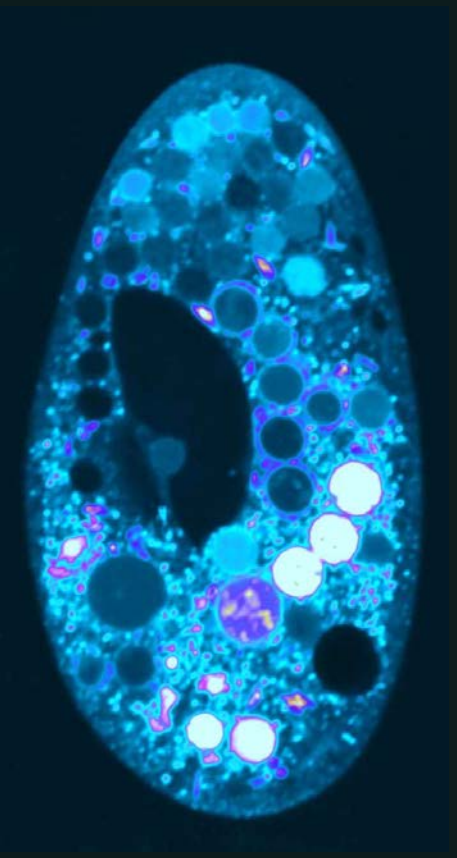
Evolution of the Mutation Rate



- The mutation rate is the only trait for which we have a general theory of evolution across the Tree of Life.

- The mutation rate scales across phylogenetic groups, among tissues, and among polymerases within cells, in predictable ways with an ~10^4-fold range of variation.

- No evidence that mutation rates have been optimized to maximize the long-term rate of adaptive evolution.

- No evidence that the replication fidelity has been pushed to the limits of molecular perfection in any species.

- **The Drift Barrier to mutation-rate reduction**: Once the selective advantage of lowering the mutation rate is less than the power of drift, $1/(2N_e)$, the mutation rate has reached its minimum possible value.

- Nothing exceptional about the current human mutation rate.

Drake's (1991) Conjecture:
A Constant Rate of 0.003 Mutations per Genome per Cell Division in Microbes

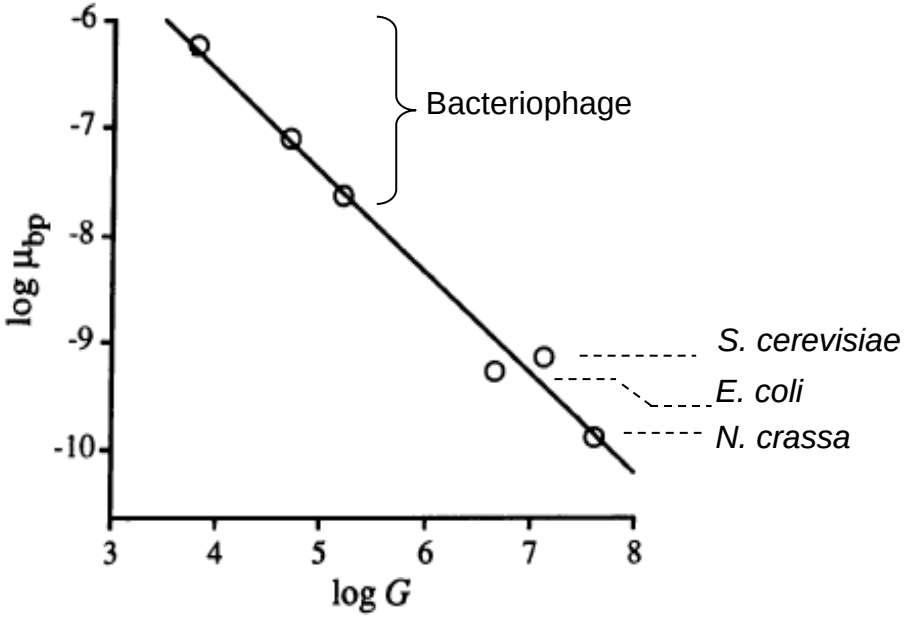
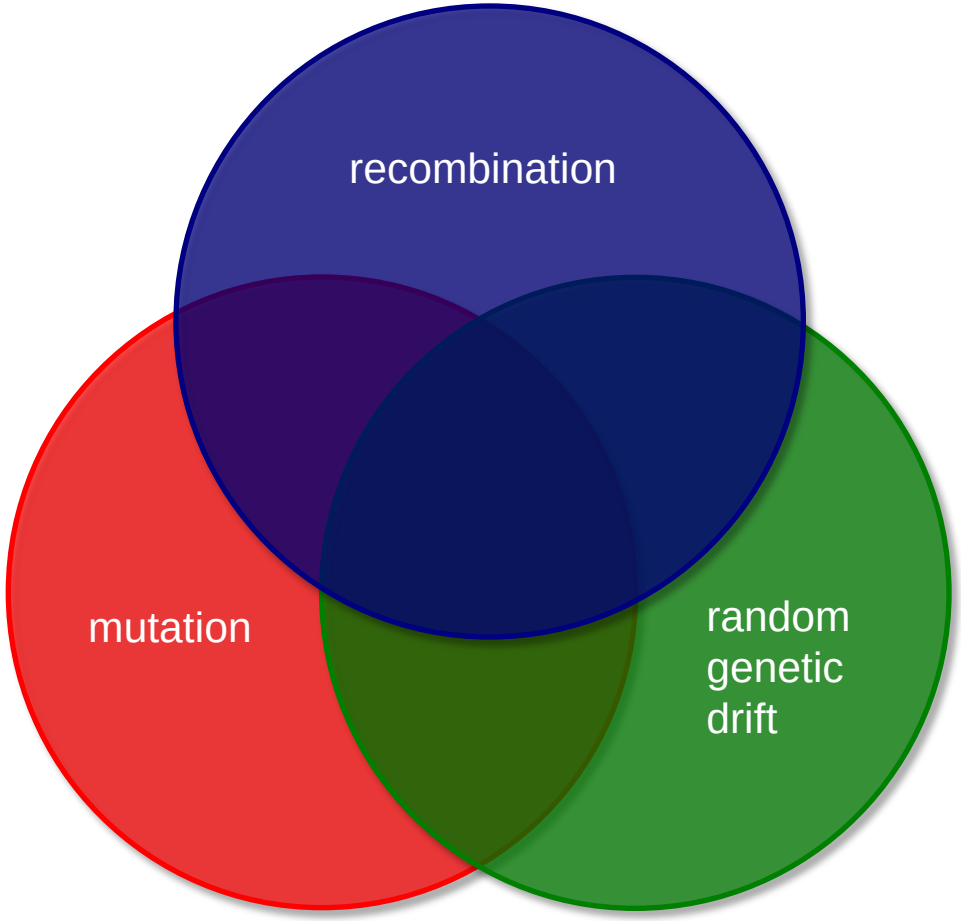


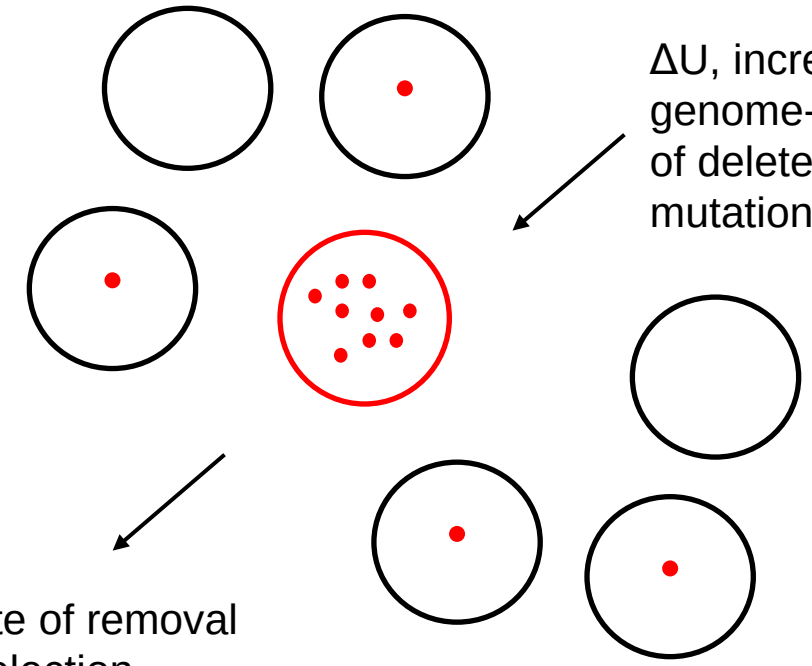
FIG. 1. Average mutation rate μ_{bp} per base pair as a function of genome size G in bp. The logs of the rates for each organism were averaged and all 13 values are included. Phages T2 and T4 were treated as a single organism.

