Marine Biology (2002) 141: 377–386
DOI 10.1007/s00227-002-0785-x

D. Erpenbeck · J.A.J. Breeuwer · H.C. van der Velde
R.W.M. van Soest

Unravelling host and symbiont phylogenies of halichondrid sponges (Demospongiae, Porifera) using a mitochondrial marker

Received: 26 June 2001 / Accepted: 13 December 2001 / Published online: 12 April 2002

Abstract We present the first comparative phylogenetic analysis of a selected set of marine sponges and their bacterial associates. The Halichondrida form an important order in demosponge systematics and are of a particular interest due to the production of secondary metabolites. We sequenced a fragment of the cytochrome oxidase subunit 1 (*CO1*) gene of the sponges and their bacterial associates, compared the reconstructed phylogenies and found evidence for radiation in coevolution. The tree of six host-species associations showed four supported cospeciation events between the sponges and the bacteria. In addition, we present the first gene tree of sponges based on a mitochondrial marker. The tree shows major congruences with previous morphological studies and suggests the applicability of a mitochondrial marker in sponge molecular systematics.

Introduction

Sponges are sessile organisms that sequester food particles by filter feeding. Their outer cell tissue (pinacoderm) is, in contrast to other filter feeders, not completely sealed off from the surrounding medium. This may facilitate the formation of various types of associations with other organisms; some of these associations may be more permanent than others. They can be intracellular as well as extracellular (Wilkinson 1978), although fitness effects and the permanence of these relationships remain largely unknown. Several symbiotic associations in sponges have been described (cf. Vacelet 1970). In this article we shall use the term "symbiosis" in its literal definition: "living in association", regardless of particular amensalistic, commensalistic or mutualistic effects). These associations are known to harbour a wide variety of prokaryotes, including Archaebacteria (Preston et al. 1996), Cyanobacteria (Unson et al. 1994) and Proteobacteria (cf. Schumann-Kindel et al. 1997). Sponges produce a variety of molecules that may play a role in chemical defence against predators, in competition for substrate, and as protection against fouling (see Sarma et al. 1993 for an overview), but in some cases symbiotic prokaryotes may be the actual producers of these compounds (cf. Stierle et al. 1988). This suggests that symbionts are important in sponge ecology and allows for the evolution of a more permanent association with cospeciation (Wilkinson 1983, 1984). Ultrastructural studies have demonstrated vertical transmission of bacterial symbionts with sponge larvae (Gallissian and Vacelet 1976; Lévi and Lévi 1976). A common method to detect evidence for cospeciation is the comparison of host and symbiont phylogenies (cf. Casiraghi et al. 2001). If symbionts and host coevolve, their phylogenies are expected to be congruent (Page 1994). If, on the other hand, symbionts have been acquired independently many times or if associations are not specific, congruent phylogenies are not expected. In several marine systems, such congruent phylogenetic patterns have been observed (Distel et al. 1994), but records for sponge–symbiont coevolution are lacking.

Sponges are among the most primitive Metazoa (see Ax 1995). However, their position in the tree of life, as well as their internal phylogenetic relationships, are still a matter of discussion (Borchiellini et al. 2001). This is largely due to the fact that the pool of suitable morphological characters is not sufficient to allow a thorough phylogenetic analysis. Cytological features have recently been developed (cf. Boury-Esnault et al. 1994), but they can only be employed for cytologically diverse groups. Chemical compounds have been suggested as

Communicated by O. Kinne, Oldendorf/Luhe

D. Erpenbeck (✉) · J.A.J. Breeuwer · H.C. van der Velde
R.W.M. van Soest
Institute for Biodiversity and Ecosystem Dynamics,
University of Amsterdam, P.O. Box 94766,
1090 GT Amsterdam, The Netherlands

E-mail: erpenbeck@science.uva.nl
Tel.: +31-20-5256532
Fax: +31-20-5255402

well (Soest and Braekman 1999). However, since the original producer, sponge or symbiont, is in many cases difficult do detect, chemical-compound data should be used with care.

Sponge gene trees, entirely based on nuclear ribosomal genes, such as 18S- and 28SrDNA, have recently been constructed (see Borchiellini et al. 2000 for an overview). Incongruity with some morphological hypotheses has already been observed (cf. Alvarez et al. 2000), as the result of either the ambiguous morphological-character analysis or the particular properties of the sequenced gene (non-protein coding, multicopy, relatively slow evolution) or both. Clearly, we need additional gene trees to resolve poriferan phylogeny (e.g. mitochondrial genes).

In this study we sequenced a fragment of the mitochondrial cytochrome oxidase subunit 1 (CO1) of halichondrid demosponges together with the homologue of a bacterial associate, and we present the first comparative phylogenetic analysis of sponges and their symbionts. The halichondrids form a crucial and pivotal order of the demosponges, the major sponge class. Recent studies (cf. Soest et al. 1990) assign taxa of previously different families and orders to this group. They are, moreover, of particular interest because the composition of secondary metabolites is influenced by the presence of prokaryote symbionts (cf. Althoff et al. 1998). Our approach will help to understand the evolution of the symbiosis between sponges and prokaryotes and the role of symbionts in speciation and adaptability of sponges.

An additional concern is the common practice of using conserved primers to amplify gene regions of interest. Comparison with molecular databases is usually enough to establish whether the obtained gene sequences are of the correct phylogenetic group and not a contaminant. However, in symbiont-rich groups, such as the sponges, which have been little studied from the molecular standpoint, the origin of the sequences needs to be checked.

Materials and methods

DNA samples and extraction

We studied 23 morphologically representative individuals of the genera *Agelas*, *Amorphinopsis*, *Axinella*, *Axinyssa*, *Ciocalypta*, *Didiscus*, *Halichondria*, *Haliclona*, *Hymeniacidon*, *Liosina*, *Reniochalina*, *Scopalina*, *Stylissa* and *Suberites* from material freshly collected during the SYMBIOSPONGE project (EU-MAS3CT 97-0144) and from collection material at the Zoological Museum, Amsterdam. The investigated specimens are listed in Table 1.

Up to three specimens per species were examined. Freshly collected samples were transferred immediately into 100% ethanol which was exchanged several times to avoid dilution with intercellular sea water. For each specimen, we kept tissue for DNA and

Table 1. Listing of the specimens, species, authors, sample locations, PCRs, clones and the yielded fragments per DNA extraction (distinguished in their sequence signatures). "Signature 1" sequences were subsequently identified as sponge DNA sequences; "signature 2", as bacterial associates; "signature 3" are assumed to be of crustacean origin. Sequence numbers in brackets: digits indicate the specimen number, "a" or "b" the different genotypes of sequenced clones. For GenBank-accession numbers, see Results and Appendix

