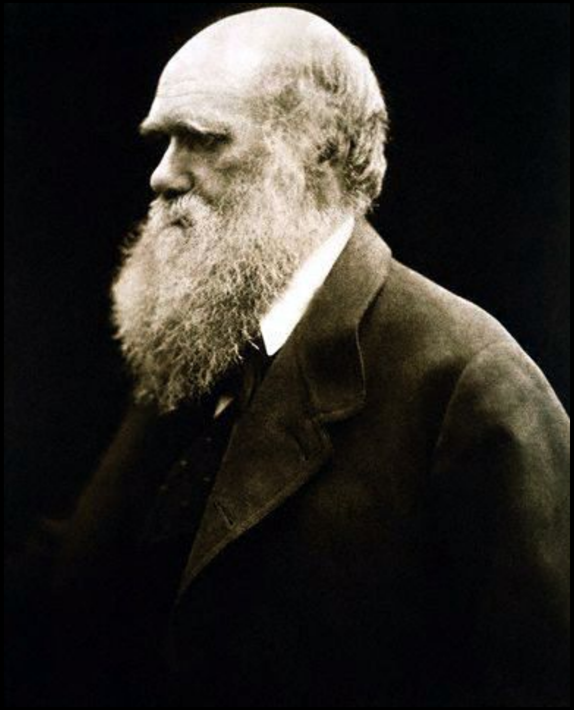
Популяционная генетика - 6

Факторы эволюции

- мутации
- отбор
- половое размножение
- генетико-автоматические процессы (генетический дрейф)
- миграция

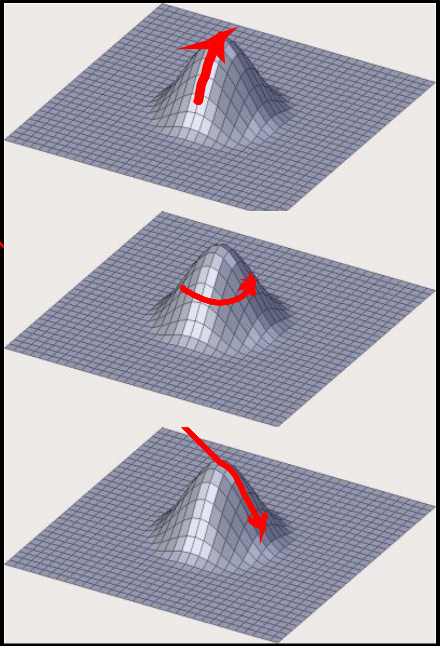
Как отличить Дарвина от Кимуры?



