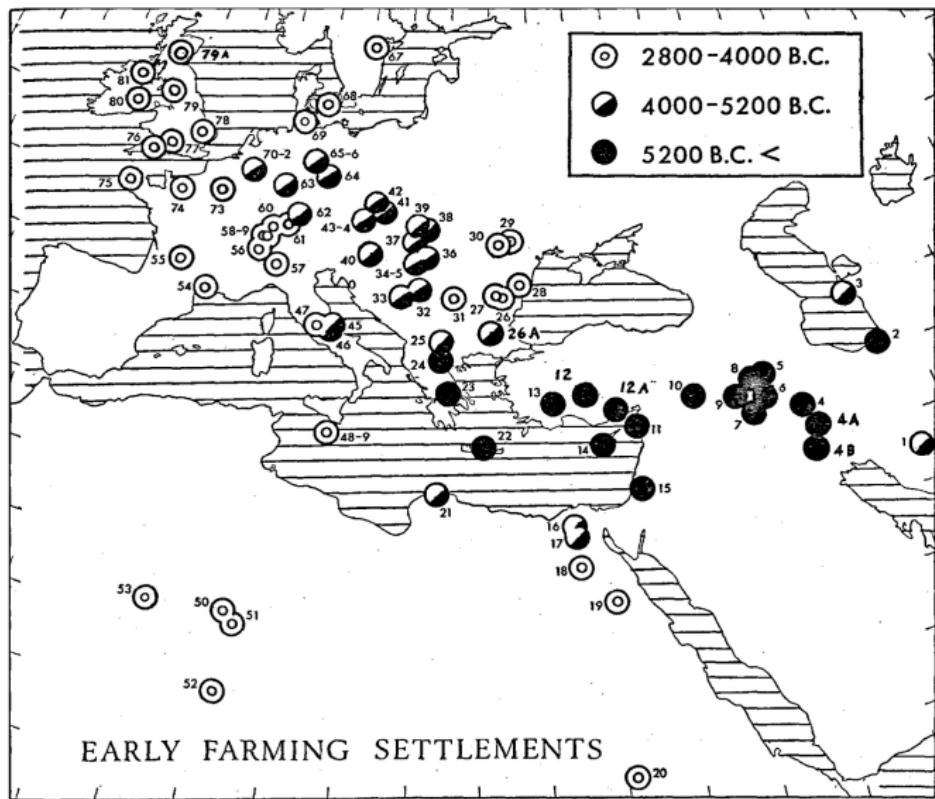


# Outline

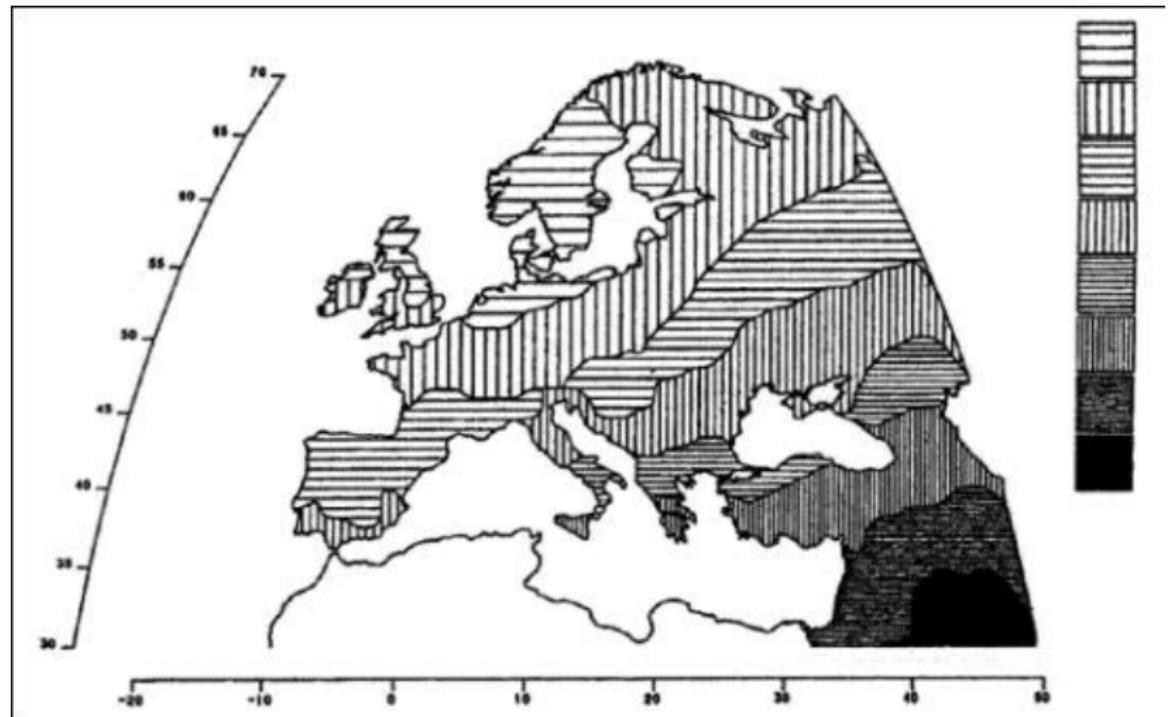
- ▶ The European Neolithic: a movement of peoples or of technology?
- ▶ Linkage disequilibrium (LD)
- ▶ How LD responds to changes in population size.
- ▶ The history of European population size.

# 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?

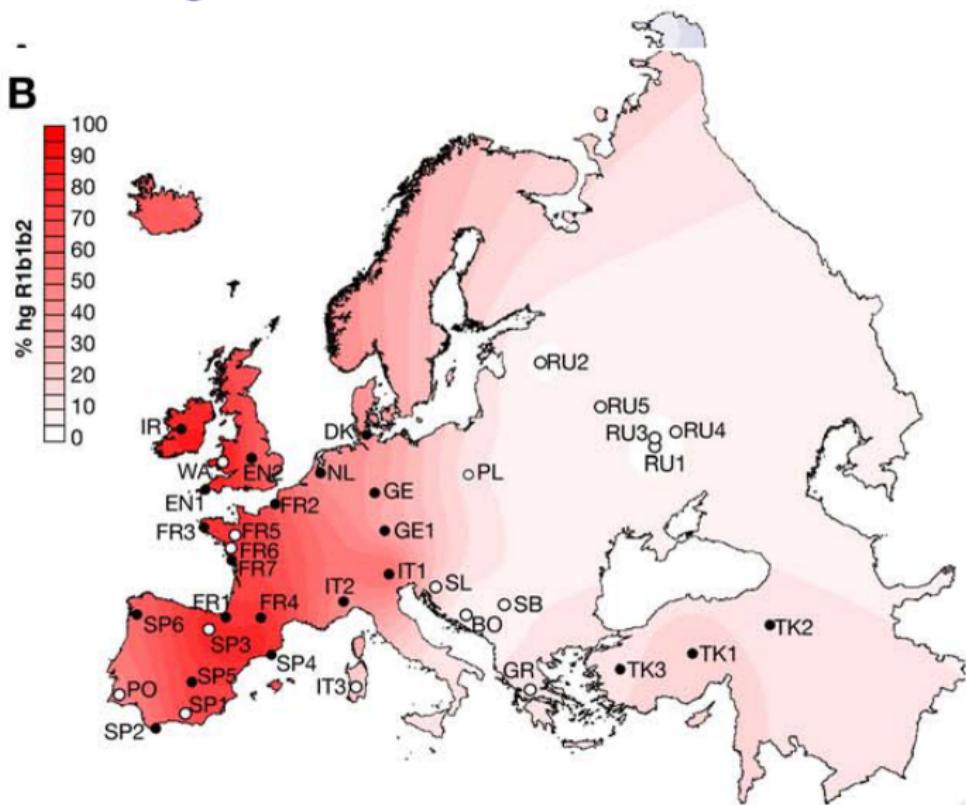
Local hunter-gatherers contributed less than 30% in the original settlements. This finding leads us to reject a predominantly cultural transmission of agriculture.

