EUROPEAN JOURNAL OF PHYCOLOGY
2024, VOL. 59, NO. 1, 94–111
https://doi.org/10.1080/09670262.2023.2249073

British Phycological Society
Understanding and using algae
Taylor & Francis
Taylor & Francis Group

OPEN ACCESS Check for updates

Genetic and phenotypic intra-species diversity of alga *Tisochrysis lutea* reveals original genetic structure and domestication potential

Grégory Carrier[a], Jérémy Berthelier[b], Agathe Maupetit[a], Elodie Nicolau[a], Martial Marbouty[c], Nathalie Schreiber[a], Aurélie Charrier[a], Clémence Carcopino[d], Laura Leroi[d] and Bruno Saint-Jean[a]

[a]IFREMER, Unité Physiologie et Toxines des Microalgues Toxine, PHYTOX, Nantes, F-44311, France; [b]Plant Epigenetics Unit, Okinawa Institute of Science and Technology Graduate University (OIST), Okinawa, 904-0495, Japan; [c]Institut Pasteur, Unité Régulation Spatiale des Génomes, UMR 3525, CNRS, Paris, F-75015, France; [d]IFREMER, Service de Bioinformatique de l'Ifremer, SeBiMER, Plouzané, F-29280, France

ABSTRACT
Oceanic phytoplankton species are generally composed of many strains, with intra-species diversity consisting of genetic and phenotypic variability. Despite its importance in ecological and biotechnological contexts, this intra-species diversity and variation among strains has been little studied. We investigated the intra-species diversity of the microalga *Tisochrysis lutea*, a haptophyte of the Isochrysidales order. Inter-strain diversity of *T. lutea* was studied because of the economic importance of the species as a feed in aquaculture and for antioxidant metabolite production, particularly fucoxanthin and other carotenoids, which have health benefits. We analysed Tara Ocean datasets which revealed that *T. lutea* was present in the Pacific, Atlantic and Indian Oceans but not in the Arctic or Austral Oceans. We next made phenotypic and genotypic comparisons of 11 strains of *T. lutea* from worldwide algal collections. All strains were cultivated in the same controlled conditions for one week, and several phenotypic traits were measured, notably antioxidant content. In parallel, the genomes of each strain were sequenced, and genetic variants identified. At the genetic and phenotypic levels, the strains were distinct from each other and our analysis revealed natural trait variations of interest in relation to further exploitation in domestication programmes. A large number of genetic variations were identified among the strains, but no major differences in genome size were observed. Moreover, limited genetic structure was observed among these strains, which could be a consequence of the complex life history of species within the Isochrysidales. Our study provides new knowledge on the intra-species diversity that should be considered in future environmental studies and breeding programmes.

HIGHLIGHTS
- *Tisochrysis lutea* is found in many parts of the world's oceans.
- *T. lutea* has high inter-strain phenotypic and genomic variation.
- Genetic structure of strains from culture collections is limited.

ARTICLE HISTORY Received 14 November 2022; Revised 3 August 2023; Accepted 12 August 2023

KEYWORDS Algae; antioxidant molecules; genetic structure; Haptophyta; intra-species diversity; *Tisochrysis lutea*

Introduction

Photosynthetic aquatic algae or cyanobacteria that inhabit the planktonic region of freshwater or marine bodies play key roles in climatic and geochemical processes and in biodiversity (Arrigo, 2005). For instance, phytoplankton contribute 45% of atmospheric oxygen and reduce carbon dioxide concentrations through the biological carbon pump mechanism (Geider *et al.*, 2001). Phytoplankton also support biodiversity as an essential primary food in ocean ecosystems and the first link in the food web of many marine ecosystems (Field *et al.*, 1998). Recently, with the help of oceanic expeditions, the diversity of phytoplankton was inventoried at around 35 000 prokaryotic operational taxonomic units (OTUs) and 150 000 eukaryotic OTUs (Bork *et al.*, 2015; Sunagawa *et al.*, 2020). Among these, at least 30% of the eukaryotic species are unknown (Vargas *et al.*, 2015).

Within phytoplankton diversity, each species generally consists of various strains or sub-populations which display phenotypic and genetic differences (Simon *et al.*, 2009). In the present study, we examined the intra-species diversity of the marine microalga *Tisochrysis lutea*. This species is a member of the Haptophyta. Haptophyta is an ecologically dominant phylum in the planktonic photic realm (Bendif *et al.*, 2013). Data gathered on the Tara Oceans expeditions indicate that haptophytes are the second most abundant group of eukaryotic phytoplankton after diatoms (Penot *et al.*, 2022; Pierella Karlusich *et al.*, 2023). *T. lutea* belongs to Isochrysidales, an order which represents approximately 2% of the total relative abundance of haptophytes (Penot *et al.*, 2022; Pierella Karlusich *et al.*, 2023). The Isochrysidales are divided into three families: Isochrysidaceae, Noelaerhabdaceae and Prinsiaceae. *T. lutea* belongs to the Isochrysidaceae, which is characterized by non-calcifying organisms with

CONTACT Grégory Carrier Grégory.Carrier@ifremer.fr

© 2023 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.
This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. The terms on which this article has been published allow the posting of the Accepted Manuscript in a repository by the author(s) or with their consent.

Published online 29 Sep 2023

an as yet undescribed life history (Edvardsen *et al.*, 2000; Bendif *et al.*, 2013).

Historically, *T. lutea* has been studied due to its widespread use in aquaculture as a feed for shellfish, oysters and shrimps (Marchetti *et al.*, 2012; Ippoliti *et al.*, 2016). In the last decade, *T. lutea* has become a microalgal model because it produces high quantities of fucoxanthin, one of the most economically valuable carotenoids due to its health benefits (Gao *et al.*, 2020; Mohamadnia *et al.*, 2022), and docosahexaenoic acid (DHA), an important lipid for healthy foetal development (Premaratne *et al.*, 2021; Thurn *et al.*, 2022). Because of these biotechnological interests, considerable genomic knowledge has also already been accumulated for this species (Garnier *et al.*, 2014; Berthelier *et al.*, 2018; Carrier *et al.*, 2018; Hernández Javier *et al.*, 2018). However, genomic studies have mainly focused on a single strain *Tisochrysis lutea* CCAP 927/14, and undertaken in aquaculture rather than wild populations (Borowitzka, 1997; Bougaran *et al.*, 2003). Several other strains of this species were harvested and conserved in culture collections (Bendif *et al.*, 2013), but information on their origins is limited, and analysis of inter-strain diversity is scarce. As with breeding programmes for crop species, domestication programmes have been conducted to select the domesticated strains of microalgae with the best traits (Bougaran *et al.*, 2012; Almutairi, 2020; Gachelin *et al.*, 2021). However, inter-strain diversity, life history and breeding method all need to be understood to properly conduct domestication programmes, necessitating further research (Larkum *et al.*, 2012; Rumin *et al.*, 2020; Chen *et al.*, 2022). Only a few studies of intra-species diversity in microalgae have been reported in the literature. High inter-strain diversity and strong genetic structure have been reported for the diatoms *Ditylum brightwellii* (Rynearson & Virginia Armbrust, 2004), *Phaeodactylum tricornutum* (Rastogi *et al.*, 2020) and *Pseudo-nitzschia pungens* (Casteleyn *et al.*, 2010), the dinophyte *Alexandrium* sp. (Le Gac *et al.*, 2016; Paredes *et al.*, 2019), and for the chlorophytes *Ostreococcus* species (Rodríguez *et al.*, 2005) and *Chlamydomonas reinhardtii* (Gallaher *et al.*, 2015). Within the Haptophyta, only the inter-strain diversity of *E. huxleyi* has been studied and contrary to other microalgal species, only weak genetic stratification was observed in the studied populations (Read *et al.*, 2013; Filatov, 2019). Deciphering the intra-species diversity of *T. lutea* is necessary to optimize domestication programmes by identifying the strains with the most suitable traits for selection or improvement.

In this study, the main objective was to evaluate the range of genetic and phenotypic variation among the strains currently available in culture collections worldwide. First, we investigated the distribution of *T. lutea* in the world's oceans using data from the Tara Ocean campaigns (de Vargas *et al.*, 2015). Knowledge about the origin, stratification, phenotypic and genetic properties of strains of this species was limited. Ideally, this study should have been carried out on strains directly isolated from their respective biomes. However, such an approach would require substantial resources. We, therefore, performed this study using *T. lutea* strains that had been isolated from various marine environments and maintained in algal collections for decades (Table 1). In this study, 11 strains were compared phenotypically, focusing on their pigment and lipid profiles. The genomes of each *T. lutea* strain were also sequenced to investigate the genetic variations, genome organization and genetic structure of this species. The collected data have been summarized and made available in an easy-to-use genome browser for the scientific community (https://genomes-catalogue.ifremer.fr).

Material and methods

Distribution of Tisochrysis lutea *in the world's oceans*

To estimate the distribution of *T. lutea* in the oceans, we used Tara Ocean data from the Ocean Atlas (https://tara-oceans.mio.osupytheas.fr/). Operational Taxonomic Units (OTUs) corresponding to the Isochrysidaceae family were collected from the Ocean Barcode Atlas (Vernette *et al.*, 2021) eukaryote databases (18SVv9 vV2) with a similarity search within 99%. The OTU CCMP1244 was closer to *T. lutea* (87% identity) than another *Isochrysis* species known (77% identity for the second closest). During

Table 1. Probable origins of the strains used in the experiment.

Tisochrysis lutea strains	Probable origin	Probable year of isolation	Orders in culture collections	Preservation method
CCAP 927/14	Tahiti archipelago	1977 by Haines	CCAP	Cryopreserved since 2003
CCMP 463	Little Water Cay	1984 by Glazer	CCMP	Cryopreserved since 1999
RCC 179	Unknown	Unknown	RCC	maintained in culture
RCC 1344	Atlantic coast;	Probert	RCC	maintained in culture
RCC 3691	Tahiti archipelago	1978 by Martin	RCC	maintained in culture
RCC 3692	Tahiti archipelago	1978 by Martin	RCC	maintained in culture
RCC 3693	Tahiti archipelago	1978 by Martin	RCC	maintained in culture
RCC 3699	North Sea	1962 Amundsen	RCC	maintained in culture
NIVA 4–91	North Sea	1989 University of Oslo	NIVA	maintained in culture
IFMG	near to Kategate, North Sea	Unknown	NORCCA	maintained in culture
Argenton	near to Argenton, Atlantic cost	1998	IFREMER	maintained in culture

the Tara Ocean campaign, sample separations were made according to the organism size. To retrieve *T. lutea*, the size fraction selected was around 5 µm, corresponding to the size of this species (Bendif *et al.*, 2013). The extracted data (OTU, origin locations and associated temperature) are given in Supplementary table S1.

Microalga strains and culture conditions

From algal collections worldwide, 18 strains named *T. lutea* or *Isochrysis* sp. were ordered and cultivated to sequence the 18S gene. Based on 18S results, 11 strains were confirmed to be *T. lutea* and became the strains studied in the present paper: CCAP927-14, CCMP463, RCC179, RCC1344, RCC3691, RCC3692, RCC3693, RCC3699, NIVA4-91, IMFG and Argenton (Table 1). All these strains were grown in 200 ml flasks of natural seawater enriched with sterile Conway–Walne medium (Walne, 1966) reset every 20 days. All cultures were maintained with bubbled 0.22 mm filtered air, at a constant temperature of 21°C, under a constant light irradiance of 50 µmol photons m^{-2} s^{-1}. To phenotype the strains, 300 000 cells ml^{-1} of each were inoculated in photobioreactors (Berard *et al.*, 2021) containing 150 ml natural seawater enriched with Conway–Walne medium with a modified nitrogen content (150 µM). Light irradiance (100 µmol photons m^{-2} s^{-1}), temperature (21°C) and pH (8.2) were controlled to keep culture conditions identical. Samples of all strains were harvested after 7 days of culture to be phenotyped (pigment and lipid measurements). After exhaustion of the nitrogen, which occurred after 5 days, the growth of the strain was stopped for all strains. The uptake of all nitrogen was confirmed by analysis of the culture medium. A 2 ml sample was collected from each culture, microalgae were eliminated by filtration (0.2 µm, Minisart, Sartorius), and the culture medium then analysed by spectroscopy at 220 nm as described in Collos *et al.* (1999).

