“Must we geneticists become bacteriologists, physiological chemists and physicists, simultaneous with being zoologists and botanists? Let us hope so.”

H. J. Muller (American Naturalist, 1922)

Drift, Mutation Pressure, and the Origins of Complex Cellular Features

- Extension of the mutational-hazard hypothesis: the evolution of complex cellular features by effectively neutral processes.

- The origin of multimeric protein complexes: homomers and heteromers.

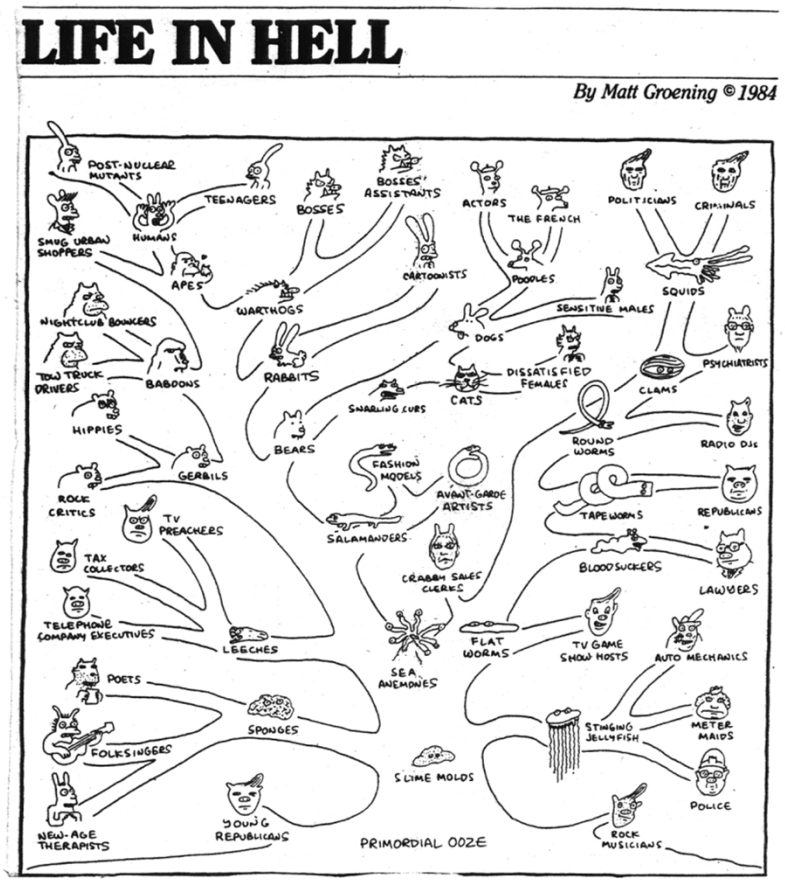
- Extension of the drift-barrier hypothesis: the evolution of layered mechanisms of cellular surveillance, and the myth of the adaptiveness of robustness.

Myth: Natural selection promotes the evolution of organismal complexity.

Reality:

- There is no evidence at any level of biological organization that natural selection encourages complexity. Substantial evidence exists that a reduction in the efficiency of selection promotes the evolution of genomic complexity, whereas an increase in the efficiency of selection can favor simplicity.

- Larger organisms with more complex morphologies have higher historical extinction rates.



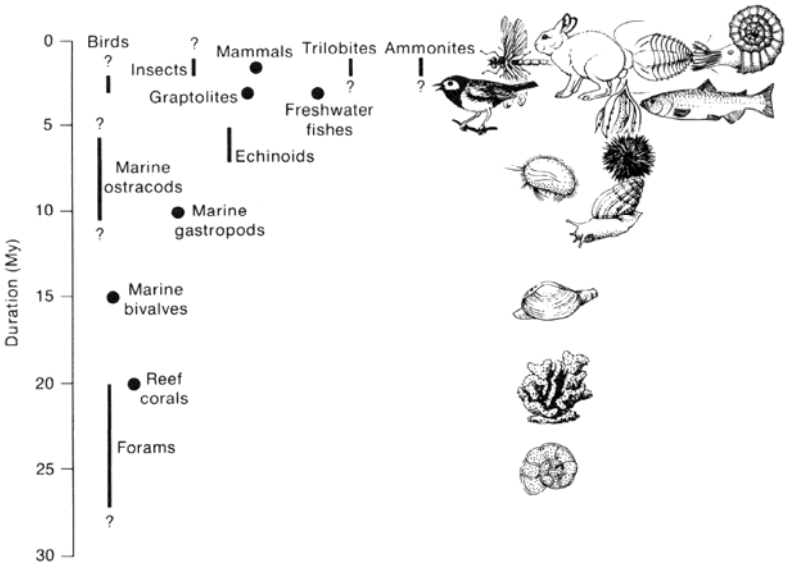


FIGURE 9-2
Hierarchy of average species durations estimated for major taxa. See text for sources of data.

From: Stanley (1985).

Neutral Constructive Evolution: Can Complex Structures Arise by Neutral Processes Rather Than Being Promoted For Their Selective Advantages?

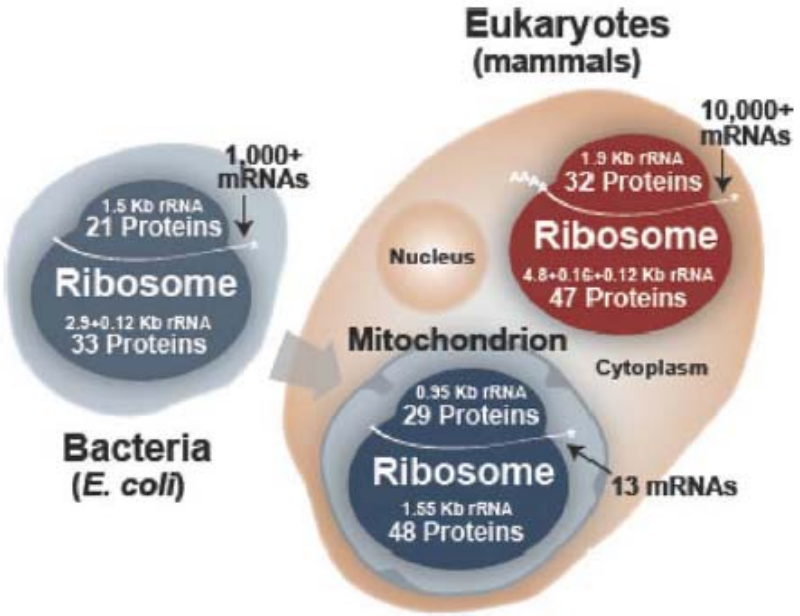


Figure 3. Ribosome complexity in bacteria and eukaryotes. The cartoon on the left summarizes the complexity of the ribosome of *Escherichia coli*, on the right, the human cytoplasmic and mitochondrial ribosomes. In each case, the number of proteins comprising the small and large ribosomal subunits is provided, as is the approximate size and number of ribosomal RNA (rRNA) species and the number of messenger RNAs (mRNAs) translated.

Gain in ribosomal proteins in eukaryotes.

Stoltzfus (J. Mol. Evol., 1999)
Archibald et al. (IUBMB Life, 2011)

Fortuitous interaction with B suppresses deleterious mutational effects in A, making A dependent on B.

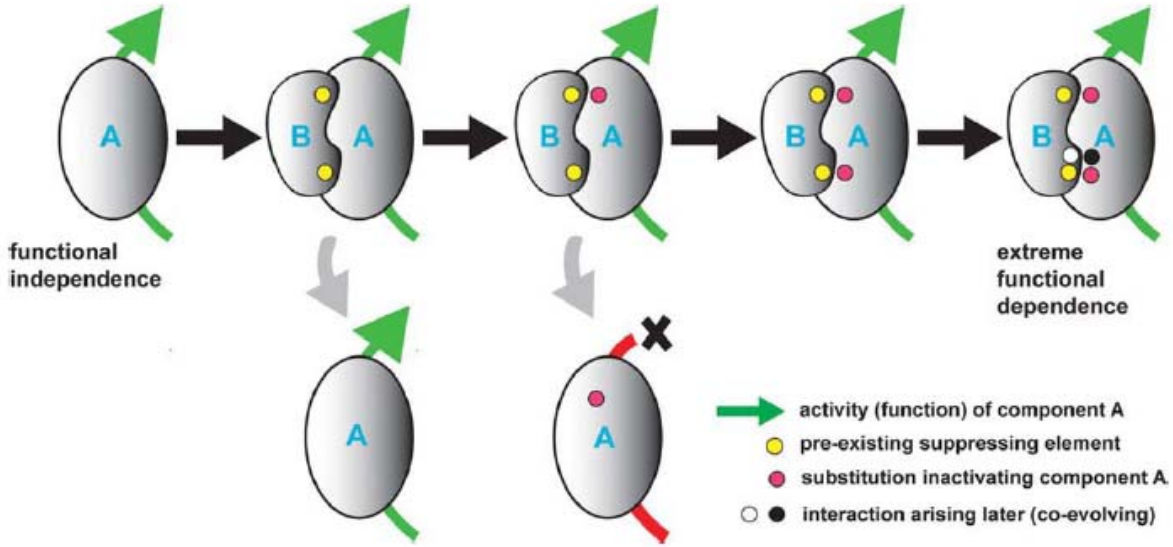


Figure 1. Constructive neutral evolution of biochemical complexity. Schematic depicts (i) a generic enzymatic reaction carried out by cellular component A, (ii) fortuitous (and presuppressing) neutral interactions (yellow dots) with component B, (iii) mutation in A (red dot) that inactivates its activity but that is suppressed by existing interaction with B, (iv) additional mutation in A that is also presuppressed by interaction with B, and (v) coevolving A:B interaction arising later. At stage (ii), A is able to function whether or not B is present and interacting with it, but at stages (ii) and beyond, A is not able to function in the absence of B.

Archibald et al. (IUBMB Life, 2011)

Passive Evolution of Complexity in a Multimeric Structure by Duplication and Complementary Loss of Subfunctions: the process of *subfunctionalization*.

Homo-oligomer

Initial stage of structurally distinct, but functionally equivalent hetero-oligomers

Accumulation of mutations that prevent self-binding leads to alternating subunit types

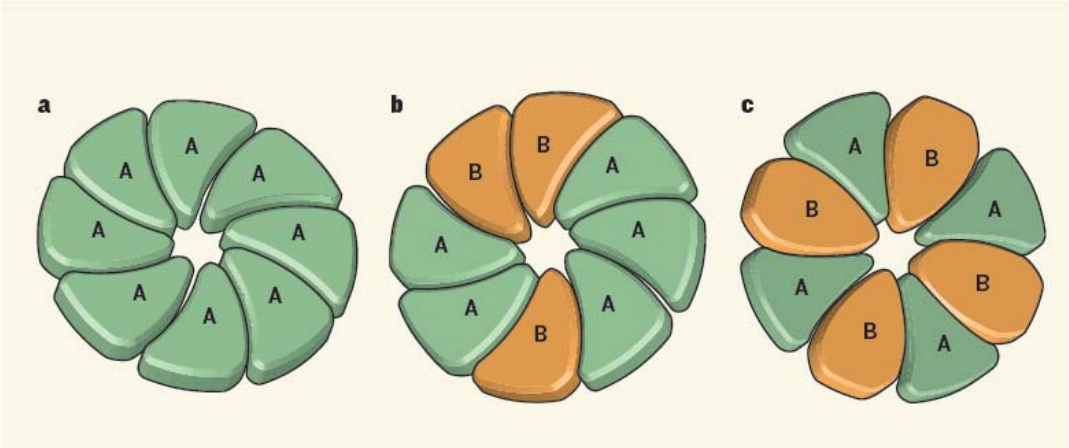
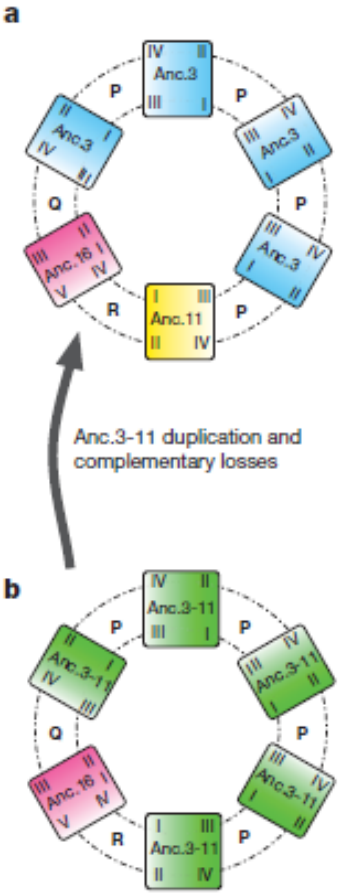


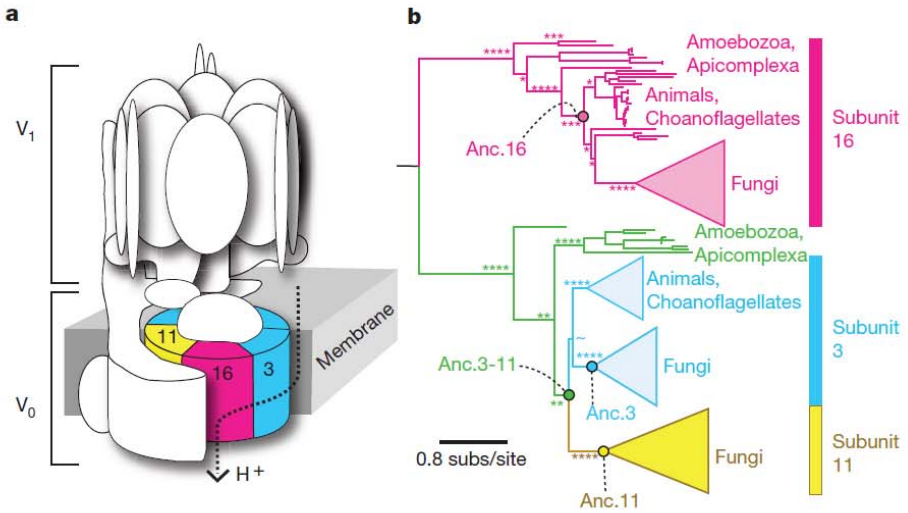
Figure 1 | Evolution of complexity in a protein ring. **a**, Homo-oligomeric protein complexes contain identical subunits (A) that are typically encoded by a single gene. **b**, If the gene is duplicated, the two gene copies can diverge through the accumulation of neutral mutations, generating structurally distinguishable but functionally unaltered subunits A and B, which can form functional hetero-oligomeric rings by random mixture. **c**, If additional mutations prevent subunits from binding to others of their own type, functional rings could still be formed by alternating subunit types. As further mutations accumulate, the probability of returning to the initial homo-oligomeric situation becomes very small. Thus, the subunit composition of a protein association may become complex in the absence of initial selective advantages. Thornton and colleagues[1] provide experimental evidence suggesting that this type of process has occurred in the evolution of a six-membered protein ring that forms part of the vacuolar H^+-ATPase enzymes in fungi.

