

Article

# The Landscape of Genome-Wide and Gender-Specific Microsatellites in Indo-Pacific Humpback Dolphin and Potential Applications in Cetacean Resource Investigation

Yao Ming <sup>1</sup>, Xueying Yu <sup>2</sup>, Wei Liu <sup>3</sup>, Jingzhen Wang <sup>2,\*</sup> and Wenhua Liu <sup>1,4,\*</sup>

<sup>1</sup> Guangdong Provincial Laboratory of Marine Biotechnology, Institute of Marine Science, Shantou University, 243 Daxue Road, Shantou 515063, China; myyf128@126.com

<sup>2</sup> Guangxi Key Laboratory of Marine Disaster in the Beibu Gulf, Beibu Gulf University, Qinzhou 535011, China; junzhi45@163.com

<sup>3</sup> Department of Technology and Resource Management, Guangdong Jiangmen Chinese White Dolphin Provincial Nature Reserve Management Bureau, Jiangmen 529020, China; sunmoon365@163.com

<sup>4</sup> Southern Marine Science and Engineering Guangdong Laboratory, Guangzhou 511458, China

\* Correspondence: wangjingzhen-1@163.com (J.W.); whliu@stu.edu.cn (W.L.); Fax: +86-754-86500614 (W.L.)

**Abstract:** Microsatellites are one of the important genome characterizations that can be a valuable resource for variety identification, genetic diversity, phylogenetic analysis, as well as comparative and conservation genomics research. Here, we developed comprehensive microsatellites through genome-wide mining for the threatened cetacean Indo-Pacific humpback dolphin (*Sousa chinensis*). We found 87,757 microsatellites with 2–6 bp nucleotide motifs, showing that about 32.5 microsatellites per megabase comprises microsatellites sequences. Approximately 97.8% of the markers developed in this study were consistent with the published identified markers. About 75.3% microsatellites were with dinucleotide motifs, followed by tetranucleotide motifs (17.4%), sharing the same composition pattern as other cetaceans. The microsatellites were not evenly distributed in the *S. chinensis* genome, mainly in non-coding regions, with only about 0.5% of the markers located in coding regions. The microsatellite-containing genes were mainly functionally enriched in the methylation process, probably demonstrating the potential impacts of microsatellites on biological functions. Polymorphic microsatellites were developed between different genders of *S. chinensis*, which was expected to lay the foundation for genetic diversity investigation in cetaceans. The specific markers for a male Indo-Pacific humpback dolphin will provide comprehensive and representative male candidate markers for sex identification, providing a potential biomolecular tool for further analysis of population structure and social behavior of wild populations, population trend evaluation, and species conservation management.

**Keywords:** *Sousa chinensis*; SSRs; motif; genetic biodiversity; gender identification markers; cetacean



**Citation:** Ming, Y.; Yu, X.; Liu, W.; Wang, J.; Liu, W. The Landscape of Genome-Wide and Gender-Specific Microsatellites in Indo-Pacific Humpback Dolphin and Potential Applications in Cetacean Resource Investigation. *J. Mar. Sci. Eng.* **2022**, *10*, 834. <https://doi.org/10.3390/jmse10060834>

Academic Editor: Nguyen Hong Nguyen

Received: 7 April 2022

Accepted: 9 June 2022

Published: 20 June 2022

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Microsatellites, or simple sequence repeats (SSRs), are short tandem repeats of 1–6 nucleotides motifs, which are ubiquitous in prokaryotes and eukaryotes. Compared with other genetic molecular markers, such as restriction fragment length polymorphism, random amplified polymorphic DNA, amplified fragment length polymorphism, sequence-related amplified polymorphism and target region amplification polymorphism, microsatellites are characterized by high distribution frequency, codominant inheritance, reproducibility and high polymorphism [1]. Recent studies have demonstrated that microsatellites play important roles in affecting gene activity, chromatin organization, and DNA metabolic processes [2]. In addition, microsatellites are valuable tools for population genetic analyses since they are a rich source of hypervariable codominant markers [3,4]. They are currently the most important codominant markers and have been extensively used in quantitative

trait loci (QTL) mapping, genetic diversity studies, marker-assisted selective breeding, and evolutionary studies [1,5–7]. However, the development of novel markers is still a labor-intensive and time-consuming process. Most recently, advances in next generation sequencing (NGS) have sped up microsatellite markers development from the genome level, making it fast and cost-effective compared with traditional methods.

Microsatellites mining has been widely applied and improved in eukaryotic genomes, such as gossypium, prunus, vigna species, red deer, buffalo, common carp, etc. [4,8–14]. Furthermore, microsatellites extensively identified across 719 eukaryotes revealed the SSR distribution across evolutionarily related species, including protists, plants, fungi, invertebrates and vertebrates [15]. In cetaceans, valuable microsatellites have been developed and broadly used in the study of cetacean genetic diversity [16–21]. Published PCR primer pairs were used to amplify alleles to determine the haplotype relationships and genetic diversity of bottlenose dolphins (*Tursiops truncatus*) living together in a Japanese aquarium [22]. Nineteen novel tetra-nucleotide microsatellite markers were isolated from the *Tursiops aduncus* in order to improve genotyping accuracy for applications in large-scale population-wide paternity and relatedness assessments [19]. In addition, microsatellites were also applied to understand the genetic differentiation and speciation processes in Chilean dolphins [21]. However, currently, few SSR markers have been reported for *S. chinensis* [18,23,24]. Genome-wide characterization of microsatellites is imperative for better species identification and conservation research, particularly for the endangered *S. chinensis*. The identification and characterization of microsatellites, and thus the establishment of a database, will contribute to the study of genetic diversity and population structure of *S. chinensis* and other cetaceans.

The threatened *S. chinensis* generally inhabit shallow coastal waters and are vulnerable to human activities. However, their population assessments mainly depend on wild shipping surveys, photo identification, and population models assumptions [25–28]. Here, based on the genome sequences of *S. chinensis* that we previously published [29], we focused on unraveling the characteristics, distribution, and function effects of *S. chinensis* microsatellites. The comprehensive molecular markers will provide a useful database and biological tools for genetic diversity assessment, sex identification, population structure analysis and the delineation of management units of *S. chinensis*. Moreover, common features and composition trends of microsatellites among the cetacean species will be explored. It also has important implications for genetic diversity surveys and the conservation of cetaceans.

## 2. Materials and Methods

### 2.1. Microsatellites Detection and Primer Designing

The female genome sequence of *S. chinensis* used in this study was from our previous genome research (GeneBank: GCA\_003521335.2 (female)) [29]. The male genome sequence was from the published genome sequence in the China National GeneBank DataBase (CNCBdb) (CNCBdb: CNP0000397 (male)) [30]. MISA (v1.0) [31] was used to scan the female and male genome of *S. chinensis* and detect microsatellites with the parameters: length of SSR  $\geq 22$  bp, flank length of SSR = 150 bp, motif length: 2 to 6 bp, repetition minimum: 5. Primer3 (v2.2.2) [32] was applied to design primers targeting the sequence regions containing SSRs. Microsatellites were also isolated from the genome of *Balaena mysticetus* ([www.bowhead-whale.org](http://www.bowhead-whale.org)), *Balaenoptera acutorostrata* (GeneBank: GCF\_000493695.1), *Orcinus orca* (GeneBank: GCF\_000331955.2), *Tursiops truncatus* (GeneBank: GCF\_011762595.1), *Delphinapterus leucas* (GeneBank: GCF\_002288925.1), *Lipotes vexillifer* (GeneBank: GCF\_000442215.1) and *Neophocaena asiaorientalis* (GeneBank: GCF\_003031525.1) with the same method and parameters as those used in the *S. chinensis* genome.

