Journal of Oceanology and Limnology

https://doi.org/10.1007/s00343-023-3060-2

Genome-wide SNP markers provided insights into the reproductive strategy and genetic diversity of the green tide causative species *Ulva prolifera* in China*

Song SUN[1, 2, 3, 4], Weiwei MA[1, 4], Nan WANG[1, 2], Song FENG[1, 2], Yan SUN[1, 2, **]

[1] *CAS Key Laboratory of Marine Ecology and Environmental Sciences, Institute of Oceanology, Chinese Academy of Sciences, Qingdao 266071, China*

[2] *Laboratory for Marine Ecology and Environmental Science, Laoshan Laboratory, Qingdao 266237, China*

[3] *Jiaozhou Bay Marine Ecosystem Research Station, Chinese Academy of Sciences, Qingdao 266071, China*

[4] *University of Chinese Academy of Sciences, Beijing 100049, China*

Received Apr. 3, 2023; accepted in principle May 9, 2023; accepted for publication May 15, 2023

© Chinese Society for Oceanology and Limnology, Science Press and Springer-Verlag GmbH Germany, part of Springer Nature 2023

Abstract *Ulva prolifera* is the causative species of the annually occurring large-scale green tides in China since 2007. Its specific biological features on reproductivity strategies, as well as intra-species genetic diversity, are still largely unknown, especially at the genome level, despite their importance in understanding the formation and outbreak of massive green tides. In the present study, the restriction site-associated DNA genotyping approach (2b-RAD) was adopted to identify the genome-wide single-nucleotide polymorphisms (SNPs) of 54 individual thalli including samples collected from Subei Shoal in 2019 and Qingdao coast from 2019 to 2021. SNPs genotype results revealed that most of the thalli in 2019 and 2020 were haploid gametophytes, while only half of the thalli were gametophytes in 2021, indicating flexibility in the reproductive strategies for the formation of the green tides among different years and the dominance of asexual and vegetative reproductive mode for the floating period. Besides, population analysis was conducted and revealed a very low genetic diversity among samples from Subei Shoal and the Qingdao coast in the same year and a higher divergence among samples in different years. The results showed the efficiency of 2b-RAD in the exploration of SNPs in *U. prolifera* and provided the first genome-wide scale evidence for the origin of the large-scale green tides on the Qingdao coast. This study improved our understanding of the reproductive strategy and genetic diversity of the green tide causative species and will help in further revealing the biological causes of the green tide in China.

Keyword: green tide; *Ulva prolifera*; 2b-RAD; single-nucleotide polymorphism (SNP); reproductive strategy; genetic diversity

1 INTRODUCTION

Green tide is a harmful ecological phenomenon commonly caused by the explosive proliferation and aggregation of floating large green algae under favorable conditions (Li et al., 2016; Cao et al., 2023). Although the algae are nontoxic, their blooms can dramatically influence the aquatic environment and result in serious ecological impact and economic losses (Liu and Zhou, 2018; Zhang et al., 2019; Yuan et al., 2022). In recent years, green tides have occurred more frequently in coastal areas, estuaries, and lakes throughout the world and become a cosmopolitan problem (Morand and Briand, 1996; Ye et al., 2011; Smetacek and Zingone, 2013). In China, large-scale green tides occurred in the Yellow Sea every year since 2007, and have attracted extensive studies to the physiological and biological character of the dominant species, as well as the origin, formation, and development of the world's largest macroalgal bloom (Liu et al., 2010, 2013; Pang et al., 2010; Duan et al., 2012;

* Supported by the Laoshan Laboratory (No. LSKJ202204005), the Mount Tai Scholar Climbing Plan to Song SUN, and the Open Fund of CAS Key Laboratory of Marine Ecology and Environmental Sciences, Institute of Oceanology, Chinese Academy of Sciences (No. KLMEES201801)

** Corresponding author: ysun@qdio.ac.cn

Wang et al., 2015). According to these analyses, *Ulva prolifera* has been confirmed as the causative species of the Yellow Sea green tides (Zhang et al., 2015; Zhao et al., 2015; Liu and Zhou, 2018). The green tide caused by *U. prolifera* was mostly from the Subei Shoal along the Jiangsu coast and drifted more than 200 km northward in the Yellow Sea to the Shandong coast, driven by surface currents and southwest and southeast winds (Keesing et al., 2011; Xing et al., 2018).

The formation of the green tide in the Yellow Sea is associated with complex events including human activities (*Porphyra* aquaculture and eutrophication), and natural geohydrodynamic and climatic conditions (sand shoals, currents, temperature, and wind), which create suitable environments for the green algal blooms (Liu and Zhou, 2018). However, except for the chemical-physical conditions, the distinct biological traits of the causative species, *U. prolifera*, were the most indispensable causal for the green tide outbreaks. *U. prolifera* has high ecological adaptations with a broad tolerance to temperature, salinity, and irradiance (Dan et al., 2002; Han et al., 2013). It can grow fast in nutrition-rich conditions with a growth rate of 36% per day (Liu et al., 2010; Keesing et al., 2011). Furthermore, its extraordinary complexity in reproduction is also an important reason that makes it possible to outbreak in certain circumstances. It has been reported that *U. prolifera* have a variety of reproductive pathways including sexual, asexual, and vegetative reproduction (Lin et al., 2008). Sexual reproduction involves the fusion of opposite mating types, minus (mt^-) and plus (mt^+), which form zygotes and develop into diploid sporophytes (Hiraoka et al., 2003; Yamazaki et al., 2017). The asexual reproductive mode includes parthenogenesis and apomeiosis. In the former mode, gametes fail to mate and develop into haploid gametophytes, while in the latter mode, the asexual thalli release diploid biflagellate or quadriflagellate zoosporoids that directly develop into diploid thalli (Liu et al., 2015; Hiraoka and Higa, 2016; Ichihara et al., 2019). Additionally, various vegetative reproductive ways, including regeneration from segments, protoplasts, isolated cells, and in-situ somatic germination, have also been reported (Lin et al., 2008; Wu et al., 2018; Liu et al., 2022). Despite its importance in the rapid colonization and the outbreak of green tide, field investigations focused on the reproductive strategy of *U. prolifera* are limited, due to the difficulties in distinguishing the isomorphic haploid gametophytes and diploid sporophytes.