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

The Population-genetic Environment



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

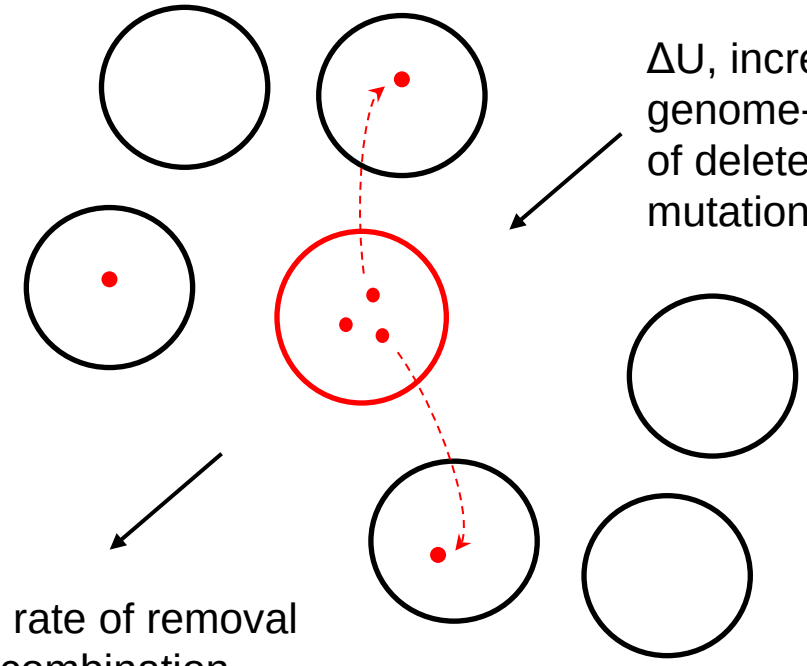
Effect / mutation = s

↓

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

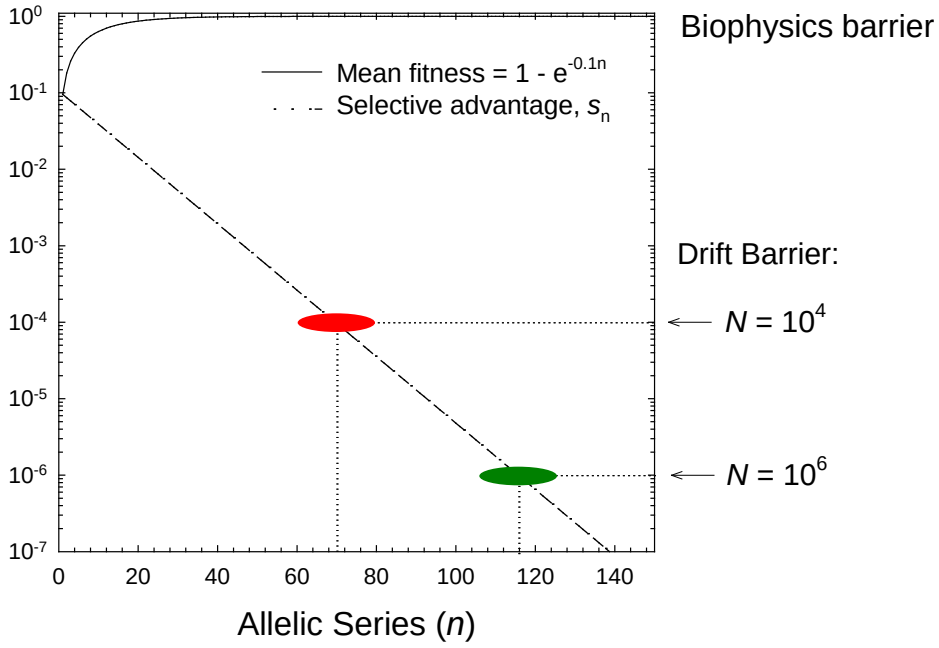
Effect / mutation = s

↓

Total effect on fitness = 2 s ΔU

The Drift-barrier Hypothesis for a Single Trait

Asymptotically Increasing Perfection in an Allelic Series



Biophysics barrier

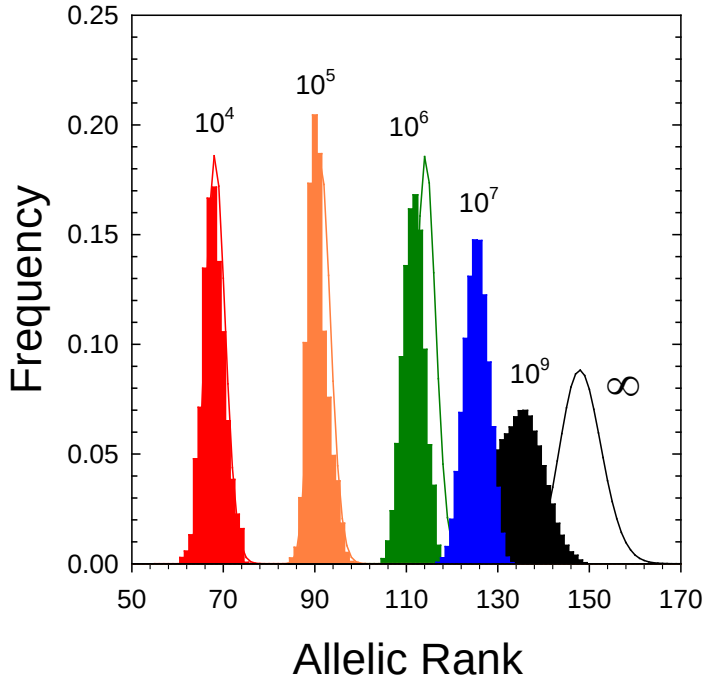
Drift Barrier:

← $N = 10^4$

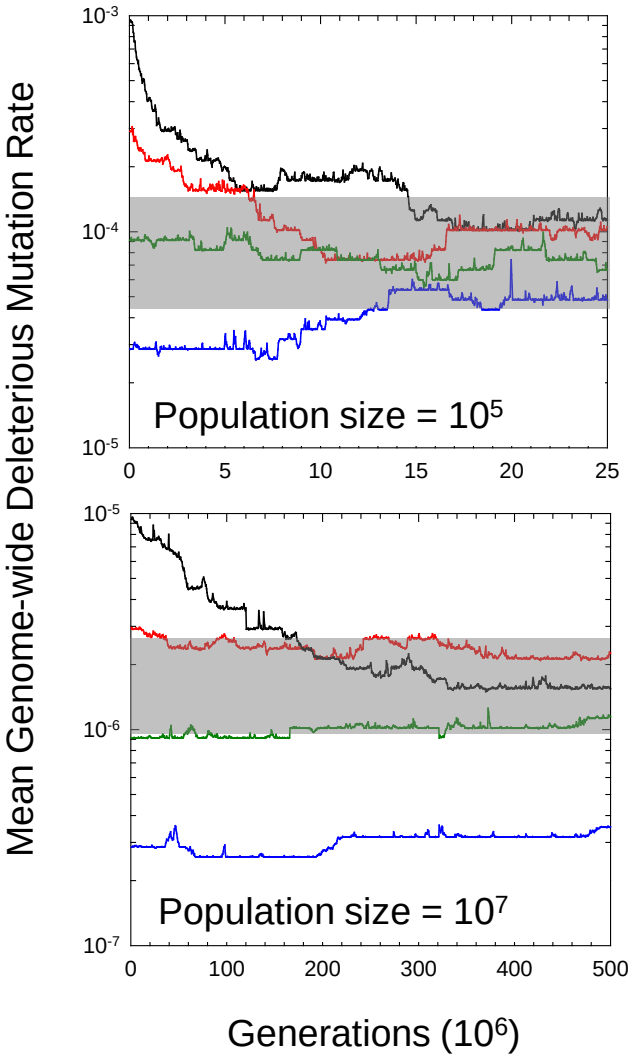
← $N = 10^6$

downward mutation bias

Equilibrium Allele-frequency Distributions with Increasing Population Sizes



Quasi-equilibrium Mutation Rates Resulting From Deleterious-mutation Load

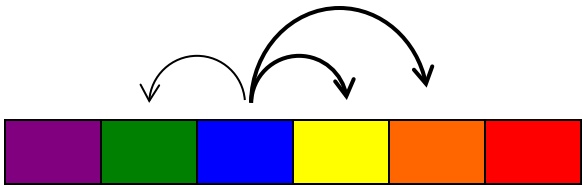


Effective selection for antimutators

DRIFT BARRIER

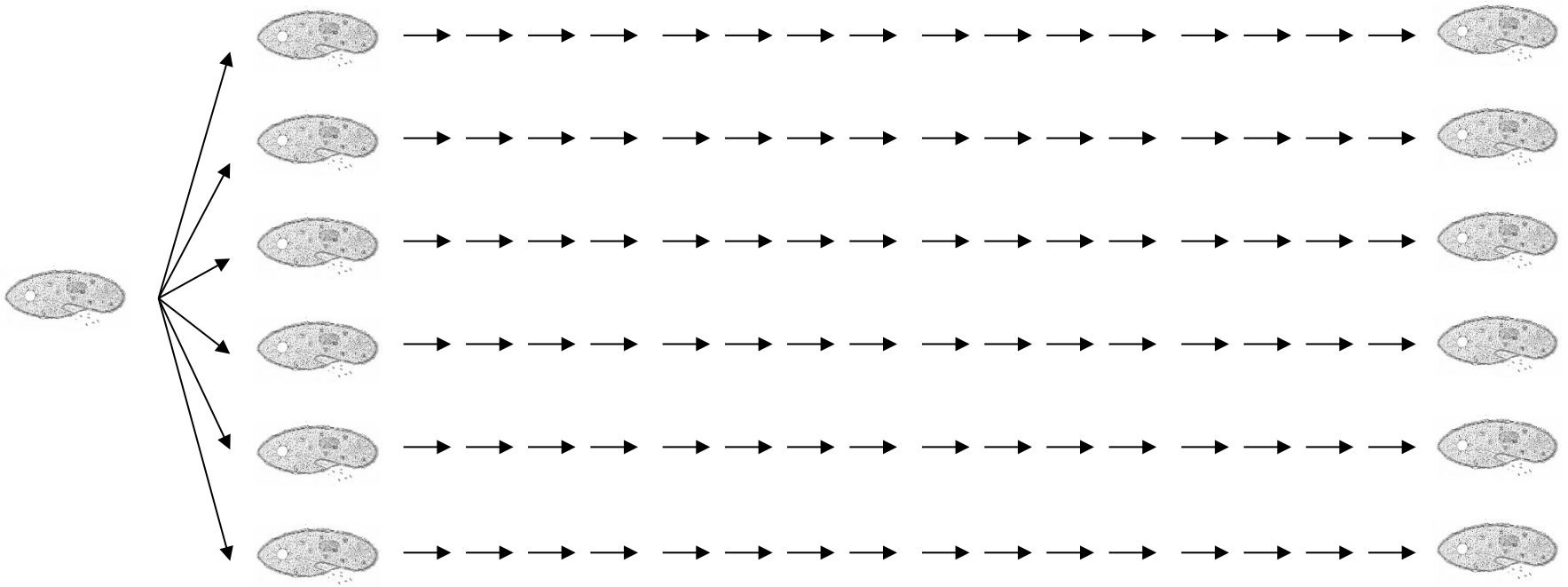
Biased production of mutators

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

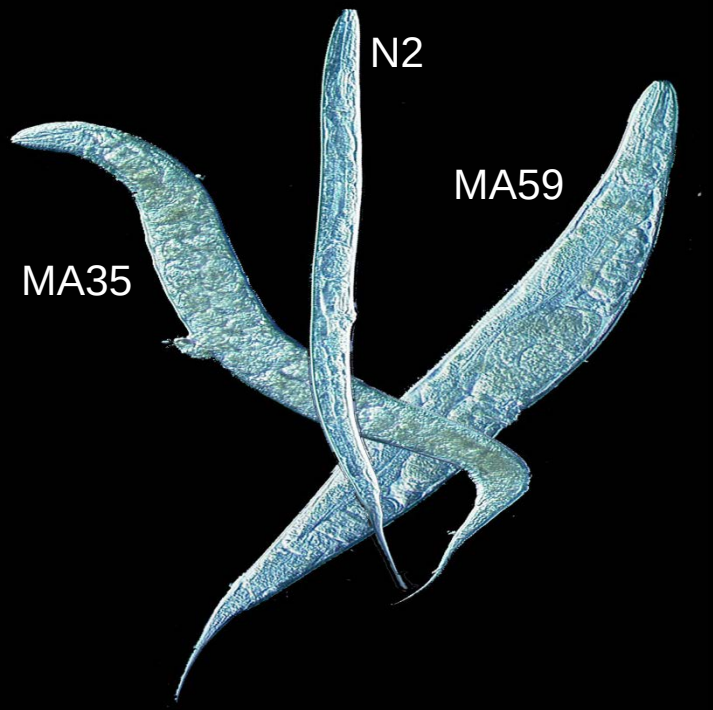


Analysis of Genome Stability with a Mutation-accumulation Experiment:

- 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.

