# nature portfolio

Corresponding author(s):	Sascha H. Duttke
Last updated by author(s):	Jun 3, 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 <u>Editorial Policies</u> and the <u>Editorial Policy Checklist</u>.

<.	トつ	ıŤ١	ıc:	ŀι	CS
J	ιc	I L	I.O.	L I	LJ

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.
X	A description of all covariates tested
$\boxtimes$	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>
$\boxtimes$	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
$\times$	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 was used

Data analysis

Data was aligned using Hisat2 v 2.2.1

Data was analyzed using HOMER v.4.11 software (http://homer.ucsd.edu/homer/) and all commands including parameters used are listed within the methods section. Some data was analyzed using HOMER2 software which is "in principle offered to be published" at nature (2022-03-04864B). What was analyzed utilizing this software was noted within the methods section.

METASCAPE (https://metascape.org/gp/index.html) software was utilized for gene ontology analysis. MEIRLOP (https://github.com/npdeloss/meirlop) software was utilized for covariant centric motif analysis.

The following python libraries were also used to run the code: matplotlib (https://matplotlib.org/) v 3.9 pandas (https://pandas.pydata.org/) v 2.2.2 numpy v. 1.25 scipy v 1.7.3 statsmodel v 0.13.2 sys v 3.7.12 seaborn v 0.12.2

Any custom code used for analysis was listed within the methods section with a description of how it was done. This code requires Python 3 or higher to run.

All custom code and commands from HOMER software are described in detail for each of the steps it was involved within. These sections are: csRNA-seq Data Analysis, 5'GRO-seq and GRO-seq analysis, RNA-seq analysis, ChIP-seq analysis, ATAC-seq analysis, Motif correlation of stable and unstable TSRs, Transcript stability switch analysis, Mapping stats calculation, Histograms and annotation of TSS to captured reads, Tag distribution histograms, Hexamer Analysis, RNA processing-related motif finding, Pausing index, Gene Ontology Analysis, and STARR-seq

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 raw and processed data generated for this study can be accessed at NCBI Gene Expression Omnibus (GEO; https://www.ncbi.nlm.nih.gov/eo/) accession number GSE233927. All data can be browsed at: https://labs.wsu.edu/duttke/mcdonaldbr\_ernaplants\_2024/ All data generated and analyzed are summarized in Supplementary Table 3.

#### 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 <u>nature.com/documents/nr-reporting-summary-flat.pdf</u>			

### Life sciences study design

il studies must disclose on these points even when the disclosure is negative.			
Sample size	No sample-size calculations were utilized. We aimed to get at least two samples for each species, however, our samples were based upon what others were willing to donate to us.		
Data exclusions	No data was excluded.		
Replication	Whenever possible, two biological replicates for each species analyzed.		
Randomization	No randomization was utilized. This study involved looking at raw sequencing data, identifying transcription start sites, and comparing characteristics across eukaryotic species, as such no covariates were present and needed to be controlled for.		
Blinding	No blinding was utilized. This study involved looking at raw sequencing data, identifying transcription start sites, and comparing characteristics across eukaryotic species, as such there were not variables that could present bias in analysis and needed to be controlled for.		

## 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		ystems Methods		
n/a Involved in the study		n/a Involved in the study		
Antibodies		ChIP-seq		
Eukaryotic cell lines		Flow cytometry		
Palaeontology and archaeology		———		
Animals and other o	rganism	S		
Clinical data  Dual use research of	fconcer	n		
Plants	Concer	'		
Eukaryotic cell lin	es			
Policy information about ce	ell lines	and Sex and Gender in Research		
Cell line source(s)		Drosophila S2 - male sex, purchased from ATCC		
Authentication		Morphology and replication rate was compared to parameters provided by ATCC		
Mycoplasma contaminati	on	Cell lines were not tested for mycoplasma		
Commonly misidentified (See ICLAC register)	lines	No commonly misidentified lines were used.		
A				
Animals and othe	r res	earch organisms		
Policy information about <u>st</u> <u>Research</u>	<u>udies ir</u>	nvolving animals; ARRIVE guidelines recommended for reporting animal research, and Sex and Gender in		
Laboratory animals	N/A			
Wild animals	N/A			
Reporting on sex	N/A			
Field-collected samples N/A				
Ethics oversight	N/A			
_	he appro	oval of the study protocol must also be provided in the manuscript.		
Dual use research	of c	oncern		
Policy information about <u>dual use research of concern</u>				
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 Ecosystems				

Any other significant area

Experiments of concern					
Does the work involve any of these experiments of concern:					
No Yes	o Yes				
Demonstrate how	Demonstrate how to render a vaccine ineffective				
	to thera <sub>l</sub>	peutically useful antibiotics or antiviral agents			
		pathogen or render a nonpathogen virulent			
Increase transmiss					
	Alter the host range of a pathogen				
	-	cic/detection modalities  of a biological agent or toxin			
		of a biological agent of toxin			
Z L / w// series potential	an, nam				
DI .					
Plants					
Seed stocks		ana - Col-0			
		Physcomitrium (Physcomitrella) patens - Gransden Selaginella moellendorffii was purchased from Plant Delights Nursery			
Novel plant genotypes	Not ap	plicable			
Authentication	Describe any authentication procedures for each seed stock used or novel genotype generated. Describe any experiments used to assess the effect of a mutation and, where applicable, how potential secondary effects (e.g. second site T-DNA insertions, mosiacism,				
	off-target gene editing) were examined.				
ChIP-seq					
Data deposition					
	w and fi	nal processed data have been deposited in a public database such as <u>GEO</u> .			
		sited or provided access to graph files (e.g. BED files) for the called peaks.			
Data access links May remain private before publi	cation. N/A				
Files in database submiss	sion	N/A			
Genome browser session					
(e.g. <u>UCSC</u> )	ı	N/A			
Methodology					
Replicates	Describ	be the experimental replicates, specifying number, type and replicate agreement.			
Sequencing depth	Describe the sequencing depth for each experiment, providing the total number of reads, uniquely mapped read whether they were paired- or single-end.				
		be the antibodies used for the ChIP-seq experiments; as applicable, provide supplier name, catalog number, clone name, and nber.			
Peak calling parameters Specused		the command line program and parameters used for read mapping and peak calling, including the ChIP, control and index files			
Data quality	Describ	Describe the methods used to ensure data quality in full detail, including how many peaks are at FDR 5% and above 5-fold enrichment.			
		be the software used to collect and analyze the ChIP-seq data. For custom code that has been deposited into a community cory, provide accession details.			