

Article

Whole-Genome Survey and Microsatellite Marker Detection of Antarctic Crocodile Icefish, *Chionobathyscus dewitti*

Jinmu Kim ¹, Seung-Jae Lee ¹, Euna Jo ^{1,2}, Eunkyung Choi ¹, Minjoo Cho ¹, Soyun Choi ¹, Jeong-Hoon Kim ² and Hyun Park ^{1,*}

¹ Division of Biotechnology, College of Life Sciences and Biotechnology, Korea University, Seoul 02841, Korea

² Korea Polar Research Institute (KOPRI), Incheon 21990, Korea

* Correspondence: hpark@korea.ac.kr; Tel.: +82-(23)-290-3051

Simple Summary: Crocodile icefish inhabit the deep sea around the Southern Ocean and belong to the family Channichthyidae. The species lacks hemoglobin and has evolved an antifreeze protein, unlike other teleosts. In this study, the whole-genome survey and microsatellite motifs were analyzed, which provide relevant information on genetic diversity, population genetics, and the genomic study of crocodile icefish.

Abstract: The crocodile icefish, *Chionobathyscus dewitti*, belonging to the family Channichthyidae, is an endemic species of the Southern Ocean. The study of its biological features and genetics is challenging as the fish inhabits the deep sea around Antarctic waters. The icefish, the sole cryopelagic species, shows unique physiological and genetic features, unlike other teleosts. It lacks hemoglobin and has evolved antifreeze proteins. Here, we report the genome sequencing data of crocodile icefish produced using the Illumina Novaseq 6000 platform. The estimated genome size was 0.88 Gb with a K-value of 19, and the unique sequence, heterozygosity, error, and duplication rates were 57.4%, 0.421%, 0.317%, and 0.738%, respectively. A genome assembly of 880.69 Mb, with an N50 scaffold length of 2401 bp, was conducted. We identified 2,252,265 microsatellite motifs from the genome assembly data, and dinucleotide repeats (1,920,127; 85.25%) had the highest rate. We selected 84 primer pairs from the genome survey assembly and randomly selected 30 primer pairs for validation. As a result, 15 primer pairs were validated as microsatellite markers.

Keywords: *Chionobathyscus dewitti*; crocodile icefish; Illumina Novaseq 6000; Jellyfish; GenomeScope; K-value; microsatellite



Citation: Kim, J.; Lee, S.-J.; Jo, E.; Choi, E.; Cho, M.; Choi, S.; Kim, J.-H.; Park, H. Whole-Genome Survey and Microsatellite Marker Detection of Antarctic Crocodile Icefish, *Chionobathyscus dewitti*. *Animals* **2022**, *12*, 2598. <https://doi.org/10.3390/ani12192598>

Academic Editor: Piero Cossu

Received: 29 August 2022

Accepted: 26 September 2022

Published: 28 September 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

Over the past 40 million years, the Antarctic shelf has been affected by a series of structural and oceanographic events that have changed the composition of fish fauna and initiated fauna substitution [1]. At the core of these various events, the Drake Passage between southern South America and the Antarctic Peninsula dates back 23.5 million to 32.5 million years, dividing notothenioid stock into Magellan (non-Antarctic) and Antarctic components [2]. Over time, water temperature drops, ice forms, and salinity changes in the Drake Passage [3], making it harder to capture specimens around the polar regions. Because the Antarctic Peninsula has seen the most severe local consequences of climate change in the polar regions [3], it is critical to determine whether Antarctic fish are experiencing any loss of genetic variety.

The variety of adaptations that evolved in Antarctic fish enabling their survival in freezing temperatures represents the extremes of cold adaptation in vertebrates [4]. The Antarctic sea, perennially at or near freezing, makes the oxygen concentration higher than that in other oceans [4]. Species belonging to the Channichthyidae (Teleosteri: Perciformes) family inhabit the Antarctic sea and are white-blooded icefish known for their unique

physiological characteristics because of the absence of hemoglobin, an oxygen-binding protein in the blood [5]. Dissolved oxygen can be freely transported in solution, but most animals carry oxygen into the respiratory tissue by one or more respiratory proteins. Like other vertebrates, in fish, hemoglobin carries oxygen to cells [1]. Hemoglobin is highly temperature sensitive, and its structural and functional properties may in part reflect its environmental history, which may benefit many biological questions [1]. High-Antarctic notothenioids have evolved to reduce the hemoglobin concentration and multiplicity, and even erythrocyte numbers, compared with temperate and tropical species [6,7].

The crocodile icefish, *Chionobathyscus dewitti*, with various unique biological features, belongs to the family Channichthyidae and inhabits the Southern Ocean. It is known as one of the significant prey of the Antarctic toothfish, *Dissostichus mawsoni* [8]. *C. dewitti* occurs in the deep sea (up to 2000 m) around the Southern Ocean, but its exact habitat is unknown [9]. As a result, this species may be an excellent model for explaining the distribution and evolution of Antarctic fish. There have been no relevant genetic markers established for this species, nor has there been any research to understand the population genetic structure to investigate the evolution of this species.