### 2.2. Microsatellites Statistics

The motif types of microsatellites in eight cetaceans (*S. chinensis*, *B. mysticetus*, *B. acutorostrata*, *O. orca*, *T. truncatus*, *D. leucas*, *L. vexillifer* and *N. asiaorientalis*) include

dinucleotides, trinucleotides, tetranucleotides, pentanucleotides, and hexanucleotides. The number of microsatellites of each type was classified and counted respectively for each species.

### 2.3. Microsatellites Distribution and GO (Gene Ontology) Function Enrichment

As the length of the microsatellites was shorter than the gene length in the genome, the start and end sites of microsatellites were calculated and it was determined whether these sites were located in the gene coding regions based on the genome gene annotation locations. The genes containing microsatellites were obtained and the GO function enrichment was conducted. The GO enrichment methods were as follows: Firstly, mapping all target genes to GO terms in the database (<http://www.geneontology.org/> (accessed on 20 December 2017)), then, calculating gene numbers per GO term, and finally, using a hypergeometric test to compare the genomic background to find significantly enriched GO terms in the target genes. The calculating formula was:

$$P = 1 - \sum_{i=0}^{m-1} \frac{\binom{M}{i} \binom{N-M}{n-i}}{\binom{N}{n}}$$

where  $N$  is the number of all genes with GO annotations;  $n$  is the number of target genes in  $N$ ;  $M$  is the number of all genes annotated to the certain GO terms; and  $m$  is the number of target genes in  $M$ . The calculated  $p$ -value was Bonferroni corrected, with a  $p$ -value  $< 0.05$  as the threshold value. GO terms fulfilling this condition were defined as significantly enriched GO terms in target genes.

### 2.4. Microsatellite Markers Identification

The 46 published SSR sequences of *S. chinensis* were derived from previous studies. These microsatellites were identified via PCR amplification in *S. chinensis*. We aligned the microsatellites sequences to the *S. chinensis* genome sequence with an e-value  $< 1 \times 10^{-5}$  by BLAT (v. 36) (-t = dna, -q = dna, -out = blast8) [33]. The location and motifs of the published markers were compared with our results, and the consistent markers were considered to be verified. Other markers were shorter SSRs, which were verified by searching SSRs with parameters: flank length of SSRs = 20 bp, length of SSRs  $\geq 8$  bp.

### 2.5. Male *S. chinensis* Specific Markers Development

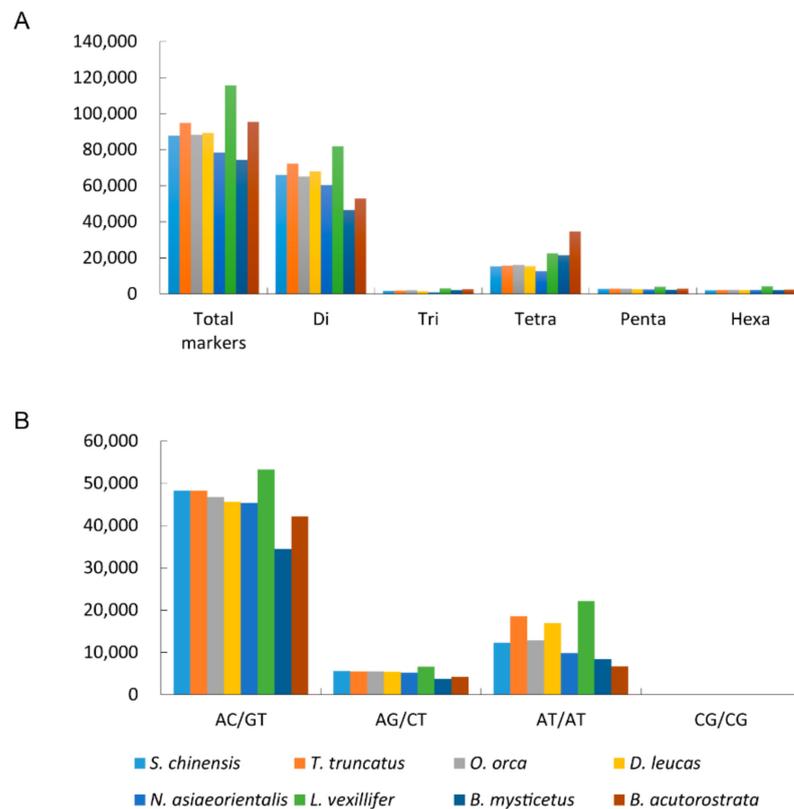
As the published male *S. chinensis* genome has no Y chromosome information, researchers have linked these unknown scaffolds to a chromosome named "chrUN" in the genome [30]. In order to obtain male-specific microsatellites from the *S. chinensis* genome, we extracted the Y chromosome unlocalized genomic scaffolds (NW\_022983135.1, NW\_022983136.1, NW\_022983137.1, NW\_022983138.1) from the published bottlenose dolphin genome (GeneBank: GCF\_011762595.1). Subsequently, we aligned the above Y chromosome scaffolds to the published "chrUN" sequences [30] to identify the Y chromosome candidate regions in the male *S. chinensis* genome.

We separated the "chrUN" chromosome sequences into contigs (32,373 contigs) (chrUN\_1 ~ chrUN\_32373) by splitting the sequences at the N positions before alignment. Based on the alignment by *Mummer* (v4.0.0) with default parameters, we filtered the results with a minimum alignment length of 10 kb. Then, we obtained candidate contigs on the Y chromosome and performed male-specific microsatellites searches using the candidate contigs sequences.

### 3. Results and Discussion

#### 3.1. Isolation and Characterization of Genomic Microsatellites in *S. chinensis*

In order to isolate microsatellite sequences, we scanned the *S. chinensis* genome [29] to identify microsatellites with motif types, such as dinucleotides, trinucleotides, tetranucleotides, pentanucleotides, and hexanucleotides. Around 87,758 microsatellites were detected, with an average frequency of 32.5 microsatellites per mega base (Tables S1 and S2). Among these 87,758 microsatellites, dinucleotides were the most abundant (66,040), accounting for 75.25%, followed by tetranucleotides (15,257; 17.39%), pentanucleotides (2774; 3.16%), hexanucleotides (1961; 2.23%), and trinucleotides (1726; 1.97%) (Figure 1A, Table S3). It was observed that the dinucleotides and tetranucleotides were the major microsatellites distributed in the *S. chinensis* genome (Figure 1A), which was also identified in other cetacean species [34]. Among the dinucleotide microsatellites, AC/GT motifs (73.03%) were the most recurrent, followed by AT/AT (18.57%), AG/CT (8.37%), and CG/CG (0.03%) motifs (Table S3). AC/GT motifs were the most abundant in the dinucleotide microsatellites (Figure 1B). On the whole, the composition of the microsatellites of the *S. chinensis* genome was consistent with that in mammals such as humans, primates, rats, camels, and pigs [35–40].