Specimen number	Taxon	Author	Family	Location	PCRs	Clones	Sequence signatures with sequence number (bold, in brackets)
1	*Agelas oroides*	Schmidt 1864	Agelasidae	Spain/Catalonia	3	5	Signature 1 (**1a**) + signature 3 (**1b**)
2	*Amorphinopsis excavans*	Carter 1887	Halichondriidae	Oman	2	2	Signature 1 (**2**)
3	*Axinella damicornis*	Schmidt 1862	Axinellidae	France/Roscoff	3	9	Signature 1 (**3a**) + signature 2 (**3b**)
4	*Axinella verrucosa*	Schmidt 1862	Axinellidae	Spain/Catalonia	4	4	Signature 2 (**4**)
5	*Axinyssa ambrosia*	Von Lendenfeld 1897	Halichondriidae	Dutch Antilles	3	5	Signature 2 (**5**)
6	*Ciocalypta penicillus*	Bowerbank 1864	Halichondriidae	France/Roscoff	2	2	Signature 1 (**6**)
7	*Didiscus oxeata*	Hechtel 1983	Desmoxyidae	Dutch Antilles	2	2	Signature 1 (**7**)
8	*Halichondria bowerbanki*	Burton 1930	Halichondriidae	The Netherlands	1	2	Signature 1 (**8**)
9	*Halichondria bowerbanki*	Burton 1930	Halichondriidae	The Netherlands	1	1	Signature 1 (**9**)
10	*Halichondria bowerbanki*	Burton 1930	Halichondriidae	The Netherlands	3	3	Signature 2 (**10**)
11	*Halichondria panicea*	Pallas 1766	Halichondriidae	The Netherlands	1	3	Signature 2 (**11**)
12	*Halichondria panicea*	Pallas 1766	Halichondriidae	France/Roscoff	2	4	Signature 1 (**12**)
13	*Haliclona xena*	De Weerdt 1986	Chalinidae	The Netherlands	1	2	Signature 1 (**13**)
14	*Hymeniacidon perlevis*	Montagu 1818	Halichondriidae	France/Roscoff	3	3	Signature 2 (**14**)
15	*Hymeniacidon perlevis*	Montagu 1818	Halichondriidae	France/Roscoff	1	2	Signature 1 (**15**)
16	*Liosina paradoxa*	Thiele 1899	Dictyonellidae	Sulawesi/Indonesia	3	4	Signature 1 (**16b**) + signature 3 (**16a**)
17	*Reniochalina* sp.		Axinellidae	Oman	4	5	Signature 2 (**17**)
18	*Scopalina lophyropoda*	Schmidt 1862	Dictyonellidae	Spain/Catalonia	3	6	Signature 1 (**18b**) + signature 2 (**18a**)
19	*Stylissa flabelliformis*	Hentschel 1912	Dictyonellidae	Sulawesi/Indonesia	3	3	Signature 2 (**19**)
20	*Suberites suberia*	Montagu 1818	Suberitidae	France/Roscoff	2	2	Signature 2 (**20**)
21	*Suberites suberia*	Montagu 1818	Suberitidae	France/Roscoff	2	2	Signature 1 (**21**)

for morphological work separately. Additionally, small chopped pieces of the sponge were embedded in silica gel (particle size 0.0063–0.004 mm^3) as an alternative preservation method for DNA extraction (Alvarez et al. 2000). Total DNA was extracted from the choanosome to reduce the chance of amplifying DNA templates from outside the sponge. The phenol/chloroform extraction method was used on most specimens. We homogenised the tissue in 400 µl of lysis buffer [100 mM Tris HCl, 10 mM EDTA (pH 8.0), 100 mM NaCl, 50 mM DTT, 0.5% SDS; after Hadrys et al. 1992], which contained 2 µl of 20 mg/ml proteinase K, and subsequently incubated it at 56°C for 1 h. DNA was extracted once with phenol:chloroform:isoamylalcohol (25:24:1) and then precipitated with ethanol using standard protocols. Alternatively, we used a DNA extraction kit (Quiamp DNA Mini Kit, Quiagen) and followed the manufacturer's protocol.

Primer design

Primers that amplify mitochondrial cytochrome oxidase subunit 1 (CO1) were designed from non-sponge sequences, because no poriferan homologues were available. Anthozoa are currently thought to be the earliest diverging group of cnidarians (Ax 1995) and therefore among the closest metazoan relatives to sponges. We used the CO1 homologues of the anthozoa *Metridium senile* and *Sarcophyton glaucum* (for the GenBank accession number, see Appendix) as templates to design a reverse primer (*C1-Npor 2760*, TCTAGGTAATCCA GCTAAACC). This primer was used in combination with a universal metazoan forward primer (*C1-J 2165*, GAAGTTTATATTTTAATTTTACC(AGT)GG, = *Nancy reverse*, designed by Misof et al. 2000) to amplify a 539–542 base-pair fragment of the *CO1* gene. Multiple PCRs were performed for most of the specimens (see Table 1).

PCR, cloning and sequencing

PCR amplifications (Saiki et al. 1988) were carried out on a Hybaid Thermocycler using the following 50 µl reaction mixture: 1 U SuperTaq Polymerase (Promega), 10 mM Tris HCl, 3 mM $MgCl_2$, 2 mM dNTPs (Gibco), 50 mM KCl (pH 8.3), 10 ng DNA template and primers, to reach a final concentration of 4 pmol each (Misof et al. 2000). The SuperTaq/H_20 mix was added 1 min after starting the initial denaturation. Following an initial 2-min denaturation step at 92°C, each reaction underwent 30 cycles of 92°C for 30 s, 50°C for 30 s, and 72°C for 30 s, followed by a final 5 min at 72°C. PCR products were run on a 1% agarose gel, excised and purified using GenecleanII Kit (Bio101) prior to ligation in a pGEM vector (pGEM T-easy System II, Promega), and cloned in *Escherichia coli* JM 109 (Promega). After isolation (Sambrook et al. 1989), the plasmids underwent cycle sequencing with infrared-labelled M13 primers (Biolegio) using a Deaza-Termination Kit (Amersham). Forward and reverse strands of up to nine clones (see Table 1) were sequenced with a LICOR DNA Analyzer Gene Reader IR 400 (Licor) automated sequencer. A clustal alignment was performed with AlignIR 1.2 (Licor) under default settings.

Validation of CO1 sequences

Two methods were used to confirm the taxonomic origin of the sequences: we performed a BLAST search of GenBank and we checked the position of the sequences in a phylogenetic tree.

The BLAST search provided a first estimation, but was not sufficient for our purposes: BLAST algorithms (blastn and blastp) score phenetic similarities only between taxon pairs and are highly dependent on the abundance of suitable subjects. The most important parameters are the length of the sequence and the match frequency per sequence length. As a result, shorter sequences with only a few mismatches can have a lower score than a longer sequence with more mismatches.

This systematic error made it necessary to determine the relationship of our sequences to those of other taxa, using a phylogenetic tree. We added homologous GenBank sequences from a wide range of representative taxa of other phyla to our data set (see Appendix). Fragments with more than 40% missing characters were not included, to prevent inconsistent results. We performed this analysis at the amino-acid level to keep random noise and long-branch attraction low. The sequences were translated into amino acids by subsequent use of all published genetic codes currently present in GenBank. These translated data were used as input to construct a phylogenetic tree using maximum-parsimony (MP) and minimum-evolution (ME) algorithms.

