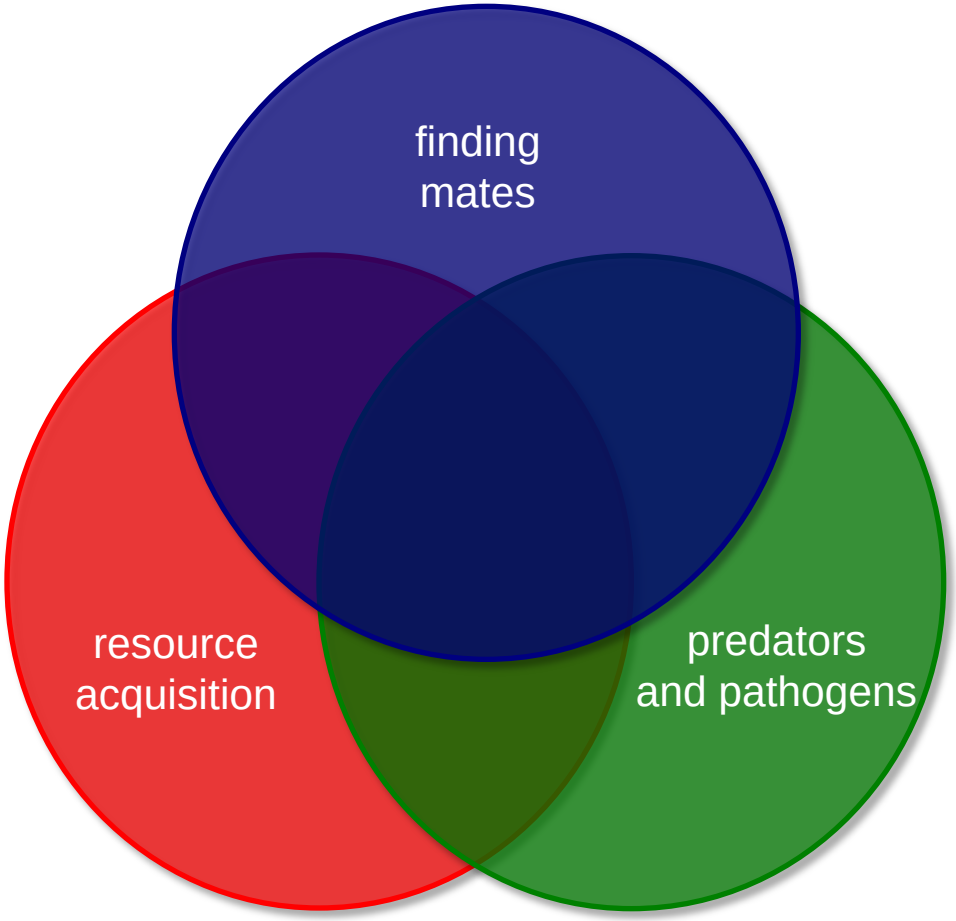
Is Evolution Simply a Matter of the External Environment?



Evolutionary cell biology: Two origins, one objective

Michael Lynch[a,1], Mark C. Field[b,2], Holly V. Goodson[c,2], Harmit S. Malik[d,e,2], José B. Pereira-Leal[f,2], David S. Roos[g,2], Aaron P. Turkewitz[h,2], and Shelley Sazer[i,1]

[a]Department of Biology, Indiana University, Bloomington, IN 47405; [b]Division of Biological Chemistry and Drug Discovery, University of Dundee, Dundee DD1 5EH, United Kingdom; [c]Department of Chemistry and Biochemistry, University of Notre Dame, Notre Dame, IN 46556; [d]Division of Basic Sciences and [e]Howard Hughes Medical Institute, Fred Hutchinson Cancer Research Center, Seattle, WA 98195; [f]Instituto Gulbenkian de Ciência, P-2781-901 Oeiras, Portugal; [g]Department of Biology, University of Pennsylvania, Philadelphia, PA 19143; [h]Department of Molecular Genetics and Cell Biology, University of Chicago, Chicago, IL 60637; and [i]Verna and Marrs McLean Department of Biochemistry and Molecular Biology, Baylor College of Medicine, Houston, TX 77030

Evolution ary Cell Biology

- The origin of all aspects of biodiversity ultimately resides at the cellular level.
  - To what extent do the internal workings of cells constrain the evolution of "external" phenotypes? Are there enough degrees of freedom that the cellular details don't matter?

- Evolutionary biology is not simply comparative biology, but will require comparative studies at both the within- and among-species levels – unicellular species, prokaryotes and eukaryotes.

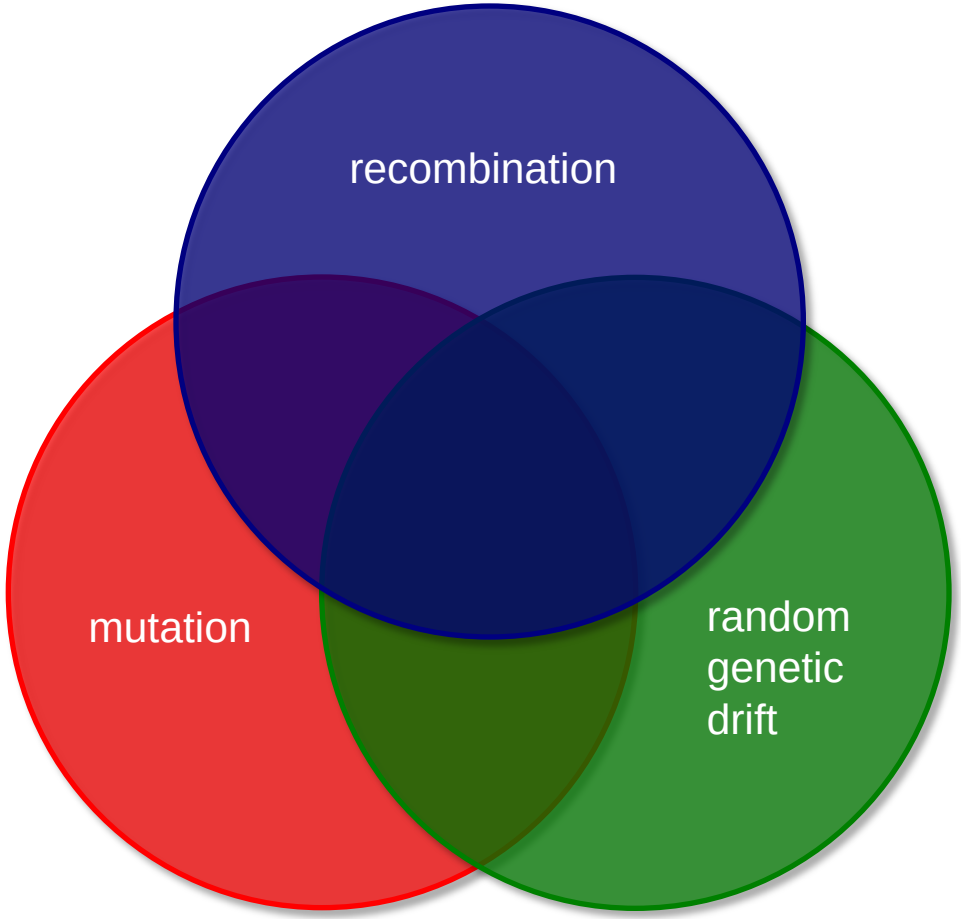
- What are cell biology's scaling laws, and how do we explain them?

- Potential for developing a mechanistic, integrative understanding of evolution:

  Biophysics ----- Population Genetics ----- Biochemistry

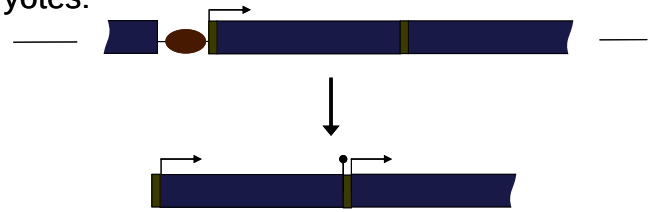
- Evolution is not a simple matter of natural selection – how much of cellular evolution is driven by nonadaptive processes?

The Population-genetic Environment

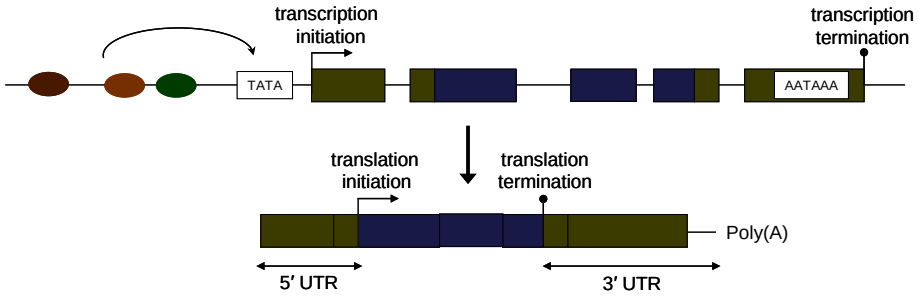


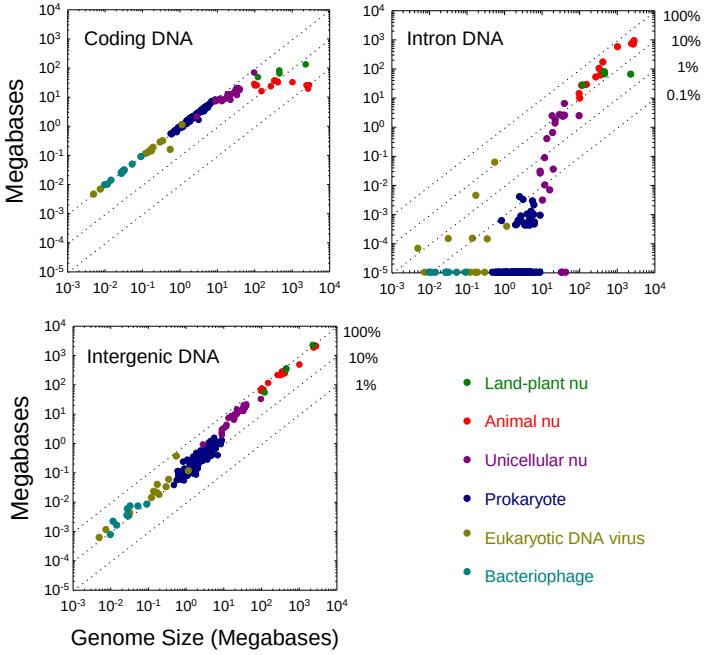
The Origin of Gene-structure Complexity by Nonadaptive Mechanisms

Prokaryotes:



Eukaryotes:

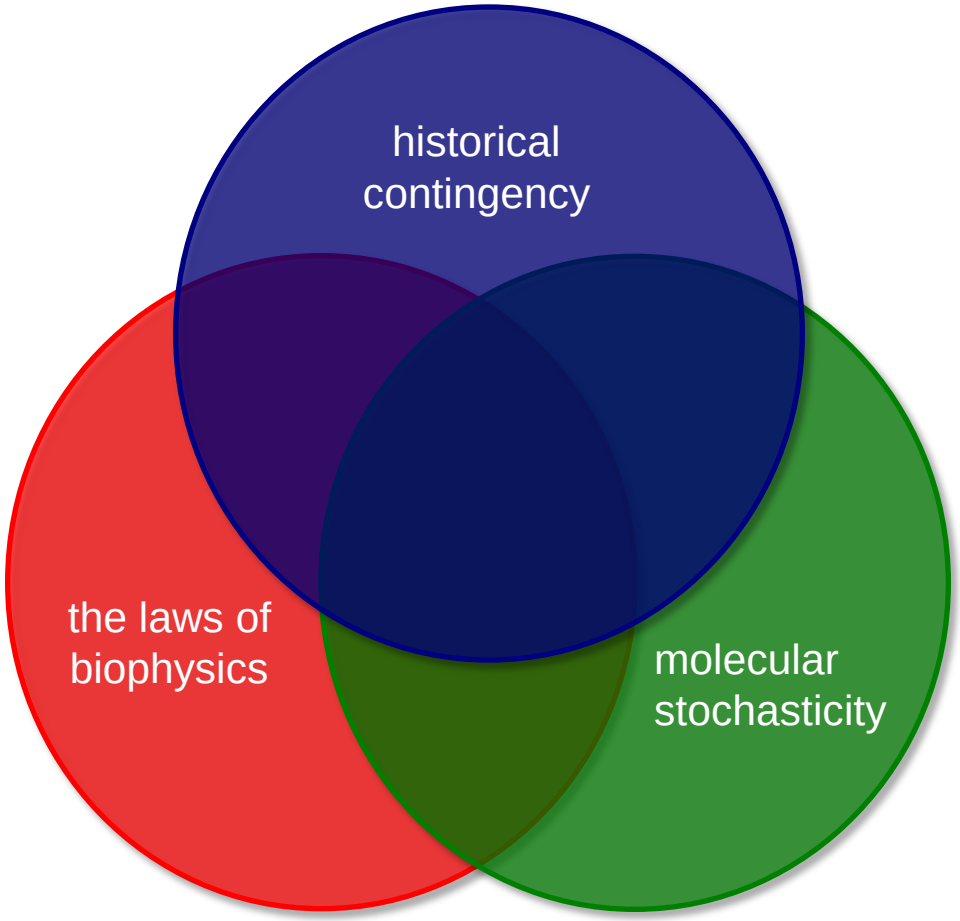




- Nearly all embellishments to gene structure impose weak mutational disadvantages. While these can be efficiently removed by selection in prokaryotes with large effective population sizes, they can accumulate in an effectively neutral fashion in eukaryotes experiencing relatively high levels of random genetic drift.

Can these general principles help explain structural features of proteins and cellular diversity?

The Cellular Environment



Three Vingettes:

- Some cell biological scaling features.
- Intraspecific diversity in cellular features explained by variation in the power of random genetic drift.
- Unsolved issues on the higher-order structure of proteins.

Was the Increase in Energy Produced by Mitochondria a Pre-requisite for the Evolution of Complex Cells?

The energetics of genome complexity

Nick Lane[1] & William Martin[2]

Nature, 2010

- Unclear why the appropriate total currency is the energetic cost of running a gene.
- Genes can be selectively promoted for reasons that have nothing to do with energy acquisition.
- Need for baseline information on the lifetime energetic requirements of a cell, and the contributions from various cellular features.
- Need for an evolutionarily meaningful cost measure.

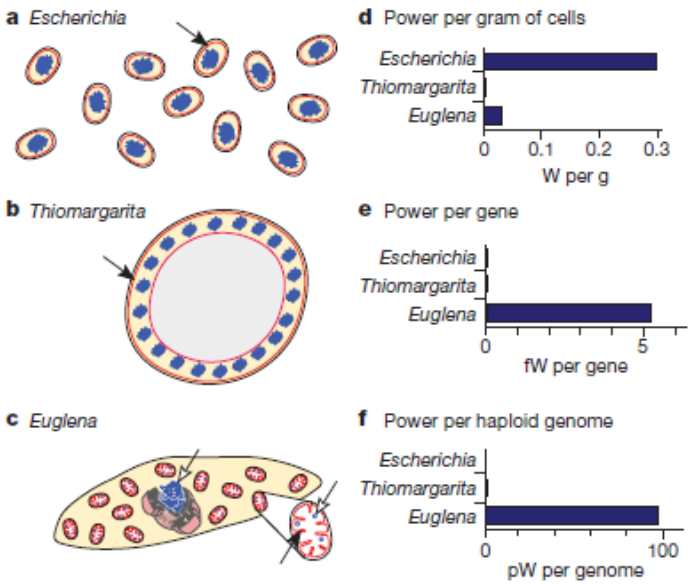


Figure 2 | The cellular power struggle. **a**–**c**, Schematic representations of a medium sized prokaryote (*Escherichia*), a very large prokaryote (*Thiomargarita*), and a medium-sized eukaryote (*Euglena*). Bioenergetic membranes across which chemiosmotic potential is generated and harnessed are drawn in red and indicated with a black arrow; DNA is indicated in blue. In **c**, the mitochondrion is enlarged in the inset, mitochondrial DNA and nuclear DNA are indicated with open arrows. **d**–**f**, Power production of the cells shown in relation to fresh weight (**d**), per haploid gene (**e**) and per haploid genome (power per haploid gene times haploid gene number) (**f**). Note that the presence or absence of a nuclear membrane in eukaryotes, although arguably a consequence of mitochondrial origin[70], has no impact on energetics, but that the energy per gene provided by mitochondria underpins the origin of the genomic complexity required to evolve such eukaryote-specific traits (see text).

Three Levels for the Cost of a Gene:

Georgi Marinov

1) **Chromosome**: synthesis of nucleotides for replication, and amino acids for nucleosomes.

