J Mol Evol (2006) 63:222–230
DOI: 10.1007/s00239-005-0294-x

JOURNAL OF MOLECULAR EVOLUTION

Conservation and Phylogeny of a Novel Family of Non-*Hox* Genes of the *Antp* Class in Demospongiae (Porifera)

Evelyn Richelle-Maurer,[1] Nicole Boury-Esnault,[2] Valeria B. Itskovich,[3] Michaël Manuel,[4] Shirley A. Pomponi,[5] Gisèle Van de Vyver,[1] Carole Borchiellini[2]

[1] Laboratoire de Physiologie moléculaire de la Cellule, Université Libre de Bruxelles, CP 300, 50 av. F. Roosevelt, 1050 Brussels, Belgium
[2] Centre d'Océanologie de Marseille, Université de la Méditerranée, UMR-CNRS 6540, Station marine d'Endoume, rue de la Batterie des Lions, 13007 Marseille, France
[3] Limnological Institute of the Siberian Branch of RAS, Irkutsk, Russia
[4] Evolution et Developpement, UMR CNRS UPMC IRD ENS MNHN 7138 "Systématique, Adaptation, Evolution," Université Pierre et Marie Curie (Paris 6), 9 quai St Bernard, 75 005 Paris, France
[5] Harbor Branch Oceanographic Institution, Inc., 5600 U.S. 1 North, Fort, Pierce, FL 34946, USA

Received: 7 December 2005 / Accepted: 14 March 2006 [*Reviewing Editor:* Dr. Nicolas Galtier]

Abstract. A survey across the most basal animal phylum, the Porifera, for the presence of homeobox-containing genes led to the isolation of 24 partial or complete homeobox sequences from 21 sponge species distributed in 15 families and 6 orders of Demospongiae. All the new sequences shared a high identity/similarity with *EmH-3* (*Ephydatia muelleri*), a non-*Hox* gene from the *Antp* class. The Demox sequences, *EmH-3*, and related homeodomains formed a well-supported clade with no true affinity with any known bilaterian family, including the Tlx/Hox11 family, suggesting that the *EmH-3* family of genes, comprising 31 members, represents a novel family of non-*Hox* genes, called the Demox family, widespread among Demospongiae. The presence of the Tlx/Hox11 specific signature in the Demox family and common regulatory elements suggested that the Demox and Tlx/Hox11 families are closely related. In the phylogenetic analyses, freshwater Haplosclerida appeared as monophyletic, and Haplosclerida and Halichondrida as polyphyletic, with a clade comprising *Agelas* species and *Axinella corrugata*. As for their expression, high levels of Demox transcripts were found in adult tissues. Our data add to the number of published poriferan homeobox sequences and provide independent confirmation of the current Demospongiae phylogenies.

Key words: Sponges — Demospongiae — Homeobox genes — Expression — Phylogeny

Correspondence to: Evelyn Richelle-Maurer; *email:* emaurer@telenet.be

Introduction

Over the last decade, increasing attention has been paid to the study of basal metazoans to establish their phylogenetic relationship with bilaterians. In particular, the analysis of homeobox genes in early animals is of special interest for the understanding of the evolutionary pathway and the origin of metazoans. In this context, Porifera, at the base of the metazoan phylogenetic tree, represents a key group for developmental and evolutionary studies. In Porifera, a variety of homeobox-containing genes has been identified, all of them having bilaterian counterparts except the *Sycox* genes, so far identified only in calcareous sponges (Manuel and Le Parco 2000). Published sequences are members of the classes Irx (Iroquois) (Perovic et al. 2003), Lim (Wiens et al. 2003), Pax (Hoshiyama et al. 1998), POU (Seimiya

et al. 1997), and the *Antp* class of *non-Hox* genes, including the families Bar/Bsh (Hill et al. 2004), Msx (Seimiya et al. 1994), NK-2 (Seimiya et al. 1994; Manuel and Le Parco 2000), and the *EmH-3*/*prox2* demosponge homeobox genes, previously assigned to the Tlx/Hox11 family (Coutinho et al. 1994; Seimiya et al. 1994; Degnan et al. 1995; Richelle-Maurer et al. 1998, 2004; Wiens et al. 2003). No *Hox* genes or para-*Hox* genes have yet been reported in Porifera, suggesting that these genes appeared after the Porifera branching and that the non-*Hox* genes are the most ancestral ones (Galliot 2000; Gauchat et al. 2000).

This paper is dedicated to the study of the group of demosponge homeobox genes comprising *EmH-3*/*EflH-1* (*Ephydatia muelleri, E. fluviatilis*) and the closely related sponge sequences, *prox2* (*E. fluviatilis*), *HOXa1* (*Suberites domuncula*), and *Spox-Ta* (*Tethya aurantia*). All these genes form a well-supported orthology group, which we propose to call the Demox family. In expression analyses, the *EmH-3* gene appeared to be involved in cell-fate decisions and in the control of cell proliferation/differentiation. The expression of *EmH-3* is time- and cell-specific and is necessary for the differentiation of archaeocytes into choanocytes and hence for the completion of a functional sponge (Richelle-Maurer and Van de Vyver 1999a, b; Nikko et al. 2001). It is worth noting that the *EmH-3* promoter is functionally active in mammalian cells, implying common regulatory programs with higher metazoans (Coutinho et al. 1998, 2003). *EmH-3* expression is associated with the undifferentiated cell state in the transfected mammalian hematopoietic cells as it is in the young *Ephydatia* sponges (Coutinho et al. 2003). The presence of *EmH-3* homologues in several freshwater sponges (Richelle-Maurer et al. 2004) and of an *EmH-3*-like gene in two marine sponges (Degnan et al. 1995; Wiens et al. 2003) suggested that this type of gene could be widespread among Porifera.

To address this issue, and to ascertain the position of the Demox genes within the *Antp* class, we surveyed different orders and classes of Porifera for the presence of an *EmH-3*-like gene by means of PCR with different sets of primers. Expression analyses using RT-PCR were performed for several species.

Materials and Methods

Sponge Samples

The sponge samples were collected in Florida (Fort Pierce deep and shallow reef, Indian River lagoon, July 2002) and in the Bahamas during a Harbor Branch Oceanographic Institution (HBOI) Research Expedition aboard the R/V *Seward Johnson* (November 8–25, 2002). Sponges were collected by scuba diving or by the HBOI *Johnson-Sea-Link* manned submersible. Immediately after collection, sponges were dissociated into single cells using nylon mesh, Ca^{2+}/Mg^{2+}-free seawater, and centrifugation as described elsewhere (Richelle-Maurer et al. 2001). The external surface of the sponge was avoided to minimize potential contamination. The cell pellets were stored at –80°C. Vouchers in 70% ethanol were deposited at HBOI. A total of 28 sponge samples corresponding to 25 species, 19 families, 9 orders, and 2 classes of Porifera were investigated (Table 1). Taxonomic identification and classification of the samples were based on the Systema Porifera (Hooper and van Soest 2002).

