

Genetic Determinants of Orofacial Shape and Relationship to Cleft Lip/Palate



Richard Spritz & Stephanie Santorico
University of Colorado Denver



Fernando Pardo-Manuel de Villena
University of North Carolina

Benedikt Hallgrímsson / Mange Manyama
University of Calgary / BUCHS (Tanzania)

Ophir Klein
University of California San Francisco

Washington Mio
Florida State University

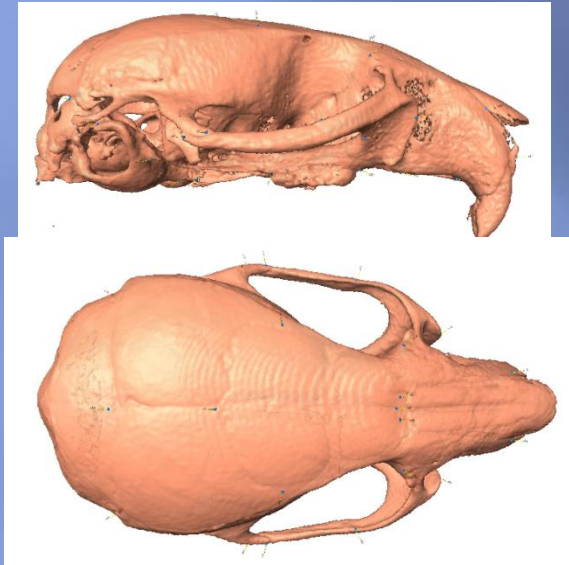


Specific Aims

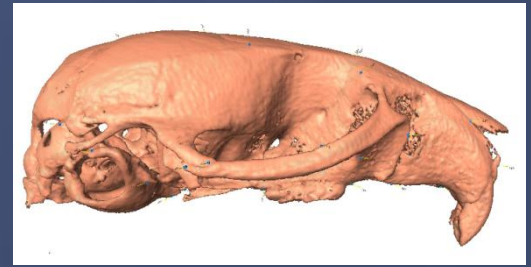
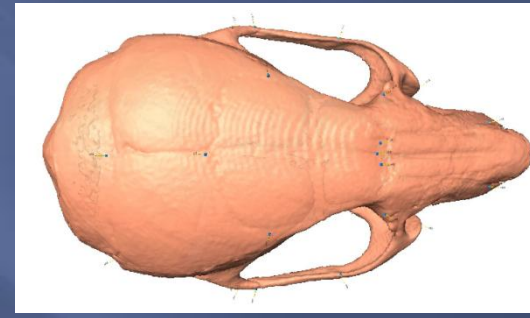
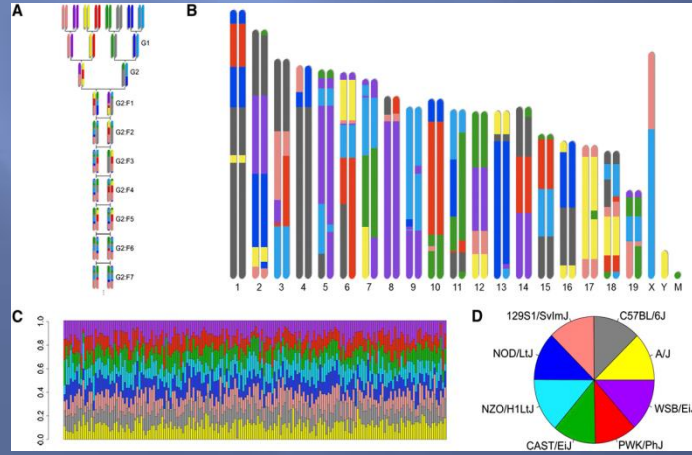
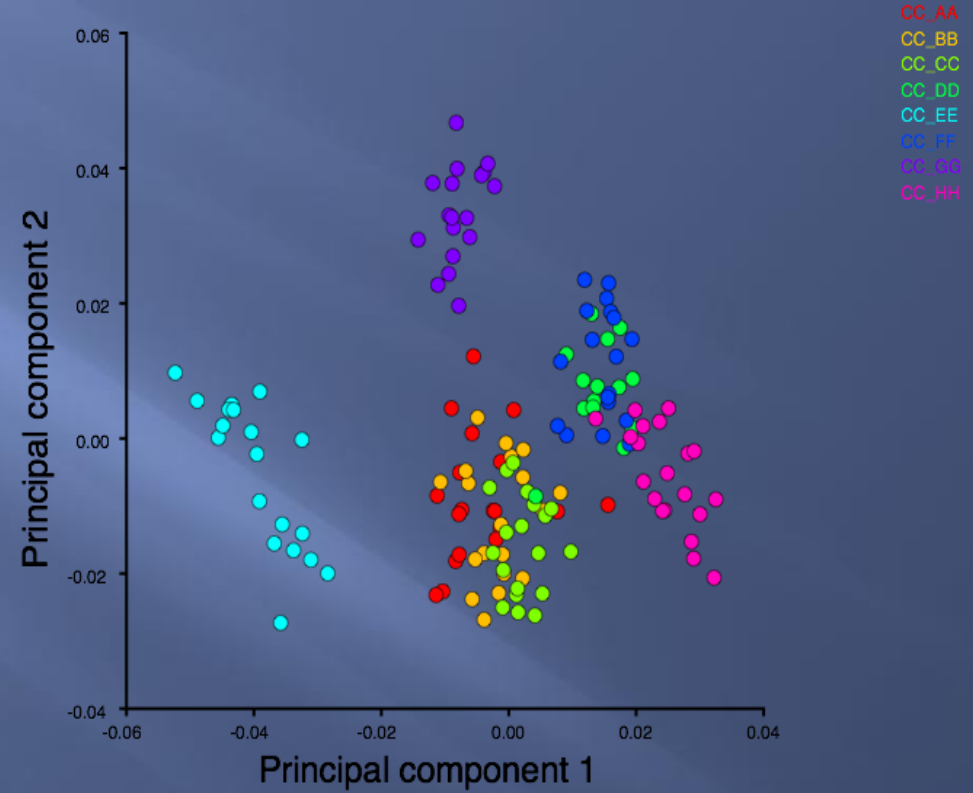


1. Determination of heritable midfacial morphometric variation among 8 mouse strains of the mouse Collaborative Cross
2. Fine-map major quantitative trait loci (QTLs) responsible for those major heritable midfacial morphometric phenotypes using the mouse Collaborative Cross > Instant mouse models
3. Carry out GWAS of four homologous human midfacial morphometric phenotypes in Tanzanian children (n=3700). Carry out replication study in Tanzanian children (n=2600). [Test association of confirmed loci in USA EUR, USA Hispanic/Latino, (Asian) populations]

The Collaborative Cross



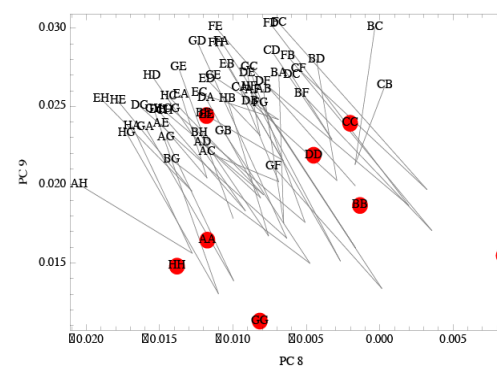
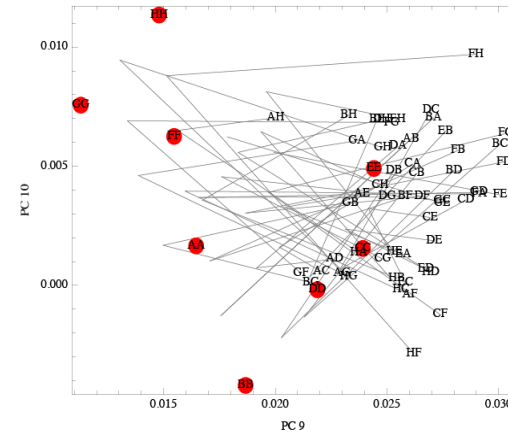
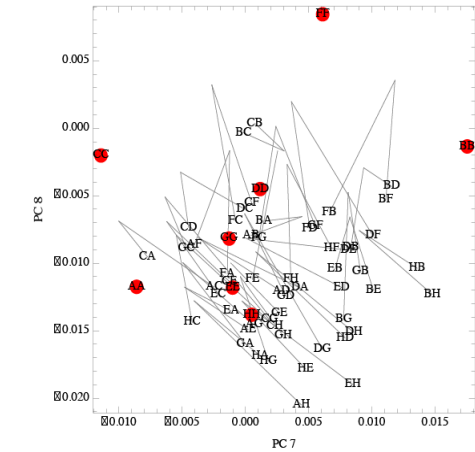
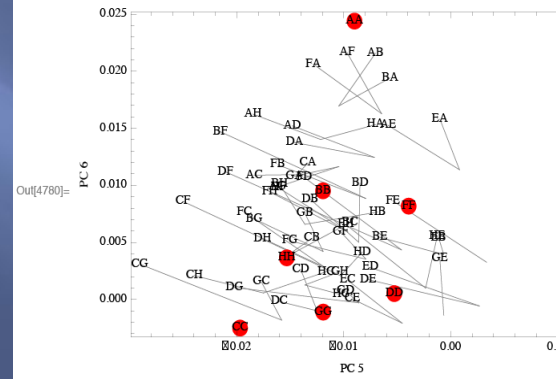
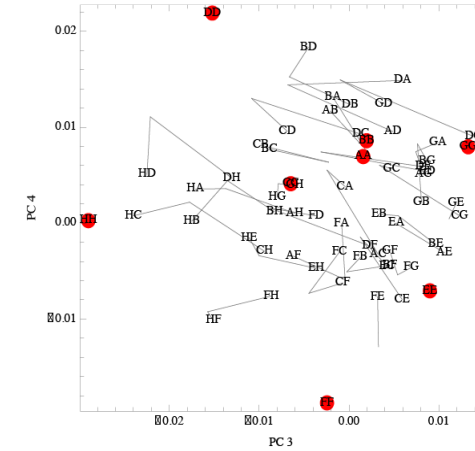
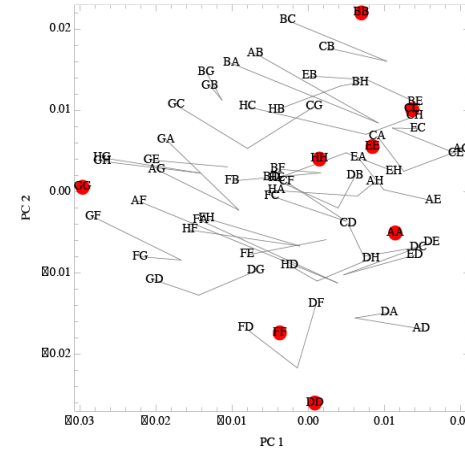
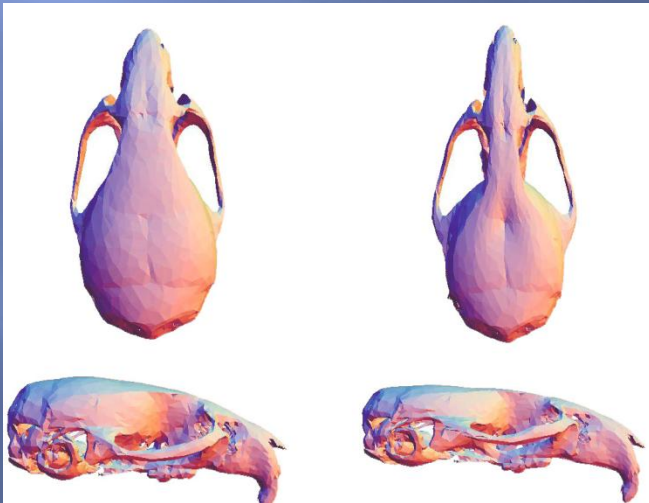
Analysis of the Parental and F1 Collaborative Cross Mice



