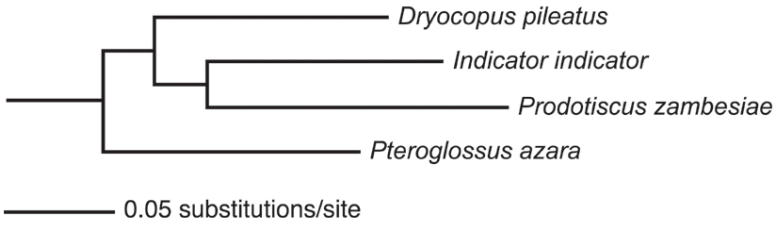
Problems with dating…

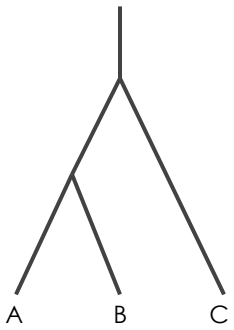
- uncertainty in calibration points
- fossil evidence provides lower bound on age only
- variance of genetic distance estimates
- "saturation" of genetic distances
- extrapolation outside of calibrated range
- ancestral polymorphism
- **variation in substitution rate among lineages**



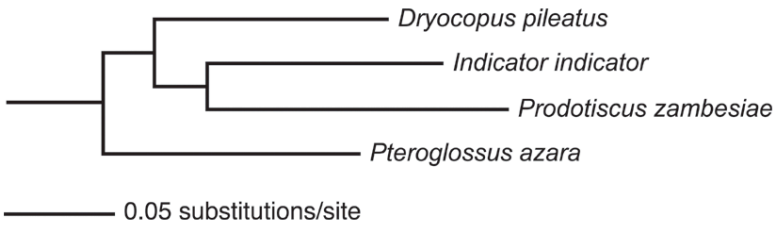
Dryocopus *Indicator* *Prodotiscus* *Pteroglossus*

Relative Rates Test

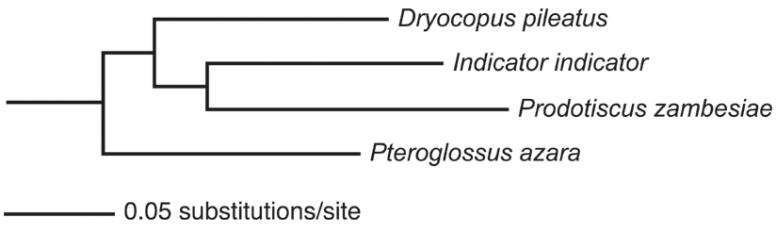
- compares genetic distances between two taxa (A, B) and an outgroup (C)
- if evolutionary rate is constant, distances should be equal
- $d_{AC} = d_{BC}$



one taxon
many taxa
one locus
many loci



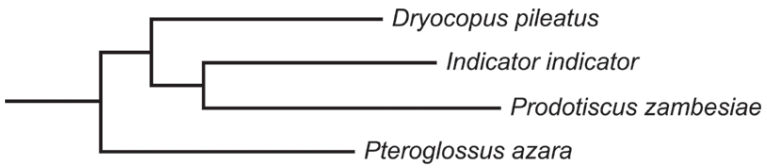
Dryocopus *Indicator* *Prodotiscus* *Pteroglossus*

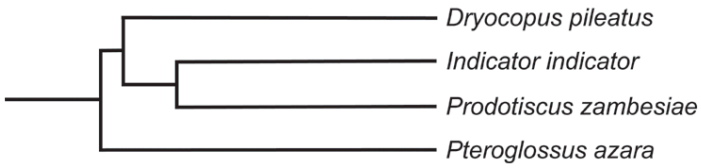


		Differences							
Comparison	Sites	**AG**	**CT**	**AC**	**AT**	**CG**	**GT**	All	TVs
Dryocopus vs *Indicator*	8991	323	754	360	149	61	30	1677	600
Dryocopus vs *Prodotiscus*	8991	322	772	458	157	78	44	1831	737

Likelihood ratio test for rate constancy

- compare the likelihood (probability) of the data when a molecular clock is enforced versus the likelihood when all branches are free to vary in length (product of time and mutation rate)

