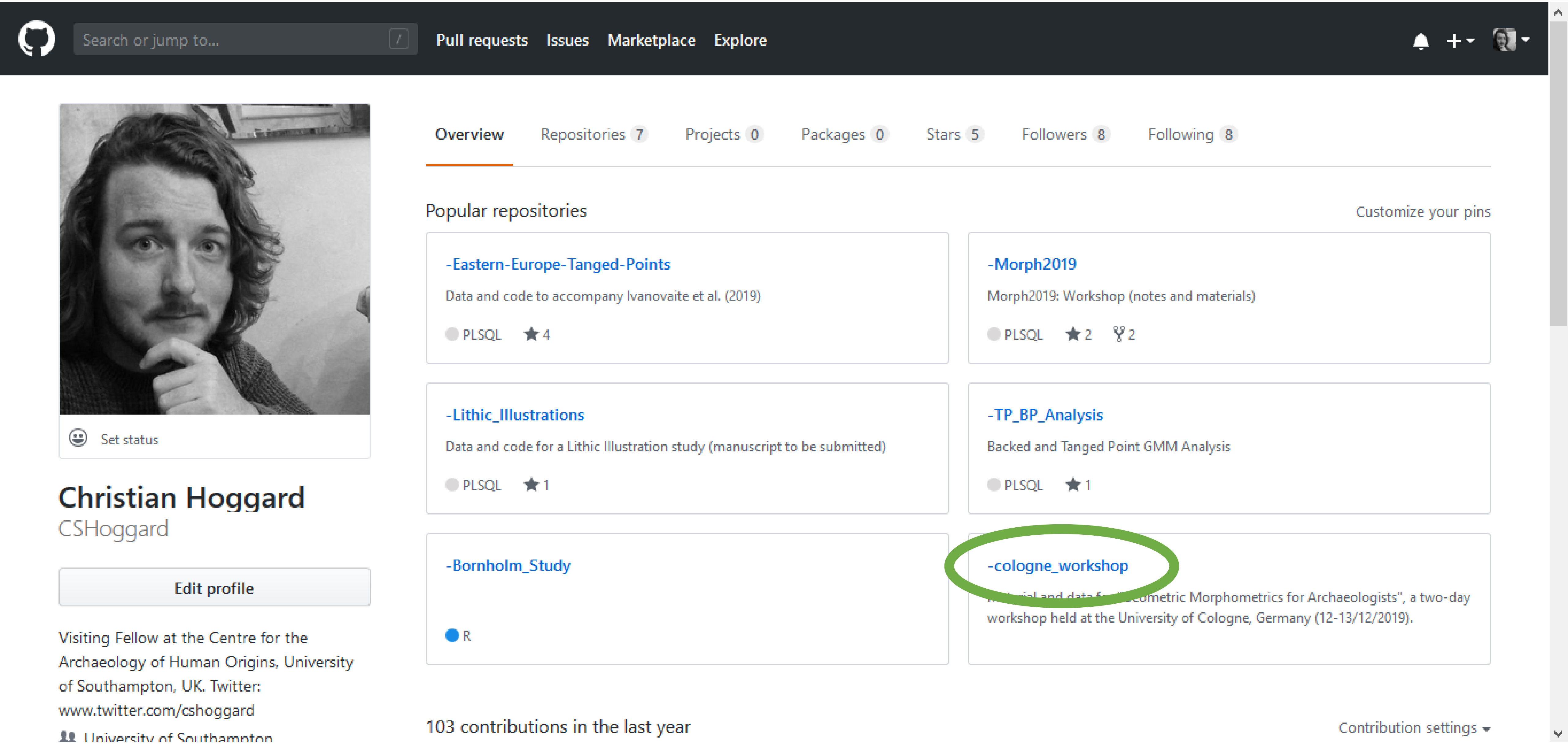
Geometric Morphometrics for Archaeologists

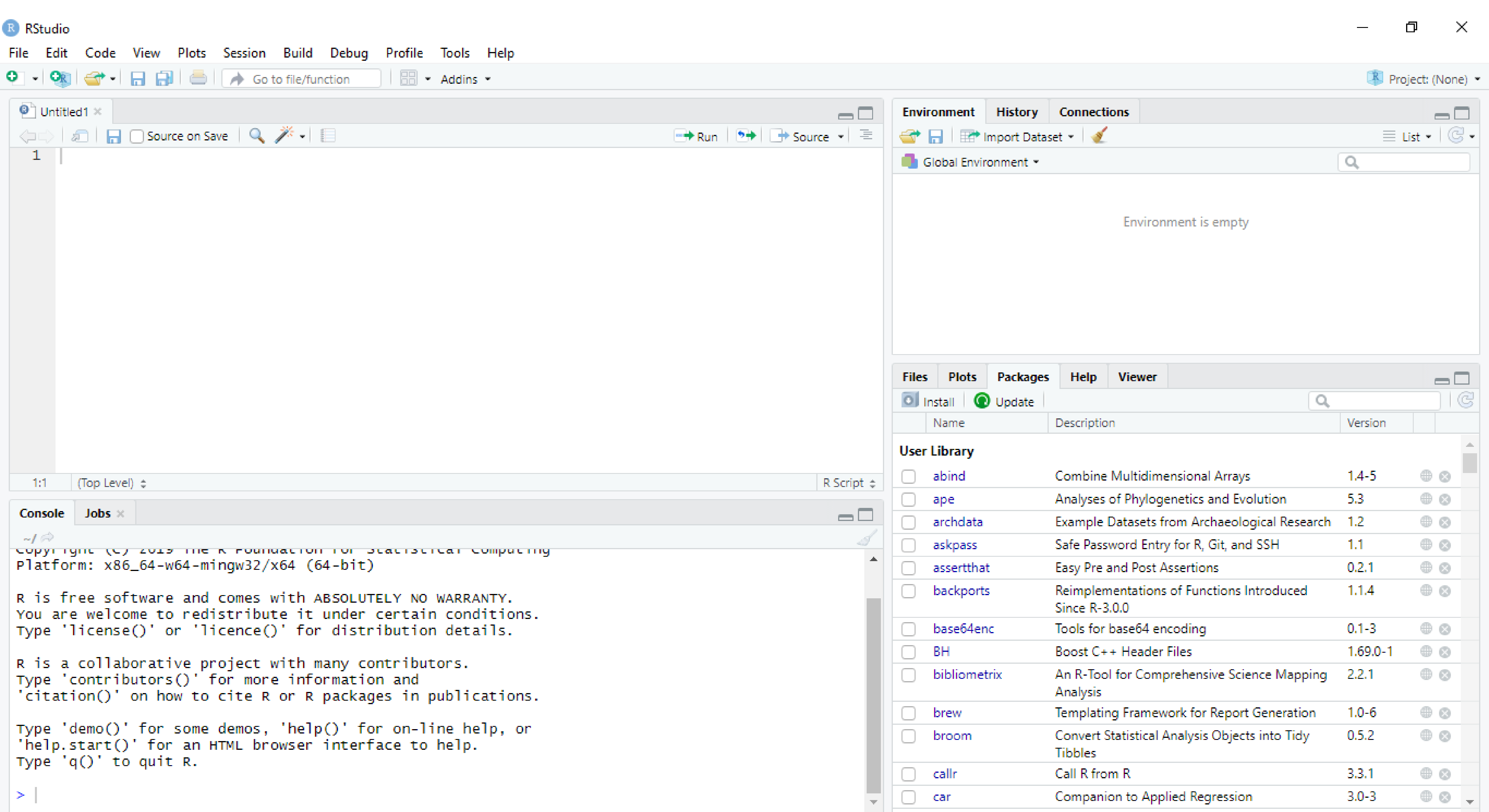
Dr. Christian Hoggard

CSHoggard

Welcome!

1. Get connected to the Wi-Fi
2. Find the workshop materials
   https://github.com/CSHoggard/-cologne_workshop
3. Ensure R/Rstudio are on your laptop
4. Stuck? Please ask for help!


Search or jump to...
Pull requests
Issues
Marketplace
Explore
Set status
Christian Hoggard
CSHoggard
Edit profile
Visiting Fellow at the Centre for the Archaeology of Human Origins, University of Southampton, UK. Twitter: www.twitter.com/cshoggard
University of Southampton
Overview
Repositories 7
Projects 0
Packages 0
Stars 5
Followers 8
Following 8
Popular repositories
Customize your pins
-Eastern-Europe-Tanged-Points
Data and code to accompany Ivanovaite et al. (2019)
PLSQL 4
-Morph2019
Morph2019: Workshop (notes and materials)
PLSQL 2 2
-Lithic_Illustrations
Data and code for a Lithic Illustration study (manuscript to be submitted)
PLSQL 1
-TP_BP_Analysis
Backed and Tanged Point GMM Analysis
PLSQL 1
-Bornholm_Study
R
-cologne_workshop
ometric Morphometrics for Archaeologists", a two-day workshop held at the University of Cologne, Germany (12-13/12/2019).
103 contributions in the last year
Contribution settings


RStudio
File Edit Code View Plots Session Build Debug Profile Tools Help
Go to file/function
Addins
Project: (None)
Untitled1
Source on Save
Run
Source
1
1:1
(Top Level)
R Script
Console
Jobs
~/
Platform: x86_64-w64-mingw32/x64 (64-bit)
R is free software and comes with ABSOLUTELY NO WARRANTY.
You are welcome to redistribute it under certain conditions.
Type 'license()' or 'licence()' for distribution details.
R is a collaborative project with many contributors.
Type 'contributors()' for more information and
'citation()' on how to cite R or R packages in publications.
Type 'demo()' for some demos, 'help()' for on-line help, or
'help.start()' for an HTML browser interface to help.
Type 'q()' to quit R.
>
Environment
History
Connections
Import Dataset
List
Global Environment
Environment is empty
Files
Plots
Packages
Help
Viewer
Install
Update
Name
Description
Version
User Library
abind Combine Multidimensional Arrays 1.4-5
ape Analyses of Phylogenetics and Evolution 5.3
archdata Example Datasets from Archaeological Research 1.2
askpass Safe Password Entry for R, Git, and SSH 1.1
assertthat Easy Pre and Post Assertions 0.2.1
backports Reimplementations of Functions Introduced Since R-3.0.0 1.1.4
base64enc Tools for base64 encoding 0.1-3
BH Boost C++ Header Files 1.69.0-1
bibliometrix An R-Tool for Comprehensive Science Mapping Analysis 2.2.1
brew Templating Framework for Report Generation 1.0-6
broom Convert Statistical Analysis Objects into Tidy Tibbles 0.5.2
callr Call R from R 3.3.1
car Companion to Applied Regression 3.0-3

Dr. Christian S. Hoggard

Position: Visiting Fellow (University of Southampton)
Previous: Assistant Professor / Postdoctoral Researcher (Aarhus University)

Interests: R Stats / GMM (2D) / Quantitative Archaeology / Palaeolithic Artefact Studies

Recent publications

Ivanovaite, L., Swertka, K., Hoggard, C.S., Sauer, F. and Riede, F. (accepted). All these fantastic cultures? Research history and regionalisation in the Late Palaeolithic tanged point cultures of Eastern Europe. European Journal of Archaeology.

Vestergaard, C. and Hoggard, C.S. (2019). A Novel Geometric Morphometric (GMM) Application to the Study of Bronze Age Tutuli. Danish Journal of Archaeology.

Riede, F., Hoggard, C.S. and Shennan, S. (2019). Reconciling material cultures in archaeology with genetic data requires robust cultural evolutionary taxonomies. Nature: Palgrave Communications..

Hoggard, C.S., McNabb, J. and Cole, J.N. (2019). The application of elliptic Fourier analysis in understanding biface shape and symmetry through the British Acheulean. Journal of Paleolithic Archaeology.

Hoggard, C.S. and Stade, C.M. (2018). The efficiency Of Middle Palaeolithic technological blade strategies: an experimental investigation. Lithics: the journal of the Lithic Studies Society.

Format for the two-day workshop…

Day One (Morning):
- Introductory lecture (What is Geometric Morphometrics? What are Landmarks? How do we do it?)
- Some basics in R (environment and generating data)
- Organising data for morphometrics
- Geometric Transformations

Day One (Afternoon):
- Procrustes Superimposition
- Other imposition methods (e.g. Resistant Fit)
- Representing shape differences (warp grids, arrows, etc.)
- Exploring shape variation and testing hypotheses (PCA, multivariate regression, MANOVA, LDA and cladistics)

Day Two (Morning):
- Acquiring outline data in R
- Fourier Analysis (principles)
- Elliptic Fourier Analysis
- Variants of Fourier Analysis
- Outline data exploration

Day Two (Afternoon):
- Advanced analytical frameworks (e.g. Maximum Likelihoods and cluster techniques)
- Advanced outline methods (symmetry and rectilinearity extraction)
- Concluding remarks

“Art of the possible”

The Sticky Note…

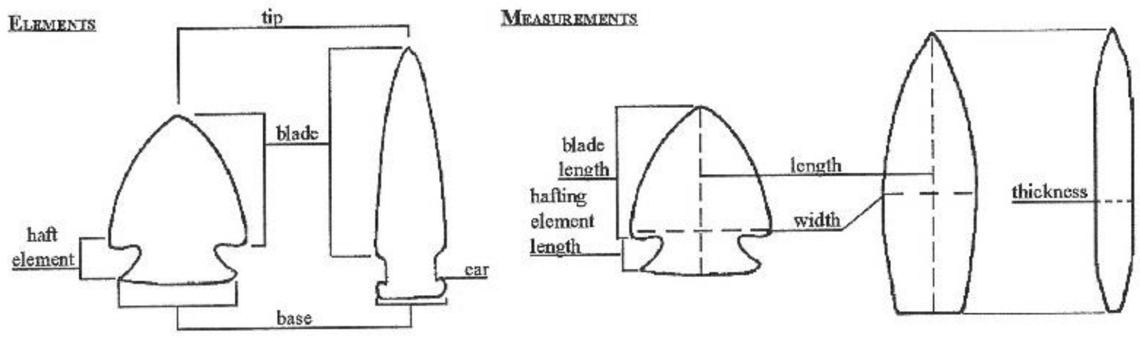
Day One…

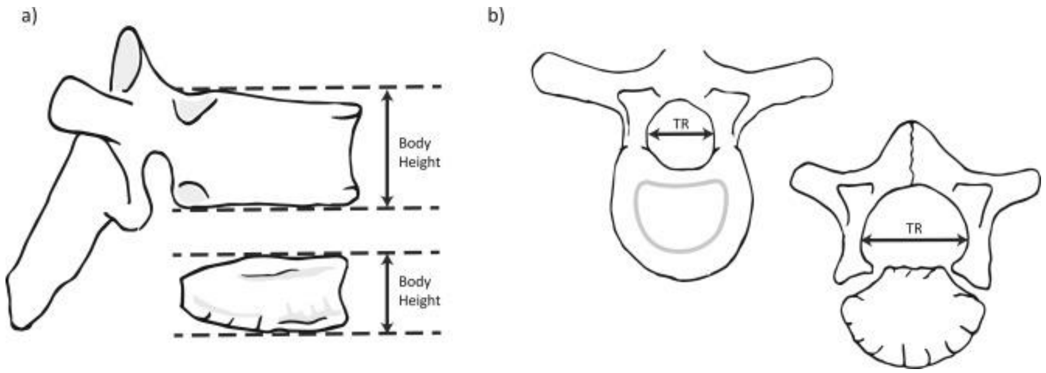
Morphometrics and Geometric Morphometrics...

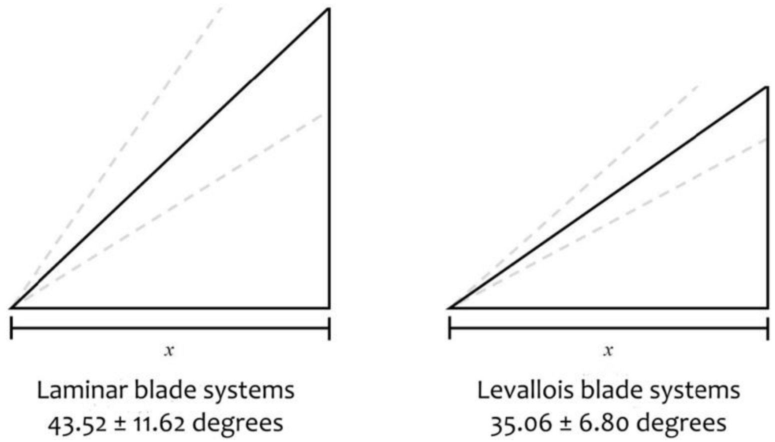
Morphometrics 101

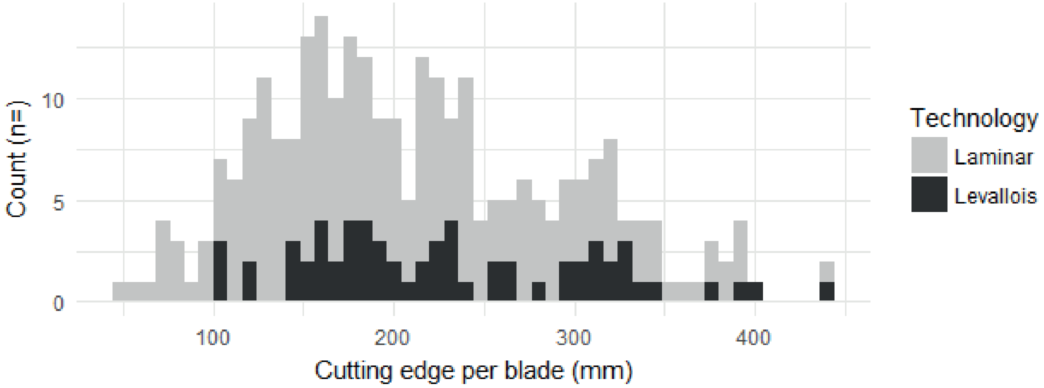
- First coined by Professor of Zoology (UCD) Robert Blackith in 1957
- **Quantitative study of shape, shape variation and shape covariation**
- Two types of morphometric studies:
  - **Traditional morphometrics** (e.g. length measurements, angles, ratios…)
  - **Geometric morphometrics** (e.g. landmarks, outlines and curves, surfaces…)
- Note: Geometric Morphometrics $\ngtr$ Traditional Morphometrics

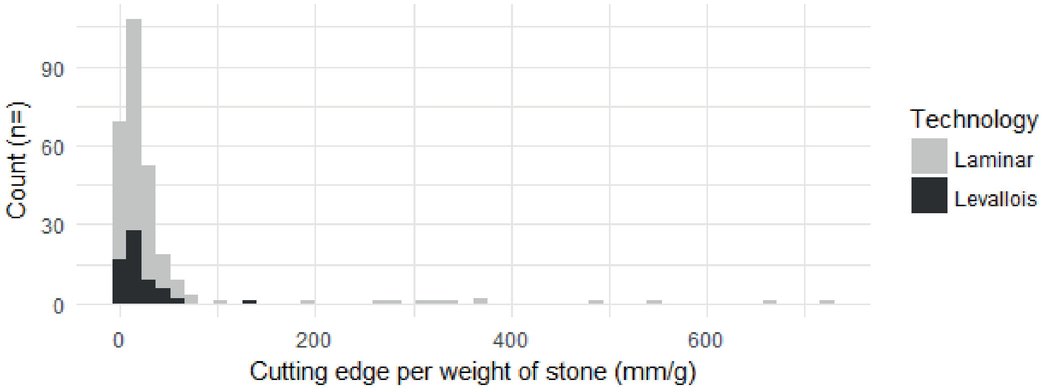
Blackith, E.R. (1957). Polymorphism in Some Australian Locusts and Grasshoppers. Biometrics., 13.: 183.

