Acta Protozool. (2014) 53: 63–75
http://www.eko.uj.edu.pl/ap
doi:10.4467/16890027AP.13.0021.1118

ACTA
PROTOZOOLOGICA

Special issue:
Marine Heterotrophic Protists
Guest editors: John R. Dolan and David J. S. Montagnes

Review paper

Problematic Biases in the Availability of Molecular Markers in Protists: The Example of the Dinoflagellates

Fernando GÓMEZ

Laboratory of Plankton Systems, Oceanographic Institute, University of São Paulo, São Paulo, Brazil

Abstract. Dinoflagellates (Alveolata, Dinophyceae) are protists with a truly remarkable diversity in lifestyles (free-living, parasites and mutualistic symbionts), habitats (marine, freshwater, plankton, benthos), and trophic modes (heterotrophic, plastid-containing). Here dinoflagellates are used to evaluate biases in the availability of molecular markers in relation to the variety of functional and ecological characteristics of protists. A large number of dinoflagellate sequences are available in GenBank, at least one for 56% of the 264 described genera. The most common marker is the small ribosomal subunit ribosomal DNA (49%). At the species level, SSU rDNA or the large subunit rDNA are available for 15% of the 2,386 described species. Availability of sequences of the internal transcribed spacers (ITS) and cytochrome oxidase I (COI) show a strong bias towards cultivable species. Relative to trophic mode, while about half of the known dinoflagellates are heterotrophic, only 12% of them have been sequenced compared to 29% of the plastid-containing species. For the COI marker availability is 10 times greater for plastid-containing compared to heterotrophic species. Freshwater species are underrepresented (13%) relative to the marine forms (22%). A high proportion of benthic species have been sequenced (46%) reflecting interest in *Symbiodinium* and harmful epiphytic taxa. Most of the relatively few described mutualistic species have been sequenced (> 80%). In contrast, only 17% of the described parasitic species have been sequenced, and most of the available sequences were not identified at the species level. In recent years, new species have been described mostly from coastal blooms or cultures. These studies are favored by the availability of abundant material for detailed studies of ultrastructure and multi-gene molecular phylogenies. Many methods are difficult to apply for the scarce specimens available from the open ocean. The requirement of these protocols, easy to apply with cultured species, is an obstacle in our knowledge of the open ocean diversity because it discourages studies based on sparse material. Consequently, in recent years descriptions of new species from the open ocean have declined considerably.

Key words: Alveolate, Dinophyceae, Dinophyta, DNA barcoding, microbial diversity, molecular phylogeny, unicellular eukaryotes.

Abbreviations: COI – cytochrome oxidase I; DNA – deoxyribonucleic acid; ITS – internal transcribed spacers; LSU rDNA – large subunit ribosomal DNA; PCR – polymerase chain reaction; SSU rDNA – small subunit ribosomal DNA.

Address for correspondence: Fernando Gómez, Laboratory of Plankton Systems, Oceanographic Institute, University of São Paulo, Praça do Oceanográfico 191, sala 100, São Paulo, SP 05508-120, Brazil; E-mail: fernando.gomez@fitoplancton.com

INTRODUCTION

The use of molecular methods has played a major role in advancing our understanding of microbial diversity. The technique of DNA taxonomy or DNA barcoding, a short standardized stretch of DNA sequence, may be used to identify species (Blaxter 2004, Miller 2007). The most extended molecular markers in DNA taxonomy are the mitochondrial markers [cytochrome oxidase I (COI), cytochrome oxidase B, etc.], and the nuclear markers of the ribosomal DNA operon. COI is the most extended marker for the DNA barcoding in animals (Hebert *et al*. 2003), and macroalgae (Le Gall and Saunders 2010), while it is considered too conserved for higher plants (Newmaster *et al*. 2006). The utility of COI for DNA barcoding is controversial in protists, with less documented attempts and variable success (Ehara *et al*. 2000, Evans *et al*. 2007). It has been pointed out that a barcoding system based on mitochondrial markers such as COI or cytochrome oxidase B will not provide a universal solution to protist identifications because anaerobic species lack mitochondria (Henze and Martin 2003).

The other main group of markers for phylogenetical analysis is the ribosomal DNA operon. These genes are present throughout the living world and they evolve relatively slowly, both of these traits enable comparison of distantly related organisms. The repetitive arrangement within the genome provides enough amounts of template DNA for PCR, even in smallest organisms (Hillis and Dixon 1991). These markers are the 18S rDNA or small subunit ribosomal DNA (SSU rDNA), the 28S rDNA or large subunit rDNA (LSU rDNA), the internal transcribed spacers 1 (ITS1) and 2 (ITS2) regions, and the 5.8S rDNA. There are other markers used in molecular phylogenies, mainly based on protein coding genes: α- and β-tubulin, actin, cytochrome oxidase B, heat-shock protein 90, and others (Pochon *et al*. 2012). These alternative markers are not well represented in public databases. The vast majority of the sequences are from photosynthetic strains available in culture collections and the phylogenies do not include enough taxa to be widely useful.

The dinoflagellates are an ancient alveolate group of about ~ 2,400 extant described species and other ~ 2,000 fossil species (Taylor 1987, Gómez 2012a). Dinoflagellates possess numerous unique morphological and ultrastructural attributes (Hackett *et al*. 2004) such as huge genome sizes of 0.2–200 pg DNA per cell (LaJeunesse *et al*. 2005, Lin 2006) and many genes with high copy numbers (Rowan *et al*. 1996). This feature has favored single-cell sequencing (Lynn and Pinheiro 2009). The single-cell PCR technique allows sequencing of rare uncultured species, including those found only in low abundances, from different habitats and ecological niches.

While the size of the dinoflagellate genome has, to date, prevented whole genome sequencing projects, several genes have been sequenced. The first sequence of a dinoflagellate gene was the 5S rDNA of *Crypthecodinium cohnii* (Hinnebusch *et al*. 1981). The most extensive markers are parts of the rDNA array, the SSU rDNA (McNally *et al.* 1994, Saunders *et al*. 1997), LSU rDNA (Lenaers *et al*. 1991, Daugbjerg *et al*. 2000), and the ITS regions (LaJeunesse 2001). The differences in the evolutionary rates in the ribosomal genes of dinoflagellates, especially SSU rDNA and also in the domains of the LSU rDNA, has yielded phylogenetic trees with a characteristic structure: a large group of short branched sequences, the so-called Gymnodiniales, Prorocentrales and Peridiniales complex (GPP complex), and a group with longer branches including Gonyaulacales and other taxa (Saunders *et al*. 1997, Saldarriaga *et al*. 2004).

The phylogenies based on SSU and LSU rDNA markers do not resolve the relations between the main classical orders due to extremely low divergence rates. As a general rule, the SSU rDNA marker tends to be suitable to infer phylogenies at family or genus level while the LSU rDNA marker is more effective in discriminating species. The ITS marker is present in multiple distinct copies, with the possibility that high intra- and intergenomic variation and the presence of indels that can make direct sequencing challenging and alignment difficult. It is really difficult to design a dinoflagellate-specific primer set to amplify the ITS region for all or most of the dinoflagellate taxa (Litaker *et al*. 2007, Pochon *et al*. 2012, Stern *et al*. 2012). The COI marker can be useful for identifying single cells or monospecific cultures and successfully used to distinguish different genotypes of the coral symbiont *Symbiodinium* (Takabayashi *et al*. 2004). However, its utility as a dinoflagellate barcode is questionable due to repetitive failure of the primer sets to amplify all dinoflagellate species in natural assemblages, and the lower resolving power with generally lower bootstrap support than other genes examined (Zhang *et al*. 2007, Lin *et al*. 2009). Plastid genes have not been used to infer the phylogeny of dinoflagellates

for several reasons. For example, they are only present in about 50% of taxa, and in those species, they may have been lost and gained on multiple occasions (Saldarriaga *et al*. 2001, Hackett *et al*. 2004).