Phenotype measurements

After 7 days of culture, samples of each strain were harvested. Common culture traits were measured: growth rate, cell size and amount of carbon per cell. The growth of microalgae was automatically monitored by optical density measurements. The formula used to calculate the growth rate was $\mu = (\mathrm{Ln}(X_2)\text{-}\mathrm{Ln}(X_1))/(t_2 - t_1)$. To measure the carbon content in each microalgal cell, approximately 100 million algal cells were filtered on a 25 mm GF/C fibreglass filter (Whatman). The filters were deposited in a LIMP glass, placed in a steam room and dried at 75°C for 24 h, analyses were performed with a carbon elemental analyser (Thermoelectron) with methionine, aspartic acid and nicotinamide standards for calibration. Cell size and number were assessed using a Coulter Counter Multisizer 3 (Beckman Coulter, High Wycombe, UK). Before measurement, samples were diluted to 1/50 with sterile seawater. Cell size, given as equivalent sphere diameter, was then calculated using MS-Multisizer 3 software (Beckman Coulter, High Wycombe, UK). From the perspective of biotechnological applications, we focused on the antioxidant capacity of these strains and several lipid and pigment molecules of potential biotechnological interest were measured. For the lipid profiles, samples were harvested at 7 days of culture for all strains. Approximately 150 million algae cells were filtered on a 25 mm GF/F fibreglass filter (Whatman) and suspended in 6 ml Folch solution (chloroform: methanol 2:1) as described in Folch *et al.* (1957). Lipid class separation was performed by column chromatography according to the method of Soudant *et al.* (1995), with a BPX-70 capillary column (60 m long, 0.25 mm internal diameter, 0.25 µm film thickness; SGE, Austin, Texas, USA) containing a polar stationary phase (cyanopropyl-siloxane). The upper organic phases containing fatty acid methyl esters (FAMEs) were collected and assayed by gas chromatography coupled with flame-ionization detection (GC-FID). FAME quantification was compared with the C17 internal standard (Sigma-Aldrich, St. Louis, Missouri, USA) by GC-FID using a gas chromatograph (Autosystem Gas Chromatograph; Perkin-Elmer, Waltham, Massachusetts, USA). Total fatty acids (TFAs) were calculated as the sum of saturated fatty acids (SFAs), polyunsaturated fatty acids (PUFAs), including in particular docosahexaenoic acid (DHA), and monounsaturated fatty acids (MUFAs).

For pigment analysis, samples were harvested after 7 days of culture for all strains. Approximately 80 million algal cells in 8 ml were filtered on a 25 mm GF/F fibreglass filter (Whatman) and immediately frozen at −80°C. Extraction was realized with 2 ml of cold acetone containing 5% of water with vitamin E at 2.5 ng l^{-1} as the internal standard. The solution was vortexed and sonicated for 10 min and macerated for 24 h at −80°C, then filtered on a 0.2 µm PTFE filter (Phenomenex, France). The extract was analysed by HPLC (Agilent LC 1200) with a DAD detector at 436 and 450 nm. Chromatographic conditions were as described in Pajot *et al.* (2023). Quantification was performed using external calibration against the pigment standards Chlorophyll c2 (Chl c2); Fucoxanthin (Fx); Chlorophyll a (Chl a); Diadinoxanthin (Dd); Diatoxanthin (Dt); Zeaxanthin (Zx); Echinenone (Echin); Violaxanthin (Vx); Phaeophytin a (Pheo a); β-carotene (β-Car). Chlorophyll c1 (Chl c1) was quantified with Chl c2 standard. 3-hydroxyechinenone (HEchin) was quantified with Echin standard. Cis-fucoxanthin (Cis-Fuco) was quantified with Fx standard. All standards were purchased from DHI Lab

products, Denmark. To measure antioxidant capacity, samples were harvested at 7 days of culture for all strains, with approximately 20 million algae cells in 2 ml. The microalgal culture was mixed with 100 µl dihydrorhodamine 123 (DHR 123) for 2 min following the protocol of Yazdani (2015). Measurements were taken on an Amnis ImageStreamX Mk II Imaging Flow Cytometer following standard recommendations. A Spearman correlation matrix was realized between phenotypic traits measured for all strains using R software (version 4.0.4). A heat map showing the correlation matrix was then constructed. The coefficient of correlation was measured and the correlation considered significant at a p value threshold of 0.05.

Reference genome assembly using Hi-C technology

Total DNA of the CCAP927/14 strain was extracted using a phenol-chloroform protocol (Hu *et al.*, 2004). DNA quality and concentration were assessed with gel electrophoresis and Qubit fluorometric quantification. Hi-C libraries were constructed following the protocol described in Baudry *et al.* (2020). Briefly, DNA was digested using the GC-neutral restriction enzyme DpnII by overnight incubation at 37°C. Digested DNA was labelled with biotin. DNA fragments with biotin and blunt-end ligation were selected. The Hi-C libraries were sequenced with an Illumina NextSeq 550 system (2 × 35 bp, paired-end), generating approximately 2.7 Gb of data. Scaffolding was performed using the latest version of instaGRAAL (https://github.com/koszullab/instaGRAAL), an algorithm that uses chromosome conformation data to re-scaffold contigs and improve genome assembly (Baudry *et al.*, 2020). The original genome was first split into sequences digested *in silico* with DpnII. Then, the reads from the Hi-C library were mapped on these digested fragments to compute an initial contact matrix. InstaGRAAL reorders the fragments digested following the contact matrix and standard polymer physics model. To obtain the best genome model, 100 iterations were performed.

Reference genome annotation

The coding gene region was annotated with the MAKER2 pipeline (Holt & Yandell, 2011). This pipeline combines several approaches and finds a consensus from the results to obtain the best annotation possible. Genes were identified based on proteomic data, transcriptomic data (Garnier *et al.*, 2014), and gene prediction with AUGUSTUS software trained on *T. lutea* (Stanke *et al.*, 2006) and SNAP software using *Arabidopsis thaliana* model genes (Korf, 2004).

The annotation of repeated elements, particularly transposable elements, was made by submitting a previously obtained TE library to TEannot, implemented in the PiRATE pipeline (Berthelier *et al.*, 2018). The annotations of 88 transcription factors were improved by adding previously curated annotation data (Thiriet-Rupert *et al.*, 2018) (88 genes). Chloroplast and mitochondrial genomes had been characterized in a previous study (Méndez-Leyva *et al.*, 2019) and similar results were found here.

Genome sequencing

Microalgal cultures were treated with antibiotics (Sigma No. A5955) before harvesting to minimize bacterial contamination. Microalgal harvesting was realized from a new culture that was grown for 6 days. Total DNA was extracted using a phenol-chloroform method as previously described by Hu *et al.* (2004). Illumina sequencing was performed at the GeT-PlaGe INRAE platform using an Illumina HiSeq3000. Libraries were built following the Illumina TrueSeq protocol. Paired-end sequencing was done (2 × 150 bases) using a bar code corresponding to each strain. On average, 7 Gbp were obtained for each strain (36 million reads for a length of 150 bases). The raw sequencing data obtained were filtered and trimmed to conserve only reads of sufficient quality. TrimGalore was used on the raw data to eliminate Illumina residual adapters and artefact sequences and only conserve paired reads with quality scores higher than Q30 for 150 bases. Oxford Nanopore Technology sequencing was performed with a MinION-101B with R9.4 flow cells and SQK-LSK108 library kits. Guppy v3.1 was used for the basecalling. On average, 1.1 Gbp were obtained for each strain. Porechop (https://github.com/rrwick/Porechop) and NanoFilt (De Coster *et al.*, 2018) were used to eliminate residual adapters and select sequences above 2000 bases with a quality score higher than Q8.

Polymorphism identification

For each strain, reads were independently mapped on the new reference genome assembly. We obtained an average 58X genome depth for short reads and 10X for long reads (Supplementary table S1). From these mapped data, genetic variations were investigated for each strain using bioinformatics tools. Three types of genetic variation were looked for: (i) short mutations corresponding to single nucleotide polymorphisms (SNPs) and short (10 bases) insertions or deletions (indels), (ii) large mutations related to large indels (>100 bases) or indels caused by transposable elements (TEs), and

(iii) gene presence/absence variation (PAV). A genetic variation detected in one strain means that there are two alleles for one locus in this strain. Each genetic variation identified could be shared among all strains or be specific to only one or some of them. Genetic variations shared by all studied strains were called 'core variations', while those not shared by all strains because they were found in only one or some of the strains were called 'specific variations'.

Detection of SNPs and short indels was performed among all 11 strains. Short reads were mapped independently on the new genome assembly using BWA and Mosaik software (Li & Durbin, 2009; Lee *et al.*, 2013) (see Supplementary data S1). SNPs and short indels were detected with Freebayes and GATK (Garrison & Marth, 2012; Poplin *et al.*, 2018), crossing all data. SNPs and short indels were validated if detected by at least two programs and with a minimum allele frequency of 10%.

Detection of large insertion variations, gene presence/absence and genome size variation between strains required the construction of an anchor genome. An anchor genome assembly was thus made containing all the insertion variants identified in each strain. Identification of large insertions was realized by crossing the detections made by the three programs Sniffle v1.0.11 (Sedlazeck *et al.*, 2018), Mobster v2.41 (Thung *et al.*, 2014) and MindTheGap V2.2 (Rizk *et al.*, 2014) (see Supplementary data S1). Sniffle software uses mapping of long reads on a genome assembly to identify insertion variants. It detects different types of insertions without a priori using evidence from split-read alignments. Long reads were independently mapped on the new reference genome using MiniMap 2.17 (Li, 2018) and ModularAlignement v1.1.1 (Schmidt *et al.*, 2019). Using two alignment methods made it possible to limit mapping biases and obtain a more exhaustive result. Results obtained were crossed for each strain and the redundant deductions were eliminated. MindTheGap software was applied independently with Ilumina short reads of each strain to detect insertions of any size and without a priori. MindTheGap results were pooled, detected insertions were sorted and redundancies eliminated. Mobster detected TE insertions from the Ilumina short-read data and a library of known TEs (Thung *et al.*, 2014). In the same way, Mobster was used independently for each strain and insertions of TEs were pooled, sorted and the redundancies eliminated. Finally, insertion variants identified for each strain from these three programs were pooled, sorted and the redundancies eliminated. VCFtools and VCFlibs (Danecek *et al.*, 2011) were used to create an anchor genome assembly containing all insertions.

Gene presence/absence variants among strains were identified. Reads of each strain were mapped on the previously built anchor genome assembly. Read counting was performed for each strain, considering the presence of a gene as a function of the read number mapped (minimum 50% coverage and 20 reads depth). The presence/absence of large insertions such as TEs among strains was established in the same way. Short reads of each strain were also mapped on the anchor genome. The genome size of each strain was obtained from the mapped reads for each strain on the anchor genome (BAM to FASTA file conversion) if a nucleobase had a minimum of 20 reads depth.

