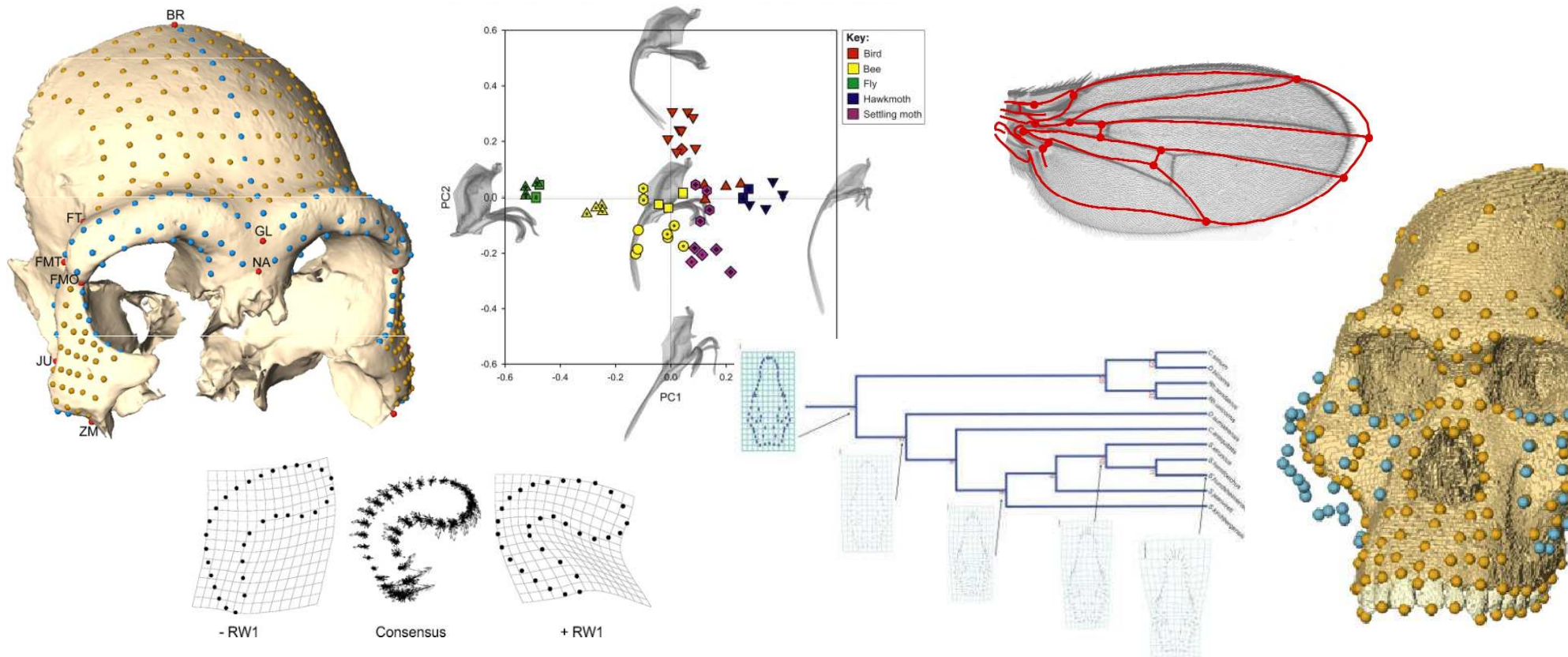


ANALYSIS OF PHENOTYPE



Genetic methods and morphology

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

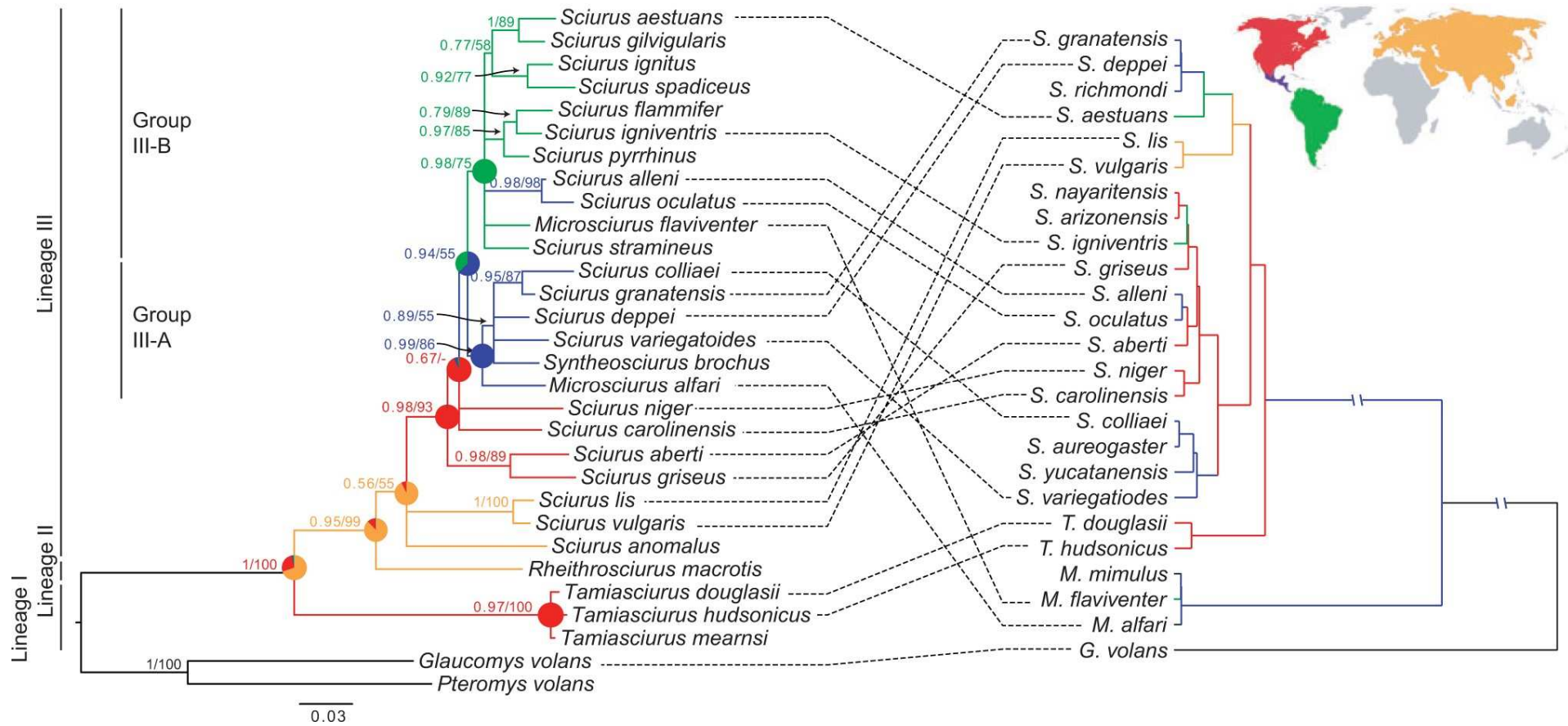
Trait variation in time (phylogenesis)

Trait variation depending on other factors

Some methods are shared (e.g. PCA)

DNA

morfometrie



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 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 traits

Mendelian inheritance, 1 – few genes
mutations in *D. melanogaster*
mutations of *Hox* genes:
Antennapedia, *Ultrabithorax*

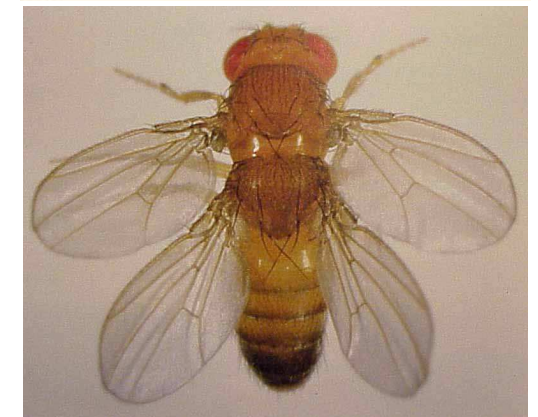
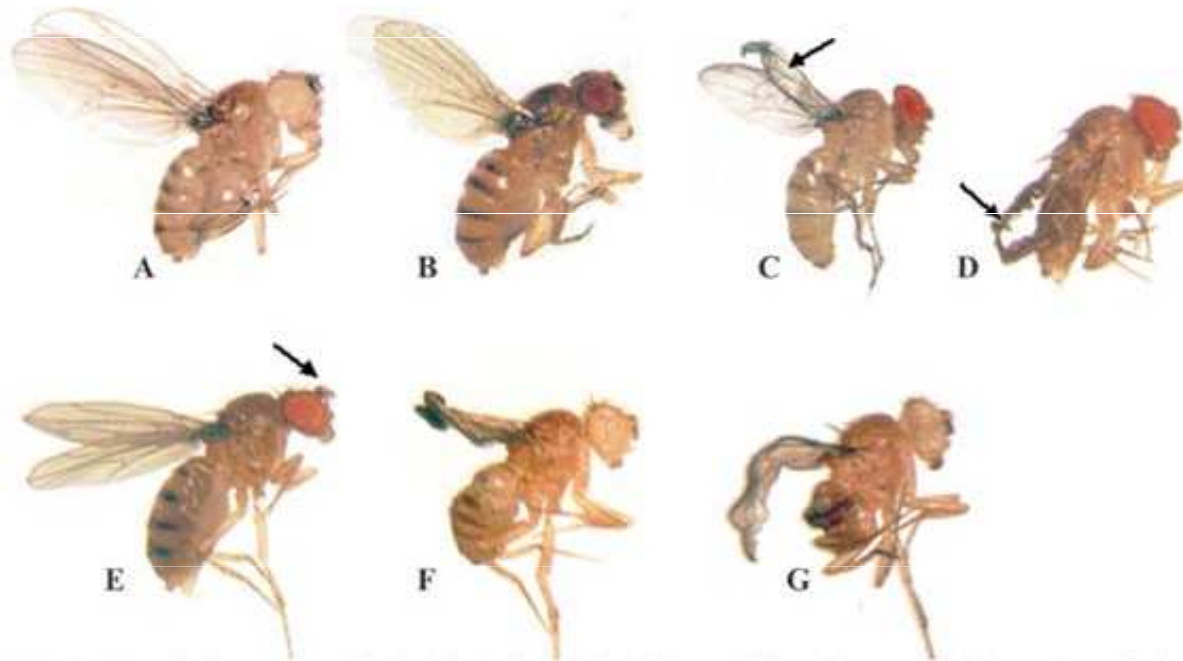


Fig. 2. Altered phenotypes obtained in individuals of *Drosophila willistoni* 17A2 isostrain submitted to temperature stress. Picture shows female individuals with altered morphologies: A, *white* (white eyes); B, *sepia* (brown eyes); C, *blistered* (arrow indicates the presence of blisters on the wings); 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 traits



colouration:

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

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



pigments:

eumelanin, phaeomelanin, carotens,
haemoglobin, trichosiderin

Qualitative traits

main allelic series:

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 = *dilute* – *brown* - non-*agouti*



albino

Qualitative traits

main allelic series :

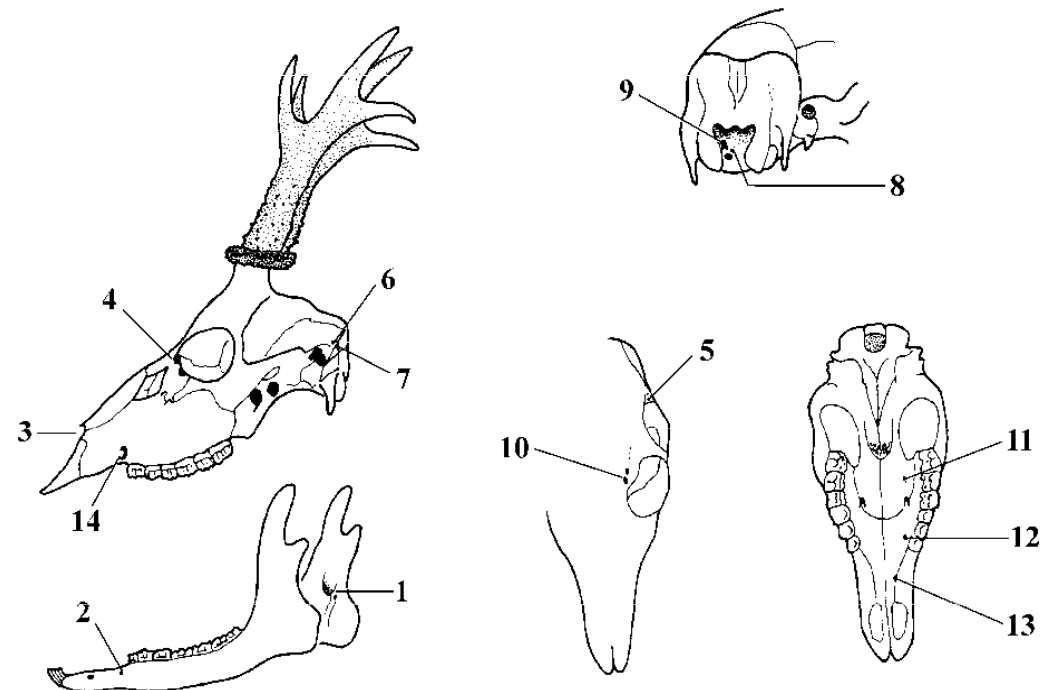
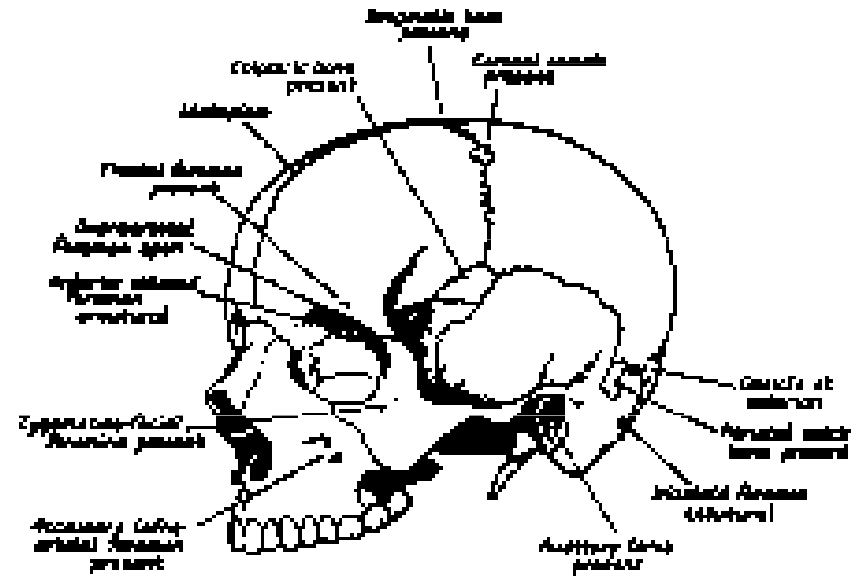
cat: *A-agouti*, *T-tabby*, *B-brown*, *O-orange*,
S-white spotting, *W-white*, *L-long hairs*



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

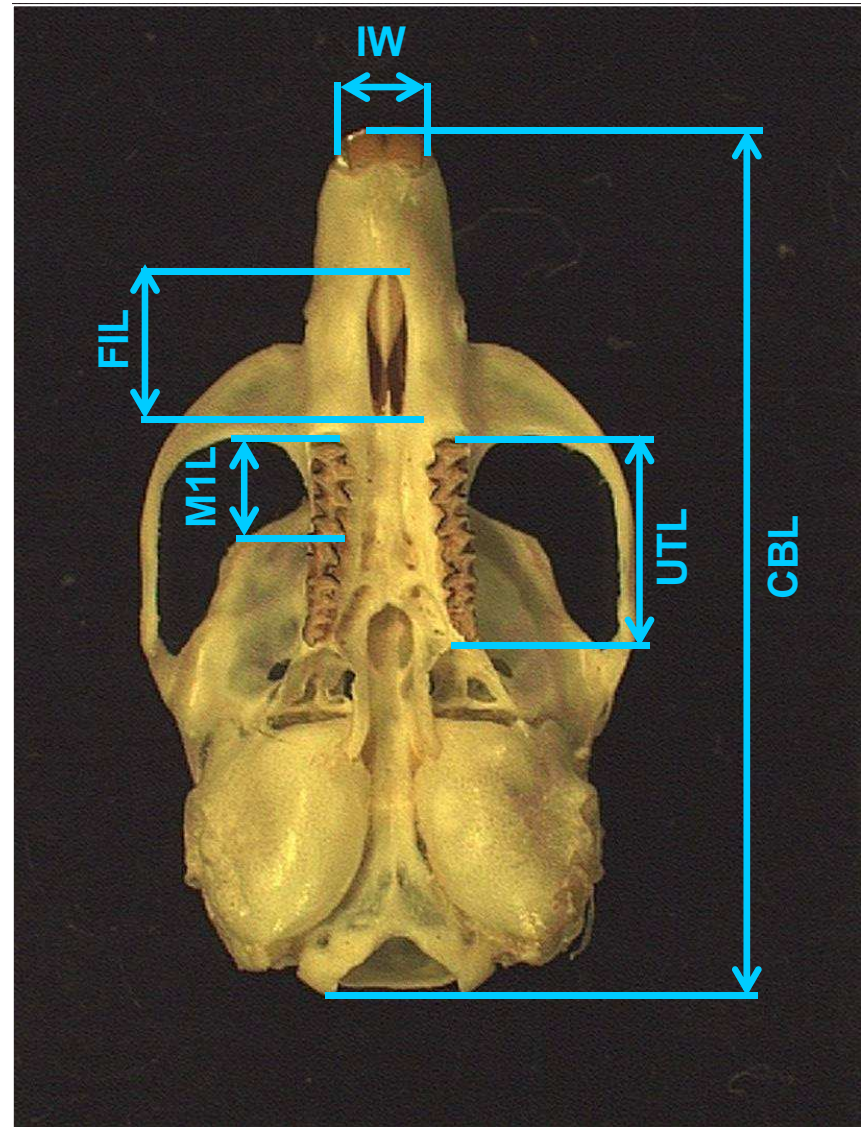
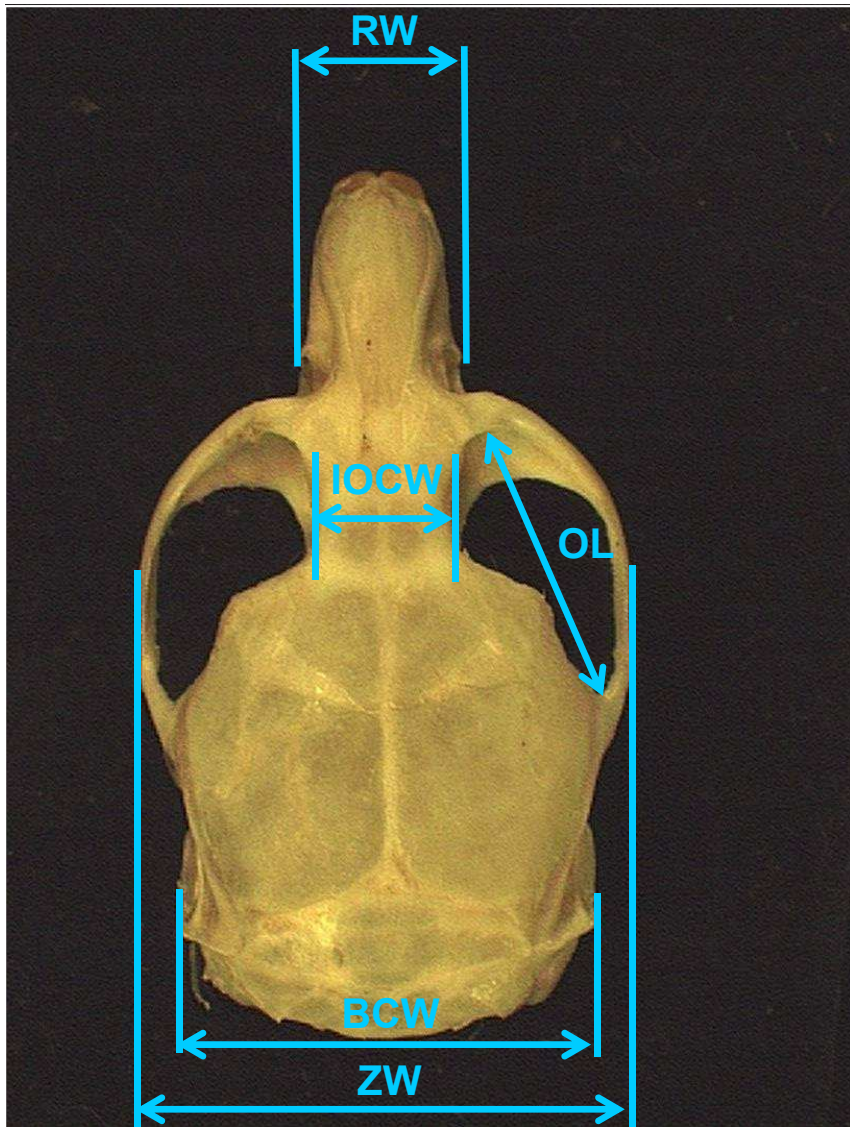
says, to what extent phenotypic variance has genetic basis

True heritability:

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

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

Traditional morphometrics

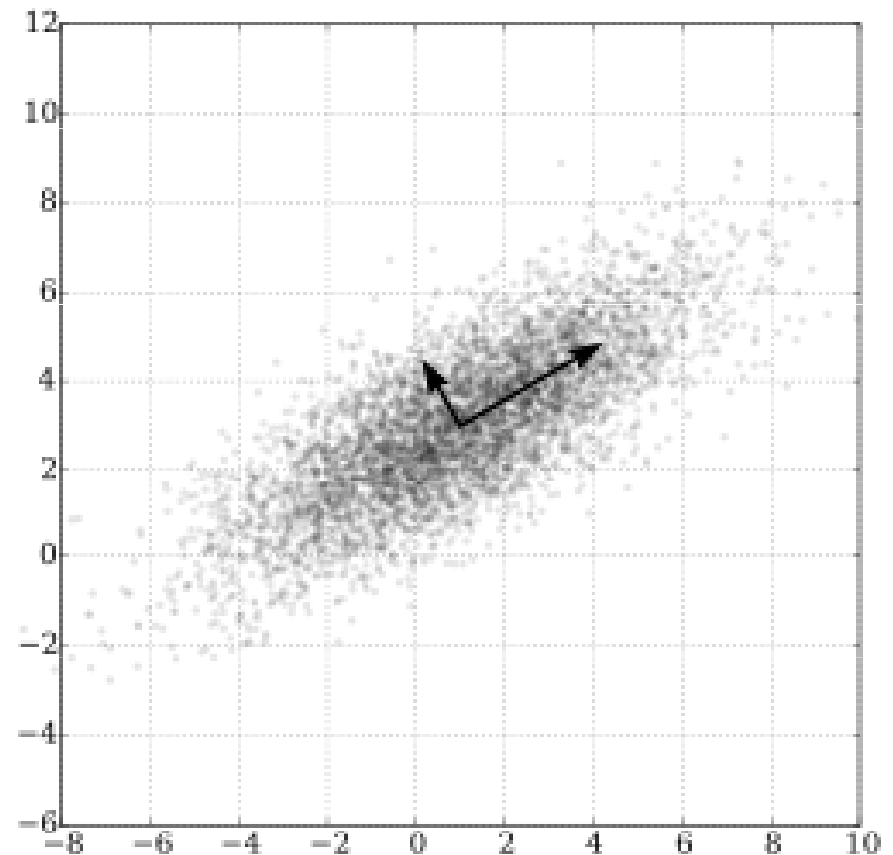


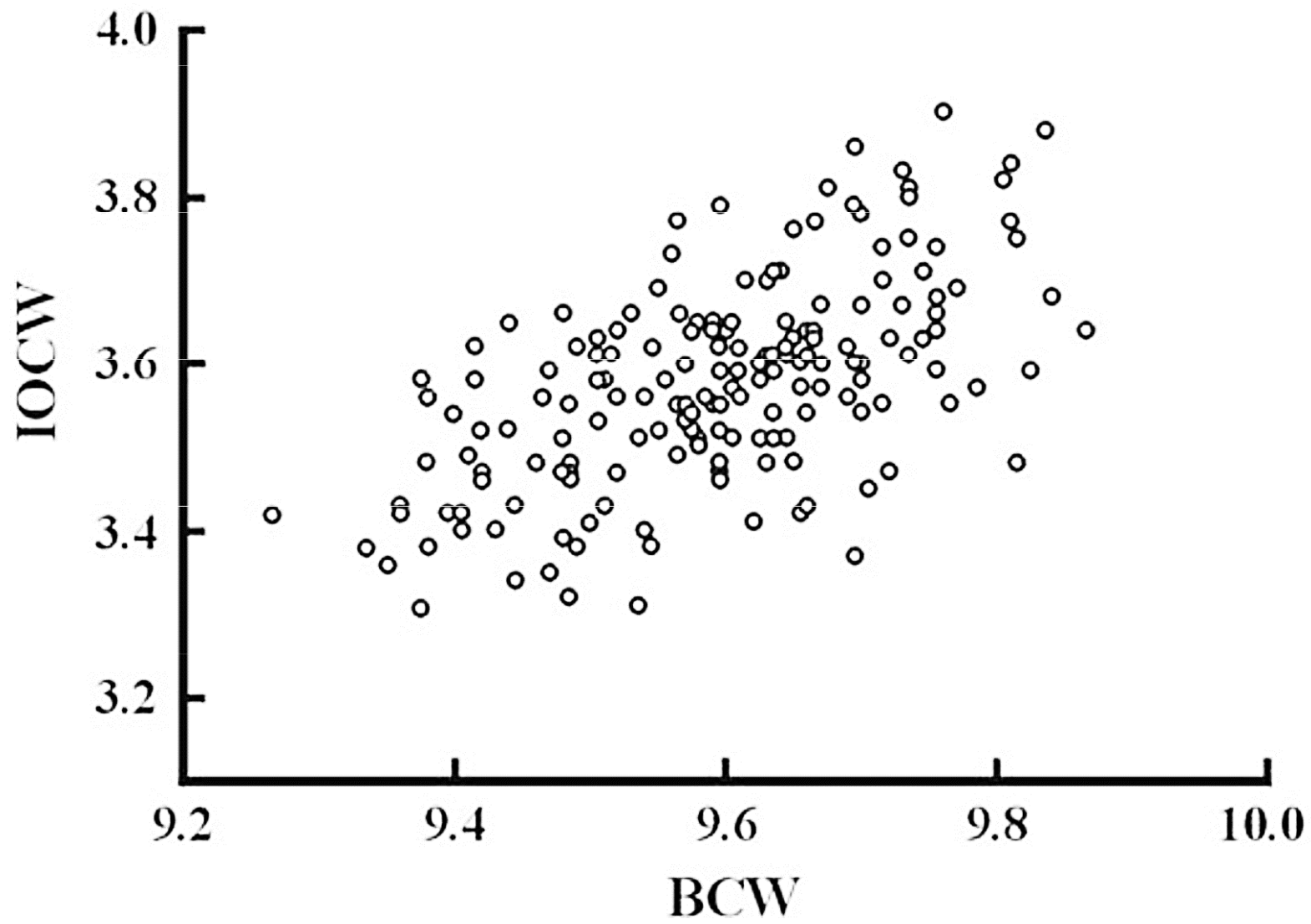
Principal components analysis (PCA)

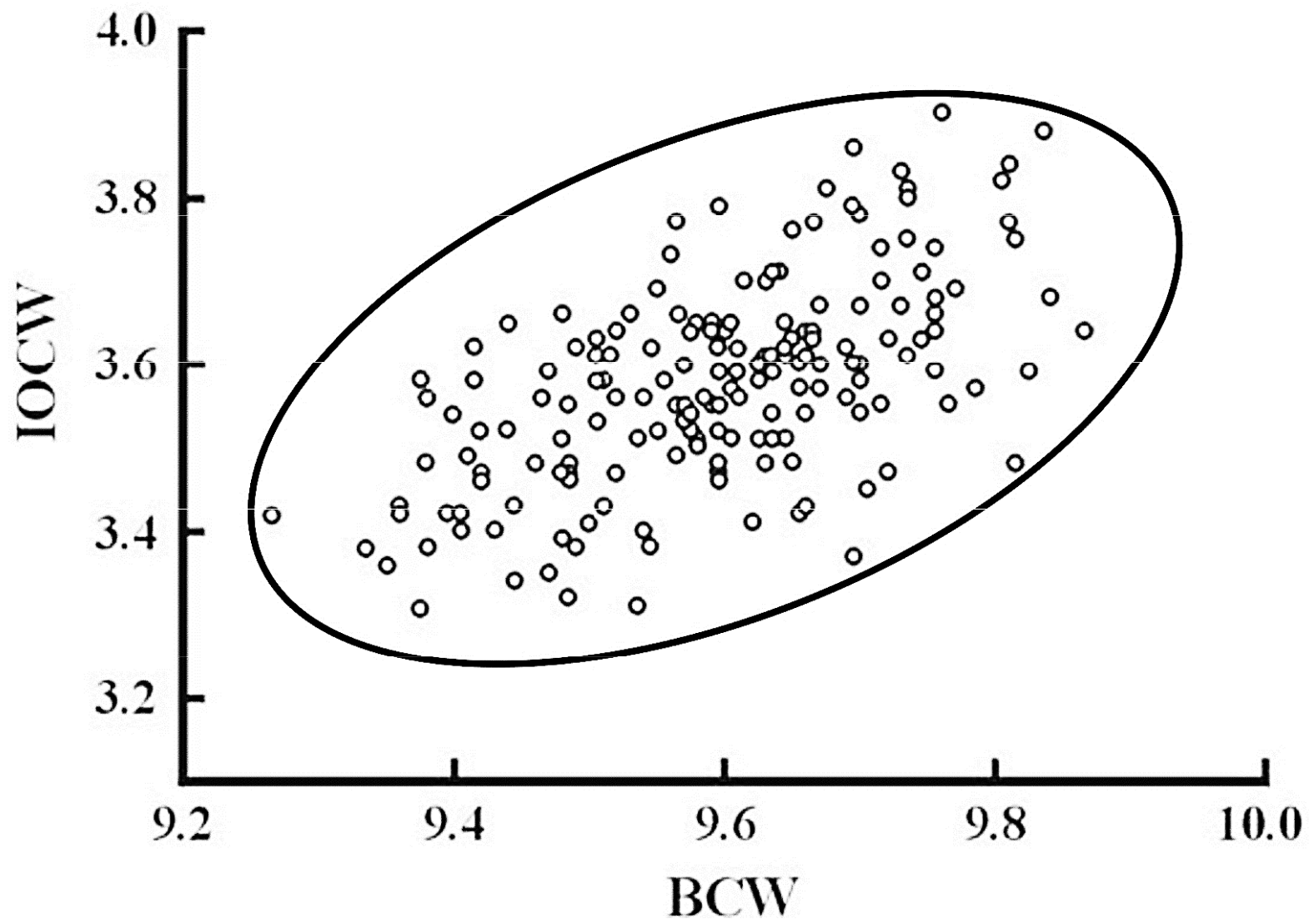
snížení dimenze dat s co nejmenší ztrátou informace

