Computational Genetics Spring 2013 Lecture 11

Eleazar Eskin

University of California, Los Angeles

(slides from Michael Palmer, Serafim Batzoglou, Jeff Wall, and Alan Mann)

Human Origins

Lecture 11.

February25th, 2013.

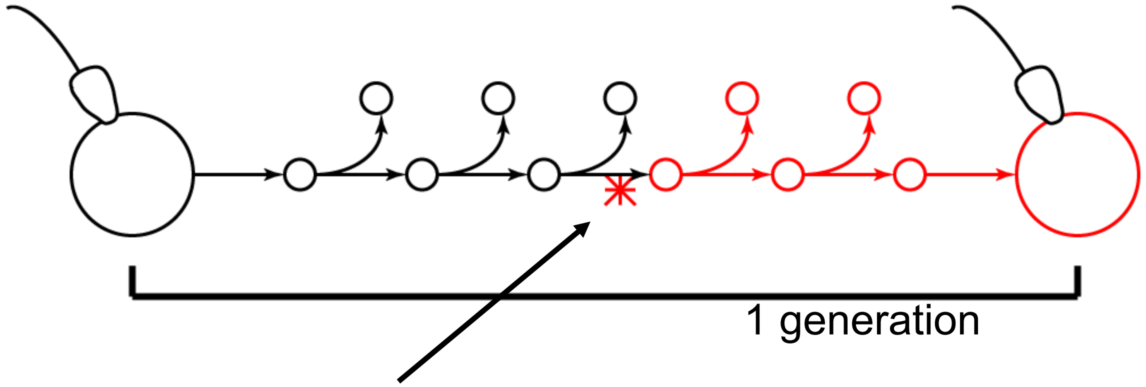
Overview

- History
- Some Population Genetics
  - **origins of genetic variation**
  - **evolutionary timescales**
  - **selection and drift**
  - **neutral theory**
- Detection of Selection in Humans with SNPs
- Some more Population Genetics
  - **Migration**
  - **Wright's F_{ST}**
- Inference of Human Phylogenetic Tree
- Time to Most Recent Common Ancestor (TMRCA)
- Unique Origin vs. Multiregional Evolution Models
- Geographic Origin of Humans

History of Study of Human Variation

- Blood proteins (ABO gene, 1919)
- Radioisotopes to study DNA
- Polymerase Chain Reaction (PCR), 1986
  - **method to “amplify” (copy) a piece of DNA**
  - **led to an explosion of DNA sequence data**
- Almost every protein has genetic variants
- These variants are useful markers for population studies

Origins of Genetic Variation



Number of cell divisions from one generation to next

	Mouse	Human
Male	~40	~400
Female	~20	~23

How often does *this* happen per generation? (germ line matters,

Rate of Genetic Events (avg) in Mammals

Point substitution (nuc)	~0.5 x 10^{-8} per bp
Microdeletion (1-10bp)	about 1/20 of point
Microinsertion (1-10bp)	about half of μdel
Recombination	~10^{-10}
Mobile element ins’ n	~10^{-11}
Inversion	?? much rarer

Exceptions

Hypermutable sites

C->T = 10x avg point rate

Simple Sequence Repeats

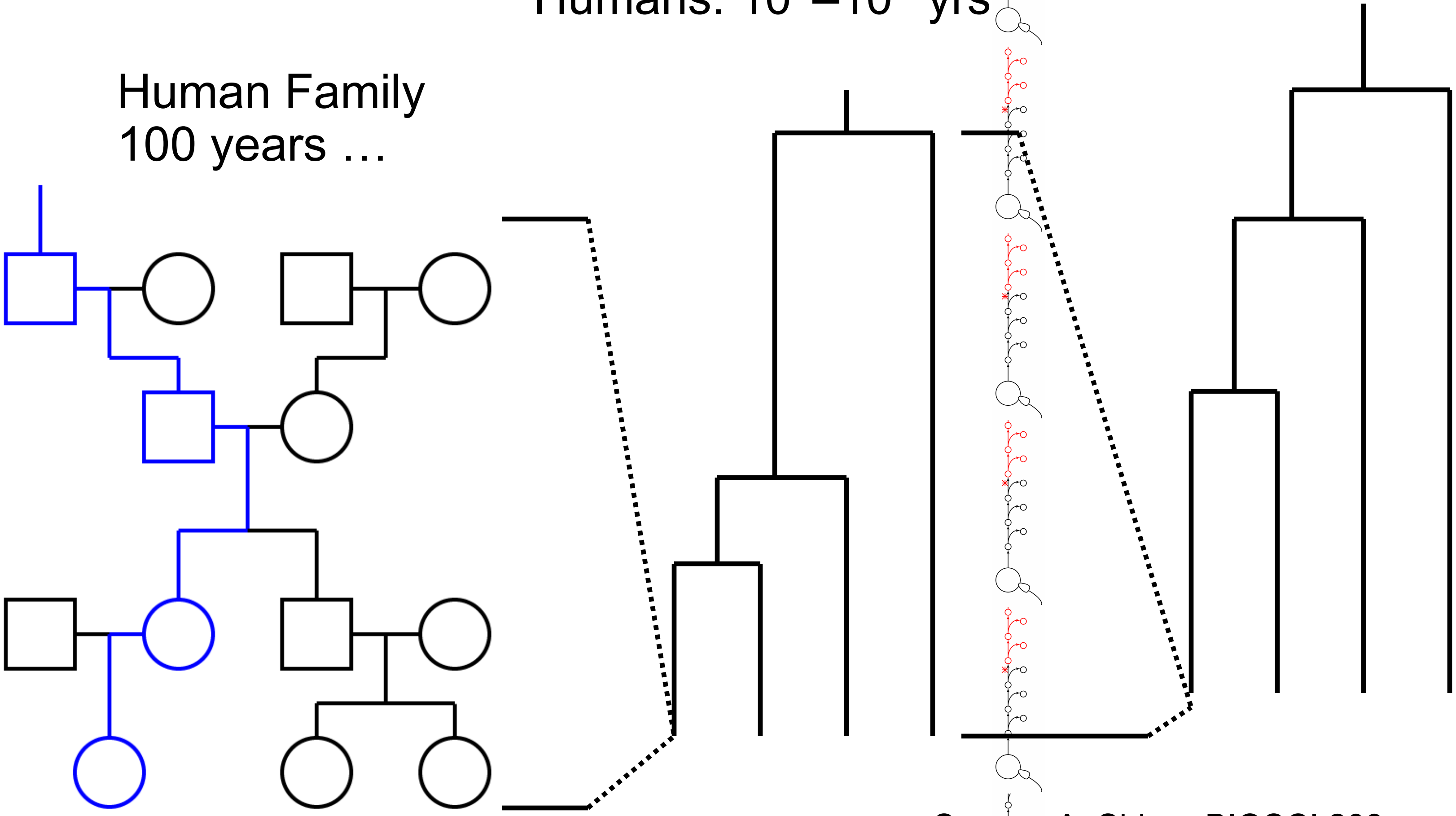
10-1000x indel rate (some 10^{-4}!)

mitochondrial DNA

10-100x nuclear point rate

Source: A. Sidow, BIOSCI 203

Accumulation of Variation over Time



Human Family
100 years …

Mammals:10^8 years

Apes: 10^7 years

Humans: 10^5–10^6 yrs

Eukaryotes:
10^9 years

Source: A. Sidow, BIOSCI 203

Drift and Selection

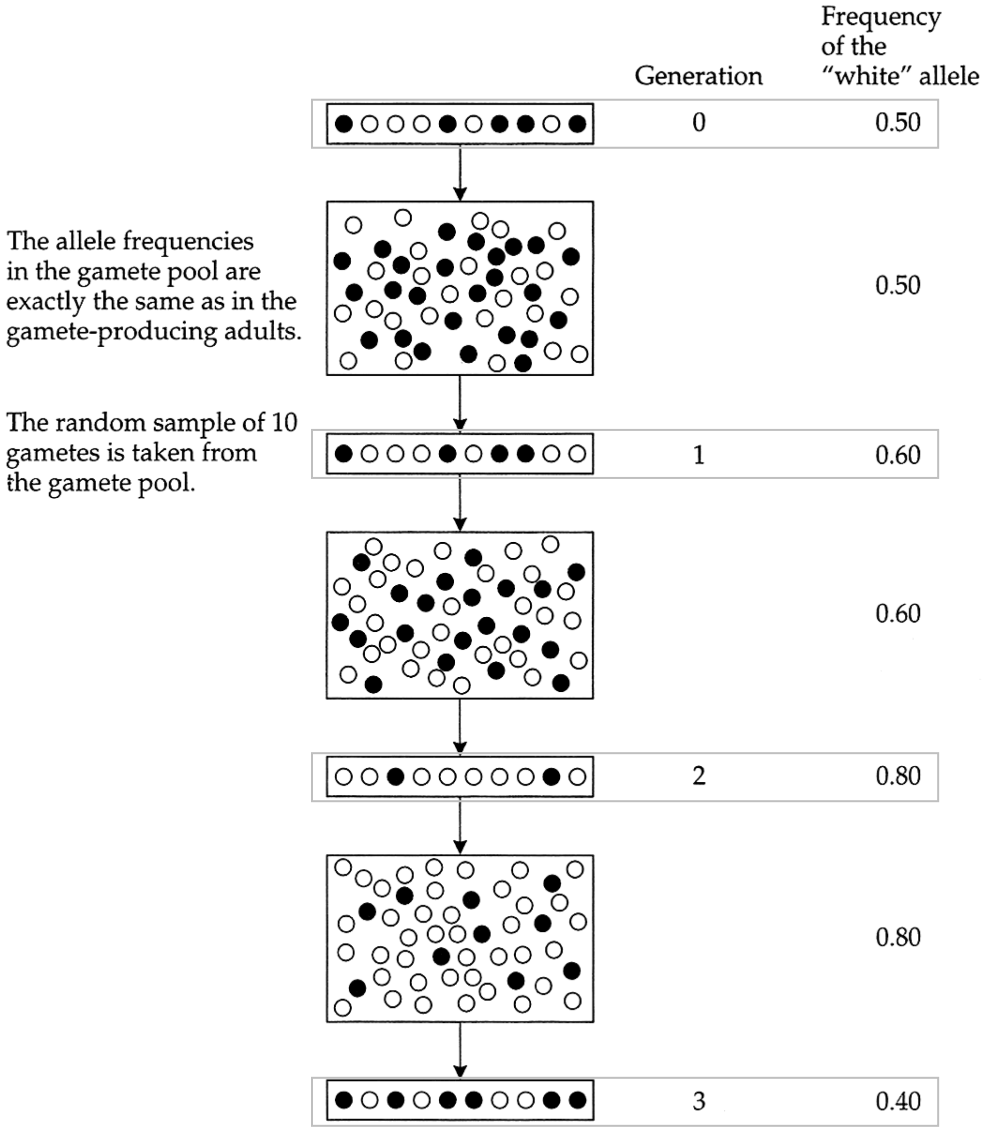
The two forces that determine the fate of alleles in a population

- Drift
  - □ **Change in allele frequencies due to <u>sampling</u>**
  - □ **a ‘stochastic’ process**
  - □ **Neutral variation is subject to drift**

- Selection
  - □ **Change in allele frequencies due to <u>function</u>**
  - □ **‘deterministic’**
  - □ **Functional variation may be subject to selection (more later)**