Analysis of the genetic structure of Tisochrysis lutea *strains*

Strain stratification was analysed for each type of genetic variation studied (short, large and gene PAVs) using the sNMF tool with Kmin = 2 and Kmax = 10 (Frichot *et al.*, 2014). A principal component analysis (PCA) was also performed for each type of genetic variation studied, using R software (version 4.0.4). To finish, for each type of genetic variation, a distance tree based on genetic similarity was built using the FastME method (Desper & Gascuel, 2002) with 100 bootstrap values (100 used).

Results

Distribution of Tisochrysis lutea *in the oceans*

The Tara Ocean expeditions data provide an overall idea of the distribution and abundance of phytoplankton in the oceans. These data revealed that *T. lutea* was low in abundance in the oceans, corresponding to around 1% among the Haptophyta. For the Isochrysidales, OTUs from *T. lutea* species and *E. huxleyi* species were of similar abundance: 23% for each species. However, half the species representing this order are still undescribed (Fig. 1b). Moreover, OTUs of *T. lutea* are present in the surface waters of the Atlantic, Pacific and Indian Oceans in surface water, although absent from the Arctic and Austral Oceans (beyond the 60th north meridian and 50th south meridian, approximately) (Fig. 1a). *Tisochrysis lutea* was found to inhabit regions with an average temperature of 21.2 ± 5.8°C (SE) and mean salinity of 35.9 ± 4.8 g l^{-1} (SE) (Supplementary table S1).

Collection of Tisochrysis lutea *strains*

Given that it was not possible to isolate strains directly from the world's oceans (Table 1), to build the most complete collection of *T. lutea* strains possible, we investigated and collected 18 strains that could correspond to *T. lutea* or *Isochrysis* sp. from algal collections available worldwide (Table 1). We confirmed that 11 of the 18 strains that were thought to be *T. lutea* were this

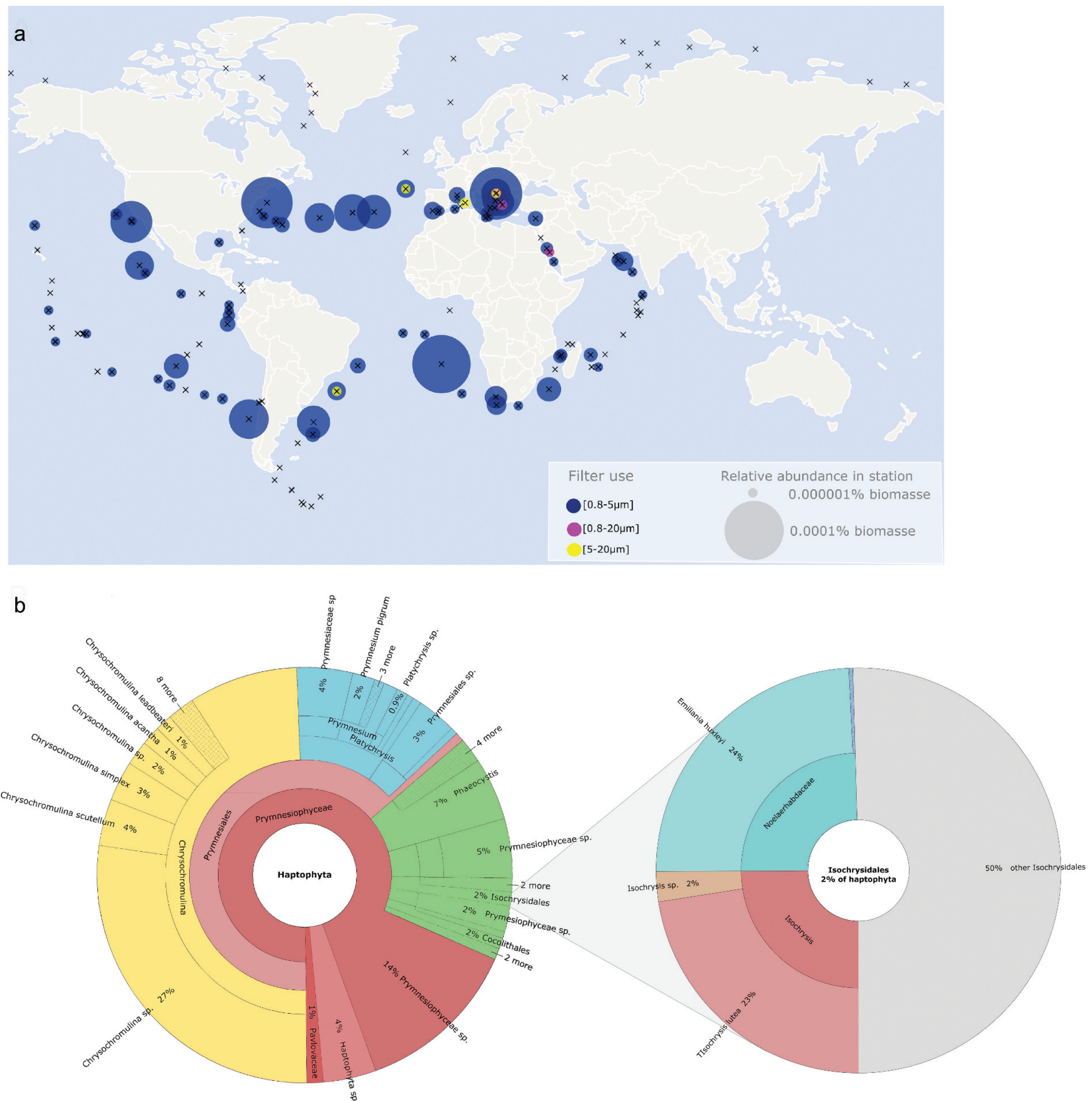

Fig. 1. Analysis of *Tisochrysis lutea* from Tara Ocean data. (A) Distribution and relative abundance of *T. lutea* sampled during the expeditions. (B) Relative species abundance of Haptophyta from Tara Ocean expeditions.

species by analysing their 18S gene sequences (Supplementary fig. S1). The 11 cultivated strains came from diverse geographic locations (Table 1), although their origins were only approximate because they were isolated a long time ago and unfortunately were poorly documented at the time (Table 1). These strains had also been maintained in laboratory cultures conditions or cryopreserved since the 2000s (Table 1).

Phenotype comparisons of Tisochrysis lutea *strains*

We first evaluated the intra-species variation of the 11 *T. lutea* strains at the phenotype level. We found a 1.3× fold change in growth rate, a 1.2× fold change in cell size and 1.8× fold change in the amount of carbon (Fig. 2a–c). The strain with the best growth rate was NIVA4-91 (μ = 0.83 day^{-1} in these culture conditions). For the antioxidant traits measured, the CCAP927-14 strain had the highest content of antioxidant molecules (DHR123 score, Fig. 2d) and pigment molecules (Fig. 2j–q), especially carotenoids (fucoxanthin, echinenone, diadinoxanthin or β carotene). Moreover, a comparison of the measured antioxidant response to nitrogen limitation conditions indicated high variation (2.0× fold) (Fig. 2d). Carotenoids with antioxidant properties, such as echinenone (1.5× fold) and fucoxanthin (1.3× fold), docosahexaenoic acid lipid (1.4× fold) and polyunsaturated fatty acids (1.6× fold) also displayed high variation. To summarize, none of the strains had similar phenotypic profiles for any of the phenotypic traits measured. In the principal component analysis (PCA) from pigment and lipid profiles, the variation

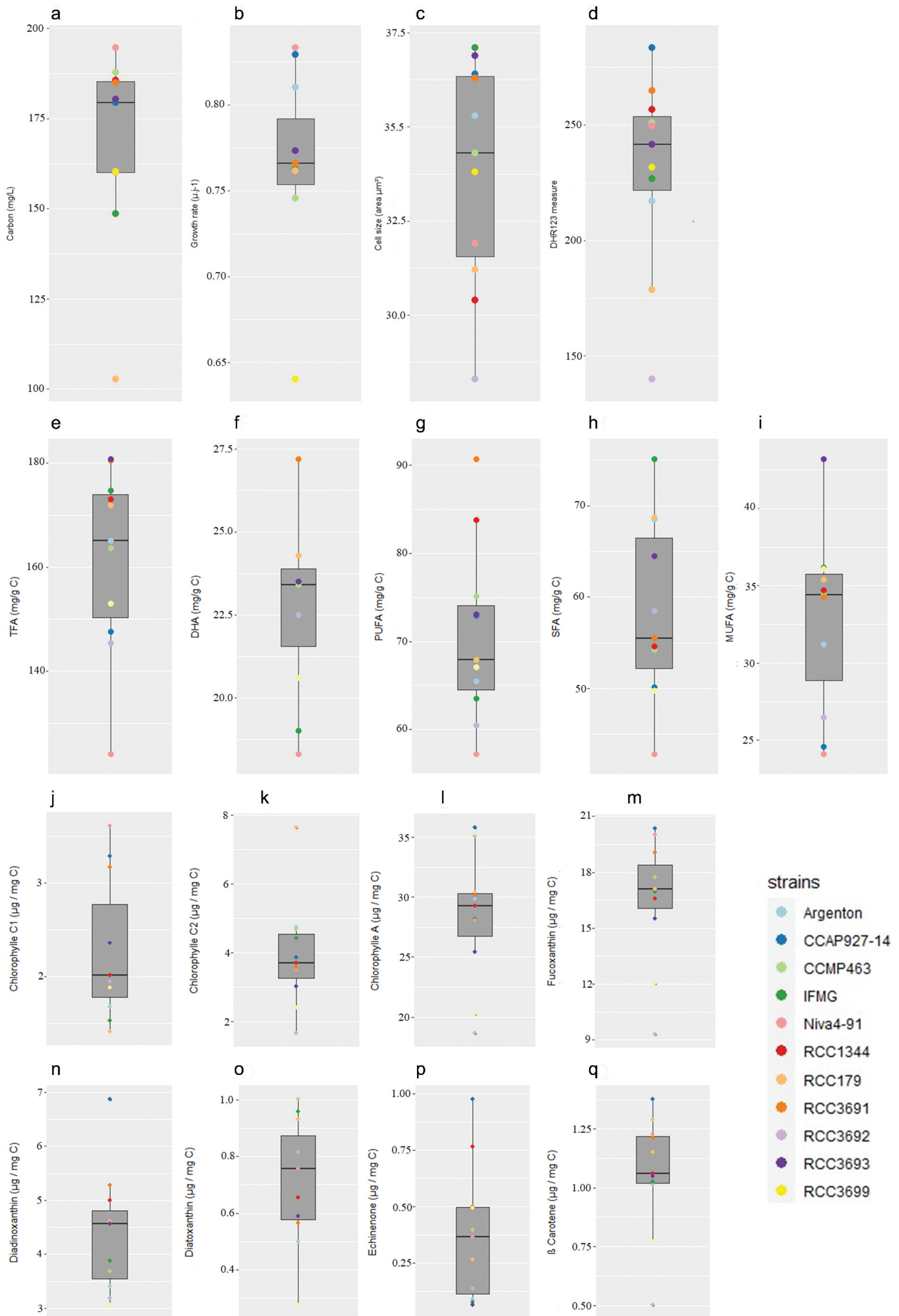

Fig. 2. Boxplots for each phenotypic trait measured on all the strains analysed.

explained by the first dimension was greater than 50% and revealed significant variation among strains (Fig. 3). All strains were well dispersed and no clusters were observed, supporting their heterogeneity. Regarding the lipid PCA analysis, the SFAs (saturated fatty acids) and PUFAs (polyunsaturated fatty acids) had opposite profiles, highlighting that strains containing high SFA amounts contained less PUFA and vice versa (Fig. 3a). For the pigment PCA (Fig. 3b), diadinoxanthin and diatoxanthin had opposite trends, supporting the known xanthophyll cycle of algae which consists of the de-epoxidation of diadinoxanthin to diatoxanthin (Vershinin & Kamnev, 1996). Chlorophyll c1 and c2 also displayed opposing trends in PCA that might suggest a putative cycle between them. In the pairwise correlations, the DHR123 measurements were generally significantly correlated with traits recognized to have an antioxidant function (Dd, β-car, Fuco and PUFA lipids) (Fig. 4). Chlorophyll *c1* and echinenone were also found to be correlated, suggesting that they may both have an antioxidant function.

