

# Cladistic analyses of molecular characters: The good, the bad and the ugly

Maximilian J. Telford

*University Museum of Zoology, Department of Zoology, Downing Street, Cambridge CB2 3EJ, UK;*  
email: [mt281@cam.ac.uk](mailto:mt281@cam.ac.uk)

Keywords: molecular synapomorphy, phylogeny, rare genomic change, RGC, Metazoa, cladistics

## Abstract

Molecular cladistics is an emerging discipline in which any heritable molecular characteristic can be treated in the same way that a traditional cladist would treat a morphological character. Taxa that share specific derived molecular characters (synapomorphies) are recognized as more closely related to each other than they are to other taxa without these characters. Herein, I point out that molecular characters are susceptible to the same problems of homoplasy and uncertain polarity as morphological characters and illustrate these problems (and point towards a general solution) using examples from the Metazoa.

## Contents

Introduction	93
Homoplasy: real and apparent	94
Mitochondrial genetic codes	94
Mitochondrial gene order in birds	95
Indels in the EF1alpha gene and the position of the acoel flatworms	95
Contradictory stories from gene fusions	95
Introns	96
Apparent homoplasy due to secondary loss of a character: novel domain combinations	96
Polarity	
Mitochondrial genetic codes: the echinoderms and hemichordates	97
Hox gene signatures: the dicyemid mesozoans and the priapulids	98
Conclusions	99
Acknowledgments	99
References	99

## Introduction

The past 15 years have seen the steady expansion of the field of molecular phylogenetics. Molecular genetic data are accumulating ever faster and methods of phylogenetic analysis are steadily improving. One strand of molecular phylogenetics that is becoming increasingly appreciated does not deal with aligned sequences of nucleotides or amino acids in the customary manner. Rather, this complementary approach treats heritable characteristics of the genome – and the variety of characters used is impressive – in the same way that a cladist would use a morphological character. In such molecular cladistic studies, taxa sharing a novel or derived character (synapomorphies) are inferred to be more closely related to each other than they are to taxa lacking this character. As in morphology-based cladistics, the sharing of a primitive character (symplesiomorphy) cannot be taken as an indicator of close relationship.

Examples of the diverse molecular characteristics used to date have been covered in an excellent review by Rokas and Holland (2000), who also touched on some of their shortcomings. I want to examine in more detail the potential pitfalls when using Molecular Synapomorphies (referred to as Rare Genomic Changes – RGCs – by Rokas and Holland). I will show that, perhaps unsurprisingly, some suffer from the same problems as morphological characteristics but that their strength, relative to morphological characters, lies in the ease of homology assessment of molecular characters.

### Homoplasy: real and apparent

For the purposes of this discussion I would like to make a distinction between real and apparent homoplasy. Homoplasy entails the parallel evolution of a particular character state in two unrelated lineages through convergence, parallelism or reversion. In truth, because most morphological characters arise in an immensely complicated fashion involving genotypic changes leading to alterations in developmental pathways and finally to a phenotype, it is highly unlikely that a *truly* identical morphological character could arise independently in separate lineages. Characters only appear identical but are in fact unrelated leading to apparent homoplasy. The homoplasy problem arises because we can rarely know enough about a morphological character (for example we cannot investigate the molecular developmental basis for every diagnostic bristle on a fly's leg) to ascertain identity. Morphological analyses generally have to rely on overall strength of phylogenetic signal within the totality of the data set to reveal homoplasy in (hopefully) the minority of characters.

Traditional molecular phylogenetic efforts suffer from the opposite problem; it is usually easy to know with certainty that one is dealing with an identical character – a particular amino acid at a particular position within a protein, for example. However, due to the low diversity of potential character states (4 nucleotides or 20 amino acids) real homoplasy is prevalent because, given independent changes at a specific position in two taxa, there is a 1 in 4 or 1 in 20 chance, respectively, that the novel state will be identical.

The great strength of molecular synapomorphies ought to be that they avoid both of these problems. Being molecularly based it is easy to be certain of identity between characters in different species – with molecular characters, unlike morphological characters with their hidden genotypic and developmental layers, “What You See Is What You Get” –, moreover, because these characters are of a higher order of complexity than simple substitutions between nucleotides or amino acids, real homoplasy should not be a problem. What we find, however, are various surprising cases of real homoplasy affecting these molecular characters and I discuss some of these below.

