Off-axis symbiosis found: characterization and biogeography of bacterial symbionts of Bathymodiolus mussels from Lost City hydrothermal vents

Off-axis Symbiosis Found: Characterization and Biogeography of Bacterial Symbionts of *Bathymodiolus* Mussels from Lost City Hydrothermal Vents

DeChaine[1], E. G., A. E. Bates[2], T. M. Shank[3], & C. M. Cavanaugh[1*]

[1]Department of Organismic and Evolutionary Biology, Harvard University, 16 Divinity Ave., Biolabs 4081, Cambridge, Massachusetts, 02138, USA

[2]Department of Biology, University of Victoria, Victoria, V8W2Y2, CA

[3]Biology Department, Woods Hole Oceanographic Institute, Woods Hole, Massachusetts, 02543, USA

*Corresponding author. Tel: (617) 495-2177; Fax: (617) 496-6933; E-Mail: cavanaug@fas.harvard.edu

Keywords: bacterial biogeography, *Bathymodiolus* mussels, chemoautotroph, deep-sea hydrothermal vents, endosymbionts, methanotroph

Running title: Biogeography of Lost City Symbionts

Summary

Organisms at hydrothermal vents inhabit discontinuous chemical "islands" along mid-ocean ridges, a scenario that may promote genetic divergence among populations. The 2003 discovery of mussels at the Lost City Hydrothermal Field provided a means of evaluating factors that govern the biogeography of symbiotic bacteria in the deep-sea. The unusual chemical composition of vent fluids, the remote location, and paucity of characteristic vent macrofauna at the site, raised the question of whether microbial symbioses existed at the extraordinary Lost City. And if so, how did symbiotic bacteria therein relate to those hosted by invertebrates at the closest known hydrothermal vents along the Mid-Atlantic Ridge (MAR)? To answer these questions, we performed microscopic and molecular analyses on the bacteria found within the gill tissue of *Bathymodiolus* mussels (Mytilidae, Bathymodiolinae) that were discovered at the Lost City. Here we show that Lost City mussels harbour chemoautotrophic and methanotrophic endosymbionts simultaneously. Furthermore, populations of the chemoautotrophic symbionts from the Lost City and two sites along the MAR are genetically distinct from each other, which suggests spatial isolation of bacteria in the deep-sea. These findings provide new insights into the processes that drive diversification of bacteria and evolution of symbioses at hydrothermal vents.

Introduction

Recent evidence suggests that microbial populations in spatially and chemically fragmented habitats exhibit geographic structure (Whitaker et al., 2003; Papke et al., 2003) rather than being distributed ubiquitously as previously hypothesized (see Finlay 2002; Fenchel 2003). The patchy mosaic of populations in heterogeneous environments restricts gene flow, while promoting genetic differentiation and local adaptation (Slatkin 1987). Due to the heterogeneous nature of hydrothermal vent environments, chemosynthetic bacteria inhabiting vents probably have geographically structured populations as well. If so, this would have direct implications for how topographic features of the seafloor, deep-ocean currents, and chemically variable environments impact the evolution and diversity of bacteria, the origin and evolution of bacteria-vent invertebrate symbioses, and the assemblage of hydrothermal vent communities.

The fragmented distribution of deep-sea hydrothermal vents lies in stark contrast to the uniform conditions of the marine abyssal zone (Tunnicliffe, 1988, 1991; Tunnicliffe and Fowler, 1996; Van Dover, 2000). Discrete hydrothermal vent fields are comparable to islands, distributed in a spatially, chemically, and temporally patchy chain along the deep-sea ridges and remote, off-axis sites (Tunnicliffe, 1988, 1991; Tunnicliffe and Fowler, 1996; Tunnicliffe et al., 1998; Van Dover et al., 2002). Differences between ridges in geography, tectonic activity, age of spreading center, and connectedness of ridge segments likely play a major role in regulating gene flow among populations (Vrijenhoek, 1997; Van Dover et al., 2002; Hurtado et al., 2003), the distribution of vent macrofauna (Van Dover, 1995; Tunnicliffe and Fowler, 1996; Juniper and Tunnicliffe, 1997), and the composition of ecological communities (Tunnicliffe, 1991). For example, 'fracture zones' (Fig. 1) likely inhibit dispersal of

larvae by separating ridge segments that are undergoing independent volcanic evolution (Van Dover et al. 2002). Though associations between chemosynthetic bacteria and their invertebrate hosts provide the basis for macrofaunal production at deep-sea hydrothermal vents, almost nothing is known about the distribution of genetic variation in the symbionts and how population structure of bacteria affects ecological interactions and the evolution of symbioses at vents.

Most dominant vent macrofauna host endosymbiotic bacteria for the capture of chemical energy, yielding a constant food source in this stochastic environment (Fisher, 1990; Cavanaugh et al., 2005; Stewart et al., 2005). Typically, the host provides the symbionts with simultaneous access to oxygen and reduced compounds, and the bacteria, in turn, supply the host with fixed carbon generated from C_1 compounds. Along the Mid-Atlantic Ridge (MAR), vents are inhabited by two species of mussels, *Bathymodiolus azoricus* and *B. puteoserpentis* (Mytilidae; Bathymodiolinae), that host chemoautotrophic (energy source: reduced compounds such as H_2S; carbon source: CO_2) and methanotrophic (energy and carbon source: CH_4) γ-Proteobacteria within their gill tissue (Cavanaugh et al., 1992; Distel et al., 1995; Nelson et al., 1995; Fiala-Médioni et al., 2002). The unique capacity of some bathymodioline mussels to house dual endosymbionts permits the host to utilize multiple compounds for energy acquisition, allowing colonization of diverse environments (Distel et al., 1995; Fiala-Médioni et al., 2002; DeChaine and Cavanaugh, 2005).

The off-axis location, shallow depth (~800 m), distinct chemical environment, and scarcity of known symbiont-hosting invertebrates at the Lost City Hydrothermal Field (LC) suggested the possibility that mussels discovered there (*B.* aff. *azoricus*, T. Shank unpublished data), host endosymbionts that are different from those on the MAR.

For example, the vent fluids of the off-axis Lost City are relatively cool (10-90°C), alkaline (pH≥10), methane-rich (0.13-0.28 mmol kg^{-1}), and lie in stark contrast to the acidic, sulfide-rich effluent of the MAR vents (200-360°C, pH=3-5) that are ~15 km east and 2200 m deeper (Kelley, et al., 2001, 2005). The abundance of methane and hydrogen and low availability of H_2S (due to the high pH of the fluids, given the pKa1 of H_2S = 7.04; Budavari 1996) at the LC suggested that mussels found therein might host primarily methanotrophs, and not thioautotrophs, which would be novel given that all known vent mussels in the Atlantic Ocean host dual symbionts. Furthermore, the apparent remoteness of the Lost City provided a unique setting to resolve whether symbiotic bacterial populations at hydrothermal vents are ubiquitous or structured. Answers to these questions provide a foundation for understanding the forces that promote genetic divergence among symbiont populations and thus determine the biogeography and evolution of bacteria.

The objectives of this study were to 1) determine if the bathymodioline mussels inhabiting the LC hosted symbiotic bacteria and 2) how the symbionts were related, both phylogenetically and demographically, to those hosted by mussels on the MAR. First, we employed transmission electron microscopy (TEM) to determined the presence and morphology of putative symbionts within the Lost City mussel gill tissue and then resolved the relationship of the Lost City symbionts with known symbiotic and free-living bacteria using sequence data from the conserved 16S rRNA gene (Woese, 1987). The bathymodioline chemoautotrophs from the Lost City, and from two MAR fields, Lucky Strike and Snake Pit, were selected for additional population genetic analyses because: 1) bathymodiolines apparently acquire their chemoautotrophic symbionts from the environment each generation (Won et al., 2003), and thus serve as sampling vessels

of the free-living bacterial population, 2) the chemoautotrophs only have one ribosomal RNA operon, precluding concerns over non-orthologous genetic variation (Won et al., 2003), and 3) chemoautotrophs are more widespread among invertebrate hosts at hydrothermal vents than methanotrophic symbionts (Cavanaugh et al., 2005), thus permitting broad genetic comparisons. An intraspecific phylogeny and the demographic history of each chemoautotroph population (defined by location) were inferred from sequence data of the rapidly evolving 16S-23S rRNA internal transcribed spacer (ITS; Antón et al., 1998). By employing both conserved and highly variable markers, the phylogenetic position of Lost City mussel symbionts was resolved at two scales, within the γ-Proteobacteria and among populations of chemoautotrophs hosted by vent-endemic mussels along the northern MAR.

