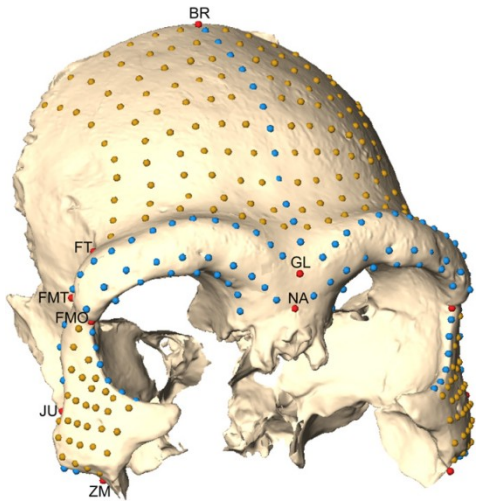
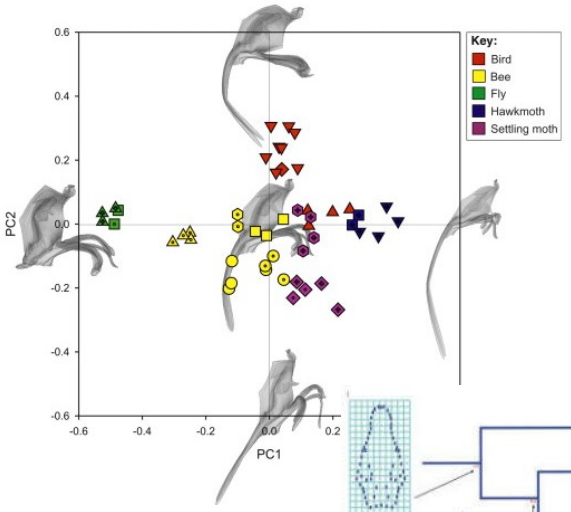
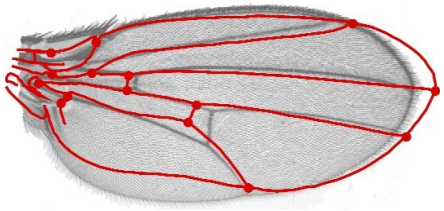
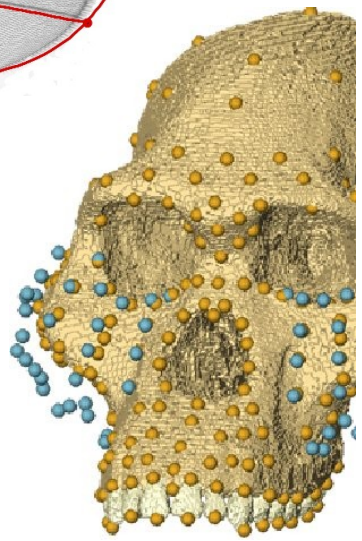
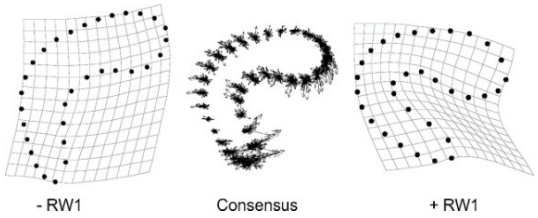
ANALYSIS OF PHENOTYPE

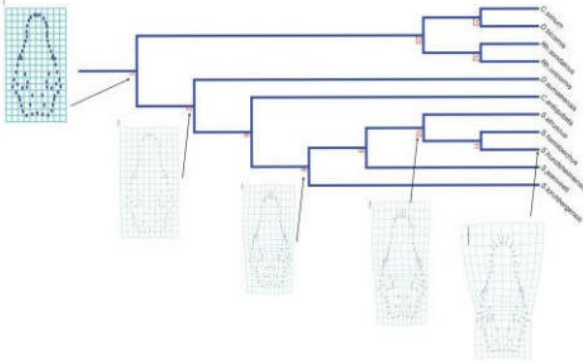












esf evropský sociální fond v ČR
EVROPSKÁ UNIE
MINISTERSTVO ŠKOLSTVÍ, MLÁDEŽE A TĚLOVÝCHOVY
OP Vzdělávání pro konkurenceschopnost
INVESTICE DO ROZVOJE VZDĚLÁVÁNÍ

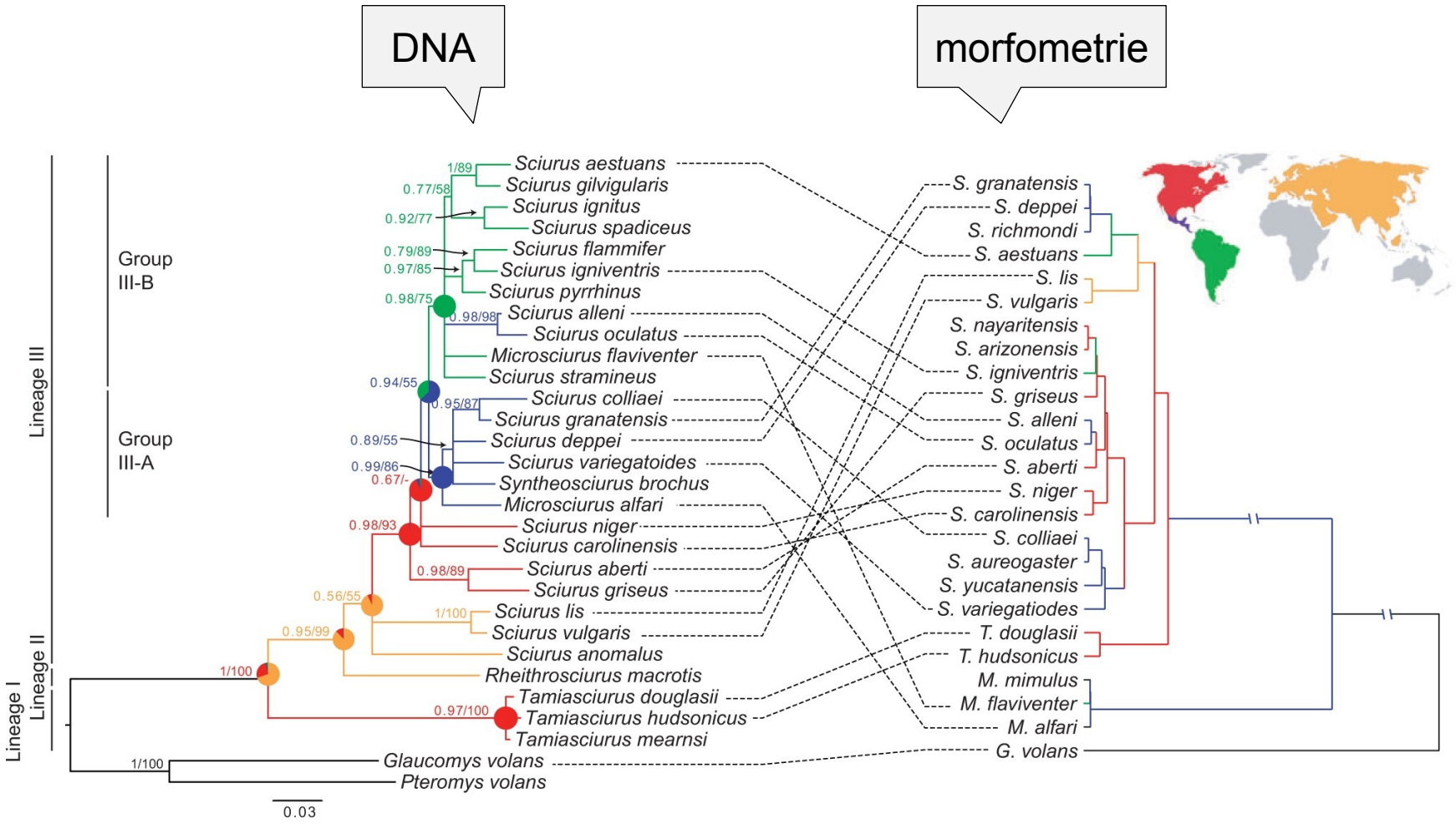
Genetic methods and morphology

What is the genetic basis of a morphological trait?
(quantitative trait loci = QTL)

Trait variation in time (phylogenesis)

Trait variation depending on other factors

Some methods are shared (e.g. PCA)



Pečnerová et al. *Syst. Biol*. (2015)

Molecular vs. morphological traits

amount (10^3 - $>10^6$ vs. 10^2)

independence

phylogenetic scale
(with mol. traits we can compare e.g. bacteria and vertebrates)

larger number of taxa

usually represent many genes
(vs. e.g. mtDNA, cpDNA)

we can also study museum/fossil material

traits	variation	genetic determination
qualitative	discrete	one to a few genes of large effects
quantitative – plastic	continuous	many genes of small effects + non-genetic influence
quantitative – meristic	continuous scale of discrete traits	many genes of small effects + non-genetic influence (threshold traits)

Many so-called qualitative traits have, in fact, *quantitative* basis!

Analysis of phenotype

qualitative traits

epigenetic traits

traditional morphometrics

geometric morphometrics

Qualitative trais

Mendelian inheritance, 1 – a few genes

mutations in *D. melanogaster*

mutations of *Hox* genes:

Antennapedia, *Ultrabithorax*

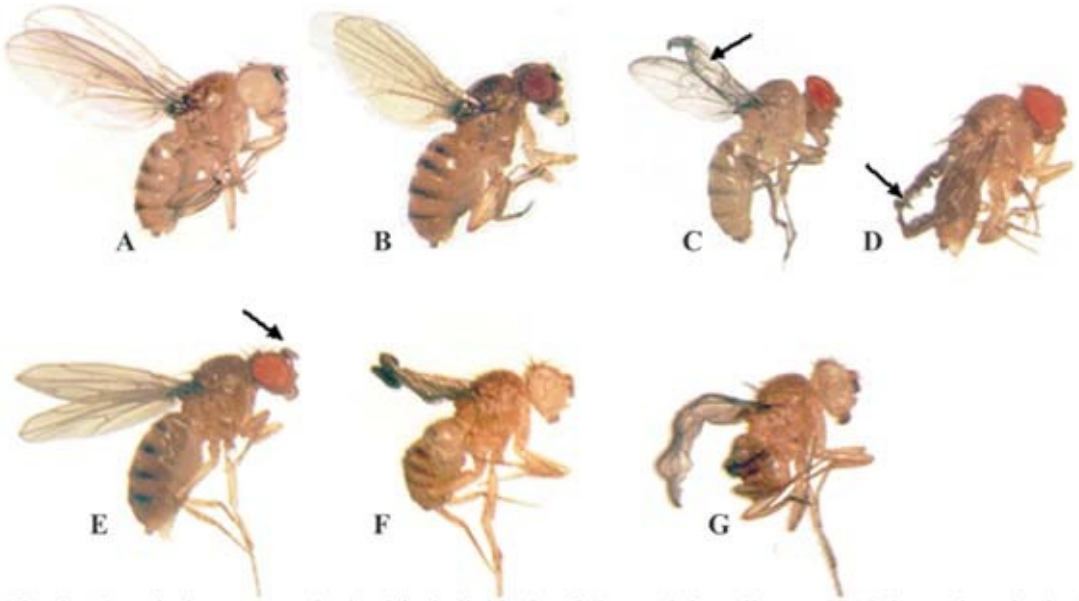
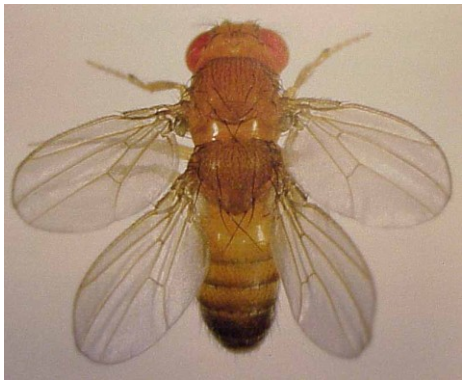


Fig. 2. Altered phenotypes obtained in individuals of *Drosophila willistoni* 17A2 isostrain submitted to temperature stress. Picture shows female individuals with altered morphologies: A. *white* (white eyes); B. *sepia* (brown eyes); C. *blistered* (arrow indicates the presence of blisters on the wings); D. *Curly* (curved wings indicated by the arrow); E. with apparently fused antennae (indicated by the arrow) - see figure 5 for details of the mutant structure; F. Female *white* and *Curly*; G. *white* and *blistered*.



Qualitative trais

colouration:

scarlet tiger moth (přástevník hluchavkový, *Callimorpha dominula*), grove snail (páskovka hajní, *Cepaea nemoralis*), beetle elytrons

mammals: ~15 domestic and laboratory species - cat, mouse, guinea pig, weasel, leopard, mink, horse

pigments:

eumelanin, phaeomelanin, carotens, haemoglobin, trichosiderin

Qualitative trais

main allelic series – mouse:

A = *agouti* (colour structure along hair)
B = *brown* (protein component of pigment granules)
C = *albino* (reduction of number of pigmented lesion)
D = *dilute* (aggregation of pigmented lesions)
E = *extension* (changes in amount of eumelanin)

agouti

dilute

DBA = *d*ilute–*b*rown–non-*a*gouti

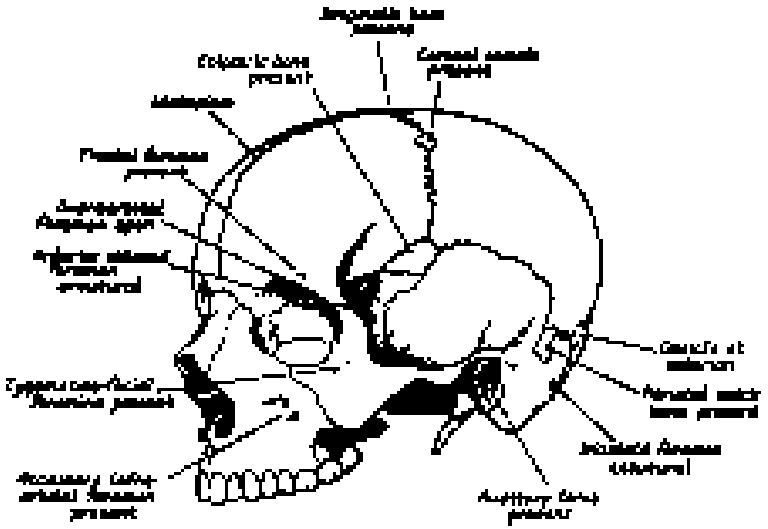
albino

Qualitative trais

main allelic series – cat:

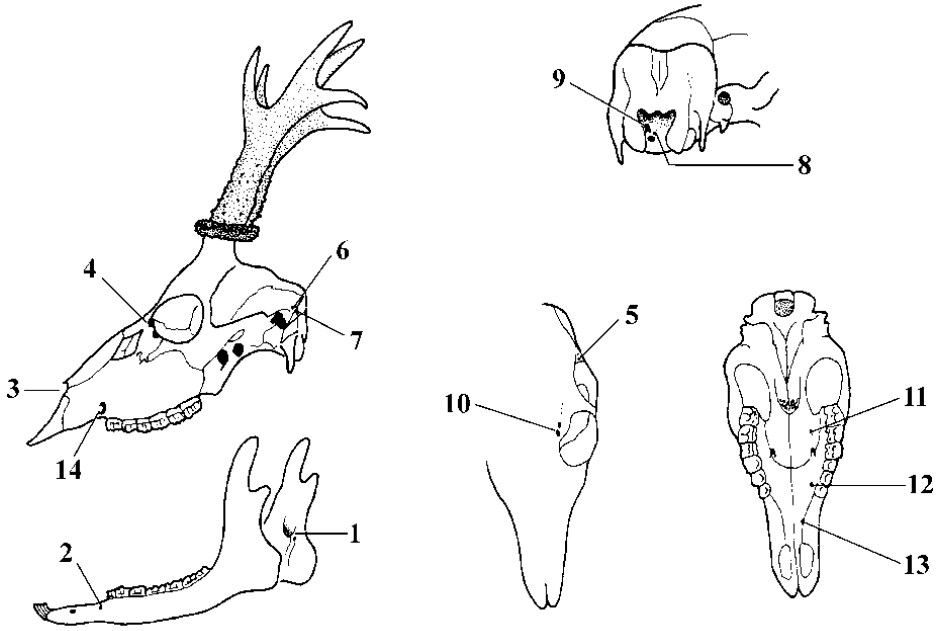
A-agouti, T-tabby, B-brown, O-orange,
S-white spotting, W-white, L-long hairs

Epigenetic traits



Epigenesis = developmental interactions over/outside of alteration of genes

basic criterion = absence of correlation between the trait and its size



Quantitative traits

Relationship of genotype and phenotype:

$V_P = V_G + V_E$

V_P = total phenotype variance

V_G = genotype variance

V_E = variance caused by environment

$V_G = V_A + V_D + V_I$

A = additivity; D = dominance; I = epistasis

Quantitative traits

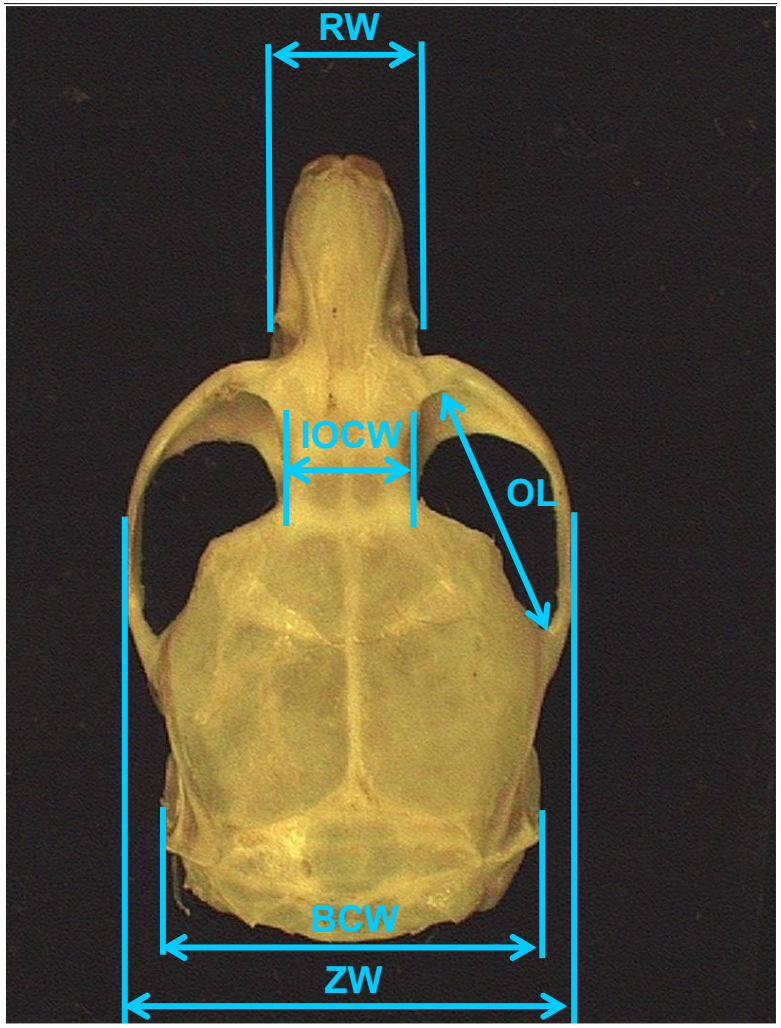
Heritability, h^2:
= measure of heritable part of phenotypic variability

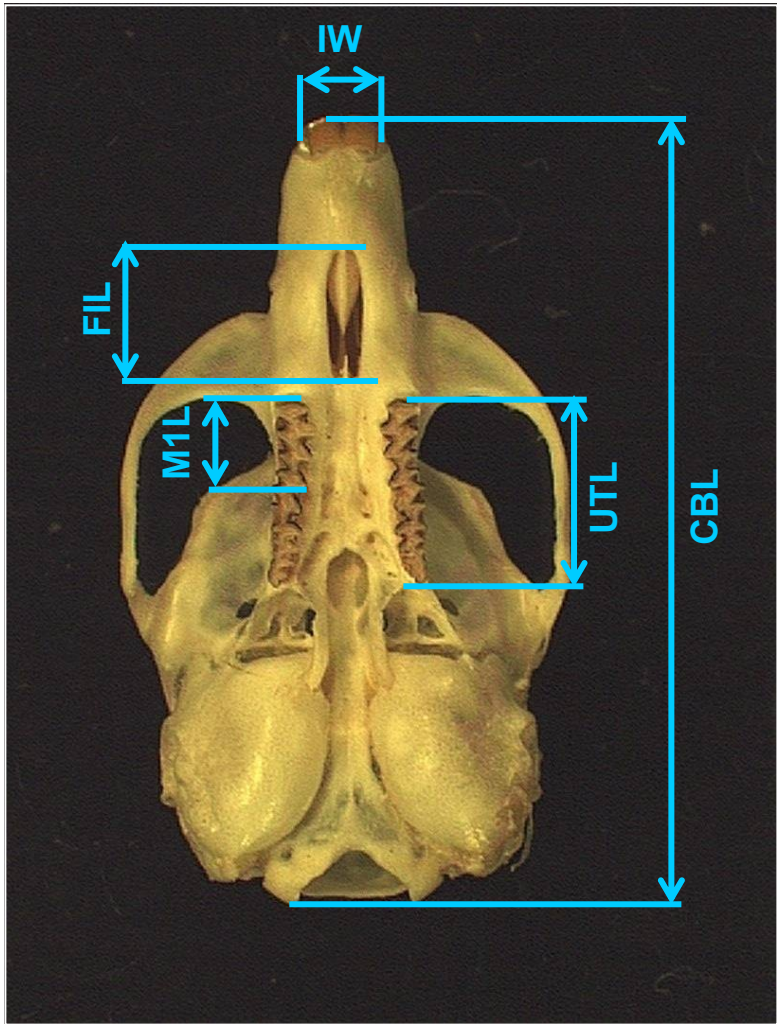
says, to what extent phenotypic variance has genetic basis

True heritability:

in narrow sense $h^2 = V_A / V_P$
in broad sense $h^2 = V_G / V_P$

Traditional morphometrics



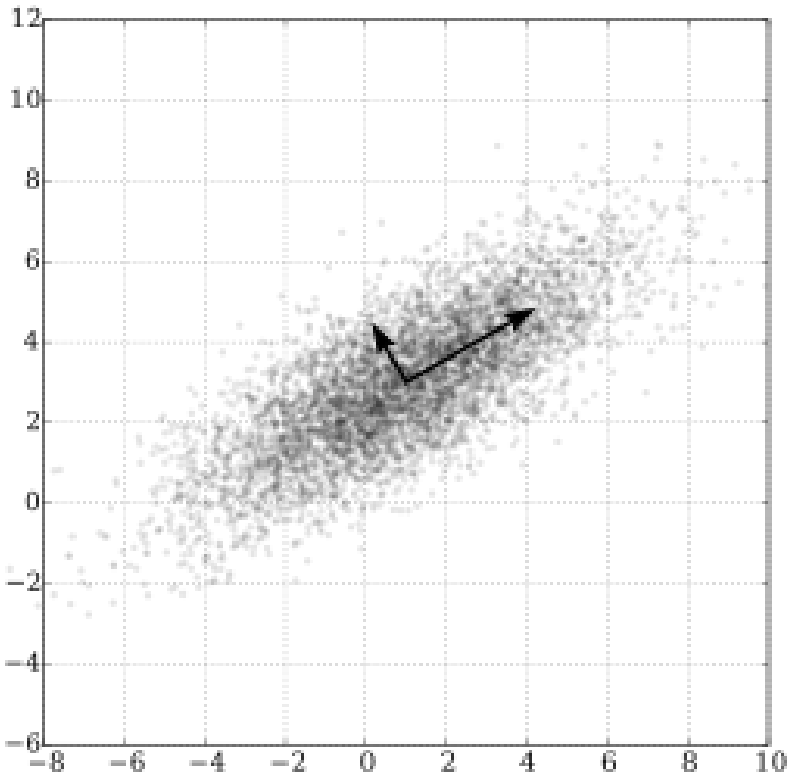


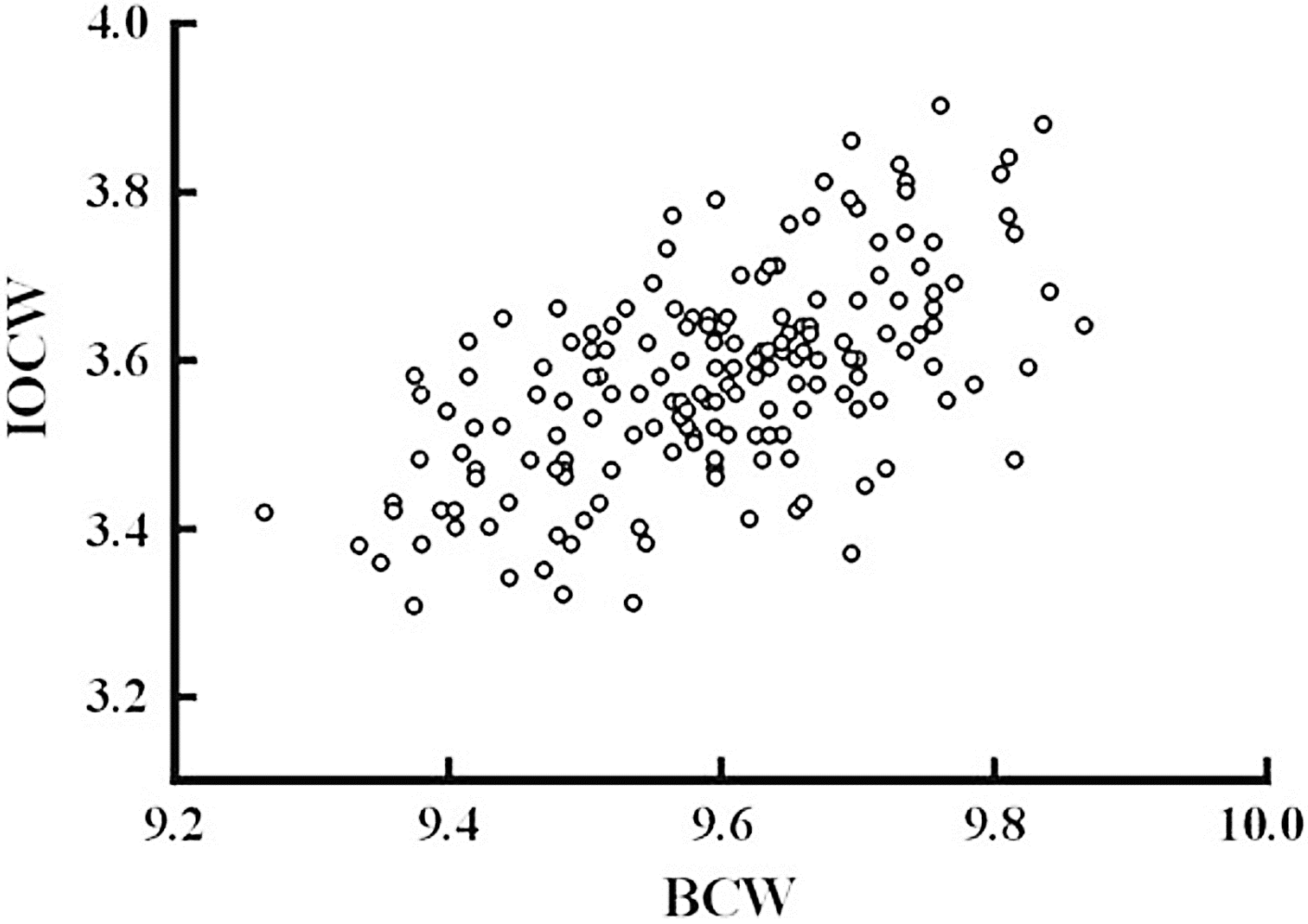
Principal components analysis (PCA)

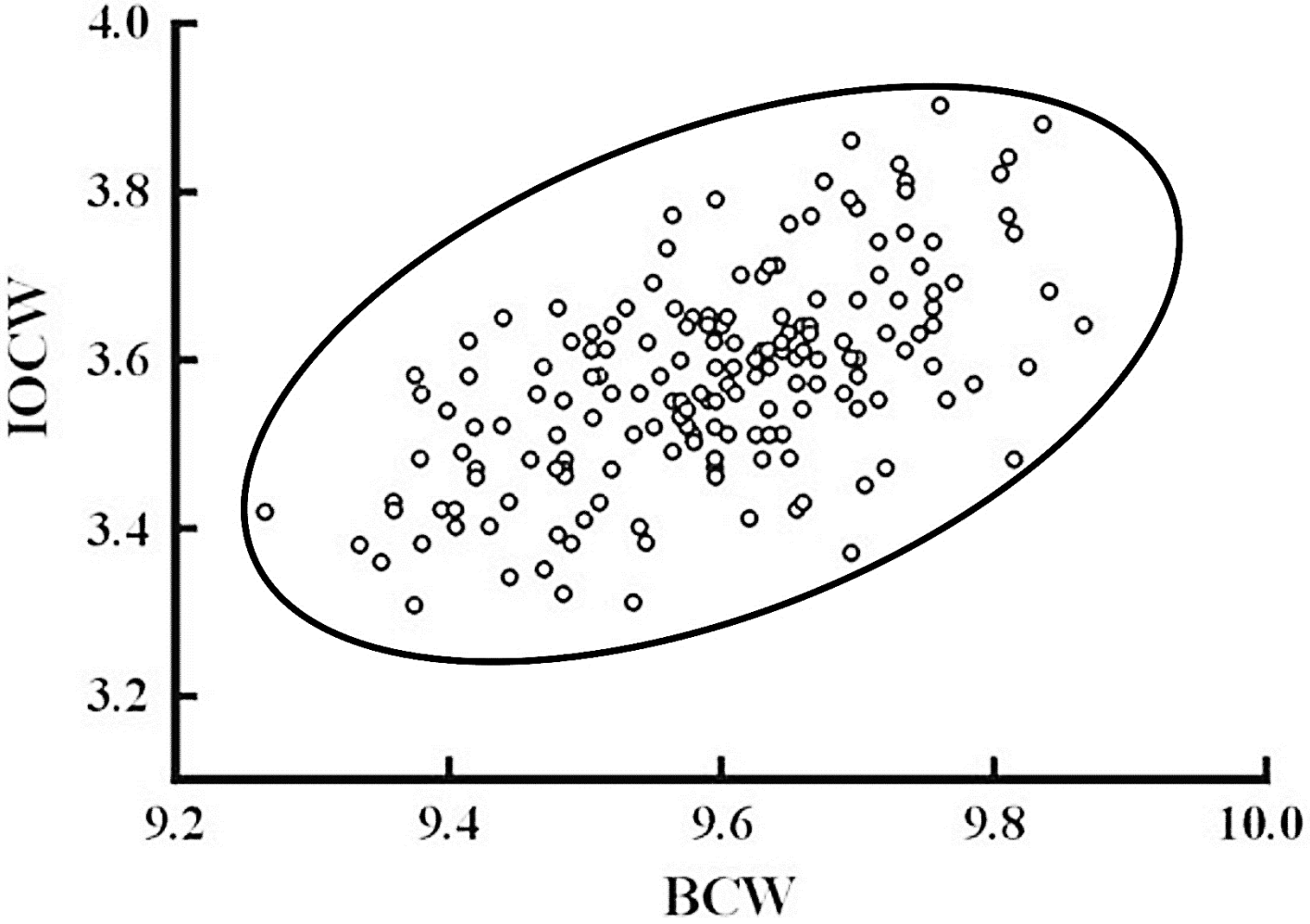
reduction of data dimensionality with as low information loss as possible

