

RESEARCH ARTICLE

High genetic differentiation of Indo-Pacific humpback dolphins (*Sousa chinensis*) along the Asian Coast of the Pacific Ocean

Yufei Dai^{1,2,3}  | Watchara Sakornwimon⁴ | Rachawadee Chantra⁵ | Liyuan Zhao^{1,2,3} | Fuxing Wu^{1,2,3} | Reyilamu Aierken¹ | Kongkiat Kittiwattanawong⁶ | Xianyan Wang^{1,2,3}

¹Laboratory of Marine Biology and Ecology, Third Institute of Oceanography, Ministry of Natural Resources, Xiamen, China

²Key Laboratory of Marine Ecological Conservation and Restoration, Ministry of Natural Resources, Xiamen, China

³Fujian Provincial Key Laboratory of Marine Ecological Conservation and Restoration, Xiamen, China

⁴Marine and Coastal Resources Research Center, The Central Gulf of Thailand, Chumphon, Thailand

⁵Marine and Coastal Resources Research Center, The Upper Gulf of Thailand, Samut Sakhon, Thailand

⁶Phuket Marine Biological Research Center, Phuket, Thailand

Correspondence

Kongkiat Kittiwattanawong, Phuket Marine Biological Research Center, Phuket 83000, Thailand.

Email: kkongkiat@gmail.com

Xianyan Wang, Laboratory of Marine Biology and Ecology, Third Institute of Oceanography, Ministry of Natural Resources, Xiamen 361005, China.

Email: wangxianyan@tio.org.cn

Funding information

China-ASEAN Maritime Cooperation Fund, Grant/Award Number: HX04-210901; National Natural Science Foundation for Young Scholars of China, Grant/Award Number: 41806135; Natural Science Foundation of Fujian Province, Grant/Award Number: 2021J01510; Scientific Research Foundation of Third Institute of Oceanography, MNR, Grant/Award Number: 2020017 and 2019018

Abstract

The Indo-Pacific humpback dolphin (*Sousa chinensis*) is a vulnerable marine mammal species that inhabits shallow, coastal waters from Southeast China, southward throughout Southeast Asia, and westward around the Bay of Bengal to eastern India. Polymorphic microsatellites are useful for elucidating ecological and population genetics-related questions. Here, 18 new polymorphic microsatellites were developed from *S. chinensis* genomic DNA by Illumina MiSeq sequencing. Population genetic analyses were conducted on 42 *S. chinensis* individuals from three geographic locations, including the Xiamen Bay of China, the Western Gulf of Thailand, and Andaman Sea. Our microsatellite data revealed a strong and significant population structure among the three sampling regions (overall $F_{ST} = 0.371$, $p = .001$). Pairwise mutual information index also demonstrated high levels of genetic differentiation between different region pairs (values range from 0.272 to 0.339, $p < .001$). Moreover, Structure analysis inferred three genetic clusters, with the high assignment probabilities of 95.92%, 99.47%, and 99.68%, respectively. Principal coordinate analysis plots of individuals divided entire genotypes into three clusters, indicating high level of genetic differentiation. Our results indicated the strong genetic structure in *S. chinensis* populations is a result of geographic distances. Other factors such as environmental variables, anthropogenic interference, and social behavior may also have contributed to population differentiation.

KEYWORDS

cetacean, genetic diversity, humpback dolphin, microsatellite genotyping, population differentiation

TAXONOMY CLASSIFICATION

Conservation genetics

Yufei Dai and Watchara Sakornwimon contributed equally to this work.

This is an open access article under the terms of the [Creative Commons Attribution](https://creativecommons.org/licenses/by/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2022 The Authors. *Ecology and Evolution* published by John Wiley & Sons Ltd.

1 | INTRODUCTION

The Indo-Pacific humpback dolphin (*Sousa chinensis*) is widely distributed in shallow, coastal waters from Southeast China, southward throughout Southeast Asia, and westward around the Bay of Bengal to eastern India (Jefferson & Curry, 2015; Jefferson & Rosenbaum, 2014; Mendez et al., 2013). Recently, this species has been classified as “vulnerable” on the International Union for Conservation of Nature’s RedList of Threatened Species based on an inferred population size reduction (Jefferson et al., 2017). Primary threats to this vulnerable species are incidental mortality caused by intensive fishing efforts using entangling gear, as well as ongoing habitat loss and degradation due to coastal development. Vessel collisions and environmental contamination may also be significant threats in some areas (Jefferson & Smith, 2016; Jefferson et al., 2017). For long-lived animals with late maturation and low reproductive rates such as *S. chinensis*, these threats often have resulted in priority conservation status being afforded to a number of small and fragmented populations (Brown et al., 2014).

Understanding genetic diversity and population structure is essential for the assessment of conservation status and effective management of a species, especially for inshore dolphins whose isolated populations are highly affected by human activities (Brown et al., 2014; Jefferson et al., 2009; Mace & Lande, 1991). Current information of population parameters is available for only a few sites in China, Malaysia, Thailand, and Bangladesh (Jefferson et al., 2017). Photo-identification catalogues that would allow the identification of individuals and comparisons between these regions individuals are lacking. Although previous genetic studies revealed strong population structure in different *S. chinensis* communities (Amaral et al., 2017; Mendez et al., 2013), the analysis of genetic samples from the Asian coast of the Pacific and the Indian Ocean is still lacking. Recent population-level analyses based on a single locus of the mitochondrial DNA (mtDNA) control region have detected significant genetic differentiation between most of the geographic populations in both Chinese and Thai waters (Zhao et al., 2021). However, limited conclusions can be drawn when relying on a single mtDNA locus as a marker, because mtDNA is maternally inherited and often has higher mutation rates than nuclear DNA (Mendez et al., 2013). Analyses using additional molecular markers should be conducted.

Microsatellites are widely used and have gradually become an important genetic marker, because highly polymorphic microsatellites are useful for elucidating molecular ecology- and population genetics-related issues such as migration rates, bottlenecks, and kinship (De Barba et al., 2017; Selkoe & Toonen, 2006). These markers are usually short in length (100–300 bp), and they can still be amplified with polymerase chain reaction (PCR), even when using some poor-quality samples caused by DNA degradation (Taberlet et al., 1999). Microsatellite studies on cetaceans have successfully used genomic DNA extracted from sloughed skin or tissues collected from decomposing stranded animals (Valsecchi & Amos, 1996). However, using low-quality DNA samples may lead to low amplification success rates and high rates of genotyping errors, such as

allelic dropouts and other allele-like artifacts that are generated by amplification (Bonin et al., 2004; Pompanon et al., 2005).

Traditional methods for microsatellite isolation include construction of an enriched library followed by cloning and Sanger sequencing, which are both expensive and extremely laborious and time-consuming (Zane et al., 2002). With the development of next-generation sequencing technologies, isolation of species-specific microsatellite loci has become more convenient and efficient (Kumar & Kocour, 2017; Vaini et al., 2019). Paired-end sequencing on the Illumina platform is currently the most commonly used approach for microsatellite isolation (González-Castellano et al., 2018).

Several genetic and genomic studies have been published on humpback dolphins (Gui et al., 2013; Jia et al., 2019; Ming, Jian, Yu, Yu, et al., 2019; Ming, Jian, Yu, Wang, et al., 2019; Zhang et al., 2020), but only a limited number of microsatellite sequences are reported in *S. chinensis*, and some of those are shown to have low polymorphism (Chen & Yang, 2009; Lin et al., 2012). In addition, there are more dinucleotide microsatellites among the existing markers. Using such short tandem repeat motifs may produce a large amount of strand slippage during PCR and increase the likelihood of stutter bands and genotyping errors (Pei et al., 2018; Zalapa et al., 2012). Compared to dinucleotide markers, tetrameric and pentameric markers have lower stutter slippage efficiency and clearer peak discrimination during PCR amplification and genotyping (Gill et al., 2005; Pei et al., 2018).

Therefore, the objectives of this study were to: (1) isolate tetra-, penta-, and hexa-nucleotide microsatellites from *S. chinensis* genome sequences using Illumina sequencing; (2) evaluate genetic diversity of *S. chinensis* using samples obtained from three different sampling sites along the Asian coast of the Pacific and the Indian Ocean; and (3) infer population structure among the three sampled locations.

2 | MATERIALS AND METHODS

2.1 | Ethics permits

All fieldwork was conducted under permits from the Ministry of Agriculture and Rural Affairs of China, and with approval from the Department of Marine and Coastal Resources of Thailand. The relevant CITES Permits (No. 2018CN/IC000475/CH) and HS Code (4103909010) were obtained for import of samples. All samples were collected from dead, stranded individuals.