Addition of a third component



Experimental Demonstration of an Increase in Complexity Via Complementary Degenerative Mutations

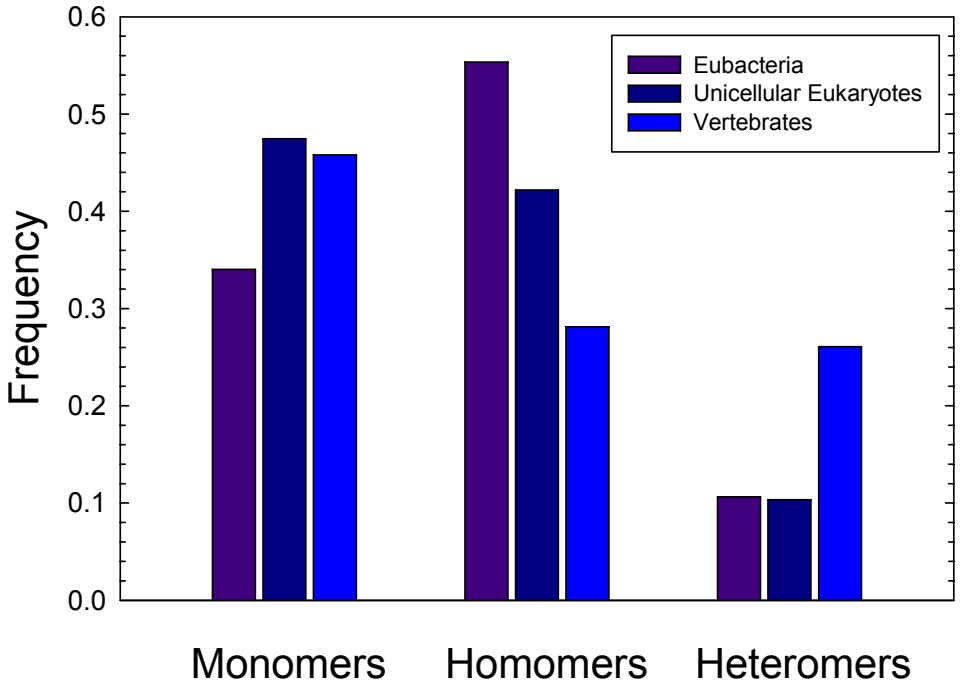
- Vacuolar H^+ ATPase is a multi-subunit complex that pumps protons across membranes to acidify vesicles.
- In most eukaryotes, the membrane ring of vacuolar ATP synthase consists of five subunits of subunit Vma3 and one of its duplicate Vma16.
- In fungi, Vma3 was duplicated, and now an additional subunit Vma11 has been added in such a way that it must lie between Vma3 and Vma16.
- This implies partial degeneration of both Vma3 and Vma11, as the common ancestor must have been able to bind to Vma16 on both sides.
- Experimental reconstruction and replacement of the inferred Vma3/11 ancestral gene implies that the ancestral copy did indeed impose a 5+1 structure in the remainder of today's eukaryotes.
- No evidence that the addition of the Vma11 subunit led to novel functions or increased efficiency in the fungi complex.

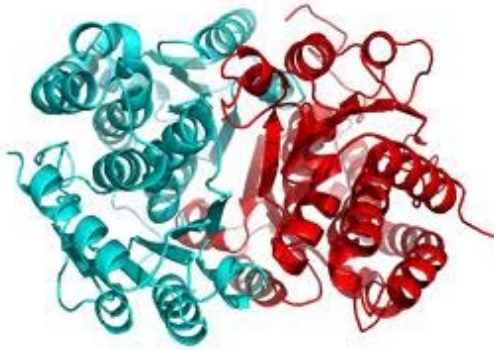


Finnigan et al. (Nature, 2012)
Doolittle (Nature, 2012)

Figure 1 | Structure and evolution of the V-ATPase complex. **a**, In *S. cerevisiae*, the V-ATPase contains two subcomplexes: the octameric V_1 domain is on the cytosolic side of the organelle membrane, and the hexameric V_0 ring is membrane bound. Protein subunits Vma3, Vma11 and Vma16 are labelled and coloured. **b**, Maximum likelihood phylogeny of V-ATPase subunits Vma3, Vma11 and Vma16. All eukaryotes contain subunits 3 and 16, but Fungi also contain subunit 11. Circles show ancestral proteins reconstructed in this study. Colours correspond to those of subunits in panel **a**; unduplicated orthologues of Vma3 and Vma11 are green. Asterisks show approximate likelihood ratios for major nodes: ****, $>10^3$; ***, $>10^2$; **, >10; *, <10; ~, <2. The complete phylogeny is presented in Supplementary Information, section 2.

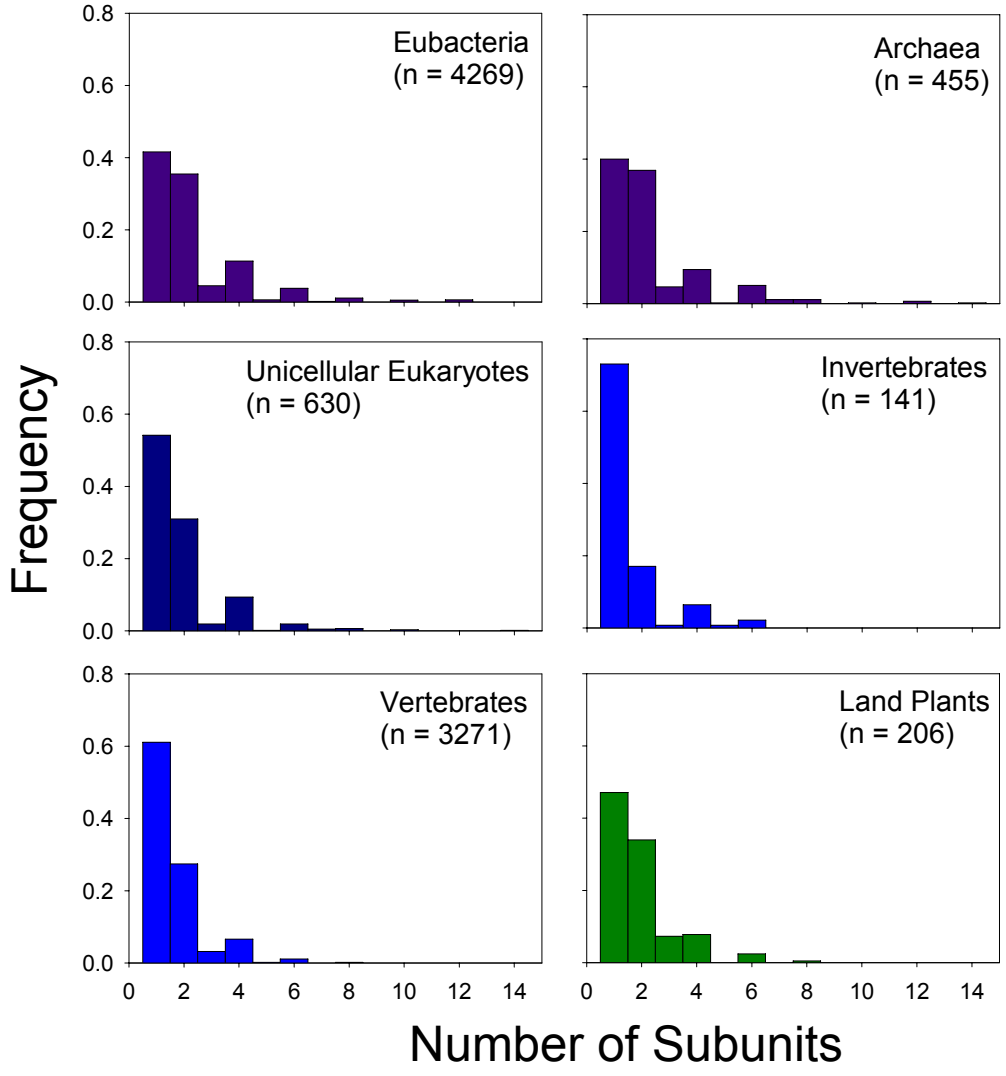
Distribution of Complex Types

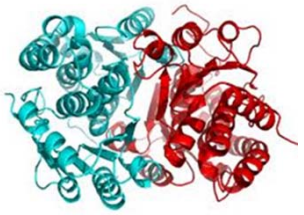




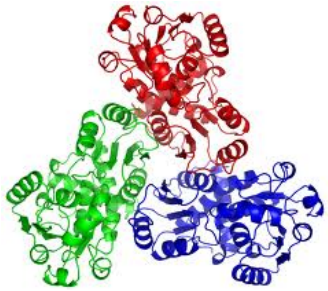
- ~60% of proteins with known structures exist as dimers or higher-order complexes.
- Monomers are slightly more common in eukaryotes.
- For complexes, homomers are ~4x more frequent than heteromers in unicellular species, but equally frequent in vertebrates.

Distribution of Homomeric Types

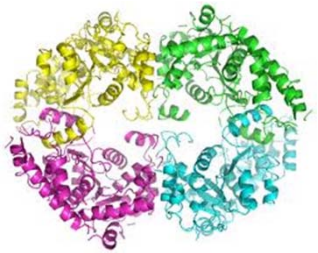




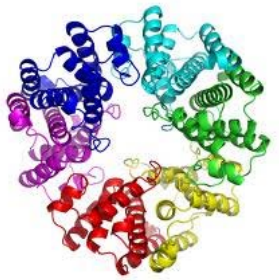
dimer



trimer



tetramer

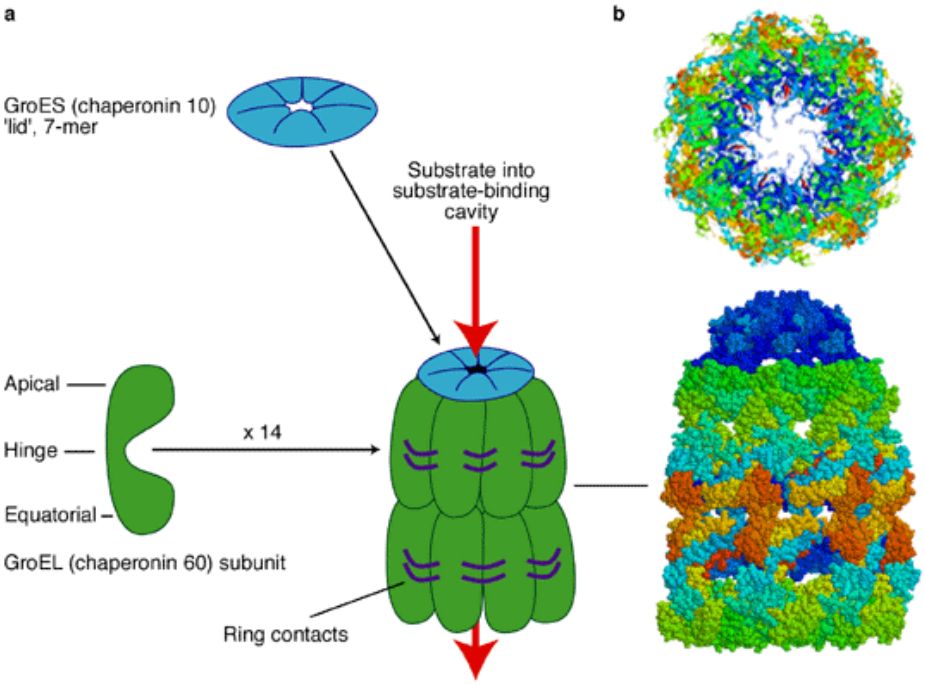


hexamer

From: 3D Complex.org (Levy et al. 2008)

The Eubacterial Chaperonin Complex

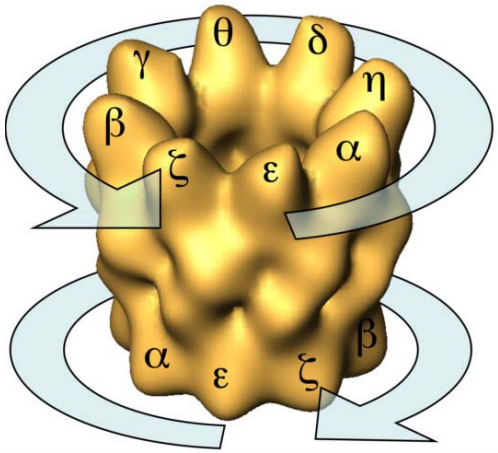
- Most Eubacteria have a **GroEL chaperonin** system – a homo-tetradecamer (14 subunits, all the same), oligomerized into cages. With the help of GroES, unfolded proteins are moved into the interior for proper folding.
- Mitochondria and chloroplasts also have GroEL.
- Several *Mycoplasma* and *Ureaplasma* species (intracellular pathogens) have lost chaperonins entirely.



Structure of chaperonin 10 and chaperonin 60

The Eukaryotic Chaperonin Complex

- CCT (chaperonin containing tailless complex) proteins are hetero-hexadecamers (16 proteins in the total barrel; sometimes 18) restricted to Eukaryotes and Archaea.
- The eight components diverged following a series of ancient gene duplications prior to LECA (Archibald et al. 2000).
- Each component is thought to have a specialized binding function, and sites known to be involved in binding seem to be under positive selection in Eukaryotes (Fares and Wolfe 2003).
- In Archaea, the subunits do not appear to have become specialized, and there are only two to three types (as opposed to eight in Eukaryotes).
- A classical case of a homomer becoming a heteromer through duplication, degeneration, and complementation.



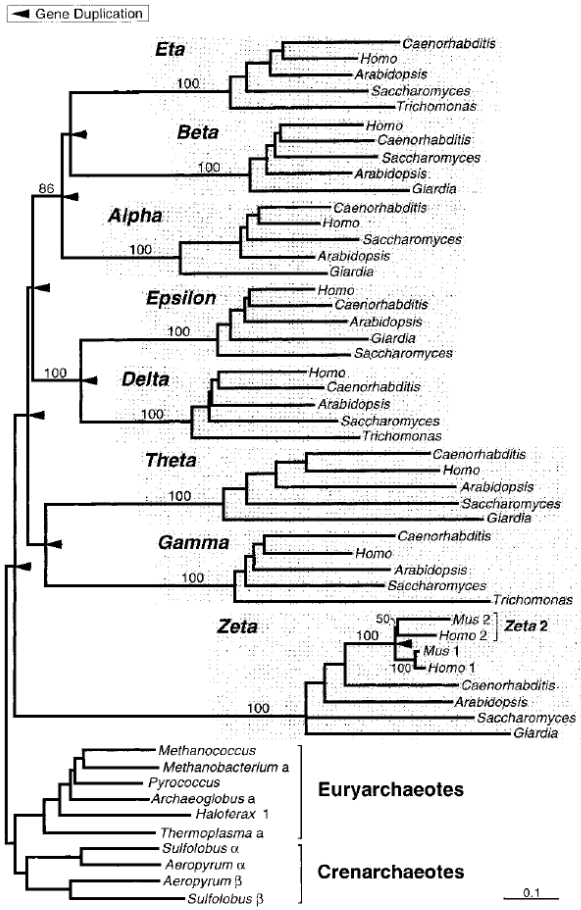


FIG. 2. Phylogeny of eukaryotic CCTs. The tree shown was constructed using the Fitch–Margoliash distance algorithm (Felsenstein, 1995) from an alignment of 53 sequences and 355 unambiguously aligned amino acid sites. The eight different CCT subunit families found in eukaryotes are highlighted in gray. For each CCT subunit family, 5 representative sequences were chosen to represent the full spectrum of eukaryotic diversity: two animals (*Homo* and *Caenorhabditis*) and one from each of fungi (*Saccharomyces*), plants (*Arabidopsis*), and protists (*Giardia* or *Trichomonas*). For CCTζ, the ζ-1 and ζ-2 subunits of *Homo* and *Mus* were also analyzed. The tree is rooted with a representative sample of archaeal chaperonin sequences. Support values for important nodes on the tree are given above the branches and were calculated by bootstrapping with 100 resampling replicates. The scale bar represents the estimated number of substitutions per amino acid site.

Gene Duplication and Chaperonin Evolution

- Eubacterial chaperonins have a single subunit; Archaea have 1 to 3; and Eukaryotes have 8.
- Parallel duplications leading to heteromeric structures have occurred in the Archaea, and reversions to homomers have also occurred.
- Eubacterial rings have seven members, while those in Archaea and Eukaryotes have eight or nine.

