

## 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  |
| <input type="checkbox"/>            | <input checked="" type="checkbox"/> The exact sample size ( $n$ ) for each experimental group/condition, given as a discrete number and unit of measurement  |
| <input type="checkbox"/>            | <input checked="" type="checkbox"/> A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly  |
| <input type="checkbox"/>            | <input checked="" type="checkbox"/> The statistical test(s) used AND whether they are one- or two-sided<br><i>Only common tests should be described solely by name; describe more complex techniques in the Methods section.</i>   |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> A description of all covariates tested  |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons   |
| <input type="checkbox"/>            | <input checked="" type="checkbox"/> 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) |
| <input type="checkbox"/>            | <input checked="" type="checkbox"/> For null hypothesis testing, the test statistic (e.g. $F$ , $t$ , $r$ ) with confidence intervals, effect sizes, degrees of freedom and $P$ value noted<br><i>Give <math>P</math> values as exact values whenever suitable.</i>                            |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings  |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes  |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> 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 | MGISEQ-2000 platform   |
| Data analysis   | TrimGalore (v0.6.1), Burrows-Wheeler Aligner (BWA v0.7.17-r1188), Genome Analysis Toolkit (GATK, v4.0.12.0)88 and Samtools (v1.9) for SNP detection. Whatshap (v0.18) and SHAPEIT4 (v4.2.2) for reference panel construction. Minimac4 for imputation test. GCTA (v1.93.2), MEGA (v11), emmax (beta-07Mar2010), ADMIXTURE (v1.3.0) and VCFtools (v0.1.17) for population structure analyses. ADMIXTTTOOLS (v7.0.1), DFOIL, Treemix (v1.13) and ALDER (v1.03) for admixture analyses. MitoZ (v2.4), MitoToolPy (v1.0) and PopART (v1.7) for whole-mitochondrial and paternal analyses. PLINK (v1.90) and GEMMA (v0.98.5) for GWAS analysis. Delly2 (v0.8.1), Lumpy (v0.3.1), Manta (v1.6.0) and Wham (v1.7.0) for SV detection. |

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 raw sequencing datasets generated in this study have been deposited in NGDC/GSA under accession code PRJCA030193 (<https://ngdc.cncb.ac.cn/gsa>). The reference panel and genomic variants datasets have been deposited to the Figshare (<https://doi.org/10.6084/m9.figshare.25847761.v1>).*

## 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  | <i>not applicable</i> |
| Reporting on race, ethnicity, or other socially relevant groupings | <i>not applicable</i> |
| Population characteristics   | <i>not applicable</i> |
| Recruitment  | <i>not applicable</i> |
| Ethics oversight   | <i>not applicable</i> |

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](https://www.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     | <i>1,011 individuals from 50 representative pig populations were used. These breeds were selected based on their phylogenetic relationship and represented genetic diversity across Eurasian pigs.</i> |
| Data exclusions | <i>No samples were excluded in this study. Filters applied to eliminate low-quality genetic variants and linkage disequilibrium pruning were described in the Methods section.</i>                     |
| Replication     | <i>not applicable</i>  |
| Randomization   | <i>The samples in this study were randomized.</i>  |
| Blinding        | <i>Not applicable. For an animal genome and population study, blinding is not necessary.</i>   |

## Behavioural & social sciences study design

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

|                   |   |
|-------------------|---|
| Study description | <i>Briefly describe the study type including whether data are quantitative, qualitative, or mixed-methods (e.g. qualitative cross-sectional, quantitative experimental, mixed-methods case study).</i>  |
| Research sample   | <i>State the research sample (e.g. Harvard university undergraduates, villagers in rural India) and provide relevant demographic information (e.g. age, sex) and indicate whether the sample is representative. Provide a rationale for the study sample chosen. For studies involving existing datasets, please describe the dataset and source.</i> |

|                   |  |
|-------------------|--|
| Sampling strategy | <i>Describe the sampling procedure (e.g. random, snowball, stratified, convenience). Describe the statistical methods that were used to predetermine sample size OR if no sample-size calculation was performed, describe how sample sizes were chosen and provide a rationale for why these sample sizes are sufficient. For qualitative data, please indicate whether data saturation was considered, and what criteria were used to decide that no further sampling was needed.</i> |
| Data collection   | <i>Provide details about the data collection procedure, including the instruments or devices used to record the data (e.g. pen and paper, computer, eye tracker, video or audio equipment) whether anyone was present besides the participant(s) and the researcher, and whether the researcher was blind to experimental condition and/or the study hypothesis during data collection.</i>  |
| Timing            | <i>Indicate the start and stop dates of data collection. If there is a gap between collection periods, state the dates for each sample cohort.</i>   |
| Data exclusions   | <i>If no data were excluded from the analyses, state so OR if data were excluded, provide the exact number of exclusions and the rationale behind them, indicating whether exclusion criteria were pre-established.</i>  |
| Non-participation | <i>State how many participants dropped out/declined participation and the reason(s) given OR provide response rate OR state that no participants dropped out/declined participation.</i>   |
| Randomization     | <i>If participants were not allocated into experimental groups, state so OR describe how participants were allocated to groups, and if allocation was not random, describe how covariates were controlled.</i>   |

