75. Hsieh, C. L. & Lieber, M. R. CpG methylated minichromosomes become inaccessible for V(D)J recombination after undergoing replication. *EMBO J.* **11**, 315–325 (1992).
76. Cherry, S. R., Beard, C., Jaenisch, R. & Baltimore, D. V(D)J recombination is not activated by demethylation of the κ-locus. *Proc. Natl Acad. Sci. USA* **97**, 8467–8472 (2000).
77. Ji, Y., Zhang, J., Lee, A. I., Cedar, H. & Bergman, Y. A multistep mechanism for the activation of rearrangement in the immune system. *Proc. Natl Acad. Sci. USA* **100**, 7557–7562 (2003).
78. Gilbert, D. M. Nuclear position leaves its mark on replication timing. *J. Cell Biol.* **152**, F11–F15 (2001).
79. Li, F. *et al.* The replication timing program of the Chinese hamster β-globin locus is established coincident with its repositioning near peripheral heterochromatin in early G1 phase. *J. Cell Biol.* **154**, 283–292 (2001).
80. Dhar, V., Skoultchi, A. I. & Schildkraut, C. L. Activation and repression of a β-globin gene in cell hybrids is accompanied by a shift in its temporal regulation. *Mol. Cell. Biol.* **9**, 3524–3532 (1989).
81. Forrester, W. C., Thompson, C., Elder, J. T. & Groudine, M. A developmentally stable chromatin structure in the human β-globin gene cluster. *Proc. Natl Acad. Sci. USA* **83**, 1359–1363 (1986).
82. Groudine, M., Kohwi-Shigematsu, T., Gelinas, R., Stamatoyannopoulos, G. & Papayannopoulou, T. Human fetal to adult hemoglobin switching: changes in chromatin structure of the β-globin gene locus. *Proc. Natl Acad. Sci. USA* **80**, 7551–7555 (1983).
83. Busslinger, M., Hurst, J. & Flavell, R. A. DNA methylation and the regulation of the globin gene expression. *Cell* **34**, 197–206 (1983).
84. Thiel, G., Lietz, M. & Hohl, M. How mammalian transcriptional repressors work. *Eur. J. Biochem.* **271**, 2855–2862 (2004).
85. Brown, K. E. *et al.* Association of transcriptionally silent genes with Ikaros complexes at centromeric heterochromatin. *Cell* **91**, 845–854 (1997).
86. Hahm, K. *et al.* Helios, a T cell-restricted Ikaros family member that quantitatively associates with Ikaros at centromeric heterochromatin. *Genes Dev.* **12**, 782–796 (1998).
87. Lunyak, V. V. *et al.* Corepressor-dependent silencing of chromosomal regions encoding neuronal genes. *Science* **298**, 1747–1752 (2002).
88. Francis, N. J., Kingston, R. E. & Woodcock, C. L. Chromatin compaction by a polycomb group protein complex. *Science* **306**, 1574–1577 (2004).
89. Baumann, M., Mamais, A., McBlane, F., Xiao, H. & Boyes, J. Regulation of V(D)J recombination by nucleosome positioning at recombination signal sequences. *EMBO J.* **22**, 5197–5207 (2003).
90. Smith, K. T., Coffee, B. & Reines, D. Occupancy and synergistic activation of the *FMR1* promoter by Nrf-1 and Sp1 *in vivo*. *Hum. Mol. Genet.* **13**, 1611–1621 (2004).
91. de Vries, B. B. *et al.* Variable *FMR1* gene methylation of large expansions leads to variable phenotype in three males from one fragile X family. *J. Med. Genet.* **33**, 1007–1010 (1996).
92. Rousseau, F., Robb, L. J., Rouillard, P. & Der Kaloustian, V. M. No mental retardation in a man with 40% abnormal methylation at the *FMR-1* locus and transmission of sperm cell mutations as premutations. *Hum. Mol. Genet.* **3**, 927–930 (2000).
93. Smeets, H. J. *et al.* Normal phenotype in two brothers with a full *FMR1* mutation. *Hum. Mol. Genet.* **4**, 2103–2108 (2000).
94. Bardoni, B. & Mandel, J. L. Advances in understanding of fragile X pathogenesis and FMRP function, and in identification of X linked mental retardation genes. *Curr. Opin. Genet. Dev.* **12**, 284–293 (2002).
95. Bird, A. P. & Wolffe, A. P. Methylation-induced repression — belts, braces, and chromatin. *Cell* **99**, 451–454 (1999).
96. Eden, S., Hashimshony, T., Keshet, I., Thorne, A. W. & Cedar, H. DNA methylation models histone acetylation. *Nature* **394**, 842–843 (1998).
97. Razin, A. & Cedar, H. Distribution of 5-methylcytosine in chromatin. *Proc. Natl Acad. Sci. USA* **74**, 2725–2728 (1977).
98. Gruenbaum, Y., Naveh-Many, T., Cedar, H. & Razin, A. Sequence specificity of methylation in higher plant DNA. *Nature* **292**, 860–862 (1981).
99. Gruenbaum, Y., Stein, R., Cedar, H. & Razin, A. Methylation of CpG sequences in eukaryotic DNA. *FEBS Lett.* **123**, 67–71 (1981).
100. Rountree, M. R., Bachman, K. E. & Baylin, S. B. DNMT1 binds HDAC2 and a new co-repressor, DMAP1, to form a complex at replication foci. *Nature Genet.* **25**, 269–277 (2000).
101. Kerem, B. S., Goitein, R., Diamond, G., Cedar, H. & Marcus, M. Mapping of DNase I sensitive regions on mitotic chromosomes. *Cell* **38**, 493–499 (1984).

Acknowledgements
We would like to acknowledge the support of the US National Institutes of Health, the Israel Science Foundation, the Israel Cancer Research Fund, the Belfer Foundation and the Rosetrees Trust.

Competing interests statement
The authors declare no competing financial interests.

Online links

DATABASES
The following terms in this article are linked online to:
Entrez: http://www.ncbi.nlm.nih.gov/entrez
DNMT 1 | *FMR1* | NRF1 | *Pou5f1* | *RAG1* | *RAG2* | TBP | *Xist* | ZFPN1A1
OMIM: http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=OMIM
fragile X syndrome

FURTHER INFORMATION
The Cedar laboratory web site: http://www.md.huji.ac.il/depts/humangenetics/cedar/cedar.html
Access to this interactive links box is free online.

TIMELINE

Molecular clocks: four decades of evolution

Sudhir Kumar

Abstract | During the past four decades, the molecular-clock hypothesis has provided an invaluable tool for building evolutionary timescales, and has served as a null model for testing evolutionary and mutation rates in different species. Molecular clocks have also influenced the development of theories of molecular evolution. As DNA-sequencing technologies have progressed, the use of molecular clocks has increased, with a profound effect on our understanding of the temporal diversification of species and genomes.

The idea that molecular evolution occurs at an approximately uniform rate over time, known as the molecular-clock hypothesis, was put forward in the early 1960s, remarkably only a few years after DNA was established as the hereditary material and the first protein (insulin) was completely sequenced in the mid-1950s (REFS 1,2). The molecular-clock hypothesis recognized the similarity of protein evolutionary rates among morphologically diverse species — even those with vastly different life-history traits. For this reason, the early history of molecular clocks, filled with numerous fundamental innovations, was also rife with many concerns about their accuracy and general applicability — concerns that have been reiterated over the past four decades. Here I provide a glimpse into this history and describe how early proposals and innovations have been refined and used in the modern age of genomics (TIMELINE). For the sake of brevity, and because the use of molecular clocks began with the analysis of mammalian proteins, the focus here will be on molecular clocks in mammals and their close relatives.

A protein molecular clock

The use of molecular clocks began in 1962 when Zuckerkandl and Pauling[3], in order to date the origins of different globins, assumed that there is a uniform rate of molecular evolution among species and duplicated proteins (BOX 1). This informal proposal of a molecular clock[4] was followed by a formal statement the following year by Margoliash[5]. "It appears that the number of residue differences between cytochrome *c* of any two species is mostly conditioned by the time elapsed since the lines of evolution leading to these two species originally diverged." Margoliash compared sequences of species within a group (the ingroup species) with those of an external reference species (the outgroup species), and found that the numbers of amino-acid differences were similar between outgroup–ingroup pair comparisons (FIG. 1a). This indicated uniformity in the rate at which differences accumulated among ingroup species, as evolution has occurred for exactly the same amount of time in these organisms (see FIG. 1a for details).