**Figure 1.** Statistics of microsatellites with different repeat types in eight cetacean genomes (Sousa chinensis, Orcinus orca, Balaenoptera acutorostrata, Balaena mysticetus, Tursiops truncatus, Delphinapterus leucas, Lipotes vexillifer, and Neophocaena asiaeorientalis). (A) Statistics of all types of microsatellite repeats: Di (Dinucleotides), Tri (Trinucleotides), Tetra (Tetranucleotides), Penta (Pentanucleotides), and Hexa (Hexanucleotides); (B) Statistics of dinucleotides motif microsatellites in cetacean species. Dinucleotides motifs contain four types: AT, CG, AC/GT and AG/CT (AT = TA, CG = GC, AC = CA = GT = TG, AG = GA = CT = TC).

To date, genome-wide microsatellites have been deciphered and mined for many aquatic species such as jellyfish, catfish and eel fish, etc. For instance, 142,616 microsatellites were found in cannonball jellyfish (*Stomolophus* sp.) by next-generation sequencing,

which laid the foundation for further population genetics studies [41]. Similarly, 24 polymorphic microsatellite loci in the Neotropical freshwater fish *Ichthyoelephas longirostris* were identified, 19 of which were used to assess their genetic diversity and structure in three Colombian rivers of the Magdalena basin [42]. The population structure and genetic diversity in yellow catfish (*Pelteobagrus fulvidraco*) were assessed using eight microsatellites in combination with capillary electrophoresis [43]. Recently, SSRs markers were detected in swamp eels, and the results showed that dinucleotide microsatellites were the most abundant type, accounting for 71% of all microsatellites loci, whereas the AC-rich motifs were the most abundant type [44]. Moreover, comprehensive microsatellites assays were performed on the genomes of 14 fish species; the abundance and frequency of the repeats were counted, and (AC) and (AT)-rich repeats were found to be dominant in different fish [45]. These results were similar to our observations of *S. chinensis* in this study. It has also been reported that (AC)<sub>n</sub> is the most common dinucleotide motif in vertebrate genomes, with a frequency of 2.3 times that of the second most common dinucleotide type (AT)<sub>n</sub> [46]. This study is the first comprehensive exploration of microsatellites within the genome of *S. chinensis* and will help facilitate genetic diversity and conservation genetics studies.

### 3.2. The Microsatellites Characteristics Comparison among Cetacean Genomes

*B. mysticetus*, *B. acutorostrata*, *O. orca*, *T. truncatus*, *D. leucas*, *L. vexillifer*, and *N. asiaeorientalis* genomes were applied for microsatellites development (Table S4). The results represented similar characteristics to *S. chinensis* in terms of microsatellite composition (Figure 1A). Odontoceti possessed more total dinucleotide microsatellites than Mysticeti (Figure 1A). Besides, the number of dinucleotide microsatellites with different motifs was also higher in Odontoceti than Mysticeti (Figure 1A). On the other hand, Mysticeti had more tetranucleotide microsatellites than Odontoceti (Figure 1A). The total number of microsatellites in cetaceans ranged from 70,000 to 120,000 (Figure 1A, Table S1). AC/GT and AT/AT-rich repeats were the main types of dinucleotide microsatellites (Figure 1B, Table S5).

### 3.3. Microsatellites Distribution in the *S. chinensis* Genome

We investigated the distribution of microsatellites in the genome's coding and non-coding regions. Most microsatellites were located in non-coding regions (87,347; 99.5%), and only about 0.5% (410) were located in coding regions, which was expected, as coding regions represent only a very small fraction of the genome. Based on the former reports, microsatellites were not evenly distributed throughout the whole genome [46–49]. Numerous microsatellites were distributed in non-coding regions and rarely in coding regions of eukaryotic species [50], e.g., about 7–10% in higher plants [51], and 9–15% in vertebrates [36,52]. In the genome of a *swamp eel*, about 1% of genes contain microsatellites in the coding regions of the genome [44]. In our study, only 0.8% of the genes contained microsatellites, a lower percentage than in vertebrates such as primates and rabbits [36,52], but one that is close to that of the *swamp eel* [44].

Next, GO function terms of the genes containing microsatellites in the coding regions were analyzed based on the genome annotation information [29]. Microsatellites were distributed in 195 genes, mainly enriched in 53 GO terms. Methylation processes such as histone acetyltransferase activity (GO: 0004402), histone acetylation (GO: 0016573), histone modification (GO: 0016570), histone methyltransferase complex (GO: 0035097), histone methyltransferase activity (H3-K4 specific) (GO: 0042800), and histone-lysine N-methyltransferase activity (GO: 0018024) were enriched (Table 1). Others were associated with metabolic and biosynthetic process, such as the regulation of the RNA metabolic process (GO: 0051252), the regulation of the nitrogen compound metabolic process (GO: 0051171), the regulation of the primary metabolic process (GO: 0080090), the regulation of the cellular metabolic process (GO: 0031323), the RNA metabolic process (GO: 0016070), etc (Table 1). The GO term annotation results suggest that the microsatellites were located in functional gene regions and may have impacts on gene function, such as the metabolic and methylation process. DNA methylation is readily detected in mam-

mals, which is critical for species' normal development and genome stability. Repetitive sequences have been regarded as 'epigenetic elements' that can regulate gene expression [53,54]. In plants and mammals, a large proportion of methylated cytosines are found in repeat elements [55,56]. The methylation of repeat elements is thought to have been compiled persistently during the evolution process and plays an important role in the epigenetic regulation of genes [57]. Most studies have focused on transposable elements, such as Alu elements (Alu), and long interspersed element-1 (LINE-1) [58,59]. The relationship between methylation and simple repeats still needs to be further studied. Thus, the analysis of genes containing microsatellites and their functions could probably be used to reveal potential mechanisms and characteristics of gene expression and regulation in *S. chinensis*.