2) **Transcription**: synthesis of ribonucleotides for steady-state number of transcripts.

2) **Protein**: synthesis of amino acids for steady-state number.

- All measured relative to the total energy budget of the cell in units of ATP hydrolyses.

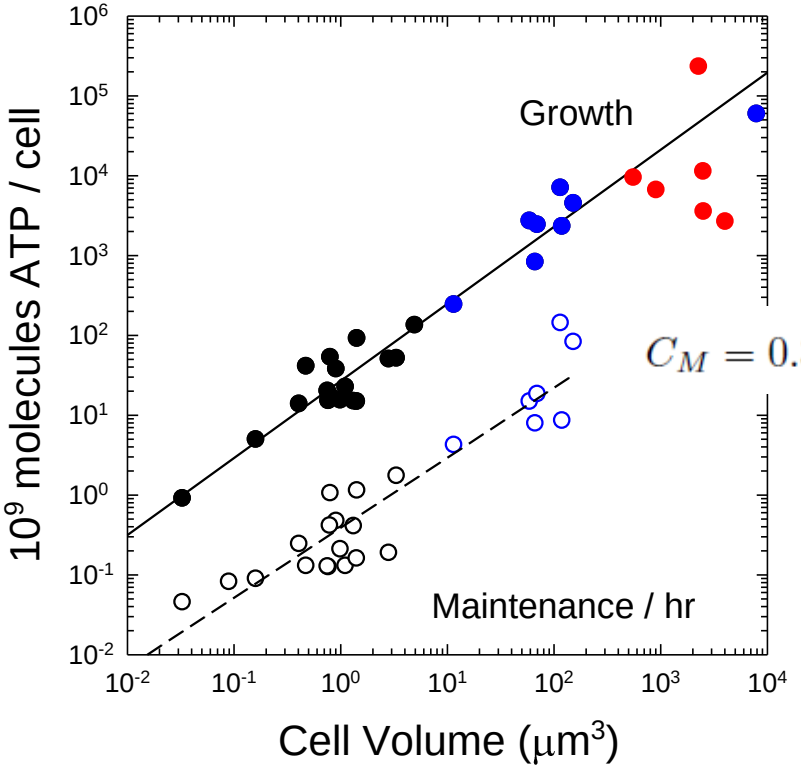
Evolutionary consequences:

Total baseline cost: $s_c = s_{DNA} + s_{RNA} + s_{PRO}$

Net selective advantage: $s_n = s_p - s_c$

- If $|s_i| < 1/N_e$ (N_e = the effective population size), selection is unable to eradicate or promote the feature – effective neutrality.

Lifetime Energy Requirement of a Cell

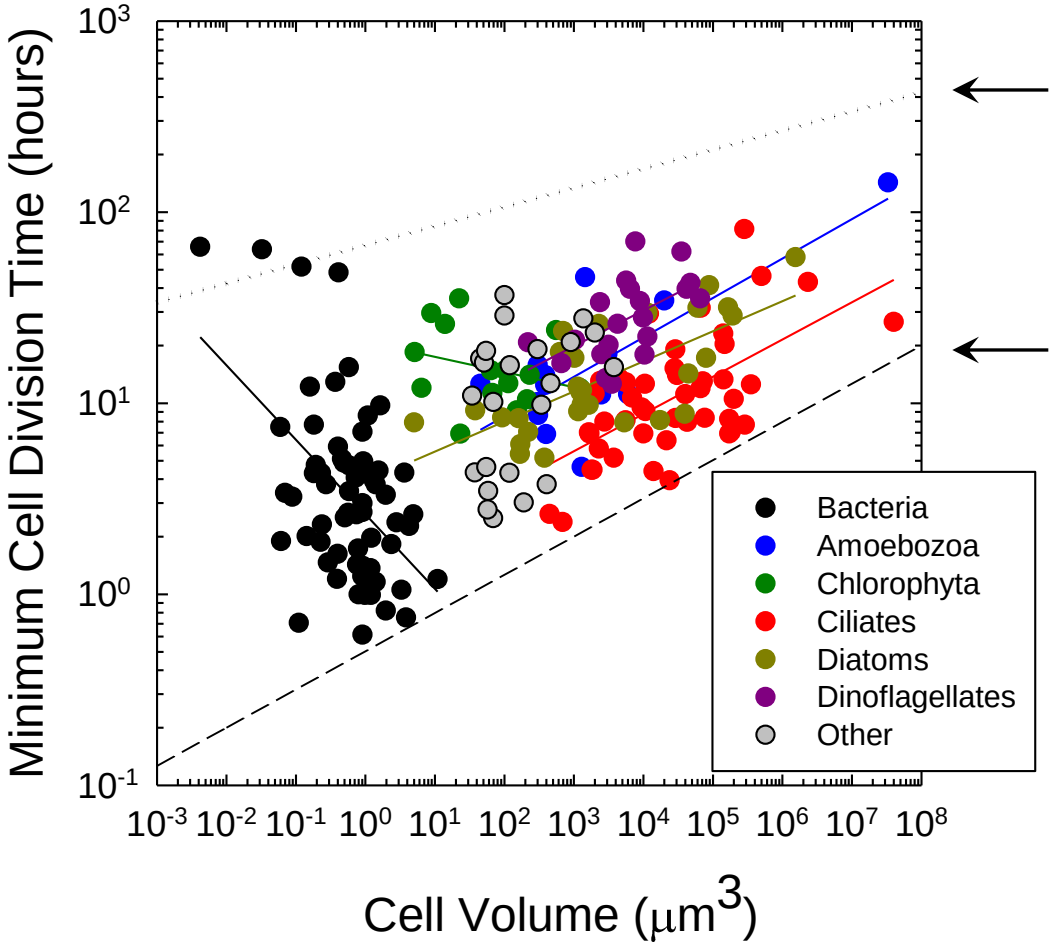


- Scaling is nearly isometric.
- Scaling is continuous across the prokaryote-eukaryote divide.
- It takes ~27 x 10^9 ATP hydrolyses to build 1 μm^3 of cell volume (an *E. coli* cell).
- What dictates the slopes and intercepts of these functions?

- Total ATP consumption / cell division: $C_T = C_G + tC_M$, where t = cell division time (hours).

  If $t < 69V^{0.09}$ hours (20 C), contribution from cell growth dominates.

Scaling of Cell-division Time With Cell Size



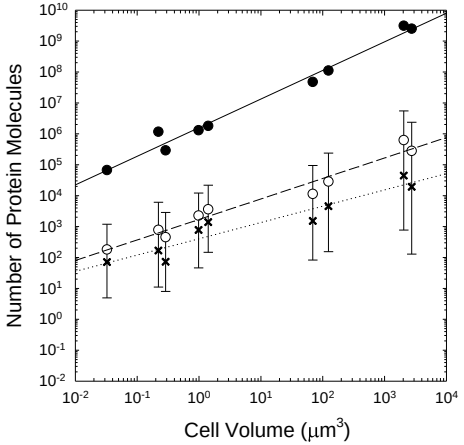
Equal expenditure on maintenance and parts replacement

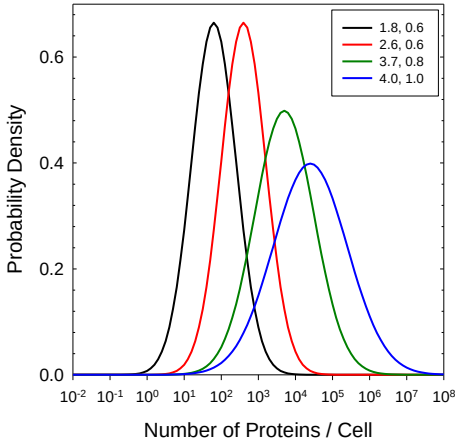
Lower limit to cell division times,

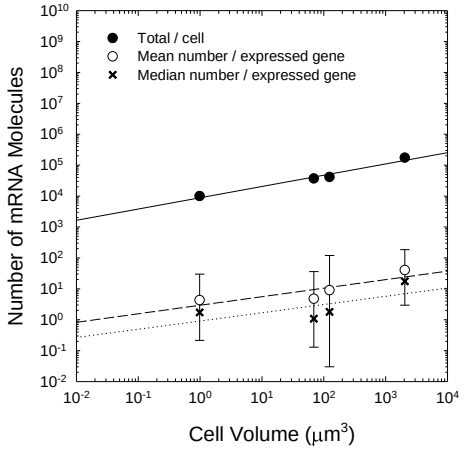
$t_{min} \approx 0.5V^{0.2}$ hours (at 20 C)

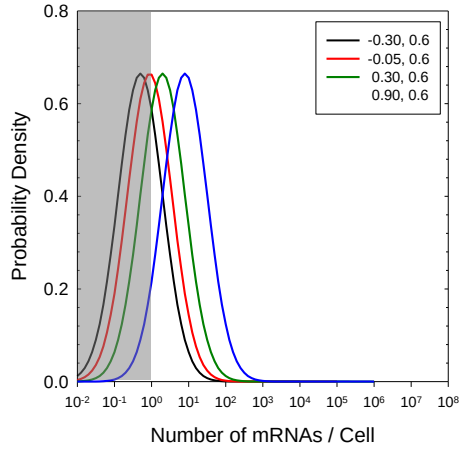
- Bacterial growth rates scale negatively with cell size, despite having larger numbers of genes.
- What defines the growth-rate speed limit?
- What dictates the scaling of the speed limit with cell size?

Scaling of Steady-state Numbers of mRNAs and Proteins With Cell Volume

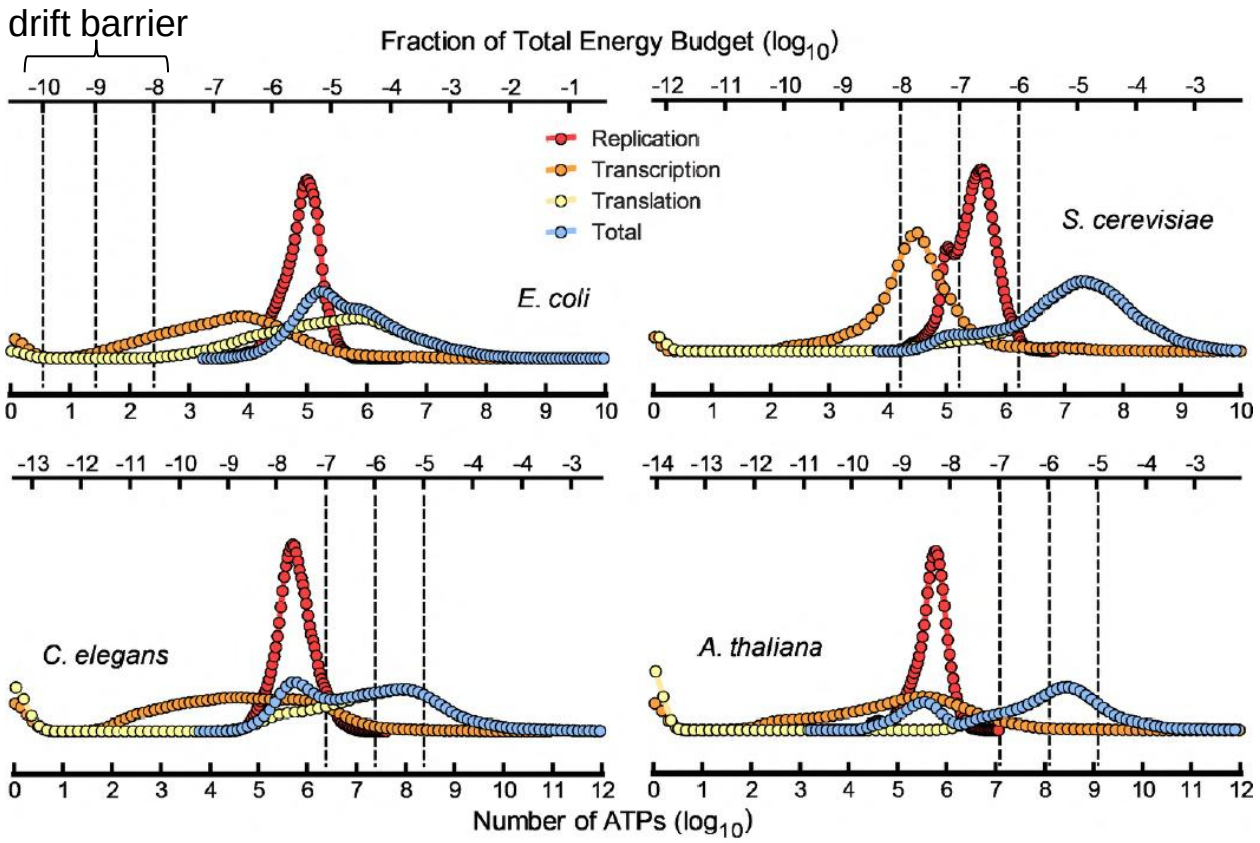






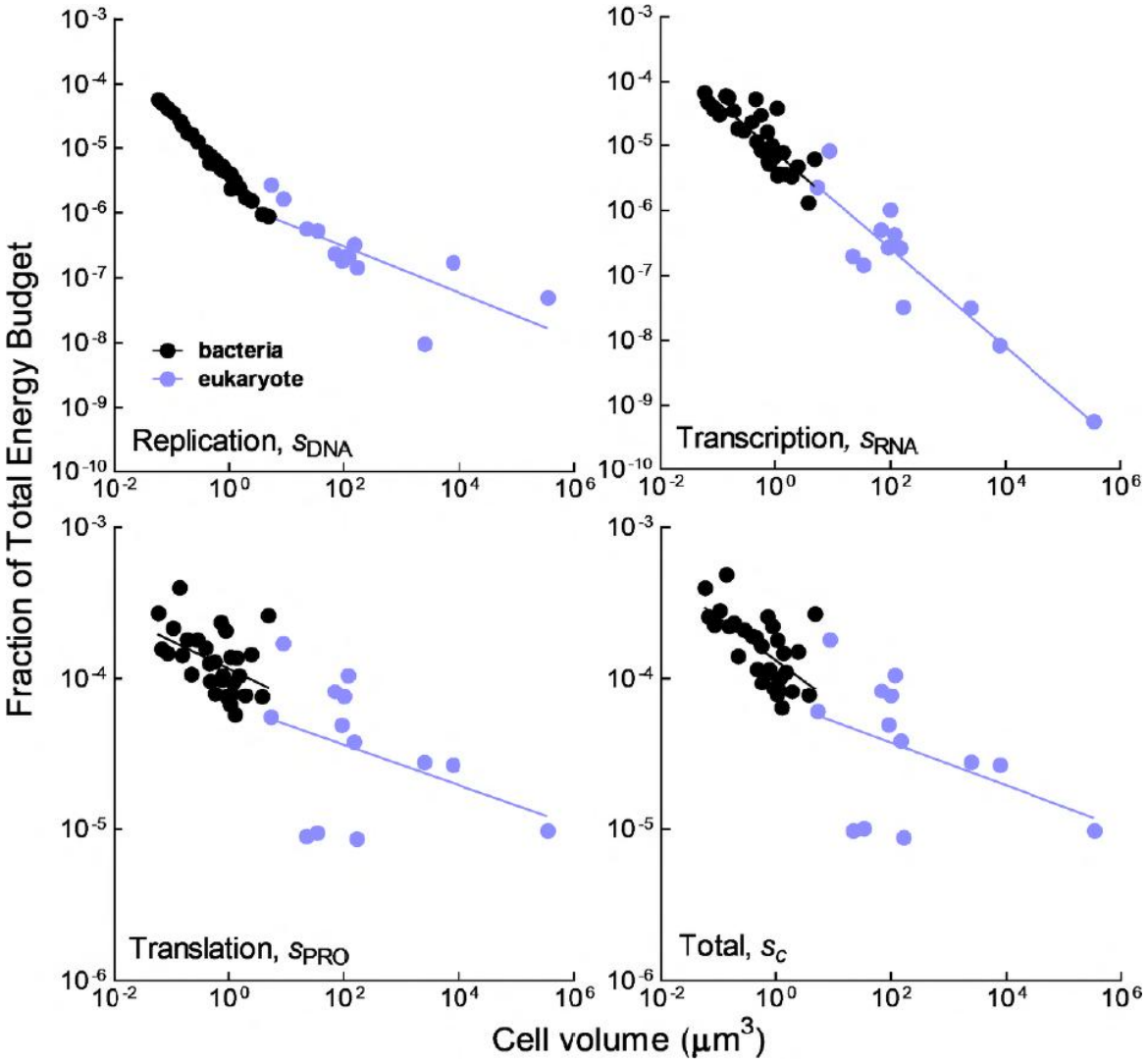


Distribution of the Costs for All Genes in Four Species

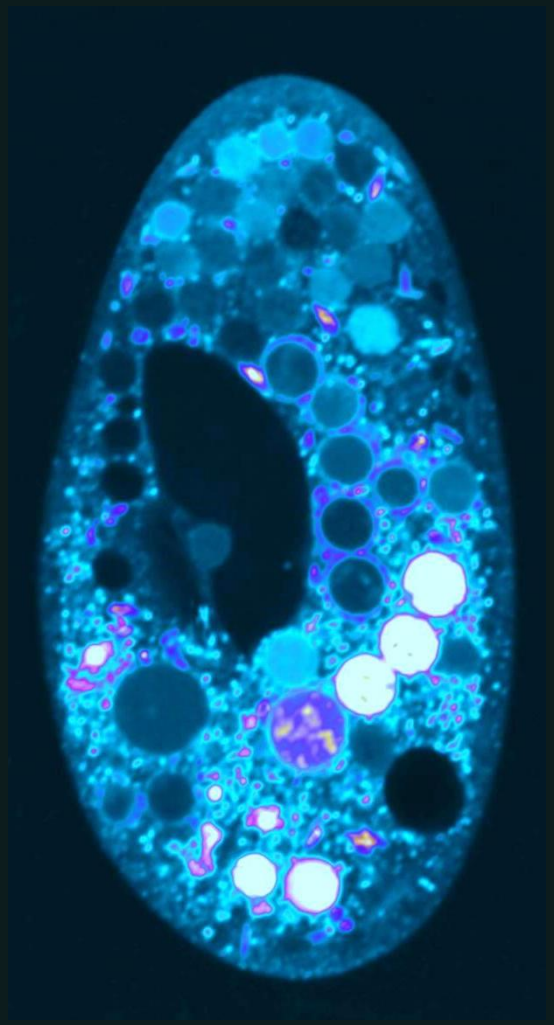


- Bacteria – costs are visible to natural selection at all three levels.
- Multicellular eukaryotes – costs are often one to two orders of magnitude higher than in bacteria, but at the DNA and RNA levels are often still too small to be perceived by selection.