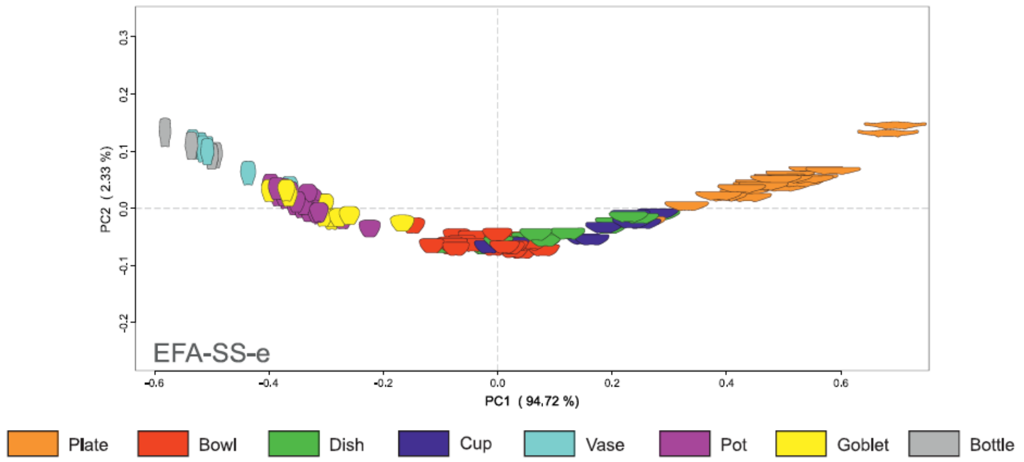

ELEMENTS
tip
blade
haft element
base
ear
MEASUREMENTS
blade length
hafting element length
length
width
thickness

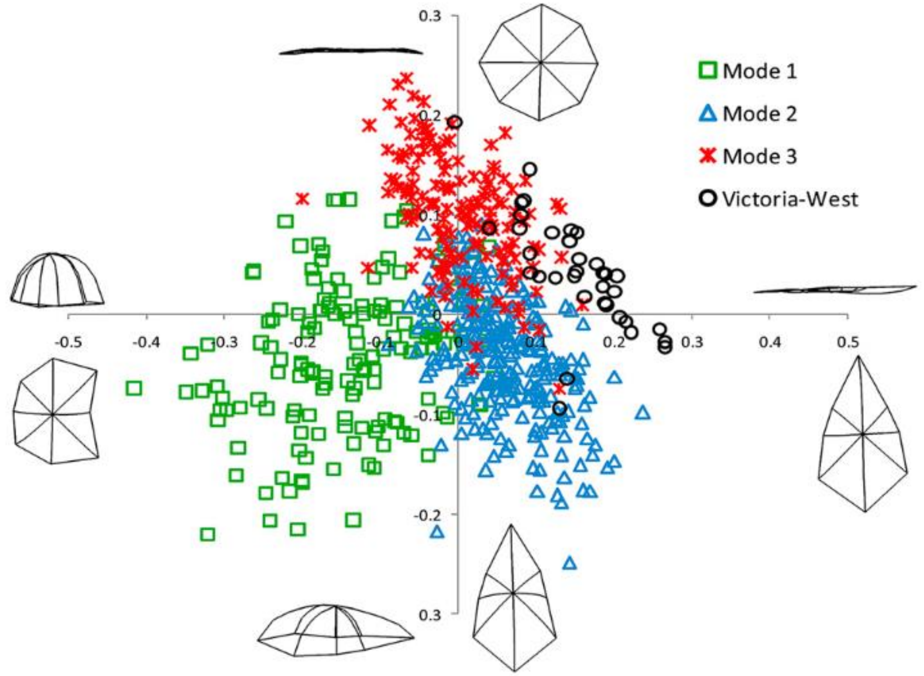

a)
Body Height
Body Height
b)
TR
TR

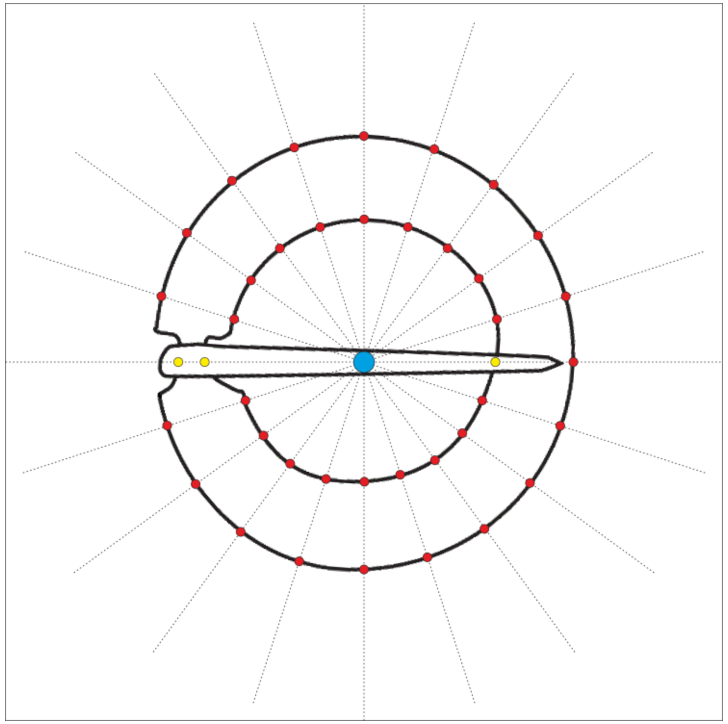

x
Laminar blade systems
43.52 ± 11.62 degrees
x
Levallois blade systems
35.06 ± 6.80 degrees

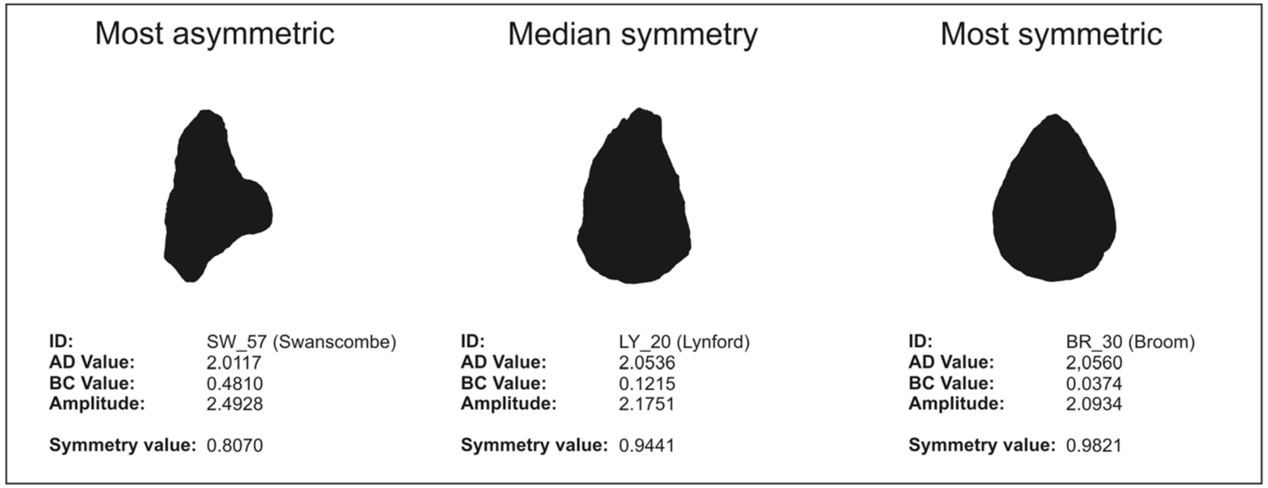

Count (n=)
10
5
0
100
200
300
400
Cutting edge per blade (mm)
Technology
Laminar
Levallois


Count (n=)
90
60
30
0
0
200
400
600
Cutting edge per weight of stone (mm/g)
Technology
Laminar
Levallois


EFA-SS-e
PC1 (94.72 %)
PC2 (2.33 %)
Plate
Bowl
Dish
Cup
Vase
Pot
Goblet
Bottle


Mode 1
Mode 2
Mode 3
Victoria-West




Most asymmetric
Median symmetry
Most symmetric
ID: SW_57 (Swanscombe)
AD Value: 2.0117
BC Value: 0.4810
Amplitude: 2.4928
Symmetry value: 0.8070
ID: LY_20 (Lynford)
AD Value: 2.0536
BC Value: 0.1215
Amplitude: 2.1751
Symmetry value: 0.9441
ID: BR_30 (Broom)
AD Value: 2,0560
BC Value: 0.0374
Amplitude: 2.0934
Symmetry value: 0.9821

What do we mean by shape and size

Shape

"In general terms, the shape of an object, dataset or image that can be defined as the **total of all information is invariant under translation, rotation, and isotropic rescalings**"

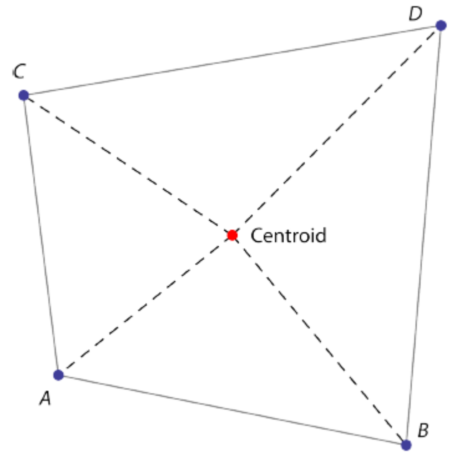
Small, C. (1996). The statistical theory of shape. New York: Springer.

Size

Vague term: what to use? Length? Circumference?

Preferred term...

Centroid size: square root of the summed squared lengths of the dashed lines



Mitteroecker, P., Gunz, P., Windhager, S., Schaefer, K. (2013). A brief review of shape, form, and allometry in geometric morphometrics, with applications to human facial morphology. Hysterix, the Italian Journal of Mammalogy. pp. 59-66.

Aim of Geometric Morphometrics

1. Determine whether two assemblages are different in terms of their shape?
2. Determine how are these related to size? time? raw material? hominin?
3. Determine whether differences correspond to a hypothesis or a model?
4. Determine on an assemblage level: what is the mean shape?
5. Determine with respect to size: is there an allometric relationship?

Advantages

Powerful method of documenting shape change

Less information is lost in comparison to lineal measurements

Can easily be collected from a variety of methods (e.g. photographs)

Abstraction and registration method permits an analysis of exclusively shape

Disadvantages

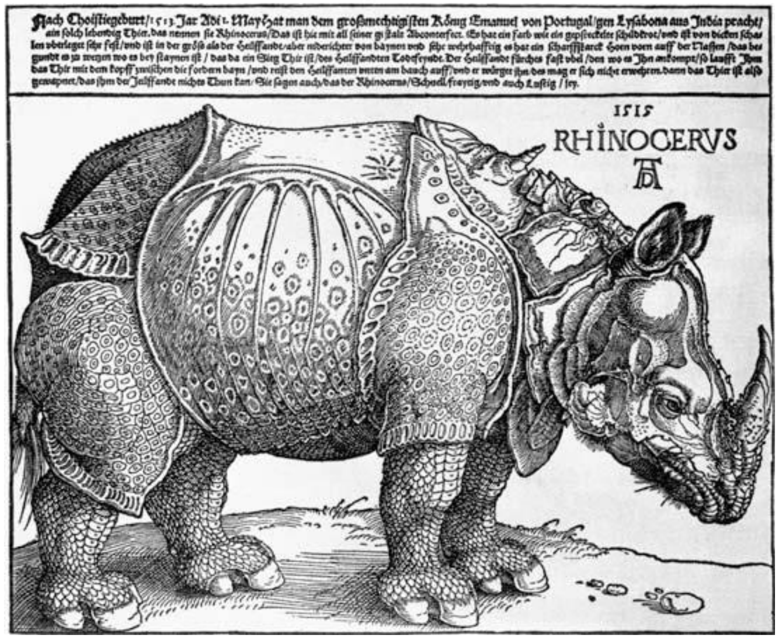
Size is often removed which may be of biological importance (can be reintegrated)

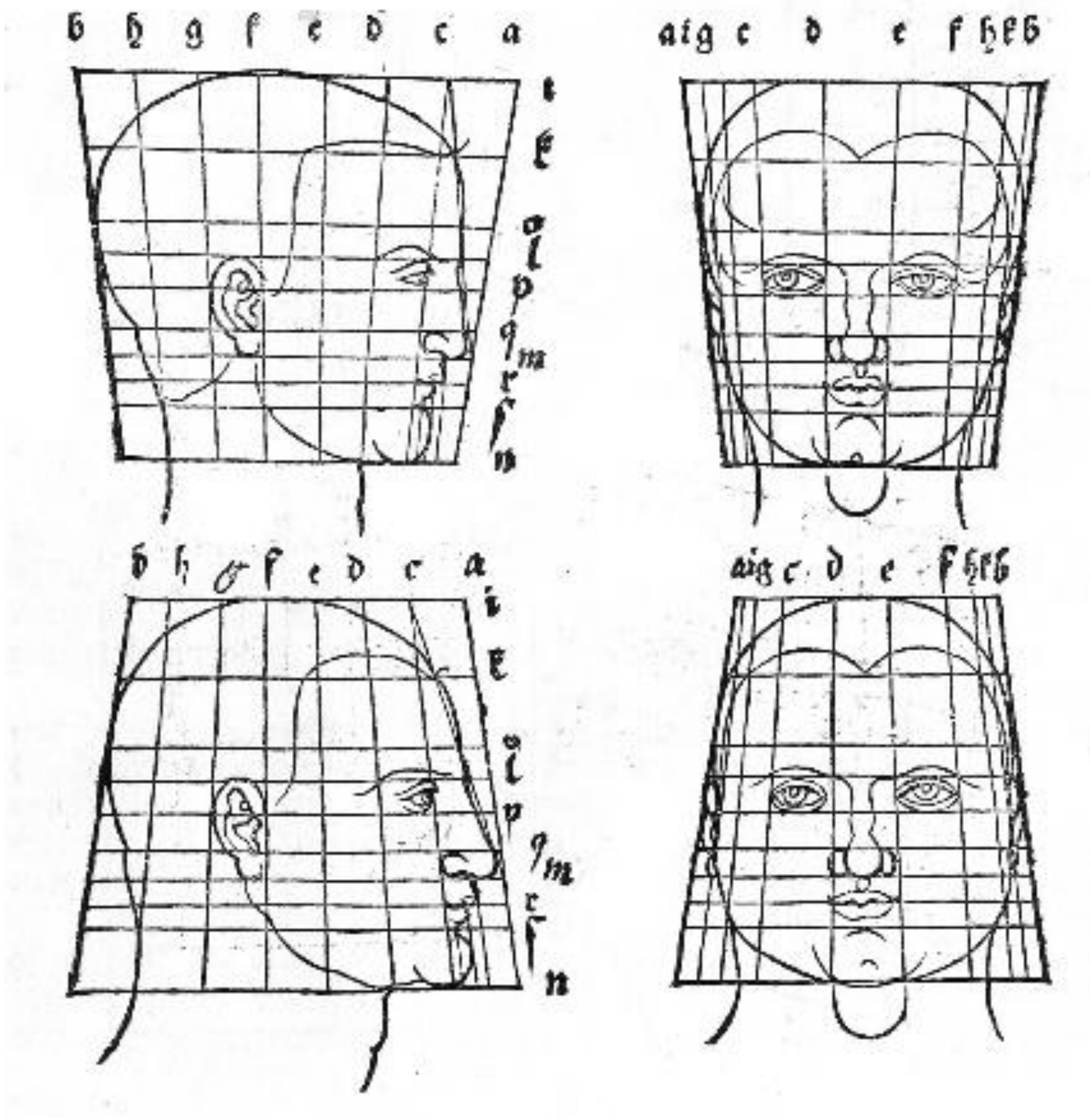
Skill competency: often requires specialised software and knowledge

A short history of geometric morphometrics...

Albrecht Dürer (1471-1528)

- Painter, printmaker and theorist
- Founder of descriptive geometry working on helices, conchoids and epicycloids
- Investigated the Delian Problem (doubling the cube)
- Used shape transformations of the human head in his studies on human proportions (providing the basis of determining shape change)





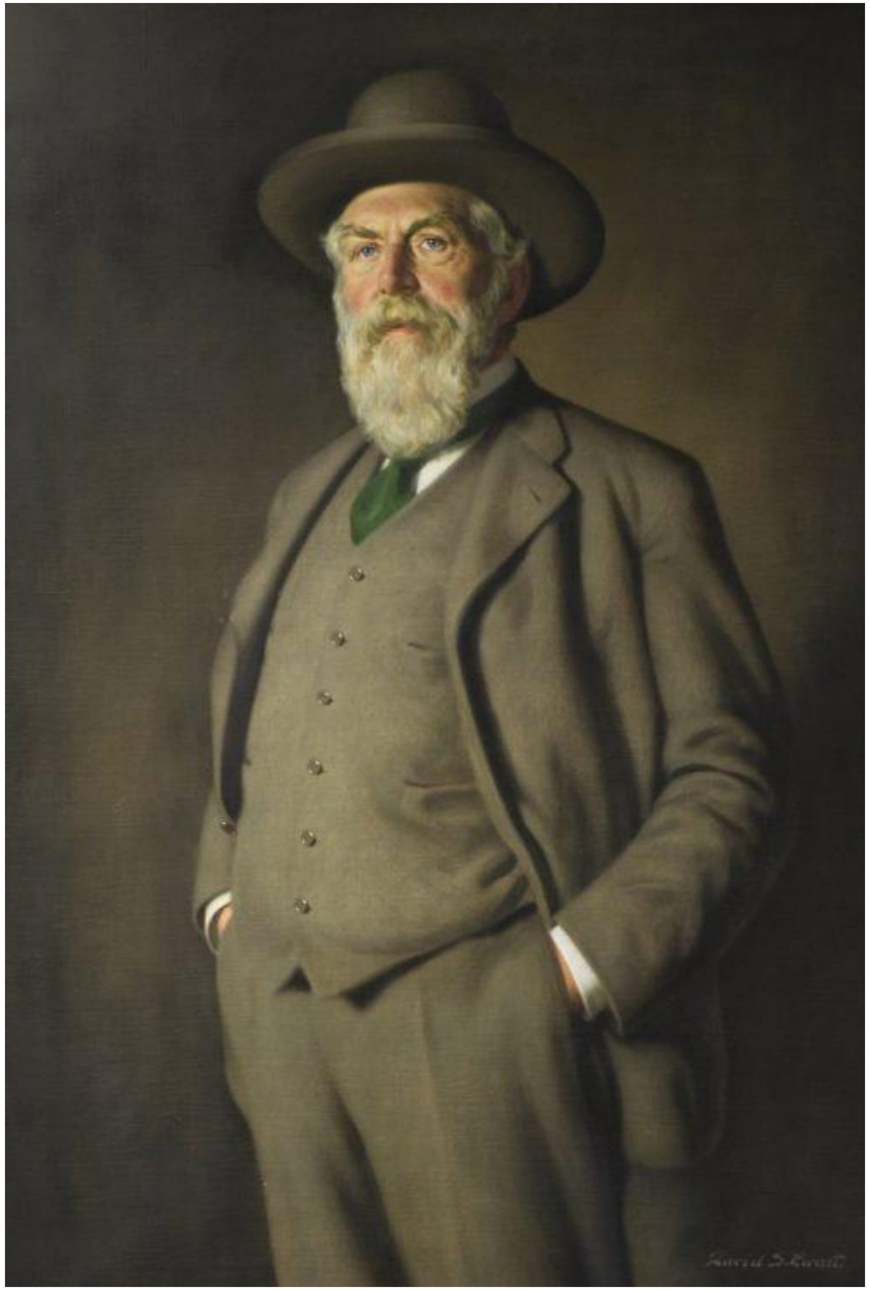
Dü rer, A. (1528). De Symetria Partium in Rectis Formis Humanorum Corporum Libri, German edn. Patav. Nuremberg: Parisiis Curvi ac recti proportio.

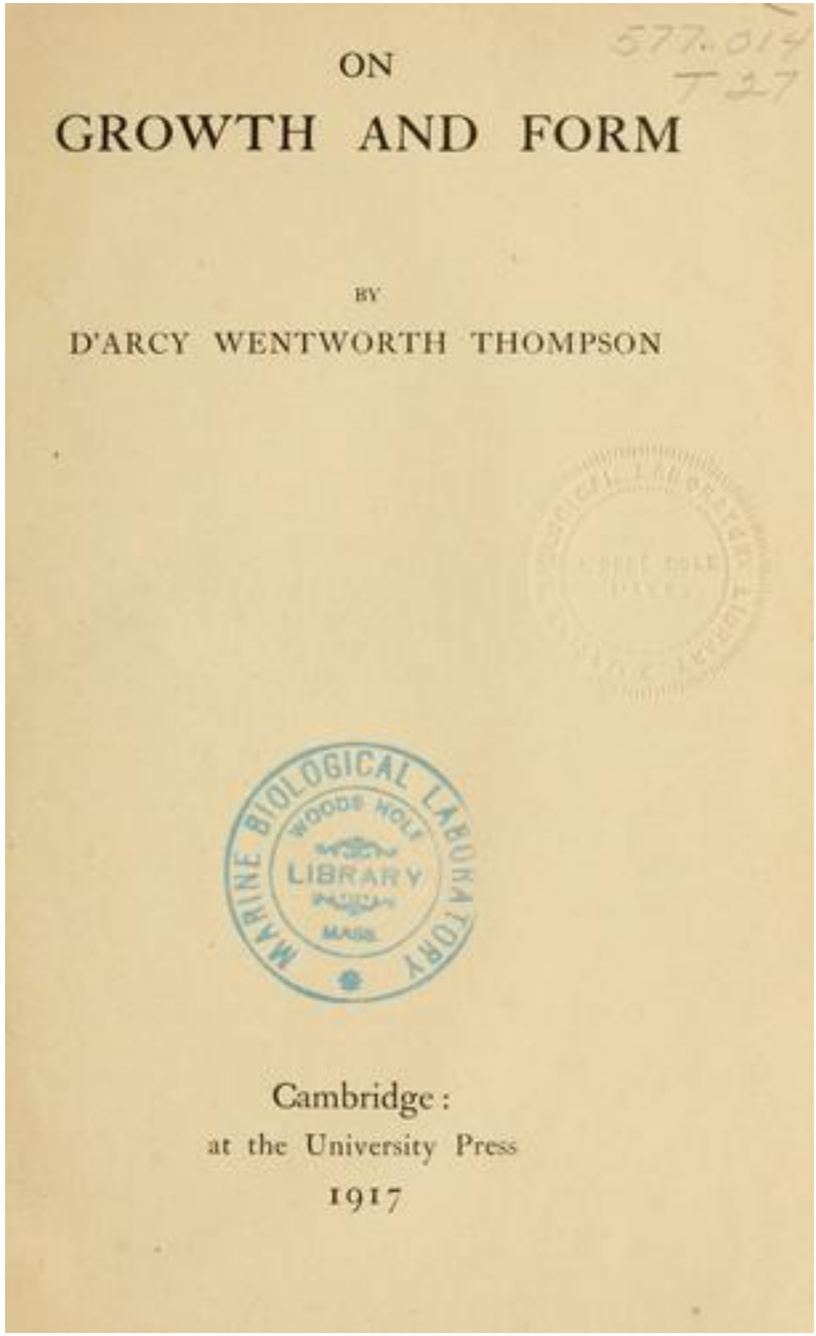
Sir D’Arcy Wentworth Thompson (1860-1948)

- Biologist, mathematician and classics scholar
- Famous for his quotes on the mathematical beauty of nature (inspired Huxley, Turing, Lévi-Straussand van der Rohe)

On Growth and Form (1917)

- Fundamental book documenting the process of body structures formed in plants and animals
- Emphasis on mathematical structures accounting for biological diversity

