The ISME Journal (2012) 6, 352–362

www.nature.com/ismej

ORIGINAL ARTICLE

Phylogeny and fitness of *Vibrio fischeri* from the light organs of *Euprymna scolopes* in two Oahu, Hawaii populations

Michael S Wollenberg and Edward G Ruby
Department of Medical Microbiology and Immunology, University of Wisconsin-Madison, Madison, WI, USA

The evolutionary relationship among *Vibrio fischeri* isolates obtained from the light organs of *Euprymna scolopes* collected around Oahu, Hawaii, were examined in this study. Phylogenetic reconstructions based on a concatenation of fragments of four housekeeping loci (*recA*, *mdh*, *katA*, *pyrC*) identified one monophyletic group ('Group-A') of *V. fischeri* from Oahu. Group-A *V. fischeri* strains could also be identified by a single DNA fingerprint type. *V. fischeri* strains with this fingerprint type had been observed to be at a significantly higher abundance than other strains in the light organs of adult squid collected from Maunalua Bay, Oahu, in 2005. We hypothesized that these previous observations might be related to a growth/survival advantage of the Group-A strains in the Maunalua Bay environments. Competition experiments between Group-A strains and non-Group-A strains demonstrated an advantage of the former in colonizing juvenile Maunalua Bay hosts. Growth and survival assays in Maunalua Bay seawater microcosms revealed a reduced fitness of Group-A strains relative to non-Group-A strains. From these results, we hypothesize that there may exist trade-offs between growth in the light organ and in seawater environments for local *V. fischeri* strains from Oahu. Alternatively, Group-A *V. fischeri* may represent an example of rapid, evolutionarily significant, specialization of a horizontally transmitted symbiont to a local host population.
The ISME Journal (2012) **6,** 352–362; doi:10.1038/ismej.2011.92; published online 21 July 2011
Subject Category: microbe-microbe and microbe-host interactions
Keywords: biogeography; intraspecific; mutualism; phylogenetics; symbiosis

Introduction

The bioluminescent, heterotrophic marine bacterium *Vibrio fischeri* (Gammaproteobacteria: *Vibrionaceae*) maintains a mutualistic, binary relationship with bobtail squids of the genera *Euprymna*, *Sepiola* and *Rondeletiola* (Cephalopoda: Sepiolidae), and fishes of the genera *Monocentris* and *Cleiodopus* (Actinopterygii: Beryciformes). The host animals are hypothesized to receive a behavioral benefit from bacterial bioluminescence, whereas the bacteria are hypothesized to receive a metabolic benefit from the host (Stabb and Millikan, 2009). The Hawaiian bobtail squid, *Euprymna scolopes*, in particular, has been used for more than two decades as a model to study host–microbe interactions (McFall-Ngai and Ruby, 1991; Visick and Ruby, 2006). One characteristic of the squid/*Vibrio* model is the ease of laboratory cultivation of the host. Adult *E. scolopes* will mate in aquaria and lay clutches of eggs that produce hundreds of hatchlings. Another characteristic is the potential to manipulate symbiosis initiation in these newly hatched juveniles. *V. fischeri* are horizontally transmitted between squid generations and juvenile squid emerge from eggs uncolonized by *V. fischeri* (Wei and Young, 1989). As a result of this mode of transmission, *V. fischeri* populations are known to exist both in a specific organ in the host (the light organ) as well as the seawater environment (Lee and Ruby, 1992).

We recently analyzed the population structure of mutualistic *V. fischeri* collected from the light organs of adult *E. scolopes* from two populations on Oahu, Hawaii (Wollenberg and Ruby, 2009). Host light organs were found to contain polyclonal symbiont populations—population structure was assessed with a combination of physiological and genetic assays, including a novel *V. fischeri* DNA-fingerprinting technique (VfRep-PCR). Interestingly, strains identified by one fingerprint type (VfRep-PCR type-I) were found to be significantly more abundant than strains identified by other fingerprint types in the light organs of hosts collected from a specific environment (Maunalua Bay).

This former study was confined to describing patterns of *V. fischeri* population structure in the

Correspondence: EG Ruby, Department of Medical Microbiology and Immunology, University of Wisconsin-Madison, Madison, WI 53706-1521, USA.
E-mail: egruby@wisc.edu
Received 28 March 2011; revised 6 June 2011; accepted 6 June 2011; published online 21 July 2011

light organs of particular host populations. Here, we bridge the gap between the observation of a single, highly abundant fingerprint type in the Maunalua Bay host population and hypothesis testing of ideas about both the evolutionary history and the contemporary fitness of this group in different niches. The study described below consisted of two distinct, but complementary analyses. First, a phylogenetic scheme was used to reconstruct the evolutionary relationship among *V. fischeri* strains with distinct VfRep-PCR types. Second, different fitness assays were performed on representatives of these types in two Maunalua Bay environments: the squid host and natural seawater. The phylogenetic scheme and fitness assays were used to test the following hypotheses:

(i) Strains with distinct VfRep-PCR types also form distinct (and/or monophyletic) clades within an evolutionary reconstruction of *V. fischeri*, and
(ii) Strains identified as VfRep-PCR type-I are more fit in both the Maunalua Bay host and seawater environments than other VfRep-PCR types.

Materials and methods

V. fischeri *strains*

V. fischeri strains were selected for use in this study to satisfy two goals: (1) The strains represent a diverse group of VfRep-PCR types based on our previous analysis (Wollenberg and Ruby, 2009) of symbiont population diversity on Oahu; and (2) the strains represent an ecologically, temporally and geographically broad sampling of *V. fischeri*. A core group of 45 *V. fischeri* strains was selected from our laboratory collection (Figure 1 and Supplementary Table S1). Two closely related outgroup strains, *V. salmonicida* LFI1238 and *V. logei* SA6, were used to root phylogenetic reconstructions and provide an outgroup when required by common population genetics statistics.

Molecular phylogenetic analyses

A multi-locus sequence analysis scheme was used to organize the molecular data for the strains used in this study. A previously published *V. fischeri* scheme (Mandel *et al.*, 2009) that included three housekeeping loci (*recA*, *mdh* and *katA*) was expanded by adding one additional locus (*pyrC*); two loci involved with bioluminescence (*ainS*, *luxA*) were analyzed for comparative purposes. New sequences from this analysis have been submitted to the GenBank database under accession numbers JF509753–JF509942. Specific information on the multi-locus sequence analysis scheme is summarized in the Supplementary Information and Supplementary Table S2.

Multiple sequence alignments for each collection of locus fragments were generated by using the MUSCLE (Edgar, 2004) algorithm as implemented in SeaView 4.2.8 (Gouy *et al.*, 2010). Similarly to previous descriptions (Mandel *et al.*, 2009; Bose *et al.*, 2011), phylogenetic reconstructions assuming a tree-like topology were created by using three common methods: maximum parsimony; maximum likelihood (ML) and Bayesian inference (Bayes). Specifics of these methods, as well as descriptive analyses of molecular sequence data, are detailed in the Supplementary Information and Supplementary Table S3.

Phylogenetic reconstructions not assuming a tree-like topology were conducted by using the programs SplitsTree 4.10 (Huson and Bryant, 2006) to highlight and clarify inconclusive recombination results,

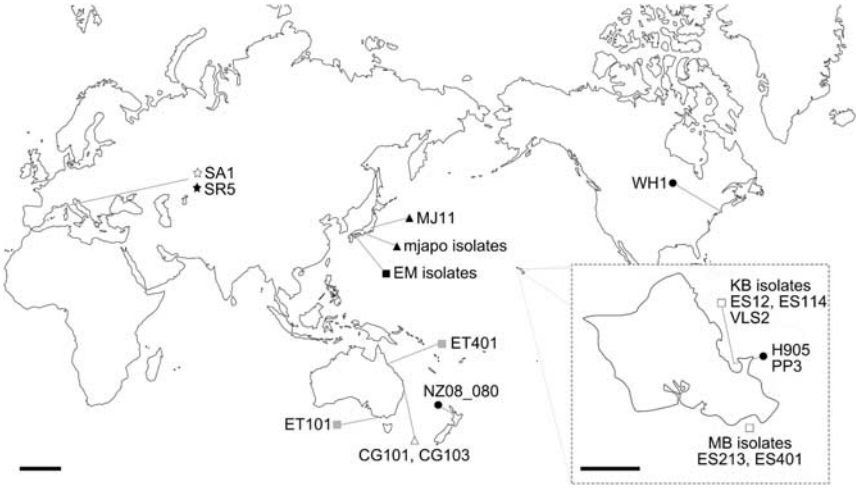

Figure 1 Geographic distribution of the *V. fischeri* strains used in this study. The majority of these strains were sampled from locations in the Pacific Ocean, specifically the Hawaiian Island of Oahu (inset). The shapes next to strain names denote the strains' ecological habitat: Filled black circle, planktonic (seawater); triangle or square or star, light-organ symbiont. The triangle, square and star fill colors represent host organisms as follows: White triangle, *C. gloriamarus*; black triangle, *M. japonicus*; white square, *E. scolopes*; gray square, *E. tasmanica*; black square, *E. morsei*; white star, *S. affinis*; black star, *S. robusta*. The bars in the main figure and inset represent 2000 and 20 km, respectively. The background continent and Oahu outlines were taken from maps made available freely at http://d-maps.com/.

and ClonalFrame 1.1 (Didelot and Falush, 2007) to address conflicting signals among the different gene fragments. For each concatenated data set, three independent runs of ClonalFrame were performed; in each run, the 50% majority rule consensus genealogy was estimated from the posterior distribution of 200 000 generations (thinning interval of 100) following a burn-in of 100 000 generations. Markov Chain Monte Carlo convergence was judged to be satisfactory by using the Gelman-Rubin test (Gelman and Rubin, 1992). The resulting posterior distribution of 6000 samples (three runs of 2000 samples each) from each concatenated data set was analyzed by using a consensus network (Holland *et al.*, 2004) with mean edge weights at a threshold of 0.2 and an equal-angle splits transformation as implemented in SplitsTree. Nodes with 95% posterior probability (from the 6000 original ClonalFrame samples) were identified in each network graph.