Costs for Average Genes in 44 Species: continuity of negative scaling between bacteria and eukaryotes.



How far can natural selection drive an adaptation?

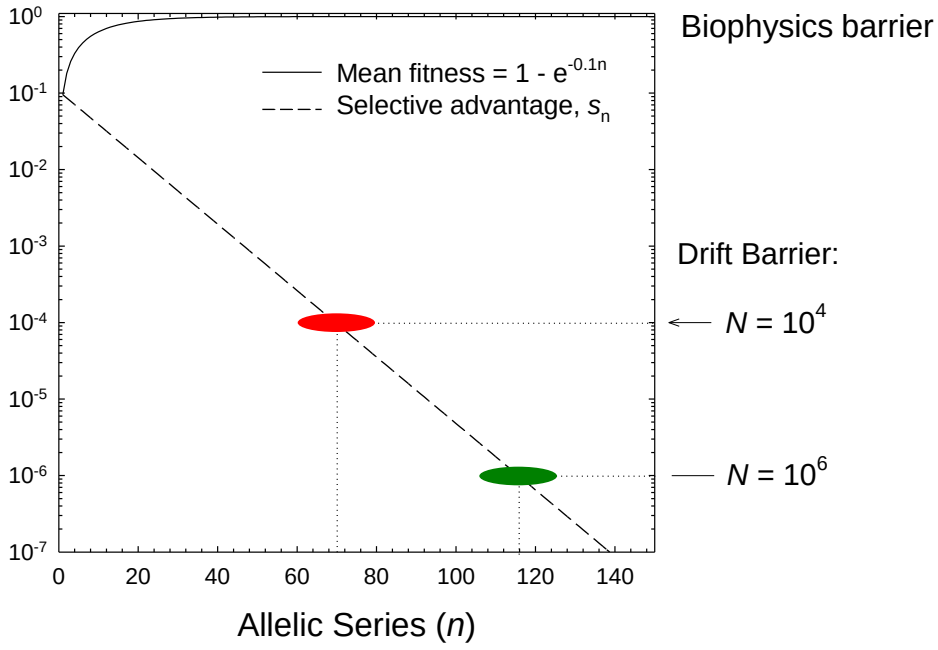


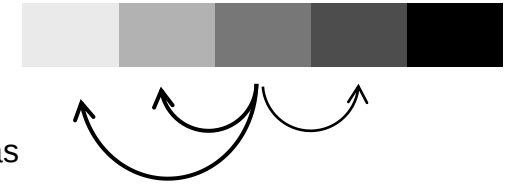
- Do cellular adaptations hit the Biophysics Barrier – the absolute limits of molecular perfection?

- **The Drift Barrier to Achieving Adaptive Perfection**: Once the selective advantage of improving a trait is less than the power of drift, $1/(2N_e)$, no further improvement in fitness can be sustained.

The Drift-barrier Hypothesis for a Single Trait

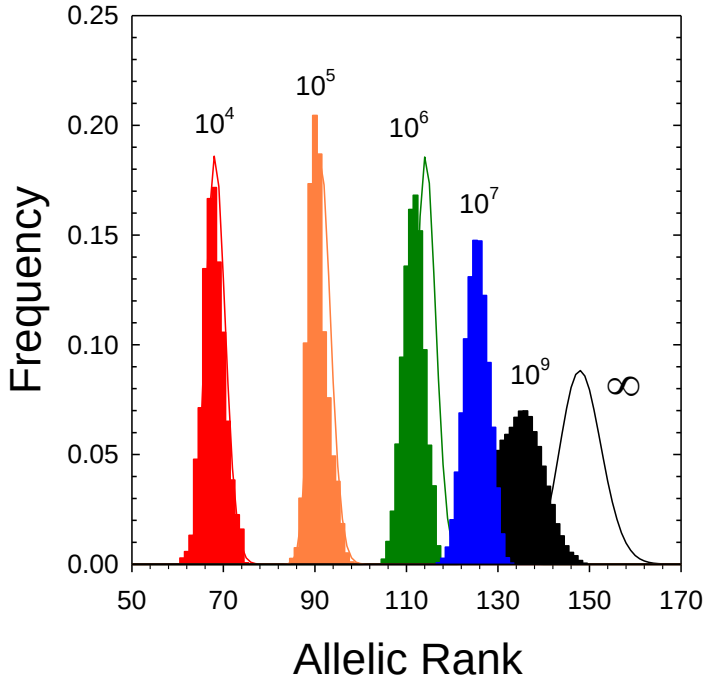
Asymptotically Increasing Perfection in an Allelic Series



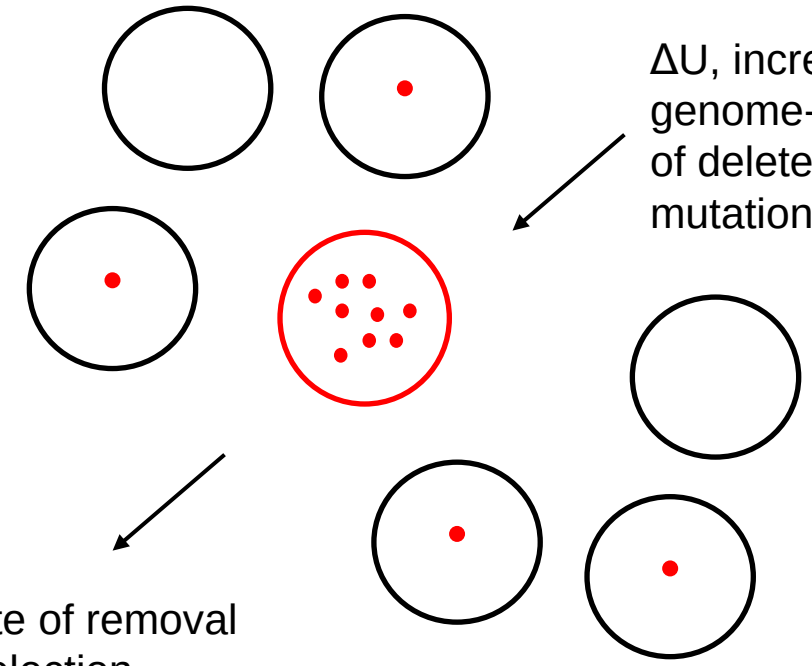


downward mutation bias

Equilibrium Allele-frequency Distributions with Increasing Population Sizes



The Magnitude of Selection Operating to Improve Replication Fidelity



ΔU, increase in genome-wide rate of deleterious mutation

s, rate of removal by selection

Excess number of mutations at equilibrium = ΔU / s

↓

X

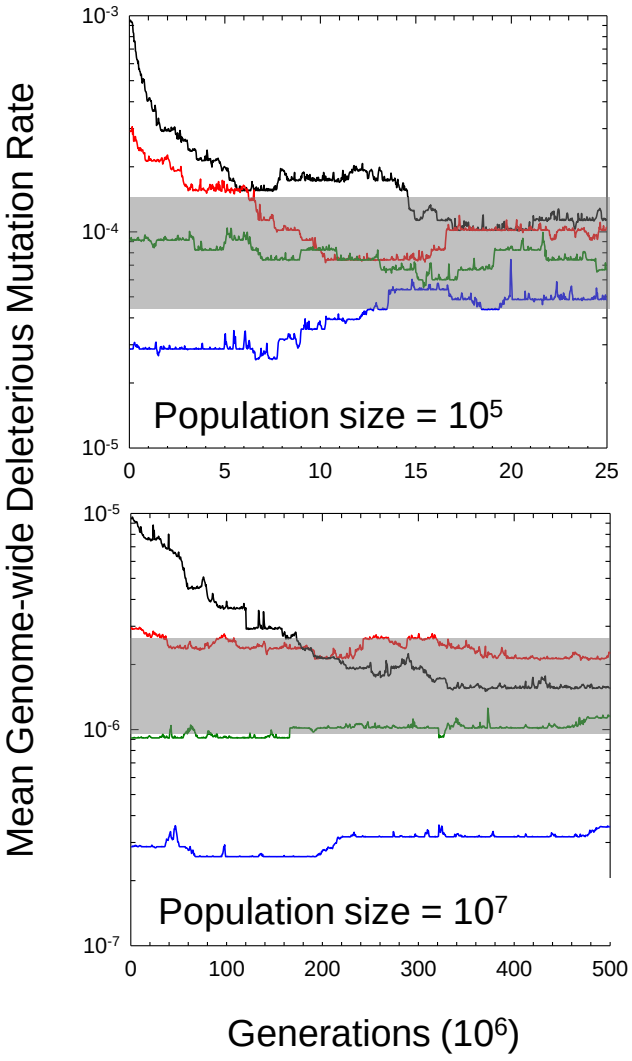
Effect / mutation = s

↓

Total effect on fitness = ΔU

- Selective disadvantage of a mutator = increase in genome-wide deleterious mutation rate

Quasi-equilibrium Mutation Rates Resulting From Deleterious-mutation Load

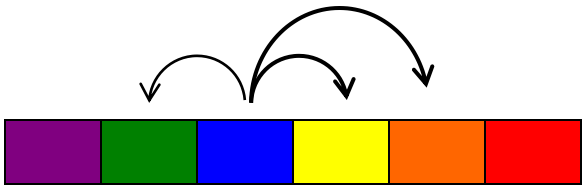


Effective selection for antimutators

DRIFT BARRIER

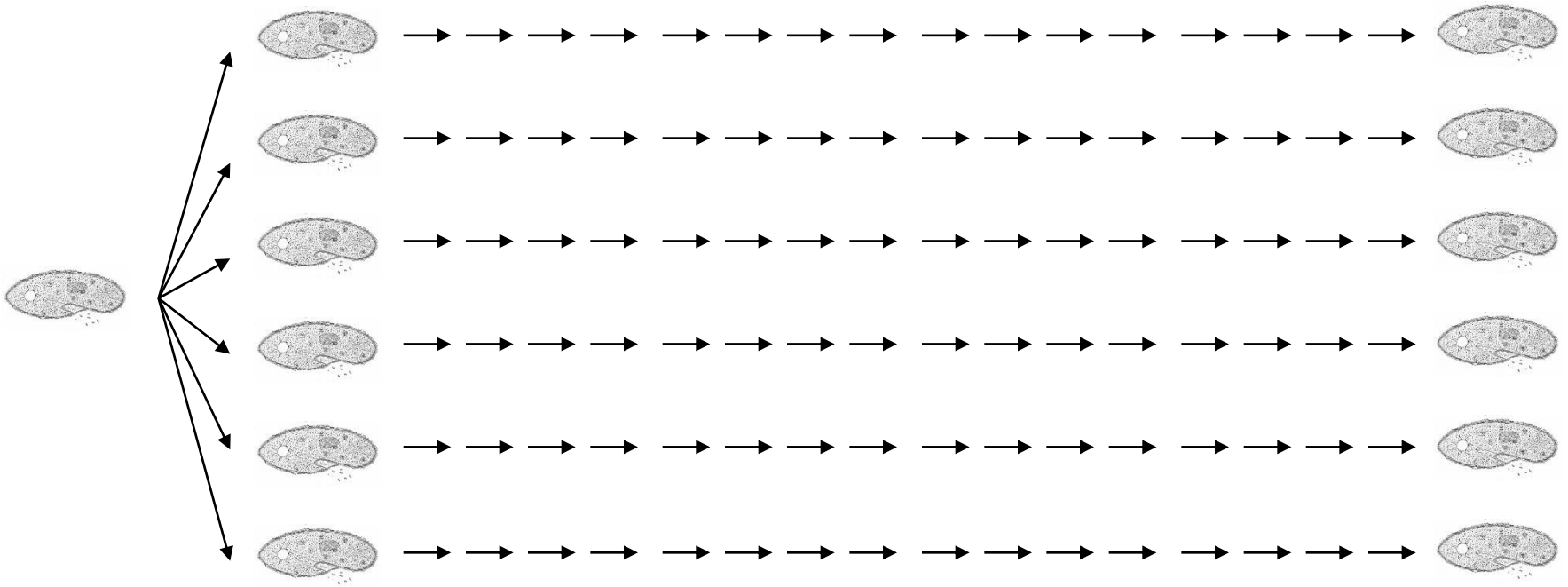
Biased production of mutators

- Equilibrium mutation rate is inversely proportional to the effective population size.



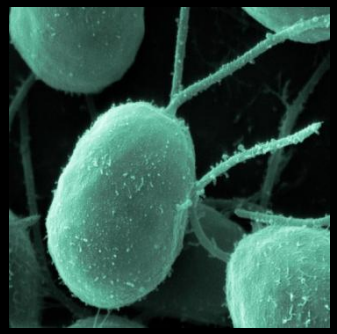
Analysis of Genome Stability with a Mutation-accumulation Experiment:

- Starting with a single stem cell, sublines are maintained by single-progeny descent, preventing selection from removing spontaneous mutations.
- Continue for thousands of cell divisions.
- Characterize by whole-genome sequencing.



