

Reporting Summary

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 [Authors & Referees](#) 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
<i>Only common tests should be described solely by name; describe more complex techniques in the Methods section.</i> |
| ☐ | ☒ 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
<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 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

- ZEN 2 Pro (Carl Zeiss Microscopy GmbH, 2011; version: 2.0.14283.302; 64-bit)
- Fiji (ImageJ) (version 2.0.0 or later, <https://imagej.net/Fiji/Downloads>)
- Microsoft Excel (version 15.15 or later)
- BBEdit (version: 13 or later, <https://www.barebones.com/products/bbedit/download.html>)

Data analysis

- R (version 3.6.1 or later, <https://cran.r-project.org/mirrors.html>)
- RStudio (version 1.2.5001 or later, <https://rstudio.com/products/rstudio/download/>)
- geomorph (R package version 3.2.1, <https://cran.r-project.org/package=geomorph>)
- Morpho (R package version 2.8, <https://cran.r-project.org/web/packages/Morpho/index.html>)
- cluster (R package version 2.1.0, <https://cran.r-project.org/web/packages/cluster/index.html>)
- mclust (R package version 5.4.6, <https://cran.r-project.org/web/packages/mclust/index.html>)
- factoextra (R package version 1.0.7, <https://cran.r-project.org/web/packages/factoextra/index.html>)
- rptR (R package version 0.9.22, <https://cran.r-project.org/web/packages/rptR/index.html>)

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/reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research [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 list of figures that have associated raw data
- A description of any restrictions on data availability

All data generated and analyzed in this study is included in the paper or its supplements. Upon reasonable request, the raw data that was used to generate the example results is available from the corresponding authors. The supplementary data files 1-3 contain raw data (i.e. landmark coordinates of example data sets).

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](https://www.nature.com/documents/nr-reporting-summary-flat.pdf)

Ecological, evolutionary & environmental sciences study design

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

Study description

This manuscript provides a standardized and ready-to-use protocol for geometric morphometric and cluster analysis in microscopic organisms. GM utilizes two- or three-dimensional landmark coordinates to quantify morphological shapes. First, a set of landmark configurations containing information about the shape, size, orientation and location of homologous anatomical traits is acquired for all specimens. In the second step, general Procrustes analysis (GPA) is performed to remove all parameters of non-shape variation from the dataset. Afterwards, multivariate statistical analysis is performed on the shape coordinates to identify biologically relevant differences between categorical variables like sex, age, culture conditions or species. Finally, patterns of shape (co-)variation are described and visualized using various tools like PCA analysis or deformation grids. Additionally, we use clustering approaches in order to identify group structures in the data set and we describe group-specific allometric trends. We use various nematode species (including well establish model systems like *C. elegans* and *P. pacificus*) to generate example data sets for the proposed analysis (see below).

Research sample

Our example data sets contain (i) the two different adult morphs of the diplogastrid nematode *Pristionchus pacificus*, (ii) adults of the monomorphic diplogastrid species *Pristionchus bucculentus* and *Pristionchus elegans* and (iii) the two monomorphic rhabditid species *Caenorhabditis elegans* and *Caenorhabditis briggsae*. All biological specimen can be cultured in the lab, and came from laboratory stocks, which are available upon request. Generally only young adult hermaphrodites were considered in this study. We chose *P. pacificus* and *C. elegans* as both of these nematodes represent well-established model organisms in evo-devo studies. We chose both morphs of *P. pacificus* to show that our protocol can reliably differentiate shape differences within the same species. We chose to compare these two morphs to two monomorphic species of *Pristionchus* to show that this protocol for geometric morphometrics can reliably identify identical morphs across species (here: *P. bucculentus* and the eurytomatous morph of *P. pacificus*), as well as intermediate morphs of other species (here: *P. elegans* vs. the two morphs of *P. pacificus*). Additionally we chose *C. elegans* and compared it to *C. briggsae* as both of these model systems are described as superficially similar in morphology. This shows that, by defining other landmark configurations, the proposed protocol for geometric morphometrics and cluster analysis can easily be applied to other nematodes (or microscopic animals in general).

