Archaic Genes in Modern Humans

Alan R. Rogers

April 6, 2023

Outline

- History of thought about archaic admixture
- Inferring admixture from derived allele patterns and aDNA.
- Inferring admixture from LD in modern DNA.
- Multiple Denisovan populations

Hypotheses of modern human origins

Multiregional Modern humans evolved in a worldwide population united by gene flow (Wolpoff 1989).

Replacement "A single origin, outward migration of separate stirps, like the sons of Noah, and an empty world to occupy, with no significant threat of adulteration by other gene pools or even evaporating gene puddles" (Howells 1976).

Admixture Expansion from a single origin, involving "encounters between populations of modern man and of other forms, with consequent gene flow" (Howells 1976; also: Brauer 1984, Smith et al 1989, Trunkhaus 2005).

Geneticists on admixture

Consensus among geneticists during 1990s: little or no admixture between moderns and archaics.

Eswaran et al 2002, 2005: argued for admixture

Hawks et al 2006: Small amounts of admixture can have a big effect, because advantageous alleles are most likely to introgress.

Green et al 2010: sequenced Neanderthal genome; showed that 1.3–2.7% of European genes came from Neanderthals. [current estimate 1.5–2.1%]

Outline

- ○ History of thought about archaic admixture
- ▶ Inferring admixture from derived allele patterns and aDNA.
- ▶ Inferring admixture from LD in modern DNA.
- ▶ Multiple Denisovan populations

Nucleotide site patterns

Ancestral allele (0) is shared with chimp.

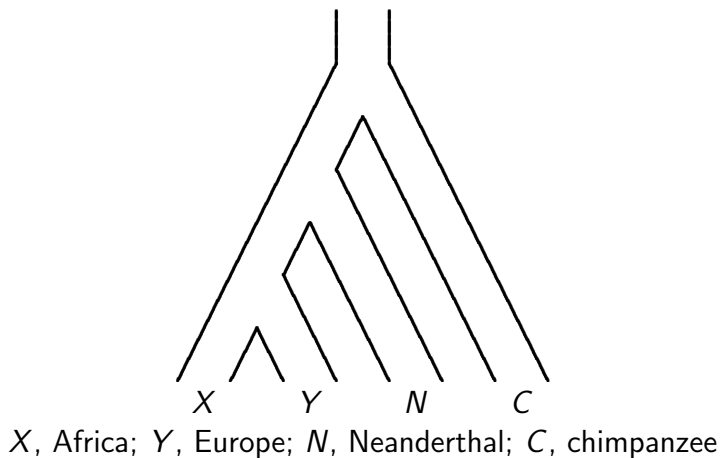Mutant allele (1) shared by two human populations.

	Nucleotide Site Pattern		
	xy	*yn*	*xn*
Afr (X)	1	0	1
Eur (Y)	1	1	0
Nea (N)	0	1	1
Chi	0	0	0
#	303,340	103,612	95,347

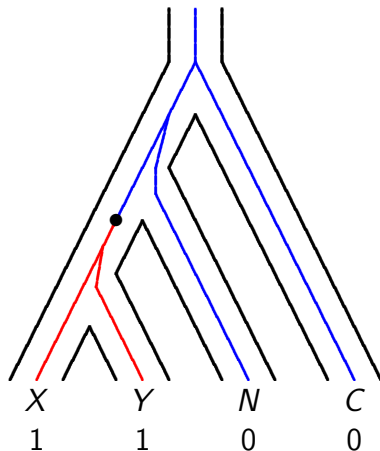Pattern *xy*: most common; reflects history of population splits.

Patterns *yn* & *xn*: how do they arise?

Why does *yn* exceed *xn*?

