nature portfolio

Corresponding author(s): Christian Dina
Richard Redon

Last updated by author(s): Jan 31, 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	For PREGO data, we used Analysis Power Tools (APT https://www.thermofisher.com/fr/fr/home/life-science/microarray-analysis/microarray-analysis-partners-programs/affymetrix-developers-network/affymetrix-power-tools.html), a tool dedicated to Axiom arrays to call genotyping. We used SNPolisher http://tools.thermofisher.com to apply array specific QC. For FranceGenRef whome genome sequence determination was performed using : GATK 3.8 https://software.broadinstitute.org/gatk/best-practices PLINK1.90 https://www.cog-genomics.org/plink/ vcftools v0.1.12a https://vcftools.github.io/man_0112a.html bcftools https://samtools.github.io/bcftools/bcftools.html
Data analysis	PLINK1.90 https://www.cog-genomics.org/plink/ ADMIXTURE vs1.3 https://dalexander.github.io/admixture/index.html smartpca (EIG-6.1.4) https://alkesgroup.broadinstitute.org/EIGENSOFT/OLD fineSTRUCTURE vs.2, CHROMOPAINTER & GLOBETROTTER https://people.maths.bris.ac.uk/~madjl/finestructure/finestructure_info.html SNPolisher http://tools.thermofisher.com SHAPEIT v2.r790 https://mathgen.stats.ox.ac.uk/genetics_software/shapeit/shapeit.html RefinedIBD https://faculty.washington.edu/browning/refined-ibd.html IBDNe https://faculty.washington.edu/browning/ibdne.html Admixtools https://reich.hms.harvard.edu/software bwa-aln version 0.7.17 https://bio-bwa.sourceforge.net/bwa.shtml

ANGSD http://www.popgen.dk/angsd/index.php/Genotype_calling
schmutzi https://grenaud.github.io/schmutzi/
ARGON https://github.com/pierpal/ARGON
R statistical package https://www.r-project.org

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

We obtained public databases representing European (and world) populations. The ancient DNA dataset (Human Origins Array (HOA) dataset V42.4) is available at https://reich.hms.harvard.edu/allen-ancient-dna-resource-aadr-downloadable-genotypes-present-day-and-ancient-dna-data.
The new WGS data on ancient DNA (bam files) are available from the European Nucleotide Archive (ENA), under the accession number PRJEB71835. The HOA SNP genotypes for the Viking samples are now available in the version 54.1 of the Allen Ancient DNA Resource (https://reich.hms.harvard.edu/allen-ancient-dna-resource-aadr-downloadable-genotypes-present-day-and-ancient-dna-data). Genetic data on contemporary human individuals are subject to the French regulation on the protection of identifiable personal data. WGS data from the FranceGenRef panel will therefore be submitted to the French Centralised Data Centre of the France Medicine Genomic Plan, which is currently under construction. Requests for the use of these data can be sent to richard.redon@inserm.fr. Access to the PREGO array genotyping data will be organised on request under the terms of a data access agreement as described at https://umr1087.univ-nantes.fr/prego-biobank.
The HGDP dataset is available at https://www.hagsc.org/hgdp/files.html. Finally, the EGAD00000000120 from the The International Multiple Sclerosis Genetics Consortium and the Wellcome Trust Case Control Consortium 2 and the EGAD00010000124 from the Genetic Analysis of Psoriasis Consortium & the Wellcome Trust Case Control Consortium 2, and the EGAD00010000632 from the Peopling of the British Isles are available at EGA repository.