Dinoflagellates exhibit extreme diversity in almost all ecological characteristics, including their modes of nutrition (heterotrophic, plastid-containing), habitat distribution (marine, freshwater, plankton, benthos), and lifestyle (free-living species, parasites and mutualistic symbionts) (Taylor 1987, Gómez 2012b). For these reasons, dinoflagellates are an ideal case to evaluate biases in the availability of the molecular markers in relation to functional and ecological characteristics of protist taxa. It appears that the advances in the molecular phylogeny show a high taxonomic selectivity, and may even have adverse consequences for the knowledge of certain groups.

MATERIALS AND METHODS

The primary source used was the checklist of living dinoflagellates by Gómez (2012a) that listed 2,377 described species. Twelve other species have been added (recently described or missing in the previous checklist). Each species was classified with regard to distribution and habitat (marine or freshwater, and planktonic or benthic), lifestyle (free-living, parasitic or mutualistic symbionts) as well as trophic modes (heterotrophic or plastid-containing) following the criteria reported in Gómez (2012b). The taxa identified at the species level (binomials) were surveyed in international nucleotide sequence databases (DDBJ/EMBL/GenBank), and labeled according to the availability of at least one nucleotide sequence of the most extensive molecular markers (SSU, LSU, ITS and COI). The database is provided as supplementary material (supplementary material Table S1). The alternative molecular markers (actin, α- and β-tubulin, cytochrome oxidase B, heat-shock protein 90, etc.) were largely underrepresented, and the vast majority of the sequences were from photosynthetic strains available in culture collections. Almost all the species names reported in GenBank were represented by at least one nucleotide sequence of the SSU, LSU, ITS or COI markers, while the only binomials lacking any of these markers were *Symbiodinium fittii* (only microsatellite sequence available), *Prorocentrum nanum* (only cytochrome oxidase B), and *Pyrocystis fusiformis* (only luciferase).

It is often difficult to determinate whether the binomials reported in GenBank were correctly identified. There is no formal documentation of taxonomic identifications (e.g. photographs, collection sources for cultures, or information on the individual who performed identifications). Some sequences of other groups are erroneously reported under dinoflagellate names [*Exuviaella pusilla* (#DQ388459) or *Prorocentrum minimum* (#EF017804)]. It can be assumed that the cultures were contaminated, and the authors did not first check the BLAST (Basic Local Alignment Search Tool) entries of their sequences. For example, a sequence of *Neoceratium furca* (#AY027908) corresponds with an *Alexandrium* species. In addition to contaminations, culture collections are sometimes subjected to mislabeling or misidentifications. For example sequences of the armoured genus *Heterocapsa* were named under the unarmoured genus *Gymnodinium* [#*Gymnodinium* sp. CCMP424 (#EF492492) or *Gymnodinium* sp. UTEX1653 (#EF492494)] (see additional examples in Stern *et al.* 2012). According to GenBank the sequence of *Peridinium centenniale* was restricted to the cytochrome oxidase B gene. The SSU rDNA sequences of this species may be under other names [Dinophyceae sp. CCAC0002 (#EF058236), *Glenodinium inaequale* (#EF058237)]. There were sequences from cultures identified as 'Scrippsiella inulfa' that may correspond to *Calcigonellum infula*. The misspelling of the epithet did not facilitate the identification of the correct species name. In other cases, there is no documentation to resolve doubts. The epithets of 'Prorocentrum donggang' or 'P. tainan' resemble valid binomials. However, these species have not been formally described and the apparent epithet may refer to places where the cells were isolated in Taiwan. 'Gymnodinium falcatum' (#AY320049) has been never formally described. It may refer to *Gyrodinium faltacum*. The species is named *Pseliodinium vaubanii*, one of its synonyms, because the morphology and the sequence did not branch with *Gymnodinium* or *Gyrodinium sensu stricto*. These are examples of the difficulties to assign a proper species name to the sequences.

RESULTS

Overview of the availability of the main molecular markers

To date, for the 264 described extant genera, at least one nucleotide sequence is available in GenBank for 149 (56%) of dinoflagellate genera (Table 1). The most common marker is SSU rDNA for 131 genera (49%), LSU rDNA for 108 genera (41%), ITS for 69 genera (26%), and COI for 48 genera (18%). At least one nucleotide sequence identified at the species level is available for 493 dinoflagellates (Table 1). This corresponded to 20% of the 2,386 extant described species. The SSU rDNA marker is available for 345 species (14% of 2,386), LSU for 358 species (15%), ITS for 184 species (7%), and COI for 97 species (4%).

Among the most speciose dinoflagellate genera (> 11 species per genus), the highest percentage of sequenced species concerns *Gambierdiscus* (100%), *Symbiodinium* (87%, only lacking *S. tridacnorum*), and *Alexandrium* (77%), followed by other marine plastid-containing genera such as *Karenia*, *Karlodinium*, *Blastodinium*, *Heterocapsa*, *Prorocentrum* and *Scrippsiella,* with sequences of more than 40% of the species (Fig. 1). All these taxa are available in culture, with the exception of the parasite *Blastodinium*. The genus

Table 1. Number and percentage of dinoflagellate genera and species with at least one nucleotide sequence available in DDBJ/EMBL/GenBank in January 2013. SSU – small subunit rDNA, LSU – large subunit rDNA, ITS – internal transcribed spacers, COI – cytochrome oxidase I, CHL – plastid-containing, HET – heterotrophic, FRE – free-living, PAR – parasite, SYM – mutualistic symbiont, MAR – marine, FS – freshwater/ continental, PLK – plankton, BEN – benthos.

	Total	Any marker	SSU	LSU	ITS	COI
Genera	264	149 (56%)	131 (49%)	108 (41%)	69 (26%)	48 (18%)
Species	2,386	493 (20%)	345 (14%)	358 (15%)	184 (7%)	97 (4%)
CHL	1,204	348 (29%)	237 (20%)	269 (22%)	161 (13%)	89 (7%)
HET	1,182	145 (12%)	108 (10%)	89 (7%)	23 (2%)	8 (0.7%)
FRE	2,200	447 (20%)	307 (14%)	331 (15%)	156 (7%)	91 (4%)
PAR	165	29 (17%)	29 (17%)	18 (11%)	18 (11%)	1 (0.6%)
SYM	21	17 (81%)	9 (43%)	9 (43%)	10 (47%)	5 (24%)
MAR	1,964	438 (22%)	305 (15%)	310 (16%)	163 (8%)	93 (5%)
FW	422	55 (13%)	40 (9%)	48 (11%)	21 (5%)	4 (0.9%)
PLK	2,180	397 (18%)	277 (13%)	301 (14%)	154 (7%)	81 (3%)
BEN	206	96 (46%)	68 (33%)	57 (27%)	30 (14%)	16 (7%)

with the highest number of sequenced species, 28, is *Neoceratium* (Fig. 1). The percentage of sequenced species ranged from 5–8% among the most speciose genera (*Protoperidinium*, 268 species; *Gymnodinium*, 216 species; *Gyrodinium*, 153 species). There are no sequences labeled with a proper species name for the speciose genera *Warnowia*, *Oxytoxum*, *Lissodinium*, *Centrodinium* or *Corythodinium* (Fig. 1). The genus *Dinophysis* comprises both plastid-containing and heterotrophic species. *Dinophysis sensu lato* is the genus with the higher number of sequenced heterotrophic species. Sequences of eight heterotrophic species of the *Dinophysis hastata*-group and other six heterotrophic species of other clades (*D. apicata*, *D. argus*, *D. braarudii*, *D. brevisulcus*, *D. expulsa*, *D. similis*) are available in GenBank (supplementary material Table S1). The exclusively heterotrophic genus with the higher number of sequenced species is *Ornithocercus* (6 of the 15 described species, 40%). No sequences are available for *Glenodinium*, the most speciose freshwater genus (the sequence labeled *Glenodinium inaequale* corresponded to other genus). The speciose freshwater genus with the higher percentage of sequenced species is *Peridinium* (36%), while only one species is sequenced for the genera *Cystodinium* and *Peridiniopsis* (Fig. 1).

