scientific reports

OPEN

Genetic population structure of Japanese freshwater crab, *Geothelphusa dehaani* species complex using genome wide SNPs

Taiga Kunishima[1,2✉], Kento Takata[1,3✉], Kanto Nishikawa[3,4], Kodai Gibu[5,6], Miyuki Nishijima[6] & Akira Iguchi[6,7]

The Japanese freshwater crab *Geothelphusa dehaani* species complex is distributed widely across the Japanese Archipelago. Despite its suggested high genetic and morphological diversity, key aspects such as nuclear DNA (nuDNA) population structure and relationship between body color patterns and genetic populations remain unclear. To address these gaps, this study analyzed genome-wide single nucleotide polymorphisms (SNPs) in nuDNA and mitochondrial DNA (mtDNA) cytochrome oxidase subunit 1 (COI) markers in samples from Hokkaido to the Tokara Islands, Japan. ADMIXTURE analysis identified five distinct populations with significant geographic boundaries. These populations exhibited unique geographical patterns, spanning across islands and enclave distribution, indicating that *G. dehaani* populations have been shaped by complex factors, including sea level changes and volcanic activity. Regional body color variations partially aligned with SNP clades. Further, combining body color with collection locality data could help identify the specimen populations. Contrasting patterns between mtDNA and nuDNA suggest historical gene flow and introgression, emphasizing the need for caution when interpreting earlier phylogenetic studies based on combined mtDNA and nuDNA sequences. Our findings provide a foundational baseline for future research into the taxonomy, phylogeny, and population dynamics of the *G. dehaani* species complex, advancing our understanding of its evolutionary history.

The genetic population structure of a species is shaped by interactions between geological events, climatic changes, and ecological dynamics[1–3]. These processes influence population divergence, gene flow, and adaptation, contributing to current biodiversity. Primary freshwater crabs exhibit a unique evolutionary history, characterized by a land-locked life history and direct development without a planktonic larval stage, and these traits often lead to genetic differentiation and speciation among islands[4–6]. This relationship between life history traits and genetic structure, shaped by the geographical history, makes these crabs a valuable model for understanding genetic divergence mechanisms.

In particular, the genus *Geothelphusa* is the northernmost and diverse freshwater crab group in East Asia, spanning from Taiwan to the Japanese Archipelago[6], exhibiting island-specific differentiation in some species[7–9]. *Geothelphusa dehaani* (Japanese name, Sawagani), the most commonly observed freshwater crab in Japan, is usually found in freshwater habitats from coastal areas to high-altitude mountains, from Nakanoshima in the northern Ryukyu Islands (29.9°N) to southern Hokkaido (41.5°N)[9–12]. This wide distribution makes it ideal for

[1]Present address: Wakayama Prefectural Museum of Natural History, 370-1 Funoo, Kainan, Wakayama 642-0001, Japan. [2]Present address: Faculty of Agriculture, Setsunan University, 45-1 Nagaotoge, Hirakata, Osaka 573-0101, Japan. [3]Graduate School of Human and Environmental Studies, Kyoto University, Sakyo, Kyoto 606-8501, Japan. [4]Graduate School of Global Environmental Studies, Kyoto University, Sakyo, Kyoto 606-8501, Japan. [5]Department of Ecosystem Studies, Graduate School of Agricultural and Life Sciences, The University of Tokyo, Yayoi, Tokyo 113-0032, Japan. [6]Geological Survey of Japan, National Institute of Advanced Industrial Science and Technology (AIST), Tsukuba, Ibaraki 305-8567, Japan. [7]Integrated Research Center for Nature Positive Technology, National Institute of Advanced Industrial Science and Technology (AIST), 1-1-1 Higashi, Tsukuba, Ibaraki 305-8567, Japan. ✉email: taiga.kunishima@setsunan.ac.jp; k.takata013@gmail.com

investigating how the geological history of the Japanese Archipelago has shaped the genetic structures of these freshwater crab.

Previous genetic studies on *G. dehaani* were limited in scope, using samples from restricted areas and examining partial regions of enzyme or mitochondrial DNA sequences[13–16]. Takenaka et al.[17] were the first to comprehensively analyze its phylogeny, revealing 10 populations across the Japanese Islands and identifying island-specific genetic differentiation using a combination of partial mitochondrial DNA (mtDNA) markers, namely cytochrome oxidase subunit 1 (COI) and 16S rRNA and nuclear DNA (nuDNA), namely internal transcribed spacer (ITS) and histone H3. However, the discordance between mtDNA and nuDNA phylogenies, often observed in other taxa, e.g., the geotrupid dung beetle *Phelotrupes auratus*[18] or the Japanese fire-bellied newt *Cynops pyrrhogaster*[19], indicates the need for further analysis. Such discordance can result from differing evolutionary rates or adaptive mtDNA introgression[20]. If this applies to *G. dehaani*, it is necessary to examine each genetic structure in detail to understand its phylogeny and evolutionary processes.