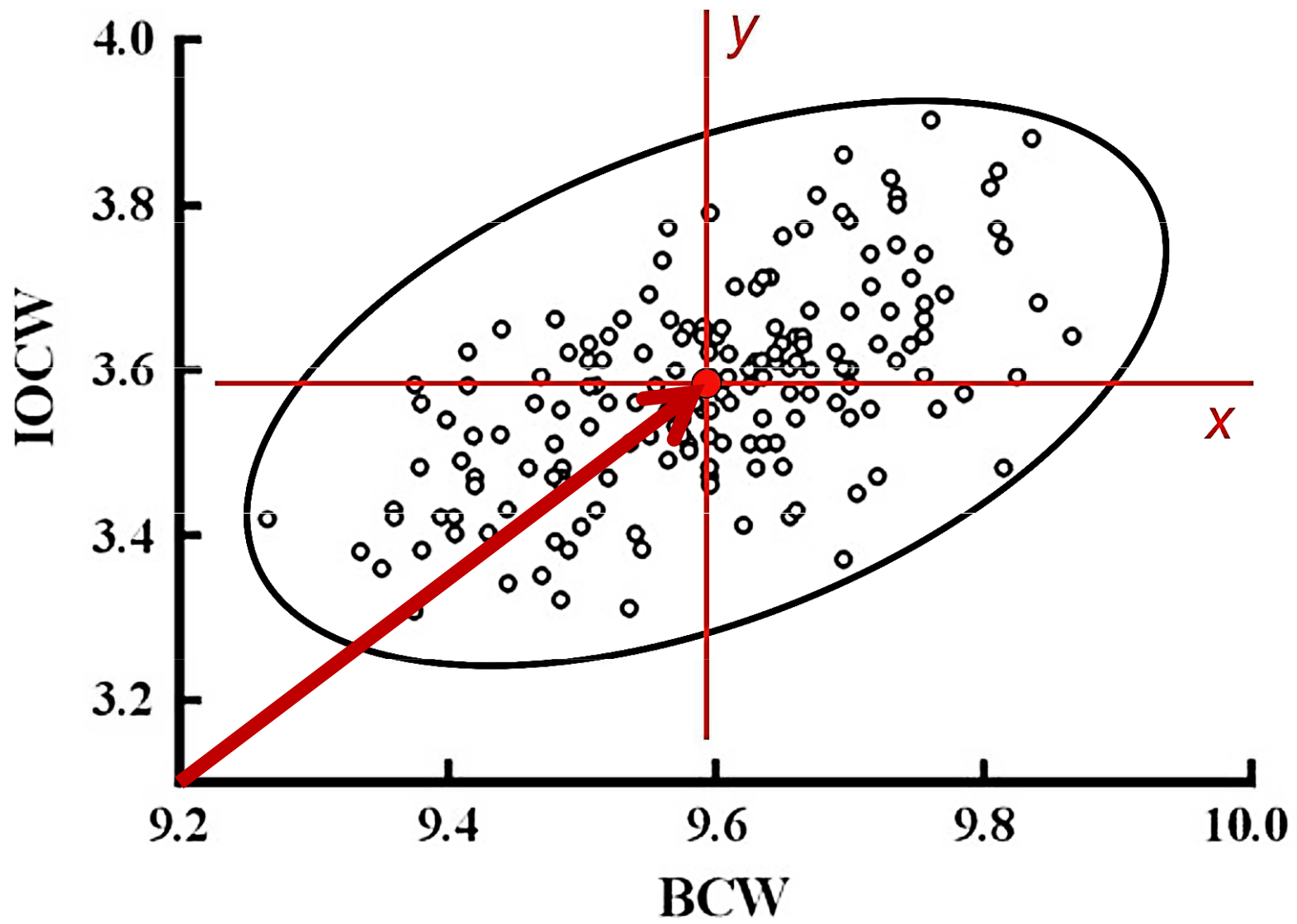
exploratory data analysis

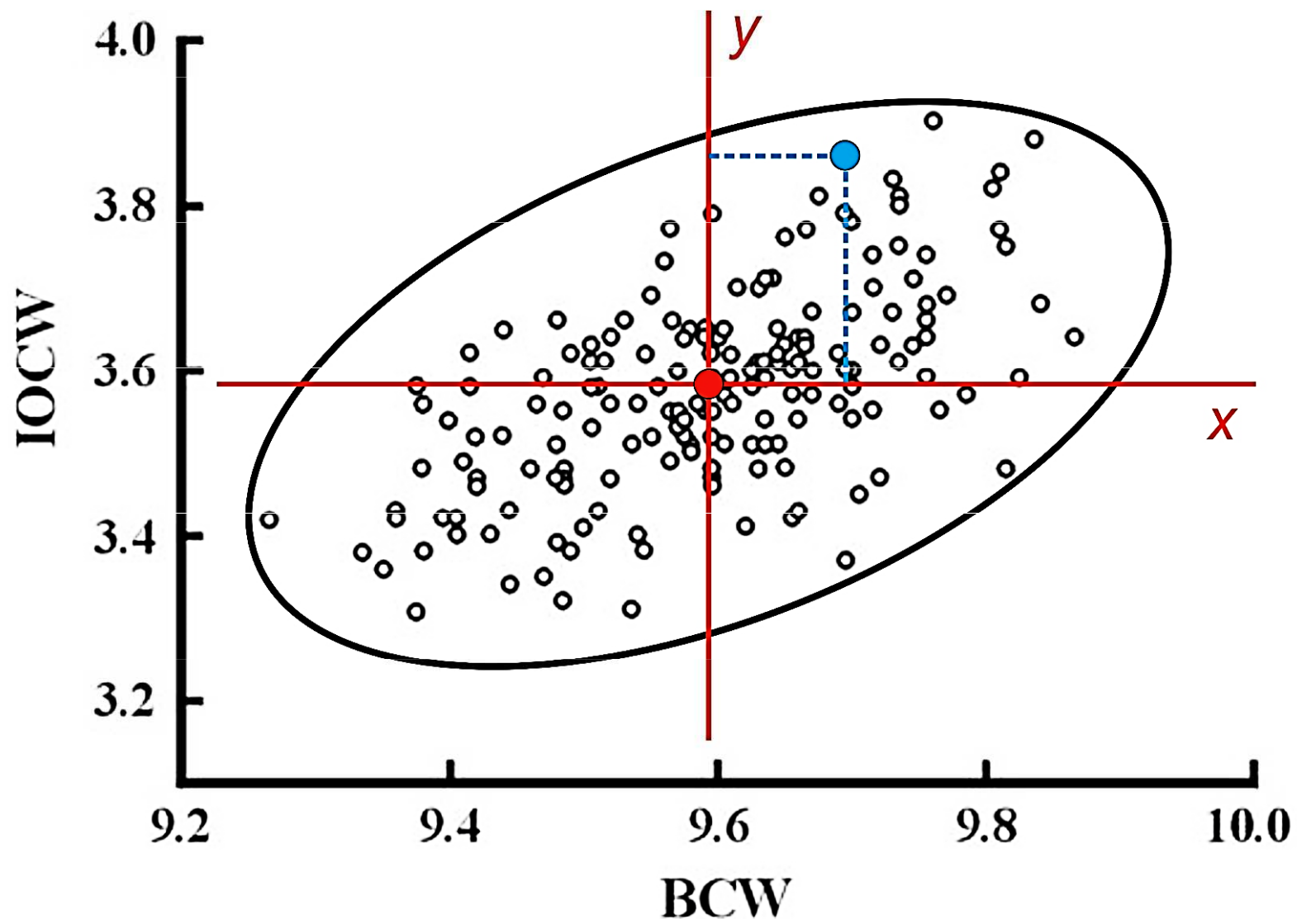
making predictive models

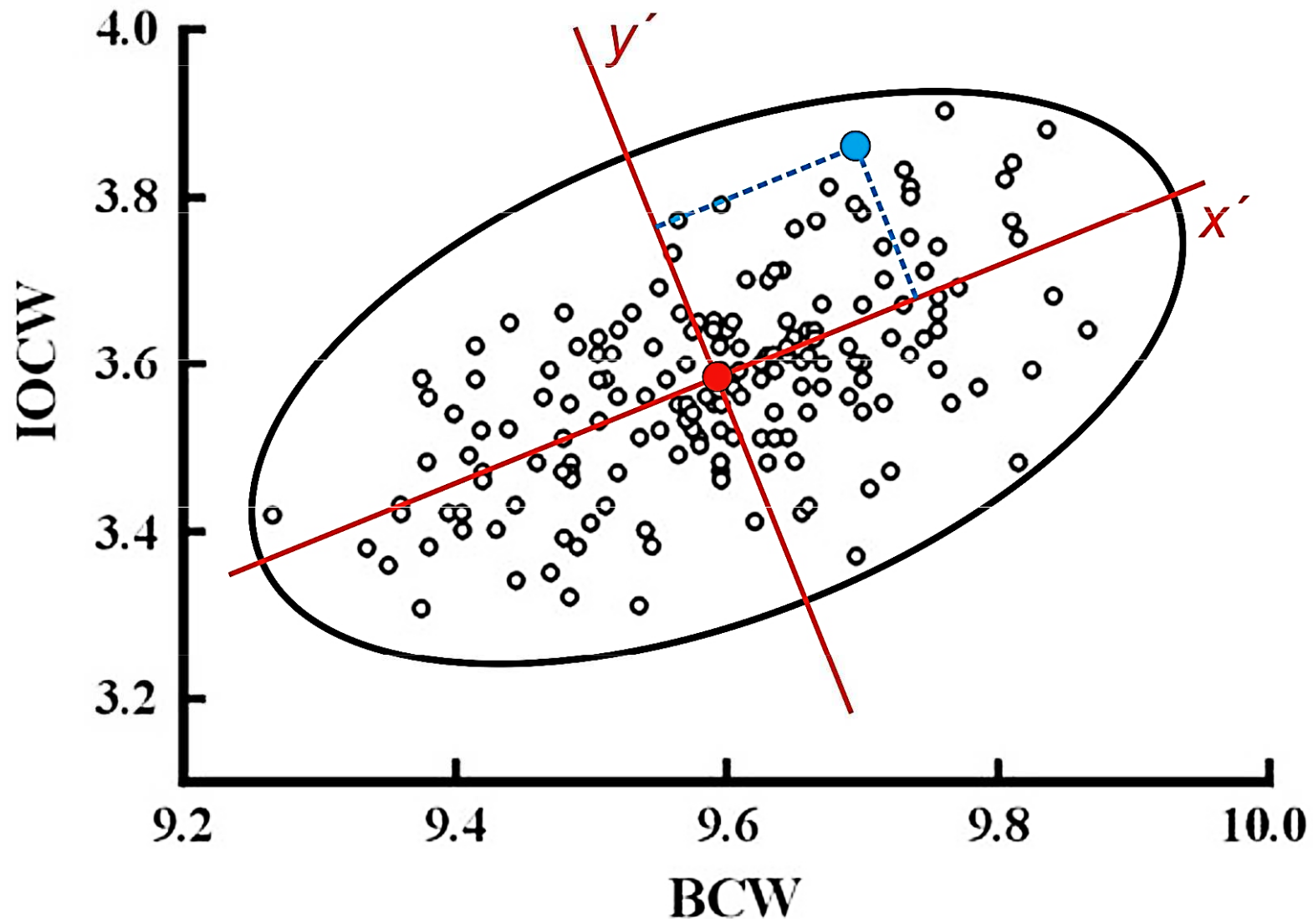












n individuals
 p variables



correlation or covariance matrix

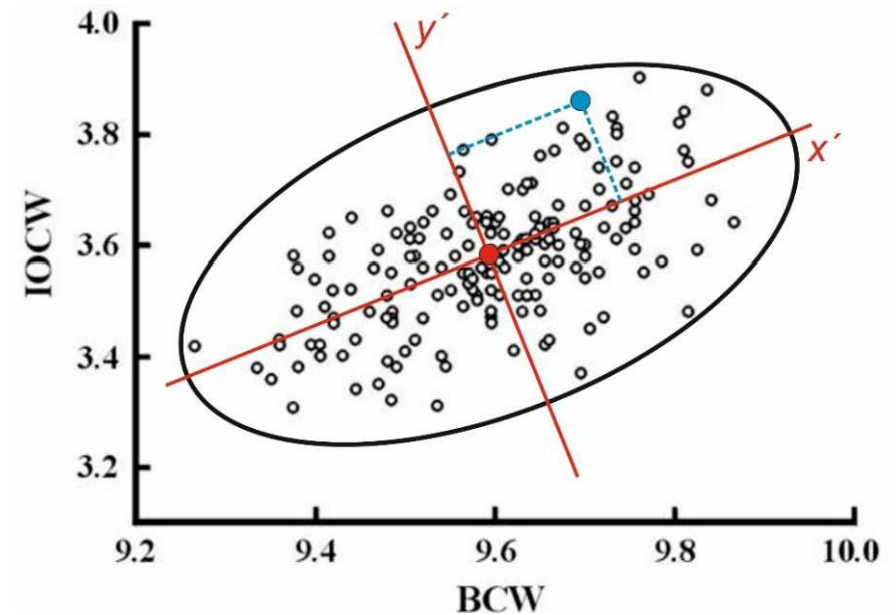


eigenvalue = latent root (latentní kořen)

eigenvector = latent vector (latentní vektor)

PC1: vysvětluje největší podíl variability

PC2: druhý největší podíl variability atd.



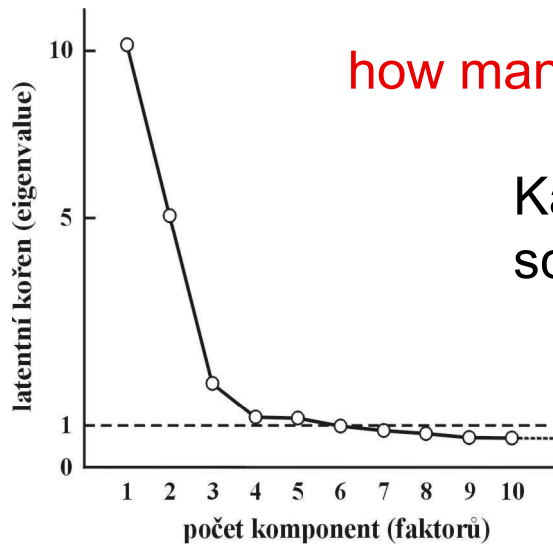
size vector

var	PC1	PC2	PC3	var	PC1	PC2	PC3
URW	0,3126	-0,3902	-0,0933	LN	0,4438	-0,6583	0,2404
ZPW	0,4818	0,0162	-0,1758	LAL	0,1440	0,1699	-0,6798
CBL	0,5590	-0,6997	0,2391	LM1L	0,6450	0,5211	-0,0603
OL	0,3831	-0,6762	0,3076	LM1W	0,7188	0,3311	0,0261
RL	0,4454	-0,6119	0,0486	LM2L	0,5420	0,4698	0,2232
IOCW	0,5883	-0,1422	-0,5204	LM2W	0,7936	0,2996	0,1126
BCW	0,5758	-0,2945	-0,5128	LM3L	0,3572	0,4141	0,2567
ZW	0,6409	-0,6029	-0,1885	LM3W	0,7147	0,3775	0,0767
BCH	0,3083	-0,3053	-0,3544	LML	0,7815	0,4879	0,1092
LFI	0,3956	-0,5874	0,2949	UM1L	0,6716	0,4054	-0,0434
I1W	0,4973	-0,6556	0,0238	UM1W	0,7315	0,2869	-0,0957
FP	0,5640	-0,2443	-0,1547	UM2L	0,4596	0,4304	0,2878
FPT	0,3218	-0,1715	-0,3899	UM2W	0,8023	0,2360	-0,0642
PM	0,5434	-0,2574	-0,0485	UM3L	0,6621	-0,0026	0,1481
LD	0,3898	-0,8068	0,2282	UM3W	0,6975	0,2560	0,1133
				UML	0,7739	0,4258	0,1353
eigen	10,137	6,059	2,028				
% var	32,701	19,544	6,542				
kum %	32,701	52,245	58,793				

loadings (zátěže)

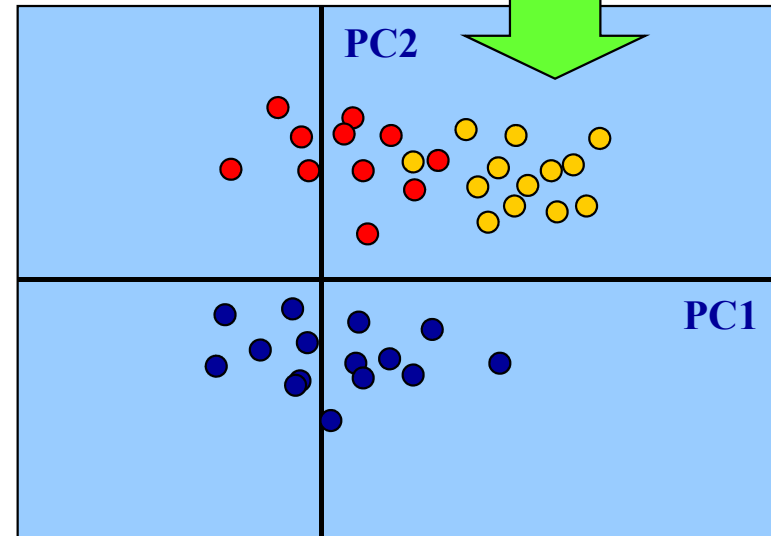
Pop.	PC1	PC2	PC3
1	-0,4107	0,3434	-0,7138
1	-0,3473	0,0747	-0,4280
1	-0,4107	10,2478	-0,6631
1	-0,2454	0,5547	-0,9033
1	-10,1552	0,9977	-10,6618
1	-10,5334	10,3502	10,7457
1	-0,5520	20,1285	10,2274
2	-0,9484	0,0695	-0,4729
2	0,7837	0,7317	0,6073
2	0,0757	-0,8404	0,2365
2	-10,4157	-10,4619	-0,1617
2	-0,8787	-10,6109	10,0195
3	0,5419	-0,9804	0,5811
3	10,4301	0,3881	0,4764
3	0,6310	0,3449	-0,0335
3	0,2657	0,0851	10,2535
3	0,8796	0,6772	0,2254
3	10,7009	-0,2075	0,4952
3	-0,0263	0,3242	0,9853
3	0,9681	10,2721	-0,0680

PC scores



how many components?

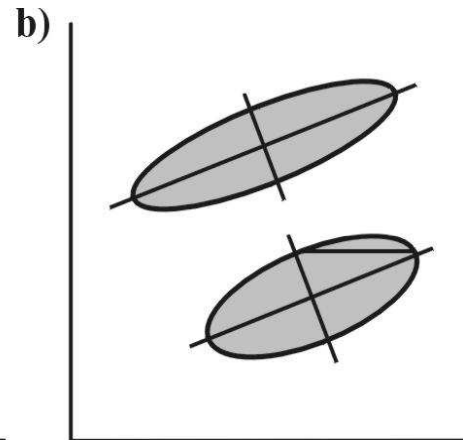
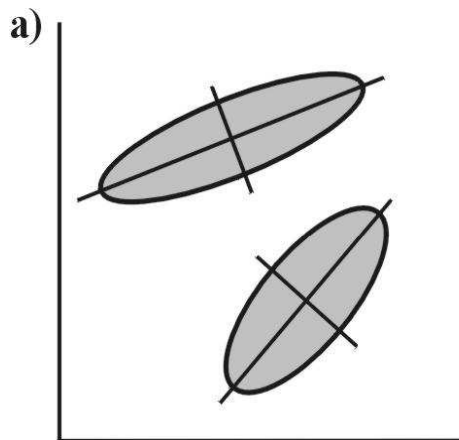
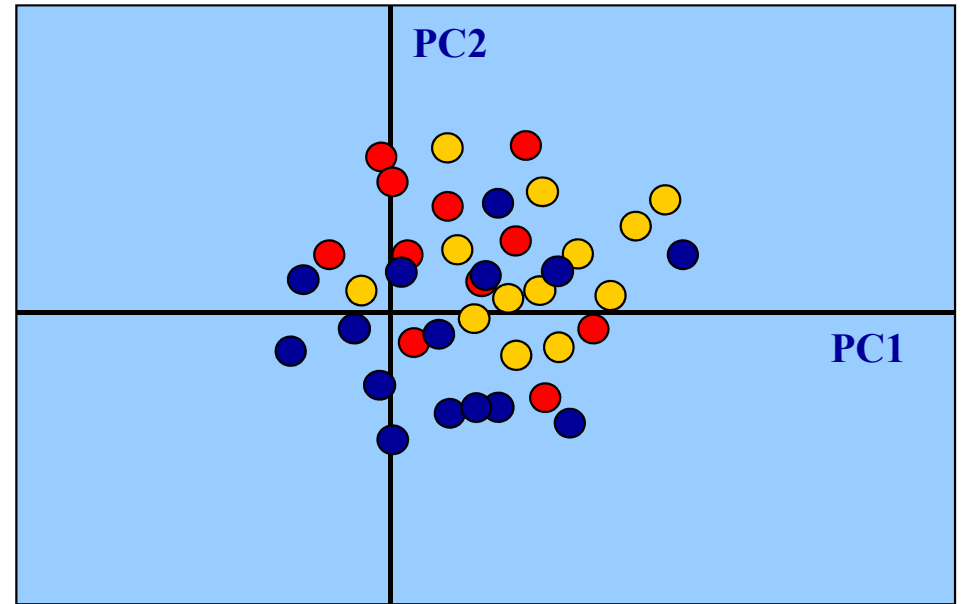
Kaiser's criterion
scree graph



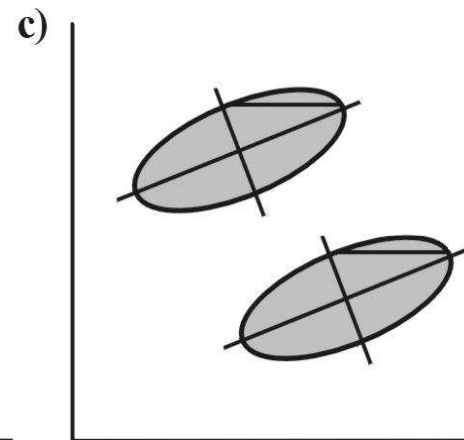
Problem of multiple groups:

CPCA (common PCA)

MGPCA (multiple-group PCA)



CPCA

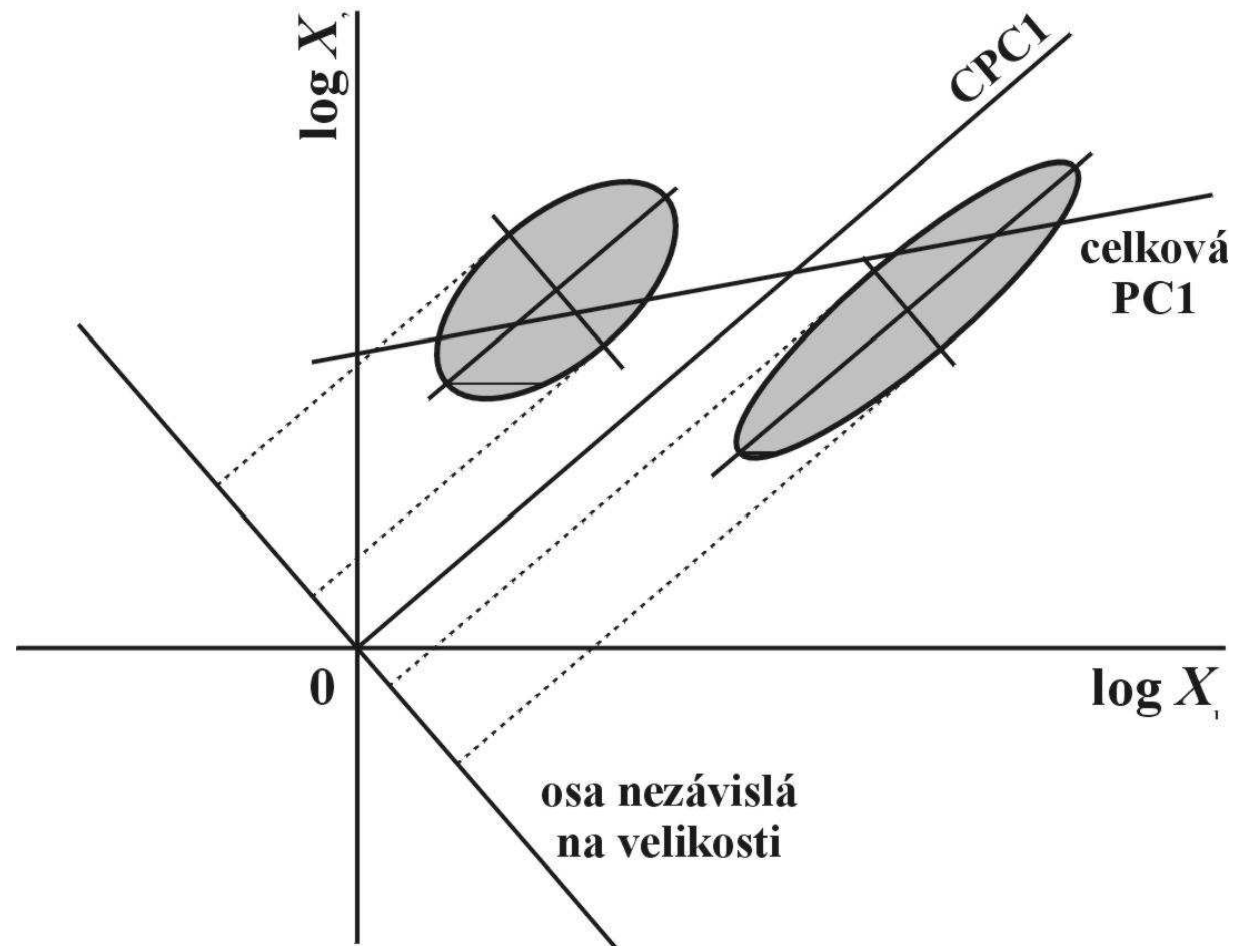


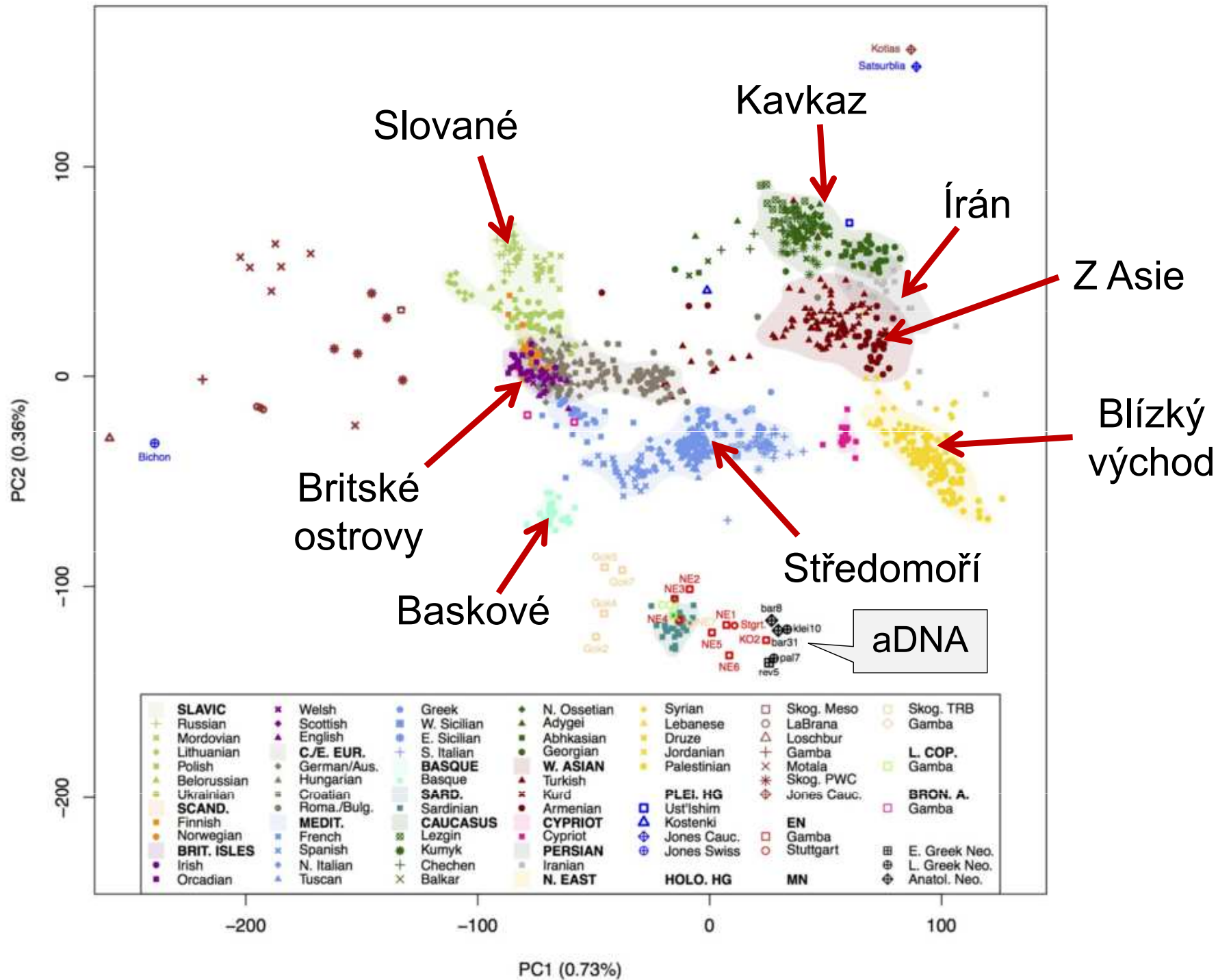
MGPCA

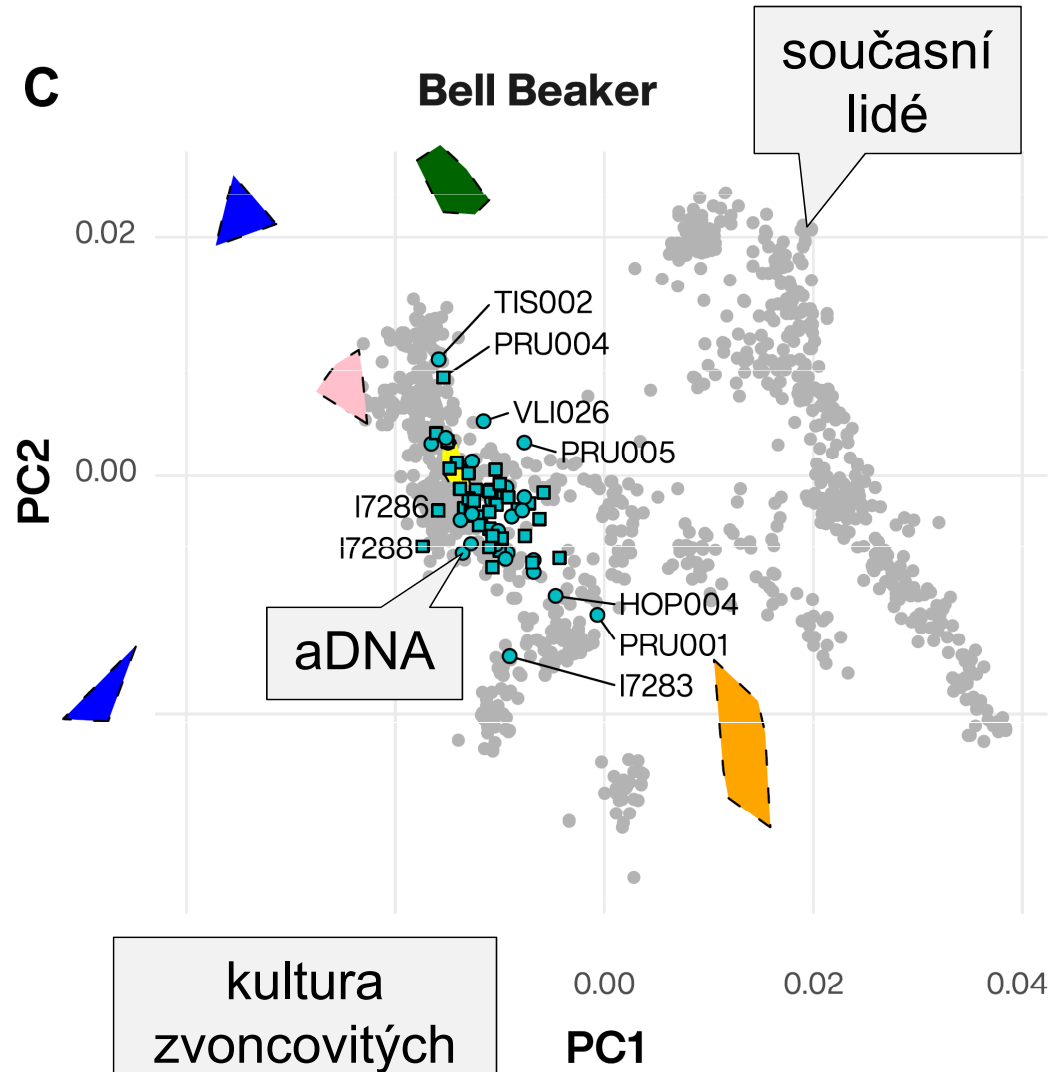
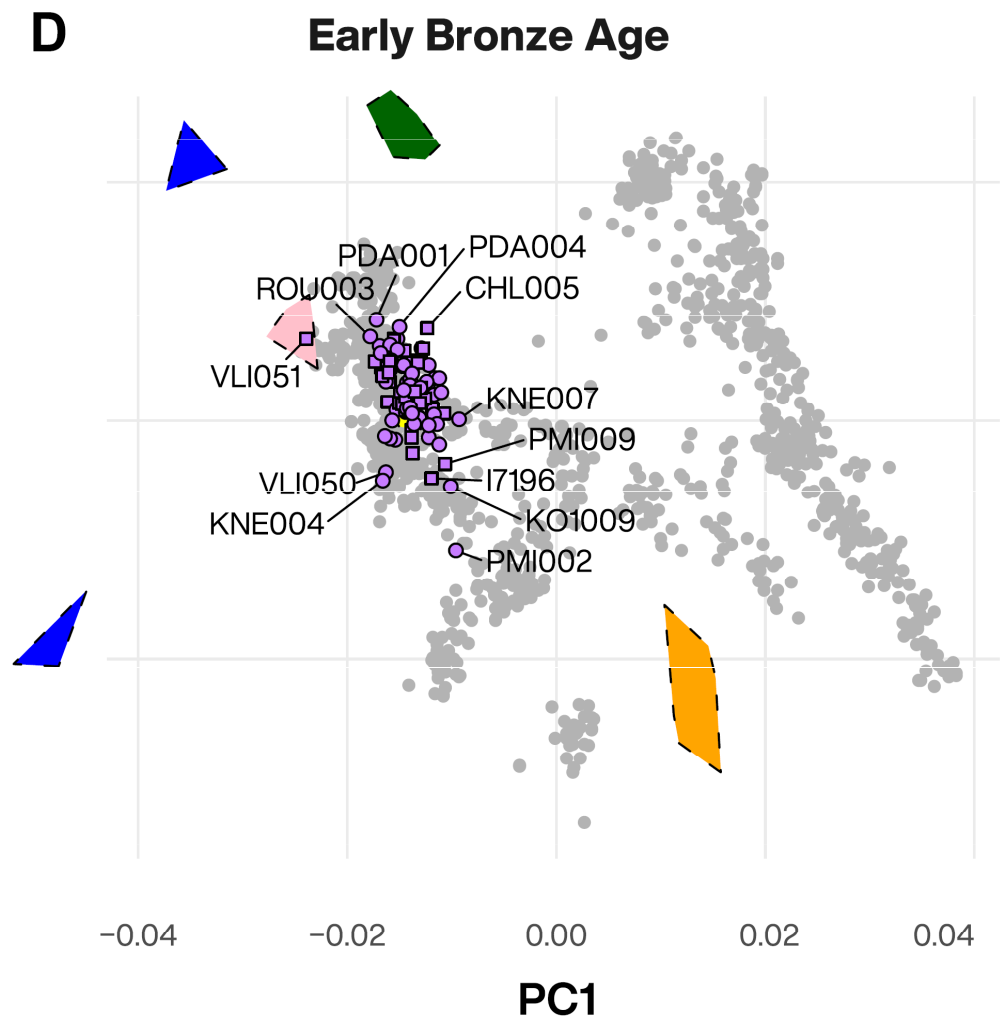
Problem of size:

omitting PC1

Burnaby's adjustment





C**D**

■ Bell Beaker
 ■ Early Bronze Age
 ⋮
 □ Male
 ○ Female

Multidimensional scaling (MDS):

not only correlation/covariation matrix, all types, 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

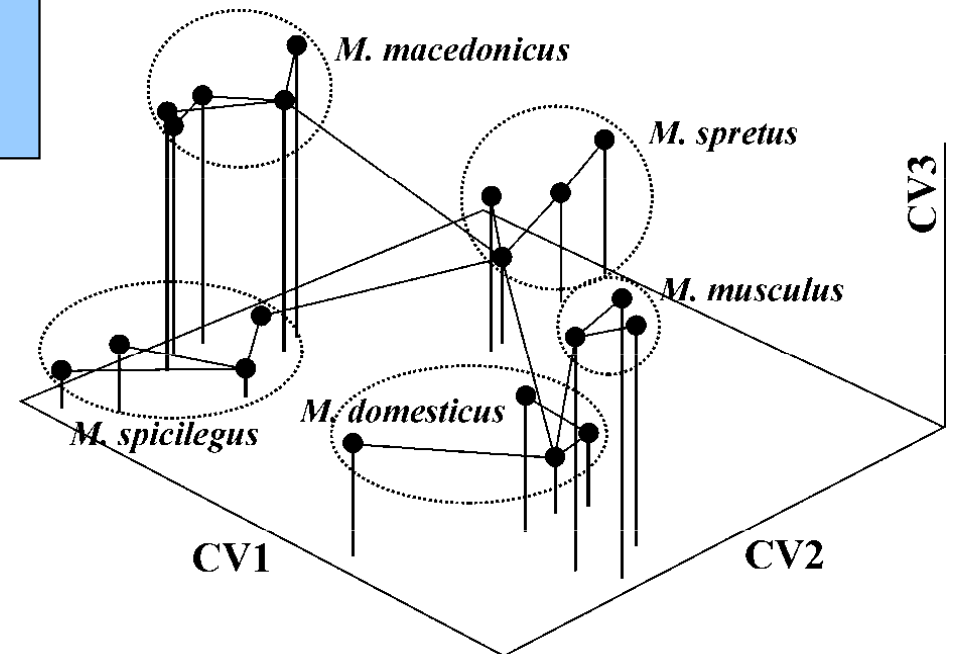
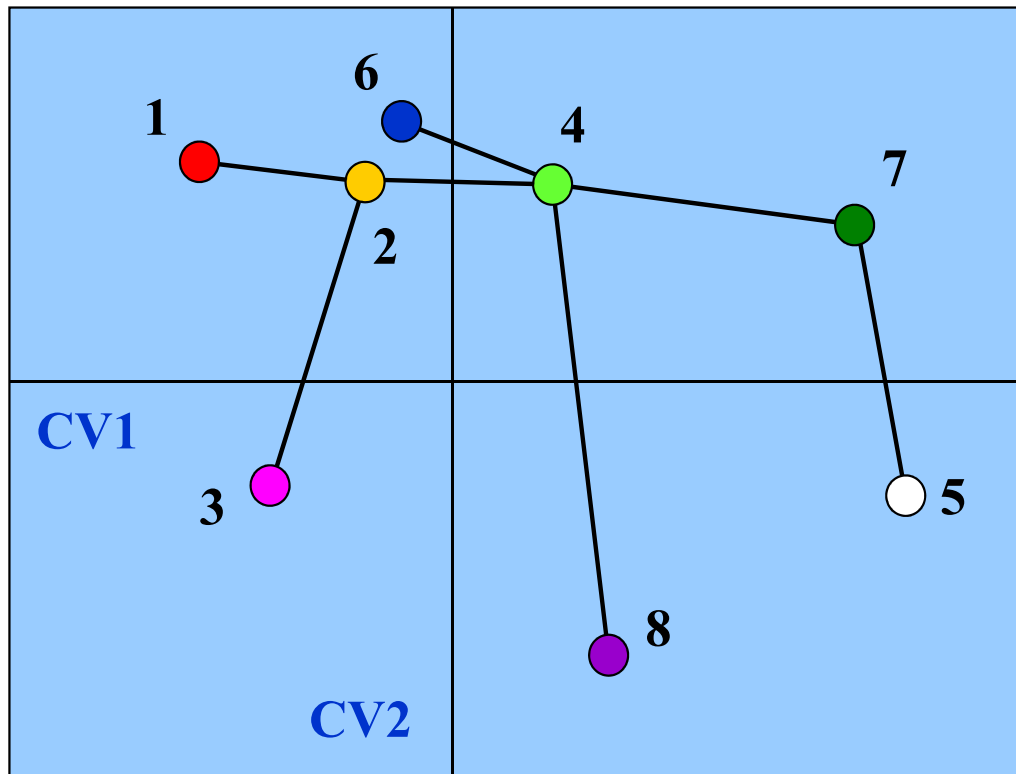
MANOVA (Wilk's Lambda, Pillai's trace)

Hotelling T^2 test

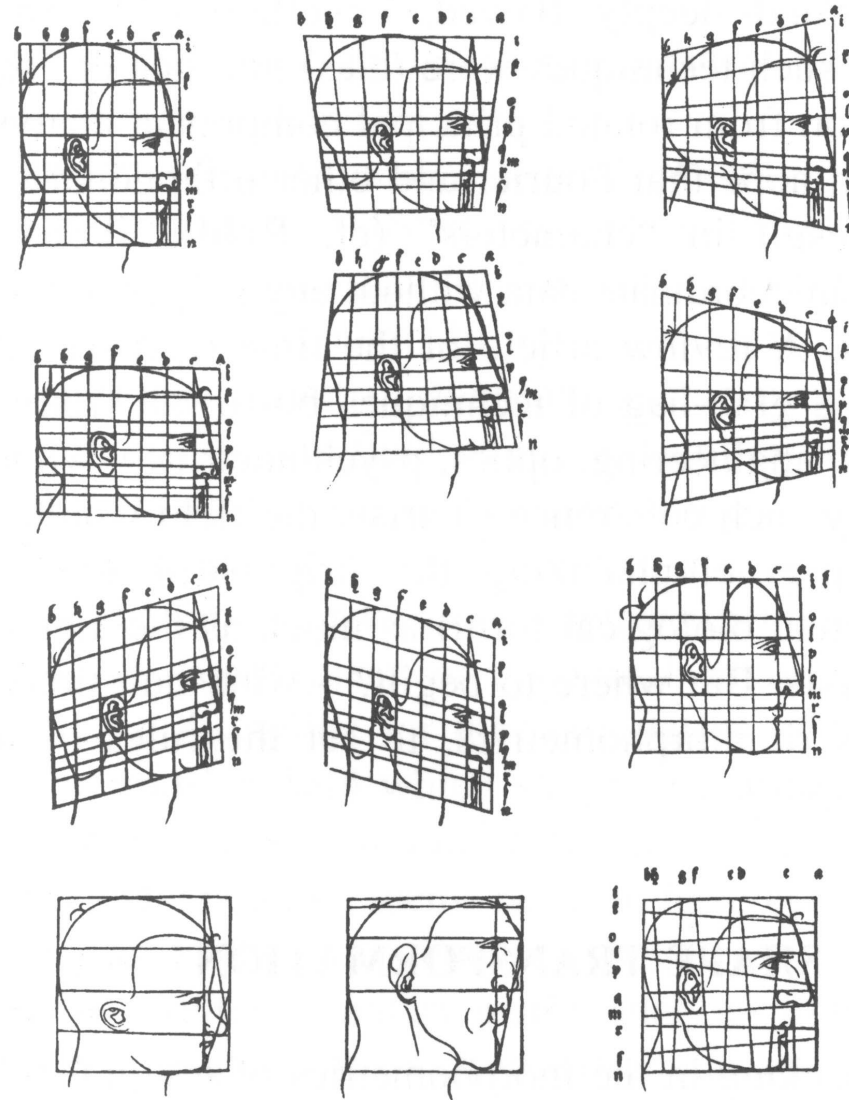
stepwise DFA

Cluster analysis

Minimum spanning tree (MST)



Geometric morphometrics

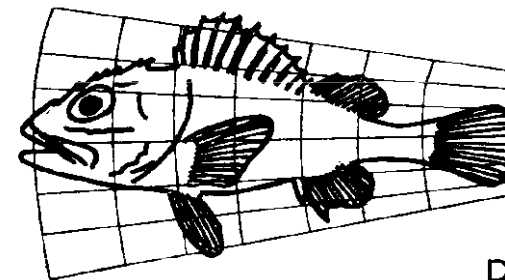
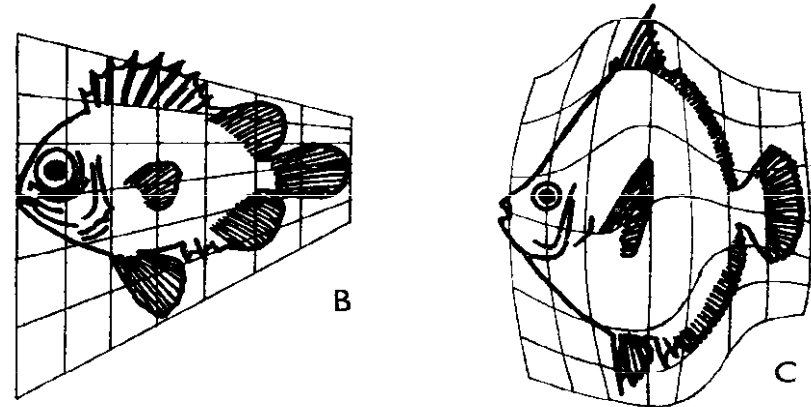
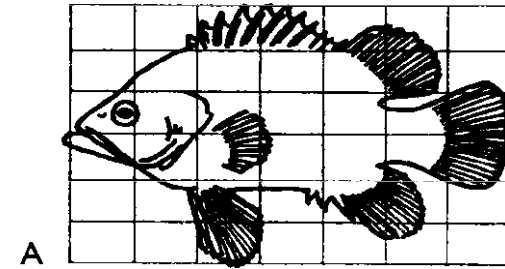


A. Dürer (1524): Vier Bücher von Menlicher Proportion.

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...

