nature portfolio

Corresponding author(s): Hong-Ye Li and Da-Wei Li

Last updated by author(s): Feb 9, 2024

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	1. RNA-seq was performed with Illumina NovaSeq 6000 at the Novogene Bioinformatics Technology (Peking, China). 2. CUT&Tag DNA libraries were sequenced on an Illumina NovaSeq platform (BGI, Shenzhen, China). 3. Quantitative real time PCR (qPCR) were performed on a Bio-Rad CFX96 real-time PCR system (Bio-RAD, USA). 4. Cell morphology was visualized using a confocal laser scanning microscope (LSM 880 with AiryScan, Carl Zeiss, Germany) and a scanning electron microscope (SEM, HITACHI, Japan). 5. Cell length and cell width were quantified by Fiji ImageJ tool.
Data analysis	1. The resulting RNA-seq reads were quality checked by using FastQC and mapped onto Phaeodactylum tricornutum reference genome. Differential gene expression was performed using the HISAT2-Stringtie-DESeq2 pipeline. A corrected FDR≤0.001 and |log2Ratio|≥1 were set as threshold for significant differential expression gene (DEGs) analysis. 2. Paired-end reads of CUT&Tag libraries were aligned using Bowtie2 version 2.2.5 (https://bowtie-bio.sourceforge.net/bowtie2/index.shtml). Peak calling was executed with Model-based Analysis of ChIP Seq 2 (MACS2). The promoter database of P. tricornutum was obtained by hiPromoters (Patent No.ZL202110822106.X), and the reads in peaks were pinpointed to promoters using BLAST. Gene ontology (GO) enrichment analysis o was performed using eggnog 5.0. Potential target genes were assigned into functional pathways using Kyoto Encyclopedia of Genes and Genomes (KEGG) Mapper tool (https://www.genome.jp/kegg/mapper/ mapper/). Binding motifs of PtHSF2 were analyzed using Multiple EM for Motif Elicitation (MEME).

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 RNA-seq data, along with the CUT&Tag sequencing data, were uploaded to The National Genomics Data Center (NGDC) under the accession code PRJCA023558 for the submission. Source data were 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	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	The specific sample sizes and results of statistical analysis were described in the Figure legends. Sample sizes were based on similar studies published previously and were sufficient to show statistical differences between the experimental group and the control group.
Data exclusions	No data was excluded from the analysis.
Replication	All the data were based on at least 3 independent biological replicates to ensure the reliability of the results, and the number of replicates is indicated in the corresponding figure legend and/or in the corresponding material and method section, with the exception of the sequencing of CUT&Tag experiment, which was undertaken in mixed samples.
Randomization	Samples were selected randomly.
Blinding	Noblinding was done as none of the experiments described in this study involve group allocation during data collection or analyses.

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

Plants

Seed stocks	Phaeodactylum tricornutum (CCMP2561) was obtained from the National Center for Marine Algae and Microbiota (NCMA, USA).
Novel plant genotypes	N/A
Authentication	N/A