Improvement of the Tisochrysis lutea *reference genome*

The new reference genome assembly of *T. lutea* presented in this study is 82.5 Mbp in length and comprised 55 supercontigs. The N50 is 3.0 Mb and completion of eukaryote core genes with BUSCO gave 255/303 completed genes (Supplementary fig. S3). Among the 55 contigs, 28 corresponded to 97% of the size of the genome and 27 contigs of small size (less 1 Mb) corresponded to 3% of the size of the genome. The latter were probably repeated regions of the genome that were difficult to sequence (Supplementary fig. S2). The new reference genome was annotated to identify regions corresponding to genes and repeats. Overall, 45% of the genome corresponded to genes and 33% to repeated DNA, among which 21% corresponded to TEs. The number of genes is 19 913 and the number of probable isogenes is 38 930. All these data are available on a genome browser at https://genomes-catalogue.ifremer.fr/.

Genome size variations

Genomes of the 11 strains were sequenced and we first examined how genome size varied among them (Supplementary table S1). The core genome, consisting of common genetic regions of these *T. lutea* strains, was estimated to be 79.6 Mb. The dispensable genome, consisting of genetic regions presented in only one or several *T. lutea* strains but not in all, was estimated to be 1.5 Mb on average for each strain, corresponding to 1.8% of the whole genome size. The smallest genome belongs to RCC3699 (80 Mbp) and the two largest genomes to CCAP927-14 and RCC3692 (82.5 Mbp).

Short variations

Among the different types of genetic variations investigated, 172 579 short polymorphisms (126 627 SNPs and 46 476 short indels) were identified among the strains of *T. lutea*. Across all strains, SNPs revealed a relatively balanced ratio of transitions (Ts) to transversions (Tv) (Ts/Tv = 0.86). The core short variations among these 11 strains numbered 36 758 (Fig. 5 SNPS-1). Conversely, 103 103 short variations were considered strain-specific, meaning they were found to be present in one or several strains but not all. On average, each strain was found to contain 7364 specific short variations (standard deviation: 1.84) in addition to the 36 758 core short variations. Specific short variations of particular strains were variable; RCC3693 had the least (4592) and CCAP927-24 had the most (11 332). Moreover, 61.9% of specific short variations were specific to one strain and 16.0% were present in only two. To confirm this distribution, Tajima's D score (Nielsen, 2001) was measured and an average score of −2.19 was obtained (Supplementary table S1). This score confirmed that this population of strains has an excess of rare alleles.

Transposable element variations

Among the large indels detected in these strains, a large majority (96.5%) were generated by TEs that had been previously identified (Berthelier *et al.*, 2018). Hat2_Ace was the TE family that induced the most variations among strains (11.5%) followed by Gypsy1 (10.4%) and Harbinger5 (8.0%) (Supplementary table S1). Concerning the distribution among strains, 6893 variations of TEs were shared among all strains and 13 542 were strain specific (Fig. 5 TE-1). On average, 1231 specific TE variations were present in each genome of each strain, in addition to core TE variations. We identified twice the number of specific insertion variations related to TEs in RCC3692 and CCAP927-14 than in the other strains (Fig. 5. TE-1). Moreover, the specific insertion variations of TEs detected in these two strains were not specific to only one or two TE families (Supplementary table S1). The larger number was not related to a burst of a specific TE family but to an overall increase of TE copies.

Gene presence/absence variations

The core genes, consisting of genes shared by all the *T. lutea* strains studied, numbered 19 494. Gene

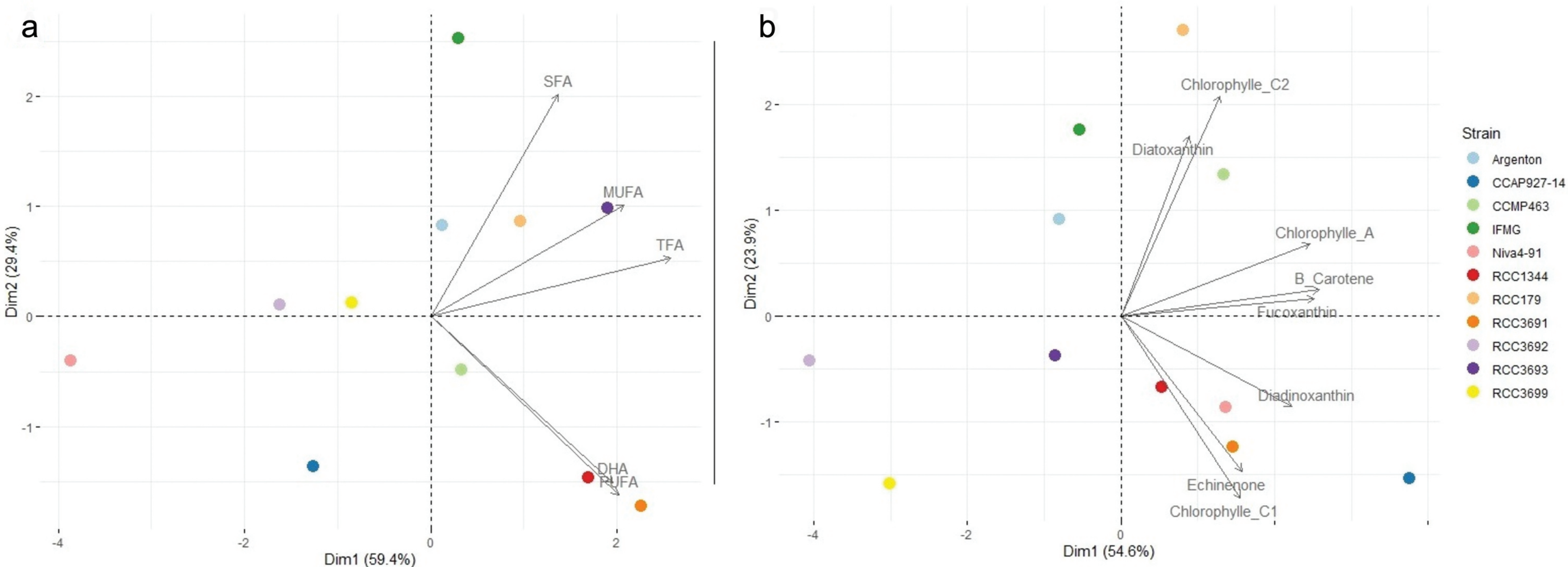

Fig. 3. Principal component analysis for all *Tisochrysis lutea* strains for phenotypic traits of (A) lipids and (B) pigments. Each trait measured was normalized by cellular carbon content. TFA: total fatty acids, DHA: docosahexaenoic acid, PUFA: polyunsaturated fatty acids, SFA: saturated fatty acids, MUFA: monounsaturated fatty acids.

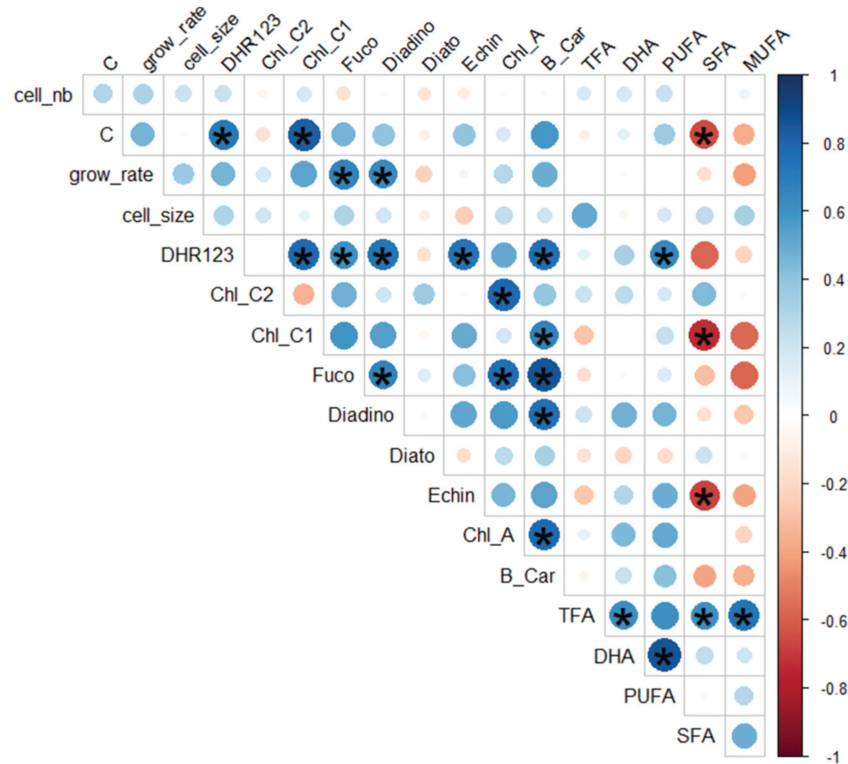

Fig. 4. Spearman correlation matrix between phenotypic traits measured for all strains. The colour scale indicates the correlation score: red indicates a negative correlation, blue shows a positive correlation, and white indicates an absence of correlation. A star within a circle indicates a significant correlation (p value > 0.05). cell_nb: cell count, C: carbon per cell, cell_size: the size of cell, DHR123: dihydrorhodamine 123, chl_c2: chlorophyll *c2*, chl_c1: chlorophyll *c1*, Fuco: fucoxanthin, Diadino: diadinoxanthin, Diato: diatoxanthin, Echin: echinenone, Chl_A: chlorophyll *a*, B_Car: β carotene, TFA: total fatty acids, DHA: docosahexaenoic acid, PUFA: polyunsaturated fatty acids, SFA: saturated fatty acids, MUFA: monounsaturated fatty acids.

presence or absence variation (PAV) among the strains is shown in Fig. 5 (PAV-1). An average of 334 specific genes were detected in the genome of each strain. Strains IFMG and RCC1344 contained the least specific genes (115) and strain RCC3692 had the most (634). In total, 1078 genes were specific or rare and, among these, 325 (30.1%) were unique to a single strain while 753 (69.9%) were present in two or more strains.

Genetic structure of the strains

Genetic structure of the studied strains analysed for each variation type (short, large and gene variations) showed no difference among strains (Supplementary fig. S4). A principal component analysis (PCA) was also performed to assess the genetic stratification for each variation type (Fig. 5, SNP-TE-PAV-2.). Overall, the majority strains clustered together and variations explained by the first dimension of the PCA was lowest (35%), confirming the weak genetic structure observed for these strains. However, for the short and large variations, the two strains CCAP927-14 and RCC3692 were outside of the main group of strains. Concerning gene variation, only RCC3692 was outside of the other strains. To finish, for each type of variation, a distance tree of genetic similarity was built to examine the relations among the strains with a different method (Fig. 5, SNP-TE-PAV-3). Consistently with the PCA, most of the strains were clustered together. However, CCAP927-1 was positioned outside of the main group of strains for short (SNPs) and large (TEs) genetic variations (Fig. 5, SNP-TE-3). In addition, the strain RCC3692 was also outside the main group of strains for gene presence/absence variations (Fig. 5, TE-PAV-3).

