EVOLUTION & DEVELOPMENT 9:5, 405–408 (2007)

Poriferan ANTP genes: primitively simple or secondarily reduced?

Kevin J. Peterson[a,*] and Erik A. Sperling[b]

[a]Department of Biological Sciences, Dartmouth College, Hanover, NH 03755, USA
[b]Department of Geology and Geophysics, Yale University, PO Box 208109, New Haven, CT 06520, USA
*Author for correspondence (email: kevin.j.peterson@dartmouth.edu)

Using in silico techniques, Larroux et al. (2007) have recently shown that the demosponge *Amphimedon queenslandica* has only seven different NK genes, but no Hox, Parahox, or EGH genes, consistent with all previous PCR surveys that have been performed on demosponges (reviewed in Garcia-Fernàndez 2005). They further argued that this reduced set of ANTP genes is primary, and is not the result of secondary loss. Here, we argue instead that their gene tree, as does ours, supports the opposite conclusion, namely that demosponges in particular, and possibly most basal animals in general, have secondarily lost myriad transcription factors including many NK genes, and at least one Hox/Parahox/EGH gene. This observation has implications for our understanding of the genetic complexity of the last common ancestor of all living metazoans, the mono- versus paraphyly of sponges, and the early evolutionary history of animals.

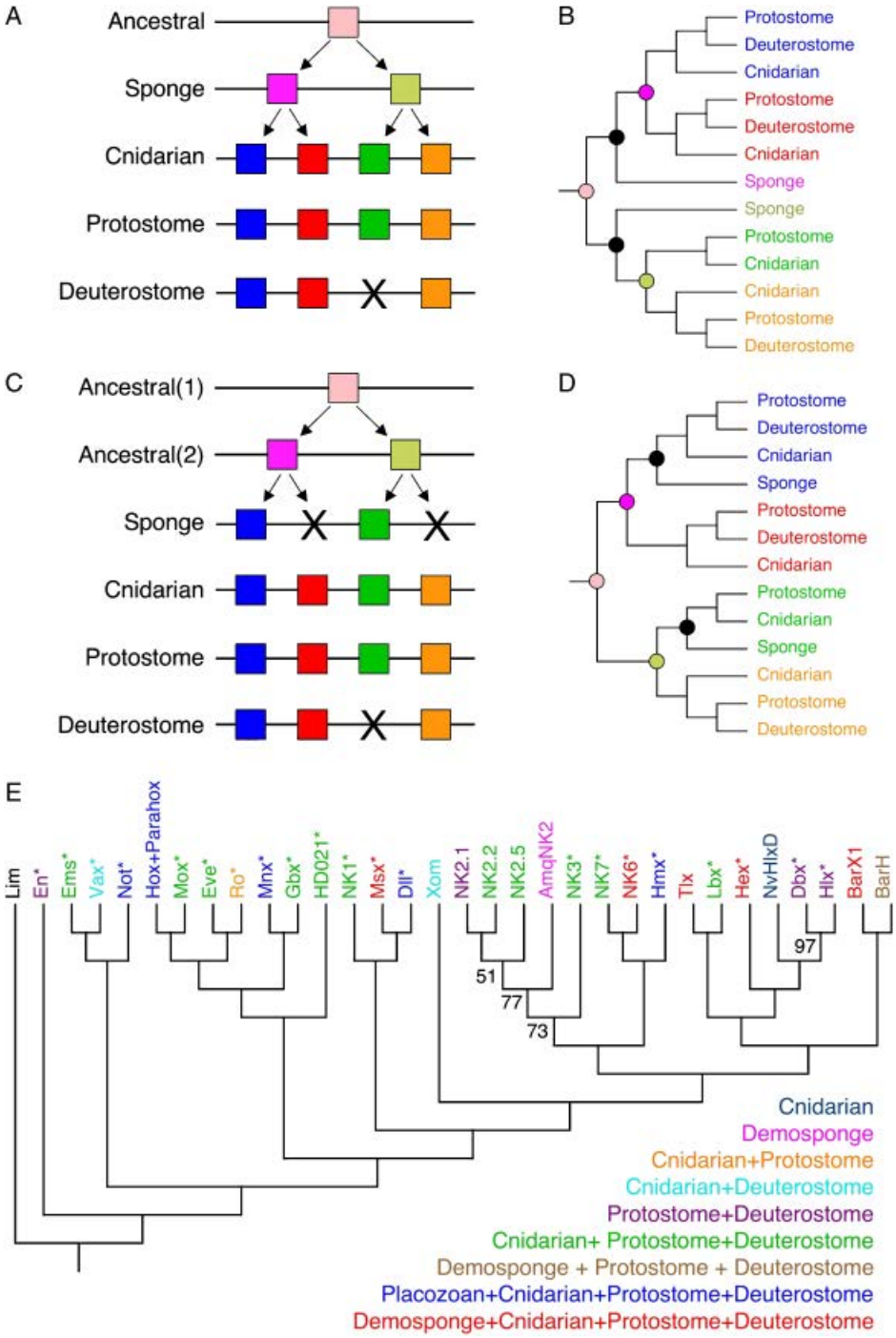
In order to decide between the two competing hypotheses, primitively simple versus secondarily reduced, it is necessary to distinguish between the gene duplication events that gave rise to the individual paralogy groups versus the speciation events that led to extant demosponges and eumetazoans (cnidarians, protostomes, and deuterostomes). As outlined by Simionato et al. (2007; see also Ryan et al. 2006), and as shown here (Fig. 1, A–D), each hypothesis makes different predictions about the order of gene duplication versus speciation events. In the former hypothesis, that of primitive simplicity, gene duplication events follow the speciation event between sponges and eumetazoans giving two genes per eumetazoan taxon (primitively) for each single demosponge gene (Fig. 1A). When drawn as a gene phylogeny, with the nodes representing either gene duplication events (magenta and lime green) or speciation events (black), the two gene duplication events clearly follow the speciation event (Fig. 1B). If, however, the gene duplication events preceded the speciation event between sponges and eumetazoans, and was followed by secondary loss of some of these genes in the sponge, as illustrated in Fig. 1C, then the resulting gene tree should show sponge orthologues to some, but not all, of the eumetazoan genes (Fig. 1D). This is made especially clear when one of the deuterostomes genes is secondarily lost (indicated on Fig. 1, A and C with an "X")—note that the cnidarian green gene is now the sister gene of its protostome orthologue, and given the species tree there is no way that deuterostomes are primitive with respect to the absence of this gene.