(Lounès Chikhi et al. 2002)

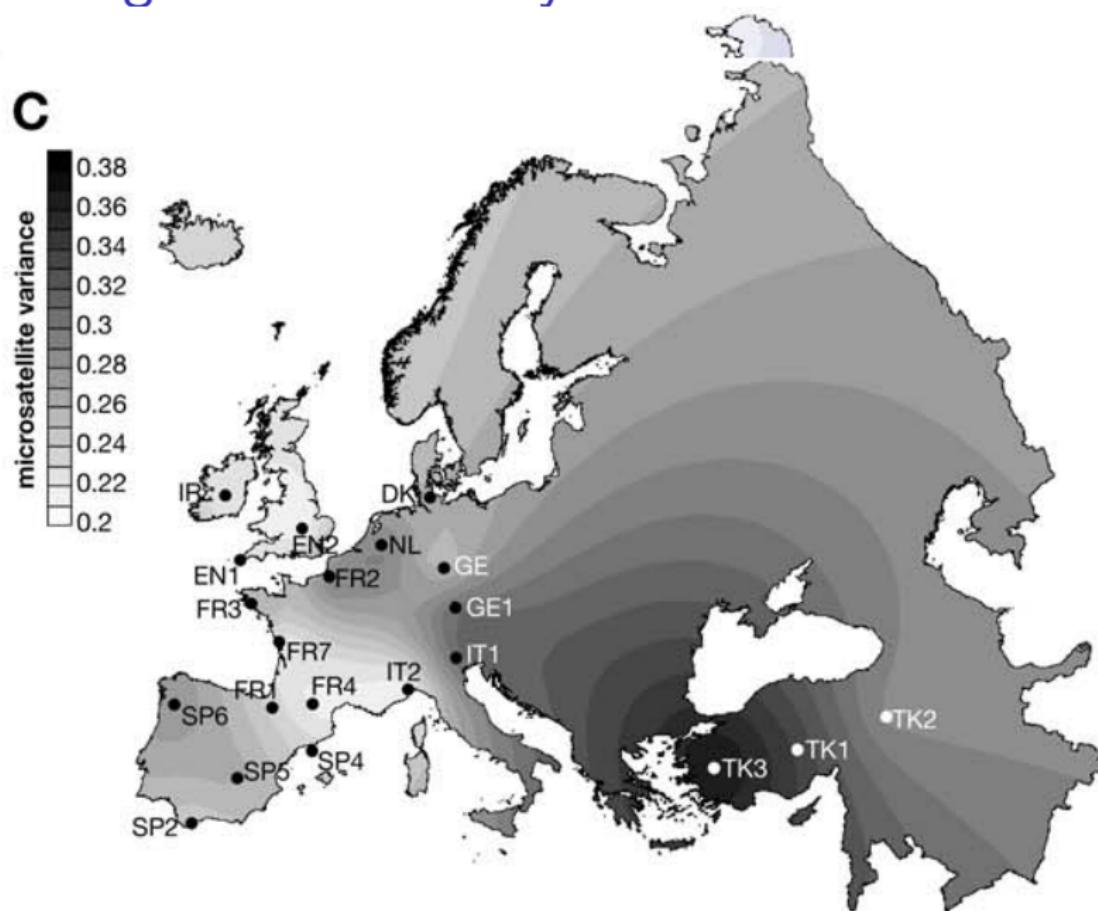
Both mitochondrial DNA and Y chromosome analyses have indicated a contribution of Neolithic Near Eastern lineages to the gene pool of modern Europeans of around a quarter or less. This suggests that dispersals bringing the Neolithic to Europe may have been demographically minor.

(Martin Richards 2003)

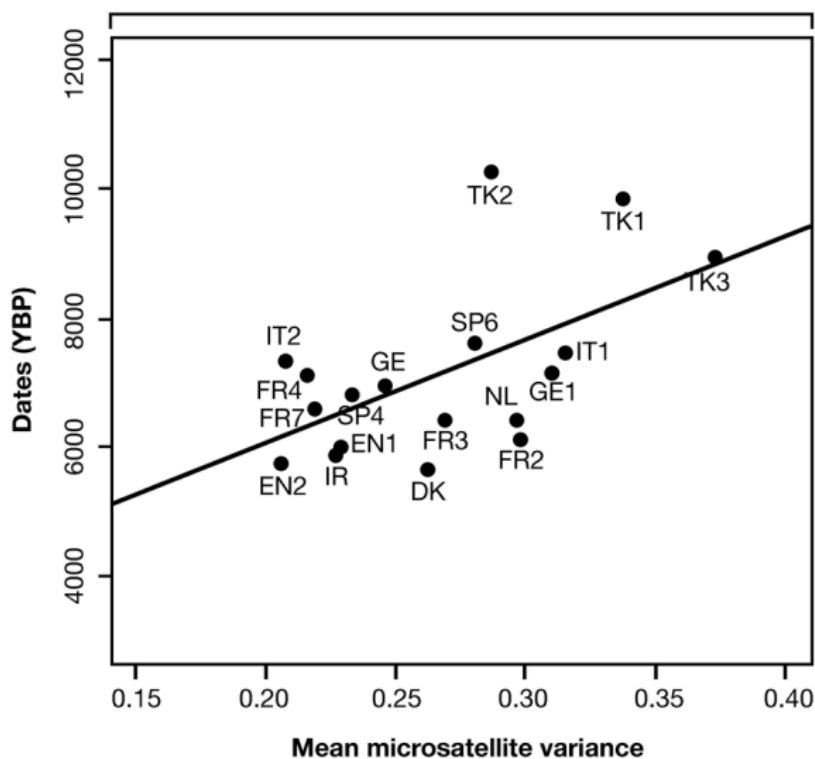
# Y haplogroup R1b1b2 most common in Ireland: Mesolithic origin?



