

Homology evolving

David P. Mindell and Axel Meyer

The defining criterion for homology has been common ancestry ever since darwinian evolution replaced fidelity to archetype as the explanation for similarity among traits of organisms. Although, in some cases, an alternative criterion of similarity in developmental control or other underlying traits can be helpful in diagnosing homology, evidence for dissociation among traits at different levels of organization (e.g. genotype and phenotype) and for frequency of reticulate (lateral) evolution accentuate the problems inherent in using similarity as the defining criterion. Growing awareness that character history and organismal lineage history can and do differ puts homology definition and assessment center stage in our understanding of the evolution of genomes.

The term 'HOMOLOGY' (see Glossary) has been in use for over 150 years, with different meanings over time for different groups of biologists. Although first used to describe 'the same organ under every variety of form and function', most biologists now use homology to denote hypotheses of evolutionary relationship among traits of organisms. This is justified by the fact that darwinian evolution has replaced FIDELITY TO ARCHETYPE as the explanation for similarity among traits in different species. Relationships of homology among traits stem from both lineage diversification (speciation or CLADOGENESIS) and character evolution within individuals (e.g. duplication of genes). Relationships of homology are crucial in evolutionary studies, because they denote alternative results of evolutionary processes in different lineages. In theory, all kinds of traits can be assessed for homology relationships, including the structure and function of molecular sequences, chromosomes, genomes, cells, body parts and behaviors, as well as their mechanisms of development and integration. In practice, however, the task remains difficult.

Assessing homology for characters involves a set of related questions. What is a character? What defines character homology? How is homology best diagnosed? Which characters are associated across levels of organization through cause–effect relationships? How variable are the cause–effect relationships across species (taxa)? Resolution of these questions requires empirical analyses as well as agreement among biologists on a defining criterion to facilitate communication and discovery. The so-called 'homology problem' has long animated biologists largely because formulations and answers to these questions have been controversial. Researchers with different interests, analytical tools and expertise might answer the questions differently, and the answers can change over time.

We discuss the criterion for homology and the significance of both RETICULATE EVOLUTION and DISSOCIATION among traits at different levels of

organization for homology assessment. We point out the central importance of common ancestry as the defining criterion for homology. Similarity, an alternative criterion, has multiple potential causes (including common ancestry, convergence, parallel evolution and reversal) and therefore yields multiple possible explanations for homology. Common ancestry as a defining criterion entails a single explanation and is not restricted to any particular set of traits or taxa. Focus on and discovery of the histories of characters (and thereby homology relationships) is increasingly important to broad biological understanding, as awareness grows that the 'tree of life' includes networks among the roots and branches.

The criterion for homology

Historically, the alternatives for a defining criterion for homology have been (1) similarity among traits; and (2) inheritance of traits from a most recent common ancestor, with the similarity criterion predating Darwin and the acceptance of evolution. Homology is still used by some to denote general similarity, for example when describing 'percent homology' between two DNA sequences, although this practice has been repeatedly discouraged¹. Homology as a measure of overall similarity or matching is also inherent in descriptions of recombination among 'homologous' or, less often, 'nonhomologous' chromosomes. As noted recently by Egel², this is a traditional usage among geneticists and might explain the resonance of homology as similarity for some molecular biologists. A 'biological homology' definition posits morphological structures as homologous 'if they share a set of developmental constraints, caused by locally acting self-regulatory mechanisms or organ differentiation'³. Although this bases homology assessment on similarity in developmental mechanisms of trait formation, proposed steps for assessing 'biological homology' use phylogenetic analyses and a common ancestry criterion in identifying putative homologs for subsequent comparison of developmental mechanisms⁴.

'Biological homologs' therefore become homologous traits (in the evolutionary sense) whose shared variation stems from shared developmental mechanisms, and the similarity and common ancestry criteria are overlain. The biological homology concept can help in understanding constraints on morphological traits and the maintenance of character identity over time; however, its use is limited to morphological traits whose mechanisms of development can be determined, and it lacks the generality and greater applicability of homology as defined by common ancestry.

In considering the criterion for homology, it is important to distinguish the two operations of defining and diagnosing homology. Although the defining criterion has changed to common ancestry in keeping with the explanation for sharing of traits

David P. Mindell*
Dept of Ecology and
Evolutionary Biology, and
Museum of Zoology,
University of Michigan,
Ann Arbor, MI 48109-
1079, USA.
*e-mail:
mindell@umich.edu

Axel Meyer
Dept of Biology,
University of Konstanz,
D-78457 Konstanz,
Germany.