## Ecological, evolutionary & environmental sciences study design

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

|                          |   |
|--------------------------|---|
| Study description        | <i>Briefly describe the study. For quantitative data include treatment factors and interactions, design structure (e.g. factorial, nested, hierarchical), nature and number of experimental units and replicates.</i>   |
| Research sample          | <i>Describe the research sample (e.g. a group of tagged <i>Passer domesticus</i>, all <i>Stenocereus thurberi</i> within Organ Pipe Cactus National Monument), and provide a rationale for the sample choice. When relevant, describe the organism taxa, source, sex, age range and any manipulations. State what population the sample is meant to represent when applicable. For studies involving existing datasets, describe the data and its source.</i> |
| Sampling strategy        | <i>Note the sampling procedure. Describe the statistical methods that were used to predetermine sample size OR if no sample-size calculation was performed, describe how sample sizes were chosen and provide a rationale for why these sample sizes are sufficient.</i>  |
| Data collection          | <i>Describe the data collection procedure, including who recorded the data and how.</i>   |
| Timing and spatial scale | <i>Indicate the start and stop dates of data collection, noting the frequency and periodicity of sampling and providing a rationale for these choices. If there is a gap between collection periods, state the dates for each sample cohort. Specify the spatial scale from which the data are taken</i>  |
| Data exclusions          | <i>If no data were excluded from the analyses, state so OR if data were excluded, describe the exclusions and the rationale behind them, indicating whether exclusion criteria were pre-established.</i>  |
| Reproducibility          | <i>Describe the measures taken to verify the reproducibility of experimental findings. For each experiment, note whether any attempts to repeat the experiment failed OR state that all attempts to repeat the experiment were successful.</i>  |
| Randomization            | <i>Describe how samples/organisms/participants were allocated into groups. If allocation was not random, describe how covariates were controlled. If this is not relevant to your study, explain why.</i>   |
| Blinding                 | <i>Describe the extent of blinding used during data acquisition and analysis. If blinding was not possible, describe why OR explain why blinding was not relevant to your study.</i>  |

Did the study involve field work?  Yes  No

## Field work, collection and transport

|                        |   |
|------------------------|---|
| Field conditions       | <i>Describe the study conditions for field work, providing relevant parameters (e.g. temperature, rainfall).</i>  |
| Location               | <i>State the location of the sampling or experiment, providing relevant parameters (e.g. latitude and longitude, elevation, water depth).</i>   |
| Access & import/export | <i>Describe the efforts you have made to access habitats and to collect and import/export your samples in a responsible manner and in compliance with local, national and international laws, noting any permits that were obtained (give the name of the issuing authority, the date of issue, and any identifying information).</i> |
| Disturbance            | <i>Describe any disturbance caused by the study and how it was minimized.</i>   |

# 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   |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Antibodies                             |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Eukaryotic cell lines                  |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Palaeontology and archaeology          |
| <input type="checkbox"/>            | <input checked="" type="checkbox"/> Animals and other organisms |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Clinical data                          |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Dual use research of concern           |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Plants                                 |

## Methods

- |                                     |   |
|-------------------------------------|---|
| n/a                                 | Involved in the study                           |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> ChIP-seq               |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Flow cytometry         |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> MRI-based neuroimaging |

## Antibodies

Antibodies used

Validation

## Eukaryotic cell lines

Policy information about [cell lines and Sex and Gender in Research](#)

Cell line source(s)

Authentication

Mycoplasma contamination

Commonly misidentified lines (See [ICLAC](#) register)

## Palaeontology and Archaeology

Specimen provenance

Specimen deposition

Dating methods

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

Ethics oversight

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

## Animals and other research organisms

Policy information about [studies involving animals](#); [ARRIVE guidelines](#) recommended for reporting animal research, and [Sex and Gender in Research](#)

