



Published in final edited form as:

Bioessays. 2009 September ; 31(9): 912–920. doi:10.1002/bies.200900017.

Genomic mutation rates: What high-throughput methods can tell us

K. T. Nishant, Nadia D. Singh, and Eric Alani

Department of Molecular Biology and Genetics, Cornell University, Ithaca, NY 14853

Summary

High-throughput DNA analyses are increasingly being used to detect rare mutations in moderately sized genomes. These methods have yielded genome mutation rates that are markedly higher than those obtained using pre-genomic strategies. Recent work in a variety of organisms has shown that mutation rate is strongly affected by sequence context and genome position. These observations suggest that high-throughput DNA analyses will ultimately allow researchers to identify trans-acting factors and cis sequences that underlie mutation rate variation. Such work should provide insights on how mutation rate variability can impact genome organization and disease progression.

Keywords

high-throughput DNA analysis; DNA sequencing; mutation rate; variability; genome

Introduction

Mutations play important roles in disease progression (e.g. 1) and in shaping genome evolution and architecture. Such events can affect gene size, organization, and expression level, and can alter genetic interactions that act in recombination and sex (2–5). Mutations also play direct roles in phenotypic evolution (6). For example, gene duplication followed by divergence of the duplicated genes through mutational processes is thought to be a major mechanism for evolving novel gene functions (7).

In general, mutations fall into one of three categories: single nucleotide mutations, insertions/deletions, and chromosome rearrangements. Insertions and deletions can involve single base pairs, entire genes, or larger chromosomal regions. Single nucleotide mutations can result from exposure of the genome to endogenous and exogenous mutagens (8). These DNA damage events can occur at high frequency; for example, a single mammalian cell accumulates ~10,000 abasic (apurinic) lesions per day (9). Most DNA lesions induced by endogenous and exogenous means are recognized and repaired by well characterized DNA repair systems (10,11). In their absence, the replication of DNA containing damaged bases (e.g. oxidation, deamination, alkylation) can generate mutational events at a high rate, primarily through the loss of template information and/or the recruitment of low fidelity DNA polymerases that display error rates as high as 1 per 100–1000 nucleotides incorporated (12). Mutations in the genome can also arise as a result of errors during DNA replication. These events are rare because the nucleotide selectivity and proofreading functions of DNA polymerases, combined with post-replicative mismatch repair systems, greatly reduce the error rate to roughly 1 per 10^9 nucleotides incorporated (12).

Mutations occur in coding and non-coding regions and can be broadly classified as lethal, deleterious, neutral, or beneficial based on their fitness effects. Neutral mutations have negligible effects on fitness and are thus invisible to natural selection. Beneficial mutations are favored by positive selection and may range from mildly to highly adaptive (13–15), while deleterious mutations impose a fitness cost and tend to be removed from the population by natural selection. The relative proportion of these mutations varies between species (16) due to factors including a species' effective population size and genome organization (e.g. the percent of the genome containing coding and repetitive sequences). For example, 20–30% of amino acid mutations appear neutral in humans (17,18,19); this value is significantly lower (2.8%) in enteric bacteria (20).

Work in a number of organisms suggests that mutation rates vary as a function of genome position (21–24). In baker's yeast the mutation rate of a microsatellite reporter placed at a variety of chromosomal positions varied by 16-fold (21). A two-fold difference in substitution rate was observed in human chromosomes measured at one megabase pair resolution (24); because these rates were estimated using non-functional transposable elements, it is likely that this variation is largely due to mutation rate variation among genomic locations. Regional variation in mutation rates has been correlated with differences in base composition (23,25), local recombination rate and gene density (26), transcriptional activity (27,28), variations in repair efficiency at different sites in the genome (21), chromatin structure (29), nucleosome position (30) and replication timing (31). Some of the mutational variation described above is likely to occur in repetitive parts of the genome that are subject to less selection and appear to be less stable (32,33).

Obtaining a measure of mutation rate variation using more accurate and comprehensive high-throughput methods, combined with bioinformatic approaches, will allow for a better understanding of the cis and trans-acting factors that affect mutation rate and will likely be relevant to understanding genome evolution, organization and disease progression. This review will provide an overview of pre- and post-genomic methods to determine mutation rates and how post-genomic technologies can be used to measure mutation rate variation.

Pre-genomic determinations of mutation rates

Mutation rates are typically represented as μ_n , the mutation rate per base, or U , the mutation rate per genome. In general these rates are measured with respect to single nucleotide mutations. U can often be indicated as U_T , the total mutation rate per genome, or U_D , the deleterious mutation rate per genome based on fitness (34,35). μ_n and U are presented as the number of mutations/base/division, and the number of mutations/genome/division, respectively. For multicellular organisms, these units are typically expressed per generation instead of per division. Unless otherwise noted, all estimates for U_T and U_D are for diploid genomes. Pre-genomic mutation rate measurements in bacteria and yeast have generally been obtained using fluctuation tests (36). In these assays, a large number of parallel cultures are started with a relatively low number of wild-type cells, grown under non selective conditions, and then plated onto selective media to identify mutants. The total number of cells at the end of the growth period is determined by plating an appropriate dilution on non-selective media. The number of mutations that arise in each culture should follow a Poisson distribution that can be used to estimate the mutation rate by several methods including the commonly used method of the median (36). At the *CAN1* locus in the baker's yeast *S. cerevisiae*, the average mutation rate per base (μ_n) from the analysis of several fluctuation tests was estimated to be 1.7×10^{-10} (37). A more recent study in *S. cerevisiae* (38) provided several refinements to this analysis; μ_n was estimated, based on analysis of the *URA3* and *CAN1* loci, to be 3.80×10^{-10} and 6.44×10^{-10} , respectively. These data also support the idea that mutation rates vary as a function

of genome position, and provide further motivation to implement high-throughput methods to accurately quantify mutation rates across the genome.

The fluctuation test estimate of μ_n in *S. cerevisiae* relies on detecting mutational events in genetic markers using selection based on reversion to function or loss of function. For example, the reversion rate at three different loci was used to measure rates of spontaneous mutation during mitosis and meiosis (39). While these early estimates of mutation rates were important in terms of obtaining initial values across organisms, they were based on measurements made at a few loci that were then extrapolated to the entire genome. Such an approach is likely to be inaccurate because mutation rates can vary according to chromosome position (21 and see below) and it can often fail to detect synonymous mutations and mutations in non-coding regions.