Molecular biotechnologies played a very important role in the previous research on the Yellow Sea green tide, such as distinguishing the bloom-forming algae, investigating their distribution and seasonal variation, and exploring their reproductive mode (Han et al., 2013; Wang et al., 2018). Genetic markers, including ITS, rbcL, and 5S sequence, have been widely used for molecular identification of causative species. Besides, a sequence characterized amplified region (SCAR) marker highly specific to the floating ecotype of *U. prolifera* was developed and used to track the origin of the green tide (Zhao et al., 2015; Zhang et al., 2018). A recent study also developed two mating type-specific makers (mt^- and mt^+) that could detect the ploidy or sex of thalli and uncover their reproductive modes in an efficient way (Liu et al., 2022). In addition to the identification of species, more highly variable molecular markers, such as inter-simple sequence repeat (ISSR) and simple sequence repeat (SSR), were used to investigate the population genetic diversity of *Ulva* populations (Zhao et al., 2011, 2015; Zhang et al., 2014b; Li et al., 2016). These studies suggested that floating *U. prolifera* samples had a close genetic relationship and were separated from their relevant attached species (Zhao et al., 2011; Zhang et al., 2014b; Li et al., 2016). Analysis of genetic diversity can help to correctly identify the origin of the blooms and provide valuable information for understanding the annual outbreak of the massive green tides. However, most of the previous studies used a limited number of markers from a narrow region of the genome and not enough genetic data was described to reveal diversity at the genome level.

The advances in high-throughput next-generation DNA sequencing (NGS) technology have provided a more convenient framework for detecting high-information whole-genome markers. 2b-RAD, a restriction site-associated DNA (RAD) genotyping method based on sequencing the uniform fragments produced by type IIB restriction endonucleases (for example, BsaXI and AlfI), is a representative cost-effective approach that can explore the single-nucleotide polymorphisms (SNPs) covering the whole genome range and provide high-resolution population genomics data for population genetic analyses (Wang et al., 2012, 2016). So far, 2b-RAD has been widely applied to assess genetic diversity in many marine species (Wang et al., 2021; Fifer et al., 2022; Ruocco et al., 2022). In the present study, 2b-RAD was used to genotype a total of 54 individual thalli collected from the Subei Shoal and Qingdao in three consecutive years from 2019 to

2021. The objectives were to detect and genotype single-nucleotide polymorphisms (SNPs) at a genome-wide level and to characterize the genetic diversity among different populations. Our present study will provide useful information for the utility of 2b-RAD on the genetic analysis of *Ulva* species. Moreover, the results may also shed light on the reproduction strategies of the bloom-forming alga, provide evidence in determining the origin of the algae blooms in the Yellow Sea, and contribute to the management of green tides.

2 MATERIAL AND METHOD

2.1 Collection of samples

The algal samples were collected during the green-tide event in three consecutive years from 2019 to 2021. When collected, the samples were cleaned simply in situ and kept in an ice box. After being transported to the laboratory, the algal samples were washed three times in sterile seawater and separated carefully into individual thalli (single specimens). The individual thalli were stored at -80 °C before DNA extraction. Sampling sites were shown in Fig.1. And a summary of the collected individuals used in this study was shown in Table 1. Briefly, a total of 37 individuals were collected at Subei Shoal and at the coastal areas of Qingdao in 2019; 9 and 8 individuals were collected at Qingdao in 2020 and 2021, respectively.

2.2 DNA extraction and identification of individual thallus

Genomic DNA was extracted from individual thallus using the Plant Genomic DNA Kit (TIANGEN, China) according to the manufacturer's protocol. The integrity of DNA was confirmed on 1% agarose gel. Molecular identification of *U. prolifera* was based on the PCR amplification of ITS region and 5S rDNA spacer region. The primer sequences for PCR amplification and the PCR reaction programs were following the descriptions by Zhang et al. (2018).

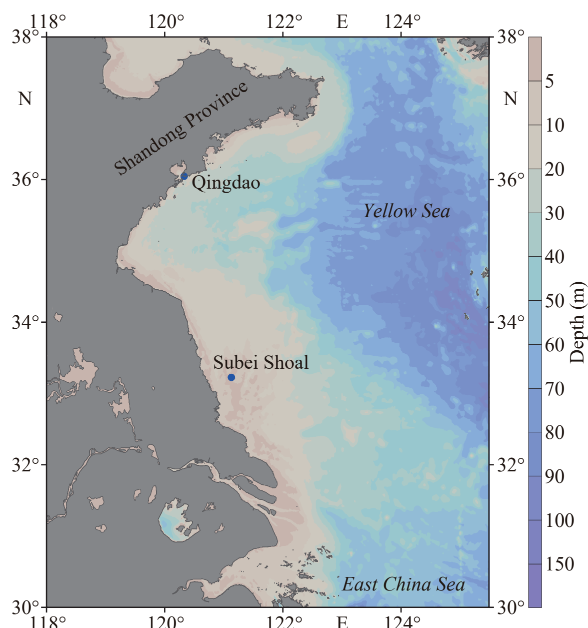

Fig.1 Sampling sites of the *U. prolifera*

The map is created by Ocean Data View (Schlitzer, 2023). The blue dots indicated the locations of the two sampling sites in this study.

The PCR products were confirmed by electrophoresis in 1% agarose gel, and sequenced by Sangon Biotech (Shanghai) Co., Ltd. The ITS sequences of the genera *Ulva* and *Blidingia*, as well as 5S sequences of *U. prolifera* and *U. linza* were downloaded from GenBank and aligned with the sequences obtained in the present study using ClustalW. The phylogenetic tree was constructed by neighbor-joining (NJ) method using MEGA7 (Kumar et al., 2016). The reliability of branches was estimated with non-parametric bootstrapping (1 000 replicates). In the present study, only samples collected in 2019 were used for PCR amplification of ITS region and 5S rDNA spacer region, and samples collected at Qingdao in the years 2020 and 2021 were identified based on their morphology.

2.3 2b-RAD library construction and sequencing

The 2b-RAD libraries were prepared as described by Wang et al. (2016). Briefly, genomic DNA from

Table 1 Summary of the *U. prolifera* individuals used in this study

Group	Sample	Number of individuals	Collection date	Sampling area	Longitude (E)	Latitude (N)
SS19	SS19-1 to -12	12	May 31, 2019	Subei Shoal	121°08′	33°14′
QD19	QD19-1 to -15	15	June 28, 2019	Qingdao	120°20′	36°03′
	QD19-16 to -25	10	July 29, 2019	Qingdao	120°20′	36°03′
QD20	QD20-1 to -9	9	June 24, 2020	Qingdao	120°20′	36°03′
QD21	QD21-1 to -8	8	June 18, 2021	Qingdao	120°20′	36°03′

each individual thallus was digested with type IIB restriction enzymes BsaXI (New England Biolabs, USA) at 37 °C for 45 min. The digestion products of every five individuals were set as one group for adaptor ligation and PCR amplification to introduce specific adaptors containing SapI restriction sites (five pairs of adaptors per five samples). Then the PCR products of the five individuals were purified and mixed. The mixture was digested with SapI (New England Biolabs, USA) at 37 °C for 30 min. After purification through magnetic beads, the digested DNA fragments from five individuals were concatenated in a predefined order using T4 DNA ligase (New England Biolabs, USA). The ligation products were purified with 8% polyacrylamide gel again, and barcodes were introduced by PCR with barcode-bearing primers. The PCR products were purified by the MinElute PCR Purification Kit (Qiagen, Germany) and pooled for sequencing using the Illumina Hiseq 2500 platform.