Simionato et al. (2007) showed a clear example of gene duplication following the poriferan/eumetazoan speciation with the basic Helix–Loop–Helix (bHLH) genes: *ARNT* and *Bmal*. In their example (see their Fig. 6), the single gene from *A. queenslandica* is equally related to the two eumetazoan genes, *ARNT* and *Bmal*. This is seen with other bHLH genes including the *TF4*/*MLX* group, and the *HIF*/*Sim*/*Trh* group. A similar pattern is found with the NK2 cluster. Eumetazoans primitively have three NK2 genes: *NK2.1*, *NK2.2*, and *NK2.5*, whereas the demosponge has a single gene equally related to all three (Fig. 1E, magenta; see also Fig. S3 in Larroux et al. 2007). Hence, rather than sponges losing NK2 genes, it appears that the gene duplication events followed the speciation event between demosponges and eumetazoans, as illustrated in Fig. 1, A and B, and the presence of a single NK2 gene in calcisponges (Manuel and Le Parco 2000) is consistent with this hypothesis.

However, *Amphimedon* also shows clear examples of secondary loss of both bHLH genes and ANTP genes. With respect to the NK genes, demosponge orthologues of *Msx*, *NK6*, and *Hex* are supported with varying degrees of precision, whereas three others, *Tlx*, *BarX1*, and *BarH*, are found but not supported $>50\%$ by bootstrap analysis or $>95\%$ by posterior probabilities (Fig. 1E; Fig. S3 in Larroux et al. 2007). Nonetheless, in both our analysis and in the analysis of Larroux et al. (2007), each individual demosponge NK gene clusters with a single eumetazoan NK gene, and in no case does the sponge gene group with more than one eumetazoan NK gene (Fig. 1E, red and brown). The point is well exemplified with NK2 and NK3. Because the presence of a single NK2 gene in the demosponge is moderately well supported, as is the sister grouping between the NK2 and NK3 families (Fig. 1E, and Larroux et al. 2007), the sponge must have secondarily lost the NK3 gene because the gene duplication event giving rise to NK2 and NK3 preceded the speciation event between demosponges and eumetazoans (see Fig. 1, C and D). Similarly, given that an *Msx* gene is found in the

demosponge, and if *Msx* is closely related to *Dll* through gene duplication, as is found here and in Larroux et al. (2007), then the demosponge must have secondarily lost the *Dll* orthologue as well. Hence, with respect to the NK genes, the most parsimonious interpretation of these data is that the demosponge *A. queenslandica* lost the orthologue of *Dll*, *NK1*, *NK3*, *NK7*, *Xom*, *Hmx*, *Lbx*, and *Hlx*/*Dbx*. It must also have lost at least one gene related to the *Ems*/*Vax*/*Not* group, and possibly the *engrailed* orthologue as well.