Mutation accumulation assays have been used in several non-mammalian model systems to directly measure genome-wide mutation rates (40–43). In these assays, a set of initially isogenic lines are maintained and allowed to accumulate mutations by minimizing the effects of selection. Selection is minimized by frequent bottlenecks, where minimum effective population sizes are maintained to allow even deleterious mutations to accumulate. Different lines will independently accumulate different numbers of mutations, leading to loss of fitness (ΔM) compared to controls, and an increase in variance for fitness (ΔV) among the lines. Fitness measures commonly used in these assays are growth rate and reproductive success. Because organismal fitness is controlled by a very large number of loci, it offers the widest mutational target, allowing the recovery of most mutational events (44). From the ΔM and ΔV parameters, one can infer mutation rates computationally using the Bateman-Mukai (BM; 42), maximum likelihood (ML; 45) or minimum distance (MD; 46) methods. A comparison of these methods is described in Garcia-Dorado and Gallego (47). As shown in Table 1, computational analysis of mutational accumulation assays in *Drosophila melanogaster* has resulted in estimates of U_D ranging from 0.01 to 0.17 (43,46,48). In *C. elegans*, analogous methods have estimated U_D to be 0.005 per haploid genome (49). In *S. cerevisiae*, estimates of U_D range from 6.3×10^{-5} per haploid genome (14) to 9.5×10^{-5} per diploid genome (50).

The computational models described above assume that all mutations have similar effects on fitness. This assumption, which is almost certain to be incorrect, will cause bias in mutation rate estimates. Methods have been developed that do not make this assumption. For instance, in *S. cerevisiae*, all four meiotic products can be recovered, which facilitates directly linking a single deleterious mutational event and its fitness effect (44). Diploid clones that have acquired a single deleterious mutation can be sporulated and all four haploid products can be recovered; the growth rates of the two wild type haploids relative to the two mutant haploids are used to estimate the fitness effects of the novel mutation. Using this strategy U_D was determined in *S. cerevisiae* to be 1.1×10^{-3} per diploid genome (44). However, even this direct method can underestimate U_D as deleterious mutations with small fitness effects may not show observable growth differences among the haploid progeny, especially under laboratory conditions.

Indirect methods are often used to infer mutation rates in mammals where mutation accumulation studies and fluctuation tests are impractical. These methods measure neutral sequence differences between related species to infer mutation rates (51). Neutral substitutions are considered to be: 1. The four-fold degenerate synonymous sites in open reading frames of protein-coding sequences, 2. Pseudogene loci, 3. Repetitive DNA sequences and 4. Noncoding non-repetitive DNA. Mutation rate estimates in mammals based mostly on indirect methods and a few direct estimates are shown in Table 2.

While indirect methods have improved our estimation of mutation rates, they also have their limitations. Estimating mutation rates indirectly from phylogenetic comparisons of DNA sequences is dependent upon accurate estimates of the generation length and divergence time of a species; these measures, however, are difficult to obtain. In addition, the four-fold degenerate synonymous sites may not be neutral, as has been suggested (52,53), and non-coding DNA may be subject to high levels of selective constraint and may also evolve under positive selection, at least in some systems (e.g. 54,55).

Development of new high-throughput methods to measure mutation rates

The above pre-genomic approaches typically determine mutation rates based on limited sequence analyses at a few loci or fitness based assays. Since most mutations that confer phenotypes are thought to be deleterious, these small-scale approaches can skew the distribution towards observable mutations (16,40,48,50,56–59). In addition, the heterogeneity in the fitness effects of mutations makes it difficult to accurately infer mutation rate in fitness based mutation accumulation assays (60), where the fitness effects of all mutations are assumed to be the same. High-throughput genome wide measurements of mutation can reduce concerns about skewed mutation distributions and variable fitness effects because mutations are directly detected.

High-throughput measurements can be performed using traditional sequencing methods such as Sanger sequencing or by other methods that detect sequence variants such as single strand conformation polymorphism (SSCP) and denaturing high performance liquid chromatography (DHPLC; 61). However, these methods are in general labor and cost intensive. An alternative is to use new sequencing and microarray technologies that provide rapid, accurate, and cost effective mutational profiling at a genomic scale. Three new sequencing technologies are commercially available that produce short sequence reads (25–200 bp) using a massively parallel approach. In general, a reference genome, which is available for almost all model organisms, is required to assemble these reads. These technologies have been developed by Illumina (Genome Analyzer), Roche (Genome Sequencer FLX) and ABI (SOLiDS). These and other emerging technologies are reviewed exhaustively (62,63). Microarray approaches have been developed that offer a viable alternative to whole genome sequencing by identifying mutations based on differential hybridization to oligonucleotide tiling arrays. Such arrays allow the entire genome to be interrogated at single base level for mutations which can be identified by re-sequencing (64).

High-throughput DNA analysis approaches have led to a considerable amount of new information on single base mutation rates in model organisms. In particular, estimates of mutation parameters μ_n , U_T , and U_D have been significantly revised upwards; these increases most likely reflect the greater sensitivity of high-throughput approaches for detecting mutation events (Table 3). Post-genomic estimates of U_D appear high in *D. melanogaster* and *C. elegans* and are similar to values determined in mammals using pre-genomic estimates (Tables 2 and 3). This suggests that high deleterious mutation rates are not unique to mammalian genomes.

In *S. cerevisiae*, μ_n was determined to be 3.3×10^{-10} by pyrosequencing the genomes of four mutation accumulation lines to a 5-fold average genome coverage (65). U_T was estimated to be approximately 0.32 per haploid genome and mostly comprised homopolymeric mutations (0.30) that were shown to have very high mutation rates. Although comparison with pre-genomic estimates of U_D that are in the range of 10^{-5} to 10^{-3} suggests that only a very small percentage (0.1%) of mutations confer fitness defects that can be detected in laboratory assays, it is not clear what fraction of the homopolymeric mutations are deleterious (see 66).

