UNIVERSIDADE FEDERAL DO ESTADO DO RIO DE JANEIRO

PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIAS BIOLÓGICAS

MESTRADO EM CIÊNCIAS BIOLÓGICAS

Amanda Santos Goulart

Caracterização genética de cepas brasileiras e filogeografia do complexo de espécies *Prorocentrum lima* (Dinophyceae).

Rio de Janeiro

2022

Amanda Santos Goulart

Caracterização genética de cepas brasileiras e filogeografia do complexo de espécies *Prorocentrum lima* (Dinophyceae).

Dissertação apresentada ao Programa de Pós-Graduação em Ciências Biológicas (Biodiversidade Neotropical) da Universidade Federal do Estado do Rio de Janeiro como requisito para obtenção do grau de Mestre em Ciências Biológicas.

Orientador: Prof. Dr. Fabiano Salgueiro

Coorientadora: Profª. Drª. Silvia Mattos Nascimento.

Rio de Janeiro

2022

Catalogação informatizada pelo(a) autor(a)

Goulart, Amanda Santos
G694 Caracterização genética de cepas brasileiras e filogeografia do complexo de espécies *Prorocentrum lima* (Dinophyceae) / Amanda Santos Goulart. -- Rio de Janeiro, 2022.
53f

Orientador: Fabiano Salgueiro.
Coorientadora: Silvia Mattos Nascimento.
Dissertação (Mestrado) - Universidade Federal do Estado do Rio de Janeiro, Programa de Pós-Graduação em Ciências Biológicas, 2022.

1. Prorocentraceae. 2. Oceano Atlântico Sul. 3. rDNA. 4. Dinofisistoxinas. 5. Ácido Ocadaico. I. Salgueiro, Fabiano, orient. II. Nascimento, Silvia Mattos, coorient. III. Título.

Amanda Santos Goulart

CARACTERIZAÇÃO GENÉTICA DE CEPAS BRASILEIRAS E FILOGEOGRAFIA DO COMPLEXO DE ESPÉCIES PROROCENTRUM LIMA (DINOPHYCEAE)

Dissertação apresentada ao Programa de Pós-graduação em Ciências Biológicas (Biodiversidade Neotropical) da Universidade Federal do Estado do Rio de Janeiro, como requisito para obtenção do título de Mestre(a) em Ciências Biológicas.

Aprovada em 20 de junho de 2022.

Banca Examinadora

Dr(a). Fabiano Salgueiro (Orientador)
(Universidade Federal do Estado do Rio de Janeiro/UNIRIO)

Dr(a). Paulo Paiva
(Universidade Federal do Rio de Janeiro/UFRJ)

FRAGA RIVAS, SANTIAGO (FIRMA)

Firmado digitalmente por FRAGA RIVAS, SANTIAGO (FIRMA)
Fecha: 2022.06.21 18:19:22 +02'00'

Dr(a). Santiago Fragas Rivas
(Instituto Español de Oceanografia /IEO)

Agradecimentos

Gostaria de agradecer aos meus pais, minha irmã e meu noivo, que sempre estiveram ao meu lado, me apoiando e incentivando em todos os momentos.

Agradeço a Deus e a Nossa Senhora das Graças por terem me dado forças, fé e esperança em todos os momentos difíceis.

Gostaria de agradecer aos meus orientadores, Fabiano Salgueiro e Silvia Mattos Nascimento por todo o ensinamento, auxilio e orientação especialmente nestes momentos difíceis de pandemia.

Agradeço aos colegas de laboratório por toda a ajuda e apoio, especialmente ao Vinicius, Karol, Felipe, Fernanda e Lohayne. Agradeço, também, aos colegas de mestrado que dividiram comigo todos os momentos difíceis e desesperadores do mestrado, especialmente a minha querida amiga Raquel Marques.

Por fim, agradeço a FAPERJ e a UNIRIO pelo financiamento ao projeto e ao PPGBIO por possibilitar o desenvolvimento deste trabalho.

Sumário:

Resumo

O dinoflagelado bentônico *Prorocentrum lima* é encontrado em latitudes temperadas e tropicais de todo o mundo, sendo considerado um complexo de espécies ("*Prorocentrum lima* species complex" - PLSC). As cepas do complexo sintetizam toxinas como o ácido ocadaico (OA) e dinofisistoxinas (DTXs), responsáveis pela síndrome gastrointestinal conhecida como DSP (*Diarrhetic Shellfish Poisoning*). O presente trabalho visa revisitar o PLSC, incluindo cepas isoladas da costa e ilhas oceânicas brasileiras, integrando análises filogenéticas e filogeográficas. Os loci ITS (*Internal Transcripted Space*), LSU (*Large SubUnit*) e SSU (*Small Subunit*) do DNA ribossomal (rDNA) de 23 cepas brasileiras foram sequenciados e analisados em conjunto com sequencias do rDNA de 191 cepas do PLSC disponíveis no Genbank. Nas análises filogenéticas realizadas usando Máxima Verossimilhança e Inferência Bayesiana as sequências brasileiras agruparam-se em três (clados 1, 2 e 3) dos quatro clados do PLSC observados nas árvores baseadas nos loci ITS e LSU e em dois (clados 1+2 e 3) dos três clados nas reconstruções filogenéticas baseadas no locus SSU. As distâncias genéticas entre os quatro clados do PLSC, principalmente no marcador ITS, foram maiores do que as observadas entre outras espécies de *Prorocentrum*. A análise bayesiana da estrutura populacional (BAPS) baseada nos loci ITS, LSU e SSU revelou 4, 5 e 3 *clusters*, respectivamente. Estes *clusters* correspondem aos principais clados e subclados do PLSC observados nas reconstruções filogenéticas. Considerando as análises filogeográficas, foram encontrados 18 haplótipos do marcador ITS, 21 do LSU e cinco do SSU. De acordo com os mapas de distribuição haplotípica, os haplótipos referentes aos clados 1 e 3 (ITS e LSU) e clado 1+2 e 3 (SSU) apresentaram uma ampla distribuição, ocorrendo em regiões tropicais e subtropicais dos Oceanos Atlântico, Pacífico e Indico. Os haplótipos referentes ao clado 2 (ITS e LSU) ocorreram exclusivamente no Oceano Atlântico. Os haplótipos referentes ao clado 4 foram observados em regiões temperadas do Oceano Atlântico Norte, Mar Mediterrâneo e Oceano Pacífico Sul. Levando-se em consideração todas as análises realizadas, conclui-se que o PLSC é composto por pelo menos três linhagens genéticas distintas, com distribuições biogeográficas parcialmente sobrepostas, que podem corresponder à diferentes espécies.

Abstract

The benthic dinoflagellate *Prorocentrum lima* is found in temperate and tropical latitudes around the world, being considered a species complex ("*Prorocentrum lima* species complex" - PLSC). The strains of the complex synthesize toxins such as okadaic acid (OA) and dinophysistoxins (DTXs), responsible for the gastrointestinal syndrome known as DSP (Diarrhetic Shellfish Poisoning). The present work aims to revisit the PLSC, including strains isolated from the coast and Brazilian oceanic islands, integrating phylogenetic and phylogeographic analyses. The ITS (Internal Transcripted Space), LSU (Large SubUnit) and SSU (Small Subunit) loci of ribosomal DNA (rDNA) from 23 Brazilian strains were sequenced and analyzed together with rDNA sequences from 191 PLSC strains available in Genbank. Phylogenetic analyses were performed using Maximum Likelihood and Bayesian Inference, the Brazilian sequences clustered into three (clades 1, 2 and 3) of the four PLSC clades in phylogenetic reconstructions based on the ITS and LSU loci and in two (clades 1+2 and 3) of the three clades in the phylogenetic reconstructions based on the SSU locus. The genetic distances between the four PLSC clades, mainly in the ITS marker, were greater than those observed among other *Prorocentrum* species. Bayesian population structure analysis (BAPS) based on the ITS, LSU and SSU loci revealed 4, 5 and 3 clusters, respectively. These clusters correspond to the main PLSC clades and subclades observed in the phylogenetic reconstructions. Considering the phylogeographic analyses, 18 haplotypes were found for the ITS marker, 21 for the LSU and five for the SSU. According to the haplotypic distribution maps, the haplotypes referring to clades 1 and 3 (ITS and LSU) and clade 1+2 and 3 (SSU) showed a wide distribution, occurring in tropical and subtropical regions of the Atlantic Oceans, Pacific and Indian. The haplotypes referring to clade 2 (ITS and LSU) occurred exclusively in the Atlantic Ocean. The haplotypes referring to clade 4 were observed in temperate regions of the North Atlantic Ocean, Mediterranean Sea and South Pacific Ocean. Taking into account all the analyses carried out, it is concluded that the PLSC is composed of at least three distinct genetic lineages, with partially overlapping biogeographic distribution, which may correspond to different species.

Introdução

O gênero de dinoflagelado *Prorocentrum* foi descrito por Ehrenberg (1834), com *Prorocentrum micans* como a espécie-tipo. Atualmente, apresenta 84 espécies descritas e consideradas válidas (Guiry e Guiry, 2022). A classificação das espécies de *Prorocentrum* é baseada em características morfológicas, sendo elas, a análise da forma e do tamanho das células; a ornamentação da superfície da teca, a banda intercalar e as minúcias arquitetônicas da área periflagelar. Entretanto, essas características podem ser variáveis e as diferenças entre espécies podem ser sutis e podem apresentar plasticidade fenotípica (Hoppenrath et al., 2013). Logo, as análises genéticas são importantes para auxiliar na identificação das espécies (Hoppenrath et al., 2013).

Dessa maneira, sequências de genes do DNA ribossomal (rDNA) têm sido muito utilizadas na identificação e compreensão das relações filogenéticas entre espécies do gênero *Prorocentrum* (Nagahama et al., 2011; Zhang et al., 2015; Nascimento et al., 2016; Nascimento et al., 2017; Nishimura et al., 2020). Nestas análises é importante o uso de mais de um marcador molecular, devido as suas taxas evolutivas distintas. Por exemplo, as sequências do loci ITS (*Internal Transcribed Space*) do rDNA mostraram-se bastante variáveis, revelando uma alta diversidade inter e até mesmo intra-específica no gênero *Prorocentrum* (Stern et al., 2012). Já as sequências do loci LSU (*Large SubUnit*) e SSU (*Small Subunit*) do rDNA, por serem mais conservadas, são mais adequadas para avaliar a distinção de espécies (Boopathi et al., 2015).

As espécies de *Prorocentrum* estão distribuídas mundialmente em habitats planctônicos e bentônicos. Dentre as espécies bentônicas do gênero, *Prorocentrum lima* (Ehrenberg) F. Stein, 1878 possui distribuição cosmopolita, ocorrendo desde ambientes tropicais até temperados (Nagahama et al., 2011). *Prorocentrum lima* é produtor de toxinas como o ácido ocadáico (AO) e as dinofisistoxinas (DTXs) (e.g., Bravo et al., 2001; Hoppenrath et al., 2013; Nascimento et al., 2016; Nishimura et al., 2020) que podem causar problemas de saúde pública como a intoxicação diarreica por molusco (*Diarrhetic Shellfish Poisoning* – DSP) (Murata et al., 1982).

A descrição original de *P. lima* foi feita por Ehrenberg (1860) sob o nome de *Cryptomonas lima*. A sua redescrição foi feita por Nagahama e Fukuyo (2005) considerando as características morfológicas do organismo coletado na localidade tipo, o Golfo de Nápoles, Itália. Entretanto, devido à grande variação da forma das células de *P. lima* provenientes de diferentes localidades, Aligizaki et al. (2009) propôs o uso do termo "complexo de espécies *P. lima"* ("*Prorocentrum lima* species complex" - PLSC). Em seguida, Nagahama et al. (2011) sugeriu que a espécie *Prorocentrum arenarium* (FAUST, 1994), de forma mais arredondada que *P. lima*, que tem forma oval, poderia ser considerada como uma variação morfológica de *P. lima* e sinonimizou as duas espécies.

Nos últimos anos diversas reconstruções filogenéticas baseadas nos loci ITS e LSU do rDNA tem revelado que o complexo de espécies *P. lima* é formado por quatro clados principais, com moderado-elevado suporte (Nagahama et al., 2011; Nascimento et al., 2016; Nascimento et al., 2017; Nishimura et al., 2020). Além disso, esses trabalhos têm sugerido que existe certa relação entre o local de ocorrência dos espécimes e alguns desses clados. Assim, o presente trabalho visa revisitar o complexo de espécies *P. lima*, usando sequências de todo o mundo, incluindo novas cepas isoladas do Oceano Atlântico Sul (Brasil), empregando uma abordagem que integra análises filogenéticas e filogeográficas.

Capítulo Único

Genetic characterization of Brazilian strains and phylogeography of the *Prorocentrum lima* species complex (Dinophyceae).

Amanda S. Goulart[1]*, Silvia M. Nascimento[2], Vinicius C. P. Magalhães[1], Fabiano Salgueiro[1]

1. Laboratório de Biodiversidade e Evolução Molecular, Universidade Federal do Estado do Rio de Janeiro (UNIRIO), Av. Pasteur, 458, 512, Rio de Janeiro, RJ 22290-240, Brazil
2. Laboratório de Microalgas Marinhas, Universidade Federal do Estado do Rio de Janeiro (UNIRIO), Av. Pasteur, 458, 314-B, Rio de Janeiro, RJ 22290-240, Brazil

*Corresponding author
AS Goulart
E-mail address: amandagoulart@edu.unirio.br
Telephone contact: +552132155509

Abstract

The benthic dinoflagellate *Prorocentrum lima* is found in temperate and tropical latitudes around the world, being considered a species complex ("*Prorocentrum lima* species complex" - PLSC). The strains of the complex synthesize toxins such as okadaic acid (OA) and dinophysistoxins (DTXs), responsible for the gastrointestinal syndrome Diarrhetic Shellfish Poisoning. The present work aims to revisit the PLSC, including strains isolated from the coast and Brazilian oceanic islands, integrating phylogenetic and phylogeographic analyses. The ITS (Internal Transcribed Space), LSU (Large SubUnit) and SSU (Small Subunit) loci of ribosomal DNA (rDNA) from 23 Brazilian strains were sequenced and analyzed together with rDNA sequences from 191 PLSC strains available in Genbank. Phylogenetic analyzes were performed using Maximum Likelihood and Bayesian Inference. The Brazilian sequences clustered in three clades (1, 2 and 3) of the four PLSC clades in phylogenetic reconstructions based on the ITS and LSU loci and in two (clades 1+2 and 3) of the three clades in the phylogenetic reconstructions based on the SSU locus. The genetic distances between the four PLSC clades, mainly in the ITS marker, were greater than those observed among other *Prorocentrum* species. Bayesian population structure analysis (BAPS) based on the ITS, LSU and SSU loci revealed 4, 5 and 3 clusters, respectively. These clusters corresponded to the main PLSC clades and subclades observed in the phylogenetic reconstructions. Considering the phylogeographic analyses, 18 haplotypes were found for the ITS marker, 21 for the LSU and five for the SSU. According to the haplotypic distribution maps, the haplotypes referring to clades 1 and 3 (ITS and LSU) and clade 1+2 and 3 (SSU) showed a wide distribution, occurring in tropical and subtropical regions of the Atlantic, Pacific and Indian Oceans. The haplotypes corresponding to clade 2 of both ITS and LSU reconstructions occurred exclusively in the Atlantic Ocean. The haplotypes corresponding to clade 4 were observed in temperate regions of the North Atlantic Ocean, Mediterranean Sea and the South Pacific Ocean. We concluded that the PLSC is composed of at least three distinct genetic lineages, with partially overlapping biogeographic distribution, which may correspond to different species.

KEYWORDS: Prorocentraceae, Epibenthic dinoflagellate, okadaic acid, dinophysistoxins, rDNA, Phylogeny .

1. Introduction

The benthic dinoflagellate *Prorocentrum lima* (Ehrenberg) F. Stein, 1878 was described from the Gulf of Naples in the Mediterranean Sea and has a cosmopolitan distribution, being found from tropical to temperate coastal waters (Nagahama et al., 2011). In the tropics, *P. lima* is found throughout the year (Nascimento et al., 2016) while in temperate environments the species can be found at relatively high densities during the summer and autumn (Foden et al., 2005; Aligizaki at al., 2009). *Prorocentrum lima* cells produces okadaic acid (OA) and dinophysistoxin-1 (DTX-1) in varying quantities (e.g. Bravo et al., 2001; Hoppenrath et al., 2013; Nishimura et al., 2020). These toxins can accumulate in marine fauna and cause diarrheic shellfish poisoning (DSP) in human consumers of contaminated shellfish (Murata et al., 1982).

The identification of *Prorocentrum* species is based on morphological features, like cell shape and size; the ornamentation of the theca surface, the intercalary band and architectural details of the periflagellar area (Hoppenrath et al., 2013). However, some of these characteristics can be variable or the difference among species may be subtle (Hoppenrath et al., 2013). Thus, genetic analyzes are essential to assist in species identification. Ribosomal DNA (rDNA) gene sequences have been widely used to better understand the phylogenetic relationships between *Prorocentrum* species (Nagahama et al., 2011; Zhang et al., 2015; Nascimento et al., 2017; Chomerat et al., 2018; Nishimura et al., 2020; Cembella et al., 2021).

