fortune Journals

Research Article

Genome-Wide Microsatellite Development for Genetic Assessment in *Takifugu obscurus* Restocking Programs

Jiyoung Woo[1], Jihye Park[1], Mi-Jin Yim[1], Seok-Chun Ko[1], Biet Thanh Tran[2], Keun-Yong Kim[2], Jeehyo Song[2], Moo-Sang Kim[3], Kyung Hoon Hahm[4], Dong-Hoon Lee[5], Moongeun Yoon[1*]

Abstract

The obscure pufferfish, *Takifugu obscurus*, is an economically important species in South Korea. It has become endangered due to overfishing and habitat degradation during the last decades. A national restocking program has been initiated to restore wild populations by releasing cultured juveniles into river systems. To ensure the success of this effort, maintaining genetic diversity and understanding population structure in both wild and aquaculture populations are critical. In this study, we successfully developed 14 polymorphic microsatellite loci for *T. obscurus*, demonstrating their high amplification success, strong polymorphism (PIC = 0.45–0.84), low null allele frequency, using next-generation sequencing (NGS) from three whole-genome shotgun libraries. These loci were used to assess the genetic diversity and population structure of 150 *T. obscurus* individuals from five populations: two wild adult populations from Busan (BAW) and Gyeonggi-do (GAW), and three aquaculture juvenile populations from Busan (BJA), Chungnam (CJA), and Gyeonggi-do (GJA). The genetic diversity analysis indicates that the two wild populations, BAW and GAW exhibit high genetic diversity, with a mean number of alleles per locus (Na) of 6.286 for both populations, mean number of effective alleles (Ne) of 3.453 and 3.478, mean observed heterozygosity (Ho) of 0.658 and 0.657, mean expected heterozygosity (He) of 0.680 and 0.677, and polymorphic information content (PIC) of 0.637 and 0.634, respectively. This high level of genetic diversity could serve as a crucial benchmark for selecting broodstock and monitoring the genetic health of restocked populations. Meanwhile, one aquaculture population, GJA showed similar genetic diversity and clustered closely with the wild populations, suggesting its suitability for restocking. In contrast, the other two, BJA and CJA displayed lower genetic diversity and significant genetic differentiation from the wild adult populations, raising concerns about their compatibility for restocking efforts. This information is vital for guiding future management actions and ensuring the long-term success and sustainability of restocking programs of *T. obscurus*.

Keywords: Microsatellites; Obscure pufferfish; *Takifugu obscurus*; Population structure; Next-generation sequencing; Restocking program

Introduction

The obscure pufferfish, *Takifugu obscurus* (Abe, 1949) is a euryhaline and anadromous species found in coastal and inland waters of China and the Korean Peninsula [1,2]. While it thrives in marine environments, it migrates to brackish and freshwater areas to spawn [1,2]. The juveniles grow in

Affiliation:
[1]National Marine Biodiversity Institute of Korea, Seocheon 33662, Republic of South Korea
[2]Genetic Analysis Team, AquaGenTech Co., Ltd., Busan 48228, Republic of South Korea
[3]Department of Bioinformatics, theMOAGEN, Daejeon, Republic of South Korea
[4]Busan Marine Fisheries Resource Research Institute, Busan 46763, Republic of South Korea
[5]Gyeonggi Province Maritime and Fisheries Research Institute, Yangpyeong 12513, Republic of South Korea

***Corresponding author:**
Moongeun Yoon, National Marine Biodiversity Institute of Korea, Seocheon 33662, Republic of South Korea.

Citation: Jiyoung Woo, Jihye Park, Mi-Jin Yim, Seok-Chun Ko, Biet Thanh Tran, Keun-Yong Kim, Jeehyo Song, Moo-Sang Kim, Kyung Hoon Hahm, Dong-Hoon Lee, Moongeun Yoon. Genome-Wide Microsatellite Development for Genetic Assessment in *Takifugu Obscurus* Restocking Programs. Journal of Biotechnology and Biomedicine. 8 (2025): 72-81.

Received: January 10, 2025
Accepted: January 27, 2025
Published: March 26, 2025

fortune Journals
Woo J, et al., J Biotechnol Biomed 2025
DOI:10.26502/jbb.2642-91280180

inland waters and return to sea after a few months or in the following spring [1]. *T. obscurus* is highly valued in China and South Korea for its high-quality meat and market price [2,3]. However, overfishing and environmental pollution have significantly reduced its wild populations, leading South Korea during the last decades to classify it as a protected species [4,5]. In response, South Korea has implemented a national restocking program, releasing cultured juveniles (approximately 5 cm in length) into river systems to support population recovery and conservation efforts [2,5]. This program aims to restore fishery stocks, increase spawning biomass, and promote conservation. Genetic variation and population structure are critical factors influencing the success of restocking programs [6]. Genetic variation serves as a foundation of adaptability within species and populations, enabling *T. obscurus* to respond to environmental changes and ensuring its survival and evolutionary potential [7]. Maintaining high genetic diversity in stocking programs is essential, as it ensures that released individuals have the adaptability needed to survive and reproduce in the wild. This genetic diversity enables populations to respond effectively to environmental changes and prevents the negative effects of inbreeding, which can reduce fitness and survival [7]. Conversely, population structure refers to the distribution of genetic variation within and between populations, often shaped by geographic, environmental, or reproductive barriers [8]. In aquaculture, divergent breeding practices can also create distinct population structures that differ significantly from those of wild populations. Such differences can reduce compatibility between restocked and wild fish, potentially disrupting local adaptations or causing outbreeding depression [9,10.11]. Therefore, investigating genetic variation and population structure is crucial for designing effective restocking programs for *T. obscurus* that ensure the long-term viability of both wild and restocked populations.

