

## Identification and Characterization of Polymorphic Microsatellite DNA Markers Using Next-generation Sequencing in *Parapristipoma trilineatum*

Chun Mae Dong<sup>1</sup>, Mi-Nan Lee<sup>1</sup>, Jae Koo Noh<sup>2</sup>, Jin Woo Park<sup>2</sup>, Young-Ok Kim<sup>1</sup> and Eun-Mi Kim<sup>1\*</sup>

<sup>1</sup>Biotechnology Research Division, National Institute of Fisheries Science, Busan 46083, Korea

<sup>2</sup>Jeju Fisheries Research Institute, National Institute of Fisheries Science, Jeju 63068, Korea

Received July 3, 2023 / Revised August 12, 2023 / Accepted August 14, 2023

This study was conducted to develop microsatellite markers in *Parapristipoma trilineatum* using next-generation sequencing. A total of 402,244,934 reads were generated on the Illumina Hiseq X Ten System, yielding 60,738,985,034 bp of sequences. The de novo assembly resulted in 1,320,995 contigs. A total of 952,326 contigs (0.016%) including 151 microsatellite loci were derived from the 1,320,995 contigs longer than 640 bp. A total of 34 primer sets were designed from the 151 microsatellite loci. As a result, 15 microsatellite loci were chosen and used for assuming population genetic parameters in the wild and farmed populations. The mean number of effective alleles was 12, ranging from 6 to 25. The observed heterozygosity ( $H_o$ ) and the expected heterozygosity ( $H_e$ ) ranged between 0.530 and 0.873, with an average of 0.750, and from 0.647 to 0.895, with an average of 0.793, respectively. According to these results, the developed set of 15 microsatellite markers is expected to be useful for the analysis of genetic characteristics in the population of *P. trilineatum* in Korea. There are requirements now for further genetic information, fishery resource management, breeding guidelines, support with the selection of breeds and studies on the effects of release, all of which will improve species conservation, and through future research, we aim to offer genetic foundational data with that goal.

**Key words :** Genetic variability, microsatellite loci, microsatellite markers, next generation, *Parapristipoma trilineatum*

### 서 론

벤자리(*Parapristipoma trilineatum*)는 농어목(Perciformes) 하스돔과(Haemulidae)에 속하는 온대성 연안어종으로 쿠로시오 난류의 영향을 받는 연안 깊은 바위 지대에서 서식하며, 우리나라에서는 제주도과 추자도 남부 해역에서 주로 서식하는 제주 특산어종이다[21]. 형태적으로 몸 중앙을 따라 검은색 세 줄이 뚜렷하게 그려져 있으며, 이 세 줄은 성장 또는 서식지에 따라 희미해지거나 소실되는 특징을 가지고 있다[18].

벤자리는 우리나라와 일부 아시아 국가에서 최근 들어 미래 고부가가치 자원으로 각광받고 있지만, 여름철에 주로 어획되는 희소성 어종으로 생산량이 적고, 해양환경의 변화 및 오염 등으로 인해 서식지 분포 변화가 가속화되고 있다[20, 37]. 이에 일본과 중국의 경우 1999년대부터

벤자리 어업 자원량을 효율적으로 관리하기 위해 양식산업 및 종묘생산 등에 관한 연구가 활발히 이루어져 왔으며[13, 43], 2002년대부터는 유전학적 통계분석을 통해 벤자리의 인공종묘생산을 보다 체계적으로 관리하고 있다[2, 25]. 반면, 우리나라에서는 2014년에 제주도 해역에서 포획한 벤자리의 어미로부터 종묘를 생산하였으며, 2020년대 이후부터 벤자리의 대량 인공종묘생산 및 중간양성 기술개발 등의 연구가 이루어지고 있다.

현재까지 벤자리에 대한 국내 연구는 사육수의 고수온 스트레스가 벤자리에 미치는 생리학적 영향[21], 벤자리의 핵형 분석에 관한 연구[33] 등 생태학적 특성 및 자원생물학적 연구에 불과하며 유전학적 통계분석을 통해 벤자리 자원을 보다 체계적으로 관리할 수 있는 연구는 아직 미비한 실정이다.

벤자리의 유전학적 관리가 중요한 이유는 소규모 양식산 집단에서 어미들의 유전자형을 고려하지 않고 무작위적으로 교배하여 유전자형의 손실이 일어날 수 있고[36], 이로 인해 근교약세(inbreeding depression)로 이어져 생산된 자손의 성장 및 생존에 영향을 미칠 수 있기 때문이다[22]. 또한, 유전적 다양성의 부족으로 인한 병목현상, 기형 증가, 성장 저하, 저항성 감소, 집단 근친도의 증가, 질병 등 부정적인 영향을 미칠 수 있다[1, 12].

#### \*Corresponding author

Tel : +82-51-720-2462, Fax : +82-51-720-2456

E-mail : ocean0629@korea.kr

This is an Open-Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/3.0>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

대표적 양식어종인 연어[30], 넙치[19], 참돔[34] 및 전복[9] 등의 경우에도 한정된 어미로부터 종묘생산이 이루어져 왔기 때문에 자연산 집단에 비해 양식산 집단에서 대립유전자 수 및 유전자 빈도가 크게 감소되었다고 보고된 바 있다[23]. 따라서, 우리나라 벤자리의 인공종묘생산에 있어 체계적인 어미 관리를 위해 벤자리의 유전적 다양성 유지에 관한 연구가 필요하다.

초기의 분자 마커 종류로는 Randomly Amplified Polymorphic DNA; RAPD, Amplified Fragment Length Polymorphism; AFLP, Restriction Fragment Length Polymorphism; RFLP, Single Nucleotide Polymorphism; SNP, Simple Sequence Repeat; SSR 이 주로 사용되었지만[4], 이러한 기법들은 상대적으로 정보량이 적고 많은 노동력과 연구비용이 많이 드는 단점이 있다[44].