Source: A. Sidow, BIOSCI 203

Genetic Drift 1



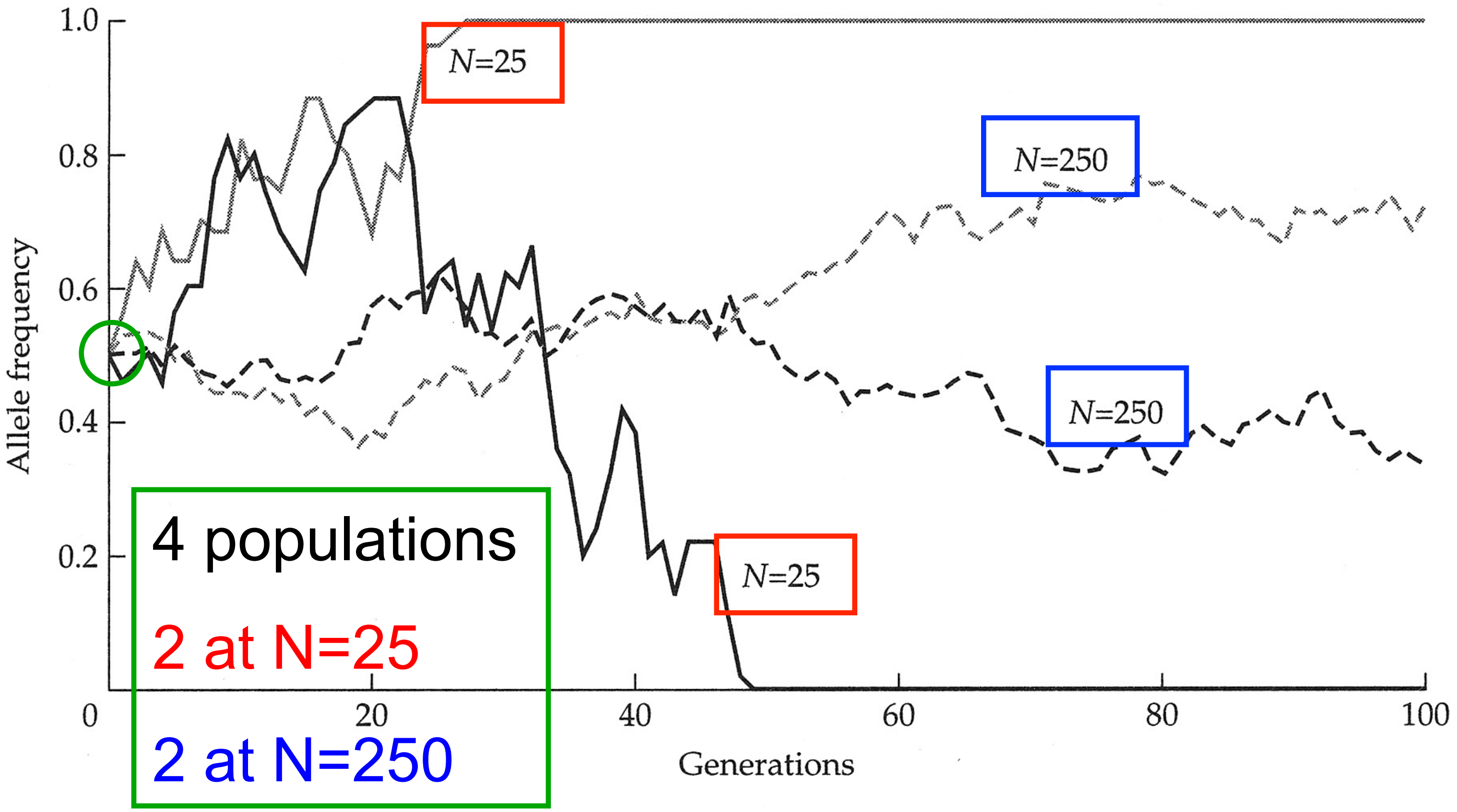
The allele frequencies in the gamete pool are exactly the same as in the gamete-producing adults.

The random sample of 10 gametes is taken from the gamete pool.

Figure 2.3 Random sampling of gametes. Allele frequencies in the gamete pools (large boxes) in each generation are assumed to reflect exactly the allele frequencies in the adults of the parental generation (small boxes). Since the population size is finite, allele frequencies fluctuate up and down. Modified from Bodmer and Cavalli-Sforza (1976).

From Li (1997) Molecular Evolution, Sinauer Press, via A. Sidow BIOSCI 203

Genetic Drift 2: Population Size Matters



4 populations

2 at N=25

2 at N=250

Figure 2.4 Changes in frequencies of alleles subject to random genetic drift in populations of different sizes (*N*). In each generation, 2*N* genes were sampled with replacement from the previous generation. For each population size, two replicates are presented. It is assumed that the effective population size N_e is equal to the actual size *N*.

From Li (1997) Molecular Evolution, Sinauer Press, via A. Sidow BIOSCI 203

Genetic Drift over time – expected values

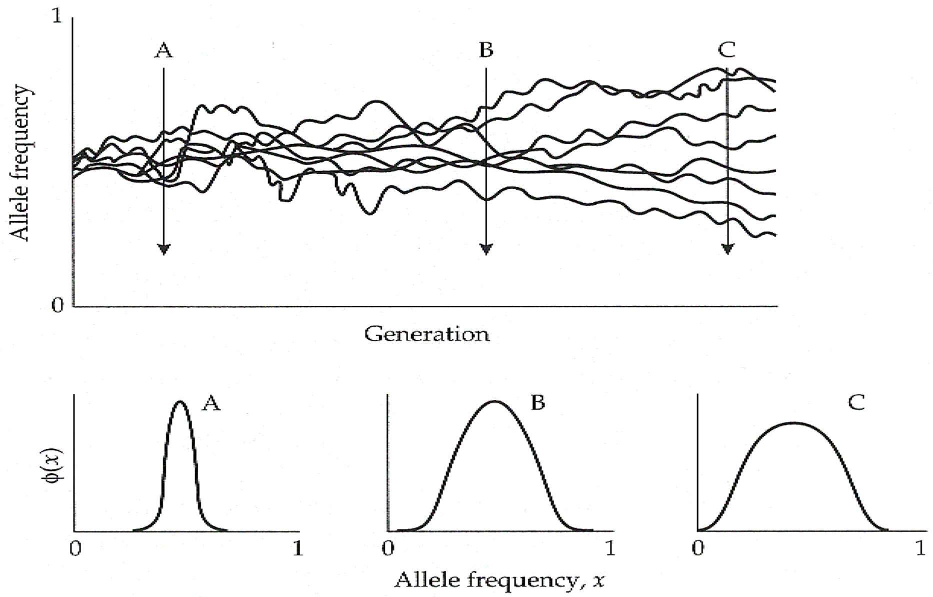


Figure 7.3 The model of random genetic drift can be seen by imagining a large collection of populations undergoing the process of repeated sampling. As the top part of the figure indicates, the populations' allele frequencies change erratically, and tend to drift apart. At time intervals, a snapshot of the populations would produce distributions of allele frequencies whose variance increases over time.

Principles of Population Genetics, Hartl and Clark

Genetic Drift over time – expected values

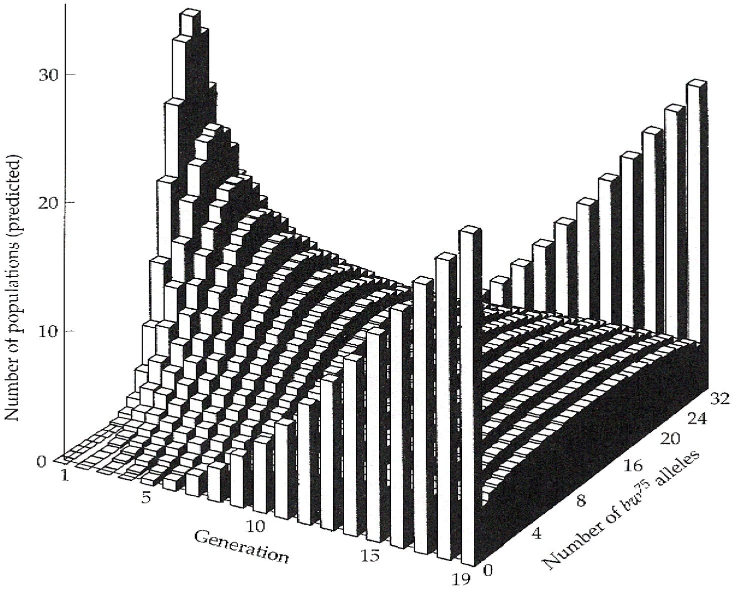
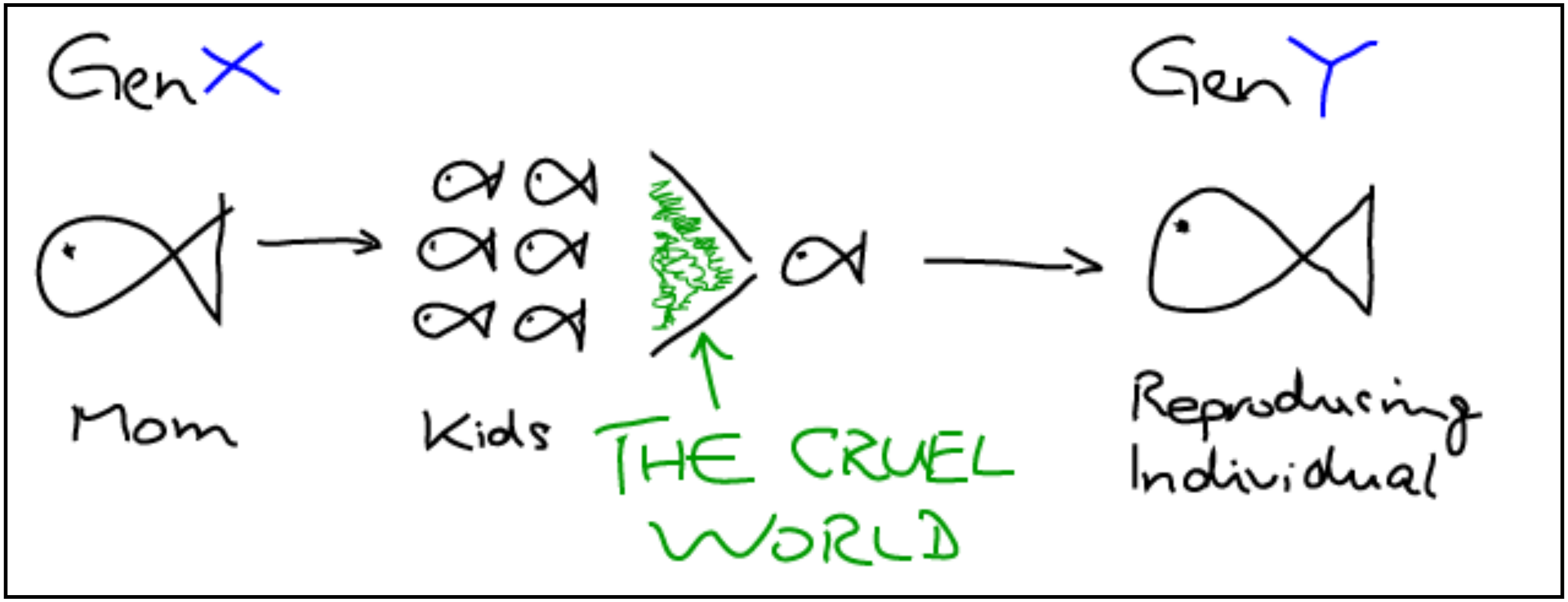


Figure 7.5 Prediction of the Wright-Fisher model for the distribution $\phi(x,t)$ of populations of size $N = 16$ with allele frequency x at generation t, for 20 generations after an initial frequency of 0.5. The values of $\phi(x,t)$ were generated using the Markov transition probability matrix, whose terms are given by the binomial distribution. The model with $2N = 32$ predicts that fewer populations have fixed by generation 19 than actually did go to fixation in the experiment in Figure 7.4. This is because the effective population size is smaller than the observed count (see Figure 7.12).

Principles of Population Genetics, Hartl and Clark

Selection 1: Fitness



- viability = chance of survival to reproductive age
  - **one measure of fitness**
- If fitness depends on genotype, then we have selection
  - **if organisms live/die independent of genotype, that's drift**

Source: A. Sidow, BIOSCI 203

Effective population size N_e

- Sewall Wright (1931, 1938)
- “The number of breeding individuals in an idealized population that would show the same amount of dispersion of allele frequencies under random genetic drift or the same amount of inbreeding as the population under consideration".
- Usually, $N_e < N$ (absolute population size)
- N_e != N can be due to:
  - □ **fluctuations in population size**
  - □ **unequal numbers of males/females**
  - □ **skewed distributions in family size**
  - □ **age structure in population**

Selection vs Drift 1: |s| and Pop Size

If $|s| < 1/N_e$,

then selection is ineffective and the alleles are solely subject to drift: the alleles are "effectively neutral"

What is the probability of fixation?