Recent and Current Eukaryotic Targets of Study

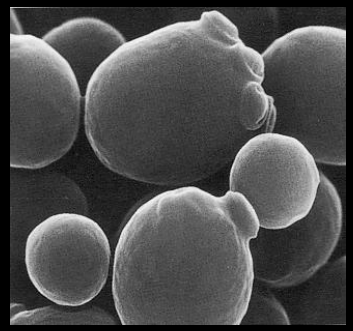
Arabidopsis



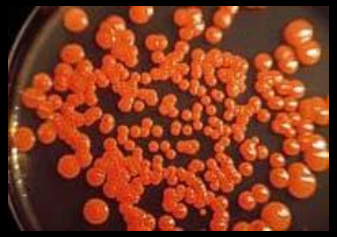
Chlamydomonas

Phaeodactylum

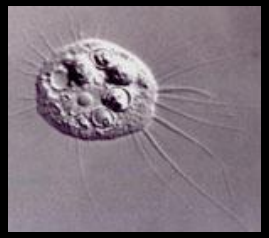
Dictyostelium



Saccharomyces



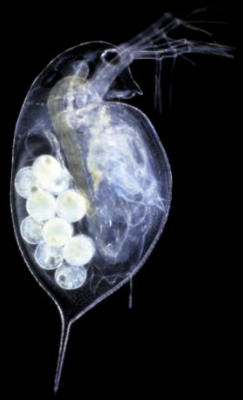
Rhodotorula



Ichthyosporean

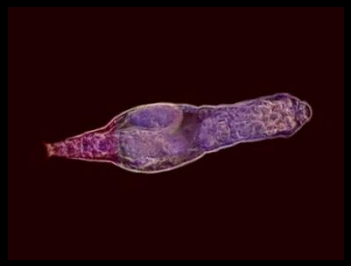
Naegleria

Paramecium

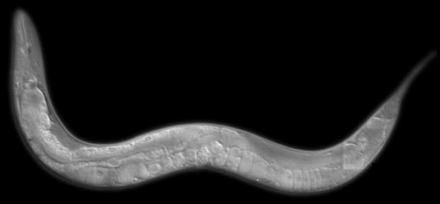


Daphnia

Drosophila



Adineta



Caenorhabditis

Mutation-accumulation Studies in Prokaryotes

Group	Species	Genome Size (Mb)	G/C %
Bacteria:			
Acidobacteria	*Acidobacterium capsulatum*	4.1	61.0
Actinobacteria	*Kineococcus radiotolerans*	5.0	74.2
Actinobacteria	*Mycobacterium smegmatis*	7.2	65.2
Actinobacteria	*Mycobacterium* sp.	7.2	65.2
Alpha-proteobacteria	*Agrobacterium tumefaciens*	5.7	59.0
Alpha-proteobacteria	*Caulobacter crescentus*	4.0	67.2
Alpha-proteobacteria	*Rhodobacter sphaeroides*	4.5	68.2
Beta-proteobacteria	*Burkolderia cenocepacia*	7.8	66.8
Beta-proteobacteria	*Janthinobacterium* sp.	6.0	61.1
Gamma-proteobacteria	*Photorhabdus luminescens*	5.7	42.8
Gamma-proteobacteria	*Pseudomonas fluorescens**	7.1	63.3
Gamma-proteobacteria	*Shewanella putrefaciens*	4.7	44.5
Gamma-proteobacteria	*Teredinibacter turnerae*	5.2	50.9
Gamma-proteobacteria	*Vibrio cholerae**	4.1	47.5
Gamma-proteobacteria	*Vibrio fischeri**	4.3	38.3
Cyanobacteria	*Synechococcus elongatus*	2.7	55.5
Deino-Thermus	*Deinococcus radiodurans**	3.2	66.6
Firmicute	*Bacillus subtilis**	4.2	43.5
Firmicute	*Staphylococcus epidermidis*	2.6	32.0
Flavobacteria	*Flavobacterium* sp.	6.1	34.1
Lactobacillale	*Lactobacillus* sp.	2.9	46.4
Planctomycete	*Gemmata obscuriglobus*	9.2	67.2
Tenericute	*Mesoplasma florum*	0.8	27.0
Archaea:			
Euryarchaeota	*Haloferax volcanii*	4.0	65.5

* = concurrent study with mismatch-repair deficient lines

Major recent contributors:

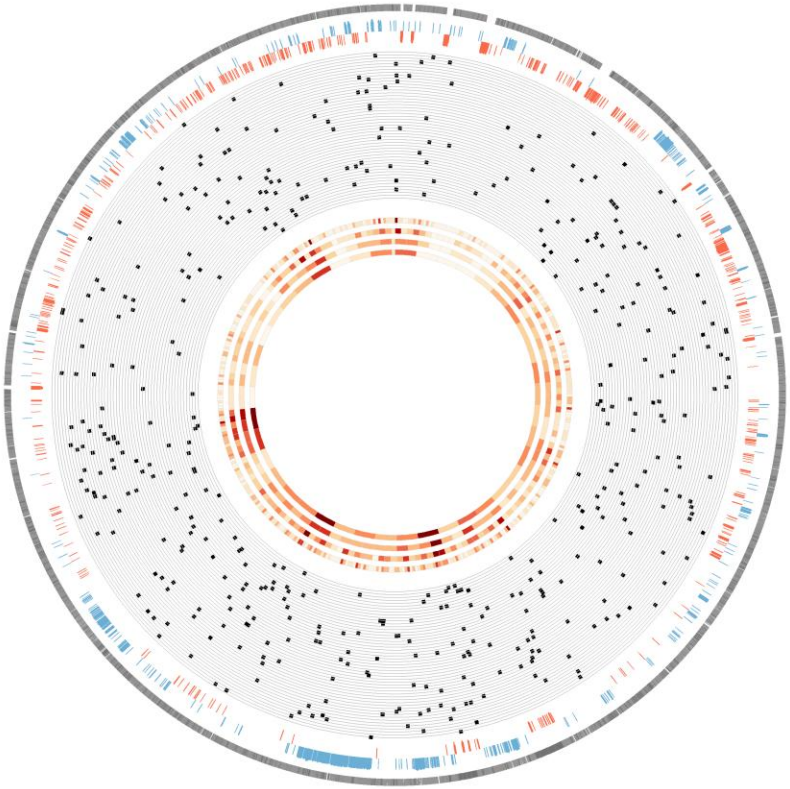
Indiana University:
- Matthew Ackerman
- Tom Doak
- Pat Foster
- Jean-Francois Gout
- Matthew Hahn
- Nate Keith
- Weiyi Li
- Hongan Long
- Sam Miller
- Ron Pearson
- Dan Schrider
- Way Sung
- Abe Tucker
- Emily Williams

University of New Hampshire:
- Vaughn Cooper
- Marcus Dillon
- Kelley Thomas

Hacettepe University:
- Sibel Kucukyildirim

Universidade Federal do Rio de Janeiro:
- Carlos Suarez

Mutation in Small vs. Large Genomes

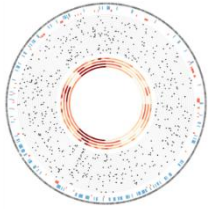


Index:

Outer Rings
- Gene Density
- High G/C Region
- High A/T Region

Intermediate Rings
- Mutations

Inner Rings
- Mutation Density
  - Window Size (1k, 5k, 25k, 100k)



Bacillus subtilis 3610

Genome size: 4,214,598 bp

GC content: 43.5%

50 lines - 450 mutations - 5000 generations

Mutation Rate : 3.27×10^{-10}/site/gen.

Mesoplasma florum L1

Genome size: 793,224 bp

GC content: 27.0%

50 lines – 599 mutations - 2000 generations

Mutation Rate : 1.14×10^{-8}/site/gen.

Drake's (1991) Conjecture:
A Constant Rate of Mutation per Genome per Cell Division in Microbes

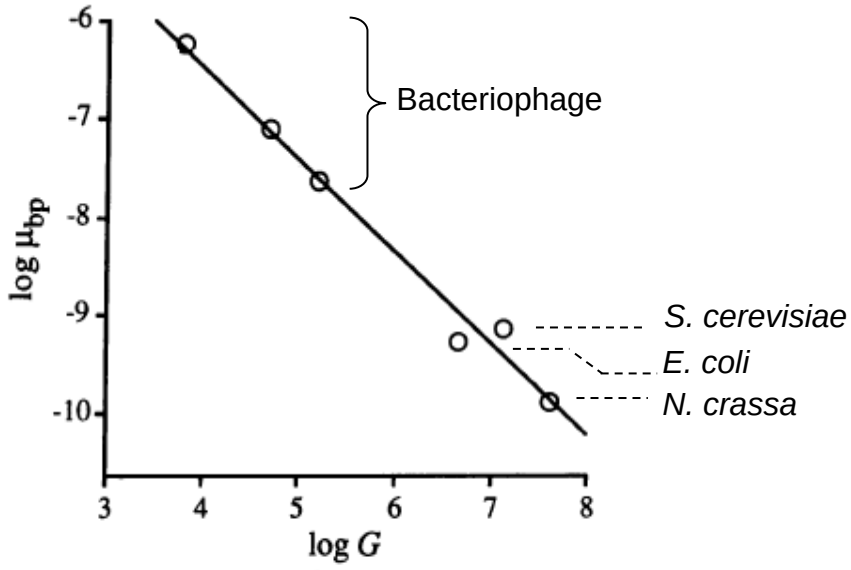
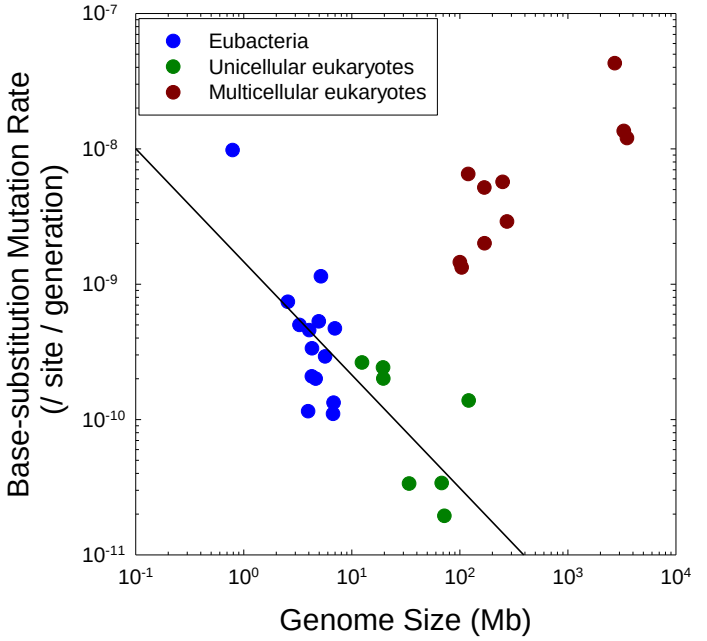


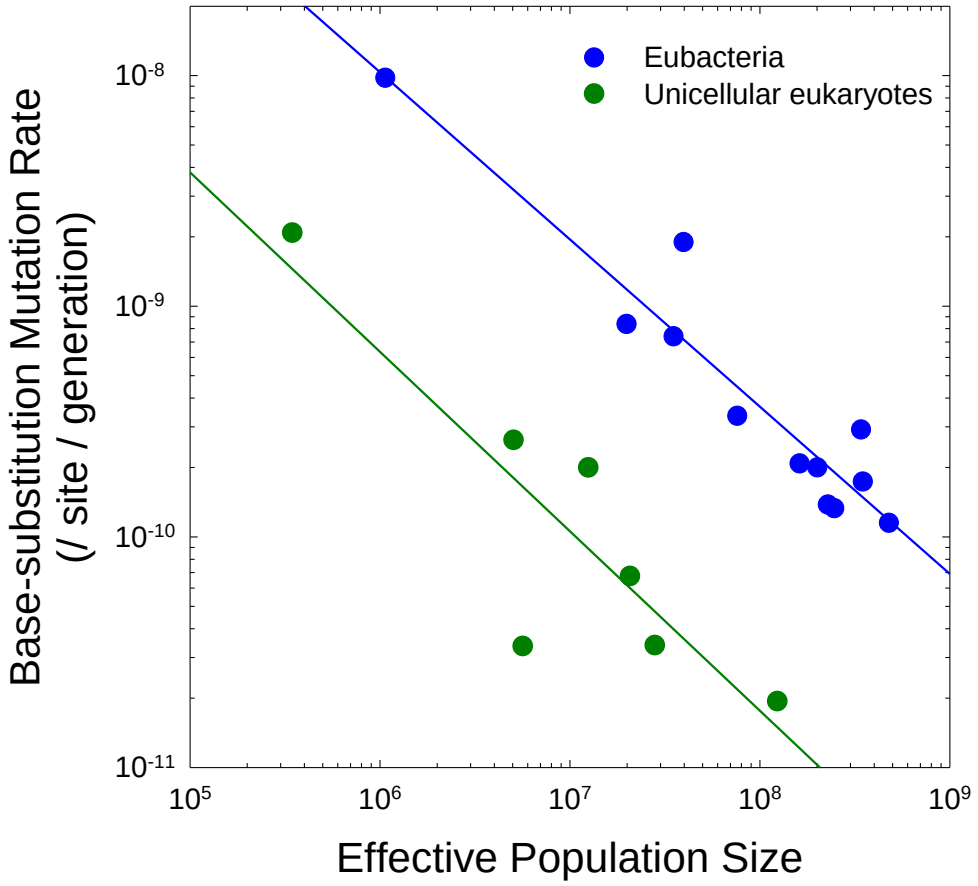
FIG. 1. Average mutation rate μ_{bp} per base pair as a function of genome size G in bp. The logs of the rates for each organism were averaged and all 13 values are included. Phages T2 and T4 were treated as a single organism.



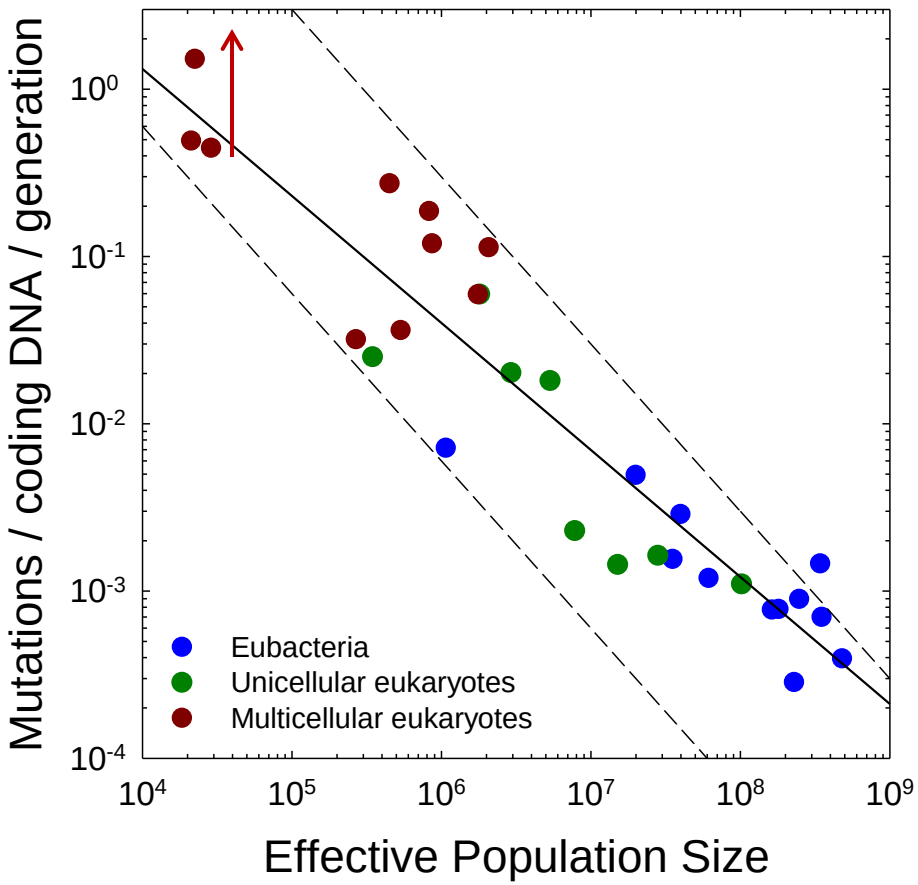
"Because this rate is uniform in such diverse organisms, it is likely to be determined by deep general forces."

The Mutation Rate / Nucleotide Site Is Inversely Proportional to the Average Effective Population Size of a Species

For a given magnitude of genetic drift, selection is capable of driving the mutation rate down further in eukaryotes than prokaryotes.



A Universal Inverse Scaling Between the Genome-wide Deleterious Mutation Rate and N_e Across the Tree of Life



- The mutation rate per nucleotide site scales inversely with both the effective population size and the amount of functional DNA in the genome (the total target size for deleterious mutations).

$uG_e \sim 1 / N_e \quad \rightarrow \quad u \sim 1 / (G_e \cdot N_e)$

u = mutation rate / site / generation

G_e = amount of functional DNA (sites)

N_e = effective population size

Eubacteria
Unicellular eukaryotes
Multicellular eukaryotes

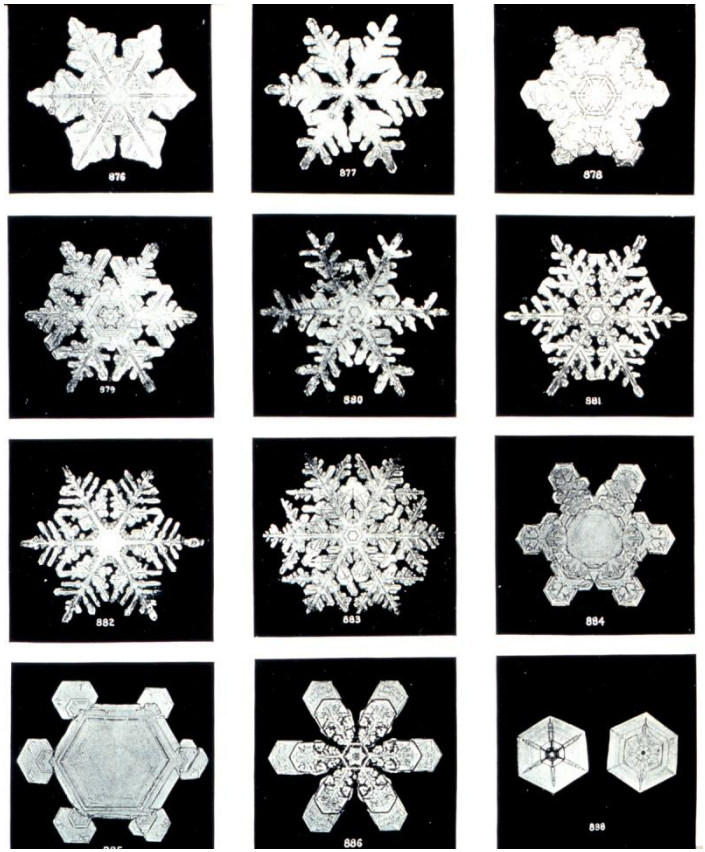
SUMMARY

- Replication fidelity is the only trait for which we have detailed phenotypic measurements across the entire Tree of Life.

