Phylogenetic relationships of hydrothermal vent mussels (Bathymodiolinae) and their symbionts

The following supplement accompanies the article

Phylogenetic relationships of hydrothermal vent mussels (Bathymodiolinae) and their symbionts

Kristina M. Fontanez[1,2], Colleen M. Cavanaugh[1,*]

[1]**Department of Organismic and Evolutionary Biology, Harvard University, Cambridge, Massachusetts 02138, USA**
[2]***Present address:*** **Department of Civil and Environmental Engineering, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA**

*Corresponding author. Email: cavanaug@fas.harvard.edu

Marine Ecology Progress Series 474: 147–154 (2013)

Supplement. Supplementary material contains additional detail about phylogenetic analysis methods and supplementary tables and figures.

Prior to Bayesian analyses of symbiont and host loci, MrModeltest (Nylander 2004) was used to test a set of nested models of sequence evolution that are restrictions of the general time-reversible model with rate variation among sites (Yang 1994). The best-fit model of sequence evolution for each gene was evaluated based on the Akaike information criterion using standard procedures in PAUP v4.0b10 (Posada & Crandall 2001, Swofford 2003) and then chosen according to Akaike weights (Akaike 1974) (Table S2). Diffuse priors were used for all analyses. Priors for the following parameters were used: topology (all topologies equally weighted), branch lengths (exponential (10) prior), instantaneous rate matrices (Dirichlet 1,1,1,1,1,1 prior), and equilibrium base frequencies (Dirichlet 1,1,1,1 prior).

For models with gamma distributed rates and/or a proportion of invariant sites, a uniform (0.05,50) prior and a uniform (0,1) prior, respectively, were assumed. All Markov chain Monte Carlo (MCMC) analyses in MrBayes were conducted using Metropolis coupling with 20 or more parallel chains. Swap rates between adjacent chains were considered adequate if they exceeded 20%. For each analysis, at least 5 independent MCMC runs that were each iterated at least 5.0×10^7 times were conducted until MCMC convergence was reached. A burn-in of at least 500 samples was removed, and the remaining samples from all independent runs were combined to construct majority rule consensus phylograms. MCMC convergence was assessed using the CODA package in R (R Development Core Team 2007). As a further test of convergence, the medians of the posterior distributions of model parameters for each locus were compared to maximum likelihood estimates computed using PAUP to ensure their close agreement (Table S2).

Table S1. Overview of the genetic, location, and ultrastructural evidence used to reclassify the Juan de Fuca (JdF) bathymodioline specimens examined in this study. COI: Cytochrome *c* oxidase; ND4: NADH dehydrogenase subunit 4; TEM: Transmission electron microscopy images showed bacteriocytes with coccoid bacteria that appeared to be extracellular; –: not determined

				Bathymodioline				Symbiont		
Specimen	**JdF vent segment**	**Lat/Long**	**Depth (m)**	**Morph-ology**	**Genes**			**TEM**	**16S rRNA**	**Reference**
					COI	**18S**	**ND4**			
JdF2008	Endeavor	47° 58'N, 129° 05' W	2189	–	✓	✓	✓	✓	✓	This study
JdF1999 (*Bathymodiolus* sp.)	Endeavor	47° 56' N, 129° 06 W	2200	–	✓	–	–	✓	✓	(McKiness et al. 2005)
JdF1990 (*Adipicola* MV)	Middle Valley	48° 27.5' N, 128° 42.5' W	2420	✓	✓	–	–	✓	–	(Juniper et al. 1992, Southward 2008)

Table S2. Primers for host and symbiont genes examined in this study. Polymerase Chain Reaction (PCR) conditions: initial denaturation for 5 min at 98°C; 30 cycles of 10 s at 98°C, 30 s at annealing temperature, 1 min at 72°C; extension for 10 min at 72°C. PCR mixture: 1X HF buffer (Finnzymes), 0.2 mM dNTPs, 0.5 mM each primer, 1 unit of Phusion DNA polymerase (Finnzymes). Cytochrome *c* oxidase (COI) primers based on an alignment of COI genes from *Benthomodiolus lignicola* (AY275545), *Bathymodiolus* sp. (DQ077892), and *Benthomodiolus geikotsucola* (AB257513). *18S* primers are overlapping across the gene

