Genetics and the Neolithic of Europe
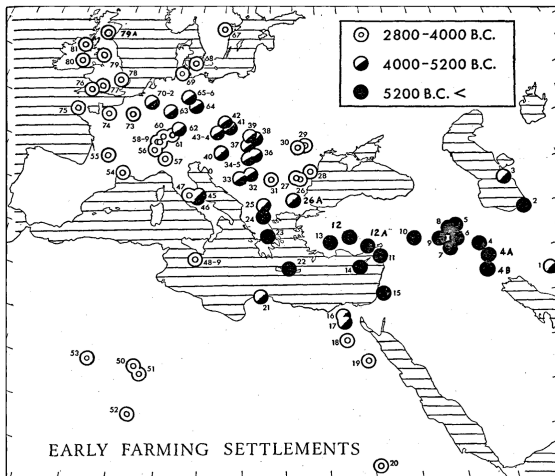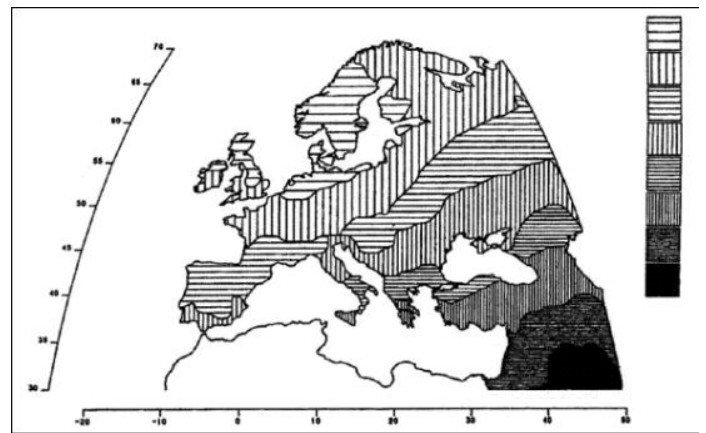
Alan R. Rogers

November 23, 2013

Outline

- ▶ The European Neolithic: a movement of peoples or of technology?
- ▶ Linkage disequilibrium (LD)
- ▶ How LD responds to changes in population size.
- ▶ The history of European population size.

Spread of farming across Europe

(Grahame Clark, 1965, *Proc. Prehist. Soc.*)

Major axis of genetic variation in Europe

95 genes (Cavalli-Sforza, 1994, p. 292)

Movement of people or of technology?

Local hunter-gatherers contributed less than 30% in the original settlements. This finding leads us to reject a predominantly cultural transmission of agriculture.

(Lounès Chikhi et al. 2002)

Both mitochondrial DNA and Y chromosome analyses have indicated a contribution of Neolithic Near Eastern lineages to the gene pool of modern Europeans of around a quarter or less. This suggests that dispersals bringing the Neolithic to Europe may have been demographically minor.

(Martin Richards 2003)

24 ky old burial from Mal'ta, Siberia
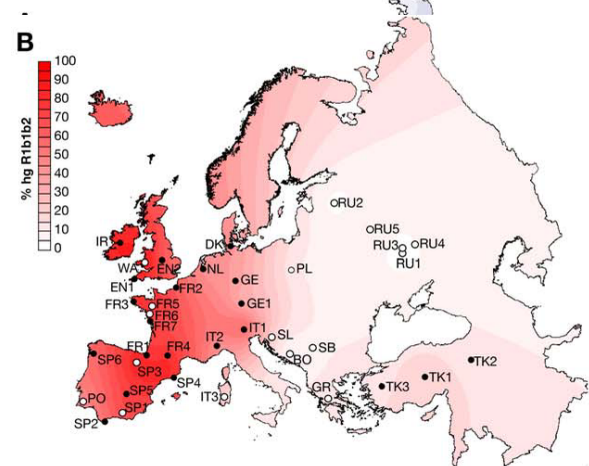

(Raghavan et al. 2013)

24 ky old burial from Mal'ta, Siberia


(Raghavan et al. 2013)

- ► 1/3 of ancestry shared with Native Americans and Europeans
- ► European mitochondrial DNA

Y haplogroup R1b1b2 most common in Ireland: Mesolithic origin?

