Evol Biol (2007) 34:144–151
DOI 10.1007/s11692-007-9010-7

ESSAY

Understanding Neutral Genomic Molecular Clocks

Soojin V. Yi

Received: 24 July 2007 / Accepted: 27 August 2007 / Published online: 13 October 2007
© Springer Science+Business Media, LLC 2007

Abstract The molecular clock hypothesis is a central concept in molecular evolution and has inspired much research into why evolutionary rates vary between and within genomes. In the age of modern comparative genomics, understanding the neutral genomic molecular clock occupies a critical place. It has been demonstrated that molecular clocks run differently between closely related species, and generation time is an important determinant of lineage specific molecular clocks. Moreover, it has been repeatedly shown that regional molecular clocks vary even within a genome, which should be taken into account when measuring evolutionary constraint of specific genomic regions. With the availability of a large amount of genomic sequence data, new insights into the patterns and causes of variation in molecular clocks are emerging. In particular, factors such as nucleotide composition, molecular origins of mutations, weak selection and recombination rates are important determinants of neutral genomic molecular clocks.

Keywords Molecular clock · Generation time effect · Regional heterogeneity

Introduction

The idea that evolutionary rates are constant among lineages, the so-called "constant molecular clock" hypothesis (Margoliash 1963; Zuckerkandl and Pauling 1962), has held enormous promise for evolutionary studies. If indeed rates were constant among lineages, inferring the timing of evolutionary events could be reduced to a straightforward problem of comparing evolutionary distances among lineages, using protein or DNA sequence data.

After over four decades of study, alas, the idea of a 'global' molecular clock, namely, constant evolutionary rates among distantly related species, has now been repeatedly refuted (Bromham and Penny 2003; Kumar 2005; Li 1997). However, the molecular clock hypothesis has had a strong influence on the field of molecular evolution, by (i) motivating data collection, (ii) instigating theoretical advances for understanding rate variation of mutations and fixation, and (iii) bringing methodological advances in molecular phylogenetics.

The arrival of the modern genomics era has undoubtedly affected many issues in molecular evolution, including the molecular clock hypothesis. In particular, understanding the 'background' *neutral* rate at which genomic sequences evolve has become an important topic in comparative genomics: one of the main goals of modern comparative genomics is to identify hidden, functional regions of different genomes, especially of the human genome. A widely used approach is to compare the molecular clocks of different genomic regions. A region evolving at a significantly slower rate than expected under neutrality is considered to be under purifying selection (Dermitzakis et al. 2002; Keightley et al. 2005). Conversely, fastly evolving regions may be under positive selection for their adaptive significance (Pollard et al. 2006). Therefore, to confidently infer natural selection on specific genomic regions, it is critical to understand how neutral evolutionary rates vary within and between different genomes.

In this review, we discuss several historical issues related to the molecular clock hypothesis that have been

S. V. Yi (✉)
School of Biology, Georgia Institute of Technology,
310 Ferst Drive, Atlanta, GA 30332, USA
e-mail: soojinyi@gatech.edu

re-examined under the new lights of modern genomics, as well as some new ideas emerging from the analysis of abundant genomic data. The focus of this article is on neutral molecular clocks in the absence of natural selection. For the sake of brevity, we will limit discussions to issues related to mammalian molecular clocks, which has been a topic of intensive investigation.

How to Study Neutral Molecular Clocks?

The neutral theory of molecular evolution (Kimura 1983) posits that the rate of substitution (evolution) at a neutral site is equal to the rate of mutation. Thus, molecular clocks at neutral sites will run at the rate of mutations. For this reason, we can compare neutral molecular clocks among different lineages to understand patterns and causes of lineage-specific mutation rates.

A constant molecular clock across diverse lineages implies that mutation rates are constant over time. Similarly, we can study molecular clocks of different regions in a genome to investigate regional heterogeneity of mutation rates.

While the neutral theory provides a theoretical underpinning for the study of the molecular clock hypothesis, finding sites that can be classified as 'neutral' has been a challenge. Earlier studies considered synonymous sites (sites that do not change resulting amino acids following substitution) as neutral. In particular, the positions at which any nucleotide substitution can occur without changing the corresponding amino acid, the so-called 'fourfold degenerate sites', have been widely used to infer neutral molecular clocks (for example, Kumar and Subramanian 2002). With the advance of DNA sequencing technology, studies began to use sites that do not encode proteins, known as 'non-coding' sites, including introns (Castresana 2002; Shimmin et al. 1993) and intergenic regions (Chen and Li 2001; Yi et al. 2002).

In the last few years, we have witnessed an explosive accumulation of genome-scale sequence data from mammalian species. As of July 2007, whole genome assemblies of 21 mammalian species are available in public databases (cf. http://www.ensembl.org). In addition, many BAC-derived genomic sequences of other species that are not yet slated toward complete genome sequencing are also available. Since mammalian genomes mostly consist of non-coding regions, the availability of large amounts of genomic data has greatly relived the problem of having to collect data for analyses of neutral molecular clocks. Utilizing newly available data, some historical problems in the molecular clock hypothesis have recently been re-visited.

