Homology and developmental genes

The concept of homology lies at the heart of comparative biology[1]. Recent advances in developmental biology have created the need to clarify the application of this concept to comparisons of gene expression among taxa. By definition, features are homologous if they share a common evolutionary origin. Yet many investigators have recently interpreted similar patterns of regulatory gene expression as sufficient evidence to establish homology among structures. This limits attention to a single source of evidence, and ignores the evolutionary histories of the genes and of the structures in which they are expressed. Molecular biology provides powerful tools for recognizing homologies among structures. If we are to draw meaningful conclusions when making cross-taxonomic comparisons, however, we must use these tools critically and apply the concept of homology consistently.

For well over a century, a common evolutionary origin has been the central idea encapsulated by the term homology, although evolutionary biologists have debated many aspects of the concept[1–3]. With the explosion of molecular data, it became clear that homology is a concept that applies not only to morphology, but also to genes and developmental processes[1,4–6]. More recently, it has also become clear that homology at one level does not necessitate homology at another[1–7]. Therefore, we must be clear at exactly what level we are inferring homology: genes, their expression patterns, their developmental roles, or the structures to which they give rise. Errors can occur when comparisons are conflated across these levels of biological organization and when gene expression patterns are used as the primary criterion of homology. We discuss below three errors common in the current literature, and outline practical solutions to them.

The first type of error arises when orthology (gene copies derived from speciation) and paralogy (gene copies derived from duplication) are not clearly distinguished. Evolutionary inferences based on comparing expression patterns of paralogous genes are misleading, because the wrong genes are compared. While no one would make the mistake of comparing *Drosophila abdA* to locust *Ubx*, it is easier to fall into a trap when nomenclature is less distinct. For instance, *Dlx-2* from *Xenopus* and *Dlx-2* from zebrafish are not orthologous genes[8], and a direct comparison would be just as inappropriate. A more difficult problem arises when gene duplications have occurred since the species being compared have diverged (e.g. *Drosophila hedgehog* versus zebrafish *indian, desert, tiggy-winkle* and *sonic hedgehogs*; Ref. 9). Here, it is not clear which (if any) of the four genes in zebrafish most accurately corresponds to the ancestral gene before duplication. The general solution to avoiding this first type of error is to begin by reconstructing the evolutionary history of the gene family in all species under comparison in order to identify the timing of gene duplications relative to divergences, and then to restrict comparisons of gene expression to true orthologs (e.g. Ref. 10). In cases where gene duplications occurred after divergence of the species being compared (as in the *hedgehog* example above), one-to-one comparisons can only be made with caution. If we do not view the expression patterns of regulatory genes in a phylogenetic framework we run the risk of comparing the wrong genes.

FIGURE 1. The same means to different ends. The gene *distal-less* is expressed during the development of appendages of arthropods, echinoderms and chordates[17]. Although the gene is clearly homologous in these organisms, and although it might even be playing a similar functional role, the limbs themselves are not homologous[19]. This, and many other cases, illustrate the importance of clearly distinguishing between the evolutionary history of genes, their developmental roles, their expression domains and the structures to which they give rise. [Images kindly supplied by C. Janson (beetle and frog) and C. Lowe (seastar).]

The second type of error involves the notion of 'functional homology', which confuses similarity due to a common evolutionary origin (homology) with similarity due to functional convergence (analogy)[1,2]. This distinction is crucial because functions of homologous genes (orthologous or paralogous) can either diverge or converge on the functions of unrelated genes through evolutionary time[11]. Although it is frequently assumed that conservation of gene function is more frequent than convergence, this is a largely untested assumption, particularly for developmental regulatory genes. Striking similarities exist in the functional roles of regulatory genes between invertebrates and vertebrates[12,13]. The solution to determining which of these similarities are homologous is to reconstruct the evolutionary history of the genes (see above), their roles, and the structures in which they are expressed[14,15]. Because clearly homologous structures or genes can have different functions, similarity of function is not a valid criterion for the determination of homology of either genes or structures. In this context, successful gene swapping experiments[16] do not necessarily strengthen the case for homology among the structures in which these genes function.

PII: S0168-9525(97)01271-7

The third type of error has recently received much attention, and is perhaps the most deceptive[5,7,17]. The phenomenon of recruitment (co-option) can lead to situations where truly orthologous genes are expressed in non-homologous structures during development. Most regulatory genes play several distinct roles during development[12,18]; for instance, no one considers *EN1* expression in chick somites and mouse brain as evidence that these are homologous structures[18]. A potential for confusion, however, arises in cases where a homologous gene has been independently recruited to superficially similar roles. For example (Fig. 1), *distal-less* is expressed in the distal portion of appendages during their outgrowth in arthropods, echinoderms and chordates[17]. Although the domains of gene expression are strikingly similar in all three phyla, and might reflect a homologous role specifying proximodistal axes, the appendages themselves are clearly not homologous[19]. This and other cases demonstrate that orthologous regulatory genes can be expressed in structures that have independent evolutionary origins – emphasizing the importance of distinguishing between homology among genes, developmental mechanisms and structures.

