



Research Article

Development of microsatellite markers and evaluation of the genetic diversity of the edible sea anemone *Paracondylactis sinensis* (Cnidaria, Anthozoa) in China

Junyuan Li[‡], Xuyi Yang[‡], Zifeng Zhan[§], Juan Feng[‡], Tinghui Xie[‡], Yang Li[§]

[‡] College of Agriculture and Bioengineering, Taizhou Vocational College of Science and Technology, Taizhou, China

[§] Laboratory of Marine Organism Taxonomy and Phylogeny, Qingdao Key Laboratory of Marine Biodiversity and Conservation, Shandong Province Key Laboratory of Experimental Marine Biology, Institute of Oceanology, Chinese Academy of Sciences, Qingdao, China

Corresponding author: Yang Li (liyang@qdio.ac.cn)

Academic editor: Danwei Huang

Received: 09 Aug 2024 | Accepted: 18 Sep 2024 | Published: 25 Sep 2024

Citation: Li J, Yang X, Zhan Z, Feng J, Xie T, Li Y (2024) Development of microsatellite markers and evaluation of the genetic diversity of the edible sea anemone *Paracondylactis sinensis* (Cnidaria, Anthozoa) in China.

Biodiversity Data Journal 12: e134363. <https://doi.org/10.3897/BDJ.12.e134363>

Abstract

Paracondylactis sinensis Carlgren, 1934 is a sea anemone with economic value in China. The wild population of *P. sinensis* has been shrinking due to overfishing and environmental pollution, which have caused price instability. In winter, the price of *P. sinensis* can reach 25 USD per kilogram. Up to now, there are no genetic markers developed for *P. sinensis*, preventing a further exploration of their population genetic diversity. In this study, the full-length transcriptome of *P. sinensis* was sequenced and microsatellite DNA markers (simple sequence repeats [SSRs]) were developed from those transcripts. A total of 52 primer pairs, which can amplify specific polymorphic bands in PCR experiments, were designed for the SSR markers. Genetic diversity and population genetics were analysed for *P. sinensis* populations collected from the coasts of Taizhou and Rizhao using six microsatellite DNA loci. While inbreeding was detected in both populations ($F_{is} > 0$), the overall number of alleles ($N_a = 11.3$) and bottleneck analysis suggested that the genetic diversity of *P. sinensis* has not been greatly impacted.

Clustering analyses using STRUCTURE, principal coordinate analysis and unweighted pair group method with arithmetic mean tree revealed that the Taizhou population diverged from the Rizhao population; however, the genetic differentiation between the populations was moderate. Human-mediated commercial activities may be the principal reasons for the gene flow between the populations. Our study provides the first evaluation of the genetic resources of wild *P. sinensis* populations in China, which can serve as a useful reference for future comparative studies on population genetics and may guide policy-makers in initiating strategies for germplasm conservation and artificial breeding.

Keywords

Anthozoa, microsatellite DNA loci, population genetics, genetic diversity, ISO-seq

Introduction

The sea anemone *Paracondylactis sinensis* Carlgren, 1934 (Cnidaria, Anthozoa) is distributed in the Indo-West Pacific Ocean and has been discovered mainly in countries such as Japan, China, India and Vietnam. *P. sinensis* lives in sandy beaches in intertidal or shallow subtidal zones. In China, *P. sinensis* has become an economic species and is widely accepted as a delectable delicacy, especially by local residents of Zhejiang Province (Li et al. 2023). However, industrial aquaculture for *P. sinensis* has not been established. The wild population of *P. sinensis* has decreased in recent years because of overfishing and other human activities. Accordingly, *P. sinensis* has become more expensive, with market prices fluctuating depending on the season (Li et al. 2023). Taizhou and Rizhao are cities in south China and north China where wild *P. sinensis* are typically caught for commercial use. However, no effective genetic markers are available for *P. sinensis*, which hinders access to the genetic diversity of this species or further investigations of its population genetics.

Up to now, only the mitochondrial genome and some mitochondrial barcodes are accessible in GenBank for *P. sinensis* (Li et al. 2023). Due to the slow evolutionary rates of the mitochondrial sequences in the anthozoans, mitochondrial DNA markers are limited in resolving genetic diversity at the population level for these organisms (Shearer et al. 2008). Microsatellite DNA markers are composed of simple sequence repeats (SSRs). As a result of their high degree of polymorphism, robust reproducibility, co-dominance and abundance in the genome, SSRs have become useful tools that are applicable in population genetics, parentage analysis, genetic linkage studies and molecular marker-assisted breeding (Park et al. 2009). With the development of high-throughput sequencing, SSRs can be easily and efficiently isolated from the genome or transcriptome databases of non-model organisms. SSRs for other species of sea anemones have been identified through computational searches of DNA databases derived from Illumina sequencing (Sherman et al. 2007, Gatins et al. 2016, Bocharova et al. 2018, Lane et al. 2020, Ryan et al. 2021). However, assemblies from short reads can

introduce errors and the gene coding regions with useful application values are usually interrupted by introns or repetitive sequences. In this study, we used PacBio long reads (polished by Illumina short reads of RNA) (Tan et al. 2023, Teng et al. 2024) to generate a full-length transcriptome for *P. sinensis*. Using the obtained transcripts, SSRs were identified by repeat motif searches and a total of 52 polymorphic SSR markers were screened by PCR. Genetic diversity and population genetics were analysed for *P. sinensis* populations collected from the coasts of Taizhou and Rizhao. The results of this study provide the first information about the genetic resources for *P. sinensis* in China and serve as a reference for policy-makers to implement strategies associated with natural population conservation and breeding of improved genetic seeds.