If $|s| < 1/N_e$, then $P(\text{fix}) = q$

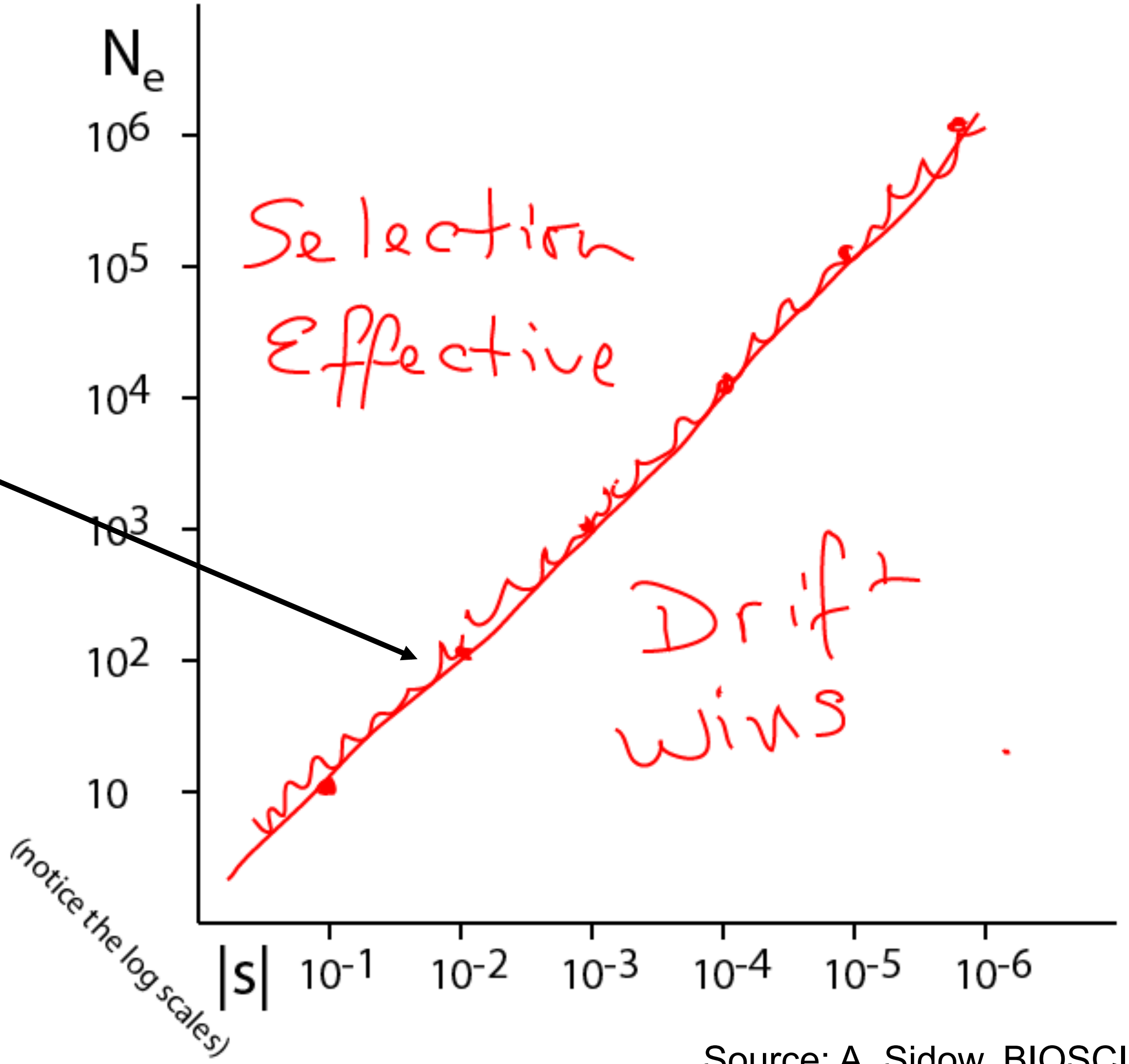
If $|s| > 1/N_e$, then $P(\text{fix}) = \dfrac{1 - e^{-4 N_e s q}}{1 - e^{-4 N_e s}}$

N_e = effective pop size
s = selection coefficient
q = allele frequency

Source: A. Sidow, BIOSCI 203

Selection vs Drift 2: |s| and Pop Size

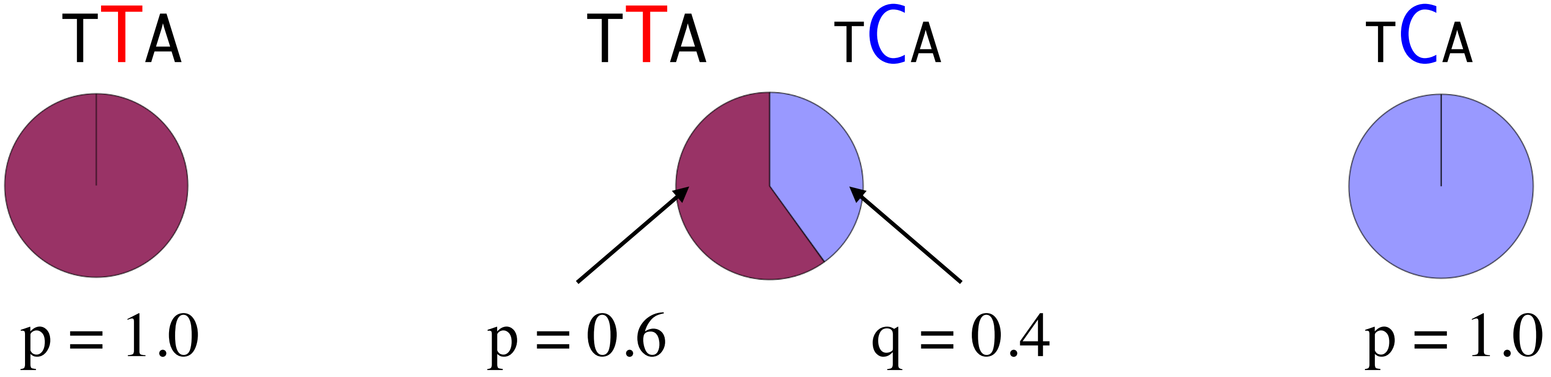
Around the diagonal, where inverse of pop size is close to |s|, selection and drift are in a tug-of-war.

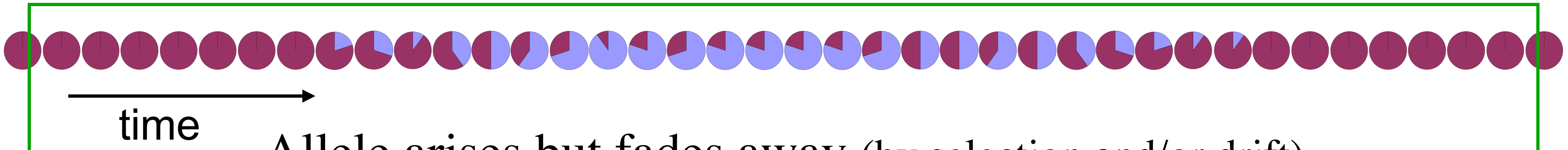


Source: A. Sidow, BIOSCI 203

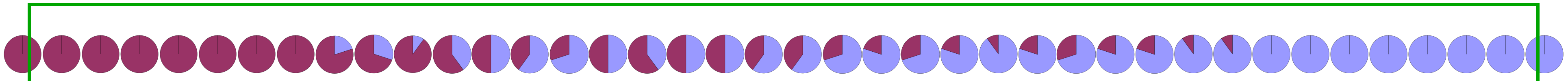
Evolutionary Change (fixation)

Let’s look at a single nucleotide site in the genome





Allele arises but fades away (by selection and/or drift)



Allele arises and moves to fixation (by selection and/or drift)

Source: A. Sidow, BIOSCI 203

Neutral theory (Kimura)

- How do mutation & drift interact, in absence of selection?

- Probability of eventual fixation (of a neutral allele at frequency p_0)
  - □ $\mathbf{p_0}$
  - □ **E.g., for a new mutation in diploid pop: $p_0=1/2N_e$**
- Average time to fixation of a neutral allele
  - □ **$4N_e$ generations**
- Rate at which neutral mutations are fixed (mutation rate is μ)
  - □ μ **(does not involve N_e)**
- Average time between consecutive neutral substitutions
  - □ **$1/\mu$**
- Average homozygosity at equilibrium, using infinite alleles model
  - □ **$1/(4N_e\mu+1)$**

Detection of Selection in Humans with SNPs

Large-scale SNP-survey looked at:

106 Genes in an average of 57 human individuals

60,410 base pairs of noncoding sequence (UTRs, introns, some promoters)

135,823 base pairs of coding sequence

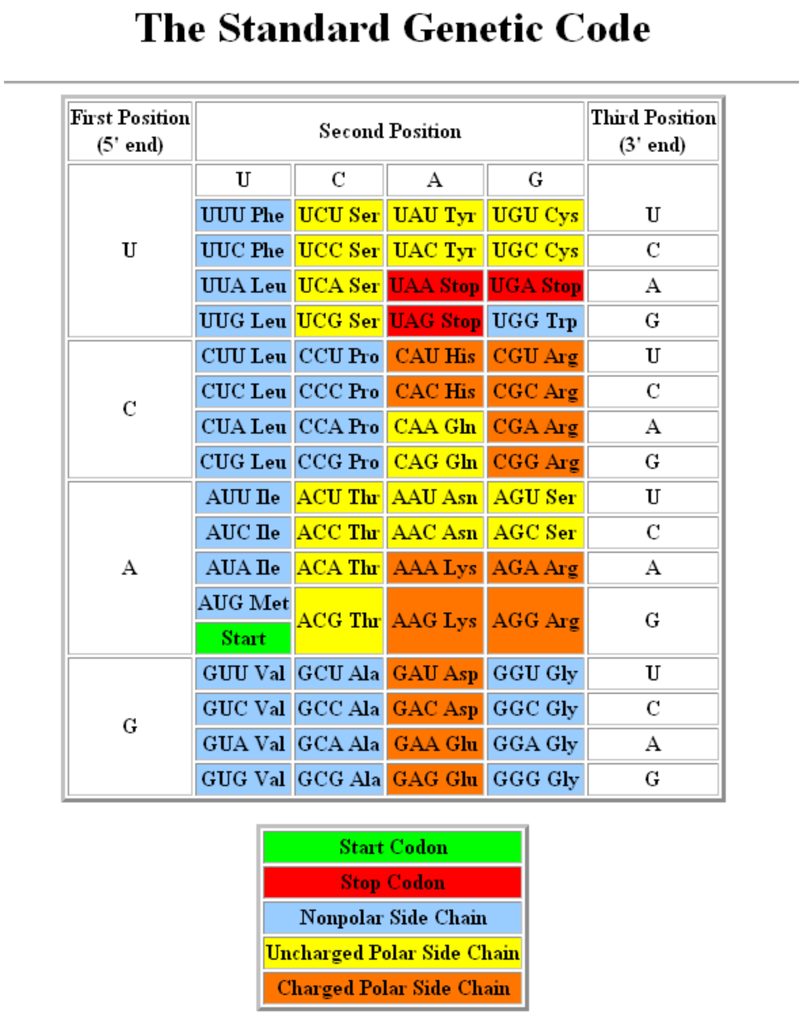
Some salient points:

- Because survey is snapshot of *current* frequencies, evidence for selection or drift is indirect
- This is about bulk properties, not about individual genes

We will discuss only polymorphisms in coding sequence (cSNPs)

Data from Cargill et al, Nature Genetics 1999 vol 22:231

The Degenerate Genetic Code



The Standard Genetic Code

First Position (5' end)	Second Position				Third Position (3' end)
	U	C	A	G	
U	UUU Phe	UCU Ser	UAU Tyr	UGU Cys	U
	UUC Phe	UCC Ser	UAC Tyr	UGC Cys	C
	UUA Leu	UCA Ser	UAA Stop	UGA Stop	A
	UUG Leu	UCG Ser	UAG Stop	UGG Trp	G
C	CUU Leu	CCU Pro	CAU His	CGU Arg	U
	CUC Leu	CCC Pro	CAC His	CGC Arg	C
	CUA Leu	CCA Pro	CAA Gln	CGA Arg	A
	CUG Leu	CCG Pro	CAG Gln	CGG Arg	G
A	AUU Ile	ACU Thr	AAU Asn	AGU Ser	U
	AUC Ile	ACC Thr	AAC Asn	AGC Ser	C
	AUA Ile	ACA Thr	AAA Lys	AGA Arg	A
	AUG Met Start	ACG Thr	AAG Lys	AGG Arg	G
G	GUU Val	GCU Ala	GAU Asp	GGU Gly	U
	GUC Val	GCC Ala	GAC Asp	GGC Gly	C
	GUA Val	GCA Ala	GAA Glu	GGA Gly	A
	GUG Val	GCG Ala	GAG Glu	GGG Gly	G