**Table 1.** GO function terms of the genes containing microsatellites in the genomic coding regions.

GO ID	GO Term	GO Class	Corrected <i>p</i> -Value
GO:0010468	regulation of gene expression	BP	$1.76 \times 10^{-19}$
GO:0051252	regulation of RNA metabolic process	BP	$1.73 \times 10^{-17}$
GO:2000112	regulation of cellular macromolecule biosynthetic process	BP	$2.08 \times 10^{-17}$
GO:0006355	regulation of transcription, DNA-templated	BP	$2.41 \times 10^{-17}$
GO:0006351	transcription, DNA-templated	BP	$2.41 \times 10^{-17}$
GO:0051171	regulation of nitrogen compound metabolic process	BP	$1.77 \times 10^{-16}$
GO:0080090	regulation of primary metabolic process	BP	$2.23 \times 10^{-16}$
GO:0031323	regulation of cellular metabolic process	BP	$4.73 \times 10^{-16}$
GO:0016070	RNA metabolic process	BP	$5.11 \times 10^{-14}$
GO:0090304	nucleic acid metabolic process	BP	$2.73 \times 10^{-11}$
GO:0005634	nucleus	CC	$8.34 \times 10^{-9}$
GO:0006725	cellular aromatic compound metabolic process	BP	$2.00 \times 10^{-8}$
GO:1901360	organic cyclic compound metabolic process	BP	$3.74 \times 10^{-8}$
GO:0003676	nucleic acid binding	MF	$9.51 \times 10^{-8}$
GO:0010467	gene expression	BP	$1.57 \times 10^{-7}$
GO:0005488	binding	MF	$5.25 \times 10^{-7}$
GO:0003677	DNA binding	MF	$8.27 \times 10^{-7}$
GO:0044451	nucleoplasm part	CC	$8.87 \times 10^{-6}$
GO:0031981	nuclear lumen	CC	$1.07 \times 10^{-5}$
GO:0034645	cellular macromolecule biosynthetic process	BP	$1.11 \times 10^{-5}$
GO:0043227	membrane-bounded organelle	CC	$1.93 \times 10^{-5}$
GO:0043565	sequence-specific DNA binding	MF	$2.44 \times 10^{-5}$
GO:0044271	cellular nitrogen compound biosynthetic process	BP	$2.59 \times 10^{-5}$
GO:0050789	regulation of biological process	BP	$3.32 \times 10^{-5}$
GO:0004402	histone acetyltransferase activity	MF	$4.15 \times 10^{-5}$
GO:0043231	intracellular membrane-bounded organelle	CC	$6.23 \times 10^{-5}$
GO:0050794	regulation of cellular process	BP	$6.46 \times 10^{-5}$
GO:0140110	transcription regulator activity	MF	$1.21 \times 10^{-4}$
GO:0044428	nuclear part	CC	$1.74 \times 10^{-4}$
GO:0065007	biological regulation	BP	$2.80 \times 10^{-4}$
GO:0016573	histone acetylation	BP	$2.80 \times 10^{-4}$
GO:0006325	chromatin organization	BP	$3.29 \times 10^{-4}$
GO:0016570	histone modification	BP	$3.43 \times 10^{-4}$
GO:0035097	histone methyltransferase complex	CC	$5.69 \times 10^{-4}$
GO:0009058	biosynthetic process	BP	$6.57 \times 10^{-4}$
GO:0018205	peptidyl-lysine modification	BP	$7.91 \times 10^{-4}$
GO:0034641	cellular nitrogen compound metabolic process	BP	$8.80 \times 10^{-4}$
GO:1901363	heterocyclic compound binding	MF	$1.27 \times 10^{-3}$
GO:0097159	organic cyclic compound binding	MF	$1.30 \times 10^{-3}$
GO:0003712	transcription coregulator activity	MF	$1.43 \times 10^{-3}$
GO:0044260	cellular macromolecule metabolic process	BP	$1.49 \times 10^{-3}$
GO:0042800	histone methyltransferase activity (H3-K4 specific)	MF	$4.83 \times 10^{-3}$

**Table 1.** Cont.

GO ID	GO Term	GO Class	Corrected <i>p</i> -Value
GO:0016286	small conductance calcium-activated potassium channel activity	MF	$4.83 \times 10^{-3}$
GO:0043170	macromolecule metabolic process	BP	$7.22 \times 10^{-3}$
GO:0005515	protein binding	MF	$9.68 \times 10^{-3}$
GO:0016592	mediator complex	CC	0.015
GO:1990234	transferase complex	CC	0.019
GO:0003700	DNA-binding transcription factor activity	MF	0.023
GO:0018024	histone-lysine N-methyltransferase activity	MF	0.023
GO:0003713	transcription coactivator activity	MF	0.024
GO:0005516	calmodulin binding	MF	0.029
GO:0071339	MLL1 complex	CC	0.029

Note: Biological Process, BP; Cellular Component, CC; Molecular Function, MF.

### 3.4. Validation of Microsatellites in *S. chinensis*

Several accurate and widely used techniques can be applied to identify cetacean species, such as morphology, fingerprinting, and DNA barcoding [60,61]. However, it remains challenging to investigate the genetic diversity of cetaceans, due to the difficulty of obtaining accurate samples, especially for species threatened with extinction. *S. chinensis* is currently listed as Vulnerable by IUCN. It is also classified as a first-class protected animal by the Chinese National Key Protected Wild Aquatic Animals List. Live animal experiments are illegal and harmful to them. Cetacean stranding provides a unique opportunity to gain biological and ecological knowledge of cetaceans [62]. Previous studies have obtained microsatellites loci from stranded *S. chinensis* and identified them via PCR amplification [18,24]. We aligned the previous identified microsatellite markers to our genome. Of these total 46 markers, all could be mapped to the genome sequence. Then, we checked the positions and motif types between the published markers and the markers we isolated from the genome. Overall, 39 published markers were found in our markers that had the same location on the genome and the same motif types (+/−), and 7 published markers (SGATA47, SGATA35, SGATA28, SGATA25, SGATA18, SGATA13, SCA30) were not found in our markers at the corresponding position of the genome (Table 2). However, we extracted the sequences of the above seven marker-locatable genomic regions and developed the microsatellites using looser parameters: flank length of SSRs = 20 bp, and length of SSRs ≥ 8 bp in the genome. The results showed that the motif and the repeat numbers of these six microsatellites (※) were the same as those reported, except for SGATA18 (#) (Table 2). As a result, 45 of 46 markers could be confirmed at the genomic scale, with an overall verification rate of about 97.8%, indicating that the method of microsatellites isolation from the whole-genome sequence was comprehensive and accurate.

**Table 2.** Identification of microsatellites using published markers.

Published Marker	Repeat Type	SSR ID in <i>S. chinensis</i>	Repeat Type	Note
SGATA42	(GATA)11	scaffold428.1_57256	(GATA)11	+
SCA6	(CA)17	scaffold113.1_88883	(CA)19	+
SGATA47	(TATC)5	/	(TAGA)5	※
SGATA45	(AGAT)8	scaffold2.1_32807	(ATCT)9	−
SGATA35	(TCTA)4	/	(GATA)4	※
SGATA33	(ATAG)7	scaffold470.1_58268	(TCTA)8	−
SGATA30	(AGAT)7	scaffold162.1_47075	(TATC)7	−
SGATA28	(AT)4	/	(AT)4	※
SGATA25	(GATA)5	/	(AGAT)5	※
SGATA21	(TAGA)10	scaffold35.1_72245	(ATAG)10	+
SGATA18	(GATA)3GAT(GATA)3AATA(GATA)2	/	/	#
SGATA13	(TCTA)4	/	(TCTA)4	※
SCA66	(CA)11	scaffold85.1_89945	(GT)11	−

Table 2. Cont.