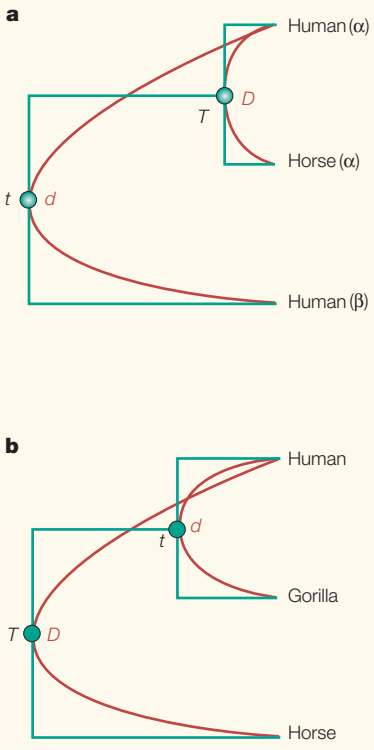
This approach of comparing evolutionary rates among ingroup species with reference to an outgroup species is called the relative-rate approach, as it does not require us to know the timing of species divergence. In 1964, in further support of the molecular-clock hypothesis, Doolittle and Blombäck[6] explicitly compared protein sequence identity with time of species divergence and found an inverse correlation between the two (FIG. 1b).

The validity of the hypothesis that clocks run at the same speed in different species was immediately questioned[7,8]. According to Mayr in 1965 (REF. 9), "evolution is too complex and too variable a process, connected with too many factors, for the time dependence of the evolutionary process at the molecular level to be a simple function" (see REF. 4 for a review on this debate). The globin data that were analysed by Zuckerkandl and Pauling were also questioned, because fewer protein differences were observed between human and kangaroo sequences than between human and horse[9], contradictory to the known evolutionary relationships between these species. The stochastic nature of the evolutionary process, which is now well established, had not been realized at that time, and this observation was taken as evidence not only against molecular clocks, but also against molecular phylogenetics. However, in 1965, Zuckerkandl and Pauling[9] explained the statistical properties of the relationships of evolutionary distance and geological time, which showed that smaller protein differences between the human and kangaroo can occur by chance even if evolutionary rates are constant. They also stated that, although any single protein would not accurately reflect evolutionary distance, the sensitivity of a protein in measuring evolutionary time would increase with protein length, time of evolutionary divergence (in millions of years) of the species being analysed and the evolutionary rate of the protein. In the same work, the authors first coined the term 'molecular evolutionary clock'.

Further challenges to the validity of molecular clocks came from Goodman and colleagues in the 1960s. They reported a much lower rate of evolution of albumin protein in humans than in other species, a phenomenon called the 'hominoid slowdown'[10–12]. This observation was contrary to the requirement for the molecular-clock hypothesis that similar evolutionary rates occur in different species. However, in 1967, Sarich and Wilson[13] used a relative-rate approach to argue that the observed patterns of albumin evolution were consistent with a molecular clock. The discrepancy occurred because Goodman and colleagues had computed evolutionary rates by dividing the inferred number of amino-acid changes by the contemporary estimates of divergence times. However, the divergence time for humans and chimpanzees was then thought to be 30 million years — about five times greater than the age accepted today — resulting in a large underestimate of evolutionary rate in the hominoid lineage. Wilson *et al.*[14] and Easteal *et al.*[15] provide excellent summaries of how time estimates drawn from the fossil record, which since have been considerably revised, led to significant protraction of the molecular-clock debate.

The scientists who initially developed the idea of protein clocks used radically different evolutionary mechanisms to explain them. In keeping with the contemporary supposition of evolution by natural selection, Margoliash and Smith[5,16] offered a selectionist explanation in which the rate of evolution increases owing to positive selection only over short time intervals, producing similarity when rates are averaged over long time spans. Selectionist models continued to be put forward into the 1970s and beyond (see REFS 17,18 for examples). By contrast, Zuckerkandl and Pauling[3,9] argued that most observed substitutions are mutations that have little or no effect on protein function (neutral mutations), which have been fixed by random chance alone in the population (random genetic drift[19]). As a result, the extent of functional differences between proteins is not proportional to the

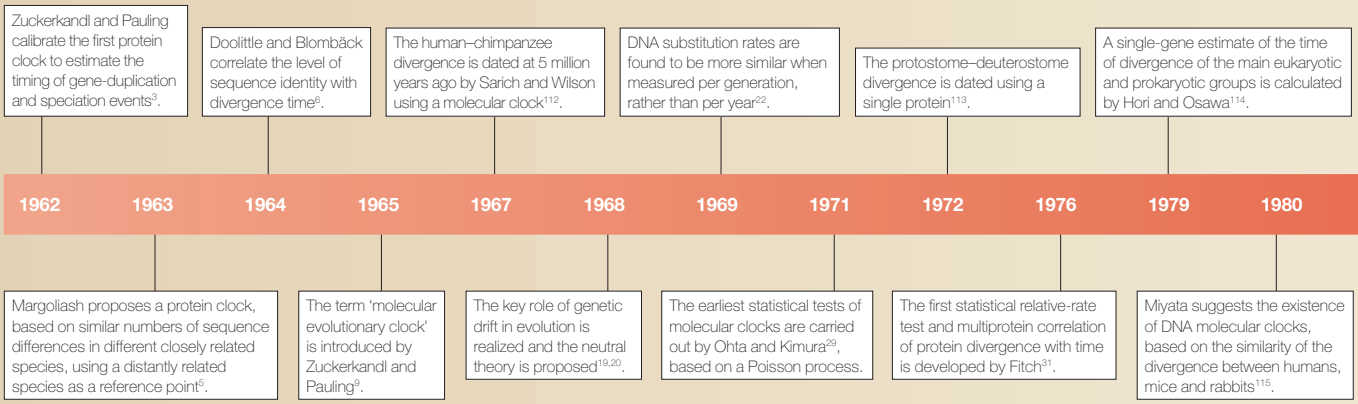
Box 1 | The earliest uses of the molecular clock



In 1962, Zuckerkandl and Pauling[3] estimated the time of divergence of four members of the haemoglobin gene family (α, β, γ and δ) by assuming an approximate molecular clock. This was calibrated using the number of observed sequence differences (*D*) between the horse and human α-haemoglobin proteins and the divergence time between the two species (*T*), which is based on the fossil record. They took a pair-wise approach to estimating divergence times, which is shown schematically for α- and β-haemoglobin in panel a. The molecular-clock calibration was carried out by dividing twice the known divergence time by the amount of sequence divergence (2*T*/*D*); the factor of 2 is used here because *D* is equal to the sum of divergence from the common ancestor to the two descendents. This calibration was then used to convert other measurements of protein sequence differences to time. For example, the formula $t = d\,(T/D)$ gives the time when the α- and β-chains diverged, where *d* is the amount of sequence difference between α- and β-chains in humans. The time estimate obtained will have the same units as the time used for clock calibrations (in this case, millions of years).

Zuckerkandl and Pauling also estimated the timing of the human–gorilla divergence using α- and β-chains separately (panel b). They calculated the molecular-clock calibration to be 11 to 18 million years (Myr) per amino-acid substitution, based on the observation of 18 differences between human and horse α-haemoglobin proteins and the assumption that these two species diverged 100–160 million years ago (Mya). Using an average calibration of 14.5 Myr per substitution, the human–gorilla divergence was dated to have occurred 14.5 and 7.25 Mya by α- and β-chains, because human and gorilla show two and one differences in these chains, respectively. Therefore, Zuckerkandl and Pauling[3] reported a mean date of 11 Mya for the human–gorilla divergence from an analysis of the two proteins. One year later, Margoliash[5] used the same calibration point to estimate multiple species divergence times. These estimates were based on single, slowly evolving proteins and were therefore not very accurate. In 1965, Zuckerkandl and Pauling[9] predicted that the accuracy of molecular clocks would be improved by using many proteins of different types. Over the past decade, a large number of proteins have been analysed to estimate divergence times among the principal groups of mammals and among animal phyla[35,39,60].

Timeline | **Four decades of molecular clocks**



number of amino-acid substitutions in each protein. Their observations helped to lay the foundation for the now widely accepted neutral theory of molecular evolution[20,21].