Molecular Clock Variation Among Lineages: Generation Time Effect

We will first discuss the so-called generation time effect. The original molecular clock hypothesis, which posits that evolutionary rates are constant over absolute time, has instigated much concern from geneticists, because geneticists are used to measuring genetic and evolutionary events using generations. More specifically, the observation that rates are constant over time is incompatible with the presumed molecular origins of germline mutations (namely, errors during DNA replication). Since the number of DNA replication events per unit time is proportional to the number of generations, it is predicted that evolutionary rates should be related to the number of generations per time, rather than the absolute time. Furthermore, one can predict that species with more generations per unit time (i.e., with shorter generation times) should accumulate more mutations than species with fewer generations in the same amount of time (i.e., longer generation times). This prediction is referred to as the 'generation-time effect' (Kohne 1970; Laird et al. 1969).

Comparison of Molecular Clocks in Rodents and Primates

Tests of the generation time effect hypothesis have led to controversial results. One of the first studies that used DNA sequence comparison to investigate the generation time effect is Wu and Li (1985). They used sequence data from 11 genes in mouse and human to compare evolutionary rates in primates and rodents. Generation time effects predict that the rodent lineage should harbor more substitutions, because they have much shorter generation time. To investigate this, they used a relative rate test, which compares evolutionary rates of two lineages without the usage of independent calibration (i.e., fossil records) by employing an outgroup sequence that has diverged before the split of the two lineages considered (see Fig. 1A). Alternatively, one can use a likelihood-ratio test to compare evolutionary rates of different lineages (Fig. 1B).

Wu and Li (1985) found that evolutionary rates were faster in the rodent lineage than in the primate lineage, and the pattern was strongest in synonymous sites and untranslated regions. Because synonymous sites and untranslated regions should be free from natural selection relative to nonsynonymous sites, their conclusion provides a strong support of generation time effect.

One caveat of Wu and Li (1985) was that the choice of carnivores as the outgroup to rodents and primates is controversial (e.g., see Cannarozzi et al. 2007; Murphy

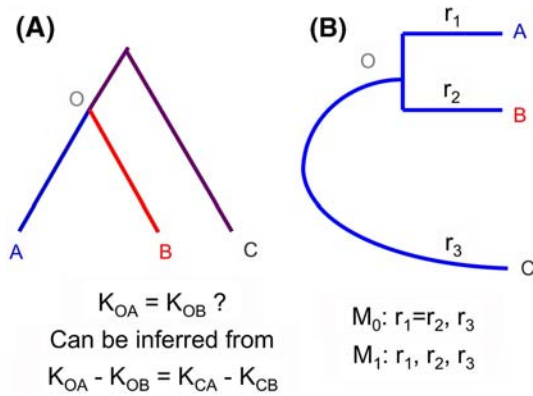

Fig. 1 Methods to compare molecular clocks between lineages A and B using an outgroup C. (**A**) Relative rate test. The branch lengths of OA and OB correspond to lineage-specific numbers of substitutions since their divergence form the common ancestor O. Because the sequence of the common ancestor O is unknown, the difference between OA and OB will be inferred from the difference between their divergence from an outgroup C. (**B**) Likelihood ratio test. Competing hypotheses M_0 (identical rates on lineages A and B) and M_1 (different rates on lineages A and B) are compared by considering the ratio of the maximum likelihood under the two hypotheses

et al. 2001, for conflicting conclusions for the place of carnivores and rodents in mammalian phylogeny). Nevertheless, subsequent studies generally supported a faster rate of evolution of the rodent lineage compared to the primate lineage (Gu and Li 1992; Huttley et al. 2007), although the degree of rate difference varies between studies.

Primate Genomic Molecular Clocks and Generation Time Effect

Another well-known example of the generation time effect is the phenomenon of 'hominoid rate slowdown' (Goodman 1961, 1962). Since humans and other apes have markedly longer generation time than Old World monkeys, it is proposed that humans and apes should show a slower rate of evolution. Earlier studies supported this hypothesis (Goodman 1962; Li et al. 1987) but some of these studies were criticized due to errors in calibration or bias in their data (Easteal 1991; Sarich and Wilson 1967).

Recent analyses employing genomic sequence data have largely supported the hominoid rate slowdown hypothesis (Kim et al. 2006; Steiper and Young 2006; Steiper et al. 2004; Yi et al. 2002). Furthermore, New World monkeys, which have shorter generation times than hominoids and Old World monkeys, show even faster molecular clocks (Steiper and Young 2006). Therefore, at least within primates, evolutionary rate variation among lineages follows the predictions of the generation time effect hypothesis.

