

## 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

Data analysis

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 data for the PacBio HiFi reads, RNA sequencing reads, and re-sequencing Illumina short reads have been deposited in the Genome Sequence Archive (GSA)75 database at the National Genomics Data Center, Beijing Institute of Genomics, Chinese Academy of Sciences / China National Center for Bioinformation under BioProject PRJCA012695. The genome assembly, genome annotation, pan-TE library, graph pan-genome, gene family and gene presence/absence matrices files have been deposited in the Figshare database ([https://figshare.com/articles/dataset/32\\_ecotypes\\_Arabidopsis\\_thaliana\\_genomes\\_gene\\_annotation\\_pan-TE\\_library\\_graph\\_pan-genome\\_gene\\_family\\_and\\_gene\\_presence\\_absence\\_matrices\\_files\\_/21673895](https://figshare.com/articles/dataset/32_ecotypes_Arabidopsis_thaliana_genomes_gene_annotation_pan-TE_library_graph_pan-genome_gene_family_and_gene_presence_absence_matrices_files_/21673895)). Public RNA-seq data were downloaded from the NCBI SRA database under BioProject PRJNA187928, PRJEB15161, and PRJNA319904. 1,135 individuals re-sequencing data were downloaded from PRJNA273563. 19 BIOCLIM and SRTM elevation data used in this study were download from WorldClim v2.1 ([www.worldclim.org](http://www.worldclim.org)). The global UV-B radiation data was download from <https://www.ufz.de/gluv>. Source data are provided with this paper.

## 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	<input type="text" value="No human research in this study."/>
Reporting on race, ethnicity, or other socially relevant groupings	<input type="text" value="No human research in this study."/>
Population characteristics	<input type="text" value="No human research in this study."/>
Recruitment	<input type="text" value="No human research in this study."/>
Ethics oversight	<input type="text" value="No human research in this study."/>

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://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	<input type="text" value="A total representative 32 individuals including 32 ecotypes of Arabidopsis thaliana were sequenced by HiFi long-reads sequencing technology."/>
Data exclusions	<input type="text" value="During the function annotation and following comparative analysis, we chose the longest transcript to represent each gene. As our article does not study on alternative splicing."/>
Replication	<input type="text" value="The genome sequences were taken and sequenced with more than 15-60 fold coverage. And the RT-qPCR and Dual-LUC activity assays experiments are independently performed three times. All attempts at replication were successful."/>
Randomization	<input type="text" value="In this study, all samples were collected based on their phenotypic characteristics and kinships, excluding closely related individuals."/>
Blinding	<input type="text" value="Blinding is not applicable in our study because it does not involve subjects which receive different treatments."/>

## 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 &amp; experimental systems

## Methods

n/a	Involvement
<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 checked="" type="checkbox"/>	<input 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 type="checkbox"/>	<input checked="" type="checkbox"/> Plants

n/a	Involvement
<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

## Plants

Seed stocks

Seed materials of AH-7 ecotype are collected from the wild. The other 31 seed materials are collected from Arabidopsis Biological Resource Center (ABRC). Supplementary Table 1 provides descriptions of the names, sampling locations, and CS number for the 32 ecotypes.

Novel plant genotypes

No novel plant genotypes were produced.

Authentication

N/A