Differences in genetic structure often reflect phenotypic traits, such as body coloration. Geographic variation in body color is well-documented in *G. dehaani*[15,21–24], with individuals traditionally classified into three main color types: dark (DA), red (RE), and blue (BL)[21,22]. These body color types exhibit distinct geographic distributions, with DA being widespread and RE and BL more localized[21–25]. Body color could serve as a population marker if there were a relationship between body color and genetic structure; in this case, *G. dehaani* could be used as a model for studying gene flow, selection, and demographic history underlying color variation, as animal coloration is a key adaptive trait with diverse functions[18,26]. However, previous studies on color–genetic relationships yielded conflicting results, largely due to reliance on partial mtDNA data[14,15,17,24,27]. As mtDNA data alone offers limited insights into the demographic history and geographic division of color forms, genome-wide sequence data are necessary to elucidate the genetic basis of geographic color variation[18].

Genome-wide approaches using large datasets of single nucleotide polymorphisms (SNPs) in nuDNA help uncovering insights not visible by analyzing a region of mtDNA, nuDNA, or morphological analyses[28]. Further, genome-wide approaches provide a fundamental baseline to tackle taxonomic issues in the *G. dehaani* species complex, which may include cryptic species or distinct genetic groups. Molecular phylogenetic studies based on partial mtDNA regions have suggested multiple clades and potential undescribed species within *G. dehaani*[14,15,17,24,27]. In fact, Naruse and Ng[9] recently described two new species based on morphological differences, *G. mutsu* and *G. amakusa*, from Aomori in Honshu Island and Amakusa Island in Nagasaki Prefecture in Kyushu, respectively. They also redesignated the lectotype of *G. dehaani* sensu lato from its type locality in Nagasaki City. However, Naruse and Ng[9] did not describe the phylogenetic relationship among these three species, and Takenaka et al.[17] found no genetic differences between specimens from Amakusa and those near the *G. dehaani* type locality (Clade 7). This situation emphasizes the need to clarify the phylogenetic relationship within the *G. dehaani* species complex, including *G. amakusa* and *G. mutsu*.

Addressing these issues requires extensive sampling across the *G. dehaani* species complex distribution and advanced molecular tools, such as next-generation sequencing[17,27]. Accordingly, this study aimed to investigate the genetic population structure of the *G. dehaani* species complex using genome-wide SNP data. We comprehensively collected 504 specimens from 217 sites across the Japanese Islands to examine the phylogenetic relationships between nuDNA and mtDNA and assess a potential discordance. Moreover, we used approximate Bayesian computational (ABC) analysis to infer its demographic history and analyzed the relationship between body color and genetic structure to evaluate its potential as a population marker. Finally, we have discussed the evolutionary patterns of the *G. dehaani* species complex in the Japanese Islands.

Results
Body color distributions
In this study, we identified four body color types, following Chokki[22] and Naruse and Ng[9]: dark purplish brown carapace and thoracic legs (DA), dark brown carapace, sometimes with orange or brown color on the posterolateral portion, and reddish orange thoracic legs (RE), grayish blue or bluish green carapace, occasionally including dark reddish purple, with light grayish or light yellow thoracic legs (BL), and a sky blue carapace with whitish, brownish, or reddish thoracic legs (amakusa). We also found the OC type, which did not align with descriptions by Chokki[22] and Naruse and Ng[9], featured a reddish purple carapace with a milky white wedge pattern in the posterior two-thirds or a dark brown anterior half with a dark orange pattern on the anterior margin and posterior half, all with whitish bases on the thoracic legs (Fig. 1, Table S1). As expected, the distribution patterns in the Japanese Archipelago differed among body color types (Fig. 1): DA was widespread in Hokkaido, Honshu, Shikoku, Kyushu, and adjacent islands (Sadogashima, Okinoshima, and Goto islands); BL was locally concentrated along the Pacific coast, particularly in Honshu from the Boso Peninsula to the Izu Peninsula and the southeastern Kii Peninsula, southern part of Shikoku, southern-west part of Kyushu, and Osumi Islands (Yakushima and Tanegashima islands); RE was primarily found in southern Japan, including the southern Kii Peninsula, southwestern Shikoku, and Kyushu; amakusa was located in the Amakusa Islands; and OC in Nagasaki and Oita in Kyushu.