тест *Dn/Ds*

вид 1 ATG TCC CTA TAC GGA GCT

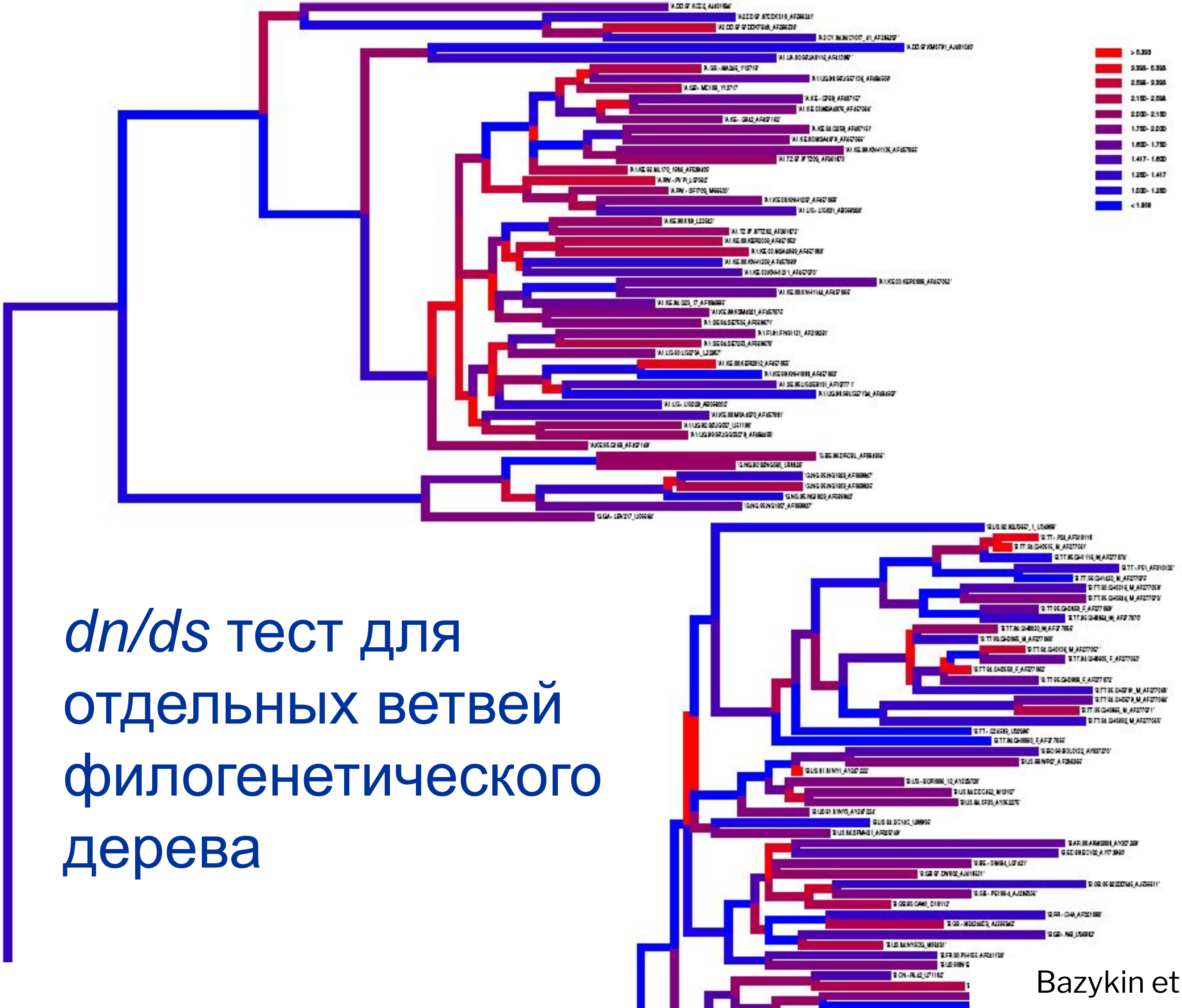
вид 2 ATG TCT CAT TAT AGA GAT



положительный отбор: *Dn*/*Ds* > 1

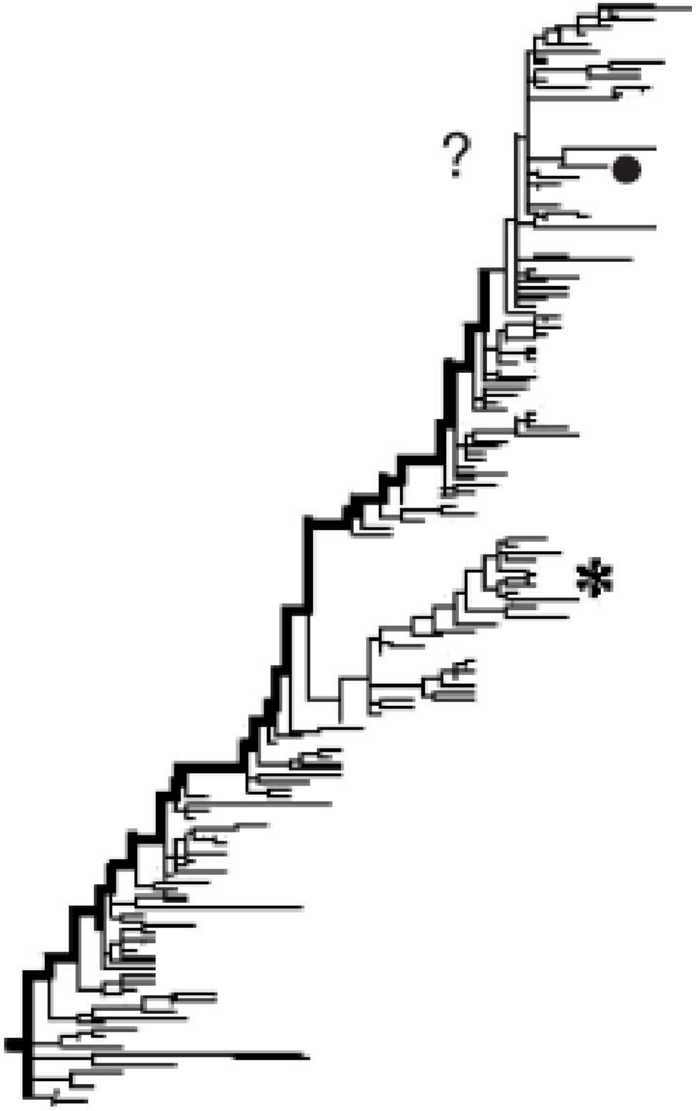
нейтральная эволюция: *Dn*/*Ds* = 1

отрицательный отбор: *Dn*/*Ds* < 1



dn/ds тест для отдельных ветвей филогенетического дерева

Bazykin et al. *PNAS*



A. 1993-94 test tree

Bush et al. *Science*

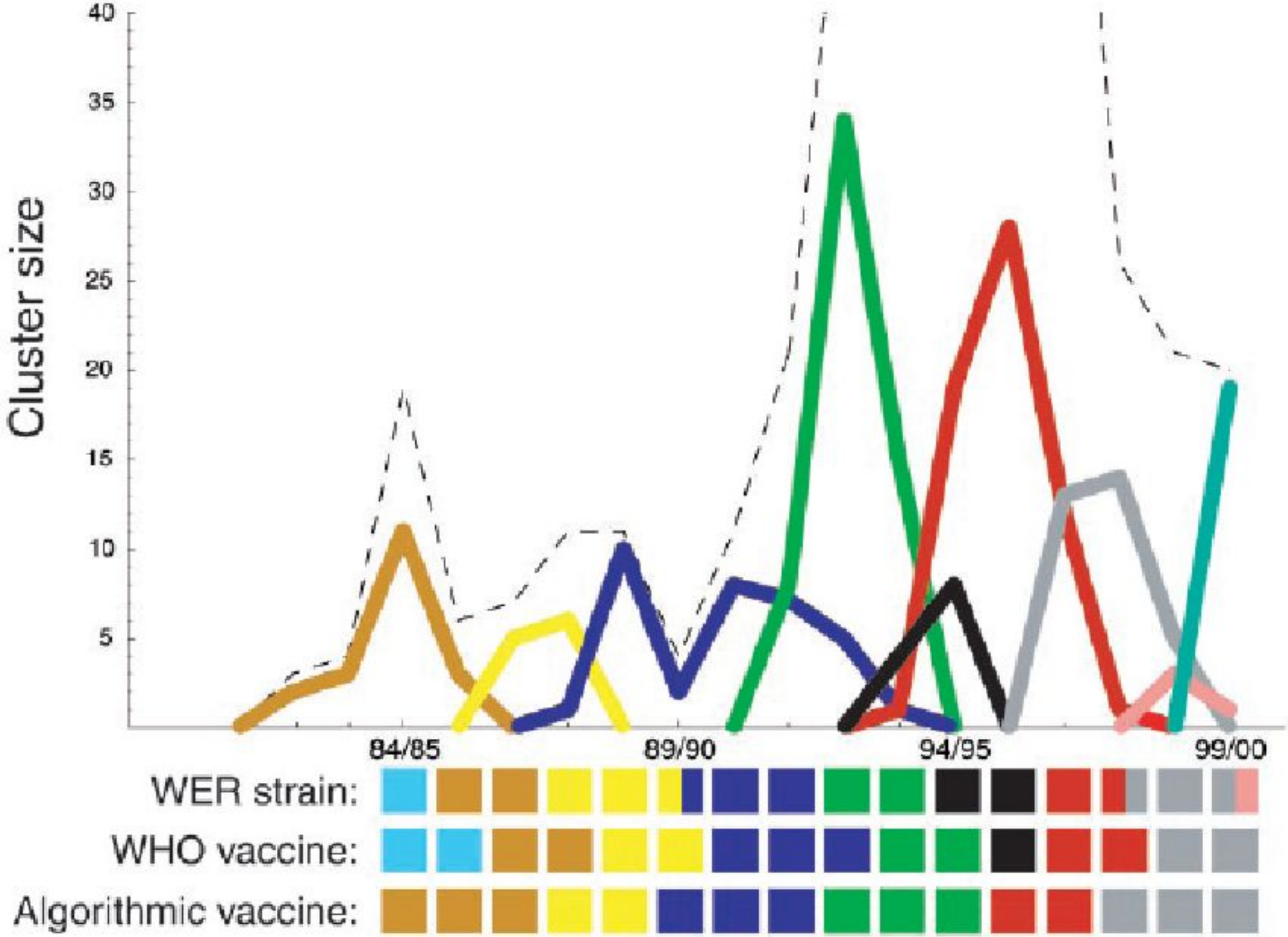
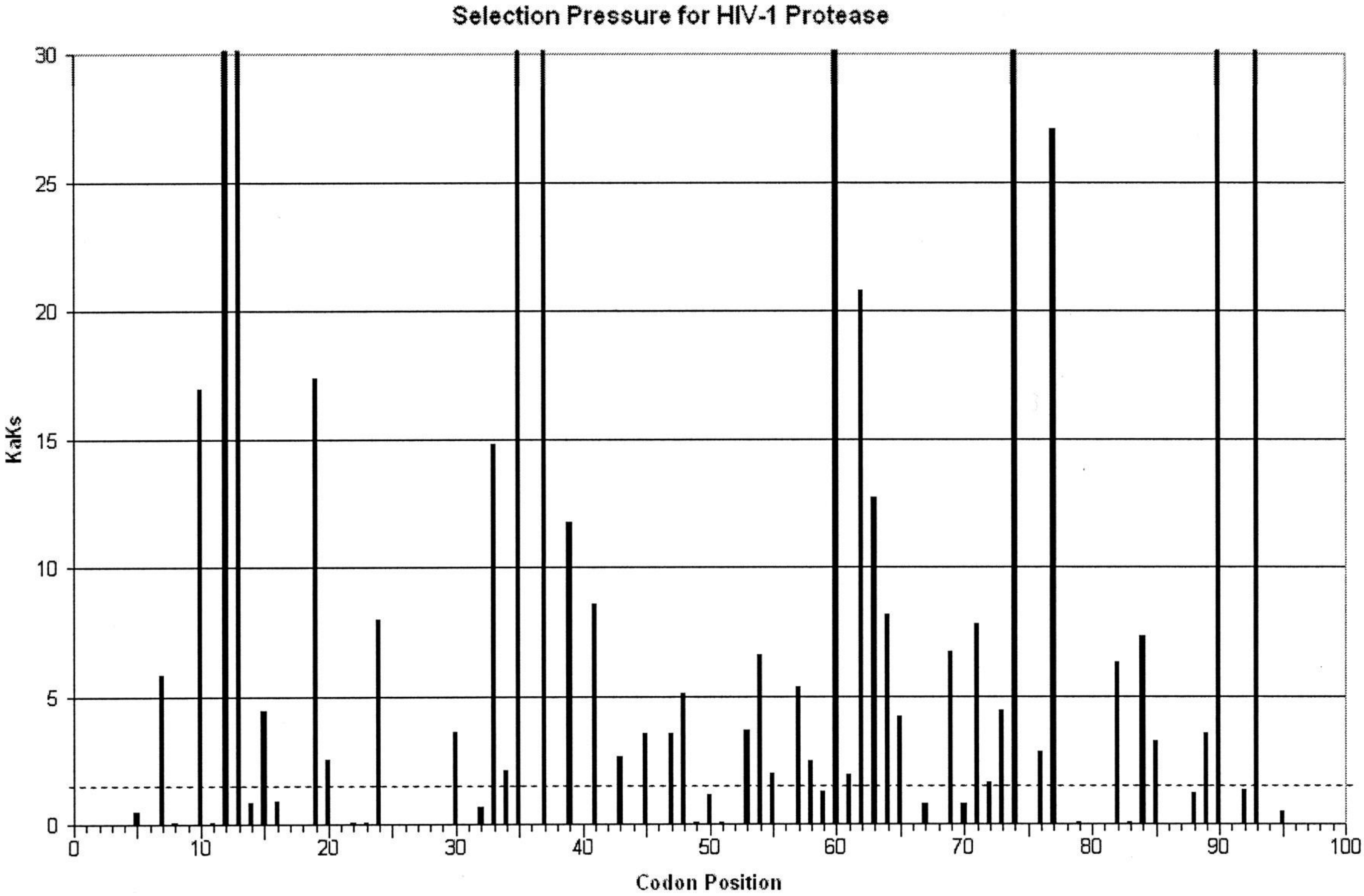
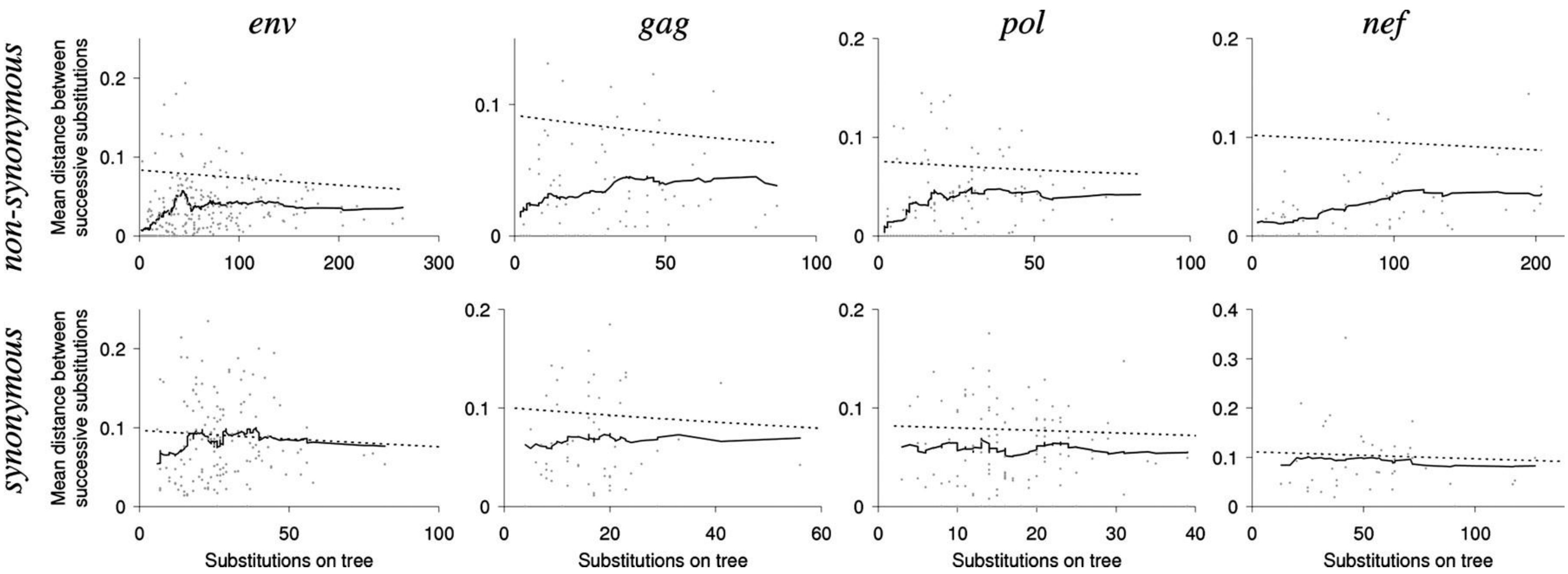


Fig. 3. The number of HA1 sequences within each cluster plotted as a function of influenza season. The graph shows the eight largest clusters as well as any other clusters that contain sequences used in a WHO vaccine. The tiles denoted "WHO vaccine" indicate the 'color' of the WHO-recommended vaccine in each season, e.g., the color of the cluster corresponding to the strain on which each vaccine was based. The tiles denoted "Algorithmic vaccine" indicate the color of vaccine prescribed each season by the algorithm proposed in the main text (the dominant cluster from the previous season). The tiles denoted "WER strain" indicate the color of the dominant antigenic type, based on HI assays, as reported by the WHO in its *Weekly Epidemiological Record* (40–56). (Note that one of the three strains reported in WER for 1999–2000 is missing from the Los Alamos sequence database.) Both vaccines tend to match the WER strains well; in some seasons the WHO vaccine matches better, and in some seasons the algorithmic vaccine matches better.