- Mutation-rate evolution appears to obey scaling laws based on fundamental population-genetic principles, most notably the power of random genetic drift.

- If evolutionary cell biology is to advance beyond comparative biology and adaptive story telling, these principles will need to be explored with other cellular features.

Mesmerizing Beauty, Diversity, and the Adaptationist Paradigm

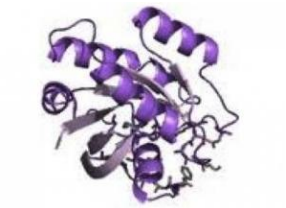


"….. from so simple a beginning endless forms most beautiful and most wonderful have been, and are being, evolved." Charles Darwin

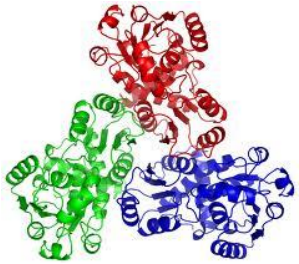
The Origin of Variation in Molecular Complexes:

Driven by adaptive processes unique to individual lineages?

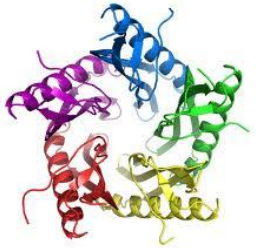
Or a consequence of biased mutation pressure and biophysical factors?



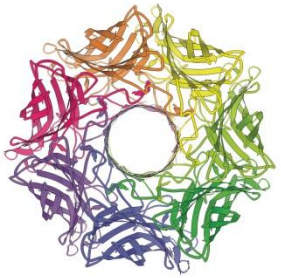
monomer



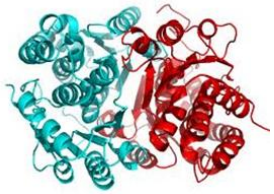
trimer



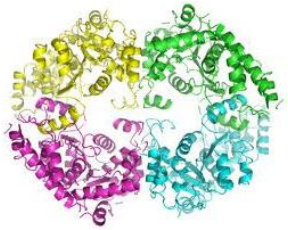
pentamer



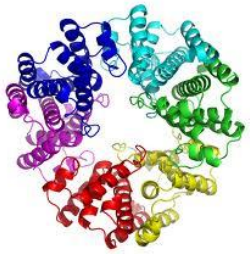
heptamer



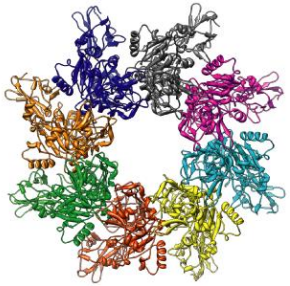
dimer



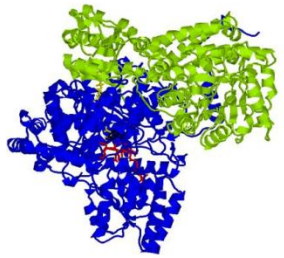
tetramer



hexamer



octamer

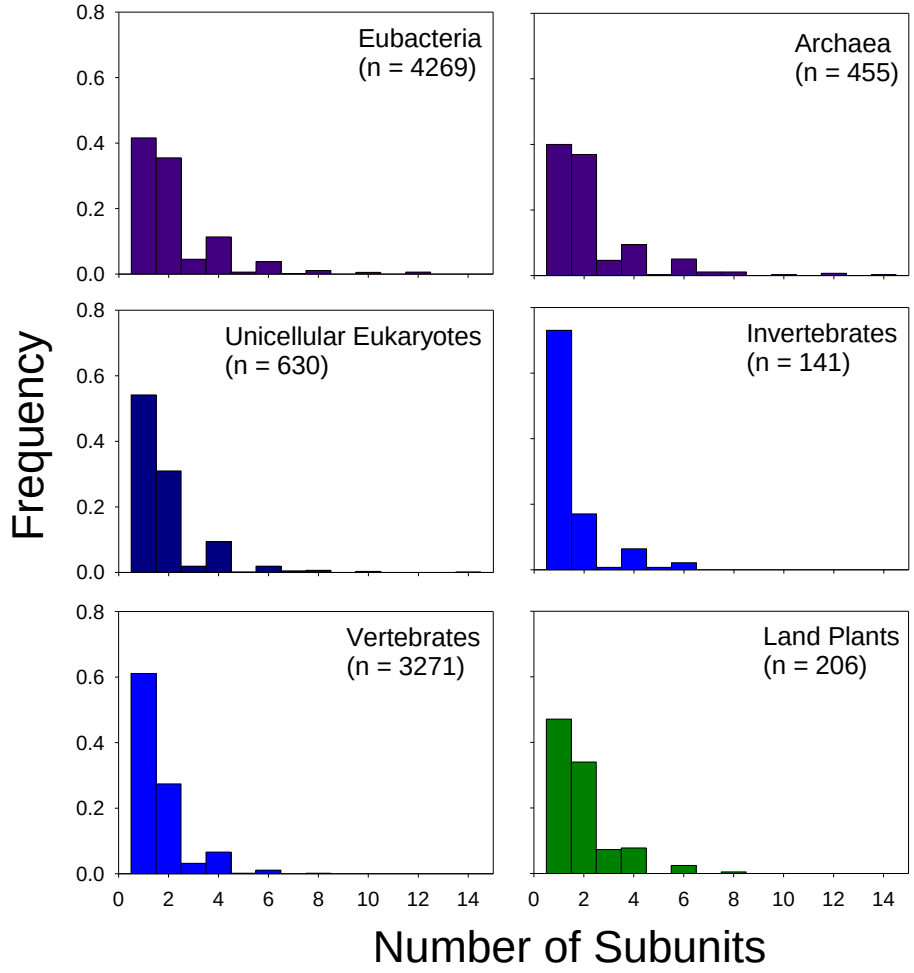


- **Potential advantages to complex formation**:
  - increased structural diversity,
  - reduced surface area increases productive encounter rate with substrate,
  - reduced problems of folding single large proteins,
  - reduced vulnerability to denaturation and/or engagement in promiscuous interactions,
  - reduced molecular motion at the catalytic site increases substrate specificity,
  - increased flexibility for allosteric regulation.

- **Compensation for structural deficiencies in monomeric subunits?**

- **Proteins with an affinity to oligomerize also come at a cost:**
  - Elevated production levels necessary for a critical encounter rate for successful multimerization.
  - Problems with harmful interactions between heterotypic molecules in heterozygotes in the establishment phase.

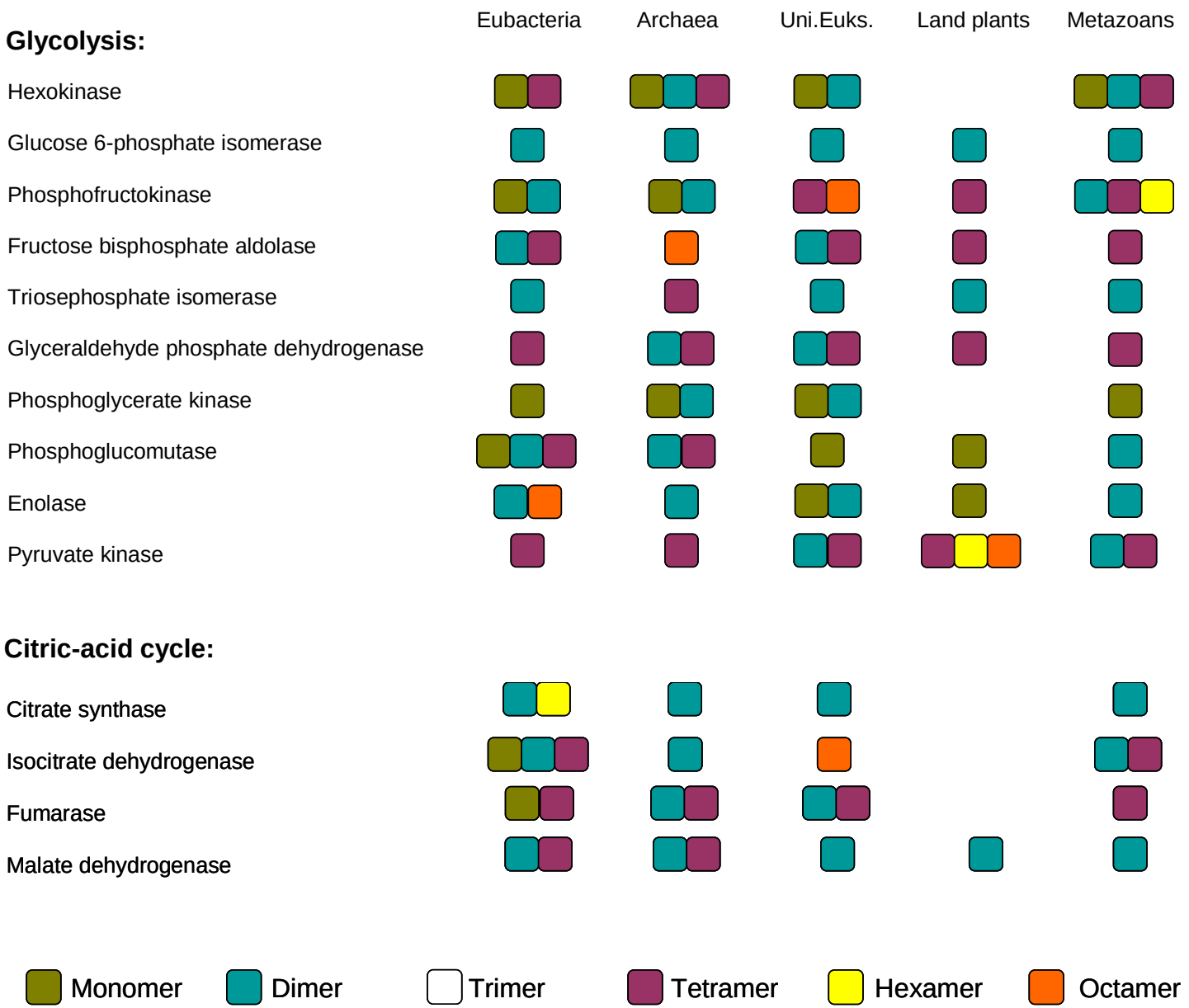
Distribution of Homomeric Types: approximate constancy across the Tree of Life.



- Roughly two thirds of proteins are multimeric, independent of phylogenetic lineage.
- Roughly two thirds of multimers are dimers.
- ~15% are tetramers, most of which are “dimers of dimers,” most likely arising via an intermediate dimeric state.
- Odd-mers are greatly under-represented.

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

Known Oligomerization Structures for the Enzymes of Central Metabolism



Enzymes with Identical Multimeric States Need Not Have the Same Structural Basis

Dihydrodipicolinate synthase (involved in lysine synthesis)

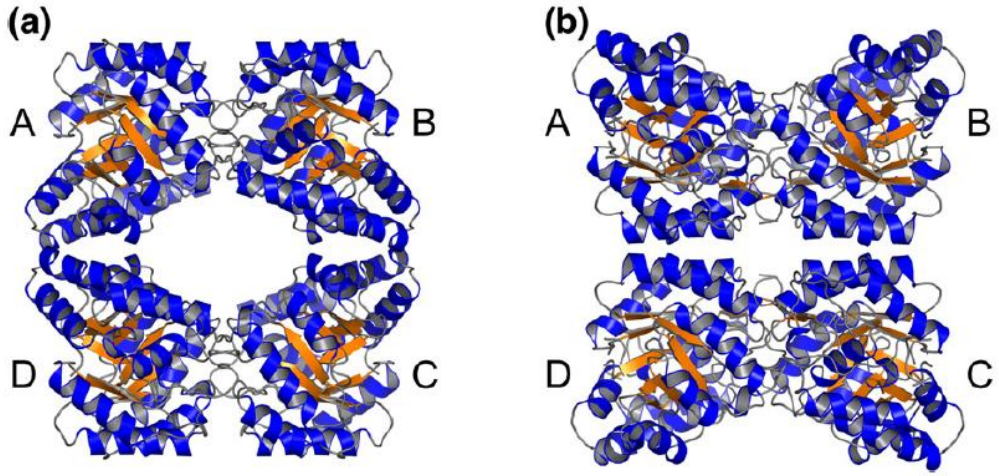
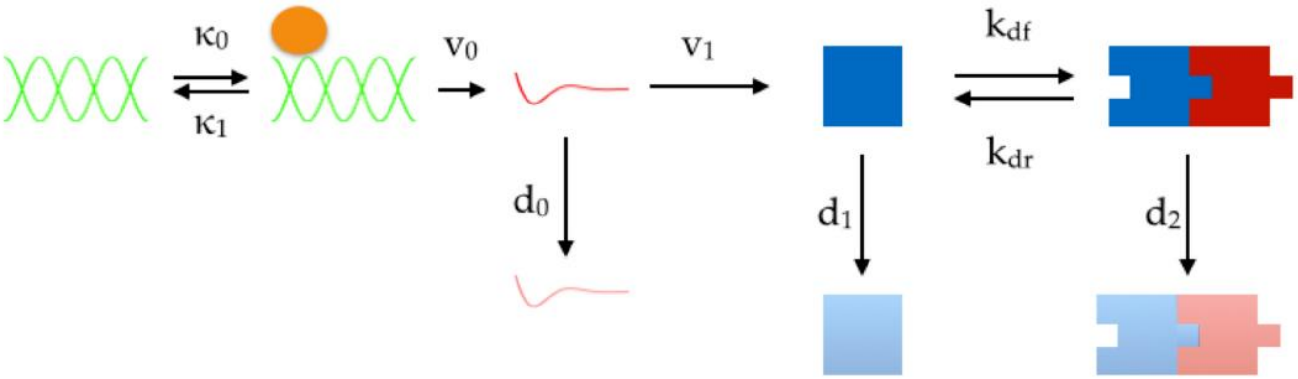


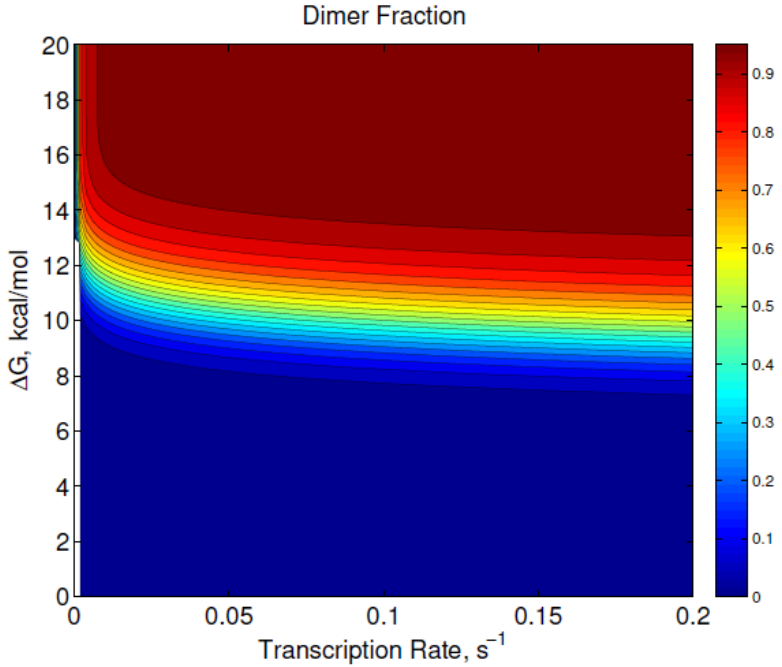
Fig. 1. The X-ray crystal structures of DHDPS from (a) *E. coli*[18,19] and (b) *N. sylvestris*.[23] Each enzyme is a homotetramer of $(\beta/\alpha)_8$-barrels composed of two tight-dimer units (A–B and C–D), but the arrangement of the two dimeric units is different.

