Evolution of symbiosis in the *Vibrionaceae*: a combined approach using molecules and physiology

Michele K. Nishiguchi and Vinod S. Nair

Department of Biology, New Mexico State University, Box 30001, MSC 3AF, Las Cruces, NM 88003-8001, USA

Correspondence
Michele K. Nishiguchi
nish@nmsu.edu

The family *Vibrionaceae* is considered to be one of the most diverse and well-studied groups of bacteria. Here, evolution is assessed within the *Vibrionaceae* to determine whether multiple origins of eukaryotic associations have occurred within this diverse group of bacteria. Analyses were based on a large molecular dataset, along with a matrix that consisted of 100 biochemical and restriction digest characters. By using direct optimization methods to analyse both datasets individually and in combination, a total-evidence cladogram has been produced, which supports the hypothesis that several important symbionts (both mutualistic and pathogenic) within the *Vibrionaceae* are not monophyletic. This leads us to consider that symbiosis (and subsequently, associations with *Eukarya*) has evolved multiple times within the *Vibrionaceae* lineage.

INTRODUCTION

The evolution of microbial pathogens from mutualistic and free-living bacteria has been a targeted area of interest for determining the roles of virulence and specificity among closely related species. Previous studies have used either biochemical or molecular data to elucidate relationships among selected groups of taxa, but none have clarified the transition between free-living and symbiotic species (Bang *et al.*, 1981; Baumann *et al.*, 1983; Nearhos & Fuerst, 1987). The methods by which pathogens evolve among various clades of micro-organisms and whether particular virulence factors are transferred and then regulated within and between species are common themes that underlie the evolution of pathogenesis. The *Vibrionaceae* are a group of highly divergent bacteria, which contains both free-living and symbiotic species (Nishiguchi & Jones, 2003). Members of this family are globally important due to their prevalence in numerous ecological niches. Many species of the *Vibrionaceae* possess the ability to bioluminesce via the *lux* operon, although its biological role in bacterial cells remains unknown (Czyż *et al.*, 2000). Many of these luminescent species form symbiotic relationships, both pathogenic and cooperative, with marine organisms. They are found increasingly to be the cause of diseases in many secondary hosts, which consume primary marine hosts that contain concentrated numbers of vibrios (Ruby & Lee, 1998). Recently, it has been determined that specific virulence factors in pathogenic *Vibrio* species are exhibited in the colonization of eukaryotic tissue by non-pathogenic *Vibrio fischeri* (Reich & Schoolnik, 1994, 1996; Reich *et al.*, 1997; Ruby, 1999b; Stabb & Ruby, 2003) and may even be required for colonization (Nyholm *et al.*, 2000; Stabb *et al.*, 2001). To colonize an animal host, both pathogenic and mutualistic bacteria must conquer defence mechanisms of the host that discourage the growth of unacceptable micro-organisms. These defences may include environmental conditions within the host, active immune responses and general host defence (Koropatnick *et al.*, 2001; Stewart *et al.*, 2001). As a result, there is a growing number of similarities between the tissue colonization of benign species, such as *V. fischeri*, and that of pathogenic vibrios, such as *Vibrio cholerae* and *Vibrio parahaemolyticus* (Makino *et al.*, 2003). Due to these similarities, it has been suggested that many of the characteristics of host-tissue colonization in pathogenic and benign species may have arisen from a common ancestral origin (McFall-Ngai, 1999; Ruby, 1999a).

In this study, we have attempted to complete a phylogenetic survey of the *Vibrionaceae* by using DNA sequence data from three molecular loci [16S rRNA, the intergenic region of the *lux* operon (*luxRI*) and glyceraldehyde phosphate dehydrogenase (*gapA*)] and 100 non-sequence characters (biochemical and restriction digest patterns), in order to determine which evolutionary patterns are prevalent and influential for radiation among this species-rich family of *γ-Proteobacteria*. Total sequence data comprise approximately 3·6 kb per complete strain. These loci were selected due to their ability to resolve some of the internal clades

Published online ahead of print on 6 June 2003 as DOI 10.1099/ijs.0.02792-0.

Abbreviation: TBR, tree bisection and reconnection.

A table showing all biochemical and restriction digest data that were used in the analysis is available as supplementary material in IJSEM Online.

within the *Vibrionaceae* (*gapA* and *luxRI*), as well as to distinguish among distantly related clades (16S rRNA). This combination of molecular loci should delineate phylogenetic history at various hierarchical levels; thus, the resolution of the analyses should distinguish among the span of species (by using 16S rRNA) and strains (by using *gapA* and *luxRI*). With the knowledge that colonization of host species is a particularly important aspect of the life-history strategy of some species of the *Vibrionaceae*, we combined the molecular and morphological attributes of 75 representative biotypes to determine whether there was an ancestral split among different symbiotic biotypes of *Vibrionaceae*, leading to several species that have evolved the ability to associate with a specific eukaryotic host.

METHODS

Isolation of bacteria. Bacterial isolates were obtained from squid or fish light organs by previously described methods (Nishiguchi *et al.*, 1998; Nishiguchi, 2000, 2002). Briefly, bacterial strains were isolated from loliginid, sepiolid or monocentrid light organs, or were obtained from various laboratories (Table 1). Light organs were dissected out of the eukaryotic host and homogenized in sterile sea water, then subsequently diluted and plated on Luria–Bertani high-salt (LBS) [32 parts per thousand (p.p.t.)] agar plates (Ruby & Asato, 1993). Colonies were isolated the next day and grown in 5 ml LBS medium overnight at their optimum temperature (18–28 °C) and frozen as stocks until they were utilized. Squid and fish symbionts were identified according to previously described methods (Nishiguchi *et al.*, 1998; Nishiguchi, 2000).