In addition, the demosponge must have also lost at least one gene related to the Hox/Parahox and extended Hox group (which includes the genes *Eve*, *Rough*, *Mnx*, and *Gbx*). Ignoring the presence of *Gsx* in the placozoan for the moment (see below), and the fact that Larroux et al. (2007) analyzed



an unrooted gene tree, which is strange considering that their argument depends upon assessing the polarity of evolutionary change, the only way sponges could be primitive with respect to the complete absence of any *Hox*-like genes is if this group of genes was derived from a eumetazoan NK gene. More specifically, the presence/absence pattern of NK genes in demosponges would seem to demand, according to the hypothesis of Larroux et al. (2007), that the Hox family was derived from the eumetazoan *Msx* gene, given that this is the only near relative found in the demosponge (see Fig. 1E), and thus making the *Msx* gene family paraphyletic. However, to date, no gene tree has found any NK gene, including *Msx*, paraphyletic. In fact, usually what is recovered are reciprocally monophyletic (but never strongly supported) NK and extended Hox gene families (see Fig. 1E and numerous other analyses including Bürglin 1995; Ryan et al. 2006). In principle, although Larroux et al. (2007) are most likely correct in stating that the NK cluster predates the Hox cluster, it is clear from the phylogeny that at least one *Hox*-like gene must have been lost early in the evolutionary history of demosponges, in addition to numerous NK genes. Of course, if placozoans are basal to demosponges in the metazoan tree (Dellaporta et al. 2006), then it is likely that many *Hox*-related genes (including possibly a cluster) must have been lost in the demosponge lineage given the possession of GSX in the placozoan (Monteiro et al. 2006).

An interesting observation is that the demosponge *A. queenslandica* and the placozoan *Trichoplax* have completely nonoverlapping sets of ANTP genes. Monteiro et al. (2006) found that the placozoan has *Not*, *Mnx*, *Dll*, and *Hmx* orthologues, as well as the Parahox gene *Gsx*, which is confirmed here as well (Fig. 1E, blue). The placozoan does not have *Msx*, *NK2*, *NK6*, *Tlx*, *Hex*, or the two *Bar* genes present in demosponges, or any other ANTP-type gene. It remains possible that more ANTP-type genes will be found in the placozoan, given the completed genome sequence. However, considering that all but *Hex* were already known from a variety of demosponges using standard PCR and library screening protocols (Seimiya et al. 1994; Coutinho et al. 2003; Hill et al. 2004; Richelle-Maurer et al. 2006), the chances that the placozoan has a substantially larger complement of ANTP genes seem remote. This raises a very interesting question—why would numerous genes, critical to development in bilaterians, be lost independently in these two basal clades of animals, but retained in the stem organisms leading to crown-group Eumetazoa? What is it about the development and/or ecology of the eumetazoan lineage that resulted in the maintenance of these genes in these species, as opposed to the lineages leading to crown-group demosponges and possibly crown-group placozoans where a tremendous amount of gene loss occurred? This question becomes especially interesting when one takes into account that molecular phylogenetics has indicated that the last common ancestor of all animals, as well as the earliest eumetazoans, was constructed like a modern sponge (Borchiellini et al. 2001; reviewed and discussed in Sperling et al. 2007). Why would one group of sponges, those leading to modern eumetazoans, retain all ANTP genes, whereas another sponge lineage, the one leading to modern demosponges, lose so many of them?

