Stability across the Whole Nuclear Genome in the Presence and Absence of DNA Mismatch Repair

Scott Alexander Lujan and Thomas A. Kunkel *

Abstract: We describe the contribution of DNA mismatch repair (MMR) to the stability of the eukaryotic nuclear genome as determined by whole-genome sequencing. To date, wild-type nuclear genome mutation rates are known for over 40 eukaryotic species, while measurements in mismatch repair-defective organisms are fewer in number and are concentrated on Saccharomyces cerevisiae and human tumors. Well-studied organisms include Drosophila melanogaster and Mus musculus, while less genetically tractable species include great apes and long-lived trees. A variety of techniques have been developed to gather mutation rates, either per generation or per cell division. Generational rates are described through whole-organism mutation accumulation experiments and through offspring-parent sequencing, or they have been identified by descent. Rates per somatic cell division have been estimated from cell line mutation accumulation experiments, from systemic variant allele frequencies, and from widely spaced samples with known cell divisions per unit of tissue growth. The latter methods are also used to estimate generational mutation rates for large organisms that lack dedicated germlines, such as trees and hyphal fungi. Mechanistic studies involving genetic manipulation of MMR genes prior to mutation rate determination are thus far confined to yeast, Arabidopsis thaliana, Caenorhabditis elegans, and one chicken cell line. A great deal of work in wild-type organisms has begun to establish a sound baseline, but far more work is needed to uncover the variety of MMR across eukaryotes. Nonetheless, the few MMR studies reported to date indicate that MMR contributes 100-fold or more to genome stability, and they have uncovered insights that would have been impossible to obtain using reporter gene assays.

Keywords: mutation accumulation; mutagenesis; mutation rate; DNA replication; genome stability; DNA mismatch repair; whole-genome sequencing; DNA repair; Eukarya

1. Overview

We considered 123 independent nuclear genome mutation rate measurements, gathered from over 90 studies performed in either wild-type or MMR-deficient strains of 48 eukaryotic species. We confined the analysis to whole-genome studies that explicitly report rates or, rarely, to studies from which rates may be easily calculated. We present mean mutation rates for similar systems that are either MMR-proficient (Table 1) or MMR-deficient (Table 2). Studies are listed in Table 3. Granular details and notes on each study may be found in Table S1. Where available, rates per generation and per cell division are both presented. We classify each estimate as using either germline or somatic cells, although there is no distinction for most unicellular eukaryotes, and some organisms, e.g., hyphal fungi and many plants, lack dedicated germlines. We highlight trends and extremes, and then comment on how whole-genome rates elucidate mechanisms of MMR.

2. A Brief History

Mutation accumulation (MA) experiments are a venerable approach for estimating spontaneous mutation rates (reviewed in [1]). Theorized in the 1920s and first implemented...
in the 1960s, MA experiments use replicate lines derived from an ancestral population that can evolve independently. The population is subjected to periodic artificial bottlenecks to fix mutations regardless of their effects on selective fitness. Originally, mutations were selected via phenotypic changes due to mutations in reporter loci. Sequencing a reporter locus in the final population allowed for mutation detection and counting, resulting in mutation spectra and mutation rate estimates for that locus. However, no reporter locus can simulate all possible contexts, transcription states, chromatin states, replication times, or proximity to various genomic features. The advent of whole-genome sequencing bypassed these restrictions by making the entire genome the reporter. Mutations may be called by comparing the parental sequence to the sequences of progeny populations.

The first successful whole-genome MA experiments were published in 2008, first in \textit{Saccharomyces cerevisiae} (baker’s yeast) \cite{2} and then in the bacterium \textit{Salmonella typhimurium} \cite{3}. Wild-type, whole-genome mutation rates were previously compared across kingdoms \cite{1}, and therefore here we confine the discussion to the eukaryotic nuclear genome. Lynch et al. found 33 mutations in four wild-type haploid \textit{Saccharomyces cerevisiae} lines that had been each propagated for approximately 4800 cell divisions \cite{2}. They estimated the whole-genome mutation rate at 0.33 Gbp$^{-1}$ division$^{-1}$. The race was on to find rates in as many diverse species as possible. By the end of 2010, the list included model organisms such as \textit{Drosophila melanogaster} (fruit or vinegar fly; 0.1 Gbp$^{-1}$ division$^{-1}$; \cite{4}), \textit{Caenorhabditis elegans} (a roundworm; 0.32 Gbp$^{-1}$ division$^{-1}$; \cite{5}), and \textit{Arabidopsis thaliana} (thale cress; 0.22 Gbp$^{-1}$ division$^{-1}$; \cite{6}).

The first whole-genome mutation rate estimates for genetically manipulated eukaryotes were also published in 2010. Zanders et al. performed the first estimates for DNA mismatch repair (MMR)-deficient organisms, a baker’s yeast strain with a temperature-sensitive variant of the MMR gene \textit{MLH1} (\textit{mlh1-7ts}; 3.7 Gbp$^{-1}$ division$^{-1}$ \cite{7}). Comparison with the wild-type rates of Lynch et al. implied MMR repair of over 90% of replication errors (MMR$^{-}$/MMR$^{+}$ = correction efficiency; 3.7/0.33 = 11.2). This comported well with prior reporter locus estimates in \cite{8}. Larrea et al. then used MMR-deficient (\textit{msh2}Δ) baker’s yeast with a variant of DNA polymerase (Pol) $\delta$ (\textit{pol3-L612M}; \cite{9}). The known mutation bias of \textit{pol3-L612M}, found in previous experiments in vitro \cite{10}, showed the bulk of Pol $\delta$ synthesis to occur on the nascent lagging strand. This extended results from previous mutation accumulations in reporter genes \cite{11}. Thus, whole-genome mutation collections were shown to be useful for revealing cellular mechanisms.

A study in 2010 also reported the first whole-genome mutation rate estimate for humans (11 Gbp$^{-1}$ generation$^{-1}$ \cite{12}). This estimate could not come from whole-genome MA experiments. Baker’s yeast can reproduce through budding, a form of binary fission. Baker’s yeast, roundworms, and thale cress can reproduce through selfing. Vinegar flies neither bud nor self-fertilize, but they can be inbred in order to fix mutations. None of these options are available for humans, and therefore Roach et al. sequenced the genomes of a nuclear family and inferred mutations by comparing children to parents. Such parent-offspring sets are now a standard method for finding whole-genome mutation rates in outcrossing species, including wild populations.

The following decade saw scores of whole-genome rate measurements, plus a host of mutation frequencies and spectra from tumor genomes (e.g., \cite{13}). Note that tumor studies often use similar terminology and technology to the experiments listed here, but, lacking cell division counts, they may report mutation frequencies rather than rates. This restriction has been circumvented somewhat by measurements in cancer cell lines (e.g., chicken DT40 tumor line \cite{14,15} and human cell line RPE1 \cite{16}, and by raising organoids from tumor samples \cite{17}). The latter is also useful for estimating mutation rates in normal somatic tissues \cite{17,18}.

Some progress has been made in calculating mutation rates given incomplete knowledge of ancestral states or generation counts. Where complete pedigrees are unknown or ancestral samples are unavailable but little selective pressure is expected, mutations may be inferred by deriving the genotype of the last common ancestor. This technique, known
as identity by descent, limits analysis to certain highly conserved segments [19]. Likewise, the number of cell divisions in the stem line for a particular tissue may be unknown. Given a representative sample of the whole tissue, the mutation rate in the first few rounds of replication may be inferred from variant allele frequencies (VAF). VAF methods are easiest with blood and require sophisticated modelling to account for unequal contributions of early embryonic cells [20].

3. Nuclear Mutation Rates in MMR-Proficient Germ Cells

Mutation rates are, by necessity, conditional. There is little ab initio reason to expect mutation rates to remain constant across differing species, environmental conditions, stressors, exposures, tissues, and germline versus somatic status. For instance, mutation rates may vary with organismal, tissue, or parental ages. Human mutation counts increase with parental age, particularly paternal age, which affects the mutation rate per generation [21]. Wherever necessary, Table S1 uses assumed average parental age, as defined by the authors of the study in question. Somatic rates are averaged across estimates, including across tissues [18] and growth conditions [22], where rates vary little. Rates are not combined if conditions are known to cause large rate differences, such as different ploidies [23] or homozygous versus hybrid or otherwise highly heterozygous individuals [24]. Mutation rates are also conditional across individual genomes, a subject we address below in our discussion of MMR.

Wild-type generational mutation rates range from 0.0076 Gbp⁻¹ generation⁻¹ in the ciliate *Tetrahymena thermophila* [25] to 3380 Gbp⁻¹ generation⁻¹ in the hyphal fungus *Neurospora crassa* (red bread mold; [26]). These extremes are largely explainable by how these organisms transmit their genetic code through generations. Ciliates such as *Tetrahymena* and *Paramecium tetraurelia* [27] keep dozens of working copies of their genome in transcriptionally active compartment called the macronucleus while protecting a germ copy in a protected micronucleus. In contrast, red bread mold has no separate germ line, undergoing an average of 300 asexual divisions per sexual generation [26]. The asexual rate is listed as “somatic” in Table 1, although this definition is debatable. However, for reasons that are not entirely clear, most mutations per sexual generation occur in the last few divisions, perhaps only during meiosis. Is this the case with meiosis in other organisms?

Table 1. Nuclear genome mutation rates from whole-genome experiments (MMR-proficient).

| ct. | Species                        | Supergroup | Lower Clade | Cellularity | Ploidy | Germ V. Soma | Mutation Rates | Lines | Mutations |
|-----|--------------------------------|------------|-------------|-------------|--------|--------------|----------------|-------|-----------|
| 1   | *Phaeodactylum tricornutum*    | TSAR Group | Stramenopiles | uni-        | 2n     | g            | 0.49           | 36    | 156       |
| 1   | *Paramecium tetraurelia*      | TSAR Group | Ciliophora  | uni-        | 2n     | g            | 0.030          | 7     | 29        |
| 1   | *Tetrahymena thermophila*     | TSAR Group | Ciliophora  | uni-        | 2n     | g            | 0.0076         | 8     | 5         |
| 1   | *Plasmodium falciparum*       | TSAR Group | Apicomplexa  | uni-        | 1n     | g            | 0.25           | 279   | 85        |
| 1   | *Batrachococcus prasinus*     | Archeplastida | Chlorophyta | uni-        | 1n     | g            | 0.44           | 37    | 32        |
| 3   | *Chlamydomonas reinhardtii*   | Archeplastida | Chlorophyta | uni-        | 1n     | g            | 0.18           | 91    | 6890      |
| 1   | *Micromonas pusilla*          | Archeplastida | Chlorophyta | uni-        | 1n     | g            | 0.98           | 36    | 85        |
| 1   | *Ostreococcus mediterraneus*  | Archeplastida | Chlorophyta | uni-        | 1n     | g            | 0.59           | 37    | 65        |
| 1   | *Ostreococcus tauri*          | Archeplastida | Chlorophyta | uni-        | 1n     | g            | 0.48           | 40    | 104       |
| 5   | *Arabidopsis thaliana*        | Archeplastida | Embryophyta | multi-      | 2n     | g            | 6.7            | 156   | 2324      |
| 1   | *Arabidopsis thaliana*        | Archeplastida | Embryophyta | multi-      | 2n     | (het.) g     | 27             | -     | 299       |
| 1   | *Eucalyptus melliodora*       | Archeplastida | Embryophyta | multi-      | 2n     | g            | 62             | -     | 1         |
| 1   | *Lemna minor*                 | Archeplastida | Embryophyta | multi-      | 2n     | g            | 0.087          | -     | 16        |
| 1   | *Oryza sativa*                | Archeplastida | Embryophyta | multi-      | 2n     | g            | 3.2            | -     | 5         |
| 1   | *Oryza sativa*                | Archeplastida | Embryophyta | multi-      | 2n     | (het.) g     | 11             | -     | 11        |
| 1   | *Pinus sitchensis*            | Archeplastida | Embryophyta | multi-      | 2n     | s            | 27             | -     | 20        |
| 1   | *Populus trichocarpa*         | Archeplastida | Embryophyta | multi-      | 2n     | g            | 2.0            | -     | 2         |

3 of 15
Another hyphal fungus, the fairy ring mushroom *Marasmius oreades*, has the lowest measured mutation rate per cell division at 0.0038 Gbp⁻¹ division⁻¹ [26]. This is even lower than in the ciliates, but without an obvious mechanistic explanation. Nonetheless, with over 19,000 divisions per generation, this still yields a relatively high mutation rate of 73 Gbp⁻¹ generation⁻¹. How is the low rate per division maintained and how is the high rate per generation tolerated? Is this situation common among hyphal fungi? The cell divisions per generation for the fairy ring mushroom in Table 1 are estimated from the ratio of rates, per-generation divided by per-division. This would be incorrect if it has a sexual rate dominated by mutations in later cell divisions, as in red bread mold. Perhaps clarity will emerge through testing more organisms with more diverse lifestyles and genetic architectures. The highest wild-type “germline” mutation rate per division is 0.98 Gbp⁻¹ division⁻¹, in the haploid unicellular alga *Micromonas pusilla* [28]. How does this organism deal with a rate per division over 250-times higher than in the fairy ring mushroom? This rate is in turn dwarfed by those in animal somatic cells.

### Table 1. Cont.

| Clade | Species | Supergroup | Lower Clade | Cellularity | Ploidy | Germ V. Soma | Mutation Rates | Lines | Mutations |
|-------|---------|------------|-------------|-------------|--------|--------------|---------------|-------|-----------|
|       |         |            |             |             |        |              | Gbp⁻¹ gen⁻¹ | Gbp⁻¹ div⁻¹ |           |
| 1     | *Prunus hybrid* | Archaeplastida | Embryophyta | multi- (-het.) | 2n | g | 14 | - | 30 | 171 |
| 1     | *Prunus persica* | Archaeplastida | Embryophyta | multi- | 2n | g | 8.6 | - | 32 | 114 |
| 1     | *Quercus robur* | Archaeplastida | Embryophyta | multi- | 2n | s | 47 | - | 1 | 17 |
| 1     | *Silene latifolia* | Archaeplastida | Embryophyta | multi- | 2n | g | 7.3 | - | 10 | 39 |
| 2     | *Spirodea polyrhiza* | Archaeplastida | Embryophyta | multi- | 2n | g | 0.082 | - | 47 | 46 |
| 1     | *Dictyostelium discoideum* | Amoebozoa | Mycetozoa | alternates | 1n | g | 0.029 | 0.029 | 3 | 1 |
| 1     | *Neospora crassa* | Opisthokonta | Ascomycota | multi- | 1n | g | 3400 | - | 268 | 10,493 |
| 1     | *Neospora crassa* | Opisthokonta | Ascomycota | multi- | 1n | s | - | 0.60 | 10 | 93 |
| 5     | *Saccharomycetes cerevisiae* | Opisthokonta | Ascomycota | uni- | 1n | g | 0.39 | 0.35 | 68 | 475 |
| 9     | *Saccharomycetes cerevisiae* | Opisthokonta | Ascomycota | uni- | 2n | g | 0.23 | 0.23 | 392 | 3194 |
| 3     | *Schizosaccharomyces pombe* | Opisthokonta | Ascomycota | uni- | 1n | g | 0.37 | 0.37 | 180 | 1308 |
|       | *Marasmius oreades* | Opisthokonta | Basidiomycota | multi- | 2n | s | 73 | 0.0038 | 40 | 111 |
| 1     | *Schizopyllum commune* | Opisthokonta | Basidiomycota | multi- | 2n | g | 20 | - | 17 | 9 |
| 1     | *Schizopyllum commune* | Opisthokonta | Basidiomycota | multi- | 2n | s | - | 0.02 | 24 | 300 |
| 4     | *Caenorhabditis elegans* | Opisthokonta | Nematoda | multi- | 2n | g | 3.1 | 0.57 | 57 | 3535 |
| 1     | *Caenorhabditis species* | Opisthokonta | Nematoda | multi- | 2n | g | 1.3 | 0.12 | 25 | 448 |
| 1     | *Pristionchus pacificus* | Opisthokonta | Nematoda | multi- | 2n | g | 2.0 | - | 22 | 802 |
| 1     | *Apis mellifera* | Opisthokonta | Arthropoda | multi- | 1n | g | 4.5 | - | 46 | 35 |
| 1     | *Bombus terrestris* | Opisthokonta | Arthropoda | multi- | 1n | g | 3.9 | - | 32 | 23 |
| 1     | *Chironomus riparius* | Opisthokonta | Arthropoda | multi- | 2n | g | 4.2 | - | 10 | 51 |
| 2     | *Daphnia pulex* | Opisthokonta | Arthropoda | multi- | 2n | g | 3.1 | - | 30 | 1210 |
| 6     | *Drosophila melanogaster* | Opisthokonta | Arthropoda | multi- | 2n | g | 5.1 | 0.13 | 175 | 3539 |
| 1     | *Helicocinus melipomene* | Opisthokonta | Arthropoda | multi- | 2n | g | 2.9 | 0.073 | 30 | 9 |
| 1     | *Aotus nancymae* | Opisthokonta | Chordata | multi- | 2n | g | 8.1 | - | 8 | 283 |
| 1     | *Canis lupus* | Opisthokonta | Chordata | multi- | 2n | g | 4.5 | - | 4 | 27 |
| 1     | *Chlorocebus aethiops* | Opisthokonta | Chordata | multi- | 2n | g | 9.4 | - | 3 | 8 |
| 1     | *Clupea harengus* | Opisthokonta | Chordata | multi- | 2n | g | 2.0 | - | 12 | 19 |
| 1     | *Ficedula albicollis* | Opisthokonta | Chordata | multi- | 2n | g | 4.6 | - | 7 | 55 |
| 2     | *Galus gallus domesticus* | Opisthokonta | Chordata | multi- | 2n | s | - | 0.91 | 6 | 384 |
| 1     | *Gorilla gorilla* | Opisthokonta | Chordata | multi- | 2n | g | 11 | - | 8 | 161 |
| 13    | *Homo sapiens* | Opisthokonta | Chordata | multi- | 2n | s | 12 | 0.17 | 3062 | 156,475 |
| 8     | *Homo sapiens* | Opisthokonta | Chordata | multi- | 2n | s | - | 8.6 | 388 | 86,157 |
| 1     | *Macaca mulatta* | Opisthokonta | Chordata | multi- | 2n | g | 5.8 | - | 14 | 307 |
| 3     | *Mus musculus* | Opisthokonta | Chordata | multi- | 2n | g | 5.1 | 0.11 | 50 | 1614 |
| 2     | *Mus musculus* | Opisthokonta | Chordata | multi- | 2n | s | - | 4.2 | 30 | 3697 |
| 3     | *Pan troglodytes* | Opisthokonta | Chordata | multi- | 2n | g | 13 | - | 7 | 283 |
| 1     | *Papa anubis* | Opisthokonta | Chordata | multi- | 2n | g | 6.2 | - | 12 | 475 |
| 1     | *Pongo abellii* | Opisthokonta | Chordata | multi- | 2n | g | 17 | - | 1 | 51 |

Rates are averaged (mean) over all experimental estimates (unweighted). Rates are rounded to two significant digits. Color code: “supergroup” and “lower clade” columns are colored to highlight related clades; green saturation increases linearly with experiment counts in column “ct.”; a gradient from blue to red was applied across “Gbp⁻¹ div⁻¹” columns of Tables 1 and 2, with blue indicating the lowest rates and red the highest. Abbreviations: ct. = number of independent estimates; g = germline; s = somatic cells; gen. = generation; div. = cell division; - = not determined.
Table 2. Nuclear genome mutation rates from whole-genome experiments (MMR-deficient).

| ct. | Species                     | Supergroup       | Lower Clade | Cellularity | Ploidy | Germ Rates | Mutation Rates | MMR Efficiency |
|-----|-----------------------------|------------------|-------------|-------------|--------|------------|----------------|----------------|
| 2   | Arabidopsis thaliana        | Archaeplastida   | Embryophyta | multi-      | 2n     | g          | 810            | 27             | 14             | 8902           | 120 ± 100 b   |
| 3   | Saccharomyces cerevisiae    | Opisthokonta     | Ascomycota  | uni-        | 1n     | g          | 31             | 31             | 6              | 1840           | 79 ± 89       |
| 4   | Saccharomyces cerevisiae    | Opisthokonta     | Ascomycota  | uni-        | 2n     | g          | 13             | 13             | 25             | 3684           | 57 ± 57       |
| 1   | Schizosaccharomyces pombe   | Opisthokonta     | Ascomycota  | uni-        | 1n     | g          | 19             | 19             | 5              | 2597           | 51 ± 51       |
| 2   | Caenorhabditis elegans      | Opisthokonta     | Nematoda    | multi-      | 2n     | g          | 72             | 9              | 9110           | -              | 130           |
| 1   | Gallus gallus domesticus    | Opisthokonta     | Chordata    | multi-      | 2n     | s          | 47             | 2              | 6531           | -              | 52            |

Rates are averaged (mean) over all experimental estimates (unweighted). Rates and correction efficiencies are rounded to two significant digits. Color code: “supergroup” and “lower clade” columns are colored to highlight related clades; green saturation increases linearly with experiment counts in column “ct.”; a gradient from blue to red was applied across “Gbp⁻¹ div⁻¹” columns of Tables 1 and 2, with blue indicating the lowest rates and red the highest. Notes: ± = efficiencies calculated from mutation rates per generation; b = efficiencies calculated from mutation rates per cell division. Abbreviations: ct. = number of independent estimates; g = germline; s = somatic cells; gen. = generation; div. = cell division; - = not determined.

4. Nuclear Mutation Rates Trends in MMR-Proficient Organisms

Three trends in nuclear mutation rates appear in the data. First, as previously stated for humans, mutation rates increase with parental age. Second, in plants, highly heterozygous lines have higher mutation rates than homozygous lines. Third, in animals, somatic mutation rates exceed germline rates. Mutation counts increase with parental age in many species. In humans, paternal age has a particularly strong effect on offspring mutation counts, commensurate with continuing cell divisions in the male germline (reviewed [21]). However, maternal age is also a factor, which is more difficult to explain. Although outside the scope of this review, whole mitochondrial genome sequencing studies also show age-dependent increases in both point mutations [29] and large deletions [30]. The situation is even more extreme in large, long-lived hyphal fungi [31,32] and trees [33–37]. Because they grow outward linearly, lack a dedicated germline, and tend to fruit near their maximum extent, each consecutive fruiting results in more offspring mutations. Will whole-genome mutation rate studies ever find age-related increases in shorter lived or unicellular eukaryotes?

Only two whole-genome mutation rate studies were found that compared homozygous lines with highly heterozygous lines. Both were in plants, encompassing three species. Yang et al. found 3.6-fold higher rates in heterozygous thale cress and 3.4-fold higher rates in heterozygous rice (Oryza sativa) [24]. Likewise, Xie et al. found a more modest 1.6-fold increase in a hybrid peach tree (Prunus davidiana × P. persica) versus a weakly heterozygous peach tree (P. persica) [33]. Both studies concluded that highly heterozygous lines have higher mutation rates than homozygous lines. The idea that heterozygosity is tied to plant mutation rates has been discussed [33] and is supported by previous reporter locus assays (e.g., [38]). Will the results of these few experiments be recapitulated in other plants or in other eukaryotic clades?

One study measured comparable somatic and germline mutation rates per cell division in two organisms: humans and house mice [39]. The highest measured wild-type mutation rate per cell division belongs to house mouse fibroblasts at 8.1 Gbp⁻¹ division⁻¹, roughly 70-fold higher than in the germline. Likewise, human fibroblasts rates were 2.7 Gbp⁻¹ division⁻¹, roughly 80-fold higher than in the germline. Is this a general feature of multicellular organisms other than hyphal fungi, or is it limited to just animals or to mammals only? How are lower mutation rates maintained in the germline? Does MMR play a part or is it only a matter of protection from insult exposure? More information is needed in other animals and multicellular fungi, plants, and stramenopiles (e.g., kelp).
5. Nuclear Mutation Rates in MMR-Deficient Cells

Table 2 lists overall mutation rates in MMR-deficient cells. These come from baker’s yeast, fission yeast (Schizosaccharomyces pombe), thale cress, roundworms (C. elegans), and an immortalized chicken cell line (Gallus gallus domesticus DT40). The mean rates have non-overlapping ranges: MMR-proficient with $0.23-0.91 \text{ Gbp}^{-1} \text{ division}^{-1}$, and MMR-deficient with $13-72 \text{ Gbp}^{-1} \text{ division}^{-1}$. Correction efficiencies are remarkably consistent, ranging from 50- to 130-fold, despite disparate species, ploidies, cellular lineages (i.e., somatic versus germline), and methods for ablating MMR (see Table S1 for genotypes and notes). The correction efficiencies are bimodally distributed, with fission yeast, chicken cells, and diploid baker’s yeast clustered at 51–57× and thale cress, roundworms, and haploid baker’s yeast efficiencies from 100–130×. Is this a coincidental artefact of the few systems studied? Regardless, these whole-genome rate measurements have clearly shown that MMR is highly efficient, repairing at least 98% of replication errors. Indeed, this is probably an underestimate (see Section 8).

6. Genome-Wide Mutations and the Mechanisms of MMR

For long-lived organisms, reporter locus experiments are an inefficient way to collect mutations. For shorter-lived organisms, given the expense of whole-genome sequencing and the time required for mutation accumulation experiments (ideally hundreds of generations), why not use reporter loci? First, reporter loci do not adequately model the sequence complexity of the genome (as discussed above). Second, reporter loci cannot replicate the diversity of selective pressures across the genome. Both factors are essential for the study of MMR.

For example, the baker’s yeast genome is GC-poor, but certain AT-rich features are concentrated outside of regions that are translated into proteins (like most reporter loci). AT homopolymer tracts, particularly long tracts, are concentrated in untranslated regions (UTRs) that flank most genes [40]. This leads reporter locus assays to underestimate the rates of deletions in long homopolymers and the rates of multi-base insertions and deletions (indels) [41]. Whole-genome mutation accumulations show that these regions become indel hotspots upon removal of MMR [40], with rates and indel sizes increasing with tract length [40,42]. In fact, the shape of the curve of rate versus tract length is diagnostic of the degree to which mismatch extension is favored over proofreading. Extension could be driven by a proofreading defect [43] or by alteration of nucleotide concentrations [44].

Unlike in yeast, AT homopolymers in humans are concentrated in genes, where cancer genomes indicate strong transcriptional strand asymmetry for indels [45,46]. Studies of tumors with Pol δ proofreading defects suggest that MMR repairs about threefold more mismatches produced during lagging strand replication compared with leading [45]. Massive studies of cancer genomes have allowed the construction of mutation spectrum signatures that are diagnostic of such processes as MMR [47,48]. Tumors with mutations in DNA polymerase (Pol) ε have mutation spectra that resemble spectra from cell lines with defects in both Pol ε and MMR [49]. This suggests that MMR is somehow suppressed in those tumors. Conversely, there appears to be a mutational hotspot in the gene that encodes the catalytic subunit of Pol ε in MMR-deficient mouse lymphomas [50]. Spectra in MMR-deficient chicken cells allowed Németh et al. to collapse six MMR-associated COSMIC signatures into two [15]. They found no correlation between these signatures and the identity of the defective MMR genes in the tumors (i.e., MSH2, MSH6, or MLH1). This suggests that either modulation of transcription or translation or some form of inhibition are to blame for the MMR defects in these tumors. This is a profound revelation, given that MMR-deficient cancers generate mutant neoantigens that make them sensitive to immune checkpoint blockade [51]. Thus, whole-genome mutation rate experiments may affect cancer diagnosis and treatment.

Whole-genome experiments have revealed that MMR preferentially protects many genome features. In baker’s yeast, it protects UTRs and inter-nucleosome linkers from indels, translated gene bodies from point mutations, and sequence-encoded nucleosome
positions from substitutions [40]. Much of this is recapitulated in thale cress [52], and in humans, MMR selectively protects exons relative to introns [53]. In fission yeast, MMR selectively protects euchromatin [54]. Baker’s yeast strains have slightly higher rates in early as opposed to late replicating regions, with some indication of higher MMR efficiency early in replication [40]. Likewise, variable human MMR is thought to cause elevated mutation rates in late replicating heterochromatin compared to early replicating euchromatin [55]. Are MMR proteins depleted or in some other way impaired later in replication? In humans, some MMR proteins are differentially expressed across the cell cycle [56]. In mice, histone modifications can target MMR to transcriptionally active regions [57], both locally and globally [58]. The extent of targeting elsewhere and in other organisms is unknown. Unfortunately, those these trends point in the same direction, only a few of these studies report rates [15,40,52,54], making it difficult to compare effects across organisms in a quantitative manner.

Why does MMR appear to selectively protect some features over others? Perhaps the extent of MMR targeting, as in mice, is underappreciated. Alternatively, MMR may operate at a similar rate across each genome, but some contexts are simply more mutable. This would be expected if natural selection effectively erases mutations missed by MMR. Over evolutionary timescales, mutable sequences would disappear in regions under little selection. Depletion of MMR would then reveal the fingerprints of past selection (discussed in [40]).

7. Summary

Herein, we have gathered known whole-genome mutation rates, encompassing 90 studies (Table 3). We hope that future researchers will expand the list and use the information to uncover new insights into the patterns of mutagenesis across eukaryotes and beyond. We have also outlined some advances in the understanding of mutagenesis since the advent of whole-genome experiments. These advances reveal variation in eukaryotic DNA mismatch repair mechanisms that were invisible to most reporter locus assays. Further progress requires more breadth in the organisms, tissues, and conditions. In particular, new strains are required to uncover the interplay between mismatch repair and other nuclear systems, such as nucleotide pool maintenance, exonucleolytic proofreading, and ribonucleotide excision repair.

Table 3. List of whole-genome mutation rate experiments.

| First Author | Year | Reference | Species | MMR Genotype |
|--------------|------|-----------|---------|--------------|
| Lynch        | 2008 | [2]       | Saccharomyces cerevisiae | WT |
| Keightley    | 2009 | [4]       | Drosophila melanogaster | WT |
| Denver       | 2009 | [5]       | Caenorhabditis elegans | WT |
| Ossowskis    | 2010 | [6]       | Arabidopsis thaliana | WT |
| Roach        | 2010 | [12]      | Homo sapiens | WT |
| Zanders      | 2010 | [7]       | Saccharomyces cerevisiae | mlh1-7ts |
| Nishant      | 2010 | [59]      | Saccharomyces cerevisiae | WT |
| Conrad       | 2011 | [60]      | Homo sapiens | WT |
| Denver       | 2012 | [61]      | Caenorhabditis species | WT |
| Ma           | 2012 | [62]      | Saccharomyces cerevisiae | mlh1-7ts |
| Kong         | 2012 | [63]      | Homo sapiens | WT |
| Ness         | 2012 | [64]      | Chlamydomonas reinhardtii | WT |
| Saxer        | 2012 | [65]      | Dictyostelium discoideum | WT |
| Sung         | 2012 | [27]      | Chlamydomonas reinhardtii | WT |
| Sung         | 2012 | [27]      | Paramecium tetraurelia | WT |
| Michaelson   | 2012 | [66]      | Homo sapiens | WT |
| Schrider     | 2013 | [67]      | Drosophila melanogaster | WT |
| Lang         | 2013 | [68]      | Saccharomyces cerevisiae | WT |
| Lang         | 2013 | [68]      | Saccharomyces cerevisiae | msh2A |
| Li           | 2014 | [69]      | Homo sapiens | WT |
| First Author | Year | Reference | Species | MMR Genotype |
|--------------|------|-----------|---------|--------------|
| Keightley    | 2014 | [70]      | *Drosophila melanogaster* | WT |
| Stirling     | 2014 | [71]      | *Saccharomyces cerevisiae* | WT |
| Weller       | 2014 | [72]      | *Pristionchus pacificus*  | WT |
| Serero       | 2014 | [73]      | *Saccharomyces cerevisiae* | WT |
| Serero       | 2014 | [73]      | *Saccharomyces cerevisiae* | msh2Δ |
| Zhu          | 2014 | [74]      | *Saccharomyces cerevisiae* | WT |
| Venn         | 2014 | [75]      | *Pan troglodytes*         | WT |
| Meier        | 2014 | [76]      | *Caenorhabditis elegans*  | WT |
| Behjati      | 2014 | [17]      | *Mus musculus*            | WT |
| Lujan        | 2014 | [40]      | *Saccharomyces cerevisiae* | WT |
| Lujan        | 2014 | [40]      | *Saccharomyces cerevisiae* | msh2Δ |
| Jiang        | 2014 | [22]      | *Arabidopsis thaliana*    | WT |
| Keightley    | 2015 | [77]      | *Heliconius melpomene*    | WT |
| Francioli    | 2015 | [78]      | *Homo sapiens*            | WT |
| Uchimura     | 2015 | [79]      | *Mus musculus*            | WT |
| Baranová     | 2015 | [80]      | *Schizophyllum commune*   | WT |
| Yang         | 2015 | [24]      | *Apis mellifera*          | WT |
| Yang         | 2015 | [24]      | *Arabidopsis thaliana*    | WT |
| Yang         | 2015 | [24]      | *Oryza sativa*            | WT |
| Ness         | 2015 | [81]      | *Chlamydomonas reinhardtii* | WT |
| Farlow       | 2015 | [82]      | *Schizosaccharomyces pombe* | WT |
| Keith        | 2016 | [83]      | *Daphnia pulex*           | WT |
| Rahbari      | 2015 | [84]      | *Homo sapiens*            | WT |
| Haye         | 2015 | [85]      | *Saccharomyces cerevisiae* | msh6Δ |
| Behringer    | 2016 | [86]      | *Schizosaccharomyces pombe* | WT |
| Sharp        | 2016 | [87]      | *Drosophila melanogaster*  | WT |
| Huang        | 2016 | [88]      | *Drosophila melanogaster*  | WT |
| Sun          | 2016 | [54]      | *Schizosaccharomyces pombe* | WT |
| Sun          | 2016 | [54]      | *Schizosaccharomyces pombe* | msh6Δ |
| Smeds        | 2016 | [89]      | *Ficedula albicollis*     | WT |
| Long         | 2016 | [25]      | *Tetrahymena thermophila* | WT |
| Zámborszky  | 2017 | [14]      | *Gallus gallus domesticus* | WT |
| Watson       | 2016 | [90]      | *Arabidopsis thaliana*    | MSH2−/− |
| Flynn        | 2017 | [91]      | *Daphnia pulex*           | WT |
| Xie          | 2017 | [33]      | *Prunus persica*          | WT |
| Xie          | 2017 | [33]      | *Prunus hybrid*           | WT |
| Besenbacher  | 2016 | [92]      | *Homo sapiens*            | WT |
| Hamilton     | 2017 | [93]      | *Plasmodium falciparum*   | WT |
| Liu          | 2017 | [94]      | *Bombus terrestris*       | WT |
| Ju           | 2017 | [20]      | *Homo sapiens*            | WT |
| Krascovec    | 2017 | [28]      | *Bathycoccus prasinos*    | WT |
| Krascovec    | 2017 | [28]      | *Micromonas pusilla*      | WT |
| Krascovec    | 2017 | [28]      | *Ostreococcus mediterraneus* | WT |
| Krascovec    | 2017 | [28]      | *Ostreococcus tauri*      | WT |
| Milholland   | 2017 | [39]      | *Homo sapiens*            | WT |
| Milholland   | 2017 | [39]      | *Mus musculus*            | WT |
| Feng         | 2017 | [95]      | *Clupea harengus*         | WT |
| Maretty      | 2017 | [96]      | *Homo sapiens*            | WT |
| Dutta        | 2017 | [97]      | *Saccharomyces cerevisiae* | WT |
| Jónsson      | 2017 | [98]      | *Homo sapiens*            | WT |
| Pfeifer      | 2017 | [99]      | *Chlorocebus aethiops*    | WT |
| Assaf        | 2017 | [100]     | *Drosophila melanogaster*  | WT |
| Tatsumoto    | 2017 | [101]     | *Pan troglodytes*         | WT |
| Schmid-Siegert| 2017 | [34]      | *Quercus robur*           | WT |
| Belfield     | 2018 | [52]      | *Arabidopsis thaliana*    | Atmsh2-1 |
| Meier        | 2018 | [102]     | *Caenorhabditis elegans*  | WT |
Table 3. Cont.

| First Author | Year | Reference | Species | MMR Genotype |
|--------------|------|-----------|---------|--------------|
| Meier        | 2018 | [102]     | Caenorhabditis elegans | mlh-1        |
| Meier        | 2018 | [102]     | Caenorhabditis elegans | pms-2        |
| Sharp        | 2018 | [23]      | Saccharomyces cerevisiae | WT           |
| Krasovec     | 2018 | [103]     | Silene latifolia | WT           |
| Thomas       | 2018 | [104]     | Aotus nancymae | WT           |
| Oppold       | 2017 | [105]     | Chironomus riparius | WT           |
| Brody        | 2018 | [16]      | Homo sapiens | WT           |
| Weng         | 2019 | [106]     | Arabidopsis thaliana | WT           |
| Besenbacher  | 2019 | [107]     | Pan troglodytes | WT           |
| Besenbacher  | 2019 | [107]     | Gorilla gorilla | WT           |
| Besenbacher  | 2019 | [107]     | Pongo abelii | WT           |
| Xu           | 2019 | [108]     | Spiraleda polyrhiza | WT           |
| Konrad       | 2019 | [109]     | Caenorhabditis elegans | WT           |
| Krasovec     | 2019 | [110]     | Phaeodactylum tricornutum | WT           |
| Koch         | 2019 | [111]     | Canis lupus | WT           |
| Williams     | 2019 | [112]     | Saccharomyces cerevisiae | WT           |
| Hanlon       | 2019 | [35]      | Picea sitchensis | WT           |
| Hiltunen     | 2019 | [31]      | Marasmius oreades | WT           |
| Lindsay      | 2019 | [113]     | Mus musculus | WT           |
| Tian         | 2019 | [19]      | Homo sapiens | WT           |
| Németh       | 2020 | [15]      | Gallus gallus domesticus | WT           |
| Németh       | 2020 | [15]      | Gallus gallus domesticus | MSH2−/−   |
| Orr          | 2020 | [36]      | Eucalyptus melliodora | WT           |
| Bezmenova    | 2020 | [32]      | Schizophyllum commune | WT           |
| Wang         | 2020 | [26]      | Macaca mulatta | WT           |
| Wang         | 2020 | [26]      | Neurospora crassa | WT           |
| Wu           | 2020 | [114]     | Papio anubis | WT           |
| Wu           | 2020 | [114]     | Homo sapiens | WT           |
| Sandler      | 2020 | [115]     | Spiraleda polyrhiza | WT           |
| Sandler      | 2020 | [116]     | Lemma minor | WT           |
| Hofmeister   | 2020 | [37]      | Populus trichocarpa | WT           |
| Sui          | 2020 | [117]     | Saccharomyces cerevisiae | WT          |
| Zhou         | 2021 | in review | Saccharomyces cerevisiae | msh6Δ      |

Experiments are arranged by publication date. A study with multiple measurements in the same species with the same MMR genotype is listed only once. More details about each measurement are available in Table S1. Abbreviations: MMR = DNA mismatch repair.

8. More Future Questions

In addition to questions throughout this review, others arise due to the following. MMR efficiency calculations presented here assume that all mutations are due to replication and are subject to mismatch repair. The veracity of these assumptions is an outstanding question. For instance, most spontaneous mutations in wild-type yeast could be due to mutagenic repair of spontaneous lesions [118], which may not be amenable to MMR. Indeed, 40–85% of mutations in the wild-type baker’s yeast CAN1 reporter are attributable to errors made by DNA polymerase ζ [119–122]. Is this true across the genome, in other organisms, other conditions, or in various tissues? How much of the remaining wild-type mutation rate is due to other assumption-breaking processes? Is MMR dependent on other systems, such that a mutation that affects MMR also alters, say, polymerase proofreading or ribonucleotide excision repair, thus causing additional complicating mutagenesis? Until such questions are answered, all MMR efficiency calculations are likely to be minimum estimates and should be treated as provisional.
Supplementary Materials: The following are available online at https://www.mdpi.com/article/10.3390/cells10051224/s1, Table S1: Individual nuclear genome mutation rates from whole-genome experiments. Abbreviations: concat = concatenation of select columns; PMID = PubMed identification number; g = germline; s = somatic cells; w = wild type; m = MMR-deficient. Genome reference sizes represent NCBI Genome median assembly lengths, where available. Otherwise, they are median reference lengths from cited studies.

Funding: This study was supported by Project Z01 ES065070 to T.A.K from the Division of Intramural Research of the NIH, NIEHS.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Katju, V.; Bergthorsson, U. Old Trade, New Tricks: Insights into the Spontaneous Mutation Process from the Partnering of Classical Mutation Accumulation Experiments with High-Throughput Genomic Approaches. Genome Biol. Evol. 2019, 11, 136–165. [CrossRef] [PubMed]
2. Lynch, M.; Sung, W.; Morris, K.; Coffey, N.; Landry, C.R.; Dopman, E.B.; Dickinson, W.J.; Okamoto, K.; Kulkarni, S.; Hartl, D.L.; et al. A genome-wide view of the spectrum of spontaneous mutations in yeast. Proc. Natl. Acad. Sci. USA 2008, 105, 9272–9277. [CrossRef]
3. Lind, P.A.; Andersson, D.I. Whole-Genome mutational biases in bacteria. Proc. Natl. Acad. Sci. USA 2008, 105, 17878–17883. [CrossRef] [PubMed]
4. Keightley, P.D.; Trivedi, U.; Thomson, M.; Oliver, F.; Kumar, S.; Blaxter, M.L. Analysis of the genome sequences of three Drosophila melanogaster spontaneous mutation accumulation lines. Genome Res. 2009, 19, 1195–1201. [CrossRef] [PubMed]
5. Denver, D.R.; Dolan, P.C.; Wilhelm, I.J.; Sung, W.; Lucas-Lledo, J.L; Howe, D.K.; Lewis, S.C.; Okamoto, K.; Thomas, W.K.; Lynch, M.; et al. A genome-wide view of Caenorhabditis elegans base-substitution mutation processes. Proc. Natl. Acad. Sci. USA 2009, 106, 16310–16314. [CrossRef] [PubMed]
6. Ossowski, S.; Schneeberger, K.; Lucas-Lledo, J.L; Warthmann, N.; Clark, R.M.; Shaw, R.G.; Weigel, D.; Lynch, M. The Rate and Molecular Spectrum of Spontaneous Mutations in Arabidopsis italiana. Science 2010, 327, 92–94. [CrossRef]
7. Zanders, S.; Ma, X.; RoyChoudhury, A.; Hernandez, R.D.; Demogines, A.; Barker, B.; Gu, Z.; Bustamante, C.D.; Alani, E. Detection of Heterozygous Mutations in the Genome of Mismatch Repair Defective Diploid Yeast Using a Bayesian Approach. Genetics 2010, 186, 493–503. [CrossRef]
8. Kunkel, T.A.; Erie, D.A. Eukaryotic Mismatch Repair in Relation to DNA Replication. Annu. Rev. Genet. 2015, 49, 291–313. [CrossRef]
9. Larrea, A.A.; Lujan, S.A.; McElhinny, S.A.N.; Mieczkowski, P.A.; Resnick, M.A.; Gordenin, D.A.; Kunkel, T.A. Genome-wide model for the normal eukaryotic DNA replication fork. Proc. Natl. Acad. Sci. USA 2010, 107, 17674–17679. [CrossRef]
10. McElhinny, S.A.N.; Stith, C.M.; Burgers, P.M.; Kunkel, T.A. Inefficient proofreading and biased error rates during inaccurate DNA synthesis by a mutant derivative of Saccharomyces cerevisiae DNA polymerase delta. J. Biol. Chem. 2007, 282, 2324–2332. [CrossRef]
11. McElhinny, S.A.N.; Gordenin, D.A.; Stith, C.M.; Burgers, P.M.; Kunkel, T.A. Division of Labor at the Eukaryotic Replication Fork. Mol. Cell 2008, 30, 137–144. [CrossRef] [PubMed]
12. Roach, J.C.; Glusman, G.; Smit, A.F.A.; Huff, C.D.; Hubley, R.; Shannon, P.T.; Rowen, L.; Pant, K.P.; Goodman, N.; Bamshad, M.; et al. Analysis of Genetic Inheritance in a Family Quartet by Whole-Genome Sequencing. Science 2010, 328, 636–639. [CrossRef] [PubMed]
13. Alexandrov, L.B.; Nik-Zainal, S.; Wedge, D.C.; Aparicio, S.A.J.R.; Behjati, S.; Biankin, A.V.; Bignell, G.R.; Bolli, N.; Borg, A.; Borresen-Dale, A.-L.; et al. Signatures of mutational processes in human cancer. Nature 2013, 500, 415–421. [CrossRef] [PubMed]
14. Žamborszky, J.; Szikriszt, B.; Gervai, J.Z.; Pipek, O.; Póti, A.; Krzystanek, M.; Ribi, D.; Szalai-Gindl, J.M.; Csabai, I.; Szallasi, Z.; et al. Loss of BRCA1 or BRCA2 markedly increases the rate of base substitution mutagenesis and has distinct effects on genomic deletions. Oncogene 2017, 36, 746–755. [CrossRef] [PubMed]
15. Németh, E.; Lovrics, A.; Gervai, J.Z.; Seki, M.; Rosgo, G.; Bardelli, A.; Szüts, D. Two main mutational processes operate in the absence of DNA mismatch repair. DNA Repair 2020, 89, 102827. [CrossRef] [PubMed]
16. Brody, Y.; Kimerling, R.J.; Maruvka, Y.E.; Benjamin, D.; Elacqua, J.J.; Haradhvala, N.J.; Kim, J.; Mouw, K.W.; Frangak, K.; Koren, A.; et al. Quantification of somatic mutation flow across individual cell division events by lineage sequencing. Genome Res. 2018, 28, 1901–1918. [CrossRef]
17. Behjati, S.; Huch, M.; Van Boxtel, R.; Karthaus, W.; Wedge, D.C.; Tamuri, A.U.; Martincorena, I.; Petljak, M.; Alexandrov, L.B.; Gundem, G.; et al. Genome sequencing of normal cells reveals developmental lineages and mutational processes. Nature 2014, 513, 422–425. [CrossRef]
18. Blokzijl, F.; De Ligt, J.; Jager, M.; Sasselli, V.; Roerink, S.; Sasaki, N.; Huch, M.; Boymans, S.; Kuijk, E.; Prins, P.; et al. Tissue-Specific mutation accumulation in human adult stem cells during life. Nature 2016, 538, 260–264. [CrossRef]
43. Lujan, S.A.; Clark, A.B.; Kunkel, T.A. Differences in genome-wide repeat sequence instability conferred by proofreading and mismatch repair defects. *Nucleic Acids Res.* 2015, 43, 4067–4074. [CrossRef]

44. Watt, D.I.; Buckland, R.J.; Lujan, S.A.; Kunkel, T.A.; Chabes, A. Genome-Wide analysis of the specificity and mechanisms of replication infidelity driven by imbalanced dNTP pools. *Nucleic Acids Res.* 2016, 44, 1669–1680. [CrossRef]

45. Andrianova, M.A.; Bazykin, G.A.; Nikolaev, S.I.; Seplyarskiy, V.B. Human mismatch repair system balances mutation rates between strands by removing more mismatches from the lagging strand. *Genome Res.* 2017, 27, 1336–1343. [CrossRef]

46. Georgakopoulos-Soares, I.; Koh, G.; Momen, S.E.; Jiricny, J.; Hemberg, M.; Nik-Zainal, S. Transcription-coupled repair and mismatch repair contribute towards preserving genome integrity at mononucleotide repeat tracts. *Nat. Commun.* 2020, 11, 1–9. [CrossRef][PubMed]

47. Forbes, S.; Bhamra, G.; Bamford, S.; Dawson, E.; Kok, C.; Clements, J.; Menzies, A.; Teague, J.; Futreal, P.; Stratton, M. The Catalogue of Somatic Mutations in Cancer (COSMIC). *Curr. Protoc. Hum. Genet.* 2008, 57. [CrossRef]

48. Tate, J.G.; Bamford, S.; Jubbe, H.C.; Sondka, Z.; Beare, D.M.; Bindal, N.; Boutselakis, H.; Cole, C.G.; Creature, C.; Dawson, E.; et al. COSMIC: The Catalogue of Somatic Mutations in Cancer. *Nucleic Acids Res.* 2018, 47, D941–D947. [CrossRef]

49. Hodel, K.P.; De Borja, R.; Henninger, E.E.; Campbell, B.B.; Ungerleider, N.; Light, N.; Pursell, Z.F. Explosive mutation accumulation triggered by heterozygous human Pol epsilon proofreading-deficiency is driven by suppression of mismatch repair. *eLife* 2018, 7, e32692. [CrossRef]

50. Gladbach, Y.S.; Wiegele, L.; Hamed, M.; Merkenschläger, A.M.; Fuellen, G.; Junghanss, C.; Maletzki, C. Unraveling the Heterogeneous Mutational Signature of Spontaneously Developing Tumors in MLH1(-/-) Mice. *Cancers* 2019, 11, 1485. [CrossRef][PubMed]

51. Le, D.T.; Durham, J.N.; Smith, K.N.; Wang, H.; Bartlett, B.R.; Aulakh, L.K.; Lu, S.; Kemberling, H.; Wilt, C.; Luber, B.S.; et al. Mismatch repair deficiency predicts response of solid tumors to PD-1 blockade. *Science* 2017, 357, 409–413. [CrossRef][PubMed]

52. Bellfield, E.J.; Ding, Z.J.; Jamieson, F.J.; Visscher, A.M.; Zheng, S.J.; Mithani, A.; Harberd, N.P. DNA mismatch repair preferentially protects genes from mutation. *Genome Res.* 2018, 28, 66–74. [CrossRef][PubMed]

53. Frigola, J.; Sabariniathan, R.; Mularoni, L.; Muñoz, F.; Gonzalez-Perez, A.; López-Bigas, N. Reduced mutation rate in exons due to differential mismatch repair. *Nat. Genet.* 2017, 49, 1684–1692. [CrossRef][PubMed]

54. Sun, L.; Zhang, Y.; Zhang, Z.; Zheng, Y.; Du, L.; Zhu, B. Preferential Protection of Genetic Fidelity within Open Chromatin by the Mismatch Repair Machinery. *J. Biol. Chem.* 2016, 291, 17692–17705. [CrossRef][PubMed]

55. Supek, F.; Lehner, B. Differential DNA mismatch repair underlies mutation rate variation across the human genome. *Nat. Cell Biol.* 2015, 521, 81–84. [CrossRef][PubMed]

56. Reyes, G.X.; Zhao, B.; Schmidt, T.T.; Gries, K.; Kloor, M.; Hombauer, H. Identification of MLH2/hPMS1 dominant mutations that prevent DNA mismatch repair function. *Commun. Biol.* 2020, 3, 1–14. [CrossRef]

57. Li, F.; Mao, G.; Tong, D.; Huang, J.; Gu, L.; Yang, W.; Li, G.-M. The Histone Mark H3K36me3 Regulates Human DNA Mismatch Repair through Its Interaction with Mutyx. *Cell* 2013, 153, 590–600. [CrossRef][PubMed]

58. Aska, E.-M.; Dermadi, D.; Kauppi, L. Single-Cell Sequencing of Mouse Thymocytes Reveals Mutational Landscape Shaped by Replication Errors, Mismatch Repair, and H3K36me3. *iScience* 2016, 590–600. [CrossRef][PubMed]

59. Nishant, K.T.; Wei, W.; Mancera, E.; Argueso, J.L.; Schlattl, A.; Delhomme, N.; Alani, E. The baker’s yeast diploid genome is remarkably stable in vegetative growth and meiosis. *PloS Genet.* 2016, 6, e1001109. [CrossRef]

60. Conrad, D.F.; Keebler, J.E.; Depristo, M.A.; Lindsay, S.; Zhang, Y.; Cassals, F.; Idaghdour, Y.; Hartl, C.L.; Torroja, C.; Garimella, K.V.; et al. Variation in genome-wide mutation rates within and between human families. *Nat. Genet.* 2011, 43, 712–714. [CrossRef]

61. Denver, D.R.; Wilhelm, L.J.; Howe, D.K.; Gafner, K.; Dolan, P.C.; Baer, C.F. Variation in base-substitution mutation in experimental and natural lineages of Caenorhabditis nematodes. *Genome Biol. Evol.* 2012, 4, 513–522. [CrossRef][PubMed]

62. Ma, X.; Rogacheva, M.V.; Nishant, K.T.; Zanders, S.; Bustamante, C.D.; Alani, E. Mutation Hot Spots in Yeast Caused by Long-Range Clustering of Homopolymeric Sequences. *Cell Rep.* 2012, 1, 36–42. [CrossRef]

63. Kong, A.; Frigge, M.L.; Masson, G.; Besenbacher, S.; Sulem, P.; Magnusson, G.; Gudjonsson, S.A.; Sigurdsson, A.; Jonasdottir, A.; Jonasdottir, A.; et al. Rate of de novo mutations and the importance of father’s age to disease risk. *Nature* 2013, 488, 471–475. [CrossRef][PubMed]

64. Ness, R.W.; Morgan, A.D.; Colegrave, N.; Keightley, P.D. Estimate of the Spontaneous Mutation Rate in Chlamydomonas reinhardtii. *Genet.* 2012, 192, 1447–1454. [CrossRef][PubMed]

65. Saxer, G.; Havlak, P.; Fox, S.A.; Quance, M.A.; Gupta, S.; Fofanov, Y.; Strassmann, J.E.; Queller, D.C. Whole Genome Sequencing of Mutation Accumulation Lines Reveals a Low Mutation Rate in the Social Amoeba Dictyostelium discoideum. *PLOS ONE* 2012, 7, e46759. [CrossRef]

66. Michaelsson, J.J.; Shi, Y.; Gujral, M.; Zheng, H.; Malhotra, D.; Jin, X.; Jian, M.; Liu, G.; Greer, D.; Bhandari, A.; et al. Whole-Genome Sequencing in Autism Identifies Hot Spots for De Novo Germline Mutation. *Cell* 2012, 151, 1431–1442. [CrossRef]

67. Schrider, D.R.; Houle, D.; Lynch, M.; Hahn, M.W. Rates and genomic consequences of spontaneous mutational events in Drosophila melanogaster. *Genetics* 2013, 194, 937–954. [CrossRef]

68. Lang, G.I.; Parsons, L.; Gammie, A.E. Mutation Rates, Spectra, and Genome-Wide Distribution of Spontaneous Mutations in Mismatch Repair Deficient Yeast. *G3 Genes Genomes Genet.* 2013, 3, 1453–1465. [CrossRef]
69. Li, R.; Montpetit, A.; Rousseau, M.; Wu, S.Y.M.; Greenwood, C.M.T.; Spector, T.D.; Pollak, M.; Polychronakos, C.; Richards, J.B. Somatic point mutations occurring early in development: A monozygotic twin study. *J. Med. Genet.* **2013**, *51*, 28–34. [CrossRef] [PubMed]

70. Keightley, P.D.; Ness, R.W.; Halligan, D.L.; Haddrill, P.R. Estimation of the Spontaneous Mutation Rate per Nucleotide Site in a Drosophila melanogaster Full-Sib Family. *Genetics* **2014**, *196*, 313–320. [CrossRef] [PubMed]

71. Stirling, P.C.; Shen, Y.; Corbett, R.; Jones, S.J.M.; Hieter, P. Genome Destabilizing Mutator Alleles Drive Specific Mutational Trajectories in Saccharomyces cerevisiae. *Genetics* **2014**, *196*, 403–412. [CrossRef]

72. Weller, A.M.; Rödelsperger, C.; Eberhardt, G.; Molnar, R.I.; Sommer, R.J. Opposing Forces of A/T-Biased Mutations and G/C-Biased Gene Conversions Shape the Genome of the Nematode Pristionchus pacificus. *Genetics* **2014**, *196*, 1145–1152. [CrossRef]

73. Serero, A.; Jubin, C.; Loellet, S.; Leogoix-Né, P.; Nicolas, A.G. Mutational landscape of yeast mutator strains. *Proc. Natl. Acad. Sci. USA* **2014**, *111*, 1897–1902. [CrossRef]

74. Zhu, Y.O.; Siegal, M.L.; Hall, D.W.; Petrov, D.A. Precise estimates of mutation rate and spectrum in yeast. *Proc. Natl. Acad. Sci. USA* **2014**, *111*, E2310–E2318. [CrossRef] [PubMed]

75. Venn, O.; Turner, I.; Mathieson, I.; De Grooth, N.; Bontrop, R.; McVean, G. Strong male bias drives germline mutation in chimpanzees. *Science* **2014**, *344*, 1272–1275. [CrossRef]

76. Meier, B.; Cooke, S.L.; Weiss, J.; Bailly, A.P.; Alexandrov, L.B.; Marshall, J.; Campbell, P.J.C. elegans whole-genome sequencing reveals mutational signatures related to carcinogens and DNA repair deficiency. *Genome Res.* **2014**, *24*, 1624–1636. [CrossRef]

77. Keightley, P.D.; Pinharanda, A.; Ness, R.W.; Simpson, F.; Dasmahapatra, K.K.; Mallet, J.; Davey, J.W.; Jiggins, C.D. Estimation of the Spontaneous Mutation Rate in Heliconius melpomene. *Mol. Biol. Evol.* **2015**, *32*, 239–243. [CrossRef]

78. Francioli, L.C.; Genome of the Genome of the Netherlands Consortium; Polak, P.P.; Koren, A.; Menelaou, A.; Chun, S.; Renkens, I.; Van Duijn, C.M.; Swertz, M.A.; Wijmenga, C.; et al. Genome-Wide patterns and properties of de novo mutations in humans. *Nat. Genet.* **2015**, *47*, 822–826. [CrossRef] [PubMed]

79. Uchimura, A.; Higuchi, M.; Minakuchi, Y.; Ohno, M.; Toyoda, A.; Fujiyama, A.; Miura, I.; Wakana, S.; Nishino, J.; Yagi, T. Germline mutation rates and the long-term phenotypic effects of mutation accumulation in wild-type laboratory mice and mutator mice. *Genome Res.* **2015**, *25*, 1125–1134. [CrossRef]

80. Baranova, M.A.; Logacheva, M.D.; Penin, A.A.; Seplyarskiy, V.B.; Safonova, Y.Y.; Naumenko, S.A.; Klepikova, A.V.; Gerasimov, E.S.; Bazykin, G.A.; James, T.Y.; et al. Extraordinary Genetic Diversity in a Wood Decay Mushroom. *Mol. Biol. Evol.* **2015**, *32*, 2773–2783. [CrossRef]

81. Ness, R.W.; Morgan, A.D.; Vasanthakrishnan, R.B.; Colegrave, N.; Keightley, P.D. Extensive de novo mutation rate variation between individuals and across the genome of *Chlamydomonas reinhardtii*. *Genome Res.* **2015**, *25*, 1739–1749. [CrossRef]

82. Farlow, A.; Long, H.; Arnoux, S.; Sung, W.; Doak, T.G.; Nordborg, M.; Lynch, M. The Spontaneous Mutation Rate in the Fission Yeast Schizosaccharomyces pombe. *Genetics* **2015**, *201*, 737–744. [CrossRef]

83. Keith, N.; Tucker, A.E.; Jackson, C.E.; Sung, W.; Lledó, J.I.L.; Schrider, D.R.; Schaeck, S.; Dudycha, J.L.; Ackerman, M.S.; Younge, A.J.; et al. High mutational rates of large-scale duplication and deletion in *Daphnia pulex*. *Genome Res.* **2015**, *25*, 60–69. [CrossRef]

84. Rahbabi, R.; Wuster, A.; Lindsay, S.; Hardwick, R.J.; Alexandrov, L.B.; Al Turki, S.; Dominiczak, A.F.; Morris, A.D.; Porteous, D.; Smith, B.H.; et al. Timing, rates and spectra of human germline mutation. *Nat. Genet.* **2016**, *48*, 126–133. [CrossRef]

85. Haye, J.E.; Gammie, A.E. The Eukaryotic Mismatch Recognition Complexes Track with the Replisome during DNA Synthesis. *Mol. Biol. Evol.* **2016**, *32*, 126–133. [CrossRef]

86. Behringer, M.G.; Hall, D.W. The repeatability of genome-wide mutation rate and spectrum estimates. *Curr. Genet.* **2016**, *62*, 507–512. [CrossRef]

87. Sharp, N.P.; Agrawal, A.F. Low Genetic Quality Alters Key Dimensions of the Mutational Spectrum. *PLoS Biol.* **2016**, *14*, e1002419. [CrossRef]

88. Huang, W.; Lyman, R.F.; Lyman, R.A.; Carbone, M.A.; Harbison, S.T.; Magwire, M.M.; Mackay, T.F. Spontaneous mutations and the origin and maintenance of quantitative genetic variation. *eLife* **2016**, *5*, e14625. [CrossRef]

89. Smeds, L.; Qvarnström, A.; Ellegren, H. Direct estimate of the rate of germline mutation in a bird. *Genome Res.* **2016**, *26*, 1211–1218. [CrossRef]

90. Watson, J.M.; Platzer, A.; Kazda, A.; Akimcheva, S.; Valuchova, S.; Nizhynska, V.; Nordborg, M.; Riha, K. Germline replications and somatic mutation accumulation are independent of vegetative life span in Arabidopsis. *Proc. Natl. Acad. Sci. USA* **2016**, *113*, 12226–12231. [CrossRef]

91. Flynn, J.M.; Chain, F.J.; Schoen, D.J.; Cristescu, M.E. Spontaneous Mutation Accumulation in *Daphnia pulex* in Selection-Free vs. Competitive Environments. *Mol. Biol. Evol.* **2017**, *34*, 160–173. [CrossRef]

92. Besenbacher, S.; Sulem, P.; Helgason, A.; Helgason, H.; Kristjansson, H.; Jonasdottir, A.; Jonasdottir, A.; Magnusson, O.T.; Thorsteinsdottir, U.; Masson, G.; et al. Multi-nucleate de novo mutations in Humans. *PLoS Genet.* **2016**, *12*, e1006315. [CrossRef]

93. Hamilton, W.L.; Claessens, A.; Otto, T.D.; Kekre, M.; Fairhurst, R.M.; Rayner, J.C.; Kwiatkowski, D. Extreme mutation bias and high AT content in Plasmodium falciparum. *Nucleic Acids Res.* **2017**, *45*, 1889–1901. [CrossRef]
94. Liu, H.; Jia, Y.; Sun, X.; Tian, D.; Hurst, L.D.; Yang, S. Direct Determination of the Mutation Rate in the Bumblebee Reveals Evidence for Weak Recombination-Associated Mutation and an Approximate Rate Constancy in Insects. Mol. Biol. Evol. 2017, 34, 119–130. [CrossRef]

95. Feng, C.; Pettersson, M.; Lamichhaney, S.; Rubin, C.-J.; Rafati, N.; Casini, M.; Folkvord, A.; Andersson, L. Moderate nucleotide diversity in the Atlantic herring is associated with a low mutation rate. eLife 2017, 6, 6. [CrossRef]

96. Maretty, L.; Jensen, J.M.; Petersen, B.; Sibbesen, J.A.; Liu, S.; Villesen, P.; Skov, L.; Belling, K.G.-I.; Have, C.T.; Gonzalez-Izarzagaza, J.M.; et al. Sequencing and de novo assembly of 150 genomes from Denmark as a population reference. Nat. Cell Biol. 2017, 548, 58–91. [CrossRef]

97. Dutta, A.; Lin, G.; Pankajam, A.V.; Chakraborty, P.; Bhat, N.; Steinmetz, L.M.; Nishant, K.T. Genome Dynamics of Hybrid Saccharomyces cerevisiae During Vegetative and Meiotic Divisions. G3 Genes Genomes Genet. 2017, 7, 3669–3679. [CrossRef]

98. Jonsson, H.; Sulem, P.; Kehr, B.; Kristmundsdottir, S.; Zink, F.; Hjartarson, E.; Hardarson, M.T.; Hjorleifsson, K.E.; Eggertsson, H.P.; Gudjonsson, S.A.; et al. Parental influence on human germline de novo mutations in 1548 trios from Iceland. Nature 2017, 549, 519–522. [CrossRef]

99. Pfeifer, S.P. Direct estimate of the spontaneous germ line mutation rate in African green monkeys. Evolution 2017, 71, 2858–2870. [CrossRef]

100. Assaf, Z.J.; Tilk, S.; Park, J.; Siegal, M.L.; Petrov, D.A. Deep sequencing of natural and experimental populations of Drosophila melanogaster reveals biases in the spectrum of new mutations. Genome Res. 2017, 27, 1988–2000. [CrossRef]

101. Tatsumoto, S.; Go, Y.; Fukuta, K.; Noguchi, H.; Hayakawa, T.; Tomonaga, M.; Hirai, H.; Matsuzaawa, T.; Agata, K.; Fujiyama, A. Direct estimation of de novo mutation rates in a chimpanzee parent-offspring trio by ultra-deep whole genome sequencing. Sci. Rep. 2017, 7, 1–12. [CrossRef][PubMed]

102. Meier, B.; Volkova, N.V.; Hong, Y.; Schofield, P.; Campbell, P.J.; Gerstung, M.; Gartner, A. Mutational signatures of DNA mismatch repair deficiency in C. elegans and human cancers. Genome Res. 2018, 28, 666–675. [CrossRef]

103. Krasovec, M.; Chester, M.; Ridout, K.; Filatov, D.A. The Mutation Rate and the Age of the Sex Chromosomes in Chironomus riparius. Evol. Lett. 2017, 1, 86–92. [CrossRef]

104. Oppold, A.-M.; Pfenninger, M. Direct estimation of the spontaneous mutation rate by short-term mutation accumulation lines in Chironomus riparius. Evol. Biol. 2018, 28, 1832–1838.e4. [CrossRef]

105. Thomas, G.W.; Wang, R.J.; Puri, A.; Harris, R.A.; Raveendran, M.; Hughes, D.S.; Murali, S.C.; Williams, L.E.; Doddapaneni, H.; Muzny, D.M.; et al. Reproductive Longevity Predicts Mutation Rates in Primates. Curr. Biol. 2018, 28, 3193–3197.e5. [CrossRef]

106. Weng, M.-L.; Becker, C.; Hildebrandt, J.; Neumann, M.; Rutter, M.T.; Shaw, R.G.; Weigel, D.; Fenster, C.B. Fine-Grained Analysis of Spontaneous Mutation Spectrum and Frequency in Arabidopsis thaliana. Genetics 2019, 211, 703–714. [CrossRef]

107. Besenbacher, S.; Hvilsom, C.; Marques-Bonet, T.; Mailund, T.; Schierup, M.H. Direct estimation of mutations in great apes reconciles phylogenetic dating. Nat. Ecol. Evol. 2019, 3, 286–292. [CrossRef][PubMed]

108. Xu, S.; Stapley, J.; Gablenz, S.; Boyer, J.; Appenroth, K.J.; Sree, K.S.; Huber, M. Low genetic variation is associated with low mutation rate in the giant duckweed. Nat. Commun. 2019, 10, 1243. [CrossRef]

109. Konrad, A.; Brady, M.J.; Berghorsson, U.; Katju, V. Mutational Landscape of Spontaneous Base Substitutions and Small Indels in Experimental Caenorhabditis elegans Populations of Diverging Size. Genetics 2019, 212, 837–854. [CrossRef]

110. Krasovec, M.; Sanchez-Brosseau, S.; Piganeau, G. First Estimation of the Spontaneous Mutation Rate in Diatoms. Genome Biol. Evol. 2019, 11, 1829–1837. [CrossRef]

111. Koch, E.M.; Schweizer, R.M.; Schweizer, T.M.; Stahler, D.R.; Smith, U.W.; Wayne, R.K.; Novembre, J. De novo mutation rate estimation in wolves of known pedigree. Mol. Biol. Evol. 2019, 36, 2536–2547. [CrossRef]

112. Williams, J.S.; Lujan, S.A.; Zhou, Z.-X.; Burkholder, A.B.; Clark, A.B.; Fargo, D.C.; Kunkel, T.A. Genome-wide mutagenesis resulting from topoisomerase 1-processing of un repaired ribonucleotides in DNA. DNA Repair 2019, 84, 102641. [CrossRef]

113. Lindsay, S.J.; Rahbarni, R.; Kaplanis, J.; Keane, T.; Hurles, M.E. Similarities and differences in patterns of germline mutation between mice and humans. Nat. Commun. 2019, 10, 1–12. [CrossRef]

114. Wu, F.L.; Strand, A.I.; Cox, L.A.; Ober, C.; Wall, J.D.; Moorjani, P.; Przeworski, M. A comparison of humans and baboons suggests germline mutation rates do not track cell divisions. PLoS Biol. 2020, 18, e3000838. [CrossRef]

115. Sandler, G.; Bartkowska, M.; Agrawal, A.F.; Wright, S.I. Estimation of the SNP Mutation Rate in Two Vegetatively Propagating Species of Duckweed. G3 Genes Genomes Genet. 2020, 10, 4191–4200. [CrossRef]

116. Sui, Y.; Qi, L.; Wu, J.-K.; Wen, X.-P.; Tang, X.-X.; Ma, Z.-J.; Wu, X.-C.; Zhang, K.; Kokoska, R.J.; Zheng, D.-Q.; et al. Genome-wide mapping of spontaneous genetic alterations in diploid yeast cells. Proc. Natl. Acad. Sci. USA 2020, 117, 28191–28200. [CrossRef]

117. Quah, S.K.; von Borstel, R.C.; Hastings, P.J. The origin of spontaneous mutation in Saccharomyces cerevisiae. Genetics 1980, 96, 819–839. [CrossRef][PubMed]

118. Roche, H.; Gietz, R.D.; Kunz, B.A. Specificity of the yeast rev3 delta antitumour and REV3 dependency of the mutator resulting from a defect (rad1 delta) in nucleotide excision repair. Genetics 1994, 137, 637–646. [CrossRef]

119. Pavlov, Y.I.; Shcherbakova, P.V.; Kunkel, T.A. In Vivo consequences of putative active site mutations in yeast DNA polymerases alpha, epsilon, delta, and zeta. Genetics 2001, 159, 47–64. [CrossRef]
Cells 2021, 10, 1224

120. Kraszewska, J.; Garbacz, M.; Jonczyk, P.; Fijalkowska, I.J.; Jaszczur, M. Defect of Dpb2p, a noncatalytic subunit of DNA polymerase ε, promotes error prone replication of undamaged chromosomal DNA in Saccharomyces cerevisiae. Mutat. Res. Mol. Mech. Mutagen. 2012, 737, 34–42. [CrossRef]

121. Garbacz, M.; Araki, H.; Flis, K.; Bebenek, A.; Zawada, A.E.; Jonczyk, P.; Makiela-Dzbenska, K.; Fijalkowska, I.J. Fidelity consequences of the impaired interaction between DNA polymerase epsilon and the GINS complex. DNA Repair 2015, 29, 23–35. [CrossRef] [PubMed]

122. Garbacz, M.A.; Cox, P.B.; Sharma, S.; Lujan, S.A.; Chabes, A.; Kunkel, T.A. The absence of the catalytic domains of Saccharomyces cerevisiae DNA polymerase ε strongly reduces DNA replication fidelity. Nucleic Acids Res. 2019, 47, 3986–3995. [CrossRef] [PubMed]