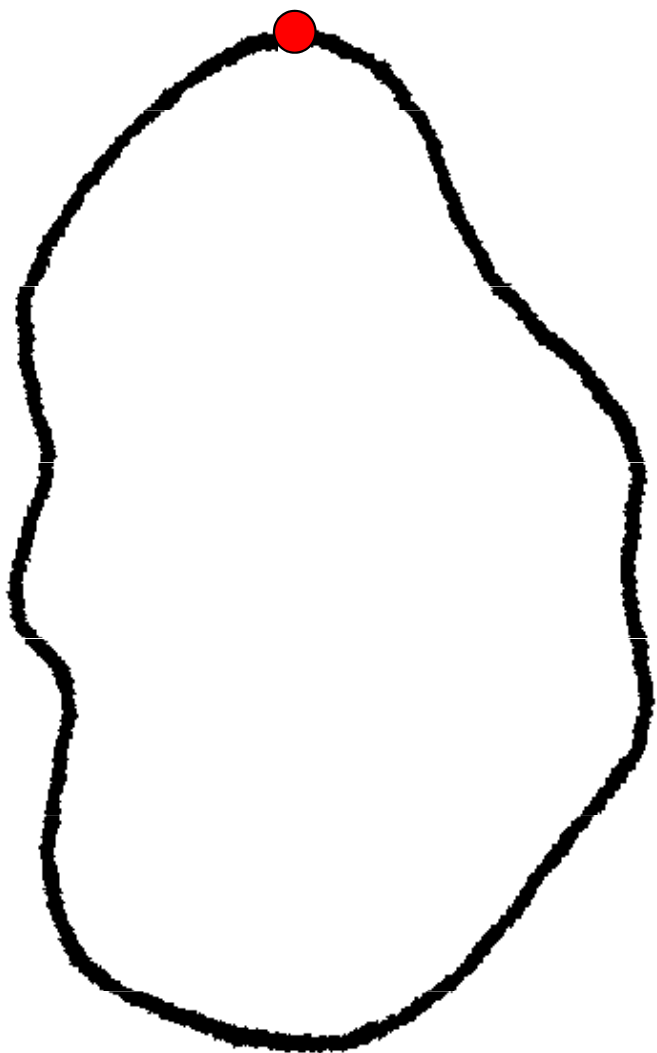
linear measurements, weights, angles, surfaces...

PCA, DFA, CVA, FA, PCoA, cluster a.

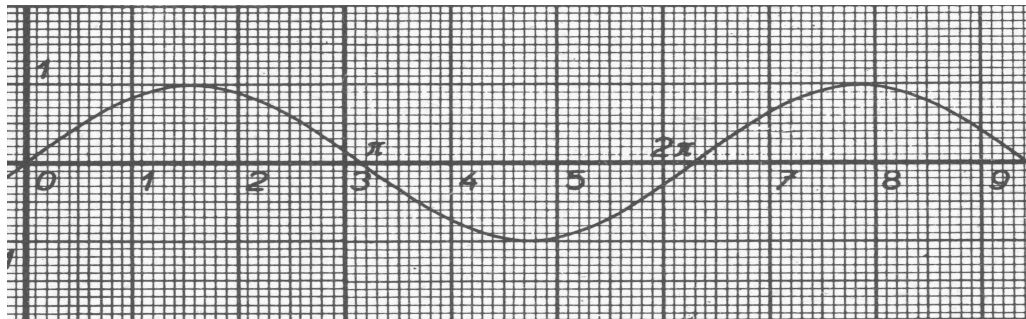
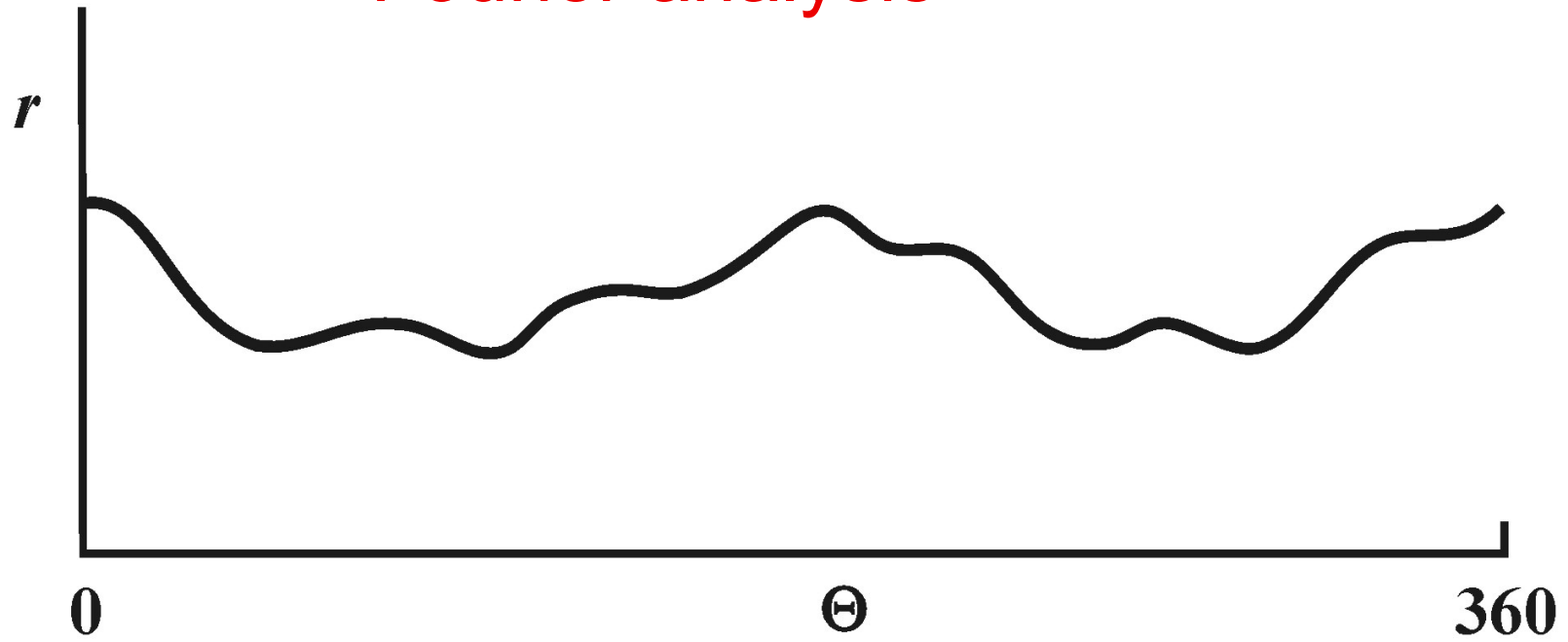
Absence of any information on shape (morphometrics)!

Geometric morphometrics I.

Analysis of closed curves

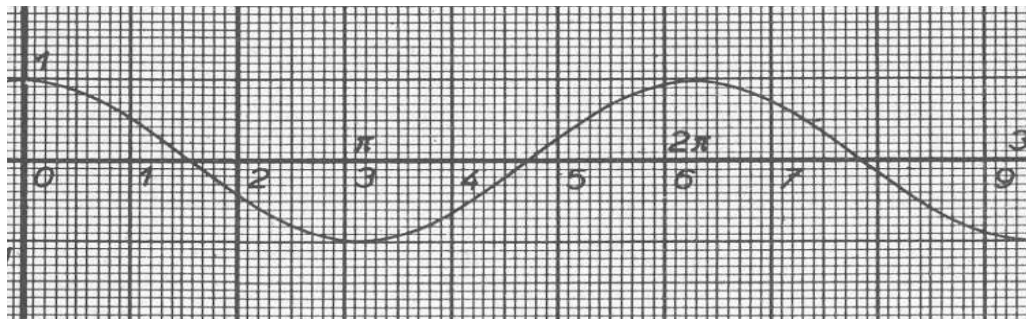


Fourier analysis



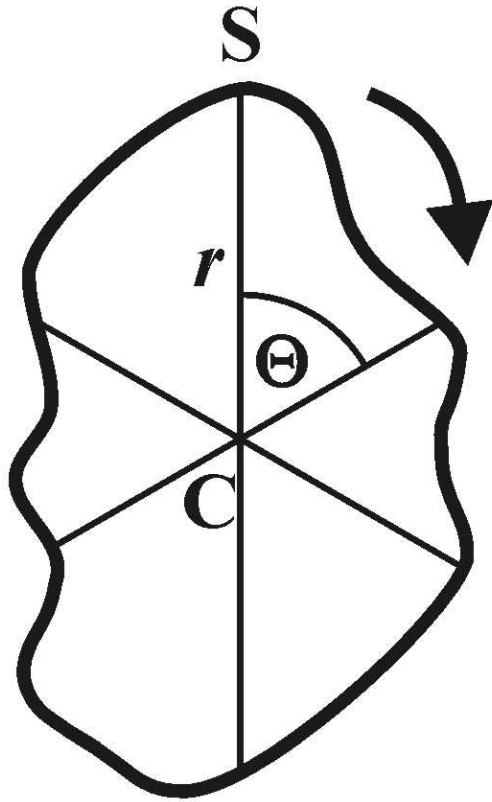
sin

harmonics, coefficients



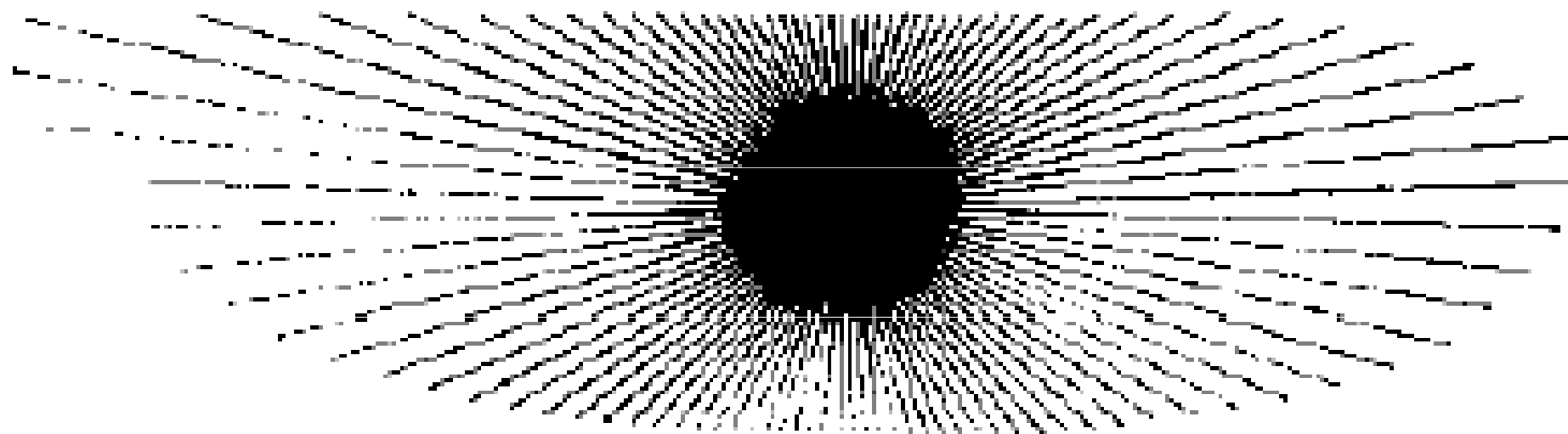
COS

Traditional Fourier a.

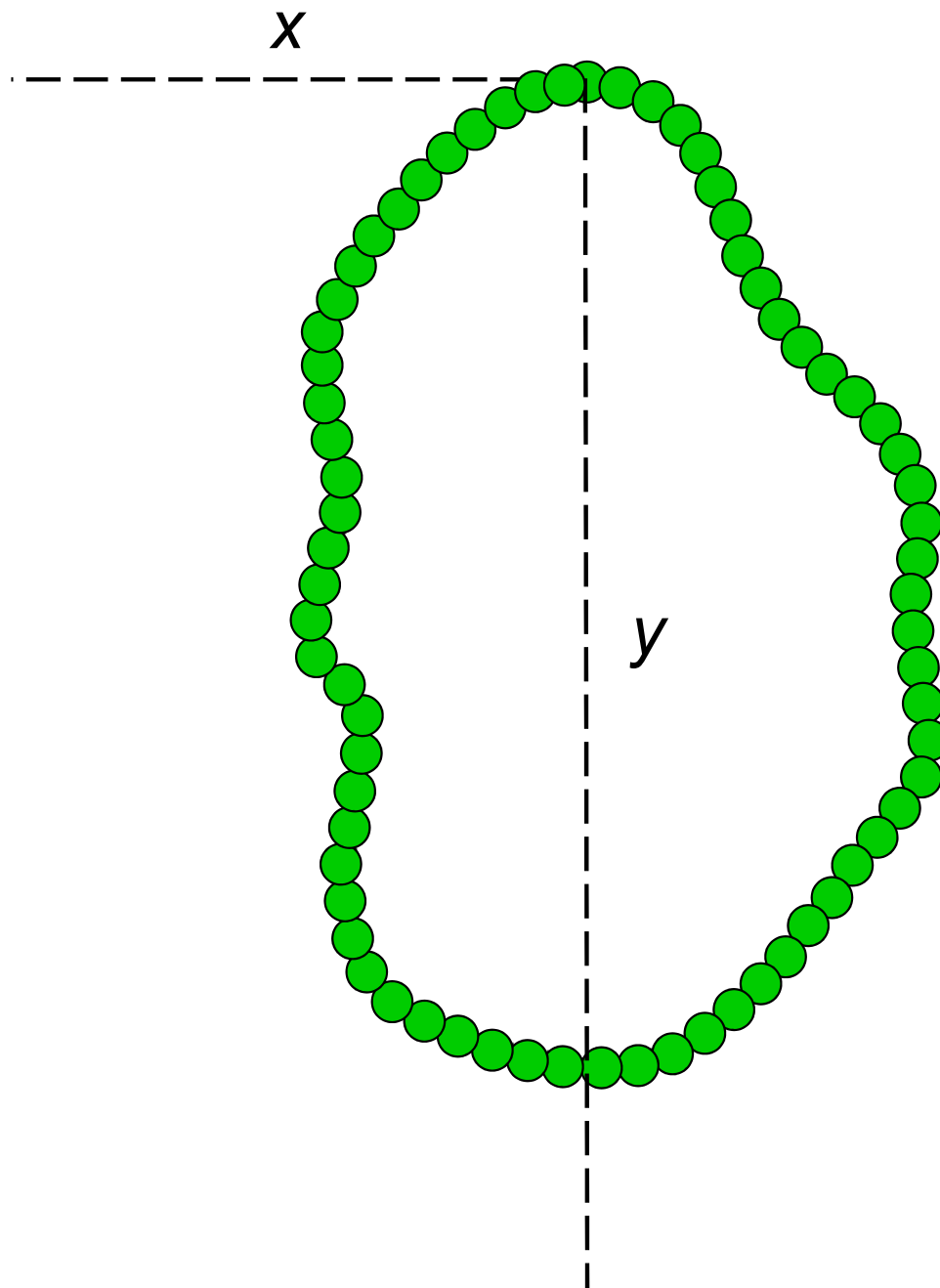


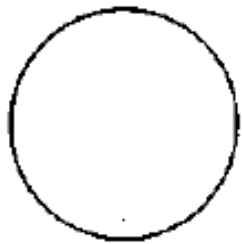
$$r_{\theta} = a_0 + \sum_{i=1}^k (a_i \cos_i \theta + b_i \sin_i \theta)$$

$$a_0 = \sqrt{\frac{1}{p} \sum_{i=1}^p \theta_i} \quad a_i = \sqrt{\frac{2}{p} \sum_{i=1}^p \theta_i \cos_i \theta_i} \quad b_i = \sqrt{\frac{2}{p} \sum_{i=1}^p \theta_i \sin_i \theta_i}$$



Elliptic Fourier a.





•



•



•



•



|

o

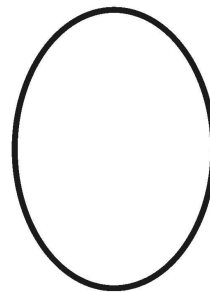
•

•

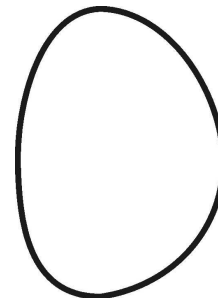
•

•

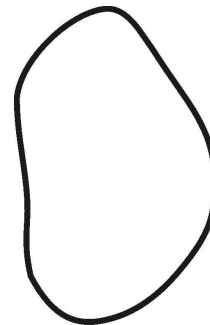
|



n=1



n=2



n=5



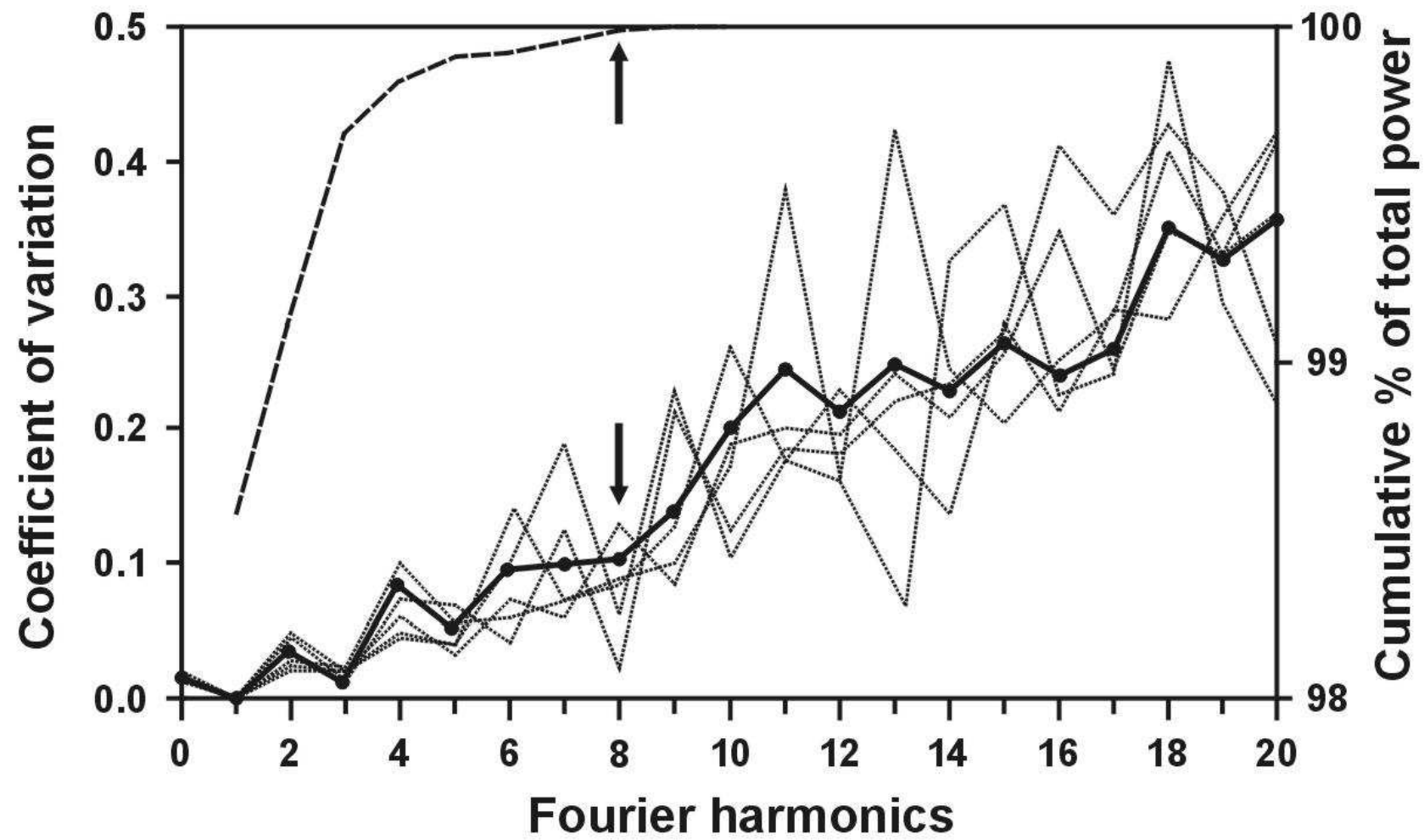
n=8



n=10



n=20



Geometric morphometrics II.

Analysis of landmarks

landmarks

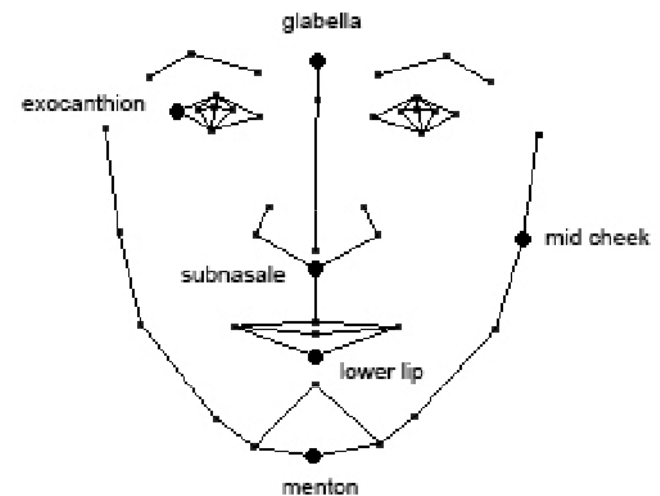
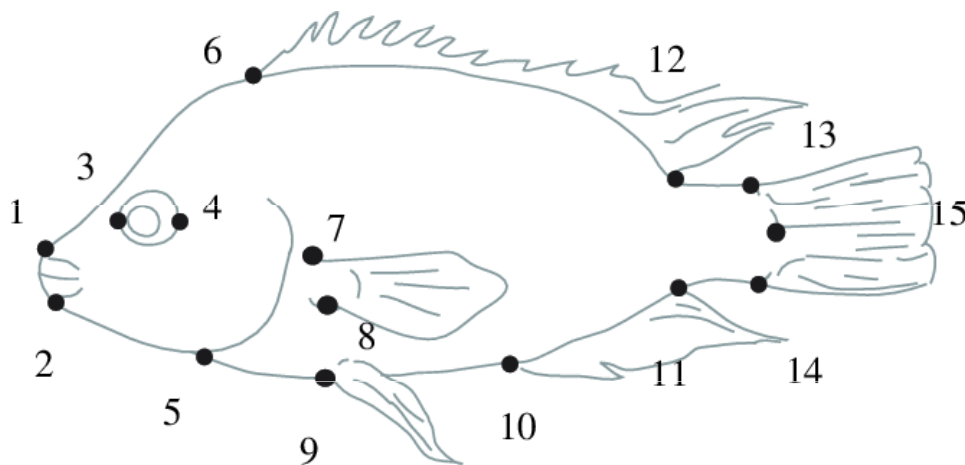
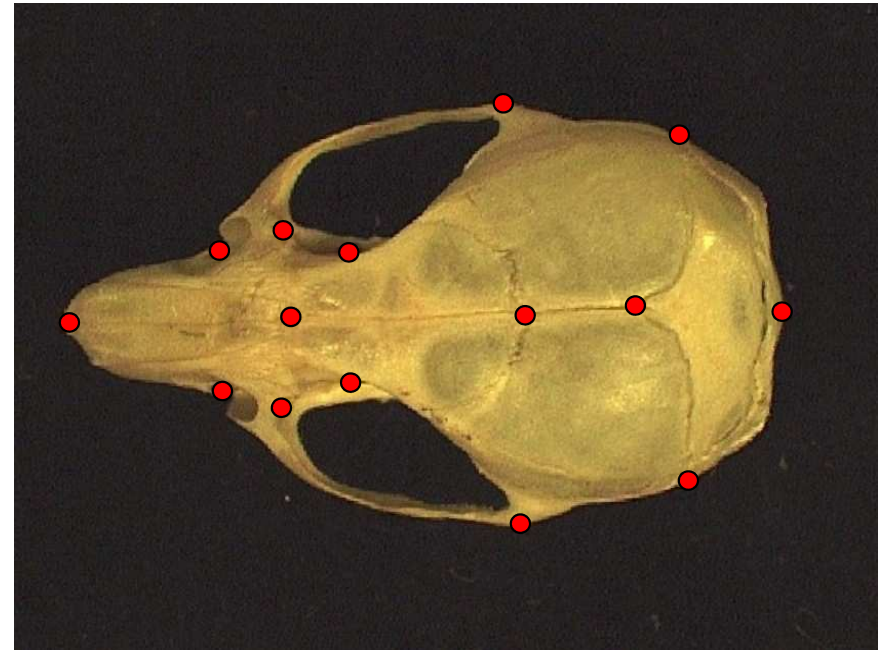
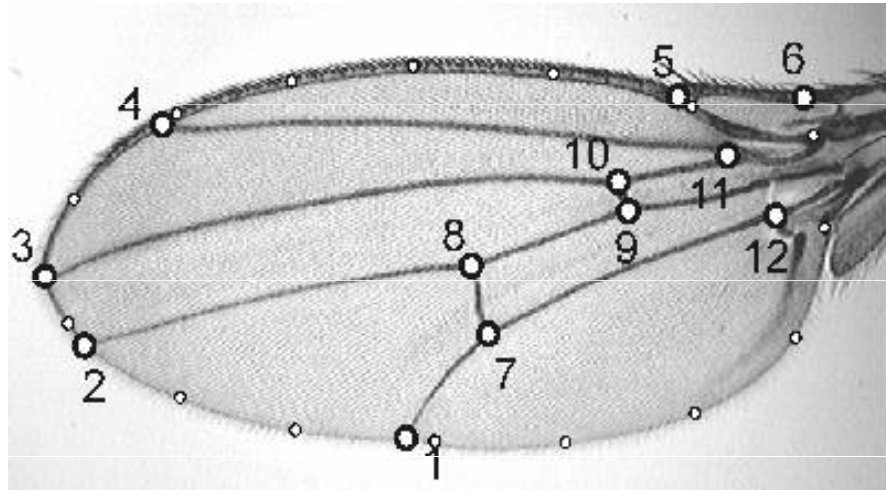
quantification of shape using shape coordinates, distinction between different shape components

information on shape maintained during the whole analysis

size standardization and ability of independent work with size vector

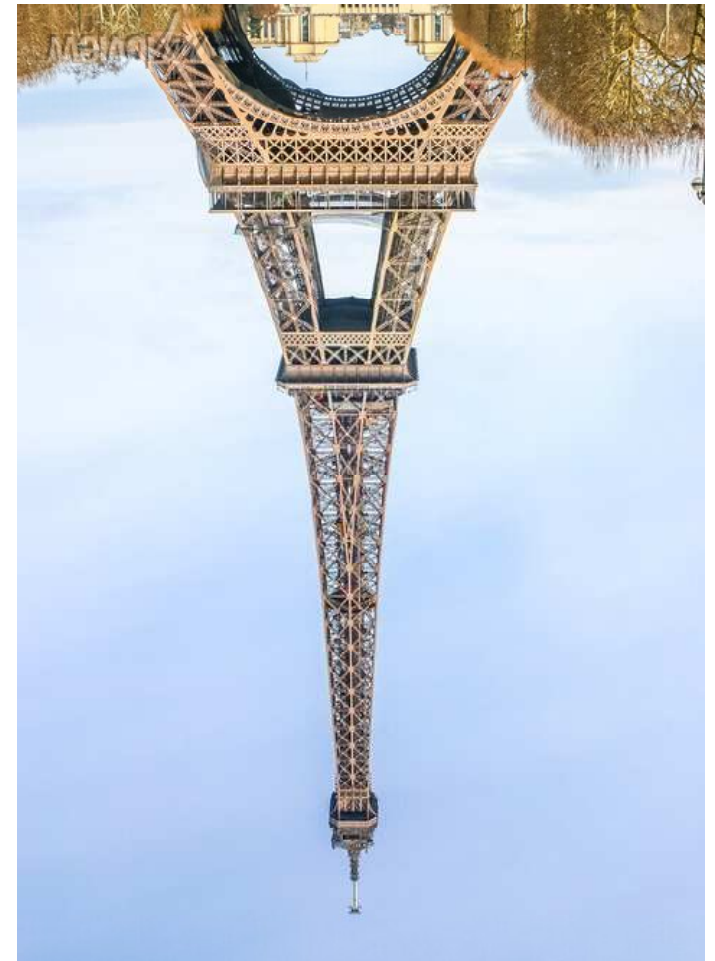
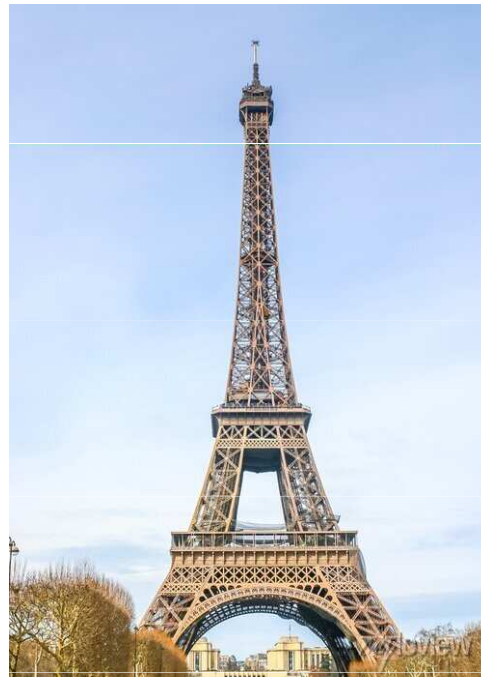
data processing with traditional morphometric methodology

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



Procrustes superposition = GLS (Generalized Least Squares)