Laboratory animals

|                         |  |
|-------------------------|--|
| Wild animals            | not applicable   |
| Reporting on sex        | The main purpose of this study is to construct a reference panel and analyze the genome architecture of Chinese domestic pigs. The sex-biased problem for constructing the reference pane was considered in our study, and we reported the sex information of each sample in Supplementary Data 1. Except for the haplogroup analysis of the Y chromosome, other analyses had little influence by sex. |
| Field-collected samples | All experimental pigs were sampled under standard procedures within 30 min at room temperature.  |
| Ethics oversight        | This study was conducted in strict accordance with the protocol approved by the Institutional Animal Care and Use Committee (IACUC) of China Agricultural University (Beijing, People's Republic of China; Approval No. AW60604202-1-1).   |

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

## Clinical data

Policy information about [clinical studies](#)

All manuscripts should comply with the ICMJE [guidelines for publication of clinical research](#) and a completed [CONSORT checklist](#) must be included with all submissions.

|                             |  |
|-----------------------------|--|
| Clinical trial registration | <i>Provide the trial registration number from ClinicalTrials.gov or an equivalent agency.</i>                            |
| Study protocol              | <i>Note where the full trial protocol can be accessed OR if not available, explain why.</i>                              |
| Data collection             | <i>Describe the settings and locales of data collection, noting the time periods of recruitment and data collection.</i> |
| Outcomes                    | <i>Describe how you pre-defined primary and secondary outcome measures and how you assessed these measures.</i>          |

## Dual use research of concern

Policy information about [dual use research of concern](#)

### Hazards

Could the accidental, deliberate or reckless misuse of agents or technologies generated in the work, or the application of information presented in the manuscript, pose a threat to:

- | No                       | Yes                      |                            |
|--------------------------|--------------------------|----------------------------|
| <input type="checkbox"/> | <input type="checkbox"/> | Public health              |
| <input type="checkbox"/> | <input type="checkbox"/> | National security          |
| <input type="checkbox"/> | <input type="checkbox"/> | Crops and/or livestock     |
| <input type="checkbox"/> | <input type="checkbox"/> | Ecosystems                 |
| <input type="checkbox"/> | <input type="checkbox"/> | Any other significant area |

### Experiments of concern

Does the work involve any of these experiments of concern:

- | No                       | Yes                      |   |
|--------------------------|--------------------------|---|
| <input type="checkbox"/> | <input type="checkbox"/> | Demonstrate how to render a vaccine ineffective                             |
| <input type="checkbox"/> | <input type="checkbox"/> | Confer resistance to therapeutically useful antibiotics or antiviral agents |
| <input type="checkbox"/> | <input type="checkbox"/> | Enhance the virulence of a pathogen or render a nonpathogen virulent        |
| <input type="checkbox"/> | <input type="checkbox"/> | Increase transmissibility of a pathogen                                     |
| <input type="checkbox"/> | <input type="checkbox"/> | Alter the host range of a pathogen  |
| <input type="checkbox"/> | <input type="checkbox"/> | Enable evasion of diagnostic/detection modalities                           |
| <input type="checkbox"/> | <input type="checkbox"/> | Enable the weaponization of a biological agent or toxin                     |
| <input type="checkbox"/> | <input type="checkbox"/> | Any other potentially harmful combination of experiments and agents         |

## 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, mosaicism, off-target gene editing) were examined.   |

## ChIP-seq

### Data deposition

- Confirm that both raw and final processed data have been deposited in a public database such as [GEO](#).
- Confirm that you have deposited or provided access to graph files (e.g. BED files) for the called peaks.

|  |   |
|--|---|
| Data access links<br><i>May remain private before publication.</i> | For "Initial submission" or "Revised version" documents, provide reviewer access links. For your "Final submission" document, provide a link to the deposited data.   |
| Files in database submission                                       | Provide a list of all files available in the database submission.   |
| Genome browser session<br>(e.g. <a href="#">UCSC</a> )             | Provide a link to an anonymized genome browser session for "Initial submission" and "Revised version" documents only, to enable peer review. Write "no longer applicable" for "Final submission" documents. |

### Methodology

|                         |   |
|-------------------------|---|
| Replicates              | Describe the experimental replicates, specifying number, type and replicate agreement.  |
| Sequencing depth        | Describe the sequencing depth for each experiment, providing the total number of reads, uniquely mapped reads, length of reads and whether they were paired- or single-end. |
| Antibodies              | Describe the antibodies used for the ChIP-seq experiments; as applicable, provide supplier name, catalog number, clone name, and lot number.                                |
| Peak calling parameters | Specify the command line program and parameters used for read mapping and peak calling, including the ChIP, control and index files used.                                   |
| Data quality            | Describe the methods used to ensure data quality in full detail, including how many peaks are at FDR 5% and above 5-fold enrichment.  |
| Software                | Describe the software used to collect and analyze the ChIP-seq data. For custom code that has been deposited into a community repository, provide accession details.        |

## Flow Cytometry

### Plots

Confirm that:

- The axis labels state the marker and fluorochrome used (e.g. CD4-FITC).
- The axis scales are clearly visible. Include numbers along axes only for bottom left plot of group (a 'group' is an analysis of identical markers).
- All plots are contour plots with outliers or pseudocolor plots.
- A numerical value for number of cells or percentage (with statistics) is provided.