Results

Characterization of the Lost City symbiosis

This characterization, which constitutes the first description of a symbiosis from the LC, revealed two morphologically distinct Gram negative bacteria in the mussel bacteriocytes, gill epithelial cells specialized for housing symbiotic bacteria (Fig. 2a). As in other vent symbioses, the bacteriocytes were separated by symbiont-free intercalary cells. Vacuoles within a bacteriocyte harbored either several coccoid bacteria (~0.3 μm in diameter) or a single, large bacterium (~1.5-2.0 μm) exhibiting intracytoplasmic membranes typical of type I methanotrophs (Fig 2b). Based on the characteristics of endosymbionts in other bathymodioline mussels (Cavanaugh et al., 1992; Fiala-Médioni et al., 2002; Robinson et al., 1998), the small and large bacteria were inferred to be chemoautotrophs and methanotrophs, respectively.

Phylogenetic analyses of 16S rRNA sequence data corroborated the TEM observations of dual endosymbionts in the Lost City mussels. First, sequence alignments revealed that the symbiont phylotypes from the Lost City (Genbank accession numbers A and B) were identical to two phylogenetically distinct lineages of γ-Proteobacteria, a chemoautotroph and a methanotroph, previously found in both *B. azoricus* and *B. puteoserpentis* on the MAR. The presence of both phylotypes in the gill tissue of these two MAR mussel species has been verified through *in situ* hybridization with phylotype-specific probes (Distel et al., 1995; Duperron et al., 2005). In our analyses, chemoautotrophic and methanotrophic symbionts of the mussels formed separate, well-supported monophyletic clades that were nested with chemoautotrophs of vent-endemic vesicomyid clams and free-living methanotrophs, respectively (Fig. 3). This finding demonstrates the tight ecological and historic specificity of the interaction between the mussels and two distinct subsets of the γ Proteobacteria. While both Bayesian and maximum parsimony analyses inferred similar tree topologies, the deep relationships among the mussel methanotroph, mussel and clam chemoautotroph, and other symbiont clades remain uncertain (for $\alpha = 0.05$). Finally, the occurrence of both symbiont phylotypes across the 2-3 mussel species demonstrated a lack of host fidelity and implied that both methanotrophs and chemoautotrophs were acquired from the local environment.

All mussel individuals harboured several, distinct chemoautotrophic symbiont ITS-genotypes (Genbank accession numbers X through Y). Twenty-four percent of the ITS-genotypes from Lucky Strike and 38% from Snake Pit were shared among host individuals within each of those localities (Fig. 4). In contrast, the two Lost City host mussels had no chemoautotroph ITS-genotypes in common with each other, possibly

157 owing to the small population size of mussels at that location. The occurrence of
158 multiple, geographically restricted chemoautotroph ITS-genotypes within an individual
159 host reinforced the contention that each individual mussel acquired its symbionts from
160 the local, free-living bacterial community as shown by the distribution of 16S rRNA
161 phylotypes in this study and previous analyses of ITS variation (Won et al., 2003).
162
163 *Biogeography of bathymodioline chemoautotrophic endosymbionts*
164 Analyses of ITS sequence data showed that the chemoautotrophic symbionts of
165 bathymodioline mussels were not distributed ubiquitously, but rather exhibited
166 population structure associated with geographic location. This finding, which contrasts
167 with the observed ubiquity of the 16S rRNA phylotype (above and Duperron et al.,
168 2005), underscores the need to use highly variable markers in analyses at the population
169 level. The genetic variation in the ITS region (1.05 % average pair-wise sequence
170 divergence) permitted resolution of evolutionary relationships among populations of
171 chemoautotrophic symbionts at hydrothermal fields. Two distinct ITS-clades of
172 chemoautotrophs were separated by 13 nucleotide substitutions: the *Bathymodiolus*
173 *puteoserpentis* (Snake Pit) symbionts and the *B. azoricus* - *B.* aff. *azoricus* clade, which
174 included symbionts from both Lucky Strike and the Lost City (Fig. 4). Furthermore, the
175 overall estimates of θ (= 6.9) and T (= 1.6) from MDIV imply that the northern and
176 southern populations of *Bathymodiolus* chemoautotrophic symbionts in the north
177 Atlantic (as defined by the ITS-clades) are large and historically have been separated
178 from one another.
179 Our genetic analyses revealed that populations of chemoautotrophic symbionts
180 inhabiting different hydrothermal vent fields were isolated and experienced independent

181 demographic histories. First, populations of chemoautotrophs at the Lost City and
182 Lucky Strike were more genetically diverse, as estimated by θ based on the number of
183 segregating sites (W) and the average pair-wise nucleotide diversity (π) for haploid
184 genomes, than the population at Snake Pit (Table 1). Moreover, Tajima's D tests of
185 neutrality suggest that the populations at the Lost City and at Lucky Strike have been
186 demographically stable, whereas the symbionts at Snake Pit likely experienced a
187 population bottleneck (a reduction in population size followed by rapid population
188 growth; Table 1). We cannot rule out, however, the possibility of a selective sweep for
189 Snake Pit symbionts, because Tajima's D does not effectively differentiate between
190 population processes and selection (Tajima 1989). Finally, based on F_{ST} estimates of
191 isolation, our analyses revealed genetic divergence among populations of
192 chemoautotrophic symbionts at all study locations, irrespective of host species (Table 2)
193 or distance between sites (no isolation-by-distance, $p = 0.9$).
194
195 **Discussion**
196 Though the LC lies distantly off-axis and has a novel chemical environment
197 (Kelley, et al., 2001, 2005), mussels in the genus *Bathymodiolus* found at the Lost City
198 host dual symbionts, a methanotroph and a chemoautotroph, with identical 16S rRNA
199 phylotypes as those along the Mid-Atlantic Ridge (MAR). This result is unexpected
200 given the paucity of H_2S (due to the high pH) in the effluent of Lost City vents (D.
201 Butterfield pers. comm.) and raises the question of whether the chemoautotrophs are
202 using sources of energy other than sulfur compounds, such as hydrogen that is abundant
203 at the vent fluids. Indeed, alternate energy sources may be used by many symbionts, as
204 only *B. thermophilus* found along the Eastern Pacific Rise have been shown to use

205 sulfur (Belkin et al. 1986; Nelson et al. 1995). The discovery and characterization of the
206 Lost City bathymodioline symbionts, in light of the diversity of chemical environments
207 inhabited by mussels, underscores the ecological and evolutionary stability of the dual
208 symbiosis.

209 The occurrence of single phylotypes, for both the chemoautotroph and the
210 methanotroph, across different host species demonstrated that neither of the symbiont
211 types was host species-specific. A similar lack of host-species fidelity was shown for
212 the chemoautotrophic endosymbionts of hydrothermal vent tubeworms that were
213 inferred to be environmentally transmitted (Feldman et al., 1997; Nelson and Fisher,
214 2000; reviewed in Cavanaugh et al., 2005). Our analyses revealing the broad
215 distribution of symbiont phylotypes across multiple host species suggest that *both*
216 methanotrophs and chemoautotrophs of mussels in the northern Atlantic are acquired
217 from the environment, rather than being transmitted from mother to offspring each
218 generation as for the closely related chemoautotrophs of another vent bivalve,
219 *Calyptogena magnifica* (Cary and Giovannoni, 1993). This finding implies that mussels
220 acquire symbionts from the local community when they colonize a site and has
221 implications for local adaptation of symbionts to that environment.