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



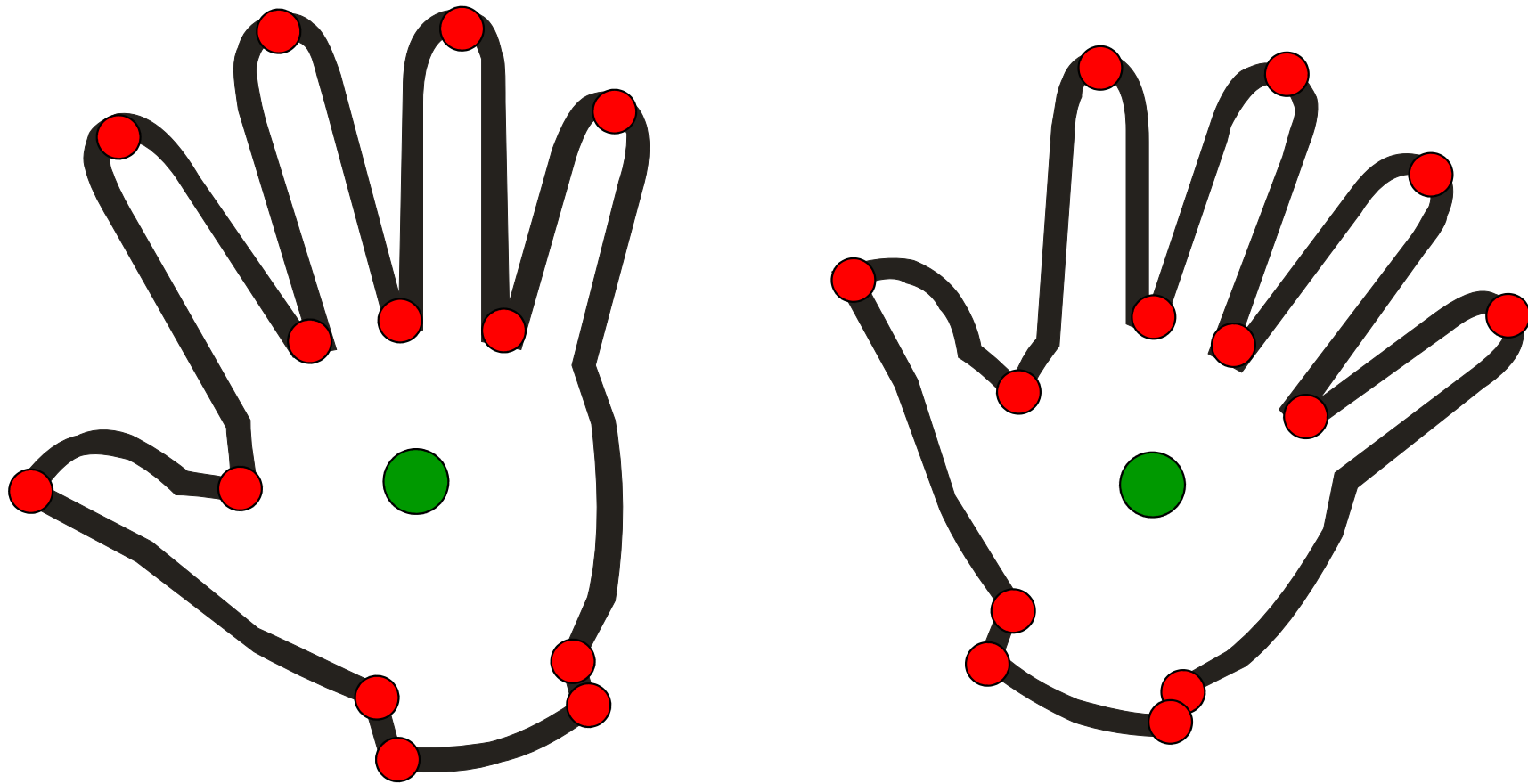
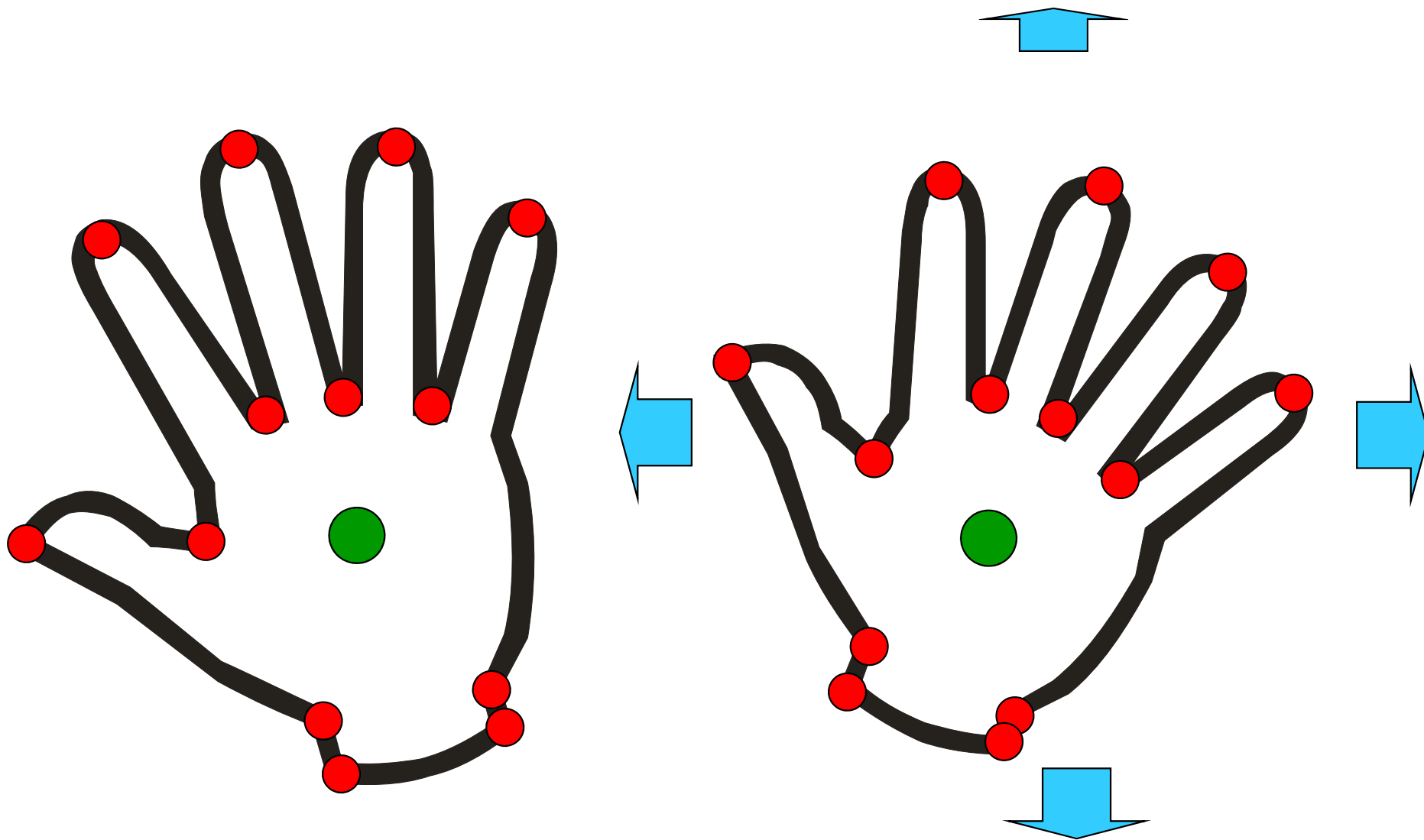
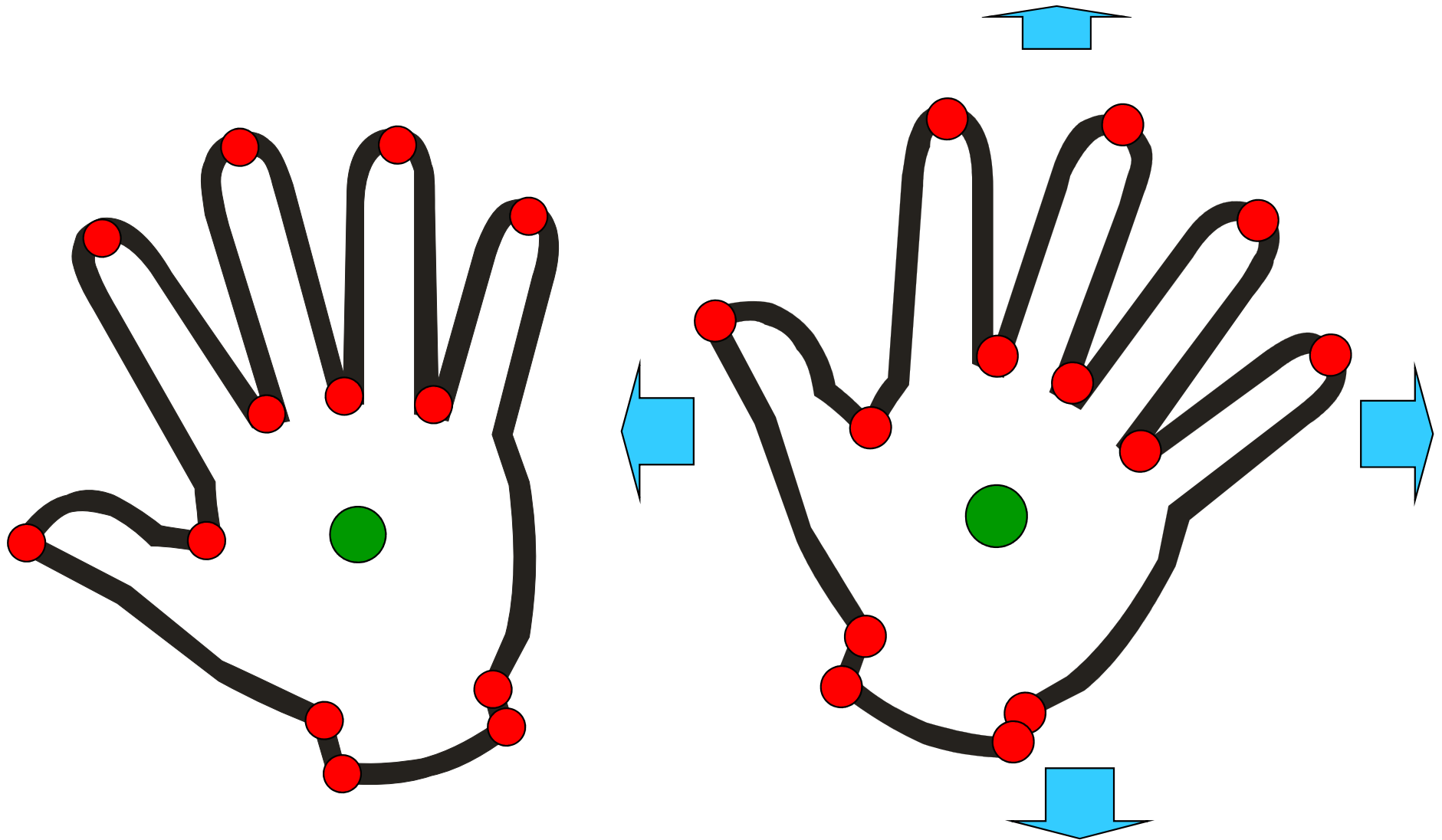


IMAGE SPACE: p – dimensions, k – landmarks; $n = pk$

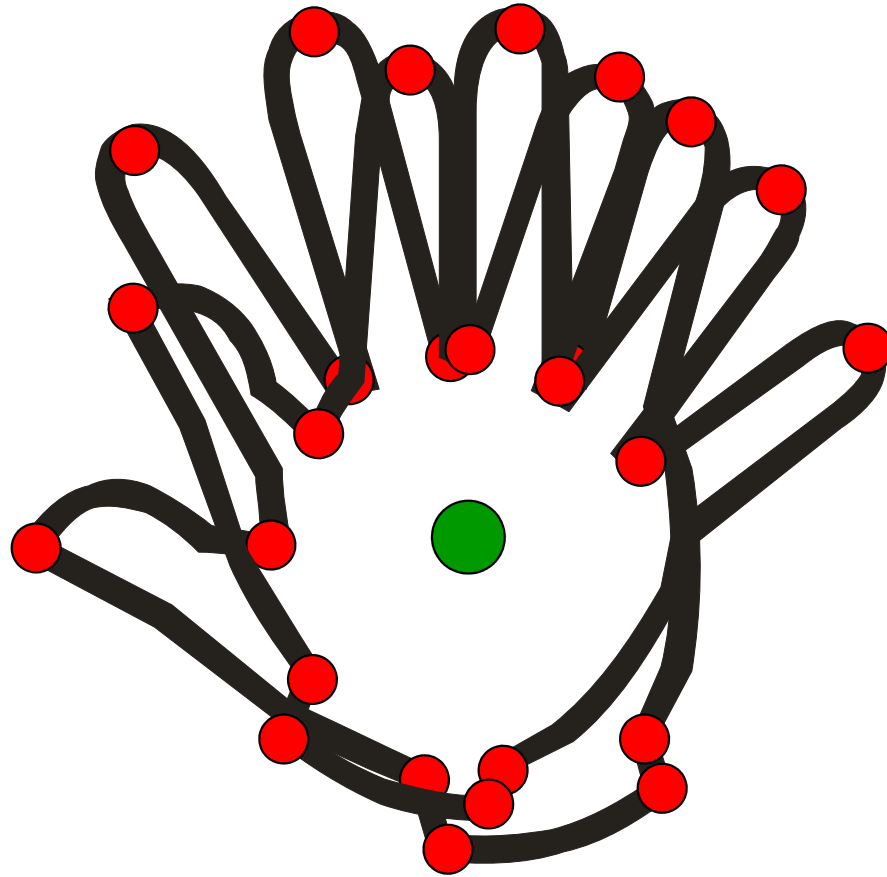
Procrustes superposition 1



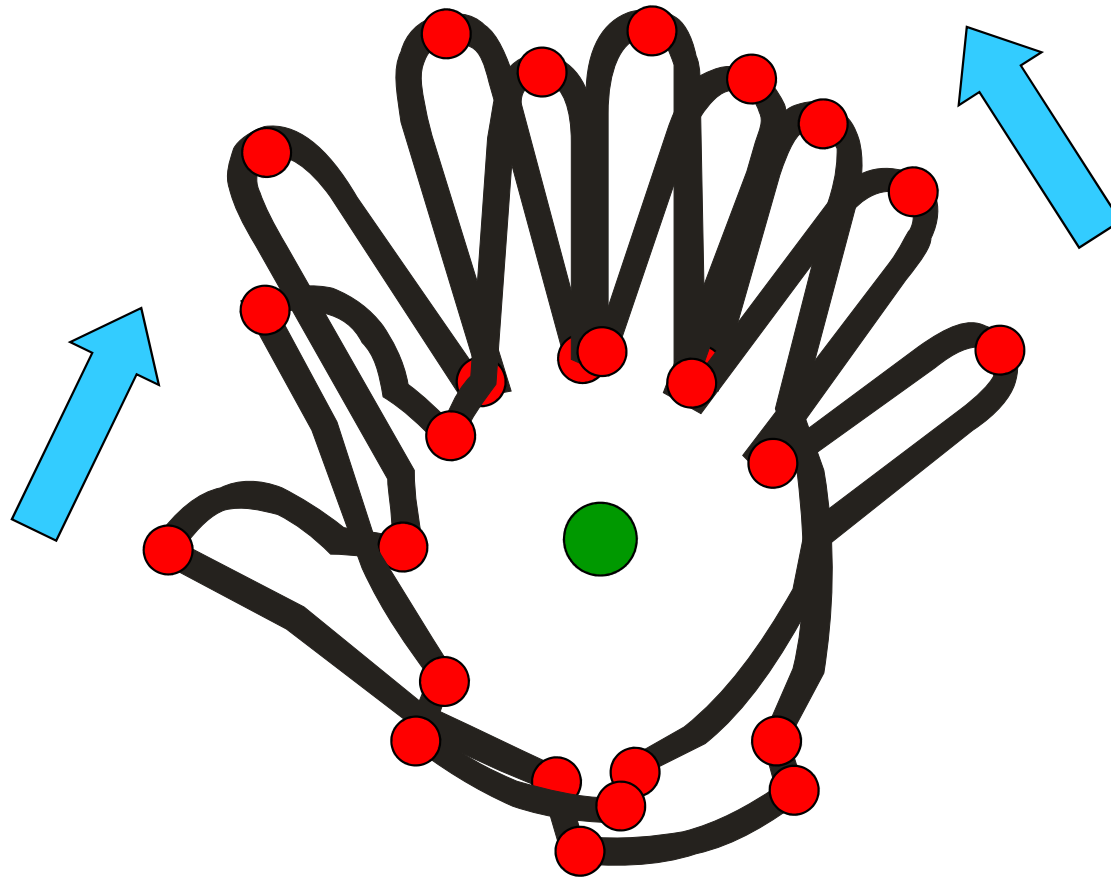
Procrustes superposition 1



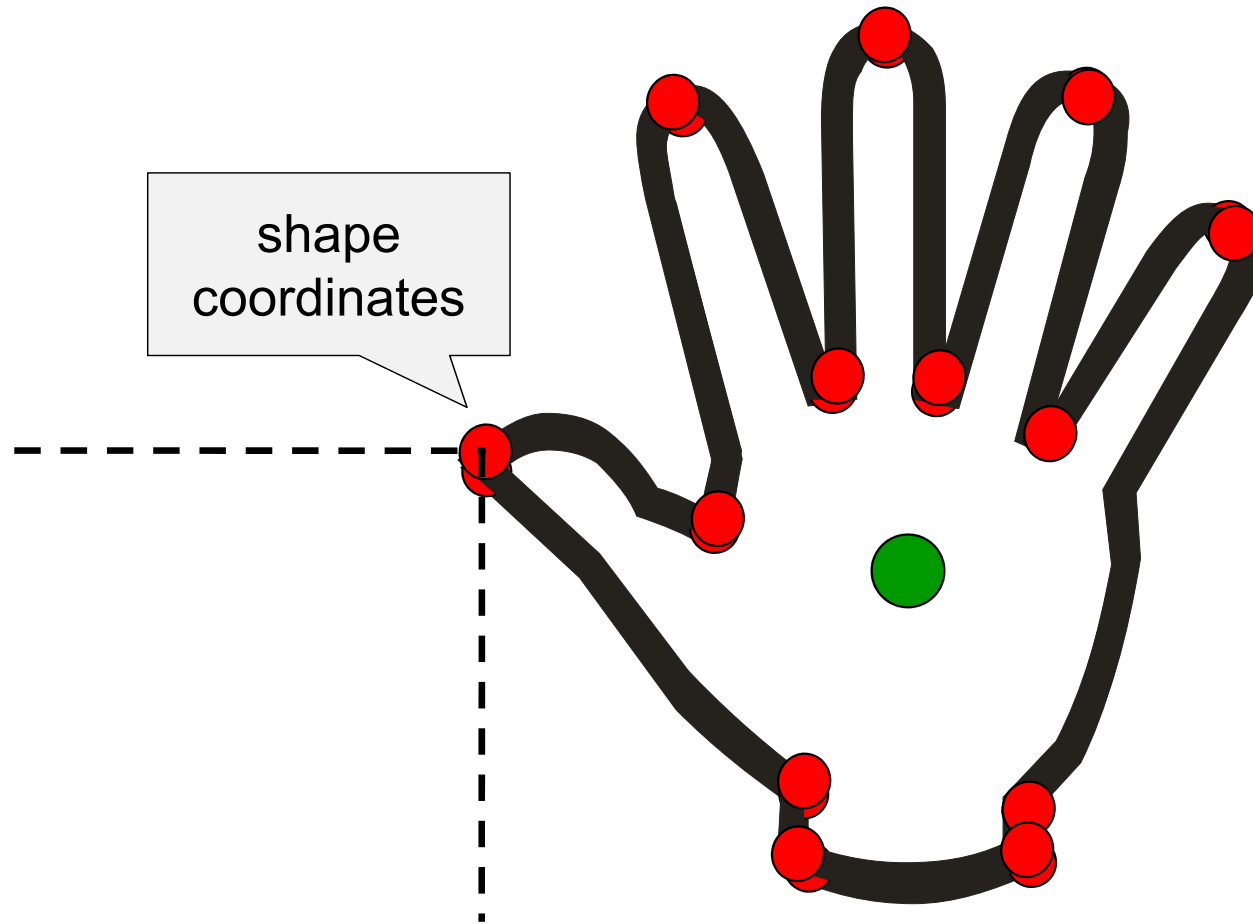
Procrustes superposition 2



Procrustes superposition 3



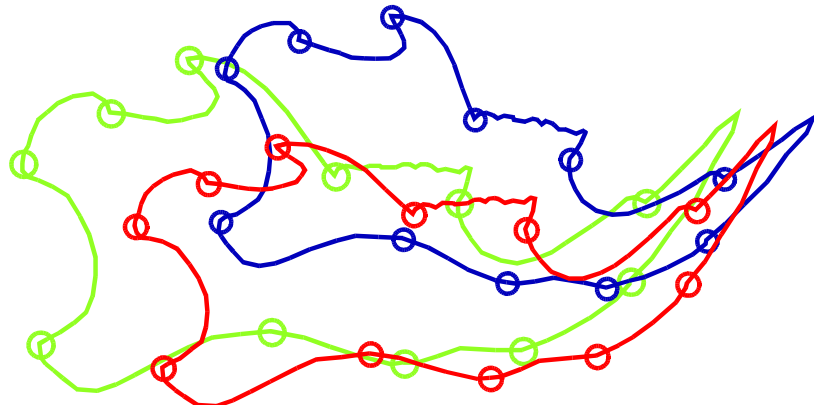
Procrustes superposition 3



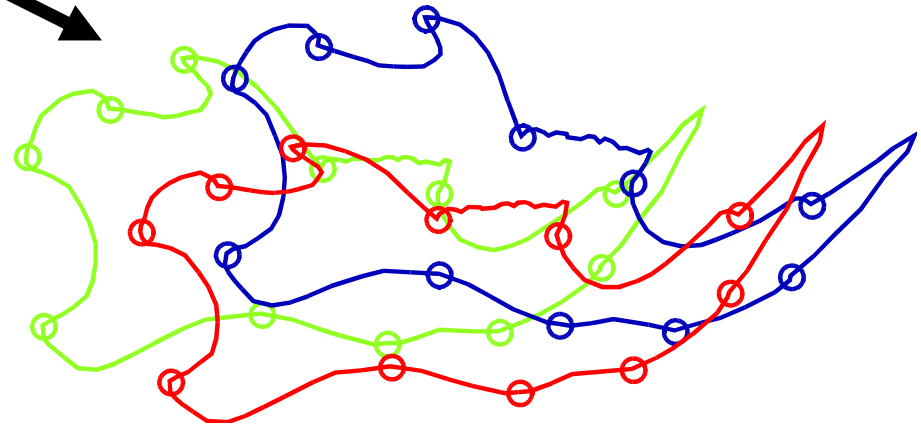
SHAPE SPACE: $n = pk - k - k(k-1)/2 - 1$

Extracting shape information: Procrustes superposition

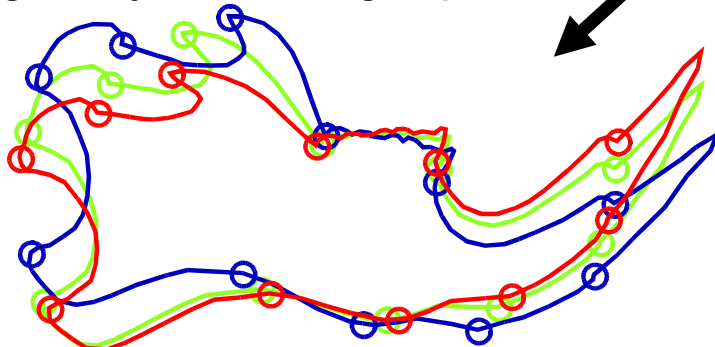
Original landmark configurations



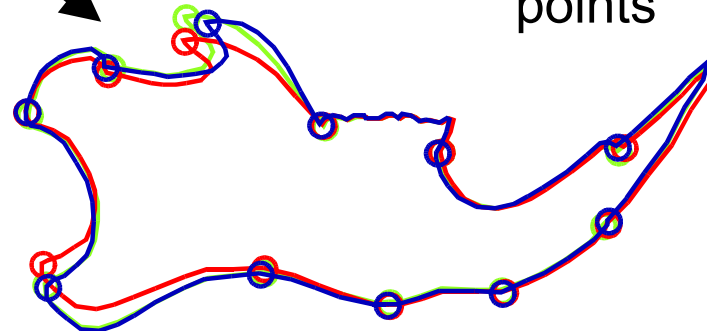
1. Change scale so that all configurations have the same size

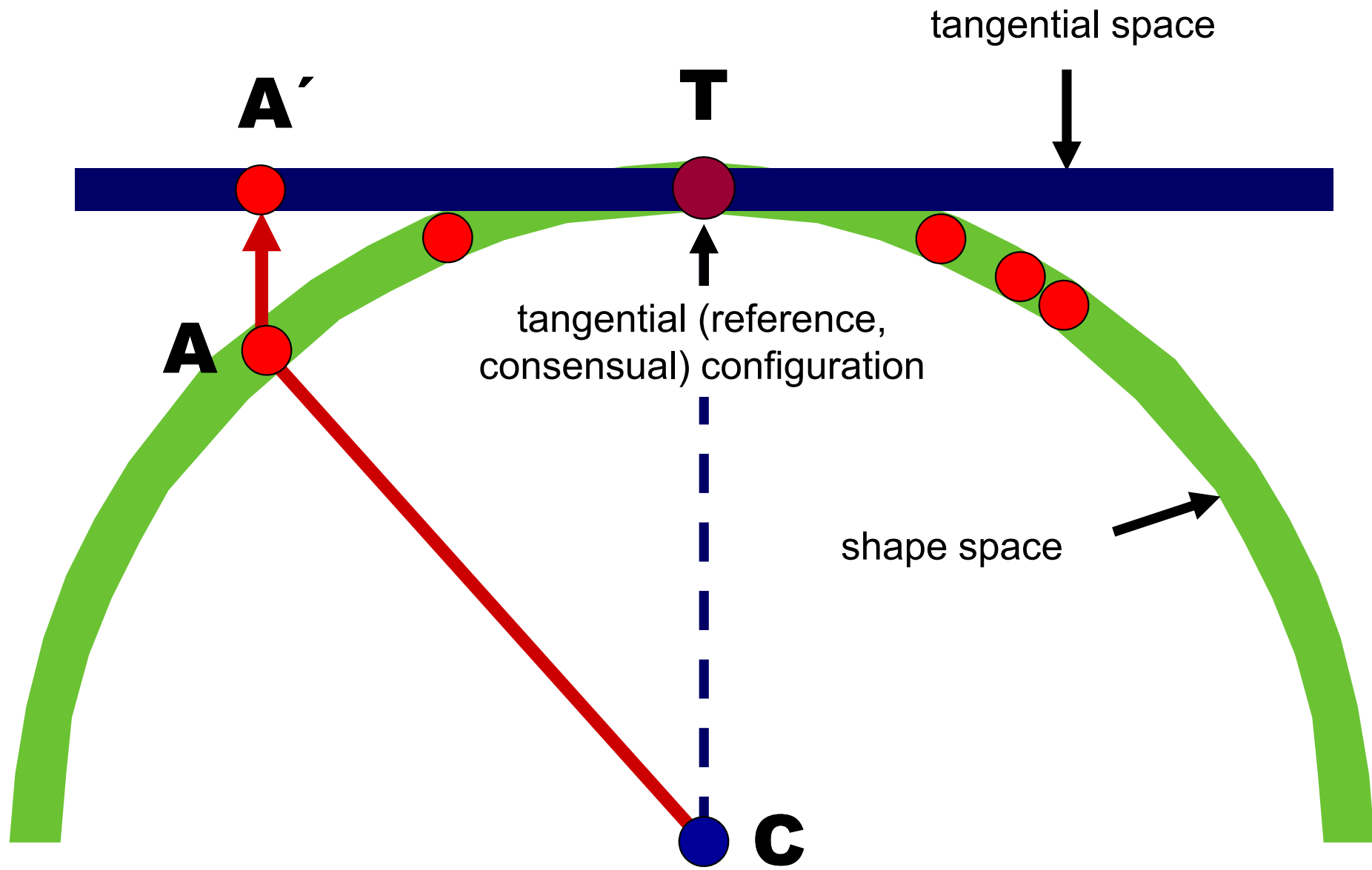


2. Superposition of the centers of gravity on a single point

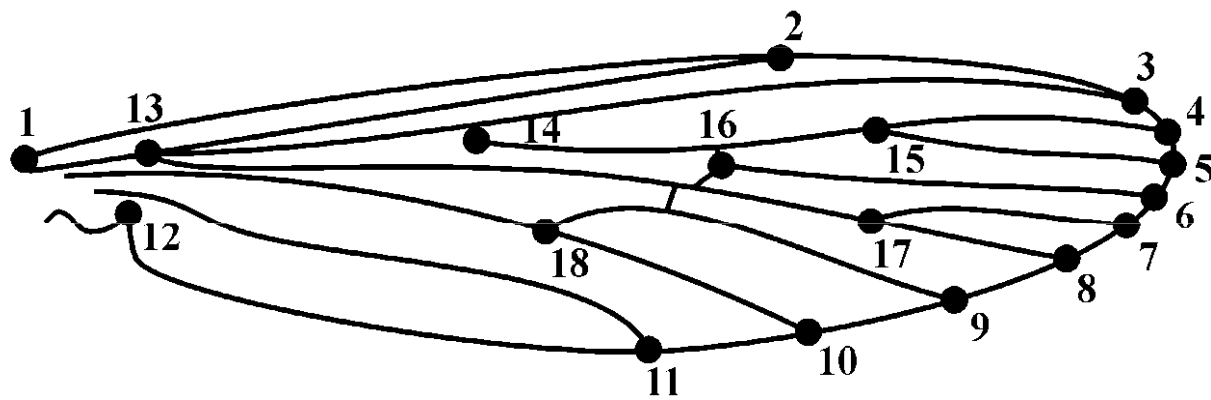


3. Rotation to minimize the dispersion of corresponding points

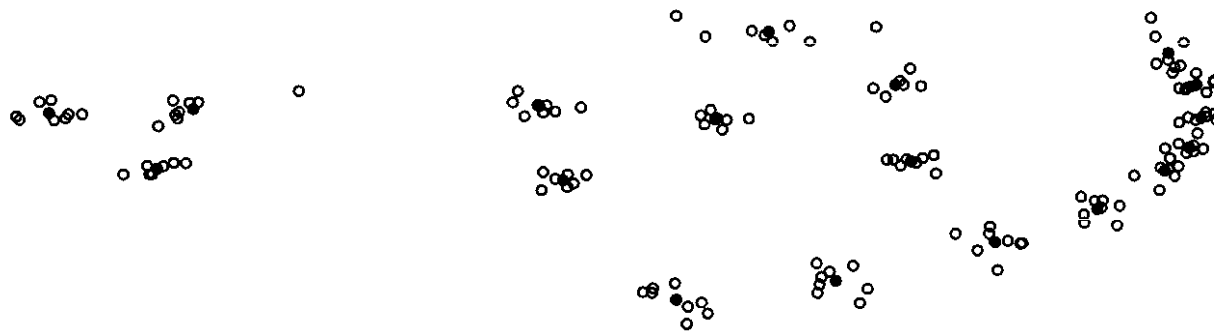




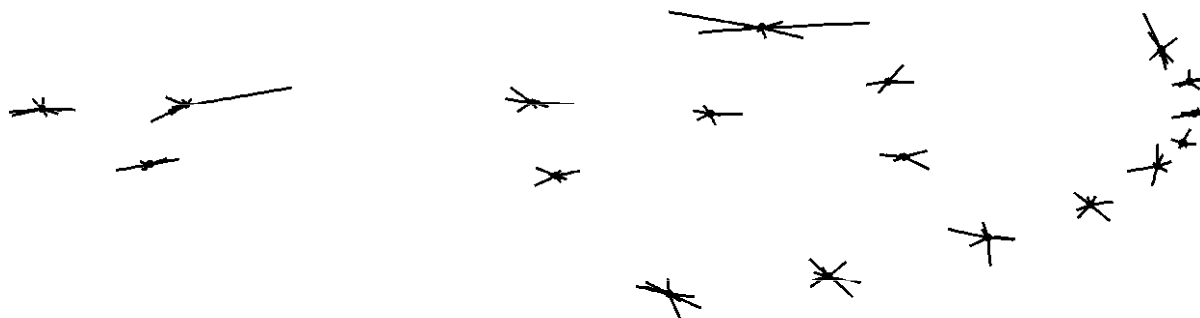
a)



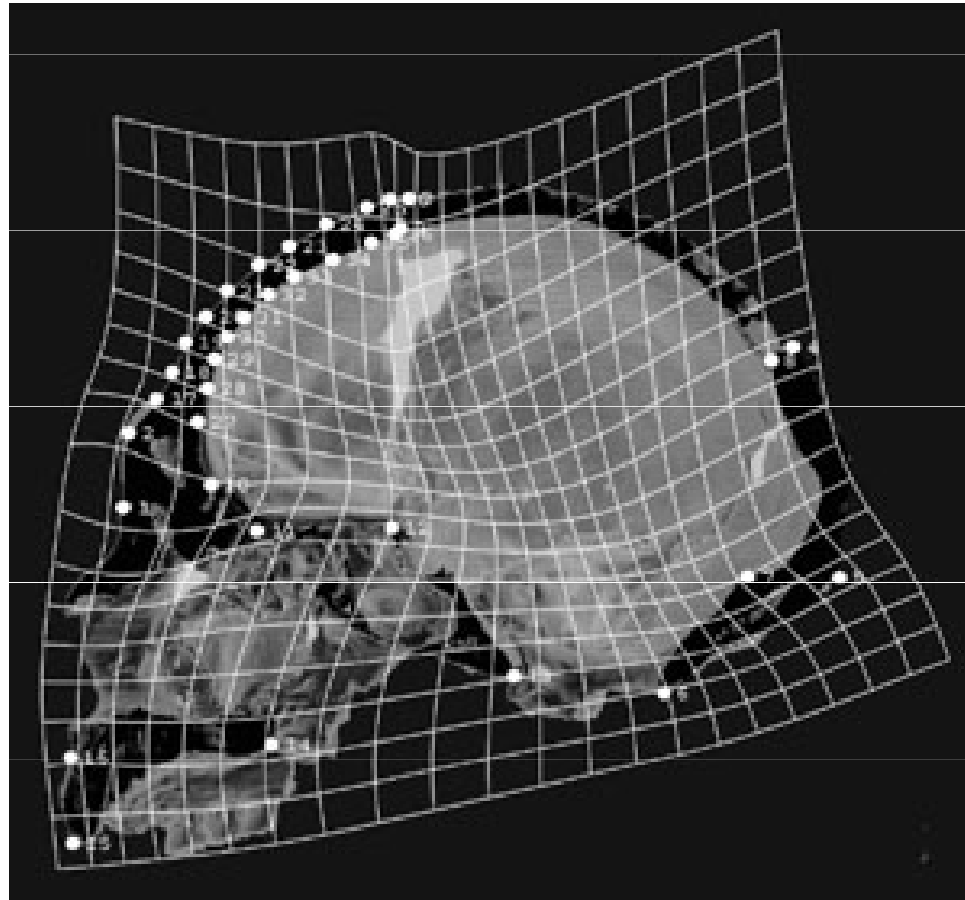
b)



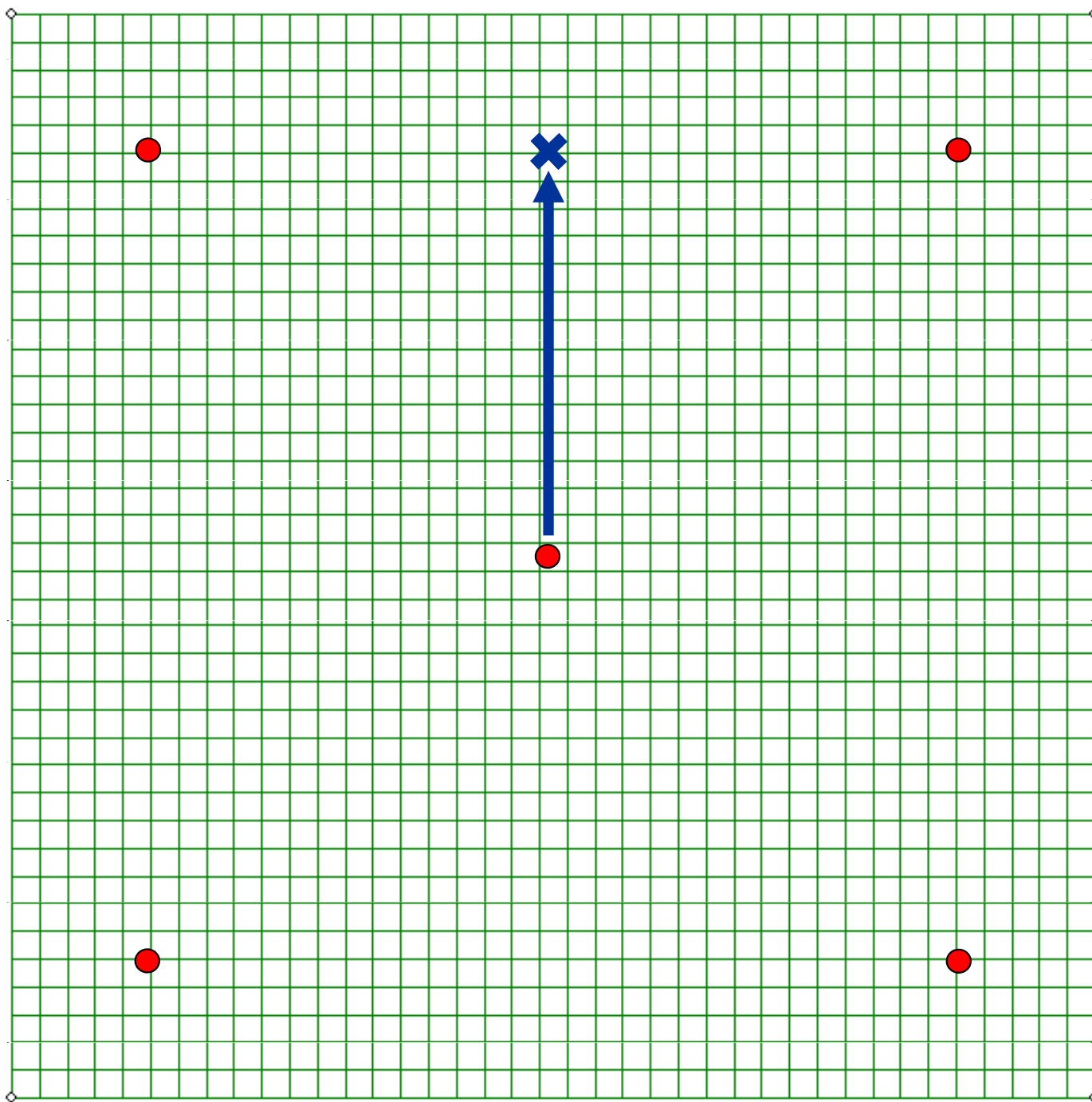
c)