There are only 59 species of the 2,386 described species (2%) with all of the four main markers (SSU, LSU, ITS, COI) available in GenBank. These species belong to the genera responsible for harmful algal blooms such as *Prorocentrum* (11 species), *Alexandrium* (9 species), and *Dinophysis* (5 species). There is not any representative of a parasitic form. The 'well-sequenced' group is largely dominated by plastid-containing species, with only seven heterotrophic species: *Crypthecodinium*, *Noctiluca*, *Oxyrrhis*, three pfiesterid species (*Cryptoperidiniopsis*, *Pfiesteria*, *Pseudopfiesteria*), and *Phalacroma rotundatum*. All the species with the four main markers are planktonic, with the exception of seven species of symbiotic *Symbiodinium*, and epiphytic species of *Prorocentrum* and *Coolia monotis*. The only freshwater species with the four main markers are *Parvodinium inconspicuum* and *Peridinium willei*.

The SSU rDNA marker is the only available for 111 species, with 53 heterotrophic species (48% heterotrophic species). The LSU rDNA marker alone is available for 95 species, with 36 heterotrophic species (38%). The ITS marker is the only sequence for 14 species, mainly the genera *Scrippsiella* (5 species) and *Heterocapsa* (3 species). These are plastid-containing species, with the exception of ITS sequence of *Protoperidinium tricingulatum* (supplementary material Table S1). The COI marker is the only molecular marker for three species, the plastid-containing *Neoceratium macroceros*, *Gonyaulax hyalina* and *Prorocentrum pusillum* (the SSU rDNA sequence of the latter did not correspond to a dinoflagellate).

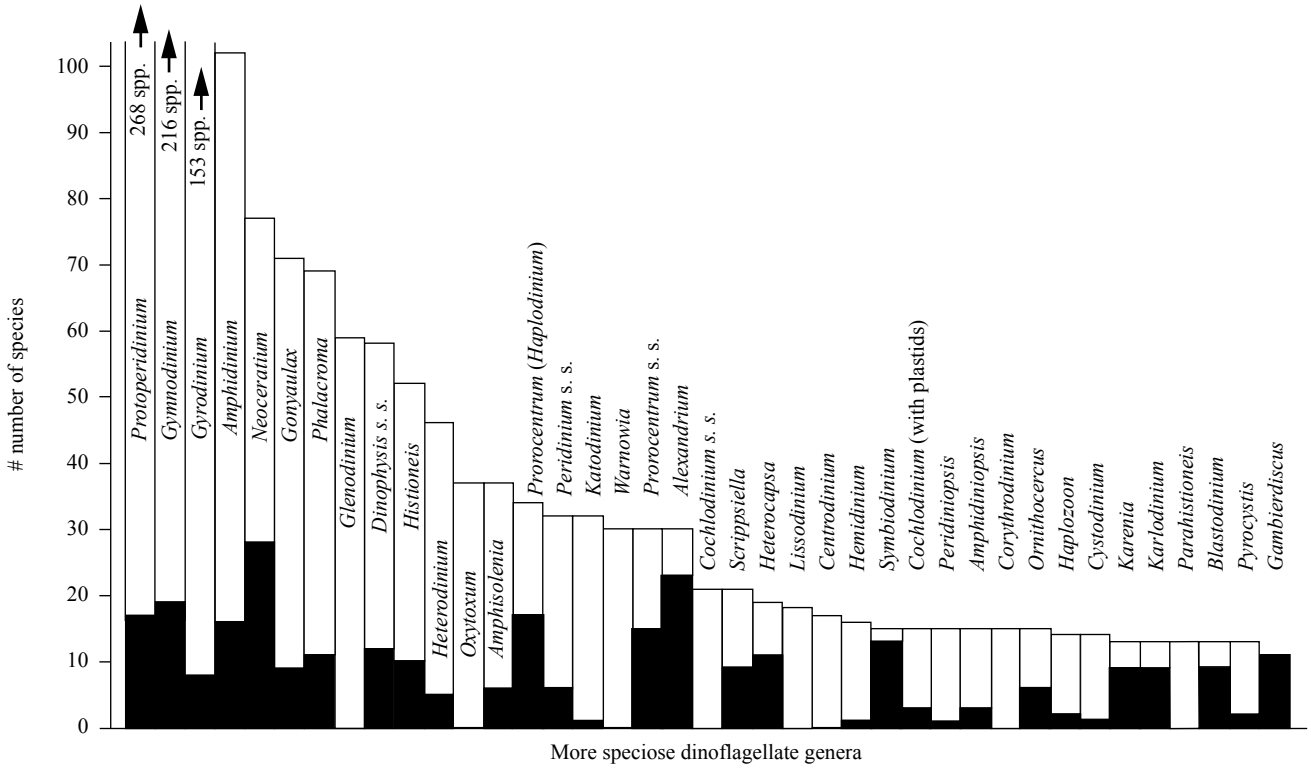

Fig. 1. Number of species of the most speciose dinoflagellate genera (> 11 species per genus). The empty bars represented the number of described species based on Gómez (2012a). The black bars represent the number of species with, at least, one nucleotide sequence available in DDBJ/EMBL/GenBank in January 2013.

Sequences by habitat

From marine waters a total of 1,964 dinoflagellate species (82% of the total 2,386 species) have been described compared to 422 species (18%) from freshwater environments. Of the described 1,964 marine species, there is at least one sequence for 438 species or 22% (Table 1). SSU rDNA sequences are available for 305 species (15% of the total marine species), LSU for 310 species (16%), ITS for 163 species (8%), and COI for 93 species (4%).

A total of 55 species of the 422 freshwater species (13%) have been sequenced for at least one molecular marker. The SSU marker is available for 40 species (9%), LSU for 48 species (11%), and ITS for 21 species (5%). COI sequences are only available for four freshwater species (< 1%, *Ceratium hirundinella*, *Parvodinium inconspicuum*, *Peridinium willei* and *Piscinoodinium limneticum*). The percentage of heterotrophic species described from freshwaters is low (11%). Only the sequences of two heterotrophic species, *Gyrodinium helveticum* and *Tyrannodinium edax*, are available in GenBank (supplementary material Table S1).

Most of the described dinoflagellate species are planktonic, 91% of the total 2,386 species, with only 204 (9%) of the species benthic. With regard to sequences, there is at least one nucleotide sequence for 399 plankton species (18%, Table 1). An SSU rDNA sequence is available for 279 species (13%), LSU for 301 species (14%), ITS for 154 species (7%), and COI for 81 species (3%). For the 204 benthic species, there are sequences for 94 species (46%). This relative overrepresentation is due to the numerous studies on *Symbiodinium*, and the harmful epiphytic species of *Prorocentrum* (20 species), *Gambierdiscus* (11 species), *Ostreopsis* (4 species) and *Coolia* (3 species). The SSU rDNA was the most common marker (66 species, 32%), followed by the LSU rDNA (57 species, 28%), ITS (30 species, 14%), and COI (16 species, 8%).