Juvenile squid experiments

Adult *E. scolopes* were collected from Maunalua Bay (MB), Oahu, Hawaii, and transported to aquaria in Madison, Wisconsin. Matings between MB adults and subsequent egg laying, development and hatching took place in these same aquaria. Hatchlings from MB adult matings were immediately moved into bowls containing 50–200 ml of SFTW (0.2-μm sterile-filtered tank water) taken from an aquarium. All aquarium water was made from Instant Ocean (Spectrum Brands, Madison, WI, USA). For each experiment, approximately equal concentrations of two *V. fischeri* strains from different VfRep-type groups were inoculated into bowls containing 50–200 ml of SFTW; dilutions of this mixture were spread on SWT (seawater–tryptone) agar plates (Boettcher and Ruby, 1990) to assess the initial concentrations of bacteria. Newly hatched juveniles ($n = 5$–20) were added to a bowl; after either 3 or 18 h, the juveniles were moved to individual glass vials containing 3 ml of fresh SFTW. The animals were kept in these vials for an additional 30–45 h, with a water change at approximately 20 h. Next, juvenile squid were anesthetized on wet ice, frozen at −80 °C and individually resuspended and homogenized in SFTW. Dilutions from this homogenate were then spread on SWT agar plates for enumeration of the relative proportions of the two strains by using their distinctive yellow or white colony pigmentations (Wollenberg and Ruby, 2009). Each competition experiment was evaluated by using the logarithm of the relative competitive index (RCI) to the base 10 (Stabb and Ruby, 2003). If the expression X_Y is defined as the concentration of bacterial strain X after time Y, the RCI of Strain-A versus B after 48 h in the host is defined by the following equation: $RCI_{AB} = [(A_{48}/B_{48})/(A_0/B_0)]$.

For each single-strain colonization experiment, approximately 1000 colony-forming units (c.f.u.) ml^{-1} of a single *V. fischeri* strain were inoculated into a bowl containing SFTW. Approximately 2.5 ml of this SFTW/*V. fischeri* mixture were added to each of 45 separate glass scintillation vials. Forty-five newly hatched juveniles were added individually to each of the vials. The animals were kept in these vials for an additional 24–72 h, with a water change at approximately 24 h and 48 h. At 24 h, 48 h and 72 h after addition, 15 animals were anesthetized on wet ice for 5 min, frozen, homogenized and plated on SWT agar as described above. Data were summarized by recording the c.f.u./squid for each successful colonization event at each time point (per 15 squid). Because we are more concerned with the means of these data, and less with the variation among the individual data points, we report 95% confidence interval and s.e.m., rather than standard deviation. Single-strain colonization experiments were performed with six strains (ES213, ES114, MB11B1, KB2B1, MB14A3 and MB15A4) representing the three major Hawaiian clades found in phylogenetic reconstructions (Figure 2 and Supplementary Figure S2).

For all squid experiments, aposymbiotic juveniles were placed in SFTW as a negative control. Briefly, 5–10 newly hatched juvenile squid were placed in a

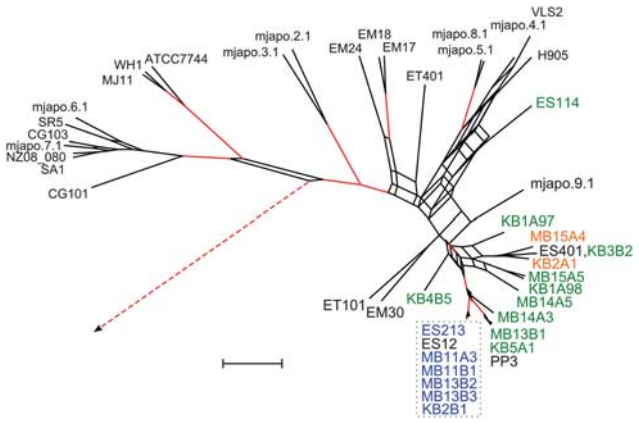

Figure 2 *V. fischeri* VfRep-PCR type-I strains form a monophyletic group within a phylogenetic reconstruction of global *V. fischeri*. A consensus network inferred from 6000 phylograms produced by ClonalFrame analyses of all taxa using a concatenation of four loci: *recA*, *mdh*, *katA* and *pyrC*. ClonalFrame was used to infer genealogical relationships from posterior distributions of 2000 samples from three independent runs ($3 \times 2000 = 6000$ total samples). This aggregate ClonalFrame posterior distribution was used by SplitsTree to build a consensus network. For this network, reticulate relationships found in at least 20% of the sample may be visualized by parallel branches in parallelograms; nodes supported by a posterior probability greater than 95% (as summarized by ClonalFrame) are at the terminus of red-colored branches. The root (dashed red branch with arrow) has been truncated to expand the branches of the main *V. fischeri* group; the root contains both outgroup taxa: *V. salmonicida* LFI1238 and *V. logei* SA6. VfRep-PCR patterns have been mapped onto *E. scolopes* symbionts with colors: identically colored strains all share the same VfRep-PCR pattern (or type) (Supplementary Table S1): Blue, type-I (boxed); orange, type-II; green, type-III. The strains within the dotted box have been designated as 'Group-A'. The black bar is equivalent to a branch length of 0.1 substitutions per site.

bowl containing only SFTW immediately after hatching. After 72 h, the squid were sacrificed as above. In all experiments, none of these squid were colonized by more than 28 c.f.u.s of *V. fischeri* per squid (the lowest limit of detection assayed; data not shown).

Hawaiian seawater experiments

Six *V. fischeri* strains carrying the kanamycin-resistant (Kan^R), green fluorescent protein-expressing plasmid pVSV102 were constructed by triparental mating as described previously (Wollenberg and Ruby, 2009). These six strains (ES213, ES114, MB11B1, KB2B1, MB14A3 and MB15A4) represented the three major Hawaiian clades found in phylogenetic reconstructions (Figure 2 and Supplementary Figure S2). It was necessary to create Kan^R, green fluorescent protein strains because of the high background of naturally occurring, SWT-culturable bacteria in fresh MB seawater (see Supplementary Information for a more detailed description).

For growth experiments, the six labeled strains constructed above were individually grown in fresh, MB seawater-containing microcosms (Lee and Ruby, 1994). Briefly, a frozen stock culture of each strain was streaked onto three separate SWT agar plates containing 100 μg ml^{-1} kanamycin and grown for 24 h at 22–24 °C. A colony scrape from each of these plates was diluted in 0.22-μm-filtered fresh MB seawater to a concentration of $\sim 10^3$ c.f.u. ml^{-1}. These three (biological) replicates were performed for statistical purposes. Each diluted seawater culture was used as the primary inoculum of two experimental microcosms containing 3 ml of either 0.22-μm-filtered or unfiltered MB seawater. The microcosms were shaken at 28 °C in an enclosed orbital shaker (New Brunswick Scientific, Edison, NJ, USA) at 225 r.p.m. for 96 h with periodic sampling and plating on SWT agar plates containing 100 μg ml^{-1} kanamycin. Colony abundance data were transformed into concentration data and analyzed by two-way, repeated-measures analysis of variance. The two factors of interest were the strains' group membership and time in hours (and their interaction). Correlation of repeated measures over time was accounted for by using random effects for strains (nested in groups) and for replicates (nested in strains). Analyses were conducted by using the statistical software R (R_Development_Core_Team, 2010) and the following statistical model: CFU/ml = Group * Hours + Error (Strain/Replicate). Statistical significance was evaluated with F-tests.

Results

Estimation of evolutionary parameters and ML gene phylogenies from individual locus fragments

We compared sequence alignment characteristics and ML-estimated evolutionary parameters for each of the six locus fragments from the 45 *V. fischeri* strains analyzed in this study. The two locus fragments associated with quorum signaling and/or bioluminescence (*ainS*, *luxA*) contained nearly twice the percentage of variable and parsimony-informative sites, per locus, when compared with the other four loci (Supplementary Table S3). For each locus fragment, the substitution parameters were complex, with four or more rate parameters as estimated by hierarchical likelihood ratio tests. Accounting for among-site rate variation significantly improved the model fit for each locus fragment ($P < 0.001$) (Supplementary Table S3).

Single-locus, ML phylogenetic reconstructions of all strains demonstrated that few general similarities or statistically significant nodes were common among tree topologies (Supplementary Figure S1). Notably, the reconstruction based on the *luxA* fragment contained nearly twice as many statistically supported nodes in comparison to each of the reconstructions based on the other five locus fragments. Upon inspection, this relatively high resolution could be explained by two factors: (i) 23 different *luxA* alleles exist among the strains, the second-lowest number of alleles among the six loci analyzed; and (ii) among these 23 alleles, the *luxA* fragments analyzed contain the lowest percentage of constant sites of any of the six loci analyzed. In other words, the *luxA* tree appears to have been well-resolved relative to the other five loci because its reconstruction was based on few alleles containing many nucleotide differences (Supplementary Table S3).

To further understand patterns of nucleotide polymorphism among these loci, more detailed test statistics were calculated (Supplementary Table S4). For all of these metrics, *luxA* stood out from the other five loci. First, the *luxA* locus fragment has a much larger number of total mutations than segregating sites, meaning that this locus generally has more than one polymorphism at each polymorphic site. Next, the average number of pairwise difference (k) and nucleotide diversity (π) metrics have their highest values for *luxA*. Finally, estimations of both recombination and mutation parameters by various methods demonstrated that *luxA* has the lowest ratio of recombination rate to mutation rate (c/μ) of all loci, reflecting both a low recombination parameter and a high mutation parameter relative to the other five loci (Supplementary Table S4). Together, these metrics indicate that, of the six locus fragments analyzed for these *V. fischeri* strains, the *luxA* fragment has accumulated a relatively large number of polymorphisms through mutations, but has also been simultaneously constrained to a small number of absolute alleles. One hypothesis for this observation is that selection has been acting on *luxA* within the *V. fischeri* metapopulation. To evaluate this hypothesis, tests of the neutral equilibrium model of molecular evolution (Kimura, 1983) were conducted.

When patterns of nucleotide substitutions for these six loci were analyzed under a neutral equilibrium model, this model was strongly rejected for *luxA* and weakly rejected for *ainS* (Table 1, Tajima's *D*). Because both selection and/or population demographics may explain statistically significant rejections of the null model for any single test (Tajima, 1989), multiple tests were employed to arrive at a clearer idea of the underlying processes responsible for the observed polymorphism in the data. Test statistics of Tajima's *D* were significantly positive for both *ainS* and *luxA*, indicating an excess of polymorphisms relative to polymorphic sites, as implied by $\varnothing_\pi > \varnothing_W$ (Table 1). A significantly positive value of *D* is indicative of balancing selection on a particular locus and/or a population bottleneck. Tajima's *D* was calculated for each sequence by using both nucleotide and amino-acid sequences—the latter, as noted (Vos and Velicer, 2006), may be more relevant in assessing evidence for selection because only non-synonymous polymorphisms are considered.