Plotkin et al. *PNAS*

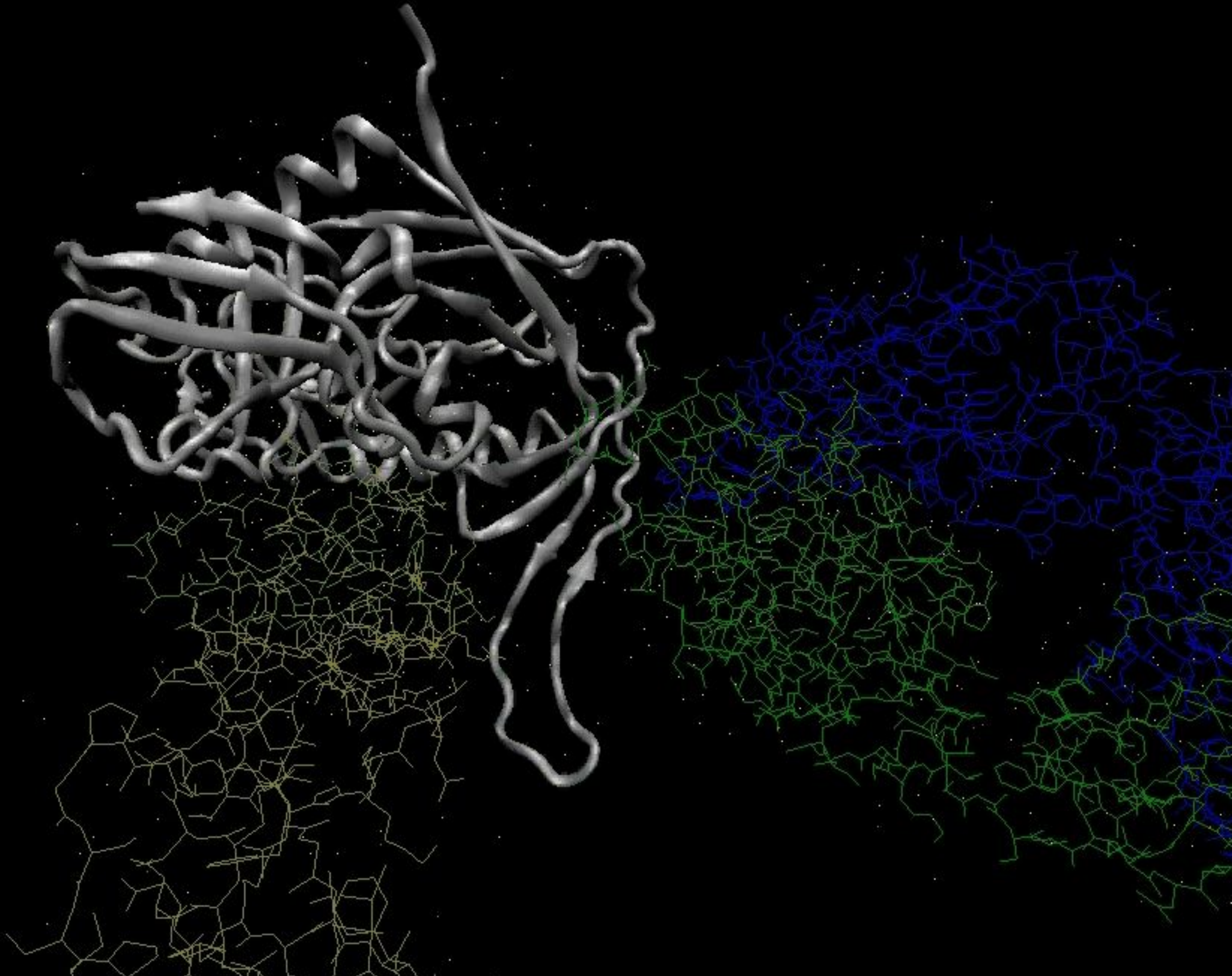


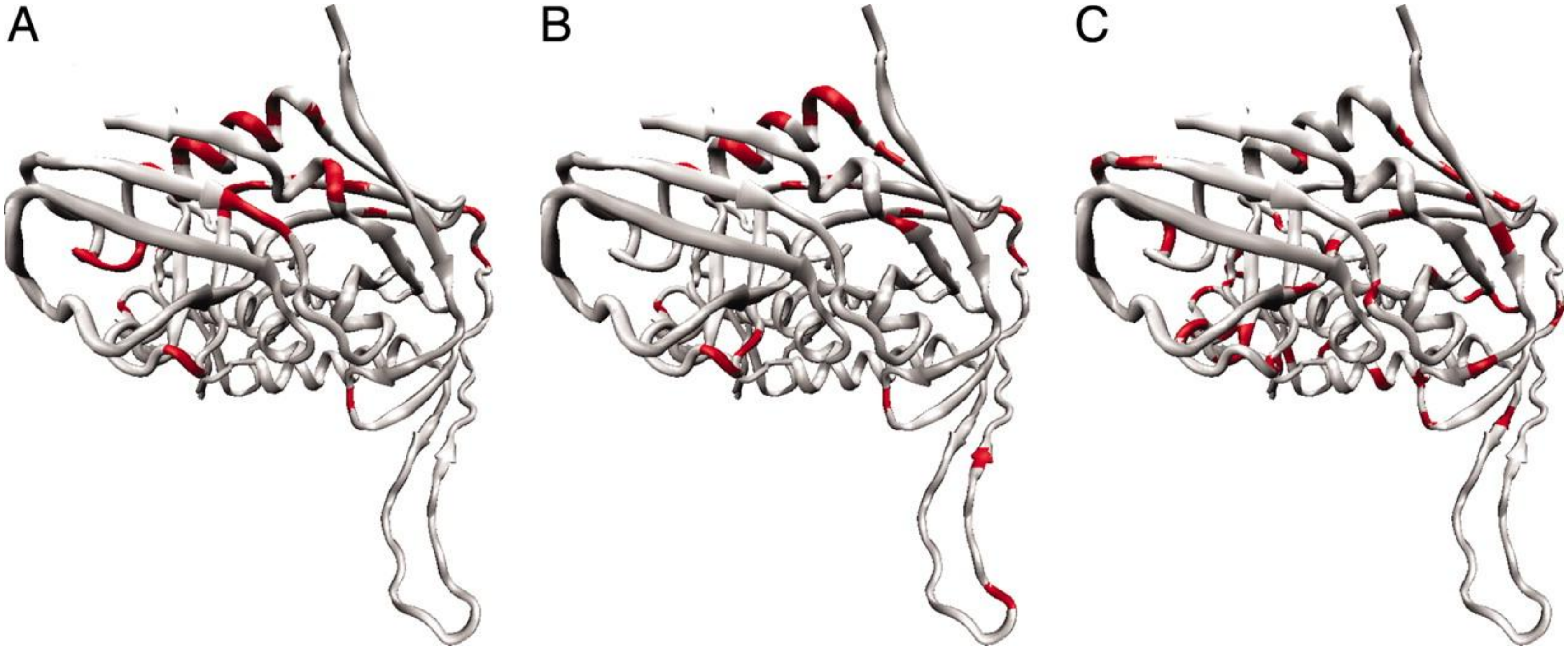
Positive selection mapping of HIV-1 protease from 40,000 patient samples. The *Ka*/*Ks* value represents the greatest selection pressure among all the individual amino acid mutations at each codon. The dotted line indicates the *Ka*/*Ks* value of 1 (Ka = d_N, Ks = d_S)



Distances between successive nonsynonymous (*Upper*) and synonymous (*Lower*) substitutions on the phylogenetic trees. The mean distance between successive nonreversing substitutions is shown for codon sites with each total number of substitutions on the tree (dots). Solid lines present mean distances between successive substitutions within each sliding 30-site window. Dashed lines show mean distances between independent substitutions obtained in simulations.

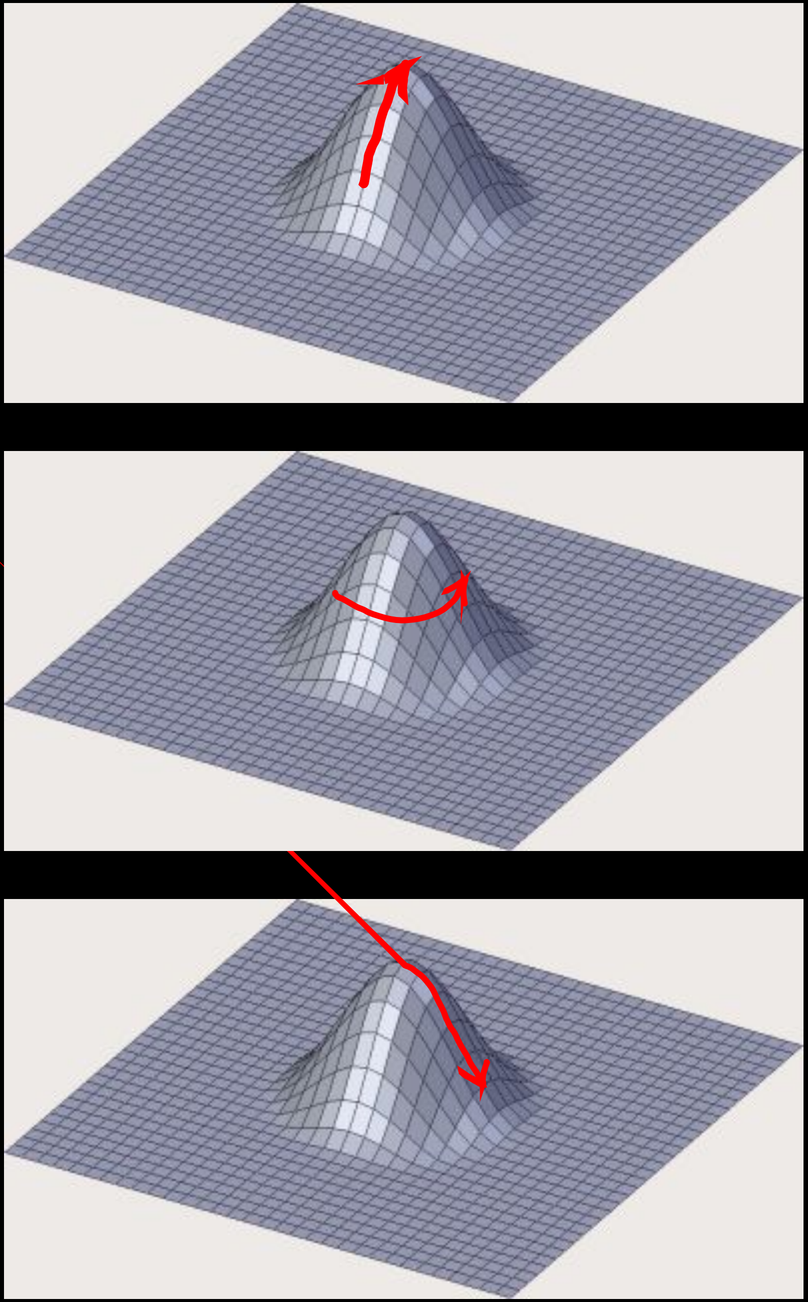
Bazykin et al. *PNAS*





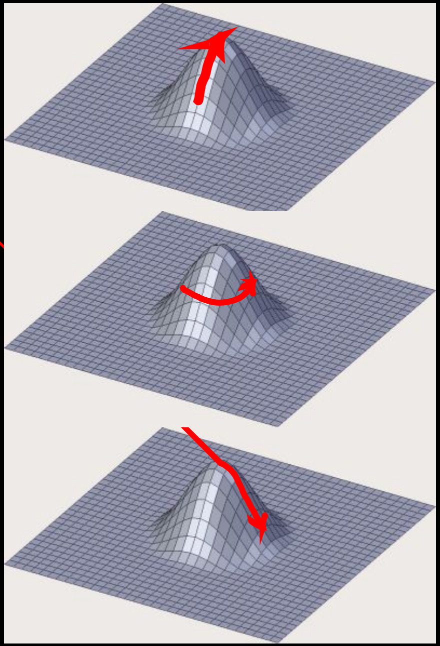
Amino acid sites inferred to be under positive selection in HIV-1 gp120. (*A*) Amino acid sites with >80 replacements. (*B*) Rapidly evolving sites previously inferred to be under positive selection. (*C*) Conservative sites (<80 replacements) with strongly clumped substitutions.

Bazykin et al. *PNAS*



тест *Dn/Ds*

вид 1 ATG TCC CTA TAC GGA GCT
вид 2 ATG TCT CAT TAT AGA GAT



положительный отбор: *Dn*/*Ds* > 1

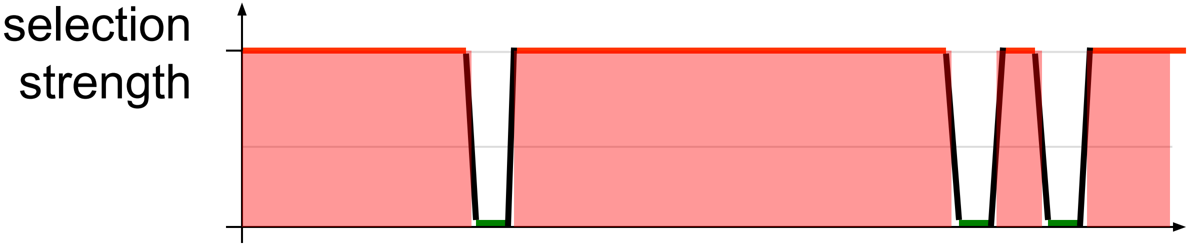
нейтральная эволюция: *Dn*/*Ds* = 1

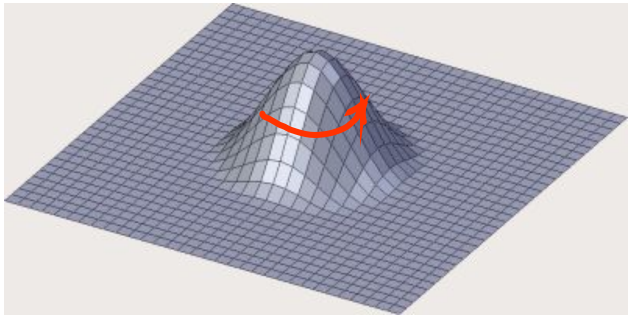
отрицательный отбор: *Dn*/*Ds* < 1

- ***Dn/Ds* = 0.1: что это значит?**


REEVDNKLKGIMENIYTTCRD
REEVDAKLKGIMENIYNTSRD

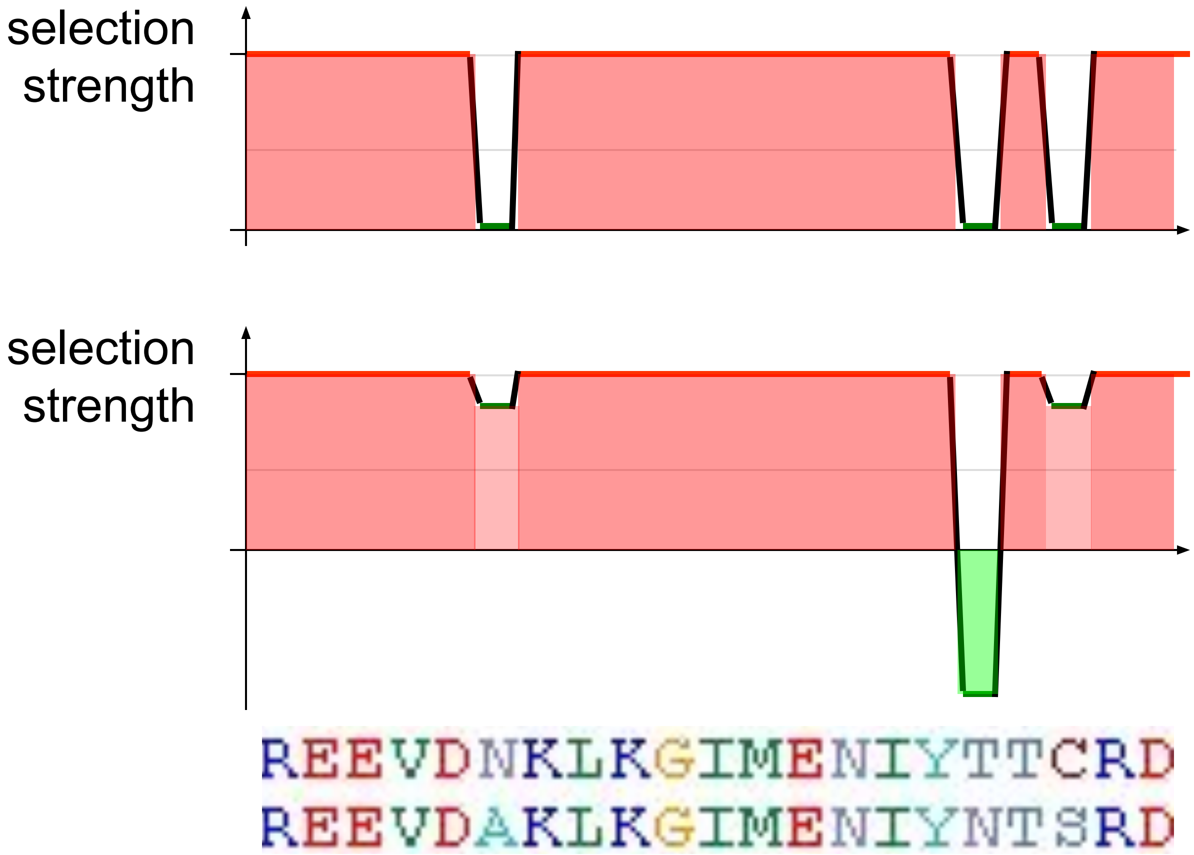
- ***Dn/Ds*** **= 0.1: что это значит?**