222 Chemoautotrophic symbiont populations hosted by bathymodioline mussels
223 were inferred to be isolated from each other because no ITS-genotypes were shared
224 among the three hydrothermal fields. Since *B. azoricus* individuals at the Broken Spur
225 hybrid zone on the MAR (just south of the Lost City) harboured symbionts from both
226 northern (*B. azoricus*) and southern (*B. puteoserpentis*) ITS-clades (Won et al., 2003), it
227 is unlikely that the host organisms affected the distribution of ITS genotypes, though the
228 host may have selected for certain bacterial phylotypes from the local free-living

229 population. Because the chemical environment of the Lost City is drastically different
230 from both vent sites on the MAR (Kelley et al. 2001), and symbiont populations from
231 Lost City and Lucky Strike are closely related phylogenetically while those from the
232 two MAR sites are not, the chemical environment may not be a large factor in
233 governing the distribution of bathymodioline symbionts. Rather, geography probably
234 played a major role in generating isolation among populations.

235 The three sites in this study are separated by fracture zones, depth, distance, and
236 deep ocean currents, all of which have been implicated as dispersal barriers that could
237 promote genetic divergence among populations (Van Dover et al., 2002). Though the
238 distance between the Lost City and Lucky Strike to the north (1253 km) is greater than
239 the distance between the Lost City and Snake Pit to the south (832 km), the Lost City
240 chemoautotrophs cluster with those in the north (Fig. 4). Thus, we inferred that distance
241 did not have as large an effect on isolation as did topographic features that likely
242 influence deep-ocean currents. For instance, the two ITS-clades (Fig. 4) are
243 geographically separated by many transform faults that offset the spreading axis,
244 including the ~6000 m deep Atlantis Fracture Zone, just to the south of the Lost City
245 and the Kane Fracture Zone just north of Snake Pit (Fig. 1). Smaller fracture zones, such
246 as the Oceanographer Fracture Zone to the north, are apparently not as strong of barriers
247 to dispersal, but this remains to be evaluated. Thus, understanding the biogeographic
248 history of bacteria that inhabit hydrothermal vents provides an empirical basis for and
249 an independent means of assessing models of deep-ocean currents.

250 Isolation among hydrothermal vent fields has likely led each population of
251 chemoautotrophic symbionts to experience independent demographic histories, as
252 inferred through differences in θ, Tajima's D tests of neutrality, and the high levels of

253 isolation estimated by F_{ST}. Because the 16S-ITS-23S spacer is mostly comprised of
254 seemingly functionless regions (Antón et al., 1998), the possibility that selection caused
255 the observed patterns of genetic variation is unlikely. Rather, the demographic history of
256 a symbiont population may depend on the tectonic activity at the site, which, in addition
257 to supplying the bacteria with reduced compounds for energy production, could
258 decimate the population in an intense eruption. For example, we inferred that the
259 population of chemoautotrophic symbionts at Snake Pit was unstable, while the other
260 two populations were at equilibrium. A long-lived hydrothermal vent field, such as the
261 Lost City (Früh-Green et al., 2003) may maintain a heterogeneous and stable population
262 of chemoautotrophs, while shorter-lived or more eruptive sites may generate greater
263 fluctuations in population size and thus reduce genetic diversity.
264 Our findings fit with the biogeographic model for macrofauna larvae outlined by
265 Van Dover et al. (2002), which states that the greater degree of faulting along slow-
266 spreading ridges (e.g., the MAR) should serve to isolate populations. Since symbionts
267 are acquired from the local environment each generation, the host likely plays little role
268 in determining the distribution of genetic variation in bacterial populations among
269 locations. Rather, the strong divergence between northern and southern symbiont
270 populations and the lack of isolation-by-distance among localities demonstrated that
271 geographic barriers to dispersal, such as faulting, depth, and other topographic features
272 of the seafloor, divide bacterial populations. Also, though off-axis sites may be remotely
273 located, they may be connected (or have a historic connection) via deep-ocean currents
274 with sites along the ridge, as indicated by the Lost City populations clustering with
275 those of Lucky Strike. We conclude that topography is a major influence on the
276 distribution of diversity among populations of symbiotic bacteria at hydrothermal vents,

277 and that additional research is needed to clarify how differences among ridges in
278 tectonic activity, geography, and physical oceanography have impacted the population
279 structure of symbiotic bacteria and at what scale.

280 Resolving how populations of bacterial endosymbionts are structured has
281 important implications for microbial biogeography, bacterial diversity and evolution,
282 the origin and evolution of prokaryote-eukaryote symbioses, and the ecology and
283 evolution of life at deep-sea hydrothermal vents. First, studies in microbial
284 biogeography have revealed that limits to gene flow might yield geographic structure
285 within microbial taxa (Papke et al. 2003; Whitaker et al. 2003; Kirchman et al., 2005).
286 Population subdivision implies an increased potential for local adaptation and lineage
287 diversification. Until now, genetic structure and potential for local adaptation in
288 chemosynthetic endosymbionts have remained uncertain. Environmentally transmitted
289 endosymbionts are expected to respond to abiotic selective forces in the environment as
290 well as experience gene transfer with the hydrothermal vent free-living bacterial
291 community. This will not only impact the genetic diversity of symbionts, but may
292 ultimately affect the fitness of the invertebrate host. This and future studies on the
293 biogeography of symbionts inhabiting deep-sea hydrothermal vents, including
294 comparisons with the free-living bacterial community, host biogeography, and among-
295 site variation in environmental factors, will provide a basis for understanding the
296 processes responsible for the diversification of bacteria and symbioses on this planet.
297

298 **Experimental Procedures**

299 *Specimen collection*

300 Mussels were collected using DSV Alvin from the off-axis Lost City
301 hydrothermal vent field (30°07.40'N, 42°07.24'W; 800 m deep) and from the Lucky
302 Strike (37°17.26'N, 32°16.50'W; 1693 m deep) and Snake Pit (23°22.10'N,
303 44°56.91'W; 3492 m deep) vent sites on the MAR (Fig. 1). Specimens were preserved
304 for ultrastructural analysis or stored at -80^{o}C. Symbiont-bearing gill tissue was fixed,
305 embedded, and examined by transmission electron microscopy (Distel et al., 1995).
306 DNA was extracted from the frozen gill tissue of the two Lost City mussels, 20
307 individuals of *Bathymodiolus azoricus* from Lucky Strike and 20 of *B. puteoserpentis*
308 from Snake Pit with DNeasy Tissue Extraction Kits (Qiagen, Valencia, CA).
309
310
311 *Genetic sampling and analyses*
312 To resolve evolutionary relationships, the symbiont(s) 16S rRNA gene was
313 amplified using the universal bacteria primers 27f and 1492r (Weisburg et al., 1991),
314 from multiple specimens of the three vent sites, gel purified (Qiagen Gel Extraction
315 Kit), and cloned (TOPO TA Cloning Kit; Invitrogen Corp., Frederick, MD). Thirty-two
316 clones per host population were analyzed (16 for each of the two mussels from the Lost
317 City and two clones for each of 16 mussels at both the Lucky Strike and Snake Pit
318 sites). The legitimacy of point mutations in all unique phylotypes was evaluated using
319 ARB (Ludwig et al., 2004) by assessing complementary base pairing on the 16S rRNA
320 secondary structure and by following the sequence conservation rule (Acinas et al.,
321 2004).
322 To estimate within- and among-population genetic variation, sequence data from
323 the polymorphic 16S-ITS-23S region of the chemoautotrophs was used. The marker was

amplified using two symbiont-specific primer combinations: Sym-ITS-830F and Sym-ITS-23SR; Sym-ITS-1322F and Sym-ITS-23SR (Won et al., 2003). The former primer set was used to confirm symbiont species identification, because it yielded an 1800 nucleotide sequence including approximately 600 bp of 16S rRNA. The latter pair provided the ITS sequences for population genetic analyses. Ninety clones from the Sym-ITS-1322F and Sym-ITS-23SR amplicons were sequenced from each of the two Lost City mussels. For both the Lucky Strike and Snake Pit populations, 143 clones were sequenced from 20 host individuals.