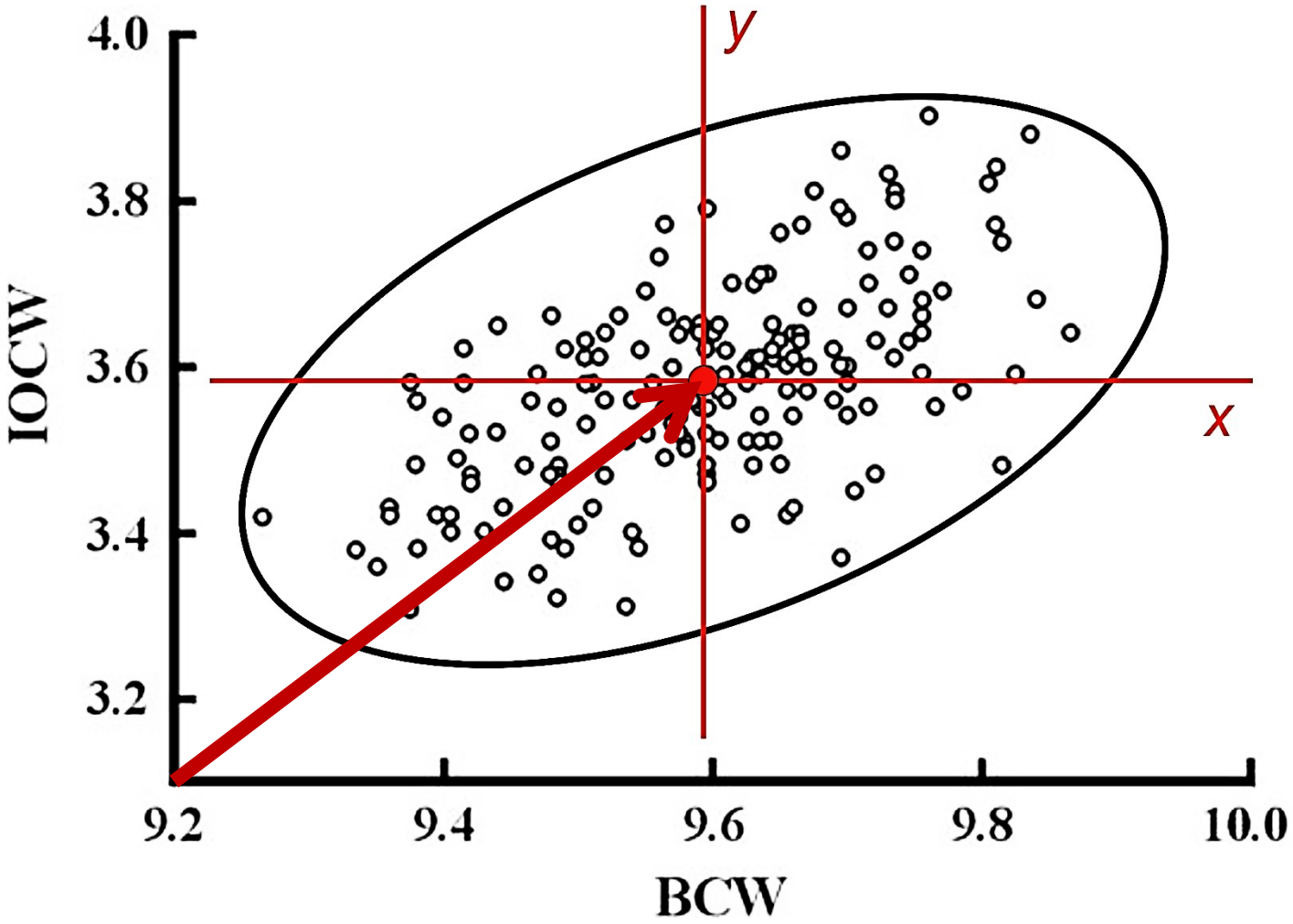
exploratory data analysis

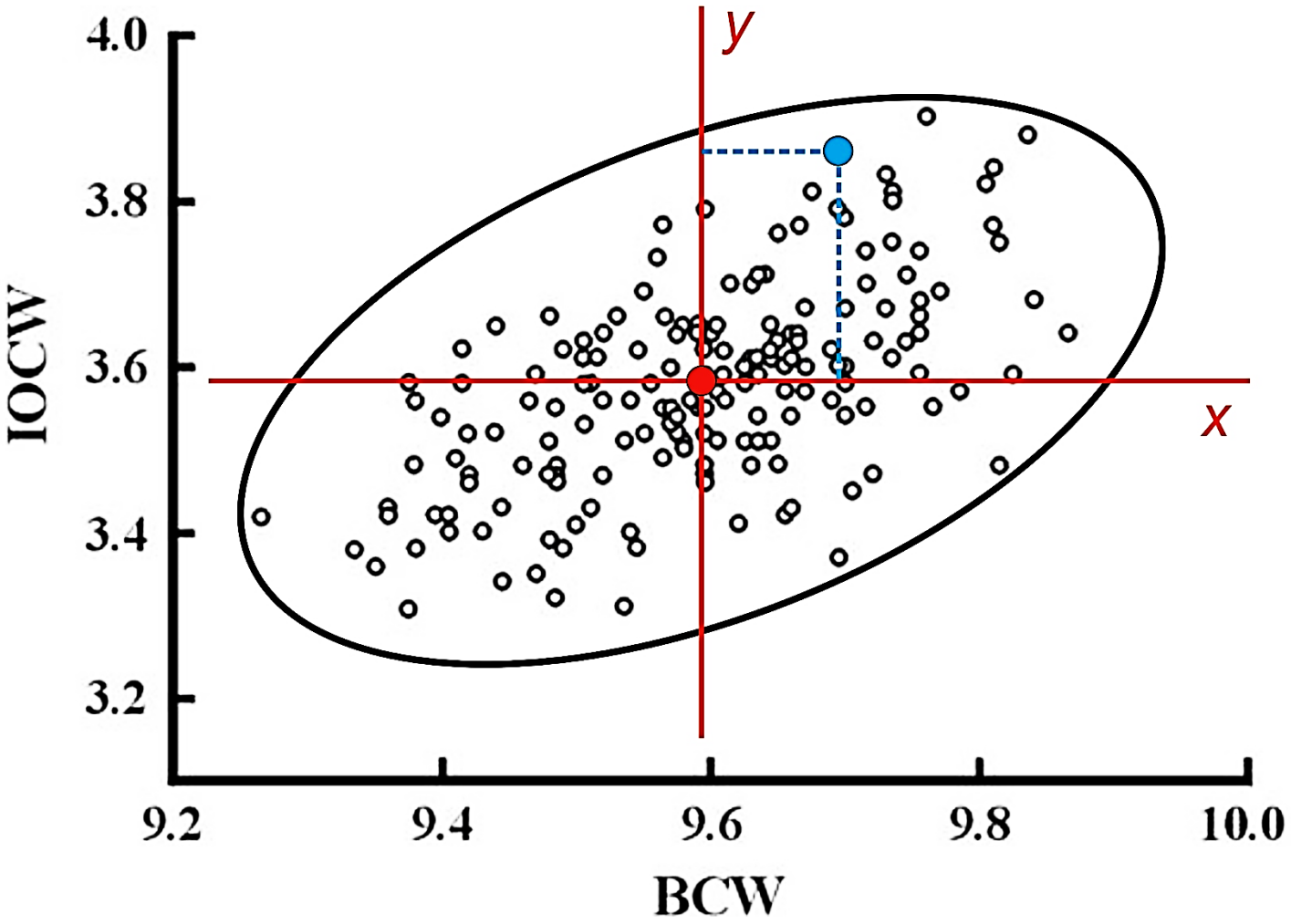
making predictive models

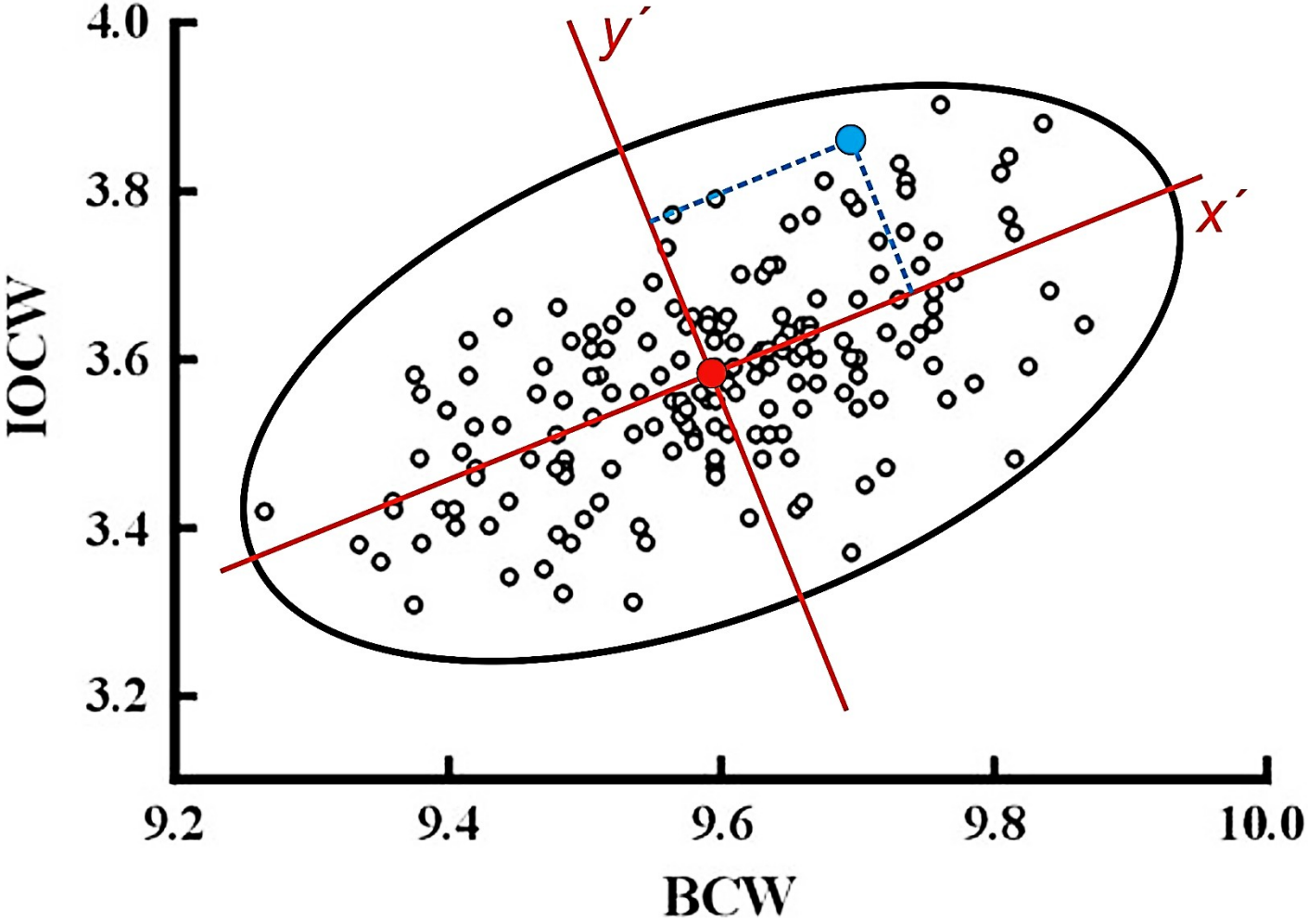



4.0
3.8
3.6
3.4
3.2
IOCW
9.2
9.4
9.6
9.8
10.0
BCW

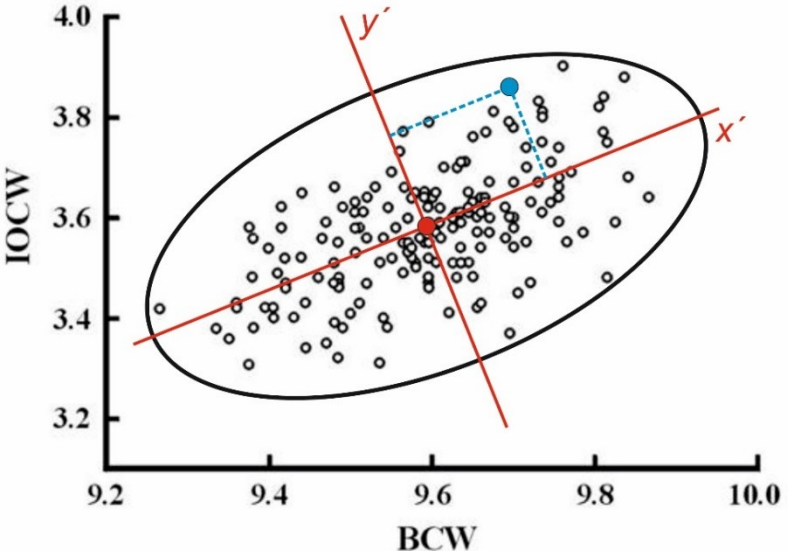

4.0
3.8
3.6
3.4
3.2
IOCW
9.2
9.4
9.6
9.8
10.0
BCW


y
x
4.0
3.8
3.6
3.4
3.2
IOCW
9.2
9.4
9.6
9.8
10.0
BCW


y
x
4.0
3.8
3.6
3.4
3.2
IOCW
9.2
9.4
9.6
9.8
10.0
BCW


4.0
3.8
3.6
3.4
3.2
IOCW
9.2
9.4
9.6
9.8
10.0
BCW
y´
x´

n individuals
p variables



correlation or covariance matrix

eigenvalue = latent root (latentní kořen)

eigenvector = latent vector (latentní vektor)

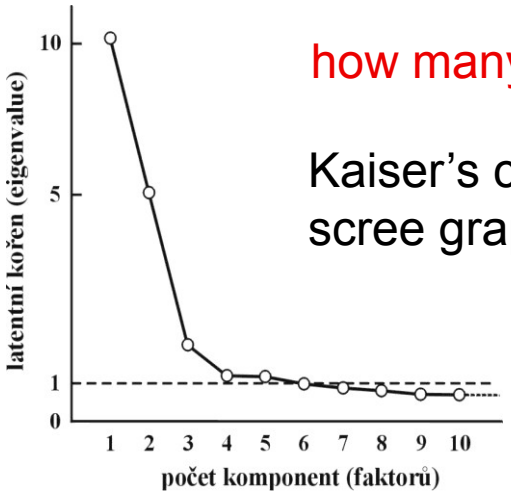
PC1: vysvětluje největší podíl variability

PC2: druhý největší podíl variability atd.

size vector

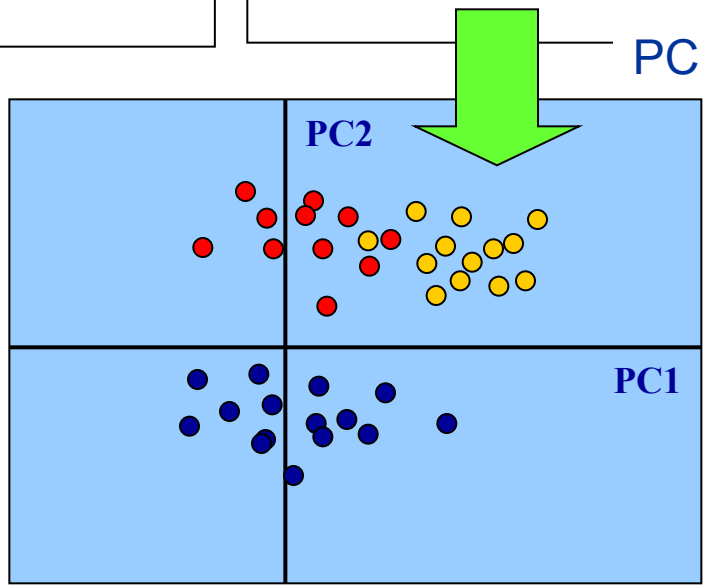
loadings (zátěže)

PC scores



how many components?

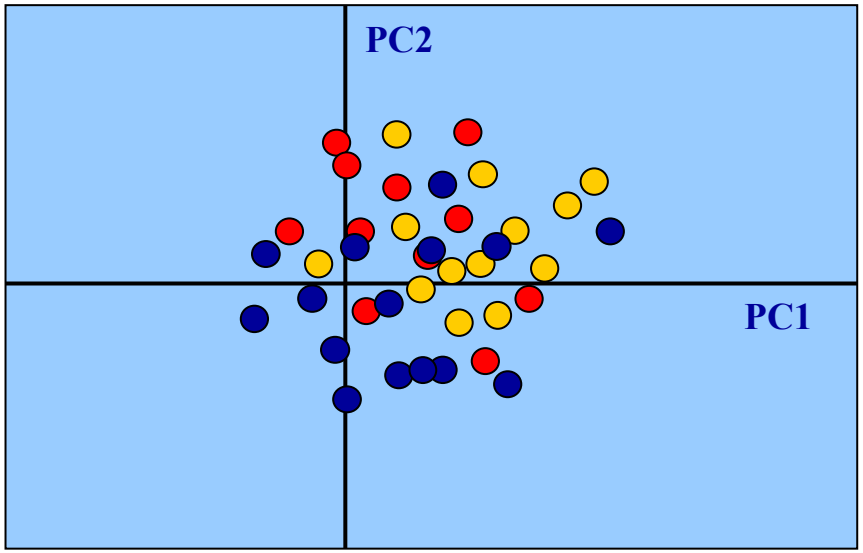
Kaiser's criterion: 5 PCs
scree graph: 3–4 PCs



Problem of multiple groups:

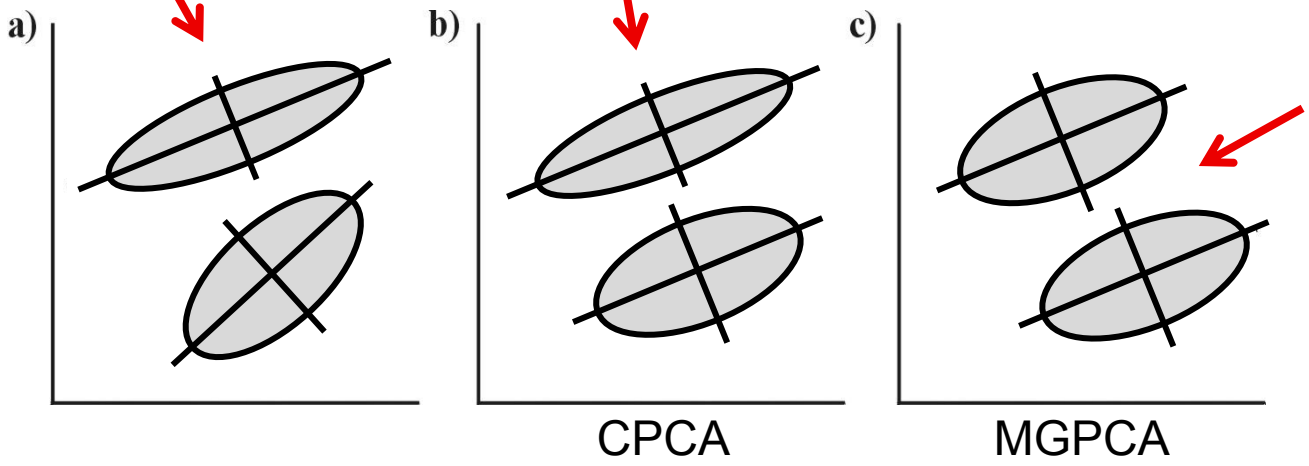
CPCA (common PCA)

MGPCA (multiple-group PCA)



different variation
different direction

different variation
same direction

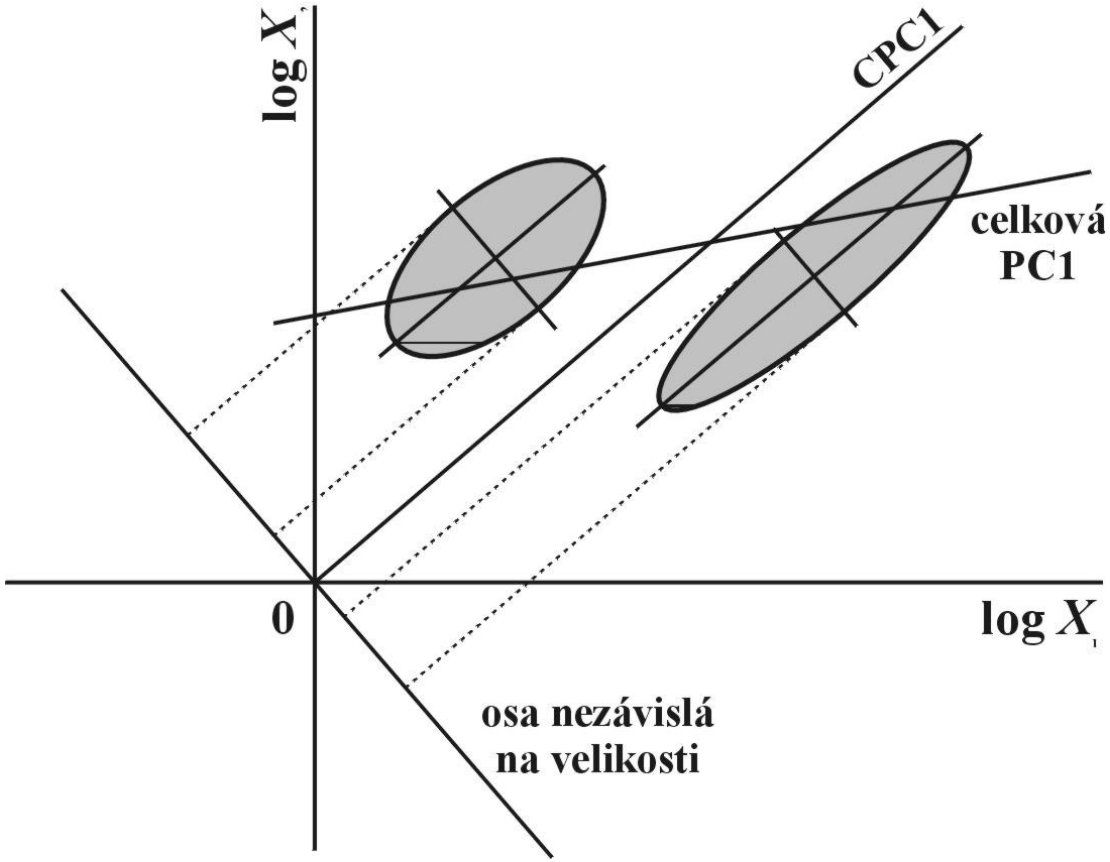


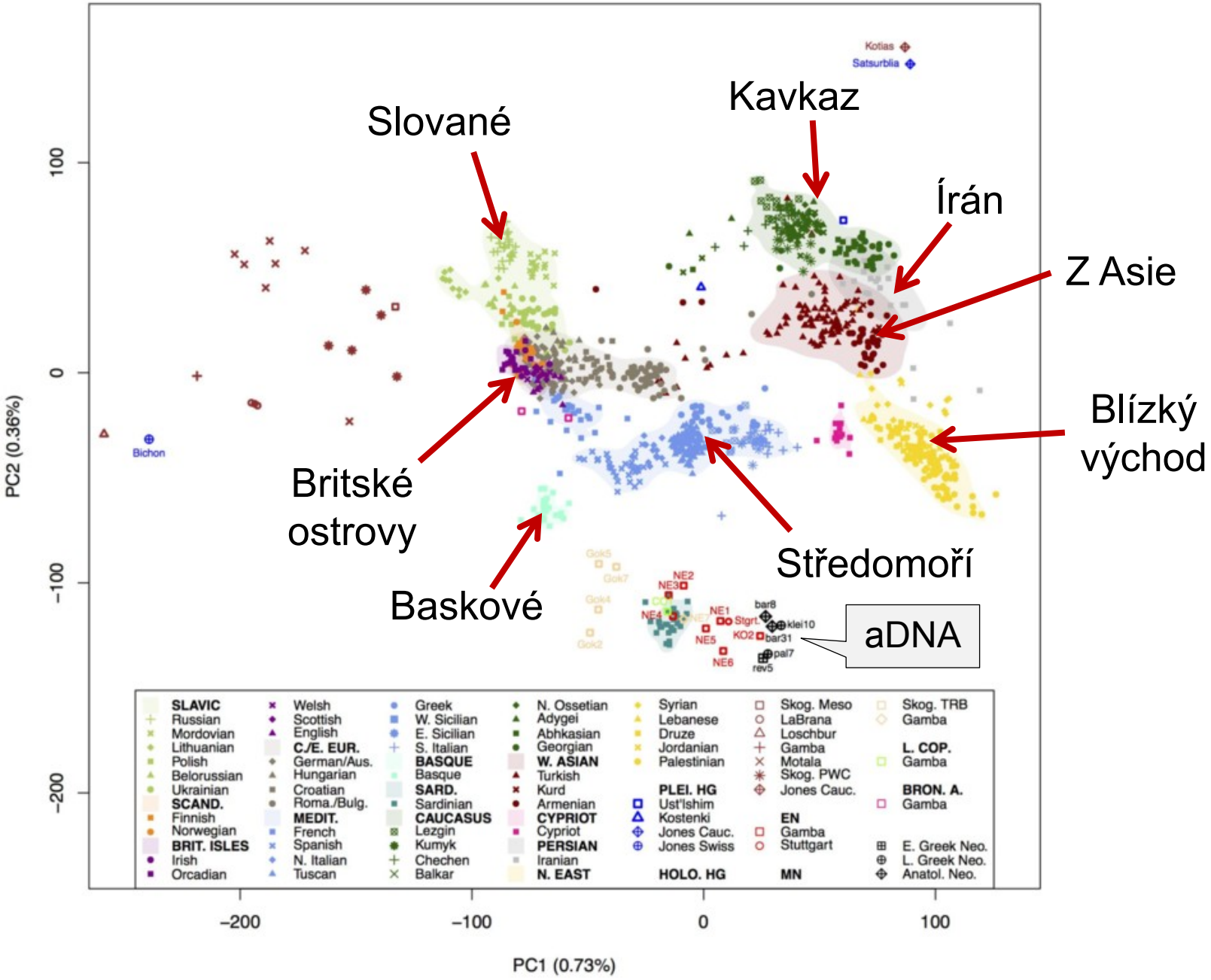
same variation
same direction

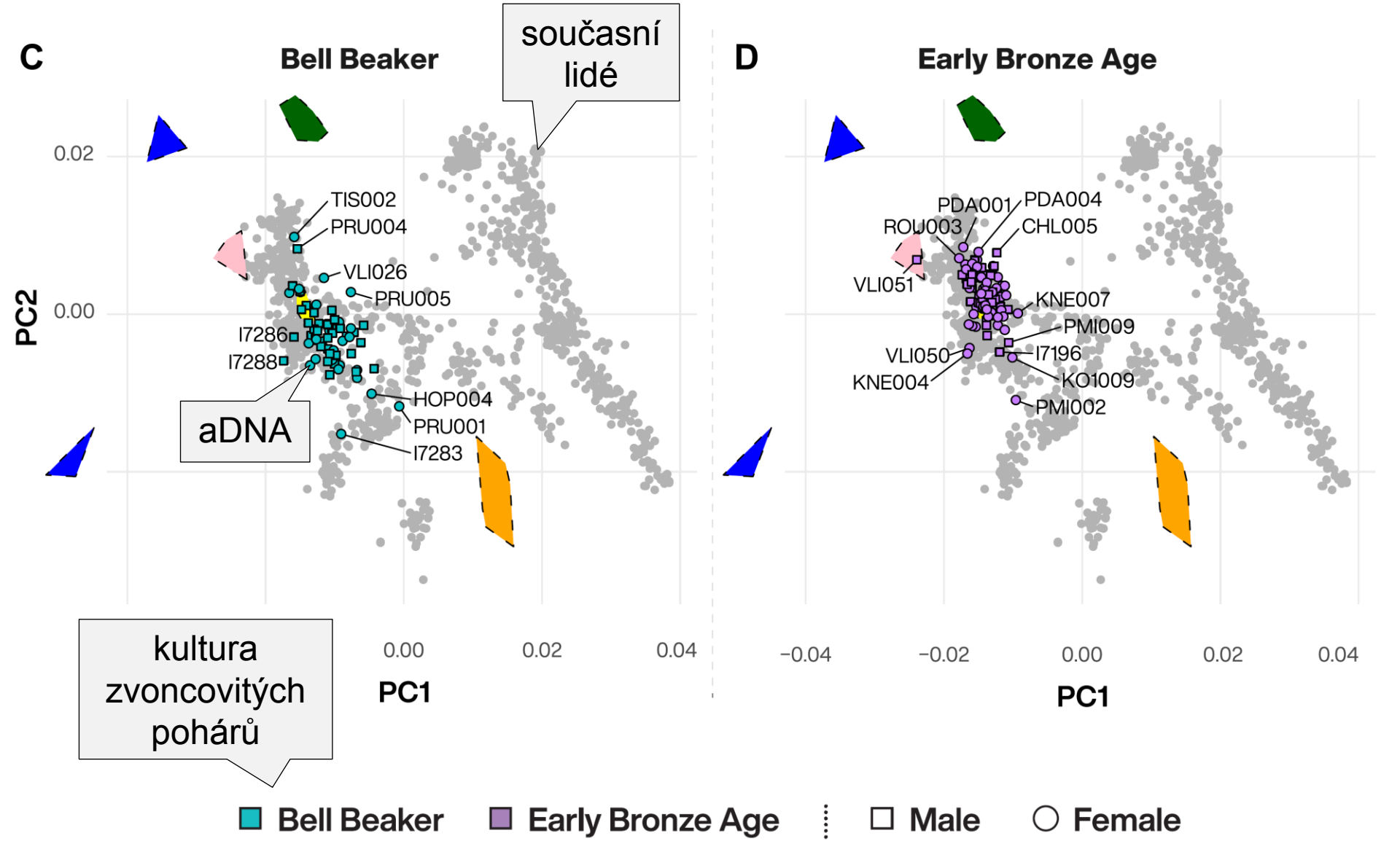
Problem of size:

omitting PC1

Burnaby's adjustment




Slované
Kavkaz
Írán
Z Asie
Blízký
východ
Britské
ostrovy
Baskové
Středomoří
aDNA
Kotias
Satsurblia
Bichon
PC2 (0.36%)
PC1 (0.73%)
100
0
−100
−200
−200
−100
0
100
SLAVIC
Russian
Mordovian
Lithuanian
Polish
Belorussian
Ukrainian
SCAND.
Finnish
Norwegian
BRIT. ISLES
Irish
Orcadian
Welsh
Scottish
English
C./E. EUR.
German/Aus.
Hungarian
Croatian
Roma./Bulg.
MEDIT.
French
Spanish
N. Italian
Tuscan
Greek
W. Sicilian
E. Sicilian
S. Italian
BASQUE
Basque
SARD.
Sardinian
CAUCASUS
Lezgin
Kumyk
Chechen
Balkar
N. Ossetian
Adygei
Abhkasian
Georgian
W. ASIAN
Turkish
Kurd
Armenian
CYPRIOT
Cypriot
PERSIAN
Iranian
N. EAST
Syrian
Lebanese
Druze
Jordanian
Palestinian
PLEI. HG
Ust'Ishim
Kostenki
Jones Cauc.
Jones Swiss
HOLO. HG
Skog. Meso
LaBrana
Loschbur
Gamba
Motala
Skog. PWC
Jones Cauc.
EN
Gamba
Stuttgart
MN
Skog. TRB
Gamba
L. COP.
Gamba
BRON. A.
Gamba
E. Greek Neo.
L. Greek Neo.
Anatol. Neo.


C
Bell Beaker
současní lidé
D
Early Bronze Age
0.02
0.00
PC2
TIS002
PRU004
VLI026
PRU005
I7286
I7288
aDNA
HOP004
PRU001
I7283
PDA001
PDA004
ROU003
CHL005
VLI051
KNE007
PMI009
VLI050
I7196
KNE004
KO1009
PMI002
0.00
0.02
0.04
−0.04
−0.02
0.00
0.02
0.04
PC1
PC1
kultura zvoncovitých pohárů
Bell Beaker
Early Bronze Age
Male
Female

Multidimensional scaling (MDS):

not only correlation/covariation matrix, all types of matrix, e.g. similarity/dissimilarity matrix

Classical MDS = principal coordinates analysis (PCoA)

Metric MDS

Non-metric MDS

Generalized MDS

Discriminant function analysis (DFA) and canonical analysis (CVA):

a priori groups

minimization of within-group variation

maximization of among-group variation

Mahalanobis (generalized) distances

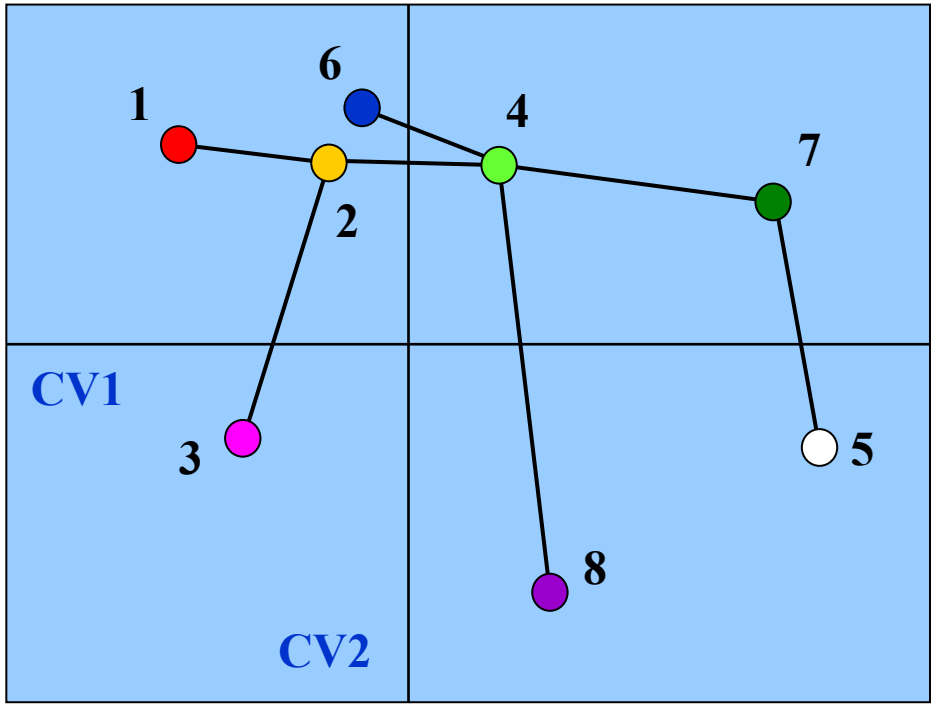
MANOVA (Wilk's Lambda, Pilai's trace)

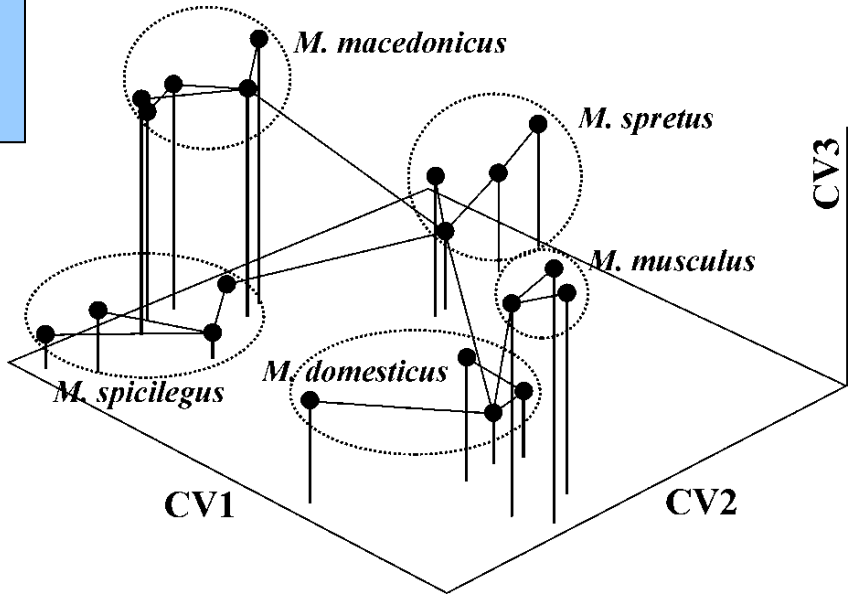
Hotelling T^2 test

stepwise DFA

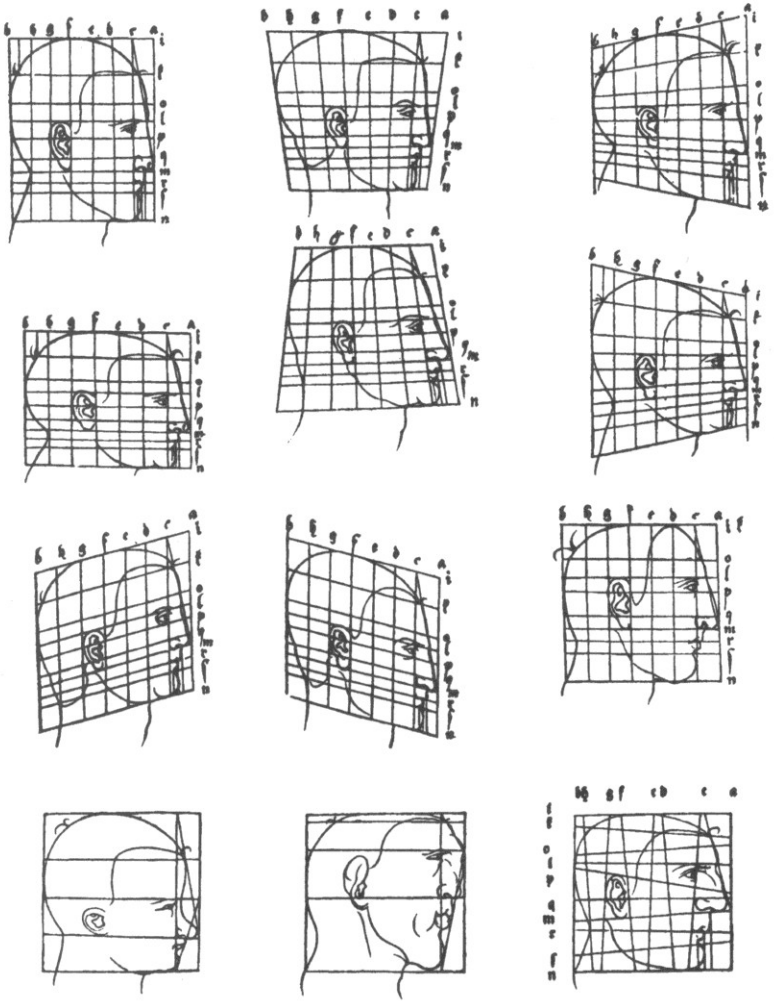
Cluster analysis

Minimum spanning tree (MST)





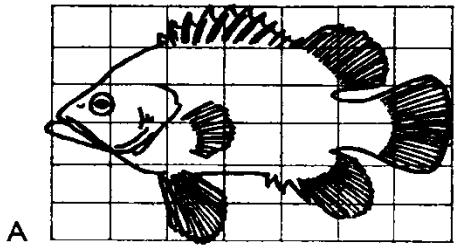
Geometric morphometrics



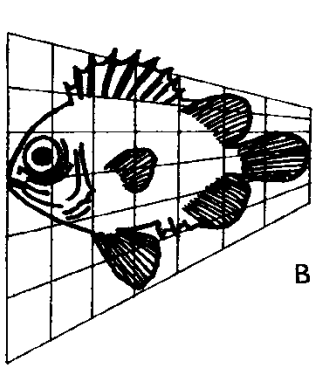
A. Dürer (1524): Vier Bücher von Menlicher Proportion.