Genetic population structure based on mtDNA
The amplified mtDNA sequences ranged from 557 to 590 base pairs, yielding 195 haplotypes (Table S2). The *G. dehaani* species complex was divided into three monophyletic populations (clades pop1–3) and 14 subclades in the neighbor network (Fig. 2). Pop1 comprised specimens from Shikoku and the southeastern Kii Peninsula, further divided into three subpopulations: 1a (central and western Shikoku), 1b (southeastern Kinki), and 1c (southeastern Shikoku). Pop2 included specimens from Shikoku and Kyushu, comprising 2a (Kyushu) and 2b (Shikoku and Kyushu). Pop3 was separated into nine subpopulations: 3a (Hokkaido, Honshu from Tohoku to eastern Chugoku, and eastern Shikoku), 3b (western Chugoku and northwestern Shikoku), 3c (western
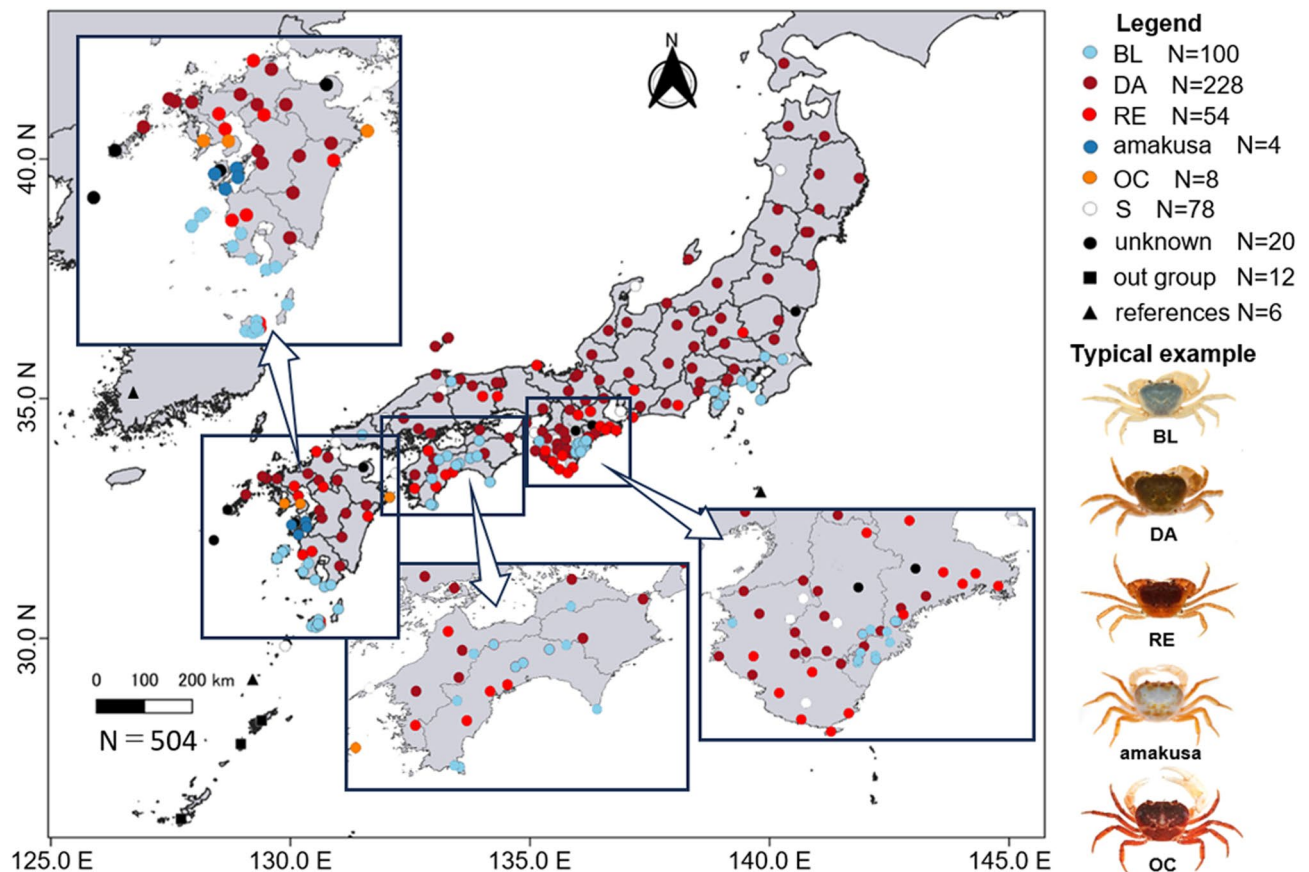
Fig. 1. Map showing the geographical distribution of the analyzed samples. Symbol shapes represent species and data sources; circles, the *Geothelphusa dehaani* species complex; squares, outgroups; triangles: reference data of *G. dehaani* sensu lato from the literature. Symbol colors represent body coloration types; sky blue, BL type; brown, DA type; red, RE type; blue, amakusa type: orange, OC type; white, unidentified specimens due to small individuals; and black, unknown.

Kyushu, including *G. amakusa*, and South Korea), 3d (northeastern Shikoku and Kyushu), 3e (Yakushima and Nakanoshima islands), 3f (Tanegashima Islands), 3 g (Koshikishima Islands), 3 h (southern Kanto and Eastern Shizuoka Prefecture), and 3i (Danjo Islands in Oshima Islands).

The haplogroup diversity of populations based on mtDNA was highest in pop3, nucleotide diversity was highest in pop2 and both were lowest in pop1 (Table S3). Tajima's *D* values and Fu's *Fs* values showed significant values ($P < 0.01$) only in pop3; both were negative. Pairwise fixation index (F_{ST}) values were significant between all sites, in the following order: pop1 - pop2 < pop1 - pop3 < pop2 - pop3 (Table S4).