For each locus, forward and reverse strands were cycle sequenced using the M13 primer pair, the BigDye Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems, Atlanta, GA), cleaned with Performa DTR 96-well Std. Plate Kit (Edge BioSystems, Gaithesburg, MD), and sequenced on an ABI 3730 Gene Analyzer. Sequences were edited in Sequencher 4.1.2 (Gene Codes Corp.), aligned in ClustalX (Thompson et al., 1997), and alignments were manually edited in MacClade 4.0 (Maddison and Maddison, 2003).

The phylogenetic relationships among the 1303 bp portion of the 16S rRNA phylotypes from this and previous studies of bathymodioline symbionts (with *Escherichia coli* as an outgroup; see Table 3 for Genbank accession numbers) were inferred using maximum parsimony in PAUP 4.10b (Swofford, 2003) and Bayesian posterior probabilities implemented with MrBayes v3.0b4 (Huelsenbeck and Ronquist, 2001). Maximum parsimony trees were generated on PAUP 4.10b (Swofford, 2003), with heuristic searches, random sequence addition with 100 replicates, and TBR branch swapping. Significance was determined from a 1000 replicate bootstrap analysis using the same search parameters. From the Bayesian analysis, using four-chain Metropolis-

348 coupled Markov chain Monte Carlo (MCMCMC) analysis, a consensus tree of 11,000
349 post burn-in sampled trees was generated in PAUP 4.10b (Swofford, 2003). Both
350 Bayesian and parsimony analyses yielded similar inferences of evolutionary history.
351 Intraspecific phylogenies were inferred from the ITS sequences for the
352 chemoautotrophic symbionts using parsimony implemented in the TCS software
353 package (Clement et al., 2000) and a combination of Bayesian and maximum likelihood
354 analyses. Eighty-five nucleotides of tRNA-Ala and 77 bp of tRNA-Ile occurred within
355 the 16S-ITS-23S sequence. The 1200 bp of ITS included five indels at positions 361-
356 364, 542-543, 711, 937-939, and 966-982, which were each transformed into a single
357 polymorphic position (Widmer and Baltisberger, 1999). After converting the indels to
358 one base substitution each, 48 polymorphic sites were described for the remaining 1178
359 bp, of which 36 sites were parsimony informative. For the Bayesian analysis, post burn-
360 in trees were imported into PAUP 4.10b (Swofford, 2003) and sorted to choose the
361 maximum likelihood tree. The parsimony and maximum likelihood trees were similar
362 and the few differences did not affect any conclusions.
363 Within-population genetic variation and among-population genetic
364 differentiation were estimated to test the relationships between the Lost City
365 chemoautotrophic endosymbionts and the two populations on the MAR. All measures
366 were averaged across individuals from the population to account for potential PCR bias.
367 First, an overall measure of genetic diversity for haploid genomes ($\theta = 2Ne\mu$) for all
368 populations and the amount of genetic divergence (T) between northern and southern
369 clades (see Results) were estimated using MDIV (Nielsen and Wakeley 2001) assuming
370 the HKY finite sites model and running the coalescent simulations three times for each
371 species to evaluate convergence for each parameter. Within-population genetic diversity

372 was estimated based on the average pair-wise nucleotide diversity (θ_π) and the number
373 of segregating sites (θ_W) for haploid genomes (e.g., Herbeck et al., 2003). In addition,
374 both estimators of θ should be equivalent in a population at equilibrium that is evolving
375 neutrally. Tajima's D was used to compare the two estimators of θ and examine whether
376 populations were at equilibrium (Tajima, 1989). To test whether or not populations of
377 bacterial symbionts were isolated, the mean pair-wise differences and degree of
378 differentiation (F_{ST}) among locations were estimated (e.g., Whitaker et al., 2003).
379 Isolation-by-distance was tested (Rousset, 1997), with distances among sites as follows:
380 Lost City to Lucky Strike (1253 km), Lost City to Snake Pit (832 km), and Lucky Strike
381 to Snake Pit (2037 km). All analyses were performed using Arlequin 2.0 (Schneider et
382 al., 2000).
383

Acknowledgements

We express our deep appreciation to the captain and crews of the R/V Atlantis and DSV Alvin for their immeasurable assistance in specimen collection (OCE 0136871, T. Shank). The sequencing and analyses were funded by an NSF Microbial Biology Postdoctoral Fellowship for E. G. DeChaine (DBI-0400591) and NSF grants for C. M. Cavanaugh (OCE-0453901, DEB-0089738). For sample collections and unpublished sequence data the authors would like to thank Z. McKiness. We would also like to thank D. Stahl and three anonymous reviewers for their constructive feedback.

392 **References**

393 Acinas, S. G., V. Klepac-Ceraj, D. E. Hunt, C. Pharino, I. Ceraj, D. L. Distel, M. F.
394 Polz. 2004. Fine-scale phylogenetic architecture of a complex bacterial
395 community. *Nature* **430**: 551-554.

396 Antón, A. I., A. J. Martínez-Murcia, F. Rodríguez-Valera. 1998. Sequence diversity in
397 the 16S-23S intergenic spacer region (ISR) of the rRNA operons in
398 representatives of the *Escherichia coli* ECOR collection. *J. Mol. Evol*. **47**: 62-72.

399 Belkin, S., D. C. Nelson, and H. W. Jannasch. 1986. Symbiotic assimilation of CO_2 in
400 two hydrothermal vent animals, the mussel *Bathymodiolus thermophilus* and the
401 tubweorm *Riftia pachyptila*. *Biol*. *Bull*. **170**: 110-121.

402 Budavari, S. 1996. *The Merck Index - An Encyclopedia of Chemicals, Drugs, and*
403 *Biologicals*. Whitehouse Station, NJ: Merck and Co., Inc., p. 823.

404 Cary, S. C. and S. J. Giovannoni. 1993. Transovarial inheritance of endosymbiotic
405 bacteria in clams inhabiting deep-sea hydrothermal vents and cold seeps. *Proc.*
406 *Natl. Acad. Sci. USA*. **90**: 5695-5699.

407 Cavanaugh, C. M., C. Wirsen, and H. J. Jannasch. 1992. Evidence for methylotrophic
408 symbionts in a hydrothermal vent mussel (Bivalvia: Mytilidae) from the Mid-
409 Atlantic Ridge. *Appl. Environ. Microbiol*. **58**: 3799-3803.

410 Cavanaugh, C.M., Z.P. McKiness, I.L.G. Newton, and F.J. Stewart. 2005. Marine
411 chemosynthetic symbioses. In M. Dworkin et al., Eds., *The Prokaryotes: An*
412 *Evolving Electronic Resource for the Microbiological Community*, Springer-
413 Verlag, New York.

414 Clement, M., D. Posada, K. A. Crandall. 2000. TCS: a computer program to estimate
415 gene genealogies. *Molec. Ecol*. **9**: 1657-1659.

416 DeChaine, E. G. and C. M. Cavanaugh. 2005. Symbioses of methanotrophs and deep-
417 sea mussels (Mytilidae: Bathymodiolinae). In J. Overmann, Ed., *Molecular*
418 *Basis of Symbiosis*, Sinauer Assoc. In press.
419 Distel, D., H. K. Lee, and C. M. Cavanaugh. 1995. Intracellular coexistence of
420 methano- and thioautotrophic bacteria in a hydrothermal vent mussel. *Proc.*
421 *Natl. Acad. Sci. USA.* **92**: 9598-9602.
422 Duperron, S. C. Bergin, F. Zielinski, Z. P. McKiness, E. G. DeChaine, M. Sibuet, C. M.
423 Cavanaugh, and N. Dubilier. 2005. A dual symbiosis shared by two
424 bathymodioline mussels (Bivalvia: Mytilidae) from the Mid-Atlantic Ridge.
425 *Environ. Microbiol*. In review.
426 Feldman, R., Black, M., Cary, C., Lutz, R., and R. Vrijenhoek. 1997. Molecular phylogenetics
427 of bacterial endosymbionts and their vestimentiferan hosts. *Molec. Mar. Biol. Biotech.*
428 **6**: 268-277.
429 Fenchel, T. 2003. Biogeography for bacteria. *Science* **301**: 925-926.
430 Fisher, C. R. 1990. Chemoautotrophic and methanotrophic symbioses in marine
431 invertebrates. *Rev. Aqua. Sci.* **2**: 399-436.
432 Fiala-Médioni, A., Z. McKiness, P. Dando, J. Boulegue, A. Mariotti, A. Alayse-Danet,
433 J. Robinson, and C. Cavanaugh. 2002. Ultrastructural, bioghemical, and
434 immunological characterization of two populations of a new species of Mytilid
435 mussel, *Bathymodiolus azoricus*, from the Mid-Atlantic Ridge: evidence for a
436 dual symbiosis. *Mar. Biol.* **141**: 1035-1043.
437 Finlay, B. J. 2002. Global dispersal of free-living microbial eukaryote species. *Science* **296**:
438 1061-1063.