### Methodology

|                    |  |
|--------------------|--|
| Sample preparation | Describe the sample preparation, detailing the biological source of the cells and any tissue processing steps used.  |
| Instrument         | Identify the instrument used for data collection, specifying make and model number.  |
| Software           | Describe the software used to collect and analyze the flow cytometry data. For custom code that has been deposited into a community repository, provide accession details. |

Cell population abundance

Describe the abundance of the relevant cell populations within post-sort fractions, providing details on the purity of the samples and how it was determined.

Gating strategy

Describe the gating strategy used for all relevant experiments, specifying the preliminary FSC/SSC gates of the starting cell population, indicating where boundaries between "positive" and "negative" staining cell populations are defined.

Tick this box to confirm that a figure exemplifying the gating strategy is provided in the Supplementary Information.

## Magnetic resonance imaging

### Experimental design

Design type

Indicate task or resting state; event-related or block design.

Design specifications

Specify the number of blocks, trials or experimental units per session and/or subject, and specify the length of each trial or block (if trials are blocked) and interval between trials.

Behavioral performance measures

State number and/or type of variables recorded (e.g. correct button press, response time) and what statistics were used to establish that the subjects were performing the task as expected (e.g. mean, range, and/or standard deviation across subjects).

### Acquisition

Imaging type(s)

Specify: functional, structural, diffusion, perfusion.

Field strength

Specify in Tesla

Sequence &amp; imaging parameters

Specify the pulse sequence type (gradient echo, spin echo, etc.), imaging type (EPI, spiral, etc.), field of view, matrix size, slice thickness, orientation and TE/TR/flip angle.

Area of acquisition

State whether a whole brain scan was used OR define the area of acquisition, describing how the region was determined.

Diffusion MRI

 Used

 Not used

### Preprocessing

Preprocessing software

Provide detail on software version and revision number and on specific parameters (model/functions, brain extraction, segmentation, smoothing kernel size, etc.).

Normalization

If data were normalized/standardized, describe the approach(es): specify linear or non-linear and define image types used for transformation OR indicate that data were not normalized and explain rationale for lack of normalization.

Normalization template

Describe the template used for normalization/transformation, specifying subject space or group standardized space (e.g. original Talairach, MNI305, ICBM152) OR indicate that the data were not normalized.

Noise and artifact removal

Describe your procedure(s) for artifact and structured noise removal, specifying motion parameters, tissue signals and physiological signals (heart rate, respiration).

Volume censoring

Define your software and/or method and criteria for volume censoring, and state the extent of such censoring.

### Statistical modeling & inference

Model type and settings

Specify type (mass univariate, multivariate, RSA, predictive, etc.) and describe essential details of the model at the first and second levels (e.g. fixed, random or mixed effects; drift or auto-correlation).

Effect(s) tested

Define precise effect in terms of the task or stimulus conditions instead of psychological concepts and indicate whether ANOVA or factorial designs were used.

Specify type of analysis:

Whole brain

ROI-based

Both

Statistic type for inference

Specify voxel-wise or cluster-wise and report all relevant parameters for cluster-wise methods.

(See [Eklund et al. 2016](#))

Correction

Describe the type of correction and how it is obtained for multiple comparisons (e.g. FWE, FDR, permutation or Monte Carlo).

## Models & analysis

- n/a | Involved in the study
- Functional and/or effective connectivity
  - Graph analysis
  - Multivariate modeling or predictive analysis

Functional and/or effective connectivity

*Report the measures of dependence used and the model details (e.g. Pearson correlation, partial correlation, mutual information).*

Graph analysis

*Report the dependent variable and connectivity measure, specifying weighted graph or binarized graph, subject- or group-level, and the global and/or node summaries used (e.g. clustering coefficient, efficiency, etc.).*

Multivariate modeling and predictive analysis

*Specify independent variables, features extraction and dimension reduction, model, training and evaluation metrics.*