Genomic Analyses of Mammalian Molecular Clocks

While the above example of primates exhibits an excellent relationship between evolutionary rates and generation times, comparisons of more distantly related species have shown weaker generation time effect. Kumar and Subramanian (2002) analyzed a large amount of fourfold degenerate sites from a variety of mammals, and concluded that the rate difference between rodents and primates is approximately similar to the rate difference between humans and Old World monkeys. Because the generation time difference between rodents and primates is much greater than that between humans and Old World monkeys, this observation suggests that the generation time effect is not a strong determinant of evolutionary rates between distantly related species. Notably, Huttley et al. (2007) recently analyzed genomic sequence data alignments between opossum, human, mouse, rat and dog and determined that the rate difference between human and mouse is ~11–14%, and that eutherian lineages altogether are evolving approximately 30% slower than the opossum lineage. These comparisons suggest that replication errors may not be the sole source of mutations in these species and/or that rates of replication errors vary greatly between distantly related species.

Variation of Molecular Clocks Among Different Regions within a Genome

Molecular clocks also vary substantially among different genomic regions (Filipski 1988; Wolfe et al. 1989). Regional variation of molecular clocks includes variation within chromosomes as well as between sex chromosomes and autosomes. The X chromosome tends to show lower evolutionary divergence than autosomes, while the Y chromosome typically exhibits the highest divergence within the genome. This difference between sex chromosomes and autosomes may be caused by the same mechanism as the generation time effect, as explained in the following section.

Different Rates Between the X Chromosome and Autosomes

Sex chromosomes and autosomes spend different amounts of time in males and females. While the Y chromosome only inhabits the male germline, the X chromosome and autosomes spend 1/3 and half of their times in the male germline, respectively. In mammals, males typically undergo a greater number of germ cell divisions, hence more replications. According to the generation time effect

hypothesis, the number of mutations is proportional to the number of replication events. Therefore, the generation time effect hypothesis predicts higher evolutionary rates for the Y chromosome than autosomes, and slower rate for the X chromosome. This phenomenon is referred to as 'male-driven evolution' (Li et al. 2002; Shimmin et al. 1993).

Evolutionary analyses using primate introns generally supported male-driven evolution (Nachman and Crowell 2000; Shimmin et al. 1993). Interestingly, the degree of male-driven evolution is weaker in rodents, in which the difference of generation times between males and females is smaller than that in primates (Chang et al. 1994; Li et al. 1996). Male-driven evolution was also observed in birds (Ellegren and Fridolfsson 1997), which have female heterogamety (females have heterologous sex chromosomes, unlike the X/Y chromosomal system in mammals). These observations suggest that the difference in the molecular clocks of the sex chromosome and autosomes is consistent with the generation time effect hypothesis.

Genomic Analysis of Male-driven Evolution

With the advance of DNA sequencing technology, the question of male-driven evolution has been re-visited, using a large amount of non-coding, intergenic sequence data. In contrast to earlier results, Bohossian et al. (2000) found that male-driven evolution in the human genome is much weaker than previously proposed. As a response, Makova and Li (2002) analyzed over 10 kbs of non-coding intergenic DNA in several primates and found that the degree of male driven evolution is much stronger than reported in Bohossian et al. (2000), and that between closely related species such as humans and chimpanzees, the presence of ancestral polymorphism can obscure evolutionary effects. Interestingly, Patterson et al. (2006) recently analyzed a large amount of sequences from human, gorilla, orangutan and macaque and suggested that the rate difference between the X chromosome and autosomes in humans is more exaggerated than that in gorillas. Therefore, the degree of male driven evolution in different lineages needs to be further examined in a variety of taxa.

Causes of Regional Variation of Molecular Clocks

Rates also vary between different regions within a chromosome, and between different autosomes. Even though such 'regional variability' of molecular clocks has been recognized for several decades (Filipski 1988; Wolfe et al. 1989), what causes such variability is poorly understood. In this section, we will discuss our current knowledge on two important determinants of neutral genomic molecular clocks, GC content and recombination.

It should be noted that there are other potential factors of neutral genomic molecular clocks. For example, an earlier hypothesis proposed that differences in the efficacy of DNA repair across the genome causes regional variability of evolutionary rates (Britten 1986). If the fidelity of DNA repair system varies between different genomic regions, those with high fidelity would show a slower molecular clock than the regions with a less efficient repair system. Even though this hypothesis is an appealing explanation for the regional variability of genomic molecular clocks, a direct test of this hypothesis on a genome scale remains as a challenge.

Availability of large-scale comparative genomic data and genomic linkage maps has allowed detailed investigation of the roles of GC content and recombination on regional variability of neutral molecular clocks. In particular, the role of recombination on neutral mammalian molecular clock has received considerable interest in the last few years.

