Molecular Phylogenetics and Evolution 63 (2012) 702–713

Contents lists available at SciVerse ScienceDirect

Molecular Phylogenetics and Evolution

journal homepage: www.elsevier.com/locate/ympev

ALG11 – A new variable DNA marker for sponge phylogeny: Comparison of phylogenetic performances with the 18S rDNA and the COI gene

Frida Belinky [a,1], Amir Szitenberg [a,1], Itay Goldfarb [a,1], Tamar Feldstein [a], Gert Wörheide [b,c], Micha Ilan [a], Dorothée Huchon [a,d,*]

[a] Department of Zoology, George S. Wise Faculty of Life Sciences, Tel-Aviv University, Tel-Aviv 69978, Israel
[b] Department of Earth- and Environmental Sciences & GeoBioCenterLMU, Ludwig-Maximilians-Universität München, Richard-Wagner-Str. 10, 80333 München, Germany
[c] Bavarian State Collections of Palaeontology and Geology, Richard-Wagner-Str. 10, 80333 München, Germany
[d] National Evolutionary Synthesis Center, 2024 W. Main Str., Suite A200 Durham, NC 27705-4667, USA

ARTICLE INFO

Article history:
Received 21 March 2011
Revised 13 February 2012
Accepted 14 February 2012
Available online 25 February 2012

Keywords:
ALG11
COI
18S rDNA
Porifera
Cnidaria
Molecular phylogeny
Saturation

ABSTRACT

Phylogenetic relationships within sponge classes are highly debated. The low phylogenetic signal observed with some current molecular data can be attributed to the use of few markers, usually slowly-evolving, such as the nuclear rDNA genes and the mitochondrial COI gene. In this study, we conducted a bioinformatics search for a new molecular marker. We sought a marker that (1) is likely to have no paralogs; (2) evolves under a fast evolutionary rate; (3) is part of a continuous exonic region; and (4) is flanked by conserved regions. Our search suggested the nuclear ALG11 as a potential suitable marker. We next demonstrated that this marker can indeed be used for solving phylogenetic relationships within sponges. Specifically, we successfully amplified the ALG11 gene from DNA samples of representatives from all four sponge classes as well as from several cnidarian classes. We also amplified the 18S rDNA and the COI gene for these species. Finally, we analyzed the phylogenetic performance of ALG11 to solve sponge relationships compared to and in combination with the nuclear 18S rDNA and the COI mtDNA genes. Interestingly, the ALG11 marker seems to be superior to the widely-used COI marker. Our work thus indicates that the ALG11 marker is a relevant marker which can complement and corroborate the phylogenetic inferences observed with nuclear ribosomal genes. This marker is also expected to contribute to resolving evolutionary relationships of other apparently slow-evolving animal phyla, such as cnidarians.

1. Introduction

Sponges are highly diversified with an estimated 15,000 living species (Hooper and Van Soest, 2002). Because sponges are both morphologically simple and diverse, sponge morphological systematics is known to be exceedingly difficult (Boury-Esnault, 2006; Hooper and Van Soest, 2002). As a case in point, numerous sponge orders that had been defined based on morphological similarities (e.g., Poecilosclerida, Halichondrida and Haplosclerida) have since been recovered as polyphyletic in molecular phylogenies (Erpenbeck et al., 2005a, 2007a,b; McCormack et al., 2002). DNA sequences have thus recently become a critical source of information in the clarification of sponge evolutionary relationships (Wörheide and Erpenbeck, 2007).

Previous sponge phylogenetic studies beneath the class level considered relatively large sponge samplings, but relied on only a few markers (usually the nuclear rDNA genes: 18S and 28S rDNA, or the mitochondrial COI gene) (Cárdenas et al., 2010; Erpenbeck et al., 2005a, 2007b; Nichols, 2005; Voigt et al., 2008). Both nuclear rDNAs and the COI gene have their advantages. For example, the rDNA genes possess both conserved stems and variable loop regions which provide signals for different levels of phylogenetic inference (Voigt et al., 2008). Moreover, their amplification is facilitated by their high copy number in the genome and suitable PCR primers are available for a variety of taxa. Consequently, rDNA genes, and the 18S rDNA in particular, have been widely used to infer the evolutionary relationships among Metazoa (Aguinaldo et al., 1997; Collins, 1998; Medina et al., 2001; Peterson and Eernisse, 2001; Voigt et al., 2008; Zrzavý et al., 1998). Similar to the rDNA genes, mitochondrial markers have the advantage of being present in high copy number in each cell. They also present an independent source of information when compared to the nuclear genome. Among the 13 mitochondrial protein-coding genes shared by all animal mitochondrial genomes, the COI gene is the most

* Corresponding author at: Department of Zoology, George S. Wise Faculty of Life Sciences, Tel-Aviv University, Tel-Aviv 69978, Israel. Fax: +972 3 640 9403.
E-mail address: huchond@post.tau.ac.il (D. Huchon).
[1] These authors contributed equally to this work.

1055-7903/$ - see front matter © 2012 Elsevier Inc. All rights reserved.
doi:10.1016/j.ympev.2012.02.008

sequenced, as indicated by the number of entries in the nucleotide database of NCBI (Sayers et al., 2011). Utilization of this gene was promoted by the existence of universal primers, which facilitate its amplification in most metazoan species (Folmer et al., 1994). Additionally, the low sequence variability observed within species when compared to the higher variability among species has made it the marker of choice for the barcoding of animals in general (Hebert et al., 2003, 2004) and sponges in particular (Pöppe et al., 2010; Solé-Cava and Wörheide, 2007; Wörheide and Erpenbeck, 2007) but see (Erpenbeck et al., 2006; Huang et al., 2008). Given the rather large, and constantly increasing, number of sequences available for COI, this gene is often included as a marker in phylogenetic analyses of sponge evolutionary relationships (Addis and Peterson, 2005; Cárdenas et al., 2010, 2011; Erpenbeck et al., 2007a, 2008; Nichols, 2005).

There are nonetheless several disadvantages to these two markers. It is generally assumed that all rDNA copies are homogenized. However, several cases of divergent 18S rDNA copies within the same genome have been reported in Metazoa (Carranza et al., 1996; Krieger and Fuerst, 2002; Papillon et al., 2006), thus rendering the rDNA marker less reliable than previously thought. In sponges, ribosomal intragenomic variation has been shown to affect phylogenetic inferences based on internal transcribed spacer (ITS) sequences (Alvarez et al., 2007; Redmond and McCormack, 2009; Wörheide et al., 2004) but not 18S rDNA or 28S rDNA inferences. Additionally, although the 18S rDNA provides high statistical support for many metazoan clades, being a slow-evolving marker it may fail to provide enough phylogenetic signal to resolve relationships at the family level among slow-evolving members of phyla such as Porifera and Cnidaria (Berntson et al., 2001; Dohrmann et al., 2006). The COI gene is also a slow evolving marker in sponges and anthozoans (Huang et al., 2008) and, indeed, COI sometimes fails to robustly resolve sponge relationships when used as a phylogenetic marker (Addis and Peterson, 2005; Erpenbeck et al., 2007a; Nichols, 2005). Generally, the phylogenetic relevance of COI should be evaluated for sponges, in particular since hexactinellid sponges possess a different mitochondrial genetic code than demosponges, which could affect COI-based phylogenetic reconstructions (Haen et al., 2007). Additionally, the presence of a mitochondrial intron inserted at the same location as the reverse barcoding primer complicates its amplification in sponges (Rot et al., 2006; Szitenberg et al., 2010). Finally, evolutionary relationships inferred from a single gene marker may be affected by various types of phylogenetic artifacts. However, not all genes are expected to be affected by the same biases. For example, one marker may be sensitive to the long-branch attraction artifact (Felsenstein, 1978), while a second marker may be much less sensitive to this bias. This can happen, for example, when a specific marker is fast-evolving within a specific clade because of gene-specific relaxation of selection. The above arguments suggest that more than one phylogenetic marker should be used to infer a species tree, and incongruences among markers should be studied for potential biases. Indeed, the combination of several gene markers generally improves the inference of evolutionary relationships when compared to the analysis of separate markers (Huchon et al., 2002; Mitchell et al., 2000; Russo et al., 1996).

Three approaches can be used to increase the number of gene markers: (1) large-scale sequencing approaches, such as EST projects (e.g., Dunn et al., 2008; Philippe et al., 2009; Pick et al., 2010); (2) the sequencing of specific genomic DNA markers (e.g., Borchiellini et al., 1998; Erpenbeck et al., 2005b); and (3) the sequencing of specific cDNA markers (Sperling et al., 2007, 2009). In the first approach, data originate from high throughput genomic or EST sequencing. Indeed, recent studies aiming to clarify the position of sponges among metazoans, as well as the relationships among the four sponge classes (Hexactinellida, Calcarea, Demospongiae and Homoscleromorpha), have considered numerous genes, albeit from a relatively small sampling of the sponge diversity (usually less than 10 sponge species (Belinky et al., 2010; Dunn et al., 2008; Philippe et al., 2009; Pick et al., 2010; Rokas et al., 2005; Schierwater et al., 2009)). While such approaches are clearly powerful, they are usually more costly than classical genomic DNA sequencing, and thus they often come at the expense of large taxonomic sampling. In addition, these high throughput approaches cannot always be applied to all samples as they require both high quality and high quantity of DNA or RNA samples. In particular, these methods are often inapplicable for most museum specimens, in which RNA, and to a lesser extent DNA, is often degraded.

The second and third approaches are to sequence independently specific markers that, in combination with existing markers, will improve the resolution of sponge phylogeny. Identification of a novel phylogenetic marker for sponges is not an easy task. Phylogenetic markers should be long enough (minimum 500 bp), additionally, they should neither be too conserved nor too variable. Since intronic sequences and intergenic regions are usually too variable to be used as informative markers, except at the intra-specific level (Bentlage and Wörheide, 2007; Wörheide et al., 2008), the second approach is to sequence long exonic regions for phylogenetic reconstructions above the species level. An alternative by which to avoid intronic sequence is that of the third approach: cDNA sequencing (Sperling et al., 2007, 2009). However, similar to the large-scale sequencing approach, cDNA sequencing is not possible for most museum samples for which RNA is not preserved.

In all three methods, the markers considered should be unlikely to have paralogous copies, since erroneous inference of orthology can mislead phylogenetic inference (Martin and Burg, 2002). Finally, for the two last approaches the marker should also be flanked by conserved regions to allow the design of universal sponge primers. Searching for such markers is a challenging task given the limited current knowledge regarding the sponge nuclear genome, apart from *Amphimedon queenslandica* (Srivastava et al., 2010). Notably, while databases of orthologous genes exist for representatives of the mammalian diversity (Ranwez et al., 2007), no such tool exists for sponges. As a case in point, no sponge species have yet been included in major comparative genome databases such as HomoloGene (Sayers et al., 2011), Ensembl (Flicek et al., 2011) or the UCSC Genome Browser (Fujita et al., 2011).