최근 기술의 발달로 차세대 염기서열(Next-Generation Sequencing, NGS) 분석방법을 사용하여 대상 종의 염기서열 정보를 대량으로 확보하고, 이를 통해 microsatellite 영역을 직접 탐색하여 전통적인 microsatellite 마커 개발법[31] 보다 재현성 및 다형성이 높은 microsatellite 마커를 개발하는 방법이 주로 사용되고 있다[14, 27, 42, 45]. 이러한 NGS 방법으로 개발된 microsatellite 마커는 높은 변이를 나타내는 유전자 표식으로 유전자 지도, 개체식별 및 친자확인, 집단 유전학 등을 분석하는데 활용되고 있다[32].

수산자원을 대상으로 NGS 방법을 사용한 연구로는 보리새우[46], 상어[11], 방어[8], 및 굴[7] 등 유전학적 특성 연구가 활발히 수행되었지만, 벤자리를 대상으로 NGS 방법을 적용한 유전학적 특성 연구는 현재까지 이루어지지 않았다.

따라서 본 연구는 벤자리의 인공종묘생산에 있어 체계적인 어미 관리를 위해 NGS 방법을 사용하여 새로운 microsatellite 영역을 탐색하고 이를 기반으로 재현성과 다형성이 높은 microsatellite 마커를 개발하였다. 새로 개발한 microsatellite 마커를 사용하여 벤자리 집단의 유전적 다양성 및 유연관계를 분석하여 향후, 벤자리의 유전학적 기초정보를 확보하는데 활용할 것이다.

## 재료 및 방법

### 시료 확보 및 genomic DNA 추출

본 연구에 사용된 벤자리 시료는 총 274개체이며, 서귀포 위미시험포 양식산 집단(2020년, n=102)과 서귀포 위미항 자연산 집단(2020년, n=122) 및 제주 한림 자연산 집단(2021년, n=50)을 분석에 사용하였다.

Genomic DNA의 순수분리와 정제는 DNeasy® 96 Blood & Tissue Kit (Qiagen GmbH, Germany)를 사용하였다. 분석 대상 시료 20 mg을 E-tube에 담아 DNeasy® 96 Blood &

Tissue Kit로 제조사의 분석방법에 따라 ATL buffer 180 µl와 Proteinase K 20 µl를 첨가한 후, 혼합하여 56°C에서 24시간 동안 반응시켰다. AL buffer 180 µl와 99% ethanol 180 µl을 첨가하여 S-Blocks의 DNeasy 96 plates에 옮겨 8,000 rpm (6,000× g)으로 1분간 원심분리 하였다. Column plates를 교체한 후, AW 1 buffer 450 µl 첨가하여 8,000 rpm (6,000× g)으로 1분간 원심분리 하였다. 이전과 같은 방법으로 column plates를 교체한 후, AW 2 buffer 450 µl 첨가하여 20,000 rpm (14,000× g)으로 3분간 원심분리 하였다. Ethanol을 제거하기 위해 공기 중에 말린 다음, AE buffer 100 µl를 첨가하여 genomic DNA를 회수하였다. 회수한 genomic DNA는 1.8% (w/v) agarose gel로 전기영동(E-Graph Gel Documentation System ATTO, Korea)하여 genomic DNA 밴드의 유무를 확인한 후, NanoPhotometer N60 Touch (Implen GmbH, Germany)으로 농도를 측정하였다.

### 차세대 염기서열 분석(NGS) 및 microsatellite 탐색

차세대 염기서열 분석(NGS)을 사용하여 벤자리의 microsatellite 마커 개발을 위해 서귀포 위미항 자연산 집단(2020년, n=122) 중 2개체를 사용하였다. 추출된 DNA는 정성적 및 정량적 품질을 검사한 후, 라이브러리 제작과 Illumina 시퀀스 분석(Hiseq X ten, USA)을 진행하였다. Short reads (WES data)의 전처리 과정은 PCR duplicate reads를 제거하고, cutadapt [28]로 adaptor 서열을 제거한 이후, quality trimming을 위해 SolexaQA (v.1.13) [5]의 DynamicTrim과 LengthSort 프로그램을 사용하여 염기서열 전처리 과정(pre-processing)을 진행하였다.

DynamicTrim의 phred score는  $\geq 20$ 으로 설정하여 short read의 양끝에 bad quality 염기를 1차로 제거하였으며[15], LengthSort의 short read length는  $\geq 25$  bp로 설정하고 DynamicTrim 과정에서 너무 많은 염기가 잘린 read를 최종 제거하였다[29]. 과정을 통과한 염기서열을 대상으로 SOAPdenovo2 (version 2.04) [26] 프로그램을 사용하였으며, paired-end 염기서열 정보를 각각 *de novo* assembled하였다. Assembled contig 서열에서 MISA (v1.0) [39] 프로그램을 사용하여 raw SSR을 탐색하였다. 분석 샘플 간 SSR의 비교 분석을 수행하기 위해, reference 기준으로 비교 샘플간의 SSR size matrix를 작성하였다. Reference 서열에서 찾은 SSR 영역을 기준으로 앞뒤 20 bp의 flanking sequence를 추출하고, flanking sequence를 reference와 분석 샘플의 reference genome에 *in silico* PCR을 수행하여 각 샘플 별 예상 SSR size를 계산하였다. Reference genome 서열에서 찾은 SSR을 대상으로 primer를 제작하였다. Primer의 제작은 Primer3 (v2.3.5) [40] 프로그램을 사용하여, primer 길이는 20-24 bp, 증폭 산물의 크기는 100-300 bp, GC의 함량은 50% 내외를 기준으로 하였으며, Tm 값은 48-62

℃로 설정하였다.