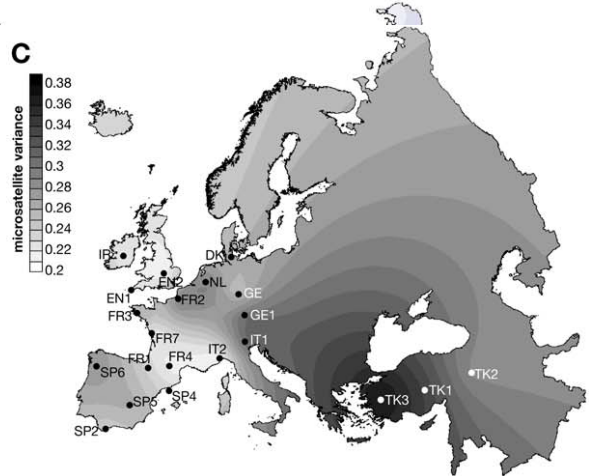
Microsatellite loci

- ► Microsatellite: a short DNA sequence repeated several times, adjacent on chromosome
  Example:
  ATTC ATTC ATTC ATTC (4 copies)
  ATTC ATTC ATTC ATTC ATTC ATTC (6 copies)
- ► High mutation rate.
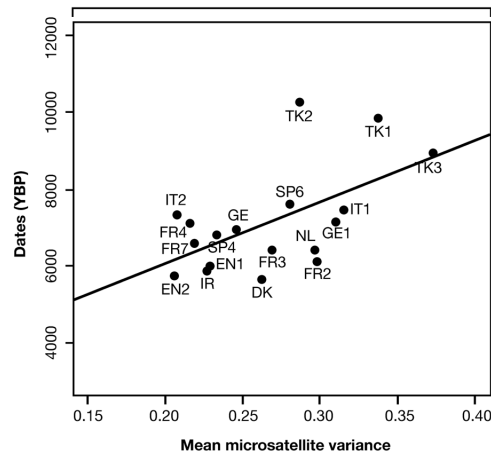- ► Variance among chromosomes increases with time.
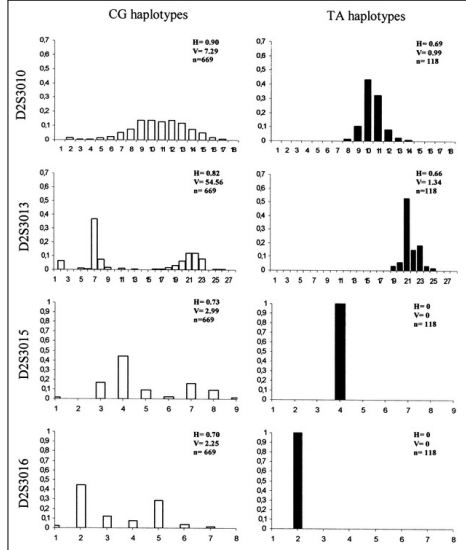
Microsatellite variation measures age

- ► Rows are different microsatellites
- ► CG haplotype is old
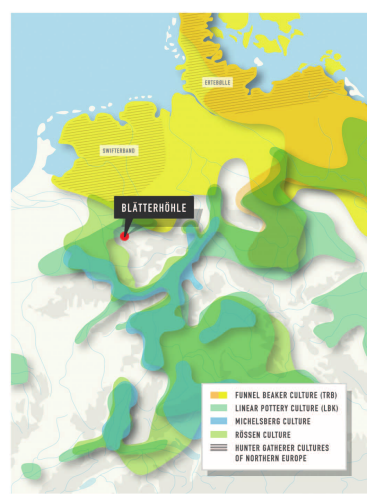- ► TA haplotype is younger: about 10,000 years