In this study, we conducted a bioinformatics search for a novel sponge nuclear marker. Three markers were selected, however, successful PCR amplification could only be achieved for the asparagine-linked glycosylation 11 protein (ALG11, also known as the alpha-1, 2-mannosyltransferase in yeast). We here compare the phylogenetic performance of the ALG11 gene with those of the 18S rDNA and COI gene.

2. Materials and Methods

2.1. Screen for a novel gene marker

In order to identify a fast-evolving nuclear exon, we selected homology groups from the National Center for Biotechnology Information (NCBI) HomoloGene database build 46 that fitted two criteria. First, the group had to include exactly one representative each from *Homo sapiens*, *Drosophila melanogaster*, *Caenorhabditis elegans* and *Saccharomyces cerevisiae*. This criterion provides a rough means by which to eliminate duplicated gene families. Second, the nucleotide coding sequences of *H. sapiens*, *D. melanogaster* and *C. elegans* had to include at least one exon longer than 500 bp. Since we aimed to amplify an exonic region, this second criterion was used to filter genes harboring only short exons. After this initial search, 29 homology groups remained.

We furthered narrowed the list of candidate markers by keeping only genes that are present in *Nematostella vectensis* and harbor a long exon (this step was performed before publication of the *A. queenslandica* genome). To this end, BLASTp searches were conducted against all *N. vectensis* predicted protein sequences available in StellaBase (Sullivan et al., 2006) using the *H. sapiens* sequence of each of the 29 homology groups as query. Following Hall (2004, p. 16)_, only sequences with an *E*-value $\leqslant 10^{-5}$ were assigned to a HomoloGene group index according to the best hit it received in the similarity search. A reciprocal BLASTp search, against all human protein sequences present in the NCBI protein database (http://www.ncbi.nlm.nih.gov/sites/entrez?db=protein), was then performed using as query the 29 *N. vectensis* sequences selected in the previous step. This step ensured that no other human sequence was more closely related to one of the 29 *N. vectensis* sequences than the one in the HomoloGene group.

To determine if the genes selected could possess a long intronless region in Porifera, protein sequences of each homology group were aligned. Intron positions were extracted from StellaBase for *N. vectensis* and from NCBI for all other species, and converted into positions in the aligned protein sequences using Perl scripts (see Supplementary material). The protein alignments were examined manually. Conserved proteins, highly variable proteins, and proteins for which no intronless region was shared among species were eliminated. Three candidate markers were selected at the end of this search: the asparagine-linked glycosylation 11 homolog (ALG11), the asparagine-linked glycosylation 2 homolog (ALG2), and the solute carrier family 33 (acetyl-CoA transporter), member 1 (SLC33A1).

For each candidate a BLASTp search was conducted using the Human sequence as query and a phylogenetic tree was built to confirm that the markers were indeed single copy in all animals. The result of such BLASTp search is presented, for the ALG11 gene, in Supplementary material (Fig. S1). PCR primers were designed for each marker based on the multiple sequence alignment of fungi, choanoflagellate and animal sequences (see Supplementary material, Fig. S2 for the ALG11 gene). Only the ALG11 gene could be successfully amplified in species representatives of the sponge diversity (see below).

2.2. DNA extraction, amplification, and sequencing

The source of the DNA and tissue samples used in this work are listed in Table 1. Tissue samples, cleaned of visible epibionts, were reduced to powder in a mortar with liquid nitrogen. Genomic DNA was extracted from 95% ethanol-preserved tissue samples following the procedure of Fulton et al. (1995).

Amplifications of the ALG11 sequences were performed in two steps. A first amplification was performed with the primers ALG11-D1 and ALG11-R1, followed by a re-amplification of the initial PCR

Table 1
Origin of the tissue samples. Fresh tissue samples are samples that were kept in ethanol at −80 °C and extracted few months after collection. Old tissue samples are samples that were kept in ethanol at −80 °C and extracted 2 or 4 years after collection. Museum samples are samples from museum origin. In our case, these samples had been collected and kept frozen in ethanol or DMSO at least a few years before receiving them. All museum samples were shipped by mail oversea at room temperature, and placed at −80 °C upon reception. Poecilosclerida is a polyphyletic order, *Biemna* and *Ectyoplasia* belong to families that are not related to other Poecilosclerida families (Erpenbeck et al., 2007a,b). Voucher specimens are indicated by a *. BELUM, Ulster Museum (Belfast, Northern Ireland); DH, Dorothee Huchon lab collection; GW, Gert Wörheide lab collection; QM, Queensland Museum (Brisbane, Australia); SMF; Senckenberg Museum (Frankfurt am Main, Germany); TAU, Zoological Museum, Tel Aviv University (Tel-Aviv, Israel); UCMP, University of California Museum of Paleontology (Berkeley, USA).

Phylum (class)/order	Species	Tissue number/voucher number*	Origin	Donor/collector	Tissue sample quality
Cnidaria (Anthozoa)	*Stylophora pistillata*	DH GS3	Red Sea, Eilat, Israel	Alain Daniel	DNA sample
Cnidaria (Hydrozoa)	*Hydractinia echinata*	DH S188	Lab culture	Sharon Gild	DNA sample
Cnidaria (Scyphozoa)	*Pelagia noctiluca*	TAU CO35436*	Mediterranean Sea, Ashdod, Israel	Chana Feldstein	Fresh tissue sample
	Rhopilema nomadica	DH S143	Mediterranean Sea, Akko, Israel	Tamar Feldstein	Fresh tissue sample
Porifera (Demospongiae)					
Chondrosida	*Chondrosia reniformis*	TAU 25193*	Mediterranean Sea, Achziv, Israel	Chagai Rot	Old tissue sample
Chondrosida	*Chondrilla* sp. (*Cf. nucula*)	DH S013	Mediterranean Sea, Achziv, Israel	Chagai Rot	Old tissue sample
Dictyoceratida	*Ircinia* sp.	TAU 25496*	Mediterranean Sea, Achziv, Israel	Tamar Feldstein	Fresh tissue sample
Dictyoceratida	*Sarcotragus spinosulus*	TAU 25501*	Mediterranean Sea, Neve-Yam, Israel	Tamar Feldstein	Fresh tissue sample
Poecilosclerida	*Desmapsamma anchorata*	UCMP WC1660*	Bocas del Toro, Panama	Scott A. Nichols	Museum sample
Poecilosclerida	*Negombata magnifica*	TAU 25198*	Red Sea, Eilat, Israel	Chagai Rot	Old tissue sample
"Poecilosclerida"	*Biemna fistulosa*	TAU 25567*	Bahamas	Micha Ilan	DNA sample
"Poecilosclerida"	*Ectyoplasia ferox*	UCMP WC970*	Galeta, Panama	Scott A. Nichols	Museum sample
Spirophorida	*Cinachyrella levantinensis*	TAU 25456*	Mediterranean Sea, Achziv, Israel	Chagai Rot	Fresh tissue sample
Spirophorida	*Craniella* sp.	QM G316372*	South Norfolk Ridge (Norfanz cruise)	John N.A. Hooper	Museum sample
Porifera (Homoscleromorpha)	*Plakinastrella* sp.	UCMP WC938*	Bocas del Toro, Panama	Scott A. Nichols	Museum sample
	Oscarella sp.	DH S157	Mediterranean Sea, Achziv, Israel	Dorothee Huchon	Fresh tissue sample
Porifera (Hexactinellida)					
Lyssacinosida	*Walteria leuckarti*	SMF 10522*	Sagami bay, Japan	Dorte Janussen	Museum sample
Lyssacinosida	*Caulophacus arcticus*	SMF 10520*	Arctic Ocean	Dorte Janussen	Museum sample
Hexactinosida	*Aphrocallistes vastus*	TAU 25566*	Barkley Sound (Canada)	Sally P. Leys	Museum sample
Hexactinosida	*Heterochone sp.*	SMF 10523*	Sagami bay, Japan	Dorte Janussen	Museum sample
Porifera (Calcarea)					
Calcaronea	*Grantessa* sp.	GW 974	Great Barrier Reef (Australia)	Gert Wörheide	Museum sample
Calcaronea	*Leuconia nivea*	BELUM Mc3329*	Lee's wreck, Northern Ireland	Bernard E. Picton	Museum sample
Calcinea	*Leucetta chagosensis*	GW 945	Great Barrier Reef (Australia)	Gert Wörheide	Museum sample
Calcinea	*Clathrina* sp.	DH S134	Bahamas	Micha Ilan	Old tissue sample

product using the nested primers ALG11-D2 and ALG11-R2. Primer sequences are indicated in Table 2. There were, however, two exceptions to this protocol. For *Craniella* sp. 3878 (Demospongiae), the first amplification was performed with the primers ALG11-D1 and ALG11-R1 but the re-amplification was done with the primers ALG11-D2-new and ALG11-R2-new. For Calcarea species the first amplification was performed with the primers ALG11-D1 and ALG11-R2(t) followed by a re-amplification of the initial PCR product into two overlapping fragments using the primers ALG11-D2 and ALG11-R3 for the first fragment and the primers ALG11-F3 and ALG11-R2(t) for the second fragment.

The 18S rDNA gene was PCR-amplified using the primer sets 18S1/18S2 (Borchiellini et al., 2001). Since the quantity of DNA obtained was usually insufficient for direct sequencing, the PCR products obtained were re-amplified in two overlapping fragments of approximately 1 kb and 1.2 kb using the primer-pairs 18S1/18S6 and 18S7/18S2 (Table 2). For *Craniella* sp. 3878, the 18S rDNA was amplified with the primers 18S-poe-D and 18S-poe-R (Table 2), yielding a slightly shorter fragment.

Following Rot et al. (2006), the COI gene from sponges was amplified with the primer LCO1490 and COX1R1, with a few exceptions. Irciniid COI was amplified with the newly-designed primers Cox_Irc_D1 and Cox_Calc_R1 followed by a re-amplification of the initial PCR product using the nested primers Cox_Irc_LCO and Irc_Cox1R1 (Table 2). The *Walteria* (Hexactinellida) sequence was amplified with the primers HEX_COX1D1 and Cox_Calc_R1. The PCR products obtained were then re-amplified in two semi-nested reactions with the internal primer set HEX_COX1D1/HEX_COX1D3 and Cox_Calc_R1/Cox_Calc_R2 (Table 2). Unfortunately, no COI sequences could be obtained for any Calcarea samples. Amplifications of Cnidaria COI were performed in two steps. A first amplification was performed with the primers Clath_COX_F1 and Cnidr_R1_COX, followed by a re-amplification of the initial PCR product using the nested primers Cnidaria_LCO and Cnidr_R2_COX (or with the reverse primer Clath_COX_R1 for *Hydractinia echinata*). Primer sequences are indicated in Table 2.

PCR products were purified using polyethylene glycol (PEG) in saline (NaCl). Most of the products were directly sequenced using Big Dye Terminator v1.1 (Applied Biosystems) on an ABI 310 sequencer. In a few cases, PCR products were ligated into the pSC-A vector. Ligation products were then transformed into StrataClone™ SoloPack® competent cells (Stratagene). In such cases, five clones per species were sequenced. All genes were completely sequenced on both strands. The 52 new sequences (Table 3) have been deposited in the European Nucleotide Archive of EMBL under Accession Numbers FR819667–FR819693 and HE591446–HE591470.

