Genetic diversity and population structure of razor clam *Sinonovacula constricta* in Ariake Bay, Japan, revealed using RAD-Seq SNP markers.

Ryo Orita[1,*], Yukio Nagano[2], Yoshio Kawamura[1], Kei Kimura[1], and Genta Kobayashi[1]

[1]*Faculty of Agriculture, Saga University, 1 Honjo-machi, Saga, 840-8502, Japan*

[2]*Analytical Research Center for Experimental Sciences, Saga University, 1 Honjo-machi, Saga, 840-8502, Japan*

*Corresponding author: Ryo Orita, E-mail: ss7427@cc.saga-u.ac.jp

Abstract

The razor clam *Sinonovacula constricta* is a commercially important bivalve in Japan. The current distribution of this species in Japan is limited to Ariake Bay, where the fishery stock is declining. It is necessary to understand the genetic population structure of this species in order to restore the fishery stock while preserving the genetic diversity of the clam. Here, we report for the first time the genetic population structure of *S. constricta* in Ariake Bay, Japan. Paired-end restriction site-associated DNA sequencing (RAD-Seq) analyzed samples of *S. constricta* collected from seven mudflats located along Ariake Bay. Two different genetic populations exist in Ariake Bay, one inhabiting wild habitats and the other inhabiting the transplanted area of artificial seedlings. Our results suggest that genetic differentiation occurred between these two populations (F_{st} value = 0.052), and a high level of genetic differentiation is maintained between the two groups. In the future, monitoring the interbreeding status of the two genetically distinct populations and the genetic differentiation within each population is important for conserving the genetic diversity of *S. constricta* in Japan.

Introduction

The technique of artificially producing seedlings and releasing them into fishing grounds is an effective method of aquaculture enhancement. This technique has been implemented with a variety of fishery species (e.g., finfish [1, 2], flounder [3], shellfish [4-7]). In artificial seed production, large numbers of seedlings are often produced by a few broodstocks and released into natural seawater, which reduces the genetic diversity of wild populations in the field. Therefore, when seedlings are produced and released into natural seawater, there is a need to understand the genetic structure within the wild population in order to preserve genetic diversity.

The razor clam *Sinonovacula constricta* (Lamarck, 1818) is an important economic species in East Asian coastal waters [8-10] and is widely farmed in China [10]. This species was once distributed from central to western Japan [11-12] but is now restricted to the mudflats around Ariake Bay on the west coast of Kyushu in western Japan. In Japan, razor clams are not farmed, and the catch depends on the amount of

41 resources in natural seawaters. According to a report by Ministry of the Environment,
42 Japan [13], the annual catch of *S. constricta* in Ariake Bay often recorded more than
43 1000 tons per year from the 1900s to the early the 1920s, but it decreased to less than 500
44 tons in most years from the latter half of the 1920s to the latter half of the 1980s, and no
45 catches have been made since 1994 because of the marked decline in the abundance of
46 this species. In order to restore the *S. constricta* resources, the Saga Prefectural Ariake
47 Fisheries Research and Development Center has developed seed production techniques
48 [14, 15], and since 1988, artificial seedlings of *S. constricta* have been released into the
49 transplanting areas established in natural seawaters [16] and have been conserved in
50 transplanting areas as populations to produce subsequent generations. However, the
51 genetic population structure of *S. constricta* within Ariake Bay has not been examined.
52 In recent years, technological innovations have made the genome-wide analysis of
53 non-model organisms more widely available, and these techniques are useful for
54 examining the genetic population structure of non-model aquatic organisms [17-20]. In
55 the present study, to clarify the genetic population structure of *S. constricta* in Ariake Bay,

56 we conducted field sampling surveys at 11 mudflats located along the bay in 2019 (Fig. 1)

57 and used restriction-site-associated DNA sequencing (RAD-Seq) to analyze population

58 genetics and phylogenetics. We propose guidelines concerning the optimal seed

59 production and seed release methods for the conservation of the genetic diversity of *S.*

60 *constricta* in Japan.

61

62 **Results**

63 **Sampling data.** We collected 28 razor clams from seven of the 11 mudflats and found

64 no razor clams on the remaining four mudflats (Table 1, Fig. 1). The mean CPUE of the

65 seven mudflats where individuals could be collected was 3.2 (± 3.0 SD) ind. person^{-1}

66 hour^{-1}, and the range of CPUE at each site was 0.3 to 8.5 ind. person^{-1} hour^{-1}. The shell

67 length of the collected individuals was 39.7 to 94.6 mm, implying that they contained

68 specimens from 1 to 5 years old. The mean shell length of individuals collected from the

69 transplanted area was 86.0 (± 4.5 SD) mm for the 3-year-old seedlings produced in 2016

70 (i.e., sample ID: 2016.1-2016.7) and 92.9 (± 1.1 SD) mm for the 5-year-old seedlings

71 produced in 2014 (i.e., sample ID: 2014.1-2014.5).

72

73 **RAD-Seq data.** We used 30 individuals for RAD-Seq analysis, including 16

74 individuals collected from wild habitats, 12 individuals collected from the transplanted

75 area of artificial seedlings, and 2 individuals stored at the Saga Prefectural Ariake

76 Fisheries Research and Development Center (Table 1). In all, 401,109,332 raw reads

77 were generated from the HiSeq X run, and 249,040,001 reads were retained after

78 discarding reads with low quality and adapters. The final dataset included 137,246 loci,

79 with an average of 679.38 bp (SE ± 0.18) of genotyped sites per locus.