In *C. elegans*, a total of 4 Mb was sequenced from mutation accumulation lines at different generations using the Sanger method (67). μ_n was estimated to be 2.1×10^{-8} per base, which is an order of magnitude higher than pre-genomic estimates based on laboratory fitness assays (34, Table 3). U_T was estimated to be 2.1 per haploid genome which, when compared to pre-genomic haploid estimates of U_D (0.005), suggested that most mutations (99%) in these lines have fitness defects that are not easily seen in the laboratory. U_D was inferred to be 0.48 per haploid genome which is two orders of magnitude higher than the pre-genomic estimates and highlights the drawbacks of inferring them based on laboratory fitness assays. More than half of the mutations were small insertions or deletions (17 insertion-deletion events out of 30 mutations observed) instead of single base mutation events as assumed earlier (68). A high estimate of U_D (0.14 per diploid genome per generation for protein coding genes) was also inferred by Davies et al. (69) by comparing the number of deleterious mutations detected at the molecular level in forward and reverse mutation assays with fitness based assays. This comparison suggested that greater than 96% of the deleterious mutations fixed in the mutation accumulation lines have fitness effects too subtle to be detected based on laboratory fitness assays.

In *Drosophila melanogaster* the mutation rate per base (μ_n) was estimated to be 8.4×10^{-9} by scanning 20 Mb of DNA from three sets of mutation accumulation lines using denaturing high-performance liquid chromatography (70). This estimate is about 24 fold higher than pre-genomic estimates (34). U_D was estimated to be 1.2, which is again higher than pre-genomic estimates from computational analysis of mutation accumulation lines (48). Significant heterogeneity in the mutation rate was also seen between the three lines. μ_n was 4.8×10^{-9} for the Madrid line, 17.2×10^{-9} for the Florida-33 line, and 6.8×10^{-9} for the Florida-39 line. The ability to detect mutation rate variation between lines/individuals is one of the advantages of using high-throughput approaches as opposed to the population-based estimates that are obtained from pre-genomic methods. Unlike the case in *C. elegans*, single base substitution mutations comprised the majority of the mutations.

The post-genomic estimates of μ_n appear somewhat similar in *C. elegans* and *D. melanogaster* (2.5 fold difference, Table 3) but are markedly lower in *S. cerevisiae* (25 and 63 fold lower relative to *D. melanogaster* and *C. elegans*, respectively). Mutation rate estimates in multicellular organisms such as *C. elegans* and *D. melanogaster* can appear magnified because the values reflect per generation estimates (number of mutations per division multiplied by the number of germ line divisions per generation) rather than per division as determined for unicellular organisms. Even after taking this into account, it is likely that multicellular organisms bear an increased mutational load that makes them more susceptible to the effect of deleterious mutations (71; U_D estimates in Table 3). It is fascinating that higher mutation rates in multicellular organisms do not appear to interfere with the evolution of complexity in multicellular organisms (72). One explanation for the appearance of higher mutation rates in multicellular organisms is that the lower effective population sizes of these organisms contribute to higher mutation rates by increasing the role of genetic drift in fixing mutator alleles (2).

Two problems associated with the new high-throughput sequencing technologies are the inability to obtain mappable sequence information from repetitive regions of the genome, and the high error rates associated with sequence detection (73). Repetitive regions are significantly underrepresented in the useable output of current high-throughput sequencing methods due to the difficulties in mapping repetitive sequences to unique chromosomal positions using short read data. As described below this can be a major concern because mutation rates in repetitive regions are likely to be significantly higher than in non-repetitive regions. In addition, mutations found in only a subset of the repeat copies are often masked as sequencing errors because it is difficult to determine the origin of any particular sequence read (74). Also, DNA

sequencing errors, which are often specific to the new technology used, can make it challenging to identify heterozygous mutations in diploid genomes (64,75). Lack of adequate genome wide coverage during sequencing can also contribute to sequence errors that appear as mutational events. At present the overall impact of false negatives and positives on mutation rate is unclear.

Many of the sequencing errors associated with the new technologies cannot be resolved by programs designed for the Sanger sequencing method such as Polyphred; therefore, observed mutations have to be carefully analyzed (76). New algorithms are being developed to provide base quality scores for these new sequencing methods (e.g. 77), and to map short repeat sequence reads to a reference genome (78). Also, some of the base calling errors characteristic to the new sequencing technologies, such as higher error rates towards the end of sequencing reads, have been estimated (79) and methods to better detect these errors are being developed (80,81). Since most new mutations that arise in diploid organisms are heterozygous, detecting them using these new sequencing technologies is challenging, but will likely be overcome with increased sequence coverage and verification through other methods such as Sanger sequencing.

Applying improved mutation rate estimates to understand molecular evolution

Accurate estimates of mutation rate parameters μ_n and U are crucial from a molecular evolution perspective because they are used to fix baseline mutation rates within a species. These values are in turn useful for mutation rate comparisons under altered environmental and growth conditions. Improvements in mutation rate measurements will be useful for accurately determining the rate of deleterious mutations that are not efficiently removed by selection and can thus contribute to mutational loads that accumulate and cause the extinction of small populations (82). Because humans pass on roughly 100 new mutations to their offspring (83), knowing how many of these mutations are beneficial, neutral or deleterious has implications for the long term fitness of a species that produces few offspring (84).

More accurate estimates of mutation rate are crucial in an evolutionary context as well (51). For instance, more informed estimates of mutational load are of primary importance from a theoretical perspective in terms of understanding the evolution of sex and the evolution of recombination. In addition, because the number of neutral substitutions per site (K) is a function of time and mutation rate ($K = 2\mu T$), increasing the accuracy of mutation rate (μ) estimates will improve our estimates of divergence times among species under the assumption of a molecular clock. Finally, the most fundamental population genetic parameter is arguably $\theta = 4N_e\mu$, where N_e is the effective population size. This equation describes the level of neutral variation in a population. This parameter, which depends critically on mutation rate, is indispensable in population genetic models, which are used to infer patterns of selection and demography based on extant population-level sequence data. Improved estimates of mutation rate will help inform our understanding of the strength and frequency of adaptive events, the distribution of selective constraint across the genome, and the effects of population history on population-level variation.

The challenge ahead: Using high-throughput methods to determine the scale of mutation rate variation

As described above, a major drawback of most high-throughput DNA analyses is that they are unable to detect mutations in highly repetitive sequences. These include simple repeat sequences that have repeat lengths of less than 300 bp such as microsatellites, minisatellites, *Alu* repeats and telomeric repeats as well as much longer repeats (several kb) like LINE elements and rDNA repeats. Repetitive DNA comprises as much as 50% of the human genome

(85) and a substantial portion (~17–57%) of the genome in most model organisms (e.g. 86, 87). Repetitive regions are prone to a wider range of mutational events such as insertion-deletion mutations and rearrangements (88). In addition, single nucleotide substitution events are known to be higher near insertion/deletion mutations (89). Studies performed primarily in bacteria, yeast and *Drosophila* have shown that frameshift mutations in repetitive regions can occur at up to several orders of magnitude higher than in non-repetitive regions (e.g. 90–92). These mutagenic events are likely to dramatically affect the fitness of an organism if they occur in the open reading frame of a gene. For example, Heck et al. (66) searched the *S. cerevisiae* genome and identified greater than 600 seven-nucleotide repeat runs in essential genes and calculated that strains grown for 160 generations display a 7×10^{-4} probability of acquiring a mutation in one or more of these runs. Thus, essential genes containing simple sequence repeats within open reading frames are at risk for disruption. Simple repetitive sequences have also been identified within developmental genes (93). The repetitive sequences in these genes are thought to contribute to rapid morphological evolution by contributing to localized genetic variation without causing a general increase in mutational load (94,95).