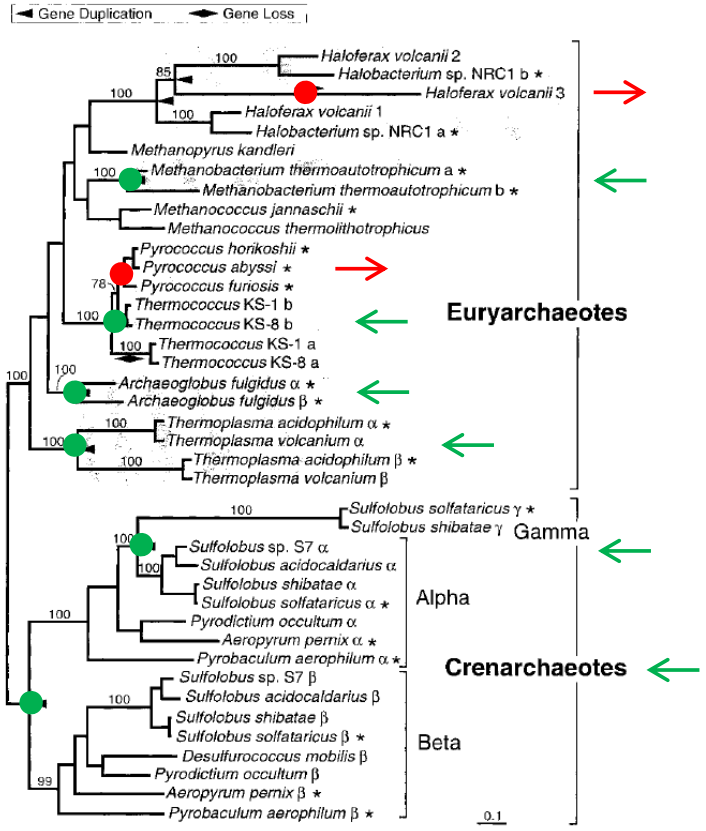


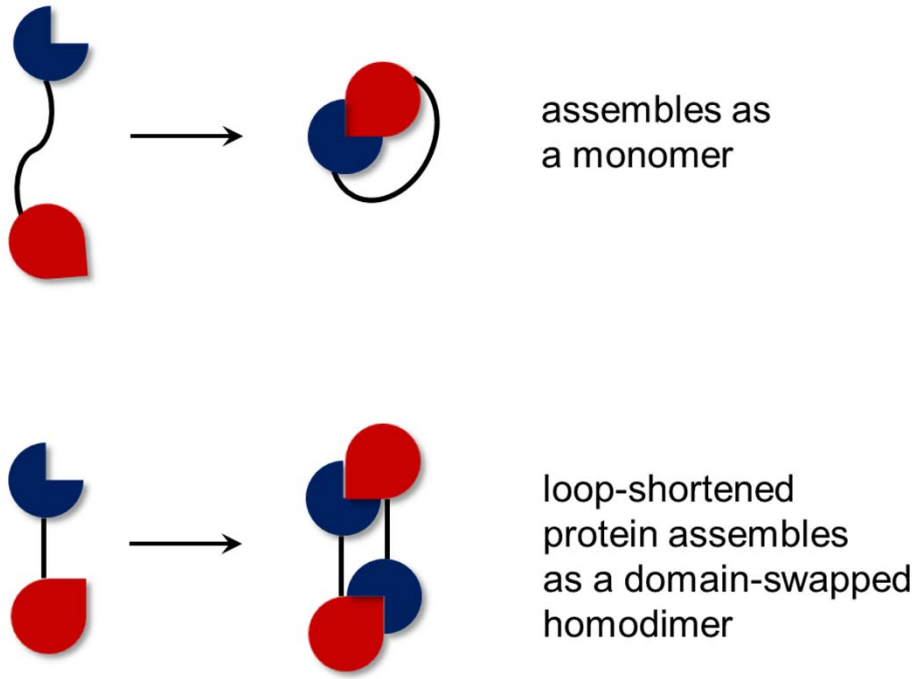
FIG. 1. Phylogenetic analysis of archaeal chaperonins. The tree shown is a ML tree (lnL −15614.72) inferred from a chaperonin protein sequence alignment containing 40 sequences and 452 unambiguously aligned amino acid positions. The two recognized kingdoms within Archaea (euryarchaeotes and crenarchaeotes) are labeled, and inferred gene duplications and gene losses are indicated (see text). Within euryarchaeotes, regions of the tree in which lineage-specific gene duplications have occurred are shaded. For crenarchaeotes, the three different gene/subunit families (α, β, and γ) are indicated. Asterisks appear next to sequences from organisms whose genomes have been completely sequenced. Statistical support values for significant nodes appear above the branches (ML RELL bootstrap values; inferred from a heuristic search of 1000 trees in protML) (Adachi and Hasegawa, 1996) (see text). The scale bar represents the estimated number of amino acid substitutions per site.

Archibald et al. (2001, J. Struct. Biol.)

The origin of protein interactions and allostery in colocalization

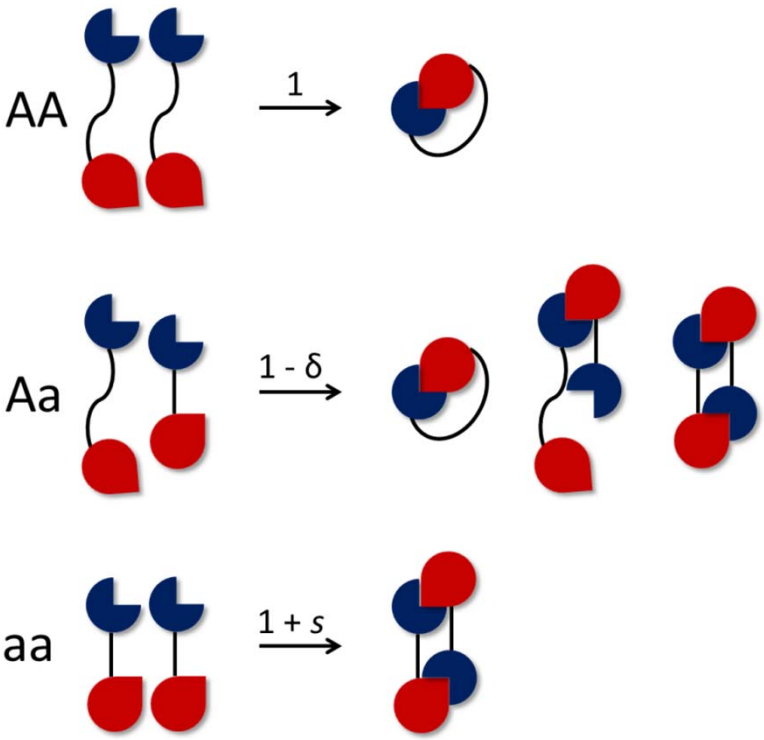
John Kuriyan[1,2] & David Eisenberg[3]

NATURE|Vol 450|13 December 2007|doi:10.1038/nature06524



The Domain-swapping Model

The Population-genetic Conditions for the Origin of Domain Swapping



Pure Monomer

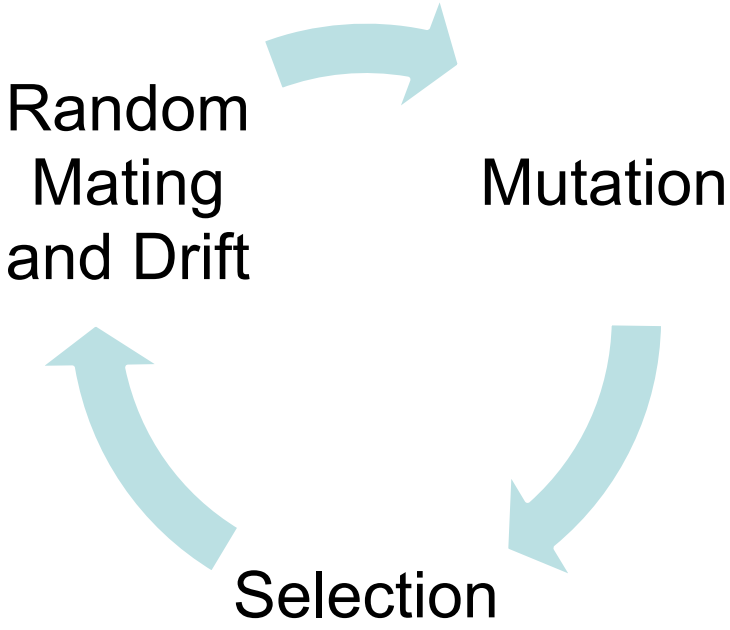
Heterozygote with mixture of structures

Pure Homodimer

- Advantages: preadapted to complexation, and only requires a single mutation.

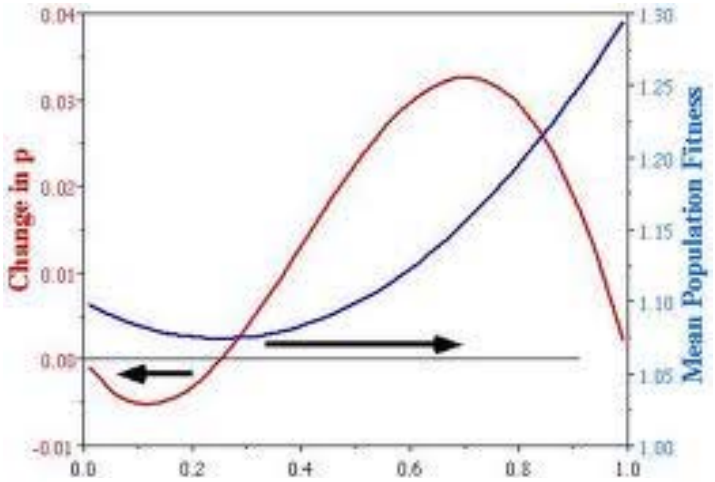
- Disadvantage: reduced heterozygote fitness may impose a strong barrier to fixation;
  the homozygote might also be weakly disadvantageous due to the diffusion barrier to assembly.

With Recurrent Mutational Introduction of the Domain-swapping Allele, How Long Does It Take To Establish (Go To Fixation) In Populations?



- Function of the population size, the mutation rate, and the fitness effects of the domain-swapping allele in heterozygotes and homozygotes

  = rate of input (2Nu) x probability of fixation.



Frequency of domain-swapping allele

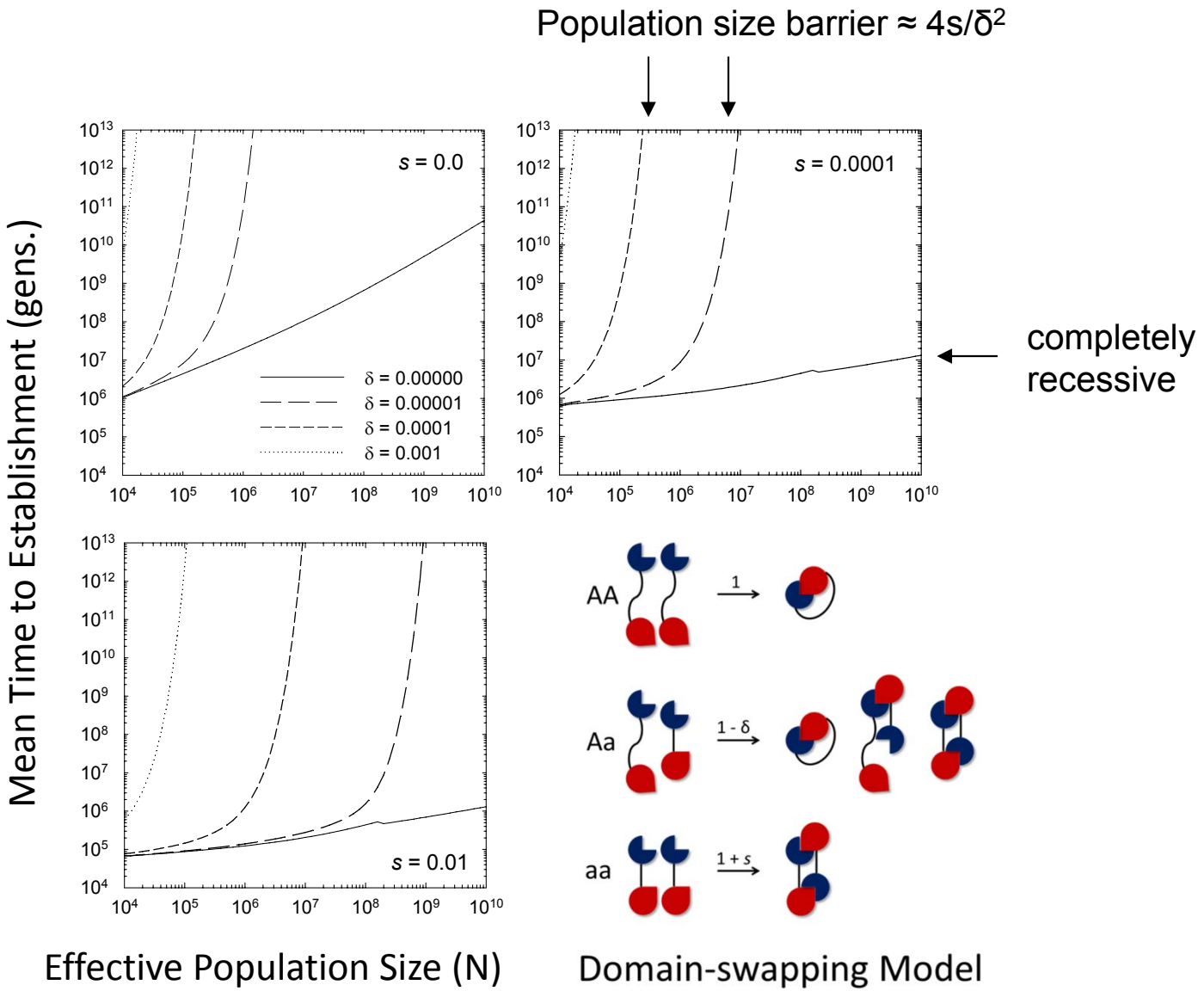
Probability of fixation of an underdominant mutation =

$$\frac{\mathrm{erf}\{[p_0-(0.5/(1+\omega))]\sqrt{4\theta(1+\omega)}\}+\mathrm{erf}\{\sqrt{\theta/(1+\omega)}\}}{\mathrm{erf}\{[1-(0.5/(1+\omega))]\sqrt{4\theta(1+\omega)}\}+\mathrm{erf}\{\sqrt{\theta/(1+\omega)}\}},$$

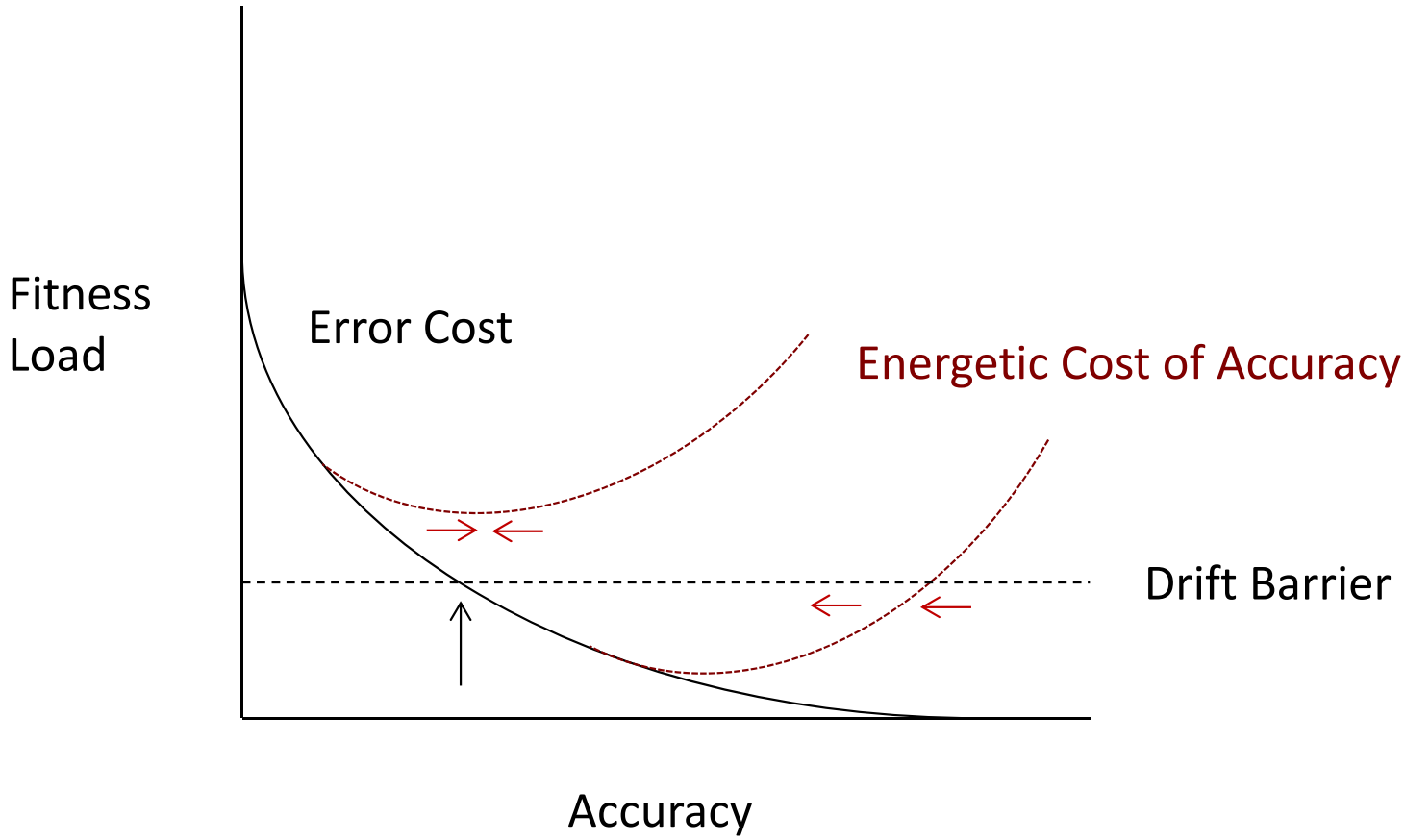
where $\theta = Ns$, $\omega = s/(2\delta)$ (Walsh 1982), and

$$\mathrm{erf}(x) = \int_0^x e^{-y^2}\,dy,$$

Evolution of Domain-swapping Homodimers is Strongly Inhibited in Large Populations, Unless the Heterozygote Disadvantage is Extremely Weak



The Limits to Molecular Perfection



The Evolution of Neutrality for the Efficiency of an Enzymatic Function: an inevitable outcome of natural selection (Hartl et al., Genetics, 1985).