Gene ID	Gene	Forward primer (5' 3')	Reverse primer (5' 3')	Size (bp)	Annealing temp.	Reference
Host						
COI	Mitochondrial cytochrome oxidase *c* subunit I	COIdegF – 5'-GTT GGC ACG KCC AGG WAG AT-3'	COIdegR – 5'-ATA GTR ATM CCT CCA GCT ARW-3'	~ 500	58	This study
ND4	Mitochondrial NADH dehydrogenase subunit 4	Arg BL – 5'-CAA GAC CCT TGA' ITT CGG CTC A-3'	NAP 2H – 5'-TGG AGC TTC TAC GTG RGC TTT-3'	~1200	52	(Arevalo et al. 1994, Bielawski & Gold 1996)
18S	Small subunit rRNA	1F – 5'-TAC CTG GTT GAT CCT GCC AGT AG-3'	3R – 5'-AGG CTC CCT CTC CGG AAT CGA AC-3'	~450	49	(Giribet et al. 1996)
		3F – 5'-GTT CGA TTC CGG AGA GGG A-3'	18Sbi – 5'-GAG TCT CGT TCG TTA TCG GA-3'	~1000	49	(Giribet et al. 1996); (Giribet et al. 1999)
		18S2 – 5'-ATG GTT GCA AAA GCT GAA A-3'	9R – 5'-GAT CCT TCC GCA GGT TCA CCT AC-3'	~800	49	(Giribet et al. 1999); (Giribet et al. 1996)
Symbiont						
16S	Small subunit rRNA	*27*F – 5'-AGA GTT TGA TCM TGG CTC AG-3'	*1492*R – 5'-TAC GGY TAC CTT GTT ACG ACT T-3'	~ 1500	50	(Weisburg et al. 1991)

Table S3. Best-fit models and phylogenetic model parameters for host and symbiont gene analyses. p_i = equilibrium frequency of nucleotide base *i*; *k* = transition/transversion (ts/tv) rate ratio; *R(i)* = general time-reversible rate matrix parameterized as in PAUP v.4.0; *I* = proportion of invariant sites; *a* = gamma (G) shape parameter; all Bayesian estimates are the median of the marginal posterior distribution. Symbiont gene analyses utilizing multiple partitions included a rate multiplier to allow rate heterogeneity between partitions. AIC: Akaike information criterion, a model comparison criterion; lower scores indicate better models

Locus		AIC	Base frequencies				Rate matrix					Rate variation	
			p_A	p_C	p_G	p_T	*R(a)*	*R(b)*	*R(c)*	*R(d)*	*R(e)*	*a*	*I*
Host													
COI	ML	11960.39	0.27	0.10	0.19	0.43	1.02	9.36	0.67	3.50	31.02	0.32	0.26
(GTR + I + G)	Bayes		0.27	0.10	0.19	0.44	0.97	7.41	0.84	3.50	17.62	0.50	0.23
ND4	ML	15235.30	*0.23*	*0.14*	*0.26*	*0.37*	*1.20*	5.37	0.90	1.45	6.63		0.17
(GTR + I)	Bayes		0.22	0.14	0.26	0.38	1.18	5.20	0.94	1.46	6.31		0.18
18S	ML	9638.67	0.24	0.23	0.28	0.25	1.15	2.55	1.24	0.53	3.27	0.69	0.67
(GTR+I+G)	Bayes		0.24	0.24	0.28	0.25	1.11	2.67	1.24	0.40	3.09	0.09	0.65
Symbiont							*k*						
16S	ML	10802.19	0.26	0.21	0.27	0.25	1.77						
(HKY)	Bayes		0.42	0.14	0.23	0.21	0.65						