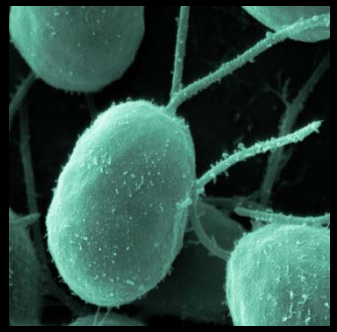


Extreme Morphological Divergence in MA lines of *C. elegans*



Recent and Current Eukaryotic Targets of Study

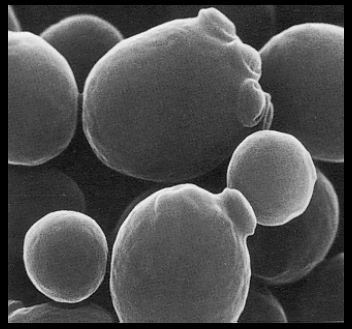
Arabidopsis



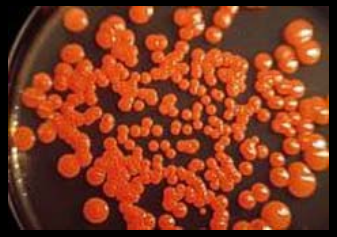
Chlamydomonas

Phaeodactylum

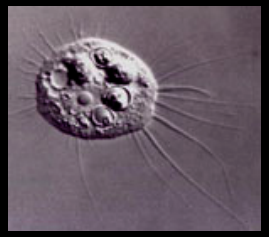
Dictyostelium



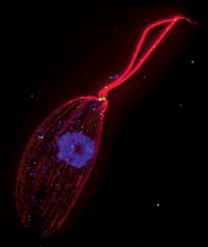
Saccharomyces



Rhodotorula

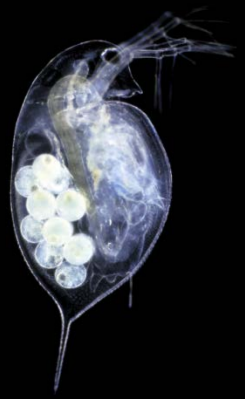


Ichthyosporean



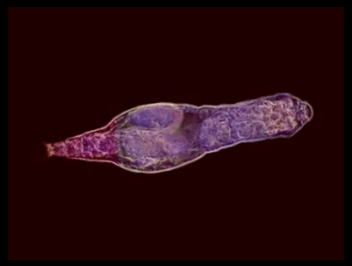
Naegleria

Paramecium

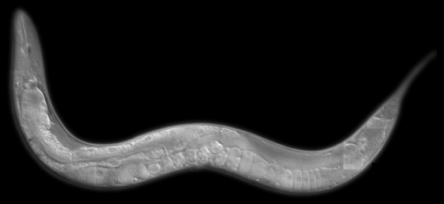


Daphnia

Drosophila



Adineta



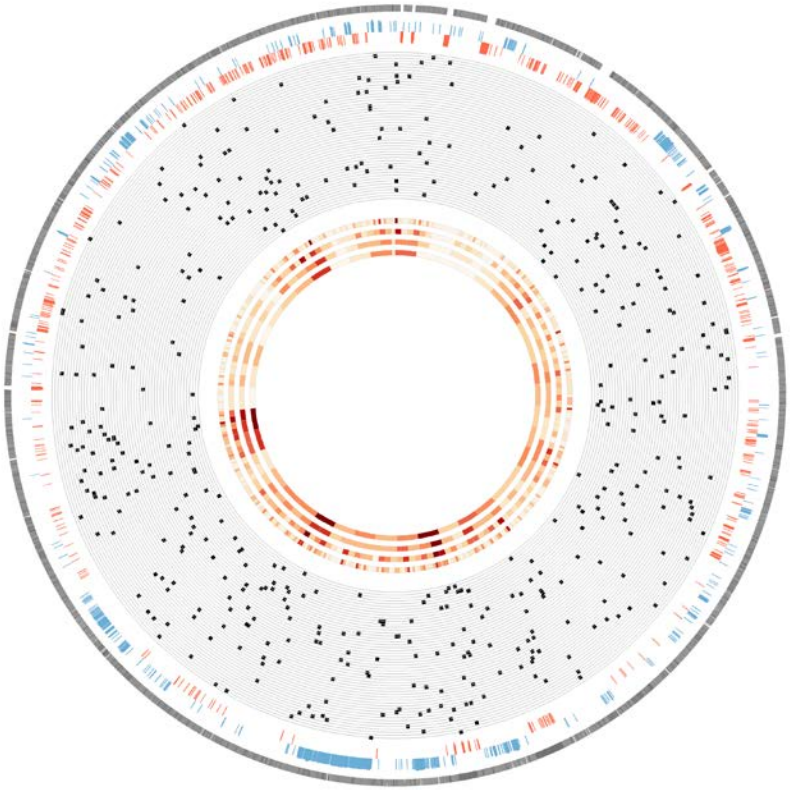
Caenorhabditis

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

Mutation in Small vs. Large Genomes

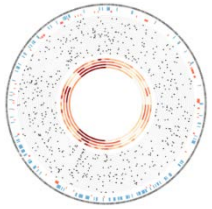


Index:

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.

Scaling of the Mutation Rate per Nucleotide Site With Genome Size

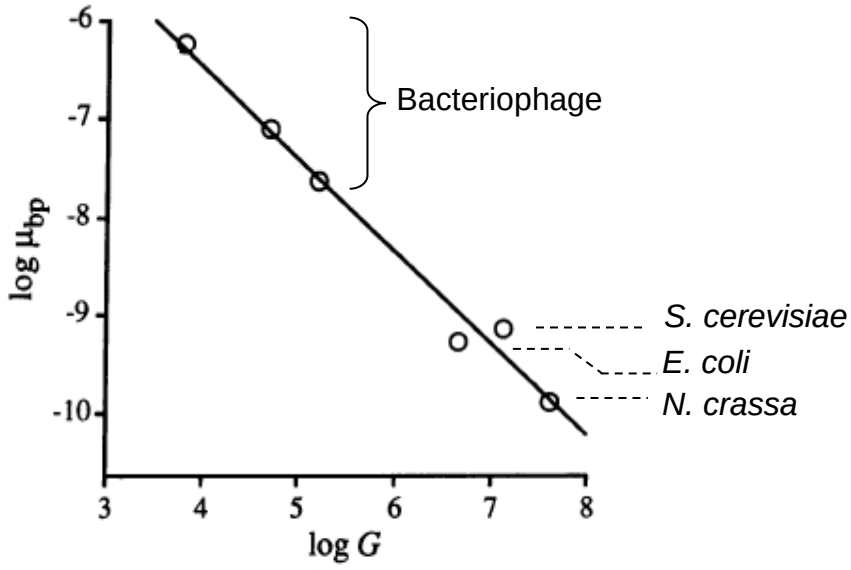
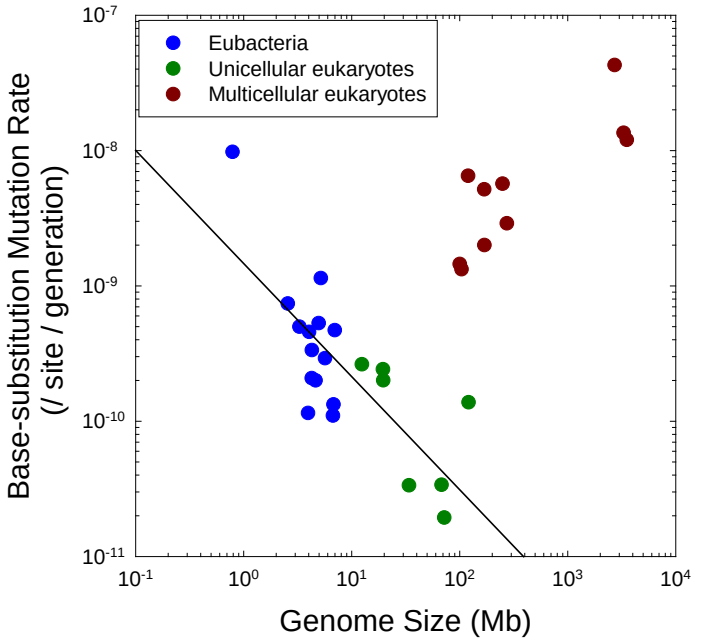
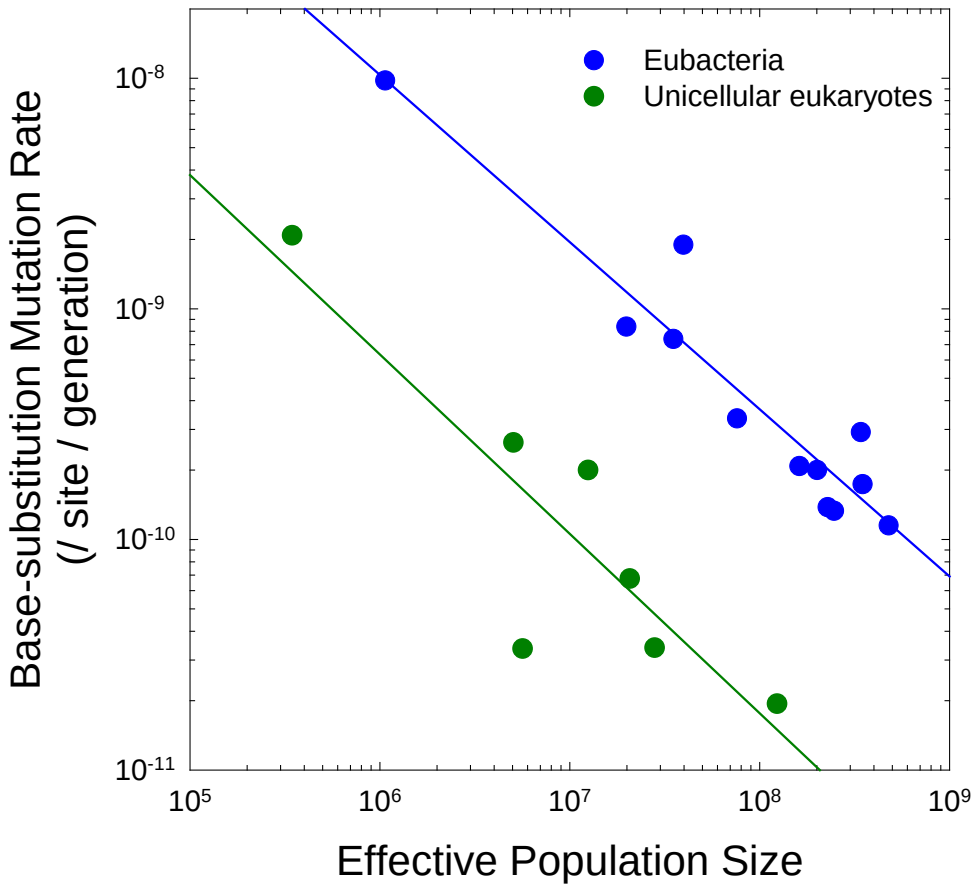


