

## Molecular Clocks and Evolutionary Relationships: Possible Distortions Due to Horizontal Gene Flow

Michael Syvanen

Department of Medical Microbiology and Immunology, School of Medicine, University of California, Davis, CA 95616, USA

**Summary.** This paper discusses recent evidence suggesting that genetic information from one species occasionally transfers to another remotely related species. Besides addressing the issue of whether or not the molecular data are consistent with a widespread influence of horizontal gene transfer, the paper shows that horizontal gene flow would not necessarily preclude a linear molecular clock or change the rate of molecular evolution (assuming the neutral allele theory). A pervasive influence of horizontal gene transfer is more than just consistent with the data of molecular evolution, it also provides a unique explanation for a number of possibly conflicting phylogenies and contradictory clocks. This phenomenon might explain why some protein clocks are linear while the superoxide dismutase clock is not, how the molecular data on the phylogeny of apes and Australian song birds are not necessarily in conflict with those based on morphology, and, finally, why the mycoplasmas have an accelerated molecular clock.

**Key words:** Relative rate test – Globin genes – Cytochromes – Pseudogenes – Coadaptation – Viruses – Viral host range – Bacteria – Songbirds – Apes

### Introduction

I explore the formal possibility that the molecular clock may be experiencing influences that limit its genealogical utility. I refer to the possible movement of genetic information from one species to another. This process is called cross-species gene exchange,

though what I will refer to is transfer of any flucleotide sequence from the germ line of one species to that of another. My discussion focuses on molecular evolution in animals, but it is notable that in plants, sequence convergence is so frequent that it is hard to construct phylogenies from molecular data (Peacock and Boulter 1975). (This convergence could reflect the operation of plant hybridization mechanisms that are nonexistent or much less active in animals. Virally mediated gene transfers may also play a role.) This paper discusses three topics: (1) possible examples of the interspecies exchange of genetic information, (2) how this exchange may influence molecular evolution and the molecular clock, and (3) the resolution of conflicts between molecular and organismal systematics, with specific examples.

### *Cross-Species Gene Transfer*

Only a few cases of whole genes being transferred across animal species barriers have been reported. For reasons given below, I do not think these cases will turn out to be the most important. The possible transfer of parts of genes may prove to be more important. This was first suggested from an analysis of four sequenced mammalian  $\beta$ -globin genes (Syvanen 1984a) where intron-mediated cross-species gene conversion events were proposed. This example is controversial because it can be argued that the conserved “silent” regions of the  $\beta$ -globin gene that I documented are still conserved by functional constraint and not by the cross-species gene conversion events. This pattern of unusually highly conserved regions in the vicinity of introns has also been reported from a comparison of the xxx-y-globin genes from human, chimpanzee, and gorilla (Scott et al. 1984; Slightom et al. 1985). The significance of the conserved gaps in this latter case is statistically

higher than in the case of the four mammalian  $\beta$ -globins (Syvanen, unpublished calculation). It should be pointed out that the authors who reported the  $\gamma$ -globin sequences have chosen to interpret these highly conserved regions in the vicinity of the introns in a different way (Slightom et al. 1985).

If the interpretation of cross-species gene transfer is correct in these cases, then it probably means that this process may significantly influence the molecular clock. This raises the possibility that cross-species gene conversion events involve very short regions of coding sequences, perhaps tens of nucleotides rather than entire genes. If very short regions are transferred during each event, then we can deal with cross-species gene conversion frequencies at a given locus as a counterpoint to mutational frequencies with the final molecular clock being the sum of the two processes. This concept is implicit throughout the following discussion and it distinguishes my treatment of cross-species gene flow from earlier discussions where the unit of transfer has been the whole gene. This is an important point; earlier work that questioned a significant role of cross-species gene transfer in bacterial evolution considered only transfer of entire genes (Ambler et al. 1979; Dickerson 1980; Woese et al. 1980).

Other possible examples of lateral gene transfers involve the retroviruses, transposable elements, and pseudogenes. It seems quite clear that retroviruses, in the course of their evolution, have picked up host genes (Bishop 1983). In a number of cases the viral *onc* gene has been shown to have a homologous prototype in the vertebrate host (Varmus 1984). The homology per se only establishes common ancestry, but given the general distribution of the genes, there is general agreement that the host *onc* gene is ancestral to the viral counterpart (Varmus 1984; Bishop 1985).