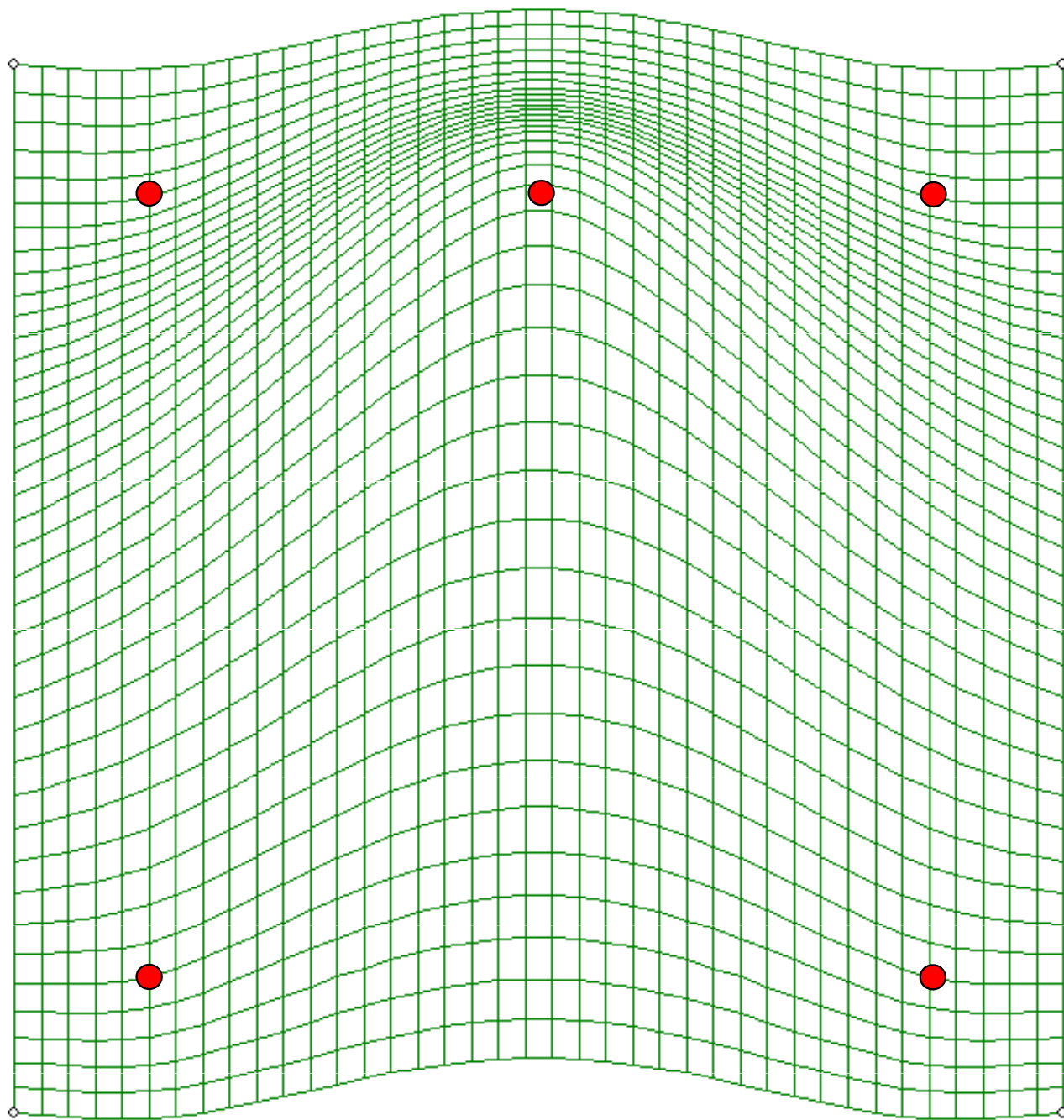


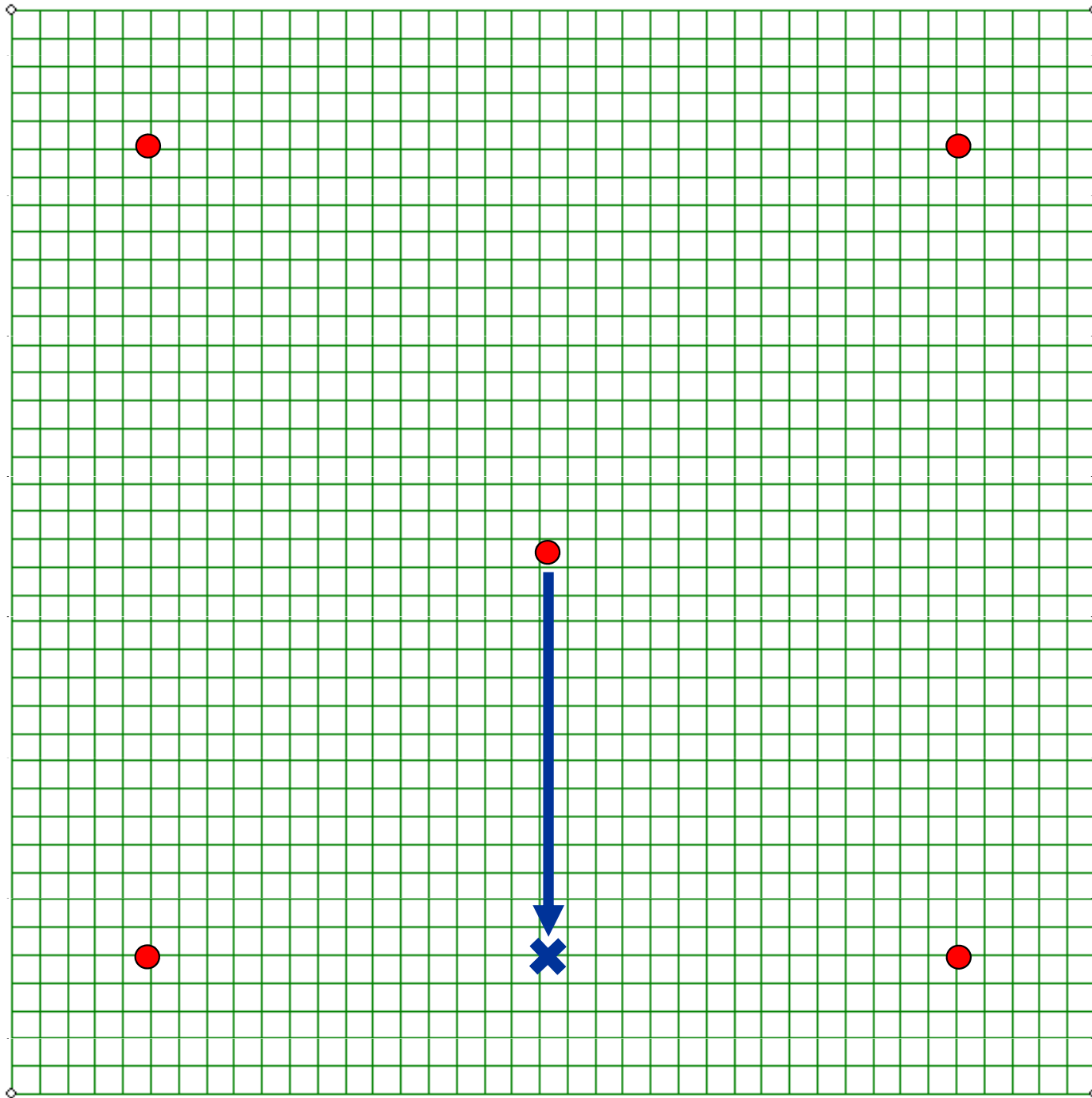
Thin-Plate Spline (TPS)

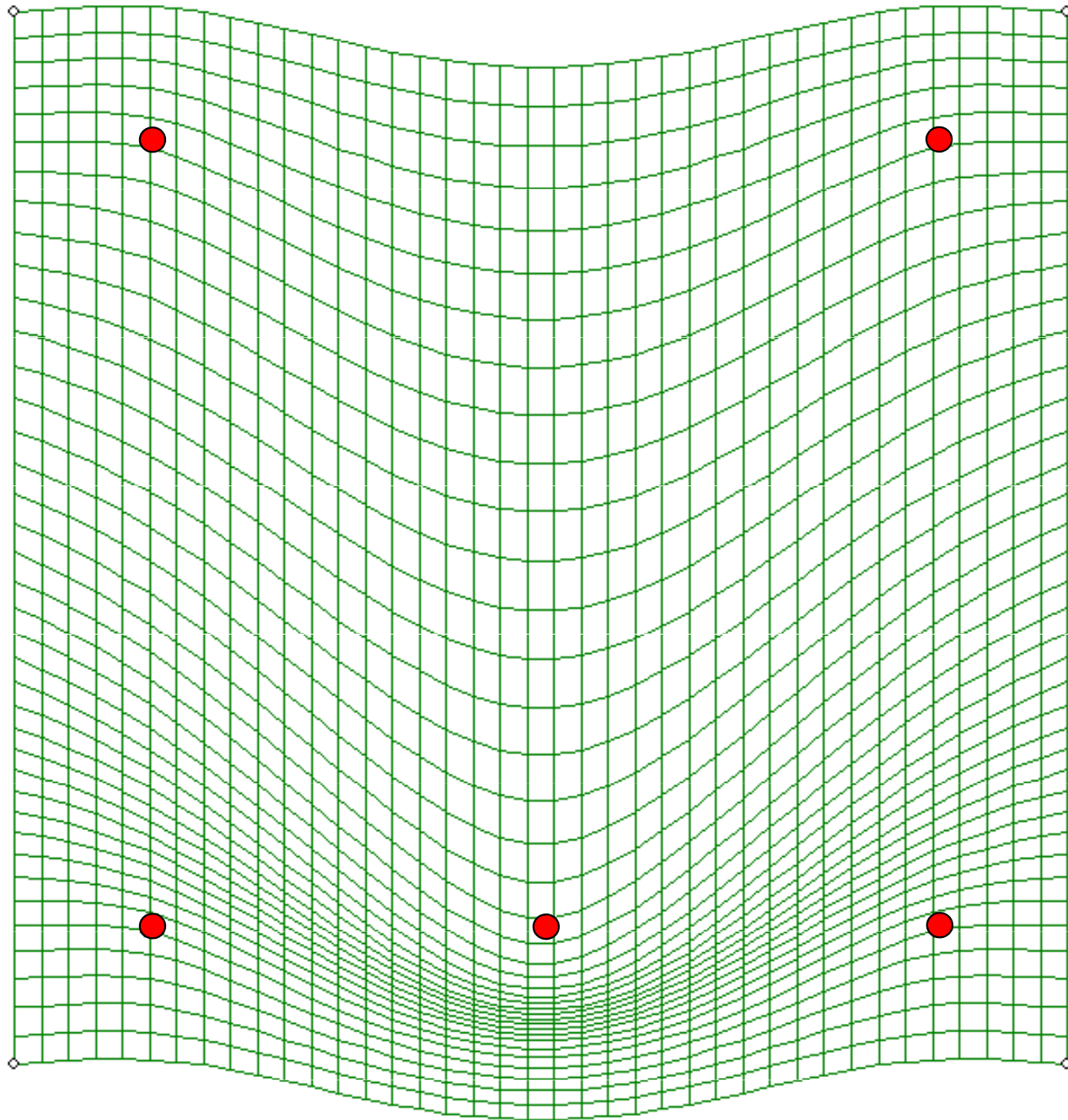


metaphor of infinitely large and infinitely thin metal plate









Thin-Plate Spline (TPS)

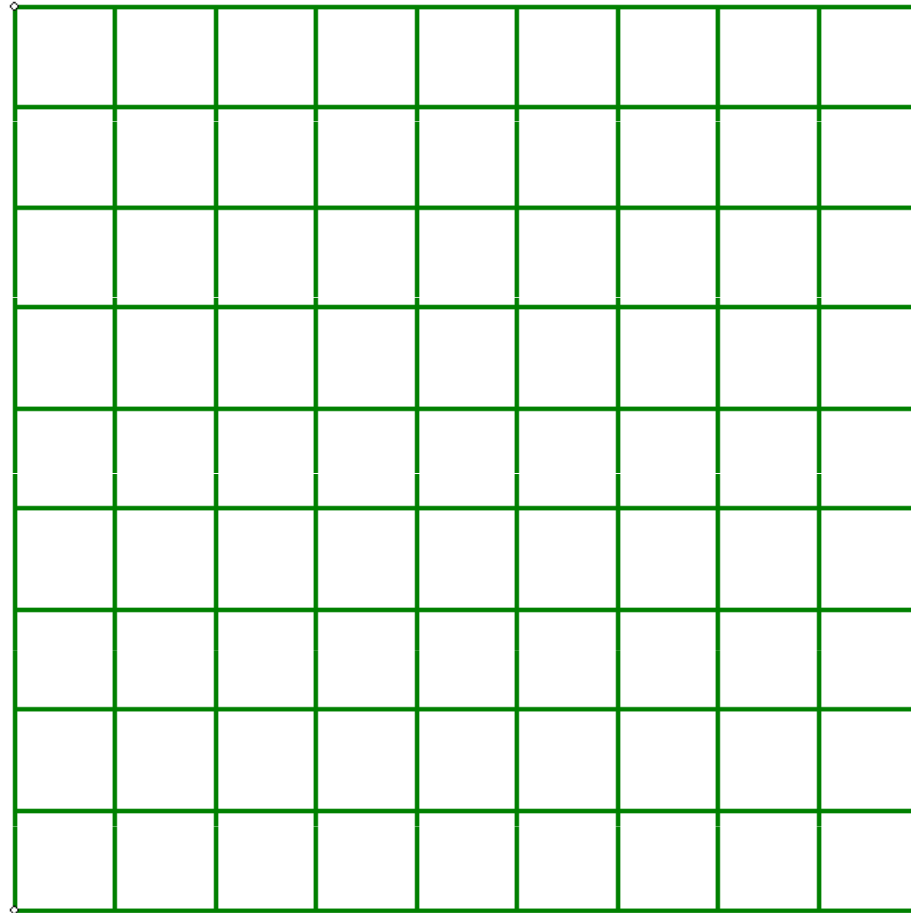
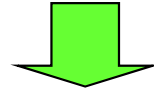
energy necessary for plate deformation = **bending energy**

differentiating between affine and nonaffine shape changes

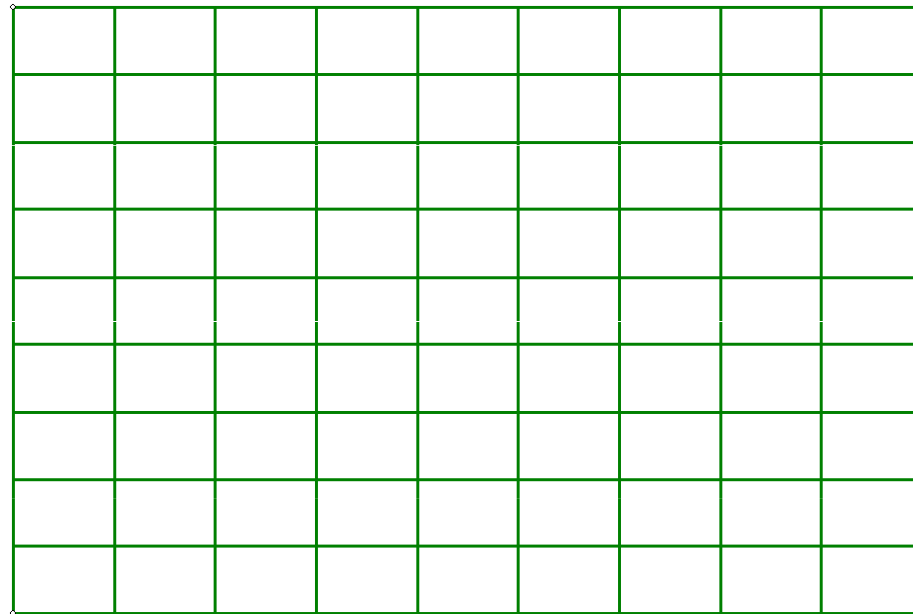
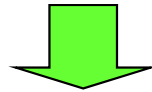
projection of latent roots of bending energy into components
= **partial warps**