FIG. 1. Average mutation rate μ_{bp} per base pair as a function of genome size G in bp. The logs of the rates for each organism were averaged and all 13 values are included. Phages T2 and T4 were treated as a single organism.

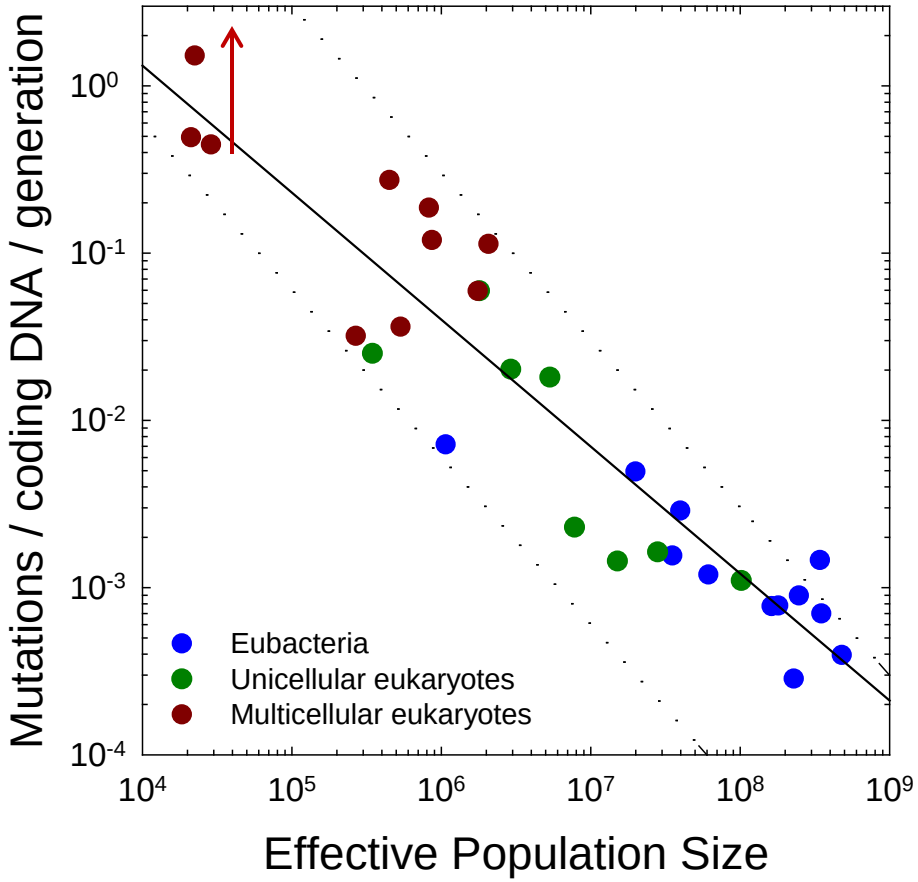


The Mutation Rate / Nucleotide Site Is Inversely Proportional to the Average Effective Population Size of a Species

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).

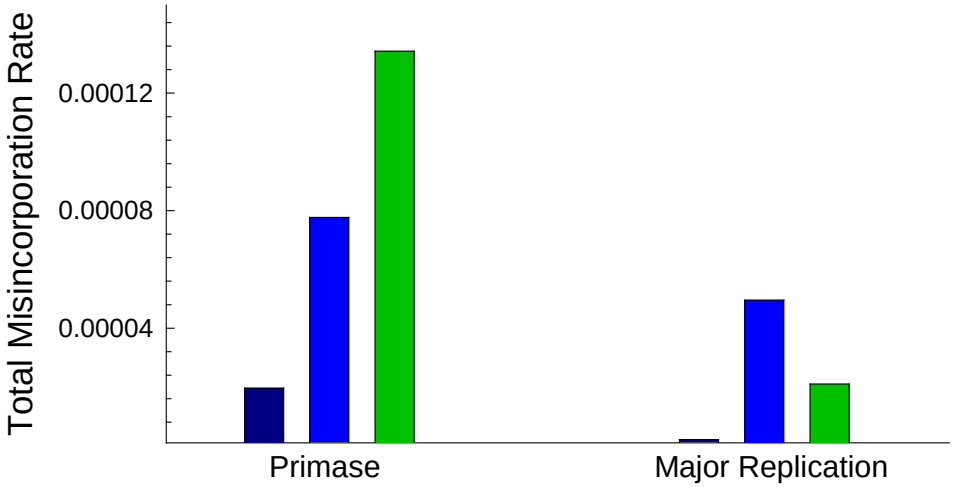
$uG_e \sim 1 / N_e \quad \rightarrow \quad u \sim 1 / (G_e \cdot N_e)$

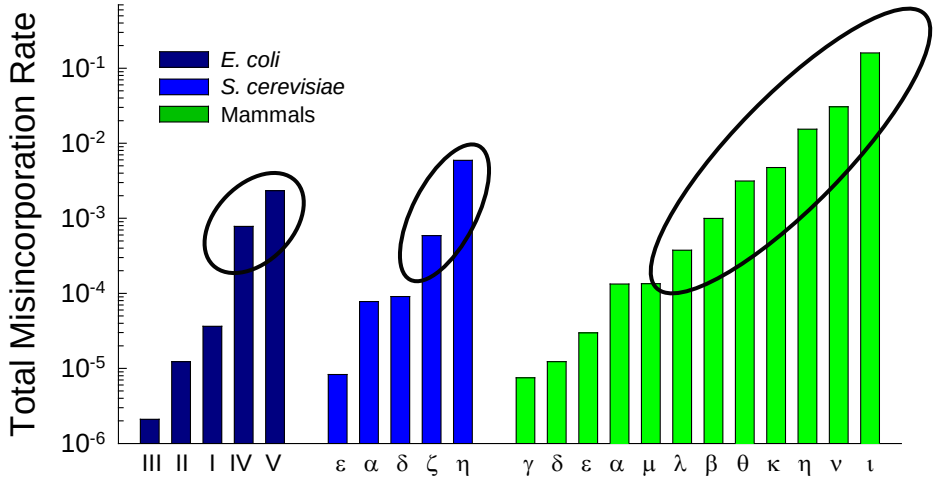
u = mutation rate / site / generation

G_e = amount of functional DNA (sites)