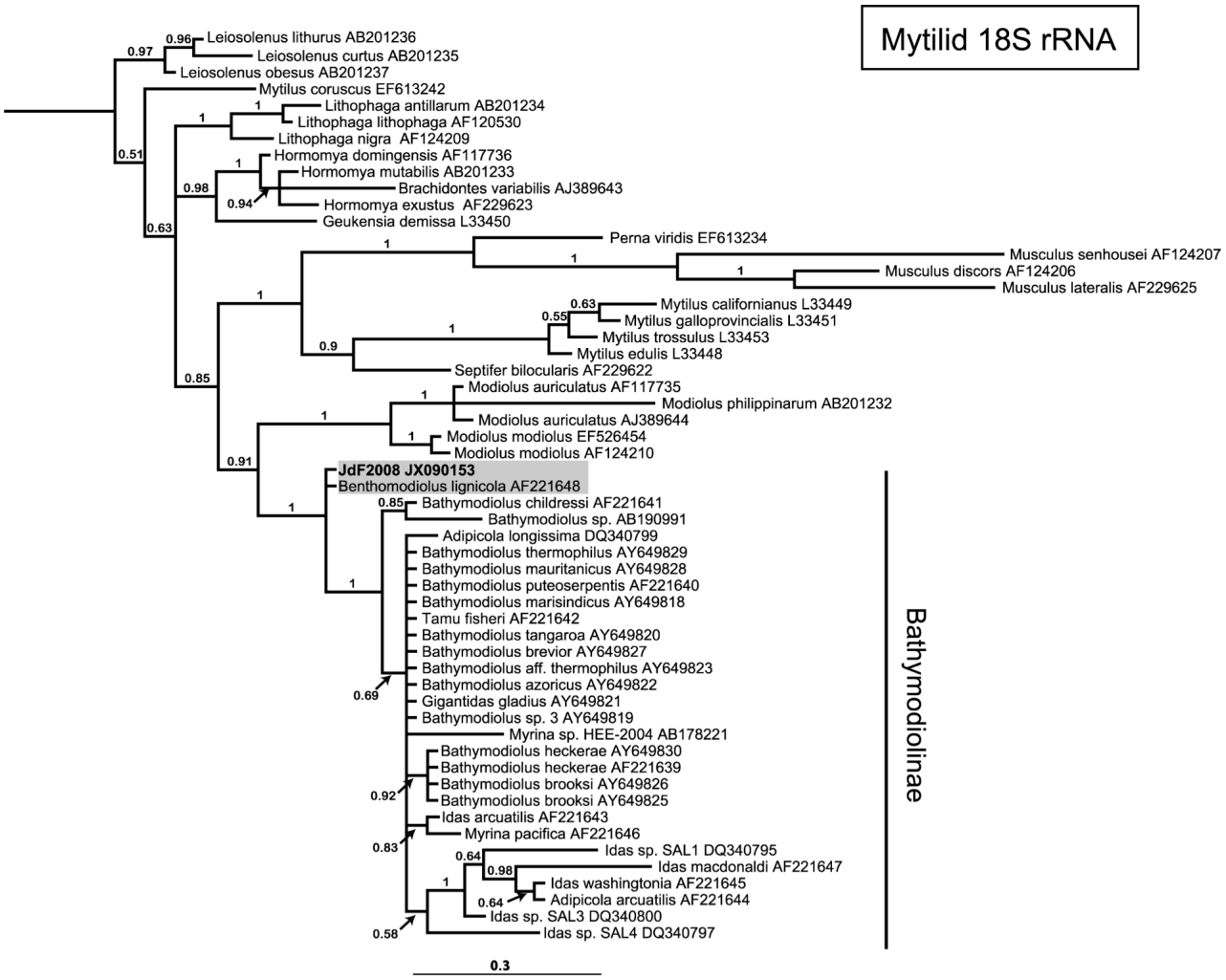

Fig. S1. Bayesian phylogeny of mytilids based on the 18S rRNA gene (1604 bp). JdF2008 (bold, this study) and *Benthomodiolus lignicola* fall basal to all other bathymodiolines. The outgroup *Crassostrea gigas* (AM12263) was removed for illustrative purposes