2.2 | Sample collection and DNA extraction

Muscle or skin tissues of 55 dead stranded *S. chinensis* were collected from 2010 to 2018, and included 15 individuals obtained along the coast of Xiamen in China (XM), 19 individuals collected from the Western Gulf of Thailand (WG), and 21 individuals collected along the Andaman Sea coast (AS) (Figure 1). Geographic coordinates for most individuals from the three sampling regions were recorded at

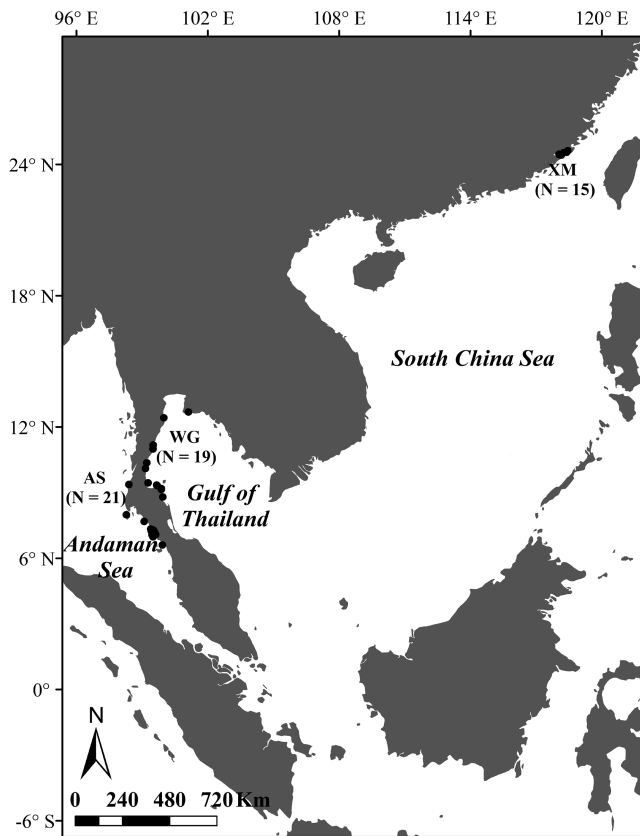


FIGURE 1 Locations of sample collection in the different geographical regions analyzed in this study. The three sampling sites are including the Xiamen Bay of China (XM, $N = 15$), the western Gulf of Thailand (WG, $N = 19$), and the Andaman Sea site (AS, $N = 21$)

the time of collection (Appendix S1). Genomic DNA from minced tissue samples were extracted using DNeasy blood and tissue extraction kits (QIAGEN) according to the manufacturer's protocol.

2.3 | Microsatellite selection and multiplex PCR design

Purified genomic DNA was quantified by TBS-380 fluorometer (Turner BioSystems Inc.). High-quality DNA ($OD_{260/280} = 1.8-2.0$, $>1 \mu\text{g}$) of a single individual collected at XM was used to generate an enriched library and sequenced on the Illumina MiSeq platform by Majorbio Bio-Pharm Technology Co., Ltd. The detailed procedures were as follows. First, the DNA sample was sheared into 400–500-bp fragments using a Covaris M220 Focused Acoustic Shearer following the manufacturer's protocol. Then, an Illumina sequencing library was prepared from the sheared fragments using the NEXTflex™ Rapid DNA-Seq Kit. Briefly, 5' ends were first end-repaired and phosphorylated. Next, the 3' ends were A-tailed and ligated to the sequencing adapters. The third step was to enrich the adapter-ligated products using PCR. Finally, the prepared library was used for paired-end Illumina sequencing ($2 \times 150 \text{ bp}$) on an Illumina HiSeq X Ten machine. After filtering low quality and

duplicated sequences and removing adapter-related reads, a total of 2,174,959 clean reads were assembled using SOAPdenovo version 2.04 software (Luo et al., 2012). At last, 1,398,738 contigs including 95,161 large contigs ($>1000 \text{ bp}$) were obtained, with an average GC content of 55.49% and contig N50 of 1513 bp (Table 1).

The assembled data were searched for tetra-, penta-, and hexa-nucleotide microsatellite motifs using MSATCOMMANDER version 0.8.2 (Faircloth, 2008). The searching parameters were a minimum of 10 repeats for tetra-nucleotide motifs, and six repeats for the other two repeat classes. Only 'perfect-type' microsatellite sequences (pure repeats) with a flanking region of at least 30 bp on each side were selected. PCR primers for 162 available microsatellites were designed using Primer3 (Rozen & Skaletsky, 2000). After primers design and PCR genotyping, a total of 18 polymorphic microsatellite loci (Table 2) with 'perfect-type' and long tandem repeat motifs were allocated into 6 multiplex PCR panels using software MPprimer (Shen et al., 2010), based on annealing temperature, complementarity of primer pairs, and allele size range (Figure 2).

2.4 | Microsatellite genotyping

The 5' end of each forward primer was labeled with a fluorescent dye (6-FAM, VIC or NED). The total PCR volume ($20 \mu\text{l}$) consisted of approximately 50 ng of genomic DNA, 1×Multiplex PCR Kit (Takara), 0.2 μM of each primer (forward and reverse), and ddH_2O added to make up the final volume (Table 2). PCR conditions involved an initial denaturation step at 94°C for 3 min, followed by 32 cycles of 94°C for 30 s, the specific annealing temperature (Table 2) for 90 s, extension at 72°C for 60 s, and a final extension for 30 min at 60°C . Fragment analysis was performed on the PCR products on an ABI 3730XL automated DNA sequencer (Applied Biosystems), using GeneScan LIZ 500 as the internal size standard. Allele sizes were automatically scored with GeneMapper version 4.1 (Applied Biosystems) and manually checked. We used the same individual as a positive control for genotyping each locus separately in each multiplex PCR panel to ensure consistent amplification of alleles. A negative control without DNA template was also used in each PCR batch to detect possible contamination during PCR amplification. Moreover, genotypes from six individuals

TABLE 1 Summary of dataset assemblies for *S. chinensis* through Illumina sequencing

Assembly statistics	Data statistics
Total number of contigs	1,398,738
Total bases of contigs	724,659,648
Total number of large contigs ($>1000 \text{ bp}$)	95,161
Total bases of large contigs	148,023,926
Contig N50 (bp)	1513
G+C content (percentage)	55.491%

TABLE 2 Multiplex design information for all the tested 18 microsatellite loci in *S. chinensis*. Primer concentration was 0.2 μ M

Multiplex panels	Locus	Repeat motif	Primer sequences (5'–3')	Primer dosage (μ l)	Annealing temperature ($^{\circ}$ C)	GenBank accession number	Fluorescent label
Multiplex 1	Sch5878	(CAAC) ₁₂	F: TCTCCAGTGTGGCTCTT R: ACATTTTGAAGGCAAGCTG	0.20 0.20	61	MK766860	6-FAM
	Sch6660	(AAGG) ₁₃	F: CTGAGTGGTCTCAAGGAG R: TCTGCTGACATGCCTCACTC	0.18 0.18		MK766861	VIC
	Sch443	(CCAT) ₁₂	F: GGACTACAAGAAGCTGGGCA R: CTGGTGCCTGTAGCTGTTGT	0.18 0.18		MK766850	NED
Multiplex 2	Sch10207	(CATC) ₁₂	F: CCCTCTTGCTCTCTCCCT R: TGTCTATTGTACAGCAGGATGGA	0.15 0.15	62	MK766870	6-FAM
	Sch843	(AAAT) ₁₁	F: GAGAAACATTTGTCTAAGTGCTCTG R: GAACGCAGATCCTAACGTCTAATTAG	0.15 0.15		MK766851	VIC
	Sch7424	(ATGG) ₁₃	F: GGAAGGGTGGATGGTTAGGT R: ATGTTCCCTGAGGATTGTGC	0.15 0.15		MK766864	NED
Multiplex 3	Sch7357	(ATGG) ₁₁	F: CAGTGCCTCGAACAGAGATTG R: AAGTATCCACACCCATCCA	0.15 0.15	61	MK766863	6-FAM
	Sch193	(AGAGA) ₁₂	F: GTATGGAAGGAAGGGAGGGA R: CAAACTAAGGAAGCAAATGCAG	0.20 0.20		MK766846	VIC
	Sch8186	(CCAT) ₁₁	F: CCCACAGAAGTCAAGCATCA R: CTGGAATCTGGGGTGAGAAAT	0.15 0.15		MK766865	NED
Multiplex 4	Sch123	(CCTAAC) ₇	F: GGAAGCCAGGTACCACTGTTG R: TGAGGACAGCACAGACCAGAG	0.15 0.15	63	MK766845	6-FAM
	Sch4657	(TTCC) ₁₁	F: TGTCCATGCAAGCGTAAATC R: AGTGTGGCATTCTCCAGC	0.15 0.15		MK766855	VIC
	Sch5373	(GATG) ₁₁	F: GGCTCCAGAGCTTGTGATCT R: GGAAGTCCATCTCCCTCTCC	0.15 0.15		MK766858	NED
Multiplex 5	Sch9144	(ATCT) ₁₄	F: TAGAGCCTGCATGAGTGTGG R: CAACAGATCAGCGTGCCT	0.15 0.15	59	MK766868	6-FAM
	Sch2513	(CATC) ₁₃	F: GGGTTTACACCTGTCTGCTGT R: TCAACACATCATTGCCGAAT	0.15 0.15		MK766854	VIC
	Sch5685	(AGGA) ₁₂	F: CATTCTCCAGATGTACGTCCA R: CCTCGGGTAAGTCCCTCTTC	0.25 0.25		MK766859	NED
Multiplex 6	Sch974	(GTTTT) ₁₃	F: GCTGAGGATATCAGGGTGGGA R: CAGGGAAGTCCCAGAAATCA	0.15 0.15	59	MK766849	6-FAM
	Sch5094	(TCTA) ₁₁	F: CTGGGGTTTCTAGCTTGCAG R: ATTCTCCAGAGGAACCAGCA	0.15 0.15		MK766857	VIC
	Sch8947	(CTAT) ₁₂	F: GGGAAAGATGCCAATCTGAA R: CGTACCGCAACAAAGAGTGA	0.25 0.25		MK766867	NED

(>10%) were randomly retested to estimate genotyping error rates for across all 18 loci. Samples of each of the six samples were reamplified and resequenced.