Published Marker	Repeat Type	SSR ID in <i>S. chinensis</i>	Repeat Type	Note
SCA64	(CA)23	scaffold77.1_66768	(CA)22	+
SCA58	(AC)12	scaffold18.1_1347	(GT)12	–
SCA56	(CA)15	scaffold284.1_42084	(CA)13	+
SCA55	(AC)14	scaffold249.1_12650	(AC)14	+
SCA54	(CA)20	scaffold199.1_4149	(AC)18	+
SCA48	(CA)16	scaffold319.1_1085	(TG)16	–
SCA39	(AC)20	scaffold196.1_39358	(AC)24	+
SCA37	(AC)23	scaffold57.1_92416	(GT)23	–
SCA33	(AC)22	scaffold34.1_17501	(AC)17	+
SCA30	(CA)9	/	(TG)9	※
SCA27	(AC)21	scaffold70.1_71178	(TG)22	–
SCA22	(CT)7TTCT(CA)36	scaffold54.1_109484	(GT)28GAGAAA(GA)7	–
SCA17	(AC)16	scaffold66.1_63598	(AC)19	+
SCA12	(AC)13	scaffold20.1_86731	(AC)13	+
SCA9	(CA)23	scaffold329.1_58395	(AC)37	–
Sch5878	(CAAC)12	scaffold108.1_36155	(TTGG)11	–
Sch6660	(AAGG)13	scaffold10.1_97587	(CCTT)13	–
Sch443	(CCAT)12	scaffold10.1_97555	(GGAT)12	–
Sch10207	(CATC)12	scaffold240.1_14114	(GATG)12	–
Sch843	(AAAT)11	scaffold25.1_94818	(TTTA)11	–
Sch7424	(ATGG)13	scaffold89.1_41509	(CCAT)13	–
Sch7357	(ATGG)11	scaffold20.1_86885	(ATGG)11	+
Sch193	(AGAGA)12	scaffold26.1_104752	(TCTCT)11	–
Sch8186	(CCAT)11	scaffold15.1_6749	(CCAT)11	+
Sch123	(CCTAAC)7	scaffold66.1_64097	(CCTAAC)7	+
Sch4657	(TTCC)11	scaffold304.1_101153	(AGGA)11	–
Sch5373	(GATG)11	scaffold91.1_67861	(TCCA)11	–
Sch9144	(ATCT)14	scaffold11.1_40258	(TAGA)15	–
Sch2513	(CATC)13	scaffold165.1_13454	(CATC)13	+
Sch5685	(AGGA)12	scaffold9.1_107855	(AGGA)13	+
Sch974	(GTTTT)13	scaffold32.1_92270	(AAAAC)10	–
Sch5094	(TCTA)11	scaffold28.1_42911	(GATA)11	–
Sch8947	(CTAT)12	scaffold11.1_40359	(CTAT)12	+

Note: The table comprises published microsatellite IDs, repeat types [18,24], the corresponding microsatellite IDs, and repeat types extracted from the *S. chinensis* genome. +/– denotes forward and reverse mapping in the genome and present in the extracting markers; ※ denotes the published markers could be mapped to the genome but not included in our results, after which these sequences were checked and found to be shorter SSRs (these markers can be isolated from the genome if looser parameters are set when scanning the genome sequence with *MISA* (v1.0) software); # denotes the markers could be mapped to the genome, but not included in our isolated microsatellites, even when changing the search parameters.

### 3.5. Polymorphic Microsatellites Detection Based on Male and Female Markers in *S. chinensis*

We reported the potential polymorphic microsatellites between male and female individuals. These markers may play an important role in the gender evolution or may be associated with sex chromosomes in mammals [63]. These microsatellites may also provide candidate markers for assessing population structure (male to female ratio) and the social behavior of *S. chinensis*. A total of 86,432 microsatellites were developed based on the male *S. chinensis* genome [30] (Table S6). Among these markers, 65,571 (75.87%) were dinucleotides microsatellites, followed by 14,677 (16.98%) tetranucleotides (Table S7). There were 15,678 candidate polymorphic microsatellites from the female and male markers, implying the presence of different PCR lengths and repeat numbers in both genomes (Table S8). Among them, 14,377 (91.71%) polymorphic microsatellites were dinucleotides, followed by 1063 (6.77%) tetranucleotides, 108 (0.69%) trinucleotides, 68 (0.43%) pentanucleotides, and 62 (0.40%) hexanucleotides (Table S9). The same trend of microsatellite composition was found in female and male *S. chinensis*. Dinucleotides and tetranucleotides were the predominant motif types among all the microsatellites isolated from the cetacean genome

sequences (Figure 1A). These markers are of great value for conducting population surveys of *S. chinensis* in the field.

### 3.6. Specific Markers of Male *S. chinensis*

Until now, no cetacean genome has been found to have intact individual Y chromosome (a single and complete Y chromosome) assembled sequences. Only the bottlenose dolphin has the candidate Y chromosome-associated scaffolds. The bottlenose dolphin is a close relative of the *S. chinensis* [64]. Therefore, to obtain Y chromosome-specific microsatellites in the genome of male *S. chinensis*, we aligned unanchored scaffold sequences other than autosomal and X chromosome sequences in male *S. chinensis* with candidate Y chromosome-associated scaffold sequences of bottlenose dolphins. We obtained candidate contigs on the Y chromosome of *S. chinensis* genome based on the alignment results. This is the first time that the Y chromosome region sequence of this protected species has been identified. Finally, 1349 microsatellites were isolated from the candidate contigs, which can be applied as candidate markers for male dolphins (Table S10). These markers will have promising applications in the sex identification of *S. chinensis* after testing on additional samples. Since cetaceans have highly conserved genomes, these male-specific markers will also be important in the sex ratio investigation in *S. chinensis* and other cetaceans.

## 4. Conclusions

We explored and identified useful microsatellite markers in *S. chinensis* from a genome-wide scale and designed primers for these markers. A comparison with the published marker sequences revealed that the markers developed in this study had high reliability. Microsatellites in the *S. chinensis* genome are mainly distributed in the non-coding regions, and only approximately 0.5% are distributed in coding regions. Their associated genes presented potential methylation impacts in dolphins. Dinucleotide repeat motif type microsatellites are the most abundant, and this motif type is dominated by the AC-rich type. Cetacean species share the same microsatellite composition characteristics, which possibly reveals the conserved nature of the cetacean genomes. The study also yielded polymorphic microsatellite data for male and female *S. chinensis*, which provided a valuable database for the investigation and demography of dolphin populations. Moreover, the male-specific marker mining provided useful candidate markers for the sex identification of *S. chinensis*, as well as other cetacean species. Given that many marine mammals are species at risk of extinction, there is an urgent need to develop diverse conservation tools. The extensive and reliable genome-wide microsatellite information we identified in this study is expected to expand to other cetacean species and will play an important role in their population genetics studies and conservation.