Sequences by trophic mode

Among described dinoflagellate species 49% (1,182 species) are purely heterotrophic, devoid of plastids, while 51% of the dinoflagellate species have been reported to contain plastids (that does not strictly imply autotrophy). In GenBank there are 493 species with at least one sequence (Table 1) and of these species only 145 species (29%) were heterotrophic. This bias towards plastid-containing species varies according to the markers. The SSU rDNA marker is available for 345 species, 107 of which (31%) are heterotrophic. The LSU rDNA sequence is available for 358 species, with 89 heterotrophic species (24%). The ITS sequence is available for 184 species, including 23 heterotrophic species (12%). The COI sequence, available for 97 species includes only 8 heterotrophic species (Table 1).

Sequences by lifestyle

Overall the species catalogue of dinoflagellates is largely dominated by free-living forms (2,200 species, 92%). There are relatively few parasitic species (165 species, 7%), or mutualistic symbionts (21 species, < 1%). Of the 2,200 free-living species, 477 species (22%) are represented by at least one molecular marker. The SSU marker is available for 307 free-living species (14%), 331 species for LSU (15%), 156 species for ITS (7%), and 91 species for COI (4%, Table 1).

For the 165 parasitic species, there are 29 species (17%) with at least one available sequence. The best sequenced genus is *Blastodinium* with 11 sequenced species. The SSU rDNA sequence was available for all the parasitic species represented in GenBank (29 species, 17% of the 165 parasites). LSU rDNA sequences were available for 18 species (11%), and ITS for 18 species (11%). Only one sequence of the COI marker is available (0.6%, *Piscinoodinium limneticum*).

From the 21 mutualistic symbiotic species, there are sequences available for 17 species (80%). Most of them belong to the genus *Symbiodinium*. The only taxa lacking nucleotide sequences were *Scrippsiella velellae*, *Endodinium chattonii*, *Symbiodinium tridacnorum* and *S. fittii*; the latter with only a microsatellite sequence available (supplementary material Table S1). For the 21 species of mutualistic symbionts, the ribosomal markers (SSU, LSU, ITS) are available for 9–10 species each one, and COI sequences were available for five species (*Symbiodinium* and *Pelagodinium*).

DISCUSSION

Habitat bias

Differences between marine and freshwater species

Most dinoflagellate species inhabit marine waters and only 17% of the described species are found in freshwater environments. The freshwater dinoflagellates are highly dominated by plastid-containing species (88%), while in marine environments there is a slight dominance of heterotrophic species (58%) (Gómez 2012b). Consequently, it is easier to establish cultures and to obtain sequences of freshwater species. Overall, the freshwater habitats are usually more accessible and require fewer infrastructures for sampling when compared to open ocean areas. For these reasons, we can expect a higher representation of the freshwater species in the nucleotide databases. However, the percentage of sequenced freshwater species (13%) was lower than for the marine species (22%, Table 1). In part, the marine species are favored by interest in the marine harmful algal blooms (*Alexandrium*, *Dinophysis*, *Gambierdiscus*, etc.). However, the low percentage of sequenced species from freshwaters could also be related an overestimation of freshwater species richness (Thessen *et al*. 2012).

The accessibility of freshwater species favors species descriptions from occasional samplings in ponds, lakes or reservoirs near laboratories. The species descriptions are often too poor to allow the organism to be re-identified and the descriptions in a highly dispersed literature, often apparently unknown to other researchers. For example, *Cystodinium cornifax*, *Gymnodinium uberrimum*, *Prosoaulax lacustris* (= *Amphidinium lacustre*) or *Woloszynskia pascheri* have been described under more than ten heterotypic synonyms. An overestimation of species is more evident for the unarmored genera. The descriptions are often based on distorted cells due to the preservation treatment or live specimens that were stressed or moribund. Under *Gymnodinium*, *Gyrodinium*, *Amphidinium* and *Katodinium* have been pooled species from different phylogenetic origins. Since Daugbjerg *et al*. (2000), the spliting of these unarmoured genera has begun based on molecular data. A critical study is required to assess the disproportionate number of new species described by some authors (Schiller, Skvortzov, van Meel, Baumestier, Conrad, Christen, Harris, Campbell, Okolodkov). The description of new species using previously occupied names by Skvortzov, Christen, van

Meel or Campbell is evidence that these authors did not check all the existing literature. Thessen *et al*. (2012) reported that 38% of the species of *Gymnodinium* as 'oncers', not ever observed since their original description. Within the context of paucity of skilled taxonomists associated with the sequencing projects, an excess of poor quality species descriptions do not facilitate the identification at the species level. Thessen *et al*. (2012) reported that there are 250 sequences in GenBank that referred to *Gymnodinium*, but only 86 (30%) are labeled with a proper species name.

The freshwater thecate genera *Cystodinium* and *Hemidinium* contain an anomalously high number of species described by Baumeister (1957) and Skvortzov (1958, 1968), respectively. The genus *Glenodinium* is plagued with poor species descriptions, mainly described by Skvortzov (1958, 1968) and Schiller (1955). We can hypothesize two explanations, perhaps not mutually exclusive: A) these authors are good examples of splitter taxonomists, ignoring intraspecific variability and/or the previous species descriptions, or B) these authors were right and there is a considerable diversity and degree of endemism of dinoflagellate in freshwater environments.

Considering the first hypothesis – for authors such as Skvortzov – we can speculate that lack of access to all the scientific literature may have hindered appropriate discrimination among species. This is not the case of Schiller who authored a complete monograph on marine and freshwater dinoflagellates (Schiller 1937). It is difficult to evaluate whether their new species correspond to known species due to the scarce detail of the line drawings, especially for unarmored dinoflagellates. *Gonyaulax*, *Prorocentrum* or *Dinophysis* are well investigated as responsible for harmful proliferations and these thecate genera are less susceptible to cell shape changes. Schiller described species of *Gonyaulax* (*G. grabrielae*, *G. matkovicii*) that are only known from the original description. Most of the species of *Prorocentrum* and *Dinophysis* described by Schiller have been synonymized (Gómez 2012a). The observations of other *Dinophysis* species (i.e. *D. biceras*) are restricted to Schiller's description. This is highly unusual because Schiller described these species from areas of the coastal Mediterranean Sea (Gulf of Trieste, Naples) that has been intensively investigated. This suggests that many of the species described by Schiller were life stages of known species, teratogenical or damaged specimens.

The other alternative is to consider that these authors were right. There is a high diversity of dinoflagellates in freshwater environments, and their observations are evidence of a high degree of endemism because the taxa have not been observed in other places. The fact is that there are very few well documented examples of endemic dinoflagellates in freshwater environments (Annenkova *et al*. 2011). Environmental sequencing surveys have revealed a large diversity of dinoflagellates in marine waters, while the abundance and diversity of clades is considerably lower in freshwater environments (Slapeta *et al*. 2005, Richards *et al*. 2005, Lefèvre *et al*. 2008). Microscopical observations suggest a relatively low number of dinoflagellate lineages in freshwater environments. For example, members of the order Dinophysales are absent in freshwaters, while they reach a high diversity in marine waters (Hastrup Jensen and Daugbjerg 2009, Gómez *et al*. 2011). Gonyaulacales or Prorocentrales are scarcely represented in freshwaters (Logares *et al*. 2007). *Ceratium* is restricted to few freshwater species, while the marine relatives reach a high diversity (Gómez *et al.* 2010a). Basal dinoflagellates such as Syndiniales are abundant and genetically diverse in marine waters, but they are unknown in freshwater environments (Guillou *et al.* 2008). Schiller (1955) described several tens of new dinoflagellates from an Alpine lake, while more recent studies from 27 Alpine lakes found only a total of 34 dinoflagellates species (Hansen and Flaim 2007). For these reasons, it is more plausible to consider that the low percentage of sequenced freshwater species is largely influenced by an overestimation of species richness due to the activity of 'splitters'.