Other test statistics, such as Fu and Li's *D* and *F*, Fu's *Fs* and Ramos-Onsis and Rozas's R_2 are more sensitive to population demographics than Tajima's *D*, and are often used to infer the role that demographic processes may have had in a given data set's null-model rejection (Simonsen *et al.*, 1995). Using these tests, there was ample evidence that departure from neutral expectations by *luxA* may be due to demographic event(s); that is, these significantly large, positive test statistics imply that a population bottleneck may be responsible for the infrequency of polymorphic sites relative to the polymorphisms in the *luxA* data set. Interestingly, in the *ainS* data set, only Fu and Li's *F* is significant, implying that demography may have less of a role in any significant violations of the neutral equilibrium model for this locus (Table 1).

Intragenic and intergenic recombination within and among loci

Evidence for intragenic recombination within each locus was analyzed by using a suite of algorithms included in the RDP software package; recombination events were only detected within the 783-bp *mdh* locus fragments. This result was mitigated by truncating the 3′-end of this fragment to remove the sequences showing evidence of intragenic recombination, resulting in a 519-bp *mdh* locus fragment (data not shown). In addition, although there were no well-supported intragenic recombination events characterized for *katA* and *pyrC*, these two fragments had a relatively high ratio of recombination rate to mutation rate (Supplementary Table S4), suggesting the presence of conflicting evidence both for and against the role of intragenic recombination within these loci.

The possibility of intergenic recombination among *V. fischeri* strains was qualitatively supported by the lack of similarity among nodes for each single-locus ML reconstruction (Supplementary Figure S1). Topological incongruence among single-locus trees is one of the proposed hallmarks of intergenic recombination (Dykhuizen and Green, 1991). The hypothesis of topological incongruence among individual locus-fragment reconstructions was found to be statistically well-supported and ubiquitous under both ML and maximum parsimony frameworks (Supplementary Table S5).

Another method for assessing evidence for intergenic recombination is the calculation of several different summary statistics of recombination, the most common being the standardized index of association (I_A^S) and the global ratio of recombination rate to point mutation rate for all loci (r/m). For the six loci studied, the START2 program calculated $I_A^S = 0.3218$ and rejected the null hypothesis of linkage equilibrium ($V_O = 0.7096$, $V_e = 0.272$,

Table 1 Selection and population growth test statistics from six loci of 45 different *V. fischeri* strains

Locus fragment	ω[a]	*Tajima's* D[b,c]			*Fu and Li's* D[c,d]	*Fu and Li's* F[c,d]	*Fu's* Fs[c,d]	R_2[e]
		S	η	AA[f]				
recA	0.0122 (0.0020, 0.0376)	−0.744	−0.860	−1.475	−1.201	−1.311	−3.578	0.083
mdh	0.0115 (0.0036, 0.0268)	0.828	0.308	−0.955	0.219	0.561	−1.156	0.137
ainS	0.1799 (0.1488, 0.2152)	2.423 (0.005)	1.736 (0.044)	2.021	0.755	1.781 (0.022)	−0.317	0.186
katA	0.0402 (0.0221, 0.0661)	0.701	0.410	0.530	0.021	0.363	−0.244	0.131
pyrC	0.0263 (0.0136, 0.0450)	0.004	−0.221	−1.476	0.267	0.228	−3.094	0.109
luxA	0.0642 (0.0525, 0.0776)	3.034 (<0.001)	1.706 (0.020)	2.890 (<0.001)	1.964 (<0.001)	2.979 (<0.001)	19.610 (<0.001)	0.207 (<0.001)

[a]K_a/K_s or d_N/d_S: Ratio of the number of non-synonymous substitutions per non-synonymous site to the number of synonymous substitutions per synonymous site, calculated using the SLAC method (95% confidence interval in parentheses).
[b]Tajima's *D* test statistic calculated using the nucleotide sequence and the total number of segregating sites (*S*), total mutations (η) or the amino-acid sequence and the total number of segregating sites (*AA*).
[c]Statistical significance of test statistic (given in parentheses if $P<0.05$) was calculated from 10^3 neutral coalescence simulations by using segregating sites with an intermediate level of recombination.
[d]Fu and Li's *D* and *F*, and Fu's *Fs* test statistics were calculated by using total segregating sites.
[e]Population growth test statistic of Ramos-Onsis and Rozas (Ramos-Onsins and Rozas, 2002); the test statistic distribution and statistical significance were calculated from 10^3 neutral coalescence simulations by using segregating sites with an intermediate level of recombination. For the *luxA* fragment, the test statistic is significantly larger than the distribution calculated by the coalescent at a level of $P<0.001$.
[f]Statistical significance of test statistic was estimated from the β distribution.

$P<0.001$) implying a relatively clonal population structure for the 45 *V. fischeri* strains analyzed. For a concatenation of the six loci, *r/m* was calculated as 1.46 (95% confidence interval = 1.06–1.90); this value is most likely an underestimation of the summary statistic because the sequence collection itself represents a global sample that may include multiple local subpopulations. A value of 1.46 is typical of a free-living bacterial species (Vos and Didelot, 2009); both of these recombination metrics support the hypothesis that recombination has an important, yet moderate, role in the generation of diversity among *V. fischeri*.

Phylogenetic reconstruction of concatenated loci and comparison with VfRep-PCR data

To assess the evolutionary significance of the *V. fischeri*-specific VfRep-PCR fingerprinting data, a multi-locus sequence analysis approach was used to analyze the phylogenetic relationships among *V. fischeri*. Multi-locus analyses have the advantage of combining the weak phylogenetic signal present in multiple housekeeping (or other) loci into a significantly well-resolved reconstruction through the aggregation of phylogenetic signal using informative sites (Hanage *et al.*, 2006). One important caveat to this approach is that loci with a relatively large number of polymorphisms have the potential to contribute a disproportionate amount of information to the final reconstruction. For the *V. fischeri* data analyzed above, single-locus analyses demonstrated that one of the locus fragments, *luxA*, had both a disproportionate amount of variable sites and strong violations of neutral theory relative to the other five loci. Furthermore, the other bioluminescence-related locus, *ainS*, had similar, but slightly less serious, shortcomings (Table 1 and Supplementary Table S4). Therefore, phylogenetic reconstructions of a concatenation made with only the four 'housekeeping' loci (*recAmdhkatApyrC*) (Figure 2 and Supplementary Figure S2) were overlaid with VfRep-PCR type data (Supplementary Table S1).

For all reconstructions, *E. scolopes* symbionts formed a well-supported, but polyphyletic clade. Within this clade, VfRep-PCR type-I strains all formed a well-supported, monophyletic clade, 'Group-A', that was distal to all other *V. fischeri* taxa (Figure 2 and Supplementary Figure S2). VfRep-PCR type-II and III strains formed a well-supported, paraphyletic clade ('non-Group-A') containing the more distal Group-A clade. The Group-A clade contained not only contemporary symbionts isolated in 2005, but also two strains (ES12, ES213) isolated from the light organs of *E. scolopes* collected in Maunalua Bay nearly 20 years earlier (Lee, 1994).

Fitness assays of Group-A strains compared with other Hawaiian V. fischeri *strains*

In a previous study (Wollenberg and Ruby, 2009), MLST Group-A (that is, VfRep-PCR type-I) strains dominated the symbiont populations in the *E. scolopes* hosts collected from Maunalua Bay, Hawaii, in 2005. With the observations of the relative homogeneity of the Group-A *V. fischeri* strains, and their abundance relative to other Hawaiian *V. fischeri* strains, we hypothesized that these strains were more fit in both the free-living and the host-associated environments encountered in Maunalua Bay, Hawaii. To test these hypotheses, growth and survival assays were performed both in MB *E. scolopes* hosts and in fresh MB seawater.

Competition assays in juvenile MB hosts demonstrated that *V. fischeri* MLST Group-A strains showed a colonization advantage over the non-Group-A strains (Figure 3). When different Group-A strains were individually competed against one of two different non-Group-A strains for 48 h, the Group-A strains dominated the majority of competition experiments. In all experiments, Group-A strains were also found to be colonizing as an apparent monoculture in a disproportionately high percentage of juveniles after 48 h of competition. Variables such as time of initial inoculation, concentration of initial inoculant and strain identity did not affect these results qualitatively (Figure 3). Additionally, another non-Group-A strain, ES114, showed similar competition results against Group-A strains after both 3 and 18-h inoculation periods (data not shown).

One possible explanation for this observation of competitive advantage of Group-A strains over non-Group-A strains in MB juveniles is an improved ability of the Group-A strains to initiate symbiosis with MB juveniles. To test this hypothesis, single-strain colonization experiments of juvenile *E. scolopes* from MB were performed (Figure 4). The results of these experiments demonstrated that individual strains varied in their abundance in MB juveniles at 24 h, 48 h and 72 h after infection, and that there was no statistically significant difference between the mean abundance of Group-A strains and non-Group-A strains at these time points ($F_{1,4} = 0.258$, $P = 0.64$). These data are not surprising, as the ability to discern patterns of difference in colonization fitness between strains often requires co-colonization experiments (Stabb and Ruby, 2003).

Incubation for 4 days in fresh, unfiltered MB seawater revealed significantly different growth and survival responses of Group-A *V. fischeri* relative to non-Group-A *V. fischeri* (Figure 5). Seawater growth/survival data from filtered seawater showed no qualitative or quantitative difference between Group-A and non-Group-A *V. fischeri* strains. For all strains, abundance increased 2- to 10-fold for the first 12 h, followed by a 10-fold decrease by 96 h. By contrast, in unfiltered seawater, Group-A strains showed a more rapid decrease in abundance after 12 h than non-Group-A *V. fischeri* strains. Results of a two-way, repeated-measures analysis of variance included a significant interaction term between Group identity and time ($F_{5,80} = 10.368$, $P < 0.001$).