2.5 | Statistical analysis

GenAIEx version 6.501 (Peakall & Smouse, 2006, 2012) was used to estimate the genetic parameters for each sampling location,

including the number of alleles per locus (N_a), effective number of alleles (eN_a), observed heterozygosity (H_o), expected heterozygosity (H_e), unbiased expected heterozygosity (uH_e), and Shannon's information index (I). GENEPOP version 4.0.7 (Rousset, 2008) was used to test departure from Hardy–Weinberg equilibrium (HWE) and linkage disequilibrium (LD) among all pairs of loci. The p -value-based multiple testing of Bonferroni sequential correction was performed with Myriads version 1.1 (Carvajal-Rodriguez, 2018). Micro-Checker version 2.2.3 (Van Oosterhout et al., 2004) was used to detect

FIGURE 2 The 18 polymorphic microsatellite markers allocated in six multiplex PCR panels. Each locus is shown by allelic range, and different colors represent different fluorescent dyes (FAM: blue, VIC: green, NED: black)

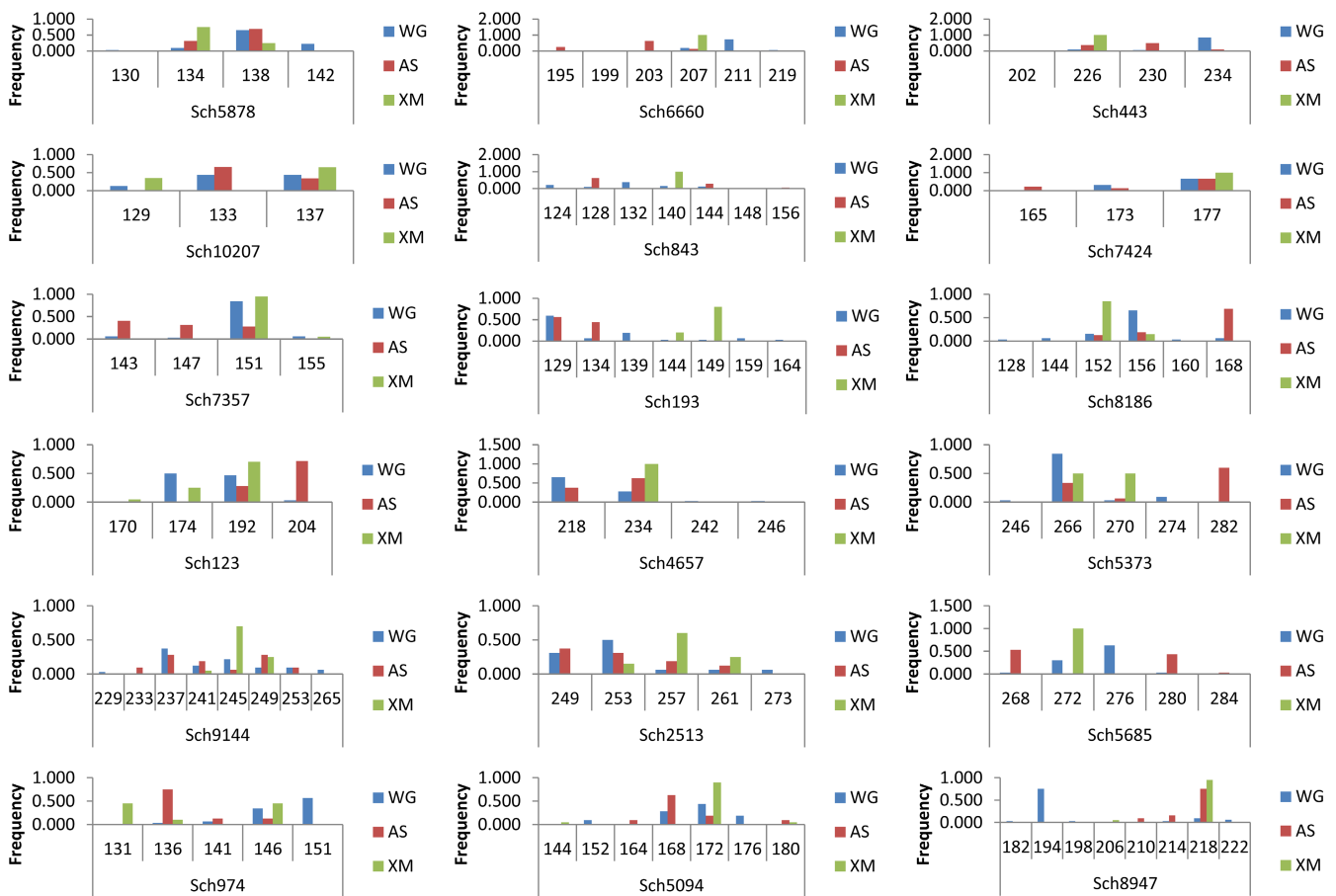
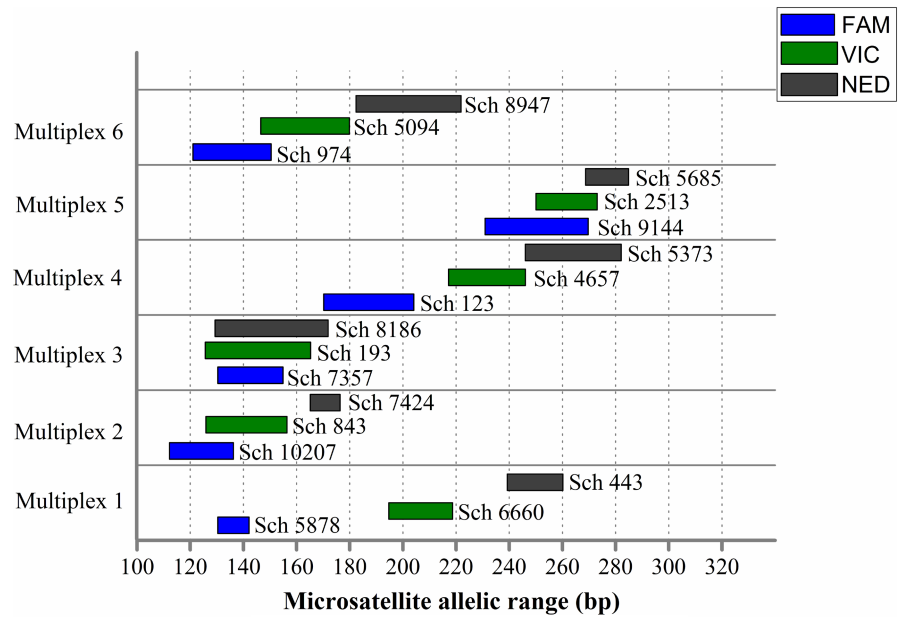


FIGURE 3 Allele Frequencies with Graphs by each sampling location and each locus for microsatellite data. Allele frequencies displayed for 18 polymorphic microsatellites, and different colors represent different sampling regions (WG: blue, AS: red, XM: green)

occurrences of null alleles, allelic dropouts, or scoring error for each locus, with 95% confidence intervals.

FSTAT version 2.9.3.2 (Goudet, 1995, update in Feb. 2002) was used to assess the estimator of genetic population differentiation, F_{ST} (Weir & Cockerham, 1984), among different sampling sites based

on 1000 permutations. In addition, the mutual information index (sHua) (Sherwin et al., 2006) between different region pairs was estimated using GenAEx software.

We inferred population structure, using Structure version 2.3.4 (Pritchard et al., 2000), which estimates the number of genetic

clusters (K) based on genotyping data generated from the six multiplex PCR panels. The length of the burn-in period was set to 10^5 iterations, followed by 10^6 in the number of Markov Chain Monte Carlo iterations. The LOCPRIOR model was chosen to infer possible weak population structure with the assistance of sample group information. The number of inferred K was set between 1 and 10, and 20 independent replicates were run for each K value. Subsequently, the Structure Harvester version 0.6.94 (Earl & VonHoldt, 2012) online tool was used to calculate the Delta K value and determine the best number of K clusters (Evanno et al., 2005). CLUMPP version 1.1.2 (Jakobsson & Rosenberg, 2007) was used to summarize the optimal alignment of the 20 replicates for the same K value. The final results were displayed graphically with Distruct version 1.1 (Rosenberg, 2004).

Moreover, a principal coordinate analysis (PCoA) was performed in GenAlEx based on the standardized covariance of the individual-by-individual genetic distance matrix. Mantel analysis (Diniz-Filho et al., 2013) was also used in GenAlEx to test isolation by distance (IBD) by testing for correlation between matrices of individual-by-individual genetic distances and geographic distances measured as the distance between two individuals calculated from sampling location coordinates. We conducted a second IBD analysis excluding the genetic and geographical data for XM samples because of the large geographic distances between XM and other regions. Both tests were run with 999 random permutations in GenAlEx.

3 | RESULTS

3.1 | Available microsatellite data for analysis

For genetic analysis, only DNA samples for which at least 16 out of the 18 loci could be genotyped were included. Therefore, 13 samples of individuals (five from XM, three from WG and five from AS) were discarded because of poor amplification success. For the remaining 42 individuals, no contamination was detected during multiple PCRs, and no genotyping errors were observed when randomly retesting the six individuals. Finally, genotypes of 18 microsatellites for 42 individuals were used in this genetic study, including 10 individuals from XM, 16 from WG, and 16 from AS. There was no evidence of allelic dropouts or scoring errors due to stuttering for any of the 18 loci in all the three sampling locations. Allele frequency distributions of different sampling regions based on 18 polymorphic loci are graphically represented in Figure 3.

3.2 | Genetic diversity

Genetic diversity estimates of N_a , eN_a , I , H_o , H_e , and uH_e values for each locus and each region are summarized in Table 3. The average estimates of N_a , eN_a , and I were highest in WG, with values of 4.500 ± 0.316 , 2.319 ± 0.204 , and 0.991 ± 0.076 , respectively. The highest mean H_o and H_e were in AS, with values 0.497 ± 0.035

and 0.525 ± 0.025 , respectively (Figure 4). ANOVA results revealed N_a values differed significantly from each other among sampled regions (ANOVA: $F_{2,51} = 25.372$, $p < .001$). There were also significant differences among eN_a ($F_{2,51} = 8.430$, $p = .001$), I ($F_{2,51} = 19.031$, $p < .001$) and H_o ($F_{2,51} = 4.850$, $p = .012$) values. Additionally, significant differences were found among H_e ($F_{2,51} = 17.556$, $p < .001$) and uH_e ($F_{2,51} = 16.600$, $p < .001$) from any of the three geographic regions. The results revealed that the levels of genetic diversity detected for *S. chinensis* in XM were the lowest.