80

81 **Population genetics structure.** The genetic relationships among sampling sites were

82 clearly divided into individuals collected from wild habitat and those from the

83 transplanted area of artificial seedlings along the axis of PC1 (Fig. 2a). The PC2 axis

84 represents the differences within each habitat type. In the PCA plot, the coordinates of the

85 transplanted samples were more widely scattered than those of the wild samples were,
86 showing more diverse genetic relationships among transplanted samples. When
87 evaluated by shell length (Fig. 2b), genetic diversity was higher in the 75–90 mm size
88 group than that in the younger age group with smaller shells. The results of admixture
89 analysis also showed a significant genetic difference between the individuals collected
90 from the wild versus those collected from the transplanted areas of artificial seedlings
91 (Fig. 3). The F_{st} value between populations consisting of individuals collected from the
92 wild and the transplanted areas was 0.052. Owing to the calculation of the pairwise F_{st}
93 values for each individual within the two populations, there was no significant difference
94 in the mean values between these populations (Fig. 4). The pairwise F_{st} values within the
95 population were more varied in the transplanted area (from 0.44 to 0.64) than those in the
96 wild habitat (from 0.53 to 0.56). There were no significant differences in shell
97 morphology between these two genetically different populations (Supplementary Fig.
98 S1).
99

100 **Phylogenetic tree analysis.** ML tree and SVDquartets analysis of RAD-Seq data
101 showed that the phylogenies were highly consistent with each other (Fig. 5a and 5b). The
102 obtained phylogenetic trees were clearly divided into individuals collected from the wild
103 habitat and those collected from the transplanted area of artificial seedlings, with the
104 exception of one sample (i.e., sample ID: 2016.7).
105

106 **Discussion**

107 **Genetic population structure of *S. constricta* in the Ariake Bay, Japan.** All our
108 analyses of the RAD-Seq data showed that individuals collected from wild habitats were
109 genetically different from those collected from sites where artificial seedlings were
110 transplanted. Only one sample collected from the area transplanted in 2016 (sample ID:
111 2016.7) showed different results from other seedlings collected in the same area, and
112 belonged to a population consisting of individuals collected from the wild habitat (Fig. 3,
113 Fig. 5a and 5b). Since this individual had a distinctly larger shell length than other
114 3-year-old seedlings in the same area, and its size was the same as that of 5-year-old

115 seedlings (Table 1), it was likely a wild individual inhabiting the transplanted area prior
116 to the 2016 transplantation. Therefore, we consider that there were two genetically
117 distinct populations in Ariake Bay in 2019.

118 The reason for the differences between these two populations is the different
119 parental origins of the seedlings. The parental origins can be divided into the Nanaura
120 mudflat and the Kashima mudflat for a number of reasons. Artificial seedlings of *S.*
121 *constricta* have been released by the Saga Prefectural Ariake Fisheries Research and
122 Development Center into transplanting areas established in several mudflats along the
123 coast of Ariake Bay. The seedlings collected from the transplanted area in this study
124 (Takori_b and Takori_c) originated from parents at the Nanaura mudflat (hereinafter
125 referred to as Nanaura lineage population), since the parents were collected from the
126 Nanaura mudflat before 2010 and released into the transplanting areas of the Takori
127 mudflat. In contrast, individuals collected from the wild habitat of Kashima, Nanaura,
128 and Takori mudflats are likely to have originated from parents of Kashima mudflat
129 (hereinafter referred to as the Kashima lineage population). In 2017 and 2018, the

130 seedlings produced by parents collected from Kashima mudflat were released into new
131 transplanting areas at Takori and Nanaura mudflats due to declining populations in wild
132 and transplanted areas in these mudflats. Therefore, the 1- to 2-year old individuals
133 collected in wild habitats in 2019 were likely the offspring of the transplanted Kashima
134 seedlings in 2017 and 2018.

135 Artificial transplantation of aquatic organisms has been reported to alter the
136 geographic genetic structure of wild populations [21-24], including that of *S. constricta*
137 (in this case, along the Chinese coast) [25]. Anthropogenic transplantation likely strongly
138 influenced the distribution of the genetic population structure of *S. constricta* in Japan.
139 Comparison of the genetic distribution of species with a life history similar to that of *S.*
140 *constricta* but those that are not exploited for fisheries can enhance our understanding of
141 anthropogenic impacts on the genetic distribution of *S. constricta* in the Ariake Bay.

142

143 **Genetic differentiation among and within populations.** For artificial seed
144 production and release of seedlings, it is necessary to understand the distribution of

145 populations with different genetic structures in natural seawater. Therefore, it is
146 necessary to select broodstocks and release areas according to the degree of genetic
147 differentiation among and within populations. F_{st} values are widely used as a measure of
148 the degree of genetic differentiation among and within populations [26-28]. According to
149 Niu et al. [25], the F_{st} value among populations of *S. constricta* at 10 sites along the coast
150 of mainland China is 0.044. The F_{st} value between the two populations in the present
151 study was 0.052, suggesting that a similar level of genetic differentiation occurred
152 between the two populations in Ariake Bay, Japan.

153 For these two populations, the genetic differentiation within each population was
154 maintained at a high level, and its variation was greater within the population consisting
155 of individuals collected from the transplanted area (Fig. 4). There are three possible
156 reasons why the variation in the degree of genetic differentiation within populations
157 differed between the two populations. The first is that the degree of variation in genetic
158 differentiation in these populations originally may have been different. Second, the
159 methods of seedling production may lead to greater variation of genetic differentiation

160 within populations in the transplanted area. In seedling production, broodstocks were
161 randomly selected with closely related individuals collected from the same transplanted
162 area or distantly related individuals collected from different transplanted areas. After
163 stimulating the mature individuals to spawn, 20 to 30 individuals were placed in the same
164 tank to produce fertilized eggs. Thus, the degree of genetic differentiation may have been
165 more varied in the transplanted populations than in the wild habitat population. Third, the
166 sample size may have been too small to accurately estimate the population. However, in
167 the current situation in which the number of wild populations is small, it is difficult to
168 verify the validity of a larger sample size.
169
170 **Guidelines concerning the best seed production and seed release methods.** The
171 current abundance of *S. constricta* is estimated to be very low based on the CPUE results
172 of this study (Table 1). These results suggest that artificial seed production and release of
173 seedlings could be an effective strategy for the recovery of *S. constricta* abundance in
174 Ariake Bay, Japan. The present study suggested genetic differentiation between the two

175 populations of *S. constricta* in Ariake Bay, and a high level of genetic differentiation
176 within each population was maintained. Therefore, maintaining the two populations
177 without interbreeding seems to be an appropriate conservation strategy. As of 2019, the
178 Kashima lineage population is widely distributed in wild habitats, whereas the Nanaura
179 lineage population is limited to the Takori mudflats. In order to avoid hybridization with
180 the Kashima lineage population, the release of seedlings of the Nanaura lineage
181 population should continue to be restricted to the mudflats within the current cove, with a
182 transplanting area.