Among various molecular markers used to investigate genetic variation and population structure in fisheries, microsatellite markers are powerful tools due to their high polymorphism, codominant inheritance, ability to detect fine-scale genetic differences, abundance in many fish genomes, and ease of use [12,13]. For microsatellite markers to be effective, they must be PCR-amplifiable and contain length-altering polymorphisms within the repeat motif sequence. Traditional methods for screening and isolating polymorphic microsatellite markers are labor-intensive, costly, and often yield a limited number of loci [14,15]. Recently, microsatellite marker development using next-generation sequencing (NGS) from whole-genome shotgun libraries has become a fast and cost-effective alternative, enabling the simultaneous identification of tens of thousands of markers across the genome of a target species [16]. This NGS-based approach is increasingly applied to develop markers in various species, including carp [17], salmon [18], shark [19], and mackerel [20]. However, these studies often rely on genomic data from a single individual to identify and mine microsatellite markers, followed by the random selection of a few hundred markers for polymorphism detection through PCR experiments. The design, selection, and testing of polymorphic microsatellite markers thus remain labor-intensive and frequently yield low success rates [14].

Optimization of the NGS-based approach for detecting highly polymorphic microsatellite loci prior to experimental validation can be achieved using bioinformatics software to genotype multiple individuals with resequencing data [14,16]. Sequence reads from all individuals are aligned to a reference genome, enabling identifications of polymorphic microsatellite loci before primer development. This optimized method significantly improves efficiency of the developed microsatellite markers and reduces experimental efforts [17]. Wang et al. [21] successfully applied this pipeline to identify highly polymorphic microsatellite loci in a globally threatened species, brown eared-pheasant (*Crossoptilon mantchuricum*), known for its low genetic diversity. Despite the importance of genetic markers in studying population structure, supporting conservation, and restocking programs for economically significant species, *T. obscurus* remains underrepresented in available genetic resources. Currently, only a limited number of markers are available for this species [22], highlighting the need for further research to develop and characterize additional microsatellite markers. This study aims to (1) develop and validate 14 polymorphic microsatellite loci from three whole-genome shotgun libraries of *T. obscurus*; (2) use these loci to assess the genetic diversity and population structure of 150 *T. obscurus* individuals from five populations, including two wild adault populations and three aquaculture juvenile populations for restocking from South Korea; and (3) evaluate genetic characterization of the wild and aquaculture populations of *T. obscurus* to assess their suitability for use in restocking programs.

Materials and Methods

Sample collection and genomic DNA (gDNA) extraction

A total of 150 *T. obscurus* individuals were collected from five populations in South Korea, including two wild adult populations from Busan (BAW) and Gyeonggi-do (GAW), and three aquaculture juvenile populations from Busan (BJA), Chungnam (CJA), and Gyeonggi-do (GJA) (**Table 1**). The pelvic fin of each individual was preserved in absolute ethanol for subsequent gDNA extraction following the protocol of Asahida et al. [23]. gDNA quality was assessed with a NanoDrop One Microvolume UV-Vis Spectrophotometer (Thermo Fisher Scientific Inc., Wilmington, DE, USA). Each sample with the highest gDNA concentration from BAW, GAW, and CJA populations was selected for library preparation, sequencing, and microsatellite screening.

fortune Journals
Woo J, et al., J Biotechnol Biomed 2025
DOI:10.26502/jbb.2642-91280180

Table 1: Sampling details for *Takifugu obscurus* populations in South Korea

Sampling site	Abbreviation	Sampling size	Developmental stage	Source
Busan	BAW	24	Adult	Wild
Busan	BJA	30	Juvenile	Aquaculture
Gyeonggi-do	GAW	30	Adult	Wild
Chungnam	CJA	30	Juvenile	Aquaculture
Gyeonggi-do	GJA	36	Juvenile	Aquaculture

Whole-genome sequencing and *de novo* assembly

gDNA was fragmented using an S220 Ultra Sonicator (Covaris, Woburn, MA) under the following conditions: peak incident power at 175 W, duty factor of 5%, 200 cycles per burst, and a sonication duration for 35 s. Libraries were prepared using MGIEasy Duplex UMI Adapters Kit (MGI Tech, Wuhan, China), following the manufacturer's protocol. Library quality and fragment size distribution were assessed using an Agilent 2100 Bioanalyzer (Agilent Technologies, Santa Clara, CA). Quantification of the libraries was performed with the PicoGreen™ dsDNA Assay Kit (Thermo Fisher Scientific). After normalization, sequencing was carried out on an MGISEQ-2000 platform (MGI Tech) to generate 150 bp paired-end reads. Raw sequencing reads were processed by trimming adapter sequences, filtering out low-quality reads (< Q20) and short-length reads (< 50 bp), and removing duplicate reads using Cutadapt v3.1. Cleaned reads were then *de novo* assembled in CLC Genomics Workbench (CLC bio, Aarhus, Denmark) using default parameters, with a similarity threshold of 0.8. Consensus sequences from the assembled contigs were subsequently used for microsatellite identification.

Microsatellite identification and design

Genome-wide microsatellites in *T. obscurus* were identified using the MISA-web (http://misaweb.ipk-gatersleben.de/), which detects perfect mono-, di-, tri-, tetra-, penta-, and hexa-nucleotide repeat motifs under default settings. To enhance amplification success and polymorphism of the microsatellite markers, only microsatellites identified across three resequencing libraries with varied motif lengths were retained. Flanking regions of the selected microsatellites were extracted for primer design in Primer3 (https://primer3.ut.ee/), using the following criteria: perfect repeat motifs with ≥ 5 tandem repeats; primer length between 18 and 23 nucleotides, with an optimal length of 20 nucleotides; melting temperature (T_m) between 58°C and 64°C, with 60°C as optimal; PCR product sizes from 100 to 400 bp; GC content between 30% and 70%, with 50% as optimal; and no significant complementarity between primers.