DNA and RNA Extraction and (RT)PCR Amplification

Genomic DNA was isolated from the dissociated sponge cells using DNAzol reagent (Invitrogen) or GTC buffer followed by phenol/chloroform extraction and isopropanol precipitation. PCR amplifications were performed with different sets of degenerate and nondegenerate primers designed against *EmH-3* sequence (Table 2). Ev1 and Ev3 corresponded to the highly conserved homeodomain (HD) sequences ELEKKF and WFQNRR. ELEKKF was used instead of ELEKEF, usually utilized to amplify *Hox* genes (Manuel and Le Parco 2000), as this sequence was found in *EmH-3*, *Demox-Namo*, and *Demox-Ndig*. Furthermore, when ELEKEF/WFQNRR was used in the freshwater sponge *Ephydatia fluviatilis*, no homeobox encoding ELEKEF was recovered but two homeoboxes encoding ELRRF (*prox1*) and ELEKKF (*prox2*) were recovered (Seimiya et al. 1994).

For freshwater sponges, all the primer combinations (Table 2) except those including Em5 gave positive results. On the contrary, for marine sponges, only the degenerate primers Ev1 and Ev3 gave positive amplification with the exception of two Haplosclerida species, *Niphates amorpha* (*Demox-Namo*) and *N. digitalis* (*Demox-Ndig*), which also gave positive results with the primer sets Em9, Ev9/Em6, Ev6. PCR reactions were carried out in a 50-µl total volume containing 1–5 ng DNA, Taq PCR buffer, 10 mM dNTPs, 25 pmol of each primer, and 2.5 units of the Taq DNA polymerase (Amersham). They were run in a Techne Unit Genius Cycler under the following conditions: 2 min at 95°C, then 35–40 cycles at 94°C for 1 min, 45°–51°C for 1 min 30, 72°C for 2 min, and a final extension step at 72°C for 10 min.

Total RNA was extracted from several samples using TRIzol reagent (Invitrogen) and was amplified using the Promega Access RT-PCR System according to the manufacturer's protocol. RT-PCR conditions were those described by Richelle-Maurer et al. (2004), with an annealing temperature of 50°C. RT-PCR experiments were carried out with the Em9/Em6 set of primers for *Niphates amorpha* and *N. digitalis* (Haplosclerida) as well as with primers designed against *Demox-Namo* and *Demox-Ndig* sequences (Sp10F and Sp10R). For the other samples, *Spheciospongia vesparium* (Hadromerida), *Mycale* sp. (Poecilosclerida), and *Haliclona* sp. (Haplosclerida), the Ev1/Ev3 set of primers was used. Ten percent of the RT-PCR products were analyzed on a 1.5% agarose gel.

Reactions without template and with DNA or RNA extracted from *E. muelleri* adult sponges were performed as negative and positive controls, respectively.

(RT)PCR products of the expected size (≈120–250 bp) were either directly purified using the QIAquick PCR Purification Kit (Qiagen) or extracted from the agarose gel and purified on GenElute Agarose Spin Columns (Sigma).

Sequencing and Analysis

The purified DNA fragments were directly sequenced on both strands using the ABI PRISM BigDye Terminator v3.0 Cycle Sequencing Ready Reaction Kit and an ABI PRISM 3100 automated sequencer according to the recommended protocol (Applied

Table 1. List of species, their current classification, collection site, gene name, and accession number

Order	Family	Species	Voucher	Collection	Homeobox	Accession No.
Class Demospongiae						
Homosclerophorida	Plakinidae	*Plakortis angulospiculatus*	N 6	Bahamas	—	
Astrophorida	Geodiidae	*Geodia neptuni*	N 7	Bahamas	—	
	Corallistidae	*Corallistes* sp.	N 20	Bahamas	—	
Hadromerida	Clionaidae	*Spheciospongia vesparium*	Sp 14	Florida	Demox-Sves1	AY864744
		Spheciospongia vesparium	N 29	Bahamas	Demox -Sves2	AY864745
	Placospongiidae	*Placospongia melobesioides*	Sp 15	Florida	Demox-Pmel	AY864746
Agelasida	Agelasidae	*Agelas conifera*	N 1	Bahamas	Demox-Acon	AY864747
		Agelas clathrodes	N 24	Bahamas	Demox-Acla	AY864748
		Agelas dispar	N 11	Bahamas	Demox-Adis	AY864749
Poecilosclerida						
S.O. Myxillina	Iotrochotidae	*Iotrochota birotulata*	N 19	Bahamas	Demox-Ibir	AY864750
	Tedaniidae	*Tedania ignis*	N 18	Bahamas	Demox-Tign	AY864751
S.O. Mycalina	Mycalidae	*Mycale* sp.	Sp 12	Florida	Demox-Msp1	AY864752
		Mycale sp.	Sp 16	Florida	Demox-Msp2	AY864753
Halichondrida	Axinellidae	*Axinella corrugata*	N 2	Bahamas	Demox-Acor	AY864754
		Pseudaxinella lunaecharta	Sp 8	Florida	Demox-Plun	AY864755
		Ptilocaulis gracilis	N 8	Bahamas	Demox-Pgra	AY864756
	Halichondriidae	*Halichondria melanodocia*	Sp 3	Florida	Demox-Hmel	AY864757
Haplosclerida						
S.O. Haplosclerina	Chalinidae	*Haliclona* sp.	Sp 5	Florida	Demox-Hsp	AY864758
	Niphatidae	*Niphates digitalis*	N 26	Bahamas	Demox-Ndig	AY864759
		Niphates amorpha	Sp 10	Florida	Demox-Namo	AY864760
S.O. Petrosina	Petrosiidae	*Xestopongia muta*	N 23	Bahamas	Demox-Xmut1	AY864761
		Xestopongia muta	N 30	Bahamas	Demox-Xmut2	AY864762
S.O. Spongillina	Potamolepidae	*Potamolepis* sp.		Africa	Demox-Psp	AY864765
Freshwater sponges	Lubormirskiidae	*Lubomirskia baikalensis*	364	Baikal	Demox-Lbai	AY864766
		Baikalospongia intermedia	467	Baikal	Demox-Bint	AY864767
Dictyoceratida	Ircinidae	*Ircinia felix*	Sp 7	Florida	Demox-Ifel	AY864763
	Dysideidae	*Dysidea*	N 27	Bahamas	Demox-Dsp	AY864764
Class Calcarea						
Clathrinida	Leucettidae	*Leucetta* sp.	N 5	Bahamas	—	

Biosystems). The sequences were edited using the program Sequencer 3.0 (Applied Biosystems). Chromatograms were examined and amino acid sequences were deduced using BioEdit v.7.0.4. Searches for related sequences in other metazoans were performed in the GenBank database using the BLAST network service.