Biochemical analyses of isolates, utilization of carbon type and RFLP scoring. Bacterial strains from frozen or fresh stock were cultured on LBS agar plates overnight at their optimum temperatures (18–28 °C). Single c.f.u. of each strain were obtained the following day and liquid cultures were grown to an OD_{600} of 0·5 in 10 ml LBS medium. The culture was then centrifuged in 50 ml Sorvall tubes at 5000 r.p.m. for 10 min in a refrigerated centrifuge (5 °C). The pellet was resuspended gently in 5 ml PBS (pH 7·3) and centrifuged at 5000 r.p.m. for 10 min. This step was repeated to ascertain complete removal of LBS medium. The pellet was then resuspended in 10 ml PBS (pH 7·3). A 100 µl aliquot of this suspension was added to each of the 96 wells of a GN2 Biolog plate and incubated overnight at their optimum temperature (18–28 °C). Plates were scored colorimetrically for the carbon source present in each well (see Supplementary Table in IJSEM Online). Each Biolog plate was completed in duplicate to ensure reproducibility of the results. Biochemical phenotypes were recorded in a data matrix by using MacClade 4.05 (Maddison & Maddison, 2002), with a score of 0 for no utilization and 1 for utilization of each substrate in the Biolog plate. In total, 95 carbon sources were measured (see Supplementary Table in IJSEM Online). Here we treated each phenotypic feature, as measured by the Biolog plate, as our primary homology statement *sensu* de Pinna (1991). It is phylogenetic analysis that decides ultimately whether the primary homology statement survives the homology test (secondary homology). For RFLPs, we scored the number of fragments per restriction digest from previously reported data (Urakawa *et al.*, 1999). The fragments comprised the last five of the 100 characters used in the biochemical/RFLP data matrix. These characters were scored by the number of fragments produced by each restriction enzyme (i.e. if an enzyme produced four fragments from one strain, the score in the matrix for that RFLP would be 4).

Isolation, amplification and sequencing of bacterial loci. Isolation of DNA was completed by initially growing pure cultures of each strain in 5 ml LBS broth overnight. A sample of each culture (2 ml) was centrifuged for 1 min and the medium was removed. DNA from each culture was then isolated by using a DNeasy isolation kit (Qiagen). Once purified DNA was obtained, 1–10 ng template DNA was used for PCR amplifications. For the 16S rRNA locus, four sets of primers (each amplifying approximately 420 bp) were used to complete the entire 16S rRNA gene sequence (~1600 bp). These consisted of: 1F (5′-AGAGTTTGATCMTGGCTCAG-3′) and 4R (5′-AGGCCTTCTTCATACACGCG-3′); 2F (5′-GCAAGCCTGATGCAGCCATG-3′) and 3R (5′-ATCGTTTACGGCGTGGACTA-3′); 3F (5′-AAACAGGATTAGATACCCTG-3′) and 2R (5′-CTGGTCGTAAGGGCCATGAT-3′); and 4F (5′-AGGTGGGGATGACGTCAAGT-3′) and 1R (5′- AAGGAGGTGWTCCARCC-3′). 16S rRNA gene amplification consisted of one cycle at 94 °C for 2 min, followed by 25 cycles of the following: 94 °C for 15 s, 49 °C for 15 s and 72 °C for 15 s. Reactions were terminated by holding samples at 72 °C for 7 min. For the *luxRI* locus, primers used to amplify the fragment were LuxRIF (5′-CAGCGGTTAGTTGTATTGAG-3′) and LuxRIR (5′-AGCAAAACGRCTTAATTC-3′), which amplified a gene product of approximately 1200 bp. *luxRI* amplification consisted of one cycle at 94 °C for 2 min, followed by 25 cycles of the following: 94 °C for 15 s, annealing for 15 s at temperatures from 42 to 49 °C and extension at 72 °C for 30 s. The reaction was terminated by holding samples at 72 °C for 7 min after amplification was completed. For the *gapA* locus, primers were gapAF (5′-GGATTTGGCCGCATCGGCCG-3′) and gapAR (5′-CCGAACTCGTTGTCGTACCA-3′), which amplified a gene product of ~900 bp. Samples were amplified by using the following procedure: one cycle at 94 °C for 2 min, followed by 25 cycles of 94 °C for 15 s, 36–38 °C for 30–40 s and 72 °C for 30 s. The reactions were terminated by holding samples at 72 °C for 7 min. Two internal primers were used for each of the *luxRI* and *gapA* loci, to obtain complete double-stranded sequences. These were: gapAFInt (5′-CCTGAATTGCTYGGWSAACA-3′) and gapARInt (5′-CACCGAGATGCGTTAACAATG-3′); and LuxRIFInt (5′-GAATGAGGAGACTACCTTAC-3′) and LuxRIRInt (5′-GCCGTAAATACAGAAAGTTTG-3′). All PCRs had similar concentrations of magnesium (2·5 mM), dNTPs (200 µM each; Promega), primers (0·2 µM of each forward and reverse primer), 1 × reaction buffer (10 mM Tris/HCl, pH 9·0, 50 mM KCl, 0·1 % Triton X-100) and 0·2 U *Taq* polymerase (Promega) or AmpliTaq (Applied Biosystems). Once all PCR templates were amplified, reactions were cleaned with a GENECLEAN II DNA purification kit (BIO 101). All samples were pre-sequenced by using Applied Biosystems BigDye (version 3.1) and excess fluorescently labelled dNTPs were removed by spin column or plate (Edge Biosystems). Samples were sequenced on an Applied Biosystems 3100 automated capillary sequencer. Double-stranded sequences were combined and checked for chromatograph errors by using the program Sequencher (version 4.0; Gene Codes). All *luxRI* and *gapA* sequences and approximately 80 % of the 16S rRNA loci were sequenced in our laboratory. 16S rRNA gene sequences were also used to confirm the identity of light organ isolates by comparing the full sequence to those in the Ribosomal Database Project (http://rdp.cme.msu.edu/html).