Both species make homotetramers, but the dimer-dimer interfaces are completely nonoverlapping, face to face in the former, and back to back in the latter (Griffin et al. 2008).

A Biophysical / Biochemical Approach to the Problem



Kyle Hagner

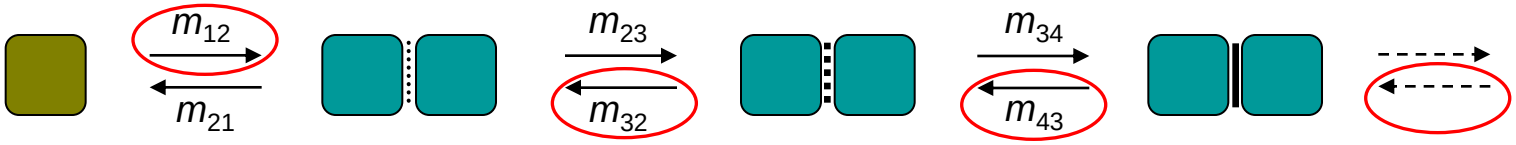


How to define fitness?

What is the proper form of a neutral model?

$$W = f\left(\tilde{M} + w_D \tilde{D}\right) \equiv f\left(\tilde{E}\right)$$

Evolution of a Dimeric Structure

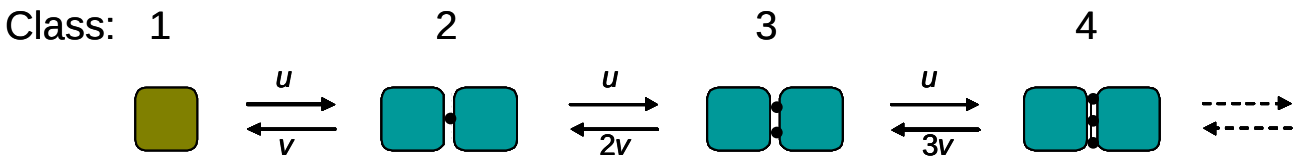


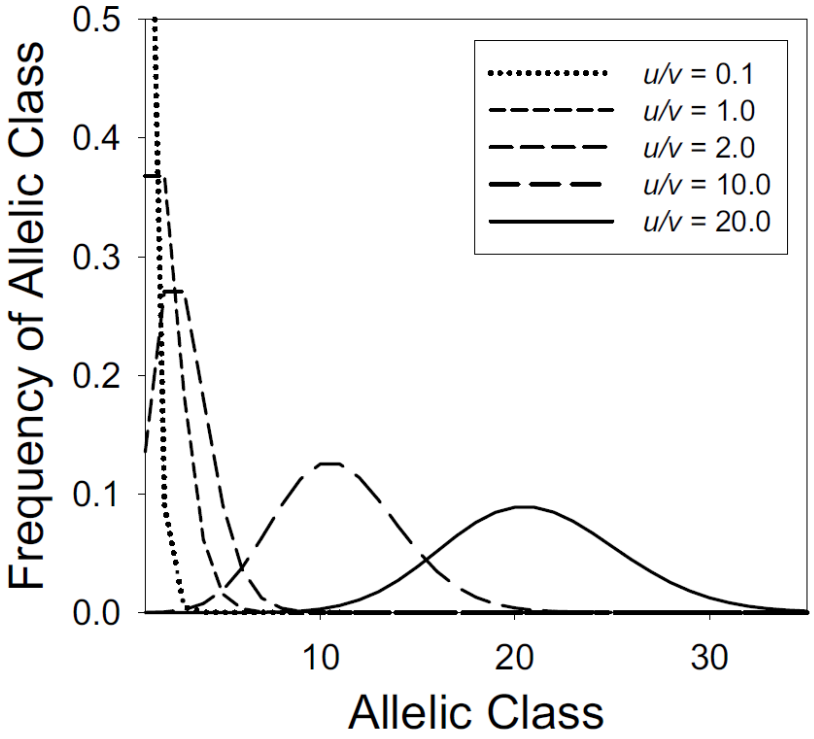
- Each transition rate is equal to the product of the number of relevant mutations arising per generation and the fixation probability.

- At steady state, the flux rate must be equal in both directions. This means that the net rate of establishment of dimers from monomers must equal the reverse rate.

- The equilibrium probability of each state is simply proportional to the product of the total set of transition rates towards the state from both directions.

The Neutral Expectation: the steady-state distribution of alternative allelic states is Poisson, a simple function of the ratio of upward and downward mutation rates, independent of population size.





$$\widetilde{P}_i = \frac{(u/v)^{i-1}}{(i-1)!C}$$

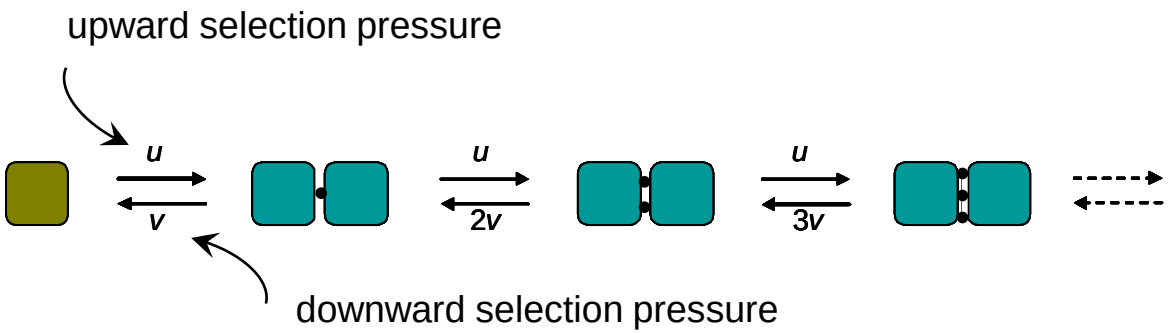
$$C = \sum_{i=0}^{\infty}(u/v)^i/i! = e^{u/v}$$

u/v is the mutation bias.

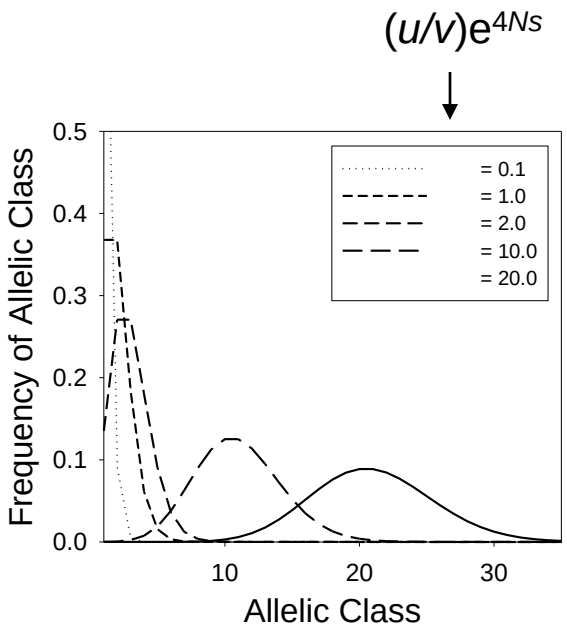
Expected frequency of monomers = $e^{-u/v}$

Adding in Selection:

- *s* is the selective advantage (or disadvantage) of each incrementing allele.
- e^{4Ns} is the ratio of fixation probabilities for beneficial vs. deleterious mutations.
- 4*Ns* is the ratio of the power of selection to random genetic drift.



- The distribution is again Poisson, but now the key parameter is $(u/v)e^{4Ns}$.
- The effects of selection, drift, and mutation bias cannot be disentangled from observations on the steady-state distribution alone.



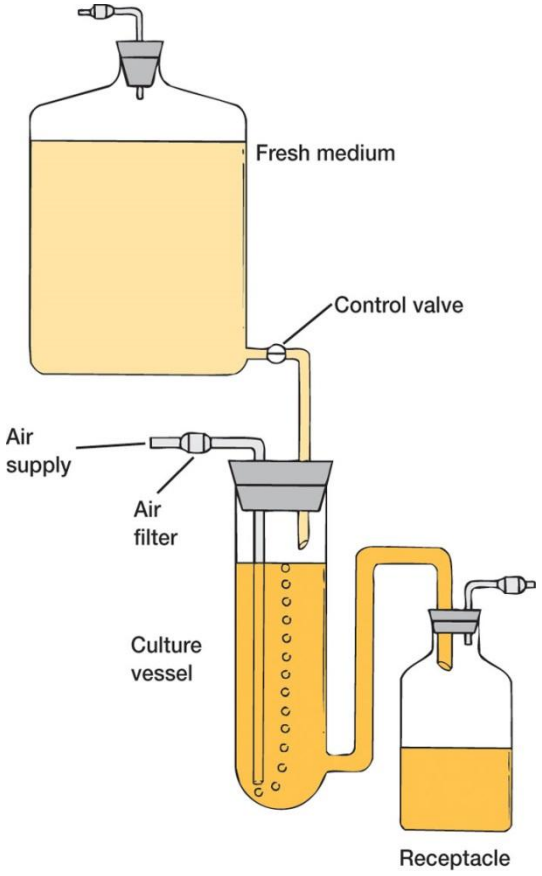
General Conclusions on Multimer Evolution

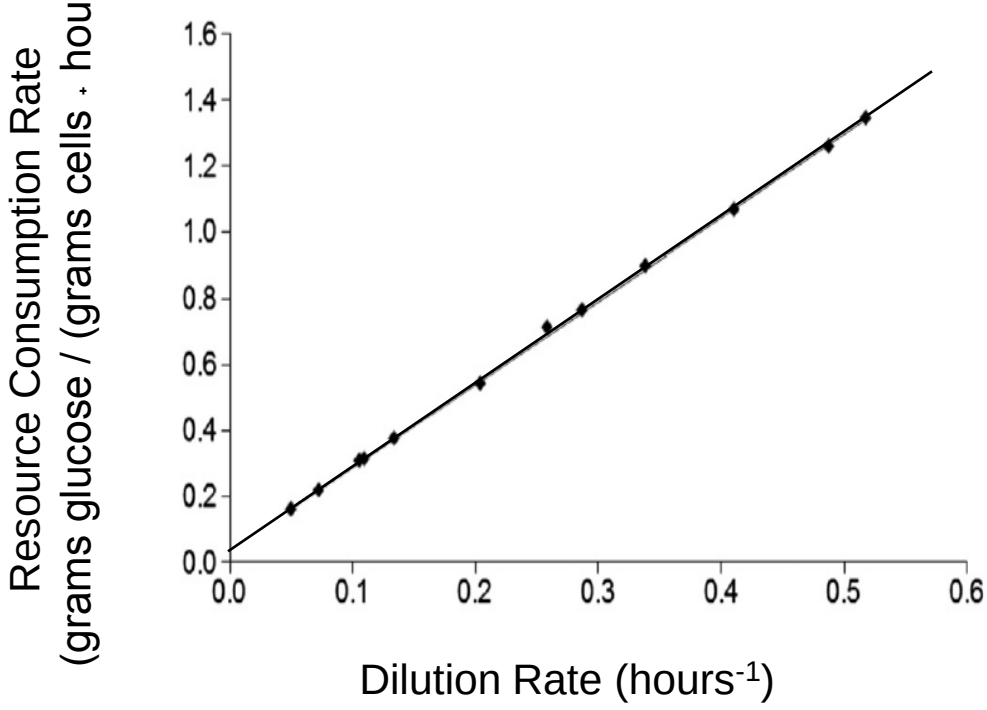
- Substantial phenotypic variation can arise among lineages, **even when selection and mutation is operating in an identical manner in all lineages.**

- The most common molecular state is not necessarily the optimum – **even with *negative* selection against multimers, they will still be common provided the mutational bias towards binding affinity is sufficiently large.**

- If the ratio of the power of selection and drift is < 1.0, the phenotypic distribution is entirely driven by mutation bias – **effective neutrality**.

Measuring Cellular Maintenance and Growth Requirements With a Chemostat





Effectively Neutral Evolution at the Level of Cellular Features?

- Complete rewiring of regulatory pathways (transcription factors and their binding sites) in different yeast species – ribosomal proteins; mating type; galactose utilization.

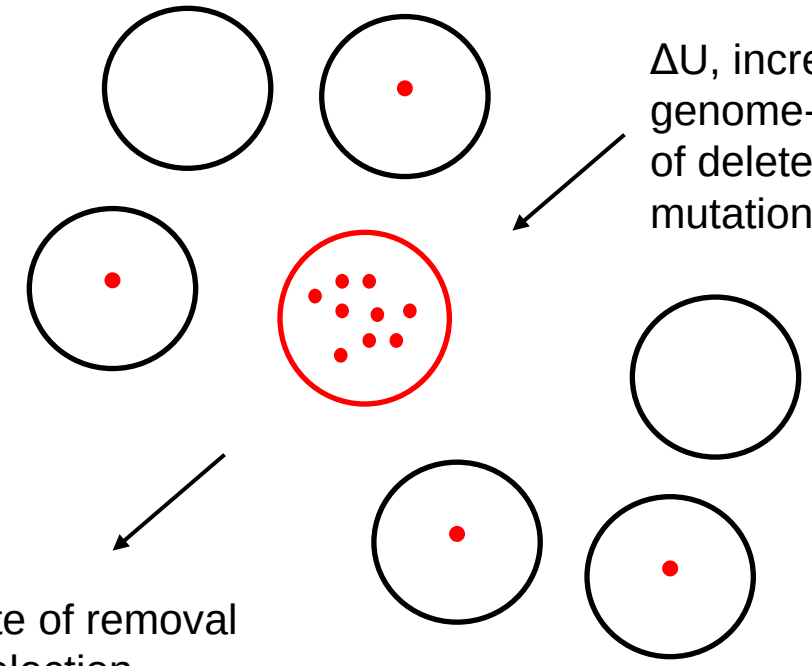
- Transcription-factor binding site variation in sequence motifs is extensive.

- Enzyme reaction rates are orders of magnitude lower than the diffusion limit, and enzyme promiscuity is the rule.

- Variation in the multimeric nature of proteins is independent of organismal complexity.

- Replication fidelity is reduced in species with smaller effective population sizes.

The Magnitude of Selection Operating to Improve Replication Fidelity



ΔU, increase in genome-wide rate of deleterious mutation

s, rate of removal by selection

Excess number of mutations at equilibrium = ΔU / s

↓

X

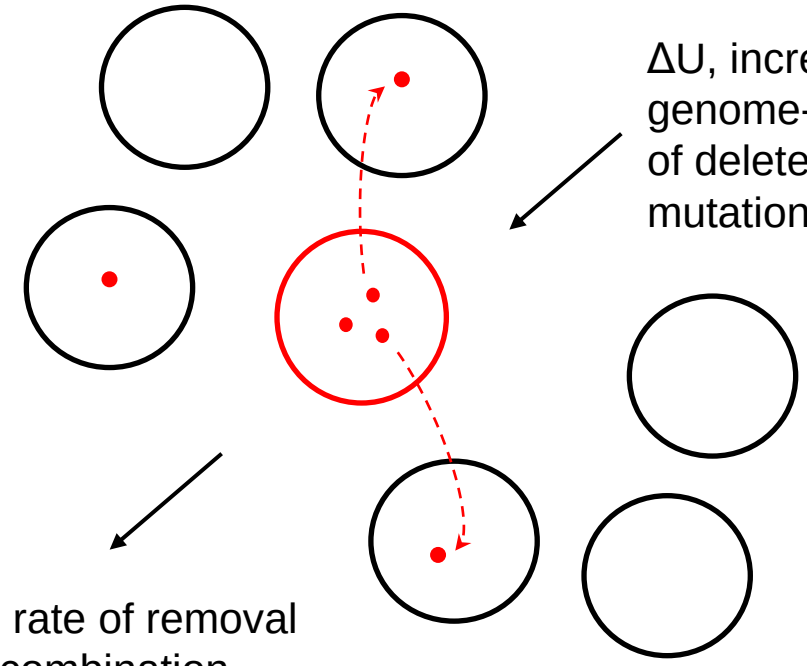
Effect / mutation = s

↓

Total effect on fitness = ΔU

- Selective disadvantage of a mutator in an asexual population
  = increase in genome-wide deleterious mutation rate

The Force of Selection to Improve Replication Fidelity is Greatly Reduced in Sexual Populations



ΔU, increase in genome-wide rate of deleterious mutation

~1/2, rate of removal by recombination

Excess number of mutations at equilibrium = ΔU / (1/2)