Acquiring material from the open sea usually requires more infrastructures than that needed for sampling in freshwater environments. Samples from the classical ocean expeditions were analyzed by skilled taxonomists (Kofoid, Jørgensen, Paulsen, Pavillard, etc.). The results were published in monographs or reports of scientific expeditions. This avoids the species re-descriptions due to previous records appearing only in many separate and difficult to access publications. Despite this, the phenomenon of a likely overestimation of species richness is not restricted to freshwater habitats. The most speciose dinoflagellate genus, *Protoperidinium*, contains 269 marine species that are easy to collect by net sampling and they are relatively well-preserved. The available sequences are restricted to 24 species that represent but 9% of the described species (Yamaguchi *et al*. 2006, Gribble and Anderson 2006). Numerous species described by Böhm, Gaarder, Mangin, Matzenauer, Meunier, Schiller and Wailes have been scarcely ever reported by other authors. Balech and Abé also described numerous species with detailed

studies of the tabulation. None of the species described by these authors has been sequenced (*Protoperidinium thulesense* and *P. steidingerae* are new names for previously known species). The exclusion of these doubtful species will provide more realistic values of the percentage of sequenced species.

The lack of sequences identified at the species level may be related to the difficulties in recognizing species during the isolation under the microscope. For example, Carbonell-Moore described twenty-five species and three genera of the family Podolampadaceae based on diagnostic characters that are only visible using scanning electron microscopy (Carbonell-Moore 1994). These species are present in low abundance in the under sampled warm waters of the open ocean. The diagnostic morphological characters used for species delimitation, even genera, are not easily discernible under light microscopy. These cells resemble small transparent 'balls', and the sequenced specimens are pooled as *Blepharocysta* sp. (Gómez *et al*. 2010b).

Takano and Horiguchi (2006) published an article entitled 'Acquiring scanning electron microscopical, light microscopical and multiple gene sequence data from a single dinoflagellate cell'. These authors concluded: 'This technique can be applied to both photosynthetic and heterotrophic dinoflagellates and will accelerate biodiversity studies'. According to this technique, a single dinoflagellate cell is fixed (Lugol, glutaraldehyde), postfixed (osmium tetroxide), followed by an alcohol dehydration series, critical point drying, sputter coating, vacuum, electron radiation, and later the DNA is extracted from the single cell to get sequences for several molecular markers. Many holotypes deposited in scanning electron microscopy stubs of museums could theoretically benefit from this technique. However, neither Takano and Horiguchi (2006) nor any author(s) have ever successfully applied this technique to a single cell from a natural sample. However, as a consequence of such reports purporting that a single cell can yield a wealth of data, studies of rare open ocean dinoflagellates can be easily rejected by critical reviewers using the false argument that one can easily obtain several different molecular markers and images using scanning electron microscopy from a single cell.

Differences between plankton and benthic species

Dinoflagellates are dominated by planktonic species, while benthic forms represent but 8% of the total species (Gómez *et al*. 2012b). The type of habitat, plankton or benthos, has also an influence on the availability of molecular markers. At least one sequence is available for 18% of the plankton species, while the percentage increases to 46% for the benthic species (Table 1). Several harmful groups (the epiphytes *Prorocentrum*, *Gambierdiscus*, *Ostreopsis*, *Coolia* and some species of *Amphidinium*), and the symbiont *Symbiodinium* are significant constituents of the sequenced species. With the exception of some insufficiently known freshwater benthic genera (*Cystodinium*, *Stylodinium*, *Tetradinium*), the benthic forms appear to be wellrepresented in GenBank compared to planktonic species.

In the earlier taxonomical studies, the descriptions of benthic species were sporadic (i.e. *Prorocentrum lima*), and often from accidental observations of resuspended material. Epiphytic dinoflagellates began to receive attention after Fukuyo in late 1970's, and Faust described numerous species in the 1990's. After a few pioneering studies on sand-dwelling dinoflagellates (Herdman, Balech), the number of genera and species has largely increased in the last years (Dodge, Hoppenrath, Horiguchi, Murray, Selina, Yoshimatsu). The sampling coverage of benthic species is nearly restricted to shallow waters as epiphytes of macrophytes, coral reefs or sandy beaches. Environmental sequencing surveys in deep ocean sediments reveal an unknown diversity of benthic dinoflagellates (López-García *et al*. 2007).

Differences with trophic mode

One half of the species of dinoflagellates described are heterotrophic (Gómez 2012b). However, while about 29% of the plastid-containing species have at least one sequence available, the figure is but 12% for heterotrophic species (Table 1). Fortunately, sequences of the SSU and LSU rDNA markers can be obtained from single cells that facilitate investigation of uncultivable species (Lynn and Pinheiro 2009). For other alternative markers, the bias continues towards the photosynthetic species available in cultures. For the COI marker, the number of sequenced plastid-containing species is 10 times greater than for the heterotrophic ones. Each time that a new molecular marker is presented as the ideal barcode marker, it is tested with the species available in cultures (Stern *et al*. 2012). These species constitute less than 10% of described species, and they do not represent the functional and ecological diversity of the dinoflagellates.

Difference with lifestyle

Relative representation in sequence availability also varies with dinoflagellate lifestyle. The relatively few

described symbiotic dinoflagellates associated with coral reefs such as *Symbiodinium* have received considerable attention (LaJeunesse *et al*. 2001, Pochon *et al*. 2012). *Symbiodinium* displays considerable genetic diversity, which suggests a high number of undescribed species (McNally *et al*. 1994). The diversity of mutualistic symbiotic dinoflagellates associated with planktonic Rhizaria (Acantharia, Foraminifera and Radiolaria) remains underestimated (Siano *et al*. 2010, Anderson 2014).

A total of 165 dinoflagellates species have been described as parasites (7%). The dinokaryotic parasites (86 species) are usually identified to the species level in the molecular studies (Coats *et al*. 2010, Gómez *et al*. 2009), and especially for *Blastodinium* (Skovgaard *et al.* 2012, Skovgaard 2014). The diversity of the basal dinoflagellates (79 species) is probably largely underestimated. Most of the sequences of basal parasitic dinoflagellates (*Amoebophrya*, *Euduboscquella*, *Hematodinium*, *Ichthyodinium* and *Syndinium*) have not been not identified to the species level (Guillou *et al*. 2008). Furthermore, the epithets of numerous parasitic forms are based on the host names such as *Amoebophrya cerati*, *A. leptodisci* (the dinoflagellates *Neoceratium* or *Leptodiscus*), *A. sticholonchae* (*Sticholonche*, Acantharea) or *A. tintinni* (ciliate tintinnids) and the use of the host identity is not a valid criterion for the identification of a parasitic species (Bachvaroff *et al*. 2012). Comparisons between host and parasite phylogenies do not suggest a simple pattern of host or parasite specificity. A single parasite species may infect different hosts, and a single host can be infected by several parasites (Bachvaroff *et al*. 2012). An effort is needed to find diagnostic characters for the identification of the parasitic basal dinoflagellates. This will contribute to reduce the excessive number of sequences that are not identified with the proper species name.