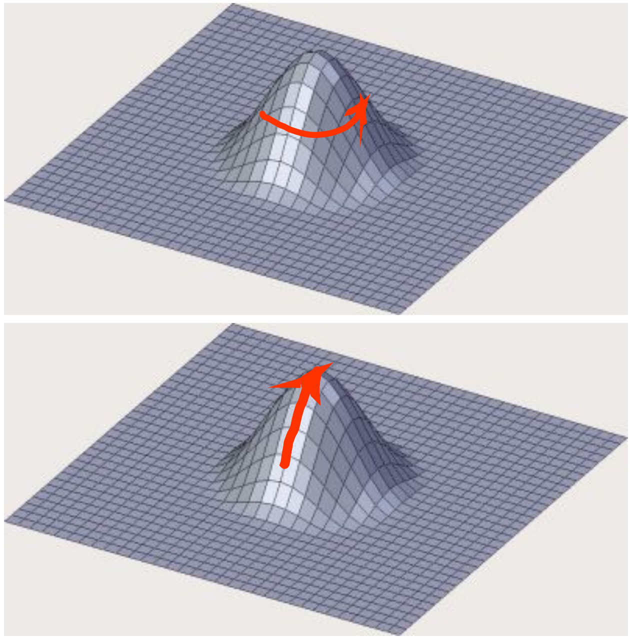




REEVDNKLKGIMENIYTTCRD
REEVDAKLKGIMENIYNTSRD

- ***Dn/Ds*** **= 0.1: что это значит?**





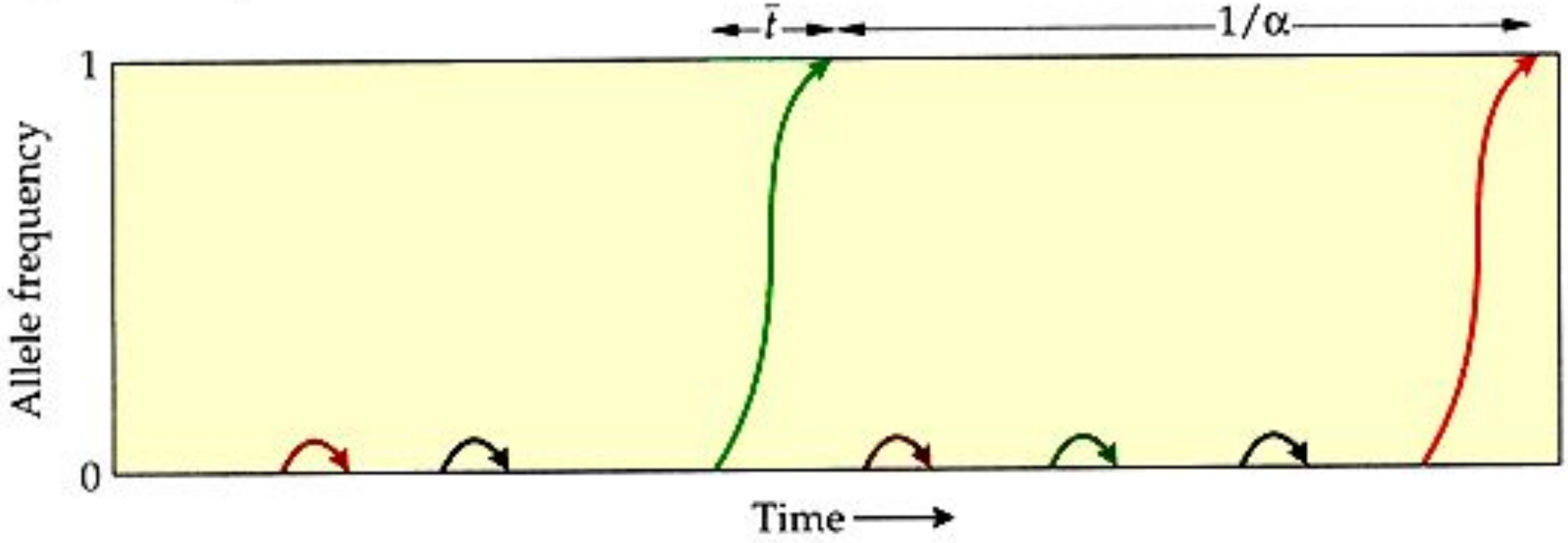
Adaptive protein evolution at the *Adh* locus in *Drosophila*

John H. McDonald & Martin Kreitman

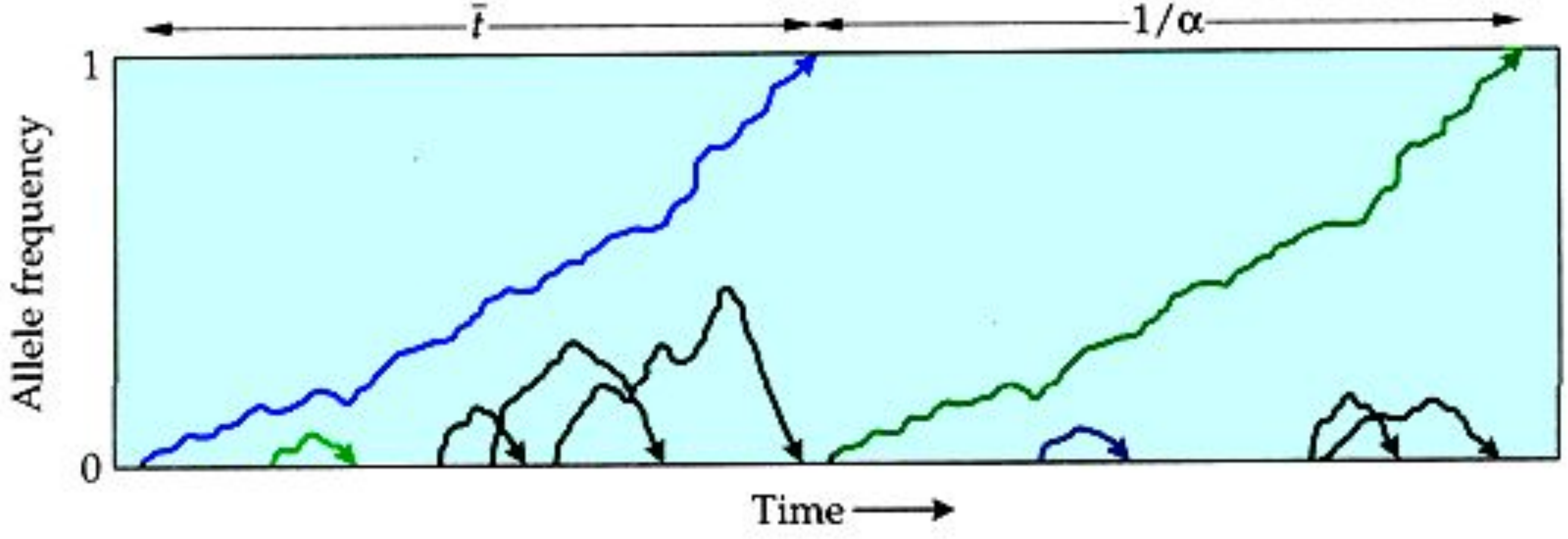
Department of Ecology and Evolutionary Biology, Princeton University,
Princeton, New Jersey 08544, USA

PROTEINS often differ in amino-acid sequence across species. This difference has evolved by the accumulation of neutral mutations by random drift, the fixation of adaptive mutations by selection, or a mixture of the two. Here we propose a simple statistical test of the neutral protein evolution hypothesis based on a comparison of the number of amino-acid replacement substitutions to synonymous substitutions in the coding region of a locus. If the observed substitutions are neutral, the ratio of replacement to synonymous fixed differences between species should be the same as the ratio of replacement to synonymous polymorphisms within species. DNA sequence data on the *Adh* locus (encoding alcohol dehydrogenase, EC 1.1.1.1) in three species in the *Drosophila melanogaster* species subgroup do not fit this expectation; instead, there are more fixed replacement differences between species than expected. We suggest that these excess replacement substitutions result from adaptive fixation of selectively advantageous mutations.

(a) Advantageous mutations



(b) Neutral mutations



Тест Макдональда-Крейтмана

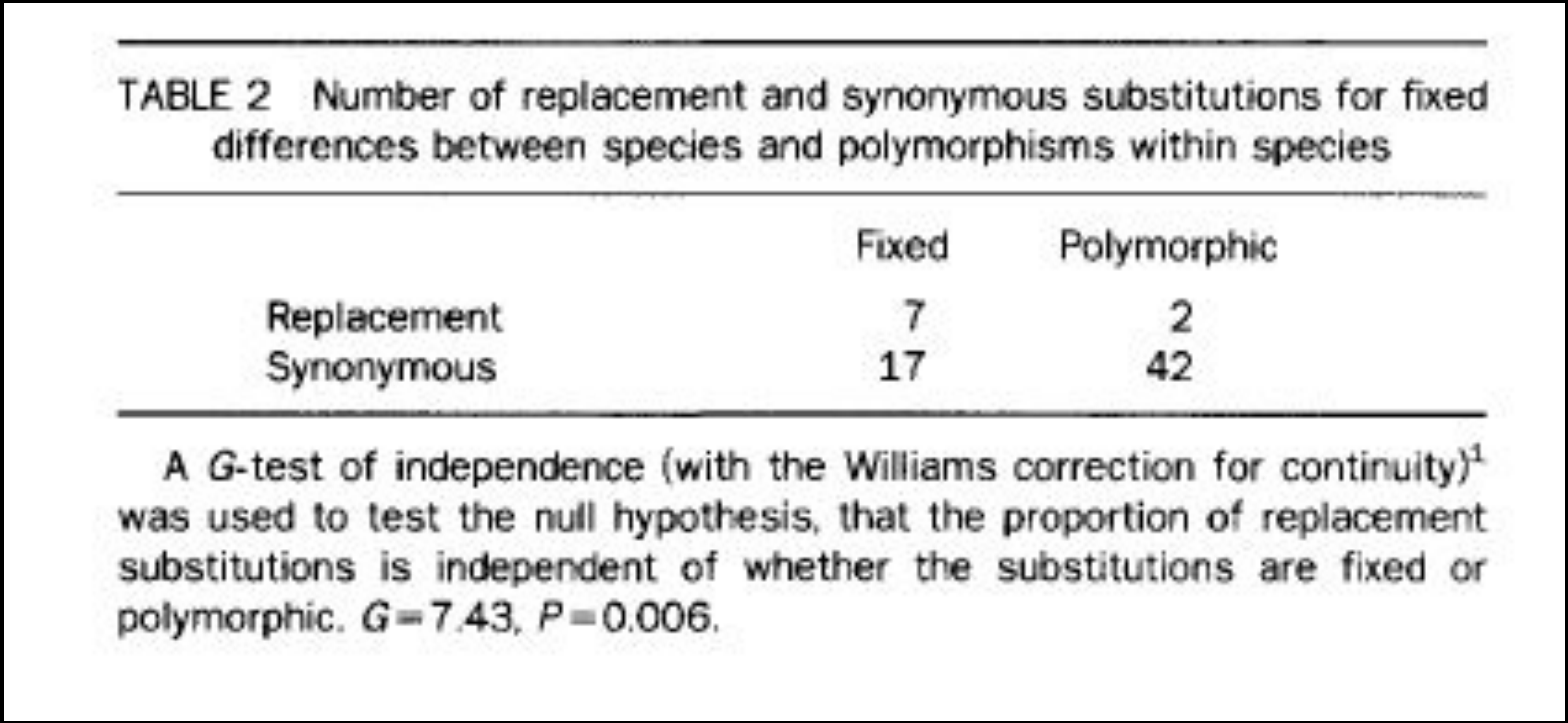
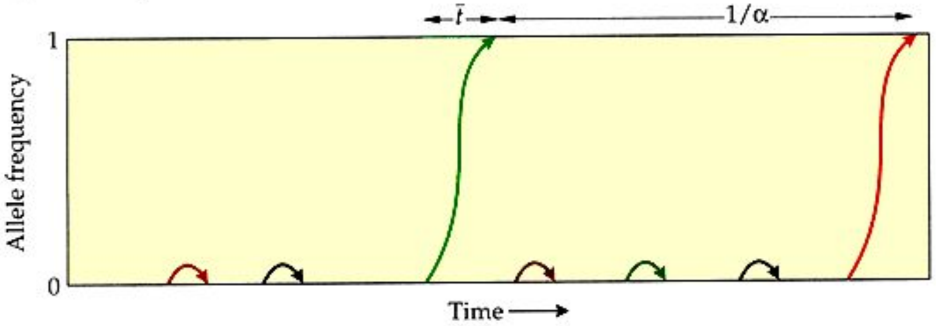


TABLE 2 Number of replacement and synonymous substitutions for fixed differences between species and polymorphisms within species

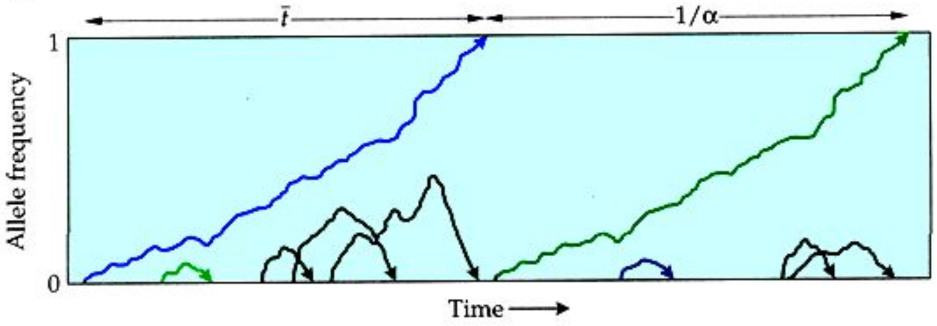
	Fixed	Polymorphic
Replacement	7	2
Synonymous	17	42

A *G*-test of independence (with the Williams correction for continuity)[1] was used to test the null hypothesis, that the proportion of replacement substitutions is independent of whether the substitutions are fixed or polymorphic. $G = 7.43$, $P = 0.006$.

