Week 7

Ancestry & 23andMe

Notes:
- 1-on-1 meetings with each person over next two weeks
- Thurs Domain Hours this week only: 10-10:50 (not 10:30-11:30)

Assignment 2: Making progress

- Used test data only?
- Separation of populations?
- How to map back to which population a sample was from?
- Discuss: "map"-ness of output

The rest of the quarter...

Week 8: Impacts & Ethics (Final Readings)

Week 9: Related Questions in the Domain

Week 10: Proposals (Elevator Pitch)
- Groups: 2-4
- Start brainstorming projects

Discussion Questions

- What is 23andme?
- What data/techniques/approaches were used in this paper? How/why was each used?
- What is admixture? Why is it important when working with genetic data?
- What is the relationship between self-reported ancestry and genetic ancestry?
- What are two (2) things you learned from this reading?

What y'all learned

- Genetic ancestry is not the same as self-reported ancestry
- Genome-wide ancestry estimates across America
- can trace back admixture and ancestry to historical events
- lower estimates of African Ancestry in self-reported than genetic ancestry
- self-reporting is influenced by social biases due to the importance of "race" in today's society
- Importance of asking the 'right' question (survey terminology)
- Effects of LD on confounding admixture mapping & PCA
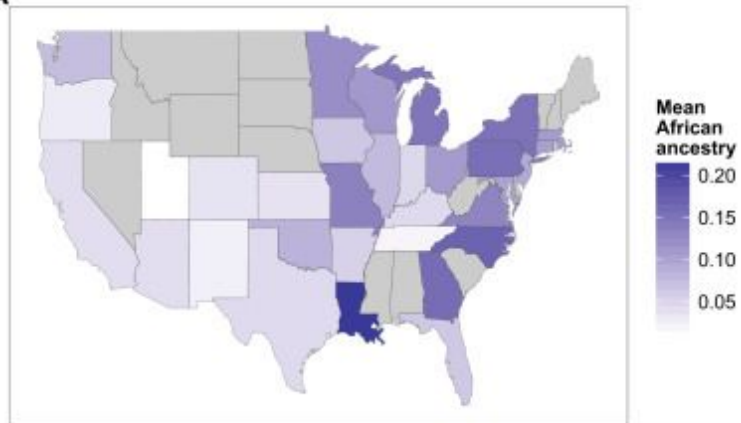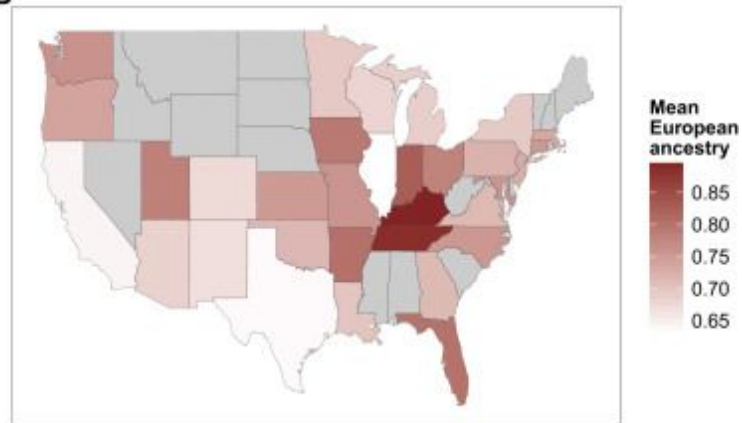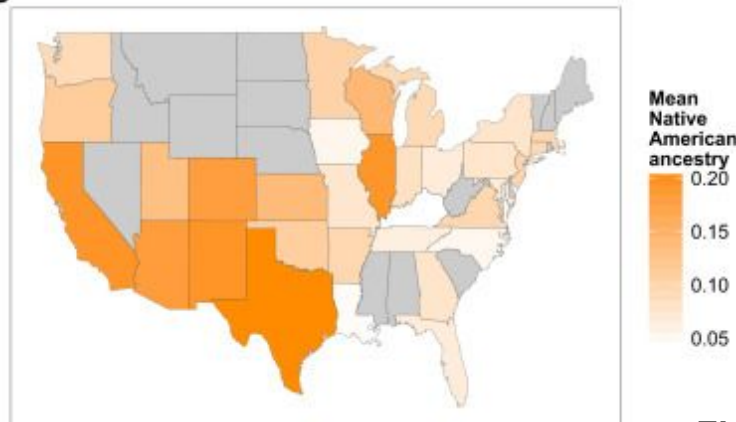- Bias in the underlying data (NY & CA)
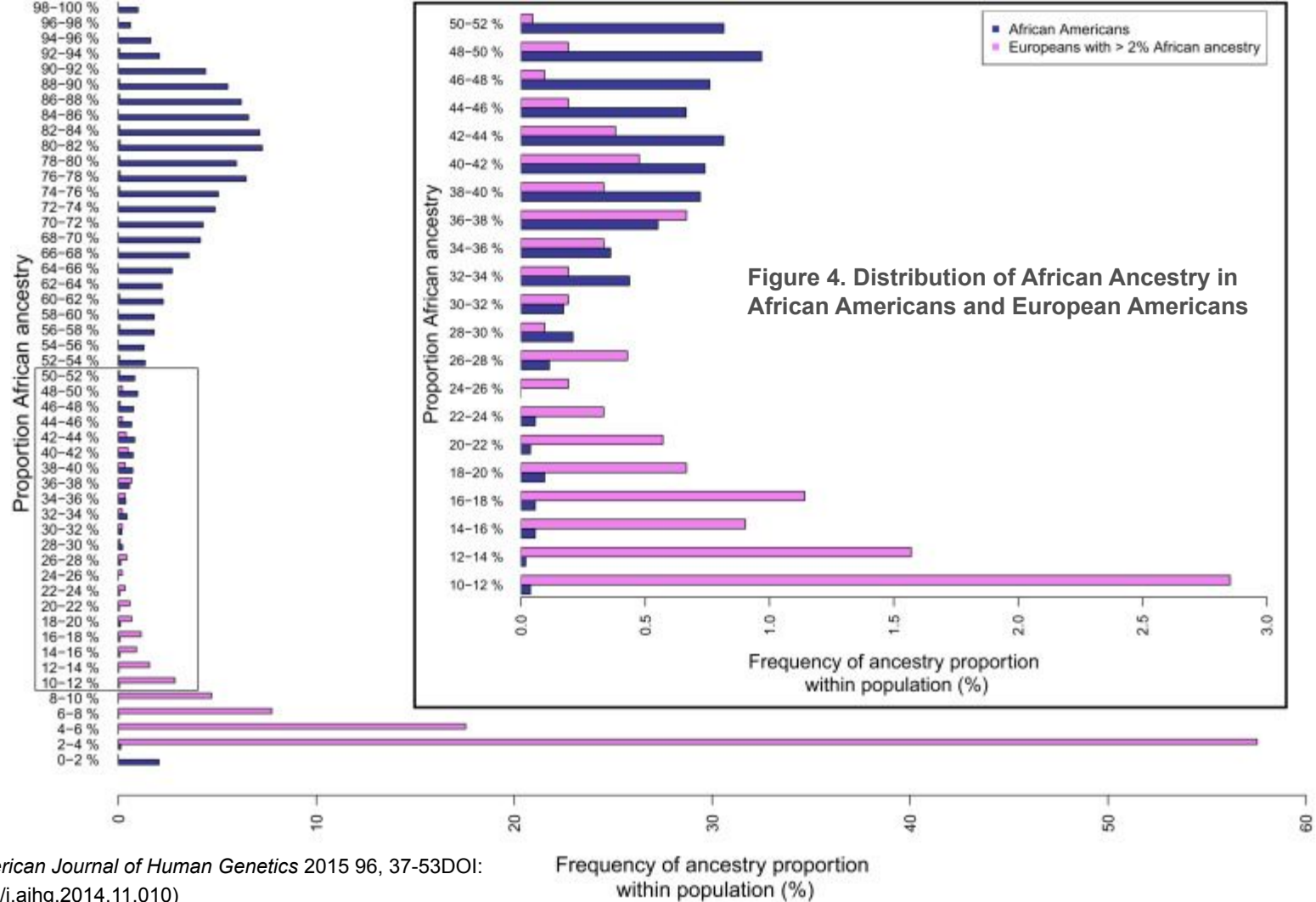- Used 1000G data for validation