Variation tends to be highly structured

F1 crosses deviate non-randomly from the mid-parental values

Total Dominance Variance (magnified 7X)



Project Collection Sites



Tanzania 3D Photo / Data / Saliva Collection



Summary of Enrollment

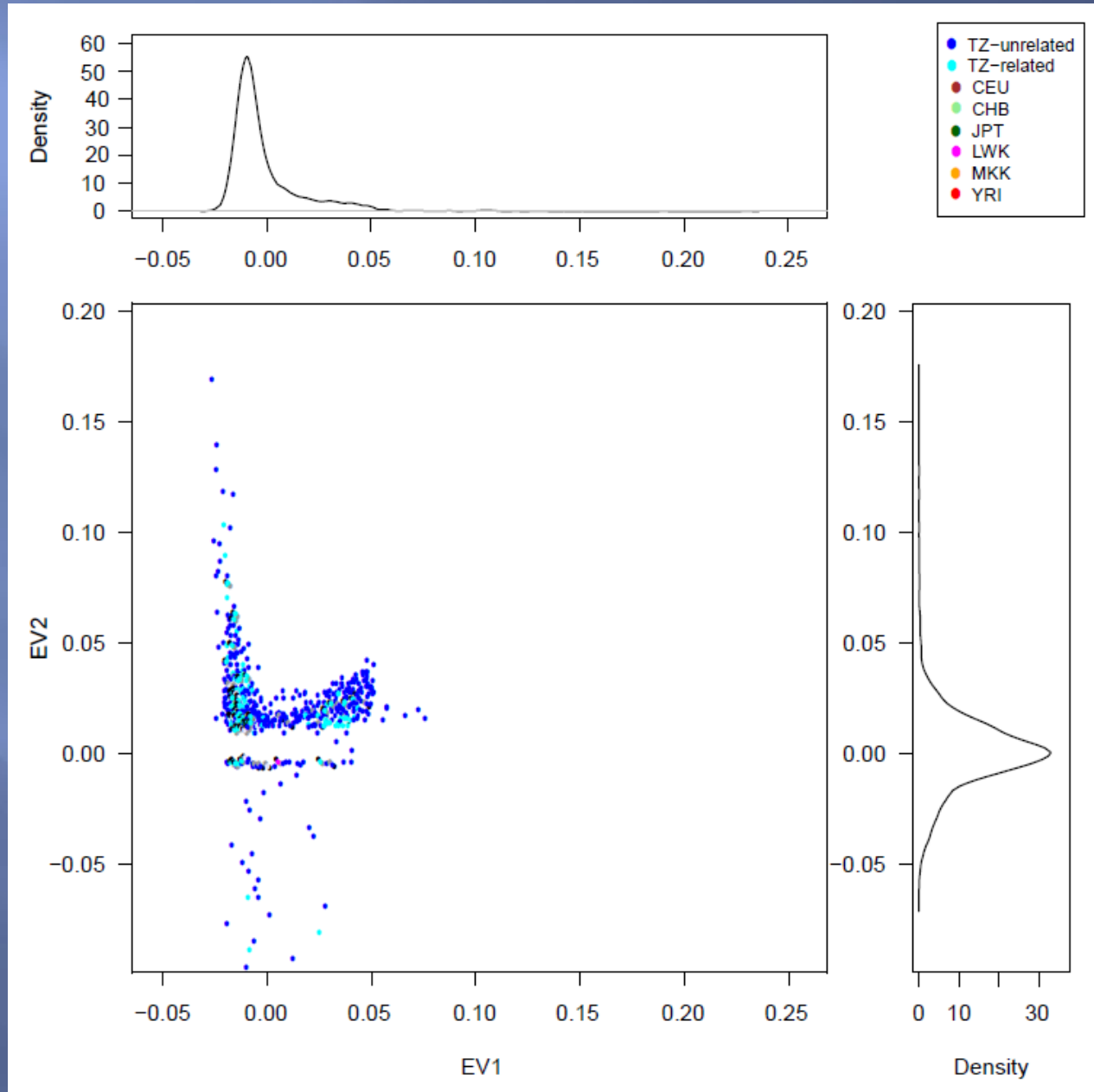


| | Opened | Status | Enrollment |
|-------------------------|---------------|---------------|-------------------|
| Tanzania African | 5/09 | Closed | 7336 |
| USA | 5/10 | Closed | |
| Caucasian | | | 659 |
| Hispanic | | | 138 |
| Asian | | | 32 |
| Other/mixed | | | 7 |

Genotyping Summary

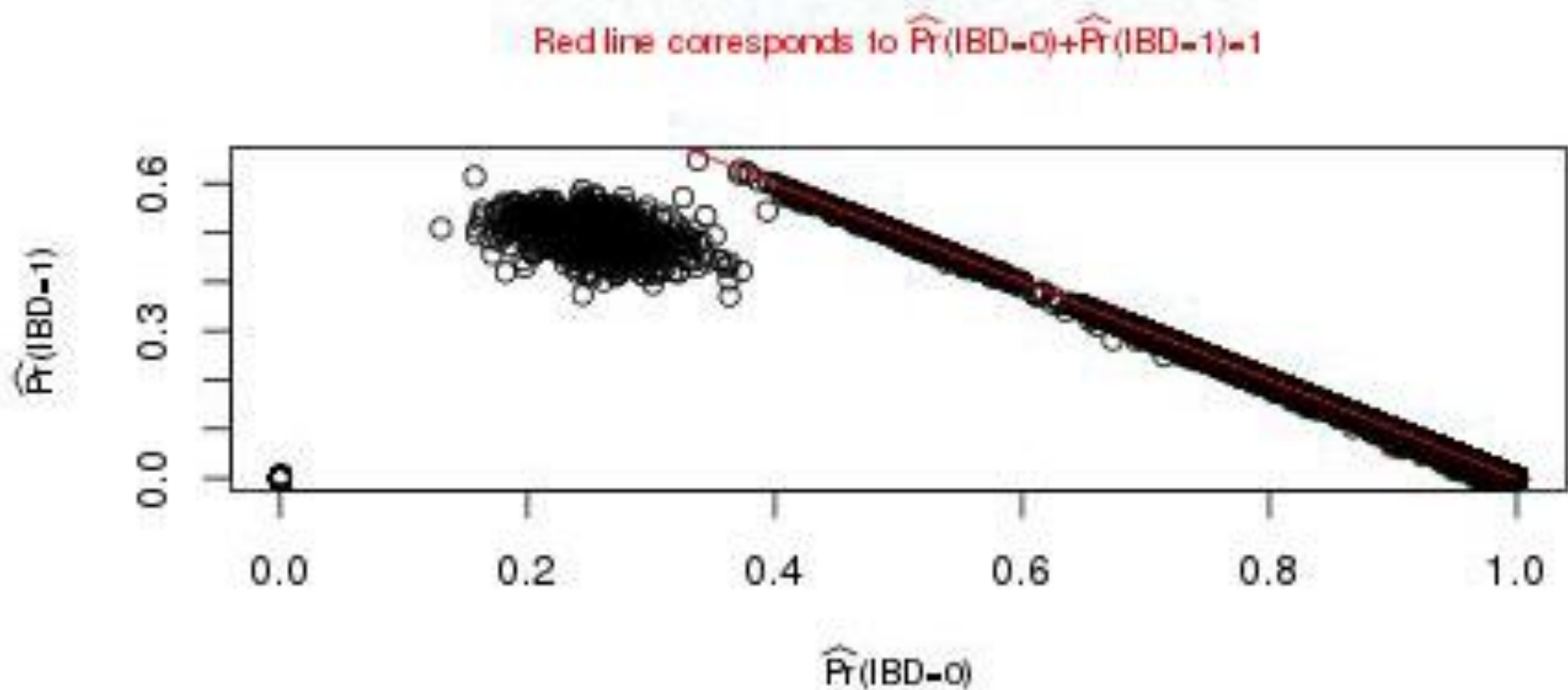
1. 3701 Tanzanian Bantu children selected for genotyping (including duplicates) at CIDR using Illumina HumanOmni25Exome-8v1_A array (2,567,845 variants); 3671 passed CIDR QC
2. 3557 unique Tanzanian children genotyped, data sent to CBS at Univ Washington (Cathy Laurie)
3. Median call rate 99.9%; error rate $\sim 1.4E-05$ (duplicates)
4. Two XXY individuals identified and filtered
5. Final subjects 2199 unrelated individuals ($KC < 0.98$; = half-sibs) + 563 families (424 w/2, 97 w/3, 42 w/ ≥ 4 relatives) discovered by analyses of relatedness
6. Minimal other population structure, few outliers, inbreeding coefficient very low
7. No batch effects or other significant artifacts; 99.8% of SNPs pass QC filters
8. 28% SNPs MAF $< 2\%$ for autosomes, 0.5% for X-chr
9. 17.9% of SNPs filtered (mostly as monomorphic)
10. Genomewide imputation (Impute2) pending