183 To produce seedlings to slow down the decline in genetic diversity within a
184 population, mathematical models have evaluated five strategies: I) avoidance of line
185 breeding, II) selective use of wild-born individuals, III) line breeding, IV) avoidance of
186 artificial breeding of males, and V) selective use of wild-born males [29]. Among them,
187 strategy II has been considered the most effective [29, 30]. The best way to produce
188 seedlings of the Kashima lineage population would be to follow strategy II; all parental
189 individuals would be randomly collected from wild habitats and previously produced

190 seedlings would not be used as parents. In contrast, the Nanaura lineage population has
191 already been maintained by a similar approach to strategy I; all parental individuals will
192 be randomly collected from transplanted areas without excluding released seedlings. This
193 approach gradually reduces the genetic diversity within the population. Indeed, lower
194 pairwise F_{st} values were detected in the transplanted population (Fig. 4). Therefore, in the
195 Nanaura lineage population, countermeasures should be taken to increase the number of
196 parents used for seed production or to avoid using the produced seedlings as parents of
197 the next generation.

198 In conclusion, we detected two genetically distinct populations of *S. constricta* in
199 the Ariake Bay, Japan. Additionally, our results suggests that genetic differentiation
200 within each population is maintained at a high level. In the future, monitoring the
201 interbreeding status of the two genetically distinct populations and the genetic
202 differentiation within each population is important to enhance the stock of *S. constricta* in
203 Japan by producing and releasing seedlings, while preserving genetic diversity. As there
204 was no significant difference in the morphology of the shells of these genetically distinct

205 populations (Supplementary Fig. S1), DNA markers may be effective for discriminating

206 between the two populations.

207

208 **Methods**

209 **Sampling and genomic DNA extraction.** Sampling of the razor clam *S. constricta*

210 was conducted at 11 mudflats located along Ariake Bay (Fig. 1). Information on the

211 sampling day and habitat types at each site is shown in Table 1. Habitats are divided into

212 two types: transplanted area of artificial seedlings, and wild habitat. As mentioned above,

213 artificial seedlings of *S. constricta* have been released into the transplanting areas

214 established in natural seawaters since 1988. Therefore, artificially produced individuals

215 live within the transplanted area, and the wild habitat would be populated by natural

216 individuals and offspring of seedlings from the transplanted areas. The ages of

217 individuals in the transplanted areas were estimated by calculating backwards from the

218 year of release. At each site, razor clams were explored by hand and their abundance was

219 assessed using the catch per unit effort index (CPUE), calculated as the number of clam

220 individuals collected divided by the number of explorers and sampling time (i.e., ind.

221 person^{-1} hour^{-1}). The samples were stored at -80°C until genomic DNA extraction. In

222 addition to the collected individuals, the oldest seedlings (produced in 2005, $n = 2$) stored

223 at the Saga Prefectural Ariake Fisheries Research and Development Center, were

224 included in the RAD-Seq analysis for comparison with seedlings produced in the past.

225 Genomic DNA was extracted from finely chopped foot tissue of the specimens using

226 DNAs-ici!-F (RIZO, Tsukuba, Japan), and the concentration of the DNA was assessed

227 using a Qubit 3.0 Fluorometer (Thermo Fisher Scientific).

228

229 **RAD-Seq analysis.** The preparation for *the Eco*RI RAD-Seq library has been

230 described previously [28], based on the established method [31]. The library was

231 sequenced with 150 bp paired-end reads in one lane of an Illumina HiSeq X (Illumina,

232 San Diego, CA, USA). The RAD-Seq data were analyzed using the Stacks program

233 (version 2.52) [32]. The data were quality filtered using the process_radtags program in

234 the Stacks package (-r -c -q, - adapter_1, adapter_2, and adapter_mm2). Filtered

235 sequences were registered in the DRA (DRA accession number: DRA010922). The data
236 were aligned with the reference genome of *S. constricta* [33] (NCBI/DDBJ BioProject
237 number: PRJNA508451) using BWA-MEM (version 0.7.17) [34]. Aligned data were
238 analyzed using the ref_map.pl script of the Stacks package.
239
240 **Population genetics analysis.** Principal component analysis (PCA) was conducted
241 based on a variant call format (VCF) file using the SNPrelate program (v.1.18.1) [35].
242 The VCF file was created by the populations program of the Stacks package
243 (--rm-pcr-duplicates, -R 0.7, --write-single-snp, --min-maf 0.05) [36]. The admixture
244 analysis was conducted based on the PLINK file using ADMIXTURE (v. 1.3.0) [37]. The
245 PLINK file was created with the same Stacks populations program that created the VCF
246 file [38]. The genetic diversity of razor clams among and within populations was assessed
247 using AMOVA F_{st} values obtained from 'populations.fst_files', which were calculated
248 by the Stacks populations program. To visualize the distribution of all data points,

249 pairwise F_{st} values between individuals in each population are shown in boxplot with

250 beeswarm using R (package: 'beeswarm') [39].

251

252 **Phylogenetic tree analysis.** Phylogenetic tree analysis was performed using two

253 different approaches: the maximum likelihood (ML) method and SVDquartets analysis.

254 Multiple alignments within the cluster were obtained using the Stacks package

255 populations program mentioned earlier. The phylogenetic tree based on the ML method

256 was constructed using the RAxML program (version 8.2.12) [40] (-f = a, -x = 12,345, -p =

257 12,345, -N (bootstrap value) = 1,000, and -m = GTRGAMMA). The SVDquartets

258 analysis based on the coalescent model was computed as implemented in PAUP ver. 4.0a

259 167 using multiple alignments [41, 42]. The number of bootstrap analyses replicates was

260 1,000. In each analysis, the oldest samples (i.e., sample ID, 2005.1; 2005.2), were used as

261 the root.

262

263 **References**

1. Fushimi, H. Production of juvenile marine finfish for stock enhancement in Japan. *Aquaculture* **200,** 33–53. https://doi.org/10.1016/S0044-8486(01)00693-7 (2001).

