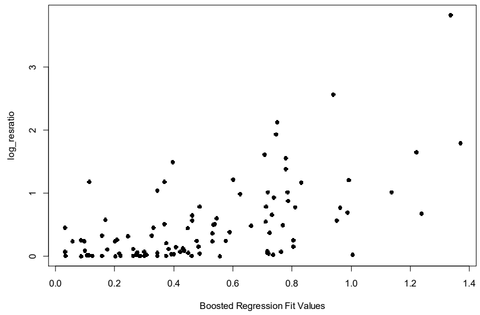
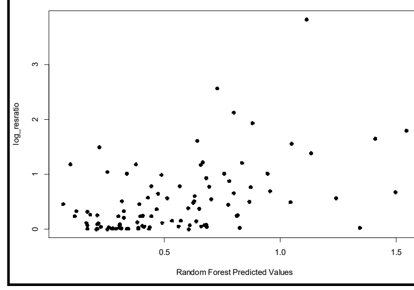
- Does BRT return a predicted group membership or value for continuous variable?
  - Yes, in the $fit object
  - In this dataset, R^2=0.346



- Is BRT the same as random forest?
  - Very similar, randomforest package in R
  - Machine/ensemble learning, Works by constructing multiple decision trees
  - Here, R^2=0.23

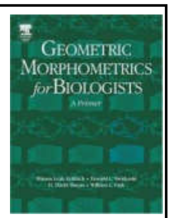


Top 5 Variables

Boosted Reg.	Random Forest
Carniv	Carniv
Migra	Migra
Med_riv	Mud
Broad	Var_sub
Var_sub	run

Shape Analyses

- Morphometrics
  - Traditional study of distances, angles or ratios.
- Long history of morphometrics in various areas of biology



Zelditch, M. L., Swiderski, D. L., Sheets, H. D. & Fink, W. L. 2004 Geometric morphometrics for biologists: A primer. San Diego, CA: Elsevier Academic Press.

Geometric Morphometrics

- Traditional morphological approach
  - Measure a series of lengths or angles from definable points

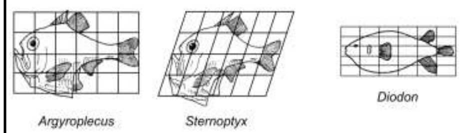
individual	Mass	tail_len	head_height	dorsal_height	head_len	snout_dorsal	girth	max_depth
BG_1	0.395247	0.429605	0.192973	0.37148	0.333247	0.421596	0.152932	0.437663
BG_2	0.423397	0.426089	0.197259	0.361233	0.323544	0.451052	0.152961	0.437956
BG_3	0.827264	0.426891	0.184874	0.36564	0.35014	0.400373	0.158543	0.411492
BG_4	1.267028	0.438864	0.215657	0.415559	0.323158	0.414574	0.177745	0.403918
BG_5	1.099403	0.409378	0.226027	0.411837	0.30137	0.431331	0.170355	0.400058
BG_6	2.535706	0.403897	0.229242	0.425079	0.32825	0.416243	0.173687	0.378885
BG_7	3.185518	0.390746	0.219366	0.432948	0.316731	0.420094	0.171915	0.368349
BG_8	3.057318	0.385773	0.206244	0.414739	0.324053	0.398158	0.15998	0.357274
BG_9	3.407385	0.433962	0.215865	0.440657	0.312031	0.419558	0.171536	0.366454
BG_10	4.617698	0.406656	0.239666	0.491752	0.35164	0.437221	0.172521	0.36621

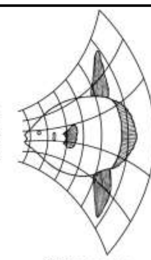
  - Apply univariate or multivariate (ordination, clustering) tests
  - Problems
    - Measures are not independent
    - Analyses of individual characters alone are not as meaningful
    - Not an analysis of **shape**

Shape

- What is shape?
  - Geometrical description of an object
  - Geometric properties of an object invariant of scale, rotation or size.
  - The geometric difference between objects after removing the effects of scale, rotation and size.

