

Corresponding author(s): Fei He, Hong-Qing Ling

Last updated by author(s): Oct 21, 2025

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

no software was used

Data analysis

Salmon version 1.8.0
 OrthoFinder version 2.5.4
 DESeq2
 TBtools
 SnpEff version 5.0e
 InterProScan version 5.66-98.0
 GCTA version 1.94.1
 transcriptome-wide association studies (TWAS) (<http://gusevlab.org/projects/fusion/>)
 VCFtools version 0.1.16
 Python-based composite likelihood approach (XP-CLR) (<https://github.com/hardingnj/xpclr>)
 Gephi (<https://gephi.org/>)
 PLINK version 1.9
 R version 4.3.1
 JCVI

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 sequence data were deposited in Genome Sequence Archive (<https://bigd.big.ac.cn/gsa>) under accession number CRA022106 and CRA022107. The results, including RNA-seq quantification, the pan-gene atlas, orthologous genes and eQTLs, have been uploaded to the Figshare database (<https://figshare.com/s/0e907a409ac93a4b4fea>)

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

Use the terms sex (biological attribute) and gender (shaped by social and cultural circumstances) carefully in order to avoid confusing both terms. Indicate if findings apply to only one sex or gender; describe whether sex and gender were considered in study design; whether sex and/or gender was determined based on self-reporting or assigned and methods used.

Provide in the source data disaggregated sex and gender data, where this information has been collected, and if consent has been obtained for sharing of individual-level data; provide overall numbers in this Reporting Summary. Please state if this information has not been collected.

Report sex- and gender-based analyses where performed, justify reasons for lack of sex- and gender-based analysis.

Reporting on race, ethnicity, or other socially relevant groupings

Please specify the socially constructed or socially relevant categorization variable(s) used in your manuscript and explain why they were used. Please note that such variables should not be used as proxies for other socially constructed/relevant variables (for example, race or ethnicity should not be used as a proxy for socioeconomic status).

Provide clear definitions of the relevant terms used, how they were provided (by the participants/respondents, the researchers, or third parties), and the method(s) used to classify people into the different categories (e.g. self-report, census or administrative data, social media data, etc.)

Please provide details about how you controlled for confounding variables in your analyses.

Population characteristics

Describe the covariate-relevant population characteristics of the human research participants (e.g. age, genotypic information, past and current diagnosis and treatment categories). If you filled out the behavioural & social sciences study design questions and have nothing to add here, write "See above."

Recruitment

Describe how participants were recruited. Outline any potential self-selection bias or other biases that may be present and how these are likely to impact results.

Ethics oversight

Identify the organization(s) that approved the study protocol.

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

This study is a quantitative experimental research combining RNA sequencing, eQTL mapping, and phenotypic analysis to explore the contribution of gene expression diversity to modern wheat improvement.

Research sample

The research sample comprised 328 accessions of hexaploid common wheat (*Triticum aestivum* L.), including 172 landraces, 92 Chinese modern cultivars (MCCs), and 64 U.S. modern cultivars (USMCs). These accessions were selected to represent the major genetic diversity and breeding history of common wheat across different geographic and breeding backgrounds. Seedlings were grown under controlled conditions, and leaf tissues were collected at the two-week stage for RNA sequencing. The dataset integrates previously published whole-genome resequencing data (<https://doi.org/10.1093/plcell/koad229>) and newly generated RNA sequencing data from this study. In addition, RNA sequencing data were included for six samples of rye (*Secale cereale* L.) and seven samples of *Thinopyrum ponticum*, both representing introgression donor species widely used in wheat breeding. Leaf tissues of these species were also collected from seedlings at the two-week stage under standard cultivation conditions. Furthermore, phenotypic and genomic data from the KN9204 mutant library were recorded and provided by Xigang Liu and colleagues. This sample set was

	designed to capture population-scale gene expression variation and to represent the genetic and expression diversity underlying modern wheat improvement.
Sampling strategy	Varieties with sufficient representation from each subgroup (modern Chinese cultivars, modern United States cultivars, and landraces) were selected to maximize the diversity of genotypes and phenotypes. Landraces were broadly sampled from global wheat-growing regions to capture ancestral diversity, while modern cultivars from China and the United States were chosen to represent distinct breeding programs and improvement histories. Sampling was based on previously characterized germplasm panels and whole-genome resequencing data (https://doi.org/10.1093/plcell/koad229), ensuring broad geographic and breeding coverage. A total of 328 accessions were included to provide comprehensive representation of genetic and expression diversity in hexaploid wheat.
Data collection	A total of 328 wheat samples, six rye (<i>Secale cereale</i> L.) samples, and seven <i>Thinopyrum ponticum</i> samples were grown in trays under controlled conditions, and leaf tissues were collected at the two-week seedling stage for RNA sequencing. Phenotypic data for 34 agronomic traits were obtained from multi-year field experiments conducted at a farm in Zhao County, Shijiazhuang City, Hebei Province, China. These field phenotypes were manually recorded by Shusong Zheng and colleagues using pen and paper and subsequently digitized for analysis. Data for eight seedling powdery mildew resistance traits were collected in a greenhouse setting by Caihong Zhao and colleagues through manual assessment and paper-based recording, followed by digital entry and calculation of mean phenotypic values. In addition, field phenotypic data from the KN9204 mutant library were recorded and provided by Xigang Liu and colleagues.
Timing and spatial scale	Field phenotypic data for 34 agronomic traits were collected from 328 wheat accessions planted during the winter growing seasons over three consecutive years (2013–2016) at a farm in Zhao County, Shijiazhuang City, Hebei Province, China (38°05'N, 114°52'E). Field plots were arranged in a randomized design with row and column spacing of 110 cm × 25 cm, with three independent replicates. For each accession, five plants from the center of each plot were selected for trait evaluation. Seedling-stage powdery mildew resistance was assessed in a greenhouse using tray-grown seedlings. Each accession was evaluated with three biological replicates, and phenotypic observations were recorded manually on paper before being digitized for analysis.
Data exclusions	1699B line was removed from the genetic variance partitioning analyses due to low proportion of mapped RNA-seq reads to the reference genome.
Reproducibility	The field data in this study are multi-year with high repeatability and reliability.
Randomization	For the field experiments, a randomized block design was employed to ensure the random allocation of different wheat varieties to distinct plots within each experimental block. Each block consisted of multiple plots, and the placement of genotypes was randomized within these plots to account for any spatial variation. The experimental design allowed for the collection of multi-year, multi-location phenotypic data, ensuring that the results were not biased by environmental factors. This randomization helps to control for systematic errors and enhances the reliability and generalizability of the findings.
Blinding	Yes. Phenotypic data collection was performed in a blinded manner to minimize observer bias. The investigators who recorded the traits were unaware of the sample identities during measurement and data recording.
Did the study involve field work?	☒ Yes ☐ No

Field work, collection and transport

Field conditions	Field trials were conducted five consecutive years under rainfed and irrigated conditions at a research station in Zhao County, Shijiazhuang, Hebei Province, China.
Location	Zhao County, Shijiazhuang City, Hebei Province, China (38°05'N, 114°52'E)
Access & import/export	Not applicable.
Disturbance	Not applicable.

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	Involvement 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	Involvement in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging

Plants

Seed stocks	All seeds used in this study are stored in the seed repository of the Institute of Genetics and Developmental Biology, Chinese Academy of Sciences.
Novel plant genotypes	This study did not generate new genotype data; all genotypes were obtained from previously published resequencing data (https://doi.org/10.1093/plcell/koad229).
Authentication	Not applicable.