Microsatellite loci validation

To develop microsatellite markers, 50 primer pairs were initially selected as candidates. Their amplification performance was evaluated using gDNA from five *T. obscurus* individuals, each sourced from a distinct population. Fifteen primer pairs with the strongest PCR amplification, as visualized by gel electrophoresis, were retained. Forward primers were fluorescently labeled, and primer pairs were divided into two assay sets (A and B). These assays were further validated with 150 *T. obscurus* individuals from five populations. Multiplex PCR for each set was conducted in 25 µL reactions using *AccuPower*® Gold Multiplex PCR PreMix (Bioneer, Daejeon, South Korea) on a ProFlex PCR System (Thermo Fisher Scientific). The fluorescent PCR products were size-separated on an ABI 3730*xl* DNA Analyzer (Thermo Fisher Scientific), with allele sizes calibrated against GeneScan™ 500 LIZ™ Size Standard (Thermo Fisher Scientific) and analyzed using GeneMapper v4.0 (Thermo Fisher Scientific). For each microsatellite locus, the number of alleles per locus (Na), observed heterozygosity (Ho), expected heterozygosity (He), polymorphic information content (PIC), and null allele frequency (F_{null}) were calculated using Cervus v3.0.7 [24] and GenAlEx v6.5 [25,26]. Microsatellite quality control, including scoring errors, allelic dropout, and null allele detection, was performed in Micro-Checker v2.2.3 [27]. Linkage disequilibrium (LD) between loci was tested in Genepop v4.7.5 [28,29], with significance thresholds adjusted by the Bonferroni correction.

Diversity analysis within population

Genetic diversity in five *T. obscurus* populations was analyzed by calculating the number of alleles (Na), effective number of alleles (Ne), observed heterozygosity (Ho), expected heterozygosity (He), polymorphic information content (PIC), and the inbreeding coefficient (F_{is}) using Cervus v3.0.7 and GenAlEx v6.5. Deviations from Hardy–Weinberg equilibrium (HWE) were assessed in Genepop v4.7.5 with default parameters, including 1,000 dememorizations, 100 batches, and 1,000 iterations per batch.

Population genetic structure

To determine the most appropriate measure of genetic differentiation between populations, an allele size randomization test was conducted using SPAGeDi v1.5 [30] with 10,000 permutations. The test indicated no significant mutation effect across all loci (R_{ST} = 0.073, P = 0.602; supplementary materials, **Table S1**), supporting F_{ST} as a more

Citation: Jiyoung Woo, Jihye Park, Mi-Jin Yim, Seok-Chun Ko, Biet Thanh Tran, Keun-Yong Kim, Jeehyo Song, Moo-Sang Kim, Kyung Hoon Hahm, Dong-Hoon Lee, Moongeun Yoon. Genome-Wide Microsatellite Development for Genetic Assessment in *Takifugu Obscurus* Restocking Programs. Journal of Biotechnology and Biomedicine. 8 (2025): 72-81.

fortune Journals
Woo J, et al., J Biotechnol Biomed 2025
DOI:10.26502/jbb.2642-91280180

reliable metric than R_{ST} for assessing genetic differentiation among five *T. obscurus* populations. Subsequent analyses used pairwise F_{ST} values, calculated in ARLEQUIN v3.1 [31] with 1,000 permutations. A genetic distance matrix based on Nei's method (Nei 1972) was generated in GenAlEx v6.5 and imported into MEGA v10.2.2 [32] to construct a cluster dendrogram using the Unweighted Pair Group Method with Arithmetic Mean (UPGMA). General genetic relationships among populations were visualized using principal coordinates analysis (PCoA) in GenAlEx v6.5. To analyze genetic structure across five *T. obscurus* populations, STRUCTURE v2.3.4 [33] was used with 10 independent runs for each K value (ranging from 1 to 10), applying 1,000,000 Markov Chain Monte Carlo iterations after a burn-in period of 100,000. The optimal K value was identified using STRUCTURE HARVESTER [34].

Results

Microsatellite abundance and distribution

A total of 67,980,814, 82,764,911, and 77,500,391 raw paired-end reads with Q30 scores > 93% were generated from three whole-genome libraries from each one individual from BAW, GAW, and CJA populations of *T. obscurus*, respectively; supplementary materials, **Table S2**). After bioinformatic processing, 7,898,266, 10,010,451, and 8,930,526 high-quality sequences were retained for microsatellite screening, resulting in the identification of 1,501,751, 1,934,234, and 1,694,218 microsatellite markers in each library, respectively (supplementary materials, **Table S3**). The distributions of five microsatellite types (di-, tri-, tetra-, penta-, and hexa-nucleotides) were consistent across all three libraries, with tri-nucleotide motifs being the most abundant (611,912, 799,783, and 701,752, respectively), followed by di-, tetra-, penta-, and hexa-nucleotide motifs (supplementary materials, **Table S4**). Among di-nucleotide motifs, AC/GT was the most frequent (26.37%, 27.07%, and 27.06%, respectively), while AGC/CTG, AAAT/ATTT, and AGAGG/CCTCT were the most common for tri- (10.34%, 10.63%, and 10.35%), tetra- (1.19%, 1.08%, and 1.11%), and penta-nucleotide motifs (0.21%, 0.22%, and 0.22%), respectively (supplementary materials, **Figure S1**).

Microsatellite loci validation

A total of 783,029 potential microsatellite loci of *T. obscurus* were analyzed for primer design. Fifteen optimal loci were selected and characterized from them, with allele sizes ranging from 118 to 402 bp (supplementary materials, **Table S5**). Genotyping success was high, with 150 out of 150 individuals successfully genotyped for most loci, except Tob11 (146), Tob10 (144), and Tob32 (146). The number of alleles per locus (Na) ranged from 5 to 14, observed heterozygosity (Ho) from 0.48 to 0.78, expected heterozygosity (He) from 0.50 to 0.85, and the polymorphic information content (PIC) from 0.45 to 0.84.

Genotyping quality was verified with MICRO-CHECKER (supplementary materials, **Table S5**), which detected no scoring errors such as stuttering or large allele dropout. While null alleles associated with an excess of homozygotes were detected at loci such as Tob24, Tob28, Tob38, Tob21, and Tob13, their frequencies were low ($F_{null} < 0.2$) in all cases except Tob38 ($F_{null} = 0.22$), leading to its exclusion from further analysis. No significant linkage disequilibrium was detected after applying the sequential Bonferroni correction.