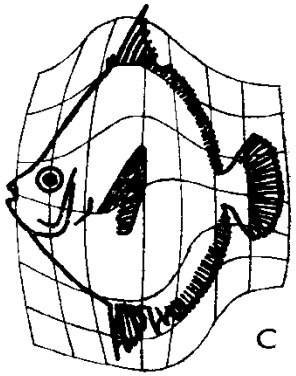
In the past, there were two different strategies in study of shape of biological objects:

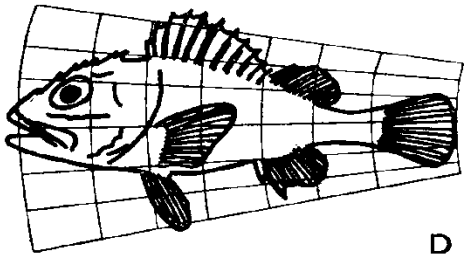
1. W. D'Arcy Thompson



Absence of quantification of shape changes!







W. A. Thompson (1917): On Growth and Form

In the past, there were two different strategies in study of shape of biological objects:

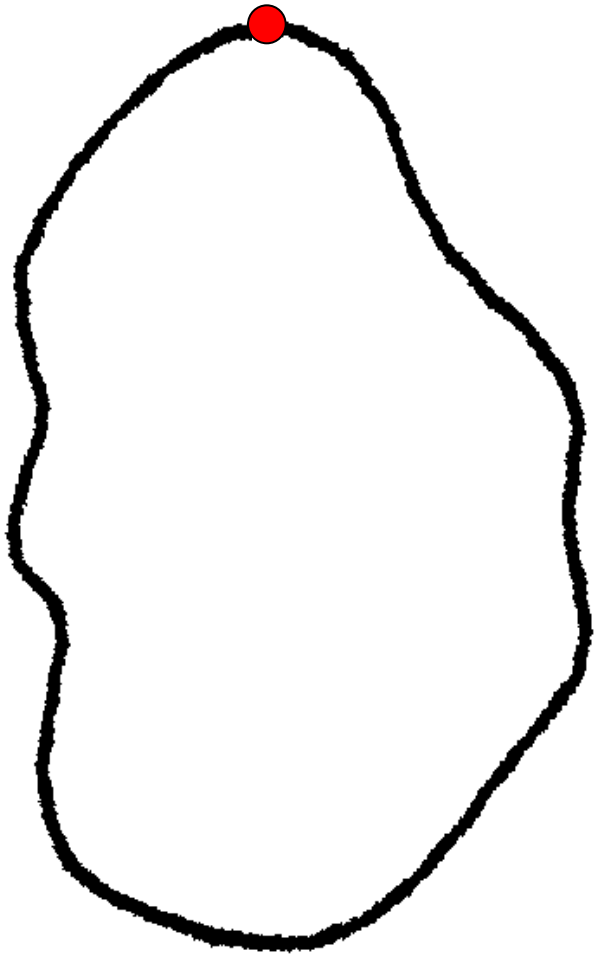
2. Traditional morfometrics:

F. Galton, K. Pearson, R.A. Fisher, S. Wright, H. Hotelling...

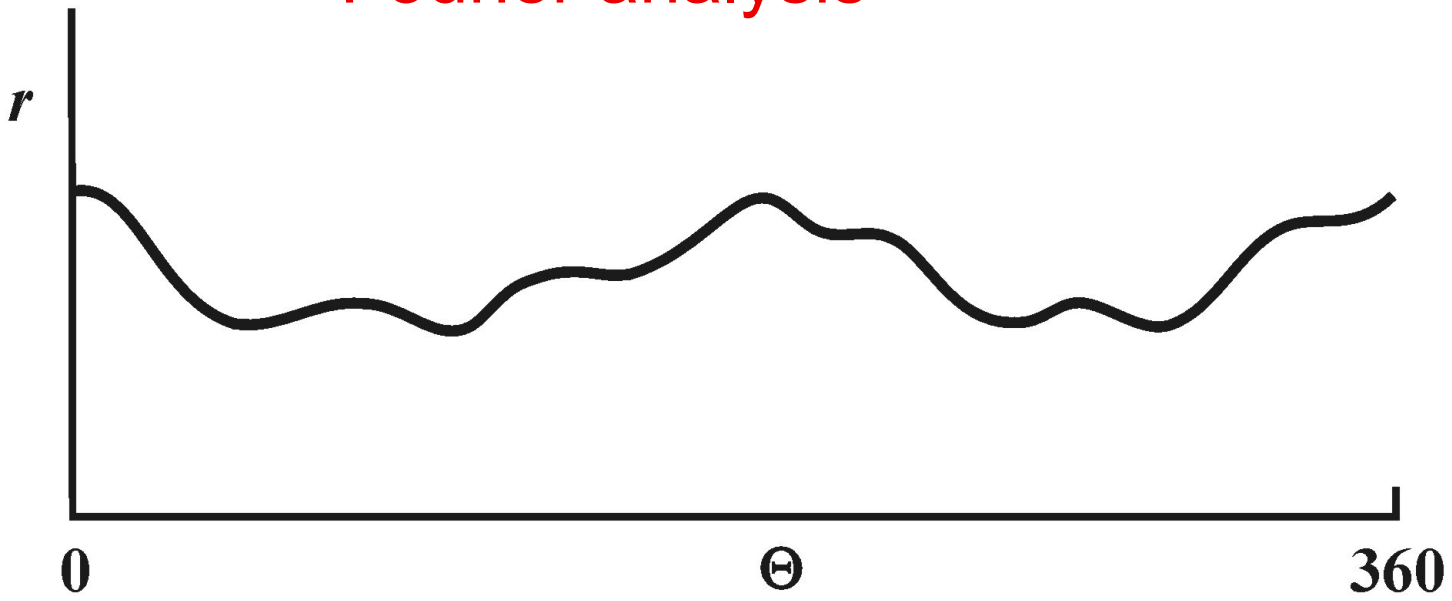
linear measurements, weights, angles, surfaces...
PCA, DFA, CVA, FA, PCoA, cluster a.

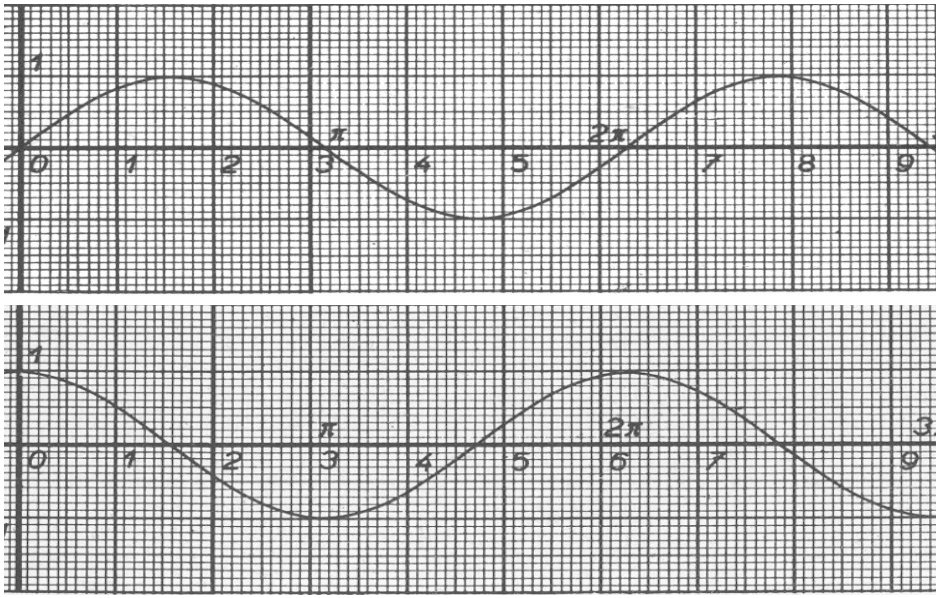
Absence of any information on shape (morphometrics)!

Geometric morphometrics I.
Analysis of closed curves



Fourier analysis

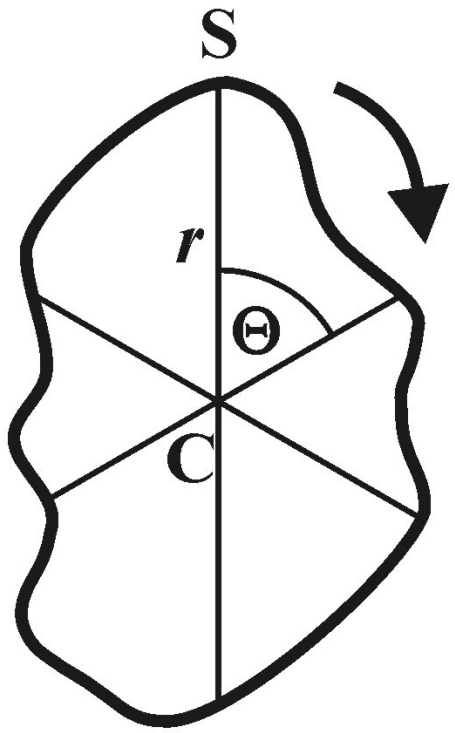




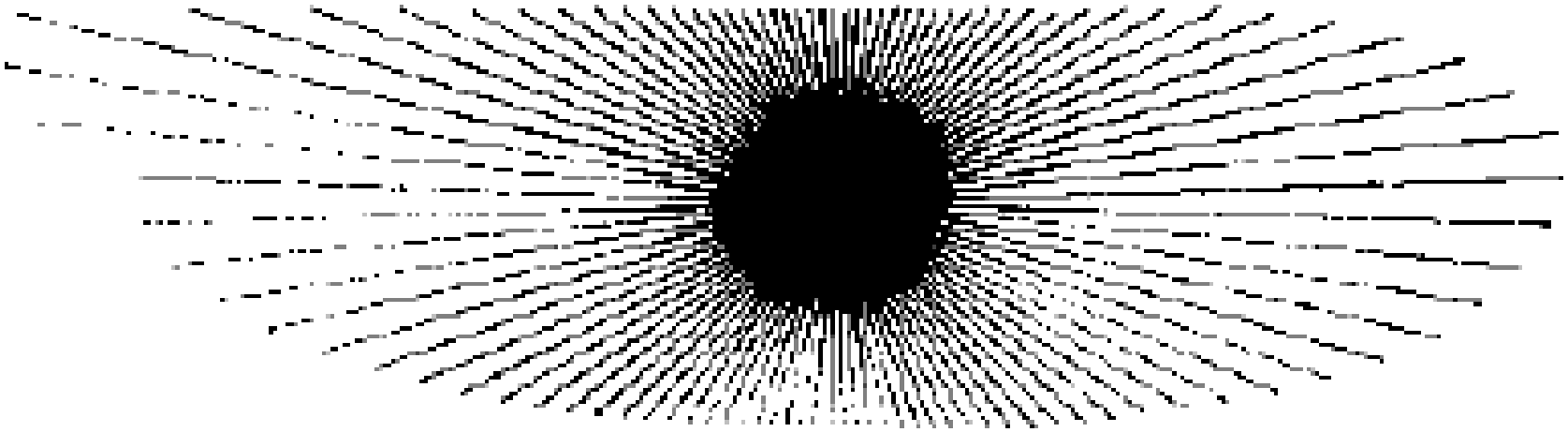
sin

harmonics, coefficients

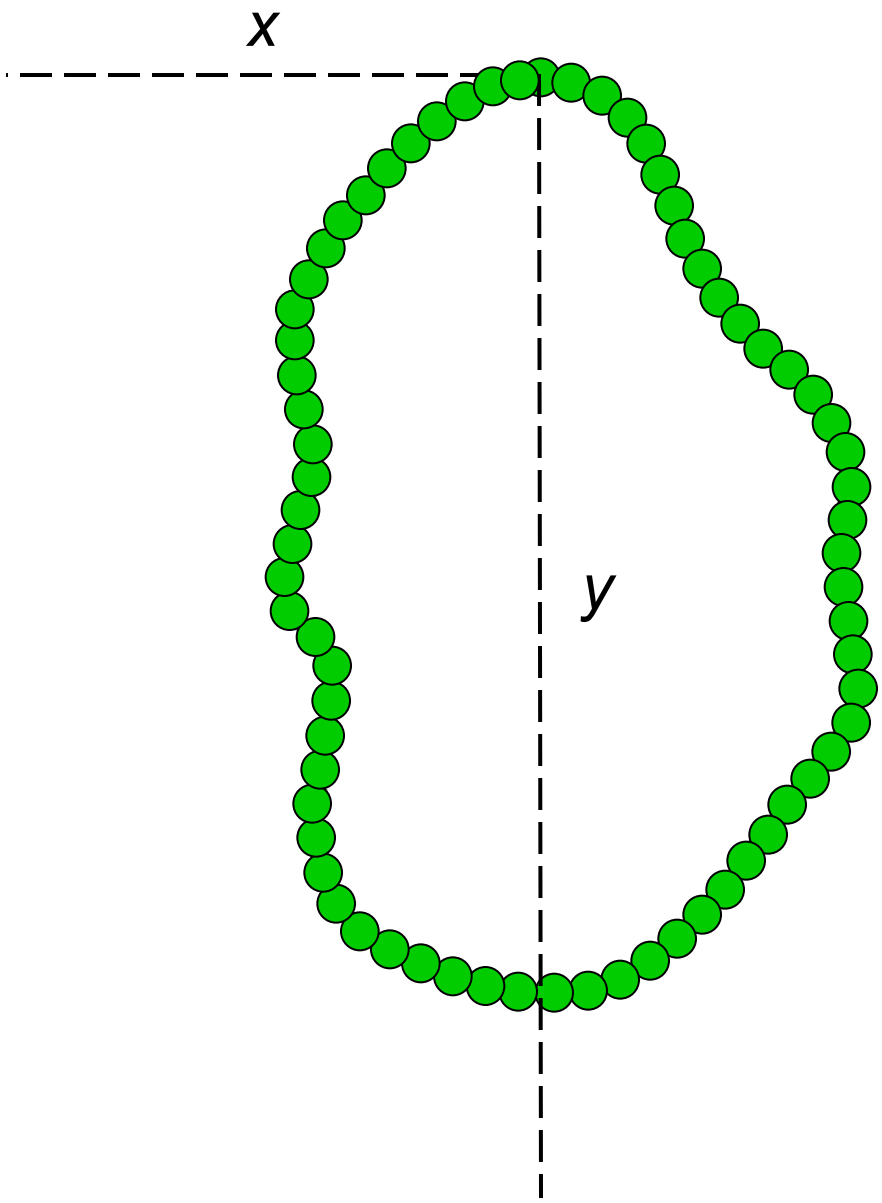
cos

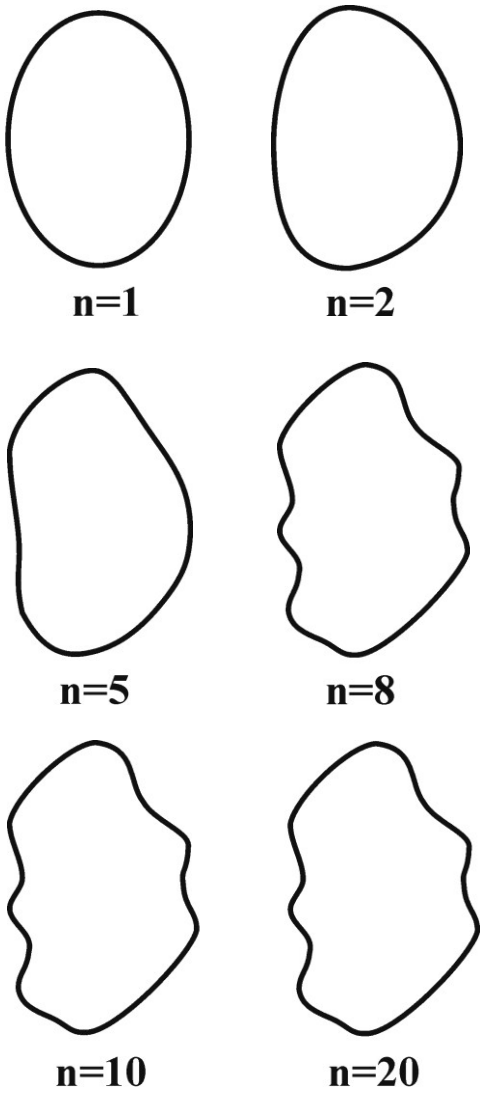


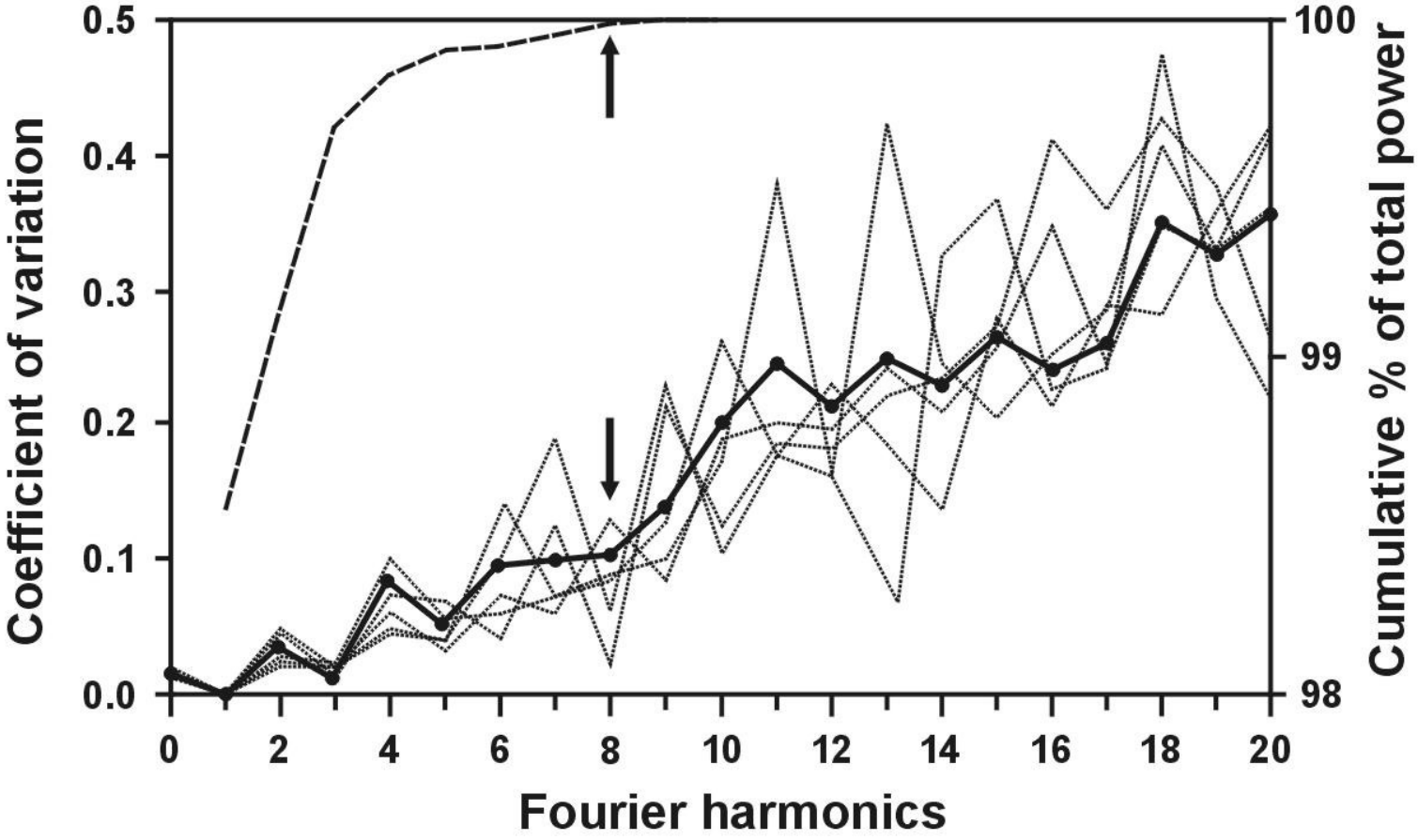
Traditional Fourier a.



Elliptic Fourier a.




n=1
n=2
n=5
n=8
n=10
n=20


0.5
0.4
0.3
0.2
0.1
0.0
Coefficient of variation
100
99
98
Cumulative % of total power
0
2
4
6
8
10
12
14
16
18
20
Fourier harmonics

Geometric morphometrics II.
Analysis of landmarks

landmarks

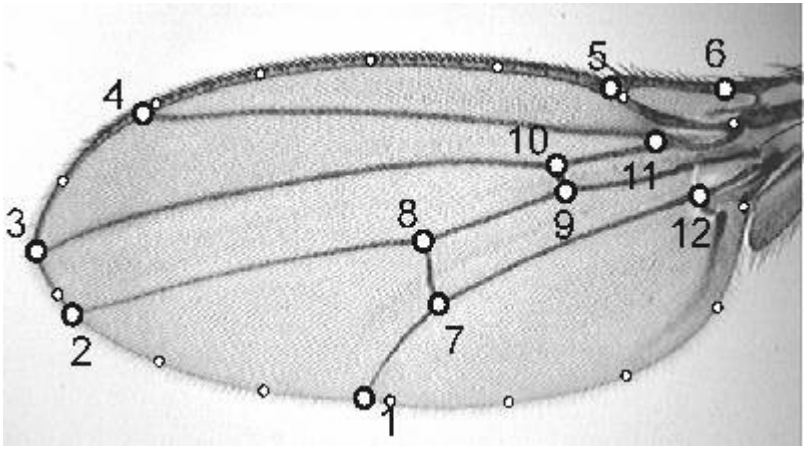
quantification of shape using shape coordinates, distinction between different shape components

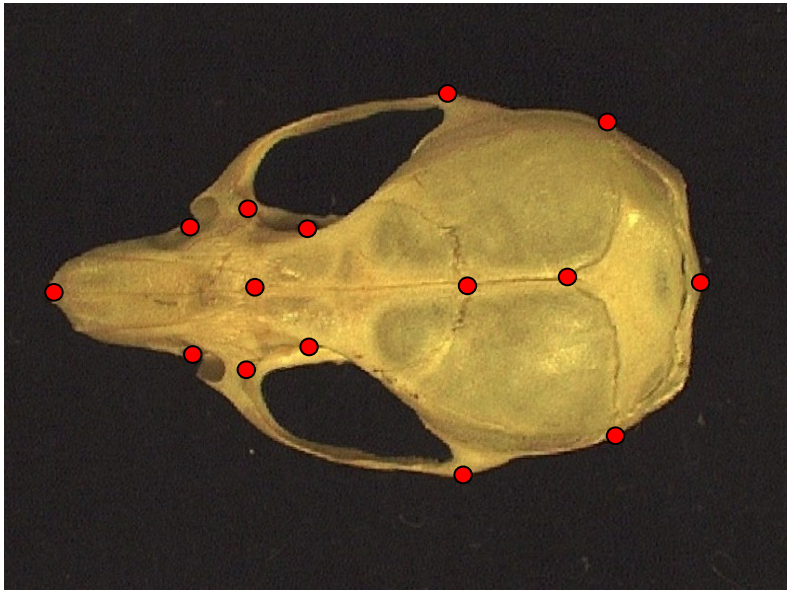
information on shape maintained during the whole analysis

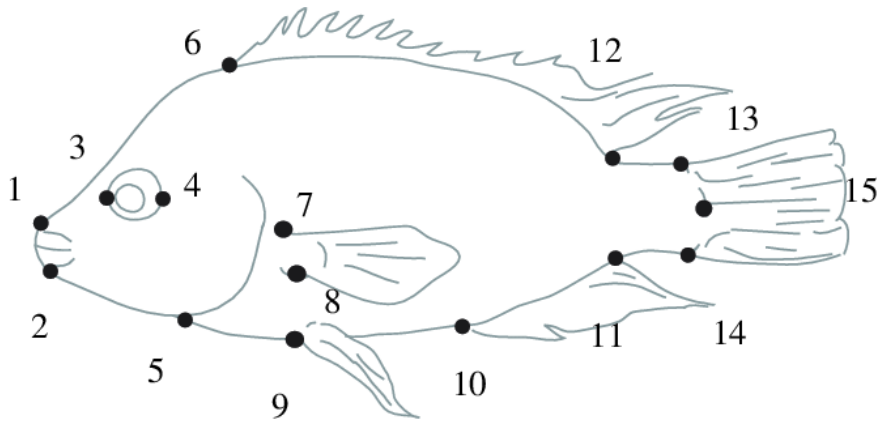
size standardization and ability to work independently with the size vector

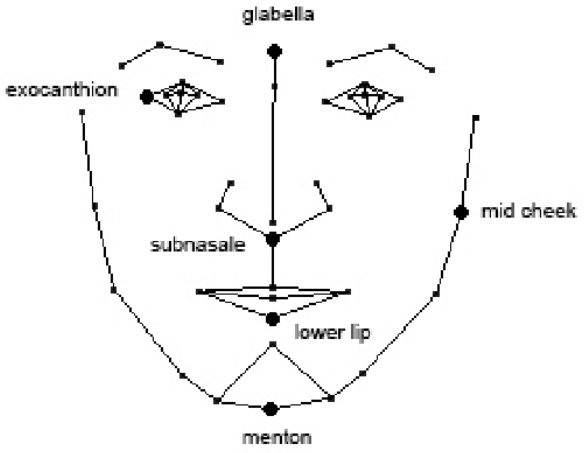
data processing with traditional morphometric methodology

landmarks = points that can be accurately localized and which are – at least in a geometric sense – homologous among the objects



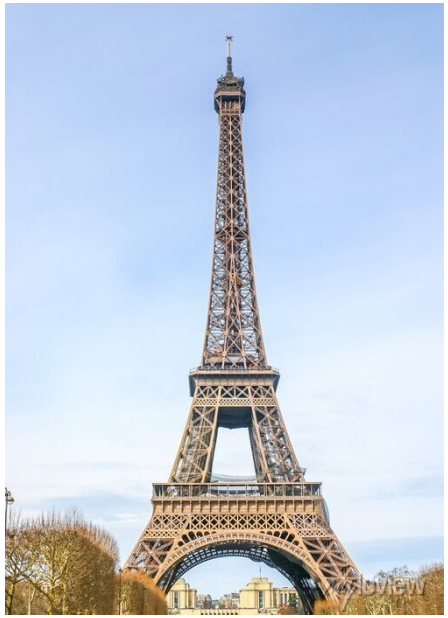


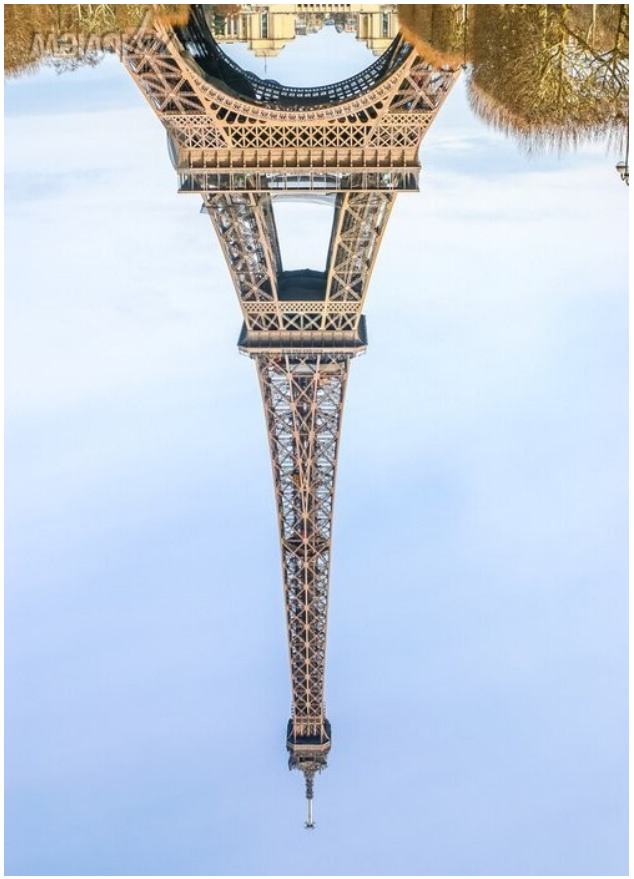




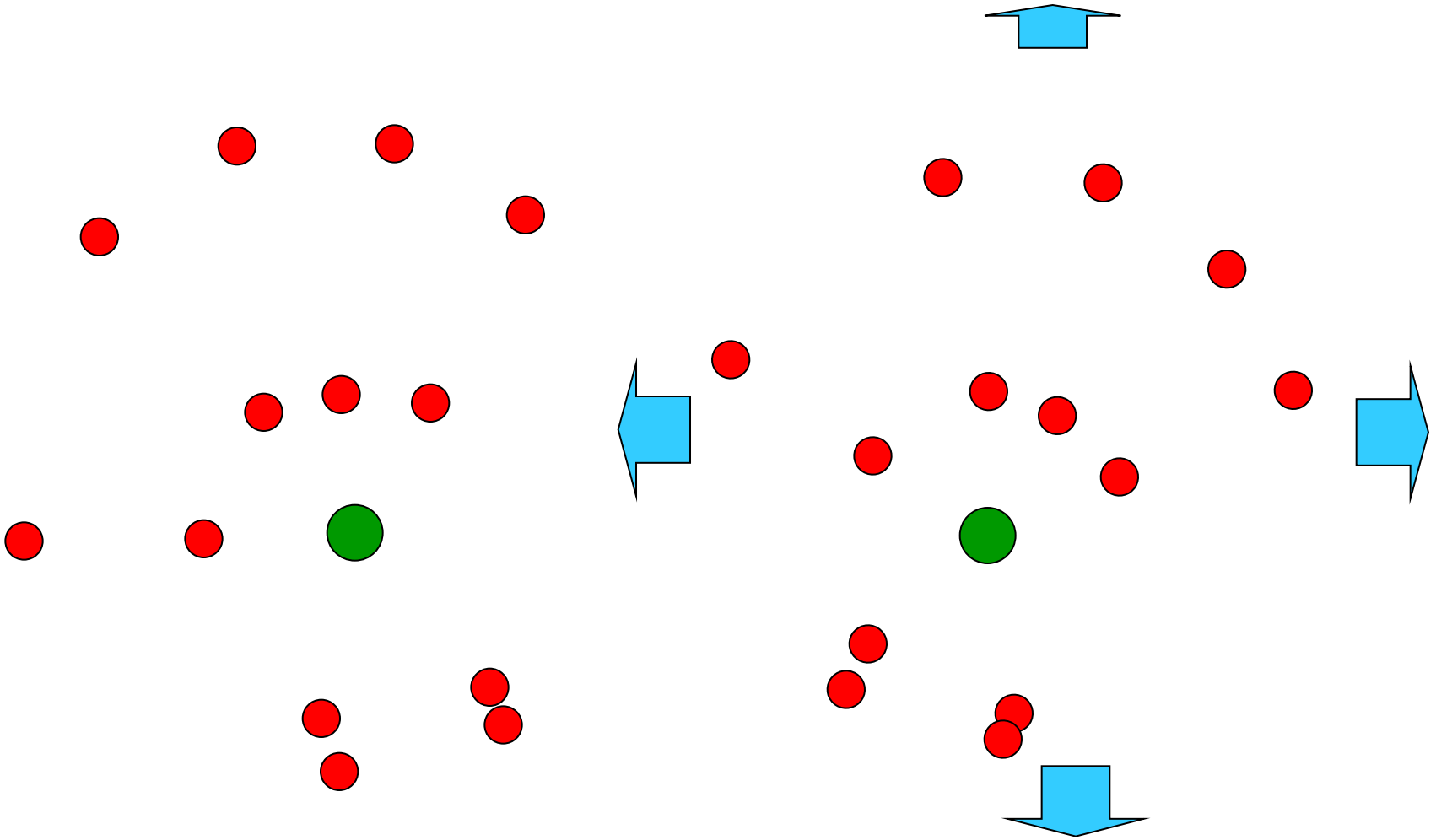
Procrustes superposition
= GLS (Generalized Least Squares)

shape = everything except information on size, position, and orientation of objects