**Microsatellite 마커 선별**

NGS 방법으로 탐색된 microsatellite 영역의 증폭 여부를 확인하기 위해 서귀포 위미항 자연산 집단(2020년, n=122) 중 8개체를 대상으로 PCR 분석을 수행하였다. PCR 조성은 10X Ex Taq buffer (Mg2+ plus, 20 mM, 1 ml) 1 µl, dNTP Mixture (2.5 mM each, 800 µl) 0.8 µl, TaKaRa Ex Taq (5 U/µl, 250 U) 0.1 µl, forward primer와 reverse primer 각각 0.3 µl (10 mM), template DNA 0.5 µl를 넣고, 총 10 µl가 되도록 멸균된 증류수를 넣은 후, PCR (Veriti™96-Well Fast Thermal Cycler, Applied Biosystems™, US)을 진행하였다. PCR 조건은 95℃에서 11분간 DNA를 사전변성(preincubation) 후, 95℃에서 50초 변성(denaturation)하였다. 54℃에서 50초 primer 합성(annealing), 72℃에서 50초 DNA 합성(extension)하여, 총 35회 반복한 후, 최종 DNA 합성(full extension)을 72℃에서 7분간 실시하였다. PCR 과정이 끝난 후, 전기영동을 사용하여 1.8% agarose gel 상에서 증폭된 DNA 밴드 유무를 확인하였다.

PCR 과정을 통해 증폭여부 등을 확인한 후, 선별된 후미 microsatellite 마커의 forward 서열 정방향에 형광물질 (6-FAM, HEX 및 TAMRA)을 합성하여 위와 같은 방법으로 PCR을 다시 진행하였다(Table 4). PCR 증폭산물에 GeneScan 400HD ROX size standard (Applied Biosystems, USA)와 Hi-Di Forma-mide를 혼합하여 95℃에서 2분간 변성한 후, ABI PRISM 3730XL Genetic Analyzer (Applied Biosystems, USA)로 유전자 단편 크기(fragment size)를 확인하였다.

**Microsatellite 마커의 효율성 검증**

최종 선별된 microsatellite 마커로 서귀포 위미시험포 양식산 집단(2020년, n=102)과 서귀포 위미항 자연산 집단(2020년, n=122) 및 제주 한림 자연산 집단(2021년, n=50)을 대상으로 마커의 효율성을 검증하였다.

Microsatellite의 다형성정보지수(Polymorphic Information Content, PIC)를 확인하기 위해 Cervus ver. 3.0.7 [17]의 대립인자빈도 분석법을 적용하여 산출하였으며, 대립유전자 크기가 결정된 데이터 정보는 Micro-Checker [41] 프로그램을 사용하여 null allele에 의한 genotyping error를 검증 하였다. Arlequin version 3.1 software [10] 및 GENEPOP version 4.0 computer package [35] 프로그램을 사용하여 대립유전자 수(Number of Alleles, N<sub>A</sub>) 및 유전자좌위별 이형접합도(Heterozygosity)의 관찰치 이형접합률(Observed Heterozygosity, H<sub>O</sub>)과 Hardy-Weinberg Equilibrium (HWE)에서의 기대치 이형접합률(Expected Heterozygosity, H<sub>E</sub>)을 계산하였다.

Table 1. Summary of NGS sequencing results

Sample	Statistics of Sequencing raw data				Statistics of PCR duplicate removal data				Statistics of Final trimmed data				
	No. of reads	Total length (bp)	GC (%) <sup>*1</sup>	Q20 (%) <sup>*2</sup>	Genome cov. <sup>*3</sup>	No. of reads	Total length (bp)	Trimmed/ raw (%) <sup>*4</sup>	Genome cov. <sup>*3</sup>	Total length (bp)	Total length (bp)	Trimmed/ raw (%) <sup>*4</sup>	Genome cov. <sup>*3</sup>
Benjari-1	201,122,467	30,369,492,517	35.7	95.64	80.66	197,587,304	29,835,682,904	98.24%	79.24	177,388,005	23,379,291,440	76.98%	58.83
Benjari-2	201,122,467	30,369,492,517	35.45	95.19		197,587,304	29,835,682,904	98.24%		177,388,005	20,921,446,551	68.89%	
Total	402,244,934	60,738,985,034	35.58	95.42		395,174,608	59,671,365,808			354,776,010	44,300,737,991		

\*<sup>1</sup> GC (%): GC content.

\*<sup>2</sup> Q20 (%): Ratio of bases that have phred quality score of over 20.

\*<sup>3</sup> Genome cov. : The total read length of each sample divided by the expected genome size (753 Mb).

\*<sup>4</sup> Trimmed/Raw (%): (Total length of trimmed reads / Total length of raw reads)\*100

Table 2. Statistics of assembled contig

Sample	Hash length ( <i>k-mer</i> ) <sup>*1</sup>	Num. of contigs	Length (bp) of contigs				
			Total length	MIN	MAX	AVG	N50 <sup>*2</sup>
Benjari	69	1,320,995	705,613,658	200	66,252	534	758

\*<sup>1</sup> K-mer: assembly k-mer length.

\*<sup>2</sup> N50: Contig and scaffold length corresponding to 50% of the total number of total nucleotide sequences produced when accumulating in descending order for the generated contigs and scaffolds.