Furthermore, there is a growing number of cases indicating that a virus can contribute genetic information to a host species. In fact, the first report of horizontal gene transfer was the probable transfer of a retroviral sequence from Old World monkeys to a recent ancestor of domestic cats (Benveniste and Todaro 1974). Evidence of this nature has continued to accumulate (Duesberg 1983). Even more striking examples have recently been revealed from the complete nucleotide sequence of *Drosophila's copia* element and yeast's Ty element. *Copia* is a mobile element found in insects (*Drosophila* but probably in others (Miller and Miller 1982)] that was identified as a normally functioning mobile element performing as those elements do. Saigo and coworkers (1984; Emori et al. 1985) found that *copia* is homologous to vertebrate retroviruses, containing the homologous *gag*, *pol*, and *env* sequences. Also the Ty element of yeast was initially seen as a mobile genetic element whose subsequent

characterization indicates extensive homology to vertebrate retroviruses (Boeke et al. 1985; Clare and Farabaugh 1985; Garfinkle et al. 1985; Hauber et al. 1985). Conservation of the basic gene organization of these retrovirus-like elements might indicate some horizontal transmission of these sequences in the lineages leading to modern yeast, *Drosophila*, and vertebrates. These examples of transposable elements, with their ability to induce elaborate regulatory mutations (Syvanen 1984b)—now turning out to be closely related to viruses themselves—are important to the new evolutionary theory that incorporates horizontal gene flow.

Other kinds of mobile elements also appear to have experienced transfers across species boundaries. The P-element is found in modern populations of *Drosophila melanogaster* and is responsible for hybrid dysgenesis. This element appears to have been introduced into *Drosophila melanogaster* within the past 50 years and presumably came from a different species (Kidwell 1983).

Possibly relevant to retrovirus-mediated horizontal gene transfer are the structures and inferred evolution of processed pseudogenes. Analyses by Miyata and Hayashida (1981) and Li et al. (1981) of evolutionary distances between pseudogenes and their functional counterparts have clearly shown that sites within pseudogenes evolve much more rapidly than do sites within the functional genes. In general, both the previous coding and synonymous sites within the pseudogene evolve at the same high rate, as would be expected if all functional constraint had been lost to the pseudogene. The mouse pseudo-aglobin gene (called xxxa3) behaves differently (Miyata and Yasuga 1981). Apparently, for part of its history after splitting from the ancestral aglobin gene, xxxa3 diverged under selection for a functional xxxa-globin protein. This behavior, as interpreted by these authors, indicates that after the duplication event, the pseudogene remained active for 7 million years (Myr), presumably through selection, after which the gene became dispensable. An interpretation I offer is that the original splitting event was not a gene duplication event within the species but rather reproductive isolation of the mouse's ancestor into two populations. Then, 7 Myr after divergence, a processed transcript from one group transposed into the ancestor of modern mice. According to this model, there is no need to postulate that the processed pseudogene retained activity after its formation. Other examples are possibly seen with the cytochrome-c pseudogenes in rat (Scarpulla et al. 1982; Scarpulla 1984). These workers isolated eight different clones from a phage lambda charon 4A rat genomic library that cross-hybridized with the yeast cytochrome-c gene. Four of these were either the rat

cytochrome-c gene or closely related pseudogenes. However, the other four isolates hybridized only to the yeast gene (and not to that from rat) and the sequence of at least one appears to represent a highly diverged copy of a yeast-related gene in the rat genome. Finally, Gruskin et al. (1986) have proposed that the “reverse transcriptase calmodulin” gene from chicken [which was initially proposed to have evolved from a pseudogene (Stein et al. 1983)] originated from a remotely related lineage.

### *Molecular Evolution*

These examples may represent either rare occurrences or highly atypical DNA sequences. However I would like to pursue the possibility that they are conspicuous examples of a more subtle and wide spread phenomenon. A number of arguments have been presented suggesting that cross-species gene exchange may be frequent enough to be a major factor in speciation (Hartman 1976; Erwin and Valentine 1984) or macroevolutionary change (Reaney 1976; Syvanen 1985). If these speculations turn out to be correct, then it seems probable that exchange of genetic information among different species should influence the rate at which homologous genes diverge from one another. I will assume in the following discussion that molecular evolution operates primarily through the fixation of neutral mutations (Kimura 1968). In this discussion, “molecular evolution” refers to nucleotide substitution that accumulate through real time for a given lineage. The term “molecular clock” will refer to the rate of divergence between contemporary species which is the empirically determined molecular distance divided by two times the time of divergence.