Start Codon

Stop Codon

Nonpolar Side Chain

Uncharged Polar Side Chain

Charged Polar Side Chain

Null Hypothesis for SNP Survey

•In the average coding region, about 30% of possible point muts are **silent**

•*Silent* substitutions – don' t change the aa

•*Replacement* substitutions – do change the aa

•*conservative* substitutions – a functionally similar aa

•*nonconservative* substitutions – a functionally different aa

If there had been no selection in population history, we would expect

70% of coding region polymorphisms to be replacement and

30% to be silent

But consider:

1. Silent changes usually produce no phenotype and are therefore unlikely to be subject to selection -- neutral assumption holds
2. Replacement changes can produce a phenotype, if only subtle or in synthetic combination -- neutral assumption may not hold
3. Far more replacement changes are deleterious than advantageous

Source: A. Sidow, BIOSCI 203

Results of SNP Survey

1. *Silent polymorphisms outnumber replacement polymorphisms*

	Total	Silent	Replacement
Observed	392	207	185
Expected if no selection	392	118	274

2. *Conservative replacements outnumber nonconservative replacements*

	Total	Conservative	Nonconservative
Observed	185	119	66
Expected if no selection	185	~92	~93

- Implication: selection against deleterious mutations
  - penalizes replacements
  - especially penalizes nonconservative replacements

Source: A. Sidow, BIOSCI 203

Overview

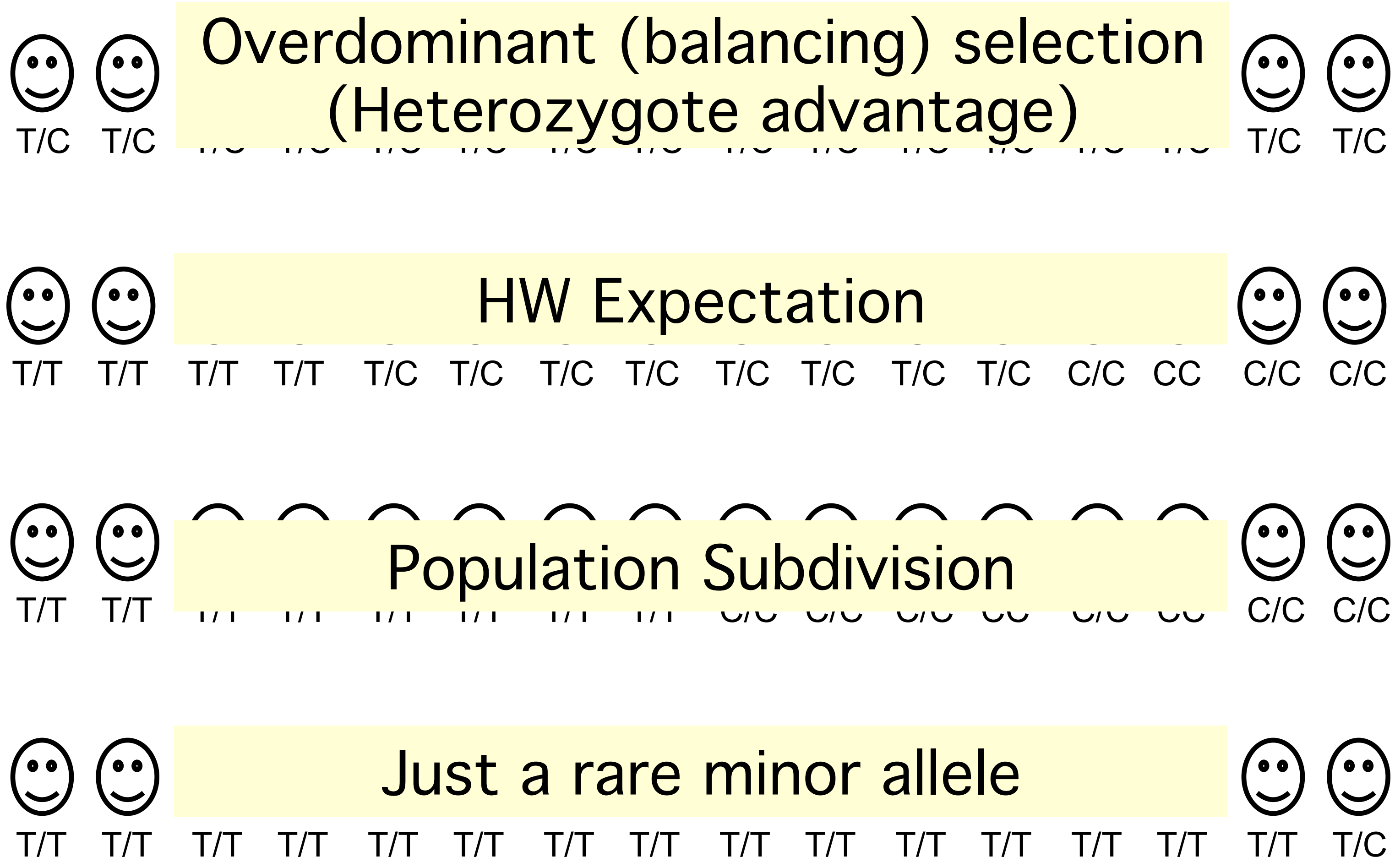
- History
- Some Population Genetics
  - **origins of genetic variation**
  - **evolutionary timescales**
  - **selection and drift**
  - **neutral theory**
- Detection of Selection in Humans with SNPs
- Some more Population Genetics
  - **Migration**
  - **Wright's F_{ST}**
- Inference of Human Phylogenetic Tree
- Time to Most Recent Common Ancestor (TMRCA)
- Unique Origin vs. Multiregional Evolution Models
- Geographic Origin of Humans

Migration: another source of allele frequency change

- In a subdivided population, drift and varied selection result in diversity among subpopulations
- Migration limits genetic divergence
  - **Lack of migration can allow speciation to occur**
- Only 1 migrant per generation is enough to keep drift partially in check (prevent complete fixation of alleles) !

Allele frequencies and population history

What are the allele frequencies vs. heterozygosities?



Pop1

Overdominant (balancing) selection
(Heterozygote advantage)

T/C T/C … T/C T/C

Pop2

HW Expectation

T/T T/T T/T T/T T/C T/C T/C T/C T/C T/C T/C T/C C/C CC C/C C/C

Pop3

Population Subdivision

T/T T/T … C/C C/C

Pop4

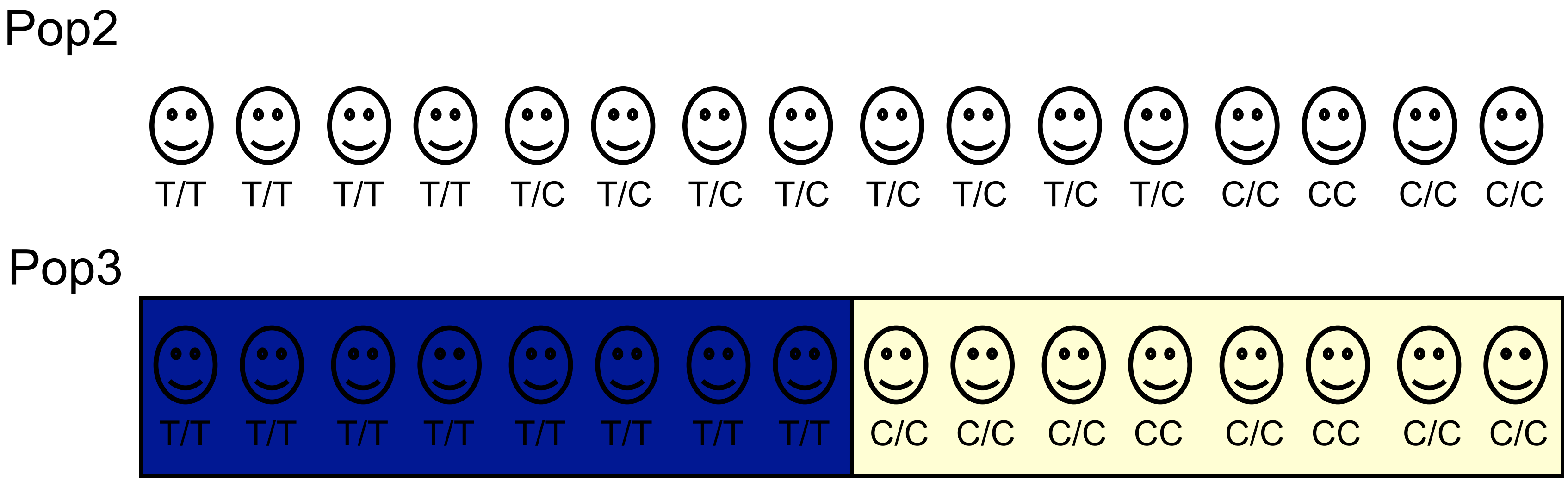
Just a rare minor allele

T/T T/T T/T T/T T/T T/T T/T T/T T/T T/T T/T T/T T/T T/T T/T T/C

Source: A. Sidow, BIOSCI 203

Population Subdivision

Wright's F-statistics (F_{ST}, etc) are measures of genetic diversity
Indicates population subdivision



Pop2

T/T T/T T/T T/T T/C T/C T/C T/C T/C T/C T/C T/C C/C CC C/C C/C

Pop3

T/T T/T T/T T/T T/T T/T T/T T/T C/C C/C C/C CC C/C CC C/C C/C

Maybe: Pop3a (Oahu) Pop3b (Kauai)

F_{ST} measures a reduction of average heterozygosity between the subpopulations and the total population.

Source: A. Sidow, BIOSCI 203

Wright's F_{ST}: a measure of genetic diversity among populations

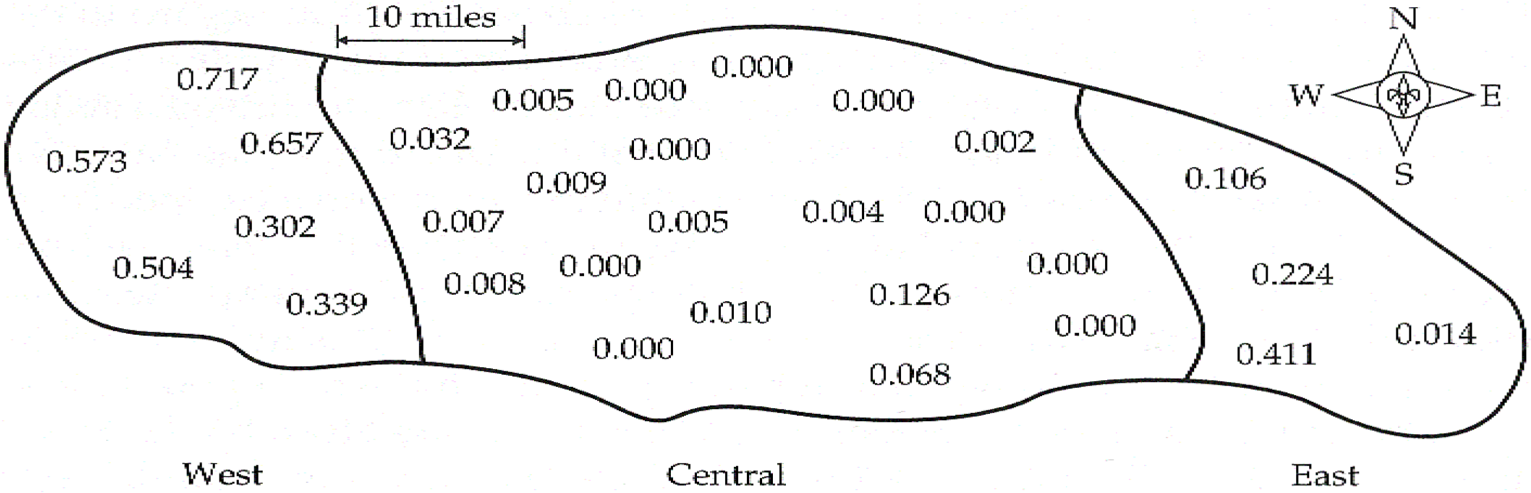


Figure 4.2 Estimated frequency of a recessive allele for blue flower color in populations of *Linanthus parryae* in an area of approximately 900 square miles in the Mohave desert. Each allele frequency is based on an examination of approximately 4000 plants over an area of about 30 square miles. (After Wright 1943a.)