Population tree

X, Africa; Y, Europe; N, Neanderthal; C, chimpanzee

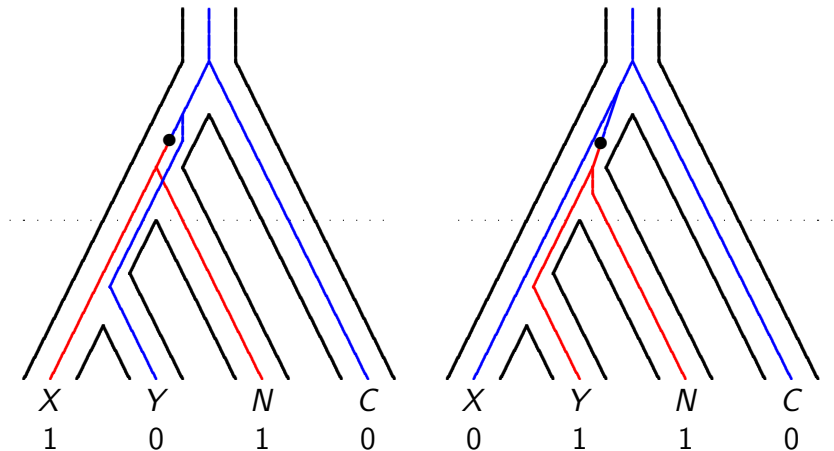# Embedded gene genealogy with mutation

- ▶ Genealogy of 4 genes shown in color.
- ▶ Bullet (•) marks mutation from allele 0 to allele 1.
- ▶ Descendants of mutant have allele 1; others have 0.
- ▶ Gene genealogy matches phylogeny
- ▶ Mutant allele shared by closest relatives, X and Y.

Incomplete lineage sorting

Pattern *xn*

Pattern *yn*

X	Y	N	C		X	Y	N	C
1	0	1	0		0	1	1	0

These two should be equally common

Nucleotide site patterns again

	Nucleotide Site Pattern		
	xy	*yn*	*xn*
African (X)	1	0	1
European (Y)	1	1	0
Neanderthal (N)	0	1	1
Chimpanzee	0	0	0
# sites	303,340	103,612	95,347

Common pattern (*xy*) reflects history of population splits.
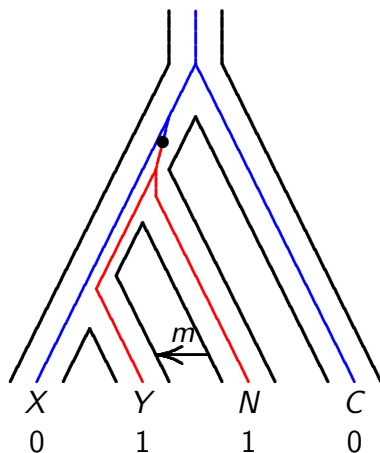
Absent admixture, the other two should be equally common

Why does *yn* exceed *xn*?

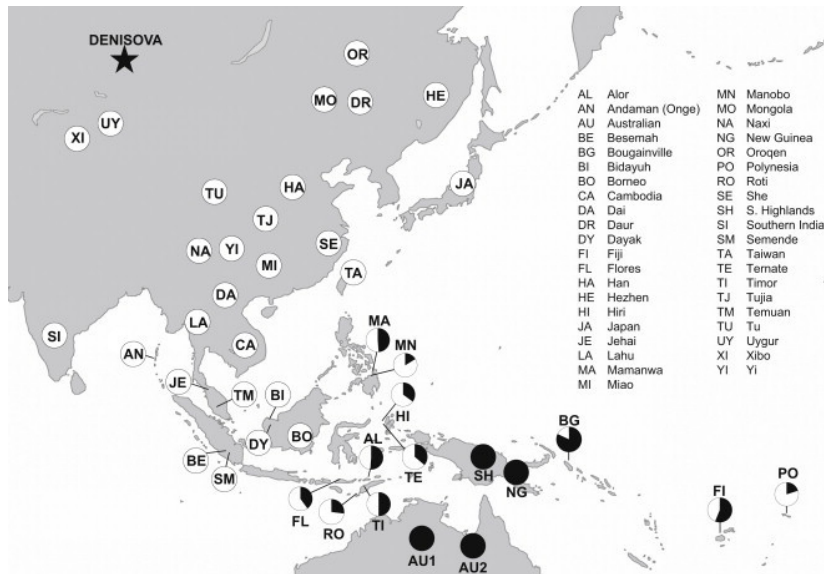# Neanderthal admixture inflates *yn* site pattern

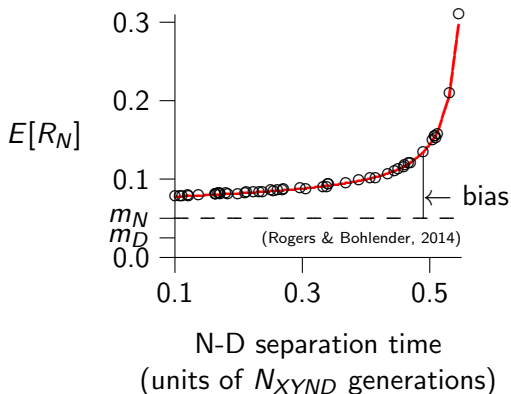Key: *X*, Africa; *Y*, Europe; *N*, Neanderthal; *C*, chimpanzee.