Principal Components Analysis of Population



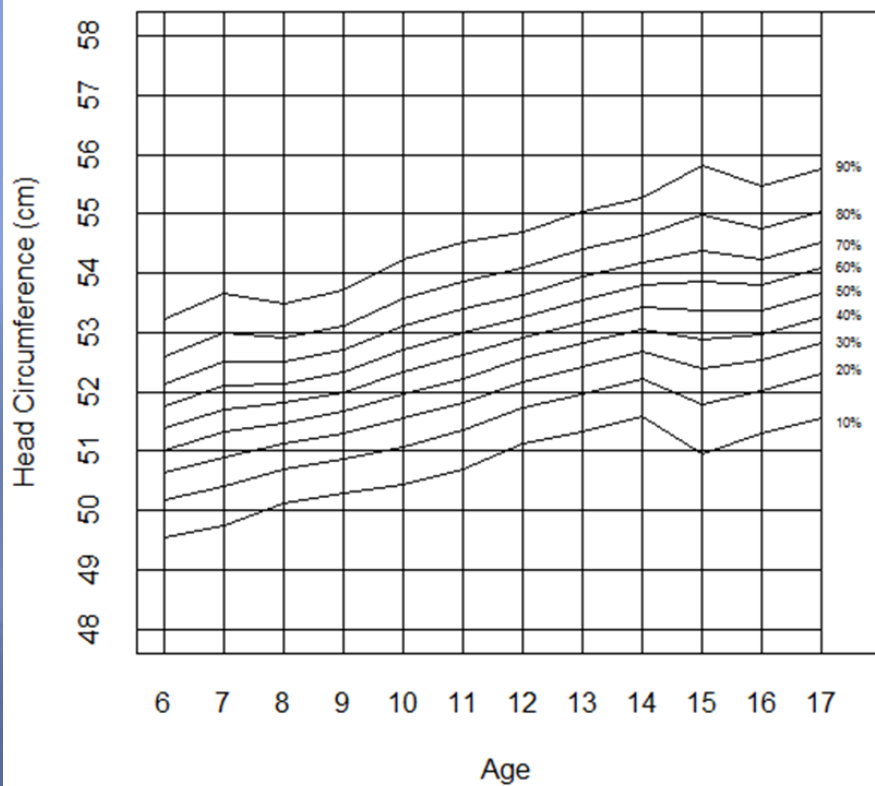
Analysis of Relatedness in Population

IBD Estimates

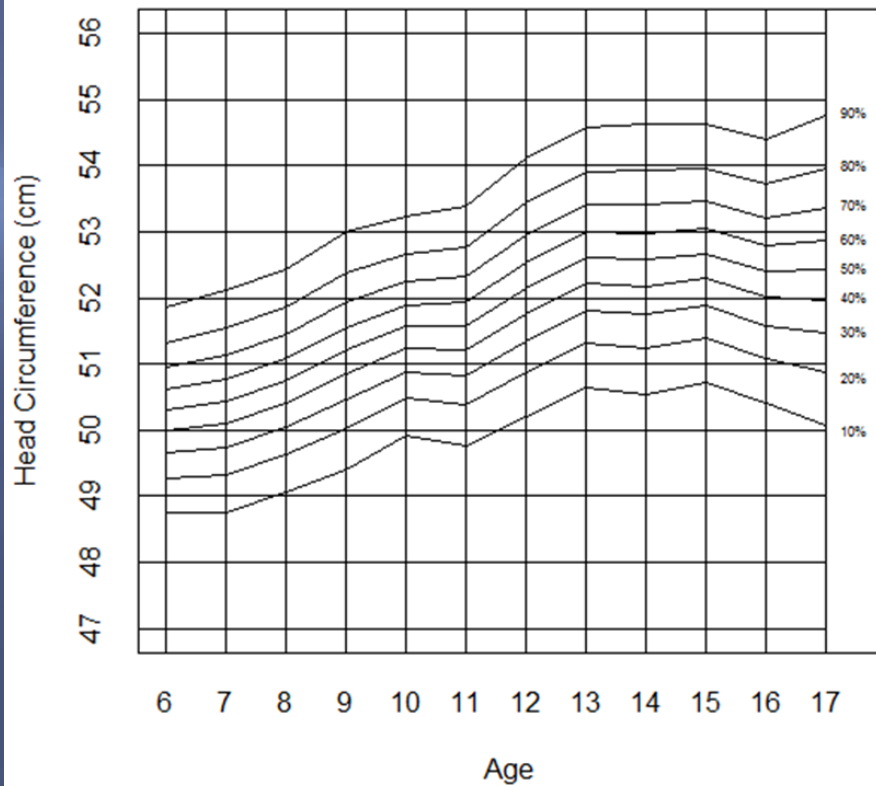


Head Circumference Data

Head Circumference for Tanzanian males Age 6-17



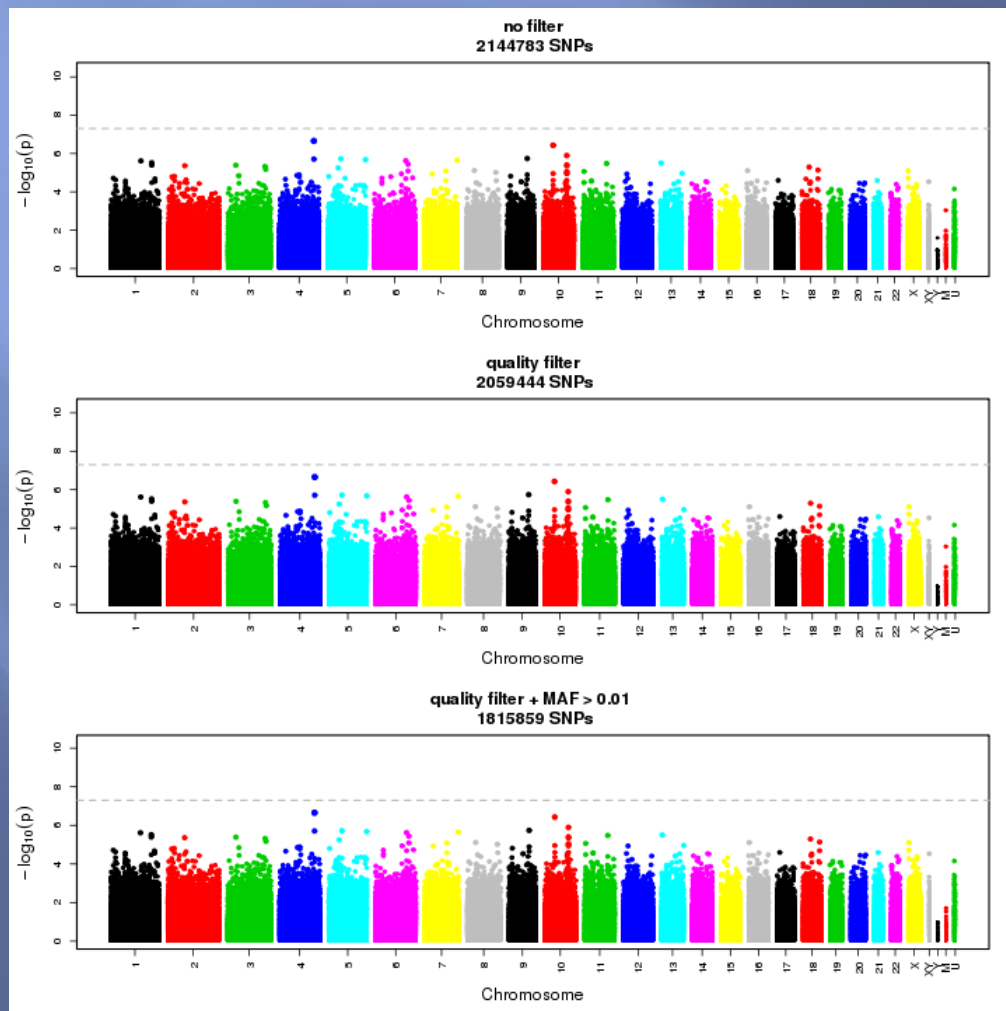
Head Circumference Tanzanian Females Age 6-17



Preliminary Association Analysis

Head Circumference (n=1906)