3.3 | Hardy–Weinberg equilibrium and linkage disequilibrium

The results of the HWE tests showed no significant deviations ($p > .05$) in XM. Six loci (Sch443, Sch843, Sch4657, Sch5685, Sch6660, and Sch7424) were monomorphic. All the 18 loci showed polymorphism in WG and AS. One locus (Sch5373) in AS and four loci (Sch4657, Sch6660, Sch7357, and Sch8947) in WG showed significant heterozygosity deficits after sequential Bonferroni correction (Table 3). No significant LD was found in any of the 153 pairs of the 18 tested loci after Bonferroni sequential correction for multiple tests.

3.4 | Genetic differentiation

Genetic differentiation among different sampled locations was estimated by comparing pairwise F_{ST} and sH_{ua} values that were calculated based on genetic data from the 18 microsatellites. The estimated F_{ST} revealed significant and strong genetic differentiation (overall $F_{ST} = 0.371$, $p = .001$) among the three geographic regions. The pairwise sH_{ua} ranged from 0.272 to 0.339, which also demonstrated high levels of genetic differentiation between different region pairs ($p < .001$). Estimates of pairwise F_{ST} and sH_{ua} values are presented in Table 4. AMOVA results for the degree of variance in *S. chinensis* individuals are summarized in Table 5. There was 37.08% genetic variance among the three geographic regions, 6.85% variance among individuals within each region, and 56.07% variance within individuals.

Structure Harvester analysis showed a clear peak for Delta K at $K = 3$ (Figure 5a), indicating that there were three clusters based on 42 genotypes of *S. chinensis* individuals in the regions that were sampled. The graphical output by Distruct suggested a division into three distinct clusters with no admixture among populations (Figure 5b). For $K = 3$, a strong genetic structure for the three inferred genetic clusters (XM, WG, and AS) was apparent, with the high assignment probabilities of 95.92%, 99.47%, and 99.68%, respectively. No possible partitions within cluster were detected. PCoA plots of individuals based on standardized covariance of the genetic distance matrix also divided all genotypes into three clusters, indicating a high level of genetic differentiation (Figure 5c). Axis coordinates 1 and 2 accounted for 24.93% and 22.10% of the total variance, respectively.

TABLE 3 Genetic diversity parameters in the three geographic regions. *N* is the sample size, *N_a* is the number of alleles, *eN_a* is the number of effective alleles, *I* is the Shannon's information index, *H_o* is the observed heterozygosity, *H_e* is the expected heterozygosity, *uH_e* is the unbiased expected heterozygosity, *P_{HWE}* is the *p* value of Hardy-Weinberg equilibrium test, * indicates significant departure from Hardy-Weinberg equilibrium after sequential Bonferroni correction (*p* < .003), ND represents not done

Pop	Locus	<i>N</i>	<i>N_a</i>	<i>eN_a</i>	<i>I</i>	<i>H_o</i>	<i>H_e</i>	<i>uH_e</i>	<i>P_{HWE}</i>
WG	Sch5878	16	4.000	2.048	0.939	0.438	0.512	0.528	0.096
	Sch6660	16	4.000	1.796	0.833	0.188	0.443	0.458	0.001*
	Sch443	16	3.000	1.380	0.539	0.250	0.275	0.284	0.246
	Sch10207	16	3.000	2.510	0.983	0.438	0.602	0.621	0.299
	Sch843	16	6.000	4.197	1.580	0.813	0.762	0.786	0.457
	Sch7424	16	3.000	1.889	0.748	0.375	0.471	0.486	0.414
	Sch7357	16	4.000	1.388	0.598	0.063	0.279	0.288	0.000*
	Sch193	16	7.000	2.510	1.295	0.500	0.602	0.621	0.022
	Sch8186	16	6.000	2.151	1.130	0.500	0.535	0.552	0.005
	Sch123	16	3.000	2.124	0.810	0.563	0.529	0.546	0.747
	Sch4657	16	4.000	1.954	0.850	0.500	0.488	0.504	0.000*
	Sch5373	16	4.000	1.384	0.582	0.313	0.277	0.286	0.997
	Sch9144	16	7.000	4.414	1.686	0.813	0.773	0.798	0.144
	Sch2513	16	5.000	2.783	1.230	0.500	0.641	0.661	0.021
	Sch5685	15	4.000	2.027	0.877	0.333	0.507	0.524	0.603
	Sch974	16	4.000	2.276	0.972	0.625	0.561	0.579	0.008
	Sch5094	16	4.000	3.180	1.254	0.625	0.686	0.708	0.267
	Sch8947	16	6.000	1.730	0.936	0.125	0.422	0.435	0.000*
Mean		15.944	4.500	2.319	0.991	0.442	0.520	0.537	
AS	Sch5878	16	2.000	1.753	0.621	0.375	0.430	0.444	0.611
	Sch6660	16	3.000	2.133	0.900	0.563	0.531	0.548	0.948
	Sch443	16	4.000	2.498	1.045	0.625	0.600	0.619	0.289
	Sch10207	16	2.000	1.822	0.643	0.563	0.451	0.466	0.324
	Sch843	16	4.000	2.107	0.932	0.500	0.525	0.542	0.448
	Sch7424	16	3.000	2.024	0.869	0.625	0.506	0.522	0.652
	Sch7357	16	3.000	2.926	1.086	0.625	0.658	0.679	0.548
	Sch193	16	2.000	1.969	0.685	0.375	0.492	0.508	0.341
	Sch8186	16	3.000	1.910	0.831	0.375	0.477	0.492	0.311
	Sch123	16	2.000	1.679	0.594	0.563	0.404	0.417	0.118
	Sch4657	16	2.000	1.882	0.662	0.500	0.469	0.484	0.790
	Sch5373	15	3.000	2.103	0.853	0.133	0.524	0.543	0.002*
	Sch9144	16	6.000	4.655	1.645	0.813	0.785	0.810	0.063
	Sch2513	16	4.000	3.459	1.305	0.438	0.711	0.734	0.007
	Sch5685	16	3.000	2.107	0.806	0.500	0.525	0.542	0.718
	Sch974	16	3.000	1.684	0.736	0.313	0.406	0.419	0.217
	Sch5094	16	4.000	2.256	1.051	0.563	0.557	0.575	0.851
	Sch8947	16	3.000	1.679	0.728	0.500	0.404	0.417	0.620
Mean		15.944	3.111	2.258	0.888	0.497	0.525	0.542	
XM	Sch5878	10	2.000	1.600	0.562	0.500	0.375	0.395	0.292
	Sch6660	10	1.000	1.000	0.000	0.000	0.000	0.000	ND
	Sch443	10	1.000	1.000	0.000	0.000	0.000	0.000	ND
	Sch10207	10	2.000	1.835	0.647	0.700	0.455	0.479	0.089
	Sch843	10	1.000	1.000	0.000	0.000	0.000	0.000	ND

(Continues)

TABLE 3 (Continued)

Pop	Locus	N	Na	eNa	I	Ho	He	uHe	P _{HWE}
	Sch7424	10	1.000	1.000	0.000	0.000	0.000	0.000	ND
	Sch7357	10	2.000	1.105	0.199	0.100	0.095	0.100	0.868
	Sch193	10	2.000	1.471	0.500	0.400	0.320	0.337	0.429
	Sch8186	10	2.000	1.342	0.423	0.300	0.255	0.268	0.577
	Sch123	10	3.000	1.802	0.746	0.600	0.445	0.468	0.607
	Sch4657	10	1.000	1.000	0.000	0.000	0.000	0.000	ND
	Sch5373	10	2.000	2.000	0.693	0.600	0.500	0.526	0.527
	Sch9144	10	3.000	1.802	0.746	0.400	0.445	0.468	0.873
	Sch2513	10	3.000	2.247	0.938	0.600	0.555	0.584	0.343
	Sch5685	10	1.000	1.000	0.000	0.000	0.000	0.000	ND
	Sch974	10	3.000	2.410	0.949	0.600	0.585	0.616	0.989
	Sch5094	10	3.000	1.227	0.394	0.200	0.185	0.195	0.989
	Sch8947	10	2.000	1.105	0.199	0.100	0.095	0.100	0.868
	Mean	10.000	1.944	1.441	0.389	0.283	0.239	0.252	

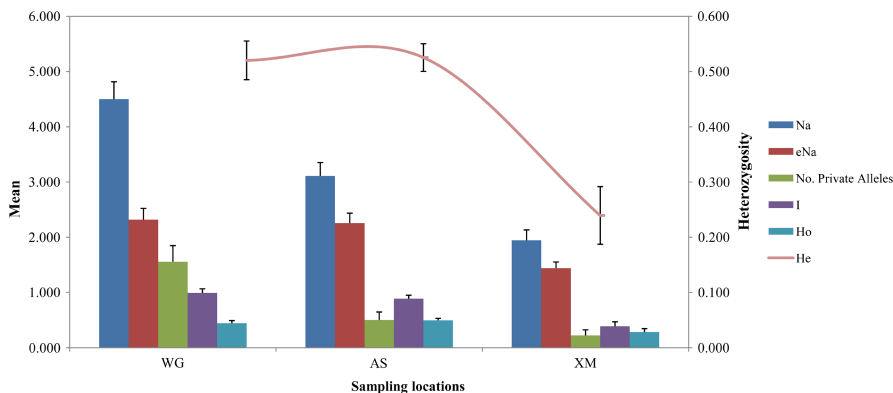


FIGURE 4 Allelic patterns across populations in the geographic distribution of *S. chinensis*. Na is the number of different alleles, eNa is the effective number of alleles, I is the Shannon's information index, Ho is the observed heterozygosity, and He represented by the curve is the expected heterozygosity. The bars represent Standard Error (SE) values

TABLE 4 Matrix of pairwise mutual information index (sHua, above diagonal) and F_{ST} (below diagonal) estimates among the three geographic regions based on microsatellites

Sampling region	XM	WG	AS
XM	-	0.293***	0.339***
WG	0.422***	-	0.272***
AS	0.445***	0.290***	-

*** $p < .001$.

The Mantel tests revealed a positive and significant correlation ($R^2 = .2939$, $p = .0001$) between the individual-by-individual genetic distances and the geographic distances, indicating a pattern of IBD among the three geographic regions (Figure 6a). The result remained positive when geographic coordinates and microsatellite data for XM individuals were removed ($R^2 = .3939$, $p = .0001$; Figure 6b).