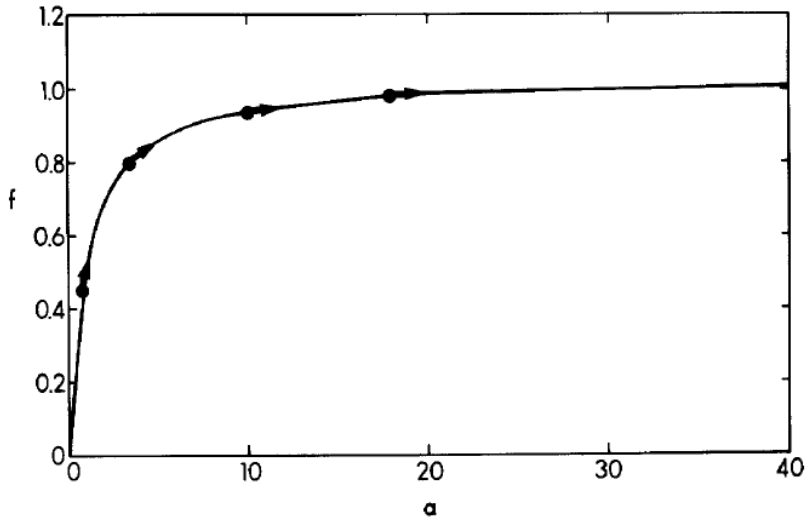


FIGURE 1.—Standardized Michaelis-Menten saturation equation $f(a) = a/(1 + a)$ scaled to equal at $a = 30$.

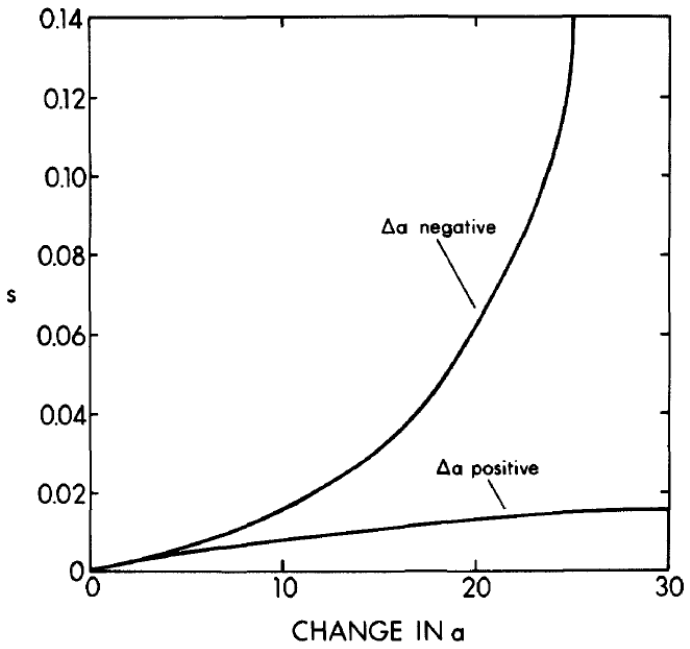
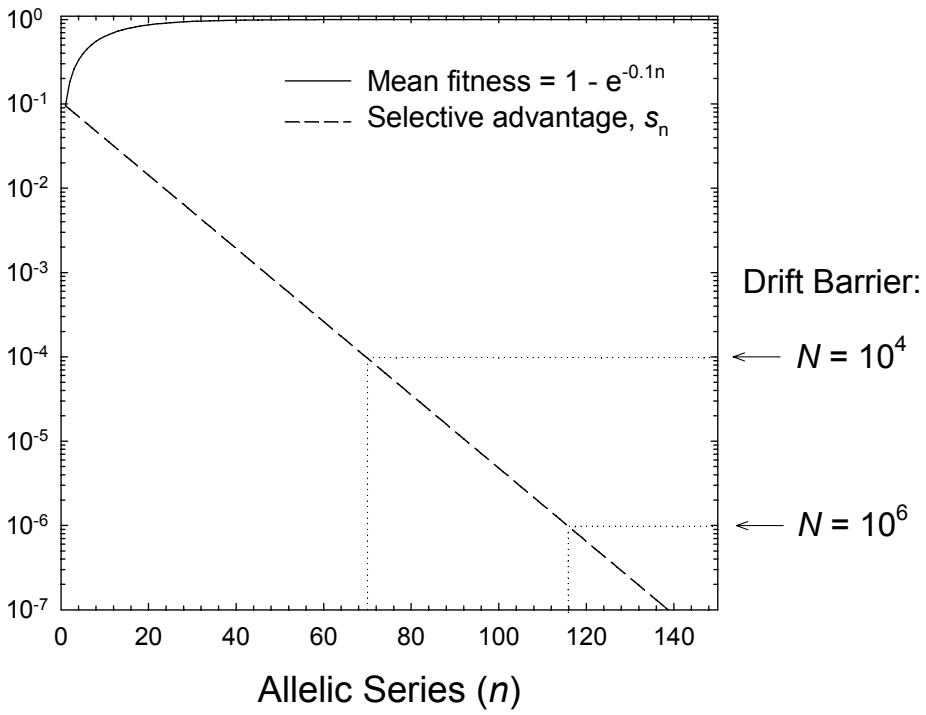
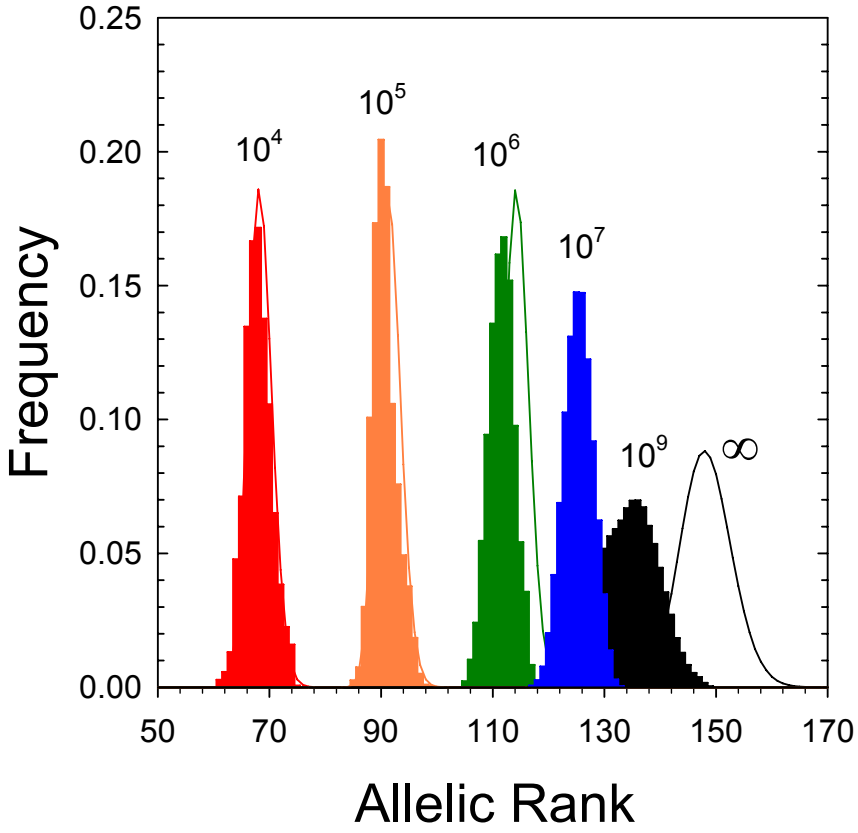


FIGURE 2.—Selection coefficient (s) resulting from a given positive or negative change in enzyme activity (Δa), when initially $a = 30$.

Asymptotically Increasing Perfection in an Allelic Series



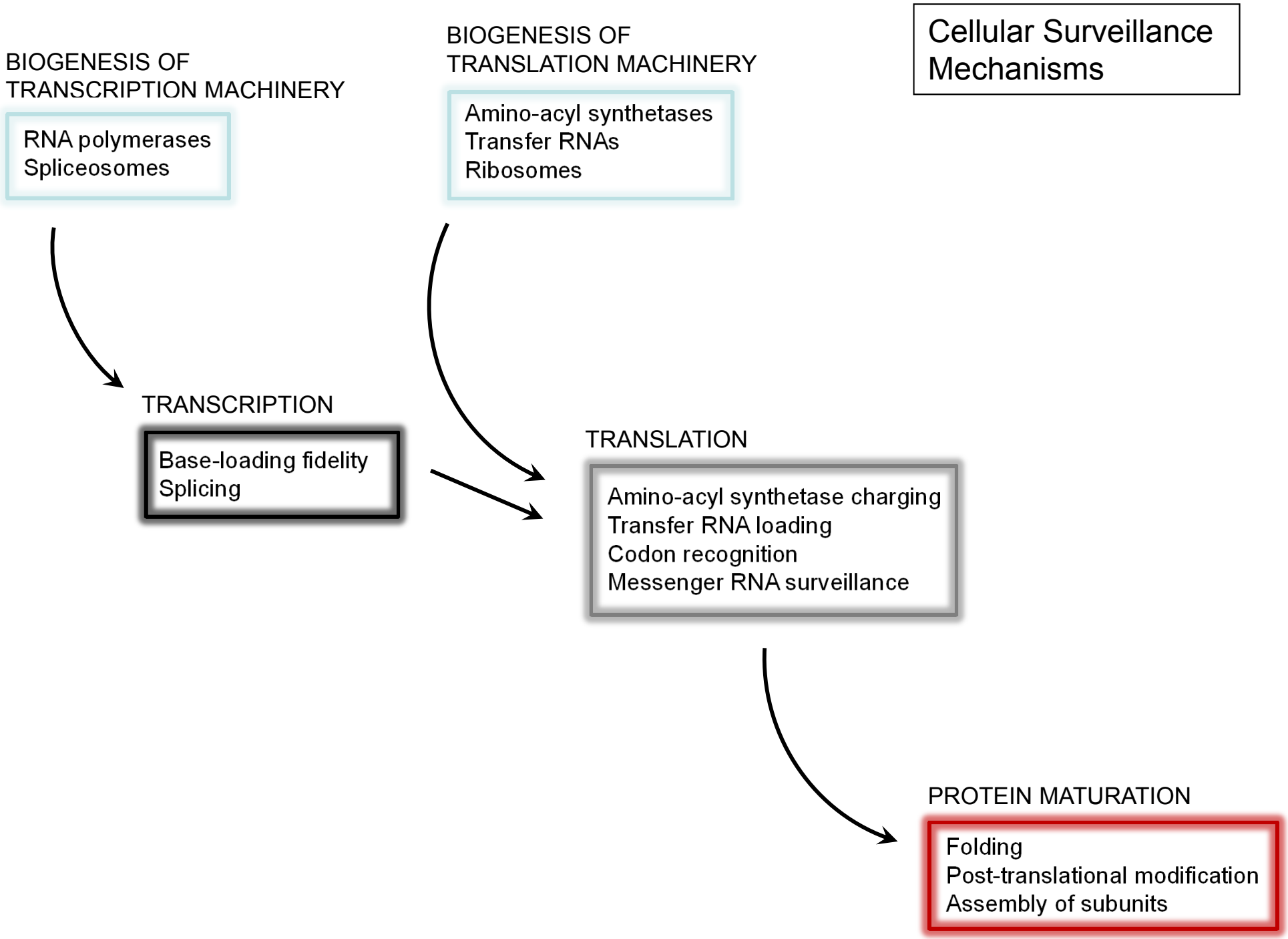
Equilibrium Frequencies of Alleles with Increasing Population Sizes



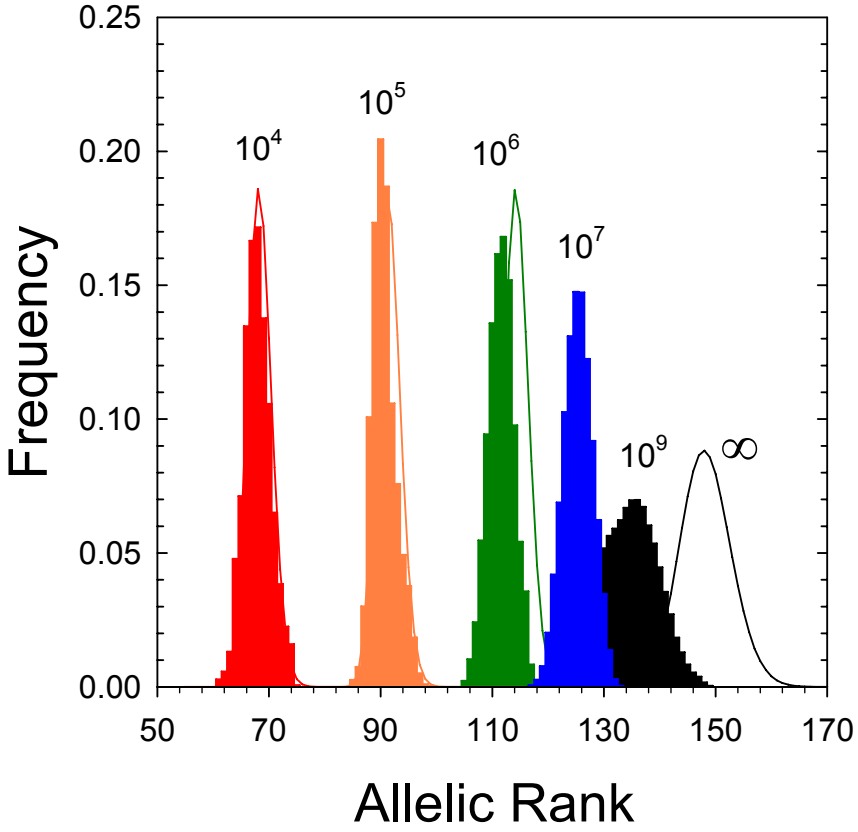
Evolutionary Layering and the Limits to Molecular Perfection:

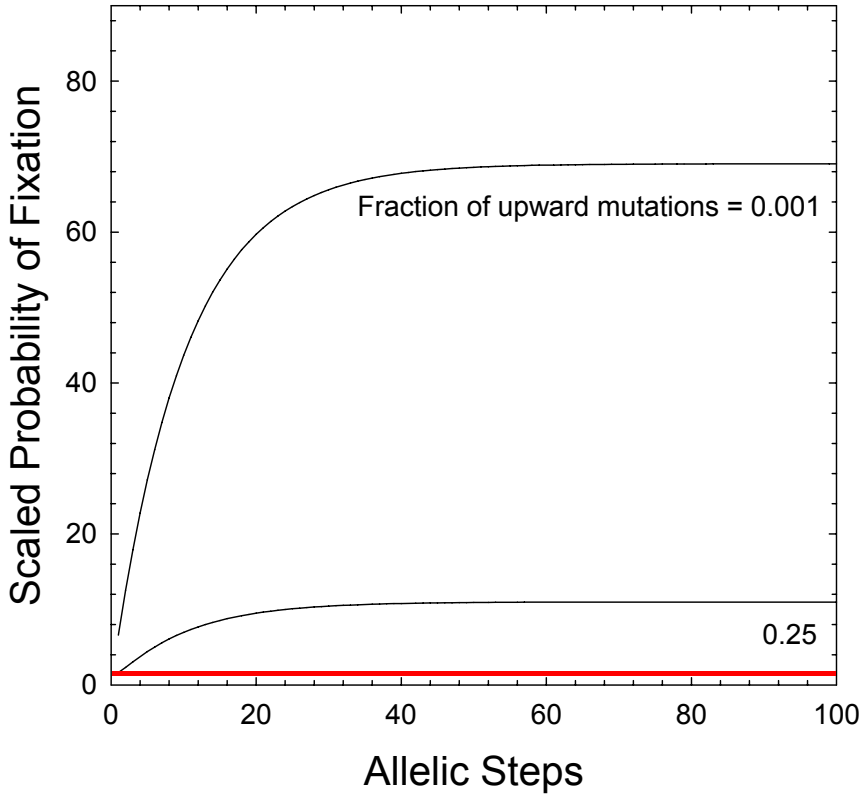
1) Can a secondary layer of defense be added that breaks the drift barrier?