These investigations necessitate the use of microsatellites, which have proven to be useful and reliable genetic markers in genetic studies [10,11]. In the past, the most popular method for isolating microsatellites was to identify repeat-containing sequences from libraries of genomic DNA enriched for microsatellite motives (Type II loci, [11,12]). Currently, these markers are isolated from sequencing libraries acquired by next-generation sequencing (NGS) techniques, which have proven to be more cost effective than the previous labor-intensive and costly approach [13]. Various pipelines are still being developed to select accurate microsatellite marker using the NGS technique, and we classified by in silico microsatellite detection and the primer design of crocodile icefish using QDD pipeline version 3.1.2 [14], and this information will help improve the conservation measures of crocodile icefish.

C. dewitti genome analysis is required to identify the biological and genetic features that enable it to live in extremely cold environments. Using NGS, we estimated the genomic characteristics of *C. dewitti* and investigated the occurrence of microsatellite. These results will aid further whole-genome sequencing of *C. dewitti* and the development of new molecular markers for population genetics.

2. Materials and Methods

2.1. Sample Collection and Genome Sequencing

C. dewitti specimens (length: ~30 cm) were captured from the Southern Ocean (65°05'S, 170°30' E on CCAMLR Subarea 88.1) in Antarctica. The samples were then transferred to a freezer and transported to the laboratory. Muscle tissue from a single female specimen was anatomized from the sample and isolated using the conventional phenol-chloroform method to extract pure genomic DNA. The quality and quantity of the DNA were confirmed using a Bioanalyzer (Agilent Technologies, Santa Clara, CA, USA) and a Qubit 2.0 Fluorometer (Invitrogen, Life Technologies, CA, USA). Morphological classification and mitochondrial COI markers were used to identify the species [15].

DNA (1 µg) was sheared into 350 bp fragments using a Covaris S2 system (Covaris, Woburn, MA, San Diego, CA, USA) in accordance with the manufacturer's protocol. The *C. dewitti* DNA library was prepared using the Illumina TruSeq DNA PCR-Free Library Prep protocol (Illumina San Diego, CA, USA). The quality check was performed using a Bioanalyzer (Agilent Technologies) and was used to cluster on the Illumina cBOT station and sequence paired ends for 101 cycles using the Illumina Novaseq 6000 platform (Illumina).

2.2. Data Analysis, Genome Assembly, and Microsatellite Detection

The quality values of Q20 (percentage of bases whose base call accuracy exceeds 99%) and Q30 (percentage of bases whose base call accuracy exceeds 99.9%) and the GC

content were estimated by seqtk version 1.3 (Available on-line: <http://github.com/lh3/seqtk>, accessed on 22 February 2022) with the primary Illumina paired-end data. Illumina paired-end data were used to estimate the genome size of *C. dewitti* using Jellyfish version 2.1.4 [16]. K-values of 17, 19, and 25 were used, and the 19-mer was completely visualized and the unique sequence, heterozygosity, duplication, and error rates were determined using GenomeScope [17]. The de novo draft genome assembly was conducted using Maryland Super-Read Celera Assembler (MaSuRCA) version 3.3.4 [18], and the contig-level assembly statistics were calculated using the `assemblathon_stats.pl` script (available online: <https://github.com/ucdavis-bioinformatics/assemblathon2-analysis> (accessed on 5 March 2022)). We checked the assessment of the completeness of the MaSuRCA assembly using Benchmarking Universal Single-Copy Orthologs (BUSCO) version 5.4.2 [19] with the Eukaryota odb 10 database.

The crocodile icefish draft genome assembly was used to identify microsatellite repeat units using the QDD pipeline version 3.1.2 [14]. Using default parameters, three QDD pipeline steps were performed. The three QDD steps that were added were contig 1 (step 1), `make_cons 0` (step 2), and contig 1 (step 3). The output of the final step was used to choose the best primer pairs for microsatellite validation. The selected candidate primer pairs were filtered using the following parameters: forward and reverse flanking regions between the simple sequence repeat, primer sequences ≥ 20 , maximum primer alignment score of 5 or ≥ 7 motif repeats, extracting singletons, and high-quality primer design [20].

For the microsatellite validation, the pure genomic DNA was isolated using one specimen with phenol-chloroform method. Randomly selected, 30 primer pairs were used for validation via PCR carried out in a Thermal Cycler Dice Touch (Takara Bio, Shiga, Japan). The PCR tube (50 μ L) contained 1 μ L genomic DNA (50 ng/ μ L), 0.5 μ L (20 pmol/L) of each forward and reverse primer, 25 μ L of 2X EmeraldAmp PCR Master Mix (Takara Bio, Kusatsu, Japan), and 23 μ L of double-distilled water. The PCR conditions were as follows: 2 min at 94 $^{\circ}$ C; 35 cycles at 94 $^{\circ}$ C for 30 s, 60 $^{\circ}$ C for 30 s, and 70 $^{\circ}$ C for 1 min; and a final extension at 72 $^{\circ}$ C for 10 min. Using the PCR product, three replicates were performed and the final validations were visualized by 4% agarose gel electrophoresis for 30 min at full volt and were displayed using a 20 bp DNA ladder (Takara Bio, Kusatsu, Japan).

