MOLECULAR PHYLOGENY AND THE GEOGRAPHIC ORIGIN OF HALIOTIDAE TRACED BY HAEMOCYANIN SEQUENCES

KLAUS STREIT1, DANIEL L. GEIGER2 AND BERNHARD LIEB1

1Institute of Zoology, Johannes Gutenberg-University, D-55128 Mainz, Germany;
2Santa Barbara Museum of Natural History, 2539 Puesta del Sol Road, Santa Barbara, CA 93105, USA

(Received 4 March 2005; accepted 4 July 2005)

ABSTRACT

Phylogenetic relationships of 12 species of the vetigastropod family Haliotidae were analysed on the basis of partial haemocyanin coding genes. Both protein-coding and genomic sequence data, of which the latter includes one ancient intron, were analysed. We were able to distinguish two genetically different monophyletic groups with high bootstrap support: (i) the abalones from Europe, South Africa, Australia, Taiwan and Japan group together and are separated from (ii) five California and two Japanese abalones. The usefulness of haemocyanin as a phylogenetic character and the origin of the Haliotidae are discussed.

INTRODUCTION

Molluscan haemocyanins are large oxygen-binding proteins with molecular masses ranging from 3500 kDa to 8 MDa and more (for detailed review, see van Holde & Miller, 1995). The origin of haemocyanin has been calculated to lie within the Precambrian ca 700–800 Ma (Lieb, Altenhein & Markl, 2000) before all molluscan classes diverged. Thus, this protein was identified as a promising one to trace molluscan evolution. However, haemocyanin can only be found within cephalopods, chitons, gastropods and protobranch bivalves; it does not occur in other molluscs, e.g. Aplacophora, most bivalves and scaphopods. Our previous work focused on the molecular and biochemical characterization of the haemocyanins of Megathura crenulata (Sowerby, 1825) and Haliotis tuberculata Linnaeus, 1758. More recently, we started to analyse the evolutionarily informative content of these vetigastropod haemocyanins. Both species express two immunologically distinguishable isoforms, which probably evolved by duplication ca 340 ± 50 Ma (Keller et al., 1999; Lieb & Markl, 2004), before Haliotidae and Fissurellidae separated; the two isoforms can clearly be assigned as orthologous or paralogous proteins (see Results). As a consequence, we investigated the hypothesis that the haemocyanins might be a good character to trace the phylogeny of Vetigastropoda.

The clade Vetigastropoda comprises several families: the hydrothermal vent families Lepetodrilidae, Peltopsiridae, Neomphalidae, Tmnocinclidae, Sutilizonidae and Glypeosectidae, as well as the more shallow water Trochoidea (Trocchidae, Turbinidae and allied families), Fissurellidae, Pleurotomariidae, Scissurellidae s.l. and Haliotidae. There are some indications that (part of) Cocculiniformia may also belong in Vetigastropoda (Hedegaard, 1997; Colgan, Ponder & Eglger, 2000; Colgan et al., 2003; Geiger & Thacker, 2005), while others have found them more closely related to Patello gastropoda (McArthur & Harasewych, 2003). Haliotidae are of significant economic importance; over 15 species are fished and farmed (see Shepherd, Tegner & Guzman del Frío, 1992; Shepherd, McShane & Wells, 1997; Fleming & Hone, 1996; Cook et al., 1998; Leighton, 2000). Abalone have also served as model organisms in larval biology (Crofts, 1929, 1937; Lewis, Leighton & Vacquier, 1980; Morse, 1990; Jaeckle & Manahan, 1992; Naganuma et al., 1994) and biomineralization (Dauphin et al., 1989; Fritz et al., 1994).

Beside this economic role, Haliotidae are also the subject of intense investigations regarding their phylogeny and origin (for review see Geiger & Poppe, 2000). A number of attempts have been made to discriminate species and to resolve their phylogeny, using morphology, allozymes, new genetic data and also microsatellite data (Brown, 1991, 1993, 1995; Lee & Vacquier, 1992, 1995; Lee, Ota & Vacquier, 1995; Vacquier & Lee, 1993; Geiger, 1999, 2000). The radula, traditionally used as a character to trace intra-familial phylogeny (Thiele, 1931, but see Geiger & Jansen, 2004; Geiger & Thacker, 2005, for problems within Vetigastropoda), has limited information for the resolution of relationships within Haliotidae (Geiger, 1999). Two factors obscure the family-level phylogenetic signal in the radula; the plasticity of the radula to feeding requirements, and extensive effects of heterochrony in the various vetigastropod lineages (Geiger & Jansen, 2004; Geiger & Thacker, 2005; Warren, 1990). Other morphological character suits, such as the epipodium and the hypobranchial gland, have been found inadequate to resolve the species-level phylogeny within Haliotidae (Geiger, 1999).