A molecular clock for DNA

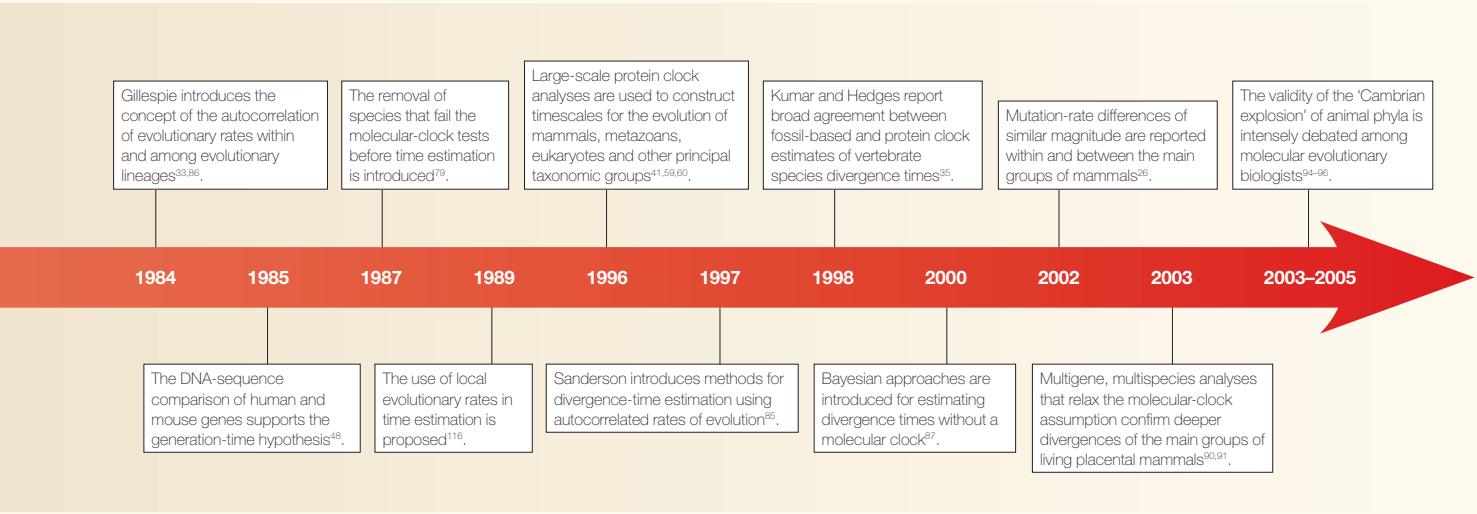
Following the proposal and discussion of protein molecular clocks in the 1960s, it was natural to investigate whether this idea could be extended to DNA. However, direct comparison of DNA sequences was not possible until the mid-to-late 1970s, when DNA-sequencing techniques became available, and therefore indirect methods were initially used. In 1969, Laird *et al.*[22] estimated evolutionary divergence between species by measuring the strength of heterologous DNA–DNA duplexes formed between single-copy genomic DNA from different species[15]. Single-copy DNA was used because the amount of repetitive DNA was known to be vastly different even among closely related mammals[23], which might lead to biased estimates of evolutionary divergence. These methods were inferior to the direct comparison of protein sequences as individual sequence changes could not be directly counted. However, they did provide genome-wide sequence-divergence estimates (at least for non-repetitive DNA), rather than estimates that were based on comparisons of only one or a few proteins[15,24], which do not constitute a representative sample for the whole genome.

Using the DNA–DNA hybridization technique, Laird *et al.*[22] inferred a 10-fold difference in DNA mutation rates per year between murid rodents (for example, mice) and artiodactyls (for example, cows) and determined that the rates would be markedly more similar if they were measured in units of generation time (in years). This marks the beginning of the consideration of generation time in dictating the rate of molecular evolution. One year later, in 1970, Kohne[25] used a relative-rate approach to demonstrate a lower rate of DNA mutation in humans than in the great apes. Remarkably, the 20% difference observed using DNA-association kinetics was close to that obtained in later studies that were based on a direct comparison of DNA sequences[26,27]. Building on the generation-time hypothesis of Laird *et al.*, Kohne linked the mutation rate with the number of germ-cell divisions in each generation. This provided a biological mechanism for the observed patterns of slower evolutionary rates in species with longer generation times, as their germline cells tend to undergo fewer replications (in which mutations can occur) per year. This evidence also indicated that errors in DNA replication are the primary source of mutation.

Variations in the protein clock

Although no substantial progress was made in DNA molecular-clock research in the 1970s, the sequencing of many proteins from diverse species provided opportunities to examine the global (relating to different species) and universal (relating to different proteins) nature of the protein clock. By 1971, it was clear that different proteins evolve at vastly different rates; for example, fibrinopeptides evolve very fast and cytochrome *c* evolves very slowly[28]. By contrast, assessing the similarity of protein evolutionary rates in different species was not straightforward. The possibility of directly observing a global clock for a protein had already been excluded, because of the stochastic nature of the evolutionary process[9]. In 1971, Ohta and Kimura[29] suggested that if evolution is taking place at a constant rate, then the stochastic nature of evolutionary change dictates that the average rate in different species and the variance of these rates will be equal when a Poisson distribution (which models events as occurring independently of each other) is used to model the occurrence of the number of the substitutions. They found that the observed variance was larger than the average rate, which was contrary to their expectation and to the existence of a molecular clock. Similar results were reported by Langley and Fitch[30] soon after.

In 1976, Fitch[31] proposed a formal relative-rate test (BOX 2) in an effort to conduct a calibration-free statistical test of molecular evolutionary clocks. He applied this test to a combined analysis of seven mammalian proteins (cytochrome *c*, fibrinopeptides A and B, α- and β-haemoglobins, insulin C-peptide and myoglobin) and found that the null hypothesis of uniform accumulation of all substitutions over time could be rejected. However, the number of protein-altering nucleotide substitutions showed an excellent linear relationship with species divergence time (FIG. 1c). These contrasting observations led him to conclude that "the clock, at least for amino-acid changing nucleotide substitutions, is not the stochastic timepiece that radioactive decay is." He also explained that the observed linear relationship of evolutionary distance with time could, among other possibilities, be due to the uniformity of the combined rate of change over time of all proteins. This was consistent with Zuckerkandl and Pauling's[9] suggestion a decade earlier that during the phases of rapid organismal diversification, only a few biological systems and the proteins involved were likely to undergo an enhanced rate of evolution; most proteins were likely to evolve at their usual rates. Therefore, although we might see discrepancies if only a few proteins are analysed, these will be minimized if many proteins are used.



In 1979, Gillespie and Langley[32] suggested that the assumptions made in the Ohta–Kimura[29] test, which argued against the existence of a molecular clock, might not be correct. They concluded that the "available data may not be incompatible with a constant-rate neutral allele model of evolution." However, Gillespie[33] reversed his position in 1984 and stated that "the inferred dynamics of molecular evolution appear to be much more erratic than suggested by neutral allele models and by the molecular-clock hypothesis." He even proposed models for the evolution of the evolutionary rate itself, which have recently formed the basis of sophisticated methods for estimating times of evolutionary divergence (see below). However, soon after Gillespie's study, Takahata showed that the higher variance in evolutionary rates between species than would be expected from the molecular-clock hypothesis need not be attributed to erratic evolutionary rates or non-neutral evolution[34].

It is now clear that the molecular clock can be statistically rejected for a substantial proportion of proteins in comparisons of the main taxonomic groups of vertebrates[35–37], invertebrates[38–40] and other eukaryotes and prokaryotes[41,42]. However, it still provides a useful means of estimating evolutionary time. There are also significant differences in overall proteome evolutionary rates among species (see REFS 36,37,42 for some examples), but these disparities are much smaller than the differences in the morphological and life-history traits among these groups. Researchers have used these observations to argue for the decoupling of protein evolutionary rates from morphological evolutionary rates, which vary dramatically among species and taxonomic groups[7,14,43–45].

Two decades of DNA clock controversy

By the beginning of the 1980s, the development of DNA-sequencing techniques allowed the sequencing of many genes for which protein sequences had previously been analysed. At the same time, the neutral theory of molecular evolution matured, providing a framework for estimating the rate of mutation from the rate of neutral substitutions[20,46] (BOX 3). These two developments triggered investigations into the similarity of DNA mutation rates among species.

In 1980, Miyata *et al.*[47] calculated that DNA mutation rates among mammals were very similar. However, in the same year, Bonner *et al.*[20,46] reported a significantly lower rate of evolution in Malagasy primates compared with other extant primates. Five years later, Wu and Li[48] provided significant evidence for a large mutation-rate difference between humans and murid rodents. They suggested that this was due to a generation-time effect, as previously indicated by Laird *et al.*[22] and Kohne[25]. In 1986, Britten[49] also showed a much faster rate of mutation in some rodents than in humans, but he suggested that it was not differences in generation times, but changes in repair mechanisms that were the cause. Reviews by Li[50] in 1993 and Easteal *et al.*[15] in 1995 discussed how these controversies remained unresolved even after nearly a decade of intense research following the studies by Wu and Li[48] and Britten[49] in the mid-1980s. This stemmed from disagreements about the evolutionary relationships between humans and rodents and from the use of species divergence times that were based on a fragmentary fossil record in estimating absolute rates of mutation.

In 1993, an inverse relationship between the rate of DNA evolution and body size was observed for some genes and species[51,52], although it was generally believed that body size itself did not directly affect the rate of molecular evolution, but was correlated with other life-history traits. These included generation time, which is generally longer for larger organisms, and metabolic rate, which is generally slower for larger organisms. In 1994, Rand suggested that a high metabolic rate produces an increased concentration of mutagenic oxygen radicals as a result of aerobic respiration, thereby influencing mutation rates in mitochondrial DNA (mtDNA)[117]. This might explain anomalies such as the slow rate of evolution of shark mtDNA[51,53], because sharks have lower metabolic rates, and the slower rates of evolution of the poikilotherms ('cold-blooded' animals) compared with the endotherms ('warm-blooded' animals), as the former have relatively low metabolic rates. However, a study by Bromham *et al.*[54] in 1996 did not find any effect of mammalian metabolic rate on the rate of neutral DNA evolution beyond what might be explained by its covariation with generation time.