4 | DISCUSSION

In this study, 18 polymorphic microsatellites with tetra-, penta-, and hexa-nucleotide repeats were isolated from the genomic DNA of

S. chinensis and used for genetic analysis of 42 *S. chinensis* individuals from three geographic locations. Dimeric and trimeric microsatellite loci were not considered in this work, because stutter bands caused by slipped strand mispairing during PCRs might have occurred at short tandem repeat motifs (Hauge & Litt, 1993; Murray et al., 1993). Compared with dimeric microsatellites, tetrameric and pentameric loci are shown to have lower stutter slippage efficiency and clearer peak discrimination during PCR amplification and genotyping (Pei et al., 2018). Here, we provided 18 novel polymorphic microsatellite markers, which could be useful for future molecular genetics studies on *S. chinensis* and other closely related species.

Estimating the levels of genetic diversity in natural populations provides important information for evaluating species viability (An et al., 2014). Our microsatellite data showed the levels of genetic diversity detected for *S. chinensis* in XM were significantly lower than those in the other two populations. The low genetic diversity in XM may be related to the small sample size. However, previous studies have revealed low levels of molecular diversity in *S. chinensis* in Chinese waters (Chen et al., 2008, 2010; Lin et al., 2010, 2012; Zhang et al., 2020; Zhao et al., 2021). Moreover, mtDNA data of humpback dolphins in China have displayed the lowest genetic diversity, despite having the largest sample size among all sampled

TABLE 5 Analysis of molecular variance of *S. chinensis* in the three sampled regions. df is the degrees of freedom, SS is the sums of squares, MS is the mean squares, Est. Var. is the estimated variance within and among populations

Source of variation	df	SS	MS	Est. Var.	%
Among populations	2	148.144	74.072	2.527	37.078
Among individuals with population	39	185.475	4.756	0.467	6.854
Within individuals	42	160.500	3.821	3.821	56.068
Total	83	494.119		6.816	100

FIGURE 5 Population structure and Principal Coordinates (PCoA) plot among 42 genotypes of *S. chinensis* individuals based on 18 microsatellite markers. (a) Optimal K value determined by Structure Harvester online program; (b) Results of clustering ($K = 3$) calculated by Structure program based on the individual Q -matrix. Black vertical lines separate the three sampling locations, and different colors represent the possible genetic clusters; (c) PCoA plot of individuals based on the standardized covariance of genetic distance matrix. Different shapes and colors represent different sampling regions (WG: red circle, AS: blue square, XM: green triangle)

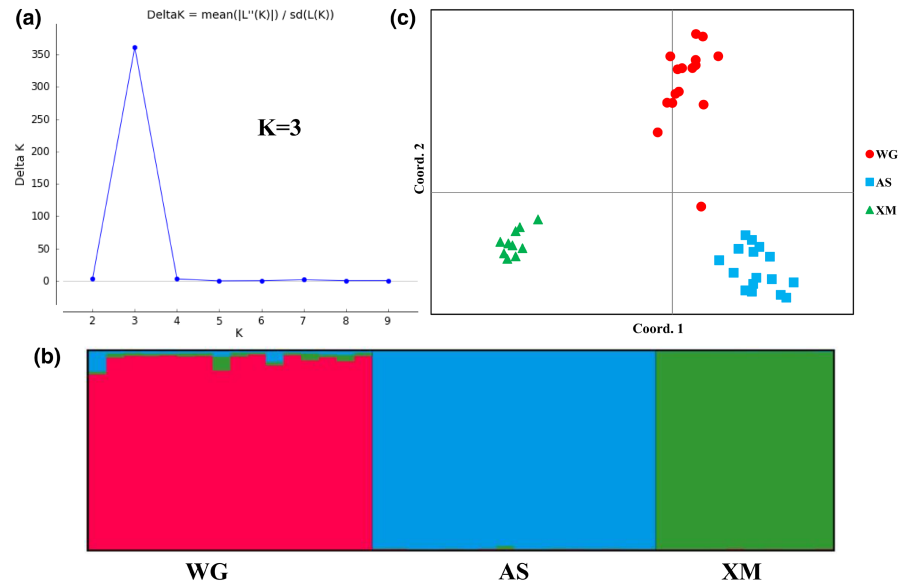
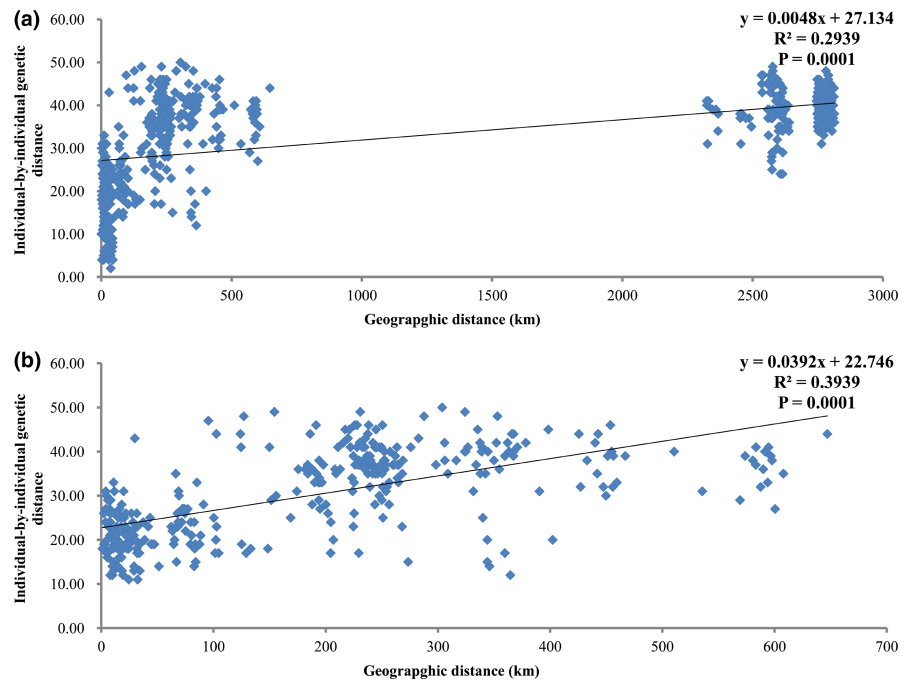


FIGURE 6 Isolation by distance plots using individual-by-individual genetic distances and geographic distances (km) among (a) the three (WG, AS, and XM), and (b) the two (WG and AS) sampling regions. Geographic and microsatellite data of 35 individuals were included in analysis, including 11 individuals from WG, 16 from AS, and 8 from XM



regions (Amaral et al., 2017; Mendez et al., 2013). Historical bottlenecks that lead to a reduction in population size may explain such low genetic diversity in humpback dolphin communities in China (Lin et al., 2010). Besides, habitat loss caused by human development has been largely responsible for the decline of *S. chinensis* populations in southern China (Lin et al., 2016; Wang et al., 2017). In Xiamen Bay,

photo-identification surveys have recorded only a few *S. chinensis* individuals (approximately 60 individuals) across seasons from 2010 to 2015 (Wang et al., 2016; Zeng et al., 2020). By contrast, the minimum population size of *S. chinensis* off Donsak in the Western Gulf of Thailand was estimated to be 193 during the survey period from 2011 to 2013 (Jutapruet et al., 2015). The low genetic diversity in

S. chinensis in XM may be associated with random genetic drift in this small population. More attention should be paid to conservation management of *S. chinensis* in XM, because genetic deterioration and stochastic events can significantly increase the risk of random extinction in the small populations (Karczmarski et al., 2017).

The results revealed a significant strong genetic differentiation in *S. chinensis* among different regions for both F_{ST} and $sHua$ estimates. A previous study revealed that F_{ST} was more appropriate than other estimators when numbers of loci and sample size were limited (Balloux & Goudet, 2002). Compared with F_{ST} , the $sHua$ value is known to be a better estimator of genetic differentiation, and can robustly reflect dispersal over a wide range of population sizes and dispersal rates (Manlik et al., 2019; Sherwin et al., 2017). Similarly, our results of the estimated $sHua$ also showed a high degree of genetic differentiation, which indicates a strong population structure among *S. chinensis* communities of these three sampled regions. Relatively higher degrees of genetic differentiation with higher pairwise F_{ST} or Φ_{ST} values (>0.5) have been reported for *S. chinensis* between populations in China and Thailand based on mtDNA sequence data (Amaral et al., 2017; Mendez et al., 2013; Zhao et al., 2021).

