nature portfolio

Corresponding author(s):	nidia.alvarez.armada@gmail.com	
Last updated by author(s):	2024/07/18	

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.

_		4.0			
\cdot	tっ	ıtı	C	۲ı	CS
.)	La	ш		u	1

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
	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>
	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 was collection on statistics for histories contains articles on many of the points above

Software and code

Policy information about availability of computer code

The beam line used a toroidal mirror to focus the beam and a Si (111) monochromator for energy selection, with the storage ring containing 500 mA in top-off mode at 3.0 GeV. Scans were monitored during analysis using the software packages MicroScanGUI and uXAS GUI.

SRS-XRF data were analyzed using Sam's Microprobe Analysis Toolkit. Mean hex (RGB) values were calculated for each ROI using the RGB Measure.java pluqin for ImageJ. The statistical significance of variation in elemental concentrations was assessed using MANOVA at 95% confidence intervals in the packages "stats" and "MASS" in R. All software packages described here are open source.

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 dataset associated with this study is available from the FigShare digital repository: https://doi.org/10.6084/m9.figshare.26086177. The data are provided in the electronic supplementary material.

Research involving human participants, their data, or biological material

Policy information about studies with <u>human participants or human data</u>. See also policy information about <u>sex, gender (identity/presentation)</u>, <u>and sexual orientation</u> and <u>race</u>, 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 belo	w that is the best fit for your research.	If you are not sure, read the appropriate sections before making your selection.
X 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 extant specimens analyzed (n = 78) represent 55 species, 27 families and eight orders and were selected for analysis because they meet the following criteria: (i) the color pattern on the wings or elytra comprises large, sharply defined blocks, bands or spots; (ii) the cuticle lacks prominent bosses, tubercles and setae (thus avoiding topographic artefacts in SRS-XRF maps); (iii) the pigment(s) responsible for producing the patterning is known for that species or closely related species in the same family; (iv) the family is represented in the fossil record; and (v) specimens are readily available (Fig. S1 and Tables S1, S2). In all specimens, the color pattern was analyzed on the wings or elytra only. The fossil specimens analyzed (n = 15) represent 13 species and four orders and were selected for analysis because they meet the following criteria: (i) the cuticle includes sharply defined blocks, bands or spots that differ in visual tone; (ii) the cuticle is preserved in both dark- and light-toned regions; and (iii) the taxa have extant representatives at family level (Fig. S8 and Tables S5, S6). All fossils analyzed are preserved as carbonaceous impressions.

Data exclusions

Data for eleven specimens in Figs. S3 and S6 were calibrated against a different set of standards to the rest of the dataset and are thus analyzed separately.

Replication

In modern insects, regions of interest (ROIs) were selected for cuticle regions with different colors (n = 1-3 ROIs per color). In fossil insects, ROIs were selected from cuticle regions with different tones and the adjacent sedimentary matrix (n = 1-3 ROIs per cuticle tone and sediment). More than one ROI per cuticular tone in modern and fossil insects in addition to the sediment background accounts for replication of each traze elemental measurement.

Randomization

Concentration data were centered and log-transformed to minimize variation relating to absolute concentrations. Trends in the data were assessed using various approaches. LDA is a supervised approach that reduces the dimensions of multivariate data into linear combinations of variables to maximize separation among pre-defined data classes. Principal Component Analysis (PCA) is similar to LDA but is an unsupervised approach that does not require data categories to be defined a priori.

Blinding

All individuals in the population were blinded.

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).

Research sample

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.

Sampling strategy

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

Data exclusions

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.

Reproducibility	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.
Randomization	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.
Blinding	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	
Field Work, collec	tion 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 & 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.
Materials & experime n/a Involved in the study Antibodies Eukaryotic cell lines X Palaeontology and a X Animals and other o Clinical data Dual use research o Plants	n/a Involved in the study ChIP-seq Flow cytometry archaeology MRI-based neuroimaging organisms
Antibodies	
Antibodies used	Describe all antibodies used in the study; as applicable, provide supplier name, catalog number, clone name, and lot number.
Validation	Describe the validation of each primary antibody for the species and application, noting any validation statements on the manufacturer's website, relevant citations, antibody profiles in online databases, or data provided in the manuscript.
Eukaryotic cell lin	es
Policy information about <u>ce</u>	ell lines and Sex and Gender in Research
Cell line source(s)	State the source of each cell line used and the sex of all primary cell lines and cells derived from human participants or vertebrate models.
Authentication	Describe the authentication procedures for each cell line used OR declare that none of the cell lines used were authenticated.

Confirm that all cell lines tested negative for mycoplasma contamination OR describe the results of the testing for Mycoplasma contamination mycoplasma contamination OR declare that the cell lines were not tested for mycoplasma contamination.

Commonly misidentified lines (See ICLAC register)

Name any commonly misidentified cell lines used in the study and provide a rationale for their use.

