Cohesive molecular genetic data delineate species diversity in the dinoflagellate genus *Symbiodinium*

E. M. SAMPAYO,*‡ S. DOVE* and T. C. LAJEUNESSE†
Centre for Marine Studies, University of Queensland, St. Lucia, Qld 4072, Australia, †Department of Biology, Pennsylvania State University, 208 Mueller Laboratory, University Park, PA 16802, USA

Abstract

The diversity of symbiotic dinoflagellates (*Symbiodinium*) in pocilloporid corals originating from various reef habitats surrounding Heron Island, southern Great Barrier Reef, was examined by targeting ribosomal, mitochondrial, and chloroplast genes using six methods that analyse for sequence differences. The ability of each of 13 genetic analyses to characterize eight ecologically distinct *Symbiodinium* spp. was dependent on the level of conservation of the gene region targeted and the technique used. Other than differences in resolution, phylogenetic reconstructions using nuclear and organelle gene sequences were complementary and when combined produced a well-resolved phylogeny. Analysis of the ribosomal internal transcribed spacers using denaturing gradient gel electrophoresis fingerprinting in combination with sequencing of dominant bands provided a precise method for rapidly resolving and characterizing symbionts into ecologically and evolutionarily distinct units of diversity. Single-stranded conformation polymorphisms of the nuclear ribosomal large subunit (D1/D2 domain) identified the same number of ecologically distinct *Symbiodinium* spp., but profiles were less distinctive. The repetitive sequencing of bacterially cloned ITS2 polymerase chain reaction amplifications generated numerous sequence variants that clustered together according to the symbiont under analysis. The phylogenetic relationships between these clusters show how intragenomic variation in the ribosomal array diverges among closely related eukaryotic genomes. The strong correlation between phylogenetically independent lineages with different ecological and physiological attributes establishes a clear basis for assigning species designations to members of the genus *Symbiodinium*.

Keywords: coral symbionts, diversity, species, *Symbiodinium*, taxonomy, zooxanthellae

Received 16 May 2008; revised 27 October 2008; accepted 4 November 2008

Introduction

Molecular genetic analyses have led to an exponential increase in discoveries concerning the diversity, ecology and evolution of micro-organisms (Morris *et al.* 2002). It is now clear that the biosphere contains a diversity of microbes orders of magnitude greater than was once perceived (e.g. Pace 1997; DeLong & Pace 2001; Rappé & Giovannoni 2003). Assessing the diversity and ecology of most prokaryotes and unicellular eukaryotes remains challenging even with the availability of current molecular-based techniques.

Correspondence: Eugenia M. Sampayo, Fax: +61 733654755;
E-mail: emn13@psu.edu
‡Present address: Department of Biology, Pennsylvania State University, University Park, PA16802, USA

Amid a vast amount of ever-growing sequence data, there is considerable debate surrounding the choice of what genes should be sequenced and what amount of sequence divergence justifies classifying ecologically and evolutionarily distinct units of diversity (i.e. delineating species) (e.g. Hey 2001; Green & Bohannen 2006; Cohan & Perry 2007; Koeppel *et al.* 2008). While there are many ongoing attempts to improve and standardize molecular genetic classifications, this investigation endeavours to accurately, and precisely, measure diversity among dinoflagellate endosymbionts in the genus *Symbiodinium*.

Dinoflagellates in the genus *Symbiodinium* (Freudenthal 1962) occur at high concentrations in the tissues of many cnidarians (reef-building corals, soft corals, fire corals and anemones; Trench 1993; LaJeunesse *et al.* 2003), molluscs

(giant clams and nudibranchs; Rudman 1982; Baillie *et al.* 2000a; Loh *et al.* 2006), flatworms (Barneah *et al.* 2007), sponges (Sarà & Liaci 1964; Schönberg & Loh 2005), and some large protists (foraminifera and ciliates; Pochon *et al.* 2001; Lobban *et al.* 2002). The lack of clear morphological differences as well as technical difficulties in culturing *Symbiodinium* spp. made initial attempts at describing their diversity and ecology slow and challenging (Schoenberg & Trench 1980a, b, c; Trench & Blank 1987). The application of molecular genetic analyses helped to eliminate the long-standing paradigm that a single species, *Symbiodinium microadriaticum*, occurred among all species of host (Taylor 1974).

The application of sequence-based analyses has significantly changed perceptions of diversity, and revealed many biogeographical, ecological, and evolutionary relationships among *Symbiodinium* spp. and their hosts (Baker 2003; Coffroth & Santos 2005; LaJeunesse 2005). Phylogenetic reconstructions based on the small (SSU or 18S) and large subunit (LSU or 28S) ribosomal DNA separated the genus *Symbiodinium* into eight major clades (A–H) (Coffroth & Santos 2005). The existence of these evolutionarily divergent lineages has been supported by phylogenetic reconstructions using organelle (mitochondrial and chloroplast) DNA sequences (Santos *et al.* 2002; Takabayashi *et al.* 2004; Pochon *et al.* 2006). The combined molecular genetic evidence describes an ancient genus whose origins extend back many tens of millions of years (Tchernov *et al.* 2004; LaJeunesse 2005; Pochon *et al.* 2006).

Analyses of more rapidly evolving DNA sequences such as portions of the D1/D2 domain of the LSU rDNA (Baker & Rowan 1997; Loh *et al.* 2001; Rodriguez-Lanetty *et al.* 2001; Chen *et al.* 2005), internal transcribed spacer regions of the rDNA (ITS1 and ITS2; Hunter *et al.* 1997; Baillie *et al.* 2000a; LaJeunesse 2001; van Oppen *et al.* 2001), hypervariable regions of domain V of the chloroplast 23S rDNA (Santos *et al.* 2003b) and microsatellite flanking regions (Santos *et al.* 2004) indicate that *Symbiodinium* 'clades' are crude taxonomic designations. The ecological breadth and broad geographical distributions of *Symbiodinium* populations offers numerous opportunities for speciation. Not surprisingly, sequence comparisons of faster evolving genes further subdivide *Symbiodinium* diversity into units that reveal important ecological, biogeographical and phylogenetic patterns that are not evident at the 'clade' level (van Oppen *et al.* 2001; LaJeunesse *et al.* 2003; Santos *et al.* 2003a, 2004; LaJeunesse *et al.* 2004a, b; Thornhill *et al.* 2006; Warner *et al.* 2006; Sampayo *et al.* 2007, 2008). However, efforts to resolve and classify diversity below the cladal level have created some confusion because different research laboratories apply different screening techniques targeting different DNA regions. The reporting of sequences from different genes and/or use of different techniques rarely allows for cross-comparison between published findings. A universal

sub-cladal taxonomy, based on the consistent application of an accepted set of molecular genetic analyses would significantly advance coral–algal symbioses research in the areas of comparative physiology, ecology, biogeography and evolution. Determining the consistency and resolution of the many genetic markers and methods currently used in the study of *Symbiodinium* is therefore of critical importance.

The *Symbiodinium* diversity in *Stylophora pistillata*, *Pocillopora damicornis* and *Seriatopora hystrix* from the southern Great Barrier Reef was recently characterized using ITS2-denaturing gradient gel electrophoresis (DGGE) fingerprinting (Sampayo *et al.* 2007). Various symbiont 'types' were described that corresponded to particular host species, certain depth ranges, and differences in physiological tolerance to thermal stress (Sampayo *et al.* 2007, 2008). The well-characterized genetic, ecological, and physiological relationships described for this symbiont community provided a model system to evaluate the effectiveness and overlap of various molecular genetic analyses for investigating the ecology of animal–algal symbioses. Thirteen genetic analyses were applied to all samples. Genes and spacer regions of the ribosomal array (SSU, LSU, ITS1 and 2) as well as chloroplast (cp23S domain V), and mitochondrial (cytochrome–b mtDNA) genes were targeted and analysed using several techniques including restriction fragment length polymorphism (RFLP) digests, distance of migration in polyacrylamide gels, single-stranded conformation polymorphisms (SSCP), DGGE, extensive sequencing of bacterially cloned rDNA and direct sequencing (Fig. 1). When relevant, advantages and/or limitations associated with each methodology are discussed in the context of their ability to adequately resolve ecologically meaningful diversity. What emerges from these findings is a clear and robust phylogenetic delineation among *Symbiodinium* possessing distinct ecological and physiological attributes.

Materials and methods

Sample collection and processing

Three scleractinian host species, *Stylophora pistillata, Pocillopora damicornis*, and *Seriatopora hystrix* were sampled extensively in October 2004 at Heron Island on the southern Great Barrier Reef, Australia. *Stylophora pistillata* and *P. damicornis* associate with genetically distinct host-specific *Symbiodinium* depending on the reef habitat and location where individual colonies were collected while *S. hystrix* associates with a single symbiont at all depths (Sampayo *et al.* 2007). Fragments from these three host species were collected from the uppermost part of colonies growing in shallow reef habitats (3–5 m) and colonies living at their deepest distribution (16–19 m).
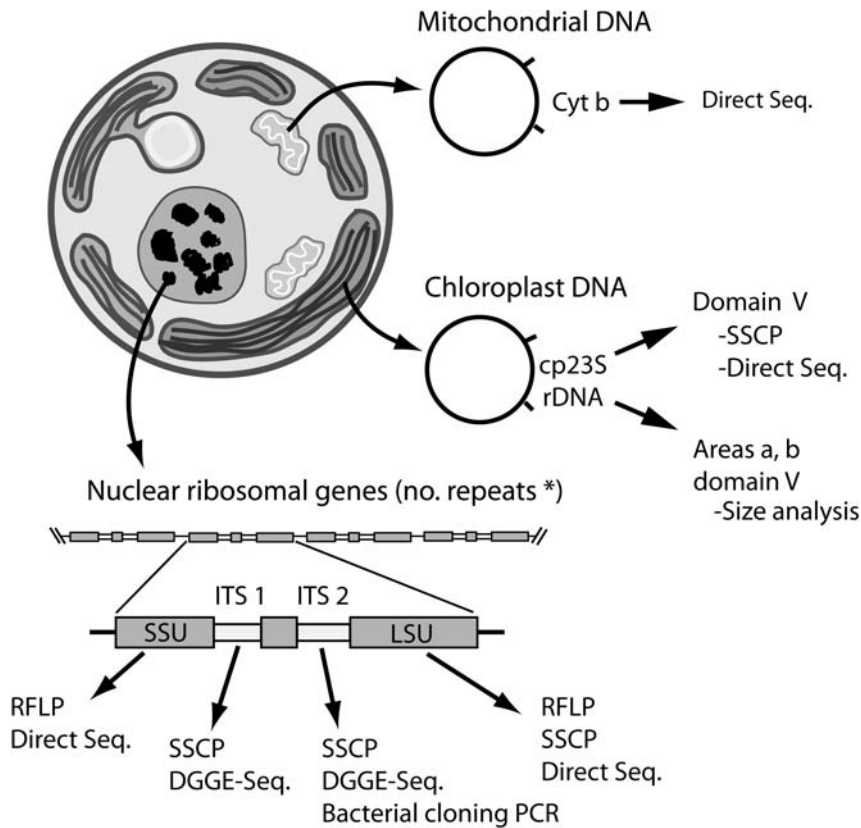
Fig. 1 Schematic of nuclear and organelle genomes targeted in the symbiotic dinoflagellate *Symbiodinium*. Nuclear ribosomal genes and spacer regions, including the small subunit (SSU or 18S), large subunit (LSU or 28S), and internal spacer regions (ITS1 and 2), as well as the chloroplast 23S rDNA domain V and hypervariable areas a and b of domain V, and mitochondrial cytochrome b (cyt *b*) gene were analysed using a variety of techniques including SSCP, DGGE, RFLP, fragment size analysis electrophoresis through polyacrylamide, direct sequencing and cloning/sequencing of PCR amplifications.

Tissue from each coral fragment was removed using an airgun attached to the low-pressure outlet of an underwater dive regulator. The slurry was centrifuged (6000 g, 10 min), the supernatant discarded and the algal pellet preserved in 20% dimethyl sulphoxide buffer (Seutin *et al.* 1991) and stored at –20 °C until further processing. Before DNA extraction, the preserved cells were pelleted by centrifugation (6000 g, 5 min), the preservation buffer removed and the pellet washed three times using DNAB. DNA was extracted using a QIAGEN Plant Mini Kit with inclusion of all optional steps suggested by the manufacturer.

RFLP of the small and large subunit rDNA genes

RFLP uses enzymes that digest the polymerase chain reaction (PCR) amplifications at specific recognition sites and produce differently sized fragments that are separable on agarose gels. The SSU rDNA (~1600–1700 bases) was amplified using primers ss5z (equimolar mix of ss5z1 and ss5z2) and ss3z following conditions of Rowan & Powers (1991; Table 1). The D1/D2 region of the LSU rDNA (~700 bases) was amplified with primer pair 28S-forward and 28S-reverse (Zardoya *et al.* 1995) following PCR cycling conditions of Loh *et al.* (2001; Table 1). For RFLP analysis of the SSU and LSU, 10 µL of each PCR amplification was digested with 1 µL *Taq*I enzyme (New England Biolabs)

for 1 h at 65 °C in 10 µL dH_2O and 2 µL enzyme buffer solution. Two additional restriction enzymes, *Hha*I or *Xho*I (New England Biolabs), were used separately on LSU rDNA amplifications and digested at 37 °C. For each digest, restriction fragments were separated by electrophoresis on 2% agarose gels for 2.5 h at 60 V alongside a 1 Kb plus DNA size ladder (Invitrogen), and post-stained with ethidium bromide.