Wright's F_{ST}: a measure of genetic diversity among populations

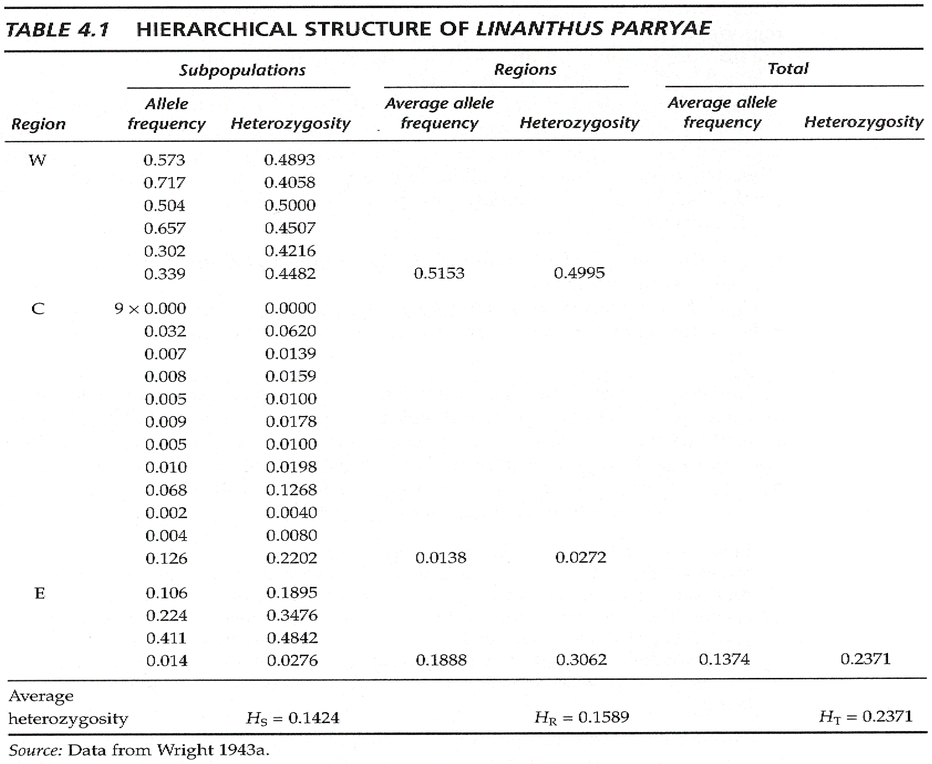


TABLE 4.1 HIERARCHICAL STRUCTURE OF *LINANTHUS PARRYAE*

	Subpopulations		*Regions*		*Total*	
Region	*Allele frequency*	*Heterozygosity*	*Average allele frequency*	*Heterozygosity*	*Average allele frequency*	*Heterozygosity*
W	0.573	0.4893				
	0.717	0.4058				
	0.504	0.5000				
	0.657	0.4507				
	0.302	0.4216				
	0.339	0.4482	0.5153	0.4995		
C	9 × 0.000	0.0000				
	0.032	0.0620				
	0.007	0.0139				
	0.008	0.0159				
	0.005	0.0100				
	0.009	0.0178				
	0.005	0.0100				
	0.010	0.0198				
	0.068	0.1268				
	0.002	0.0040				
	0.004	0.0080				
	0.126	0.2202	0.0138	0.0272		
E	0.106	0.1895				
	0.224	0.3476				
	0.411	0.4842				
	0.014	0.0276	0.1888	0.3062	0.1374	0.2371
Average heterozygosity		$H_S = 0.1424$		$H_R = 0.1589$		$H_T = 0.2371$

Source: Data from Wright 1943a.

"Decrease of heterozygosity"

$F_{ST} = (H_T - H_S)/H_T$

(0.2371 – 0.1424)/0.2371 = **0.3993**

(indicates high overall diversity of subpopulations)

- **0 – 0.05: little genetic differentiation**
- **0.05-0.15: moderate**
- **0.15-0.25: great**
- **> 0.25: very great**

$F_{SR} = (H_R - H_S)/H_R$

(0.1589 – 0.1424)/0.1589 = **0.1036**

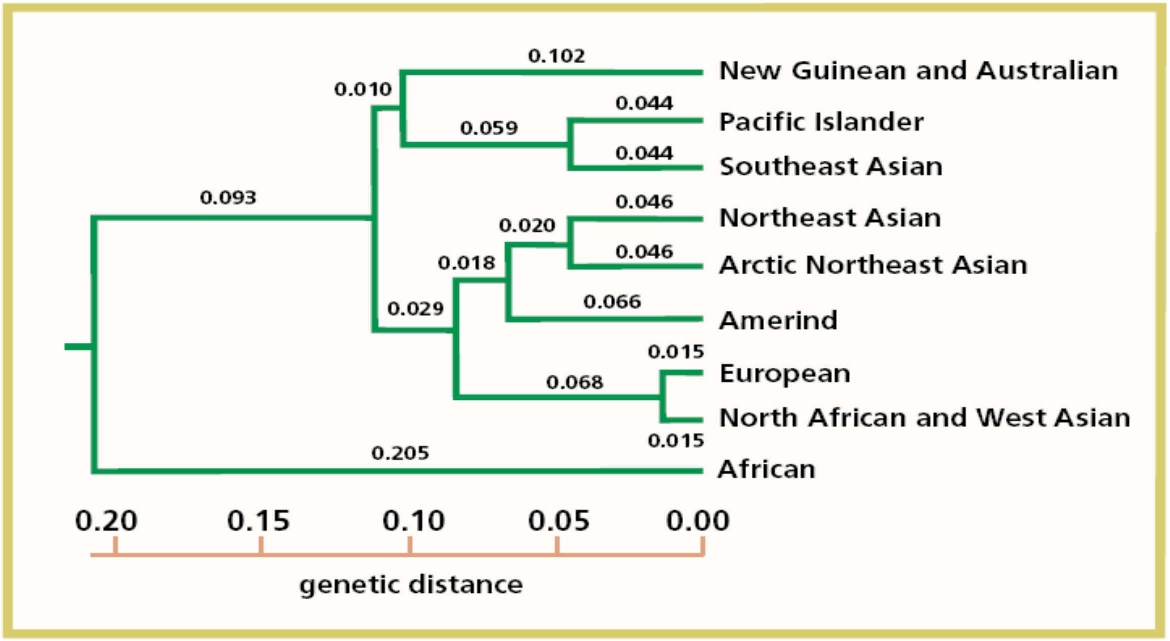
Variation among subpops within each region

$F_{RT} = (H_T - H_R)/H_T$

(0.2371 – 0.1589)/0.2371 = **0.3299**

Variation among regions within total pop (greater than variation within regions – regions capture population structure)

Inference of Human Phylogenetic Tree



BOB CRIMI

Fig. 1 Summary tree of world populations. Phylogenetic tree based on polymorphisms of 120 protein genes in 1,915 populations grouped by continental sub-areas and F_{st} genetic distances[14]. Root placed assuming a constant rate of evolution.

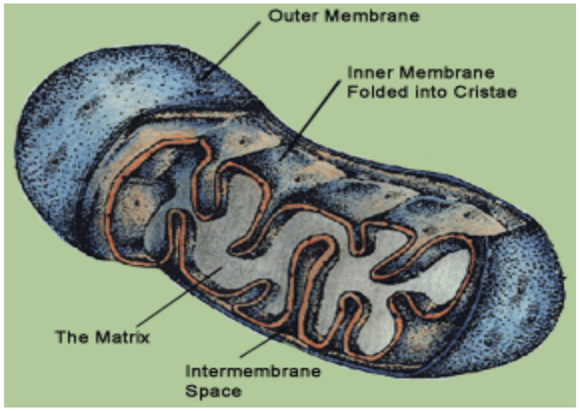
nature genetics supplement • volume 33 • march 2003

Time to Most Recent Common Ancestor (TMRCA)

- Archeological evidence
  - **origin in Africa 50-100kya**
  - **spread to rest of world, 50-60kya**
- What does genetic evidence say?
- What about the location?

Mitochondrial DNA

- An organelle of the animal cell
- Kreb's Cycle (aerobic respiration) takes place here
- Transmitted only along female lineage
- Haploid genome, independent from human "host"
- High mutation rate



Mitochondrial “Eve”

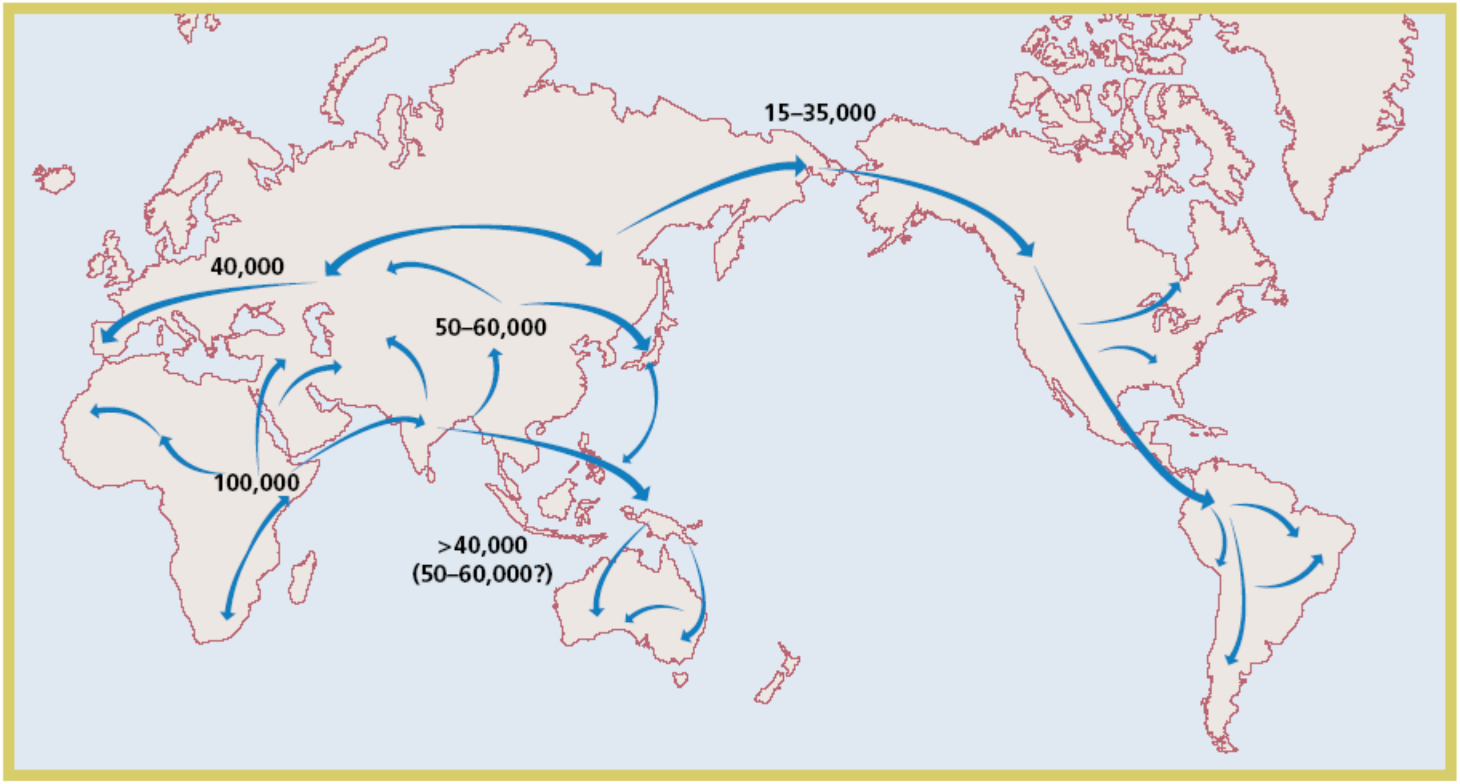
- Most recent *matrilineal* common ancestor of all living humans
- All our mitochondria are descended from hers
- Does *not* mean she was the only human female alive at the time
  - **Consider the set S of all humans alive today**
  - **Take the set S’ = *mothers-of*(S). (now all female)**
  - **Size(S’) ≤ Size (S)**
  - **…continue until you have one member: that’s Eve**
- Members of S have other female ancestors, but Eve is the only one with an unbroken matrilineal line to all of S
- She lived ~230kya
- She was not Eve during her own lifetime
  - **Title of Eve depends on current set of people alive**
  - **as matrilineal lines die out, you get a more recent Eve**
- Difficult to determine if she was *Homo sapiens*

Y-chromosome “Adam”

- Part of the Y chromosome does not recombine
- Hence we can do a similar trick
  - **However, only men (XY) carry the Y chromosome**
  - **So we can only identify the most recent patrilineal common ancestor of all *men* living today:**
- Estimated to live ~100kya
  - **never met “Eve”!**

- Why are mtDNA and Y chromosome TMRCA dates so different?
  - **lower N_E for males than for females?**
    - polygyny more frequent than polyandry?
    - higher male mortality rates?
    - higher male variability in reproductive success?
  - **patrilocal marriage more common than matrilocal?**
  - **mtDNA mutation rates variable, causing error?**

Tracking Human Migrations



BOB CRIMI

Fig. 3 The migration of modern *Homo sapiens*. The scheme outlined above begins with a radiation from East Africa to the rest of Africa about 100 kya and is followed by an expansion from the same area to Asia, probably by two routes, southern and northern between 60 and 40 kya. Oceania, Europe and America were settled from Asia in that order.