The Mantel analysis showed a pattern of IBD among different sampling regions, which indicated that the observed genetic differentiation in *S. chinensis* communities is a result of geographic distances. Phylogenetic analysis of the *Sousa* genus based on both mtDNA and nuclear DNA markers demonstrated that the Chinese and Thai haplotypes represented one assemblage, although morphological evidence revealed a clear distinction between the two sampling regions (Mendez et al., 2013). Theoretically, there is still potential genetic exchange or contact between *S. chinensis* communities in China and Thailand. However, humpback dolphins are known to prefer shallow waters (less than approximately 25 m in depth) and reside only in coastal (generally within 1–2 km off shore) and estuarine waters (Hung & Jefferson, 2004; Jefferson, 2000; Jefferson et al., 2009). This species shows minimal linear distance movement, with a maximum dispersal distance of 300 km in Chinese waters (Jefferson & Hung, 2004; Wang et al., 2016). Therefore, exchange of *S. chinensis* individuals seems to more likely occur between adjacent communities in China. The geographic distance from the northern South China Sea to the Gulf of Thailand is over 3000 km, which is much greater than the *S. chinensis* individual dispersal distance. Our microsatellite data also revealed that the genetic structure of *S. chinensis* in China and Thailand follows an IBD model, which can explain the strong genetic differentiation among the three sampling locations in this study.

The possible mechanisms that drive the population differentiation and even result in the species boundaries of humpback dolphins include distribution patterns, environmental factors, and behavioral processes (Mendez et al., 2013). Barriers to individual dispersal and gene flow can exist among different *S. chinensis* communities. In the previous studies, significant genetic differentiations of *S. chinensis* between the neighboring regions were also detected in Chinese waters (see Zhang et al., 2020; Zhao et al., 2021). Geographic barriers have been found among humpback

dolphin populations and were caused by different oceanographic features, such as ocean currents, upwelling, bathymetry, sea surface temperature, primary productivity, and salinity (Amaral et al., 2017; Mendez et al., 2011, 2013). Genetic evidence for Indo-Pacific marine fauna has shown distinct genetic lineages of several species in the east and the west, including Indo-Pacific bottlenose dolphins (*Tursiops aduncus*) and humpback dolphins (e.g. Amaral et al., 2017; Keyse et al., 2014; Zhao et al., 2021). Our microsatellite analysis also showed a significant genetic differentiation between WG and AS populations. Historically, the crustal movement of the continental plates and some climatic events may prevent the individual dispersal and gene flow of marine organisms across the Malacca Strait (Zhao et al., 2021). Besides these environmental variables, development of coastal areas may lead to anthropogenic barriers to dispersal and produce the isolated population fragments of inshore dolphins (Brown et al., 2014). It is reported that the coastal development projects have been increasing continuously in the Western Gulf of Thailand (Jutapruet et al., 2015). Therefore, anthropogenic impacts in the coastal habitats of humpback dolphins may be associated with the significant genetic differentiation detected in the Western Pacific and Eastern Indian Ocean sides of the present sampled regions.

Owing to the difficulties associated with biological sample collection, molecular genetics studies on *S. chinensis* have often been restricted to using highly degraded DNA samples from museum specimens or dead stranded individuals, which leads to small sample sizes from limited locations. Recently, an increasing number of molecular genetic studies have obtained genetic materials from humpback dolphins and other small coastal dolphins by minimally invasive sampling methods (e.g., Amaral et al., 2017; Brown et al., 2014; Manlik et al., 2019; Raudino et al., 2019; Zhang et al., 2020; Zhao et al., 2021). Future studies may use these methods to increase the sample size and sampling regions, which can better elucidate the genetic patterns and gene flows among different *S. chinensis* communities.

5 | CONCLUSION

In this study, 18 new microsatellite markers with pure and long tandem repeat motifs were isolated from *S. chinensis* genomic DNA using Illumina MiSeq sequencing. These polymorphic microsatellites were allocated into 6 multiplex PCR panels and successfully obtained genetic data of 42 *S. chinensis* individuals from the Xiamen Bay of China, the Western Gulf of Thailand, and Andaman Sea coast. Our microsatellite evidence, together with mtDNA sequence data reported in the present study area (Zhao et al., 2021), indicate that there are high genetic differentiation among *S. chinensis* communities along the Asian coast of the Pacific and the Indian Ocean. The strong genetic structure in *S. chinensis* populations may be associated with multiple factors such as geographical distribution patterns, environmental variables, anthropogenic interference, and social behavior. These novel polymorphic microsatellite markers will

be useful for future molecular genetics studies on this endangered species and other closely related species.

ACKNOWLEDGMENTS

This work was financially supported by the National Natural Science Foundation for Young Scholars of China (No. 41806135), the Fujian Provincial Natural Science Foundation (No. 2021J01510), the Scientific Research Foundation of Third Institute of Oceanography, MNR (No. 2020017 and 2019018), and the China-ASEAN Maritime Cooperation Fund (No. HX04-210901). The authors thank Mallory Eckstut, PhD, from Liwen Bianji, Edanz Editing China (www.liwenbianji.cn/ac), for editing the English text of a draft of this manuscript. Finally, we express our sincere thanks to editors and anonymous reviewers for their constructive comments that significantly improved our manuscript.

CONFLICT OF INTEREST

The authors declare no conflicts of interest.

AUTHOR CONTRIBUTIONS

Yufei Dai: Conceptualization (lead); formal analysis (lead); methodology (lead); software (lead); writing – original draft (lead); writing – review and editing (equal). **Watchara Sakornwimon:** Formal analysis (equal); investigation (equal); methodology (lead); software (lead); writing – original draft (lead). **Rachawadee Chantra:** Investigation (lead); methodology (equal); resources (equal). **Liyuan Zhao:** Formal analysis (equal); software (equal); writing – review and editing (equal). **Fuxing Wu:** Formal analysis (equal); software (equal); writing – review and editing (equal). **Reyilamu Aierken:** Formal analysis (equal); software (equal); writing – review and editing (equal). **Kongkiat Kittiwattanawong:** Conceptualization (lead); project administration (lead); supervision (lead); validation (equal). **Xianyan Wang:** Conceptualization (lead); funding acquisition (lead); project administration (lead); supervision (lead); validation (equal).

DATA AVAILABILITY STATEMENT

Sequences containing the polymorphic microsatellite loci reported in this paper have been deposited into the GenBank database under the following Accession Numbers: [MK766845–MK766870](https://doi.org/10.1007/s10592-016-0913-7).

ORCID

Yufei Dai  <https://orcid.org/0000-0002-6250-4453>

REFERENCES

- Amaral, A. R., Smith, B. D., Mansur, R. M., Brownell, R. L., & Rosenbaum, H. C. (2017). Oceanographic drivers of population differentiation in Indo-Pacific bottlenose (*Tursiops aduncus*) and humpback (*Sousa* spp.) dolphins of the northern Bay of Bengal. *Conservation Genetics*, 18, 371–381. <https://doi.org/10.1007/s10592-016-0913-7>
- An, H. S., Kang, H. W., Han, H. S., Park, J. Y., Myeong, J. I., & An, C. M. (2014). Population genetic structure of the tongue sole (*Cynoglossus semilaevis*) in Korea based on multiplex PCR assays with 12 polymorphic microsatellite markers. *Genes & Genomics*, 36, 539–549. <https://doi.org/10.1007/s13258-014-0190-0>
- Balloux, F., & Goudet, J. (2002). Statistical properties of population differentiation estimators under stepwise mutation in a finite island model. *Molecular Ecology*, 11(4), 771–783. <https://doi.org/10.1046/j.1365-294X.2002.01474.x>
- Bonin, A., Bellemain, E., Bronken, E. P., Pompanon, F., Brochmann, C., & Taberlet, P. (2004). How to track and assess genotyping errors in population genetics studies. *Molecular Ecology*, 13(11), 3261–3273. <https://doi.org/10.1111/j.1365-294X.2004.02346.x>
- Brown, A. M., Kopps, A. M., Allen, S. J., Bejder, L., Littleford-Colquhoun, B. L., Parra, G. J., Cagnazzi, D. D., Thiele, D., Palmer, C., & Frere, C. H. (2014). Population Differentiation and hybridisation of Australian Snubfin (*Orcaella heinsohni*) and Indo-Pacific Humpback (*Sousa chinensis*) dolphins in north-western Australia. *PLoS One*, 9(7), 1–14. <https://doi.org/10.1371/journal.pone.0101427>
- Carvajal-Rodríguez, A. (2018). Myriads: P-value-based multiple testing correction. *Bioinformatics*, 34(6), 1043–1104. <https://doi.org/10.1093/bioinformatics/btx746>
- Chen, H., Zhai, K., Chen, J., Chen, Y., Wen, H., Chen, S., & Wu, Y. (2008). A preliminary investigation on genetic diversity of *Sousa chinensis* in the Pearl River Estuary and Xiamen of Chinese waters. *Journal of Genetics and Genomics*, 35(8), 491–497. [https://doi.org/10.1016/S1673-8527\(08\)60067-3](https://doi.org/10.1016/S1673-8527(08)60067-3)
- Chen, L., Caballero, S., Zhou, K., & Yang, G. (2010). Molecular phylogenetics and population structure of *Sousa chinensis* in Chinese waters inferred from mitochondrial control region sequences. *Biochemical Systematics and Ecology*, 38(5), 897–905. <https://doi.org/10.1016/j.bse.2010.09.009>
- Chen, L., & Yang, G. (2009). A set of polymorphic dinucleotide and tetranucleotide microsatellite markers for the indo-pacific humpback dolphin (*Sousa chinensis*) and cross-amplification in other cetacean species. *Conservation Genetics*, 10, 697–700. <https://doi.org/10.1007/s10592-008-9618-x>
- De Barba, M., Miquel, C., Lobréaux, S., Quenette, P. Y., Swenson, J. E., & Taberlet, P. (2017). High-throughput microsatellite genotyping in ecology: Improved accuracy, efficiency, standardization and success with low-quantity and degraded DNA. *Molecular Ecology Resources*, 17(3), 492–507. <https://doi.org/10.1111/1755-0998.12594>
- Diniz-Filho, J., Soares, T., Lima, J., Dobrovolski, R., Landeiro, V., Telles, M., Rangel, T., & Bini, L. (2013). Mantel test in population genetics. *Genetics and Molecular Biology*, 36(4), 475–485. <https://doi.org/10.1590/S1415-47572013000400002>
- Earl, D. A., & VonHoldt, B. M. (2012). STRUCTURE HARVESTER: A website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation Genetics Resources*, 4, 359–361. <https://doi.org/10.1007/s12686-011-9548-7>
- Evanno, G., Regnaut, S., & Goudet, J. (2005). Detecting the number of clusters of individuals using the software STRUCTURE: A simulation study. *Molecular Ecology*, 14(8), 2611–2620. <https://doi.org/10.1111/j.1365-294X.2005.02553.x>
- Faircloth, B. C. (2008). MSATCOMMANDER: detection of microsatellite repeat arrays and automated, locus-specific primer design. *Molecular Ecology Resources*, 8, 92–94. <https://doi.org/10.1111/j.1471-8286.2007.01884.x>
- Gill, P., Curran, J., & Elliot, K. (2005). A graphical simulation model of the entire DNA process associated with the analysis of short tandem repeat loci. *Nucleic Acids Research*, 33(2), 632–643. <https://doi.org/10.1093/nar/gki205>
- González-Castellano, I., Perina, A., González-Tizón, A. M., Torrecilla, Z., & Martínez-Lage, A. (2018). Isolation and characterization of 21 polymorphic microsatellite loci for the rockpool shrimp *Palaemon elegans* using Illumina Miseq sequencing. *Scientific Reports*, 8, 17197. <https://doi.org/10.1038/s41598-018-35408-1>