Genetic diversity within population

The fourteen microsatellite loci finally selected in this study were used to assess the genetic diversity of 150 *T. obscurus* individuals from five populations in South Korea, comprising two wild adult populations (BAW and GAW) and three aquaculture juvenile populations for restocking (BJA, CJA, and GJA) (Table 2). Mean diversity values indicated that BAW, GAW, and GJA populations exhibited high genetic diversity, with mean Na values of 6.286, 6.286, and 6.357; mean Ne values of 3.453, 3.478, and 3.513; mean Ho values of 0.658, 0.657, and 0.620; mean He values of 0.680, 0.677, and 0.682; and mean PIC values of 0.637, 0.634, and 0.639, respectively (**Table 2**). In contrast, BJA and CJA populations displayed lower genetic diversity across all parameters, with mean Na of 4.643 and 4.071; mean Ne of 2.821 and 2.344; mean Ho of 0.707 and 0.549; mean He of 0.594 and 0.554; and mean PIC of 0.537 and 0.498, respectively. The mean F_{is} values indicated low or slightly positive inbreeding levels in most populations, with BJA showing the lowest F_{is} (–0.19), indicating an excess of heterozygotes, GJA having a moderate F_{is} (0.08), and CJA showing a near-zero F_{is} (0.004). Deviations from HWE were most prominent in BJA, CJA, and GJA, with significant deviations in ten, seven, and eight loci, respectively, while BAW and GAW showed deviations in only two and three loci, respectively. Loci such as Tob11, Tob34, Tob24, and Tob21 deviated from HWE in more than three populations, while Tob20, Tob07, and Tob32 deviated in fewer than two populations.

Population genetic structure

Pairwise F_{ST} analysis indicated moderate genetic differentiation between CJA population and the other four, with significant F_{ST} values of 0.099 (BAW), 0.159 (BJA), 0.121 (GAW), and 0.112 (GJA) ($P < 0.05$) across 14 microsatellite loci (**Table 3**). In contrast, genetic differentiation among the remaining populations (BAW, BJA, GAW, and GJA) was lower, with pairwise F_{ST} values ranging from 0.001 to 0.056, indicating closer genetic similarity within this group.

The UPGMA dendrogram, constructed using Nei's standard genetic distance[35], showed a distinct clustering of CJA population, clearly separating it from the other four (**Figure 1**). The remaining four populations (BAW, BJA, GAW, and GJA) formed a second cluster, with BJA further separating as a distinct subgroup from BAW, GAW, and GJA.

fortune Journals

Woo J, et al., J Biotechnol Biomed 2025
DOI:10.26502/jbb.2642-91280180

Table 2: Genetic diversity indices for five *Takifugu obscurus* populations in South Korea (BAW, BJA, GAW, CJA, and GJA) based on 14 microsatellite markers

Pop	Tob14	Tob11	Tob34	Tob24	Tob22	Tob28	Tob39	Tob20	Tob07	Tob29	Tob21	Tob10	Tob13	Tob32	Mean
BAW (24)															
Na	6	11	5	5	8	7	4	6	6	4	9	3	8	6	6.286
Ne	3.827	5.647	3.182	2.549	3.84	4.251	2.318	3.815	2.304	1.837	4.235	2.327	4.702	3.512	3.453
Ho	0.833	0.833	0.75	0.417	0.833	0.625	0.625	0.708	0.5	0.417	0.833	0.375	0.667	0.792	0.658
He	0.739	0.823	0.686	0.608	0.74	0.765	0.569	0.738	0.566	0.456	0.764	0.57	0.787	0.715	0.68
PIC	0.692	0.801	0.642	0.531	0.71	0.727	0.518	0.691	0.531	0.413	0.74	0.488	0.763	0.666	0.637
Fis	-0.128	-0.013	-0.094	0.314	-0.127	0.183	-0.099	0.04	0.117	0.086	-0.091	0.342	0.153	-0.107	0.041
HWE	0.802	**0.003**	0.197	0.078	0.954	0.144	0.641	0.223	0.145	0.336	0.6	**0.043**	0.057	0.307	
BJA (30)															
Na	5	7	4	4	4	5	4	4	3	4	7	3	6	5	4.643
Ne	3.303	5.158	2.459	1.854	1.978	3.34	1.556	3.93	1.703	2.187	4.688	1.795	2.748	2.799	2.821
Ho	0.967	0.967	0.8	0.333	0.567	0.967	0.433	0.933	0.533	0.733	0.867	0.6	0.6	0.6	0.707
He	0.697	0.806	0.593	0.461	0.494	0.701	0.357	0.746	0.413	0.543	0.787	0.443	0.636	0.643	0.594
PIC	0.647	0.778	0.521	0.393	0.424	0.649	0.321	0.698	0.359	0.471	0.753	0.358	0.566	0.576	0.537
Fis	-0.386	-0.199	-0.348	0.276	-0.146	-0.38	-0.213	-0.252	-0.292	-0.351	-0.102	-0.355	0.057	0.067	-0.188
HWE	**0**	**0**	**0.005**	**0.007**	**0.002**	**0**	0.757	**0.002**	0.325	**0.015**	**0.002**	**0.025**	0.231	0.541	
GAW (30)															
Na	6	12	5	5	8	6	4	6	4	5	8	3	9	7	6.286
Ne	3.704	5.341	3.291	3.377	4.615	3.607	2.007	3.468	1.985	2.043	5.882	2.171	3.719	3.488	3.478
Ho	0.733	0.867	0.733	0.533	0.8	0.7	0.467	0.6	0.533	0.533	0.733	0.667	0.667	0.633	0.657
He	0.73	0.813	0.696	0.704	0.783	0.723	0.502	0.712	0.496	0.511	0.83	0.539	0.731	0.713	0.677
PIC	0.687	0.791	0.654	0.658	0.757	0.68	0.453	0.658	0.443	0.47	0.808	0.438	0.707	0.675	0.634
Fis	-0.005	-0.066	-0.053	0.242	-0.021	0.032	0.07	0.157	-0.075	-0.045	0.116	-0.236	0.088	0.112	0.023
HWE	0.408	0.856	**0.01**	**0.004**	0.982	0.147	0.748	0.527	1	0.692	**0.021**	0.317	0.132	0.534	
CJA (30)															
Na	3	6	4	5	4	3	3	4	3	4	4	4	5	5	4.071
Ne	2.323	2.571	2.02	2.699	3.12	1.674	2.532	1.861	1.653	1.881	2.927	1.989	2.394	3.18	2.344
Ho	0.6	0.733	0.233	0.633	0.933	0.3	0.7	0.567	0.5	0.6	0.367	0.633	0.167	0.724	0.549
He	0.569	0.611	0.505	0.629	0.679	0.403	0.605	0.463	0.395	0.468	0.658	0.497	0.582	0.685	0.554
PIC	0.505	0.566	0.459	0.586	0.618	0.363	0.531	0.416	0.347	0.427	0.587	0.415	0.527	0.624	0.498
Fis	-0.054	-0.2	0.538	-0.006	-0.374	0.255	-0.157	-0.224	-0.266	-0.281	0.443	-0.274	0.714	-0.056	0.004
HWE	0.418	0.157	**0**	0.749	**0.003**	0.075	**0.005**	0.693	0.402	0.503	**0**	**0**	**0**	**0.031**	
GJA (38)															
Na	6	12	4	5	7	6	6	5	4	4	10	4	8	8	6.357
Ne	3.323	6.187	2.997	3.57	3.687	3.625	2.164	3.108	2.254	1.882	5.503	2.459	4.438	3.989	3.513
Ho	0.75	0.531	0.75	0.583	0.778	0.639	0.5	0.694	0.556	0.444	0.667	0.567	0.556	0.667	0.62
He	0.699	0.838	0.666	0.72	0.729	0.724	0.538	0.678	0.556	0.469	0.818	0.593	0.775	0.749	0.682
PIC	0.649	0.823	0.62	0.668	0.701	0.675	0.492	0.612	0.513	0.421	0.801	0.521	0.743	0.71	0.639
F_{is}	-0.073	0.366	-0.126	0.19	-0.067	0.118	0.07	-0.024	0.001	0.052	0.185	0.045	0.283	0.11	0.081
HWE	**0.034**	**0**	**0**	**0.003**	0.881	**0.009**	**0.001**	0.847	0.442	0.749	**0**	0.106	**0**	0.202	

