

Corresponding author(s):	DBPR NCOMMS-18-17955E
--------------------------	-----------------------

Last updated by author(s): Jul 24, 2019

# **Reporting Summary**

X Life sciences

Nature Research 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 Research policies, see <u>Authors & Referees</u> and the <u>Editorial Policy Checklist</u>.

Statistics				
For all statistical analy	ses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.			
n/a Confirmed				
☐ ☐ The exact sar	mple size $(n)$ for each experimental group/condition, given as a discrete number and unit of measurement			
	on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly			
The statistical Only common	al test(s) used AND whether they are one- or two-sided tests should be described solely by name; describe more complex techniques in the Methods section.			
A description	A description of all covariates tested			
A description	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons			
	tion of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) n (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)			
For null hypo	othesis testing, the test statistic (e.g. $F$ , $t$ , $r$ ) with confidence intervals, effect sizes, degrees of freedom and $P$ value noted as exact values whenever suitable.			
For Bayesian	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings			
For hierarchi	cal 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 <u>statistics for biologists</u> contains articles on many of the points above.			
Software and	code			
Policy information abo	out <u>availability of computer code</u>			
Data collection	Provide a description of all commercial, open source and custom code used to collect the data in this study, specifying the version used OR state that no software was used.			
Data analysis	Provide a description of all commercial, open source and custom code used to analyse the data in this study, specifying the version used OR state that no software was used.			
	stom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers. e deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.			
Data				
- Accession codes, u - A list of figures that	out <u>availability of data</u> : include a <u>data availability statement</u> . This statement should provide the following information, where applicable: nique identifiers, or web links for publicly available datasets t have associated raw data y restrictions on data availability			
The RNA sequencing dat	ta were available from the GEO database under accession number GSE112859			
Field-spec	ific 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.

Ecological, evolutionary & environmental sciences

Behavioural & social sciences

# Life sciences study design

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

Sample size No statistical methods were used to predetermine sample size. The sample size was chosen according to research in the same field.

Data exclusions No data were excluded for analysis.

Replication All experiments were carried out with at least three biological replicates.

Randomization Samples were allocated randomly.

Research sample

Sampling strategy

Blinding The investigators were not blinded to allocation during experiments.

## Behavioural & social sciences study design

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

Study description Briefly describe the study type including whether data are quantitative, qualitative, or mixed-methods (e.g. qualitative cross-sectional,

quantitative experimental, mixed-methods case study).

State the research sample (e.g. Harvard university undergraduates, villagers in rural India) and provide relevant demographic information (e.g. age, sex) and indicate whether the sample is representative. Provide a rationale for the study sample chosen. For studies involving

existing datasets, please describe the dataset and source.

Describe the sampling procedure (e.g. random, snowball, stratified, convenience). Describe the statistical methods that were used to predetermine sample size OR if no sample-size calculation was performed, describe how sample sizes were chosen and provide a rationale for why these sample sizes are sufficient. For qualitative data, please indicate whether data saturation was considered, and what criteria

were used to decide that no further sampling was needed.

Data collection

Provide details about the data collection procedure, including the instruments or devices used to record the data (e.g. pen and paper, computer, eye tracker, video or audio equipment) whether anyone was present besides the participant(s) and the researcher, and whether

the researcher was blind to experimental condition and/or the study hypothesis during data collection.

Timing Indicate the start and stop dates of data collection. If there is a gap between collection periods, state the dates for each sample cohort.

Data exclusions | If no data were excluded from the analyses, state so OR if data were excluded, provide the exact number of exclusions and the rationale

behind them, indicating whether exclusion criteria were pre-established.

Non-participation State how many participants dropped out/declined participation and the reason(s) given OR provide response rate OR state that no

participants dropped out/declined participation.

Randomization | If participants were not allocated into experimental groups, state so OR describe how participants were allocated to groups, and if

allocation was not random, describe how covariates were controlled.

# Ecological, evolutionary & environmental sciences study design

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

Study description

Briefly describe the study. For quantitative data include treatment factors and interactions, design structure (e.g. factorial, nested, hierarchical), nature and number of experimental units and replicates.

Research sample

Describe the research sample (e.g. a group of tagged Passer domesticus, all Stenocereus thurberi within Organ Pipe Cactus National Monument), and provide a rationale for the sample choice. When relevant, describe the organism taxa, source, sex, age range and any manipulations. State what population the sample is meant to represent when applicable. For studies involving existing datasets,

describe the data and its source.

Sampling strategy

Note the sampling procedure. Describe the statistical methods that were used to predetermine sample size OR if no sample-size calculation was performed, describe how sample sizes were chosen and provide a rationale for why these sample sizes are sufficient.

Data collection Describe the data collection procedure, including who recorded the data and how.

