nature research

Corresponding author(s): Bryan Briney, Dennis R. Burton

Reporting Summary

Nature Research 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 Research policies, see Authors & Referees and the Editorial Policy Checklist.

Statistical parameters

When statistical analyses are reported, confirm that the following items are present in the relevant location (e.g. 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
☐	☒	An indication of 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 statistics 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
☐	☒	Clearly defined error bars *State explicitly what error bars represent (e.g. SD, SE, CI)*

Our web collection on statistics for biologists may be useful.

Software and code

Policy information about availability of computer code

Data collection	N/A
Data analysis	abcloud 0.1.0, abstar 0.3.4, abtools 0.2.0, abutils 0.0.6, biopython 1.70, boto3 1.6.3, celery 4.1.0, ipython 6.2.1, jupyter-core 4.4.0, jupyterlab 0.32.1, matplotlib 2.1.2, natsort 5.2.0, numpy 1.14.2, pandas 0.22.0, paramiko 2.4.0, pymongo 3.6.1, python 3.6.4, scikit-bio 0.5.1, scipy 1.0.0, seaborn 0.8.1, weblogo 3.6.0

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/reviewers upon request. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research 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 list of figures that have associated raw data
- A description of any restrictions on data availability

Sequence data that support the findings in this study are available at the NCBI Sequencing Read Archive (www.ncbi.nlm.nih.gov/sra) under BioProject number PRJNA406949. Raw and processed datasets, as well as code for data processing and figure generation, are available at www.github.com/briney/grp_paper.

Field-specific reporting

Please select 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/authors/policies/ReportingSummary-flat.pdf

Life sciences study design

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

Sample size	Sample size was determined by estimating intra-subject repertoire differences using data from prior studies.
Data exclusions	No data were excluded
Replication	Multiple biological replicates (distinct cell aliquots) and technical replicates (duplicated processing of the same biological replicate) were analyzed for each subject.
Randomization	N/A. This study did not divide subjects into experimental groups.
Blinding	N/A. Blinding was not relevant to this study, as subjects were not divided into experimental groups.

Reporting for specific materials, systems and methods

Materials & experimental systems

n/a	Involved in the study
☒	☐ Unique biological materials
☒	☐ Antibodies
☒	☐ Eukaryotic cell lines
☒	☐ Palaeontology
☒	☐ Animals and other organisms
☐	☒ Human research participants

Methods

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

Human research participants

Policy information about studies involving human research participants

Population characteristics	All subjects were healthy adults between the ages of 18-35. The subject cohort contained an balanced mix of gender (5 males and 5 females) and contained equal numbers of self-identified Caucasian and African-American subjects. Additionally, all subjects reported no acute illness within the 14 days prior to leukapheresis.
Recruitment	Subjects were recruited by our clinical partner (HemaCare, Inc). Researchers involved in the study did not participate in subject recruitment beyond establishment of exclusion criteria. All samples were de-identified prior to delivery.