Biorthogonal Grids



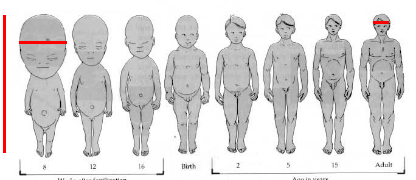


- Birth of modern analysis of shape

Thompson, D 1917 On growth and Form. Cambridge University Press, Cambridge.

- Mapped body shape onto grids
- First mathematical application to the problem of quantifying shape (quantified as warping, compressing or extending the body)

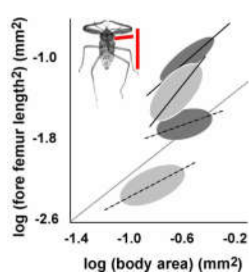
Allometry and isometry



- Changes in shape with increasing size.
- Universal pattern in most (all?) biological shape analyses.

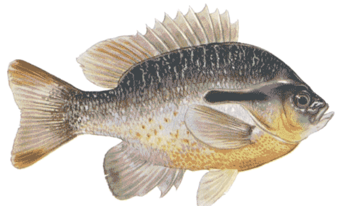
Allometry and isometry

- **Isometric** – a traits relative size or shape does not change with body size.
- Most traits are **allometric** meaning their relative size or shape changes with body size



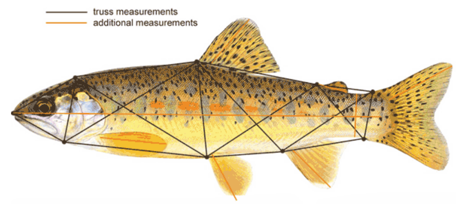
Modern morphometrics

- Approaches to analyses
  - Truss
  - Landmark
  - Outline and Surface
- Biological shapes tend to be 3D
  - Simplest applications work in 2D, math extends easily to 3D



Truss Systems

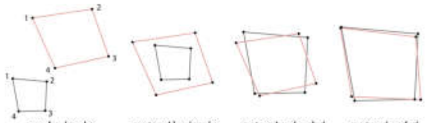
- Strauss and Bookstein (1982). The truss: body form reconstruction in morphometrics. Systematic Zoology 31:113-135.
- Measure all possible distances among series of points (should be homologous)
- Traditional univariate or multivariate statistics applied



Geometric Morphometrics

- Modern multivariate approach
- For each specimen, a series of landmarks are digitized (x,y and z coordinates)
- Procrustes analysis used to align, rotate and scale landmarks across specimens
- Produces Procrustes shape coordinates (shape variables)

Fig. 3 The three steps of Procrustes superimposition: translation to a common origin, scaling to unit centroid size, and rotation to minimize the sum of squared Euclidean distances among the homologous landmarks. The resulting landmark coordinates are called Procrustes shape coordinates



Landmarks

- Landmarks should be homologous (comparable among different taxa), fewer than # of specimens
  - **True landmark** – homologous points
  - **Semilandmark** – relative to another homologous landmark (midway between landmark A and B)
  - **Pseudolandmark** – relative positions (point of maximum body depth)

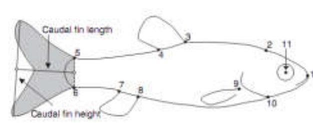


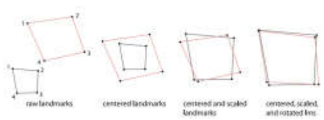
Fig. 2 Landmarks and measurements used for analysis of body morphology and caudal fin form of *Barbus neumayeri*. Open symbols represent midpoints used in the measurement of caudal fin length. The shaded region depicts the caudal fin surface area used in the calculation of caudal fin aspect ratio.

```
Landmark data – x,y
coordinates on grid of
arbitrary dimensions or
scale.
```

Analyzing shape

- **Obtain "mean shape"** – sometimes determined iteratively, sometimes called "consensus".
- **Translation** – align all points to a central frame of reference (typically the centroid of each shape)
- **Scaling** – scale all the landmarks so that distances to centroids are directly comparable among specimens
- **Rotation** – rotate all specimens so that landmarks are properly aligned
- **Landmark Configuration** – Procrustes superimposed data for an individual.