Genetic differences among populations based on MIG-seq

Table S5 shows the number of SNPs and CV error values for datasets 0–6. For set 0 (All individuals included in the SNP analysis), the lowest CV error value was obtained for K = 5, dividing the population into five groups: HO (Hokkaido, Honshu from Tohoku to eastern Chugoku, and Shikoku), SHI (southeast Kinki and Shikoku), nKC (western Chugoku and northern Kyushu), cK (central Kyushu, including *G. amakusa*), and sKK (southern Kyushu, southern Kanto, and eastern Shizuoka) (Fig. 3; Fig. S1, S2). These groups partially corresponded to our observed color types, excluding juveniles (S) and individuals that could not be color-discriminated (un), with HO primarily comprising DA (89% of all specimens) and RE (11%); SHI primarily comprising BL (71%); nKC primarily comprising DA (80%) and RE (7%); cK primarily comprising DA (48%), RE (24%), and amakusa (14%); and sKK predominantly comprising BL (100%) (Figs. 1 and 3). In the SHI population, only four specimens corresponded to DA and RE types (specimen numbers 62, 276, 432, and 527). A few OC specimens were observed in nKC and cK populations. Among the 140 samples in sets 1–6, 128 had a Q value (Q) of > 95% (Fig. S3, S4). Twelve samples had Q values between 5% and 95%, including set 1 (Kanto region: Nos. 104, 105), set 2 (Kinki region: No. 166), set 4 (Shikoku region & Awaji Island: No. 346), set5 (Northern Kyushu: Nos. 422, 423) and set 6 (southern Kyushu: Nos. 422, 423, 457, 458, 465, 483) (Fig. S4). Except for No. 483, these samples were collected near geographical population boundaries.

The results for both variant positions and all positions were consistent in the five populations (Table 1). The highest nucleotide diversity (π) value was observed in sKK, whereas HO and SHI had the lowest. nKC had the highest observed heterozygosity (*Ho*), and sKK the highest expected heterozygosity (*He*). All groups exhibited
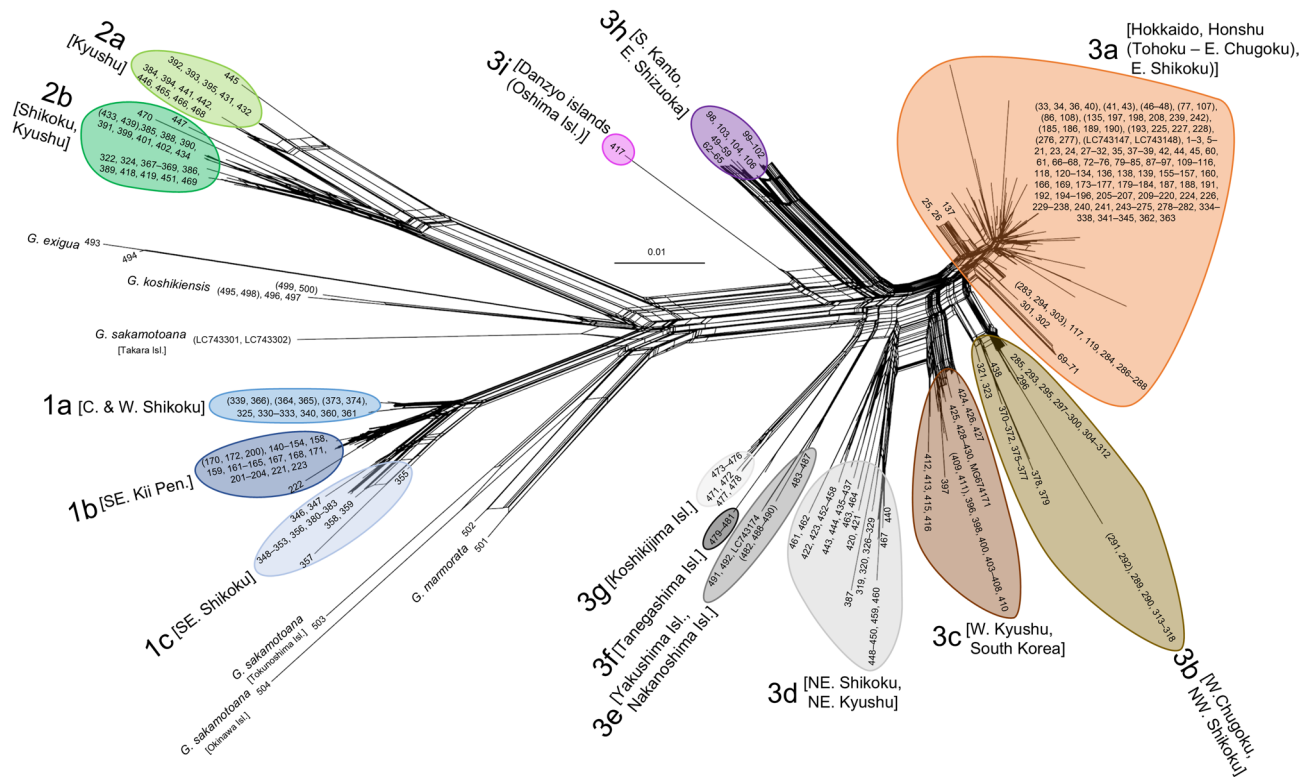
Fig. 2. Neighbor-net phylogenetic network based on the 557–590 base pairs of the mtDNA COI region.

higher He than Ho, indicating homozygous excess, with all inbreeding coefficient (F_{IS}) > 0. Pairwise F_{ST} and corrected F_{ST} (F'_{ST}) values were all high and significant (Table 2, F_{ST} = 0.3375–0.5746; F'_{ST} = 0.3382–0.5684, all $P < 0.01$), indicating deep genetic divergence among populations. The F_{ST} values were consistent across populations, being larger in SHI-HO, SHI-cK, and HO-sKK pairs, and smallest in the HO-nKC pair (Table 2).