Materials and Methods

Full-length transcriptome sequencing

One *Paracondylactis sinensis* individual collected from the coast of Taizhou (121°39.5'E, 28°20.1'N) was selected for RNA extraction. The tentacle, body column and mesentery tissues were dissected and the TRIzol method (Invitrogen, USA) was used to isolate RNA. Contaminating DNA was removed using a TURBO DNA-free™ Kit (Ambion, USA). The quantity and integrity of the extracted RNA were assayed with a Qubit fluorometer (Thermo Scientific, USA) and an Agilent Bioanalyzer 2100 (Agilent, USA), respectively. Equal amounts of RNA from different tissues were pooled to construct a library for PacBio sequencing. The ISO-Seq library was constructed following the protocols of the Preparing ISO-Seq® libraries using SMRTbell® prep kit 3.0 (PacBio, USA). SMRT sequencing was performed on a PacBio Sequel IIe sequencer (PacBio, USA). For Illumina sequencing, RNA from each tissue sample was used to construct sequencing libraries separately. Paired-end libraries with insert sizes of 300 bp were prepared from RNA using an NEBNext Poly(A) mRNA Magnetic Isolation Module of the NEBNext® Ultra™ II RNA Library Prep Kit (NEB, USA). The libraries originating from the different tissues were sequenced on an Illumina NovaSeq 6000 sequencer (2 × 150 bp paired-end reads) (Illumina, USA). All sequencing services in this study were provided by Novogene Bioinformatics Technology Co., Ltd. (Tian Jin, China).

The long reads generated by the PacBio sequencer were processed according to the PacBio ISO-Seq pipeline (<https://github.com/PacificBiosciences/IsoSeq/blob/master/isodeq-clustering.md>). To polish the transcript consensus sequences clustered from PacBio long reads, both Illumina data (cleaned by Trimmomatic 0.36 (Bolger et al. 2014)) and PacBio data were used to improve the sequencing accuracy by NextPolish2 (Hu et al. 2024). Finally, CD-HIT v. 4.8.1 was used to remove redundancy in the transcripts with the following parameters: -c 0.90 -n 9 -T 0 -M 80,000 (Li and Godzik 2006). The coding regions of the transcripts were predicted by TransDecoder v.5.5.0 software (<http://transdecoder.sourceforge.net/>) and the resulting open reading frames (ORFs) were annotated by the NCBI NR protein database by the BLASTx tool v.2.2.31 + with an e value of less than 1×10^{-5} .

Microsatellite marker development

SSRs were detected with the MicroSatellite identification tool (MISA) software (Beier et al. 2017). The identified repeat types included mononucleotides, dinucleotides, trinucleotides, tetranucleotides, pentanucleotides and hexanucleotides. The primers used for each detected SSR were designed by the Batch Target Region Primer Design plugin of TBtools software (Chen et al. 2020). To select polymorphic and specific SSR markers, PCR experiments were performed on five *P. sinensis* individuals. The PCR mixture consisted of 15 to 30 ng of total genomic DNA, 2.0 μ l of each primer (10 μ M), 10.0 μ l of 5 \times buffer, 4.0 μ l of dNTPs (each 10 mM), 1.0 μ l of PrimeSTAR HS DNA Polymerase (Takara, Japan) and PCR-grade water to a total volume of 50 μ l. The PCR amplification programme was as follows: an initial denaturation step at 95°C for 5 minutes, followed by 25 cycles of denaturation at 95°C for 30 seconds, annealing starting at 65°C and decreasing by 1°C per cycle to 60°C for 30 seconds for each annealing step and extension at 72°C for 30 seconds. The reaction was terminated with a final extension step at 72°C for 5 minutes. The PCR products were detected using an SDS-PAGE gel stained by silver.

Genetic diversity and population structure analyses

Sixteen individuals of *Paracondylactis sinensis* were collected from the coasts of Taizhou, Zhejiang (121°39.5'E, 28°20.1'N) and Rizhao, Shandong (119°37.2'E, 35°30.3'N) respectively (Fig. 1Suppl. material 1) and all the 32 individuals were used for subsequent genetic analyses. Six primers were randomly selected from the microsatellite marker primer set developed in this study. Forward primers labelled with the fluorescent dye 6-FAM were synthesised. Combined with the corresponding unlabelled reverse primers, PCR was performed for all 32 *P. sinensis* individuals. The PCR system and programme were the same as those described above. The PCR products were then detected via capillary electrophoresis on a 3730XL DNA Analyzer (Applied Biosystems, USA). Allele sizes were visualised using GeneMapper software (Dash et al. 2020).

The genetic diversity indicators, including the number of (effective) alleles, Shannon index, observed heterozygosity, expected heterozygosity and inbreeding coefficient, were calculated by GenAEx 6.5 (PEAKALL and SMOUSE 2005). BOTTLE NECK v. 1.2.02 (Piry et al. 1999) was utilised to detect recent reductions in effective population size. The two-phase model (TPM) is more appropriate for identifying bottleneck effects for microsatellite loci data (Cornuet and Luikart 1996), so it was chosen as the mutational model. The Wilcoxon test was employed to assess statistical significance. The Bayesian clustering approach, implemented using STRUCTURE v.2.3.4 (Pritchard et al. 2000), with 2×10^5 Markov Chain Monte Carlo iterations and a burn-in period of 2×10^4 was employed to infer population structure. The number of clusters (K) was evaluated from 1 to 10 and the results were analysed using StructureSelector (Li and Liu 2017) to determine the optimal K value. Cluster analysis was also performed using NTSYS-pc 2.10e software (Rohlf 1987). A dendrogram was created using the unweighted pair group method with arithmetic mean (UPGMA), which employs similarity matrices derived from

the simple matching coefficient. Principal coordinate analysis (PCoA) was used to illustrate the multidimensional distributions of *P. sinensis* in a scatter plot. The source of genetic differentiation was analysed using an analysis of molecular variance (AMOVA) on the microsatellite dataset. The AMOVA analysis was conducted at three hierarchical levels: variances between the populations from Taizhou and Rizhao, variances amongst individuals within each population and variances within individuals. The fixation index (F_{st}) and effective number of migrants (N_m) between the populations were calculated by GenAEx 6.5 (PEAKALL and SMOUSE 2005).

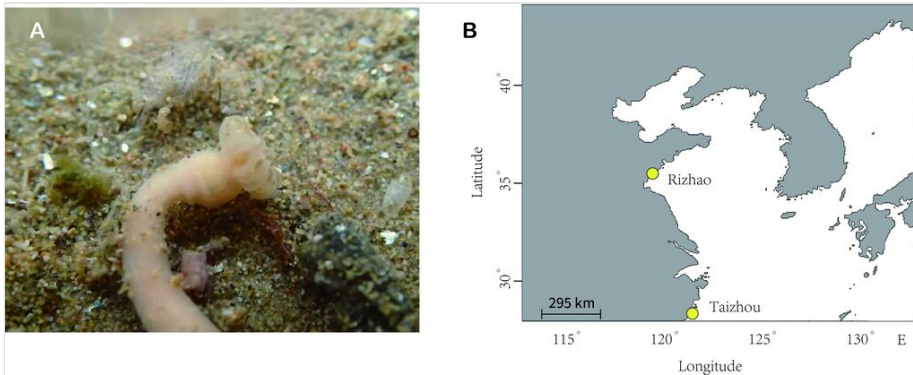


Figure 1. [doi](#)

In situ picture of a *Paracondylactis sinensis* settling on sand (A) and sampling sites of the populations analysed in this study (B).