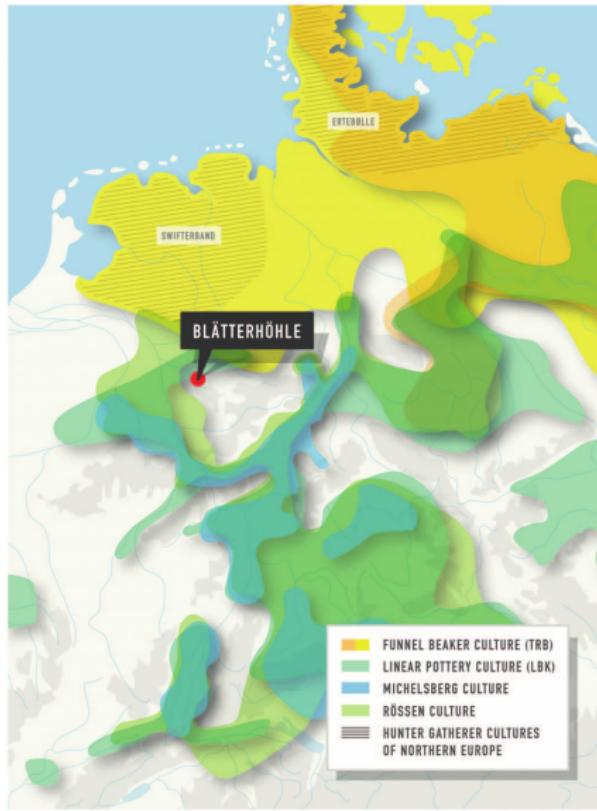
No: it originated in Turkey



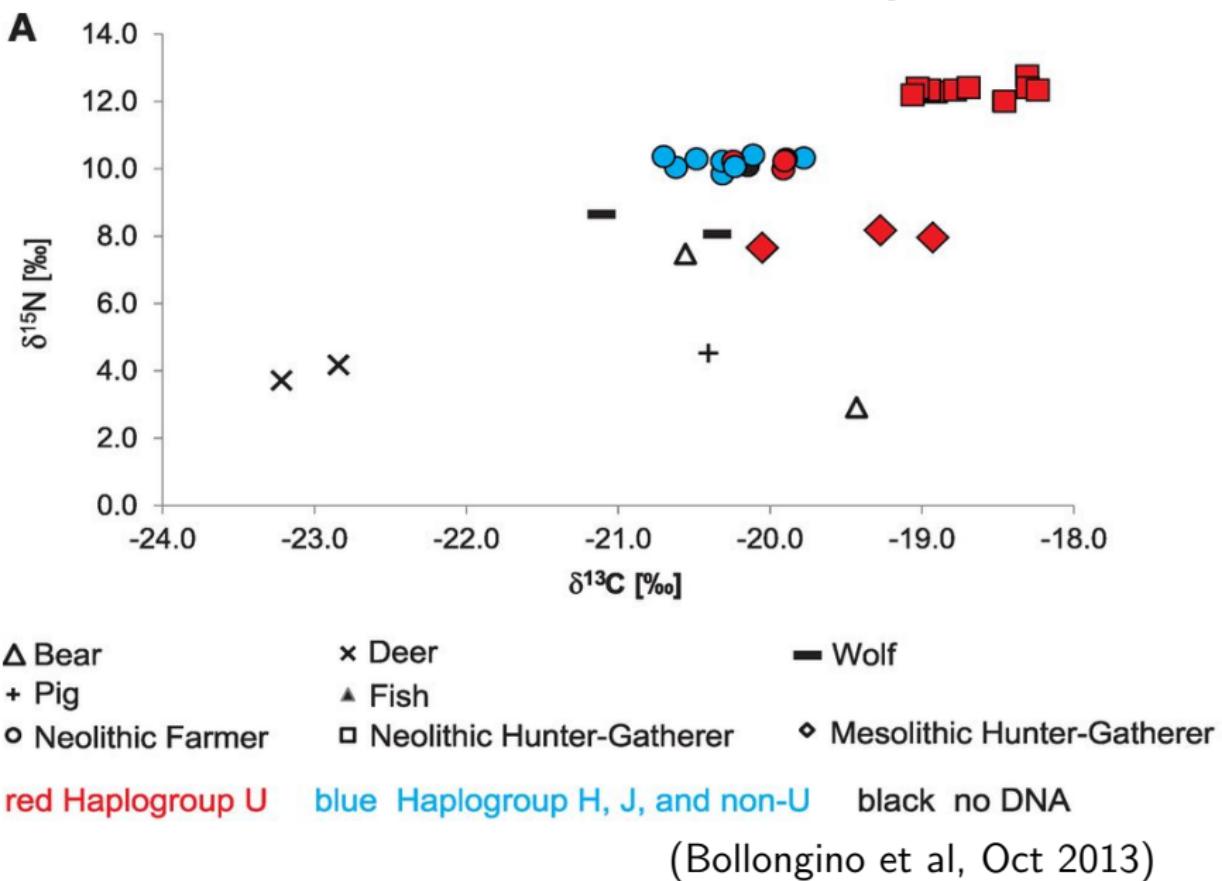
# Microsatellite variance vs. earliest Neolithic dates



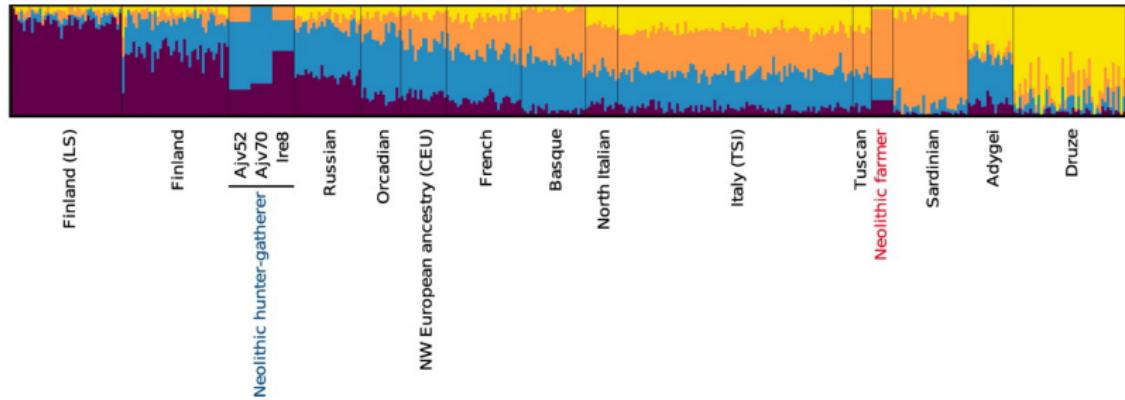
# The Blätterhöhle site in Germany



# mtDNA of Neolithic farmers and foragers

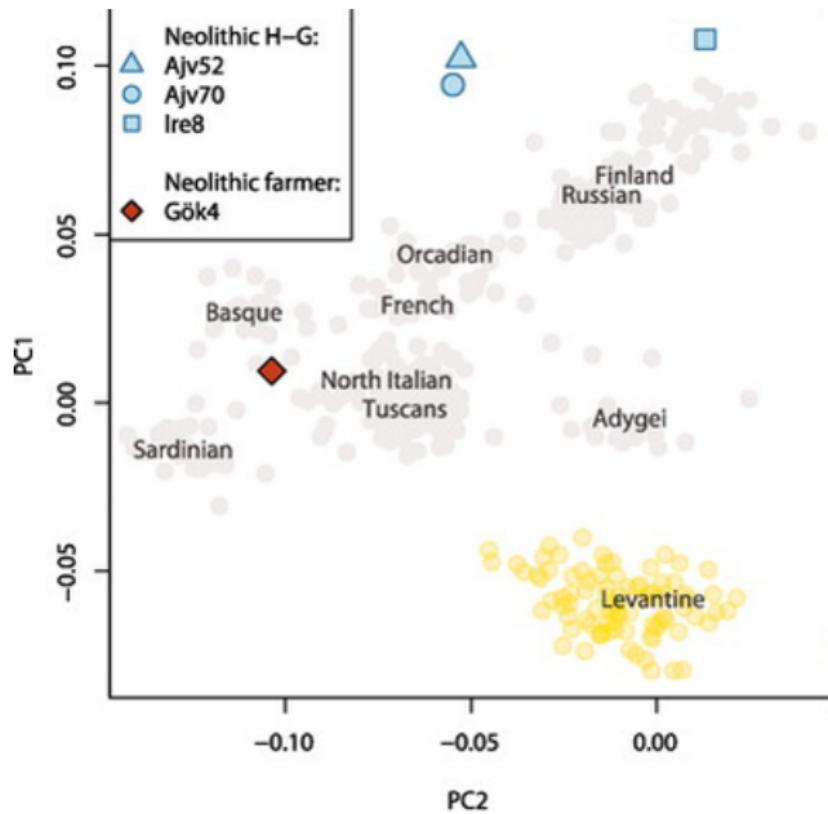


# Nuclear genes of Neolithic farmers and foragers



(Skoglund et al, 2012)

# Nuclear genes of Neolithic farmers and foragers



(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)

# Linkage disequilibrium (LD)

Gamete	Locus	
	1	2
1	A	B
2	A	B
3	A	B
4	A	B
5	A	B
6	A	b
7	a	B
8	a	B
9	a	b
10	a	b

A	a
B	5 2
b	1 2
	6 4
	7 3 10

- ▶  $B$  is more common among  $A$ -gametes than  $a$ -gametes.
- ▶  $A$  is more common among  $B$ -gametes than  $b$ -gametes.

# Linkage disequilibrium (LD)

Gamete	Locus	
	1	2
1	A	B
2	A	B
3	A	B
4	A	B
5	A	B
6	A	b
7	a	B
8	a	B
9	a	b
10	a	b

A	a
B	5
b	1
6	4
7	3
10	

- ▶  $B$  is more common among  $A$ -gametes than  $a$ -gametes.
- ▶  $A$  is more common among  $B$ -gametes than  $b$ -gametes.

# Linkage equilibrium (LE)

Gamete	Locus	
	1	2
1	A	B
2	A	B
3	A	B
4	A	B
5	A	b
6	A	b
7	a	B
8	a	B
9	a	b

	A	a
B	4	2
b	2	1
	6	3
		9

- ▶  $B$  is equally common among  $A$ -gametes and  $a$ -gametes.
- ▶  $A$  is equally common among  $B$ -gametes and  $b$ -gametes.