## 결과 및 고찰

### Microsatellite 탐색 및 다형성 확인

벤자리 microsatellite 마커 개발을 위해 서귀포 위미항 자연산 집단(2020년, n=122)의 2개체를 NGS (Illumina HiSeq X ten, USA) 기법으로 분석하였다. Pair-end library는 151 bp로 제작하여 402,244,934개의 read를 얻었으며, 결과적으로 생산된 총 염기서열은 약 61 Gb (60,738,985,034 bp)이다(Table 1). 염기서열 정확도의 99%를 나타내는 Q20에서 벤자리 2개체의 평균값은 95.42%로 높은 수치를 나타냈으며, GC의 비율 평균값은 35.61%로 확인되었다 [24, 38]. NGS 기법으로 분석된 염기서열들은 짧은 read를 제거하기 위해서 전처리 과정(pre-processing)을 진행하였다. 전처리 과정은 PCR이 중복되는 read를 제거한 후, Cutadapt [28]로 adaptor 서열을 제거하고, 이후 품질 정비를 위해 SolexaQA (v.1.13) [5] 패키지의 DynamicTrim과 LengthSort 프로그램을 사용하였다. DynamicTrim 과정은 phred score를  $\geq 20$ 으로 설정해 short read 양쪽 끝의 bad quality 서열을 제거하고 양질의 cleaned read로 정제하였으며, LengthSort 과정은 short read length  $\geq 25$  bp로 설정하여 [15] DynamicTrim 과정에서 서열이 짧은 read를 제거하는 과정을 수행하였다[29].

PCR이 중복되는 read를 1차로 제거한 염기서열은 약 60 Gb (59,671,365,808 bp)로 미가공 데이터(raw data)의 98.24%였으며, Cutadapt 및 SolexaQA package (Dynamic Trim, LengthSort)로 최종 전처리 과정을 거친 염기서열은 약 44 Gb (44,300,737,991 bp)로 미가공 데이터의 72.94%였다. 각 개체별로는 20,921,446,551-23,379,291,440 bp였다(Table 1).

Draft genome 분석은 전처리 과정을 통과해 평균 염기서열 44 Gb (44,300,737,991 bp)를 대상으로 SOAPdenovo2 (v.2.04) [26] 프로그램을 사용하여 벤자리 2개체의 염기서열 데이터를 각각 *de novo* assembly 하였다.

K-mer를 비교하여 assembly를 수행한 결과, 69-mer로 진행했을 때 가장 hash length가 길어 분석에 사용하였으며, 제작된 contig는 1,320,995개가 생성되었다(Table 2).

탐색된 SSR 구간을 각 개체별로 서열을 탐색한 결과, di-nucleotide 228,526 (24.0%)개, tri-nucleotide 35,795 (3.8%)개, tetra-nucleotide 21,936 (2.3%)개, penta-nucleotide 29,418

(3.1%)개, hexa-nucleotide 384,630 (40.4%)개, hepta-nucleotide 133,843 (14.1%)개, octa-nucleotide 70,644 (7.4%)개, nona-nucleotide 30,823 (3.2%)개, deca-nucleotide 16,711 (1.8%)개였다. 이 중 hexa-nucleotide의 비율이 가장 높았으며, deca-nucleotide가 가장 낮은 비율을 나타냈다(Table 3). 이는 NGS 방법을 사용하여 수산자원인 방어를 대상으로 연구한 결과와 비교해 보면 분포되는 양상은 비슷하지만 벤자리에서 탐색된 SSR 분포 개수와 프라이머 쌍의 비율이 더 높게 나타났다[8].

### Microsatellite 마커 제작

탐색된 SSR 구간에서 서로 중복되는 SSR motif가 없도록 하였으며, 반복 수 길이가 10 bp 이상, 증폭산물 크기 100-300 bp, primer 크기 20-24 bp, 적정 온도 (T<sub>m</sub>) 48-62°C, GC 함량 40-60%의 조건으로 microsatellite 마커 후보군 952,326개를 선별하였다(Table 3). 선별된 952,326개의 서열 중 반복 염기의 종류, primer의 크기, 증폭산물의 위치 등을 고려하여 151개의 microsatellite 마커를 1차 선별하였다. 1차로 선별된 151개의 후보서열 중 PCR 증폭 여부 및 증폭산물의 크기 등을 확인하여 2차 선별하였다. 2차로 선별된 34개의 microsatellite 마커에 형광물질(dye)을 합성한 후, 유전자형(genotyping) 분석을 통해 대립유전자 수, 크기 등을 고려하여 최종적으로 15개의 microsatellite 마커를 선별하였다(Table 4).

### Microsatellite 마커의 효율성 검증

개발된 15개의 microsatellite 마커의 효율성 검정을 위

Table 3. Statistics of SSR detection

Motif type	No. of primer designed SSR
P2 (di-nucleotide)	228,526
P3 (tri-nucleotide)	35,795
P4 (tetra-nucleotide)	21,936
P5 (penta-nucleotide)	29,418
P6 (hexa-nucleotide)	384,630
P7 (hepta-nucleotide)	133,843
P8 (octa-nucleotide)	70,644
P9 (nona-nucleotide)	30,823
P10 (deca-nucleotide)	16,711
Total	952,326