Results

Statistics of transcriptome data

A total of 1,112,498 polymerase reads were obtained using the PacBio Sequel II platform. After preprocessing, 142.47 Gb of subreads with an N50 length as long as 2,843 bp were obtained. The Illumina sequencer generated 8.68, 6.30 and 8.28 Gb of clean data for the tentacle, body column and mesentery tissues, respectively. The sequencing data produced in this study have been submitted to the National Center for Biotechnology Information Database under BioProject number PRJNA1086232. The clustering procedure for the PacBio long reads generated 54,259 transcripts, which included 54,219 high-quality transcripts and 40 low-quality transcripts. After polishing and redundant sequences removed, 16,490 non-redundant transcripts remained. The total length of these final transcripts was 47,089,935 bp, with an N50 length of 3,059 bp. The length distribution pattern of the transcripts indicated that most transcripts ranged from 2,000 bp to 3,000 bp in length (Fig. 2). Functional annotation revealed 14,829 transcripts with homologues in the NCBI NR protein database. Protein entries of sea anemones (*Actinia tenebrosa* 9.10%, *Exaiptasia diaphana* 8.54% and *Nematostella vectensis* 8.21%) and corals (*Pocillopora damicornis* 7.34%, *Stylophora pistillata* 7.30%, *Acropora millepora* 7.09% and *Orbicella faveolata* 6.74%) were most closely related (Fig. 3).

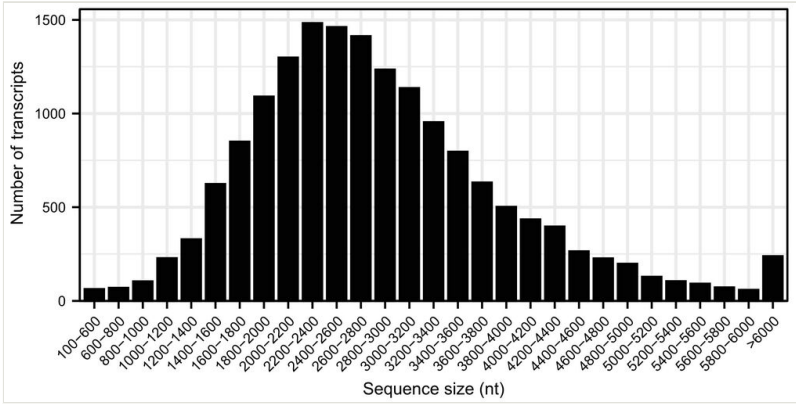


Figure 2. [doi](#)

Length distribution of the transcripts of *Paracondylactis sinensis*.

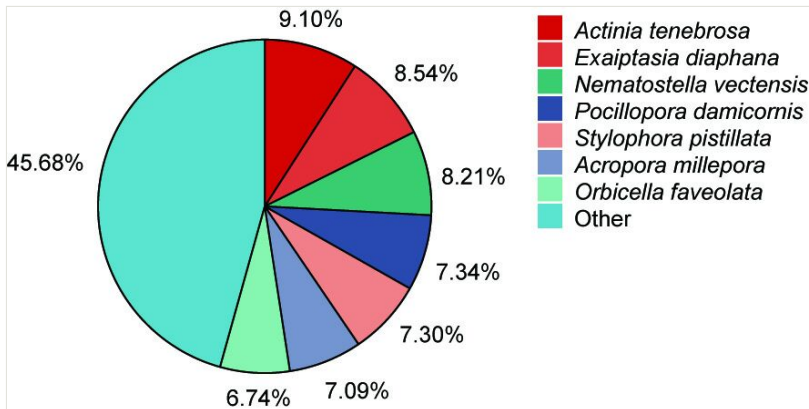


Figure 3. [doi](#)

Species distribution for the annotation with the NR database.

Identification and screening of SSR markers

MISA software revealed 7,596 SSRs distributed within 4,867 transcripts. The distribution frequency of SSRs was 46.06% and the average interval between SSRs was 6.20 kb. The major SSR motifs were mononucleotide, trinucleotide and dinucleotide, accounting for 47.76%, 30.27% and 11.37% of the total number of repeats, respectively (Table 1). The number of repeats within each SSR ranged from 5 to 102. In general, for each type of repeat, as the number of repeats increased, the number of that type of repeat decreased accordingly (Table 1). To screen primers for SSRs, a total of 150 pairs of primers were randomly selected from the SSR primer sets designed with TBtools (Chen et al. 2020). After detection by electrophoresis with an SDS-PAGE gel, the primers amplifying unspecific or non-polymorphic bands were excluded (Suppl. material 2). Finally, 52 pairs of primers that generated specific and polymorphic amplified results were identified

(Table 2). Six primers of the primer set were used for the subsequent assessment of genetic diversity (Table 2).

