/study.cgi?study_id=phs002857.v3.p1).

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 focuses on high-grade serous ovarian cancer, which is inherently a female-specific disease as it affects the female reproductive system (ovaries and fallopian tubes).

Reporting on race, ethnicity, or other socially relevant groupings

No categorization variables such as race, ethnicity or other socially relevant groupings were used in this study.

Population characteristics

The study cohort (MSK SPECTRUM) includes 40 women with newly diagnosed, treatment-naïve high-grade serous ovarian cancer (HGSOC). Patients between the ages of 39 and 81 at diagnosis (median age: 59 years). 12 out of 40 cases had BRCA1 mutations (30%) and 7 out of 40 cases had a BRCA2 mutation (17%).

Recruitment

All enrolled patients were consented to an institutional biospecimen banking protocol and a protocol to perform targeted panel sequencing (MSK-IMPACT). All analyses were performed per a biospecimen research protocol. All protocols were approved by the Institutional Review Board (IRB) of Memorial Sloan Kettering Cancer Center (MSKCC). Patients were consented following the IRB-approved standard operating procedures for informed consent. Written informed consent was obtained from all patients before conducting any study-related procedures. This study was conducted in accordance with the Declaration of Helsinki and the Good Clinical Practice guidelines (GCP).

Ethics oversight

Institutional Review Board (IRB) at Memorial Sloan Kettering Cancer Center (MSKCC).

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	Quality-filtered study datasets (MSK SPECTRUM cohort): Quality-filtered study datasets (MSK SPECTRUM cohort): • scWGS: 41 patients, 70 samples, 30,260 cells • scRNA-seq: 32 patients, 125 samples, 751,300 cells • IF: 37 patients, 102 samples, 20,988,413 primary nuclei, 896,042 ruptured micronuclei • Bulk tumor WGS: 33 patients, 33 samples • Bulk normal WGS: 41 patients, 41 samples • Tumor-normal targeted panel sequencing (MSK-IMPACT): 41 patients, 41 samples • cfDNA duplex sequencing: 24 patients, 152 samples
Data exclusions	Low-quality cells were removed from scWGS and scRNA analyses. Clone frequency estimates were not performed if tumor fraction in plasma samples was < 0.0001.
Replication	Reproducibility of cfDNA assay was confirmed by assaying multiple samples from the same patient when available, for n=18 patients between 3-13 samples from different timepoints were available. We also applied patient specific probes to off-target patients to assess error rates as shown in figure 2.
Randomization	The approach of our study was to investigate in depth clonal trajectories alongside detailed clinical annotations in a small subset of patients (n=18) so randomization was not appropriate. Randomization was not appropriate as we were not treating patients based on cfDNA measurements, only observing evolution in response to treatment.
Blinding	The approach of our study was to investigate in depth clonal trajectories alongside detailed clinical annotations in a small subset of patients (n=18) so blinding was not feasible or appropriate. Data processing and analysis was performed uniformly without consideration of any patient covariates.

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

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	Not applicable.
Study protocol	The clinical study was conducted under the purview of MSKCC institutional tissue banking protocol 06-107 titled "Storage and research use of human biospecimens". Data generation and data analysis were carried out under MSKCC protocol 15-200 titled "Chemotherapy, somatic mutations, neoantigens, and the immune environment in ovarian cancer". Protocol 15-200 operates by using specimens banked under protocol 06-107. Due to MSKCC standard operating policies, internal protocol documents cannot be shared publicly but are available upon request to the IRB.
Data collection	Patients were enrolled on the study at MSKCC in New York, USA. Sample collection took place between January 2019 and December 2024.
Outcomes	Recurrence dates are defined by "progression of disease" (POD), a patient without improvement after treatment or while on maintenance therapy based on CT scan. Improvement or lack thereof is determined based on CT scan impressions (e.g. an increase in a lymph node or unchanged tumor implants). We define patients as "alive with disease" (AWD) if they have not achieved remission but have also opted out of new treatment lines and/or are on observation.

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.