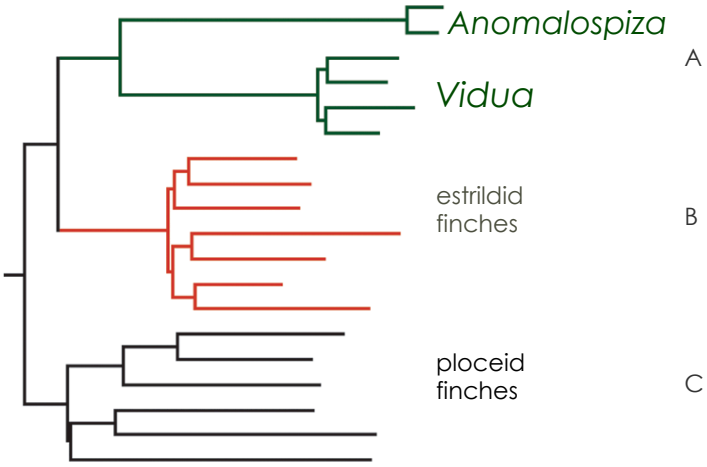




No clock: -ln L = 27859.36
Clock: -ln L = 27904.29
2 x ΔlnL = 89.86, $p < 0.0001$

- test statistic: 2 * Δ ln *L* is distributed approximately as X^2 (chi-square) with *n*-2 degrees of freedom, where *n* = number of terminal taxa
  - unconstrained tree: 2*n*-3 branch lengths
  - constrained tree: *n*-1 branch lengths

Relative Rates Test w/ Multiple Taxa



if evolutionary rates are constant, $d_{AC} = d_{BC}$

d_{AC} = *0.0925**; d_{BC} = *0.0738**

*calculated using the method of Steel et al. 1996 Syst. Biol.

$$d_{AC} = T_2 r_1 + (T_2 - T_1) r_1 + T_1 r_2 = 0.0925$$

$$d_{BC} = 2 T_2 r_1 = 0.0738$$

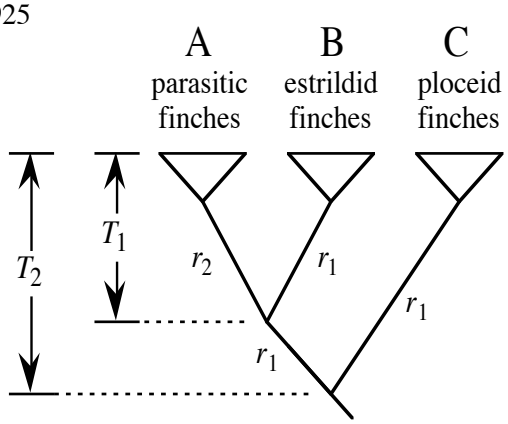
$$d_{AB} = T_1 (r_1 + r_2) = 0.0798$$

setting $r_1 = 1$,

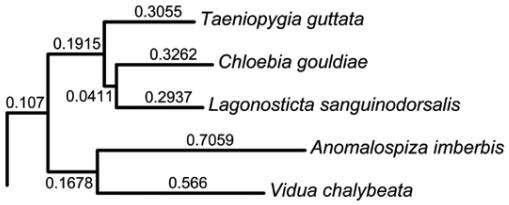
$$T_2 = d_{BC} / 2 r_1 = 0.0369$$

$$T_1 = \frac{d_{AB} + d_{BC} - d_{AC}}{2 r_1} = 0.0306$$

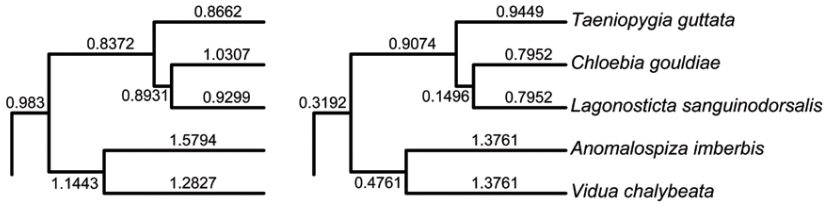
$$r_2 = \frac{d_{AB} + d_{AC} - d_{BC}}{2 T_1} = 1.61$$



MrBayes: branch lengths (product of time and rate)



BEAST: separate estimates of rate and time



Variation in Evolutionary Rate

- rates may vary among lineages due to…
  - differences in life history
    - especially generation time, metabolic rate
  - diversifying natural selection
    - but likely limited to few sites in few genes
  - population history
    - the rate of neutral evolution does not depend on population size
    - the rate of nearly neutral evolution does!

