Lecture 13. Mutation

Mutation, together with selection, is one of only two factors that are absolutely necessary for Darwinian evolution.

What is mutation, mechanistically? - A set of processes that generate mutations, changes of genotypes. There are three such processes:

1) DNA replication, due to errors in it

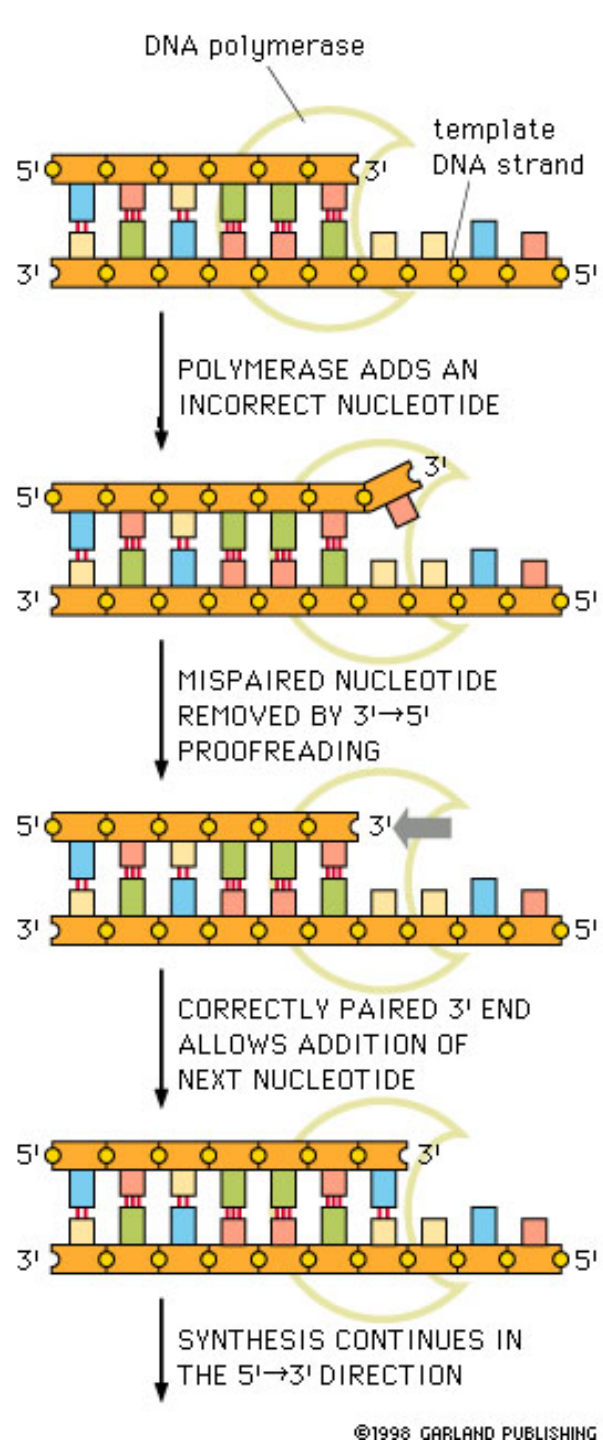
2) DNA repair, due to errors in it

3) Cell division, due to errors in it

DNA replication struggles to be as precise as physically feasible.

DNA repair deals with damages. The difference between a mutation and a damage must be clearly understood - damages violate integrity of DNA, mutations change its sequence. A damage needs to be repaired, and the repair is performed correctly, no mutation appears.

Errors can occur in the course of both mitosis and meiosis, leading to large-scale mutations.

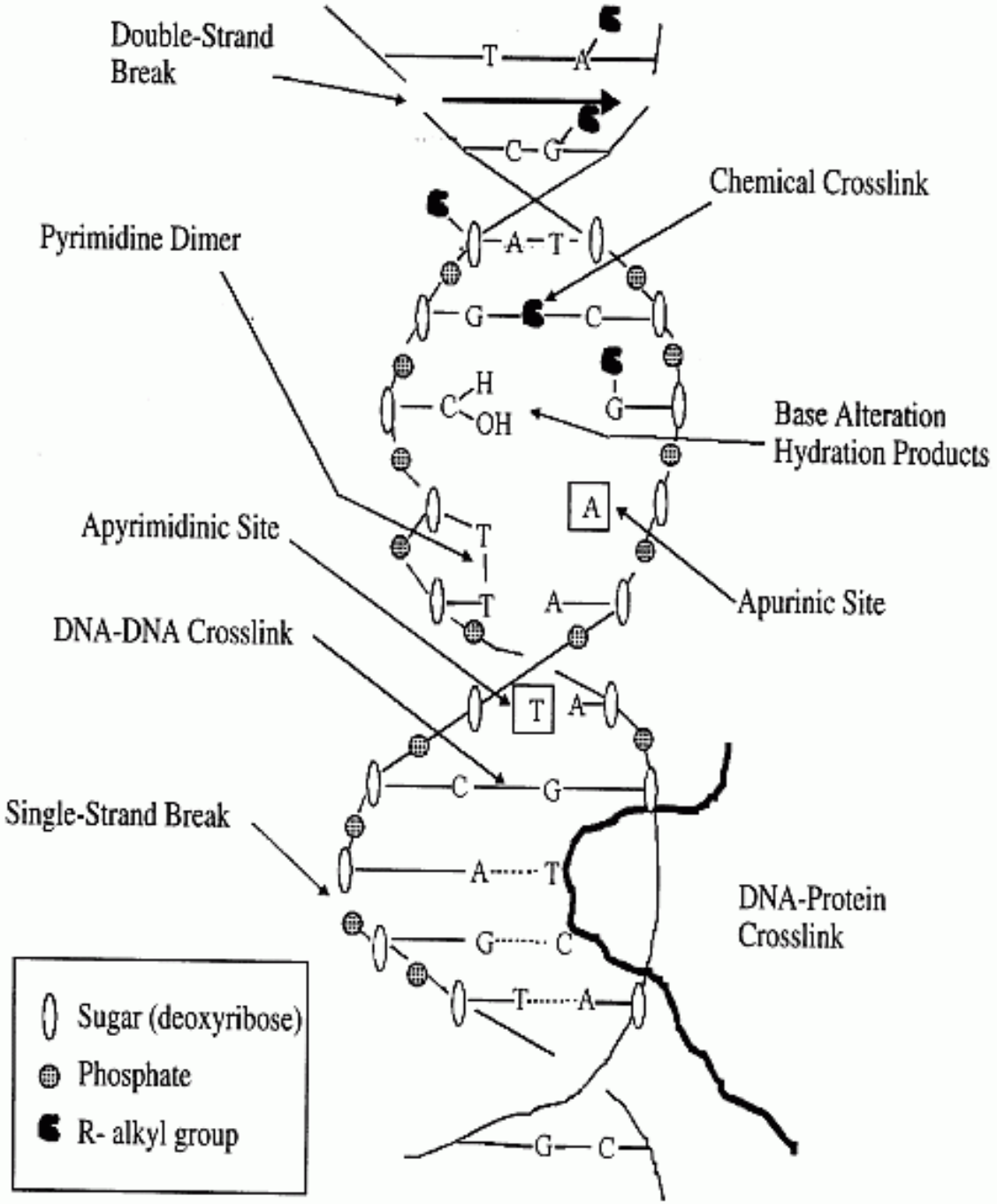


Struggle for fidelity of DNA replication.

Most of DNA polymerases also possess 3'->5' "proof-reading" exonuclease activity. Initially, an incorrect nucleotides is attached to the growing strand with probaility $10^{-4\text{-}5}$. However, an incorrectly attached nucleotide is almost always removed in the course of proof-reading, and this does not happen only with probability $10^{-4\text{-}5}$. As the result, the per nucleotide mutation rate is only ~10^{-9}.

Because no process can be perfectly selective, the proof-reading exonuclease also removes a fraction (up to ~50%) of correctly attached nucleotide, so that fidelity is involved with a cost.

RNA polymerases do not have a proof-reading activities, leading to very high mutation rates (~$10^{-4\text{-}5}$) in viruses which use RNA to store their genetic information (at least during a part of their life cycle).