Table 4. Characteristics of 15 microsatellite loci developed in *P. trilineatum*

Locus	Motif	Primer sequence (5'→3')	AT(°C)	Dye	Observed Size ranges (bp)	Null
PT2-6	(TG)14	F-TTCCTTTGTGTGTTGGTGTGTG R-TGTGAAGCAGATTCCTCCTCAG	60°C	TAMRA	91-115	-
PT2-31	(GA)15	F-GCAGCAAGGGAGAGAAGTGA R-GACACTTCTCACGGGCTGAA	60°C	HEX	172-220	-
PT2-48	(AC)14	F-CACTGTCACCGCTAGAGCAA R-CCGGGTGTGTTTGTGAGAC	60°C	FAM	180-266	-
PT2-49	(AC)15	F-ACAGAGCCGCTAAATCTCCG R-GGCAATCTGTACCCACTCCC	60°C	FAM	184-200	-
PT2-58	(GT)14	F-CCGAACCAGAGGTGTCACAA R-ACTGACTGGGACAAAGAGCG	60°C	FAM	215-287	-
PT2-66	(GT)11	F-TCTCTCCCACACACTCCCAA R-ACTGAGAAACAACAGATGGGGG	60°C	TAMRA	236-296	-
PT3-23	(TGC)9	F-GCCGGGACTCACGTCTAAAG R-TCCTCAGCAGTCGCGACAAG	60°C	HEX	182-227	-
PT3-28	(GAG)10	F-TGAGGAGAGCGAAGTAGTGGA R-TGAATCATTGGCCAGGGACC	60°C	HEX	141-174	-
PT3-35	(TAC)8	F-AAGACTTCAGGCAGGTGCAC R-CGTGCAAGTGTTTTAGCGCA	60°C	HEX	138-210	-
PT3-36	(ATA)10	F-ATTCTGCTCATCCTGGGACG R-ACTGTGATTTCCGGTGCTCCT	60°C	HEX	143-182	-
PT3-44	(CTG)10	F-ACAGAGAGGGGTGTCACTCA R-CGAGTGGCCACAGAAGTTGA	60°C	FAM	183-246	-
PT3-57	(CAG)10	F-TCACTCCCTGCGAGGATACA R-CAGAAACCTGTGTGTGCTGC	60°C	FAM	218-245	-
PT4-4	(TATG)8	F-AAAGTGCAGCAAGGACAAGGT R-CGCCCGTGCAAAGAATCTTG	60°C	TAMRA	122-182	-
PT4-15	(TTTA)9	F-CATGTTGACTGTGGGGAGGA R-TCGAGGCCCTCTTTATTCGC	60°C	FAM	193-241	-
PT4-21	(CATC)9	F-AGCTCCACTCAAACCTGAACAGA R-GACCACCTGACCTCCTCTGA	60°C	HEX	224-300	-

해 벤자리 3집단(서귀포 위미시험포 양식산 집단(2020년, n=102), 서귀포 위미항 자연산 집단(2020년, n=122), 제주 한림 자연산 집단(2021년, n=50))을 대상으로 유전학적 특성 분석을 수행하였다.

Microsatellite 마커의 다형성 정보량을 나타내는 PIC 값이 0.5 이상일 경우 충분한 개체식별력을 갖는 마커로 보고하고 있으며[3], 본 연구에서 개발한 15개의 microsatellite 마커 PIC 값은 0.578-0.882 범위로 평균 0.762의 수치로 나타나 개발한 마커가 높은 개체식별력을 나타내는 것을 의미한다(Table 5). 또한, null 대립유전자 분석 결과에서도 15개의 microsatellite 마커 모두 null allele, scoring errors, large allele dropout는 존재하지 않아 벤자리 집단의 유전적 변이를 분석하는데 유용할 것으로 판단된다(Table 4).

벤자리 양식산과 자연산 집단 내의 산출된 대립유전자 수( $N_A$ )는 평균 12개이며, 각 집단의 관찰이형접합률( $H_o$ )

과 기대이형접합률( $H_e$ ) 값은 평균 0.750과 0.793으로 나타났다. 이는 Barinova *et al.* (2002)[2]가 보고한  $N_A$  평균 6.6개,  $H_o$ 과  $H_e$  평균 0.613-0.608 값으로 보고된 결과 보다 높게 나타났으며 또한, 해산 어류 평균인 0.79[6]와 유사한 결과로 새로 개발된 15개의 microsatellite 마커가 벤자리의 유전적 다양성을 분석하는데 유용한 마커로 판단된다 (Table 5).

본 연구는 NGS 방법을 이용하여 벤자리의 새로운 염기서열 정보를 확보하였으며, 확보된 염기서열에서 최종적으로 15개의 microsatellite 마커를 선별하였다. 최종 선별된 마커를 벤자리의 양식산 집단과 자연산 집단에 적용한 결과, 충분한 다형성 및 재현성을 나타내었으며, 유전적 다양성을 분석하는데 유용한 마커로 판단되었다. 향후, 새롭게 개발한 15개의 microsatellite 마커를 사용하여 벤자리의 유전적 다양성 및 집단구조 분석 등의 유전학적

Table 5. Genetic variabilities at 15 microsatellite markers in the parents and offspring population of *P. trilineatum*