partial warp 0 ~ affine component

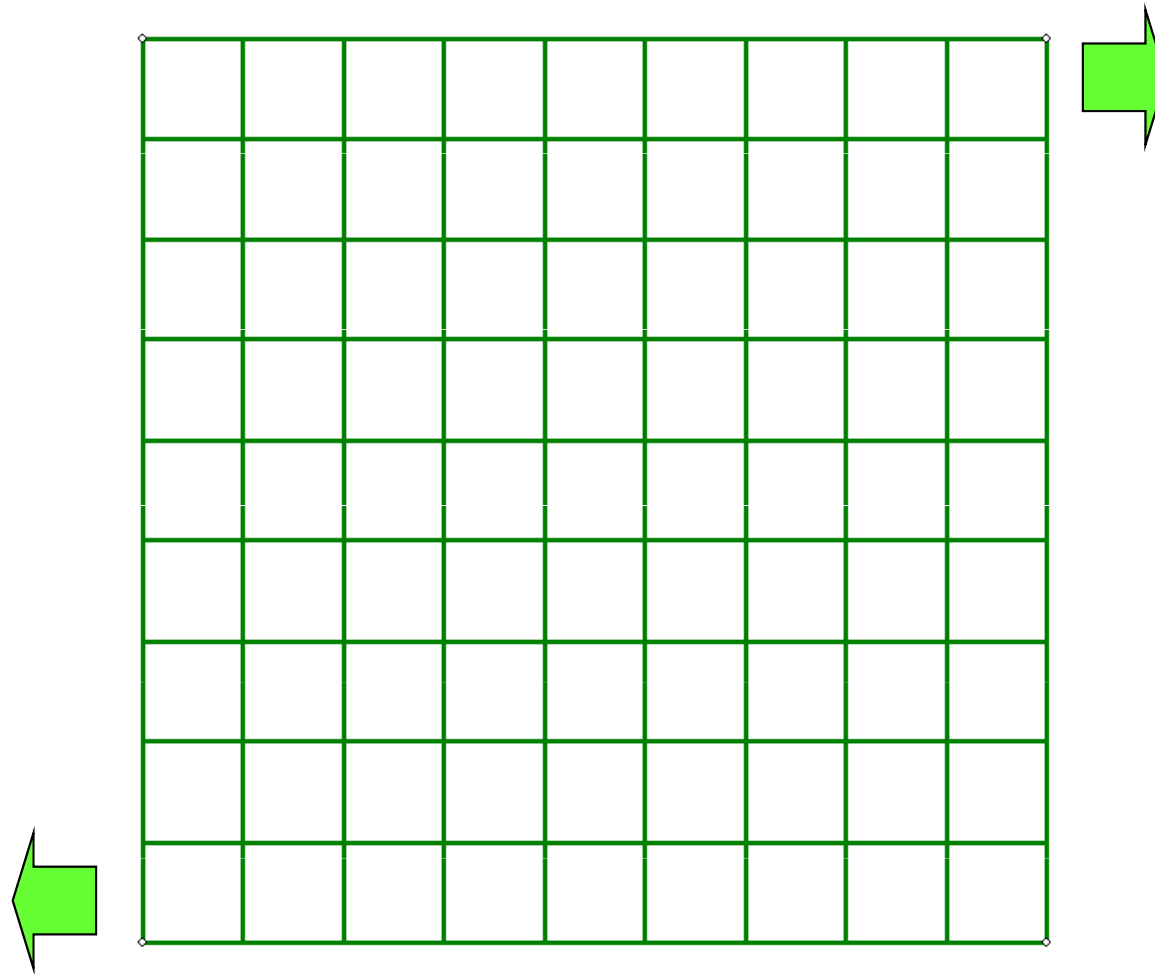
Afine shape change



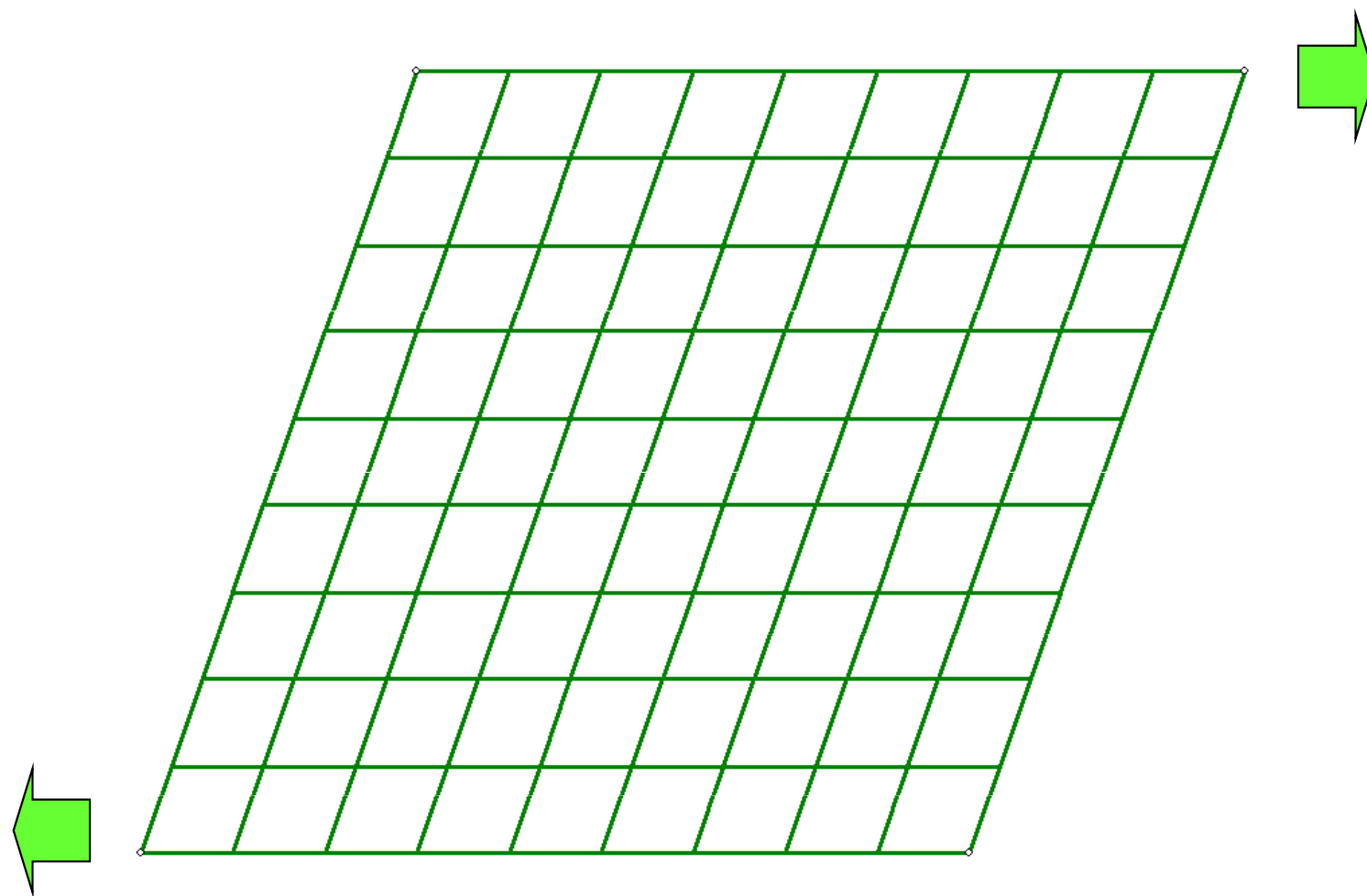
A fine shape change



Afine shape change

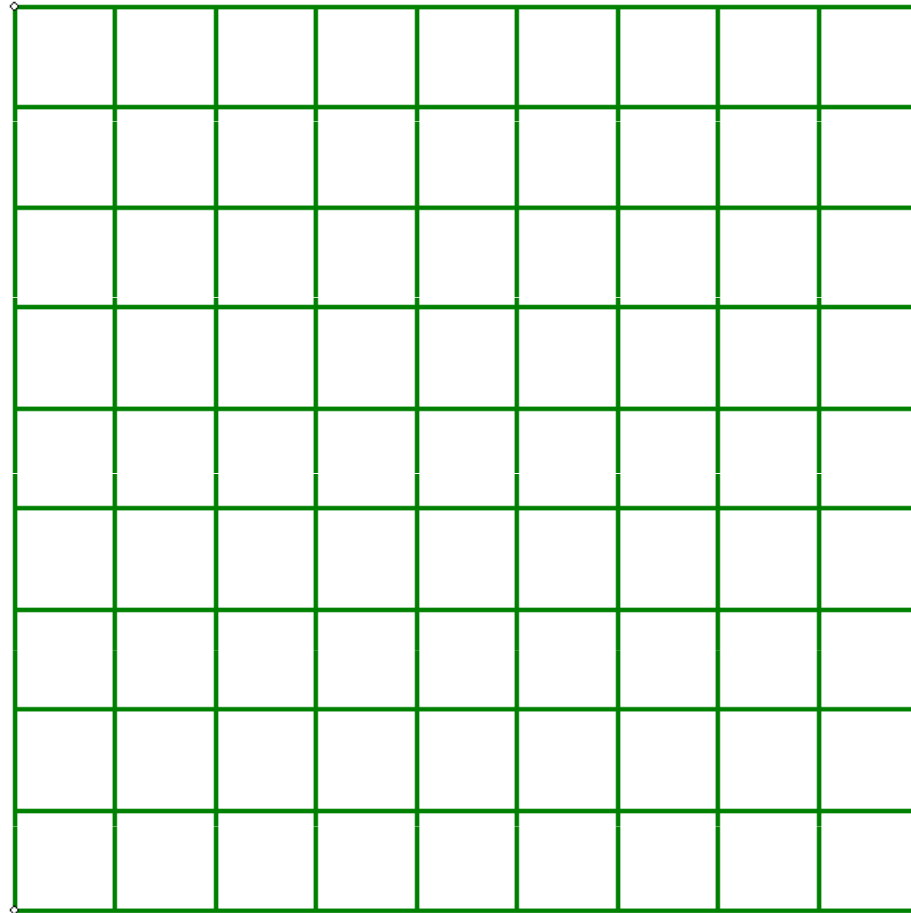


Afine shape change

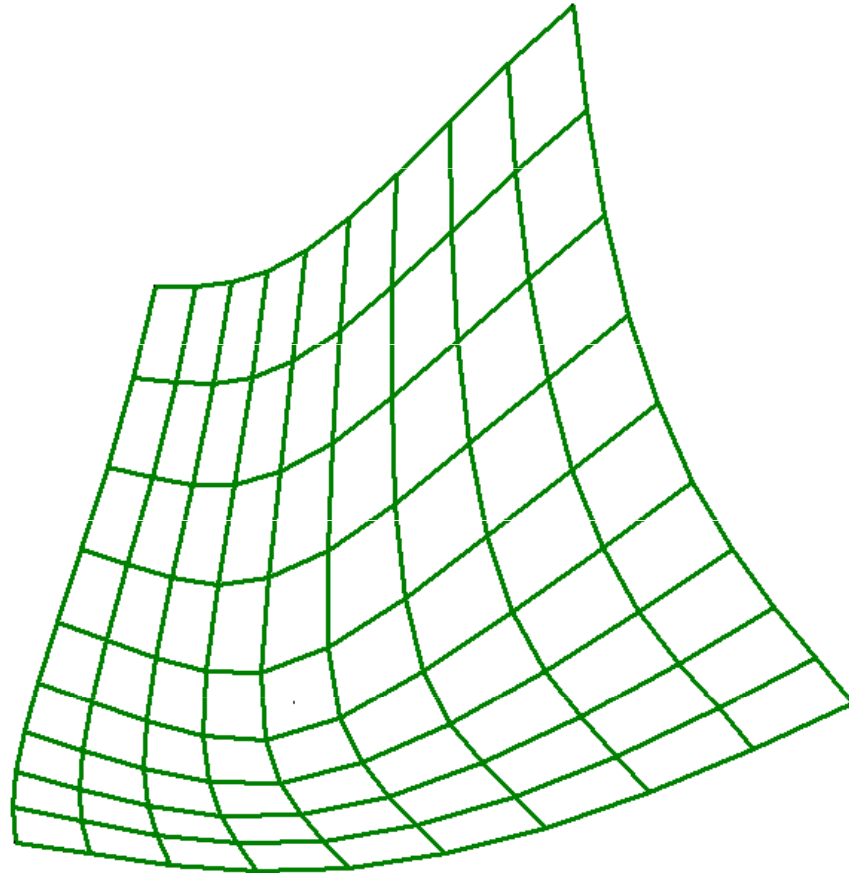


parallel lines remain parallel

Nonafine shape change

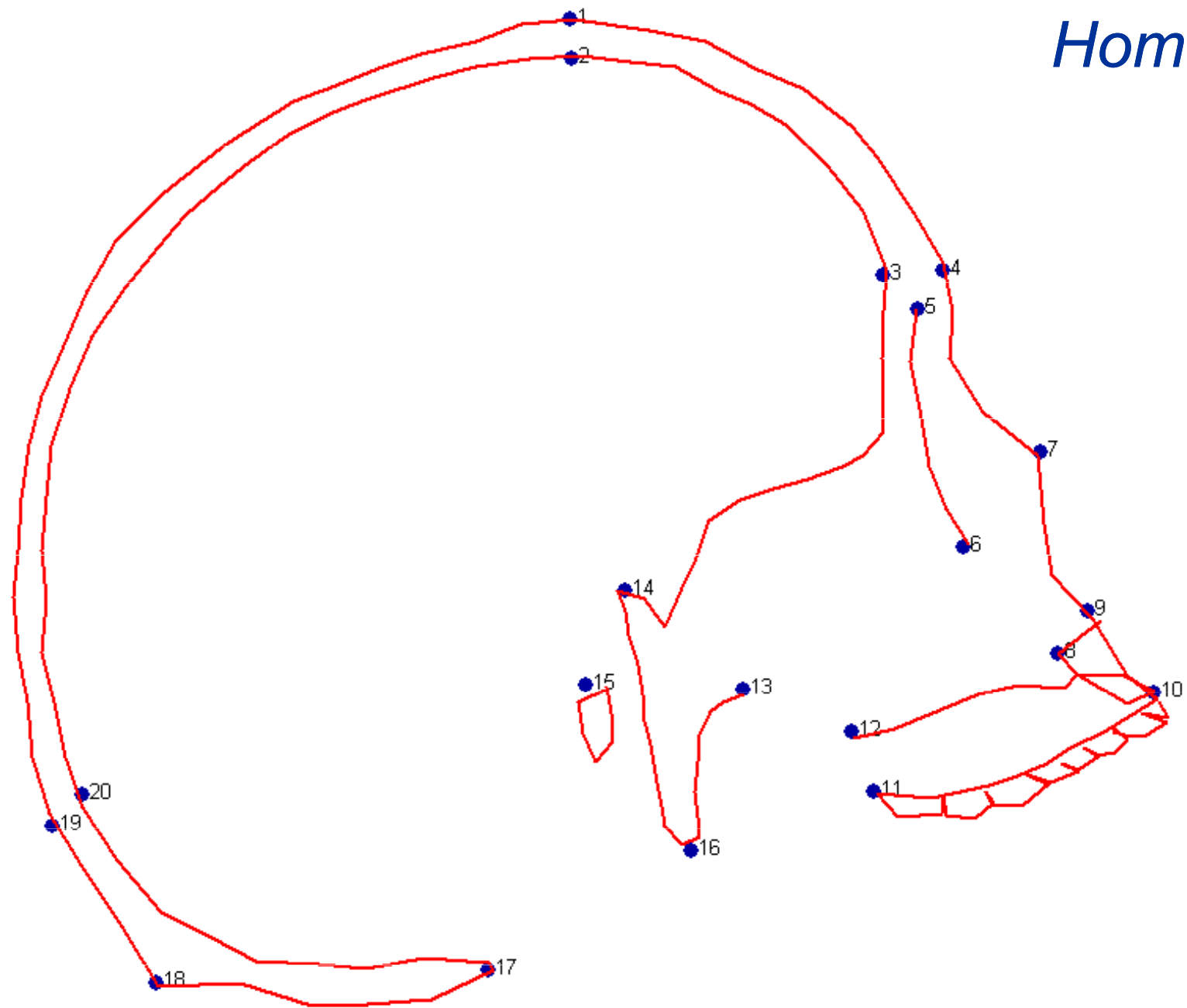


Nonafine shape change

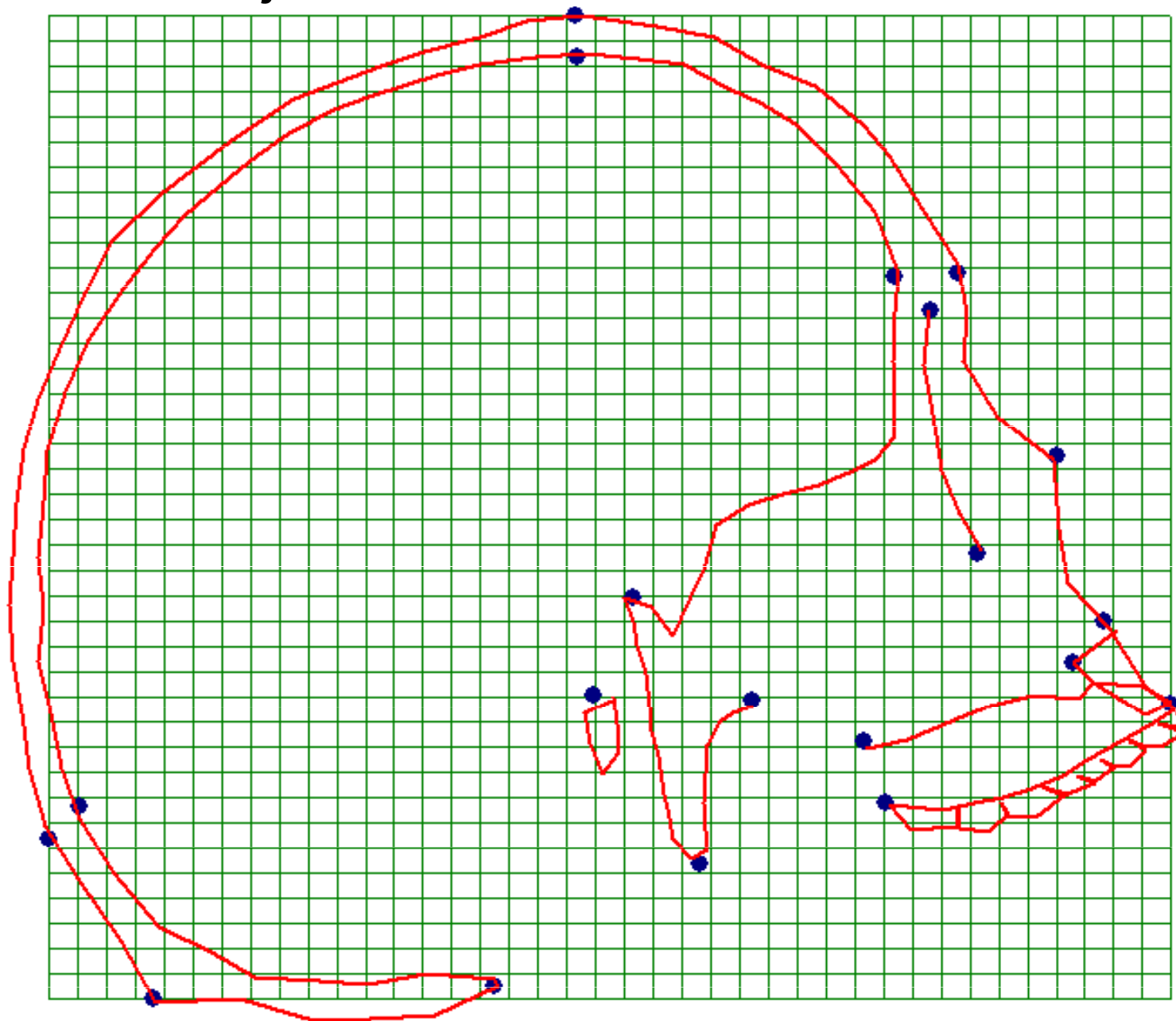


parallel lines are not parallel any more

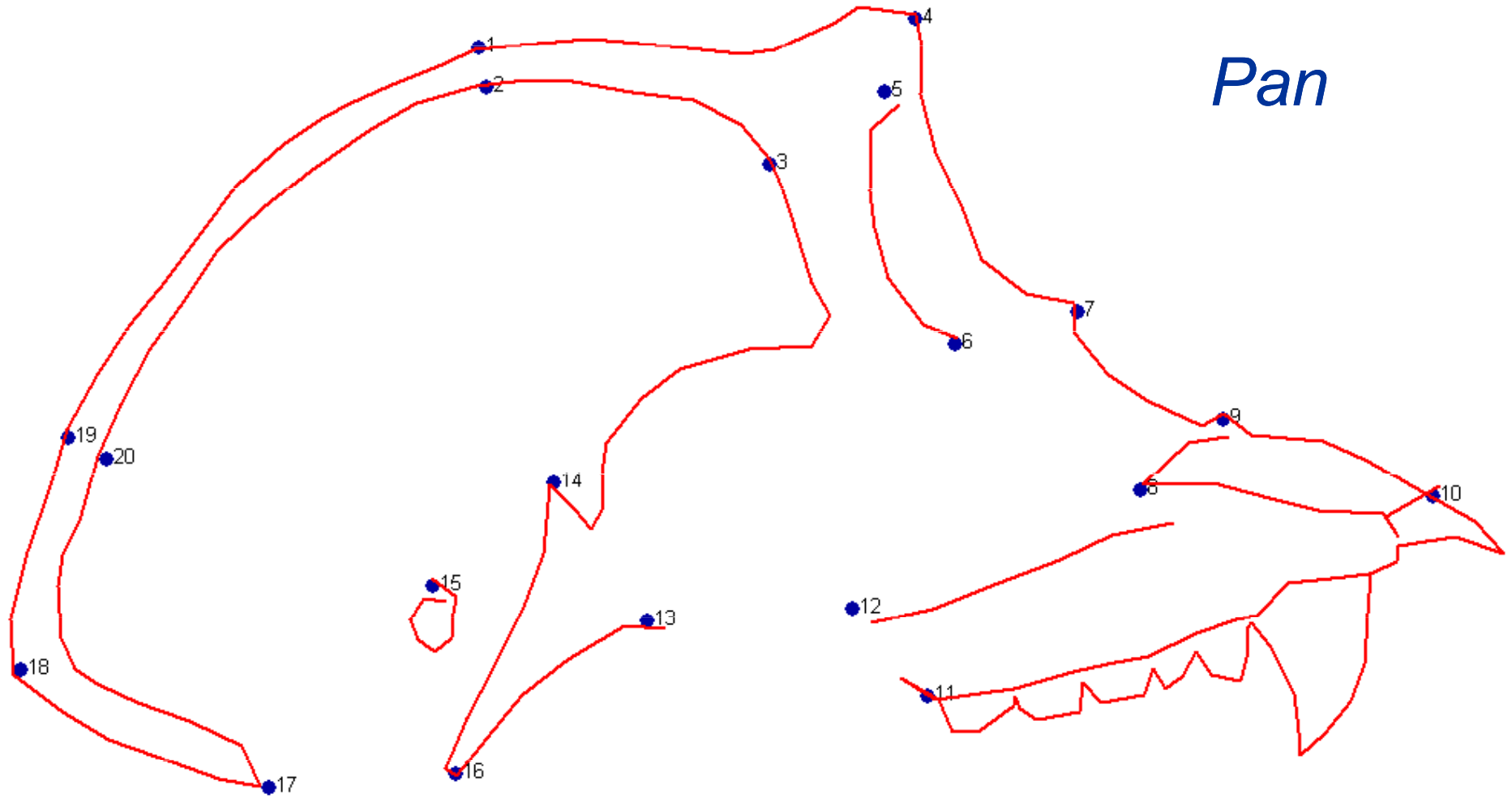
Homo



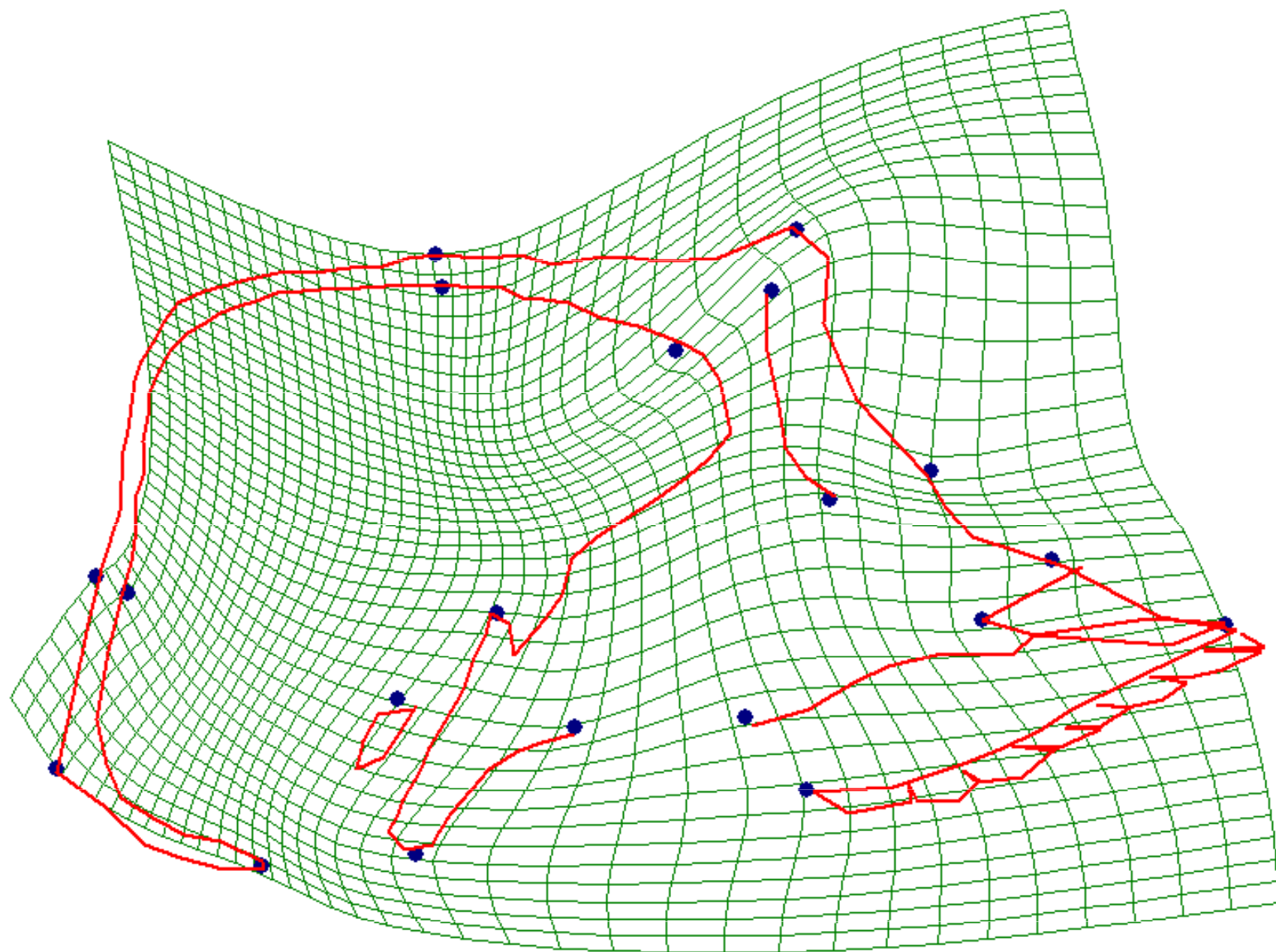
reference object

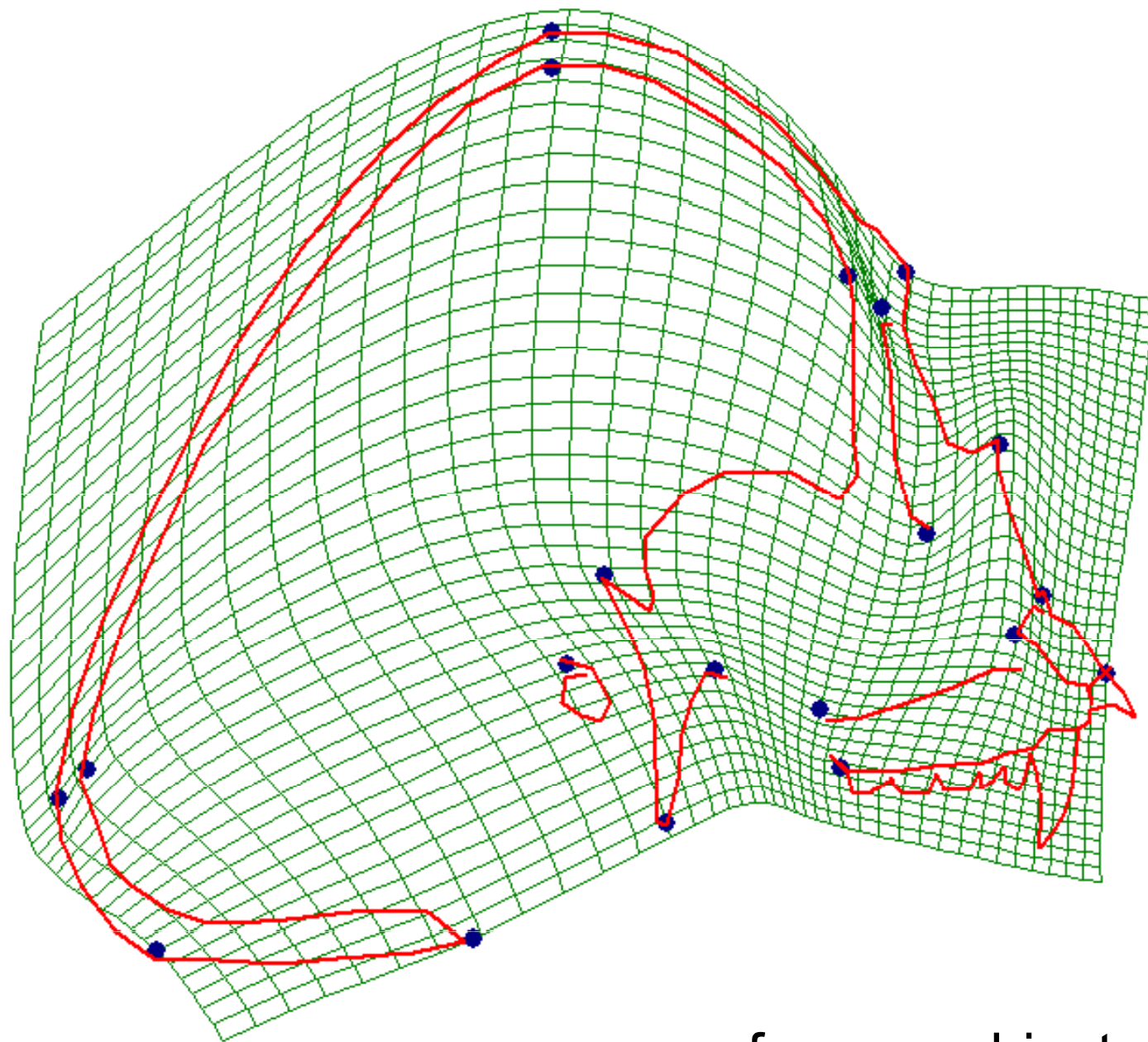


Pan



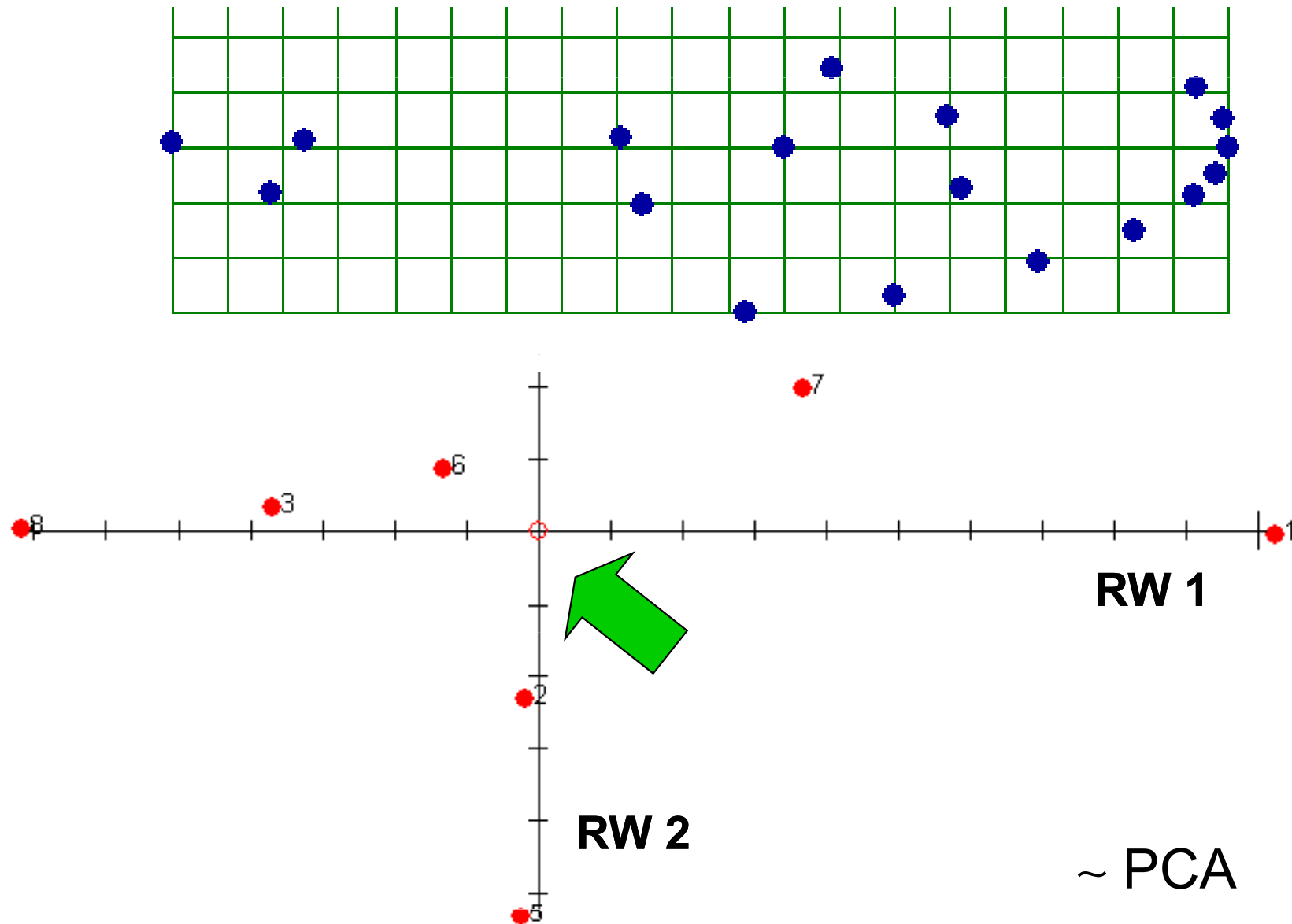
reference object = *Homo*

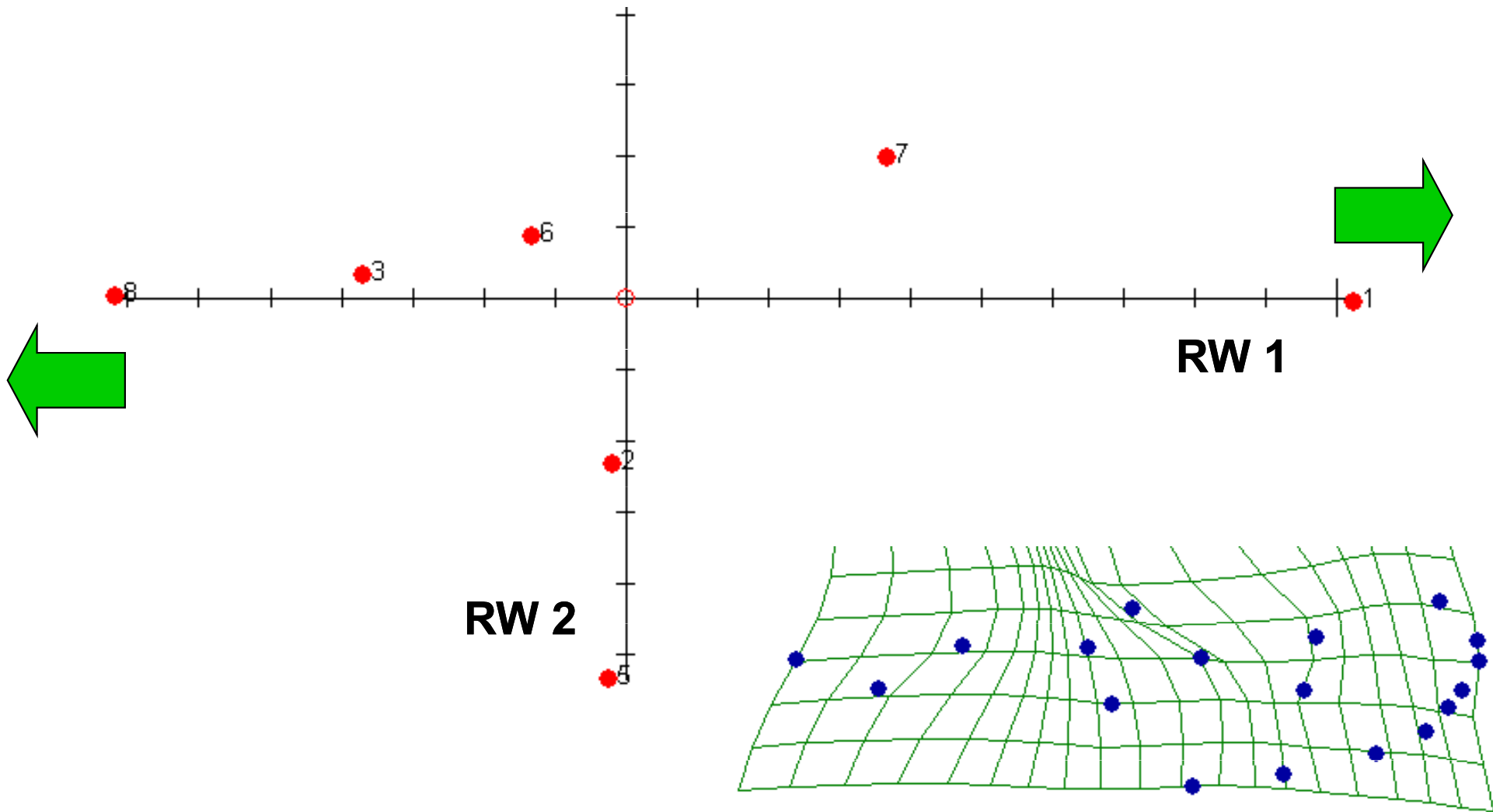
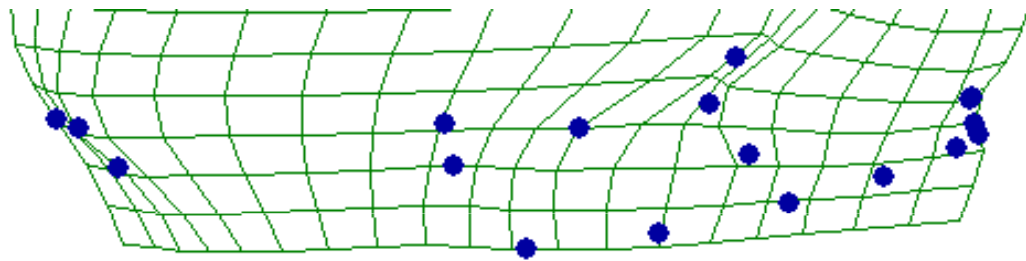




reference object = *Pan*

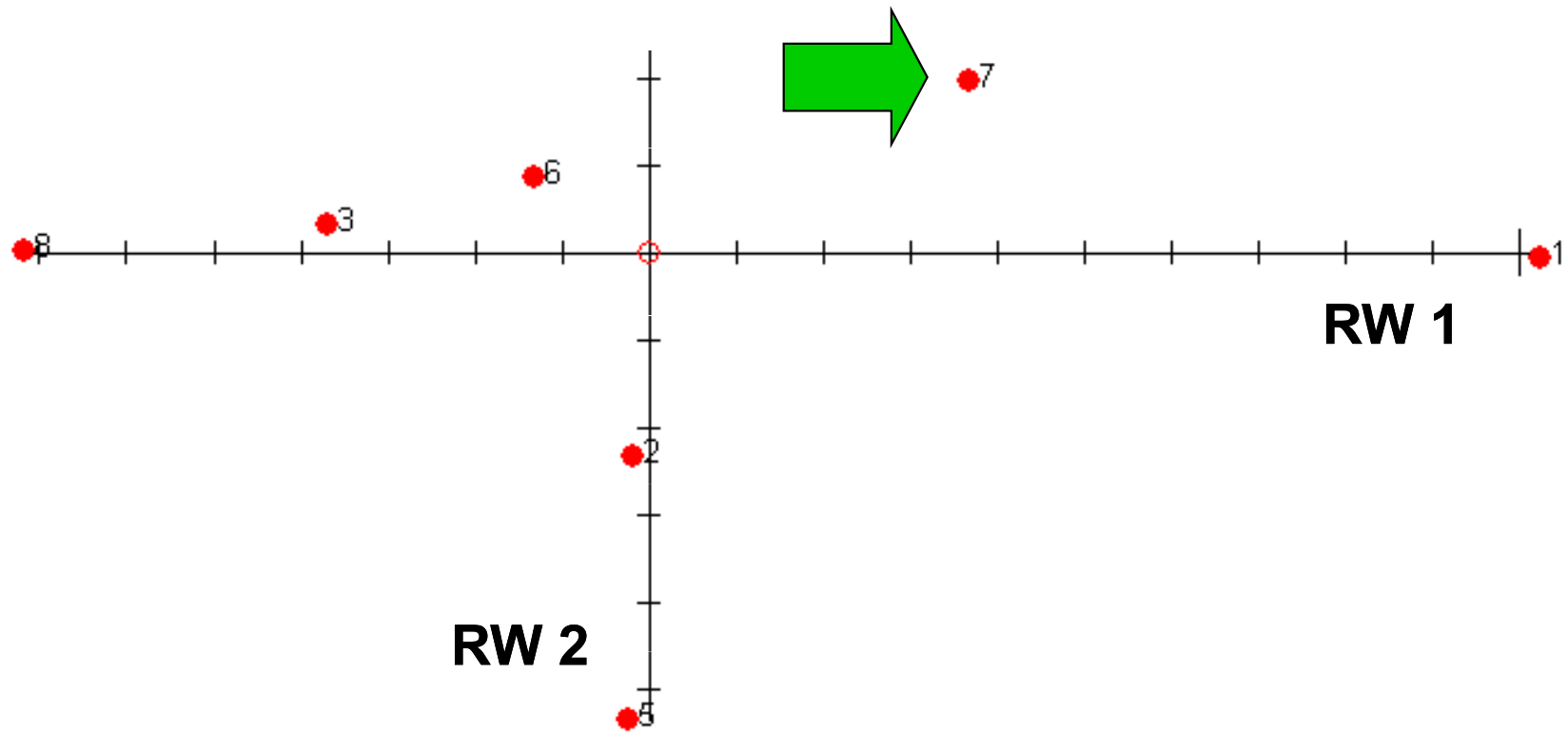
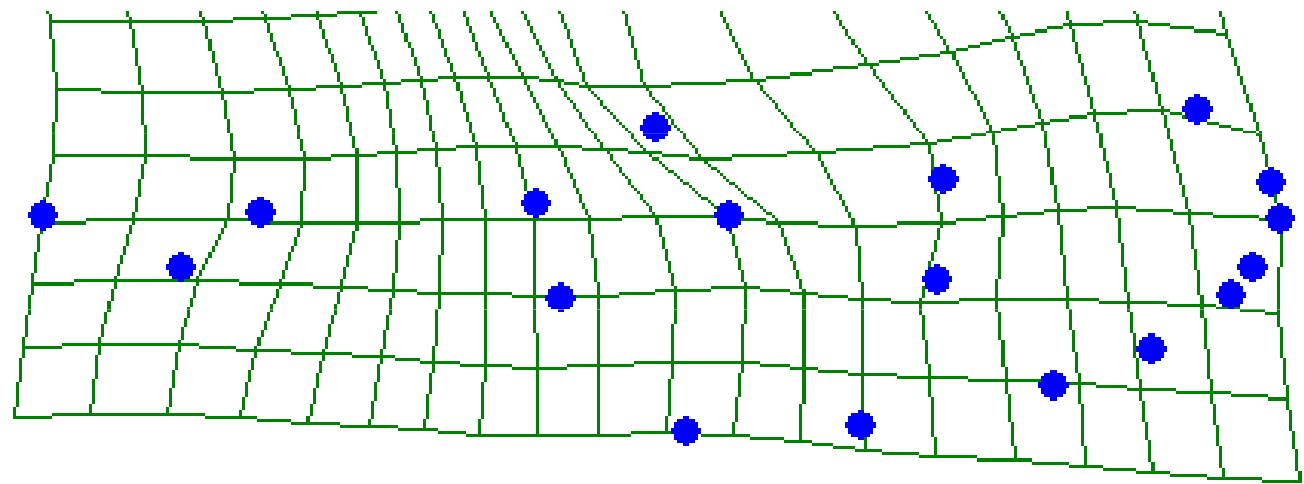
>2 samples: Thin-Plate Spline Relative Warps (TPSRW)





RW 2

RW 1



Software:

tpsDig: úprava obrázků, digitalizace bodů, měření rozměrů

tpsSplin: TPS

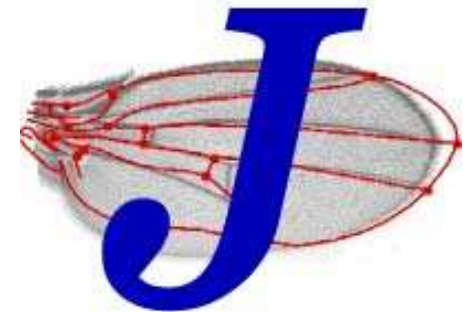
tpsRelw: TPS Relative Warps

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



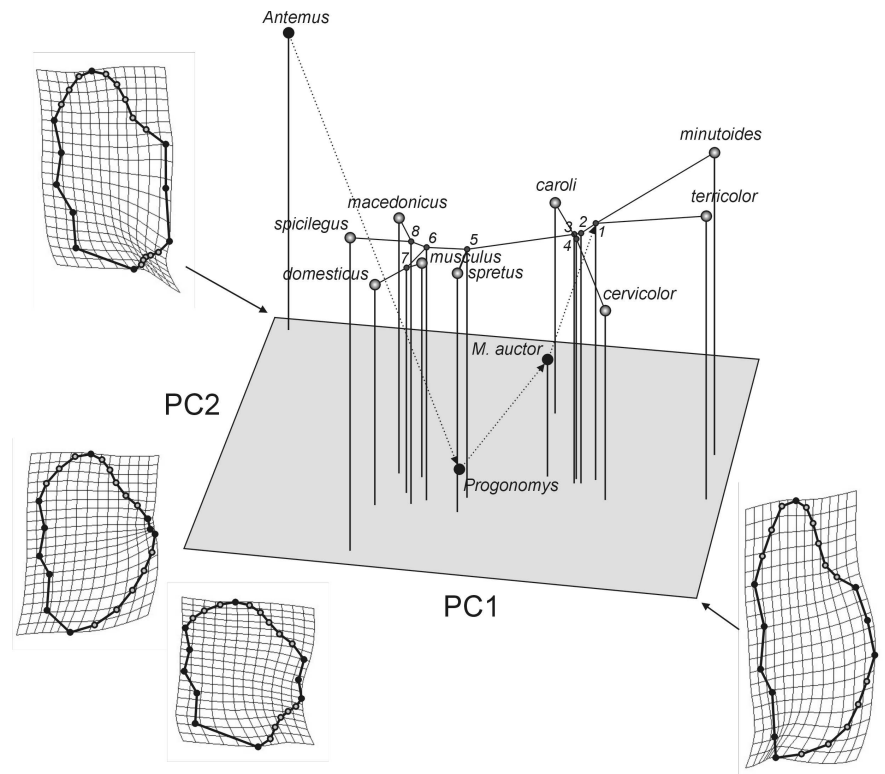
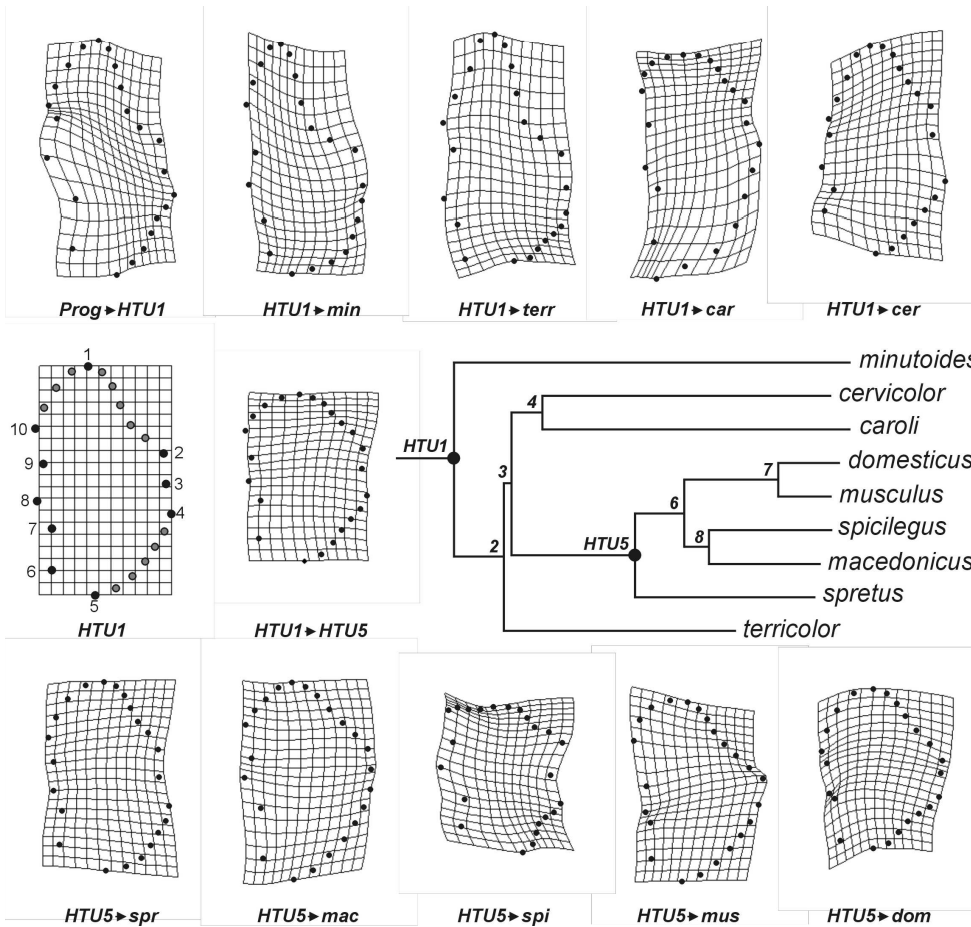
MorphoJ