Until the late 1990s, owing to the paucity of sequence data, many studies involved the comparison of only one or a few species from a small number of the main groups of mammals. The continued development of high-throughput DNA-sequencing technology, following the advent of PCR in 1985, led to a rapid growth in genetic data by the end of the 1990s (FIG. 2). Using these data in 2002, Kumar and Subramanian[26] were able to characterize mutation-rate differences within and among the main groups of mammals. Although they found an excellent linear relationship between neutral evolutionary distance and fossil-based time estimates, they reported substantial

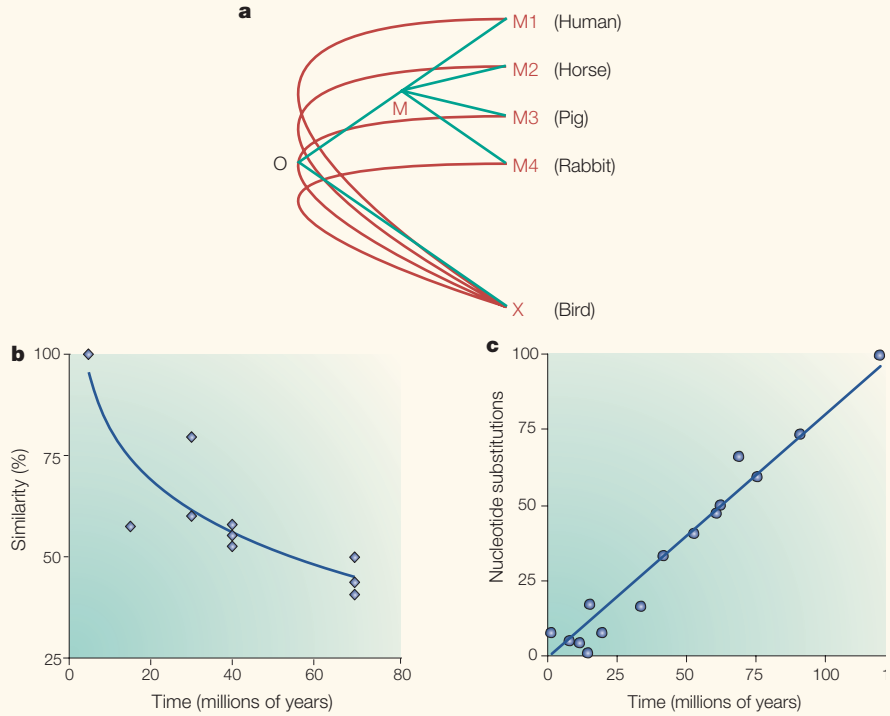


Figure 1 | **Assessing the similarity of evolutionary rates among lineages. a** | A relative-rate comparison for mammalian species (M1–M4; human, horse, pig, rabbit), which are known to have a most recent common ancestor (M), and another species (X; bird). Species X is the outgroup species and is distantly related to species M1–M4, the ingroup species. Evolutionary lineages leading to species M1–M4 separated from the lineage leading to X at the same point, O. Furthermore, species M1–M4 are products of an evolutionary process that has run for exactly the same amount of time, because they share a common ancestor. Therefore, if a given protein is equally different when we compare the same bird protein with proteins from different mammals, then the rate at which differences accumulate is similar among mammals (M1–M4). This approach does not require knowledge of when the common mammalian ancestor existed, which is needed for the approach shown in panel **b**. **b** | A scatter graph showing the negative relationship between the amount of protein sequence identity between species and the times of species divergence, as drawn by Doolittle and Blombäck[6] in 1964. However, because amino-acid substitutions accumulate with time, species that are distantly related temporally show larger evolutionary distances, even if the evolutionary rates are not constant. Furthermore, such diagrams emphasize, both visually and statistically, the oldest divergences over the more recent ones. Therefore, they provide only weak evidence for the existence of molecular clocks. **c** | The linear relationship between the combined number of nucleotide substitutions from 7 proteins and the species divergence time[31]. Each point represents a pair of mammalian species. Such graphs might exaggerate the positive relationship between evolutionary distance and time of divergence, because many species pairs are not independent, as all mammals are related to each other through an underlying phylogeny. Therefore, a test of the significance of the observed correlation requires the use of sophisticated statistical methods[59]. Panel **b** is modified, with permission, from *Nature* REF. 6 © (1964) Macmillan Magazines Ltd. Panel **c** is modified, with permission, from REF. 31 © (1976) Sinauer Associates.

mutation-rate differences between hamsters and mice (20%), cows and pigs (14%), cats and dogs (23%), and humans and Old World monkeys (22%). Interestingly, the magnitude of the rate difference between primate species (22%) was found to be as large as that observed when primates and rodents were compared (18%; see BOX 3 for further details and a discussion of the controversies that arose on this subject).

Because generation times, physiological attributes and other life-history traits are generally more similar within groups than between groups, it seems that that replication-independent processes — such as DNA methylation, recombination and repair mechanisms — have a greater role as a source of mutation[26,55,56] than previously predicted by others, including Kohne[25] and Wu and Li[48]. In summary however, it is clear that mutation rates in different mammals are not identical. As a result, the emphasis of current research has changed from testing the existence of a global DNA clock in mammals to quantifying the extent to which there are rate differences among species and determining their causes[26,57].

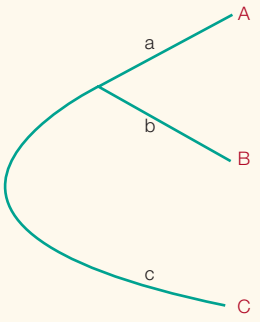
Divergence times from large data sets

In the second half of the 1990s, the rapid growth in the availability of sequence data allowed three independent research groups to begin large-scale, multi-protein analyses for estimating species divergence times. Doolittle *et al.*[41] and Feng *et al.*[58] in 1996 and 1997, respectively, reported the divergence times of the main groups of eukaryotes and prokaryotes, Wray *et al.*[59] dated the diversification of metazoan phyla in 1996, and Hedges *et al.*[60] and Kumar and Hedges[35] in 1996 and 1998, respectively, proposed a timescale for vertebrate evolution. These results sparked a new set of debates by challenging the prevailing hypotheses about the adaptive radiations of mammals that are proposed to have occurred at the Cretaceous–Tertiary (K/T) boundary (~65 million years ago (Mya)) and those of animal phyla that are proposed to have occurred at the beginning of the Cambrian period (~500–600 Mya). In both cases, fossil-based estimates for the age of the most recent common ancestors of extant species are about almost half those that are calculated using molecular clocks (see the review in REF. 61). The fossil record yields minimum dates for species divergence, which partly explains this difference. However, the discrepancy between molecular and fossil times from these studies is unusually large, especially given that many other molecular times agree closely with fossil-based estimates[35], and indicates that there are large gaps in the fossil record.

From 1996 to 2000, independent molecular-clock analyses supported the conclusion that the main 'supergroups' of extant placental mammals diversified before the K/T boundary[62], and the continental-breakup hypothesis[60] was proposed to explain this earlier speciation[63,64]. Some more recent fossil discoveries[65,66] have also led to remarks that "the fossil evidence supports the argument that there were some superordinal clades of extant placental mammals present by the Late Cretaceous [~90 Mya]"[67], which is more consistent with the evidence from molecular-clock studies. In addition, the existence of a gap in the fossil record for the main supergroups of placental mammals that would be needed to support the molecular-clock data has been considered plausible by some palaeontologists[67–70], although it has been rejected by others[71–73].

In contrast to mammalian timescales, the suggestion of a Precambrian diversification of the main metazoan phyla by Wray *et al.*[59], which also disputed previous estimates from

Box 2 | The chi-square test for the molecular clock



The first relative-rate test for examining the null hypothesis that the amount of evolutionary change in two lineages is equal was proposed by Fitch[31]. It works as follows, with reference to the diagram shown in the figure: if the number of sequence differences between species A, B and C are known, values for the amount of change (*a*, *b* and *c*) can be assigned to the various branches of the phylogenetic tree. For example[31], if there are 8, 19 and 17 differences between sequences A–B, A–C and B–C, then three simple equations can be written: $a + b = 8$, $a + c = 19$ and $b + c = 17$. If the left- and right-hand sides of the first two equations are added up, then $a + b + a + c = 8 + 19$, which is simplified to $2a + b + c = 27$. Because $b + c$ is equal to 17 (according to the third equation above), a must be 5. This procedure can be used to compute the value of b as well, which is 3 in this case. As both A and B evolved from a common ancestor at the same time, the time elapsed on each lineage is the same. Therefore, testing the difference between a and b is equivalent to directly testing the difference in evolutionary rates between lineages A and B. That is, the null hypothesis is $a = b$ under the molecular-clock principle.