The presence of localized regions of the genome with different mutation rates can have important consequences for the evolution field. Recent work examining cryptic mutation hotspots indicates that mutation rate variation has been grossly underestimated in the human genome (96). The substantial mutation rate variation within a genome makes previous calculations of average mutation rates μ_n and U , initially measured to be relatively constant across species (34), less useful in terms of estimating population history and species divergence time. The existence of such variation also makes it more difficult to determine the roles of selection and mutation in maintaining conserved genomic regions. Identifying regional variation in mutation rates using bioinformatic analyses of high-throughput data is likely to be useful in terms of identifying sequences/regions that correlate with both high and low mutation rates, and should allow us to distinguish between the role played by selection or low mutation pressure in maintaining conserved genomic regions (24). Already such approaches have borne fruit. For example, recent work in yeast suggests that linker DNA has a 10–15 % lower substitution rate than nucleosomal DNA (30). Future work in this field will likely involve an analysis of the relative contributions of mutation rate versus selective constraint in establishing nucleosome positioning (30). Regional variation in mutation rate has also been hypothesized to influence the spatial distribution of genes (3).

Genome wide mutation rate measures and local heterogeneity in mutation rates are relevant from a human disease perspective. Knowledge of the mutation rate associated with different tumors can be useful for clinical therapy. Greenman et al. (97) re-sequenced 274 Mb of DNA corresponding to coding exons of 518 protein kinase genes in 210 human cancers. Strikingly, the prevalence of mutations in different types of cancers was different. For example lung and gastric cancers showed the highest prevalence of mutations (4.21 and 2.10 mutations per Mb, respectively) while testis and breast cancers showed lower levels (0.12 and 0.19 mutations per Mb, respectively). Tumors with higher mutation rates are predicted to require treatment with multiple drugs in order to avoid drug resistance (98). Chromosomal regions with higher mutation rates may also be more susceptible to DNA damage. Such sites could be responsible for chromosome instability that is associated with tumorigenesis (99).

Lastly, the new DNA sequencing methods will likely identify patterns of mutagenesis that would be difficult to detect by measuring mutations at only a few loci. For example, mutations have been reported to occur in showers in systems ranging from viruses (100) to mice (101). Mutational showers, defined as regions containing mutations at levels higher than predicted by mutation rate and random distribution, are hypothesized to occur due to a transient hypermutable state of a fraction of the population (102,103). Such clustering of mutations in genomic space was seen in transcription-associated mutagenesis in bacteria and yeast (27,

104). Mutations mediated by error-prone DNA double strand break repair pathways are also thought to create mutation clusters around double-strand break sites (103,105–108). The new sequencing technologies offer an efficient way to characterize subpopulations with different mutation rates because the expense of sequencing multiple populations is relatively modest.

Concluding Remarks

We are looking forward to the development of more accurate high-throughput methods that can be used to identify *de novo* mutation events in both unique and repetitive regions of a diploid genome. Such measurements, performed on a large scale, will provide more accurate estimates of mutation rates that can ultimately be used to identify trans acting factors and cis sequences that affect mutation rate variation. Work by Lynch *et al.* (65), Denver *et al.* (67) and Haag-Liautard *et al.* (70) provide excellent examples of how this work will be pursued. Traditional approaches that seek to summarize mutation rate information over a genome are being replaced by high-throughput approaches that will provide a better estimation of mutation rate variation that result from distinct mutation formation mechanisms (89,94,96). Such achievements will ultimately allow scientists to measure mutation rate variation associated with different drug treatments, biological processes (e.g. different stages of the cell cycle), environmental conditions, and nutritional and disease states (e.g. 97). Ultimately the high-throughput technologies will allow one to determine mutation rate variation between individuals at specific regions in the genome. This is particularly relevant in the coming era of personalized medicine for estimating genetic disease risk.

Abbreviations

BM	Bateman-Mukai
ML	maximum likelihood
MD	minimum distance
μ_n	mutation rate per base
U	the mutation rate per genome

Acknowledgments

We are grateful to Charles Aquadro and members of the Alani laboratory for discussions and comments on the manuscript. E. A. and K. T. N. were supported by NIH grant GM53085. N. D. S. is supported by NIH grant number 1F32GM080944-01 to N. D. S., Charles F. Aquadro, and Andrew G. Clark.

References

1. Vogelstein B, Kinzler KW. Cancer genes and the pathways they control. *Nat Med* 2004;10:789–799. [PubMed: 15286780]
2. Lynch M. The origins of eukaryotic gene structure. *Mol Biol Evol* 2006;23:450–468. [PubMed: 16280547]
3. Chuang JH, Li H. Functional bias and spatial organization of genes in mutational hot and cold regions in the human genome. *PLoS Biol* 2004;2:E29. [PubMed: 14966531]
4. Keightley PD, Otto SP. Interference among deleterious mutations favours sex and recombination in finite populations. *Nature* 2006;443:89–92. [PubMed: 16957730]
5. Rifkin SA, Houle D, Kim J, White KP. A mutation accumulation assay reveals a broad capacity for rapid evolution of gene expression. *Nature* 2005;438:220–223. [PubMed: 16281035]
6. Nei M. The new mutation theory of phenotypic evolution. *Proc Natl Acad Sci USA* 2007;104:12235–12242. [PubMed: 17640887]