1. Two XXY individuals excluded
2. Four outliers excluded (probable measurement errors)



Top hits:

kgp6832430 chr10; $P=4.46E-07$

3 SNPs chr 4; $P=5.03E-07$

kgp6930620 chr6; $P=7.72E-07$

kgp6397840 chr5; $P=1.05E-06$

2 SNPs chr 10; $P=1.92E-06$

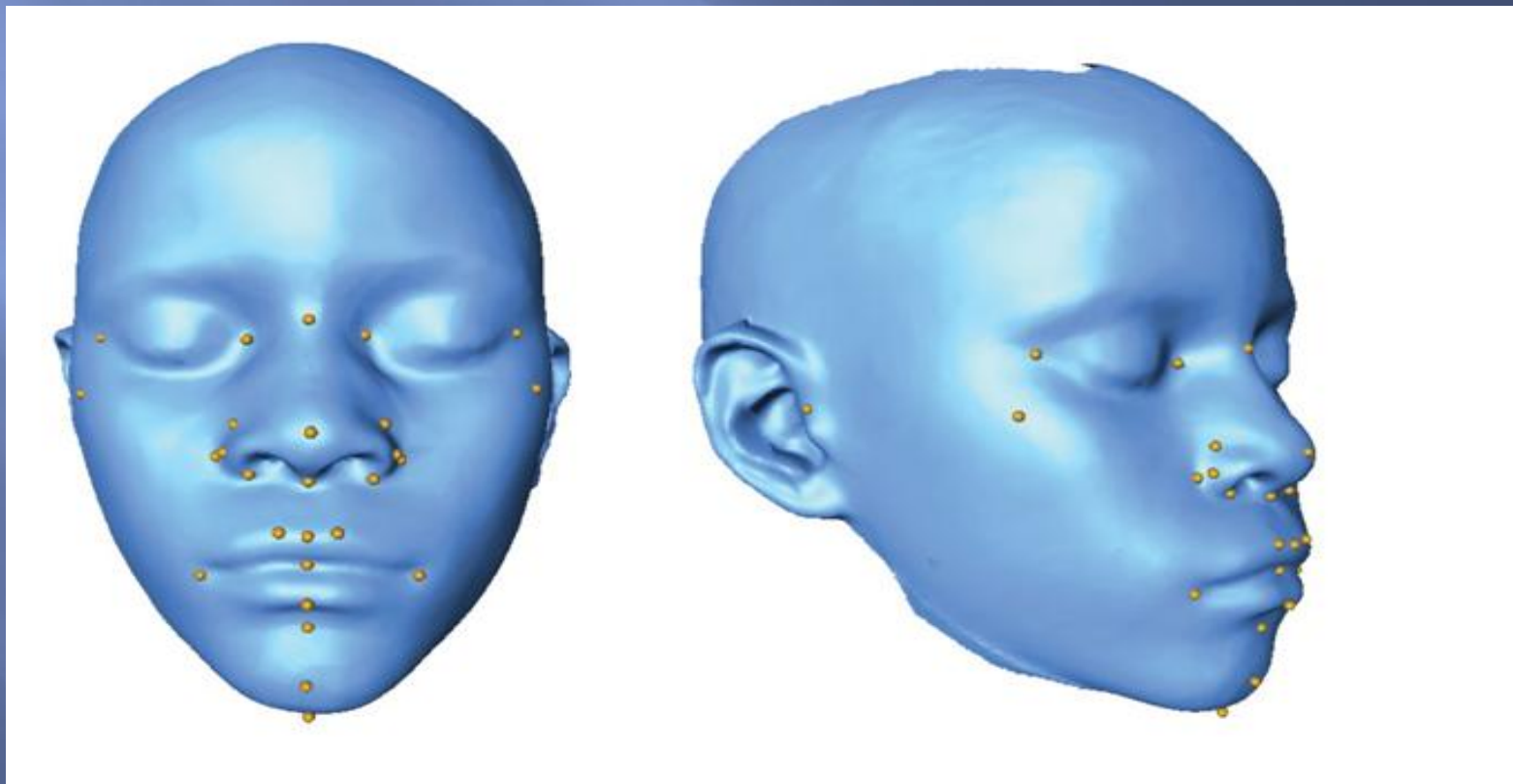
Phenotyping

A 3D camera captures six digitized facial scans of each individual



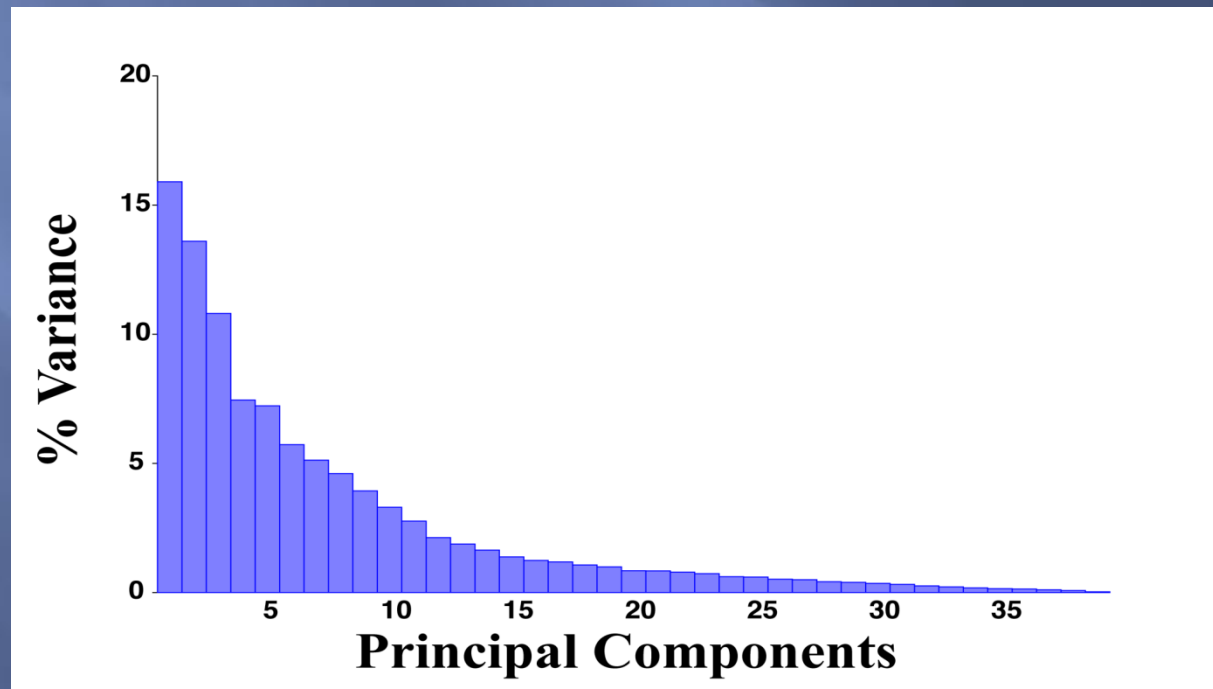
Phenotyping

- ▣ Scans are compiled into a single mesh object for landmarking
- ▣ Landmarks are placed for 29 common morphometric facial points



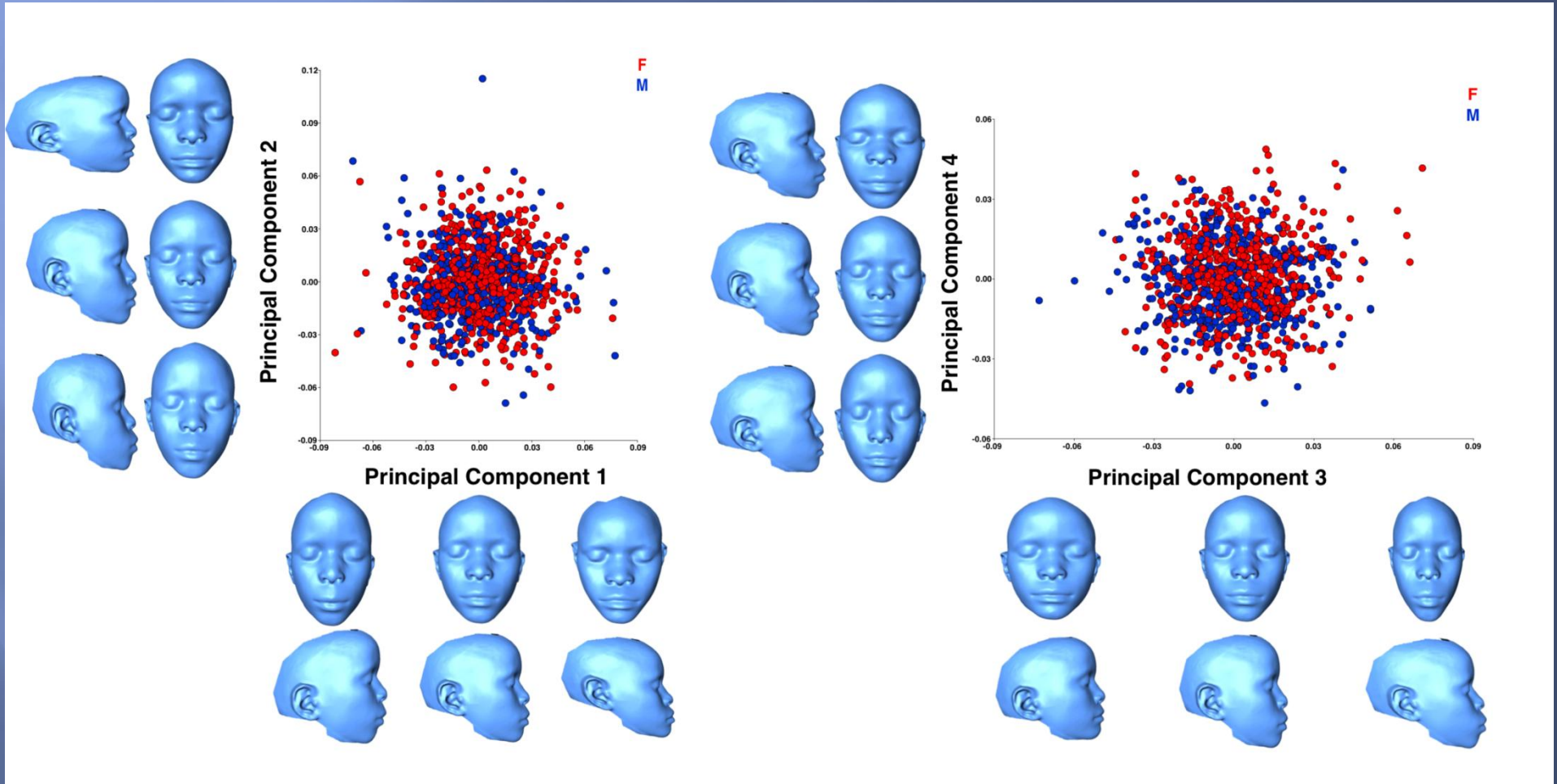
Principal Components Analysis

- ▣ X, Y, Z variable coordinates for each landmark
- ▣ Calculate the covariance between each set of variables
- ▣ The top principal components (PC) represent the axes of variation with the largest magnitudes