(a) Advantageous mutations



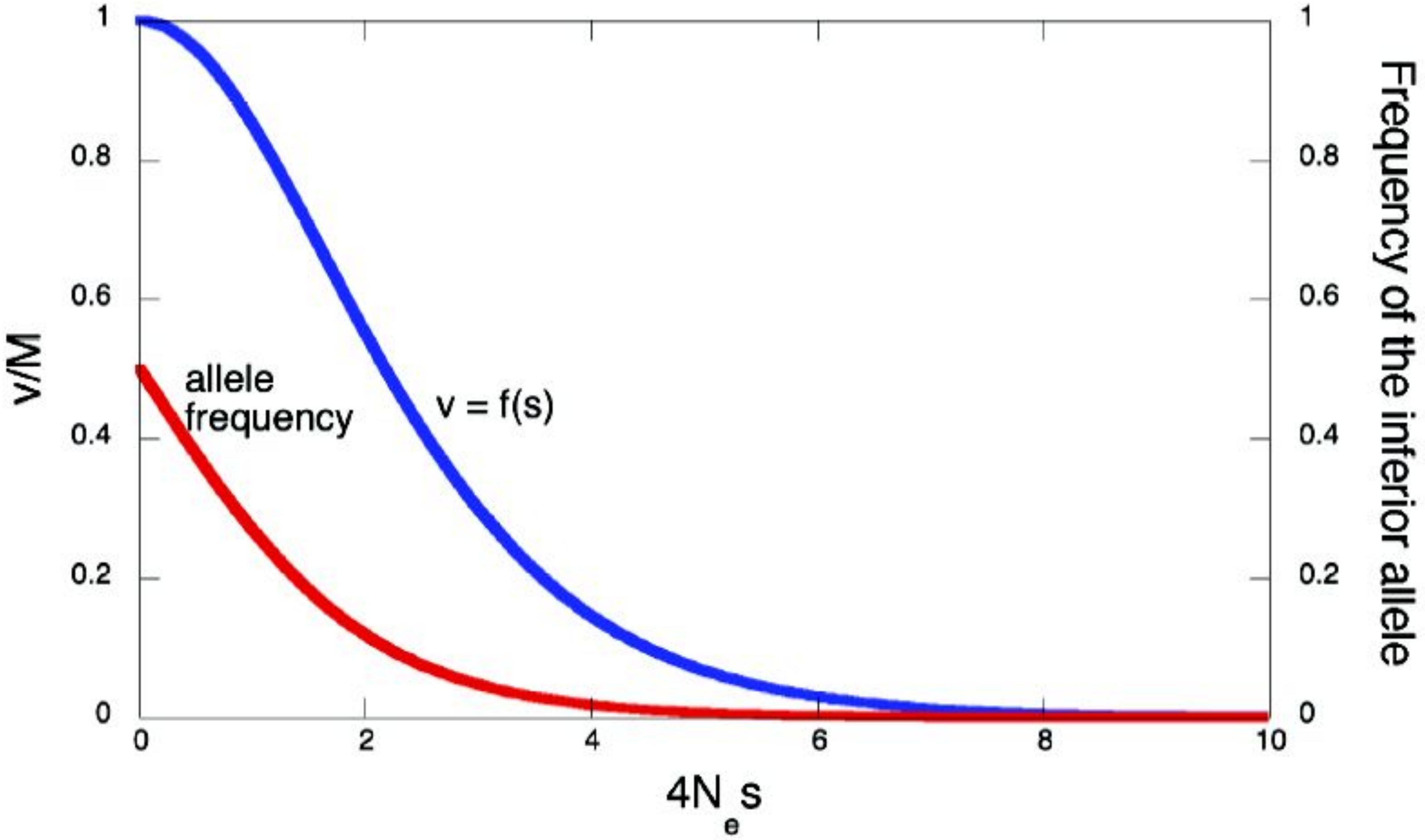
(b) Neutral mutations



$$k_{\text{полезная}} \sim 2Nus$$

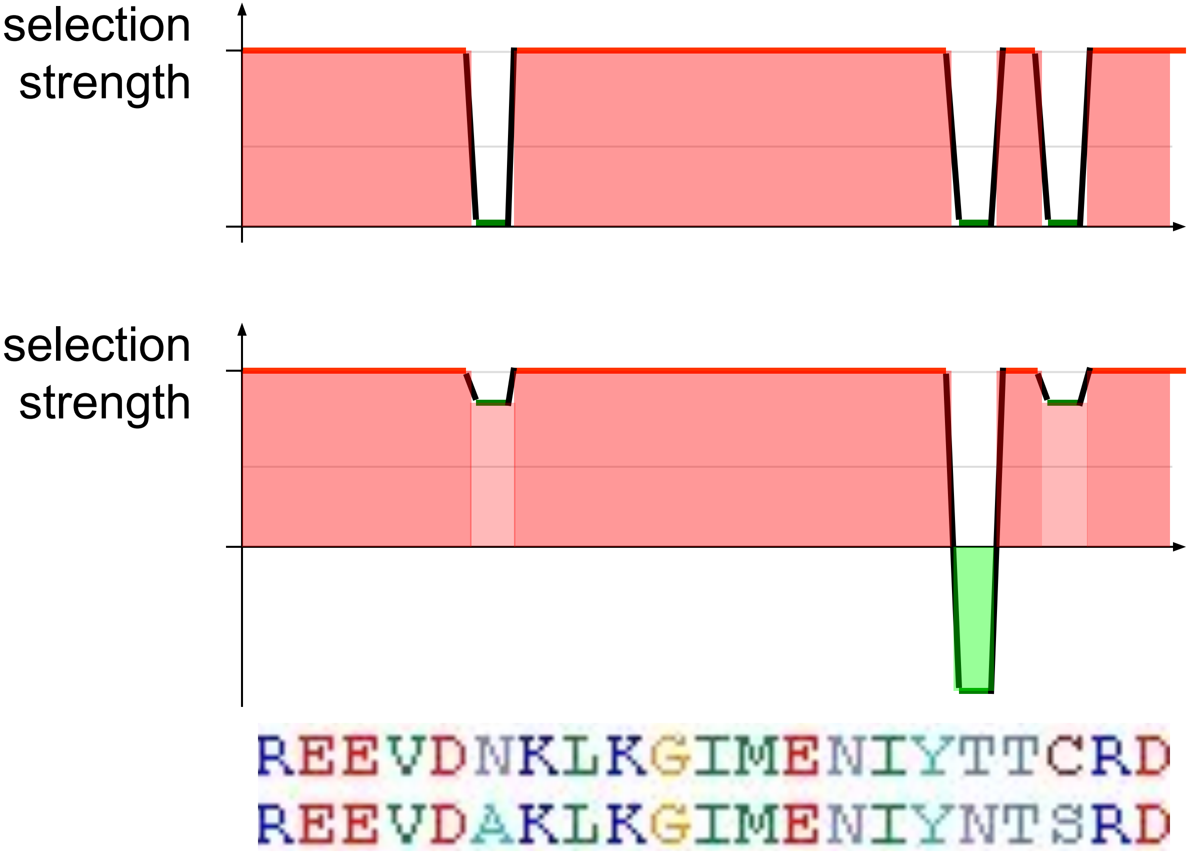
$$k_{\text{нейтральная}} \sim u$$

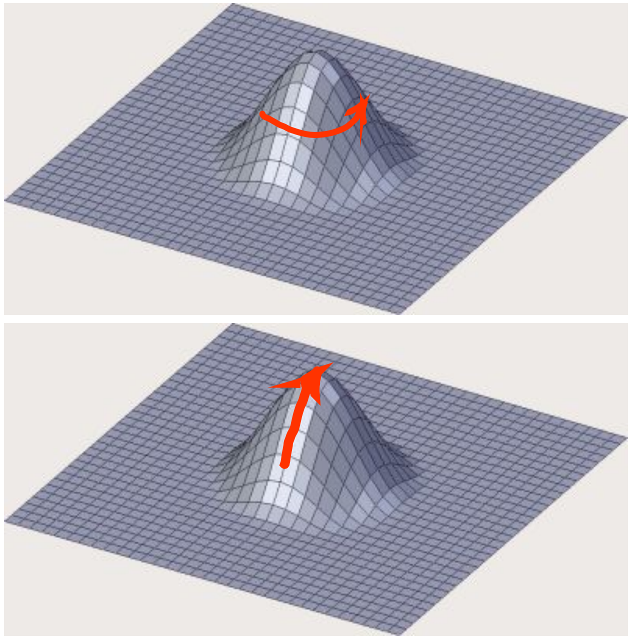
$$k_{\text{вредная}} = \frac{2Nus}{e^{2Ns}-1}$$