2.3. Alignment generations

In addition to the 52 sequences obtained as described above, 18S rDNA, COI and ALG11 sequences of species representing the major animal clades were retrieved from GenBank and trace data(-Table 3). Protein sequences of the COI and ALG11 genes and DNA sequences of the 18S rDNA gene were aligned with MAFFT under the L-INS-I strategy (Katoh et al., 2005) using the GUIDANCE webserver (Penn et al., 2010). Unreliable positions (i.e., position with a GUIDANCE threshold below 0.93) were excluded from the alignment as well as positions with 50% or more of missing data. The DNA sequences of COI and ALG11 were then aligned according to the protein sequence alignments using the program PAL2NAL (Suyama et al., 2006). For the ALG11 and the COI genes two datasets were considered: the first dataset includes all codon positions, in the second dataset the third codon positions were excluded using PAUP* v4.0b10 (Swofford, 2000). Sequence alignments are available as Supplementary material and in the Dryad repository: doi:10.5061/dryad.20617h5g.

2.4. Phylogenetic analyses

Phylogenetic analyses were conducted on each gene separately (with and without third codon positions for the ALG11 and COI

Table 2
Primer sequences. Name, sequence and direction of the primers used to amplify the ALG11, 18S rDNA and COI genes. Unless otherwise stated, primers were designed in this study. (1) Borchiellini et al. (2001), (2) Folmer et al. (1994), (3) Rot et al. (2006).

Gene	Primer name	Sequence	Direction	Usage	Ref.
ALG11	ALG11-D1	5′-TTYCAYCCNTAYTGYAAYGCNGGNGG-3′	Forward	External	
	ALG11-R1	5′-ATNCCRAARTGYTCRTTCCACAT-3′	Reverse	External	
	ALG11-D2	5′-TGYAAYGCNGGNGGNGGNGGNGA-3′	Forward	Re-amplification	
	ALG11-R2	5′-CCRAARTGYTCRTTCCACATNGTRTG-3′	Reverse	Re-amplification	
	ALG11-R3	5′-GTCCANGANGARTTNACCAT-3′	Reverse	Re-amplification	
	ALG11-F3	5′-CAYTAYCCNACNATHWSNACNGAYATG-3′	Forward	Re-amplification (Calcarea)	
	ALG11-D2-new	5′-GGNGGNGARMGNGTNYT-3′	Forward	Re-amplification	
	ALG11-R2-new	5′-ATGWDSRAYGARCAYTTYGG-3′	Reverse	Re-amplification	
	ALG11-R2(t)	5′-CAYACNATGTGGAAYGARCAYTTYGG-3′	Reverse	External (Calcarea)	
18S rDNA	18S1	5′-AACCTGGTTGATCCTGCCA-3′	Forward	External	
	18S2	5′-TGCAGGTTCACCTACAGAA-3	Reverse	External	(1)
	18S6-sponge	5′-CCTTCCGTCAATTCCTTTAAGT-3′	Reverse	Re-amplification	(1)
	18S7-sponge	5′-CGAAAGCATTTGCCAAGGATGTT-3′	Reverse	Re-amplification	
	18S-poe-D	5′-AGTCATATGCTTGTCTCAAAG-3′	Forward	Re-amplification	
	18S-poe-R	5′-CCTTGTTACGACTTTTACTTCCT-3′	Forward	Re-amplification	
COI	LCO1490	5′-GGTCAACAAATCATAAAGATATTGG-3′	Forward	External	(2)
	COX1 R1	5′-TGTTGRGGGAAAAARGTTAAATT-3′	Reverse	External	(3)
	COX1 D5	5′-CARCAYTTATTTTGATTYTTTGG-3′	Forward	Re-amplification	
	HEX COX1D1	5′-GCHTTTATAGGAACNTCTYTRAG-3	Forward	External	
	HEX COX1D3	5′-CACATAATCCCYTACYTAACAGG-3′	Forward	Re-amplification	
	Cox_Calc_R2	5′-CCWGTTARWCCGCCTATKGTRAATA-3′	Reverse	Re-amplification	
	Cox_Calc_R1	5′-AARAARTGTTGRGGGAARAADGT-3′	Reverse	External	
	Cox_Irc_D1	5′-TRTTTTCCACTAAYCAYAARGAYAT-3′	Forward	External	
	Irc_Cox1R1	5′-TGTTGMGGGAAAAAAGTYAARTT-3′	Reverse	Re-amplification	
	Clath-COX-F1	5′-TGRTTYACNWSNAAYCAYAARGANAT-3′	Forward	External/re-amplification	
	Clath-COX-R1	5′-ACRTARTGRAARTGNGCNACNACRTARTA-3′	Reverse	Re-amplification	
	Cnidaria_LCO	5′-TTYTCWACNAAYCAYAARGAYATHGG-3′	Forward	Re-amplification	
	Cnidr-R1-COX	5′-CCYARRAARTGYTGDGGRAARAA-3′	Reverse	External	
	Cnidr-R2-COX	5′-ARRAARTGYTGDGGRAARAADGT-3′	Reverse	Re-amplification	

Table 3
Accession numbers of the sequences. (1) COI *Nematostella* sp.; ALG11, 18S rDNA *N. vectensis* (2) COI *Plakortis angulospiculatus*; ALG11, 18S rDNA *Plakinastrella* sp. (3) COI *Oscarella carmella*; ALG11, 18S rDNA *Oscarella* sp. Sequences in bold are newly obtained sequences.

	Species	ALG11	COI	18S rDNA
Deuterostomia				
Chordata	*Homo sapiens*	NM_001004127	X93334	X03205
Cephalochordata	*Branchiostoma floridae*	Trace data	NC_000834	M97571
Echinodermata	*Strongylocentrotus purpuratus*	XM_782184	NC_001453	Trace data
Hemichordata	*Saccoglossus kowalevskii*	Trace data	NC_007438	Trace data
Protostomia				
Arthropoda (Hymenoptera)	*Apis mellifera*	XM_624909	NC_001566	AY703484
Arthropoda (Coleoptera)	*Tribolium castaneum*	XM_964389	NC_003081	Trace data
Chelicerata	*Ixodes scapularis*	DS735445	Trace data	Trace data
Mollusca	*Aplysia californica*	Trace data	NC_005827	AY039804
Mollusca	*Crassostrea gigas*	EST data	NC_001276	AB064942
Annellida	*Capitella* sp.	Trace data	Trace data	Trace data
Placozoa				
	Trichoplax adhaerens	XM_002112768	NC_008151	AY652578
Cnidaria				
Anthozoa	*Stylophora pistillata*	**FR819667**	**FR819681**	**FR819686**
Anthozoa	*Nematostella* sp. (1)	XM_001626758	NC_008164	Trace data
Hydrozoa	*Hydra magnipapillata*	XM_002162520	NC_011221	Trace data
Hydrozoa	*Hydractinia echinata*	**HE591446**	**HE591456**	**HE591463**
Scyphozoa	*Pelagia noctiluca*	**HE591447**	**HE591457**	**HE591464**
Scyphozoa	*Rhopilema nomadica*	**HE591448**	**HE591458**	**HE591465**
Porifera (Demospongiae)				
Chondrosida	*Chondrosia reniformis*	**FR819670**	AM076986	**FR819689**
Chondrosida	*Chondrilla* sp. (*Cf. nucula*)	**FR819671**	**FR819682**	**FR819690**
Dictyoceratida	*Ircinia* sp.	**HE591449**	**HE591459**	**HE591466**
Dictyoceratida	*Sarcotragus spinosulus*	**HE591450**	**HE591460**	**HE591467**
Haplosclerida	*Amphimedon queenslandica*	Trace data	NC_008944	Trace data/(Voigt et al., 2008)
Poecilosclerida	*Desmapsamma anchorata*	**HE591451**	**HE591461**	**HE591468**
Poecilosclerida	*Negombata magnifica*	**FR819668**	NC_010171	**FR819687**
"Poecilosclerida"	*Biemna fistulosa*	**FR819669**	AM076982	**FR819688**
"Poecilosclerida"	*Ectyoplasia ferox*	**HE591452**	**HE591462**	EU702415
Spirophorida	*Cinachyrella levantinensis*	**FR819672**	AM076987	HM629802
Spirophorida	*Craniella* sp.	**HE591453**	HM032748	**HE591469**
Porifera (Homoscleromorpha)				
Homoscleromorpha	*Plakinastrella* sp. (2)	**FR819673**	NC_010217	**FR819691**
Homoscleromorpha	*Oscarella* sp. (3)	**FR819674**	NC_009090	**FR819692**
Porifera (Hexactinellida)				
Lyssacinosida	*Walteria leuckarti* (Sagami bay)	**FR819675**	**FR819683**	AM886399
Lyssacinosida	*Caulophacus* sp. (art XXIII-1 Artic exp-05)	**FR819676**	**FR819684**	AM886395
Hexactinosida	*Aphrocallistes vastus*	**HE591454**	NC_010769	AM886406
Hexactinosida	*Heterochone* sp. (Sagami bay)	**FR819677**	**FR819685**	AM886404
Porifera (Calcarea)				
Calcaronea	*Grantessa* sp.	**FR819678**	–	**FR819693**
Calcaronea	*Leuconia nivea*	**HE591455**	–	**HE591470**
	Leucetta chagosensis	**FR819679**	–	**FR819694**
Calcinea	*Clathrina* sp.	**FR819680**	–	**FR819695**
Choanoflagellida				
	Monosiga brevicollis	XM_001747535	NC_004309	AF100940
	Monosiga ovata	EST data	EST data	AF271999
Fungi				
Basidiomycota	*Cryptococcus neoformans*	XM_767960	NC_004336	BR000310
Basidiomycota	*Ustilago maydis*	XM_756263	NC_008368	X62396
Ascomycota	*Neurospora crassa*	XM_958527	X01850	AY046271
Ascomycota	*Schizosaccharomyces pombe*	NM_001022703	NC_001326	X58056

genes) and on the concatenated dataset (with third codon positions). For each dataset two approaches were used: ML and Bayesian inference.

Bayesian analyses were performed using PhyloBayes v 3.2e with the site-heterogeneous CAT–GTR model (Lartillot et al., 2009). Because the mixture models available in PhyloBayes do not allow combining DNA and protein sequences, all analyses were based on DNA sequences. For each dataset (separated and combined), two independent chains were run for 154,296–628,402 cycles depending on the dataset. Convergence of the two chains was verified using the bpcomp and tracecomp programs of the PhyloBayes package. The sampling was performed every ten cycles after discarding 20% of the cycles as burnin. For each chain, the total number of trees considered was higher than 8000. As recommended by the PhyloBayes manual, we verified that for each run the parameters "maxdiff" were less than 0.1, "rel_diff" were less than 0.1 and "effsize" were higher than 100. To verify that the CAT model would better explain the evolution of the 18S rDNA gene than doublet (e.g., Savill et al., 2001) or homogeneous substitution models, following Tsagkogeorga et al. (2009), we analyzed the posterior distribution of the number of inferred profiles (k) provided by PhyloBayes (Fig. 1). For the 18S rDNA gene, the number of profiles was always higher than 20, and the average profile number was 63.6. Our results thus agree with Tsagkogeorga et al.'s (2009) conclusions that homogeneous and doublet models might not capture the extent of evolutionary constraints affecting this molecule.