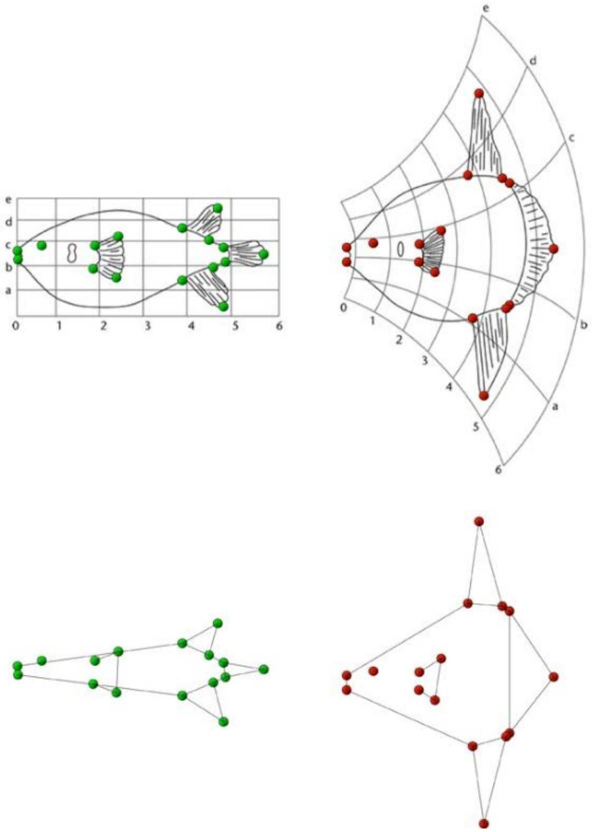




ON
GROWTH AND FORM

BY
D'ARCY WENTWORTH THOMPSON

Cambridge:
at the University Press
1917



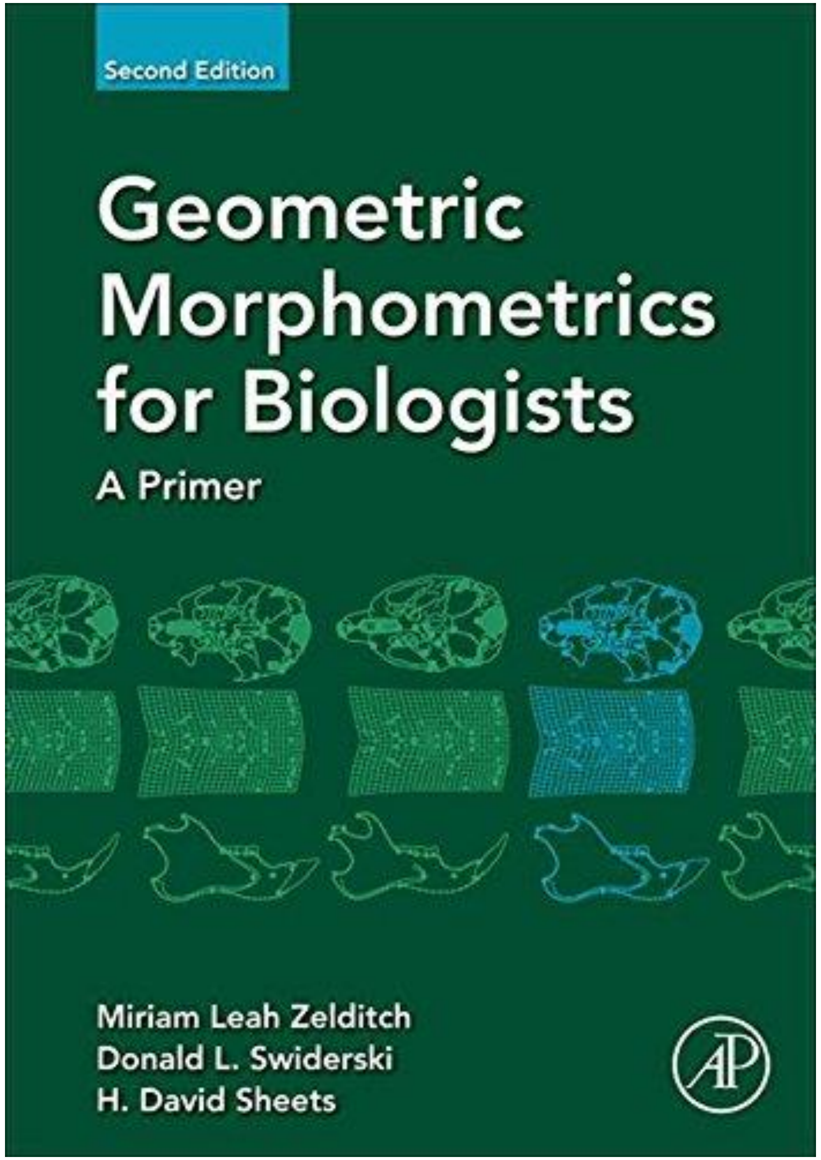
"The harmony of the world is made manifest in Form and Number, and the heart and soul and all the poetry of Natural Philosophy are embodied in the concept of mathematical beauty."

Thompson, D.W. (1917). On Growth and Form. Cambridge University Press.

Towards a statistical framework...

From the 1960s onwards...

- Fred Bookstein
- Dennis Slice
- Miriam Zelditch
- Norman MacLeod
- Ian Dryden and Kanti Mardia
- James Rohlf



How do we ‘do’ geometric morphometrics

(Following the design study...)

Stage 1: Dataset creation

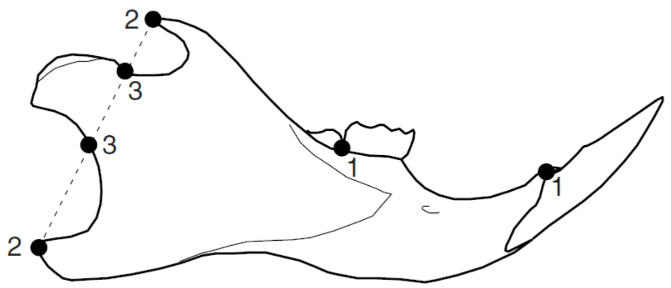
Methods include...

- **CT scanning**
- **Photogrammetry** and structure from motion (**SfM**) recording
- **Microphotogrammetry**
- **Microscribe**
- **3D scanners** (e.g. NextEngine)
- Data obtained from **drawings** and **photographs**

Note: Considered the error associated with each technique

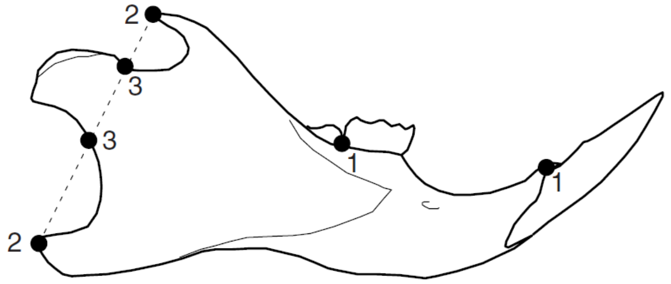
Stage 2: Landmark choice

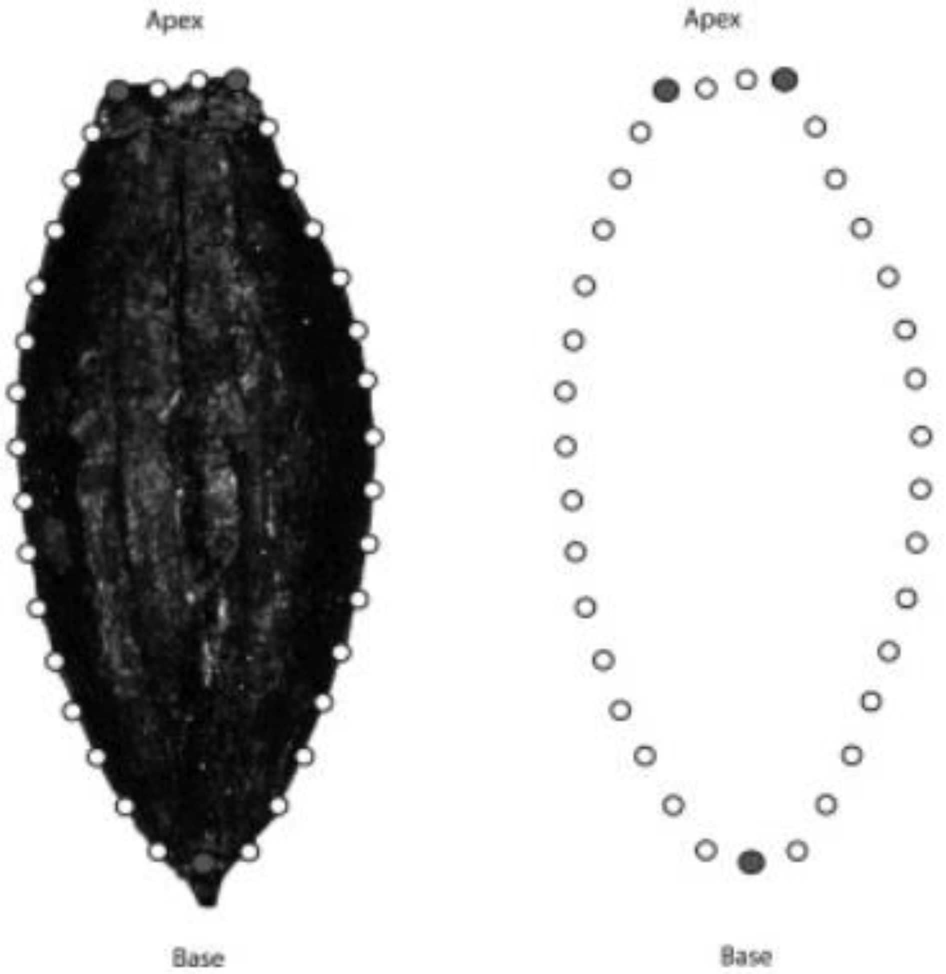
- Central to geometric morphometrics are landmarks
- **Landmark**: coordinate point used to represent a shape and/or a homologous point on a structure
- Quantifiable as Cartesian coordinates (x , y / z coordinates)
- Variety of different ways of approaching what type of landmarks are necessary
- Can be treated as individual points or converted (using various techniques) into curves and outlines



Stage 2: Landmark choice (types)

- **Various types of landmarks**
  - **Type I**: Homologous biological structures
  - **Type II**: Geometric definition e.g. greatest curvature
  - **Type III**: Point with reference to another point
- A special example: **semilandmarks**
  - Placed using an algorithm
  - Equidistant and placed between one or two end-points
  - A special **Type III landmark**
  - See also sliding semilandmarks





Ros, J., Evin, A., Bouby, L. & Marie-Pierre, R. (2013). Geometric morphometric analysis of grain shape and the identification of two-rowed barley (Hordeum vulgare subsp. distichum L.) in southern France. Journal of Archaeological Science. 41.

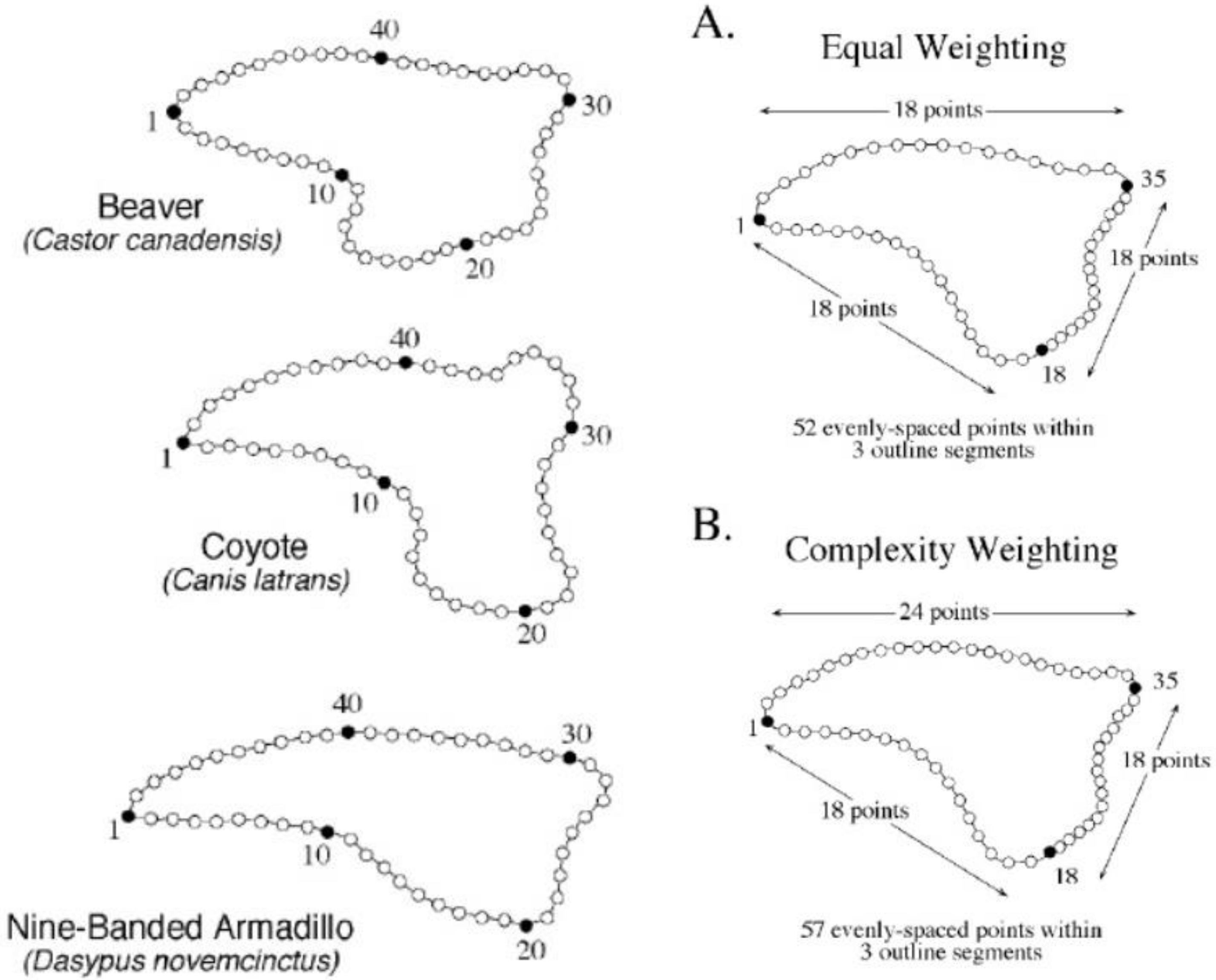
Stage 2: Landmark choice (position)

Quantity and coverage of landmarks:

- Landmarks should sample aspects of shape which are of **archaeological interest**
- Landmarks should be **repeatable** and **identifiable** on all examples if possible
- Landmarks should cover **as much of the shape as possible** (issues of distortion during GPA)
- Sufficient as to **not increase the 'weighting of areas'**
- Combination of **both landmarks and semilandmarks considered 'optimal'**

Ordering:

- Landmarks should always be **plotted in the same order** (otherwise the math wouldn't work!)



Norman MacLeod's PalaeoMath 101 Series
(http://www.palass.org/publications/newsletter/palaeomath-101)

Stage 2: Landmark choice (final comments)

For…

Bioarchaeologists

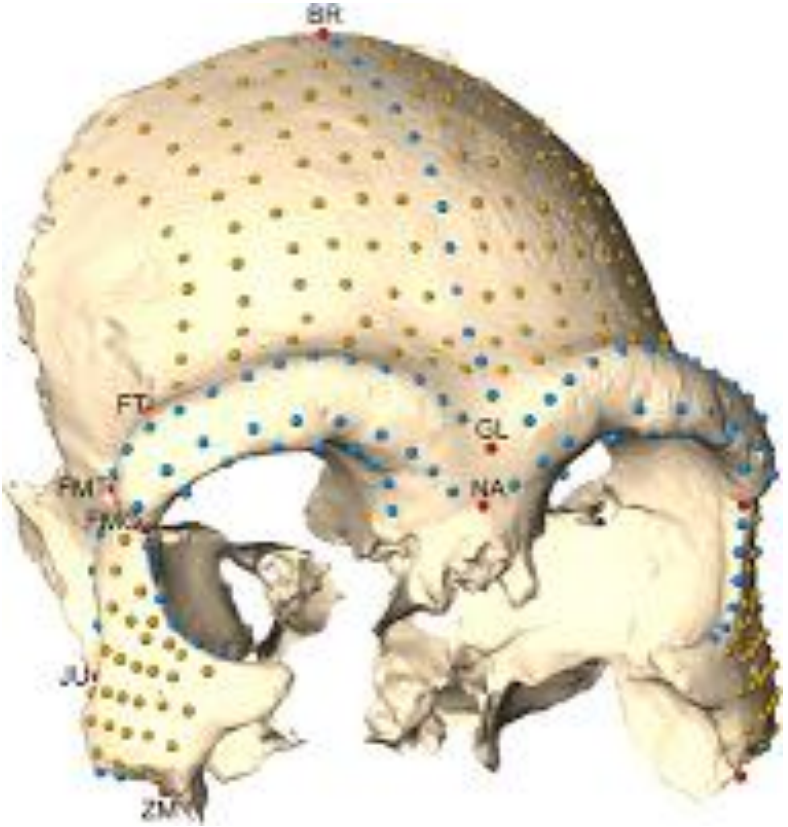
- Greater number of points of morphological correspondence on specimens
- Easier to study three-dimensional shape (orientation is less of an issue)

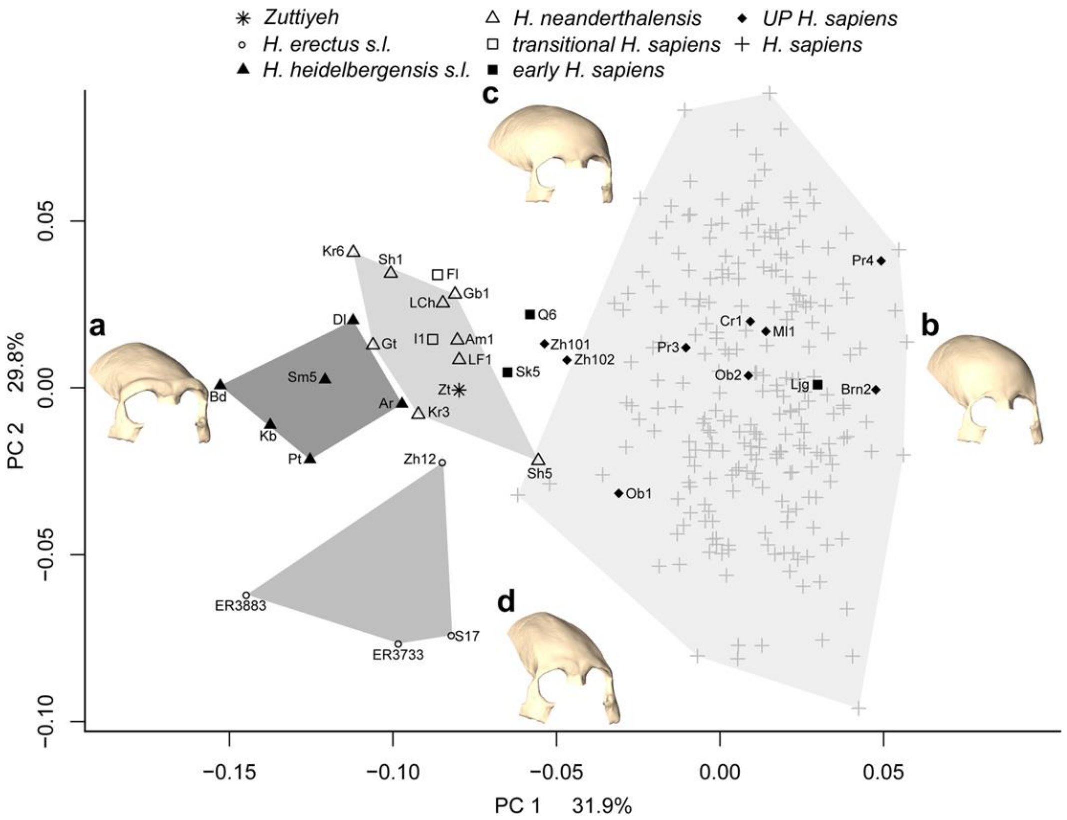
Archaeologists studying non-biological material

- Greater creativity and thought is needed in orienting specimens and placing landmarks
- Fewer case studies to compare geometric morphometric methodologies

All archaeologists

- Fragmentation and sample size (representation of a population) need to be considered

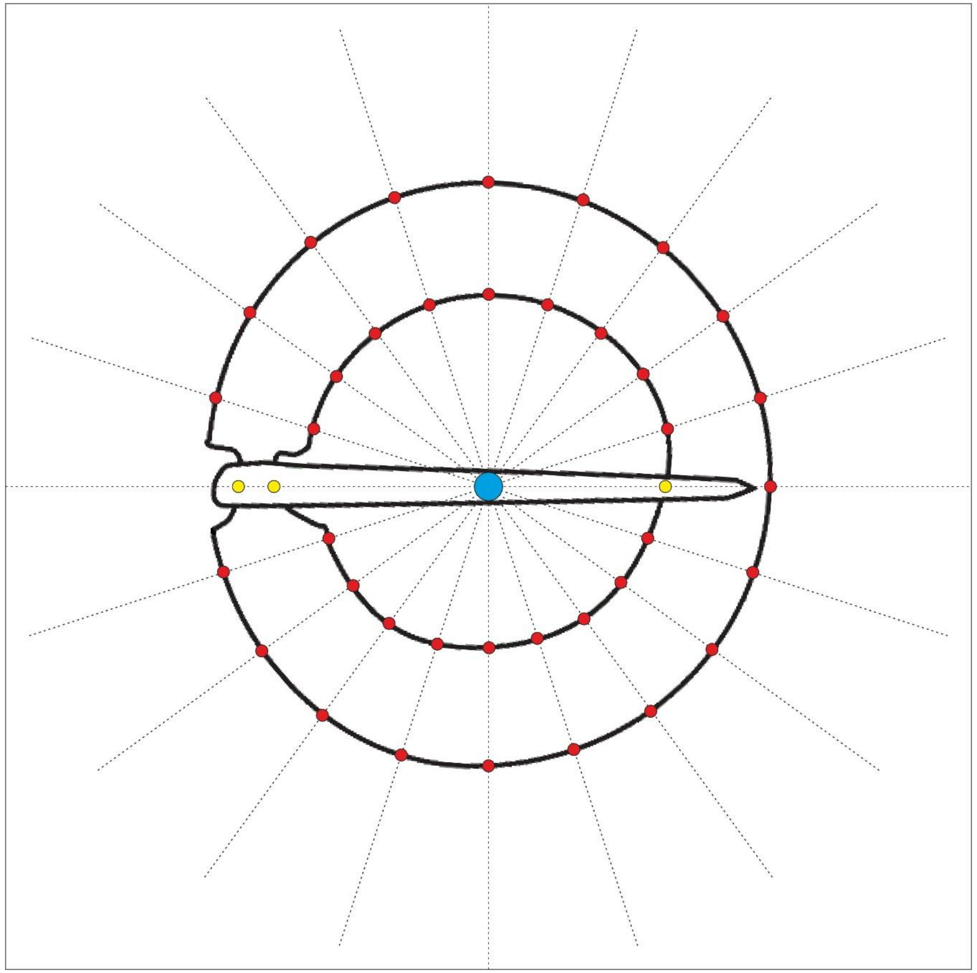




Freidline S.E, Gunz P., Janković I., Harvati K., Hublin J-J. (2012). A comprehensive morphometric analysis of the frontal and zygomatic bone of the Zuttiyeh fossil from Israel.. Journal of Human Evolution: 225-241.