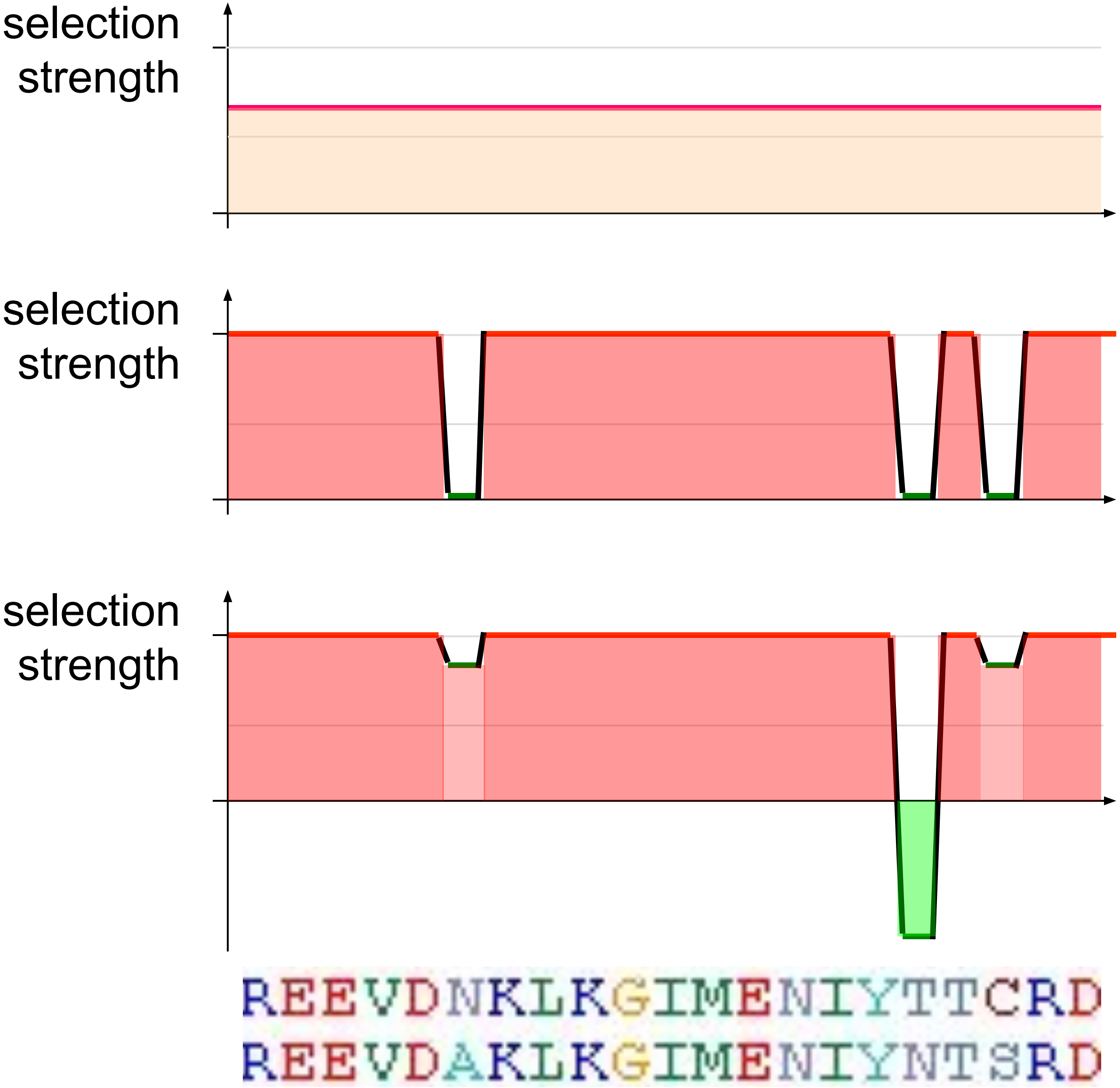
Bazykin et al. *Biology Direct*

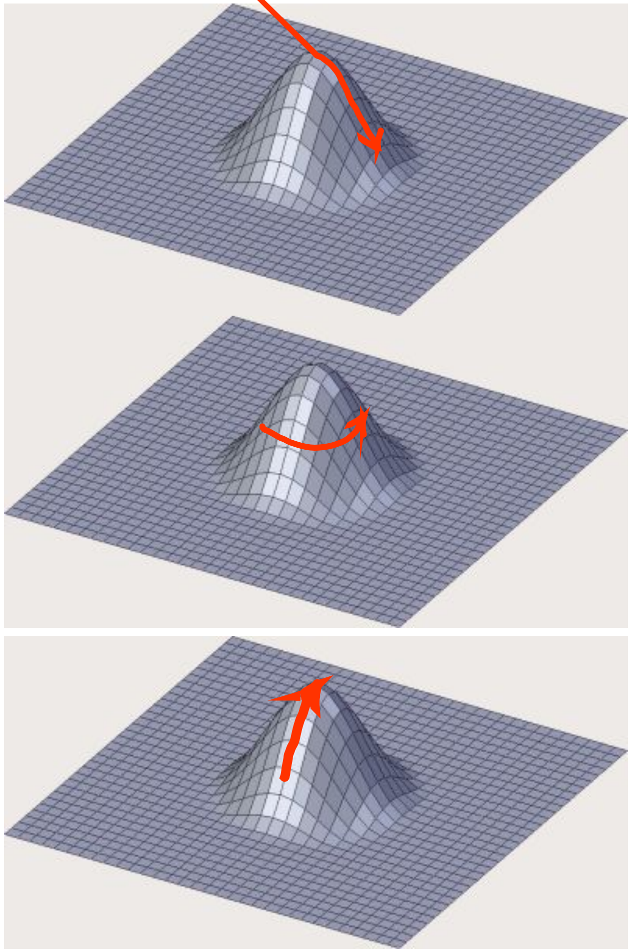
- *Dn/Ds* = 0.1: что это значит?

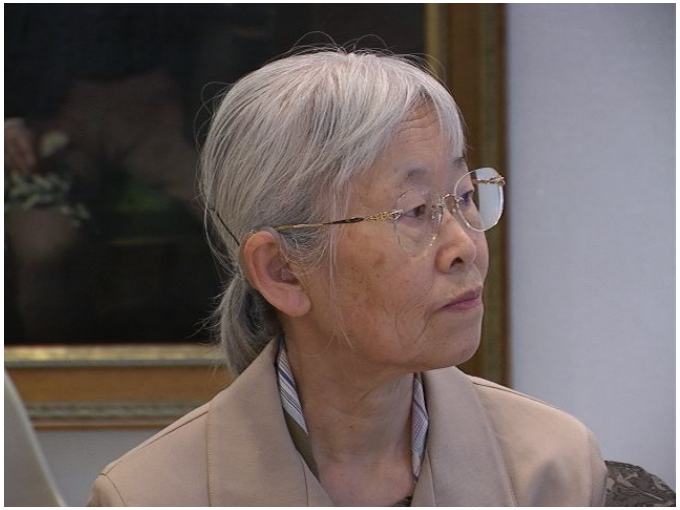


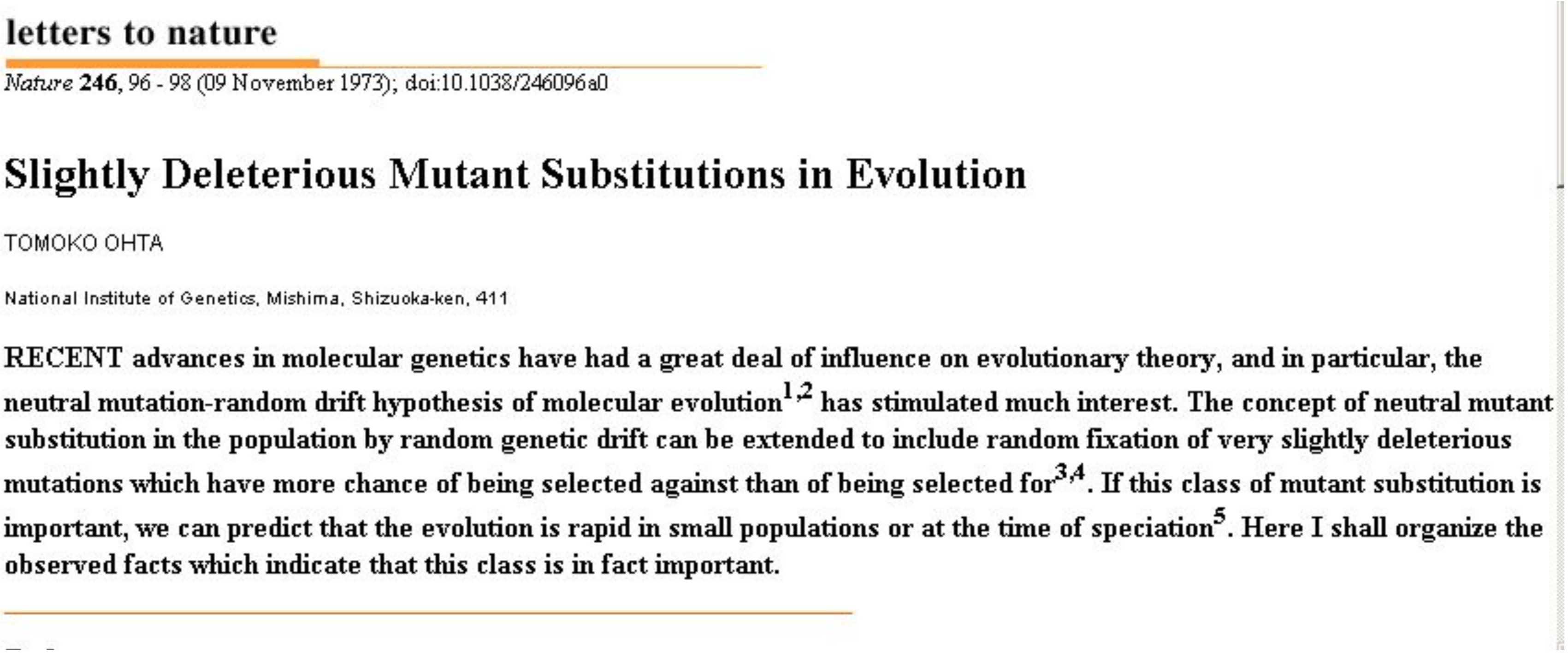


- ***Dn/Ds*** **= 0.1: что это значит?**









letters to nature

Nature **246**, 96 - 98 (09 November 1973); doi:10.1038/246096a0

Slightly Deleterious Mutant Substitutions in Evolution

TOMOKO OHTA

National Institute of Genetics, Mishima, Shizuoka-ken, 411

RECENT advances in molecular genetics have had a great deal of influence on evolutionary theory, and in particular, the neutral mutation-random drift hypothesis of molecular evolution[1,2] has stimulated much interest. The concept of neutral mutant substitution in the population by random genetic drift can be extended to include random fixation of very slightly deleterious mutations which have more chance of being selected against than of being selected for[3,4]. If this class of mutant substitution is important, we can predict that the evolution is rapid in small populations or at the time of speciation[5]. Here I shall organize the observed facts which indicate that this class is in fact important.

Тест Макдональда-Крейтмана

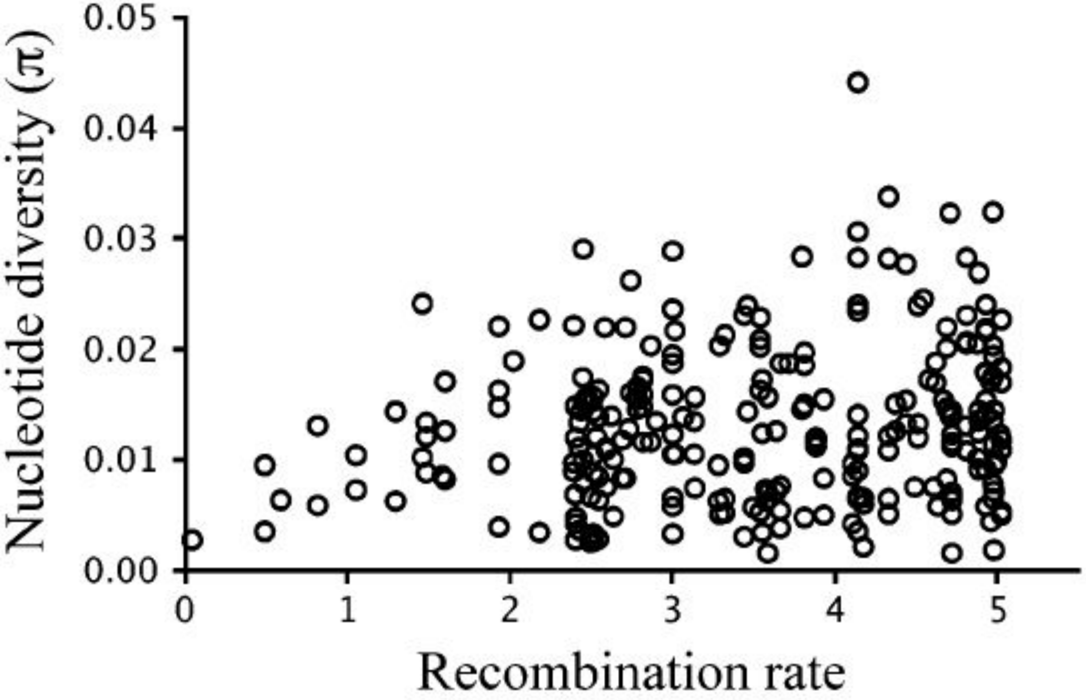
Основное объяснение:

- движущий отбор

Альтернативное объяснение:

- усиление стабилизирующего отбора (напр. из-за увеличения размера популяции)

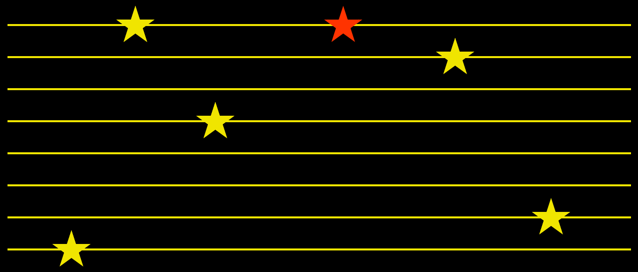
➨ этот тест чувствителен к особенностям популяционной демографии



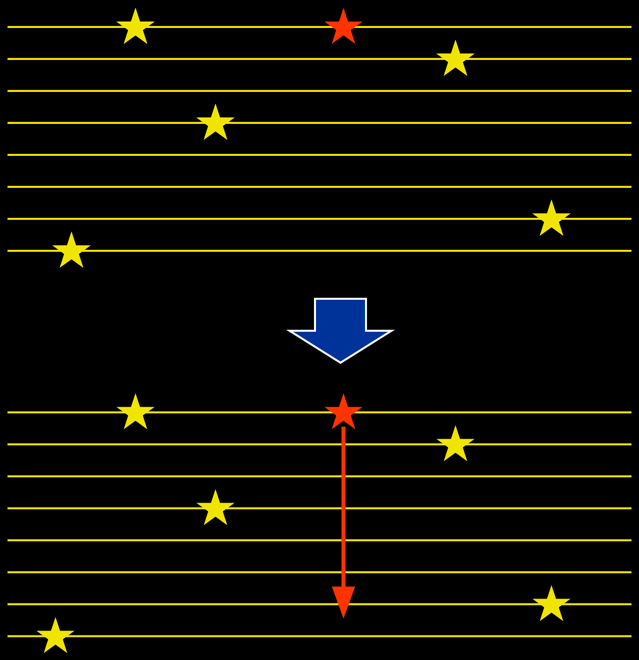
Ometto et al. *Mol. Biol. Evol.*

«Выметание» отбором

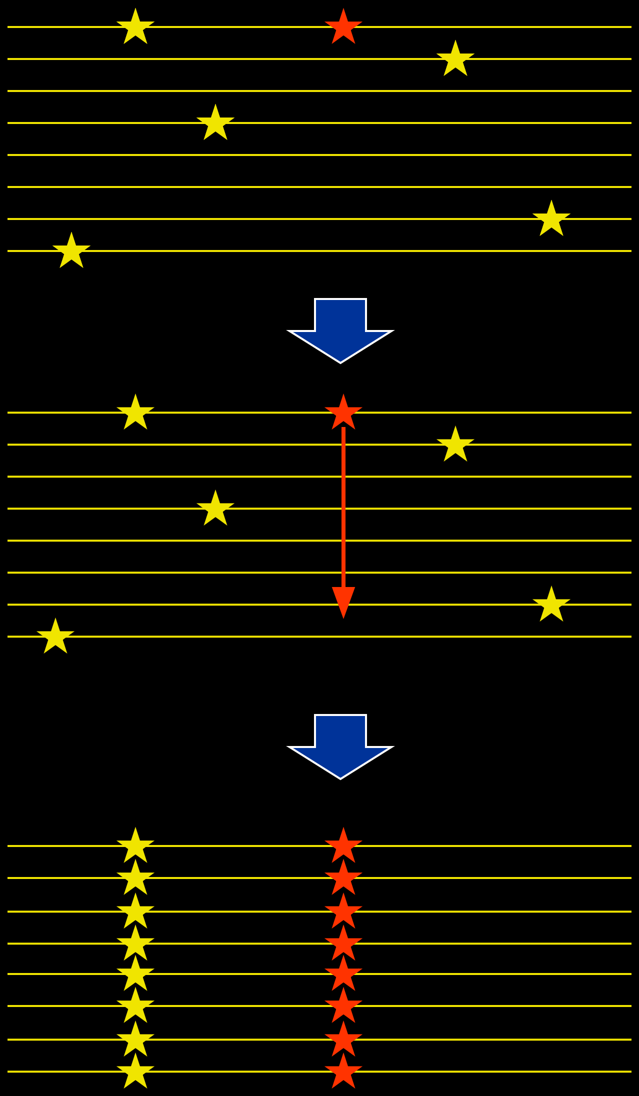
«Выметание» отбором



«Выметание» отбором



«Выметание» отбором



Вокруг сайта под отбором изменчивости не будет!