The links between genes and phenotype variation

Among the short mutations, 94.8% were found to be present in intergenic regions and were predicted not to significantly affect the function of any genes, 3.4% were located in coding regions but predicted not to affect translation, 1.6% may have an impact on amino acids but with a low impact on protein formation, and 0.2% (161) might have a high impact on proteins (to generate a stop codon or modify the start or stop codon). Among potentially affected proteins, we found only 11 (7%) to have a homologous protein with a known putative protein domain, but none with

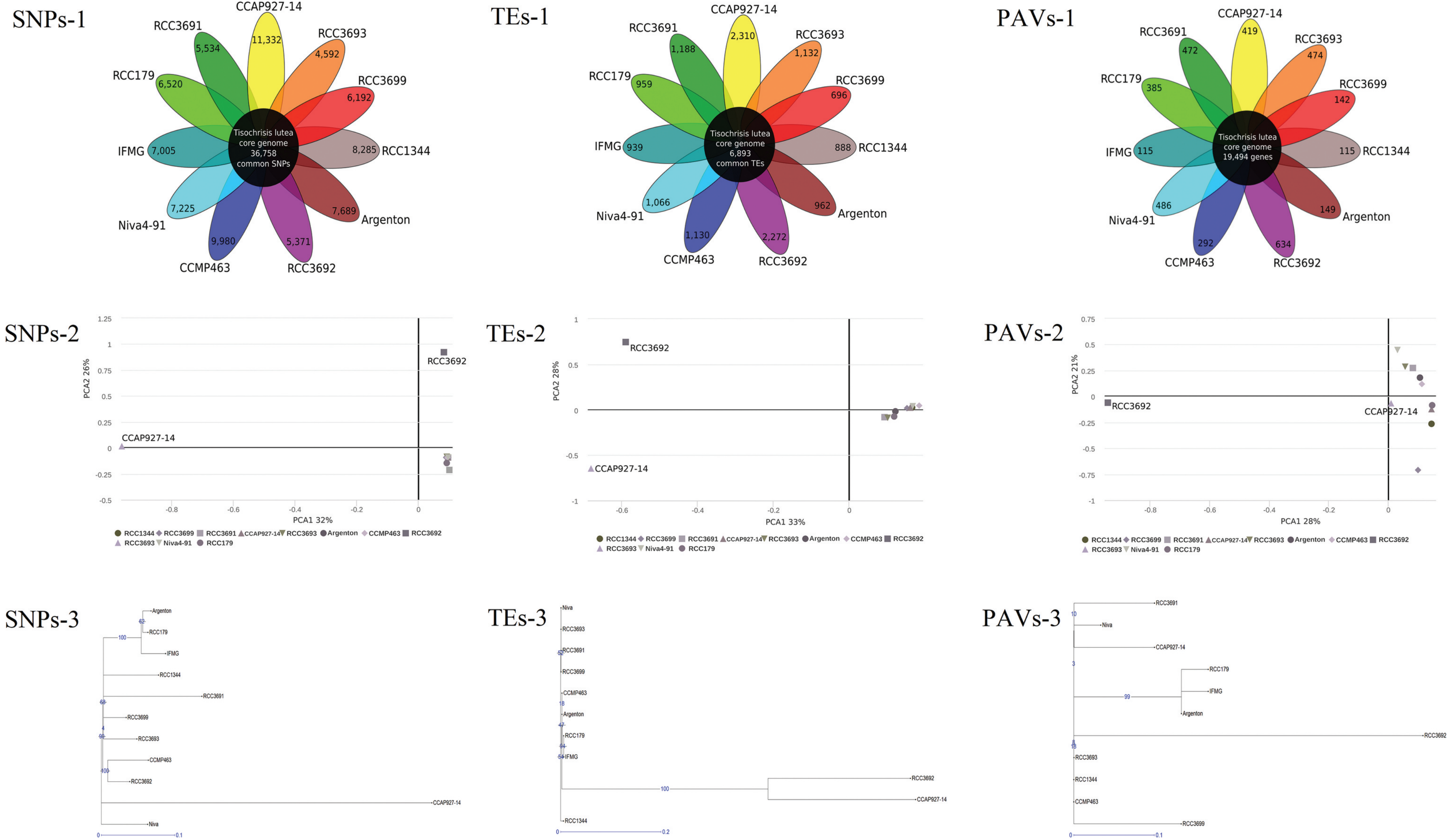

Fig. 5. The 'SNPs' sub-figures present results on short variations, 'TEs' on transposable element variations and 'PAVs' on gene presence/absence variations. Sub-figures SNPs-1, TEs-1 and PAVs-1 show the different types of genetic variations in more detail, with values in centres of the 'flowers' giving the numbers of these variations shared by all strains and values in the petals of the 'flowers' giving the strain-specific numbers. Figures SNPs-2, TEs-2 and PAVs-2 are principal component analyses (PCA) on the variations found in the strains. Figures SNPs-3, TEs-3 and PAVs-3 are trees of similarity among strains built from variation data using FastME (Desper & Gascuel, 2002). The numbers in blue correspond to the scores of bootstrap values (100 used).

a clear role in pigment or lipid metabolism. Overall, 453 genes were potentially affected by TE insertions. Among these, 117 (25.8%) have a homologous protein. Considering PAV analysis among strains, of the 1078 polymorphic genes, only 92 (8.5%) have a homologous protein. Nevertheless, none of these genes containing polymorphisms potentially affecting the function of translated proteins have a protein domain or a homologous protein sufficiently characterized in *T. lutea* to be clearly linked to lipid or pigment metabolism.

Discussion

Over the last decade, worldwide oceanic expeditions, such as Tara Ocean, have enabled the first global sampling of the marine microbial world, opening a new era for the inventory of phytoplanktonic diversity (Bork *et al*., 2015). Barcoding approaches with a part of the 18S gene (V9) were used to identify OTUs and estimate their abundance at each sampling station visited by the Tara Ocean expedition (Pierella Karlusich *et al*., 2020). This inventory was not perfect because these data contained several biases, including only one sample per station and identification of organisms based only on 18SvV9 gene sequencing, which allows identification of the genus but not always the species. Despite these biases, Tara Ocean data highlight the global distribution of *T. lutea* across the Atlantic, Pacific and Indian Oceans. The biomass of *T. lutea* is estimated to be around 0.3% of eukaryotes in the plankton, which is far from insignificant (de Vargas *et al*., 2015; Pierella Karlusich *et al*., 2020; Penot *et al*., 2022). This observation suggests that *T. lutea* has a wide distribution across the world as has already been observed in several microalgal taxa including diatoms (Chepurnov *et al*., 2008; Evans *et al*., 2009; De Luca *et al*., 2019; Rastogi *et al*., 2020) dinoflagellates (Le Gac *et al*., 2016; Paredes *et al*., 2019) and other Isochrysidales such as *E. huxleyi* (Read *et al*., 2013), correlated with a high inter-strain diversity.

The variations observed between the *T. lutea* strains in this study might come from the isolation of the strains and partly from their maintenance in the microalgae collection for decades. Genetic drift in algal collections has been little studied but a selection effect and genetic drift have probably impacted the 11 strains of *T. lutea* used in this study, as previously observed in other organisms (Driscoll, 2023; Nelson, 2023). In an industrial context, evaluations of the conservation of improved traits of domesticated microalgae are crucial to ensure that no phenotypic drift will occur over time. The stability of the lipid-improved phenotype of a domesticated strain was recently confirmed after seven years of maintenance in culture conditions despite the genetic drift effect (Berthelier *et al*., 2023). Ideally, the cryopreservation of strains would ensure conservation of the genotype and phenotype properties (Rhodes *et al*., 2006; Nakanishi *et al*., 2012). Cryopreservation is used in algal collections, and among the 11 strains of *T. lutea* studied here, two had been cryopreserved for 10 years (Table 1). In the future, the proper isolation of strains with a modern traceability and cryopreservation protocol would be necessary to link genetic and phenotypic traits of strains according to their habitats.

In this study, we investigated the genetic variations and genetic structure of 11 strains of *T. lutea*. Several types of genetic polymorphism were identified among these strains: short variations (SNPs, indels), TE indels, gene presence/absence variations (PAV) and variation in genome size. Specific variations identified in each strain showed differences between strains. Curiously, the two strains CCAP927-14 and RCC3692 displayed a higher number of TE insertions, which were around two-fold higher than in the nine other strains. The variation observed between these two strains was not the consequence of the increase in the number of copies of a single TE family but resulted from an overall increase in the number of copies of every TE family. The higher number of TEs could suggest that laboratory culture conditions of these two strains may have favoured TE propagations (Craig *et al*., 2021) or could be the result of an ancient response of TEs faced with environmental stress (Lisch, 2013). Previous results from 1000 sequenced ecotypes of *Arabidopsis thaliana* demonstrated that TE indels can be significant actors of genetic variation in populations by affecting genes and improving the host fitness in changing environmental conditions (Baduel *et al*., 2021). At the phenotypic level, the strain RCC3692 showed no particular specificities. However, the CCAP927-14 strain was the most enriched in several pigments (chlorophyll *a*, fucoxanthin, diadinoxanthin, echinenone and β carotene). These results led to a search for specific genetic variation that could be related to this enriched pigment content. However, no genes known to be involved in pigment synthesis were found to be impacted by genetic variations.

On the basis of the genetic variations identified, we detected no genetic structure, demonstrating the genetic singularity of each strain. These inter-strain genetic variations arise from the sampling of these strains in specific habitats at specific times. Furthermore, some of the genetic variations measured could have been produced by evolutionary effects during culture of collections maintained in ideal laboratory growth conditions: genetic drift in culture conditions enhancing the inter-strain divergence. In contrast to this study on *T. lutea*, high genetic structure has been

observed in the diatoms *Ditylym brightwellii* (Rynearson & Virginia Armbrust, 2004) and *Phaeodactylum tricornutum* (Rastogi *et al.*, 2020), the dinoflagellates *Alexandrium catenella* (Paredes *et al.*, 2019) and *Alexandrium minutum* (Le Gac *et al.*, 2016), and the chlorophyte *Ostreococcus tauri* (Blanc-Mathieu *et al.*, 2014). The dispersal ability of *T. lutea* in the ocean could be high, in contrast to other cryptic species revealed by marine studies in diatoms and dinoflagellates (Adams *et al.*, 2009; Godhe & Härnström, 2010). Another known example of a microalgal species with a very weak genetic structure is the Haptophyte *Emiliana huxleyi* (Filatov, 2019). Both *E. huxleyi* and *T. lutea* are haptophytes of the Isochrysidales order, which seems be characterized by strains with a weak genetic structure (Marc *et al.*, 2017; Filatov, 2019). It was previously suggested that the singular life cycle of the Isochrysidales, including predominantly asexual reproduction and a weak mutation rate, could be one of the underlying causes (Marc *et al.*, 2017; Filatov, 2019). However, neither the life cycle of *T. lutea* nor its mutation rate have yet been characterized and future experiments will be necessary to investigate these aspects.