Prorocentrum lima presents a cell shape that varies from ovoid, ellipsoid or round (Aligazaki et al., 2009; Nagahama et al., 2011; Nascimento et al., 2017). Due to the high variability found in cell size and shape and shape of pores present at the theca surface Aligizaki et al. (2009) proposed the use of the term "*P. lima* species complex" (PLSC). After that, Nagahama et al. (2011) suggested that the species *Prorocentrum arenarium* M. A. Faust 1994, that presents a wider cell shape than the typical oval shape of *P. lima* cells, should be considered a morphological variation of *P. lima* and synonymized the two species. His assumption was based on SSU phylogenetic analysis which indicated that the *P. arenarium* clade was in between the *P. lima* species complex clades.

In recent years, several phylogenetic reconstructions based on ribosomal DNA sequences have revealed that the *P. lima* species complex is divided in four main clades based on ITS and/or LSU sequences (Nagahama et al., 2011; Nascimento et al. al., 2016;

Nascimento et al., 2017; Nishimura et al., 2020). Nishimura et al. (2020) analyzed strains of the *P. lima* complex from Japan and considered, based on LSU sequences, that one of the four clades showed genetic divergence at the species level while the other three clades did not show clear divergence and that further studies were needed. In the present work, the term PLSC will be used to refer to the four main clades of the *P. lima* species complex, in accordance with Nishimura et al. (2020).

The present work aims to revisit the relationships between PLSC clades using an approach that integrates phylogenetic and phylogeographic analysis using sequences from all over the world and including new strains isolated from the South Atlantic Ocean.

2. Material and methods

2.1. Isolation of strains and establishing cultures

Twenty-three new strains of the *Prorocentrum lima* species complex were established from macroalgae samples collected from five locations at the Brazilian coast and two oceanic islands from the South Atlantic Ocean (Fig. 1). Cultures were established as described in Nascimento et al. (2020). Two strains isolated from Spain (VGO776 and PL27V) were kindly provided by the Spanish Institute of Oceanography (IEO) and were kept in culture. The geographic origin of each strain is presented in Supplementary Table S1. All stock cultures were maintained in a temperature-controlled cabinet at 24 ± 2°C, with a 12 h light: 12 h dark cycle and a photon flux density of 60 µmol m^{-2} s^{-1} provided by cool-white, fluorescent tubes.

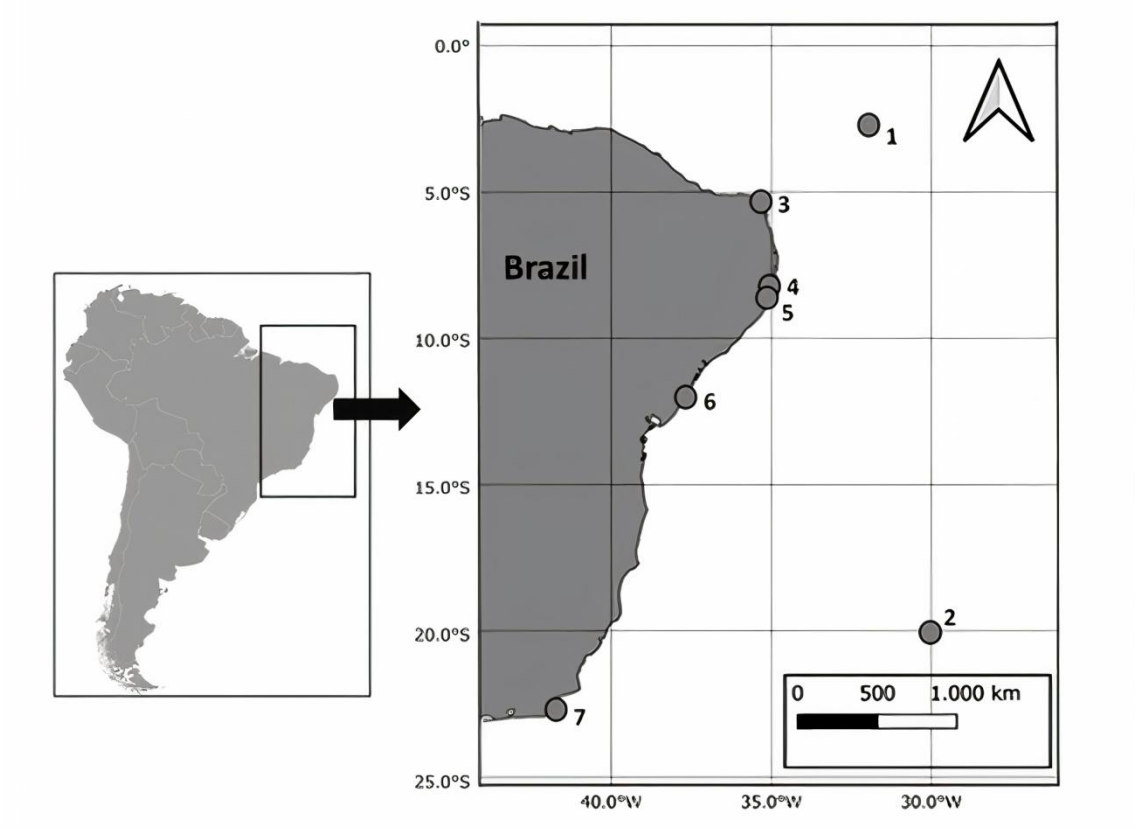

Figure 1. Map of South America showing the seven locations where PLSC strains were isolated from Brazil. 1- Fernando de Noronha Archipelago; 2- Trindade and Martim Vaz Archipelago; 3- Barra do Cunhaú; 4- Praia dos Carneiros; 5-Maragogi; 6- Praia do Forte; 7- Armação dos Búzios.

2.2. DNA extraction, PCR and Sequencing

Exponentially growing cells of PLSC strains were harvested in 1.5 mL microtubes by centrifugation (Eppendorf 5424R centrifuge) at 5000 g for 15 min to settle the cells into pellets. The supernatant was discarded, and the cell pellets were stored at -80°C for further analysis. Genomic DNA was extracted from the pellets using Nucleo Spin Plant II kit (Machery-Nagel, Germany) following the manufacturer's instructions and then stored at -20°C

Three ribosomal DNA (rDNA) loci were analyzed: the Internal Transcribed Spacer (ITS = ITS1-5.8S-ITS2), the D1-D3 region of the large subunit (LSU), and a partial sequence (~900nt) of the small subunit (SSU). Initially several primer pairs available in the literature were tested, showing low amplification and sequencing efficiency: ITSA x ITSB (Sato et al., 2011), ITS1 x ITS4 (Scholin et al., 1994), D1R x D3C (Scholin et al., 1994), D1RxLSUB (Scholin et al., 1994; Litaker et al., 2003), D1R x D3B (Scholin et al., 1994;

Nunn et al., 1996), 16S1N x 16S2N (Grzebyk et al., 1998) and 18ScomF1 x 18ScomR1 (Zhang et al., 2005). Thus, new pairs of primers were designed based on ITS, LSU and SSU sequences of the PLSC available from GenBank using the standard parameters of the Geneious Prime v2020.1 software (www.geneious.com). Forward and reverse primers were designed within conserved regions of the multiple sequence alignment of each rDNA locus. The ITS, LSU and SSU regions were amplified and sequenced using the primer pairs presented in table 1.

Table1. Primers used to amplify and sequence the ITS, LSU and SSU loci of the PLSC strains.

Loci	Primer Sequence (5'-3')
ITS	PLSC_ITS_F1
	GTTGATTACGTCCCTGCCCT
	PLSC_ITS_R1
	ATGAAAGCCACCACCACCTT
	PLSC_ITS_F2
	TGCACATCAGGGCACATTAT
	PLSC_ITS_R2
	TTCACTGGCCTAACATCGTG
LSU	PLSC_D1D3_F1
	CAGGATTCCGTGAGCCAACA
	PLSC_D1D3_R1
	AGGGAAACTTCGGAGGGAAC
	PLSC_D1D3_F2
	TCAGTAATGGCGAATGAACG
	PLSC_D1D3_R2
	TCGGAGGGAACCAGCTACTA
SSU	PLSC_SSU_F1
	TGACCTATCAGCTTCCGACG
	PLSC_SSU_R1
	ACTCATTGGGCGCATCAGTG
	PLSC_SSU_F2
	CGCAAATTACCCAATCCTGA
	PLSC_SSU_R2
	GCAGCCCAGAACATCTAAGG

The amplification reaction mixture (25 µL) contained 1 unit (U) Taq DNA polymerase (ThermoScientific Inc., USA), 1x reaction buffer with NH_4SO_4, 2.5 mM $MgCl_2$, 0.8 mg of Bovine Serum Albumin (BSA), 0.8 mM dNTP's (ThermoScientific Inc., USA), 8 pmol of each primer and approximately 10 µg of genomic DNA. The PCR cycling comprised an initial 5 min heating step at 95°C, followed by 40 cycles of 95°C for 1 min, 45°C (for ITS) or 58°C (for LSU) for 1 min, 72°C for 1 min, and a final extension at 72°C

for 5 min. For SSU, PCR cycling comprised an initial 5 min heating step at 95°C, followed by 10 cycles of 95°C for 1 min, touchdown of 60°C-55°C (Δ = -0,5 ºC per cycle) for 1 min, 72°C for 1 min and 30 sec, then, after these 10 initial cycles, were started 35 cycles of 95 °C for 1 min, 55 °C for 1 min, 72 °C for 1 min and 30 sec and a final extension at 72 °C for 5 min. A 5 µL aliquot of each PCR reaction was checked by electrophoresis in a 1% agarose gel stained with GelRed (Biotium Inc., USA). PCR products were purified and sequenced by Macrogen Inc. (Seoul, Korea) in both directions using the same PCR primers through the traditional capillary sequencing method.

2.3. Phylogenetic analyses

Sequence reads were manually checked and edited using the software Chromas together with the software Mega 7 (Kumar et al., 2016). Sequences were BLAST searched against the GenBank database (www.ncbi.nlm.nih.gov/blast) to test for sequence homology with non-target taxa. Individual consensus sequences were aligned using MAFFT v7 with default settings. (Katoh and Standley, 2013). Poorly aligned positions and divergent regions were eliminated using default settings in Gblocks v0.91b (Castresana, 2000). Phylogenetic analyses were performed separately for the ITS, LSU and SSU loci. The sequences obtained in the present study were analyzed together with all PLSC sequences publicly available from Genbank for each of the three markers (total of 153 ITS, 145 LSU and 33 SSU sequences). Sequences from species closely related to PLSC (Chomérat et al., 2018; Nishimura et al., 2020) were included: *Prorocentrum* sp. type 3 (Morphotype 5 of Zhang et al. (2015), *Prorocentrum* sp. type 1, *Prorocentrum* sp. type 2, *P. caipirignum* and *P. hoffmannianum*. Sequences with less than 450bp (ITS), 540 bp (LSU) and 880 bp (SSU) were removed of the analysis. Sequences of *Prorocentrum levis*, *Prorocentrum bimaculatum* and *Prorocentrum consutum* were used as outgroup in the phylogenetic reconstructions based on ITS, LSU and SSU loci, respectively. The outgroup were selected based on previous phylogenetic analyses of the genus *Prorocentrum* (eg. Boopathi et al., 2015; Luo et al., 2017; Chomérat et al., 2018; Lim et al., 2019) and BLAST analyses, due to the similarity to the PLSC - *P. caipirignum* - *P. hoffmannianum* clade.

The MEGA7 software was used to select the best fit model of nucleotide substitution (K2+I for ITS, K2+G for LSU and HKY+G for SSU) and to construct Maximum Likelihood (ML) phylogenetic trees with 1000 bootstrap replications (BS). The

phylogenetic relationships were also examined using Bayesian Inference (BI) with MrBayes v3.2 (Ronquist et al., 2012). To sample across nucleotide substitution models, the command ‘‘lset nst = mixed’’ was used before running the analysis. Markov Chain Monte Carlo procedure consisted of two independent trials with four chains each. Each chain was run for 1.000.000 generations and sampled every 100th cycle. Posterior probability (PP) values for the resulting 50% majority rule consensus tree were estimated after discarding the first 25% of trees as burn-in.

2.4. Distance analyses

Distance analyses for each of the markers used (ITS, LSU and SSU) were obtained by calculating p-distance in MEGA7. The alignments were the same used for the phylogenetic analyses (153 sequences and 324 positions for the ITS; 145 sequences and 467 positions for the LSU and 33 sequences and 884 positions for the SSU). The PLSC sequences were separated into three (SSU) and four (ITS and LSU) groups according to the main clades formed in our phylogenetic trees, usually retrieved in the literature (Chomérat et al., 2018; Nishimura et al., 2020; Cembella et al., 2021). Then, the calculation of intraspecies and interspecies distances and the calculation of intraclades and interclades distances of the PLSC were performed.

2.5. Bayesian Analysis of Population Structure (BAPS)

To investigate the presence of genetic clusters within the sequences of the PLSC (135 ITS, 119 LSU and 26 SSU sequences), a Bayesian analysis of population structure was performed in BAPS 6.0 (Corander & Tang, 2007; Corander et al., 2008), which allows the assignment of the sequences of each strain to genetic clusters without defining a priori information about their geographic localities. Sequences were aligned using MAFFT v7 (Katoh and Standley, 2013) and treated using Gblocks v0.91b (Castresana, 2000). Population mixture analysis was performed using the module “clustering with linked loci,” and different values of k were tested, from k = 1 to 12.

2.6. Construction of a genealogical network of haplotypes

For the phylogeographic analyses, the same alignments of the BAPS analyses were used, only with PLSC sequences. The alignments were inserted into the DnaSP software (Librado and Rozas, 2009) to define the haplotypes, sites with gaps/missing were considered and invariable sites were removed. The haplotype network was created in the NETWORK software (Bandelt et al., 1999) using the Median-joining network method. The

data concerning the identified haplotypes and their geographic localities were introduced in Microsoft Excel spreadsheets. The relative frequencies of the haplotypes were calculated by location and the graphics were plotted on a World Map. The collection sites of the analyzed strains (Supplementary Table S2) were obtained from Genbank, literature or culture collections.

3. Results

3.1. Phylogenetic analyses

A total of 49 sequences from 23 PLSC strains from Brazil and Spain were generated in the present study and used in the phylogenetic analyses (ITS = 22, LSU = 16, SSU = 11 sequences). These were compared with sequences of the PLSC and related species available from GenBank and the outgroup (ITS = 132, LSU = 130, SSU = 23 sequences) using Maximum Likelihood and Bayesian Inference. The final alignments used for phylogenetic inferences included 154 sequences and 324 positions for ITS; 146 sequences and 467 positions for LSU and 34 sequences and 884 positions for SSU. The topology of the BI and ML trees based on ITS (Fig. 2), LSU (Fig. 3) and SSU (Fig. 4) loci agreed in general with previous studies encompassing the PLSC (Nagahama et al., 2011; Nascimento et al., 2016; Chomérat et al., 2018; Moreira-González et al., 2018; Nishimura et al., 2020; Cembella et al., 2021).

The phylogenetic analyses revealed four (ITS and LSU) and three (SSU) main clades for the PLSC, named according to Nishimura et al. (2020). The phylogenetic trees based on the ITS (Fig. 2) and LSU (Fig.3) loci presented moderate-high support values for the four PLSC clades (ITS: BS = 69-100 / PP = 0.58-1.0; LSU: BS = 68-99 / PP = 0.72-1.0). The ITS and LSU trees were congruent for the PLSC, presenting the closest clades 1 and 2, followed by clade 3 and 4, respectively (Fig. 2 and 3). The topology of the phylogenetic tree based on SSU sequences (Fig. 4) did not recover the same four PLSC clades as ITS and LSU analyses (Fig. 2 and 3). Clades 3 and 4 observed in the ITS and LSU trees (Fig. 2 and 3) were also recovered when SSU sequences were considered (Fig. 4). However, clades 1 and 2 observed in the ITS and LSU trees (Fig. 2 and 3) were not recovered separately, being called clade 1+2 in the SSU tree (Fig. 4). Moreover, relationships between the PLSC clades were more weakly supported in the SSU analyses (BS = 34-98 / PP = 0.66-1.0) (Fig. 4).

The phylogenetic reconstructions based on ITS (Fig. 2), LSU (Fig. 3) and SSU (Fig. 4) rDNA loci grouped the 36 sequences from the South Atlantic Ocean (Brazil) in three of the four clades of the PLSC (clades 1, 2 and 3); one isolate from Tenerife, Spain (VGO776) grouped in clade 3 and one isolate from Ria de Vigo, Spain (PL27V) grouped in clade 4 (Fig. 2, 3 and 4). Based on ITS phylogenetic analyses (Fig. 2) 12 isolates from Brazil grouped in clade 1, with strains originating from several regions of the Atlantic and Pacific Oceans. Nine strains grouped in clade 2, which includes strains from the South Atlantic Ocean, the Gulf of Mexico and the Caribbean and Sargasso Sea. One of the strains was grouped in clade 3, with strains from the Atlantic and Pacific Oceans. The same pattern was found for LSU (Fig. 3).Nine isolates from Brazil grouped in clade 1, four strains grouped in clade 2, on strain from Brazil and one strain from Tenerife, Spain grouped in clade 3 and one strain from Vigo, Spain, grouped in clade 4. Based on SSU (Fig. 4), only three clades were formed and clades 1 and 2, found in the phylogenetic analyzes based on ITS and LSU, formed a unique clade in the analyses based on SSU sequences. Nine isolates from Brazil grouped in clade 1+2, and two strains grouped in clade 3.