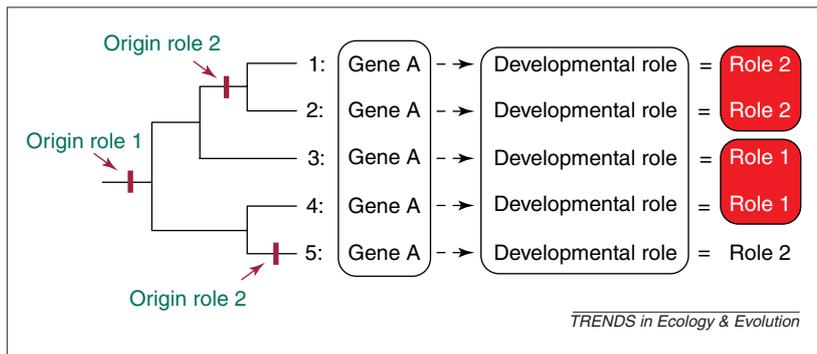


Fig. 1. Application of common ancestry as the criterion for homology in a hypothetical case of dissociation between a regulatory gene and its particular role in development. Homologous characters and states are enclosed in boxes. Regulatory gene A first plays a developmental role (role 1) in the common ancestor for taxa 1–5. Dissociation events occur such that role 2 is substituted for role 1 in the most recent common ancestor for taxa 1 and 2 and again in taxon 5. The following homology relationships can be described. ‘Gene A’ as a character is homologous in all five taxa. ‘Developmental role’ as a character is also homologous in all five taxa. However, the state of the character ‘developmental role’ changes such that character state 1 (role 1) is homologous in taxa 3 and 4 and character state 2 (role 2) is homologous in taxa 1 and 2. Character state 2 (role 2) in taxon 5 is not homologous to that in taxa 1 and 2, as character state 2 is not shared owing to inheritance of character state 2 (role 2) from the most recent common ancestor of taxa 1 and 2. Note that homologous characters can have non-homologous character states. Just as gene A could have non-homologous states at any particular nucleotide position in different taxa, the ‘developmental role’ character might have non-homologous states in different taxa.

among taxa, diagnoses of homology are not always based on phylogeny. This might be because of a lack of pre-existing phylogenetic analyses, a lack of suitable data for the taxa of interest, a lack of confidence in phylogenetic hypotheses, or potential differences in the evolutionary history of characters and organismal lineages. When homology hypotheses are not based on phylogenetic hypotheses (as in constructing data matrices before phylogenetic analyses), various aspects of similarity are used, including similarity in structure, position, developmental origins and function. Questions then arise regarding the degree to which different kinds of similarity reflect common ancestry, and the association of characters at different levels of biological organization.

At different times, virtually every class of data or trait has been advocated as being key in determining historical homology relationships, including morphological structure, the nervous system, developmental patterns, genetic data and connections among traits⁵. Researchers have tended to emphasize the importance of traits at a different, underlying level of organization as being key in determining homology for their particular traits of interest. For example, some say that homology is demonstrated for two morphological structures when they have the same embryological origin⁶ or when they are encoded by homologous genes, or that behavioral homology is demonstrated when it is based on homologous structures. However, this general approach entails a problematic assumption of association among traits at different levels of biological organization, for example, that homologous genes always encode homologous phenotypic traits.

As our understanding of the evolution of particular characters grows, more instances of dissociation

between traits over evolutionary time are seen. Various homologous regulatory genes in different species have been found to add, lose or change developmental roles and to give rise to different morphological structures. The transcription factors ‘distal-less’, ‘engrailed’ and ‘orthodenticle’ each have orthologs involved in patterning very different structural features in different metazoan taxa⁷. Alternatively, the genetic basis for important developmental processes can both change and vary among taxa. This is seen in the case of the homeodomain transcription factor *even-skipped* having an important role in pattern formation in *Drosophila* embryos, but not in the locust *Schistocerca americana*, or the wasp *Aphidius ervi*, in spite of homologous structures (segments) being present in all three groups^{8,9}.

Evolutionary dissociation is also evident where similar functions arise independently in distantly related gene families. For example, among the ABC transporters, iron uptake is a function of members of two distinct subfamilies¹⁰. At the morphological level, many consider the embryonic stage known as ‘gastrulae’ to be homologous across vertebrates, although developmental processes producing gastrulae are significantly different¹¹. Similarly, the ‘tailbud’ stage in growth of vertebrate embryos, once promoted by Haeckel as nearly identical and homologous across vertebrates, entails extensive variation in form among vertebrates as a result of allometry, heterochrony and differences in body plan and somite number¹².