439 Früh-Green, G. L., D. S. Kelley, S. M. Bernasconi, J. A. Karson, K. A. Ludwig, D. A.
440 Butterfield, C. Boschi, and G. Proskurowski. 2003. 30,000 years of
441 hydrothermal activity at the Lost City vent field. *Science* **301**: 495-498.
442 Herbeck, J. T., D. J. Funck, P. H. Degnan, & J. J. Wernegreen. 2003. A conservative
443 test of genetic drift in endosymbiotic bacterium Buchnera: Slightly deleterious
444 mutations in the chaperonin groEL. *Genetics* **165**: 1651-1660.
445 Huelsenbeck, J. P. & Ronquist, F. 2001. MrBayes: Bayesian inference of phylogenetic
446 trees. *Bioinformatics* **17**, 754-755.
447 Hurtado, L. A., m. Mateos, R. A. Lutz, and R. C. Vrijenhoek. 2003. Coupling of bacterial
448 endosymbiont and host mitochondrial genomes in the hydrothermal vent clam
449 *Calyptogena magnifica*. *Applied Environ. Microbiol.* **69**: 2058-2064.
450 Juniper, S. K. and V. Tunnicliffe. 1997. Crustal accretion and the hot ecosystem. *Philos. Trans.*
451 *R. Soc. Lond. A* **355**: 450-474.
452 Kelley, D. S., J. A. Karson, D. K. Blackman, G. L. Früh-Green, D. A. Butterfield, M.
453 D. Lilley, E. J. Olson, M. O. Schrenk, K. K. Roe, G. T. Lebon, P. Rivizzigno,
454 and the AT3-60 Shipboard Party. 2001. An off-axis hydrothermal vent field near
455 the Mid-Atlantic Ridge at 30°N. *Nature* **412**: 145-149.
456 Kelley, D. S., Karson, D. K., Früh-Green, G. L., Yoerger, D. R., Shank, T. M.,
457 Butterfield, D. A., Hayes, J. M., Schrenk, M. O., Olson, E. J., Proskurowski, G.,
458 Jakuba, M., Bradley, A., Larson, B., Ludwig, K., Glickson, D., Buckamn, K.,
459 Bradley, A. S., Brazelton, W. J., Roe, K., Elend, M. J., Delacour, A.,
460 Baernasconi, S. M., Lilley, M. D., Baross, J. A., Summons, R. E., and S. P.
461 Sylva. 2005. A serpentinite-hosted ecosystem: The Lost City Hydrothermal
462 Field. *Science* **307**: 1428-1434.

463 Kirchman, D. L., Dittel, A. I., Malmstrom, R. R., and M. T. Cottrell. 2005. Biogeography of
464 major bacterial groups in the Delaware estuary. *Limnol. Oceanog*. **50**: 1697-1706.
465 Ludwig, W., O. Strunk, R. Westram et al. 2004. ARB: a software environment for
466 sequence data. *Nucleic Acid Res*. **32**: 1363-1371.
467 Maddison, D. R. and W. P. Maddison. 2000. *MacClade 4*. Sinauer Assoc. Inc.
468 Sunderland, MA.
469 Nelson, D. C. and C. R. Fisher. 2000. Absence of cospeciation in deep-sea
470 vestimentiferan tubeworms and their bacterial endosymbionts. *Symbiosis*. **28**: 1-
471 15.
472 Nelson, D. C., K. D. Hagan, and D. B. Edwards. 1995. The gill symbiont of the
473 hydrothermal vent mussel *Bathymodiolus thermophilus* is a psychrophilic,
474 chemoautotrophic, sulfur bacterium. *Mar. Biol*. **121**: 487-495.
475 Nielsen, R., and J. W. Wakeley. 2001. Distinguishing migration from isolation: and
476 MCMC approach. *Genetics* **158**: 885-896.
477 Papke, R. T., N. B. Ramsing, M. M. Bateson, and D. M. Ward. 2003. Geographical
478 isolation in hot spring cyanobacteria. Environ. Microbiol. 5: 650-659.
479 Robinson, J. J., M. F. Polz, A. Fiala-Médioni, and C. M. Cavanaugh. 1998.
480 Physiological and immunological evidence for two distinct C-1-utilizing
481 pathways in *Bathymodiolus puteoserpentis* (Bivalvia: Mytilidae), a dual
482 endosyumbiotic mussel from the Mid-Atlantic Ridge. *Mar. Biol.* **132**: 625-633.
483 Rousset, F. 1997. Genetic differentiation and estimation of gene flow from F-statistics
484 under isolation by distance. *Genetics* **145**: 1219-1228.

485 Schneider, S., D. Roessli, and L. Excoffier. 2000. *ARLEQUIN, a Software Package for*
486 *Population Genetics Data Analysis, Version 2.0.* Genetica and Biometry
487 Laboratory, Univ. of Geneva. Geneva, Switzerland.
488 Slatkin, M. 1987. Gene flow and the geographic structure of natural populations.
489 *Science* 236: 787-792.
490 Stewart, F. J., I. L. G. Newton, and C. M. Cavanaugh. 2005. Chemosynthetic
491 endosymbioses: adaptations to oxic-anoxic interfaces. *Trends Microbiol*. **13**:
492 439-448.
493 Swofford, D. L. 2003. *PAUP*. Phylogenetic Analysis Using Parsimony (*and Other*
494 *Methods).* Sinauer Associates, Sunderland, MA.
495 Tajima, F. 1989. The effect of change in population size on DNA polymorphism.
496 *Genetics* **123**: 597-601.
497 Thompson, J. D., T. J. Gibson, F. Plewniak, F. Jeanmougin, and D. G. Higgins. 1997.
498 The ClustalX windows interface: flexible strategies for multiple sequence
499 alignment aided by quality analysis tools. *Nucl. Acid Res*. **24**: 4876-4882.
500 Tunnicliffe, V. 1988. Biogeography and evolution of hydrothermal-vent fauna in the eastern
501 Pacific Ocean. *Proc. R. Soc. Lond. B* **233**: 347-366.
502 Tunnicliffe, V. 1991. The biology of hydrothermal vents: Ecology and evolution. *Oceanogr.*
503 *Mar. Biol. Annu. Rev*. **29**: 319-407.
504 Tunnicliffe, V. and C. Fowler. 1996. Influence of sea-floor spreading on the global
505 hydrothermal vent fauna. *Nature* **379**: 531-533.
506 Tunnicliffe, V., A. G. McArthur, and D. Mchugh. 1998. A biogeographical perspective of the
507 deep-sea hydrothermal vent fauna. *Advances in Marine Biology* **34**: 353-442.
508 Van Dover, C. L. 1995. Ecology of Mid-Atlantic Ridge hydrothermal vents. In: Parson,

