

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 (<i>n</i>) 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 <i>Only common tests should be described solely by name; describe more complex techniques in the Methods section.</i>
☐	☒ 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. <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>
☐	☒ 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 <i>d</i> , Pearson's <i>r</i>), 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	We did not use any software for Data collection.
Data analysis	We used publicly available software for the data analysis (PLINK 1.9 and 2.0, GCTA(1.93.3beta2), Eagle version2, Minimac3 (2.0.1), AdmixTools (version 7.0.2), PRS-CS (version Jun 4, 2021), SHAPEIT (4.2.1), Minimac4 (1.0.1), METAL (version 2020-05-05), R (4.1.0). The software is described in the Methods section of the manuscript.

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](#)

Genotype data of the Biobank Japan Project (BBJ) are available at the NBDC Human Database with the research ID hum0014 and hum0311. GWAS data of the Nagahama cohort study are available at the NBDC Human Database with the research ID hum0012. The analysis of UK Biobank (UKB) GWAS data was conducted via

the application number 47821. All ancient genomic data were compiled in the previous study (Cooke et al. Sci Adv. 2021). The summary statistics from the genome-wide association analysis for Jomon ancestry have been deposited in the National Bioscience Database Center (NBDC) Human Database (<https://humandbs.dbcls.jp/en/>) under accession code hum0197 (<https://humandbs.dbcls.jp/en/hum0197-latest>).

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	Since biological sex influences phenotypic variations, we used biological sex as covariates in our study. We obtained sex information for BBJ participants from their medical records. We obtained sex information for UKBB participants from their self-reported information. We listed the detailed sex information for BBJ participants in Supplementary Table 1.
Reporting on race, ethnicity, or other socially relevant groupings	For the participants in the Biobank Japan and the Nagahama cohort study, we determined their genetic ancestry based on PCA from their genomic information along with 1000 Genomes Project. Furthermore, we classified the participants into the subpopulations based on the registered location information or genetic PCA information. For the UK Biobank participants, we defined the East Asian individuals based on the genetic PCA. Additionally, we divided the UKB EAS participants into subpopulations based on the self-reported ethnicity provided by UK Biobank. Details are described in Method section.
Population characteristics	BBJ is a hospital-based cohort, and participants have the diagnosis of at least one of 47 common diseases. The Nagahama cohort study is a community-based study and recruited healthy individuals from Nagahama City, Shiga Prefecture, Japan. UKB is a population-based cohort, enrolling healthy volunteers aged between 40 and 69 years.
Recruitment	BBJ collaboratively collected DNA, serum samples, and clinical information from more than 250,000 participants across 12 medical institutions in Japan. Participants in the first cohort were recruited between 2003 and 2008, while those in the second cohort were recruited between 2013 and 2018. The Nagahama cohort study has recruited participants since 2007. UKB is a population-based cohort study that recruited approximately 500,000 individuals from 2006 to 2010 from across the United Kingdom.
Ethics oversight	All the participants provided written informed consent approved from ethic committee of the Institute of Medical Sciences, the University of Tokyo. This study was approved by the ethical committees of Osaka University Graduate School of Medicine and Graduate School of Medicine, the University of Tokyo.

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	BBJ is a biobank cohort that collaboratively collected DNA and serum samples from 12 medical institutions in Japan and recruited more than 250,000 participants. After the stringent quality control, we used 171,287 participants from the first cohort of BBJ (age ≥ 18, female = 78,319). As a validation study, we used 72,695 individuals from the second cohort of BBJ, which is an independent cohort from the first BBJ cohort. We also used 2,993 individuals from the Nagahama cohort. As a replication study, we extracted 2,286 East Asian individuals from UK Biobank using PCA. We used all available ancient Japanese and continental Asia genomes (n = 22). Sample size calculation was not carried out, but we used maximum samples in BBJ, whose genotype and phenotype data were available.
Data exclusions	In BBJ, we excluded individuals with the age under 18, low genotype call rate (< 99%), closely related (GRM ≥ 0.178), and genetic ancestry other than Japanese (based on PCA plot) for quality control as described in Sakaue and Kanai et al. Nat Genet. 2021.
Replication	We performed the validation analysis of tripartite ancestral model and predictive power of Jomon-related variants using two independent Japanese cohorts, the Nagahama cohort study (n = 2,993) and the second cohort of BioBank Japan (n = 72,695). We also replicated the association of Jomon ancestry with the increasing risk of BMI using East Asian individuals from the UK Biobank (n = 2286).
Randomization	In leave-one-group-out method for polygenic prediction in BMI, we randomly split the entire participants of BBJ into 5 groups.
Blinding	We did not apply blinding of the samples because no intervention was conducted in our study.

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 |