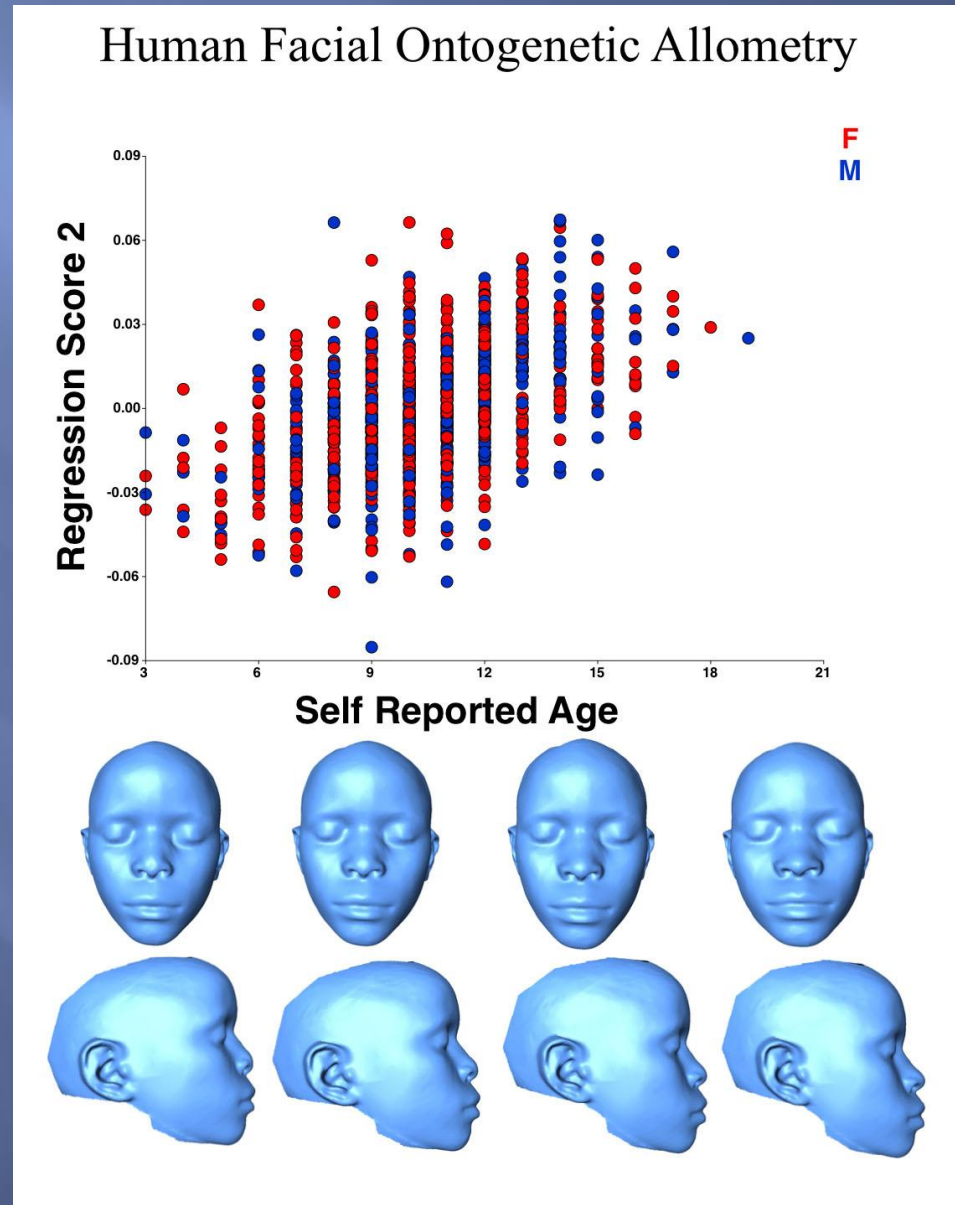
Preliminary Analysis

PCs 1-4; PC1 accounts for 15% of total variance in shape



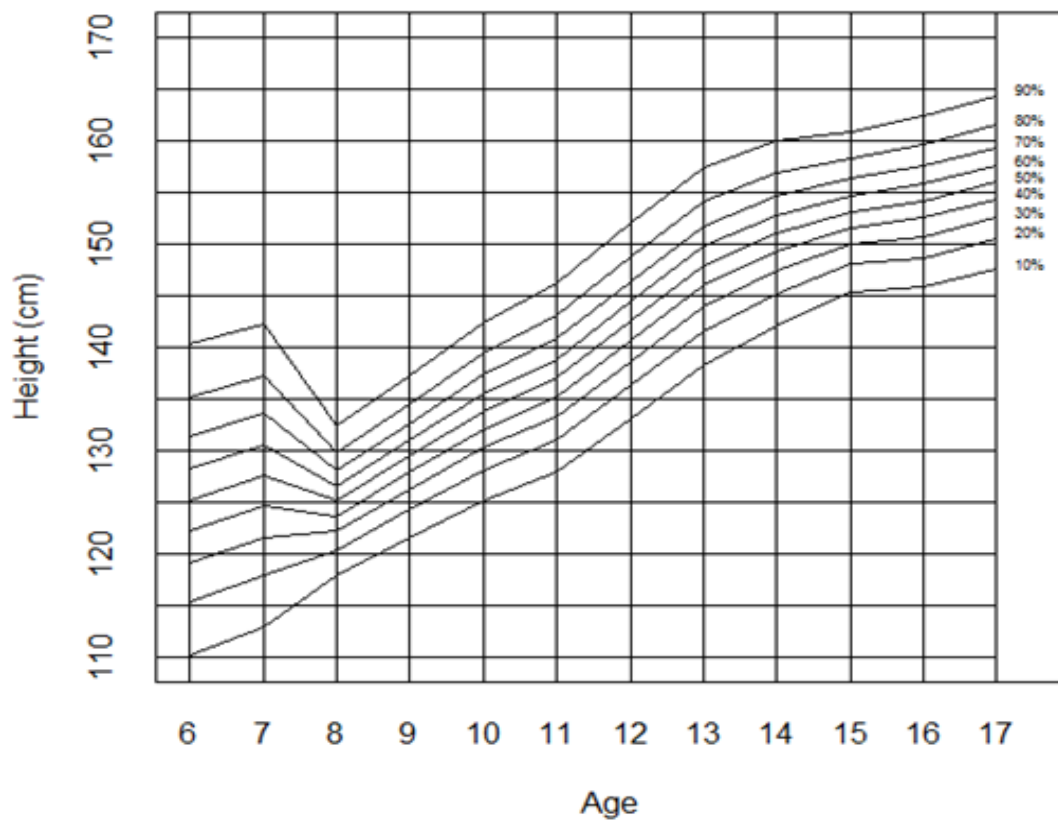
Multivariate Regression

Quantification of shape variation associated with age

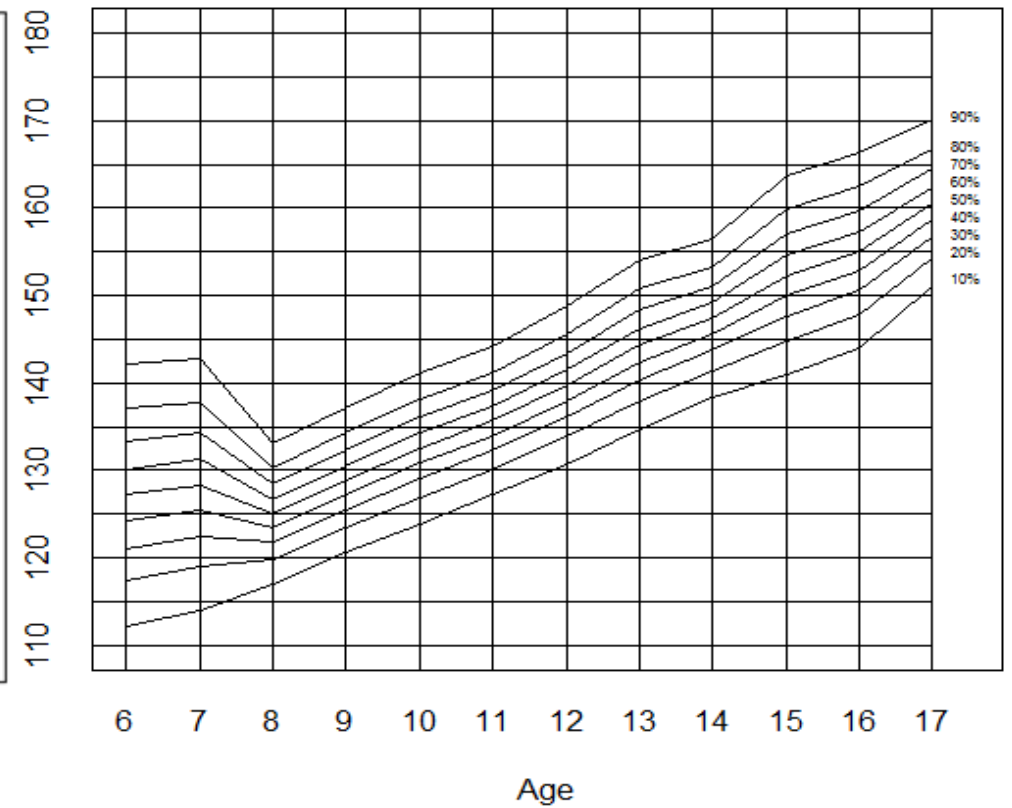