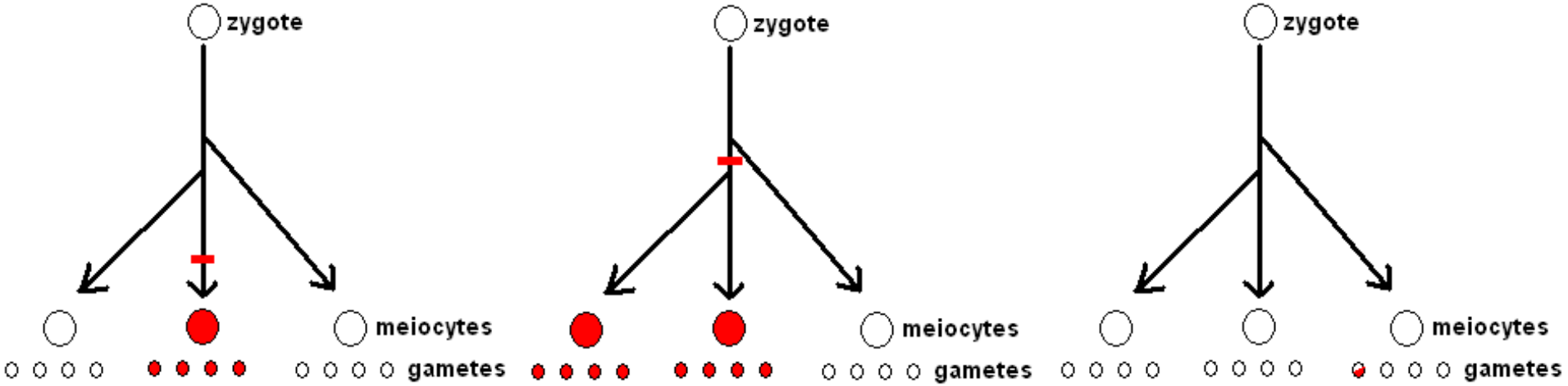
Some common DNA damages.

There are many ways, in which chemical integrity of a DNA molecule can be violated.

In each human cell, every day, many thousands of "spontaneous" DNA damages occur - and all of them must be repaired.

No wonder, that some of these damages are repaired inprecisely, producing mutations.

In multicellular organisms, the timing of a mutation affects the number of mutants.



Germline mutations occuring in a diploid multicellular male.

(left) A mutation that occurred late will be present only in one or a small number of gametes, and will result a single mutant offspring (singleton).

(center) A mutation that occurred earlier will be present in many gametes and will result in several mutant offspring (cluster).

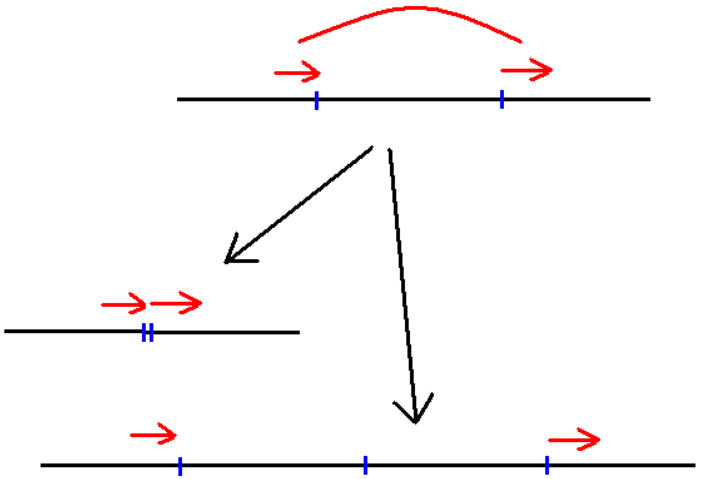
(right) A damage that affected only one DNA strand in the gamete can be transmitted unrepaired to the zygote and lead to a mutation in half of cells in the offspring.

What is a mutation, genetically?

Any change of the genotype could be called a mutation. However, changes caused by reciprocal meiotic recombination, due to independent assortment of non-homologous chromosomes and to crossing-over, are traditionally excluded from mutation and viewed as a separate phenomenon, since they occurs regularly and do not produce really novel genotypes.

reciprocal, meiotic recombination

In contrast, irregular changes which result from non-reciprocal (ectopic) genetic exchanges between different genome segments, as well as gene conversion, are traditionally considered mutations, although their molecular mechanism are closer to that of reciprocal recombination, than to mutation *sensu stricto*.



What is a role of mutation in evolution?

Being the only process that can generate really novel genotypes, mutation is a *sine qua non* of evolution: if mutation were to cease, evolution would eventually stop, after all the existing variation is used up by positive selection.

"The power of selection . . . absolutely depends on the variability of organic beings" (Darwin).

However, the vast majority of selectively non-neutral mutations are unconditionally deleterious, because the space of genotypes contains a huge number of unfit genotypes and only a tiny proportion of fit genotypes. Thus, without being checked by negative selection which preserves *status quo*, mutation rapidly destroys a lineage.

Beneficial mutations, favored by positive selection, provide the necessary raw material for adaptive evolution. However, they are rare, and their availability may constraint Macroevolution.

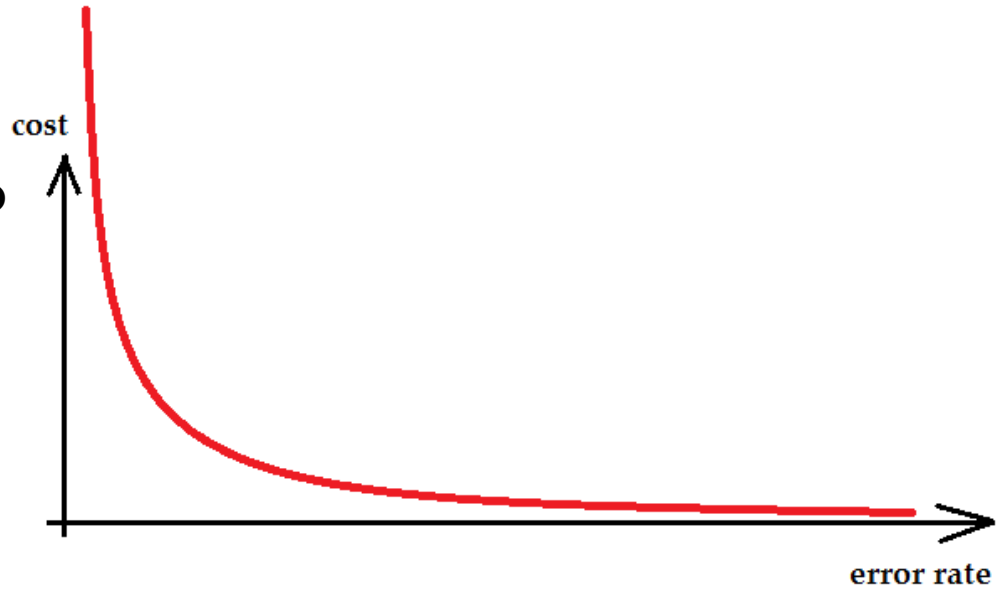
Why does mutation occur?

1. "Everything consisting of parts crumbles ..." (Gautama Buddha, ~500 BCE).

"Mutations are accidents, and accidents will happen" (Alfred Sturtevant, 1938).

2. Alternatively, mutation can be an adaptation, that enables organisms to occasionally produce improve offspring.

To some extent, Buddha and Sturtevant are certainly right: laws of physics do not allow a perfect fidelity of DNA handling. If an organism tries to reduce its mutation rate to zero, the cost, in terms of both time and energy, of DNA handling would approach infinity.



Still, we do not know what would happen if there were no cost of fidelity - mutation may or may not cease.

A mutation is described by a pair of sequence segments S1 and S2, of lengths L1 and L2, such that the ancestral allele carries S1 and the derived allele carries S2 instead. It is convenient to recognize four kinds of mutations (the ancestral allele is presented first):

1) L1 = L2 = 1: a single-nucleotide substitution (SNP), *e. g.*, atCca and atGca.

Substitutions of a purine with a purine and, thus, of a pyrimidine with a pyrimidine, if the opposite DNA strand is considered (AAAGAAA > AAAAAAA; AAACAAA > AAATAAA), are called transitions, and purine > pyrimidine (AAAGAAA > AAATAAA) and pyrimidine > purine (AAATAAA > AAACAAA) substitutions are called transversion. Transitions are usually 2-3 times more common than transversions.