2) If such a genomic addition is assimilated, what are the long-term consequences for the refinements of the previous layer, the new layer, and the combined effects of both?

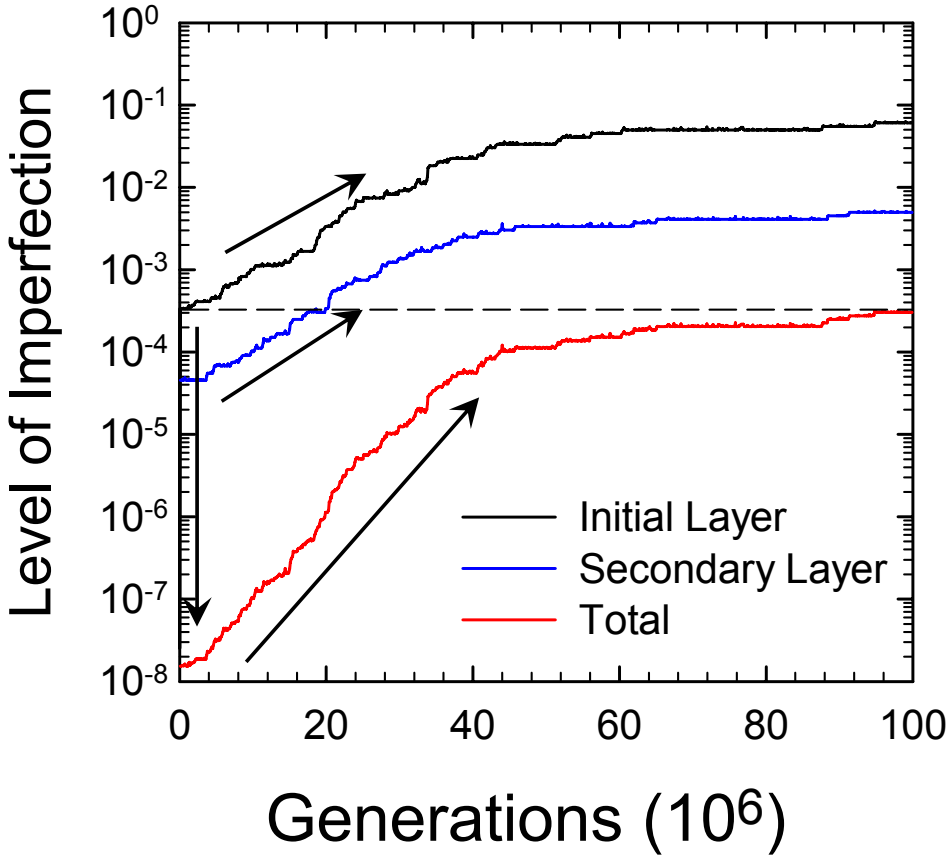

Cellular Surveillance Mechanisms
BIOGENESIS OF TRANSCRIPTION MACHINERY
RNA polymerases
Spliceosomes
BIOGENESIS OF TRANSLATION MACHINERY
Amino-acyl synthetases
Transfer RNAs
Ribosomes
TRANSCRIPTION
Base-loading fidelity
Splicing
TRANSLATION
Amino-acyl synthetase charging
Transfer RNA loading
Codon recognition
Messenger RNA surveillance
PROTEIN MATURATION
Folding
Post-translational modification
Assembly of subunits

Scaled Probability of Establishment of a Secondary Layer of Surveillance
(relative to the neutral expectation)





The Response to the Addition of a Layer of Accuracy Is Transient



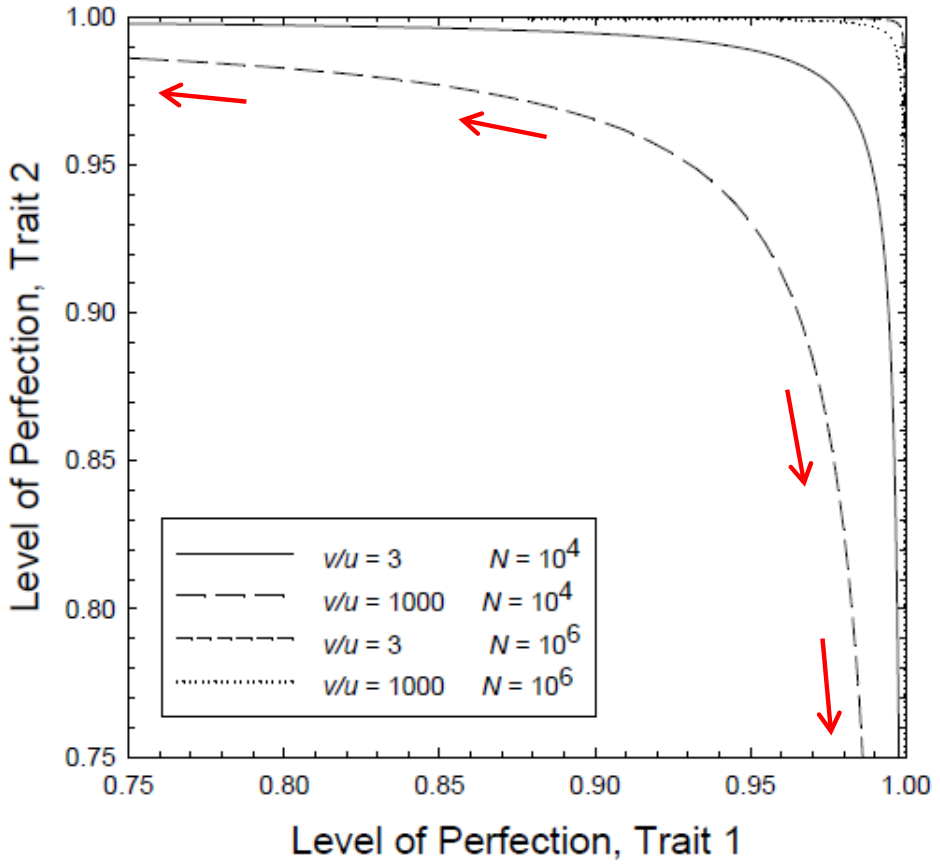
- Rapid improvement accompanies establishment of a new layer of protection.
- Both layers then gradually become less efficient.
- The level of overall performance returns to that for the single-layered state.

- The “Paradox of Robustness” (S. Frank, PLoS One): a more complex system evolves, but nothing is gained in the long run.
- Something has been lost: sensitivity of the system to mutational breakdown has increased.

A Bivariate Drift Barrier:

• Selection will operate to drive the joint effects of two traits down to the limits imposed by drift.

• There is a ridge along which the population can freely drift, even to the extent of losing one trait.



Limiting Rate of Evolution of Simple Adaptations: single-site changes with additive effects

N = actual population size
N_e = effective population size
s = selective advantage
u = mutation rate to beneficial allele

- Number of new mutations entering the population per generation = $2Nu$

- As $N_e \rightarrow \infty$, the probability of fixation $\rightarrow 2sN_e / N$

  As $N_e \rightarrow 0$, the probability of fixation $\rightarrow 1 / (2N)$

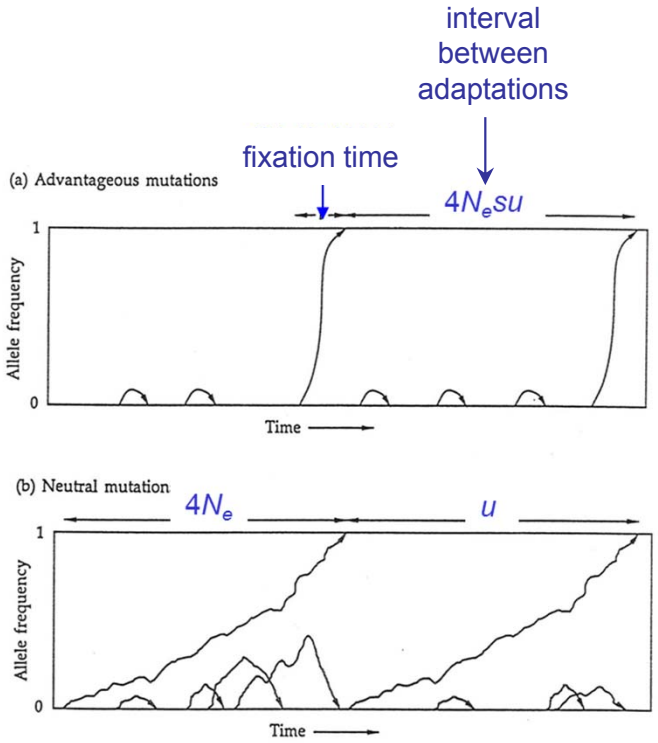


Figure 6. Dynamics of gene substitution for (a) advantageous and (b) neutral mutations. Advantageous mutations are either quickly lost from the population or quickly fixed, so that their contribution to genetic polymorphism is small. The frequency of neutral alleles, on the other hand, changes very slowly by comparison, so a large amount of transient polymorphism is generated. $\bar{t}$ is the conditional fixation time and $1/\alpha$ is the mean time between consecutive fixation events. From Nei (1987).

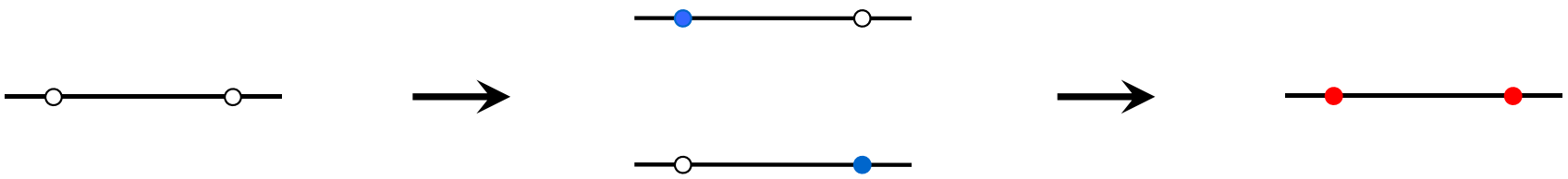
- Selection prevails if $4N_es >> 1$

- Drift prevails if $4N_es << 1$

- Long-term rate of adaptive evolution in large populations = $4N_esu$

If most molecular evolution reflected adaptive mutations, larger populations would evolve more rapidly.

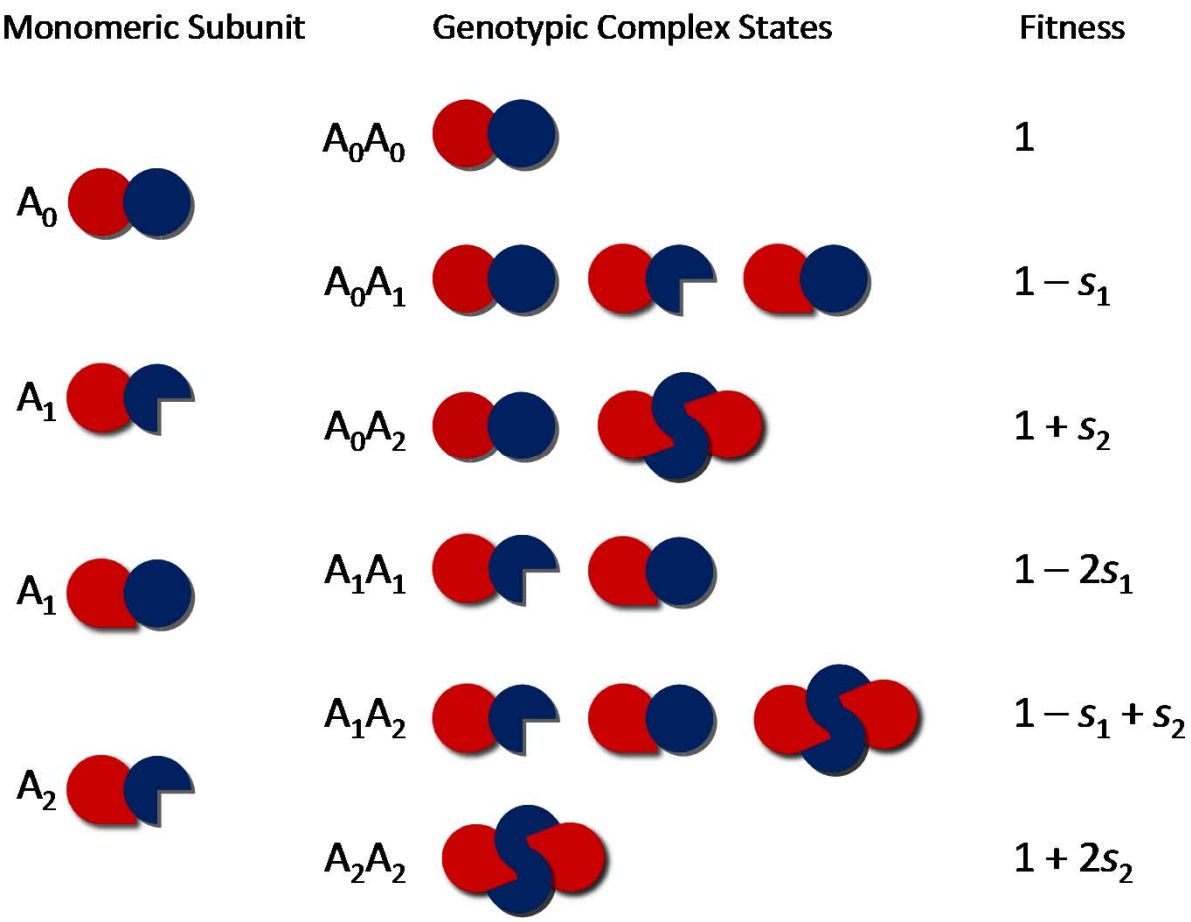
Evolution of a Complex Adaptation Through Neutral / Deleterious Intermediates