2. Masuda, R. & Tsukamoto, K. Stock enhancement in Japan: review and perspective. *Bull. Mar. Sci.* **62,** 337–358. https://www.ingentaconnect.com/content/umrsmas/bullmar/1998/00000062/00000002/art00005 (1998).

3. Sekino, M., Saitoh, K., Yamada, T., Hara, M. & Yamashita, Y. Genetic tagging of released Japanese flounder (*Paralichthys olivaceus*) based on polymorphic DNA markers. *Aquaculture* **244**, 49–61. https://doi.org/10.1016/j.aquaculture.2004.11.006 (2005).

4. Arnold, W. S. Bivalve enhancement and restoration strategies in Florida, USA. *Hydrobiologia* **465,** 7–19. https://doi.org/10.1023/A:1014596909319 (2001).

5. Castell, L. L., Naviti, W. & Nguyen, F. Detectability of cryptic juvenile *Trochus niloticus* Linnaeus in stock enhancement experiments. *Aquaculture* **144,** 91–101. https://doi.org/10.1016/S0044-8486(96)01320-8 (1996).

6. McCormick, T. B., Herbinson, K., Mill, T. S. & Altick, J. A review of abalone seeding, possible significance and a new seeding device. *Bull. Mar. Sci.* **55,** 680–693. https://www.ingentaconnect.com/content/umrsmas/bullmar/1994/00000055/f0020002/art00035 (1994).

7. Zohar, Y. *et al.* The Chesapeake Bay blue crab (*Callinectes sapidus*): A multidisciplinary approach to responsible stock replenishment. *Rev. Fish. Sci.* **16,** 24–34. https://doi.org/10.1080/10641260701681623 (2008).

8. Funge-Smith, S., Briggs, M. & Miao, W. Regional Overview of Fisheries and Aquaculture in Asia and the Pacific 2012. *RAP Publication (FAO)* (2012).

9. Mao, Y., Lin, F., Fang, J., Fang, J., Li, J. & Du, M. Chapter 4. Bivalve production in China. In *Goods and Services of Marine Bivalves.* (eds Smaal, A. C. *et al.*) 51–72 (Springer, Cham, 2019).

10. Ran, Z. *et al.* Fatty acid and sterol changes in razor clam *Sinonovacula constricta* (Lamarck 1818) reared at different salinities. *Aquaculture* **473,** 493–500. https://doi.org/10.1016/j.aquaculture.2017.03.017 (2017).

11. Suzuki, T., Inoue, K. & Ozawa, T. Environmental degradation in Ise and Mikawa Bays after 1960's as viewed from intertidal molluscan community. *Boll. Nagoya Univ. Museum* **22,** 31–64. http://doi.org/10.18999/bulnum.022.04 (2006).

12. Nakamura, T., Nishida, I., Takada, H., Okuno, M., Minami, M., & Oda, H. Marine reservoir effect deduced from 14C dates on marine shells and terrestrial remains at archeological sites in Japan. *Nucl. Instrum. Methods Phys. Res., B* **259**, 453–459. https://doi.org/10.1016/j.nimb.2007.01.186 (2007).

13. Ministry of the Environment, Japan. *Ariake Sea and Yatsushiro Sea Comprehensive Survey Evaluation Committee Report.* https://www.env.go.jp/council/20ari-yatsu/rep061221/all.pdf (2006).

14. Ito, S., Eguchi, T. & Kawahara, I. Rearing experiment on planktonic larvae of the JackKnife clam, *Sinonovacula constricta*. *Boll. Saga Prefect. Ariake Fish. Res. Dev. Cent.* **20,** 49–53. https://agriknowledge.affrc.go.jp/RN/2010813185.pdf (2001).

15. Ohkuma, H., Ymagachi, T., Kawahara, I., Eguchi, T. & Ito, S. A study on the development of techniques for mass production of seeds of jackknife clam,

Sinonovacula constricta. *Boll. Saga Prefect. Ariake Fish. Res. Dev. Cent.* **22,** 47–54. https://agriknowledge.affrc.go.jp/RN/2030813210.pdf (2004).

16. Tsukuda, M. *et al.* Variation in the Distribution of the JackKnife Clam, *Sinonovacula constricta*, on the Muddy Tidal flat of Ariake Sound off Saga Prefecture. *Boll. Saga Prefect. Ariake Fish. Res. Dev. Cent.* **28,** 47–49. https://agriknowledge.affrc.go.jp/RN/2010925685.pdf (2017).

17. Holman, L. E., Garcia de la serrana, D., Onoufriou, A., Hillestad, B. & Johnston, I. A. A workflow used to design low density SNP panels for parentage assignment and traceability in aquaculture species and its validation in Atlantic salmon. *Aquaculture* **476,** 59–64. https://doi.org/10.1016/j.aquaculture.2017.04.001 (2017).

18. Li, Y. H. & Wang, H. P. Advances of genotyping-by-sequencing in fisheries and aquaculture. *Rev. Fish Biol. Fish.* **27,** 535–559. https://doi.org/10.1007/s11160-017-9473-2 (2017).

19. Robledo, D., Palaiokostas, C., Bargelloni, L., Martínez, P. & Houston, R. Applications of genotyping by sequencing in aquaculture breeding and genetics. *Rev.*

Aquac **10,** 670–682. https://doi.org/10.1111/raq.12193 (2018).

20. You, X., Shan, X. & Shi, Q. Research advances in the genomics and applications for molecular breeding of aquaculture animals. *Aquaculture* **526,** 735357. https://doi.org/10.1016/j.aquaculture.2020.735357 (2020).

21. Arthington, A. H. Ecological and genetic impacts of introduced and translocated freshwater fishes in Australia. *Can. J. Fish. Aquat. Sci.* **48,** 33–43. https://doi.org/10.1139/f91-302 (1991).

22. Habtemariam, B. T., Arias, A., García-Vázquez, E. & Borrell, Y. J. Impacts of supplementation aquaculture on the genetic diversity of wild *Ruditapes decussatus* from northern Spain. *Aquacult. Environ. Interact.* **6,** 241–254. https://doi.org/10.3354/aei00128 (2015).