Dissociation among traits, including co-option of genes for different functions, are important facilitators of change and their recognition can be key in understanding evolutionary processes. Increasingly, dissociation will be considered as a hypothesis potentially reconciling conflict between characters that appear to support alternative phylogenies. In a recent example, a dissociation has been hypothesized between developmental mechanisms and digit primordia in the avian hand, with the mechanisms acting on primordia 1–3 in theropods potentially having been dissociated (shifted) to act on primordia 2–4 in birds¹³. This could explain why the digits do not appear to be phylogenetically homologous in theropod dinosaurs and birds, in conflict with many other characters that suggest that they are sister taxa. Other researchers say that there is no evidence directly supporting the dissociation, and putative character homologies for birds and theropods remain contested¹⁴.

For a general concept such as homology, a shared defining criterion is needed for effective communication among biologists. This criterion should not be restricted to any particular type of trait or set of taxa to avoid assumptions of identical association among traits (e.g. genotype and phenotype) across taxa, and to facilitate homology assessment for more attributes of diverse organisms. Common ancestry

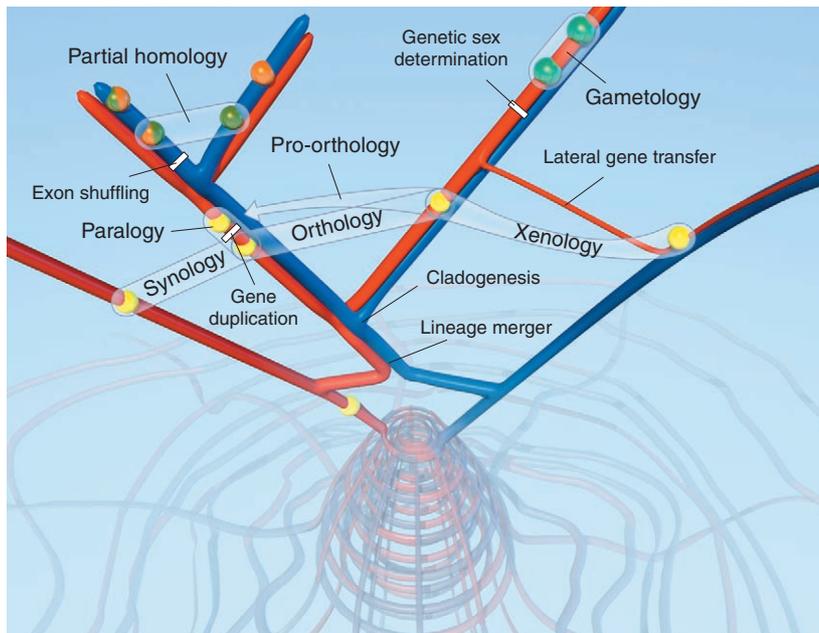


Fig. 2. Idealized phylogenetic tree illustrating seven forms of homology and their associated evolutionary processes (see Table 1). Spheres with the same color depict homologous traits (e.g. genes) among taxa, and the clear tubes connecting pairs of traits found in different taxa denote the various forms of homology, as labeled. Different forms of homology are associated with one or more different evolutionary processes: orthology with organismal lineage splitting (cladogenesis); paralogy with gene duplication; xenology with lateral gene transfer; partial homology with exon shuffling or other recombination mechanisms; gametology with origination of genetic sex determination and barriers to recombination between sex chromosomes; pro-orthology with duplication of a particular gene in one lineage but not in another; and synology with fusion of formerly independent lineages. Note that a single trait can have several different homology relationships, as seen for the trait whose relationships of orthology, synology, paralogy, and xenology are shown in yellow. Red and blue branches denote distinct lineages from which representative taxa have merged, by means of endosymbiosis, over time. The lattice structure depicts the early stages of the evolution of life, before the existence of highly integrated cells, when lateral transfer of molecular traits appears to have been particularly common²².

provides the more general and broadly applicable criterion, because it defines homology in terms of historical pattern of descent covering change for all traits across taxa, rather than similarity for any single process or structure, which might not be present in all taxa and might not be associated with other organismal traits in the same way across taxa.

An example of how the criterion of common ancestry can be applied in homology assessment of a 'dissociated' trait is presented in Fig. 1. Although useful as an aid in diagnosing homology, similarity is ill suited as a defining criterion. This is because there are multiple causes for character similarity [convergent evolution, character reversal and parallel evolution (as for 'Role 2' in taxa 1 and 5 of Fig. 1)], rendering shared similarity for traits ambiguous or empirically empty as a defining criterion for homology¹⁵, and of little practical use for evolutionary studies.