A priori analyses

Our aim was to reconstruct and compare the gene trees of sponges and associate organisms. We performed five different tests prior to the reconstructions of both gene trees, to prove hierarchical structure and detect noise due to long-branch attraction:

1. Permutation-tail probability test (PTP test)
2. Signal-content test
3. Taxon-variance analysis
4. Saturation test
5. Base-pair homogeneity test.

We employed a PTP test (Faith 1991) as implemented in PAUP*4.0 (Swofford 2000) with 100 replicates based on a branch-and-bound search to determine whether the information of our data sets arose by chance.

Phenetic and cladistic similarities of all taxon pairs were compared to estimate the phylogenetic signal of the data set (signal-content test) and to identify long-branch attraction (taxon-variance analysis) using the RASA 3.0.2.b programme (Lyons-Weiler 2001).

Saturated substitutions can bias a phylogenetic signal. We detected these positions by plotting substitution events for every taxon pair (calculated with PAUP*4.0) against p-distance (Lehmann 1998). Substitutions with a saturated distribution were excluded (zero-weighted) from parsimony analyses. We screened all characters of the data set, and the third codon position separately, for saturation in transversions and both types of transitions.

Additionally, the base frequencies were tested for homogeneity with a χ^2 test of all taxa, as implemented in PAUP*4.0.

Phylogenetic analyses

We employed the PAUP*4.0 programme package for MP, maximum-likelihood (ML) and ME analyses. Gaps were treated as missing characters and multistate characters in amino-acid sequences (based on ambiguous bases) were coded as uncertainties.

We used the branch-and-bound algorithm for smaller data sets (< 25 taxa) to reconstruct ML and MP trees. Their bootstrap tests were carried out with 1,000 replicates. On larger data sets (> 25 taxa), we performed a heuristic search with random addition of sequences with random replicates and TBR branch-swapping. Prior to the ML analysis, we determined the appropriate model of character evolution, using MODELTEST 3.0 (Posada and Crandall 1998).

Minimum-evolution analyses were based on distances that assumed ML models and compared with uncorrected p-distances.

A posteriori analyses

If there was incongruence or weak support in the phylogenetic reconstructions, alternative topologies were tested with a Kishino Hasegawa Test, as implemented in PAUP*4.0. We searched for evidence of coevolution between host and symbionts by comparing the CO1 gene trees. TreeMap 1.0b (e.g. Page 1994) was employed to plot the host tree against the symbiont tree. The search for coevolution, duplication and host-switch events was performed with the "exact search" option. Random replicates ($n = 10{,}000$) of host and species tree were created and the number of coevolution

Table 2. Pairwise standard distances between sequenced taxa. Below diagonal: total character differences. Above diagonal: mean character differences (standard differences, adjusted for missing data, calculated with PAUP 4.0). The lines separate different sequence signatures. *Grey shaded areas* indicate distances between taxa assigned to the same signature. Sequence numbers refer to Table 1

	Taxon / Number	signature 1													signature 2										sig. 3	
		13	3a	12	21	1a	2	7	8	9	15	6	16b	18b	3b	5	4	10	19	20	18a	17	14	11	1b	16a
signature 1	*Haliclona xena* 13		0.17	0.19	0.09	0.18	0.19	0.16	0.20	0.20	0.18	0.19	0.16	0.19	0.31	0.32	0.30	0.29	0.31	0.31	0.29	0.32	0.30	0.29	0.33	0.33
	Axinella damicornis 3a	91		0.21	0.18	0.16	0.18	0.19	0.22	0.21	0.17	0.18	0.20	0.16	0.32	0.32	0.30	0.28	0.32	0.30	0.32	0.32	0.29	0.29	0.32	0.34
	Halichondria panicea 12	105	114		0.19	0.21	0.16	0.20	0.01	0.02	0.16	0.15	0.19	0.22	0.30	0.31	0.29	0.30	0.30	0.31	0.30	0.30	0.31	0.31	0.33	0.37
	Suberites suberia 21	49	96	104		0.20	0.20	0.17	0.20	0.21	0.19	0.20	0.18	0.21	0.30	0.32	0.30	0.29	0.30	0.31	0.30	0.32	0.29	0.29	0.33	0.34
	Agelas oroides 1a	99	86	114	108		0.19	0.19	0.21	0.21	0.17	0.19	0.21	0.01	0.31	0.34	0.31	0.30	0.31	0.31	0.32	0.34	0.30	0.30	0.31	0.35
	Amorphinopsis excavans 2	101	95	87	109	104		0.18	0.16	0.16	0.09	0.12	0.17	0.20	0.32	0.31	0.30	0.30	0.32	0.30	0.29	0.31	0.30	0.30	0.34	0.35
	Didiscus oxeata 7	84	101	106	91	105	96		0.20	0.20	0.17	0.17	0.16	0.20	0.30	0.31	0.28	0.28	0.30	0.30	0.31	0.31	0.28	0.28	0.31	0.36
	Halichondria bowerbanki 8	109	117	8	110	116	89	108		0.02	0.16	0.16	0.20	0.23	0.31	0.32	0.30	0.31	0.31	0.32	0.31	0.31	0.32	0.32	0.33	0.36
	Halichondria bowerbanki 9	101	106	8	104	107	79	102	8		0.16	0.15	0.20	0.23	0.32	0.32	0.31	0.32	0.32	0.32	0.31	0.31	0.32	0.32	0.33	0.38
	Hymeniacidon perlevis 15	95	92	85	103	92	50	93	87	80		0.07	0.17	0.18	0.31	0.30	0.29	0.29	0.31	0.31	0.30	0.30	0.30	0.29	0.31	0.34
	Ciocalypta penicillus 6	101	99	83	108	101	64	94	84	77	38		0.18	0.19	0.32	0.30	0.29	0.30	0.32	0.31	0.31	0.29	0.30	0.30	0.34	0.36
	Liosina paradoxa 16b	88	109	104	99	111	93	88	106	98	92	95		0.21	0.30	0.29	0.28	0.28	0.29	0.30	0.27	0.29	0.28	0.29	0.32	0.33
	Scopalina lophyropoda 18b	102	89	120	112	5	108	109	122	113	96	105	116		0.32	0.35	0.31	0.30	0.32	0.32	0.33	0.35	0.31	0.31	0.32	0.35
signature 2	*Axinella damicornis* 3b	166	173	163	163	168	174	160	168	158	168	172	160	172		0.19	0.17	0.15	0.01	0.15	0.12	0.18	0.16	0.15	0.34	0.40
	Axinyssa ambrosia 5	173	173	165	172	184	168	168	170	158	162	159	158	189	99		0.17	0.17	0.18	0.15	0.18	0.01	0.18	0.18	0.36	0.41
	Axinella verrucosa 4	159	161	158	159	165	163	151	164	153	157	157	153	169	89	89		0.12	0.16	0.14	0.17	0.16	0.12	0.13	0.35	0.40
	Halichondria bowerbanki 10	157	153	163	155	159	162	148	169	158	156	160	149	163	81	91	64		0.15	0.14	0.15	0.16	0.02	0.04	0.35	0.40
	Stylissa flabelliformis 19	165	172	163	162	167	173	159	168	158	168	171	159	171	3	97	87	79		0.15	0.12	0.17	0.16	0.15	0.34	0.39
	Suberites suberia 20	167	163	166	166	167	164	163	171	157	165	169	164	172	80	79	77	74	80		0.18	0.14	0.14	0.14	0.36	0.43
	Scopalina lophyropoda 18a	155	172	160	158	173	157	164	166	153	160	164	143	177	66	98	91	79	64	94		0.18	0.16	0.16	0.34	0.40
	Reniochalina sp 17	171	171	163	170	184	167	166	168	156	161	158	154	189	95	6	85	87	93	75	94		0.17	0.17	0.36	0.40
	Hymeniacidon perlevis 14	163	157	165	157	161	164	152	171	160	160	164	153	165	86	97	66	10	84	77	83	93		0.05	0.36	0.41
	Halichondria panicea 11	154	159	166	158	163	162	152	171	161	154	161	154	167	81	95	70	22	79	73	85	89	28		0.34	0.40
sig. 3	*Agelas oroides* 1b	143	142	146	144	137	148	135	146	146	136	151	141	140	148	156	152	151	149	158	146	156	155	150		0.25
	Liosina paradoxa 16a	176	184	199	180	187	190	191	196	188	183	194	179	190	212	217	215	214	211	230	211	217	218	216	107	