Tanzanian Growth Charts: Height-for-age

Females

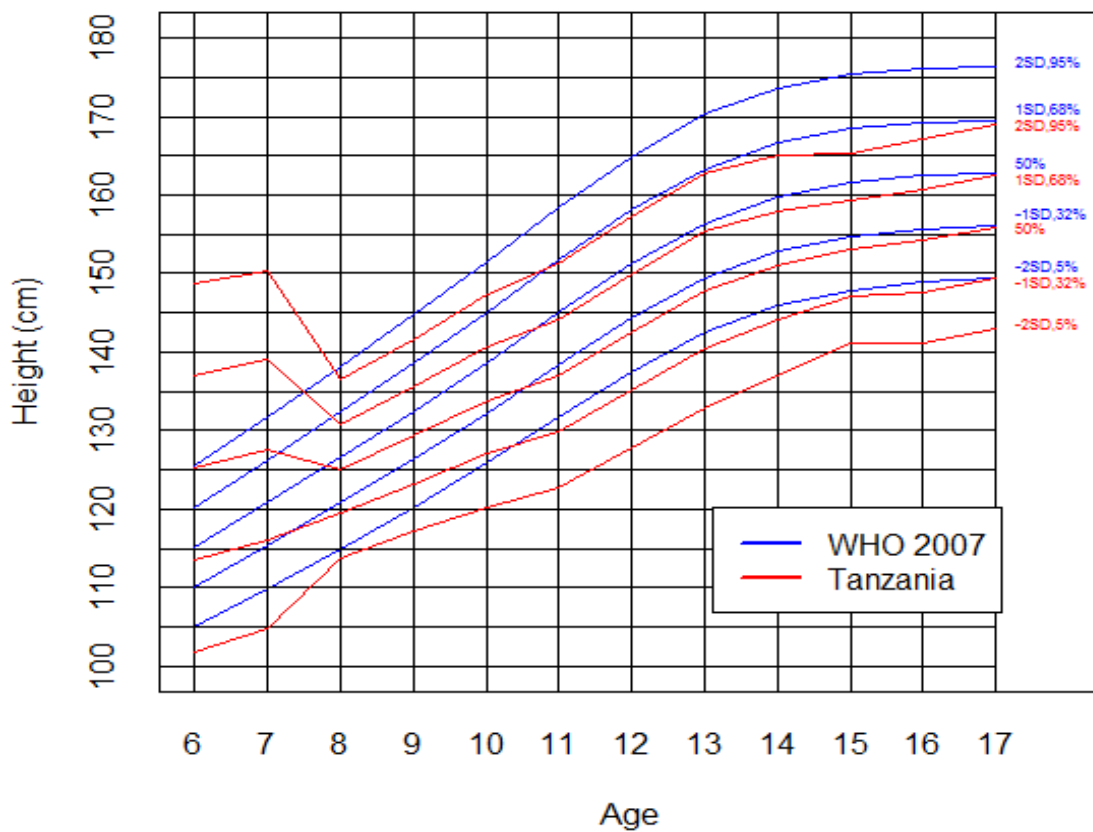


Males

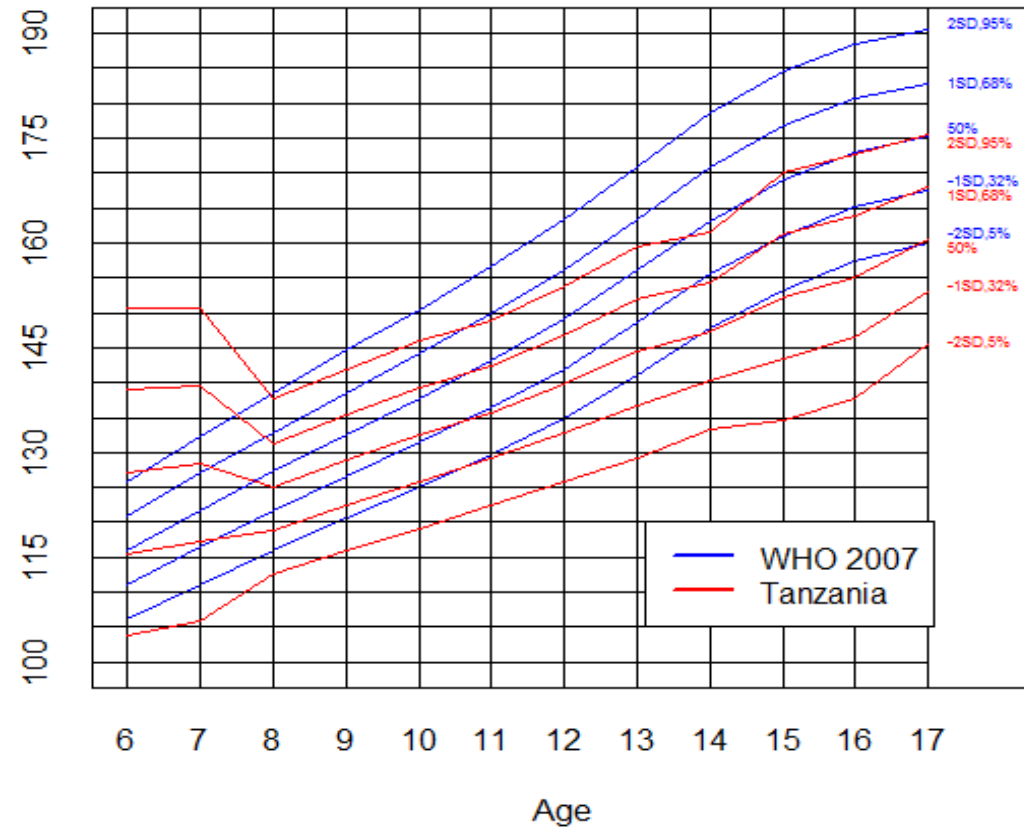


Tanzanian vs. WHO 2007 Growth Charts: Height-for-age

Females

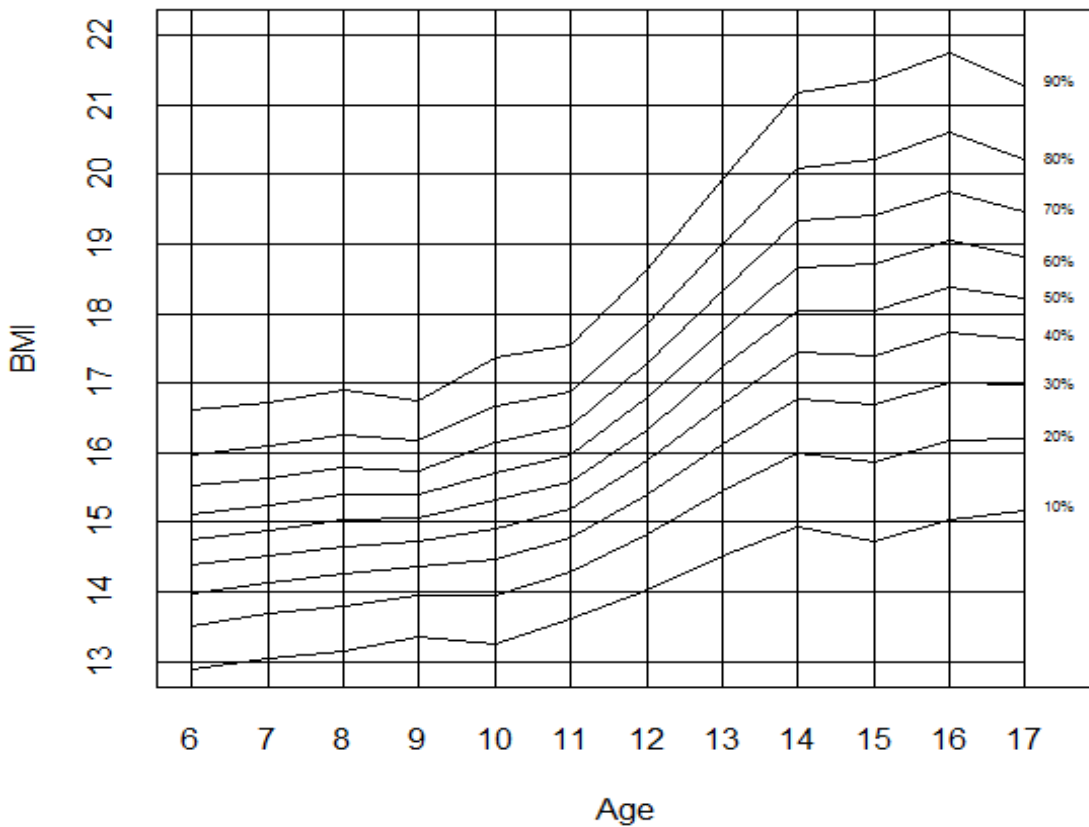


Males

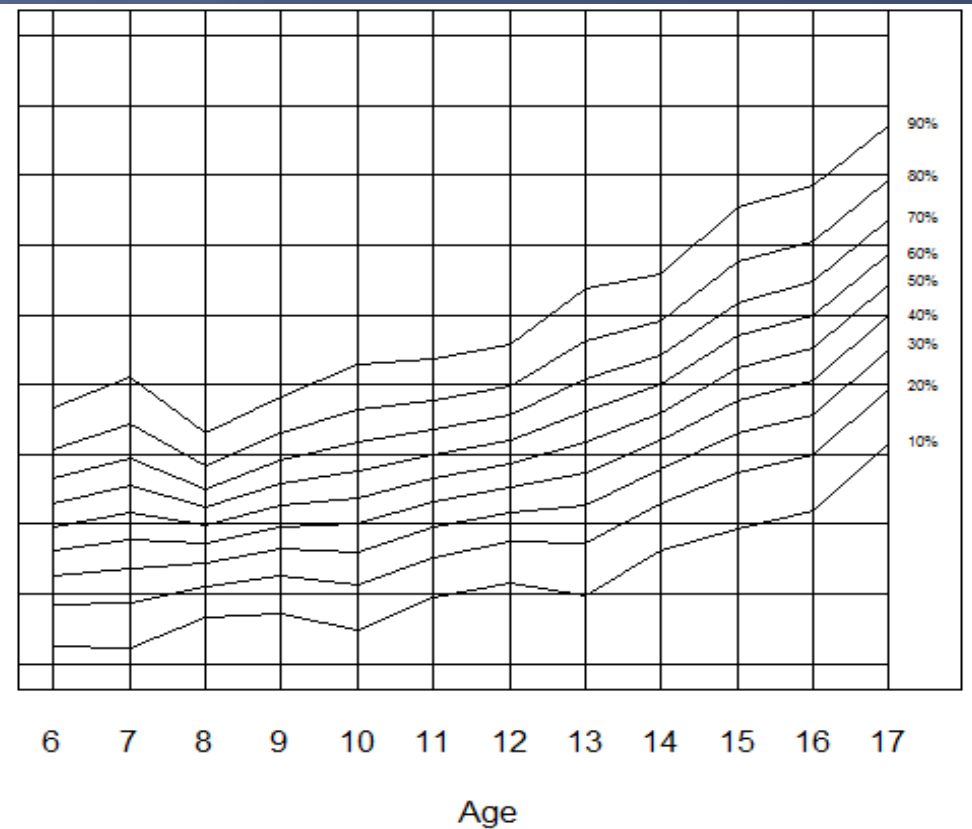