2.4 Sequence data processing and genotyping

The raw paired-end (PE) reads were first assembled using PEAR software (Zhang et al., 2014a). Then the data were filtered by removing reads with more than 8% ambiguous bases (N), of poor quality (15% nucleotide positions with a Phred quality <30). The obtained high-quality reads containing tags from five individuals were divided into single-tag datasets using the Perl scripts according to Wang et al. (2016), and tags without restriction sites were removed.

The BsaXI tags of *U. prolifera* genome (NCBI accession number GCA_023078555.1) were extracted based on the enzyme's recognition site, which served as a reference for SNP discovery. High-quality tags of each individual were aligned to the reference genome using SOAP2 (Li et al., 2009) with the following parameters: *r*=0 (discard all multi-reads), *M*=4 (find the best match), and *v*=2 (allow two mismatches on a read). Genotyping analysis of sequencing data was performed using the RADtyping software (Fu et al., 2013). To obtain robust results in the subsequent analyses, the SNP markers were further filtered. The selective rules of markers included polymorphism, sufficient genotype rate (available genotyping over 80% of individuals), and minor allele frequency (MAF) higher than 0.01. Then, the qualified markers were used for downstream analysis.

2.5 Identification and validation of haploid individuals

The SNP markers were used to identify haploid individual thallus. Numbers of different SNP loci types (homozygous, heterozygous, or missing) among individual thallus were statistically studied based on the variant call format (VCF) file using an in-house python script. In this study, we considered the individuals with the majority of their markers homozygous as haploid gametophytes (proportion of heterozygous markers <0.05), while others were diploids or possible cross-contamination. To confirm the results of SNP markers, we also conducted PCR amplification of the mating type-related markers (Liu et al., 2022) for validation of the ploidy of the haploid gametophytes. The PCR products of mating type-related markers were detected by electrophoresis in 1.2% agarose gel. In the study, only individuals collected in Subei Shoal (SS19) were employed for mating-type detection, other samples were not validated due to the limitation of quantity and quality of their DNA extracted.

2.6 Population genetic analysis

The haploid individual thalli were used for population genetic analysis. First, genotyping results of the haploid thalli were extracted from the VCF file using vcftools (Danecek et al., 2011). The pairwise genetic differentiation F_{ST} values among the studied populations were calculated by R package hierfstat (Goudet, 2005). Genetic distance based on the proportion of different loci was calculated and visualized using the Poppr package (Kamvar et al., 2014). Analysis of molecular variance (AMOVA) that quantified the proportion of variation was also performed using the Poppr package.

3 RESULT

3.1 Molecular identification of *U. prolifera* was based on ITS and 5S sequences

Altogether 37 individual thalli collected in the year 2019 were employed for ITS and 5S rDNA spacer region sequencing. Phylogenetic analysis of ITS sequences demonstrated that all the samples were clustered into the *U. linza-procera-prolifera* (LPP) clade (Fig.2a). To gain an insight into the taxonomic status of the LPP complex, all samples were further analyzed based on 5S rDNA spacer sequence. According to the 5S sequence, all the samples were attributed to the *U. prolifera* 5S-B group.

3.2 2b-RAD sequencing and genotyping

2b-RAD sequencing of the 54 individual thalli in the three consecutive years from 2019 to 2021

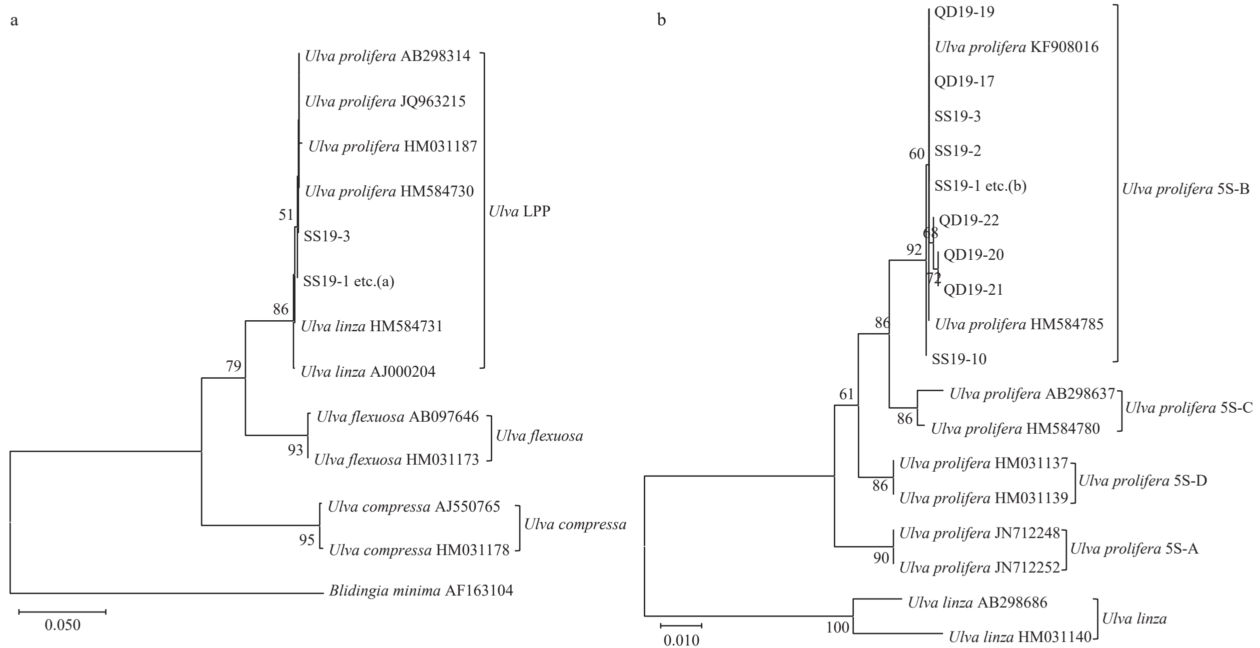

Fig.2 Neighbor-joining phylogeny tree (NJ tree) based on the sequences of the ITS region (a) and 5S rDNA spacer region (b)