In the case illustrated, a and b are not equal (5 and 3, respectively), so to check if this difference is significant, a chi-square test can be carried out (with 1 degree of freedom), using the equation $(a - b)^2/(a + b)$. If the result is greater than 3.841, the molecular-clock hypothesis can be rejected at a 5% significance level. In the example shown, this result is only 0.5, so the clock hypothesis is not rejected. More powerful three-sequence relative-rate tests have been developed recently using evolutionary distances[48], likelihood ratios[104] and non-parametric constructs[37,76]. Tests also have been developed for analysing multiple species (for a review of this see REFS 81,105).

the fossil record, was contested immediately in an independent molecular-clock analysis by Ayala and colleagues[74] in 1998. They used more genes than the original study, and reported consistency between fossil- and molecular-based time estimates. However, other molecular-clock analyses[39,58,75] at that time used an even larger number of proteins and favoured the main conclusions made by Wray *et al.*, although they reported dates that were 20–30% earlier.

Naturally, the validity of the assumptions that underlie molecular clocks is questioned whenever fossil and molecular times disagree, and concerns have been raised not only by palaeontologists, but also by molecular evolutionary biologists. These debates have arisen despite the fact that researchers who have used molecular clocks have made efforts to guard against possible distortions. This has been done by examining the robustness of estimates with and without assuming the existence of a global protein clock[35,41,58–60], and by conducting tests to remove proteins for which evolutionary rates differ significantly between species[35,60].

For example, in 2000, Bromham and colleagues expressed concerns because tests of the molecular clock often have a low power to reject proteins that do not show a clock-like pattern of evolutionary change. This is especially true if the sequence divergence between proteins in different species is small (caused by slow rates of evolution and/or short divergence times). This is because of the lack of a sufficient number of substitutions to statistically distinguish between the presence of true rate difference in different species and the chance occurence of a different number of changes in different species, even when the molecular clock exists[76–78]. However, in 1998, Kumar and Hedges[35] had already attempted to ameliorate some of these problems by increasing the stringency of relative-rate tests to remove proteins that are even mildly non-clock-like. They were able to reject up to 90% of proteins when using stringent molecular-clock tests, but found that divergence-time estimates using proteins that passed increasingly more stringent relative tests were very similar. This was interpreted to show a lack of directional (lineage-specific) rate differences when many proteins are compared for the same set of species[61]. These results supported the predictions of Zuckerkandl and Pauling[9] and Fitch[31] that proteome clocks are more reliable for estimating divergence times, because protein-specific rate differences in different species are expected to average out when many proteins are used.

Relaxed and local clocks

Contemporaneously with clock studies that used genomic datasets, many methodological developments provided ways to 'relax' the assumptions of molecular clocks when estimating species divergence times. These approaches modelled variations in evolutionary rates among lineages, rather than using only those genes and species that passed rate-constancy tests, as was done in the 1980s and 1990s (REFS 35,60,79–81). These new methods were developed to avoid the use of relative-rate tests or the assumption of a global molecular clock when estimating time. The use of such 'local clocks' began in 1989, when Hasegawa and colleagues[82,83] suggested that the molecular clock should be allowed to 'tick' at different rates in different groups of species in a phylogenetic tree.

The use of these methods continued into the mid-1990s, when Uyenoyama[84] and Takezaki *et al.*[80] applied local clocks to estimate species divergence times in a lineage-specific manner, using evolutionary distances between pairs of sequences. The absence of *a priori* knowledge of the parts of the phylogeny that evolve with different rates initially precluded the widespread use of most methods of this type. This was overcome in 1997, when Sanderson[85] used the autocorrelation of evolutionary rates suggested by Gillespie[33,86] to automatically determine changes in evolutionary rates in different lineages, which allowed evolutionary rates to vary from lineage to lineage (relaxed clocks). Gillespie[33,86] had suggested that evolutionary rates themselves might evolve, and that evolutionary rates in ancestral and descendent lineages might be more similar than in more distantly related lineages. To optimize the assignment of the best rate to each lineage, Sanderson[85] proposed that the evolutionary rate difference between ancestral and descendent lineages should be minimized (called the penalized-likelihood method). In 1998, Thorne *et al.*[87] provided new ways of estimating the variation of evolutionary rate among lineages by using a Bayesian framework, which allows for the incorporation of prior information on minimum and maximum divergence times into time-estimation procedures, on the basis of the fossil record (see reviews in REFS 61,88). These and other similar developments[89] helped to usher in a new level of sophistication in estimating times of species divergence and evolutionary rates in different species.

In 2003, Springer *et al.*[90] and Hasegawa *et al.*[91] used relaxed molecular-clock methods to confirm the divergence of mammalian orders before the K/T boundary. These analyses also reconciled the molecular and fossil-based times for the divergence of the mouse and rat, which had been estimated at

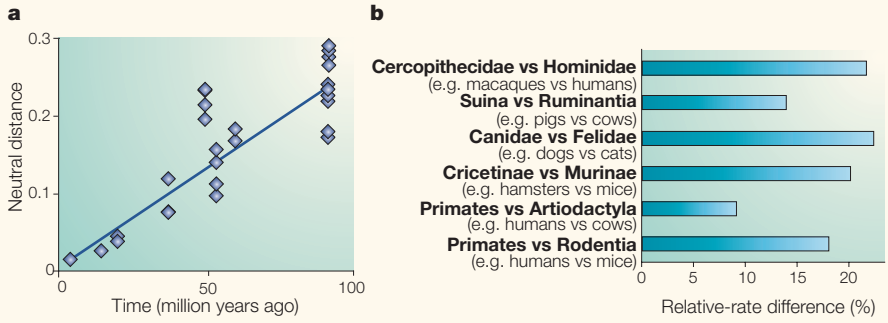
more than 30 Mya in previous large-scale molecular-clock analyses, as compared with a 12-Mya date that is based on the fossil record[15,35,36]. By contrast, the timing of the Cambrian explosion of animal phyla continues to be disputed by palaeontologists[92,93] and molecular evolutionary biologists. For example, Aris-Brosou and Yang[94] and Peterson *et al.*[95] in 2003 and 2004, respectively, used methods that incorporate models that allow evolutionary rates to vary from lineage to lineage, and reported results that are consistent with the Cambrian explosion hypothesis. This contradicted the previous suggestion of a Precambrian diversification that was made by Wray *et al.*[59] on the basis of molecular-clock methods.

However, these studies have proved controversial. Recently, Blair and Hedges[96] have argued that "young time estimates were obtained because fossil calibrations were used as maximum limits rather than as minimum limits" in maximum likelihood analyses. This would have biased the results towards younger estimates in the study by Peterson *et al.*[95]. Blair and Hedges[96] were also critical of the model used by Aris-Brosou and Yang[94] to describe the rate variation, because their analyses yielded many incorrect time estimates for some well-established divergences.

In 2004 and 2005, Hedges and colleagues[97,98] also calculated divergence times of animal phyla using the methods of Thorne *et al.*[99] and Sanderson[100] to relax the assumptions of the molecular clock. Their time estimates are consistent with those obtained a decade earlier by Wray *et al.*[59] using the simpler molecular-clock methods available at that time. This congruence does not imply that there is a global molecular clock, because significant differences in genomic and proteomic evolutionary rates are known to exist among animal phyla, vertebrate classes and mammalian orders, as discussed earlier. Instead, it indicates that there is an unexpected robustness of most inferred species divergence times to violations of the molecular clock, as long as either the rate variation is modelled among species or genes that show detectable departures from molecular clocks are removed. This might explain why most time estimates that were inferred in the early, large-scale multi-protein studies have been validated (with some modifications) by methods that use local or relaxed molecular clocks[61,91,101].

Factors that explain the disparity between divergence times that are estimated from molecular and fossil evidence, but are independent of the validity of molecular clocks, include the number and quality of the fossil-based calibration points, consideration of fossil-calibration uncertainty in determining confidence intervals and potential biases in methods that are used to convert evolutionary distances into time. Scientific discourse on the validity of these concerns and their impact on building reliable molecular timescales is not discussed here, as it is not directly relevant to the discussion of the existence of molecular clocks (see REFS 61,101–105 for more details). In summary, it is clear that the molecular-clock hypothesis is strengthened when molecular- and

Box 3 | **Inferring mammalian mutation rates from DNA substitution rates**



The neutral theory of molecular evolution[20] predicts that the rate at which substitutions accumulate at a particular position in the genome is equal to the mutation rate, as long as all the mutations at that position are strictly neutral (that is, they have no affect on organismal fitness). Third-codon positions for which no mutation can change the encoded amino acid are considered to be strictly neutral positions in mammals[20,81] and are called fourfold-degenerate sites[106]. For the past two decades, the numbers of substitutions at these sites have frequently been used to estimate mutation rates.