509 L. M., C. L. Walker, and D. R. Dixon (eds.). *Hydrothermal Vents and Processes*. *Geol*.
510 *Soc*. *Spec*. *Publ*. **87**: 257-294.
511 Van Dover, C. L. 2000. *The Ecology of Deep-sea Hydrothermal Vents*. Princeton Univ. Press.
512 Princeton, N. J.
513 Van Dover, C. L., C. R. German, K. G. Speer, L. M. Parson, and R. C. Vrijenhoek.
514 2002. Evolution and biogeography of deep-sea vent and seep invertebrates.
515 *Science* **295**: 1253-1257.
516 Vrijenhoek, R. C. 1997. Gene flow and genetic diversity in naturally fragmented
517 metapopulations of deep-sea hydrothermal vent animals. *J*. *Heredity* **88**: 285-293.
518 Weisburg, W. G., S. M. Barns, D. A. Pelletier, and D. J. Lane. 1991. 16S ribosomal
519 DNA amplification for phylogenetic study. *J. Bacteriol*. **173**: 697-703.
520 Whitaker, R. J., S. W. Grogan, and J. W. Taylor. 2003. Geographic barriers isolate
521 endemic populations of hyperthermophilic archaea. *Science* **301**: 976-978.
522 Widmer, A. and M. Baltisberger. 1999. Extensive intraspecific chloroplast DNA
523 (cpDNA) variation in the alpine *Draba aizoides* L. (Brassicaceae): haplotype
524 relationships and population structure. *Molec. Ecol.* **8**: 1405-1415.
525 Woese, C. R. 1987. Bacterial evolution. *Microbiol. Reviews* **51**: 221-271.
526 Won, Y., S. J. Hallam, G. D. O'Mullan, I. L. Pan, K. R. Buck, and R. C. Vrijenhoek.
527 2003. Environmental acquisition of thiotrophic endosymbionts by deep-sea
528 mussels of the genus *Bathymodiolus*. *Appl. Env. Microbiol*. **69**: 6785-6792.
529

529 Table 1. Summary statistics for ITS sequences from the three populations of
530 chemoautotrophic symbionts.
531

532 Location	Host Species	θ_π (sd)	θ_W (sd)	Tajima's D
533 Lucky Strike	*B. azoricus*	2.4 (1.4)	2.8 (1.0)	-0.63
534 Lost City	*B.* aff. *azoricus*	2.3 (1.4)	2.5 (0.8)	-0.23
535 Snake Pit	*B. puteoserpentis*	0.7 (0.6)	1.9 (0.7)	-1.67*

536
537 Estimates of genetic diversity (θ) based on the average pair-wise sequence divergence
538 (π) and number of segregating sites (W) are shown, with standard deviations (sd).
539 Estimates of Tajima's D are given for each population (*denotes $p < 0.01$).
540

Table 2. Pair-wise comparisons of populations of chemoautotrophs hosted by bathymodioline mussels.

		Mean pair-wise differences	
Location	Lucky Strike	Lost City	Snake Pit
Lucky Strike		1.2	11.7
Lost City	0.34		12.4
Snake Pit	0.89	0.89	
	Population pair-wise F_{ST}		

Mean pair-wise differences for chemoautotroph ITS sequences are shown above the diagonal, and population pair-wise F_{ST} values are given below. All F_{ST} values are significant.

553 Table 3. List of bacteria and Genbank accession numbers used to generate the 16S
554 rRNA phylogeny for γ-Proteobacteria (Fig. 3).

555 Environment	Species	Genbank accession no.
556 Free-living bacteria		
557	*Achromatium oxaliferum*	L48227
558	*Beggiatoa alba*	L40994
559	*Escherichia coli*	J01695
560	*Halomonas elongata*	X67023
561	*Hydrogenovibrio marinus*	D86374
562	*Methylobacter whittenburyi*	X72773
563	*M. capsulatus*	L20843
564	*M. luteus*	AF304195
565	*M. vinelandii*	L20841
566	*Methylomicrobium agile*	X72767
567	*M. pelagicum*	L35540
568	*Methylomonas methanica*	AF150806
569	*M. rubra*	AF150807
570	*Pseudomonas mendocina*	AF232713
571	*Rhabdochromatium marinum*	X84316
572	*Thiocystis gelatinosa*	Y11317
573	*Thiomicrospira thyasirae*.	AF16046
574		
575 Chemoautotrophic symbionts		
576 Host Taxonomy	Host species	

Phylum Annelida			
	Oligochaeta	*Inanidrilus leukodermatus*	U24110
		Olavius loisae	AF104472
	Vestimentifera	*Escarpia spicata*	U77482
		Lamellabrachia columna	U77481
		Ridgeia piscesae	U77480
		Riftia pachyptila	M99451
Phylum Mollusca			
	Bivalvia		
	Lucinidae	*Codakia orbicularis*	X84979
		Lucina nassula	X84980
		Lucinoma aequizonata	M99448
	Mytilidae	*Bathymodiolus* aff. *brevior*	DQ077891
		B. puteoserpentis	U29163
		B. azoricus - puteoserpentis	AM083974 and this study
		B. septemdierum	AB036709
		B. thermophilus	M99445
		B. sp. Gabon Margin	AJ745718
		B. sp. Juan de Fuca	Z. McKiness unpub. data
	Thyasiridae	*Thyasira flexuosa*	L01575
	Vesicomyidae	*Calyptogena elongata*	AF035719
		C. fossajaponica	AB044744
		C. phaseoliformes	AF035724
		C. kilmeri	AF035720

601		*C. magnifica*	AF035721
602		*C. pacifica*	AF035723
603		*Ectenogena extenta*	AF035725
604		*Vesicomya gigas*	AF035726
605	Phylum Nematoda		
606	Desmodoridae	*Laxus* sp.	U241110
607			
608	Methanotrophic symbionts		
609	Host Taxonomy	Host species	
610	Phylum Mollusca		
611	Bivalvia		
612	Mytilidae	*B. puteoserpentis* M	U29164
613		*B. azoricus - puteoserpentis* M	AM083950 and this study
614		*B. childressi*	U05595
615		*B. japonicus*	AB036711
616		*B. platifrons*	AB036710
617		*B.* sp. Gabon Margin M	AJ745717
618			

618 **Figure Legends**

619 Figure 1. Map of study sites and the Mid-Atlantic Ridge (MAR). The collection sites,
620 including the Lost City, Lucky Strike, and Snake Pit vent fields, are designated by white
621 circles and labelled. The MAR and its dominant fracture zones (F. Z.) are highlighted
622 by black lines. Several fracture zones relevant to the discussion are also labelled to
623 emphasize the geographically discontinuous nature of the MAR.

624

625 Figure 2. Transmission electron micrographs of endosymbionts within the gill tissue of
626 a Lost City mussel. A. Chemoautotrophic (C) and type I methanotrophic (M) symbionts
627 within the apical portion of two bacteriocytes (bc) separated by a symbiont-free
628 intercalary cell (ic). Scale bar = 2 μm. B. Higher magnification of the two symbiont
629 morphotypes; note intracytoplasmic membranes of the type I methanotroph. Scale bar =
630 0.5 μm.