3. Results and Discussion

3.1. Genome Sequencing Data Analysis and Genome Survey

A total of 60.43 Gb of sequencing data was generated using the paired-end library method on the Illumina Novaseq 6000 platform. Q20 and Q30 should be at least 90% and 85%, respectively, for highly accurate sequencing data on the Illumina NGS platform [21]. The Q20 and Q30 values of the primary Illumina paired-end data were 96.1% and 91.0%, respectively, using the seqtk script. The GC content of *C. dewitti* was 49.9% (Table 1), which did not affect the genome assembly.

Table 1. Genome survey sequencing statistics of *Chionobathyscus dewitti*.

Raw Data (bp)	Total Reads	Q20 (%)	Q30 (%)	GC Content (%)
60,426,455,804	400,175,204	96.1	91.0	49.9

Genome sequencing data (Table 1) were used to predict the genome size via K-mer analysis. K-values of 17, 19, and 25 were used, and the 19-mer was best estimated. Based on the 19-mer distribution, the estimated genome size was 881.7 Mb in total, and the unique sequence, heterozygosity, error, and duplication rates were 57.4%, 0.421%, 0.317%, and 0.738%, respectively (Figure 1). Other estimated K-values used and the genome size, heterozygosity, error, and duplication rates are displayed in Supplementary Table S1.

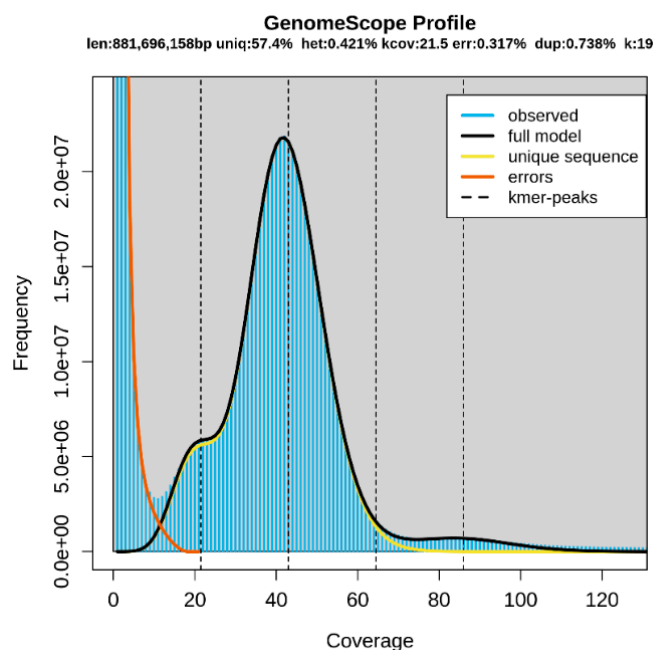


Figure 1. K-mer analysis of *Chionobathyscus dewitti* with GenomeScope (K = 19).

The graph shows the overall K-mer and genome sizes of *C. dewitti*. Len, estimated total genome length; Uniq, unique portion of the genome (not repetitive); Het, heterozygosity rate; Kcov, K-mer coverage for the heterozygous bases; Err, error rate; Dup, duplication rate. The blue bar in the graph represents the observed K-mer, and the yellow and orange lines in the graph represent the errors and unique sequences, respectively. The statistics of the overall analysis are displayed in the graph.

3.2. Genome Assembly and Microsatellite Detection

Genome sequencing data of *C. dewitti* were assembled into contigs using MaSuRCA. Using the `assemblathon_stats.pl` script, we obtained 695,247 scaffolds with a total size of 897,784,561 bp. The longest scaffold was 51,375 bp and the N50 scaffold length was 2401 bp. The GC content of the genome assembly was 41.72% (Table 2). Using BUSCO and the Eukayota database, we validated assembly contiguity and evaluated the completeness of the assembled genome. Among the 255 total BUSCO groups searched, 112 (43.9%) BUSCO core genes were identified as complete BUSCO profiles. Of these, 110 (43.1%) were single-copy and 2 (0.8%) were duplicated BUSCOs. The fragmented and missing BUSCOs were 115 (45.1%) and 28 (11%), respectively (Table S2). The results of the assembled genome provide preliminary data for whole-genome studies to accomplish accurate assembly and chromosomal-level scaffolding using long-read sequencing and chromosome conformation capture technologies.

Table 2. Statistics of *Chionobathyscus dewitti* assembly with MaSuRCA.

	MaSuRCA
Number of scaffolds	695,247
Total size of scaffolds	897,784,561
Longest scaffold	51,375
Number of scaffolds > 1 K nt	222,128 (31.9%)
Number of scaffolds > 10 K nt	6622 (1.0%)
N50 scaffold length	2401
L50 scaffold count	87,813
GC contents (%)	41.72

The assembled genome was used to identify microsatellites using QDD version 3.1.2. A total of 2,252,265 microsatellite motifs were identified. Among these, dinucleotide repeats (1,920,127; 85.25%) had the highest rate, followed by trinucleotide repeats (275,584; 12.24%), tetranucleotide repeats (46,099; 2.05%), pentanucleotide repeats (6676; 0.30%), and hexanucleotide repeats (3779; 0.16%; Figure 2). The most frequent motif in dinucleotides was AC/GT (72.23%), followed by AG/CT (16.41%), AT/AT (11.30%), and CG/CG (0.06%). The most frequent motifs in the trinucleotide repeats were AAT/ATT (25.32%), followed by AGG/CCT (22.85%), AAC/GTT (17.76%), AGC/GCT (12.72%), AAG/CTT (8.06%), ATC/GAT (8.17%), AAC/GGT (2.09%), ACT/AGT (2.03%), CCG/CGG (0.53%), and ACG/CGT (0.47%). Among tetranucleotide, pentanucleotide, and hexanucleotide repeats, the most frequent motifs were AGGG/CCCT (13.50%), AGAGG/CCTCT (28.30%), and AACCTT/AGGGTT (39.01%). These results confirm that the high mutation rate is closely related to long mutations, which is consistent with the observation that the repeat frequency decreases with repeat length [22].