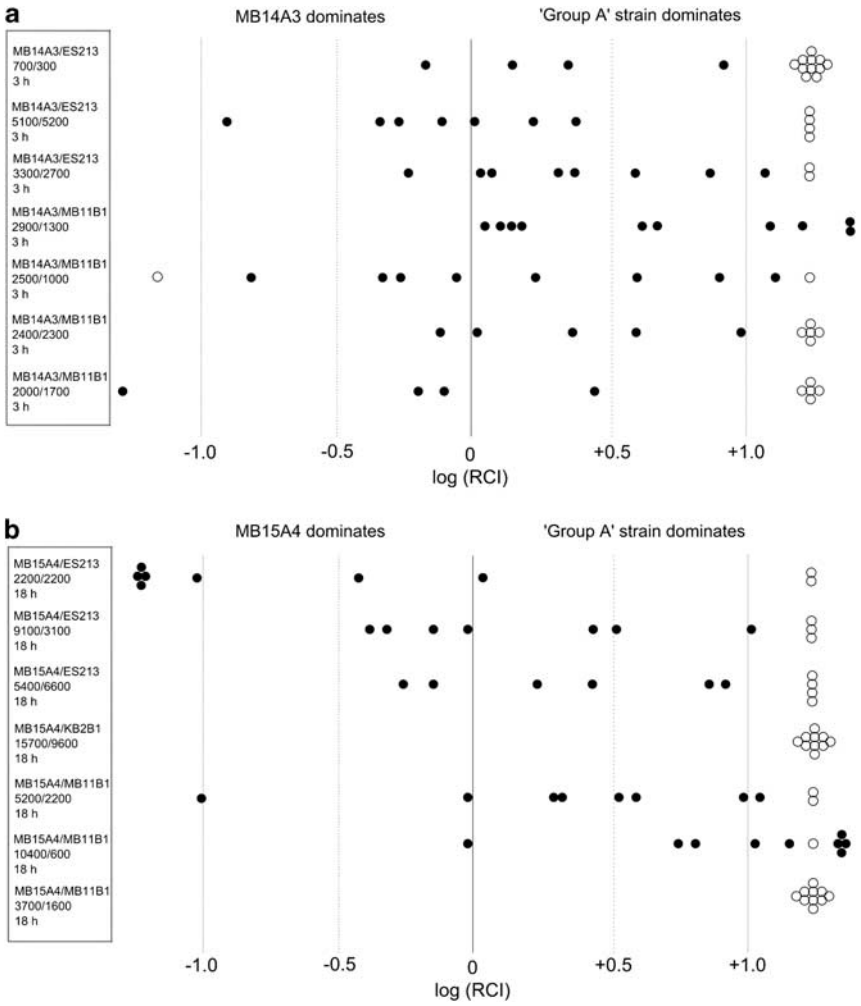

Figure 3 Dual-strain, juvenile *E. scolopes* experiments competing single *V. fischeri* 'Group-A' strains against either (**a**) MB14A3 or (**b**) MB15A4. In both graphs, each row represents a single experiment, summarized by the box on the left as follows: First line, strains' identities; second line, initial mean concentration of each strain in c.f.u./ml (the s.d. was less than 50% of this mean value for all experiments); third line, duration of inoculation. All experiments were concluded at 48 h after hatching and involved 9–15 juvenile squid born from mating adults collected from Maunalua Bay, Hawaii. The unfilled circles represent individual squid that were found to contain only one of the strains (that is, the other strain was below the limit of detection—approximately 500 times less abundant than the strain present). The unfilled circles are arbitrarily positioned at 1.25 and −1.25 units on each axis. Each filled circle represents an individual squid that contained both strains, and is positioned on the graph as the logarithm of the RCI (logRCI) for the two strains. Filled circles with a logRCI value greater than 1.25 or less than −1.25 are placed either to the right or to the left of these values (and unfilled circles), respectively, because of graphical limitations.

This test result was interpreted as follows: a significant difference exists between the mean abundance profile of the Group-A strains and the non-Group-A strains over the duration of the experiment. Qualitatively, this difference appears as a sharp decline after 48 h for the Group-A strains, compared with a slow decline for the other strains. In filtered MB seawater, results of two-way, repeated-measures analysis of variance included a non-significant interaction term between Group identity and time ($F_{5,80} = 1.497$, $P = 0.20$).

Discussion

This and another recent study (Wollenberg and Ruby, 2009) have focused on the population biology of symbiotic *V. fischeri* from a particular ecological and geographical setting: the Hawaiian bobtail squid *E. scolopes* from the island of Oahu. In the previous study, phenotypic assays and molecular fingerprinting revealed that the symbiotic populations of adult *E. scolopes* light organs are oligoclonal. Here, we advance these ecological observations with a study of these light-organ symbionts by using evolutionary inference and experimental fitness assays. We show here that a monophyletic group of *V. fischeri* (Group-A) is found disproportionately in hosts from Maunalua Bay, Oahu, Hawaii (MB). This new observation led to an additional question: Do members of Group-A, when compared with other symbiotic *V. fischeri* not in this group, demonstrate increased fitness in the Maunalua Bay environments?

Experiments using the offspring of MB squid suggest that *V. fischeri* strains from Group-A are more fit in the MB *E. scolopes* light-organ niche than the non-Group-A strains (Figure 3). These results demonstrate the competitive advantage of a collection of closely related, symbiotic *V. fischeri* at a previously unstudied scale: a local, intra-island distance (<20 km) between bays on Oahu, Hawaii. In the past, *V. fischeri* strains from squid populations separated by larger geographic distances (>2000 km) have been shown to exhibit a competitive advantage colonizing juvenile hosts from the strain's local environment (Nishiguchi *et al.*, 1998; Nishiguchi, 2002). Although the molecular determinants of host specificity are beginning to be reported in *V. fischeri* (Mandel, 2010), and there exists evidence to suggest that the response regulator *rscS* may have some role in determining the host range (that is, squid or fish) of *V. fischeri* symbionts (Mandel *et al.*, 2009), to date there has been no work defining what molecule(s) might mediate the competitive advantage of one strain of *V. fischeri* over another in the same host species at such a local geographic scale.

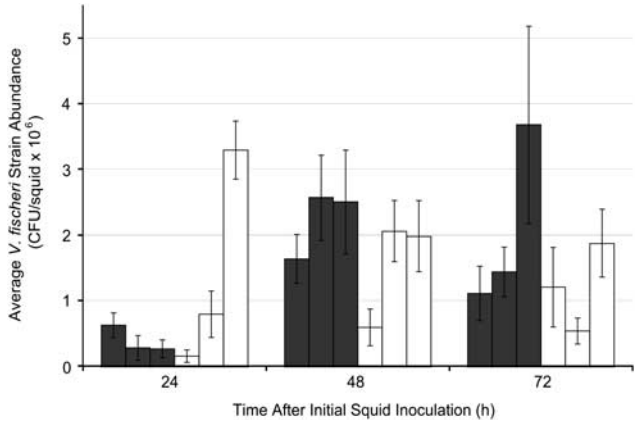

Figure 4 Single-strain *V. fischeri* inoculations of juvenile squid demonstrated no clear difference between the levels of symbiosis initiation by Group-A and non-Group-A *V. fischeri*. Each bar represents the mean c.f.u./squid of 10–15 squid inoculated with a single isolate of *V. fischeri* and sacrificed at 24, 48 or 72 h after inoculation. The error bars represent 95% confidence intervals. The histogram bars relate to *V. fischeri* strains as follows: Non-'Group-A' *V. fischeri* strains are represented by the three black bars (left, ES114; center, MB14A3; right, MB15A4), whereas 'Group-A' *V. fischeri* strains are represented by the three white bars (left, ES213; center, MB11B1; right, KB2B1).

The Group-A and non-Group-A strains showed equivalent growth and persistence in 0.2-μm-filtered MB seawater (Figure 5a), suggesting that the innate ability of different *V. fischeri* strains to survive starvation is comparable. As shown previously (Lee and Ruby, 1994), incubation of *V. fischeri* cells in unfiltered, fresh seawater led to decreased (culturable) cell counts (c.f.u.s) after the first 12–24 h of initial growth. Interestingly, in contrast to their competitive dominance in juvenile *E. scolopes*, the Group-A strains were at a statistically significant disadvantage in persisting in fresh, unfiltered MB seawater after 12 h when compared with the non-Group-A strains (Figure 5b). Taken together, results from these studies suggest a link between (biotic or abiotic) particulate factors in Maunalua Bay seawater and the abundance and survival of *V. fischeri* cells in the planktonic environment.

How might future studies make sense of both the phylogenetic and the contrasting growth/survival characteristics of *V. fischeri* strains from Oahu? One attractive conceptual framework for the evolution of the Group-A strains involves the hypothesis that host population biogeography has driven local symbiont evolution—in other words, *V. fischeri* evolution in response to selection by the local host population. Another (possibly mutually exclusive) working hypothesis consistent with these results is that the evolution of the Group-A strains has involved a more general trade-off between growth in the light-organ environment and survival in the seawater environment.

Limited evidence exists in support of the former hypothesis that host biogeography drives microbial symbiont evolution at an intra-island scale. First, molecular and phenotypic data suggest that the

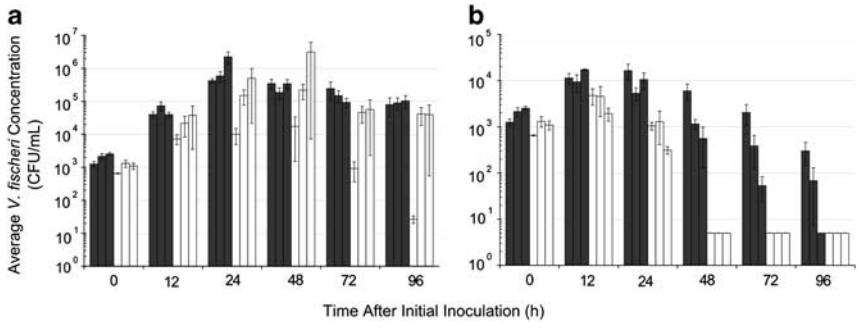

Figure 5 Single-strain *V. fischeri* growth experiments in microcosms filled with either (**a**) 0.22-μm-filtered or (**b**) unfiltered fresh seawater from Maunalua Bay, Hawaii. In both graphs, each bar represents the mean c.f.u./ml of three biological replicates of a single strain, with error bars indicating the s.e.m. The histogram bars relate to *V. fischeri* strains as follows: Non-'Group-A' *V. fischeri* are represented by the three black bars at each time point (left, ES114; center, MB14A3; right, MB15A4), whereas 'Group-A' *V. fischeri* are represented by the white bars at each time point (left, ES213; center, MB11B1; right, KB2B1). In the unfiltered seawater data (**b**), the mean abundance values that are indicated at 3 c.f.u. ml^{-1} for the 48- to 96-h time points were below this limit of detection.