Phylogenetic analysis. Sequence data were analysed by using the direct optimization method described by Wheeler (1996) and implemented in the computer program POY (Wheeler *et al.*, 2002). This method directly assesses the number of DNA sequence transformations (evolutionary events) required by a phylogenetic topology, without the use of multiple sequence alignment. This is accomplished by generalization of existing character optimization procedures, including insertion and deletion events (indels) in addition to base substitutions. The crux of the method is the treatment of indels as processes, as opposed to the patterns implied by multiple

Table 1. List of all *Vibrionaceae* strains used in the phylogenetic analyses, host species or culture collection accession number (if applicable), genes sequenced for each strain and GenBank accession numbers for these sequences

Strain	Source	Location	16S rRNA	*gapA*	*luxRI*
Escherichia coli			J01859.1		
Listonella pelagia	ATCC 25916^T		X74722		
Salinivibrio costicola	ATCC 33508^T		X74699		
Photobacterium damselae	ATCC 33539^T		X74700		
Photobacterium phosphoreum	Anomalopid fishes		AY292916		AY292963
Photobacterium leiognathi	Free-living		AY292917		
Photobacterium leiognathi LN101	*Uroteuthis noctiluca*		AY292944		AY292981
Photobacterium leiognathi RM1	*Rondeletiola minor*	France (Banyuls sur Mer)	AY292947		AY292983
Photobacterium leiognathi SN2B	*Sepiolina nipponensis*	Japan (Tosa Bay)	AY292951		
Psychrobacter sp. EH201	*Euprymna hyllebergi*	Thailand (Rayong)	AY292940		
Vibrio aerogenes	FG1, CCRC 17041^T		AF124055		
Vibrio aestuarianus	ATCC 35048^T		X74689		
Vibrio anguillarum	ATCC 19264^T		X16895		
'*Vibrio aspartigenicus*'	GSP-1		M98446		
Vibrio alginolyticus	ATCC 17749^T		X56576		
Vibrio campbellii	ATCC 25920^T		X56575		
Vibrio carchariae/*Vibrio harveyi*	*Paralichthys dentatus*		AF134581		
Vibrio carchariae/*Vibrio harveyi* EH701	*Euprymna hyllebergi*	Thailand (Rayong)	AY292941		
Vibrio cholerae	Environmental isolate		AY292952		AY292964
Vibrio cincinnatiensis	ATCC 35912^T		X74698		
Vibrio diabolicus	HE800^T		X99762		
Vibrio diazotrophicus	ATCC 33466^T		X56577		
Vibrio fischeri	ATCC 7744^T		AY292938		AY292977
Vibrio fischeri CG101	*Cleidopus gloriamaris*	Australia (Townsville)	AY292939	AY292958	AY292978
Vibrio fischeri EB12	*Euprymna berryi*	Japan (Tosa Bay)	AY292921	AY292954	AY292968
Vibrio fischeri EM17	*Euprymna morsei*	Japan (Tokyo Bay)	AY292922	AF034846	AY292969
Vibrio fischeri ES114	*Euprymna scolopes*	Hawaii, USA (Kaneohe Bay)	AY292919	AF034845	AY292966
Vibrio fischeri ES915	*Euprymna scolopes*	Hawaii, USA (Paiko)	AY292920	AY292953	AY292967
Vibrio fischeri ET101	*Euprymna tasmanica*	Australia (Melbourne)	AY292923	AF034847	AY292970
Vibrio fischeri ET301	*Euprymna tasmanica*	Australia (Sydney)	AY292942	AY292959	AY292979
Vibrio fischeri ET401	*Euprymna tasmanica*	Australia (Townsville)	AY292943	AY292960	AY292980
Vibrio fischeri MDR7	Free-living	California, USA (Marina del Rey)	AY292945		
Vibrio fischeri MJ101	*Monocentrus japonicus*	Japan	AY292946		AY292982
Vibrio fischeri SA1G	*Sepiola affinis*	France (Banyuls sur Mer)	AY292924	AF034848	AY292971
Vibrio fischeri SI1D	*Sepiola intermedia*	France (Banyuls sur Mer)	AY292948	AY292961	AY292984
Vibrio fischeri SL518	*Sepiola ligulata*	France (Banyuls sur Mer)	AY292950	AY292962	AY292985
Vibrio fischeri SR5	*Sepiola robusta*	France (Banyuls sur Mer)	AY292926	AF034851	AY292972
Vibrio fischeri WH1	Free-living	Massachusetts, USA (Woods Hole)	AY292930	AY292955	AY292973
Vibrio fluvialis	NCTC 11327^T		X76335		
Vibrio furnissii	ATCC 35016^T		X74704		
Vibrio gazogenes	ATCC 29988^T		X74705		
Vibrio harveyi	ATCC 14126^T		X56578		
Vibrio hollisae	ATCC 33564^T		X56583.1		
Vibrio iliopiscarius	ATCC 51760^T		AB000278		
Vibrio lentus Sat101	*Sepiola atlantica*	Spain (Vigo)	AY292935		
Vibrio lentus Sat201	*Sepiola atlantica*	Spain (Vigo)	AY292936		
Vibrio lentus SI01	*Sepiola intermedia*	France (Banyuls sur Mer)	AY292927		
Vibrio logei	AHL vs15		AY292931		
Vibrio logei	ATCC 15382		AY292932	AY292956	AY292974
Vibrio logei	ATCC 35077		AY292933		
Vibrio logei SA112	*Sepiola affinis*	France (Banyuls sur Mer)	AY292925		
Vibrio logei SI1E	*Sepiola intermedia*	France (Banyuls sur Mer)	AY292949		
Vibrio logei SL101	*Sepiola ligulata*	France (Banyuls sur Mer)	AY292928		

Table 1. cont.