Analysis of sequence length within the cp23S domain V hypervariable regions

Insertions and deletions (indels) cause size differences within the hypervariable domain (areas a and b) of the chloroplast large subunit (cp23S) domain V. Using polyacrylamide gels, small size differences can be detected and this methodology has been able to distinguish closely related members of clade A, B and C *Symbiodinium* (Santos *et al.* 2003a,b). The hypervariable domain (~200 bases) was amplified using primers $23S_{HYPERUP}$, $23S_{HYPERDNM13}$, and IRD800 fluorescently labelled M13-reverse primer (LICOR Biotechnology) following PCR conditions in Santos *et al.* (2003b; Table 1). PCR amplifications were electrophoresed (as outlined in Santos *et al.* 2003b) and screened on ultra-thin polyacrylamide gels (0.25 mm) using a LICOR NEN Global IR2 DNA Sequencer System and a LICOR IRDye 50–350 bp DNA marker for size reference across the gel.

Table 1 Gene regions targeted for analyses, primer pairs used for PCR, approximate size of amplified DNA fragment, and annealing temperature used to examine diversity in the genus *Symbiodinium*. For analysis of ITS regions using denaturing gradient gel electrophoresis, a GC-rich area (clamp) is attached to the reverse primer (underlined)

Region	Primer	Primer sequence (5′ to 3′)	Size	T_m
cp23S domain V areas a and b	23S$_{HYPERUP}$*	TCAGTACAAATAATATGCTG	~200 bp	55 °C
	23S$_{HYPERDNM13}$*	GGATAACAATTTCACACAGGTTATCGCCCCAATTAAACAGT		
cp23S domain V	23S1M13†	CACGACGTTGTAAAACGACGGCTGTAACTATAACGGTCC	~600 bp	55 °C
	23S2M13†	GGATAACAATTTCACACAGGCCATCGTATTGAACCCAGC		
Cytochrome *b*	cob1 (forward)‡	ATGAAATCTCATTTACAWWCATATCCTTGTCC	~950 bp	55 °C
	dinocob1 (reverse)‡	TCTCTTGAGGKAATTGWKMACCTATCCA		
18S rDNA	ss5z1§	GCAGTTATAATTTATTTGATGGTCACTGCTAC	~1600 bp	56 °C
	ss5z2§	GCAGTTATAGTTTATTTGATGGTTGCTGCTAC		
	ss3z§	AGCACTGCGTCAGTCCGAATAATTCACCGG		
28S rDNA	28S-forward¶	CCCGCTGAATTTAAGCATATAAGTAAGCGG	~630 bp	66 °C
D1/D2 domain	28S-reverse¶	GTTAGACTCCTTGGTCCGTGTTTCAAGA		
ITS1 rDNA — SSCP	symITSFP**	CTCAGCTCTGGACGTTGYGTTGG	~380 bp	59 °C
	symITSRP**	TATCGCRCTTCRCTGCGCCCT		
ITS1 rDNA — DGGE	ITS1CLAMP††	<u>CGCCCGCCGCGCCCCGCGCCCGTCCCGCCGCCCCCG</u>	~420 bp	62–52 °C (TD)
		CCCGGGATCCGTTTCCGTAGGTGAACCTGC		
	ITS1intrev2††	TTCACGGAGTTCTGCAAT		
ITS2 rDNA — SSCP	ITSintfor2‡‡	GAATTGCAGAACTCCGTG	~320 bp	59 °C
	ITS2-reverse‡‡	GGGATCCATATGCTTAAGTTCAGCGGGT		
ITS2 rDNA — DGGE	ITS2CLAMP‡‡	<u>CGCCCGCCGCGCCCCGCGCCCGTCCCGCCGCCCCCG</u>	~360 bp	62–52 °C (TD)
		CCCGGGATCCATATGCTTAAGTTCAGCGGGT		
	ITSintfor2‡‡	As above		

*Santos *et al.* 2003b; †2002; ‡Zhang *et al.* 2005; §Rowan and Powers 1991; ¶Zardoya *et al.* 1995; **van Oppen *et al.* 2005; ††LaJeunesse *et al.* 2008; ‡‡LaJeunesse 2002; TD, touchdown PCR protocol.

SSCP analysis of the cp23S domain V as well as the LSU D1/D2 domain and ITS (1 and 2) regions of the ribosomal genes

SSCP relies on the conformational (secondary structure) properties of denatured single-stranded DNA. Point mutations can change this conformation thus altering the migration dynamics of the fragment in polyacrylamide gels. The cp23S domain V (~700 bp) was amplified using the primers 23S1M13 and 23S2M13 under conditions reported by Santos *et al.* (2002). PCR products were mixed in a 1:1 ratio with formamide loading dye (9.5 mL formamide, 0.4 mL 0.5 м EDTA and 0.25% bromophenol blue/xylene cyanol), denatured at 94 °C for 5 min, and placed on ice before loading. Of each sample, 10 to 15 µL was loaded onto 6% polyacrylamide (37.5:1 acrylamide/bis) gels containing 2.5% glycerol. Single-stranded DNA fragments were electrophoresed for 5 h at 80 V in 0.5× TBE on a Hoefer SE 260 mini vertical electrophoresis unit (Hoefer Inc., CA) and post-stained with ethidium bromide.

SSCP was also employed to analyse the D1/D2 domain of the LSU (amplified as described in the RFLP section) and each of the internal transcribed spacer (ITS1 and ITS2) regions of the rDNA. The ITS1 rDNA was amplified using primers symITSFP and symITSRP (van Oppen *et al.* 2005; Table 1) under conditions outlined by van Oppen *et al.*

(2001). The ITS2 rDNA region was amplified with primer pair ITSintfor2 and ITS2-reverse (LaJeunesse & Trench 2000; Table 1) using PCR conditions described in Sampayo *et al.* (2007) but with a 5-min final extension step. Amplified fragments for both the ITS1 and ITS2 were mixed in a 1:1 ratio with formamide loading dye, denatured for 5 min at 94 °C and placed on ice before loading. Of each sample, 1 µL was loaded onto 4% polyacrylamide gel (37.5:1 acrylamide/bis) followed by a 30 to 45 second pulse after which wells were flushed with running buffer. Electrophoreses were conducted on a GelScan 2000 (Corbett Life Sciences, Australia) in 0.6× TBE at constant temperature (22 °C) and run for 45 min at 1200 V. The D1/D2 domain of the LSU was analysed using both the Hoefer and Gelscan 2000 SSCP systems under conditions mentioned above for the cp23S domain V (Hoefer system) and the ITS rDNA regions (GelScan 2000), respectively. On all SSCP gels, a nondenatured sample was run as a mobility reference for double-stranded DNA.

DGGE analysis of the ITS1 and 2 rDNA

DGGE separates double-stranded DNA and relies on the changes that single base-pair mutations cause in the dissociation (melting) point, which in turn influences the migration characteristics of DNA fragments. Dissociation

of double-stranded DNA is achieved by increasing gradients of denaturing chemicals in the direction of migration (Muyzer 1999). The ITS1 and ITS2 regions were amplified separately for analysis by DGGE with a modified primer containing a 40-bp GC-rich sequence ('GC-clamp') added to the 5′ end of one of the two normal primers (Table 1). The primer pair ITSintfor and ITS2CLAMP was used to amplify the ITS2 region under conditions specified by LaJeunesse *et al.* (2003; Table 1). Amplified ITS2 fragments were separated on a DGGE DCode system (Biorad) using 8% poly-acrylamide gels (37.5:1 acrylamide/bis) with an internal gradient of 30–60% denaturants (formamide and urea). Wells were rinsed with running buffer after which 10-μL sample was loaded with DCode loading dye (1:1 ratio). Gels were electrophoresed at 100 V for 14 h. The primer pair 'ITS1 clamp' and 'ITSintrev2' were used to analyse the ITS1 region under conditions in LaJeunesse *et al.* (2008; Table 1). This primer pair amplifies a region similar to that of symITSFP and symITSRP used in SSCP (van Oppen *et al.* 2005). ITS1 PCR amplifications were run on a C.B.S. Scientific DGGE system using 8% polyacrylamide gels (40:1 acrylamide/bis) with an internal gradient of 50% to 90% denaturants and separated by electrophoresis at 115 V for 15 h. All DGGE gels were stained for 20 min with 1× SYBR Green (AMRESCO or Invitrogen) in 1× TAE running buffer.

Bacterial cloning of PCR-amplified ITS2

The cloning of PCR amplifications of rDNA has been used extensively to evaluate genetic diversity of environmental microbial samples (mainly the 16S rDNA; reviewed in Morris *et al.* 2002; Rappé & Giovannoni 2003) and has recently been suggested as a screening method for *Symbiodinium* diversity (Apprill & Gates 2007). Here we tested the use of bacterial cloning of ITS2 rDNA to characterize diversity within samples and to compare these results to the other analyses performed in this investigation. Samples of each distinct *Symbiodinium* spp. as well as a single sample containing two distinct symbionts (based on investigations described above) were analysed using bacterial cloning of the ITS2 region. The ITS2 region was amplified using primers ITSintfor and ITS-reverse as described in Table 1. PCR products were cloned using the pGEM-T easy vector system (Promega) and TOPO One Shot competent cells (Invitrogen). At least 15 clones per sample were grown in Luria-Bertani (LB) medium, purified (UltraClean Mini Plasmid Prep Kit, Molecular Biology Laboratories) and directly sequenced at the Australian Genome Research Facility using vector specific forward and/or reverse primers.

The processes of PCR and cloning can generate a number of sequenced artefacts (Suzuki & Giovannoni 1996; Speksnijder *et al.* 2001). Here we excluded clones of heteroduplexes and chimeras from subsequent analysis by screening each clone on DGGE. Clones that contained heteroduplex PCR fragments appeared as a quartet on the DGGE gels and were not sequenced. Chimeras were excluded from sequence alignments of clones obtained from a single sample (as outlined in Thornhill *et al.* 2007). Sequence changes among variants were not analysed to determine whether they disrupted secondary structure and the data set may therefore include potential pseudo-genes as well as point mutations produced during the cloning process (Speksnijder *et al.* 2001; Thornhill *et al.* 2007). Analyses of the sequence database obtained from cloning of the ITS2 rDNA is outlined below.

Direct sequencing of ribosomal DNA, cp23S rDNA domain V, and mitochondrial cytochrome b *genes*

Three to four examples from each distinct ITS2-DGGE fingerprint identified by Sampayo *et al.* (2007) were analysed using direct bidirectional sequencing of the D1/D2 domain of the LSU rDNA (Loh *et al.* 2001), the cp23S domain V (Santos *et al.* 2002) and the mitochondrial cytochrome *b* gene (amplified using primer pair Dinocob1F and Dinocob1R following PCR cycling conditions outlined in Zhang *et al.* 2005; Table 1). The SSU was also sequenced but, due to its length, only partial sequences were obtained with bi-directional sequencing. SSU sequences were therefore only used to verify the *Symbiodinium* cladal designation and not for phylogenetic analysis. Direct sequencing of the ITS1 and ITS2 rDNA regions was carried out from excised bands from DGGE fingerprints. Dominant bands in each profile were excised, re-amplified using the non-GC reverse primer (LaJeunesse *et al.* 2003; Table 1) and PCR purified (UltraClean PCR purification Kit, Molecular Biology Laboratories). All samples were sequenced in separate runs at the Australian Genome Research Facility (AB3730xl capillary sequencer) using the forward or reverse primers (for ITS1 and ITS2 using the non-GC primers). Chroma-tograms were visually inspected (SEQMAN) and sequences were aligned using Clustal X (version 1.83; Thompson *et al.* 1997).

Phylogenetic analyses

Phylogenetic reconstructions based on direct sequences from the cp23S domain V, cytochrome *b* mtDNA, the LSU D1/D2 domain, the ITS1 and ITS2 rDNA (from excised dominant bands) as well as the sequences obtained from cloning of the ITS2 were constructed using PAUP (version 4.0b10; Swofford 2000) under the maximum parsimony criterion (delayed transition). Deletions were incorporated as a fifth character state and indels were represented as a single evolutionary event. To determine whether the data sets could be analysed in concatenation, a 'partition-homogeneity' or 'incongruence length difference (ILD)' test

was run. Phylogenetic incongruence indicates that different partitions (or gene regions) undergo a distinct evolutionary history (Farris *et al.* 1995). Although the ILD test has limitations (Ramírez 2006), it showed phylogenetic incongruence of all DNA regions ($P = 0.01$) and supported the concatenated analyses. Maximum likelihood analysis was performed on all gene regions separately as well as the concatenated analyses. The results provided similar topologies to the maximum parsimony analysis (see also LaJeunesse 2005; Loh *et al.* 2006). Therefore, and because maximum parsimony analysis allows for the inclusion of evolutionarily relevant indels, only data from maximum parsimony analyses are presented. All bootstrap values were calculated based on 1000 replicates. Finally, uncorrected genetic distance was calculated (PAUP) within cloned ITS2 sequences of a single *Symbiodinium* as well as between distinct *Symbiodinium* 'types' using the most derived dominant repeat.