Demographic history

To explore the demographic history of the five genetic clades, we analyzed individuals composed of single genetic elements based on the ADMIXTURE results (see "Materials and methods"), using DIYABC-RF based on 313 SNPs. Principal component analysis (PCA) pre-scenario checks indicated that the observed dataset aligned well with the simulated dataset (Fig. 4a), suggesting that the analysis conditions were suitable for random forest analysis.

Among the 32 hypothetical divergence scenarios tested (Fig. S5), Scenario 30 was selected by 52.0% votes (Table S6, 520 out of a total 1000 votes) as the best fit, with a mean posterior probability of 0.831 and prior and posterior error rates (i.e., global and local errors) of 0.133 and 0.169, respectively. Scenario 30 proposed a divergent history where the SHI population diverged first, followed by HO, with nKC, sKK, and cK populations differentiating around the same time (Fig. 4b). The expected time of divergence events for the SHI population at t3 and for the HO population at t2 were estimated as 2.78×10^5 (95% confidence interval (CI): 1.35×10^5–4.32×10^5) and 1.94×10^5 (95% CI: 1.04×10^5–3.00×10^5) generations ago, respectively. Finally, the expected time of divergence events in the Kyushu at t1 was estimated as 1.06×10^5 (95% CI: 5.59×10^4–1.59×10^5) generations ago. Assuming the generation time to be four years in the *G. dehaani* species complex[29], divergent events were expected to have occurred in the SHI population at 1.11 (95% CI: 0.54–1.72) MYA, in the HO population at 0.77 (0.41–1.20) MYA, and for the nKC, cK, and sKK populations at 0.42 (0.22–0.63) MYA. Scenarios diverging by body color (Scenarios 1–3) received only two votes.

Discussion
Genetic population structure and evolutionary history

In this study, the ADMIXTURE analysis identified five populations across the Japanese Archipelago, delineated by distinct geographical boundaries. The ABC results suggest that the divergence order is firstly SHI, then HO, and lastly three Kyusyu populations. In addition, a clear differentiation between populations was demonstrated in F'_{ST} (Table 2; Fig. 4). In this study, we treat all populations as intra-specific to avoid taxonomic confusion, because at the present stage, we cannot judge whether those are cryptic species without morphological observations.

