Multidisciplinary Research Journal
ISSN (online): *Applied for*

Volume 1, Issue 4, 2025
October - December 2025

Review

Base Substitutions in Genomes Due to Deamination and Oxidation of DNA Bases, Favoring Genome Compositional Biases

Nishita Deka [1,5], Pratyush Kumar Beura [1], Monika Jain [1], Najima Ahmed [4], Ramesh Chandra Deka [2,4,5], Siddhartha Shankar Satapathy [3,4,5], Suvendra Kumar Ray [1,5]*

[1] Department of Molecular Biology and Biotechnology, Tezpur University, Tezpur, Assam, INDIA-784028
[2] Department of Chemical Sciences, Tezpur University, Tezpur, Assam, INDIA-784028
[3] Department of Computer Science & Engineering, Tezpur University, Tezpur, Assam, INDIA-784028
[4] Centre for Multidisciplinary Research, Tezpur University, Tezpur, Assam, INDIA-784028
[5] Centre for Bioinformatics & Computational Biology, Tezpur University, Tezpur, Assam, INDIA-784028
* Correspondence: suven@tezu.ernet.in

Scopus Author ID 8347695700

Received: 14 August 2025; Accepted: 30 October 2025; Published: 23 December 2025

Abstract: The genome G+C content of bacteria varies widely, from 13% to 75%, which is influenced by both environmental and internal mutation pressure; however, the precise determinants of this variability remain unresolved. Mutation-based models, such as Sueoka's directional mutation hypothesis, suggest that G+C content arises from mutational pressures within an organism without providing any specific advantage to it. Though there are several advantages associated with genome G+C%, there is limited evidence, favoring any selection mechanism for G+C% evolution. Hence, the genome G+C% in organisms is largely studied under the neutral theory of evolution. Cytosine deamination and guanine oxidation are recognized as major contributors to A/T mutational bias, producing frequent substitutions such as C→T transitions and G→T transversions, respectively. While these mechanisms leading to A+T enrichment have been well studied, counteracting processes that promote G+C enrichment in organisms are comparatively less understood. This review mainly highlights adenine as an underexplored contributor: its deamination and oxidation yield A→G and A→C substitutions, respectively, both biased toward increased G+C content. We further consider how the efficiency of DNA repair mechanisms may shape G+C content across evolutionary timescales. Together, these perspectives address a gap in the current understanding of the mutational forces influencing genome composition.

Keywords: Genome G+C %; Base substitution mutation; Deamination; Oxidative lesions; Transition; Transversion; DNA repair

1. Introduction

G+C content of genomes refers to the percentage of guanine (G) and cytosine (C) bases in the chromosome of an organism. The genomic G+C content of bacteria varies widely, ranging from 13% to approximately 75% [1]. Remarkably, different bacterial species can maintain a relatively fixed G+C content over evolutionary time. It is observed that closely related species often share similar G+C content, linked to phylogeny, but it is not strict, referred to as the "phylogenetic G+C content paradox." For instance, while Actinomycetes have a high G+C content (>55%) and Firmicutes have a low G+C content (~43%), both are Gram-positive bacteria. In the case of Gram-negative bacteria, α-proteobacteria and γ-proteobacteria have a wide range of genome G+C%, ranging from high to low, but β proteobacterium has a high

genome G+C% [2], and Spirochaetes have a G+C% of 40.6%. Closely related species can exhibit markedly different G+C contents. Factors include mutation bias (e.g., differential rates of A:T→G:C versus G:C→A: T mutations), variations in DNA repair efficiency that affect the fixation of specific nucleotide changes, and horizontal gene transfer, which can introduce DNA segments with distinct base compositions from unrelated taxa. Together, these processes contribute to shaping the genomic G+C content, explaining why evolutionary divergence in G+C content does not always align with phylogenetic relatedness. Xylella fastidiosa (recently named as *Xanthomonas fastidiosa*) and *Xanthomonas oryzae* and other *Xanthomonas* are both Gram-negative bacteria classified within the same genus; however, X. fastidiosa has a comparatively low genomic G+C content of ~50%, whereas members of *Xanthomonas* typically possess a higher G+C content of ~65% [3]. In contrast, both *Ralstonia solanacearum*, a plant pathogen, and *Cupriavidus taiwanensis*, a rhizosphere-associated nitrogen-fixing symbiotic bacterium, belong to the class β-Proteobacteria and display similar genome G+C contents, despite their markedly different lifestyles [4]. The factors that contribute to an organism's genome G+C% are still not well understood. Understanding the factors that determine and maintain genome G+C content remains a critical area of evolutionary research. Two main theories have provided some explanations regarding the wide range of variability in G+C content across the prokaryotic genomes: first, the selectionist theory and the mutationist theory [5]. One selectionist hypothesis suggested that the relatively high G+C content observed in some soil-surface bacteria might help minimize the risk of thymidine dimer formation under UV exposure, as the probability of formation of thymidine dimer in AT-rich genomes is likely to be more than that in GC rich because the chance of two consecutive thymidines is higher in AT genomes than in GC rich genomes. However, subsequent studies indicated that bacteria inhabiting the soil surface and those residing beneath it do not differ significantly in their genomic G+C content. Therefore, while the hypothesis was thought-provoking, it has not received broad support [6,7], Similarly, the correlation between thermophilic bacteria and higher G+C content was initially noted because G: C base pairs possess an extra hydrogen bond and a more stable stacking pattern compared to A: T pairs, contributing to greater thermal stability [8],[9]. Researchers proposed the correlation between the preferential usage of amino acids coded by GC-rich codons in thermophilic bacteria, but this hypothesis was not accepted. Higher G+C content stabilizes RNA secondary structures, improving gene expression and translation, particularly under stress conditions [10]. For example, bacteria in high-temperature environments tend to have G+C-rich coding regions, promoting stable codon-anticodon pairing during translation. However, certain bacteria, including *Thermoanaerobacter* sp. and *Caldicellulosiruptor saccharolyticus*, inhabit high-temperature environments while maintaining A+T-rich genomes [2]. It is interesting to note that two microbes living in the same environmental habitat may have different genome G+C content; for example, the G+C% of all microbes residing in the human gut is not the same [11], which implies that genome G+C% is not completely determined by environmental factors.

In 1962, Noboru Sueoka introduced a model for understanding changes in GC content [12], suggesting that such changes resulting from the conversion of base pairs between AT (adenine-thymine) and GC (guanine-cytosine) contribute to the mutationist view. Sueoka denoted the AT/TA pair as the "α" pair and the GC/CG pair as the "γ" pair, proposing that conversion rates between α and γ are relatively uniform across the genome. Hence, the equilibrium between G/C ←→ A/T mutational patterns varies among all the bacterial species.

This ratio of nucleotide conversion is termed GC mutational pressure [13]. This GC mutational pressure has been known to be influenced by the overall conversion between the nitrogenous bases. However, only those conversions that persist and are passed on to future generations affect the overall GC content, leading to what he termed the effective base conversion rate [12]. Later, in 1988, Sueoka proposed the directional mutation theory, which suggests that mutation pressure is not random but instead pushes genomes toward either higher or lower G+C content [14]. Further, Hershberg and Petrov identified a consistent bias toward AT content in bacterial genomes, where *de novo* G/C →A/T mutations occur more frequently than the reverse [15]. Rocha and Feil stated that the mutation pattern cannot explain the genome composition and gave additional insights into mechanisms influencing genome G+C content. They proposed that factors such as purifying selection and biased gene conversion (BGC) can counteract AT bias over time, especially in GC-rich organisms [16]. They also mentioned the antagonistic nature of mutation and selection in the persistence of GC composition in bacterial genomes, where mutation has a significant impact on shaping genome composition [16]. However, the low GC composition in organisms can be confronted by the AT-biased mutation concept. However, a rapid decline of GC content was observed in an experiment by Falk Hildebrand and his co-workers (2010) [5], which indicated an AT-biased mutation pattern [17]. Undoubtedly, such an AT-biased mutation study supports the concept of a rapid decline in GC composition across bacterial genomes; however, the role of selection cannot be ignored. In fact, both the mutationist and selectionist views remain a topic of debate among molecular evolutionary biologists.

Interestingly, the majority of the prokaryotic genome is composed of coding sequences, making it challenging to study GC mutational pressure. Within coding regions, codon degeneracy plays a central role in shaping nucleotide composition. The second codon position is under strong purifying selection because it largely determines amino acid identity. In contrast, the third codon position, particularly at four-fold degenerate sites, is often more tolerant to change. These synonymous sites are less constrained and therefore more directly influenced by mutational biases, making them valuable indicators of GC mutational pressure [18]. However, codon degeneracy is not evolutionarily neutral. Codon usage bias, arising from the unequal use of synonymous codons, has functional consequences beyond the genetic code's degeneracy. The availability of tRNA influences codon selection, which in turn impacts co-translational protein folding, translational fidelity, and efficiency. Rare codons may slow translation and aid in appropriate protein folding, whereas preferred codons that match abundant tRNAs improve the speed and precision of protein synthesis. Therefore, while codon degeneracy offers redundancy, it also serves as the basis for selection pressures that influence the regulation of gene expression and the composition of the genome [19],[20].

Considering the above observations, there is scope for further research to gain a deeper understanding of the mechanism behind genome composition in bacteria.

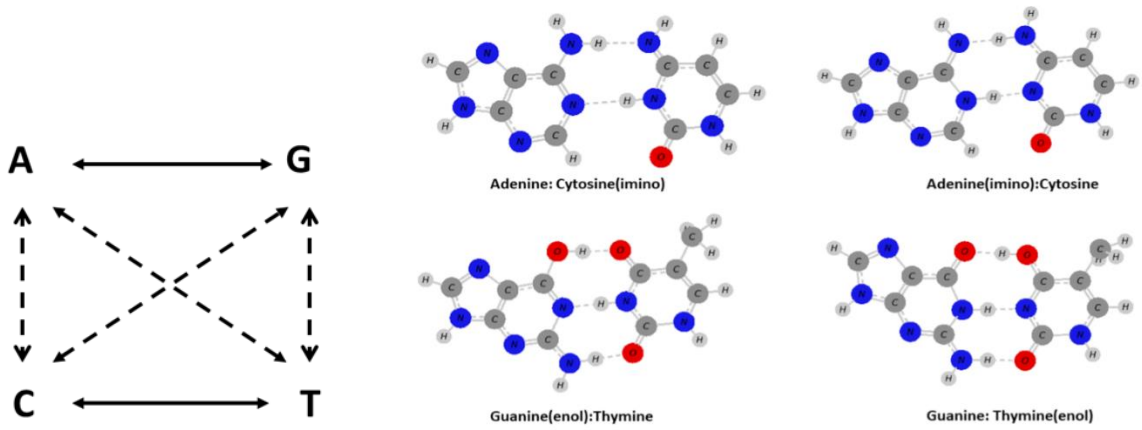
2. Base Substitution Mutation in Genomes

Among various types of mutations, base substitution mutation is one of the main causes for shaping various genomic features like genome GC content, codon usage bias, strand asymmetry, etc. [14],[21]. Base substitution mutations play a significant role in shaping GC content during genome evolution. These mutations can either increase or decrease the frequency of G and C nucleotides in the genome, depending on selective pressures and

mutational biases [22]A transition (ti) is a base substitution occurring within the same class of nitrogenous bases (purine–purine or pyrimidine–pyrimidine), whereas a transversion (tv) occurs between classes (purine–pyrimidine). There is a total of 12 types of substitutions possible (Table 1), out of which 8 are tv and 4 are ti [23],[24] (Figure 1 (a)). Chemical modifications such as base deamination and oxidation can induce transition and transversion mutations, respectively. Cytosine deamination and guanine oxidation are well-established contributors to genome G+C content. However, the potential roles of adenine deamination and oxidation in shaping G+C content remain underexplored.

Table 1. Base substitution and their related mis-pairs

Sl. no	Base substitution	Resulting mismatch	Mismatch due to chemical modification of bases
1	A->C	A:G or C:T	8-oxo-A(syn): G
2	A->G	G:T or A:C	Hypoxanthine: C
3	A->T	A:A or T:T	
4	C->A	T:C or G:A	
5	C->G	G:G or C:C	
6	C->T	G:T or A:C	U:G
7	G->A	G:T or A:C	
8	G->C	C:C or G:G	
9	G->T	T:C or A:G	8-oxo-G(syn): A
10	T->A	A:A or T:T	
11	T->C	A:C or G:T	
12	T->G	A:G or C:T	



(a) (b)

Figure 1. (a) Transition and Transversion substitution between bases. (The solid lines denote transitions (4 ti) and dashed lines denote transversions (8 tv)); **(b)** Mis pairs due to tautomerization

3. The Chemistry behind Base Substitutions

In organisms, base tautomerization and chemical modification of DNA are major driving forces of base substitution, which occurs during replication due to base mispairing (Figure 1 (b)). In DNA, the amino and keto tautomeric forms are generally more stable than the imino and enol forms. Density functional theory (DFT) studies further support this, showing that the relative stabilities of rare tautomers differ among bases: the enol form of guanine and the imino form of cytosine are more stable than the corresponding rare forms of thymine and

adenine. Importantly, these differences in stability have biological consequences. Rare but sufficiently stable tautomers can pair in non-canonical ways, generating mismatches such as G·T or A·C. If not corrected by proofreading or repair mechanisms, these mismatches are propagated during replication, leading to transition mutations (C→T and G→A), which are among the most frequent substitutions observed in genomes. Thus, the greater the stability of a rare tautomer-induced mispair, the higher the probability it will persist long enough to escape repair and contribute to the overall mutation rate.

4. The Rate Difference between Transition and Transversion across Genomes

The frequency of different base substitutions varies because individual bases differ in their susceptibility to deamination and oxidation [25]. Transition (ti) and transversion (tv) rates also differ markedly: ti events are generally two or more times more frequent than tv across organisms [26],[27]. Among transitions, C→T (or G→A) occurs more frequently than T→C (or A→G) [28]. Similarly, among the eight possible transversions, G→T (or C→A) is typically observed at a higher frequency than the others [29].

Several molecular mechanisms underlie these biases. For example, cytosine deamination produces uracil, leading to the frequent C→T transition [24], while oxidative damage of guanine to 8-oxo-guanine favors G→T transversions [30]. Despite extensive characterization of these processes, their cumulative impact on genomic G+C composition remains insufficiently explored.

In addition to spontaneous lesions, other chemical modifications also influence substitution patterns. For instance, methylation of bases such as guanine (O6-methylguanine) and thymine (O4-methylthymine) can cause G:C→A:T and T:A→C:G mutations, respectively [31]. These reactions occur enzymatically through S-adenosylmethionine (SAM)-dependent methylases at specific DNA sequences. Likewise, ultraviolet (UV) radiation induces pyrimidine dimers, which can promote substitution mutations. However, because methylation and UV damage are not random, they are considered secondary to the spontaneous processes of deamination and oxidation that are the main focus of this review.

5. Damages to Nitrogenous Bases

5.1. Deamination of Bases

Of the four DNA bases, cytosine, adenine, and guanine contain amino groups. Deamination is the removal of an amino group, converting one base into another. This process occurs in all organisms and represents a major source of DNA damage. It takes place through temperature- and pH-dependent reactions and produces uracil, hypoxanthine, and xanthine from cytosine, adenine, and guanine, respectively [32] (Figure 2). If unrepaired, these deamination events can generate transition mutations, contributing to mutagenesis.

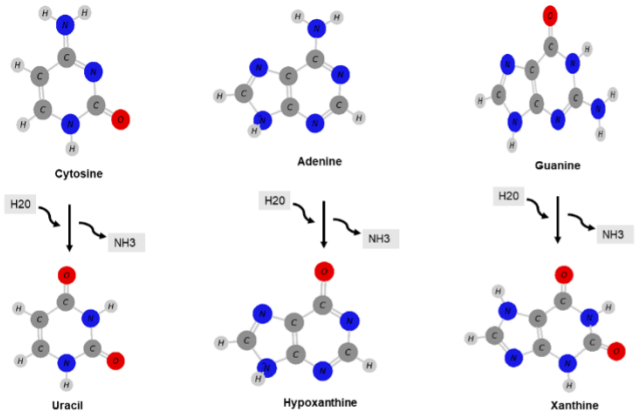
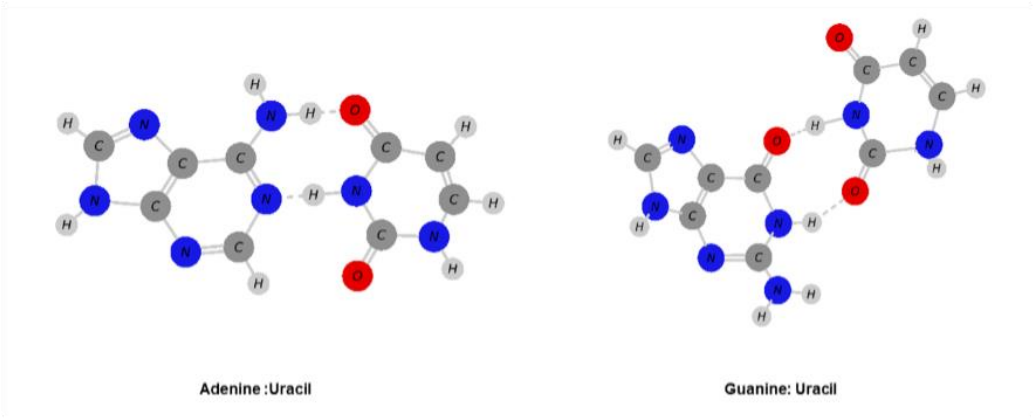
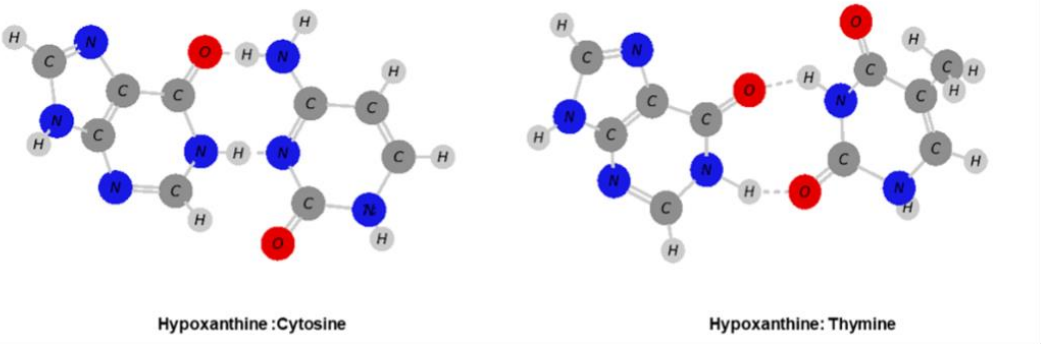


Figure 2. Cytosine, adenine, and guanine upon deamination give rise to uracil, hypoxanthine, and xanthine, respectively.

5.2 Implications of Deamination on Genomes



(a)

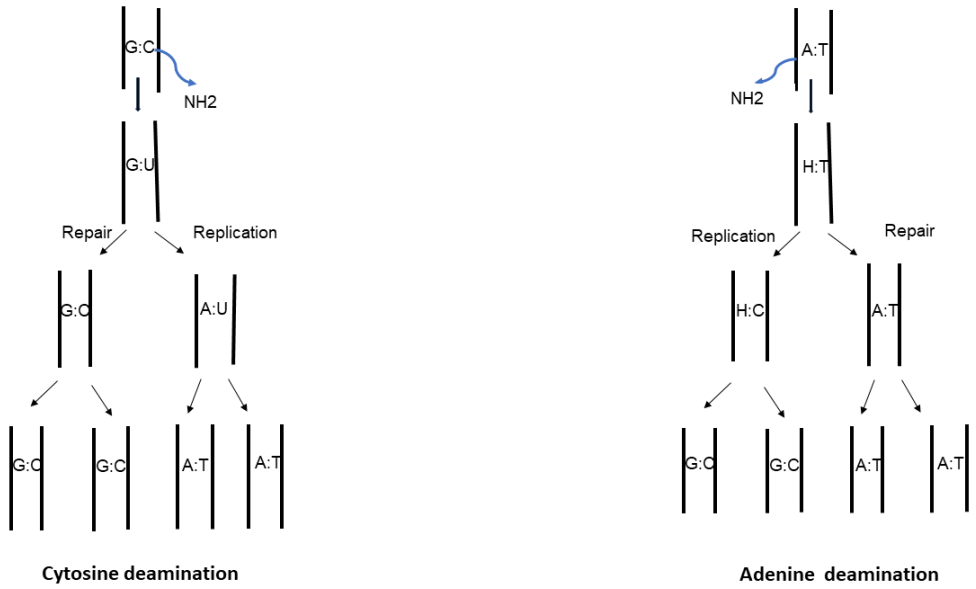


(b)

Figure 3. (a) Uracil can pair with Adenine and Guanine. Uracil: Guanine mis pair causes G:C→A: T transitions; **(b)** Hypoxanthine pairing with cytosine and thymine, respectively. Hypoxanthine pairs with cytosine and does A: T→G:C transitions

When cytosine undergoes deamination, it is converted into uracil. Due to cytosine deamination the original C:G pair during replication got altered by U:G pair (C→U)as shown in Figure 3 (a). During the next round of DNA replication, the uracil in this mismatch can pair with adenine, leading to the incorporation of an A opposite the U. If the lesion is not repaired, this results in the permanent substitution of the original C:G base pair with an A:T base pair,

producing a G:C → A:T transition [33]. Similarly, deamination of adenine produces hypoxanthine (HX), which preferentially pairs with cytosine rather than thymine. During DNA replication, an unrepaired HX: C pairing can replace an A: T pair with a G: C pair, generating an A: T→G: C transition mutation (Figure 3 (b)). Such errors are typically corrected by base excision repair (BER) or Alternate Excision Repair (AER), but if left unrepaired, they contribute to transition mutation. Deamination of guanine yields xanthine, which retains the ability to pair efficiently with cytosine. Consequently, its effect is considered minor compared to cytosine and adenine deamination [34],[35],[36],[37],[38],[39]. Figures 4 (a) and 4 (b) summarize the overall impact of base deamination on genomic nucleotide composition.



(a) (b)

Figure 4. (a) Cytosine deamination results in G:C→A: T type of transition mutation if not repaired; (**b)** Adenine deamination results in A: T→G:C type of transition mutation if not repaired

5.3 Factors Influencing the Deamination of Bases

Cytosine deamination is the most common type of base deamination, occurring at an estimated ~100 events per cell per day, whereas adenine deamination is much rarer, with frequencies ~50-fold lower [32].[34]. At 37 °C and physiological pH, the half-life of guanine against deamination is ~10^6–10^7 years per base. The estimated half-life corresponds to approximately 1 deamination event per 10^6–10^7 guanines per day per cell. [37],[38]. However, Wang and Hu reported that in unbuffered solutions, adenine deamination occurs at a significantly higher rate than cytosine, challenging earlier assumptions. Cytosine deamination is also accelerated under mildly acidic conditions (pH 5–6) and in buffered solutions, while adenine and guanine deamination are generally absent under these circumstances [37],[40]. Several factors influence the rate of deamination. Cytosine has been extensively studied in this context, while adenine and guanine deamination remain less characterized.

5.3.1. Single-stranded vs. Double-stranded DNA

Cytosine deamination in single-stranded DNA proceeds with a half-life of ~200 years, compared to ~30,000 years in double-stranded DNA [41]. The increased susceptibility of

single-stranded DNA arises from greater accessibility of the N3 position of cytosine, which is shielded by Watson–Crick base pairing in the double helix [42].

5.3.2. Mismatched vs. Canonical Base Pairs

Mis-paired cytosines are particularly prone to deamination, with rates 10–100 times higher in C: C or C: T mismatches than in canonical C: G pairs. Cytosines in C: C mismatches deaminate ~3-fold faster than those in C: T mismatches, especially at elevated temperatures (e.g., 60°C). In general, mismatched cytosines exhibit 8–26-fold higher deamination rates, approaching those observed in single-stranded DNA at 37°C. Pyrimidine–pyrimidine mismatches (C: C, T: T) are especially conducive to deamination due to helix destabilization and formation of "open" base pairs that favor hydrolytic attack [42]

5.3.3. pH and Temperature

Free cytosine and cytidine undergo rapid deamination upon heating in weakly acidic buffers, whereas adenosine and guanosine show no detectable changes under similar conditions. Under alkaline conditions, cytosine deamination is the predominant form of DNA degradation, proceeding ~10 times faster than other alkali-catalyzed modifications [40]. Within DNA, the rate of cytosine deamination is slowest at pH 8.0–8.5. Elevated temperatures also enhance deamination: at 95°C, denaturation exposes cytosine residues and accelerates their deamination, while at 70°C, partial renaturation slows the process [43].

Overall, cytosine deamination represents the dominant spontaneous base modification in DNA, with adenine deamination occurring much less frequently. The rate differences across structural, chemical, and environmental contexts highlight how deamination contributes to mutation bias and may influence genomic GC content.

5.4 Oxidation of DNA Bases

Oxidative DNA damage arises from reactive oxygen species (ROS) generated during normal cellular metabolism or from external sources such as UV radiation, chemicals, and pollutants. ROS, including superoxide, hydrogen peroxide, and hydroxyl radicals, can modify DNA bases, thereby compromising genomic integrity [32],[44]. Among the bases, guanine has the lowest standard reduction potential, making it the most susceptible to oxidation. This leads to the formation of 8-oxo-7,8-dihydroguanine (8-oxo-G) [45] (Figure 5).

8-oxo-G is widely used as a biomarker of oxidative stress and is associated with aging, cardiovascular disease, neurodegenerative disorders, and cancer [46]. It also has the lowest ionization potential among the DNA bases, making it the preferred target for one-electron oxidizing agents. Such agents can attack guanine directly or via hole transfer from other radical cation sites. Additionally, singlet oxygen, generated by UV sunlight, reacts specifically with guanine [47].

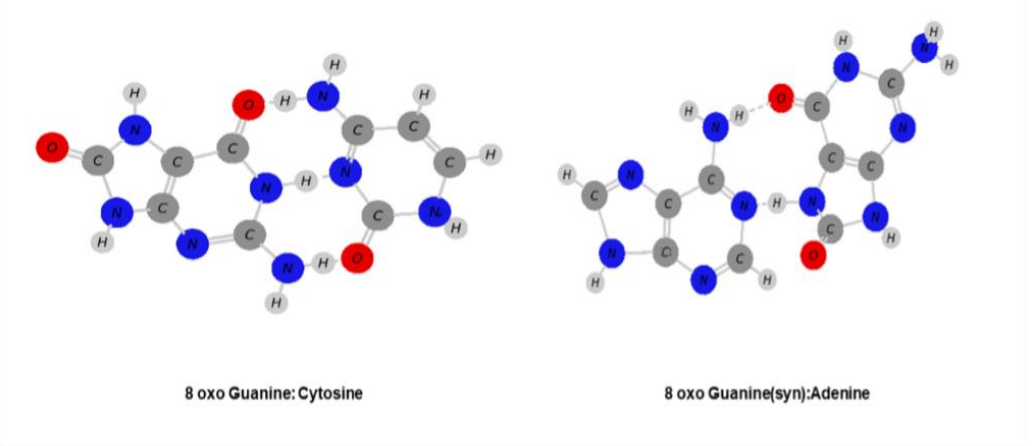
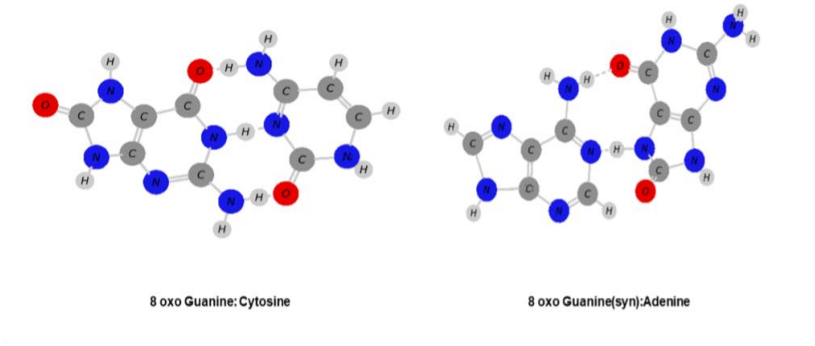


Figure 5. Guanine and Adenine upon oxidation, get converted to 8-oxo-Guanine and 8-oxo-Adenine, respectively

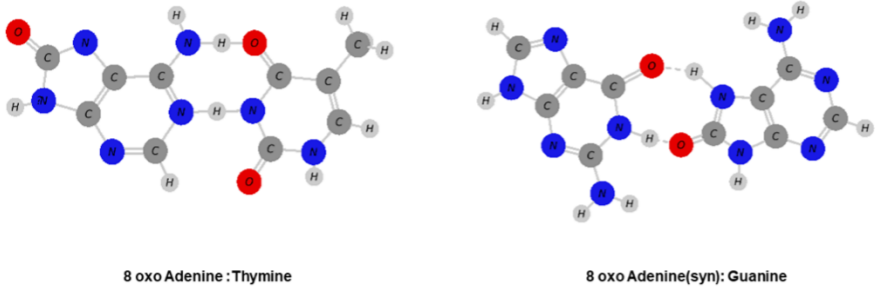
Although studied less extensively, adenine is also susceptible to oxidative modification, resulting in the production of 8-oxo-adenine (8-oxo-A). This lesion has mutagenic potential, as it can mispair with guanine during replication, contributing to transversion mutations. While its biological significance is less characterized compared to 8-oxo-G, adenine oxidation represents an additional pathway by which oxidative stress influences genomic stability [48].

5.5. Implications of Oxidative Damage to Bases in Genomes

8-oxo-guanine, the oxidative adduct of guanine, is a potent mutagen that induces G→T transversion mutations in both eukaryotes and prokaryotes (Figure 7 (a)). In its syn conformation, 8-oxo-G uses the Hoogsteen edge to base pair with adenine, while in its anti-conformation, it can still pair with cytosine, as does unmodified guanine (Figure 6 (a)) [49]. DNA polymerases in both eukaryotes and prokaryotes preferentially incorporate adenine opposite 8-oxo-G rather than cytosine, thereby promoting G:C→T: A substitution [32].



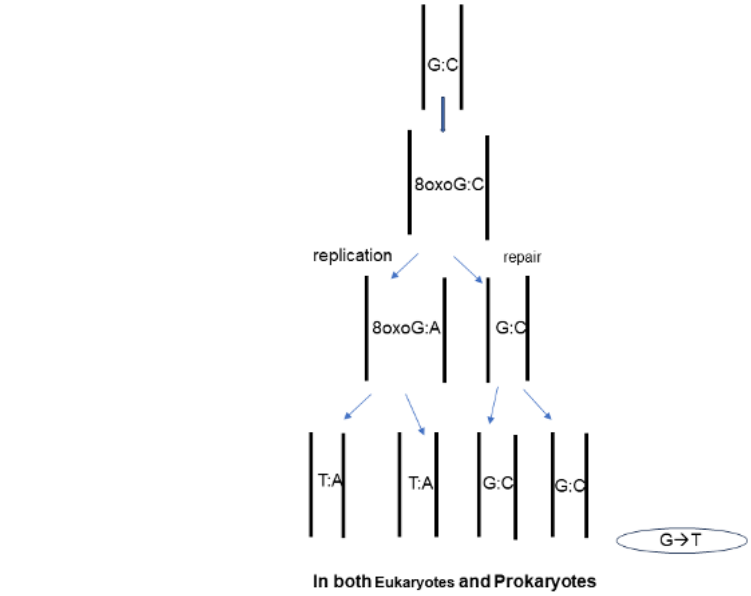
(a)



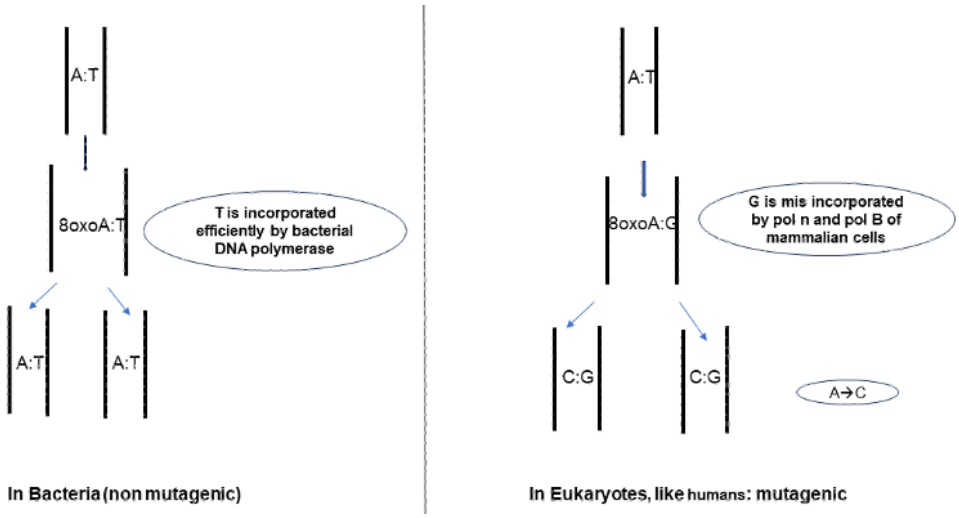
(b)

Figure 6. (a) Base pairing of 8-oxoG with cytosine; **(b)** Base pairing of 8-oxoA with thymine and Guanine

Adenine oxidation generates 8-oxo-adenine (8-oxo-A), which is also mutagenic but underexplored. In higher eukaryotes, 8-oxo-A contributes to A→C transversions (Figure 7 (b)) [30]. It is considerably less mutagenic than 8-oxo-G, with an estimated <10% relative mutagenicity [45]. In prokaryotes, DNA polymerases generally incorporate thymine opposite 8-oxo-A, allowing error-free bypass. By contrast, in mammalian cells, DNA polymerases such as Polη and Polβ can insert guanine opposite 8-oxo-A (Figure 6 (b)) [50], [51], thereby generating A: T→C: G mutations [50].



(a)



(b)

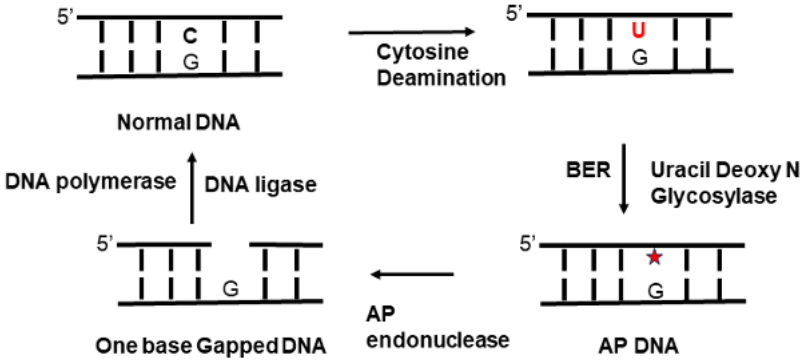
Figure 7. (a) Implication of Guanine oxidation on mutation bias. Oxidative lesions of Guanine give G:C→T: A type of transversion mutation in both prokaryotes and eukaryotes; **(b)** Implication of adenine oxidation on mutation in both prokaryotes and eukaryotes. Oxidation of adenine is found to be mostly non-mutagenic to bacteria; however, in higher eukaryotes, it can cause A→C transversions.

6. Repair of DNA damage

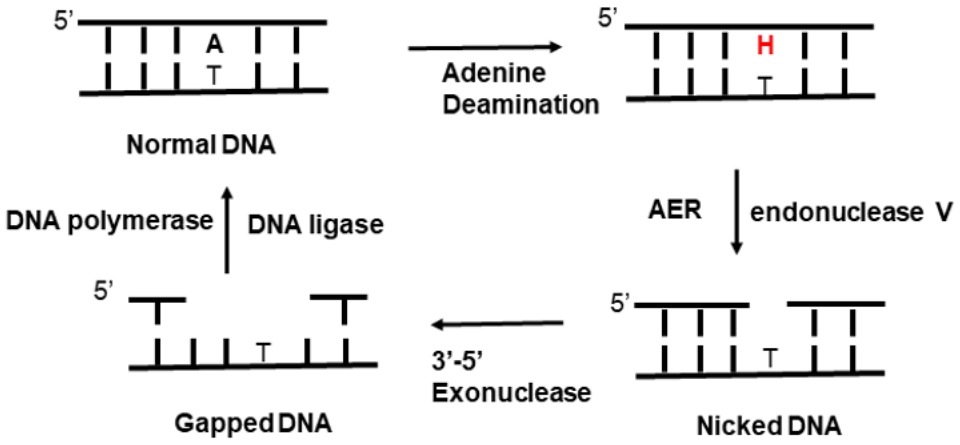
6.1. Repair of Deaminated Bases in DNA

In prokaryotes, uracil is excised through the base excision repair (BER) pathway by uracil-DNA glycosylase (UDG) (Figure 8 (a)). Hypoxanthine can be removed, though less efficiently, by AlkA (alkyladenine DNA glycosylase). Endonuclease V (EndoV) provides an additional mechanism, recognizing deoxyinosine (dI, the deoxynucleotide form of hypoxanthine) and cleaving the second phosphodiester bond on the 3′ side of the lesion within double-stranded DNA [31],[52]. This activity, conserved across Bacteria, Archaea, and Eukaryotes, has been proposed to initiate an alternative excision repair pathway for deaminated purines (Figure 8 (b)) [53].

Recent work by Shino and colleagues identified a novel enzyme, endonuclease Q (EndoQ), present in archaea and some bacteria. Unlike EndoV, EndoQ exhibits dual specificity, cleaving immediately 5′ of uracil (dU) or inosine (dI) within both single- and double-stranded DNA. Although its structure and mechanism differ from those of EndoV, the functional relevance of EndoQ in higher eukaryotes remains unclear and requires further investigation [54].



(a)



(b)

Figure 8. (a) Removal and repair of Uracil from DNA by Uracil N glycosylase enzyme via Base excision repair pathway (*denotes AP site); **(b)** Removal of hypoxanthine from DNA by endonuclease V enzyme via alternate excision repair pathway.

6.2. Repair of Oxidative Bases in DNA

The oxidative guanine adduct, 8-oxo-guanine, is repaired through the conserved "GO repair" pathway. In bacteria, this involves MutT, MutM, and MutY, with human homologs MTH1, OGG1, and MUTYH, respectively (Figure 9). MutT hydrolyzes 8-oxo-dGTP to prevent misincorporation. MutM (OGG1) excises 8-oxoG from 8-oxoG:C pairs, initiating BER. If replication occurs before repair, MutY (MUTYH) removes adenine mispaired with 8-oxoG, thereby preventing G:C→T: A mutations [55]. Importantly, defects in OGG1 or MUTYH are strongly associated with genomic instability and have been linked to colorectal and other cancers, underscoring the clinical significance of this pathway.

Repair of oxidized adenine (8-oxo-adenine), on the other hand, is less clear. 8-oxo-A can be removed from mispairs like 8-oxoA: T, 8-oxoA:G, and 8-oxoA:C by human thymine DNA glycosylase (TDG) and E. coli mismatch-specific uracil DNA glycosylase (MUG). Limited activity against 8-oxoA:C has been demonstrated in vitro by other enzymes, such as hOGG1 and NEIL1. But for 8-oxo-A, no specific repair pathway comparable to the GO system has been found. Despite its potential as a mutagen, adenine oxidation is still poorly understood, which highlights the need for more research [48], [56].

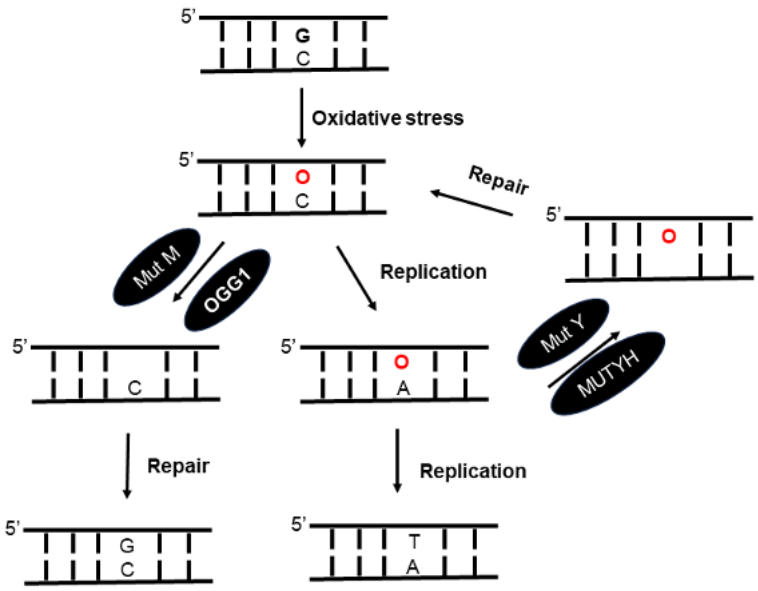


Figure 9. Removal of 8-oxo-guanine from DNA by "GO pathway" in both eukaryotes and prokaryotes

7. Conclusions and Future Perspectives

Deamination and oxidation of DNA bases represent two major sources of base substitution mutations in genomes. While deamination predominantly results in transition substitutions, oxidation frequently gives rise to transversion substitutions. Among these processes, cytosine deamination and guanine oxidation contribute to an overall increase in genomic A+T content, whereas adenine deamination and oxidation contribute towards a decrease in genomic G+C%.

Reactive oxygen species (ROS) frequently damage DNA, particularly guanine. Oxidation of guanine generates 8-oxoG, which tends to mispair with adenine during

replication, resulting in G:C → T:A transversions. Oxidation of adenine, although less well studied, produces 8-oxoA, which can mispair with Guanine and lead to A: T → C: G transversions, thereby contributing to an increase in GC content. In summary, oxidation of guanine cause mutation that favour A+T bias, whereas oxidation of adenine causes mutation that favors G+C bias in the genome. Depending on the frequency of mutational events and the efficiency of DNA repair mechanisms, these processes can significantly influence and help maintain the overall genomic G+C content

Most research on base deamination has focused on cytosine, with comparatively less attention given to adenine and guanine. Similarly, studies on the effects of pH, temperature, and DNA structural context have emphasized cytosine. Cytosine deamination is widely recognized as a major contributor to mutations, primarily driving C→T transitions, but its precise rate across organisms has not been comprehensively reported. This difficulty arises because C→T transitions can also result from tautomerization, repair efficiency, and the activity of enzymes such as DNA cytosine methyltransferases (DCM), making it hard to isolate the contribution of deamination alone. A similar challenge applies to oxidative damage: guanine and adenine oxidation are important sources of G→T and A→C substitutions, yet systematic comparative studies across organisms remain limited.

Although cytosine deamination has been thought of as mutagenic, it also has beneficial biological functions, such as AID-mediated hypermutation, which promotes antibody diversification [57,58]; APOBEC enzymes, which provide antiviral defense [59]; and 5-methylcytosine deamination, which is used in zebrafish to reprogram epigenetics [60]. Similarly, by offering a constant source of genetic variability, base substitution mutations brought about by deamination or oxidative damage support long-term genomic trends. Such substitutions in prokaryotes allow for adaptation to environmental stress [61], the development of antibiotic resistance [62], and antigenic variation that facilitates immune evasion, which in turn propels the evolution of pathogens [63]. Differential deamination and oxidation patterns can therefore be regarded as significant evolutionary forces that connect molecular mutagenesis to more general processes of pathogenicity, adaptation, and genome plasticity.

At the same time, these mutation-driven processes must be understood within the broader framework of genome evolution. While cytosine deamination and guanine oxidation contribute to A+T enrichment, adenine deamination and oxidation can promote G+C enrichment, thereby counteracting this bias. Natural selection additionally influences these mutational inputs, favoring variants that improve stability, efficiency, or adaptability. Therefore, a viewpoint that harmonizes both mutationist and selectionist influences offers the most thorough insight into bacterial genome structure and evolution. Future work should prioritize structural and mechanistic analyses of all amino group–containing bases to elucidate the chemical basis for differential deamination susceptibilities. Such insights would advance our understanding of the intrinsic biases governing base damage and repair, as well as their broader implications for genome stability.

7.1. Diet, Oxidative Stress, and Mutagenesis: A Translational Perspective

Insights from microbial studies on base deamination and oxidation can also inform translational perspectives in humans. Cellular acidity is linked to the Warburg effect, where acidic microenvironments promote altered metabolism and potentially increase mutation rates [64],[65]. Diet has been shown to influence systemic acid–base balance. Bahrami & Greiner

(2021) reported that alkaline-forming foods help maintain higher blood oxygen saturation and may reduce the risks associated with long-term acidosis. Such effects are relevant, since acidic conditions are known to facilitate base deamination and DNA instability.

Alkaline-forming foods include lemons, dates, almonds, spinach, onions, apples, oranges, and green beans, many of which contain mannose, a sugar reported to impair tumor growth in preclinical models [66]. While these associations are promising, it is important to note that causality has not been firmly established, and dietary effects should be viewed as supportive rather than deterministic.

In parallel, oxidative DNA damage caused by reactive oxygen species (ROS) can be counteracted by antioxidants. Numerous studies highlight the role of plant-derived compounds in neutralizing ROS and limiting mutagenic events [67]. Foods rich in antioxidants include broccoli, carrots, tomatoes, legumes, cherries, citrus fruits, garlic, ginger, cloves, cinnamon, saffron, curry leaves, amla, wheatgrass, and soybeans [68]. These diets may complement endogenous defenses such as glutathione, melatonin, and catalase.

Taken together, diets enriched in alkaline-forming and antioxidant-rich foods may contribute to reducing the risks of DNA damage by mitigating acid-induced deamination and ROS-induced oxidation. However, current evidence remains preliminary, and further mechanistic and clinical studies are needed before firm recommendations can be made. This perspective should therefore be regarded as a potential translational direction rather than a direct conclusion.

Multidisciplinary Domains

This research covers the domains: (a) Molecular Biology, (b) Molecular Evolution, (c) Computational Chemistry, and (d) Bioinformatics

Funding

This research received no external funding

Acknowledgment

We are grateful to both anonymous reviewers for their critical comments that helped us to improve the quality of the manuscript significantly. ND is thankful to the DBT, GoI-funded Centre for Bioinformatics and Computational Biology, Tezpur University (BT/PR40253/BTIS/137/52/2022) for the fellowship of Scientific Administrative Assistant. ND is also thankful to the computation facility of CMDR of TU and the NNP grant awarded to SKR and SSS.PKB is grateful to Tezpur University and the DBT project grant (BT/PR40231/BTIS/137/63/2023) for the fellowship. MJ is thankful to Tezpur University and CSIR-HRDG for the fellowship. NA is thankful to Tezpur University for the fellowship.

Conflicts of Interest

The authors declare no conflict of interest.

Declaration on AI Usage

The authors declare that the article has been prepared without the use of AI tools.

References

[1] Lassalle, F.; Périan, S.; Bataillon, T.; Nesme, X.; Duret, L.; Daubin, V. GC-content evolution in bacterial genomes: the biased gene conversion hypothesis expands. *PLoS Genetics*, **2015,** 11(2): e1004941. https://doi.org/10.1371/journal.pgen.1004941

[2] Satapathy, S.S.; Dutta, M.; Ray, S.K. Variable correlation of genome GC% with transfer RNA number as well as with transfer RNA diversity among bacterial groups. *Microbiological Research*, **2010,** 165(3): 232–242. https://doi.org/10.1016/j.micres.2009.06.003

[3] Bansal, K.; Kumar, S.; Kaur, A.; Singh, A.; Patil, P.B. Deep phylo-taxono genomics reveals Xylella as a variant lineage of plant-associated Xanthomonas. *Genomics*, **2021,** 113(6): 3989–4003. https://doi.org/10.1016/j.ygeno.2021.09.019

[4] Guidot, A.; Jiang, W.; Ferdy, J.B.; Thébaud, C.; Barberis, P.; Gouzy, J.; et al. Multihost experimental evolution of Ralstonia solanacearum. *Molecular Biology and Evolution*, **2014,** 31(11): 2913–2928. https://doi.org/10.1093/molbev/msu236

[5] Hildebrand, F.; Meyer, A.; Eyre-Walker, A. Evidence of selection upon genomic GC-content in bacteria. *PLoS Genetics*, **2010, 6**(9): e1001107. https://doi.org/10.1371/journal.pgen.1001107

[6] Chuckran, P.F.; Flagg, C.; Propster, J.; Rutherford, W.A.; Sieradzki, E.T.; Blazewicz, S.J.; et al. Edaphic controls on genome size and GC content. *Soil Biology and Biochemistry*, **2023,** 178: 108935. https://doi.org/10.1016/j.soilbio.2023.108935

[7] Schell, M.A.; Karmirantzou, M.; Snel, B.; Vilanova, D.; Berger, B.; Pessi, G.; et al. The genome sequence of Bifidobacterium longum reflects its adaptation to the human gastrointestinal tract. *Proceedings of the National Academy of Sciences of the USA*, **2002,** 99(22): 14422–14427. https://doi.org/10.1073/pnas.222409599

[8] Kushiro, A.; Shimizu, M.; Tomita, K. Molecular cloning and sequence determination of the tuf gene of Thermus thermophilus. *European Journal of Biochemistry*, **1987,** 170(1–2): 93–98. https://doi.org/10.1111/j.1432-1033.1987.tb13693.x

[9] Kagawa, Y.; Nojima, H.; Nukiwa, N.; Ishizuka, M.; Nakajima, T.; Yasuhara, T.; et al. High GC content in the third codon position of an extreme thermophile. *Journal of Biological Chemistry*,**1984,** 259(5): 2956–2960.

[10] Galtier, N.; Lobry, J.R. Relationships between genomic G+C content, RNA secondary structures, and optimal growth temperature. *Journal of Molecular Evolution*, **1997,** 44(6): 632–636. https://doi.org/10.1007/PL00006186

[11] Nayfach, S.; Shi, Z.J.; Seshadri, R.; Pollard, K.S.; Kyrpides, N.C. New insights from uncultivated genomes of the global human gut microbiome. *Nature*, **2019,** 568(7753): 505–510. https://doi.org/10.1038/s41586-019-1058-x

[12] Sueoka, N. On the genetic basis of variation and heterogeneity of DNA base composition. *Proceedings of the National Academy of Sciences of the USA*, **1962,** 48(4): 582–592. https://doi.org/10.1073/pnas.48.4.582

[13] Xia, X.; Xie, Z.; Li, W.-H. Effects of GC content and mutational pressure on exon lengths. *Journal of Molecular Evolution*, **2003,** 56(3): 362–370. https://doi.org/10.1007/s00239-002-2413-y

[14] Sueoka, N. Directional mutation pressure and neutral molecular evolution. *Proceedings of the National Academy of Sciences of the USA*, **1988,** 85(8): 2653–2657. https://doi.org/10.1073/pnas.85.8.2653

[15] Hershberg, R.; Petrov, D.A. Evidence that mutation is universally biased towards AT in bacteria. *PLoS Genetics*, **2010,** 6(9): e1001115. https://doi.org/10.1371/journal.pgen.1001115

[16] Rocha, E.P.C.; Feil, E.J. Mutational patterns cannot explain genome composition. *PLoS Genetics*, **2010,** 6(9): e1001104. https://doi.org/10.1371/journal.pgen.1001104

[17] Akashi, M.; Yoshikawa, H. Relevance of GC content to the conservation of DNA polymerase III. *Frontiers in Microbiology*, **2013,** 4: 266. https://doi.org/10.3389/fmicb.2013.00266

[18] Graur, D. Molecular and Genome Evolution. Sinauer Associates, Sunderland, MA, USA, **2015**.

[19] Hanson, G.; Coller, J. Codon optimality, bias and usage in translation and mRNA decay. *Nature Reviews Molecular Cell Biology*, **2018,** 19(1): 20–30. https://doi.org/10.1038/nrm.2017.91

[20] Liu, Y. A code within the genetic code: codon usage regulates co-translational protein folding. *Cell Communication and Signaling*, **2020,** 18(1): 145. https://doi.org/10.1186/s12964-020-00642-6

[21] Osawa, S.; Ohama, T.; Yamao, F.; Muto, A.; Jukes, T.H.; Ozeki, H.; et al. Directional mutation pressure and tRNA choice. *Proceedings of the National Academy of Sciences of the USA*, **1988,** 85(4): 1124–1128. https://doi.org/10.1073/pnas.85.4.1124

[22] Abdel-Haleem, H. The origins of genome architecture. *Journal of Heredity*, **2007,** 98(6): 633–634. https://doi.org/10.1093/jhered/esm071

[23] Sankoff, D.; Cedergren, R.J.; Lapalme, G. Frequency of indels and substitutions in 5S rRNA. *Journal of Molecular Evolution*, **1976,** 7(2): 133–149. https://doi.org/10.1007/BF01732261

[24] Beura, P.K.; Sen, P.; Aziz, R.; Satapathy, S.S.; Ray, S.K. Nucleotide polymorphism in intergenic regions. *Journal of Genetics*, **2023,** 102(1): 1–7. https://doi.org/10.1007/s12041-023-01431-4

[25] Vogel, F.; Kopun, M. Higher frequencies of transitions among point mutations. *Journal of Molecular Evolution*, **1977,** 9(2): 159–180. https://doi.org/10.1007/BF01732668

[26] Collins, D.W.; Jukes, T.H. Rates of transition and transversion. *Genomics*, **1994,** 20(3): 386–396. https://doi.org/10.1006/geno.1994.1208

[27] Aziz, R.; Sen, P.; Beura, P.K.; Das, S.; Tula, D.; Dash, M.; et al. Improved estimation of synonymous and nonsynonymous sites. *DNA Research*, **2022,**29(4): dsac023. https://doi.org/10.1093/dnares/dsac023

[28] Kreutzer, D.A.; Essigmann, J.M. Oxidized, deaminated cytosines as a source of C→T transitions. *Proceedings of the National Academy of Sciences of the USA*, **1998**, 95(7): 3578–3582. https://doi.org/10.1073/pnas.95.7.3578

[29] Hollstein, M.; Sidransky, D.; Vogelstein, B.; Harris, C.C. p53 mutations in human cancers. *Science*, **1991,** 253(5015): 49–53. https://doi.org/10.1126/science.1905840

[30] Tan, X.; Grollman, A.P.; Shibutani, S. Mutagenic properties of oxidized DNA lesions. *Carcinogenesis*, **1999,** 20(12): 2287–2292. https://doi.org/10.1093/carcin/20.12.2287

[31] Chatterjee, N.; Walker, G.C. Mechanisms of DNA damage, repair and mutagenesis. *Environmental and Molecular Mutagenesis*, **2017,**58(5): 235–263. https://doi.org/10.1002/em.22087

[32] Andrés, C.M.C.; de la Lastra, J.M.P.; Juan, C.A.; Plou, F.J.; Pérez-Lebeña, E. Oxidative and nitrative DNA modifications. *International Journal of Molecular Sciences*, **2023,** 24(20): 15240. https://doi.org/10.3390/ijms242015240

[33] Coulondre, C.; Miller, J.H.; Farabaugh, P.J.; Gilbert, W. Molecular basis of base substitution hotspots. *Nature*, **1978,** 274(5673): 775–780. https://doi.org/10.1038/274775a0

[34] Parry, T.E. On the mutagenic action of adenine. *Leukemia Research*, **2007,** 31(12): 1621–1624. https://doi.org/10.1016/j.leukres.2007.03.011

[35] Sidorkina, O.; Saparbaev, M.; Laval, J. Nitrous acid-induced mutagenesis. *Mutagenesis*, **1997,** 12(1): 23–28. https://doi.org/10.1093/mutage/12.1.23

[36] Freese, E. Molecular mechanisms of mutations. *Molecular Genetics*, **1963,** 207: 1–14.

[37] Lindahl, T. Instability and decay of DNA. *Nature*, **1993,** 362(6422): 709–

[38] Shapiro, R.; Pohl, S.H. Reaction of ribonucleosides with nitrous acid. *Biochemistry*, **1968,** 7(1): 448–455. https://doi.org/10.1021/bi00842a051

[39] Jung, H.; Hawkins, M.A.; Lee, S. Bypass of deaminated purines by TLS polymerases. *Biochemical Journal*, **2020,** 477(24): 4797–4810. https://doi.org/10.1042/BCJ20200549

[40] Shapiro, R.; Klein, R.S. Deamination of cytidine and cytosine. *Biochemistry*, **1966,** 5(7): 2358–2362. https://doi.org/10.1021/bi00871a041

[41] Gates, K.S. Chemical processes that damage DNA. *Chemical Research in Toxicology*, **2009,** 22(11): 1747–1760. https://doi.org/10.1021/tx900136h

[42] Frederico, L.A.; Kunkel, T.A.; Shaw, B.R. Cytosine deamination in mismatched base pairs. *Biochemistry*, **1993,** 32(26): 6523–6530. https://doi.org/10.1021/bi00078a014

[43] Lindahl, T.; Nyberg, B. Heat-induced cytosine deamination. *Biochemistry*, **1974,** 13(16): 3405–3410. https://doi.org/10.1021/bi00713a035

[44] Kino, K.; Hirao-Suzuki, M.; Morikawa, M.; Sakaga, A.; Miyazawa, H. Guanine oxidation products. *Genes & Environment*, **2017,** 39(1): 21. https://doi.org/10.1186/s41021-017-0084-0

[45] Wood, M.L.; Esteve, A.; Morningstar, M.L.; Kuziemko, G.M.; Essigmann, J.M. Genetic effects of oxidative DNA damage. *Nucleic Acids Research*, **1992,** 20(22): 6023–6032. https://doi.org/10.1093/nar/20.22.6023

[46] Singh, A.; Kukreti, R.; Saso, L.; Kukreti, S. Oxidative stress and guanine tracts. *International Journal of Molecular Sciences*, **2019**, 20(17): 4258. https://doi.org/10.3390/ijms20174258

[47] Cadet, J.; Douki, T.; Gasparutto, D.; Ravanat, J.L. Oxidative damage to DNA. *Mutation Research*, **2003,** 531(1–2): 5–23. https://doi.org/10.1016/j.mrfmmm.2003.09.009

[48] Kruchinin, A.A.; Kamzeeva, P.N.; Zharkov, D.O.; Aralov, A.V.; Makarova, A.V. 8-Oxoadenine in mammals. *International Journal of Molecular Sciences*, **2024,** 25(2): 1342. https://doi.org/10.3390/ijms25021342

[49] Hahm, J.Y.; Park, J.; Jang, E.S.; Chi, S.W. 8-Oxoguanine: from damage to regulation. *Experimental & Molecular Medicine*, **2022,** 54(10): 1626–1642. https://doi.org/10.1038/s12276-022-00862-5

[50] Koag, M.C.; Jung, H.; Lee, S. Mutagenesis mechanism of oxidative adenine lesions. *Nucleic Acids Research*, **2020,** 48(9): 5119–5134. https://doi.org/10.1093/nar/gkaa230

[51] Koag, M.C.; Jung, H.; Lee, S. Mutagenic replication of oxidized adenine. *Journal of the American Chemical Society*, **2019,** 141(11): 4584–4596. https://doi.org/10.1021/jacs.8b12987

[52] Kow, Y.W. Repair of deaminated bases in DNA. *Free Radical Biology and Medicine***, 2002**, 33(7): 886–893. https://doi.org/10.1016/S0891-5849(02)00967-3

[53] Cao, W. Endonuclease V and DNA deamination repair. *Cellular and Molecular Life Sciences*, **2013,** 70(17): 3145–3156. https://doi.org/10.1007/s00018-013-1321-4

[54] Shi, K.; Moeller, N.H.; Banerjee, S.; McCann, J.L.; Carpenter, M.A.; Yin, L.; et al. Recognition of deaminated DNA by endonuclease Q. *Proceedings of the National Academy of Sciences of the USA*, **2021,** 118(10): e2021120118. https://doi.org/10.1073/pnas.2021120118

[55] David, S.S.; O'Shea, V.L.; Kundu, S. Base-excision repair of oxidative DNA damage. *Nature*, **2007,** 447(7147): 941–950. https://doi.org/10.1038/nature05978

[56] Rozelle, A.L.; Cheun, Y.; Vilas, C.K.; Koag, M.C.; Lee, S. DNA interstrand cross-links induced by oxidative adenine lesions. *Nature Communications*, **2021,** 12(1): 1897. https://doi.org/10.1038/s41467-021-22246-0

[57] Muramatsu, M.; Kinoshita, K.; Fagarasan, S.; Yamada, S.; Shinkai, Y.; Honjo, T. Activation-induced cytidine deaminase in hypermutation. *Cell*, **2000,** 102(5): 553–563. https://doi.org/10.1016/S0092-8674(00)00078-7

[58] Di Noia, J.M.; Neuberger, M.S. Molecular mechanisms of antibody somatic hypermutation. *Annual Review of Biochemistry*, **2007,** 76: 1–22. https://doi.org/10.1146/annurev.biochem.76.061705.090740

[59] Harris, R.S.; Dudley, J.P. APOBECs and virus restriction. *Virology*, **2015**, 479–480: 131–145. https://doi.org/10.1016/j.virol.2015.03.012

[60] Rai, K.; Huggins, I.J.; James, S.R.; Karpf, A.R.; Jones, D.A.; Cairns, B.R. DNA demethylation via deamination. *Cell*, **2008,** 135(7): 1201–1212. https://doi.org/10.1016/j.cell.2008.11.042

[61] Andersson, D.I.; Hughes, D. Gene amplification and adaptive evolution. *Annual Review of Genetics*, **2009,** 43: 167–195. https://doi.org/10.1146/annurev-genet-102108-134805

[62] Davies, J.; Davies, D. Origins and evolution of antibiotic resistance. *Microbiology and Molecular Biology Reviews*, **2010,** 74(3): 417–433. https://doi.org/10.1128/MMBR.00016-10

[63] Moxon, R.; Bayliss, C.; Hood, D. Bacterial contingency loci and adaptation. *Annual Review of Genetics*, **2006,** 40: 307–333. https://doi.org/10.1146/annurev.genet.40.110405.090442

[64] Bahrami, H.; Greiner, T. The alkaline diet and the Warburg effect. *World Nutrition*, **2021,** 12: 20–39.

[65] Hamanaka, R.B.; Chandel, N.S. Warburg effect and redox balance. *Science*, **2011,** 334: 1215637. https://doi.org/10.1126/science.1215637

[66] Gonzalez, P.S.; O'Prey, J.; Cardaci, S.; Barthet, V.J.A.; Sakamaki, J.; Beaumatin, F.; et al. Mannose impairs tumour growth. *Nature*, **2018,** 563(7733): 719–723.https://doi.org/10.1038/s41586-018-0729-3

[67] Lobo, V.; Patil, A.; Phatak, A.; Chandra, N. Free radicals, antioxidants and functional foods. *Pharmacognosy Reviews*, **2010,** 4(8): 118–126. https://doi.org/10.4103/0973-7847.70902

[68] Wang, H.; Cao, G.; Prior, R.L. Total antioxidant capacity of fruits. *Journal of Agricultural and Food Chemistry*, **1996,** 44(3): 701–705. https://doi.org/10.1021/jf950579y