

# On: Paleolithic to Bronze Age Siberians Reveal Connections with First Americans and across Eurasia, by Yu et al

Alan R. Rogers

8 June 2020

# Overview

- ▶ Nineteen Upper Paleolithic to Early Bronze Age genomes from the Lake Baikal region of Siberia. One Upper Paleolithic genome shares the admixed ancestry that gave rise to all non-Arctic Native Americans.
- ▶ Populations of the Bronze Age in this region formed from those of the Neolithic by prolonged admixture.
- ▶ Long-range human and *Y. pestis* mobility across Eurasia during the Early Bronze Age.

# Samples (from 10 archaeological sites)

- ▶ 1 Upper Paleolithic individual (14,050–13,770 BP)
- ▶ 4 Early Neolithic individuals (7,320–6,500 BP)
- ▶ 14 Late Neolithic to Early Bronze Age (LNBA) individuals (4,830–3,570 BP)

# Sequencing

## **Samples with poor DNA preservation**

- ▶ Capture array targeting a set of 1.24 million variable sites achieves coverage of 0.04–2.07X
- ▶ Pseudo-haploid genotypes called at 34–886k SNPs.

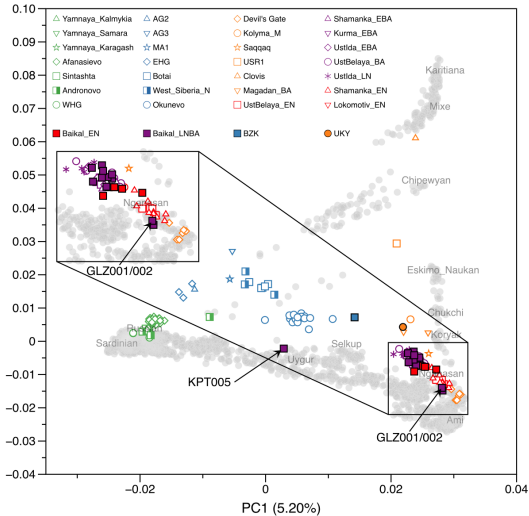
## **8 samples with better preservation**

- ▶ Deep shotgun sequencing, achieving 0.1–1.9X coverage.
- ▶ Imputed diploid genotypes at 386–518k SNPs from the Human Origins SNP array.

## **Possible sources of bias**

- ▶ Ascertainment bias, resulting from choice of SNPs to capture or from genotyping only SNPs in Human Origins Array.
- ▶ Is it easier (or harder) to call heterozygous genotypes?

# Principal Components Map

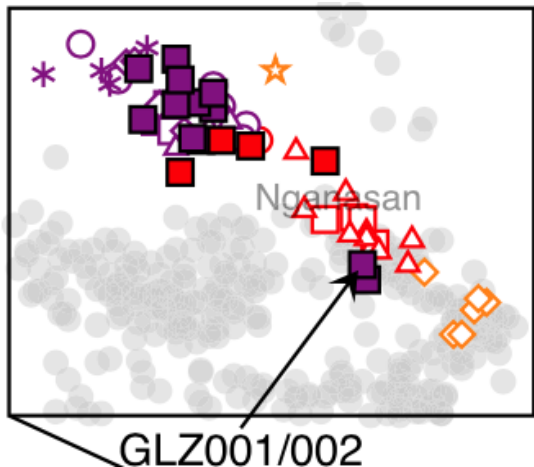


Ancient Siberians similar  
to modern ones.

Along the line from  
“Northeast Asian”  
(lower right) and  
“Ancient North  
Eurasians” (upper left).

New Paleolithic  
genome: ●

# Principal Components Inset

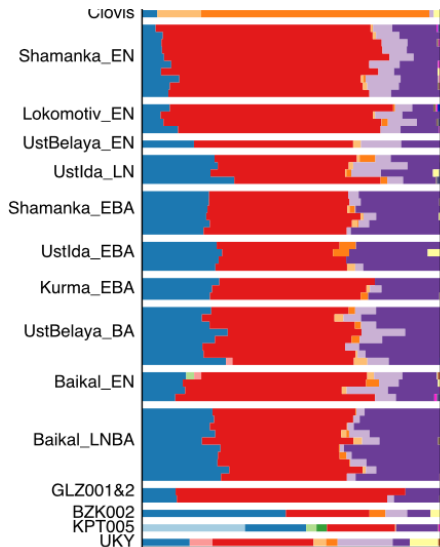


Early Neolithic closer to  
Northeast Asian.

Late Neolithic/Bronze  
Age closer to Ancient  
North Eurasian (off map  
to NW).

Suggests gene flow from  
ANE into Baikal region  
after Neolithic.

# Admixture Analysis



Mixture Ancient North Eurasian, Northeast Asian, and Nganasan.