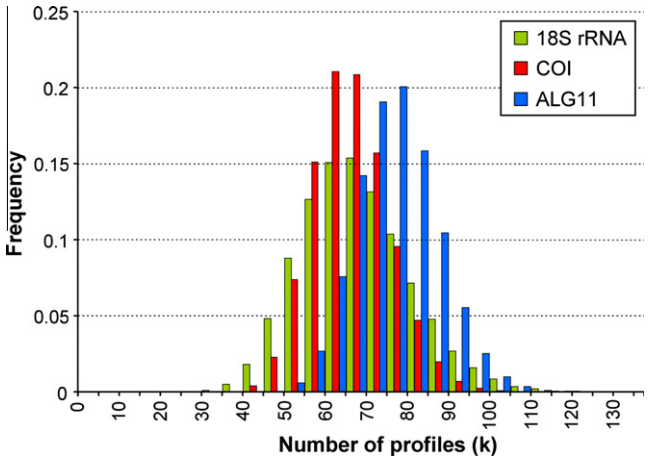

Fig. 1. Estimated number of profiles k under the CAT–GTR mixture model. Frequencies of the number of different profiles (k) estimated through the MCMC runs at the stationary stage for each dataset. The third codon positions were included in the COI and ALG11 datasets.

While the COI gene shows a distribution of number of inferred profiles similar to that of the 18S rDNA, the ALG11 gene appears to be even more heterogeneous across sites (Fig. 1). A possible explanation is that the gene contains a lower fraction of invariant positions than the 18S rDNA.

Maximum likelihood analyses of the separated and concatenated datasets were performed using RAxML 7.2.6 (Stamatakis, 2006) with the GTR + Γ4 model. Bootstrap percentages (BPs) were computed for each dataset based on 100 replicates. For the ALG11 and the COI genes each codon position was treated as a different partition with different α-shape parameters, GTR-rates, and empirical base frequencies. In the combined analysis, similarly, different parameters were assumed for the 18S rDNA partition and for each codon position of the ALG11 and COI genes.

2.5. PABA analysis

To better evaluate the performance of each marker, we conducted a partition addition bootstrap alteration (PABA) analysis (Struck et al., 2006). For nodes of interest the PABA approach computes the alteration of BP (δ) when a partition is added. In our case, three gene partitions were considered in the PABA analysis: the 18S rDNA, the ALG11 and the COI gene. To compute, for example, the δ value of the ALG11 on the demosponge monophyly, when the ALG11 gene is added as third marker, we first computed the BP of this node when the 18S rDNA and COI are combined (in this case BP = 32). We then computed the BP of demosponge monophyly in the combined analysis with the three markers 18S rDNA, COI and ALG11 (BP = 76). The alteration of bootstrap support created by the addition of ALG11 is thus $\delta = 76 - 32 = 44$. Consequently, positive δ values indicate that a partition contributes to the support of a node while a negative value indicates a conflict between partitions. To compute the average δ value of the ALG11 for the demosponge monophyly, when the ALG11 gene is added as second marker, the same computation was performed twice: once to calculate the difference in BP support values resulting from the addition of the ALG11 dataset to the 18S rDNA dataset, and once to calculate the difference in BP support values resulting from the addition of the ALG11 dataset to the COI dataset. These computations were repeated for each marker and for 12 nodes of interest. Because COI sequences of Calcarea could not be determined, Calcarea were excluded from the PABA analysis so that all datasets were compared based on the same species sampling. Phylogenetic reconstructions were conducted using the ML criterion for each marker separately and for all possible combinations of markers (i.e., ALG11 + COI, ALG11 + 18S rDNA, COI + 18S rDNA, and ALG11 + COI + 18S rDNA). Trees were reconstructed with RAxML 7.2.8 as indicated above assuming different among-site rate variation and substitution rate parameters for each partition and each codon position. The third codon positions of the ALG11 and COI gene were included in the PABA analysis.

3. Results and Discussion

3.1. The ALG11 gene – a novel variable marker

Our motivation was to determine a new phylogenetic marker that could easily be amplified in all sponge classes and would also be more variable than the 18S rDNA and the COI genes. To identify such a marker, a high throughput bioinformatics search was conducted. In this search, specific criteria were established, such as evolutionary rate, lack of introns, phylogenetic profile and lack of paralogs (see Section 2). Among the three best selected markers, only the ALG11 gene could be amplified from the four sponge classes, Hexactinellida, Demospongiae, Homoscleromopha Calcarea and from the cnidarian classes Anthozoa, Hydrozoa and Scyphozoa.

The ALG11 gene is a mannosyltransferase that belongs to the N-linked protein glycosylation pathway; a pathway that is highly conserved in eukaryotes (O'Reilly et al., 2006). In human (NG_028038) and fly (NT_037436 (Hoskins et al., 2007)), the ALG11 gene consists of four and five exons, respectively. However, in the cnidarians *Nematostella vectensis* (NW_001834198 (Putnam et al., 2007)) and *Hydra magnipapillata* (NW_002161912 (Chapman et al., 2010)), as well as in the placozoan *Trichoplax adhaerens* (NW_002060947 (Srivastava et al., 2008)) and in the sponge *A. queenslandica* (trace data (Srivastava et al., 2010)), this gene is intronless, suggesting that the gain of introns occurred after the divergence of bilaterians. The region amplified in this work corresponds to the end of exon 2 and most of exon 3 of the human gene. This region was chosen visually since its sequence variability seems to be at a level appropriate for resolving the desired level of sponge relationships and it is flanked by highly conserved regions at the protein level, which facilitate primer design (Supplementary material, Fig. S2). Finally, preliminary phylogenetic analysis of ALG11 sequences indicated that this gene is a single copy gene in animals and fungi (Supplementary material, Fig. S1).

The marker was successfully amplified from genomic DNA of all sponge and cnidarian samples for which we could amplify the 18S rDNA and the COI gene, including museum samples that had been preserved for several years in ethanol (Table 1). The region amplified was about 930 bp and none of the sequences obtained were found to include an intron. This result confirms that the ALG11 marker can be used for museum samples from which cDNA cannot be obtained.

Our second goal – to determine a variable marker – was also achieved. After removing non-reliable characters from the alignment, the ALG11 dataset comprised 882 positions, of which 104 were constant, 41 were variable uninformative and 737 were parsimony informative. By comparison, the 18S rDNA dataset, after removing ambiguously aligned regions, comprised 1591 positions, of which 834 were constant, 172 were variable uninformative and 585 were parsimony informative. Finally, the COI dataset (also after removing non-reliable characters from the alignment), comprised 1200 positions, of which 336 were constant, 98 were variable uninformative, and 766 were parsimony informative. Therefore, although the ALG11 dataset was the shortest of the three, it appears to contain the same number of variable positions as the COI and more than the 18S rDNA. The degree of variation in the ALG11 dataset was considerable when observing pairwise

differences between sequences. For example, among the 11 demosponge species sequenced, the *p*-distances ranged from 0.160 to 0.509 (140–449 substitutions, Supplementary Table S1), while for the 18S rDNA the *p*-distances ranged from 0.008 to 0.132 (12–209 substitutions, Supplementary Table S2) and for the COI they ranged from 0.047 to 0.303 (56–360 substitutions, Supplementary Table S3). Clearly, the ALG11 is more variable than the other two markers. It is worth noting that this was also true when the third codon positions were excluded. In this case the *p*-distances among demosponges ranged for the ALG11 gene from 0.078 to 0.416 (46–246 substitutions, Supplementary Table S4), while for the COI gene they ranged from 0.014 to 0.151 (11–117 substitutions, Supplementary Table S5).

To evaluate the level of saturation of the ALG11 gene, following Philippe and Forterre (Philippe and Forterre, 1999), we plotted in Fig. 2 the average number of observed nucleotide substitutions per site (i.e., the *p*-distances) as a function of the average number of inferred nucleotide substitutions per site based on the Maximum Likelihood (ML) criterion (i.e., the patristic ML distances or the sum of lengths of all branches linking two sequences on the ML tree). Using this approach, mutational saturation is inferred when *p*-distances remain constant while ML distances increase (i.e., presence of a "plateau") (Philippe and Forterre, 1999). For the ALG11 and COI genes the saturation plots were drawn either with the three codon positions (Fig. 2A and B) or excluding the third codon positions (Fig. 2C and D). The saturation plot involving all possible species pairs of our dataset, and third codon positions (Fig. 2A), shows that the ML tree distances are the largest for the COI gene and that *p*-distances above 0.3 are likely to be saturated for this marker. These results illustrates the extremely high mutation rate of the third codon positions in mitochondrial genes. By comparison, the ALG11 marker appears to be more variable and less saturated although *p*-distances above 0.5 are likely to be saturated. No saturation is noted for the 18S rDNA gene which is much less variable. When the third codon positions were removed (Fig. 2C) the tree distances were much reduced for both the ALG11 and the COI gene. The saturation plot suggests that *p*-distances above 0.4 are likely to be saturated for the ALG11 gene while *p*-distances above 0.2 are likely to be saturated for the COI gene. Again the COI gene appears to be less variable and more saturated than the ALG11 gene.

To evaluate the performance of each marker to solve sponge relationships, saturation plots were also drawn considering only distances among members of the same sponge class (Fig. 2 B and D). The ALG11 gene does not appear to be saturated, with or without the third codon positions (Fig. 2B). Consequently, third codon positions were conserved in subsequent phylogenetic analyses. The COI gene, conversely, seems less saturated, when third codon positions are excluded. However, the ML tree reconstructed based on the first and second codon positions of the COI gene supported less sponge clades than the tree reconstructed based on the three codon positions (Supplementary material, Fig. S3). Since our goal is to compare marker performances within sponge classes, all three codon positions of the COI gene were also considered in subsequent analyses.

More generally, the saturation plots confirm the high variability of the ALG11 marker when compared to the 18S rDNA and the COI gene, as the number of observed and inferred substitutions between sequence pairs is usually the highest for the ALG11 gene. This is particularly true when considering distances between two members of the same sponge class (Fig. 2B).