Fig. 3 The three steps of Procrustes superimposition: translation to a common origin, scaling to unit centroid size, and rotation to minimize the sum of squared Euclidean distances among the homologous landmarks. The resulting landmark coordinates are called Procrustes shape coordinates

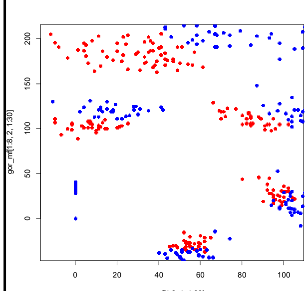


Shape Analyses in R

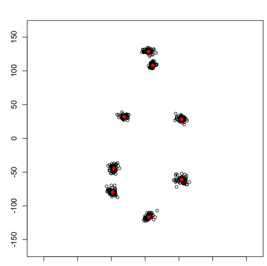
- Sample dataset – eight landmarks on gorilla skulls
  - 29 males, 30 females
- 3 dimensional arrays (gorm.dat and gorf.dat) in shapes package

```
, , 1  <----  Individual #1

     [,1] [,2]
[1,]   53  220  ┐
[2,]   46  -35  │
[3,]    0    0  │
[4,]    0   37  ├ x,y coordinates for 8 landmarks
[5,]   12  122  │
[6,]   58  204  │
[7,]   93  117  │
[8,]  103   28  ┘

, , 2

     [,1] [,2]
[1,]   57  219
[2,]   50  -43
[3,]    0    0
[4,]    0   37
[5,]   13  119
[6,]   61  198
[7,]  102  110
[8,]  104   20
```

Procrustes superimposition

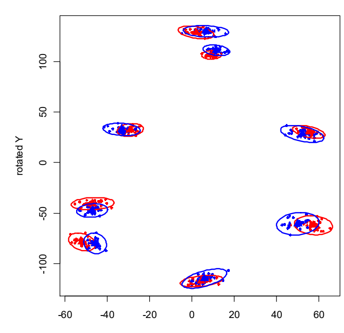


Raw data:
Male and Female gorilla data
8 landmarks, 59 individuals



Same data after Procrustes superimposition.
Red = mean shape (consensus)

Procrustes superimposition



Procrustes coordinates color coded by sex with conf. ellipse.
These are the data that we use in the PCA.
In this case, 59 individuals and 16 shape variables

Analysis of shape

- PCA (sometimes cannonical variates analysis) of shape variables.
  - The difference here is important, recall what a cannonical analysis does.
  - Matrix is samples by landmarks x2 (x,y for each landmark)
  - Details and terminology
    - PCA – relative warp analysis
    - Relative warps – PCA axes
    - Warp scores – specimens projected into ordination space
    - Bending energy – shape difference between individual and mean or consensus

Evol Biol (2009) 36:235–247
DOI 10.1007/s11692-009-9055-x

SYNTHESIS

Advances in Geometric Morphometrics

Philipp Mitteroecker · Philipp Gunz

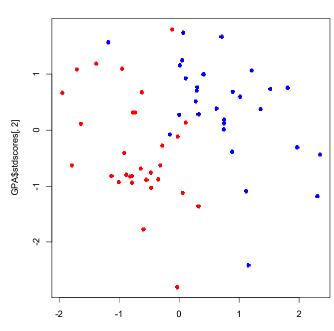
Table 1 Some key terms in geometric morphometrics and their synonyms