Type of repeat	Repeat times							Total	Ratio/%	Distribution Frequency/%
	5	6	7	8	9	10	>10			
Mononucleotide	\	\	\	\	\	\	3628	3628	47.76	22.00
Dinucleotide	\	284	162	116	68	37	197	864	11.37	5.24
Trinucleotide	1119	471	286	139	96	55	133	2299	30.27	13.94
Tetranucleotide	184	89	46	17	24	18	95	473	6.23	2.87
Pentanucleotide	19	7	1	5	1	6	28	67	0.88	0.41
Hexanucleotide	51	44	27	16	16	12	99	265	3.49	1.61
Total	1373	895	522	293	205	128	4180	7596	100	46.06
Ratio/%	18.08	11.78	6.87	3.86	2.70	1.69	55.03			

Marker name	SSR motif	Forward primer (5'-3')	Reverse primer (5'-3')	Product size (bp)
SS1	(AG)10	CCCAGGGAGTTGCCATTCT	TCCCCAAATCTCCATCTGCT	280
SS2	(TG)11	GGCAAATCCCAGCTCCC	TGTCGGCAAATGTTTTGACAG	278
SS3	(GT)14	GCTGTCTGCATTGCTTCAGT	GGGCACACGTGACAAGG	271
SS4	(TC)20	ACGCCTTCTATAGCTCGCG	GGCGACCAACAGATGCG	265
SS5*	(GA)20	AGCCGTAGTAGACCCCGT	GCTCTGACGTACAGGC	258
SS6	(TG)37	TGCCCTGAGATCAGCCC	ACATTACAGCTGGTTGGAGG	257
SS7	(GT)12	GCGCCCCACCTATACTG	GCATCTGCAGTAAGCGCT	253
SS8*	(GA)14	AACACGTCCCTATCGGAGT	AACGGTCGAAAGGGGTCC	245
SS9*	(TATG)26	CGCCACTCATGCTTGCC	ACACTTGGAAGACCTTTTGTCT	279
SS10	(TGT)9	CCAAGCCACGAAATCCTTGG	TCCGAGTCCCTGGCTGTT	279
SS11	(TTG)10	TGACGATGCGTGCAAGGT	ACAGCGAACGTTGACTTCT	278
SS12	(CAA)10	ACGCAAGCAATCCGTGG	GGCGCTGCCTGGATGTT	277
SS13	(TAGA)14	GGCCCAAATACGCTAACCG	CGCGGAACAGGGGTAGG	274

Marker name	SSR motif	Forward primer (5'-3')	Reverse primer (5'-3')	Product size (bp)
SS14	(CCA)11	GCTCAGCACTTCGTACACCT	TCATGAACCCTGCTCCATC	272
SS15*	(TCT)8	TCAACTTCTCGCCGGCAG	GGGGAGGAGGGAAGGGAG	271
SS16	(AAGG)31	AGGCTAAGGACTGTGCGAG	CCGTCGCAAAACGTCTGG	269
SS17	(TTG)8	GGCCATGTTGTGAGAGACCT	AGGGCCGGTGGGAATAG	261
SS18	(GAT)13	AGGACAAATGCCACCGAG	GGTTTCCGCCGTCTTCC	258
SS19	(TTCT)13	AGCATGCGGTCTGTGAC	GCATTCTGGTCTAGCTGGGG	256
SS20	(GAT)10	AAAAGGAAGCCTGGCAAG	TGAGGGTCCAGCCTTGACT	255
SS21	(AGG)8	TAAGCGCGTCTGTGC	ACTCGCGAGGGCTTCTTG	254
SS22	(ATG)13	ATGTCATACGGCAAATGGC	TGGGTGGCATTCTTCGGT	253
SS23	(TGG)8	TGGTGGGAGGGTGATG	TCACACATGCCCTTTGAC	252
SS24	(GACA)9	TCTTGTCCCGCTGCATG	CCCGTGGCAGACATCCAT	251
SS25	(ATG)12	AGTTGGCAGGGCAAATAAC	TGCCTTGTCAAAAATGTCCG	244
SS26	(GCTT)8	CCGGTAGGAACTTCCCTGG	TAAGTGTCAACCCGCGCG	242
SS27	(AAC)9	CACCTACACCGCAAGCC	CACGAAACGCAAGCAGC	240
SS28	(TTCC)8	AGCTCCGTCTTTCCTTTCCT	GCATTACCACGTAAAATGCG	235
SS29	(TCTT)8	AGCCGAAGAGCTCTGGGT	TGCTTGTCTTGTTCAGTGTGG	233
SS30	(AAC)9	TGCAACAGCAAACGGGAAG	TCACATCACTGGTGAGGC	226
SS31*	(ACA)9	GGCTTGCAACCCCTTCG	GCCAGTGCCTTTCCTTTC	222
SS32	(AGA)8	GTGGTGGGCAACATGGGT	GCGTCAGTCGTGCCACT	222
SS33*	(TTG)8	GCTCGACTCAGCTGCGT	TGACAGTAGACAGCTACCAC	219
SS34	(AAC)8	ACCCCTAACAACAGTGGGC	ATCTCGGGTCGCCAAACC	214
SS35	(CGT)9	CGGGGTGGCTATGCGATC	AGATCACTAAAGTGCAGAC	197
SS36	(CAA)9	ACAGCAGGCACGTTTCC	AGGGTGGAGGTCGCATCT	185
SS37	(TTTC)15	AGGCATTAGGTCATTTCGGACT	TTCGCACGGGGTCTTTC	176
SS38	(TTG)8	TAAGCGCAAGGCCAC	ACCTTGTGCACTGCTAGCT	174
SS39	(GTT)11	TGCCGAGTTTCACAGCG	GCGCGTCTTTGTTGTTGC	147
SS40	(ATG)8	TGAAAGTGCGCCGAAGT	GGTGATGGTACGTAGTCAGT	147
SS41	(AGA)8	CCGTCTGCTCTTGCCAG	TCACAACGTAACGGACAGC	109
SS42	(TCA)8	CGGATCGTCACGTCACC	AGCAACACCTTTGTTTTGTGT	108
SS43	(TTTAC)8	ATGACGGCCGAAACCACG	GGGAAAAGTTTGGTACTCGGT	272

Marker name	SSR motif	Forward primer (5'-3')	Reverse primer (5'-3')	Product size (bp)
SS44	(TTTTG) [*] 24	CGAGAACTCGTTTGTGTTGTT	AGCTGCTTCACTTTGGTCTT	262
SS45	(T) [*] 20	CCTGAGTCAGTTACGCAAAGG	AGCCATCAAAGTTACAGGCT	229
SS46	(TGA) [*] 11	CCTCCAGGGATGAAGGCG	TTGGCCCAGTACAGGTC	127
SS47	(TC) [*] 20	TGCTTCACTCACCCATGGG	AGCCTCTCCTACTCACAGCT	129
SS48	(TG) [*] 32	ACCGACGCGTTGAAAGG	CGTTCTCACATCCAAAACGGT	149
SS49	(CT) [*] 15	GAAGTTGGCCCTAGCGC	TGGACCAAGTTACTGGACAC	170
SS50	(TC) [*] 13	CCCTTCTCAGTGGTTGGC	ATCTCGCGGGAGGAGAGT	174
SS51	(TG) [*] 19	TGTGAGGATTTGGAGGTTTCG	GCATAGCAAACCAGGCAC	184
SS52	(GAT) [*] 14	GCAACCATGGATGATACCGC	TGCTGTCTTCATCGTGGCT	187

Note: the marker name labelled with a star indicates its following usage in fluorescent primer synthesis and population genetic analysis.