2) L1 > 0, L2 = 0: a deletion (DP), *e. g.*, atGca and atca.

3) L1 = 0, L2 > 0: an insertion (IP), *e. g.*, atca and atACca.

4) All other cases are complex mutations (CPs), *e. g.*, atGca and atACca.

Over 99% of all mutations fit into these four categories, with L1, L2 < 20.

~1% are large-scale deletions, insertions, inversions, etc.

Very occasionally, really complex mutations, referred to as closely spaced multiple mutations (CSMMs) occur:

A

CCAC~~GCTG~~CTGGACGCCATG~~TCCCGGGAGGCTGGGG~~CGCGCGGGGACCTTCCTGGGGGAGG
GGA T CT G T

B

tccagcttattttacagttttttttgttgttgtttggctgggtttttttgttttttttt-
gg

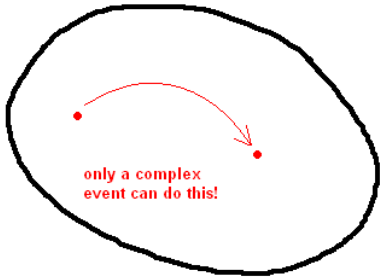
tgttgttgttgttgtattgttttttgttttttt~~t~~gacagagtcactgtcgcctaggctgg
g g g g gt g

C

ATCCTTGTTATTGGAGGAGGAGCAACAGGAAGTGGCTGTGCGCTAGATGCTGTCACCAGAGgtaagtc
A A A T

Three extreme CSMMs. Barred sequences denote deleted nucleotides whereas nucleotides substitutions are indicated below the wild-type sequence. Exonic sequence is denoted by upper case letters, whereas intronic sequence is shown in lower case. The dash in the sequence of mutation B with a ‘‘g’’ below is indicative of the insertion of a single guanine in the mutant allele.

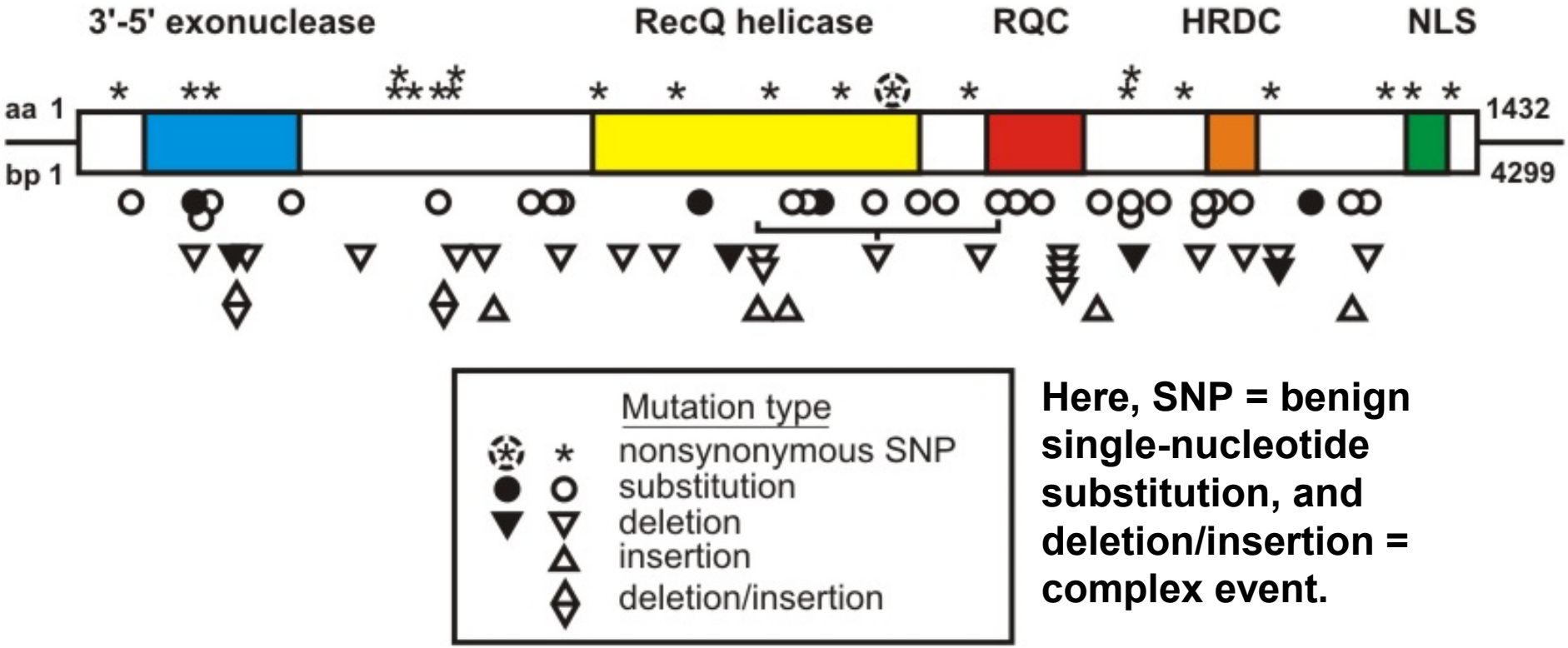
Still, really complex mutations are too rare to matter much, which constrains the course of evolution.



A lot of data on mutation come from patients suffering from Mendelian diseases.

Mutation Number	Mutation	Consequence	Exon	Domain
062	c.95A>G	p.K32R	3	
063	c.107G>A	p.R36Q	3	
064	c.123delA	p.E41fsX47	3	
001	c.171C>G	p.Y57X	3	
065	c.340G>A	p.V114I	3	
002	c.356_366del11	p.S118fsX125	5	exonuclease
051	c.356-2A>C	exon 5 skip	5	exonuclease
003	c.375A>T	p.K125N	5	exonuclease
004	c.403A>G	p.K135E	5	exonuclease
066	c.406G>A	p.A136T	5	exonuclease
052	c.474delT	p.F158fsX161	5	exonuclease
005	c.487_489delGATinsC	p.T162fsX166	5	exonuclease
006	c.502_503delAA	p.K167fsX177	5	exonuclease
007	c.655-1G>A r.655_724del70	p.Y218fsX227	7	exonuclease
008	c.867_874delAGAAAATC	p.I288fsX301	9	
067	c.970A>G	p.T324A	9	
068	c.986A>G	p.Q329R	9	
069	c.1027G>A	p.E343K	9	
009	c.1105C>T	p.R369X	9	
010	c.1123_1124delGAinsC	p.F374fsX378	9	
070	c.1147G>T	p.L383F	9	
071	c.1161G>A	p.M387I	9	
011	c.1165delA	p.E388fsX392	9	
012	c.1250_1253delTTGC	p.D416fsX436	9	
013	c.1278_1279insATCT	p.S246fsX430	10	
014	c.1389T>A	p.Y463X	11	
015	c.1462G>T	p.E488X	12	

Example: summary of mutations that cause Werner syndrome



Here, SNP = benign single-nucleotide substitution, and deletion/insertion = complex event.