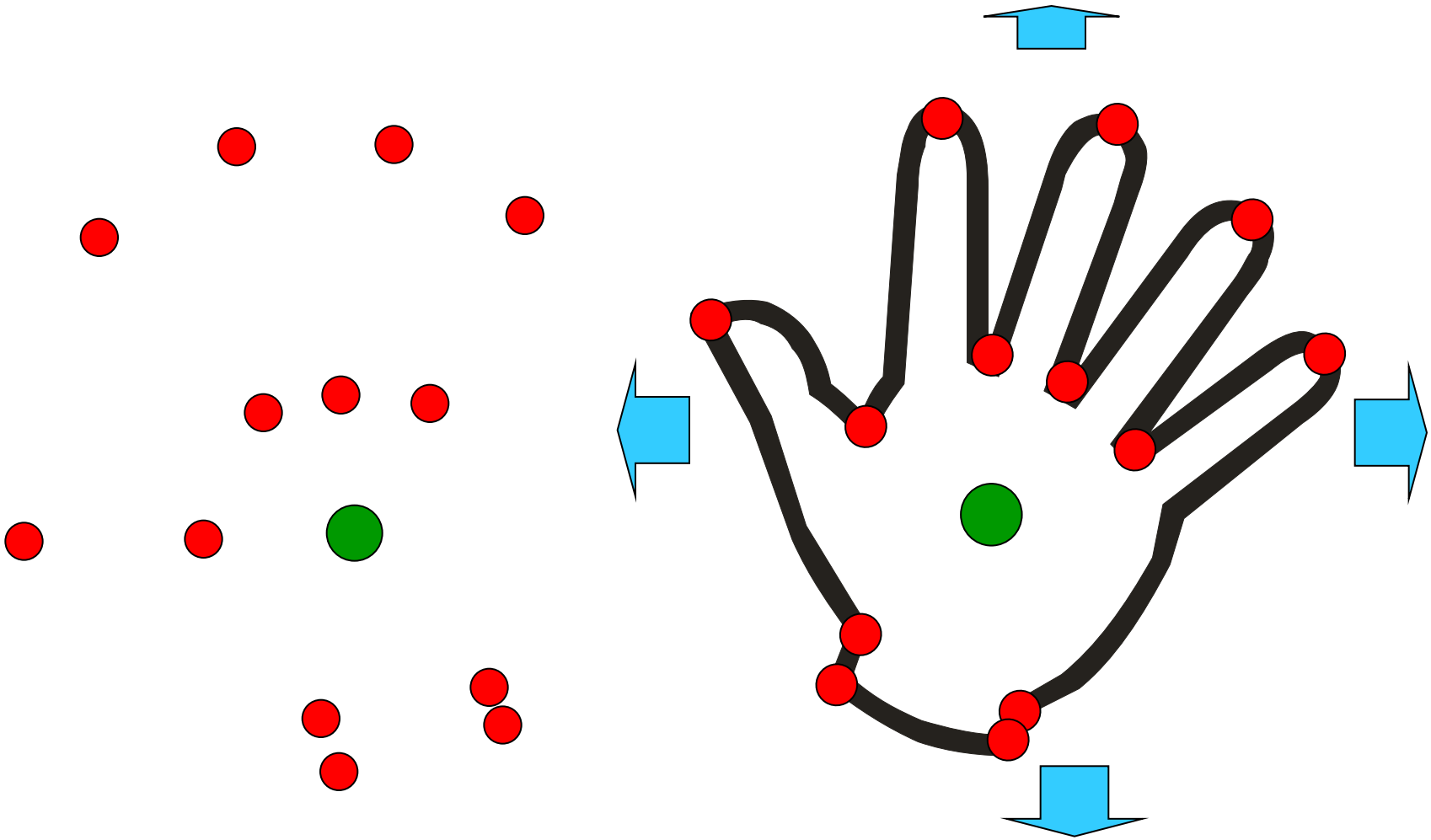




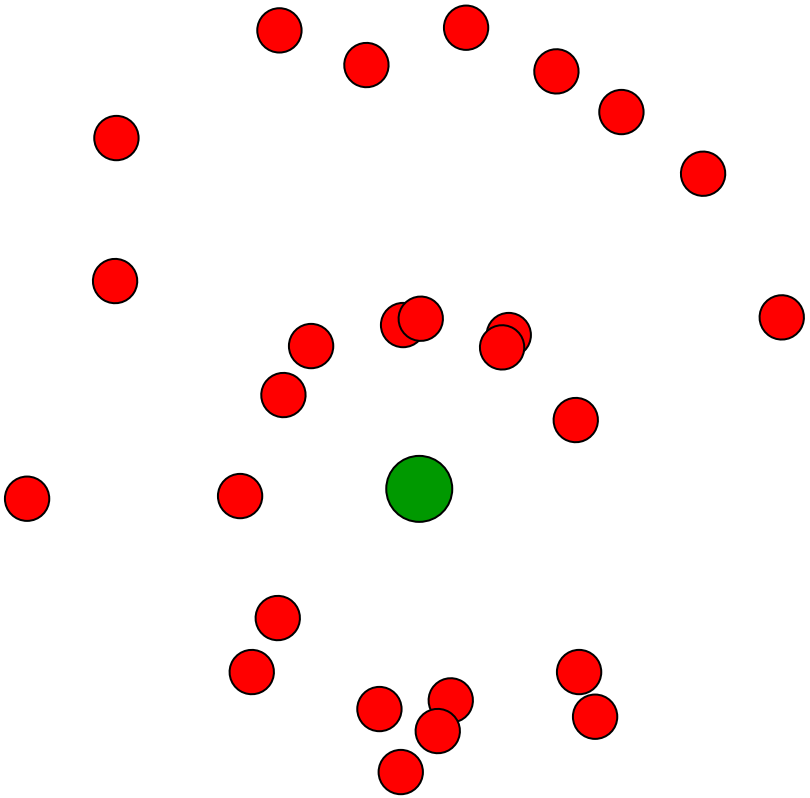
1) Size adjustment: unit centroid size



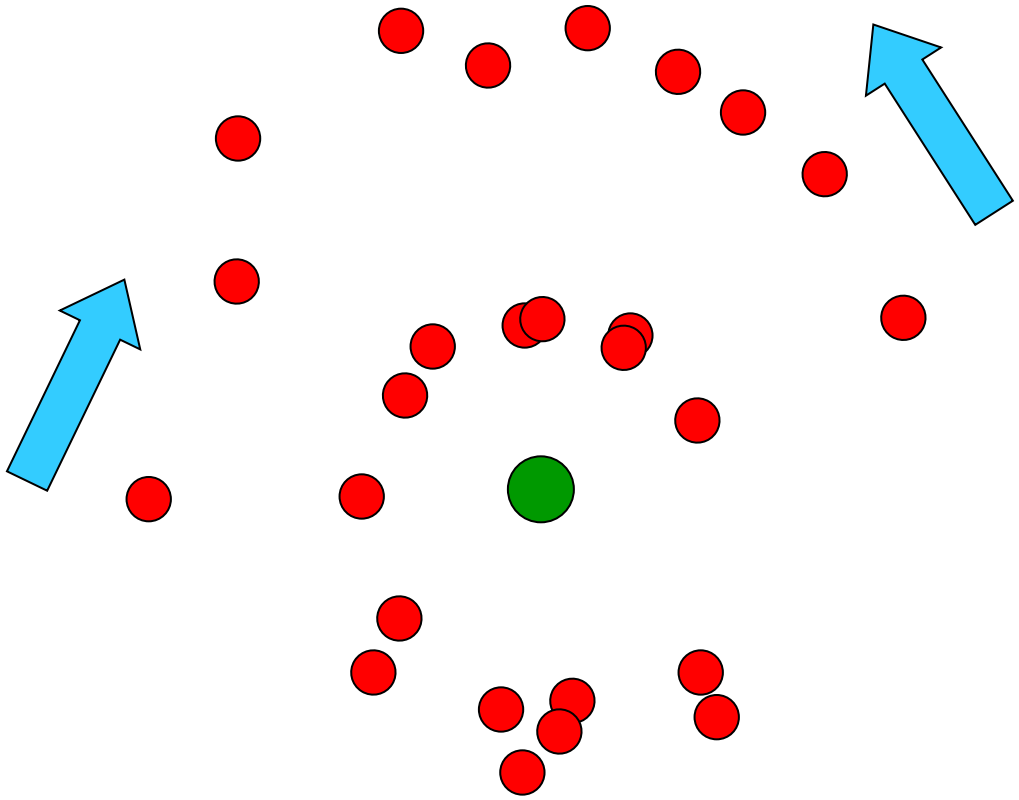
1) Size adjustment: unit centroid size



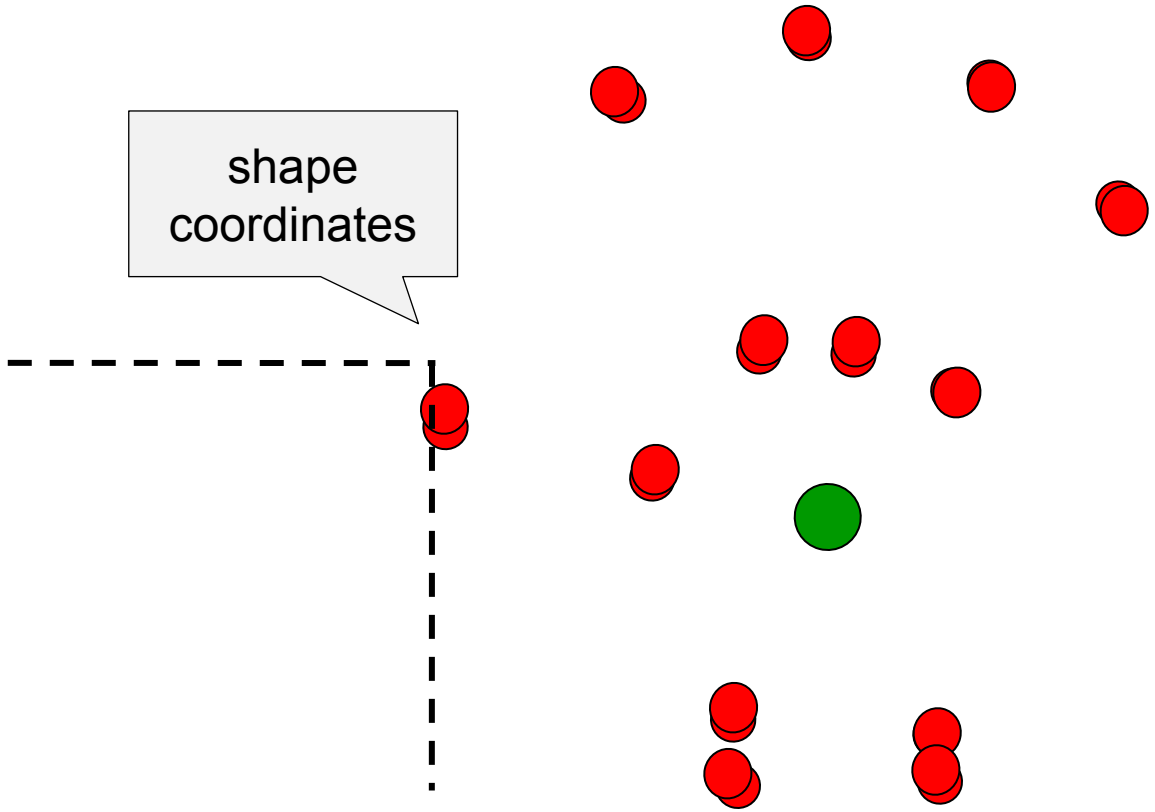
2) Transformation: same centroid position



3) Rotation: minimization of distances between homologous landmarks



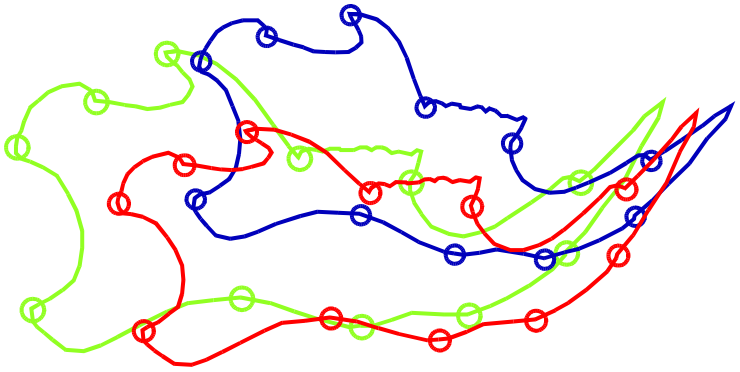
3) Rotation: minimization of distances between homologous landmarks



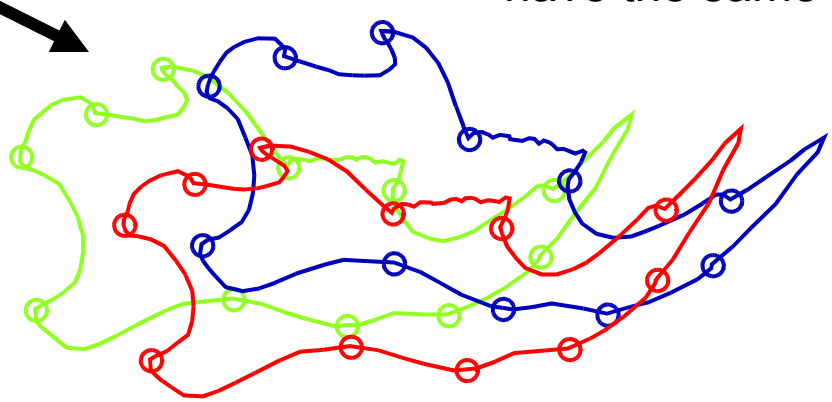
SHAPE SPACE: $n = pk - k - k(k-1)/2 - 1$

Extracting shape information: Procrustes superposition

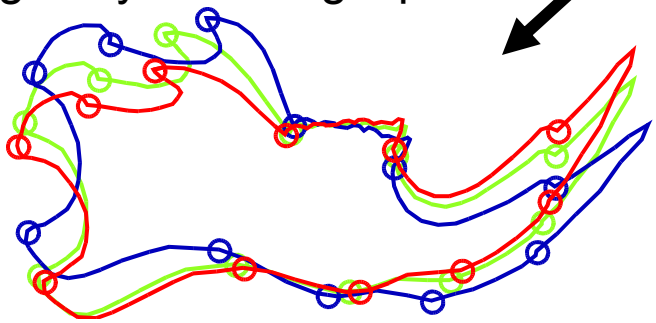
Original landmark configurations



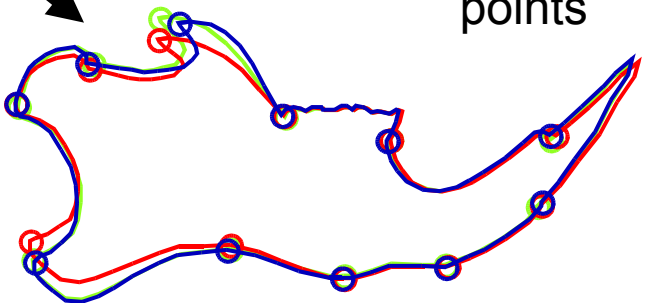
1. Change scale so that all configurations have the same size

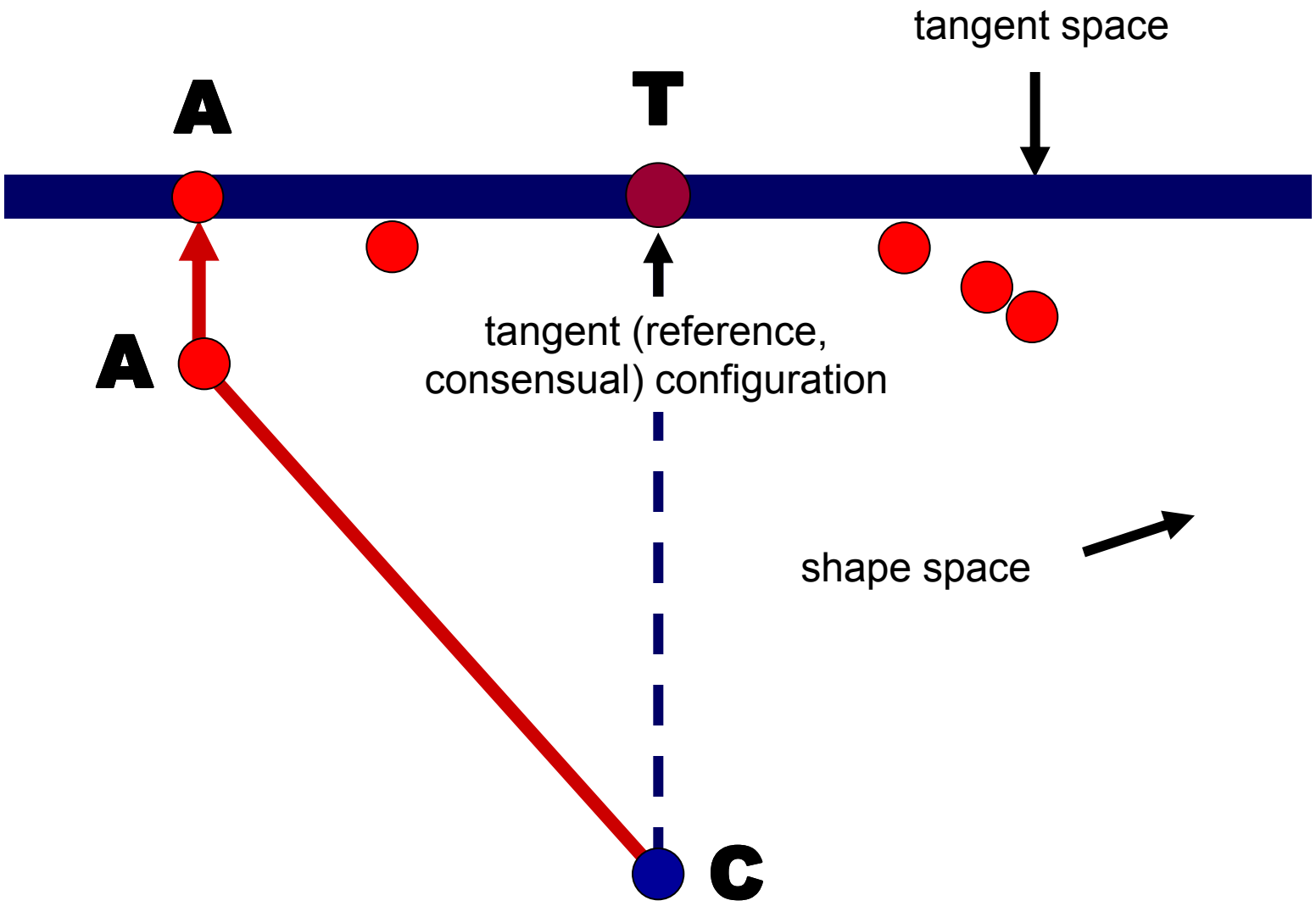


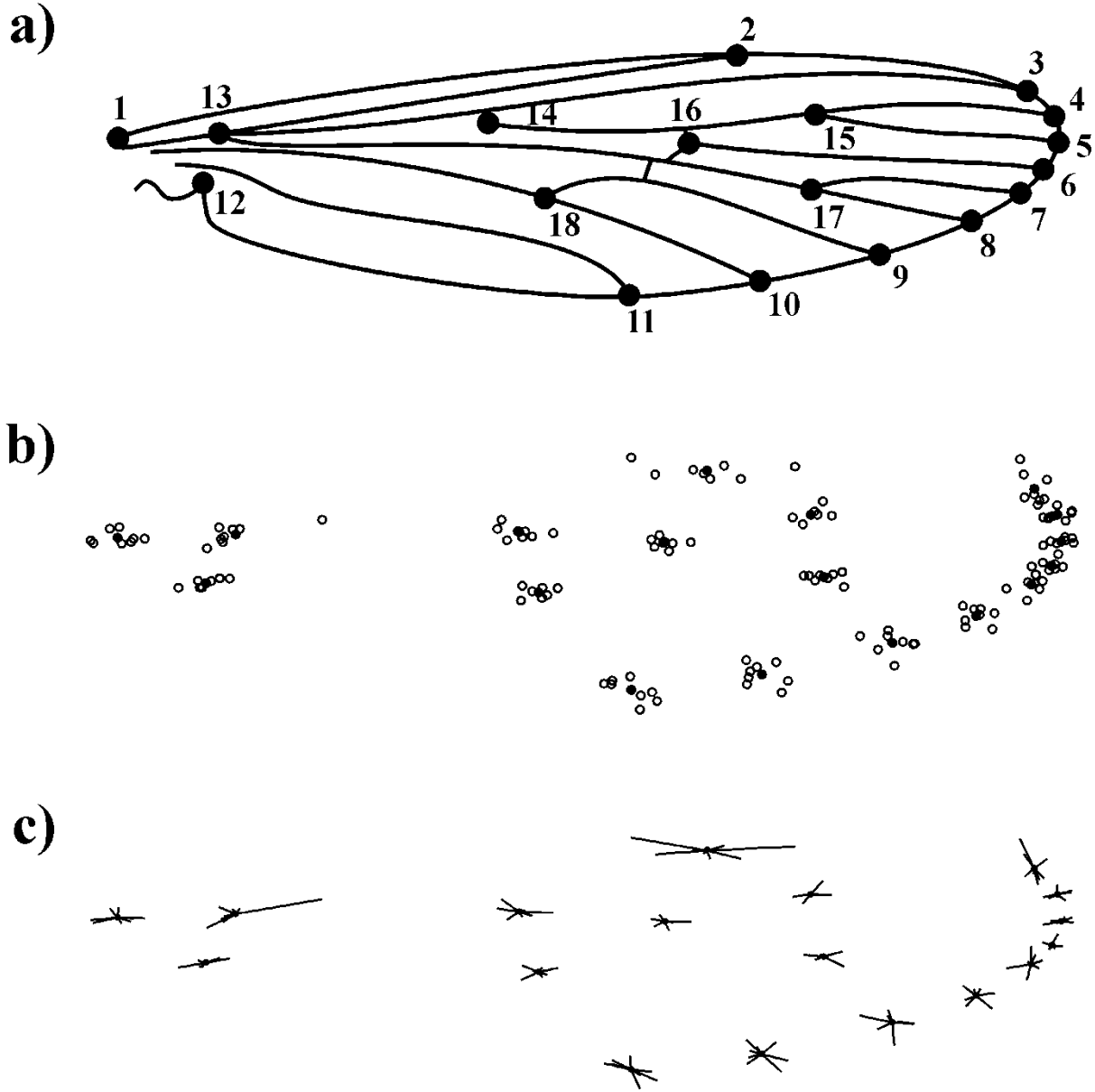
2. Superposition of the centers of gravity on a single point



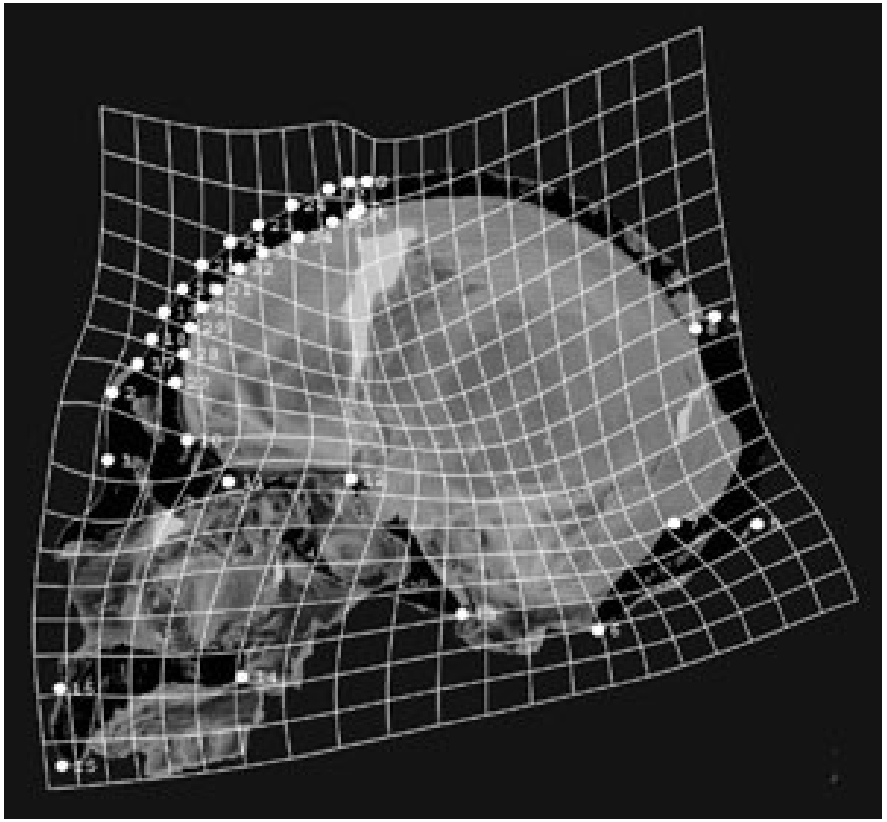
3. Rotation to minimize the dispersion of corresponding points



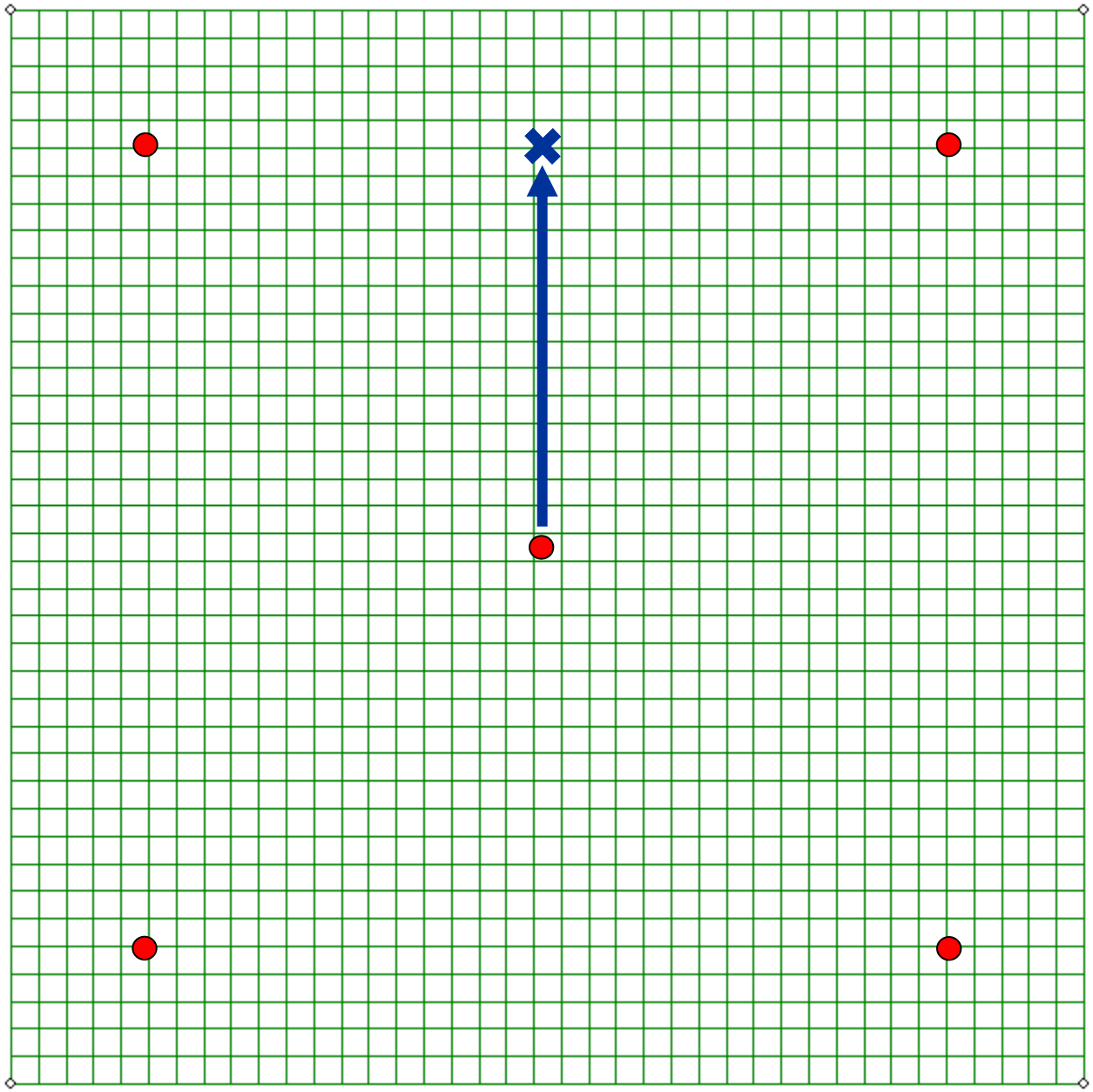

tangent space
A
T
A
tangent (reference,
consensual) configuration
shape space
C

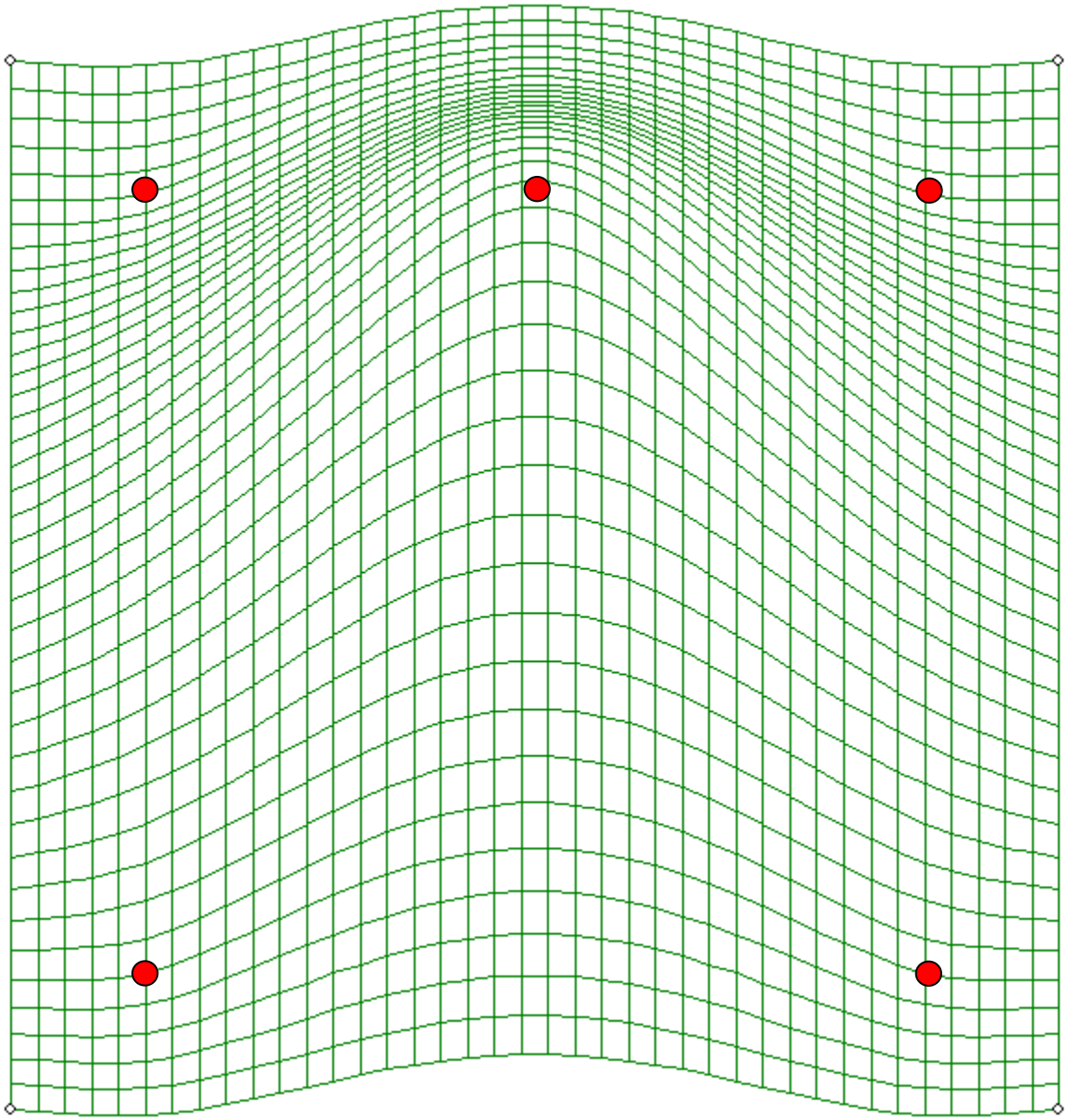

a)
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
b)
c)

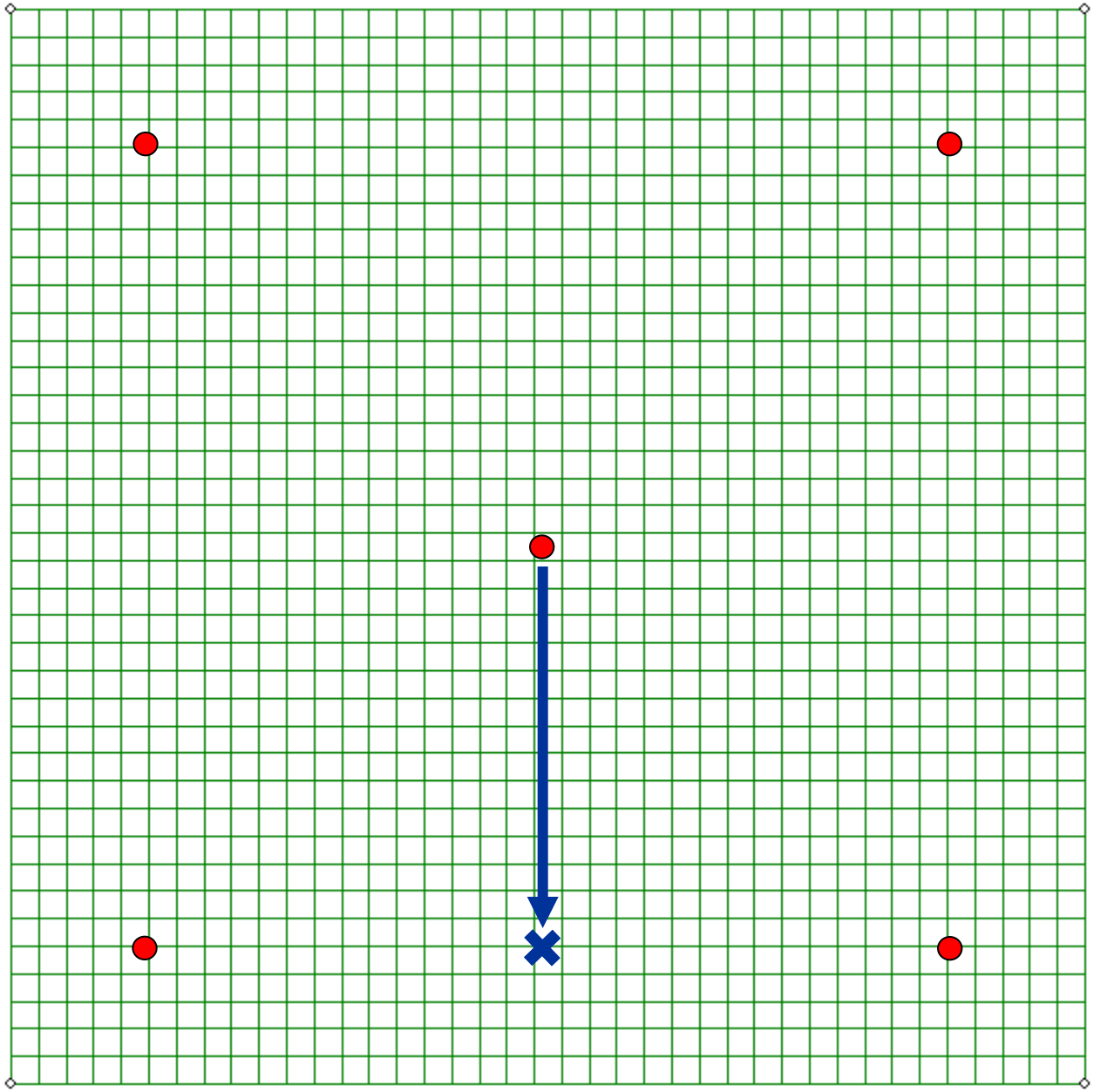
Thin-Plate Spline (TPS)