### Mitochondrial genetic codes

Changing a genetic code, even that of the small genome of the eukaryotic mitochondrion, is a hugely significant event. It cannot happen suddenly, for example by changing the specificity of a tRNA synthetase or the anticodon of a tRNA, as tens or hundreds (or in the case of a nuclear genome many thousands) of amino acids would be swapped and this would affect almost every protein and would certainly be lethal. The more likely mechanism for change is, rather, the gradual, complete loss of a certain codon followed by its reassignment and gradual reintroduction in its new guise (see e.g. Castresana et al., 1998). Even so, the process is clearly immensely complex and seems unlikely to happen often. To make *convergent* codon reassignment even less likely, there are 64 codons with the potential to change and each can be reassigned from the incumbent amino acid (or stop codon) to any of 19 others (or stop codon). If all changes were equally likely, even allowing that a change has happened, there is only a 1 in 1280 chance that the same change would occur in two independent lineages.

Considering the unlikeliness of convergent codon reassignment, it is surprising to discover that there are multiple instances of this within eukaryotic mitochondria and even nuclei. Perhaps one of the most striking examples is the convergent reassignment of *two* codons in the echinoderms and the rhabditophoran flatworms (Telford et al., 2000). In both taxa the same reassignments are seen: the codon AAA has been reassigned from coding for Lysine to coding for Asparagine and AUA reassigned from Methionine to Isoleucine. Other clear instances of convergence include UAA and UAG codons being independently reassigned from STOP to Glutamine in the nuclear genomes of diplomonads, several ciliates and the green alga *Acetabularia acetabulum* and UGA reassigned from STOP to Tryptophan five separate times in the mitochondria of various eukaryotic groups (Knight et al., 2001).

Fortunately, thanks to other sources of phylogenetic information, we are not misled into believing that the echinoderms and the rhabditophoran flatworms, for example, are sister-groups, but consideration of mitochondrial genetic codes highlights

the need to understand how a trait evolves if one is to use it as a molecular synapomorphy. Clearly all changes in genetic code are not equal; due to whatever constraint or for whatever adaptive reasons, certain changes are more likely to occur than others.

#### *Mitochondrial gene order in birds*

Convergent evolution of novel mitochondrial gene arrangements is also, on the face of it, an unlikely event. There are 13 proteins, 2 rRNAs, and 22 tRNAs in the circular metazoan mitochondrial genome, which can theoretically be arranged in  $2 \times 10^{52}$  different ways. Despite the vanishingly small probability of independent adoption of any specific novel arrangement, it has been shown that a rearrangement of three contiguous genes (Proline-tRNA, ND6, and Glutamic acid-tRNA) relative to the mitochondrial control region has occurred on at least 4 separate occasions during the evolution of birds (Mindell et al., 1998). Parallel inversions of sections of plant chloroplast genomes have also been reported (Hoot and Palmer, 1994). Of course, shifting a single fragment from position A to position B has a much higher likelihood of convergence than suggested by the probability quoted above, but the tendency for repeated evolution of a particular novel arrangement might also point to an underlying constraint. Importantly, this constraint could only be inferred through sufficiently dense sampling and through some prior knowledge of bird phylogeny. An *a priori* assumption that parallel changes in mitochondrial gene order are hugely unlikely could easily have caused us to reconstruct an incorrect phylogeny.

#### *Indels in the EF1alpha gene and the position of the acoel flatworms*

The acoel flatworms have caused a lot of controversy in recent years. Although, in common with other flatworms, the acoels lack a coelom and complete gut, the two groups share no other obvious features. Ribosomal RNA phylogenies suggest the two groups are unrelated and Telford et al. (2000) showed that the acoels do not share the rhabditophoran flatworm mitochondrial genetic code changes

discussed above instead sharing the standard invertebrate code with most other animals. The absence of the rhabditophoran novelties demonstrates clearly that the acoels are not derived from within the Rhabditophora. In direct contradiction of this result, Berney et al. (2000) found a short peptide motif in the EF1alpha gene shared by one group of rhabditophoran flatworms and an acoel, *Convoluta roscoffensis*. Based on this observation these authors suggested that the acoels are in fact derived from within the Rhabditophora.

Once again, it turns out that wider sampling was able to resolve this contradiction and again showed the importance of understanding the evolution of the character. Littlewood et al. (2001) sequenced the same region of EF1alpha from 3 further species of acoel and found that all three lacked the peptide motif. On the other hand, a menagerie of other metazoans (molluscs, annelids, nematodes and chordates) did have the character or a close approximation. This particular character, although initially persuasive of a link between acoels and rhabditophoran flatworms, proved unreliable due to homoplasy.