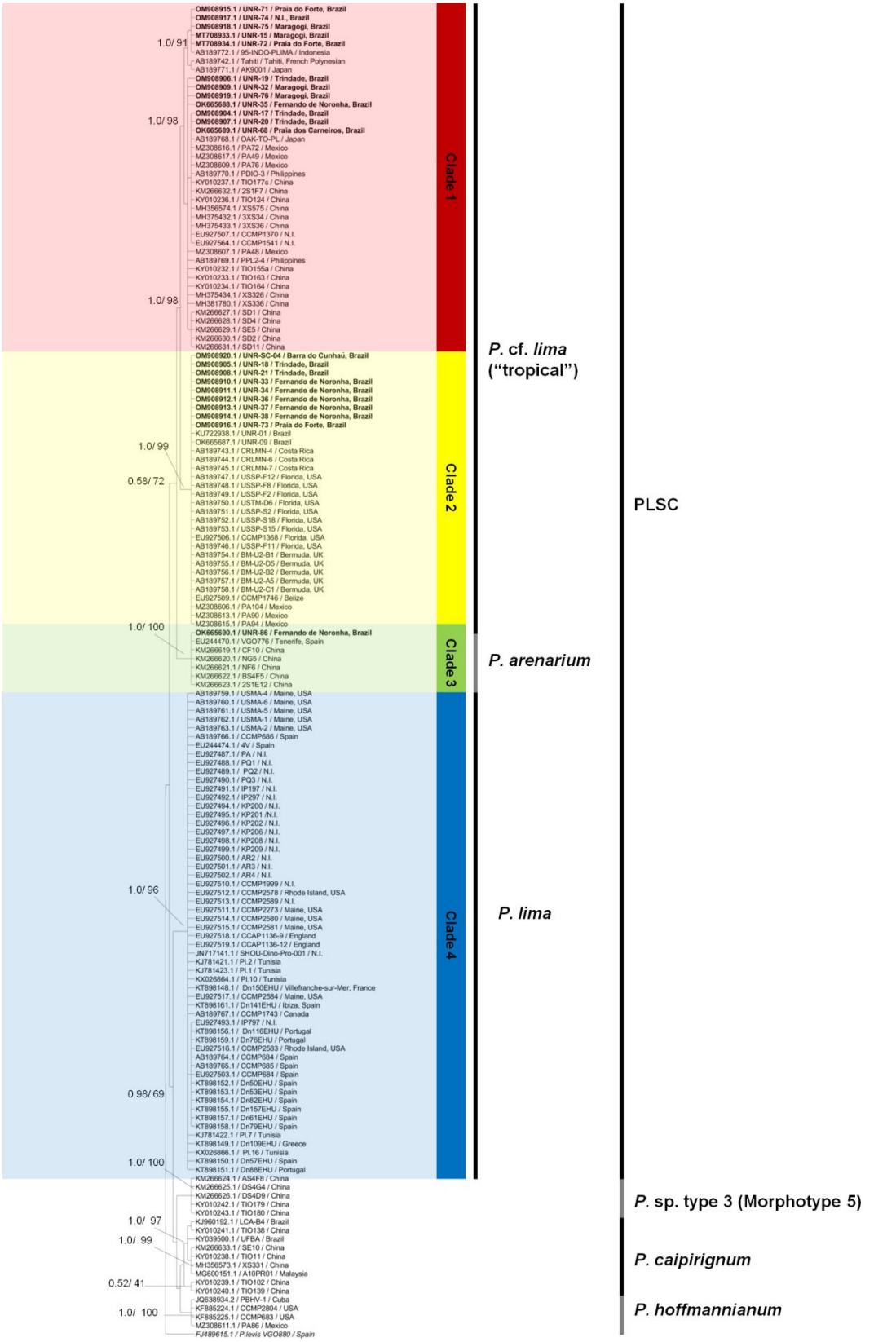

Fig. 2. Bayesian Inference phylogenetic tree based on ITS sequences. Operational taxonomic units (OTUs) are identified by: GenBank accession number/strain code/geographic origin. Numbers at nodes are posterior probability values from BI and bootstrap values from ML analysis respectively (cut-off = 50% for both analyses). The

colored vertical bars represent the four PLSC clades according to Nishimura et al. (2020). N.I.:Not informed. New sequences published in this study are displayed in bold.

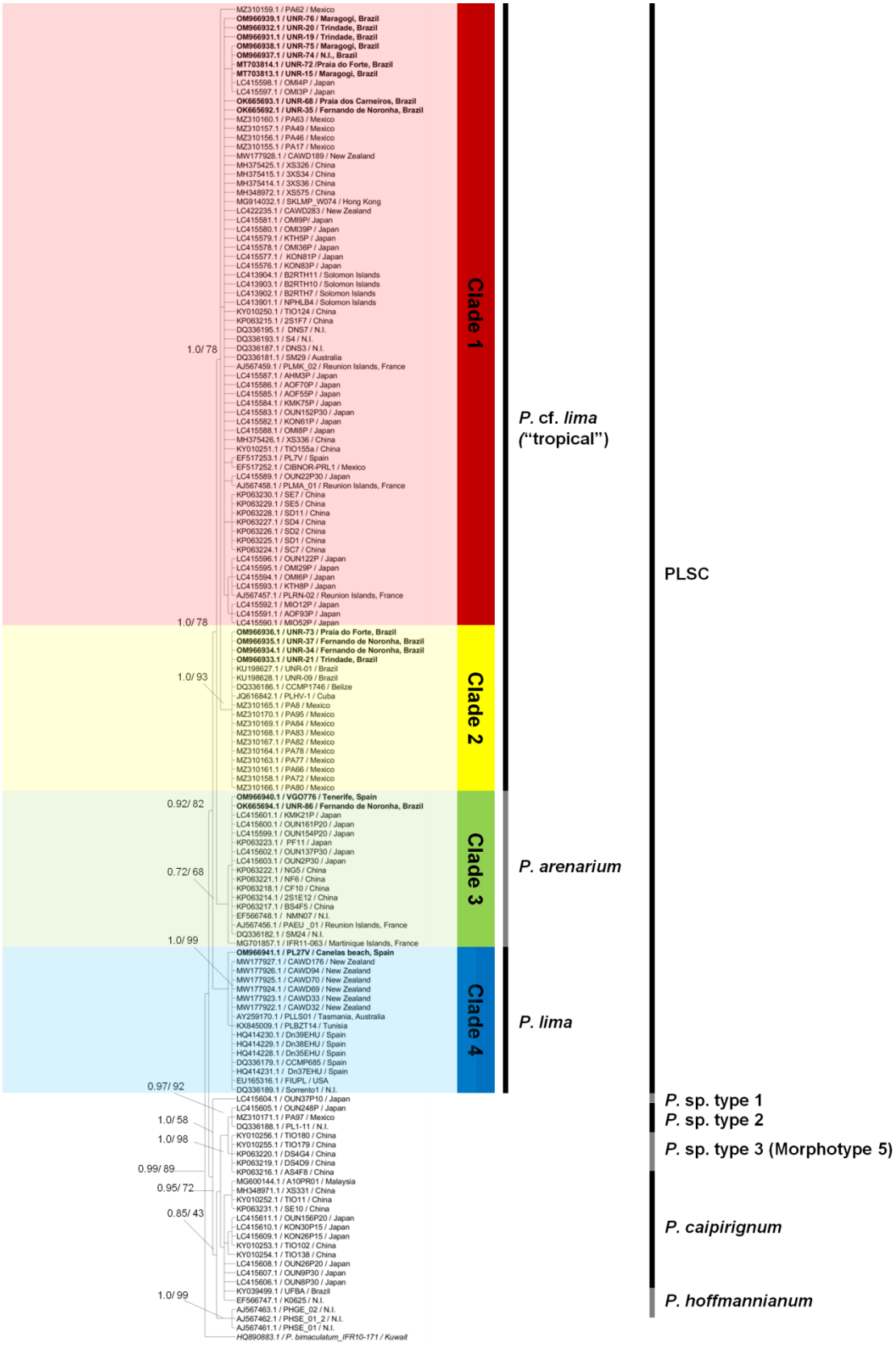

Fig. 3. Bayesian Inference phylogenetic tree based on LSU sequences. Operational taxonomic units (OTUs) are identified by: GenBank accession number/strain code/ geographyc origin. Numbers at nodes are posterior probability values from BI and bootstrap values from ML analysis respectively (cut-off = 50% for both analyses). The colored vertical bars represent the four PLSC clades according to Nishimura et al. (2020). N.I.-Not informed. New sequences published in this study are displayed in bold.

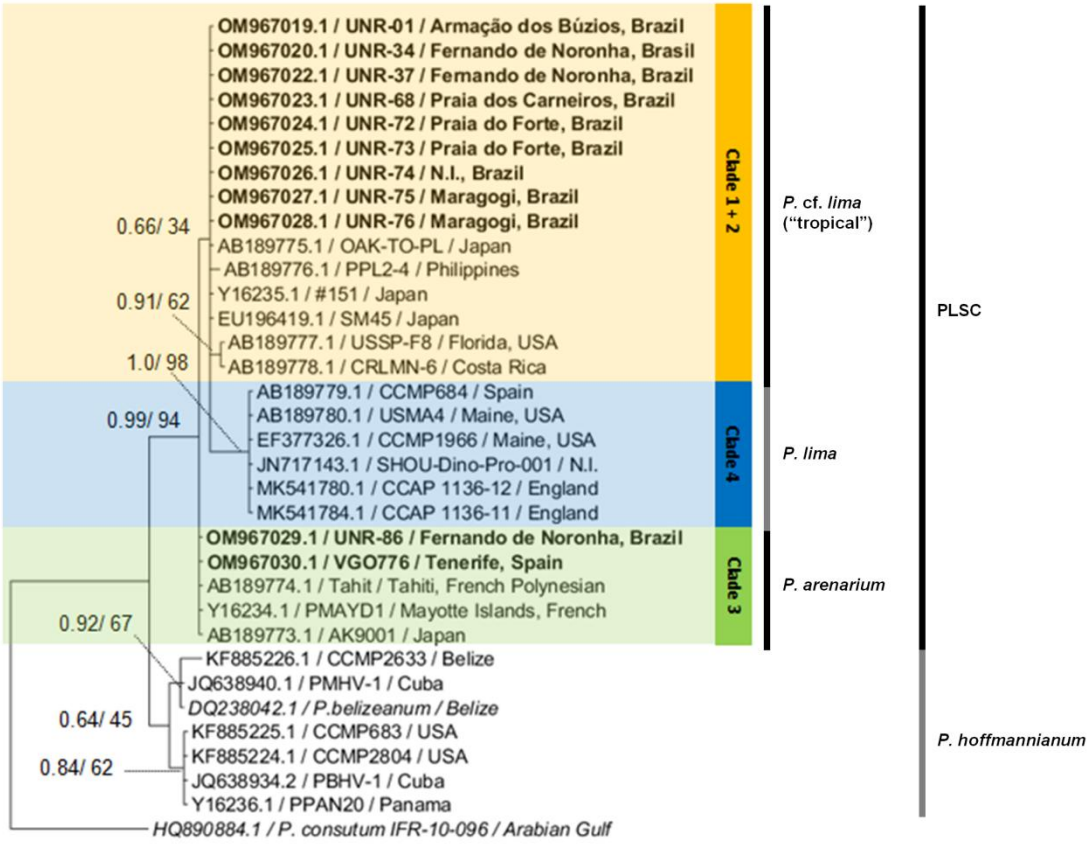

Fig. 4. Bayesian Inference phylogenetic tree based on SSU sequences. Operational taxonomic units (OTUs) are identified by: GenBank accession number/strain code/geographic origin. Numbers at nodes are posterior probability values from BI and bootstrap values from ML analysis respectively (cut-off = 50% for both analyses). The colored vertical bars represent the four PLSC clades according to Nishimura et al. (2020). N.I.-Not informed. New sequences published in this study are displayed in bold.

3.2. Distance analyses

The p-distance within clades in the PLSC complex ranged from 0 to 0.0085 for ITS sequences (Table 2), from 0 to 0.0030 for LSU sequences (Table 3) and from 0 to 0.0004 for SSU sequences (Table 4). Considering the ITS and LSU loci, the p-distance was lower between PLSC clades 1 x 2 (ITS=0.068; LSU=0.0012) and higher between clades 2 x 4 (ITS=0.148; LSU=0.048). The SSU marker showed lower values between PLSC clades

1+2 x 3 (0.001) and higher values between clades 1+2 x 4 (0.007) and 3 x 4 (0.008) (Table 4).

Table 2. Average p-distances between PLSC clades and *Prorocentrum* species based on ITS sequences (319 positions). N° = Number of sequences analyzed. The bold diagonal line represents the intraclade and intraspecies p-distance.

Clades	N°	[1]	[2]	[3]	[4]	[5]	[6]	[7]
[1] *P.* cf. *lima* ("tropical")-Clade 1	40	**0.0085**						
[2] *P.* cf. *lima* ("tropical")-Clade 2	32	0.068	**0.0016**					
[3] *P. arenarium*-Clade 3	7	0.105	0.102	**0**				
[4] *P. lima*-Clade 4	56	0.142	0.148	0.123	**0.0005**			
[5] *P. hoffmannianum*	4	0.153	0.159	0.126	0.121	**0**		
[6] *P. caipirignum*	9	0.145	0.147	0.127	0.130	0.066	**0.0109**	
[7] *P.* sp. type 3	5	0.135	0.138	0.109	0.109	0.056	0.071	**0**

Table 3. Average p-distances between PLSC clades and *Prorocentrum* species based on LSU sequences (452 positions). NA- not available; N° = Number of sequences analyzed. The bold diagonal line represents the intraclade and intraspecies p-distance.

Clades	N°	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]
[1] *P.* cf. *lima (*"tropical")-Clade 1	67	**0.003**								
[2] *P.* cf. *lima* ("tropical")-Clade 2	19	0.012	**0.004**							
[3] *P. arenarium*-Clade 3	17	0.020	0.025	**0.0009**						
[4] *P. lima*-Clade 4	16	0.042	0.048	0.035	**0.0002**					
[5] *P. caiprignum*	14	0.050	0.053	0.045	0.067	**0.0030**				
[6] *P.* sp. type 3	5	0.054	0.058	0.049	0.071	0.013	**0**			
[7] *P.* sp. type 2	3	0.046	0.049	0.040	0.062	0.009	0.009	**0**		
[8] *P.* sp. type 1	1	0.046	0.045	0.040	0.062	0.013	0.018	0.013	**NA**	
[9] *P. hoffmannianum*	3	0.055	0.058	0.049	0.070	0.027	0.031	0.027	0.027	**0.0044**

Regarding the genetic divergence between PLSC and related species, the p-distance between some of the PLSC clades is similar or even greater than the distance observed between other valid species, for example, in analyses based on ITS sequences between *P. caipirignum* x *P. hoffmannianum* (0.066). The p-distance between *P. caipirignum, P. hoffmannianum* and *Prorocentrum* sp. type 3 (*P. lima* morphotype 5 according to Zhang et al. 2015) ranged between 0.056 to 0.071 for ITS and from 0.013 to 0.031 for LSU (Table 2 and 3). The p-distance based on the SSU marker was not evaluated for *P. caipirignum* and for *Prorocentrum* sp. type 3, as there are no SSU sequences available for these species on Genbank (Table 4).

Table 4. Average p-distances between PLSC clades and *Prorocentrum* species based on SSU sequences (884 positions). N° = Number of sequences analyzed. The bold diagonal line represents the intraclade and intraspecies p-distance.

Clades	N°	[1]	[2]	[3]	[4]
[1] *P.* cf. *lima (*"tropical")-Clades 1+2	15	**0.0004**			
[2] *P. arenarium*-Clade 3	5	0.001	**0**		
[3] *P. lima*-Clade 4	6	0.007	0.008	**0**	
[4] *P. hoffmannianum*	7	0.013	0.012	0.015	**0.002**

3.3. Bayesian analysis of population structure (BAPS)

Bayesian analysis of the population structure based on the ITS sequences indicated the existence of four genetic clusters (Fig. 5A). Cluster A was formed exclusively by sequences from temperate regions of the North Atlantic Ocean, equivalent to clade 4 in phylogenetic analyzes (Fig. 2); cluster B included sequences from tropical and subtropical regions of the Atlantic and Pacific Oceans, equivalent to clade 3; cluster C was formed by sequences from tropical and subtropical regions of Atlantic and Pacific Oceans, equivalent to clade 1 and cluster D was composed solely of lineages from tropical regions of the Atlantic Ocean, equivalent to clade 2.

BAPS analysis based on LSU sequences indicated five genetic clusters (Fig. 5B). Cluster A and cluster B were equivalent to the clusters A and B found at BAPS analyses based on the ITS, and cluster E were equivalent to clade D. Cluster C and D, equivalent to clade 1 was formed by sequences from tropical and subtropical regions of the Atlantic, Pacific and Indian Oceans.

Considering the SSU sequences, the BAPS analysis identified three genetic clusters (Fig. 5C). Cluster A and B were equivalent to clusters A and B found at BAPS analyses based on ITS and LSU loci. Clusters C was formed by sequences from tropical and subtropical regions of the Atlantic, Pacific and Indian oceans, equivalent to clades 1+2 (Fig. 3).