Perry and Ashworth (1999) reported a dramatic example of the role of recombination on molecular clock. They analyzed molecular evolution of the F_{XY} gene in several mammalian species. This gene straddles the pseudoautosomal boundary in the domestic mouse *Mus musculus* so that the 3′ end of the F_{XY} gene is located in the highly recombining pseudoautosomal regions, and the 5′ end is X-linked. In the outgroup species, including in *M. spretus* which has diverged from *M. musculus* within the last few million years, the F_{XY} gene is entirely X-linked. Despite the short evolutionary time since its relocation, the 3′ end of the F_{XY} gene has experienced a remarkably fast molecular clock compared to the 5′ end. Perry and Ashworth (1999) estimated that the synonymous substitution rate at the 3′ end of the F_{XY} gene has increased over 170-fold compared to the 5′ end. This observation led to the conclusion that recombination could directly affect neutral molecular clock. Soon after, several other studies have found that both within-species nucleotide polymorphism and between-species divergence are strongly correlated with the local recombination rates in the human genome (Fullerton et al. 2001; Hellmann et al. 2003; Lercher and Hurst 2002).

While Perry and Ashworth (1999) proposed that recombination itself is mutagenetic, detailed analyses of human and mouse pseudoautosomal boundary region cast some doubt on this proposal (Huang et al. 2005; Yi et al. 2004). Interestingly, recombination and GC content are correlated in a variety of taxa (Birdsell 2002; Jensen-Seaman et al. 2004; Kong et al. 2002; Meunier and Duret 2004). One explanation for the relationship between GC content and recombination is that when mismatch repair

between GC and AT alleles occurs, recombination favors the spread of GC-alleles (i.e., there is a fixation bias toward GC alleles) via biased gene conversion (Galtier et al. 2001). The substitution patterns of the above cited example of the F_{XY} gene also fits the biased gene conversion hypothesis (Montoya-Burgos et al. 2003).

The analysis of the human and chimpanzee genomes revealed that the effects of GC content and recombination on regional molecular clock are complicated and vary between different regions (Chimpanzee Sequencing and Analysis Consortium 2005). In highly recombining regions (where recombination rates are greater than 2 cM/Mb), GC-rich regions tend to have a faster molecular clock than GC-poor regions. However, in regions with low recombination rates (less than 0.8 cM/Mb), GC-poor regions tend to have faster molecular clock than GC-rich regions. Furthermore, some regions of the genome, corresponding to cytogenetically defined Giemsa dark bands (G bands), had an approximately 10% faster molecular clock than the genome wide average, despite the fact that both GC content and recombination rates were lower in those regions than the genome-wide average.

Therefore, while GC content and recombination are two important determinants of neutral molecular clocks, genomic analyses have shown that neither can solely explain regional variability of mammalian molecular clocks. Elucidating the relative contributions of these and other mechanisms on regional heterogeneity is an important topic for future molecular clock studies.

Other Potential Determinants of Genomic Molecular Clocks

In this section, we will discuss some new ideas on neutral genomic molecular clocks that have emerged within the last few years, enabled by the availability of large amount of genomic sequence data. These directions may have far-reaching implications for the study of genomic molecular clocks.

Rapidly Changing Recombination Hotspots May Change Neutral Molecular Clocks

Above we discussed how recombination could affect neutral molecular clocks. Furthermore, new knowledge on the physical distribution of recombination on genomes has opened up new possibilities for recombination to cause rapid change of neutral molecular clocks. In the course of the last few years, it has been established that in the human genome (and presumably in other mammalian genomes) recombination occurs not randomly, but clustered in small ($\sim$kb scale) regions, called 'recombination hotspots' (Jeffreys et al. 2001; Jeffreys et al. 2004; McVean et al. 2004; Myers et al. 2005).

Intriguingly, recombination hotspots seem to appear and disappear rather rapidly on an evolutionary timescale. Ptak et al. (2005) and Winckler et al. (2005) have shown that in the genomes of humans and chimpanzees, the locations of a substantial number of recombination hotspots do not coincide. Yi and Li (2005) also concluded that recombination hotspots are evolutionarily transient, based upon the pattern of molecular evolution of several recombination hotspots in some primates.

A rapid change of recombination hotspots can potentially cause regional neutral molecular clocks to run at different speeds among closely related species. It is known that recombination rates are not strongly conserved between different mammalian species such as human, mouse and rat (Jensen-Seaman et al. 2004). Recombination maps of human and baboon are also known to differ (Rogers et al. 2000). The relationship between the rapid evolution of recombination hotspots and changes of gross recombination rate estimates between genomes needs to be pursued further as detailed data on recombination rates between closely related species become available.

Mutations Caused by DNA Methylation and Heterogeneity of Genomic Molecular Clocks

DNA methylation is an important epigenetic mechanism, involved in several regulatory mechanisms including chromatin compaction, X chromosome inactivation, and imprinting (Jones and Takai 2001; Li 2002). DNA methylation also affects neutral molecular clocks by changing mutational bias and frequency. DNA methylation in mammals occurs at the cytosine in 5′CG3′ context (so-called 'CpG dinucleotides'). Methylated cytosines undergo deamination to become thymine. As a result, DNA methylation can generate a CpG to TpG (or CpA on the complementary strand) mutation. This type of mutation occurs over ten times more frequently than other types of point mutations (Chimpanzee Sequencing and Analysis Consortium 2005; Nachman and Crowell 2000).