(Coelho et al 2005)

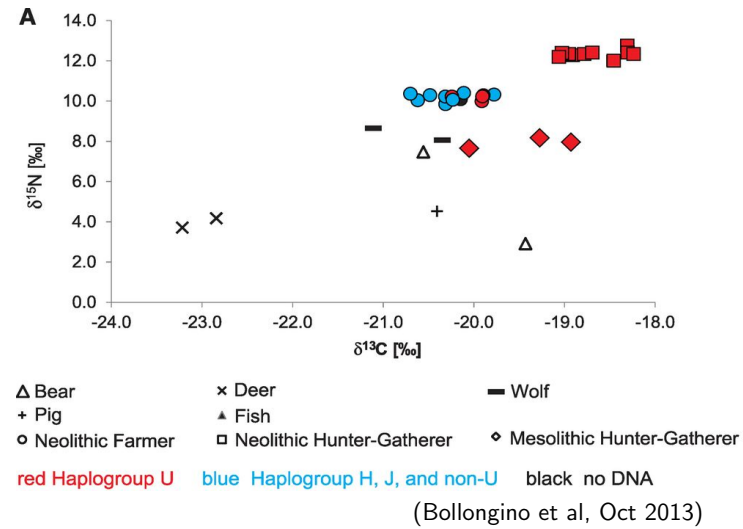## On Y haplogroup R1b1b2, microsatellite variation is greatest in Turkey

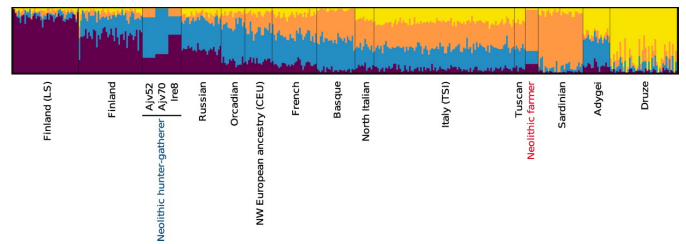## Microsatellite variance vs. earliest Neolithic dates

The Blätterhöhle site in Germany

mtDNA of Neolithic farmers and foragers
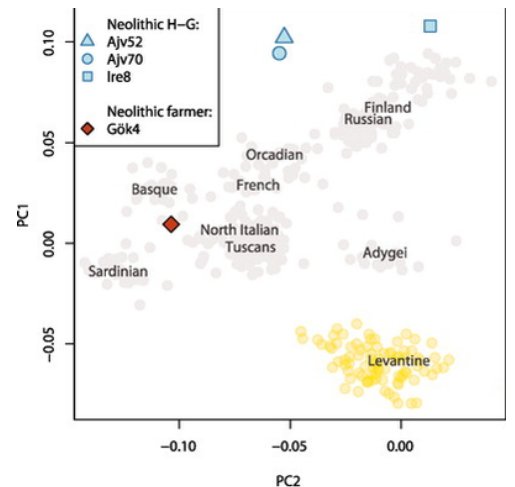

(Bollongino et al, Oct 2013)

Nuclear genes of Neolithic farmers and foragers

(Skoglund et al, 2012)

During Neolithic, farmer DNA like modern Sardinians.
Forager DNA like modern Finns.

Neolithic farmers and foragers had different DNA


(Skoglund et al, 2012)

Outline

- ○ The European Neolithic: a movement of peoples or of technology?
- ▶ Linkage disequilibrium (LD)
- ▶ How LD responds to changes in population size.
- ▶ The history of European population size.

