

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	No software/codes were used for data collection.
Data analysis	De novo genome assembly of Ultra-long reads and PromethION reads: Nextdenovo (v2.3.1). Gap filling: LR_Gapcloser. Polish the assembled genome: T2T-Polish(https://github.com/arangrhie/T2T-Polish), NextPolish (v1.4.1). Assembly quality evaluation: BUSCO (v5.1.2), LTR_retriever (v2.9.8), Merqury (v1.3). Transposable element annotation: RepeatMasker (http://repeatmasker.org/), LTR_retriever (v2.9.8), TEsorter (v1.4.6). Gene prediction: BRAKER2 (v2.1.5), HISAT2 (v2.2.1), StringTie (v2.2.1), GenomeThreader (v1.7.3), EVidenceModeler (v2.1.0), IGV-GSAm (v0.6.76). Gene domain prediction: InterProScan (v5.70-102.0), egg-nog-mapper (v2). Phylogenetic tree analysis: OrthoFinder (v2.5.5), MUSCLE (v5.1), IQ-TREE (v1.6.12), FigTree (v1.4.3), MCMCtree (v4.10.7). CENH3 ChIP-seq analysis: BWA (v0.7.17), MACS2 (v2.2.9.1). Identification of centromere satellite sequences: TAREAN, RepeatExplorer2. Evolutionary rate test: MEGAX. Collinearity analysis: jcv (v1.5.7).

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 (ONT, Illumina paired-end and RNA-seq) and genome assembly generated in this study have been deposited in the National Genomics Data Center under BioProject PRJCA041880.

Publicly available data used in this paper - GenBank accessions of reference genome sequences: *B. distachyon* (GCA_000005505.4), *O. granulate* (GWHAAEL000000000), *O. brachyantha* (GCA_000231095.2), *O. alta* (GWHAZTO 000000000), *O. eichingeri* (https://datacommons.cyverse.org/browse/iplant/home/shared/commons_repo/curated/Shenton_OofficinalisAnnotation_2019), *O. rhizomatis* (https://datacommons.cyverse.org/browse/iplant/home/shared/commons_repo/curated/Shenton_OofficinalisAnnotation_2019), *O. officinalis* (<https://figshare.com/s/32d0ac68cee7f647e1e3>), *O. minuta* (<https://figshare.com/s/32d0ac68cee7f647e1e3>), *O. punctata* (GCA_000573905.1), *O. meridionalis* (GCA_000338895.2), *O. longistaminata* (GCA_000789195.1), *O. glumipatula* (GCA_000576495.1), *O. glaberrima* (GCA_000147395.1), *O. barthii* (GCA_000182155.2), *O. nivara* (GCA_000576065.1), *O. rufipogon* (GCA_000817225.1), *O. sativa* (NIP: GWHDRIN000000000; MH63: CP054676–CP054688). SRA accessions of genome Illumina sequences: *O. granulate* (SRR6175405), *O. brachyantha* (DRR053294), *O. alta* (CRR123261), *O. officinalis* (DRR014203), *O. punctata* (DRR056215) and NIP (SRR25203306)

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	N/A
Reporting on race, ethnicity, or other socially relevant groupings	N/A
Population characteristics	N/A
Recruitment	N/A
Ethics oversight	N/A

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	No statistical methods were required to establish sample size in this study.
Data exclusions	No data were excluded from the analysis. Filtering parameters for raw sequencing data and sequence alignments were described in the methods.
Replication	For each phylogenetic analysis, we performed 1,000 bootstrap replicates to assess the robustness of the inferred tree topology. Each replicaterepresents an independent reconstruction of the same dataset, and the analysis was performed once with 1,000 replicates. All bootstrap attempts were successfully completed.
Randomization	Group allocation was not relevant to this study and therefore randomization was not necessary.
Blinding	Group allocation was not relevant to this study and therefore blinding was not necessary.

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

Antibodies

Antibodies used

anti-OsCENH3 antibody

Validation

The anti-OsCENH3 antibody was custom-prepared in rabbits using synthetic peptides (C-AEPKKKLQFERSPR) and diluted to 1:200 in TNB buffer (0.5 m Tris-HCl, pH 7.5, 0.75 m NaCl, and 2.5% blocking reagent). Having been validated by immunoassays, the anti-OsCENH3 antibody demonstrated its effectiveness for ChIP-seq in *Oryza australiensis* by yielding peaks with highly specific enrichment in pericentromeric regions, where they precisely co-localized with CentO satellite repeats.

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
☒	☐ Public health
☒	☐ National security
☒	☐ Crops and/or livestock
☒	☐ Ecosystems
☒	☐ Any other significant area

Experiments of concern

Does the work involve any of these experiments of concern:

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

Plants

Seed stocks	Single seed descent (SSD) germplasm for <i>O. australiensis</i> (IRGC ACC 105278) was obtained from the International Rice Research Institute (IRRI, Philippines) under the Standard Material Transfer Agreement.
Novel plant genotypes	No novel plant genotypes were generated in this study.
Authentication	Not applicable

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
May remain private before publication.

NGDC under BioProject CRR1936087 and CRR2575982

Files in database submission

NGDC under BioProject PRJCA041880

Genome browser session
(e.g. [UCSC](#))

NA

Methodology

Replicates	2
Sequencing depth	10X
Antibodies	Anti-CENH3
Peak calling parameters	Peaks for CENH3 were called using MACS2 (parameters: "-f BAM -g 9e8 --broad --broad-cutoff 0.1"). Subsequently, peaks with ChIP signals greater than 50 were selected and merged within 150 kb using Bedtools.
Data quality	Raw reads were first trimmed using skewer (parameters: "-q 30") and subsequently aligned to <i>O. australiensis</i> genome using bwa. SAMtools was used to filter multi-mapped reads. Picard MarkDuplicates was used to remove PCR duplicate reads.
Software	skewer, bwa, samtools, MACS2, Picard