However, it is well known that substitution rates at these sites are not equal to mutation rates if there are other factors involved, such as codon-usage biases[40,107] and differences in GC content. Although the effect of codon-usage bias in mammalian genomes is known to be small[108], Kumar and Subramanian[26] found that many human genes have significantly different GC contents to the corresponding mouse genes[109], and that this is true for many mammalian species pairs[26]. These genes are not good candidates for estimating mutation rates. The rate of substitution at the neutral sites might not be equal to the mutation rate if the process of nucleotide substitution was not the same in the evolutionary lineages leading to humans and mice. When this distinction was made (that is, genes with significantly different GC content were removed from the data set), an 18% difference in mutation rate was found between the human and mouse[26]. However, the exclusion of genes on the basis of GC-content difference is not accepted as a valid approach by all molecular evolutionary biologists[27,57].

In general, the analysis of large, genome-scale data sets has led to the observation of similar patterns of rate variation among species by independent research groups. For example, Kumar and Subramanian[26] found that the evolutionary-rate difference between humans and mice exceeded 68% for genes that show significantly different GC contents. This is similar in magnitude to the rate reported by Wu and Li in 1985 (REF. 48), who included all genes and did not test for GC-content differences, although this difference is smaller than that reported by Waterston *et al.*[110], which was based on a genome-wide analysis of the evolutionary rates of ancestral repeat families. Another example is the 20–30% rate difference between humans and Old World monkeys, reported by Yi *et al.*[27] in 2002, which is similar to other reports of 22% (REF. 26) and 30% (REF. 111) rate differences in 2002 and 1996, respectively.

Panel a shows the broad correlation between time of divergence and neutral sequence divergence per lineage between species[26]. However, the extent of rate differences between species is also clear, even when using a large number of genes, as many estimates for the same species divergence times show significant variation. Panel b shows the rate variation between specific groups of species, but with an unexpected twist: the magnitude of rate differences between the main mammalian groups (for example, between primates and rodents) is similar to that seen within these groups (for example, within the primates, between macaques and humans)[26]. At present, there are no biological or theoretical considerations that explain why these differences arise even in closely related species.

Panels a and b are based on data from REF. 26.

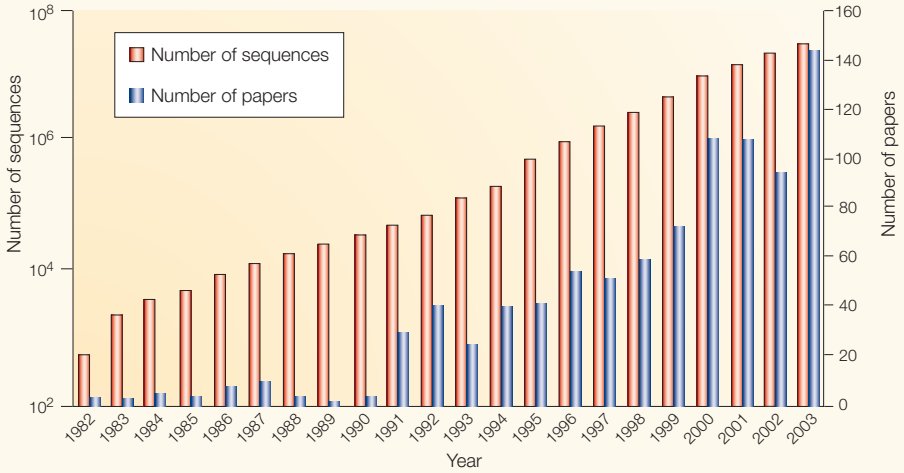


Figure 2 | **Growth in the size of the DNA databanks and the use of the molecular clock.** The graph shows the growth of the US National Institutes of Health GenBank database from 1982 to 2003 (data are from the National Center for Biotechnology Information web site) and the number of research publications that have used molecular-clock methods between 1982 and 2003. Data were obtained by searching the Web of Science online resource with the search term 'molecular clock'. Abstracts and titles of all resulting publications were examined to ensure their appropriateness to molecular clocks. Although the absolute numbers of papers that were identified are shown, it is advisable to focus on the slope of the trend rather than on the absolute numbers, because a larger number of other publications have used time estimates that are based on molecular clocks. For example, 4 large-scale, multi-protein studies published in 1996–1998 (REFS 35,41,59,60) have been cited in more than 150 research publications every year in the past 5 years; not all of these papers are captured with the search criteria used.

fossil-based estimates of species divergence agree[35,61]. However, discrepancies provide the impetus for the development of models to explain the evolution of the evolutionary rate itself, and for determining the relative contributions of genomic and other biological attributes of species to perturbations of the molecular clock.

Still ticking

Molecular clocks have revolutionized evolutionary biology. They have provided a framework for estimating the times of divergence of populations and species, the diversification of gene families and the origin of sequence variations. In the absence of fossil or biogeographical records, molecular-clock techniques remain the only way to infer the timing of gene duplications — which have not been discussed here — and speciation events. The growing impact of the molecular-clock concept is reflected in the exponential increase of the number of research publications that have used this method between 1980 and the present time (FIG. 2). This impact reflects the fact that the power of molecular clocks can be harnessed in purpose-specific ways. For example, fast-evolving mitochondrial genomes and hypermutable nuclear DNA are used to construct rapidly ticking clocks for the fine resolution of events over relatively short timescales, including the evolution of populations and closely related species. Slowly ticking clocks, which are based on nuclear DNA, are used to time deeper divergences, whereas very highly conserved proteins are used to establish the timing of the earliest divergences in the tree of life. With the continuing sequencing of genomes and a better understanding of the magnitude and variability of evolutionary and mutation rates, molecular clocks will continue to have a major effect on the study of evolution.

Sudhir Kumar is at the Center for Evolutionary Functional Genomics, The Biodesign Institute and School of Life Sciences, Arizona State University, Tempe, Arizona 85287-5301, USA.

