nature portfolio

Corresponding author(s): Samuel H. Sternberg

Last updated by author(s): Aug 14, 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	All commercial, open source and/or custom code was used for bioinformatic and data analysis
Data analysis	All commercial, open source and/or custom code used for analyses of the data are indicated in manuscript. This includes CD-HIT (v4.8.1), MAFFT (v7.508), trimAL (v1.4.rev15) , IQ-Tree 2 (v2.1.4), Infernal (v1.1.4), CMFinder (v0.4.1.9), LocARNA (v1.9.1), MUSCLE (v3.8.1551), Prodigal, HMMR Suite (v3.3.2), FastTree2 (2.1.11), cutPrimers(v4.2), UMI-tools (v1.1.4), SAMtools (v1.17), deepTools2 (v3.5.1), Integrative Genomics Viewer, MACS, BEDTools, MEME Suite, WebLogo (v2.8), Geneious Prime (2023.0.1), Prism (9.3.1). All Custom code used in this study can be found at https://github.com/sternberglab/Meers_et_al_2023

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

Next-generation sequencing data are available in the National Center for Biotechnology Information (NCBI) Sequence Read Archive: SRX19058888-SRX19058905, SRR23476356-SRR23476358 and SRR24994123 (BioProject Accession: PRJNA925099) and the Gene Expression Omnibus (GSE223127). The published genome used for ChIP-seq analyses was obtained from NCBI (GenBank: CP001509.3). The published genome used for bioinformatics analyses of the Geobacillus stearothermophilus genome was obtained from NCBI (GenBank: NZ_CP016552.1). Datasets generated and analyzed in the current study are available from the corresponding author upon reasonable request.

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

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	Sample sizes are reported in the figure legends and generally encompassed three biological replicates.
Data exclusions	No data were excluded.
Replication	All data could be reproduced, and most experiments and analyses presented were the result of two or three independent biological replicates.
Randomization	Samples were not randomized as it was not applicable for the design of this study.
Blinding	Investigators were not blinded as it was not applicable for the design of this study.

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

Antibodies

Antibodies used: Monoclonal ANTI-FLAG M2 antibody produced in mouse, Sigma Aldrich, Catalogue number: F1804, clone M2

Validation: According to the manufacturer, the "monoclonal ANTI-FLAG® M2 may be used in IP [immunoprecipitation] procedures when used in conjunction with an insoluble carrier matrix, such as a Protein G resin" (https://www.sigmaaldrich.com/deepweb/assets/sigmaaldrich/product/documents/175/747/f1804bul-ms.pdf). As suggested, the ANTI-FLAG M2 antibody was used together with Dynabeads Protein G resin (Thermo Fisher) in this study. Furthermore, ANTI-FLAG M2 antibody was used in a ChIP-seq study by Partridge et al., Nature (2020), titled "Occupancy maps of 208 chromatin-associated proteins in one human cell type" and by Hoffmann, Kim, Beh et al., Nature (2022), titled "Selective TnsC recruitment enhances the fidelity of RNA-guided transposition."

Plants

Seed stocks: *Report on the source of all seed stocks or other plant material used. If applicable, state the seed stock centre and catalogue number. If plant specimens were collected from the field, describe the collection location, date and sampling procedures.*

Novel plant genotypes: *Describe the methods by which all novel plant genotypes were produced. This includes those generated by transgenic approaches, gene editing, chemical/radiation-based mutagenesis and hybridization. For transgenic lines, describe the transformation method, the number of independent lines analyzed and the generation upon which experiments were performed. For gene-edited lines, describe the editor used, the endogenous sequence targeted for editing, the targeting guide RNA sequence (if applicable) and how the editor was applied.*

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

☒ 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.

Raw sequencing reads, processed sequencing reads and MACS3 peak call files (all files listed below) were uploaded to GEO (accession: GSE223127). The GEO reviewer accession token for this submission is: wxsjmwmyrtcpvoz. All data deposited in GEO will be publicly accessible.

