Population Genetic Analyses of a Burrowing Mayfly, *Ephoron shigae* (Ephemeroptera: Polymitarcyidae), from Korea and Japan

Ji Hyoun Kang[1P], Jeong Mi Hwang[1], Soon-Jik Kwon[2], Sekine, Kazuki[3], Sun-Jae Park[4], Min Jeong Baek[4], and Yeon Jae Bae[1,5c]

[1]Korean Entomological Institute, Korea University, Seoul, Republic of Korea
[2]Ecoplex, Seoul, Republic of Korea
[3]Geo-Environmental Science, Rissho University, Saitama 360-0194, Japan
[4]Division of Environmental Science and Ecological Engineering, College of Life Sciences, Korea University, Seoul, Republic of Korea
[5]Animal Resources Division, National Institute of Biological Resources, Incheon, Republic of Korea

National Institute of Biological Resources

ABSTRACT

Ephoron shigae is a burrowing mayfly renowned with its extremely short adult period (1-2 h) and population-specific asexual reproduction (i. e. parthenogenesis). It is also well known as a nuisance insect due to its frequently reported mass emergence patterns across Japan. Molecular analyses on this mayfly species have focused on its phylogeny and phylogeography with mainly single mitochondrial marker; however, the population genetic structure using highly polymorphic and co-dominant marker like microsatellite loci, have not been investigated so far. Here, we developed a total of ten novel microsatellite markers based on the 222,890 microsatellite loci isolated from *E. shigae* genomic DNA sequences, for the first time, for *Ephoron species* and analysed the genetic structure of 220 individuals from 11 *E. shigae* populations in Korea and Japan. This information of the population genetic structures and the level of genetic diversity with genome-wide microsatellites of *E. shigae* would provide insight into the present and past microevolutionary process of the populations of *E. shigae*, a potentially nuisance insect in South Korea.

INTRODUCTION

- *Ephoron shigae* (Takahashi), a burrowing polymitarcyid mayfly, which is widely distributed in Korea, Japan, northeast China, Far East Russia (Ishiwata 1996), has received considerable attention due to its intriguing life cycle characters such as highly synchronously emergence with extremely short adult stage (1 to 2 h) and asexual reproduction (Sekiné and Tojo 2019) (Figure 1).

- Its univoltine life cycle through embryonic diapause has been identified, and the simultaneous emergence and swarming in large number were also frequently reported in Japan (Sekiné et al. 2015).

- Large number of adults are easily attracted to the car and streetlight owing to its strong phototaxis and thus massive pile of dead body disturbed people by interfering traffic even causing car accident (Sekiné et al. 2013).

- Microsatellites, which are single sequence repeats, have been commonly used in population genetic studies on various insect taxa to identify genotypic diversity and population genetic differentiation (Kim et al. 2017; Chen and Dorn 2010). Currently, microsatellite markers have not been identified for *Ephoron* species.

- Here, we report the development of ten novel microsatellite markers of *E. shigae* using high throughput sequencing methods for the first time on this species and anlyzed population genetic structures of eleven populations in Korea and Japan.

Figure 1. Nymph and adult of *Ephoron shigae* in Korea (A and B) and Japan (C and D)

MATERIALS AND METHODS

Sample collection
A total of 220 adult individuals from the six populatios in South Korea and five in Japan was used.

Newly developed microsatellite markers using high-throughput sequencing
Sequencing: MGISEQ2000 platform (Illumina, USA); microsatellite marker identification: Krait v1.3.3. (Du et al. 2018); primer design: mfeprimer v3.2.2. (Wang et al. 2019).

Population genetic analyses
ARLEQUIN v.3.5.(Excoffier et al. 2010), GENEPOP v4.0 (Rousset 2008), STRUCTURE v. 2.3.1 (Evanno et al. 2005)

RESULTS AND DISCUSSION

Table 1. Summary statistics of generated read sequence data and assembly (A), perfect microsatellites (B), and microsatellite marker identification (from Di-mer to Hexa-mer) of *E. shigae* genome (C).