Evolutionary process and forms of homology

Terms

Homology relationships are a result of common ancestry. ORTHOLOGY, PARALOGY and XENOLOGY are well known as arising from processes of organismal lineage splitting (or cladogenesis), gene duplication and horizontal transfer of genetic material,

respectively (Table 1, Fig. 2). Recently proposed terms 'pro-, semi- and trans-orthology' are applicable to three different evolutionary relationships among genes involving both cladogenesis and gene duplication¹⁶. Whereas xenologous relationships stem from horizontal transfer of genes, synologous relationships stem from fusion of formerly independent lineages to form a new taxon¹⁷. GAMETOLOGY denotes the relationship between homologous genes on nonrecombining, opposite sex chromosomes¹⁸. Like paralogs, gametologs can potentially be used to root phylogenetic analyses where suitable organismal outgroups are not available or are too distantly related to in-group taxa.

These additional terms allow greater precision in describing various character relationships, and can be used where the relationship is known or assumed. A specialized form of homology commonly recognized for phenotypic characters is serial homology, denoting repeated units within an individual (e.g. arthropod segments, chordate vertebrae and limbs). Presumed derivations of serially homologous phenotypic traits from homologous genes and developmental processes remain difficult to demonstrate. However, studies showing regulatory gene control of the differences in serially homologous traits provide some support for these associations¹⁹.

Although the term 'functional homology' is misused where it is intended to denote mere similarity of function for traits, particular functions might be homologous as characters if shared among taxa owing to common ancestry. In Fig. 1, gene function 'Role 2' is homologous for taxa 1 and 2, but not taxa 1 and 5. Potential confusion regarding the level at which homology is implied by authors can be avoided by stating explicitly whether the homology relationship pertains to genic, structural, functional, behavioral, or some other type of trait²⁰.

Reticulate evolution

Because most studies of phylogeny focus on bifurcating trees, identification of homology relationships among traits has focused on those arising via lineage splitting, including cladogenesis for taxa and duplication events for genes. However, reticulate evolution is increasingly seen as an important phenomenon in the history of life^{21,22}. To the extent that past reticulate evolution events can be inferred, they introduce horizontal branches to bifurcating trees for taxa and for genes, and the potential to recognize the additional homology relationships as indicated by those horizontal branches (e.g. SYNOLOGY and xenology, Table 1, Fig. 2).

Early views of genotype evolution occurring primarily by point substitution have given way to recognition of a major role for natural genetic engineering events, resulting in the proliferation, combining and swapping of DNA sequences within and among genomes. These events can represent reticulate evolution for molecular sequences where

Table 1. Forms of homology

Form	Associated evolutionary process	Examples	Refs
Orthology	Cladogenesis (organismal lineage splitting)	Relationship between characters that are derived from the same character in the most recent common ancestor: cytochrome C in humans and chimps	1
Pro-orthology	Cladogenesis and gene duplication	Relationship between a singleton gene and a duplicate of its ortholog in another lineage: <i>AmphiOtx</i> in amphioxus is pro-orthologous to both <i>Otx1</i> and <i>Otx2</i> in mouse	16
Semi-orthology	Cladogenesis and gene duplication	Reverse of pro-orthology: <i>Otx1</i> and <i>Otx2</i> in mouse are semi-orthologous to <i>AmphiOtx</i> in amphioxus	16
Paralogy	Gene duplication	Relationship between characters arising from a gene duplication: 1 hemoglobin and myoglobin genes in humans	1
Xenology	Lateral gene transfer between taxa	Relationship between homologous characters whose history includes lateral transfer of genetic material, following divergence from the most recent common ancestor: aminoacyl-tRNA synthetase (TyrRSs) in select Archaea and eukaryotes	1,39
Partial homology	Exon shuffling or other recombination of genetic material yielding a composite gene with differing relationships of homology for different regions	Relationship between a composite gene and the homolog for some but not all of its parts: tissue plasminogen activator and fibronectin genes in humans	23
Gametology	Origin of genetic sex determination and barriers to sex chromosome recombination	Homologous genes on opposite sex chromosomes: <i>CHDW</i> and <i>CHDZ</i> in birds are gametologs	18
Synology	Fusion of formerly independent lineages, as in endosymbiosis	Relationship between homologous genes in different taxa or in different genomes within an individual resulting from lineage fusion: homologous genes in plant chloroplasts and cyanobacteria	17

nonhomologous genes or gene segments are combined in some manner to give rise to a heritable mosaic. Different domains or regions within mosaic genes have different origins, and the entire gene is reasonably described as having PARTIAL HOMOLOGU to various other genes^{1,23}.