Results

The 13 genetic analyses applied to samples from three species of pocilloporid corals identified a maximum of eight genetically distinct *Symbiodinium* spp. exhibiting ecological and functional differences (Sampayo *et al.* 2007, 2008). However, the ability to separate each symbiont was dependent on the technique employed and gene region analysed. The extent of diversity characterized by each analysis is summarized in Fig. 2. DGGE of the ITS2 region, SSCP or direct sequencing of the D1/D2 domain of the LSU rDNA as well as ITS1-DGGE provided the highest and clearest level of genetic resolution. These analyses resolved four symbionts specific to *Stylophora pistillata* (C8a, C35a, C78a, C79), three specific to *Pocillopora damicornis* (C33, C33a, C42a) and one specific to *Seriatopora hystrix* (C3nt). ITS2-DGGE, 28S-SSCP, and bacterial cloning identified mixed symbionts present in one of the *P. damicornis* colonies (containing 'types' C33a and C42a).

In order to reduce confusion when comparing the results of each analysis, an alphanumeric designation based on ITS2-DGGE nomenclature was applied retrospectively (cf LaJeunesse 2002) and 'type' designations correspond to taxonomic terms used in a previous publication (Sampayo *et al.* 2007; C35a = Sp-1; C78a = Sp-2; C8a = Sp-3; C79 = Sp-4; C42a = Pd-2, C33a = Pd-4, 4*; C33 = Pd-5, 5*; C3nt = Sh). The use of the ITS2-DGGE nomenclature places these symbionts in the context of other *Symbiodinium* 'types'[1] that have been described in this manner around the world.

---

[1] Symbiont 'type' or ITS2 type: a genetic designation well below the taxonomic level of *Symbiodinium* clade. 'Type' nomenclature is derived from the sequence, or sequences, of a band, or bands, that are diagnostic of a repeatable PCR-DGGE fingerprint. Based on ecological, physiological, biogeographical, and genetic data (this paper included), these 'types' are probably distinct species.

RFLP of the small and large subunit rDNA genes

RFLP analysis of the complete SSU rDNA identified four distinct profiles; two unique to *S. pistillata*, one unique to *P. damicornis*, and a fourth profile that is shared between the three host species *S. pistillata*, *P. damicornis* and *S. hystrix* (Fig. S1a,b,h, Supporting information). Only one profile directly matched the characteristic clade C restriction pattern (*sensu* Rowan & Powers 1991; Fig. S1a, b, h, supporting information) while another profile was similar to that of *Symbiodinium* clade A (C35a in Fig. S1a compare to clade A pattern in Rowan & Knowlton 1995; Santos *et al.* 2002). Partial sequencing showed that all samples belonged to clade C *Symbiodinium*. SSU-RFLPs resolved *S. pistillata* symbionts C35a and C79 as well as *P. damicornis* symbiont C42a but did not distinguish C8a, C78a, C33, C33a, and C3nt from each other (Fig. 2).

Restriction digests using *Taq*I, *Xho*I and *Hha*I on amplified *Symbiodinium* LSU rDNA identified a total of three distinct profiles (Fig. 2; Fig. S1c–h, supporting information). Results for the three enzymes combined distinguished in total five symbionts (Fig. 2). Analyses of LSU restriction digests using only one enzyme provided mixed results (Fig. S1c–h, supporting information) indicating that the choice of enzyme affects resolution of ecological patterns. The only symbiont that was consistently resolved independent of the enzyme used was *P. damicornis* symbiont C42a. Depending on the enzyme, either *S. pistillata* symbionts C35a, C78a, or C8a were resolved by *Taq*I, *Xho*I or *Hha*I, respectively (Fig. S1c–h, supporting information). Combining the results from all three restriction enzymes still could not distinguish certain *Symbiodinium* spp. (cf. Sampayo *et al.* 2007) found in *S. pistillata* (C79) from those found in *P. damicornis* (C33, C33a) and *S. hystrix* (C3nt) (Fig. 2).

Analysis of sequence length within the cp23S domain V hypervariable regions

No apparent length differences were detected among amplifications of the hypervariable region (areas a and b) within the cp23S domain V (Fig. 2; Fig. S2a, b, Supporting information). Size was calculated at 198 bases and later verified from direct sequencing of the complete cp23S domain V (see section below). Based on this analysis, it would incorrectly appear that a single *Symbiodinium* sp. occurs in the three host species at all depths and reef locations surveyed at Heron Island (Sampayo *et al.* 2007).

SSCP analysis of the cp23S domain V as well as the LSU D1/D2 domain and ITS1 and 2 regions of the ribosomal genes

No differences were observed in the SSCP analysis of amplifications of the chloroplast large subunit (Fig. 2,
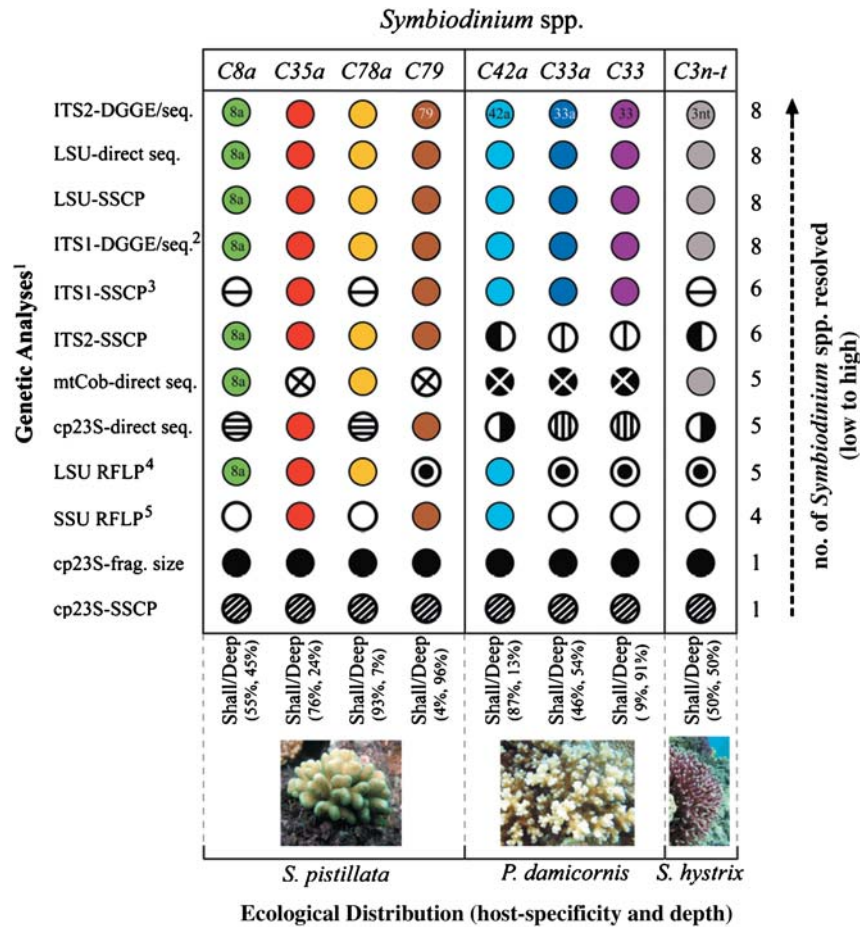
Fig. 2 Summary of genetic resolution provided by the analysis of rDNA, chloroplast, and mitochondrial genes using a variety of screening techniques. Each distinct *Symbiodinium* was given the 'type' designation based on ITS2-DGGE nomenclature. Distinct *Symbiodinium* 'types' exhibit ecological differentiation by associating with specific hosts or occupying a distinct depth range. Ecological data were obtained from Sampayo *et al*. 2007 and relative abundances of symbionts are shown in parenthesis by each depth range (percentage values are the relative abundance of a given symbiont at shallow or deep depths). For *Stylophora pistillata* symbionts C8a, C35a, C78a and C79 (n = 77, 41, 28, 71, respectively) percentages were based on data from three sites; for *Seriatopora hystrix* symbiont C3nt (n = 56) and *Pocillopora damicornis* symbionts C42a, C33a, and C33 (n = 31, 41, 11, respectively) values were based on two sites [Sampayo *et al*. 2007; data from Pd1 and Pd2 (C42a) as well as Pd4, 4* (C33a) and Pd5, 5* (C33) were grouped]. A colour circle is featured when the genetic analysis resolves a symbiont as being unique from the others. For each row, circles with identical black and white patterns indicate that the technique did not resolve differences between them. For example, sequences from the mitochondrial cytochrome *b* gene did not resolve C35a from C79 or between C33, C33a, and C42a. Therefore, this marker identified only five distinctive taxa. [1]Note that the cloning of ITS2-PCR fragments is not shown here but is represented in Fig. 3 as a separate analysis.[2] ITS1-DGGE in combination with sequencing only separated C42a from C3nt and C35a from C79 based on sequence identity. [3]ITS1 primers differed between SSCP and DGGE protocols. [4]LSU RFLP based on the combined results from three independent restriction digests using *Taq*I, *Xho*I and *Hha*I. [5]SSU RFLP based on restriction digests using *Taq*I.

Fig. S3a, Supporting information). The larger size of this amplified product (~585 bases) may limit the technique's ability to resolve sequence differences among the particular sequences analysed (Sunnucks *et al*. 2000).

SSCP analyses of the LSU, ITS1 and ITS2 amplifications resolved most of the eight *Symbiodinium* 'types' (Fig. S3b–e, supporting information) but resolution varied depending on the rDNA region targeted (Fig. 2). Fingerprint patterns from ITS1-SSCP (Fig. S3d), and LSU-SSCP (Fig. S3e) consistently identified three distinct *Symbiodinium* in *P. damicornis* (C42a, C33, and C33a) while ITS2-SSCP detected only two

(joining C33 and C33a, Fig. 2; Fig. S3e). Furthermore, ITS2-SSCP could not distinguish *Symbiodinium* C3nt from C42a (Fig. 2; Fig. S3e) and ITS1-SSCP could not distinguish C8a, C78a and C3nt (Fig. 2; Fig. S3d) from each other despite numerous sequence differences between the two (Fig. 4a, b).

DGGE-sequencing analysis of the ITS1 and ITS2 rDNA regions

ITS2-DGGE and ITS1-DGGE fingerprinting in combination with direct sequencing of excised bands resolved eight

Table 2 Summary of cloned ITS2 sequence variants obtained from separate PCR amplifications conducted on representative samples of eight different *Symbiodinium* spp. Cloned heteroduplexes and chimeric sequences were identified and removed from further analysis as described by Thornhill *et al.* (2007). Remaining clones were used for phylogenetic analysis in Fig. 4 and calculations of the percentage of unique variants retrieved for each distinct *Symbiodinium* spp. Average uncorrected genetic distance is calculated from clones retrieved from each *Symbiodinium* spp. separately and is based on the ITS2 region alone (195 aligned bases). HD, heteroduplexes.

Host species	*Symbiodinium* spp. (initial no. of clones)	Percentage of cloned HD (n)	Percentage of chimeras (n)	Clones analysed (n)	Percentage of unique variants (n)	Uncorrected genetic distance (P) ($\pm$ SD)
S. pistillata	C8a (21)	33 (7)	0 (0)	14	50 (7)	0.0140 ($\pm$ 0.0081)
	C35a (24)	17 (4)	0 (0)	20	60 (12)	0.0132 ($\pm$ 0.0161)
	C78 (16)	0 (0)	19 (3)	13	54 (7)	0.0185 ($\pm$ 0.0118)
	C79 (22)	14 (3)	0 (0)	19	63 (12)	0.0124 ($\pm$ 0.0066)
P. damicornis	C33 (18)	5 (1)	17 (3)	14	43 (6)	0.0114 ($\pm$ 0.0073)
	C33a (18)	0 (0)	11 (2)	16	38 (6)	0.0145 ($\pm$ 0.0071)
	C42a (57)	9 (5)	5 (3)	49	49 (24)	0.0144 ($\pm$ 0.0097)
	C42a+C33a (35)	9 (3)	14 (5)	27	44 (12)	0.0175 ($\pm$ 0.0086)
S. hystrix	C3nt (32)	8 (3)	8 (3)	26	42 (11)	0.0117 ($\pm$ 0.0069)

genetically distinct symbionts from the three pocilloporid species analysed (Fig. 2; Fig. S4c, Supporting information). However, the fingerprints produced from ITS1-DGGE were not as distinct as those from ITS2-DGGE (compare Fig. S4a with S4c). Although ITS2-DGGE produced markedly different fingerprints for all eight *Symbiodinium* spp., ITS1-DGGE produced similar profiles for *Symbiodinium* C35a and C79 as well as C42a and C3nt (Fig. S4a). Upon close examination of the profiles, however, faint banding differences are visible between the ITS1-fingerprints of different 'types' (Fig. S4b). Excision and direct sequencing of the dominant bands from each profile confirmed that these were different *Symbiodinium* (see also Fig. 4a, b).