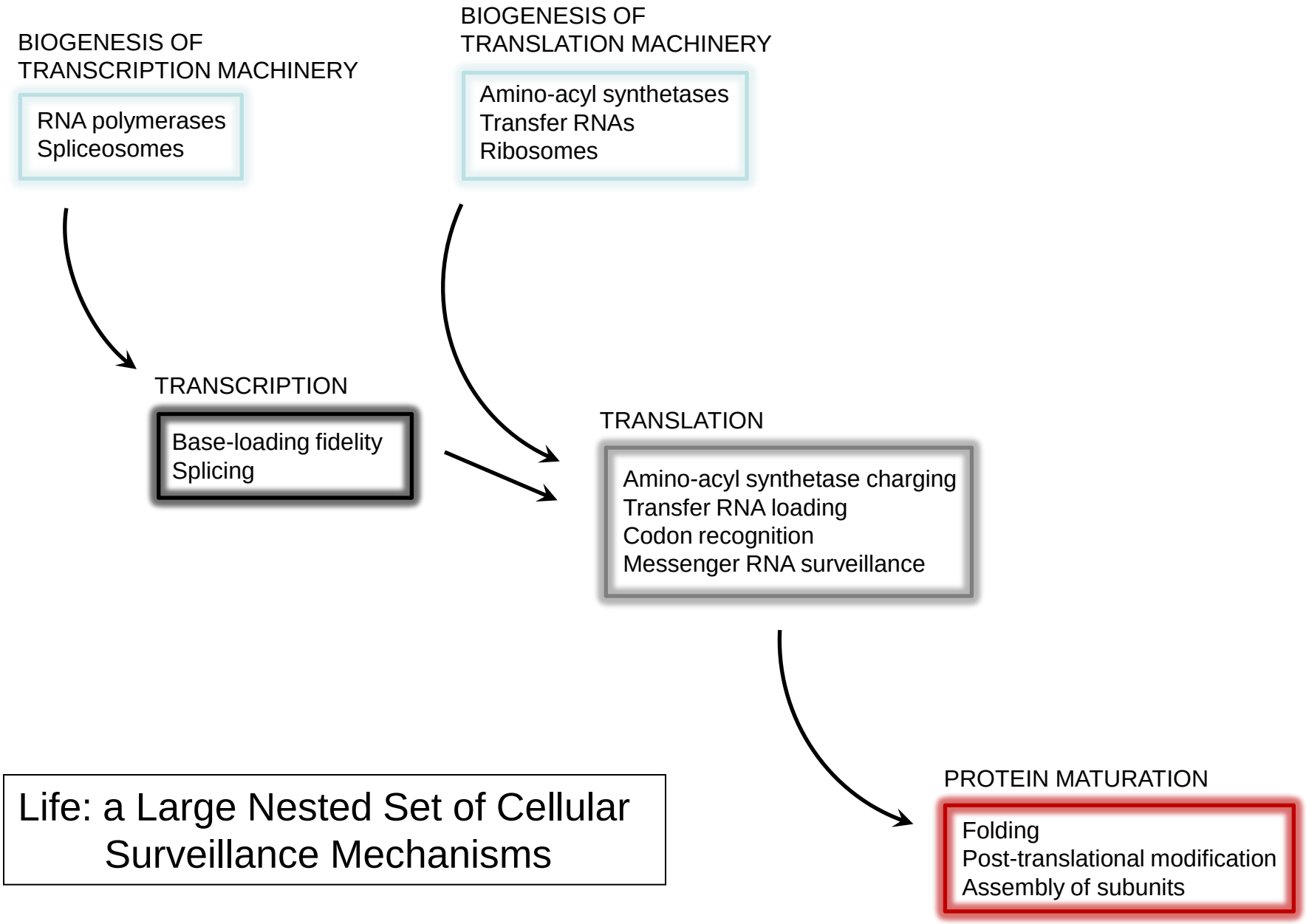
↓

X

Effect / mutation = s

↓

Total effect on fitness = 2 s ΔU


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
Life: a Large Nested Set of Cellular
Surveillance Mechanisms

Selection on the Replication Error Rate in Sexual Populations:

the selective disadvantage of a mutator allele is $\Delta u \cdot 2 \cdot G_e \cdot s$

Mutations remain linked to a mutator allele for an average of 2 generations

Number of nucleotides in the genome subject to selection

Heterozygous effect of a deleterious mutation

Selection on the Transcription Error Rate:

selective disadvantage of a transcriptional mutator is $\Delta u \cdot 1 \cdot T_e \cdot s \cdot d$

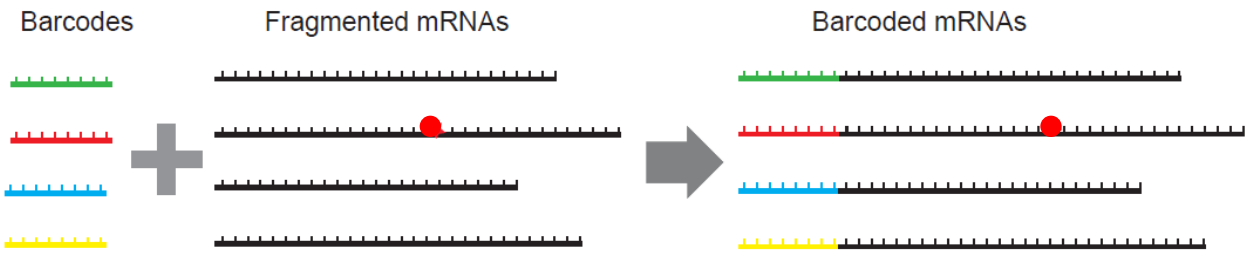
The pool of errors remains associated with the mutator for just one generation

Number of nucleotides in the transcriptome subject to selection

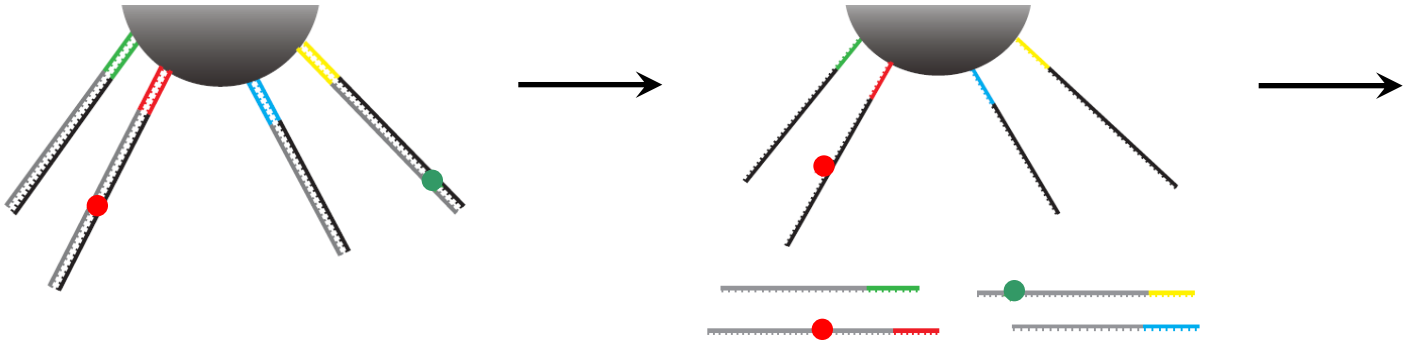
Heterozygous effect of a deleterious mutation

Dilution effect (<< 1.0)

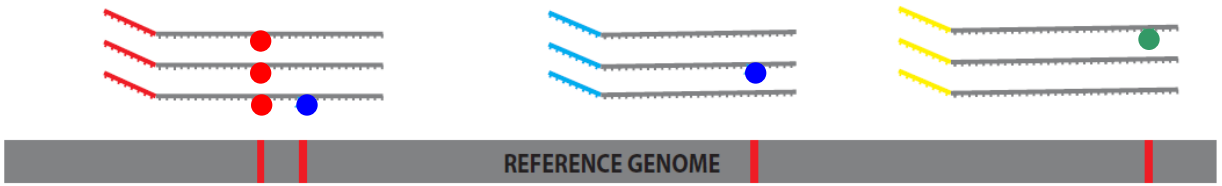
Estimation of the *in vivo* Transcription-error Rate From an RNA Library (Gout et al., PNAS, 2013)



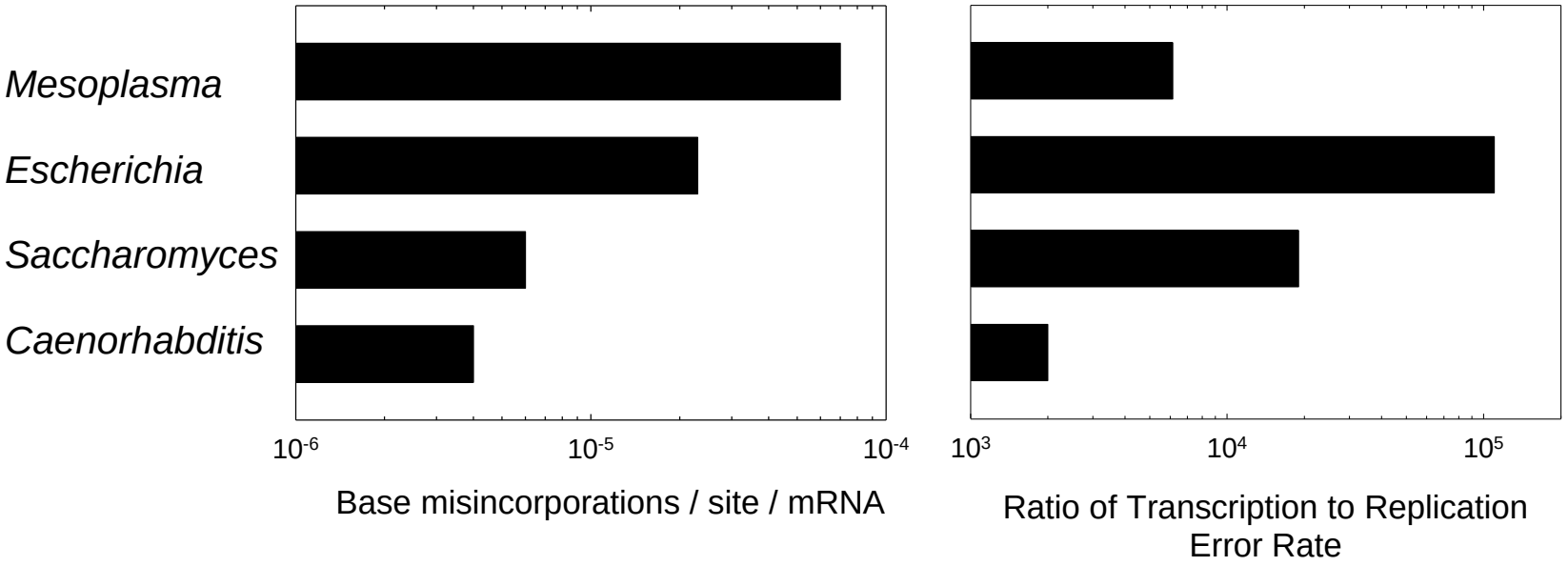
Capture fragments on beads; reverse transcribe; isolate cDNAs; repeat to obtain replicates:



Sequence to high depth; sort into uniquely coded families; search for consistent errors;

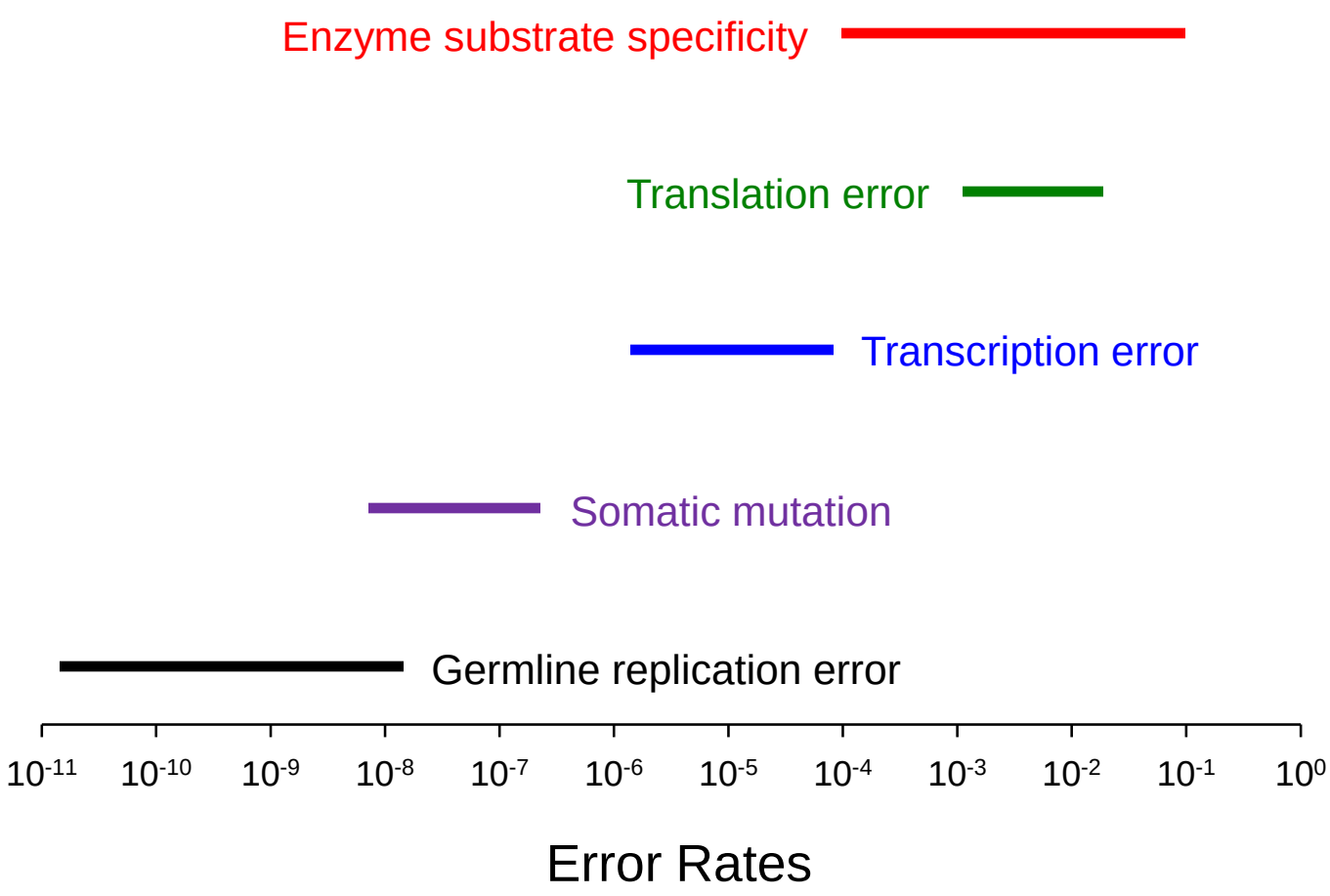


Transcription-error Rates Are Orders of Magnitude Higher Than Replication-error Rates