Hoggard, C.S., Lauridsen, L. and Witte, K.B. (2019). The Potential of Geometric Morphometrics for Danish Archaeology: Two Case Studies. *Arkæologisk Forum,* 40: 30-42.

Article: (http://www.archaeology.dk/16738/Nr.%2040%20-%202019).
OSF: https://osf.io/en5d2/.



Stage 3: Landmark digitisation

Variety of different programs including:

- TPS Suite (TpsUtil and TpsDig2)*
- R Packages (geomorph, StereoMorph and GUImorph)*
- Landmark Editor (IDAV)
- PhyloNimbus

Variety of different output files created including:

- .tps*
- .nts
- .csv
- .text

*using in today's practicals

Stage 3: Landmark digitisation

Ask yourself...

Do all my specimens have the correct number of points?

Are all my landmarks in the correct order?

Are the ID labels correct?

Are they to scale? (For size-integrated analyses)

```
LM=28
66.00000 340.00000
72.00000 318.00000
75.00000 297.00000
78.00000 274.00000
84.00000 253.00000
86.00000 232.00000
88.00000 209.00000
83.00000 187.00000
80.00000 166.00000
76.00000 155.00000
75.00000 147.00000
57.00000 144.00000
37.00000 146.00000
34.00000 155.00000
33.00000 166.00000
32.00000 187.00000
32.00000 209.00000
30.00000 231.00000
30.00000 253.00000
30.00000 274.00000
33.00000 296.00000
37.00000 318.00000
46.00000 340.00000
50.00000 350.00000
52.00000 357.00000
57.00000 361.00000
61.00000 355.00000
63.00000 350.00000
IMAGE=Abri_Mannlefels_R.png
ID=Abri_Mannlefels_R
SCALE=0.105240
```

Two-dimensional shape analysis (a breakdown)

1. **Landmark Analysis**
2. **Outline Analysis**
    - Semi-landmark Analysis
    - Fourier-based Analysis
        - Radial Fourier Analysis (RFA)
        - Elliptic Fourier Analysis (EFA)
        - Discrete Cosine Transform (DCT)
    - Eigenshape Analysis*
3. **Other methods** e.g. Polynomial Curve Fitting*

* Not covered here

Format for the two-day workshop…

Day One (Morning):
- ~~Introductory lecture (What is Geometric Morphometrics? What are Landmarks? How do we do it?)~~
- Some basics in R (environment and generating data)
- Organising data for morphometrics
- Geometric Transformations

Day One (Afternoon):
- Procrustes Superimposition
- Other imposition methods (e.g. Resistant Fit)
- Representing shape differences (warp grids, arrows, etc.)
- Exploring shape variation and testing hypotheses (PCA, multivariate regression, MANOVA, LDA and cladistics)

Day Two (Morning):
- Acquiring outline data in R
- Fourier Analysis (principles)
- Elliptic Fourier Analysis
- Variants of Fourier Analysis
- Outline data exploration

Day Two (Afternoon):
- Advanced analytical frameworks (e.g. Maximum Likelihoods and cluster techniques)
- Advanced outline methods (symmetry and rectilinearity extraction)
- Concluding remarks

GMM in the R Environment

- With an increasing number of packages for creating, manipulating and analysing shape coordinates, and a more code-literate discipline, R is the ideal environment for archaeologists.

- Permits a transparent, repeatable and reproducible GMM analytical workflow.

- Note: you may find certain stages easier to process outside of the R environment e.g. digitisation (and that is fine!)

Exercise 1: R Environment (setup)

1. Download **R** and **Rstudio**
   R: https://www.r-project.org/
   RStudio: https://rstudio.com/products/rstudio/

2. Load up RStudio and familiarise yourself with the different windows

3. Download all files from the GitHub repository (https://github.com/CSHoggard/-cologne_workshop)

4. Set the working directory to the location of the downloaded files (Session…)

5. Run the **setup.R** code using the ‘run’ button or through the shortcut (Ctrl + Enter)

Landmark data in the R environment (examples)

Digitisation within R

- For outlines: `import_jpg` and `import_Conte` in **Momocs**
- For 2D landmarks: `digitize2d` in **Geomorph**
- For 3D landmarks: `digitsurface` in **Geomorph** / **GUImorph** (not on CRAN yet!)

- Note: `readland.ply` function in **Geomorph**
- Note: `buildtemplate` function in **Geomorph**

Importing landmark data into R

- .tps: `readland.tps` function in **Geomorph** / `import_tps` function in **Momocs**
- .nts: `readland.nts` function in **Geomorph**
- SM: `readland.shapes` function in **Geomorph**

Exercise 2: Creating/Reading Landmark Data #1

1. Load Geomorph v.3.1.3 (Beta) and Momocs in RStudio
2. Import the skull_1.ply file into RStudio using the read.ply() function and label it "SK1"
3. Use the `buildtemplate()` function to create a template for (23 LM / 200 SLM)
4. Experiment with plotting landmarks for this or any other skull using `digitsurface()`
5. Use the `readmulti.nts()` function to read in the already digitised files
6. Add in the metadata through the `read.csv()` function

Use the ? Function to read more about the functions used e.g. `?readmulti_nts()`

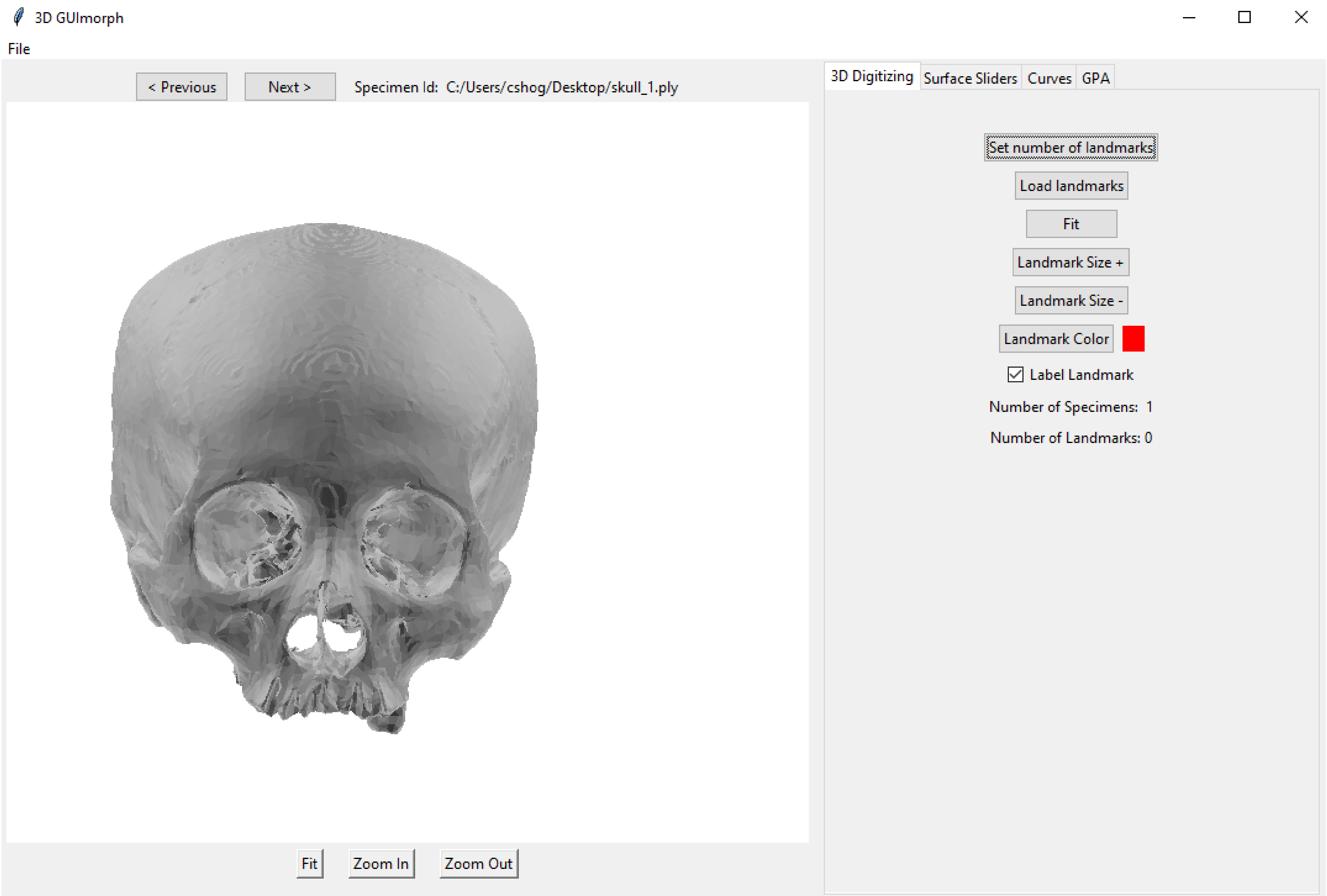
Exercise 3: Creating/Reading Landmark Data #2

1. Download tpsDig2 and tpsUtil
   tpsUtil: https://life.bio.sunysb.edu/morph/soft-utility.html
   tpsDig2: https://life.bio.sunysb.edu/morph/soft-dataacq.html

2. Collate all the Ellensbanke and Eskebjerg tanged points into one .tps file in tpsUtil using the “Build tps file from images” - save the file as “tanged.tps”

3. Load the new .tps file in TpsDig2 and place 6 landmarks at the extremities, tang extremities and shoulder mid-points

4. Save the .tps file and import into R using the `import_tps` function in Momocs


3D GUImorph
File
< Previous
Next >
Specimen Id: C:/Users/cshog/Desktop/skull_1.ply
3D Digitizing
Surface Sliders
Curves
GPA
Set number of landmarks
Load landmarks
Fit
Landmark Size +
Landmark Size -
Landmark Color
Label Landmark
Number of Specimens: 1
Number of Landmarks: 0
Fit
Zoom In
Zoom Out

Format for the two-day workshop…

Day One (Morning):

- ~~Introductory lecture (What is Geometric Morphometrics? What are Landmarks? How do we do it?)~~
- ~~Some basics in R (environment and generating data)~~
- ~~Organising data for morphometrics~~
- Geometric Transformations

Day One (Afternoon):

- Procrustes Superimposition
- Other imposition methods (e.g. Resistant Fit)
- Representing shape differences (warp grids, arrows, etc.)
- Exploring shape variation and testing hypotheses (PCA, multivariate regression, MANOVA, LDA and cladistics)

Day Two (Morning):

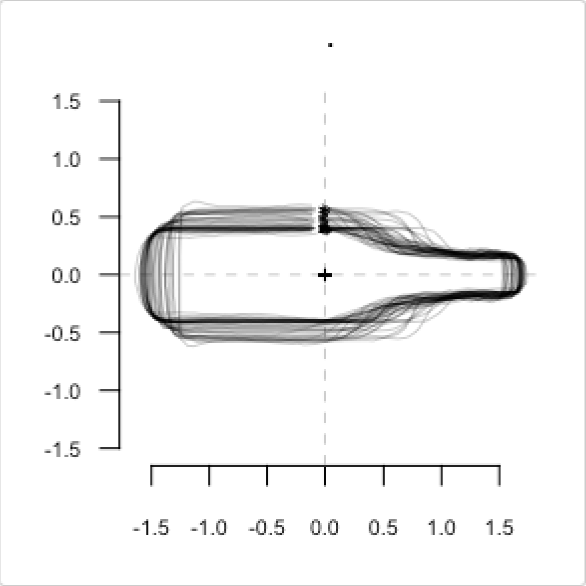
- Acquiring outline data in R
- Fourier Analysis (principles)
- Elliptic Fourier Analysis
- Variants of Fourier Analysis
- Outline data exploration

Day Two (Afternoon):

- Advanced analytical frameworks (e.g. Maximum Likelihoods and cluster techniques)
- Advanced outline methods (symmetry and rectilinearity extraction)
- Concluding remarks

Geometric Transformations

```
bot %>%
  coo_center %>% coo_scale %>%
  coo_alignxax() %>% coo_slidedirection("N") %T>%
  print() %>% stack()
```



- In a number of different packages, the landmark configurations and outlines can be transformed (for visual or analytical purposes)
- Momocs: coo_center, coo_scale etc.
- Geomorph: fixed.angle, rotate.coords

Format for the two-day workshop…

Day One (Morning):
~~• Introductory lecture (What is Geometric Morphometrics? What are Landmarks? How do we do it?)~~
~~• Some basics in R (environment and generating data)~~
~~• Organising data for morphometrics~~
~~• Geometric Transformations~~

Day One (Afternoon):
• Procrustes Superimposition
• Other imposition methods (e.g. Resistant Fit)
• Representing shape differences (warp grids, arrows, etc.)
• Exploring shape variation and testing hypotheses (PCA, multivariate regression, MANOVA, LDA and cladistics)

Day Two (Morning):
• Acquiring outline data in R
• Fourier Analysis (principles)
• Elliptic Fourier Analysis
• Variants of Fourier Analysis
• Outline data exploration

Day Two (Afternoon):
• Advanced analytical frameworks (e.g. Maximum Likelihoods and cluster techniques)
• Advanced outline methods (symmetry and rectilinearity extraction)
• Concluding remarks

Landmark analysis:
Generalised Procrustes Analysis (GPA)

Also known as...

- Procrustes Superimposition
- Procrustes Analysis
- Procrustes Fitting
- Generalised Least Squares

Procedure to isolate shape from a number of sometimes related variables, specifically **rotation, size and translation**

Generalised Procrustes Analysis (GPA)

- In Greek Mythology…
  - Procrustes was a son of Poseidon and lived on a sacred way between Athens and Eleusis
  - There he had a bed, in which he invited every passer-by to spend the night, and where he set to work on them with his blacksmith's hammer, to stretch them to fit
  - In later tellings, if the guest proved too tall, Procrustes would amputate the excess length; nobody ever fitted the bed exactly
  - Procrustes continued his reign of terror until he was captured by Theseus, travelling to Athens along the sacred way, who "fitted" Procrustes to his own bed

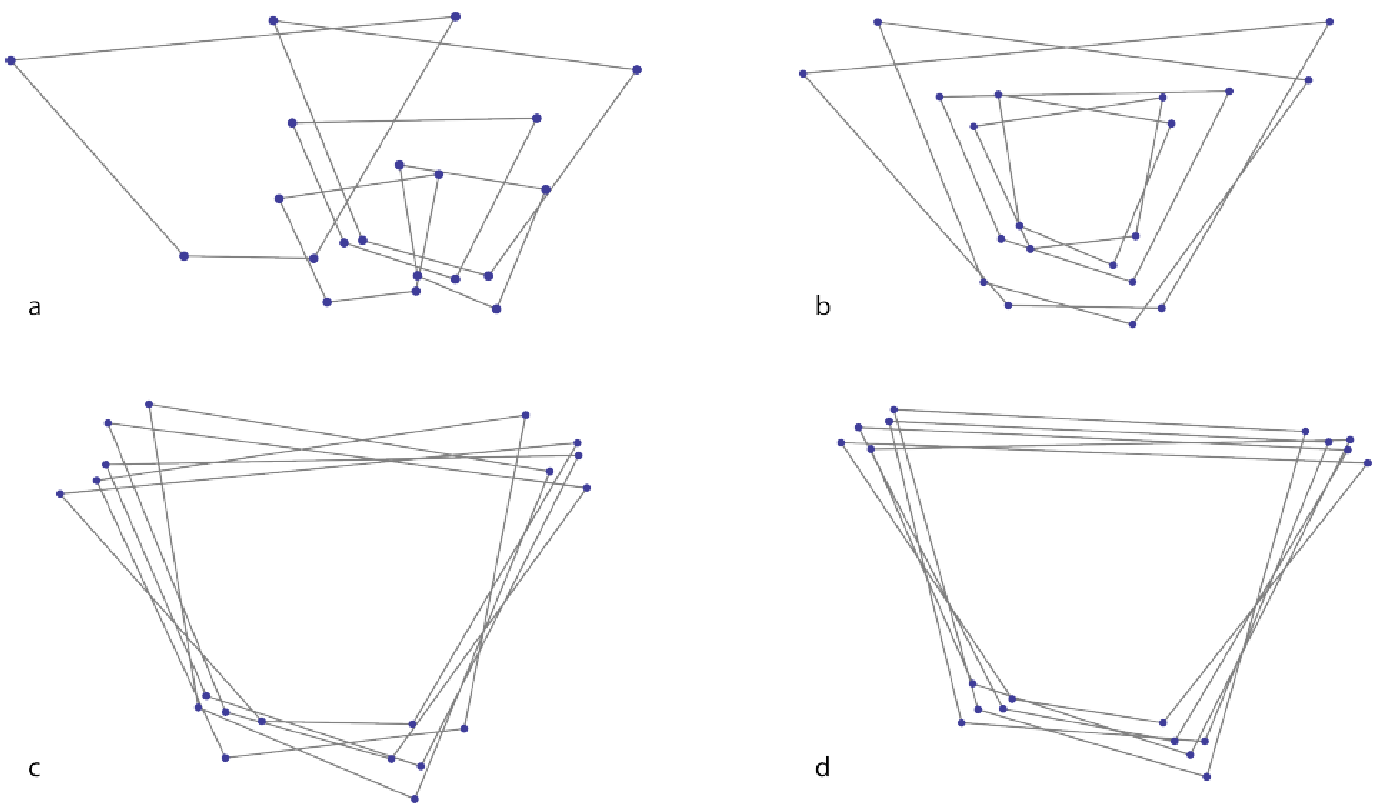
Generalised Procrustes Analysis (GPA)

- The raw coordinates are…
  - Translated to a **common centroid**
  - Scaled to the **same centroid size**
  - Rotated to **minimise the summed square distances** between corresponding landmarks

- Outcome: **Procrustes coordinates**

- The Procrustes coordinates describe shape *per se*
- The **Procrustes distance**: sum of distance between corresponding transformed landmarks
- These are used in subsequent statistical analyses

Generalised Procrustes Analysis (GPA)



Mitteroecker, P., Gunz, P., Windhager, S. & Schaefer, K. (2013). A brief review of shape, form, and allometry in geometric morphometrics, with applications to human facial morphology. Hysterix, the Italian Journal of Mammalogy. pp. 59-66.

Exercise 4: Generalised Procrustes Analysis (GPA)

1. Using both the multi.nts file (skull) and the Ldk file (tanged.points) perform a Full Generalised Procrustes alignment for all examples.