events compared, to show that the observed events in our data were not based on chance alone. For the randomisations, we used the "Yule (Markovian)" model and the "proportional-to-distinguishable" model.

Results

The sequenced data and their origin

We amplified DNA fragments of 542–545 bp length. Since the sequences were coding for a protein, they could be aligned unambiguously by their codons and translated into a functional 180 amino-acid protein sequence of the cytochrome oxidase subunit 1 for the symbiont and a 181 amino-acid protein sequence for the sponge. The sequences are published in GenBank under the accession numbers AF437294-AF437314 and AY061880. The alignment is available in the EMBL-Align database (ALIGN_000244).

Our universal primer set amplified DNA from sponge and associate from the sponge sample (see Table 1). We obtained, simultaneously, sponge and non-sponge CO1 sequences from DNA extracts of *Axinella damicornis, Agelas oroides, Halichondria bowerbanki*, *Halichondria panicea*, *Hymeniacidon perlevis, Liosina paradoxa, Scopalina lophyropoda* and *Suberites suberia* (Table 1). The fact that different sequences were obtained from the same sponge specimen demonstrates that it is extremely important to sequence multiple clones from PCR-amplified gene regions if conserved primers are used. If not, molecular phylogenies can be misleading, particularly in symbiont-rich taxa, such as the Porifera.

The sequences obtained could be differentiated into three distinct groups, defined by characteristic sequence patterns (signatures, see Table 2). They were derived from three different taxonomic groups. In the following text, these three groups will be called "signature 1", "signature 2" and "signature 3".

We found no evidence for a paralogous nuclear origin of the sequences (pseudogenes; e.g. insertions/deletions) that could lead to frame shifts or point mutations, which code for translation stops. All sequences coded for a functional cytochrome oxidase subunit 1 fragment.

To determine the taxonomic source of all sequences, we included 63 CO1 sequences from GenBank which originated from a representative set of higher taxa (referring to the GenBank taxonomy browser http://www.ncbi.nlm.nih.gov/Taxonomy/taxonomyhome.html/index.cgi). The extended data set was used in the minimum-evolution and MP analyses to infer the phylogenetic position of our clones. We forced the additional taxa to cluster in monophyletic Archea, Eubacteria and Eukaryota, since this data set did not seem to support the hypotheses of the

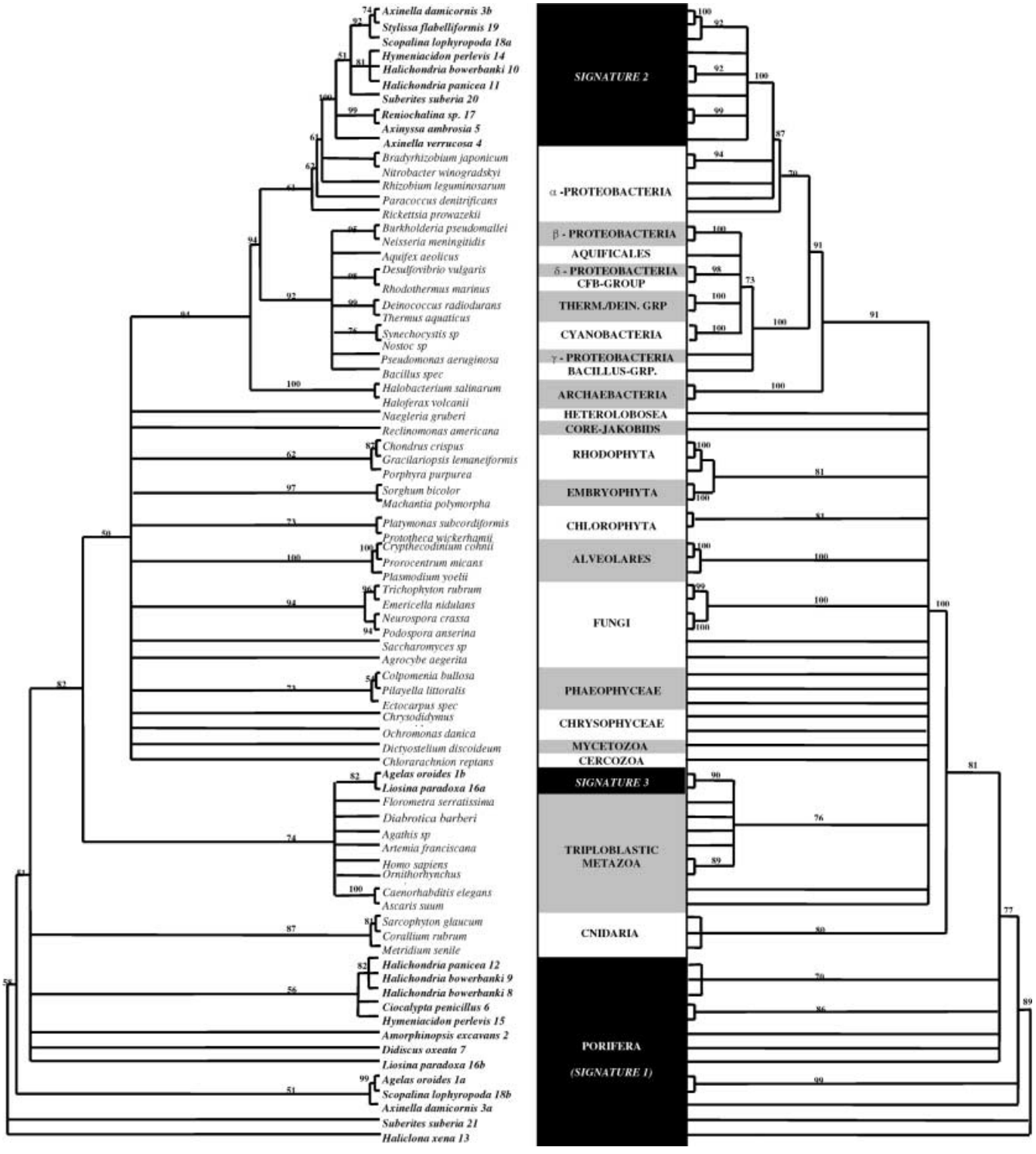
►

Fig. 1. Comparison of bootstrap 50%-majority-rule maximum-parsimony (*left*) and 70%-majority-rule-minimum-evolution (*right*) tree of sequenced taxa (*bold*) and representatives of other clades (taken from GenBank; accession numbers are given in the Appendix). The trees were reconstructed with the "three domains of life" hypothesis as "backbone constraint". Taxonomic classification is adapted from the NCBI-GenBank taxonomy browser (http://www.ncbi.nlm.nih.gov/Taxonomy/taxonomy home.html/index.cgi). Sequence numbers at taxon labels refer to Table 1; numbers at branches are bootstrap values. MP-tree values: length 2,183; CI 0.429; RI 0.522. Taxon names in brackets indicate the host of the sequenced organism