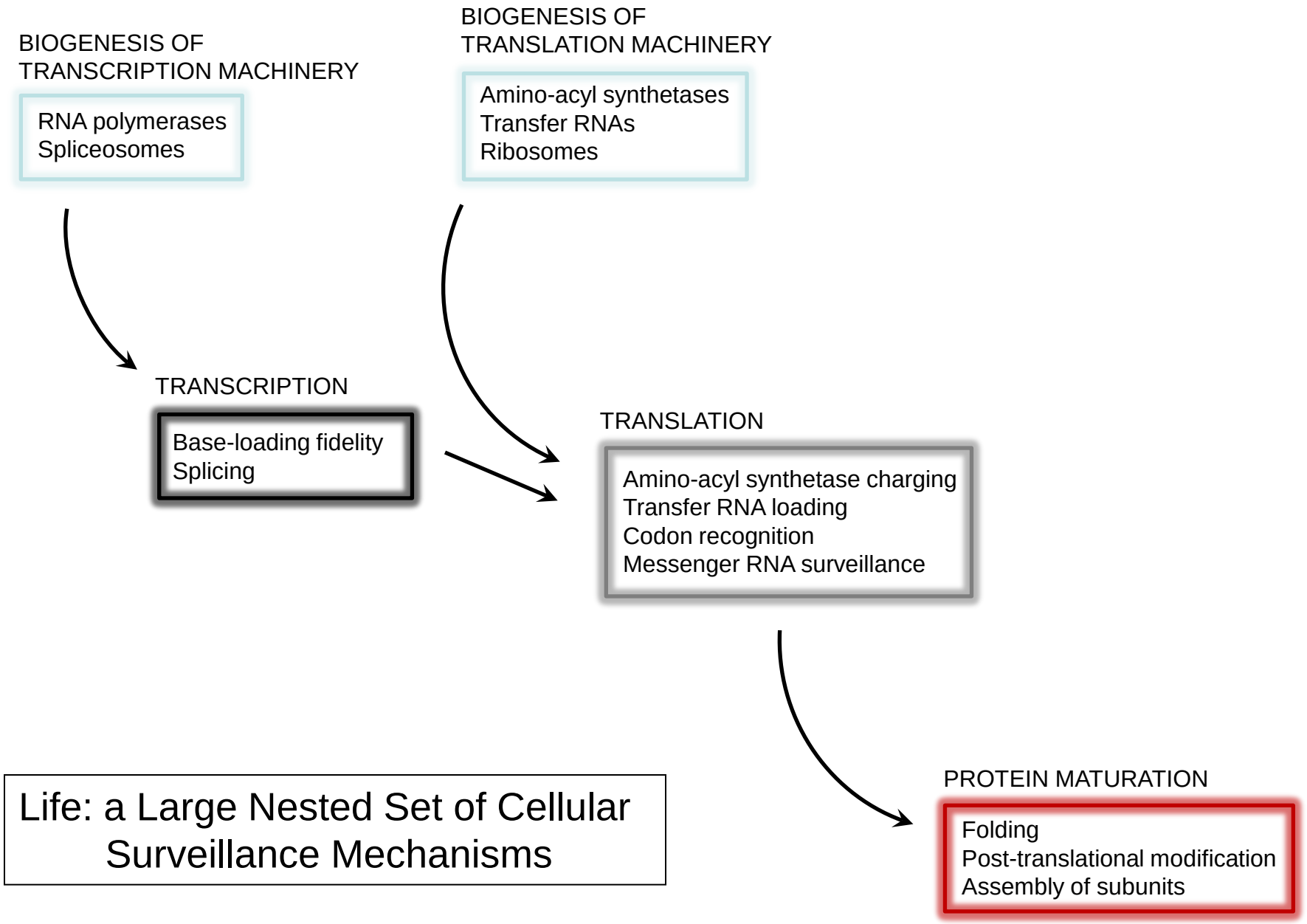
N_e = effective population size

Polymerase Error Rates Are Magnified in Eukaryotes and in Enzymes Involved in Fewer Nucleotide Transactions





Polymerases used in DNA repair are highly error prone, consistent with the drift hypothesis: enzymes involved in fewer nucleotide transactions experience less selection for fidelity.


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 $\Delta u \cdot 1 \cdot T_e \cdot s \cdot d$

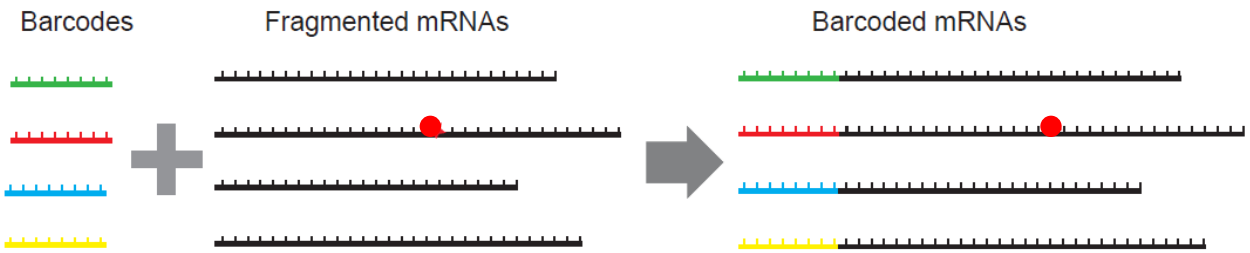
The pool of errors remains associated with the mutator for just one generation

Number of nucleotides in the transcriptome subject to selection

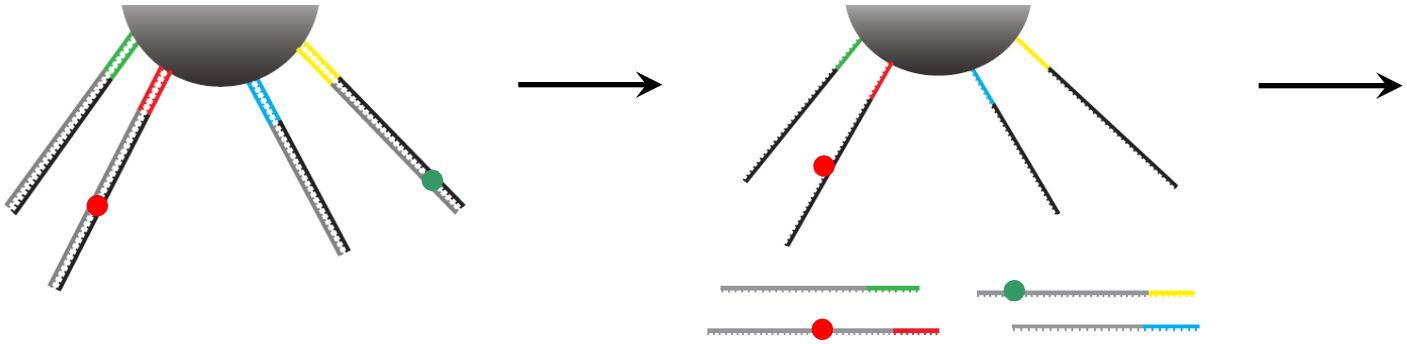
Heterozygous effect of a deleterious mutation

Dilution effect (<< 1.0)

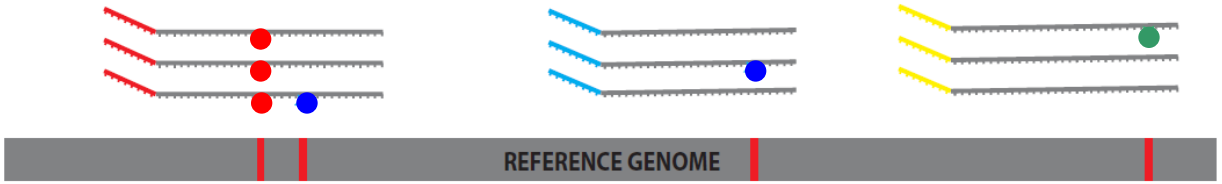
Estimation of the *in vivo* Transcription-error Rate From an RNA Library (Gout et al., PNAS, 2013)