Bending energy	The amount of "bending" of a deformation between two landmark configurations as quantified by the TPS function
Centroid size	The measure of scale or overall size used in Procrustes analysis
Consensus configuration, mean shape, Procrustes average	Coordinate-wise average of the Procrustes coordinates
Deformation grid	A deformed regular grid illustrating the shape differences between two landmark configurations
Form	The geometric properties of an object that are invariant to rotation and translation (shape plus overall size)
Kendall shape space	The (nonlinear) space induced by a set of shape coordinates
Procrustes distance	Euclidian distance between two configurations of Procrustes coordinates as a metric measure of shape difference
Procrustes form space, size-shape space	The space induced by the Procrustes shape coordinates augmented by the natural logarithm of Centroid Size
Procrustes residuals	Deviation of a configuration of Procrustes coordinates from the consensus configuration
Procrustes shape coordinates	Landmark coordinates after Procrustes superimposition representing the shape of an object
Procrustes superimposition, Procrustes fit, GPA	Superimposition of landmark configurations to compute shape coordinates
Relative warp analysis	Principal component analysis of Procrustes shape coordinates
Semilandmarks, sliding landmarks	Landmarks on smooth curves/surfaces with positions along the curvature that cannot be identified and that are thus estimated
Shape	The geometric properties of an object that are invariant to scale, rotation, and translation
Shape regression	Multivariate regression of the shape coordinates on some other variable
Singular warp analysis	Partial least squares analysis of Procrustes shape coordinates
Tangent space	Euclidean approximation of Kendall shape space
Thin-plate spline (TPS)	An interpolation function modeling the differences between two shapes as a smooth deformation. It is used to compute deformation grids

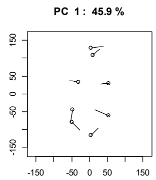
Common Visualization Techniques

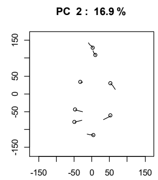
- Ordination of Procrustes superimposition scores
  - Individuals plotted in morphological space
- Thin Plate Spline (TPS)
  - A "transformation grid" comparison of two configurations (specimens) or two arbitrary points along PCA axes.
    - Compare mean configuration for two taxa or experimental groups
    - Compare individuals to the mean configuration

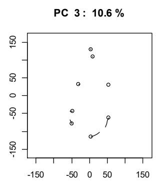
- PCA of shape variables (warp scores)
- Male and female gorilla skulls



PCA – how the 8 landmarks change in PCA space





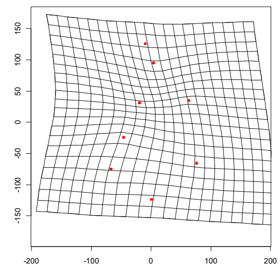


These result from the loadings (eigenvectors).

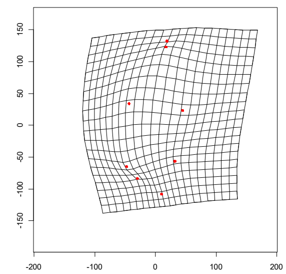
Each plot represents how the eight landmarks change across that axis.

Vector length ~ strength of relationship.

TPS Grids

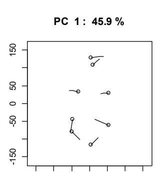


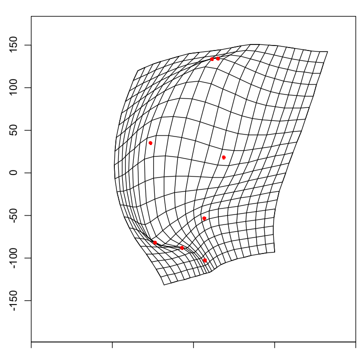
Consensus vs. mean female shape
Magnified 6 times



Consensus vs. mean female shape
Magnified 6 times

TPS Grids

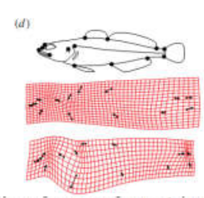




Mean male vs. mean female shape magnified 6 times.