Other molecular data, e.g. allozymes (Brown, 1991, 1993, 1995), the sperm protein lysin (Lee & Vacquier, 1992, 1995), or microsatellites (Evans, White & Elliott, 2000; Selvamani et al., 2000) also seem to resolve some phylogenetic or species-specific relationships, but may prove difficult to expand. Particularly, the homology/orthology of the functional lysin protein in Trochidae and Haliotidae is doubtful (Hellberg & Vacquier, 1999). This problem can be avoided using haemocyanins, of which orthologous or paralogous proteins can readily be identified. We introduce here haemocyanin as a promising new phylogenetic marker, show its utility at the species level, and discuss its prospects for further studies at higher systematic levels.

MATERIAL AND METHODS

DNA extraction

DNA was extracted using the E.Z.N.A. invertebrate DNA extraction kit (PeqLab, Germany) according to the manufacturer’s instructions or kindly provided by V. Vacquier (Scripps Institute, San Diego, USA).

Correspondence: B. Lieb; e-mail: lieb@uni-mainz.de
PCR amplification

PCR was performed using a standard three-step-protocol and haemocyanin-specific primers (HtH1 for: GTTTTCCATGCC-TGATGAA and HtH1g rev: CCATGGCCTGATGGCCAC) that were designed from the cDNA sequence of *H. tuberculata* haemocyanin type 1 (HtH1, GenBank Acc #: AJ 252741). 50 ng of species-specific genomic DNA were used as template in a 50 μl final volume reaction supplied with 1 μM primer, 200 μM dNTPs, 3 mM MgCl2, 3 μl of 10 times concentrated PCR buffer and 2.5 units recombinant Taq polymerase (Invitrogen, Germany). After recombinant plasmids were isolated by using 10 amplification cycles (10 s at 94°C, 2 min at 72°C) and final extension for 10 min at 72°C.

Gel electrophoresis, DNA purification and sequencing reactions

PCR products were electrophoresed in a 1.5% agarose gel (1 x TBE buffer) and gel-extracted using the Qiagen gel-extraction kit from Qiagen (Germany). The isolated fragments were cloned into TOPO-XL vectors (Invitrogen, Germany) and transformed into Top10 (Invitrogen, Germany). After recombinant plasmids were isolated by using Miniprep spin columns (Peqlab, Germany), both strands of the inserts were sequenced with standard M13rev or M13for primers, respectively, using the Taq Dye Terminator system (Perkin Elmer).

Sequence analyses

Alignments were created by ClustalX (V1.81; Thompson et al., 1997) and optimized manually using GeneDoc (http://www.psc.edu/biomed/genedoc/). Database searches and further analyses were performed using standard sequence analysis tools, e.g. BLASTN and BLASTX (Altschul et al., 1990; Bairoch & Apweiler, 1999) available at NCBI GenBank-biocomputing site, MrBayes V3.0 (http://mrbayes.ebc.uu.se/mrbayes/info.php) or the Phylip package (http://evolution.genetics.washington.edu/).

RESULTS

In our studies we used haemocyanin fragments of 12 members of the Haliotidae to evaluate the usefulness of haemocyanin sequence data for the analysis of their intra-familial phylogenetic relationships. Since it is well known that haemocyanin mostly exists as two different isoforms (van Holde & Miller, 1995; Gebauer et al., 1994) we first had to demonstrate that (i) in all species at least one haemocyanin type exists and (ii) we were able to sequence the orthologous haemocyanin coding sequences.

Nothing was known about the conservation of haemocyanin sequences between species within any molluscan family. The only molecular data known about haemocyanin came from studies of Enterocoposcus dofleini (Wulker, 1910) in comparison to *H. tuberculata* haemocyanins (Miller et al., 1998; Keller et al., 1999, Lieb et al. 2001). Based on these data, haemocyanins probably evolved prior to the radiation of all mollusc classes, ca 700 Ma within the Precambrian area, and still show a mean similarity of 70% between cephalopods and gastropods (Lieb et al., 2000, 2001). Additionally, it is known that the gene architecture of the two *H. tuberculata* isoforms (HtH1 and HtH2) is highly conserved in terms of the localization and phases of introns [Altenhein et al., 2002]. Thus, we speculated that, within Haliotidae, the primary structure of the coding regions as well as intron positions are highly conserved. Therefore, we

only a small genomic segment of one isoform, namely isoform 1 (Fig. 1). Using *H. tuberculata* haemocyanin-specific primers we were able to amplify single fragments spanning 522–657 bp. All fragments were cloned, sequenced and *H. tuberculata*-specific primers were omitted from phylogenetic analyses.