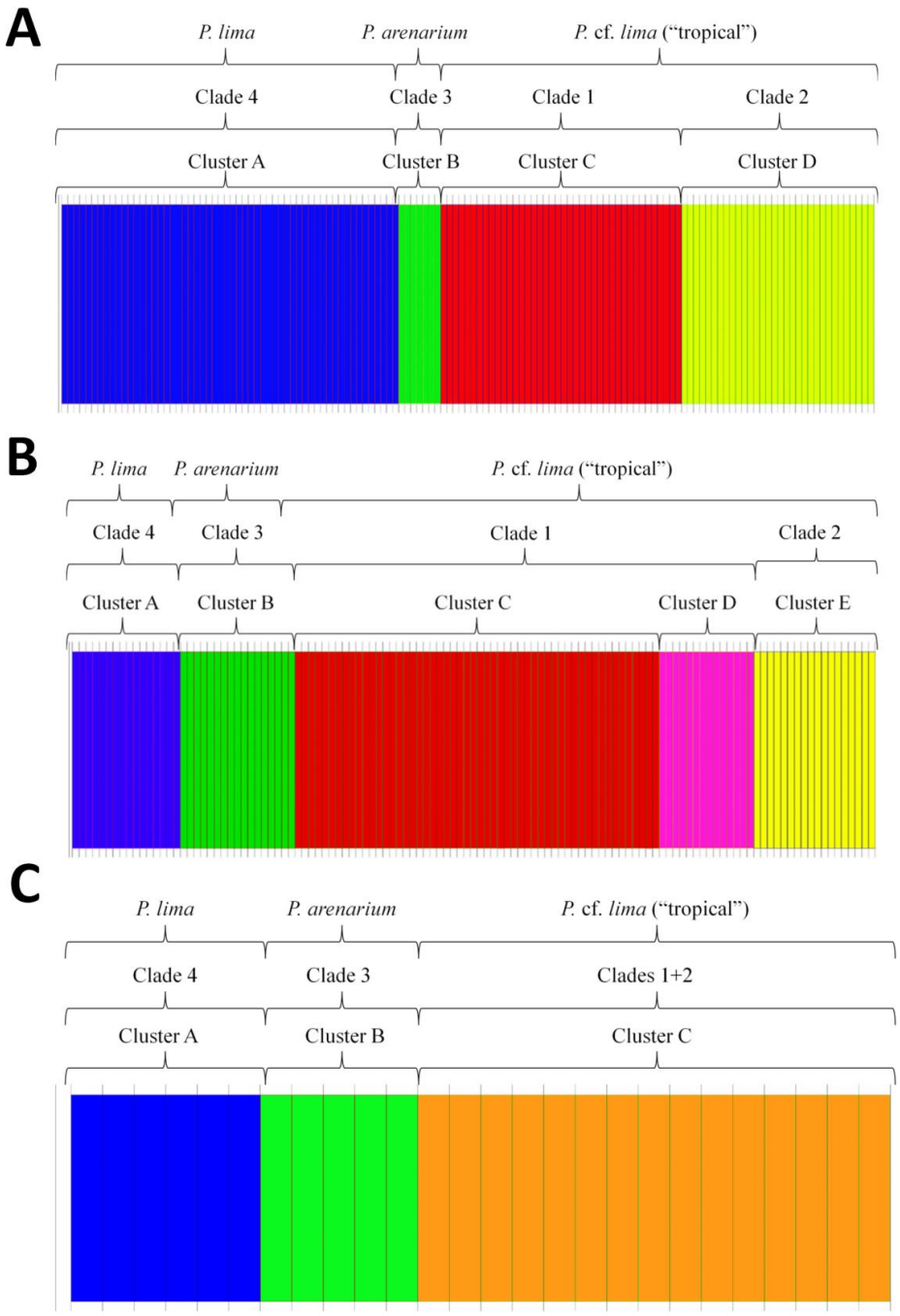

Fig. 5. Bayesian clustering analysis of PLSC rDNA sequences, considering (A) ITS sequences; (B) LSU sequences and (C) SSU sequences. Colors represent the different genetic clusters identified. Each vertical bar represents a sequence.

3.4. Phylogeographic analyses

The analyses of the PLSC rDNA sequences allowed the identification of 18 haplotypes based on the ITS locus (Fig. 6A), 21 haplotypes based on the LSU locus (Fig. 6B) and five haplotypes based on the SSU locus (Fig. 6C). In the PLSC haplotype network

based on ITS locus (Fig. 6A), clade 1 (Cluster C) was formed by seven haplotypes (H1-H7); Clade 2 (Cluster D) was formed by three haplotypes (H8- H11); Clade 3 (Cluster B) was formed by a single haplotype (H12) and Clade 4 (Cluster A) was formed by six haplotypes (H13-H18).

Considering the PLSC haplotype network based on LSU locus (Fig. 6B), clade 1 (Cluster C and D) was formed by ten haplotypes (H1-H10); Clade 2 (Cluster E) was formed by four haplotypes (H11-H14); Clade 3 (Cluster B) included five haplotypes (H15-H19) and clade 4 (Cluster A) was formed by two haplotypes (H20, H21). Regarding the PLSC haplotype network based on SSU locus (Fig. 6C), clade 1 (Cluster C) was formed by two haplotypes (H1 and H2) and clades 2 (Cluster C), 3 (Cluster B) and 4 (cluster A) was formed by one haplotype each, H3, H4 and H5 respectively.

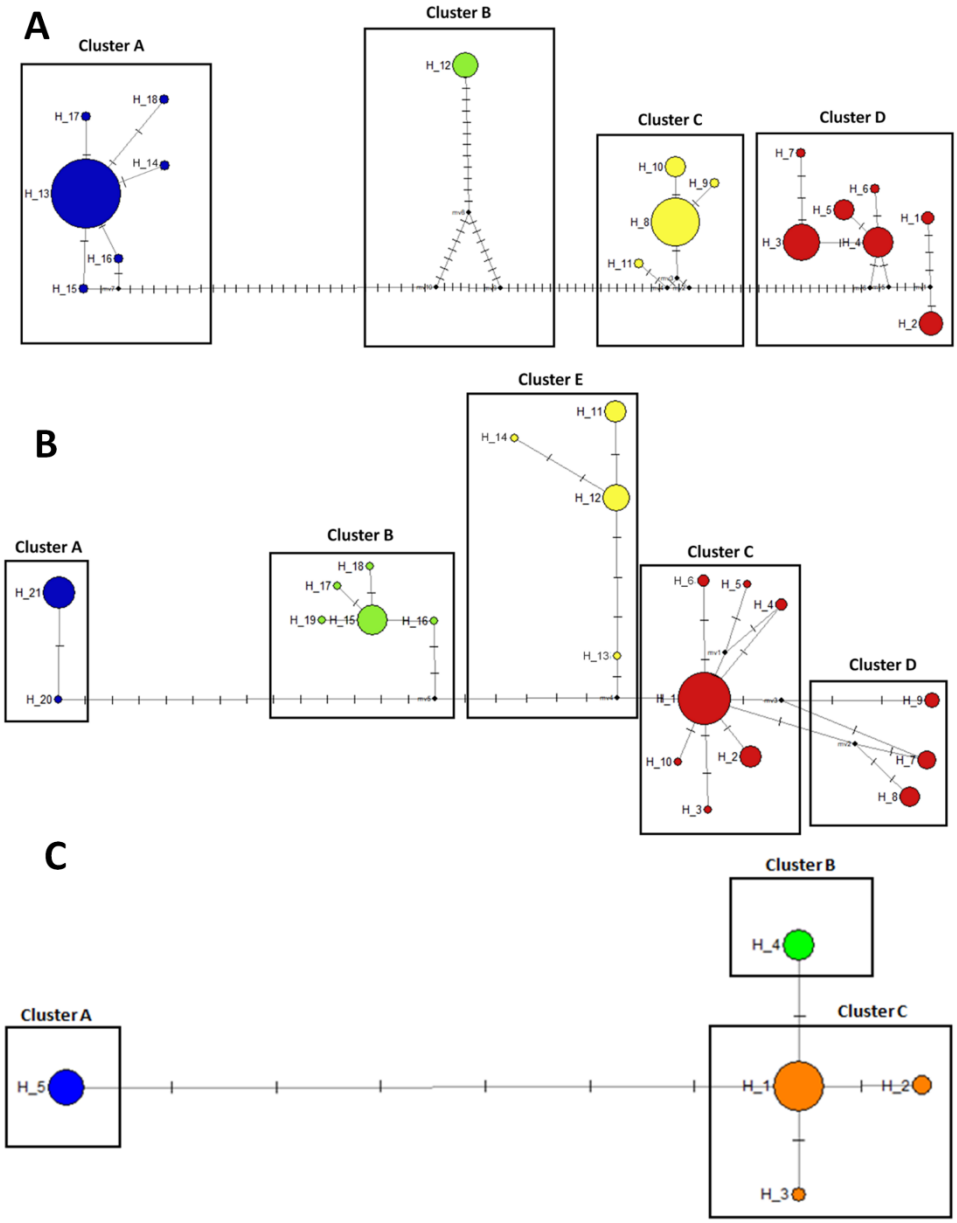

Fig. 6. Haplotype networks based on ITS (A) LSU (B) and SSU (C) PLSC rDNA sequences. The haplotypes were represented using the same colors as the clades in the phylogenetic trees (Red = clade 1; Yellow = clade 2; Green = clade 3; Blue = clade 4). Rectangles identify the main groups found in BAPS analyses. Cross-lines on the branches correspond to mutations in the alignment and the circles represent the haplotypes and the mean vectors. The diameter of each haplotype circle is proportional to the relative haplotype frequency.

According to the maps based on ITS, LSU and SSU sequences (Fig. 7A, B and C respectively), haplotypes from strains in Clade 1 have a wide distribution, occurring in tropical and subtropical regions, between latitudes 42°N and 12°S of the Atlantic Ocean,

between latitudes 40°N and 29°S of the Pacific Ocean and at latitudes 21°S of the Indian Ocean (Reunion Island). The haplotypes corresponding to the strains from Clade 2 occurred exclusively in the tropical Atlantic Ocean between latitudes 32°N and 22°S. Haplotypes corresponding to Clade 3 also had a wide distribution, occurring in tropical and subtropical regions between latitudes 14°N and 3°S in the Atlantic Ocean, between latitudes 33°N and 19°N in the Pacific Ocean (this distribution may be broader, since the strain SM24 were described for tropical Australia, Genbank code DQ336182.1, but without information about its geographic coordinates) and at latitude 22°S in the Indian Ocean. The haplotypes from the strains grouped in Clade 4 are present solely in temperate and subtropical regions between latitudes 33°N and 45°N of the Atlantic Ocean and Mediterranean Sea and between latitudes 34°S and 41°S of the Pacific Ocean (Tasmania and New Zealand).

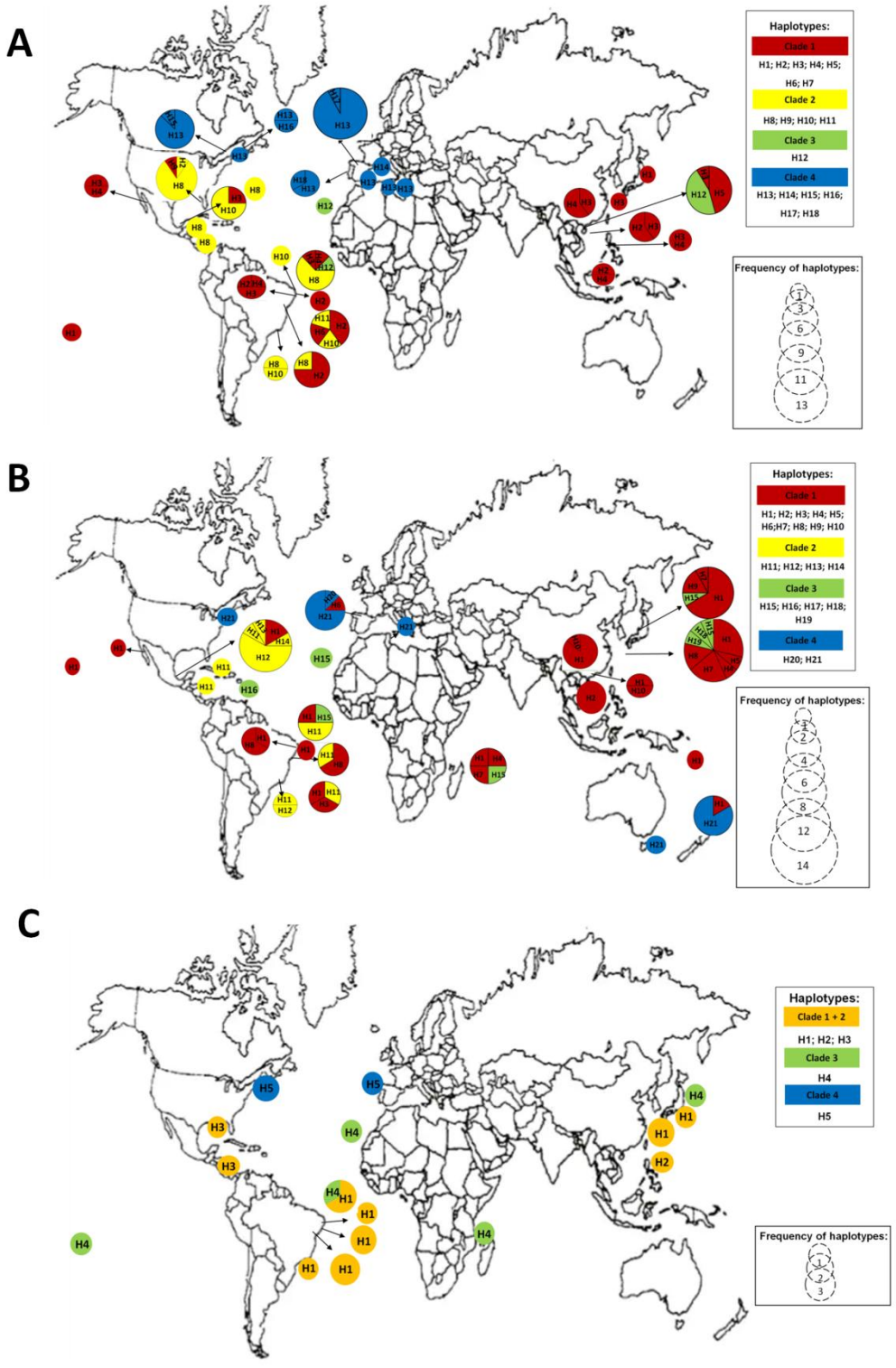

Fig. 7. Haplotypic Diversity Maps based on ITS (A), LSU (B) and SSU (C) sequences of strains encompassing the PLSC.

4. Discussion

Prorocentrum lima was described by Ehrenberg (1860) under the name of *Cryptomonas lima*. The redescription of the species was carried out considering morphological characters of the organism collected in the type locality, Gulf of Naples, Italy, without molecular analysis (Nagahama and Fukuyo., 2005). Subsequently, the use of the term "*P. lima* species complex" (PLSC) was proposed due to the high variability found in the size and shape of cells and in the shape of the pores present on the theca surface (Aligizaki et al., 2009). Then, *P. arenarium* and *P. lima* were synonymized based on phylogenetic analysis of SSU sequences (Nagahama et al., 2011). In recent years, with the expansion and availability of molecular data, many studies have been developed to better understand the limits of PLSC (Zhang et al., 2015; Nascimento et al., 2017, Chomérat et al., 2018; Nishimura et al. al., 2020).

In the present work, the molecular phylogenetic analyses, including a greater number of sequences, recovered tree topologies based on ITS, LSU and SSU loci that are in agreement with previous published phylogenies (Zhang et al., 2015; Nascimento et al., 2016; Moreira-González et al., 2018; Chomérat et al., 2018; Nishimura et al., 2020; Cembella et al., 2021). The phylogenetic analyses based on the ITS and LSU markers (Figs. 2 and 3) are congruent, with the four clades of the PLSC presenting moderate-high posterior probability (BI) and Bootstrap (ML) values that support phylogenetic relationships. The topology of the phylogeny based on SSU rDNA sequences (Fig. 4) was consistent with that of ITS and LSU regions, but presented three, instead of four, PLSC clades. Besides that, the relationships between some of the PLSC clades are unclear as some nodes had low support values.

Molecular markers of rDNA genes have different evolutionary rates, ITS sequences are quite variable, revealing high inter-specific and even intra-specific diversity (Stern et al., 2012), while LSU and SSU sequences are more conserved, but LSU sequences evolve faster than SSU sequences in dinoflagellates, providing better phylogenetic resolution (Boopathi et al., 2015). Within the genus *Prorocentrum*, different genetic distance values were observed at intra and interspecific levels (Nascimento et al., 2017). In the present work, the lowest values of ITS p-distances were observed between PLSC clades 1x2 (0.068), clades herein called as *P.* cf. *lima* “tropical” and the highest distance values were

observed between clade 4 (*P. lima*) and the remaining clades 1x4 (0.142), 2x4 (0.148) and 3x4 (0.123), followed by distance values between clade 3 (*P. arenarium*) and remaining clades 1x3 (0.105), 2x3 (0.102), 3x4 (0.123). These values reveal a high genetic divergence among PLSC clades, which far exceeds the limit proposed by Litaker et al. (2007) of 4% (0.04) p-distance based on the ITS loci to recognize most free-living dinoflagellate species.. Regarding the LSU sequences, the lowest values of genetic distance were also observed between PLSC clades 1x2 (0.012), corresponding to *P.* cf. *lima* "tropical". While the highest values were found between clade 4 (*P. lima*) and the remaining clades, 1x4 (0.042), 2x4 (0.048) and 3x4 (0.035), followed by values between clade 3 (*P. arenarium*) and the other clades 1x3 (0.020), 2x3 (0.025) and 3x4 (0.035). That is in accordance with what was found for ITS, but with lower distance values.