Files in database submission

Raw sequencing files:
Cas12a_NT_ChIP-seq_paired_raw_rep1_1.fastq.gz
Cas12a_NT_ChIP-seq_paired_raw_rep1_2.fastq.gz
Cas12a_NT_ChIP-seq_paired_raw_rep2_1.fastq.gz
Cas12a_NT_ChIP-seq_paired_raw_rep2_2.fastq.gz
Cas12a_targeting_ChIP-seq_paired_raw_rep1_1.fastq.gz
Cas12a_targeting_ChIP-seq_paired_raw_rep1_2.fastq.gz
Cas12a_targeting_ChIP-seq_paired_raw_rep2_1.fastq.gz
Cas12a_targeting_ChIP-seq_paired_raw_rep2_2.fastq.gz
Cas9_NT_ChIP-seq_paired_raw_rep1_1.fastq.gz
Cas9_NT_ChIP-seq_paired_raw_rep1_2.fastq.gz
Cas9_NT_ChIP-seq_paired_raw_rep2_1.fastq.gz

Cas9_NT_ChIP-seq_paired_raw_rep2_2.fastq.gz
Cas9_targeting_ChIP-seq_paired_raw_rep1_1.fastq.gz
Cas9_targeting_ChIP-seq_paired_raw_rep1_2.fastq.gz
Cas9_targeting_ChIP-seq_paired_raw_rep2_1.fastq.gz
Cas9_targeting_ChIP-seq_paired_raw_rep2_2.fastq.gz
Input_Cas9_targeting_ChIP-seq_paired_raw_1.fastq.gz
Input_Cas9_targeting_ChIP-seq_paired_raw_2.fastq.gz
Input_TnpB3_targeting_ChIP-seq_paired_raw_1.fastq.gz
Input_TnpB3_targeting_ChIP-seq_paired_raw_2.fastq.gz
IscB_NT_ChIP-seq_paired_raw_rep1_1.fastq.gz
IscB_NT_ChIP-seq_paired_raw_rep1_2.fastq.gz
IscB_NT_ChIP-seq_paired_raw_rep2_1.fastq.gz
IscB_NT_ChIP-seq_paired_raw_rep2_2.fastq.gz
IscB_targeting_ChIP-seq_paired_raw_rep1_1.fastq.gz
IscB_targeting_ChIP-seq_paired_raw_rep1_2.fastq.gz
IscB_targeting_ChIP-seq_paired_raw_rep2_1.fastq.gz
IscB_targeting_ChIP-seq_paired_raw_rep2_2.fastq.gz
TnpB3_NT_ChIP-seq_paired_raw_rep1_1.fastq.gz
TnpB3_NT_ChIP-seq_paired_raw_rep1_2.fastq.gz
TnpB3_NT_ChIP-seq_paired_raw_rep2_1.fastq.gz
TnpB3_NT_ChIP-seq_paired_raw_rep2_2.fastq.gz
TnpB3_targeting_ChIP-seq_paired_raw_rep1_1.fastq.gz
TnpB3_targeting_ChIP-seq_paired_raw_rep1_2.fastq.gz
TnpB3_targeting_ChIP-seq_paired_raw_rep2_1.fastq.gz
TnpB3_targeting_ChIP-seq_paired_raw_rep2_2.fastq.gz