One of the exciting challenges in trying to understand the homology and evolution of molecular characters is in assessing the historical significance of TRANSPOSABLE ELEMENTS (TEs). TEs include transposons that relocate autonomously via recombination and retrotransposons that relocate indirectly via RNA intermediates. Retrotransposons can be further divided into elements with the enzyme reverse transcriptase (retroviruses, LTR retrotransposons and LINES) and elements without reverse transcriptase (SINES and processed retropseudogenes). As an indication of their abundance, the human *Homo sapiens* and corn *Zea mays* genomes are estimated to be at least 45% and 50% TEs, respectively^{24,25}. TEs can inactivate genes via insertion, alter gene regulation, assemble new genes from transposed components and add new sequences within genes, potentially providing raw material for new functions. Their dynamic capabilities are increasingly seen as a major factor in genome and organismal evolution²⁶ and this, in turn, indicates their important potential influence on character homology relationships. Accurate homology assessment for molecular traits will increasingly require understanding of TE activity. Aside from their role in reticulate evolution within genomes, some TEs are conserved over evolutionary time, appear to be

homologous across taxa and provide potentially useful traits for phylogenetic analyses, as in the case of SINES among some mammalian taxa²⁷; however, convergence in these traits is now known, indicating the need for extensive sampling and careful interpretation of findings²⁸.

Viruses integrating their own DNA and potentially DNA from a previous host into the nuclear genome of their current host can effect reticulate evolution between genomes from different taxa. Consequences of such horizontal transfer, yielding xenologous trait relationships, can range from the hypothetical first transfer of DNA polymerase to eukaryotes²⁹ to the transfer of oncogenes. Horizontal transfer can also be caused by PLASMIDS, as in the transfer of antibiotic resistance genes among bacteria, and can also occur between genomes within a taxon having multiple genomes. Examples include the many instances of mitochondrial and chloroplast genes moving to the nucleus in various animals and plants. Tracing these different kinds of historical events is largely an exercise in molecular character homology assessment. Some molecular traits within an individual organism will be seen to have different histories of evolution compared with other molecular traits and with the organismal lineage itself.

Studies of virus evolution are increasing as sequence data become available, and reticulate evolution appears to be an important theme in the evolution of some lineages. Reticulation can result from viral recombination events in which transcription begins on one virus template, but is completed on another. Recombination among

phylogenetically distinct HIV-1 subtypes appears to have given rise to successful new forms on one or more occasions³⁰. Influenza virus genomes have a set of eight RNA segments that can reassort, exchanging whole segments from different lineages when different viral lineages co-infect a host cell. The influenza virus responsible for six human deaths in Hong Kong in 1997 might represent a reassortment event, with one segment being well adapted to domestic poultry and other segments being from wild birds³¹. If so, different segments from a single virus will have orthologs in different influenza lineages, an outcome of reticulate evolution that can yield a new, recombinant lineage capable of eluding immune system defenses in newly colonized hosts.

Homology assessment is pushed to a temporal extreme in analyses of the universal tree of life. The conventional view has been that phylogenetic analysis of orthologous genes for Eukarya, Archea and Bacteria would yield an accurate hierarchical tree with only a few reticulations, such as those representing endosymbiotic events giving rise to mitochondria and chloroplasts. Recent evidence indicates, however, a pervasive role for HORIZONTAL GENE TRANSFER in the early stages of the evolution of life, before the existence of stable and highly integrated cells^{21,22}. As a consequence, many organisms retain a mixture of genes from the earliest lineages of life, and analyses of different sets of orthologous genes can yield incongruent trees, indicating again that, for some genes, the phylogeny of life is more like a network than a strictly bifurcating tree. The gain, loss and substitution of functionally important genetic components among organisms provides further evidence that what is generally viewed as organismal history is only a subset of the history of all its components or parts. Homology assessment for these parts is crucial to discovering details of the universal network of life.

Uses for homology

Initial hypotheses of homology are used in constructing character matrices for phylogenetic analyses, and are generally based on similarity in features such as structure, position and function for morphological and molecular traits, or assessment of similarity in preliminary alignments for molecular sequence traits. Initial homology hypotheses can also be based on previous phylogenetic analyses of different data sets. These initial hypotheses can be revised following phylogenetic analysis, with homologs denoted as shared derived characters for various monophyletic groups when character changes are plotted on trees.