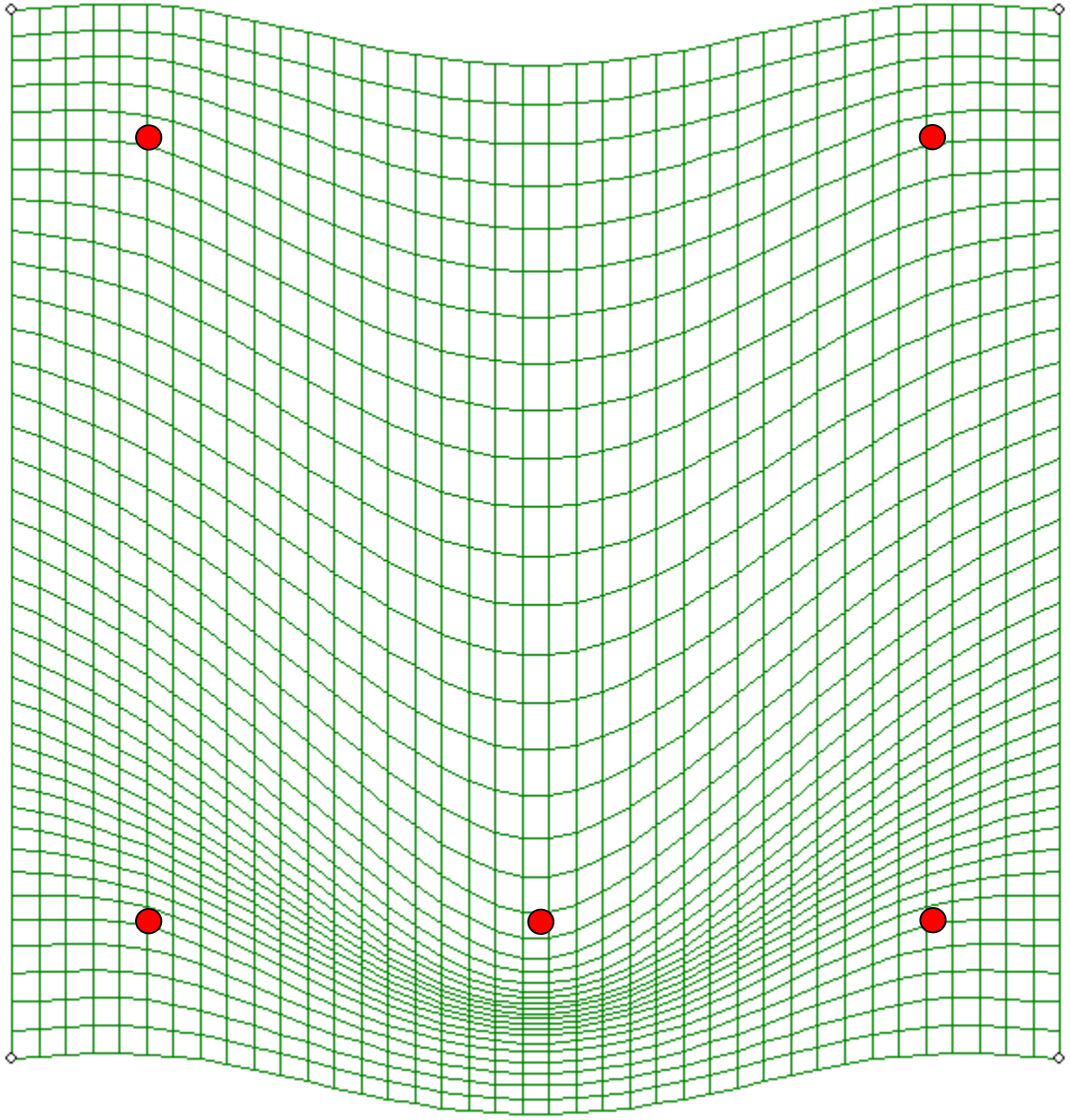


metaphor of infinitely large and infinitely thin metal plate









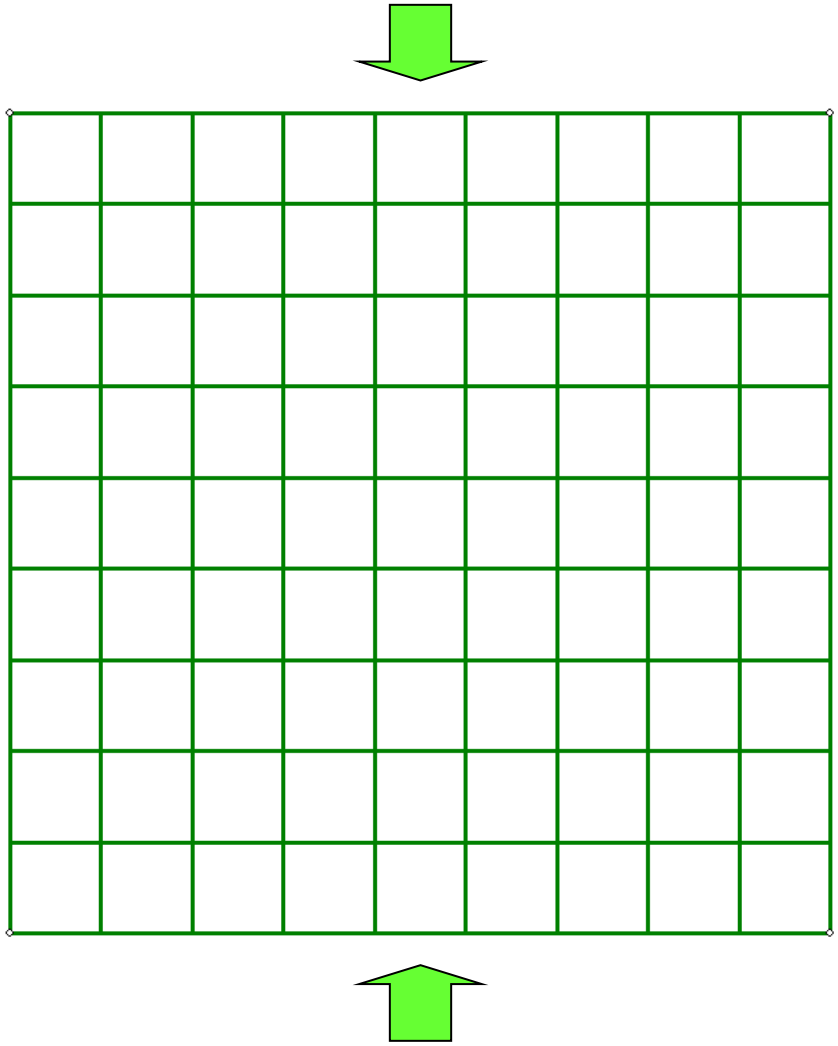
Thin-Plate Spline (TPS)

energy necessary for plate deformation = bending energy

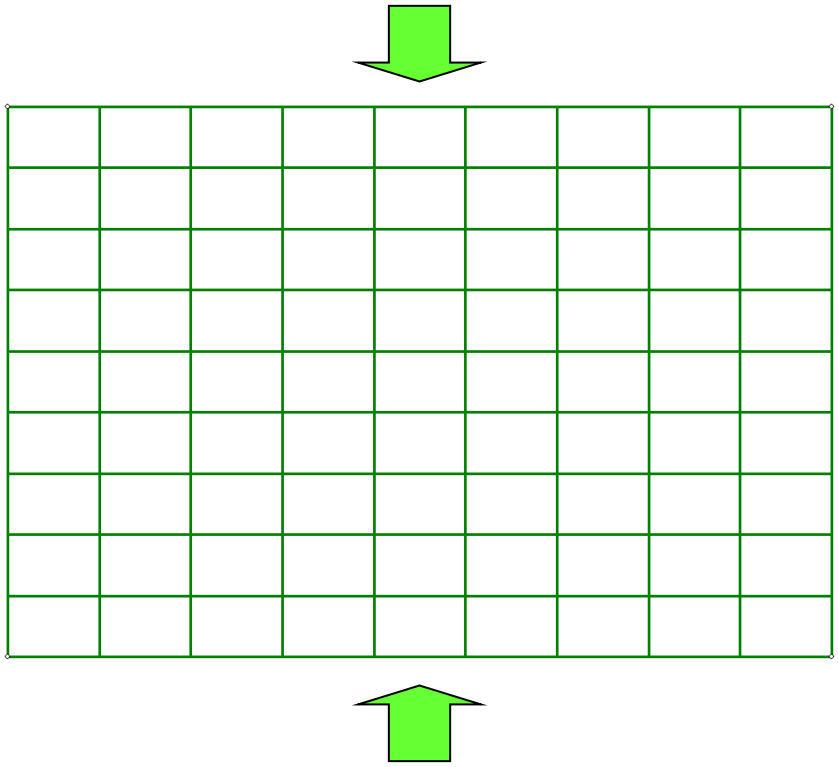
differentiating between afine and nonafine shape changes

projection of latent roots of bending energy into components
= partial warps
partial warp 0 ~ afine component