It could be argued that the secondary loss of ANTP genes in demosponges occurred gradually over geologic time. Nevertheless, two observations militate against this suggestion. The first is that the same complement of ANTP genes, and no more, are consistently found in all demosponges analyzed during homeobox PCR surveys; as mentioned above, all but *Hex* have been found in a variety of demosponges

Fig. 1. Using gene trees to distinguish between primitively simple versus secondarily reduced complements of ANTP genes. (A–B) Speciation preceding gene duplication. If the gene duplication event giving rise to the red/blue genes and the green/orange genes occurred after the speciation event between sponges and eumetazoans (A), then sponges should have single copy genes (magenta, lime green) equally related to the two copies in eumetazoans (red/blue and green/orange, respectively) (B). Note that in the case of a clear secondary loss, that of the green gene in deuterostomes (A, indicated with an "X"), the cnidarian green gene is the sister gene of its protostome orthologues. (C–D) Gene duplication preceding speciation. If, however, the speciation event between sponges and eumetazoans followed the gene duplication events (C), then sponges should have clear orthologues of the eumetazoan genes (blue and green) (D). Note that the secondary loss of the red and orange genes (X) phylogenetically resembles the situation with the secondary loss of the deuterostome green gene. (E) Simplified phylogeny of ANTP genes. All potential orthologues of every ANTP gene were phylogenetically analyzed from a deuterostome (primarily the sea urchin *Strongylocentrotus purpuratus*), a protostome (primarily the dipteran fly *Drosophila melanogaster*), the cnidarian *Nematostella vectensis*, the placozoan *Trichoplax adhaerans*, and the demosponge *Amphimedon queenslandica*, giving a total of 125 sequences analyzed. Phylogenetic analyses of the amino acid sequences were performed using PAUP v. 4.0b10 (Swofford 2002) for Macintosh. Three *Lim* genes were chosen as the outgroups. Distance analysis used minimum evolution as the optimality criterion (heuristic search with tree-bisection-reconnection [TBR]), and mean character difference as the distance measure. Addition sequence used 100 random replications with one tree held at each step. Bootstrap analysis used 1000 replicates. The paralogy groups are color coded according to their presence across these five taxa (e.g., red indicates that the gene is found in the demosponge, cnidarian, protostome, and deuterostome, but not the placozoan, whereas blue indicates its presence in all taxa except the demosponge). Note that no gene is present in all five taxa. An asterisk behind the gene name indicates that the monophyly of the paralogy group is supported $>50\%$, and is very similar to the results of Larroux et al. (2007). Internal bootstrap numbers are given where there is a clear (*NK2*) or potential (*Dbx*/*Hlx*) case of secondary duplication following speciation. This phylogeny, except for NK2, is congruent with gene duplication preceding speciation, followed by secondary loss in the sponge, as illustrated in Fig. 1, C and D, and not with sponges being primitively simple, as illustrated in Fig. 1, A and B, and as argued by Larroux et al. (2007).

(Seimiya et al. 1994; Coutinho et al. 2003; Hill et al. 2004; Richelle-Maurer et al. 2006), suggesting that this is the complement of the crown group. Second, Peterson and Butterfield (2005) dated the origin of the demosponge crown group to be approximately 630 Ma, which is further supported by the direct geological record in the form of biomarkers (Love et al. 2006). Given that Peterson and Butterfield (2005) dated the origin of crown-group Metazoa at 664 Ma, a large portion of the ANTP complement of genes was lost very early in demosponge history, but the remaining genes have been conserved for almost 630 Myr across a wide range of demosponge taxa.

It could also be argued that sponges are simply not paraphyletic, but monophyletic as classically recognized. This then would imply that many genes were lost early in the evolutionary acquisition of the poriferan body plan, but maintained on the stem-lineage leading to eumetazoans, and possibly suggesting that the earliest metazoans were more eumetazoan-like than sponge-like. Interestingly, the independent loss of ANTP genes presents a fairly powerful way to test the paraphyly of Porifera. If it turns out that calcisponges and homoscleromorphs have the same or a reduced complement of ANTP genes as compared with the demosponge complement, this would be consistent with their monophyly, as two independent reductions leading to the same subset would be highly unlikely, as shown in Fig. 1E with the placozoan and demosponge. Although at least one of the demosponge genes is found in calcisponges (the NK2 gene, Manuel and Le Parco 2000), the fact that a potential *Hmx* orthologue (originally identified as *Hlx* by Manuel and Le Parco, but these authors did not include *Hmx* orthologues in their analysis) was also amplified from the calcisponge *Sycon* and in our analysis groups with eumetazoan *Hmx* genes (data not shown) is consistent with Porifera being a paraphyletic grade, although this does not refute monophyly. We predict that different NK genes will be discovered in the homoscleromorphs as compared with the demosponges, consistent with the identification of two bHLH genes found in the homoscleromorph *Oscarella caremla* not found in the demosponge *A. queenslandica* (Simionato et al. 2007), and thus consistent with the paraphyly of Porifera. But even if sponges are shown to be monophyletic, this intriguing paradox is still very much unresolved, given that no analysis has ever suggested that placozoans are closely related to sponges, strongly indicating that two independent and basal lineages of animals independently lost most of their ANTP genes. To us, why there were multiple independent lines of gene loss in these basal groups, but full retention of these genes along the stem leading to crown-group Eumetazoa, remains one of the most fascinating but largely unexplored questions about early animal evolution.