- Goudet, J. (1995). FSTAT (version 1.2): A computer program to calculate F-statistics. *Journal of Heredity*, 86(6), 485–486. <https://doi.org/10.1093/oxfordjournals.jhered.a111627>
- Gui, D., Jia, K., Xia, J., Yang, L., Chen, J., Wu, Y., & Yi, M. (2013). De novo assembly of the Indo-Pacific humpback dolphin leucocyte transcriptome to identify putative genes involved in the aquatic adaptation and immune response. *PLoS One*, 8(8), e72417. <https://doi.org/10.1371/journal.pone.0072417>
- Hauge, X. Y., & Litt, M. (1993). A study of the origin of 'shadow bands' seen when typing dinucleotide repeat polymorphisms by the PCR. *Human Molecular Genetics*, 2(4), 411–415. <https://doi.org/10.1093/hmg/2.4.411>
- Hung, S. K., & Jefferson, T. A. (2004). Ranging patterns of Indo-Pacific humpback dolphins (*Sousa chinensis*) in the Pearl River Estuary, People's Republic of China. *Aquatic Mammals*, 30(1), 159–174. <https://doi.org/10.1578/AM.30.1.2004.159>
- Jakobsson, M., & Rosenberg, N. A. (2007). CLUMPP: A cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. *Bioinformatics*, 23(14), 1801–1806. <https://doi.org/10.1093/bioinformatics/btm233>
- Jefferson, T. A. (2000). Population biology of the Indo-Pacific humpbacked dolphin in Hong Kong waters. *Wildlife Monographs*, 144, 1–65.
- Jefferson, T. A., & Curry, B. E. (2015). Humpback dolphins: A brief introduction to the genus *Sousa*. *Advances in Marine Biology*, 72, 1–16. <https://doi.org/10.1016/bs.amb.2015.04.001>
- Jefferson, T. A., & Hung, S. K. (2004). A review of the status of the Indo-Pacific humpback dolphin (*Sousa chinensis*) in Chinese waters. *Aquatic Mammals*, 30(1), 149–158. <https://doi.org/10.1578/AM.30.1.2004.149>
- Jefferson, T. A., Hung, S. K., & Würsig, B. (2009). Protecting small cetaceans from coastal development: Impact assessment and mitigation experience in Hong Kong. *Marine Policy*, 33(2), 305–311. <https://doi.org/10.1016/j.marpol.2008.07.011>
- Jefferson, T. A., & Rosenbaum, H. C. (2014). Taxonomic revision of the humpback dolphins (*Sousa* spp.), and description of a new species from Australia. *Marine Mammal Science*, 30(4), 1494–1541. <https://doi.org/10.1111/mms.12152>
- Jefferson, T. A., & Smith, B. D. (2016). Re-assessment of the conservation status of the Indo-Pacific humpback dolphin (*Sousa chinensis*) using the IUCN red list criteria. *Advances in Marine Biology*, 73, 1–26. <https://doi.org/10.1016/bs.amb.2015.04.002>
- Jefferson, T. A., Smith, B. D., Braulik, G. T., & Perrin, W. (2017). *Sousa chinensis* (errata version published in 2018). *The IUCN RedList of Threatened Species*, e.T82031425A123794774. <https://doi.org/10.2305/IUCN.UK.2017-3.RLTS.T82031425A50372332.en>
- Jia, K., Bian, C., Yi, Y., Li, Y., Jia, P., Gui, D., Zhang, X., Lin, W., Sun, X., Lv, Y., Li, J., You, X., Shi, Q., Yi, M., & Wu, Y. (2019). Whole Genome Sequencing of Chinese White Dolphin (*Sousa chinensis*) for High-Throughput Screening of Antihypertensive Peptides. *Marine Drugs*, 17(9), 504. <https://doi.org/10.3390/md17090504>
- Jutapruet, S., Huang, S. L., Li, S., Lin, M., Kittiwattanawong, K., & Pradit, S. (2015). Population Size and Habitat Characteristics of the Indo-Pacific Humpback Dolphin (*Sousa chinensis*) Off Donsak, Surat Thani. *Thailand. Aquatic Mammals*, 41(2), 129–142. <https://doi.org/10.1578/AM.41.2.2015.129>
- Karczmarski, L., Huang, S. L., & Chan, S. C. (2017). Threshold of long-term survival of a coastal delphinid in anthropogenically degraded environment: Indo-Pacific humpback dolphins in Pearl River Delta. *Scientific Reports*, 7, 42900. <https://doi.org/10.1038/srep42900>
- Keyse, J., Crandall, E. D., Toonen, R. J., Meyer, C. P., Trembl, E. A., & Riginos, C. (2014). The scope of published population genetic data for Indo-Pacific marine fauna and future research opportunities in the region. *Bulletin of Marine Science*, 90(1), 47–78. <https://doi.org/10.5343/bms.2012.1107>
- Kumar, G., & Kocour, M. (2017). Applications of next-generation sequencing in fisheries research: A review. *Fisheries Research*, 186(1), 11–22. <https://doi.org/10.1016/j.fishres.2016.07.021>
- Lin, W., Chang, L., Céline, H., Zhou, R., Chen, J., Chen, X., & Wu, Y. (2012). Differentiated or not? An assessment of current knowledge of genetic structure of *Sousa chinensis* in China. *Journal of Experimental Marine Biology and Ecology*, 416–417, 17–20. <https://doi.org/10.1016/j.jembe.2012.02.002>
- Lin, W., Karczmarski, L., Xia, J., Zhang, X., Yu, X., & Wu, Y. (2016). Increased human occupation and agricultural development accelerates the population contraction of an estuarine delphinid. *Scientific Reports*, 6, 35713. <https://doi.org/10.1038/srep35713>
- Lin, W., Zhou, R., Porter, L., Chen, J., & Wu, Y. (2010). Evolution of *Sousa chinensis*: A scenario based on mitochondrial DNA study. *Molecular Phylogenetics and Evolution*, 57(2), 907–911. <https://doi.org/10.1016/j.ympev.2010.07.012>
- Luo, R., Liu, B., Xie, Y., Li, Z., Huang, W., Yuan, J., He, G., Chen, Y., Pan, Q. I., Liu, Y., Tang, J., Wu, G., Zhang, H., Shi, Y., Liu, Y., Yu, C., Wang, B. O., Lu, Y., Han, C., ... Wang, J. (2012). SOAPdenovo2: an empirically improved memory-efficient short-read *de novo* assembler. *GigaScience*, 1, 30. <https://doi.org/10.1186/2047-217X-1-18>
- Mace, G. M., & Lande, R. (1991). Assessing extinction threats: Toward a reevaluation of IUCN threatened species categories. *Conservation Biology*, 5(2), 148–157. [https://doi.org/10.1016/0006-3207\(92\)91254-P](https://doi.org/10.1016/0006-3207(92)91254-P)
- Manlik, O., Chabanne, D., Daniel, C., Bejder, L., Allen, S. J., & Sherwin, W. B. (2019). Demography and genetics suggest reversal of dolphin source-sink dynamics, with implications for conservation. *Marine Mammal Science*, 35(3), 732–759. <https://doi.org/10.1111/mms.12555>
- Mendez, M., Jefferson, T. A., Kolokotronis, S.-O., Krützen, M., Parra, G. J., Collins, T., Minton, G., Baldwin, R., Berggren, P., Särnblad, A., Amir, O. A., Peddemors, V. M., Karczmarski, L., Guissamulo, A., Smith, B., Sutaria, D., Amato, G., & Rosenbaum, H. C. (2013). Integrating multiple lines of evidence to better understand the evolutionary divergence of humpback dolphins along their entire distribution range: a new dolphin species in Australian waters? *Molecular Ecology*, 22(23), 5936–5948. <https://doi.org/10.1111/mec.12535>
- Mendez, M., Subramaniam, A., Collins, T., Minton, G., Baldwin, R., Berggren, P., Särnblad, A., Amir, O. A., Peddemors, V., Karczmarski, L., Guissamulo, A., & Rosenbaum, H. C. (2011). Molecular ecology meets remote sensing: Environmental drivers to population structure of humpback dolphins in the Western Indian Ocean. *Heredity*, 107, 349–361. <https://doi.org/10.1038/hdy.2011.21>
- Ming, Y., Jian, J., Yu, F., Yu, X., Wang, J., & Liu, W. (2019). Molecular footprints of inshore aquatic adaptation in Indo-Pacific humpback dolphin (*Sousa chinensis*). *Genomics*, 111(5), 1034–1042. <https://doi.org/10.1016/j.ygeno.2018.07.015>
- Ming, Y., Jian, J., Yu, X., Wang, J., & Liu, W. (2019). The genome resources for conservation of Indo-Pacific humpback dolphin, *Sousa chinensis*. *Scientific Data*, 6, 68. <https://doi.org/10.1038/s41597-019-0078-6>
- Murray, V., Monchawin, C., & England, P. R. (1993). The determination of the sequences present in the shadow bands of a dinucleotide repeat PCR. *Nucleic Acids Research*, 21(10), 2395–2398. <https://doi.org/10.1093/nar/21.10.2395>
- Peakall, R., & Smouse, P. E. (2006). GenAlEx 6: Genetic analysis in Excel. Population genetic software for teaching and research. *Molecular Ecology Notes*, 6(1), 288–295. <https://doi.org/10.1111/j.1471-8286.2005.01155.x>
- Peakall, R., & Smouse, P. E. (2012). GenAlEx 6.5: Genetic analysis in Excel. Population genetic software for teaching and research – An update. *Bioinformatics*, 28(19), 2537–2539. <https://doi.org/10.1093/bioinformatics/bts460>

