First publ. in: Trends in Genetics 12 (1996), 12, pp. 496-497

Evolution and orthology of *hedgehog* genes

Members of the conserved hedgehog (*hh*) gene family of secreted proteins fulfill a number of important regulatory functions during development. A newly discovered member of this gene family, called *echidna hedgehog* (*ehh*)[1], has temporal and spatial expression patterns and functions in muscle development of zebrafish that differ from those of *sonic hedgehog* (*shh*). Based on comparisons of sequence similarity with other *hh* genes and because of its distinct functions[1], it is has been suggested that *ehh* is a member of a new ortholog class of the vertebrate *hh* gene family[1].

Comparisons of sequence similarity do not clearly distinguish between different levels of relatedness. We conducted evolutionary analyses of vertebrate *hh* genes and showed that *ehh* is not a member of an entirely new branch of the *hh* gene family but, rather, that it is the zebrafish ortholog of the previously known *Indian hedgehog* (*Ihh*) genes (Fig. 1). Hence, if one were interested in studying the evolution of the structure and function of the *ehh* gene in zebrafish and other model systems it might best be compared with its orthologs, the *Ihh* genes. *Ihh* genes have been sequenced and their functions in development determined in mouse[2], human[3], frog (termed *banded hedgehog*, *bhh*)[4], and most recently in chicken[5]. The phylogenetic analysis suggests that *Ihh* genes are evolutionarily more closely related to *shh* genes than either of these are to *Desert hedgehog* (*Dhh*) genes (Fig. 1). In invertebrates, only a single *hh* gene is found, therefore, the *hh* gene family seems to have undergone two major gene duplication events during the evolution of vertebrates[6,7]. Additionally, a more recent duplication of the *shh* gene resulted in the origin of the *tiggy-winkle hh* (*twhh*) gene in the zebrafish[8] and other cyprinid fishes[7]. Independently, and perhaps more recently, the *dhh* (*4 hh*) has been duplicated[4], probably due to an increase in ploidy in *Xenopus*; these duplications occurred, possibly repeatedly, in other groups of tetraploid frogs (Fig. 1).

Interestingly, each of the *Ihh* orthologs (zebrafish *ehh*, frog *bhh*, mouse *Ihh*, human *Ihh*, chicken *Ihh*) seem to have somewhat dissimilar developmental functions. For example, mouse *Ihh* is expressed in gut and cartilage[9], whereas frog *bhh* is expressed in the neural plate[4]. In chicken, *Ihh* has similar, but distinct, biological properties from *shh*, in the regulation of chondrogenesis[5]. The gene tree (Fig. 1) highlights that orthologous developmental control genes, such as *Ihh*, can take on a multitude of developmental regulatory functions despite the fact that all *Ihh* genes are evolutionarily more closely related to each other than they are to other members of the conserved *hh* gene family. This observation supports the suggestion that, in general, genes can be easily co-opted into new functions during evolution. Evolutionary co-option[10] of morphological structures or behaviors for functions other than the one for which they were selected originally is an evolutionary phenomenon that, hence, might also apply to conserved regulatory genes in development.

Homology is a statement about evolutionary relatedness due to shared evolutionary history[11,12], hence, it can only be determined by gene-tree phylogenetic analyses[13]. The functions of homologous genes might not be similar, because functions of genes can change (e.g. diverge or converge) in evolution. Importantly, despite its erroneous usage[14], similarity in function has never been part of the definition of