The genetic resources of a species such as *T. lutea* include a dispensable part of the genome, comprising specific alleles, genes or other genomic elements, present in one or all of the ecotypes, that enable the species to adapt to different environments (Marroni *et al.*, 2014; Tranchant-Dubreuil *et al.*, 2019). We found that the proportion of dispensable genes is relatively small for *T. lutea* (5.5% PAV in total), and similar results were obtained in other pangenomic algal studies, such as those on *P. tricornutum* (5% PAV, Rastogi *et al.*, 2020) or *Chlamydomonas reinhardtii* (5% PAV, Gallaher *et al.*, 2015), as well as for *Prochlorococcus* (13% PAV, Delmont & Eren, 2018). The other well-studied haptophyte alga, *E. huxleyi* (Read *et al.*, 2013), contains more PAV (around 25%) than *T. lutea*. The number of dispensable genes in microalgal pangenome studies appears limited compared with pangenome analysis on plants (Bayer *et al.*, 2020) or cyanobacteria (Beck *et al.*, 2012), where PAV is often close to 50%. The small number of pangenome studies on microalgae mean that it is still difficult to generalize on whether limited dispensable genes are characteristic of microalgae.

Despite the lack of available knowledge regarding the function of genes, we attempted to find links between genetic variation and the phenotypic traits observed. Few functional genomics studies have been performed in microalgae and only 25% of *T. lutea* genes have sufficient similarities to previously described proteins to assign them potential functions (Villar *et al.*, 2018). Studies investigating gene-by-gene function require knockout mutants, tools whose development has only just begun in *T. lutea* (Bo *et al.*, 2020). However, this is a challenging process that is less effective in non-model microalgae than in model species. An interesting approach to overcome this lack of knowledge on the gene-phenotype link, would be quantitative genetic strategies such as quantitative trait loci (QTL). Some QTL studies have already been performed in microalgae but remain very marginal because of their technical complexity (Avia *et al.*, 2017; Yu *et al.*, 2020). Such approaches could be promising with *T. lutea*, because the genetic and phenotypic variations seem to be appropriate.

Comparison of the phenotypes of these 11 strains showed that all had phenotypic traits that differentiated them. Each strain was characterized by a specific phenotype, highlighting large inter-strain variation. For example, the comparison of their respective growth rates showed that all these strains have different growth rates under identical culture conditions, suggesting that each strain has specific fitness according to environmental conditions (Flynn *et al.*, 2010). In our culture conditions, CCAP927-14 and RCC3691 strains stand out because of their higher contents of fucoxanthin and DHA, respectively. Concerning echinenone pigments, their role in photoprotection has been widely described in cyanobacteria (Steiger *et al.*, 1999). Echinenone binds to orange carotenoid protein to dissipate energy in cyanobacteria (Sedoud *et al.*, 2014). To date, this pigment has rarely been identified in eukaryote organisms and its function is still unclear (Mulders *et al.*, 2013). Nevertheless, a homologous protein of orange carotenoid protein was also retrieved in *T. lutea* (Pajot *et al.*, 2023). The characterization of echinenone function in algae and its associated metabolic pathway are of interest because they could provide a source of antioxidant molecules for various biotechnological applications. The large phenotypic range observed among these strains is of major interest to the scientific community working on microalgal domestication. Implementing microalgal domestication programmes still represents a considerable challenge to obtain industrial strains with traits of interest for economically viable industrial production of microalgae (Vigani *et al.*, 2015; Maréchal, 2021). Despite the outstanding potential of microalgae for various biotechnological applications, the commercialization of added-value compounds (e.g. pigment or DHA) or biomass itself is still restricted to a high-value niche market. Consequently, microalgal improvement programmes have been conducted in response to the growing demand for biomolecules or biomass productivity (Choi *et al.*, 2008; Xue *et al.*, 2015; López-Smalley *et al.*, 2020; Sánchez *et al.*, 2022; Trovão *et al.*, 2022). The success of domestication programmes relies partly on the exploitation of the phenotypic and genetic diversity of the organism. In this study, the inter-strain genetic and

phenotypic variability observed in *T. lutea* opens the way for obtaining improved strains for industrial marketing. For instance, our results suggest that the strains CCAP927-14 and CC3691, which accumulated the most fucoxanthin and DHA, respectively, among the studied strains, would be interesting candidates for domestication programmes to further improve these traits.

Acknowledgements

We acknowledge culture collections (RCC, CCAP, CCMP, NIVA, NORCCA) the platform Genotoul GeT-PlaGe for the Illumina sequencing of *T. lutea* strains and Helen McCombie (BTU-UBO) for English corrections.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This work was supported by the French National Research Agency, grant DynAlgue project [ANR-16-CE20-16] and by the French Research Institute for Exploitation of the Sea (IFREMER).

Supplementary materials

The following supplementary material is accessible via the Supplementary Content tab on the article's online page at https://doi.org/10.1080/09670262.2023.2249073

Supplementary fig. S1. Phylogeny tree to several strains of Isochrysidaceae family.
Supplementary fig. S2. Final contact map of *T. lutea* reference genome.
Supplementary fig. S3. Results of BUSCO analysis
Supplementary fig. S4. Results of genetic structuring
Supplementary data. S1. Software workflow to genome analysis
Supplementary table. S1. List of strains in the collection, sequencing result, size of genomes, transposable elements distribution, Tajima D score.

Author contributions

GC supervised the study and performed the genetic analyses, wrote the manuscript and made the figures. JB performed transposable element analyses and participated in drafting the manuscript. AM phenotype analyses. EN pigment measure and supervised phenotype analyses. MM HIC-seq. NS genotyping of strains. AC algal cultures and phenotype measure. CC and LL genome browser. BS supervised the study and participated in drafting the manuscript.

ORCID

Grégory Carrier http://orcid.org/0000-0001-6859-5707

References

Adams, N.G., Trainer, V.L., Rocap, G., Herwig, R.P. & Hauser, L. (2009). Genetic population structure of *Pseudonitzschia pungens* (Bacillariophyceae) from the Pacific Northwest and the North Sea. *Journal of Phycology*, **45**(5): 1037–1045.

Almutairi, A.W. (2020). Improvement of chemical composition of *Tisochrysis lutea* grown mixotrophically under nitrogen depletion towards biodiesel production. *Molecules*, **25**(20): 4609.

Arrigo, K.R. (2005). Marine microorganisms and global nutrient cycles. *Nature*, **437**(7057): 349–355.

Avia, K., Coelho, S.M., Montecinos, G.J., Cormier, A., Lerck, F., Mauger, S., Faugeron, S., Valero, M., Cock, J. M. & Boudry, P. (2017). High-density genetic map and identification of QTLs for responses to temperature and salinity stresses in the model brown alga ectocarpus. *Scientific Reports*, **7**: 43241.

Baduel, P., Leduque, B., Ignace, A., Gy, I., Gil, J., Loudet, O., Colot, V. & Quadrana, L. (2021). Genetic and environmental modulation of transposition shapes the evolutionary potential of *Arabidopsis thaliana*. *Genome Biology*, **22**(1): 138.

Baudry, L., Guiglielmoni, N., Marie-Nelly, H., Cormier, A., Marbouty, M., Avia, K., Mie, Y.L., Godfroy, O., Sterck, L., Cock, J.M., Zimmer, C., Coelho, S.M. & Koszul, R. (2020). instaGRAAL: Chromosome-level quality scaffolding of genomes using a proximity ligation-based scaffolder. *Genome Biology*, **21**(1): 148.

Bayer, P.E., Golicz, A.A., Scheben, A., Batley, J. & Edwards, D. (2020). Plant pan-genomes are the new reference. *Nature Plants*, **6**(8): Article 8.

Beck, C., Knoop, H., Axmann, I.M. & Steuer, R. (2012). The diversity of cyanobacterial metabolism: Genome analysis of multiple phototrophic microorganisms. *BMC Genomics*, **13**(1): 56.

Bendif, E., Probert, I., Schroeder, D. & de Vargas, C. (2013). The description of *Tisochrysis lutea* gen. nov. sp. nov. and *Isochrysis nuda* sp. nov. in the Isochrysidales, and the transfer of Dicrateria to the Prymnesiales (Haptophyta). *Journal of Phycology*, **25**(6): 1763–1776.

Berard, J.-B., Bougaran, G. & Carrier, G. (2021). *System and method for culture of phytoplankton* (inpi patent FR3103497). https://data.inpi.fr/brevets/FR3103497?q=ifremer#FR3103497

Berthelier, J., Casse, N., Daccord, N., Jamilloux, V., Saint-Jean, B. & Carrier, G. (2018). A transposable element annotation pipeline and expression analysis reveal potentially active elements in the microalga *Tisochrysis lutea*. *BMC Genomics*, **19**(1): 378.

Berthelier, J., Saint-Jean, B., Casse, N., Bougaran, G. & Carrier, G. (2023). Phenotype stability and dynamics of transposable elements in a strain of the microalga *Tisochrysis lutea* with improved lipid traits. *PLOS ONE*, **18**(4): e0284656.

Blanc-Mathieu, R., Verhelst, B., Derelle, E., Rombauts, S., Bouget, F.-Y., Carré, I., Château, A., Eyre-Walker, A., Grimsley, N., Moreau, H., Piégu, B., Rivals, E., Schackwitz, W., Van de Peer, Y. & Piganeau, G. (2014). An improved genome of the model marine alga *Ostreococcus tauri* unfolds by assessing Illumina de novo assemblies. *BMC Genomics*, **15**(1): 1103.

Bork, P., Bowler, C., de Vargas, C., Gorsky, G., Karsenti, E. & Wincker, P. (2015). Tara oceans studies plankton at planetary scale. *Science*, **348**(6237): 873.

Borowitzka, M.A. (1997). Microalgae for aquaculture: Opportunities and constraints. *Journal of Applied Phycology*, **9**(5): 393–401.

Bougaran, G., Le DÃ©an, L., Lukomska, E., Kaas, R. & Baron, R. (2003). Transient initial phase in continuous culture of *Isochrysis galbana affinis* Tahiti. *Aquatic Living Resources*, **16**(4): 389–394.

Bougaran, G., Rouxel, C., Dubois, N., Kaas, R., Grouas, S., Lukomska, E., Le Coz, J.-R. & Cadoret, J.-P. (2012). Enhancement of neutral lipid productivity in the microalga *Isochrysis affinis* galbana (T-Iso) by a mutation-selection procedure. *Biotechnology and Bioengineering*, **109**(11): 2737–2745.

Bo, Y., Wang, K., Wu, Y., Cao, H., Cui, Y. & Wang, L. (2020). Establishment of a chloroplast transformation system in *Tisochrysis lutea. Journal of Applied Phycology*, **32**(5): 2959–2965.

Carrier, G., Baroukh, C., Rouxel, C., Duboscq-Bidot, L., Schreiber, N. & Bougaran, G. (2018). Draft genomes and phenotypic characterization of *Tisochrysis lutea* strains. Toward the production of domesticated strains with high added value. *Algal Research*, **29**:1–11.

Casteleyn, G., Leliaert, F., Backeljau, T., Debeer, A.-E., Kotaki, Y., Rhodes, L., Lundholm, N., Sabbe, K. & Vyverman, W. (2010). Limits to gene flow in a cosmopolitan marine planktonic diatom. *Proceedings of the National Academy of Sciences of the United States of America*, **107**(29): 12952–12957.

Chen, Y., Liang, H., Du, H., Jesumani, V., He, W., Cheong, K.-L., Li, T. & Hong, T. (2022). Industry chain and challenges of microalgal food industry-a review. *Critical Reviews in Food Science and Nutrition*, 1–28.