- Pei, J., Bao, P., Chu, M., Liang, C., Ding, X., Wang, H., Wu, X., Guo, X., & Yan, P. (2018). Evaluation of 17 microsatellite markers for parentage testing and individual identification of domestic yak (*Bos grunniens*). *PeerJ*, 6, e5946. <https://doi.org/10.7717/peerj.5946>
- Pompanon, F., Bonin, A., Bellemain, E., & Taberlet, P. (2005). Genotyping errors: Causes, consequences and solutions. *Nature Reviews Genetics*, 6, 847–859. <https://doi.org/10.1038/nrg1707>
- Pritchard, J. K., Stephens, M., & Donnelly, P. (2000). Inference of population structure using multilocus genotype data. *Genetics*, 155(2), 945–959. <https://doi.org/10.1093/genetics/155.2.945>
- Raudino, H. C., Tyne, J. A., Smith, A., Ottewell, K., McArthur, S., Kopps, A. M., Chabanne, D., Harcourt, R. G., Pirota, V., & Waples, K. (2019). Challenges of collecting flow from small cetaceans. *Ecosphere*, 10(10), e02901. <https://doi.org/10.1002/ecs2.2901>
- Rosenberg, N. A. (2004). DISTRUCT: A program for the graphical display of population structure. *Molecular Ecology Notes*, 4(1), 137–138. <https://doi.org/10.1046/j.1471-8286.2003.00566.x>
- Rousset, F. (2008). GENEPOP' 007: A complete re-implementation of the GENEPOP software for Windows and Linux. *Molecular Ecology Resources*, 8(1), 103–106. <https://doi.org/10.1111/j.1471-8286.2007.01931.x>
- Rozen, S., & Skaletsky, H. (2000). Primer3 on the WWW for general users and for biologist programmers. *Methods of Molecular Biology*, 132, 365–386. <https://doi.org/10.1385/1-59259-192-2-365>
- Selkoe, K. A., & Toonen, R. J. (2006). Microsatellites for ecologists: A practical guide to using and evaluating microsatellite markers. *Ecology Letters*, 9(5), 615–629. <https://doi.org/10.1111/j.1461-0248.2006.00889.x>
- Shen, Z., Qu, W., Wang, W., Lu, Y., Wu, Y., Li, Z., Hang, X., Wang, X., Zhao, D., & Zhang, C. (2010). MPprimer: A program for reliable multiplex PCR primer design. *BMC Bioinformatics*, 11, 143. <https://doi.org/10.1186/1471-2105-11-143>
- Sherwin, W. B., Chao, A., Jost, L., & Smouse, P. E. (2017). Information theory broadens the spectrum of molecular ecology and evolution. *Trends in Ecology and Evolution*, 32(12), 948–963. <https://doi.org/10.1016/j.tree.2017.09.012>
- Sherwin, W. B., Jabot, F., Rush, R., & Rossetto, M. (2006). Measurement of biological information with applications from genes to landscapes. *Molecular Ecology*, 15(10), 2857–2869. <https://doi.org/10.1111/j.1365-294X.2006.02992.x>
- Taberlet, P., Waits, L. P., & Luikart, G. (1999). Noninvasive genetic sampling: Look before you leap. *Trends in Ecology and Evolution*, 14, 323–327. [https://doi.org/10.1016/S0169-5347\(99\)01637-7](https://doi.org/10.1016/S0169-5347(99)01637-7)
- Vaini, J. O., Mota, K. G., Ojeda, A. P., Barreiros, J. P., Moreira, R. G., & Hilsdorf, A. W. S. (2019). Development and characterization of 20 polymorphic microsatellite markers for *Epinephelus marginatus* (Lowe, 1834) (Perciformes: Epinephelidae) using 454 pyrosequencing. *Genetics and Molecular Biology*, 42(1), 74–79. <https://doi.org/10.1590/1678-4685-GMB-2018-0067>
- Valsecchi, E., & Amos, W. (1996). Microsatellite markers for the study of cetacean populations. *Molecular Ecology*, 5(1), 151–156. <https://doi.org/10.1111/j.1365-294X.1996.tb00301.x>
- Van Oosterhout, C., Hutchinson, W. F., Wills, D. P. M., & Shipley, P. (2004). Micro-checker: Software for identifying and correcting genotyping errors in microsatellite data. *Molecular Ecology Notes*, 4(3), 535–538. <https://doi.org/10.1111/j.1471-8286.2004.00684.x>
- Wang, X., Wu, F., Chang, W., Hou, W., Chou, L., & Zhu, Q. (2016). Two separated populations of the Indo-Pacific humpback dolphin (*Sousa chinensis*) on opposite sides of the Taiwan Strait: Evidence from a larger-scale photo-identification comparison. *Marine Mammal Science*, 32(1), 390–399. <https://doi.org/10.1111/mms.12257>
- Wang, X., Wu, F., Zhu, Q., & Huang, S. L. (2017). Long-term changes in the distribution and core habitat use of a coastal delphinid in response to anthropogenic coastal alterations. *Aquatic Conservation Marine and Freshwater Ecosystems*, 27(3), 643–652. <https://doi.org/10.1002/aqc.2720>
- Weir, B. S., & Cockerham, C. C. (1984). Estimating F-statistics for the analysis of population structure. *Evolution*, 38(6), 1358–1370. <https://doi.org/10.1111/j.1558-5646.1984.tb05657.x>
- Zalapa, J. E., Cuevas, H., Zhu, H., Steffan, S., Senalik, D., Zeldin, E., McCown, B., Harbut, R., & Simon, P. (2012). Using next-generation sequencing approaches to isolate simple sequence repeat (SSR) loci in the plant sciences. *American Journal of Botany*, 99(2), 193–208. <https://doi.org/10.3732/ajb.1100394>
- Zane, L., Bargelloni, L., & Patarnello, T. (2002). Strategies for microsatellite isolation: A review. *Molecular Ecology*, 11(1), 1–16. <https://doi.org/10.1046/j.0962-1083.2001.01418.x>
- Zeng, Q., Lin, W., Dai, Y., Zhong, M., Wang, X., & Zhu, Q. (2020). Modeling demographic parameters of an edge-of-range population of Indo-Pacific humpback dolphin in Xiamen bay, china. *Regional Studies in Marine Science*, 40, 101462. <https://doi.org/10.1016/j.rsma.2020.101462>
- Zhang, P., Zhao, Y., Li, C., Lin, M., Dong, L., Zhang, R., Liu, M., Li, K., Zhang, H., Liu, X., Zhang, Y., Yuan, Y., Liu, H., Seim, I., Sun, S., Du, X., Chang, Y., Li, F., Liu, S., & Li, S. (2020). An Indo-Pacific humpback dolphin genome reveals insights into chromosome evolution and the demography of a vulnerable species. *iScience*, 23(10), 101640–<https://doi.org/10.1016/j.isci.2020.101640>
- Zhao, L., Sakornwimon, W., Lin, W., Zhang, P., Chantra, R., Dai, Y., Aierken, R., Wu, F., Li, S., Kittiwattanawong, K., & Wang, X. (2021). Early divergence and differential population histories of the Indo-Pacific humpback dolphin in the Pacific and Indian Oceans. *Integrative Zoology*, 16(4), 612–625. <https://doi.org/10.1111/1749-4877.12527>

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of the article at the publisher's website.

How to cite this article: Dai, Y., Sakornwimon, W., Chantra, R., Zhao, L., Wu, F., Aierken, R., Kittiwattanawong, K., & Wang, X. (2022). High genetic differentiation of Indo-Pacific humpback dolphins (*Sousa chinensis*) along the Asian Coast of the Pacific Ocean. *Ecology and Evolution*, 12, e8901. <https://doi.org/10.1002/ece3.8901>