N, sampling size; Na, number of alleles per locus; Ne, number of effective alleles; Ho, observed heterozygosity; He, expected heterozygosity; PIC, polymorphic information content; F_{is}, inbreeding coefficient. HWE, Hardy–Weinberg equilibrium probability test. Bold values indicate significant deviation of Hardy–Weinberg equilibrium probability test ($P < 0.05$)

Woo J, et al., J Biotechnol Biomed 2025
DOI:10.26502/jbb.2642-91280180

Table 3: Average pairwise F_{ST} values (lower diagonal) and associated statistical significance levels (upper diagonal) based on 14 microsatellite loci across five *Takifugu obscurus* populations in South Korea

Population	BAW	BJA	GAW	CJA	GJA
BAW	-	0.000*	0.418	0.000*	0.071
BJA	0.041	-	0.000*	0.000*	0.000*
GAW	0.001	0.047	-	0.000*	0.005*
CJA	0.099	0.159	0.121	-	0.000*
GJA	0.008	0.056	0.013	0.112	-

Statistically significant (*$P < 0.05$).

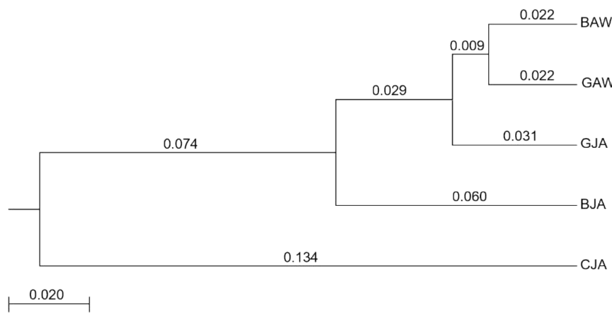

Figure 1: UPGMA dendrogram illustrating the genetic relationships among five *Takifugu obscurus* populations in South Korea, based on Nei's standard genetic distance calculated from 14 microsatellite loci.

The PCoA revealed that the first, second, and third axes explained 10.48%, 6.82%, and 5.86% of the genetic variation, respectively (Figure 2). The first axis demonstrated clear separation between CJA population and the other four. Although some overlap was observed among BAW, BJA, GAW, and GJA, the second axis suggested subtle differentiation of BJA population from the other three.

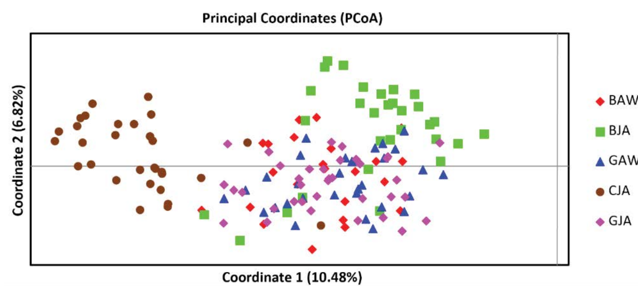

Figure 2: Biplot of principal coordinates analysis (PCoA) of five *Takifugu obscurus* populations in South Korea.

Using STRUCTURE HARVESTER and the Evanno method, the most likely number of genetic clusters was identified as K = 3 (supplementary materials, **Table S6**, **Figure S2**). At K = 3, individuals from the five populations clustered into three groups with high assignment probabilities; Cluster I comprised BJA, Cluster II included CJA, and Cluster III grouped BAW, GAW, and GJA (**Figure 3**). Some individuals exhibited admixture, suggesting possible gene flow or outliers across these clusters.