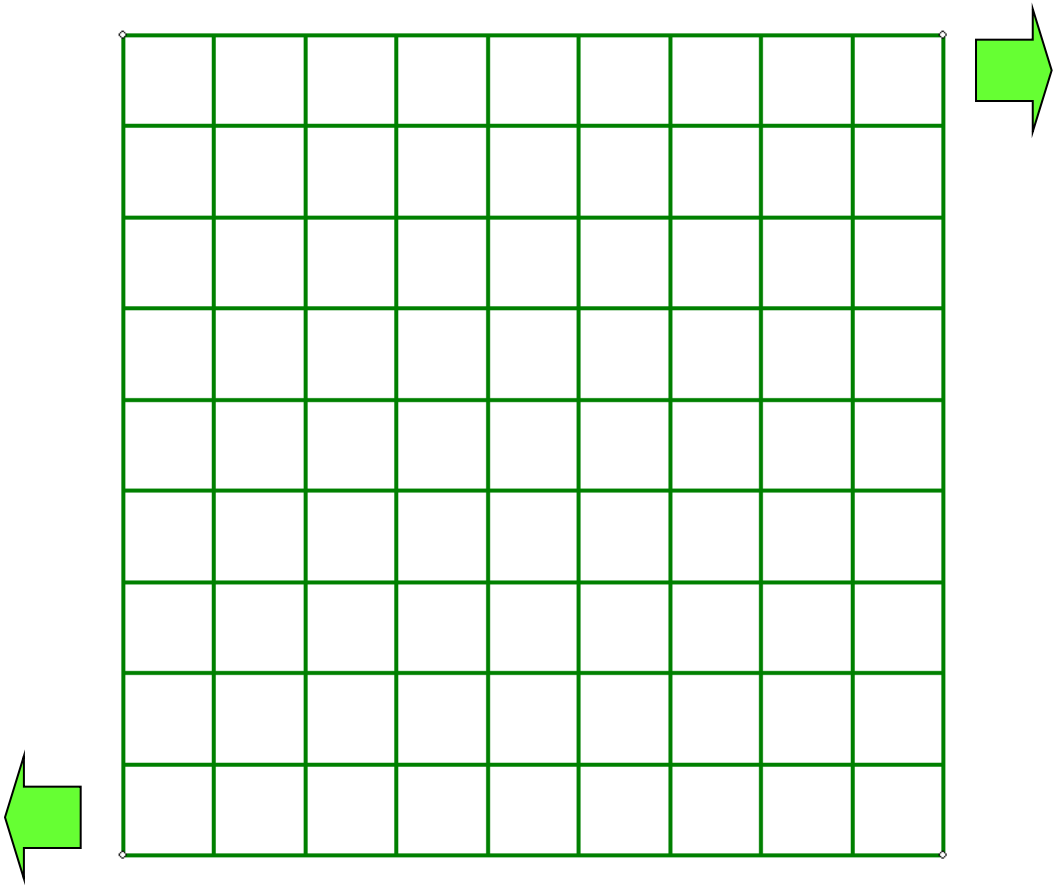
Afine shape change



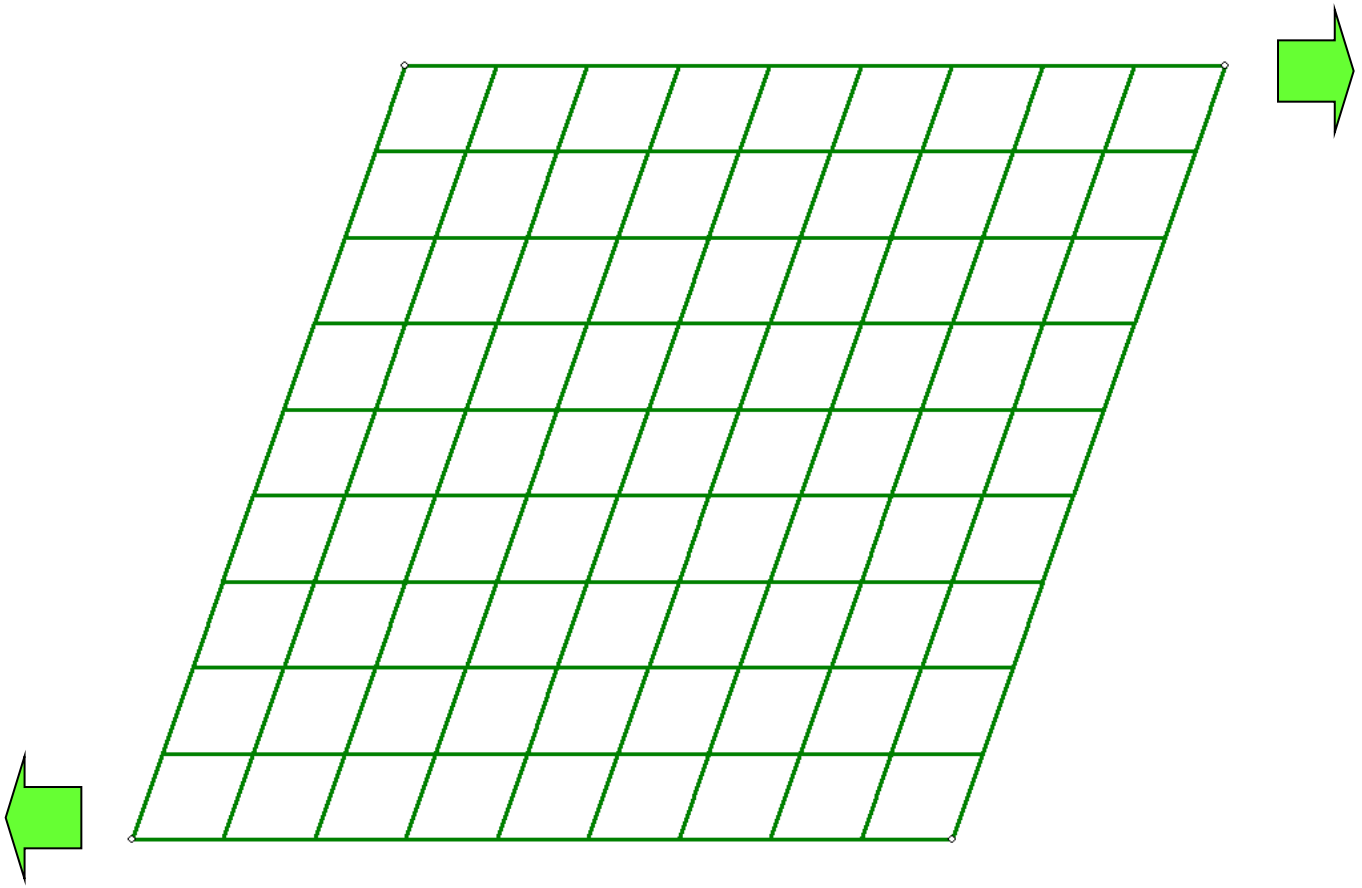
Afine shape change



Afine shape change

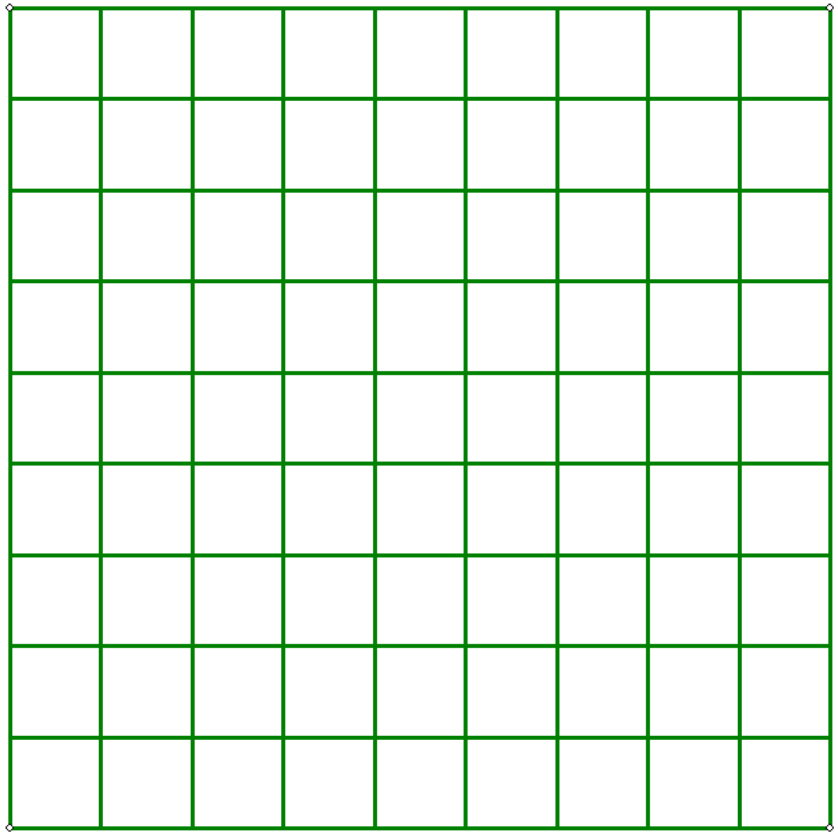


Afine shape change

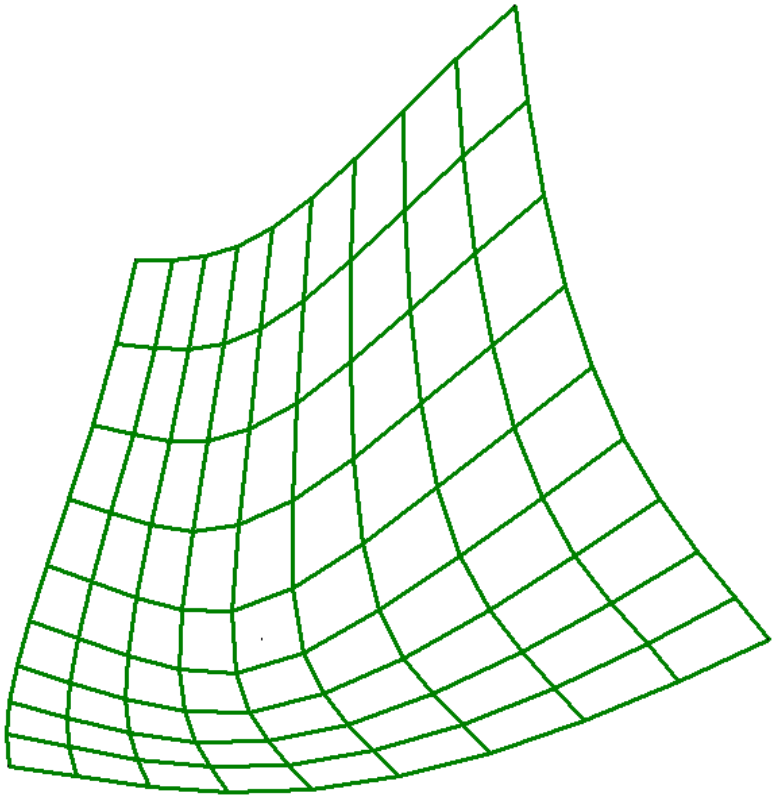


parallel lines are still parallel

Nonafine shape change

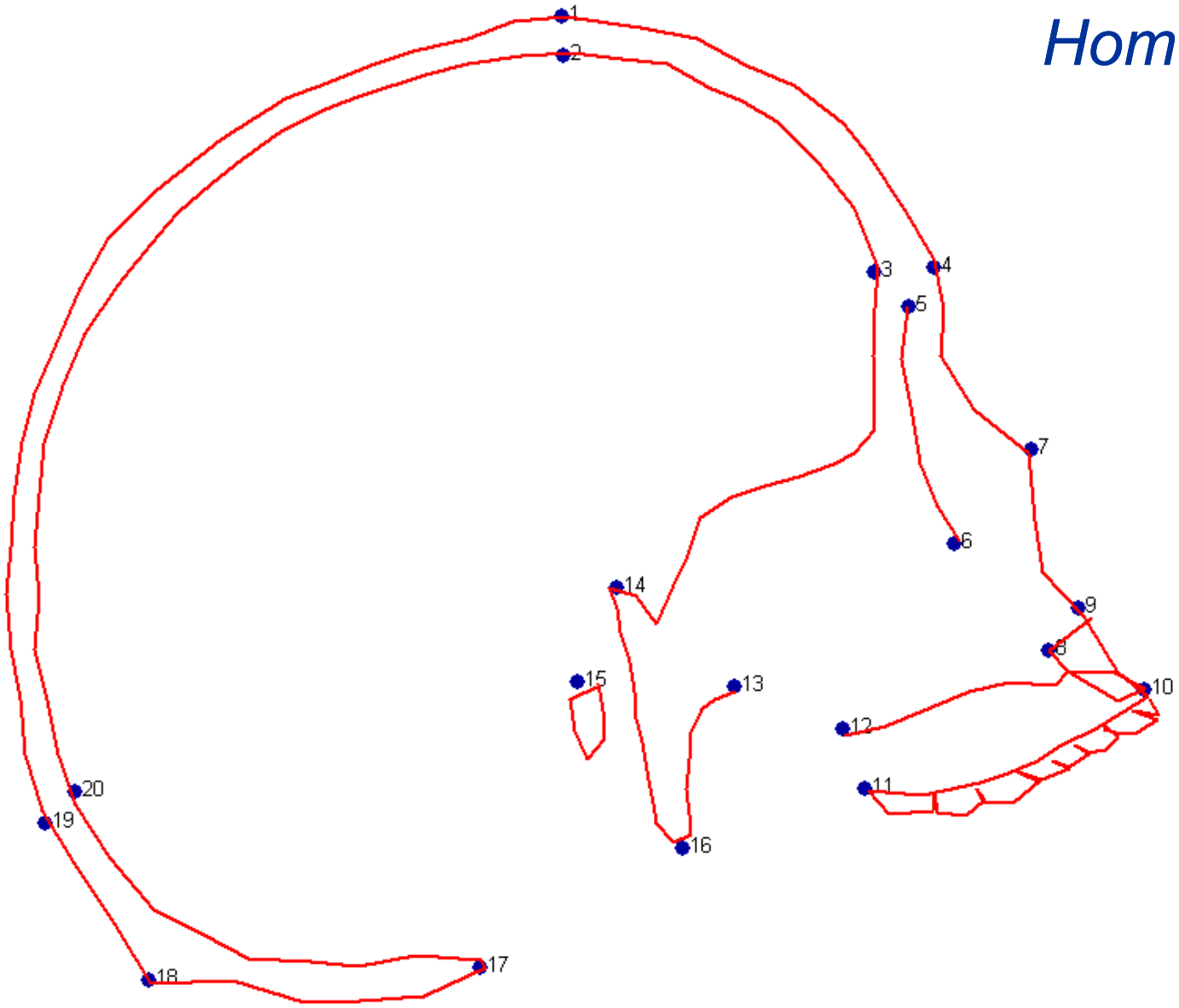


Nonafine shape change

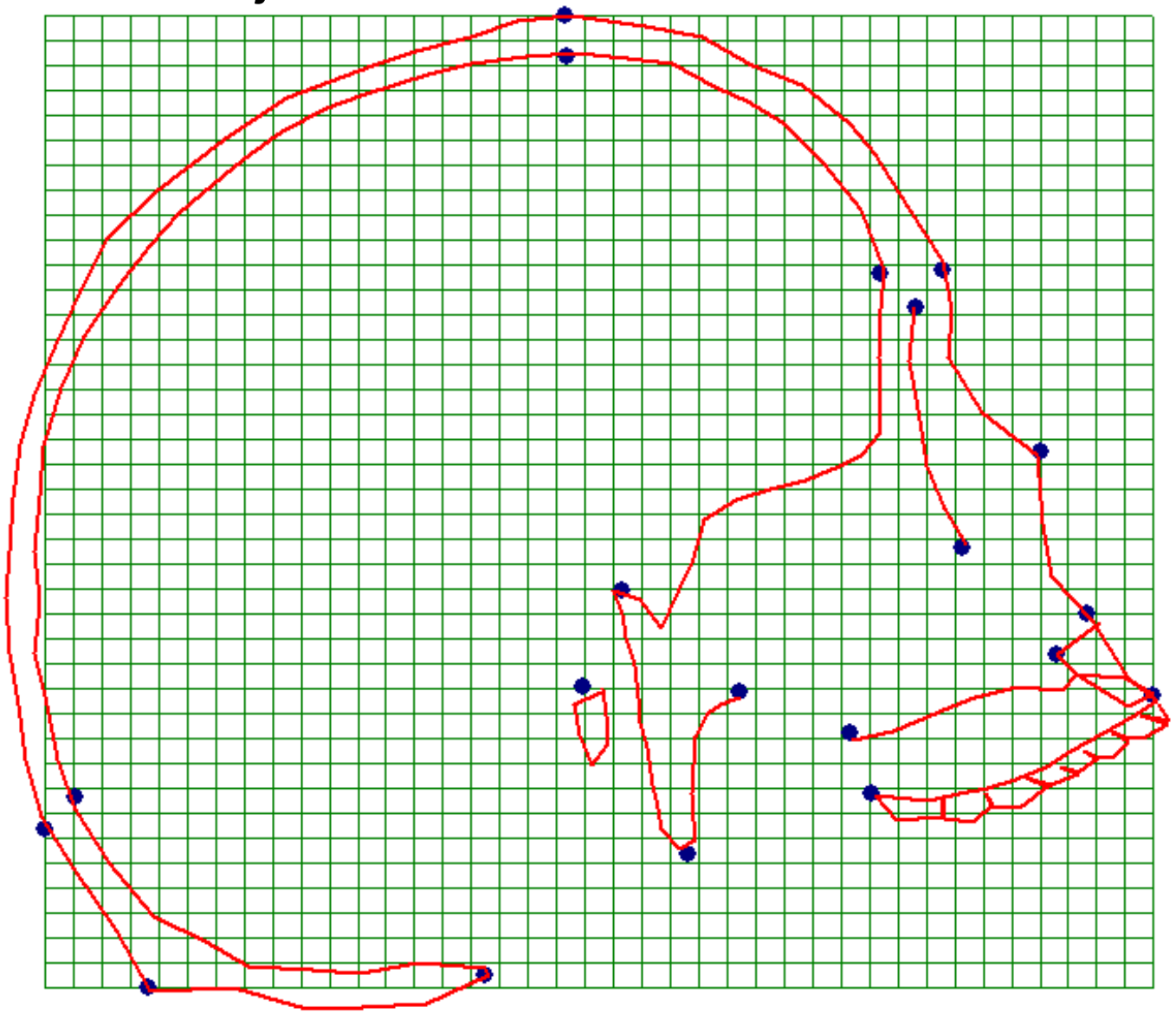


parallel lines are not parallel any more

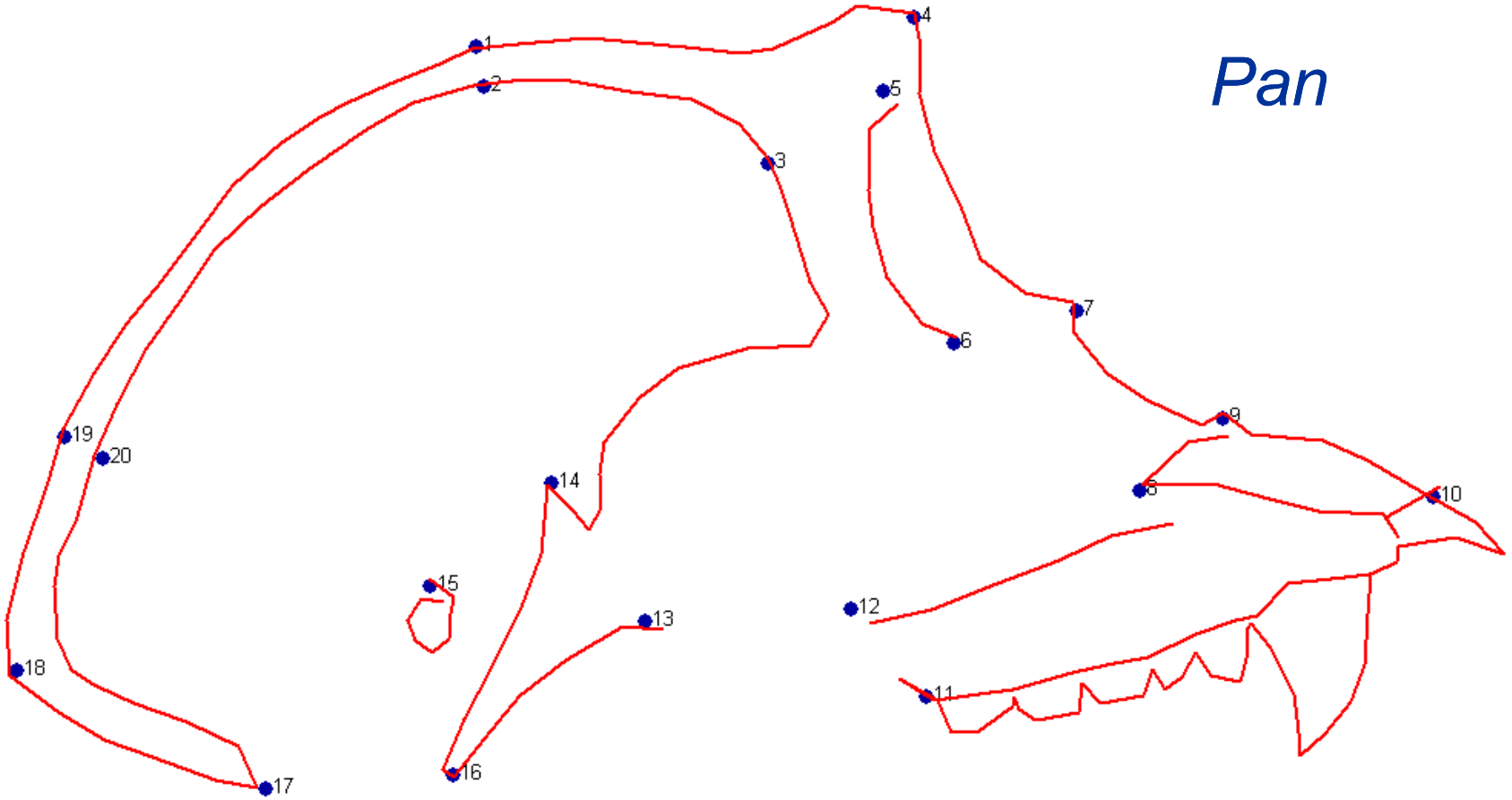
Homo



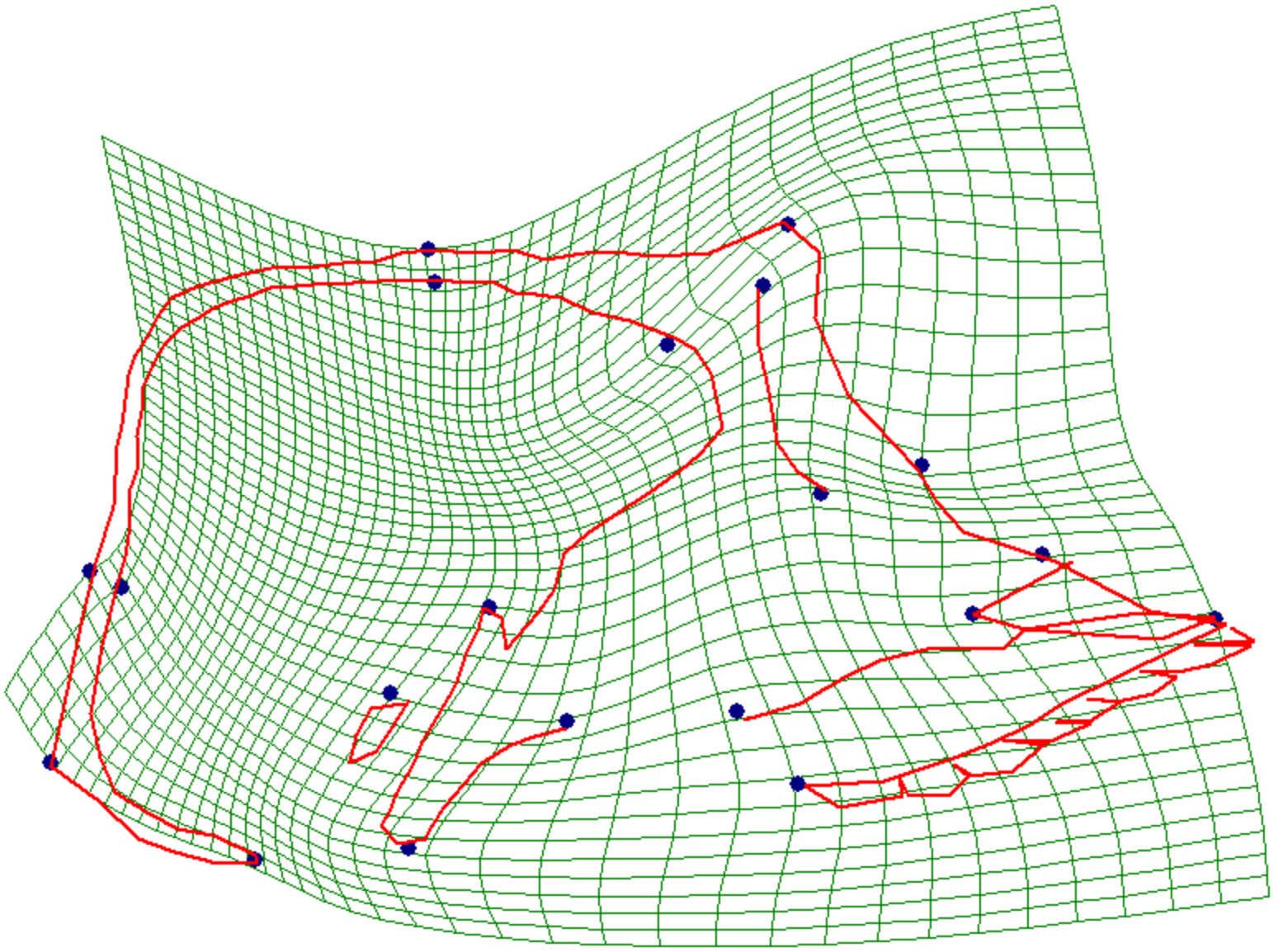
reference object

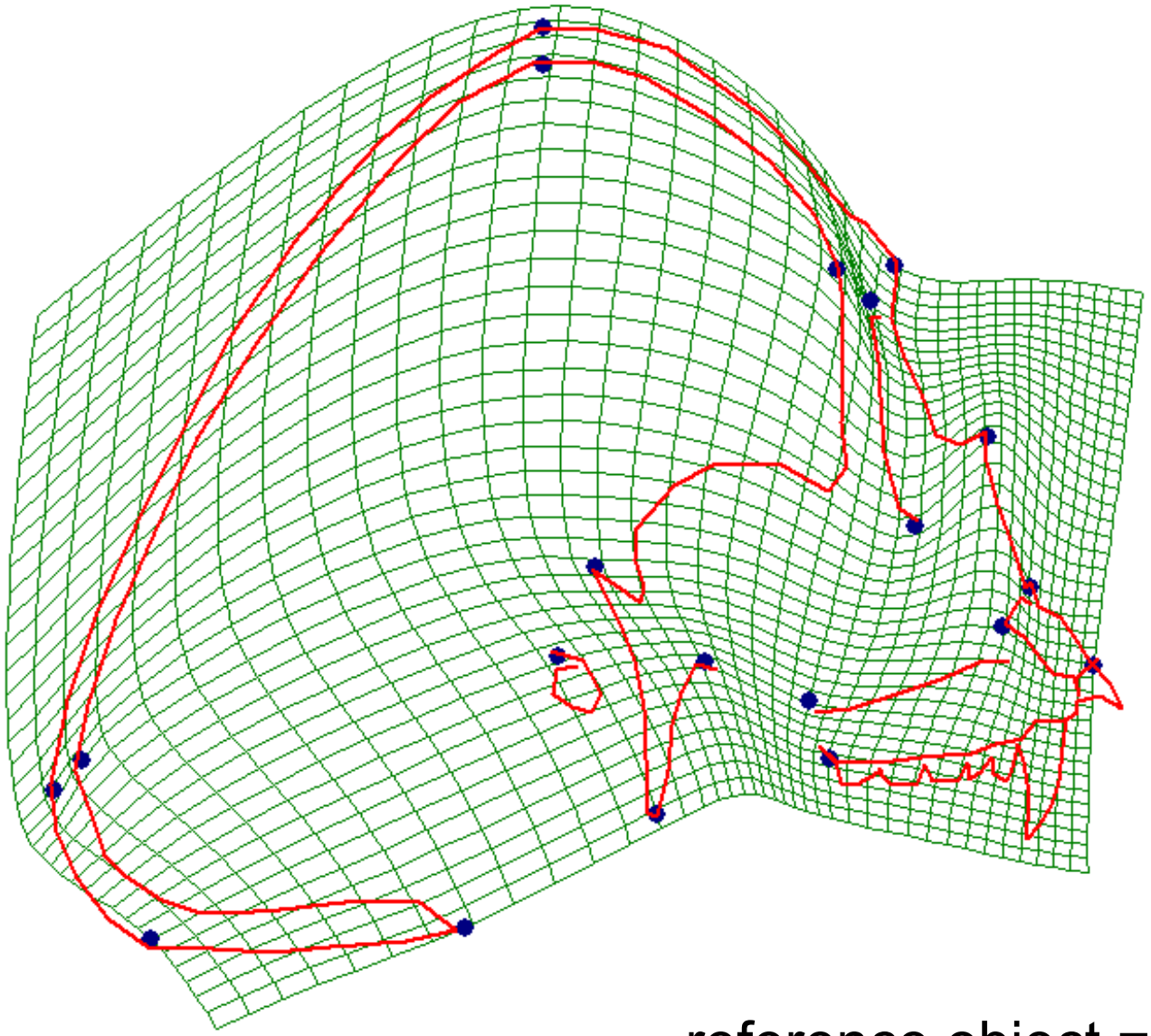


Pan



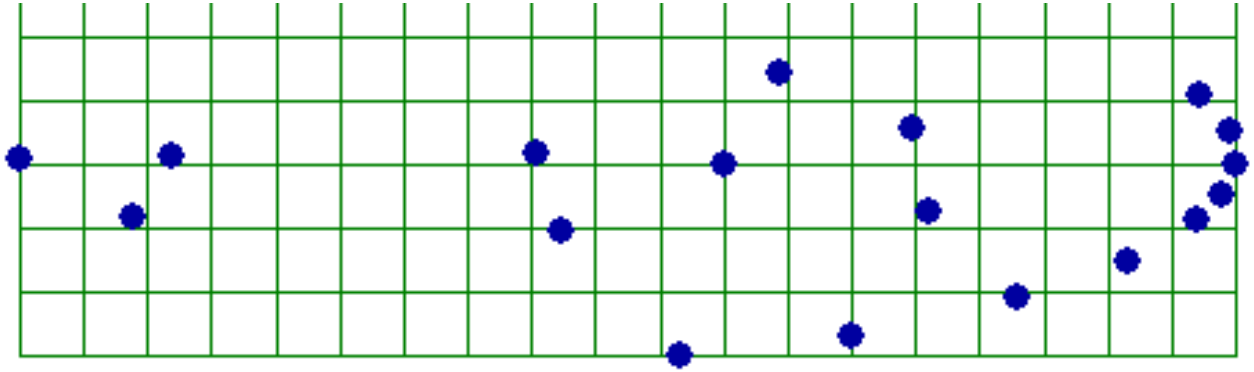
reference object = *Homo*

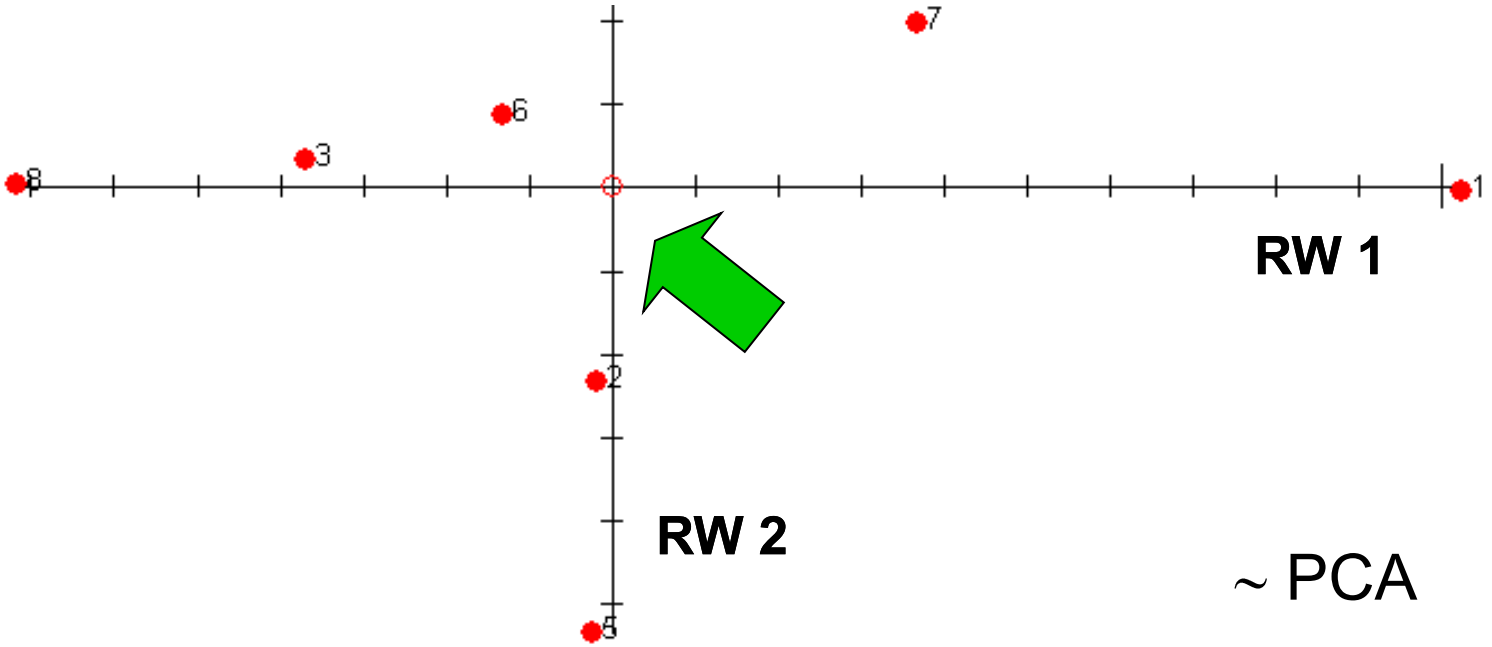




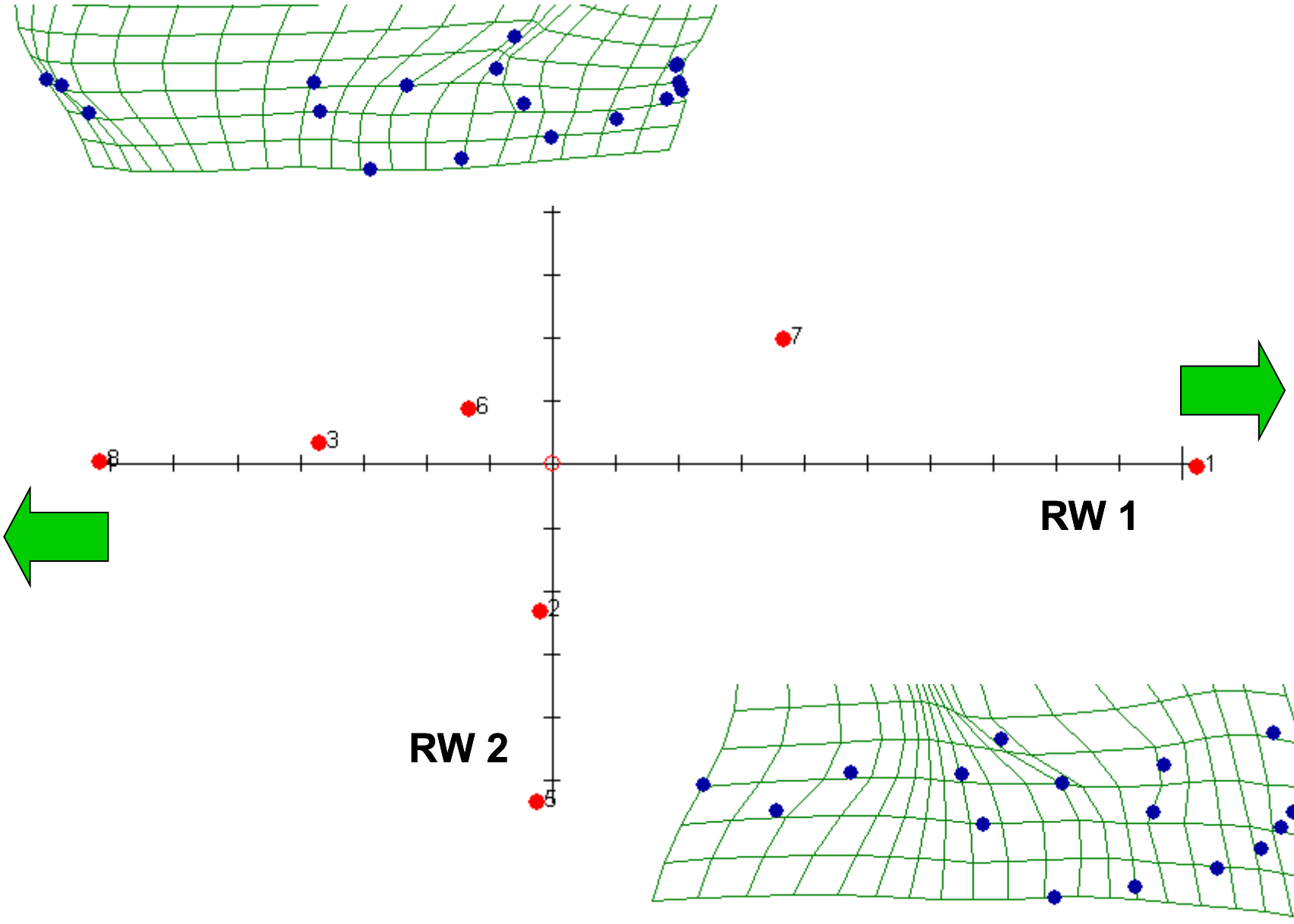
reference object = *Pan*

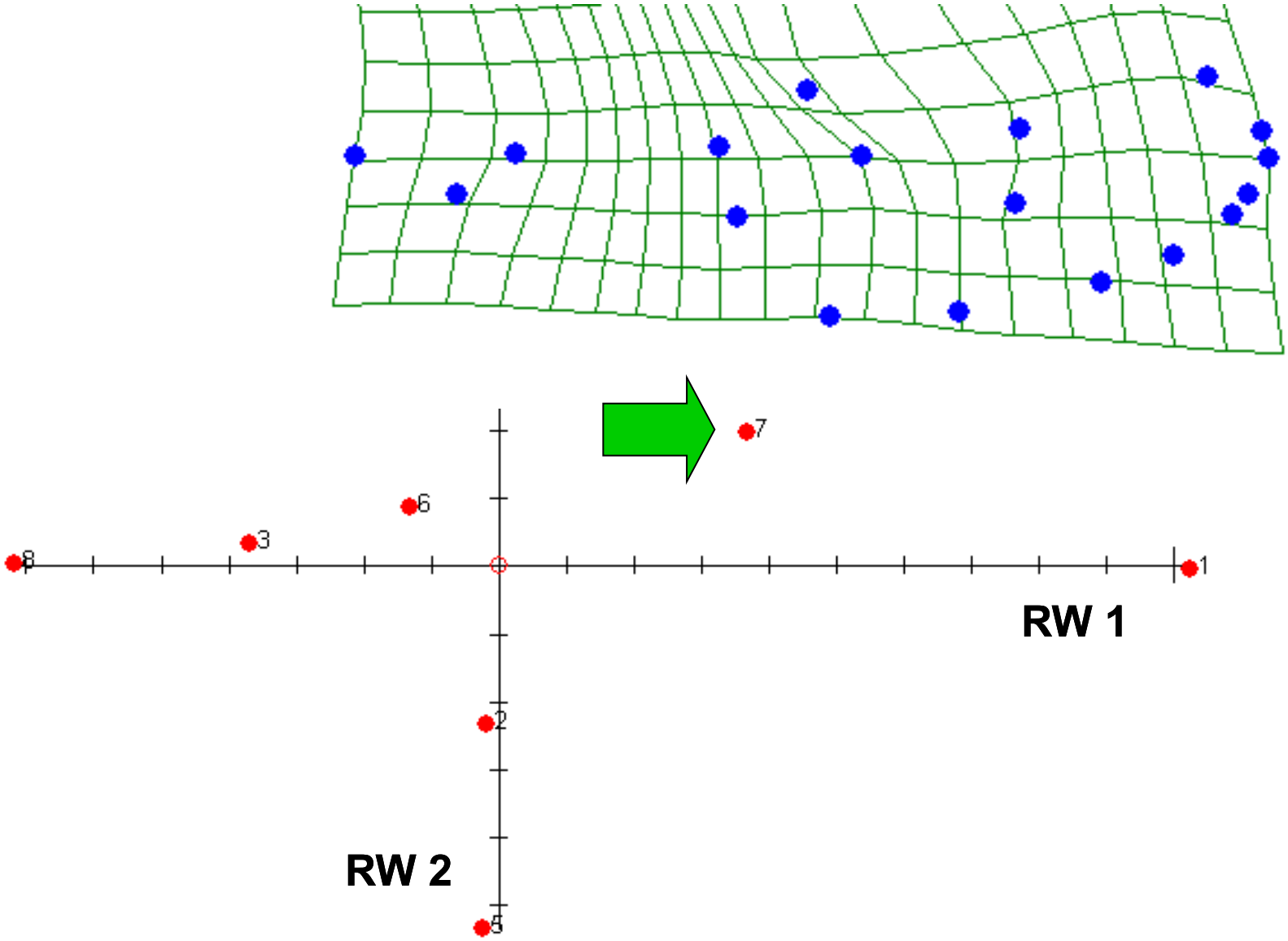
>2 samples: Thin-Plate Spline Relative Warps (TPSRW)





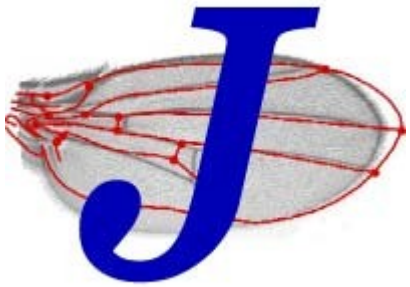
~ PCA


7
6
3
8
1
RW 1
2
RW 2
5


7
6
3
8
1
RW 1
2
RW 2
5

Software:

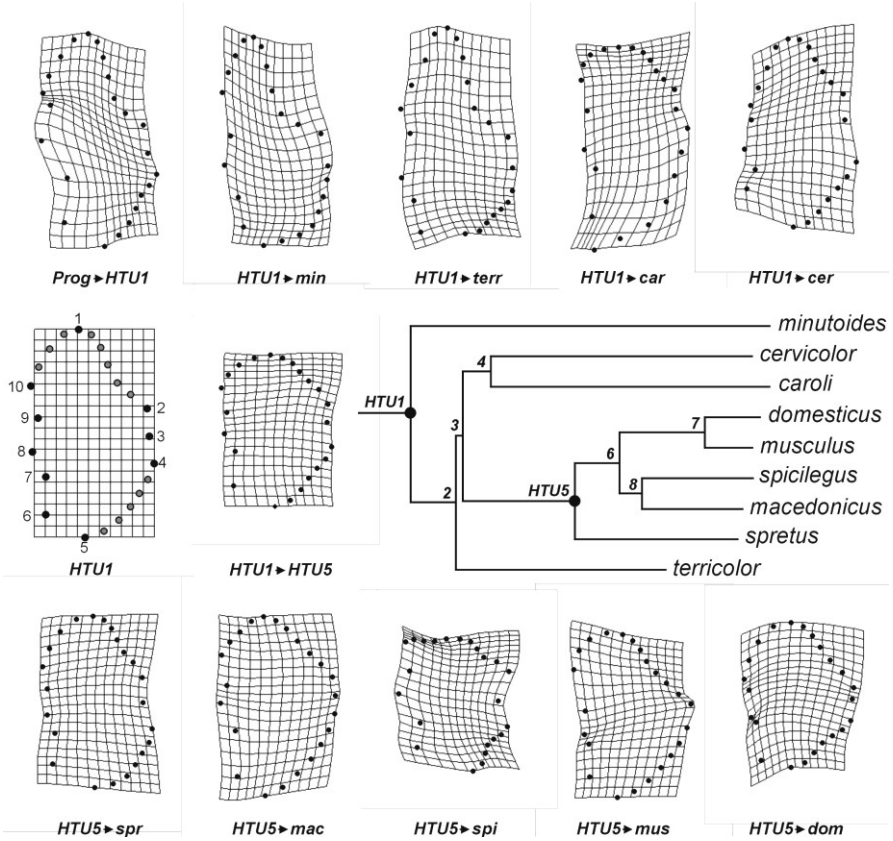
tpsDig: úprava obrázků, digitalizace bodů, měření rozměrů
tpsSplin: TPS
tpsRelw: TPS Relative Warps
tpsRegr: regrese na nezávislou proměnnou
tpsPLS: metoda parciálních nejmenších čtverců
(např. korelace 2 sad bodů)
tpsSuper: deformace obrázků („unwarping“)
tpsTree: analýza tvarových změn podél větví
fylogenetického



MorphoJ

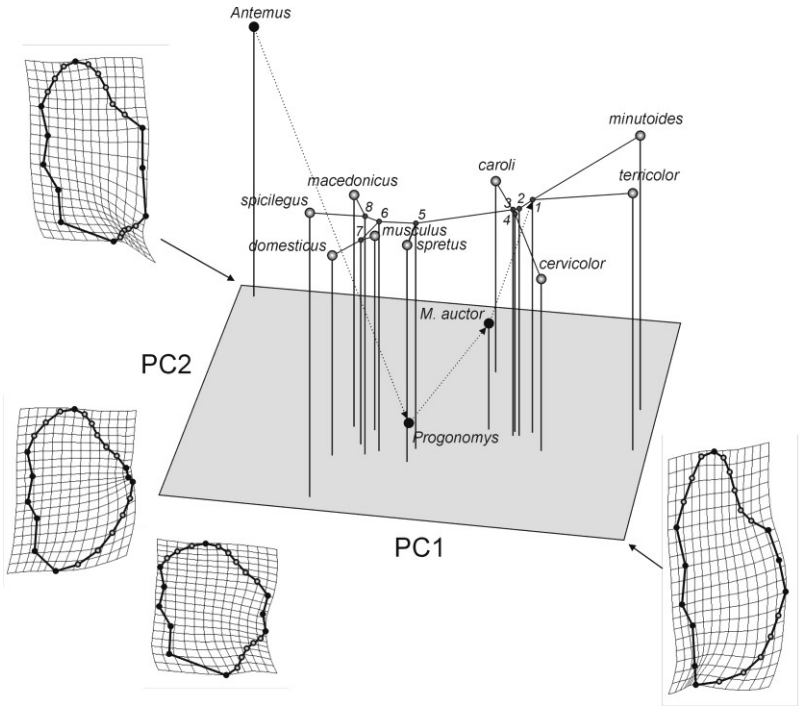
http://life.bio.sunysb.edu/morph/

Morfometrics and phylogenesis

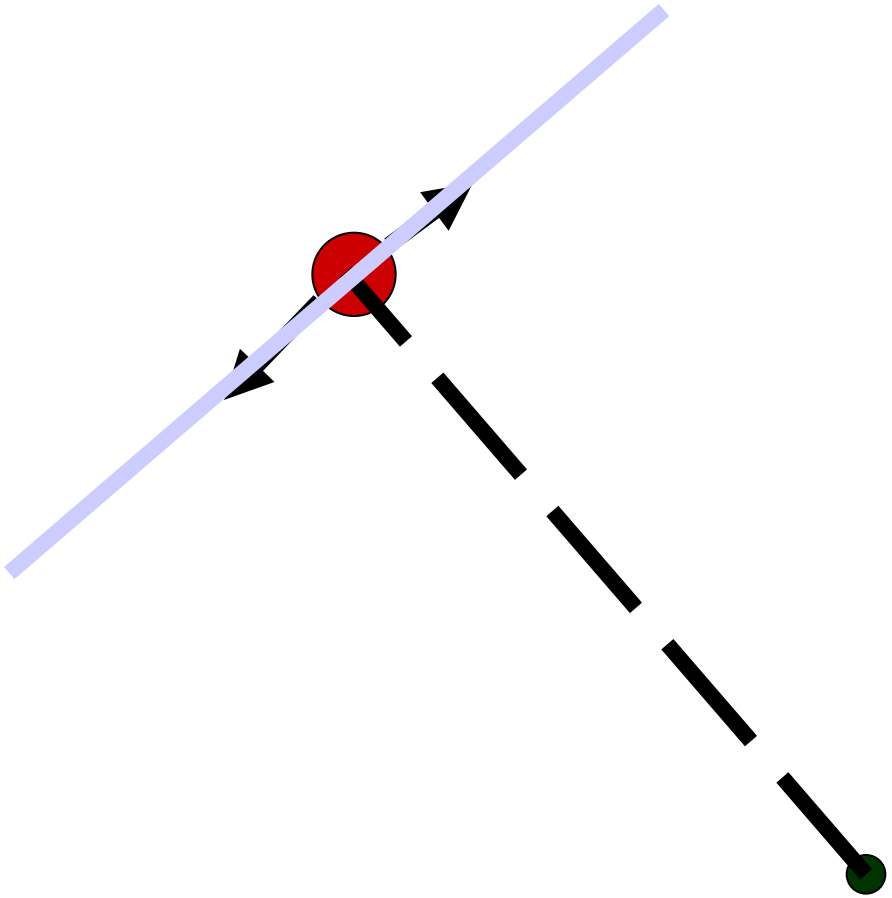


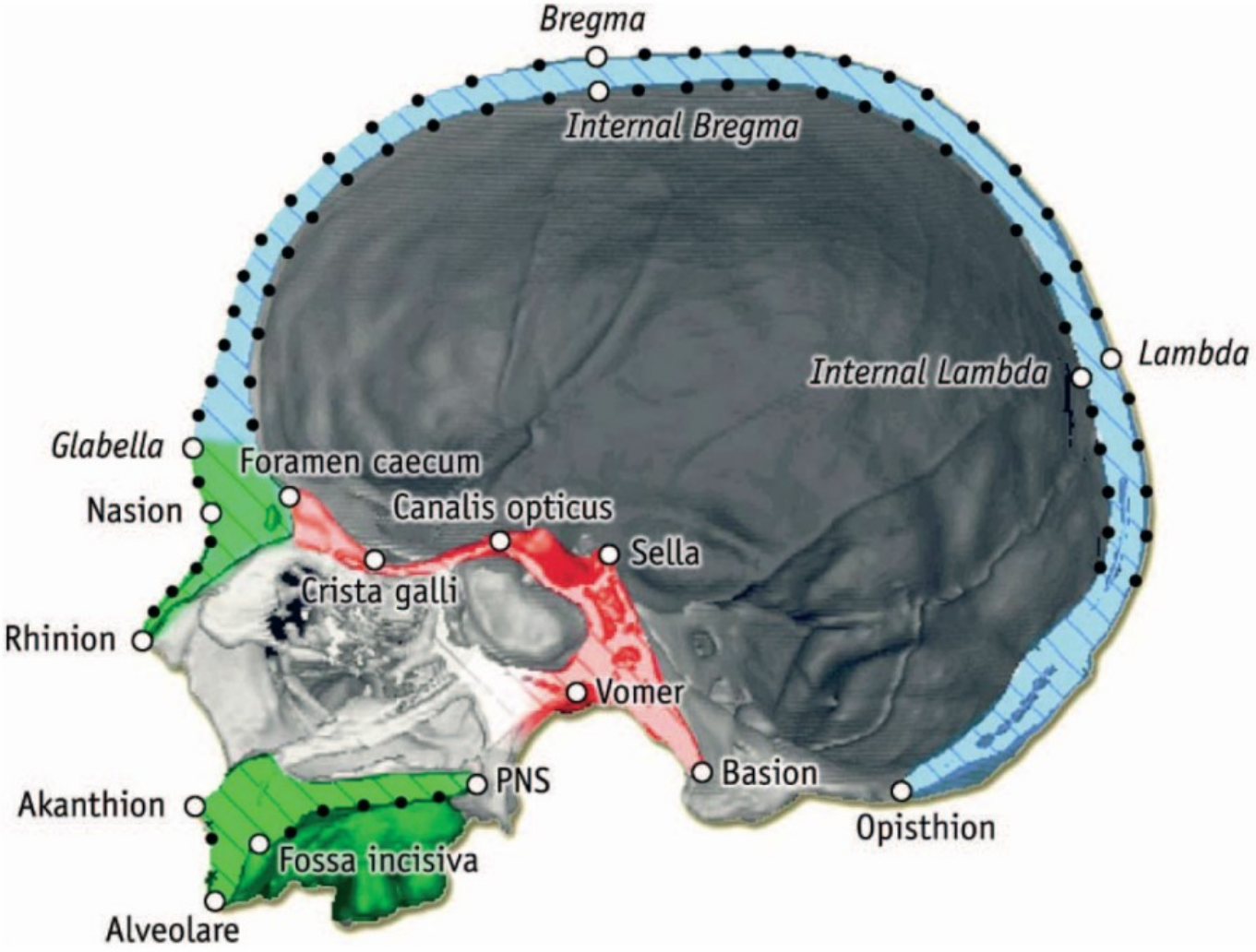
tpsTree

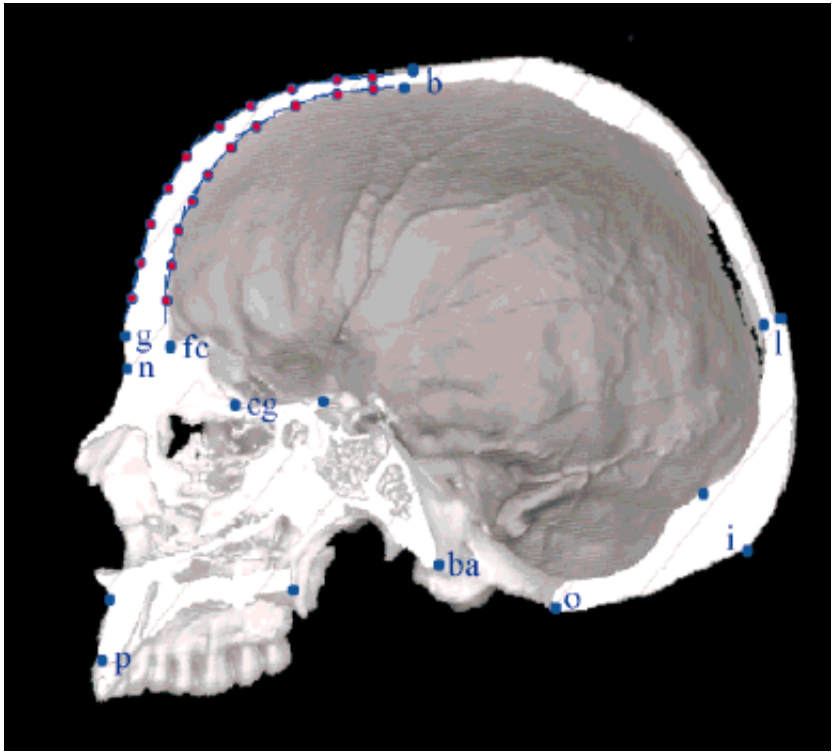
mapping of shape changes

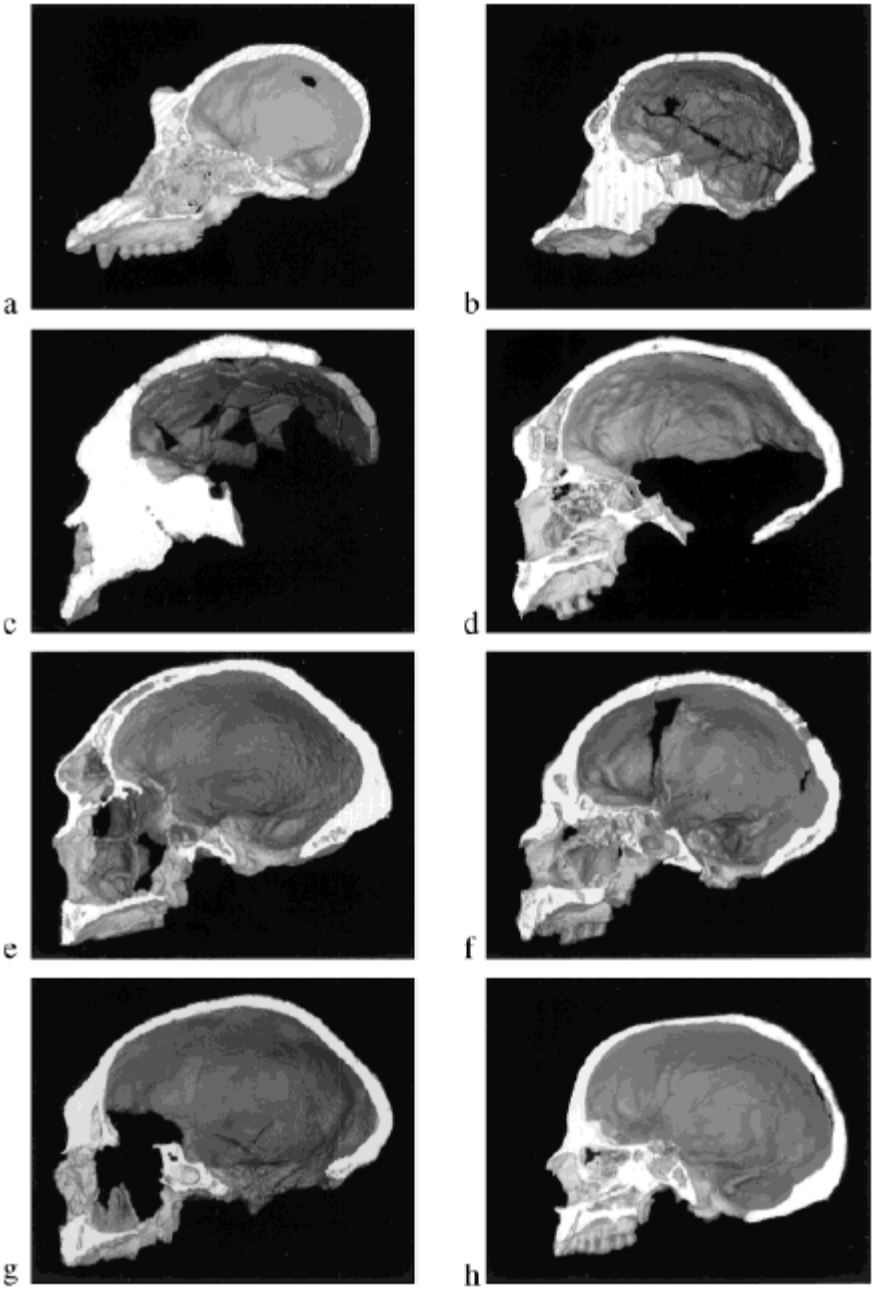


Landmark based methods without landmarks – „sliding semilandmarks“

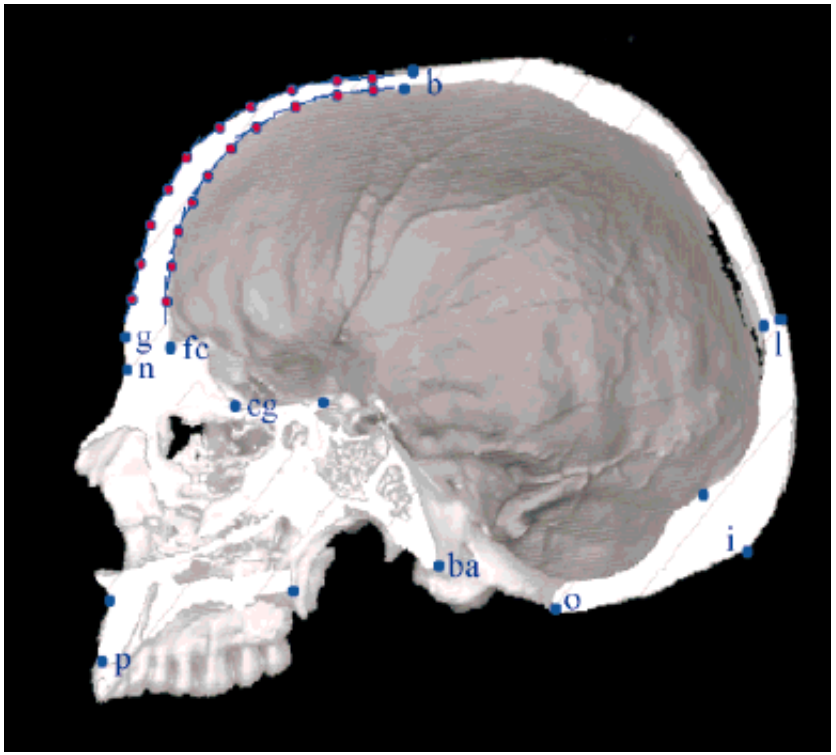


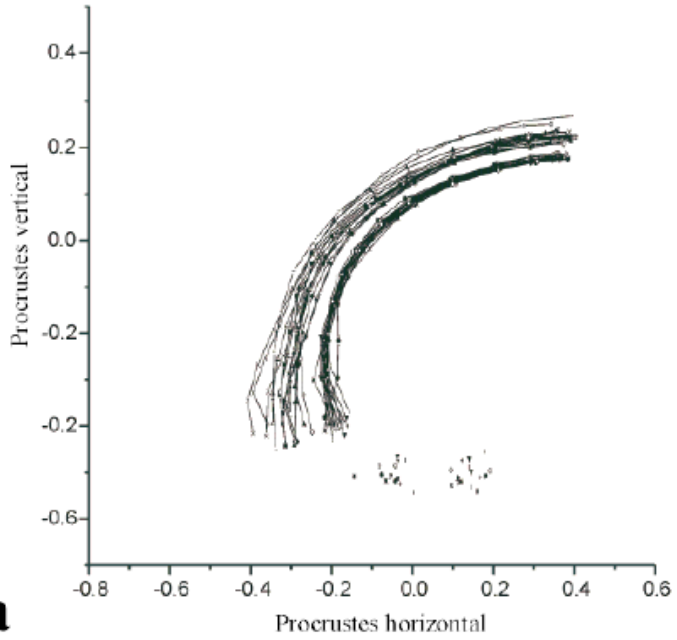

Bregma
Internal Bregma
Lambda
Internal Lambda
Glabella
Foramen caecum
Nasion
Canalis opticus
Sella
Crista galli
Rhinion
Vomer
Basion
PNS
Akanthion
Opisthion
Fossa incisiva
Alveolare