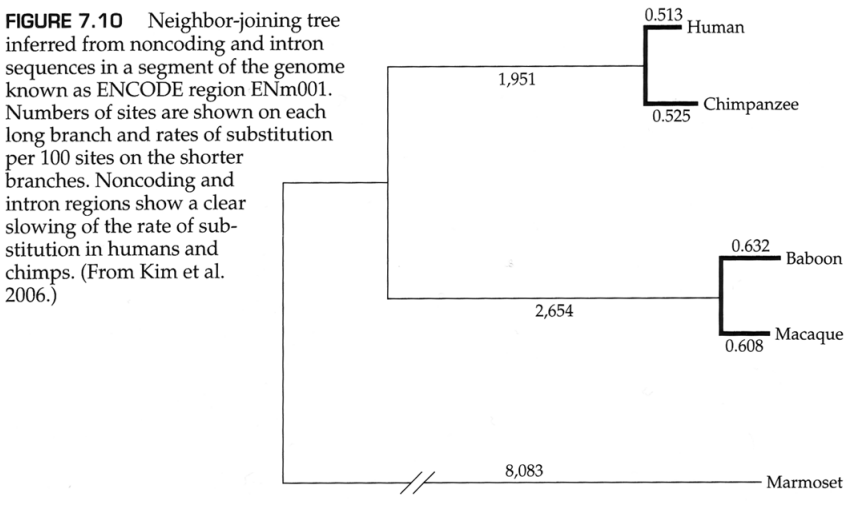


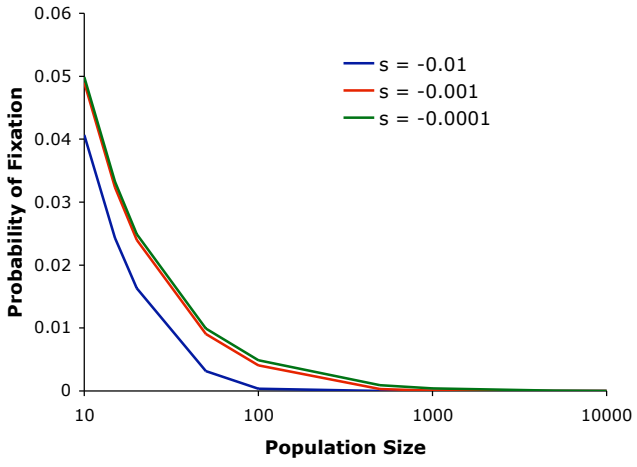
FIGURE 7.10 Neighbor-joining tree inferred from noncoding and intron sequences in a segment of the genome known as ENCODE region ENm001. Numbers of sites are shown on each long branch and rates of substitution per 100 sites on the shorter branches. Noncoding and intron regions show a clear slowing of the rate of substitution in humans and chimps. (From Kim et al. 2006.)

Nearly Neutral Theory

- what happens in small populations when selection is weak?
  - changes in allele frequency due to drift and selection are approximately equal $|2Ns| \approx 1$
- probability of fixation for a new, "nearly neutral" allele:

$$\Pr(A \text{ fixed}) = \frac{2s}{1 - e^{-4Ns}}$$

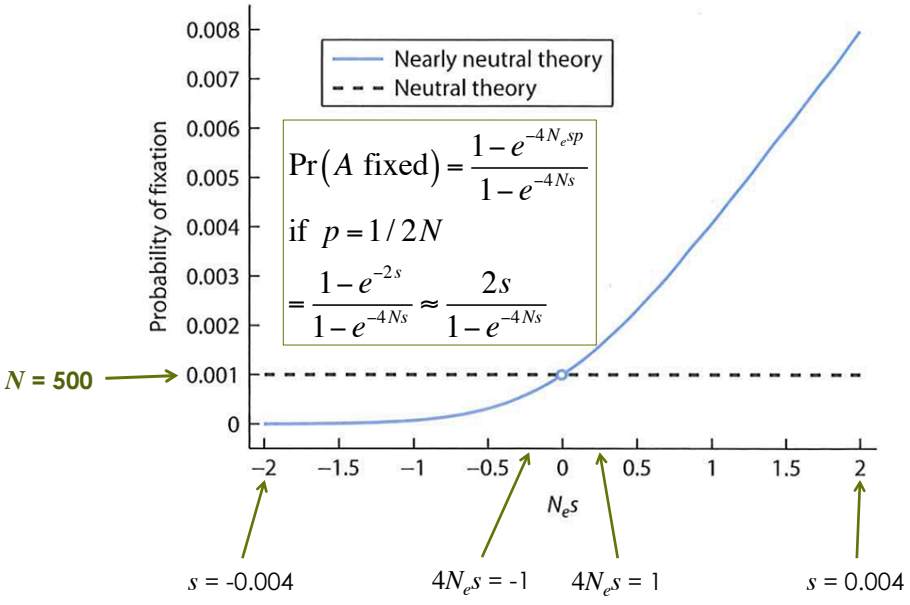
$$w_{AA} = 1 + s, \quad w_{Aa} = 1 + s/2, \quad w_{aa} = 1$$



$$\Pr(A \text{ fixed}) = \frac{2s}{1 - e^{-4Ns}}$$

What qualifies as nearly neutral?

- Hamilton: $2s = 1/2N_e$ or $4N_e s = 1$
  - value at which "the processes of genetic drift and selection are **equal**"
- Hartl & Clark: $|2Ns| \approx 1$
- Hedrick: $s < 1/(2N)$ or $2Ns < 1$
- Ohta & Gillespie (1996): $s \approx 1/N$ or $Ns \approx 1$



Nearly Neutral Theory - Summary

- the rate of neutral evolution is independent of population size
  - substitution rate equals mutation rate $2N\mu \times \frac{1}{2N} = \mu$
- in contrast, the fate of nearly neutral mutations depends on population size
  - when N is small, the effect of genetic drift can be comparable to that of selection, making slightly deleterious mutations "effectively neutral" $|2Ns| \approx 1$
- thus, lineages experiencing small population size should accumulate both neutral and nearly neutral mutations, leading to a faster rate of sequence evolution

Testing the Nearly Neutral Theory

- how to distinguish neutral and nearly neutral mutations?
  - synonymous (silent) versus non-synonymous (replacement) substitutions?
  - synonymous likely to be neutral
  - non-synonymous more likely to be deleterious

- Ohta (1994) - generation time effect differs between synonymous and non-synonymous mutations

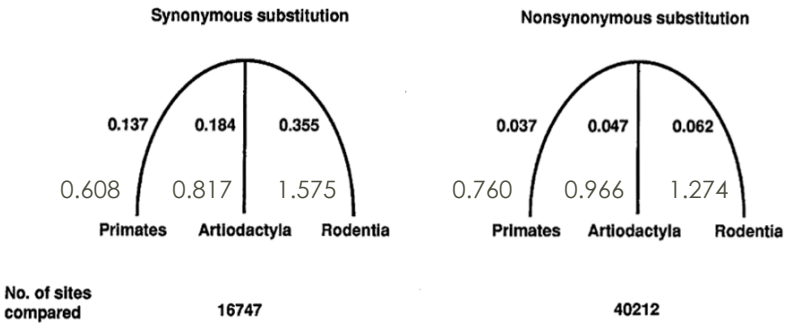
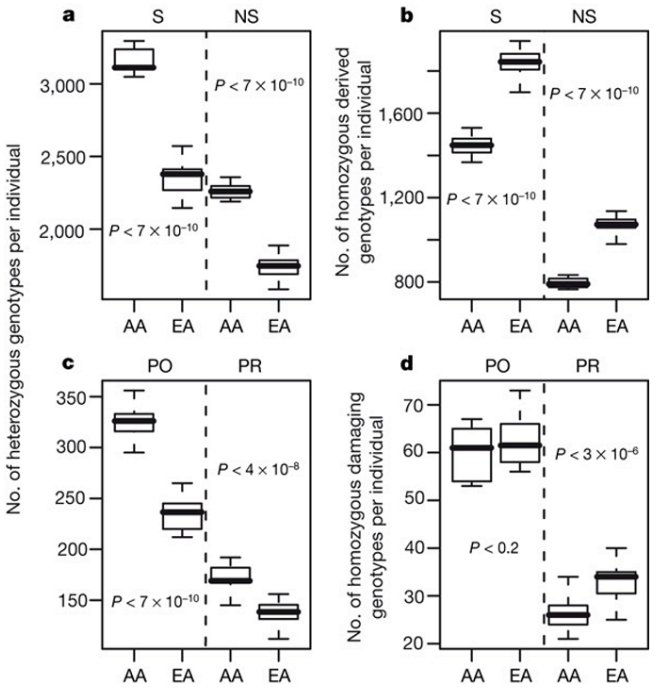