Chepurnov, V.A., Mann, D.G., von Dassow, P., Vanormelingen, P., Gillard, J., Inzé, D., Sabbe, K. & Vyverman, W. (2008). In search of new tractable diatoms for experimental biology. *Bioessays*, **30**(7): 692–702.

Choi, G.-G., Bae, M.-S., Ahn, C.-Y. & Oh, H.-M. (2008). Enhanced biomass and gamma-linolenic acid production of mutant strain *Arthrospira platensis. Journal of Microbiology and Biotechnology*, **18**(3): 539–544.

Collos, Y., Mornet, F., Sciandra, A., Waser, N., Larson, A. & Harrison, P. (1999). An optical method for the rapid measurement of micromolar concentrations of nitrate in marine phytoplankton cultures. *Journal of Applied Phycology*, **11**(2): 179–184.

Craig, R.J., Hasan, A.R., Ness, R.W. & Keightley, P.D. (2021). Comparative genomics of Chlamydomonas. *The Plant Cell*, **33**(4): 1016–1041.

Danecek, P., Auton, A., Abecasis, G., Albers, C.A., Banks, E., DePristo, M.A., Handsaker, R.E., Lunter, G., Marth, G.T., Sherry, S.T., McVean, G. & Durbin, R., 1000 Genomes Project Analysis Group. (2011). The variant call format and VCFtools. *Bioinformatics (Oxford, England)*, **27**(15): 2156–2158.

De Coster, W., D'Hert, S., Schultz, D.T., Cruts, M. & Van Broeckhoven, C. (2018). NanoPack: Visualizing and processing long-read sequencing data. *Bioinformatics*, **34**(15): 2666–2669.

De Luca, D., Kooistra, W.H.C.F., Sarno, D., Gaonkar, C.C. & Piredda, R. (2019). Global distribution and diversity of chaetoceros (Bacillariophyta, Mediophyceae): Integration of classical and novel strategies. *PeerJ*, **7**: e7410.

Desper, R. & Gascuel, O. (2002). Fast and accurate phylogeny reconstruction algorithms based on the minimum-evolution principle. *Journal of Computational Biology: A Journal of Computational Molecular Cell Biology*, **9**(5): 687–705.

de Vargas, C., Audic, S., Henry, N., Decelle, J., Mahé, F., Logares, R., Lara, E., Berney, C., Bescot, N.L., Probert, I., Carmichael, M., Poulain, J., Romac, S., Colin, S., Aury, J.-M., Bittner, L., Chaffron, S., Dunthorn, M., Engelen, S. & Karsenti, E. (2015). Eukaryotic plankton diversity in the sunlit Ocean. *Science*, **348**(6237): 1261605.

Delmont, T.O. & Eren, A.M. (2018). Linking pangenomes and metagenomes: The Prochlorococcus metapangenome. *PeerJ*, **6**: e4320.

Driscoll, C. (2023). Book review: Conservation and the genomics of populations: Third edition. *Journal of Heredity*, **114**(2): 195–197.

Edvardsen, B., Eikrem, W., Green, J.C., Andersen, R.A., Moon-van der Staay, S.Y. & Medlin, L.K. (2000). Phylogenetic reconstructions of the Haptophyta inferred from 18S ribosomal DNA sequences and available morphological data. *Phycologia*, **39**(1): 19–35.

Evans, K.M., Chepurnov, V.A., Sluiman, H.J., Thomas, S.J., Spears, B.M. & Mann, D.G. (2009). Highly differentiated populations of the freshwater diatom *Sellaphora capitata* suggest limited dispersal and opportunities for allopatric speciation. *Protist*, **160**(3): 386–396.

Field, C.B., Behrenfeld, M.J., Randerson, J.T. & Falkowski, P. (1998). Primary production of the biosphere: Integrating terrestrial and oceanic components. *Science*, **281**(5374): 237–240.

Filatov, D.A. (2019). Extreme lewontin's paradox in ubiquitous marine phytoplankton species. *Molecular Biology and Evolution*, **36**(1): 4–14.

Flynn, K.J., Greenwell, H.C., Lovitt, R.W. & Shields, R.J. (2010). Selection for fitness at the individual or population levels: Modelling effects of genetic modifications in microalgae on productivity and environmental safety. *Journal of Theoretical Biology*, **263**(3): 269–280.

Folch, J., Lees, M. & Sloane Stanley, G.H. (1957). A simple method for the isolation and purification of total lipides from animal tissues. *The Journal of Biological Chemistry*, **226**(1): 497–509.

Frichot, E., Mathieu, F., Trouillon, T., Bouchard, G. & François, O. (2014). Fast and efficient estimation of individual ancestry coefficients. *Genetics*, **196**(4): 973–983.

Gachelin, M., Boutoute, M., Carrier, G., Talec, A., Pruvost, E., Guihéneuf, F., Bernard, O. & Sciandra, A. (2021). Enhancing PUFA-rich polar lipids in *Tisochrysis lutea* using adaptive laboratory evolution (ALE) with oscillating thermal stress. *Applied Microbiology and Biotechnology*, **105**(1): 301–312.

Gallaher, S.D., Fitz-Gibbon, S.T., Glaesener, A.G., Pellegrini, M. & Merchant, S.S. (2015). Chlamydomonas genome resource for laboratory strains reveals a mosaic of sequence variation, identifies true strain histories, and enables strain-specific studies. *The Plant Cell*, **27**(9): 2335–2352.

Gao, F, Teles (Cabanelas, I., Iago, Wijffels, R. H., & Barbosa, M. J. (2020). Process optimization of fucoxanthin production with *Tisochrysis lutea. Bioresource Technology*, **315**: 123894.

Garnier, M., Carrier, G., Rogniaux, H., Nicolau, E., Bougaran, G., Saint-Jean, B. & Cadoret, J.P. (2014). Comparative proteomics reveals proteins impacted by nitrogen deprivation in wild-type and high lipid-accumulating mutant strains of *Tisochrysis lutea. Journal of Proteomics*, **105**: 107–120.

Garrison, E. & Marth, G. (2012). Haplotype-based variant detection from short-read sequencing. ArXiv, 1207.3907 [q-Bio].

Geider, R.J., Delucia, E.H., Falkowski, P.G., Finzi, A.C., Grime, J.P., Grace, J., Kana, T.M., Roche, J.L., Long, S. P., Osborne, B.A., Platt, T., Prentice, I.C., Raven, J.A., Schlesinger, W.H., Smetacek, V., Stuart, V., Sathyendranath, S., Thomas, R.B., Vogelmann, T.C., Williams, P. & Woodward, F.I. (2001). Primary productivity of planet earth: Biological determinants and physical constraints in terrestrial and aquatic habitats. *Global Change Biology*, **7**(8): 849–882.

Godhe, A. & Härnström, K. (2010). Linking the planktonic and benthic habitat: genetic structure of the marine diatom *Skeletonema marinoi*. *Molecular Ecology*, **19**(20): 4478–4490.

Hernández Javier, L., Benzekri, H., Gut, M., Claros, M.G., van Bergeijk, S., Cañavate, J.P. & Manchado, M. (2018). Characterization of iodine-related molecular processes in the marine microalga *Tisochrysis lutea* (Haptophyta). *Frontiers in Marine Science*, **5**: 134.

Holt, C. & Yandell, M. (2011). MAKER2: An annotation pipeline and genome-database management tool for second-generation genome projects. *BMC Bioinformatics*, **12**(1): 491.

Hu, Z., Zeng, X., Wang, A., Shi, C. & Duan, D. (2004). An efficient method for DNA isolation from red algae. *Journal of Applied Phycology*, **16**(3): 161–166.

Ippoliti, D., González, A., Martín, I., Sevilla, J.M.F., Pistocchi, R. & Acién, F.G. (2016). Outdoor production of *Tisochrysis lutea* in pilot-scale tubular photobioreactors. *Journal of Applied Phycology*, **28**(6): 3159–3166.

Korf, I. (2004). Gene finding in novel genomes. *BMC Bioinformatics*, **5**(1): 59.

Larkum, A.W.D., Ross, I.L., Kruse, O. & Hankamer, B. (2012). Selection, breeding and engineering of microalgae for bioenergy and biofuel production. *Trends In Biotechnology*, **30**(4): 198–205.

Lee, W.-P., Stromberg, M., Ward, A., Stewart, C., Garrison, E. & Marth, G.T. (2013). MOSAIK: A hash-based algorithm for accurate next-generation sequencing read mapping. *ArXiv, 1309.1149 [q-Bio]*.

Le Gac, M., Metegnier, G., Chomérat, N., Malestroit, P., Quéré, J., Bouchez, O., Siano, R., Destombe, C., Guillou, L. & Chapelle, A. (2016). Evolutionary processes and cellular functions underlying divergence in *Alexandrium minutum*. *Molecular Ecology*, **25**(20): 5129–5143.

Li, H. (2018). Minimap2: pairwise alignment for nucleotide sequences. *Bioinformatics*, **34**(18): 3094–3100.

Li, H. & Durbin, R. (2009). Fast and accurate short read alignment with burrows–wheeler transform. *Bioinformatics*, **25**(14): 1754–1760.

Lisch, D. (2013). How important are transposons for plant evolution? *Nature Reviews Genetics*, **14**(1): 49–61.

López-Sánchez, A., Silva-Gálvez, A.L., Aguilar-Juárez, Ó., Senés-Guerrero, C., Orozco-Nunnelly, D.A., Carrillo-Nieves, D. & Gradilla-Hernández, M.S. (2022). Microalgae-based livestock wastewater treatment (MbWT) as a circular bioeconomy approach: Enhancement of biomass productivity, pollutant removal and high-value compound production. *Journal of Environmental Management*, **308**: 114612.

Marc, K., Adam, E.-W., Sophie, S.-F. & Piganeau, G. (2017). Spontaneous mutation rate in the smallest photosynthetic eukaryotes. *Molecular Biology and Evolution*, **34**(7): 1770–1779.

Marchetti, J., Bougaran, G., Le Dean, L., Mégrier, C., Lukomska, E., Kaas, R., Olivo, E., Baron, R., Robert, R. & Cadoret, J.P. (2012). Optimizing conditions for the continuous culture of *Isochrysis affinis galbana* relevant to commercial hatcheries. *Aquaculture*, **326–329**: 106–115.

Maréchal, E. (2021). Grand challenges in microalgae domestication. *Frontiers In Plant Science*, **12**: 764573.

Marroni, F., Pinosio, S. & Morgante, M. (2014). Structural variation and genome complexity: Is dispensable really dispensable? *Current Opinion In Plant Biology*, **18**: 31–36.

Méndez-Leyva, A.B., Guo, J., Mudd, E.A., Wong, J., Schwartz, J.-M. & Day, A. (2019). The chloroplast genome of the marine microalga *Tisochrysis lutea*. *Mitochondrial DNA Part B*, **4**(1): 253–255.

Mohamadnia, S., Tavakoli, O. & Faramarzi, M.A. (2022). Production of fucoxanthin from the microalga *Tisochrysis lutea* in the bubble column photobioreactor applying mass transfer coefficient. *Journal of Biotechnology*, **348**: 47–54.

Mulders, K.J.M., Weesepoel, Y., Lamers, P.P., Vincken, J.-P., Martens, D.E. & Wijffels, R.H. (2013). Growth and pigment accumulation in nutrient-depleted *Isochrysis* aff. *galbana* T-ISO. *Journal of Applied Phycology*, **25**(5): 1421–1430.

Nakanishi, K., Deuchi, K. & Kuwano, K. (2012). Cryopreservation of four valuable strains of microalgae, including viability and characteristics during 15 years of cryostorage. *Journal of Applied Phycology*, **24**(6): 1381–1385.