Sampling strategy

Each of the three example data sets was based on 80 specimens and all groups in these data sets were represented equally (e.g. 80 animals in the *Pristionchus pacificus* data set with 40 animals per morph). A general rule in geometric morphometrics is that for proper statistical analysis (using distance-based PERMANOVA on shape) one should have two to four times as many individuals as variables. Our proposed landmark sets contain 18 landmarks and therefore 36 landmark coordinates (i.e. variables). Four of these variables are lost during the GPA procedure. Thus we have 32 variables. We chose 80 individuals per example data set as this number is 2.5 times as large as the amount of variables, and is easily divisible by the number of groups in each data set (e.g. strains or species), allowing equality in group sizes.

Data collection

For microscopic imaging, all worms were mounted on object slides with 5% noble agarose pads, which contained 10 mM sodium azide to sedate the animals. All specimens were examined using a 100x DIC objective together with immersion oil. All image stacks were taken using the same settings (8.33 x 7.12 inches; 800 x 684 px; each image being 0.27 μ m thick in z direction) and saved in ".czi" or ".tif" format. All (semi-)landmarks were obtained using the multi-point and measurement tools of FIJI and saved in an Excel spreadsheet. Finally this spreadsheet was edited and converted into ".txt" file using BBEdit or a R routine (Suppl. data 6). Sampling and data acquisition was performed by TT, SSW and MSW.

Timing and spatial scale

The raw data (i.e. the DIC images) were taken in a period of time from June 2019 to August 2019. No temporal gaps occurred during sampling. The time scale is largely explained by the number of different strains or species used in the study, as we subsequently thawed frozen strains/species after finishing sampling of one strain/species. Establishment of new laboratory cultures took up to 2 weeks. Landmark data was collected from Juli to September (in parallel to ongoing sampling).

Data exclusions

No data was excluded from the analysis.

Reproducibility

The reproducibility of landmark coordinates is evident from our high biological replicate number, data visualization (PCA), and

Reproducibility	statistical inference (permanova). Nevertheless, to verify the reproducibility of our data sets, we calculated the variances of each landmark across specimen of the same species (Supp. Fig. 2). Furthermore, we estimated the repeatability and percent measurement error for shapes of <i>P. pacificus</i> using 10 randomly picked biological and 3 technical replicates per morph (generated with 20 minute breaks in between labeling worms), over a course of several days. This was done by three annotators with varying degree of training in landmark annotation (Supp. fig. 6 and 7).
Randomization	All specimens have been picked randomly from agar plates. Once all image stacks of the desired number of properly orientated individuals was reached, no more individuals were considered.
Blinding	Specimen IDs have been encrypted during the replication process.
Did the study involve field work?	☐ Yes ☒ No

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	Involvement in the study
☒	☐ Antibodies
☒	☐ Eukaryotic cell lines
☒	☐ Palaeontology
☐	☒ Animals and other organisms
☒	☐ Human research participants
☒	☐ Clinical data

Methods

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

Animals and other organisms

Policy information about [studies involving animals](#); [ARRIVE guidelines](#) recommended for reporting animal research

Laboratory animals	We used the African strain (RS5205) of <i>Pristionchus pacificus</i> , the RS5596 strain of <i>P. bucculentus</i> , the RS5720 strain of <i>P. elegans</i> , the N2 reference strain of <i>C. elegans</i> and the AF16 reference strain of <i>C. briggsae</i> . All specimens used in the study are hermaphrodites and only young adults have been considered.
Wild animals	This study did not involve wild animals.
Field-collected samples	This study did not involve animals which were collected during fieldwork.
Ethics oversight	Nematodes are not included in the animal welfare act of Germany. Therefore, no ethical approval or guidance is required to work with any of the selected strains/species.

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