Genetic diversity and population structure analysis

The number of alleles (N_a), number of effective alleles (N_e) and inbreeding coefficient (F_{is}) of the *P. sinensis* population collected from Taizhou were 11.167, 7.713 and 0.085, respectively, which were relatively smaller than the corresponding values of the population collected from Rizhao (11.500, 7.833 and 0.118). The Shannon index (I), observed heterozygosity (H_o) and expected heterozygosity (H_e) were slightly greater in the Taizhou population (2.157, 0.781 and 0.856) than in the Rizhao population (2.107, 0.740 and 0.838) (Table 3). In a bottleneck analysis, populations with a significant excess or deficiency of heterozygosity are considered to have undergone a recent genetic bottleneck. According to the TPM, the bottleneck analysis indicated no genetic bottleneck effect in either population (p values for the Wilcoxon tests for both populations greater than 0.05). STRUCTURE analysis of the data, based on six microsatellite loci, indicated that the most likely number of genetic clusters was $K = 2$ ($\ln P(D) = -1024.3$; variance $\text{Var}[\ln P(D)] = 149.4$; $\Delta K = 8.49$). This result suggested the presence of two distinct genetic groups amongst the 32 *P. sinensis* individuals in the present study. Moreover, the populations from Taizhou and Rizhao displayed apparent spatial structuring, with the 16 individuals of the Taizhou population clustering into one genetic group and the 16 individuals of the Rizhao population forming another distinct group (Fig. 4A). PCoA based on the codominant genotypic distance matrix of the six microsatellite markers also revealed that the Taizhou population was separate from the Rizhao population. The first and second axes of the PCoA explained 20.3% and 13.6% of the total molecular variance, respectively and the p value for the differentiation level between the two populations was smaller than 0.01 (Fig. 4B). An additional confirmation of the genetic separation between the Taizhou and Rizhao populations was provided by a dendrogram constructed using UPGMA analysis, based on genetic similarity coefficients. The result

also indicated that the Taizhou and Rizhao populations were two distinct genetic groups (Fig. 4C), which was consistent with the results of the STRUCTURE and PCoA analyses. AMOVA revealed significant genetic variance between the two populations (9%, $p < 0.01$), amongst individuals within population (12%, $p < 0.01$) and within individuals (78%, $p < 0.01$) (Table 4). A comparison between the Taizhou and Rizhao populations indicated that the genetic differentiation characterised by the F_{st} value was 0.068 and the N_m between the two populations was 3.419.

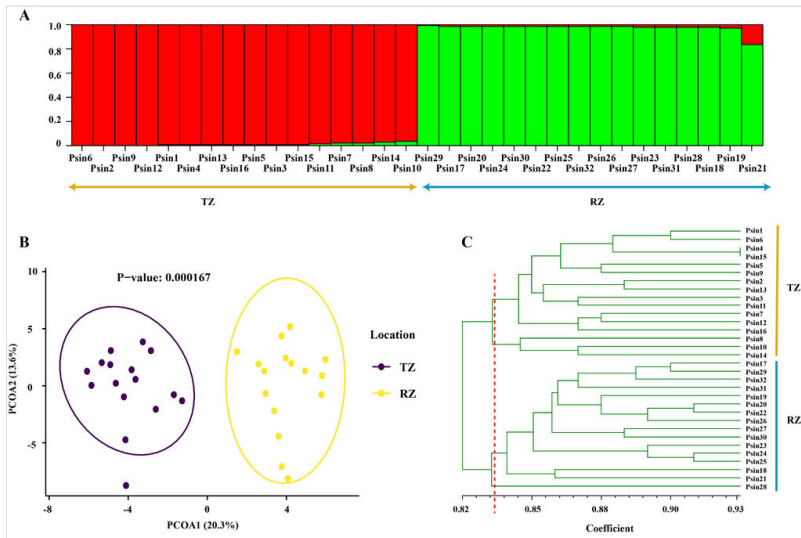


Figure 4. [doi](#)

Population analyses of 32 *Paracondylactis sinensis* individuals collected from Taizhou (16 individuals) and Rizhao (16). (A) Bayesian STRUCTURE clustering results; each colour represents the proportion of inferred ancestry from $K (=2)$ ancestral populations, with each bar representing an individual sample; (B) Two-dimensional projection of the PCoA for 32 *P. sinensis* samples along the first two principal axes; (C) Relationships of *P. sinensis* populations based on genetic similarities derived from polymorphism patterns of SSR markers, displayed in a UPGMA dendrogram. TZ: Taizhou; RZ: Rizhao; Psin: *P. sinensis*.

Table 3.

Genetic diversity parameters for the populations of *Paracondylactis sinensis* from Taizhou (TZ) and Rizhao (RZ).

Population	N	Na	Ne	I	Ho	He	F_{is}
TZ	16	11.167	7.713	2.157	0.781	0.856	0.085
RZ	16	11.500	7.833	2.107	0.740	0.838	0.118
Mean	16	11.333	7.773	2.132	0.760	0.847	0.101

Note: N: number of samples; Na: number of different alleles; Ne: number of effective alleles; I: Shannon's information index; Ho: observed heterozygosity; He: expected heterozygosity; F_{is} : inbreeding coefficient.

Table 4.

Analysis of molecular variance for the *Paracondylactis sinensis* from Taizhou and Rizhao, based on six microsatellite loci data.