E. scolopes population on Oahu shows a local structure between Maunalua and Kaneohe Bays (Kimbell *et al.*, 2002). Second, the total number of *V. fischeri* cells in the planktonic environment has been estimated to be several (Jones *et al.*, 2007) to many (Lee and Ruby, 1995) orders of magnitude lower than in the host light organs, suggesting that the seawater niche may have a relatively minor role in *V. fischeri* evolution. Finally, recent studies of *V. fischeri* populations in *E. scolopes* have demonstrated that the light-organ environment can show strong selective pressure on the physiology of *V. fischeri* (Schuster *et al.*, 2010). Taken together, this evidence suggests many key features related to the host *E. scolopes* (that is, local population structure; strong, selective environment; high symbiont population density) that might potentiate the evolution of local *V. fischeri* strains. Studies of other horizontal, 'binary' symbioses have argued that local host populations may have some influence on the local symbiont population structure and evolution (see, for example, Santos *et al.*, 2003; Sicard *et al.*, 2004; Cafaro *et al.*, 2010; Maneesakorn *et al.*, 2011), although other arguments also exist (for example, Martínez-Romero, 2009; Mueller *et al.*, 2010), even within the squid/*Vibrio* symbiosis (Nishiguchi *et al.*, 1998; Dunlap *et al.*, 2007). For instance, using different phylogenetic markers and approaches, the latter two citations arrive at opposing conclusions concerning the evidence for parallel cladogenesis between species of squid and their light-organ symbionts.

Alternatively, our results might best be addressed by the second hypothesis that general trade-offs exist between *V. fischeri* growth/survival in the two environments studied (seawater and the squid host). Phenotypic and/or genotypic trade-offs have been hypothesized to explain the life-history differences observed among or within evolutionary lineage(s) in two or more ecological niches (Stearns, 1992). Contemporary microbial ecology research has sometimes focused on the trade-offs microorganisms experience when grown in environments either sufficient or lacking in metabolites (Velicer and Lenski, 1999). In the case of *V. fischeri*, the *E. scolopes* light organ is thought to provide a dynamic, nutrient-rich habitat (Ruby *et al.*, 2005; Wier *et al.*, 2010) in contrast to the relatively oligotrophic planktonic environment (Giovannoni and Stingl, 2007). In ecological terms, then, symbiotic *V. fischeri* around Oahu might inhabit a heterogeneous environment composed of two physiologically distinct habitats: seawater and the host light organ.

Additional experiments, both in juvenile *E. scolopes* and in the Hawaiian planktonic environment, will be needed to evaluate the relative robustness of the local host specificity and the host/seawater trade-off hypotheses. For instance, experimental evolution from a single parental genotype has provided a highly rigorous approach to measuring differences in bacterial traits associated with environmental trade-offs (Bennett and Lenski, 2007). Because experimental evolution can be performed in the squid host (Schuster *et al.*, 2010), similar studies in *E. scolopes* from a certain Oahu population (for example, MB), using a known squid symbiont (for example, the non-Group-A strain *V. fischeri* ES114), could examine trade-offs between the light-organ and seawater habitats. Similarly, it will be important to determine whether the planktonic habitat represents an important non-symbiotic reservoir for *V. fischeri* proliferation, or simply a carbon-poor environment that must be survived on the way to a more nutritive niche (for example, the squid light organ).

In summary, the evolutionary concepts of trade-offs and local-host/symbiont co-evolution and specificity will provide useful frameworks for further study of sympatric *V. fischeri* collected from Oahu, Hawaii. With the help of additional observational and experimental work both in the host and in Hawaiian seawater, the squid/*Vibrio* association will continue to be a model system in which the study of an inter-kingdom mutualism enriches knowledge of bacterial ecology and evolution.

Acknowledgements

We thank the members of the Ruby and McFall-Ngai labs for discussing this research and the resulting manuscript. We are also grateful for the statistical advice of C Ané and the critical editing of H Goodrich-Blair. R Gates provided support and laboratory resources at the Hawaiian Institute for Marine Biology (HIMB) during collection. MSW thanks both H and T Putnam for help in the logistical aspects of the Hawaii fieldwork. MSW was supported by NIH Molecular Biosciences (5T32GM007215-35) and NIH Microbes in Health and Disease (2T32AI055397-07) Training Grants through University of Wisconsin-Madison. Research support was provided by NSF IOS 0841507 to M McFall-Ngai and EGR, and NIH R01 RR12294 to EGR and M McFall-Ngai.

References

Bennett AF, Lenski RE. (2007). An experimental test of evolutionary trade-offs during temperature adaptation. *Proc Natl Acad Sci USA* **104**: 8649–8654.

Boettcher KJ, Ruby EG. (1990). Depressed light emission by symbiotic *Vibrio fischeri* of the sepiolid squid *Euprymna scolopes*. *J Bacteriol* **172**: 3701–3706.

Bose JL, Wollenberg MS, Colton DM, Mandel MJ, Septer AN, Dunn AK *et al.* (2011). Contribution of rapid evolution of the *luxR-luxI* intergenic region to the diverse bioluminescence output of *Vibrio fischeri* strains isolated from different environments. *Appl Environ Microbiol* **77**: 2445–2457.

Cafaro MJ, Poulsen M, Little AE, Price SL, Gerardo NM, Wong B *et al.* (2010). Specificity in the symbiotic association between fungus-growing ants and protective *Pseudonocardia* bacteria. *Proc Biol Sci* **278**: 1814–1822.

Didelot X, Falush D. (2007). Inference of bacterial microevolution using multilocus sequence data. *Genetics* **175**: 1251–1266.
Dunlap PV, Ast JC, Kimura H, Fukui A, Yoshino T, Endo H. (2007). Phylogenetic analysis of host-symbiont specificity and codivergence in bioluminescent symbioses. *Cladistics* **23**: 507–532.
Dykhuizen DE, Green L. (1991). Recombination in *Escherichia coli* and the definition of biological species. *J Bacteriol* **173**: 7257–7268.
Edgar RC. (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res* **32**: 1792–1797.
Gelman A, Rubin DB. (1992). Inference from iterative simulation using multiple sequences. *Stat Sci* **7**: 457–472.
Giovannoni S, Stingl U. (2007). The importance of culturing bacterioplankton in the 'omics' age. *Nat Rev Microbiol* **5**: 820–826.
Gouy M, Guindon S, Gascuel O. (2010). SeaView version 4: a multiplatform graphical user interface for sequence alignment and phylogenetic tree building. *Mol Biol Evol* **27**: 221–224.
Hanage WP, Fraser C, Spratt BG. (2006). Sequences, sequence clusters and bacterial species. *Philos Trans R Soc Lond B Biol Sci* **361**: 1917–1927.
Holland BR, Huber KT, Moulton V, Lockhart PJ. (2004). Using consensus networks to visualize contradictory evidence for species phylogeny. *Mol Biol Evol* **21**: 1459–1461.
Huson DH, Bryant D. (2006). Application of phylogenetic networks in evolutionary studies. *Mol Biol Evol* **23**: 254–267.
Jones BW, Maruyama A, Ouverney CC, Nishiguchi MK. (2007). Spatial and temporal distribution of the *Vibrionaceae* in coastal waters of Hawaii, Australia, and France. *Microb Ecol* **54**: 314–323.
Kimbell JR, McFall-Ngai MJ, Roderick GK. (2002). Two genetically distinct populations of bobtail squid, *Euprymna scolopes*, exist on the island of Oahu. *Pac Science* **56**: 347–355.
Kimura M. (1983). *The Neutral Theory of Molecular Evolution*. Cambridge University Press: Cambridge.
Lee K, Ruby EG. (1995). Symbiotic role of the viable but nonculturable state of *Vibrio fischeri* in Hawaiian coastal seawater. *Appl Environ Microbiol* **61**: 278–283.
Lee KH, Ruby EG. (1992). Detection of the light organ symbiont, *Vibrio fischeri*, in Hawaiian seawater by using *lux* gene probes. *Appl Environ Microbiol* **58**: 942–947.
Lee KH. (1994). *Ecology of Vibrio fischeri, the Light Organ Symbiont of the Hawaiian Sepiolid Squid Euprymna scolopes*. University of Southern California: Los Angeles.
Lee KH, Ruby EG. (1994). Effect of the squid host on the abundance and distribution of symbiotic *Vibrio fischeri* in nature. *Appl Environ Microbiol* **60**: 1565–1571.
Mandel MJ, Wollenberg MS, Stabb EV, Visick KL, Ruby EG. (2009). A single regulatory gene is sufficient to alter bacterial host range. *Nature* **458**: 215–218.
Mandel MJ. (2010). Models and approaches to dissect host–symbiont specificity. *Trends Microbiol* **18**: 504–511.
Maneesakorn P, An R, Daneshvar H, Taylor K, Bai X, Adams BJ *et al.* (2011). Phylogenetic and cophylogenetic relationships of entomopathogenic nematodes (Heterorhabditis: Rhabditida) and their symbiotic bacteria (*Photorhabdus*: Enterobacteriaceae). *Mol Phylogenet Evol* **59**: 271–280.
Martínez-Romero E. (2009). Coevolution in *Rhizobium*–legume symbiosis? *DNA Cell Biol* **28**: 361–370.
McFall-Ngai MJ, Ruby EG. (1991). Symbiont recognition and subsequent morphogenesis as early events in an animal–bacterial mutualism. *Science* **254**: 1491–1494.
Mueller UG, Ishak H, Lee JC, Sen R, Gutell RR. (2010). Placement of attine ant-associated *Pseudonocardia* in a global *Pseudonocardia* phylogeny (Pseudonocardiaceae, Actinomycetales): a test of two symbiont-association models. *Antonie Van Leeuwenhoek* **98**: 195–212.
Nishiguchi MK, Ruby EG, McFall-Ngai MJ. (1998). Competitive dominance among strains of luminous bacteria provides an unusual form of evidence for parallel evolution in Sepiolid squid–*Vibrio* symbioses. *Appl Environ Microbiol* **64**: 3209–3213.
Nishiguchi MK. (2002). Host–symbiont recognition in the environmentally transmitted sepiolid squid–*Vibrio* mutualism. *Microb Ecol* **44**: 10–18.
Ramos-Onsins SE, Rozas J. (2002). Statistical properties of new neutrality tests against population growth. *Mol Biol Evol* **19**: 2092–2100.
R_Development_Core_Team (2010). *R: a Language and Environment for Statistical Computing*. R Foundation for Statistical Computing: Vienna, Austria.
Ruby EG, Urbanowski M, Campbell J, Dunn A, Faini M, Gunsalus R *et al.* (2005). Complete genome sequence of *Vibrio fischeri*: a symbiotic bacterium with pathogenic congeners. *Proc Natl Acad Sci USA* **102**: 3004–3009.
Santos SR, Gutiérrez-Rodríguez C, Lasker HR, Coffroth MA. (2003). *Symbiodinium* sp. associations in the gorgonian *Pseudopterogorgia elisabethae* in the Bahamas: high levels of genetic variability and population structure in symbiotic dinoflagellates. *Marine Biol* **143**: 111–120.
Schuster BM, Perry LA, Cooper VS, Whistler CA. (2010). Breaking the language barrier: experimental evolution of non-native *Vibrio fischeri* in squid tailors luminescence to the host. *Symbiosis* **51**: 85–96.
Sicard M, Ferdy JB, Pages S, Le Brun N, Godelle B, Boemare N *et al.* (2004). When mutualists are pathogens: an experimental study of the symbioses between *Steinernema* (entomopathogenic nematodes) and *Xenorhabdus* (bacteria). *J Evol Biol* **17**: 985–993.
Simonsen KL, Churchill GA, Aquadro CF. (1995). Properties of statistical tests of neutrality for DNA polymorphism data. *Genetics* **141**: 413–429.
Stabb EV, Ruby EG. (2003). Contribution of *pilA* to competitive colonization of the squid *Euprymna scolopes* by *Vibrio fischeri*. *Appl Environ Microbiol* **69**: 820–826.
Stabb EV, Millikan DS. (2009). Is the *V. fischeri–E. scolopes* symbiosis a defensive mutualism?. In: White JF, Torres MS (eds) *Defensive Mutualism in Microbial Symbiosis*. CRC Press: New York City, pp 85–98.
Stearns SC. (1992). *The Evolution of Life Histories*. Oxford University Press: New York, NY.
Tajima F. (1989). The effect of change in population size on DNA polymorphism. *Genetics* **123**: 597–601.
Velicer GJ, Lenski RE. (1999). Evolutionary trade-offs under conditions of resource abundance and scarcity: experiments with bacteria. *Ecology* **80**: 1168–1179.
Visick KL, Ruby EG. (2006). *Vibrio fischeri* and its host: it takes two to tango. *Curr Opin Microbiol* **9**: 632–638.
Vos M, Velicer GJ. (2006). Genetic population structure of the soil bacterium *Myxococcus xanthus* at the centimeter scale. *Appl Environ Microbiol* **72**: 3615–3625.