Figure 2. The Distribution of Ancestry of Self-Reported Latinos across the US

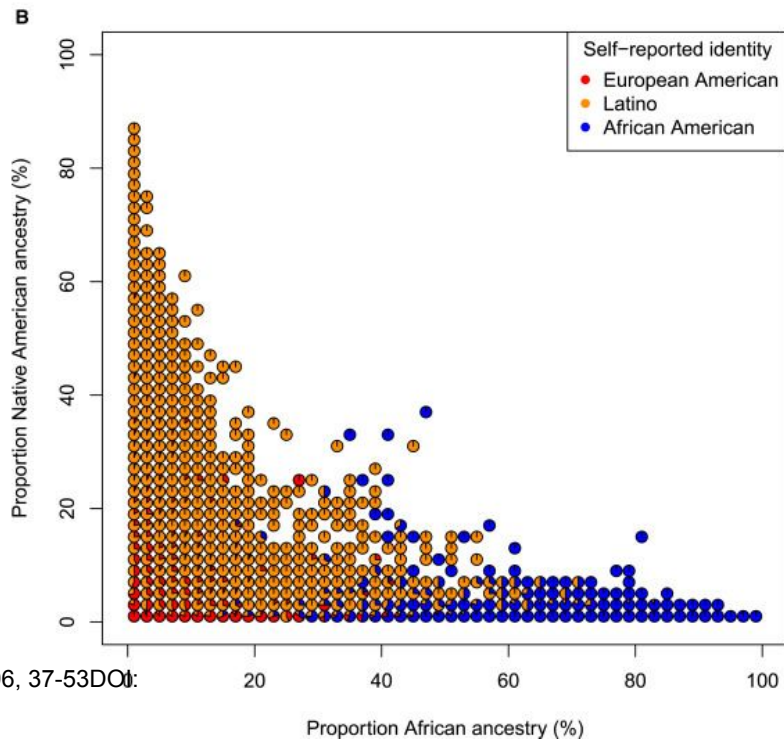Figure 4. Distribution of African Ancestry in African Americans and European Americans