Fig. 1. Star phylogenies of 49 genes. Figures beside each branch are the estimated numbers of substitutions per site.

  - interpreted as consequence of nearly neutral evolution
  - inverse correlation between population size and body size/generation time

Lohmueller et al. 2007. Proportionally more deleterious genetic variation in European than in African populations. *Nature* 451: 994-997.

- used protein structure prediction to estimate the number of functionally consequential SNPs carried by each of 15 African Americans (AA) and 20 European Americans (EA)
- higher heterozygosity in AA, but…
- the proportion of SNPs that are non-synonymous is significantly higher in the EA sample (55.4%) than in the AA sample (47.0%)
- same result for SNPs that were inferred to be 'probably damaging' (15.9% in EA; 12.1% in AA)

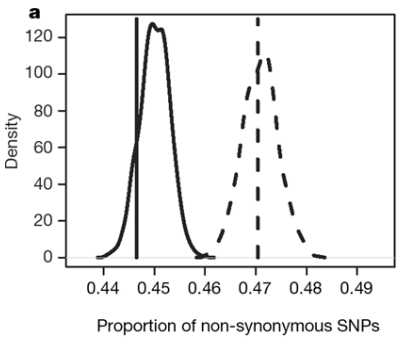


Lohmueller et al. 2007. Proportionally more deleterious genetic variation in European than in African populations. *Nature* 451: 994-997.

Table 1 | Distribution of Applera SNPs by population and functional class

Category	Shared	Private AA	Private EA	Mean derived frequency	
				AA*	EA†
Synonymous	8,056 (58.3%)	8,958 (53.0%)	3,879 (44.6%)	0.211	0.266
Non-synonymous	5,771 (41.7%)	7,950 (47.0%)	4,826 (55.4%)	0.174	0.202
Benign	4,448 (78.6%)	5,260 (67.7%)	2,928 (62.1%)	0.200	0.238
Possibly damaging	795 (14.0%)	1,572 (20.2%)	1,035 (22.0%)	0.113	0.119
Probably damaging	422 (7.4%)	942 (12.1%)	749 (15.9%)	0.099	0.108

* Average frequency from SNPs segregating in the AA sample. No correction for ancestral misidentification was used.
† Average frequency from SNPs segregating in the EA sample. No correction for ancestral misidentification was used.



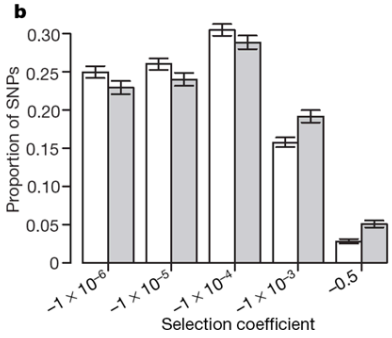


Figure 2 | Demography and selection can cause a proportional excess of non-synonymous SNPs in Europeans. **a**, **b**, Results of forward simulations of a population that expanded (AA 2 in Supplementary Table 2), to represent the AA population and a population that experienced a bottleneck to represent the EA population (EA 1 in Supplementary Table 2). **a**, Distribution of the proportion of non-synonymous SNPs segregating in samples simulated under European (dashed curve) and African (solid curve) demographic models. Vertical lines show the observed proportions in the Applera data set. **b**, Distribution of selection coefficients for simulated SNPs in the AA (white bars) and the EA (grey bars) samples. The labels on the *x* axis are the more negative limits of the bins. Error bars denote 95% intervals on the proportion of SNPs in each group.

Lohmueller et al. 2007. Proportionally more deleterious genetic variation in European than in African populations. *Nature* 451: 994-997.