«Выметание» отбором

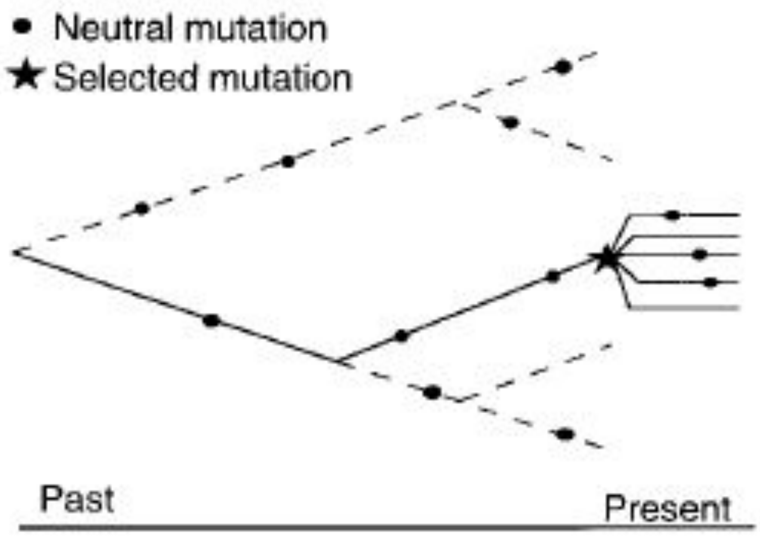
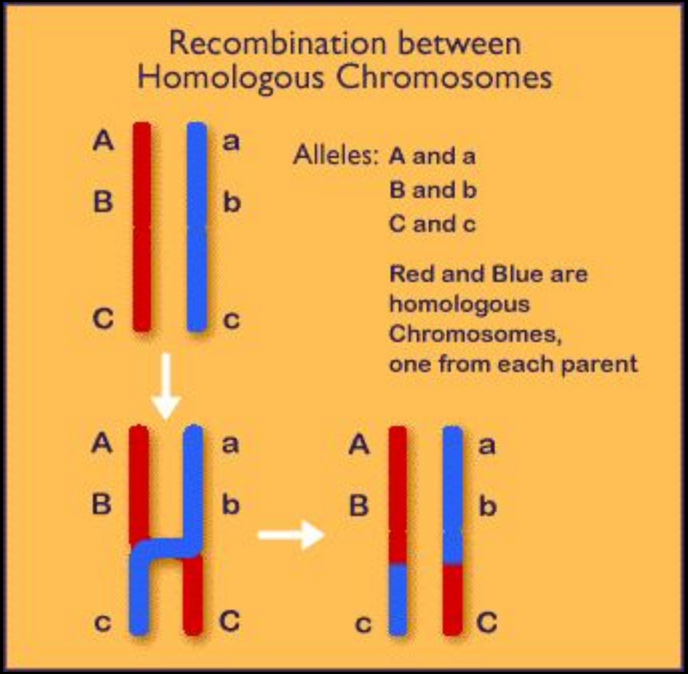


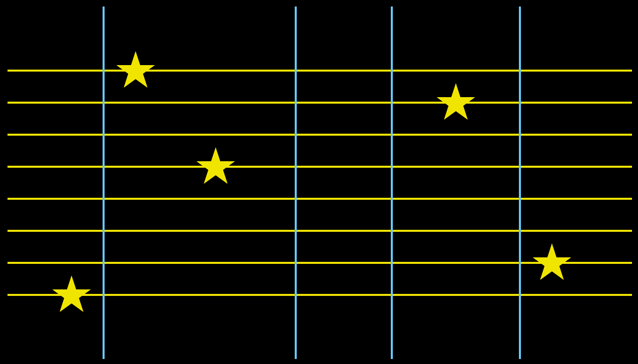
Figure 1 Genealogy of a genomic region affected by a simple selective sweep. The dashed lines indicate lineages that are lost during the sweep. Note the difference in shape of the genealogy before and after the sweep.