Linkage disequilibrium (LD) is one of those unfortunate terms that does not reveal its meaning. As every instructor of population genetics knows, the term is a barrier not an aid to understanding... Detecting LD does not ensure either linkage or a lack of equilibrium.
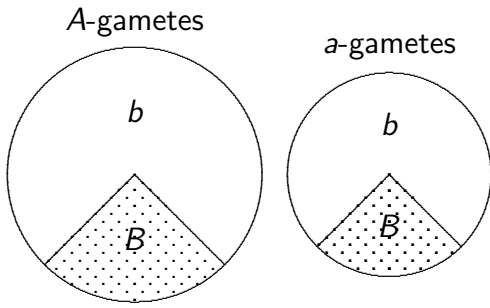
(Montgomery Slatkin, 2008)

Gamete types at two genetic loci

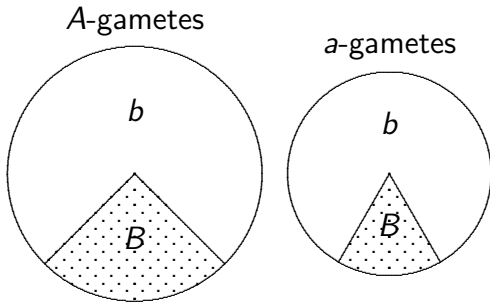| a	b
a	B
A	b
A	B
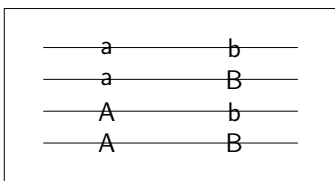
One locus has alleles *A* and *a*.

The other has alleles *B* and *b*.

There are four possible types of gamete.

Gametes that carry *A* are "*A*-gametes."

Linkage Equilibrium $\Longleftrightarrow$ shaded fractions equal

A-gametes

a-gametes

LE: Knowledge about one locus tells nothing about other.

Here, *B* is equally common among *A*-gametes and *a*-gametes.

LD $\Longleftrightarrow$ shaded fractions unequal

A-gametes

a-gametes

LD: Knowledge about one locus helps predict the other.

Here, *B* is more common among *A*-gametes than *a*-gametes.

You can see LD in sequence data
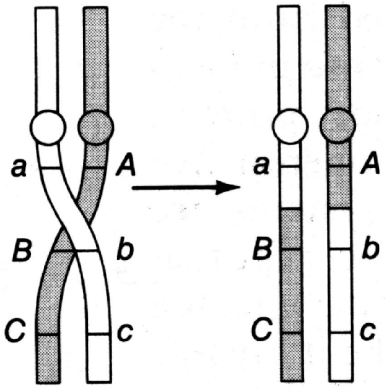
		Nucleotide position												
			1	1	1	1	1	1	1	2	2	2	
		3	8	2	3	3	6	6	7	9	0	2	3
	1	9	4	6	4	4	3	3	5	0	2	6	0
	2	7	7	1	3	4	4	9	3	1	0	3	4
Orang	T	G	C	A	T	G	T	A	A	C	G	C	T
Chimp	T	G	C	A	T	G	T	A	A	T	G	C	T
A	.	.	.	.	A	.	.	.	.	G	A	A	.
B	.	.	.	.	A	.	.	.	.	G	A	.	.
C	.	.	.	.	.	T	.	G	.	.	.	.	C
D	C	.	.	.	.	.	G	G	.	.	.	C
E	C	C	.	.	.	.	G	G	.	.	.	C
F	C	.	.	.	.	C	G	.	.	.	.	C
G	C	.	T	.	.	.	C	G	.	.	.	C
H	C	.	T	G	.	.	C	G	.	.	.	C

(GARRIGAN ET AL 2004)

- Dots: identical to chimp sequence.
- Sites not independent.
- A at site 1343 predicts G at 1951
- This is linkage disequilibrium (LD).