Nelson, R. (2023). Lessons from a soybean collection. *Crop Science*, **63**(3): 1050–1058.

Nielsen, R. (2001). Statistical tests of selective neutrality in the age of genomics. *Heredity*, **86**(6): 641–647.

Pajot, A., Lavaud, J., Carrier, G., Lacour, T., Marchal, L. & Nicolau, E. (2023). Light-response in two clonal strains of the haptophyte *Tisochrysis lutea*: Evidence for different photoprotection strategies. *Algal Research*, **69**: 102915.

Paredes, J., Varela, D., Martínez, C., Zúñiga, A., Correa, K., Villarroel, A. & Olivares, B. (2019). Population genetic structure at the northern edge of the distribution of *Alexandrium catenella* in the patagonian fjords and its expansion along the open pacific ocean coast. *Frontiers in Marine Science*, **5**: 532.

Penot, M., Dacks, J.B., Read, B. & Dorrell, R.G. (2022). Genomic and meta-genomic insights into the functions, diversity and global distribution of Haptophyte algae. *Applied Phycology*, **3**(1): 340–359.

Pierella Karlusich, J.J., Ibarbalz, F.M. & Bowler, C. (2020). Phytoplankton in the Tara Ocean. *Annual Review of Marine Science*, **12**(1): 233–265.

Pierella Karlusich, J.J., Pelletier, E., Zinger, L., Lombard, F., Zingone, A., Colin, S., Gasol, J.M., Dorrell, R.G., Henry, N., Scalco, E., Acinas, S.G., Wincker, P., de Vargas, C. & Bowler, C. (2023). A robust approach to estimate relative phytoplankton cell abundances from metagenomes. *Molecular Ecology Resources*, **23**(1): 16–40.

Poplin, R., Ruano-Rubio, V., DePristo, M.A., Fennell, T.J., Carneiro, M.O., der Auwera, G.A.V., Kling, D.E., Gauthier, L.D., Levy-Moonshine, A., Roazen, D., Shakir, K., Thibault, J., Chandran, S., Whelan, C., Lek,

M., Gabriel, S., Daly, M.J., Neale, B., MacArthur, D.G. & Banks, E. (2018). Scaling accurate genetic variant discovery to tens of thousands of samples. *bioRxiv*, 201178.

Premaratne, M., Liyanaarachchi, V.C., Nimarshana, P.H.V., Ariyadasa, T.U., Malik, A. & Attalage, R.A. (2021). Co-production of fucoxanthin, docosahexaenoic acid (DHA) and bioethanol from the marine microalga *Tisochrysis lutea*. *Biochemical Engineering Journal*, **176**: 108160.

Rastogi, A., Vieira, F.R.J., Deton-Cabanillas, A.-F., Veluchamy, A., Cantrel, C., Wang, G., Vanormelingen, P., Bowler, C., Piganeau, G., Hu, H. & Tirichine, L. (2020). A genomics approach reveals the global genetic polymorphism, structure, and functional diversity of ten accessions of the marine model diatom *Phaeodactylum tricornutum*. *The ISME Journal*, **14**(2): 347–363.

Read, B.A., Kegel, J., Klute, M.J., Kuo, A., Lefebvre, S.C., Maumus, F., Mayer, C., Miller, J., Monier, A., Salamov, A., Young, J., Aguilar, M., Claverie, J.-M., Frickenhaus, S., Gonzalez, K., Herman, E.K., Lin, Y.-C., Napier, J., Ogata, H., Sarno, A.F., Shmutz, J., Schroeder, D., de Vargas, C., Verret, F., von Dassow, P., Valentin, K., Van de Peer, Y., Wheeler, G., Dacks, J.B., Delwiche, C. F., Dyhrman, S.T., Glöckner, G., John, U., Richards, T., Worden, T.Z., Zhang, X. & Grigoriev, I.V. (2013). Pan genome of the phytoplankton *Emiliania* underpins its global distribution. *Nature*, **499**(7457): 209–213.

Rhodes, L., Smith, J., Tervit, R., Roberts, R., Adamson, J., Adams, S. & Decker, M. (2006). Cryopreservation of economically valuable marine micro-algae in the classes Bacillariophyceae, Chlorophyceae, Cyanophyceae, Dinophyceae, Haptophyceae, Prasinophyceae, and Rhodophyceae. *Cryobiology*, **52**(1): 152–156.

Rizk, G., Gouin, A., Chikhi, R. & Lemaitre, C. (2014). MindTheGap: Integrated detection and assembly of short and long insertions. *Bioinformatics*, **30**(24): 3451–3457.

Rodríguez, F., Derelle, E., Guillou, L., Le Gall, F., Vaulot, D. & Moreau, H. (2005). Ecotype diversity in the marine picoeukaryote Ostreococcus (Chlorophyta, Prasinophyceae). *Environmental Microbiology*, **7**(6): 853–859.

Rumin, J., Nicolau, E., Gonçalves de Oliveira Junior, R., Fuentes-Grünewald, C. & Picot, L. (2020). Analysis of scientific research driving microalgae market opportunities in Europe. *Marine Drugs*, **18**(5): Article 5.

Rynearson, T.A. & Virginia Armbrust, E. (2004). Genetic differentiation among populations of the planktonic marine diatom *Ditylum brightwellii* (Bacillariophyceae). *Journal of Phycology*, **40**(1): 34–43.

Schmidt, M., Heese, K. & Kutzner, A. (2019). Accurate high throughput alignment via line sweep-based seed processing. *Nature communications*, **10**(1939).

Sedlazeck, F.J., Rescheneder, P., Smolka, M., Fang, H., Nattestad, M., von Haeseler, A. & Schatz, M.C. (2018). Accurate detection of complex structural variations using single-molecule sequencing. *Nature Methods*, **15** (6): 461–468.

Sedoud, A., López-Igual, R., Ur Rehman, A., Wilson, A., Perreau, F., Boulay, C., Vass, I., Krieger-Liszkay, A. & Kirilovsky, D. (2014). The cyanobacterial photoactive orange carotenoid protein is an excellent singlet oxygen quencher. *The Plant Cell*, **26**(4): 1781–1791.

Simon, N., Cras, A.-L., Foulon, E. & Lemée, R. (2009). Diversity and evolution of marine phytoplankton. *Comptes Rendus Biologies*, **332**(2–3): 159–170.

Smalley, T., Fields, F.J., Berndt, A.J., Ostrand, J.T., Heredia, V. & Mayfield, S.P. (2020). Improving biomass and lipid yields of *Desmodesmus armatus* and *Chlorella vulgaris* through mutagenesis and high-throughput screening. *Biomass and Bioenergy*, **142**: 105755.

Soudant, P., Marty, Y., Moal, J. & Samain, J. (1995). Separation of major polar lipids in *Pecten maximus* by high-performance liquid chromatography and subsequent determination of their fatty acids using gas chromatography. *Journal of Chromatography B: Biomedical Sciences and Applications*, **673**(1): 15–26.

Stanke, M., Tzvetkova, A. & Morgenstern, B. (2006). AUGUSTUS at EGASP: Using EST, protein and genomic alignments for improved gene prediction in the human genome. *Genome Biology*, **7**(1): S11.

Steiger, S., Schäfer, L. & Sandmann, G. (1999). High-light-dependent upregulation of carotenoids and their antioxidative properties in the cyanobacterium *Synechocystis* PCC 6803. *Journal of Photochemistry and Photobiology B: Biology*, **52**(1): 14–18.

Sunagawa, S., Acinas, S.G., Bork, P., Bowler, C., Eveillard, D., Gorsky, G., Guidi, L., Iudicone, D., Karsenti, E., Lombard, F., Ogata, H., Pesant, S., Sullivan, M.B., Wincker, P. & de Vargas, C. (2020). Tara Oceans: Towards global ocean ecosystems biology. *Nature Reviews Microbiology*, **18**(8): 428–445.

Thiriet-Rupert, S., Carrier, G., Trottier, C., Eveillard, D., Schoefs, B., Bougaran, G., Cadoret, J.-P., Chénais, B. & Saint-Jean, B. (2018). Identification of transcription factors involved in the phenotype of a domesticated oleaginous microalgae strain of *Tisochrysis lutea*. *Algal Research*, **30**: 59–72.

Thung, D.T., de Ligt, J., Vissers, L.E., Steehouwer, M., Kroon, M., de Vries, P., Slagboom, E.P., Ye, K., Veltman, J.A. & Hehir-Kwa, J.Y. (2014). Mobster: Accurate detection of mobile element insertions in next generation sequencing data. *Genome Biology*, **15**(10): 488.

Thurn, A.-L., Stock, A., Gerwald, S. & Weuster-Botz, D. (2022). Simultaneous photoautotrophic production of DHA and EPA by *Tisochrysis lutea* and *Microchloropsis salina* in co-culture. *Bioresources and Bioprocessing*, **9**(1): 130.

Tranchant-Dubreuil, C., Rouard, M. & Sabot, F. (2019). Plant pangenome: Impacts on phenotypes and evolution. *Annual Plant Reviews Online*, **2**(2): 453–478.

Trovão, M., Schüler, L.M., Machado, A., Bombo, G., Navalho, S., Barros, A., Pereira, H., Silva, J., Freitas, F. & Varela, J. (2022). Random mutagenesis as a promising tool for microalgal strain improvement towards industrial production. *Marine Drugs*, **20**(7): 40.

Vernette, C., Henry, N., Lecubin, J., de Vargas, C., Hingamp, P. & Lescot, M. (2021). The Ocean barcode atlas: A web service to explore the biodiversity and biogeography of marine organisms. *Molecular Ecology Resources*, **21**(4): 1347–1358.

Vershinin, A.O. & Kamnev, A.N. (1996). Xanthophyll Cycle In Marine Macroalgae. *Botanica Marina*, **39**(1–6): 421–426.

Vigani, M., Parisi, C., Rodríguez-Cerezo, E., Barbosa, M.J., Sijtsma, L., Ploeg, M. & Enzing, C. (2015). Food and feed products from micro-algae: Market opportunities and challenges for the EU. *Trends In Food Science & Technology*, **42**(1): 81–92.

Villar, E., Vannier, T., Vernette, C., Lescot, M., Cuenca, M., Alexandre, A., Bachelerie, P., Rosnet, T., Pelletier, E., Sunagawa, S. & Hingamp, P. (2018). The Ocean gene

atlas: Exploring the biogeography of plankton genes online. *Nucleic Acids Research*, **46**(W1): W289–W295.

Walne, P. (1966). Experiments in the large-scale culture of the larvae of *Ostrea edulis* L. *Journal of Ministry of Agriculture, Fisheries Invest*, **2**: 25–53.

Xue, J., Niu, Y.-F., Huang, T., Yang, W.-D., Liu, J.-S. & Li, H.-Y. (2015). Genetic improvement of the microalga *Phaeodactylum tricornutum* for boosting neutral lipid accumulation. *Metabolic Engineering*, **27**: 1–9.

Yazdani, M. (2015). Concerns in the application of fluorescent probes DCDHF-DA, DHR 123 and DHE to measure reactive oxygen species in vitro. *Toxicology in Vitro: An International Journal Published in Association With BIBRA*, **30**(1 Pt B): 578–582.

Yu, X., Wang, L., Xu, K., Kong, F., Wang, D., Tang, X., Sun, B. & Mao, Y. (2020). Fine mapping to identify the functional genetic locus for red coloration in *Pyropia yezoensis* thallus. *Frontiers in Plant Science*, **11**: 867.