Values at branch nodes represent NJ bootstrap probability (>50%). For the sequenced individuals, only different haplotypes were presented. Alphabet a includes SS19-2, SS19-4, SS19-5, SS19-6, SS19-7 SS19-8, SS19-9, SS19-10, SS19-11, SS19-12, QD19-1, QD19-2, QD19-3, QD19-4, QD19-5, QD19-6, QD19-7, QD19-8, QD19-9, QD19-10, QD19-11, QD19-12, QD19-13, QD19-14, QD19-15, QD19-16, QD19-17, QD19-18, QD19-19, QD19-20, QD19-21, QD19-22, QD19-23, QD19-24, QD19-25. Alphabet b includes SS19-4, SS19-5, SS19-6, SS19-7 SS19-8, SS19-9, SS19-11, SS19-12, QD19-1, QD19-2, QD19-3, QD19-4, QD19-5, QD19-6, QD19-7, QD19-8, QD19-9, QD19-10, QD19-11, QD19-12, QD19-13, QD19-14, QD19-15, QD19-16, QD19-18, QD19-23, QD19-24, and QD19-25.

obtained a total of 370 265 821 high-quality clean reads, with an average of 6 856 774 for each sample (Supplementary Table S1). The clean reads were mapped to the reference genome of *U. prolifera*, resulting in an average of 49.28% mapping rate. After clustering, the tags (depth>3X and <500X) of all samples ranged from 32 023 to 35 089 with an average of 34 078. The average coverage depth per tag among samples varied from 41.91× to 167.69× (mean 95.35×; Supplementary Table S2). A total of 2 407 high-quality SNPs were kept and used for subsequent analysis. These SNPs were selected by loci genotyped in >80% of the individuals and showed a MAF >0.01.

3.3 Identification of haploid individuals

Proportions of different SNP loci types (homozygous, heterozygous, or missing) among individual thalli were statistics according to their genotyping results (Fig.3). Samples with the majority of their markers homozygous were considered haploid gametophytes (proportion of heterozygous markers <0.05), while others were diploids or possible cross-contamination. According to the results, a total of 47 individuals were presumptive to be haploids. The proportions of the haploids among the four groups were different. All the individuals in the SS19 group were detected to be haploid gametophytes; the proportions of haploids in the QD19 and QD20 groups were slightly lower, accounting for 92% and 89%, respectively; contrastively, the proportion of haploids detected in the QD21 groups was significantly lower compared with the other groups, only half of the total samples were presumptive to be haploid gametophytes.

The mating type-related markers of *U. prolifera* (Liu et al., 2022) were used for validation of the ploidy of the haploid gametophytes. Due to the limitation in the quantity and quality of DNA extracted, only thalli collected from Subei Shoal (SS19) were used for mating-type detection. The results revealed all the individual thalli from the SS19 group could be amplified with only one of the mating-type related markers (Fig.4), demonstrating that they were either mt^+ or mt^- gametophytes, which was comparable with their ploidy detected based on the SNP genotyping results.

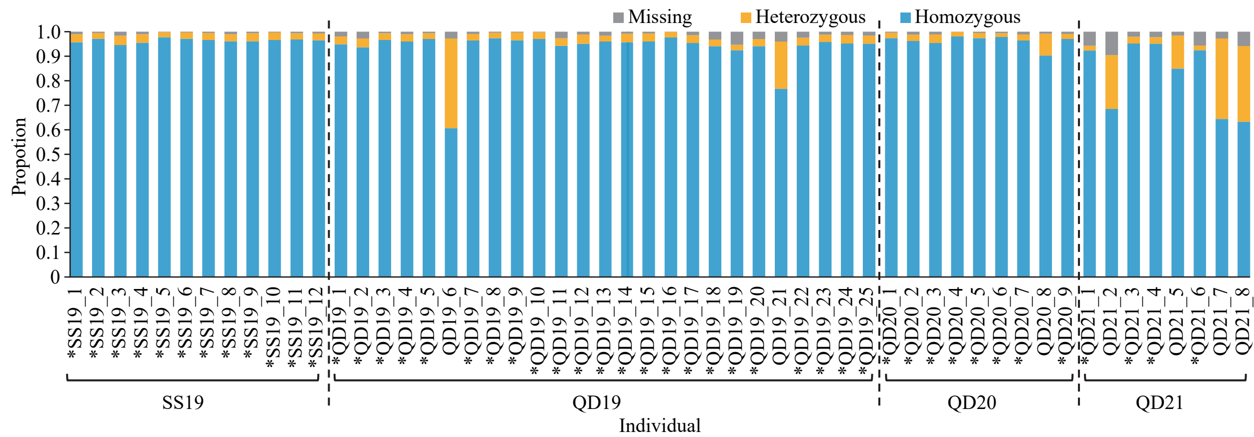

Fig.3 Statistics of SNP type in the sequenced individual thalli

The presumptive haploid individuals with a proportion of heterozygous markers <0.05 are marked with an asterisk (*).

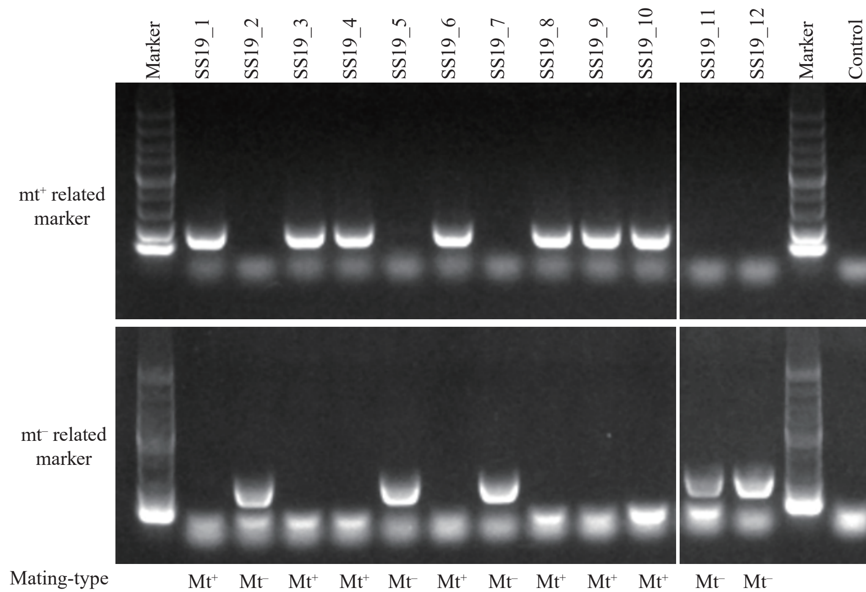

Fig.4 Ploidy validation of *U. prolifera* thalli using mating type-related markers

3.4 Population genetic diversity

The haploid individuals were used to analyze population genetic diversity among different groups. Pairwise F_{ST} values between the four groups ranged from 0.020 to 0.177 (Table 2). The F_{ST} value between the two groups collected in the year 2019 was the lowest, whereas the F_{ST} value between different years was higher, suggesting that only estimates of pairwise population differentiation between samples in different years. AMOVA analysis revealed that about 82% of the variation was within groups (Table 3), which was much higher than that between groups, indicating that the genetic variation of these samples was mainly caused by differentiation within groups. In addition, a dendrogram based on the genetic distance of the 47 haploid *U. prolifera* individuals was generated using UPGMA clustering method (Fig.5). The result did not show clustering relationships of samples from different groups.