Subsequent hypotheses of homology are used, implicitly or explicitly, in nearly all comparative analyses of evolutionary patterns (phylogeny) and processes (mechanisms of change). Conventionally, natural selection has been inferred where nonhomologous similarity for traits in different taxa

is associated with shared environmental and ecological constraints. Major limitations of the comparative approach include taxon-sampling problems, uncertainty associated with phylogeny and a lack of knowledge about historical change in environmental influences. HOMOPLASY is the recurrence of character similarity in evolution, and, like homology, is best discovered via phylogenetic analyses. The uses of homology and homoplasy, as differing explanations for relatedness among traits, are therefore very closely linked^{32,33}.

A current, biotechnology industry application of homology concerns identification of orthologous genes as an indicator of possible functional equivalence in different taxa. Finding functionally equivalent orthologs provides markers in humans and model organisms for studies of disease genes origins, variability, regulation, expression and heritability. Depending on the details, similarity searches of databases might identify nonhomologous genes with convergent or parallel sequence similarity lacking functional equivalency, or the searches could identify orthologous sequence lacking functional equivalency across taxa owing to an evolutionary dissociation between gene and gene function. Clearly, homology of genes does not guarantee their equivalency in function. Eisen³⁴ has described an approach to predicting function of uncharacterized genes by assessing their phylogenetic position relative to orthologs of known function.

A recent, complementary approach in assessment of similarity in gene function is the use of microarray data, detailing tissue-specific gene expression profiles. For example, Walker *et al.*³⁵ examined the expression of 40 000 human genes in 522 cDNA libraries, and discovered several hundred genes of previously unknown function associated with various diseases. The majority of the genes discovered showed no sequence similarity to genes of known function, and thus could not have been identified as to potential functional similarity in the usual manner of looking at sequence similarity scores. Studies showing differences in expression profiles of orthologs in diseased versus healthy human tissue and among different types of disease (e.g. leukemias³⁶) indicate dissociation between traits at different levels of organization (genes and gene expression) within a single species, over time and in different tissues. Thus, well-corroborated understanding of homology for genes can be useful in learning about specific mutations underlying various diseases³⁷.

Conclusion

The defining criterion for homology has been common ancestry ever since darwinian evolution replaced fidelity to ARCHETYPE as the explanation for similarity among traits of organisms. The alternative criterion of similarity (whether for developmental control or other 'underlying' traits) can be helpful in diagnosing, but not defining homology, because it lacks connection to

Acknowledgements

We thank Walter Fitch, Jaime García-Moreno, Brian Hall, Joshua Rest, Yves Van de Peer, Gunter Wagner and David Wake for valuable discussion or comments on this article. John Megahan provided expert assistance in drawing Fig. 2. DPM was supported by the US National Science Foundation (DBI-9974525 and DEB-9726427) and a Guest Professorship from the University of Konstanz. AM was supported by the Deutsche Forschungsgemeinschaft.

Glossary

Archetype: concept of an idealized, basic model of vertebrate design.

Cladogenesis: evolutionary diversification and branching of lineages on a phylogenetic tree.

Dissofiation: change in the linkage or effects that particular traits from different levels of biological organization have on each other over evolutionary time. For example, in Fig. 1, the hypothetical change in developmental role for orthologous genes in taxa 1 and 5 results from two dissociation events; the resultant, identical 'role 2' in taxa 1 and 5 is an instance of homoplasy. Dissociation is seen in the case of the homeodomain transcription factor *even-skipped* having an important role in pattern formation in *Drosophila* embryos, but not in the locust *Schistocerca americana* or the wasp *Aphidius ervi*, in spite of homologous structures (segments) being present in all three groups.

Fidelity to archetype: a pre-darwinian explanation for similarity among traits of organisms.

Homology: relationship between traits of organisms that are shared as a result of common ancestry.

Homoplasy: relationship between identical traits of organisms that are shared as a result of reversal, convergence, or parallel evolution.

Horizontal gene transfer: transfer of genetic material from one genome to another in different individuals, including those from different species; also called lateral gene transfer.

LINES: long interspersed elements; LINES are retroposons resulting from the activity of RNA polymerase II. They are regarded as processed pseudogenes and their function is unknown.

Orthology, Paralogy, Xenology, Synology, Gametology, Partial homology: forms of homology for traits resulting from: cladogenesis; gene duplication; interspecific (horizontal) transfer of genetic material; lineage fusion, origin of genetic sex determination, and recombination (e.g. exon shuffling), respectively (see Table 1, Fig. 2).