Estimate from Neandertal DNA

- DNA of modern Eurasians is 1.5–2.1% Neandertal
  (Prüfer et al 2014).

- Same is true for modern people of east Asia and Papua New Guinea, but not Africa.   Green et al (2010)

- Admixture must have occurred *after* moderns left Africa but *before* they expanded throughout the world.

- Eurasian introgressed segments most similar to Neanderthal from Caucasus.   (Prüfer et al 2014)

Denisovan DNA most common in Australia, NG, and Oceania

Some estimators have large biases

$E[R_N]$

N-D separation time
(units of N_{XYND} generations)

(Rogers & Bohlender, 2014)

Bias is pronounced in populations that received gene flow from Denisovans as well as Neanderthals.

Worse in East Asia than in Europe.

Worse still in Melanesia.

All of the estimates reviewed above are suspect.

Legofit

Our own estimator, Legofit, is an estimator that avoids this bias. I'll discuss it in a separate lecture.

Outline

- ○ History of thought about archaic admixture
- ○ Inferring admixture from derived allele patterns and aDNA.
- ▶ Inferring admixture from LD in modern DNA.
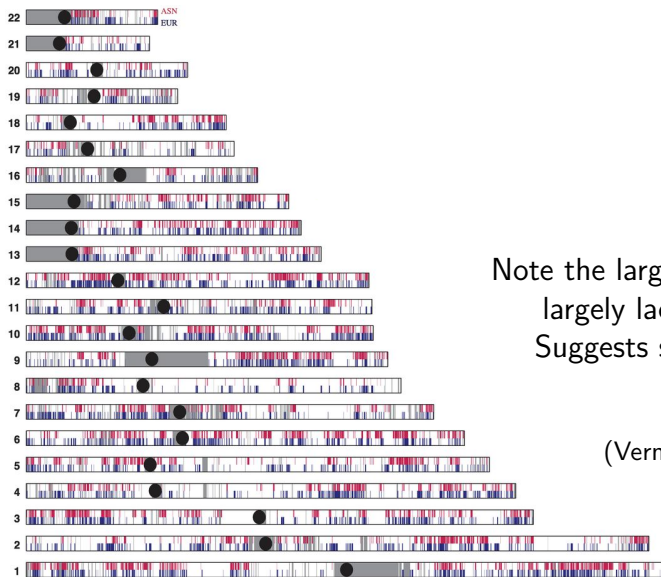- ▶ Multiple Denisovan populations

Using LD to discover admixed chromosome segments

1. Recent introgression $\rightarrow$ modern genomes should contain long segments archaic chromosome.
2. These segments should differ a lot, because of the long separation time.

Method of Vernot and Akey (2014)