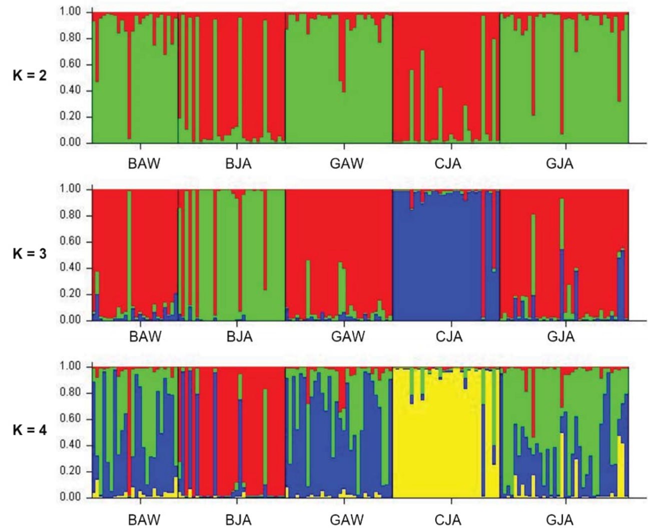

Figure 3: Bar plots showing the genetic structure of five *Takifugu obscurus* populations in South Korea, based on 14 microsatellite markers. Each individual is represented by a vertical line, with color segments indicating the probability of membership in each inferred cluster at K = 2, 3, and 4.

Discussion

The vast amount of data generated from NGS analysis in this study, which allowed for the direct mining of more than 1.5 million microsatellites across the *T. obscurus* genome, enabled the comprehensive selection of markers applicable to a wide range of genetic studies. This large-scale data mining greatly enhances the chances of identifying highly polymorphic and species-specific markers, essential for non-model organisms, such as *T. obscurus*. Unlike traditional methods, this pipeline automates both marker identification and primer design using tools like MISA and Primer3, making the process faster, more cost-effective, and scalable. Additionally, mining microsatellite markers using resequencing data from the three *T. obscurus* individuals, rather than relying on a single genome, enhance the quality and reliability of the developed markers. The effectiveness of this approach was validated in this study; 14 microsatellite markers selected exhibited high amplification success (100% in the eleven loci, > 96% in the remaining three), high polymorphism (PIC = 0.45–0.84), low null allele frequency ($F_{null} < 0.2$), no scoring errors, and no allelic dropout across five *T. obscurus* populations. These results highlight the effectiveness of the pipeline in generating robust, reliable markers for assessing the population genetics of both wild adult and aquaculture juvenile populations of *T. obscurus* in this study.

fortune Journals

Woo J, et al., J Biotechnol Biomed 2025
DOI:10.26502/jbb.2642-91280180

Genetic variation is crucial in restocking programs, as it provides key insights into the genetic health and viability of both wild and hatchery-reared populations. In this study, the genetic diversity of both wild adult and aquaculture juvenile populations of *T. obscurus* was found to be moderate-to-high. The wild adult populations collected from Busan (BAW) and Gyeonggi-do (GAW) exhibited particularly high genetic diversity, as indicated by their high numbers of alleles, heterozygosity, and polymorphic information content. These results are consistent with Ma (2009), making them essential reference points for broodstock selection and genetic monitoring. Such diversity ensures that broodstock used in hatcheries retains genetic variability comparable to that of natural populations, which is crucial for the adaptability and long-term survival of restocked fish. Using BAW and GAW populations as genetic benchmarks helps safeguard against genetic erosion in hatchery populations, thereby preserving the genetic health of both restocked and wild populations. Minor deviations from HWE were observed (two loci in BAW and three in GAW), along with slightly positive inbreeding coefficients (F_{is} of 0.04 for BAW and 0.023 for GAW) in the two wild adult populations suggesting that they maintain overall genetic stability. However, ongoing monitoring of BAW and GAW is necessary to prevent further inbreeding and to ensure they continue to serve as robust genetic reservoirs. Maintaining high genetic diversity and monitoring breeding patterns in these wild populations is crucial for their long-term viability and successful restocking efforts.

Among the three aquaculture juvenile populations for restocking (BJA, CJA, and GJA), GJA population showed genetic high diversity comparable to the two wild adult populations. This high level of genetic diversity is crucial for ensuring the adaptability, resilience, and long-term survival of restocked fish in the wild, making GJA a strong candidate for restocking programs. However, eight of the 14 microsatellite loci in GJA population showed deviations from HWE, along with a positive inbreeding coefficient (F_{is} = 0.081). A positive F_{is} suggests some degree of inbreeding or non-random mating, leading to a deficiency of heterozygotes and thus causing deviations from HWE. However, heterozygote deficiency can also be attributed to null alleles, technical artifacts, or small sampling size [36]. Despite these factors, the overall genetic profile of GJA, with its high diversity, makes it suitable for use in restocking programs, as long as ongoing management and monitoring are implemented to maintain genetic health and prevent further inbreeding. Although the other two populations (BJA and CJA) exhibited moderate genetic diversity, they fell short compared to the higher genetic diversity observed in the two wild adult populations, which can be a concern for restocking programs. Lower genetic diversity in these aquaculture juvenile populations can limit the adaptability of restocked fish to environmental changes, diseases, or other stressors in the wild. In addition, they showed a considerable number of loci deviating from HWE (ten loci in BJA and seven in CJA). The lower overall diversity and the relatively high number of loci deviating from HWE of compared to the wild adult populations make them less suitable for restocking programs, unless steps are taken to manage and improve their genetic profiles.

The aim of the restocking program for *T. obscurus* is to restore and enhance wild populations that have declined due to overfishing, habitat degradation, and environmental changes, while minimizing any adverse effects on the local gene pool. Reintroducing genotypes from genetically similar populations is expected to have smaller negative impacts compared to introducing genetically dissimilar populations, which can result in genetic homogenization, outbreeding depression, or disruption of local adaptations [9,37]. For instance, Cooke and Philipp [37] demonstrated that mixing two genetically distinct stocks of largemouth bass (*Micropterus salmoides*) led to a reduction in overall fitness, particularly affecting cardiovascular performance. In this study, population structure analysis using F_{ST}, UPGMA, PCoA, and STRUCTURE consistently showed that an aquaculture juvenile population, GJA clusters closely with two wild adult populations, BAW and GAW. This indicates that GJA shares a similar genetic background with the wild populations, making it a strong candidate for restocking programs. Its genetic similarity ensures that restocked individuals will likely integrate smoothly into the wild gene pool, preserving both natural genetic diversity and local adaptations. In contrast, the other two aquaculture juvenile populations, CJA and BJA are genetically distinct from each other and from GJA, BAW, and GAW cluster. This differentiation suggests that CJA and BJA may have experienced genetic drift or divergent breeding practices, making them less compatible with the wild populations. Introducing fish from these two populations into the wild without careful management could disrupt the natural genetic structure, potentially leading to maladaptation or reduced fitness in restocked populations.