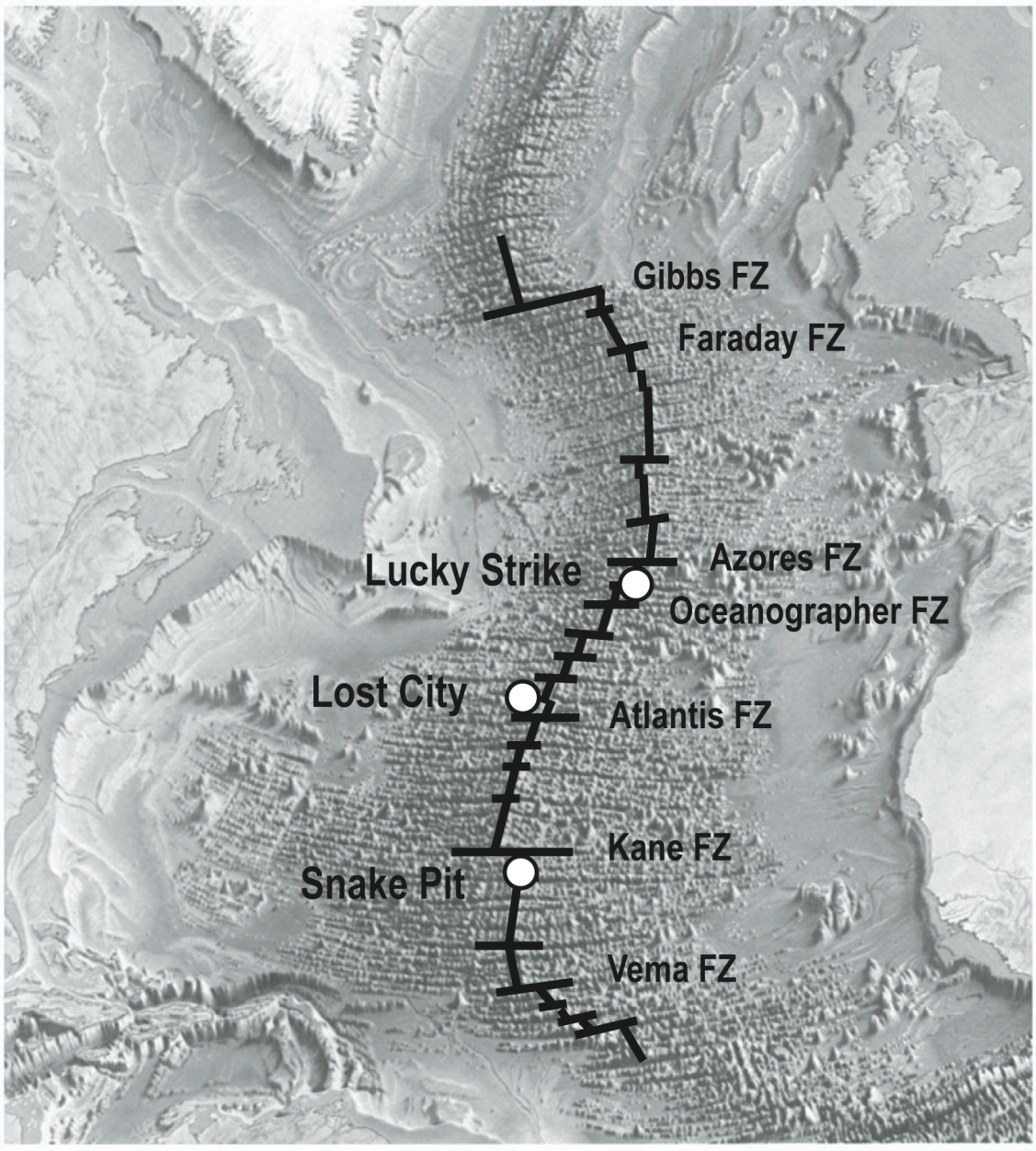
631

632 Figure 3. Phylogeny of chemoautotrophic and methanotrophic endosymbionts hosted by
633 bathymodioline mussels and free-living γ-Proteobacteria, inferred from 16S rRNA gene
634 sequences (1303 nucleotides). Posterior probabilities from 11,000 bayesian trees are
635 shown above branches (significant ≥ 95) and bootstrap values based on 1000 maximum
636 parsimony replicates are given below the branches. The two phylotypes in this study (*B.*
637 *azoricus* - *B. puteoserpentis*) are boxed in gray and lettered (C and M) for
638 chemoautotrophs and methanotrophs, respectively. Furthermore, the two clades that

include mussel symbionts are boxed and labelled (C and M). All symbiotic bacteria are labelled 'symbiont' while free-living bacteria are designated by taxonomic name alone.

Figure 4. Parsimony network inferred from 1200 nucleotides of rrn internal transcribed spacer (ITS rRNA) genotypes from chemoautotrophic symbionts of *Bathymodiolus* mussels collected from the Lost City, Lucky Strike, and Snake Pit hydrothermal fields. ITS-genotypes are shown as circles, with size indicating relative frequency. Shading denotes location and the distribution of genotypes within the host mussel population as follows: Lucky Strike (black = genotypes found in >1 host individual, black checkerboard = genotypes restricted to only one host individual), Lost City (gray, no symbionts were shared between the two host individuals), and Snake Pit (white = genotypes found in >1 host individual, gray checkerboard = genotypes restricted to only one host individual). Lines connecting genotypes are one nucleotide difference. Small black dots represent unsampled, hypothetical ancestors. LC1 and LC2 designate genotype clades from the two individual Lost City mussels. Finally, the northern and southern 'clades' are boxed and labelled N and S, respectively.

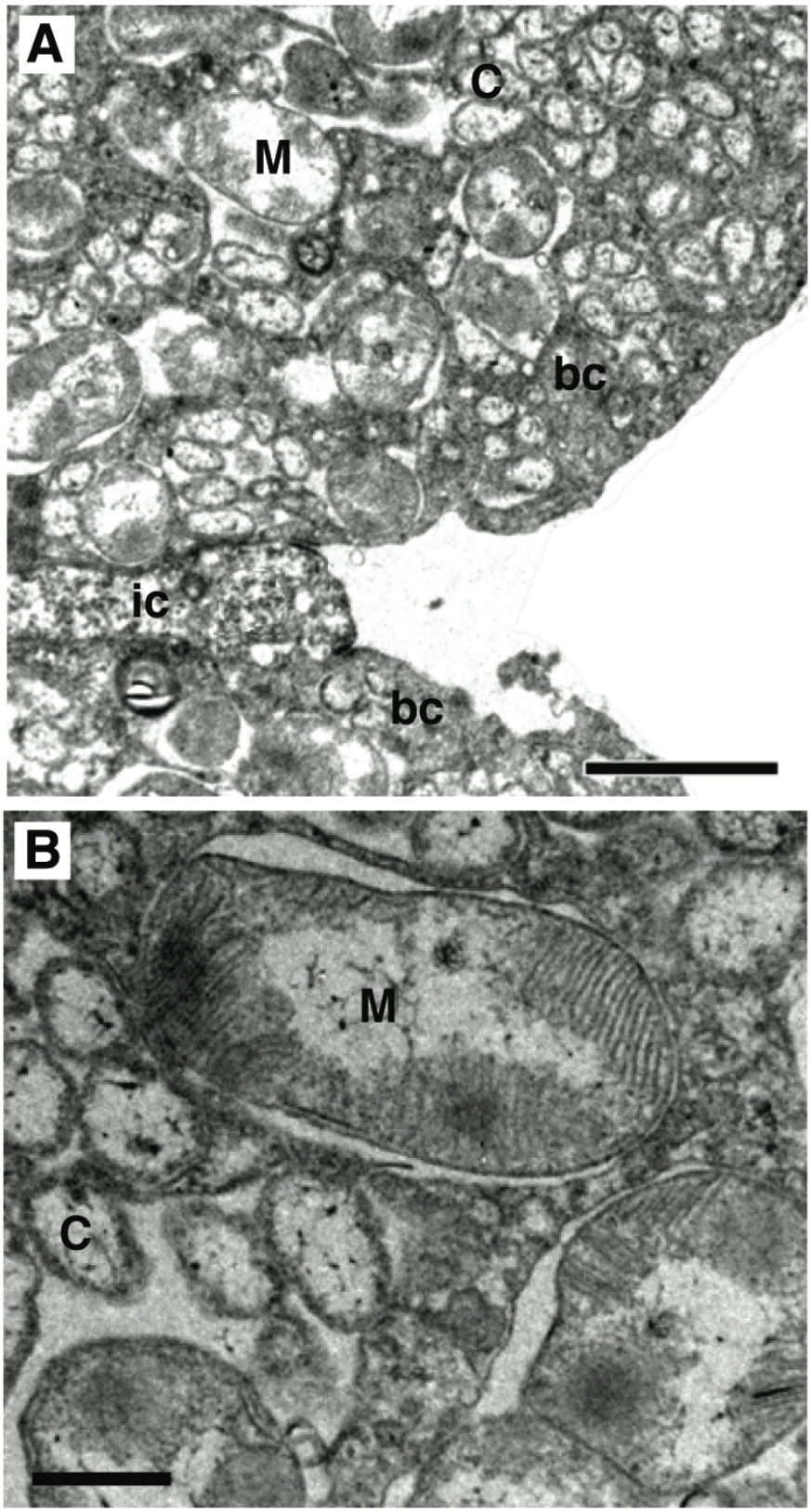
655 Figure 1.