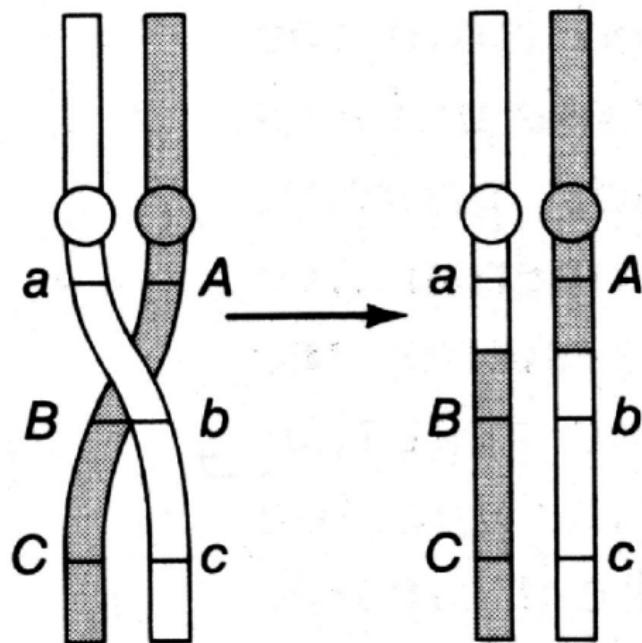
# You can see LD in sequence data

	Nucleotide position											
	1	1	1	1	1	1	1	1	2	2	2	2
	3	8	2	3	3	6	6	7	9	0	2	3
	1	9	4	6	4	4	3	3	3	5	2	0
	2	7	7	1	3	4	4	9	3	1	0	3
Orang	T	G	C	A	T	G	T	A	A	C	G	C
Chimp	T	G	C	A	T	G	T	A	A	T	G	C
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.

## Conventional measure of LD

$$D = P_{11} - ab$$

where  $P_{11}$  is frequency of gamete type  $A_1B_1$ ,  $a$  is frequency of allele  $A_1$  and  $b$  is that of  $B_1$ .

(Lewontin & Kojima, 1960)

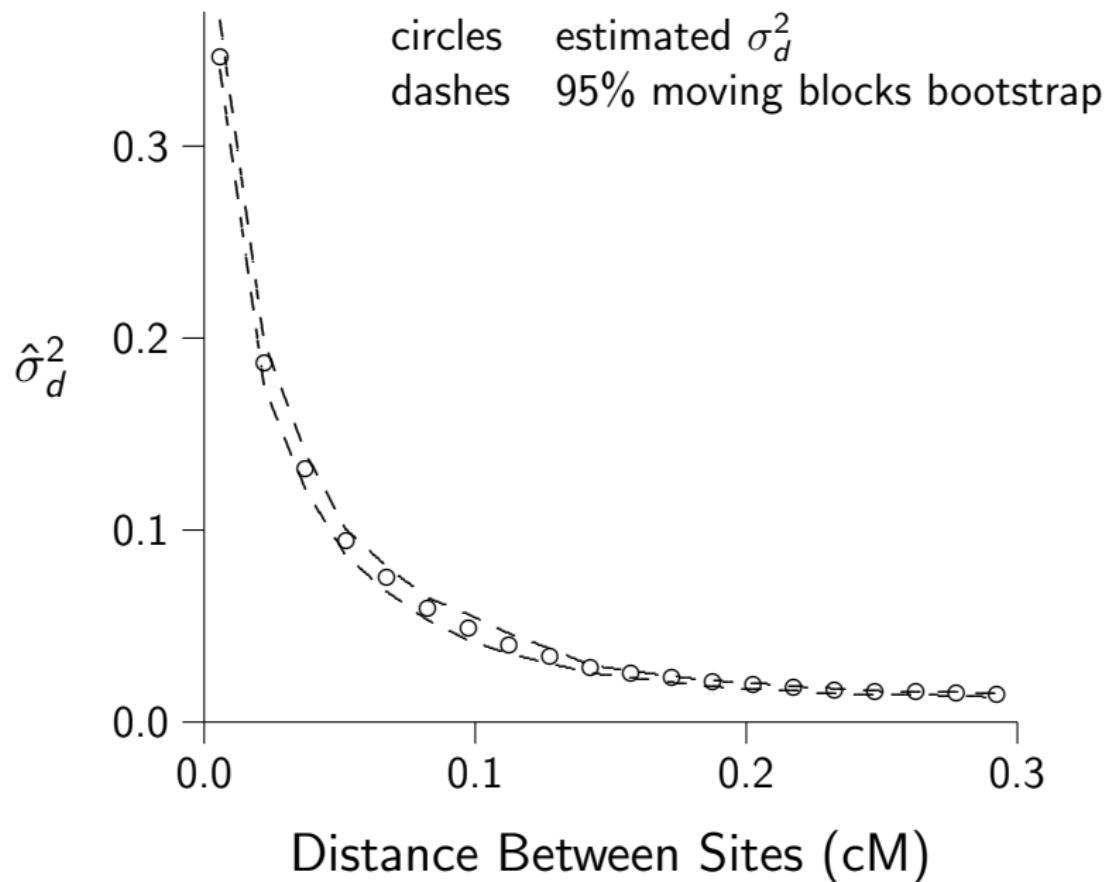
## An alternative measure

$$\sigma_d^2 = \frac{E[D^2]}{E[a(1-a)b(1-b)]}$$

(Ohta and Kimura, 1969)

- ▶ Sensitive to loci with deep gene trees.
- ▶ Allows us to see farther into the past.
- ▶ Insensitive to sequencing error.
- ▶ Deterministic theory relating  $\sigma_d^2$  to population history.

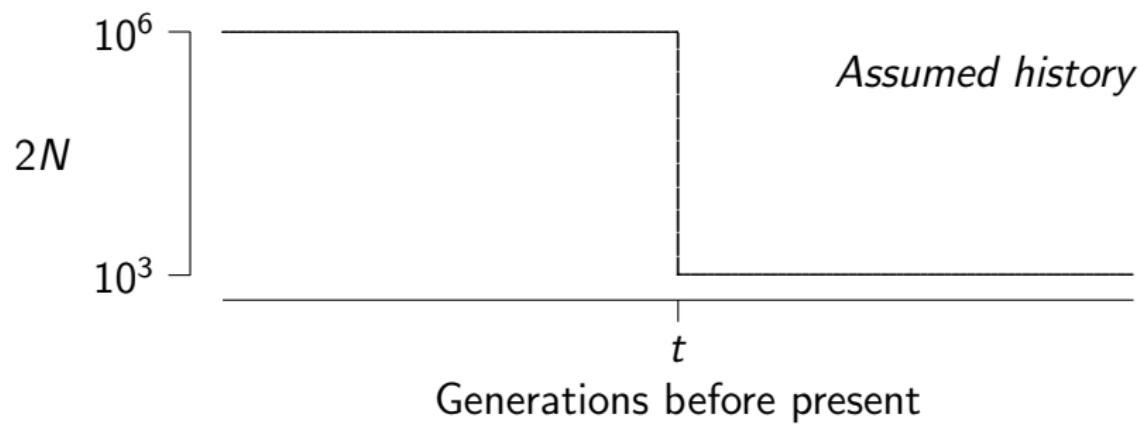
# European LD curve for chromosome 1



# Outline

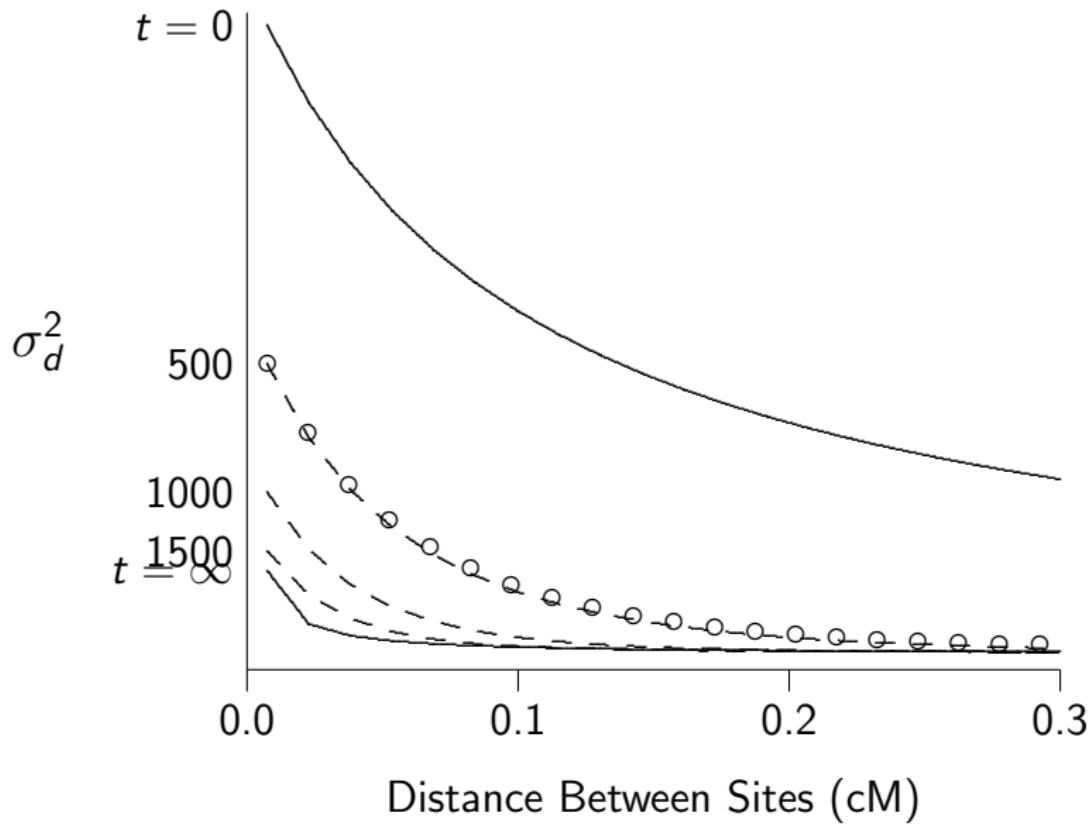
- The European Neolithic: a movement of peoples or of technology?
- Linkage disequilibrium (LD)
- ▶ How LD responds to changes in population size.
- ▶ The history of European population size.