3.2. Separate phylogenetic analyses

Each DNA dataset described above was analyzed separately using the Bayesian inference scheme, with the CAT–GTR model of sequence evolution, and using the ML criterion with the

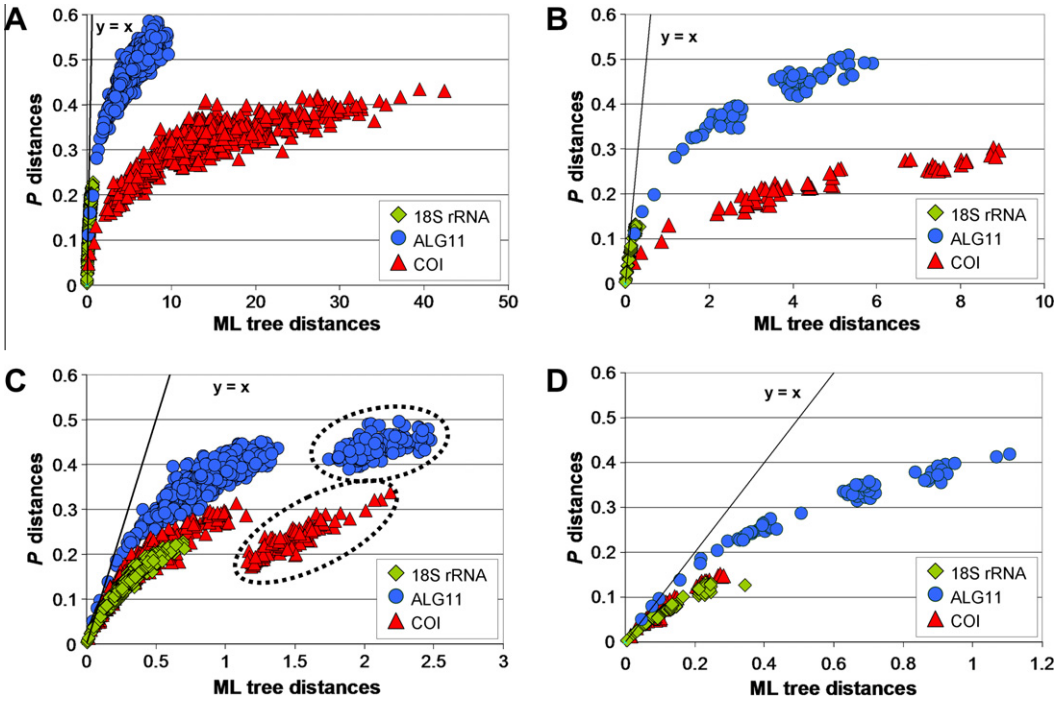

Fig. 2. Saturation plots of character sets. A, B: Saturation plots based on the three codon positions of the ALG11 and COI genes. C, D: Saturation plots based on the first two codon positions of the ALG11 and COI genes. A, C: Saturation plots between all pairs of species present in each dataset the 18S rRNA distances are the same in the two plots. B, D: Saturation plots within sponge classes (i.e., between pairs of demosponge species, between pairs of calcarea species, between pairs of homoscleromorph species, and between pairs of hexactinellid species), the 18S rRNA distances are the same in the two plots. The straight line ($y = x$) represents the situation for which there is no homoplasy in the data (i.e., the number of inferred substitutions equals the number of observed differences). The doted ellipse indicates pairwise comparisons which involve metazoan and fungi, and, for which, the ALG11 and COI genes are inferred to be saturated (plateau).

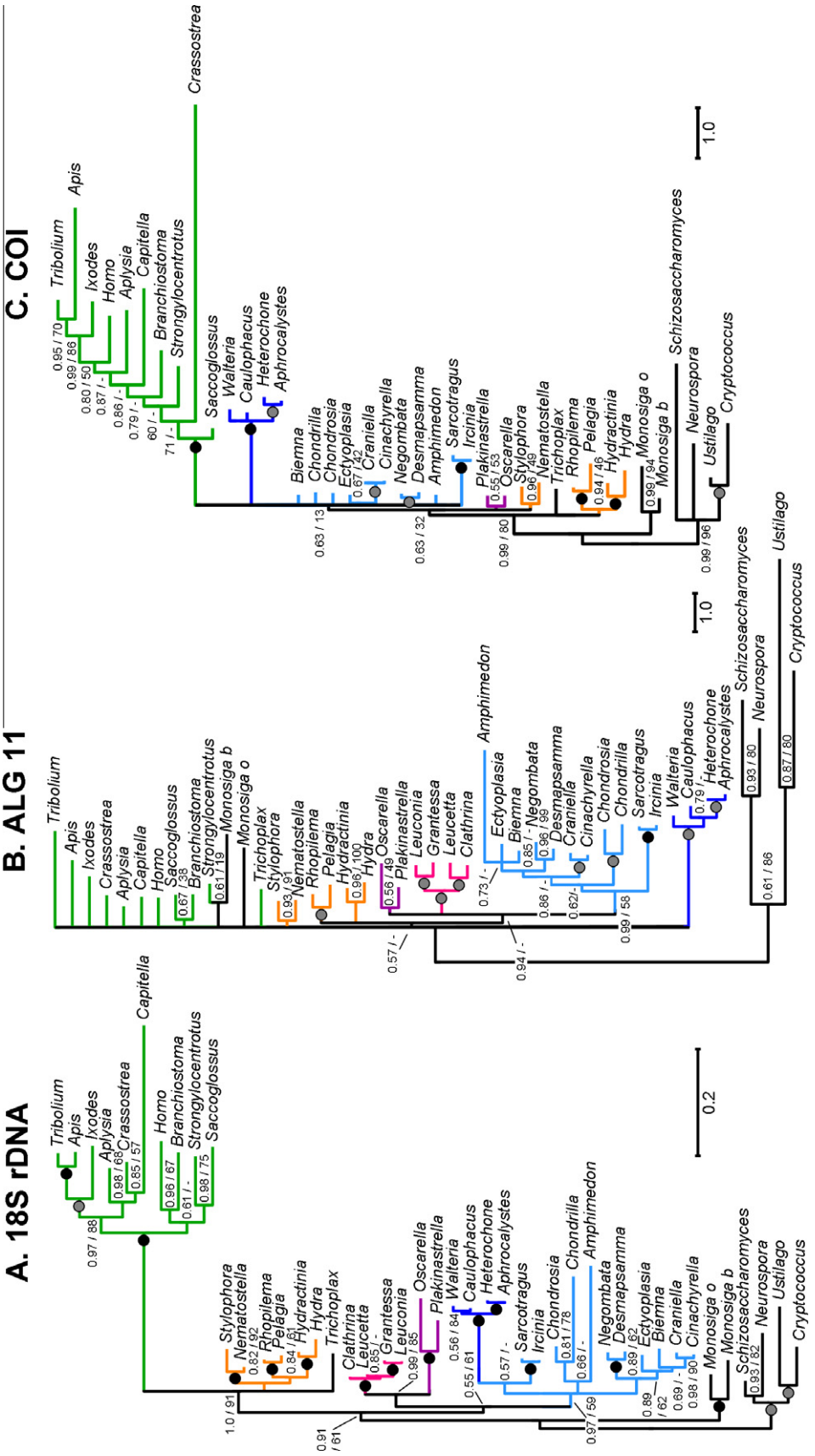

Fig. 3. Bayesian consensus trees reconstructed for each marker using the CAT–GTR model. A: Phylogenetic reconstruction based on 18S rDNA. B: Phylogenetic reconstruction based on ALG11 DNA sequences (all three codon positions). C: Phylogenetic reconstruction based on COI DNA sequences (all three codon positions). Only nodes supported by a posterior probability above 0.50 are presented. Posterior probabilities/ML bootstrap supports are given near the corresponding node. "–" indicates that the corresponding node is absent from the bootstrap consensus tree. Solid circles indicate branches with maximal support values (Bayesian PP = 1.0 and ML BP = 100); gray circles indicate branches with high support values ($0.99 \leqslant PP \leqslant 1.0$ and $98 \leqslant BP \leqslant 100$).

Table 4
Alteration of bootstrap support δ to nodes with the addition of data partitions. NA – not applicable due to alteration from 100 to 100. Negative values are indicated with a gray background. The values of the genes that contributed the most to the support of a node are indicated in bold. G1, G2 and G4 are demosponge clades as defined by Borchiellini et al. (2004).

Clades	BP in combined tree	18S rDNA		ALG11		COI	
		2nd	3rd	2nd	3rd	2nd	3rd
Bilateria	100	**50.5**	NA	0.5	NA	**50**	**NA**
Cnidaria	64	**70**	**57**	2	8	−12	−20
Placozoa + Cnidaria	53	**66.5**	**30**	22	−3	8.5	−27
Placozoa + Cnidaria + Bilateria	98	**99**	**87**	7	0	6	−2
Hexactinellida	100	NA	NA	NA	NA	NA	NA
Homoscleromorpha	100	**63**	**19**	29	NA	15	NA
Demospongiae	76	−7	1	**34.5**	**44**	−7.5	3
Demospongiae + Hexactinellida	32	**44.5**	**3**	14	−34	−4.5	−32
G1	100	NA	NA	NA	NA	NA	NA
G2	94	18	−6	**46**	**30**	−7	−5
G4	100	**74**	**47**	32	2	19	10

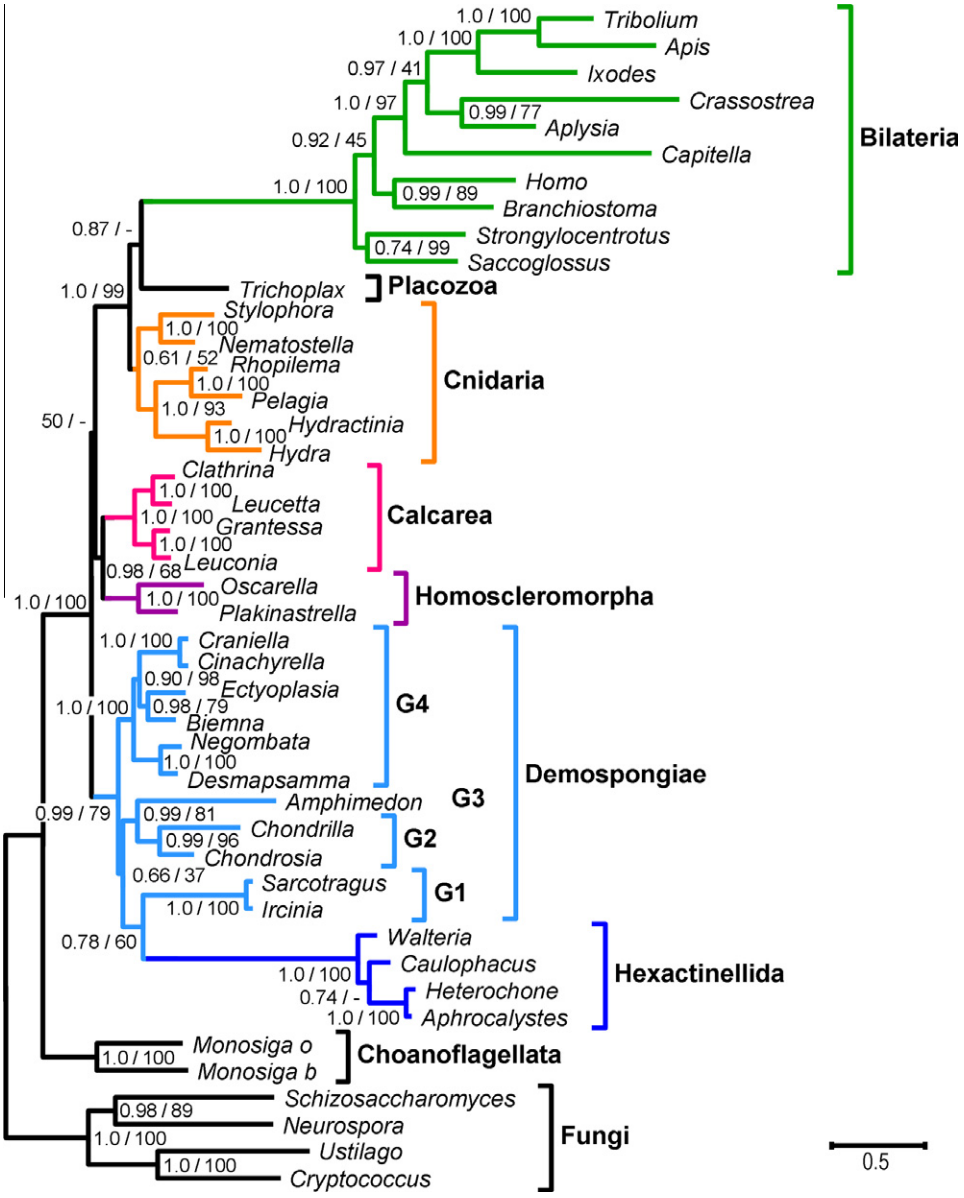

Fig. 4. Bayesian consensus tree reconstructed using the CAT–GTR model for the concatenated DNA sequences of 18S rDNA, ALG11, and COI. Posterior probabilities/ML bootstrap supports are given near the corresponding node. "–" indicates that the corresponding node is absent from the bootstrap consensus tree.