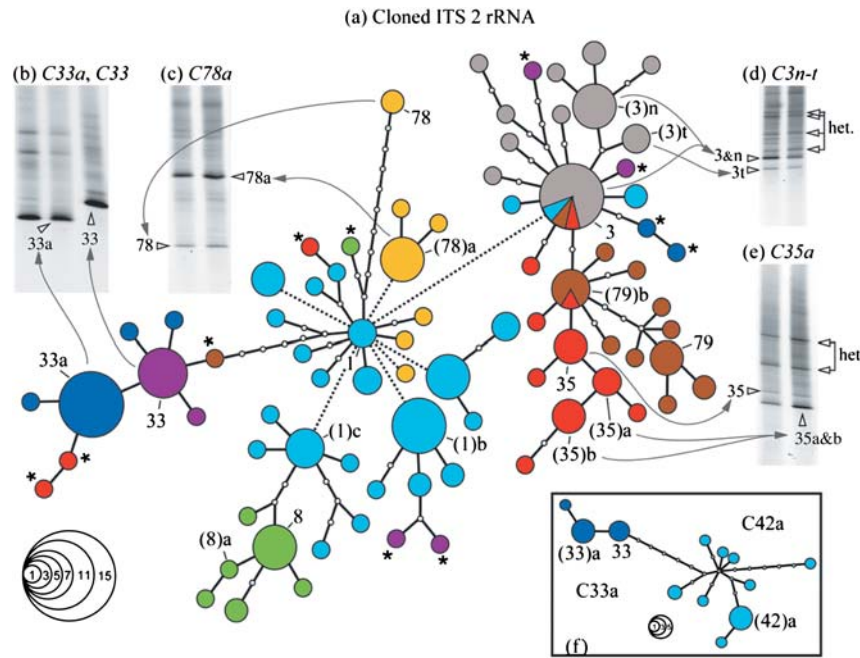### Bacterial cloning and sequencing of ITS2 PCR amplifications

Bacterial cloning of ITS2 PCR amplifications produced numerous sequence variants. For each distinct *Symbiodinium* 'type' characterized by the methods described above, between 16 to 57 clones were sequenced. Chimeric sequences (Thornhill *et al.* 2007) or heteroduplexes produced during PCR were removed from the database before phylogenetic analysis. The remaining sequences showed that 38% to 63% of retrieved clones within each *Symbiodinium* 'type' were unique (Table 2). The percentage of unique variants did not appear to be influenced by the total number of clones sequenced (Table 2) suggesting that the discovery of new sequences from each clone library had not reached saturation. The average uncorrected genetic distance (p) within a single *Symbiodinium* 'type' ranged from 0.0117 to 0.0185 calculated from the ITS2 region alone (195 aligned bases after removing partially amplified 5.8S and LSU; LaJeunesse 2005; Table 2).

For *Symbiodinium* 'types' C35a, C79, C3nt, C78a, C33a and C33, the most frequently found clone sequence, or sequences, corresponded to the dominant band, or bands, of the DGGE fingerprint profile (Fig. 3b–e). Differences in brightness among the dominant bands also appeared to correspond with the frequency that a particular sequence was recovered during cloning. The only exceptions were *Symbiodinium* C8a and C42a. It appears that C42, especially, possesses a poorly homogenized genome and contains many codominant and rare variants. For 'type' C8a, additional sequencing of clones may have recovered more of the codominant sequence variant, 'a'.

The brightest band in the gel profiles of C35a, C79 and C3nt actually contained two co-migrating bands. These represent codominant variants that differed by 1, 4, and 2 bases, respectively (Fig. 3d, e). Co-migrating bands were identified based on (i) the existence and sequencing of heteroduplexes found higher in the gel profile, (ii) the occurrence of ambiguous base positions from direct sequencing of the band, and (iii) when analysing re-amplifications of the excised band, or from the analysis of cloned fragments, using DGGE (see below). If co-migrating bands or heteroduplexes were excised and rerun or heteroduplexes were cloned the DGGE profile produced a triplet involving the co-migrating homoduplexes and their two accompanying heteroduplexes.

The maximum parsimony analysis of cloned sequences identified eight distinct sequence clusters (as represented by different colours; Fig. 3) that corresponded with each distinct *Symbiodinium* 'type' identified by ITS2-DGGE, ITS1-DGGE and LSU-SSCP (Fig. 2). For each distinct 'type', the most commonly retrieved sequence variant(s) occupied the centre of a cluster and were surrounded by rare singletons (i.e. Fig. 3b–e). The degree of sequence differences and the number of codominant sequences within each cluster,

Fig. 3 (a) Unrooted phylogeny (maximum parsimony) of ITS2 sequence variants obtained through bacterial cloning of PCR amplifications from eight genetically distinct *Symbiodinium* spp. (Fig. 2). (GenBank Accession nos FJ529570–FJ529667). Sequences obtained from each are indicated by a different colour: C35a = red, C78a = yellow, C8a = green, C79 = brown, C42a = light blue, C33a = dark blue, C33 = purple, C3nt = grey. Each 'type' is ecologically and physiologically different as each exhibits differences in host specificity, depth zonation, and thermal tolerance (LaJeunesse 2002; LaJeunesse *et al*. 2003; Sampayo *et al*. 2007; Sampayo *et al*. 2008). Sequence variants generated through cloning show the extent of phylogenetic overlap of high- and low-copy number intragenomic variants. Stippled branches represent a 1 base change and circle size represents the number of variants recovered with an identical sequence. Sequence variants with asterisks indicate anomalous positions in the unrooted phylogeny. (b, c, d, e) ITS2-DGGE analyses of *Symbiodinium* C33a, C33, C78a, C3nt and C35a show how this technique translates the presence of dominant intragenomic variation into a distinctive gel fingerprint, providing a visual barcode distinctive of each symbiont. Direct sequencing of the dominant band(s) in a DGGE fingerprint profile typically corresponds to the sequence of the numerically dominant clone (Thornhill *et al*. 2007). Heteroduplexes (het) are produced when intragenomic codominant sequence variants, differing by one or two base changes, anneal to each other after the disassociation phase of the PCR. (f) Analysis of cloned sequences from one mixed sample that contained fingerprint profiles for both C33a and C42a.

however, was variable depending on the symbiont analysed (Fig. 3, Table 2). *Symbiodinium* C42a contained the highest number of frequently retrieved codominant variants, including C1, C1c, C1b and C42a, and many unique sequences (singletons). Coincidentally, this particular 'type' produced an ITS2-DGGE fingerprint with the most codominant bands (Fig. S4c).

There was limited phylogenetic overlap observed among sequence variants from each *Symbiodinium* 'type' (determined by each colour in Fig. 3). The sequence clusters of C35a and C79 showed some overlap but only involved shared ancestral sequences. In several instances, certain variants were anomalously positioned well outside the sequence cluster of the sample (designated by asterisks, Fig. 3) but these were all singletons and did not overlap with any of the more numerically dominant retrieved sequences from other *Symbiodinium* spp. in this study. As such, these rare variants are not necessarily diagnostic of the existence of another symbiont 'type' present in the

sample and may represent pseudogenes or distant rare intragenomic sequences.

Phylogenetic analysis of directly sequenced ribosomal and organellar genes

Attempts at direct sequencing of the ITS region had limited success. For this reason, ITS1 and ITS2 sequences derived from excised diagnostic bands of a DGGE profile and representative of the dominant ribosomal repeat (see discussion and Fig. 3b–e) were used in the phylogenetic analyses. In many cases, one symbiont 'type' is represented in the phylogeny by one to three codominant repeats in the ITS region. Direct sequencing of the LSU was mostly successful and a total of eight different sequences were produced that corresponded to each of eight distinct LSU-SSCP profiles (Fig. 2, Fig. S3b). The presence of codominant variants or samples with two or more symbionts will produce ambiguous sequence reads when direct

sequencing is attempted. Sequence chromatograms of the large subunit from *Symbiodinium* C42a contained many background peaks and made portions of these sequences difficult to read. This was overcome by comparing multiple sequences from different samples of this 'type.' Through this comparison, and because some samples sequenced better than others, the dominant sequence was determined.

Organelle genes (cytochrome *b* mtDNA, cp23S) sequenced cleanly with direct sequencing from PCR products, suggesting that the amplified product was a homogeneous template. Sequences of the cp23S domain V verified that no length variability (heteroplasmy) existed within the hypervariable areas (a and b) but phylogenetically informative sequence differences were present outside this region among the *Symbiodinium* 'types' examined. A pseudogene containing a large 380 base deletion of the cp23S domain V was co-amplified along with the 'normal' cp23S in samples with C35a and C79, supporting the close phylogenetic relationship between these *Symbiodinium* spp. (see discussion). Sequences of the cp23S rDNA and cytochrome *b* mtDNA resolved about half the total diversity revealed by rDNA regions (compare Fig. 4a–c to Fig. 4d, e).

The branch topologies among independent phylogenetic reconstructions, based on organelle and nuclear gene regions, were congruent (Fig. 4a–e). In all cases, *Symbiodinium* spp. estimated to be closely or distantly related to each other by one gene sequence were supported by sequences from other gene regions. Sequences of genes that did not resolve ecologically distinct *Symbiodinium* spp. often were ancestrally positioned in the phylogeny. In situations where the chloroplast phylogeny failed to resolve between ecologically distinct symbionts, the mitochondrial gene resolved them, and vice versa. Therefore, when these organelle gene sequences were combined, they produced a phylogeny identical to a phylogenetic reconstruction based on the combined ITS1 and ITS2 regions (Fig. 5a). An evolutionary reconstruction based on the combined sequence data produced a deeply branched phylogeny (Fig. 5b). The congruence among phylogenetic reconstructions based on sequences from nuclear and organelle genomes and the sharing of the same sequences among representative samples of each 'type' indicate that genetic recombination does not occur among the eight *Symbiodinium* spp. analysed in this study.

Discussion

Sequence comparisons can provide a precise way to distinguish among morphologically similar but ecologically distinct and evolutionarily diverged micro-organisms. Depending on their level of conservation, not all gene sequences and the techniques used to analyse them exhibit the same resolution. Various approaches are currently used to analyse and quantify the diversity of dinoflagellate endosymbionts (Coffroth & Santos 2005). However, there are few standardized methods for accurately classifying this diversity into ecologically and evolutionarily relevant units. Concerns about the lack of consistency and desires to unify the community of investigators in employing similar approaches are important issues in the research of microbes (Konstantinidis & Tiedje 2005; Green & Bohannen 2006; Cohan & Perry 2007; De Queiroz 2007; Alverson 2008; Koeppel *et al*. 2008).

In order to build consensus among researchers of animal–algal symbioses, numerous molecular genetic analyses targeting nuclear, chloroplast and mitochondrial gene sequences were employed to investigate their ability to discern diversity among dinoflagellate endosymbionts in the genus *Symbiodinium*. The gene regions and methods of analysis employed differed in their ability to resolve important ecological patterns among *Symbiodinium* sampled from different hosts and reef environments (Fig. 2). While the distinct methodologies distinguished different levels of diversity, each was complementary with others that provided the same or greater resolution. While each technique has advantages and limitations that are summarized partially in Table 3 and discussed later in more detail, some were clearly better than others for purposes of ecological research and species delineation.

In summary, several independent approaches converged on the same estimate of diversity. DGGE of the ribosomal ITS1 and ITS2 spacers in combination with direct sequencing of the dominant bands, as well as SSCP and direct sequencing of the D1/D2 domains of the LSU provided consistent resolution among ecologically and physiologically different *Symbiodinium* spp. (Sampayo *et al*. 2007, 2008). Sequences from the chloroplast 23S rDNA domain V and the mitochondrial cytochrome *b* were congruent with nuclear ribosomal sequence data (Fig. 5a), but when analysed independently were too conserved to discern all the ecologically relevant diversity. Most importantly, when all sequence data were combined, phylogenetic divergence increased without additional differentiation (Fig. 5b). These data, combined with ecological and physiological attributes, indicate that a lineage-based species concept (de Queiroz 2007; Alverson 2008) for identifying and classifying *Symbiodinium* spp. is possible using existing molecular genetic approaches.