#### *Contradictory stories from gene fusions*

In the example I describe below, we have two contradictory molecular characters, one of which logically must be homoplastic. On the one hand, the  $\beta$ -thymosin gene, which exists as a single short peptide in the majority of metazoans, has been shown to be triplicated and serially linked in nematodes and arthropods (Manuel et al., 2000). This unusual character seemingly gives strong support to the Ecdysozoa Hypothesis that postulates a clade of moulting animals including nematodes and arthropods. On the other hand, the Glutamyl and Prolyl aminoacyl tRNA synthetases, which are separate genes in most taxa including the yeast *Saccharomyces* and the plant *Arabidopsis* and also in nematodes, are found to be fused into a single, bifunctional protein in both arthropods and vertebrates (Berthonneau and Mirande, 2000 and unpublished observations). This finding seems to suggest that, contrary to the Ecdysozoa Hypothesis, flies are more closely related to the vertebrates than they are to the nematodes.

Clearly one of these two characters must be homoplastic. Perhaps arthropods and nematodes are indeed ecdysozoan sister-groups, in which case either nematodes have reverted to a primitive, unfused state for the two RNA synthetases or arthropods and vertebrates have fused their genes convergently. On the other hand, if the coelomate arthropods and coelomate vertebrates are more closely related than either is to the pseudocoelomate nematodes, then the triplication of  $\beta$ -thymosin is either convergent in flies and worms or has been secondarily lost in vertebrates. Further sampling of sister taxa is needed to discover which one of these two characters is homoplastic. Meanwhile, the point is made that, although both characters seem on the face of it to be the result of very rare genetic events, one or other of them must indeed have occurred convergently.

### *Introns*

Perhaps less surprising than the previous examples, it is becoming increasingly clear that using the presence or absence of an intron as a molecular synapomorphy is not always reliable. In some cases (but by no means all) when adequately sampled, it becomes clear that introns are readily and repeatedly lost and hence prove less reliable as indicators of phylogenetic relationship than might once have been thought (Krzywinski and Besansky, 2002; Wada et al., 2002).

### *Apparent homoplasy due to secondary loss of a character: Novel domain combinations*

King and Carroll (2001) have demonstrated that the joining of the EGF and the Tyrosine Kinase domains in a single protein is a combination found only in the clade comprising the Metazoa and their possible sister-group; the choanoflagellates. This observation supports the presumed link between the choanoflagellates and Metazoa. This work led us (unpublished collaboration with Rob Russell and Patrick Aloy, EMBL, Heidelberg) to look for novel combinations of protein domains in the completely sequenced genomes of human, fly and nematode in the hope that unique domain combinations might

provide characters to test the Ecdysozoa Hypothesis (see above). If the Ecdysozoa Hypothesis is correct then we might expect flies and worms to share some unique combinations of protein domains not seen in humans or out-groups (fungi and plants). If the older, Coelomate Hypothesis is correct, which links flies and humans to the exclusion of nematodes, then flies and humans might share certain unique domain combinations.

What we find is a large number of domain combinations common to flies and humans and absent from worms and the out-groups mentioned. If we look for all combinations of pairs of protein domains present in two out of the three metazoans and absent in the out-groups we find 20 shared by humans and worms, 29 shared by flies and worms, and 276 shared by humans and flies. If we look for all combinations of three domains then we find none shared by humans and worms, 3 shared by flies and worms, and 12 shared by humans and flies. Should this destroy our confidence in the Ecdysozoa Hypothesis? We think probably not, at least not without closer examination of these results.

The problem here lies with *Caenorhabditis elegans*, which is certainly a very derived animal. *C. elegans* is a model species selected to reproduce very rapidly in the laboratory. It is small, has a small and constant cell number and has an atypical mode of development and an atypical (derived) genome associated with these lifestyle constraints. Examination of the genes in its genome reveals various oddities. Of its genes with clear homologs in other taxa, most are fast evolving; it has lost several of its *Hox* genes and, on the other hand, it has evolved a large number of new genes, for example those associated with chemoreception which, it has been postulated, were probably evolved in order to make up for in hardwiring what the worm lacks in brainpower.

The point of these observations is that, upon reflection, we can see that the genes and the genome of *C. elegans* are highly unusual and highly derived. The significance of this is that many of its genes are likely to have been secondarily lost and this fact makes the analyses I presented above uninformative and, at present, unreliable because the distribution of character states we observe could reasonably be explained by the secondary loss of

many of these characters in the ancestry of the nematode.