Table 2 Pairwise F_{ST} values across four groups of *U. prolifera*.

	SS19	QD19	QD20	QD21
SS19	–	–	–	–
QD19	0.020	–	–	–
QD20	0.158	0.142	–	–
QD21	0.102	0.116	0.177	–

Table 3 AMOVA analysis among populations and within populations

Source of variation	Degree of freedom	Sum of squares	Mean squares	Percentage of variation (%)
Between groups	3	4 367.992	1 455.997 5	18.00
Within groups	43	19 164.438	445.684 6	82.00
Total	46	23 532.430	511.574 6	100

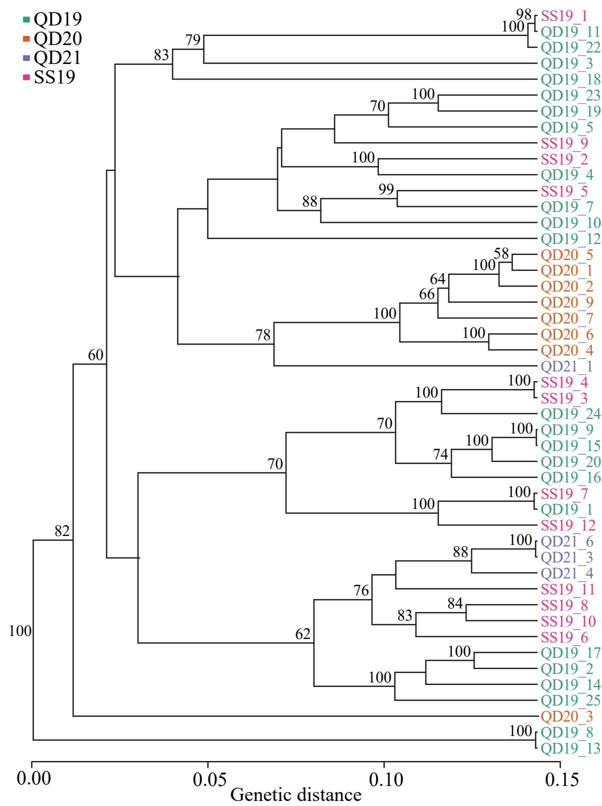

Fig.5 Dendrogram of *U. prolifera* individual thalli based on genetic distance

The dendrogram tree was constructed by UPGMA method with 1 000 bootstrap replicates (node values greater than 50% are shown).

4 DISCUSSION

Green tide caused by *U. prolifera* is a serious environmental problem in China and even in the whole world. Genetic analysis of bloom-forming alga is important in tracking their origin and understanding the biological feature and reproductive strategies which lead to their extremely high biomass during green tides. Previous genetic studies of the population diversity of *U. prolifera* were mainly based on less variable molecular markers restricted in narrow genome regions. Besides, many of them focused on limited sampling sites or specific sampling times. Therefore, genome-wide studies and systematic investigation for samples from different regions and outbreak years are needed. In the current study, we reported the first identification of genome-wide SNPs in *U. prolifera* of the Subei Shoal population collected in 2019 and Qingdao populations collected in three consecutive years from 2019 to 2021. Our results provided insights into the reproductive strategy and the population diversity of the bloom-forming alga.

4.1 Implications for the reproductive strategies of *U. prolifera*

Ulva prolifera has an extraordinary complexity of reproductive pathways including sexual, asexual, and vegetative reproduction. Despite varied in ploidy or mating type, all the progeny generated from different reproductive types share an identical morphology. Using the genome-wide SNP genotyping results, we determined the life history types of the sampled individual thalli. According to our results, the reproductive patterns in different years were different. Most of the samples collected in 2019 and 2020 were haploid gametophytes (>89%), suggesting that the gametophytes were dominant in the years 2019 and 2020 (Fig.3). On the contrary, the proportion of haploids detected in the samples in 2021 (50%) was significantly lower compared with the previous two years (Fig.3). Therefore, gametophytes and sporophytes may coexist in 2021. Both sporophytes and gametophytes of the floating *U. prolifera* samples have been observed in previous studies. Based on the types of formed germ cells, four sporophytes and two gametophytes were identified in six samples collected in 2011 (Liu et al., 2015), whereas only sporophytes were detected based on karyotype analysis of 27 individuals collected in 2018 (Zhao et al., 2019). To date, there is no general consensus regarding the main reproductive strategies for *U. prolifera* blooms (Zhang et al., 2016; Cui et al., 2018; Zhao et al., 2019). Based on the present and previous studies, it is likely that both sexual and asexual lifestyles could be the dominant type of green tide outbreaks. The formation of the *U. prolifera* blooms began with the germination of the micropropagules (including gametes, meiospores, and zygotes) in the "seed

bank" at the Subei Shoal. Germination differences were detected in the micropropagules, with the germination rate approximately three times greater for zygotes (91.67%) and meiospores (80.29%) than that for gametes (30%) (Cui et al., 2018), implying the advantages of sexual reproductive mode. However, the crossover of different mating-type may rely on certain environmental conditions. For example, low salinity was reported to have negative effects on the success of sexual reproduction (Hiraoka and Higa, 2016). Therefore, the interannual difference in the chemical and physical conditions may influence the reproductive strategies during the formation of the green tide. The flexibility in their reproductive mode contributes to the environmental adaption in different years and might be a critical reason for the annual *U. prolifera* blooms.