Loci	Locus															Mean all loci
	PT2-6	PT2-31	PT2-48	PT2-49	PT2-58	PT2-66	PT3-23	PT3-28	PT3-35	PT3-36	PT3-44	PT3-57	PT4-4	PT4-15	PT4-21	
Farmed (2020, N=102)	N <sub>A</sub>	12	15	26	8	19	19	13	10	9	8	5	11	12	14	13
	H <sub>o</sub>	0.814	0.833	0.892	0.784	0.863	0.853	0.853	0.725	0.627	0.765	0.627	0.814	0.863	0.569	0.767
	H <sub>E</sub>	0.810	0.826	0.908	0.768	0.879	0.878	0.806	0.747	0.707	0.787	0.638	0.797	0.871	0.883	0.798
	PIC	0.783	0.801	0.896	0.737	0.863	0.866	0.777	0.720	0.658	0.751	0.570	0.765	0.853	0.867	0.769
	HW	0.605	0.335	0.012	0.140	0.634	0.039	0.578	0.381	0.021	0.383	0.308	0.001	0.565	0.000	0.311
Wild (2020, N=122)	N <sub>A</sub>	12	12	28	9	19	24	11	10	10	11	7	12	13	17	14
	H <sub>o</sub>	0.770	0.820	0.820	0.705	0.910	0.877	0.836	0.705	0.648	0.762	0.582	0.664	0.877	0.582	0.751
	H <sub>E</sub>	0.806	0.804	0.902	0.762	0.864	0.889	0.813	0.697	0.702	0.747	0.634	0.789	0.879	0.895	0.792
	PIC	0.776	0.775	0.891	0.722	0.847	0.877	0.786	0.665	0.647	0.708	0.562	0.759	0.863	0.882	0.761
	HW	0.070	0.449	0.033	0.065	0.381	0.869	0.447	0.636	0.594	0.589	0.147	0.000	0.557	0.000	0.367
Wild (2021, N=50)	N <sub>A</sub>	10	10	21	8	16	17	8	8	7	10	7	9	13	15	11
	H <sub>o</sub>	0.720	0.720	0.800	0.720	0.840	0.780	0.840	0.780	0.580	0.860	0.740	0.620	0.880	0.440	0.732
	H <sub>E</sub>	0.771	0.777	0.877	0.763	0.875	0.889	0.796	0.719	0.692	0.769	0.668	0.780	0.895	0.896	0.789
	PIC	0.734	0.742	0.859	0.721	0.853	0.870	0.760	0.687	0.633	0.727	0.603	0.742	0.875	0.877	0.754
	HW	0.103	0.141	0.026	0.532	0.266	0.043	0.913	0.300	0.027	0.011	0.908	0.003	0.680	0.000	0.267
Mean all loci	N <sub>A</sub>	11	12	25	8	18	20	11	9	9	10	6	11	13	15	12
	H <sub>o</sub>	0.768	0.791	0.837	0.736	0.871	0.837	0.843	0.737	0.618	0.796	0.650	0.699	0.873	0.530	0.750
	H <sub>E</sub>	0.796	0.802	0.895	0.764	0.873	0.885	0.805	0.721	0.700	0.768	0.647	0.789	0.882	0.891	0.793
	PIC	0.764	0.773	0.882	0.727	0.854	0.871	0.774	0.691	0.646	0.729	0.578	0.755	0.864	0.875	0.762
	HW	0.259	0.308	0.024	0.245	0.427	0.317	0.646	0.439	0.214	0.328	0.289	0.001	0.601	0.000	0.315

N<sub>A</sub>=number of alleles per locus; H<sub>o</sub>=observed heterozygosity; H<sub>E</sub>=expected heterozygosity; PIC=polyomorphic information content.\*Not in conformity with Hardy-Weinberg Equilibrium (*p*<0.003, Bonferroni-corrected value)

특성 분석을 통해 벤자리 인공종묘생산에 있어 체계적으로 어미를 관리하는데 활용될 것이며 더 나아가 유전학적 특성, 교배지침, 선발육종, 방류효과조사 등 종 보존을 위한 유전학적 기초자료로 제공하고자 한다.

### 감사의 글

이 논문은 2023년도 국립수산물과학원 수산과학연구소(2023018, 수산 유전자원의 탐색 및 활용)의 지원으로 수행된 연구입니다.

### The Conflict of Interest Statement

The authors declare that they have no conflicts of interest with the contents of this article.