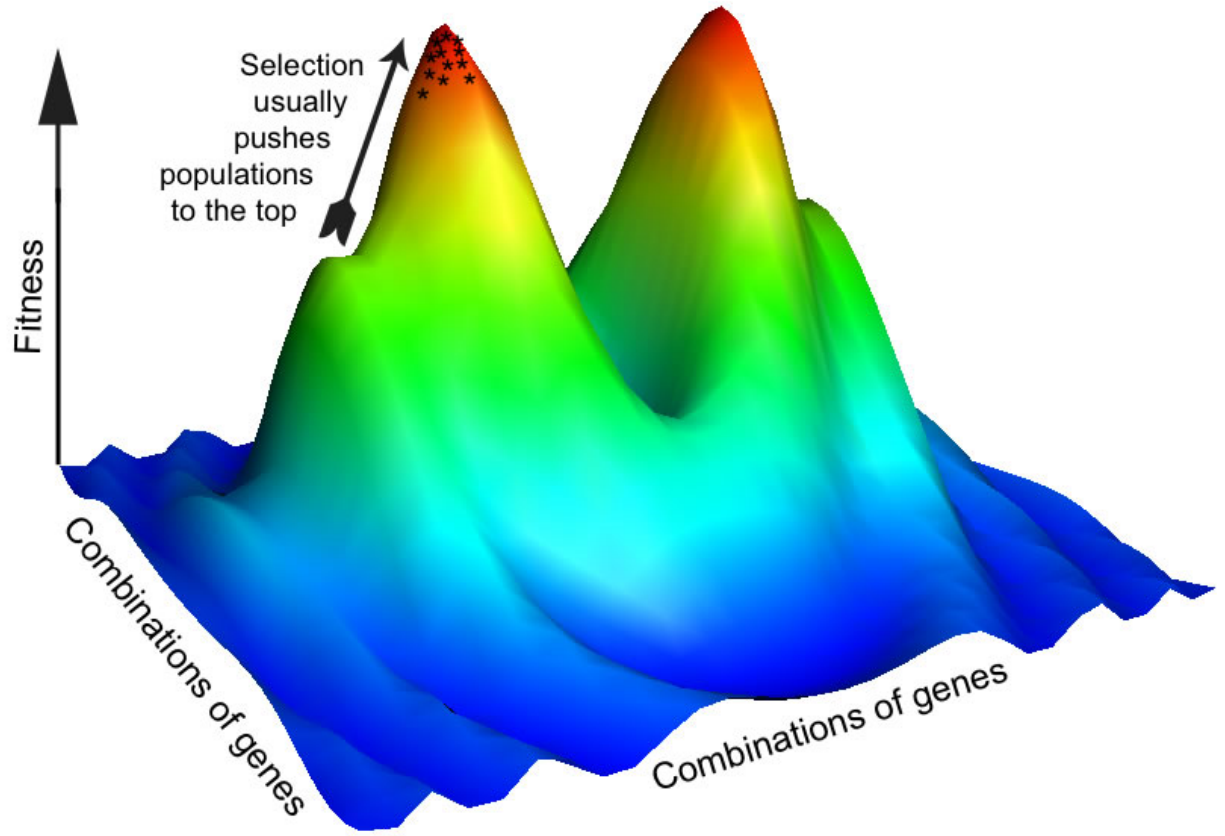
Fitness: 1 $1 - s_1$ $1 + s_2$

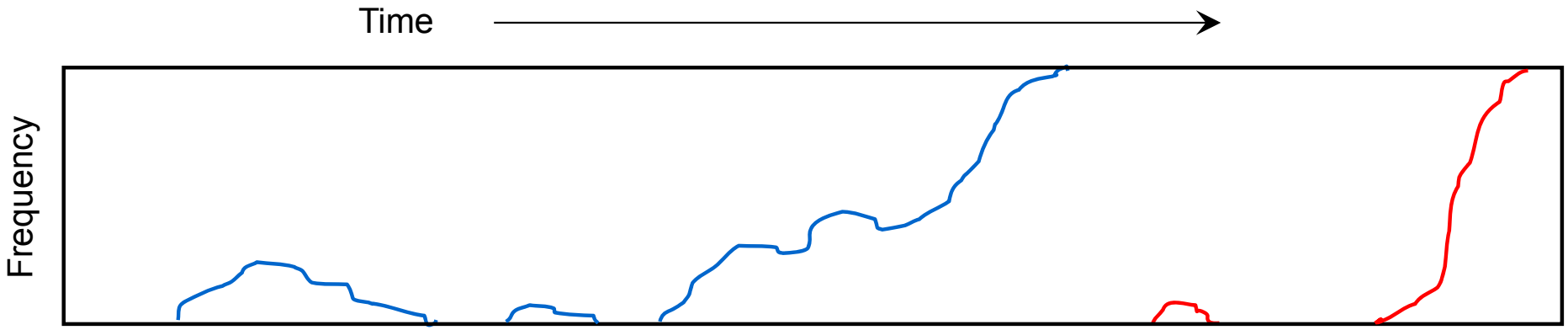
A Two-site Model

Evolution of a Homodimer by Compensatory Mutation

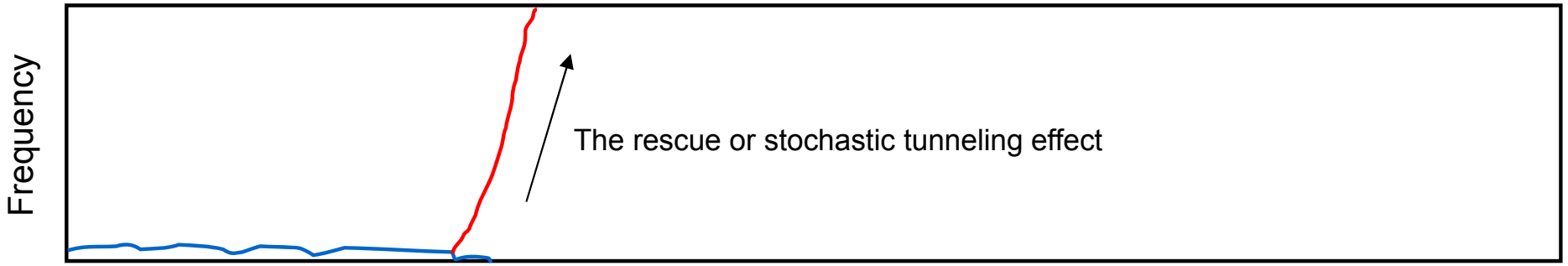


A common view: selection cannot take a population from one adaptive peak to another, unless the population size is small enough to allow maladaptive drift across the fitness valley.



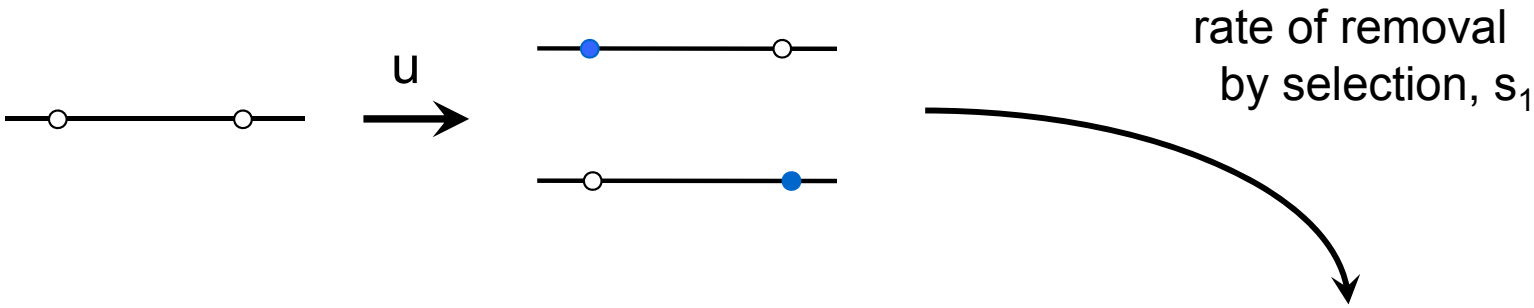


- Small population sizes – under the sequential model, adaptation proceeds in a stepwise fashion, which can necessitate a sojourn through a mean-population fitness bottleneck.



- "Large" population sizes – intermediate deleterious alleles need never be fixed, but are kept at low frequencies by selection-mutation balance, serving as launching pads for the final adaptation.

Maintenance of an Intermediate-step Deleterious Allele by Selection/Mutation Balance



- Equilibrium frequency = u / s_1
- Equilibrium number of copies = $2Nu / s_1$
- Half-life of a newly arisen deleterious allele = $1 / s_1$

When does the origin-fixation scheme break down?

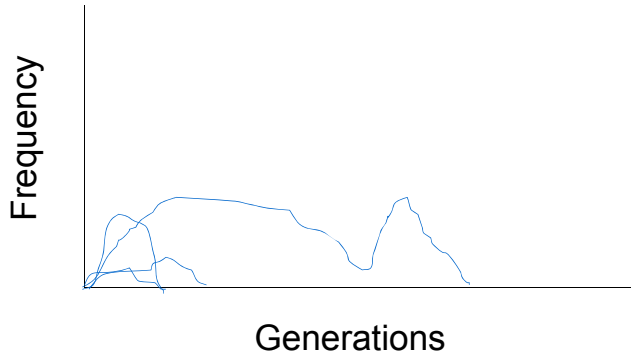
Assuming first-step neutrality, mean time to fix the first (neutral) mutation = 4N generations.

Rate of origin of second-step mutations during this period < 2Nu.

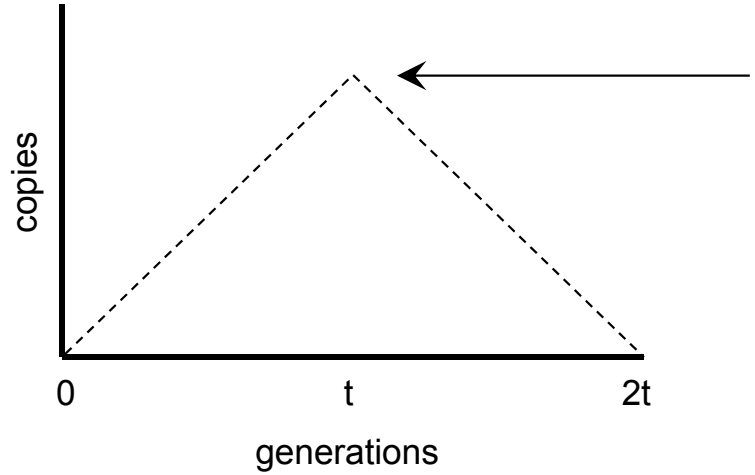
$(4N)(2Nu) < 1$ requires $N < (8u)^{-1/2}$.

Example. With $u = 10^{-9}$, $N < 100{,}000$.

How likely is it that a first-step mutation will acquire a second-step mutation destined to pull it to fixation prior to being lost by drift?



The Rescue Effect With a Neutral Intermediate-state Allele



Essentially all new mutations are destined to loss by drift.

Probability of surviving to generation t is ~$2/t$,

in which case the average number of copies is ~$t/2$.

- With an average $t/4$ copies of the allele being present over this time span of $2t$ generations, a lineage of age t will yield an average ~$(2/t) \cdot (t/4) \cdot 2t = t$ opportunities for the arrival of second-site mutations.

- The cumulative number of targets for second-step mutations grows as ~t^2.

- The solution to $t^2 \cdot u \cdot 2s = 1$ yields the approximate mean arrival time for a second-step mutation destined to fixation. The rescue rate is then $\sqrt{(u \cdot 2s)}$.

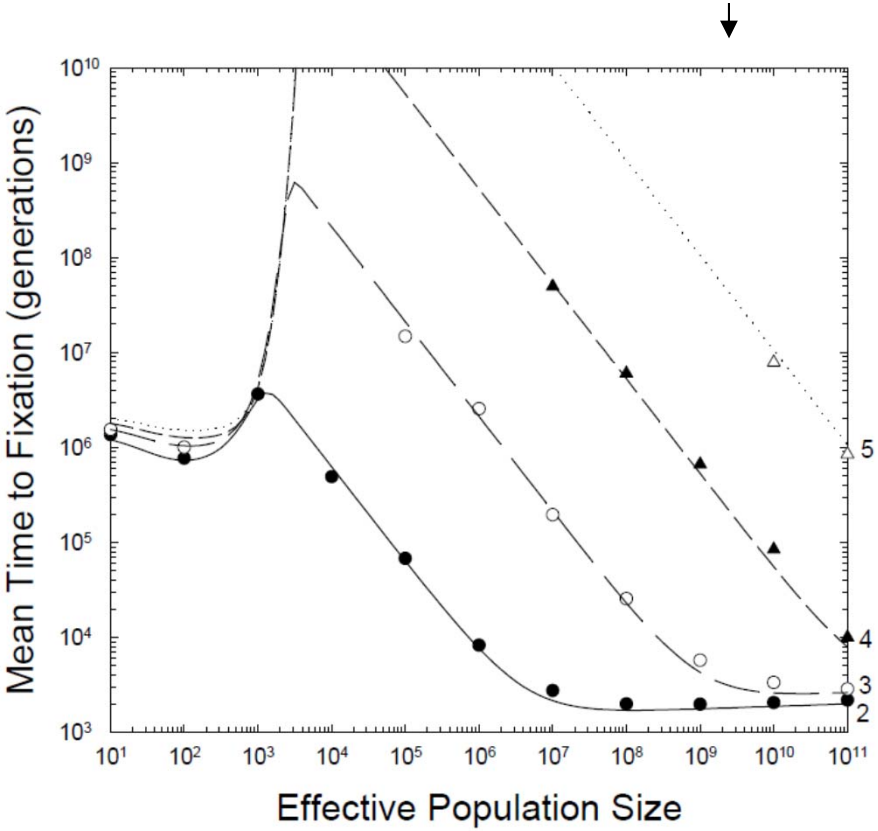
- Because $2Nu$ first-step mutations arise per generation, the expected rate of appearance of the adaptation is $2Nu \cdot \sqrt{(u \cdot 2s)}$.

Recall for a one-site adaptation, the rate is $4Nsu$.

Evolution of a Complex Adaptation Through Deleterious Intermediates: complete linkage

- Large populations acquire the adaptation much more rapidly than small populations, and can do so quite quickly.
- Maladapted individuals are the source of this kind of adaptation.

Rate of establishment = $4Nd!(u/s_1)^d(s_2/s_1)$



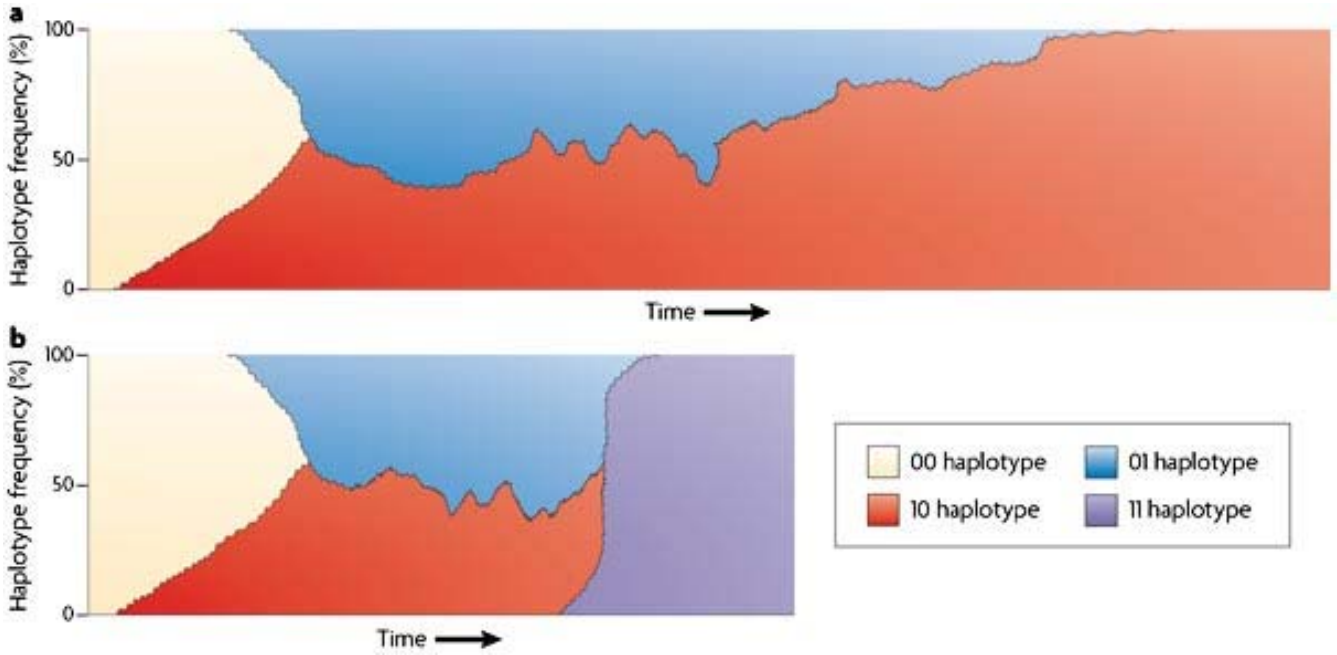
Small populations:

d-1 intermediate steps are *effectively* neutral.

Large populations:

- Deleterious first-step alleles are present at selection-mutation equilibrium copy number $2N(u / s_1)$;
- Each such newly arisen allele has a half life of $1 / s_1$ generations, and must acquire d-1 additional mutations for adaptation prior to elimination, so the rate of origin of adaptive alleles scales with $2Nd!(u / s_1)^d$;
- Probability of fixation of adaptive mutant is $2s_2$.