Tanzanian Growth Charts: BMI-for-age

Females

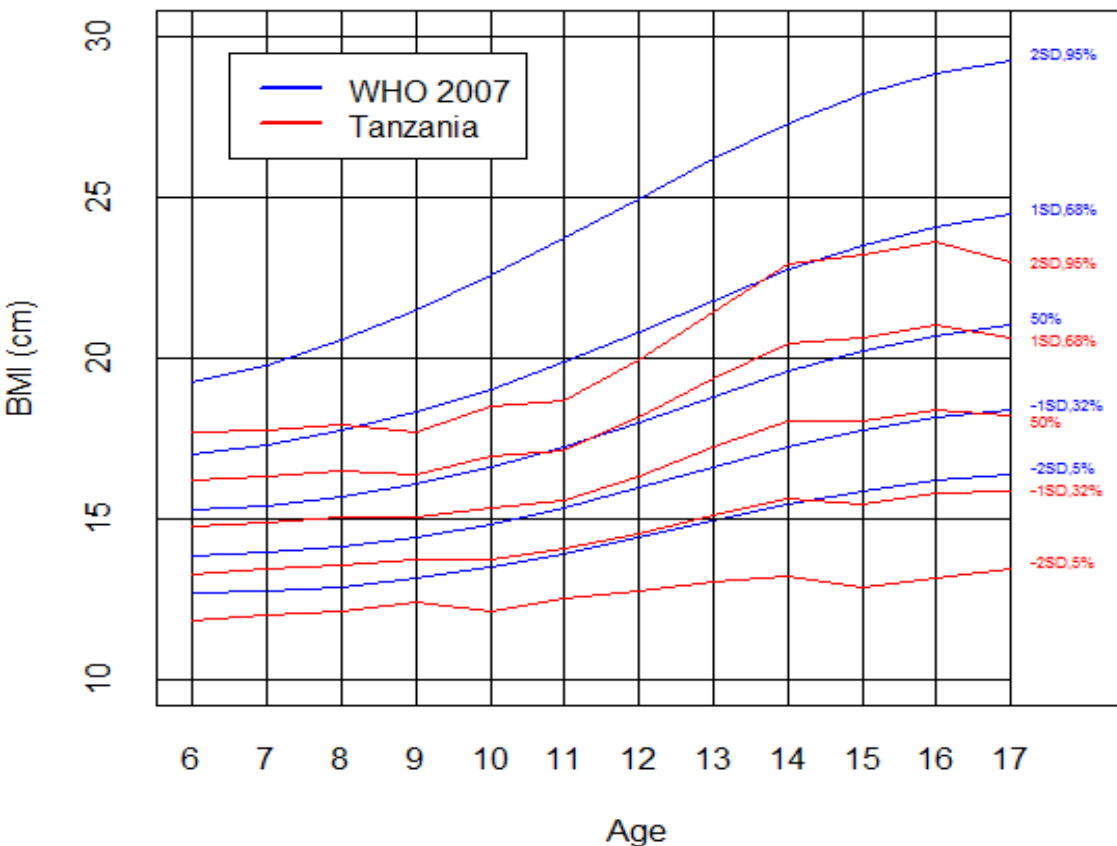


Males



Tanzanian vs. WHO 2007 Growth Charts: BMI-for-age

Females



Males

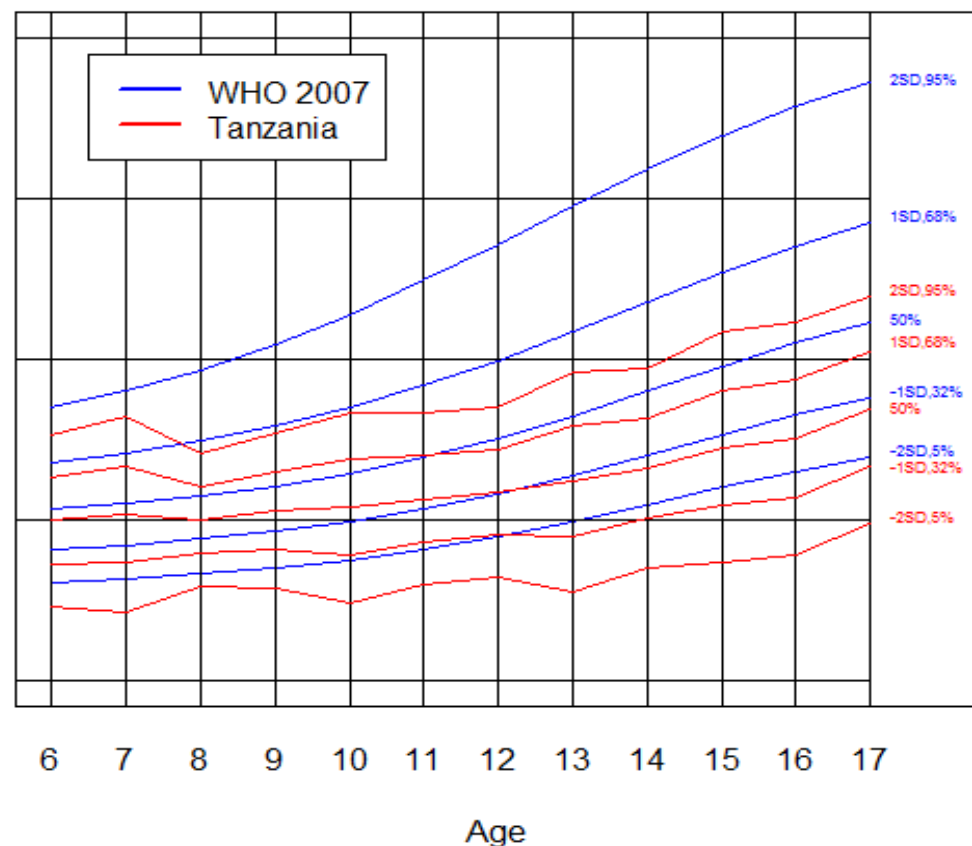
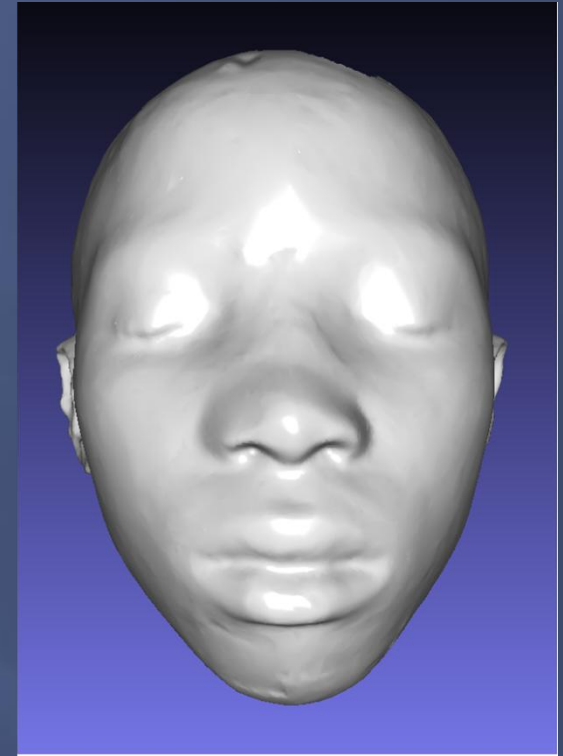
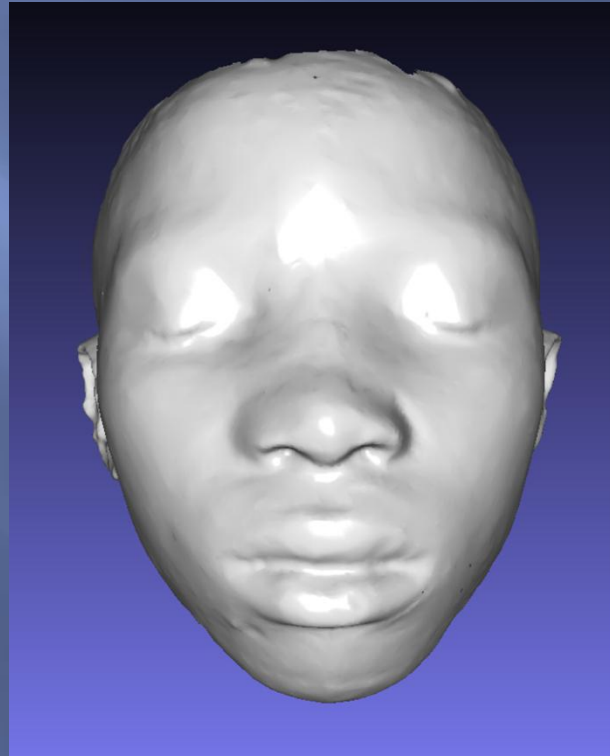