Our findings indicate that the genetic population structure of the *G. dehaani* species complex was shaped by multiple geological events throughout its complex evolutionary history. Notably, most populations, except the cK population distributed in Kyushu, span multiple islands. Takenaka et al.[17] identified genetic differentiation among nine major populations according to island, based on a combination of partial mtDNA and nuDNA
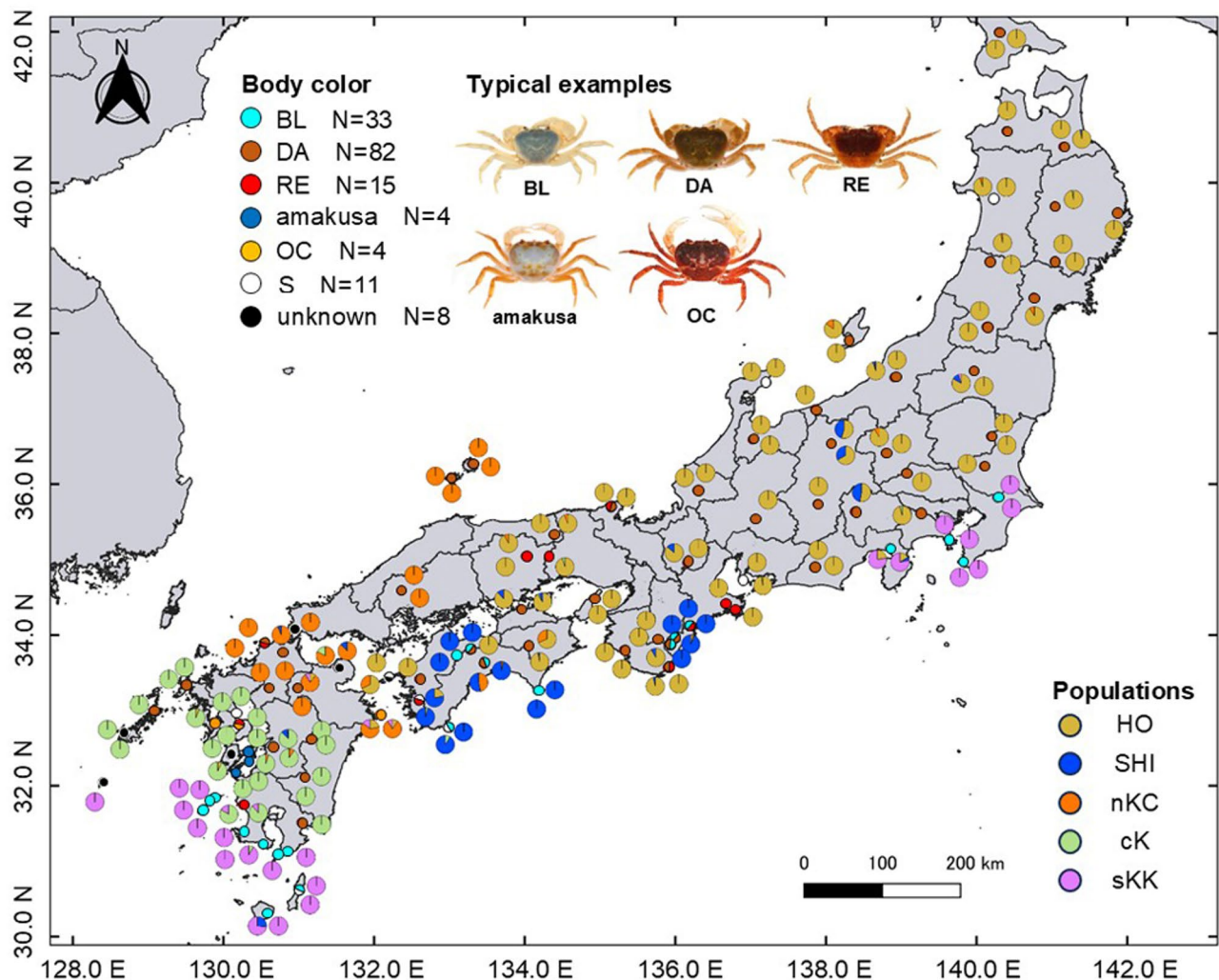
Fig. 3. Map showing the results of the ADMIXTURE analysis for each location (set 0) and the distribution of color types. Large pie chart, genetic elements of each individual; small pie chart, color types of individuals at each location used in this analysis. Color types are indicated as in Fig. 1. The colors in the pie charts corresponding to each group and genetic elements are as follows: HO (brown), Hokkaido, Honshu from Tohoku to eastern Chugoku, Shikoku; SHI (blue), southeast Kinki and Shikoku; nKC (orange), western Chugoku and northern Kyushu; cK (green), central Kyushu, including *Geothelphusa amakusa*; and sKK (pink), southern Kyushu, southern Kanto and eastern Shizuoka. The colors of the ADMIXTURE graph correspond to each genetic element on the map, and the sample numbers refer to Table S1. Distribution map with sample numbers represents as Fig. S2.

regions (COI, 16S rRNA, ITS, and histone H3), suggesting that straits between islands act as major genetic barriers for *G. dehaani*. In contrast, our results highlight the importance of geological events, including volcanic activity and climate change, in shaping the genetic structure of this species group.

For example, the co-occurrence of three populations (nKC, cK, and sKK) on Kyushu Island cannot be explained solely by the geographical barrier posed by the strait, as these populations are separated along a North–South axis with distinct geographical boundaries. ABC analysis suggested that they diverged nearly simultaneously from a common ancestor, implying a large-scale or continuous event disrupted gene flow across Kyushu. Moreover, the identified boundaries appear to be influenced by sea level fluctuations and volcanic activity. For example, the nKC population spans the Kanmon Strait into central Chugoku and northern Kyushu. Despite its unclear boundary with the HO population, it aligns with the distribution boundaries of other species, such as calopterygid damselflies and small salamanders[30,31], hinting at historical events in the Chugoku region that impeded gene flow. Major rivers, such as Gohnokawa and Ota, which flow across the Chugoku Mountains to the Sea of Japan, may have once served as barriers connected to the Seto Inland Sea[32]. In contrast, the southern boundary between the nKC and cK distributed in Kyushu lacks a singular geographical barrier, suggesting that multiple factors influenced their genetic divergence. For example, during past interglacial periods, the Tsukushi and Fukuoka plains in northern Kyushu, which were submerged by rising sea levels, restricted the movement of freshwater organisms between this plain and the adjacent Sefuri Mountains and other mountain ranges[27,33].

	Variant positions				All positions (variant and fixed)			
	π	Ho	He	F_{IS}	π	Ho	He	F_{IS}
HO	0.05981	0.03748	0.05902	0.12375	0.00034	0.00022	0.00034	0.00071
SHI	0.05715	0.03481	0.05058	0.06024	0.00033	0.00020	0.00029	0.00035
nKC	0.08578	0.05196	0.07679	0.07603	0.00049	0.00030	0.00044	0.00044
cK	0.08015	0.04769	0.07528	0.10178	0.00046	0.00028	0.00044	0.00059
sKK	0.10926	0.04516	0.09996	0.18168	0.00063	0.00026	0.00058	0.00105

Table 1. Overall population genomic statistics for the *Geothelphusa dehaani* species complex based on 354 SNP loci (variant positions) and 9058 SNP loci (all positions). Abbreviations: observed heterozygosity (*Ho*), expected heterozygosity (*He*), nucleotide diversity (π), inbreeding coefficient (F_{IS}). Population abbreviations were based on the results of ADMIXTURE: HO (Hokkaido, Honshu from Tohoku to Eastern chugoku, and Shikoku), SHI (southeast Kinki and Shikoku), nKC (western Chugoku and northern Kyushu), cK (central Kyushu, including *G. amakusa*), and sKK (southern Kyushu, southern Kanto, and Eastern Shizuoka).