Timing and spatial scale Indicate the start and stop dates of data collection, noting the frequency and periodicity of sampling and providing a rationale for these choices. If there is a gap between collection periods, state the dates for each sample cohort. Specify the spatial scale from which

the data are taken

	If no data were excluded from the analyses, state so OR if data were excluded, describe the exclusions and the rationale behind them, indicating whether exclusion criteria were pre-established.	
,	Describe the measures taken to verify the reproducibility of experimental findings. For each experiment, note whether any attempts to repeat the experiment failed OR state that all attempts to repeat the experiment were successful.	
	Describe how samples/organisms/participants were allocated into groups. If allocation was not random, describe how covariates were controlled. If this is not relevant to your study, explain why.	
O	Describe the extent of blinding used during data acquisition and analysis. If blinding was not possible, describe why OR explain why blinding was not relevant to your study.	
Did the study involve field	work? Yes No	
Field work, collect	ion and transport	
Field conditions	Describe the study conditions for field work, providing relevant parameters (e.g. temperature, rainfall).	
Location	State the location of the sampling or experiment, providing relevant parameters (e.g. latitude and longitude, elevation, water depth).	
Access and import/export	Describe the efforts you have made to access habitats and to collect and import/export your samples in a responsible manner and in compliance with local, national and international laws, noting any permits that were obtained (give the name of the issuing authority, the date of issue, and any identifying information).	
Disturbance	Describe any disturbance caused by the study and how it was minimized.	
We require information from au	n/a Involved in the study  ChIP-seq  Flow cytometry  MRI-based neuroimaging	
Antibodies used	QKI6 (AB9906, Merck Millipore); ACVR1B (ab109300, Abcam); GAPDH (60004-1-lg, Proteintech); Ki-67 (9449T, Cell Signaling	
Antibodies used	Technology); Caspase-3 (9664T, CST); anti-Ago2 antibody (ab32381, Abcam); rabbit isotype IgG (12-370, Merck Millipore)	
Validation	The validation of each primary antibody was provided in manuscript.	
Eukaryotic cell line	es established	
Policy information about <u>cel</u>	l lines	
Cell line source(s)	293T and gastric cancer cell line, MGC-803 and HGC-27 were obtained from the ATCC.	
Authentication	none of these cell lines was authenticated.	
Mycoplasma contaminatio	on Cell lines were tested negative for mycoplasma contamination.	
Commonly misidentified li (See <u>ICLAC</u> register)	No misidentified cell lines were used in the study.	

## Palaeontology

Specimen provenance

Provide provenance information for specimens and describe permits that were obtained for the work (including the name of the

Specimen provenance	issuing authority, the date of issue, and any identifying information).
Specimen deposition	Indicate where the specimens have been deposited to permit free access by other researchers.
Dating methods	If new dates are provided, describe how they were obtained (e.g. collection, storage, sample pretreatment and measurement), where they were obtained (i.e. lab name), the calibration program and the protocol for quality assurance OR state that no new dates are provided.
Tick this box to confirm	that the raw and calibrated dates are available in the paper or in Supplementary Information.
nimals and other	organisms
olicy information about stud	dies involving animals; ARRIVE guidelines recommended for reporting animal research
Laboratory animals	4-6 week old female BALB/c and NOD/SCID mice
Wild animals	this study did not involve wild animals.
Field-collected samples	this study did not involve field collected samples.
Ethics oversight	All animal experiments were performed with the approval of the Research Ethics Committee of Institute of Basic Medical Science.
ote that full information on the	e approval of the study protocol must also be provided in the manuscript.
luman research p	articipants
olicy information about stud	dies involving human research participants
Population characteristics	Describe the covariate-relevant population characteristics of the human research participants (e.g. age, gender, genotypic information, past and current diagnosis and treatment categories). If you filled out the behavioural & social sciences study designations 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.
te that full information on the	e approval of the study protocol must also be provided in the manuscript.
linical data	
olicy information about <u>clin</u> manuscripts should comply w	ical studies vith the ICMJE <u>guidelines for publication of clinical research</u> and a completed <u>CONSORT checklist</u> must be included with all submission
Clinical trial registration	Provide the trial registration number from ClinicalTrials.gov or an equivalent agency.
Study protocol	Note where the full trial protocol can be accessed OR if not available, explain why.
	Describe the settings and locales of data collection, noting the time periods of recruitment and data collection.
Data collection	
	Describe how you pre-defined primary and secondary outcome measures and how you assessed these measures.
Data collection Outcomes ChIP-sea	
Outcomes ChIP-seq	
Outcomes  ChIP-seq ata deposition	

Data access links

May remain private before publication.

For "Initial submission" or "Revised version" documents, provide reviewer access links. For your "Final submission" document, provide a link to the deposited data.

Files in database submission

Provide a list of all files available in the database submission.

Genome browser session (e.g. <u>UCSC</u>)

Provide a link to an anonymized genome browser session for "Initial submission" and "Revised version" documents only, to enable peer review. Write "no longer applicable" for "Final submission" documents.