2. For the multi.nts file use the `gpagen` function in Geomorph (call `?gpagen`)

3. For the tanged.points file use the `fgProcrustes` function in Geomorph (call `?fgProcrustes`)

4. Explore visualisation styles for both gpa files (`?stack`) and Momocs (`?plot`)

Variants of the Procrustes process... (Momocs)

fgProcrustes: Full Generalized Procrustes alignment between shapes

fProcrustes: Full Procrustes alignment between two shapes

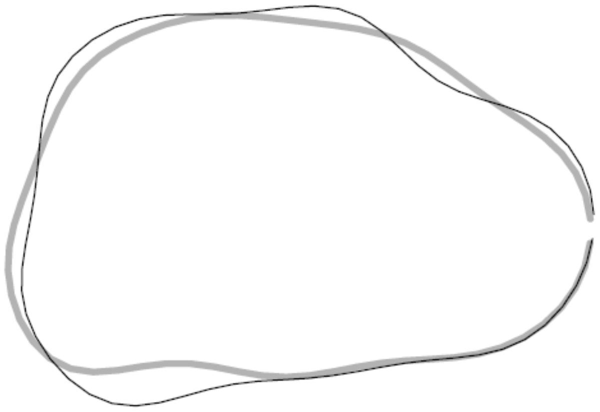
fgsProcrustes: Full Generalized Procrustes alignment between shapes with sliding landmarks

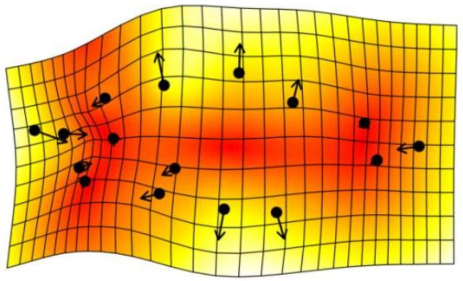
pProcrustes: Partial Procrustes alignment between two shapes

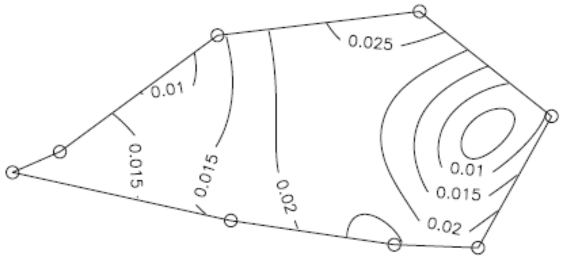
Representing shape change

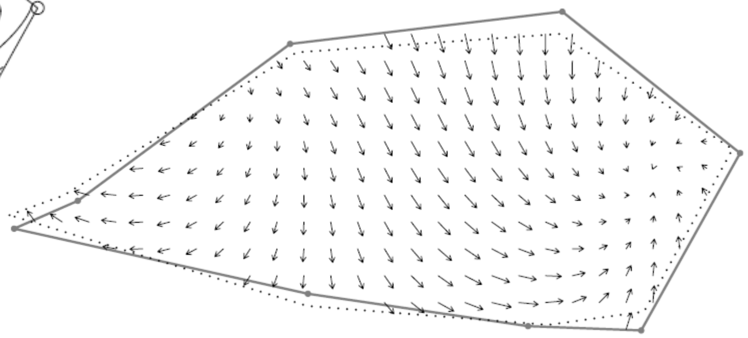
Various methods of visualising shape change:

- Deformation grids
- Principal strains
- Lollipop sticks
- Vectors
- Contours





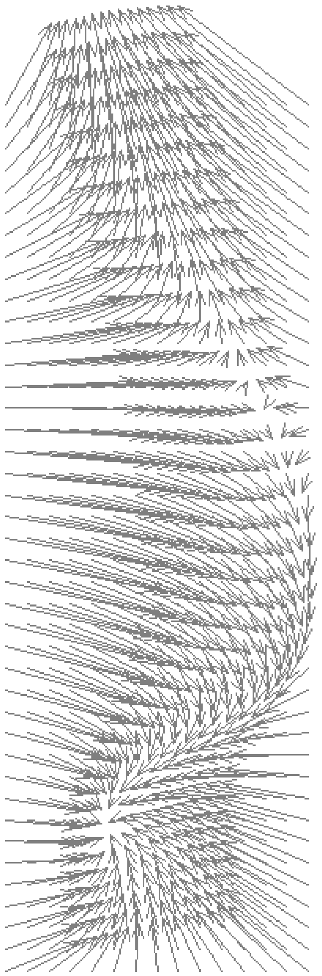




Exercise 5: Visualising Shape Change

- Explore the `tps_grid()`, `tps_iso()`, `tps_raw()` and `tps_arr()` functions in Momocs to compare between two examples (your choice!)

Note: these visualisations work better with outline data (tomorrow!)



Format for the two-day workshop…

Day One (Morning):
- ~~Introductory lecture (What is Geometric Morphometrics? What are Landmarks? How do we do it?)~~
- ~~Some basics in R (environment and generating data)~~
- ~~Organising data for morphometrics~~
- ~~Geometric Transformations~~

Day One (Afternoon):
- ~~Procrustes Superimposition~~
- ~~Other imposition methods (e.g. Resistant Fit)~~
- ~~Representing shape differences (warp grids, arrows, etc.)~~
- Exploring shape variation and testing hypotheses (PCA, multivariate regression, MANOVA, LDA and cladistics)

Day Two (Morning):
- Acquiring outline data in R
- Fourier Analysis (principles)
- Elliptic Fourier Analysis
- Variants of Fourier Analysis
- Outline data exploration

Day Two (Afternoon):
- Advanced analytical frameworks (e.g. Maximum Likelihoods and cluster techniques)
- Advanced outline methods (symmetry and rectilinearity extraction)
- Concluding remarks

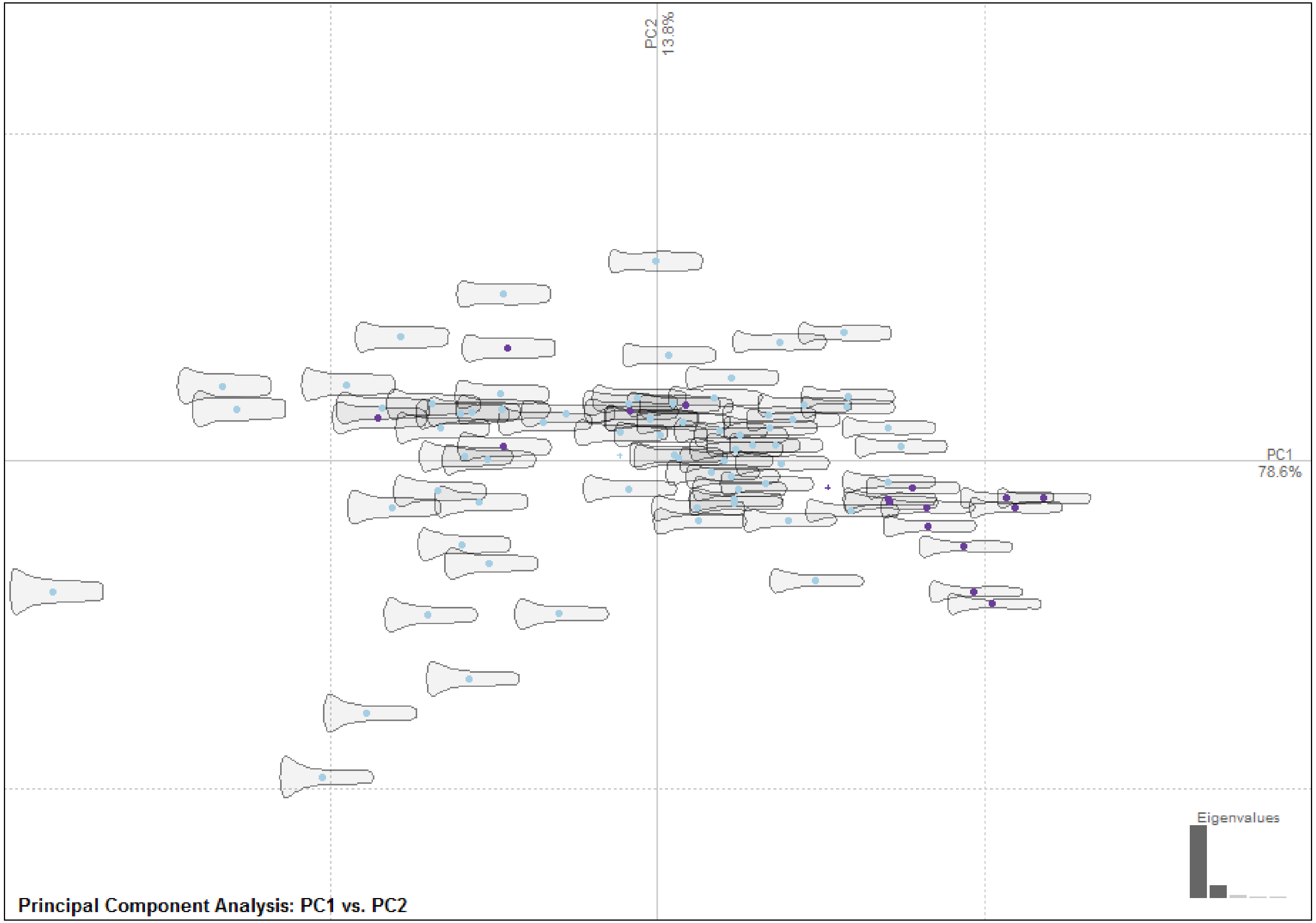
Examining shape variance...

- **Principal Component Analysis** (PCA)
- **Discriminant Analysis** (LDA/CVA)
- **MANOVA** / **Procrustes ANOVA**
- **Mean shape examination**
- **Regression** (Lineal/Multiple)
- **Cluster Analysis** (Hierarchical etc.)
- **Maximum Likelihood** (ML) tree building exercises
- **Kohonen Self-organising Maps (SoMs) / Artificial Neural Networks**
- **Multiple Factor Analysis** (MFA)

#1: Principal Component Analysis (PCA)

`PCA() / plotTangentSpace()`

- Often the first method of analysis following normalisation
- Principal Component space = shape space
- Every point represents a landmark configuration
- Consensus shape (mean shape) = origin (0,0)
- Weighting of the components = % shape variance
- Principal axes represent sources of theoretical shape change
- Similar practice to PCA of other quantitative information
- The scores produced are often used for statistical analysis



Hoggard, C.S., Lauridsen, L. and Witte, K.B. (2019). The Potential of Geometric Morphometrics for Danish Archaeology: Two Case Studies. Arkæologisk Forum, 40: 30-42. (http://www.archaeology.dk/16738/Nr.%2040%20-%202019). OSF: https://osf.io/en5d2/

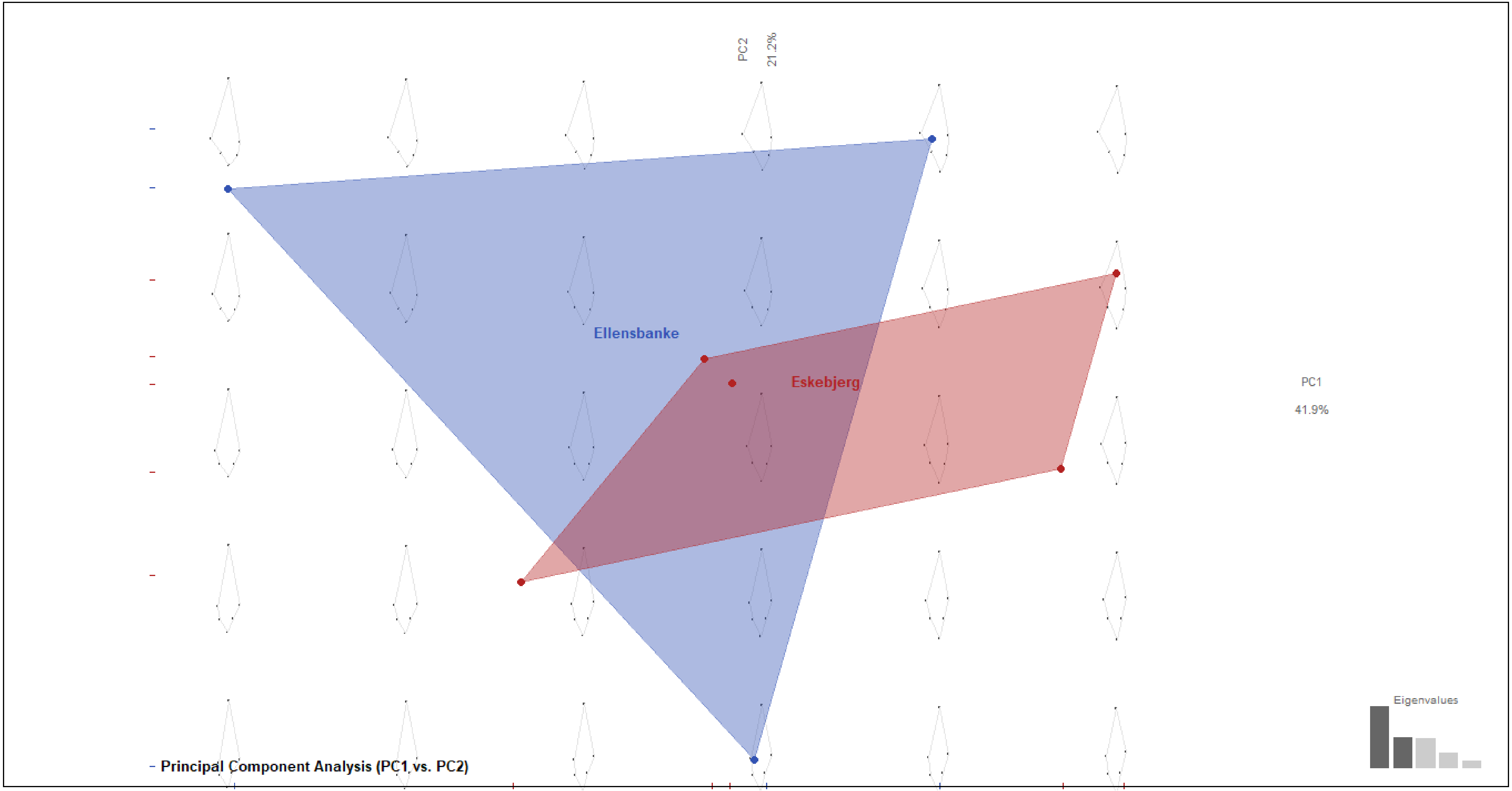
PCA: Customisation in Momocs

```
## S3 method for class 'PCA'
plot(x, fac, xax = 1, yax = 2, points = TRUE,
  col = "#000000", pch = 20, cex = 0.5, palette = col_solarized,
  center.origin = FALSE, zoom = 1, xlim = NULL, ylim = NULL,
  bg = par("bg"), grid = TRUE, nb.grids = 3, morphospace = TRUE,
  pos.shp = c("range", "full", "circle", "xy", "range_axes", "full_axes")[1],
  amp.shp = 1, size.shp = 1, nb.shp = 12, nr.shp = 6, nc.shp = 5,
  rotate.shp = 0, flipx.shp = FALSE, flipy.shp = FALSE, pts.shp = 60,
  border.shp = col_alpha("#000000", 0.5), lwd.shp = 1,
  col.shp = col_alpha("#000000", 0.95), stars = FALSE, ellipses = FALSE,
  conf.ellipses = 0.5, ellipsesax = FALSE, conf.ellipsesax = c(0.5, 0.9),
  lty.ellipsesax = 1, lwd.ellipsesax = sqrt(2), chull = FALSE,
  chull.lty = 1, chull.filled = TRUE, chull.filled.alpha = 0.92,
  density = FALSE, lev.density = 20, contour = FALSE, lev.contour = 3,
  n.kde2d = 100, delaunay = FALSE, loadings = FALSE,
  labelspoints = FALSE, col.labelspoints = par("fg"),
  cex.labelspoints = 0.6, abbreviate.labelspoints = TRUE,
  labelsgroups = TRUE, cex.labelsgroups = 0.8, rect.labelsgroups = FALSE,
  abbreviate.labelsgroups = FALSE, color.legend = FALSE, axisnames = TRUE,
  axisvar = TRUE, unit = FALSE, eigen = TRUE, rug = TRUE,
  title = substitute(x), box = TRUE, old.par = TRUE, ...)
```

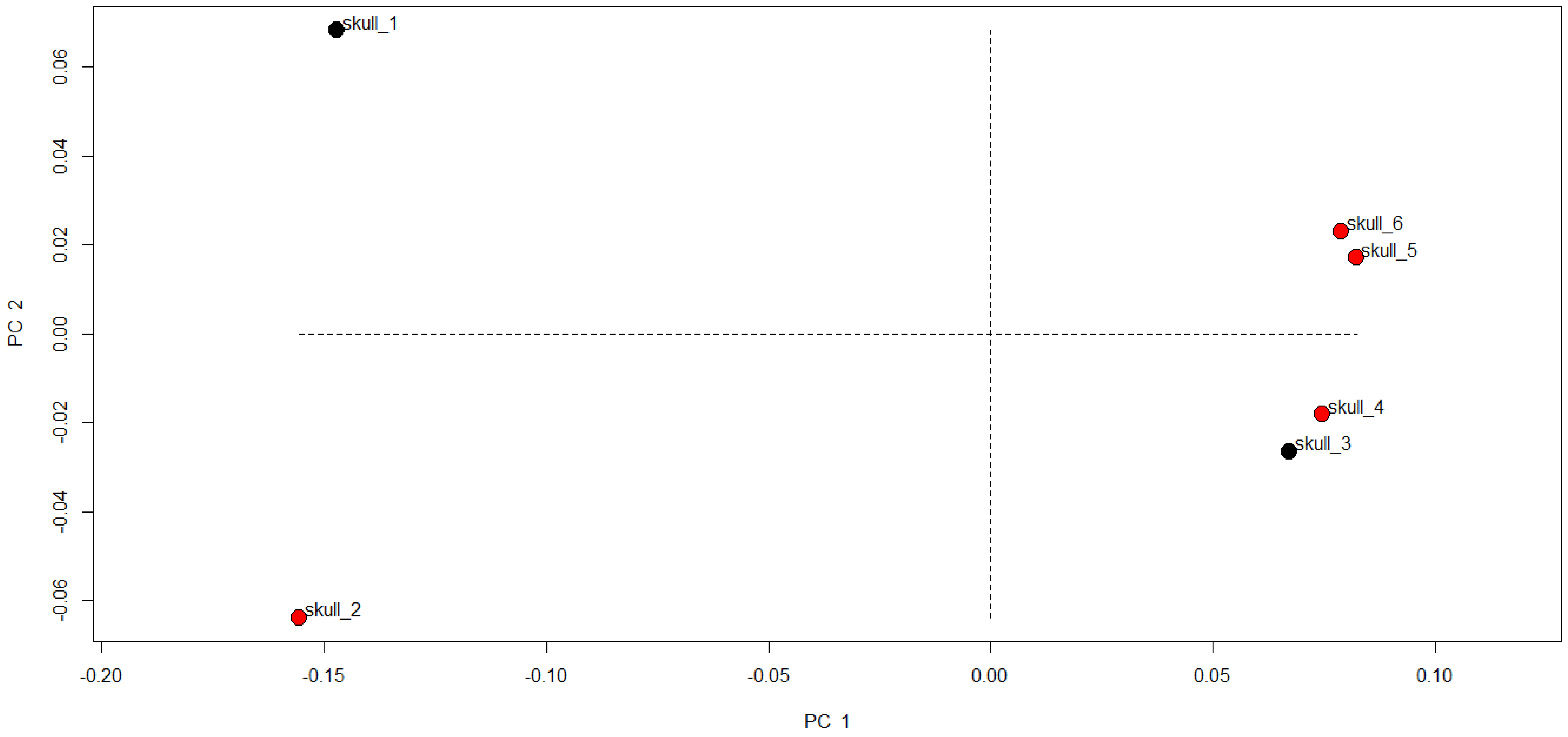
Data from Momocs and Geomorph can also be converted to a data-frame and visualised in the tidyverse package!

Exercise 6: Principal Component Analysis

- Explore the main sources of theoretical shape variation and the distribution of points through the plotting of a principal component space for both datasets.
- Momocs: `PCA()` function
- Geomorph: `plotTangentSpace()` function
- Produce a scree table and scree plot for the tanged points (`?scree / ?scree_plot`)
- Customise where possible (`?plot.PCA()`) and include a variable



PC2
21.2%
Ellensbanke
Eskebjerg
PC1
41.9%
Eigenvalues
Principal Component Analysis (PC1 vs. PC2)



skull_1
skull_2
skull_3
skull_4
skull_5
skull_6
PC 1
PC 2
-0.20
-0.15
-0.10
-0.05
0.00
0.05
0.10
-0.06
-0.04
-0.02
0.00
0.02
0.04
0.06

#2: Discriminant Function Analysis (LDA/CVA)

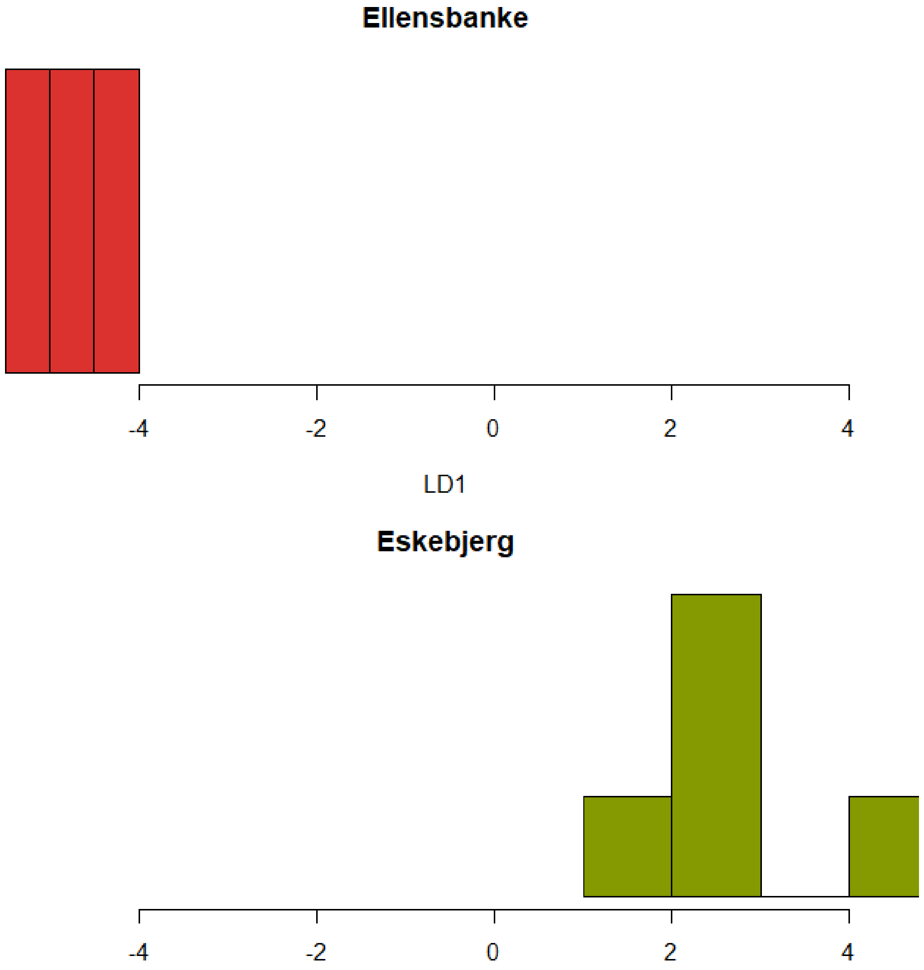
LDA()

- Use of discriminant analysis to determine whether groups defined *a priori* can be distinguished based on their maximum group separation (as through their PCA scores or landmark configurations)

- Similarly to normal discriminant analysis methods a training dataset can be created and two classification scores can be obtained:
  - An initial percentage outlining the success of the separation as based on group data
  - A jack-knifed value based on leave-one-out cross validation (useful for determining whether an artefact can be categorised solely on its shape

Exercise 7: Discriminant Analysis

- Use the `LDA()` function in Momocs to create a discriminant analysis for the site variable
- Plot the resulting graphic and examine the LDA file to extract scores
- Customise the Lineal Discriminant plot



#3 MANOVA / Procrustes ANOVA

`MANOVA() / procD.lm()`

- PCA and DA act as exploratory devices for looking at shape difference
- MANOVA and Procrustes ANOVA provide a statistical framework for examining shape
- Procrustes ANOVA: performed in Geomorph (with landmark data)
- MANOVA: performed in Momocs (with landmark and outline data)
- Null hypothesis: same populations / shape

Exercise 8: MANOVA / Procrustes ANOVA

- Using the `MANOVA()` and `procD.lm()` functions test for difference between:
  - Tanged points vs. Site
  - Skull vs. sex
  - Skull vs. location

- For geomorph the data will need to be in a two-dimensional array format!

```
## S3 method for class 'PCA'
MANOVA(x, fac, test = "Hotelling", retain = 0.99, drop)

procD.lm(f1, iter = 999, seed = NULL, RRPP = TRUE, SS.type = c("I",
  "II", "III"), effect.type = c("SS", "MS", "Rsq", "F", "cohen"),
  int.first = FALSE, Cov = NULL, data = NULL,
  print.progress = TRUE, ...)
```

#4 Mean shapes

`mshapes() / mshape()`

- Mean shapes are useful heuristics for conveying shape information for particular circumstances or groups
- Median shapes are also possible to extract using the `FUN()` function. May be more meaningful if there are a number of outliers!)