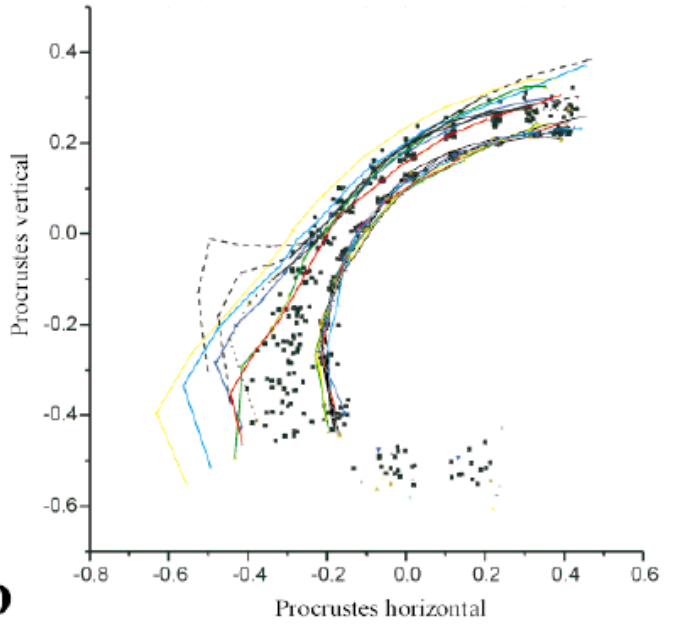


Bookstein et al.,
Anat. Record (1999)


b
g
n
fc
l
cg
i
ba
o
p

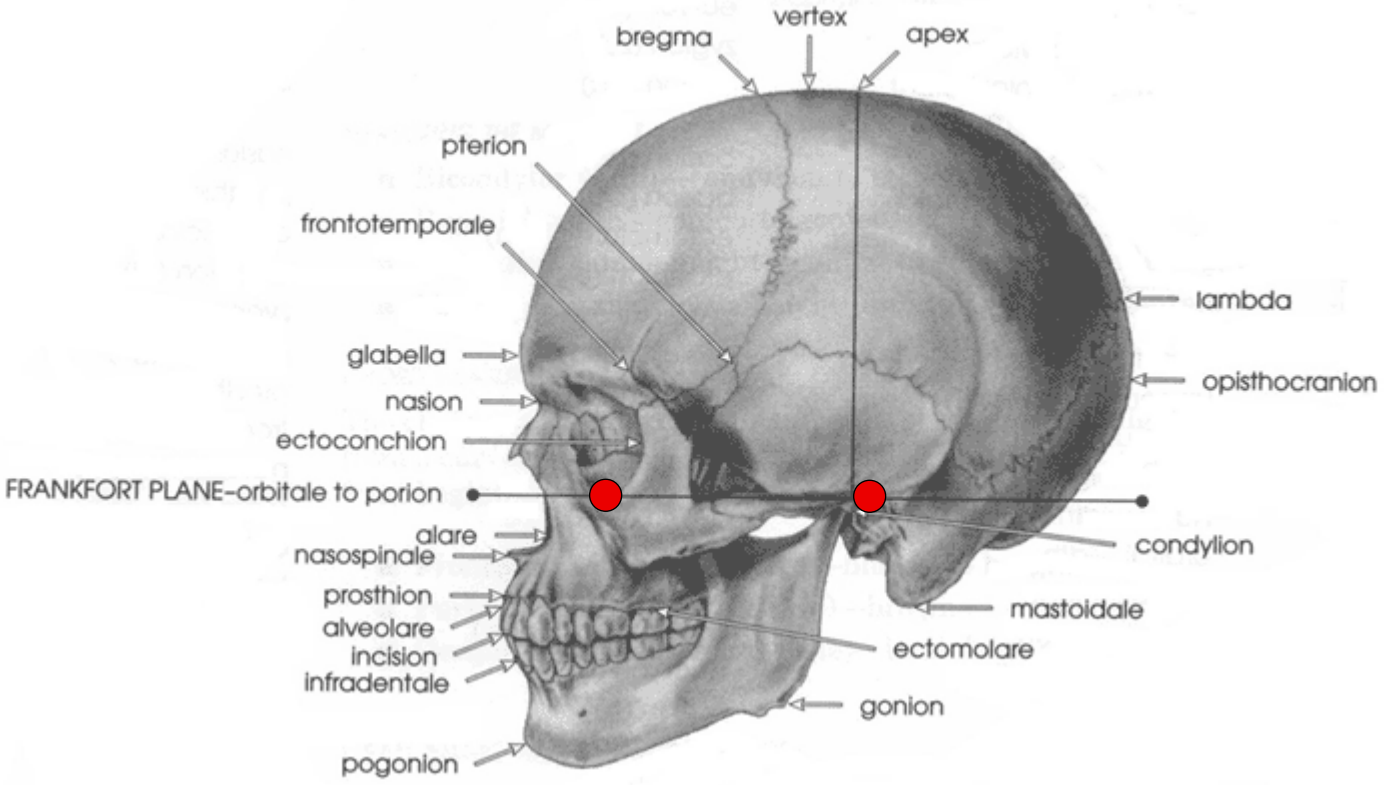

Procrustes vertical
0.4
0.2
0.0
-0.2
-0.2
-0.6
-0.8
-0.6
-0.4
-0.2
0.0
0.2
0.4
0.6
Procrustes horizontal

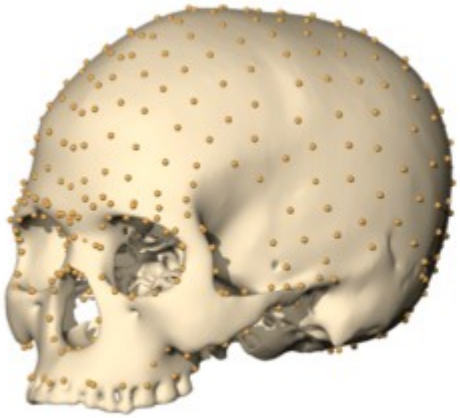
a

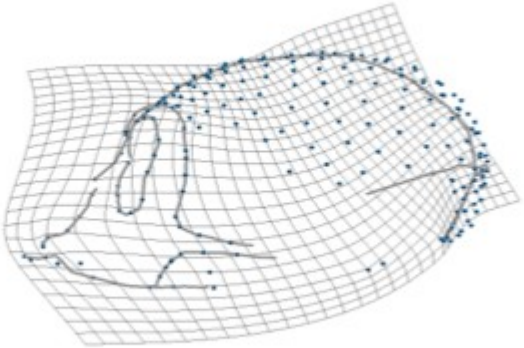

Procrustes vertical
0.4
0.2
0.0
-0.2
-0.4
-0.6
-0.8
-0.6
-0.4
-0.2
0.0
0.2
0.4
0.6
Procrustes horizontal

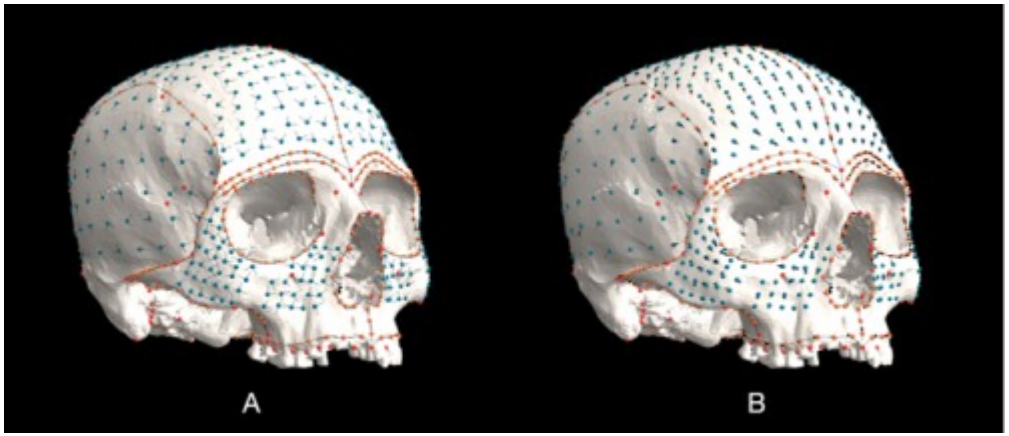
b

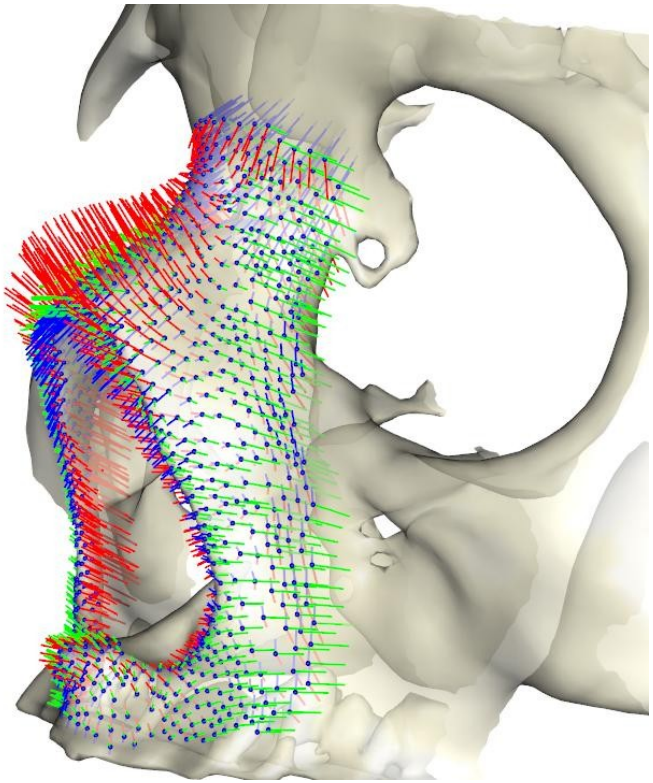
Craniometric Points, Lateral View

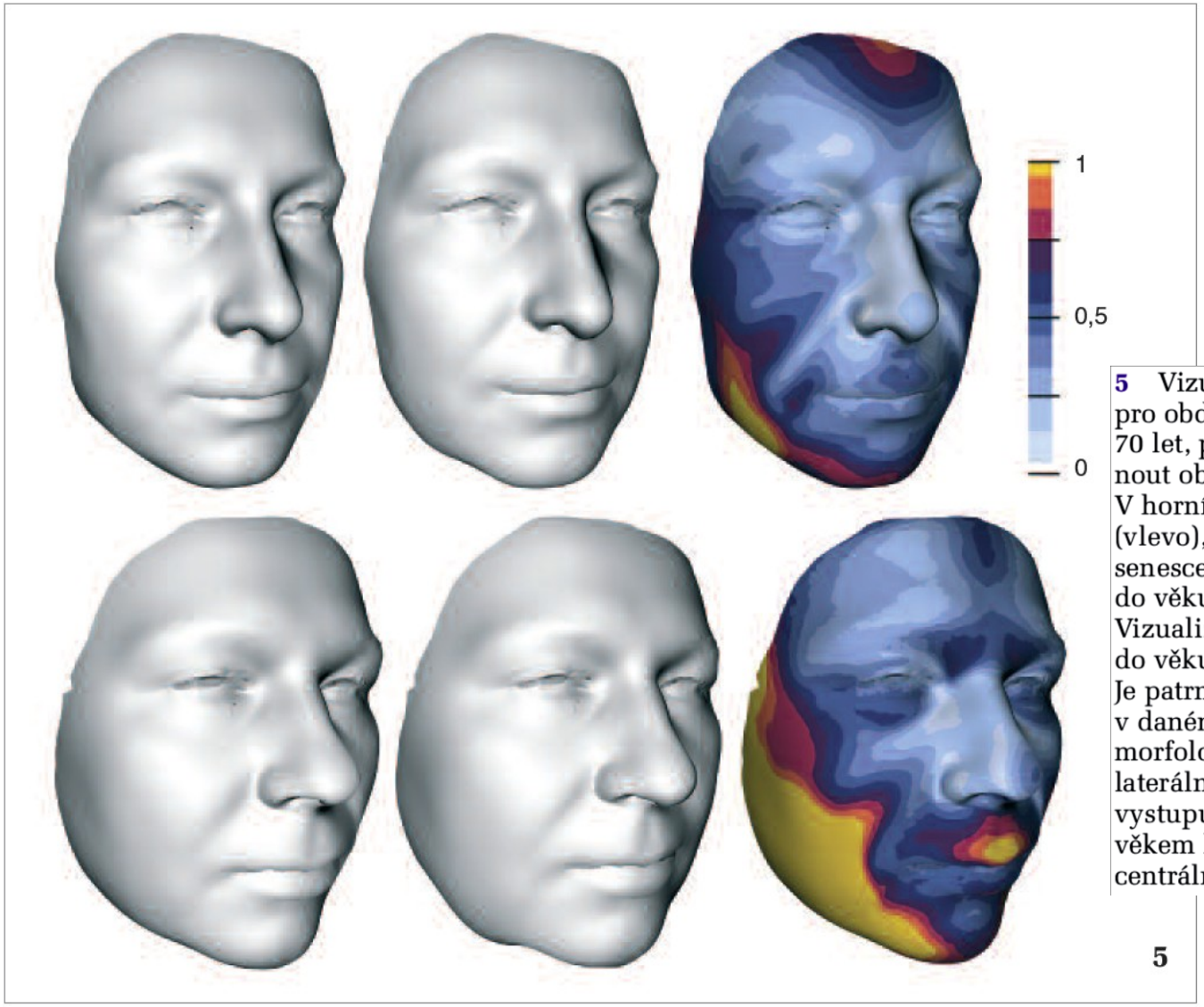






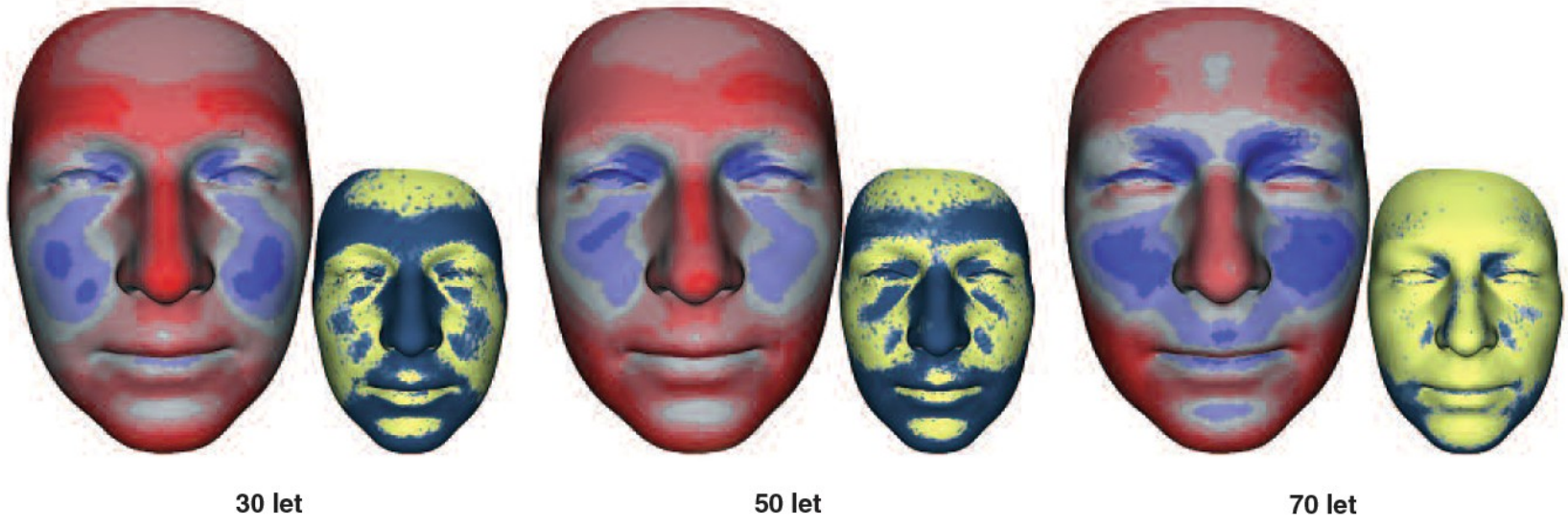

A
B





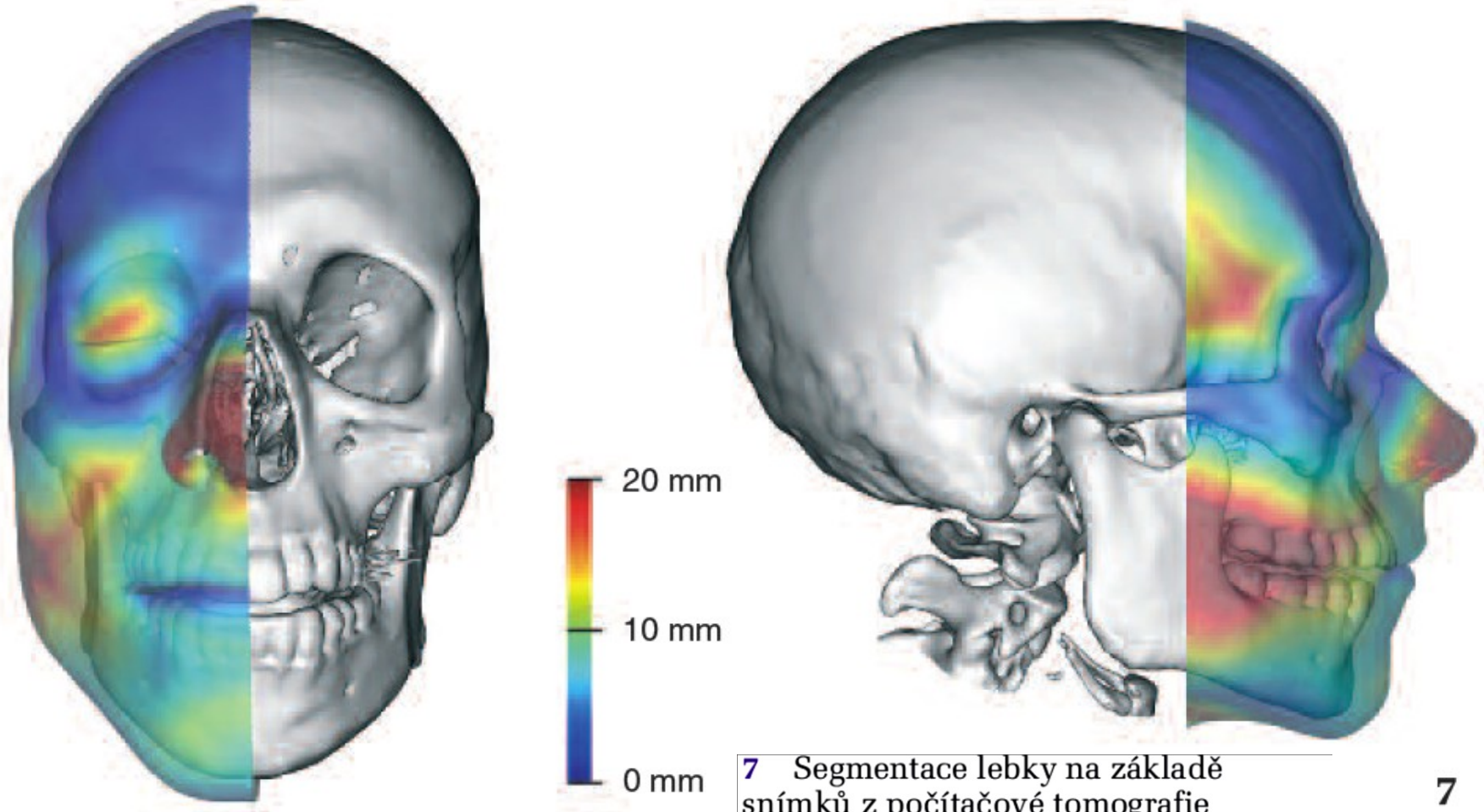
5 Vizualizace predikčního algoritmu pro období dospělého věku od 20 do 70 let, podle kterého lze omladit/zestárnout obličej v daném věkovém rozmezí. V horní řadě obličej ženy ve věku 29 let (vlevo), který podle algoritmu senescence žen (vpravo) zestárnul do věku 70 let života (uprostřed). Vizualizaci zestárnutí 23letého muže do věku 70 let ukazuje spodní řada. Je patrné, že muži prodělávají v daném věkovém intervalu výraznější morfologické změny než ženy – laterální rozšíření obličeje, méně vystupující horní ret. U žen dochází věkem k zešikmení především centrální oblasti čela.

Velemínská, Dupej, Živa 5/2016



6 Vizualizace vývoje sexuálního dimorfismu v průměrném věku 30, 50 a 70 let. Větší obličeje ukazují znaky pohlavně dimorfní ve prospěch mužů (červená barva) a žen (modrá). Menší obličeje zobrazují mapy signifikance, kde odstíny modré barvy znázorňují oblasti s průkaznými rozdíly mezi mužským a ženským pohlavím. Z těchto žlutomodrých map zřetelně vidíme, že sexuální dimorfismus obličeje se s věkem výrazně snižuje.

Velemínská, Dupej, Živa 5/2016

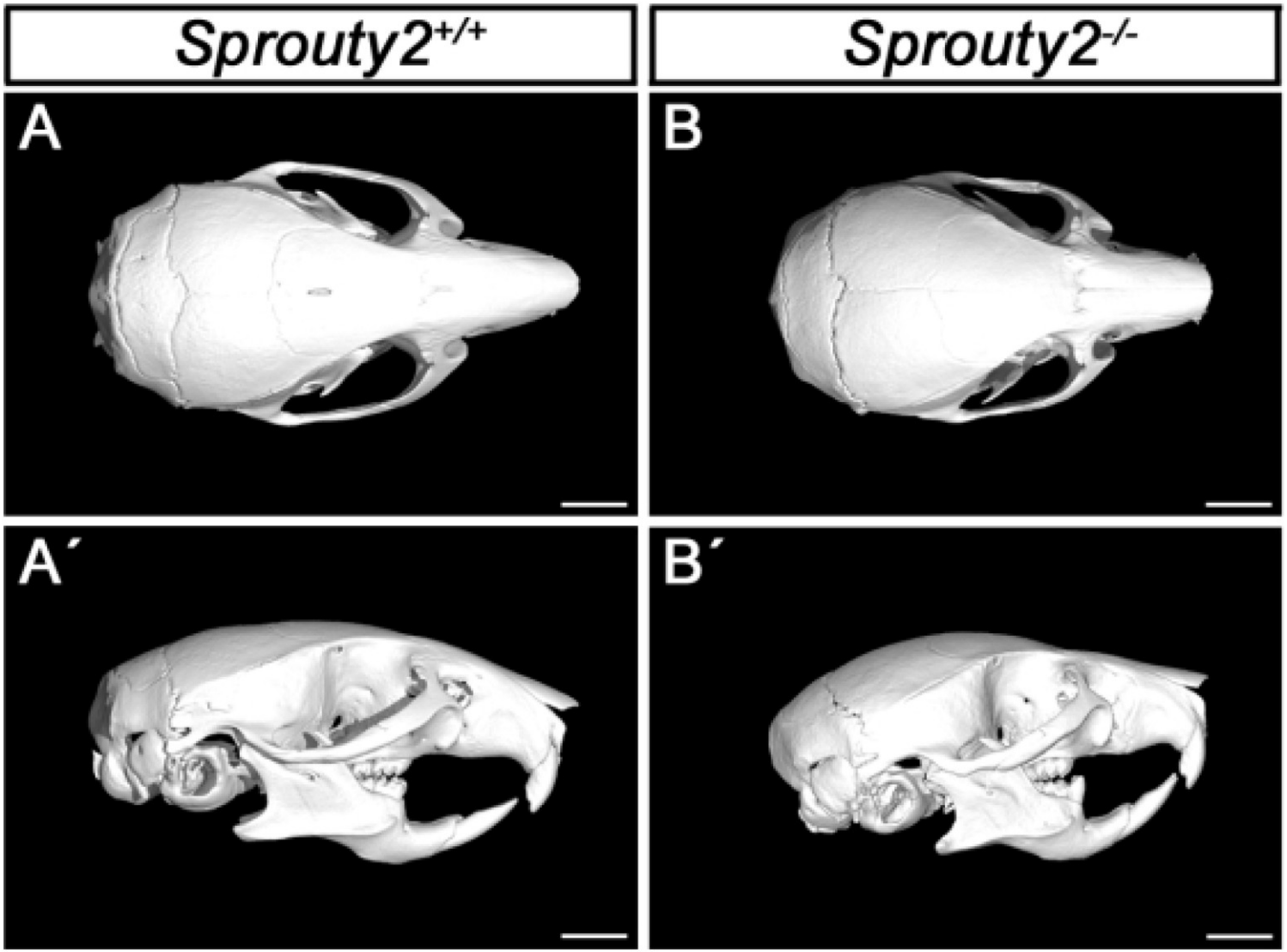


7 Segmentace lebky na základě snímků z počítačové tomografie hlavy člověka. Měkké tkáně jsou znázorněny průsvitně jen na části lebky a pomocí barevné mapy je odstupňována tloušťka měkkých tkání, zásadní pro rekonstrukci obličeje podle lebky (vlevo pohled zepředu, vpravo ze strany).
Všechny orig.: J. Dupej a J. Velemínská

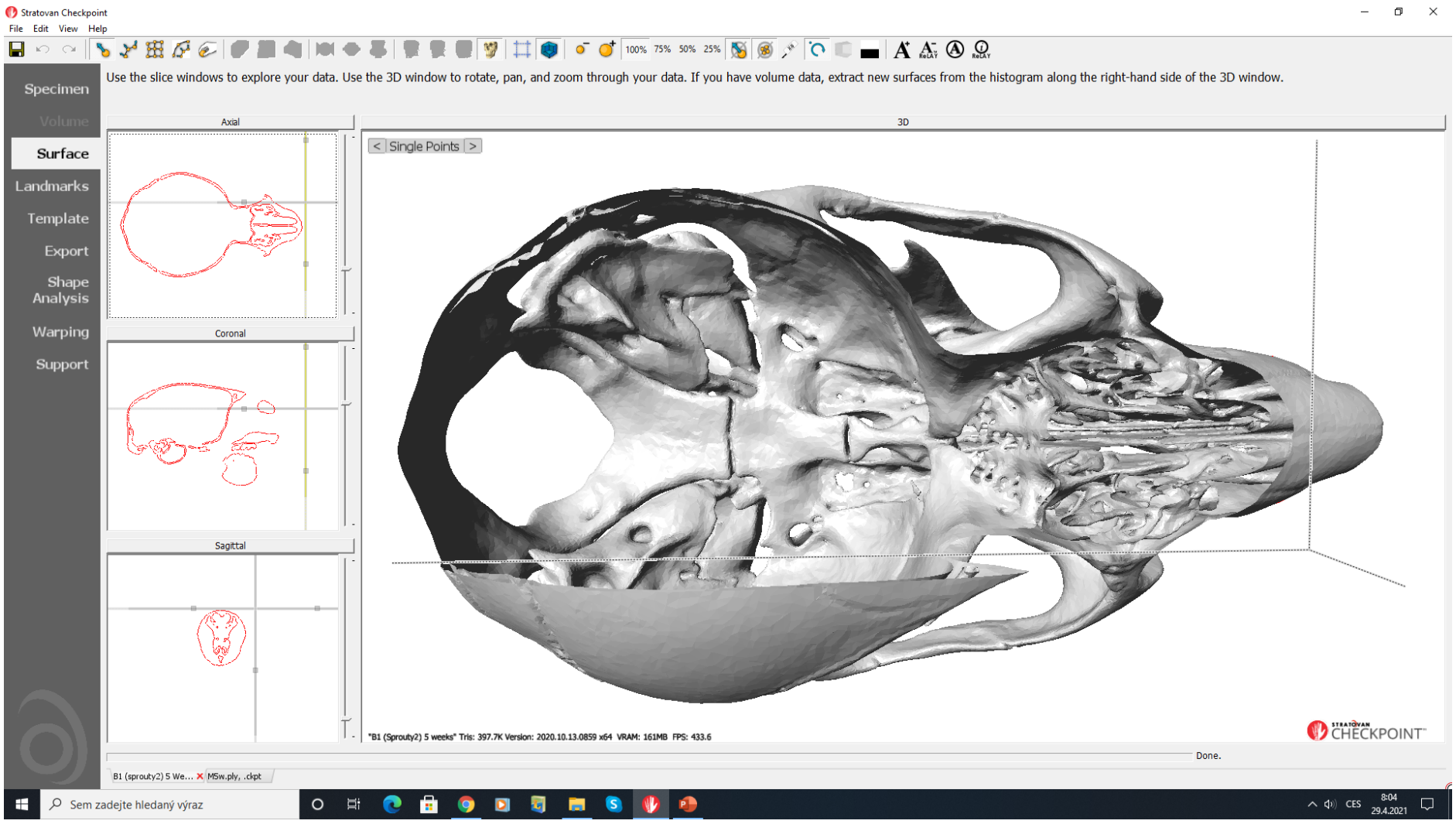
7

Velemínská, Dupej, Živa 5/2016

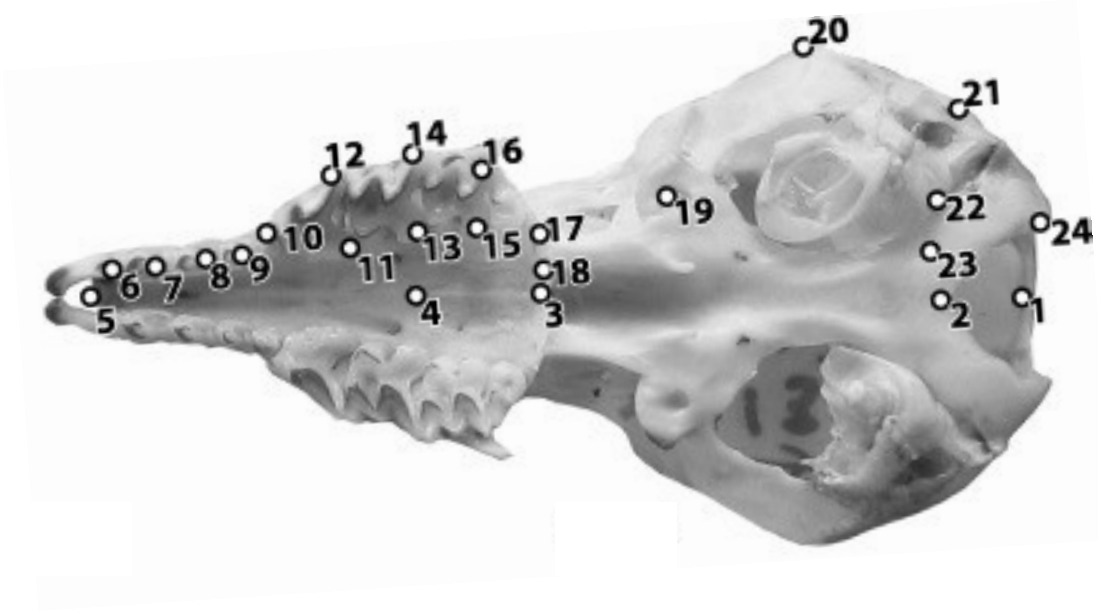
Example: knockout of gene *Sprouty*



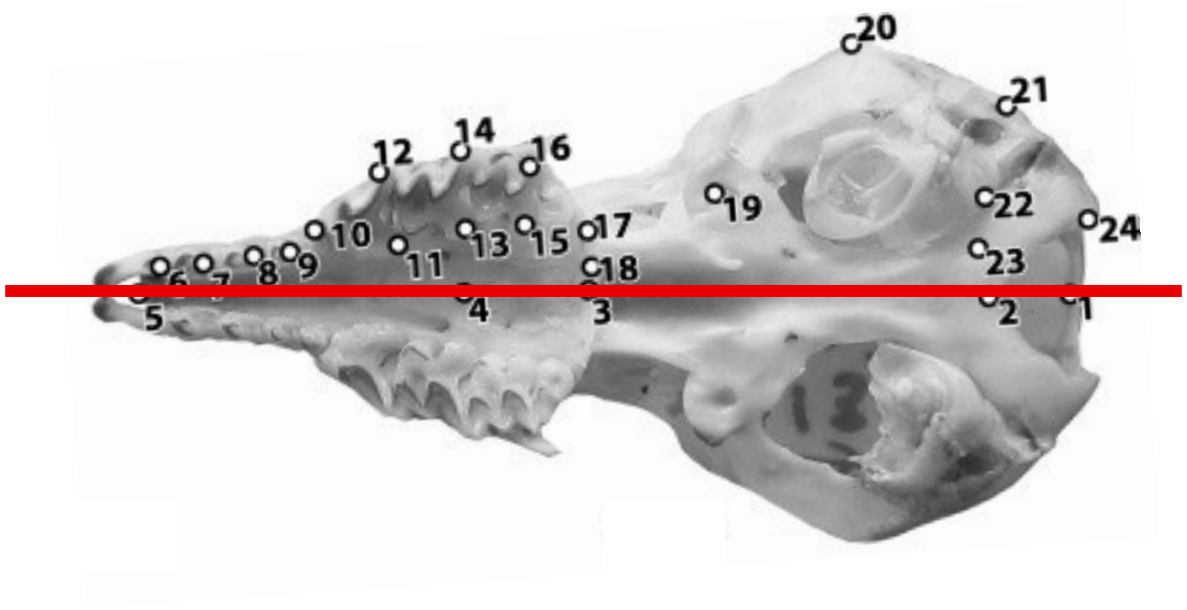
Program Checkpoint (Stratovan):