7. Teichmann SA, Babu MM. Gene regulatory network growth by duplication. *Nat Genet* 2004;36:492–496. [PubMed: 15107850]
8. Friedberg EC, McDaniel LD, Schultz RA. The role of endogenous and exogenous DNA damage and mutagenesis. *Curr Opin Genet Dev* 2004;14:5–10. [PubMed: 15108798]
9. Lindahl T, Nyberg B. Rate of depurination of native deoxyribonucleic acid. *Biochemistry* 1972;11:3610–3618. [PubMed: 4626532]
10. Iyer RR, Pluciennik A, Burdett V, Modrich PL. DNA mismatch repair: functions and mechanisms. *Chem Rev* 2006;106:302–323. [PubMed: 16464007]
11. Wyman C, Kanaar R. DNA double-strand break repair: all's well that ends well. *Annu Rev Genet* 2006;40:363–383. [PubMed: 16895466]
12. McCulloch SD, Kunkel TA. The fidelity of DNA synthesis by eukaryotic replicative and translesion synthesis polymerases. *Cell Res* 2008;18:148–161. [PubMed: 18166979]
13. Shaw FH, Geyer CJ, Shaw RG. A comprehensive model of mutations affecting fitness and inferences for *Arabidopsis thaliana*. *Evolution* 2002;56:453–463. [PubMed: 11989677]
14. Joseph SB, Hall DW. Spontaneous mutations in diploid *Saccharomyces cerevisiae*: more beneficial than expected. *Genetics* 2004;168:1817–1825. [PubMed: 15611159]
15. Giraud A, Matic I, Tenaillon O, Clara A, Radman M, et al. Costs and benefits of high mutation rates: adaptive evolution of bacteria in the mouse gut. *Science* 2001;291:2606–2608. [PubMed: 11283373]
16. Eyre-Walker A, Keightley PD. The distribution of fitness effects of new mutations. *Nat Rev Genet* 2007;8:610–618. [PubMed: 17637733]
17. Fay JC, Wyckoff GJ, Wu CI. Positive and negative selection on the human genome. *Genetics* 2001;158:1227–1234. [PubMed: 11454770]
18. Boyko AR, Williamson SH, Indap AR, Degenhardt JD, Hernandez RD, et al. Assessing the evolutionary impact of amino acid mutations in the human genome. *PLoS Genet* 2008;4:e1000083. [PubMed: 18516229]
19. Keightley PD, Eyre-Walker A. Joint inference of the distribution of fitness effects of deleterious mutations and population demography based on nucleotide polymorphism frequencies. *Genetics* 2007;177:2251–2261. [PubMed: 18073430]
20. Charlesworth J, Eyre-Walker A. The rate of adaptive evolution in enteric bacteria. *Mol Biol Evol* 2006;23:1348–1356. [PubMed: 16621913]
21. Hawk JD, Stefanovic L, Boyer JC, Petes TD, Farber RA. Variation in efficiency of DNA mismatch repair at different sites in the yeast genome. *Proc Natl Acad Sci USA* 2005;102:8639–8643. [PubMed: 15932942]
22. Wolfe KH, Sharp PM, Li WH. Mutation rates differ among regions of the mammalian genome. *Nature* 1989;337:283–285. [PubMed: 2911369]
23. Matassi G, Sharp PM, Gautier C. Chromosomal location effects on gene sequence evolution in mammals. *Curr Biol* 1999;9:786–791. [PubMed: 10469563]
24. Arndt PF, Hwa T, Petrov DA. Substantial regional variation in substitution rates in the human genome: importance of GC content, gene density, and telomere-specific effects. *J Mol Evol* 2005;60:748–763. [PubMed: 15959677]
25. Hardison RC, Roskin KM, Yang S, Diekhans M, Kent WJ, et al. Covariation in frequencies of substitution, deletion, transposition, and recombination during eutherian evolution. *Genome Res* 2003;13:13–26. [PubMed: 12529302]
26. Lercher MJ, Hurst LD. Human SNP variability and mutation rate are higher in regions of high recombination. *Trends Genet* 2002;18:337–340. [PubMed: 12127766]
27. Datta A, Jinks-Robertson S. Association of increased spontaneous mutation rates with high levels of transcription in yeast. *Science* 1995;268:1616–1619. [PubMed: 7777859]
28. Lippert MJ, Chen Q, Liber HL. Increased transcription decreases the spontaneous mutation rate at the thymidine kinase locus in human cells. *Mutat Res* 1998;401:1–10. [PubMed: 9639662]
29. Teytelman L, Eisen MB, Rine J. Silent but not static: accelerated base-pair substitution in silenced chromatin of budding yeasts. *PLoS Genet* 2008;4:e1000247. [PubMed: 18989454]
30. Washietl S, Machné R, Goldman N. Evolutionary footprints of nucleosome positions in yeast. *Trends Genet* 2008;24:583–587. [PubMed: 18951646]