# What does population growth do to the LD curve?

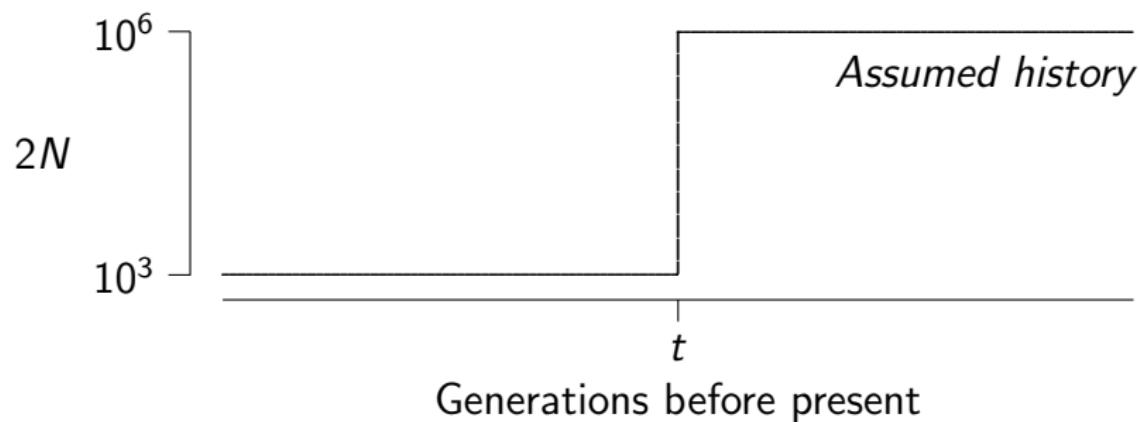


$t$  generations ago, the population ( $2N$ ) grew from  $10^3$  to  $10^6$ .

After population growth, right side delines fastest

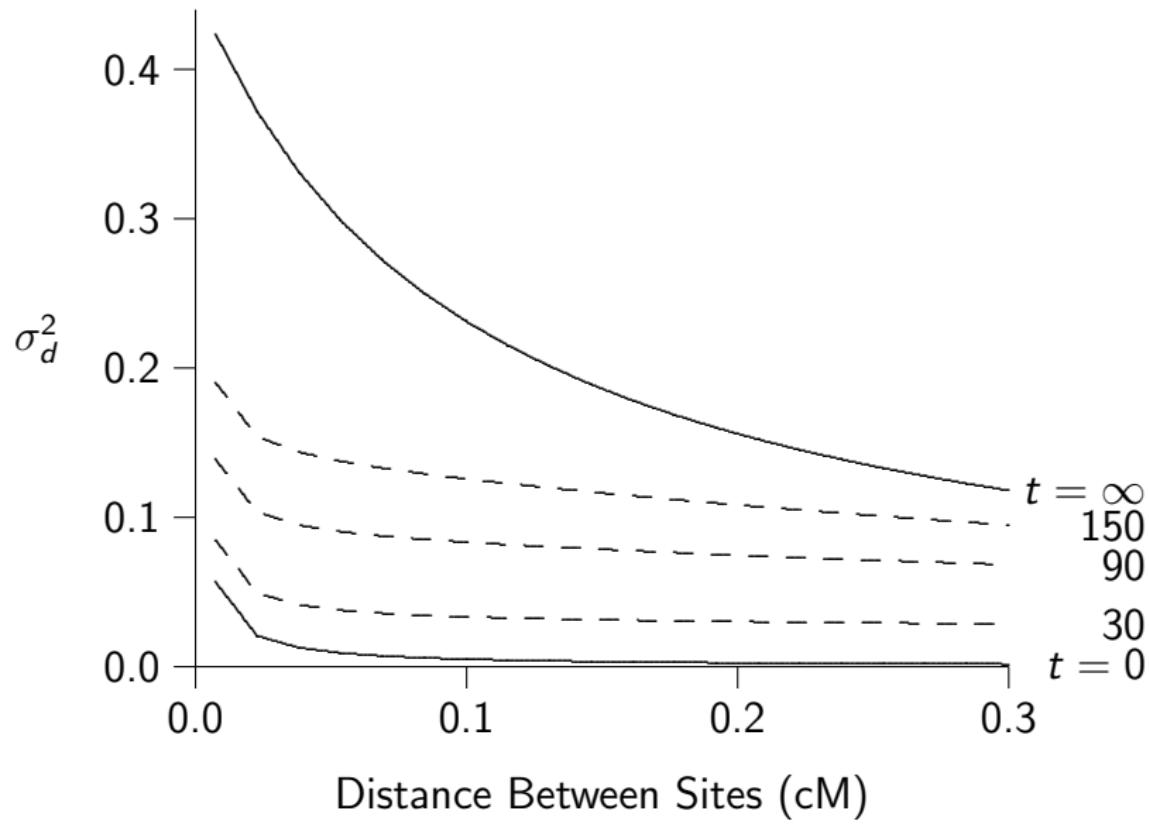


## What does population decline do to the LD curve?

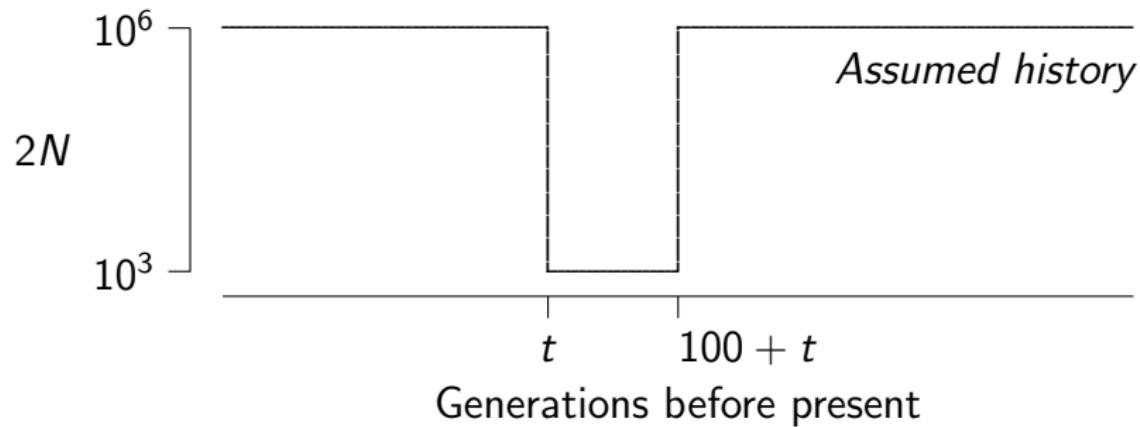


$t$  generations ago, the population ( $2N$ ) shrank from  $10^6$  to  $10^3$ .