656

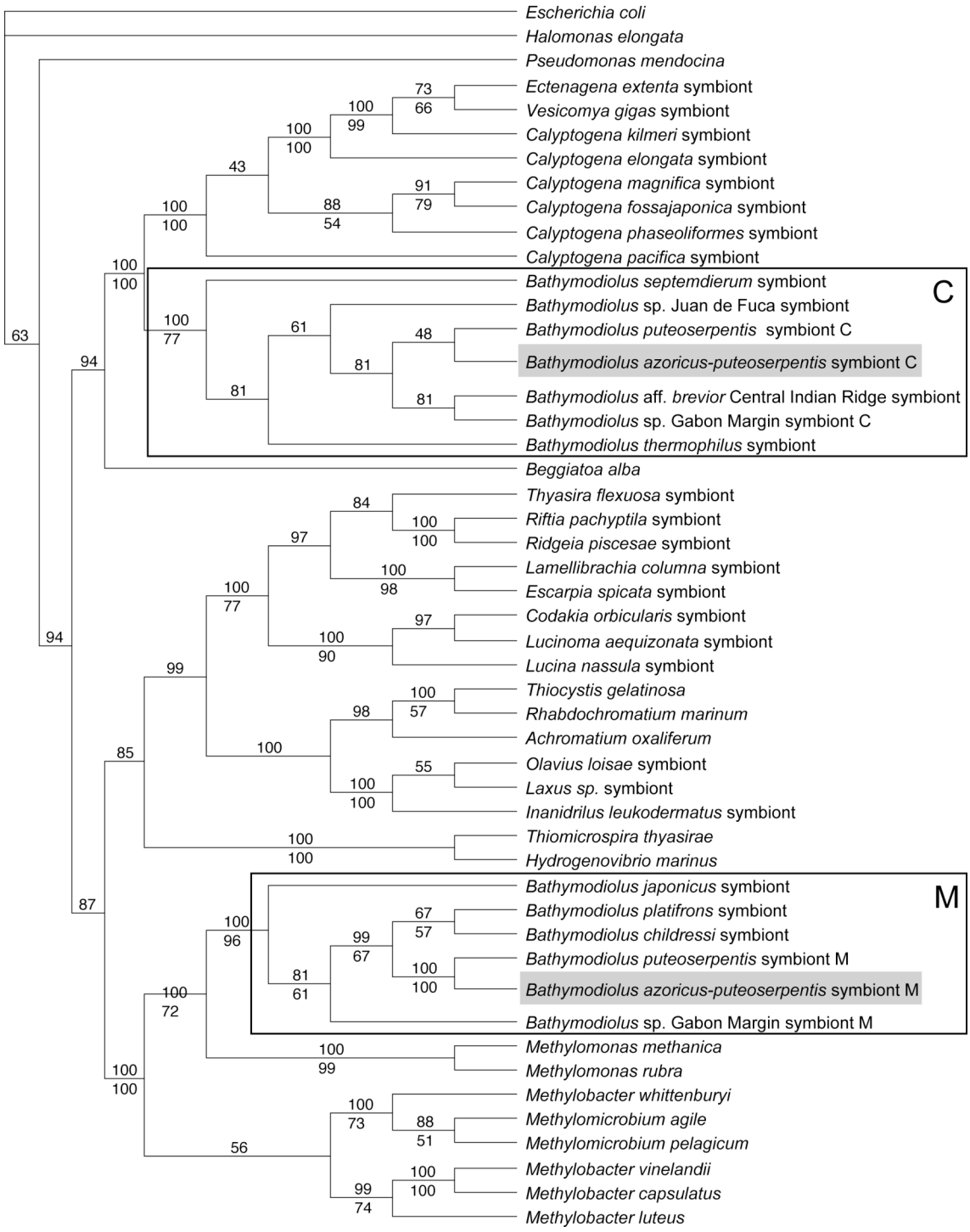
657

657 Figure 2.



658

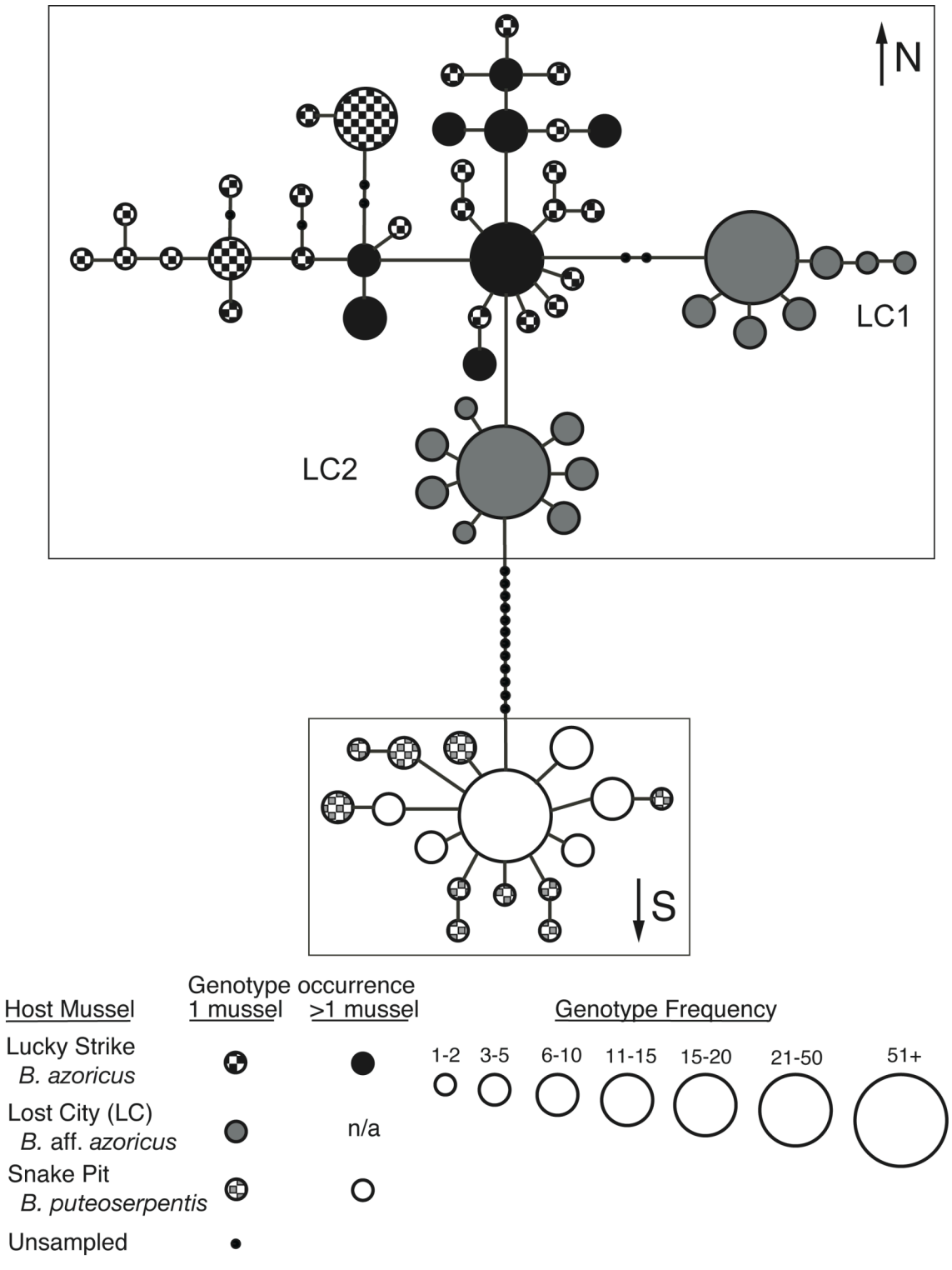
659 Figure 3.



660

661

661 Figure 4.



662

663