#### Methodology

**Antibodies** 

Replicates Describe 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 reads, length of reads and whether they were paired- or single-end.

Describe the antibodies used for the ChIP-seq experiments; as applicable, provide supplier name, catalog number, clone

name, and lot number.

Peak calling parameters

Specify the command line program and parameters used for read mapping and peak calling, including the ChIP, control and index files used.

Data quality

Describe the methods used to ensure data quality in full detail, including how many peaks are at FDR 5% and above 5-fold enrichment.

Software

Describe the software used to collect and analyze the ChIP-seq data. For custom code that has been deposited into a community repository, provide accession details.

### Flow Cytometry

### Plots

### Methodology

Sample preparation Describe the sample preparation, detailing the biological source of the cells and any tissue processing steps used.

**Instrument** Identify the instrument used for data collection, specifying make and model number.

Software Describe the software used to collect and analyze the flow cytometry data. For custom code that has been deposited into a

community repository, provide accession details.

Cell population abundance Describe the abundance of the relevant cell populations within post-sort fractions, providing details on the purity of the samples and how it was determined.

Describe the gating strategy used for all relevant experiments, specifying the preliminary FSC/SSC gates of the starting cell

population, indicating where boundaries between "positive" and "negative" staining cell populations are defined.

Tick this box to confirm that a figure exemplifying the gating strategy is provided in the Supplementary Information.

### Magnetic resonance imaging

#### Experimental design

Gating strategy

Design type Indicate task or resting state; event-related or block design.

Design specifications

Specify the number of blocks, trials or experimental units per session and/or subject, and specify the length of each trial or block (if trials are blocked) and interval between trials.

Behavioral performance measures

State number and/or type of variables recorded (e.g. correct button pre

State number and/or type of variables recorded (e.g. correct button press, response time) and what statistics were used to establish that the subjects were performing the task as expected (e.g. mean, range, and/or standard deviation across subjects).

Acquisition		
Imaging type(s)	Specify: functional, structural, diffusion, perfusion.	
Field strength	Specify in Tesla	
Sequence & imaging parameters	Specify the pulse sequence type (gradient echo, spin echo, etc.), imaging type (EPI, spiral, etc.), field of view, matrix size, slice thickness, orientation and TE/TR/flip angle.	
Area of acquisition	State whether a whole brain scan was used OR define the area of acquisition, describing how the region was determined.	
Diffusion MRI Used	Not used	
Preprocessing		
Preprocessing software	Provide detail on software version and revision number and on specific parameters (model/functions, brain extraction, segmentation, smoothing kernel size, etc.).	
Normalization	If data were normalized/standardized, describe the approach(es): specify linear or non-linear and define image types used for transformation OR indicate that data were not normalized and explain rationale for lack of normalization.	
Normalization template	Describe the template used for normalization/transformation, specifying subject space or group standardized space (e.g. original Talairach, MNI305, ICBM152) OR indicate that the data were not normalized.	
Noise and artifact removal	Describe your procedure(s) for artifact and structured noise removal, specifying motion parameters, tissue signals and physiological signals (heart rate, respiration).	
Volume censoring	Define your software and/or method and criteria for volume censoring, and state the extent of such censoring.	
Statistical modeling & inferenc	e	
Model type and settings	Specify type (mass univariate, multivariate, RSA, predictive, etc.) and describe essential details of the model at the first and second levels (e.g. fixed, random or mixed effects; drift or auto-correlation).	
Effect(s) tested	Define precise effect in terms of the task or stimulus conditions instead of psychological concepts and indicate whether ANOVA or factorial designs were used.	
Specify type of analysis: Whol	e brain ROI-based Both	
Statistic type for inference (See <u>Eklund et al. 2016</u> )	Specify voxel-wise or cluster-wise and report all relevant parameters for cluster-wise methods.	
Correction	Describe the type of correction and how it is obtained for multiple comparisons (e.g. FWE, FDR, permutation or Monte Carlo).	
Models & analysis		
n/a Involved in the study  Functional and/or effective co  Graph analysis  Multivariate modeling or pred		
Functional and/or effective connect	Report the measures of dependence used and the model details (e.g. Pearson correlation, partial correlation, mutual information).	
Graph analysis	We chose the appropriate tests according to the data distributions. Statistical differences in tissue miRNA expression levels between cancer and normal sample sources were determined using two-sided Mann-Whitney II test. Student's tast (two-tailed) was performed for two-group data and three-group data were	

analyzed using one-way analysis of variance. The Spearman's rank correlation test was conducted for statistical correlations. All data were analyzed using GraphPad Prism 5.0 software (GraphPad Software, Inc., USA) or SPSS 16.0 software (SPSS Inc., Chicago, IL, USA) and presented as means  $\pm$  SD. Specify independent variables, features extraction and dimension reduction, model, training and evaluation

metrics.

Multivariate modeling and predictive analysis