three life domains. This procedure had no major influence on the phylogenetic position of sequences, except that the bootstrap support was weaker. Eight sequences, representing Acanthamoebidae, Cryptophyta, Euglenozoa, Haptophyceae, Malawimonidae, Bacillariophyta, Xanthophyta and Mesozoa, resulted in collapsed branches or could not be aligned unambiguously. All these taxa did not cluster with the target sequences and were excluded from further analyses. This step had no influence on the results.

The phylogenetic position of our 25 sequences is shown in the trees in Fig. 1. The minimum-evolution bootstrap tree supports the results of the MP analysis. In both cases, our sequence signatures fall into three distinct groups.

Sequences assigned to "signature 1" were identified as the sponge sequences. They cluster with the cnidarian



CO1 sequences that form a monophyletic clade separate from sponges (bootstrap probability: MP 87%; ME 80%), supported by a bootstrap value of 82 (MP) and 100 (ME), respectively (Fig. 1). A more representative library of diploblast CO1 sequences is lacking.

The taxa of "signature 2" form a well supported monophyletic group (bootstrap support: MP and ME 100%). They were identified as bacteria, since they cluster in the eubacterial domain with a bootstrap probability of 94% in the MP tree and 91% in the ME tree. In our trees, they fall within the Proteobacteria of the α subdivision, supported with a bootstrap probability up to 87% in the ME tree, but less in the MP tree. Inclusion or exclusion of our sequences did not influence the phylogenetic position of the α-Proteobacteria relative to the other eubacteria, so we can reject an influence of long-branch attractors. We tested a taxa set consisting exclusively of prokaryotes and "signature 2" sequences separately from the previous analyses, to avoid bias from other taxa. The resulting MP tree (not shown here) supported our findings of a monophyletic clade comprising α-Proteobacteria and "signature 2" sequences, with a bootstrap probability of 97%, separately from all other "bacteria" of the data set.

Both "signature 3" sequences cluster distantly from the sponges ("signature 1") and α-Proteobacteria ("signature 2"). They group within the higher metazoa clade (bootstrap support: MP 74%; ME 76%), and, in a separate analysis, clustered with arthropods. This might indicate a potential contamination of the sponge extract with crustacean DNA and could be explained by the common occurrence of arthropod sponge associates.

Phylogenetic relationships of the bacterial symbionts

Comparative analyses require reliable gene trees. Therefore, we first carried out a separate phylogenetic analysis of the bacterial symbionts, to exclude confounding effects of long-branch attraction (e.g. Lyons-Weiler 2001). We determined whether a reliable outgroup relative to the sponge bacteria could be found. In theory, the best outgroup in this data set would have been the closest related α-Proteobacterium in our amino-acid tree (Fig. 1), but CO1 nucleotide sequences of *Paracoccus denitrificans, Rhizobium leguminosarum, Nitrobacter winogradskyi* and *Bradyrhizobium japonicum* were not available. *Rickettsia prowazekii*, the next closest taxon, was on a long branch in the phylogenetic tree that resulted in a significant reduction of phylogenetic signal (RASA, $t_{RASA} = 1.0378$, $df = 41$). Therefore, we performed further phylogenetic analyses without outgroup rooting.

The data set contained a significant phylogenetic signal (RASA, $t_{RASA} = 4.11$, $df = 32$, $P < 0.005$) which was significantly different from that of random data (PTP test, $P < 0.01$). Some 129 (35%) of the 363 characters in the symbiont data set were parsimony-informative. Of these, 108 belonged to the third codon position. CT transitions of their third codon position tended to slight saturation and were excluded from the further parsimony analyses. The general time-reversible model with gamma rates (GTR + Γ; Rodríguez et al. 1990) fitted the data sets best for ML reconstructions (following the likelihood-ratio test and the Akaike Information Criterion; Akaike 1974).

Tree topologies under all three different phylogenetic reconstruction methods were identical. The unrooted tree is displayed in Fig. 2. The bacteria cluster in three well supported clades. Only the phylogenetic position of the *Axinella verrucosa* symbiont could not be resolved unambiguously. We determined the root of this symbiont tree in an additional analysis of the amino-acid data. We included the symbiont and all other α-proteobacterial sequences. In this tree, *Reniochalina* sp., *Axinyssa ambrosia* and *Suberites suberia* are basal in relation to the other symbionts and split up first (data not shown).

Sponge phylogeny

The sponge data set consisted of 12 taxa with 542 characters, of which 167 (31%) were parsimony-informative. Some 131 informative characters were from the third codon position, of which, CT transitions tended to slight saturation. They were excluded from the further parsimony analyses. The data set contained a strong phylogenetic signal (RASA, $t_{RASA} = 5.9927$, $df = 62$, $P < 0.005$), was free of taxa with significant long branches (RASA taxon-variance analysis, *F*-test) and different from random data (PTP test, $P < 0.01$). MODELTEST estimated the TVM + Γ model (Rodríguez et al. 1990; likelihood-ratio test) as being

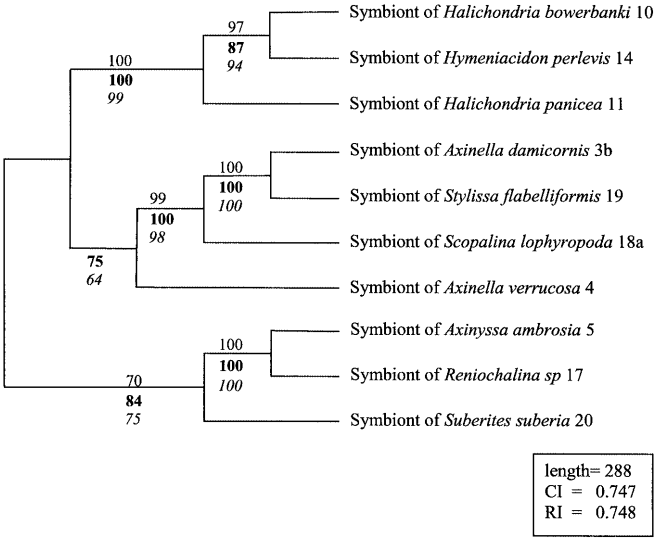


Fig. 2. Strict-consensus maximum-parsimony gene tree of the sequenced α-Proteobacteria with plotted bootstrap support values > 50: top to bottom: maximum parsimony; maximum likelihood (*bold*, GTR + Γ model); and minimum evolution under GTR + Γ distances (*italics*). Numbers at taxon labels refer to Table 1

more suitable than the HKY85 model (Hasegawa et al. 1985; AIC). The resulting tree is shown in Fig. 3: MP, minimum evolution (under p and MLdistances) and ML tree under an assumption that both substitution models are identical.