In addition, we also detected that the reproductive pattern was consistent in individual thalli collected from different sampling sites in the same year. All the individuals collected from the Subei Shoal in 2019 were detected to be haploid gametophytes, and most of the samples from Qingdao in the same year were also gametophytes except two individuals (Fig.3). Field surveys and satellite image analyses have revealed that the macroalgae floating mats formed in the Subei Shoal along the Jiangsu coast in early May (Liu and Zhou, 2018). They drifted more than 200 km northward in the Yellow Sea to the coast of Shandong from May to July (Keesing et al., 2011; Xing et al., 2018). As indicated in our results, the dominant reproductive pattern of *U. prolifera* during their floating period lasted for three months might be asexual and vegetative reproduction. These results are consistent with the previous analysis, which enabled the rapid expansion and large-scale increase of green tide biomass (Cui et al., 2018). Besides, floating *U. prolifera* has densely branched morphology with a high surface area/volume ratio and high rates of nutrient uptake, and its high tolerance to temperature and irradiance resulted in an extremely high growth rate (Hiraoka and Oka, 2008; Liu et al., 2010; Luo et al., 2012; Cui et al., 2015). To sum up, asexual, vegetative reproduction, and the high growth rate of thalli was the main reason accounting for the high biomass of green tide during the floating period.

The results of the present study indicated that the reproductive pattern of the green tides was determined by their formation period in the Subei Shoal where micropropagules germinated and developed into thalli. The total biomass of the Yellow Sea green tides varies over years, with a maximum coverage area of 55 699, 18 237, and 60 594 km^2, and a maximum distribution area of 508, 192, and 1 746 km^2 in 2019, 2020, and 2021, respectively (Sun et al., 2022). The variation of biomass over different years was thought to be affected by the different implementations of source-control measures and the competition effect with other bloom-forming species, such as *Sargassum* (Song et al., 2022). In the current study, we also observed differences in the dominant reproductive mode among different years. Although the sample size in our study was relatively small, our results suggest that the reproductive strategies of *U. prolifera* in different years warrant continuous attention. Further investigations with larger samples are needed to reveal whether the reproductive strategies are related to the scale of the green tide. This will aid in the advancement of the management of the Yellow Sea green tides in the future.

4.2 Genetic diversity of *U. prolifera* among different sampling sites and different years.

In order to assess the genetic diversity among populations, pairwise F_{ST} values were estimated using genotype results of the haploid gametophytes (Table 2). Besides, AMOVA analyses were performed and a UPGMA dendrogram was generated (Table 3; Fig.5). The F_{ST} values between samples collected from Subei Shoal and Qingdao in 2019 implied a very low genetic divergence among the two groups in the same year. Previous studies using single genetic markers (e.g., ITS, SCAR) have demonstrated that *U. prolifera* from the Subei Shoal was the major source of floating green algae in the Yellow Sea (Pang et al., 2010; Zhang et al., 2018). Our result was comparable to the previous studies, and for the first time, provided genome-wide scale evidence for the origin of the large-scale green tide on the Qingdao coast.

Although AMOVA analyses suggested that the genetic variation of sequenced samples was mainly caused by differentiation within groups, pairwise F_{ST} values indicated higher genetic divergence among samples collected in different years. According to the UPGMA dendrogram, despite there being no clear general clustering pattern of samples from different years, most of the individuals collected in 2020 and 2021 were clustered together. Compared

with samples in 2019, fewer individual thalli were obtained in 2020 and 2021. The relatively small sampling size of the two years might not be representative of the genetic diversity of the entire population, therefore, might be a potential factor leading to the overestimate of the genetic divergence of *U. prolifera* among different years. Additionally, human control measures of the green tide, such as the employment of source prevention practices and collection of floating green algae in the major waterways in the Subei Shoal and Yellow Sea (Xiao et al., 2021; Song et al., 2022; Sun et al., 2022), might be another important reason related to the estimated genetic divergence among samples over years. These control measures account for different biomass of the green tide (Song et al., 2022; Sun et al., 2022), therefore, may lead to the change in the genetic population structures over different years. In future studies, investigation with a larger sample size and complete spatial scale are needed to fully assess the population structure and genetic diversity of *U. prolifera*.

5 CONCLUSION

Ulva prolifera is an algal species with great significance in the research area of marine ecology because its specific biological feature of high environmental adaption and advantages in reproductivity which lead to extremely high biomass in green tide. The current study is the first report to identify genome-wide SNPs for *U. prolifera* collected in three consecutive years from 2019 to 2021. Our results have revealed that the reproductive patterns of thalli were various in different years and might be related to the interannual difference in chemical and physical conditions during the formation of the green tide. The flexibility in the reproductive mode contributes to the environmental adaption in different years and might be a critical reason for the annual bloom of *U. prolifera*. The reproductive patterns of thalli in the same year were identical among different sites, implying that asexual, vegetative reproduction and the high growth rate of thalli was the main reason accounting for the high biomass of green tide during the floating period. Population analysis based on SNP genotype results indicated a very low genetic diversity among samples from Subei Shoal and the Qingdao coast in the same year and a higher divergence among samples in different years. These results provide genome-wide scale evidence for the origin of the large-scale green tide on the Qingdao coast, and suggest that future studies of investigation with a larger sample size and complete spatial scale are needed to fully assess the population structure and genetic diversity of *U. prolifera*.

6 DATA AVAILABILITY STATEMENT

This sequencing data have been deposited at the NCBI Nucleotide Short Read Archive (SRA) under accession No. PRJNA948913.

7 ACKNOWLEDGMENT

We would like to thank all colleagues for their assistance with sample collections.

References

Cao J X, Liu J L, Zhao S et al. 2023. Advances in the research on micropropagules and their role in green tide outbreaks in the Southern Yellow Sea. *Marine Pollution Bulletin*, **188**: 114710, https://doi.org/10.1016/j. marpolbul. 2023.114710.

Cui J J, Shi J T, Zhang J H et al. 2018. Rapid expansion of *Ulva* blooms in the Yellow Sea, China through sexual reproduction and vegetative growth. *Marine Pollution Bulletin*, **130**: 223-228, https://doi.org/10.1016/j. marpolbul. 2018.03.036.

Cui J J, Zhang J H, Huo Y Z et al. 2015. Adaptability of free-floating green tide algae in the Yellow Sea to variable temperature and light intensity. *Marine Pollution Bulletin*, **101**(2): 660-666, https://doi.org/10.1016/j.marpolbul.2015. 10.033.

Dan A, Hiraoka M, Ohno M et al. 2002. Observations on the effect of salinity and photon fluence rate on the induction of sporulation and rhizoid formation in the green alga *Enteromorpha prolifera* (Müller) J. Agardh (Chlorophyta, Ulvales). *Fisheries Science*, **68**(6): 1182-1188, https://doi.org/10.1046/j.1444-2906.2002.00553.x.

Danecek P, Auton A, Abecasis G et al. 2011. The variant call format and VCFtools. *Bioinformatics*, **27**(15): 2156-2158, https://doi.org/10.1093/bioinformatics/btr330.