What we see in all of these cases – and this is equally true of morphological characters – is that two factors are important in order not to be fooled by homoplasy, real or apparent. First and foremost, it is desirable to sample widely. Only through sampling many lineages within the eukaryotes can we see that certain changes in mitochondrial genetic code happen relatively frequently, and only through a broad sampling of the Metazoa can we understand where and when these code changes occurred in this clade. Secondly, and hopefully this follows on from the first, one must understand the character and its evolution. We should ask whether there are any underlying molecular reasons why we might expect a certain character to evolve repeatedly in unrelated lineages or, as in the case of the *C. elegans* genome, to be secondarily lost?

The case of the secondarily lost genes of *C. elegans* highlights a further important consideration when performing such studies; a character that can be assessed in both a primitive and a derived state is generally preferable to one that is coded as absent or present. Molecular characters seem likely to be particularly susceptible to secondary loss (cases of missing genes are commonplace) and furthermore, unless dealing with completely sequenced genomes, it is often all but impossible to *prove* a certain character is absent in a genome because, as pointed out by Rokas and Holland (2000), absence of evidence is not the same as evidence of absence. The inability to clone a certain DNA sequence does not demonstrate it does not exist.

### Polarity

As emphasised above, only the sharing of *derived* characters is informative regarding phylogenetic relationships. Shared *primitive* characters simply place those species that possess them within the larger clade of all taxa that have (or had and subsequently lost) those characters. To make this point absolutely clear to anyone not well versed in cladistics there follows an example. Consider the relationship between a rabbit, a horse and a lizard about which we know only that the rabbit shares a

five-toed foot with the lizard and warm bloodedness with the horse.

Considering first the five-toed foot; this is in fact a primitive character that is shared by all quadrupeds: rabbits, lizards, frogs and even the ancestor of horses. This character cannot help in determining the relationship of the rabbit to any of these taxa in particular. Warm bloodedness, on the other hand, is a derived character peculiar to the rabbit and the horse and absent in the lizard. We can see that the five toed foot is a primitive character because it is present in the frog which, being a more ancient lineage than the three we are considering, we assume is primitive. Following the same reasoning, warm bloodedness should be a derived character because it is absent in the frog. The use of the early diverging frog to see which character state (cold or warm bloodedness) is primitive and which (therefore) derived is a procedure known as out-group comparison. This is simply an argument from parsimony because, in this example, if warm bloodedness were the primitive state it would have to have evolved once and then been lost twice; once in frogs and once in lizards; the alternative interpretation requires that it has evolved just once in mammals.

Character polarity is just as important when using molecular characters and below I give two examples of when this has been problematic.

### *Mitochondrial genetic codes: The echinoderms and hemichordates*

The hemichordates, as their name suggests, were long thought to be more closely related to the chordates than to the third group of deuterostomes; the echinoderms. Recent molecular phylogenies deny this close relationship between chordates and hemichordates and instead position the hemichordates as the sister-group of the echinoderms (Bromham and Degnan, 1999; Halanych, 1996). One of the changes in mitochondrial genetic code discussed above has been cited in support of this alternative idea of a hemichordate/echinoderm clade: the re-assignment in echinoderms of the codon AUA from coding for Methionine (Met) to coding for Isoleu-

cine (Ile) has recently been shown to co-occur in hemichordates (Castresana et al., 1998).

Upon closer inspection, however, it is not clear that this character (AUA = Ile) is most parsimoniously interpreted as a derived character. Knowing that AUA also codes for Ile in the Cnidaria, a close out-group of the Bilateria, one can readily show that it is equally parsimonious to assume that either AUA = Met, or AUA = Ile is the primitive state. Each of these solutions requires 2 changes as follows: (i) if *AUA = Met* is primitive within the Bilateria this requires the change AUA = Ile to AUA = Met at base of the Bilateria and the reversal AUA = Met to AUA = Ile in the echinoderms, (ii) if *AUA = Ile* is primitive within the Bilateria this requires the convergent evolution of AUA = Ile to AUA = Met twice; once in the chordates and once in the protostomes (Telford et al., 2000). In short, the co-occurrence of the character AUA = Ile in echinoderms and in hemichordates has not been conclusively shown to be a shared *derived* character providing additional evidence linking these two groups.