The argument has been made that transfer of DNA sequences cannot be important because of the data base upon which molecular systematics and the molecular clock are based. This objection has been raised in particular with respect to the use of molecular data in the classification of bacteria (Sanderson 1976; Wilson et al. 1977; Ambler et al. 1979; Dickerson 1980; Woese et al. 1980). That is, if different macromolecules yield the same phylogeny, then the occurrence of cross-species gene transfer must be unlikely. And, of course, the basic observation has been that different macromolecules yield roughly the same phylogenetic picture. However, this objection should not be a serious problem if cross-species gene conversion events occur uniformly throughout an entire genome and if, furthermore, parts of genes, not entire genes, are the usual unit of conversion. That is, failure of the phylogenetic congruency test (Wilson et al. 1977) should not necessarily be a precondition for suspecting the occurrence of cross-species gene transfer.

Another problem for any hypothesis entertaining an extensive influence of cross-species gene exchange concerns the linearity of the molecular clock. If genes could cross species boundaries, it has been argued, then for any given protein, molecular distances versus time of divergence should scatter widely. One might predict that molecular distance for a pairwise comparison would depend on the last time a gene transferred between the two and not on the divergence time of their respective species. Even though this last supposition must be true if cross-species gene transfer had occurred, I will show that the molecular clock data could still be linear.

### *Viral Host Range*

One explanation for this concerns the transfer mechanism itself, i.e., the species host range of the viral vectors. In general very few viruses have the ability to infect highly unrelated species, though many have the ability to infect a number of closely related species. There appears to be a decreasing probability of cross-species viral infection with increasing evolutionary distance. This can be seen more clearly if we consider that group of viruses that are important in human disease (Table 1). Considerable effort has been spent ascertaining the host ranges of these viruses. This information has been obtained either in efforts to find suitable laboratory conditions whereby the viruses can be prepared for study or vaccine production or in epidemiological studies during the search for disease reservoirs. The list of viruses can be considered a representative sampling of animal viruses at least for the purpose of gaining a rough picture of the host-range gradient; it is random in the sense that the sample was compiled independently of the question that I am illustrating. The percentages of these human viruses able to grow in monkeys, rodents, and chickens are 71%, 46% and 12%, respectively.

Therefore, if each animal in nature is linked to its relatives through many viruses that display similar host-range gradients, then this would mean that an animal would have a higher probability of exchanging genes with a closely related species than with an unrelated species. The effect this would have on the molecular clock would be to slow it down, not scatter the data points. This effect slows down the rate of divergence, not the rate of evolution because the clock is measured by comparing contemporary species, not ancestral with descendant species.

### *Coadaptation*

In addition to the constraints imposed by viral host range to lateral gene flow, the problem of coadaptation should also impose constraints. The problem of coadaptation was first introduced by popu-

Human virus	Monkey	Rodent	Chicken	Comments
Epstein-Barr	-	-	-	Transforms monkey cells; infects marmoset
Chicken pox	-	-	-	
Polyomavirus	-	-	-	Some varieties may also grow in other primates; adapted to grow in rodents
Hepatitis type A	-	-	-	Grows in marmoset and chimpanzee
Papillomavirus (wart)	-	-	-	25 varieties; no cross-species growth
Adenovirus	-	-	-	Grows in cultured monkey cells; transforms hamster cells
Rhinovirus (cold)	-	-	-	Grows in chimpanzee
Polio, Types I and III	+	-	-	Some Type II strains grow in rodents
Measles	+	-	-	Adapted to grow on a variety of embryos; can cause persistent infections in hamsters
Echovirus	+	-	-	Nonpathogenic; 5 of 30 serotypes infect mice
Respiratory syncytial	+	-	-	Grows in ferrets
Rubella	+	-	-	
Small pox	+	-	-	Chick embryos
Coxsackie	+	+	-	Grows in newborn mice
Herpes simplex	+	+	-	Grows in chick embryos
Parainfluenza	+	+	-?	
Mumps	+	+	-	Adapted to grow in chick and mouse embryos
Yellow fever	+	+	-	Primary reservoir is monkeys; insect-borne
(see encephalitis)				
Arenevirus (LCM)	+	+	-	Primary reservoir is mouse
Reovirus	+	+	-	Grows in most mammals
Vaccinia	+	+	-	Grows in chick embryos
Influenza type A	+	+	+	
Rabies	+	+	+	Infects all mammals
Encephalitis	+	+	+	Includes about 70 varieties; primary reservoir is birds, rodents, etc.; insect-borne