Plasmids: DNA molecules that are inherited and replicated independently of chromosomal DNA. Some cause horizontal gene transfer, as when plasmids convey drug resistance genes from one bacterium to another.

Reticulate evolution: evolutionary events yielding horizontal (anastomosing) branches, rather than a strictly bifurcating tree, within a phylogeny; caused by exchange of genetic material among taxa as in merger of formerly distinct lineages (e.g. endosymbiosis), horizontal gene transfer, hybridization among species and recombination among viral lineages.

SINES: short interspersed elements; SINES are retroposons generated by RNA polymerase III. They are regarded as processed pseudogenes and are of unknown function. They are derived from genes that encode small cytoplasmic RNAs, including tRNAs.

Transposable elements: segments of DNA capable of moving from one location in a genome to another in the same genome.

explanation. Although some researchers prefer similarity as a criterion for this same lack of connection³⁸, many others do not, because the multiple

causes for similarity would make homology with this defining criterion uninformative for studies of evolution. New knowledge of dissociation among traits at different levels of organization (e.g. nonhomologous regulatory genes with equivalent functions) further highlights the problems inherent in using similarity for particular 'underlying' traits, as a means for diagnosing homology.

Abundant evidence for reticulate evolution via plasmids, viruses, endosymbioses, hybridization among taxa, lateral gene transfer early in the history of life and gene transfer between organelle and nuclear genomes, all indicates that the history of life is better depicted as a network than as a strictly bifurcating tree. Comprehensive phylogeny for the network must attempt to integrate the sometimes incongruent histories of traits (particularly molecular ones) and taxa. Indeed, the differences between traits and taxa become less distinct as evolution, character homology and phylogeny are contemplated for a particular genome.

Accurate assessment of homology relationships, defined by common ancestry, for molecular traits moving laterally among organismal lineages, offers some of the most profound recent insights into evolution. However, having a clear defining criterion in common ancestry and generally robust methods for its diagnosis does not mean that the task of homology determination is always tractable. As more characters are being recognized in more taxa, it seems many or most homology relationships might be unrecoverable owing to extinction of taxa, high levels of character divergence and recurrence of similarity. However, some important homology relationships are recoverable, and that should be reward enough.

References

- Fitch, W.M. (2000) Homology: a personal view on some of the problems. *Trends Genet.* 16, 227–231
- Egel, R. (2000) How 'homology' entered genetics. *Trends Genet.* 16, 437–439
- Wagner, G.P. (1989) The biological homology concept. *Annu. Rev. Ecol. Syst.* 20, 51–69
- Wagner, G.P. (1999) A research programme for testing the biological homology concept. In *Homology* (Bock, G.R. and Cardew, G., eds), pp. 125–134, John Wiley & Sons
- Lauder, G.V. (1994) Homology, form and function. In *Homology: The Hierarchical Basis of Comparative Biology* (Hall, B.K., ed.), pp. 152–196, Academic Press
- Katz, P.S. and Tazaki, K. (1992) Comparative and evolutionary aspects of the crustacean stomatogastric system. In *Dynamic Biological Networks: The Stomatogastric Nervous System* (Harris-Warrick, R.M. et al., eds), pp. 221–261, Massachusetts Institute of Technology Press
- Lowe, C.J. and Wray, G.A. (1997) Radical alterations in the roles of homeobox genes during echinoderm evolution. *Nature* 389, 718–721
- Patel, N.H. et al. (1992) Changing role of *even-skipped* during the evolution of insect pattern formation. *Nature* 357, 339–342
- Grbic, M. and Strand, M.R. (1998) Shifts in the life history of parasitic wasps correlate with pronounced alterations in early development. *Proc. Natl. Acad. Sci. U. S. A.* 95, 1097–1101
- Henikoff, S. et al. (1997) Gene families: the taxonomy of protein paralogs and chimeras. *Science* 278, 609–614
- Hall, B.K. (1994) Homology and embryonic development. *Evol. Biol.* 28, 1–37
- Richardson, M.K. et al. (1997) There is no highly conserved embryonic stage in the vertebrates: implications for current theories of evolution and development. *Anat. Embryol.* 196, 91–106
- Wagner, G.P. and Gauthier, J.A. (1999) 1,2,3 = 2,3,4: a solution to the problem of the homology of the digits in the avian hand. *Proc. Natl. Acad. Sci. U. S. A.* 96, 5111–5116
- Feduccia, A. (1999) 1,2,3 = 2,3,4: accommodating the cladogram. *Proc. Natl. Acad. Sci. U. S. A.* 96, 4740–4742
- Rieppel, O. and Grande, L. (1994) Summary and comments on systematic pattern and evolutionary process. In *Interpreting the Hierarchy of Nature* (Grande, L. and Rieppel, O., eds), pp. 248–255, Academic Press
- Holland, P.W.H. (1999) The effect of gene duplication on homology. In *Homology* (Bock G.R. and Cardew, G., eds), pp. 226–236, John Wiley & Sons
- Gogarten, J.P. (1994) Which is the most conserved group of proteins? Homology—orthology, paralogy, xenology and the fusion of independent lineages. *J. Mol. Evol.* 39, 5441–5443
- García-Moreno, J. and Mindell, D.P. (2000) Using homologous genes on opposite sex chromosomes (gametologs) in phylogenetic analysis: a case study with avian CHD. *Mol. Biol. Evol.* 17, 1826–1832
- Weatherbee, S.D. et al. (1998) *Ultrathorax* regulates genes at several levels of the wing-patterning hierarchy to shape the development of the *Drosophila* haltere. *Genes Devel.* 12, 1474–1482
- Abouheif, E. et al. (1997) Homology and developmental genes. *Trends Genet.* 13, 432–433
- Doolittle, W.F. (1999) Phylogenetic classification and the universal tree. *Science* 284, 2124–2128
- Woese, C.R. (2000) Interpreting the universal phylogenetic tree. *Proc. Natl. Acad. Sci. U. S. A.* 97, 8392–8396
- Hillis, D.M. (1994) Homology in molecular biology. In *Homology: The Hierarchical Basis of Comparative Biology* (Hall, B.K., ed.), pp. 339–368, Academic Press
- Lander, E.S. et al. (2001) Initial sequencing and analysis of the human genome. *Nature* 409, 860–921
- Bennetzen, J.L. (2000) Transposable element contributions to plant gene and genome evolution. *Plant Mol. Biol.* 42, 251–269