After population decline, whole curve rises

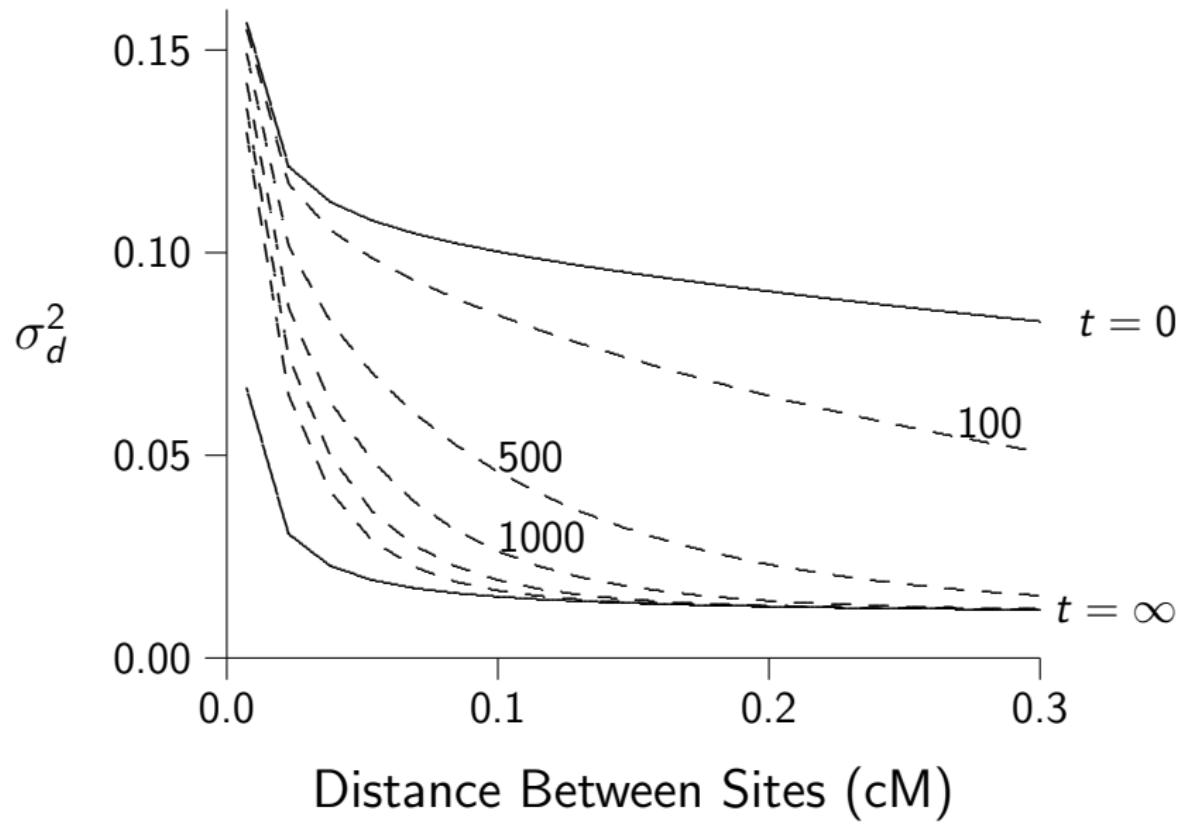


# What does a “bottleneck” do to the LD curve?



$t$  generations ago, the population recovered from a 100-generation bottleneck.

## Bottleneck makes curve very steep



## Effect of population history on LD curve

- ▶ After growth, right side of curve falls fastest
- ▶ After bottleneck, this effect is exaggerated
- ▶ After population decline, curve rises but stays flat
- ▶ We can use these principles to infer population history.

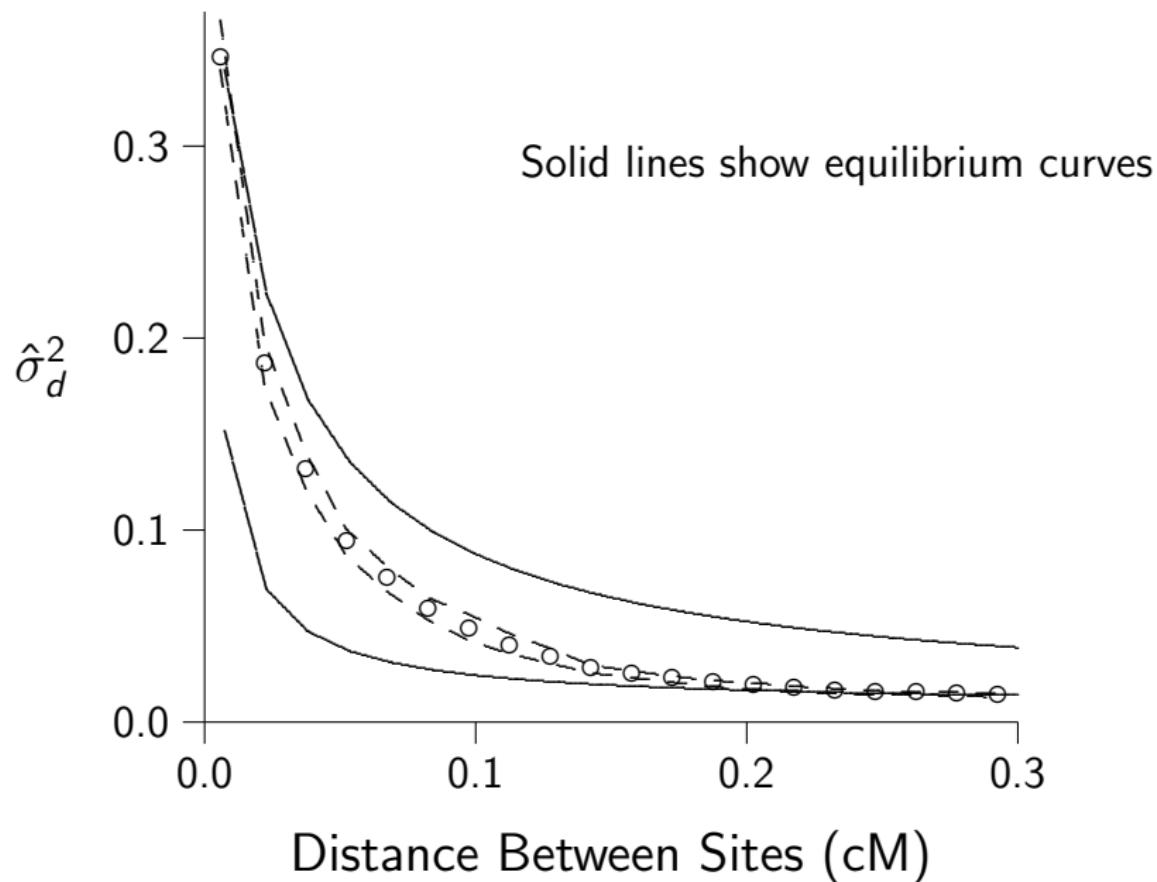
## Effect of population history on LD curve

- ▶ After growth, right side of curve falls fastest
- ▶ After bottleneck, this effect is exaggerated
- ▶ After population decline, curve rises but stays flat
- ▶ We can use these principles to infer population history.

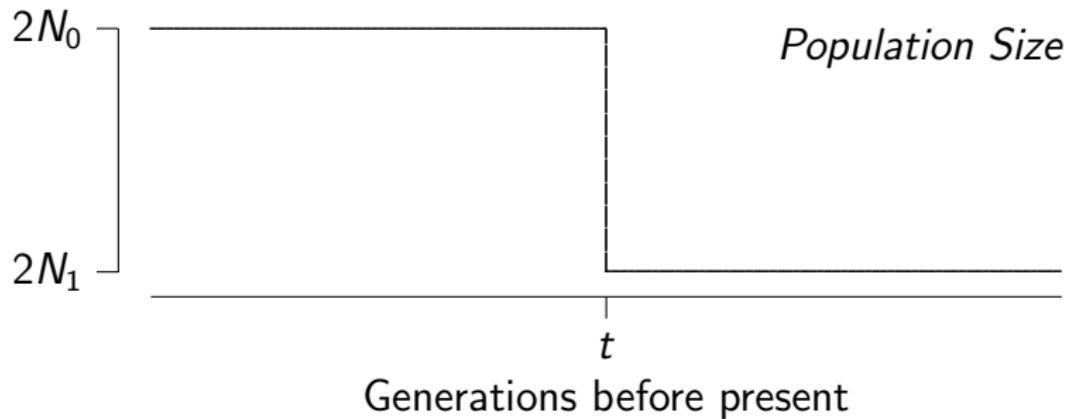
# Outline

- 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.