Current consensus: ~1,000 individuals (a tribe) left Africa 1(

nature genetics supplement • volume 33 • march 2003

Human microsatellite data

- 1052 individuals; 52 populations; 377 autosomal microsatellite markers

“microsatellite” or Short Tandem Repeat (STR) = 2-6 bases repeated several times

e.g., TCTA TCTA TCTA TCTA TCTA TCTA TCTA TCTA

- “indigenous populations” only; all individuals’ grandparents lived in same place

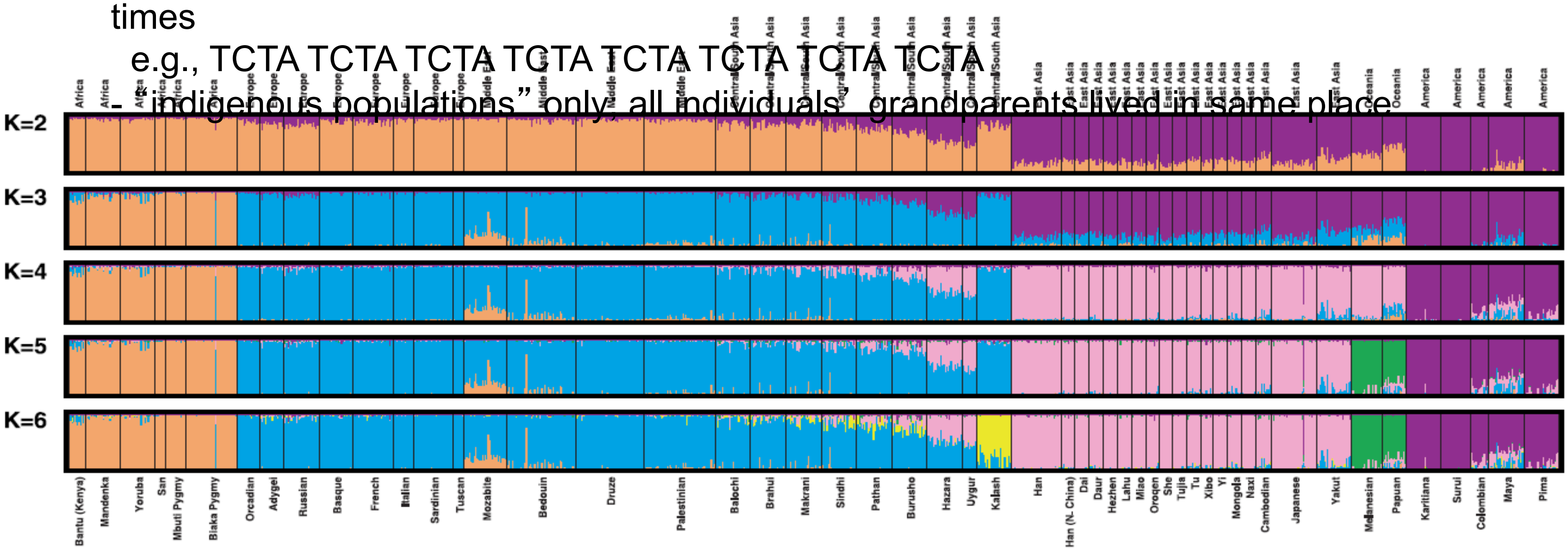


Fig. 1. Estimated population structure. Each individual is represented by a thin vertical line, which is partitioned into *K* colored segments that represent the individual’s estimated membership fractions in *K* clusters. Black lines separate individuals of different populations. Populations are labeled below the figure, with their regional affiliations above it. Ten *structure* runs at each *K* produced nearly identical individual membership coefficients, having pairwise similarity coefficients above 0.97, with the exceptions of comparisons involving four runs at *K* = 3 that separated East Asia instead of Eurasia, and one run at *K* = 6 that separated Karitiana instead of Kalash. The figure shown for a given *K* is based on the highest probability run at that *K*.

Rosenberg *et al., Science* **298**:2381-2385.

F_{ST} versus distance in Humans

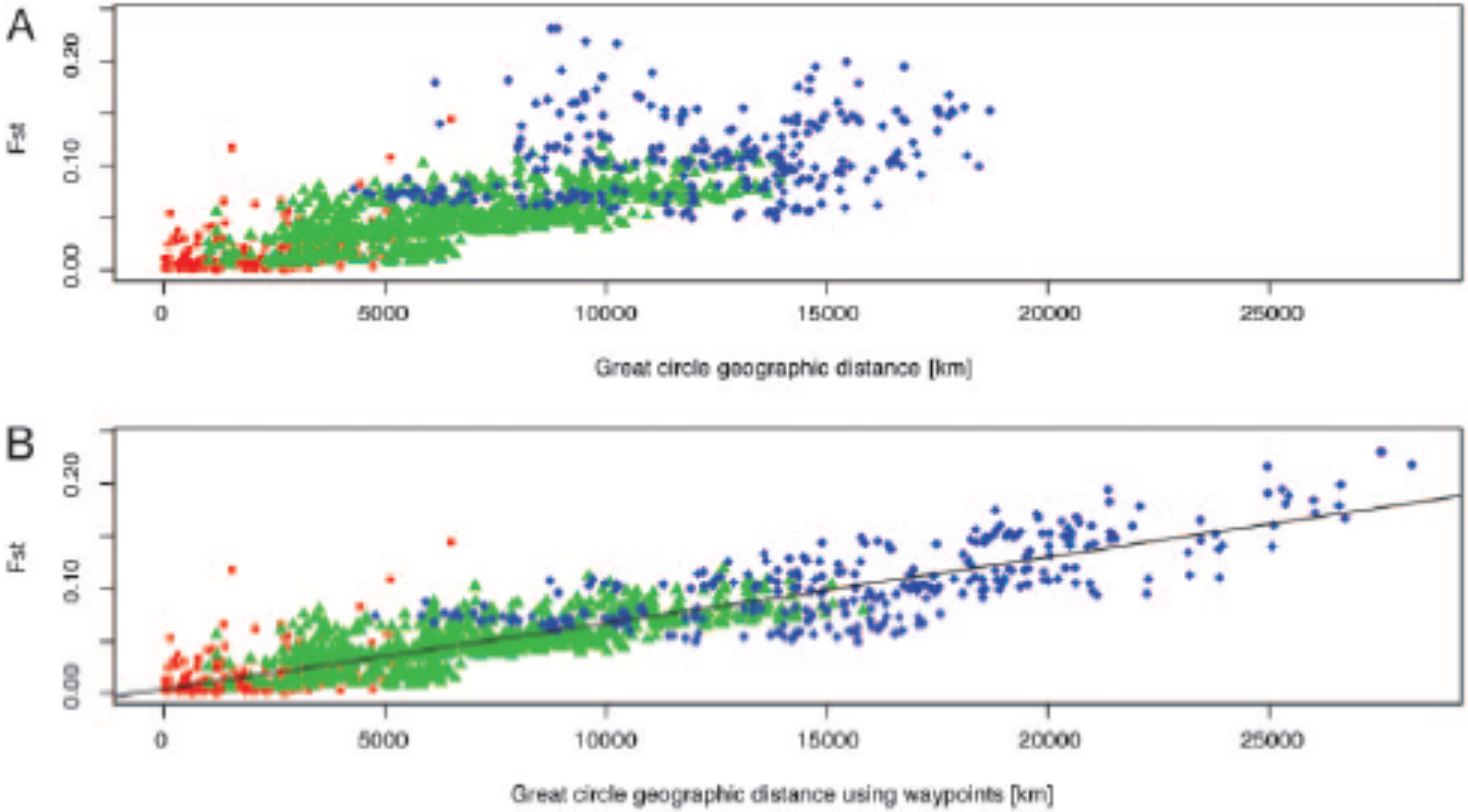


Fig. 1. Scatterplot of F_{ST} and geographic distance. Red dots denote within-region comparisons, green triangles indicate comparisons between populations in Africa and Eurasia, and blue diamonds represent comparisons with America and Oceania. (*A*) The relationship between F_{ST} and geographic distance computed using great circle distances. R^2 for the linear regression of genetic distance on geographic distance is 0.5882. (*B*) The correction for large bodies of water produces a different scatterplot ($R^2 = 0.7835$). The regression line fitted to the data [$\widehat{F_{ST}} = 4.35 \times 10^{-3} + (6.28 \times 10^{-6}) \times$ (geographic distance in kilometers)] is drawn in black.

Ramachandran et al.,*PNAS* **102**(44)

Geographic Origin of Humans

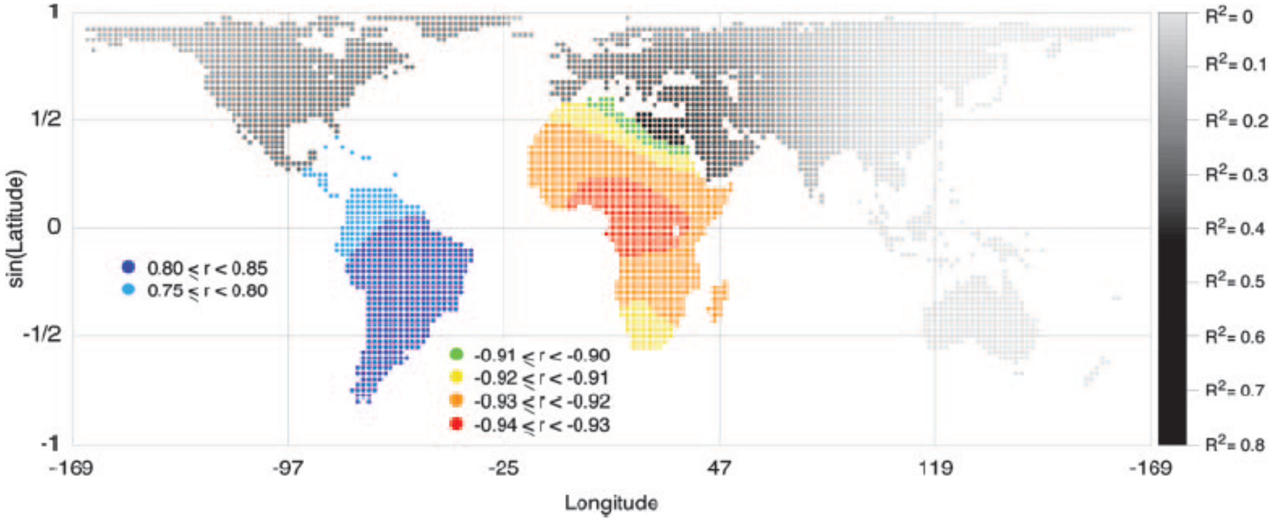


Fig. 5. The origin of the human expansion. The color or shade of each of the 4,210 locations (shown as dots) indicates either a correlation coefficient r or an R^2 value for the regression of expected heterozygosities in 53 HGDP-CEPH populations on geographic distance (corrected for large bodies of water) to the location displayed. Note that, for a simple linear regression, $r^2 = R^2$. Grayscale points indicate R^2 values, as shown by the gradient on the right, and correlation coefficients r are displayed in Africa and South America to reflect the sign of the relationship between heterozygosity and geographic distance to locations in these continents. R^2 values range from 0.757 to 0.870 in Africa and from 0.519 to 0.659 in South America. The maximum value of r ($\approx$0.812) is observed when the origin is (30S, 50.2W); the minimum value of r (approximately $-$0.933) is observed when the origin is (4.3N, 12.8E).