Duan W J, Guo L X, Sun D et al. 2012. Morphological and molecular characterization of free-floating and attached green macroalgae *Ulva* spp. in the Yellow Sea of China. *Journal of Applied Phycology*, **24**(1): 97-108, https://doi.org/10.1007/s10811-011-9654-7.

Fifer J E, Yasuda N, Yamakita T et al. 2022. Genetic divergence and range expansion in a western north Pacific coral. *Science of the Total Environment*, **813**: 152423, https://doi.org/10.1016/j.scitotenv.2021.152423.

Fu X T, Dou J Z, Mao J X et al. 2013. RADtyping: an integrated package for accurate de novo codominant and dominant RAD genotyping in mapping populations. *PLoS One*, **8**(11): e79960, https://doi.org/10.1371/journal.pone.0079960.

Goudet J. 2005. HIERFSTAT, a package for R to compute and test hierarchical *F*-statistics. *Molecular Ecology Notes*,

5(1): 184-186, https://doi.org/10.1111/j.1471-8286.2004.00828.x.

Han W, Chen L P, Zhang J H et al. 2013. Seasonal variation of dominant free-floating and attached *Ulva* species in Rudong coastal area, China. *Harmful Algae*, **28**: 46-54, https://doi.org/10.1016/j.hal.2013.05.018.

Hiraoka M, Dan A, Shimada S et al. 2003. Different life histories of *Enteromorpha prolifera* (Ulvales, Chlorophyta) from four rivers on Shikoku Island, Japan. *Phycologia*, **42**(3): 275-284, https://doi.org/10.2216/i0031-8884-42-3-275.1.

Hiraoka M, Higa M. 2016. Novel distribution pattern between coexisting sexual and obligate asexual variants of the true estuarine macroalga *Ulva prolifera*. *Ecology and Evolution*, **6**(11): 3658-3671, https://doi.org/10.1002/ece3.2149.

Hiraoka M, Oka N. 2008. Tank cultivation of *Ulva prolifera* in deep seawater using a new "germling cluster" method. *Journal of Applied Phycology*, **20**(1): 97-102, https://doi.org/10.1007/s10811-007-9186-3.

Ichihara K, Yamazaki T, Miyamura S et al. 2019. Asexual thalli originated from sporophytic thalli via apomeiosis in the green seaweed *Ulva*. *Scientific Reports*, **9**(1): 13523, https://doi.org/10.1038/s41598-019-50070-x.

Kamvar Z N, Tabima J F, Grünwald N J. 2014. *Poppr*: a R package for genetic analysis of populations with clonal, partially clonal, and/or sexual reproduction. *PeerJ*, **2**: e281, https://doi.org/10.7717/peerj.281.

Keesing J K, Liu D Y, Fearns P et al. 2011. Inter- and intra-annual patterns of *Ulva prolifera* green tides in the Yellow Sea during 2007-2009, their origin and relationship to the expansion of coastal seaweed aquaculture in China. *Marine Pollution Bulletin*, **62**(6): 1169-1182, https://doi.org/10.1016/j.marpolbul.2011.03.040.

Kumar S, Stecher G, Tamura K. 2016. MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Molecular Biology and Evolution*, **33**(7): 1870-1874, https://doi.org/10.1093/molbev/msw054.

Li R Q, Yu C, Li Y R et al. 2009. SOAP2: an improved ultrafast tool for short read alignment. *Bioinformatics*, **25**(15): 1966-1967, https://doi.org/10.1093/bioinformatics/btp336.

Li Y, Huang H J, Li H Y et al. 2016. Genetic diversity of *Ulva prolifera* population in Qingdao coastal water during the green algal blooms revealed by microsatellite. *Marine Pollution Bulletin*, **111**(1-2): 237-246, https://doi.org/10.1016/j.marpolbul.2016.07.001.

Lin A P, Shen S D, Wang J W et al. 2008. Reproduction diversity of *Enteromorpha prolifera*. *Journal of Integrative Plant Biology*, **50**(5): 622-629, https://doi.org/10.1111/j.1744-7909.2008.00647.x.

Liu D Y, Keesing J K, He P M et al. 2013. The world's largest macroalgal bloom in the Yellow Sea, China: formation and implications. *Estuarine*, *Coastal and Shelf Science*, **129**: 2-10, https://doi.org/10.1016/j.ecss.2013.05.021.

Liu D Y, Zhou M J. 2018. Green tides of the yellow sea: massive free-floating blooms of *Ulva prolifera*. *In*: Glibert P M, Berdalet E, Burford M A et al eds. Global Ecology and Oceanography of Harmful Algal Blooms. Springer, Cham. p.317-326, https://doi.org/10.1007/978-3-319-70069-4_16.

Liu F, Pang S J, Xu N et al. 2010. *Ulva* diversity in the Yellow Sea during the large-scale green algal blooms in 2008-2009. *Phycological Research*, **58**(4): 270-279, https://doi.org/10.1111/j.1440-1835.2010.00586.x.

Liu Q, Yu R C, Yan T et al. 2015. Laboratory study on the life history of bloom-forming *Ulva prolifera* in the Yellow Sea. *Estuarine*, *Coastal and Shelf Science*, **163**: 82-88, https://doi.org/10.1016/j.ecss.2014.08.011.

Liu Q C, Wu C H, Xie W F et al. 2022. Validation of mating type-related markers in *Ulva prolifera* (Ulvophyceae, Chlorophyta) and their detection during various reproductive modes. *Algal Research*, **62**: 102611, https://doi.org/10.1016/j.algal.2021.102611.

Luo M B, Liu F, Xu Z L. 2012. Growth and nutrient uptake capacity of two co-occurring species, *Ulva prolifera* and *Ulva linza*. *Aquatic Botany*, **100**: 18-24, https://doi.org/10.1016/j.aquabot.2012.03.006.

Morand P, Briand X. 1996. Excessive growth of macroalgae: a symptom of environmental disturbance. *Botanica Marina*, **39**(1-6): 491-516, https://doi.org/doi:10.1515/botm.1996.39.1-6.491.

Pang S J, Liu F, Shan T F et al. 2010. Tracking the algal origin of the *Ulva* bloom in the Yellow Sea by a combination of molecular, morphological and physiological analyses. *Marine Environmental Research*, **69**(4): 207-215, https://doi.org/10.1016/j.marenvres.2009.10.007.

Ruocco M, Jahnke M, Silva J et al. 2022. 2b-RAD Genotyping of the seagrass *Cymodocea nodosa* along a latitudinal cline identifies candidate genes for environmental adaptation. *Frontiers in Genetics*, **13**: 866758, https://doi.org/10.3389/fgene.2022.866758.