GTR + Γ4 model. The Bayesian trees are shown in Fig. 3. These trees were congruent with the ones obtained under the ML criterion. Only nodes with Bayesian posterior probability (PP) above 0.5 are shown. Similarly, given the low species sampling, relationships among Bilateria are not addressed since they have already been the subject of numerous studies (Dunn et al., 2008; Lartillot and Philippe, 2008; Philippe et al., 2005; Rokas et al., 2005).

Phylogenetic reconstructions show that, although it is the least variable, the 18S rDNA tree is the most resolved (Fig. 3A). Indeed the analysis of 18S rDNA sequences recovers the monophyly of: Bilateria (PP = 1.0; ML bootstrap percentage BP = 100), Cnidaria (PP = 0.82, BP = 92), Homoscleromorpha (PP = 1.0, BP = 100), Calcarea (PP = 1.0, BP = 100), Hexactinellida (PP = 1.0, BP = 100), Homoscleromorpha + Calcarea (PP = 0.99, BP = 85), Silicea (Hexactinellida + Demospongiae PP = 0.97, BP = 59), Metazoa (PP = 0.91, BP = 61), Choanoflagellida (PP = 1.0, BP = 100), and Fungi (PP = 1.0, BP = 99). Surprisingly, Demospongiae were not found to be monophyletic since Hexactinellida were placed as the sister clade of the Dictyoceratida species (*Sarcotragus* and *Ircinia*). However, Demosponge paraphyly is not supported (PP = 0.57, BP < 50).

The Bayesian ALG11 and COI trees are much less resolved. The ALG11 tree (Fig. 3B) only recovers the monophyly of the sponge clades Calcarea (PP = 0.99, BP = 99), Homoscleromorpha (PP = 0.56, BP = 49), Hexactinellida (PP = 0.99, BP = 100), Demospongiae (PP = 0.99, BP = 58), and Fungi (PP = 0.61, BP = 86) while the COI tree recovers the monophyly of Bilateria (PP = 1.0, BP = 100), Metazoa (PP = 0.99, BP = 80), Choanoflagellida (PP = 0.99, BP = 94), Fungi (PP = 0.99, BP = 96) and Hexactinellida (PP = 1.0, BP = 100).

The low support observed for Metazoa and Bilateria monophyly with the ALG11 marker can be explained by the fast evolutionary rate of this gene. This dataset is most probably saturated by the high number of substitutions that occurred along the long-branch leading to the outgroup and the long-branch of the fast-evolving Bilateria. Such saturation is clearly visible when third codon positions are removed (Fig. 2B). Given the short length of the marker (only 882 characters), even advanced phylogenetic methods are unable to resolve the evolutionary relationships of these clades. However, as hypothesized, this gene performed better than the COI gene in resolving relationships among slow-evolving branches such as the Demospongiae branches of this study. The good phylogenetic performance of the 18S rDNA gene at most phylogenetic levels confirms that it is indeed a marker of choice in solving animal relationships. The better performance of the 18S rDNA might be due to it being a non-coding gene and a much longer marker (1591 positions), almost twice the length of the ALG11 gene.

3.3. PABA analyses

In agreement with the phylogenetic results obtained on separate markers, the PABA analysis indicates that the 18S rDNA gene contributes the most to determining the animal phylogeny (Table 4). The 18S rDNA is the gene showing the highest alteration of bootstrap support in favor of the monophyly of Cnidaria, the grouping of Bilateria + Cnidaria + Placozoa, the monophyly of Homoscleromorpha, the grouping of Hexactinellida and Demospongiae, and the monophyly of the G4 clade *sensus* Borchiellini et al. (2004) (in our case *Biemna*, *Cinachyrella*, *Craniella*, *Desmapsamma*, *Ectyoplasia*, and *Negombata*). Our results indicate that for relationships among sponges the ALG11 gene performs as well as the 18S rDNA gene. In particular, when combined, the ALG11 marker provides the highest alteration of booststrap support in favor of the monophyly of Demospongiae, or in favor of the monophyly of clade G2 *sensus* Borchiellini et al. (2004) (in our case the Chondrosida genera *Chondrilla* and *Chondrosia*). Finally, our results indicate that the COI gene is the one that contributes the least to the branch support, and that this marker tends to conflict with other markers. For most of the nodes (e.g., Cnidaria monophyly, Demospongiae monophyly, or Demospongiae + Hexactinellida) the COI gene provides negative alterations of booststrap support. Because these negative support alterations are rather weak (most δ values are above -10), these δ values most probably indicate lack of phylogenetic signal rather than genuine incongruence. Indeed, there is no reason to assume that the evolutionary history of the COI gene differed from that of the 18S rDNA. Preliminary analyses indicated that the removal of the third codon positions in the COI and ALG11 datasets does not alter these conclusions.

Although the ALG11 tree was not fully resolved, the PABA analysis indicates that, when combined with other markers, this gene improves branch support and thus tree resolution. Consequently, the ALG11 gene appears to be a good marker when combined with the 18S rDNA. Concerning the COI and 18S rDNA performances, our results agree with Struck et al. (2006), who show that for Eunicida (Annelida) the 18S rDNA has the largest positive influence on δ values, and COI has the least influence.

3.4. Combined analysis

A phylogenetic tree was reconstructed based on the combined dataset of the three markers. The resulting tree is presented in Fig. 4. It mostly agrees with the 18S rDNA tree, as it recovers with maximal support (PP = 1.0, BP = 100) the monophyly of Bilateria, Hexactinellida, Homoscleromorpha, Calcarea, Metazoa, Choanoflagellida, and Fungi. It also recovers Silicea (Hexactinellida + Demospongiae PP = 0.99, BP = 79) as well as Homoscleromorpha + Calcarea (PP = 0.98, BP = 68).

Similar to the 18S rDNA result Demospongiae were found to be paraphyletic, with Hexatinellida placed as sister clade of Dictyoceratida species. This weird positioning is most probably an artifact since most phylogenetic analysis support the monophyly of demosponges (Pick et al., 2010; Sperling et al., 2009). Additional species and markers are thus needed to establish the relationships among the four main demosponge lineages and Hexactinellida.

4. Conclusions

Our goal was to design a novel phylogenetic marker to investigate relationships within sponge classes. We successfully amplified a new variable exonic marker from the ALG11 gene for a few representatives of each sponge class. The ALG11 was found to be more informative regarding sponge relationships than the COI gene. The PABA analysis further shows that ALG11 increases the phylogenetic signal when combined with other markers. The ALG11 gene is thus a promising marker to improve our understanding of sponge relationships. Since no introns were found in the cnidarian species considered, it is probable that this marker can also be applied to solve debated nodes within the Cnidaria.

Acknowledgments

We would like to thank Naomi Paz for editing the text, Tal Pupko for his comments, and Nir Ezra for sequencing *Desmapsamma anchorata*. This work would not have been possible without the help of Alain Daniel, John Hooper and the Queensland Museum, Dorte Janussen and the Senckenberg Museum, Sally Leys, Scott Nichols and the University of California Museum of Paleontology, Bernard Picton and the Ulster Museum, and Chagai Rot, who provided tissue samples. Revital Ben-David-Zaslow and Sigal Shefer provided invaluable help with the collections of the Zoological Museum at Tel Aviv University. We also would like to thank the Israeli Ministry of Science Culture & Sport for supporting the National Collections of Natural History at Tel-Aviv University. This

research was supported by the Israel Science Foundation (Grant No. 600/06 to D.H). D.H. was supported by the National Evolutionary Synthesis Center (NESCent), NSF #EF-0905606. G.W. acknowledges funding by the German Science Foundation (DFG) through the priority programme 1174 "Deep Metazoan Phylogeny" (Project Wo896/6, 1-4). Last but not least, we thank the two anonymous reviewers and Prof. Neil Blackstone for their valuable comments and suggestions.

Appendix A. Supplementary material

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.ympev.2012.02.008. As well as in the Dryad Digital Repository at http://datadryad.org/doi:10.5061/dryad.20617h5g.

References

Addis, J.S., Peterson, K.J., 2005. Phylogenetic relationships of freshwater sponges (Porifera, Spongillina) inferred from analyses of 18S rDNA, COI mtDNA, and ITS2 rDNA sequences. Zool. Scr. 34, 549–557.

Aguinaldo, A.M.A., Turbeville, J.M., Linford, L.S., Rivera, M.C., Garey, J.R., Raff, R.A., Lake, J.A., 1997. Evidence for a clade of nematodes, arthropods and other moulting animals. Nature 387, 489–493.

Alvarez, B., Krishnan, M., Gibb, K., 2007. Analysis of intragenomic variation of the rDNA internal transcribed spacers (ITS) in Halichondrida (Porifera: Demospongiae). J. Mar. Biol. Ass. UK 87, 1599–1605.

Belinky, F., Cohen, O., Huchon, D., 2010. Large-scale parsimony analysis of metazoan indels in protein-coding genes. Mol. Biol. Evol. 27, 441–451.

Bentlage, B., Wörheide, G., 2007. Low genetic structuring among *Pericharax heteroraphis* (Porifera: Calcarea) populations from the Great Barrier Reef (Australia), revealed by analysis of nrDNA and nuclear intron sequences. Coral Reefs 26, 807–816.

Berntson, E.A., Bayer, F.M., Mcarthur, A.G., France, S.C., 2001. Phylogenetic relationships within the Octocorallia (Cnidaria: Anthozoa) based on nuclear 18S rRNA sequences. Mar. Biol. 138, 235–246.