### References

- Allendorf, F. W., Luikart, G. and Aitken, S. N. 2013. Conservation and the genetics of populations. Wiley-Blackwell, Oxford, U.K.
- Barinova, A. A., Kumagai, K., Nakajima, M. and Taniguchi, N. 2002. Identification and characterization of microsatellite DNA markers developed in threeline grunt *Parapristipoma trilineatum*. *Fish Genet Breed Sci.* **32**, 27-32.
- Botstein, D., White, R. L., Skolnick, M. and Davis, R. W. 1980. Construction of a genetic linkage map in man using restriction fragment length polymorphisms. *Am. J. Hum. Genet.* **32**, 314.
- Cho, Y., McCouch, S., Kuiper, M., Kang, M. R., Pot, J., Groenen, J. and Eun, M. 1998. Integrated map of AFLP, SSLP and RFLP markers using a recombinant inbred population of rice (*Oryza sativa L.*). *Theor. Appl. Genet.* **97**, 370-380.
- Cox, M. P., Peterson, D. A. and Biggs, P. J. 2010. SolexaQA: At-a-glance quality assessment of Illumina second-generation sequencing data. *BMC Bioinformatics* **11**, 485.
- DeWoody, J. A. and Avise, J. C. 2000. Microsatellite variation in marine, freshwater and anadromous fishes compared with other animals. *J. Fish Biol.* **56**, 461-473.
- Dong, C. M., Lee, H. M., Lee, M. N., Noh, E. S., Nam, B. H., Kim, Y. O. and Kim, E. M. 2022. Development of new microsatellite DNA markers using next-generation sequencing in pacific oyster *Crassostrea gigas*. *Kor. J. Malacol.* **38**, 129-138.
- Dong, C. M., Lee, M. N., Kim, E. M., Park, J. Y., Kim, G. D. and Noh, J. K. 2020. Development and genetic diversity analysis of microsatellite markers using next-generation sequencing in *Seriola quinqueradiata*. *J. Life Sci.* **30**, 291-297.
- Evans, B., Bartlett, J., Sweijd, N., Cook, P. and Elliott, N. G. 2004. Loss of genetic variation at microsatellite loci in hatchery produced abalone in Australia (*Haliotis rubra*) and South Africa (*Haliotis midae*). *Aquaculture* **233**, 109-127.
- Excoffier, L., Laval, G. and Schneider, S. 2005. Arlequin (ver.3.0): an integrated software package for population genetics data analysis. *Evol. Bioinform. Online* **1**, 47-50.
- Feldheim, K. A., Jabado, R. W., Chapman, D. D., Cardenosa, D. and Maddox, J. D. 2020. Microsatellite primer development in elasmobranchs using next generation sequencing of enriched libraries. *Mol. Biol. Rep.* **47**, 2669-2675.
- Frankham, R., Ballou, J. D. and Briscoe, D. A. 2004. A primer of conservation genetics. Cambridge University Press, Cambridge, U.K.
- Fukuda, Y., Okamura, A., Nishiyama, M., Kawakami, H., Kamaishi, T. and Yoshinaga, T. 2002. Granulomatosis of cultured three-line grunt *Parapristipoma trilineatum* caused by an intracellular bacterium. *Fish Pathol.* **37**, 119-124.
- Gardner, M. G., Fitch, A. J., Bertozzi, T. and Lowe, A. J. 2011. Rise of the machines—recommendations for ecologists when using next generation sequencing for microsatellite development. *Mol. Ecol. Resour.* **11**, 1093-1101.
- Garg, R., Patel, R. K., Tyagi, A. K. and Jain, M. 2011. De novo assembly of chickpea transcriptome using short reads for gene discovery and marker identification. *DNA Res.* **18**, 53-63.
- Goulet, J. 1995. FSTAT (version 1.2): a computer program to calculate F-statistics. *J. Hered.* **86**, 485-486.
- Kalinowski, S. T., Taper, M. L. and Marshall, T. C. 2007. Revising how the computer program CERVUS accommodates genotyping error increases success in paternity assignment. *Mol. Ecol.* **16**, 1099-1106.
- Kim, I. S., Lee, C. Y., Lee, C. L., Kim, Y. J., Kim, B. J. and Kim, J. H. 2011. Illustrated book of Korean fishes. Published by Kyo-Hak Publishing, Seoul, Korea. 326-327.
- Kim, J. E., Hwang, J. A., Kim, H. S. and Lee, J. H. 2020. Genetic variability comparison of wild and cultured far eastern catfish (*Silurus asotus*) of Korea using microsatellite marker. *Dev. Reprod.* **24**, 317-325.
- Kim, J. W., Kim, H. W., Huh, S. H. and Kwak, S. N. 2011. Seasonal variation and species composition of fish species in artificial reefs in the Shinyang-Ri coastal waters off Jeju island, Korea. *J. Kor. Soc. Fish. Tech.* **47**, 118-127.
- Kim, K. H., Hong, S. W., Moon, H. N. and Yeo, I. K. 2018. Physiological responses of the chicken grunt *Parapristipoma trilineatum* to high water temperature stress. *Kor. J. Fish. Aquat. Sci.* **51**, 714-719.
- Kim, K. S., Noh, C. H., Sade, A. and Bang, I. C. 2015. Effectiveness of microsatellite markers for parentage analysis of giant grouper (*Epinephelus lanceolatus*). *Kor. J. Ich.* **27**, 10-15.
- Kim, S. G., Morishima, K., Satoh, N., Fujioka, T., Saito, S. and Arai, K. 2007. Parentage assignment in hatchery population of brown sole *Pleuronectes herzensteini* by microsatellite DNA markers. *Fish. Sci.* **73**, 1087-1093.