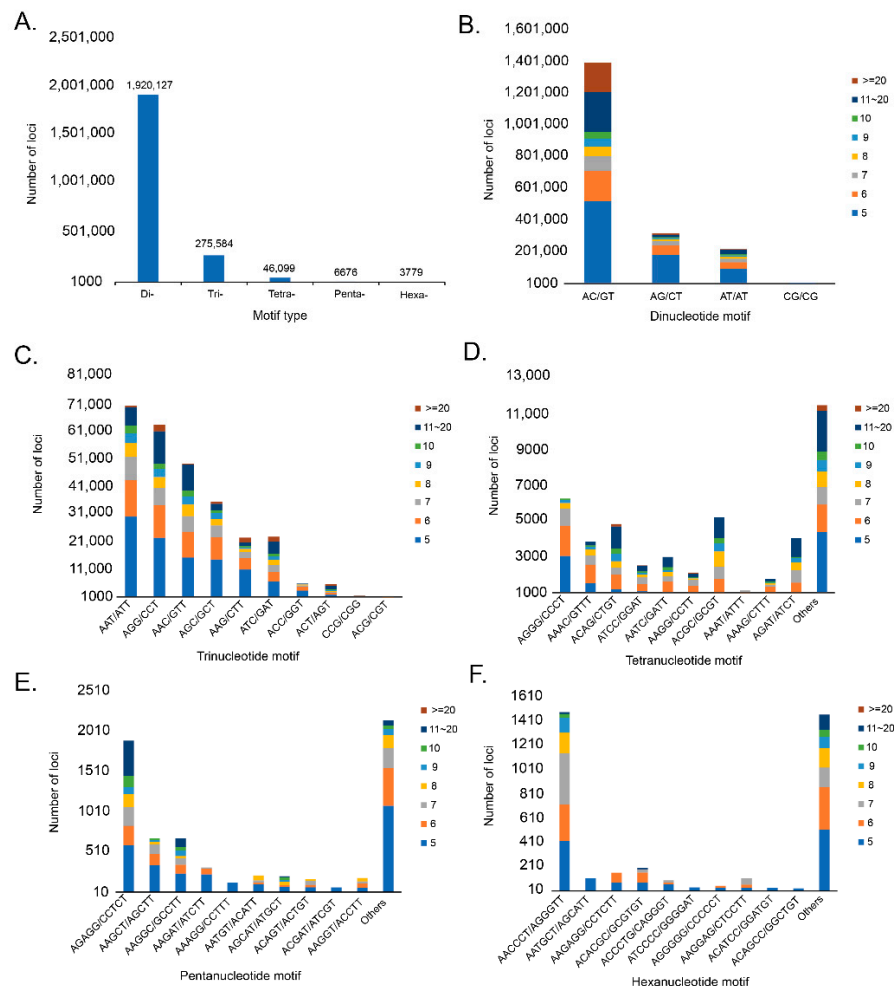


Figure 2. Type and frequency of microsatellite motifs in *Chionobathyscus dewitti* genome. (A) represents the frequency of different microsatellite motif types and (B) represents the frequency of different dinucleotide microsatellite motifs. (C–F) represent the frequency of different trinucleotide, tetranucleotide, pentanucleotide, and hexanucleotide microsatellite motifs.

Using the default parameters of the QDD pipeline and the conditions set forth in [20], we identified 84 primer pairs in the *C. dewitti* draft genome. Thirty primer pairs were randomly selected for verification and used for PCR. As a result, 15 primer pairs were confirmed to yield only a single band, and the primer sizes are shown (Figure 3 and Table 3).

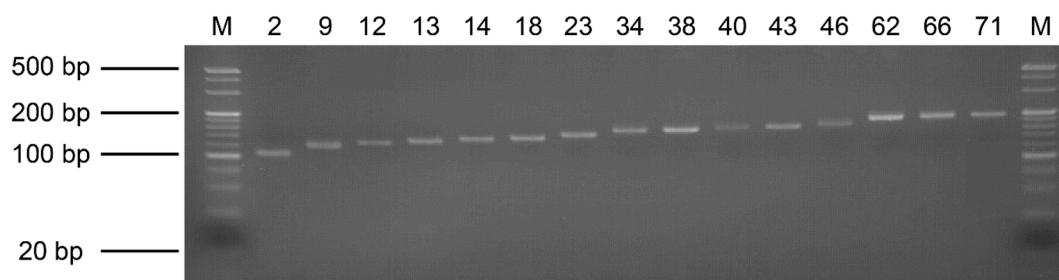


Figure 3. Validation of PCR products representing 15 microsatellite primer pairs. A 4% agarose gel was used for electrophoresis, and M represents the 20 bp DNA marker. The number of the primer pairs are in order.