While the genetic similarity of the aquaculture juvenile population, GJA to the wild adult populations suggests its suitability for restocking, potential risks associated with genetic homogenization and maladaptation warrant further consideration. Restocking programs that rely heavily on a single aquaculture population, like GJA, may inadvertently reduce genetic diversity in wild populations over time, leading to genetic homogenization. This can weaken local adaptations that are crucial for survival in varied natural environments. Additionally, while GJA may exhibit genetic profiles similar to wild populations in the study, it is possible that these cultured fish lack specific adaptive traits necessary to thrive in dynamic river systems. Managers should consider introducing fish from multiple, genetically diverse sources such as incorporating multiple broodstock lineages within GJA population or other aquaculture populations with

fortune Journals

Woo J, et al., J Biotechnol Biomed 2025
DOI:10.26502/jbb.2642-91280180

compatible genetic backgrounds, or implementing measures to maintain local genetic variation within wild populations. To optimize the genetic health of restocked populations, managers should monitor genetic diversity regularly—ideally on an annual or biennial basis—to track changes over time and detect any early signs of genetic drift or inbreeding, both during and after release into the wild. Criteria for broodstock selection should emphasize high genetic diversity, including maintaining heterozygosity levels comparable to wild populations (e.g., Ho > 0.65) and avoiding excessive use of any single aquaculture population. Thresholds for genetic diversity metrics, such as PIC, should be maintained to ensure adaptability; for instance, PIC values above 0.6 may be targeted to maintain robust genetic health. Given the limited number of markers and samples used in this study, future work should include a broader set of markers or genome-wide analyses and incorporate more samples across diverse wild and hatchery populations. This approach will enable a more comprehensive assessment of genetic diversity, population structure, and unique adaptive traits critical for *T. obscurus* resilience. Addressing these concerns and implementing systematic monitoring will help sustain the genetic resilience of *T. obscurus* populations over the long term.

Conclusions

This study highlights the successful development of 14 polymorphic microsatellite markers through NGS using three whole-genome shotgun libraries and their effective application in assessing genetic diversity and population structure in both wild adult and aquaculture juvenile populations of *T. obscurus*. The high genetic diversity observed in the two wild adult populations, BAW and GAW provides a valuable reference point for guiding restocking programs. Notably, the genetic similarity of an aquaculture population, GJA to the two wild adult populations suggests its suitability for integration into restocking programs, supporting the conservation and resilience of *T. obscurus* in natural habitats. In contrast, the lower genetic diversity and significant genetic differentiation in the other two aquaculture juvenile populations, BJA and CJA indicate potential challenges for their use in restocking unless further genetic management is applied. These findings underscore the importance of careful genetic monitoring and the selection of compatible broodstock to sustain the genetic health and adaptability of *T. obscurus* populations over the long term.

Acknowledgement

We would like to express our sincere thanks go to Chungcheongnam-do Institute of Fisheries Resources for providing the necessary resources for this study.

Funding

This Study was funded by National Marine Biodiversity Institute of Korea (MABIK) under in-house Research Programs [Development of Biomaterials derived from Marine Organisms, 2025M00500].

Data Availability

The microsatellite markers for *Takifugu obscurus* are available in GenBank (accession number PV101362–PV101375). Raw sequencing data are deposited in NCBI under Bioproject PRJNA1221654, Biosamples SAMN46748622–SAMN46748624, and SRA accessions SRS24029318–SRS24029320

Competing Interests

The authors declare that they have no competing of interests.

Compliance with Ethical Standards

This article does not contain any studies involving human participants performed by any of the authors. All experiments were performed following the guidelines of the South Korean Association for Laboratory Animals (approval no. 18-0680, 5 October 2018).

Supplementary Files:

https://fortunejournals.com/supply/JBB12122.pdf

References

1. Kato A, Doi H, Nakada T, Sakai H, Hirose S. *Takifugu obscurus* is a euryhaline fugu species very close to *Takifugu rubripes* and suitable for studying osmoregulation. BMC physiology 5 (2005): 1-11.

2. Kim JH, Wang SY, Kim IC, Ki JS, Raisuddin S, Lee JS, Han KN. Cloning of a river pufferfish (*Takifugu obscurus*) metallothionein cDNA and study of its induction profile in cadmium-exposed fish. Chemosphere 71 (2008): 1251-1259.

3. Yang Z, Chen YF. Induced ovulation in obscure puffer *Takifugu obscurus* by injections of LHRH-a. Aquaculture International 12 (2004): 215-223.

4. Chen Y, Yang Z. Diets of obscure puffer (*Takifugu obscurus*) and ocellated puffer (*Takifugu ocellatus*) during spawning migration. Journal of Freshwater Ecology 20 (2005): 195-196.

5. Kim JH, Dahms HU, Han KN. Biomonitoring of the river pufferfish, *Takifugu obscurus* in aquaculture at different rearing densities using stress-related genes. Aquaculture Research 44 (2013): 1835-1846.