Source of variation	df	Sum of squares	Estimated variance	Percentage of variance
Between populaitons	1	11.719	0.273	9%
Amongst individuals within populations	30	89.594	0.353	12%
Within individuals	32	73.000	2.281	78%
Total	63	174.313	2.907	100%

Discussion

Advantages of the SSRs developed in this study

With the development of high-throughput sequencing technology, single nucleotide polymorphism (SNP) markers based population genomics have gained widespread use in studies aimed at understanding genetic diversity within species (Díaz et al. 2021, Wang et al. 2021, Yirgu et al. 2023), considering their higher power and resolution compared to SSR markers (Zimmerman et al. 2020). However, for practical applications involving polymorphism detection, the genetic signals of SSRs can be easily and economically collected using gel-based methods (such as polyacrylamide gel electrophoresis or capillary electrophoresis), which are less resource-intensive compared to SNP detection (usually realised by approaches, such as SNP arrays or next-generation sequencing). Furthermore, the species transferability of SNP markers is typically lower than that of SSR markers, when the SSRs are derived from coding regions (Park et al. 2009). The SSR markers developed in this study were derived from transcriptome data that contained rich coding regions. The flanking sequences of these SSRs are more conserved than the flanking regions of SNPs or SSRs previously reported for other sea anemones (Sherman et al. 2007, Gatins et al. 2016, Bocharova et al. 2018, Chan et al. 2020, Lane et al. 2020, Ryan et al. 2021), as the latter two are based on DNA sequences with a great proportion of non-coding regions characterised by variable sequences. Therefore, the primers of SSR markers developed in this study can potentially be applied across closely-related species with greater success.

Genetic diversity of *Paracondylactis sinensis*

The number of alleles (N_a) is a parameter that can reflect genetic diversity and adaptability to changing environmental conditions. However, the value of N_a is subjected to the number of samples and detection methods. The value of N_a increases with the use of larger sample sizes or more sensitive detection methods. In a study by Remy Gatins (Gatins et al. 2016), the genetic diversity of four species of wild sea anemones from the Indo-Pacific was investigated using SSR polymorphisms. Even with a larger sample size

(13-40 individuals) and the same detection method used in our study (capillary electrophoresis), only the populations of *Stichodactyla gigantea* in that study presented greater Na values (12.6-13.4) than those of *P. sinensis* in China (~ 11.3). The populations of the other three species (*Entacmaea quadricolor*, *Heteractis magnifica* and *S. mertensii*) presented significantly lower values (3.3-6.69). The relatively high Na value in *P. sinensis* reflects relatively rich genetic resources, which can increase the resilience of *P. sinensis* to changing environments. The overall inbreeding coefficient (F_{is}) of *P. sinensis* is 0.101 (> 0), indicating that heterozygote deficits ($H_o < H_e$) were observed. However, the fact that the H_o was lower than the H_e may not be an indicator of low genetic diversity in the population, given that the greater number of genotypes could compensate for the lower population heterozygosity (Beardmore et al. 1997). Additionally, the absence of bottleneck effects in our results corroborates that the *P. sinensis* population has not experienced a drastic reduction in heterozygosity. Overall, the Na, H_o and H_e of *P. sinensis* did not exhibit significant signs of a decrease in genetic diversity. Nevertheless, it remains necessary to take a cautious attitude by formulating and implementing conservative strategies for wild *P. sinensis* populations in China, given the increasing severity of anthropogenic impacts, including overfishing, marine pollution and habitat destruction. These human-induced pressures, if unaddressed, may precipitate an irreversible decline in heterozygosity (a much greater F_{is}), with far-reaching consequences for the genetic resources of *P. sinensis*.

Population structure and gene flow

The STRUCTURE, PCoA cluster and UPGMA tree analyses revealed that the populations from Taizhou and Rizhao were two distinct genetic groups (Fig. 4). The difference between the genetic pools of the two populations could be associated with the different environments of Taizhou and Rizhao, cities in southern and northern China, respectively. *P. sinensis* is accustomed to warm waters and they can be collected year-round along certain coasts of southern cities. However, in northern cities, *P. sinensis* emerges along the coast only in the summer. Selective pressures (such as low temperature and low food availability) may restrict population size in northern China, which further decreases genetic diversity (reflected by lower Shannon index, H_o and H_e in the Rizhao population) and aggravates F_{is} , as shown in the results of this study (Table 3). F_{st} is a measure of genetic differentiation between populations. The F_{st} value between the populations of Taizhou and Rizhao was 0.068, indicating minimal genetic differentiation (an F_{st} less than 0.05 is often considered negligible (Curnow and Wright 1979)). This low population divergence seemed to contradict the substantial geographical distance between Taizhou and Rizhao (more than 1000 km, Fig. 1). Theoretically, the effects of genetic isolation by distance are more evident in low-mobility species, as gene flow tends to homogenise populations in some geographical areas (Avice 2008, Ferreira et al. 2014). While rapid seasonal oceanic currents could facilitate the dispersal of *P. sinensis* planktonic larvae, thus reducing population differentiation, the sedentary nature of adult *P. sinensis* compared with other mobile aquatic organisms might limit this effect. A low level of genetic divergence between geographically distant populations has been observed in other marine species (Benzie and Williams 1992,

Apolônio Silva de Oliveira et al. 2017), which may be caused by multiple factors, such as human-mediated dispersal, genetic drift and population size (Yue et al. 2010). For the *P. sinensis* populations in China, commercial activity could be the principal factor. In Taizhou and nearby cities, *P. sinensis* is sold perennially in the seafood market and seafood vendors ship *P. sinensis* to northern cities in the cold seasons when the supply of wild *P. sinensis* is low. In this way, the genetic resources of southern populations have been introduced to northern districts, thus homogenising the genetic resources of the two populations to some extent. The analysis in this study revealed that the N_m reached 3.419, indicating that the gene flow between the two populations was high (an N_m greater than 1 is considered sufficient to counteract genetic drift and maintain genetic similarity between populations (Larson et al. 1984)). This gene flow resulted in greater N_a value in the Rizhao population than in the Taizhou population (Table 3). The STRUCTURE analysis suggested that the individuals in the Rizhao population tended to contain more genetic elements from the Taizhou population than vice versa (Fig. 4A, the red proportion in the bars representing the Rizhao population was slightly larger than the green parts of the bars in Taizhou group), also congruent with the hypothesis of human-mediated dispersal. According to the AMOVA, the genetic variation between the northern and southern populations was only 9%, which was smaller than the variances amongst individuals (12%) and within individuals (78%). The small variances between populations could be associated with gene flow. The observed high variation within individuals is also apparent in other species (Karnosky 1991), which suggests that the high genetic resources in *P. sinensis* potentially serves as a genetic pool for selective breeding in the future.