e-mail: s.kumar@asu.edu

doi:10.1038/nrg1659

1. Hagen, J. B. The origins of bioinformatics. *Nature Rev. Genet.* **1**, 231–236 (2000).
2. Sanger, F. Chemistry of insulin; determination of the structure of insulin opens the way to greater understanding of life processes. *Science* **129**, 1340–1344 (1959).
3. Zuckerkandl, E. & Pauling, L. in *Horizons in Biochemistry* (eds Kasha, M. & Pullman, B.) 189–225 (Academic Press, New York, 1962).
4. Morgan, G. J. Emile Zuckerkandl, Linus Pauling, and the molecular evolutionary clock, 1959–1965. *J. Hist. Biol.* **31**, 155–178 (1998).
5. Margoliash, E. Primary structure and evolution of cytochrome *c*. *Proc. Natl Acad. Sci. USA* **50**, 672–679 (1963).
6. Doolittle, R. F. & Blomback, B. Amino-acid sequence investigations of fibrinopeptides from various mammals: evolutionary implications. *Nature* **202**, 147–152 (1964).
7. Simpson, G. G. Organisms and molecules in evolution. *Science* **146**, 1535–1538 (1964).
8. Aronson, J. D. 'Molecules and monkeys': George Gaylord Simpson and the challenge of molecular evolution. *Hist. Philos. Life. Sci.* **24**, 441–465 (2002).
9. Zuckerkandl, E. & Pauling, L. in *Evolving Genes and Proteins* (eds Bryson, V. & Vogel, H. J.) 97–166 (Academic Press, New York, 1965).
10. Goodman, M. Evolution of the catarrhine primates at the macromolecular level. *Primates Med.* **1**, 10–26 (1968).
11. Goodman, M. Serological analysis of the systematics of recent hominoids. *Hum. Biol.* **35**, 377–436 (1963).
12. Goodman, M., Moore, G. W. & Matsuda, G. Darwinian evolution in the genealogy of haemoglobin. *Nature* **253**, 603–608 (1975).
13. Sarich, V. M. & Wilson, A. C. Rates of albumin evolution in primates. *Proc. Natl Acad. Sci. USA.* **58**, 142–148 (1967).
14. Wilson, A. C., Carlson, S. S. & White, T. J. Biochemical evolution. *Annu. Rev. Biochem.* **46**, 573–639 (1977).
15. Easteal, S., Collet, C. & Betty, D. *The Mammalian Molecular Clock* (R.G. Landes, New York, 1995).
16. Margoliash, E. & Smith, E. L. in *Evolving Genes and Proteins* (eds Bryson, V. & Vogel, H. J.) 221–242 (Academic Press, New York, 1965).
17. Hartl, D. & Dykhuizen, D. A selectively driven molecular clock. *Nature* **281**, 230–231 (1979).
18. Goodman, M., Braunitzer, G., Stangl, A. & Schrank, B. Evidence on human origins from haemoglobins of African apes. *Nature* **303**, 546–548 (1983).
19. Kimura, M. Evolutionary rate at the molecular level. *Nature* **217**, 624–626 (1968).
20. Kimura, M. *The Neutral Theory of Molecular Evolution* (Cambridge Univ. Press, 1983).
21. King, J. L. & Jukes, T. H. Non-Darwinian evolution. *Science* **164**, 788–798 (1969).
22. Laird, C. D., McConaughy, B. L. & McCarthy, B. J. Rate of fixation of nucleotide substitutions in evolution. *Nature* **224**, 149–154 (1969).
23. Britten, R. J. & Kohne, D. E. Repeated sequences in DNA. Hundreds of thousands of copies of DNA sequences have been incorporated into the genomes of higher organisms. *Science* **161**, 529–540 (1968).
24. Diamond, J. M. Taxonomy by nucleotides. *Nature* **305**, 17–18 (1983).
25. Kohne, D. E. Evolution of higher-organism DNA. *Q. Rev. Biophys.* **3**, 327–375 (1970).
26. Kumar, S. & Subramanian, S. Mutation rates in mammalian genomes. *Proc. Natl Acad. Sci. USA* **99**, 803–808 (2002).
27. Yi, S., Ellsworth, D. L. & Li, W. H. Slow molecular clocks in Old World monkeys, apes, and humans. *Mol. Biol. Evol.* **19**, 2191–2198 (2002).
28. Dickerson, R. E. The structures of cytochrome *c* and the rates of molecular evolution. *J. Mol. Evol.* **1**, 26–45 (1971).
29. Ohta, T. & Kimura, M. On the constancy of evolutionary rate of cistron. *J. Mol. Evol.* **1**, 18–25 (1971).
30. Langley, C. H. & Fitch, W. M. An examination of the constancy of the rate of molecular evolution. *J. Mol. Evol.* **3**, 161–177 (1974).
31. Fitch, W. M. in *Molecular Evolution* (ed. Ayala, F. J.) 160–178 (Sinauer Associates, Sunderland, Massachusetts, 1976).
32. Gillespie, J. H. & Langley, C. H. Are evolutionary rates really variable? *J. Mol. Evol.* **13**, 27–34 (1979).
33. Gillespie, J. H. The molecular clock may be an episodic clock. *Proc. Natl Acad. Sci. USA* **81**, 8009–8013 (1984).
34. Takahata, N. On the overdispersed molecular clock. *Genetics* **116**, 169–179 (1987).
35. Kumar, S. & Hedges, S. B. A molecular timescale for vertebrate evolution. *Nature* **392**, 917–920 (1998).
36. Nei, M., Xu, P. & Glazko, G. Estimation of divergence times from multiprotein sequences for a few mammalian species and several distantly related organisms. *Proc. Natl Acad. Sci. USA* **98**, 2497–2502 (2001).
37. Gu, X. & Li, W. H. Higher rates of amino acid substitution in rodents than in humans. *Mol. Phylogenet. Evol.* **1**, 211–214 (1992).
38. Ayala, F. J. Vagaries of the molecular clock. *Proc. Natl Acad. Sci. USA.* **94**, 7776–7783 (1997).
39. Wang, D. Y., Kumar, S. & Hedges, S. B. Divergence time estimates for the early history of animal phyla and the origin of plants, animals and fungi. *Proc. R. Soc. Lond. B* **266**, 163–171 (1999).
40. Tamura, K., Subramanian, S. & Kumar, S. Temporal patterns of fruit fly (*Drosophila*) evolution revealed by mutation clocks. *Mol. Biol. Evol.* **21**, 36–44 (2004).