For all species, we were able to identify two exons separated by one ancestral [Altenhein, Markl & Lieb, 2002, and Discussion] intron (see Fig. 1) of which 3’ and 3’ splice junctions are highly conserved with respect to position and phase. The deduced primary structures encoded by these exons do not differ in size and are highly conserved; the similarity at the nucleotide level ranges from 80% to 99%. The obtained intron sequences separating the two exons differ significantly in size, ranging from 263 to 399 bp in length, 3’ and 3’ splice sites are highly conserved, but more 'internally' the intron similarity varies from 58% to 95% (Fig. 1). The smallest intron is represented by the *H. asinina* haemocyanin gene fragment, which is located most upstream to the 3’ splice site due to an indel region (Fig. 1). In the other species all indels are located more downstream from the 3’ splice site, but are found most often in the first half of all introns (Fig. 1).

To verify that we had amplified only orthologous type 1 haemocyanin genes of all studied members of the Haliotidae, we constructed different phylogenetic trees. We used the Bayesian inference method (ngen = 60,000; burnin = 4,000) with different model settings, e.g. ‘jones’, ‘dayhoff’, ‘wag’ etc. The resulting consensus trees based upon multiple sequence alignments of the *Haliotis* haemocyanin gene sequences, or even exclusively protein coding sequences, all showed the same topology with slightly different supporting values for different nodes (data not shown). Best values for *Haliotis* nodes resulting from gene data analyses are indicated in Figure 2. For rooting the tree, *Mega- thura crenulata* haemocyanin type 2 (KLH2)- and *H. tuberculata* haemocyanin type 2 (HtH2)-fragments were used as outgroup. Introns were excluded because they are highly derived and cannot homologize; thus the resulting support value (Fig. 2, white arrow) is derived from exon data exclusively. Nevertheless, the resulting consensus tree shows that the orthologous type 2 haemocyanins group together with high support values and are clearly separated from the paralogous type 2 haemocyanins of *M. crenulata* and other Haliotidae.

DISCUSSION

In our studies we used haemocyanin fragments of 12 members of the Haliotidae to evaluate the usefulness of haemocyanin sequence data for the analysis of their intra-familial phylogenetic relationships. Since it is well known that haemocyanin mostly exists as two different isoforms ([van Holde & Miller, 1995; Gebauer et al., 1994]) we first had to demonstrate that (i) in all species at least one haemocyanin type exists and (ii) we were able to sequence the orthologous haemocyanin coding sequences.

Nothing was known about the conservation of haemocyanin sequences between species within any molluscan family. The only molecular data known about haemocyanin came from studies of Enterocoposcus dofleini ([Wulker, 1910]) in comparison to *H. tuberculata* haemocyanins ([Miller et al., 1998; Keller et al., 1999, Lieb et al. 2001]). Based on these data, haemocyanins probably evolved prior to the radiation of all mollusc classes, ca 700 Ma within the Precambrian area, and still show a mean similarity of 70% between cephalopods and gastropods (Lieb et al., 2000, 2001). Additionally, it is known that the gene architecture of the two *H. tuberculata* isoforms (HtH1 and HtH2) is highly conserved in terms of the localization and phases of introns ([Altenhein et al., 2002]). Thus, we speculated that, within Haliotidae, the primary structure of the coding regions as well as intron positions are highly conserved. Therefore, we

List of taxa and GenBank accession numbers:

*Haliotis cracherodii* AJ884596
*Haliotis wallydamsis* AJ884596
*Haliotis gigantea* AJ749641
*Haliotis asinina* AJ749642
*Haliotis diversicolor aquatilis* AJ749643
*Haliotis discus hannai* AJ749644
*Haliotis discus superstata* AJ749645
*Haliotis fulgens* AJ749646
*Haliotis rufescens* AJ749647
*Haliotis corrugata* AJ749648
*Haliotis midae* AJ749649
*Haliotis tuberculata* HtH1AJ 252741
*Haliotis tuberculata* HtH2 AJ 297475
*Megathura crenulata* KLH1 AJ698339
*Megathura crenulata* KLH2 AJ698340.
used HtH1-specific primers spanning an intron (see Fig. 1) to amplify the corresponding gene fragments within 11 other species. For all species we could obtain specific PCR-fragments, which varied slightly in length. Additionally, untypically amplified byproducts were not obtained in any reaction, supporting our assumption that the haemocyanins of Haliotidae are highly conserved. That conclusion was further supported by more detailed analyses showing that the distances of all obtained haemocyanin sequences, at least for coding regions, ranges between 80–95% at DNA-level and 79–97% for the deduced amino acid sequences. This high degree of identity, at least at the level of amino acids, probably would not be very useful to resolve the phylogenetic relationships of all 12 taxa. However, using the whole genomic DNA-fragments including an ‘ancient’ intron, which is conserved in phase and location throughout all studied Haliotidae (Fig. 1), was very helpful and provided much higher support values for our analyses. It is notable that, according to specific indels, the phylogeny of Haliotidae traced by haemocyanin intron/exon sequences can be subdivided into two major lineages: (i) ‘short’ intron possessing Haliotidae of the European–Australasian area; and (ii) ‘long’ intron possessing Haliotidae, predominantly the North Pacific species. However, these two major clades can also be further subdivided by indels that are located in 3’ splice site direction within the European–Australasian clade and 5’ splice site directed within the North Pacific species, and probably represents the separation caused by their disjunct geographic distribution.