This CO1 tree is highly congruent with previously published morphological analyses (Soest et al. 1990; Soest and Lehnert 1997; Alvarez et al. 2000). *Hymeniacidon perlevis* and *Ciocalypta penicillus* cluster together and form a monophyletic clade with *Amorphinopsis excavans*. These form the sister group to the monophyletic *Halichondria* species. The sequences of *Agelas oroides* and *Scopalina lophyropoda* cluster with *Axinella damicornis* outside this clade. All these branches are well supported. The positions of *Didiscus oxeata* and *Liosina paradoxa* are more basal to the other halichondrids. *Suberites suberia* branches in this tree first from the ingroup, with 100% bootstrap probability.

Search for potential cospeciation events between sponges and the symbionts

The topologies of the *CO1* gene trees of the sequenced sponges and bacteria are highly congruent (Fig. 4). TreeMap found two equally probable reconstructions with four cospeciaton events, one sorting event and one host switch. A host switch from *Halichondria bowerbanki* to *Hymeniacidon perlevis* is equally parsimonious to a switch vice-versa. Nevertheless, the randomisation tests confirmed that the occurrence of four cospeciation events is not based on chance ($P < 0.02$).

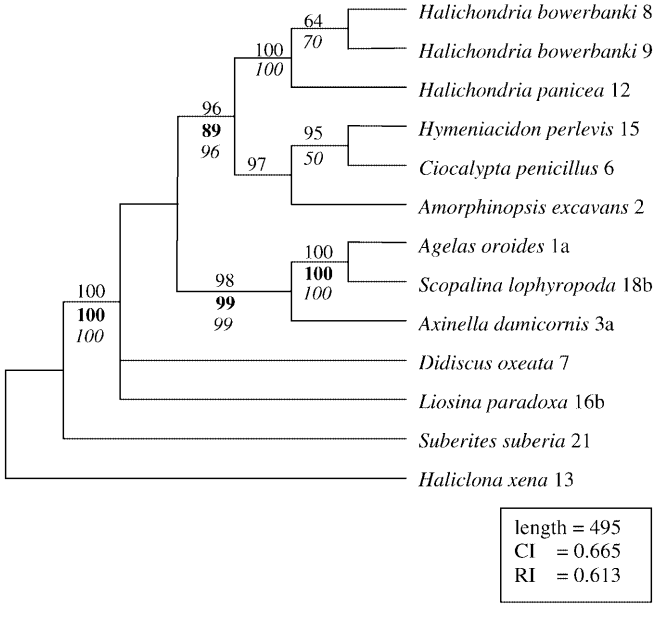


Fig. 3. Strict-consensus maximum-parsimony gene tree of the sequenced sponges with plotted bootstrap support values > 50. Above the branches, maximum parsimony; below the branches, maximum likelihood (*bold*, TVM + Γ model) and minimum evolution under TVM + Γ distances (*italics*). Numbers at taxon labels refer to Table 1

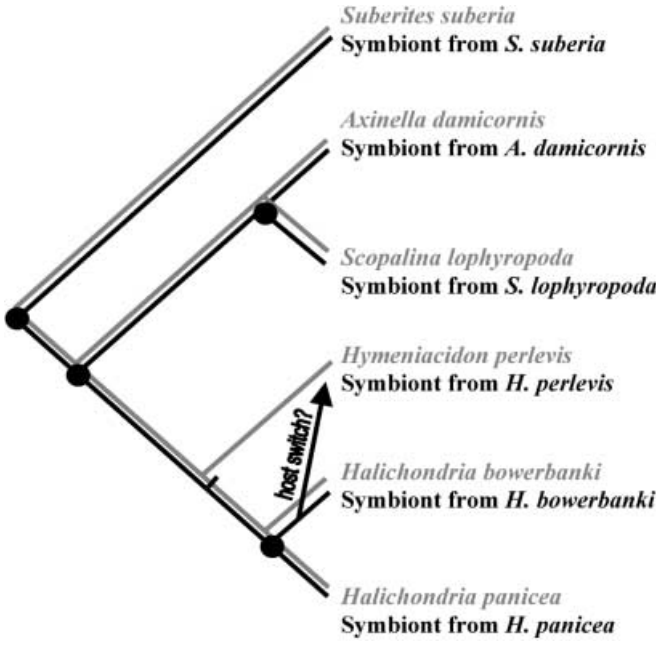


Fig. 4. One of the two "best" scenarios of coevolutionary relationships between sponges and their bacterial symbionts resulting from an "exact" analysis with TREEMAP. *Grey lines* represent sponge phylogeny; *black lines* represent the symbiont gene tree; *black dots* represent assumed cospeciation events of the symbionts; *small branches* represent sorting events; the *arrow* represents the assumed host switch (note that a host switch from *Hymeniacidon perlevis* to *Halichondria bowerbanki* is equally parsimonious)

Discussion

We obtained three different groups of sequences from the DNA that was extracted from sponge material. We can exclude nuclear mitochondrial-like sequences (NUMTS) as the source of the sequences. Frame-shift mutations or stop codons – clear but not implicit attributes of pseudogenes (Zhang and Hewitt 1996) – were not detected and all sequences could be translated from a part of the amino-acid sequence of a CO1 protein. Furthermore, we would not expect that NUMTS would show congruent phylogenetic patterns after acquiring some random mutations.

In our attempt to identify the taxonomic origin of our sequences, we have been able to identify and to separate non-sponge from sponge sequences and narrow down their taxonomic origin. Although a *CO1* gene tree may not be suitable to reconstruct a "tree of life", it will provide us, up to a certain level, a clear hint of the taxonomic origin of the sequences. One group of sequences clearly represents a part of the sponge CO1 gene, based on the fact that this group roots with the other diploblasts at the base of the Metazoa (Fig. 1). The sequences of "signature 3" are likely to be of crustacean origin. The sequences of "signature 2" can be assigned as bacterial associates. A more precise taxonomic position of the bacteria could not be determined,

owing to the lack of published α-proteobacterial CO1 sequences. To narrow down the phylogenetic position of the bacterial demosponge symbionts, additional sequence information (preferably 16SrDNA sequences) would have to be recruited. These methods may harbour a systematic error, but we extracted the necessary information from a combination of cladistic and phenetic essays. We show that, up to a certain level, our comprehensive phylogenetic analysis is suitable to detect the raw taxonomic origin of the sequences, prior to using them as determined data.

We show, on the level of CO1 sequence evolution, additional evidence for cospeciation of the bacteria with sponges. The gene tree of the sequenced bacteria is highly congruent with the reconstructed phylogeny of halichondrid sponges. That indicates that sponges and their symbionts radiate in similar patterns. Four putative cospeciation events in a six-taxa tree were found, and we prove this amount to be significantly different from random data.