Exercise 9: produce the mean and median shapes for the tanged points (according to site)

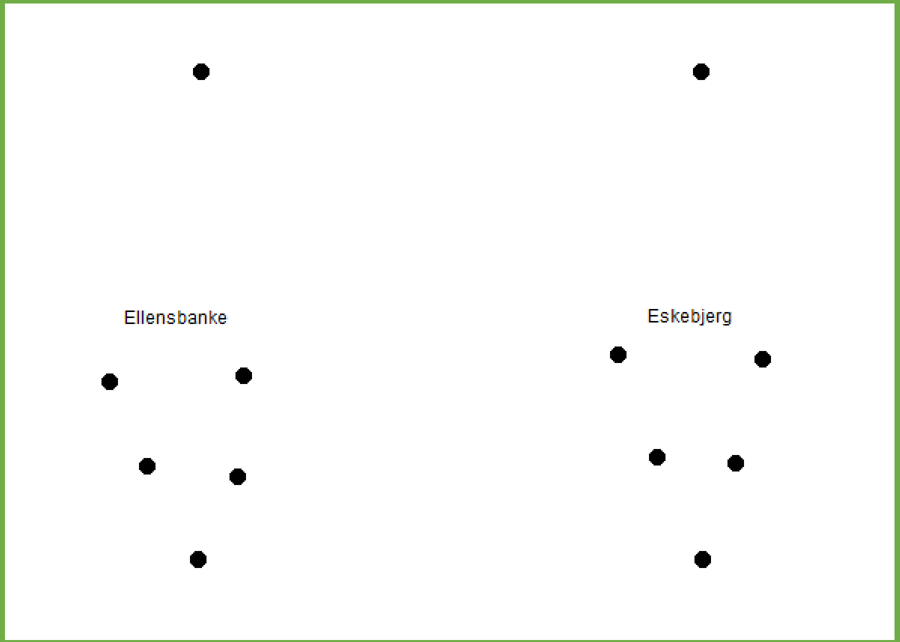
```
mshapes(x, ...)

## S3 method for class 'list'
mshapes(x, FUN = mean, ...)

## S3 method for class 'array'
mshapes(x, FUN = mean, ...)

## S3 method for class 'Ldk'
mshapes(x, FUN = mean, ...)

## S3 method for class 'OutCoe'
mshapes(x, fac, FUN = mean, nb.pts = 120, ...)

## S3 method for class 'OpnCoe'
mshapes(x, fac, FUN = mean, nb.pts = 120, ...)

## S3 method for class 'LdkCoe'
mshapes(x, fac, FUN = mean, ...)

## S3 method for class 'PCA'
mshapes(x, fac, ...)

MSHAPES(x, ...)
```

#4 Mean shapes

`mshapes() / mshape()`

- Mean shapes are useful heuristics for conveying shape information for particular circumstances or groups
- Median shapes are also possible to extract using the `FUN()` function. May be more meaningful if there are a number of outliers!)

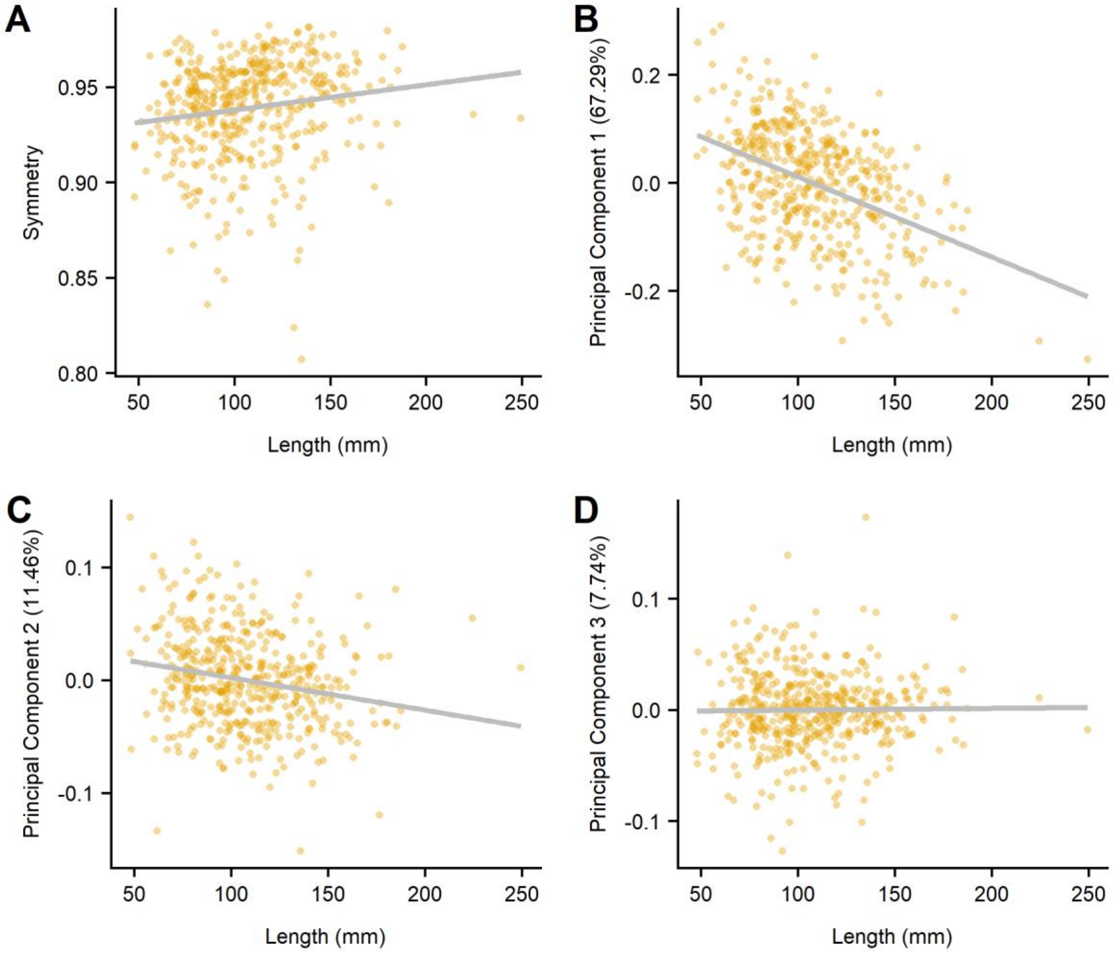
Exercise 9: produce the mean and median shapes for the tanged points (according to site)



#5 Regression-based methodologies

`lm()`

- Useful for examining shape against hypotheses involving quantitative data e.g. size (as length or centroid size) or symmetry (as extracted from outlines)
- The Principal Component scores can be extracted from a PCA file and filtered through to a lineal or multiple regression using the `lm()` function.
- Example: tomorrow with outline data (symmetry)

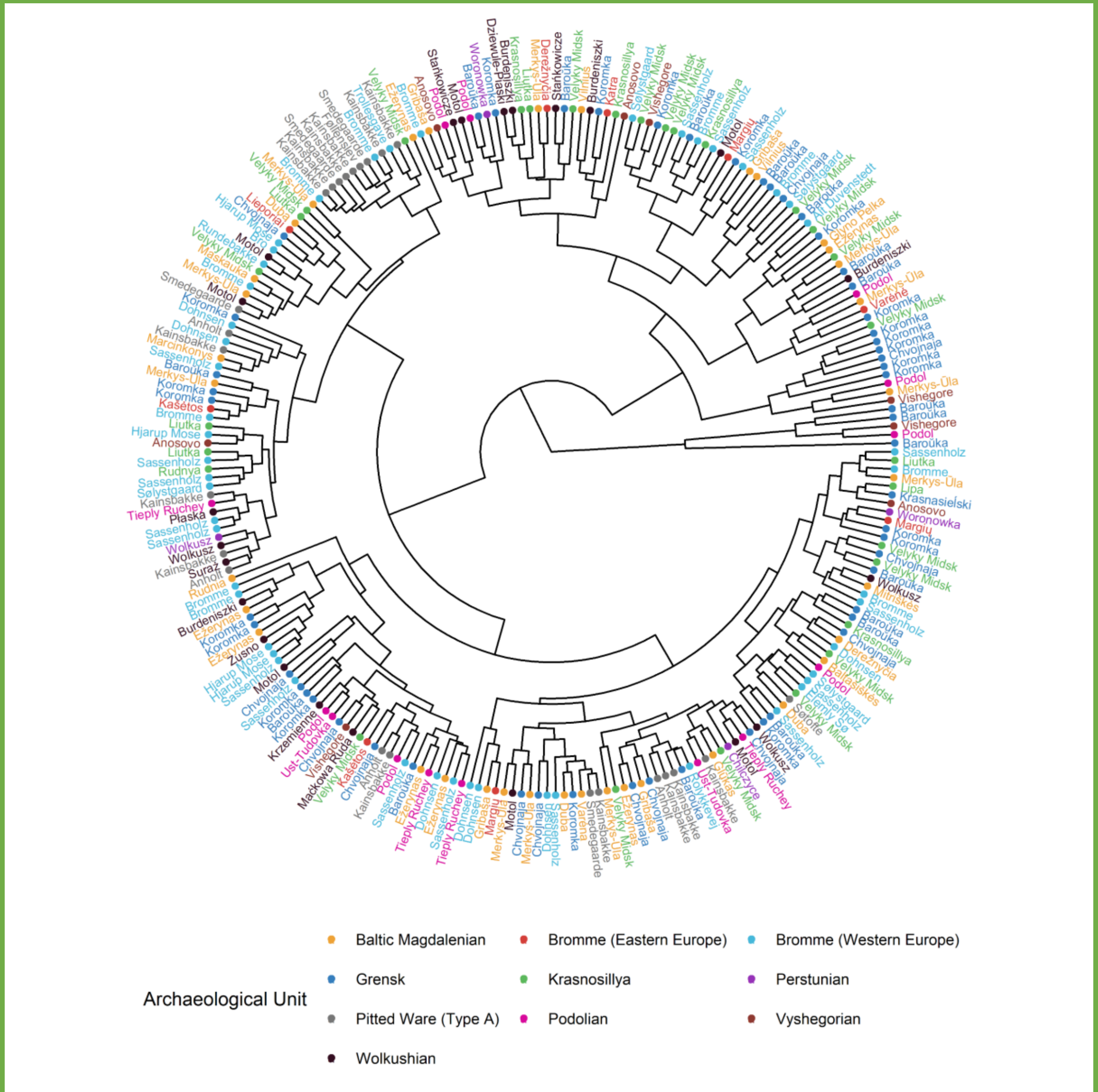


Hoggard, C.S., McNabb, J. and Cole, J.N. (2019). The application of elliptic Fourier analysis in understanding biface shape and symmetry through the British Acheulean. Journal of Paleolithic Archaeology, 2 (2): 115-133. (https://doi.org/10.1007/s41982-019-00024-6). OSF: https://osf.io/td92j/.

#6 Cluster analysis

CLUST()

- Cluster analysis is a useful method for examining a posteriori classifications based on shape data (PCA scores)
- Hierarchical clustering methodologies are implemented through CLUST()
- These .phylo class trees can then be further customised through the ggtree package (a must for anyone interesting in tree-based analyses!)
- **Maximum Likelihood**: an extension of cluster analysis (computer intensive!)



Ivanovaite, L., Swertka, K., **Hoggard, C.S.**, Sauer, F. and Riede, F. (accepted). All these fantastic cultures? Research history and regionalisation in the Late Palaeolithic tanged point cultures of Eastern Europe. European Journal of Archaeology.
OSF: https://osf.io/agrwb/.

#6 Cluster analysis

CLUST()

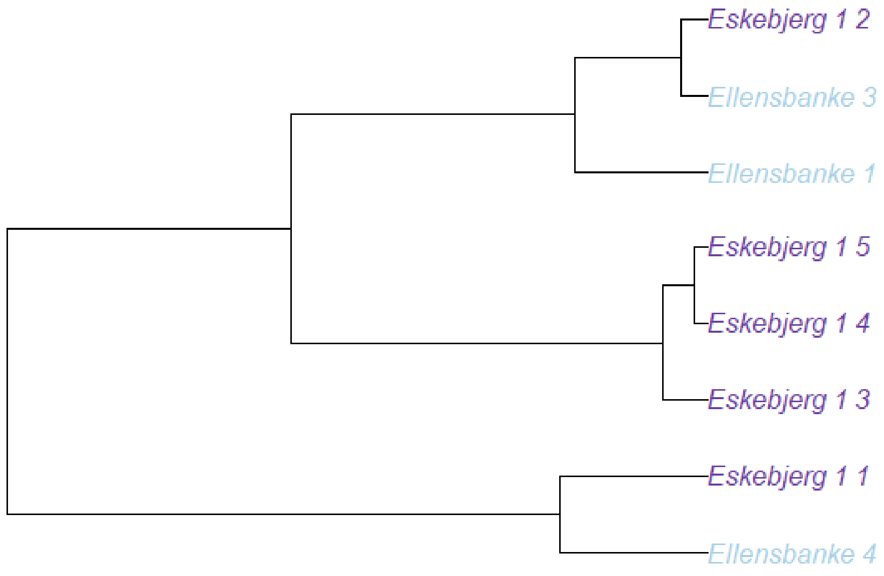
`x`	a PCA object (Coe method deprecated so far)
`fac`	the id or column name or formula for columns to use from $fac.
`type`	to pass to `ape::plot.phylo`'s `type` argument, one of "cladogram", "phylogram", "radial", "unrooted" or "fan" (by default)
`dist_method`	to feed dist's `method` argument, one of "euclidean", "maximum", "manhattan", "canberra", "binary" or "minkowski".
`hclust_method`	to feed hclust's method argument, one of "ward.D", "ward.D2", "single", "complete" (default), "average", "mcquitty", "median" or "centroid".
`retain`	number of axis to retain from the PCA as a range of number eg `1:5` to retain the first 5 PCs. If a number <= 1 is passed, then the number of PCs retained will be enough to capture this proportion of variance.
`tip_labels`	the id or column name in $fac to use as tip_labels rather than rownames. Note that you can also pass a character (or a factor) with the same number of rows of x$x
`palette`	a color palette to use (`col_qual` by default). If `NULL`, `par("fg")` is used
`...`	additional parameters to feed `plot.phylo`

Exercise 10: Cluster analysis

- Using the `CLUST()` function produce a phylo class object with the following conditions:
  - Distance method: “Euclidean”
  - Hclust method: “complete”
  - Site as a factor

- Visualise different tree types (cladogram, phylogram and fan!) and play the visual tools

- Optional:
  - Install ggtree and explore the phylo class object
  - Experiment with the `KMEANS()` function



Format for the two-day workshop…

Day One (Morning):

- ~~Introductory lecture (What is Geometric Morphometrics? What are Landmarks? How do we do it?)~~
- ~~Some basics in R (environment and generating data)~~
- ~~Organising data for morphometrics~~
- ~~Geometric Transformations~~

Day One (Afternoon):

- ~~Procrustes Superimposition~~
- ~~Other imposition methods (e.g. Resistant Fit)~~
- ~~Representing shape differences (warp grids, arrows, etc.)~~
- ~~Exploring shape variation and testing hypotheses (PCA, multivariate regression, MANOVA, LDA and cladistics)~~

Day Two (Morning):

- Acquiring outline data in R
- Fourier Analysis (principles)
- Elliptic Fourier Analysis
- Variants of Fourier Analysis
- Outline data exploration

Day Two (Afternoon):

- Advanced analytical frameworks (e.g. Maximum Likelihoods and cluster techniques)
- Advanced outline methods (symmetry and rectilinearity extraction)
- Concluding remarks

Day Two…

Format for the two-day workshop…

Day One (Morning):
~~• Introductory lecture (What is Geometric Morphometrics? What are Landmarks? How do we do it?)~~
~~• Some basics in R (environment and generating data)~~
~~• Organising data for morphometrics~~
~~• Geometric Transformations~~

Day One (Afternoon):
~~• Procrustes Superimposition~~
~~• Other imposition methods (e.g. Resistant Fit)~~
~~• Representing shape differences (warp grids, arrows, etc.)~~
~~• Exploring shape variation and testing hypotheses (PCA, multivariate regression, MANOVA, LDA and cladistics)~~

Day Two (Morning):
• Acquiring outline data in R
• Fourier Analysis (principles)
• Elliptic Fourier Analysis
• Variants of Fourier Analysis
• Outline data exploration

Day Two (Afternoon):
• Advanced analytical frameworks (e.g. Maximum Likelihoods and cluster techniques)
• Advanced outline methods (symmetry and rectilinearity extraction)
• Concluding remarks

Outline analysis

- **Outline coordinates:** XY coordinates along an open or closed curve
- Useful for structures which are comparable in a geometric sense but where individual homologous landmarks are difficult to pinpoint e.g. tooth perimeters
- Three main types of outline or curve analysis:
  - **Semilandmark analysis**
  - **Eigenshape analysis** (not as widely used)
  - **Fourier-based analyses (EFA/RFA/DCT)**

About the data used in outline analysis

- For **semilandmark** analysis **we need the same** amount of points (the process is the same as with landmark analysis)
- For **Fourier-based** analysis **we do not need the same** amount of points

- For semilandmark analysis **we need to start** at exactly the same position
- For Fourier-based analysis **we do not need to start** at exactly the same position

Landmark count during Fourier Analysis?

Outline data and R

Two methods are suggested here:

1. tpsDig2 and `import_tps()` (external and through Momocs)

2. For outlines: `import_jpg()` and `import_Conte()` (Momocs)

Extract outlines coordinates from an image silhouette

Description

Provided with an image 'mask' (i.e. black pixels on a white background), and a point form where to start the algorithm, returns the (x; y) coordinates of its outline.

Usage

```
import_Conte(img, x)
```

Arguments

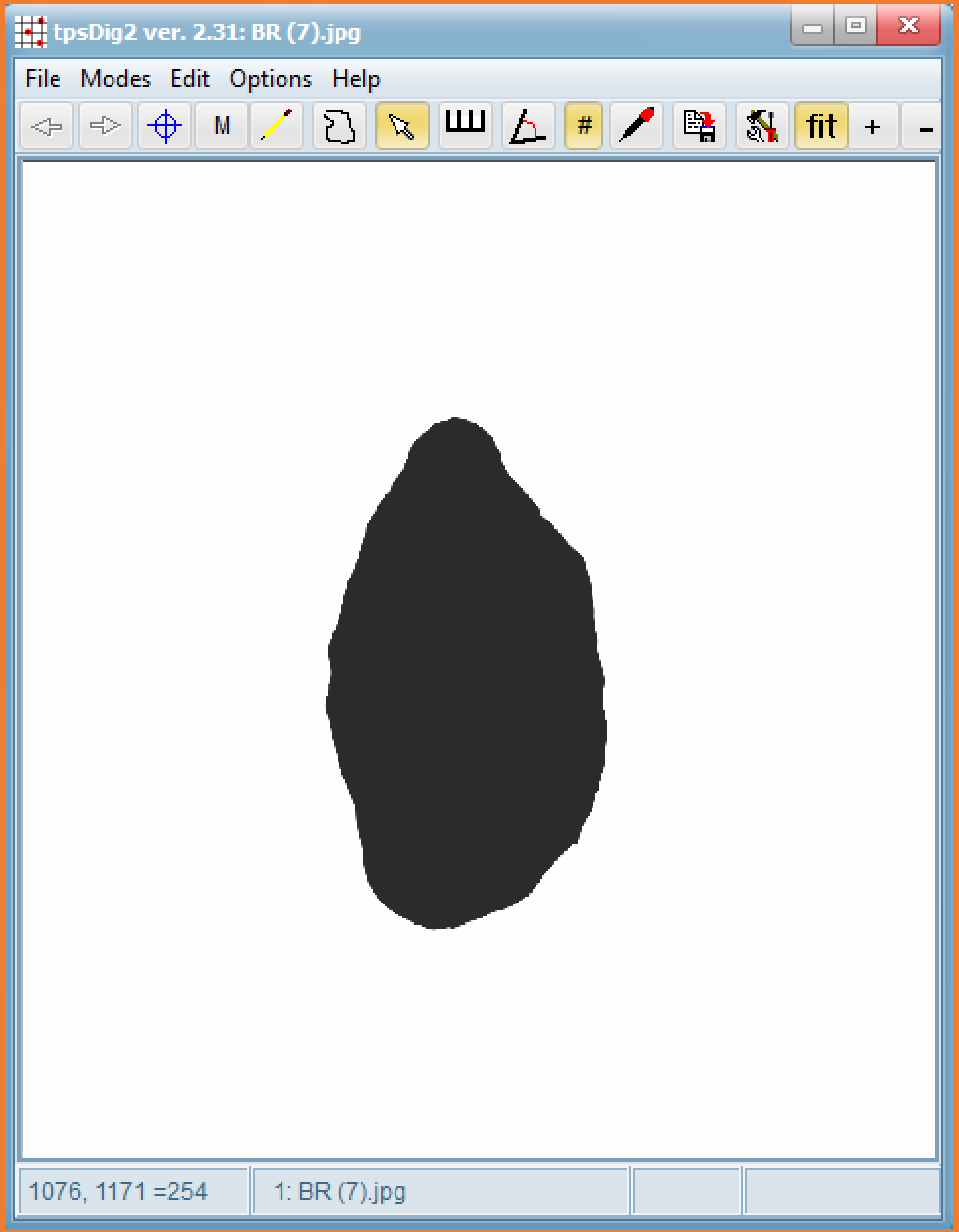
`img`	a matrix of a binary image mask.
`x`	numeric the (x; y) coordinates of a starting point within the shape.