Strain	Source	Location	16S rRNA	*gapA*	*luxRI*
Vibrio logei SR181	*Sepiola robusta*	France (Banyuls sur Mer)	AY292934	AY292957	AY292975
Vibrio logei Srond101401	*Sepiola rondeleti*	France (Banyuls sur Mer)	AY292929		
Vibrio marinus	NCIMB 1144^T		X82142		
Vibrio mediterranei	CIP 103203^T		X74710		
Vibrio metschnikovii	CIP 69.14^T		X74711		
Vibrio mimicus	ATCC 33653^T		AY292937		AY292976
Vibrio mytili			X99761		
Vibrio natriegens	ATCC 14048^T		X56581		
Vibrio navarrensis	CIP 103381^T		X74715		
Vibrio nereis	ATCC 25917^T		X74716		
Vibrio nigripulchritudo	ATCC 27043^T		X74717		
Vibrio ordalii	ATCC 33509^T		X74718		
Vibrio orientalis	ATCC 33934^T		X74719		
Vibrio parahaemolyticus	ATCC 17802^T		AY303756		AY292965
Vibrio pectenicida	Ifremer A365^T		Y13830		
Vibrio proteolyticus	ATCC 15338^T		X56579		
Vibrio salmonicida	*Salmo salar*		AY292918		
Vibrio scophthalmi	*Scophthalmus maximus*		U46579		
Vibrio shiloi	AK1^T		AF007115		
Vibrio splendidus	KAT16		AF025329		
Vibrio tapetis	CECT 4600^T		Y08430		
Vibrio tubiashii	ATCC 19109^T		X74725		
Vibrio vulnificus	C7184		X76334		

sequence alignment (Wheeler, 1998). The results of this procedure are directly compatible with parsimony-based tree lengths and the method appears to generate more efficient (and therefore simpler) explanations of sequence variation than multiple sequence alignment (Wheeler, 1996). Direct optimization, although computationally intense, is much less demanding than parsimony-based multiple sequence alignment algorithms. The method has also been demonstrated to yield more congruent results than multiple sequence alignments when congruence among partitions is used as a criterion (Wheeler & Hayashi, 1998).

Character transformations were weighted differentially to observe how they affect phylogenetic conclusions [sensitivity analysis *sensu* Wheeler (1995)]. A parameter space of two analytical variables was examined: insertion–deletion cost ratio and transversion–transition cost ratio. When the transversion–transition ratio was set at a value other than unity, the insertion–deletion cost was set according to the cost of transversions. In total, 20 combinations of parameters were employed in the analysis [insertion–deletion ratios of 1, 2, 4 and 8; transversion–transition ratios of 1, 2, 4, 8 and infinite (transversion parsimony)]. This method allows discernment between stable (those supported by analysis using a wide range of parameters) and unstable (those that only appear under particular parameter sets) relationships.

Molecular data analysis. The three molecular partitions were analysed independently and combined directly, with all characters weighted equally without regard to source. These datasets are referred as 16S (16S rRNA dataset alone), *gapA* (*gapA* dataset alone), *luxRI* (*luxRI* dataset alone) and molecular (16S + *gapA* + *luxRI*). The tree-search strategy adopted combined SPR (subtree pruning and regrafting) and TBR (tree bisection and reconnection) branch-swapping on the best of 50 random-addition replicates (-random, 50), holding 100 trees per round (-maxtrees, 100) and performing one round of tree-fusing (Goloboff, 1999). The commands '-slop 5' and '-checkslop 10' were used; these commands are intended to check all cladogram lengths that are within *n*-tenths of a percentage of the current minimum value. A slop value of 10 would check all cladograms found within 1 % of the minimum tree length. This option slows down the search, but is less affected by the heuristics of the tree-length calculation shortcuts. The 'implied alignments' obtained with POY were checked with PAUP* 4b (Swofford, 2002) and identical tree lengths were obtained. These implied alignments were used to compute jackknife values by performing 1000 random-addition replicates.

Biochemical and restriction digest data analysis. Parsimony analysis of the biochemical/restriction digest dataset was performed with POY. The tree-search strategy adopted involved a heuristic algorithm with random-addition sequence and TBR branch-swapping, similar to the methods described for the molecular data analysis (Wheeler *et al.*, 2002).

Combined analysis. Morphological and total molecular data were combined directly and analysed by using direct optimization for the same 20 parameter sets that were applied to each one of the individual molecular partitions and to the combined molecular analysis. Morphological transformations were weighted equally to the highest of the molecular costs (indels), to diminish the potentially overwhelming effect of molecular data versus morphology. Bremer support values were estimated by using the heuristics procedure implemented in POY (-bremer -constrain 'filename' -topology 'treetopology-in-parenthetical-notation'). In total, we analysed seven partitions and 20 parameter sets per partition (140 analyses) that were executed in a 866 MHz Pentium III processor (RDRAM, 256 Mb).

