Recent genetic ancestors, and how we found them

**Fact:** You have 2^2/3 = 3.67 genetic grandparents, but only ~ 33 x 22 genetic grandparents. Therefore, most genealogical common ancestors are not genetic common. G. Wabl, P. Primig. A genomic segment shared IBD by two relatives is a stretch of chromosome inherited from a single common ancestor without intervening recombination.

**Methods:** We used fastIBD in BEAGLE (Browning and Browning, 2011) to identify long (>5.45) blocks shared IBD by pairs among 2,247 Europeans (POPRES Nelson et al., 2008), and used this to infer aspects of recent history and demography. We estimated our power to detect segments of IBD and false positive rate empirically; both were quite good above 4cM, and there was still ample signal above 1cM.

We found ~1.9 million IBD blocks (831 per individual, 0.7 per pair). The IBD rate within populations varied from 14 blocks per pair (Albania) to 0.6 (Italy).

Geography of recent relatedness

Rates of IBD (mean numbers of IBD blocks shared per pair) show both continuous decay with geographic distance (due to local migration), as well as large- and small-scale regional variation (due to historical events).

When did they live?

Each IBD block comes from one genetic common ancestor. We infer \( \mu(t) \), the distribution of ages of these most recent common genetic ancestors (the coalescent time within the pedigree, in meioses) from the distribution of IBD block lengths. \( N(x) \), the number of IBD blocks of length at least \( x \) is Poisson with mean

\[
E[N(x)] = \sum_{t+x} \mu(t)(1 + (G - x)e^{-G}).
\]

We infer \( \mu \), but with a large degree of theoretically unavoidable uncertainty. Keywords: ill-conditioned inversion, ridge regression.

Substructure within populations

Populations (mostly, countries) group together based on rates of IBD, but most have significant substructure. Here are two of the most striking examples.

Conclusions

**Historical:**
1. Any pair of modern Europeans share (many!) common ancestors from the last 1,000 years.
2. Patterns of shared ancestry were shaped by both local migration (marrying in the next village) and by large-scale demographic shifts (history).
3. The Hunnic and/or Slavic expansions of ~400–800CE had a large genealogical impact. Tentatively, so did the Norse expansion (Vikings).
4. The Italian and Iberian peninsulas (and to a lesser extent, France) have stable substructure lasting over the last 2,000 years.
5. Linguistic isolates can have much higher rates of recent shared ancestry (Albanian, tentatively Latvian), but not necessarily (Hungarian).

**Methodological:**
1. Long IBD can be used to infer rates of recent shared ancestry.
2. The distribution of ages of these ancestors is obtainable, but with unavoidable uncertainty.
3. There is ample historical and demographic signal remaining in these data.

References