To confirm that our results are genuine and not due to cloning and sequencing different or paralogous isoforms, we used M. crenulata haemocyanin (KLH1/2) as well as H. tuberculata haemocyanin (HtH2) as outgroups. Since the isoform split occurred before the families diverged, the two corresponding functional units of two different species are much more closely related than the two corresponding regions of two isoforms within the same species (Keller et al., 1999; Lieb & Markl, 2004). Consequently, we could show that we only sequenced haemocyanin type 1 fragments (Fig. 2) for our analyses, which can be seen clearly within the alignment and is also mirrored by the high support values of 99–100 (Fig. 2, arrows). For the rooting of the tree we used KLH1 and also included type 2 haemocyanins to ensure that we used only orthologous sequences. Introns of KLH and HtH were excluded because they show a very high divergence rate and, therefore, cannot be aligned accurately (see Results).

This fact is not surprising, because in our earlier studies, we calculated that the divergence time of Haliotidae and Fissurellidae was ca 260 Ma (Lieb & Markl, 2004), which is in good accord with the fossil record of these families (Benton, 1993; Bandel, 1998) and therefore intron sequences had a long time to evolve in different ways. Thus, we are able to distinguish two different clades of Haliotidae, the North Pacific clade and European–Australasian clade using haemocyanin sequences. Within these clades, H. tuberculata still possess a basal position whereas the North Pacific clade seems to be a younger offshoot of a last common ancestor, probably shared by the Fissurellidae and Haliotidae ca 260 Ma.

Within these clades different species can be distinguished unambiguously the farther these species are separated geographically, e.g. H. midae (Africa) is grouped together with H. asinina (Australia) and two Asian species (H. diversicolor and H. discus supertexta), which are supported by 89%, whereas the phylogeny of the two Californian groups is less supported. This is due to a more recent divergence of these North Pacific species as indicated by the short branch lengths, rather than to data conflict (Fig. 1).

The topology obtained here is largely congruent with results from previous studies. The two major clades identified are the North Pacific clade comprising California and Japanese
species, and the European–Australasian clade. These two clades have been recovered using allozymes (Brown, 1993), lysin gene (Lee & Vacquier, 1995; Lee et al., 1995), and the combination of allozymes and lysin sequences (Geiger, 2000). Within the North Pacific clade, the basal position of *H. fulgens* and the nested Japanese species also agree with the lysin data and the combined analysis, although the allozyme tree shows somewhat different relationships. The taxon sampling within the European–Australasian clade is still limited and detailed comparisons between the various studies would be inappropriate. The results indicate that haemocyanin is suitable for species level phylogenetic investigations and should be considered as a marker in evolutionary studies. Future work on haemocyanin might include longer fragments, taxon sampling

Figure 2. The *Haliotis tuberculata* and *Megathura crenulata* haemocyanin isoforms type 2 (KLH2 and HtH2) are much more closely related than the two haemocyanin isoforms within one species, thus we can speculate that they evolved by gene duplication before the two families separated and in consequence are clearly orthologous proteins. All other sequenced fragments group together as isoform 1 orthologues, while KLH1 forms a basal fissurellid haemocyanin. However, the haemocyanin of *H. tuberculata*, *H. midae*, *H. asinina*, *H. diversicolor* and *H. discus* form a sistergroup to the remaining Haliotidae. *Haliotis tuberculata* probably represents the most basal haemocyanin. For details of methods used to reconstruct the tree, see text.

Figure 3. Our findings support that the Haliotidae originated within the Tethys Sea. Starting from there they probably spread around the world eastwards, radiating within the Asian and American Pacific regions.
should be increased, and additional outgroups among Vetigastropoda lineages, e.g. Pleurotomariidae, should be included to make the tree more robust.

ACKNOWLEDGMENTS
We thank Bernard Degnan (Queensland University, Australia) and Victor Vacquier ( Scripps Institute, USA) for kindly providing Haliotis DNA and Jürgen Markl (Johannes Gutenberg University, Germany) for working facilities. This work was supported by the DFG Li998/1–1 and Feldbausch Foundation.

REFERENCES