Fst / Fst'	HO	SHI	nKC	cK	sKK
HO		0.5237*	0.3382*	0.5501*	0.5684*
SHI	0.4919* 0.3783–0.5845		0.4520*	0.5636*	0.4963*
nKC	0.3375* 0.2400–0.4284	0.3982* 0.2929–0.4940		0.4376*	0.4962*
cK	0.5299* 0.4313–0.6152	0.5110* 0.4171–0.6010	0.3747* 0.2828–0.4620		0.4654*
sKK	0.5746* 0.4907–0.6478	0.3983* 0.2894–0.4913	0.4319* 0.3400–0.5090	0.4262* 0.3258–0.5182	

Table 2. F_{ST} and F'_{ST} values among the five *Geothelphusa dehaani* species complex populations based on 354 SNP loci. Bottom left: pairwise fixation index (F_{ST}), top right: corrected F_{ST} value (F'_{ST}). Population abbreviations were based on the results of ADMIXTURE: HO (Hokkaido, Honshu from Tohoku to Eastern chugoku, and Shikoku), SHI (southeast Kinki and Shikoku), nKC (western Chugoku and northern Kyushu), cK (central Kyushu, including *G. dehaani* in lato type locality and *G. amakusa* distribution), and sKK (southern Kyushu, southern Kanto, and Eastern Shizuoka). Asterisk indicates the significant at $P < 0.01$. Lower values in F_{ST} represent 95% CI.

Similarly, nKC and CK populations, which are distributed adjacent to each other along the Tsukushi and Fukuoka plains, may have diverged for the same reason. Additionally, volcanic activity between 5–6 million years ago (MYA) and 0.5–2.0 MYA along the Beppu–Shimabara Graben, which runs East–West through central Kyushu[34], likely shaped local barrier and limited gene flow. In fact, the latter period activity may have influenced the amphibian distribution[35]; the estimated divergence age for the Kyushu population in this study (0.42 MYA) is consistent with this hypothesis. Such disturbances promoted local isolation and genetic differentiation. Similarly, the cK–sKK boundary in southern Kyushu appears to be influenced by volcanic activity, as noted by Suzuki & Tsuda[23].

The HO, SHI, and sKK populations exhibited a unique distribution pattern, with the latter two clearly displaying enclave distributions. Two scenarios could explain such distribution: a stable contact zone scenario, in which a population persists as a remnant patch of a historically widespread distribution, or a moving contact zone scenario, in which a population migrates into new areas[36]. The SHI population distributed in Shikoku likely aligns with the first scenario for several reasons: (1) geographic and ecological barriers make recent colonization unlikely, (2) exhibits low genetic diversity ($\pi = 0.057$), suggesting a historical bottleneck, and (3) evidence of rapid expansion in the HO population by low genetic diversity in SNPs ($\pi = 0.059$) and neutrality tests, including a negative Tajima's *D* (-1.69) and Fu's *Fs* (-23.7). ADMIXTURE analysis further supports this hypothesis, revealing the SHI population share ancestry with the HO population in both the Kii Peninsula and Shikoku Island (Fig. S1). These findings suggest that the SHI population once extended across these regions but got fragmented as the rapidly expanding HO population isolated SHI individuals into enclaves. In contrast, presence of the HO population in northwestern Shikoku represents a moving contact zone scenario driven by its rapid expansion. Notably, the HO population exhibits the widest distribution among all, extending as far as Hokkaido and crossing the Blakiston Line, a key faunal boundary. Our results revealed no significant genetic differences in mtDNA and nuDNA between the Hokkaido specimens and other HO specimens, indicating either recent expansion or human-mediated introduction, as proposed by Sugime et al.[11]. These findings highlight the dynamic role of the HO population in shaping *G. dehaani* population distribution through rapid expansion, and its influence on the enclave distribution of other populations.