Cross-overs shuffle DNA

- occur during reproduction.
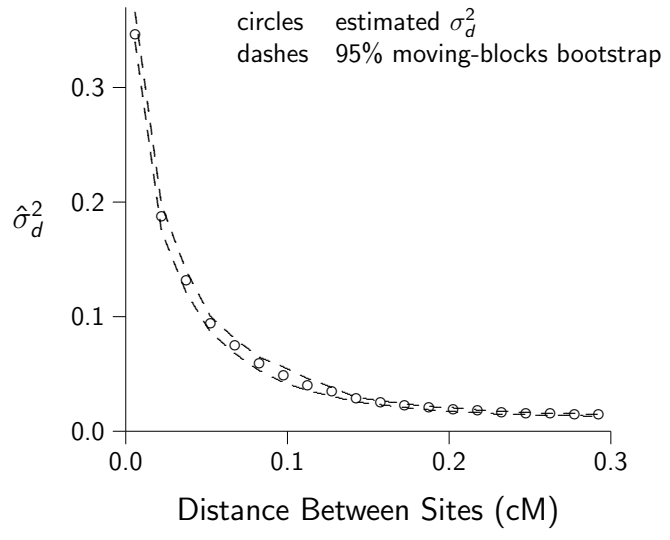- shuffle parental chromosomes.
- sites far apart shuffled more.
- destroys LD

Why population size affects LD

- Small populations have short genealogies.
- Little time for recombination to happen.
- Lots of LD.

σ_d^2 is a measure of LD

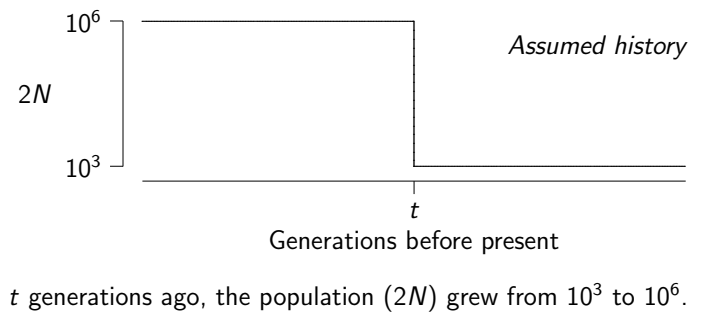- ▶ Insensitive to sequencing error.
- ▶ Sensitive to ancient history.

European LD curve for chromosome 1

circles   estimated σ_d^2
dashes   95% moving-blocks bootstrap

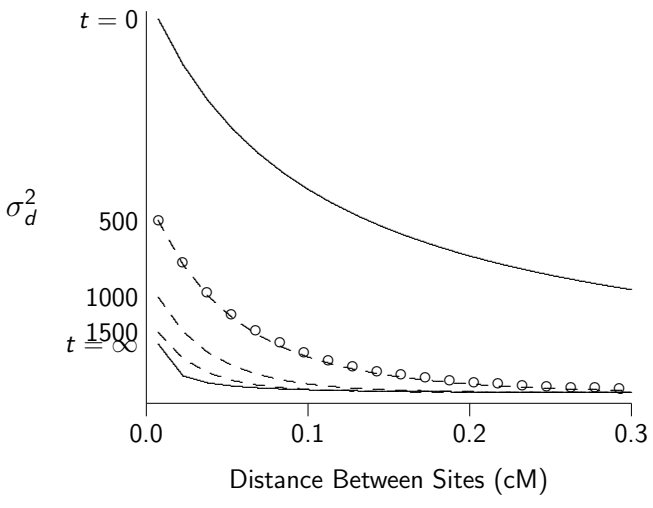$\hat{\sigma}_d^2$

Distance Between Sites (cM)