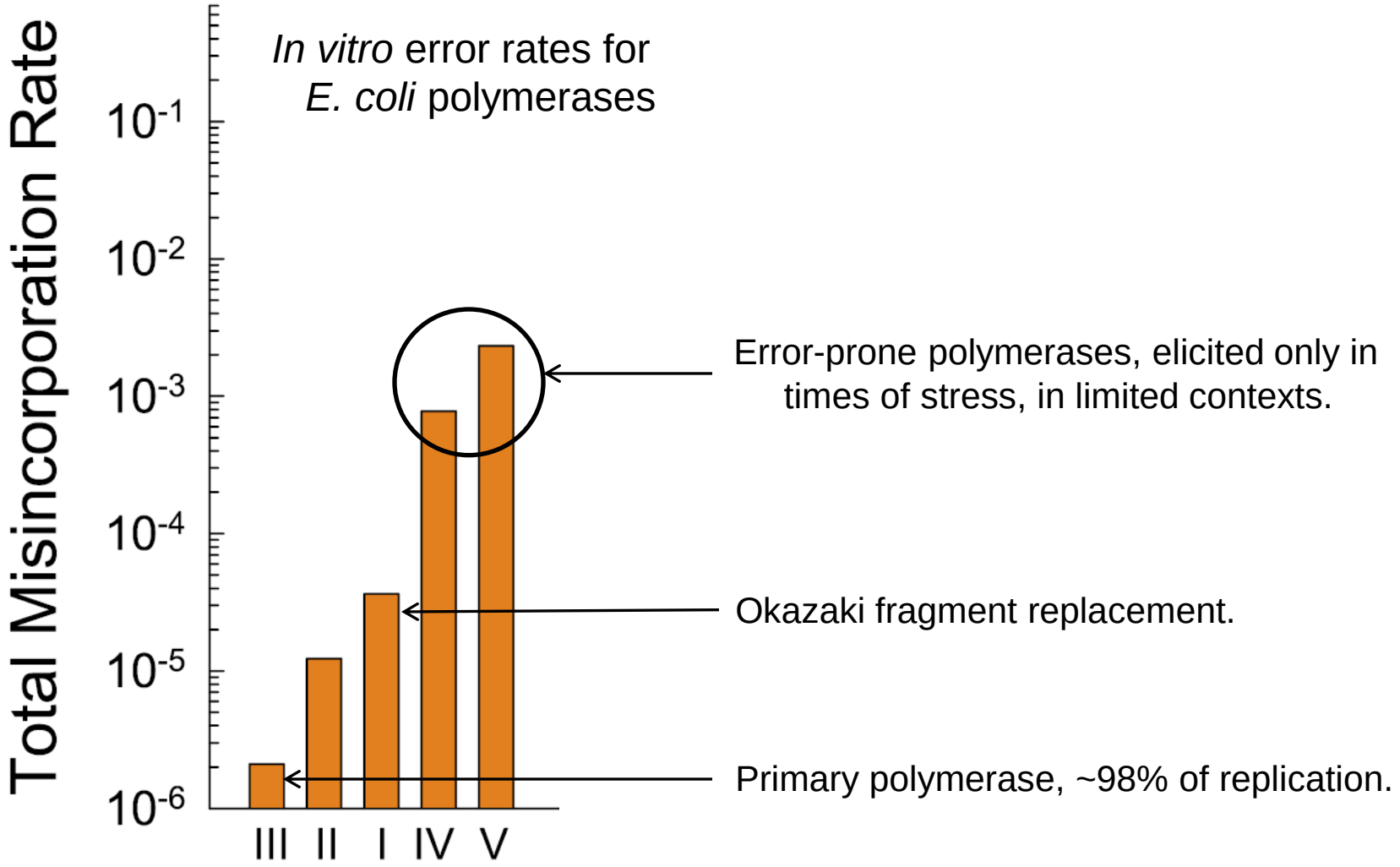


- ~1 to 5% of transcripts contain errors

Error Rates Roughly Scale With the Degree of Transience of Effects



Error Rates Are Magnified in Polymerases Involved in Fewer Nucleotide Transactions

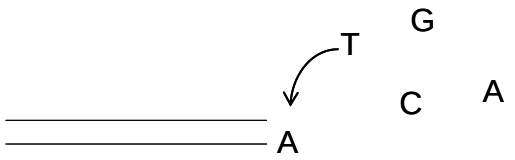


Evolutionary Layering and the Limits to Molecular Perfection:

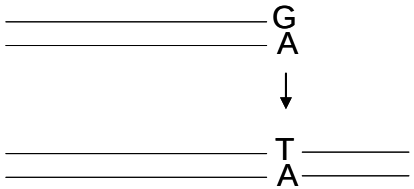
1) Can a secondary layer of defense be added that breaks the drift barrier?

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

1) Polymerase base-incorporation fidelity:

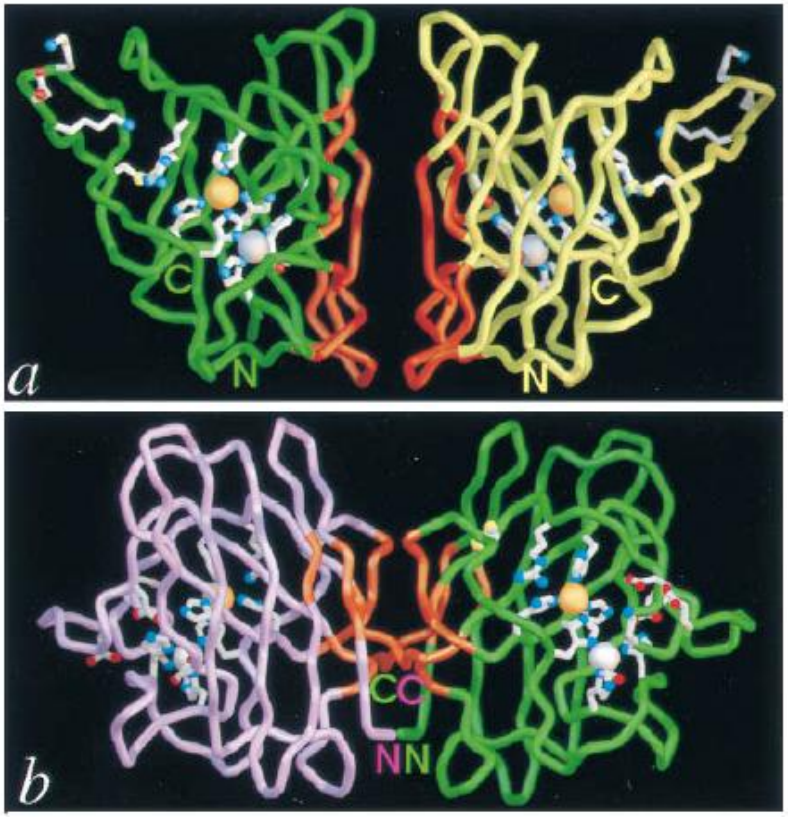


2) Polymerase proofreading:



Cu,Zn Superoxide Dismutase:

Dimer interfaces in *Photobacterium* (above) and cow (below) are constructed from diametrically opposite beta-barrel elements (Bourne et al. 2008).



- Dayhoff et al. (2010) estimate that about two-thirds of protein families containing homomers exhibit phylogenetic variation in the binding interfaces.

Comparative biology alone does not equate to evolutionary biology – it only tells us what questions are worth asking.

Nevertheless, if we are to understand how protein complexes and higher-order cellular features evolve, we must start with a comparative analysis of organisms with close enough relationships that the likely steps of cellular divergence can be deciphered and, ideally, reconstructed and studied in an experimental setting.

- Only with closely related taxa is it possible to order the history of single-step mutations and their cumulative consequences for complex traits. Unfortunately, most of today's cell biology is restricted to just a tiny fraction of cellular diversity, and the small number of existing model species are so divergent that there is no hope of confidently reconstructing ancestral states.

- Numerous cell biological features are known where moderately related lineages deploy nonorthologous proteins for the same function, e.g., amino-acid synthesis pathways, licensing of DNA replication origins, and regulation of histones and ribosomal proteins.

How often is structural homology at the cellular level not matched by orthology of the underlying genetic architecture, i.e., because of convergent evolution or underlying gene replacement?

Mechanisms of Evolution

Cells occupy a location in the hierarchy of life that is pivotal to understanding the mechanisms of evolution.

- The further a biological feature is from the target of selection (the phenotype), the more likely it is to be influenced by nonadaptive mechanisms of evolution.
    - The diversification of a wide variety of genomic features (e.g., introns and intergenic spacer DNA) among lineages appears to have arisen by differential forces of mutation and random genetic drift, which can sometimes completely overwhelm the power of selection.
    - Plausible arguments have been made that various aspects of cellular infrastructure (e.g., the ribosome) may have also originated by effectively neutral processes.
    - Although it is easy to marvel at the numerous features devoted to surveillance of internal cellular problems and their contribution to organismal robustness (e.g., DNA-replication proof-reading, decay of erroneous mRNAs, and chaperone guidance of protein folding), the establishment of layers of complexity need not have any long-term benefit. One obvious disadvantage of a complex feature is that it is a larger target for mutational inactivation relative to a simpler trait carrying out the same task.

Thus, a major challenge for evolutionary biology is to determine the extent to which the infrastructure upon which organisms are built is driven by adaptive vs. nonadaptive processes, or combinations thereof.

Resolution of these issues will play a central role in the field of ECB for the simple reason that confidence in any adaptive arguments for the evolution of cellular features must remain suspect unless the hypothesis of effective neutrality can be ruled out.

The Limits to Evolutionary Perfection

Are the constraints on the evolved levels of molecular perfection a function of population genetics or biophysics?

The argument is often made that selection is capable of refining molecular attributes until they encounter constraints imposed by principles of physics and/or chemistry (e.g., Albery and Knowles 1976).

But it is also known that once an adaptation approaches a high level of refinement, further improvements can be thwarted by the power of random genetic drift, which scales with the inverse of the effective population size (Hartl et al. 1985; Lynch 2011).

This raises the possibility that the types of cellular change that are open to evolutionary exploration are defined by the population-genetic environment.

- Many cellular processes have very high error rates (e.g., 10^{-5} to 10^{-4} per nucleotide for transcription).
- Many proteins exhibit substantial promiscuity in function.
- The efficiency of most enzymatic reaction rates is orders of magnitude below the diffusion limit.

How much of cellular evolution arises as a response to challenges from the external environment, as opposed to internal genomic / cellular threats?

- The massively complex spliceosomes.
- The nuclear envelope and the pore constituents.
- Messenger RNA surveillance mechanisms – nonsense-mediated decay.
- The evolution of multimeric proteins – novel function or making the best of compromised capacities of individual molecules.
- Reliance on chaperones for proper protein folding.
- The evolution of meiosis and the origin of selfish centromeres.

Some useful **theoretical** starting points:

- Models for the evolution of complex adaptations help define how the rates of acquisition depend on key population-genetic parameters – rates of mutation, recombination, and random genetic drift.

- We know how these primary “nonadaptive” forces of evolution scale with organism size.

- Models for the evolution of duplicate genes.

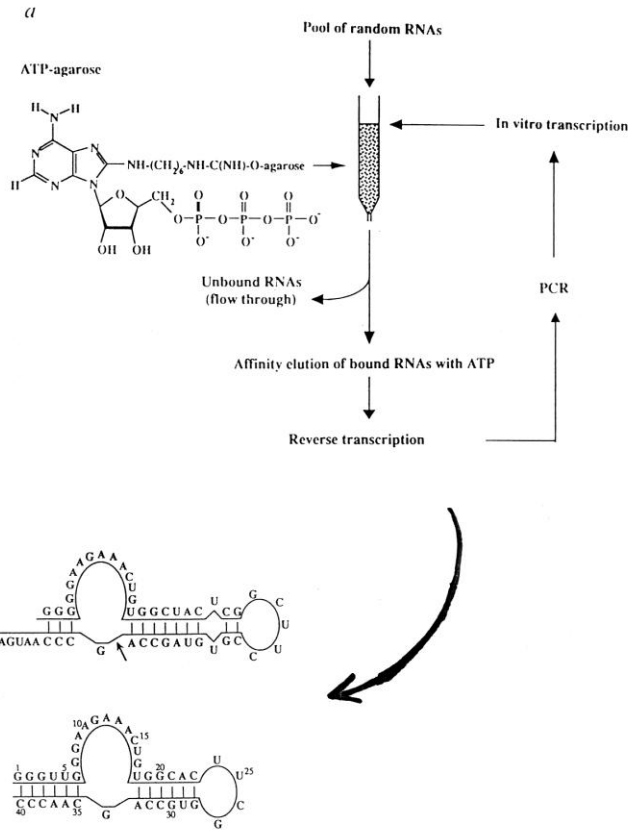
- Models for the emergence of protein-protein interactions (networks).

Some useful **empirical** starting points:

- Microfluidics provides a powerful platform for monitoring the response to selection in well-defined contexts.
- How does the outcome of selection depend on population size, mutation, and recombination rates.
- How replicable are the solutions to a specific challenge.

In Vitro RNA Evolution Experiment

Example. Selection for an RNA motif that binds ATP.
M. Sassanfar and J. W. Szostak. 1993. Nature 364: 550-553.



Three Levels for the Cost of a Gene:

1) **Chromosome**: synthesis of nucleotides for replication, and amino acids for nucleosomes.

2) **Transcription**: synthesis of ribonucleotides for steady-state number of transcripts.

3) **Protein**: synthesis of amino acids for steady-state number.

- All measured relative to the total energy budget of the cell in units of numbers of ATP hydrolyses.

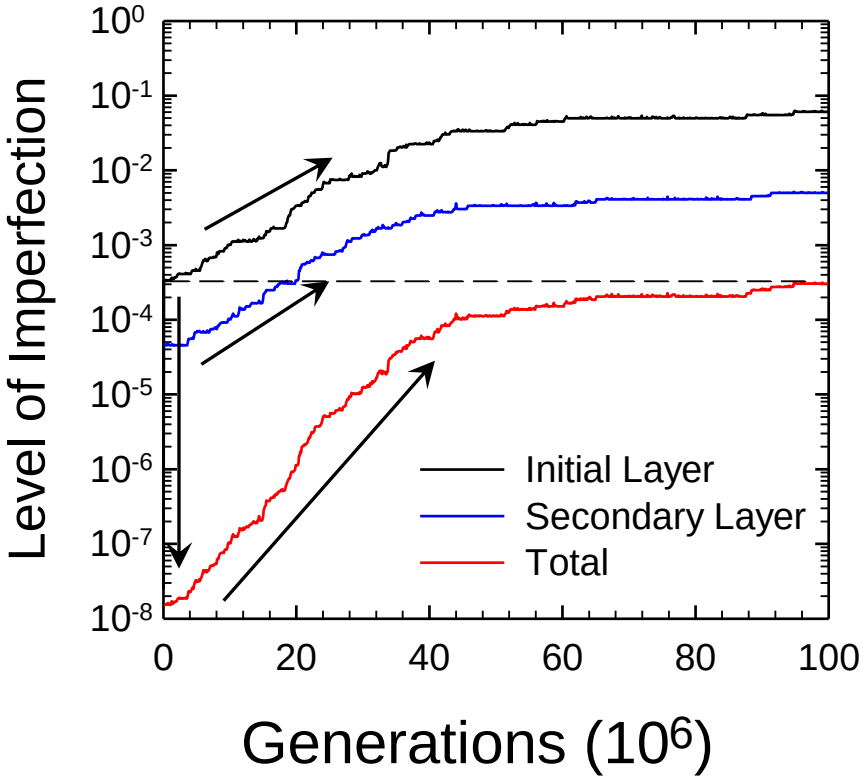
Evolutionary consequences:

Total baseline cost: $s_c = s_{DNA} + s_{RNA} + s_{PRO}$

Net selective advantage: $s_n = s_p - s_c$

- If $|s_i| < 1/N_e$, selection is unable to eradicate or promote the gene.

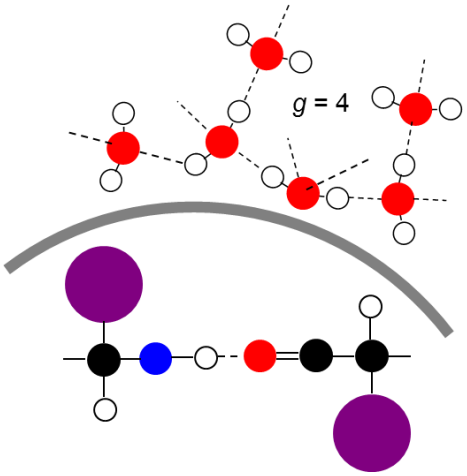
The Fitness Boost From the Addition of a Layer of Surveillance Is Transient



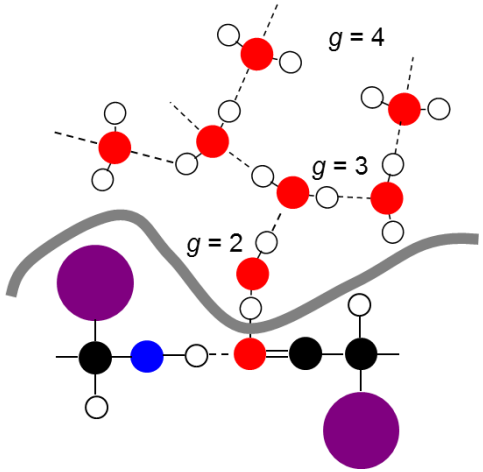
- Rapid improvement accompanies establishment of a new layer of protection.
- Both layers then gradually become less efficient.
- The level of overall performance returns to that for the single-layered state.

- 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.

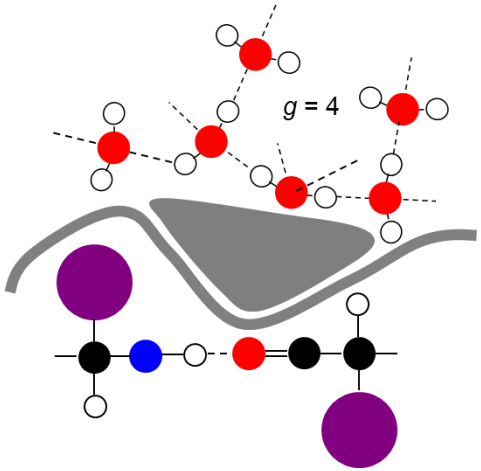
Can Nonadaptive Processes Lead to the Evolution of Protein Complexity?



Well-protected



Exposed



Tension relieved