A plus in this table indicates that the virus has been demonstrated to grow in that host. The virulence of viral growth may range from rabies, which is an infectious disease in most mammals, to the enteroviruses, for which humans are the only known natural host but the virus will grow

lation geneticists and empirical evidence has been provided for it by molecular biologists working with hybrid phages (Furth and Yates 1978; Susskind and Botstein 1978), from reversion studies of bacterial mutants (Ebright et al. 1985), and from interspecific hybrids of the mammalian cytochrome chain (Ferguson-Miller et al. 1976; Osheroff et al. 1983). Coadaptation is seen with genes encoding proteins that physically interact with specific DNA sequences or with other proteins, and that are functionally constrained by their partner's primary sequences. The range of possible amino acid substitutions in such a protein will, in part, be determined by its partner's primary sequence and will in turn influence its partner. This will lead to codivergence of an interacting pair such that they continue to be coadapted with each other. Cann et al. (1984) have presented rate data directly supporting the notion that cytochrome-c and cytochrome oxidase coevolved. As a homologous pair diverges from its common ancestor, a point will be reached where transfer of a functional protein from one species to another will be forbidden. Thus, a coadapted

protein would be expected to show continual divergence with time.

Proteins frequently cited as supporting a constant molecular clock hypothesis include cytochrome-c,  $\alpha$ -globin, and  $\beta$ -globin, all of which are physically complexed to other proteins. This can be contrasted with the superoxide dismutase clock, recently described by Lee et al. (1985), which is grossly non-linear. Superoxide dismutase is an enzyme that is involved in detoxifying  $O_2$  radicals. As such it is a soluble cytoplasmic enzyme for which there is no evidence to suggest, nor reason to believe, that this enzyme specifically interacts with other proteins. Thus, superoxide dismutase from any one species is probably unconstrained to function in other organisms.

#### The Relative Rate Test

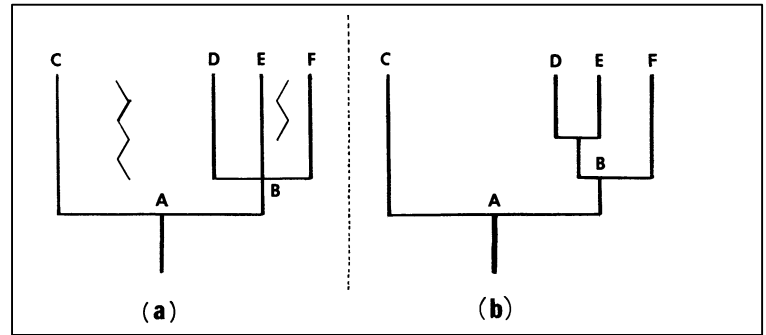
A constant rate of molecular evolution in many different lineages has been supported by the results of the relative rate test. One might argue from this that cross-species gene exchange cannot strongly in-

fluence molecular evolution because, as I have just mentioned, the effect of cross-species gene exchange upon the molecular clock would be to slow it down for those species that are free to exchange genes, but would have no effect upon those that are not free to exchange (for example, should they be geographically isolated). However, there is no contradiction. If cross-species gene exchange is pervasive, it is the rate of divergence, not the actual rate of evolution, which is influenced.

This point can be illustrated if I describe the relative rate test that has been used to measure the uniform average rate and explain how cross-species gene exchange would influence this test. The relative rate test (Sarich and Wilson 1967a) can be illustrated with the aid of the evolutionary tree in Fig. 1a. If the rate of evolution in the four lineages is constant, then the distance of the outside standard C from D, E, and F should be the same. That is, the measured value of percent difference between C and D should be the same as C and E and C and F. Now let us imagine that the lineage leading to the reference species C is unable to exchange genes with the lineages on the right half of the figure (say because of geographical isolation) but that lineages BD and BE have this ability. As I mentioned above, if this situation were occurring, then species D and E would appear more closely related than would be expected by the tree in Fig. 1a. But the question is: how would gene exchange between lineage BD and BE influence the rate of divergence from CA? The answer is: not at all. This would be equivalent to increasing the effective population sizes occupying either lineage BD or BE and, as Kimura (1968) demonstrated when he provided a rational explanation for the molecular clock, population size simply does not influence the rate of fixation of neutral substitutions.