**Supplementary Materials:** The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/jmse10060834/s1>, Table S1: The composition and classification of microsatellites in cetaceans; Table S2: The microsatellites in female Indo-Pacific humpback dolphin; Table S3: The composition and classification of dinucleotide microsatellites in female Indo-Pacific humpback dolphin; Table S4: The microsatellites in cetacean genomes; Table S5: The composition and classification of dinucleotide microsatellites in cetaceans; Table S6: The microsatellites in male Indo-Pacific humpback dolphin; Table S7: The composition and classification of microsatellites in male Indo-Pacific humpback dolphin; Table S8: The polymorphic microsatellites between female and male Indo-Pacific humpback dolphin; Table S9: The composition and classification of polymorphic microsatellites between male and female Indo-Pacific humpback dolphin; Table S10: Specific markers in male Indo-Pacific humpback dolphin.

**Author Contributions:** Y.M. and W.L. (Wenhua Liu) designed the entire research; X.Y. and W.L. (Wei Liu) prepared the samples; Y.M. analyzed the data; Y.M. wrote the manuscript with significant contributions from X.Y., W.L. (Wei Liu), J.W. and W.L. (Wenhua Liu). All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was supported by the National Natural Science Foundation of China (NSFC, 41776174), Key Special Project for Introduced Talents Team of Southern Marine Science and Engineering Guangdong Laboratory (Guangzhou) (grant number GML2019ZD0606), 2019 Guangxi One Thousand Young and Middle-Aged College and University Backbone Teachers Cultivation Program, the Ministry of Agriculture and Rural Affairs of the People's Republic of China (Chinese White Dolphin Conservation Action, 171821301354051004 and 171721301354051016), and CNOOC.

**Institutional Review Board Statement:** As we only used the published genome sequences data for the analysis, the study did not require ethical approval. We choose to exclude this statement.

**Informed Consent Statement:** We choose to exclude this statement because the study did not involve humans.

**Data Availability Statement:** We used the published data for the analysis in this study. The data resources were explained and described in the manuscript. The results and maker sequences created by this study have all been attached in the supplementary materials. We choose to exclude this statement.

**Acknowledgments:** We thank Edmond Sanganyado for polishing the manuscript writing.

**Conflicts of Interest:** The authors declare that they have no conflict of interest.

## References

- Powell, W.; Machray, G.C.; Provan, J. Polymorphism revealed by simple sequence repeats. *Trends Plant Sci.* **1996**, *1*, 215–222. [[CrossRef](#)]
- Li, Y.C.; Korol, A.B.; Fahima, T.; Beiles, A.; Nevo, E. Microsatellites: Genomic distribution, putative functions and mutational mechanisms: A review. *Mol. Ecol.* **2002**, *11*, 2453–2465. [[CrossRef](#)] [[PubMed](#)]
- Abdul-Muneer, P.M. Application of microsatellite markers in conservation genetics and fisheries management: Recent advances in population structure analysis and conservation strategies. *Genet. Res. Int.* **2014**, *2014*, 691759. [[CrossRef](#)] [[PubMed](#)]
- Vieira, M.L.C.; Santini, L.; Diniz, A.L.; Munhoz, C.d.F. Microsatellite markers: What they mean and why they are so useful. *Genet. Mol. Biol.* **2016**, *39*, 312–328. [[CrossRef](#)] [[PubMed](#)]
- Li, S.-F.; Wang, B.-X.; Guo, Y.-J.; Deng, C.-L.; Gao, W.-J. Genome-wide characterization of microsatellites and genetic diversity assessment of spinach in the Chinese germplasm collection. *Breed. Sci.* **2018**, *68*, 455–464. [[CrossRef](#)] [[PubMed](#)]
- Pandey, G.; Misra, G.; Kumari, K.; Gupta, S.; Parida, S.K.; Chattopadhyay, D.; Prasad, M. Genome-wide development and use of microsatellite markers for large-scale genotyping applications in foxtail millet [*Setaria italica* (L.)]. *DNA Res. Int. J. Rapid Publ. Rep. Genes Genomes* **2013**, *20*, 197–207. [[CrossRef](#)]
- Khajuria, Y.P.; Saxena, M.S.; Gaur, R.; Chattopadhyay, D.; Jain, M.; Parida, S.K.; Bhatia, S. Development and Integration of Genome-Wide Polymorphic Microsatellite Markers onto a Reference Linkage Map for Constructing a High-Density Genetic Map of Chickpea. *PLoS ONE* **2015**, *10*, e0125583. [[CrossRef](#)]
- Abdelkrim, J.; Robertson, B.C.; Stanton, J.-A.L.; Gemmell, N.J. Fast, cost-effective development of species-specific microsatellite markers by genomic sequencing. *BioTechniques* **2009**, *46*, 185–192. [[CrossRef](#)]
- Sharma, P.C.; Grover, A.; Kahl, G. Mining microsatellites in eukaryotic genomes. *Trends Biotechnol.* **2007**, *25*, 490–498. [[CrossRef](#)]
- Taheri, S.; Lee Abdullah, T.; Yusop, M.R.; Hanafi, M.M.; Sahebi, M.; Azizi, P.; Shamshiri, R.R. Mining and Development of Novel SSR Markers Using Next Generation Sequencing (NGS) Data in Plants. *Molecules* **2018**, *23*, 399. [[CrossRef](#)]
- Wang, Q.; Fang, L.; Chen, J.; Hu, Y.; Si, Z.; Wang, S.; Chang, L.; Guo, W.; Zhang, T. Genome-Wide Mining, Characterization and Development of Microsatellite Markers in Gossypium Species. *Sci. Rep.* **2015**, *5*, 10638. [[CrossRef](#)] [[PubMed](#)]
- Shukla, N.; Kuntal, H.; Shanker, A.; Sharma, S.N. Mining and analysis of simple sequence repeats in the chloroplast genomes of genus *Vigna*. *Biotechnol. Res. Innov.* **2018**, *2*, 9–18. [[CrossRef](#)]
- Dettori, M.T.; Micali, S.; Giovinazzi, J.; Scalabrin, S.; Verde, I.; Cipriani, G. Mining microsatellites in the peach genome: Development of new long-core SSR markers for genetic analyses in five *Prunus* species. *SpringerPlus* **2015**, *4*, 337. [[CrossRef](#)] [[PubMed](#)]
- Wang, D.; Liao, X.; Cheng, L.; Yu, X.; Tong, J. Development of novel EST-SSR markers in common carp by data mining from public EST sequences. *Aquaculture* **2007**, *271*, 558–574. [[CrossRef](#)]
- Srivastava, S.; Avvaru, A.K.; Sowpati, D.T.; Mishra, R.K. Patterns of microsatellite distribution across eukaryotic genomes. *BMC Genom.* **2019**, *20*, 153. [[CrossRef](#)]
- Valsecchi, E.; Amos, W. Microsatellite markers for the study of cetacean populations. *Mol. Ecol.* **1996**, *5*, 151–156. [[CrossRef](#)]
- Shinohara, M.L.; Domingourou, X.; Takenaka, O. Microsatellites in the bottlenose dolphin *Tursiops truncatus*. *Mol. Ecol.* **1997**, *6*, 695–696. [[CrossRef](#)]
- Chen, L.; Yang, G. A set of polymorphic dinucleotide and tetranucleotide microsatellite markers for the Indo-Pacific humpback dolphin (*Sousa chinensis*) and cross-amplification in other cetacean species. *Conserv. Genet.* **2008**, *10*, 697–700. [[CrossRef](#)]