Palaeontology and Archaeology

Specimen provenance

Modern and fossil specimens were acquired from museum collections with a broad taxonomical, geographical and age dispersation. This information is included in Table S1 (modern) and Table S5 (fossil) within the Supplementary file.

Specimen deposition

Modern specimens were collected from the life science collections of the Oxford University Museum of Natural History (UK), Natural History Museum of Oslo (Norway) and Yale Peabody Museum of Natural History (US). Specimens of two of the modern species were commerially sourced from Blades Biological (UK) and five specimens were collected by the authors of this paper at the Stanford Synchrotron Radiation Lightsource (US) grounds after finding them dead while conducting SRS-XRF data collection. Details of modern specimens can be found in Table S1 in the Supplementary file. Fossil specimens were borrowed from Museo Nacional de Ciencias Naturales, Madrid (Spain), Yale Peabody Museum of Natural History (US), Harvard Museum of Camparative Zoology (US), Russian Academy of Sciences, Moscow (Russia), Capital Normal University, Beijing (China) and University of Alaska, Fairbanks (US). Details of fossil specimens can be found in Table S5 in the Supplementary file.

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.

Ethics oversight

No ethical approval or guidance was seeked for the completion of this study because all modern specimens were sourced from already exixting museum collections and any other specimen was collected after natural death had occurred.

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Animals and other research organisms

Policy information about studies involving animals; ARRIVE guidelines recommended for reporting animal research, and Sex and Gender in Research

Laboratory animals

For laboratory animals, report species, strain and age OR state that the study did not involve laboratory animals.

Wild animals

Provide details on animals observed in or captured in the field; report species and age where possible. Describe how animals were caught and transported and what happened to captive animals after the study (if killed, explain why and describe method; if released, say where and when) OR state that the study did not involve wild animals.

Reporting on sex

Indicate if findings apply to only one sex; describe whether sex was considered in study design, methods used for assigning sex. Provide data disaggregated for sex where this information has been collected in the source data as appropriate; provide overall numbers in this Reporting Summary. Please state if this information has not been collected. Report sex-based analyses where performed, justify reasons for lack of sex-based analysis.

Specimens collected in the field had died of natural causes.

Ethics oversight

No ethical approval or quidance was seeked for the completion of this study because all modern specimens were sourced from already exixting museum collections and any other specimen was collected after natural death had occurred.

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Clinical data

Policy information about clinical studies

All manuscripts should comply with the ICMJEguidelines for publication of clinical research and a completed CONSORT checklist must be included with all submissions.

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.

Data collection

Describe the settings and locales of data collection, noting the time periods of recruitment and data collection.

Outcomes

Describe how you pre-defined primary and secondary outcome measures and how you assessed these measures.

Dual use research of concern

Policy information about dual use research of concern

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 livest	tock		
Ecosystems			
X Any other significa	ant area		
Experiments of concer	rn		
Does the work involve an	ny of these experiments of concern:		
No Yes			
Demonstrate how	to render a vaccine ineffective		
Confer resistance t	to therapeutically useful antibiotics or antiviral agents		
Enhance the virule	ence of a pathogen or render a nonpathogen virulent		
Increase transmiss	sibility of a pathogen		
Alter the host rang			
	diagnostic/detection modalities		
	nization of a biological agent or toxin		
X Any other potentia	ally harmful combination of experiments and agents		
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. I plant specimens were collected from the field, describe the collection location, date and sampling procedures.		
Novel plant genotypes Authentication	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. 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 final processed data have been deposited in a public database such as GEO.		
	e deposited or provided access to graph files (e.g. BED files) for the called peaks.		
Data access links May remain private before public	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 submiss	sion 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			
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.		
Antibodies	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

10W Cytoffictry	
Plots	
Confirm that:	
The axis labels state the mark	ker and fluorochrome used (e.g. CD4-FITC).
The axis scales are clearly visi	ible. Include numbers along axes only for bottom left plot of group (a 'group' is an analysis of identical markers).
All plots are contour plots with	th outliers or pseudocolor plots.
A numerical value for numbe	er of cells or percentage (with statistics) is provided.
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.
Gating strategy	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.
Magnetic resonance ir	naging
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 measure	es 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 & infe	rence		
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 ANO or factorial designs were used.		
Specify type of analysis:	Whole brain ROI-based Both		
Statistic type for inference	Specify voxel-wise or cluster-wise and report all relevant parameters for cluster-wise methods.		
(See Eklund et al. 2016)			
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	nnectivity Report the measures of dependence used and the model details (e.g. Pearson correlation, partial correlation mutual information).		
Graph analysis	Report the dependent variable and connectivity measure, specifying weighted graph or binarized graph, subject- or group-level, and the global and/or node summaries used (e.g. clustering coefficient, efficiency, etc.).		
Multivariate modeling and pre	edictive analysis Specify independent variables, features extraction and dimension reduction, model, training and evaluation metrics.		