Table 3. 15 microsatellite primer sets of *Chionobathyscus dewitti* used in validation.

Primer No.	Motif	PCR Product Size (bp)	Forward Primer (5' → 3')	T _m (°C)	Reverse Primer (5' → 3')	T _m (°C)
1	AT 9	105	AACTGCACAGAACAGGGTCA	59.823	AAGCCTCGGTATGCAGAGTG	59.454
2	AC 14	120	GTGAAGA AACTTGGTGCTGCC	59.897	CAGGTACGCTAAGCCAGGAG	59.688
3	AC 7	128	AAGTGGGCTGGTGA AACTTGT	59.187	TCCCGTGATGTCTTACCGAC	59.741
4	AC 8	133	TAAACCCAGGAGCGTAAGGC	59.824	GCCTCTTGCTCCGACACATA	59.749
5	AC 7	138	CCCTGAATGGCTTCCAACAC	59.964	ACGAGCCTTAATGCACACCA	59.109
6	AC 11	141	AGCATGTGTTCTTGCCAGGA	59.307	CACAGGTGGAGGCTACCTTT	59.889
7	AC 10	149	TGTGTTAAGAGCCGTGCTGT	59.821	CCTCGCTGCCTTTGAGGTAT	59.894
8	AC 17	159	CACCAGGCAGCTCGTTACTA	59.061	GTGTTTACGAACAGCGGGTT	59.469
9	AC 12	161	ACAGCCGGTTGACTGAAACT	59.179	ATCTCCAGTTCTCTGCGTCC	59.82
10	AT 9	163	CATGAGCAATGTTCCGTCGG	59.046	CCTGTTGGAGACACAAAGCC	59.625
11	AC 11	166	CCACAGATGTTGACTTGGCG	59.746	GGCCCCGTAACACCCTGTATT	59.481
12	AC 22	171	AGGCATTTAACCTCGGCACA	59.451	GGCTCCTATTTCACCCAGCT	59.962
13	AC 12	190	GAAAGCAGGCACTCAGATGC	59.753	TCATCCAGCACACTCTCCG	59.549
14	AC 14	192	GTGCATCTTTCTACCGCTGC	59.537	CTGCTCGACCCTGATGACAT	59.625
15	AC 9	194	CGGGTAAACGCTATGGAGGT	59.751	CGGACTCCATACTGTTGGCA	59.537

The figure shows the overall type and frequency of microsatellite motifs. The x-axis represents the type of motif, the y-axis represents the number of loci, and the legend displayed on the right represents the number of repeats by different colors. (A) shows different type of motifs in *C. dewitti* genome and the numbers are displayed above the bar chart. (B) shows dinucleotide motif and the repeat motif 'AC/GT' was the highest. (C) represents trinucleotide motif and the repeat motif 'AAT/ATT' have the highest. (D) represents tetranucleotide motif and among the 10 repeat motifs displayed, 'AGGG/CCCT' was the highest, and others were identified as a total of 11,468. (E) shows pentanucleotide motif and among the 10 repeat motifs displayed, 'AGAGG/CCTCT' was the highest, and others were identified as a total of 2141. (F) represents hexanucleotide motif and among the 10 repeat motifs displayed, 'AACCCCT/AGGGTT' was the highest, and others were identified as total of 1453.

All 15 species of icefish (family Channichthyidae) that live in the Southern Ocean lack hemoglobin [23], and because of the extreme aerobic mode of metabolism, they are thought to have scarce myoglobin, an intracellular respiratory pigment [23]. Both environmental and physiological properties help explain why hemoglobin and myoglobin losses are not fatal at the level of individual organisms. Advance in NGS technology and bioinformatics tools, genome surveys, and K-mer analysis developed rapidly, making these tools efficient in predicting the genome size and characteristics of non-model species [24]. Using flow cytometry, previous studies have estimated that the genome sizes of species living in the Antarctic are approximately 0.7–1.4 Gb [25]. However, with the current advancement of NGS technology, the genome sizes of Antarctic fishes are projected to be 0.6–1 Gb [26–28].

The genome size of *C. dewitti* was 0.88 Gb in this study, which was within the range of the previously reported genome sizes of Antarctic fishes. As a result, *C. dewitti*'s genome size is an acceptable outcome.

The GC content is one of the factors affecting sequence bias and between 30% and 50% does not have a significant effect on genome sequence quality [29,30]. The GC content of the assembled genome was 41.72% (Table 2), which was within the range of 30–50%. The assembled genome provides a preliminary understanding of the genomic characteristics of *C. dewitti*. However, future studies on obtaining high-quality genome assemblies, long-read sequencing technology, and chromosomal-level scaffolding technology are needed.