The Classical View That Recombination Enhances the Rate of Adaptive Evolution

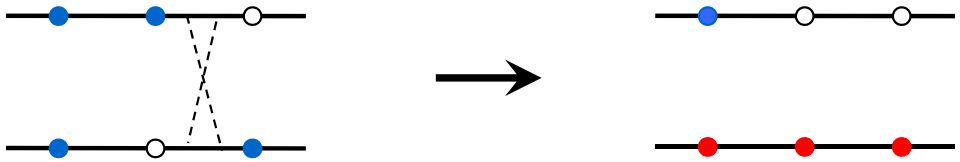


No recombination:
the two single-site mutations "compete" for fixation, and the final two-site adaptation must evolve by sequential substitutions.

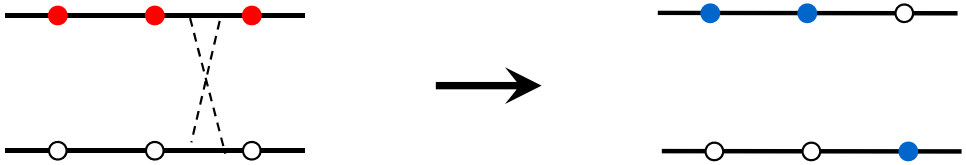
With recombination:
the emergence and fixation of the final adaptation is accelerated.

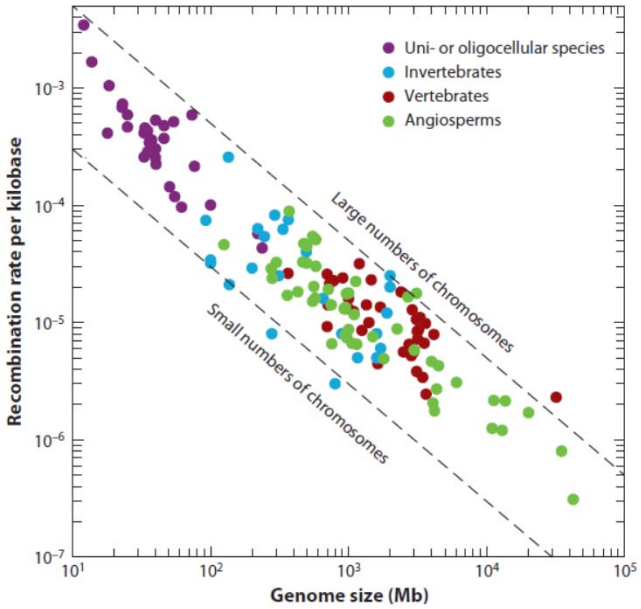
How Does Recombination Influence the Evolution of Complex Traits?

Recombination can facilitate the arrival of an adaptive combination,



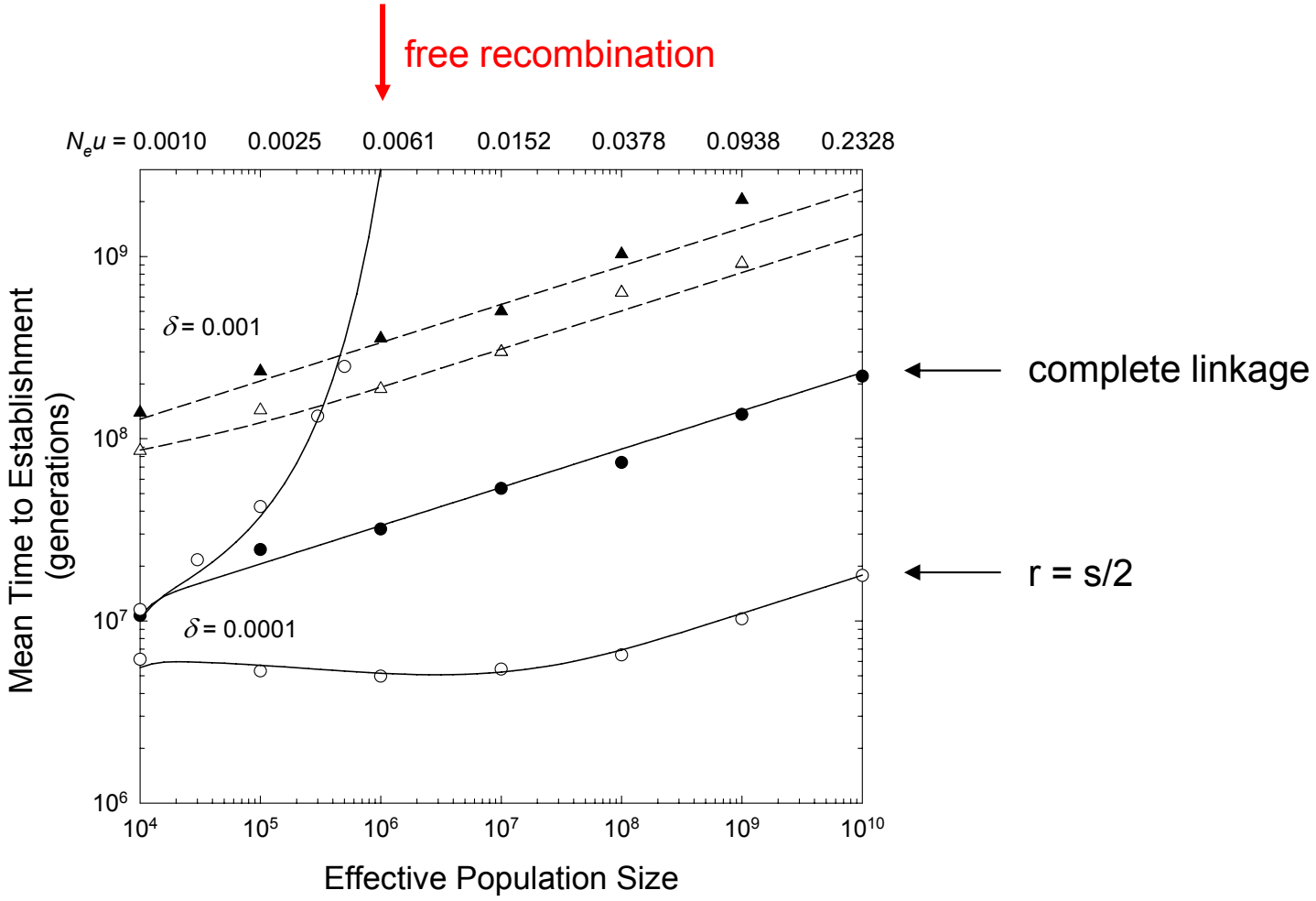
but it also inhibits the fixation of the adaptive allele,





Scaling of the Time to Establishment With Deleterious Intermediate-state Alleles

- Free recombination imposes a barrier to the evolution of the adaptation even at moderate population sizes.
- Suggests that microbial eukaryotes, which have high recombination rates and large population sizes are are unlikely to be the evolutionary sources of complex adaptations involving deleterious intermediates.



- The population-genetic mechanisms of the origin of molecular adaptations.

- The scaling of the rate of adaptation with population size.

    - Are large populations always capable of more rapid adaptation?

“Lynch has not only a fundamental misunderstanding, but a disdain for natural selection. His ideas should never be released the general public.”

Influence of the Recombination Rate When Intermediate-state Alleles Are Neutral

- There is a threshold recombination rate (r), approximately equal to the selective advantage (s) of the adaptive allele, beyond which the rate of adaptation is inhibited.

  ** This occurs because the rate of advancement of the adaptation (s) is exceeded by the rate of breakdown.

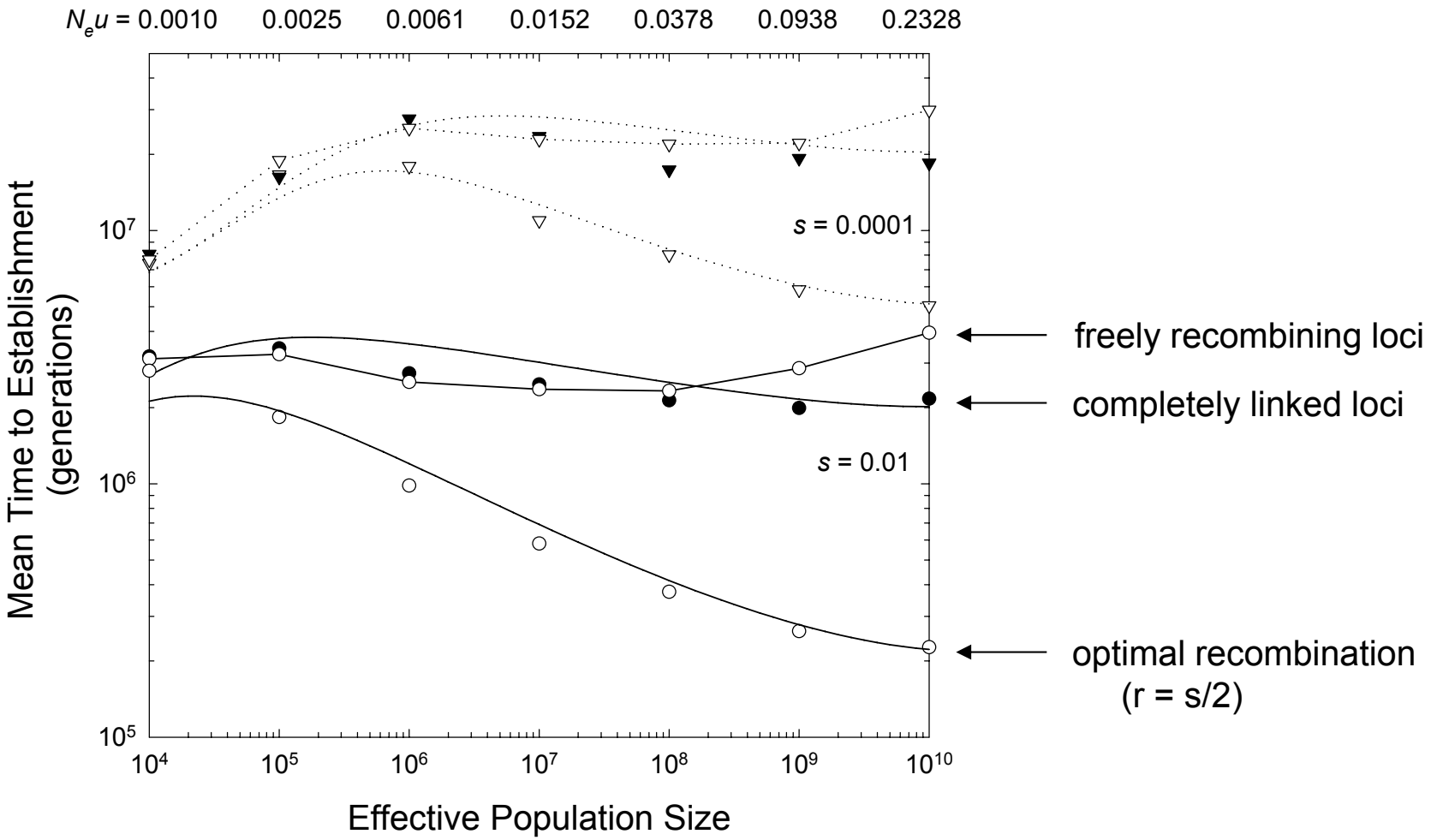
- The “optimal” recombination rate is equal to half the selective advantage:

  1) the rate of production of AB alleles by Ab/aB heterozygotes is proportional to r;
  2) the net selective advantage of AB is reduced to (s - r);
  3) the product r(s-r) is maximized when r = s/2.

- The effects of recombination are diminished in populations of small size because the evolution occurs sequentially, with each mutation becoming fixed before the next arrives.

Scaling of the Time to Establishment When Intermediate-state Alleles Are Neutral

• Recombination does not greatly influence the scaling of the time to establishment with population size.



- A/T mutation pressure drives proteins in the direction of more hydrophobic residues, which reduces problems with unfolding, but also increases problems with misfolding.

  Such bias seems to be greatest in endosymbiotic microbes (Bastolla et al., 2004, J. Mol. Biol.)

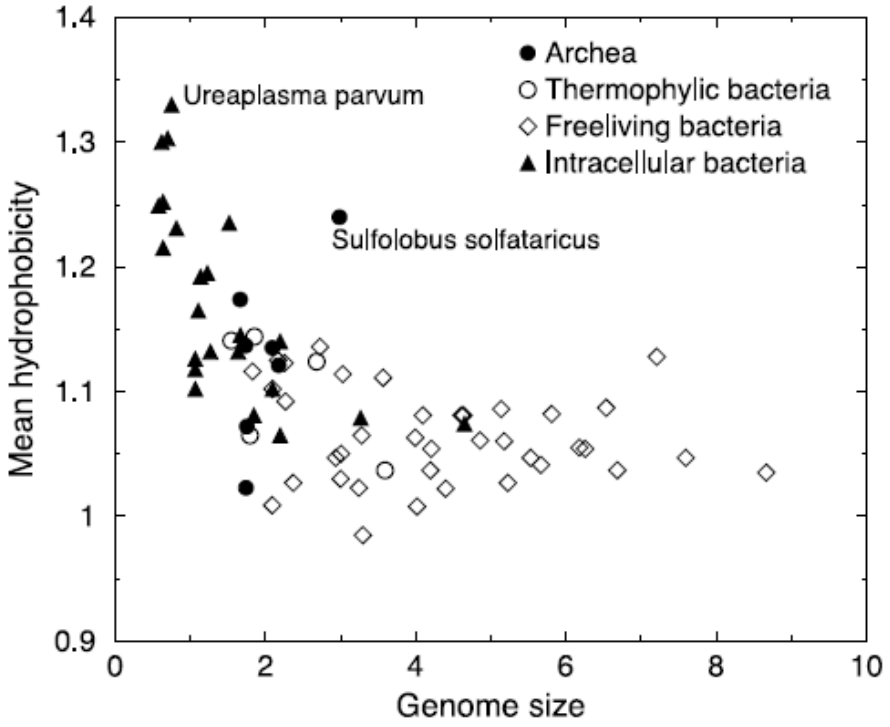


Figure 6. Mean protein hydrophobicity measured with the scale CH *versus* genome size. Each point represents a prokaryotic species.

Proteins With High Expression Rates Are Structured to Avoid Aggregation

Human

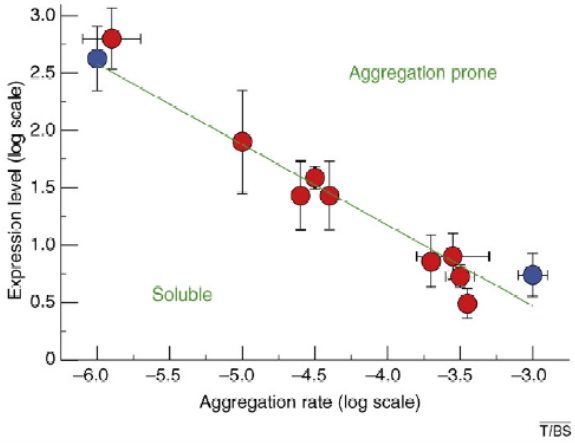
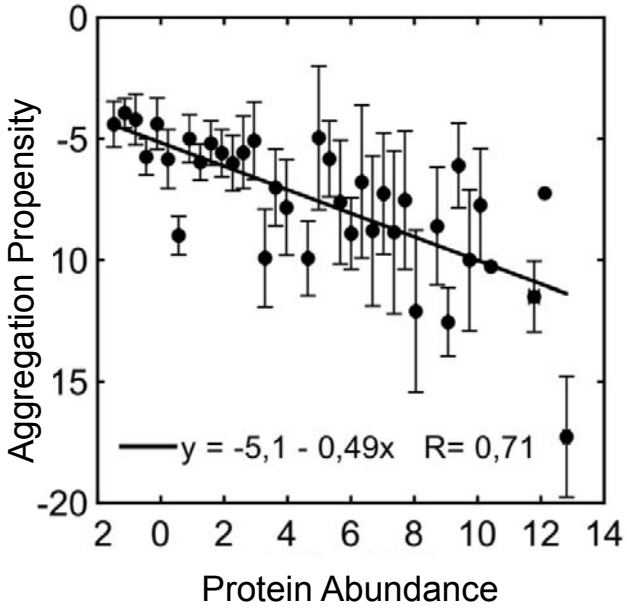


Figure 1. Correlation between expression levels and the measured aggregation rates of the corresponding proteins. The expression levels are estimates taken from measurements of the cellular mRNA concentrations [2]. All the data obtained from a comprehensive search in the databases of human expression levels and of the amyloid aggregation literature are included provided that the aggregation rates are measured at pH values between 4.0 and 8.0 (see Table 1). The standard deviations of the rates are reported only in four cases [7] because these values are not available or difficult to extract from the published data. The two proteins not involved in any known disease (acylphosphatase and glucagon) are represented by blue circles [1]. To generate a homogeneous set of variables, the expression levels of each of the 158 human samples (79 different tissues) in the database used here [2] were normalized by median scaling and between the samples by quantile normalization [17]. The expression levels were averaged over all the tissues in which a particular gene is expressed.