The 24 new homeobox sequences were deposited in the NCBI GenBank under accession numbers AY864744 to AY864767. They were designated Demox for Demospongiae and homeobox, followed by the first letter of the genus and the first three letters of the species (Table 1).

Phylogenetic Analyses

Demox HDs were aligned to representatives from all the non-*Hox* families HD sequences. Each family included at least one representative from Protostome, one representative from Deuterostome, and representatives from Cnidaria, Ctenophora, and Porifera when characterized. The alignment was unambiguous over the whole length of the sequences. The species codes are as follows: Cnidaria—*Af, Acropora formosa; Cv, Chlorohydra viridissima; Hv, Hydra vulgaris; Hs, Hydractinia symbiolongicarpus; Ms, Metridium senile;* and *Nv, Nematostella vectensis;* Ctenophora—*Pp*, *Pleurobrachia pileus*; Porifera—*Ef, Ephydatia fluviatilis; Efr, Eunapius fragilis; Em, Ephydatia muelleri; Hb, Halichondria* sp.*; Sd, Suberites domuncula; Sl, Spongilla lacustris; Sr, Sycon raphanus; Ta, Tethya aurantia;* and *Th, Trochospongilla horrida;* Arthropoda—*Dm, Drosophila melanogaster; Dv, Drosophila virilis;* and Vertebrata—*Dr, Danio rerio; Gg, Gallus gallus; Hs, Homo sapiens; Mm, Mus musculus; Rn, Rattus norvegicus;* and *Xl, Xenopus laevis.* Accession numbers for the sequences are as follows: Gbx family—*Gbx2-Xl*, U04867; *unp-Dm*, U35427; and *anth-hbxE-Nv*, AAG37789; Evx family—*eveC-Af*, 228960; *Evx1-Mm*, NM007966; and *eve-Dm*, M14767; En family—*en-Dv*, CAA28436; and *en2-Xl*, CAA44724; Not family—*flh-Dr*, NP571130; *Gnot1-Gg*, NP990685; and *Cnot-Hv*, CAB88387; Emx family—*ems-Dm*, X66270; *EMX1-Mm*, CCA48752; and *Cn-Ems-Hs*, CAA72534; Hlx family—*H2-Dm*, Y00843; and *Hlx1-Mm*, AAH50146; Sycox family—*Sycox1-Sr*, AF197139; *Sycox2-Sr*, AF197140; and *Sycox3-Sr*, AF197141; Nk2 family—*SrNkxB-Sr*, AF197144; *SrNkxA-Sr*, AF197143; *SrNkxC-Sr*, AF197145; *SrNkxD-Sr*, AF197146; *SrNkx2-Sr*, AF197142; *NK2-Dm*, S78691; *TTF1-Rn*, X53858; *CnNK2-Hv*, AF012538; *Nkx3.1-Mm*, U88542; *bagpipe-Dm*, L17133; and *prox1-Ef*, L10984; Barx family—*BarH1-Dm*, M73259; *Xbh1-Xl*, AAG14450; *BarX1-Mm*, Y07960; *Cnox3-Cv*, X64627; and *BarBsh-Hb*, AY328027; Bsx family—*Bsx-Mm*, NM178245; and *bsh-Dm*, L06475; Dlx family—*dll-Dm*, S47947; *Dlx2-Mm*, NM008553; and *Dlx-Hv*, AJ252183; Lbx family—*Lbe-Dm*, Y08821; and *Lbx1-Mm*, X90829; Hex family—*Hex-Dm*, AAF55844; *HEX-Hs*, Z21533; and *cnHex-Hv*, CAB88389; Tlx family— *Tlx-Pp*, CAEA5768; *Tlx3-Gg*, AF071875; and *Hox11-Dm*, CAA80535; Demox family—*EmH-3-Em*, U97664; *EmH-3-Efr*, AY300029; *EmH-3-Sl*, AY300030; *EmH-3-Th*, AY300031; *Prox2-Ef*, S43226; *HOXa1-Sd*, AJ493056; and *Spox-Ta*, X79265; and Msx family—*Msx3-Mm*, X96518; *Msh-Dm*, X85331; *Msh-Hv*, CAB88390; and *Prox3-Ef*, L23476. *CnOtx-Hv* (ADD20830), *Otx1-Mm* (CAA48754), *otd-Dm* (CAA41732), and the *POU* gene from *Xenopus laevis* (*Pou1-Xl*, S23248) were incorporated as outgroups.

Phylogenetic analyses of these amino acid alignments were performed using the neighbor-joining (NJ) distance method with

Table 2. Degenerate and nondegenerate primers designed against the *EmH-3* sequence comprising three exons (nt 1–192, nt 245–281, nt 539–753), the homeobox lying in the third exon between nt 591 and nt 661 (Richelle-Maurer et al. 1998)

Upstream		EmH-3	Downstream		EmH-3
Em 5	5′ATGGACAACTGCAGGGGTGA3′	nt 1–20	Em 6	5′ACTTGCATGTCTGAGAGTTT3′	nt 706–725
Em 9	5′GGAAATGATGCTGAGGACG3′	nt 559–577	Ev 6	5′ACYTGCATRTCWGASAGYTT3′	nt 706–725
Ev 9	5′GGHAAYGAYGCTGABGAYGA3′	nt 559–578	Em 3	5′CATTCTCCTATTTTGGAACC3′	nt 734–753
Ev 1	5′GARYTRGARAARAARTTY3′	nt 634–651	Ev 3	5′sYCKYCKRTTYTGRAACCA3′	nt 733–750
			Em 8	5′GGTGATTTCATTTCCATCTC3′	nt 781–800
Sp10F	5′GAAAAAGCCACGGACAGC3′	–	Sp10R	5′CGGCTAGTTCTCCTCTCTCG3′	–