Kim et al. (2006) proposed that point mutations originated by DNA methylation could cause heterogeneity of genomic molecular clocks. They argue that molecular clocks at CpG sites should exhibit a relatively constant molecular clock as opposed to generation-time dependency. The reason for this is because mutations caused by DNA methylation will accumulate continuously during the lifetime of germline cells, while other mutations accumulate only when DNA replication occurs. To test this hypothesis, they contrasted molecular clocks of CpG

dinucleotides and other sites by comparing the pairwise distance between human and chimpanzee to that between baboon and macaque (Fig. 2A). These two species pairs share a similar divergence time of 6–8 million years, based upon fossil records (Steiper et al. 2004). In contrast, hominoid rate slowdown predicts that the divergence between two Old World monkeys should be greater than the divergence between the two hominoids. Kim et al. (2006)'s analyses showed that these two types of substitutions show statistically significantly different molecular clocks. In particular, consistent with the aforementioned hypothesis, the molecular clock at CpG sites exhibit a relatively constant molecular clock while the molecular clock at other sites conform to the generation time effect (Fig. 2B).

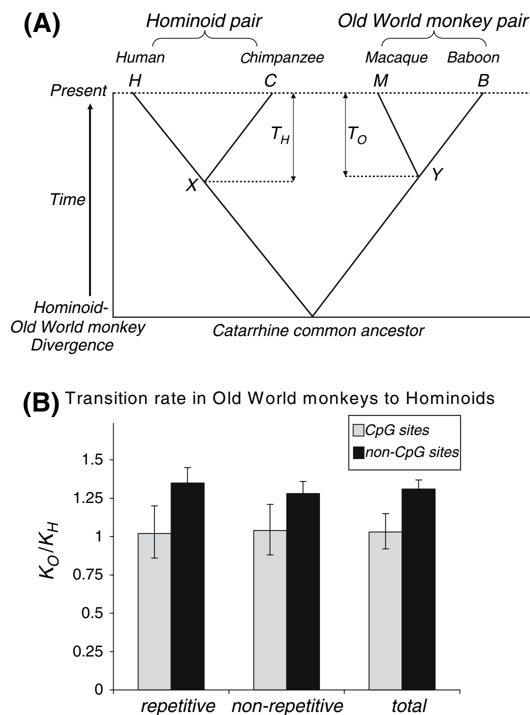

Fig. 2 (**A**) Phylogeny of the four taxa analyzed in Kim et al. (2006). T_O denotes the time since the split between the two Old World monkey species, and T_H denotes the time since the split between the two hominoids. Fossil records suggest that T_O and T_H are very close to each other. X and Y denote the common ancestors of human-chimpanzee and of macaque-baboon, respectively. The genetic divergence between the two hominoid species (K_H) is the sum of K_{HX} and K_{HC}. Likewise, K_O is the sum of K_{MY} and K_{BY}. (**B**) Contrasting molecular clocks of transitions at CpG sites versus those at non-CpG sites. The Y-axis shows the rate difference in the baboon-macaque pair to that in the human-chimpanzee pair. The Old World monkey pair has accumulated significantly more transitions in non-CpG sites, as expected by the generation time effect. In contrast, transitions at CpG sites, which are primarily of methylation-origin, show no difference between the two pairs. Figures are adapted from Kim et al. (2006)

Therefore, DNA methylation causes molecular clocks to vary within genomes. Furthermore, genomic regions harboring different amounts of CpG sites may behave differently in terms of molecular clocks. For example, coding regions have more CpG sites than non-coding regions, due to the constraint of the codon usage table. Therefore, it is likely that molecular clocks at exons are dominated by the CpG molecular clock, which is fast and relatively time-dependent. On the other hand, non-coding regions, which are often relatively devoid of CpG sites, may follow the canonical genomic molecular clock that is generation-time dependent. Indeed, it has been shown earlier that between closely related primate species, exons evolve faster than introns, because they have proportionally more CpG sites (Subramanian and Kumar 2003). This example demonstrates that taking into account the effect of CpG sites is necessary to make inference on genomic molecular clocks.

Pervasive Natural Selection on Non-coding and Synonymous Sites

The usage of non-coding regions in molecular clock studies is justified by the assumption that non-coding regions are usually neutral. However recent evidence has shown that a substantial portion of non-coding regions is subject to functional constraint. For instance, approximately 5% of mammalian genomes appears to be selectively constrained and under purifying selection (Dermitzakis et al. 2002). Also, a large portion of the human genome is actively transcribed (Cheng et al. 2005; Semon and Duret 2004).

If non-coding markers used for molecular clock analyses contain substantial amount of selectively conserved (under purifying selection) sequences, they will evolve slower than the background neutral clock. This can cause practical problems in the usage of molecular clock such as overestimation of divergence times. Alternatively, non-coding markers may be subject to different types of natural selection in different lineages. For example, non-coding markers may experience rapid adaptive evolution resulting in accelerated molecular clocks only in some lineages. Inference based upon the molecular clock in this case is unreliable.