Schlitzer R. 2023. Ocean Data View. Available from https://odv.awi.de.

Smetacek V, Zingone A. 2013. Green and golden seaweed tides on the rise. *Nature*, **504**(7478): 84-88, https://doi.org/10.1038/nature12860.

Song M J, Kong F Z, Li Y F et al. 2022. A massive green tide in the yellow sea in 2021: field investigation and analysis. *International Journal of Environmental Research and Public Health*, **19**(18): 11753, https://doi.org/10.3390/ijerph191811753.

Sun Y Q, Yao L L, Liu J L et al. 2022. Prevention strategies for green tides at source in the southern Yellow Sea. *Marine Pollution Bulletin*, **178**: 113646, https://doi.org/10.1016/j.marpolbul.2022.113646.

Wang Q C, Liu Y, Yan L et al. 2021. Genome-wide SNP discovery and population genetic analysis of *Mesocentrotus nudus* in China seas. *Frontiers in Genetics*, **12**: 717764, https://doi.org/10.3389/fgene.2021.717764.

Wang S, Liu P P, Lv J et al. 2016. Serial sequencing of isolength RAD tags for cost-efficient genome-wide profiling of genetic and epigenetic variations. *Nature Protocols*, **11**(11): 2189-2200, https://doi.org/10.1038/nprot.2016.133.

Wang S, Meyer E, McKay J K et al. 2012. 2b-RAD: a simple and flexible method for genome-wide genotyping.

Nature Methods, **9**(8): 808-810, https://doi.org/10.1038/nmeth.2023.

Wang S Y, Huo Y Z, Zhang J H et al. 2018. Variations of dominant free-floating *Ulva* species in the source area for the world's largest macroalgal blooms, China: differences of ecological tolerance. *Harmful Algae*, **74**: 58-66, https://doi.org/10.1016/j.hal.2018.03.007.

Wang Z L, Xiao J, Fan S L et al. 2015. Who made the world's largest green tide in China?—An integrated study on the initiation and early development of the green tide in Yellow Sea. *Limnology and Oceanography*, **60**(4): 1105-1117, https://doi.org/10.1002/lno.10083.

Wu C H, Jiang P, Zhao J et al. 2018. High efficiency of protoplast preparation for artificially cultured *Ulva prolifera* (Ulvophyceae, Chlorophyta). *Journal of Oceanology and Limnology*, **36**(5): 1806-1811, https://doi.org/10.1007/s00343-018-7058-0.

Xiao J, Wang Z L, Liu D Y et al. 2021. Harmful macroalgal blooms (HMBs) in China's coastal water: green and golden tides. *Harmful Algae*, **107**: 102061, https://doi.org/10.1016/j.hal.2021.102061.

Xing Q G, Wu L L, Tian L Q et al. 2018. Remote sensing of early-stage green tide in the Yellow Sea for floating-macroalgae collecting campaign. *Marine Pollution Bulletin*, **133**: 150-156, https://doi.org/10.1016/j. marpolbul. 2018.05.035.

Yamazaki T, Ichihara K, Suzuki R et al. 2017. Genomic structure and evolution of the mating type locus in the green seaweed *Ulva partita*. *Scientific Reports*, **7**(1): 11679, https://doi.org/10.1038/s41598-017-11677-0.

Ye N H, Zhang X W, Mao Y Z et al. 2011. 'Green tides' are overwhelming the coastline of our blue planet: taking the world's largest example. *Ecological Research*, **26**(3): 477-485, https://doi.org/10.1007/s11284-011-0821-8.

Yuan C, Xiao J, Zhang X L et al. 2022. A new assessment of the algal biomass of green tide in the Yellow Sea. *Marine Pollution Bulletin*, **174**: 113253, https://doi.org/10.1016/j.marpolbul.2021.113253.

Zhang J H, Kim J K, Yarish C et al. 2016. The expansion of *Ulva prolifera* O. F. Müller macroalgal blooms in the Yellow Sea, P R China, through asexual reproduction. *Marine Pollution Bulletin*, **104**(1-2): 101-106, https://doi.org/10.1016/j.marpolbul.2016.01.056.

Zhang J J, Kobert K, Flouri T et al. 2014a. PEAR: a fast and accurate Illumina paired-end reAd mergeR. *Bioinformatics*, **30**(5): 614-620, https://doi.org/10.1093/bioinformatics/btt593.

Zhang L, Wang G L, Liu C et al. 2014b. Development and utility of EST-SSR markers in *Ulva prolifera* of the south Yellow Sea. *Acta Oceanologica Sinica*, **33**(10): 105-113, https://doi.org/10.1007/s13131-014-0545-z.

Zhang Q C, Liu Q, Kang Z J et al. 2015. Development of a fluorescence *in situ* hybridization (FISH) method for rapid detection of *Ulva prolifera*. *Estuarine*, *Coastal and Shelf Science*, **163**: 103-111, https://doi.org/10.1016/j.ecss.2014.09.004.

Zhang Q C, Yu R C, Chen Z F et al. 2018. Genetic evidence in tracking the origin of *Ulva prolifera* blooms in the Yellow Sea, China. *Harmful Algae*, **78**: 86-94, https://doi.org/10.1016/j.hal.2018.08.002.

Zhang Y Y, He P M, Li H M et al. 2019. *Ulva prolifera* green-tide outbreaks and their environmental impact in the Yellow Sea, China. *National Science Review*, **6**(4): 825-838, https://doi.org/10.1093/nsr/nwz026.

Zhao J, Jiang P, Liu Z Y et al. 2011. Genetic variation of *Ulva* (*Enteromorpha*) *prolifera* (Ulvales, Chlorophyta)—the causative species of the green tides in the Yellow Sea, China. *Journal of Applied Phycology*, **23**(2): 227-233, https://doi.org/10.1007/s10811-010-9563-1.

Zhao J, Jiang P, Qin S et al. 2015. Genetic analyses of floating *Ulva prolifera* in the Yellow Sea suggest a unique ecotype. *Estuarine*, *Coastal and Shelf Science*, **163**: 96-102, https://doi.org/10.1016/j.ecss.2015.05.027.

Zhao X H, Cui J J, Zhang J H et al. 2019. Reproductive strategy of the floating alga *Ulva prolifera* in blooms in the Yellow Sea based on a combination of zoid and chromosome analysis. *Marine Pollution Bulletin*, **146**: 584-590, https://doi.org/10.1016/j.marpolbul.2019.07.018.

Electronic supplementary material

Supplementary material (Supplementary Tables S1–S2) is available in the online version of this article at https://doi.org/10.1007/s00343-023-3060-2.