RESULTS AND DISCUSSION

The use of several loci, combined with biochemical and RFLP attributes among *Vibrio* taxa, is a step forward in understanding the relationships among members of the group. Alternatives to the use of all available evidence must explicitly exclude sets of the data (Kluge, 1989; Giribet *et al.*, 2001). Once any set of characters has been excluded, a bias towards those characters included in the analysis will be evident. Combining all the evidence available allows a phylogenetic framework to be formulated with the maximum explanatory power to produce the least biased result. In this respect, our analysis constitutes the first total-evidence analysis to study bacterial relationships, a common practice in systematic zoology and botany (Giribet, 2002; Giribet & Wheeler, 2002).

Phylogenetic analysis using a combined approach that searches for the most parsimonious solution for character variation appears to be a new method that can be applied not only to metazoan phylogenies, but also to prokaryotic systematics. Molecular data were analysed by using the direct optimization procedure, which derives cladogram costs without multiple sequence alignment. This methodology accommodates sequence length variation as transformations that involve the addition, deletion and substitution of nucleotides, as opposed to the addition of unobservable 'gaps'. Direct optimization produces more parsimonious and more congruent results than multiple sequence alignment in a shorter amount of time (Giribet *et al.*, 2001). POY minimizes the weighted number of evolutionary changes over the entire tree, working in a one-step fashion as opposed to the more classical two-step analyses (alignment and tree search). Nodal support was evaluated by the jackknifing method with a deletion ratio of e^{-1} (about 0·36), in which every character has the same chance of being omitted from a given pseudoreplicate (Farris, 1970). Therefore, the single phylogeny reported here represents the most parsimonious tree with the least amount of bias.

The cladogram from the combined evidence of biochemical characters and molecular data, which minimizes overall incongruence among data partitions (utilizing character congruence as a meta-optimality criterion) (Wheeler, 1995), demonstrates the non-monophyly of the *Vibrionaceae* (Fig. 1). Several clades contain symbiotic vibrios, irrespective of the type of host (i.e. mollusc, annelid, crustacean or fish) or geographical area where they were found (Table 1). This provides evidence that symbiosis has arisen multiple times within the family *Vibrionaceae*, although support for several nodes is below the commonly used 50 % threshold. Also, some free-living species (*V. fischeri* ATCC 7744^T and *Photobacterium leiognathi*) that were included in the analyses were not monophyletic, whereas *V. fischeri* clade A, which includes both free-living and symbiotic isolates, is supported by the analyses (95 % jackknife support). Two other symbiotic strains isolated from sepiolid squids (EH201 and EH701) did not group with the clades that contained the remainder of squid host strains (clades A, B and C; Fig. 1). It is important to note that not all squid symbionts in these clades are from the same geographical location (Indo-west Pacific); in fact, several sister groups are constituted by symbiotic isolates from different host squids (Table 1). Whether *V. fischeri* ATCC 7744^T is an

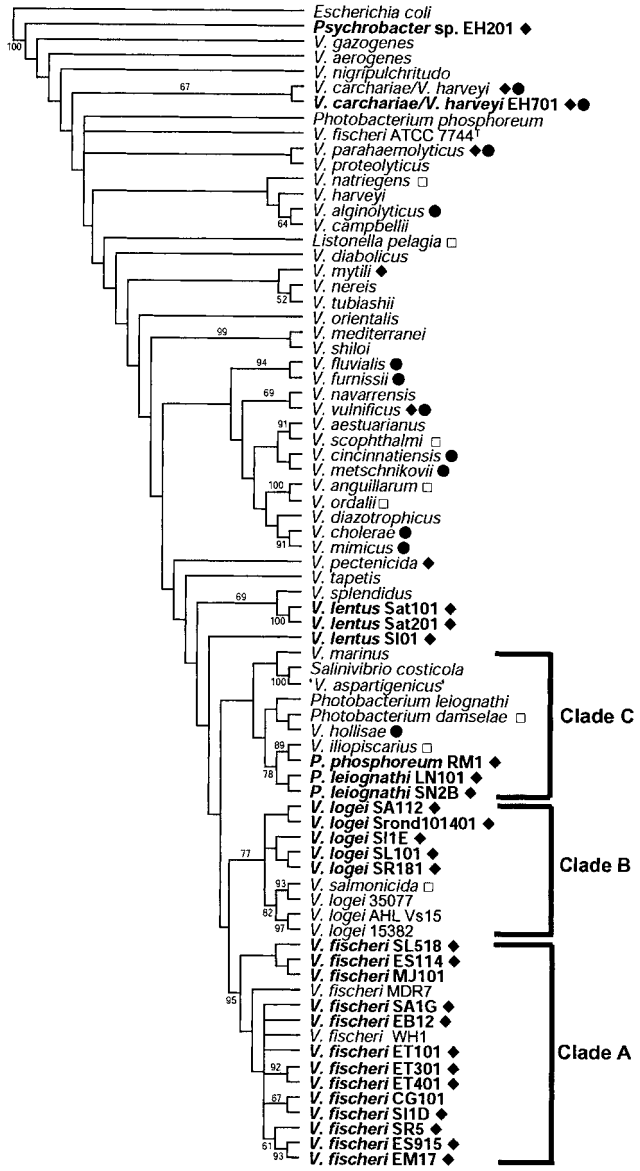

Fig. 1. Phylogenetic tree of *Vibrionaceae* lineages, based on the DNA sequences of three loci and 100 biochemical characters for the parameter set that minimizes overall incongruence among the four partitions. The single tree was obtained when all insertions and deletions equalled all other genetic changes (transitions/transversions). Jackknife values >50 % are shown at nodes. ◆, Taxa that have symbiotic relationships with molluscan hosts; □, fish pathogens; ●, taxa that are human pathogens in the environment. All light organ symbionts (fish or squid) are shown in bold.

ancestral *Vibrio* biotype that has remained separate from other *V. fischeri* environmental isolates because it is in culture remains to be determined. Future investigations should include not only strain types that have been used in previous studies (i.e. *V. fischeri* ATCC 7744^T), but also those that have recently been isolated from environmental samples, which may represent the current population of vibrios that are more likely to share a common or closely related ancestral lineage.