24. Kim, T. H., Kim, Y. K., Son, J. H., Chun, J. B. and Yoon, Y. M. 2021. Utilization of whole genome re-sequencing for large-indel markers development in malting barley. *Kor. J. Breed. Sci.* **53**, 266-276.
25. Kumagai, K., Barinova, A. A., Nakajima, M. and Taniguchi, N. 2004. Genetic diversity between Japanese and Chinese threeline grunt (*Parapristipoma trilineatum*) examined by microsatellite DNA markers. *Mar. Biotechnol.* **6**, 221-228.
26. Luo, R., Liu, B., Xie, Y., Li, Z., Huang, W., Yuan, J., He, G., Chen, Y., Pan, Q., Liu, Y., Tang, J., Wu, G., Zhang, H., Shi, Y., Liu, Y., Yu, C., Wang, B., Lu, Y., Han, C., Cheung, D. W., Yiu, S. M., Peng, S., Xiaoqian, Z., Liu, G., Liao, X., Li, Y., Yang, H., Wang, J., Lam, T. W. and Wang, J. 2012. SOAPdenovo2: an empirically improved memory-efficient short-read *de novo* assembler. *GigaScience* **1**, 18.
27. Mardis, E. R. 2008. Next-generation DNA sequencing methods. *Annu. Rev. Genom. Hum. Genet.* **9**, 387-402.
28. Martin, M. 2011. Cutadapt: cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet J.* **17**, 10-12.
29. Ness, R. W., Siol, M. and Barrett, S. C. 2011. *De novo* sequence assembly and characterization of the floral transcriptome in cross-and self-fertilizing plants. *BMC Genomics* **12**, 298.
30. Norris, A. T., Bradley, D. G. and Cunningham, E. P. 2000. Parentage and relatedness determination in farmed Atlantic salmon (*Salmo salar*) using microsatellite markers. *Aquaculture* **182**, 73-83.
31. Paetkau, D. 1999. Microsatellites obtained using strand extension: An enrichment protocol. *Biotechniques* **26**, 690-697.
32. Park, C. J., Nam, W. S., Lee, M. S., Kang, J. Y. and Kim, K. K. 2014. Microsatellite multiplex PCR method for selective breeding studies in Pacific abalone (*Haliotis discus hannai*). *Kor. J. Malacol.* **30**, 383-390.
33. Park, I. S. and Gil, H. W. 2019. Karyotypic analysis of chicken grunt, *Parapristipoma trilineatum*, small yellow croaker, *Larimichthys polyactis*, and brown croaker, *Miichthys miiuy*. *Dev. Reprod.* **23**, 73-78.
34. Perez-Enriquez, R., Takagi, M. and Taniguchi, N. 1999. Genetic variability and pedigree tracing of a hatchery-reared stock of red sea bream (*Pagrus major*) used for stock enhancement, based on microsatellite DNA markers. *Aquaculture* **173**, 413-423.
35. Rousset, F. 2008. GENEPOP'007: a complete re-implementation of the GENEPOP software for Windows and Linux. *Mol. Ecol. Resour.* **8**, 103-106.
36. Rudnick, J. A. and Lacy, R. C. 2008. The impact of assumptions about founder relationships on the effectiveness of captive breeding strategies. *Conserv Gen.* **9**, 1439-1450.
37. Seong, K. T., Hwang, J. D., Han, I. S., Go, W. J., Suh, Y. S. and Lee, J. Y. 2010. Characteristic for long-term trends of temperature in the Koreans waters. *J. Kor. Soc. Mar. Environ.* **16**, 353-360.
38. Song, W. H. and Chung, S. M. 2019. Whole genome re-sequencing and development of SSR markers in oriental melon. *J. Plant Biotechnol.* **46**, 71-78.
39. Thiel, T., Michalek, W., Varshney, R. K. and Graner, A. 2003. Exploiting EST databases for the development and characterization of gene-derived SSR-markers in barley (*Hordeum vulgare* L.). *Theor. Appl. Genet.* **106**, 411-422.
40. Untergrasser, A., Cutcutache, I., Koressaar, T., Ye, J., Faircloth, B. C., Remm, M. and Rozen, S. G. 2012. Primer3 - new capabilities and interfaces. *Nucleic Acids Res.* **40**, e115.
41. Van Oosterhout, C., Hutchinson, W. F., Wills, D. P. M. and Shipley, P. 2004. MICRO-CHECKER: software for identifying and correcting genotyping errors in microsatellite data. *Mol. Ecol. Notes* **4**, 535-538.
42. Wang, Z., Gerstein, M. and Snyder, M. 2009. RNA-Seq: a revolutionary tool for transcriptomics. *Nat. Rev. Genet.* **10**, 57-63.
43. Watanabe, K. I. and Okazaki, T. 2000. Maturation and spawning season of threeline grunt *Parapristipoma trilineatum* around mugi-oshima Island on the pacific coast of tokushima prefecture, Japan. *Nippon Suisan Gakk.* **66**, 631-638.
44. Yu, H., Xie, W., Wang, J., Xing, Y., Xu, C., Li, X., Xiao, J. and Zhang, Q. 2011. Gains in QTL detection using an ultra-high density SNP map based on population sequencing relative to traditional RFLP/SSR markers. *PLoS One* **6**, e17595.
45. Zalapa, J. E., Cuevas, H., Zhu, H., Steffan, S., Senalik, D., Zeldin, E., McCown, B., Harbut, R. and Simon, P. 2012. Using next generation sequencing approaches to isolate simple sequence repeat (SSR) loci in the plant sciences. *Am. J. Bot.* **99**, 193-208.
46. Zhong, S., Su, Y., Mao, Y., Liu, M. and Wang, J. 2015. Development of 20 microsatellite markers for kuruma shrimp (*Marsupenaeus japonicus*) using illumina sequencing. *Conserv. Genet. Resour.* **7**, 717-719.



## 초록 : 차세대 염기서열 분석법을 사용한 벤자리(*Parapristipoma trilineatum*)의 microsatellite 마커의 개발 및 유전학적 특성 분석

동춘매<sup>1</sup> · 이미란<sup>1</sup> · 노재구<sup>2</sup> · 박진우<sup>2</sup> · 김영옥<sup>1</sup> · 김은미<sup>1\*</sup>

(<sup>1</sup>국립수산과학원 생명공학과, <sup>2</sup>국립수산과학원 제주연구소)

본 연구는 차세대 염기서열 분석법(NGS)을 사용하여 벤자리의 microsatellite 마커를 개발하고, 개발한 마커를 사용하여 벤자리 집단의 유전학적 특성을 분석하기 위해 수행되었다. 차세대 염기서열 분석 장비인 Illumina Hiseq X ten을 사용하여 총 402,244,934개의 read들을 얻어 assembly를 실행한 결과, 전체의 약 0.33%에 해당하는 1,320,995개의 read들이 assembly 되었으며, 이 read들의 총 길이는 705,613,658 bp로 확인되었다. 크기가 640 bp 이상이 되는 contig는 952,326개로 나타났으며, 이 중 microsatellite 영역을 포함하는 contig를 151개(0.016%)로 1차 선별하고, PCR 증폭 여부를 통해 microsatellite 후보 34개를 2차로 선별하였다. 그 중 벤자리 집단의 마커로서 유용한 15개의 microsatellite 마커를 최종 선택하였다. 새로 개발한 15개의 microsatellite 마커를 사용하여 벤자리 집단을 대상으로 분석한 결과, 관찰된 유효 대립유전자 수( $N_A$ )는 평균 12(6~25)로 나타났다. 평균 관측치 이형접합도( $H_O$ )와 평균 기대치 이형접합도( $H_E$ )는 각각 0.750 (0.530-0.873)와 0.793 (0.647-0.895)으로 나타났으며, 이는 해산어의 평균 값인 0.79와 유사한 수치를 나타내었다. 따라서 본 연구에서 개발한 15개의 microsatellite 마커는 벤자리 집단의 유전학적 특성 분석에 유용할 것으로 사료된다.