(A)

Sequencing Data Summary	
Platform	MGISEQ 2000
Library Type	Paired-end
Read Length	[illegible]
# of Reads	[illegible]
Total bp	[illegible]
Assembled Genomic Summary	
Scaffolds #	[illegible]
[illegible]	[illegible]
[illegible]	[illegible]
[illegible]	[illegible]
Longest (Shortest) scaffolds bp	[illegible]
GC level	[illegible]
Scaffolds bp	[illegible]
Scaffolds average length	[illegible]

(B)

Item	Number
Total number of perfect SSRs	222,890
Total length of perfect SSRs (bp)	[illegible]
The average length of SSRs (bp)	23.24
The percentage of sequence covered by SSRs (%)	[illegible]
Relative abundance (loci/Mb)	[illegible]
Relative density (bp/Mb)	[illegible]

(C)

Type	Counts	Length (bp)	Percent (%)	Average Length (bp)	Relative Abundance (loci/Mb)	Relative Density (bp/Mb)
Mono	71326	1020575	32	14.31	103.35	1478.82
Di	105302	2846392	47.24	27.03	152.58	4124.44
Tri	28861	512412	12.95	17.75	41.82	742.49
Tetra	13641	545392	6.12	39.98	19.77	790.28
Penta	3545	248125	1.59	69.99	5.14	359.54
Hexa	215	6666	0.1	31	0.31	9.66

Table 2. Characteristics of 10 polymorphic microsatellite loci developed and validated in of 30 *E. shigae* individuals in this study

Locus	Repeat motif	GenBank Accession No.	Primer sequence (5'-3')	Product size	HWE	Null	PIC	N_A	H_O	H_E	AR	F_{IS}
ES043	(GT)15	OM967150	F: AAGTGATTGCGTTCAAGC R: ATGACAACACGGTAGCG	167	0.000*	0.314	0.9495	25	0.733	0.968	24.578	0.245
ES102	(GCT)6	OM967151	F: ATTGGCCTTGCAGGAAAC R: CTCTGTAGGTCTGGCTA	197	0.318	0.000	0.6461	4	0.833	0.711	4.000	0.176
ES103	(GTT)5	OM967152	F: UTGATTCTATTTTACTGAAATG R: TGAGGTTATACACTGCGC	152	0.001*	0.183	0.0901	4	0.367	0.557	3.900	0.345
ES111	(CT)9	OM967153	F: GCCTATATAAGCTTAACACTG R: GACTGGATTGACAAGTGAG	185	0.024*	0.000	0.8472	10	0.867	0.877	9.800	0.012
ES147	(TC)13	OM967154	F: GATGACAAAACTTTTCAAGACG R: CGAATTAAACCGAGTAAGCA	215	0.264	0.000	0.8916	16	0.897	0.915	15.648	-0.021
ES130	(GA)12	OM967155	F: CAAATTCTCGGGAGGTC R: GGGCTTTGCTAGAAAATCC	127	0.001*	0.139	0.7827	10	0.567	0.881	5.790	0.296
ES317	(AAT)5	OM967156	F: CTACAGGGATGAAAAG R: CAGCATGACTTTATTTTACCT	127	0.030*	0.171	0.1849	2	0.100	0.210	2.000	0.377
ES322	(ACT)5	OM967157	F: CATAACAATCTCTGCTTGCG R: CTATATAGTATATTCCTCCTCC	218	0.000*	0.071	0.9193	22	0.800	0.945	21.163	0.156
ES337	(CA)21	OM967158	F: CAACACAACTGCTGCAAC R: AGGAAATCGTAGGGGA	226	0.790	0.000	0.9379	23	0.933	0.951	22.148	0.015
ES373	(TG)17	OM967159	F: GAGACAAAAGTGCTACGG R: TGCTTCAGGTCAACAAGG	183	0.000*	0.368	0.9194	23	0.630	0.960	23.000	0.348