Vibrio logei isolates (both free-living and symbiotic) used in the analyses were found to be polyphyletic, with one clade (B) clustering with the *V. fischeri* clade of symbiotic strains (clade A). Clade B contains several isolates of *V. logei* (both symbiotic and free-living) and *Vibrio salmonicida* (a salmon pathogen), whereas a third, smaller *Vibrio lentus* clade is sister to *Vibrio splendidus* (69 % jackknife support). The finding of a well-supported clade (82 % jackknife frequency) that contains *V. logei* strains ATCC 35077, AHL Vs15 and SR181 and *Vibrio salmonicida* suggests that *V. salmonicida* might have evolved from a free-living *V. logei* congener (Wiik *et al.*, 1995). Embedded within the large clade that contains *V. logei* and *V. fischeri* are several strains of *Photobacterium leiognathi*, *Photobacterium phosphoreum* and *Photobacterium damselae* (clade C; Fig. 1), which supports previous data that suggested that *Photobacterium* and *Vibrio* may be monophyletic (MacDonell & Colwell, 1985; Ruimy *et al.*, 1994). Nesting of *Photobacterium* within the genus *Vibrio* implies that our phylogeny supports this hypothesis, but it also includes other taxa besides these two genera. More data are needed on the inclusion of other hypothesized sister taxa, which may help to further define this group of luminous bacteria.

Although much of the species relatedness differs from that reported in previous phylogenetic studies (Bryant *et al.*, 1986; Dorsch *et al.*, 1992; Kita-Tsukamoto *et al.*, 1993; Alsina & Blanch, 1994; Martin-Kearley & Gow, 1994; Urakawa *et al.*, 1997, 1999), several robust clades are apparent in our phylogeny that have been supported by other studies. *Salinivibrio costicola*, *P. damselae* and *P. leiognathi* have previously been proposed to be sister to one another (Kita-Tsukamoto *et al.*, 1993; Alsina & Blanch, 1994; Ruimy *et al.*, 1994); our results support this hypothesis, but also include other *Vibrio* species that do not cluster with *S. costicola* (Fig. 1). Another relationship supported by our phylogeny is that of human pathogens *Vibrio fluvialis*, *Vibrio furnissii*, *Vibrio cincinnatiensis*, *Vibrio metschnikovii*, *V. cholerae* and *Vibrio mimicus*, which are known to cause enteric distress (Daniels & Shafaie, 2000). Sister group relationships among these pathogenic species have high support values (*V. fluvialis*/*V. furnissii*, 94 %; *Vibrio anguillarum*/*Vibrio ordalii*, 100 %; *V. cholerae*/*V. mimicus*, 91 %) and are most likely to have arisen from a virulent strain. Again, strict monophyly among pathogenic strains is not supported in this clade. Other symbiotic species, such as the *V. fischeri*–*V. logei*–*V. salmonicida* clade, are also supported in our analyses, as in studies that used 5S rRNA (Ruimy *et al.*, 1994) and partial 16S rRNA (Kita-Tsukamoto *et al.*, 1993) gene sequences. Finally, the alliance between *V. cholerae* and *V. mimicus* was well-supported (91 % jackknife frequency), which was also shown by using 16S rRNA (Kita-Tsukamoto *et al.*, 1993) and non-sequence characters (Davis *et al.*, 1981). From this study and others, *V. mimicus* is classified as an atypical strain of *V. cholerae* and probably contains a reduced genome component from its ancestral parent. This was apparent as the 16S rRNA gene sequence was substantially smaller than those of all other vibrios sequenced in this study.

Of importance to this study are the multiple origins of pathogenic/symbiotic strains of *Vibrio* and related genera. As the association between vibrios and invertebrates existed before pathogenic *Vibrio*–human interactions, it appears that tissue-colonization mechanisms were first utilized by mutualists and were then enlisted for use in pathogenic associations (Ruby, 1999a; Schjørring & Koella, 2003). Unfortunately, very little is known about communication, regulation or co-evolution among interactive species (Nishiguchi, 2001). It is relatively unknown how host–mutualist relationships are disrupted and/or capitalized by pathogenic bacteria (Ruby, 1999b). Evolutionary transitions between mutualism and pathogenicity would produce a mosaic of types of interactions within the tree, resulting in close relationships between the two (Moran & Wernegreen, 2000), as suggested by our data. Our phylogenetic tree of *Vibrio* species is the first to combine molecular and biochemical information for this important group of bacteria and provides new insight on the evolution of both types of association. Future work should include a search for virulence/pathogenicity islands in symbiotic organisms within the genus *Vibrio*, which may help us to further our knowledge of the evolution of pathogenesis. If heterozygous islands are found in non-pathogenic species of *Vibrio*, they can then be compared to those within the genomes of pathogenic *Vibrio* species (Makino *et al.*, 2003). Similarities or differences between the two would lead to further insight in the development of pathogenicity in benign organisms.

ACKNOWLEDGEMENTS

The authors would like to thank members of the Nishiguchi lab and two anonymous referees for help with reviewing the manuscript and other forms of advice during data collection. We would especially like to thank Charity Licklite and Melanie Sanchez for their assistance in sequencing bacterial loci. Funding was provided by NSF DBI-0079820, NSF SBE-0123690 and NIH SO6-GM08136-26 to M. K. N. and grants NIH-RISE GM61222-01 and NIH-MARC GM07667-24 at NMSU for support of minority undergraduates who helped with this project.