'Molecular fashion' as an obstacle to assessing open ocean diversity

The proliferations of harmful dinoflagellates have consequences in the public health realm. These species accounted for the less than 5% of the total dinoflagellate species (Sournia 1995) but the intensive study of toxic species has largely contributed to our overall knowledge of dinoflagellates. However, the bias towards the harmful coastal species should not adversely affect studies of the other groups of dinoflagellates. The most intensively investigated species (*Alexandrium*) are photosynthetic forms available in cultures, and other species easily accessible in coastal waters. The availability of material facilitates detailed morphological, ultrastructural and molecular studies. In contrast, most of the open ocean dinoflagellates are difficult to culture, few specimens are available, and sampling is expensive because they are far from the laboratories. It is true that is possible to get a sequence from a single cell. However, the percentage of success with oceanic species is lower than from the DNA extracted from cultures.

The requirement of molecular data for any publication has adverse consequences in the description of new species from the open ocean. A total of twenty-six new species have been described since 2010. All of them are marine species, with the exception of *Prorocentrum rivalis*. Most of these new taxa were described from cultures and most of them were photosynthetic species, with only seven heterotrophs (26%). These new plankton species are barely distinctive taxa that until recently were overlooked and lumped with other known species. Six of the recently described species belong to the group of *Gymnodinium sensu stricto* (relatives of harmful species such as *Gymnodinium catenatum*). The other planktonic species were *Azadinium* and *Vulcanodinium* also investigated as potentially toxic species. Some parasites (i.e. *Amoebophrya*) are of ecological interest for the control of harmful algal blooms. The recently described genera *Tintinnophagus* and *Euduboscquella* have benefited in part from interest in parasitism. On the other hand, about one half of recent dinoflagellate descriptions (eleven species) were benthic taxa, dominated by epiphytic species (five species of *Prorocentrum*, *Gambierdiscus excentricus*, and *Coolia malayensis*), and in the last year a few species have been described from the germination of sediment cysts in paleontological studies (*Archaeperidinium saanichi*, *Scrippsiella bicarinata*, *S. kirschiae*). Of the species recently described all were from coastal waters in latitudes higher than 40º North, with the exception of benthic species from tropical waters or from the germination of cysts. The notorious high diversity of open tropical waters is nearly unexplored, while the low-diversity cold waters near the famous institutions of high latitudes are more intensively investigated.

It is noteworthy that there has not been a single description of a new species collected from the water column of the open ocean in the last three years. As an example, one can compare the temporal trends in the species descriptions in 2011 and one century ago. A total of 80 and 12 new species were described in 1911 and 2011, respectively. Most of the species described in

1911 were heterotrophs (52 species, 64%). They were collected from the open ocean expeditions, usually in tropical seas. Only one benthic species was described in 1911 (*Amphidinium herdmanii*). In 2011, nearly all the species descriptions corresponded to photosynthetic species (92%), with the exception of *Gyrodiniellum shiwhaense*. Five species of the twelve new descriptions in 2011 (41%) were benthic species.

The diversity of the open warm ocean has no economic interest; sampling is expensive because it is far from the specialized laboratories of higher latitudes and requires infrastructure (ship time). The plankton of the open waters is diverse and requires skilled taxonomists who need long apprenticeships. The phytoplankton guides are restricted to common coastal species of high latitudes. The specimens of the open waters are difficult to culture and often the few specimens available do not allow detailed descriptions. Overall, reports of biodiversity in open waters are handicapped compared to the detailed studies possible with the species available in culture and often severely treated by reviewers. This is a vicious circle, the research projects in the open sea are expensive, and the few publications resulting do not receive citations to justify the received financial support. The consequence is the decline of the knowledge of the open ocean diversity based on microscopical observations. Distinctive species remain unreported, while the new species descriptions are concentrated in well investigated regions and species and they are hardly differentiable from other known species. The criteria for admission to new species need to be more relaxed for the distinctive species of the open ocean.

Final remarks

This study reviewed the availability of molecular markers in dinoflagellates, functional and ecologically diverse of protist group. Numerous biases exist towards some groups that are easy to sequence due to the high abundance (cultivable species), while groups such as open ocean dinoflagellates are scarcely represented. The descriptions of new species based on cultures or common coastal species permit highly detailed studies (including ultrastructure, multi-gene phylogenies). However, these techniques are difficult to apply to the few specimens available from the open ocean. For these rare species, the reviewers often seem to require use of the same techniques as those used when billions of cells are available in cultures. This attitude hinders the advances and the sequences of open ocean genera remain unpublished. A recommendation is that the criteria for admission to new species need to be more relaxed for the distinctive new species of the open ocean.

The bias in sequence availability is due also to a decline in general taxonomic expertise. It is difficult to envision a solution within a context of budget reductions, and even the tentative closure of institutions with historical traditions in taxonomical studies, especially in Mediterranean Sea. Taxonomists have to be open to working in other regions that merit biodiversity exploration (i.e. Latin America or Asia). This also can contribute to reduce the bias in the geographical coverage of taxonomical studies, excessively focused on coastal waters of high latitudes of the north hemisphere (> 40º N).

From an academic point view, the financial support for the completion of the doctoral studies, often three years, is insufficient for training with regard to the recognition of the species from the open ocean. For practical and financial reason, PhD studies trends to be focused on a discrete groups, mainly easy accessible coastal species and more usually harmful species. More flexibility in study programs is needed in order to take into account the longer apprenticeships for the formation of expertise in open ocean diversity.

Courses of phytoplankton identification are highly limited in time, and it is difficult to cover the high diversity with material from the open ocean. Courses and identification guides are restricted to the harmful or common coastal species. The Internet has facilitated the diffusion of the literature such as the classical monographs that are freely available on-line. Many species remain restricted to line drawings of the original descriptions, and we have to compile that material and provide micrographs to facilitate the species recognition in update guide for the ocean dinoflagellates.

The decline in taxonomic expertise is evident based on the high number of sequences that are not identified to the species level, or often misidentified. Other evidence is the increase of the misspelling of the species names. Some journals request that authors check the species names in biodiversity websites. Often, the usefulness of these websites is measured as the number of species names that listed. Dinoflagellate lists with the correct species names are available (Gómez 2012a). The species names in the biodiversity websites are often inflated with several lexical variants that may potentially be typed by users using the search engines. The consequence may be an artificial increase of the numbers of species; also it may be difficult for the users to discern the correct spelling of the species.

'Splitter' taxonomists have inflated some groups of dinoflagellates with insufficiently described species that we may consider "literature ghosts". It is difficult in taxonomical revisions to synonymize such species with known species due to the poor original descriptions, especially in the unarmored species. In this study we argue that the apparently low number of sequenced species in some groups of dinoflagellates (*Protoperidinium*, *Gymnodinium*) is mainly due an overestimation of species richness in some groups. The huge number of species of some genera (> 100 species) likely discourages many workers from trying to make an identification to the species level. Again, more flexibility in the establishment of synonym of the "literature ghosts" is recommended to adapt the number of species to the real number of known species.

The sequences available in GenBank are not identified at the species level, or misidentified. The sequences are listed in a classification with numerous errors that will be easily corrected. The metadata associated with each sequence is not complete. In many cases, it is not easy to verify the identity of the species. Often the origin of the environmental sequences, even identified species, is labeled with non-informative data such as "ocean" that are not useful for authors working on biogeographical studies.