## European LD curve is steep

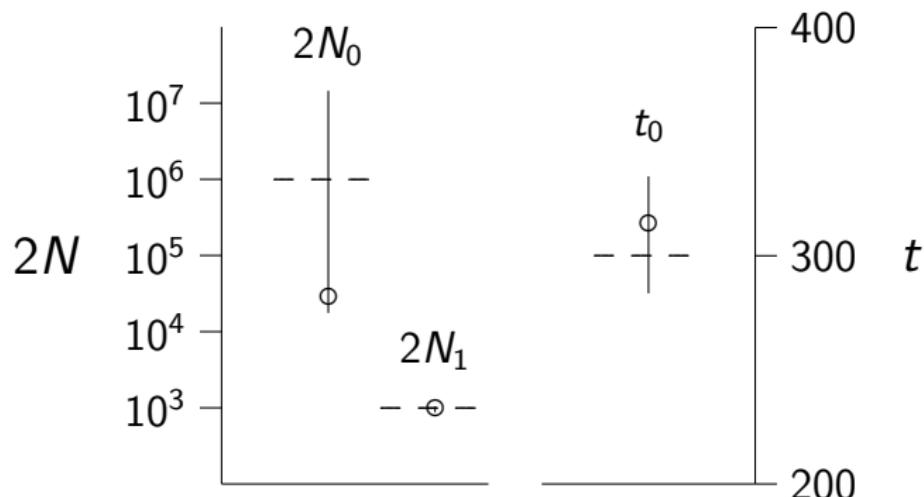


## 3-parameter model of history



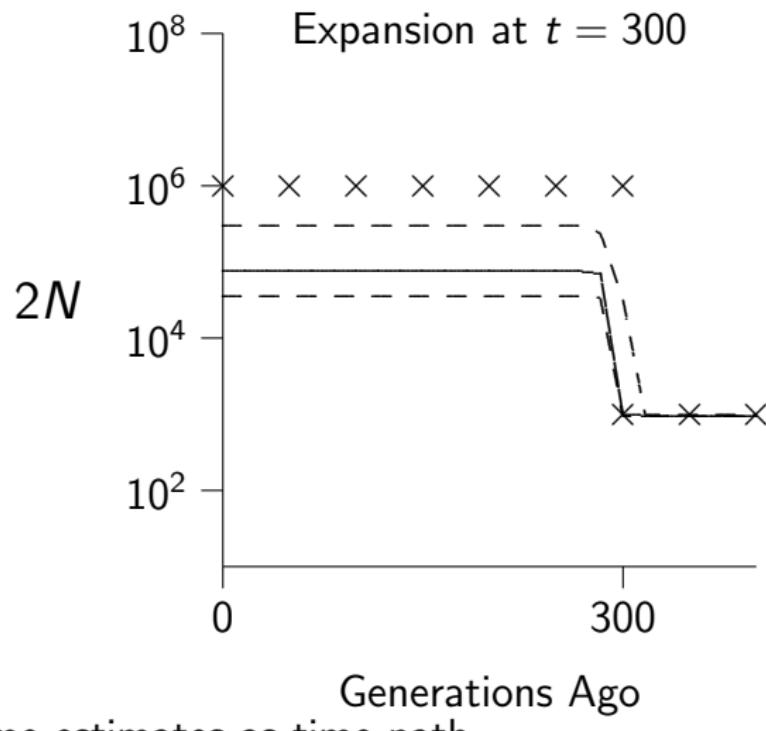
To estimate parameters, I search for values of  $2N_0$ ,  $2N_1$ , and  $t$  that provide best fit to observed LD curve.

## Estimates from simulation of expanded population

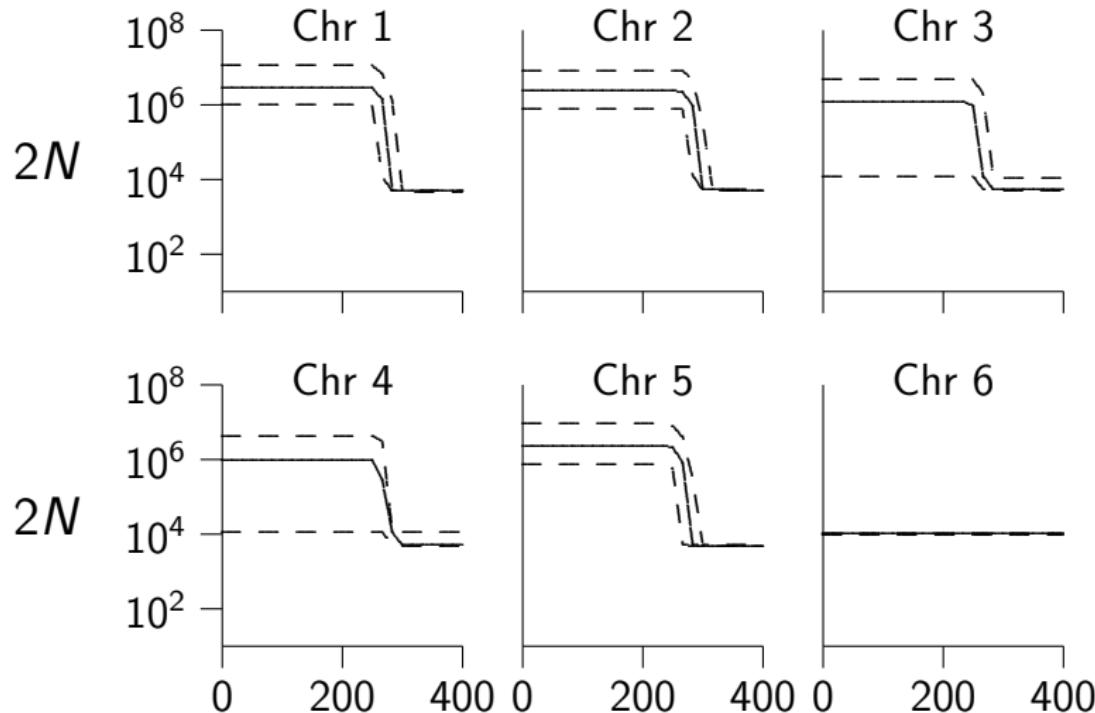


Dashed lines show simulation parameters:  $t_0 = 300$ ,  $2N_1 = 10^3$ , and  $2N_0 = 10^6$ . Open circles show parameter estimates, and vertical lines indicate 95% confidence intervals.

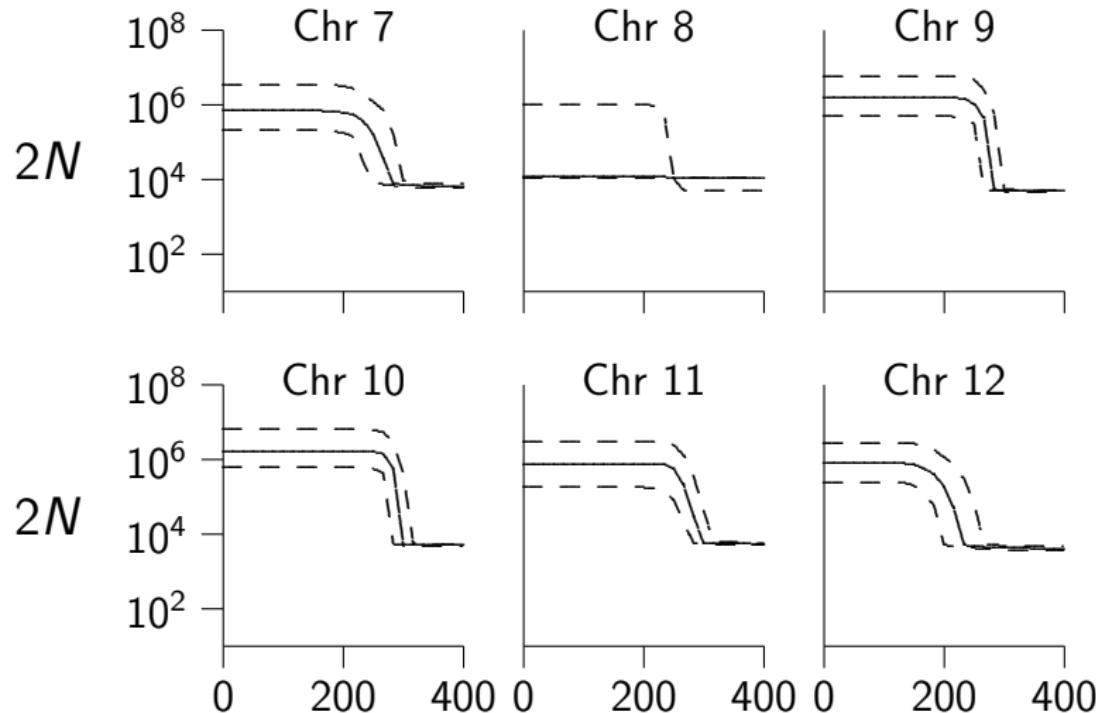
## Estimates from simulation of expanded population