6. Lopera-Barrero NM, Santos SCAD, Rodriguez-Rodriguez MDP, Fornari DC, Zancheta C, Poveda-Parra AR, et alGenetic diversity of wild populations and broodstocks

Woo J, et al., J Biotechnol Biomed 2025
DOI:10.26502/jbb.2642-91280180

of curimba for restocking programs in the Tietê, Grande, Pardo and Mogi-Guaçu rivers (Brazil). Boletim do Instituto de Pesca 41(2015): 287-304

7. Çiftci Y, Okumuş İ. Fish population genetics and applications of molecular markers to fisheries and aquaculture: I-Basic principles of fish population genetics. Turkish Journal of Fisheries and Aquatic Sciences 2 (2002): 145-155.
8. Ahmed T, Abbas K. Patterns of genetic variability in natural and hatchery populations of *Catla catla* based on microsatellite DNA markers. Pakistan Journal of Agricultural Sciences 55 (2018): 929-939.
9. Cross TF. Genetic implications of translocation and stocking of fish species, with particular reference to western Australia. Aquaculture Researc 31 (2000): 83-94.
10. Ward RD. The importance of identifying spatial population structure in restocking and stock enhancement programmes. Fisheries Research 80 (2006): 9-18.
11. Araki H, Schmid C. Is hatchery stocking a help or harm? Evidence, limitations and future directions in ecological and genetic surveys. Aquaculture 308 (2010): S2-S11.
12. Ellegren H. Microsatellites: simple sequences with complex evolution. Nature Reviews Genetics 5 (2004): 435-445.
13. Chistiakov DA, Hellemans B, Volckaert FA. Microsatellites and their genomic distribution, evolution, function and applications: a review with special reference to fish genetics. Aquaculture 255 (2006): 1-29.
14. Vukosavljev M, Esselink GD, van 't Westende WPC, Cox P, Visser RGF, Arens P, Smulders MJM. Efficient development of highly polymorphic microsatellite markers based on polymorphic repeats in transcriptome sequences of multiple individuals. Molecular Ecology Resources 15 (2015): 17-27.
15. Cui X, Huang X, Chen J, Yang X, Rong J. An efficient method for developing polymorphic microsatellite markers from high-throughput transcriptome sequencing: a case study of hexaploid oil-tea camellia (*Camellia oleifera*). Euphytica 214 (2018): 1-9.
16. Fox G, Preziosi RF, Antwis RE, Benavides-Serrato M, Combe FJ, Harris WE, et al. Multi-individual microsatellite identification: A multiple genome approach to microsatellite design (MiMi). Molecular Ecology Resources 19 (2019): 1672-1680.
17. Ji P, Zhang Y, Li C, Zhao Z, Wang J, Li J, et al. High throughput mining and characterization of microsatellites from common carp genome. International Journal of Molecular Sciences 13 (2012): 9798-9807.
18. Tsukagoshi H, Terui S, Abe S. Characterization of sixteen polymorphic microsatellite DNA loci in the chum salmon (*Oncorhynchus keta*) isolated by next-generation sequencing. Conservation Genetics Resources 7 (2015): 173-175.
19. Maduna SN, Rossouw C, Da Silva C, Soekoe M, Bester-van der Merwe AE. Species identification and comparative population genetics of four coastal houndsharks based on novel NGS-mined microsatellites. Ecology and Evolution 7 (2017): 1462-1486.
20. Joy L, Paulose S, Divya PR, Ravi C, Basheer VS, Kumar R, et al. Microsatellite marker development in Spanish mackerel *Scomberomorus commerson* using third generation sequencing technology. Molecular Biology Reports 47 (2020): 10005-10014.
21. Wang H, Gao S, Liu Y, Wang P, Zhang Z, Chen D. A pipeline for effectively developing highly polymorphic simple sequence repeats markers based on multi-sample genomic data. Ecology and Evolution 12(2022): e8705.
22. Ma H, Chen S, Liao X, Xu T, Ge J. Isolation and characterization of polymorphic microsatellite loci from a dinucleotide-enriched genomic library of obscure puffer (*Takifugu obscurus*) and cross-species amplification. Conservation Genetics 10 (2009): 955-957.
23. Asahida T, Kobayashi T, Saitoh K, Nakayama I. Tissue preservation and total DNA extraction form fish stored at ambient temperature using buffers containing high concentration of urea. Fisheries Science 62 (1996): 727-730.
24. Marshall TC, Slate JEBK, Pemberton JM. Statistical confidence for likelihood-based paternity inference in natural populations. Molecular Ecology 7 (1998): 639-655.
25. Peakall R, Smouse PE. GENALEX 6: genetic analysis in Excel. Population genetic software for teaching and research. Molecular Ecology Notes 6 (2006): 288-295.
26. Peakall R, Smouse PE. GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research-an update. Bioinformatics 28 (2012): 2537-2539.
27. Van Oosterhout C, Hutchinson WF, Wills DP, Shipley P. MICRO-CHECKER: software for identifying and correcting genotyping errors in microsatellite data. Molecular Ecology Notes 4 (2004): 535-538.
28. Raymond M, Rousset F. GENEPOP (version 1.2): population genetics software for exact tests and ecumenicism. Journal of Heredity 86 (1995): 248-249.
29. Rousset F. genepop'007: a complete re-implementation of the genepop software for Windows and Linux. Molecular Ecology Resources 8 (2008): 103-106.

30. Hardy OJ, Vekemans X. SPAGeDi: a versatile computer program to analyse spatial genetic structure at the individual or population levels. Molecular Ecology Notes 2 (2002): 618-620.

31. Excoffier L, Lischer HE. Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. Molecular Ecology Resources 10 (2010): 564-567.

32. Tamura K, Stecher G, Kumar S. MEGA11: molecular evolutionary genetics analysis version 11. Molecular Biology and Evolution 38 (2021): 3022-3027.

33. Pritchard JK, Stephens M, Donnelly P. Inference of population structure using multilocus genotype data. Genetics 155 (2000): 945-959.

34. Earl DA, VonHoldt BM. STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. Conservation Genetics Resources 4 (2012): 359-361.

35. Nei M. Genetic distance between populations. The American Naturalist 106 (1972): 283-292.

36. Valles-Jimenez R, Cruz P, Perez-Enriquez R. Population genetic structure of Pacific white shrimp (*Litopenaeus vannamei*) from Mexico to Panama: microsatellite DNA variation. Marine Biotechnology 6 (2004): 475-484.

37. Cooke SJ, Philipp DP. Influence of local adaptation and interstock hybridization on the cardiovascular performance of largemouth bass *Micropterus salmoides*. Journal of Experimental Biology 208 (2005): 2055-2062.