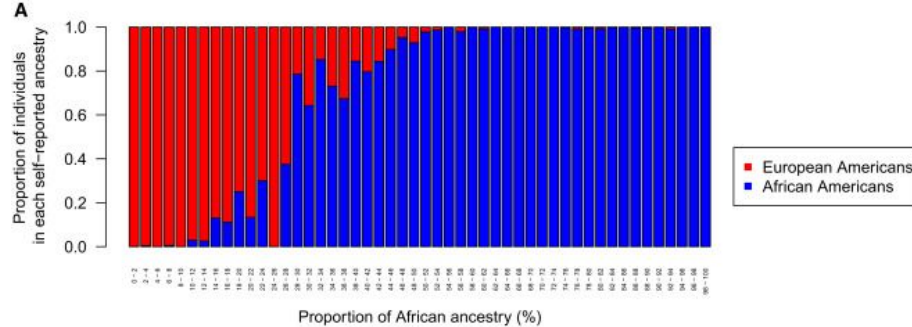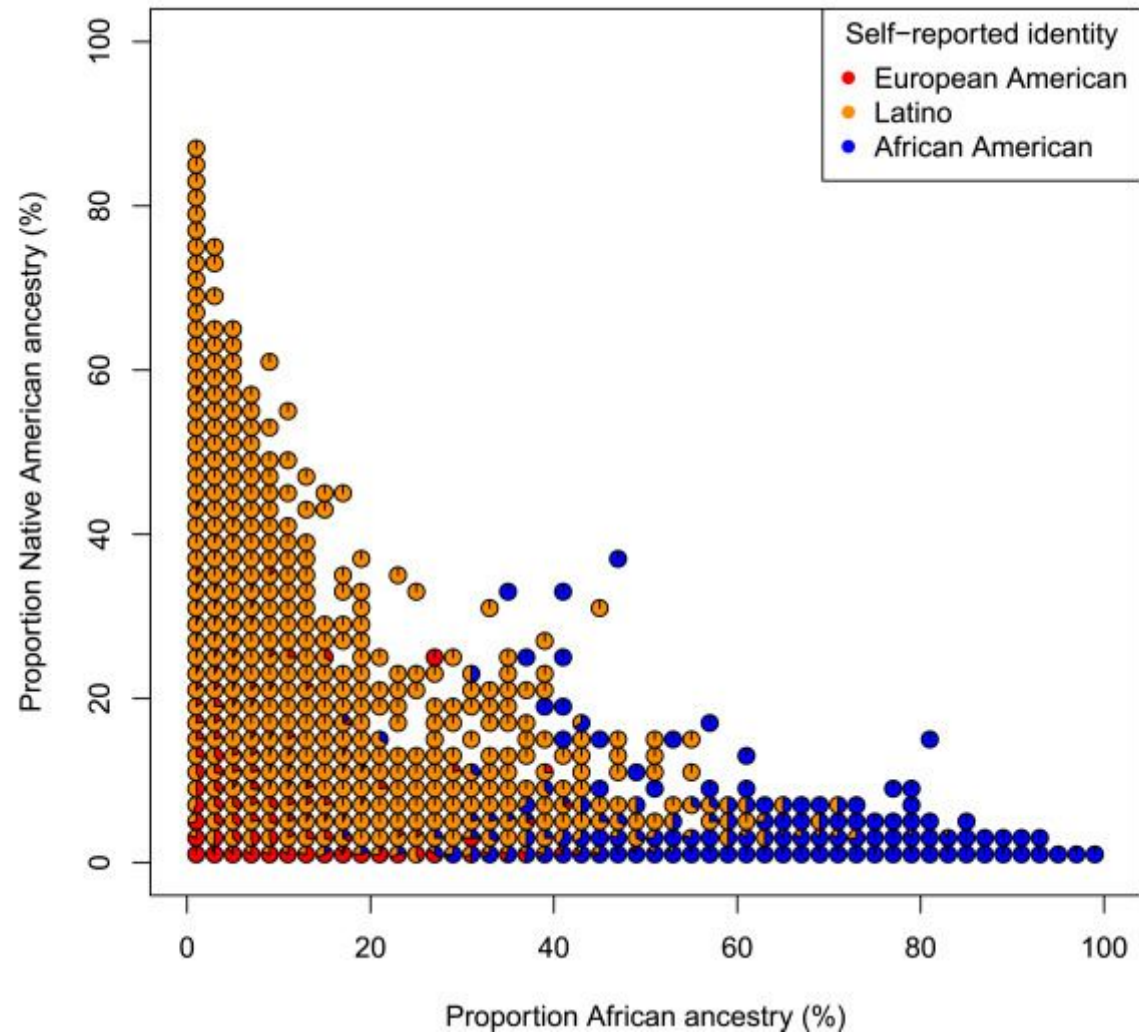
Figure 5. Proportions of Individual Self-Identities by Genome-wide Ancestry Proportions

B

What is 23andme?

What were pitfalls of/negative aspects to the paper?