23. Williams, S. L. & Orth, R. J. Genetic diversity and structure of natural and transplanted eelgrass populations in the Chesapeake and Chincoteague Bays. *Estuaries* **21,** 118–128. https://doi.org/10.2307/1352551 (1998).

24. Yamakawa, A. Y. & Imai, H. PCR-RFLP typing reveals a new invasion of Taiwanese

Meretrix (Bivalvia: Veneridae) to Japan. *Aquat. Invasions* **8,** 407–415. http://doi.org/10.3391/ai.2013.8.4.04 (2013).

25. Niu, D. H., Feng, B. B., Liu, D. B., Zhong, Y. M., Shen, H. D. & Li, J. L. Significant genetic differentiation among ten populations of the razor clam *Sinonovacula constricta* along the coast of china revealed by a microsatellite analysis. *Zool. Stud.* **51,** 406–414. http://zoolstud.sinica.edu.tw/Journals/51.3/406.pdf. (2012).

26. DeFaveri, J., Shikano, T., Ghani, N. I. A. & Merilä, J. Contrasting population structures in two sympatric fishes in the Baltic Sea basin. *Mar. Biol.* **159,** 1659–1672. https://doi.org/10.1007/s00227-012-1951-4 (2012).

27. Jeffery, N. W. *et al.* RAD sequencing reveals genomewide divergence between independent invasions of the European green crab (*Carcinus maenas*) in the Northwest Atlantic. *Ecol. Evol.* **7,** 2513–2524. https://doi.org/10.1002/ece3.2872 (2017).

28. Kato, D. *et al.* Evaluation of the population structure and phylogeography of the Japanese Genji firefly, *Luciola cruciata*, at the nuclear DNA level using RAD-Seq

354 analysis. *Sci. Rep.* **10,** 1533. https://doi.org/10.1038/s41598-020-58324-9 (2020).

355 29. Yokota, M., Harada, Y. & Iizuka, M. Genetic drift in a hatchery and the maintenance

356 of genetic diversity in hatchery-wild systems. *Fish Sci.* **69,** 101–109.

357 https://doi.org/10.1046/j.1444-2906.2003.00593.x (2003).

358 30. Harada, Y., Yokota, M., & Iizuka, M. Genetic risk of domestication in artificial fish

359 stocking and its possible reduction. *Popul. Ecol.* **40**, 311–324.

360 https://doi.org/10.1007/BF02763463 (1998).

361 31. Etter, P. D., Bassham, S., Hohenlohe, P. A., Johnson, E. A. & Cresko, W. A. SNP

362 discovery and genotyping for evolutionary genetics using RAD sequencing, in:

363 Molecular methods for evolutionary genetics. *Methods Mol Biol. Humana Press* **772,**

364 157–178. https://doi.org/10.1007/978-1-61779-228-1_9 (2011).

365 32. Rochette, N.C. & Catchen, J.M. Deriving genotypes from RAD-seq short-read data

366 using Stacks. *Nat. Protoc.* **12,** 2640–2659. https://doi.org/10.1038/nprot.2017.123

367 (2017).

368 33. Ran, Z. *et al.* Chromosome-level genome assembly of the razor clam *Sinonovacula*

369 *constricta* (Lamarck, 1818). *Mol. Ecol. Resour.* **19,** 1647–1658.
370 https://doi.org/10.1111/1755-0998.13086 (2019).
371 34. Li, H. Aligning sequence reads, clone sequences and assembly contigs with
372 BWA-MEM. arXiv preprint arXiv:1303.3997. https://arxiv.org/abs/1303.3997
373 (2013).
374 35. Zheng, X., Levine, D., Shen, J., Gogarten, S. M., Laurie, C. & Weir, B. S. A
375 high-performance computing toolset for relatedness and principal component
376 analysis of SNP data. *Bioinformatics* **28,** 3326–3328.
377 https://doi.org/10.1093/bioinformatics/bts606 (2012).
378 36. Danecek, P. *et al.* The variant call format and VCFtools. *Bioinformatics* **27,** 2156–
379 2158. http://doi.org/10.1093/bioinformatics/btr330 (2011).
380 37. Alexander, D. H., Novembre, J. & Lange, K. Fast model-based estimation of ancestry
381 in unrelated individuals. *Genome Res.* **19,** 1655–1664.
382 https://pubmed.ncbi.nlm.nih.gov/19648217/ (2009).
383 38. Purcell, S. *et al.* PLINK: a tool set for whole-genome association and

population-based linkage analyses. *Am. J. Hum. Genet.* **81,** 559–575. https://doi.org/10.1086/519795 (2007).

39. Eklund, A. R package beeswarm: the bee swarm plot, an alternative to stripchart. https://cran.r-project.org/web/packages/beeswarm/beeswarm.pdf (2016).

40. Stamatakis, A. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* **30,** 1312–1313. https://doi.org/10.1093/bioinformatics/btu033 (2014).

41. Chifman, J. & Kubatko, L. Quartet inference from SNP data under the coalescent model. *Bioinformatics* **30,** 3317–3324. https://doi.org/10.1093/bioinformatics/btu530 (2014).

42. Swofford, D. L. PAUP*. Phylogenetic analysis using parsimony (*and other methods). Version 4. (Sinauer Associates, 2002).

Acknowledgments

The authors are grateful to the Saga Prefectural Ariake Fisheries Research and Development Center for supporting the field survey and providing samples and information on seedling production. This work was supported by the "Projects for sophistication of production and utilization technology supporting local agriculture and marine industry" from Saga University.

Author contributions

R.O. designed and wrote the original manuscript draft. R.O. and Y.N. performed the RAD-Seq analysis and bioinformatic analysis. All authors reviewed drafts of the manuscript and approved its final version.

Competing interests

The authors declare no competing interests.