A similar analysis considering the p-distance values of ITS and LSU sequences from species closely related to the PLSC showed values lower than those found between PLSC clades, such as *P. caipirignum* x *P. hoffmannianum* (ITS = 0.066; LSU = 0.027), *P. caipirignum* x *P.* sp. type 3 (ITS = 0.071 LSU = 0.013) and *P. hoffmannianum* x *P.* sp. type 3 (ITS: 0.056; LSU: 0.031). Nascimento et al. (2017) found the following p-distance values between these species: *P. caipirignum* x *P. hoffmannianum*, (ITS = 0.068 – 0.091; LSU = 0.013 – 0.019), *P. caipirignum* x *P.* sp. type 3 (ITS = 0.07 - 0.078; LSU = 0.008 - 0.01). Thus, the genetic differences between the PLSC clades, particularly between clades 4 and 3 and the remaining clades, are higher than those found among other valid *Prorocentrum* species.

Regarding the SSU marker, as expected, the distance values were lower between all clades, compared to ITS and LSU values, ranging from 0.001 to 0.008. The same pattern of highest distance values between *P. lima* (clade 4) and the remaining clades was observed in this more conserved marker. The SSU phylogenetic analyzes presented a lower resolution and retrieved three PLSC clades, as clades 1 and 2 were not differentiated by this locus. The SSU has significantly lower resolution than LSU in resolving *Prorocentrum* species relationships, although less significant than the other genetic markers, the information presented by the SSU is useful to improve species identification (Boopathi et al., 2015).

The divergence between sequences of *Prorocentrum* species, *P. arabianum, P. concavum, P.* cf. *faustiae* (later synonymized with *P. concavum*), *P. lima* and *P. arenarium* (later synonymized with *P. lima*) was estimated, using a smaller dataset (five sequences of

Prorocentrum were analyzed), and revealed a variation of 0.002 - 0.198 based on the LSU locus, with a difference of 0.03 between *P. lima* and *P. arenarium* sequences (Mohammad-Noor et al., 2007). Morphological and molecular analyzes were performed of *P. lima* strains from the South China Sea based on the ITS and LSU loci of rDNA, indicating the occurrence of five morphotypes and observed genetic distance values between Morphotype 2 (clade 1 of the current work), Morphotype 3 (clade 1 of the current work), Morphotype 4 (*P. caipirignum*), Morphotype 5 (*Prorocentrum* sp. type 3) and Morphotype 1 of their work (clade 3 of the present work, *P. arenarium*) from 0 to 0.05862 in LSU locus (Zhang et al., 2015). Nascimento et al. (2017) found distances of 0.07 – 0.09 between ITS sequences of *P. caipirignum* and the two closest species (*P. hoffmannianum* and *Prorocentrum* sp. Type 3). Based on the LSU sequences, it was considered that the average p-distance between *P. lima* (clade 4) and the other clades (1, 2 and 3) presented divergence at the species level (0.034 – 0.047) and were larger than those found between *P. caipirignum* and *P. hoffmannianum* (0.027) (Nishimura et al., 2020). In the present work, using a higher number of sequences from each marker, the p-distance values between *P. lima* (clade 4) and the remaining clades (ITS: 0.0123 – 0.048 and LSU: 0.035 – 0.048) were larger than those observed between *P. caipirignum* and *P. hoffmannianum* (ITS: 0.066 and LSU: 0.027). Comparisons based on the SSU locus were hampered as there was no SSU sequence of *P. caipirignum* available in Genbank.

The results of the BAPS analyses based on ITS showed four genetic clusters. While LSU sequences were clustered in five genetic groups, the sequences corresponding to clade 1 were subdivided in two clusters (C and D) that do not show distinction in the geographic distribution (Fig. 5B). This is probably due to the larger number of sequences and higher genetic diversity of the LSU sequences from clade 1 (Fig. 3). The SSU presented three genetic clusters (A, B and C) corresponding to clades 4, 3 and 1+2 respectively, in agreement with the phylogenetic tree (Fig. 2, 3 and 4).

The PLSC is widely distributed from temperate to tropical regions in latitudes ranging from 45° N to 35° S. This great dispersion capacity may be related to their asexual reproduction by binary fission and to the production of cysts (Faust, 2004). Some studies have shown that passive transport of vegetative and/or resting cysts can occur on floating and drifting objects and plastic debris (Larsson, et al., 2018; Masó et al., 2003); besides the transport of dinoflagellate cysts by ship's ballast water (Hallegraf, 1998).

An interesting pattern observed in the haplotype distribution maps was the high genetic diversity in the populations from oceanic islands (Trindade, Fernando de Noronha, Mayotte, Okinawa and Hainan). These islands presented a higher number of haplotypes relative to coastal areas. The Fernando de Noronha Island in Brazil standouts as the only place presenting haplotypes corresponding to three of the four main clades of the PLSC (ITS: H3, H4, H8, H12; LSU: H1, H11, H15; SSU: H1, H4). In Trindade (ITS: H2, H6, H10, H11; LSU: H1, H3, H11; SSU: H1) Mayotte (LSU: H1, H7, H15), Okinawa (ITS: H3; LSU: H1, H4, H5, H7, H8, H15, H18, H19; SSU: H1) and Hainan Islands (ITS: H3, H5, H12) also presented a high number of haplotypes from two PLSC clades.

The haplotypes of *P. lima* (clade 4) (ITS: H14-H18; LSU: H20 and SSU: H5), presented a more restricted distribution, in the Temperate North Atlantic biogeographic realm in Lusitanian, Mediterranean Sea and Cold Temperate Northwest Atlantic provinces (according to the classification by Spalding et al., 2007). However, the LSU-H21 haplotype was found both in the Temperate North Atlantic and in the Temperate Australia biogeographic realms (Southern New Zealand and Southern Australian Shelf provinces), the latter probably from a recent introduction event. The haplotypes corresponding to clade 2 of *P.* cf. *lima* "tropical", (ITS: H8-H11; LSU: H11-H14 and SSU: H3) also presented a more restricted distribution and were present only in Atlantic Ocean, in Temperate Northern Atlantic (Warm Temperate Northwest Atlantic province) in the Tropical Atlantic (Tropical Northwestern Atlantic and Tropical Southwestern Atlantic), which may suggest an ongoing speciation event.

Haplotypes corresponding to clade 1 of *P.* cf. *lima* "tropical" (ITS: H1-H7; LSU: H1-H10; SSU: H1-H2) and to clade 3 (*P. arenarium*) (ITS: H12; LSU: H15-H19; SSU: H4) had a broad distribution and were mostly present in tropical and subtropical regions. Haplotipes corresponding to clade 1 were present in the Temperate Northern Atlantic (in Warm Temperate Northwest Atlantic and Lusitanian provinces) Tropical Atlantic (in Tropical Northwestern Atlantic and Tropical Southwestern Atlantic provinces), Western Indo-Pacific (Western Indian Ocean province), Temperate Northern Pacific (Cold Temperate Northwest Pacific, Warm Temperate Northwest Pacific and Warm Temperate Northeast Pacific provinces), Central Indo-Pacific (South China Sea, Western Coral Triangle and Eastern Coral Triangle provinces) and Temperate Australasia (Southern New Zealand provice) biogeographic realms of Spalding et al. (2007) (Fig. 7). Considering Haplotipes corresponding to clade 3 were present in the Tropical Atlantic (in Tropical

Northwestern Atlantic, Tropical Southwestern Atlantic and West African Transition provinces), Western Indo-Pacific (Western Indian Ocean province), Central Indo-Pacific (South China Sea and Tropical Northwestern Pacific provinces) and Eastern Indo-Pacific (Easter Island province) biogeographic realms of Spalding et al. (2007) (Fig. 7).

The present work revealed that PLSC populations comprise three evolutionary lineages: *P. lima* (clade 4), *P. arenarium* (clade 3) and *P.* cf. *lima* “tropical" (clade 1 and 2). Nagahama et al. (2011) had already suggested that allopatric speciation may be occurring and that geographically separated *P. lima* populations may have become genetically distinct. It was suggested that at least two evolutionary lineages may be inferred from their results, one from temperate climate regions, including the type locality of *P. lima* (Mediterranean Sea) that forms a monophyletic clade (equivalent to clade 4 of the present work) and a second one including strains from tropical and subtropical regions that are grouped in a monophyletic clade with two subclades (equivalent to clade 1, 2 and 3 of the present work) (Nascimento et al., 2017). The current work included a greater number of sequences and analyzed SSU, besides ITS and LSU loci, in addition to having performed phylogeographic analysis.

Thus, based on the analyzes performed here, clades 3 and 4 presents clear distinction at species level and may be considered respectively *P. arenarium* and *P. lima*. Clades 1 and 2 are clearly different from clade 3 and 4, but do not present a clear distinction at the species level between then. However our data suggest that there is a biogeographic pattern within this group, with haplotypes corresponding to clade 2 restricted to the Atlantic Ocean, while the haplotypes corresponding to clade 1 are present in the Atlantic and Pacific Oceans (Fig. 7). An investigation of the ITS secondary structure of clade 1 and 2 sequences found that they have distinct structural patterns (Cembella et al., 2021). While sequences of clade 1 have a distinct CBC (compensatory base changes) in helix 3, clade 2 sequences have two CBCs in helix 2 and one CBC in helix 3, and the authors suggested that this distinct pattern may be indicative of early stage of speciation (Cembella et al., 2021).

A wide range of cell shape variability has been reported for the PLSC, with cells almost round to oblong oval and ovoid (Hoppenrath et al., 2013). However, this high variability may be explained by the inclusion of *P. arenarium*, larger and wider, in the morphological range of *P. lima*. Zhang et al. (2015) described three PLSC morphotypes in the southern China Sea, that were differentiated by cell size, shape and surface

ornamentation (Morphotypes 1, 2 and 3) and that were genetically separated. Morphotypes 2 and 3 corresponded to clade 1 of the present work and Morphotype 1 to clade 3. Several studies have already pointed out to the wider cell shape of *P. arenarium* cells (Faust, 1994), that is a stable character in cells from clade 3. In contrast, cells from clades 1, 2 and 4 are oval or oblong in cell shape (Nascimento et al., 2017; Zhang et al., 2015; Chomérat et al., 2018; Lim et al., 2019; Silva et al., 2020).

Prorocentrum arenarium was described by Faust (1994) based on the morphological characteristics of wild cells obtained from Carrie Bow Cay, Belize (Tropical Atlantic Ocean). Subsequently, the SSU rDNA of a strain of *Prorocentrum* identified as *P. arenarium* (strain: PMAYD1; Genbank code: Y16234.1), isolated from the Island of Mayotte (tropical Indian Ocean, French territory) was sequenced and based on molecular analysis, it was suggested that *P. arenarium* would be a larger and rounder morphotype of *P. lima* (Grzebyk et al., 1998). This suggestion was based on the SSU rDNA phylogenetic analysis that showed the *P. arenarium* sequence in between the *P. lima* clades. Then, the broad-oval cell shape of *P. arenarium* was considered to be within the range of morphological variability of *P. lima* and it was suggested that *P. lima* formed a single monophyletic clade showing two subclades (Nagahama et al., 2011). So, *P. arenarium* was considered synonymous with *P. lima* (Nagahama et al., 2011). These authors used a restricted dataset, available at that time, with only 12 SSU sequences, including only one sequence considered to be of *P. arenarium* (strain: PMAYD1).

Recent studies showed that the cell shape typical of *P. arenarium* is statistically wider than that of specimens from the remaining clades (as morphotype 1 in Zhang et al., 2015, Silva et al., 2020), and it may be recognized as a stable character of all examined specimens from clade 3 (*P. arenarium*). In the present work and in published phylogenies (Zhang et al., 2015; Nascimento et al., 2016; Nascimento et al., 2017; Chomérat et al., 2018; Nishimura et al., 2020), the strains with the typical morphology of *P. arenarium* formed a well-defined clade (clade 3) in ITS, LSU and SSU phylogenies and the proposal of the synonymy of *P. lima* and *P. arenarium* is not accepted by all authors (Nascimento et al., 2017).

Unfortunately, there are no sequences available for *P. arenarium* from the type locality (Belize), and therefore the identity of this species needs to be confirmed by the examination of material from the type locality using both morphological and molecular

analysis. A material was examined from the island of Martinique, that is located 2,947 km away from Belize, and was reported, as *P. lima*, the morphology and genetic data of a cell with the morphology similar to *P. arenarium* (Fig. 3a, b in Chomerat et al., 2018) and that groups in clade 3, of *P. arenarium* (strain: IFR11-063 Genbank code: MG701857) (Chomerat et al., 2018).

Regarding *P. lima* (clade 4) and *P.* cf. *lima* "tropical" (clades 1 and 2) the strains corresponding to these groups present very similar morphology. Significant morphological differences was found regarding the central width and the ratio between the cell central width and the cell upper width, suggesting that there is morphological variation between strains from these two clades (Silva et al., 2020). However, the authors recommended that more morphometric analyzes including more strains from the different PLSC clades are performed.

Previous studies revealed that all evaluated strains from the four PLSC clades produced OA and several strains produced DTX1 (eg. Bravo et al., 2001; Rhodes et al., 2006; Hoppenrath et al., 2013; Nascimento et al., 2016; Luo et al., 2017; Moreira-González et al., 2018; Nishimura et al., 2020; Cembela et al., 2021). Regarding the production of toxins by clades of PLSC, clade 1 (*P.* cf. *lima* "tropical") has high variation in OA (0.01 pg/cell - 55.27 pg/cell) and DTX1 (not detected - 46.46 pg/cell) production (Nishimura et al., 2020). Clade 3 (*P. arenarium*) showed high variation of OA (1.19 pg/cell - 51.17 pg/cell) and no or low production of DTX1 (not detected - 0.74 pg/cell). Two strains from clade 2 (*P.* cf. *lima* "tropical") produced OA (15.2 pg/cell; 3.63 mol.cell-1) with low amount of DTX1 (0.47 pg/cell; 2.81 mol.cll-1) (Nascimento et al., 2016; Cembella et al., 2021) and clade 4 (*P. lima*) produced OA (2.40 pg/cell - 28.33 pg/cell) with lower amounts of DTX1(3.02 pg/cell – 11.62 pg/cell) (Bravo et al., 2001; Ben-Gharbia et al., 2016). Within each clade there is great variability in the amount of OA produced, but all strains of PLSC produce OA in greater amounts and little or no DTX-1 (Nishimura et al., 2020). Further analyses may elucidate whether there are differences in the toxin profile and toxin quota between the different lineages of the PLSC.

5. Conclusions

The phylogenetic and phylogeographic analyses performed in the present study revealed that the PLSC is composed of at least three distinct genetic lineages, with partially

overlapping biogeographic distribution, which may correspond to three different species. *Prorocentrum lima* (clade 4), *P. arenarium* (clade 3) and *P.* cf. *lima* "tropical" (clade 1 and 2) sequences presented high p-distance values in relation to each other, comparable to those existing among some other *Prorocentrum* species, specially for the ITS and LSU loci. Furthermore, BAPS analyses confirmed that the PLSC encompasses at least three distinct genetic clusters, corresponding to *P. lima, P. arenarium* and *P.* cf. *lima* "tropical" clades. In addition, the phylogeographic analyses revealed a restricted distribution of haplotypes belonging to *P. lima* (clade 4) and *P.* cf. *lima* "tropical" (clade 2). *Prorocentrum lima* (clade 4) haplotypes are restricted to temperate regions and there is virtually no geographic overlap between this clade and the other PLSC clades. The *P.* cf. *lima* "tropical" (clade 2) haplotypes have a distribution restricted to the tropical regions of the Atlantic Ocean, which may indicate an early stage of speciation. On the other hand, haplotypes corresponding to *P. arenarium* (clade 3) and *P.* cf. *lima* "tropical" (clade 1) presented a wide biogeographic distribution.

Acknowledgements

We acknowledge the Spanish Institute of Oceanography (IEO) for providing strainsVGO776 and PL27V from the culture collection of the Centro Oceanografico de Vigo, Spain. Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) provided financial support for the research project PELD ILOC/CNPQ 441241/2016-6 that allowed us to isolate strains from the oceanic islands. Fundação Carlos Chagas Filho de Amparo à Pesquisa do Estado do Rio de Janeiro (FAPERJ) provided support via research grant SEI-260003/002134/2021to S.M.N. and E-26/010.001828/2019 to F.S. Adicional financial support was provided by Universidade Federal do Estado do Rio de Janeiro (UNIRIO, Edital PPQ-UNIRIO nº 04/2020) to F.S. Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) provided IC scholarships to A.S.G. and to V.C.P.M through UNIRIO.

References

Aligizaki, K., Nikolaidis, G., Katikou, p., Baxevanis, a. D., Abatzopoulos, t. J., 2009. Potentially toxic epiphytic Prorocentrum (Dinophyceae) species in Greek coastal waters. Harmful Algae**,** 8 (2), 299-311.

Bandelt, H. J.; Forster, p.; Rohl, a.. 1999. Median-joining networks for inferring intraspecific phylogenies. Molecular Biology And Evolution. 16 (1), 37-48.

Ben-Gharbia, H., Yahia, O., Amzil, Z., Chomérat, N., Abadie, E., Masseret, E., Sibat, M., Triki, H. Z., Nouri, H., Laabir, M., 2016. Toxicity and Growth Assessments of Three Thermophilic Benthic Dinoflagellates (Ostreopsis cf. ovata, Prorocentrum lima and Coolia monotis) Developing in the Southern Mediterranean Basin. Toxins. 8 (10), 297.

Boopathi, T; Faria, D. G., Cheon,J.,; Youn, S. H., Ki, J., 2015. Implications of High Molecular Divergence of Nuclear rRNA and Phylogenetic Structure for the Dinoflagellate Prorocentrum (Dinophyceae, Prorocentrales**).** Journal Of Eukaryotic Microbiology. 62 (4), 519-531.