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

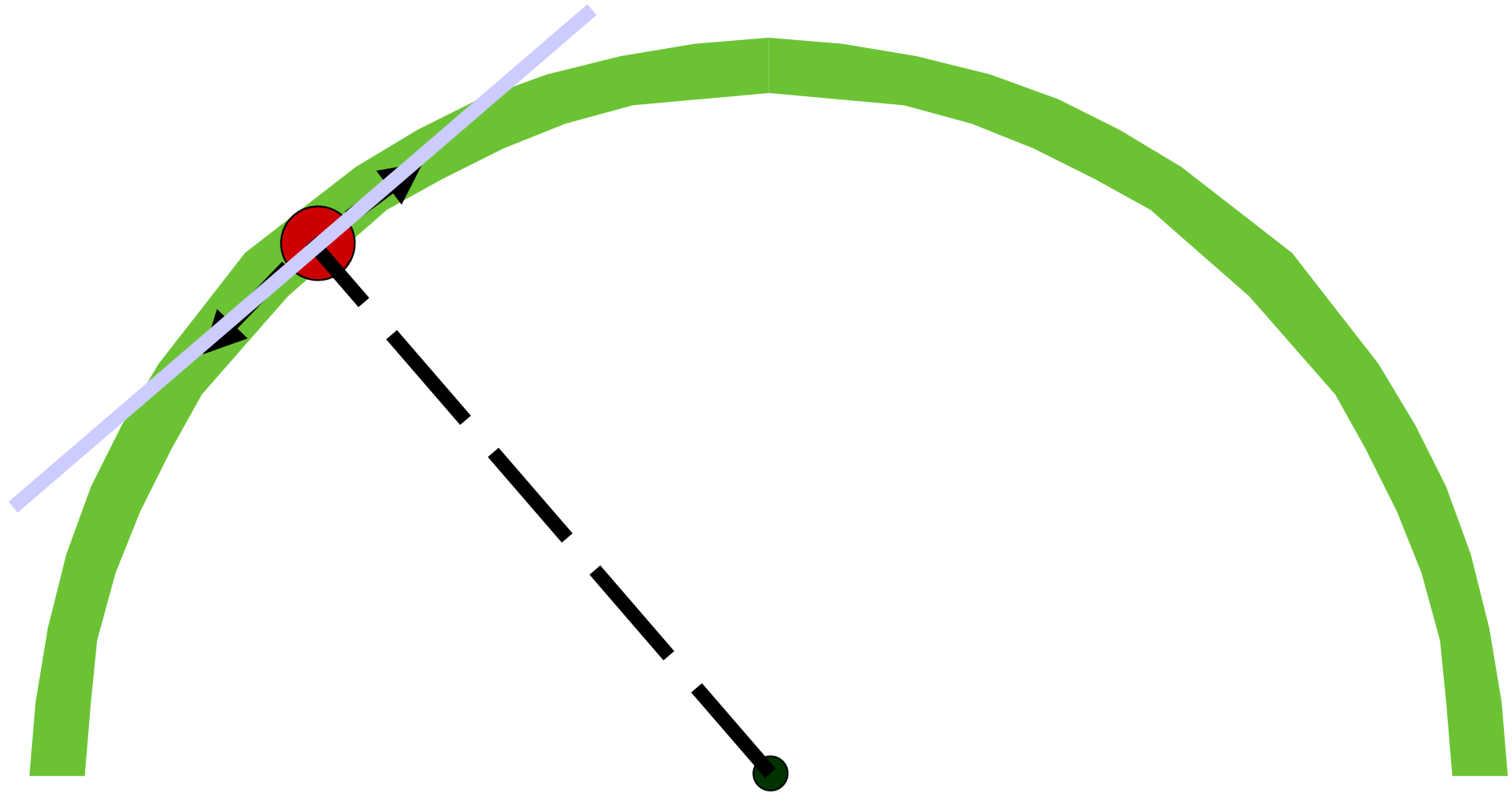
Morfometrics and phylogenesis

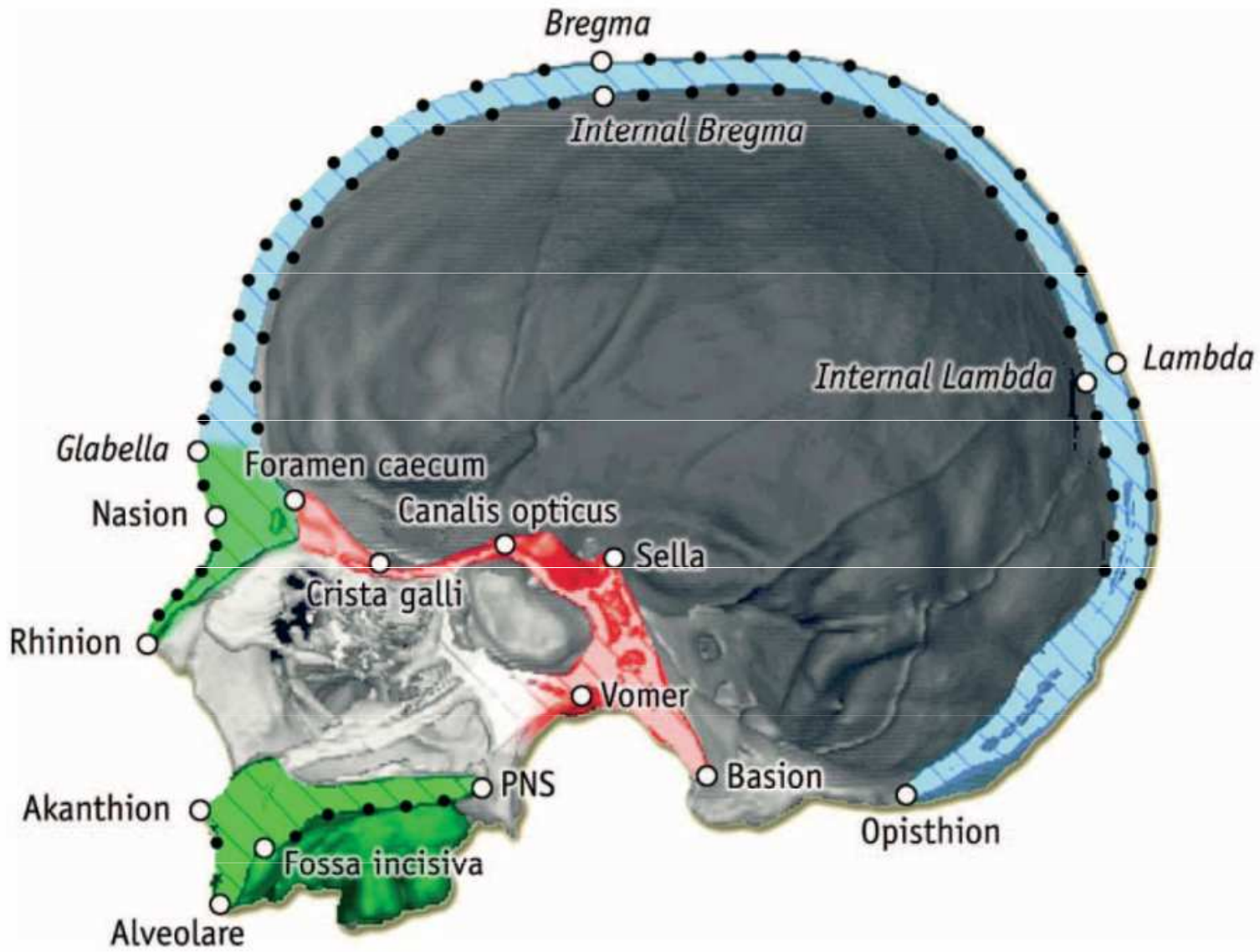
tpsTree

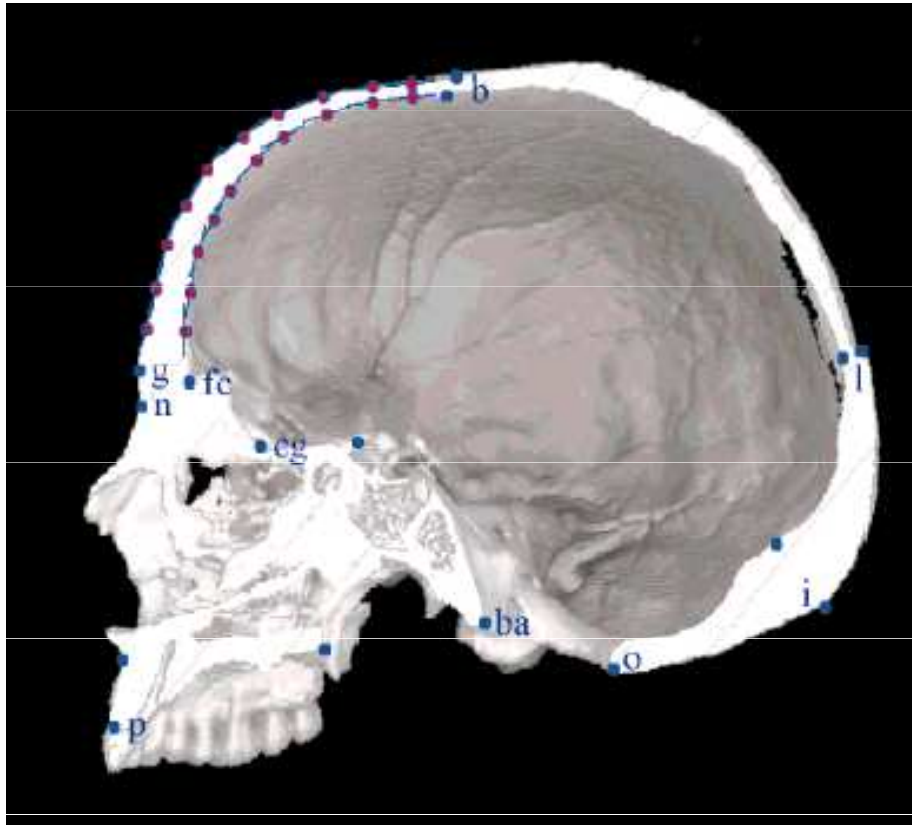
mapping of shape changes



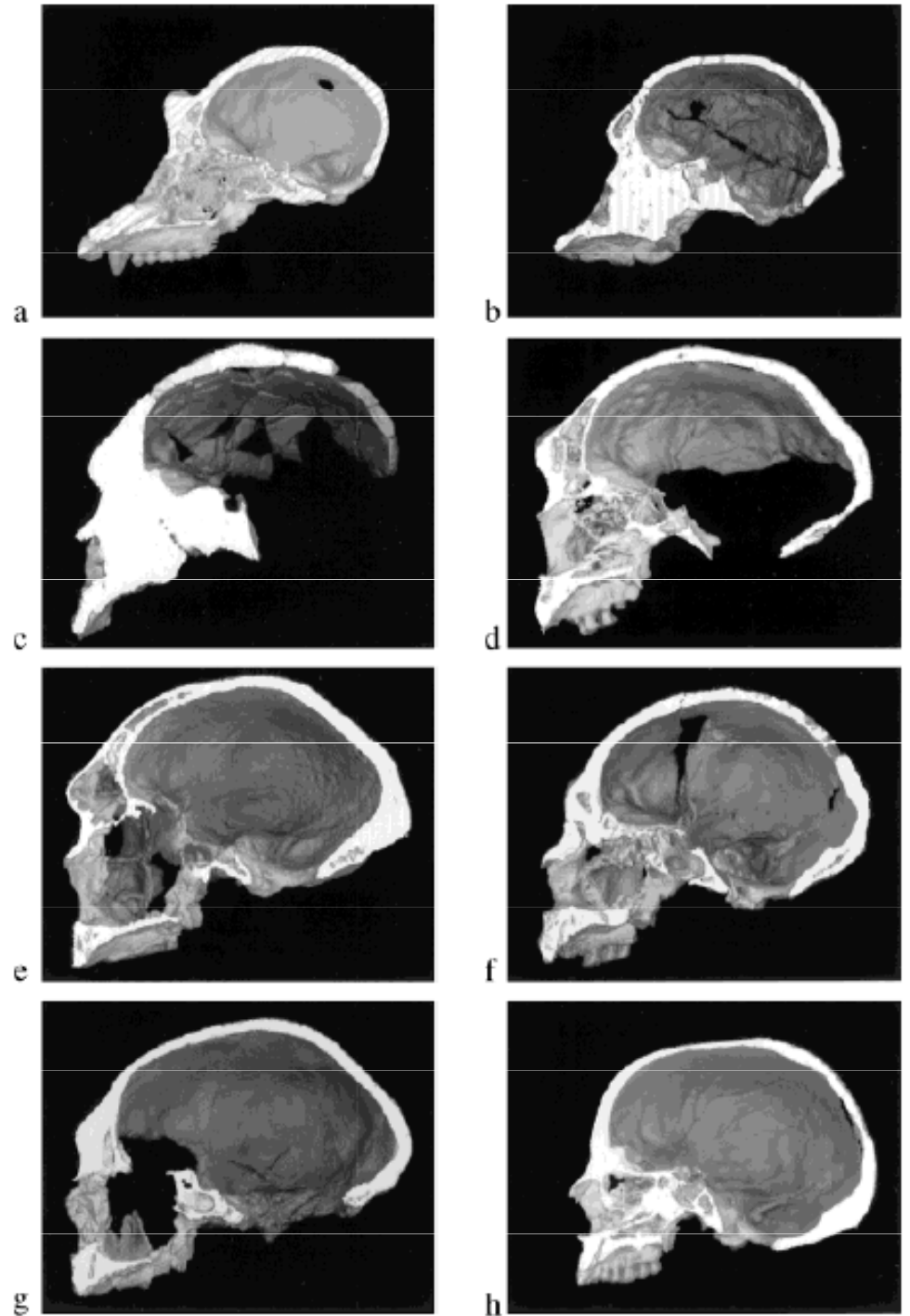
Landmark based methods without landmarks – „sliding semilandmarks“

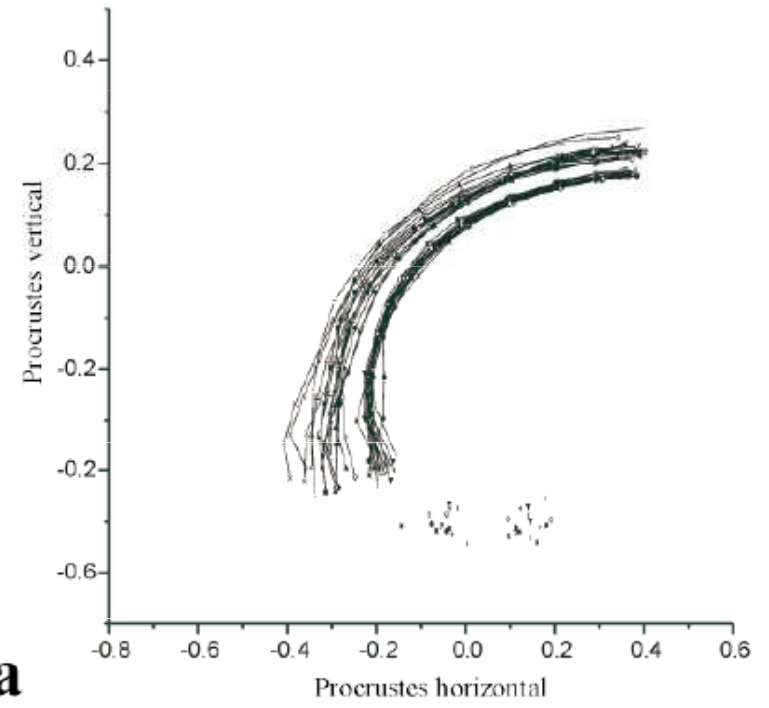
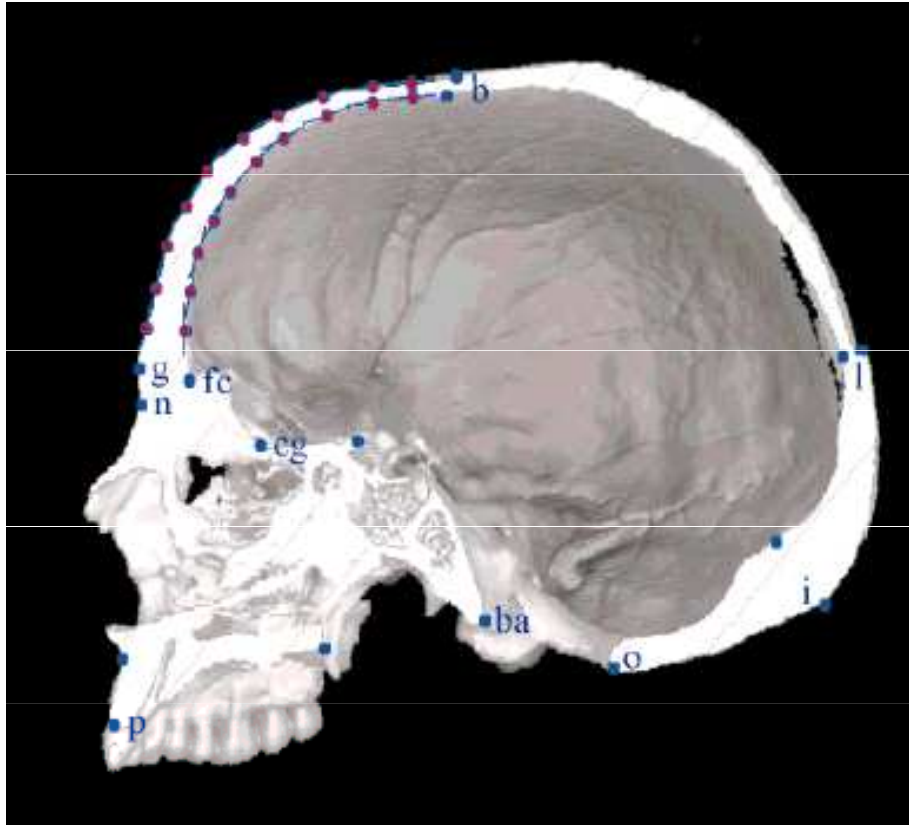




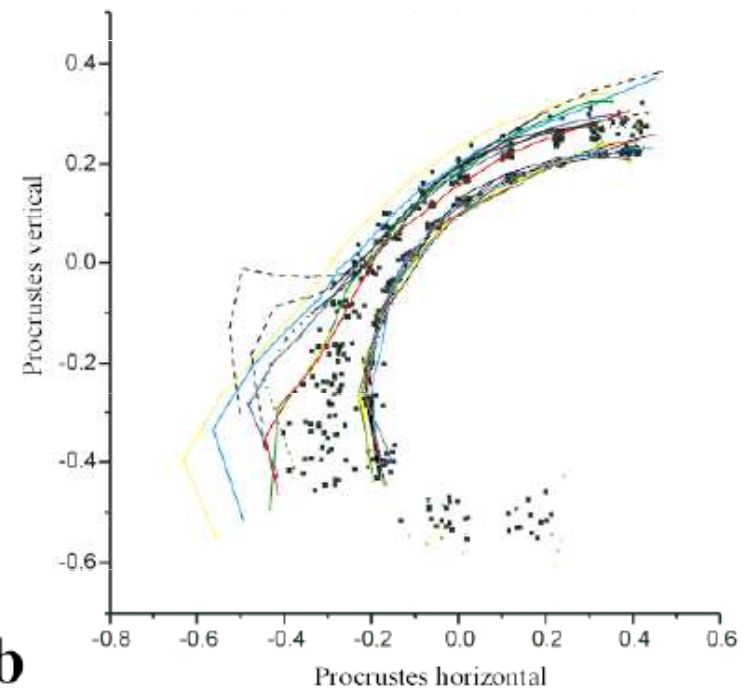


Bookstein et al.,
Anat. Record (1999)



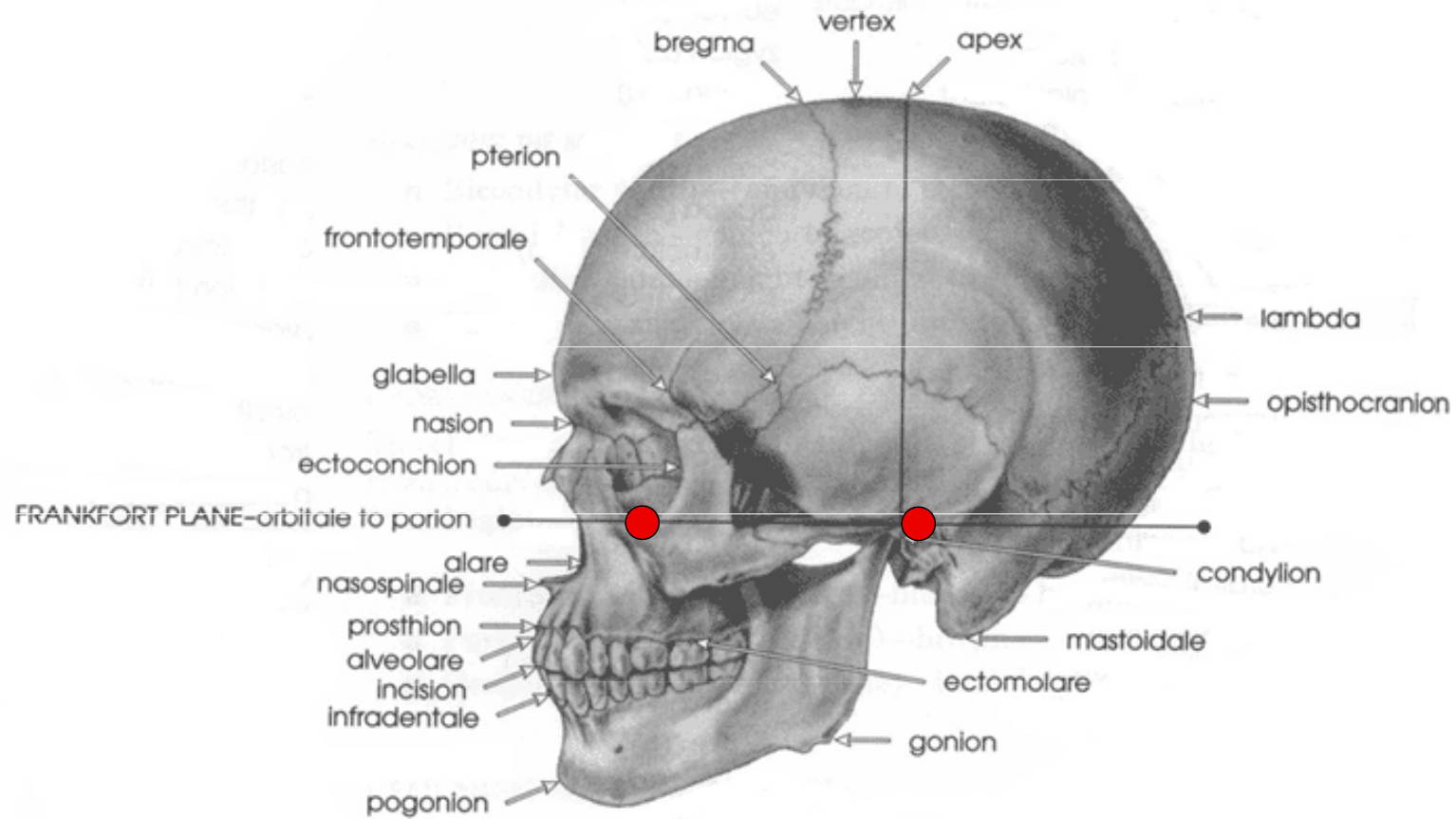


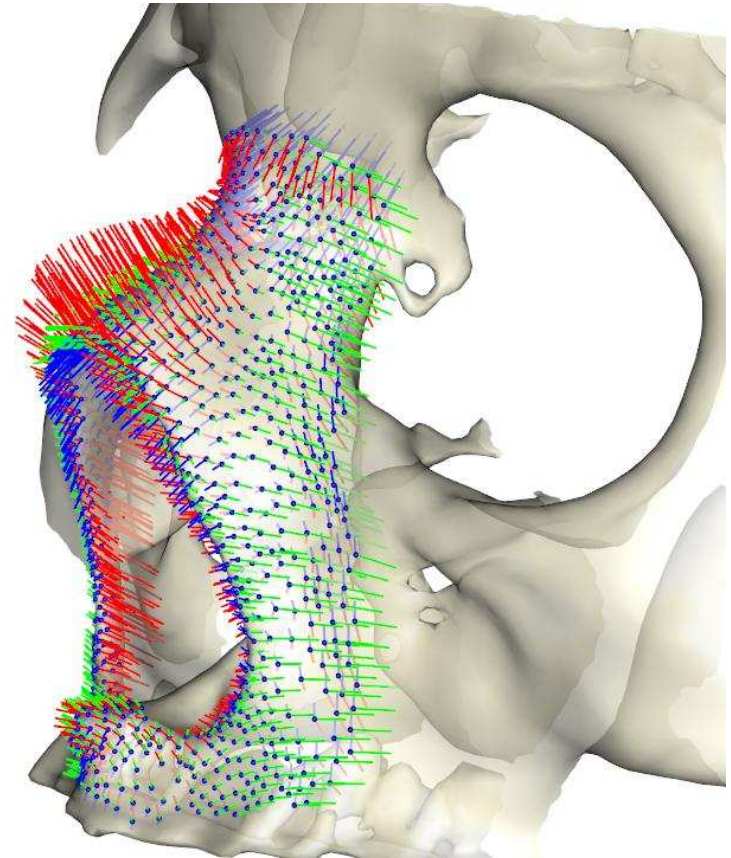
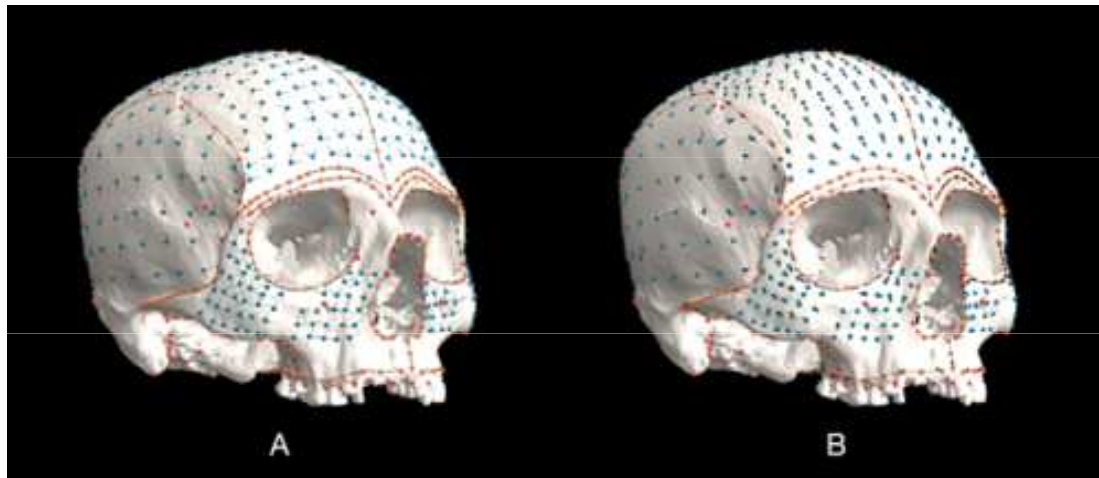
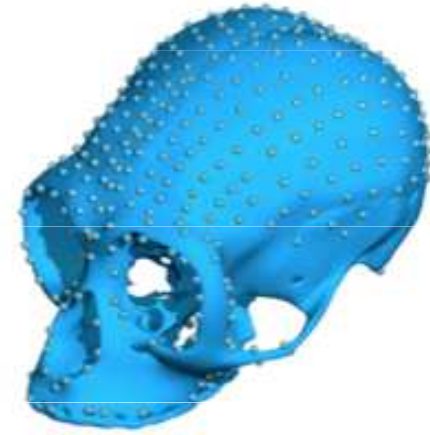
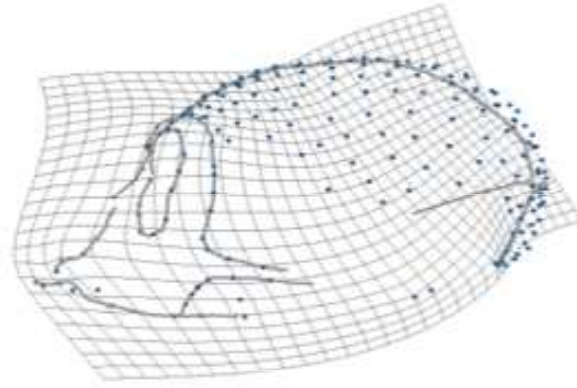
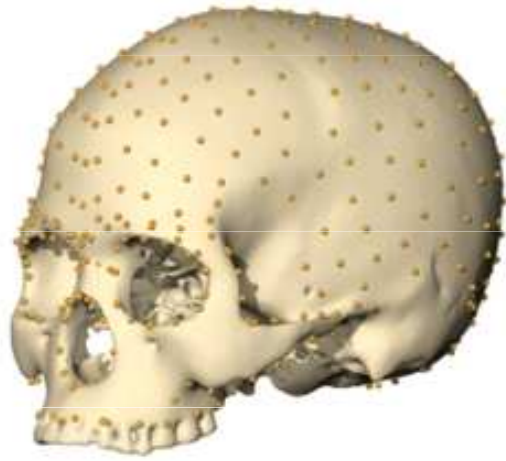
a

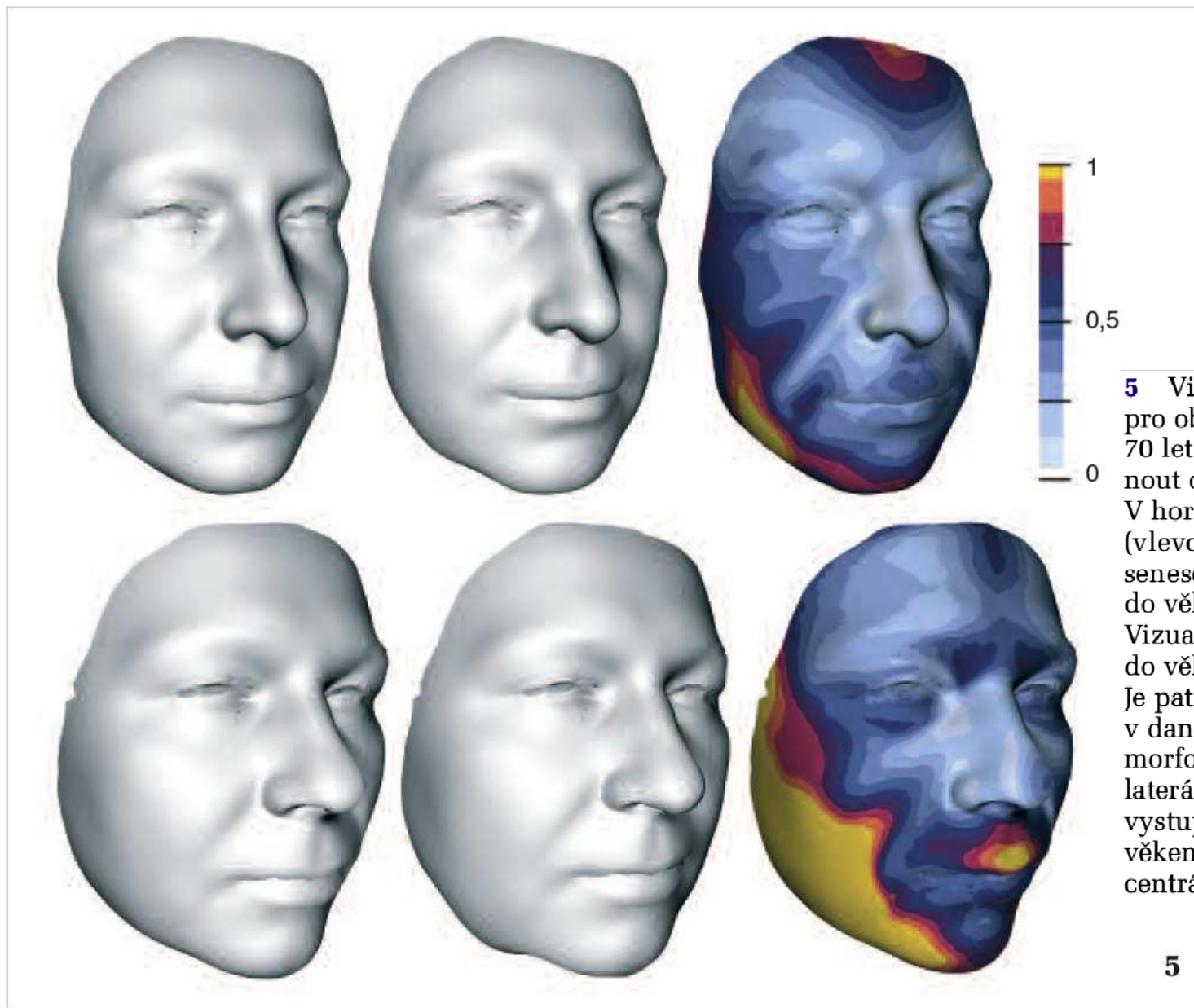


b

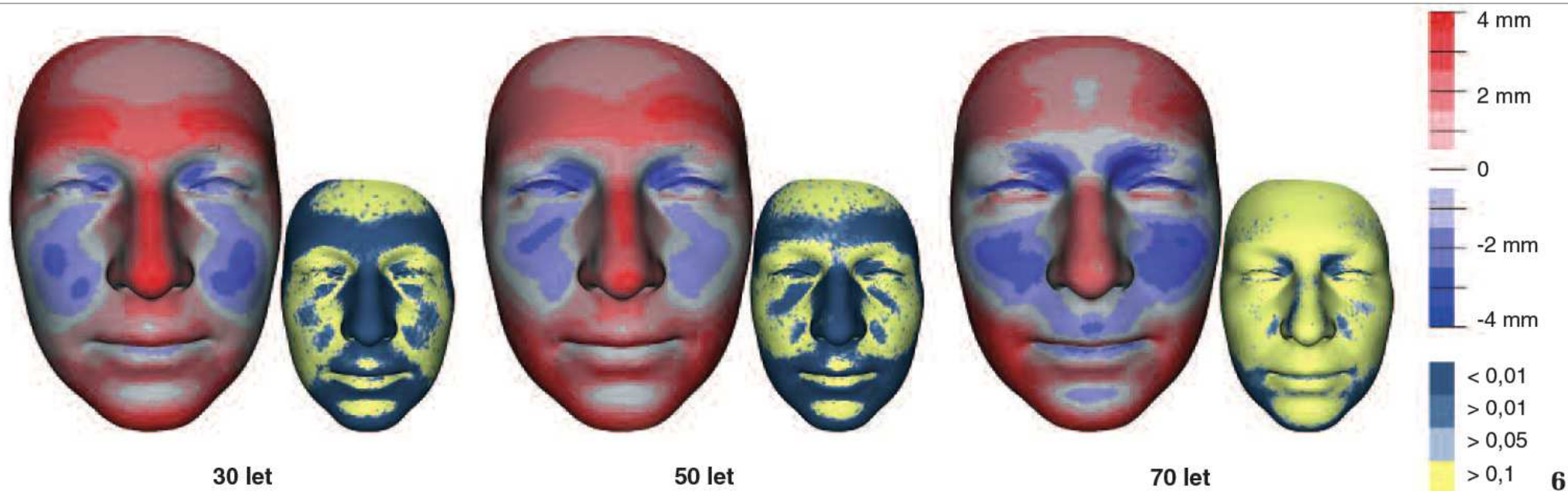
Craniometric Points, Lateral View





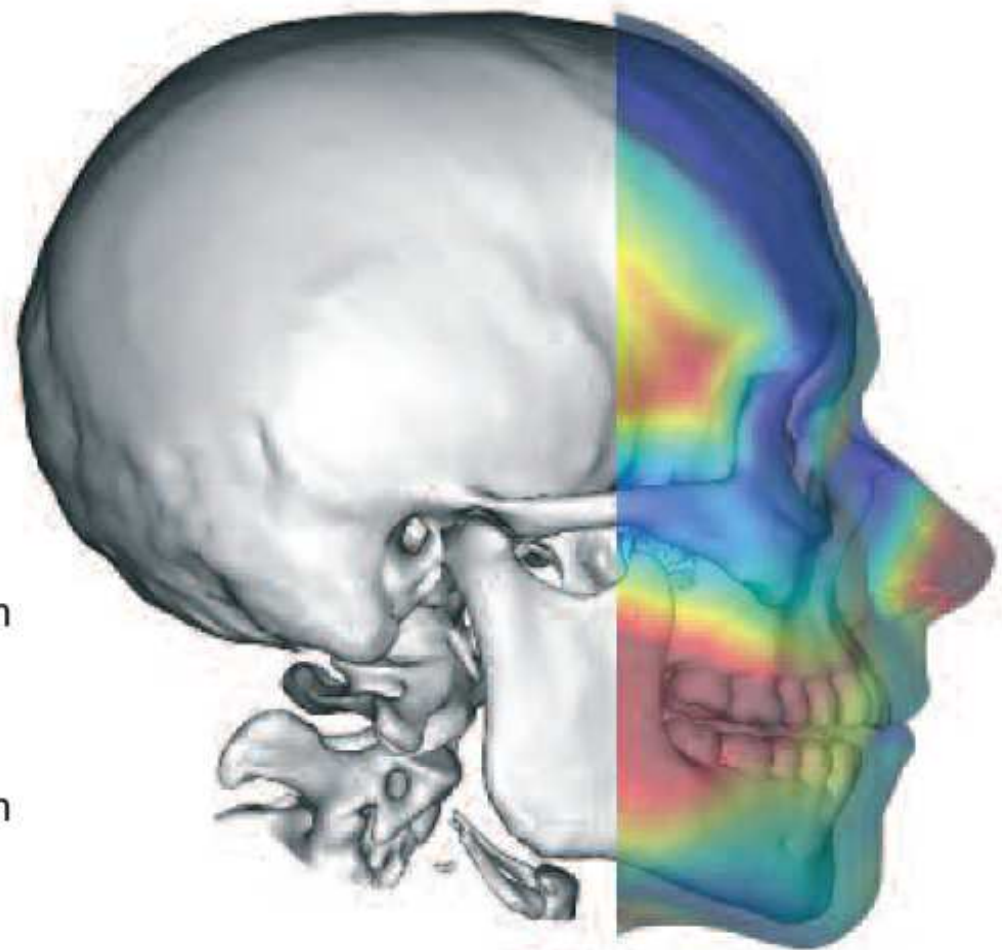
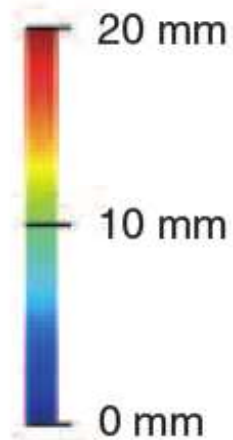
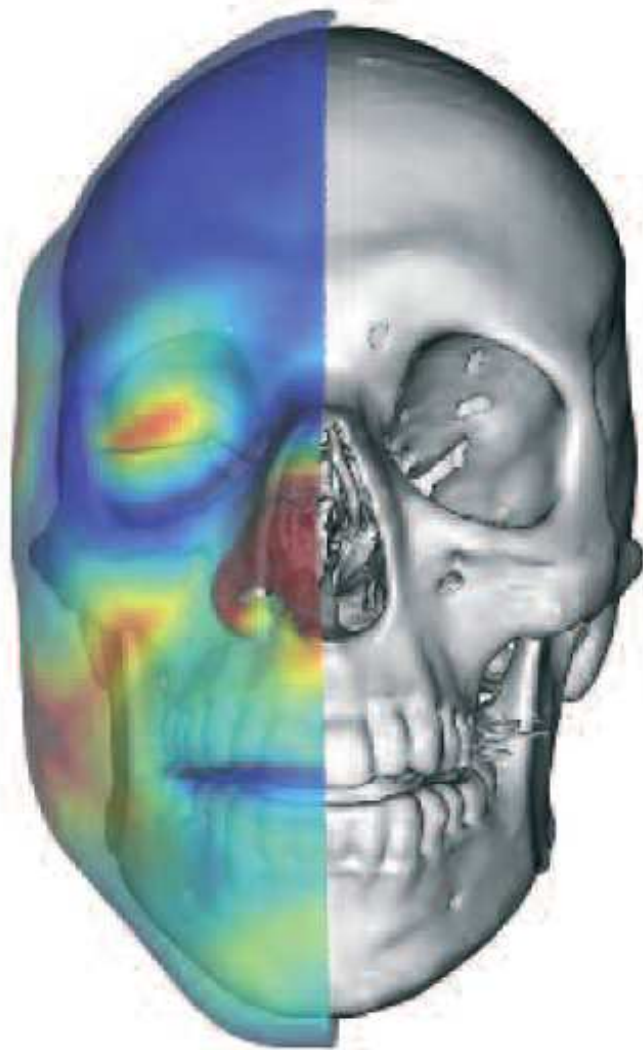


5 Vizualizace predikčního algoritmu pro období dospělého věku od 20 do 70 let, podle kterého lze omladit/zestárnout obličej v daném věkovém rozmezí. V horní řadě obličej ženy ve věku 29 let (vlevo), který podle algoritmu senescence žen (vpravo) zestárnul do věku 70 let života (uprostřed). Vizualizaci 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.



6 Vizualizace vývoje sexuálního dimorfismu v průměrném věku 30, 50 a 70 let. Větší obličej ukazuje znaky pohlavně dimorfní ve prospěch mužů (červená barva) a žen (modrá). Menší obličej zobrazuje 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 obličej se s věkem výrazně snižuje.

Velemínská, Dupej, Živa 5/2016

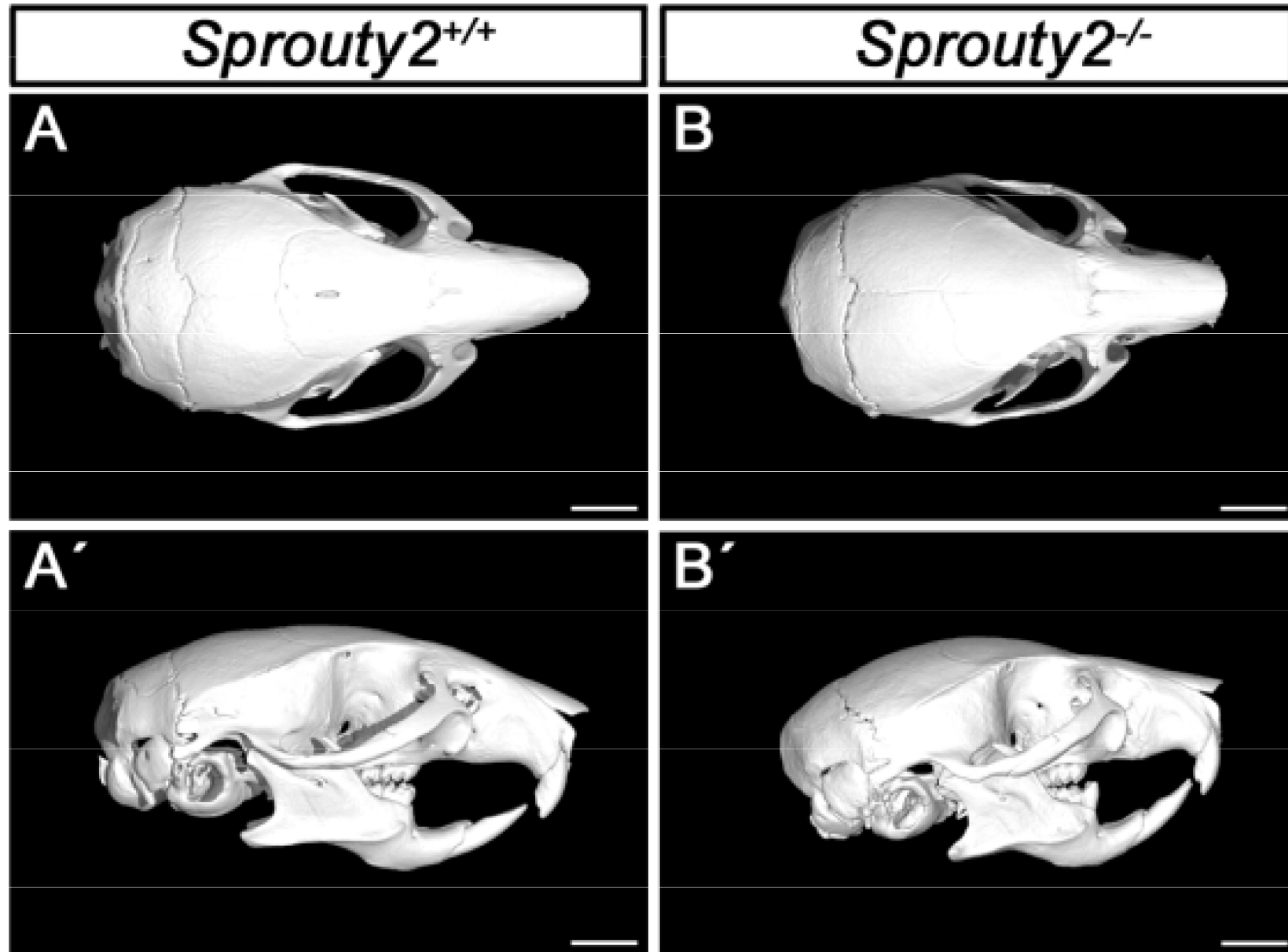


7 Segmentace lebky na základě snímků z počítačové tomografie hlavy člověka. Měkké tkáně jsou znázorněny průsvitně jen na části lebky a pomocí barevné mapy je odstupňována tloušťka měkkých tkání, zásadní pro rekonstrukci obličeje podle lebky (vlevo pohled zepředu, vpravo ze strany).