REFERENCES

Alsina, M. & Blanch, A. R. (1994). Improvement and update of a set of keys for biochemical identification of *Vibrio* species. *J Appl Bacteriol* **77**, 719–721.

Bang, S. S., Baumann, L., Woolkalis, M. J. & Baumann, P. (1981). Evolutionary relationships in *Vibrio* and *Photobacterium* as determined by immunological studies of superoxide dismutase. *Arch Microbiol* **130**, 111–120.

Baumann, P., Baumann, L. M., Woolkalis, J. & Bang, S. S. (1983). Evolutionary relationships in *Vibrio* and *Photobacterium*: a basis for a natural classification. *Annu Rev Microbiol* **37**, 369–398.

Bryant, T. N., Lee, J. V., West, P. A. & Colwell, R. R. (1986). Numerical classification of species of *Vibrio* and related genera. *J Appl Bacteriol* **61**, 437–467.

Czyż, A., Wróbel, B. & Węgrzyn, G. (2000). *Vibrio harveyi* bioluminescence plays a role in stimulation of DNA repair. *Microbiology* **146**, 283–288.

Daniels, N. A. & Shafaie, A. (2000). A review of pathogenic *Vibrio* infections for clinicians. *Infect Med* **17**, 665–685.

Davis, B. R., Fanning, G. R., Madden, J. M., Steigerwalt, A. G., Bradford, H. B., Jr, Smith, H. L., Jr & Brenner, D. J. (1981). Characterization of biochemically atypical *Vibrio cholerae* strains and designation of a new pathogenic species, *Vibrio mimicus*. *J Clin Microbiol* **14**, 631–639.

de Pinna, M. C. C. (1991). Concepts and tests of homology in the cladistic paradigm. *Cladistics* **7**, 367–394.

Dorsch, M., Lane, D. & Stackebrandt, E. (1992). Towards a phylogeny of the genus *Vibrio* based on 16S rRNA sequences. *Int J Syst Bacteriol* **42**, 58–63.

Farris, J. S. (1970). Methods for computing Wagner trees. *Syst Zool* **19**, 83–92.

Giribet, G. (2002). Relationships among metazoan phyla as inferred from 18S rRNA sequence data: a methodological approach. In *Molecular Systematics and Evolution: Theory and Practice*, pp. 85–101. Edited by R. DeSalle, G. Giribet & W. C. Wheeler. Basel, Switzerland: Birkhäuser.

Giribet, G. & Wheeler, W. C. (2002). On bivalve phylogeny: a high-level analysis of the Bivalvia (Mollusca) based on combined morphology and DNA sequence data. *Invertebr Biol* **121**, 271–324.

Giribet, G., Edgecombe, G. D. & Wheeler, W. C. (2001). Arthropod phylogeny based on eight molecular loci and morphology. *Nature* **413**, 157–161.

Goloboff, P. A. (1999). Analyzing large data sets in reasonable times: solutions for composite optima. *Cladistics* **15**, 415–428.

Kita-Tsukamoto, K., Oyaizu, H., Nanba, K. & Simidu, U. (1993). Phylogenetic relationships of marine bacteria, mainly members of the family *Vibrionaceae*, determined on the basis of 16S rRNA sequences. *Int J Syst Bacteriol* **43**, 8–19.

Kluge, A. G. (1989). A concern for evidence and a phylogenetic hypothesis of relationships among *Epicrates* (Boidae, Serpentes). *Syst Zool* **38**, 7–25.

Koropatnick, T., Estrada, A., Apicella, M. A. & McFall-Ngai, M. J. (2001). Messengers or mercenaries? Characterizing the role of macrophage-like hemocytes in the bacteria-induced morphogenesis of the juvenile light organ of *Euprymna scolopes*. *Am Zool* **41**, 1496 (abstract no. P1.24).

MacDonell, M. T. & Colwell, R. R. (1985). Phylogeny of the *Vibrionaceae*, and recommendation for two new genera, *Listonella* and *Shewanella*. *Syst Appl Microbiol* **6**, 171–182.

Maddison, D. R. & Maddison, W. P. (2002). MacClade. Sunderland, MD: Sinauer Associates.

Makino, K., Oshima, K., Kurokawa, K. & 14 other authors (2003). Genome sequence of *Vibrio parahaemolyticus*: a pathogenic mechanism distinct from that of *V. cholerae*. *Lancet* **361**, 743–749.

Martin-Kearley, J. & Gow, J. A. (1994). Numerical taxonomy of *Vibrionaceae* from Newfoundland coastal waters. *Can J Microbiol* **40**, 355–361.

McFall-Ngai, M. J. (1999). Consequences of evolving with bacterial symbionts: lessons from the squid-vibrio association. *Annu Rev Ecol Syst* **30**, 235–256.

Moran, N. A. & Wernegreen, J. J. (2000). Lifestyle evolution in symbiotic bacteria: insights from genomics. *Trends Ecol Evol* **15**, 321–326.

Nearhos, S. P. & Fuerst, J. A. (1987). Reanalysis of 5S rRNA sequence data for the *Vibrionaceae* with the Clustan Program Suite. *Curr Microbiol* **15**, 329–335.

Nishiguchi, M. K. (2000). Temperature affects species distribution in symbiotic populations of *Vibrio* spp. *Appl Environ Microbiol* **66**, 3550–3555.

Nishiguchi, M. K. (2001). Co-evolution of symbionts and hosts: the sepiolid-*Vibrio* model. In *Symbiosis: Mechanisms and Model Systems*, pp. 757–774. Edited by J. Seckbach. Dordrecht, the Netherlands: Kluwer.