1. S^* statistic (Plagnol & Wall 2006) uses modern LD to find candidate introgressed segments.
2. Accept candidate if matches Neanderthal sequence better than chance.
3. Studied 379 Europeans and 286 East Asians.
4. Found 15 Gb introgressed sequence spanning 20% of Neanderthal genome.
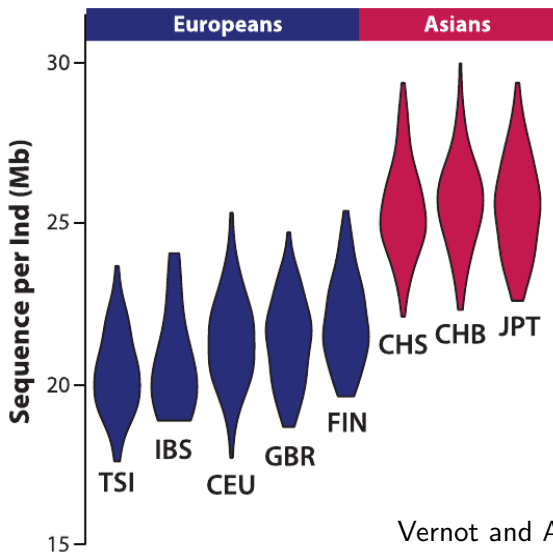
Neandertal segments in modern genomes
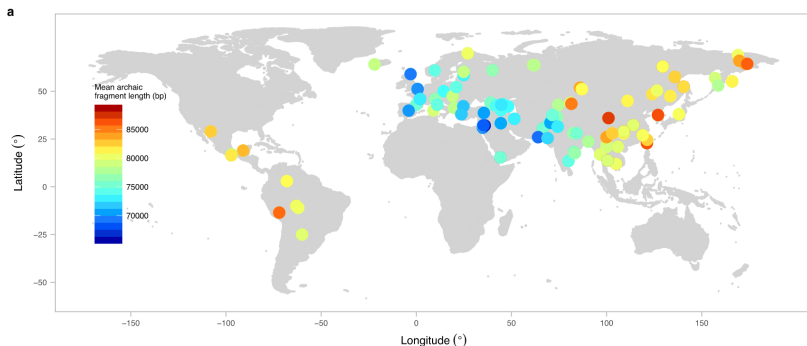

red: Asians

blue Europeans

Note the large "deserts" that largely lack archaic DNA. Suggests selection against archaic alleles.

(Vernot and Akey 2014)

Asians seem to have more Neandertal than Europeans—why?

Vernot and Akey (2014)

Asian excess may be misleading

Neanderthal fragments are longer in East Asia. Long fragments are easier to recognize. May explain East Asian excess. (Macia et al (2021))

Why are Neanderthal fragments short in Europe?

Perhaps admixture was older in Europe? Macia et al (2021) say no: the genomic regions covered by Neanderthal fragments are the same Europe and Asia. Implies a single episode of admixture.
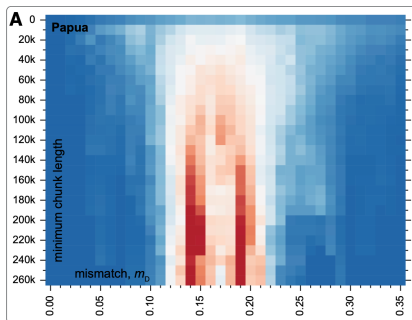
More derived alleles in Europe than Asia $\Rightarrow$ Shorter generation time in Europe $\Rightarrow$ More recombination in Europe.

So apparent excess of Neanderthal DNA in Asia may be an artifact caused by longer generation time there.

Outline

- ○ History of thought about archaic admixture
- ○ Inferring admixture from derived allele patterns and aDNA.
- ○ Inferring admixture from LD in modern DNA.
- ▶ Multiple Denisovan populations
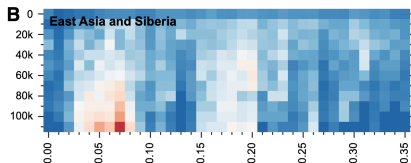
Papuans got DNA from 2 Denisovan pops

Vertical axis: length of introgressed segment

Horizontal: diff btw segment and Denisovan genome as fraction of Denisovan-modern diff.

Two Denisovan populations: one 0.15 and one at 0.2.

Jacobs et al (2019)

E Asians got DNA from 1 Denisovan pop

Vertical axis: length of introgressed segment

Horizontal: diff btw segment and Denisovan genome as fraction of Denisovan-modern diff.

One Denisovan population at 0.07.

Summary

- Neanderthal: all non-African populations have about same level (1.5–2.1%) of admixture. Suggests that admixture happened in one region.
- Denisovan: highest levels found in populations far from Africa. Suggests that admixture accumulated gradually.
- We can recognize chromosomal segments derived from Neanderthals or Denisovans, because they differ from modern haplotypes but resemble archaic haplotypes.
- Large regions of Eurasian genomes have little or no archaic introgression.
- East Asians have more archaic DNA than Europeans do.
- Papuans got DNA from 2 Denisovan pops.
- E Asians got DNA from a 3rd.