nature portfolio

Corresponding author(s): Joaquim Fort

Last updated by author(s): 15/7/24

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 software was used for data collection.
Data analysis	Software: Fortran 90, Origin 2017, Mathematica 9. Our Fortran 90 programs are publicly available at http://copernic.udg.edu/QuimFort/fort.htm, as mentioned in Methods and Supp. Methods.

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

Data S1-S5 contain the database with the haplogroups of all early Neolithic individuals (S1), mean genetic regional dates and locations (S2), error bars of frequencies of haplogroup K (S3), dates and locations of archaeological sites (S4), and haplogroups of HGs (S5).

Human research participants

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

Reporting on sex and gender	Neither sex nor gender were considered in study design, although for completeness and future reference Data S1 and S5 contain the sex of those ancient individuals for which it is known.
Population characteristics	The other relevant characteristics of the ancient individuals considered (Data S1 and S5) are the mitochondrial haplogroup, Y-chromosome haplogroup, location, archaeological context and sample ID(s).
Recruitment	We considered all European early farmers (Data S1) and hunter-gatherers (Data S5) for which we could find the mitochondrial haplogroup.
Ethics oversight	None

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	For the archaeological arrival times of the Neolithic at several regions, we picked up the oldest reliable date for each of the regions mentioned in the caption to Fig. 1. All of these samples are either cereal seeds or bones from humans or domesticated animals, thus highly reliable (Data S4). For the genetic analyses, we considered all European early farmers (Data S1) and hunter-gatherers (Data S5) for which we could find the mitochondrial haplogroup. Thus sample size was not pre-determined. We then grouped farmers in the geographical regions mentioned in Fig. 2. These regions (and thus the sub-sample sizes) were chosen so that we obtain reasonable error bars in Fig. 3, in the following sense. Too small regions would contain too few individuals and lead to too large error bars, and too large regions would reduce the number of error bars and thus our knowledge of the cline shape.
Data exclusions	Figure 3 includes regions that have at least 15 early farmers whose mt DNA haplogroup is known. The only regions excluded for this reason are Cyprus (2 individuals), Switzerland (1 individual) and Albania (1 individual), as recorded in Data S1. We have also ignored those regions whose populations were not involved in the spread of the Neolithic across Anatolia and Europe, regions with anomalously strong interactions between farmers and hunter-gatherers, and regions in which a substantial part of the sampled individuals may not be representative of the whole population. For a detailed discussion, please see Supplementary Methods, Sec. S1-D.
Replication	The error bars in Fig. 2 have been obtained by bootstrap re-sampling (80% confidence-level interval of 10,000 replicates).
Randomization	Not relevant because grouping was performed on the basis of geographic location, as necessary to analyze a genetic cline.
Blinding	Not relevant because grouping was performed on the basis of geographic location, as necessary to analyze a genetic cline.

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

Palaeontology and Archaeology

Specimen provenance	Data S1 and S5 include the original references that contain this information for all 961 early farmers and 418 hunter-gatherers, respectively. Data S4 includes the references that contain this information for the earliest Neolithic dates at 13 regions (Fig. 1).
Specimen deposition	Data S1 and S5 include the original references that contain this information for all 961 early farmers and 418 hunter-gatherers, respectively. Data S4 includes the references that contain this information for the earliest Neolithic dates at 13 regions (Fig. 1).
Dating methods	Data S1 and S5 include the original references that contain this information for all 961 early farmers and 418 hunter-gatherers, respectively. Data S4 includes the references that contain this information for the earliest Neolithic dates at 13 regions (Fig. 1).

☒ Tick this box to confirm that the raw and calibrated dates are available in the paper or in Supplementary Information.

Ethics oversight	None.

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