413 Table 1. Sampling information for *Sinonovacula constricta* in the Ariake Bay, Japan

Sampling site	Sampling day	Habitat type	CPUE* (ind. person^{-1} h^{-1})	Sample ID	Shell length (mm)
	Nov. 11, 2019	Wild	0.3	Kashima_a.1	39.7
Kashima_b	Oct. 11, 2019	Wild	1.3	Kashima_b.1	46.6
				Kashima_b.2	45.2
				Kashima_b.3	54.7
				Kashima_b.4	54.0
Kashima_c	Nov. 19, 2019	Wild	4.0	Kashima_c.1	44.7
				Kashima_c.2	51.4
				Kashima_c.3	48.1
				Kashima_c.4	67.0
	Oct. 23, 2019	Wild	N/A†	Kashima_c.5	70.2
Kashima_d	Nov. 11, 2019	Wild	0.0	–	–
Nanaura_a	Nov. 13, 2019	Wild	0.7	Nanaura_a.1	47.0
				Nanaura_a.2	53.9
Tara_a	Nov.13, 2019	Wild	0.0	–	–
Takori_a	Nov. 15, 2019	Wild	1.8	Takori_a.1	75.0
				Takori_a.2	55.7
				Takori_a.3	81.7
				Takori_a.4	94.6
Takori_b	Nov. 15, 2019	Transplanted area‡	5.5	2014.1	92.8
				2014.2	94.8
				2014.3	92.1
				2014.4	91.9
				2014.5	92.9
Takori_c	Nov. 15, 2019	Transplanted area‡	8.5	2016.1	88.4
				2016.2	87.0
				2016.3	83.7
				2016.4	82.6
				2016.5	86.7
				2016.6	79.7
				2016.7	93.6
Higashiyoka_a	Nov. 12, 2019	Wild	0.0	–	–
Higashiyoka_b	Oct. 18, 2019	Wild	0.0	–	–
N/A†	Dec. 10, 2006	Transplanted area‡	N/A†	2005.1	66.7
				2005.2	74.7

*: Catch per unit effort (individual person^{-1} hour^{-1})

†: Not available

‡: Transplanted area of artificial seedlings

Figures

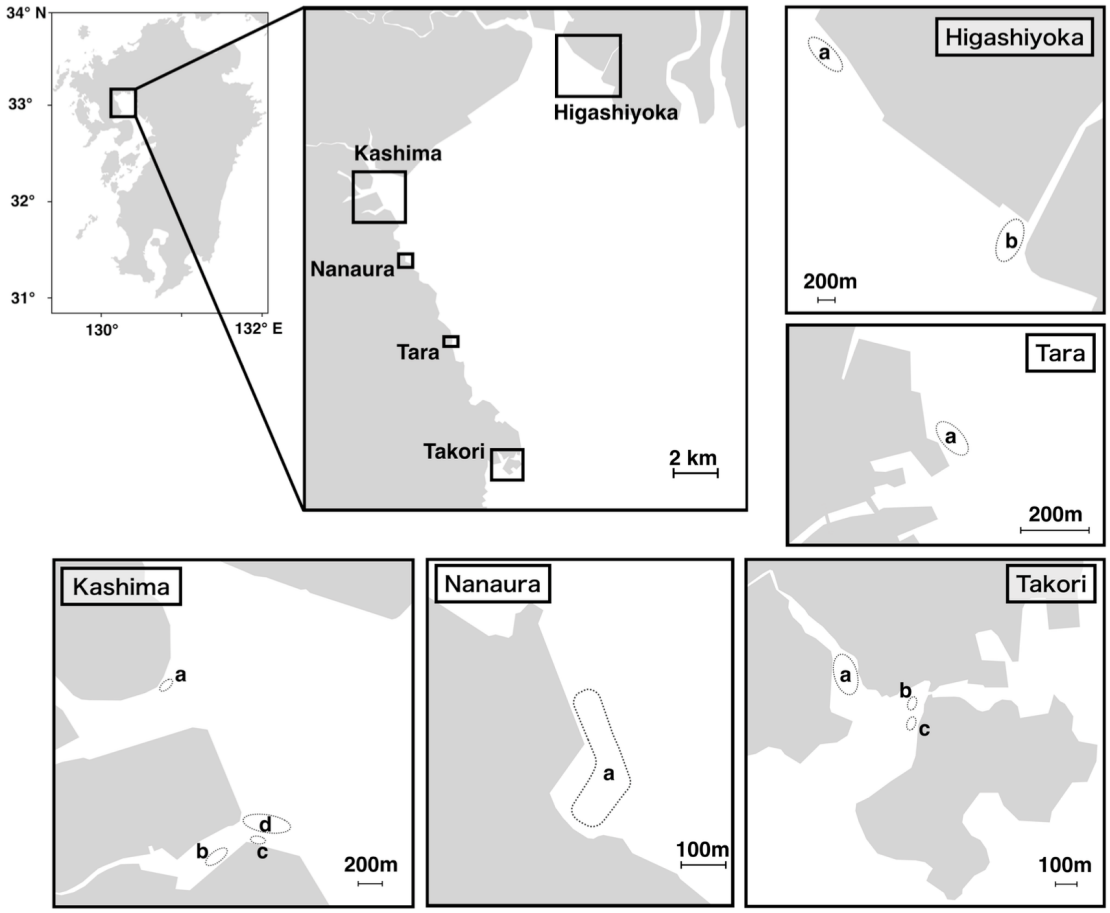

Figure 1. Sampling sites. The dotted line shows the area explored.