- Pairwise distances used to imply an ordering
- Assumes a single origin
- Assumes no intermingling after a colony founded
- A central African origin has the most explanatory power

Ramachandran et al.,*PNAS* **102**(44)

Break!

Theories of Modern Human Origins

- Two major theories attempt to explain the latter phases of human evolution and the development of modern human population variation (human ’races’)
- They view human origins very differently, with the differences based primarily on how isolated hominid populations were after spreading out from Africa around 1.8myr.
- Both theories have long histories, and in one guise or another, have been around since the recognition of the essential non-modern human qualities of the neandertals in the middle of the 19th century

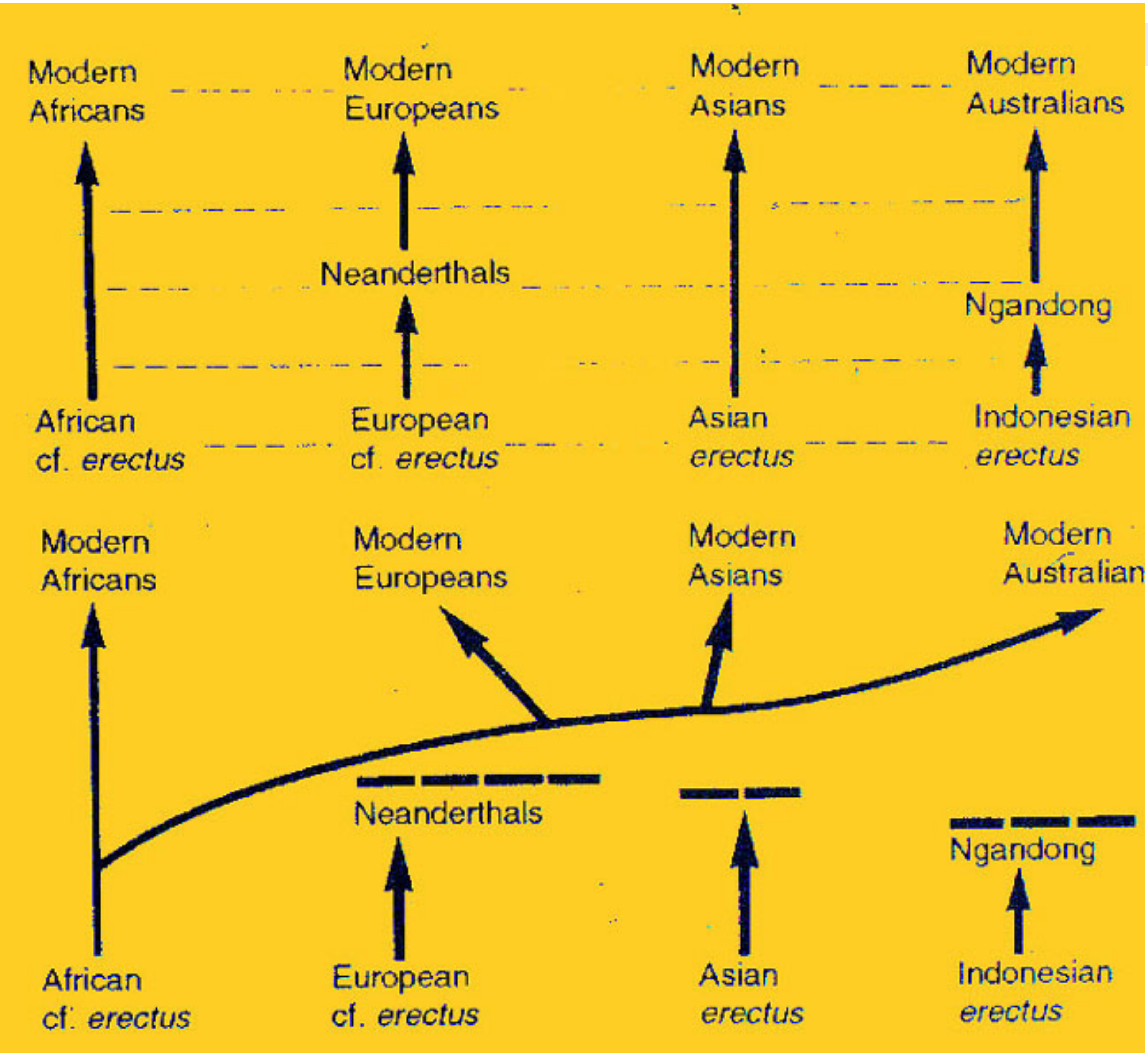
Competing Models of Human Origins

The two competing models are known as:

1. The Multi Regional Evolutionary Model.

2. The Single Origins Model (usually called “Out of Africa”).

theori


Modern Africans
Modern Europeans
Modern Asians
Modern Australians
Neanderthals
Ngandong
African cf. erectus
European cf. erectus
Asian erectus
Indonesian erectus
Modern Africans
Modern Europeans
Modern Asians
Modern Australian
Neanderthals
Ngandong
African cf. erectus
European cf. erectus
Asian erectus
Indonesian erectus

Multi Regional Evolution I

- With expansion of early *Homo* into Eurasia, hominid populations moved into new environments and began to evolve biological features for life in those places.
- In this model, hominid populations were continuously distributed over the continents, and were in more or less constant contact with other populations, thus sharing genes.
- This gene flow insured that the hominids remained one evolving species.
- By about 700,-400,00 years ago, archaic members of *H. sapiens* had appeared.

Multi Regional Evolution II

- These archaic *H. sapiens* populations in the different areas eventually evolve into living human regional populations (“races”).
- Thus, human races have a long antiquity in their local environments, having evolved from earlier archaic sapiens, and before that, from the local early *Homo* populations.
- Multi regional evolution stresses the ebb and flow of gene flow as a crucial factor in human evolution and in modern human origins.

Single Origins Theory I

- Begins in the same fashion as multi regional evolution with the spread of early *Homo* out of Africa into Eurasia. Hominid populations move into new environments and begin to evolve biological features for life in those places.
- In this theory, hominids lived in small, isolated populations and, lacking genetic contact, evolved into a number of new species.
- In Europe, this new species will eventually evolve into the neandertals, who become extinct toward the end of human evolution.

Single Origins Theory II

While in Europe these now isolated hominids evolve into a new species, the Neandertals, In Africa and Asia, other species of *Homo* were also evolving. Like the Neandertals in Europe, they also possess low sloping brain cases, and large projecting faces lacking a chin. They had large brains, often within the range of living humans.

Single Origins Theory III

- Between about 200,-100,000 years ago, modern humans, *Homo sapiens*, evolved from an earlier *Homo* ancestor.
- This evolutionary origin apparently took place in one locale, most probably somewhere in sub- Saharan Africa.
- Soon after this origin, these modern humans begin to expand out of Africa, marking a **second** expansion out of Africa.
- These modern humans move into all parts of the Old World, replacing earlier species of *Homo,* like the Neandertals, in those areas.

Single Origins Theory IV

- Thus, in this theory, modern humans, *Homo sapiens,* evolve relatively recently in one locale and spread out from there.
- Modern human races **all** have a relatively recent origin in Africa.
- Earlier humans in other parts of the Old World were separate species from modern humans. They were not part of the ancestry of modern humans but an extinct side branch, replaced by these newcomers who moved ‘out of Africa’.

Modern Human Origins

- Thus, two different theories:

  1) Multi Regional Evolution

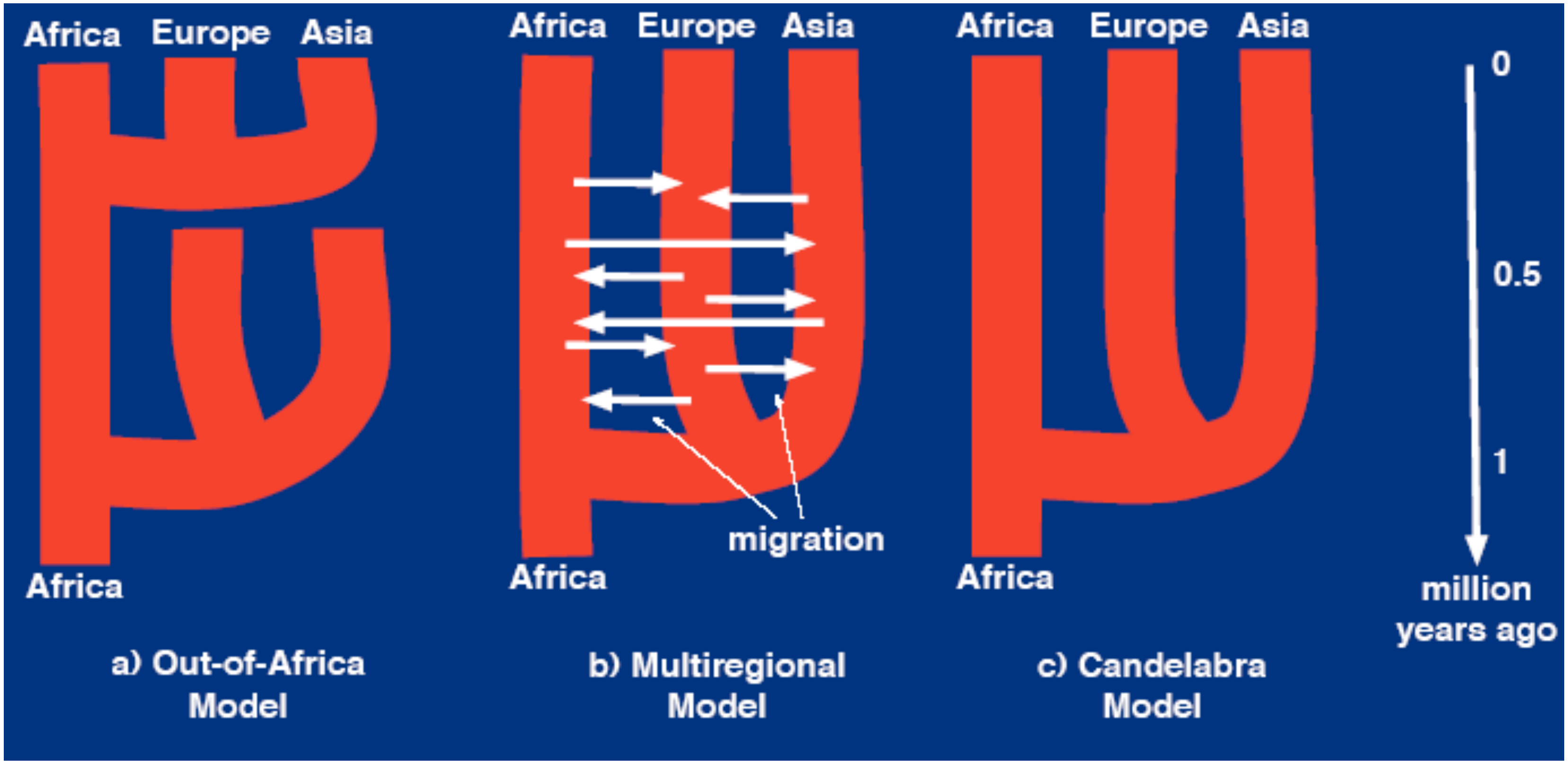
  2) Single Origins : “Out of Africa”

- Because they are amongst the most numerous of fossils, much of the emphasis of both theories centers on the Neandertals.

Single Origins Theory: Genetic Evidence

- At the moment, this is the strongest evidence for a recent origin of modern humans in Africa.
- It is based on the analysis of DNA, but not primarily the DNA found on the chromosomes in the neucleus. Other genetic material is found in structures called **mitochondria** (known as mtDNA).
- Mitochondria (singular: mitochondrion) are cell structures responsible for carrying out the conversion of the sugar glucose into a form usable to the cell for energy.