Now let us further imagine that lineages BD and BE are isolated from BF. For the same reasons, AC will still equal  $ABD = ABE = ABF$ . However, D and E will appear closer to each other than either will to F. This is where convergence induced by cross-species gene transfer will be manifested. In fact, if the degree of differences between C, D, E, and F are used to construct the tree (where independent divergence is assumed), then a diagram like the one in Fig. 1b will be obtained. To summarize, cross-species gene exchange would not be expected to affect the rate of evolution; rather, it may affect the rate of divergence.

We have seen that the relative rate test can be satisfied even if there is an influence of cross-species gene transfers occurring among the closely related species. This situation will prevail if the reference species (e.g., C in Fig. 1) is genetically isolated, either by restricted geography or by restricted viral host range. If the reference species and the two species



**Fig. 1.** a Evolutionary relationship among four contemporary species: C, D, E, and F. Wavy line indicates some physical barrier between the evolving lineages. b Deduced evolutionary relationship based upon the degree of genetic similarity if lineages BD and BE experienced genetic exchanges as described in this paper.

under comparison occupy the same geographical range and are exchanging genetic information, we would expect to see occasional failure of the relative rate test.

#### *Molecular Systematics*

Now let it be granted that neither the general outline of the molecular clock nor satisfaction of the relative rate test precludes an influence of cross-species gene transfer. Of course, these points do not encourage belief in its occurrence, but I believe the formal possibility is sufficiently strong to urge restraint in using molecular data in evolutionary classifications. I will illustrate this point with three cases.

#### *Apes*

The first controversy produced by molecular systematics was on the divergence of hominoid primates—chimpanzee, human, gorilla, gibbon, and orangutan. Paleontological evidence along with morphological comparisons had led to the notion of an adaptive radiation having occurred greater than about 15 Myr ago. When protein data were first presented, the divergence times were not only lowered, but human, gorilla, and chimpanzee were now considered to be much more closely related to each other than any were to the orangutan (Sarich and Wilson 1967b).

Since the initial Sarich and Wilson (1967b) findings, their pattern has persisted with other biochemical clocks. For example, problems associated with protein clocks—functional constraint and coadaptation—should be obviated by one of the studies of the hominoid primates. Sibley and Ahlquist (1984) have used the technique of DNA—DNA hybridization with total single-copy DNA to compare the relatedness of hominoid primates. The Sibley and Ahlquist data, as seen with their relative rate

tests, has high precision. In these studies chimpanzees and gorillas were seen to diverge from humans 6 and 8 Myr ago, respectively, whereas the divergence time of the orangutan from the others was assumed to be 13–16 Myr ago. The 16-Myr figure is based on the probable time when apes of Borneo became geographically isolated from the apes of Asia. It probably is not necessary to note that the lineages leading to human, chimpanzee, and gorilla most likely evolved in Africa (and presumably occupied overlapping ranges during their evolution), whereas the molecular data are calibrated using the geographically isolated orangutan. This is analogous to using molecular differences between C and D (Fig. 1) to calibrate the clock in order to measure the time of divergence between D and E. If cross-species gene exchange significantly affects the rate of divergence, then the type of distortion exemplified in Fig. 1b may result. If so, it is formally possible that the hominoid apes did radiate during a single period as was formerly believed. In fact, if paleontological data can eventually resolve the divergent nodes of ape evolution, then we should be able, in this one case, actually to determine the rate of cross-species gene exchange.

#### Song Birds of Australia

Probably the most dramatic result from the Sibley and Ahlquist data concerns the relationships between the passerine songbirds of Australia. Sibley and Ahlquist (1983) found a number of birds on the Australian continent that showed greater DNA similarity among themselves than they did with birds from Eurasia, though morphological systematists had assigned these birds to families on the Eurasian continent. One case involved the Australo-Papuan scrub robins *Drymodes* that had been assigned, on the basis of similar structure, as close relatives of the Eurasian turdine thrushes of the muscicapid family. Sibley and Ahlquist (1981a) showed that the scrub robins have greater molecular similarity to an endemic Australasian group than to the turdine thrushes. The similarities of morphology between the Australian scrub robins and Eurasian thrushes are attributed to convergent evolution, and not to common descent. A similar case is provided by the Australasian sittellas. Based on gross structure, organismal biologists placed these birds as being closely related to the Eurasian nuthatches, whereas the Sibley and Ahlquist (1981b) data identify them with an Australian assemblage. Again we have cases where molecular systematics apparently conflicts with morphological systematics, and the conflict is of the kind that would be produced by cross-species gene exchange among each group of species at the two geographically isolated regions.