31. Stamatoyannopoulos JA, Adzhubei I, Thurman RE, Kryukov GV, Mirkin SM, Sunyaev SR. Human mutation rate associated with DNA replication timing. *Nat Genet* 2009;41:393–395. [PubMed: 19287383]
32. Deininger PL, Moran JV, Batzer MA, Kazazian HHJ. Mobile elements and mammalian genome evolution. *Curr Opin Genet Dev* 2003;13:651–658. [PubMed: 14638329]
33. Argueso JL, Westmoreland J, Mieczkowski PA, Gawel M, Petes TD, et al. Double-strand breaks associated with repetitive DNA can reshape the genome. *Proc Natl Acad Sci USA* 2008;105:11845–11850. [PubMed: 18701715]
34. Drake JW, Charlesworth B, Charlesworth D, Crow JF. Rates of spontaneous mutation. *Genetics* 1998;148:1667–1686. [PubMed: 9560386]
35. Baer CF, Miyamoto MM, Denver DR. Mutation rate variation in multicellular eukaryotes: causes and consequences. *Nat Rev Genet* 2007;8:619–631. [PubMed: 17637734]
36. Rosche WA, Foster PL. Determining mutation rates in bacterial populations. *Methods* 2000;20:4–17. [PubMed: 10610800]
37. Drake JW. A constant rate of spontaneous mutation in DNA-based microbes. *Proc Natl Acad Sci USA* 1991;88:7160–7164. [PubMed: 1831267]
38. Lang GI, Murray AW. Estimating the per-base-pair mutation rate in the yeast *Saccharomyces cerevisiae*. *Genetics* 2008;178:67–82. [PubMed: 18202359]
39. Magni GE, Von Borstel RC. Different rates of spontaneous mutation during mitosis and meiosis in yeast. *Genetics* 1962;47:1097–1108. [PubMed: 17248123]
40. Kibota TT, Lynch M. Estimate of the genomic mutation rate deleterious to overall fitness in *E. coli*. *Nature* 1996;381:694–696. [PubMed: 8649513]
41. Korona R. Unpredictable fitness transitions between haploid and diploid strains of the genetically loaded yeast *Saccharomyces cerevisiae*. *Genetics* 1999;151:77–85. [PubMed: 9872949]
42. Mukai T. The genetic structure of natural populations of *Drosophila melanogaster*. I. Spontaneous mutation rate of polygenes controlling viability. *Genetics* 1964;50:1–19. [PubMed: 14191352]
43. Mukai T, Chigusa SI, Mettler LE, Crow JF. Mutation rate and dominance of genes affecting viability in *Drosophila melanogaster*. *Genetics* 1972;72:335–355. [PubMed: 4630587]
44. Wloch DM, Szafraniec K, Borts RH, Korona R. Direct estimate of the mutation rate and the distribution of fitness effects in the yeast *Saccharomyces cerevisiae*. *Genetics* 2001;159:441–452. [PubMed: 11606524]
45. Keightley PD. The distribution of mutation effects on viability in *Drosophila melanogaster*. *Genetics* 1994;138:1315–1322. [PubMed: 7896110]
46. García-Dorado A. The Rate and Effects Distribution of Viability Mutation in *Drosophila*: Minimum Distance Estimation. *Evolution* 1997;51:1130–1139.
47. García-Dorado A, Gallego A. Comparing analysis methods for mutation-accumulation data: a simulation study. *Genetics* 2003;164:807–819. [PubMed: 12807799]
48. Fry JD, Keightley PD, Heinsohn SL, Nuzhdin SV. New estimates of the rates and effects of mildly deleterious mutation in *Drosophila melanogaster*. *Proc Natl Acad Sci USA* 1999;96:574–579. [PubMed: 9892675]
49. Keightley PD, Bataillon TM. Multigeneration maximum-likelihood analysis applied to mutation-accumulation experiments in *Caenorhabditis elegans*. *Genetics* 2000;154:1193–1201. [PubMed: 10757763]
50. Zeyl C, DeVisser JA. Estimates of the rate and distribution of fitness effects of spontaneous mutation in *Saccharomyces cerevisiae*. *Genetics* 2001;157:53–61. [PubMed: 11139491]
51. Kimura M. Evolutionary rate at the molecular level. *Nature* 1968;217:624–626. [PubMed: 5637732]
52. Ikemura T. Correlation between the abundance of yeast transfer RNAs and the occurrence of the respective codons in protein genes. Differences in synonymous codon choice patterns of yeast and *Escherichia coli* with reference to the abundance of isoaccepting transfer RNAs. *J Mol Biol* 1982;158:573–597. [PubMed: 6750137]
53. Shields DC, Sharp PM, Higgins DG, Wright F. "Silent" sites in *Drosophila* genes are not neutral: evidence of selection among synonymous codons. *Mol Biol Evol* 1988;5:704–716. [PubMed: 3146682]

54. Bergman CM, Kreitman M. Analysis of conserved noncoding DNA in *Drosophila* reveals similar constraints in intergenic and intronic sequences. *Genome Res* 2001;11:1335–1345. [PubMed: 11483574]
55. Andolfatto P. Adaptive evolution of non-coding DNA in *Drosophila*. *Nature* 2005;437:1149–1152. [PubMed: 16237443]
56. Lynch M, Blanchard J, Houle D, Kibota T, Schultz S, et al. Spontaneous deleterious mutation. *Evolution* 1999;53:645–663.
57. Keightley PD, Lynch M. Toward a realistic model of mutations affecting fitness. *Evolution* 2003;57:683–685. [PubMed: 12703958]
58. Vassilieva LL, Lynch M. The rate of spontaneous mutation for life-history traits in *Caenorhabditis elegans*. *Genetics* 1999;151:119–129. [PubMed: 9872953]
59. Schultz ST, Lynch M, Willis JH. Spontaneous deleterious mutation in *Arabidopsis thaliana*. *Proc Natl Acad Sci USA* 1999;96:11393–11398. [PubMed: 10500187]
60. Eyre-Walker A, Woolfit M, Phelps T. The distribution of fitness effects of new deleterious amino acid mutations in humans. *Genetics* 2006;173:891–900. [PubMed: 16547091]
61. Gross E, Arnold N, Goette J, Schwarz-Boeger U, Kiechle MA. Comparison of BRCA1 mutation analysis by direct sequencing, SSCP and DHPLC. *Hum Genet* 1999;105:72–78. [PubMed: 10480358]
62. Mardis ER. The impact of next-generation sequencing technology on genetics. *Trends Genet* 2008;24:133–141. [PubMed: 18262675]
63. Metzker ML. Emerging technologies in DNA sequencing. *Genome Res* 2005;15:1767–1776. [PubMed: 16339375]
64. Gresham D, Dunham MJ, Botstein D. Comparing whole genomes using DNA microarrays. *Nat Rev Genet* 2008;9:291–302. [PubMed: 18347592]
65. Lynch M, Sung W, Morris K, Coffey N, Landry CR, et al. A genome-wide view of the spectrum of spontaneous mutations in yeast. *Proc Natl Acad Sci USA* 2008;105:9272–9277. [PubMed: 18583475]
66. Heck JA, Gresham D, Botstein D, Alani E. Accumulation of recessive lethal mutations in *Saccharomyces cerevisiae mlh1* mismatch repair mutants is not associated with gross chromosomal rearrangements. *Genetics* 2006;174:519–523. [PubMed: 16816424]
67. Denver DR, Morris K, Lynch M, Thomas WK. High mutation rate and predominance of insertions in the *Caenorhabditis elegans* nuclear genome. *Nature* 2004;430:679–682. [PubMed: 15295601]
68. Denver DR, Morris K, Thomas WK. Phylogenetics in *Caenorhabditis elegans*: an analysis of divergence and outcrossing. *Mol Biol Evol* 2003;20:393–400. [PubMed: 12644560]
69. Davies EK, Peters AD, Keightley PD. High frequency of cryptic deleterious mutations in *Caenorhabditis elegans*. *Science* 1999;285:1748–1751. [PubMed: 10481013]
70. Haag-Liautard C, Dorris M, Maside X, Macaskill S, Halligan DL, et al. Direct estimation of per nucleotide and genomic deleterious mutation rates in *Drosophila*. *Nature* 2007;445:82–85. [PubMed: 17203060]
71. Dall SR, Cuthill IC. Mutation rates: does complexity matter? *J. Theor Biol* 1999;198:283–285. [PubMed: 10339399]
72. Haygood R. Proceedings of the SBE Tri-National Young Investigators' Workshop 2005. Mutation rate and the cost of complexity. *Mol Biol Evol* 2006;23:957–963. [PubMed: 16469852]
73. Strausberg RL, Levy S, Rogers YH. Emerging DNA sequencing technologies for human genomic medicine. *Drug Discov Today* 2008;13:569–577.
74. James SA, O'Kelly MJ, Carter DM, Davey RP, van Oudenaarden A, et al. Repetitive sequence variation and dynamics in the ribosomal DNA array of *Saccharomyces cerevisiae* as revealed by whole genome resequencing. *Genome Res* 2009;19:626–635. [PubMed: 19141593]
75. Margulies M, Egholm M, Altman WE, Attiya S, Bader JS, et al. Genome sequencing in microfabricated high-density picolitre reactors. *Nature* 2005;437:376–380. [PubMed: 16056220]
76. Pop M, Salzberg SL. Bioinformatics challenges of new sequencing technology. *Trends Genet* 2008;24:142–149. [PubMed: 18262676]