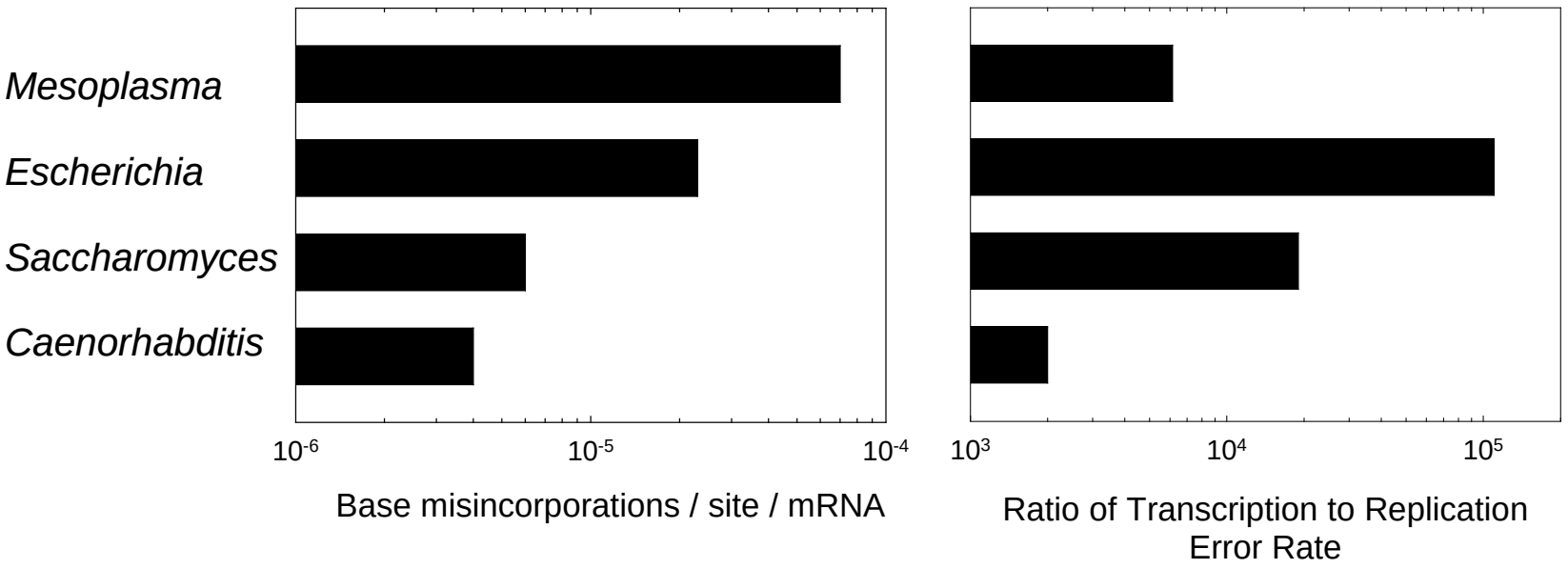
Capture fragments on beads; reverse transcribe; isolate cDNAs; repeat to obtain replicates:



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

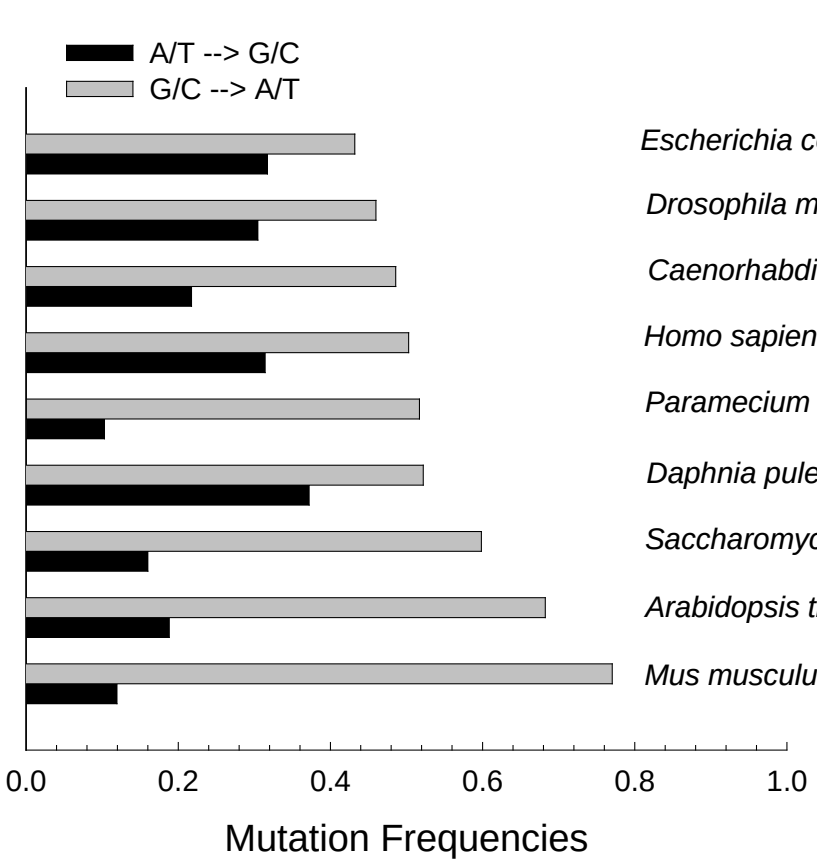


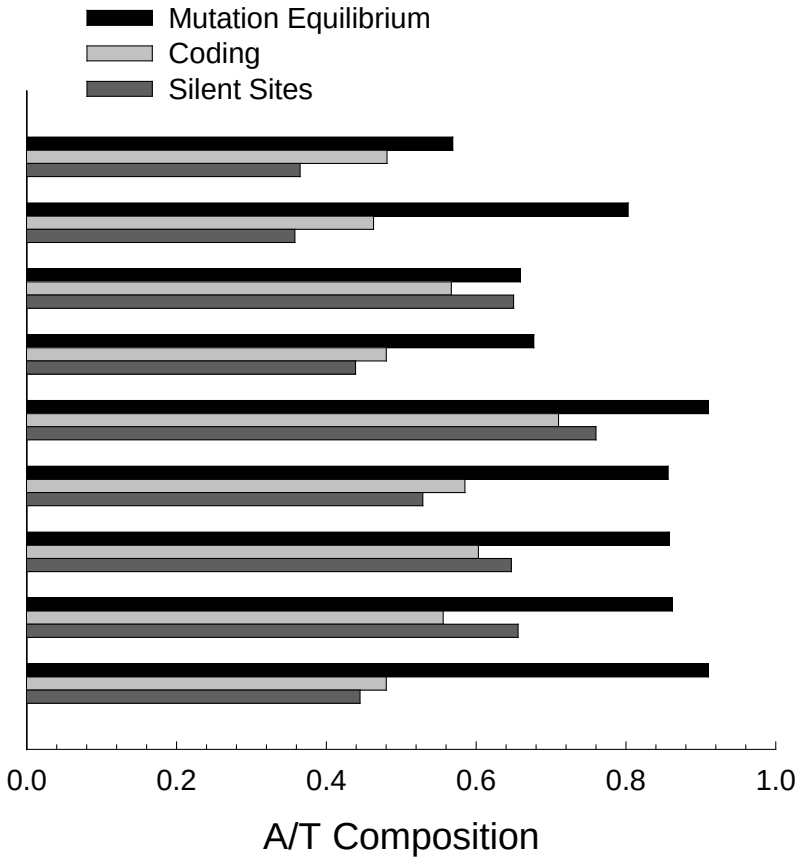
Transcript Error Rates Are Orders of Magnitude Higher Than Replication-error Rates