Processed sequencing files (bigWig), can be visualized in Integrative Genomics Viewer (IGV) using corresponding reference genome file provided on GitHub (https://github.com/sternberglab/Meers_et_al_2023), listed in Supplementary Table X:
Cas12a_NT_ChIP-seq_paired_rep1_genome-mapping_max-value-1kb-windows.bw
Cas12a_NT_ChIP-seq_paired_rep2_genome-mapping_max-value-1kb-windows.bw
Cas12a_targeting_ChIP-seq_paired_rep1_genome-mapping_max-value-1kb-windows.bw
Cas12a_targeting_ChIP-seq_paired_rep2_genome-mapping_max-value-1kb-windows.bw
Cas9_NT_ChIP-seq_paired_rep1_genome-mapping_max-value-1kb-windows.bw
Cas9_NT_ChIP-seq_paired_rep2_genome-mapping_max-value-1kb-windows.bw
Cas9_targeting_ChIP-seq_paired_rep1_genome-mapping_max-value-1kb-windows.bw
Cas9_targeting_ChIP-seq_paired_rep2_genome-mapping_max-value-1kb-windows.bw
Input_Cas9_targeting_ChIP-seq_paired_genome-mapping_max-value-1kb-windows.bw
Input_TnpB3_targeting_ChIP-seq_paired_genome-mapping_max-value-1kb-windows.bw
IscB_NT_ChIP-seq_paired_rep1_genome-mapping_max-value-1kb-windows.bw
IscB_NT_ChIP-seq_paired_rep2_genome-mapping_max-value-1kb-windows.bw
IscB_targeting_ChIP-seq_paired_rep1_genome-mapping_max-value-1kb-windows.bw
IscB_targeting_ChIP-seq_paired_rep2_genome-mapping_max-value-1kb-windows.bw
TnpB3_NT_ChIP-seq_paired_rep1_genome-mapping_max-value-1kb-windows.bw
TnpB3_NT_ChIP-seq_paired_rep2_genome-mapping_max-value-1kb-windows.bw
TnpB3_targeting_ChIP-seq_paired_rep1_genome-mapping_max-value-1kb-windows.bw
TnpB3_targeting_ChIP-seq_paired_rep2_genome-mapping_max-value-1kb-windows.bw
Cas12a_NT_ChIP-seq_paired_rep1_genome-mapping.bw
Cas12a_NT_ChIP-seq_paired_rep2_genome-mapping.bw
Cas12a_targeting_ChIP-seq_paired_rep1_genome-mapping.bw
Cas12a_targeting_ChIP-seq_paired_rep2_genome-mapping.bw
Cas9_NT_ChIP-seq_paired_rep1_genome-mapping.bw
Cas9_NT_ChIP-seq_paired_rep2_genome-mapping.bw
Cas9_targeting_ChIP-seq_paired_rep1_genome-mapping.bw
Cas9_targeting_ChIP-seq_paired_rep2_genome-mapping.bw
Input_Cas9_targeting_ChIP-seq_paired_genome-mapping.bw
Input_TnpB3_targeting_ChIP-seq_paired_genome-mapping.bw
IscB_NT_ChIP-seq_paired_rep1_genome-mapping.bw
IscB_NT_ChIP-seq_paired_rep2_genome-mapping.bw
IscB_targeting_ChIP-seq_paired_rep1_genome-mapping.bw
IscB_targeting_ChIP-seq_paired_rep2_genome-mapping.bw
TnpB3_NT_ChIP-seq_paired_rep1_genome-mapping.bw
TnpB3_NT_ChIP-seq_paired_rep2_genome-mapping.bw
TnpB3_targeting_ChIP-seq_paired_rep1_genome-mapping.bw
TnpB3_targeting_ChIP-seq_paired_rep2_genome-mapping.bw

Processed sequencing files (bed files containing MACS3 peak calls):
Cas12a_NT_rep1_macs3_summits.bed
Cas12a_NT_rep2_macs3_summits.bed
Cas12a_targeting_rep1_macs3_summits.bed
Cas12a_targeting_rep2_macs3_summits.bed
Cas9_NT_rep1_macs3_summits.bed
Cas9_NT_rep2_macs3_summits.bed
Cas9_targeting_rep1_macs3_summits.bed
Cas9_targeting_rep2_macs3_summits.bed
IscB_NT_rep1_macs3_summits.bed
IscB_NT_rep2_macs3_summits.bed

IscB_targeting_rep1_macs3_summits.bed
IscB_targeting_rep2_macs3_summits.bed
TnpB3_NT_rep1_macs3_summits.bed
TnpB3_NT_rep2_macs3_summits.bed
TnpB3_targeting_rep1_macs3_summits.bed
TnpB3_targeting_rep2_macs3_summits.bed

The Meta table uploaded to GEO contains the same read count information as Supplementary Table 5: Table_Supplement.xlsx

Genome browser session
(e.g. UCSC)

odified reference genomes (all available on GitHub: https://github.com/sternberglab/Meers_et_al_2023) were used. Normalized bigWig files can be visualized in the Integrative Genomics Viewer (IGV) using the bigWig (.bw) file of choice (normalized using bamCoverage) provided in GEO together with the respective reference genome file used for read mapping.