- 26 Kidwell, M.G. and Lisch, D.R. (2000) Transposable elements and host genome evolution. *Trends Ecol. Evol.* 15, 95–99
- 27 Lum, J.K. *et al.* (2000) Consistency of SINE insertion topology and flanking sequence tree: quantifying relationships among cetartiodactyls. *Mol. Biol. Evol.* 17, 1417–1424
- 28 Pecon Slattery, J. *et al.* (2000) Patterns of diversity among SINE elements isolated from three Y-chromosome genes in carnivores. *Mol. Biol. Evol.* 17, 825–829
- 29 Sonnhammer, E.L.L. and Wotton, J.C. (1997) Widespread eukaryotic sequences, highly similar to bacterial DNA polymerase I, looking for functions. *Curr. Biol.* 7, R463–R465
- 30 Quinones-Mateu, M.E. and Arts, E.J. (1999) Recombination in HIV-1: update and implications. *AIDS Rev.* 1, 89–100
- 31 Webster, R.G. (1999) Antigenic variation in influenza viruses. In *Origin and Evolution of Viruses* (Domingo, E. *et al.*, eds), pp. 377–420, Academic Press
- 32 Sanderson, M.J. and Hufford, L., eds (1996) *Homoplasy*, Academic Press
- 33 Meyer, A. (1999) Homology and homoplasy: the retention of genetic programmes. In *Homology* (Bock, G.R. and Cardew, G., eds), pp. 141–157, John Wiley & Sons
- 34 Eisen, J.A. (1998) Phylogenomics: improving functional predictions for uncharacterized genes by evolutionary analysis. *Genome Res.* 8, 163–167
- 35 Walker, M.G. *et al.* (1999) Prediction of gene function by genome-scale expression analysis: prostate cancer-associated genes. *Genome Res.* 9, 1198–1203
- 36 Wooster, R. (2000) Cancer classification with DNA microarrays. *Trends Genet.* 16, 327–329
- 37 Stearns, S.C., ed. (1999) *Evolution in Health and Disease*, Oxford University Press
- 38 Panchen, A.L. (1994) Richard Owen and the concept of homology. In *Homology: The Hierarchical Basis of Comparative Biology* (Hall, B.K., ed.), pp. 21–62, Academic Press
- 39 Woese, C.R. *et al.* (2000) Aminoacyl-tRNA synthetases, the genetic code, and the evolutionary process. *Microbiol. Mol. Biol. Rev.* 64, 202–236