nature portfolio

Corresponding author(s): Kenneth Birnbaum

Last updated by author(s): Feb 27, 2023

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	No additional custom code was develop in this study. Only basic function from R have been used to plot the data. Only the following softwares and packages were used: Cell Ranger 5.0.1 (10x Genomics, https://support.10xgenomics.com/single-cell-gene-expression/software/overview/welcome) was used to generate fastq files and align the reads from sequenced single cell or single nuclei Seurat package v4.0 (https://satijalab.org/seurat/) was used to integrate the expression matrices generated by Cell Ranger, according to tutorial present online. DIAMONDv2.0.6 (http://ab.inf.uni-tuebingen.de/software/diamond) and DupGen_finder.pl and DupGen_finder-unique.pl V1 (https://github.com/qiao- xin/DupGen_finder) were used to identify duplicated genes shinyGO V0.61 (http://bioinformatics.sdstate.edu/go/) was used to identify GO term enrichments FIMO algorythmv5.5.1 (https://meme- suite.org/meme/tools/fimo) was used to predict cis-regulatory elements in maize duplicates MetaNeighborv1 package in Python (https://github.com/gillislab/pyMN) was used to assess cross-species-cell transcriptional relationship scGENv2.1 (https://github.com/theislab/scgen) was used as an alternative method of single cell clustering Scanpyv1.9 (https://scanpy.readthedocs.io/en/stable/) was used calculate the nearest neighbors using scanpy.pp.neighbors, and generated a 2D projection using UMAP MINI-EXv1 (https://github.com/VIB-PSB/MINI-EX) was used to identify single cell regulatory networks ImageJ V2.9.0/1.53t

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

All reference genomes were downloaded from Arabidopsis TAIR10.38, at https://www.arabidopsis.org/, for Maize B73 v4, Sorghum bicolor v3 and Setaria viridis v2 reference genomes at https://plants.ensembl.org/.
All raw scRNA-seq and snRNA-seq data, expression matrices and analyzed R-Seurat objects are available under GEO accession (GSE225118). All data on duplicate genes are provided in Supplementary Tables. All cell specific marker genes for all species, including a shared pan library of marker genes between species are provided in Supplementary Table 4.
All data used to generate figures is available at https://figshare.com/articles/dataset/Data_for_Guillotin_et_al_/22331002, except for the following figures, for which the data can be found under GEO accession GSE225118, in the following deposited files: Arabidopsis_Cells_Nuclei_Seurat_Obj.RData.gz (Fig. 1c; Extended Data Fig. 2c,d; Extended Data Fig. 4a,b), Maize_Sorghum_Setaria_Cells_Nuclei_Seurat_Obj.RData.gz (Extended Data Fig. 3d, Extended Data Fig. 5c,d). Extended Data Figs. 2c,d; 3d; and Fig. 5c,d are clustered separately.

Human research participants

Policy information about studies involving human research participants and Sex and Gender in Research.

Reporting on sex and gender	No human related research have been used in this manuscript
Population characteristics	No human related research have been used in this manuscript
Recruitment	No human related research have been used in this manuscript
Ethics oversight	No human related research have been used in this manuscript

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 sample size calculation was performed. Each sampling was performed on a new batch of germinated seedlings, with a minimum of 30 root tips per sample. This amount was chosen to account for plant variability in root growth and minimum sample size requirement for statistical analysis. This amount was also calculated to avoid an excess of total amount of samples leading to technical difficulties in extracting cells or nuclei.
Data exclusions	No data was excluded
Replication	For each species, data were collected in the lab with at least 3 biological replicates for cells or nuclei (except for Sorghum, only one replicate was generated in single cells due to technical difficulties in digesting the cell wall of this species, as described in the manuscript, while 3 biological replicates were generated for Sorghum nuclei). Apart from the Sorghum cell-profiles for which protoplasts could not be generated (as above), all other replicates were successful. Each replicate is made by combining cells or nuclei from (30 to 200 root tips depending on the specie, as describe in Supplementary table 1). Each replicate allows the sequencing of 2000 to 7000 cells or nuclei. No technical repetitions were made in this study, ie: two single cell or single nuclei RNAseq made on the same day with the same batch of plant).
Randomization	No randomization was used in this study as plant were grown in only one control condition in the same environment, for all three biological replicates. Arabidopsis was grown in vitro and maize, sorghum and Setaria were grown in the same growth chamber in the same location at different time.
Blinding	Blinding was not possible in this study as there was no treatment effect to measure and as we compared species transcriptomes directly.

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

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging