Review

Darwin's Abominable Mystery: DNA Replication Timing, Genome Stability and Biodiversity

John Herrick *

Posted Date: 5 August 2025

doi: 10.20944/preprints202506.1612.v2

Keywords: genome stability; DNA repair; DNA replication timing program; species richness; karyotype; mutation rate; transposable elements; non-adaptive radiation

 doi:10.20944/preprints202506.1612.v2

Review

Darwin's Abominable Mystery: DNA Replication Timing, Genome Stability and Biodiversity

John Herrick

independent researcher, 3, rue des Jeûneurs, Paris, France; jhenryherrick@yahoo.fr

Abstract

Both adaptive and non-adaptive theories of evolution have been proposed to explain the process of speciation: how natural selection operates on individuals and populations. Non-adaptive theories emphasize the force of genetic drift in driving speciation while adaptive theories emphasize the force of ecological selection. Both types of theory focus on genetic variation in the organism's genotype, the set of all genes in the genome. The repeatedly observed correlation between amino acid substitution (non-synonymous nucleotide substitutions in codons, dN) and mutation rates (synonymous nucleotide substitutions in genes, dS) has remained something of a mystery since it was first observed and subsequently confirmed in multiple organisms. The following will examine the interaction between the forces of genetic drift and ecological selection in the context of two separate but interacting molecular clocks: the well established gene specific molecular clock and the largely overlooked karyotype specific or "junk" DNA clock.

Keywords: genome stability; DNA repair; DNA replication timing program; species richness; karyotype; mutation rate; transposable elements; non-adaptive radiation

Significance

The eukaryote DNA replication timing program (RT) organizes the DNA synthetic phase (S phase) of the cell cycle and coordinates genome duplication with mitosis and cell division. A complex system of DNA damage detection and repair (DDR) reinforces this organization in order to sustain and constrain mutation/substitution balance. The RT therefore has important implications for the evolution of genome architecture, karyotype diversity and species richness. To date, few studies have directly examined the role RT plays in speciation and adaptive radiations. That question will be addressed in the following.

1. Introduction

1.1.

Karyotypes evolve faster than genes, an observation frequently made since the 1970s. Maxon and Wilson, among others, reported that amphibian karyotypes evolve at several times the rate of the genes residing in the corresponding genomes; a similar observation had also been made in plants [1–3]. Karyotypes are therefore presumably under relaxed selection and evolve more randomly according to genetic drift, while genes are more subject to purifying and positive selection. Consequently, orthologous genes in different species are less divergent than the karyotypes of the respective genomes.

Additionally, the numbers of genes in vertebrates and invertebrates are very similar among different species in the respective groups [4], and evolve on average at the same constant rate. In contrast, karyotypes vary widely from species to species and across clades and lineages. Rates of karyotype evolution also vary widely across the mammalian phylogenetic clades and across the Tree

 doi:10.20944/preprints202506.1612.v2

of Life [5]. This observation is somewhat puzzling given that gene order, or synteny, is highly conserved in Mammalia, Aves and Amphibia [6], indicating that synteny is under strong purifying selection [7].

The conserved average rate of mutations in protein coding exons led to the proposal of a relatively constant molecular clock that governs the rate of gene and genotype evolution [8–10], and therefore the rate at which species diverge (speciation rate) [11,12], a hypothesis that remains to be fully confirmed [13,14]. A paradox emerges from the assumption that genes and genotypes diverge at a constant rate: assuming that macroevolution (species diversification) is linked to microevolution (genetic divergence), how can a relatively constant rate of genotype change explain the extreme differences in species richness (SR) and species evenness (SE) observed across phylogenetic clades and lineages?

1.2.

Ecological selection/speciation, it is commonly assumed, acting differentially and non-randomly on invariant rates of genetic change (standing allelic diversity) can result in environmentally determined differences in selection pressure, and therefore can explain phylogenetic differences in SR and SE [15–20]. A clear latitudinal—and therefore climatic—biodiversity gradient exists, for example, with species richness and biodiversity in biomes increasing from the poles to the equator [21–25]. Other factors including clade age also play determining roles [26–28].

The Tri-cellular model of atmospheric convection currents likewise explains differences in biome biodiversity and geographic location in climatic terms: low levels of precipitation, temperature and insolation are consistently associated with low levels of biodiversity and infertile edaphic conditions, for example, in tundra (permafrost), and in polar and desert biomes. The latitudinal location of biomes can explain SR, but SR nevertheless appears to be unlinked to diversification rates: speciation minus extinction rates [29]. Furthermore, niche rate depends on climatic factors (temperature and humidity) and correlates strongly with SR [30,31].

Other ecological features such as geographic range size (a proxy variable for effective population size, N*e*) [32] and life history traits also correlate with SR, for example r-strategists versus K-strategists and their corresponding phenotypes including body size, metabolic rate, generation time, developmental rate, fecundity and maximum lifespan. Consequently karyotype diversity (KD) and evolution, while contributing to speciation, have long been considered generally lacking in explanatory power regarding species richness and speciation rates [33–35].

1.3.

This assumption has been attributed to the apparent neutral evolution of karyotype diversity ~~KD~~ and related expansions and contractions in neutral, non-coding DNA, resulting in a wide range of genome size in any given phylogenetic lineage (Figure 1). The random expansion and contraction in a lineage's genome size [36,37], such as the salamander lineage [38], led to the proposal of a second "junk" DNA clock, or the linear increase in nuclear DNA with phylogenetic age in salamander families [39]. It should be noted, however, that the rate junk DNA changes in the genome is comparable to the respective mutation rate [40]. Why this should be so—such a widely varying species richness yet a relatively constant molecular clock rate of gene diversification—nevertheless remains an open question and an issue of continuing debate and controversy.

 doi:10.20944/preprints202506.1612.v2

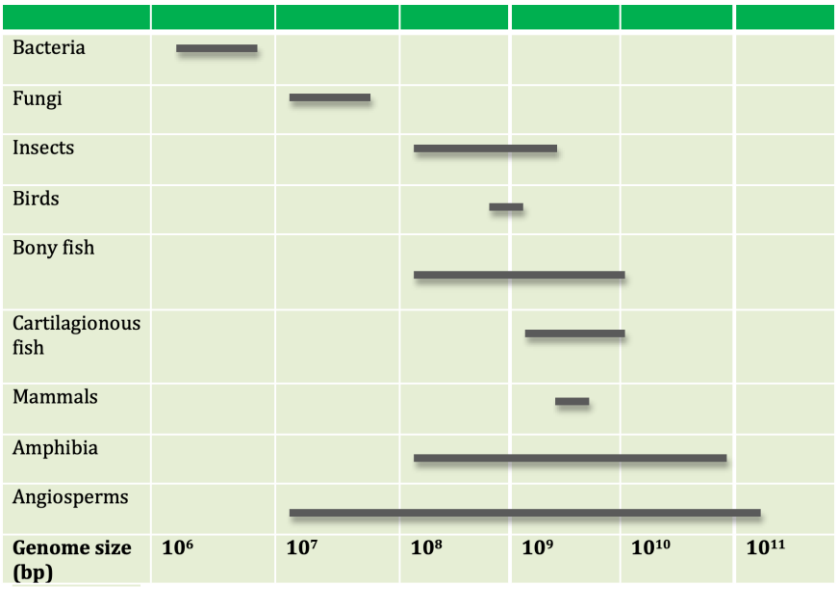

Figure 1. Schematic illustration (not to scale) depicting approximate ranges in genome size in different taxa. Amphibia and angiosperms exhibit a broad range in genome size compared to birds and mammals, resulting from large-scale DNA deletions and amplifications. Under genetic drift, genome size expansion is considered selectively neutral [183], and therefore more variable (larger range). The impact of effective population size (genetic drift) on genome size, however, has recently been called into question. [184].

Because ecological selection acts on physical phenotype and physiology independently of karyotype in any given clade or lineage, it seems unlikely a priori that ecological selection could determine or shape any given species karyotype, or that karyotype diversification could be inherently adaptive [5,41,42]. It remains certainly plausible, however, that changes in karyotype and genome architecture result in the zygotic isolation that ensues from geographic and reproductive isolation (either sympatric or allopatric) [43]; and hence ecological selection would eventually act on genetic drift-driven karyotype diversity over geological time, thus promoting the processes of speciation and adaptive radiation.

The following sequence of events might properly frame the process of ecological succession that characterizes macroevolution (44; Figure 2):

1) geographic isolation following a population split.

2) neutral (non-genic) karyotype diversification driven by genetic drift, eventually involving genes in species with small effective population sizes (microevolution).

3) reproductive (pre and post-zygotic) isolation separating diverged populations (for example, ring species).

4) ecological selection driving speciation and adaptive radiation into newly evolved niches (macroevolution).

 doi:10.20944/preprints202506.1612.v2

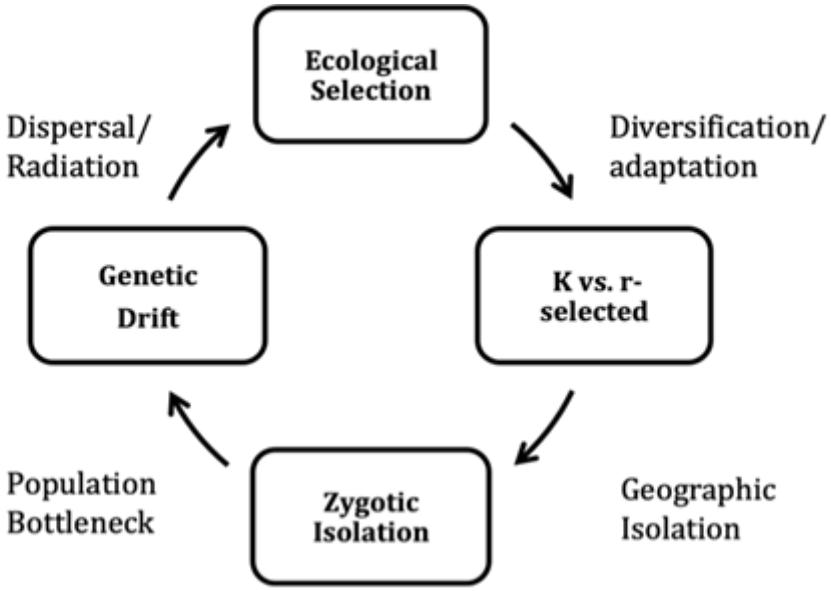

Figure 2. Generalized model of speciation (from Kimura 1991). Ecological selection/speciation (ES) is believed to depend on the standing level of allelic diversity in a population (adaptation/diversification). Smaller effective population sizes are expected to have higher levels of mutation accumulation ("Drift Barrier Hypothesis"; see 185) The figure proposes that genetic drift (GD) can promote an elevated mutation levels that can likewise result in increased levels of allelic diversity (population bottle neck). The interplay between these two forces, genetic drift and ecological selection, determines the extent to which a phenotype or population primarily experiences an adaptive radiation or a non-adaptive radiation. In the absence of strong purifying selection, lineages with high karyotype diversity have a greater probability of being relatively species rich, while species with low karyotype diversity experiencing strong selection can have either low or high levels of species richness. Birds and frogs, for example, have conserved karyotypes but are species rich, whereas salamanders have low karyotype diversity and relatively low levels of species richness. In mammals, karyotype diversity explains about 42 % of species richness (see Table 1).

1.4.

Confounding the predominant role of ES in explaining most speciation events and the highly varying adaptive radiations across clades and lineages is the observation that amino acid substitution rates in proteins are proportional to nucleotide substitution rates in genes: non-synonymous (amino acid changing) substitutions in codons correlate with synonymous (silent) substitutions in genes in all eukaryotes examined [45–47] The ratio between these two is therefore relatively constant (dN/dS proportional to 1). Any detectable deviations from neutrality (dS = dN) are interpreted either as signatures of purifying selection (dN/dS << 1) or positive selection (dN/dS > 1).

The correlation between dN and dS, though lacking a comprehensive molecular explanation, is nonetheless expected if, as commonly assumed, most amino acid substitutions in proteins are either deleterious or functionally neutral (dN and dS both reflect the mutation rate). The correlation, however, is much stronger than would be expected assuming that dN/dS simply reflects the proportion of neutral non-synonymous substitutions, and therefore the underlying mutation rate [48,49]. Moreover, the effect is uniform across all genes in a genome and phylogenetically independent of species relatedness [49–51].

Several hypotheses have been advanced to explain the correlation (see for example: 49, 52). One hypothesis proposes that the correlation is due to a positional, or genomic context effect: because the mutation rate varies across the genome from yeast to plants and animals, any positional, or regional, change in the mutation rate will impact both dN and dS indifferently and simultaneously [53]. Both categories of mutation will be affected in equal proportion: region-wide dN and dS rates will increase

 doi:10.20944/preprints202506.1612.v2

or decrease together regardless of the fact that different genes in the same genome experience significantly differing mutation rates [51].

The effect, the hypothesis predicts, will also apply to non-genic, non-coding DNA residing in the same region as coding DNA. Mutation rates, for example, in ultra-conserved elements (UCE), which are involved in vertebrate development and reside within introns or outside genes [54], correlate with dN and dS inside the exons of the adjacent genes [45]. The two categories of mutation, dN and dS, therefore remain correlated with polymorphisms in non-coding inter-genic regions, and perhaps intra-genic introns [55,56].

The eukaryotic genome is broadly partitioned into two spatial and temporal compartments: early replicating (open) euchromatin (EC) and late replicating (compact) heterochromatin (HC). EC is enriched in GC nucleotides while HC (and facultative HC) is enriched in AT nucleotides. In all species so far examined, mutation rates are significantly higher in genome regions containing HC compared to regions containing EC [57–62]. Consequently, the DNA replication timing (RT) program during DNA synthetic, or S phase, of the cell cycle will simultaneously modulate dN and dS either positively or negatively according to the genome wide variation in mutation rates: relatively low mutation rates in early replicating genes; relatively high mutation rates in late replicating genes. The following will examine the role of RT on mutation rates, genome stability, karyotype diversification and species richness.

2.1. RT and Replication Origins

The eukaryotic DNA replication timing program has been intensively studied over the last several years in yeast and metazoa [63]. Briefly, the RT program corresponds to the timing of the activation of replication origins (start sites of DNA synthesis) during the S phase of the cell cycle [64,65]. In most species, the genetic locations of replication origins are not specified by a conserved DNA sequence. Instead, origin locations depend primarily on chromatin context. Although all origins are “licensed” by an origin recognition complex (ORC and MCM helicases) in late mitotic M phase and post divisional G1 phase, only about 10% of licensed origins are activated during S phase [66]. Activation, or initiation, occurs asynchronously and with increasing density (initiations per kilobase) into mid-S phase (Figure 3), and then decreases as the cells progress toward the G2 phase prior to mitosis [67].

 doi:10.20944/preprints202506.1612.v2

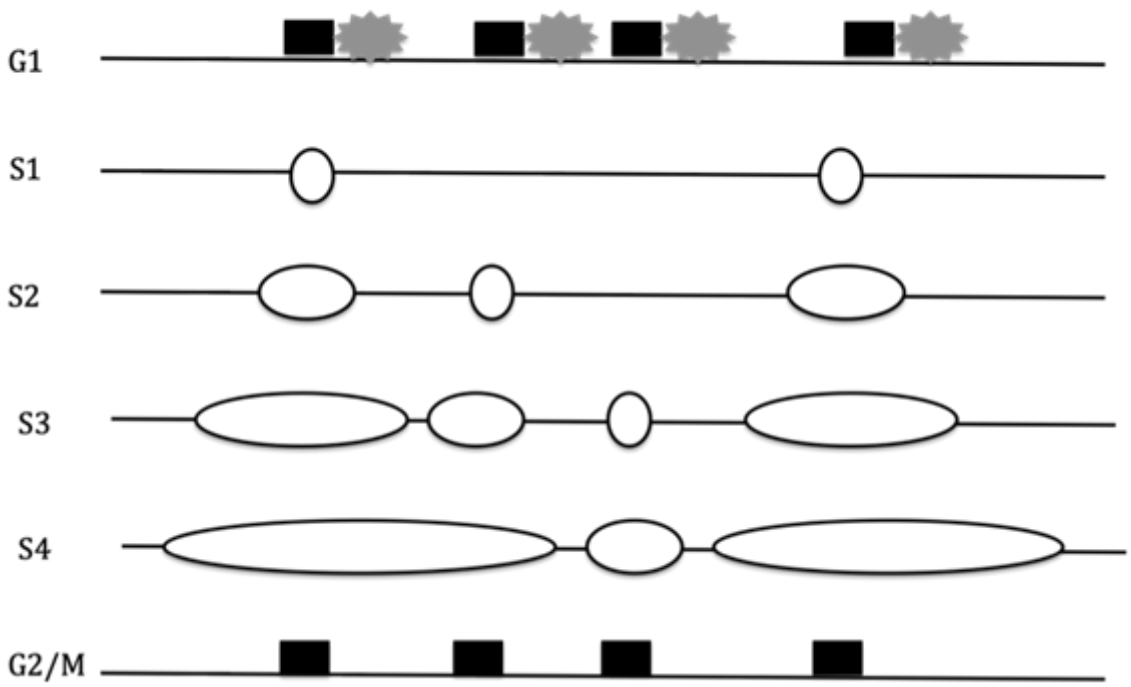

Figure 3. Schematic representation of the replication-timing program. Black boxes: origin recognition complex (ORC). Gray serrated ovals: MCM helicases. Each stage of the cell cycle is represented (G1, S, G2/M). Note that replication fork density increases to mid S-phase (S2 – S3) before decreasing in late S-phase (S4). The figure depicts the binding of ORC during the G2/M phases of the cell cycle. The chromatin bound ORC recruits the MCM helicase complex in G1 of the following cell division cycle. The fully licensed origins in G1 then recruit a complex of initiation factors that assemble the active replication forks that define S-phase.

The timing of origin activation depends on many factors that are associated with two fundamental features: 1) the efficiency, or probability, of origin activation, which is determined by the number of MCM helicases loaded at the origin and, in budding yeast, the levels of a nuclear complex of six positively acting initiation factors [68]; and 2) the strength of the intra-S phase checkpoint, which is a complex of factors that negatively regulates origin activation. Other factors, such as Rif1 in higher eukaryotes and Sir2 and Rpd3 in yeast [69,70], also play a role in establishing a late replication regime. While the biochemistry of these processes is beyond the scope of this paper, it is, in part, the opposing interplay between origin efficiency and the strength of origin inhibition that coordinates origin activation with gene transcription and establishes a late replication-timing program. It might be interesting to note that origins located near or in highly expressed genes, which are early replicating, are more prone to DNA damage [71], while at the same time early replicating regions have substantially lower rates of substitution (72; see below).

3.2. *RT and the Regulation of Gene Transcription*

Genome size in different species is an important modulator of the replication-timing program. The RT program partitions genome duplication into successive time zones and regions (replication domains) that vary in duration (45 to 60 min.) and size (1.3 to 3.6 Mbp), in addition to varying in proportion to C-value (haploid genome size measured in picograms, where one picogram = 978 Mbp) [73–76]. Consequently, the RT program regulates the duration of S phase and the cell division cycle

 doi:10.20944/preprints202506.1612.v2

according to C-value, which therefore has an important influence on generation time and related mutation rates [45].

As mentioned above, a clear relationship has been found between origin firing and gene transcription activity. Transcription start sites (TSS) correlate strongly with origin location, and introns in higher eukaryotes harbor between 40 to 60 % of replication origins depending on the species (77; Figure 4): the earlier an origin fires in the S phase, the higher the level of transcriptional activity and gene expression [78]. Highly expressed genes are for that reason early replicating. In mammals, for example, transcriptional activity is associated with origin efficiency [79–82]. In yeast, experimentally overproducing the six limiting initiation factors causes origins to fire earlier in S phase, and concomitantly increases transcriptional activity of the adjacently located genes [83].

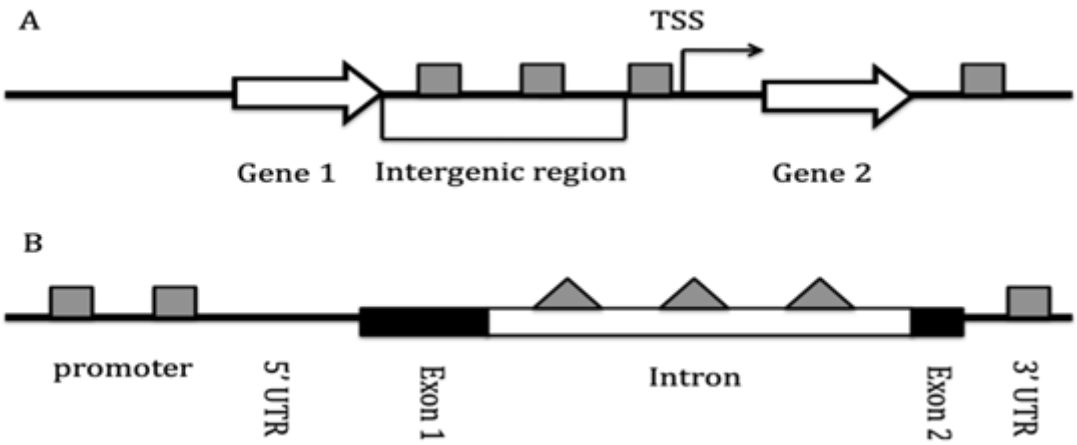

Figure 4. Schematic representation of intron/exon gene structure. **A.** Chromosome level organization. Gray boxes: licensed replication origins. TSS: transcription start site. **B.** Gene level organization. Gray triangles: origins that do not fire during an unperturbed S-phase and are clustered in introns. Figure 4A and B depict the nested structure and coordinated timing of the DNA replication program and gene transcription program (transcriptome): potential DNA replication origins, whether dormant origins or inactive origins, are irregularly spaced along the genome at approximately 15 Kb intervals. Notably, this is approximately the same interval between active origins observed in non-transcribing early stage embryos (prior to the mid-blasula transition). Origins are preferentially located externally to and between genes in part to obviate DNA polymerase-RNA polymerase collisions.

Importantly, the size of introns in genes is also frequently proportional to C-value: genes in larger genomes tend to have larger introns, but not always [84,85]. Intron size, however, can have direct consequences on transcription rates and regulation, with larger genes being transcribed more slowly and expressed generally at lower levels (intron delay) [86,87]. The global result is a longer cell cycle and correspondingly slower developmental rates, and in paedomorphs such as salamanders even fully arrested developmental programs (neoteny) [88,89].

Introns are also generally longer in later replicating genes that have tissue type specific or developmental functions, and consequently late replicating genes tend to have lower expression levels [90]. Many proteins of late replicating genes interface with the environment such as immune system genes and the olfactory gene cluster—so called “adaptive” or “speciation genes”. Later replicating adaptive genes are often essential to organism and phenotype integrity (sensory factors) [53]. Earlier replicating genes that are highly expressed, in contrast, correspond generally to genes essential for cellular function and survival, for example, the house keeping genes that maintain cellular and genetic integrity [91,92]. Interestingly, a negative correlation between protein evolutionary rate (dN) and intron size has been observed in Drosophila [93], consistent with lower mutation/substitution rates in exons (see below).

 doi:10.20944/preprints202506.1612.v2

It should also be noted that older origins in yeast, independently of efficiency, tend to be more frequently lost during the course of speciation and adaptation [94]. More efficient origins, however, are under greater selection pressure, and are conserved in preference to origins of weaker efficiency. Consequently, the distances between origins, rather than their specific genome locations outside of genes, are also under selection pressure to minimize S-phase [95].

The study on replication origin evolution in yeast, however, did not directly address the question of whether or not origin loss (or, conversely, gain) altered levels of transcription in adjacent genes, as is the case when experimentally increasing origin efficiency alters transcription factor promoter binding and stimulates gene expression levels [83]. Importantly, the same study on the evolution of the yeast replication-timing program revealed that the organization of the program evolves in a manner coordinated with protein divergence and chromosomal divergence, but without an apparent causal relationship. In humans, a causal relationship between replication timing and transcription also appears to be more subtle [96].

3.3. RT and Protein Folding Versus Protein Function

Why do essential house keeping genes and other genes that replicate early evolve more slowly? Although still controversial, one compelling explanation is that they have substantially higher expression levels, which places evolutionary constraints on the set of viable amino acid sequences and non-neutral substitutions [72,97]. It has long been assumed that the lower amino acid substitution rates were due to selection acting on the protein's essential functions. That turns out to be the case; but it is selection acting on protein folding rather than on protein function alone that appears to explain, at least in yeast, the high levels of expression in these genes—and hence their biased location in euchromatin.

Selection pressure against protein misfolding is an appealing hypothesis because it could provide another plausible explanation for the correlation between dN and dS. The correlation between gene expression level and dN in these genes is two times stronger than the correlation with dS, indicating selection for translational accuracy (slowing dS) coupled with selection for translational robustness (slowing dN) [72,98]. Selection on protein folding might therefore distinguish early replicating genes from later replicating genes, which are primarily subject to selection on function and adaptive phenotype. Biased selection on the functionality of adaptive and speciation genes thus might explain their enrichment in mutation/substitution prone late replicating DNA, with the fortuitous consequence of higher allelic diversity and polymorphism among more weakly expressed, late replicating developmental genes.

4.1. RT and Genome Stability

Based on these and other findings, it has been suggested that selection pressures have acted on origin-to-origin distances in order to minimize S-phase duration and to regulate transcriptome expression during development. Minimal origin spacing, however, must be limited in order to obviate the highly mutagenic and lethal effects of synchronous origin firing and a surfeit of multiple, simultaneously elongating DNA replication forks [99–101]. Ablating the checkpoint, for example, causes origins to fire earlier and results in massive DNA damage. Conversely, over-expressing initiation-specific proteins causes all origins to fire earlier in budding yeast and is lethal unless ribonucleotide reductase is simultaneously over-produced to supply the forks with sufficient levels of dNTPs [83].

Another related explanation for the evolution of replication timing therefore concerns the intra-S checkpoint and the associated DNA damage detection and repair system (DDR). This feature relates to dNTP supply and to the fact that replication forks are sensitive to dNTP levels, which, when imbalanced or perturbed, are highly mutagenic and a major cause of genome instability (rearrangements, amplifications and deletions, etc.). Hence, replication fork rates determine origin usage under replication stress in all organisms including bacteria. In eukaryotes, so-called

 doi:10.20944/preprints202506.1612.v2

"dormant", or auxiliary, origins are activated in response to perturbed or stalled replication forks. [102,103]

Origin usage and replication fork rates are therefore universally correlated [104]. This is likely to be the case even under non-stressed conditions [105,106], and hence arguably replication fork rates, widely varying across the genome, coordinate the RT program [107]: larger replicons (origin to origin distances) correlate with faster replication fork rates while smaller replicons correlate with slower fork rates. In this manner, the replication-timing program maintains a constant overall rate of replication in domains of differing size, an essential feature of genome stability. At the same time, coordination of fork rates and origin efficiencies serves to limit the rate of DNA damage and maintain mutation/substitution balance across the genome.

4.2. RT and Introns as Adaptations to DNA Damage

The origin of introns, which have multiple effects on gene transcription and protein diversification, has been of longstanding interest in molecular and evolutionary biology. What is their role or effect, however, on DNA replication and genome stability? Introns accumulate in evolutionary conserved genes [108], suggesting that introns might play a role in maintaining gene sequence and structural integrity. Intron gain rates, for example, correlate negatively with sequence evolution rates, while intron loss rates correlate positively [108]. At the same time, changes in C-value during evolution appear to be driving the evolution of intron size toward either longer introns (genome expansion) or shorter introns (genome contraction) [36,109–111].

Under unperturbed conditions most constitutive origins—evolutionary conserved origins—are located in inter-genic regions or are associated with transcription start sites (TSS). When S-phase is advanced experimentally by over expressing certain oncogenes, origins that normally don't fire during S-phase are induced to fire in intra-genic regions—presumably inside introns, where DNA damage is most likely to occur [112] (see Figure 4). At the same time, introns are believed to protect genes against DNA damage [113]. It therefore seems reasonable to assume that introns are enriched in anomalous intra-genic replication origins—normally inactive origins that fire only in the presence of blocked replication forks [77]. If so, longer introns will be expected to harbor proportionally more potential origins. Dormant origins, for example, are closely and regularly spaced (about 15 kb; 114). The dormant origins, as mentioned above, are believed to have evolved to ensure genome stability during replication stress [103,115].

Consequently, stress induced activation of origins within introns might serve to protect genes against DNA damage and lethal DSBs. This could explain why genes in the giant salamander genomes contain introns that are up to 5X the sizes of introns in species with smaller genomes [83], a phenomenon perhaps explained by transposon proliferation [116]: the extra, intra-genic origins are required for both gene and genome stability. Additionally, checkpoint inhibition of these supplementary origins is expected to be correspondingly stronger and mutation rates correspondingly lower. This explanation of the origin of intron size variation, however, remains to be verified.

4.3. RT and DNA Repair

The DDR employs two principal systems that respond to and repair lethal DNA double strand breaks (DSBs): "error free" homologous recombination (HR) and "error prone" non-homologous end joining (NHEJ). HR has a significantly lower mutation rate and primarily (but not exclusively) uses a homologous sister chromatid to repair DSBs. HR therefore is most active during S and G2 phases [117]. NHEJ, which has a comparatively higher mutation rate, operates throughout the cell cycle and progressively replaces HR in the last half of S-G2 phase [118,119]. This might explain why late replicating DNA has relatively higher mutation rates than early replicating DNA (Figure 5), although other factors such as error prone DNA damage polymerases play important roles [57,120].

 doi:10.20944/preprints202506.1612.v2

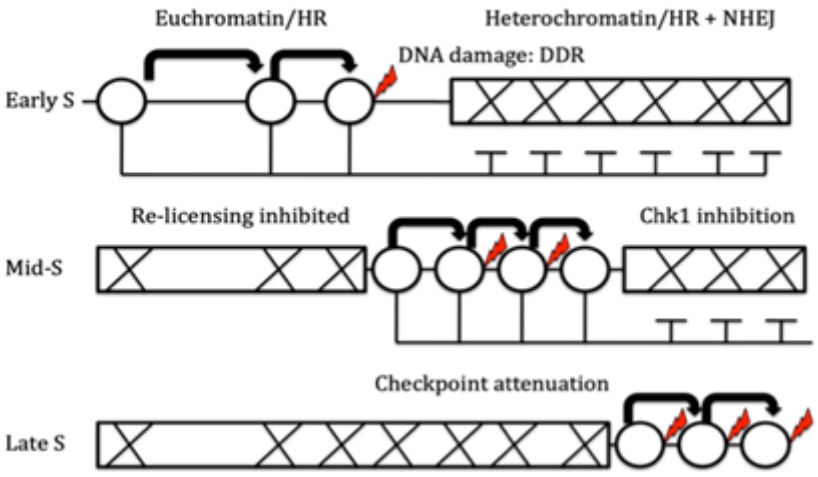

Figure 5. Chromatin organization, DNA replication timing and the DNA damage detection and repair systems (DDR). HR: homologous recombination DNA repair. NHEJ: non-homologous end-joining DNA repair. Circles: active replication origins. Arrows: origin clusters and spacing in a replication domain. Origin activation induces a cascade/domino effect associated with nearby origins firing stochastically. Crosses: inactivated origins. T-bars: checkpoint mediated origin inactivation. Lightning bolts: DNA damage at replication forks. Efficient, frequently firing origins associate with open euchromatin and fire earlier in S-phase; inefficient, infrequently firing origins tend to associate with compact heterochromatin and fire later in S-phase. Origins firing in early S-phase activate the intra-S checkpoint response (via single strand DNA), which in turn stimulates checkpoint factors (eg. Chk1), inhibiting replication initiation at the weaker, late S-phase origins. Once an origin fires and the DNA duplicated, ORC and MCM helicases are inhibited from re-licensing the replicated origins, such that any given origin fires once and only once each cell division cycle. At mid-S phase the checkpoint is attenuated, thus promoting weaker, late origins to fire. This complex system of chromatin organization, DNA replication timing and DNA repair carefully choreographs S-phase and gene expression, and limits mutation and substitution rates before the cell enters mitosis, which would have potentially lethal effects (mitotic catastrophe).

The relative ratios of these two repair systems thus directly impact mutation rates across the genome, and not surprisingly in a genome size-dependent manner: the more mutagenic NHEJ being faster and three times more efficient than homologous recombination repair [117,121]. Eukaryotes with small genomes such as yeast rely predominantly on homologous recombination repair, while species with larger genomes such as vertebrates primarily rely on NHEJ DNA repair [85,122]. Mutation rates are, therefore, expected to be anti-correlated between early and late replicating DNA in a genome size dependent manner: a weaker anti-correlation in small genome species; a stronger anti-correlation in large genome species. This remains to be directly shown, however.

If intron densities increase in more NHEJ-dependent genomes [85], then the strength of the replication-timing program might be expected to increase in parallel [123]. In yeast, for example, origins initiate throughout S-phase (early to late). The vast majority initiates in the first third of S-phase [68,124]. The pattern of origin activation during S-phase, however, varies considerably from cell to cell [125]. Species with larger genomes in contrast have a significantly less flexible and more deterministic replication-timing program [75,123]. In vertebrates, for example, origins that usually fire late rarely if ever fire early in S-phase.

5.1. RT and Genome Evolution

Although in higher eukaryotes the replication-timing program varies little in terms of the sequence of domain activation, the program itself is subject to species and tissue dependent differences in replication timing. A significant proportion in the variation occurs in facultative heterochromatin (fHC) during tissue differentiation and development [126]. The majority of the

 doi:10.20944/preprints202506.1612.v2

replication-timing variants are associated with weak, later firing origins, which have a greater probability of loss during evolution [94,96,127]. Despite the variation between tissue types, DNA replication timing is largely conserved between related species [96,128,129].

Because evolution of replication-timing programs aligns with phylogeny, evolutionary changes in replication timing represent a third class, or source, of molecular evolution and speciation that recapitulates the phylogenetic tree in primates and yeast [94,95,127,130]. Loci that are prone to mutation (sites of higher sequence divergence) such as the human accelerated region (HAR) are biased toward late replication, while sites of highly conserved sequences such as ultra-conserved elements and loss of function intolerant genes replicate early. Divergent loci and HARs are enriched in RT variant regions, which are genomic regions that have experienced an evolutionary change in the replication-timing program [96].

The evolution of the replication-timing program and the related organization of the genome into differential compartments of euchromatin, facultative heterochromatin and constitutive heterochromatin represent adaptations that have the potential to generate and maintain gene polymorphisms and allelic diversity in faster evolving late replicating genes. If so, the RT program might provide a solution of sorts to the puzzling question of why the eukaryote genome has retained rather than eliminated so much and such a variety of non-coding, largely repetitive, "junk" DNA.

The commonly held view of relaxed selection in species with low effective population size, such as salamanders, is increasingly in doubt [131–134]. This came as something of a surprise since genetic drift can play a significant role in karyotype diversity, for example among salamander phylogenetic clades [135]. Genetic drift, however, does not fully explain genetic diversity in salamander genes [136], which varies little among clades and is more subject to selection compared to other vertebrates ($dN/dS > 1$).

The hypothesis that small effective population sizes and genetic drift underlie changes in genome architecture is nonetheless supported by the observation that speciation events have been found to be associated with higher substitution rates [10,136,137]. Relaxed selection in small effective and census population sizes, accordingly, results in respectively higher levels of mutation accumulation and subsequently higher levels of standing genetic diversity on which positive selection can act [98,138,139] This scenario suggests that repeated cycles of drift during population bottlenecks followed by ecological selection act synergistically to drive speciation and rates of species accumulation in phylogenetic clades [44, 140, 141; Figure 2]

5.2. RT and the Correlation between dN and dS

Analogous to the clonal selection theory in immunology, balancing selection on DNA polymorphisms and allelic diversity has acted to multiply the adaptive opportunities and evolutionary trajectories that have led to the emergence of increasingly complex organisms. This has been shown to be the case for certain genes in the immune system-related major histo-compatibility complex (MHC). The MHC replicates in the first half of S phase. The class II elements (AT rich), compared to classes I and III (GC rich), replicate later toward the middle of S phase, when replication fork densities (and hence mutation probabilities) are highest during the cell cycle [67]. Later replication of class II genes might therefore explain the higher genetic diversity in class II genes compared to the genes in classes I and III [51].

A study on salamanders again revealed a correlation between dN and dS but instead with $dN > dS$, indicating selection for diversity [131]. In salamanders, levels of dN/dS are significantly higher than in other vertebrates while levels of dS are, paradoxically, substantially lower, reflecting stronger selection pressure or weaker genetic drift. Ecological selection might therefore "overwhelm" genetic drift in salamander species with low effective population sizes, contrary to expectations. Higher rates of selection on dN might compensate for the lower rates of mutation/substitution (dS), a plausible explanation for the relatively low species richness in most salamander family level clades: ecological selection preserves slowly evolving salamander families against extinction by genetic drift.

 doi:10.20944/preprints202506.1612.v2

In agreement with the proposal of selection for diversity, the MHC replication timing study also revealed a related correlation between dN and the amount of allelic variation. The authors concluded: "increased nucleotide substitution rate can promote allelic variation within lineages" [51]. That conclusion supports the proposal that a position or regional effect on mutation rates can explain the correlation between dN and dS in terms of mutation/substitution balance, because positive selection is not expected to act on dS unless it is acting on *locus specific* mutation rates regardless of either of the forces of genetic drift or selection on amino acid composition—or both ($dN/dS = 1$; $dN/dS \ll 1$ or $dN/dS > 1$). A locus specific, regional explanation for the correlation is consistent with the finding that dN and dS in genes are correlated with polymorphisms in proximal UCEs, the vast majority of which (77 %) are located in intergenic or intronic sequences [54].

6.1. RT and Transposable Elements

Most heterochromatin associated DNA is comprised of repetitive AT-rich transposable elements [142–144]. Since larger genomes are more prone to DNA damage, they must rely on correspondingly stronger checkpoints and more effective DDRs to maintain genome integrity and cell viability. Indeed, it has been shown that the strength of the DDR does in fact increase with genome size, as had been theoretically predicted [145,146].

Moreover, it is believed that the checkpoint evolved initially to combat and suppress the spread of DNA damaging transposable elements in the genome [142], thus fortuitously enhancing, during the course of checkpoint evolution, DNA repair and genome stability. Accordingly, the evolution of metazoa and higher eukaryotes became possible with the evolution of the strength of the intra-S checkpoint and, hence, the effectiveness of the DDR systems [147–149]. Interestingly, transposable element activation induces the checkpoint as a necessary condition for tissue and limb regeneration [150].

What effects have transposable elements had on a potential relationship between genome size evolution (and hence karyotype evolution) and clade diversification (species richness)? Darwin's "abominable mystery" addressed the geologically recent angiosperm radiation, considered the largest radiation in the terrestrial Tree of Life [151]. The topology of the angiosperm phylogenetic tree resembles that of other lineages with highly imbalanced taxonomic clades—similar to salamanders—in terms of karyotype diversity, species richness, species evenness, and, additionally, range of C-value [152, 153; see Figure 1].

The observation of a correlation between dN and dS—whether or not a species (or region of the genome) is undergoing either genetic drift or ecological selection—provides striking evidence for a positional effect influencing mutation rates associated with the replication-timing program: highly expressed, early replicating genes are selected for correct protein folding under a regime of purifying selection, while later replicating and tissue specific/developmental genes are selected for allelic diversity and phenotypic diversification (differentiation and speciation) under a regime of balancing selection. Not surprisingly then most transposable elements are associated with heterochromatin and are therefore generally replicated late.

As mentioned above, mutation/substitution rates vary between folding-selected genes (house keeping) and function-selected genes (adaptation). The relative strength of the checkpoint and effectiveness of the DDR in the respective genomic regions, or replication domains, is therefore expected to vary correspondingly according to the replication-timing program. It has been proposed here that the strength of the intra-S checkpoint can be measured in terms of genome size (C-value), because more origins necessitate stronger checkpoint inhibition of more numerous late activated origins in order to prevent them from competing for dNTPs with earlier firing origins [145,154]. Transposable element driven genome expansion, for example, would lead to a greater probability of fork stalling and DNA damage, and therefore would become maladaptive beyond a threshold C-value and whole body DNA content [148].

Maximum lifespan (MLS), in contrast, can plausibly serve as a proxy variable to measure the relative effectiveness of the DDR. Two observations support that proposal:

 doi:10.20944/preprints202506.1612.v2

1) Peto's paradox addresses the observation that maximum lifespan strongly correlates with body mass (mammal family level clades: adjusted $R^2 = 0.73$; *p = 2 x 10^{-16}; Table 1); yet, unexpectedly, long-lived large body mammals, having more cells and therefore cell division cycles, are significantly less prone to cancer and other mutation-associated disease [155, 156, 157, 158; Figure 6], an observation that might be more related to cell cycle/cell size homeostasis rather than body size [158–163].

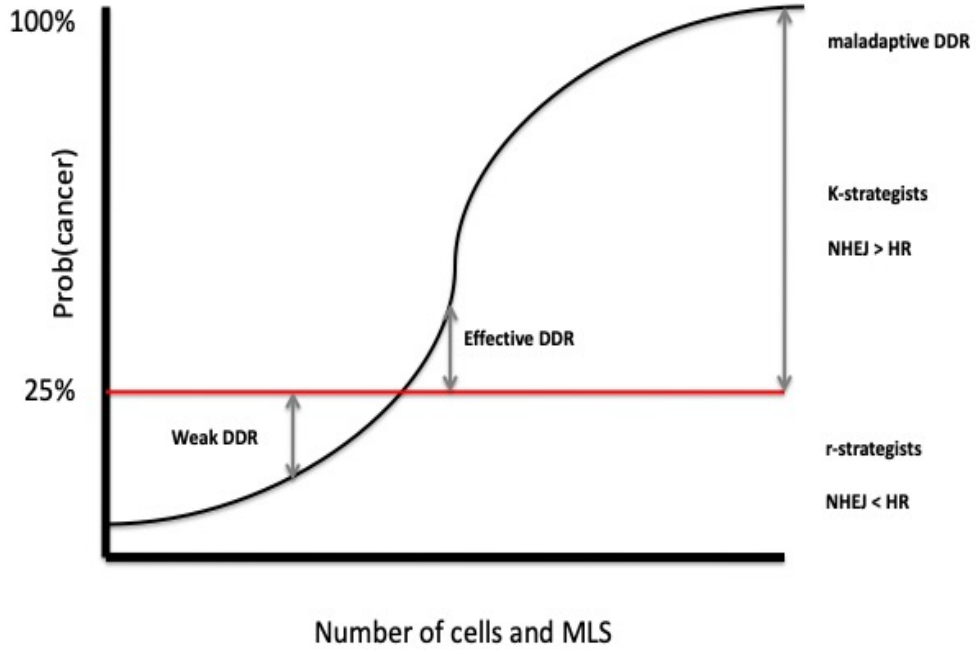

Figure 6. Illustration of Peto's paradox (adapted from 156). Y-axis: cancer risk. X-axis: body mass. Red line: observed cancer risk. Black curve: theoretically expected cancer risk. Gray arrows reflect different strengths of the DDR, from weak DDRs that are maladaptive (small observed body size; higher than expected cancer risk) to strong DDRs that are likewise maladaptive (large observed body size and lower than expected cancer risk, but elevated extinction risk). The intersection of the black and red curves indicates the threshold at which DDR efficiency switches from favoring an unstable karyotype (r-strategist) to favoring an increasingly stable karyotype (K-strategist). The ecological selection pressure in this regime therefore favors larger rather than smaller genome sizes: large genomes (higher TE density, CpG density and HC content; see references 181, 182) are fortuitous TE-driven adaptations that enhance the DDRs and, consequently, promote increases in body mass (see references 173 and 174). Transposable elements, rather than purely selfish parasites, act instead as genome commensals and mutualists benefiting from—and proliferating as a result of—the new niches and ecological adaptations TEs made possible for their respective hosts (Cope's rule: selection for K-strategists with larger genome size, body mass and longer maximum lifespan).

2) Long-lived small body species, such as the Naked Mole Rat, have enhanced NHEJ and other DDR systems, but a normal mammalian C-value/cell size of about 3 pg [164,165]. The proposal made here (MLS is a plausible proxy for the DDR—see: 164, 165—and C-value is a plausible proxy for the strength of the intra-S checkpoint) is at best tentative given that maximum lifespan and C-value are not themselves correlated (adjusted $R^2 = 0{,}007$, *p = 0.5; Table 1).

Table 1. Correlations between mammalian life history traits established using Phylogenetic Generalized Least Squares (PGLS) analysis. MLS: maximum lifespan. SR: species richness. rKD Macro: rate of karyotype evolution (genome wide structural changes: number of chromosomes and number of chromosome arms). rKD Macro reflects the level of karyotype diversity. rKD Micro: rate of sub-chromosomal changes that do not alter the number of chromosomes or the number of chromosome arms. The rKD Macro and rKD Micro data are from Martinez *et al* [5].

 doi:10.20944/preprints202506.1612.v2

	Adj R^2	*P*
MLS vs Body Mass (+)	**0.73**	**2×10^{-16}**
MLS vs C-value (+)	0,007	0,5
SR vs Body Mass (-)	**0.56**	**0.01**
MLS vs Synteny (+)	**0.48**	**0.03**
Synteny vs SR (-)	0.18	0.1
MLS vs SR (+)	**0.59**	**0.016**
rKD Macro vs SR (+)	**0.42**	**3×10^{-10}**
rKD Micro vs SR (+)	0.07	0.06

6.3. Genome Stability and Life History Traits

Although there is no a priori reason to expect a relationship between C-value and maximum lifespan, a yet to be fully investigated direct correlation might exist between the DDR and maximum lifespan, as well as a correlation between the intra-S checkpoint strength and C-value. The question of interest here concerns what specific molecular components of either system are potentially implicated (and how might they be implicated) in the established positive correlation between body mass and maximum lifespan (mammal order level: adjusted $R^2 = 8.0$; *p = 0.0006), and the negative correlation between body mass and species richness (mammal order level: adjusted $R^2 = 0.56$; *p = 0.01); and how might the relationships scale with each other, eg. linearly or as a power law?

Moreover, gestation time has a significant negative relationship with neoplasia and malignancy prevalence, while at the same time neoplasia prevalence and somatic mutation rates are closely associated: species with fewer somatic mutations exhibit lower levels of neoplasia [161]. It is well known, for example, that gestation time (which is related to body size) scales with embryonic growth rate (Table 2), suggesting slower rates of cell growth and division [166–168]. It seems reasonable then to assume that a longer S phase and slower cell cycle would allow more time to repair DNA lesions, and hence serve to enhance genetic integrity and genome stability.

Table 2. Association between gestation time and early embryo cell cycle duration. The parentheses indicate oviparous reproduction.

Table 2 Species	*Gestation time*	*Embryo Cell Cycle Duration*
Drosophila	(24 hours)	8 – 10 minutes
Frog	(6 – 21 days)	0.5 hours
Salamanders	(14) – 728 days	4 – 8 hours
Mouse	19 – 21 days	2 – 4 hours
Rabbit	30 – 32 days	5 – 8 hours
Dog	58 – 69 days	8 – 12 hours
Naked mole rat	66 -77 days	NA
Beaver	105 - 107 days	NA
Human	280 days	12 – 24 hours
Cow	279-292days	32 hours
Elephant	660 days	18 – 36 hours

Of equal interest is the positive correlation in mammals between maximum lifespan and gene synteny conservation (adjusted $R^2 = 0.48$; *p = 0.03). The conservation of synteny blocks over 180 million years of karyotype evolution in mammals [169], for example, is a clear indicator of selection acting on genome stability via physiological and adaptive functions (ecological selection, or macro-evolution impacting micro-evolution). The significant correlation with maximum lifespan, however, suggests that conservation of synteny blocks is also a feature of increased genome stability and a more effective DDR: a stable genotype imbedded in a stable karyotype that is, nevertheless, evolving much faster than the corresponding genotype.

 doi:10.20944/preprints202506.1612.v2

Notably, synteny conservation does not associate significantly with species richness ($R^2 = 0.18$; *p = 0.1), whereas maximum lifespan, in contrast, is significantly associated (negatively) with species richness (order level mammals: adjusted $R^2 = 0.59$; *p = 0.016). Taken together, these observations suggest a role for the DDR—if maximum lifespan does in fact serve as a proxy for the DDR—in enhancing genome stability and in constraining rates of speciation and therefore levels of species richness. It would appear then that evolutionary changes at the sub-cellular level (DDR) promote evolutionary changes at the level of the organism (body mass) and at the level of phylogenetic clades (species richness). This hypothesis, however, warrants further investigation.

7. Conclusions

Bengtsson first proposed in 1980 that unknown "sub-microscopic factors" might account for the correlation between karyotype diversity and species richness in mammals [186]. This observed correlation, however, was phylogenetically unsupported, but has since been confirmed using phylogenetic generalized least squares (PGLS) analysis: at the taxonomic family level in mammals, species richness strongly correlates with rates of karyotype rearrangements (genome scale changes: adjusted $R^2 = 0.42$; *p = 3.3×10^{-10}). The correlation, however, is not observed between rates of sub-chromosomal changes and species richness (adjusted $R^2 = 0.07$; *p = 0.06), and breaks down at lower taxonomic levels. Bengtsson's hypothesis that the imbalance in the Mammalian phylogenetic tree is due to the association between species richness and karyotype diversity—and therefore can be attributed to submicroscopic factors, presumably cellular and nuclear in origin—might also apply to angiosperms and all other metazoans [170–172].

In accordance with Bengtsson's hypothesis, it has been argued here that the "submicroscopic factors" that account for the karyotype diversity-species richness correlation correspond to the close coordination between the replication-timing program, the transcription program, mutation rates and the DDR, with the related interplay between genome stability and instability (mutation/substitution balance) accounting, at least in part, for the dN-dS correlation, karyotype diversity, speciation rates, species richness and species evenness across the Tree of Life. Although ecologically and molecularly independent, micro-evolutionary and macro-evolutionary processes likely interact, albeit indirectly, in mutually establishing speciation rates and species richness [173,174].

Simply stated, hyperactive checkpoints and/or DDRs are expected to result in a long term elevated lineage specific extinction risk (in contrast to a short term species specific risk; see: 175) due to a correspondingly low mutation/substitution supply within the lineage, and consequently a low standing level of genetic and allelic diversity and smaller effective population sizes. If so, an extremely low mutation/substitution rate (approaching zero), with correspondingly low adaptive potential and therefore elevated extinction risk, would likewise impose a ceiling on the evolution of genome sizes, body sizes and their related life history traits (Figure 6). That proposal, however, remains to be established.

A causal relationship between genome/karyotype stability, maximum lifespan and cancer prevalence has yet to be fully demonstrated, but the accumulating evidence is increasingly convincing [176–178]. While substantial evidence supports a role for DNA repair systems in determining maximum lifespan and other life history traits (K-strategists versus r-strategists), the roles DNA damage and repair potentially play in speciation rates can be summarized as follows:

1) genome (in)stability drives genome evolution by either increasing or decreasing rates of karyotype evolution and rates of change in genome size and chromatin organization.
2) genome evolution significantly influences rates of speciation and therefore species richness; for example, by serving as a source of the standing genomic and allelic diversity on which ecological speciation can act.

Assessing the exact roles that gene and karyotype diversifications play in adaptive and non-adaptive radiations remains an important phylogenomic challenge [179,180].

 doi:10.20944/preprints202506.1612.v2

References

1. Bush, G. L., Case, S. M., Wilson, A. C., & Patton, J. L. (1977). Rapid speciation and chromosomal evolution in mammals. *Proceedings of the National Academy of Sciences*, *74*(9), 3942-3946.
2. Maxson, L. E., & Wilson, A. C. (1979). Rates of molecular and chromosomal evolution in salamanders. *Evolution*, 734-740.
3. Levin, D. A., & Wilson, A. C. (1976). Rates of evolution in seed plants: net increase in diversity of chromosome numbers and species numbers through time. *Proceedings of the National Academy of sciences*, *73*(6), 2086-2090. doi: 10.1073/pnas.73.6.2086.
4. Prachumwat, A., & Li, W. H. (2008). Gene number expansion and contraction in vertebrate genomes with respect to invertebrate genomes. *Genome research*, *18*(2), 221-232. doi: 10.1101/gr.7046608.
5. Martinez, P. A., Jacobina, U. P., Fernandes, R. V., Brito, C., Penone, C., Amado, T. F., ... & Bidau, C. J. (2017). A comparative study on karyotypic diversification rate in mammals. *Heredity*, *118*(4), 366-373. doi.org/10.1038/hdy.2016.110
6. Griffin, D. K., Kretschmer, R., Srikulnath, K., Singchat, W., O'Connor, R. E., & Romanov, M. N. (2024). Insights into avian molecular cytogenetics—with reptilian comparisons. *Molecular cytogenetics*, *17*(1), 24. doi.org/10.1186/s13039-024-00696-y
7. Simakov, O., Marlétaz, F., Yue, J. X., O'Connell, B., Jenkins, J., Brandt, A., ... & Rokhsar, D. S. (2020). Deeply conserved synteny resolves early events in vertebrate evolution. *Nature ecology & evolution*, *4*(6), 820-830.
8. Martin, A. P., & Palumbi, S. R. (1993). Body size, metabolic rate, generation time, and the molecular clock. *Proceedings of the National Academy of Sciences*, *90*(9), 4087-4091.doi: 10.1073/pnas.90.9.4087.
9. Sarich, V. M., & Wilson, A. C. (1967). Immunological time scale for hominid evolution. *Science*, *158*(3805), 1200-1203. doi: 10.1126/science.158.3805.1200.
10. Bininda-Emonds, O. R. (2007). Fast genes and slow clades: comparative rates of molecular evolution in mammals. *Evolutionary Bioinformatics*, *3*, 117693430700300008.
11. Hedges, S. B., Marin, J., Suleski, M., Paymer, M., & Kumar, S. (2015). Tree of life reveals clock-like speciation and diversification. *Molecular biology and evolution*, *32*(4), 835-845. doi.org/10.1093/molbev/msv037
12. Lancaster, L. T. (2010). Molecular evolutionary rates predict both extinction and speciation in temperate angiosperm lineages. *BMC evolutionary biology*, *10*, 1-
13. Goldie, X., Lanfear, R., & Bromham, L. (2011). Diversification and the rate of molecular evolution: no evidence of a link in mammals. *BMC evolutionary biology*, *11*, 1-12. doi.org/10.1186/1471-2148-11-286
14. McPeek, M. A., & Brown, J. M. (2007). Clade age and not diversification rate explains species richness among animal taxa. *The American Naturalist*, *169*(4), E97-E106. doi: 10.1086/512135.
15. Rundle, H. D., & Nosil, P. (2005). Ecological speciation. *Ecology letters*, *8*(3), 336-352.
16. Futuyma, D. (2013). The origin of species by means of ecological selection. *Current Biology*, *23*(6), R217-R219.
17. Jablonski, D., & Edie, S. M. (2025). Mass extinctions and their rebounds: a macroevolutionary framework. *Paleobiology*, 1-14.
18. Scholl, J. P., & Wiens, J. J. (2016). Diversification rates and species richness across the Tree of Life. *Proceedings of the Royal Society B: Biological Sciences*, *283*(1838), 20161334. doi: 10.1098/rspb.2016.1334.
19. DeMalach, N., Ke, P. J., & Fukami, T. (2022). The effects of ecological selection on species diversity and trait distribution: predictions and an empirical test. *Ecology*, *103*(1), e03567. doi: 10.1002/ecy.3567.
20. Wellborn, G. A., & Langerhans, R. B. (2015). Ecological opportunity and the adaptive diversification of lineages. *Ecology and evolution*, *5*(1), 176-195. doi: 10.1002/ece3.1347.
21. Hillebrand, H. (2004). On the generality of the latitudinal diversity gradient. *The American Naturalist*, *163*(2), 192-211. doi: 10.1086/381004.
22. Willig, M. R., Kaufman, D. M., & Stevens, R. D. (2003). Latitudinal gradients of biodiversity: pattern, process, scale, and synthesis. *Annual review of ecology, evolution, and systematics*, *34*(1), 273-309.
23. Schemske, D. W., & Mittelbach, G. G. (2017). "Latitudinal gradients in species diversity": reflections on Pianka's 1966 article and a look forward. *The American Naturalist*, *189*(6), 599-603. doi: 10.1086/691719.

 doi:10.20944/preprints202506.1612.v2

24. Bureš, P., Elliott, T. L., Veselý, P., Šmarda, P., Forest, F., Leitch, I. J., ... & Zedek, F. (2024). The global distribution of angiosperm genome size is shaped by climate. *New Phytologist*, *242*(2), 744-759. doi: 10.1111/nph.19544
25. Pianka, E. R. (1966). Latitudinal gradients in species diversity: a review of concepts. *The American Naturalist*, *100*(910), 33-46.
26. Yu, D., & Wiens, J. J. (2024). The causes of species richness patterns among clades. *Proceedings of the Royal Society B*, *291*(2015), 20232436. doi: 10.1098/rspb.2023.2436
27. Wiens, J. J. (2023). Trait-based species richness: Ecology and macroevolution. *Biological Reviews*, *98*(4), 1365-1387. doi.org/10.1111/brv.12957
28. Eiserhardt, W. L., Hansen, L. E. S., Couvreur, T. L., Dransfield, J., Ferreira, P. D. L., Rakotoarinivo, M., ... & Baker, W. J. (2024). Explaining extreme differences in species richness among co-occurring palm clades in Madagascar. *Evolutionary Journal of the Linnean Society*, *3*(1), kzae026. doi.org/10.1093/evolinnean/kzae026
29. Tietje, M., Antonelli, A., Baker, W. J., Govaerts, R., Smith, S. A., & Eiserhardt, W. L. (2022). Global variation in diversification rate and species richness are unlinked in plants. *Proceedings of the National Academy of Sciences*, *119*(27), e2120662119.
30. Castro-Insua, Adrián, Carola Gómez-Rodríguez, John J. Wiens, and Andrés Baselga. "Climatic niche divergence drives patterns of diversification and richness among mammal families." *Scientific Reports* 8, no. 1 (2018): 8781.
31. Title, P. O., & Burns, K. J. (2015). Rates of climatic niche evolution are correlated with species richness in a large and ecologically diverse radiation of songbirds. *Ecology Letters*, *18*(5), 433-440.
32. Smyčka, J., Toszogyova, A., & Storch, D. (2023). The relationship between geographic range size and rates of species diversification. *Nature Communications*, *14*(1), 5559.
33. Kozak, K. H., & Wiens, J. J. (2010). Accelerated rates of climatic-niche evolution underlie rapid species diversification. *Ecology letters*, *13*(11), 1378-1389. doi: 10.1111/j.1461-0248.2010.01530.x.
34. Qiao, H., Peterson, A. T., Myers, C. E., Yang, Q., & Saupe, E. E. (2024). Ecological niche conservatism spurs diversification in response to climate change. *Nature Ecology & Evolution*, *8*(4), 729-738.
35. Kozak, K. H., & Wiens, J. J. (2007). Climatic zonation drives latitudinal variation in speciation mechanisms. *Proceedings of the Royal Society B: Biological Sciences*, *274*(1628), 2995-3003. doi: 10.1098/rspb.2007.1106.
36. Kapusta, A., Suh, A., & Feschotte, C. (2017). Dynamics of genome size evolution in birds and mammals. *Proceedings of the National Academy of Sciences*, *114*(8), E1460-E1469. doi: 10.1073/pnas.1616702114.
37. Elliott, T. A., & Gregory, T. R. (2015). What's in a genome? The C-value enigma and the evolution of eukaryotic genome content. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *370*(1678), 20140331. doi: 10.1098/rstb.2014.0331.
38. Decena-Segarra, L. P., Bizjak-Mali, L., Kladnik, A., Sessions, S. K., & Rovito, S. M. (2020). Miniaturization, genome size, and biological size in a diverse clade of salamanders. *The American Naturalist*, *196*(5), 634-648. doi: 10.1086/711019
39. Martin, C. C., & Gordon, R. (1995). Differentiation trees, a junk DNA molecular clock, and the evolution of neoteny in salamanders. *Journal of Evolutionary Biology*, *8*(3), 339-354. doi.org/10.1046/j.1420-9101.1995.8030339.x
40. Nuzhdin, S. V., & Mackay, T. (1995). The genomic rate of transposable element movement in Drosophila melanogaster. *Molecular Biology and Evolution*, *12*(1), 180-181. doi: 10.1093/oxfordjournals.molbev.a040188.
41. Oliveira, T. D. D., & Freitas, T. R. D. (2023). Investigating the evolutionary dynamics of diploid number variation in Ctenomys (Ctenomyidae, Rodentia). *Genetics and Molecular Biology*, *46*(3 Suppl 1), e20230180. doi: 10.1590/1678-4685-GMB-2023-0180
42. Ferguson-Smith, M. A., & Trifonov, V. (2007). Mammalian karyotype evolution. *Nature Reviews Genetics*, *8*(12), 950-962. doi: 10.1038/nrg2199
43. Escudero, M., Arroyo, J. M., Sánchez-Ramírez, S., & Jordano, P. (2024). Founder events and subsequent genetic bottlenecks underlie karyotype evolution in the Ibero-North African endemic Carex helodes. *Annals of Botany*, *133*(5-6), 871-882. doi: 10.1093/aob/mcad087
44. Kimura, M. (1991). The neutral theory of molecular evolution: a review of recent evidence. *The Japanese Journal of Genetics*, *66*(4), 367-386. doi: 10.1266/jjg.66.367.

 doi:10.20944/preprints202506.1612.v2

45. Saini, A., Usmanova, D. R., & Vitkup, D. (2025). Integrated model of the protein molecular clock across mammalian species. *bioRxiv*, 2025-01. doi.org/10.1101/2025.01.14.633008
46. Wu, C. S., & Chaw, S. M. (2016). Large-scale comparative analysis reveals the mechanisms driving plastomic compaction, reduction, and inversions in conifers II (Cupressophytes). *Genome biology and evolution*, *8*(12), 3740-3750. doi: 10.1093/gbe/evw278.
47. Marais, G., Domazet-Lošo, T., Tautz, D., & Charlesworth, B. (2004). Correlated evolution of synonymous and nonsynonymous sites in Drosophila. *Journal of molecular evolution*, *59*, 771-779. doi: 10.1007/s00239-004-2671-2
48. Wyckoff, G. J., Malcom, C. M., Vallender, E. J., & Lahn, B. T. (2005). A highly unexpected strong correlation between fixation probability of nonsynonymous mutations and mutation rate. *Trends in Genetics*, *21*(7), 381-385. doi: 10.1016/j.tig.2005.05.005.
49. Stoletzki, N., & Eyre-Walker, A. (2011). The positive correlation between d N/d S and d S in mammals is due to runs of adjacent substitutions. *Molecular biology and evolution*, *28*(4), 1371-1380. doi: 10.1093/molbev/msq320.
50. Guisinger, M. M., Kuehl, J. V., Boore, J. L., & Jansen, R. K. (2008). Genome-wide analyses of Geraniaceae plastid DNA reveal unprecedented patterns of increased nucleotide substitutions. *Proceedings of the National Academy of Sciences*, *105*(47), 18424-18429. doi: 10.1073/pnas.0806759105.
51. Garamszegi, L. Z., de Groot, N. G., & Bontrop, R. E. (2009). Correlated evolution of nucleotide substitution rates and allelic variation in Mhc-DRB lineages of primates. *BMC Evolutionary Biology*, *9*, 1-18. doi: 10.1186/1471-2148-9-73.
52. Comeron, J. M., & Kreitman, M. (1998). The correlation between synonymous and nonsynonymous substitutions in Drosophila: mutation, selection or relaxed constraints?. *Genetics*, *150*(2), 767-775. doi: 10.1093/genetics/150.2.767.
53. Chuang, J. H., & Li, H. (2004). Functional bias and spatial organization of genes in mutational hot and cold regions in the human genome. *PLoS biology*, *2*(2), e29.doi: 10.1371/journal.pbio.0020029.
54. Cummins, M., Watson, C., Edwards, R. J., & Mattick, J. S. (2024). The evolution of ultraconserved elements in vertebrates. *Molecular biology and evolution*, *41*(7), msae146. doi.org/10.1093/molbev/msae146
55. Makałowski, W., & Boguski, M. S. (1998). Evolutionary parameters of the transcribed mammalian genome: an analysis of 2,820 orthologous rodent and human sequences. *Proceedings of the National Academy of Sciences*, *95*(16), 9407-9412.doi: 10.1073/pnas.95.16.9407
56. Hodgkinson, A., & Eyre-Walker, A. (2011). Variation in the mutation rate across mammalian genomes. *Nature reviews genetics*, *12*(11), 756-766. doi.org/10.1038/nrg3098
57. Agier, N., & Fischer, G. (2012). The mutational profile of the yeast genome is shaped by replication. *Molecular biology and evolution*, *29*(3), 905-913. doi: 10.1093/molbev/msr280.
58. Chen, C. L., Rappailles, A., Duquenne, L., Huvet, M., Guilbaud, G., Farinelli, L., ... & Thermes, C. (2010). Impact of replication timing on non-CpG and CpG substitution rates in mammalian genomes. *Genome research*, *20*(4), 447-457. doi: 10.1101/gr.098947.109.
59. Weber, C. C., Pink, C. J., & Hurst, L. D. (2012). Late-replicating domains have higher divergence and diversity in Drosophila melanogaster. *Molecular biology and evolution*, *29*(2), 873-882. doi: 10.1093/molbev/msr265.
60. Staunton, P. M., Peters, A. J., & Seoighe, C. (2023). Somatic mutations inferred from RNA-seq data highlight the contribution of replication timing to mutation rate variation in a model plant. *Genetics*, *225*(2), iyad128. doi: 10.1093/genetics/iyad128
61. Stamatoyannopoulos, J. A., Adzhubei, I., Thurman, R. E., Kryukov, G. V., Mirkin, S. M., & Sunyaev, S. R. (2009). Human mutation rate associated with DNA replication timing. *Nature genetics*, *41*(4), 393-395. *Nat Genet* **41**, 393–395 (2009). doi.org/10.1038/ng.363
62. Gaboriaud, J., & Wu, P. Y. J. (2019). Insights into the link between the organization of DNA replication and the mutational landscape. *Genes*, *10*(4), 252. doi.org/10.3390/genes10040252
63. Rhind, N. (2022). DNA replication timing: Biochemical mechanisms and biological significance. *BioEssays*, *44*(11), 2200097. doi: 10.1002/bies.202200097.

 doi:10.20944/preprints202506.1612.v2

64. Boos, D., & Ferreira, P. (2019). Origin firing regulations to control genome replication timing. *Genes*, *10*(3), 199. doi.org/10.3390/genes10030199
65. Bechhoefer, J., & Rhind, N. (2012). Replication timing and its emergence from stochastic processes. *Trends in Genetics*, *28*(8), 374-381. doi: 10.1016/j.tig.2012.03.011
66. Ge, X. Q., & Blow, J. J. (2010). Chk1 inhibits replication factory activation but allows dormant origin firing in existing factories. *Journal of Cell biology*, *191*(7), 1285-1297. doi.org/10.1083/jcb.201007074
67. Goldar, A., Marsolier-Kergoat, M. C., & Hyrien, O. (2009). Universal temporal profile of replication origin activation in eukaryotes. *PLoS One*, *4*(6), e5899. doi.org/10.1371/journal.pone.0005899
68. Yang, S. C. H., Rhind, N., & Bechhoefer, J. (2010). Modeling genome-wide replication kinetics reveals a mechanism for regulation of replication timing. Molecular systems biology, 6(1), 404. doi: 10.1038/msb.2010.61.
69. Yoshida, K., Bacal, J., Desmarais, D., Padioleau, I., Tsaponina, O., Chabes, A., ... & Pasero, P. (2014). The histone deacetylases sir2 and rpd3 act on ribosomal DNA to control the replication program in budding yeast. *Molecular cell*, *54*(4), 691-697. doi: 10.1016/j.molcel.2014.04.032
70. Richards, L., Das, S., & Nordman, J. T. (2022). Rif1-dependent control of replication timing. *Genes*, *13*(3), 550. doi: 10.3390/genes13030550
71. Murat, P., Perez, C., Crisp, A., van Eijk, P., Reed, S. H., Guilbaud, G., & Sale, J. E. (2022). DNA replication initiation shapes the mutational landscape and expression of the human genome. *Science Advances*, *8*(45), eadd3686. doi: 10.1126/sciadv.add3686.
72. Drummond, D. A., Bloom, J. D., Adami, C., Wilke, C. O., & Arnold, F. H. (2005). Why highly expressed proteins evolve slowly. *Proceedings of the National Academy of Sciences*, *102*(40), 14338-14343. doi: 10.1073/pnas.0504070102.
73. Rhind, N. (2006). DNA replication timing: random thoughts about origin firing. *Nature cell biology*, *8*(12), 1313-1316. doi: 10.1038/ncb1206-1313.
74. Dimitrova, D. S., & Gilbert, D. M. (1999). The spatial position and replication timing of chromosomal domains are both established in early G1 phase. *Molecular cell*, *4*(6), 983-993. doi: 10.1016/s1097-2765(00)80227-0.
75. Sima, J., & Gilbert, D. M. (2014). Complex correlations: replication timing and mutational landscapes during cancer and genome evolution. *Current opinion in genetics & development*, *25*, 93-100. doi: 10.1016/j.gde.2013.11.022.
76. Rhind, N., & Gilbert, D. M. (2013). DNA replication timing. *Cold Spring Harbor perspectives in biology*, *5*(8), a010132. doi: 10.1101/cshperspect.a010132.
77. Tian, M., Wang, Z., Su, Z., Shibata, E., Shibata, Y., Dutta, A., & Zang, C. (2024). Integrative analysis of DNA replication origins and ORC-/MCM-binding sites in human cells reveals a lack of overlap. *Elife*, *12*, RP89548. doi: 10.1101/2023.07.25.550556.
78. Gilbert, D. M. (2002). Replication timing and transcriptional control: beyond cause and effect. *Current opinion in cell biology*, *14*(3), 377-383. doi: 10.1016/s0955-0674(02)00326-5.
79. Sequeira-Mendes, J., Díaz-Uriarte, R., Apedaile, A., Huntley, D., Brockdorff, N., & Gómez, M. (2009). Transcription initiation activity sets replication origin efficiency in mammalian cells. *PLoS genetics*, *5*(4), e1000446. doi: 10.1371/journal.pgen.1000446.
80. Müller, C. A., & Nieduszynski, C. A. (2017). DNA replication timing influences gene expression level. *Journal of Cell Biology*, *216*(7), 1907-1914. doi: 10.1083/jcb.201701061
81. Das, S. P., Borrman, T., Liu, V. W., Yang, S. C. H., Bechhoefer, J., & Rhind, N. (2015). Replication timing is regulated by the number of MCMs loaded at origins. *Genome Research*, *25*(12), 1886-1892. doi: 10.1101/gr.195305.115.
82. Liu, Y., Ai, C., Gan, T., Wu, J., Jiang, Y., Liu, X., ... & Hu, J. (2021). Transcription shapes DNA replication initiation to preserve genome integrity. *Genome biology*, *22*(1), 176.doi: 10.1186/s13059-021-02390-3.
83. Santos, M. M., Johnson, M. C., Fiedler, L., & Zegerman, P. (2022). Global early replication disrupts gene expression and chromatin conformation in a single cell cycle. *Genome Biology*, *23*(1), 217. doi.org/10.1186/s13059-022-02788-7

 doi:10.20944/preprints202506.1612.v2

84. Smith, J. J., Putta, S., Zhu, W., Pao, G. M., Verma, I. M., Hunter, T., ... & Voss, S. R. (2009). Genic regions of a large salamander genome contain long introns and novel genes. *BMC genomics*, *10*, 1-11. doi.org/10.1186/1471-2164-10-19
85. Farlow, A., Meduri, E., & Schlötterer, C. (2011). DNA double-strand break repair and the evolution of intron density. *Trends in Genetics*, *27*(1), 1-6. doi: 10.1016/j.tig.2010.10.004.
86. Swinburne, I. A., & Silver, P. A. (2008). Intron delays and transcriptional timing during development. *Developmental cell*, *14*(3), 324-330. doi: 10.1016/j.devcel.2008.02.002.
87. Heyn, P., Kalinka, A. T., Tomancak, P., & Neugebauer, K. M. (2015). Introns and gene expression: cellular constraints, transcriptional regulation, and evolutionary consequences. *Bioessays*, *37*(2), 148-154. doi: 10.1002/bies.201400138.
88. Buffenstein R, Lewis KN, Gibney PA, Narayan V, Grimes KM, Smith M, Lin TD, Brown-Borg HM. Probing pedomorphy and prolonged lifespan in naked mole-rats and dwarf mice. Physiology. 2020 Mar 1;35(2):96-111. doi: 10.1152/physiol.00032.2019.
89. De Groef, B., Grommen, S. V., & Darras, V. M. (2018). Forever young: endocrinology of paedomorphosis in the Mexican axolotl (Ambystoma mexicanum). *General and comparative endocrinology*, *266*, 194-201. doi: 10.1016/j.ygcen.2018.05.016.
90. Versteeg, R., van Schaik, B. D., van Batenburg, M. F., Roos, M., Monajemi, R., Caron, H., ... & van Kampen, A. H. (2003). The human transcriptome map reveals extremes in gene density, intron length, GC content, and repeat pattern for domains of highly and weakly expressed genes. *Genome research*, *13*(9), 1998-2004. doi: 10.1101/gr.1649303.
91. Castillo-Davis, C. I., Mekhedov, S. L., Hartl, D. L., Koonin, E. V., & Kondrashov, F. A. (2002). Selection for short introns in highly expressed genes. *Nature genetics*, *31*(4), 415-418. doi: 10.1038/ng940.
92. Francis, W. R., & Wörheide, G. (2017). Similar ratios of introns to intergenic sequence across animal genomes. *Genome biology and evolution*, *9*(6), 1582-1598.doi: 10.1093/gbe/evx103.
93. Haddrill, P. R., Charlesworth, B., Halligan, D. L., & Andolfatto, P. (2005). Patterns of intron sequence evolution in Drosophila are dependent upon length and GC content. *Genome biology*, *6*, 1-8. doi: 10.1186/gb-2005-6-8-r67
94. Agier, N., Delmas, S., Zhang, Q., Fleiss, A., Jaszczyszyn, Y., Van Dijk, E., ... & Fischer, G. (2018). The evolution of the temporal program of genome replication. *Nature communications*, *9*(1), 2199. doi.org/10.1038/s41467-018-04628-4
95. de Moura, A., & Karschau, J. (2024). Mathematical model for the distribution of DNA replication origins. *Physical Review E*, *110*(3), 034408. doi: 10.1103/PhysRevE.110.034408
96. Bracci, Alexa N., Anissa Dallmann, Qiliang Ding, Melissa J. Hubisz, Madison Caballero, and Amnon Koren. "The evolution of the human DNA replication timing program." *Proceedings of the National Academy of Sciences* 120, no. 10 (2023): e2213896120. doi: 10.1073/pnas.2213896120.
97. Yang, J. R., Liao, B. Y., Zhuang, S. M., & Zhang, J. (2012). Protein misinteraction avoidance causes highly expressed proteins to evolve slowly. *Proceedings of the National Academy of Sciences*, *109*(14), E831-E840. doi: 10.1073/pnas.1117408109.
98. Drummond, D. A., & Wilke, C. O. (2008). Mistranslation-induced protein misfolding as a dominant constraint on coding-sequence evolution. *Cell*, *134*(2), 341-352. doi: 10.1016/j.cell.2008.05.042.
99. Syljuåsen, R. G., Sørensen, C. S., Hansen, L. T., Fugger, K., Lundin, C., Johansson, F., ... & Bartek, J. (2005). Inhibition of human Chk1 causes increased initiation of DNA replication, phosphorylation of ATR targets, and DNA breakage. *Molecular and cellular biology*. doi: 10.1128/MCB.25.9.3553-3562.2005.
100. Petermann, E., Woodcock, M., & Helleday, T. (2010). Chk1 promotes replication fork progression by controlling replication initiation. *Proceedings of the National Academy of Sciences*, *107*(37), 16090-16095. doi: 10.1073/pnas.1005031107.
101. Maya-Mendoza, A., Petermann, E., Gillespie, D. A., Caldecott, K. W., & Jackson, D. A. (2007). Chk1 regulates the density of active replication origins during the vertebrate S phase. *The EMBO journal*, *26*(11), 2719-2731. doi: 10.1038/sj.emboj.7601714.
102. Ge, X. Q., Jackson, D. A., & Blow, J. J. (2007). Dormant origins licensed by excess Mcm2–7 are required for human cells to survive replicative stress. *Genes & development*, *21*(24), 3331-3341. doi: 10.1101/gad.457807.

 doi:10.20944/preprints202506.1612.v2

103. Alver, R. C., Chadha, G. S., & Blow, J. J. (2014). The contribution of dormant origins to genome stability: from cell biology to human genetics. *DNA repair*, *19*, 182-189.doi: 10.1016/j.dnarep.2014.03.012.
104. Poli, J., Tsaponina, O., Crabbé, L., Keszthelyi, A., Pantesco, V., Chabes, A., ... & Pasero, P. (2012). dNTP pools determine fork progression and origin usage under replication stress. *The EMBO journal*, *31*(4), 883-894. doi: 10.1038/emboj.2011.470.
105. Hand, R. (1975). Regulation of DNA replication on subchromosomal units of mammalian cells. *The Journal of cell biology*, *64*(1), 89-97. doi: 10.1083/jcb.64.1.89.
106. Conti, C., Sacca, B., Herrick, J., Lalou, C., Pommier, Y., & Bensimon, A. (2007). Replication fork velocities at adjacent replication origins are coordinately modified during DNA replication in human cells. *Molecular biology of the cell*, *18*(8), 3059-3067.doi: 10.1091/mbc.e06-08-0689
107. Ma, E., Hyrien, O., & Goldar, A. (2012). Do replication forks control late origin firing in Saccharomyces cerevisiae?. *Nucleic acids research*, *40*(5), 2010-2019. doi: 10.1093/nar/gkr982.
108. Carmel, L., Rogozin, I. B., Wolf, Y. I., & Koonin, E. V. (2007). Evolutionarily conserved genes preferentially accumulate introns. *Genome Research*, *17*(7), 1045-1050. doi: 10.1101/gr.5978207
109. Keane, P. A., & Seoighe, C. (2016). Intron length coevolution across mammalian genomes. *Molecular Biology and Evolution*, *33*(10), 2682-2691. doi: 10.1093/molbev/msw151.
110. Blommaert, J. (2020). Genome size evolution: towards new model systems for old questions. *Proceedings of the Royal Society B*, *287*(1933), 20201441. doi: 10.1098/rspb.2020.1441.
111. Sun, C., López Arriaza, J. R., & Mueller, R. L. (2012). Slow DNA loss in the gigantic genomes of salamanders. *Genome biology and evolution*, *4*(12), 1340-1348. doi: 10.1093/gbe/evs103.
112. Macheret, M., & Halazonetis, T. D. (2018). Intragenic origins due to short G1 phases underlie oncogene-induced DNA replication stress. *Nature*, *555*(7694), 112-116.
113. Bonnet, A., Grosso, A. R., Elkaoutari, A., Coleno, E., Presle, A., Sridhara, S. C., ... & Palancade, B. (2017). Introns protect eukaryotic genomes from transcription-associated genetic instability. *Molecular cell*, *67*(4), 608-621. doi: 10.1016/j.molcel.2017.07.002.
114. Ge, X. Q., Han, J., Cheng, E. C., Yamaguchi, S., Shima, N., Thomas, J. L., & Lin, H. (2015). Embryonic stem cells license a high level of dormant origins to protect the genome against replication stress. *Stem Cell Reports*, *5*(2), 185-194. doi: 10.1016/j.stemcr.2015.06.002.
115. Woodward, A. M., Göhler, T., Luciani, M. G., Oehlmann, M., Ge, X., Gartner, A., ... & Blow, J. J. (2006). Excess Mcm2–7 license dormant origins of replication that can be used under conditions of replicative stress. *The Journal of cell biology*, *173*(5), 673-683.doi: 10.1083/jcb.200602108.
116. Sun, C., Shepard, D. B., Chong, R. A., López Arriaza, J., Hall, K., Castoe, T. A., ... & Mueller, R. L. (2012). LTR retrotransposons contribute to genomic gigantism in plethodontid salamanders. *Genome biology and evolution*, *4*(2), 168-183.doi: 10.1093/gbe/evr139.
117. Mao, Z., Bozzella, M., Seluanov, A., & Gorbunova, V. (2008). Comparison of nonhomologous end joining and homologous recombination in human cells. *DNA repair*, *7*(10), 1765-1771. doi: 10.1016/j.dnarep.2008.06.018.
118. Shrivastav, M., De Haro, L. P., & Nickoloff, J. A. (2008). Regulation of DNA double-strand break repair pathway choice. *Cell research*, *18*(1), 134-147. doi.org/10.1038/cr.2007.111
119. Brandsma, I., & van Gent, D. C. (2012). Pathway choice in DNA double strand break repair: observations of a balancing act. *Genome integrity*, *3*, 1-10. doi: 10.1186/2041-9414-3-9
120. Ben Yamin, B., Ahmed-Seghir, S., Tomida, J., Despras, E., Pouvelle, C., Yurchenko, A., ... & Kannouche, P. L. (2021). DNA polymerase zeta contributes to heterochromatin replication to prevent genome instability. *The EMBO journal*, *40*(21), e104543.doi: 10.15252/embj.2020104543
121. Guirouilh-Barbat, J., Huck, S., Bertrand, P., Pirzio, L., Desmaze, C., Sabatier, L., & Lopez, B. S. (2004). Impact of the KU80 pathway on NHEJ-induced genome rearrangements in mammalian cells. *Molecular cell*, *14*(5), 611-623.doi: 10.1016/j.molcel.2004.05.008.
122. Scully, R., Panday, A., Elango, R., & Willis, N. A. (2019). DNA double-strand break repair-pathway choice in somatic mammalian cells. *Nature reviews Molecular cell biology*, *20*(11), 698-714. doi: 10.1038/s41580-019-0152-0.

 doi:10.20944/preprints202506.1612.v2

123. Takahashi, S., Miura, H., Shibata, T., Nagao, K., Okumura, K., Ogata, M., ... & Hiratani, I. (2019). Genome-wide stability of the DNA replication program in single mammalian cells. *Nature genetics*, *51*(3), 529-540.
124. Kelly, T., & Callegari, A. J. (2019). Dynamics of DNA replication in a eukaryotic cell. *Proceedings of the National Academy of Sciences*, *116*(11), 4973-4982.doi: 10.1073/pnas.1818680116.
125. Zhang, Q., Bassetti, F., Gherardi, M., & Lagomarsino, M. C. (2017). Cell-to-cell variability and robustness in S-phase duration from genome replication kinetics. *Nucleic acids research*, *45*(14), 8190-8198. doi: 10.1093/nar/gkx556.
126. Nakatani, T., Schauer, T., Altamirano-Pacheco, L., Klein, K. N., Ettinger, A., Pal, M., ... & Torres-Padilla, M. E. (2024). Emergence of replication timing during early mammalian development. *Nature*, *625*(7994), 401-409. doi.org/10.1038/s41586-023-06872-1
127. Massip, F., Laurent, M., Brossas, C., Fernández-Justel, J. M., Gómez, M., Prioleau, M. N., ... & Picard, F. (2019). Evolution of replication origins in vertebrate genomes: rapid turnover despite selective constraints. *Nucleic acids research*, *47*(10), 5114-5125.doi.org/10.1093/nar/gkz182
128. Ryba, T., Hiratani, I., Lu, J., Itoh, M., Kulik, M., Zhang, J., ... & Gilbert, D. M. (2010). Evolutionarily conserved replication timing profiles predict long-range chromatin interactions and distinguish closely related cell types. *Genome research*, *20*(6), 761-770.doi: 10.1101/gr.099655.109
129. Yaffe, E., Farkash-Amar, S., Polten, A., Yakhini, Z., Tanay, A., & Simon, I. (2010). Comparative analysis of DNA replication timing reveals conserved large-scale chromosomal architecture. *PLoS genetics*, *6*(7), e1001011. doi: 10.1371/journal.pgen.1001011
130. Goldberg, M. E., & Harris, K. (2022). Mutational signatures of replication timing and epigenetic modification persist through the global divergence of mutation spectra across the great ape phylogeny. *Genome Biology and Evolution*, *14*(1), evab104.doi: 10.1093/gbe/evab104.
131. Lu, B. (2023). Evolutionary insights into the relationship of frogs, salamanders, and caecilians ad their adaptive traits, with an emphasis on salamander regeneration and longevity. *Animals*, *13*(22), 3449.
132. Rios-Carlos, H., Segovia-Ramírez, M. G., Fujita, M. K., & Rovito, S. M. (2024). Genomic gigantism is not associated with reduced selection efficiency in neotropical salamanders. *Journal of Molecular Evolution*, *92*(4), 371-380. doi: 10.1007/s00239-024-10177-w.
133. Roddy, A. B., Alvarez-Ponce, D., & Roy, S. W. (2021). Mammals with small populations do not exhibit larger genomes. *Molecular Biology and Evolution*, *38*(9), 3737-3741. doi: 10.1093/molbev/msab142.
134. Mohlhenrich, E. R., & Mueller, R. L. (2016). Genetic drift and mutational hazard in the evolution of salamander genomic gigantism. *Evolution*, *70*(12), 2865-2878.
135. Sessions, S. K. (2008). Evolutionary cytogenetics in salamanders. *Chromosome Research*, *16*, 183-201. doi: 10.1007/s10577-007-1205-3
136. Liedtke, H. C., Gower, D. J., Wilkinson, M., & Gomez-Mestre, I. (2018). Macroevolutionary shift in the size of amphibian genomes and the role of life history and climate. *Nature Ecology & Evolution*, *2*(11), 1792-1799. doi.org/10.1038/s41559-018-0674-4
137. Venditti, C., & Pagel, M. (2010). Speciation as an active force in promoting genetic evolution. *Trends in Ecology & Evolution*, *25*(1), 14-20. doi: 10.1016/j.tree.2009.06.010
138. Lefébure, Tristan, Claire Morvan, Florian Malard, Clémentine François, Lara Konecny-Dupré, Laurent Guéguen, Michèle Weiss-Gayet et al. "Less effective selection leads to larger genomes." *Genome Research* 27, no. 6 (2017): 1016-1028. doi: 10.1101/gr.212589.116.
139. Fuselli, S., Greco, S., Biello, R., Palmitessa, S., Lago, M., Meneghetti, C., ... & Bertorelle, G. (2023). Relaxation of natural selection in the evolution of the giant lungfish genomes. *Molecular biology and evolution*, *40*(9), msad193. doi.org/10.1093/molbev/msad193
140. Lynch, M. (2010). Evolution of the mutation rate. *TRENDS in Genetics*, *26*(8), 345-352. doi: 10.1016/j.tig.2010.05.003.
141. Lam, D. K., Frantz, A. C., Burke, T., Geffen, E., & Sin, S. Y. W. (2023). Both selection and drift drive the spatial pattern of adaptive genetic variation in a wild mammal. Evolution, *77*(1), 221-238. doi: 10.1093/evolut/qpac014.
142. Kabi, M., & Filion, G. J. (2021). Heterochromatin: did H3K9 methylation evolve to tame transposons?. *Genome Biology*, *22*, 1-3. doi: 10.1186/s13059-021-02550-5.

 doi:10.20944/preprints202506.1612.v2

143. Janssen, A., Colmenares, S. U., & Karpen, G. H. (2018). Heterochromatin: guardian of the genome. *Annual review of cell and developmental biology*, *34*(1), 265-288. doi: 10.1146/annurev-cellbio-100617-062653.
144. Wintersberger, E. (2000). Why is there late replication?. *Chromosoma*, *109*(5), 300-307. doi: 10.1007/s004120000080.
145. Herrick, J. (2011). Genetic variation and DNA replication timing, or why is there late replicating DNA?. *Evolution*, *65*(11), 3031-3047. doi: 10.1111/j.1558-5646.2011.01407.x
146. Al Mamun, M., Albergante, L., Moreno, A., Carrington, J. T., Blow, J. J., & Newman, T. J. (2016). Inevitability and containment of replication errors for eukaryotic genome lengths spanning megabase to gigabase. *Proceedings of the National Academy of Sciences*, *113*(39), E5765-E5774. doi: 10.1073/pnas.1603241113.
147. Fortuny, A., & Polo, S. E. (2018). The response to DNA damage in heterochromatin domains. *Chromosoma*, *127*(3), 291-300. doi: 10.1007/s00412-018-0669-6
148. Al Mamun, M., Albergante, L., Blow, J. J., & Newman, T. J. (2020). 3 tera-basepairs as a fundamental limit for robust DNA replication. *Physical Biology*, *17*(4), 046002. doi: 10.1088/1478-3975/ab8c2f
149. Moreno, A., Carrington, J. T., Albergante, L., Al Mamun, M., Haagensen, E. J., Komseli, E. S., ... & Blow, J. J. (2016). Unreplicated DNA remaining from unperturbed S phases passes through mitosis for resolution in daughter cells. *Proceedings of the National Academy of Sciences*, *113*(39), E5757-E5764. doi: 10.1073/pnas.1603252113.
150. Angileri, K. M., Bagia, N. A., & Feschotte, C. (2022). Transposon control as a checkpoint for tissue regeneration. *Development*, *149*(22), dev191957. doi: 10.1242/dev.191957
151. Zuntini, A. R., Carruthers, T., Maurin, O., Bailey, P. C., Leempoel, K., Brewer, G. E., ... & Knapp, S. (2024). Phylogenomics and the rise of the angiosperms. *Nature*, *629*(8013), 843-850. doi.org/10.1038/s41586-024-07324-0
152. Borowska-Zuchowska, N., Senderowicz, M., Trunova, D., & Kolano, B. (2022). Tracing the evolution of the angiosperm genome from the cytogenetic point of view. *Plants*, *11*(6), 784. doi.org/10.3390/plants11060784
153. Schubert, I., & Vu, G. T. (2016). Genome stability and evolution: attempting a holistic view. *Trends in plant science*, *21*(9), 749-757. doi: 10.1016/j.tplants.2016.06.003.
154. Giannattasio, M., & Branzei, D. (2017). S-phase checkpoint regulations that preserve replication and chromosome integrity upon dNTP depletion. *Cellular and Molecular Life Sciences*, *74*(13), 2361-2380. doi: 10.1007/s00018-017-2474-4
155. Voskarides, K., & Giannopoulou, N. (2023). The role of TP53 in adaptation and evolution. *Cells*, *12*(3), 512. doi: 10.3390/cells12030512
156. Roche, B., Hochberg, M. E., Caulin, A. F., Maley, C. C., Gatenby, R. A., Misse, D., & Thomas, F. (2012). Natural resistance to cancers: a Darwinian hypothesis to explain Peto's paradox. *BMC cancer*, *12*, 1-4. doi: 10.1186/1471-2407-12-387.
157. Vincze, O., Colchero, F., Lemaître, J. F., Conde, D. A., Pavard, S., Bieuville, M., ... & Giraudeau, M. (2022). Cancer risk across mammals. *Nature*, *601*(7892), 263-267. doi: 10.1038/s41586-021-04224-5
158. Dart, A. (2022). Peto's paradox put to the test. *Nature Reviews Cancer*, *22*(3), 129-129. doi: 10.1038/s41568-022-00447-4
159. Bowman, J., & Lynch, V. J. (2024). Rapid evolution of genes with anti-cancer functions during the origins of large bodies and cancer resistance in elephants. *bioRxiv*.doi: 10.1101/2024.02.27.582135.
160. Maciak, S. (2022). Cell size, body size and Peto's paradox. *BMC ecology and evolution*, *22*(1), 142. doi: 10.1186/s12862-022-02096-5
161. Compton, Z. T., Mellon, W., Harris, V. K., Rupp, S., Mallo, D., Kapsetaki, S. E., ... & Boddy, A. M. (2025). Cancer prevalence across vertebrates. *Cancer discovery*, *15*(1), 227-244. doi: 10.1158/2159-8290.CD-24-0573.
162. Sulak, M., Fong, L., Mika, K., Chigurupati, S., Yon, L., Mongan, N. P., ... & Lynch, V. J. (2016). TP53 copy number expansion is associated with the evolution of increased body size and an enhanced DNA damage response in elephants. *elife*, *5*, e11994. doi: 10.7554/eLife.11994.
163. Butler, G., Baker, J., Amend, S. R., Pienta, K. J., & Venditti, C. (2025). No evidence for Peto's paradox in terrestrial vertebrates. *Proceedings of the National Academy of Sciences*, *122*(9), e2422861122. doi: 10.1073/pnas.2422861122.

 doi:10.20944/preprints202506.1612.v2

164. Tian, X., Firsanov, D., Zhang, Z., Cheng, Y., Luo, L., Tombline, G., ... & Gorbunova, V. (2019). SIRT6 is responsible for more efficient DNA double-strand break repair in long-lived species. *Cell*, *177*(3), 622-638. doi: 10.1016/j.cell.2019.03.043.
165. Popov, A. A., Petruseva, I. O., & Lavrik, O. I. (2024). Activity of DNA Repair Systems in the Cells of Long-Lived Rodents and Bats. *Biochemistry (Moscow)*, *89*(6), 1014-1023.doi: 10.1134/S0006297924060038
166. Cagan, A., Baez-Ortega, A., Brzozowska, N., Abascal, F., Coorens, T. H., Sanders, M. A., ... & Martincorena, I. (2022). Somatic mutation rates scale with lifespan across mammals. *Nature*, *604*(7906), 517-524. doi: 10.1038/s41586-022-04618-z
167. Johannes, F. (2025). Allometric scaling of somatic mutation and epimutation rates in trees. *Evolution*, *79*(1), 1-5.
168. Singh, V. P., & Singh, P. (2024). Linking DNA damage and senescence to gestation period and lifespan in placental mammals. *Frontiers in Cell and Developmental Biology*, *12*, 1480695. doi: 10.3389/fcell.2024.1480695
169. Damas, J., Corbo, M., Kim, J., Turner-Maier, J., Farré, M., Larkin, D. M., ... & Lewin, H. A. (2022). Evolution of the ancestral mammalian karyotype and syntenic regions. *Proceedings of the National Academy of Sciences*, *119*(40), e2209139119. doi: 10.1073/pnas.2209139119
170. Carta, A., & Escudero, M. (2023). Karyotypic diversity: a neglected trait to explain angiosperm diversification? *Evolution*, *77*(4), 1158-1164. doi:10.1093/evolut/qpad014.
171. Zhan, S. H., Otto, S. P., & Barker, M. S. (2021). Broad variation in rates of polyploidy and dysploidy across flowering plants is correlated with lineage diversification. *bioRxiv*, 2021-03. doi: 10.1101/2021.03.30.436382
172. Zhang, R. G., Liu, H., Shang, H. Y., Shu, H., Liu, D. T., Yang, H., ... & Ma, Y. (2025). Convergent patterns of karyotype evolution underlying karyotype uniformity in conifers. *Advanced Science*, *12*(7), 2411098. doi: 10.1002/advs.202411098.
173. Afonso Silva, A. C., Maliet, O., Aristide, L., Nogués-Bravo, D., Upham, N., Jetz, W., & Morlon, H. (2025). Negative global-scale association between genetic diversity and speciation rates in mammals. *Nature communications*, *16*(1), 1796. doi: 10.1038/s41467-025-56820-y
174. Bergeron, L. A., Besenbacher, S., Zheng, J., Li, P., Bertelsen, M. F., Quintard, B., ... & Zhang, G. (2023). Evolution of the germline mutation rate across vertebrates. *Nature*, *615*(7951), 285-291. doi: 10.1038/s41586-023-05752-y.
175. Futuyma, D. J. (2010). Evolutionary constraint and ecological consequences. *Evolution*, *64*(7), 1865-1884.
176. Schumacher, B., Pothof, J., Vijg, J., & Hoeijmakers, J. H. (2021). The central role of DNA damage in the ageing process. *Nature*, *592*(7856), 695-703. doi: 10.1038/s41586-021-03307-7.
177. Sahm, A., Cherkasov, A., Liu, H., Voronov, D., Siniuk, K., Schwarz, R., ... & Hoffmann, S. (2024). The Greenland shark (Somniosus microcephalus) genome provides insights into extreme longevity. *bioRxiv*, 2024-09. doi: https://doi.org/10.1101/2024.09.09.611499
178. Yang, K., Nishiwaki, K., Mizobata, H., Asakawa, S., Yoshitake, K., Watanabe, Y. Y., ... & Kinoshita, S. (2025). The Greenland shark genome: insights into deep-sea ecology and lifespan extremes. *bioRxiv*, 2025-02. doi: 10.1101/2025.02.19.638963
179. Kilili, H., Padilla-Morales, B., Castillo-Morales, A., Monzón-Sandoval, J., Díaz-Barba, K., Cornejo-Paramo, P., ... & Urrutia, A. O. (2025). Maximum lifespan and brain size in mammals are associated with gene family size expansion related to immune system functions. *Scientific Reports*, *15*(1), 15087. doi: 10.1038/s41598-025-98786-3.
180. Rolland, J., Henao-Diaz, L. F., Doebeli, M., Germain, R., Harmon, L. J., Knowles, L. L., ... & Schluter, D. (2023). Conceptual and empirical bridges between micro-and macroevolution. *Nature Ecology & Evolution*, *7*(8), 1181-1193.
181. Bertucci-Richter, E. M., & Parrott, B. B. (2023). The rate of epigenetic drift scales with maximum lifespan across mammals. *Nature Communications*, *14*(1), 7731.
182. Zhou, W., Liang, G., Molloy, P. L., & Jones, P. A. (2020). DNA methylation enables transposable element-driven genome expansion. *Proceedings of the National Academy of Sciences*, *117*(32), 19359-19366.
183. Lynch, M., & Conery, J. S. (2003). The origins of genome complexity. *science*, *302*(5649), 1401-1404.

 doi:10.20944/preprints202506.1612.v2

184. Marino Alba, Debaecker Gautier, Fiston-Lavier Anna-Sophie, Haudry Annabelle, Nabholz Benoit (2024) Effective population size does not explain long-term variation in genome size and transposable element content in animals eLife 13:RP100574. https://doi.org/10.7554/eLife.100574.1
185. Lynch, M., Ackerman, M. S., Gout, J. F., Long, H., Sung, W., Thomas, W. K., & Foster, P. L. (2016). Genetic drift, selection and the evolution of the mutation rate. *Nature Reviews Genetics*, *17*(11), 704-714.
186. Bengtsson, B. O. (1980). Rates of karyotype evolution in placental mammals. *Hereditas*, *92*(1), 37-47.