Microsatellites construct a meaningful distribution of eukaryotic genomes, are highly polymorphic, and have remarkable coding gene sequences in both respects [22]. In the past, it took a lot of time and effort to identify repetitive sequences in genomic DNA to select the appropriate microsatellites [11,12], but now, using the NGS technology, microsatellites can be selected more easily, conveniently, and efficiently [13]. Using assembled draft genome data of *C. dewitti*, 2 252 265 microsatellite motifs were identified, of which dinucleotide repeats (1,920,127; 85.25%) occupied the largest proportion of all repeat motifs. The frequency of di-nucleotide repeats in *C. dewitti* species was comparable to that of other fish species evaluated, such as *Pseudosciaena crocea* and *Megalobrama amblycephala* [31,32]. Furthermore, because long mutations are associated with high mutation rates, this result was consistent with data demonstrating that repeat frequency declines with increasing repeat length [22]. This finding was also consistent with previous research on microsatellite repeats in *Danio rerio*, *Oreochromis latipes*, and *Oreochromis niloticus* [33]. The overall tendency of motif frequency showed a pattern similar to that of previous studies on other fish [34]. These results could be relevant in identifying *C. dewitti* microsatellite markers.

Based on microsatellite motif analysis and select candidate primer pairs according to the parameter of QDD pipeline, we identified 84 primer pairs and randomly selected 30 primer pairs to verify by PCR. Finally, out of the 30 primer pairs, 15 primer pairs were confirmed with a single band. Recent research on the notothenioid fish *Lepidotothen nudifrons* to evaluate the genetic structure [35] and a study of population genetics in the Patagonian toothfish, *Dissostichus eleginoides* were performed using microsatellite markers [36]. The microsatellite markers that we classified provide a chance to profile the demographic events that occurred in the past and identify the characteristics of local adaptation in genetically different populations [35]. Understanding the importance of these events and the climate challenges facing polar marine organisms is essential for modeling and projecting future population viability for species management and conservation [37].

4. Conclusions

The crocodile icefish, *Chionobathyscus dewitti*, belongs to the family Channichthyidae and inhabits the deep sea around the Southern Ocean. The genome of the crocodile icefish was assembled, and microsatellite motifs were identified. The estimated genome size was 0.88 Gb based on K-mer analysis, and a total of 2,252,265 microsatellite motifs were identified. Dinucleotide repeats (1,920,127; 85.25%) had the highest rate. We selected 84 primer pairs based on the assembled genome and 15 primer pairs were validated as microsatellite markers. These genomic data provide the foundation for the development of novel molecular markers in a distinct population. However, more research is needed, such as more accurate genome assembly at the chromosomal level and more microsatellite validation experiments in the future.

Supplementary Materials: The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/ani12192598/s1>, Table S1: Genome estimation of *Chionobathyscus dewitti* using K-mer analysis; Table S2: Completeness of *Chionobathyscus dewitti* genome assembly estimated by BUSCO.

Author Contributions: H.P. and J.-H.K. conceived the study. J.K., S.-J.L., E.J., E.C., M.C. and S.C. performed genome sequencing, assembly, and analysis. J.K. and H.P. prepared the manuscript. All authors contributed to writing and editing the manuscript, collating supplementary information, and creating the figures. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by Korea Institute of Marine Science & Technology Promotion (KIMST) grant funded by the Ministry of Oceans and Fisheries (KIMST 20220547) and a Korea University Grant.

Institutional Review Board Statement: The procedures adopted in this study were performed in accordance with the guidelines of the Animal Welfare Ethical Committee and Animal Experimental Ethics Committee of the Korea Polar Research Institute (KOPRI, Incheon, Korea).

Data Availability Statement: The raw data generated by the Illumina Novaseq 6000 platform are displayed and available at NCBI (<https://www.ncbi.nlm.nih.gov/bioproject/PRJNA730390> (accessed on 28 August 2022)). The accession number for the BioProject is PRJNA730390 and for the SRA data is SRR14562209.