Note. nt, nucleotide(s). All the primers are in the third exon and/or the homeobox except Em5 at the beginning of the gene. The following primer sets were used: Em5/Em3, Ev3, Em8; Em9, Ev9/ Em6, Ev6, Ev3, Em8; and Ev1/Ev3. Sp10F and Sp10R were designed against *Demox-Namo*/ *Demox-Ndig* sequences (nt 42–59 and nt 124–146, respectively).

the PAUP 4.0b10 software (Swofford 2000), with uncorrected distances, and the maximum-likelihood (ML) method with the PhyML software (Guindon and Gascuel 2003). The chosen model of amino acid substitutions was JTT (Jones et al. 1992). Among site rate variation was estimated using a discrete approximation to the gamma distribution with eight rate categories. Parameters (proportion of invariant sites and gamma shape parameters) were estimated during the search. There were no gaps in the sequences, but some sequences were incomplete; lacking positions were scored as missing data (see the alignment, available upon request from the corresponding author). For both methods, the statistical robustness of the nodes was evaluated by bootstrapping (Felsenstein 1985) (500 bootstraps for NJ, 200 bootstraps for ML).

The relationships within the Demospongiae were inferred using nucleotide sequences. Multiple sequence alignments were generated with Clustal W (Thompson et al. 1994) and checked by eye. There were no ambiguities and no gaps, but some of the sequences were incomplete and lacking positions were scored as missing data (see the alignment, available upon request from the corresponding author). Trees were built with the NJ distance method, with uncorrected distances, and the maximum parsimony (MP) method, with the PAUP 4.0b10 software (Swofford 2000). For MP analyses, characters were treated as unordered and equally weighted. MP trees were computed using heuristic searches with 20 replicates of random taxon addition sequence and TBR branch swapping. For the ML analyses of nucleotide data, a GTR (general time-reversible) + G + I model was used. This model was chosen because it is one of the most general models available for phylogenetic reconstruction from nucleotide data using ML. Among site variation was estimated using a discrete approximation to the gamma distribution with eight rate categories. Parameters were estimated with PAUP 4.0b10 from the result of a MP heuristic search, using the command Lscores. The estimated parameters were shape = 0.485787, pinvar = 0.0855665, and rmatrix = (1.68043 4.63832 1.65267 1.04942 6.49495 1.00000). Branch support was tested with bootstrapping (500 replicates for NJ and MP). Bootstraps were not performed for the ML analysis due to excessive computing time.

Results and Discussion

Characterization of the Demox Sequences

Amplified PCR products of the expected length encoding a HD were recovered with the different sets of degenerate and nondegenerate primers. A total of 24 new homeobox-containing sequences, called Demox, were isolated, all of them from the class Demospongiae. They were identified in 15 families corresponding to the major orders of the Demospongiae *sensu stricto* (Borchiellini et al. 2004) with the exception of the order Astrophorida (Table 1). They were not found in Homoscleromorpha and Calcarea. The deduced amino acid sequences shared a high degree of identity/similarity with the *EmH-3* HD from *E. muelleri* (Richelle-Maurer et al. 1998), ranging from 95% to 100% for freshwater sponges and from 92% to 97% for marine sponges. Similarly to *EmH-3*, the Demox sequences harboured the characteristic residues of the Tlx/Hox11 family-specific signature (Fig. 1), R_1, P_4, R_{10}, K_{24}, S_{28}, K_{39}, T_{47}, Q_{59}, and S_{60} (Gauchat et al. 2000). However, identity with clear members of this family was limited (62%–66%), much lower than between Tlx/Hox11 representatives (85%–96%) or with *EmH-3* (92%–100%). Almost all marine sequences possessed an E_{22}, while freshwater sequences possessed a D_{22}, which can thus be considered as a molecular diagnostic character of the Spongillina. Marine Haplosclerida seemed to be characterized by the sequence T_{21}-E_{22}-R_{23} and Dictyoceratida by an S_{39}, while a $G_{36}R_{37}$ allied the *Agelas* species (Agelasida) with *Axinella corrugata* (Halichondrida). As many as 83% of the HD residues were invariant, corresponding mainly to the three helices. There were only a few variable residues situated at positions 1, 4, 15, 21, 22, 23, 29, 36, 37, and 57 of the HD (Fig. 1). Moreover, most of the substitutions were conservative. Identity remained high even outside the HD for the freshwater sponges but decreased for the marine sponges. However, a 12-amino acid, highly conserved sequence (100% similarity) was observed at the C-terminal of the HD in the new Demox sequences (*Demox-Bint* and *Demox-Psp*) and in the previously published genes *EmH-3*, *prox2*, and *HOXa1* (Fig. 1; R1 underlined). This region shares no similarities with any other protein but might have regulatory functions. Actually, conserved sequences downstream of the HD have been observed in Engrailed (eh5) and Tlx (TH3) homeoproteins and it has been shown that the eh5 region contributed to repression (Smith and Jaynes 1996; Logan et al. 1998). It is worth noting that in all the published *EmH-3*-related genes for which the N-terminus is available,