F: forward primer, R: reverse primer, N_A: number of allele, HWE: exact p value of Hardy-Weinberg Equilibrium (*P<0.05), Null: frequency of null allele, PIC: polymorphism information content, H_O: observed heterozygosity, H_E: expected heterozygosity, AR: average allelic richness, F_{IS}: inbreeding coefficient

Table 3. Summary of genetic diversity statistics in 11 populations of *Ephoron shigae* in Korea and Japan

Populations	Region	N	N_A	AR	H_E	H_O	F_{IS}	HWE
KNY	Korea	20	13.1	8.656	0.816	0.531	0.356	High. sign.
KYP	Korea	20	12.3	8.600	0.822	0.646	0.219	High. sign.
KCC	Korea	20	12.0	8.523	0.828	0.670	0.195	High. sign.
KGS	Korea	20	12.8	8.726	0.808	0.611	0.248	High. sign.
KOC	Korea	20	11.7	7.912	0.800	0.622	0.227	High. sign.
KGE	Korea	20	12.9	8.853	0.831	0.631	0.246	High. sign.
JHY	Japan	20	7.0	5.780	0.713	0.481	0.331	High. sign.
JYG	Japan	20	10.1	6.715	0.740	0.484	0.353	High. sign.
JAG	Japan	20	8.4	5.577	0.704	0.394	0.450	High. sign.
JKG	Japan	20	7.6	5.369	0.652	0.456	0.307	High. sign.
JCG	Japan	20	3.5	2.398	0.576	0.858	-0.511	High. sign.

N: sample size, N_A: observed mean number of alleles across nine loci, *AR*: allelic richness, H_E: expected heterozygosity, H_O: observed heterozygosity, F_{IS}: observed inbreeding coefficient, H-W tests (*P*): *P*-value for multi-locus tests for Hardy-Weinberg equilibrium

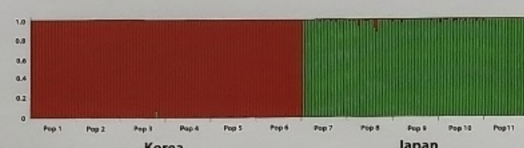

Figure 2. Population genetic structure of the 11 *E. shigae* populations determined using a Bayesian population assignment test with STRUCTURE based on nine microsatellite loci

- This study, for the first time, identified genome-wide microsatellites for *Ephoron* species. These novel microsatellite markers will be useful for investigating the genetic diversity and population genetic structure and monitoring emerging pattern of *Ephoron shigae*, a potentially nuisance insect in South Korea.

- High-throughput sequencing generated 169,793,152 reads with 628,835 scaffolds containing di-, tri-, and tetra-nucleotide repeat motifs (Table 1). The number of perfect microsatellite sequences, which were suitable for primer design, was 222,890. Ten polymorphic microsatellite markers were successfully amplified with stable and reproducible amplicon patterns and distinct peaks in capillary electrophoresis. The mean polymorphic information content (PIC) across loci was 0.7572, representing highly polymorphic markers (Table 2).

- The level of genetic diversity is much higher in the Korean populations compared to five populations in Japan. Particulary, a parthenogenetic population (JCG) from Japan showed the lowest allelic richness and the negative value of inbreeding coeffienct (Table 3).

- Distinct Korean and Japan genetic clusters with significant genetic differentiation were identified indicating low level of gene flow between two regions (Figure 2).

- Population genetic analyses on *E. shigae* using novel microsatellite markers will help to better understand the population dynamics of one of the burrowing polymitarcyid mayfly in Korean and Japan freshwater systems.

ACKNOWLEDGEMENTS

This work was supported by a grant from the National Institute of Biological Resources (NIBR), funded by the Ministry of Environment (MOE) of the Republic of Korea (NIBR202211101).

REFERENCES