Outline

- ○ The European Neolithic: a movement of peoples or of technology?
- ○ Linkage disequilibrium (LD)
- ▶ How LD responds to changes in population size.
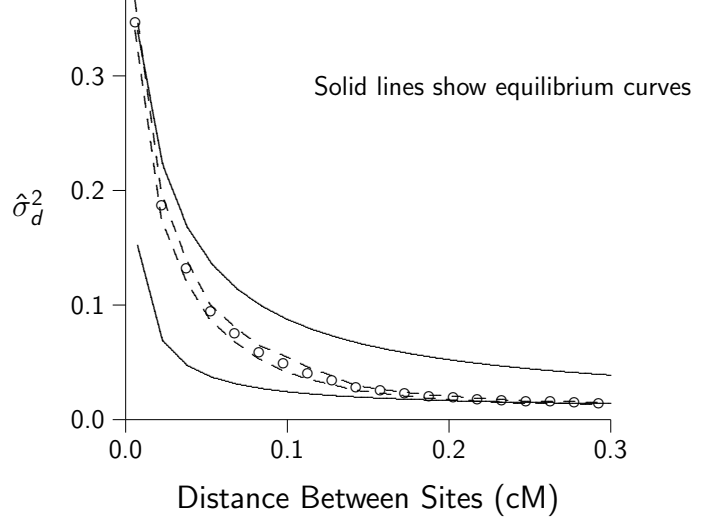- ▶ The history of European population size.
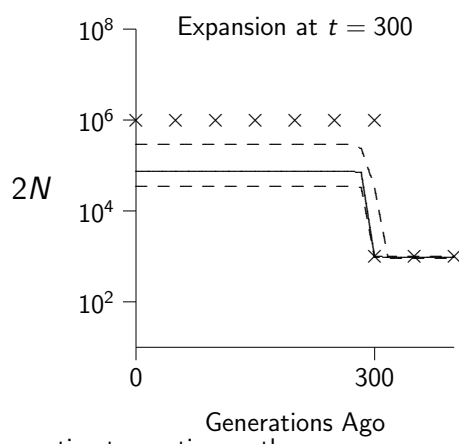
What does population growth do to the LD curve?

Assumed history

$2N$

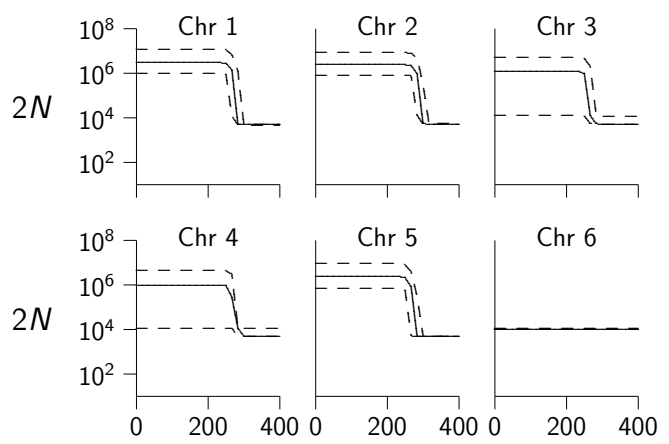Generations before present

t generations ago, the population ($2N$) grew from 10^3 to 10^6.

Population growth makes LD curve steep

$t = 0$

σ_d^2

500

1000

$t = 1500$ ∞

Distance Between Sites (cM)

European LD curve is steep

Solid lines show equilibrium curves

$\hat{\sigma}_d^2$

Distance Between Sites (cM)

Estimates from simulation of expanded population

Expansion at $t = 300$

$2N$

Generations Ago

Shows same estimates as time path.

Estimates from European chromosomes (1000-Genomes Prj)

Chr 1   Chr 2   Chr 3

Chr 4   Chr 5   Chr 6

Estimates from European chromosomes (1000-Genomes Prj)

Chr 7   Chr 8   Chr 9

Chr 10   Chr 11   Chr 12

Estimates from European chromosomes (1000-Genomes Prj)

Chr 13   Chr 14   Chr 15

Chr 16   Chr 17   Chr 18

Estimates from European chromosomes (1000-Genomes Prj)