Nuclear ribosomal genes, intragenomic diversity, and analysis techniques

Ribosomal DNA sequences are commonly used to examine diversity and infer evolutionary relationships among prokaryotic and eukaryotic lineages (Hillis & Dixon 1991; Morris *et al*. 2002; Rappé & Giovannoni 2003; Avise 2004). Indeed, ribosomal small subunit (SSU) sequencing provided the first phylogenetic (evolutionary) framework and revealed major subdivisions (i.e. clades) in the genus *Symbiodinium*
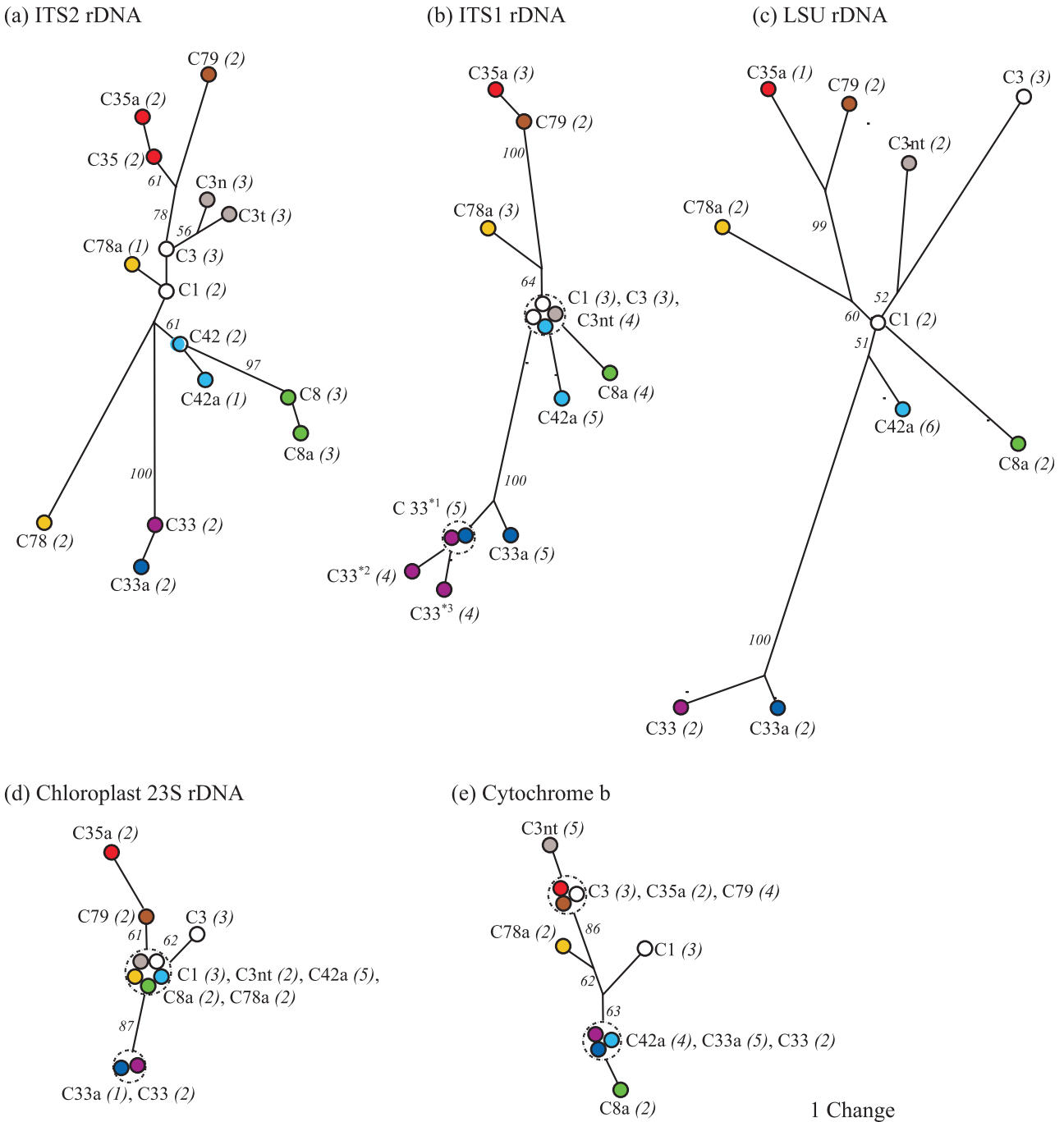
Fig. 4 Comparison of unrooted phylogenetic reconstructions based on maximum parsimony of the (a) ITS2 rDNA (obtained from PCR-DGGE fingerprints; 287 aligned bases including partial 5.8S). Many of the *Symbiodinium* spp. in this study contained similar codominant ITS2 sequence variants that are shown together in the phylogeny. (GenBank Accession nos ITS2; FJ529570, FJ529583–84, FJ529594–96, FJ529605–06, FJ529612–13, FJ529620, FJ529627, FJ529634, FJ529656); (b) ITS1 rDNA (obtained from PCR-DGGE fingerprinting; 329 aligned bases including partial 5.8S) (GenBank Accession nos ITS1; FJ529556–FJ529569). C33*1,2,3 indicate codominant rDNA repeats and identity of each is specified in Fig. S4b; (c) D1/D2 domain of the large subunit rDNA (607 aligned bases) (GenBank Accession nos cytochrome; FJ529534–FJ529544); (d) chloroplast 23S domain V (579 aligned bases) (GenBank accession nos CP23S; FJ529545–FJ529555); (e) mitochondrial cytochrome *b* (931 aligned bases) (GenBank Accession nos LSU; FJ529523–FJ529533). Numerals in parentheses indicate the total number of individual samples analysed. All branch lengths were drawn to the same scale. Distinct *Symbiodinium* 'types' are shown in separate colours corresponding to those used in Figs 2 and 3. For certain gene phylogenies, coloured circles designating different *Symbiodinium* 'types' were clustered together in instances when sequences were identical. Sequences for C1 (cultured *Symbiodinium goreaui*) and C3 from the Pacific coral *Acropora* spp. were also included for purposes of phylogenetic comparison.

(a) ITS1 & 2 rDNA vs. cp23S & mtCob
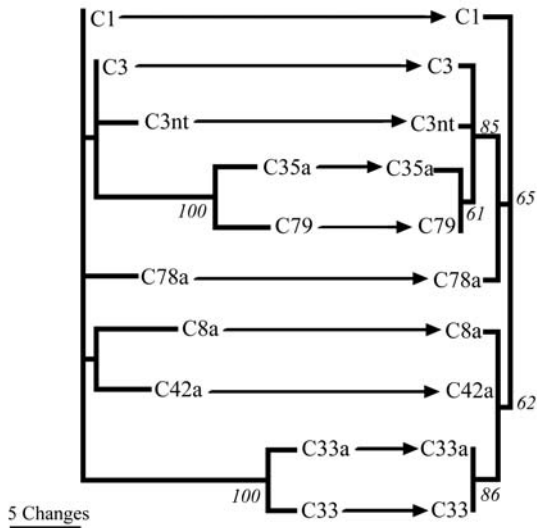
(b) Complete composite
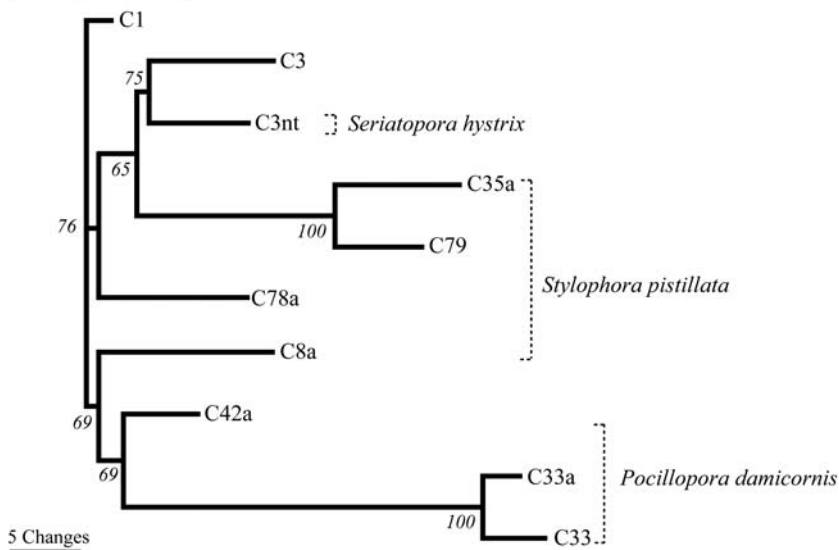
Fig. 5 (a) Similarity of maximum parsimony phylogenies obtained from concatenated (composite) nuclear (ITS1 and ITS2) vs. concatenated organellar sequences (mitochondrial cytochrome *b* and chloroplast 23S rDNA). The ITS1 and ITS2 phylogeny was based on 616 aligned bases with 59 variable characters of which 25 were parsimony informative. The mtCob and cp23S phylogeny was based on 1519 aligned bases with 14 variable characters and 7 parsimony informative characters. (b) Phylogenetic reconstruction based on maximum parsimony analysis of concatenated ITS1, ITS2, LSU rDNA, mtCob and cp23S sequences (2726 aligned bases) containing 120 variable characters of which 55 were parsimony informative. This phylogeny is largely congruent with the topology of individual rDNA gene phylogenies (Fig. 4) and contains long branches unambiguously resolving each of the *Symbiodinium* 'types' analysed in this study as distinct lineages. Bootstrap values > 50 are shown and branch length are drawn to the same scale.

(Rowan & Powers 1991). This was followed by the use of SSU and LSU RFLPs to characterize ecologically distinct *Symbiodinium* spp. in Caribbean and Eastern Pacific corals (Rowan & Knowlton 1995; Baker & Rowan 1997). The use of this technique is limited to identifying symbionts to the clade level although some additional resolution can be obtained by using multiple restriction enzymes (Fig. 2; Table 3). Nonetheless, these analyses do not resolve many ecological and biogeographical patterns especially among Indo-Pacific reef systems where cnidarian hosts are dominated by closely related members of clade C *Symbiodinium* (Baker 2003; LaJeunesse *et al.* 2003; Chen *et al.* 2005). Finally, digest profiles sometimes erroneously place a symbiont into the wrong lineage, or clade (SSU–RFLP of C35a, Fig. S1; Burnett 2002). These limitations have led

to the development of genetic approaches that use less conserved gene regions and more sensitive gel-based analyses (Baillie *et al.* 2000a; LaJeunesse 2001; Loh *et al.* 2001; van Oppen *et al.* 2001; Santos *et al.* 2003b).

Estimations of diversity among and within eukaryotic microbial groups using nuclear ribosomal DNA sequence data are potentially confounded depending on how these sequences are recovered (Thornhill *et al.* 2007). Direct sequencing compared with bacterial cloning from PCR amplifications produces results that are sometimes difficult to reconcile when carried out independently of each other (e.g. excision and direct sequencing of dominant bands from PCR-DGGE profiles compared to cloning and sequencing, LaJeunesse *et al.* 2004b; Apprill & Gates 2007). High levels of intragenomic sequence variation, including

many low-copy-number functional and nonfunctional rDNA sequence variants recovered by cloning account for much of this confusion (Thornhill *et al.* 2007).

Overcoming the confusion surrounding rDNA variation is simply a matter of understanding the molecular evolutionary processes that cause change and/or lasting stability in the dominant sequence variant(s) of this multi-copy gene region (Dover 1982; LaJeunesse & Pinzón 2007). The ribosomal genes in eukaryotes are organized into large arrays of tandem repeats. Each repeat, or cistron, contains several coding regions that comprise the SSU, 5.8S and LSU ribosomal genes and spacer regions (ITS1 and ITS2) that are spliced out of the transcribed rRNA molecule during processing of the mature ribosome (Long & Dawid 1980; Appels & Honeycutt 1986). Concerted evolution continuously homogenizes the rDNA cistrons but this process is imperfect and most eukaryotic genomes contain some degree of sequence variability, especially in less-conserved domains (e.g. Schaal *et al.* 1991; Buckler *et al.* 1997; Toller *et al.* 2001; Thornhill *et al.* 2007). Typically, one or several sequence variants, which differ usually by one or two base changes, dominate a genome (Arnheim 1983; Ohta & Dover 1983). The level of homogeneity and rate of evolutionary change to a different dominant sequence depends on the total number of repeats in the genome and whether repeats exist on different chromosomes. Ultimately, the stability of a dominant sequence among individuals of a population, or its replacement by a derived sequence variant depends on population level processes including the frequency of sexual recombination, effective population size, and rates of generational turnover (Dover 1986).

The ribosomal ITS1 and 2 regions as well as the D1/D2 domain of the LSU consistently contain fixed sequence differences among *Symbiodinium* spp. from specific hosts, water depths and geographical locations when screened using DGGE or SSCP in combination with direct sequencing (Table 2; LaJeunesse 2001, 2002; Loh *et al.* 2001; Toller *et al.* 2001; van Oppen *et al.* 2001; Barneah *et al.* 2007; Pochon *et al.* 2007; Frade *et al.* 2008; Goulet *et al.* 2008; Stat *et al.* 2008; Thornhill *et al.* 2008). Analyses of rDNA using SSCP (van Oppen *et al.* 2001; Loh *et al.* 2006; Mostafavi *et al.* 2007) or DGGE (Baillie *et al.* 2000a; LaJeunesse 2002; Thornhill *et al.* 2006) identify the dominant rDNA sequence variants and in doing so, produce repeatable fingerprints diagnostic of particular *Symbiodinium* spp. (the many rare intragenomic functional or nonfunctional variants present within a genome are not visualized on SSCP and DGGE gel profiles; LaJeunesse 2002; Thornhill *et al.* 2007). When two or more symbionts are abundant in a sample (> 10% of the population; Thornhill *et al.* 2006), the fingerprint profiles of each are identifiable in the same lane (LaJeunesse 2002). In terms of nomenclature, genetically and ecologically distinct symbionts are currently distinguished by fingerprinting techniques (such as ITS1-DGGE and ITS2-DGGE, e.g. LaJeunesse 2002; LaJeunesse

et al. 2008) based on characteristic (complete) profiles. Distinct fingerprints are assigned an alphanumeric name based on the sequence, or sequences of the dominant copy, or copies, in the ribosomal array.

Methodologies such as SSCP and DGGE have advantages and limitations, some of which are shared and some of which are method specific (Table 3). Both SSCP and DGGE are limited by the size of the DNA fragment that is analysable (Sunnucks *et al.* 2000). In comparison of each technique, the equipment required for SSCP is cheaper and electrophoresis time is shorter (see Table 3). DGGE is technically more challenging but produces considerably clearer profiles comprising bands that are resolved and separated with greater precision (compare Fig. S3 and Fig. S4) and thus DGGE analyses provided a better resolution of diversity (compare ITS1 and ITS2 SSCP with DGGE; Fig. 2). Excision and direct sequencing of prominent bands is an essential component of fingerprinting techniques (Table 3). This process is required for DGGE (e.g. LaJeunesse 2002) and is now used successfully with SSCP gels (Mostafavi *et al.* 2007; Stat *et al.* 2008). An important caveat in the use of any gel-based technique such as SSCP and DGGE, is that optimization is required of acrylamide and denaturant concentrations as well as electrophoresis times, each of which can differ depending on the gene region analysed, the taxa under investigation and the manufacturer brand of each system.