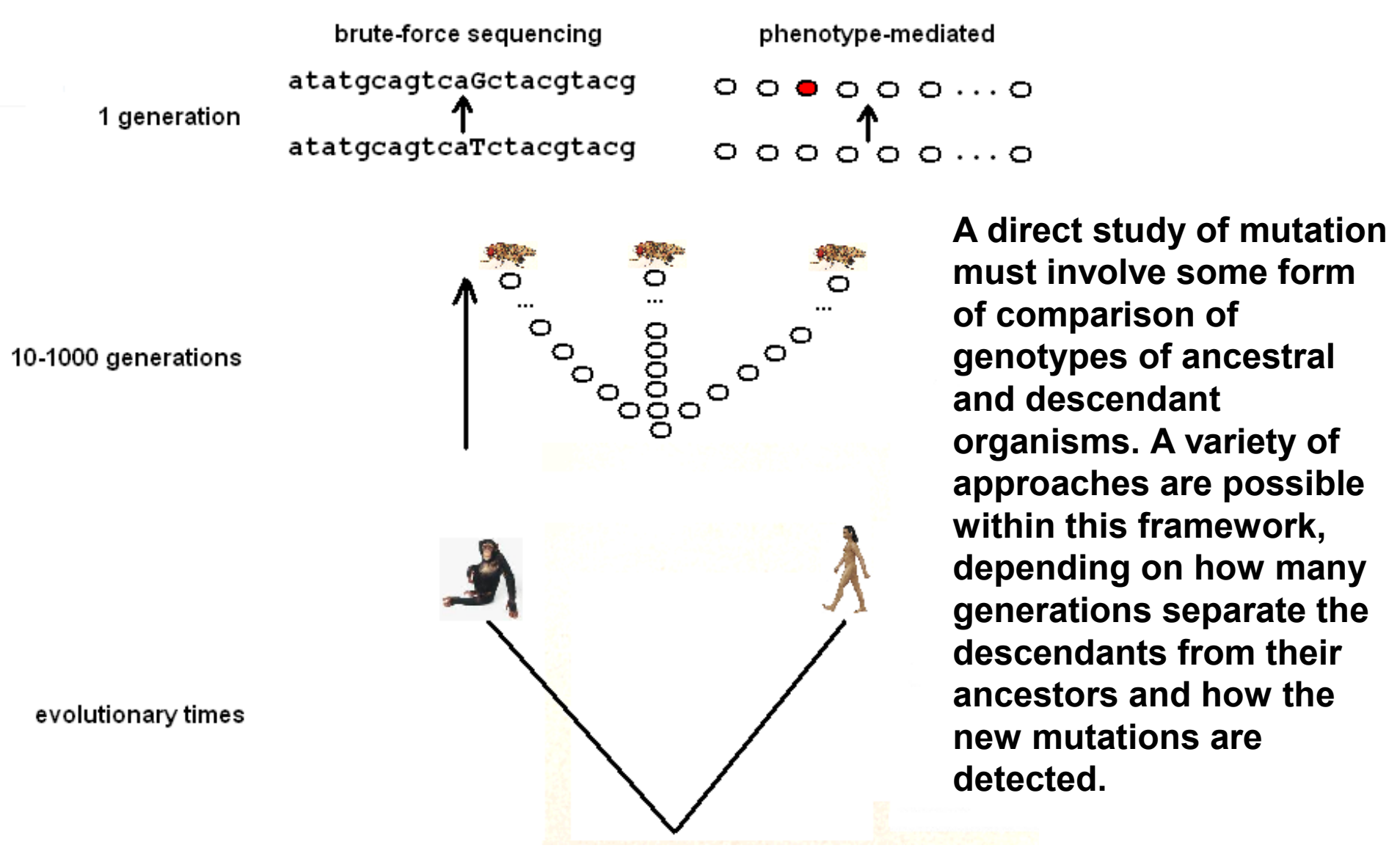
Werner syndrome (OMIM catalog # 277700) is an autosomal recessive human genetic instability syndrome whose phenotype mimics premature aging - patients appear to age rapidly after puberty. Werner syndrome appears in individuals carrying two inactive alleles of the *WRN* protein, which is a DNA helicase.

Mutation rates

Quantitatively, mutation is characterized by mutation rates. The key among them is the overall per nucleotide site per generation mutation rate μ, which can be defined as the total number of mutations of all kinds that occur ever generation in a long sequence of length L, divided by L.

Naturally, m can be subdivided into components coresponding to mutations of different kinds, such as μ_{sub}, μ_{del}, and μ_{ins}.

Multiplication by the haploid genome size G converts a per site mutation rate μ into the corresponding haploid genomic mutation rate. Thus, T, the mutation rate per the genome of an organism is $T = G\mu$ for haploids or $T = 2G\mu$ for diploids. T can be partitioned into components which correspond to molecular events of different kinds, in the same way as μ.



A direct study of mutation must involve some form of comparison of genotypes of ancestral and descendant organisms. A variety of approaches are possible within this framework, depending on how many generations separate the descendants from their ancestors and how the new mutations are detected.

Short-, millde-, and long-term direct methods for measuring mutation: parent-offspring comparison (top), comparison of an ancestral genotype(s) with those of its descendants after a moderate number of generations (middle), and comparison of genotypes of different species, separated by many generations (bottom).

Some technical details of measuring mutation rates:

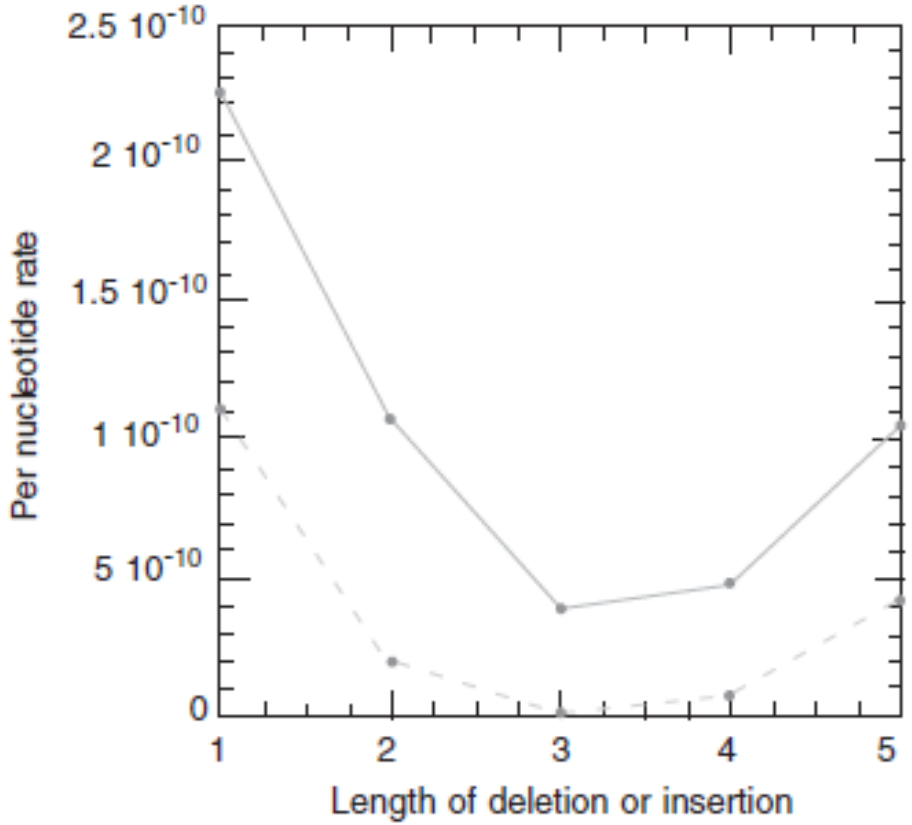
Genotypes of parents and their offspring can be compared by simply comparing their DNA sequences which, of course, requires large-scale, high-precision sequencing, because per nucleotide mutation rates are very low. Alternatively, a phenotypic screening for drastic mutations at a particular locus can be performed first. After this, the per nucleotide site mutation rate can be recovered, as long as the target size for mutations of a particular kind is known for the wild-type sequence. If, for example, at some locus the rate of loss-of-function mutations due to nonsense substitutions was measured to be 10^{-6}, and the number of all possible nucleotide substitutions which would lead to an inframe stop codon is 100, $\mu_{sub} = 3x(10^{-6}/100) = 3x10^{-8}$ (the factor of 3 appears because three different substitutions can occur at a site; we ignored a possibility that different substitutions occur at different rates).

If we consider descendants and ancestors separated by not one but a moderate number of generations (usually, 10-1000), it is crucial that mutations are allowed to accumulate freely in the course of these generations. This can be achieved by maintaining, in the laboratory, a set of MA lines or MCN populations. Obviously, this approach can only be applied to organisms with short generation times.