Conclusions

This study represents the first successful development of microsatellite markers for *Paracondylactis sinensis*, an economically valuable sea anemone in China. By employing a combination of PacBio long-read and Illumina short-read sequencing technologies, we identified 52 polymorphic SSR markers and utilised six microsatellite loci to evaluate the genetic diversity of *P. sinensis* populations from Taizhou and Rizhao. Our analyses revealed mild genetic differentiation between the two populations, with evidence of gene flow likely facilitated by human-mediated commercial activities. The absence of significant bottleneck effects and the relatively high genetic diversity, indicated by the number of alleles and heterozygosity levels, suggest that the genetic resources of *P. sinensis* have not been critically compromised. However, the detected inbreeding coefficients highlight the need for conservation strategies to safeguard the genetic integrity of this species in the face of ongoing environmental pressures. This research provides a crucial genetic baseline that can provide information for future studies and conservation efforts, offering valuable insights into the genetic management and sustainable utilisation of *P. sinensis* in China.

Data resources

The raw data of sequencing reads for the transcriptome reported in this paper have been submitted in the National Center for Biotechnology Information (NCBI) database: BioProject PRJNA1086232.

Acknowledgements

Thanks are given to the Oceanographic Data Center, IOCAS and the Center for High Performance Computing and System Simulation, Pilot National Laboratory for Marine Science and Technology (Qingdao), for providing computing power.

Funding program

This study was supported by the Taizhou Science and Technology Plan Project (No. 23nya15), the Biological Resources Programme of Chinese Academy of Sciences (No. KFJ-BRP-017-097), General Scientific Research Project of the Department of Education of Zhejiang Province (Y202250489) and Science and Technology Program of Nanji Islands National Marine Nature Reserve Administration (No. JJZB-PYCG-2021112901).

Author contributions

Junyuan Li: Conceptualisation, Investigation, Methodology, Software, Validation, Formal analysis, Writing - Original draft, Writing - Review and Editing, Funding Acquisition; Xuyi Yang: Resources, Validation, Data Curation, Writing - Review and Editing; Juan Feng: Investigation, Visualisation; Ting hui Xie: Investigation, Resources, Validation, Data Curation; Zifeng Zhan: Supervision, Project administration, Funding Acquisition; Yang Li: Conceptualisation, Validation, Investigation, Data Curation, Supervision, Project Administration, Funding Acquisition.

Conflicts of interest

The authors have declared that no competing interests exist.

References

- Apolônio Silva de Oliveira D, Decraemer W, Moens T, dos Santos GAP, Derycke S (2017) Low genetic but high morphological variation over more than 1000 km coastline refutes omnipresence of cryptic diversity in marine nematodes. *BMC Evolutionary Biology* 17 (1). <https://doi.org/10.1186/s12862-017-0908-0>
- Avise J (2008) Phylogeography: retrospect and prospect. *Journal of Biogeography* 36 (1): 3-15. <https://doi.org/10.1111/j.1365-2699.2008.02032.x>

- Beardmore JA, Mair GC, Lewis RI (1997) Biodiversity in aquatic systems in relation to aquaculture. *Aquaculture Research* 28 (10): 829-839. <https://doi.org/10.1046/j.1365-2109.1997.00947.x>
- Beier S, Thiel T, Münch T, Scholz U, Mascher M (2017) MISA-web: a web server for microsatellite prediction. *Bioinformatics* 33 (16): 2583-2585. <https://doi.org/10.1093/bioinformatics/btx198>
- Benzie JAH, Williams ST (1992) No genetic differentiation of giant clam (*Tridacna gigas*) populations in the Great Barrier Reef, Australia. *Marine Biology* 113 (3): 373-377. <https://doi.org/10.1007/bf00349161>
- Bocharova E, Sergeev A, Volkov A (2018) Identification of microsatellite loci in sea anemones *Aulactinia stella* and *Cribrinopsis albopunctata* (family Actiniidae). *F1000Research* 7 <https://doi.org/10.12688/f1000research.13724.1>
- Bolger A, Lohse M, Usadel B (2014) Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 30 (15): 2114-2120. <https://doi.org/10.1093/bioinformatics/btu170>
- Chan WWR, Tay YC, Ang HP, Tun K, Chou LM, Huang D, Meier R (2020) Reproduction in Urbanised Coastal Waters: Shallow-Water Sea Anemones (*Entacmaea quadricolor* and *Stichodactyla haddoni*) Maintain High Genetic Diversity and Panmixia. *Diversity* 12 (12). <https://doi.org/10.3390/d12120467>
- Chen C, Chen H, Zhang Y, Thomas H, Frank M, He Y, Xia R (2020) TBtools: An integrative toolkit developed for interactive analyses of big biological data. *Molecular Plant* 13 (8): 1194-1202. <https://doi.org/10.1016/j.molp.2020.06.009>
- Cornuet JM, Luikart G (1996) Description and power analysis of two tests for detecting recent population bottlenecks from allele frequency data. *Genetics* 144 (4): 2001-2014. <https://doi.org/10.1093/genetics/144.4.2001>
- Curnow RN, Wright S (1979) Evolution and the genetics of populations, Volume 4: Variability within and among Natural Populations. *Biometrics* 35 (1). <https://doi.org/10.2307/2529965>
- Dash HR, Shrivastava P, Das S (2020) Analysis of capillary electrophoresis results by GeneMapper® ID-X v 1.5 Software. *Springer Protocols Handbooks* 213-237. https://doi.org/10.1007/978-1-0716-0274-4_24
- Díaz BG, Zucchi MI, Alves-Pereira A, de Almeida CP, Moraes ACL, Vianna SA, Azevedo-Filho J, Colombo CA (2021) Genome-wide SNP analysis to assess the genetic population structure and diversity of *Acrocomia* species. *PLOS ONE* 16 (7). <https://doi.org/10.1371/journal.pone.0241025>
- Ferreira D, Galindo B, Frantini-Silva W, Almeida F, Sofia S (2014) Genetic structure of a Neotropical sedentary fish revealed by AFLP, microsatellite and mtDNA markers: a case study. *Conservation Genetics* 16 (1): 151-166. <https://doi.org/10.1007/s10592-014-0648-2>
- Gatins R, Saenz-Agudelo P, Scott A, Berumen M (2016) Development and characterization of new polymorphic microsatellite markers in four sea anemones: *Entacmaea quadricolor*, *Heteractis magnifica*, *Stichodactyla gigantea*, and *Stichodactyla mertensii*. *Marine Biodiversity* 48 (3): 1283-1290. <https://doi.org/10.1007/s12526-016-0576-0>
- Hu J, Wang Z, Liang F, Liu S, Ye K, Wang D (2024) NextPolish2: A repeat-aware polishing tool for genomes assembled using HiFi long reads. *Genomics, Proteomics & Bioinformatics* 22 (1). <https://doi.org/10.1093/gpbjnl/qzad009>