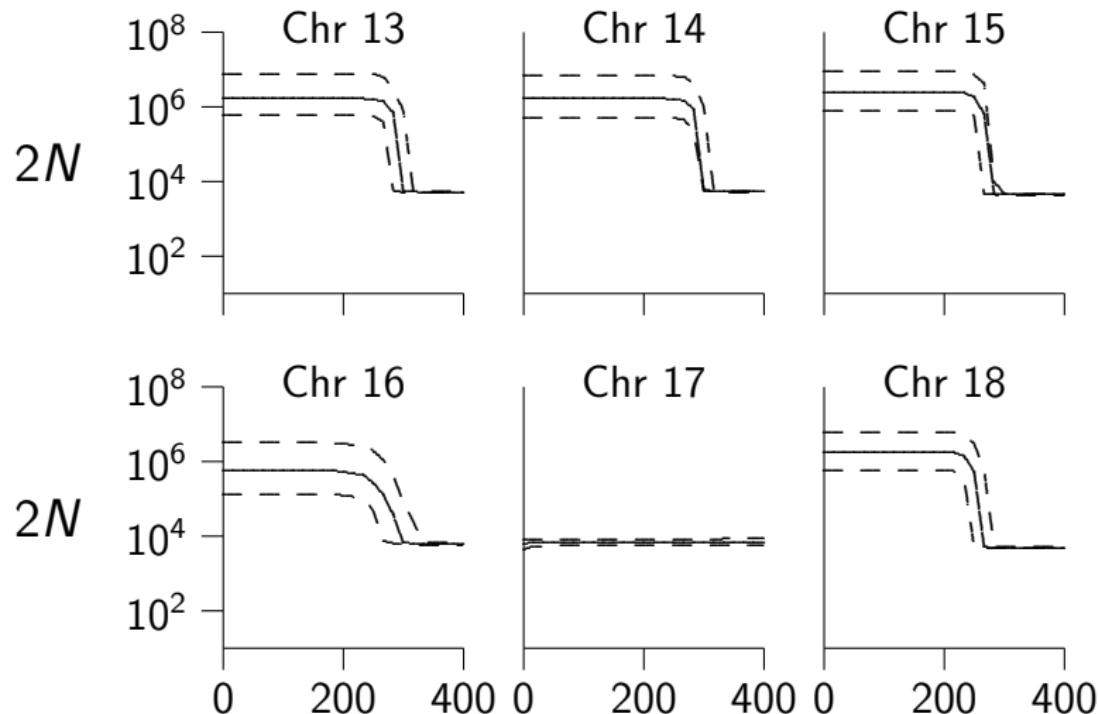
# Estimates from European chromosomes (1000-Genomes Prj)



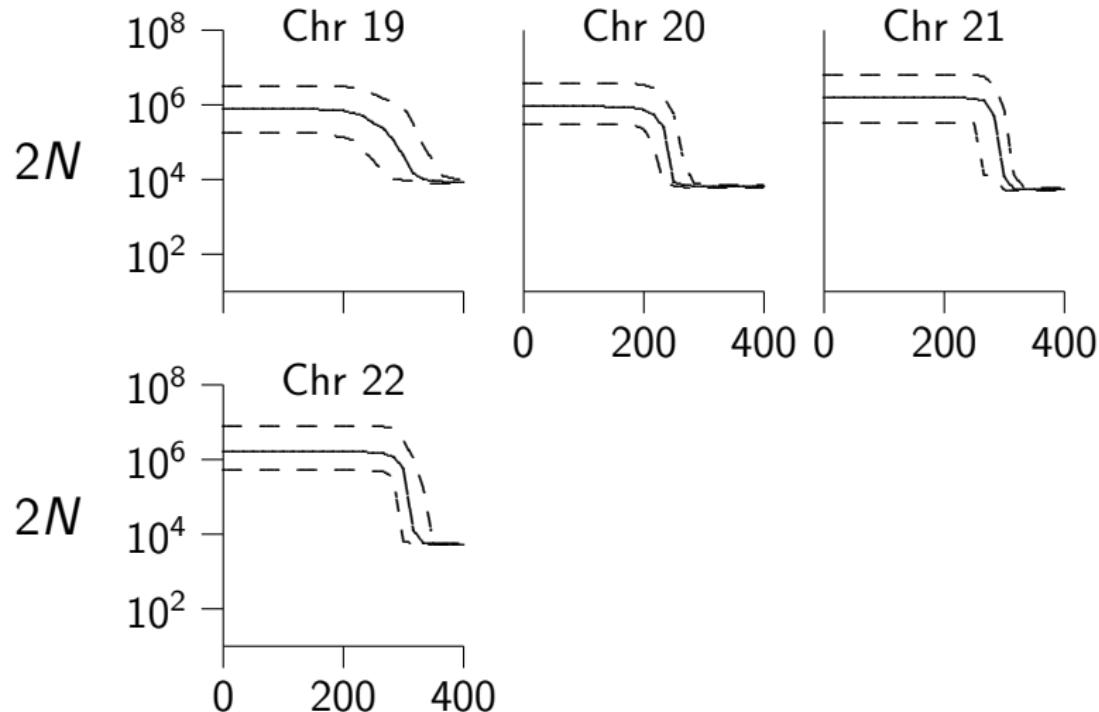
# Estimates from European chromosomes (1000-Genomes Prj)



# Estimates from European chromosomes (1000-Genomes Prj)



# Estimates from European chromosomes (1000-Genomes Prj)



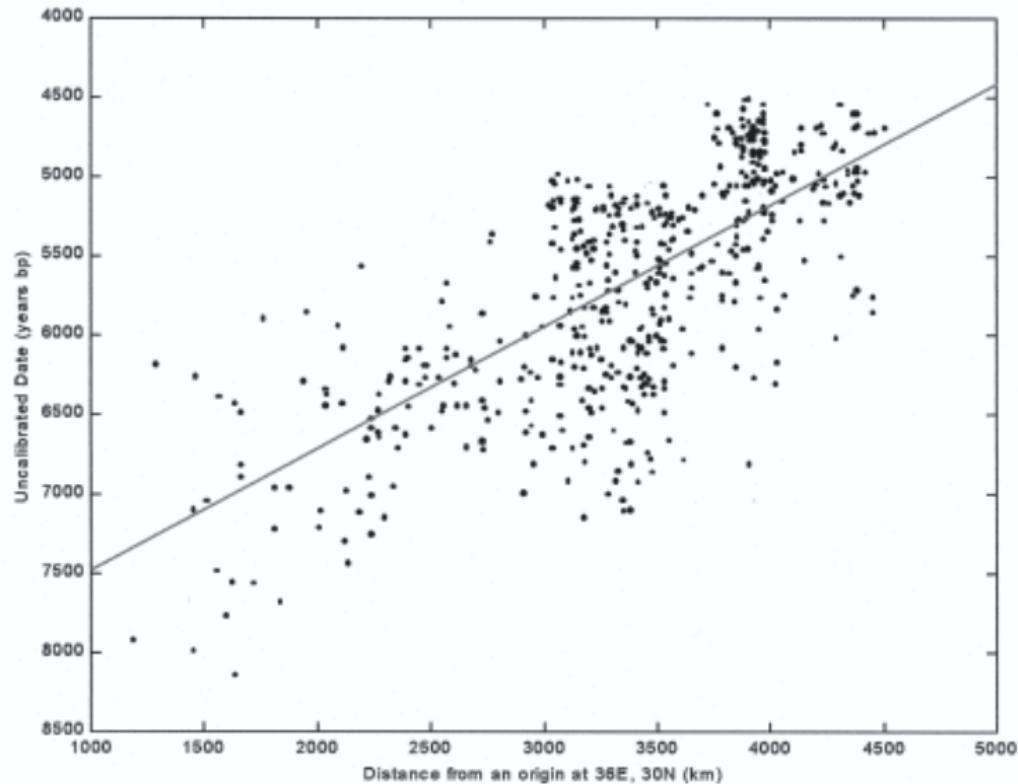
LD documents an expansion in European population size 300 generations ago.

Or roughly 8000 y ago.

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.