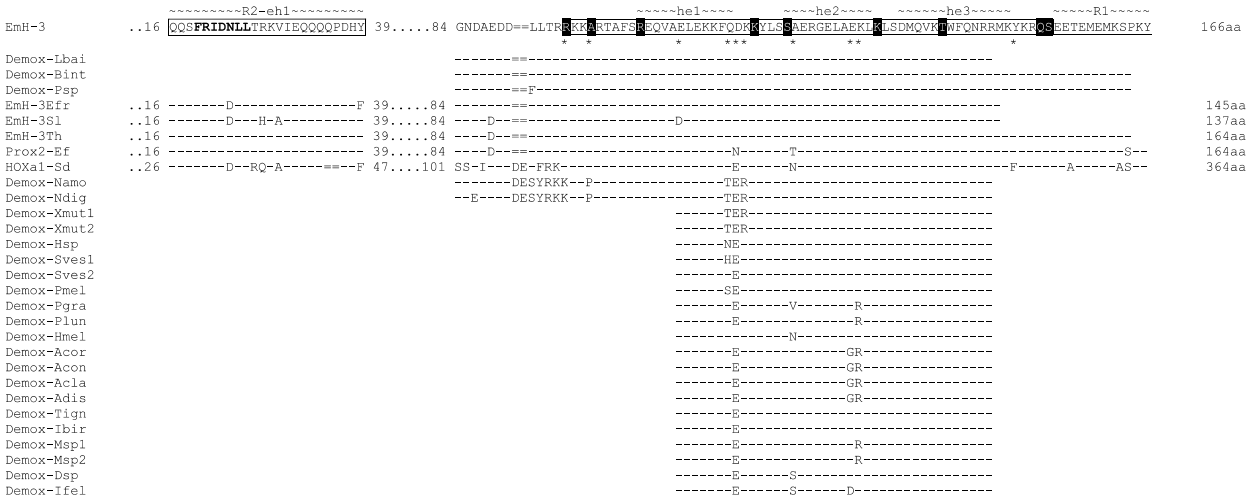

Fig. 1. Alignment of the new Demox deduced amino acid sequences to that of *EmH-3* (*Ephydatia muelleri*) allowing deletions/insertions. Demox sequences are compared to the other published related homeodomains (HDs) from the freshwater sponges *Eunapius fragilis* (*EmH-3Efr*), *Spongilla lacustris* (*EmH-3Sl*), *Trochospongilla horrida* (*EmH-3Th* [Richelle-Maurer et al. 2004]), and *E. fluviatilis* (*prox2* [Seimiya et al. 1994]) and from the marine sponge *Suberites domuncula* (*HOXa1* [Wiens et al. 2003]). The HD and the eh1 motif (R2) are boxed; the conserved sequence (R1) downstream from the HD is underlined. The 7-amino-acid core in the eh1 motive is in boldface. Amino acid identities are indicated by dashes; gaps, by equal signs (=); and variable amino acid positions in the HD, by asterisks. The residues of the Tlx/Hox11 family-specific signature are highlighted (reversed colors). The length of the published sequences is indicated numerically.

i.e., *EmH-3Efr*, *EmH-3Sl*, *EmH-3Th, prox2*, and *HOXa1*, a third conserved region (R2; Fig. 1) is found at the beginning of the gene. This region, which comprises a 7-amino-acid core, displays high similarity with the eh1 repressor domain of the En, Gsx, and Tlx/Hox11 families (Smith and Jaynes 1996; Logan et al. 1998). The presence of the eh1 motif is of particular interest as this motif is crucial for repression activity in several classes of HD proteins (Yasuo and Lemaire 2001; Bae et al. 2003; Kurata and Ueno 2003) and thus raises the possibility that *EmH-3* and related genes may act as transcriptional repressors.

Variability was higher at the nucleotide level, partly due to the degeneration of the code, so that identical amino acid sequences corresponded to different nucleotide sequences in different species except for *Lubomirskia baikalensis* and *Baikalospongia intermedia*, *Agelas clathrodes*, and *A. dispar*. Nucleotide identities varied from 82% to 100% for freshwater sponges, from 72% to 98% for marine sponges, and from 72% to 84% between freshwater and marine sponges. The Baikalian freshwater sponges *L. baikalensis* and *B. intermedia* exhibited high nucleotide identity (91%–95.6%) with *E. muelleri*, *E. fluviatilis*, and *S. lacustris*, contrary to *Potamolepis* sp. (82%–84%). High nucleotide identity (92%) was also observed between the *Agelas* species and *Axinella corrugata*.

Phylogenetic Analyses

The phylogenetic tree constructed by the ML method and based on Demox amino acid sequences and HDs from all non-*Hox* families of the *Antp* class clearly indicated that the Demox sequences together with *EmH-3* genes, *prox2*, and *HOXa1* formed a well-supported clade (ML bootstrap, 62%; NJ bootstrap, 82%) loosely related to the other non-*Hox* metazoan families (bootstrap, $< 50\%$), each family being strongly supported (Fig. 2). Interestingly, in the ML tree (Fig. 2), the Demox family was the sister group not of the Tlx family, but of a clade comprising the Tlx and Hex families (with no bootstrap support). In contrast, a clade Demox + Tlx was found in the NJ tree, but with no support (result not shown). The Sycox family (from *Sycon raphanus* [Manuel and Le Parco 2000]) is the sister group of the Hlx family (with no bootstrap support in ML and 60% of bootstraps in NJ).

Together, these data do not permit the Demox sequences to be considered as true members of Tlx/Hox11 family as previously thought (Coutinho et al. 1998, 2003; Galliot 2000; Gauchat et al. 2000; Richelle-Maurer et al. 2004). On the contrary, our data supported the view of a new family of homeobox-containing genes, designated the Demox family, with no identifiable orthologues in other metazoans and possibly specific to the Demospongiae *sensu stricto* (even if it remains to look for them in Hexactinellida and in other Homoscleromorpha and Calcispongia). Among the *Antp* class, the Demox family seemed to be evolutionarily close to the Tlx and Hex families. It can be hypothesized that some of the duplications from which these families originated have occurred in the eumetazoan branch, after the divergence of the sponge lineages, since there are no