- Karnosky DF (1991) Crop improvement (Correction) plant population genetics, breeding, and genetic resources anthony H. D. Brown Michael T. Clegg Alex L. Kahler Bruce S. Weir. *BioScience* 41 (1): 53-53. <https://doi.org/10.2307/1311547>
- Lane S, Briffa M, Wilson A, Truebano M, Foster N (2020) Evidence of fostering in an internally brooding sea anemone. *Ethology* 126 (12): 1141-1147. <https://doi.org/10.1111/eth.13088>
- Larson A, Wake DB, Yanev KP (1984) Measuring gene flow among populations having high levels of genetic fragmentation. *Genetics* 106 (2): 293-308. <https://doi.org/10.1093/genetics/106.2.293>
- Li J, Li Y, Xie T, Feng J, Yang X, Zhan Z (2023) The complete mitochondrial genome of an economic sea anemone (*Paracondylactis sinensis*) in the East China Sea. *Mitochondrial DNA B Resour* 8 (9): 977-980. <https://doi.org/10.1080/23802359.2023.2254464>
- Li W, Godzik A (2006) Cd-hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences. *Bioinformatics* 22 (13): 1658-1659. <https://doi.org/10.1093/bioinformatics/btl158>
- Li Y, Liu J (2017) StructureSelector: A web-based software to select and visualize the optimal number of clusters using multiple methods. *Molecular Ecology Resources* 18 (1): 176-177. <https://doi.org/10.1111/1755-0998.12719>
- Park Y, Lee JK, Kim N (2009) Simple Sequence Repeat Polymorphisms (SSRPs) for evaluation of molecular diversity and germplasm classification of minor crops. *Molecules* 14 (11): 4546-4569. <https://doi.org/10.3390/molecules14114546>
- PEAKALL R, SMOUSE P (2005) 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>
- Piry S, Luikart G, Cornuet J (1999) Computer note. BOTTLENECK: a computer program for detecting recent reductions in the effective size using allele frequency data. *Journal of Heredity* 90 (4): 502-503. <https://doi.org/10.1093/jhered/90.4.502>
- Pritchard JK, 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>
- Rohlf FJ (1987) NTSYS-pc: Microcomputer programs for numerical taxonomy and multivariate analysis. *American Statistician* 41 (4). <https://doi.org/10.2307/2684761>
- Ryan WH, Aida J, Krueger-Hadfield SA (2021) The contribution of clonality to population genetic structure in the sea anemone, *Diadumene lineata*. *Journal of Heredity* 112 (1): 122-139. <https://doi.org/10.1093/jhered/esaa050>
- Shearer TL, van Oppen MJH, Romano SL, Wörheide G (2008) Slow mitochondrial DNA sequence evolution in the Anthozoa (Cnidaria). *Molecular Ecology* 11 (12): 2475-2487. <https://doi.org/10.1046/j.1365-294x.2002.01652.x>
- Sherman CH, Peucker A, Ayre D (2007) Do reproductive tactics vary with habitat heterogeneity in the intertidal sea anemone *Actinia tenebrosa*? *Journal of Experimental Marine Biology and Ecology* 340 (2): 259-267. <https://doi.org/10.1016/j.jembe.2006.09.016>
- Tan K, Xu P, Lim L, Nie C, Tan K, Peng Y, Cai X, Yan X, Xu Y, Kwan KY (2023) Iso-seq and RNA-seq profiling of *Fenneropenaeus penicillatus*: Unravelling the signaling pathway and genes involved in the ovarian development. *Aquaculture Reports* 32 <https://doi.org/10.1016/j.aqrep.2023.101731>
- Teng Y, Wang Y, Zhang S, Zhang X, Li J, Wu F, Chen C, Long X, Li A (2024) Integration of full-length Iso-Seq, Illumina RNA-Seq, and flavor testing reveals potential differences in

- ripened fruits between two *Passiflora edulis* cultivars. PeerJ 12 <https://doi.org/10.7717/peerj.17983>
- Wang Q, Liu Y, Yan L, Chen L, Li B (2021) Genome-Wide SNP Discovery and Population Genetic Analysis of *Mesocentrotus nudus* in China Seas. *Frontiers in Genetics* 12 <https://doi.org/10.3389/fgene.2021.717764>
 - Yirgu M, Kebede M, Feyissa T, Lakew B, Woldeyohannes AB, Fikere M (2023) Single nucleotide polymorphism (SNP) markers for genetic diversity and population structure study in Ethiopian barley (*Hordeum vulgare* L.) germplasm. *BMC Genomic Data* 24 (1). <https://doi.org/10.1186/s12863-023-01109-6>
 - Yue GH, Li J, Bai Z, Wang CM, Feng F (2010) Genetic diversity and population structure of the invasive alien red swamp crayfish. *Biological Invasions* 12 (8): 2697-2706. <https://doi.org/10.1007/s10530-009-9675-1>
 - Zimmerman S, Aldridge C, Oyler-McCance S (2020) An empirical comparison of population genetic analyses using microsatellite and SNP data for a species of conservation concern. *BMC Genomics* 21 (1). <https://doi.org/10.1186/s12864-020-06783-9>

Supplementary materials

Suppl. material 1: Table S1. Sampling information of *Paracondylactis sinensis* used for genetic diversity and population structure analyses. [doi](#)

Authors: Li Junyuan

Data type: Table

[Download file](#) (10.76 kb)

Suppl. material 2: SDS–PAGE gel showing the electrophoresis results of three SSR primers. The results of primer 91 are specific and reveal polymorphisms amongst the five individuals. The results of primers 94 and 95 are non-polymorphic and unspecific, respectively. [doi](#)

Authors: Li Junyuan

Data type: Image

[Download file](#) (2.06 MB)