41. Doolittle, R. F., Feng, D. F., Tsang, S., Cho, G. & Little, E. Determining divergence times of the major kingdoms of living organisms with a protein clock. *Science* **271**, 470–477 (1996).
42. Hedges, S. B. *et al.* A genomic timescale for the origin of eukaryotes. *BMC Evol. Biol.* **1**, 4 (2001).
43. Mayr, E. *Animal Species and Evolution* (Cambridge Univ. Press, New York, 1963).
44. Nei, M. *Molecular Evolutionary Genetics* (Columbia Univ. Press, New York, 1987).
45. Coyne, J. A. & Orr, H. A. *Speciation* (Sinauer Associates, Sunderland, Massachusetts, 2004).
46. Kimura, M. Preponderance of synonymous changes as evidence for the neutral theory of molecular evolution. *Nature* **267**, 275–276 (1977).
47. Miyata, T. & Yasunaga, T. Molecular evolution of mRNA: a method for estimating evolutionary rates of synonymous and amino acid substitutions from homologous nucleotide sequences and its application. *J. Mol. Evol.* **16**, 23–36 (1980).
48. Wu, C. I. & Li, W. H. Evidence for higher rates of nucleotide substitution in rodents than in man. *Proc. Natl Acad. Sci. USA* **82**, 1741–1745 (1985).
49. Britten, R. J. Rates of DNA sequence evolution differ between taxonomic groups. *Science* **231**, 1393–1398 (1986).
50. Li, W. H. So, what about the molecular clock hypothesis? *Curr. Opin. Genet. Dev.* **3**, 896–901 (1993).
51. Martin, A. P. & Palumbi, S. R. Body size, metabolic rate, generation time, and the molecular clock. *Proc. Natl Acad. Sci. USA* **90**, 4087–4091 (1993).
52. Bromham, L. Molecular clocks in reptiles: life history influences rate of molecular evolution. *Mol. Biol. Evol.* **19**, 302–309 (2002).
53. Martin, A. P., Naylor, G. J. & Palumbi, S. R. Rates of mitochondrial DNA evolution in sharks are slow compared with mammals. *Nature* **357**, 153–155 (1992).
54. Bromham, L., Rambaut, A. & Harvey, P. H. Determinants of rate variation in mammalian DNA sequence evolution. *J. Mol. Evol.* **43**, 610–621 (1996).
55. Huttley, G. A., Jakobsen, I. B., Wilson, S. R. & Easteal, S. How important is DNA replication for mutagenesis? *Mol. Biol. Evol.* **17**, 929–937 (2000).
56. Hwang, D. G. & Green, P. Bayesian Markov chain Monte Carlo sequence analysis reveals varying neutral substitution patterns in mammalian evolution. *Proc. Natl Acad. Sci. USA* **101**, 13994–14001 (2004).
57. Ellegren, H., Smith, N. G. & Webster, M. T. Mutation rate variation in the mammalian genome. *Curr. Opin. Genet. Dev.* **13**, 562–568 (2003).
58. Feng, D. F., Cho, G. & Doolittle, R. F. Determining divergence times with a protein clock: update and reevaluation. *Proc. Natl Acad. Sci. USA* **94**, 13028–13033 (1997).
59. Wray, G. A., Levinton, J. S. & Shapiro, L. H. Molecular evidence for deep Precambrian divergences among metazoan phyla. *Science* **274**, 568–573 (1996).
60. Hedges, S. B., Parker, P. H., Sibley, C. G. & Kumar, S. Continental breakup and the ordinal diversification of birds and mammals. *Nature* **381**, 226–229 (1996).
61. Hedges, S. B. & Kumar, S. Genomic clocks and evolutionary timescales. *Trends Genet.* **19**, 200–206 (2003).
62. Cooper, A. & Penny, D. Mass survival of birds across the Cretaceous–Tertiary boundary: molecular evidence. *Science* **275**, 1109–1113 (1997).
63. Springer, M. S. *et al.* Endemic African mammals shake the phylogenetic tree. *Nature* **388**, 61–64 (1997).
64. Hedges, S. B. Afrotheria: plate tectonics meets genomics. *Proc. Natl Acad. Sci. USA* **98**, 1–2 (2001).
65. Archibald, J. D., Averianov, A. O. & Ekdale, E. G. Late Cretaceous relatives of rabbits, rodents, and other extant eutherian mammals. *Nature* **414**, 62–65 (2001).
66. Archibald, J. D. Fossil evidence for a Late Cretaceous origin of 'hoofed' mammals. *Science* **272**, 1150–1153 (1996).
67. Archibald, J. D. Divergence times of eutherian mammals. *Science* **285**, 2031 (1999).
68. Benton, M. J. *Vertebrate Palaeontology* (Blackwell Science, Oxford, 2000).
69. Tavare, S., Marshall, C. R., Will, O., Soligo, C. & Martin, R. D. Using the fossil record to estimate the age of the last common ancestor of extant primates. *Nature* **416**, 726–729 (2002).
70. Clarke, J. A., Tambussi, C. P., Noriega, J. I., Erickson, G. M. & Ketcham, R. A. Definitive fossil evidence for the extant avian radiation in the Cretaceous. *Nature* **433**, 305–308 (2005).
71. Asher, R. J. *et al.* Stem lagomorpha and the antiquity of Glires. *Science* **307**, 1091–1094 (2005).
72. Foote, M. & Sepkoski, J. J. Jr. Absolute measures of the completeness of the fossil record. *Nature* **398**, 415–417 (1999).
73. Feduccia, A. *The Origin and Evolution of Birds* (Yale Univ. Press, New Haven, 1999).
74. Ayala, F. J. & Rzhetsky, A. Origin of the metazoan phyla: molecular clocks confirm paleontological estimates. *Proc. Natl Acad. Sci. USA* **95**, 606–611 (1998).
75. Gu, X. Early metazoan divergence was about 830 million years ago. *J. Mol. Evol.* **47**, 369–371 (1998).
76. Tajima, F. Simple methods for testing the molecular evolutionary clock hypothesis. *Genetics* **135**, 599–607 (1993).
77. Bromham, L., Penny, D., Rambaut, A. & Hendy, M. D. The power of relative rates tests depends on the data. *J. Mol. Evol.* **50**, 296–301 (2000).
78. Bromham, L. & Penny, D. The modern molecular clock. *Nature Rev. Genet.* **4**, 216–224 (2003).
79. Li, W. H. & Tanimura, M. The molecular clock runs more slowly in man than in apes and monkeys. *Nature* **326**, 93–96 (1987).
80. Takezaki, N., Rzhetsky, A. & Nei, M. Phylogenetic test of the molecular clock and linearized trees. *Mol. Biol. Evol.* **12**, 823–833 (1995).
81. Nei, M. & Kumar, S. *Molecular Evolution and Phylogenetics* (Oxford Univ. Press, New York, 2000).
82. Hasegawa, M. & Kishino, H. Heterogeneity of tempo and mode of mitochondrial DNA evolution among mammalian orders. *Jpn. J. Genet.* **64**, 243–258 (1989).
83. Kishino, H. & Hasegawa, M. Converting distance to time: application to human evolution. *Methods Enzymol.* **183**, 550–570 (1990).
84. Uyenoyama, M. K. A generalized least-squares estimate for the origin of sporophytic self-incompatibility. *Genetics* **139**, 975–992 (1995).
85. Sanderson, M. J. A nonparametric approach to estimating divegence times in the absence of rate constancy. *Mol. Biol. Evol.* **14**, 1218–1232 (1997).
86. Gillespie, J. H. *The Causes of Molecular Evolution* (Oxford Univ. Press, New York, 1991).
87. Thorne, J. L., Kishino, H. & Painter, I. S. Estimating the rate of evolution of the rate of molecular evolution. *Mol. Biol. Evol.* **15**, 1647–1657 (1998).
88. Holder, M. & Lewis, P. O. Phylogeny estimation: traditional and Bayesian approaches. *Nature Rev. Genet.* **4**, 275–284 (2003).
89. Yang, Z. A heuristic rate smoothing procedure for maximum likelihood estimation of species divergence time. *Acta Zool. Sinica* **50**, 645–656 (2004).
90. Springer, M. S., Murphy, W. J., Eizirik, E. & O'Brien, S. J. Placental mammal diversification and the Cretaceous–Tertiary boundary. *Proc. Natl Acad. Sci. USA.* **100**, 1056–1061 (2003).
91. Hasegawa, M., Thorne, J. L. & Kishino, H. Time scale of eutherian evolution estimated without assuming a constant rate of molecular evolution. *Genes Genet. Syst.* **78**, 267–283 (2003).
92. Smith, A. B. & Peterson, K. J. Dating the time of origin of major clades: molecular clocks and the fossil record. *Annu. Rev. Earth Planet. Sci.* **30**, 65–88 (2002).
93. Morris, S. C. Evolution: bringing molecules into the fold. *Cell* **100**, 1–11 (2000).
94. Aris-Brosou, S. & Yang, Z. Bayesian models of episodic evolution support a Late Precambrian explosive diversification of the Metazoa. *Mol. Biol. Evol.* **20**, 1947–1954 (2003).
95. Peterson, K. J. *et al.* Estimating metazoan divergence times with a molecular clock. *Proc. Natl Acad. Sci. USA* **101**, 6536–6541 (2004).
96. Blair, J. E. & Hedges, S. B. Molecular clocks do not support the Cambrian explosion. *Mol. Biol. Evol.* **22**, 387–390 (2005).
97. Hedges, S. B., Blair, J. E., Venturi, M. L. & Shoe, J. L. A molecular timescale of eukaryote evolution and the rise of complex multicellular life. *BMC Evol. Biol.* **4**, 2 (2004).
98. Blair, J. E., Shah, P. & Hedges, S. B. Evolutionary sequence analysis of complete eukaryote genomes. *BMC Bioinformatics* **6**, 53 (2005).
99. Thorne, J. L. & Kishino, H. Divergence time and evolutionary rate estimation with multilocus data. *Syst. Biol.* **51**, 689–702 (2002).
100. Sanderson, M. J. r8s: inferring absolute rates of molecular evolution and divergence times in the absence of a molecular clock. *Bioinformatics* **19**, 301–302 (2003).
101. Benton, M. J. & Ayala, F. J. Dating the tree of life. *Science* **300**, 1698–1700 (2003).
102. Hedges, S. B. & Kumar, S. Precision of molecular time estimates. *Trends Genet.* **20**, 242–247 (2004).
103. Glazko, G. V., Koonin, E. V. & Rogozin, I. B. Molecular dating: ape bones agree with chicken entrails. *Trends Genet.* **21**, 89–92 (2005).
104. Muse, S. V. & Weir, B. S. Testing for equality of evolutionary rates. *Genetics* **132**, 269–276 (1992).
105. Felsenstein, J. *Inferring Phylogeny* (Sinauer Associates, Sunderland, Massachusetts, 2003).
106. Li, W. H., Wu, C. I. & Luo, C. C. A new method for estimating synonymous and nonsynonymous rates of nucleotide substitution considering the relative likelihood of nucleotide and codon changes. *Mol. Biol. Evol.* **2**, 150–174 (1985).
107. Sharp, P. M. & Li, W. H. On the rate of DNA sequence evolution in *Drosophila*. *J. Mol. Evol.* **28**, 398–402 (1989).
108. Chamary, J. V. & Hurst, L. D. Similar rates but different modes of sequence evolution in introns and at exonic silent sites in rodents: evidence for selectively driven codon usage. *Mol. Biol. Evol.* **21**, 1014–1023 (2004).
109. Kumar, S. & Gadagkar, S. R. Disparity index: a simple statistic to measure and test the homogeneity of substitution patterns between molecular sequences. *Genetics* **158**, 1321–1327 (2001).
110. Waterston, R. H. *et al.* Initial sequencing and comparative analysis of the mouse genome. *Nature* **420**, 520–562 (2002).
111. Herbert, G. & Easteal, S. Relative rates of nuclear DNA evolution in human and Old World monkey lineages. *Mol. Biol. Evol.* **13**, 1054–1057 (1996).
112. Sarich, V. M. & Wilson, A. C. Immunological time scale for hominid evolution. *Science* **158**, 1200–1203 (1967).
113. Brown, R. H., Richardson, M., Boulter, D., Ramshaw, J. A. & Jefferies, R. P. The amino acid sequence of cytochrome *c* from *Helix aspersa* Muller (garden snail). *Biochem. J.* **128**, 971–974 (1972).
114. Hori, H. & Osawa, S. Evolutionary change in 5S RNA secondary structure and a phylogenic tree of 54 5S RNA species. *Proc. Natl Acad. Sci. USA* **76**, 381–385 (1979).
115. Miyata, T., Yasunaga, T. & Nishida, T. Nucleotide sequence divergence and functional constraint in mRNA evolution. *Proc. Natl Acad. Sci. USA* **77**, 7328–7332 (1980).
116. Hasegawa, M., Kishino, H. & Yano, T. Man's place in Hominoidea as inferred from molecular clocks of DNA. *J. Mol. Evol.* **26**, 132–147 (1987).
117. Rand, D. M. Thermal habit, metabolic rate, and the evolution of mitochondrial DNA. *Trends Ecol. Evol.* **9**, 125–131 (1994).

Acknowledgements

I thank A. Filipski and S.B. Hedges for their insights and for a critical review of different versions of this manuscript. G. Valente and V. Swarna helped in locating references and compiling data; D. Desonie provided editorial support; and S. Subramanian, C. Kuslich, J.E. Blair and A. Briscoe provided scientific feedback on an earlier draft. This work was supported by a research grant from the US National Institutes of Health.

Competing interests statement

The author declares no competing financial interests.

Online links

FURTHER INFORMATION

Molecular Evolutionary Genetics Analysis (MEGA) software package: http://www.megasoftware.net
Multidivtime divergence-time estimation software: http://statgen.ncsu.edu/thorne/multidivtime.html
National Center for Biotechnology Information web site: www.ncbi.nlm.nih.gov
Phylogenetic Analysis by Maximum Likelihood (PAML) software package: http://abacus.gene.ucl.ac.uk/software/paml.html
Sudhir Kumar's Laboratory: http://www.kumarlab.net
Timescale evolutionary database: http://www.timetree.net
Web of Science online resource: http://www.isinet.com/products/citation/wos
Access to this interactive links box is free online.