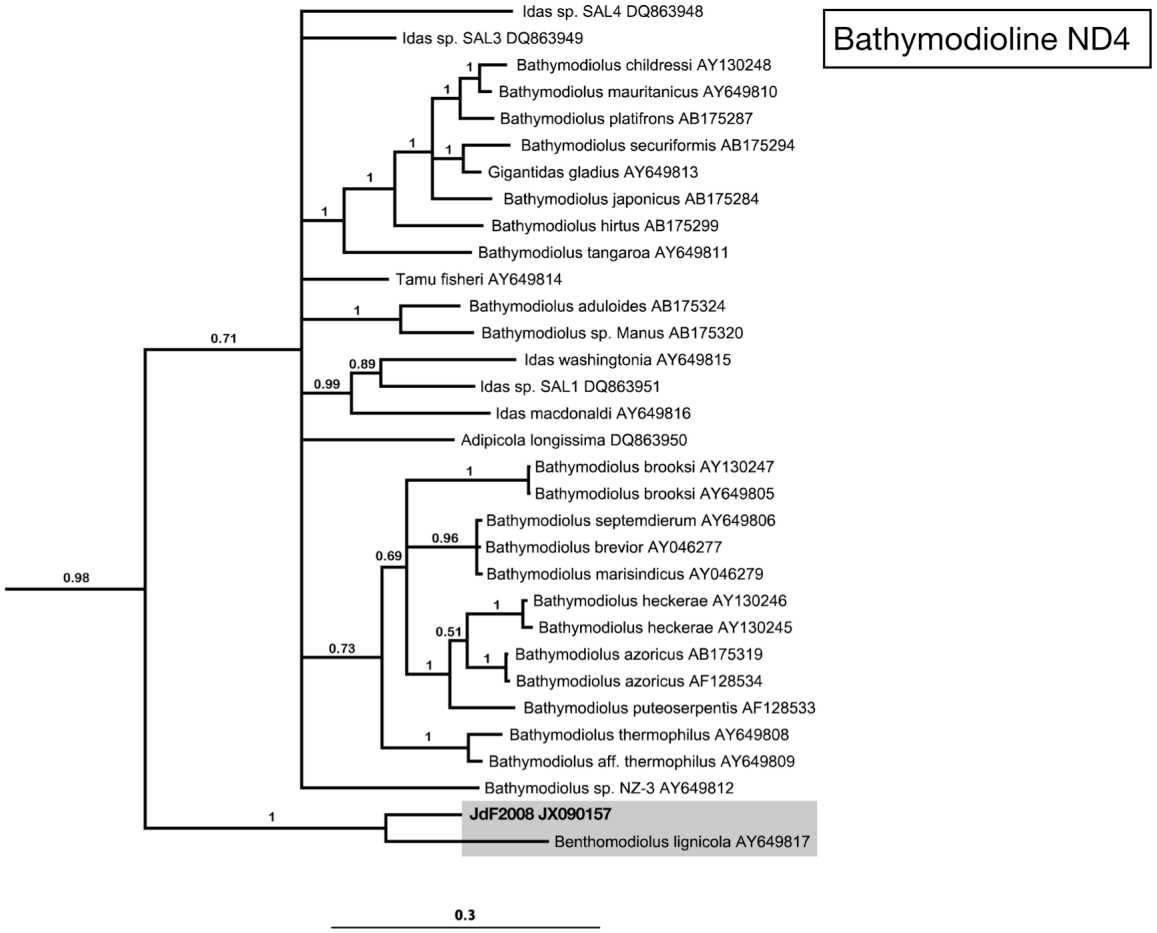

Fig. S2. Bayesian phylogeny of bathymodiolines based on NADH dehydrogenase subunit 4 gene (441 bp). JdF2008 (bold, this study) and *Benthomodiolus lignicola* fall basal to all other bathymodiolines and cluster together with strong support (Pp = 1). Outgroup taxa (*Mytilus edulis, Mytilus galloprovincialis, Mytilus trossulus, Crassostrea gigas*, and *Modiolus modiolus*) were removed for illustrative purposes

LITERATURE CITED

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

Arevalo E, Davis SK, Sites JW (1994) Mitochondrial-DNA sequence divergence and phylogenetic relationships among eight chromosome races of the Sceloporus-Grammicus complex (Phrynosomatidae) in central Mexico. Syst Biol 43:387–418

Bielawski JP, Gold JR (1996) Unequal synonymous substitution rates within and between two protein-coding mitochondrial genes. Mol Biol Evol 13:889–892

Giribet G, Carranza S, Baguna J, Riutort M, Ribera C (1996) First molecular evidence for the existence of a Tardigrada plus arthropoda clade. Mol Biol Evol 13:76–84

Giribet G, Rambla M, Carranza S, Baguna J, Riutort M, Ribera C (1999) Phylogeny of the arachnid order Opiliones (Arthropoda) inferred from a combined approach of complete 18S and partial 28S ribosomal DNA sequences and morphology. Mol Phylogenet Evol 11:296–307

Juniper SK, Tunnicliffe V, Southward EC (1992) Hydrothermal vents in turbidite sediments on a Northeast Pacific spreading centre: organisms and substratum at an Ocean Drilling Site. Can J Zool 70:1792–1809

McKiness ZP, McMullin ER, Fisher CR, Cavanaugh CM (2005) A new bathymodioline mussel symbiosis at the Juan de Fuca hydrothermal vents. Mar Biol 148:109–116

Nylander JA (2004) MrModeltest v2. Evolutionary Biology Centre, Uppsala University

Posada D, Crandall KA (2001) Selecting the best-fit model of nucleotide substitution. Syst Biol 50:580–601

R Development Core Team (2007) In: R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna

Southward EC (2008) The morphology of bacterial symbioses in the gills of mussels of the genera *Adipicola* and *Idas* (Bivalvia: Mytilidae). J Shellfish Res 27:139–146

Swofford DL (2003) PAUP: phylogenetic analysis using parsimony (and other methods). Sinauer Associates, Sunderland, MA

Weisburg WG, Barns SM, Pelletier DA, Lane DJ (1991) 16S ribosomal DNA amplification for phylogenetic study. J Bacteriol 173:697–703

Yang Z (1994) Maximum likelihood phylogenetic estimation from DNA sequences with variable rates over sites: approximate methods. J Mol Evol 39:306–314