77. Brockman W, Alvarez P, Young S, Garber M, Giannoukos G, et al. Quality scores and SNP detection in sequencing-by-synthesis systems. *Genome Res* 2008;18:763–770. [PubMed: 18212088]
78. Pomraning KR, Smith KM, Freitag M. Genome-wide high throughput analysis of DNA methylation in eukaryotes. *Methods* 2009;47:142–150. [PubMed: 18950712]
79. Dohm JC, Lottaz C, Borodina T, Himmelbauer H. Substantial biases in ultra-short read data sets from high-throughput DNA sequencing. *Nucleic Acids Res* 2008;36:e105. [PubMed: 18660515]
80. Smith AD, Xuan Z, Zhang MQ. Using quality scores and longer reads improves accuracy of Solexa read mapping. *BMC Bioinformatics* 2008;9:128. [PubMed: 18307793]
81. Dolan PC, Denver DR. TileQC: a system for tile-based quality control of Solexa data. *BMC Bioinformatics* 2008;9:250. [PubMed: 18507856]
82. Lynch M, Conery J, Burger R. Mutation accumulation and the extinction of small populations. *Am Nat* 1995;146:489–518.
83. Kondrashov AS. Direct estimates of human per nucleotide mutation rates at 20 loci causing Mendelian diseases. *Hum Mutat* 2002;21:12–27. [PubMed: 12497628]
84. Reed FA, Aquadro CF. Mutation, selection and the future of human evolution. *Trends Genet* 2006;22:479–484. [PubMed: 16857288]
85. International human genome consortium. Initial sequencing and analysis of the human genome. *Nature* 2001;409:860–921. [PubMed: 11237011]
86. Mouse genome sequencing consortium, Initial sequencing and comparative analysis of the mouse genome. *Nature* 2002;420:520–562. [PubMed: 12466850]
87. Stein LD, Bao Z, Blasiar D, Blumenthal T, Brent MR, et al. The genome sequence of *Caenorhabditis briggsae*: a platform for comparative genomics. *PLoS Biol* 2003;1:E45. [PubMed: 14624247]
88. Batzer MA, Deininger PL. Alu repeats and human genomic diversity. *Nat Genet* 2002;3:370–379.
89. Tian D, Wang Q, Zhang P, Araki H, Yang S, et al. Single-nucleotide mutation rate increases close to insertions/deletions in eukaryotes. *Nature* 2008;455:105–108. [PubMed: 18641631]
90. Schug MD, Hutter CM, Wetterstrand KA, Gaudette MS, Mackay TF, et al. The mutation rates of di-, tri- and tetranucleotide repeats in *Drosophila melanogaster*. *Mol Biol Evol* 1988;15:1751–1760. [PubMed: 9866209]
91. Tran HT, Keen JD, Krickler M, Resnick MA, Gordenin DA. Hypermutability of homonucleotide runs in mismatch repair and DNA polymerase proofreading yeast mutants. *Mol Cell Biol* 1997;17:2859–2865. [PubMed: 9111358]
92. Harfe BD, Jinks-Robertson S. Sequence composition and context effects on the generation and repair of frameshift intermediates in mononucleotide runs in *Saccharomyces cerevisiae*. *Genetics* 2000;156:571–578. [PubMed: 11014807]
93. Karlin S, Burge C. Trinucleotide repeats and long homopeptides in genes and proteins associated with nervous system disease and development. *Proc Natl Acad Sci USA* 1996;93:1560–1565. [PubMed: 8643671]
94. Kashi Y, King DG. Simple sequence repeats as advantageous mutators in evolution. *Trends Genet* 2006;22:253–259. [PubMed: 16567018]
95. Fondon JW 3rd, Garner HR. Molecular origins of rapid and continuous morphological evolution. *Proc Natl Acad Sci USA* 2004;101:18058–18063. [PubMed: 15596718]
96. Hodgkinson A, Ladoukakis E, Eyre-Walker A. Cryptic variation in the human mutation rate. *PLoS Biol* 2009;7:e1000027. [PubMed: 19192947]
97. Greenman C, Stephens P, Smith R, Dalgliesh GL, Hunter C, et al. Patterns of somatic mutation in human cancer genomes. *Nature* 2007;446:153–158. [PubMed: 17344846]
98. Komarova NL, Wodarz D. Drug resistance in cancer: principles of emergence and prevention. *Proc Natl Acad Sci USA* 2005;102:9714–9719. [PubMed: 15980154]
99. Degtyareva NP, Chen L, Mieczkowski P, Petes TD, Doetsch PW. Chronic oxidative DNA damage due to DNA repair defects causes chromosomal instability in *Saccharomyces cerevisiae*. *Mol Cell Biol* 2008;28:5432–5445. [PubMed: 18591251]
100. Malpica JM, Fraile A, Moreno I, Obies CI, Drake JW, et al. The rate and character of spontaneous mutation in an RNA virus. *Genetics* 2002;162:1505–1511. [PubMed: 12524327]