Homology is a powerful concept. In order to use it consistently when making comparisons across taxa, features should be termed homologous if, and only if, they share a common evolutionary origin. Other criteria, particularly those based on functional similarity, can be misleading. Homology is a hypothesis about the evolutionary origins of a trait, and gene expression data can be an extremely valuable source of evidence supporting homology of a morphological feature, although they cannot be the sole criteria. Any hypothesis of morphological homology based on gene expression data should include: (1) a robust phylogeny of the taxa; (2) a reconstructed evolutionary history of the genes whose expression is being compared; (3) extensive taxonomic sampling, including a broad range of evolutionarily informative species; and (4) a detailed understanding of comparative anatomy and embryology. Further, we should regard proposed homologies as falsifiable, and test the possibility that overtly similar gene expression patterns might be due to convergence or recruitment, rather than common ancestry.

These are exciting times as advances in developmental biology close the gap between genotype and phenotype. Maintaining a clear and consistent definition of homology will provide a framework for incorporating future conceptual and technological advances.

References

1 Hall, B.K., ed. (1994) *Homology: The Hierarchical Basis of Comparative Biology*, Academic Press
2 Mayr, E. (1982) *The Growth of Biological Thought*, Belknap
3 deBeer, G.R. (1971) *Homology, An Unsolved Problem*, Oxford University Press
4 Dickinson, W.J. (1995) *Trends Genet.* 11, 119–121
5 Abouheif, E. (1997) *Trends Ecol. Evol.* 12, 405–408
6 Roth, V.L. (1988) in *Ontogeny and Systematics* (Humphries, C.J., ed.), pp. 1–26, Columbia University Press
7 Bolker, J.A. and Raff, R.A. (1996) *BioEssays* 18, 489–494
8 Stock, D.W. *et al.* (1996) *Proc. Natl. Acad. Sci. U. S. A.* 93, 10858–10863
9 Zarodya, R., Abouheif, E. and Meyer, A. (1996) *Trends Genet.* 12, 496–497
10 Master, V.A., Kourakis, M.J. and Martindale, M.Q. (1996) *Dev. Dyn.* 207, 404–419
11 Koonin, E.V., Mushegian, A.R. and Bork, P. (1996) *Trends Genet.* 12, 334–336
12 Duboule, D. (1994) *Guidebook to the Homeobox Genes*, Oxford University Press
13 Salzberg, A. and Bellen, H.J. (1996) *Dev. Genet.* 18, 1–10
14 Nilsson, D.E. (1996) *Curr. Biol.* 6, 39–42
15 Peterson, K.J. (1995) *Nature* 373, 111–112
16 Tomarev, S.I. *et al.* (1997) *Proc. Natl. Acad. Sci. U. S. A.* 94, 2421–2426
17 Panganiban, G. *et al.* (1997) *Proc. Natl. Acad. Sci. U. S. A.* 94, 5162–5166
18 Patel, N.H. (1994) *Science* 266, 581–590
19 Brusca, R.C. and Brusca, G.J. (1990) *Invertebrates*, Sinauer

EHAB ABOUHEIF
ehab@life.bio.sunysb.edu

Department of Ecology and Evolution, State University of New York, Stony Brook, NY 11794, USA.

MICHAEL AKAM
akam@mole.bio.cam.ac.uk

Wellcome/CRC Institute, Tennis Court Road, Cambridge, UK CB2 1QR.

WILLIAM J. DICKINSON
dickinson@bioscience.utah.edu

Department of Biology, University of Utah, Salt Lake City, UT 841121, USA.

PETER W.H. HOLLAND
p.w.h.holland@reading.ac.uk

School of Animal and Microbial Sciences, The University of Reading, Whiteknights, Reading, UK RG6 6AJ.

AXEL MEYER
axel@life.bio.sunysb.edu

Department of Ecology and Evolution, State University of New York, Stony Brook, NY 11794, USA.

NIPAM H. PATEL
npatel@midway.uchicago.edu

Department of Organismal Biology and Anatomy, University of Chicago, Chicago, IL 60637, USA.

RUDOLF A. RAFF
rraff@sunflower.bio.indiana.edu

Department of Biology, Indiana University, Bloomington, IN 47405, USA.

V. LOUISE ROTH
vlroth@acpub.duke.edu

Department of Zoology, Duke University, Durham, NC 27708, USA.

GREGORY A. WRAY
gwray@life.bio.sunysb.edu

Department of Ecology and Evolution, State University of New York, Stony Brook, NY 11794, USA.