Bravo, I., Fernández, M. L., Ramilo, I., Martźnez, A., 2001. Toxin composition of the toxic dinoflagellate Prorocentrum lima isolated from different locations along the Galician coast (NW Spain). Toxicon. 39 (10), 1537-1545.

Castresana, J., 2000. Selection of Conserved Blocks from Multiple Alignments for Their Use in Phylogenetic Analysis. Molecular Biology and Evolution. 17 (4), 540-552.

Cembella, A. D.,. Durán-Riveroll, l. M.,. Tarazona-Janampa, U. I., Okolodkov, Y. B., García-Sandoval, R., Krock, B., Hörstmann, C., John, U., 2021. Phylogeography and Diversity Among Populations of the Toxigenic Benthic Dinoflagellate Prorocentrum From Coastal Reef Systems in Mexico. Frontiers In Marine Science**.** 8.

Chomérat, N., Bilien, G., Zentz, F., 2018. A taxonomical study of benthic Prorocentrum species (Prorocentrales, Dinophyceae) from Anse Dufour (Martinique Island, eastern Caribbean Sea). Marine Biodiversity. 49 (3), 1299-1319.

Cohen-Fernandez, E., Pedroche, F. F., Rodríguez Palacio M. C., Hernández, S. A., Meave del Castillo, M., 2010. Molecular phylogeny of Prorocentrum (Dinoflagellata) from the Pacific Coast of Mexico based on the parsimony analysis of fragments of LSUrDNA and SSUrDNA. International Journal of Plant Physiology and Biochemistry. 2. 29-37.

Corander J, Tang J. 2007. Bayesian analysis of population structure based on linked molecular information. Math Biosci. 205 (1), 19-31.

David, H., Laza-Martinez, A., Kromkamp, J., Orive, E., 2017. Physiological response of Prorocentrum lima (Dinophyceae) to varying light intensities. FEMS microbiology ecology. 94.

Ehrenberg, C.G., 1834. Dritter Beitrag zur Erkenntnis großer Organisation in der Richtung des kleinsten Raumes. Physikalische Abhandlungen. Der Königlichen Akademie Der Wissenschaften Zu Berlin., 145-336.

Ehrenber, C. G., 1860. Nachtrag zu Hrn. Ehrenbergs Mittheilung vom 8. Verzeichniss der Leuchtthierchen. Monutsberichte der Kb'niglichen Preitssischen Akademie der Wissenschaften zu Berlin. 791-793.

Faust, M. A., 1994. Three new benthic species of *Prorocentrum* (Dinophyceae) from Carrie Bow Cay, Belize: *P. sabulosum* sp. Nov., *P. sculptile* sp. Nov., and *P. arenarium* sp. Nov.. Journal Of Phycology. 30 (4), 755-763.

Faust, M. A., 2004. Morphology of ciguatera-causing Prorocentrum lima (Pyrrophyta) from widely different sites. Journal of Phycology. 27. 642-648.

Faust, M. A., Vandersea, M., Kibler, S., Tester, P., Litaker, R., 2008. Prorocentrum levis, a new benthic species (Dinophyceae) from a mangrove island, Twin Cays, Belize. Journal of Phycology. 44. 232-240.

Foden, J., Purdie, D., Morris, S., Nascimento, S., 2005. Epiphytic abundance and toxicity of Prorocentrum lima populations in the Fleet Lagoon, UK. Harmful Algae. 4.

Grzebyk, D., Sako, Y., Berland, B., 1998. Phylogenetic analysis of nine species of Prorocentrum (Dinophyceae) inferred from 18S ribosomal DNA sequences, morphological comparisons, and description of Prorocentrum panamensis, sp. nov. Journal of Phycology. 34. 1055-1068.

Guiry, M. D., Guiry, G. M., 2022. Algaebase. World-wide electronic publication, National University of Ireland, Galway. https://www.algaebase.org/search/genus/detail/?Genus_id=43640&sk=0.

Herrera-Sepúlveda, A., Medlin, L., Murugan, G., Sierra, A., Cruz, A., Hernandez Saavedra, N., 2014. Are Prorocentrum hoffmannianum and Prorocentrum belizeanum (DINOPHYCEAE, PROROCENTRALES), the same species? An integration of morphological and molecular data. Journal of Phycology.

Hallegraeff, G., 1998. Transport of toxic dinoflagellates via ships' ballast water: Bioeconomic risk assessment and efficacy of possible ballast water management strategies. Marine Ecology-progress Series. 168. 297-309.

Hoppenrath, M., Chomerat, N., Horiguchi, T., Schweikert, M., Nagahama, Y., Murray, S., 2013. Taxonomy and phylogeny of the benthic Prorocentrum species (Dinophyceae)—A proposal and review. Harmful Algae. 27. 1-28.

Katoh, K., Standley, D. M., 2013. MAFFT Multiple Sequence Alignment Software Version 7: Improvements in Performance and Usability. Molecular Biology And Evolution, 30 (4), 772-780.

Kumar, S., Stecher, G., Tamura, K., 2016. MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. Molecular Biology and Evolution. 33. 1870-1874.

Larsson, M., Laczka, O., Suthers, I., Ajani, P., Doblin, M., 2018. Hitchhiking in the East Australian Current: rafting as a dispersal mechanism for harmful epibenthic dinoflagellates. Marine Ecology Progress Series. 596.

Laza-Martinez, A., Orive, E., Miguel, I., 2011. Morphological and genetic characterization of benthic dinoflagellates of the genera *Coolia*, *Ostreopsis* and *Prorocentrum* from the

south-eastern Bay of Biscay. European Journal Of Phycology, 46 (1), 45-65. Informa UK Limited.

Librado, P., Rozas, J., 2009. Dnasp v5: a software for comprehensive analysis of DNA polymorphism data. Bioinformatics, 25 (11), 1451-1452.

Lim, Z. F., Luo, Z., Keat, L., Hii, K., Teng, S. T., Chan, L., Chomerat, N., Krock, B., Gu, H., Lim, P. T., Leaw, C. P., 2019. Taxonomy and toxicity of Prorocentrum from Perhentian Islands (Malaysia), with a description of a non-toxigenic species Prorocentrum malayense sp. nov. (Dinophyceae). Harmful Algae. 83. 95-108.

Luo, Z., Zhang, H., Krock, B., Lu, S., Yang, W., Gu, H., 2017 Morphology, molecular phylogeny and okadaic acid production of epibenthic Prorocentrum (Dinophyceae) species from the northern South China Sea. Algal Research. 22, 14-30.

Litaker, R., Vandersea, M., Kibler, S., Reece, K., Stokes, N., Steidinger, K., Millie, D., Bendis, B., Pigg, R., Tester, P., 2003. Identification of Pfiesteria piscicida (Dinophycae) and Pfiesteria-like organisms using internal transcribed spacer-specific PCR assays J. Journal of Phycology. 39. 754-761.

Litaker, R., Vandersea, M., Kibler, S., Reece, K., Stokes, N., Lutzoni, F., Yonish, B., West, M., Black, M., Tester, P., 2007. Recognizing dinoflagellate species using ITS rDNA sequences. Journal of Phycology. 43. 344-355.

Masó, M., Garcés, E., Pagès, F., Camp, J., 2003. Drifting plastic debris as a potential vector for dispersing Harmful Algal Bloom (HAB) species. Scientia Marina. 67.

Mohammad-Noor, N., Moestrup, Ø., Daugbjerg, N., 2007. Light, electron microscopy and DNA sequences of the dinoflagellate *Prorocentrum concavum* (syn. P. Arabianum) with special emphasis on the periflagellar area. Phycologia, 46 (5) 549-564.

Moreira González, A. Fernandes, L., Uchida, H., Uesugi, A., Suzuki, T., Chomerat, N., Bilien, G., Mafra, L., 2019. Variations in morphology, growth, and toxicity among strains of the Prorocentrum lima species complex isolated from Cuba and Brazil. Journal of Applied Phycology. 31.

Murata, M., Shimatani, M., Sugitani, H., Oshima, Y., Yasumoto, T., 1982. Isolation and structural elucidation of the causative toxin of the diarrhetic shellfish poisoning. Nippon Suisan Gakkaishi 48, 549–552.

Murray, S., Ip, C.L.C., Moore, R., Nagahama, Y., Fukuyo, Y., 2009. Are Prorocentroid dinoflagellates monophyletic? A study of 25 species based on nuclear and mitochondrial genes. Protist 160 (2), 245–264.

Murray, S., Nishimura, T., Fyfe, R., Harwood, T., Rhodes, L., 2018. Sampling benthic dinoflagellates in the Solomon Islands. Harmful Algae News 61, 11–12.

Nagahama, Y., Fukuyo, Y., 2005. Redescription of *Crypotomonas lima*, collected from Sorrento, Italy, the basioniym fo Prorocentrum lima. Plankton Biol. Ecol. 52 (2), 107–109.

Nagahama, Y., Murray, S., Tomaru, A., Fukuyo, Y., 2011. Species boundaries in the toxic dinoflagellate Prorocentrum lima (Dinophyceae, Prorocentrales), based on morphological and phylogenetic characters. J. Phycol. 47 (1), 178–189.

Nascimento, S. M., Salgueiro, F., Menezes, M., de Andréa Oliveira, F., Magalhães, V. C. P., De Paula, J. C., Morris, S., 2016. Prorocentrum lima from the South Atlantic: morphological, molecular and toxicological characterization. Harmful Algae 57, 39–48.

Nascimento, S. M., Mendes, M. C. Q., Menezes, M., Rodríguez, F., Alves-de-Souza, C., Branco, S., Riobó, P., Franco, J., Nunes, J. M. C., Huk, M., Morris, S., Fraga, S., 2017. Morphology and phylogeny of Prorocentrum caipirignum sp. nov. (Dinophyceae), a new tropical toxic benthic dinoflagellate. Harmful Algae 70, 73–89.

Nishimura, T., Uchida, H., Noguchi, R., Oikawa, H., Suzuki, T., Funaki, H., Ihara, C., Hagino, K., Arimitsu, S., Tanii, Y., Abe, S., Hashimoto, K., Mimura, K., Tanaka, K., Yanagida, I., Adachi, M., 2020. Abundance of the benthic dinoflagellate Prorocentrum and the diversity, distribution, and diarrhetic shellfish toxin production of Prorocentrum lima complex and P. caipirignum in Japan. Harmful Algae. 96. 101687.

Nunn, G.B., Theisen, B.F., Christensen, B., Arctander, P., 1996. Simplicity-correlated size growth of the nuclear 28S ribosomal RNA D3 expansion segment in the crustacean order Isopoda. J. Mol. Evol. 42, 211–223.

Ronquist, F., Teslenko, M., van der Mark, P., Ayres, D.L., Darling, A., Höhna, S., Larget, B., Liu, L., Suchard, M.A., Huelsenbeck, J.P., 2012. MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. Syst. Biol. 61, 539–542.

Sato, S., Nishimura, T., Uehara, K., Sakanari, H., Tawong, W., Hariganeya, N., Smith, K., Rhodes, L., Yasumoto, T., Taira, Y., Suda, S., Yamaguchi, H., Adachi, M., 2011. Phylogeography of Ostreopsis along west Pacific coast, with special reference to a novel clade from Japan. PLOS ONE 6.

Scholin, C.A., Herzog, M., Sogin, M., Anderson, D.M., 1994. Identification of group and strain-specific genetic markers for globally distributed Alexandrium (Dinophyceae). II. Sequence analysis of a fragment of the LSU rRNA gene. J. Phycol. 30, 999–1011.

Silva, R. A. F., Goulart, M. C., Neves, R. A. F., Nascimento, S. M., 2020. Morfologia de cepas de *Prorocentrum lima* e *Gambierdiscus* sp. isolados do Oceano Atlântico Sul**.** Rio de Janeiro. Unirio.

Spalding, M., Fox, H., Allen, G., Davidson, N., Ferdaña, Z., Finlayson, M., Halpern, B., Jorge, M., Lombana, A., Lourie, S., Martin, K., McManus, E., Molnar, J., Recchia, C., Robertson, J., 2007. Marine Ecoregions of the World: A Bioregionalization of Coastal and Shelf Areas. BioScience. 57. 573-583.

Stern, R., Andersen, R., Jameson, I., Küpper, F., Coffroth, M. A., Vaulot, D. Le Gall, F., Véron, B., Brand, J., Skelton, H., Kasai, F., Lilly, E., Keeling, P., 2012. Evaluating the Ribosomal Internal Transcribed Spacer (ITS) as a Candidate Dinoflagellate Barcode Marker. PloS one. 7 (8), 1-12.

Zardoya, R., Costas, E., López-Rodas, V., Garrido-Pertierra, A., Bautista, J. M., 1995. Revised dinoflagellate phylogeny inferred from molecular analysis of large-subunit ribosomal RNA gene sequences. J. Mol. Evol. 41 (5), 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, H.., Bhattacharya, D., Lin, S. A., 2007. Three-Gene Dinoflagellate Phylogeny Suggests Monophyly of Prorocentrales and a Basal Position for Amphidinium and Heterocapsa. Journal Of Molecular Evolution**,** 65 (4), 463-474..

Zhang, H., Li, Y., Cen, J., Wang, H., Cui, L., Dong, Y., Lu, S., 2015. Morphotypes of Prorocentrum lima (Dinophyceae) from Hainan Island, South China Sea: morphological and molecular characterization. Phycologia 54 (5), 503–516.

Supplementary Data

Table S1: Strains obtained in the present work, sampling site, geographic coordinates of each site. N.A.-Not available.

Strain name	Location of sampling site	Latitude, Longitude
UNR-1	Armação dos Búzios, Rio de Janeiro, Brazil	22º 44' 49" S; 41º 52' 54" W
UNR-15	Maragogi, Alagoas, Brazil	
UNR-17	Trindade and Martim Vaz archipelago, Espírito Santo, Brazil	20° 29' 22.2"S; 29° 20' 04.2" W
UNR-18	Trindade and Martim Vaz archipelago, Espírito Santo, Brazil	20° 29' 22.2"S; 29° 20' 04.2" W
UNR-19	Trindade and Martim Vaz archipelago, Espírito Santo, Brazil	20° 29' 22.2"S; 29° 20' 04.2" W
UNR-20	Trindade and Martim Vaz archipelago, Espírito Santo, Brazil	20° 29' 22.2"S; 29° 20' 04.2" W
UNR-21	Trindade and Martim Vaz archipelago, Espírito Santo, Brazil	20° 29' 22.2"S; 29° 20' 04.2" W
UNR-32	Maragogi, Alagoas, Brazil	8° 55' 18"S; 35° 09' 05.8" W
UNR-33	Fernando de Noronha archipelago, Pernambuco, Brazil	3° 50' 52.3"S; 32° 26' 31.8"W
UNR-34	F. de Noronha archipelago, Pernambuco, Brazil	3° 50' 52.3"S; 32° 26' 31.8"W
UNR-35	F. de Noronha archipelago, Pernambuco, Brazil	
UNR-36	F. de Noronha archipelago, Pernambuco, Brazil	3° 50' 52.3"S; 32° 26' 31.8"W
UNR-37	F. de Noronha archipelago, Pernambuco, Brazil	3° 50' 52.3"S; 32° 26' 31.8"W
UNR-38	F. de Noronha archipelago, Pernambuco, Brazil	3° 50' 52.3"S; 32° 26' 31.8"W
UNR-68	Praia dos Carneiros, Pernambuco, Brazil	8°42'16"S; 35°04'43"W
UNR-71	Praia do Forte, Bahia, Brazil	12°34'41.0" S; 38°00'04.7" W
UNR-72	Praia do Forte, Bahia, Brazil	12°34'41.0" S; 38°00'04.7" W
UNR-73	Praia do Forte, Bahia, Brazil	12°34'41.0" S; 38°00'04.7" W
UNR-74	Brazil	N.A.
UNR-75	Maragogi, Alagoas, Brazil	8° 55' 18"S; 35° 09' 05.8" W
UNR-76	Maragogi, Alagoas, Brazil	8° 55' 18"S; 35° 09' 05.8" W
UNR-86	F. de Noronha archipelago, Pernambuco, Brazil	3° 50' 52.3"S; 32° 26' 31.8"W
BC1	Barra do Cunhaú, Rio Grande do Norte, Brazil	6°17'18.0"S; 35°02'02.0"W
VGO 776	Tenerife, Spain	N.A.
PL 27V	Canelas beach,Vigo, Spain	42° 38’ 90.3” N; 8° 83’19.3” O

Table S2: Strains, Genbank codes for each marker used (ITS, LSU and SSU), Collection sites and references of the sequences used at the analyses. N.I.-Not Informed. XXX- Not available.