Methodology

Replicates

Two biological replicates were used for ChIP-seq samples.

Sequencing depth

The number of raw reads, uniquely mapped reads, length of reads and paired- or single-end nature are provided in Supplementary Table 5.

Antibodies

Monoclonal ANTI-FLAG M2 antibody produced in mouse, Sigma Aldrich, Catalogue number: F1804, clone M2

Peak calling parameters

#Trim reads using fastp
fastp -i "input_read1.fastq.gz" -I "input_read2.fastq.gz" -o "trimmed_output_read1.fastq.gz" -O "trimmed_output_read2.fastq.gz" -j "log".json -h "log".html

#Map reads using bowtie2 (creates a .sam output file)
#Reads were mapped to modified E. coli BL21(DE3) reference genomes (GenBank accession: CP001509.3)
bowtie2 -x “directory_to_BL21_reference_genome_file” -1 “trimmed_output_read1.fastq.gz” -2 “trimmed_output_read1.fastq.gz” -S “output.sam”

#Convert .sam into .bam file using samtools
samtools -view -b “input.sam” > “output_directory”

#Sort the .bam files using samtools
samtools sort -o output_directory “input.bam”

#Index the aligned and sorted .bam files using samtools
samtools index -b “input.bam” “output.bam.bai”

#Eliminate multi-mapping reads using samtools (retains only uniquely mapping reads);
#uses a MAPQ score of 10 as a cutoff
samtools view -bq 10 “input.bam” > “output_directory”

#Create index files for the trimmed, aligned, sorted and uniquely mapping reads using samtools
samtools index -b “input.bam” “output.bam.bai”

#Normalize reads using deepTools2 bamCoverage with the option "RPKM"
bamCoverage --normalizeUsing RPKM -bs 1 -b "not_normalized.bam" -o "normalized.bw"

#MACS3 peak calling
macs3 callpeak -t target.bam -c input.bam -n "IscB_targeting_macs3_peaks" -g 4500000 --nomodel --extsize 400 -q 0.05 -B --outdir "scB_targeting_macs3_peaks"

Control files (input files):
Input_Cas9_targeting_ChIP-seq_paired_raw.
Input_TnpB3_targeting_ChIP-seq_paired_raw

All index files generated using "samtools faidx <E_coli_BL21_reference_genome>", and are available on GitHub in the "E_coli_BL21_Reference_genomes" folder: https://github.com/sternberglab/Meers_et_al_2023.

Reference genome files as described in Supplementary Figure 5 are publicly available on GitHub: https://github.com/sternberglab/Meers_et_al_2023.

Data quality

Bowtie2 used default read quality parameters for mapping. Multi-mapping reads with a MAPQ score <10 were eliminated. All peaks called are below FDR 5%, as per MAC3 standard output.

Peaks above 5-fold enrichment:
Cas12a NT rep1: 6
Cas12a NT rep2: 9
Cas12a targeting rep1: 13
Cas12a targeting rep2: 27

Cas9 NT rep1: 0
Cas9 NT rep2: 0
Cas9 targeting rep1: 0
Cas9 targeting rep2: 1

IscB NT rep1: 44
IscB NT rep2: 32
IscB targeting rep1: 3
IscB targeting rep2: 3

TnpB3 NT rep1: 2
TnpB3 NT rep2: 2
TnpB3 targeting rep1: 2
TnpB3 targeting rep2: 3

Software

Illumina BaseSpace was used for automated read demultiplexing and adaptor trimming. All custom code for the ChIP-seq analysis was uploaded to GitHub (https://github.com/sternberglab/Meers_et_al_2023).