REFERENCES

Borchiellini, C., Manuel, M., Alivon, E., Boury-Esnault, N., Vacelet, J., and Le Parco, Y. 2001. Sponge paraphyly and the origin of Metazoa. *J. Evol. Biol.* 14: 171–179.

Bürglin, T. R. 1995. The evolution of homeobox genes. In R. Arai, M. Kato, and Y. Doi (eds.) *Biodiversity and Evolution*. The National Science Museum Foundation, Tokyo, pp. 291–336.

Coutinho, C. C., Fonseca, R. N., Mansure, J. J. C., and Borojevic, R. 2003. Early steps in the evolution of multicellularity: deep structural and functional homologies among homeobox genes in sponges and higher metazoans. *Mech. Dev.* 120: 429–440.

Dellaporta, S. L., et al. 2006. Mitochondrial genome of *Trichoplax adhaerens* supports Placozoa as the basal lower metazoan phylum. *Proc. Natl. Acad. Sci. USA* 103: 8751–8756.

Garcia-Fernàndez, J. 2005. The genesis and evolution of homeobox gene clusters. *Nat. Rev. Genet.* 6: 881–892.

Hill, A., Tetrault, J., and Hill, M. 2004. Isolation and expression analysis of a poriferan *Antp*-class *Bar-/Bsh*-like homeobox gene. *Dev. Genes Evol.* 214: 515–523.

Larroux, C., Fahey, B., Degnan, S. M., Adamski, M., Rokhsar, D., and Degnan, B. M. 2007. The NK homeobox gene cluster predates the origin of Hox genes. *Curr. Biol.* 17: 706–710.

Love, G. D., et al. 2006. Constraining the timing of basal metazoan radiation using molecular biomarkers and U-Pb isotope dating. *Geochim. Cosmochim. Acta* 70: A371.

Manuel, M., and Le Parco, Y. 2000. Homeobox gene diversification in the calcareous songe, *Sycon raphanus*. *Mol. Phylogenet. Evol.* 17: 97–107.

Monteiro, A. S., Schierwater, B., Dellaporta, S. L., and Holland, P. W. H. 2006. A low diversity of ANTP class homeobox genes in Placozoa. *Evol. Dev.* 8: 174–182.

Peterson, K. J., and Butterfield, N. J. 2005. Origin of the Eumetazoa: testing ecological predictions of molecular clocks against the Proterozoic fossil record. *Proc. Natl. Acad. Sci. USA* 102: 9547–9552.

Richelle-Maurer, E., et al. 2006. Conservation and phylogeny of a novel family of non-*Hox* genes of the *Antp* class in Demospongiae (Porifera). *J. Mol. Evol.* 63: 222–230.

Ryan, J. F., Burton, P. M., Mazza, M. E., Kwong, G. K., Mullikin, J. C., and Finnerty, J. R. 2006. The cnidarian–bilaterian ancestor possessed at least 56 homeoboxes: evidence from the starlet sea anemone, *Nematostella vectensis*. *Genome Biol.* 7: R64.

Seimiya, M., Ishiguro, H., Miura, K., Watanabe, Y., and Kurosawa, Y. 1994. Homeobox-containing genes in the most primitive metazoa, the sponges. *Eur. J. Biochem.* 221: 219–225.

Simionato, E., et al. 2007. Origin and diversification of the basic helix–loop–helix gene family in metazoans: insights from comparative genomics. *BMC Evol. Biol.* 7: 33.

Sperling, E. A., Pisani, D., and Peterson, K. J. 2007. Poriferan paraphyly and its implications for Precambrian paleobiology. *Geol. Soc., Lond., Spec. Pub.* 286: 355–368.

Swofford, D. L. 2002. *PAUP* Phylogenetic Analysis Using Parsimony (* and Other Methods) v. 4.0b10 for Macintosh*. Sinauer Associates, Sunderland, MA.