To update this story, it seems increasingly likely (see also above) that the acoelomorph flatworms are indeed the earliest branching bilaterians (Jonelius et al., 2002; Ruiz-Trillo et al., 1999). If this is true, then the observation that, in the basally branching acoels, AUA = Met suggests that the character AUA = Met is primitive. It would then follow that the character AUA = Ile could indeed be most parsimoniously interpreted as a synapomorphy linking hemichordates and echinoderms.

#### *Hox gene signatures; The dicyemid mesozoans and the priapulids*

Another instance where the polarity has not been properly considered is the case of a *Hox* gene of the dicyemid mesozoan. It had been previously shown that certain peptides found in the homeoboxes of specific *Hox* genes were associated with each of the three great clades of bilaterian: deuterostomes, ecdysozoans and lophotrochozoans (de Rosa et al., 1999). Kobayashi et al (1999) were able to show that a *Hox* gene found in a dicyemid mesozoan contained one of the peptides characteristic of the *Lox5* gene that is specific to the lophotrochozoan clade. From this they deduced that the mesozoan

was derived from within this clade rather than being the basally branching metazoan suggested by its simple morphology.

What had not been considered, however, was the polarity of these *Hox* signatures (Telford, 2000). In the absence of an out-group possessing a homolog of this gene, each of the three possible states of this character (deuterostome like, ecdysozoan-like or lophotrochozoan-like) can equally parsimoniously be considered primitive. If this lophotrochozoan *Hox* signature was in fact the primitive state for this character (the ecdysozoan and deuterostomian character states both being derived,) then the discovery of the LOX5 peptide in the mesozoan simply shows that the dicyemids are metazoans (although it does exclude them from the ecdysozoan and deuterostomian crown-groups). The metazoan status of mesozoans was not in dispute.

Fortunately, because the *Hox* genes arose by duplication, one or more of the closely related *Hox* genes can act as a proxy out-group. When this approach is taken, it is possible to identify certain derived amino acids within the *Lox5* gene that are also shared by the mesozoan gene. This approach gives limited support for the derivation of mesozoans from within the Lophotrochozoa but this result only becomes credible when the polarity of the character is established (Telford, 2000).

To emphasise the significance of this approach, one can consider the relationship of the priapulid worms within the protostomes. The priapulids clearly share a *Ubx/abdA*-like gene with other protostomes (called *Lox2* and *Lox4* in the lophotrochozoan clade) but, consistent with the Ecdysozoa Hypothesis, the priapulid gene is most similar to the arthropod *Ubx* gene. When one looks for synapomorphic 'signature' amino acids within this gene, however, one finds that the priapulids share only a single *derived* residue with the arthropods, all other amino acids being either primitive, i.e., interpreted through out-group comparison as being present in the ancestor of protostomes, or specific to the priapulid.

By contrast, the same comparisons offer overwhelming support (from polarised residues) for the notion that both Platyhelminthes (previously thought to be basal bilaterians) and Brachiopoda (previously widely believed to be related to the deuterostomes) are in fact lophotrochozoans (Telford, 2000).

## Conclusions

The theme linking all of these potential problems with molecular characters is the need to understand how they evolve. In particular, it is vital to consider the potential problems of homoplasy that are unexpected in what appear to be highly complex characters. Understanding the evolution of characters can almost always be best achieved by ensuring adequately dense sampling. The closely related issues of polarity and of synapomorphy versus symplesiomorphy are also easily overlooked as we have seen.

What I do not want to do is give the impression that molecular synapomorphies are essentially unreliable. On the contrary, I believe that they are a very valuable source of phylogenetic information, their great strengths lying firstly in the ease of ascertaining homology between features of DNA, and secondly in the almost limitless variety of features of differing complexity that can be used, only a few of which have been discussed here. Neither do I want to assert the superiority of molecular over morphological studies. What I will emphasise, however, is the great advantage of using clearly homologous characters and would suggest that shifting emphasis from quantity to quality of morphological characters might well bring rewards (see e.g. Jenner, 1999; Jenner and Schram, 1999).

This molecular cladistic approach to phylogenetics is set to become an increasingly useful tool thanks to the rapidly expanding genomic databases and the increasingly powerful bioinformatic tools we have for analysing them.

## Acknowledgements

Many thanks to Dr Chuck Cook, Prof. Frederick Schram and Dr Ronald Jenner for helpful comments on the manuscript. Thanks also to Dr Rob Russell and Dr Patrick Aloy for collaboration on the protein domain combination work referred to in this manuscript.