Vos M, Didelot X. (2009). A comparison of homologous recombination rates in bacteria and archaea. *ISME J* **3**: 199–208.

Wei SL, Young RE. (1989). Development of symbiotic bioluminescence in a nearshore cephalopod, *Euprymna scolopes*. *Marine Biol* **103**: 541–546.

Wier AM, Nyholm SV, Mandel MJ, Massengo-Tiasse RP, Schaefer AL, Koroleva I *et al.* (2010). Transcriptional patterns in both host and bacterium underlie a daily rhythm of anatomical and metabolic change in a beneficial symbiosis. *Proc Natl Acad Sci USA* **107**: 2259–2264.

Wollenberg MS, Ruby EG. (2009). Population structure of *Vibrio fischeri* within the light organs of *Euprymna scolopes* squid from two Oahu (Hawaii) populations. *Appl Environ Microbiol* **75**: 193–202.

Supplementary Information accompanies the paper on The ISME Journal website (http://www.nature.com/ismej)

1

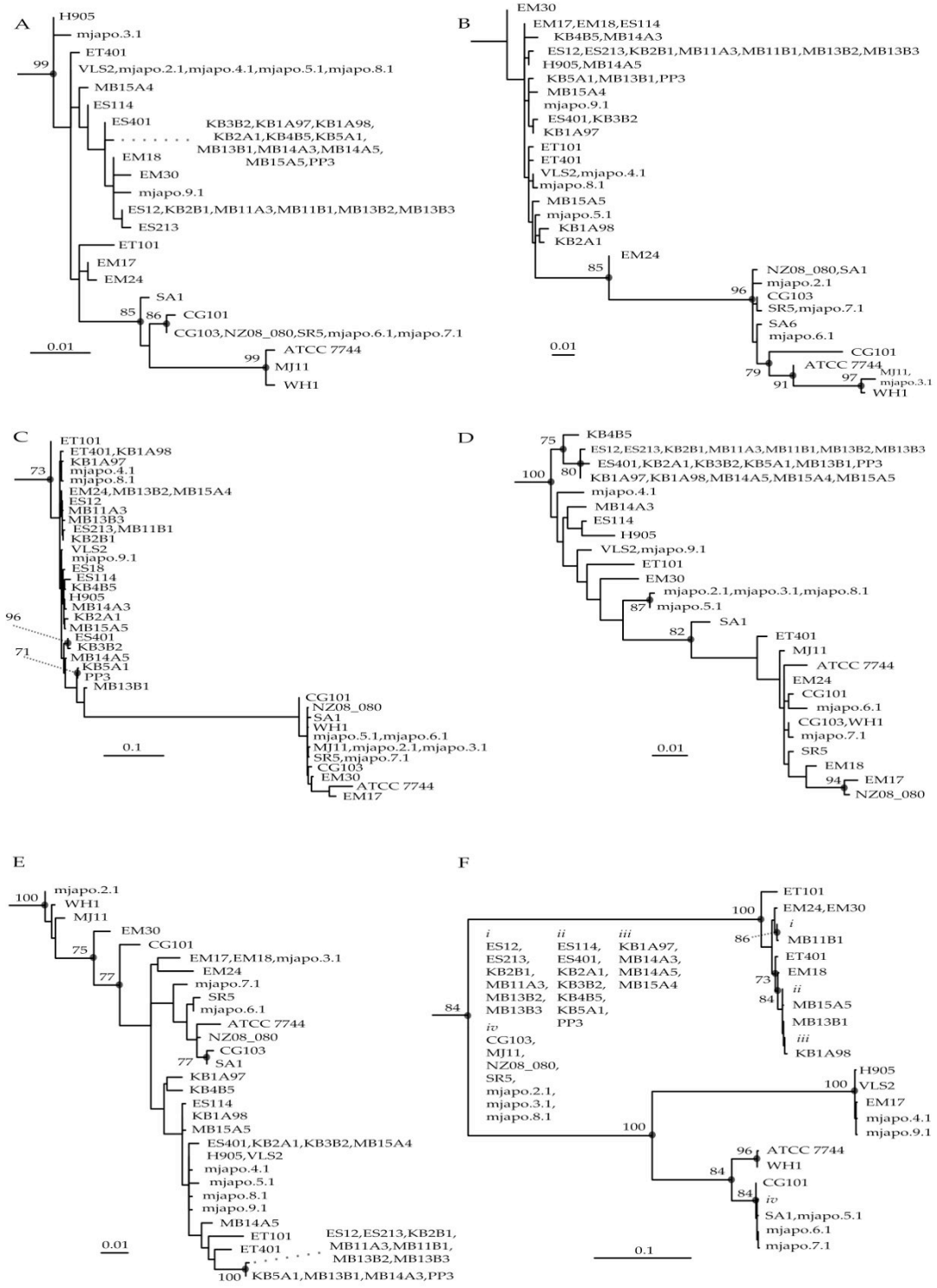

FIGURE S1**.** Maximum likelihood gene phylograms under best-fit substitution models (Table S3) inferred from (**A**) *recA,* (**B**) *mdh,* (**C**) *ainS,* (**D**) *katA,* (**E**) *pyrC,* and (**F**) *luxA* sequence data for 45 *V. fischeri* strains and outgroups (Table S1). In all reconstructions, the outgroup taxa *V. salmonicida* LFI1238 and *V. logei* SA6 have been removed and are represented by a truncated root node – the one exception is (**B**), in which SA6 appears internally. Nodes marked with dots have statistical support (the proportion of 1000 bootstrap pseudo-replicates supporting a node) greater than 70%, as listed at each node. Strains sharing identical sequences for a given locus fragment are all listed at a leaf – except in (**F**), where numbers of groups (*e.g.*, *i, ii, etc.*) of sequences are used. For each tree, black bars represent the number of substitutions/site listed above each bar.