Models of modern human origins



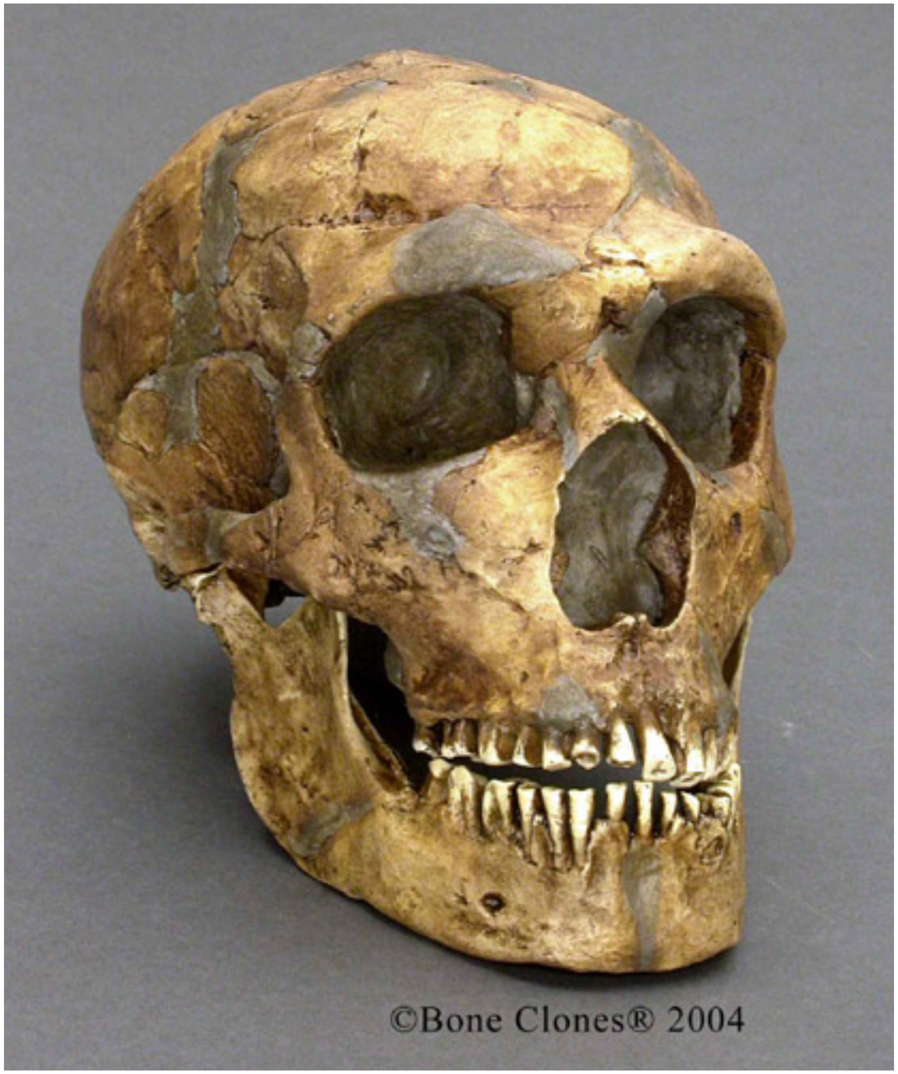
Who were the Neanderthals?

- The Neanderthals were a group of people that lived in Europe from 30,000 to 150,000 years ago.

- We have numerous stone tools and skeletal remains from Neanderthals.

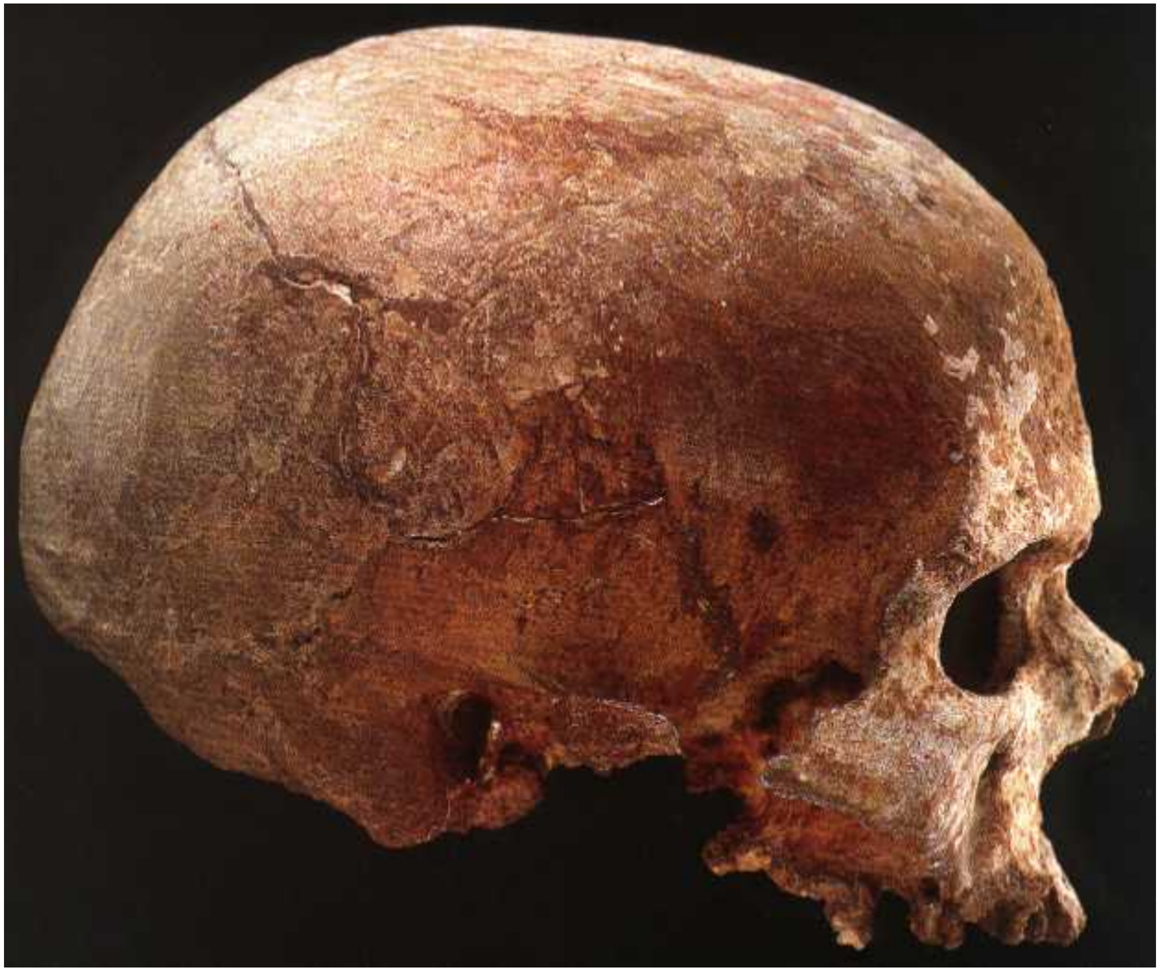
- Around 30 – 40 thousand years ago we stop seeing Neanderthal fossils and start seeing fossils that look more "modern".

Neanderthal skull



©Bone Clones® 2004

Modern human skull



Neanderthal questions

- Did the Neanderthals evolve into modern humans or did the Neanderthals die out and get replaced by modern humans?

- Where did the ancestors of modern Europeans live 50,000 years ago?

Neanderthal questions

- Did the Neanderthals evolve into modern humans or did the Neanderthals die out and get replaced by modern humans?

- Where did the ancestors of modern Europeans live 50,000 years ago?

- Another way of phrasing this question is: Did Neanderthals make any contribution to the modern gene pool?

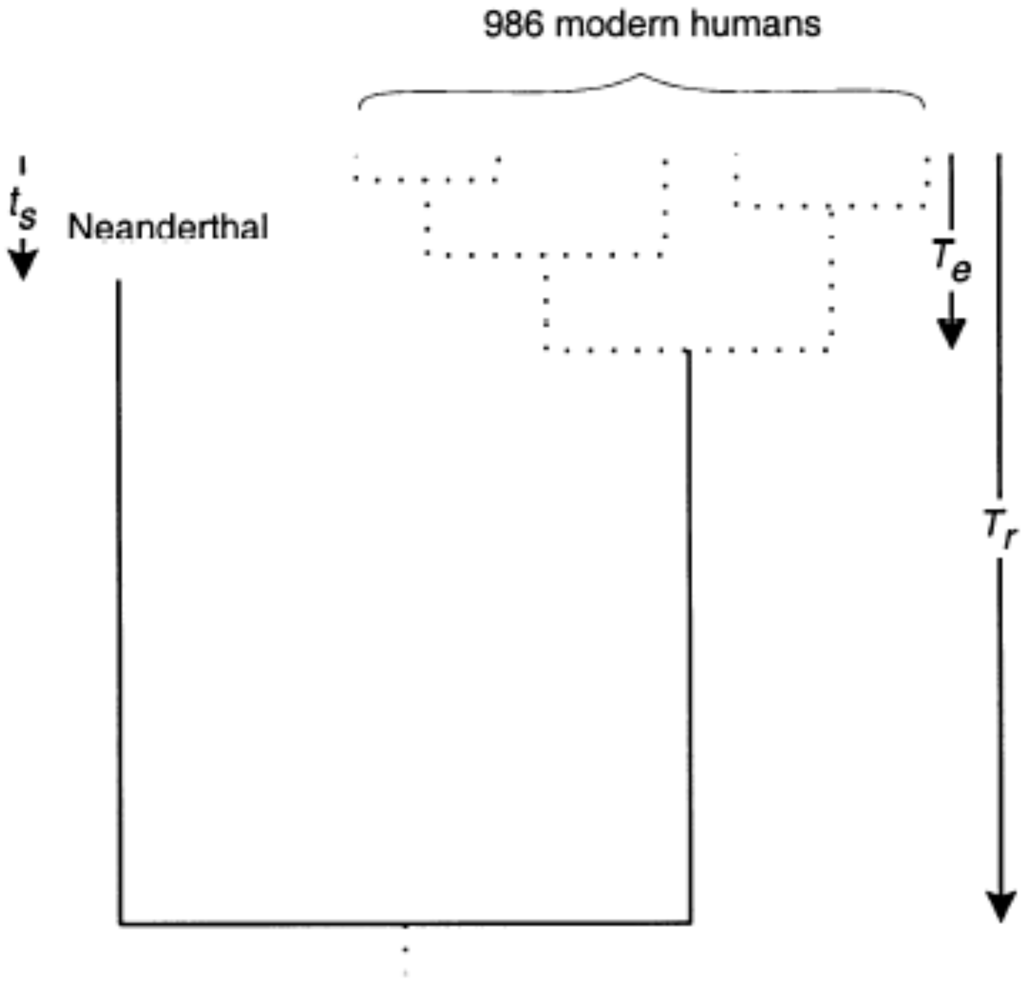
Neanderthal DNA

Recent technological advances have led to the sequencing of some Neanderthal mtDNA.

It is only possible to extract DNA from remains that are < 50,000 years old. (The older the fossil, the less likely that any of the DNA remains.)

The DNA in Neanderthal bones is highly degraded and very hard to sequence.

Neanderthal mtDNA



mtDNA has been sequenced from 5 different Neanderthals and over 1000 modern humans. The Neanderthal mtDNA looks to be substantially different.

Nordborg
1998

What can we conclude?

Neanderthals (almost certainly) did not contribute to the modern mtDNA gene pool.

What can we conclude?

Neanderthals (almost certainly) did not contribute to the modern mtDNA gene pool. This could happen because:

- Neanderthals and modern humans did not (or could not) interbreed
- Neanderthals and modern humans did mix but Neanderthal mtDNA was lost by **genetic drift**

mtDNA Results

- Comparisons based on segments of the mtDNA from a number of human populations:

  1) Documents a greater amount of mtDNA variation in Africans in comparison to human populations in other parts of the world.

  2) Discovered unique variations in Africa.

- **Conclusions drawn from this data:**

  1) Modern humans originated in Africa.

  2) There was a subsequent spread to other parts of the Old World, replacing earlier hominid populations.

Debates about mtDNA Results

- Many scientists believe that these results are simplistic and do not reflect the realities of human origins.
- Some suggest that because Africa was an optimal environment for earlier hominids, population size was always larger there than elsewhere; thus there was a greater number of mutations, and more variability.
- Others argue that if there was significant evolutionary selection on the mtDNA genes, then it would be very difficult to predict the nature of this evolution.

Things change....

- Recent sequencing of Neanderthals show that individuals outside of Africa have about 1-4% of their genome from Neanderthal!
  - **A Draft Sequence of the Neandertal Genome. Science 7 May 2010: Vol. 328 no. 5979 pp. 710-722 DOI: 10.1126/science.1188021**