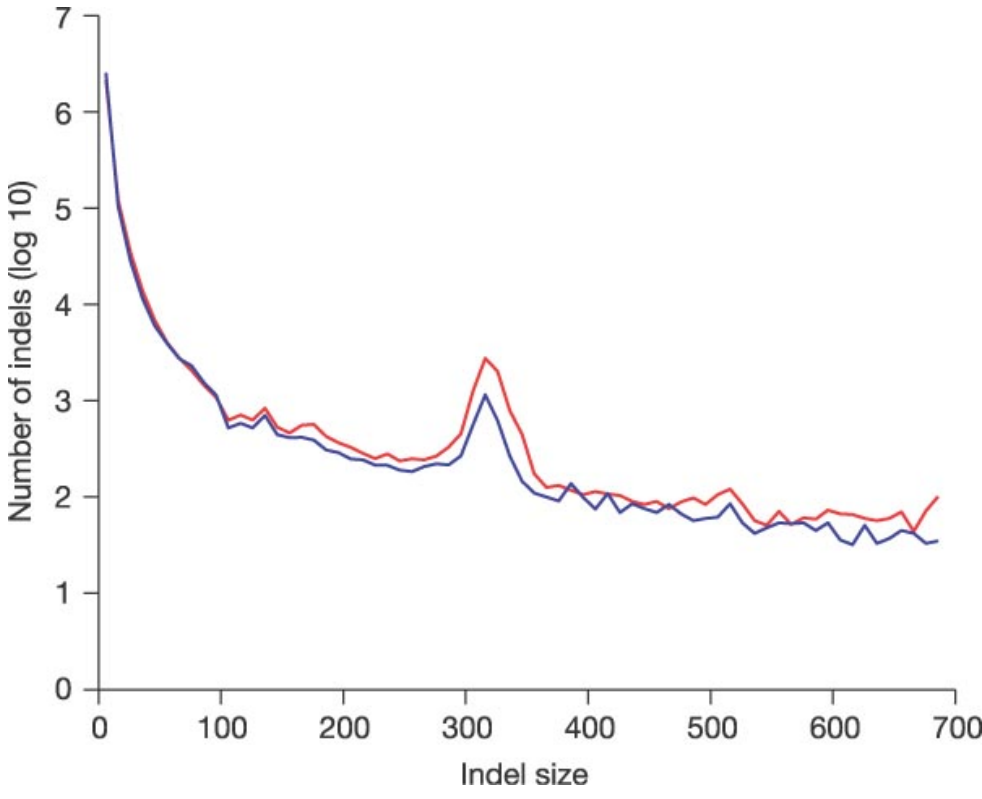
Detection of mutations, through comparison of genomes of different species, requires a precise knowledge of the overall number of generations on the evolutionary trajectory that connects the species, and complete absence of selection within the studied sequence segments. Them, $\mu_{sub} = M/G$, where M is the fraction of mismatches in the alignment of orthologous selectively neutral sequences from similar species and G is the number of generations on the path between them (this formula ignores a small contribution into interspecies divergence from variation in the ancestral population, as well as the possibility of multiple allele replacements per site).

species	rate of mutation per site per generation	comments	references
Homo sapiens	2.5×10^{-8} total (from species divergence) 1.8×10^{-8} total (phenotypic mutations analysis) 3.0×10^{-8} total (direct sequencing of Y chromosome)	indels are rare	Nachman & Crowell (2000), Kondrashov (2003), Xue *et al.* (2009)
Drosophila melanogaster	8.4×10^{-9} total 3.5×10^{-9} to 5.8×10^{-9} point mutations	some variation of rates between different lines; supports higher rate estimates	Haag-Liautard *et al.* (2007), Keightley *et al.* (2009)
Drosophila melanogaster mitochondrion	7.2×10^{-8} total 6.2×10^{-8} point mutations	high rate of G→A mutations on the major strand; higher than estimates from species divergence	Haag-Liautard *et al.* (2008)
Caenorhabditis elegans	2.1×10^{-8} total 9.1×10^{-9} point mutations	insertions very common; higher than previous estimates	Denver *et al.* (2004)
Caenorhabditis elegans mitochondrion	1.6×10^{-7} total 9.7×10^{-8} point mutations	higher than previous estimates	Denver *et al.* (2000)
Saccharomyces cerevisiae	0.33×10^{-9}	similar to previous estimates	Lynch *et al.* (2008)
S. cerevisiae mitochondrion	1.2×10^{-8} point mutations (average of two methods) 7.5×10^{-9} indels (average of two methods)		Lynch *et al.* (2008)

Insertions and deletions represent >10% of small scale mutations, and most of them are short.



Frequencies of deletions (solid line) and insertions (dotted line) of length 1, 2, 3, 4, and ≥ 5 in humans.



Indels that occurred in the course of human-chimpanzee divergence. A spike at ~300 nucleotides is due to insertions of SINE TEs.

A finer point: mutation rate is not uniform alone the genome

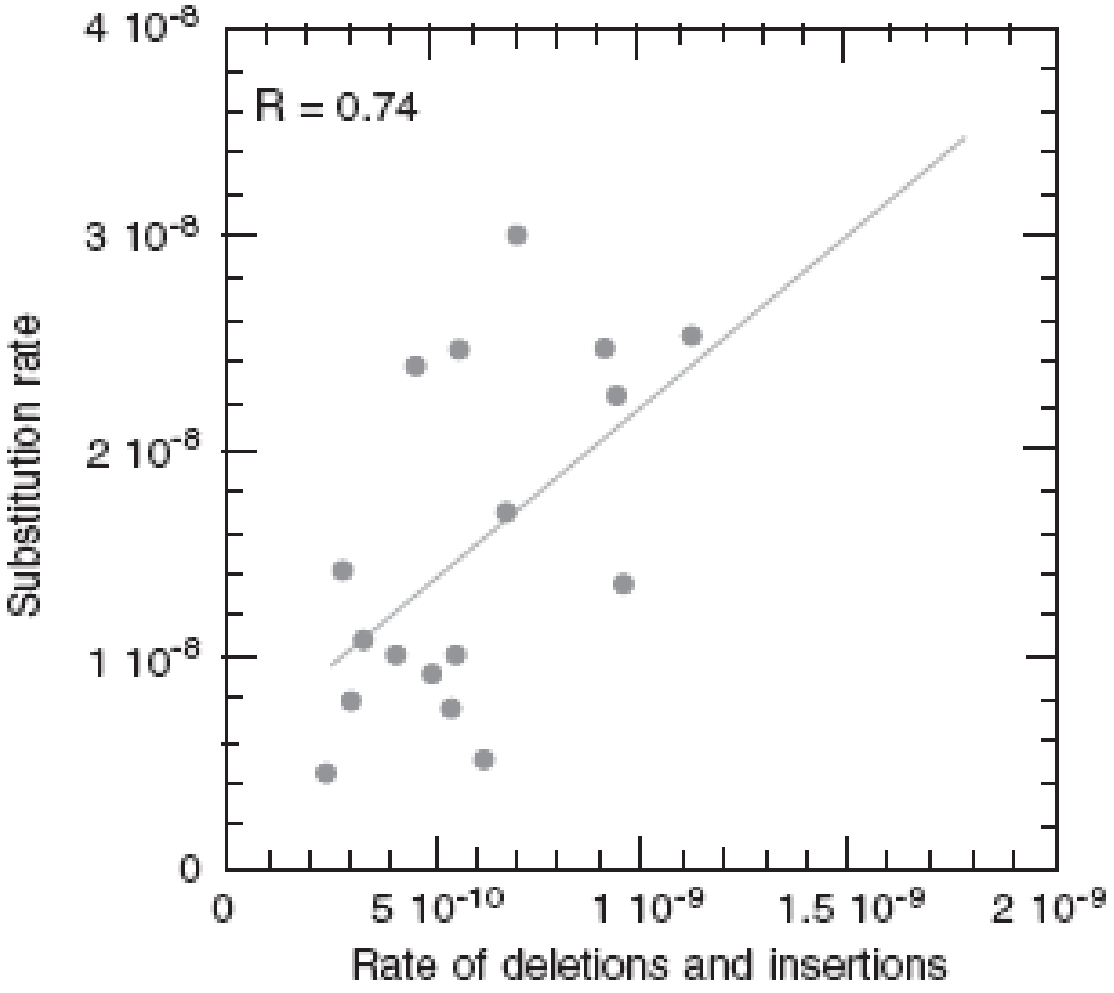
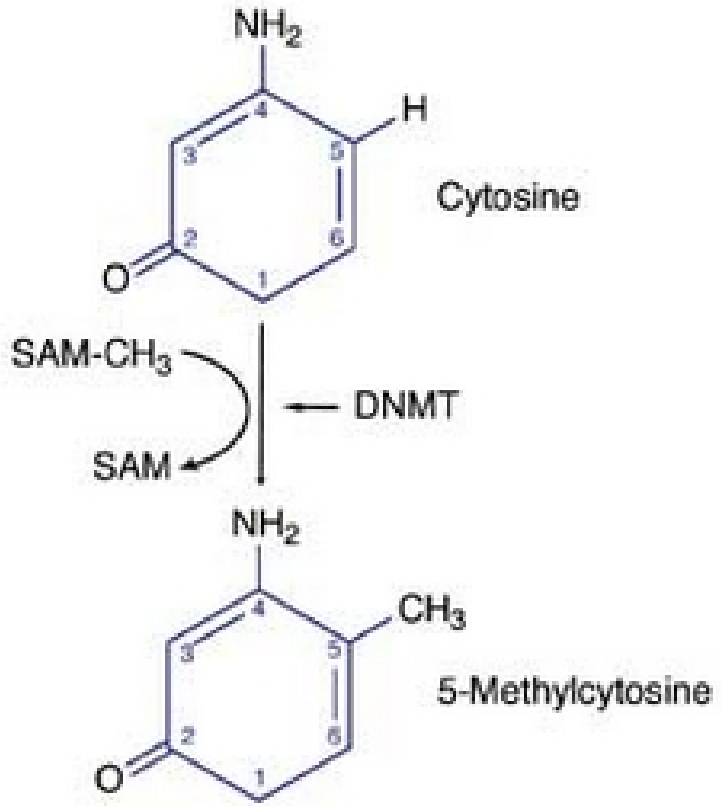