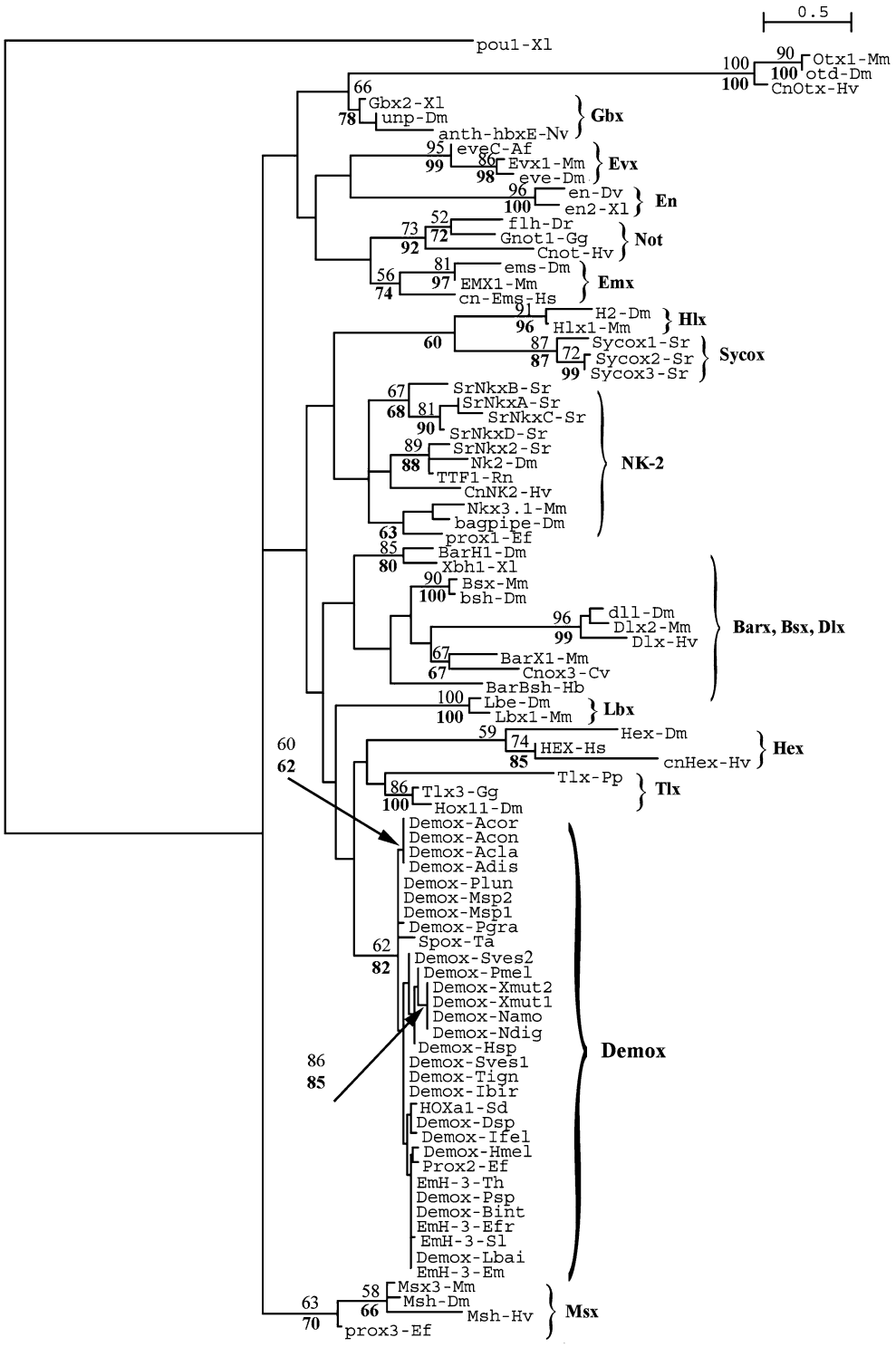

Fig. 2. Phylogenetic relationships among 87 *Antp*-class non-*Hox* genes, including all Demox sequences reported in this study, were inferred from the amino acid sequences of the full homeodomain (60 amino acids) using the maximum-likelihood (ML) method ($-\log L = 3689.18522$) (see Materials and Methods). The designation codes are listed under Materials and Methods. The tree is rooted with the diverging homeodomain *Pou1* from *Xenopus laevis*. Bootstrap values (above the branches, ML; below the branches, in boldface, neighbor joining) are indicated when > 50%. The scale bar indicates an evolutionary distance of 0.5 substitution per site.

known sponge members of the Tlx and Hex families. However, this could also simply be due to lack of data, and this question will be resolved only when complete sponge genomes become available. The presence in all the members of the Demox family of the Tlx/Hox11 family-specific signature indicates that the Demox and Tlx/Hox11 are closely related families. After the early branching of Porifera, the Demox and the Tlx/Hox11 families would have diverged, each for their part, maintaining the Tlx/Hox11 signature and some other regions necessary to their function. It should be noted that the Tlx protein from Ctenophora (*Tlx-Pp*), assigned to the Tlx/Hox11 family by Martinelli and Spring (2005), also belongs to this family in our study but with no bootstrap support. However, contrary to the Demox family, the Tlx-Pp protein does not exhibit the Tlx/Hox11 family-specific signature but, interestingly, contains an

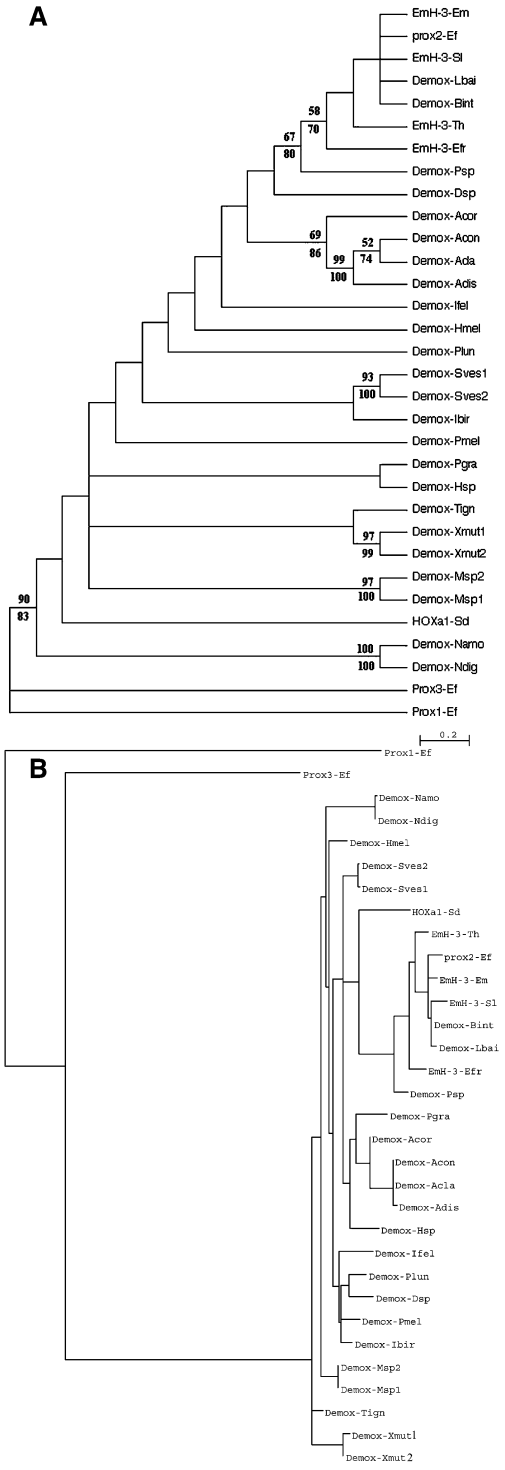

Fig. 3. Trees computed from the nucleotide sequences corresponding to amino acid positions 1 to 60 of the Demox homeodomain. The tree is rooted with *prox1* (NK-2 family) and *prox3* (Msx family) from *Ephydatia fluviatilis*. **A** Strict consensus of 21 minimal trees (390 steps) from the maximum parsimony (MP) analysis. Bootstrap values (above the branches, MP; below the branches, neighbor joining) are indicated when >50%. **B** Result from the maximum-likelihood analysis (–log$L$ = 1782.51337). The scale bar indicates an evolutionary distance of 0.2 substitution per site.

eh1 repressor motif very similar to the eh1/Tlx motif (Martinelli and Spring 2005).