Synonymous sites may be subject to natural selection more widely than commonly thought. In many species including several prokaryotes, yeast and invertebrates, synonymous sites are known to be under natural selection because different codons have different efficiency of gene expression and/or protein translation (Akashi 1994, 2001; Drummond et al. 2005). Mammalian synonymous sites, however, were considered to be free from such selective

effects and largely neutral. However, recent studies revealed the presence of several different selective mechanisms operating on mammalian synonymous sites, including selection for mRNA stability, conservation of splicing signals, and maximization of translational efficiency (Chamary et al. 2006).

Recently, Kondrashov et al. (2006) analyzed the patterns of polymorphism and divergence at synonymous sites and concluded that mammalian synonymous sites are under weak selection toward G and C nucleotides. Interestingly, they demonstrated that such selection could accelerate synonymous molecular clocks at CpG sites, while decelerating molecular clocks at non-CpG sites. In other words, selection and nucleotide composition can interact to cause variability of regional molecular clocks at synonymous sites.

Concluding remarks

The availability of large-scale sequence data has enabled detailed analyses of genomic molecular clocks. From such analyses, new insights on evolutionary rate variation within and between different genomes are emerging. Understanding variation of genomic molecular clocks will be useful for differentiating the effects of neutral evolutionary forces and natural selection on genome evolution.

References

Akashi, H. (1994). Synonymous codon usage in Drosophila melanogaster: Natural selection and translational accuracy. *Genetics, 136*(3), 927–935.

Akashi, H. (2001). Gene expression and molecular evolution. *Current Opinion in Genetics & Development, 11*, 660–666.

Birdsell, J. A. (2002). Integrating genomics, bioinformatics, and classical genetics to study the effects of recombination on genome evolution. *Molecular Biology and Evolution, 19*, 1181–1197.

Bohossian, H. B., Skaletsky, H., & Page, D. C. (2000). Unexpectedly similar rates of nucleotide substitution found in male and female hominids. *Nature, 406*, 622–625.

Britten, R. J. (1986). Rates of DNA sequence evolution differ between taxonomic groups. *Science, 231*, 1393–1398.

Bromham, L., & Penny, D. (2003). The modern molecular clock. *Nature Reviews Genetics, 4*, 216–224.

Cannarozzi, G., Schneider, A., & Gonnet, G. (2007). A phylogenomic study of human, dog and mouse. *PLoS Computational Biology, 3*(1), e2.

Castresana, J. (2002). Estimation of genetic distances from human and mouse introns. *Genome Biology, 3*(6), research0028.1–0028.7.

Chamary, J. V., Parmley, J. L., & Hurst, L. D. (2006). Hearing silence: Non-neutral evolution at synonymous sites in mammals. *Nature Reviews Genetics, 7*, 98–108.

Chang, B. H., Shimmin, L. C., Shyue, S.-K., Hewett-Emmett, D., & Li, W.-H. (1994). Weak male-driven molecular evolution in rodents. *Proceedings of the National Academy of Sciences of the United States of America, 91*, 827–831.

Chen, F. C., & Li, W.-H. (2001). Genomic divergence between humans and other hominoids and the effective population size of the common ancestor of humans and chimpanzees. *American Journal of Human Genetics, 68*, 444–456.

Cheng, J., Kapranov, P., Drenkow, J., Dike, S., Brubaker, S., Patel, S., et al. (2005). Transcriptional maps of 10 human chromosomes at 5-nucleotide resolution. *Science, 308*(5725), 1149–1154.

Chimpanzee Sequencing and Analysis Consortium. (2005). Initial sequence of the chimpanzee genome and comparison with the human genome. *Nature, 437*, 69–87.

Dermitzakis, E. T., Reymond, A., Lyle, R., Scamuffa, N., Ucla, C., Deutsch, S., et al. (2002). Numerous potentially functional but non-genic conserved sequences on human chromosome 21. *Nature, 420*, 578–582.

Drummond, D. A., Bloom, J. D., Adami, C., Wilke, C. O., & Arnod, F. H. (2005). Why highly expressed proteins evolve slowly. *Proceedings of the National Academy of Sciences of the United States of America, 102*, 14338–14343.

Easteal, S. (1991). The relative ratae of DNA evolution in primates. *Molecular Biology and Evolution, 8*(1), 115–127.

Ellegren, H., & Fridolfsson, A. K. (1997). Male-driven evolution of DNA sequences in birds. *Nature Genetics, 17*, 182–184.

Filipski, J. (1988). Why the rate of silent codon substitutions is variable within a vertebrate's genome. *Journal of Theoretical Biology, 134*(2), 159–164.

Fullerton, S. M., Bernardo-Carvalho, A., & Clark, A. G. (2001). Local rates of recombination are positively correlated with GC content in the human genome. *Molecular Biology and Evolution, 18*(6), 1139–1142.

Galtier, N., Piganeau, G., Mouchiroud, D., & Duret, L. (2001). GC-content evolution in mammalian genomes: The biased gene conversion hypothesis. *Genetics, 159*, 907–911.