FIGURE 3. The relationship between the rate of deletions and insertions $\mu_{del} + \mu_{ins}$ and the substitution rate μ_{sub}.

Apparently, per nucleotide rates of mutations of different kinds are not uniform across the human genome, and rates of mutaitons of different kinds vary, along the genome, in a correlated way.

A finer point: mutation rate at a site can strongly depend on its context

Hypermutability of 5'CpG3' dinucleotides in mammalian genomes, due to methylation of cytosine residues, within such contexts.



Mechanism of DNA methylation

The methylated cytosine may be converted to thymine by accidental deamination. The cytosine to thymine change can be corrected only by the mismatch repair which is very inefficient.

As a result, C>T transition rate if ~15 times higher for C's that are within CpG contexts, than for C's that are outside CpG contexts.

Thus, mammalian genomes are strongly depleted of CpG dinucleotides - in non-coding DNA, such dinucleotides constitute only ~1% of all dinucleotides, instead of ~6% (1/16) expected.

However, coding exons contain a much higher fraction of CpG dinucleotides. As a result, a large fraction of human pathogenic missense and nonsense mutations (~40%) occur within CpG's, mostly those that encode arginine.

Impacts of mutation on quantitaive traits

Species	Generations	Trait	ΔM ($\times 10^{-3}$)	h_m^2 ($\times 10^{-3}$)	CV_m (%)	Mutation rate	Effect
VSV	80–120	w^b	−3.1 (0.20)	–	0.64	1.6 (0.57)	−0.0022 (0.0008)
		w^b	−2.1 (0.04)	–	0.30	0.88 (0.054)	−0.0024 (0.0001)
		w^b	−2.5 (0.04)	–	0.34	1.2 (0.052)	−0.0022 (0.0001)
ϕ6 virus	~200	w	−0.92	–	0.65	0.035	−0.027
E. coli	~7500	Exponential growth rate	−0.0025	–	0.020	0.00017	−0.012
S. cerevisiae	~600	w	0.005 [−0.022/0.032]	0.48	–	–	–
	~1012	w	−0.0059	1.1	0.087	2.4e–5	−0.13
A. gloriosa	11	Number of flowers	−4.6	0.63	1.2	0.055	−0.085
A. douglasiana	11	Number of flowers	−9.7	3.7	3.2	0.059	−0.16
		Dry weight	−2.4	0.65	0.58	0.090	−0.027
A. thaliana	10	w	−8.9 [−30/−1]	3.1 [−6/10]	3.2 [−5/6.3]	0.050 [0.002/0.4]	−0.23 [−0.9/−0.02
	17	Number of fruits	−2	1.7	5.1	4e–3	−0.54
		Seeds/fruit	−0.1	1.3	0.64	1.5e–4	−0.92
C. elegans	60	Productivity	−0.30 (0.3)	1.2	1.2	0.00065	−0.46
	60	r	−0.35	–	0.42	0.0035 (0.001)	−0.10 (0.01)
	214	Productivity	−2.1	0.97	0.8 (0.3)	0.024 (0.012)	−0.088 (0.032)
		r	−1.5	1.3	3.3 (0.8)	0.0068 (0.0029)	−0.22 (0.060)
	200	w	−0.94 (0.37)	1.2	1.3	0.0042 (0.0034)	−0.25
		w	−1.2 (0.25)	1.3	1.9	0.0033 (0.0026)	−0.36
C. briggsae	200	w	−3.1 (0.26)	0.52	3.0	0.037 (0.020)	−0.10
		w	−2.4 (0.19)	2.2	2.6	0.012 (0.0055)	−0.19
O. myriophila	200	w	−1.2 (0.26)	2.7	2.1	0.0028 (0.0022)	−0.44

D. melanogaster	25	Viability[b]	−9.6[f]	0.22	1.1	0.35	−0.027 (≤0.013)
	40	Viability[b]	−11 (0.75)	0.13	2.0	0.47	−0.023 (≤0.012)
	40	Viability[b]	−4.3 [2.6/6.1]	0.30	0.87	0.14	−0.030[g]
	44	Fitness	–	–	4.1	–	–
	104–106	Viability	−1.6	0.60 (0.12)	0.91	0.02	−0.10
	210	Viability	−0.81 (0.41)	–	1.5	0.0031	−0.25
	288	w[b]	−0.30 (0.03)	–	1.1 (0.67)	0.037 (0.0018)	−0.083 (0.031)
	177–183	w[b]	−0.82 (0.43)		1.5 (0.95)	0.0015	−0.55
	27–33	Viability[b]	−6.0 [−7.8/−4.0]	–	1.2 [0.89/1.3]	0.053 [0.028/0.095]	−0.11 [0.073/0.16]
			−8.3 [−13/−1.3][f]	–	2.0 [1.2/2.6]	0.082 [0.0038/0.23]	−0.10
	31–35	Viability[b]	−8.0 [2.6/12.6][f]	–	2.2 [1.5/2.6]	0.068 [0.0085/0.19]	−0.12
	35	Viability[b]	−3.9 [−8.7/0.80][f]	–	2.4 [1.7/3.0]	0.012 [0.0/0.051]	−0.23 [−1.1/0.89]
	41	Viability[b]	−3.7 [−4.9/−2.5][f]	–	1.5 [1.1/1.7]	0.030 [0.020/0.039]	−0.11 [−0.12/−0.095]
	31	Viability[b]	−5.0 [−6.4/−3.4][f]	–	1.7 [1.4/2.0]	0.039 [0.028/0.049]	−0.11 [−0.12/−0.099]
	16	Viability[b]	−86 [−109/−65]	–	11 [9/13]	0.34 [0.20/0.55]	−0.29 [−0.40/−0.18]
	25	Viability[b]	−32 [−42/−23]	–	7 [5/8]	0.10 [0.065/0.19]	−0.33 [−0.49/−0.22]
	25	Viability[b]	−100 [−117/−87]	–	12 [10/14]	0.40 [0.26/0.66]	−0.29 [−0.40/−0.20]

VSV – vesicular stomatitis virus, w – fitness, superscript b – trait was assayed under competitive conditions, r – rate of per capita increase of the population size. Round brackets show standard errors, and square brackets show 95% confirence intervals.

For all these data, mean ΔM is -3.6×10^{-3}, mean mutational evolvability $e_m = CV_m$ is 0.017, and mean mutational heritability h^2_m is 1.3×10^{-3}. Thus, mutation introduces quantitative variation at a substantial speed: without an opposition from negative selection and drift, it will essentially destroy a trait in ~300 generations, and would double the heritable variation in <100 generations. In all the cases studied, the impact of mutation on a trait mean is much larger than what can be expected from the rate of divergence of species after a cladogenesis, testifying to preponderance of negative selection. Indeed, if a trait evolves at a rate of 1 darwin, this implies, assuming 1 generation per year, that each generation its mean value changes only by ~0.000001, which is well below the actual values of ΔM.

Estimates of the genomic deleterious mutation rate U vary widely, from 0.01 to 10.

The genomic rate of beneficial mutations is always low.

How mutation, acting alone, affects the population?