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	No sex-specific filters nor analysis have were done in our study. We focused on autosomal regions of the genome.
Reporting on race, ethnicity, or other socially relevant groupings	We used the criterion of birth place, either of each individual or of the grand parents of individuals in order to assign each participant to an administrative or historical region, city or county. Individuals from external datasets were assigned the reported country (MS/PS/HGDP) or county within United Kingdom (PoBI).
Population characteristics	Individuals newly presented in this study are sampled from a general volunteer population. Gender is well balanced and age mean is 52. No further phenotype information was used/available.
Recruitment	Project PREGO ("Population de référence du Grand Ouest", www.vacarme-project.org) collected the DNA of 5,707 healthy persons originating from western France (Pays de la Loire and Brittany regions). Individuals were recruited during 295 blood drives organised by the French Blood Service carried out between February 2014 and March 2017. Priority was given to blood drives taking place in rural areas. Only individuals whose four grandparents were born in western France and preferably within a radius of 30 km were included in this study. The FranceGenRef's individuals were sampled based on the birthplace of their grandparents, whose distance should not exceed 30 kilometres. : 50 blood donors sampled in the département of Finistère, 354 blood donors from the PREGO cohort (www.vacarme-project.org) and finally 458 individuals from the GAZEL cohort (www.gazel.inserm.fr/en), among which are individuals from five other regions of France: Normandie, Hauts-de-France, Grand East, Centre-Val de Loire and Nouvelle-Aquitaine. Our goal was to have a gene pool representative of the French population of the early/mid 20th century.
Ethics oversight	The PREGO study ("Population de référence du Grand Ouest", https://umr1087.univ-nantes.fr/prego-biobank) collected DNA from 5,707 healthy persons originating from western France (Regions Pays de la Loire and Bretagne). As pointed out in the manuscript now, the PREGO study received approvals from the local ethical committee of Nantes (Comité de Protection des Personnes), the Advisory Committee on Information Processing for Health Research (CCTIRS) and the National Commission on Informatics and Liberty (CNIL). Participants signed a written informed consent for participation in the study, inclusion in bio-resource and personal data processing. The FranceGenRef study includes 354 blood donors from the aforementioned PREGO cohort (described above) with origin in the départements of Côtes d'Armor (COT), Ille-et-Vilaine (ILL), Morbihan (MOR), Loire-Atlantique (LOI), Maine-et-Loire (MAI), Mayenne (MAY), Sarthe (SAR) and Vendée (VEN); 50 blood donors from Finistère (FIN); 458 individuals from the GAZEL cohort (www.gazel.inserm.fr/en). The GAZEL study received approvals from the National Commission for Data Processing and Freedoms (CNIL), the National Medical Council and the National Consultative Committee of Ethics. All individuals signed informed consent for genetic studies at the time of enrolment.

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	Individuals were ascertained based on the birth place of their grand-parents in order to have a representation of the genomic landscape of rural France and a set of maximally unrelated individuals as described in the methods giving us a total of 3,234 samples.
Data exclusions	Genetically related individuals were identified and excluded from the population genetics analyses as described in the methods as well as genetic outliers identified via principal component analysis.
Replication	General trends found in our study have been reported in previous studies on the French human population and the distribution of our samples is consistent with that of French samples present in the external datasets included in our study.
Randomization	Part of the analyses have been performed over the whole sample set. However, some analyses have been performed on a random subset of samples in order to keep sample sizes homogeneous across regions as variations can lead to bias in population genetic analysis.
Blinding	Blinding is not required for our study as we are not testing the effect of an administrated drug on different groups of individuals. We are investigating whether genetic relationships (which depends solely on individuals' genomic content) reflect geography (which is assumed to reflect the grand-parents "average"). In our opinion, there is therefore no way of unintentionally influencing the data upstream of the analysis.

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	The archaeological excavations in Saint-Lupien, Rezé (south shore of the Loire) took place between 2005 and 2016 and were led by the team of Mikaël Rouzic under the request of the city of Rezé. The site located in Chaussé Saint-Pierre, Angers (north shore of the Loire), was excavated by the team of Martin Pithon, from the French Institute for Preventive Archaeological Research (INRAP) and the excavations occurred between July and August 2009. Based on the archaeological remains and radiocarbon dates, the site shows evidence of occupation from the beginning of the Roman Empire to modern times. The project was initialised under the request of the city of Angers given the plan for public construction affecting the archaeological site. The archaeological study of the site was authorised by the Regional Division for Cultural Affairs (Délégation Régional des affaires culturelles) and the INRAP. Finally, the excavations in the archaeological site in Chéméré (south shore of the Loire) started in the 60s but the two individuals sequenced in this study belong to a group of 181 individuals found in the last excavations in 2007. These excavations were led by an INRAP archaeological team under a project of preventive archaeology before construction of a pavilion.
Specimen deposition	The specimens were deposited in INRAP (Institut National de Recherches Archéologiques Préventives) collections.
Dating methods	Calibrated radiocarbon dating was carried out for 3 pf the samples and the dating image is reporte in Supplementary data. When no radiocarbon datation was available, the approximate dates are based on the archaeological context described by the archaeologists of the INRAP. The context is reported in the Supplmentary data aprt dedicated to archeology.

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

Ethics oversight	The ancient DNA study is set up in the frame of the GHOST project, led by the Institut National de Recherches Archéologiques. Préventives ((https://www.inrap.fr/). This institute is granted by the French governement for all archeological studies management. The study's ID is @GIR - R110583

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