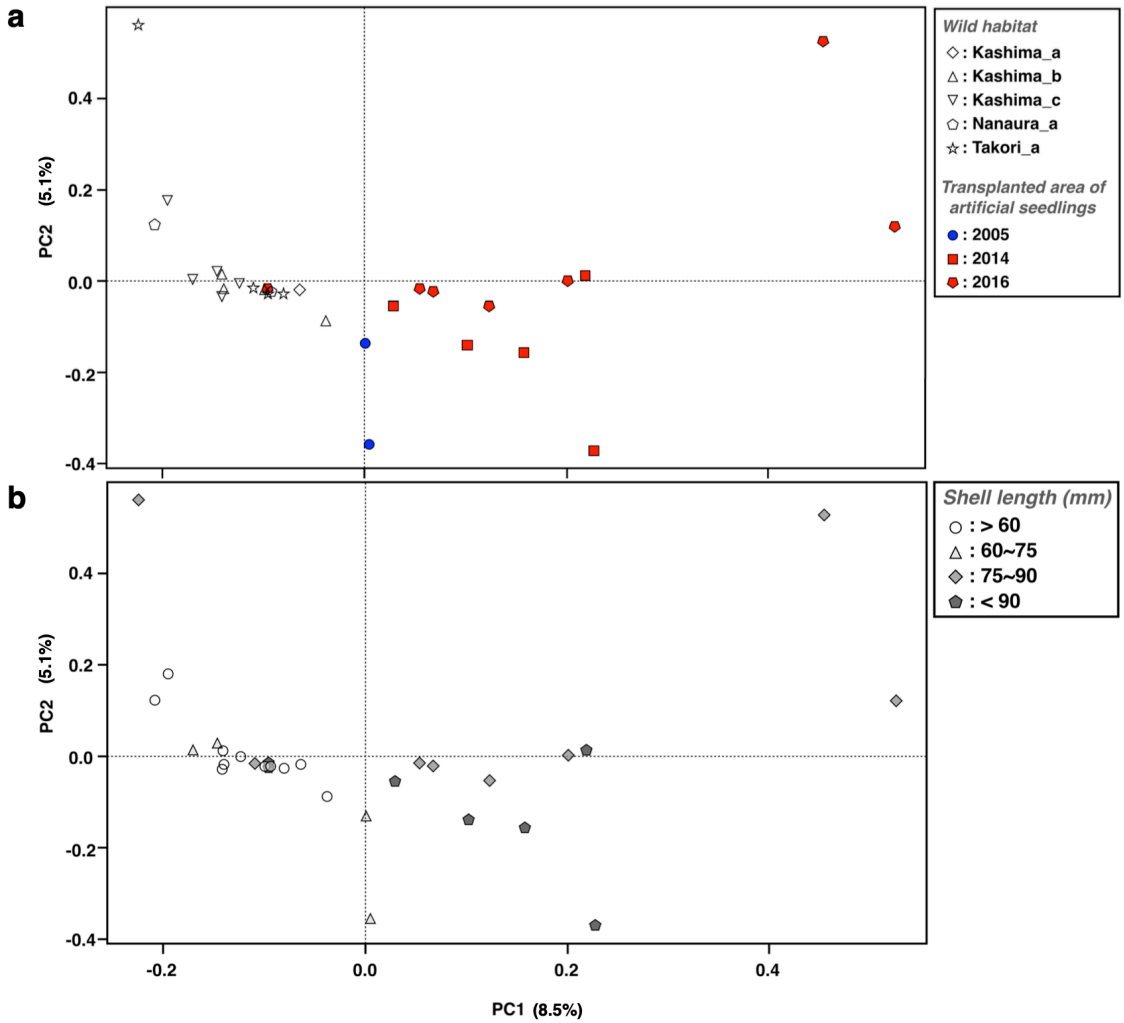

Figure 2. Biplots of principal component analysis. (a) Showing by sampling areas. (b) Showing by shell sizes.

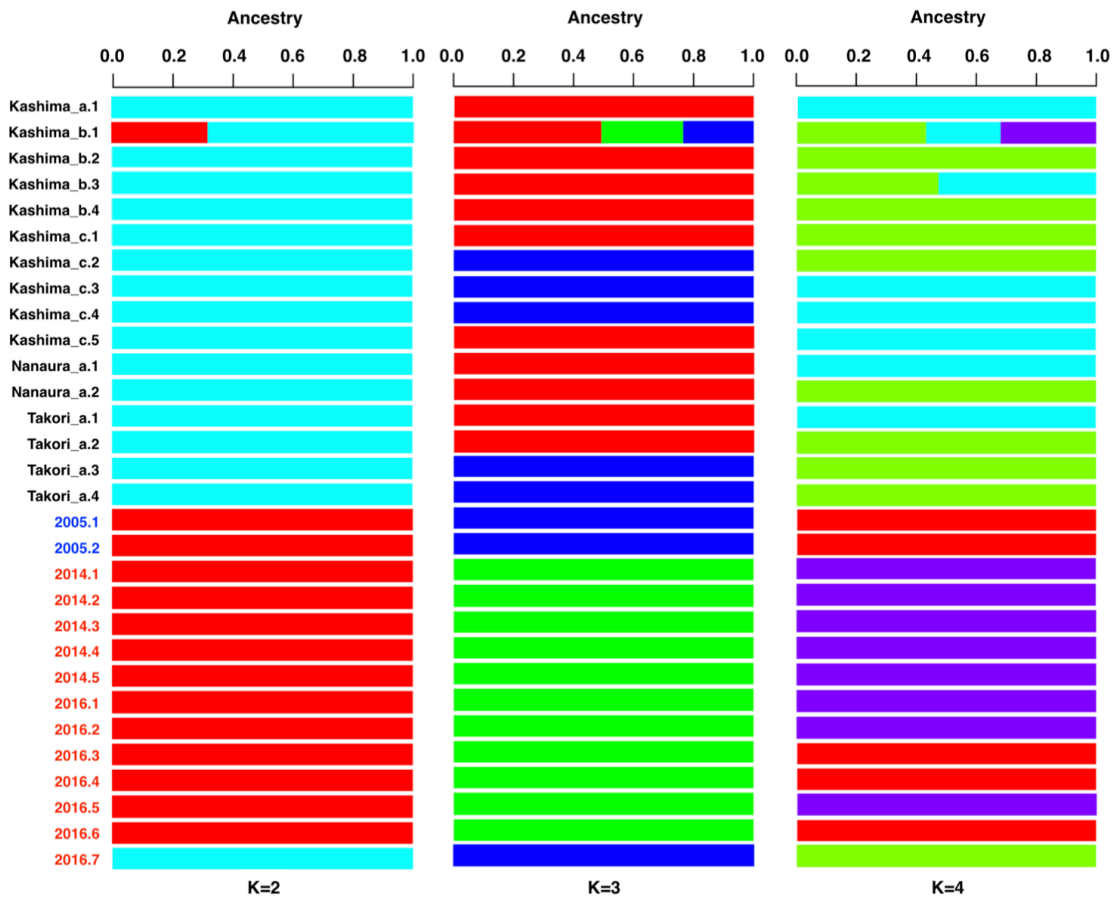

Figure 3. Admixture analysis of *S. constricta* in the Ariake Bay, Japan.