Acknowledgments: The authors gratefully acknowledge Steven J. Parker for providing samples.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Di Prisco, G.; Eastman, J.T.; Giordano, D.; Parisi, E.; Verde, C. Biogeography and adaptation of Notothenioid fish: Hemoglobin function and globin-gene evolution. *Gene* **2007**, *398*, 143–155. [[CrossRef](#)]
2. Scher, H.D.; Martin, E.E. Timing and Climatic Consequences of the Opening of Drake Passage. *Science* **2006**, *312*, 428–430. [[CrossRef](#)]
3. Toggweiler, J.R.; Bjornsson, H. Drake Passage and palaeoclimate. *J. Quat. Sci.* **2000**, *15*, 319–328. [[CrossRef](#)]
4. Verde, C.; Giordano, D.; Russo, R.; di Prisco, G. BLOOD | Erythropoiesis in Fishes. In *Encyclopedia of Fish Physiology*; Farrell, A.P., Ed.; Academic Press: San Diego, CA, USA, 2011; pp. 992–997.
5. Kim, K.-Y.; Yoon, M.; Yoo, J.S.; Lee, D.-S. The full-length mitochondrial genome of the crocodile icefish, *Chionobathyscus dewitti* (Teleostei: Perciformes: Channichthyidae). *Mitochondrial DNA Part B Resour.* **2019**, *4*, 2371–2372. [[CrossRef](#)]
6. Wells, R.M.G.; Ashby, M.D.; Duncan, S.J.; Macdonald, J.A. Comparative study of the erythrocytes and haemoglobins in nototheniid fishes from Antarctica. *J. Fish Biol.* **1980**, *17*, 517–527. [[CrossRef](#)]
7. Di Prisco, G. A study of hemoglobin in antarctic fishes: Purification and characterization of hemoglobins from four species. *Comp. Biochem. Physiol. Part B Comp. Biochem.* **1988**, *90*, 631–637. [[CrossRef](#)]
8. Yoon, T.-H.; Kang, H.-E.; Lee, S.R.; Lee, J.-B.; Baek, G.W.; Park, H.; Kim, H.-W. Metabarcoding analysis of the stomach contents of the Antarctic Toothfish (*Dissostichus mawsoni*) collected in the Antarctic Ocean. *PeerJ* **2017**, *5*, e3977. [[CrossRef](#)]
9. Sutton, C.; Manning, M. Biological parameters for icefish (*Chionobathyscus dewitti*) in the Ross Sea, Antarctica. *CCAMLR Sci.* **2008**, *15*, 139–165.
10. Agostini, C.; Patarnello, T.; Ashford, J.R.; Torres, J.J.; Zane, L.; Papetti, C. Genetic differentiation in the ice-dependent fish *Pleuragramma antarctica* along the Antarctic Peninsula. *J. Biogeogr.* **2015**, *42*, 1103–1113. [[CrossRef](#)]
11. Agostini, C.; Papetti, C.; Patarnello, T.; Mark, F.C.; Zane, L.; Marino, I.A.M. Putative selected markers in the *Chionodraco* genus detected by interspecific outlier tests. *Polar Biol.* **2013**, *36*, 1509–1518. [[CrossRef](#)]
12. Zane, L.; Bargelloni, L.; Patarnello, T. Strategies for microsatellite isolation: A review. *Mol. Ecol.* **2002**, *11*, 1–16. [[CrossRef](#)]
13. Fernandez-Silva, I.; Whitney, J.; Wainwright, B.; Andrews, K.R.; Ylitalo-Ward, H.; Bowen, B.W.; Toonen, R.J.; Goetze, E.; Karl, S.A. Microsatellites for Next-Generation Ecologists: A Post-Sequencing Bioinformatics Pipeline. *PLoS ONE* **2013**, *8*, e55990. [[CrossRef](#)]
14. Megléc, E.; Pech, N.; Gilles, A.; Dubut, V.; Hingamp, P.; Trilles, A.; Grenier, R.; Martin, J.-F. QDD version 3.1: A user-friendly computer program for microsatellite selection and primer design revisited: Experimental validation of variables determining genotyping success rate. *Mol. Ecol. Resour.* **2014**, *14*, 1302–1313. [[CrossRef](#)]
15. Ward, R.D.; Zemlak, T.S.; Innes, B.H.; Last, P.R.; Hebert, P.D. DNA barcoding Australia's fish species. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **2005**, *360*, 1847–1857. [[CrossRef](#)] [[PubMed](#)]
16. Marçais, G.; Kingsford, C. A fast, lock-free approach for efficient parallel counting of occurrences of k-mers. *Bioinformatics* **2011**, *27*, 764–770. [[CrossRef](#)]
17. Vurture, G.W.; Sedlazeck, F.J.; Nattestad, M.; Underwood, C.J.; Fang, H.; Gurtowski, J.; Schatz, M.C. GenomeScope: Fast reference-free genome profiling from short reads. *Bioinformatics* **2017**, *33*, 2202–2204. [[CrossRef](#)]
18. Zimin, A.V.; Marçais, G.; Puiu, D.; Roberts, M.; Salzberg, S.L.; Yorke, J.A. The MaSuRCA genome assembler. *Bioinformatics* **2013**, *29*, 2669–2677. [[CrossRef](#)]
19. Manni, M.; Berkeley, M.R.; Seppey, M.; Simão, F.A.; Zdobnov, E.M. BUSCO Update: Novel and Streamlined Workflows along with Broader and Deeper Phylogenetic Coverage for Scoring of Eukaryotic, Prokaryotic, and Viral Genomes. *Mol. Biol. Evol.* **2021**, *38*, 4647–4654. [[CrossRef](#)]