## References

- Berthonneau E, Mirande M. 2000. A gene fusion event in the evolution of aminoacyl-tRNA synthetases. *FEBS Lett.* 470:300-304.
- Berney C, Pawlowski J, Zaninetti L. 2000. Elongation factor 1-alpha sequences do not support an early divergence of the Acoela. *Mol. Biol. Evol.* 17:1032-1039.
- Bromham LD, Degnan BM. 1999. Hemichordates and deuterostome evolution: robust molecular phylogenetic support for a hemichordate + echinoderm clade. *Evol. Dev.* 1:166-171.
- Castresana J, Feldmaier-Fuchs G, Pääbo S. 1998. Codon reassignment and amino acid composition in hemichordate mitochondria. *Proc. Natl. Acad. Sci. USA* 95:3703-3707.
- de Rosa R, Grenier JK, Andreeva T, Cook CE, Adoutte A, Akam M, Carroll SB, Balavoine G. 1999. *Hox* genes in brachiopods and priapulids: implications for protostome evolution. *Nature*, 399:772-776.
- Halanych KM. 1996. Convergence in the feeding apparatuses of lophophorates and pterobranch hemichordates revealed by 18S rDNA: an interpretation. *Biol. Bull.* 190:1-5.
- Hoot SB, Palmer JD. 1994. Structural rearrangements, including parallel inversions, within the chloroplast genome of *Anemone* and related genera. *J. Mol. Evol.* 38:274-281.
- Jenner RA. 1999. Metazoan phylogeny as a tool for evolutionary biology: current problems and discrepancies in application. *Belg. J. Zool.* 129:245-262.
- Jenner RA, Schram FR. 1999. The grand game of metazoan phylogeny: rules and strategies. *Biol. Rev.* 74:121-142.
- Jondelius U, Ruiz-Trillo I, Bagaña J, Riutort M. 2002. The nemertodermatids are basal bilaterians and not members of the Platyhelminthes. *Zool. Scripta* 31:201-215.
- King N, Carroll SB. 2001. A receptor tyrosine kinase from choanoflagellates: Molecular insights into early animal evolution. *Proc. Natl. Acad. Sci. USA* 98:15032-15037.
- Knight RD, Freeland SJ, Landweber LF. 2001. Rewiring the keyboard: evolvability of the genetic code. *Nature Rev. Genetics* 2:49-58.
- Kobayashi M, Furuya H, Holland PWH. 1999. Dicyemids are higher animals. *Nature* 401:762.
- Krzywinski J, Besansky NJ. 2002. Frequent intron loss in the *White* gene: a cautionary tale for phylogeneticists. *Mol. Biol. Evol.* 19:362-366.
- Littlewood DTJ, Olson PD, Telford MJ, Herniou EA, Riutort M. 2001. Elongation Factor 1-alpha sequences alone do not assist in resolving the position of the Acoela within the Metazoa. *Mol. Biol. Evol.* 18:437-442.
- Manuel M, Kruse M, Müller WEG, Le Parco Y. 2000. The comparison of  $\beta$ -thymosin homologues among Metazoa supports an arthropod-nematode clade. *J. Mol. Evol.* 51:378-381.
- Mindell DP, Sorenson MD, Dimcheff DE. 1998. Multiple independent origins of mitochondrial gene order in birds. *Proc. Natl. Acad. Sci. USA* 95:10693-10697.
- Rokas A, Holland PWH. 2000. Rare genomic changes as a tool for phylogenetics. *T.R.E.E.* 15:454-459.
- Ruiz-Trillo I, Riutort M, Littlewood DTJ, Herniou EA, Bagaña J. 1999. Acoel flatworms: earliest bilaterian metazoans not members of Platyhelminthes. *Science* 283:1919-1923.
- Telford MJ. 2000. Turning *Hox* 'signatures' into synapomorphies. *Evol. Dev.* 6:360-364.
- Telford MJ, Herniou EA, Russell RB, Littlewood DTJ. 2000.

Changes in mitochondrial genetic codes as phylogenetic characters: two examples from the flatworms. *Proc. Natl. Acad. Sci. USA* **97**:11359-11364.

**Wada H, Kobayashi M, Sato R, Satoh N, Miyasaka H,**

**Shirayama Y. 2002.** Dynamic insertion-deletion of introns in deuterostome *EF1 alpha* genes. *J. Mol. Evol.* **54**:118-128.

Received: 11 June 2002