The collaboration between classical microscopists and molecular biologists have facilitated advances. Both are concerned with the big questions in the evolution of dinoflagellates, such as the interrelationships between of the major order, or to identify new clades revealed by environmental sequences. These topics are relevant for journals of high impact that help obtain funds for molecular analyses. However, these journals usually do not have space for extensive taxonomical reviews nor nomenclatural considerations that only delay publication. Some topics such as the in-group relations, the speciation inside a genus, are relevant for harmful species. However, for oceanic species the in-group relations requires a considerable effort that is not compensated in terms of citations to such articles. This discourages further studies, and many genera remain represented by a sequence.

Acknowledgements. F. G. is supported by the contract JCI-2010-08492 of the Ministerio Español de Ciencia y Tecnología, and the Brazilian contract BJT 370646/2013-14 by Conselho Nacional de Desenvolvimento Científico e Tecnológico. I would like to thank J. R. Dolan, D. J. S. Montagnes and two reviewers for their time and valuable remarks.

REFERENCES

Anderson O. R. (2014) Living Together in the Plankton: A survey of marine protist symbioses. *Acta Protozool*. **53:** 29–38

Annenkova N. V., Lavrov D. V., Belikov S. I. (2011) Dinoflagellates associated with freshwater sponges from the ancient lake Baikal. *Protist* **162:** 222–236

Bachvaroff T. R., Kim S., Guillou L., Delwiche C. F., Coats D. W. (2012) Molecular diversity of the syndinean genus *Euduboscquella* based on single-cell PCR analysis. *Appl. Environ. Microbiol.* **78:** 334–345

Baumeister W. (1957) Neue Dinococcalen aus dem niederbayerischen Hiigelland zwischen Isar und Inn (I). *Arch. Protistenk* **102:** 21–43

Blaxter M. (2004) The promise of a DNA taxonomy. *Phil. Trans. R. Soc. B* **359:** 669–679

Carbonell-Moore M. C. (1994) On the taxonomy of the family Podolampadaceae Lindemann (Dinophyceae) with descriptions of three new genera. *Rev. Palaeobot. Palynol.* **84:** 73–99

Coats D. W., Kim S., Bachvaroff T. R., Handy S. M., Delwiche C. F. (2010) *Tintinnophagus acutus* n. g., n. sp. (Phylum Dinoflagellata), an ectoparasite of the ciliate *Tintinnopsis cylindrica* Daday 1887, and its relationship to *Duboscquodinium collini* Grassé 1952. *J. Eukaryot. Microbiol.* **57:** 468–482

Daugbjerg N., Hansen G., Larsen J., Moestrup Ø. (2000) Phylogeny of some of the major genera of dinoflagellates based on ultrastructure and partial LSU rDNA sequence data, including the erection of three new genera of unarmoured dinoflagellates. *Phycologia* **39:** 302–317

Ehara M., Imaga Y., Kazuo K., Watanabe I., Ohama T. (2000) Phylogenetic analysis of diatom coxI genes and implications of a fluctuating GC content on mitochondrial code evolution. *Curr. Gen.* **37:** 29–33

Evans K. M., Wortleya A. H., Manna D. G. (2007) An assessment of potential diatom "barcode" genes (cox1, rbcL, 18S and ITS rDNA) and their effectiveness in determining relationships in *Sellaphora* (Bacillariophyta). *Protist* **158:** 349–364

Gómez F., Moreira D., López-García P. (2009) Life cycle and molecular phylogeny of the dinoflagellates *Chytriodinium* and *Dissodinium*, ectoparasites of copepod eggs. *Eur. J. Protistol.* **45:** 260–270

Gómez F., Moreira D., López-García P. (2010a) *Neoceratium* gen. nov., a new genus for all marine species currently assigned to *Ceratium* (Dinophyceae). *Protist* **161:** 35–54

Gómez F., Moreira D., López-García P. (2010b) Molecular phylogeny of the dinoflagellates *Podolampas* and *Blepharocysta* (Peridiniales, Dinophyceae). *Phycologia* **49:** 212–220

Gómez F., López-García P., Moreira D. (2011) Molecular phylogeny of dinophysoid dinoflagellates: The systematic position of *Oxyphysis oxytoxoides* and the *Dinophysis hastata* group (Dinophysales, Dinophyceae). *J. Phycol.* **47:** 393–406

Gómez F. (2012a) A checklist and classification of living dinoflagellates (Dinoflagellata, Alveolata). *CICIMAR Océanides* **27:** 65–140

Gómez F. (2012b) A quantitative review of the lifestyle, habitat and trophic diversity of dinoflagellates (Dinoflagellata, Alveolata). *Syst. Biodivers.* **10:** 267–275

Gribble K. E., Anderson D. M. (2006) Molecular phylogeny of the heterotrophic dinoflagellates, *Protoperidinium*, *Diplopsalis* and *Preperidinium* (Dinophyceae), inferred from large subunit rDNA. *J. Phycol.* **42:** 1081–1095

Guillou L., Viprey M., Chambouvet A., Welsh R. M., Kirkham A. R., Massana R., Scanlan D. J., Worden A. Z. (2008) Widespread occurrence and genetic diversity of marine parasitoids belonging to Syndiniales (Alveolata). *Environ. Microbiol.* **10:** 3349–3365

Hackett J. D., Anderson D., Erdner D. L., Bhattacharya D. (2004) Dinoflagellates: A remarkable evolutionary experiment. *Am. J. Bot.* **91:** 1523–1534

Hansen G., Flaim G. (2007) Dinoflagellates of the Trentino Province, Italy. *J. Limnol.* **66:** 107–141

Hastrup Jensen M., Daugbjerg N. (2009) Molecular phylogeny of selected species of the order Dinophysiales (Dinophyceae) – testing the hypothesis of a Dinophysioid radiation. *J. Phycol.* **45:** 1136–1152

Hebert P. D. N., Ratnasingham S., DeWaard J. R. (2003) Barcoding animal life: cytochrome c oxidase subunit 1 divergences among closely related species. *Proc. R. Soc. Lond. B* **270:** S96–S99

Henze K., Martin W. (2003) Essence of mitochondria. *Nature* **426:** 127–128

Hillis D. M., Dixon M. J. (1991) Ribosomal DNA: molecular evolution and phylogenetic inference. *Q. Rev. Biol.* **66:** 411–453

Hinnebusch A. G., Klotz L. C., Blanken R. L., Loeblich III A. R. (1981) An evaluation of the phylogenetic position of the dinoflagellate *Crypthecodinium cohnii* based on 5S rRNA characterization. *J. Mol. Evol*. **17:** 334–347

LaJeunesse T. S. (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. *J. Phycol.* **37:** 866–880

LaJeunesse T. C., Lambert G., Andersen R. A., Coffroth M. A., Galbraith D. W. (2005) *Symbiodinium* (Pyrrhophyta) genome sizes (DNA content) are smallest among dinoflagellates. *J. Phycol.* **41:** 880–886

Lefèvre E., Roussel B., Amblard C., Sime-Ngando T. (2008) The molecular diversity of freshwater picoeukaryotes reveals high occurrence of putative parasitoids in the plankton. *PLoS ONE* **3:** e2324

Le Gall L., Saunders G. W. (2010) DNA barcoding is a powerful tool to uncover algal diversity: a case study of the Phyllophoraceae (Gigartinales, Rhodophyta) in the Canadian flora. *J. Phycol.* **46:** 374–389