#### Mycoplasmas

Geographical isolation is not the only means for the total genetic isolation of a species. The mycoplasmas, a class of parasitic bacteria, may possibly have accomplished this by changing their molecular genetics. In the mycoplasmas, the codon UGA codes tryptophan, whereas this codon in the “universal” code is a translational stop sign (Yamao et al. 1985). In addition, Ryan and Morowitz (1969) have shown that one of the mycoplasmas has fewer tRNA genes than do other organisms. This may mean that mycoplasmas may have lost some of the degenerate codons. These genetic code changes quite possibly may indicate that mycoplasmas participate less actively in universal gene exchange. Negative evidence in support of this possibility has been the failure to find drug resistance plasmids in mycoplasmas, even in those that have been the target of extensive antibiotic use.

These features of mycoplasmas are relevant to the present discussion because of a second unusual feature found in these organisms. Woese et al. (1985) have reported that ribosomal RNAs of the mycoplasmas are unlike those of almost all other bacteria. The ribosomal RNA sequences have not only drifted farther from other bacterial lines, but they also show an unusual degree of variation relative to one another. These authors interpret this result to indicate an elevated rate of neutral site evolution in mycoplasmas that is the result of an elevated spontaneous mutation rate. The elevated mutation rate is part of their hypothesis; it has yet to be measured. As I have stressed throughout this paper, rate of divergence does not necessarily equal the rate of evolution when we are comparing organisms that participate to differing degrees in universal gene exchange. Because mycoplasmas may have partially dropped out of universal gene exchange, I will offer the alternative interpretation. Bacterial ribosomal RNAs of the universal code are conserved, in part, by horizontal gene flow. A gene flow mechanism of the type I have described in the present paper would answer previous arguments against this hypothesis (Woese et al. 1980). Because the mycoplasmas do not participate in this gene flow, their genes are free to diverge, constrained only by their function.

#### Conclusion

Molecular data may not be accurate in determining patterns of descent when two lineages are closely related and share the same geographical range because of the problem of cross-species polynucleotide transfer. One difficulty in accepting a pervasive influence of cross-species gene transfer, besides the entire notion being most strange, is the question of

why such a transfer mechanism would exist in the first place. In this paper the discussion has dealt with molecular clock data, which probably reflect the fixation of neutral events through the action of stochastic processes. Evolution is more than this; it is also a progression of improved forms that are fixed by natural selection. As I have argued elsewhere, natural selection acting on many separate lineages could possibly act to select indirectly for the mechanism of cross-species gene transfer simply through the increase in the rate of acquisitive evolution such a process would effect (Syvanen 1985). Since such a process would not be teleological, one would therefore expect it to act on neutral positions as well.

### Addendum

Since this paper was written there have been three reports of unusually homologous genes between highly unrelated species that possibly constitute examples of gene transfers from eukaryotes to bacteria. The prokaryotic genes involved are a glutamine synthetase II from *Bradyrhizobium japonicum* (Carlson and Chelm 1986), a hemolysin transport protein from *Escherichia coli* (Gerlach et al. 1986), and a self-processing intron from bacteriophage T4 (Michel and Dujon 1986). Another major incongruency between a minimal tree based on morphological characters and one based on protein sequences has been reported where the geographically isolated assemblages show much closer relatedness on the basis of sequence data than they do on the basis of morphological characters. This case involves the relationship of two lemur families to other African primates (Harris et al. 1986). In addition, Wyss et al. (1987) have defended the use of morphological data in tree construction by showing that in mammals, morphological character sets are not less consistent (as measured by the *consistency index*) than are protein sequences.

*Acknowledgments.* I thank Steve McCommas and Sue Greenwald for helpful discussion and for critically reading the manuscript.

### References

Ambler RP, Meyer TE, Kamen MD (1979) Cytochrome C2 sequence variation among the recognised species of purple non sulfur photosynthetic bacteria. *Nature* 278:659—660

Benveniste RE, Todaro GJ (1974) Evolution of c-type viral genes: inheritance of exogenously acquired viral genes. *Nature* 252:456—459

Bishop JM (1983) Cellular oncogenes and retroviruses. *Annu Rev Biochem* 52:301—354

Bishop JM (1985) Viral oncogenes. *Cell* 42:23—38

Boeke JD, Garfinkel DJ, Styles CA, Fink GR (1985) Ty elements transpose through an RNA intermediate. *Cell* 40:491—500

Cann RC, Brown WM, Wilson AC (1984) Polymorphic sites and the mechanism of evolution in human mitochondrial DNA. *Genetics* 106:479—499