Image Processing Artifact



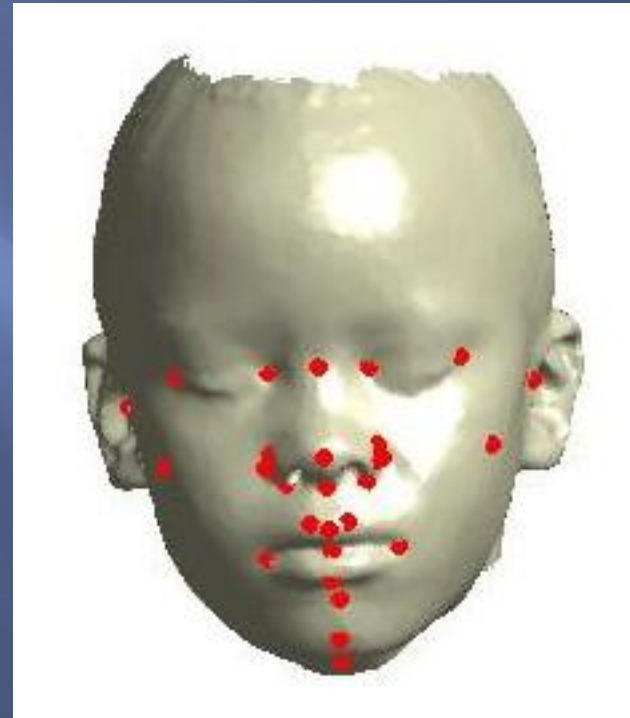
Current status: 2791 images reprocessed

Automated Landmarking

Washington Mio; Florida State University



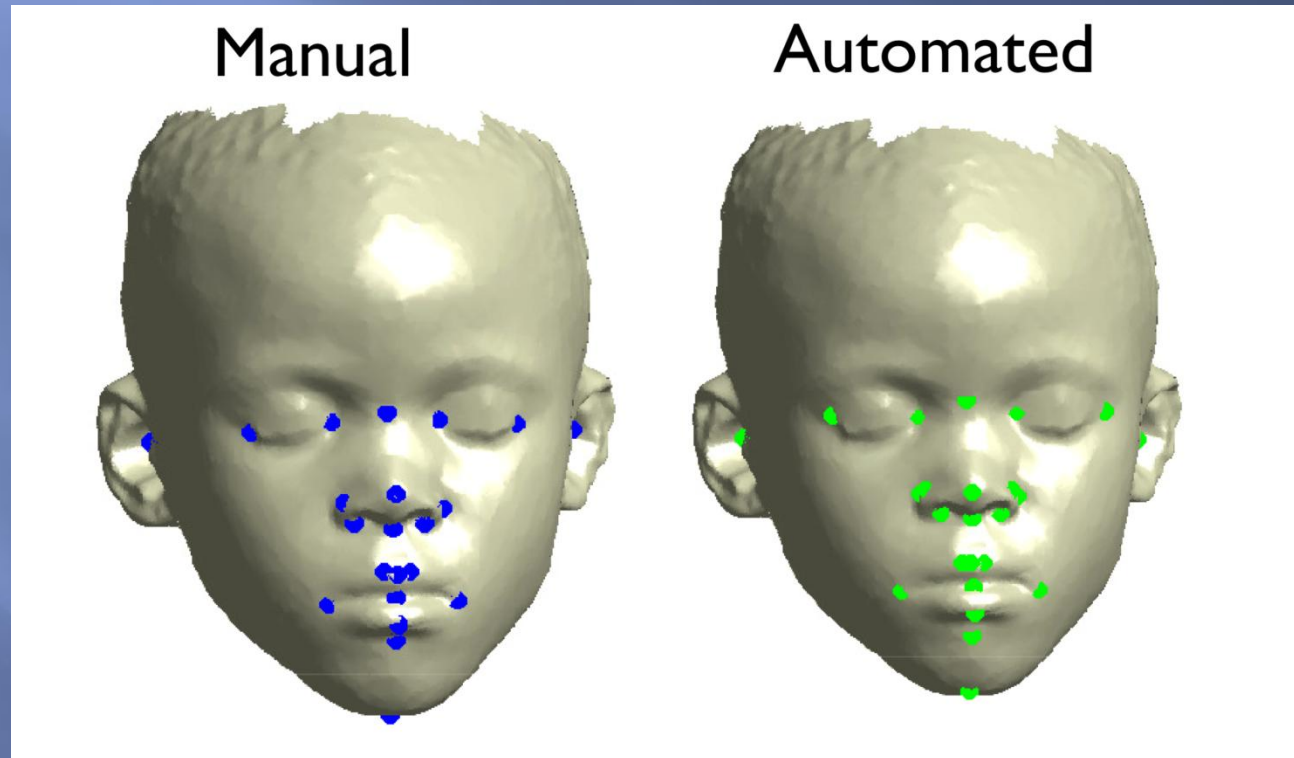
Meanshape
(n=30)



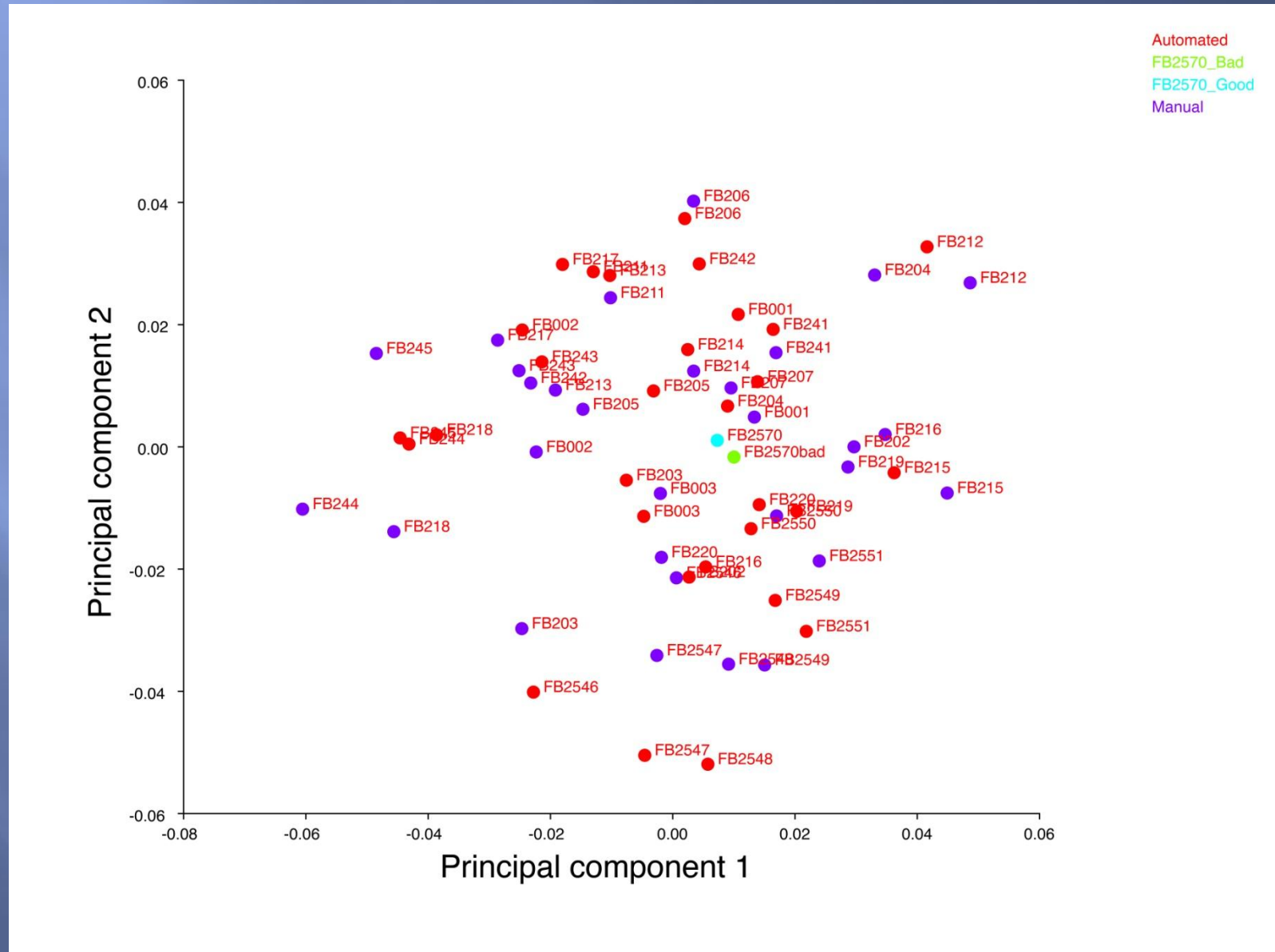
Meanshape
landmarked

Automated Landmarking

Manual and Automated Landmarking on Same Individual



Error Analysis of Automated vs. Manual Landmarking



Analytic Plans

1. Complete image processing GWAS
2. Manually review all images
3. Landmark
4. Derive principal components
5. Genetic analysis using unrelated individuals
6. Genetic analyses using all (including related) subjects
7. Complete DNA QC for replication study
8. Genotyping for replication study at CIDR
9. Do image analysis for replication study
10. Landmark for replication study
11. Derive PCs for replication study
12. Genetic analysis replication study