Sponges are known to host a variety of species other than bacteria. Since they are filter feeders and have highly permeable membranes, some of the sequences could be remnants of ingested micro-organisms or associates that live inside or among the sponge cells. This complicates the distinction between true symbionts and occasional food micro-organisms, but these bacterial "contaminations" would not show congruent phylogenetic patterns. This significant congruence provides us with strong evidence for a specific and probably permanent (following Smith 1979 and Moulder 1979) symbiotic relationship between our sponges and the detected bacteria. The detection of vertical transmission of bacterial symbionts by sponge larvae (Gallissian and Vacelet 1976; Lévi and Lévi 1976) elucidated the mechanisms for the observed phylogenetic patterns.

The evidence for coevolution is derived from the comparison of host and parasite trees. The *CO1* gene tree, resulting from the first mitochondrial data set in sponges, is congruent with previously published morphological trees (Soest et al. 1990; Soest and Lehnert 1997; Alvarez et al. 2000), but further support for the suggested hypotheses can only be gained by including additional taxa and by the reconstruction of more, independent gene trees. Support values of both gene trees are promising, but one must still be aware that those reconstructions are based on limited information: a sequence of a single fragment out of a limited number of taxa.

The data set cannot provide us with unambiguous information about the systematic position of the sequenced bacteria. But the phylogenetic trees provide some evidence for a relationship with α-Proteobacteria, although a final proof (e.g. ultrasections) is missing. Several α-Proteobacteria are known to live in association with sponges. Symbioses have been found in *Rhopaloeides odorabile* (Webster and Hill 2001), and Guan et al. (2000) extracted α-Proteobacteria from *Jaspis johnstoni* and *Plakortis lita*. There are further reports regarding halichondrid sponges: Althoff et al. 1998 discovered *Rhodobacter*-related α-Proteobacteria in a North Sea specimen of *Halichondria panicea*.

The α -proteobacterial subdivision comprises microbes that can have different effects on their host. It contains cell parasites, such as *Wolbachia* and *Rickettsia*, that manipulate host reproduction or are pathogenic (Stouthamer et al. 1999). On the other hand, a variety of agriculturally important bacteria capable of inducing nitrogen fixation in symbiosis with plants (cf. *Rhizobium* or *Nitrobacter*) is assigned to this group (Tortora et al. 2001). Mutualistic relationships with marine Metazoa have been detected as well: subcuticular α-Proteobacteria play an important role in the nitrogen metabolism of echinoderms (Burnett and McKenzie 1997). Cospeciation of symbiotic α-Proteobacteria has been reported several times. *Wolbachia* and their nematode hosts show congruent phylogenetic patterns (Casiraghi et al. 2001) and, in marine organisms, Ashen and Goff (2000) detected coevolution between putative α-Proteobacteria and marine algae.

However, the data cannot provide a distinct explanation of the relationship between the associates of *Hymeniacidon perlevis* and *Halichondria bowerbanki*. TreeMap provides three possible hypotheses. Host switches (horizontal transmission) could have occurred in combination with a sorting event from *Hymeniacidon perlevis* to *Halichondria bowerbanki* or vice-versa. The third option would include a duplication event in the Halichondriidae symbionts, instead of a host switch: one line invades *Halichondria panicea*, the other one, *Hymeniacidon perlevis* and *Halichondria bowerbanki*. The latter would require two additional sorting events and is therefore less parsimonious. Additional data are needed to answer this question.

Acknowledgements We would like to thank Matthijs van Couwelaar, Raquel Gómez, Bert W. Hoeksema, Mario J. de Kluijver, Peter Kuperus, Bernhard Misof, Frederick R. Schram, Jan J. Vermeulen and Andrew Weeks for contributions in various ways to this paper. This work was financed by the Nederlandse Organisatie voor Wetenschappelijk Onderzoek (Netherlands Scientific Research Organisation) Nr: ALW 809.34.003. All experiments comply with the current laws of The Netherlands

Appendix

GenBank accession numbers of the representative taxa

The accession numbers for the sequenced sponges and bacteria are AF437294-AF437314 and AY061880

Taxon	GenBank accession	Taxon	GenBank accession
Agathis sp.	AF078455	*Machantia polymorpha*	M68929
Agrocybe aegerita	AF010257	*Metridium senile*	AF000023

Taxon	GenBank accession	Taxon	GenBank accession
Aquifex aeolicus	AE000777	*Naegleria gruberi*	AF288092
Artemia franciscana	5835053	*Neisseria meningitidis*	AE002522
Ascaris suum	X54253	*Neurospora crassa*	M36958
Bacillus sp.	D85547	*Nitrobacter winogradskyi*	AE006502
Bradyrhizobium japonicum	X54800	*Nostoc* sp.	AF291994
Burkholderia pseudomallei	AF087002	*Ochromonas danica*	AF287134
Caenorhabditis elegans	X54252	*Ornithorhynchus anatinus*	X83427
Chlorarachnion reptans	BAA96355	*Paracoccus denitrificans*	P98002
Chondrus crispus	Z47547	*Pilayella littoralis*	Z72500
Chrysodidymus synuroideus	AF222718	*Plasmodium yoelii*	AF055587
Colpomenia bullosa	AF037995	*Platymonas subcordiformis*	Z7447795
Corallium rubrum	AF310260	*Podospora anserina*	1703265A
Crypthecodinium cohnii	T07546	*Porphyra purpurea*	AF114794
Deinococcus radiodurans	G75251	*Prorocentrum micans*	BAA24909
Desulfovibrio vulgaris	BAA06976	*Prototheca wickerhamii*	U02970
Diabrotica barberi	AF278549	*Pseudomonas aeruginosa*	AE004449
Dictyostelium discoideum	AF295546	*Reclinomonas americana*	11466559
Diplonema papillatum	AF119813	*Rhizobium leguminosarum*	Q08855
Ectocarpus sp.	AF037994	*Rhodothermus marinus*	10185096
Emericella nidulans	P00402	*Rickettsia prowazekii*	AJ 235271
Euglena gracilis	UU49052	*Saccharomyces* sp.	P 98001
Florometra serratissima	AF049132	*Sorghum bicolor*	M14453
Gracilariopsis lemaneiformis	AF118119	*Synechocystis* sp.	D90905
Halobacterium salinarum	AE005013	*Thermus aquaticus*	P98005
Homo sapiens	BAA07292	*Trichophyton rubrum*	X65223

References

Akaike H (1974) A new look at the statistical model identification. IEEE Trans Autom Contr 19:716–723

Althoff K, Schütt C, Steffen R, Batel R, Müller WEG (1998) Evidence for a symbiosis between bacteria of the genus *Rhodobacter* and the marine sponge *Halichondria panicea*: harbor also for putatively toxic bacteria? Mar Biol 130:529–536

Alvarez B, Crisp MD, Driver F, Hooper JNA, Soest RWM van (2000) Phylogenetic relationships of the family Axinellidae (Porifera: Demospongiae) using morphological and molecular data. Zool Scr 29:169–198