Carlson TA, Chelm BK (1986) Apparent eukaryotic origin of glutamine synthetase II from the bacterium *Bradyrhizobium japonicum*. *Nature* 322:568—670

Clare J, Farabaugh P (1985) Nucleotide sequence of a yeast Ty element: evidence for an unusual mechanism of gene expression. *Proc Natl Acad Sci USA* 82:2829—2833

Dickerson RE (1980) Evolution and gene transfer in purple photosynthetic bacteria. *Nature* 283:210—212

Duesberg PH (1983) Retroviral transforming genes in normal cells? *Nature* 304:219—226

Ebright RH, Cossart P, Gisquel-Sanz B, Beckwith J (1984) Mutations that alter the DNA sequence specificity of the catabolite gene activator protein of *E. coli*. *Nature* 311:232—235

Emori Y, Shiba T, Kenaya S, Inouye S, Yuki S, Saigo K (1985) The nucleotide sequence of *copia* and *copia*-related RNA in *Drosophila* virus-like particles. *Nature* 315:773—776

Erwin DH, Valentine JW (1984) Hopeful monsters, transposons and metazoan radiation. *Proc Natl Acad Sci USA* 81: 5482—5483

Evans AS (1976) *Viral infections; epidemiology and control*. Plenum, New York

Fenner F, White DO (1970) *Medical virology*. Academic Press, New York

Ferguson-Miller S, Bautigan D, Chaviano A, Margoliash E (1976) Kinetic control of electron transfer in primate and non primate cytochromes C. *Fed Proc* 35:1605

Furth ME, Yates JL (1978) Specificity determinants for bacteriophage lambda DNA replication. *J Mol Biol* 126:227—245

Garfinkel DJ, Boeke JD, Fink GR (1985) Ty element transposition: reverse transcriptase and virus-like particles. *Cell* 42:507—517

Gerlach JH, Endicott JA, Juranka PF, Henderson G, Sarangi F, Deuchars KL, Ling V (1986) Homology between P-glycoprotein and a bacterial haemolysin transport protein suggests a model for multidrug resistance. *Nature* 324:485—489

Gruskin KD, Smith TF, Goodman M (1987) Possible origin of a calmodulin gene that lacks intervening sequences. *Proc Natl Acad Sci USA* 84:1605—1608

Harris S, Thackeray JR, Jeffrey AJ, Weiss ML (1986) Nucleotide sequence analysis of the lemur  $\beta$ -globin gene family: evidence for major rate fluctuations in globin polypeptide evolution. *Mol Biol Evol* 3:465—484

Hartman H (1976) Speculations on viruses, cells and evolution. *Evol Theory* 3:159—163

Hauber J, Nelböck-Hochstetter T, Feldmann H (1985) Nucleotide sequence and characteristics of a Ty element from yeast. *Nucleic Acids Res* 13:2745—2758

Kidwell MG (1983) Evolution of hybrid dysgenesis determinants in *Drosophila melanogaster*. *Proc Natl Acad Sci USA* 80: 1655—1659

Kimura M (1968) Evolutionary rate at the molecular level. *Nature* 217:624—626

Lee YM, Friedman D, Ayala F (1985) Superoxide dismutase: an evolutionary puzzle. *Proc Natl Acad Sci USA* 82:824—828

Li WH, Gojobori T, Nei M (1981) Pseudogenes as a paradigm of neutral evolution. *Nature* 292:237—239

Michel F, Dujon B (1986) Genetic exchanges between bacteriophage T4 and filamentous fungi? *Cell* 46:323

Miller DW, Miller LK (1982) A virus mutant with an insertion of a *copia*-like transposable element. *Nature* 299:562—564

Miyata T, Hayashida H (1981) Extraordinary high evolutionary rate of pseudogenes: evidence for the presence of selective pressure against changes between synonymous codons. *Proc Natl Acad Sci USA* 78:5739—5743

Miyata T, Yasuga T (1981) Rapidly evolving mouse  $\alpha$ -globin pseudo gene and its evolutionary history. *Proc Natl Acad Sci USA* 78:450—453

Osheroff N, Speck S, Margoliash E, Veerman E, Wilms J, König B, Muijsers A (1983) The reaction of primate cytochromes C with cytochrome C oxidase. *J Biol Chem* 258:5731—5738

Peacock D, Boulter D (1975) Use of amino acid sequence data in phylogeny and evaluation of methods using computer simulation. *J Mol Biol* 95:513—527