2

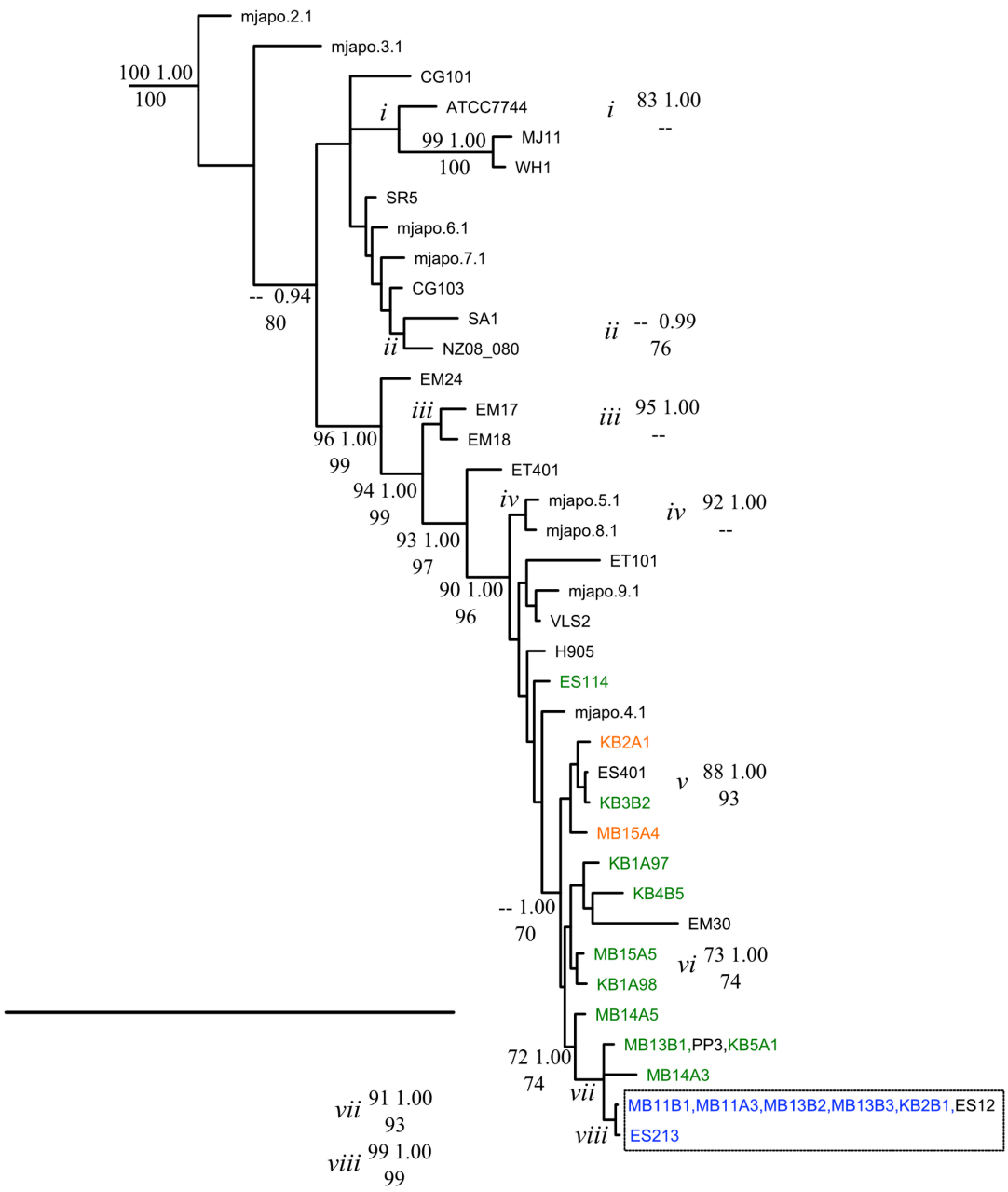

FIGURE S2. 50% majority-rule consensus Bayesian phylogram inferred with a GTR+I+Γ model of evolution (nst=6 rates=invgamma) for the concatenated gene fragments *recA*, *mdh*, *katA*, and *pyrC*. In this reconstruction, the node containing the outgroup taxa (*V. salmonicida* LFI1238 and *V. logei* SA6) is represented by the root. Statistical support is represented at each node by the following three numbers: *upper left*, percentage of 1000 bootstrap ML pseudo-replicates; *upper right*, posterior probability (of approximately 100,000-2,000,000 non-discarded samples); *middle center*, percentage of 1000 bootstrap MP pseudo-replicates. Statistical support values are listed only at nodes where more than 2 methods gave support values ≥ 70% or 0.70; support values < 70% or 0.70 are listed as '--'. Strains sharing identical sequences for a given locus fragment are all listed at that leaf; because of a lack of space, some support values have been listed either immediately to the right of their associated nodes and/or are marked with italicized lower-case roman numerals. VfRep-PCR patterns have been mapped onto the *E. scolopes* symbionts using colors – identically-colored strains all share the same VfRep-PCR pattern (or type) (Table S1): blue = type I (boxed); orange = type II; green = type III. Black bar represents 0.1 substitutions/site.

TABLE S1. General information about Vibrionaceae strains used in this study.

Strain[1]	Collection Description		Source or Reference	VfRep-PCR type	GenBank Accession Numbers					
	Geography	Ecology			*recA*	*mdh*[4]	*ainS*	*katA*	*pyrC*	*luxA*
ATCC 7744; B398	*n.i.*[2]	Planktonic	(Hendrie et al 1970)	*n.i.*[3]	JF509753	EU907965	JF509797	EU907989	JF509855	JF509899
CG101	Australia	*Cleidopus gloriamaris* (fish light organ)	(Lee 1994)	*n.i.*	HQ595306	EU907966	JF509798	EU907990	JF509856	JF509900
CG103	""	""	(Lee 1994)	*n.i.*	HQ595307	HQ595322	JF509799	HQ595331	JF509855	JF509901
EM17; ATCC 700602	Japan (Seto Sea)	*Euprymna morsei* (squid light organ)	(Lee 1994)	*n.i.*	HQ595308	EU907967	JF509800	EU907991	JF509855	JF509902
EM18	""	""	(Lee 1994)	*n.i.*	JF509754	EU907968	JF509801	EU907992	JF509855	JF509903
EM24	""	""	(Lee 1994)	*n.i.*	JF509755	EU907969	JF509802	EU907993	JF509860	JF509904
EM30	""	""	(Lee 1994)	*n.i.*	JF509756	EU907970	JF509803	EU907994	JF509861	JF509905
ES12	Oahu, HI, USA (Kaneohe Bay)	*Euprymna scolopes* (squid light organ)	(Boettcher and Ruby 1994)	*n.i.*	HQ595309	HQ595323	JF509804	HQ595332	JF509862	JF509906
ES114; ATCC 700601	""	""	(Boettcher and Ruby 1990)	III	VF_0535[5]	VF_0276[5]	VF_1037[5]	VF_A0009[5]	VF_A0412[5]	VF_A0921[5]
ES213	Oahu, HI, USA (Maunalua Bay)	""	(Boettcher and Ruby 1994)	I	HQ595310	EU907971	JF509805	EU907995	JF509863	JF509907
ES401	""	""	(Lee 1994)	*n.i.*	HQ595311	HQ595324	JF509806	HQ595333	JF509864	JF509908

ET101	Victoria, Australia (Crib Point)	*Euprymna tasmanica* (squid light organ)	(Nishiguchi 2002)	*n.i.*	HQ595312	HQ595325	JF509807	HQ595334	JF509865	JF509909
ET401	Townsville, Australia (Magnetic Island)	″″	(Nishiguchi 2002)	*n.i.*	HQ595313	HQ595326	JF509808	HQ595335	JF509866	JF509910
H905	Oahu, HI, USA (Kaneohe Bay)	Planktonic	(Lee and Ruby 1992)	*n.i.*	HQ595314	EU907972	JF509809	EU907996	JF509867	JF509911
KB1A97	″″	*E. scolopes* (squid light organ)	(Wollenberg and Ruby 2009)	III	JF509757	EU907973	JF509810	EU907997	JF509868	JF509912
KB1A98	″″	″″	(Wollenberg and Ruby 2009)	III	JF509758	JF509783	JF509811	JF509841	JF509869	JF509913
KB2A1	″″	″″	(Wollenberg and Ruby 2009)	II	JF509759	JF509784	JF509812	JF509842	JF509870	JF509914
KB2B1	″″	″″	(Wollenberg and Ruby 2009)	I	JF509760	JF509785	JF509813	JF509843	JF509871	JF509915
KB3B2	″″	″″	(Wollenberg and Ruby 2009)	III	JF509761	JF509786	JF509814	JF509844	JF509872	JF509916
KB4B5	″″	″″	(Wollenberg and Ruby 2009)	III	JF509762	JF509787	JF509815	JF509845	JF509873	JF509917
KB5A1	″″	″″	(Wollenberg and Ruby 2009)	III	JF509763	EU907974	JF509816	EU907998	JF509874	JF509918

LFI1238 (*V. salmonicida*)	Hammerfest, Norway	*Gadus morhua* (cod head kidney)	(Hjerde et al 2008)	*n.i.*	VSAL_I0634[5]	VSAL_I0359[5]	VSAL_I1158[5]	VSAL_II0215[5]	VSAL_II0468[5]	VSAL_II0962[5]
MB11A3	Oahu, HI, USA (Maunalua Bay)	""	(Wollenberg and Ruby 2009)	I	JF509764	JF509788	JF509817	JF509846	JF509875	JF509919
MB11B1	""	""	(Wollenberg and Ruby 2009)	I	JF509765	JF509789	JF509818	JF509847	JF509876	JF509920
MB13B1	""	""	(Wollenberg and Ruby 2009)	III	JF509766	JF509790	JF509819	JF509848	JF509877	JF509921
MB13B2	""	""	(Wollenberg and Ruby 2009)	I	JF509767	JF509791	JF509820	JF509849	JF509878	JF509922
MB13B3	""	""	(Wollenberg and Ruby 2009)	I	JF509768	JF509792	JF509821	JF509850	JF509879	JF509923
MB14A3	""	""	(Wollenberg and Ruby 2009)	III	JF509769	EU907975	JF509822	EU907999	JF509880	JF509924
MB14A5	""	""	(Wollenberg and Ruby 2009)	III	JF509770	EU907976	JF509823	EU907800	JF509881	JF509925
MB15A4	""	""	(Wollenberg and Ruby 2009)	II	JF509771	JF509793	JF509824	JF509851	JF509882	JF509926
MB15A5	""	""	(Wollenberg and Ruby 2009)	III	JF509772	JF509794	JF509825	JF509852	JF509883	JF509927
MJ11	Japan	*Monocentris*	(Ruby and	*n.i.*	VFMJ11_0538[5]	VFMJ11_0264[5]	VFMJ11_1119[5]	VFMJ11_A0023[5]	VFMJ11_A0452[5]	VFMJ11_A1039[5]

		japonicus (fish light organ)	Nealson 1976)							
mjapo.2.1	Wagu, Japan	""	(Ast et al 2009)	*n.i.*	JF509773	EU907978	JF509826	EU907802	JF509884	JF509928
mjapo.3.1	""	""	(Ast et al 2009)	*n.i.*	JF509774	EU907979	JF509827	EU907803	JF509885	JF509929
mjapo.4.1	""	""	(Ast et al 2009)	*n.i.*	JF509775	EU907980	JF509828	EU907804	JF509886	JF509930
mjapo.5.1	""	""	(Ast et al 2009)	*n.i.*	JF509776	EU907981	JF509829	EU907805	JF509887	JF509931
mjapo.6.1	""	""	(Ast et al 2009)	*n.i.*	JF509777	EU907982	JF509830	EU907806	JF509888	JF509932
mjapo.7.1	""	""	(Dunlap et al 2007)	*n.i.*	JF509778	EU907983	JF509831	EU907807	JF509889	JF509933
mjapo.8.1	""	""	(Mandel et al 2009)	*n.i.*	JF509779	EU907984	JF509832	EU907808	JF509890	JF509934
mjapo.9.1	""	""	(Mandel et al 2009)	*n.i.*	JF509780	EU907985	JF509833	EU907809	JF509891	JF509935
NZ08_080	N. Island, New Zealand (O'Neill Bay)	Planktonic	this study	*n.i.*	JF509781	JF509795	JF509834	JF509853	JF509892	JF509936
PP3	Oahu, HI, USA (Kaneohe Bay)	Planktonic	(Lee and Ruby 1992)	*n.i.*	HQ595317	HQ595329	JF509835	HQ595338	JF509893	JF509937
SA1	Banyuls sur Mer, France	*Sepiola affinis* (squid light organ)	(Fidopiastis et al 1998)	*n.i.*	HQ595318	EU907986	JF509836	EU908010	JF509894	JF509938
SA6	""	""	(Fidopiastis	*n.i.*	JF509782	JF509796	JF509837	JF509854	JF509895	JF509939

(*V. logei*)			et al 1998)							
SR5	Banyuls sur Mer, France	*Sepiola robusta* (squid light organ)	(Fidopiastis et al 1998)	*n.i.*	HQ595319	EU907987	JF509838	EU908011	JF509896	JF509940
VLS2	Oahu, HI, USA (Kaneohe Bay)	*E. scolopes* (squid light organ)	(Lee 1994)	*n.i.*	HQ595320	HQ595330	JF509839	HQ595339	JF509897	JF509941
WH1	Woods Hole, MA, USA	Planktonic	(Lee 1994)	*n.i.*	HQ595321	EU907988	JF509840	EU908012	JF509898	JF509942

[1] All strains are *V. fischeri* unless otherwise noted.