Ashen JB, Goff LJ (2000) Molecular and ecological evidence for species specificity and coevolution in a group of marine algal–bacterial symbioses. Appl Environ Microbiol 66:3024–3030

Ax P (1995) Das System der Metazoa. Fischer, Stuttgart

Borchiellini C, Chombard C, Lafay B, Boury-Esnault N (2000) Molecular systematics of sponges (Porifera). Hydrobiologia 420:15–27

Borchiellini C, Manuel M, Avilon E, Boury-Esnault N, Vacelet J, LeParco Y (2001) Sponge paraphyly and the origin of Metazoa. J Evol Biol 14:171–179

Boury-Esnault N, Hajdu E, Klautau E, Custodio M, Borojevic R (1994) The value of cytological criteria in distinguishing sponges at the species level: the example of the genus *Polymastia*. Can J Zool 72:795–804

Burnett WJ, McKenzie JD (1997) Subcuticular bacteria from the brittle star *Ophiactis balli* (Echinodermata: Ophiuroidea) represent a new lineage of extracellular marine symbionts in a subdivision of the class Proteobacteria. Appl Environ Microbiol 63:1721–1724

Casiraghi M, Anderson TJC, Bandi C, Bazzocchi C, Genchi C (2001) A phylogenetic analysis of filarial nematodes: comparison with the phylogeny of *Wolbachia* endosymbionts. Parasitology 122:93–103

Distel DL, Felbeck H, Cavanaugh CM (1994) Evidence for phylogenetic congruence among sulfur-oxidizing chemoautotrophic bacterial endosymbionts and their bivalve hosts. J Mol Evol 38:533–542

Faith DP (1991) Cladistic permutation test for monophyly and nonmonophyly. Syst Zool 40:366–375

Gallissian MF, Vacelet J (1976) Ultrastructure de quelques stades de l'ovogenèse de spongiaires du genre *Verongia* (Dictyoceratida). Ann Sci Nat Zool 18:381–404

Guan LL, Onuki H, Kamino K (2000) Bacterial growth stimulation with exogenous siderophore and synthetic N-acyl homoserine lactone autoinducers under iron-limited and low-nutrient conditions. Appl Environ Microbiol 66:2797–2803

Hadrys H, Balick M, Schierwater B (1992) Applications of random amplified polymorphic DNA (RAPD) in molecular ecology. Mol Ecol 1:55–65

Hasegawa M, Kishino K, Yano T (1985) Dating the human–ape splitting by a molecular clock of mitochondrial DNA. J Mol Evol 22:160–174

Lehmann AW (1998) Artbildung, akustische Kommunikation und sexuelle Selektion bei griechischen Laubheuschrecken der *Poecilimon propinquus*-Gruppe. Diss. Friedrich-Alexander-Universität, Erlangen-Nürnberg

Lévi C, Lévi P (1976) Emryogenèse de *Chondrosia reniformis* (Nardo), Démosponge ovipare, et transmission des bactéries symbiotique. Ann Sci Nat Zool 18:367–380

Lyons-Weiler J (2001) RASA 3.0 Turbo for the Macintosh. http://bio.uml.edu/LW/RASA.html

Misof B, Erpenbeck D, Sauer KP (2000) Mitochondrial gene fragments suggest paraphyly of the genus *Panorpa* (Mecoptera, Panorpidae). Mol Phyl Evol 17:76–84

Moulder JW (1979) The cell as an extreme environment Proc R Soc Lond Ser B 1155:199–210

Page RDM (1994) Parallel phylogenies: reconstructing the history of host-parasite assemblages. Cladistics 10:155–173

Posada D, Crandall KA (1998) Modeltest: testing the model of DNA substitution. Bioinformatics 14:817–818

Preston CM, Wu KY, Molinski TF, DeLong EF (1996) A psychrophilic crenarchaeon inhabits a marine sponge: *Cenarchaeum symbiosum* gen. nov., sp. nov. Proc Natl Acad Sci USA 93:6241–6246

Rodríguez F, Oliver JF, Marín A, Medina JR (1990) The general stochastic model of nucleotide substitutions. J Theor Biol 142:485–501

Saiki RK, Gekfrand DH, Stoffel SJ, Higuchi R, Horn GT, Mullis KK, Erlich HA (1988) Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. Science 239:487–491

Sambrook J, Fritsch EF, Maniatis T (1989) Molecular cloning. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.

Sarma AS, Daum T, Müller WEG (1993) Secondary metabolites from marine sponges. Akademie gemeinnütziger Wissenschaften zu Erfurt, Ullstein-Mosby, Berlin

Schumann-Kindel G, Bergbauer M, Manz W, Szewzyk U, Reitner J (1997) Aerobic and anaerobic microorganisms in modern sponges: a possible relationship to fossilization processes. Facies 36:268–272

Smith DC (1979) From extracellular to intracellular: the establishment of a symbiosis. Proc R Soc Lond Ser B 1155: 115–130

Soest RWM van, Braekman JC (1999) Chemosystematics of Porifera: a review. Mem Queensl Mus 44:569–589

Soest RWM van, Lehnert H (1997) The genus *Julavis* de Laubenfels (Porifera: Halichondrida). Proc Biol Soc Wash 110:502–510

Soest RWM van, Diaz MC, Pomponi SA (1990) Phylogenetic classification of the Halichondrids (Porifera, Demospongiae). Beaufortia 40:15–62

Stierle AC, Cardellina JH, Singleton FL (1988) A marine *Micrococcus* produces metabolites ascribed to the sponge *Tedania ignis*. Experientia 44:1021–1021

Stouthamer R, Breeuwer JAJ, Hurst GDD (1999) *Wolbachia pipientis*: microbial manipulator of arthropod reproduction. Annu Rev Microbiol 53:71–102

Swofford DL (2000) PAUP*: phylogenetic analysis using parsimony (*and other methods), version 4. Sinauer, Sunderland, Mass.

Tortora GJ, Funke BR, Case CL (2001) Microbiology: an introduction, 7th edn. Addison Wesley Longman, San Francisco

Unson MD, Holland ND, Faulkner DJ (1994) A brominated secondary metabolite synthesized by the cyanobacterial symbiont of a marine sponge and accumulation of the crystalline metabolite in the sponge tissue. Mar Biol 119:1–11

Vacelet J (1970) Description de cellules á bactéries intranucléaires chez des éponges *Verongia*. J Microsc 9:333–346

Webster NS, Hill RT (2001) The culturable microbial community of the Great Barrier Reef sponge *Rhopaloeides odorabile* is dominated by an α-Proteobacterium. Mar Biol 138:843–851

Wilkinson CR (1978) Microbial associations in sponges. II. Numerical analysis of sponge and water bacterial populations. Mar Biol 49:169–176

Wilkinson CR (1983) Phylogeny of bacterial and cyanobacterial symbionts in marine sponges. Endocytobiology 2:993–1002

Wilkinson CR (1984) Immunological evidence for the Precambrian origin of bacterial symbioses in marine sponges. Proc R Soc Lond Ser B 220:509–517

Zhang DZ, Hewitt GM (1996) Nuclear integrations: challenges for mitochondrial DNA markers. Trends Ecol Evol 11:247–251