19. Nater, A.; Kopps, A.M.; Krutzen, M. New polymorphic tetranucleotide microsatellites improve scoring accuracy in the bottlenose dolphin *Tursiops aduncus*. *Mol. Ecol. Resour.* **2009**, *9*, 531–534. [[CrossRef](#)]
20. Chen, I.; Nishida, S.; Yang, W.C.; Isobe, T.; Tajima, Y.; Hoelzel, A.R. Genetic diversity of bottlenose dolphin (*Tursiops* sp.) populations in the western North Pacific and the conservation implications. *Mar. Biol.* **2017**, *164*, 202. [[CrossRef](#)]
21. Perez-Alvarez, M.J.; Olavarria, C.; Moraga, R.; Baker, C.S.; Hamner, R.M.; Poulin, E. Microsatellite markers reveal strong genetic structure in the endemic Chilean dolphin. *PLoS ONE* **2015**, *10*, e0123956. [[CrossRef](#)]
22. Sumiyama, D.; Kitamura, S.; Terasawa, F.; Hori, Y.; Murata, K.; Kulski, J.K.; Inoko, H. Paternity determination of captive bottlenose dolphins (*Tursiops truncatus*) using microsatellite DNA analysis. *J. Vet. Med. Sci.* **2008**, *70*, 711–713. [[CrossRef](#)] [[PubMed](#)]
23. Gui, D.; Jia, K.; Xia, J.; Yang, L.; Chen, J.; Wu, Y.; Yi, M. 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* **2013**, *8*, e72417. [[CrossRef](#)] [[PubMed](#)]
24. Dai, Y.; Chantra, R.; Kittiwattanawong, K.; Zhao, L.; Sakornwimon, W.; Aierken, R.; Wu, F.; Wang, X. Genetic structure of the endangered Irrawaddy dolphin (*Orcaella brevirostris*) in the Gulf of Thailand. *Genet. Mol. Biol.* **2021**, *44*, e20200365. [[CrossRef](#)]
25. Chen, B.Y.; Zheng, D.M.; Wang, L.; Xu, X.R.; Yang, G. The northernmost distribution of Indo-Pacific humpback dolphin (*Sousa chinensis*) in the world: Evidence from preliminary survey in Ningde, China. *Pak. J. Zool.* **2012**, *44*, 1209–1214.
26. Hung, S.K.; Würsig, B.; Wang, J.; Sims, M.P.; Piwetz, M.S. *Monitoring of Marine Mammals in Hong Kong Waters (2011–12)*; Agriculture, Fisheries and Conservation Department of Hong Kong SAR Government: Hong Kong, China, 2012.
27. Xu, X.; Song, J.; Zhang, Z.; Li, P.; Yang, G.; Zhou, K. The world's second largest population of humpback dolphins in the waters of Zhanjiang deserves the highest conservation priority. *Sci. Rep.* **2015**, *5*, 8147. [[CrossRef](#)] [[PubMed](#)]
28. Wang, J.; Yang, Y.; Yang, F.; Li, Y.; Li, L.; Lin, D.; He, T.; Liang, B.; Zhang, T.; Lin, Y. A framework for the assessment of the spatial and temporal patterns of threatened coastal delphinids. *Sci. Rep.* **2016**, *6*, 19883. [[CrossRef](#)] [[PubMed](#)]
29. Ming, Y.; Jian, J.; Yu, X.; Wang, J.; Liu, W. The genome resources for conservation of Indo-Pacific humpback dolphin, *Sousa chinensis*. *Sci. Data* **2019**, *6*, 68. [[CrossRef](#)]
30. Zhang, P.; Zhao, Y.; Li, C.; Lin, M.; Dong, L.; Zhang, R.; Liu, M.; Li, K.; Zhang, H.; Liu, X.; et al. An Indo-Pacific Humpback Dolphin Genome Reveals Insights into Chromosome Evolution and the Demography of a Vulnerable Species. *iScience* **2020**, *23*, 101640. [[CrossRef](#)]
31. Thiel, T.; Michalek, W.; Varshney, R.K.; Graner, A. Exploiting EST databases for the development and characterization of gene-derived SSR-markers in barley (*Hordeum vulgare* L.). *Theor. Appl. Genet.* **2003**, *106*, 411–422. [[CrossRef](#)]
32. Untergasser, A.; Cutcutache, I.; Koressaar, T.; Ye, J.; Faircloth, B.C.; Remm, M.; Rozen, S.G. Primer3—New capabilities and interfaces. *Nucleic Acids Res.* **2012**, *40*, e115. [[CrossRef](#)] [[PubMed](#)]
33. Kent, W.J. BLAT—The BLAST-like alignment tool. *Genome Res.* **2002**, *12*, 656–664. [[PubMed](#)]
34. Bourret, V.; Macé, M.; Bonhomme, M.; Crouau-Roy, B. Microsatellites in cetaceans: An overview. *Open Mar. Biol. J.* **2008**, *2*, 38–42. [[CrossRef](#)]
35. Beckmann, J.S.; Weber, J.L. Survey of human and rat microsatellites. *Genomics* **1992**, *12*, 627–631. [[CrossRef](#)]
36. Jurka, J.; Pethiyagoda, C. Simple repetitive DNA sequences from primates: Compilation and analysis. *J. Mol. Evol.* **1995**, *40*, 120–126. [[CrossRef](#)] [[PubMed](#)]
37. Moran, C. Microsatellite repeats in pig (*Sus domestica*) and chicken (*Gallus domesticus*) genomes. *J. Hered.* **1993**, *84*, 274–280. [[CrossRef](#)]
38. Nadir, E.; Margalit, H.; Gallily, T.; Ben-Sasson, S.A. Microsatellite spreading in the human genome: Evolutionary mechanisms and structural implications. *Proc. Natl. Acad. Sci. USA* **1996**, *93*, 6470–6475. [[CrossRef](#)]
39. Manee, M.M.; Algarni, A.T.; Alharbi, S.N.; Al-Shomrani, B.M.; Ibrahim, M.A.; Binghadir, S.A.; Al-Fageeh, M.B. Genome-wide characterization and analysis of microsatellite sequences in camelid species. *Mammal Res.* **2020**, *65*, 359–373. [[CrossRef](#)]
40. Song, X.; Yang, T.; Zhang, X.; Yuan, Y.; Yan, X.; Wei, Y.; Zhang, J.; Zhou, C. Comparison of the Microsatellite Distribution Patterns in the Genomes of Euarchontoglires at the Taxonomic Level. *Front. Genet.* **2021**, *12*, 227. [[CrossRef](#)]
41. Getino-Mamet, L.N.; Valdivia-Carrillo, T.; Gomez Daglio, L.; Garcia-De Leon, F.J. Isolation and characterization of 14 tetranucleotide microsatellite loci for the cannonball jellyfish (*Stomolophus* sp.) by next generation sequencing. *Mol. Biol. Rep.* **2017**, *44*, 257–260. [[CrossRef](#)]
42. Landinez-Garcia, R.M.; Marquez, E.J. Development and characterization of 24 polymorphic microsatellite loci for the freshwater fish *Ichthyoelephas longirostris* (Characiformes: Prochilodontidae). *PeerJ* **2016**, *4*, e2419. [[CrossRef](#)]
43. Guo, W.; Guo, C.; Wang, Y.; Hu, W.; Mei, J. Population structure and genetic diversity in yellow catfish (*Pelteobagrus fulvidraco*) assessed with microsatellites. *J. Genet.* **2019**, *98*, 1–4. [[CrossRef](#)]
44. Li, Z.; Chen, F.; Huang, C.; Zheng, W.; Yu, C.; Cheng, H.; Zhou, R. Genome-wide mapping and characterization of microsatellites in the swamp eel genome. *Sci. Rep.* **2017**, *7*, 3157. [[CrossRef](#)] [[PubMed](#)]
45. Lei, Y.; Zhou, Y.; Price, M.; Song, Z. Genome-wide characterization of microsatellite DNA in fishes: Survey and analysis of their abundance and frequency in genome-specific regions. *BMC Genom.* **2021**, *22*, 421. [[CrossRef](#)]
46. Toth, G.; Gaspari, Z.; Jurka, J. Microsatellites in different eukaryotic genomes: Survey and analysis. *Genome Res.* **2000**, *10*, 967–981. [[CrossRef](#)] [[PubMed](#)]
47. Ellegren, H. Microsatellites: Simple sequences with complex evolution. *Nat. Rev. Genet.* **2004**, *5*, 435–445. [[CrossRef](#)]