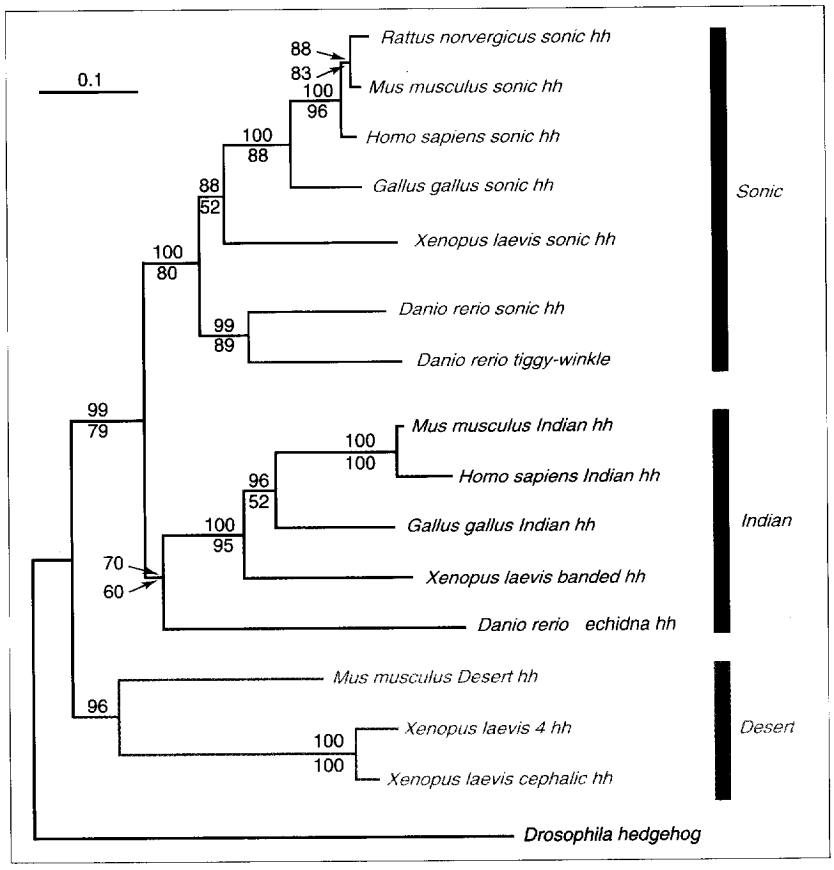

FIGURE 1. Phylogenetic relationships among members of the vertebrate *hh* gene family. A 50% majority-rule neighbor-joining[17] bootstrap tree[18] based on aligned amino acid *hh* sequences (328 characters) is shown. Ambiguously aligned positions were excluded from the phylogenetic analyses. Nomenclature of the *hh* genes follows that of the original studies. *Echidna hh* sequence (GenBank Accession No. Y08426) was kindly provided by P. Currie[1]; there is a printing error in the *ehh* sequence in Ref. 1. According to our gene tree, sequences can be classified into three orthology groups: *Dhh*, *Ihh* and *shh* as indicated by the three shaded bars. Numbers above the branches indicate percent neighbor-joining bootstrap values (1000 replications), those below are percent maximum parsimony bootstrap values (1000 replications)[19]. Branch lengths are drawn proportional to the number of inferred substitutions per site (based on Dayhoff PAM distances). All commonly used phylogenetic methods (maximum parsimony, neighbor-joining and maximum likelihood) are highly congruent and support the orthology assignment of *ehh* to the *Ihh* group and the other relationships shown. Maximum parsimony analysis[19] supports the monophyly of the *Dhh* genes with only low bootstrap support when the *Drosophila hh* is used as outgroup sequence.

homology among neither phenotypic traits nor genes[15,16]; this it particularly noteworthy because recently it has become clear that similar biochemical functions in different organisms can be performed by entirely unrelated genes[14].

Establishing that *ehh* is likely to belong to the *Indian* part of the *hh* gene family tree rather than to a completely new orthology group demonstrates that percent sequence identity and similarity or dissimilarity of function are not valid criteria for the identification of homology among genes. Once the evolutionary relationships among members of a gene family are known the nomenclature of genes would be more clear if it were based on its orthology and paralogy relationships rather than based on similarity in function. This practice would facilitate the study of evolutionary trends in changes of function for orthologous genes in different model systems.

RAFAEL ZARDOYA, EHAB ABOUHEIF AND AXEL MEYER
axelmey1@violet.berkeley.edu

Department of Ecology and Evolution and Program in Genetics, State University of New York, Stony Brook, NY 11794-5245, USA.

Acknowledgements

We thank the two reviewers for their thoughtful comments that helped to clarify this paper, and the National Science Foundation, the John Simon Guggenheim Foundation, the Max Planck Society and the Miller Institute at the University of California at Berkeley (A.M.), the Ministerio de Educación y Ciencia of Spain (R.Z.), and the Fonds pour la Formation de Chercheurs et l'Aide à la Recherche of the government of Quebec (E.A.) for financial support. This paper was written during A.M.'s tenure as a Miller Visiting Research Professor and Guggenheim Fellow at the Departments of Integrative Biology and Molecular and Cell Biology at the University of California at Berkeley. The hospitality of these departments is greatly appreciated.

References

1 Currie, P.D. and Ingham, P.W. (1996) *Nature* 382, 452–455
2 Riddle, R.D., Johnson, R.L., Laufer, E. and Tabin, C. (1993) *Cell* 75, 1401–1416
3 Marigo, V. *et al.* (1995) GenBank Accession Numbers L38517 and L38518
4 Ekker, S.C. *et al.* (1995) *Development* 121, 2337–2347
5 Vortkamp, A. *et al.* (1996) *Science* 273, 613–622
6 Kumar, S., Balczarek, K.A. and Zhi-Chun, L. (1996) *Genetics* 142, 965–972
7 Zardoya, R., Abouheif, E. and Meyer, A. (1996) *Proc. Natl. Acad. Sci. U. S. A.* 93, 13036–13041
8 Ekker, S.C. *et al.* (1995) *Curr. Biol.* 5, 944–955
9 Bitgood, M.J. and McMahon, A.P. (1995) *Dev. Biol.* 172, 126–138
10 Gould, S.J. and Vrba, E.S. (1982) *Paleobiology* 8, 4–15
11 Dickinson, W.J. (1995) *Trends Genet.* 11, 119–122
12 Meyer, A. (1996) in *New Uses for New Phylogenies* (Harvey, P.H., Leigh Brown, A.J., Maynard Smith, J. and Nee, S., eds), pp. 322–340, Oxford University Press
13 Bolker, J.A. and Raff, R.A. (1996) *BioEssays* 18, 489–494
14 Koonin, E.V., Mushegian, A.R. and Bork, P. (1996) *Trends Genet.* 12, 334–336
15 Fitch, W.M. (1970) *Syst. Zool.* 19, 99–106
16 Patterson, C. (1988) *Mol. Biol. Evol.* 5, 603–625
17 Saitou, N. and Nei, M. (1987) *Mol. Biol. Evol.* 4, 406–425
18 Felsenstein, J. (1985) *Evolution* 39, 783–791
19 Swofford, D.L. (1996) *Phylogenetic Analysis Using Parsimony* (Version 4.0d47), Smithsonian Institution