[2] *n.i.* – Collection location undocumented.

[3] *n.i.* – VfRep-PCR fingerprint not previously determined by Wollenberg and Ruby (Wollenberg and Ruby 2009).

[4] Locus tag of the longer *mdh* fragment given in table; however, a partial sequence was used for alignment (as described in Materials and Methods section, and listed in Table S2).

[5] Locus tag of the full gene from the entire genome sequence given in table; however, only partial sequence (Table S2) was used for alignment.

TABLE S2. Loci and primers utilized in *V. fischeri* multi-locus sequence analysis scheme.

Gene and Product	Chromosome and Position (*V. fischeri* ES114; *V. fischeri* MJ11)	Gene Length	Fragment Position[1]	Fragment Length	Coverage	Outer Primer – Forward Outer Primer – Reverse Inner Primer – Forward Inner Primer – Reverse
recA recombinase A	Large (577,808 – 578,854; 566,266 – 567,312)	1047	64-753	690	66%	5′-GACGATAACAAGAAAAAAGCACTGG-3′ 5′-CGTTTTCTTCAATTTCWGGAGC-3′ 5′-TGARAARCARTTYGGTAAAGG-3′ 5′-GGAGCRGCATCAGTCTCTGG-3′
mdh malate dehydrogenase	Large (288,534 – 289,469; 272,693 – 273,628)	936	85-603	519	55%	5′-AAGTAGCTGTTATTGGTGC-3′ 5′- CTTCGCCAATTTTGATATCG-3′ 5′- GGCATTGGACAAGCGTTAGC-3′ 5′-CGCCTCTTAGCGTATCTAGC-3′
ainS autoinducer synthase	Large (1,142,800 – 1,143,942; 1,177,251 – 1,178,438)	1143	184-963	780	58%	5′-GGCGGAACKATTGGAAATTTGG-3′ 5′-ATTGATTGAAAGCAGATGATGC-3′ 5′-CTTATTTTCAACATCAGAAGC-3′ 5′-TTTACATCTTGCTCACTTGTG-3′
katA catalase	Small (10,816 – 12,264; 30,031 – 31,479)	1449	580-1368	789	54%	5′-TGTCCTGTTGCACATAACC-3′ 5′-CGCTTACATCAATATCAAG-3′ 5′-CGTGGTATTCCTGCAACATAC-3′ 5′-CCGATACCTTCACCATAAGC-3′
pyrC dihydroorotase	Small (465,591 – 466,619; 513,315 – 514,343)	1029	94-975	882	86%	5′-CTGATGATTGGCATTTACAC-3′ 5′-GCCACTCAACAGCTTCACC-3′ 5′-CACTTACGTGATGGTGATGTG-3′ 5′- GCCACTCAACAGCTTCACC-3′
luxA luciferase alpha subunit	Small (1,047,714 – 1,046,650; 1,158,285 – 1,159,349)	1068	91-960	870	81%	5′-ACAAGTAYWACWGTTAARGAGCG-3′ 5′-AAGTGRTGTTCAYWWACAAARGCAG-3′ 5′-CGTATCARCCACCAGGYGAAACTC-3′ 5′-TTCWTCTTCAGTKCCATTAGC-3′

[1]Position relative to full-locus lengths from the ES114 genome.

TABLE S3. Sequence characteristics and maximum-likelihood estimates of frequency and rate parameter values for each gene fragment considered for use in the *V. fischeri* MLSA scheme[1]

Locus fragment(s)	Fragment length (in bp)	No. nt alleles/ H_d[2]	No. aa alleles	No. C sites (%)[3]	No. V sites (%)[4]	No. PI sites (%)[5]	π (A,C, G, T)[6]	Rate matrix[7]	*I*[8]	*α*[9]	Best Fit model[10]
recA	690	22/ 0.909	3	644 (93)	46 (7)	29 (4)	0.2985, 0.1774, 0.2319, 0.2922	R_a = 0.374, R_b = 2.353 R_c = 1.000, R_d = 0.374 R_e = 4.433, R_f = 1.000	NA	0.263	Custom: 4 sub. types + Γ -lnL = 1860.1488
mdh	519	28/ 0.966	5	437 (84)	82 (16)	70 (13)	0.2787, 0.1723, 0.2396, 0.3094	R_a = 1.000, R_b = 3.366 R_c = 3.366, R_d = 0.396 R_e = 8.021, R_f = 1.000	NA	0.157	Custom: 4 sub. types + Γ -lnL = 1629.4243
ainS	780	37 0.990	28	569 (73)	211 (27)	187 (24)	0.3654, 0.1978, 0.1311, 0.3056	R_a = 2.607, R_b = 9.857 R_c = 1.000, R_d = 2.607 R_e = 9.857, R_f = 1.000	0.229	0.652	Custom 3 sub. types + I + Γ -lnL = 4130.2858
katA	789	26 0.948	12	707 (90)	82 (10)	61 (8)	0.3138, 0.2097, 0.2060, 0.2705	R_a = 0.334, R_b = 2.744 R_c = 1.000, R_d = 0.334 R_e = 5.813, R_f = 1.000	0.524	0.716	Custom 4 sub. types + I + Γ -lnL = 2655.4157
pyrC	882	30 0.963	9	792 (90)	90 (10)	66 (7)	0.3374, 0.2058, 0.1608, 0.2960	R_a = 0.320, R_b = 3.037 R_c = 1.000, R_d = 0.320 R_e = 5.000, R_f = 1.000	0.501	0.680	Custom 4 sub. types + I + Γ -lnL = 3170.2015
luxA	870	23 0.934	14	584 (67)	286 (33)	279 (32)	0.3214, 0.1496, 0.2186, 0.3104	R_a = 4.176, R_b = 7.150 R_c = 2.285, R_d = 4.176 R_e = 19.791, R_f = 1.000	NA	0.298	Custom: 5 sub. types + Γ -lnL = 3821.3710

[1]ML parameters estimated from the 45 *V. fischeri* isolates listed, plus the outgroup taxa *V. salmonicida* LFI1238 and *V. logei* SA6 (with redundant sequences removed); all other parameters in the table were calculated using only the 45 *V. fischeri* isolates.

[2]Haplotype diversity.

[3]Number and percentage of constant sites (for nt data).

[4]Number and percentage of variable sites (for nt data).

[5]Number and percentage of parsimony informative sites (for nt data).

[6]Nucleotide equilibrium frequencies.

[7] Substitution rates: R_a= [A-C], R_b= [A-G], R_c= [A-T], R_d= [C-G], R_e= [C-T], R_f= [G-T].

[8] Proportion of invariant sites; NA=variable not used in model.

[9] Shape parameter of the gamma-distributed across-sites rate variation (approximated by four discrete rate categories).

[10] Best-fit model and associated –ln likelihood value.

TABLE S4. Nucleotide polymorphism summary statistics, including mutation and recombination parameters, for molecular sequence data from 45 different *V. fischeri* strains.

Locus Fragment	S[1]	η[2]	k[3]	π[4]	$\emptyset_W$[5]	$\emptyset_\Pi$[6]	C_{HW}[7]	C_H[8]	c/μ[9]
recA	46	48	8.27	0.012	0.0076	0.0060	0.0068	0.0007	0.89
mdh	82	93	23.09	0.044	0.0180	0.0222	0.0074	0.0000	0.41
ainS	211	238	80.20	0.103	0.0310	0.0520	0.0153	0.0000	0.49
katA	82	88	22.42	0.028	0.0119	0.0142	0.0141	0.0049	1.19
pyrC	90	96	20.61	0.023	0.0117	0.0117	0.0236	0.0107	2.02
luxA	286	357	119.48	0.137	0.0376	0.0687	0.0063	0.0005	0.17

[1] Number of segregating sites.
[2] Total number of mutations.
[3] Average number of pairwise nucleotide differences.
[4] Nucleotide diversity per nucleotide site.
[5] Population mutation parameter ($\emptyset$), per site, as estimated by Watterson's (Watterson 1975) method using S.
[6] Population mutation parameter ($\emptyset$), per site, as estimated by Tajima's (Tajima 1983) method using k.
[7] Population recombination parameter (C), per site, as estimated by Hey and Wakeley's (Hey and Wakeley 1997) method.
[8] Population recombination parameter (C), per site, as estimated by Hudson's (Hudson 1987) method.
[9] Recombination/mutation rate ratio; calculated using C_{HW} and $\emptyset_W$.

TABLE S5. Results of the incongruence length difference (ILD) test, and Shimodaira-Hasegawa (SH) tests for topological incongruence among data partitions.

Locus Fragment	*recA*	*mdh*	*ainS*	*katA*	*pyrC*	*luxA*
recA	-	243[1] 262[2]	618 409	324 365	359 317	661 274
mdh	285[b]	-	659 568	365 806	400 462	702 521
ainS	1352	900	-	740 1490	775 981	1077 1106
katA	430	681	650	-	481 490	783 555
pyrC	308	549	660	653	-	818 681
luxA	1816	2514	2402	2774	1991	-

[1] Value of the tree length of the original partition; for all ILD comparisons, the original partition was significantly smaller than all simulated partitions (*i.e.*, $P = 0.01$), indicating significant incongruence.

[2] Natural log likelihood difference between ML tree topology for each locus (row), and that same locus constrained by ML topologies of the other five loci (columns); for all comparisons $P < 0.01$, indicating significant incongruence.