Considering mutation alone is not very realistic: mutation is a rather slow force, so we cannot ignore other forces - selection and drift. Still, this analysis is needed for understanding more complex models. The following dynamical equation connects allele frequencies in successive generation:

$$[A]_{t+1} = [A](1-\mu) + \nu(1-[A])$$

In order to find equilibria, we substitute [A], instead of $[A]_{t+1}$, into this equation. As the result, a dynamical equation is converted into an algebraic equation:

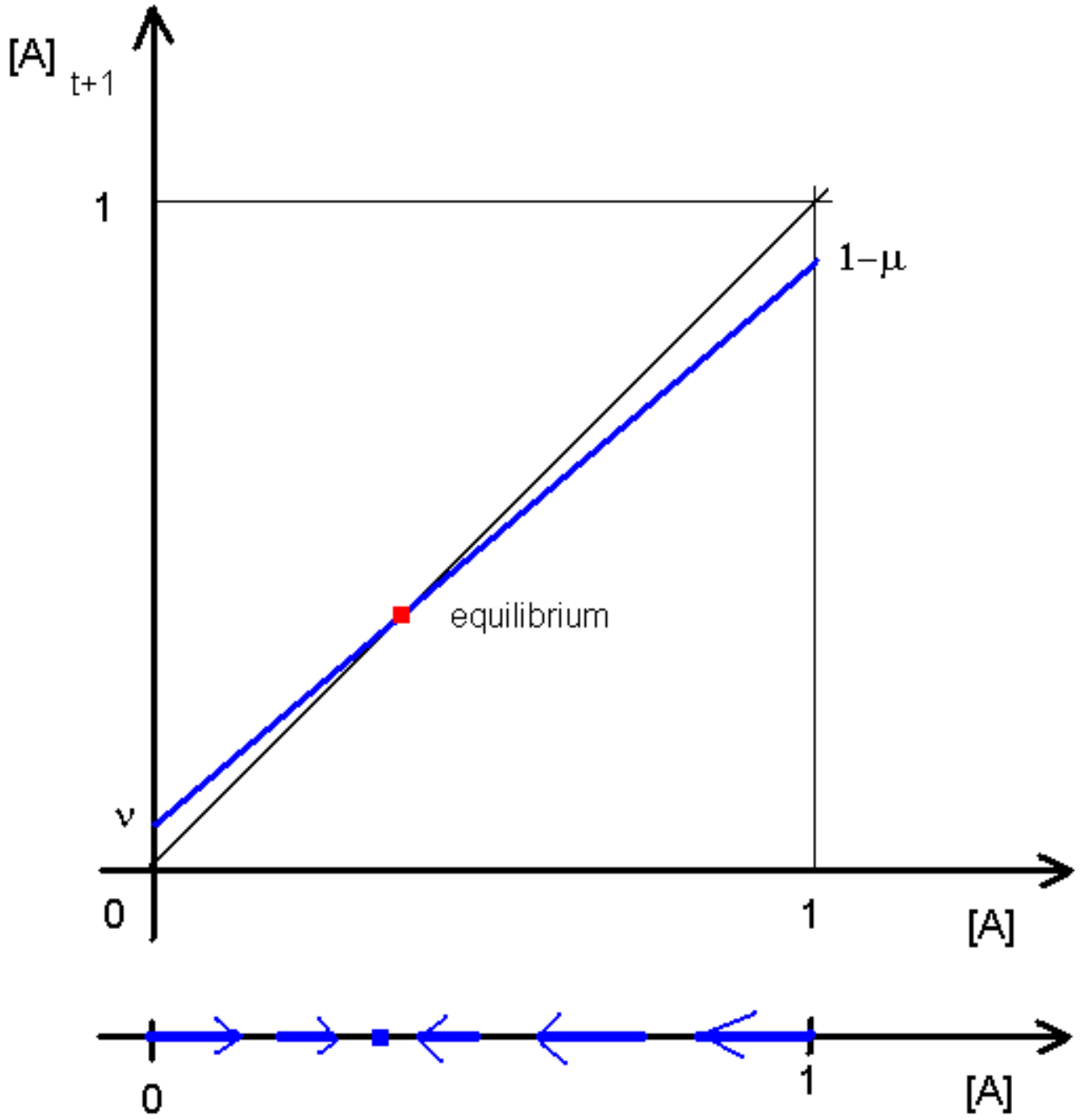
$$[A] = [A](1-\mu) + \nu(1-[A])$$

and solve it for [A]. The only solution, $[A]_{eq}$, is given by:

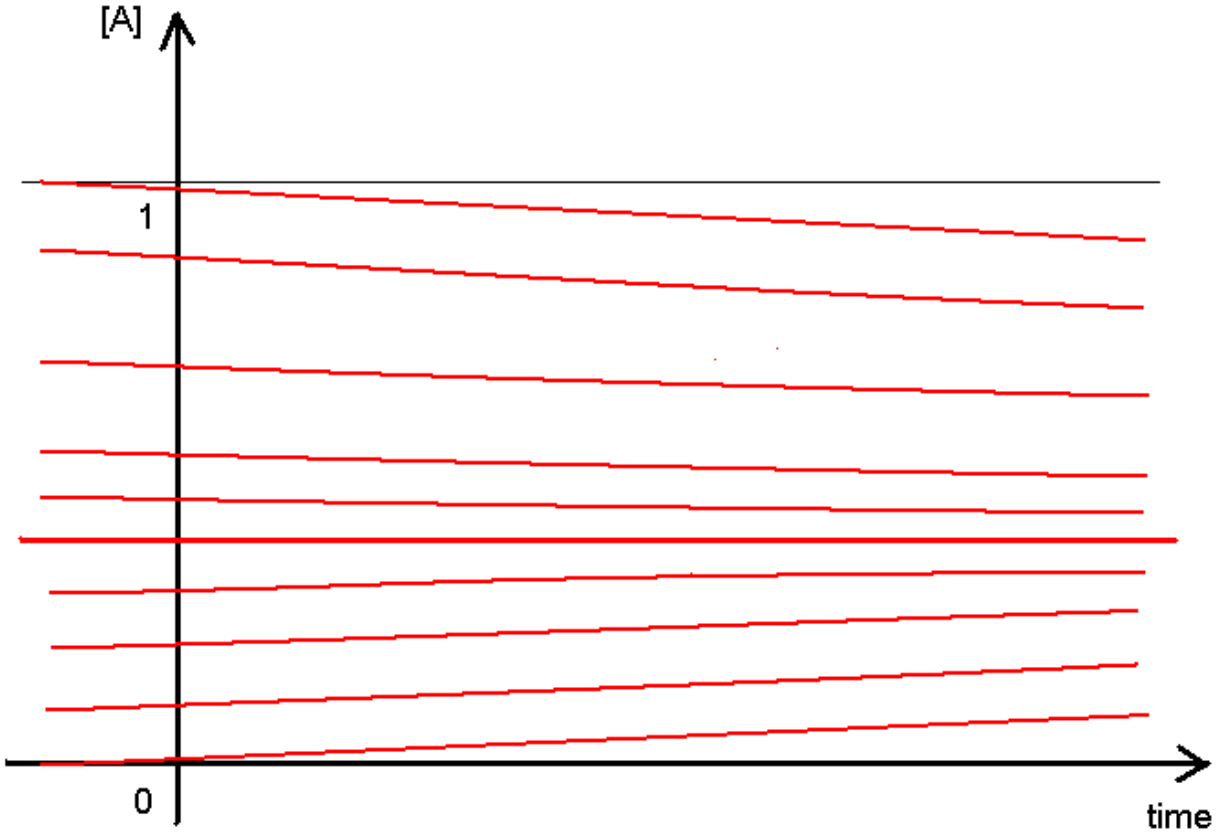
$$[A]_{eq} = \nu/(\mu+\nu)$$

Thus, if mutation acts alone, the equilibrium frequency of allele A is equal to the ratio of the mutation rate towards this allele over the sum all mutation rates. It is easy to show that this equilibrium is (globally) stable.

What is the equilibrium frequency of a?



Qualitative view on the dynamics of two alleles under mutation.