Dissimilar sequences with similar melting (DGGE) or folding (SSCP) properties will co-migrate (Sekiguchi *et al.* 2001) and this was recently suggested to limit our ability to discern *Symbiodinium* diversity (Apprill & Gates 2007). Co-migration affects both SSCP and DGGE and may confound estimates of diversity when it occurs between distinct *Symbiodinium* (compare Fig. 2 SSCP vs. DGGE of ITS1 and 2; for examples see Fig. S3d *Symbiodinium* C78a, C8a, and C3nt, Figs S3e and S4a). Co-migration may also occur between codominant repeats of a single *Symbiodinium* 'type' in which case it does not limit estimates of diversity (i.e. Fig. 3d, e). In reality, the limitations of co-migration to estimates of diversity can be reduced with the application with proper controls. For example, Pochon *et al.* (2007) found only 6% of diversity was missed due to co-migration in the analyses and classification of *Symbiodinium* diversity in more than a thousand foraminiferan samples. Appropriate controls included the sequencing of bands from similar-looking profiles that came from samples from distinct regions and different host taxa (Pochon *et al.* 2007). Co-migration was observed in the present study, but careful examination revealed that certain similar-looking fingerprints were actually different. Subsequent sequencing identified base differences between these bands. Routine sequencing and side-by-side re-analysis of samples with similar profiles is therefore recommended. Indeed, during the initial process of characterizing diversity from various species of host

Table 3 Overview of equipment requirements, cost, user expertise, throughput efficiency and resolution of various analysis techniques and gene regions used to detect diversity in the dinoflagellate genus *Symbiodinium*. All screening methods require general equipment for molecular analyses such as a centrifuge, thermocycler, agarose electrophoresis unit, gel visualization system and –20 °C freezer

|	RFLP	LICOR Size Analysis	SSCP	DGGE	Cloning	Direct sequencing
Equipment	*General	*General *Fluoresent primers *LICOR IR2 analyser†	*General *Polyacrylamide electrophoresis unit‡	*General *GC-primers *DGGE unit	*General *Competent cells *Incubator *DNA sequencer	*General *DNA sequencer
Cost per sample	Low	Moderate	Low–moderate	Moderate	High	Moderate-high
Technical expertise	Low	Moderate	Moderate	Moderate-high	Moderate-high	Moderate
Throughput	Moderate–high	High	Moderate–high	Moderate	Low	Moderate
Technique limitations for use in diversity and ecological studies	*Low resolution (analyses a small component of the entire sequence)	*Low resolution (based only on size variation) *High equipment cost	*Sequence co-migration *Two bands are created for every unique sequence	*Sequence co-migration *Heteroduplexes	*Cloning/sequencing error *Repeatability *Redundancy (intragenomic variants) *Expensive depending on sample size	*Sequence ambiguities (depending on homogenization of the gene region) *Expensive depending on sample size
Solutions to limitations	*Increase number of different restriction enzymes utilized.	*Analyses can be performed on SSCP apparatus with unlabelled primers	*Combine screening with band excisions and direct sequencing	*Examine and compare entire fingerprint to identify potential co-migrating sequences. *Commonly perform band excisions and direct sequencing	*Sequence a substantial number of clones per sample (40–50). *Use isoclonal cultures as controls.	*Ambiguities limited by direct sequencing of conserved rDNA regions or homogenous nuclear or organelle genes

†Size screening can also be carried out using long range SSCP gels with ultrathin spacers (Figure S2); ‡SSCP can also be carried out on some DGGE units.

from new regions, profiles that appear new, or familiar, are always characterized through sequencing (e.g. LaJeunesse 2002; LaJeunesse *et al*. 2003, 2004a,b).

The recovery of pseudogenes and artefacts from the cloning of PCR-amplified rDNA confounds phylogenetic relationships between organisms and introduces uncertainty to estimates of diversity (Buckler *et al*. 1997; LaJeunesse & Pinzón 2007; Thornhill *et al*. 2007). In this study, the bacterial cloning of ITS2 PCR amplifications recovered numerous sequence variants (Fig. 3) in samples that contained homogenous symbiont populations based on analyses from all of the other methodologies employed. The starlike phylogeny of clones sequenced from each *Symbiodinium* spp. (Fig. 3a) was similar to the analyses of intragenomic variation in iso-clonal cell lines of *Symbiodinium* (Thornhill *et al*. 2007). Also similar to their findings was that the most commonly cloned sequence regularly matched the sequence of the dominant band, or bands, in a PCR-DGGE fingerprint (Fig. 3b–e). These data support the concept that a small number of rDNA sequences are evolutionarily stable and that they should be targeted as markers for diversity and ecological studies (Dover 1982; Buckler *et al*. 1997; LaJeunesse & Pinzón 2007).

In several instances, cloning recovered sequences that were 'outside' the sequence cluster for a particular *Symbiodinium* (designated by asterisks in Fig. 3a). Whether they represented sequences from another *Symbiodinium* spp., or merely a divergent intragenomic variant, could not be determined. In situations where a colony was known to contain a mixed symbiont population, cloning identified each constituent in abundance (Fig. 3f). Unfortunately, the frequent recovery of pseudogenes, rare functional variants, and artefact sequences outweighs the potential of this technique for identifying low abundance background populations, especially when investigating the ecology of closely related eukaryotic taxa (Thornhill *et al*. 2007). Furthermore, biases in the cloning of certain sequence variants and the lack of repeatability may skew the identity and number of variants retrieved from trial to trial (Sampayo and LaJeunesse unpublished data). These drawbacks, combined with the high cost of reagents and labour, make cloning impractical as a primary method for analysing *Symbiodinium* diversity (Table 3).

Although cloning of the rDNA may bias estimates of diversity, these data provide evidence of speciation between closely related *Symbiodinium*. An unrooted phylogeny of the combined data from all samples produced eight spatially distinctive clusters (Fig. 3a) corresponding to each of eight *Symbiodinium* spp. defined above (Figs 4a–e and 5b). The phylogenetic relationship among clusters of cloned sequences (Fig. 3a) shows how the make-up of the ribosomal array is distinct between ecologically specialized organisms. For example, C35a and C79 are specific to *Stylophora pistillata* from the southern Great Barrier Reef
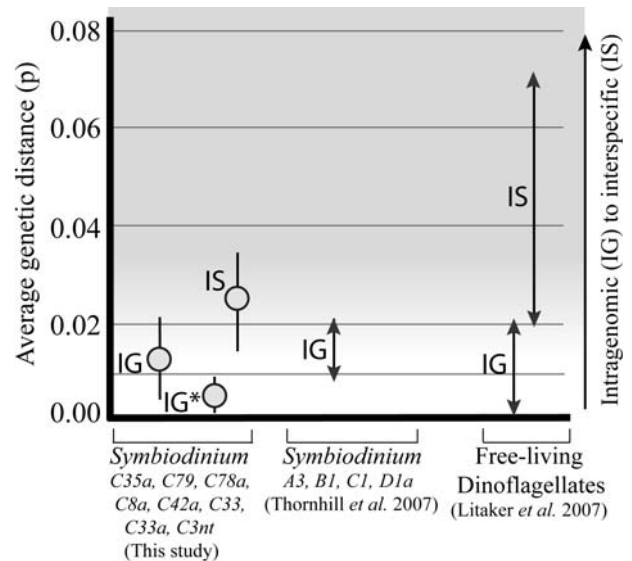
Fig. 6 Comparisons of intragenomic (IG) and interspecific (IS) ITS2 sequence variation among dinoflagellates. Sequence divergence is graphed as the average uncorrected genetic distance (*p*) and error bars represent standard deviation from the mean. IG variation among the different *Symbiodinium* 'types' investigated in this study was calculated from sequences retrieved from the cloning of PCR-amplified products. IG* was calculated using genetic distances from only the dominant sequences. For comparison, the analyses of representatives from different *Symbiodinium* clades by Thornhill *et al*. (2007) and of 81 different species of dinoflagellates spanning multiple genera by Litaker *et al*. (2007) are represented as a range rather than an average.

but exhibit different distributions influenced by the external environment. *Symbiodinium* C35a occurs in shallow colonies while C79 is found in colonies from deeper environments (Sampayo *et al*. 2007). While these symbionts are clearly ecologically and genetically distinct, they still share ancestral sequence variants. If sexual recombination between C79 and C35a were occurring, however, they would also be expected to possess some of the same derived sequence variants, but they do not (Fig. 3a). The genetic divergence between these 'types' appears to be the result of isolation through ecological specialization likely driven by irradiance (Sampayo *et al*. 2007). In time, any remaining ancestral rDNA variants would be lost through concerted evolution operating independently in each lineage. Therefore, the combined genetic and ecological data indicate that C35a and C79 are indeed different species.

Estimates of intra- vs. interspecific variation have been calculated for a wide range of dinoflagellates (Litaker *et al*. 2007). The range of variation among sequenced clones from different 'types' match with the intraspecific variation reported for many different species of dinoflagellate (Fig. 6). This range of intragenomic variation, however, is far less when only sequences identified by rDNA fingerprinting

(DGGE, SSCP) are considered. This approach conservatively focuses on only the numerically dominant and evolutionarily persistent sequence copies (LaJeunesse & Pinzón 2007). The application of these fingerprinting techniques to other dinoflagellates, and eukaryotic microbes in general, may significantly reduce estimates of intragenomic variation and improve taxonomic resolution between closely related species that may otherwise be grouped together.

Organellar chloroplast and mitochondrial genes

Chloroplast and mitochondrial genes offer an independent verification of the phylogenetic relationships among *Symbiodinium* spp. established by ribosomal DNA sequences (Santos *et al*. 2002; Takabayashi *et al*. 2004; Zhang *et al*. 2005; Pochon *et al*. 2006). The chloroplast 23S and mitochondrial cytochrome *b* sequences are conserved and did not resolve all symbionts with different host and/or depth distributions as determined by rDNA analyses (see above). When sequences from both organellar genes were combined, however, not only did they resolve seven out of the eight symbionts characterized in this study, but also their composite phylogeny was identical to reconstructions based on rDNA genes (Figs 4 and 5a; cf. Pochon *et al*. 2006). Albeit preliminary, the congruence of these phylogenetic patterns suggests that hybridization among clade C *Symbiodinium* 'types' does not occur.

Although limited information is available, it appears that dinoflagellate mitochondrial and chloroplast genomes are complex and have unusual organization (Zhang *et al*. 2002; Moore *et al*. 2003; Nash *et al*. 2007). While recombination occurs and pseudogenes have been identified in organellar genes, intragenomic variation between gene sequences is generally limited (Zhang *et al*. 2002; Santos *et al*. 2002, 2003b). Direct sequencing of these genes produces unambiguous sequences and provides a good alternative to laboratories that do not have the equipment to conduct rDNA-fingerprinting techniques using SSCP or DGGE. For purposes of rapidly screening many samples and identifying instances where multiple *Symbiodinium* spp. co-exist, however, these markers can be adapted to a gel screening analysis. Santos *et al*. (2003b) developed a potentially useful method based on size analysis of a short hypervariable area (a and b) of the cp23S rDNA domain V. Like other acrylamide-based screening techniques, it does require some expensive equipment and reagents (Table 3), but affords a high throughput capacity, is easy to use, and allows for the detection of multiple symbionts. This method was adequate in detecting diversity within clades A and B, and very sensitive in identifying mixed populations (Santos *et al*. 2003a, b), but results from this study suggest that length heterogeneity within the cp23S domain V is not common among clade C *Symbiodinium* spp. (Fig. 2,

Figs S2, S3a and S5). Many organelle genes and their control regions have yet to be examined, however, and there is a possibility that one of these could provide equal or improved resolution, if necessary, over rDNA fingerprinting.

Conclusions

With efforts to characterize microbial diversity, there are temptations to use one kind of analysis and/or establish a defined sequence divergence in the delimitation of species boundaries. Prokaryote microbial ecologists are realizing that the strict use of a percentage sequence divergence to define diversity is insufficient to describe ecologically distinct bacteria (e.g. 3% difference among 16S rDNA sequences). While broad-scale taxonomic characterizations may be appropriate to certain research questions, determining species diversity is important to understand ecological and evolutionary processes (O'Hara 1993). It is now apparent that delineating species of microbes requires finer levels of genetic differentiation integrated with ecological and physiological analyses (Konstantinidis & Tiedje 2005; Green & Bohannen 2006; Cohan & Perry 2007; Koeppel *et al*. 2008). Defining 'true' species among eukaryotic microbes is perhaps less challenging than compared to prokaryotes. Since most eukaryotes have a sexual phase in their life history, genetic analyses along with supporting ecological and physiological evidence, should allow the delimitation of species boundaries between these organisms. While growing evidence indicates that sexual recombination occurs within distinct *Symbiodinium* 'types' (Baillie *et al*. 2000b; Santos *et al*. 2003a; Coffroth & Santos 2005; D. T. Pettay and T. C. LaJeunesse, unpublished data), there is little evidence for genetic recombination between 'types' (Santos *et al*. 2004; Pettay & LaJeunesse 2007; this study). Analyses using population genetic markers such as microsatellites may further support the indication that these lineages are reproductively isolated (Santos *et al*. 2004; Pettay & LaJeunesse 2007).