Goodman, M. (1961). The role of immunologic differences in the phyletic development of human behavior. *Human Biology, 33*, 131–162.

Goodman, M. (1962). Evolution of the immunologic species specificity of human serum proteins. *Human Biology, 34*, 104–150.

Gu, X., & Li, W.-H. (1992). Higher rates of amino acid substitution in rodents than in humans. *Molecular Phylogenetics and Evolution, 1*(3), 211–214.

Hellmann, I., Ebersberger, I., Ptak, S. E., Paabo, S., & Przeworski, M. (2003). A neutral explanation for the correlation of diversity with recombination rates in humans. *American Journal of Human Genetics, 72*(6), 1527–1535.

Huang, S.-W., Friedman, R., Yu, N., Yu, A., & Li, W.-H. (2005). How strong is the mutagenicity of recombination in mammals? *Molecular Biology and Evolution, 22*(3), 426–431.

Huttley, G. A., Wakefield, M. J., & Easteal, S. (2007). Rates of genome evolution and branching order from whole genome analysis. *Molecular Biology and Evolution, 24*, 1772–1730.

Jeffreys, A. J., Holloway, J. K., Kauppi, L., May, C. A., Neumann, R., Timothy Slingsby, M., et al. (2004). Meiotic recombination hot spots and human DNA diversity. *Philosophical Transactions of the Royal Society London, B., 359*, 141–152.

Jeffreys, A. J., Kauppi, L., & Neumann, R. (2001). Intensely punctate meiotic recombination in the class II region of the major histocompatibility complex. *Nature Genetics, 29*, 217–222.

Jensen-Seaman, M. I., Furey, T. S., Payseur, B. A., Lu, Y., Roskin, K. M., Chen, C.-F., et al. (2004). Comparative recombination rates in the rat, mouse, and human genomes. *Genome Research, 14*(4), 528–538.

Jones, P. A., & Takai, D. (2001). The role of DNA methylation in mammalian epigenetics. *Science, 293*, 1068–1070.

Keightley, P. D., Lercher, M. J., & Eyre-Walker, A. (2005). Evidence for widespread degradation of gene control regions in hominoid genomes. *PLoS Biology, 3*, e42.

Kim, S.-H., Elango, N., Warden, C. W., Vigoda, E., & Yi, S. (2006). Heterogenous genomic molecular clocks in primates. *PLoS Genetics, 2*, e163.

Kimura, M. (1983). *The neutral theory of molecular evolution*. Cambridge, UK: Cambridge University Press.

Kohne, C. (1970). Evolution of higher-organism DNA. *Quarterly Reviews of Biophysics, 3*, 327–375.

Kondrashov, F. A., Ogurtsov, A. Y., & Kondrashov, A. S. (2006). Selection in favor of nucleotides G and C diversifies evolution rates and levels of polymorphism at mammalian synonymous sites. *Journal of Theoretical Biology, 240*, 616–626.

Kong, A., Gudbhartsson, D. F., Sainz, J., Jonsdottir, G. M., Gudjonsson, S. A., Richardsson, B., et al. (2002). A high-resolution recombination map of the human genome. *Nature Genetics, 31*, 241–247.

Kumar, S. (2005). Molecular clocks: Four decades of evolution. *Nature Reviews Genetics, 6*, 654–662.

Kumar, S., & Subramanian, S. (2002). Mutation rates in mammalian genomes. *Proceedings of the National Academy of Sciences of the United States of America, 99*(2), 803–808.

Laird, C. D., McConaughy, B. L., & McCarthy, B. J. (1969). Rate of fixation of nucleotide substitutions in evolution. *Nature, 224*, 149–154.

Lercher, M. J., & Hurst, L. D. (2002). Human SNP variability and mutation rate are higher in regions of high recombination. *Trends in Genetics, 18*(7), 337–340.

Li, W.-H. (1997). *Molecular evolution*. Sunderland, MA: Sinauer.

Li, E. (2002). Chromatin modification and epigenetic reprogramming in mammalian development. *Nature Reviews Genetics, 3*, 662–673.

Li, W.-H., Ellsworth, D. L., Krushkal, J., Chang, B. H.-J., & Hewett-Emmett, D. (1996). Rates of nucleotide substitution in primates and rodents and the generation-time effect hypothesis. *Molecular Phylogenetics and Evolution, 5*(1), 182–187.

Li, W.-H., Tanimura, M., & Sharp, P. M. (1987). An evaluation of the molecular clock hypothesis using mammalian DNA sequences. *Journal of Molecular Evolution, 25*, 330–342.

Li, W.-H., Yi, S., & Makova, K. (2002). Male-driven evolution. *Current Opinion in Genetics & Development, 12*(6), 650–656.

Makova, K. D., & Li, W.-H. (2002). Strong male-driven evolution of DNA sequences in humans and apes. *Nature, 416*, 624–626.

