nature portfolio

Corresponding author(s): Mark Lipson, Alfredo Coppa, Ron Pinhasi, David Reich

Last updated by author(s): Dec 20, 2024

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	BWA v0.6.1, HaploGrep v2.1.1, SeqPrep v. 1.1, contamMix v1.0.1051, ANGSD v0.923, OxCal v4.4
Data analysis	ADMIXTOOLS (qpDstat v. 1152), smartpca v. 18270, qpAdm v. 2050, ALDER v. 1.0.3, DATES v. 4555, hapROH v. 1.0, ancIBD v. 1.0, GLIMPSE v. 1.1.1

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

The aligned sequences for the newly reported individuals are available through the European Nucleotide Archive under accession number PRJEB83667. Genotype data used in analysis for the newly reported individuals are available through Dataverse (doi: 10.7910/DVN/ILWB3K). Previously published data used in our analyses are available as follows: Allen Ancient DNA Resource (doi: 10.7910/DVN/FFIDCW), western Maghreb ancient DNA data (ref. [2]; European Nucleotide Archive

accession number PRJEB59008), 1000 Genomes haplotype reference panel (http://ftp.1000genomes.ebi.ac.uk/vol1/ftp/release/20130502/), human reference genome hg19 (https://www.ncbi.nlm.nih.gov/datasets/genome/GCF_000001405.13/), mitochondrial reference genome RSRS (doi: 10.1016/j.ajhg.2012.03.002), YFull YTree phylogeny (https://www.yfull.com/tree/), and ISOGG Y-chromosome SNPs (https://ybrowse.org/).

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	n/a
Reporting on race, ethnicity, or other socially relevant groupings	n/a
Population characteristics	n/a
Recruitment	n/a
Ethics oversight	n/a

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

Ecological, evolutionary & environmental sciences study design

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

Study description	Genetic analyses were performed on genomic data obtained from ancient human skeletons. Statistical models from population genetics were fit to the data to learn about the individuals' ancestral relationships to other human groups.
Research sample	New genomic data were obtained from nine ancient individuals from the eastern Maghreb (what are now Algeria and Tunisia) who lived from roughly 6000 years ago to more than 10,000 years ago. We analyzed the new data together with a large number of previously published ancient and present-day genomes from across the Mediterranean region and beyond.
Sampling strategy	We attempted to collect data from a total of 22 skeletal samples from the four archaeological sites, obtaining working data from nine unique individuals (12 samples, three from duplicate individuals). We targeted DNA sequences covering a set of approximately 1.2 million genome-wide SNPs, which effectively cover almost all independent loci in the genome (due to linkage disequilibrium) and provide good power in population history analyses.
Data collection	DNA from the ancient remains was extracted, sequenced, and processed into SNP genotype calls (final data overseen by S.M., N.R., R.P., and D.R.)
Timing and spatial scale	Ancient individuals were sampled from the sites of Afalou Bou Rhummel (Algeria), Djebba, Doukanet el Khoutifa, and Hergla (all Tunisia).
Data exclusions	Ten of the skeletal samples did not yield working data as assessed by pre-established ancient DNA quality criteria. One individual yielded working data but with evidence of contamination, so we restricted to sequences containing damage patterns indicative of an authentic ancient origin.
Reproducibility	All attempts to reproduce were successful.
Randomization	Individuals were grouped by site, and in one case (Doukanet el Khoutifa) into two subgroups based on genome-wide ancestry.
Blinding	Analyses were performed both on group level and individual level; other sample-specific features were not relevant to results.

Did the study involve field work? ☐ Yes ☒ No

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

Palaeontology and Archaeology

Specimen provenance	Human remains from Djebba, Tunisia, were excavated under authorization (2018; directed by co-author NA). Human remains from Doukanet el Khoutifa and Hergla were collected within the framework of the Northern Tunisia Archaeological Project (NOTAP), under two scientific cooperation agreements: one between INP, CNR-ISPC, and ISMEO (2021-2024; co-direction by co-authors NA, LB, AC, and GL), and another between INP and the University of Bologna (2002-2007; co-direction: Simone Mulazzani and Ridha Boussoffara). We are grateful to the current and former INP General Directors, Tarek Baccouche and Faouzi Mahfoudh, for their support, and Adelaide Marsilio for his key role in the excavation of Doukanet el Khoutifa in 2022.
Specimen deposition	The specimens are available from the permitting authorities, and samples will not be retained at the ancient DNA laboratories.
Dating methods	We sent samples of the same skeletal elements used for ancient DNA analysis for accelerator mass spectrometry (AMS) radiocarbon dating at the Pennsylvania State University Radiocarbon Laboratory, using standard methods. We calibrated the dates using OxCal (v4.4), and the IntCal20 calibration curve.

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

Ethics oversight	This study followed the ethical protocol described in Alpaslan-Roodenberg et al. 2022 Ethics of DNA research on human remains: five globally applicable guidelines (Nature 599, 41-6), and its ethical approach was additionally reviewed as part of the requests that led to the sampling of skeletal remains from the four sites reported on here.

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

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, mosiacism, off-target gene editing) were examined.*