- ~1 to 5% of transcripts contain errors

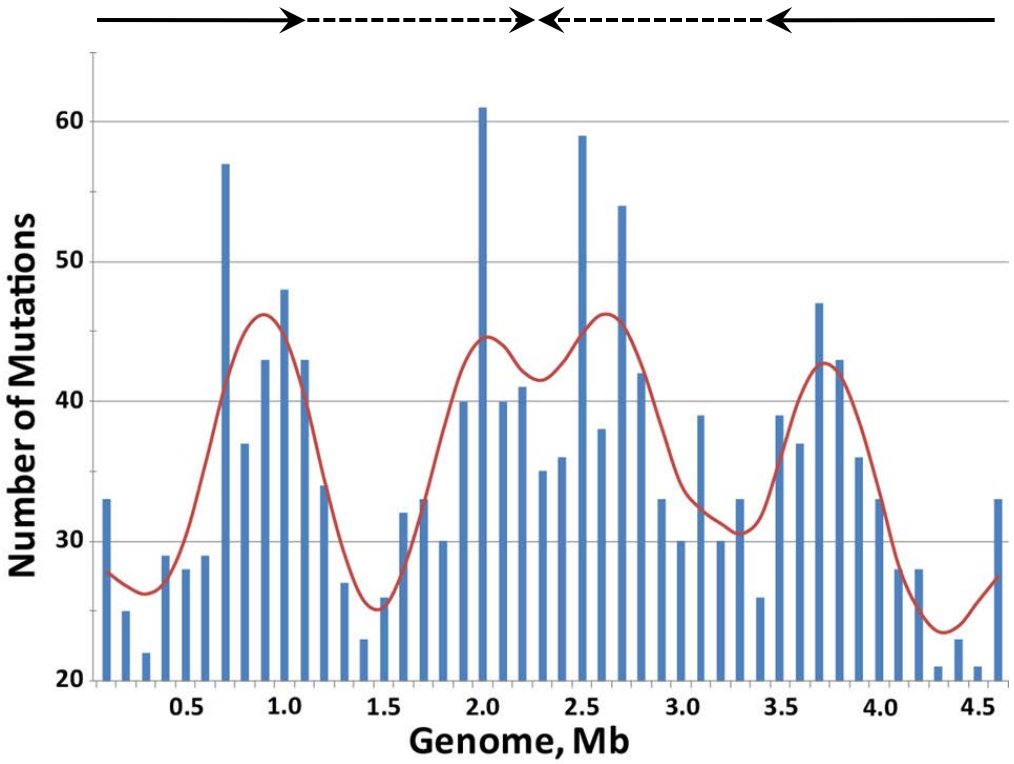
- Nearly all genomes have substantial mutation bias towards A/T production.
- Genome-wide nucleotide compositions are not in mutation equilibrium.
- The universal genomic deficit of A/T must be a result of selection and/or biased gene conversion.



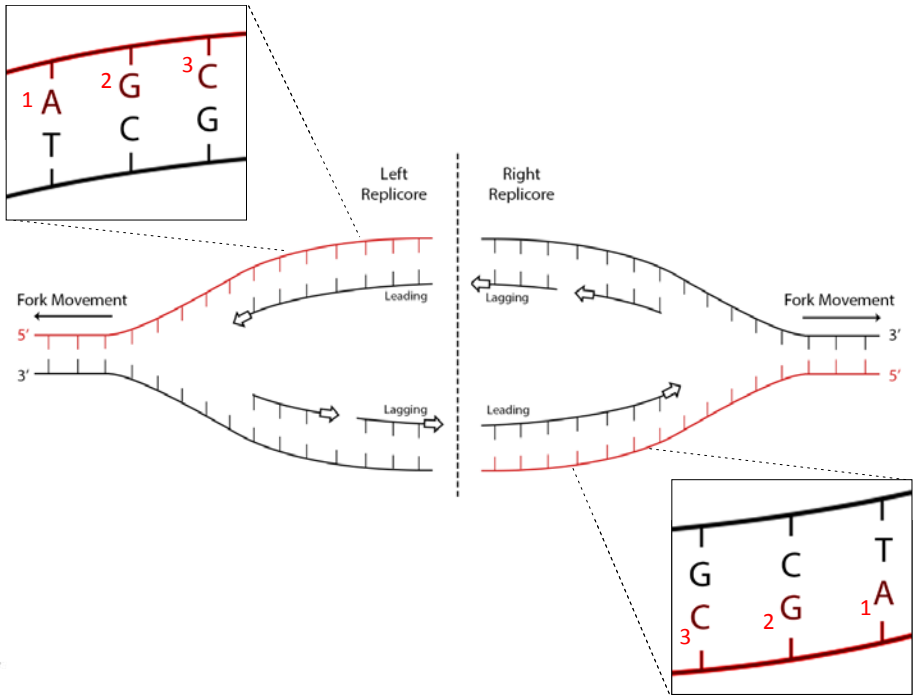


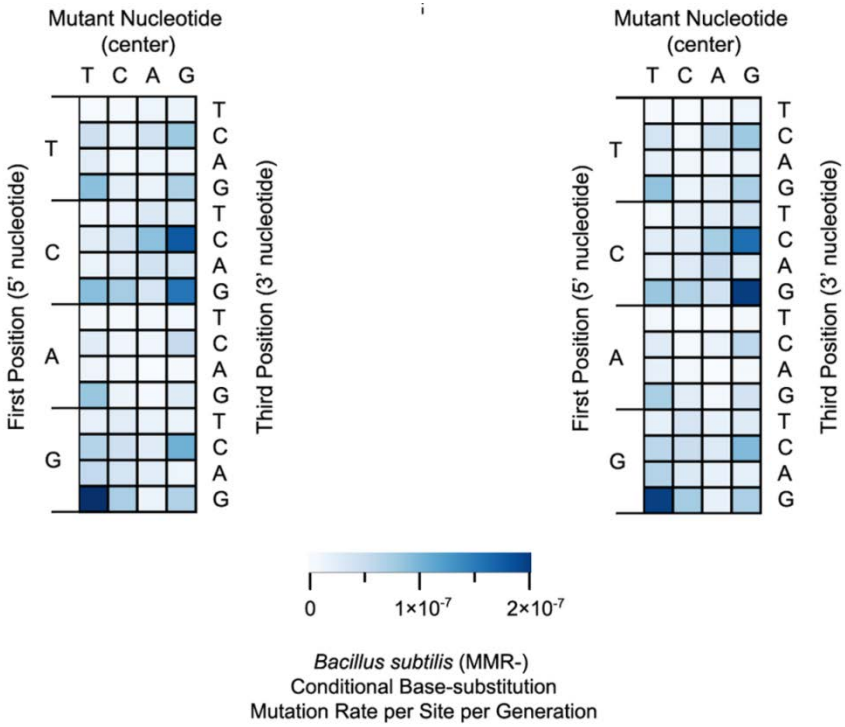
The Mutability of a Gene Depends on its Chromosomal Location (Lee et al., PNAS, 2012)

- Mutations are distributed across bacterial genomes in a large-scale, periodic pattern, repeated in mirror-image in each half of the genome.



Mutations are Context Dependent – depend on the nature of the nearest neighbors (Sung et al., in prep.)





Collaborators:

Indiana University:
Matthew Ackerman, Tom Doak, Pat Foster, Jean-Francois Gout, Matthew Hahn, Nate Keith, Iain Konigsberg, Jay Lennon, Weiyi Li, Hongan Long, Jake McKinlay, Sam Miller, Ron Pearson, Dan Schrider, Joe Shaw, Way Sung, Emily Williams, Sen Xu

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

Hacettepe University:
Sibel Kucukyildirim

Universidade Federal do Rio de Janeiro:
Marcus Senra, Carlos Suarez

University of Chicago:
Allan Drummond

University of Houston:
Becky Zufall

Arizona State University:
Reed Cartwright, David Winter

Desks, Salary:

NATIONAL INSTITUTES OF HEALTH

DEPARTMENT OF DEFENSE
UNITED STATES OF AMERICA

NSF