As for the internal phylogeny of the class Demospongiae, most of the interorder relationships could not be resolved due to the high conservation and short length of the Demox sequences leading to a large polytomy. Nevertheless, two clades were supported by bootstrap values greater than 50% (Fig. 2): the marine Haplosclerida (S.O. Haplosclerina and Petrosina) except *Demox-Hsp* (ML bootstrap, 86%; NJ bootstrap, 85%), on one hand, and the *Agelas* species (Agelasida) and *Axinella corrugata* (*Demox-Acor*, Hadromerida; bootstrap, 60% in ML and 62% in NJ), on the other hand.

Phylogenetic analyses applied to the Demox nucleotide sequences, with *prox1* and *prox3* (the two other known *Ephydatia Antp*-class homeobox genes) as outgroups, led to no better resolution of the relationships among the Demox clade (Fig. 3A). However, the monophyly of freshwater sponges (Spongillina) was confirmed, in both the MP analysis (Fig. 3A) and the ML analysis (Fig. 3B) (MP bootstrap, 67%; NJ bootstrap, 80%). Gene pairs coming from the same species (*Demox-Sves1* and *Demox-Sves2*, *Demox-Xmut1* and *Demox-Xmut2*) or from congeneric species (*Demox-Msp1* and *Demox-Msp2*, *Demox-Acon*, *Demox-Acla* and *Demox-Adis, Demox-Ndig and Demox-Namo*) formed well-supported clades. The sequence from *Axinella corrugata* grouped with the three *Agelas* sequences in the MP tree (bootstrap, 69%), the NJ tree (bootstrap, 86%), and the ML tree.

These data provide independent corroboration of the current phylogenies of Porifera based on 18S, 28SrDNA, and other molecular and biochemical data, in particular, of the close relationship between some axinellid taxa and the order Agelasida (Itskovich et al. 1999; Schröder et al. 2003; Borchiellini et al. 2004; Erpenbeck et al. 2004, 2005a, b; Nichols 2005).

Expression in Sponges

In terms of expression, RT-PCR experiments carried out with five different species from the orders Hadromerida, Poecilosclerida, and Haplosclerida revealed the presence of transcripts of the expected size in the freshly dissociated cells from adult animals (Fig. 4). Their sequence was identical to that obtained by PCR with the same primers. The levels of expression were high, similar to those obtained with the positive control gene *EmH-3* from adult sponges, especially when specific primers were used (Figs. 4c and d). These data are consistent with the high levels of expression of *EmH-3* observed at the moment of hatching to the formation of the aquiferous system and throughout the sponge's life (Richelle-Maurer and Van de Vyver 1999a). They are also in agreement with *HOXa1* expression in adult tissues as well as in

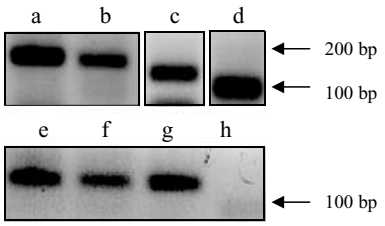

Fig. 4. Expression of the Demox genes in adult sponges (RT-PCR analysis). Gel electrophoresis shows the amplified products obtained with different sets of primers. Lane a: *EmH-3* (*Ephydatia muelleri*) used as positive control, Em9/Em6. Lane b: *Demox-Namo* (*Niphates amorpha*), Em9/Em6. Lane c: *Demox-Namo*, Em9/Sp10R. Lane d: *Demox-Namo*, Sp10F/Sp10R. Lane e: *Demox-Hsp* (*Haliclona sp.*), Ev1/Ev3. Lane f: *Demox-Msp1* (*Mycale* sp.), Ev1/Ev3. Lane g: *Demox-Sves1* (*Spheciospongia vesparium*), Ev1/Ev3. Lane h: negative control without template. Molecular sizes are indicated by arrows.

reaggregated cells (primmorphs) and during the formation of canals in the marine sponge *S. domuncula* (Wiens et al. 2003). Our data should be useful for future functional studies on the Demox genes of various demosponge species.

In conclusion, our data extend to 31 the number of published genes related to *EmH-3*. They show that the Demox sequences, *EmH-3*, *prox2*, and *HOXa1* genes are not members of the Tlx/Hox11 family but belong to a novel family of the *Antp* class, designated the Demox family. Nevertheless, the presence in all the members of the Demox family of the Tlx/Hox11 family-specific signature indicates that the Demox and Tlx/Hox11 families are closely related.

Acknowledgments. We would like to thank Amy Wright for her invitation to work in her laboratory and participate in the Bahamas Research cruise sponsored by the Harbor Branch Oceanographic Institution (HBOI), Division of Biomedical Marine Research; John K. Reed and the crew of the R/V *Seward Johnson* for their assistance in the collection of sponges; and Jane Thompson, Jose Lopez, and Laurent Gatto for their valuable advice and helpful discussions. This research was supported by an FNRS grant (Fonds National de la Recherche Scientifique, Belgium; No. 1.5.082.03).

References

Bae Y-K, Shimizu T, Yabe T, Kim C-H, Tsutomu H, Nojima H, Muraoka M, Hiramo T, Hibi M (2003) A homeobox gene, pnx, is involved in the formation of posterior neurons in zebrafish. Development 130:1853–1865

Borchiellini C, Chombard C, Manuel M, Alivon E, Vacelet J, Boury-Esnault N (2004) Molecular phylogeny of Demospongiae: implications for classification and scenarios of character evolution. Mol Phylogenet Evol 32:823–837

Coutinho CC, Vissers S, Van de Vyver G (1994) Evidence of homeobox genes in the freshwater sponge *Ephydatia fluviatilis*. In: van Soest RWM, van Kempen ThMG, Braekman JC (eds) Sponge in time and space; Biology, chemistry, paleontology. A.A. Balkema, Rotterdam, pp 385–388

Coutinho CC, Seack J, Van de Vyver G , Borojevic R, Müller WEG (1998) Origin of the metazoan bodyplan: characterization and functional testing of the promoter of the homeobox gene *EmH-3* from the freshwater sponge *Ephydatia muelleri* in mouse 3T3 cells. Biol Chem 379:1243–1251

Coutinho CC, Fonseca RN, Mansure JJ, 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

Degnan BM, Degnan SM, Giusti A, Morse DE (1995) A hox/hom homeobox gene in sponges. Gene 155:175–177