Borchiellini, C., Boury-Esnault, N., Vacelet, J., Le Parco, Y., 1998. Phylogenetic analysis of the Hsp70 sequences reveals the monophyly of metazoa and specific phylogenetic relationships between animals and fungi. Mol. Biol. Evol. 15, 647–655.

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.

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

Boury-Esnault, N., 2006. Systematics and evolution of Demospongiae. Can. J. Zool. 84, 205–224.

Cárdenas, P., Rapp, H.T., Schander, C., Tendal, O.S., 2010. Molecular taxonomy and phylogeny of the Geodiidae (Porifera, Demospongiae, Astrophorida) – combining phylogenetic and Linnaean classification. Zool. Scr. 39, 89–106.

Cárdenas, P., Xavier, J.R., Reveillaud, J., Schander, C., Rapp, H.T., 2011. Molecular phylogeny of the Astrophorida (Porifera, *Demospongiae*[P]) reveals an unexpected high level of spicule homoplasy. PLoS ONE 6, 18.

Carranza, S., Giribet, G., Ribera, C., Baguñà, J., Riutort, M., 1996. Evidence that two types of 18S rDNA coexist in the genome of *Dugesia* (*Schmidtea*) *mediterranea* (Platyhelminthes, Turbellaria, Tricladida). Mol. Biol. Evol. 13, 824–832.

Chapman, J.A., Kirkness, E.F., Simakov, O., Hampson, S.E., Mitros, T., Weinmaier, T., Rattei, T., Balasubramanian, P.G., Borman, J., Busam, D., Disbennett, K., Pfannkoch, C., Sumin, N., Sutton, G.G., Viswanathan, L.D., Walenz, B., Goodstein, D.M., Hellsten, U., Kawashima, T., Prochnik, S.E., Putnam, N.H., Shu, S., Blumberg, B., Dana, C.E., Gee, L., Kibler, D.F., Law, L., Lindgens, D., Martinez, D.E., Peng, J., Wigge, P.A., Bertulat, B., Guder, C., Nakamura, Y., Ozbek, S., Watanabe, H., Khalturin, K., Hemmrich, G., Franke, A., Augustin, R., Fraune, S., Hayakawa, E., Hayakawa, S., Hirose, M., Hwang, J.S., Ikeo, K., Nishimiya-Fujisawa, C., Ogura, A., Takahashi, T., Steinmetz, P.R.H., Zhang, X., Aufschnaiter, R., Eder, M.-K., Gorny, A.-K., Salvenmoser, W., Heimberg, A.M., Wheeler, B.M., Peterson, K.J., Böttger, A., Tischler, P., Wolf, A., Gojobori, T., Remington, K.A., Strausberg, R.L., Venter, J.C., Technau, U., Hobmayer, B., Bosch, T.C.G., Holstein, T.W., Fujisawa, T., Bode, H.R., David, C.N., Rokhsar, D.S., Steele, R.E., 2010. The dynamic genome of *Hydra*. Nature 464, 592–596.

Collins, A.G., 1998. Evaluating multiple alternative hypotheses for the origin of Bilateria: an analysis of 18S rRNA molecular evidence. Proc. Natl. Acad. Sci. USA 95, 15458–15463.

Dohrmann, M., Voigt, O., Erpenbeck, D., Wörheide, G., 2006. Non-monophyly of most supraspecific taxa of calcareous sponges (Porifera, Calcarea) revealed by increased taxon sampling and partitioned Bayesian analysis of ribosomal DNA. Mol. Phylogenet. Evol. 40, 830–843.

Dunn, C.W., Hejnol, A., Matus, D.Q., Pang, K., Browne, W.E., Smith, S.A., Seaver, E., Rouse, G.W., Obst, M., Edgecombe, G.D., Sørensen, M.V., Haddock, S.H., Schmidt-Rhaesa, A., Okusu, A., Kristensen, R.M., Wheeler, W.C., Martindale, M.Q., Giribet, G., 2008. Broad phylogenomic sampling improves resolution of the animal tree of life. Nature 452, 745–749.

Erpenbeck, D., Breeuwer, J.A.J., van Soest, R.W.M., 2005a. Implications from a 28S rRNA gene fragment for the phylogenetic relationships of halichondrid sponges (Porifera: Demospongiae). J. Zool. Syst. Evol. Res. 43, 93–99.

Erpenbeck, D., Breeuwer, J.A.J., van Soest, R.W.M., 2005b. Identification, characterization and phylogenetic signal of an elongation factor-1 alpha fragment in demosponges (Metazoa, Porifera, Demospongiae). Zool. Scr. 34, 437–445.

Erpenbeck, D., Duran, S., Rützler, K., Paul, V., Hooper, J.N.A., Wörheide, G., 2007a. Towards a DNA taxonomy of Caribbean demosponges: a gene tree reconstructed from partial mitochondrial CO1 gene sequences supports previous rDNA phylogenies and provides a new perspective on the systematics of Demospongiae. J. Mar. Biol. Ass. UK 87, 1563–1570.

Erpenbeck, D., Hooper, J.N.A., Wörheide, G., 2006. CO1 phylogenies in diploblasts and the "Barcoding of Life" – are we sequencing a suboptimal partition? Mol. Ecol. Notes 6, 550–553.

Erpenbeck, D., List-Armitage, S., Alvarez, B., Degnan, B.M., Wörheide, G., Hooper, J.N.A., 2007b. The systematics of Raspailiidae (Demospongiae: Poecilosclerida: Microcionina) re-analysed with a ribosomal marker. J. Mar. Biol. Ass. UK 87, 1571–1576.

Erpenbeck, D., Voigt, O., Gültas, M., Wörheide, G., 2008. The sponge genetree server – providing a phylogenetic backbone for poriferan evolutionary studies. Zootaxa 1939, 58–60.

Felsenstein, J., 1978. Cases in which parsimony or compatibility methods will be positively misleading. Syst. Zool. 27, 401–410.

Flicek, P., Amode, M.R., Barrell, D., Beal, K., Brent, S., Chen, Y., Clapham, P., Coates, G., Fairley, S., Fitzgerald, S., Gordon, L., Hendrix, M., Hourlier, T., Johnson, N., Kähäri, A., Keefe, D., Keenan, S., Kinsella, R., Kokocinski, F., Kulesha, E., Larsson, P., Longden, I., McLaren, W., Overduin, B., Pritchard, B., Riat, H.S., Rios, D., Ritchie, G.R.S., Ruffier, M., Schuster, M., Sobral, D., Spudich, G., Tang, Y.A., Trevanion, S., Vandrovcova, J., Vilella, A.J., White, S., Wilder, S.P., Zadissa, A., Zamora, J., Aken, B.L., Birney, E., Cunningham, F., Dunham, I., Durbin, R., Fernández-Suarez, X.M., Herrero, J., Hubbard, T.J.P., Parker, A., Proctor, G., Vogel, J., Searle, S.M., 2011. Ensembl 2011. Nucleic Acids Res. 39, D800–D806.

Folmer, O., Black, M., Hoeh, W., Lutz, R., Vrijenhoek, R., 1994. DNA primers for amplification of mitochondrial cytochrome *c* oxidase subunit I from diverse metazoan invertebrates. Mol. Mar. Biol. Biotechnol. 3, 294–297.

Fujita, P.A., Rhead, B., Zweig, A.S., Hinrichs, A.S., Karolchik, D., Cline, M.S., Goldman, M., Barber, G.P., Clawson, H., Coelho, A., Diekhans, M., Dreszer, T.R., Giardine, B.M., Harte, R.A., Hillman-Jackson, J., Hsu, F., Kirkup, V., Kuhn, R.M., Learned, K., Li, C.H., Meyer, L.R., Pohl, A., Raney, B.J., Rosenbloom, K.R., Smith, K.E., Haussler, D., Kent, W.J., 2011. The UCSC genome browser database: update 2011. Nucleic Acids Res. 39, D876–D882.

Fulton, T.M., Chunwongse, J., Tanksley, S.D., 1995. Microprep protocol for extraction of DNA from tomato and other herbaceous plants. Plant Mol. Biol. Rep. 13, 207–209.

Haen, K.M., Lang, B.F., Pomponi, S.A., Lavrov, D.V., 2007. Glass sponges and bilaterian animals share derived mitochondrial genomic features: a common ancestry or parallel evolution? Mol. Biol. Evol. 24, 1518–1527.

Hall, B.G., 2004. Phylogenetic Trees Made Easy: A How-to Manual. Sinauer Associates, Inc., Sunderland.

Hebert, P.D.N., Ratnasingham, S., deWaard, J.R., 2003. Barcoding animal life: cytochrome *c* oxidase subunit 1 divergences among closely related species. Proc. Roy. Soc. Lond. B 270 (Suppl.), S96–S99.

Hebert, P.D.N., Stoeckle, M.Y., Zemlak, T.S., Francis, C.M., 2004. Identification of birds through DNA barcodes. PLoS Biol. 2, 1657–1663.

Hooper, J.N.A., Van Soest, R.W.M., 2002. Systema Porifera. A Guide to the Classification of Sponges. In: Hooper, J.N.A., Van Soest, R.W.M. (Eds.), Systema Porifera: A guide to the classification of sponges. Kluwer Academic/Plenum Publishers, New York, pp. 1–3.

Hoskins, R.A., Carlson, J.W., Kennedy, C., Acevedo, D., Evans-Holm, M., Frise, E., Wan, K.H., Park, S., Mendez-Lago, M., Rossi, F., Villasante, A., Dimitri, P., Karpen, G.H., Celniker, S.E., 2007. Sequence finishing and mapping of *Drosophila melanogaster* heterochromatin. Science 316, 1625–1628.

Huang, D.W., Meier, R., Todd, P.A., Chou, L.M., 2008. Slow mitochondrial *COI* sequence evolution at the base of the metazoan tree and its implications for DNA barcoding. J. Mol. Evol. 66, 167–174.

Huchon, D., Madsen, O., Sibbald, M.J.J.B., Ament, K., Stanhope, M., Catzeflis, F., de Jong, W.W., Douzery, E.J.P., 2002. Rodent phylogeny and a timescale for the evolution of glires: evidence from an extensive taxon sampling using three nuclear genes. Mol. Biol. Evol. 19, 1053–1065.

Katoh, K., Kuma, K., Toh, H., Miyata, T., 2005. MAFFT version 5: improvement in accuracy of multiple sequence alignment. Nucleic Acids Res. 33, 511–518.

Krieger, J., Fuerst, P.A., 2002. Evidence of multiple alleles of the nuclear 18S ribosomal RNA gene in sturgeon (Family: Acipenseridae). J. Appl. Ichthyol. 18, 290–297.

Lartillot, N., Lepage, T., Blanquart, S., 2009. PhyloBayes 3: a Bayesian software package for phylogenetic reconstruction and molecular dating. Bioinformatics 25, 2286–2288.

Lartillot, N., Philippe, H., 2008. Improvement of molecular phylogenetic inference and the phylogeny of Bilateria. Phil. Trans. Roy. Soc. B 363, 1463–1472.

Martin, A.P., Burg, T.M., 2002. Perils of paralogy: using HSP70 genes for inferring organismal phylogenies. Syst. Biol. 51, 570–587.