In fact, this dynamical system can be investigated completely and explicitly. The frequency of A slowly approaches its equilibrium value in the following way

$$[A](t) = \nu/(\mu+\nu) + ([A]_0 - \nu/(\mu+\nu))\exp\{-(\mu+\nu)(t-t_0)\}, \quad \text{if } ([A]_0 < \nu/(\mu+\nu))$$

$$[A](t) = \nu/(\mu+\nu) - ([A]_0 - \nu/(\mu+\nu))\exp\{-(\mu+\nu)(t-t_0)\}, \quad \text{if } ([A]_0 > \nu/(\mu+\nu))$$

You do not need to remember this formula.

More complex mutational equilibria can also be investigated, using linear algebra.

Mutational matrix:

Source: \ Destination:	A	T	G	C
A	-	$\mu_{A>T}$	$\mu_{A>G}$	$\mu_{A>C}$
T	$\mu_{T>A}$	-	$\mu_{T>G}$	$\mu_{T>C}$
G	$\mu_{G>A}$	$\mu_{G>T}$	-	$\mu_{G>C}$
C	$\mu_{C>A}$	$\mu_{C>T}$	$\mu_{C>G}$	-

Dynamical model:

$$[A](\mu_{A>T}+\mu_{A>G}+\mu_{A>C}) = [T]\mu_{T>A} +[G]\mu_{G>A} +[C]\mu_{C>A}$$
$$[T](\mu_{T>A}+\mu_{T>G}+\mu_{T>C}) = [A]\mu_{A>T} +[G]\mu_{G>T} +[C]\mu_{C>T}$$
$$[G](\mu_{G>A}+\mu_{G>T}+\mu_{G>C}) = [A]\mu_{T>G} +[T]\mu_{T>G} +[C]\mu_{C>G}$$
$$[C](\mu_{C>A}+\mu_{C>T}+\mu_{C>G}) = [A]\mu_{A>C} +[T]\mu_{T>C} +[C]\mu_{G>C}$$

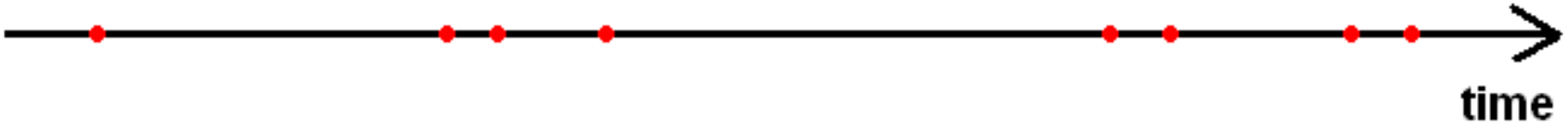
The only set of equilibirum allele frequencies can be found for this model, by solving a homogeneous system of four linear equations.

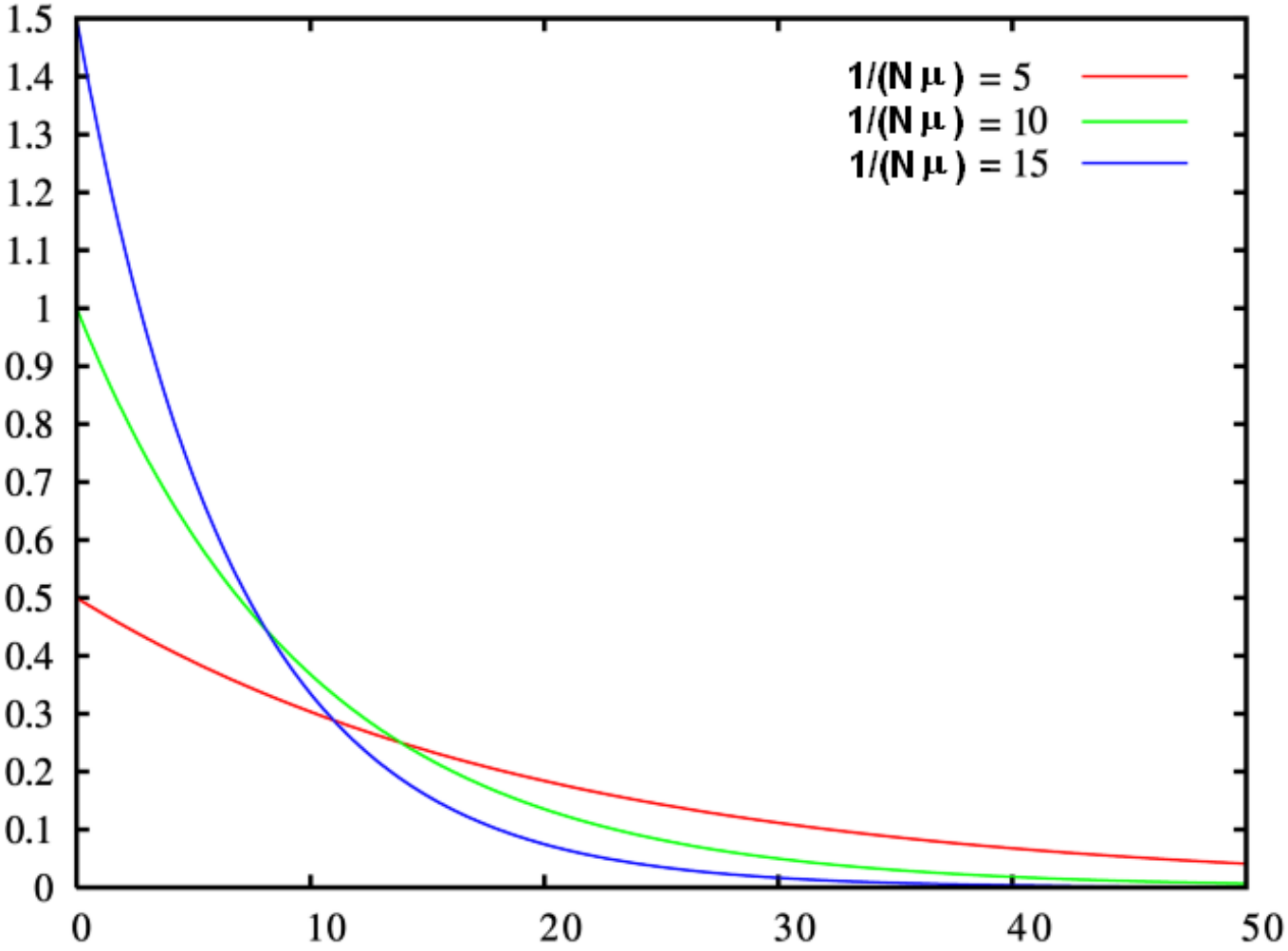
Waiting for mutation in a finite population

So far, we treated mutation as a deterministic factor. Of course, individual mutational events are, physically, quantum phenomena and, thus, are inherently stochastic. Still, the impact of mutation on a very large population can be viewed as approximately deterministic, because of the law of large numbers, analogously to the dynamics of radioactive decay. The approximation is acceptable if $N\mu >> 1$, where N is the population size, so that many mutational events are expected every generation.

However, mutation rates are so low that usually stochasticity of mutation cannot be ignored. Let us consider the opposite case of $Nm << 1$ and address the issue of origin of a currently absent allele through mutation. This issue is essential for adaptive evolution: before an adaptive replacement can happen, a beneficial mutant must appear.

When $Nm << 1$, in most of generations nothing happens, but occasionally, with probability Nm, a single new mutant appears in the population. The dynamics of these appearances are known as Poisson process. The expected waiting time t between successive appearances of mutants is $(Nm)^{-1}$, and t has an exponential probability density: $p(t) = (Nm)e^{-Nmt}$. Obviously, the necessity to wait for the appearance of a beneficial mutation can substantially impede adaptive evolution.





(top) Possible moments of appearance of rare mutations. (bottom) Exponential probability densities corresponding to expected waiting times between successive appearances of a new mutant 1/5, 1/10, and 1/15.

Quiz:

Question 1: suppose that we want to design a coding sequence that mutates slowly, and, thus, lacks CpG's. Can we always do this, given the standard genetic code?
Hint: what is the minimal and the maximal number of CpG dinucleotides in a coding sequence that encodes a pentapeptide Met Ala His Gly Arg?

Question 2: can we claim that natural coding sequences are designed in such a way that their rate of mutation is as low as possible?

Question 3: do you think that evolution should try to produce sequences with the minimal mutation rate and, generally, to reduce the mutation rate as much as possible?
Hint: nobody knows the exact answer to this – so just express you own thoughts.