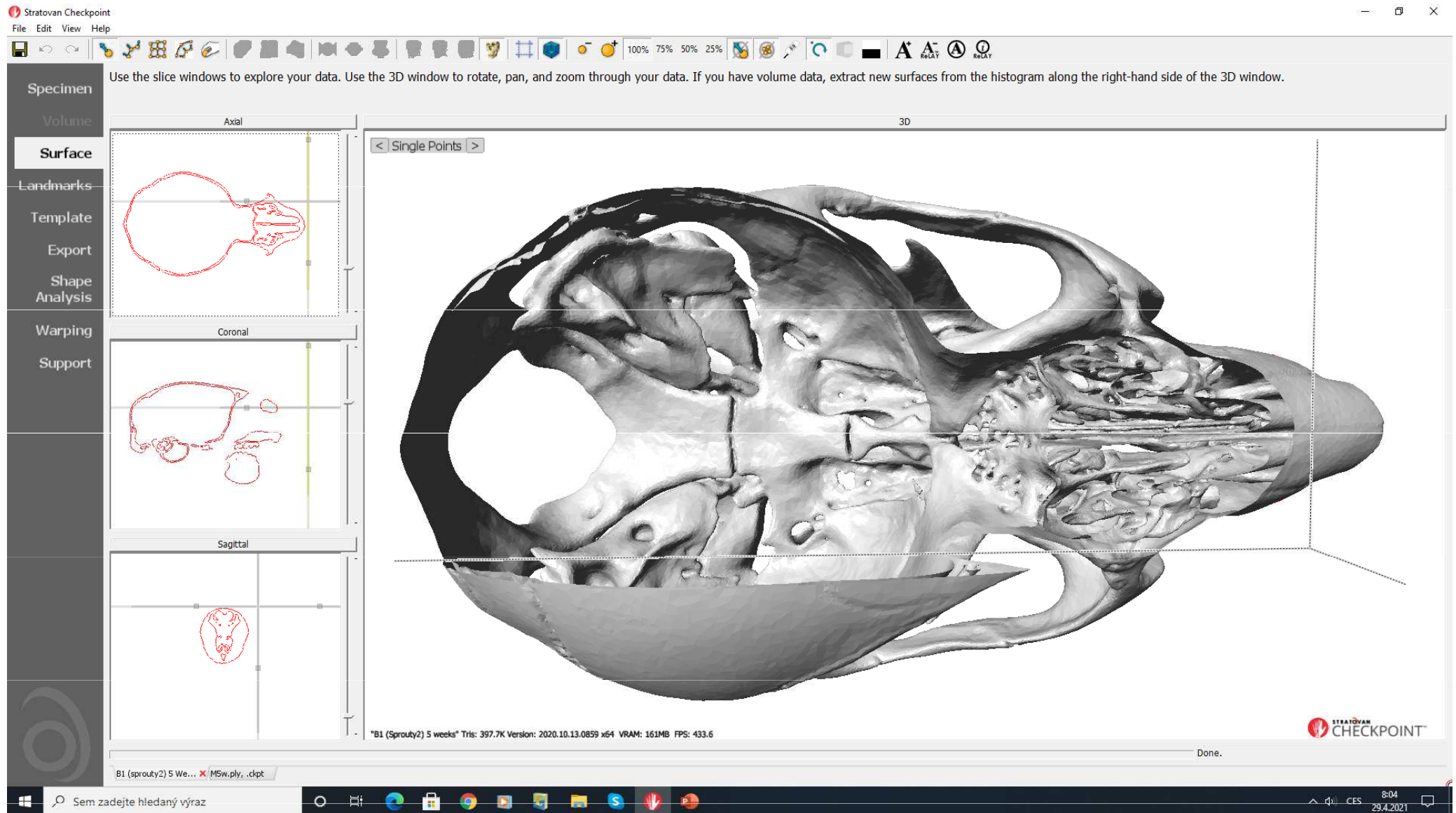
Všechny orig.: J. Dupej a J. Velemínská

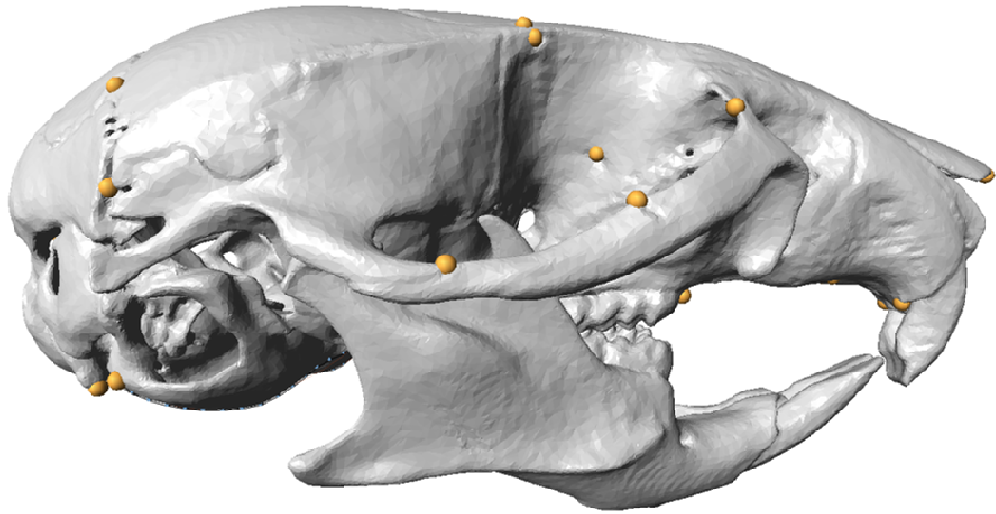
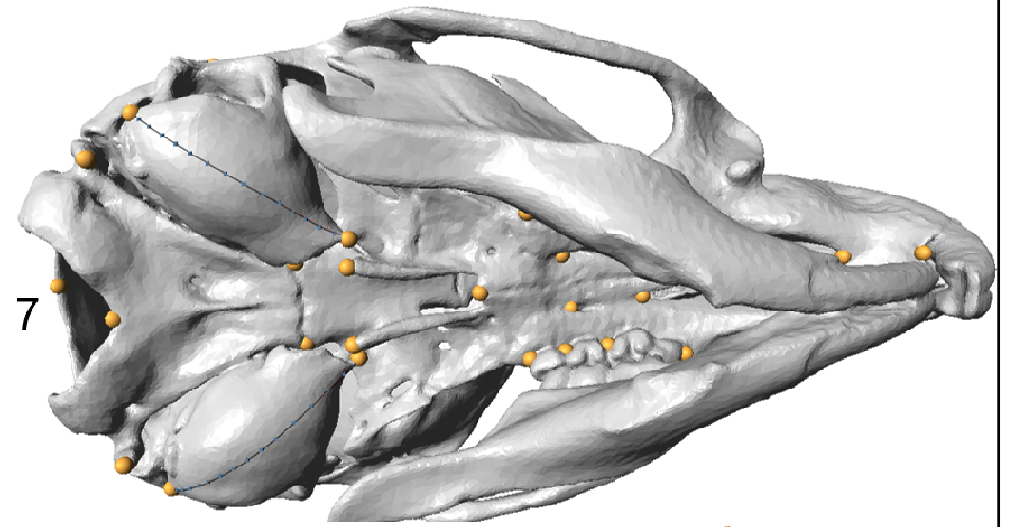
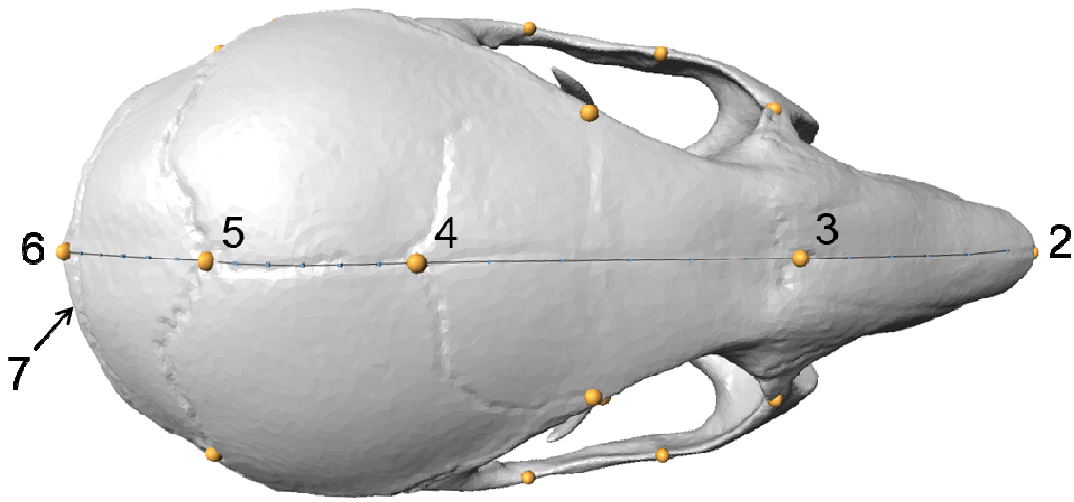
7

Example: knockout of gene *Sprouty*



Program Checkpoint (Stratovan):

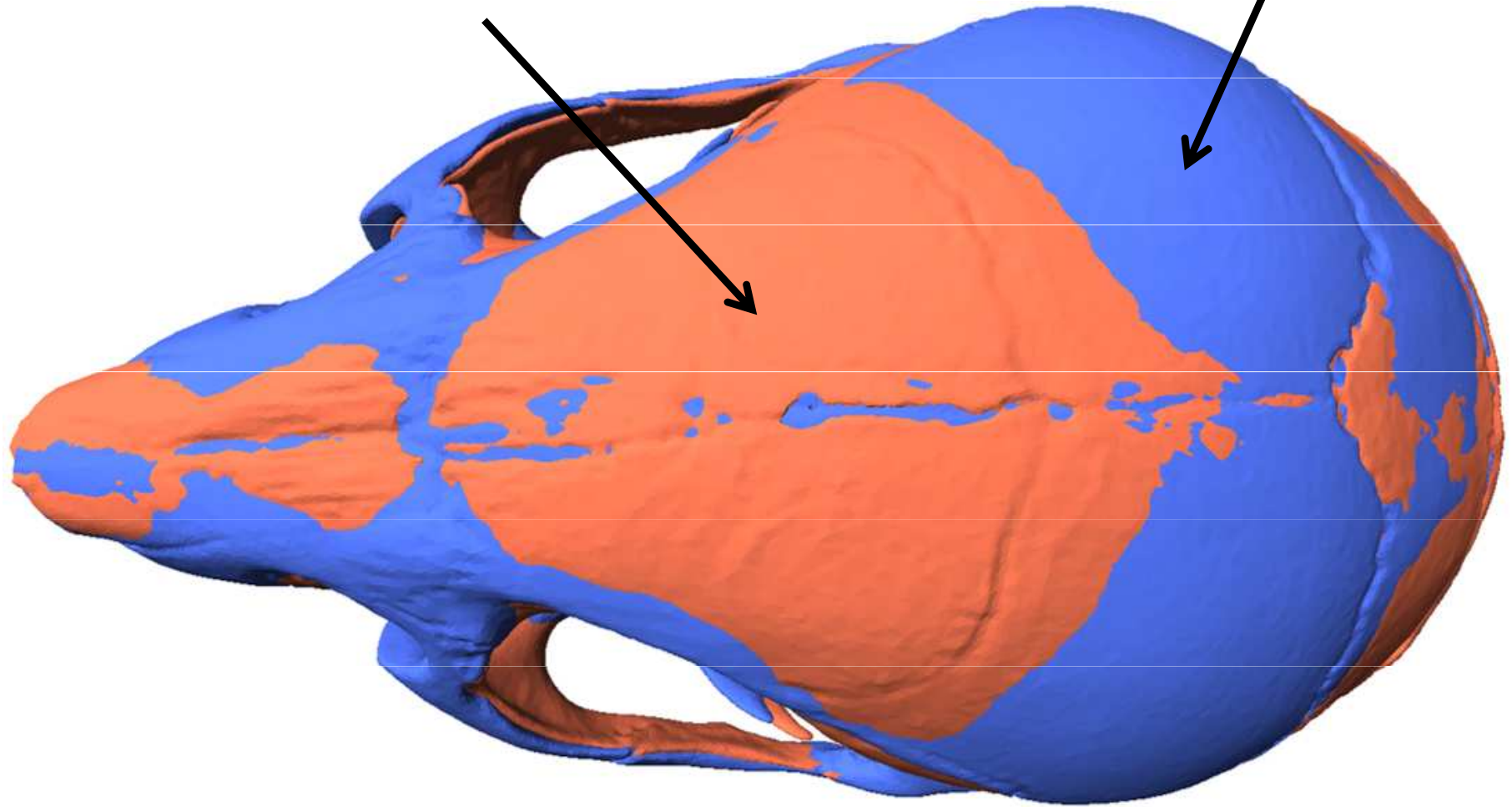




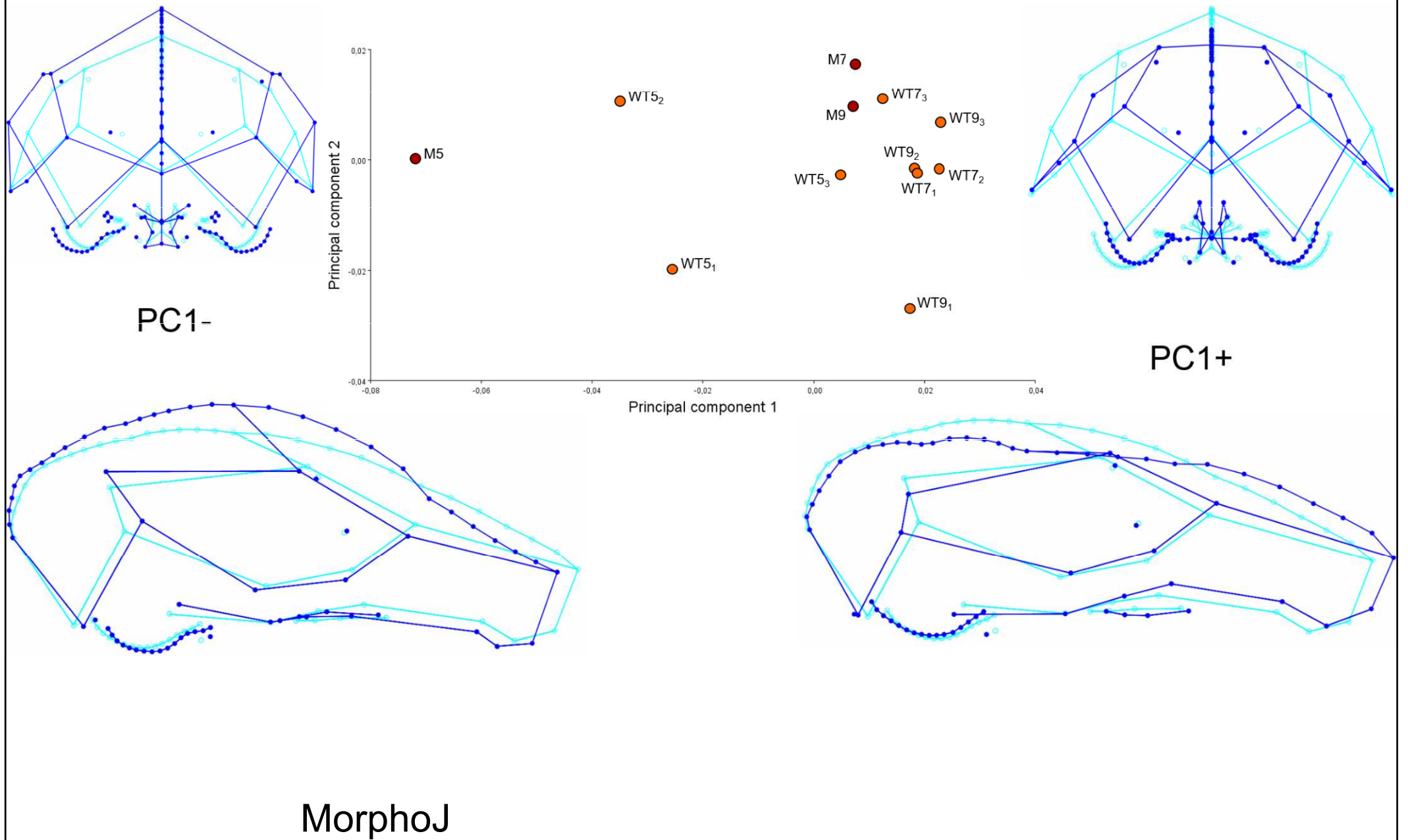
Comparison of wild-type and mutant mouse, 5 weeks

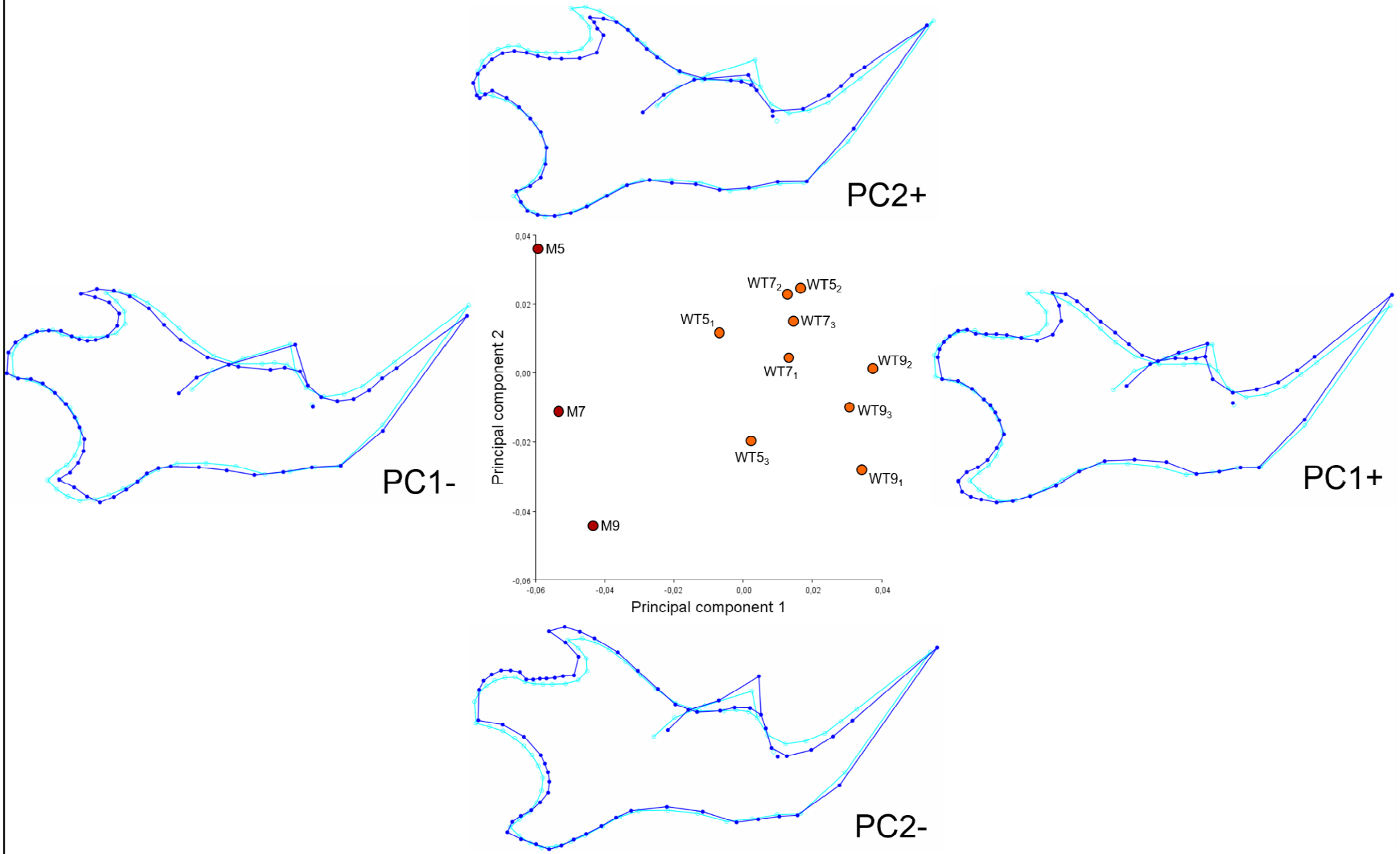
red = mutant

blue = wild-type

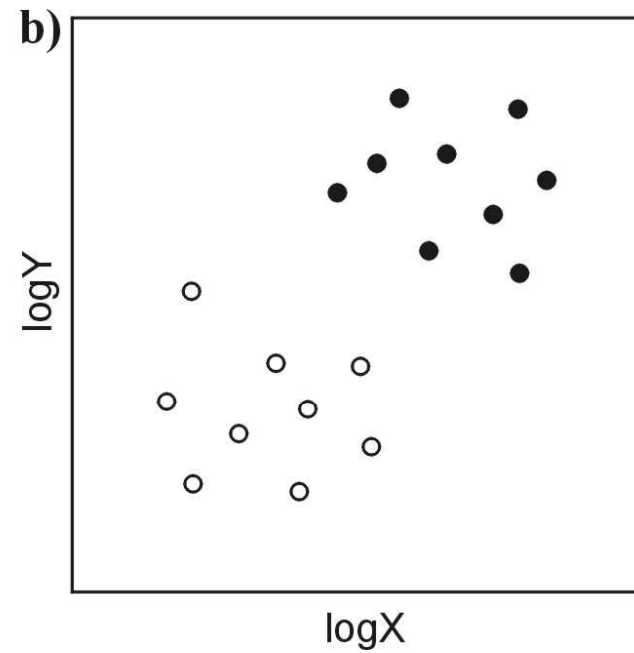
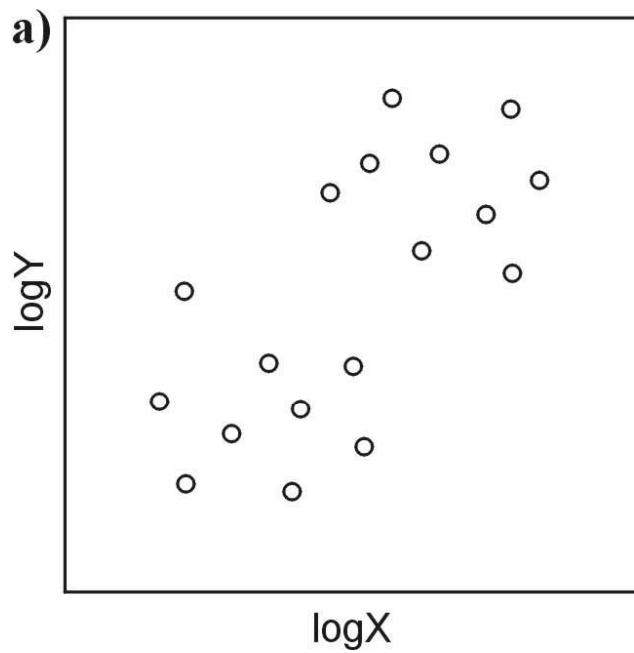
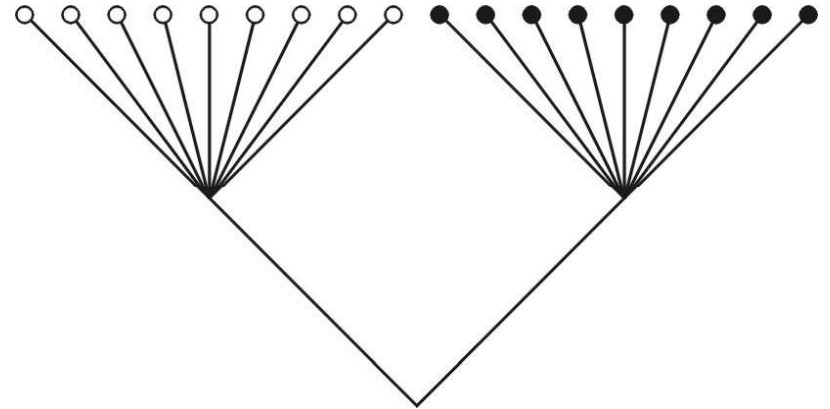
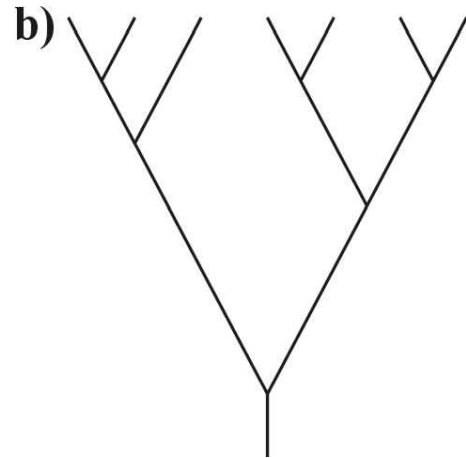
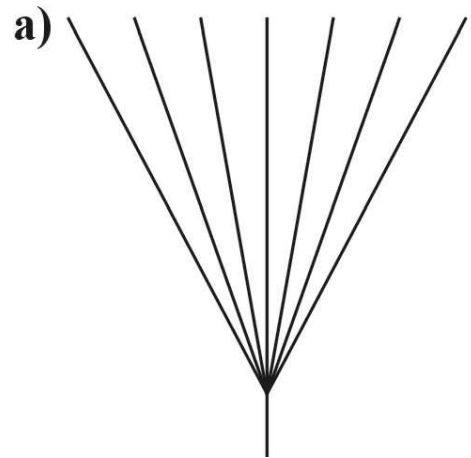


Stratovan Checkpoint



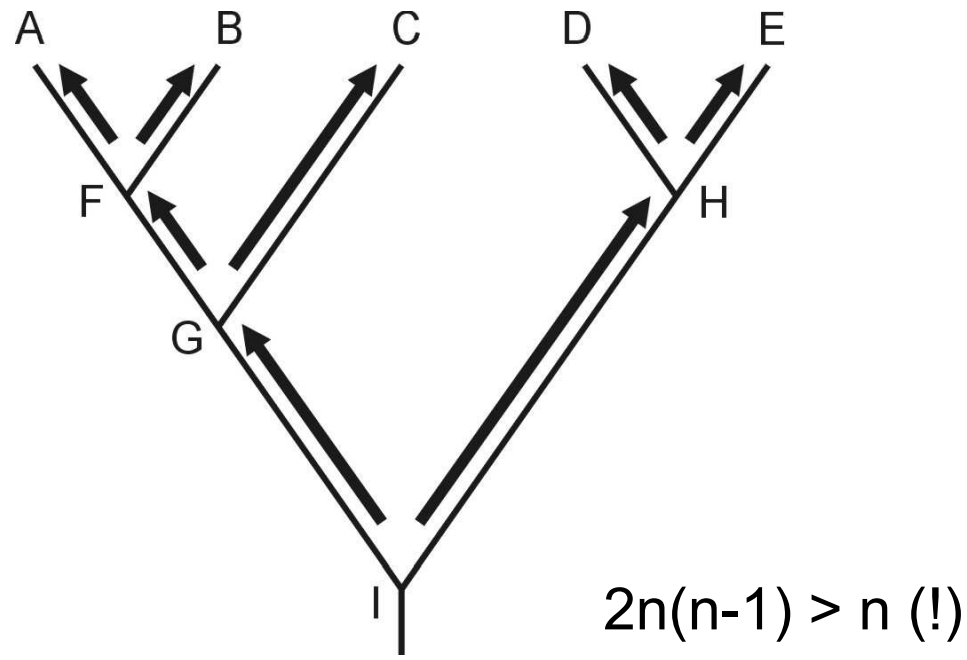


Comparative analysis



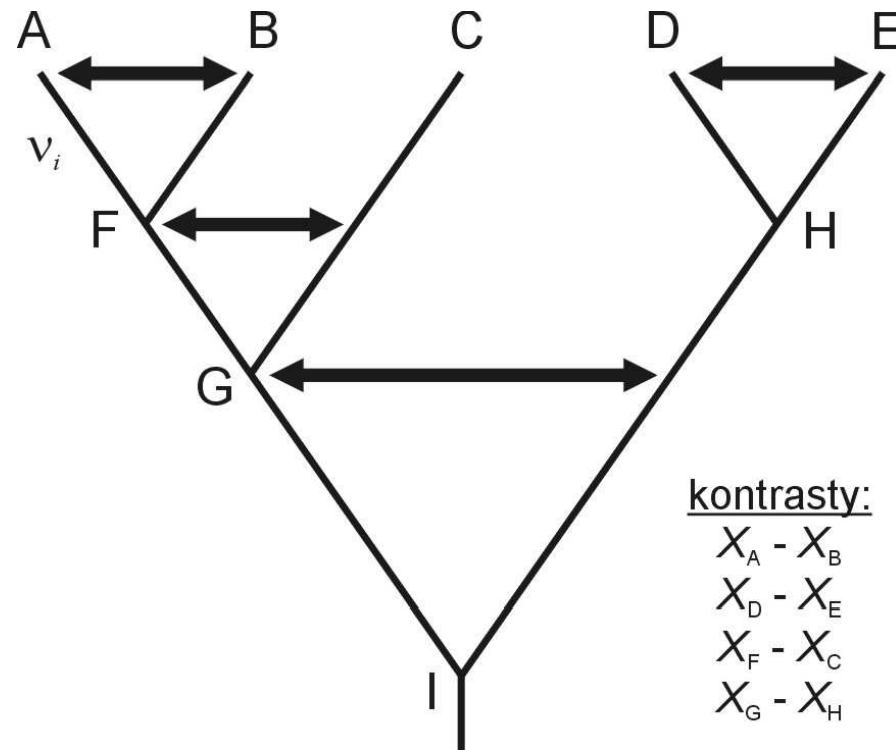
phylogenetic inertia

inference of changes along branches – usually parsimony



⇒ in fact, it doesn't solve the problem of phylogenetic independence!

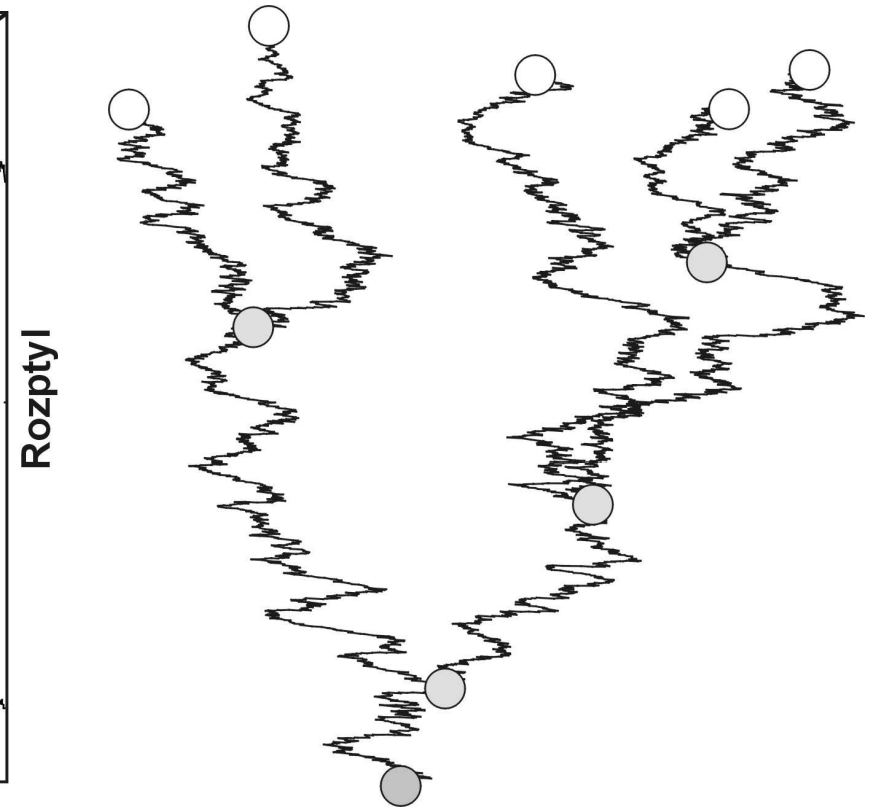
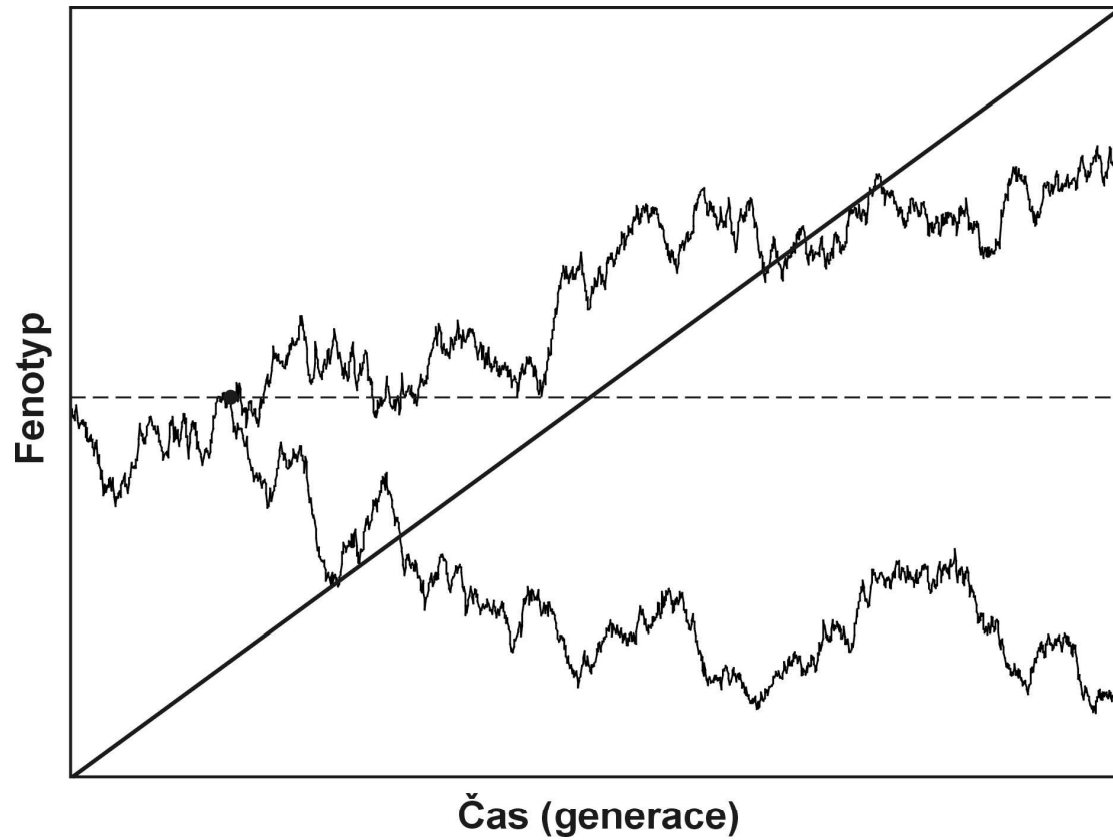
Independent contrasts



assumption: Brownian motion!

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

Brownian motion model



Ornstein-Uhlenbeck model