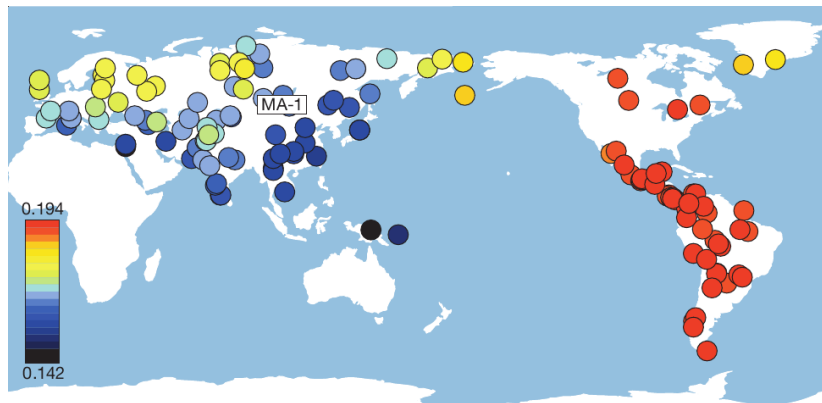
ANE ↗; NEA ↘.

KPT005 is an immigrant.

**Immigrants:** GLZ001 & 2 are early Bronze Age but have a lot of NEA. GLZ001 has non-local strontium isotope ratios.

Implies long-range mobility during this period.

# Affinities of Mal'ta (24 kya) with world populations

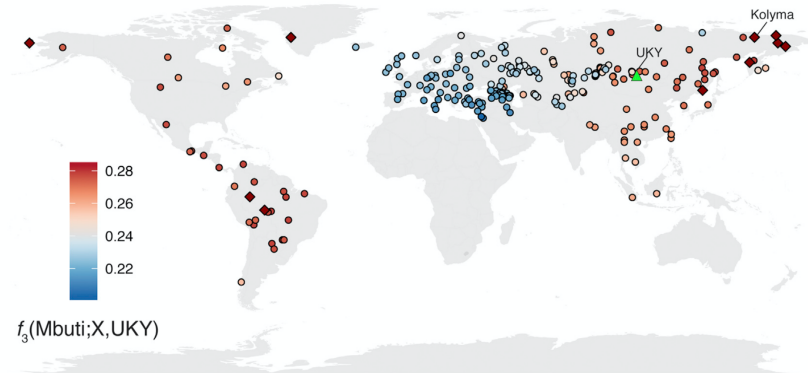


(Raghavan et al. 2013)

Similar to Amerindians and Northern European. “Ancient North Eurasians” (ANE).



# Affinities of UKY (14 kya) with world populations



(Yu et al. 2020)

Similar to Amerindians and E Asians, not to Europeans.

# Plague DNA

Found *Yersinia pestis* DNA in two Early Bronze Age individuals (GLZ001 & GLZ002, ~4,400 BP).

DNA suggests this strain of plague was not good at transmission via fleas.

# Gradual gene flow over millenia

Other studies have used genetic data to document fairly sudden and nearly complete replacement of one population by another.

This article tells a different story: the Ancient Northeast Asian DNA of Malt'a 1 (24 kya) survived in this region throughout the Upper Paleolithic, as revealed in the UKY specimen (14 kya), and even into the Early Neolithic.

Furthermore, evolutionary changes between the Early Neolithic and Bronze Age were also gradual, suggesting prolonged weak gene flow.

Under what circumstances are such changes gradual rather than sudden? We have no theory of this.

# Problem 1: Ascertainment bias

I worry that these estimates may be affected by the bias that ensues when some loci are selected for capture and others are excluded.

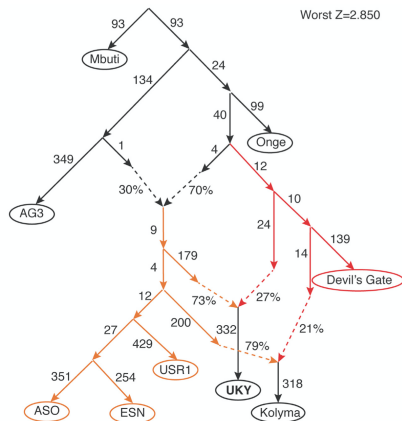
## Problem 2: Goodness of fit

In other fields, people routinely ask how well their models fit the data.

They also use residuals to diagnose problems with fit.

Anthropological geneticists seldom do this in studies of population history.

# Problem 3: Overfitting



Overfitting happens when you fit noise rather than signal and end up with the wrong model.

A real danger with complex models.

Statisticians have methods that protect against overfitting, and geneticists ought use them.

# Discussion questions

1. If the capture array overrepresents SNPs with high heterozygosity, what effect might that have on this analysis?
2. In some regions, population replacements were relatively fast and relatively complete. In others (including Baikal), they were slow and gradual. Under what circumstances should we expect one rather than the other?