- Reaney D (1976) Extrachromosomal elements as possible agents of adaptation and development. *Bacteriol Rev* 40:552—590
- Ryan JL, Morowitz HJ (1969) Partial purification of native sRNA and tRNA cistrons from *Mycoplasma* sp. (KID). *Proc Natl Acad Sci USA* 63:1282—1289
- Saigo K, Kugimiya W, Matsuo Y, Inouye S, Yoshioka K, Yuk S (1984) Identification of the coding sequence for a reverse transcriptase-like enzyme in a transposable genetic element of *Drosophila melanogaster*. *Nature* 312:659—666
- Sanderson KE (1976) Genetic relatedness in the family Enterobacteriaceae. *Ann Rev Microbiol* 30:327—349
- Sanch VM, Wilson AC (1967a) Immunological time scale for hominoid evolution. *Science* 158:1200-1204
- Sarich V, Wilson AC (1967b) Rates of albumin evolution in primates. *Proc Natl Acad Sci USA* 58:142—148
- Scarpulla RC (1984) Processed pseudogenes for rat cytochrome-c are preferentially derived from one of three alternate mRNAs. *Mol Cell Biol* 4:2279—2288
- Scarpulla RC, Agne KM, Wu R (1982) Cytochrome gene-related sequences in mammalian genomes. *Proc Natl Acad Sci USA* 79:739—743
- Scott AE, Heath P, Trusko S, Bayer SH, Prass W, Goodman M, Czelusniak J, Chang L-YE, Slightom JL (1984) The sequence of the gorilla fetal globin genes: evidence for multiple gene conversions in human evolution. *Mol Biol Evol* 1:371—389
- Sibley CG, Ahlquist JE (1981a) The relationship of the Australo-Papuan scrub-robins *drymodes* as indicated by DNA—DNA hybridization. *The Emu* 82:101—105
- Sibley CG, Ahlquist JE (1981b) The relationship of the Australo-Papuan sittellas *Daphoenositta* as indicated by DNA—DNA hybridization. *The Emu* 82:173—176
- Sibley CG, Ahlquist JE (1983) Phylogeny and classification of birds based on the data of DNA—DNA hybridization. In: Johnston RF (ed) *Current ornithology*. Plenum, New York, pp 245—292
- Sibley CG, Ahlquist JE (1984) The phylogeny of the hominoid primates, as indicated by DNA—DNA hybridization. *J Mol Evol* 20:2—15
- Slightom JL, Chang L-YE, Koop BF, Goodman M (1985) Chimpanzee fetal  $\alpha^{\gamma}$  and  $\alpha^{\delta}$  globin gene nucleotide sequences provide further evidence of gene conversions in hominoid evolution. *Mol Biol Evol* 2:370—389
- Stein JP, Munjall RP, Lagace L, Lai EC, O'Malley BH, Means AR (1983) Tissue-specific expression of a chicken calmodulin pseudogene lacking intervening sequences. *Proc Natl Acad Sci USA* 80:6485—6489
- Susskind MM, Botstein D (1978) Molecular genetics of bacteriophage P22. *Microbiol Rev* 42:385—410
- Syvänen M (1984a) Conserved regions in mammalian beta globins: could they arise by cross-species gene exchange? *J Theor Biol* 107:685—696
- Syvänen M (1984b) Evolutionary implications of mobile genetic elements. *Annu Rev Genet* 18:271—293
- Syvänen M (1985) Cross-species gene transfer; implications for a new theory of evolution. *J Theor Biol* 112:333—343
- Varmus HE (1984) The molecular genetics of cellular oncogenes. *Annu Rev Genet* 18:553—612
- Wilson AC, Carlson SS, White Ti (1977) Biochemical evolution. *Annu Rev Biochem* 46:573—639
- Woese CR, Gibson J, Box G (1980) Do genealogical patterns in purple photo synthetic bacteria reflect interspecific gene transfer? *Nature* 283:212—213
- Woese CR, Stackebrandt F, Ludwig W (1985) What are mycoplasmas: the relationship of tempo and mode in bacterial evolution. *J Mol Evol* 21:305—316
- Wyss AR, Novacek MJ, McKenna MC (1987) Amino acid sequence versus morphological data and the interordinal relationships of mammals. *Mol Biol Evol* 4:99—116
- Yamao F, Muto A, Kawauchi Y, Iwami M, Iwagami S, Azumi Y, Osawa S (1985) UGA is read as tryptophan in *Mycoplasma capricolum*. *Proc Natl Acad Sci USA* 82:2306—2309