Lenaers G., Scholin C., Bhaud Y., Saint-Hilaire D., Herzog M. (1991) A molecular phylogeny of dinoflagellate protists (Pyrrhophyta) inferred from the sequence of 24S rRNA divergent domains D1 and D8. *J. Mol. Evol.* **32:** 53–63

Lin S., Zhang H., Hou Y., Zhuang Y., Miranda L. (2009) High-level diversity of dinoflagellates in the natural environment, revealed by assessment of mitochondrial cox1 and cob genes for dinoflagellate DNA barcoding. *Appl. Environ. Microbiol.* **75:** 1279–1290

Lin S. (2011) Genomic understanding of dinoflagellates. *Res. Microbiol.* **162:** 551–569

Litaker R. W., Vandersea M. W., Kibler S. R., Reece K. S., Stokes N. A., Lutzoni F. M., Yonish B. A., West M. A., Black M. N. D., Tester P. A. (2007) Recognizing dinoflagellate species using ITS rDNA sequences. *J. Phycol.* **43:** 344–355

Logares R., Shalchian-Tabrizi K., Boltovskoy A., Rengefors K. (2007) Extensive dinoflagellate phylogenies indicate infrequent marine-freshwater transitions. *Mol. Phylogen. Evol.* **45:** 887–903

López-García P., Vereshchaka A., Moreira D. (2007) Eukaryotic diversity associated with carbonates and fluid-seawater interface in Lost City hydrothermal field. *Environ. Microbiol.* **9:** 546–554

Lynn D., Pinheiro M. (2009) A survey of Polymerase Chain Reaction (PCR) amplification studies of unicellular protists using single-cell PCR. *J. Eukaryot. Microbiol.* **56:** 406–412

McNally K. L., Govind N. S., Thomé P. E., Trench R. K. **(**1994) Small-subunit ribosomal DNA sequence analysis and a reconstruction of the inferred phylogeny among symbiotic dinoflagellates (Pyrrophyta). *J. Phycol.* **30:** 316–329

Miller S. E. (2007) DNA barcoding and the renaissance of taxonomy. *Proc. Nat. Acad. Sci. USA* **104:** 4775–4776

Newmaster S. G., Fazekas A. J., Ragupathy S. (2006) DNA barcoding in land plants: evaluation of rbcL in a multigene tiered approach. *Can. J. Bot.* **84:** 335–341

Pochon X., Putnam H. M., Burki F., Gates R. D. (2012) Identifying and characterizing alternative molecular markers for the symbiotic and free-living dinoflagellate genus *Symbiodinium*. *PLoS ONE* **7:** e29816

Richards T. A., Vepritskiy A. A., Gouliamova D. E., Nierzwicki-Bauer S. A. (2005) The molecular diversity of freshwater picoeukaryotes from an oligotrophic lake reveals diverse, distinctive and globally dispersed lineages. *Environ. Microbiol*. **7:** 1413–1425

Rowan R., Whitney S. M., Fowler A., Yellowlees D. (1996) Rubisco in marine symbiotic dinoflagellates: Form II enzymes in eukaryotic oxygenic phototrophs encoded by a nuclear multigene family. *Plant Cell* **8:** 539–553

Saldarriaga F., Taylor F. J. R., Keeling P. J., Cavalier-Smith T. (2001) Dinoflagellate nuclear SSU rRNA phylogeny suggests multiple plastid losses and replacements. *J. Mol. Evol.* **53:** 204–213

Saldarriaga J. F., Taylor F. J. R., Cavalier-Smith T., Menden-Deuer S., Keeling P. J. (2004) Molecular data and the evolutionary history of dinoflagellates. *Eur. J. Protistol.* **40:** 85–111

Saunders G. W., Hill D. R. A., Sexton J., Andersen R. A. (1997) Small subunit ribosomal RNA sequences from selected dinoflagellates: testing classical evolutionary hypotheses with molecular systematic methods. In: Origins of Algae and Their Plastids, (Ed. D. Bhattacharya). Springer, New York, pp. 237–259

Schiller J. (1937) Dinoflagellatae (Peridineae) in monographischer Behandlung. Teil 2. In: Dr. L. Rabenhorst's Kryptogamen-Flora von Deutschland, Österreich und der Schweiz. Akademische Verlagsgesellschaft, Leipzig, 589 pp.

Schiller J. (1955) Untersuchungen an den planktischen Protophyten des Neusiedlersees 1950–1954, I. Teil. *Wissenschaftliche Arbeiten aus dem Burgenland* **9:** 1–66

Siano R., Montresor M., Probert I., Not F., de Vargas C. (2010) *Pelagodinium* gen. nov. and *P. beii* comb. nov., dinoflagellate symbiont of planktonic foraminifera. *Protist* **161**: 385–399

Skovgaard A., Karpov S. A., Guillou L. (2012) The parasitic dinoflagellates *Blastodinium* spp. inhabiting the gut of marine, planktonic copepods: morphology, ecology and unrecognized species diversity. *Frontiers Microbiol.* **3:** 305

Skovgaard A. (2014) Dirty tricks in the plankton: Diversity and role of marine parasitic protists. *Acta Protozool*. **53:** 51–62

Skvortzov B. V. (1958) New and rare Flagellatae from Manchuria, Eastern Asia. *Philip. J. Sci.* **86:** 139–202

Skvortzov B. V. (1968) New and little known Peridineae from northern Manchuria, China. *Q. J. Taiwan Mus.* **21:** 79–114

Slapeta J., Moreira D., López-García P. (2005) The extent of protist diversity: insights from molecular ecology of freshwater eukaryotes. *Proc. R. Soc. B* **272:** 2073–2083

Sournia A. (1995) Red tide and toxic marine phytoplankton of the world ocean: an inquiry into biodiversity. In: Harmful Marine Algal Blooms, (Eds. Lassus *et al*.). Lavoisier, Paris, pp. 103–112

Stern R. F., Andersen R. A., Jameson I., Küpper F. C., Coffroth M. A., Vaulot D., Le Gall F., Véron B., Brand J. J., Skelton H., Kasai F., Lilly E. L., Keeling P. J. (2012) Evaluating the ribosomal internal transcribed spacer (ITS) as a candidate dinoflagellate barcode marker. *PLoS ONE* **7:** e42780

Takabayashi M., Santos S. R., Cook C. B. (2004) Mitochondrial DNA phylogeny of the symbiotic dinoflagellates (*Symbiodinium*, Dinophyta). *J. Phycol.* **40:** 160–164

Takano Y., Horiguchi T. (2006) Acquiring scanning electron microscopical, light microscopical and multiple gene sequence data from a single dinoflagellate cell. *J. Phycol*. **42:** 251–256

Taylor F. J. R. (1987) *The Biology of Dinoflagellates*. Blackwell, Oxford, 785 pp.

Thessen A. E., Patterson D. J., Murray S. A. (2012) The taxonomic significance of species that have only been observed once: The genus *Gymnodinium* (Dinoflagellata) as an example. *PLoS ONE* **7:** e44015

Yamaguchi A., Kawamura H., Horiguchi T. (2006) A further phylogenetic study of the heterotrophic dinoflagellate genus, *Protoperidinium* (Dinophyceae) based on small and large subunit ribosomal RNA gene sequences. *Phycol. Res.* **54:** 317–329

Zhang H., Bhattacharya D., Lin S. (2007) A three-gene dinoflagellate phylogeny suggests monophyly of Prorocentrales and a basal position for *Amphidinium* and *Heterocapsa*. *J. Mol. Evol.* **65:** 463–474

Received on 4th March, 2013; Revised on 18th March, 2013; accepted on 21st March, 2013