48. Rajendrakumar, P.; Biswal, A.K.; Balachandran, S.M.; Srinivasarao, K.; Sundaram, R.M. Simple sequence repeats in organellar genomes of rice: Frequency and distribution in genic and intergenic regions. *Bioinformatics* **2007**, *23*, 1–4. [[CrossRef](#)]
49. Li, Y.-C.; Korol, A.B.; Fahima, T.; Nevo, E. Microsatellites within genes: Structure, function, and evolution. *Mol. Biol. Evol.* **2004**, *21*, 991–1007. [[CrossRef](#)]
50. Metzgar, D.; Bytof, J.; Wills, C. Selection against frameshift mutations limits microsatellite expansion in coding DNA. *Genome Res.* **2000**, *10*, 72–80.
51. Wang, Z.; Weber, J.L.; Zhong, G.; Tanksley, S. Survey of plant short tandem DNA repeats. *Theor. Appl. Genet.* **1994**, *88*, 1–6. [[CrossRef](#)]
52. Van Lith, H.; Van Zutphen, L. Characterization of rabbit DNA micros extracted from the EMBL nucleotide sequence database. *Anim. Genet.* **1996**, *27*, 387–395. [[CrossRef](#)] [[PubMed](#)]
53. Pardue, M.L.; Gall, J.G. Chromosomal localization of mouse satellite DNA. *Science* **1970**, *168*, 1356–1358. [[CrossRef](#)] [[PubMed](#)]
54. McClintock, B. Chromosome organization and genic expression. In *Cold Spring Harbor Symposia on Quantitative Biology*; Cold Spring Harbor Laboratory Press: Long Island, NY, USA, 1951; pp. 13–47.
55. Suzuki, M.M.; Bird, A. DNA methylation landscapes: Provocative insights from epigenomics. *Nat. Rev. Genet.* **2008**, *9*, 465–476. [[CrossRef](#)] [[PubMed](#)]
56. Law, J.A.; Jacobsen, S.E. Establishing, maintaining and modifying DNA methylation patterns in plants and animals. *Nat. Rev. Genet.* **2010**, *11*, 204–220. [[CrossRef](#)]
57. Teixeira, F.K.; Colot, V. Repeat elements and the Arabidopsis DNA methylation landscape. *Heredity* **2010**, *105*, 14–23. [[CrossRef](#)] [[PubMed](#)]
58. Zheng, Y.; Joyce, B.T.; Liu, L.; Zhang, Z.; Kibbe, W.A.; Zhang, W.; Hou, L. Prediction of genome-wide DNA methylation in repetitive elements. *Nucleic Acids Res.* **2017**, *45*, 8697–8711. [[CrossRef](#)]
59. Bollati, V.; Galimberti, D.; Pergoli, L.; Dalla Valle, E.; Barretta, F.; Cortini, F.; Scarpini, E.; Bertazzi, P.A.; Baccarelli, A. DNA methylation in repetitive elements and Alzheimer disease. *Brain Behav. Immun.* **2011**, *25*, 1078–1083. [[CrossRef](#)]
60. Kriangwanich, W.; Buddhachat, K.; Poommouang, A.; Chomdej, S.; Thitaram, C.; Kaewmong, P.; Kittiwattanawong, K.; Nganvongpanit, K. Feasibility of melting fingerprint obtained from ISSR-HRM curves for marine mammal species identification. *PeerJ* **2021**, *9*, e11689. [[CrossRef](#)]
61. Silva, V.S.; Skueresky, N.; Lopes, F.; Koch, T.K.; Ott, P.H.; Siciliano, S.; Barreto, A.S.; Secchi, E.R.; de Meirelles, A.C.O.; Carvalho, V.L. Integrating morphology and DNA barcoding to assess cetacean diversity in Brazil. *Mammal Res.* **2021**, *66*, 349–369. [[CrossRef](#)]
62. Zhu, J.; Yu, X.; Zhang, Q.; Li, Y.; Tan, S.; Li, D.; Yang, Z.; Wang, J. Cetaceans and microplastics: First report of microplastic ingestion by a coastal delphinid, *Sousa chinensis*. *Sci. Total Environ.* **2019**, *659*, 649–654. [[CrossRef](#)]
63. Ziemniczak, K.; Traldi, J.B.; Nogaroto, V.; De Almeida, M.C.; Artoni, R.F.; Moreirafilho, O.; Vicari, M.R. In situ Localization of (GATA)<sub>n</sub> and (TTAGGG)<sub>n</sub> Repeated DNAs and W Sex Chromosome Differentiation in Parodontidae (Actinopterygii: Characiformes). *Cytogenet. Genome Res.* **2015**, *144*, 325–332. [[CrossRef](#)] [[PubMed](#)]
64. Ming, Y.; Jian, J.; Yu, F.; Yu, X.; Wang, J.; Liu, W. Molecular footprints of inshore aquatic adaptation in Indo-Pacific humpback dolphin (*Sousa chinensis*). *Genomics* **2018**, *111*, 1034–1042. [[CrossRef](#)] [[PubMed](#)]