Chr 19   Chr 20   Chr 21

Chr 22
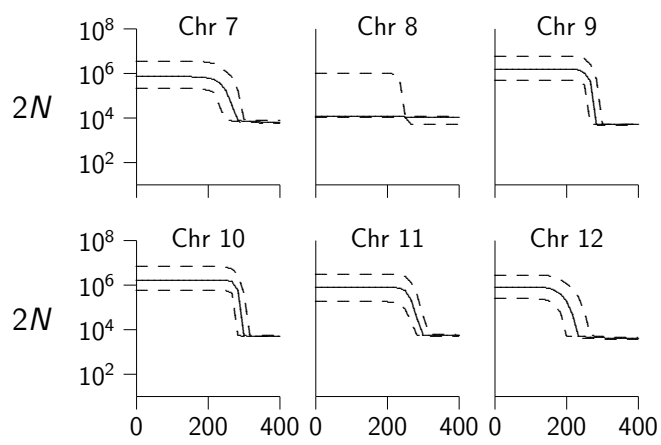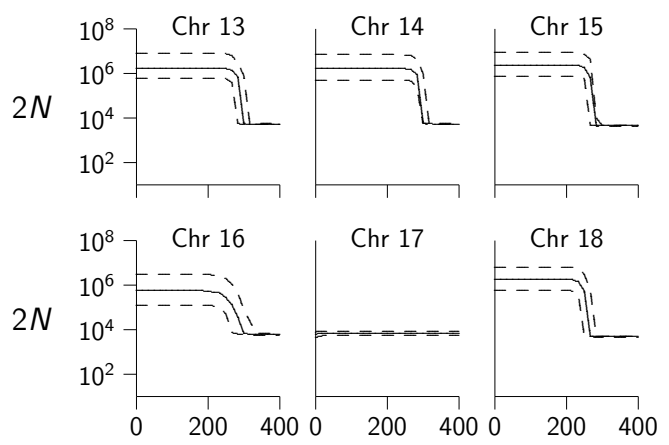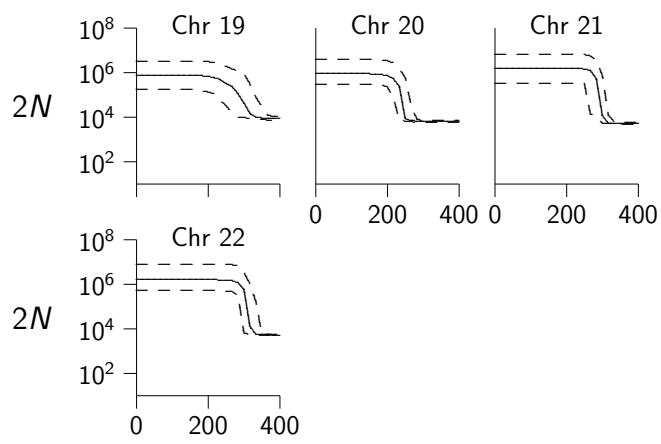
LD documents an expansion in European population size 300 generations ago.

Or roughly 8000 y ago.

Neolithic radiocarbon dates
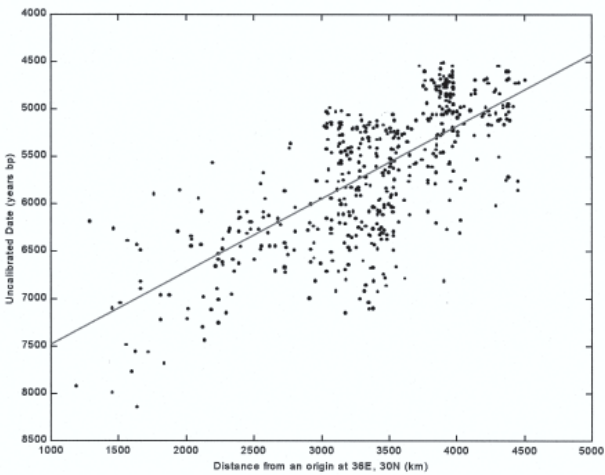
(Gkiasta et al, 2003)

Summary

- ▸ LD documents a Neolithic expansion of the European population.
- ▸ There was a movement of people during the Neolithic.