Strain	ITS	LSU	SSU	Collection site	Reference
OMI3P	XXX	LC415597.1	XXX	Japan, Okinawa, Motobu Cho, Ishikawa	(Nishimura et al., 2020)
OUN122P	XXX	LC415596.1	XXX	Japan, Okinawa, Uruma City, Nakagusuku Bay	(Nishimura et al., 2020)
OMI29P	XXX	LC415595.1	XXX	Japan, Okinawa, Motobu Cho, Ishikawa	(Nishimura et al., 2020)
OMI6P	XXX	LC415594.1	XXX	Japan, Okinawa, Motobu Cho, Ishikawa	(Nishimura et al., 2020)
OMI4P	XXX	LC415598.1	XXX	Japan, Okinawa, Motobu Cho, Ishikawa	(Nishimura et al., 2020)
KTH8P	XXX	LC415593.1	XXX	Japan, Kagoshima, Amami-oshima Island, Tatsugo Cho, Heart Rock	(Nishimura et al., 2020)
MIO12P	XXX	LC415592.1	XXX	Japan, Miyagi, Ishinomaki City, Oginohama	(Nishimura et al., 2020)
AOF93P	XXX	LC415591.1	XXX	Japan, Akita, Oga City, Funagawaminatodaishimaunosaki	(Nishimura et al., 2020)
MIO52P	XXX	LC415590.1	XXX	Japan, Miyagi, Ishinomaki City, Oginohama	(Nishimura et al., 2020)
OUN22P30	XXX	LC415589.1	XXX	Japan, Okinawa, Uruma City, Nakagusuku Bay	(Nishimura et al., 2020)
AHM3P	XXX	LC415587.1	XXX	Japan, Aomori, Hiranai Machi, Mouratsukidomari	(Nishimura et al., 2020)
OMI8P	XXX	LC415588.1	XXX	Japan, Okinawa, Motobu Cho, Ishikawa	(Nishimura et al., 2020)
AOF70P	XXX	LC415586.1	XXX	Japan, Akita, Oga City, Funagawaminatodaishimaunosaki	(Nishimura et al., 2020)
AOF55P	XXX	LC415585.1	XXX	Japan, Akita, Oga City, Funagawaminatodaishimaunosaki	(Nishimura et al., 2020)
KMK75P	XXX	LC415584.1	XXX	Japan, Kochi, Muroto City, Kannoura	(Nishimura et al., 2020)
OUN152P30	XXX	LC415583.1	XXX	Japan, Okinawa, Uruma City, Nakagusuku Bay	(Nishimura et al., 2020)
KON61P	XXX	LC415582.1	XXX	Japan, Kochi, Otsuki Cho, Nishidomari	(Nishimura et al., 2020)
OMI9P	XXX	LC415581.1	XXX	Japan, Okinawa, Motobu Cho, Ishikawa	(Nishimura et al., 2020)
OMI39P	XXX	LC415580.1	XXX	Japan, Okinawa, Motobu Cho, Ishikawa	(Nishimura et al., 2020)
KTH5P	XXX	LC415579.1	XXX	Japan, Kagoshima, Amami-oshima Island, Tatsugo Cho, Heart Rock	(Nishimura et al., 2020)
OMI36P	XXX	LC415578.1	XXX	Japan, Okinawa, Motobu Cho, Ishikawa	(Nishimura et al., 2020)
KON81P	XXX	LC415577.1	XXX	Japan, Kochi, Otsuki Cho, Nishidomari	(Nishimura et al., 2020)
KON83P	XXX	LC415576.1	XXX	Japan, Kochi, Otsuki Cho, Nishidomari	(Nishimura et al., 2020)
PPL2-4	AB189769.1	XXX	AB189776.1	Philippines, Palawan Island	(Nagahama et al.,2011)
OAK-TO-PL	AB189768.1	XXX	AB189775.1	Japan, Aka Island, Okinawa, Akashima	(Nagahama et al.,2011)
Tahiti	AB189742.1	XXX	AB189774.1	Tahiti, French Polynesian	(Nagahama et al.,2011)

PDIO-3	AB189770.1	XXX	XXX	Philippines, Dio Island	(Nagahama et al.,2011)
95-INDO-PLIMA	AB189772.1	XXX	XXX	Indonesia	(Nagahama et al.,2011)
AK9001	AB189771.1	XXX	AB189773.1	Japan, Amatsukominato, Chiba	(Nagahama et al.,2011)
SD1	KM266627.1	KP063225.1	XXX	China, Hainan Island	(Zhang et al., 2015)
SD2	KM266630.1	KP063226.1	XXX	China, Hainan Island	(Zhang et al., 2015)
SE7	XXX	KP063230.1	XXX	China, Hainan Island	(Zhang et al., 2015)
SE5	KM266629.1	KP063229.1	XXX	China, Hainan Island	(Zhang et al., 2015)
SD11	KM266631.1	KP063228.1	XXX	China, Hainan Island	(Zhang et al., 2015)
SD4	KM266628.1	KP063227.1	XXX	China, Hainan Island	(Zhang et al., 2015)
SC7	XXX	KP063224.1	XXX	China, Hainan Island	(Zhang et al., 2015)
2S1F7	KM266632.1	KP063215.1	XXX	China, Hainan Island	(Zhang et al., 2015)
#151	XXX	XXX	Y16235.1	Japan, Tokushima, Mugi Ooshima	(Grzebyk et al., 1998)
TIO124	KY010236.1	KY010250.1	XXX	China, Beihai, Guangxi	(Luo et al., 2017)
TIO177c	KY010237.1	XXX	XXX	China, Beihai, Guangxi	(Luo et al., 2017)
TIO155a	KY010232.1	KY010251.1	XXX	China, Beihai, Guangxi	(Luo et al., 2017)
TIO164	KY010234.1	XXX	XXX	China, Beihai, Guangxi	(Luo et al., 2017)
TIO163	KY010233.1	XXX	XXX	China, Beihai, Guangxi	(Luo et al., 2017)
DNS-7	XXX	DQ336195.1	XXX	N.I	Unpublished
DNS-3	XXX	DQ336187.1	XXX	N.I	Unpublished
CCMP1370	EU921507.1	XXX	XXX	N.I	Unpublished
S4	XXX	DQ336193.1	XXX	N.I	Unpublished
SKLMP_W074	XXX	MG914032.1	XXX	Hong Kong, Tai She Wan	Unpublished
CAWD283	XXX	LC422235.1	XXX	New Zealand, Rangitahua/Kermadec Islands, Raoul Island	Unpublished
XS336	MH381780.1	MH375426.1	XXX	China, Paracel Islands	Unpublished
XS326	MH375434.1	MH375425.1	XXX	China, Paracel Islands	Unpublished
3XS36	MH375433.1	MH375414.1	XXX	China, Paracel Islands	Unpublished
3XS34	MH375432.1	MH375415.1	XXX	China, Paracel Islands	Unpublished
XS575	MH356574.1	MH348972.1	XXX	China, Paracel Islands	Unpublished
B2RTH11	XXX	LC413904.1	XXX	Solomon Islands, Kinugawa Maru WWII wreck, Guadalcanal	(Murray et al., 2018)
B2RTH10	XXX	LC413903.1	XXX	Solomon Island, Kinugawa Maru WWII wreck, Guadalcanal	(Murray et al., 2018)
B2RTH7	XXX	LC413902.1	XXX	Solomon Island, Kinugawa Maru WWII wreck, Guadalcanal	(Murray et al., 2018)

NPHLB4	XXX	LC413901.1	XXX	Solomon Island ,Honiara, Guadalcanal	(Murray et al., 2018)
CIBNOR-PRL1	XXX	EF517252.1	EF517266.1	Mexico, Isla El Pardito Golfo da California	(Cohen-Fernandez et al., 2010)
PL7V	XXX	EF517253.1	EF517265.1	Spain, Ría de Pontevedra (Bueu, batea)	(Cohen-Fernandez et al., 2010) (Zardoya et al., 1995)
SM45	XXX	XXX	EU196419.1	Japan, Akajima	(Murray et al., 2009)
SM29	XXX	DQ336181.1	XXX	Australia	(Martinez et al., 2011) (Murray et al., 2009)
PLRN_02	XXX	AJ567457.1	XXX	France, Reunion Island	Unpublished
PLMA_01	XXX	AJ567458.1	XXX	France, Reunion Island	Unpublished
PLMK_02	XXX	AJ567459.1	XXX	France, Reunion Island	Unpublished
CCMP1541	EU927564.1	XXX	XXX	N.I	Unpublished
CAWD189	XXX	MW177928.1	XXX	USA, Hawaii	Direct submission
CRLMN-6	AB189744.1	XXX	AB189778.1	Costa Rica, Puerto Limon, Limon	(Nagahama et al.,2011)
USSP-F2	AB189749.1	XXX	XXX	USA, Florida, St. Pete beach	(Nagahama et al.,2011)
USSP-F11	AB189746.1	XXX	XXX	USA, Florida, St. Pete beach	(Nagahama et al.,2011)
USSP-F12	AB189747.1	XXX	XXX	USA, Florida, St. Pete beach	(Nagahama et al.,2011)
USSP-S2	AB189751.1	XXX	XXX	USA, Florida, St. Pete beach	(Nagahama et al.,2011)
USSP-S15	AB189753.1	XXX	XXX	USA, Florida, St. Pete beach	(Nagahama et al.,2011)
USSP-S18	AB189752.1	XXX	XXX	USA, Florida, St. Pete beach	(Nagahama et al.,2011)
USSP-F8	AB189748.1	XXX	AB189777.1	USA, Florida, St. Pete beach	(Nagahama et al.,2011)
BM-U2-A5	AB189757.1	XXX	XXX	Bermuda, UK	(Nagahama et al.,2011)
BM-U2-B2	AB189756.1	XXX	XXX	Bermuda, UK	(Nagahama et al.,2011)
BM-U2-C1	AB189758.1	XXX	XXX	Bermuda, UK	(Nagahama et al.,2011)
BM-U2-D5	AB189755.1	XXX	XXX	Bermuda, UK	(Nagahama et al.,2011)
BM-U2-B1	AB189754.1	XXX	XXX	Bermuda, UK	(Nagahama et al.,2011)
USTM-D6	AB189750.1	XXX	XXX	USA, Florida, Tampa Bay	(Nagahama et al.,2011)
CRLMN-7	AB189745.1	XXX	XXX	Costa Rica, Puerto Limon, Limon	(Nagahama et al.,2011)
CRLMN-4	AB189743.1	XXX	XXX	Costa Rica, Puerto Limon, Limon	(Nagahama et al.,2011)
UNR-01	KU722938.1	KU198627.1	XXX	Brazil, RJ, Praia da Tartaruga, Armação dos Búzios	(Nascimento et al., 2016)
UNR-09	KU198629.1	XXX	XXX	Brazil, RJ, Praia da Tartaruga, Armação dos Búzios	(Nascimento et al., 2016)
PLHV-1	XXX	JQ616842.1	JQ638930.1	Cuba	(Herrera-sepúlveda et al., 2012) https://www.cibnor.gob.mx/en/research/biological-collections/codimar/home
CCMP1746	EU927509.1	DQ336186.1	XXX	Belize, Southwater Cay	(Martinez etal., 2011) (Murray et al., 2009) (https://ncma.bigelow.org/CCMP1746)
CCMP1368	EU927506.1	EU165317.1	XXX	USA, Florida, Knight Key	(Scorzetti et al., 2009) (https://ncma.bigelow.org/CCMP1368)

OUN2P30	XXX	LC415603.1	XXX	Japan, Okinawa, Uruma City, Nakagusuku Bay	(Nishimura et al., 2019)
OUN137P30	XXX	LC415602.1	XXX	Japan, Okinawa, Uruma City, Nakagusuku Bay	(Nishimura et al., 2019)
KMK21P	XXX	LC415601.1	XXX	Japan, Kochi, Muroto City, Kannoura	(Nishimura et al., 2019)
CF10	KM266619.1	KP063218.1	XXX	China, Hainan Island	(Zhang et al., 2015)
PF11	XXX	KP063223.1	XXX	China, Hainan Island	(Zhang et al., 2015)
NG5	KM266620.1	KP063222.1	XXX	China, Hainan Island	(Zhang et al., 2015)
NF6	KM266621.1	KP063221.1	XXX	China, Hainan Island	(Zhang et al., 2015)
BS4F5	KM266622.1	KP063217.1	XXX	China, Hainan Island	(Zhang et al., 2015)
SM24		DQ336182.1		N.I	Unpublished
2S1E12	KM266623.1	KP063214.1	XXX	China, Hainan Island	(Zhang et al., 2015)
PMAYD1	XXX	XXX	Y16234.1	France, Mayotte Island	Direct submission
IFR11-063	XXX	MG701857.1	XXX	France, Martinique Island, Anse Dufour	(Chomérat et al., 2018)
NMN07	XXX	EF566748.1	XXX	N.I	(Mohammad-Noor et al., 2007)
VGO776	EU244470.1	KY053858.1	XXX	Spain, Pta. Hidalgo, Tenerife	Unpublished
PAEU_01	XXX	AJ567456.1	XXX	France, Reunion Island	Unpublished
CCMP685	AB189765.1;	DQ336179.1	XXX	Spain, Ria de Vido, Vigo	(Nagahama et al.,2011) (https://ncma.bigelow.org/CCMP685)
CCMP686	AB189766.1;	XXX	XXX	Spain, Ria de Vido, Vigo	(Nagahama et al.,2011) (https://ncma.bigelow.org/CCMP686)
CCMP1743	AB189767.1;	XXX	XXX	Canada, Nova Escócia, Mahone Bay	(Nagahama et al.,2011) (https://ncma.bigelow.org/CCMP1743)
CCMP684	AB189764.1	XXX	AB189779.1	Spain, Lago Cies, Vigo	(Nagahama et al.,2011) (https://ncma.bigelow.org/CCMP684)
USMA-2	AB189763.1	XXX	XXX	USA, Clam Cove, Maine	(Nagahama et al.,2011)
USMA-1	AB189762.1	XXX	XXX	USA, Clam Cove, Maine	(Nagahama et al.,2011)
USMA-6	AB189760.1	XXX	XXX	USA, Clam Cove, Maine	(Nagahama et al.,2011)
USMA-5	AB189761.1	XXX	XXX	USA, Clam Cove, Maine	(Nagahama et al.,2011)
USMA-4	AB189759.1	XXX	AB189780.1	USA, Clam Cove, Maine	(Nagahama et al.,2011)
Dn150EHU	KT898148.1	XXX	XXX	France, Villefranche-sur-Mer	(David et al., 2017)
Dn116EHU	KT898156.1	XXX	XXX	Portugal, Galé, (South of ´the Iberian Peninsula)	(David et al., 2017)
Dn141EHU	KT898161.1	XXX	XXX	Spain, Ibiza	(David et al., 2017)
PLBZT14	XXX	KX845009.1	XXX	Tunisia, Baia Bizerta	(Ben-gharbia et al., 2016)
CCMP1966	XXX	XXX	EF377326.1	USA, New Meadows River near Bath, Maine	(Zhang et al., 2007) (https://ncma.bigelow.org/CCMP1966)
PQ1	EU927488.1	XXX	XXX	Canada, New Scotland, Pomquet	Unpublished

PQ2	EU927489.1	XXX	XXX	N.I	Unpublished
PQ3	EU927490.1	XXX	XXX	N.I	Unpublished
IP197	EU927491.1	XXX	XXX	N.I	Unpublished
IP297	EU927492.1	XXX	XXX	N.I	Unpublished
IP797	EU927493.1	XXX	XXX	N.I	Unpublished
PA	EU927487.1	XXX	XXX	N.I	Unpublished
4V	EU244474.1	XXX	XXX	Spain, Galicia, Ria de Aldan	Unpublished
Dn35EHU	XXX	HQ414228.1	XXX	Spain, Biscay	(Martinez et al., 2011)
Dn37EHU	XXX	HQ414231.1	XXX	Spain, Biscay	(Martinez et al., 2011)
Dn38EHU	XXX	HQ414229.1	XXX	Spain, Biscay	(Martinez et al., 2011)
Dn39EHU	XXX	HQ414230.1	XXX	Spain, Santander	(Martinez et al., 2011)
AR2	EU927500.1	XXX	XXX	N.I	Unpublished
AR3	EU927501.1	XXX	XXX	N.I	Unpublished
KP200	EU927494.1	XXX	XXX	N.I	Unpublished
KP201	EU927495.1	XXX	XXX	N.I	Unpublished
KP202	EU927496.1	XXX	XXX	N.I	Unpublished
KP206	EU927497.1	XXX	XXX	N.I	Unpublished
KP209	EU927499.1	XXX	XXX	N.I	Unpublished
KP208	EU927498.1	XXX	XXX	N.I	Unpublished
AR4	EU927502.1	XXX	XXX	N.I	Unpublished
CCMP2273	EU927511.1	XXX	XXX	USA, New Meadows River near Bath, Maine	https://ncma.bigelow.org/CCMP2273
CCMP2578	EU927512.1	XXX	XXX	USA, Point Judith Pond, Rhode Island	https://ncma.bigelow.org/CCMP2578
CCMP2580	EU927514.1	XXX	XXX	USA, Clam Cove, Maine	https://ncma.bigelow.org/CCMP2580
CCMP2581	EU927515.1	XXX	XXX	USA, Clam Cove, Maine	https://ncma.bigelow.org/CCMP2581
CCMP2583	EU927516.1	XXX	XXX	USA, Bluff Hill Cove, Rhode Island	https://ncma.bigelow.org/CCMP2583
CCMP684	EU927503.1	XXX	XXX	Spain, Lago Cies, Vigo	https://ncma.bigelow.org/CCMP684
CCMP2589	EU927513.1	XXX	XXX	N.I	https://ncma.bigelow.org/CCMP2589
CCMP1999	EU927510.1	XXX	XXX	N.I	https://ncma.bigelow.org/CCMP1999
CCMP2584	EU927517.1	XXX	XXX	USA, New Meadows River, Maine	https://ncma.bigelow.org/CCMP2584
CCAP 1136/11	XXX	XXX	MK541784.1	Spain, Rıa de Vigo, Galicia	(Varkitzi et al., 2010) https://www.ccap.ac.uk/strain_info.php?Strain_No=1136/11
CCAP 1136/12	EU927519.1	XXX	MK541780.1	England, Marine; Gibralter Point Visitor Centre, Lincolnshire	https://www.ccap.ac.uk/strain_info.php?Strain_No=1136/12