Hamblin and Di Rienzo, 2000
Am. J. Hum. Genet. 66:1669–1679

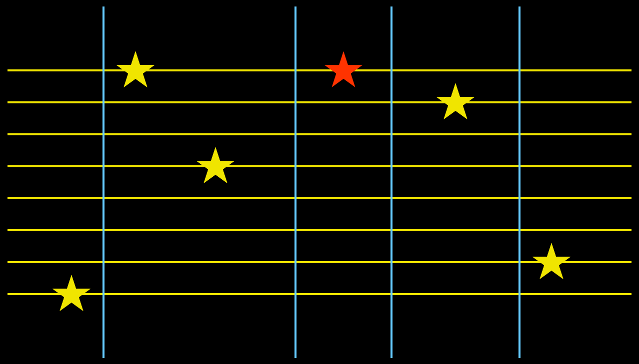
«Выметание» отбором



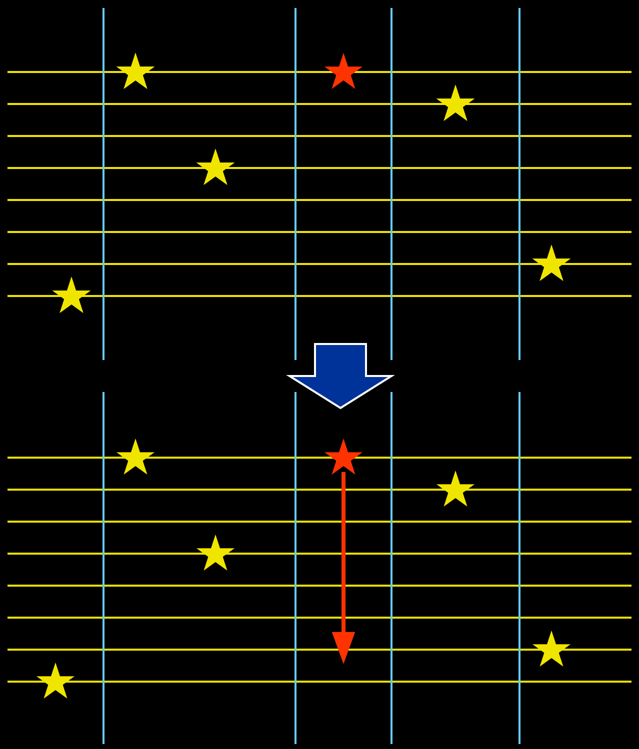
«Выметание» отбором



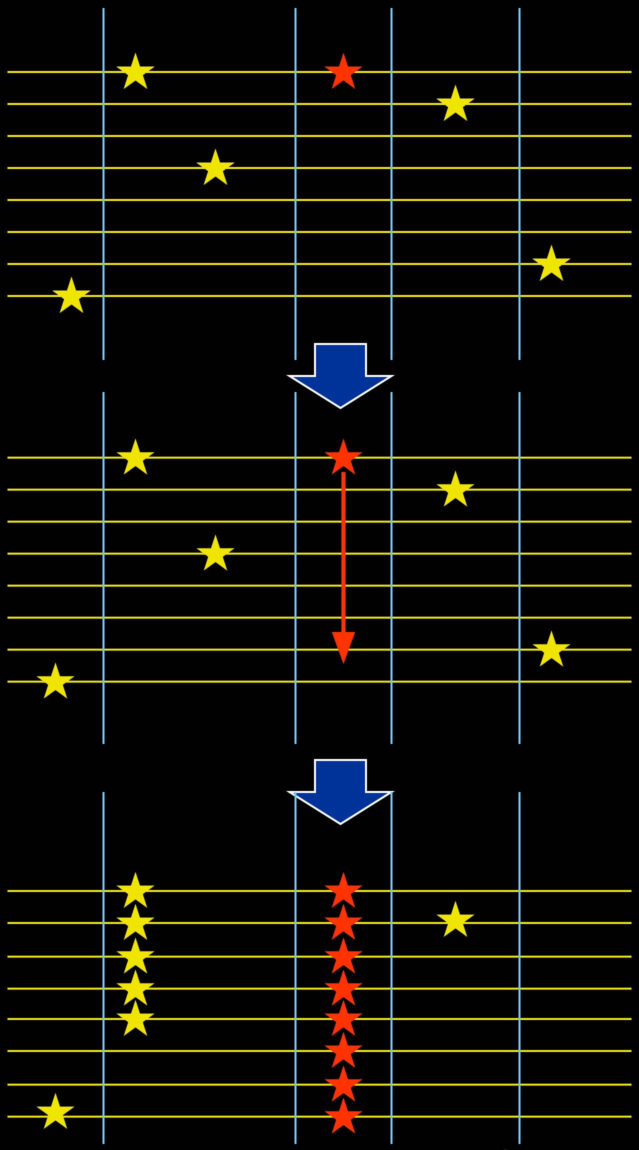
«Выметание» отбором



«Выметание» отбором

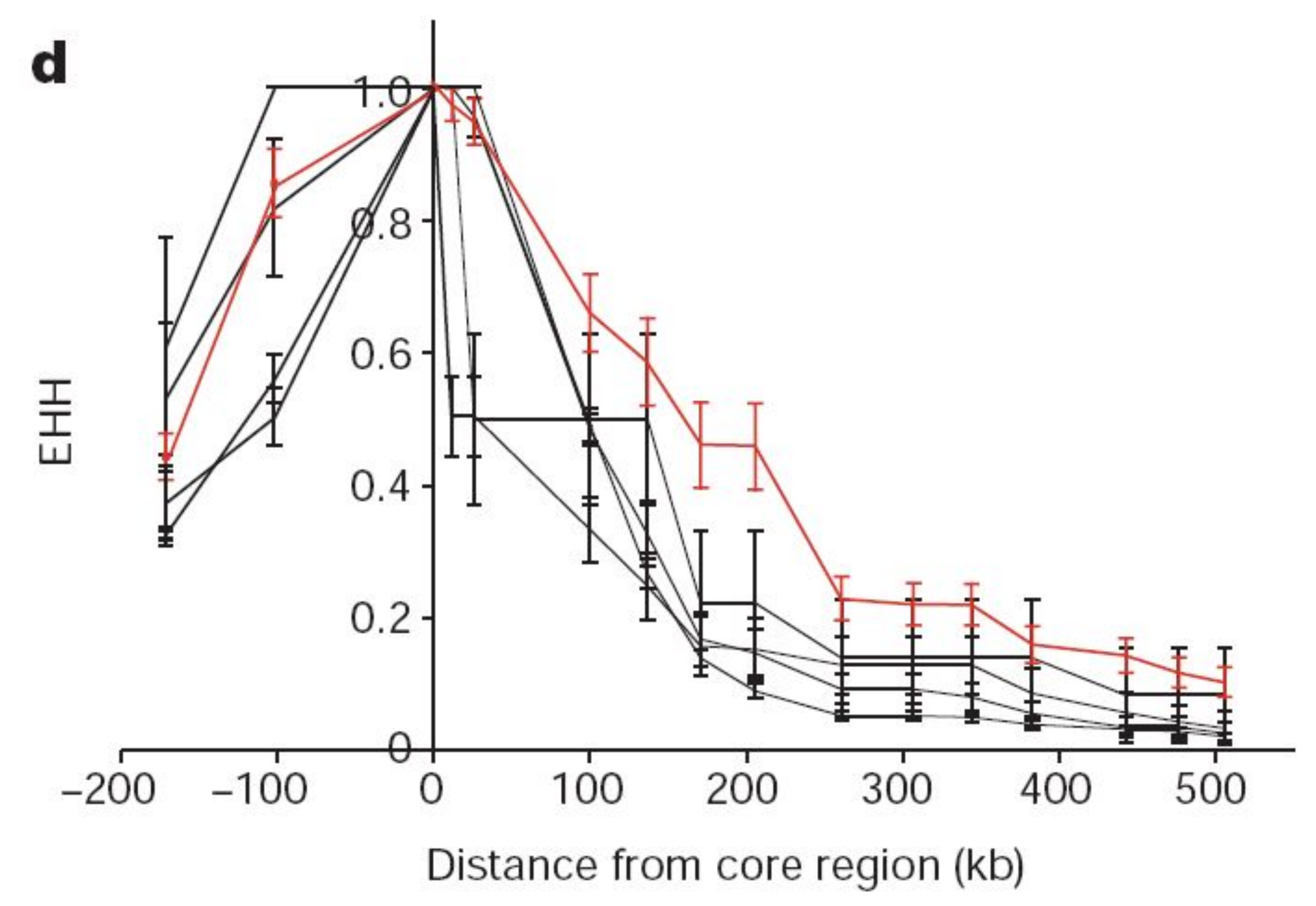


«Выметание» отбором

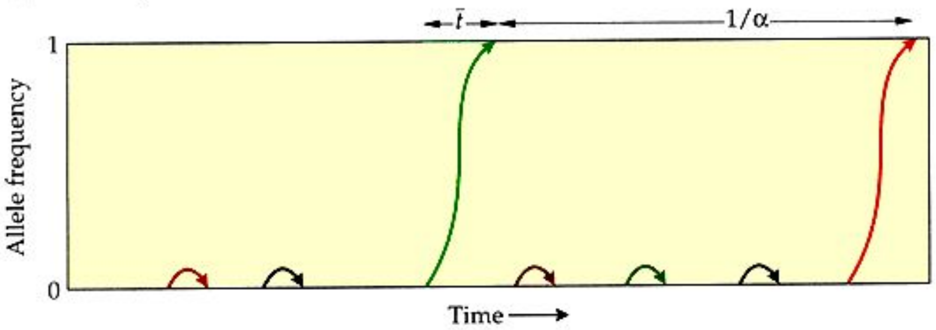


Вокруг сайта под отбором изменчивость понижена

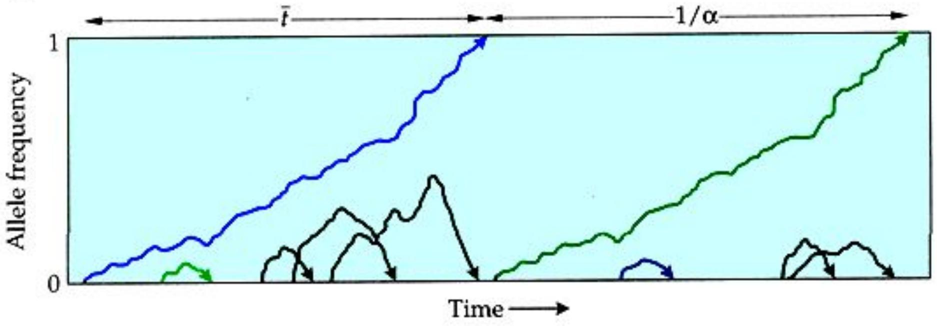
Неравновесие по сцеплению



Sabeti et al., 2002
Nature 419:832-837



(a) Advantageous mutations



(b) Neutral mutations

$$k_{\text{полезная}} \sim 2Nus$$

$$k_{\text{нейтральная}} \sim u$$

$$k_{\text{вредная}} = \frac{2Nus}{e^{2Ns} - 1}$$

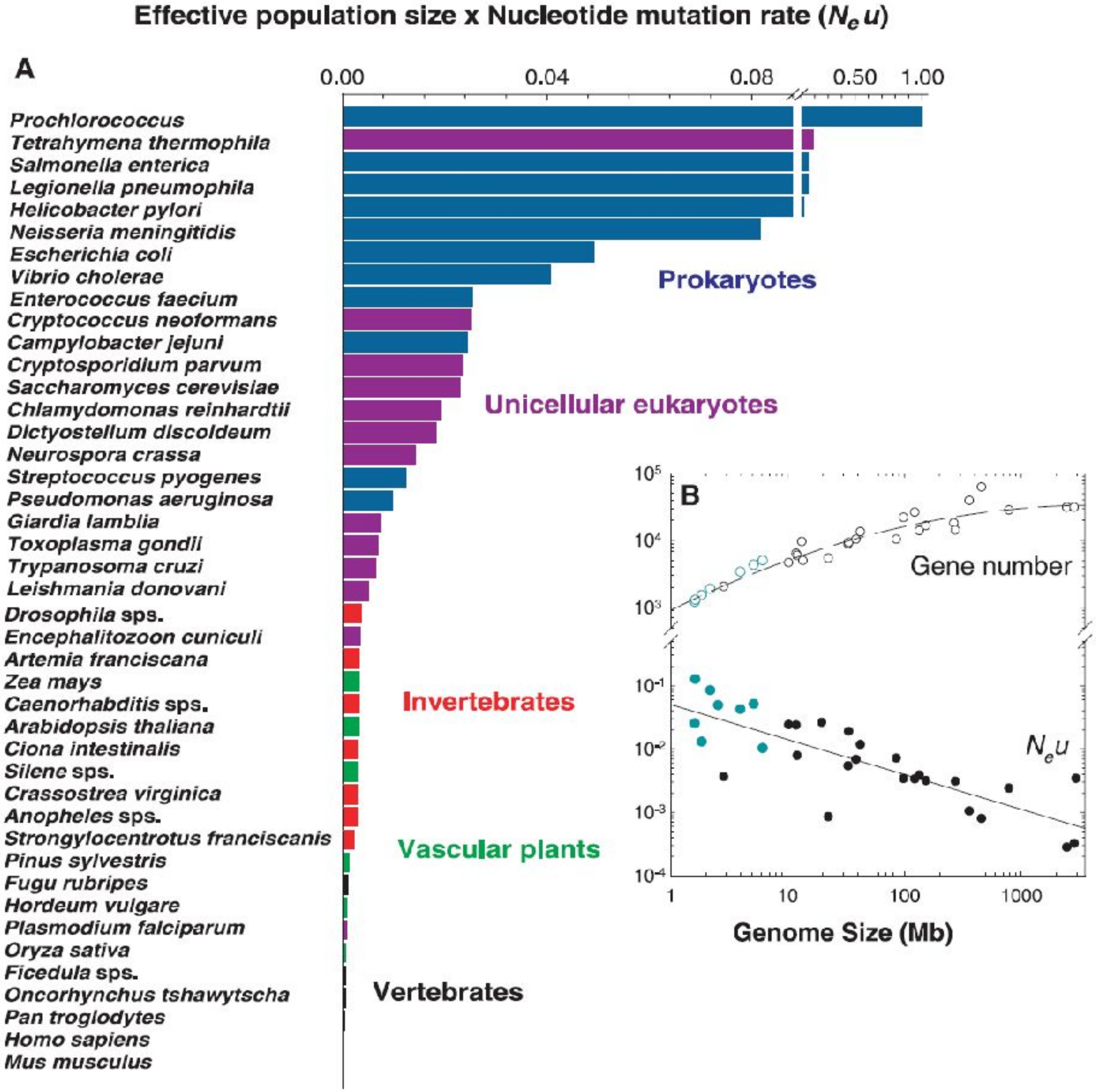


Fig. 1. (**A**) Estimates of the composite parameter $N_e u$ for a phylogenetically diverse assemblage of species. (**B**) The relationship between estimated $N_e u$, total gene number, and genome size. Data for prokaryotes are plotted in blue. The log-log regression of $N_e u$ versus genome size is highly significant, with an intercept of -1.30 ± 0.40, a slope of -0.55 ± 0.07, and $r^2 = 0.659$, df = 28 (*1*). The number of species plotted differs between graphs because genome structure information is not available for all species with $N_e u$ estimates.

Lynch et al. *Science*

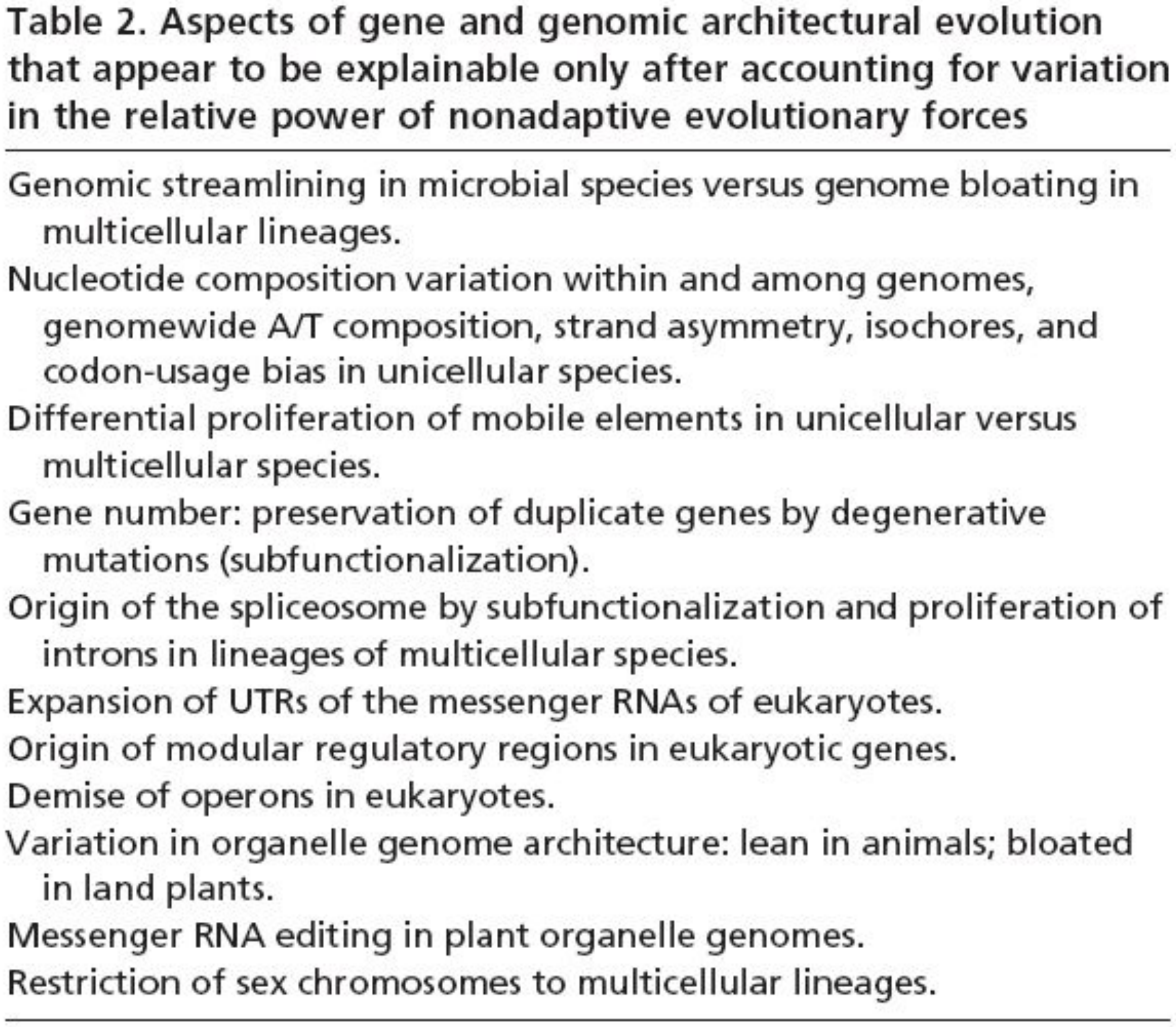


Table 2. Aspects of gene and genomic architectural evolution that appear to be explainable only after accounting for variation in the relative power of nonadaptive evolutionary forces

Genomic streamlining in microbial species versus genome bloating in multicellular lineages.
Nucleotide composition variation within and among genomes, genomewide A/T composition, strand asymmetry, isochores, and codon-usage bias in unicellular species.
Differential proliferation of mobile elements in unicellular versus multicellular species.
Gene number: preservation of duplicate genes by degenerative mutations (subfunctionalization).
Origin of the spliceosome by subfunctionalization and proliferation of introns in lineages of multicellular species.
Expansion of UTRs of the messenger RNAs of eukaryotes.
Origin of modular regulatory regions in eukaryotic genes.
Demise of operons in eukaryotes.
Variation in organelle genome architecture: lean in animals; bloated in land plants.
Messenger RNA editing in plant organelle genomes.
Restriction of sex chromosomes to multicellular lineages.

Lynch et al. *PNAS*