101. Wang J, Gonzalez KD, Scaringe WA, Tsai K, Liu N, et al. Evidence for mutation showers. *Proc Natl Acad Sci USA* 2007;104:8403–8408. [PubMed: 17485671]
102. Drake JW, Bebenek A, Kissling GE, Peddada S. Clusters of mutations from transient hypermutability. *Proc Natl Acad Sci USA* 2005;102:12849–12854. [PubMed: 16118275]
103. Galhardo RS, Hastings PJ, Rosenberg SM. Mutation as a stress response and the regulation of evolvability. *Crit Rev Biochem Mol Biol* 2007;42:399–435. [PubMed: 17917874]
104. Wright BE, Longacre A, Reimers JM. Hypermutation in derepressed operons of *Escherichia coli* K12. *Proc Natl Acad Sci USA* 1999;96:5089–5094. [PubMed: 10220423]
105. Ponder RG, Fonville NC, Rosenberg SM. A switch from high-fidelity to error-prone DNA double-strand break repair underlies stress-induced mutation. *Mol Cell* 2005;19:791–804. [PubMed: 16168374]
106. Torkelson J, Harris RS, Lombardo MJ, Nagendran J, Thulin C, Rosenberg SM. Genome-wide hypermutation in a subpopulation of stationary-phase cells underlies recombination-dependent adaptive mutation. *EMBO J* 1997;16:3303–3311. [PubMed: 9214645]
107. Heidenreich E, Novotny R, Kneidinger B, Holzmann V, Wintersberger U. Non-homologous end joining as an important mutagenic process in cell cycle-arrested cells. *EMBO J* 2003;22:2274–2283. [PubMed: 12727893]
108. Strathern JN, Shafer BK, McGill CB. DNA synthesis errors associated with double-strand-break repair. *Genetics* 1995;140:965–972. [PubMed: 7672595]
109. Ohnishi O. Spontaneous and ethyl methanesulfonate-induced mutations controlling viability in *Drosophila melanogaster*. II. Homozygous effect of polygenic mutations. *Genetics* 1977;87:529–545. [PubMed: 200526]
110. Keightley PD, Caballero A. Genomic mutation rates for lifetime reproductive output and lifespan in *Caenorhabditis elegans*. *Proc Natl Acad Sci USA* 1997;94:3823–3827. [PubMed: 9108062]
111. Kondrashov AS, Crow JF. A molecular approach to estimating the human deleterious mutation rate. *Hum Mutat* 1993;2:229–234. [PubMed: 8364591]
112. Nachman MW, Crowell SL. Estimate of the mutation rate per nucleotide in humans. *Genetics* 2000;156:297–304. [PubMed: 10978293]
113. Eyre-Walker A, Keightley PD. High genomic deleterious mutation rates in hominids. *Nature* 1999;397:344–347. [PubMed: 9950425]
114. Gaffney DJ, Keightley PD. Genomic selective constraints in murid noncoding DNA. *PLoS Genet* 2006;2:e204. [PubMed: 17166057]
115. Russell LB, Russell WL. Spontaneous mutations recovered as mosaics in the mouse specific-locus test. *Proc Natl Acad Sci USA* 1996;93:13072–13077. [PubMed: 8917546]

Table 1

Pre-genomic computational estimates of the deleterious mutation rate in model organisms

Organism	U_D	Method	Reference
<i>D. melanogaster</i>	0.14	BM	42
	0.17	BM	43
	0.06	BM	109
	0.02	BM	48
	0.01	ML	48
	0.01	ML	46
<i>C. elegans</i>	0.0026 (haploid)	ML	110
	0.05	BM	58
	0.005 (haploid)	ML	49
<i>S. cerevisiae</i>	6.3×10^{-5} (haploid)	ML	14
	9.5×10^{-5}	ML	50

All estimates for U_D , the deleterious mutation rate per genome, are for diploid genomes, unless otherwise noted. BM refers to Bateman-Mukai, and ML to Maximum Likelihood. U_D is shown as the number of deleterious mutations/genome/division (per generation for multicellular organisms).

Table 2Pre-genomic estimates of mutation rates (μ_n and U_D) in mammals

Organism	μ_n	U_D	Method	Reference
Human	10^{-8}		Indirect, pseudogene substitution rates	111
Human	2.5×10^{-8}	3	Indirect, pseudogene substitution rates	112
Human		1.6	Indirect, Synonymous/non-synonymous substitution rates	113
Human	1.8×10^{-8}		Direct, disease mutations	83
Murid	2.96×10^{-9} 3.12×10^{-9}	0.91	Indirect-substitution rates in coding, non-coding DNA	114
Mouse	6.6×10^{-6} per locus		Direct, visible markers	115

μ_n , mutation rate per base, U_D , deleterious mutation rate per diploid genome. Units for μ_n are presented as the number of mutations/base/generation; for U_D , the number of deleterious mutations/genome/generation.

Table 3

A comparison of pre-genomic and post-genomic estimates of mutation rate parameters (μ_n , U_D , U_T) in model species.

Organism	Pre-genomic estimate		Post-genomic estimate	
	μ_n	U_D	μ_n	U_T
<i>S. cerevisiae</i>	2.2×10^{-10}	10^{-5} - 10^{-3}	3.3×10^{-10}	0.32
<i>C. elegans</i>	2.3×10^{-10}	0.005	2.1×10^{-8}	2.1
<i>D. melanogaster</i>	3.4×10^{-10}	0.02-0.1	8.4×10^{-9}	1.2

Estimates for U_D (deleterious mutation rate per genome) and U_T (total mutation rate per genome) are expressed per diploid genome with the following exceptions: U_D and U_T in *C. elegans* and U_T in *S. cerevisiae* are expressed per haploid genome. For *S. cerevisiae*, pre-genomic estimates of μ_n (mutation rate per base) were obtained from Drake et al. (34) and post-genomic from Lynch et al. (65). Pre-genomic estimates of U_D were obtained from Zeyl and DeVisser (50) and Wloch et al. (44) and post-genomic estimate of U_T from Lynch et al. (65). For *C. elegans*, pre-genomic estimates of μ_n were obtained from Drake et al. (34) and post-genomic from Denver et al. (67). Pre-genomic estimates of U_D and post-genomic estimates of U_D and U_T from Denver et al. (67). For *D. melanogaster*, pre-genomic estimates of μ_n were obtained from Drake et al. (34) and post-genomic from Haag-Liautaud et al. (70). For U_D , pre-genomic were obtained from Fry et al. (48) and Mukai et al. (43) and post-genomic from Haag-Liautaud et al. (70). Units for μ_n are presented as the number of mutations/base/division; U_D , the number of deleterious mutations/genome/division; U_T the number of mutations/genome/division. All units are per generation (multicellular organisms).