Outline data and R

Two methods are suggested here:

1. tpsDig2 and `import_tps()` (external and through Momocs)

2. For outlines: `import_jpg()` and `import_Conte()` (Momocs)



Journal of Paleolithic Archaeology
https://doi.org/10.1007/s41982-019-00024-6

Check for updates

The Application of Elliptic Fourier Analysis in Understanding Biface Shape and Symmetry Through the British Acheulean

Christian Steven Hoggard[1,2] · John McNabb[2] · James Nathan Cole[3]

Published online: 21 February 2019

Abstract
Acheulean biface shape and symmetry have fuelled many discussions on past hominin behaviour in regards to the 'meaning' of biface technology. However, few studies have attempted to quantify and investigate their diachronic relationship using a substantial dataset of Acheulean bifaces. Using the British archaeological record as a case study, we first perform elliptic Fourier analysis on biface outlines to quantify and better understand the relationship between biface shape and individual interglacial periods. Using the extracted Fourier coefficients, we then detail the nature of symmetry throughout this period, before investigating both shape and symmetry in parallel. The importance of size (through biface length) as a factor in biface shape and symmetry is also considered. Results highlight high levels of symmetry from Marine Isotope Stage (MIS) 13, followed by increasing asymmetry through the British Acheulean. Other observations include a general shift to 'pointed' forms during MIS 9 and 7 and the importance of size in high biface symmetry levels. This article concludes by discussing the potential importance of secondary deposition and palimpsest sites in skewering the observed relationships throughout the Palaeolithic.

OSF: https://osf.io/td92j/

Table 1 The British Lower and Middle Palaeolithic assemblages used for this study (total: 468)

Marine Isotope Stage	*Context*	*Sample size*	*Source*	*Collection*
13 (528–474 kya)	Warren Hill	49	Marshall et al. (2002)	–
	Boxgrove	50	Marshall et al. (2002)	–
11 (427–364 kya)	Bowman's Lodge	28	Marshall et al. (2002)	–
	Elveden	29	Author (JNC)	BM (Sturge Collection)
	Swanscombe	58	Author (JM)	BM (Marston Collection)
9 (334–301 kya)	Broom	50	Marshall et al. (2002)	–
	Furze Platt	69	Author (JM)	BM (Treacher Collection)
7 (242–186 kya)	Cuxton	50	Marshall et al. (2002)	–
	Pontnewydd Cave	37	Author (JNC)	NMW
4/3 (57 kya)	Lynford	48	Author (JNC)	NAU

Exercise 11: Importing outline data

- Inspect the outline file by opening the file in notepad
- Import the handaxe outline file through the `import_tps()` function
- Import the associated metadata through the `read.csv()` function
- Use the `View()` function to open the data table
- Use the `print()` function to view the raw coordinate data
- Convert the new handaxe function to an outline file (with the data) using the `Out()` function
- Explore the file using the functions in Momocs (`?panel`)

Elliptical Fourier Analysis (EFA)

- One of a number of ‘Fourier’ based methods of **curve (de-)composition**
  - Derived from the first series by **Jean Baptiste Joseph Fourier** (1768-1830)

- EFA was developed during the Cold War, partly by Giardina and Kuhl (1977) and Kuhl and Giardina (1982): used as a method of categorising enemy aircraft

- Grounded on a set of **parametric equations to fit a curve (Fourier harmonic amplitudes) from x and y Cartesian landmarks**

- Math: **sine and cosine transformations**

- Product: **Fourier coefficients**

Elliptical Fourier Analysis (EFA)

$$A_n = \frac{T}{2n^2\pi^2} \sum_{p=1}^{k} \frac{\Delta x_p}{\Delta t_p} \left[\cos\left(\frac{2\pi n t_p}{T}\right) - \cos\left(\frac{2\pi n t_{p-1}}{T}\right) \right]$$

$$B_n = \frac{T}{2n^2\pi^2} \sum_{p=1}^{k} \frac{\Delta x_p}{\Delta t_p} \left[\sin\left(\frac{2\pi n t_p}{T}\right) - \sin\left(\frac{2\pi n t_{p-1}}{T}\right) \right]$$

$$C_n = \frac{T}{2n^2\pi^2} \sum_{p=1}^{k} \frac{\Delta y_p}{\Delta t_p} \left[\cos\left(\frac{2\pi n t_p}{T}\right) - \cos\left(\frac{2\pi n t_{p-1}}{T}\right) \right]$$

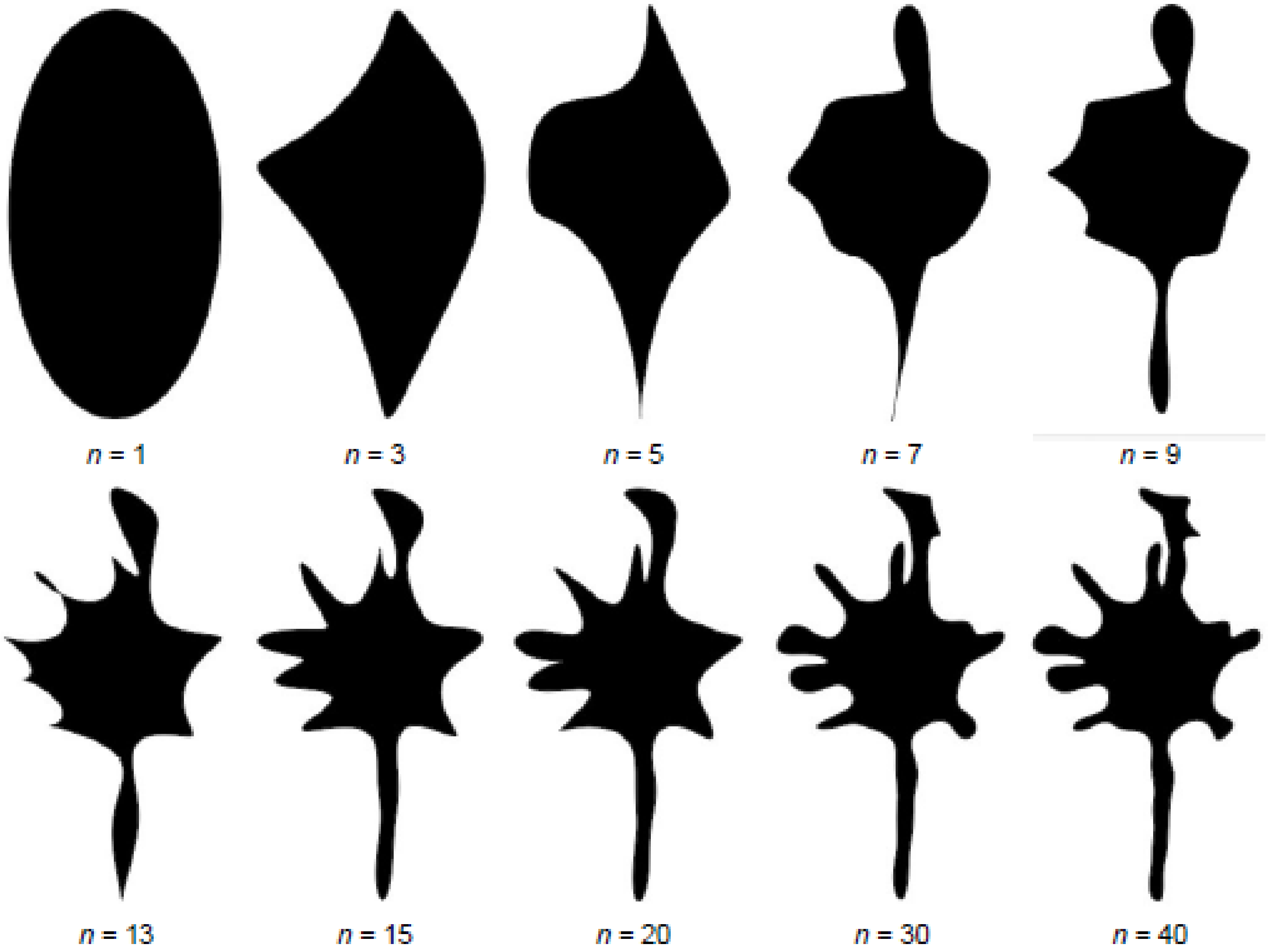
$$D_n = \frac{T}{2n^2\pi^2} \sum_{p=1}^{k} \frac{\Delta y_p}{\Delta t_p} \left[\sin\left(\frac{2\pi n t_p}{T}\right) - \sin\left(\frac{2\pi n t_{p-1}}{T}\right) \right]$$

Where: k = the total number of steps around the outline
n = the harmonic number
Δx = the displacement along the x axis between point p and $p+1$
Δt = the length of the step between point p and $p+1$
t_p = accumulated length of step segments at point p
T = sum of lengths of all steps around outline

When **summed** together, these transformations represent an **approximation of an artefact form**

Elliptical Fourier Analysis (EFA)

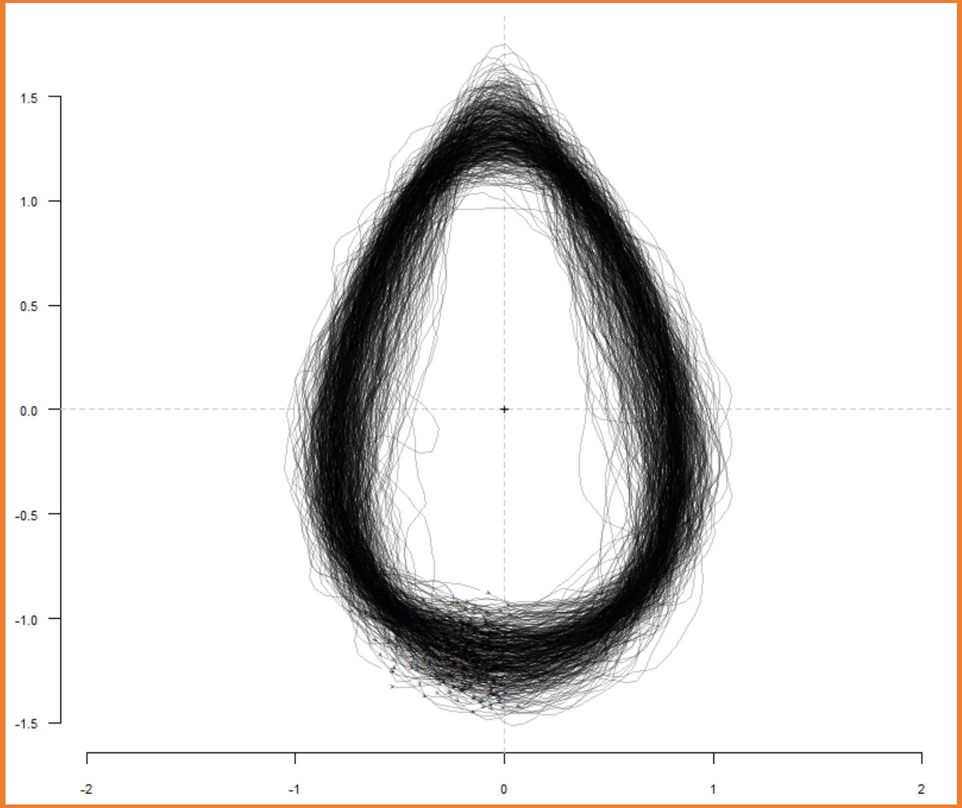
- Outline estimation process that can provide a shape of varying detail depending on the number of harmonics you used.
- **Greater number of harmonics = greater detail**
  (and more closely resembling the shape)
- **Too much detail = too much statistical noise**
- So the right level of harmonics are necessary



Norman MacLeod’s PalaeoMath 101 Series
(http://www.palass.org/publications/newsletter/palaeomath-101)

Exercise 12: Prior normalisation

- Prior normalisation (EFA), all examples should be centred, scaled and closed (if open). These can be done using functions in Momocs.
- Use the `coo_close()`, `coo_centre()` and `coo_scale()` functions and normalise the outlines.
- Check with the `stack()` function.



Calculating harmonic power

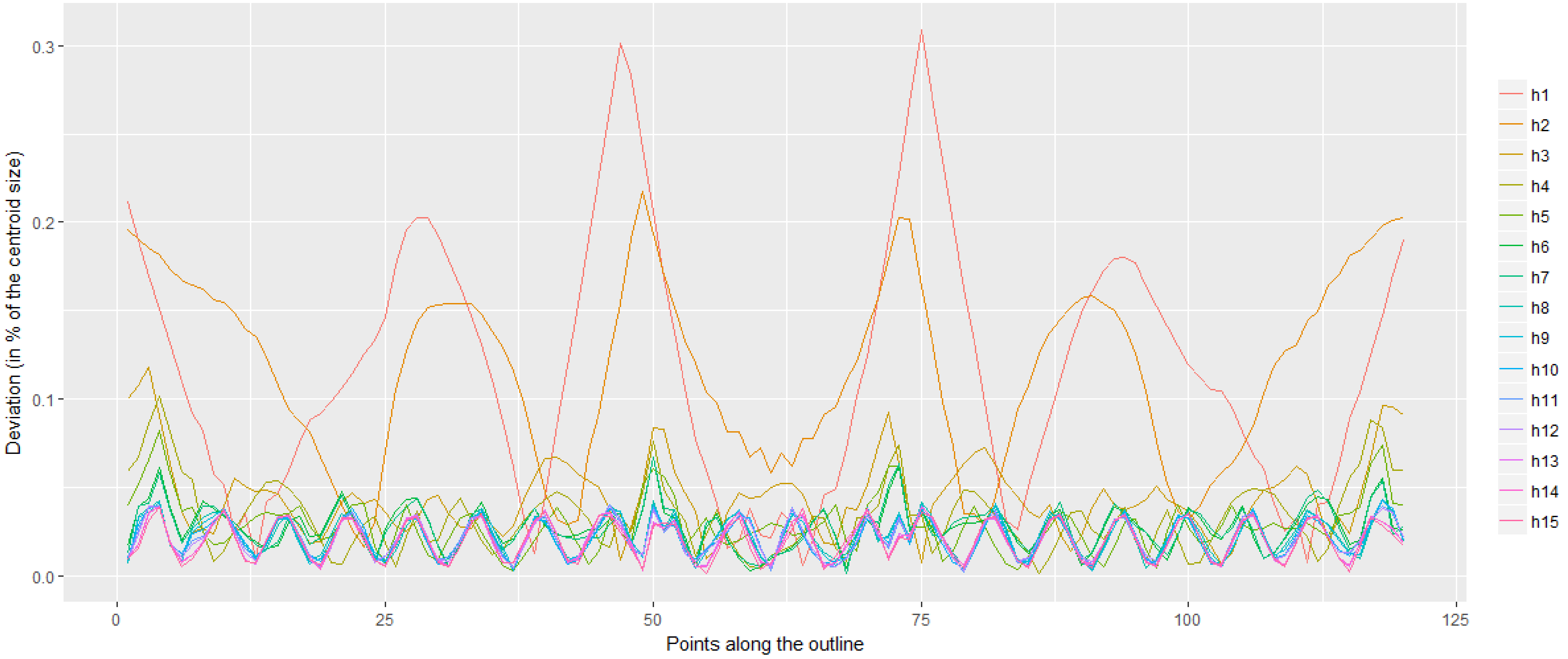
1. ‘Eyeing it’: see at what harmonic the shape looks right
2. Calculating 99% harmonic power
3. Calculating deviation in % of the centroid size
4. Calculating deviation in coefficients

For 2-4: R/Momocs: `calibrate()` functions

Norman MacLeod’s PalaeoMath 101 Series
(http://www.palass.org/publications/newsletter/palaeomath-101)

Calculating harmonic power

`calibrate_deviations_efourier()`



Exercise 13: Harmonics and EFA creation

- Use the following functions to determine the right levels of harmonics necessary for 99.9 harmonic power:
    - `calibrate_harmonicpower_efourier()`
    - `calibrate_reconstructions_efourier()`
    - `calibrate_deviations_efourier()`
- Use the `efourier()` function to create the EFA class file and examine its contents

Alternatives to Elliptic Fourier Analysis

1. Radii variation Fourier transform (RFA)
2. Tangent angle Fourier transform (TFA)
3. Radii variation Fourier transform (equally spaced curvilinear abscissa)
4. Discrete co-sinus transform (DCT)
5. Orthogonal and polynomial fits on open outlines

1-3: Closed outline techniques
4-5: Open outline techniques

See `rfourier()`, `tfourier()`, `sfourier()`, `dfourier()`, `opoly()` and `npoly()`
(and their inverse forms!)

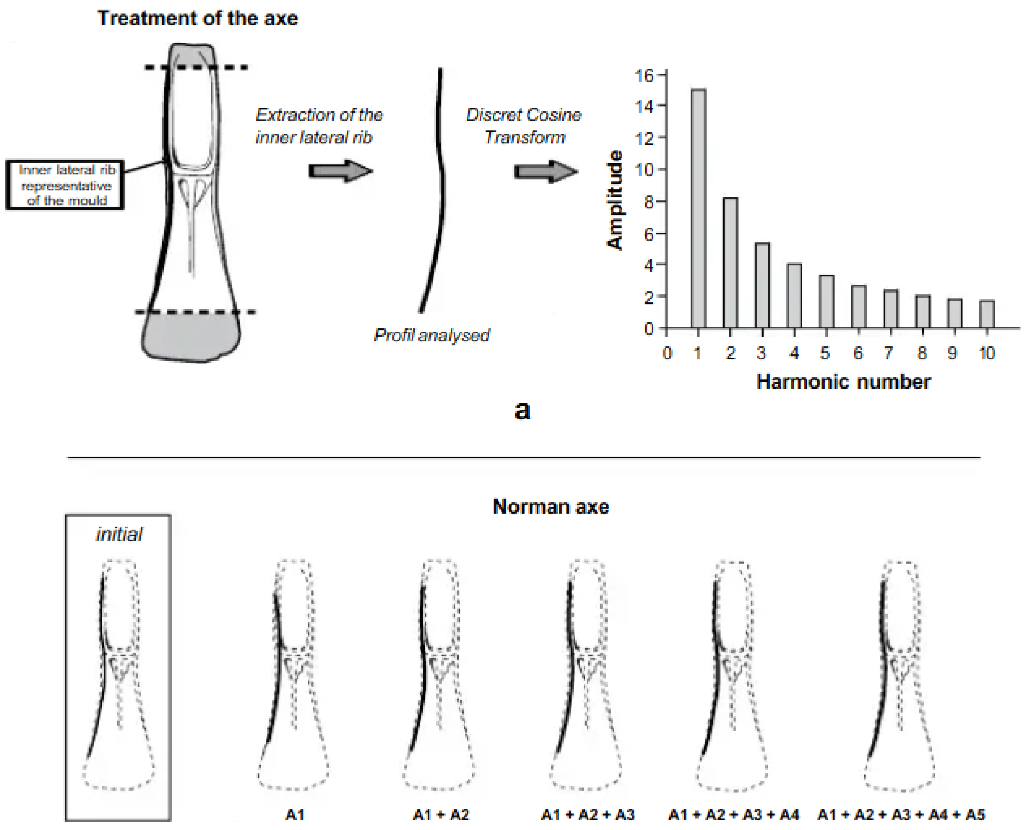
Discrete Cosine Transform (DCT)

An alternative method to EFA using cosine transformations to calculate the shape.

Useful for open outlines.

See Dommergues et al. (2007).

Note: can be performed in Momocs using the `dfourier()` functions.



Dommergues, C. H., Dommergues, J.-L., & Verrecchia, E. P. (2007). The Discrete Cosine Transform, a Fourier-related Method for Morphometric Analysis of Open Contours. Mathematical Geology, 39(8), 749-763.

Forel, B. et al. (2009). Morphometry of Middle Bronze Age palstaves by Discrete Cosine Transform. Journal of Archaeological Science., 36. 721-729.

The analysis of outline data

All the exploratory and statistical methods discussed previously can be applied to this data.

Principal Component Analysis (PCA)
What are the main shape differences? Do they differ across different Isotope Stages?