Erpenbeck D, McCormack GP, Breeuwer JAJ, van Soest RWM (2004) Order level differences in the structure of partial LSU across demosponges (Porifera): new insights into an old taxon. Mol Phylogenet Evol 32:388–395

Erpenbeck D, van Soest RWM (2005a) A survey for biochemical synapomorphies to reveal phylogenetic relationships of halichondrid demosponges (Metazoa: Porifera). Biochem Syst Ecol 31:583–616

Erpenbeck D, Breeuwer JAJ, van Soest RWM (2005b) Implications from a 28S rRNA gene fragment for the phylogenetic relationships of halichondrid sponges (Porifera:Demospongiae). J Zool Soc 43:93–99

Felsenstein J (1985) Confidence limits on phylogenies: an approach using the bootstrap. Evolution 39:783–791

Galliot B (2000) Conserved and divergent genes in apex and axis development of cnidarians. Curr Opin Genet Dev 10:629–637

Gauchat D, Mazet F, Berney C, Schummer M, Kreger S, Pawlowski J, Galliot B (2000) Evolution of the Antp-class genes differential expression of *Hydra Hox/paraHox* genes in anterior patterning. Proc Natl Acad Sci USA 97:4493–4498

Guindon S, Gascuel O (2003) A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. Syst Biol 52:696–704

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

Hooper JNA, van Soest RWM (2002) Systema Porifera: A guide to the classification of sponges. In: Hooper JNA, Soest RWMvan (eds) Systema Porifera: A guide to the classification of sponges, Vol 1. Kluwer Academic/Plenum, New York, pp 1–3

Hoshiyama D, Suga H, Iwabe N, Koyanagi M, Nikoh N, Kuma K, Matsuda F, Honjo T, Miyata T (1998) Sponge pax cDNA related to pax-2/5/8 and ancient gene duplication in the pax family. J Mol Evol 47:640–648

Itskovich VB, Belikov SI, Efremova SM, Masuda Y (1999) Phylogenetic relationships between Lubomirskiidae, Spongillidae and some marine sponges according partial sequences of 18S rDNA. Mem Queensl Mus 44:275–280

Jones DT, Taylor WR, Thornton JM (1992) The rapid generation of mutation data matrices from protein sequences. Comput Appl Biosci 8(3):275–282

Kurata T, Ueno N (2003) *Xenopus* Nbx, a novel NK-1 related gene essential for neural crest formation. Dev Biol 257:30–40

Logan C, Wingate RJT, McKay IJ, Lumsden A (1998) *Tlx-1* and *Tlx-3* homeobox gene expression in cranial sensory ganglia and hindbrain of the chick embryo: markers of patterned connectivity. J Neurosci 18:5389–5402

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

Martinelli C, Spring J (2005) T-box and homeobox genes from the ctenophore *Pleurobrachia pileus*: comparison of *brachyury, Tlbx2/3* and *Tlx* in basal metazoans and bilaterians. FEBS Lett 579(22):5024–5028

Nichols SA (2005) An evaluation of support for order-level monophyly and interrelationships within the class Demospongiae using partial data from the large subunit rDNA and cytochrome oxidase subunit I. Mol Phylogenet Evol 34:81–96

Nikko E, Van de Vyver G , Richelle-Maurer E (2001) Retinoic acid down-regulates the expression of *EmH-3* homeobox-containing gene in the freshwater sponge *Ephydatia muelleri*. Mech Age Dev 122:779–794

Perovic S, Schröder HC, Sudek S, Grebenjuk VA, Batel R, Stifanic M, Müller IM, Müller WEG (2003) Expression of one sponge Iroquois homeobox gene in primmorphs from *Suberites domuncula* during canal formation. Evol Dev 5:240–250

Richelle-Maurer E, Van de Vyver G (1999a) Expression of homeobox-containing genes in freshwater sponges. Mem Queensl Mus 44:509–514

Richelle-Maurer E, Van de Vyver G (1999b) Temporal and spatial expression of *EmH-3*, a homeobox-containing gene isolated from the freshwater sponge *Ephydatia muelleri*. Mech Age Dev 109:203–219

Richelle-Maurer E, Van de Vyver G , Vissers S, Coutinho CC (1998) Homeobox-containing genes in freshwater sponges: characterization, expression, and phylogeny. In: Müller WEG (ed) Molecular evolution: Evidence for monophyly of Metazoa, Vol 19. Springer, Berlin, pp 157–175

Richelle-Maurer E, Braekman J-C, De Kluijver MJ , Gomez R, Van de Vyver G , van Soest RWM , Devijver C (2001) Cellular location of (2*R*, 3*R*, 7*Z*)-2-aminotetradec-7-ene-1, 3-diol, a potent antimicrobial metabolite produced by the Caribbean sponge *Haliclona vansoesti*. Cell Tissue Res 306:157–165

Richelle-Maurer E, Sonet G, Van de Vyver G (2004) Identification of homologies of the *EmH-3* homeobox-containing gene in demosponges. Boll Mus Ist Biol Univ Genova 68:565–577

Schröder HC, Efremova SM, Itskovich VB, Belikov SI, Masuda Y, Krasko A, Müller IM, Müller WEG (2003) Molecular phylogeny of the freshwater sponges in Lake Baikal. J Zool Syst Evol Res 41:81–86

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

Seimiya M, Watanabe Y, Kurosawa YAD (1997) Identification of POU-class homeobox genes in a freshwater sponge and the specific expression of these genes during differentiation. Eur J Biochem 243:27–31

Smith ST, Jaynes JB (1996) A conserved region of Engrailed, shared among all en-, gsc-, Nk1-, Nk2- and msh class homeoproteins, mediates active transcriptional repression *in vivo*. Development 122:3141–3150

Swofford DL (2000) PAUP*. Phylogenetic analysis using parsimony (*and other methods), version 4. Sinauer Associates, Sunderland, MA

Thompson JD, Higgins DG, Gibson TJ (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position specific gap penalties and weight matrix choice. Nucleic Acids Res 22:4673–4680

Wiens M, Mangoni A, D'Esposito M, Fattorusso E, Korchagina N, Schröder HC, Grebenjuk VA, Krasko A, Batel R, Müller IM, Müller WEG (2003) The molecular basis for the evolution of the metazoan bodyplan: extracellular matrix-mediated morphogenesis in marine Demosponges. J Mol Evol 57:S60–S75

Yasuo H, Lemaire P (2001) Role of Goosecoid, Xnot and Wnt antagonists in the maintenance of the notochord genetic programme in *Xenopus gastrulae*. Development 128:3783–3793