Margoliash, E. (1963). Primary structure and evolution of cytochrome C. *Proceedings of the National Academy of Sciences of the United States of America, 50*, 672–679.

McVean, G. A. T., Myers, S. R., Hunt, S., Deloukas, P., Bentley, D. R., & Donnelly, P. (2004). The fine-scale structure of recombination rate variation in the human genome. *Science, 304*, 581–584.

Meunier, J., & Duret, L. (2004). Recombination drives the evolution of GC-content in the human genome. *Molecular Biology and Evolution, 21*(6), 984–990.

Montoya-Burgos, J. I., Boursot, P., & Galtier, N. (2003). Recombination explains isochores in mammalian genomes. *Trends in Genetics, 19*, 128–130.

Murphy, W. J., Eizirik, E., O'Brien, S. J., Madsen, O., Scally, M., Douady, C. J., et al. (2001). Resoluion of the early placental mammal radiation using Bayesian phylogenetics. *Science, 294*, 2348–2351.

Myers, S., Bottolo, L., Frreeman, C., McVean, G., & Donnelly, P. (2005). A fine-scale map of recombination rates and hotspots across the human genome. *Science, 310*(Oct. 14), 321–324.

Nachman, M. W., & Crowell, S. L. (2000). Estimate of the mutation rate per nucleotide in humans. *Genetics, 156*(1), 297–304.

Patterson, N., Richter, D. J., Gnerre, S., Lander, E. S., & Reich, D. (2006). Genetic evidence for complex speciation of humans and chimpanzees. *Nature, 441*, 1103–1108.

Perry, J., & Ashworth, A. (1999). Evolutionary rate of a gene affected by chromosomal position. *Current Biology, 9*, 987–989.

Pollard, K. S., Salama, S. R., Lambert, N., Lambot, M.-A., Coppens, S., Pedersen, J. S., et al. (2006). An RNA gene expressed during cortical development evolved rapidly in humans. *Nature, 443*, 167–172.

Ptak, S. E., Hinds, D. A., Koehler, K., Nickel, B., Patil, N., Ballinger, D. G., et al. (2005). Fine-scale recombination patterns differ between chimpanzees and humans. *Nature Genetics, 37*(4), 429–434.

Rogers, J., Mahaney, M. C., et al. (2000). A genetic linkage map of the baboon (Papio hamadryas) genome based on human microsatellite polymorphisms. *Genomics, 67*, 237–247.

Sarich, V. M., & Wilson, A. C. (1967). Immunological tie scale for hominid evolution. *Science, 158*, 1200–1203.

Semon, M., & Duret, L. (2004). Evidence that functional transcription units cover at least half of the human genome. *Trends in Genetics, 20*, 229–232.

Shimmin, L. C., Chang, B. H., & Li, W.-H. (1993). Male-driven evolution of DNA sequences. *Nature, 362*, 745–747.

Steiper, M. E., & Young, N. M. (2006). Primate molecular divergence dates. *Molecular Phylogenetics and Evolution, 41*, 384–394.

Steiper, M. E., Young, N. M., & Sukrarna, T. Y. (2004). Genomic data support the hominoid slowdown and an early Oligocene estimate for the hominoid-cercopithecoid divergence. *Proceedings of the National Academy of Sciences of the United States of America, 101*, 17021–17026.

Subramanian, S., & Kumar, S. (2003). Neutral substitutions occur at a faster rate in exons than in noncoding DNA in primate genomes. *Genome Research, 13*, 838–844.

Winckler, W., Myers, S. R., Richter, D. J., Onofrio, R. C., McDonald, G. J., Bontrop, R. E., et al. (2005). Comparison of fine-scale recombination rates in humans and chimpanzees. *Science, 308*(5718), 107–111.

Wolfe, K. H., Sharp, P. M., & Li, W.-H. (1989). Mutation rates differ among regions of the mammalian genome. *Nature, 337*, 283–285.

Wu, C.-I., & Li, W.-H. (1985). Evidence for higher rates of nucleotide substitution in rodents than in man. *Proceedings of the National Academy of Sciences of the United States of America, 82*, 1741–1745.

Yi, S., Ellsworth, D. L., & Li, W.-H. (2002). Slow molecular clocks in Old World monkeys, apes, and humans. *Molecular Biology and Evolution, 19*(12), 2191–2198.

Yi, S., & Li, W.-H. (2005). Molecular evolution of recombination hotspots and highly recombining pseudoautosomal regions in hominoids. *Molecular Biology and Evolution, 22*, 1223–1230.

Yi, S., Summers, T. J., Pearson, N. M., & Li, W.-H. (2004). Recombination has little effect on the rate of sequence divergence in pseudoausotomal boundary 1 among humans and great apes. *Genome Research, 14*, 37–43.

Zuckerkandl, E., & Pauling, L. B. (1962). Molecular disease, evolution, and genetic heterogeneity. In M. Kasha & B. Pullman (Eds.), *Horizons in biochemistry* (pp. 189–225). New York: Academic Press.