McCormack, G.P., Erpenbeck, D., van Soest, R.W.M., 2002. Major discrepancy between phylogenetic hypotheses based on molecular and morphological criteria within the order Haplosclerida (Phylum Porifera: Class Demospongiae). J. Zool. Syst. Evol. Res. 40, 237–240.

Medina, M., Collins, A.G., Silberman, J.D., Sogin, M.L., 2001. Evaluating hypotheses of basal animal phylogeny using complete sequences of large and small subunit rRNA. Proc. Natl. Acad. Sci. USA 98, 9707–9712.

Mitchell, A., Mitter, C., Regier, J.C., 2000. More taxa or more characters revisited: combining data from nuclear protein-encoding genes for phylogenetic analyses of Noctuoidea (Insecta: Lepidoptera). Syst. Biol. 49, 202–224.

Nichols, S.A., 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.

O'Reilly, M.K., Zhang, G., Imperiali, B., 2006. In vitro evidence for the dual function of Alg2 and Alg11: essential mannosyltransferases in N-linked glycoprotein biosynthesis. Biochemistry 45, 9593–9603.

Papillon, D., Perez, Y., Caubit, X., Le Parco, Y., 2006. Systematics of Chaetognatha under the light of molecular data, using duplicated ribosomal 18S DNA sequences. Mol. Phylogenet. Evol. 38, 621–634.

Penn, O., Privman, E., Ashkenazy, H., Landan, G., Graur, D., Pupko, T., 2010. GUIDANCE: a web server for assessing alignment confidence scores. Nucleic Acids Res. 38, W23–W28.

Peterson, K.J., Eernisse, D.J., 2001. Animal phylogeny and the ancestry of bilaterians: inferences from morphology and 18S rDNA gene sequences. Evol. Dev. 3, 170–205.

Philippe, H., Derelle, R., Lopez, P., Pick, K., Borchiellini, C., Boury-Esnault, N., Vacelet, J., Renard, E., Houliston, E., Quéinnec, E., Da Silva, C., Wincker, P., Le Guyader, H., Leys, S., Jackson, D.J., Schreiber, F., Erpenbeck, D., Morgenstern, B., Wörheide, G., Manuel, M., 2009. Phylogenomics revives traditional views on deep animal relationships. Curr. Biol. 19, 706–712.

Philippe, H., Forterre, P., 1999. The rooting of the universal tree of life is not reliable. J. Mol. Evol. 49, 509–523.

Philippe, H., Lartillot, N., Brinkmann, H., 2005. Multigene analyses of bilaterian animals corroborate the monophyly of Ecdysozoa, Lophotrochozoa, and Protostomia. Mol. Biol. Evol. 22, 1246–1253.

Pick, K.S., Philippe, H., Schreiber, F., Erpenbeck, D., Jackson, D.J., Wrede, P., Wiens, M., Alié, A., Morgenstern, B., Manuel, M., Wörheide, G., 2010. Improved phylogenomic taxon sampling noticeably affects nonbilaterian relationships. Mol. Biol. Evol. 27, 1983–1987.

Pöppe, J., Sutcliffe, P., Hooper, J.N.A., Wörheide, G., Erpenbeck, D., 2010. CO I barcoding reveals new clades and radiation patterns of Indo-Pacific sponges of the family Irciniidae (Demospongiae: Dictyoceratida). PLoS ONE 5, e9950.

Putnam, N.H., Srivastava, M., Hellsten, U., Dirks, B., Chapman, J., Salamov, A., Terry, A., Shapiro, H., Lindquist, E., Kapitonov, V.V., Jurka, J., Genikhovich, G., Grigoriev, I.V., Lucas, S.M., Steele, R.E., Finnerty, J.R., Technau, U., Martindale, M.Q., Rokhsar, D.S., 2007. Sea anemone genome reveals ancestral eumetazoan gene repertoire and genomic organization. Science 317, 86–94.

Ranwez, V., Delsuc, F., Ranwez, S., Belkhir, K., Tilak, M.K., Douzery, E.J.P., 2007. OrthoMaM: a database of orthologous genomic markers for placental mammal phylogenetics. BMC Evol. Biol. 7, 241.

Redmond, N.E., McCormack, G.P., 2009. Ribosomal internal transcribed spacer regions are not suitable for intra- or inter-specific phylogeny reconstruction in haplosclerid sponges (Porifera: Demospongiae). J. Mar. Biol. Ass. UK 89, 1251–1256.

Rokas, A., Krüger, D., Carroll, S.B., 2005. Animal evolution and the molecular signature of radiations compressed in time. Science 310, 1933–1938.

Rot, C., Goldfarb, I., Ilan, M., Huchon, D., 2006. Putative cross-kingdom horizontal gene transfer in sponge (Porifera) mitochondria. BMC Evol. Biol. 6, 71.

Russo, C.A.M., Takezaki, N., Nei, M., 1996. Efficiencies of different genes and different tree-building methods in recovering a known vertebrate phylogeny. Mol. Biol. Evol. 13, 525–536.

Savill, N.J., Hoyle, D.C., Higgs, P.G., 2001. RNA sequence evolution with secondary structure constraints: comparison of substitution rate models using maximum-likelihood methods. Genetics 157, 399–411.

Sayers, E.W., Barrett, T., Benson, D.A., Bolton, E., Bryant, S.H., Canese, K., Chetvernin, V., Church, D.M., DiCuccio, M., Federhen, S., Feolo, M., Fingerman, I.M., Geer, L.Y., Helmberg, W., Kapustin, Y., Landsman, D., Lipman, D.J., Lu, Z., Madden, T.L., Madej, T., Maglott, D.R., Marchler-Bauer, A., Miller, V., Mizrachi, I., Ostell, J., Panchenko, A., Phan, L., Pruitt, K.D., Schuler, G.D., Sequeira, E., Sherry, S.T., Shumway, M., Sirotkin, K., Slotta, D., Souvorov, A., Starchenko, G., Tatusova, T.A., Wagner, L., Wang, Y., Wilbur, W.J., Yaschenko, E., Ye, J., 2011. Database resources of the national center for biotechnology information. Nucleic Acids Res. 39, D38–D51.

Schierwater, B., Eitel, M., Jakob, W., Osigus, H.J., Hadrys, H., Dellaporta, S.L., Kolokotronis, S.O., DeSalle, R., 2009. Concatenated analysis sheds light on early metazoan evolution and fuels a modern "urmetazoon" hypothesis. PLoS Biol. 7, e1000020.

Solé-Cava, A.M., Wörheide, G., 2007. The perils and merits (or the good, the bad and the ugly) of DNA barcoding of sponges – a controversial discussion. In: Custodio, M.R., Lobo-Hajdu, G., Hajdu, E., Muricy, G. (Eds.), Porifera Research: Biodiversity, Innovation and Sustainability, Série Livros 28. Museu Nacional, Rio de Janeiro, pp. 603–612.

Sperling, E.A., Peterson, K.J., Pisani, D., 2009. Phylogenetic-signal dissection of nuclear housekeeping genes supports the paraphyly of sponges and the monophyly of Eumetazoa. Mol. Biol. Evol. 26, 2261–2274.

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

Srivastava, M., Begovic, E., Chapman, J., Putnam, N.H., Hellsten, U., Kawashima, T., Kuo, A., Mitros, T., Salamov, A., Carpenter, M.L., Signorovitch, A.Y., Moreno, M.A., Kamm, K., Grimwood, J., Schmutz, J., Shapiro, H., Grigoriev, I.V., Buss, L.W., Schierwater, B., Dellaporta, S.L., Rokhsar, D.S., 2008. The *Trichoplax* genome and the nature of placozoans. Nature 454, 955–960.

Srivastava, M., Simakov, O., Chapman, J., Fahey, B., Gauthier, M.E.A., Mitros, T., Richards, G.S., Conaco, C., Dacre, M., Hellsten, U., Larroux, C., Putnam, N.H., Stanke, M., Adamska, M., Darling, A., Degnan, S.M., Oakley, T.H., Plachetzki, D.C., Zhai, Y.F., Adamski, M., Calcino, A., Cummins, S.F., Goodstein, D.M., Harris, C., Jackson, D.J., Leys, S.P., Shu, S.Q., Woodcroft, B.J., Vervoort, M., Kosik, K.S., Manning, G., Degnan, B.M., Rokhsar, D.S., 2010. The *Amphimedon queenslandica* genome and the evolution of animal complexity. Nature 466, 720–727.

Stamatakis, A., 2006. RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. Bioinformatics 22, 2688–2690.

Struck, T.H., Purschke, G., Halanych, K.M., 2006. Phylogeny of Eunicida (Annelida) and exploring data congruence using a partition addition bootstrap alteration (PABA) approach. Syst. Biol. 55, 1–20.

Sullivan, J.C., Ryan, J.F., Watson, J.A., Webb, J., Mullikin, J.C., Rokhsar, D., Finnerty, J.R., 2006. *StellaBase*: the *Nematostella vectensis* genomics database. Nucleic Acids Res. 34, D495–D499.

Suyama, M., Torrents, D., Bork, P., 2006. PAL2NAL: robust conversion of protein sequence alignments into the corresponding codon alignments. Nucleic Acids Res. 34, W609–W612.

Swofford, D.L., 2000. PAUP*. Phylogenetic Analysis Using Parsimony (*and Other Methods). Version 4b10. Sinauer Associates, Sunderland, Massachusetts.

Szitenberg, A., Rot, C., Ilan, M., Huchon, D., 2010. Diversity of sponge mitochondrial introns revealed by *cox* 1 sequences of Tetillidae. BMC Evol. Biol. 10, 288.

Tsagkogeorga, G., Turon, X., Hopcroft, R.R., Tilak, M.K., Feldstein, T., Shenkar, N., Loya, Y., Huchon, D., Douzery, E.J.P., Delsuc, F., 2009. An updated 18S rRNA phylogeny of tunicates based on mixture and secondary structure models. BMC Evol. Biol. 9, 187.

Voigt, O., Erpenbeck, D., Wörheide, G., 2008. Molecular evolution of rDNA in early diverging Metazoa: first comparative analysis and phylogenetic application of complete SSU rRNA secondary structures in Porifera. BMC Evol. Biol. 8, 69.

Wörheide, G., Epp, L.S., Macis, L., 2008. Deep genetic divergences among Indo-Pacific populations of the coral reef sponge *Leucetta chagosensis* (Leucettidae): founder effects, vicariance, or both? BMC Evol. Biol. 8, 24.

Wörheide, G., Erpenbeck, D., 2007. DNA taxonomy of sponges – progress and perspectives. J. Mar. Biol. Ass. UK 87, 1629–1633.

Wörheide, G., Nichols, S.A., Goldberg, J., 2004. Intragenomic variation of the rDNA internal transcribed spacers in sponges (Phylum Porifera): implications for phylogenetic studies. Mol. Phylogenet. Evol. 33, 816–830.

Zrzavý, J., Mihulka, S., Kepka, P., Bezdek, A., Tietz, D., 1998. Phylogeny of the Metazoa based on morphological and 18S ribosomal DNA evidence. Cladistics 14, 249–285.