Problem of symmetric objects

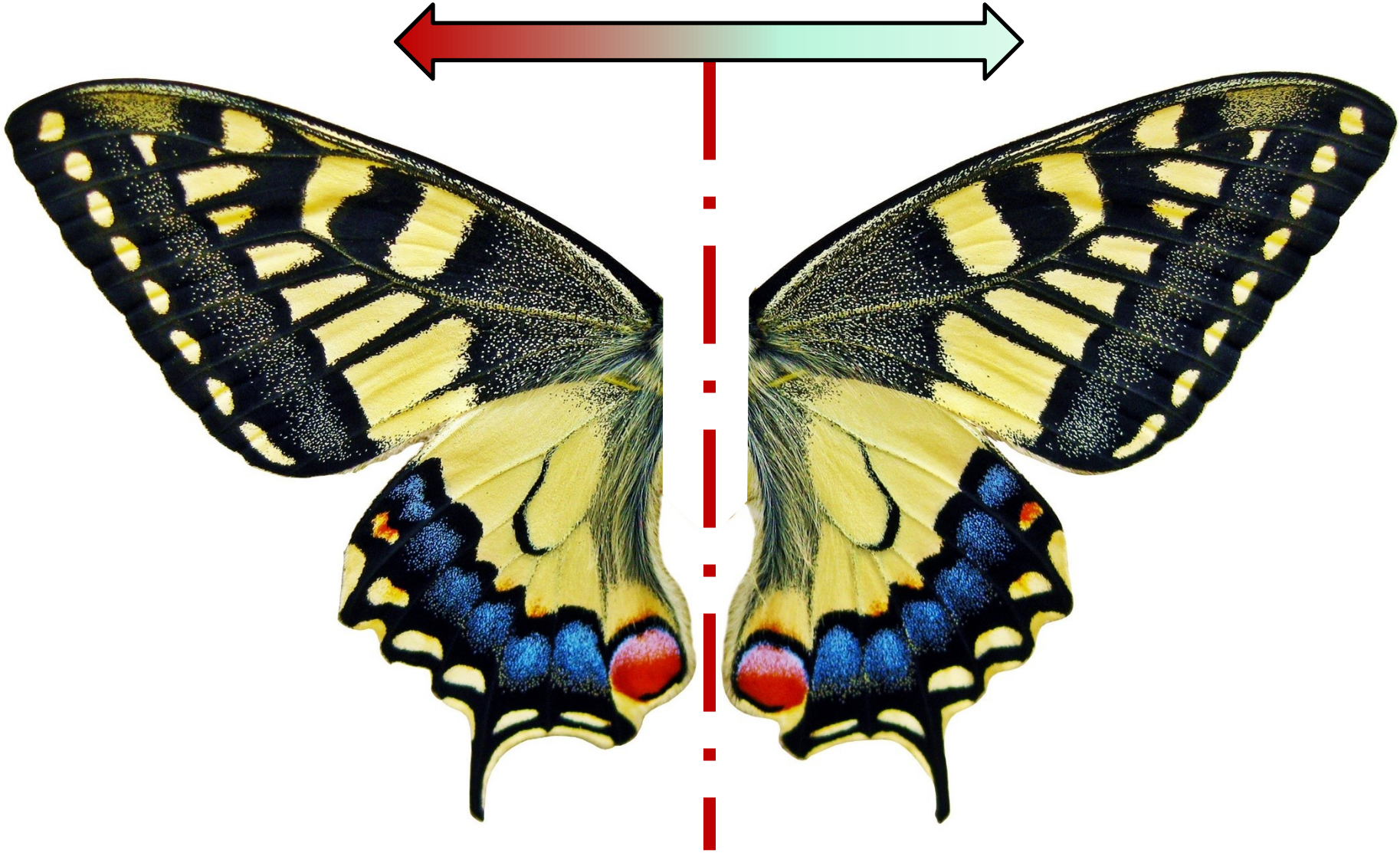


Problem of symmetric objects



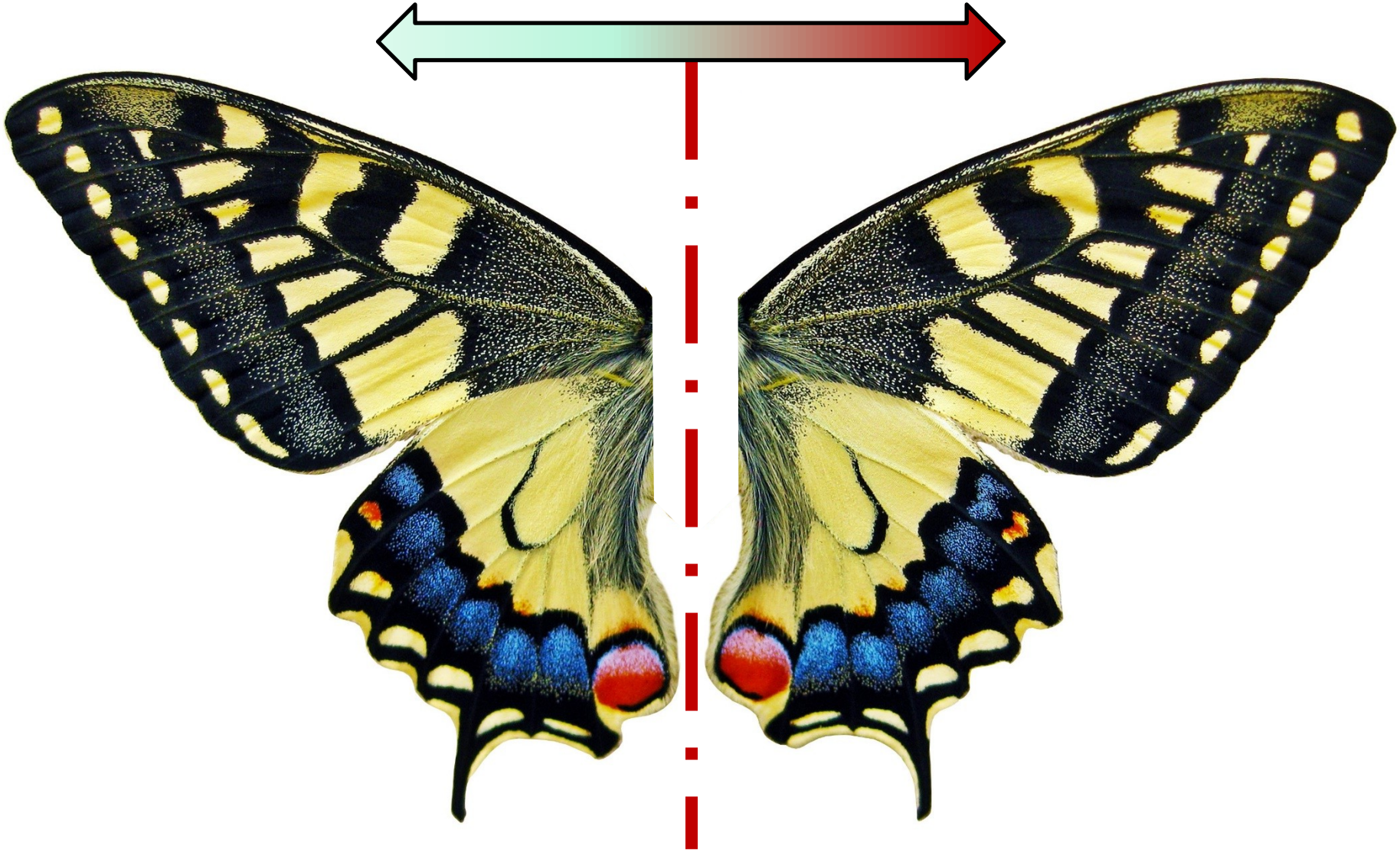
Problem of symmetric objects

matching symmetry



Klingenberg et al., Evolution (2002)

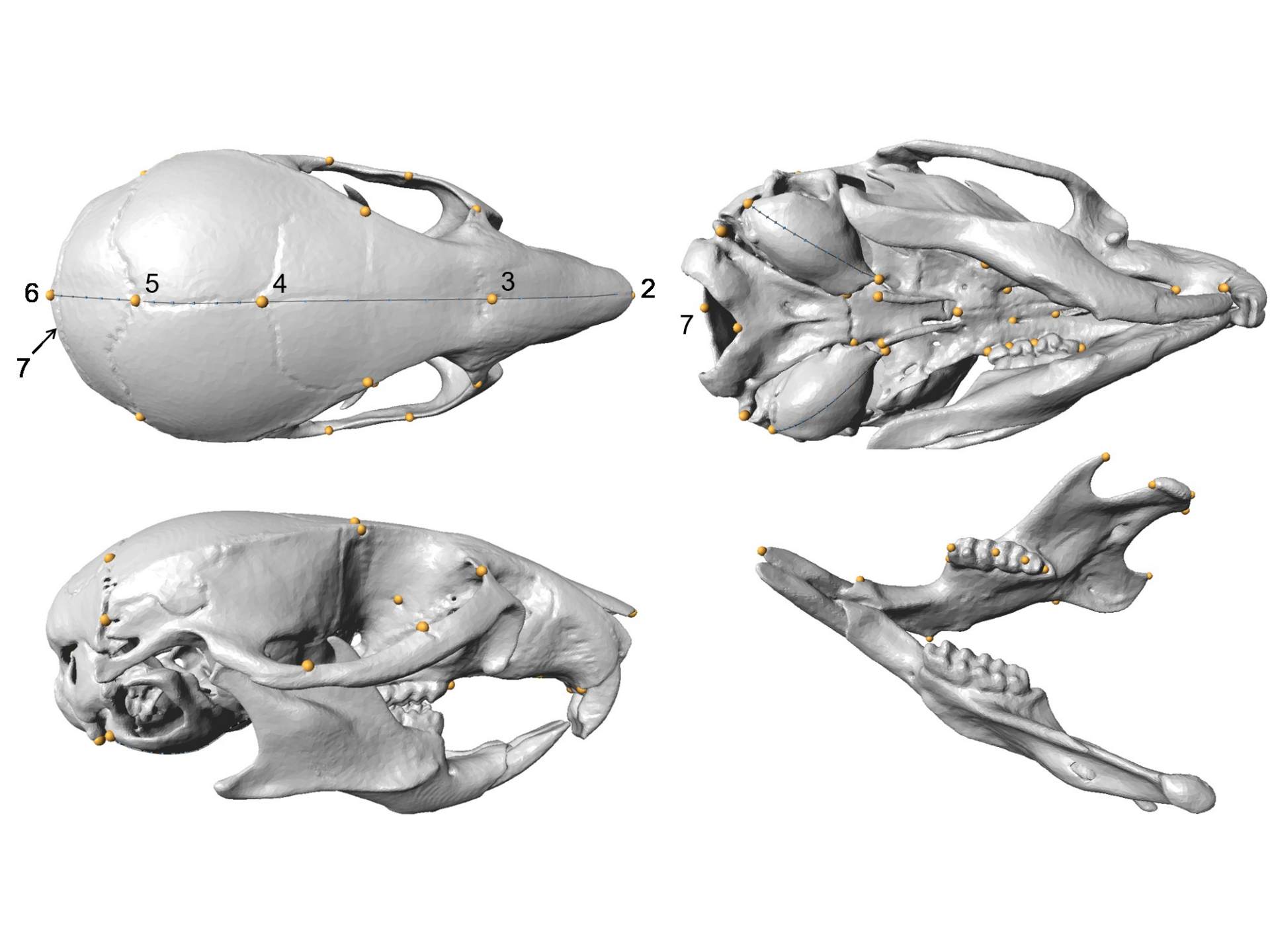
matching symmetry



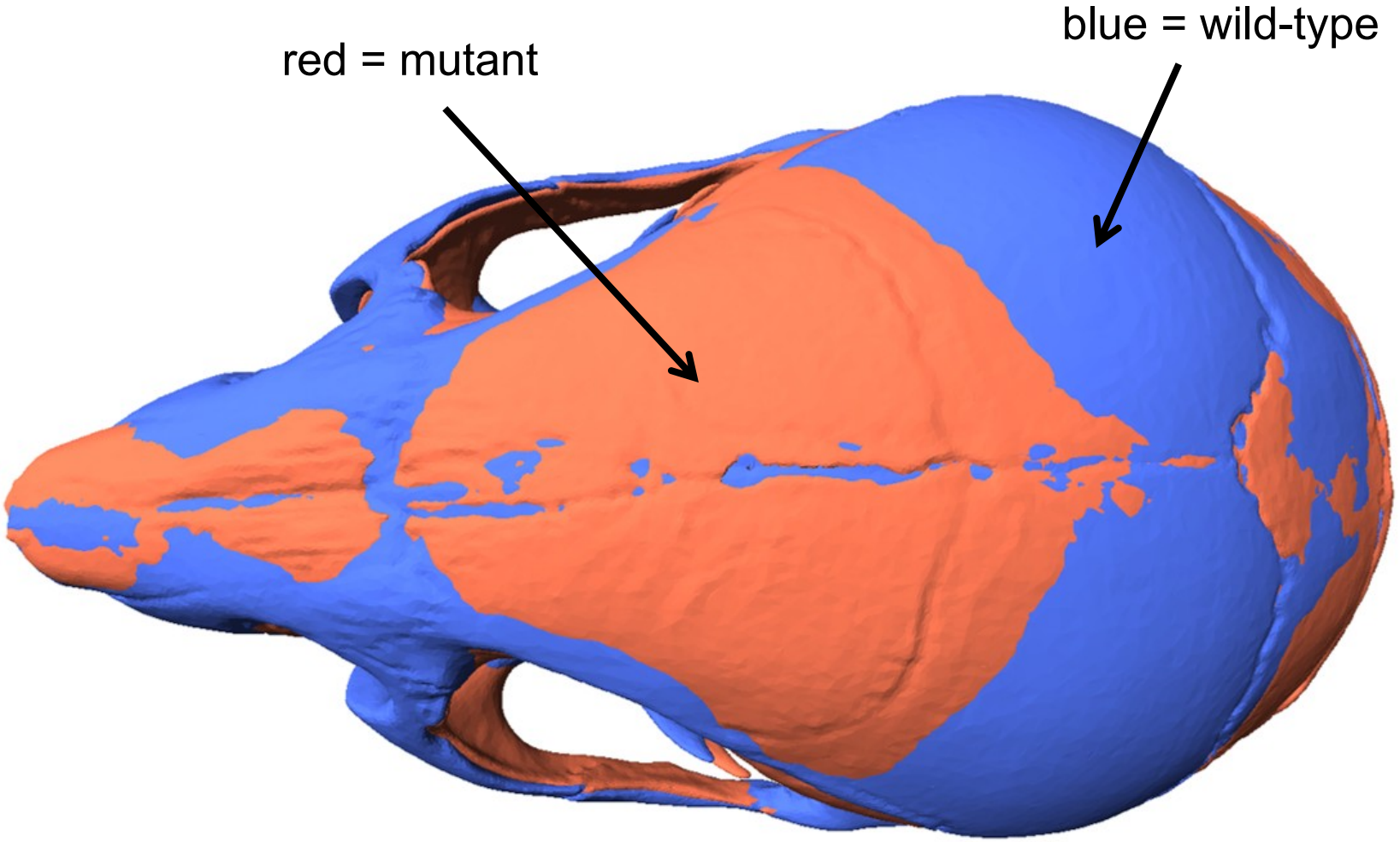
Klingenberg et al., Evolution (2002)

object symmetry

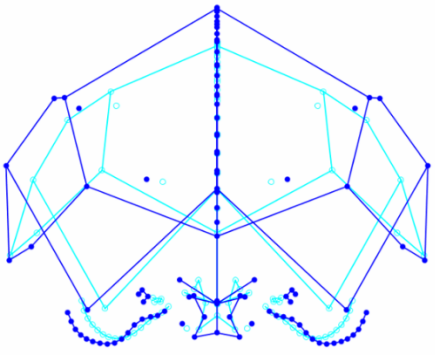
object symmetry


6
5
4
3
2
7
7

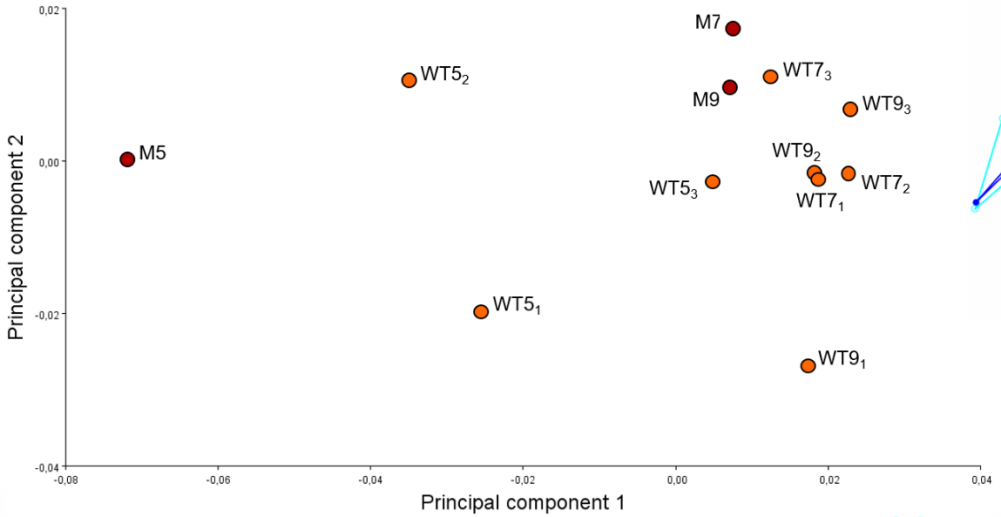
Comparison of wild-type and mutant mouse, 5 weeks

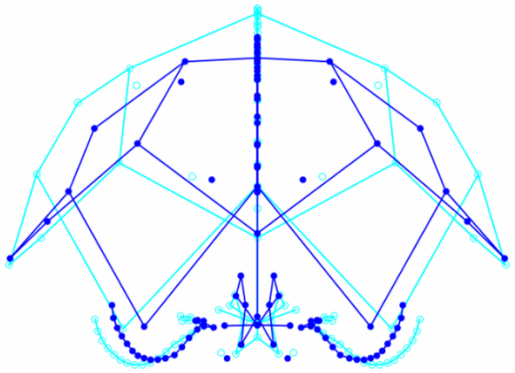


Stratovan Checkpoint

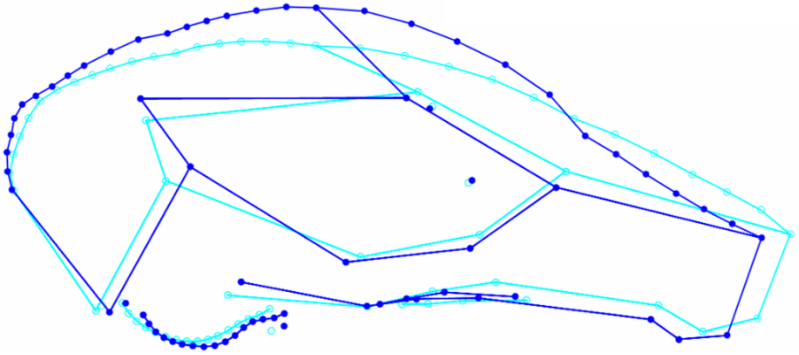


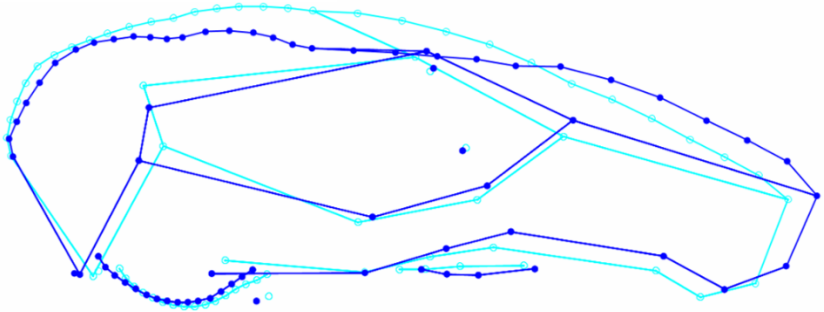
PC1-



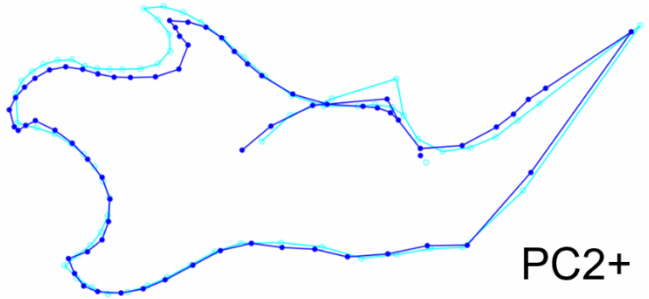


PC1+

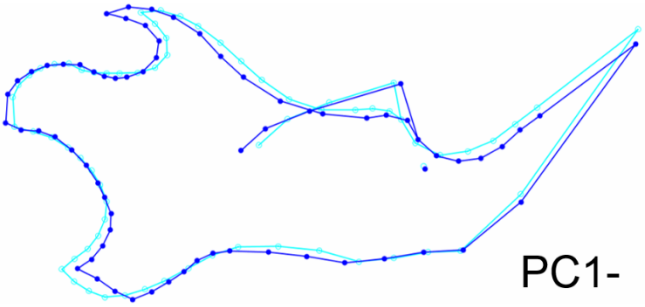




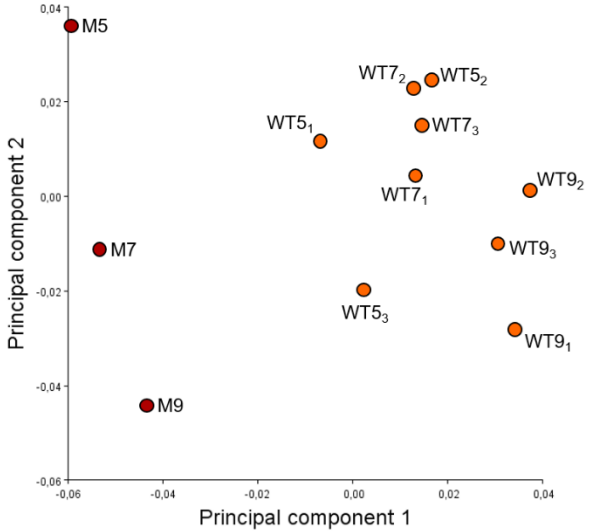
MorphoJ

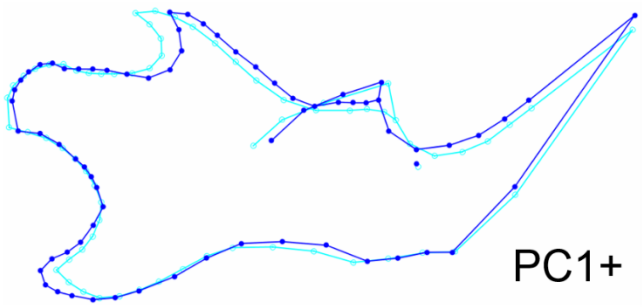


PC2+

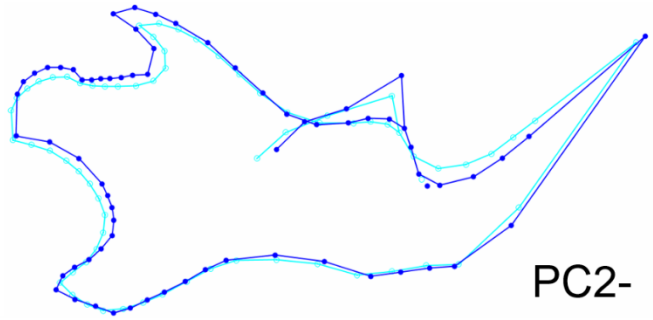


PC1-



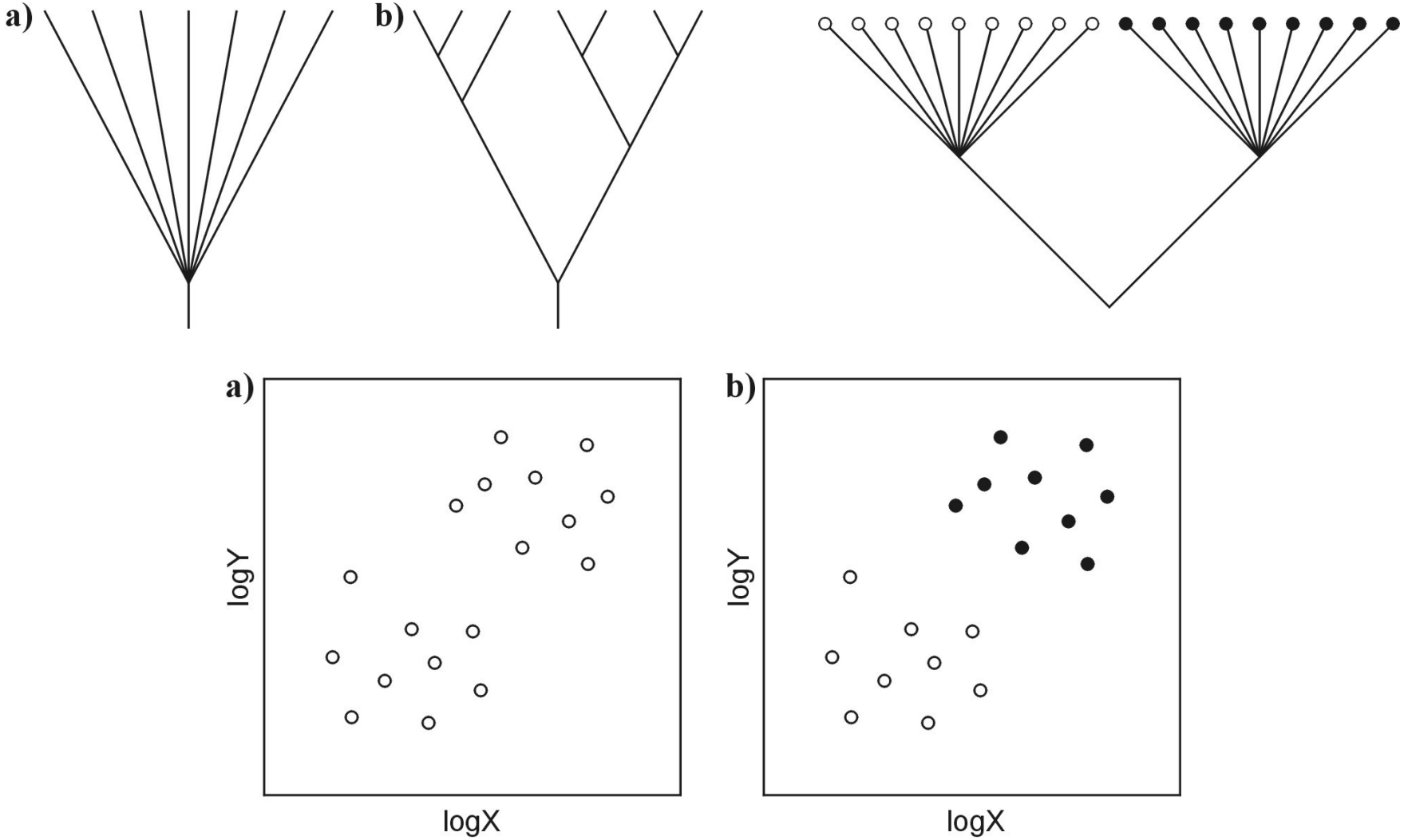


PC1+

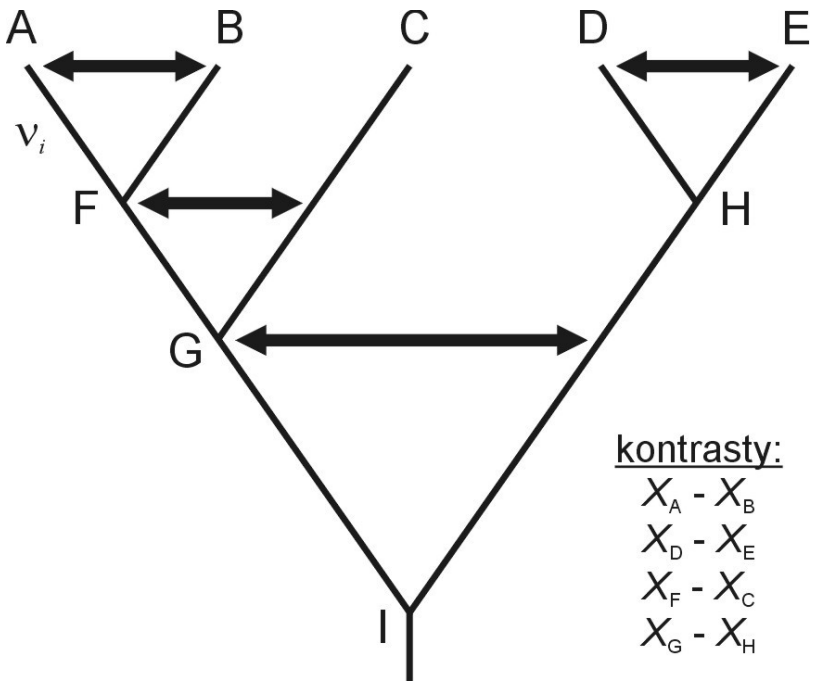


PC2-

Comparative analysis



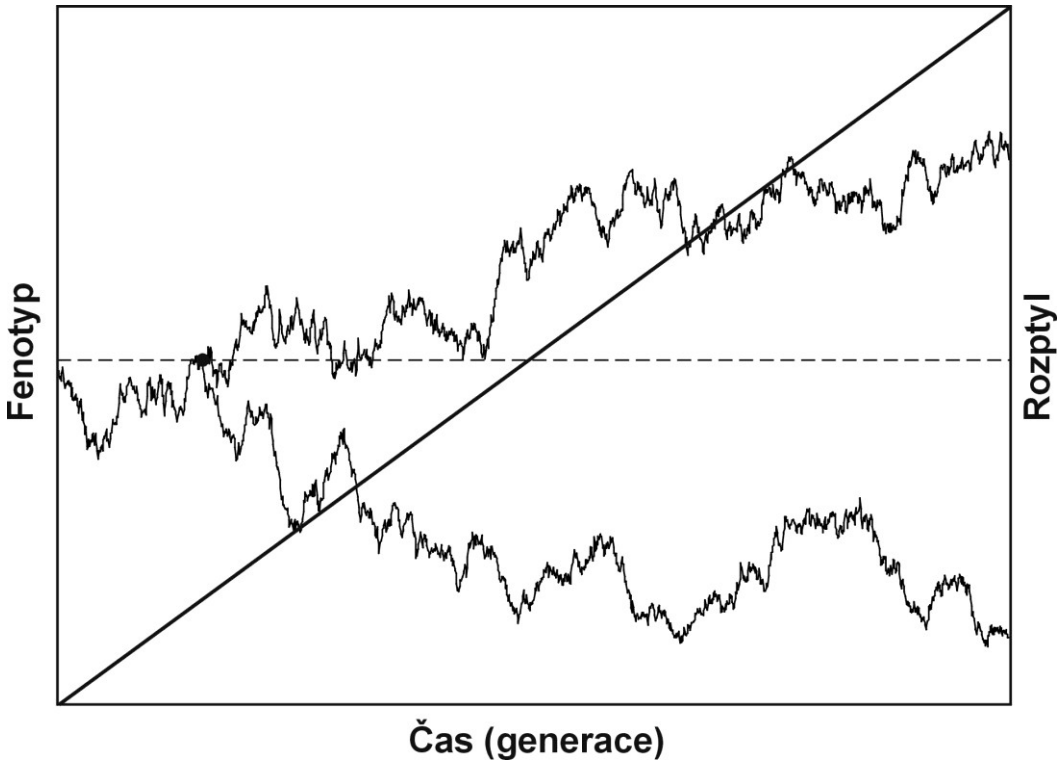
Independent contrasts

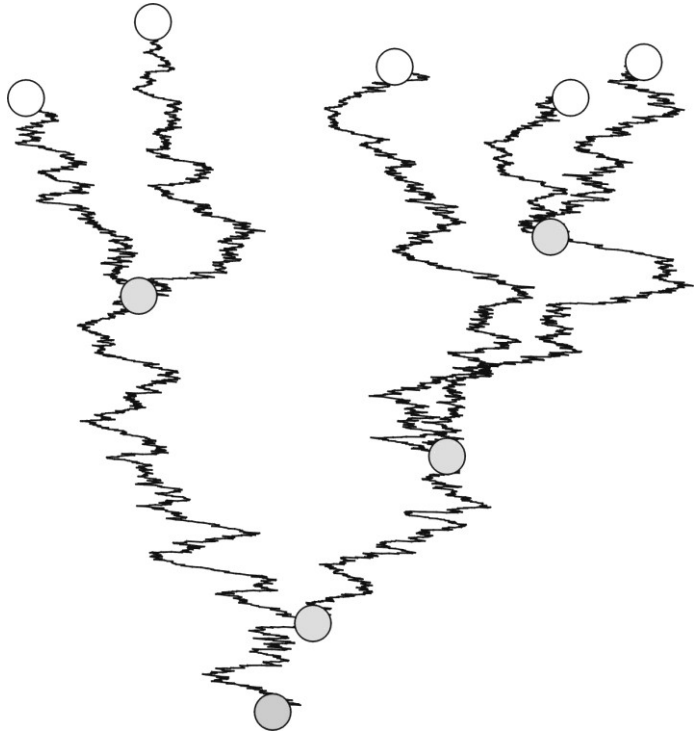
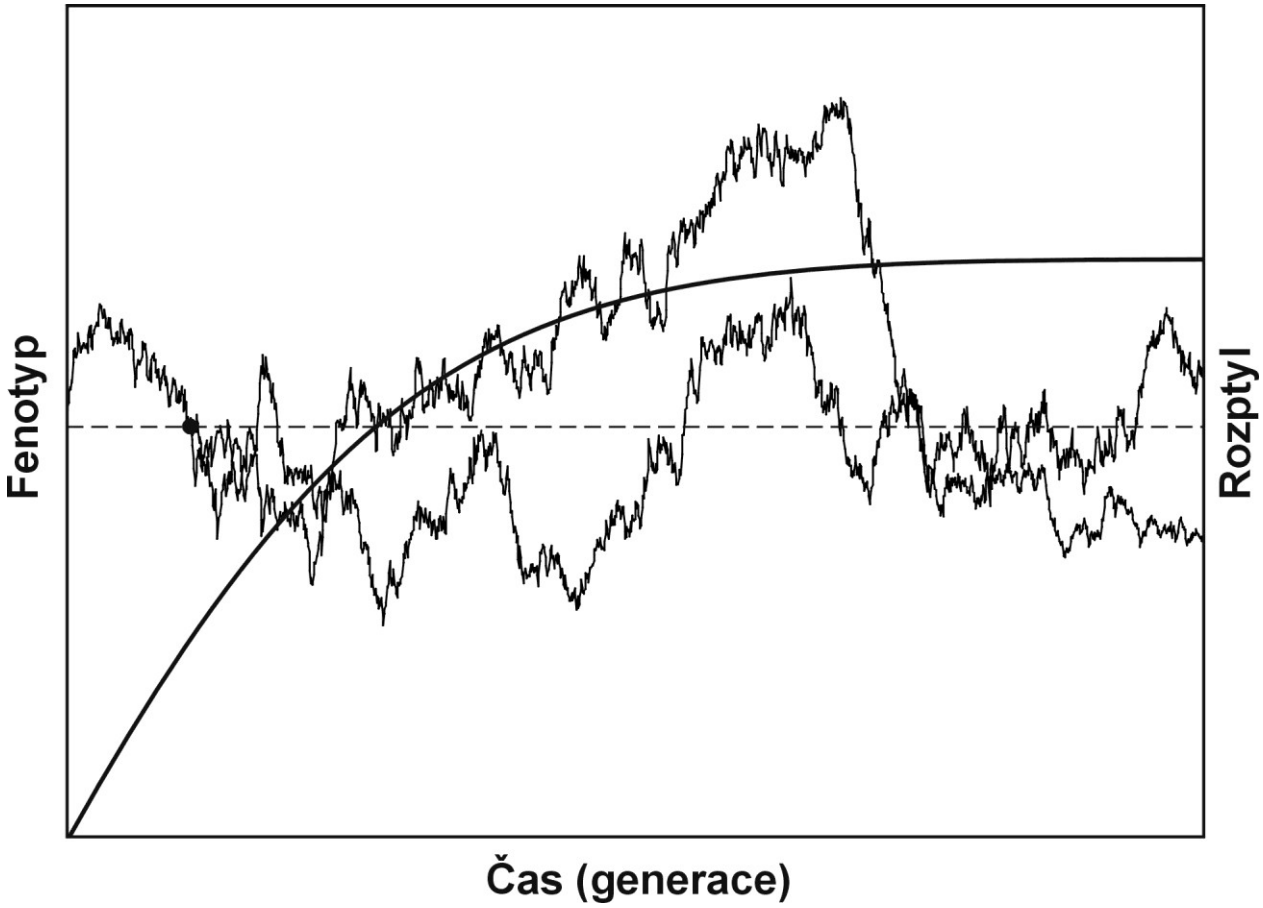


assumption: Brownian motion!

Alternatives: phylogenetic generalized least squares (PGLS)
possibility of also applying other models than Brownian motion

Brownian motion model





Ornstein-Uhlenbeck model