For important practical purposes, the institution of an accurate and widely accepted taxonomic classification of the dinoflagellate genus *Symbiodinium* would substantially advance the field of algal–animal symbiosis research. This taxonomy should reflect the reality that members of this genus are highly diverse, both ecologically and evolutionarily. The cohesiveness of nuclear and organelle gene sequence data from samples originally characterized by ITS2-DGGE confirmed and substantially improved phylogenetic resolution among these *Symbiodinium* 'types' (Fig. 5b). Distinctive ecological patterns, such as host specificity, environmental zonation, and tolerance to thermal stress correspond to evolutionarily divergent lineages of *Symbiodinium*, revealed by combining nuclear and organelle sequences. These data indicate that species of *Symbiodinium* are presently definable based on existing molecular genetic

analyses in combination with ecological data (de Queiroz 2007; Koeppel *et al.* 2008).

Acknowledgements

References

Alverson AJ (2008) Molecular systematics and the diatom species. *Protist*, **159**, 339–353.

Appels R, Honeycutt RL (1986) rDNA evolution over a billion years. In: *DNA Systematics* (ed. Dutta SK), pp. 81–135. CRC Press, Boca Raton, Florida.

Apprill AM, Gates RD (2007) Recognizing diversity in coral symbiotic dinoflagellate communities. *Molecular Ecology*, **16**, 1127–1134.

Arnheim N (1983) Concerted evolution of multigene families. In: *Evolution of Genes and Proteins* (ed. Nei M), pp. 38–61. Sinauer & Associates, Sunderland, Massachusetts.

Avise JC (2004) *Molecular Markers, Natural History and Evolution*, 2nd edn. Sinauer & Associates, Sunderland, Massachusetts.

Baillie BK, Belda-Baillie CA, Maruyama T (2000a) Conspecificity and Indo-Pacific distribution of *Symbiodinium* genotypes (Dinophyceae) from giant clams. *Journal of Phycology*, **36**, 1153–1161.

Baillie BK, Belda-Baillie CA, Silvestre V *et al.* (2000b) Genetic variation in *Symbiodinium* isolates from giant clams based on random-amplified-polymorphic DNA (RAPD) patterns. *Marine Biology*, **136**, 829–836.

Baker AC (2003) Flexibility and specificity in coral-algal symbiosis: diversity, ecology, and biogeography of *Symbiodinium*. *Annual Review of Ecology, Evolution and Systematics*, **34**, 661–689.

Baker AC, Rowan R (1997) Diversity of symbiotic dinoflagellates (zooxanthellae) in Scleractinian corals of the Caribbean and Eastern Pacific. *Proceedings of the 8th International Coral Reef Symposium*, **2**, 1301–1306.

Barneah O, Brickner I, Hooge M *et al.* (2007) Three party symbiosis: acoelomorph worms, corals and unicellular algal symbionts in Eilat (Red Sea). *Marine Biology*, **151**, 1215–1223.

Buckler SE, Ippolito A, Holtsford TP (1997) The evolution of ribosomal DNA: divergent paralogues and phylogenetic implications. *Genetics*, **145**, 821–832.

Burnett WJ (2002) Longitudinal variation in algal symbionts (zooxanthellae) from the Indian Ocean zoanthid Palythoa caesia. *Marine Ecology Progress Series*, **234**, 105–109.

Chen CA, Yang YW, Wei NV, Tsai WS, Fang LS (2005) Symbiont diversity in scleractinian corals from tropical reefs and subtropical non-reef communities in Taiwan. *Coral Reefs*, **24**, 11–22.

Coffroth MA, Santos SR (2005) Genetic diversity of symbiotic dinoflagellates in the genus *Symbiodinium*. *Protist*, **156**, 19–34.

Cohan FM, Perry EB (2007) A systematics for discovering the fundamental units of bacterial diversity. *Current Biology*, **17**, R373–R386.

De Queiroz K (2007) Species concepts and species delimitation. *Systematic Biology*, **56**, 879–886.

DeLong EF, Pace NR (2001) Environmental diversity of bacteria and archaea. *Systematic Biology*, **50**, 470–478.

Dover GA (1982) Molecular drive: a cohesive mode of species evolution. *Nature*, **299**, 111–117.

Dover GA (1986) Molecular drive in multigene families: how biological novelties arise, spread and are assimilated. *Trends in Genetics*, **2**, 159–165.

Farris JS, Kallersjo M, Kluge AG, Bult C (1995) Testing significance of incongruence. *Cladistics*, **10**, 315–319.

Frade PR, de Jongh F, Vermeulen F, van Bleijswijk J, Bak RPM (2008) Variation in symbiont distribution between closely related coral species over large depth ranges. *Molecular Ecology*, **17**, 691–703.

Freudenthal HD (1962) *Symbiodinium* gen. Nov and *Symbiodinium microadriaticum* sp. Nov a zooxanthellae: taxonomy, life cycle and morphology. *Journal of Protozoology*, **9**, 45–52.

Goulet TL, LaJeunesse TC, Fabricius K (2008) Symbiont specificity and bleaching susceptibility among soft corals during the 1998 GBR mass coral bleaching event. *Marine Biology*, **154**, 795–804.

Green J, Bohannen BJM (2006) Spatial scaling of microbial diversity. *Trends in Ecology & Evolution*, **21**, 501–507.

Hey J (2001) The mind of the species problem. *Trends in Ecology & Evolution*, **16**, 326–329.

Hillis DM, Dixon MT (1991) Ribosomal DNA: molecular evolution and phylogenetic inference. *Quarterly Review of Biology*, **66**, 411–453.

Hunter CL, Morden CW, Smith CM (1997) The utility of ITS sequences in assessing the relationships among zooxanthellae and corals. *Proceedings of the 8th International Coral Reef Symposium*, **2**, 199–1602.

Koeppel A, Perry EB, Sikorski J *et al.* (2008) Identifying the fundamental units of bacterial diversity: a paradigm shift to incorporate ecology into bacterial systematics. *Proceedings of the National Academy of Sciences, USA*, **105**, 2504–2509.

Konstantinidis KT, Tiedje JM (2005) Genomic insights that advance species definition for prokaryotes. *Proceedings of the National Academy of Sciences, USA*, **102**, 2567–2572.

LaJeunesse TC (2001) Investigating the biodiversity, ecology, and phylogeny of endosymbiotic dinoflagellates in the genus *Symbiodinium* using the ITS region: in search of a 'species' level marker. *Journal of Phycology*, **37**, 866–880.

LaJeunesse TC (2002) Diversity and community structure of symbiotic dinoflagellates from Caribbean coral reefs. *Marine Biology*, **141**, 387–400.

LaJeunesse TC (2005) Species radiations of symbiotic dinoflagellates in the Atlantic and Indo-Pacific since the Miocene-Pliocene transition. *Molecular Biology and Evolution*, **22**, 570–581.

LaJeunesse TC, Pinzón JH (2007) Screening intragenomic rDNA for dominant variants can provide a consistent retrieval of evolutionarily persistent ITS (rDNA) sequences. *Molecular Phylogenetics and Evolution*, **45**, 417–422.

LaJeunesse TC, Trench RK (2000) Biogeography of two species of *Symbiodinium* (Freudenthal) inhabiting the intertidal sea anemone *Anthopleura elegantissima* (Brandt). *Biological Bulletin*, **199**, 126–134.

LaJeunesse TC, Loh WKW, van Woesik R *et al.* (2003) Low symbiont diversity in southern Great Barrier Reef corals, relative to those of the Caribbean. *Limnology and Oceanography*, **48**, 2046–2054.

LaJeunesse TC, Bhagooli R, Hidaka M *et al.* (2004a) Closely related *Symbiodinium* spp. differ in relative dominance in coral reef host communities across environmental, latitudinal and biogeographic gradients. *Marine Ecology Progress Series*, **284**, 147–161.

LaJeunesse TC, Thornhill DJ, Cox EF *et al.* (2004b) High diversity and host specificity observed among symbiotic dinoflagellates in reef coral communities from Hawaii. *Coral Reefs*, **23**, 596–603.

LaJeunesse TC, Reyes Bonilla H, Warner ME *et al.* (2008) Specificity and stability in high latitude eastern Pacific coral-algal symbioses. *Limnology and Oceanography*, **53**, 719–727.

Litaker RW, Vandersea MW, Kibler SR *et al.* (2007) Recognizing dinoflagellate species using ITS rDNA sequences. *Journal of Phycology*, **43**, 344–355.

Lobban CS, Schefter M, Simpson AGB *et al.* (2002) Maristenor dinoferus n. gen, n. sp., a giant heterotrich ciliate (Spirotrichea: Heterotrichida) with zooxanthellae, from coral reefs on Guam, Mariana Islands. *Marine Biology*, **141**, 411–423.

Loh WKW, Loi T, Carter D, Hoegh-Guldberg O (2001) Genetic variability of the symbiotic dinoflagellates from the wide ranging coral species *Seriatopora hystrix* and *Acropora longicyathus* in the Indo-West Pacific. *Marine Ecology Progress Series*, **222**, 97–107.

Loh WKW, Cowlishaw M, Wilson NG (2006) Diversity of *Symbiodinium* dinoflagellate symbionts from the Indo-Pacific seaslug *Pteraeolidia ianthina* (Gastropoda, Mollusca). *Marine Ecology Progress Series*, **320**, 177–184.

Long EO, Dawid IB (1980) Repeated genes in eukaryotes. *Annual Review of Biochemistry*, **49**, 727–764.

Moore RB, Ferguson KM, Loh WKW, Hoegh-Guldberg O, Carter DA (2003) Highly organized structure in the non-coding region of the psbA minicircle from clade C *Symbiodinium*. *International Journal of Systematic and Evolutionary Microbiology*, **53**, 1725–1734.

Morris CE, Bardin M, Berge O *et al.* (2002) Microbial biodiversity: approaches to experimental design and hypothesis testing in primary scientific literature from 1975 to 1999. *Microbiology and Molecular Biology Reviews*, **66**, 592–616.

Mostafavi PG, Fatemi SMR, Shahhosseiny MH, Hoegh-Guldberg O, Loh WKW (2007) Predominance of clade D *Symbiodinium* in shallow-water reef-building corals off Kish and Larak Islands (Persian Gulf, Iran). *Marine Biology*, **153**, 25–34.

Muyzer GM (1999) DGGE/TGGE: a method for identifying genes from natural ecosystems. *Current Opinion in Microbiology*, **2**, 317–322.

Nash AE, Barbrook AC, Edwards-Stuart RK, Bernhardt K, Howe CJ, Nisbet RER (2007) Organization of the mitochondrial genome in the dinoflagellate Amphidium carterae. *Molecular Biology and Evolution*, **24**, 1528–1536.

O'Hara RJ (1993) Systematic generalization, historical fate, and the species problem. *Systematic Biology*, **42**, 231–246.

Ohta T, Dover GA (1983) Population genetics of multi-gene families that are dispersed into two or more chromosomes. *Proceedings of the National Academy of Sciences, USA*, **90**, 10676–10680.

van Oppen MJH, Palstra FP, Piquet AMT, Miller DJ (2001) Patterns of coral-dinoflagellate associations in Acropora: significance of local availability and physiology of *Symbiodinium* strains and host-symbiont selectivity. *Proceedings of the Royal Society B: Biological Sciences*, **268**, 2617–2617.

van Oppen MJH, Mahiny AJ, Done TJ (2005) Geographic distribution of zooxanthella types in three coral species on the Great Barrier Reef sampled after the 2002 bleaching event. *Coral Reefs*, **24**, 482–487.

Pace N (1997) A molecular view of microbial diversity and the biosphere. *Science*, **276**, 734–740.

Pettay DT, LaJeunesse TC (2007) Microsatellites from clade B *Symbiodinium* spp. specialized for Caribbean corals in the genus Madracis. *Molecular Ecology Notes*, **7**, 1271–1274.

Pochon X, Pawlowski J, Zaninetti L, Rowan R (2001) High genetic diversity and relative specificity among *Symbiodinium*-like endosymbiotic dinoflagellates in soritid foraminiferans. *Marine Biology*, **139**, 1069–1078.

Pochon X, Montoya-Burgos JI, Stadelmann B, Pawlowski J (2006) Molecular phylogeny, evolutionary rates, and divergence timing of the symbiotic dinoflagellate genus *Symbiodinium*. *Molecular Phylogenetics and Evolution*, **38**, 20–30.

Pochon X, Garcia-Cuestos L, Baker AC, Castella E, Pawlowski J (2007) One-year survey of a single Micronesian reef reveals extraordinarily rich diversity of *Symbiodinium* types in sorited foraminifera. *Coral Reefs*, **26**, 867–882.

Ramírez MJ (2006) Further problems with the incongruence length difference test: 'hypercongruence' effect and multiple comparisons. *Cladistics*, **22**, 289–295.

Rappé MS, Giovannoni SJ (2003) The uncultured microbial majority. *Annual Review of Microbiology*, **57**, 369–394.