CCAP 1136/9	EU927518.1	XXX	XXX	England, Marine; Gibralter Point Visitor Centre, Lincolnshire	(Stern et al., 2012) https://www.ccap.ac.uk/strain_info.php?Strain_No=1136/9
PLLS01	XXX	AY259170.1	XXX	Australia, Little Swanport, Tasmania	(Pearce e Hallegraeff., 2004)
Sorrento 1	XXX	DQ336189.1	XXX	N.I	Unpublished
Dn76EHU	KT898159.1	XXX	XXX	Portugal, Peniche	Unpublished
Dn82EHU	KT898154.1	XXX	XXX	Spain, A coruna	Unpublished
Dn61EHU	KT898157.1	XXX	XXX	Spain, San Sebastian	Unpublished
Dn53EHU	KT898153.1	XXX	XXX	Spain, Santander	Unpublished
Dn50EHU	KT898152.1	XXX	XXX	Spain, Zierbana	Unpublished
Dn57EHU	KT898150.1	XXX	XXX	Spain, Santona	Unpublished
Dn79EHU	KT898158.1	XXX	XXX	Spain, Vigo	Unpublished
Dn109EHU	KT898149.1	XXX	XXX	Greece, Crete	Unpublished
Dn88EHU	KT898151.1	XXX	XXX	Portugal, Cascais	Unpublished
Dn157EHU	KT898155.1	XXX	XXX	Spain, Santander	Unpublished
Pl.10	KX026864.1	XXX	XXX	Tunisia, Chebba	Unpublished
Pl.1	KJ781423.1	XXX	XXX	Tunisia, Borj Djelijel, Medenine	Unpublished
Pl.2	KJ781421.1	XXX	XXX	Tunisia, Chebba, Mahdia	Unpublished
Pl.16	KX026866.1	XXX	XXX	Tunisia, Cheik Yahia	Unpublished
Pl.7	KJ781422.1	XXX	XXX	Tunisia, Chebba, Mahdia	Unpublished
SHOU-Dino-Pro-001	JN717141.1	XXX	JN717143.1	N.I	Unpublished
FIUPL	XXX	EU165316.1	XXX	USA	(Scorzetti et al., 2009)
CAWD94	XXX	MW177926.1	XXX	New Zealand, Rangiputa	Direct submission
CAWD176	XXX	MW177927.1	XXX	New Zealand, Rangaunu Harbour	Direct submission
CAWD33	XXX	MW177923.1	XXX	New Zealand, Rangaunu	Direct submission
CAWD69	XXX	MW177924.1	XXX	New Zealand, Rangiputa	Direct submission
CAWD70	XXX	MW177925.1	XXX	New Zealand, Whatuwhiwhi	Direct submission
CAWD32	XXX	MW177922.1	XXX	Spain	Direct submission
AS4F8	KM266624.1	KP063216.1	XXX	China, Hainan Island	(Zhang et al., 2015)
DS4D9	KM266626.1	KP063219.1	XXX	China, Hainan Island	(Zhang et al., 2015)
DS4G4	KM266625.1	KP063220.1	XXX	China, Hainan Island	(Zhang et al., 2015)
TIO180	KY010243.1	KY010256.1	XXX	China, Beihai, Guangxi	(Luo et al., 2017)
TIO179	KY010242.1	KY010255.1	XXX	China, Beihai, Guangxi	(Luo et al., 2017)

SE10	KM266633.1	KP063231.1	XXX	China: Hainan Island	(Zhang et al., 2015)
UFBA	KY039500.1	KY039499.1	XXX	Brazil, Bahia, Garapuá, Tinharé island, Cairu	(Nascimento et al., 2017)
LCA-B4	KJ960192.1	XXX	XXX	Brazil, Rio de Janeiro, Arraial do Cabo	(Nascimento et al., 2017)
XS331	MH356573.1	MH348971.1	XXX	China, Paracel Islands	Unpublished
KON30P15	XXX	LC415610.1	XXX	Japan, Kochi, Otsuki Cho, Nishidomari	(Nishimura et al., 2020)
OUN9P30	XXX	LC415607.1	XXX	Japan, Okinawa, Uruma City, Nakagusuku Bay	(Nishimura et al., 2020)
OUN8P30	XXX	LC415606.1	XXX	Japan, Okinawa, Uruma City, Nakagusuku Bay	(Nishimura et al., 2020)
OUN156P20	XXX	LC415611.1	XXX	Japan, Okinawa, Uruma City, Nakagusuku Bay	(Nishimura et al., 2020)
KON26P15	XXX	LC415609.1	XXX	Japan, Kochi, Otsuki Cho, Nishidomari	(Nishimura et al., 2020)
OUN26P20	XXX	LC415608.1	XXX	Japan, Okinawa, Uruma City, Nakagusuku Bay	(Nishimura et al., 2020)
TIO138	KY010241.1	KY010254.1	XXX	China, Sanya, Hainan	(Luo et al., 2017)
TIO11	KY010238.1	KY010252.1	XXX	China, Sanya, Hainan	(Luo et al., 2017)
TIO139	KY010240.1	XXX	XXX	China, Sanya, Hainan	(Luo et al., 2017)
TIO102	KY010239.1	KY010253.1	XXX	China, Sanya, Hainan	(Luo et al., 2017)
K-0625	XXX	EF566747.1	XXX	N.I	(Mohammad-Noor et al., 2007)
PMHV-1	JQ638940.2	JQ638945.1	XXX	Cuba	(Herrera-sepulveda et al., 2012)
A10PR01	MG600151.1	MG600144.1	XXX	Malaysia, Rawa Island, Perhentian Islands Marine Park	(Lim et al., 2019)
P. belizeanum	XXX	XXX	DQ238042.1	Belize, Carrie Bow Cay	(Faust et al., 2008)
PPAN20	XXX	XXX	Y16236.1	Panama, Contadora Island	(Grzebyk et al., 1998)
CCMP683	KF885225.1	XXX	KF885225.1	USA	(Herrera-sepulveda et al., 2015)
CCMP2804	KF885224.1	XXX	KF885224.1	USA	(Herrera-sepulveda et al., 2015)
PBHV-1	JQ638934.2	XXX	JQ638934.2	Cuba	(Herrera-sepulveda et al., 2015)
PMHV-1	XXX	XXX	JQ638940.2	Cuba	(Herrera-sepulveda et al., 2015)
CCMP2633	XXX	XXX	KF885226.1	Belize, Carrie Bow Cay	(Herrera-sepulveda et al., 2015)
CCMP2904	XXX	XXX	XXX	N.I	Unpublished
PHGE_02	XXX	AJ567463	XXX	N.I	Unpublished
PHSE_01_2	XXX	AJ567462	XXX	N.I	Unpublished
PHSE_01	XXX	AJ567461	XXX	N.I	Unpublished
OUN37P10	XXX	LC415604.1	XXX	Japan, Okinawa, Uruma City, Nakagusuku Bay	(Nishimura et al., 2020)
OUN248P	XXX	LC415605.1	XXX	Japan, Okinawa, Uruma City, Nakagusuku Bay	(Nishimura et al., 2020)
PL1-11	XXX	DQ336188.1	XXX	N.I	Unpublished
VGO880	FJ489615.1	XXX	XXX	Spain	Unpublished

IFR10-171 *(P. bimaculatum)*	XXX	HQ890883.1	XXX	Kuwait	Direct submission
IFR10-096	XXX	XXX	HQ890884	Arabian Gulf	(Cembella et al.,2021)
PA8	XXX	MZ310165.1	XXX	Mexico, Veracruz	(Cembella et al.,2021)
PA66	XXX	MZ310161.1	XXX	Mexico, Veracruz	(Cembella et al.,2021)
PA77	XXX	MZ310163.1	XXX	Mexico, Veracruz	(Cembella et al.,2021)
PA78	XXX	MZ310164.1	XXX	Mexico, Veracruz	(Cembella et al.,2021)
PA82	XXX	MZ310167.1	XXX	Mexico, Veracruz	(Cembella et al.,2021)
PA83	XXX	MZ310168.1	XXX	Mexico, Veracruz	(Cembella et al.,2021)
PA84	XXX	MZ310169.1	XXX	Mexico, Veracruz	(Cembella et al.,2021)
PA95	XXX	MZ310170.1	XXX	Mexico, Veracruz	(Cembella et al.,2021)
PA80	XXX	MZ310166.1	XXX	Mexico, Veracruz	(Cembella et al.,2021)
PA62	XXX	MZ310159.1	XXX	Mexico, Veracruz	(Cembella et al.,2021)
PA17	XXX	MZ310155.1	XXX	Mexico, Veracruz	(Cembella et al.,2021)
PA46	XXX	MZ310156.1	XXX	Mexico, Baja California Sur	(Cembella et al.,2021)
PA49	MZ308617.1	MZ310157.1	XXX	Mexico, Baja California Sur	(Cembella et al.,2021)
PA63	XXX	MZ310160.1	XXX	Mexico, Veracruz	(Cembella et al.,2021)
PA97	XXX	MZ310171.1	XXX	Mexico, Veracruz	(Cembella et al.,2021)
PA76	MZ308609.1	XXX	XXX	Mexico, Veracruz	(Cembella et al.,2021)
PA72	MZ308616.1	XXX	XXX	Mexico, Veracruz	(Cembella et al.,2021)
PA48	MZ308607.1	XXX	XXX	Mexico, Baja California Sur	(Cembella et al.,2021)
PA104	MZ308606.1	XXX	XXX	Mexico, Veracruz	(Cembella et al.,2021)
PA90	MZ308613.1	XXX	XXX	Mexico, Veracruz	(Cembella et al.,2021)
PA94	MZ308615.1	XXX	XXX	Mexico, Veracruz	(Cembella et al.,2021)
PA86	MZ308611.1	XXX	XXX	Mexico, Veracruz	(Cembella et al.,2021)

Conclusões gerais:

Conclui-se, considerando as análises filogenéticas e filogeográficas realizadas no presente estudo que o PLSC é composto por pelo menos três linhagens genéticas distintas, com distribuições biogeográficas parcialmente sobrepostas, podendo corresponder a três espécies diferentes. As sequências de *P. lima* (clado 4), *P. arenarium* (clado 3) e *P.* cf. *lima* "tropical" (clados 1 e 2) apresentaram valores de distância-p elevados entre si, comparáveis aos existentes entre outras espécies de *Prorocentrum*, especialmente para os locos ITS e LSU. As análises BAPS confirmaram este padrão e revelaram que o PLSC engloba pelo menos três clusters genéticos distintos, correspondendo aos clados *P. lima*, *P. arenarium* e *P.* cf. *lima* "tropical". Além disso, as análises filogeográficas revelaram uma distribuição restrita de haplótipos pertencentes a *P. lima* (clado 4) e *P.* cf. *lima* "tropical" (clado 2). Os haplótipos de *P. lima* (clado 4) são restritos a regiões temperadas e praticamente não há sobreposição geográfica entre este clado e os outros clados do PLSC. Os haplótipos de *P.* cf. *lima* "tropicais" (clado 2) têm distribuição restrita às regiões tropicais do Oceano Atlântico, o que pode indicar um estágio inicial de especiação. Por outro lado, haplótipos correspondentes a *P. arenarium* (clado 3) e *P.* cf. *lima* "tropical" (clado 1) apresentaram ampla distribuição biogeográfica.

Referências bibliográficas:

Aligizaki, K., Nikolaidis, G., Katikou, p., Baxevanis, a. D., Abatzopoulos, t. J., 2009. Potentially toxic epiphytic Prorocentrum (Dinophyceae) species in Greek coastal waters. Harmful Algae**,** 8 (2), 299-311. http://dx.doi.org/10.1016/j.hal.2008.07.002.

Berdalet, E., Fleming, L., Gowen, R., Davidson, K., Hess, P., Backer, L., Moore, S., Hoagland, P., Enevoldsen, H., 2015. Marine harmful algal blooms, human health and wellbeing: challenges and opportunities in the 21st century. Journal of the Marine Biological Association of the United Kingdom. 1-31. http://dx.doi.org/10.1017/s0025315415001733.

Boopathi, T; Faria, D. G., Cheon,J.,; Youn, S. H., Ki, J., 2015. Implications of High Molecular Divergence of Nuclear rRNA and Phylogenetic Structure for the Dinoflagellate Prorocentrum (Dinophyceae, Prorocentrales). Journal Of Eukaryotic Microbiology. 62 (4), 519-531. http://dx.doi.org/10.1111/jeu.12206.

Bravo, I., Fernández, M. L., Ramilo, I., Martźnez, A., 2001. Toxin composition of the toxic dinoflagellate Prorocentrum lima isolated from different locations along the Galician coast (NW Spain). Toxicon. 39 (10), 1537-1545. http://dx.doi.org/10.1016/s0041-0101(01)00126-x.

Ehrenberg, C. G., 1834. Dritter Beitrag zur Erkenntnis großer Organisation in der Richtung des kleinsten Raumes. Physikalische Abhandlungen. Der Königlichen Akademie Der Wissenschaften Zu Berlin., p. 145-336.

Ehrenber, C. G., 1860. Nachtrag zu Hrn. Ehrenbergs Mittheilung vom 8. Verzeichniss der Leuchtthierchen. Monutsberichte der Kb'niglichen Preitssischen Akademie der Wissenschaften zu Berlin: 791-793.

Faust, M. A., 1994. Three new benthic species of *Prorocentrum* (Dinophyceae) from Carrie Bow Cay, Belize: *P. sabulosum* sp. Nov., *P. sculptile* sp. Nov., and *P. arenarium* sp. Nov.. Journal Of Phycology. 30 (4), 755-763. Http://dx.doi.org/10.1111/j.0022-3646.1994.00755.x.

Guiry, M.D., Guiry, G. M., 2022**.** AlgaeBase. World-wide electronic publication, National University of Ireland, Galway**.** https://www.algaebase.org/search/genus/detail/?genus_id=43640&sk=0.

Hoppenrath, M., Chomerat, N., Horiguchi, T., Schweikert, M., Nagahama, Y., Murray, S., 2013. Taxonomy and phylogeny of the benthic Prorocentrum species (Dinophyceae)—A proposal and review. Harmful Algae. 27. 1-28. Http://dx.doi.org/10.1016/j.hal.2013.03.006.

Hu, T.M., Marr, J., de Freitas, A.S.W., Quilliam, M.A., Walter, J.A., Wright, J.L.C., 1992. New diol esters isolated from cultures of the dinoflagellates Prorocentrum lima and Prorocentrum concavum. J. Nat. Prod. 55 (11), 1631–1637. http://dx.doi.org/10.1021/np50089a011.

Murata, M., Shimatani, M., Sugitani, H., Oshima, Y., Yasumoto, T., 1982. Isolation and structural elucidation of the causative toxin of the diarrhetic shellfish poisoning. Nippon Suisan Gakkaishi 48, 549–552. Http://dx.doi.org/10.2331/suisan.48.549.

Nagahama, Y., Fukuyo, Y., 2005. Redescription of *Crypotomonas lima*, collected from Sorrento, Italy, the basioniym fo Prorocentrum lima. Plankton Biol. Ecol. 52 (2), 107–109.

Nagahama, Y., Murray, S., Tomaru, A., Fukuyo, Y., 2011. Species boundaries in the toxic dinoflagellate Prorocentrum lima (Dinophyceae, Prorocentrales), based on morphological and phylogenetic characters. J. Phycol. 47 (1), 178–189. Http://dx.doi.org/10.1111/j.1529-8817.2010.00939.x.

Nascimento, S. M., Salgueiro, F., Menezes, M., de Andréa Oliveira, F., Magalhães, V. C. P., De Paula, J. C., Morris, S., 2016. Prorocentrum lima from the South Atlantic: morphological, molecular and toxicological characterization. Harmful Algae 57, 39–48. Http://dx.doi.org/10.1016/j.hal.2016.05.006.

Nascimento, S. M., Mendes, M. C. Q., Menezes, M., Rodríguez, F., Alves-de-Souza, C., Branco, S., Riobó, P., Franco, J., Nunes, J. M. C., Huk, M., Morris, S., Fraga, S., 2017. Morphology and phylogeny of Prorocentrum caipirignum sp. nov. (Dinophyceae), a new tropical toxic benthic dinoflagellate. Harmful Algae 70, 73–89. Http://dx.doi.org/10.1016/j.hal.2017.11.001.

Nishimura, T., Uchida, H., Noguchi, R., Oikawa, H., Suzuki, T., Funaki, H., Ihara, C., Hagino, K., Arimitsu, S., Tanii, Y., Abe, S., Hashimoto, K., Mimura, K., Tanaka, K., Yanagida, I., Adachi, M., 2020. Abundance of the benthic dinoflagellate Prorocentrum and the diversity, distribution, and diarrhetic shellfish toxin production of Prorocentrum lima complex and P. caipirignum in Japan. Harmful Algae. 96. 101687. Http://dx.doi.org/10.1016/j.hal.2019.101687.

Stern, R., Andersen, R., Jameson, I., Küpper, F., Coffroth, M. A., Vaulot, D. Le Gall, F., Véron, B., Brand, J., Skelton, H., Kasai, F., Lilly, E., Keeling, P., 2012. Evaluating the Ribosomal Internal Transcribed Spacer (ITS) as a Candidate Dinoflagellate Barcode Marker. PloS one. 7 (8), 1-12. Http://dx.doi.org/10.1371/journal.pone.0042780.

Zhang, H., Li, Y., Cen, J., Wang, H., Cui, L., Dong, Y., Lu, S., 2015. Morphotypes of Prorocentrum lima (Dinophyceae) from Hainan Island, South China Sea: morphological and molecular characterization. Phycologia 54 (5), 503–516. Http://dx.doi.org/10.2216/15-8.1.