Nishiguchi, M. K. (2002). Host–symbiont recognition in the environmentally transmitted sepiolid squid–*Vibrio* mutualism. *Microb Ecol* **44**, 10–18.

Nishiguchi, M. K. & Jones, B. W. (2003). Microbial biodiversity within the *Vibrionaceae*. In *Origins, Evolution, and the Biodiversity of Microbial Life*. Edited by J. Seckback. Dordrecht, the Netherlands: Kluwer (in press).

Nishiguchi, M. K., Ruby, E. G. & McFall-Ngai, M. J. (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.

Nyholm, S. V., Stabb, E. V., Ruby, E. G. & McFall-Ngai, M. J. (2000). Establishment of an animal–bacterial association: recruiting symbiotic vibrios from the environment. *Proc Natl Acad Sci U S A* **97**, 10231–10235.

Reich, K. A. & Schoolnik, G. K. (1994). The light organ symbiont *Vibrio fischeri* possesses a homolog of the *Vibrio cholerae* transmembrane transcriptional activator ToxR. *J Bacteriol* **176**, 3085–3088.

Reich, K. A. & Schoolnik, G. K. (1996). Halovibrin, secreted from the light organ symbiont *Vibrio fischeri*, is a member of a new class of ADP-ribosyltransferases. *J Bacteriol* **178**, 209–215.

Reich, K. A., Biegel, T. & Schoolnik, G. K. (1997). The light organ symbiont *Vibrio fischeri* possesses two distinct secreted ADP-ribosyltransferases. *J Bacteriol* **179**, 1591–1597.

Ruby, E. G. (1999a). Ecology of a benign ''infection'': colonization of the squid luminous organ by *Vibrio fischeri*. In *Microbial Ecology and Infectious Disease*, pp. 217–231. Edited by E. Rosenberg. Washington, DC: American Society for Microbiology.

Ruby, E. G. (1999b). The *Euprymna scolopes-Vibrio fischeri* symbiosis: a biomedical model for the study of bacterial colonization of animal tissue. *J Mol Microbiol Biotechnol* **1**, 13–21.

Ruby, E. G. & Asato, L. M. (1993). Growth and flagellation of *Vibrio fischeri* during initiation of the sepiolid squid light organ symbiosis. *Arch Microbiol* **159**, 160–167.

Ruby, E. G. & Lee, K.-H. (1998). The *Vibrio fischeri-Euprymna scolopes* light organ association: current ecological paradigms. *Appl Environ Microbiol* **64**, 805–812.

Ruimy, R., Breittmayer, V., Elbaze, P., Lafay, B., Boussemart, O., Gauthier, M. & Christen, R. (1994). Phylogenetic analysis and assessment of the genera *Vibrio*, *Photobacterium*, *Aeromonas*, and *Plesiomonas* deduced from small-subunit rRNA sequences. *Int J Syst Bacteriol* **44**, 416–426.

Schjørring, S. & Koella, J. C. (2003). Sub-lethal effects of pathogens can lead to the evolution of lower virulence in multiple infections. *Proc R Soc Lond B Biol Sci* **270**, 189–193.

Stabb, E. V. & Ruby, E. G. (2003). Contribution of *pilA* to competitive colonization of the squid *Euprymna scolopes* by *Vibrio fischeri*. *Appl Environ Microbiol* **69**, 820–826.

Stabb, E. V., Reich, K. A. & Ruby, E. G. (2001). *Vibrio fischeri* genes *hvnA* and *hvnB* encode secreted NAD^+-glycohydrolases. *J Bacteriol* **183**, 309–317.

Stewart, J. J., Nyholm, S. V. & McFall-Ngai, M. J. (2001). Squid hemocytes are able to discriminate the specific symbiont from other types of bacteria. *Am Zool* **41**, 1595–1596 (abstract no. 62.3).

Swofford, D. L. (2002). PAUP*: Phylogenetic Analysis Using Parsimony (*and other methods). Sunderland, MA: Sinauer Associates.

Urakawa, H., Kita-Tsukamoto, K. & Ohwada, K. (1997). 16S rDNA genotyping using PCR/RFLP (restriction fragment length polymorphism) analysis among the family *Vibrionaceae*. *FEMS Microbiol Lett* **152**, 125–132.

Urakawa, H., Kita-Tsukamoto, K. & Ohwada, K. (1999). Restriction fragment length polymorphism analysis of psychrophilic and psychrotrophic *Vibrio* and *Photobacterium* from the north-western Pacific Ocean and Otsuchi Bay, Japan. *Can J Microbiol* **45**, 67–76.

Wheeler, W. C. (1995). Sequence alignment, parameter sensitivity, and the phylogenetic analysis of molecular data. *Syst Biol* **44**, 321–331.

Wheeler, W. C. (1996). Optimization alignment: the end of multiple sequence alignment in phylogenetics? *Cladistics* **12**, 1–9.

Wheeler, W. C. (1998). Alignment characters, dynamic programming and heuristic solutions. In *Molecular Approaches to Ecology and Evolution*, vol. 1, pp. 243–251. Edited by R. DeSalle & B. Schierwater. Basel, Switzerland: Birkhäuser Verlag.

Wheeler, W. C. & Hayashi, C. Y. (1998). The phylogeny of extant chelicerate orders. *Cladistics* **14**, 173–192.

Wheeler, W. C., Gladstein, D. S. & DeLaet, J. (2002). POY: the optimization of alignment characters. New York: American Museum of Natural History.

Wiik, R., Stackebrandt, E., Valle, O., Daae, F. L., Rødseth, O. M. & Andersen, K. (1995). Classification of fish-pathogenic vibrios based on comparative 16S rRNA analysis. *Int J Syst Bacteriol* **45**, 421–428.