nature research

Corresponding author(s):	Daniel Segrè
Last updated by author(s):	08/15/2020

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 our Editorial Policies and the Editorial Policy Checklist.

<u> </u>				
St	·a:	רוכ	:†1	<u>ر د</u>

FOI 6	an statistical analyses, commit that the following items are present in the figure legend, table legend, main text, or Methods section.
n/a	Confirmed
\boxtimes	\Box The exact sample size (n) for each experimental group/condition, given as a discrete number and unit of measurement
\boxtimes	A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
\boxtimes	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.
\boxtimes	A description of all covariates tested
\boxtimes	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
\boxtimes	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)
\boxtimes	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>
\boxtimes	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
\boxtimes	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
\boxtimes	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 high airs contains articles on many of the points above

Software and code

Policy information about availability of computer code

Data collection

In this study we used our custom code COMETS available at https://github.com/segrelab/comets. The accompanying custom toolboxes are available at: https://github.com/segrelab/comets-toolbox

and https://github.com/segrelab/cometspy.

COMETS uses the commercial library Gurobi, under the free academic license: https://www.gurobi.com

and the following free and open source libraries: colt https://dst.lbl.gov/ACSSoftware/colt/index.html,

Commons Lang, Math and RNG libraries: https://commons.apache.org/,

 ${\sf JMatIO: https://github.com/diffplug/JMatIO,}$

Jogl: https://jogamp.org/, Junit: https://junit.org.

Data analysis

We used our custom scripts in MATLAB and Python to analyze the data. $\label{eq:matching} % \begin{subarray}{ll} \end{subarray} \begin{su$

All the scripts can be found in:

https://github.com/segrelab/comets-toolbox

and https://github.com/segrelab/cometspy.

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 Research guidelines for submitting code & software for further information.

Data

Policy information about availability of data

Clinical data

Dual use research of concern

All manuscripts must include a <u>data availability statement</u>. 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

The data in this paper was generated by computational simulations using the code that is publicly available at: https://github.com/segrelab/comets. The input files, as well as the results of the simulations are publicly available at: https://github.com/segrelab/COMETS_Protocols.

Figures 2-12 and supplementary figures and videos show the associated simulated data. There are no restrictions on the availability of the input and output simulations data, this data is distributed under the Creative Commons license. The source code is distributed under the GNU General Public License v3.0.

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. Behavioural & social sciences Ecological, evolutionary & environmental sciences For a reference copy of the document with all sections, see <u>nature.com/documents/nr-reporting-summary-flat.pdf</u> Life sciences study design All studies must disclose on these points even when the disclosure is negative. N/A Sample size Data exclusions N/A Replication N/A Randomization Blinding N/A 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 Methods Involved in the study Involved in the study n/a X Antibodies X ChIP-sea Eukaryotic cell lines Flow cytometry Palaeontology and archaeology MRI-based neuroimaging Animals and other organisms Human research participants