Tartaglia et al. (2007, Trends Biochem. Sci.)

E. coli

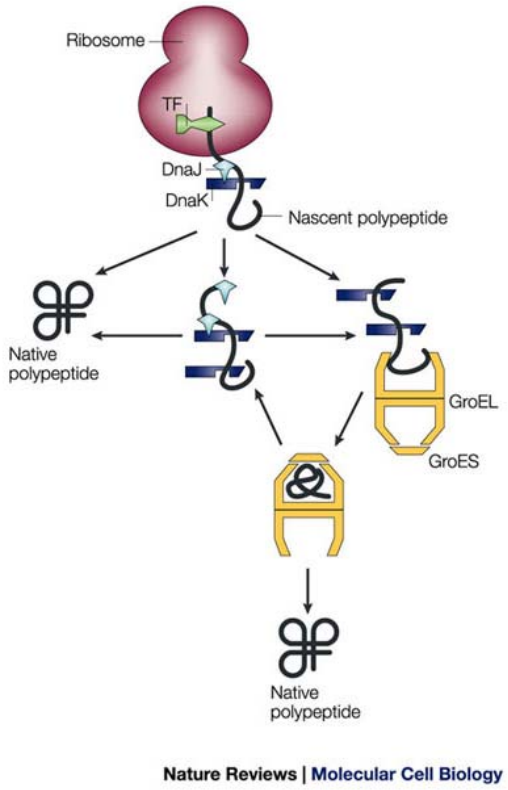


De Groot and Ventura (2010, PLoS ONE)

Defense Against Protein Misfolding: an ancient intracellular surveillance mechanism.

- Erroneous proteins that fail to acquire normal conformations are often assisted by chaperones, and when the problems cannot be solved, are removed from the cell via proteasome degradation.

- Chaperones interact with their client proteins noncovalently, preventing inappropriate aggregation by isolating them. They typically recognize hydrophobic side chains exposed by proteins in their non-native states.



- The investment in protein surveillance is large:

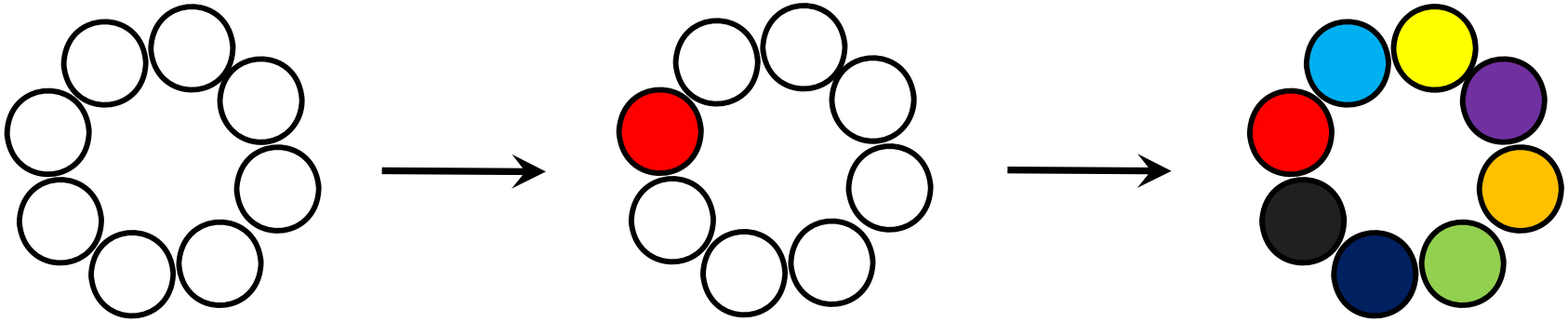
  Proteasomes constitute on the order of 1% of the total protein in mammalian cells, and chaperones constitute an investment of a similar order of magnitude.

  Up to 30% of newly synthesized proteins appear to have significant enough problems to require disposal (Schubert et al. 2000), and many proteins are incapable of folding without chaperone assistance.

  - About 250 proteins are repeatedly isolated with the chaperonin complex in *E. coli*, and at least 49 of these are absolutely dependent on chaperones for folding (Kerner et al. 2005; Fujiwara et al. 2010).

  - In the archaebacter *Methanosarcina mazei*, 333 proteins are chaperonin substrates (Hirtreiter et al. 2009).

How does a heteromeric ring evolve a specific order of subunits (such rings are also found in the nuclear pore and the proteasome)?



At each step in the process, a mechanism is required to permanently preserve the new and the old members – gain of a beneficial function, or complementary losses of subfunctions?

At each “step” in the evolutionary process, at least two mutations would seem to be required.

Does hetero-oligomerization (mixed ring types) in early stages result in “dominant-negative” effects?

Evolutionary Consequences: coevolution between protein features and chaperone expression.

- **What is the cost of chaperone mediated refolding, an ATP-dependent process?**

  By expressing a nonessential protein in yeast (e.g, YFP), Geiler-Samerotte et al. 2011) found that the misfolding of as little as 0.1% of total cellular protein can cause as much as a 3% decline in growth rate (independent of protein function), while also inducing up-regulation of chaperones. The degree to which protein misinteraction is involved is unclear.

- **Does chaperone dependence relax selection on client proteins, allowing otherwise deleterious misfolding mutations to accumulate, and/or allow the evolution of adaptations that would otherwise not be possible because of their negative effects on folding?**

  Proteins that are clients of chaperones evolve more rapidly than those that are not (Williams and Fares 2010).

  Tokuriki and Tawfik (2009) found that elevated levels of GroEL can facilitate the evolution of enhanced function of enzymes in *E. coli*. Performed a cycle of selection and mutagenesis experiments with four enzymes in presence / absence of excess GroEL production. The GroEL-dependent treatments led to twice the number of accumulated AA changes as in controls, suggesting that the elevated chaperonin levels promotes the evolution of variants with compromised folding and stability but enhanced enzyme activity.

- **A zero-sum game?**

  In *E. coli*, genes whose protein products are clients of the molecular chaperone GroEL harbor significantly lower frequencies of optimal codons (and hence are expected to experience higher rates of translational error) than do sporadic clients (Warnecke and Hurst 2010). This seems to corroborate the drift-barrier hypothesis, as there is no pleiotropic constraint preventing the joint evolution of optimal codons and use of chaperones.

Deleterious-mutation Accumulation and Compensation by Elevated Chaperone Levels

- Experiments with *E. coli* and *S. typhimurium* show that deleterious-mutation accumulation leads to a situation in which survival is enhanced by the overexpression of GroEL or DnaK (Fares et al. 2002; Maisnier-Patin et al. 2005; Van Dyk et al. 1989). This suggests that elevated mutation loads tend to select for genotypes with higher expression of chaperones.

- GroEL is the most highly expressed gene in the aphid endosymbiotic bacterium *Buchnera*, constituting about 10% of total protein (Baumann et al. 1996), and it is also upregulated in other bacterial endosymbionts (Aksoy 1995; Sato and Ishikawa 1997; Charles et al. 1997; Haines et al. 2002).

- Moran's (1996) hypothesis that up-regulation of GroEL in endosymbionts is an evolutionary innovation (apparently arising by compensatory mutation) in response to accumulated protein-folding problems that arise in species experiencing elevated levels of random genetic drift.

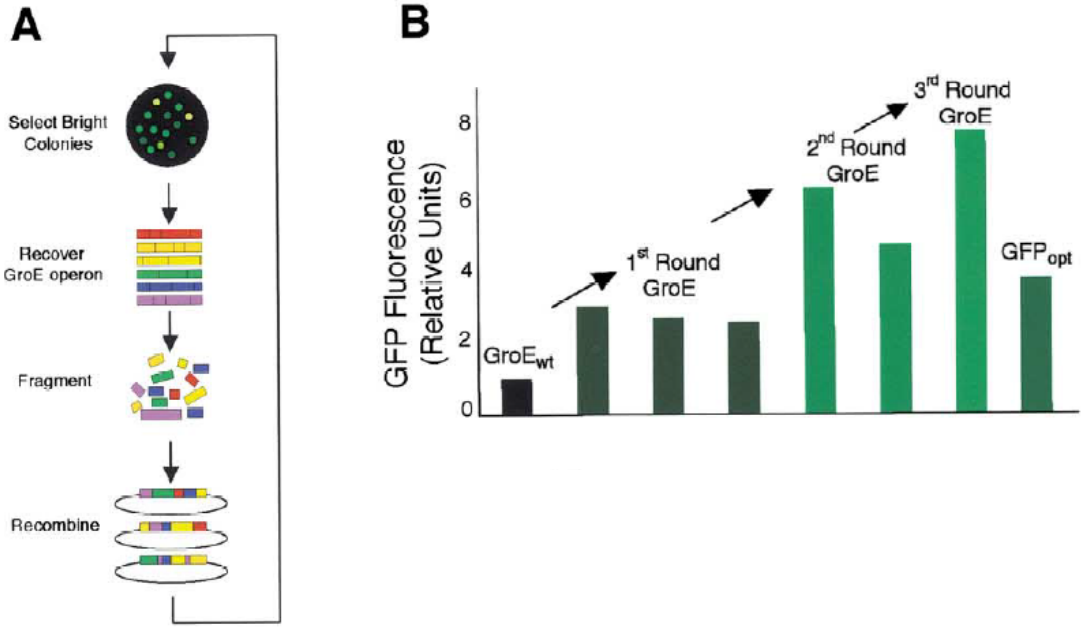
Coevolution of chamber and cargo:

- GroEL can bind and release actin and tubulin (eukaryotic-specific proteins) in an ATP-dependent manner, but does not fold them.

- CCT chaperonin does not even bind some GroEL substrates.

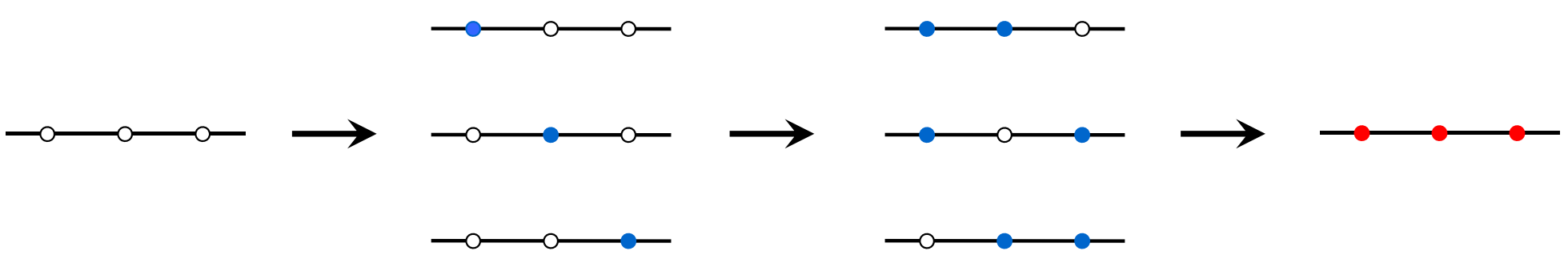
Experimental evolution of GroEL to bind a novel GFP substrate: *in vivo* selection, mutagenic PCR, and *in vitro* DNA shuffling (Wang et al., Cell, 2002).

- Improved ability to fold GFP came at the expense of the ability to bind natural substrates.



- In principle, chaperone specialization is achievable by gene duplication, divergence, and expression time/location specialization (e.g., mammalian testes).

Evolution of a Complex Adaptation Through Neutral / Deleterious Intermediates



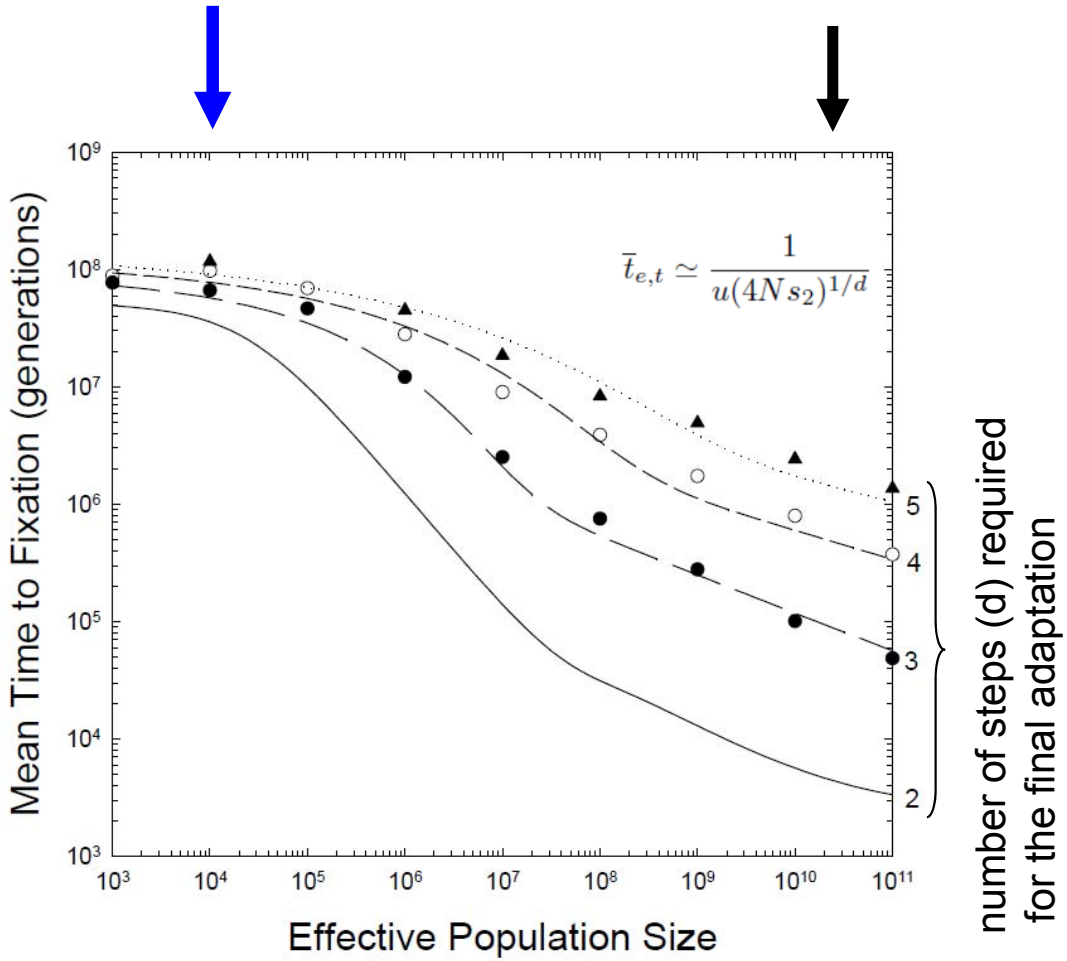
Fitness: 1 $1 - s_1$ $1 - s_1$ $1 + s_2$

A Three-site Model

Evolution of a Complex Adaptation Through Neutral Intermediates

Complete Linkage

Frequency of d-step alleles at time t = $(ut)^d$
Selection takes hold when the frequency reaches $1/(4Ns_2)$



- Small population size – rate of adaptation is nearly independent of complexity because the larger number of steps is compensated by the larger number of paths.

- Large population size – time to adaptation is inversely proportional to the mutation rate per site, not the product over all sites.

- As the number of steps increases, the relationship between N and the rate of adaptation becomes progressively flatter.

- With an evolvable mutation rate, the scaling with population size will be flatter because species with larger populations have lower mutation rates.