20. Sollars, E.S.A.; Harper, A.L.; Kelly, L.J.; Sambles, C.M.; Ramirez-Gonzalez, R.H.; Swarbreck, D.; Kaithakottil, G.; Cooper, E.D.; Uauy, C.; Havlickova, L.; et al. Genome sequence and genetic diversity of European ash trees. *Nature* **2017**, *541*, 212–216. [[CrossRef](#)]
21. Li, G.-Q.; Song, L.-X.; Jin, C.-Q.; Li, M.; Gong, S.-P.; Wang, Y.-F. Genome survey and SSR analysis of *Apocynum venetum*. *Biosci. Rep.* **2019**, *39*, BSR20190146. [[CrossRef](#)]
22. Katti, M.V.; Ranjekar, P.K.; Gupta, V.S. Differential Distribution of Simple Sequence Repeats in Eukaryotic Genome Sequences. *Mol. Biol. Evol.* **2001**, *18*, 1161–1167. [[CrossRef](#)] [[PubMed](#)]
23. Sidell, B.D.; Vayda, M.E.; Small, D.J.; Moylan, T.J.; Londraville, R.L.; Yuan, M.-L.; Rodnick, K.J.; Eppley, Z.A.; Costello, L. Variable expression of myoglobin among the hemoglobinless Antarctic icefishes. *Proc. Natl. Acad. Sci. USA* **1997**, *94*, 3420–3424. [[CrossRef](#)]
24. Choi, E.; Lee, S.J.; Jo, E.; Kim, J.; Parker, S.J.; Kim, J.-H.; Park, H. Genomic Survey and Microsatellite Marker Investigation of Patagonian Moray Cod (*Muraenolepis orangiensis*). *Animals* **2022**, *12*, 1608. [[CrossRef](#)] [[PubMed](#)]
25. Detrich, H.W.; Stuart, A.; Schoenborn, M.; Parker, S.K.; Methé, B.A.; Amemiya, C.T. Genome enablement of the notothenioid: Genome size estimates from 11 species and BAC libraries from 2 Representative Taxa. *J. Exp. Zool. Part B Mol. Dev. Evol.* **2010**, *314B*, 369–381. [[CrossRef](#)]
26. Lee, S.J.; Kim, J.-H.; Jo, E.; Choi, E.; Kim, J.; Choi, S.-G.; Chung, S.; Kim, H.-W.; Park, H. Chromosomal assembly of the Antarctic toothfish (*Dissostichus mawsoni*) genome using third-generation DNA sequencing and Hi-C technology. *Zool. Res.* **2021**, *42*, 124–129. [[CrossRef](#)]
27. Chen, L.; Lu, Y.; Li, W.; Ren, Y.; Yu, M.; Jiang, S.; Fu, Y.; Wang, J.; Peng, S.; Bilyk, K.T.; et al. The genomic basis for colonizing the freezing Southern Ocean revealed by Antarctic toothfish and Patagonian robalo genomes. *GigaScience* **2019**, *8*, giz016. [[CrossRef](#)]
28. Ahn, D.-H.; Shin, S.C.; Kim, B.-M.; Kang, S.; Kim, J.-H.; Ahn, I.; Park, J.; Park, H. Draft genome of the Antarctic dragonfish, *Parachaenichthys charcoti*. *GigaScience* **2017**, *6*, 1–6. [[CrossRef](#)] [[PubMed](#)]
29. Cheung, M.-S.; Down, T.A.; Latorre, I.; Ahringer, J. Systematic bias in high-throughput sequencing data and its correction by BEADS. *Nucleic Acids Res.* **2011**, *39*, e103. [[CrossRef](#)] [[PubMed](#)]
30. Zhou, W.; Hu, Y.; Sui, Z.; Fu, F.; Wang, J.; Chang, L.; Guo, W.; Li, B. Genome Survey Sequencing and Genetic Background Characterization of *Gracilariopsis lemaneiformis* (Rhodophyta) Based on Next-Generation Sequencing. *PLoS ONE* **2013**, *8*, e69909. [[CrossRef](#)] [[PubMed](#)]
31. Zeng, C.; Gao, Z.; Luo, W.; Liu, X.; Wang, W. Characteristics of microsatellites in blunt snout bream (*Megalobrama amblycephala*) EST sequences using 454FLX. *Acta Hydrobiol. Sin.* **2013**, *37*, 982–988. [[CrossRef](#)]
32. Li, Q.; Li, Z.; Dai, G.; Cao, Y.; Chen, X.; Chen, L.; Shangguan, J.; Ning, Y. Isolation and characterization of eleven microsatellite loci in the marbled rockfish, *Sebastes marmoratus* (Scorpaenidae). *Conserv. Genet. Resour.* **2013**, *6*, 53–55. [[CrossRef](#)]
33. 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*, 1–22. [[CrossRef](#)] [[PubMed](#)]
34. Xu, S.-Y.; Song, N.; Xiao, S.-J.; Gao, T.-X. Whole genome survey analysis and microsatellite motif identification of *Sebastes marmoratus*. *Biosci. Rep.* **2020**, *40*, BSR20192252. [[CrossRef](#)] [[PubMed](#)]
35. Papetti, C.; Harms, L.; Jürgens, J.; Sandersfeld, T.; Koschnick, N.; Windisch, H.S.; Knust, R.; Pörtner, H.-O.; Lucassen, M. Microsatellite markers for the notothenioid fish *Lepidonotothen nudifrons* and two congeneric species. *BMC Res. Notes* **2016**, *9*, 238. [[CrossRef](#)] [[PubMed](#)]
36. Garcia, K.K.; Touma, J.; Bravo, S.; Leiva, F.; Vargas-Chacoff, L.; Valenzuela, A.; Datagnan, P.; Amthauer, R.; Reyes, A.; Vidal, R. Novel microsatellite markers discovery in Patagonian toothfish (*Dissostichus eleginoides*) using high-throughput sequencing. *Mol. Biol. Rep.* **2019**, *46*, 5525–5530. [[CrossRef](#)]
37. Papetti, C.; Windisch, H.S.; La Mesa, M.; Lucassen, M.; Marshall, C.; Lamare, M.D. Non-Antarctic notothenioids: Past phylogenetic history and contemporary phylogeographic implications in the face of environmental changes. *Mar. Genom.* **2016**, *25*, 1–9. [[CrossRef](#)]