Rodriguez-Lanetty M, Loh W, Carter D, Hoegh-Guldberg O (2001) Latitudinal variability in symbiont specificity within the widespread scleractinian coral *Plesiastrea versipora*. *Marine Biology*, **138**, 1175–1181.

Rowan R, Knowlton N (1995) Intraspecific diversity and ecological zonation in coral algal symbiosis. *Proceedings of the National Academy of Sciences, USA*, **92**, 2850–2853.

Rowan R, Powers DA (1991) Molecular genetic identification of symbiotic dinoflagellates (zooxanthellae). *Marine Ecology Progress Series*, **71**, 65–73.

Rudman WB (1982) The taxonomy and biology of further aeolidacean and arminacean nudibranch molluscs with symbiotic zooxanthellae. *Zoological Journal of the Linnaean Society*, **74**, 147–196.

Sampayo EM, Franceschinis L, Hoegh-Guldberg O, Dove S (2007) Niche partitioning of symbiotic dinoflagellates. *Molecular Ecology*, **16**, 3721–3733.

Sampayo EM, Ridgway T, Bongaerts P, Hoegh-Gulberg O (2008) Bleaching susceptibility and mortality of corals are determined by fine-scale differences in symbiont type. *Proceedings of the National Academy of Sciences, USA*, **105**, 10444–10449.

Santos SR, Taylor DJ, Kinzie RA *et al.* (2002) Molecular phylogeny of symbiotic dinoflagellates inferred from partial chloroplast large subunit 23S-rDNA sequences. *Molecular Phylogenetics and Evolution*, **23**, 97–111.

Santos SR, Gutierrez-Rodriguez C, Lasker HR, Coffroth MA (2003a) *Symbiodinium* sp. associations in the gorgonian *Pseudopterogorgia elisabethae* in the Bahamas: high levels of genetic variability and population structure in symbiotic dinoflagellates. *Marine Biology*, **143**, 111–120.

Santos SR, Gutierrez-Rodriguez C, Coffroth MA (2003b) Phylogenetic identification of symbiotic dinoflagellates via length heteroplasmy in domain V of chloroplast large subunit (cp23s) ribosomal DNA sequences. *Marine Biotechnology*, **5**, 130–140.

Santos SR, Shearer TL, Hannes AR, Coffroth MA (2004) Fine scale diversity and specificity in the most prevalent lineage of symbiotic dinoflagellates (*Symbiodinium*, Dinophyta) of the Caribbean. *Molecular Ecology*, **13**, 459–469.

Sarà M, Liaci L (1964) Symbiotic associations between zooxanthellae and two marine sponges of the genus Cliona. *Nature*, **203**, 321.

Schaal BA, O'Kane SL, Rogstad SH (1991) DNA variation in plant populations. *Trends in Ecology & Evolution*, **6**, 329–332.

Schoenberg DA, Trench RK (1980a) Genetic variation in *Symbiodinium* (= Gymniodinium) microadriaticum Freudenthal, and its specificity in its symbiosis with marine invertebrates. I. Isoenzyme and soluble protein patterns of axenic. Cultures of *Symbiodinium micoradriaticum. Proceedings of the Royal Society B: Biological Sciences*, **207**, 405–427.

Schoenberg DA, Trench RK (1980b) Genetic variation in *Symbiodinium* (= Gymniodinium) microadriaticum Freudenthal, and its specificity in its symbiosis with marine invertebrates. II Morphological variation in *Symbiodinium microadriaticum. Proceedings of the Royal Society B: Biological Sciences*, **207**, 429–444.

Schoenberg DA, Trench RK (1980c) Genetic variation in *Symbiodinium* (= Gymniodinium) microadriaticum Freudenthal, and its specificity in its symbiosis with marine invertebrates. III. Specificity and infectivity of *Symbiodinium Microadriaticum. Proceedings of the Royal Society B: Biological Sciences*, **207**, 445–460.

Schönberg CHL, Loh WKW (2005) Molecular identity of the unique symbiotic dinoflagellates found in the bioeroding demosponge Cliona orientalis. *Marine Ecology Progress Series*, **299**, 157–166.

Sekiguchi H, Tomioka N, Nakahara T, Uchiyama H (2001) A single band does not always represent single bacterial strains in denaturing gradient gel electrophoresis analysis. *Biotechnology Letters*, **23**, 1205–1208.

Seutin G, White BN, Boag PT (1991) Preservation of avian blood and tissue samples for DNA analyses. *Canadian Journal of Zoology*, **69**, 82–90.

Speksnijder A, Kowalchuk GA, De Jong S, Kline E, Stephen JR, Laanbroek HJ (2001) Microvariation artifacts introduced by PCR and cloning of closely related 16S rRNA gene sequences. *Applied and Environmental Microbiology*, **67**, 469–472.

Stat M, Hoegh-Guldberg O, Fitt WK, Carter D (2008) Host symbiont acquisition strategy drives *Symbiodinium* diversity in the southern Great Barrier Reef. *Coral Reefs*, **27**, 763–772.

Sunnucks P, Wilson AC, Beheregaray LB *et al*. (2000) SSCP is not so difficult: the application and utility of single stranded conformation polymorphism in evolutionary biology and molecular ecology. *Molecular Ecology*, **9**, 1699–1710.

Suzuki MT, Giovannoni SJ (1996) Bias caused by template annealing in the amplification of mixtures of 16S rRNA genes by PCR. *Applied and Environmental Microbiology*, **62**, 625–630.

Swofford D (2000) PAUP*, *Phylogenetic analysis using parsimony (*and other methods)*, Version 4.0b10. Sinauer & Associates, Sunderland, Massachusetts.

Takabayashi M, Santos SR, Cook CB (2004) Mitochondrial DNA phylogeny of the symbiotic dinoflagellates (*Symbiodinium*, Dynophita). *Journal of Phycology*, **40**, 160–164.

Taylor DL (1974) Symbiotic marine algae: taxonomy and biological fitness. In: *Symbiosis in the Sea* (ed. Vernberg WB), pp. 245–262. Columbia University Press, New York.

Tchernov D, Gorbunov MY, de Vargas C *et al*. (2004) Membrane lipids of symbiotic algae are diagnostic of sensitivity to thermal bleaching in corals. *Proceedings of the National Academy of Sciences, USA*, **101**, 13531–13535.

Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG (1997) The Clustal_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Research*, **24**, 4876–4882.

Thornhill D, LaJeunesse TC, Kemp DW, Fitt WKW, Schmidt GW (2006) Multi-year seasonal genotypic surveys of coral-algal symbiosis reveal prevalent stability or post-bleaching reversion. *Marine Biology*, **148**, 711–722.

Thornhill DJ, LaJeunesse TC, Santos SR (2007) Measuring rDNA diversity in eukaryotic microbial systems: how intragenomic variation, pseudogenes, and PCR artefacts confound biodiversity estimates. *Molecular Ecology*, **16**, 5326–5340.

Thornhill DJ, Kemp DW, Bruns BU, Fitt WK, Schmidt GW (2008) Correspondence between cold tolerance and temperate biogeography in a western Atlantic *Symbiodinium* (Dinophyta) lineage. *Journal of Phycology*, **44**, 1126–1135.

Toller WW, Rowan R, Knowlton N (2001) Zooxanthellae of the *Montastraea annularis* species complex: patterns of distribution of four taxa of *Symbiodinium* on different reefs and across depths. *Biological Bulletin*, **201**, 348–359.

Trench RK (1993) Microalgal-invertebrate symbioses: a review. *Endocytobiosis and Cell Research*, **9**, 135–175.

Trench RK, Blank RJ (1987) *Symbiodinium* microadriaticum Freudenthal, S. goreauii sp. nov., S. kawagutii sp. nov. & S. pilosum sp. nov. gymnodinioid dinoflagellate symbionts of marine invertebrates. *Journal of Phycology*, **23**, 469–481.

Warner ME, LaJeunesse TC, Robison JE, Thur RM (2006) The ecological distribution and comparative photobiology of symbiotic dinoflagellates from reef corals in Belize: potential implications for coral bleaching. *Limnology and Oceanography*, **51**, 1887–1897.

Zardoya R, Costas E, Lopez-Rodas V, Garrido-Pertierra A, Bautista JM (1995) Revised dinoflagellate phylogeny inferred from molecular analysis of large subunit ribosomal RNA gene sequences. *Journal of Molecular Evolution*, **41**, 637–645.

Zhang H, Bhattacharya D, Lin S (2005) Phylogeny of dinoflagellates based on mitochondrial cytochrome *b* and nuclear small subunit rDNA sequence comparisons. *Journal of Phycology*, **41**, 411–420.

Zhang Z, Cavalier-Smith T, Green BR (2002) Evolution of dinoflagellate unigenic minicircles and the partially concerted divergence of their putative replicon origens. *Molecular Biology and Evolution*, **19**, 489–500.

Eugenia Sampayo is interested in the evolutionary ecology of *Symbiodinium* and the extent to which these symbiotic dinoflagellates determine host physiology and range margins. Sophie Dove is interested in the photobiology and thermal tolerance of corals. Her research is directed to understand the molecular and biochemical processes underlying coral productivity and growth under different environmental conditions. Todd LaJeunesse's research focuses on ecological and co-evolutionary processes important in animal-microbe associations as well as the response of coral-algal partnerships to major changes in climate.

Supporting information

Additional supporting information may be found in the online version of this article:

Fig. S1 SSU and LSU rDNA RFLPs of *Symbiodinium* collected from different host species, living at various depths and locations. For gels (a) through (e), those lanes labelled C35a, C78a, C8a, C79 refer to symbionts found in colonies of *Stylophora pistillata*; C42a,

C33a, C33 from colonies of *Pocillopora damicornis*. (h) RFLP analyses of symbiont C3nt from colonies of *Seriatopora hystrix*. (a, e, h) Restriction digests were performed with the enzyme *Taq*I on amplifications of the entire SSU (18S, ~1600 bp) and four distinct profiles were resolved. *Symbiodinium* 'types' C78a, C8a, C33a, C33 and C3nt had the same SSU-Taq I RFLP profile and were not resolved. Gels (c) through (g) show restriction digests of amplifications of the D1/D2 domains of the LSU (28S, ~ 650 bp) using three restriction enzymes, *Taq*I (c, d, h), *Xho*I (e, f, h) and *Hha*I (g). Each of these enzymes produced three distinct profiles but differed in which symbionts were resolved. For example, LSU-*Taq*I analysis resolved C35a and C42a as distinct, but did not distinguish the remaining 'types' from each other (c, d, h). LSU-*Xho*I resolved only C78a and C42a as being unique (e, f, h). Finally, *Hha*I resolved C8a and C42a but the remaining 'types' possessed the same profile (g). The combined analyses of all three enzymes applied to the LSU identified five genetically distinctive symbionts while analyses of the SSU identified only four (see Fig. 2).

Fig. S2 Size comparisons of the chloroplast 23S rDNA hypervariable areas a and b of domain V (Santos *et al.* 2002). Amplified fragments from samples representative of eight different *Symbiodinium* spp. (C35a, C78a, C8a, C79, C42a, C33, C33a, C3nt) were analysed on ultrathin long-range polyacrylamide gels using the (a) gelscan 2000 (GS2000) or (b) LICOR NEN analyser. Fragment size was identical across all samples, 198 bases.

Fig. S3 Analyses of PCR-amplified chloroplast (cp23S) DNA and nuclear ribosomal DNA (LSU, ITS1 and ITS2) using SSCP to distinguish among *Symbiodinium* 'types' C35a, C78a, C8a, C79, C42a, C33, C33a, C3nt. (a) SSCP of the complete chloroplast 23S rDNA domain V (~700 bp) run on a Hoefer mini-vertical gel electrophoresis system. (b) The D1/D2 domain of the large subunit of the rDNA (LSU or 28S) was separated using the Hoefer SSCP system as well as the (c) gelscan 2000 (GS2000). Finally, the internal transcribed spacer regions 1 (d) and 2 (e) regions of the rDNA were separated using the GS2000. SSCP analysis of D1/D2 domains of the LSU resolved eight profiles using both SSCP devices (GS2000 and Hoefer system). In comparison, the Hoefer System provided better resolution. Gel deformation, or 'smiling,' occurred using the GS2000 because longer running times were required to resolve the LSU fragments. Differences in equipment design may therefore affect the quality of the data generated.

Fig. S4 DGGE analyses of the internal transcribed spacer region 1 (a, b) and internal transcribed region 2 from nuclear rDNA (c). Fingerprint profiles are shown for eight symbiont 'types', C35a, C78a, C8a, C79, C42a, C33, C33a, and C3nt isolated from the three coral species, *Stylophora pistillata*, *Pocillopora damicornis*, *Seriatopora hystrix*. (a) ITS1-DGGE did not clearly distinguish between C35a and C79 or C42a and C3nt. However, fixed differences in faint background bands suggests that differences exist between the profiles. For example, further analyses of the C42a ITS1 profile identified two co-migrating bands (b). One sequence matched the ancestral C3/C1 sequence (which defines the C3nt ITS1 profile) but the sequence of the other band was two base changes different (see phylogeny in Fig. 3b) and explains why the C42a profile contained two heteroduplexes (HD). Black arrows indicate the diagnostic bands that were excised and sequenced.

Please note: Wiley-Blackwell are not responsible for the content or functionality of any supporting information supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.