Males and females differed most on the first PCA axis.

Common application

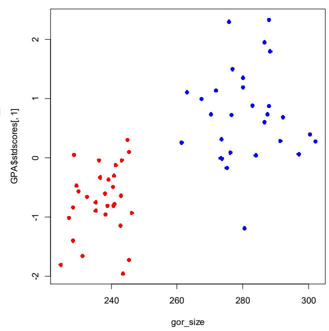


Grid deformation among two groups (species in this case). Vectors at each landmark also represented, same interpretation.

(*d*) The shape change captured by the CVA axes plotted as vectors at 14 anatomical landmarks (depicted in upper panel) on deformation grids (middle panel: invasive sculpins vs Rhine sculpins; lower panel: invasive sculpins vs Scheldt sculpins).

What about allometry?

- ProcGPA calculates centroid size for each sample.
- Should be correlated with body size.
  - Always take pictures with a scale for reference.
  - Check for correlation between centroid size and size



Strong correlation between shape and size. No overlap in size among sexes. This is a problem.

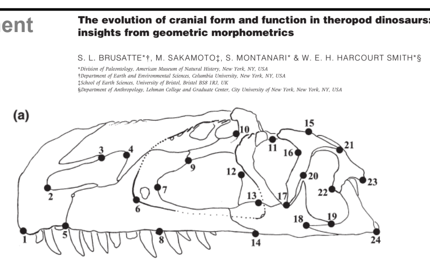
Testing for differences

- Use MANOVA, ANOSIM, MRPP or other technique to test for differences between groups of individuals in PCA space
- Allometry – need to account for changes in shape over size (sometimes this is the primary question)
  - Include separate size measure as covariate (MANCOVA)
  - If scale of all images is the same (tripod with constant zoom setting), can use centroid size.

Assignment

- Reading: none in text (review PCA, Procrustes)
- Langerhans, R. B., Layman, C. A., Shokrollahi, A. & DeWitt, T. J. 2004 Predator-driven phenotypic diversification in Gambusia affinis. *Evolution* 58, 2305–2318.
- Mitteroecker, P. & Gunz, P. 2009 Advances in Geometric Morphometrics. *Evol Biol* 36, 235–247.

Assignment



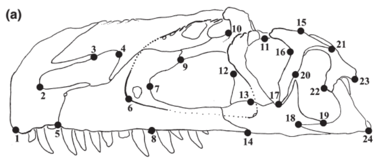
- 51 dinosaur skulls digitized at 24 landmarks
  - Data analyzed in two datasets, some missing data
    - 26 species at all 24 landmarks
    - 36 species at 13 landmarks
  - I've put together a starter script that loads the authors data, and parses it into their two datasets. You will need dino_gm.tps, dino_gm.R files.
  - I also put together files with the taxonomic groupings for the two dataset.
    - dino_taxonomy_25sp.csv, and dino_taxonomy_36sp.csv

Assignment

The evolution of cranial form and function in theropod dinosaurs: insights from geometric morphometrics

S. L. BRUSATTE*†, M. SAKAMOTO‡, S. MONTANARI* & W. E. H. HARCOURT SMITH*§

*Division of Paleontology, American Museum of Natural History, New York, NY, USA
†Department of Earth and Environmental Sciences, Columbia University, New York, NY, USA
‡School of Earth Sciences, University of Bristol, Bristol BS8 1RJ, UK
§Department of Anthropology, Lehman College and Graduate Center, City University of New York, New York, NY, USA



- Pick one dataset
- Do a geometric morphometric analysis (procGPA)
- Plot all individuals in 2D PCA space, with symbols for taxonomic group
- Describe shape changes along the first two axes
- Produce a wire diagram showing changes from the mean (concensus) to Coelophysis (tpsgrid).