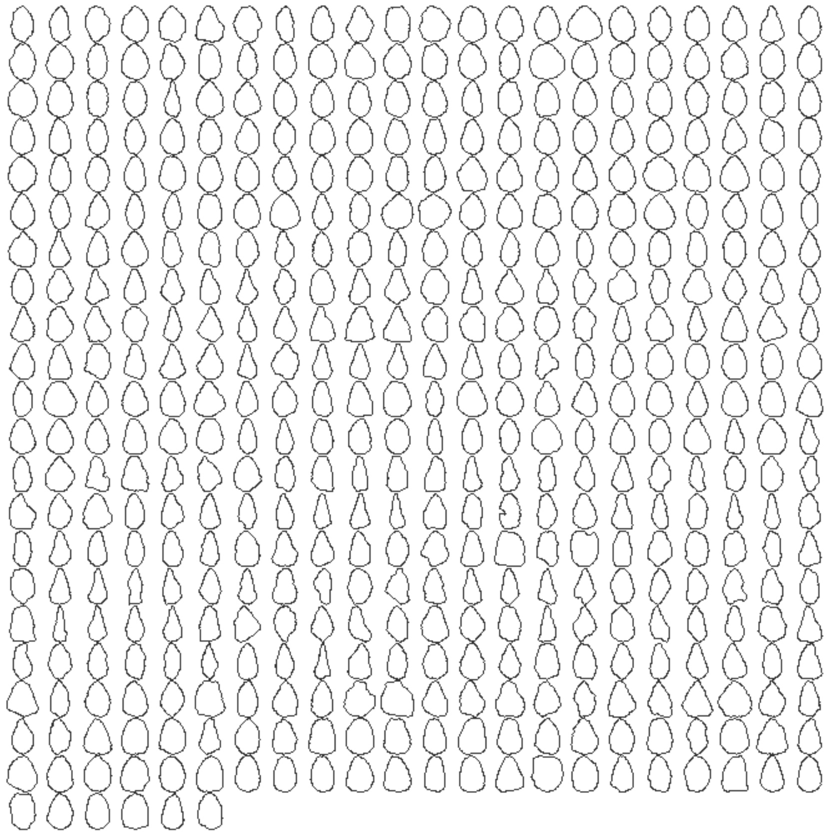
Discriminant Analysis (LDA/CVA)
To what degree do different shapes represent different sites or Isotope Stages?

MANOVA / Procrustes ANOVA
Is there a statistical difference in shape between Isotope Stages?

Mean shape examination
How do different shapes look in different stages on average?

Regression (Lineal/Multiple)
How does size relate to handaxe shape? Significant relationship?

Cluster Analysis (Hierarchical Cluster Analysis etc.)
What different groups can be suggested on the basis of the data we have?



Exercise 14: Data exploration

Using the EFA class object and the embedded data perform all of the following:

Principal Component Analysis (PCA)
What are the main shape differences? Do they differ across different Isotope Stages?

Discriminant Analysis (LDA/CVA)
To what degree do different shapes represent different sites or Isotope Stages?

MANOVA / Procrustes ANOVA
Is there a statistical difference in shape between Isotope Stages?

Mean shape examination
How do different shapes look in different stages on average?

Regression (Lineal/Multiple)
How does size relate to handaxe shape? Significant relationship?

Cluster Analysis (Hierarchical Cluster Analysis etc.)
What different groups can be suggested on the basis of the data we have?

Format for the two-day workshop…

Day One (Morning):
- ~~Introductory lecture (What is Geometric Morphometrics? What are Landmarks? How do we do it?)~~
- ~~Some basics in R (environment and generating data)~~
- ~~Organising data for morphometrics~~
- ~~Geometric Transformations~~

Day One (Afternoon):
- ~~Procrustes Superimposition~~
- ~~Other imposition methods (e.g. Resistant Fit)~~
- ~~Representing shape differences (warp grids, arrows, etc.)~~
- ~~Exploring shape variation and testing hypotheses (PCA, multivariate regression, MANOVA, LDA and cladistics)~~

Day Two (Morning):
- ~~Acquiring outline data in R~~
- ~~Fourier Analysis (principles)~~
- ~~Elliptic Fourier Analysis~~
- ~~Variants of Fourier Analysis~~
- ~~Outline data exploration~~

Day Two (Afternoon):
- Advanced analytical frameworks (e.g. Maximum Likelihoods and cluster techniques)
- Advanced outline methods (symmetry and rectilinearity extraction)
- Concluding remarks

Three (rather more complex) techniques...

1. Maximum Likelihood (ML) phylogenies

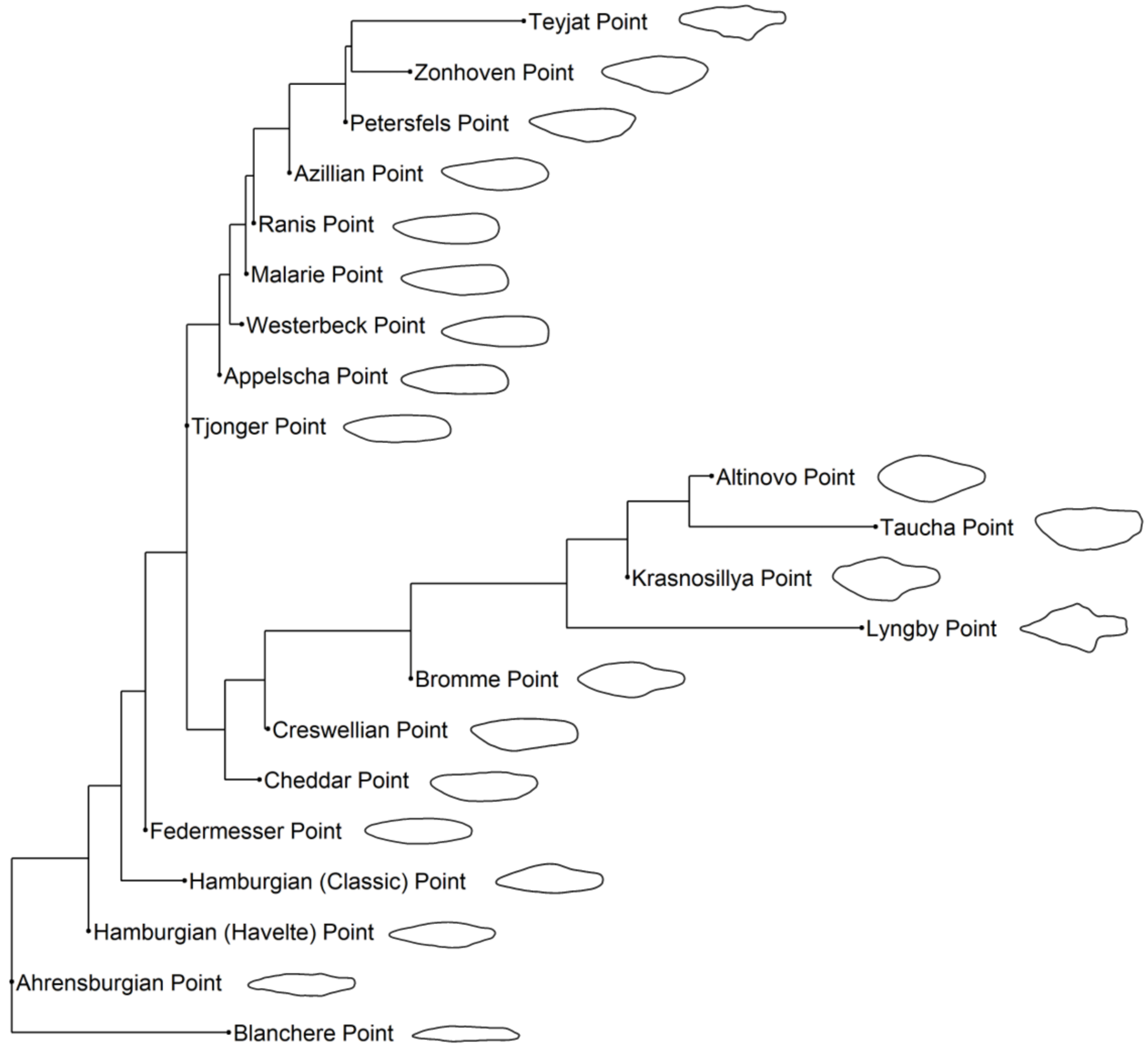
2. Unsupervised self-organising maps
   (Artificial Neural Networks)

3. Multiple Factor Analysis (MFA)

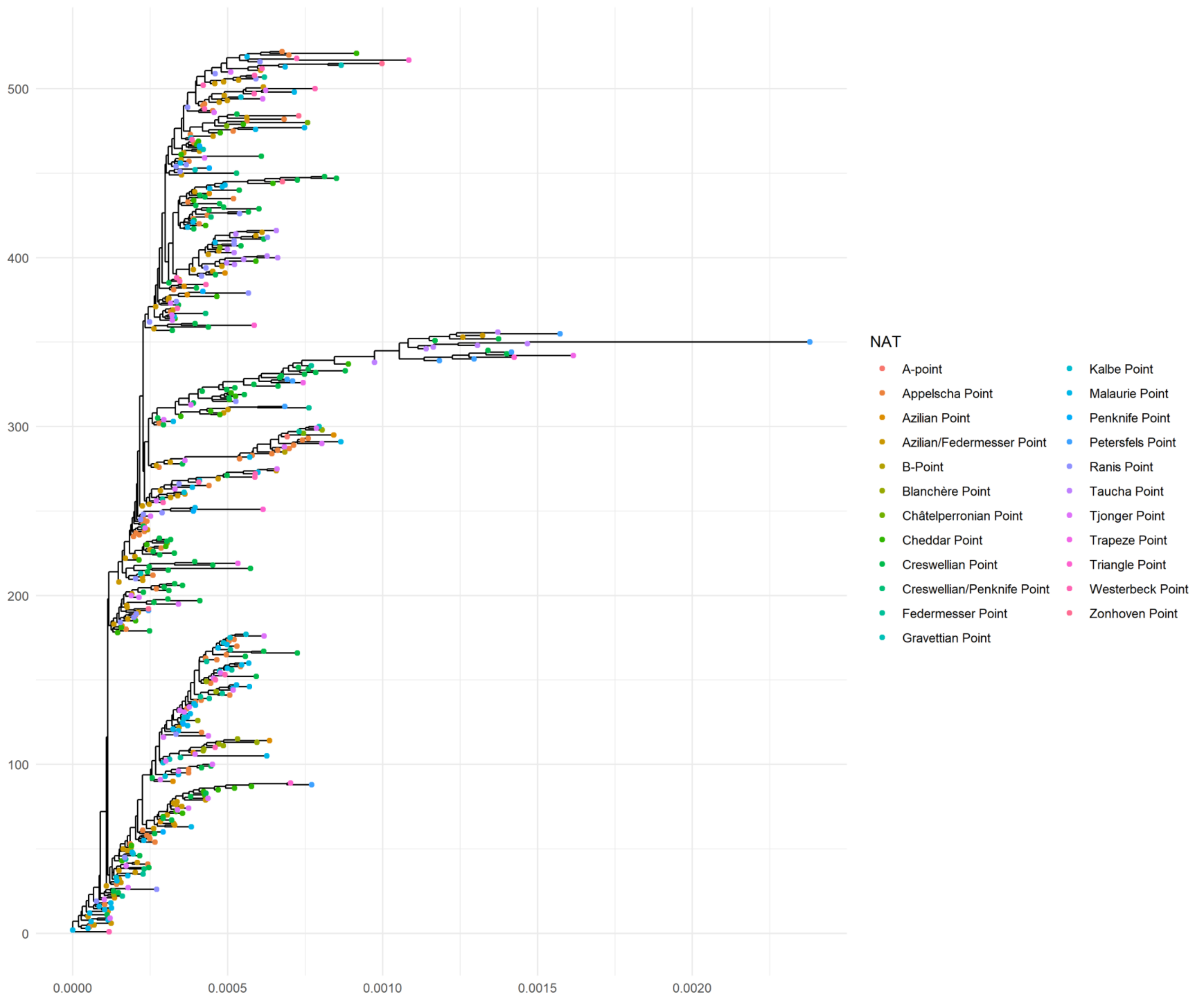
Maximum Likelihood (ML)

- A method of tree building which is more computer-intensive than traditional cluster techniques e.g. Hierarchical Cluster Analysis (HCA)
- Likelihood provides probabilities of the sequences given a model of their evolution on a particular tree
- The more probable the sequences given the tree, the more the tree is preferred.
- All possible trees are considered; computationally intense.
- Produced using the `contml()` function in RPhylip (requires Phylip to be installed!)

CONTML. Estimates phylogenies from gene frequency data by maximum likelihood under a model in which all divergence is due to genetic drift in the absence of new mutations. Does not assume a molecular clock. An alternative method of analyzing this data is to compute Nei's genetic distance and use one of the distance matrix programs. This program can also do maximum likelihood analysis of continuous characters that evolve by a Brownian Motion model, but it assumes that the characters evolve at equal rates and in an uncorrelated fashion, so that it does not take into account the usual correlations of characters.



Teyjat Point
Zonhoven Point
Petersfels Point
Azillian Point
Ranis Point
Malarie Point
Westerbeck Point
Appelscha Point
Tjonger Point
Altinovo Point
Taucha Point
Krasnosillya Point
Lyngby Point
Bromme Point
Creswellian Point
Cheddar Point
Federmesser Point
Hamburgian (Classic) Point
Hamburgian (Havelte) Point
Ahrensburgian Point
Blanchere Point



500
400
300
200
100
0
0.0000
0.0005
0.0010
0.0015
0.0020
NAT
A-point
Appelscha Point
Azilian Point
Azilian/Federmesser Point
B-Point
Blanchère Point
Châtelperronian Point
Cheddar Point
Creswellian Point
Creswellian/Penknife Point
Federmesser Point
Gravettian Point
Kalbe Point
Malaurie Point
Penknife Point
Petersfels Point
Ranis Point
Taucha Point
Tjonger Point
Trapeze Point
Triangle Point
Westerbeck Point
Zonhoven Point

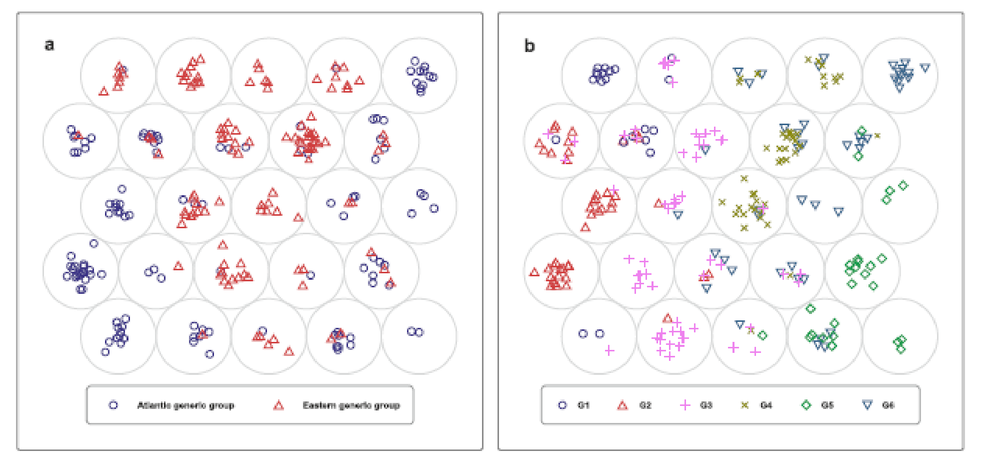
Unsupervised self-organising maps (Artificial Neural Networks)

- Method of unsupervised learning to produce a representation of the input space (shapes) in a low-dimensional map
- Sometimes referred to as a Kohonen Map (Teuvo Kohonen)
- Exploratory advantages over Principal Component Analysis (as a multidimensional scaling methodology)
- Performed in R through **kohonen**

Unsupervised model-based clustering for typological classification of Middle Bronze Age flanged axes

J. Wilczek[a,b,*], F. Monna[a], M. Gabillot[a], N. Navarro[c], L. Rusch[a], C. Chateau[d]

[a] ArTeHiS, UMR CNRS 6298, Université Bourgogne Franche-Comté, 21000 Dijon, France
[b] Ústav archeologie a muzeologie, Masarykova univerzita, 602 00 Brno, Czech Republic
[c] Biogéosciences, UMR CNRS 6282, Ecole Pratique des Hautes Etudes, Université Bourgogne Franche-Comté, 21000 Dijon, France
[d] UFR SVTE, Université Bourgogne Franche-Comté, 21000 Dijon, France



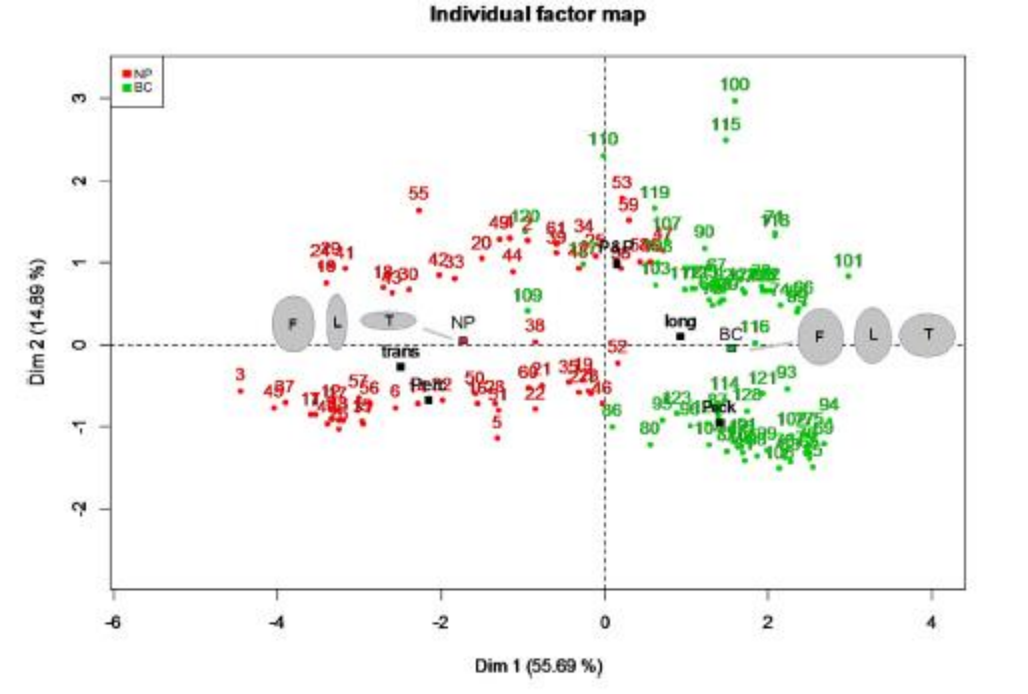
Multiple Factor Analysis (MFA)

- Another extension of the Principal Component Analysis (PCA)
- Advantage of analyzing in parallel more than one group of discrete and continuous variables simultaneously (including multiple shapes / perspectives)
- Conducted using `the FactoMineR` package

Combining morphological and metric variations in the study of design and functionality in stone weights. A comparative approach from continental and insular Patagonia, Argentina

Cardillo Marcelo [a], Scartascini Federico Luis [b], Zangrando Atilio Francisco [c]

[a] CONICET-IMHICIHU-UBA, Buenos Aires, Argentina
[b] CONICET-IMHICIHU, Argentina
[c] CONICET-CADIC-UBA, Argentina



```
MFA (base, group, type = rep("s",length(group)), excl = NULL,
    ind.sup = NULL, ncp = 5, name.group = NULL,
    num.group.sup = NULL, graph = TRUE, weight.col.mfa = NULL,
    row.w = NULL, axes = c(1,2), tab.comp=NULL)
```

`base`	a data frame with *n* rows (individuals) and *p* columns (variables)
`group`	a vector with the number of variables in each group
`type`	the type of variables in each group; four possibilities: "c" or "s" for quantitative variables (the difference is that for "s" variables are scaled to unit variance), "n" for categorical variables and "f" for frequencies (from a contingency tables); by default, all variables are quantitative and scaled to unit variance
`excl`	an argument that may possible to exclude categories of active variables of categorical variable groups. NULL by default, it is a list with indexes of categories that are excluded per group
`ind.sup`	a vector indicating the indexes of the supplementary individuals
`ncp`	number of dimensions kept in the results (by default 5)

Final note: extracting other measures

1. Area: coo_area()
2. Symmetry: `symmetry()`
3. Rectilinearity: `coo_rectilinearity()`
4. Rectangularity: `coo_rectangularity()`
5. Circularity: `coo_circularity() / coo_circularityharalick()`
6. Eccentricity: `coo_eccentricity()`
7. Elongation: `coo_elongation()`
8. Solidity: `coo_solidity()`

```
e.g. outlinefile %>% coo_elongation()
```

Format for the two-day workshop…

Day One (Morning):
~~- Introductory lecture (What is Geometric Morphometrics? What are Landmarks? How do we do it?)~~
~~- Some basics in R (environment and generating data)~~
~~- Organising data for morphometrics~~
~~- Geometric Transformations~~

Day One (Afternoon):
~~- Procrustes Superimposition~~
~~- Other imposition methods (e.g. Resistant Fit)~~
~~- Representing shape differences (warp grids, arrows, etc.)~~
~~- Exploring shape variation and testing hypotheses (PCA, multivariate regression, MANOVA, LDA and cladistics)~~

Day Two (Morning):
~~- Acquiring outline data in R~~
~~- Fourier Analysis (principles)~~
~~- Elliptic Fourier Analysis~~
~~- Variants of Fourier Analysis~~
~~- Outline data exploration~~

Day Two (Afternoon):
~~- Advanced analytical frameworks (e.g. Maximum Likelihoods and cluster techniques)~~
~~- Advanced outline methods (symmetry and rectilinearity extraction)~~
~~- Concluding remarks~~

GMM and Archaeology: The Future

1. **Greater application** of GMM in a variety of new (non biological) archaeologies
2. **Methodological applications**: automisation (recording and landmarking)
3. **Coding developments:** towards a replicable, reproducible and Shiny GMM...
4. **More powerful analytics**: e.g. Bayesian and Machine Learning techniques

Questions?

Thank you for attending!