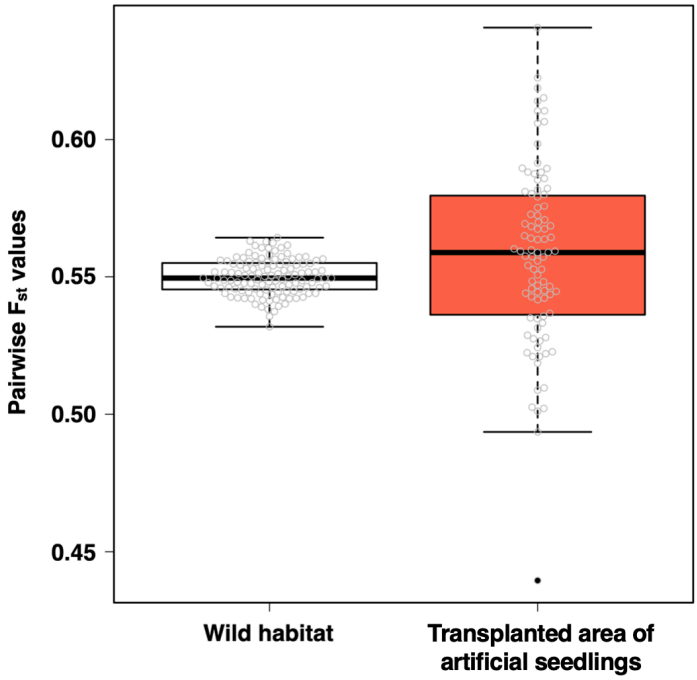

Figure 4. Pairwise F_{st} values between individuals in each population.

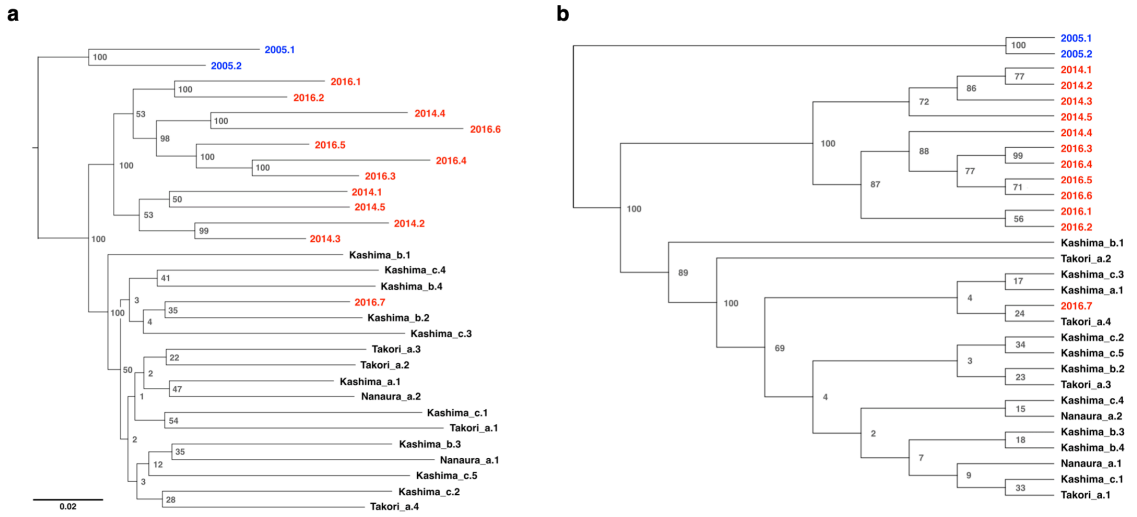

Figure 5. Phylogenetic trees of *S. constricta* inferred from RAD-Seq dataset (a) Maximum likelihood phylogeny estimated using RAxML. (b) Species tree inferred using SVDquartets. Numbers at the nodes indicate bootstrap values (1000 replicates).

455 **Supplementary material**

456

457

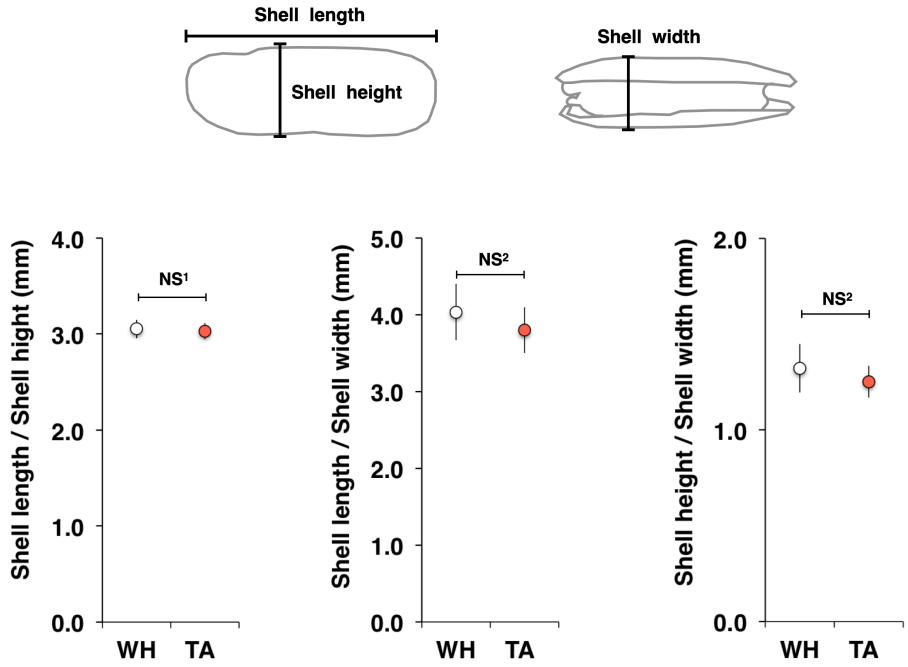

Fig. S1 Comparison of shell morphology. WH: individuals in wild habitat (n=17), TA: individuals in transplanted area (n=11). NS¹: non-significant difference between the comparison groups using the Student's t test. NS²: non-significant difference between the comparison groups using the Wilcoxon rank sum test. Bar indicate the SD.

458