On the other hand, the sKK population from the southern Kyushu, its adjacent islands, and along the coastal areas from the Izu Peninsula to the Boso Peninsula (southern Kanto), forms an enclave, with a disjunct
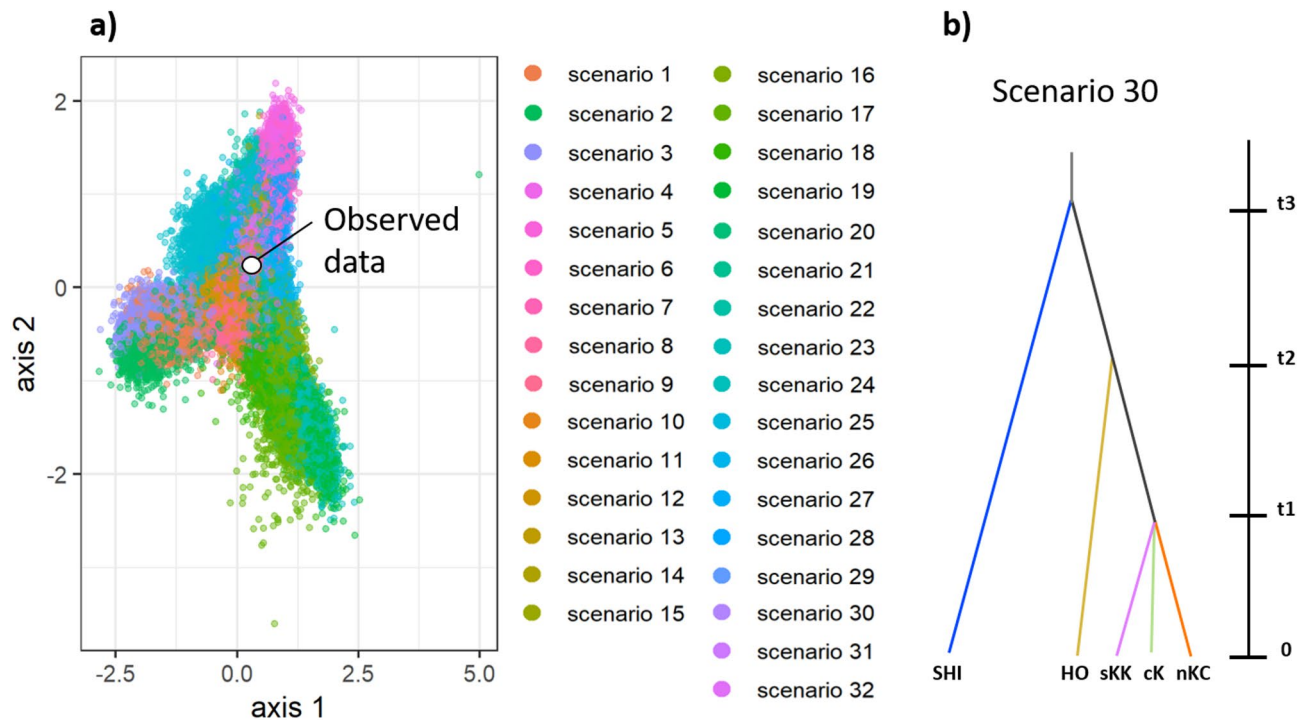
Fig. 4. Results of DIYABC-RF for the *Geothelphusa dehaani* species complex based on 313 SNPs. (**a**) Principal component analysis plots evaluating the fit between the observed data and simulated datasets for 32 demographic scenarios of the *G. dehaani* species complex in the Japanese Islands. (**b**) The best voted scenario for the divergence history of *G. dehaani* populations (Scenario 30; see Fig. S4).

distribution across Kyushu and Honshu islands that creates a unique geographic pattern. Takenaka et al.[17] proposed that the southern Kanto lineage (clades 3h in this study, see Fig. 2) expanded via oceanic dispersal from southern Kyushu when the Izu Peninsula was an island (i.e., moving contact zone scenario), and estimated their divergence at 1.1–6.0 MYA using mtDNA 16S and COI regions. Our genome-wide SNP data estimated the divergence time for the sKK population at 0.42 (0.22–0.63) MYA, earlier than their estimation. If the spread to the southern Kanto region occurred millions of years ago, the *G. dehaani* species complex, which lacks a planktonic larval stage, should have differentiated after geographical isolation and produced significant differences in our ADMIXTURE analysis, which suggests an overestimation of the divergence time for southern Kanto lineage at 1.1–6.0 MYA. Hence, we propose two alternative possibilities: (i) dispersal via marine routes after collision of the Izu Peninsula with Honshu, or (ii) persistence as remnant patches of a past widely distributed population. On the other hand, our divergence time estimation by the DIYABC is limited in its accuracy due to the absence of fossil records or secondary calibrations for *Geothelphusa* species, preventing the use of an external molecular clock. Future studies employing high-resolution methods, such as whole-genome sequencing, are necessary to clarify their intrapopulation genetic structure and provide deeper insights into the unique distribution of the sKK population.

Relationship between body color trait and genetic population structure

This study did not find a complete correspondence between body color types (DA, RE, BL, OC, and amakusa) and populations using COI and SNP analyses. However, we observed a biased distribution pattern along Pacific coast areas for the BL and RE types. Such a biased distribution of specific coloration types could arise either from repeated local adaptations to coastal climate and environment or from phenotypic plasticity that is environmentally induced along the same gradient. Although the BL type emerged twice independently in the evolutionary history of the *G. dehaani* species complex (Fig. 4, sKK and SHI populations), may be due to the repeated climate adaptation. For example, pale carapace coloration in crustaceans may play a thermoregulatory role[37]; furthermore, the range of the BL type overlaps with warm Pacific coastal areas influenced by the Kuroshio Current. In addition, BL type does not change across the body color types through rearing[24], indicating that genetic adaptation determines to blue body color trait. On the other hand, the RE type, which showed a similar distribution pattern to BL type, may be due to phenotypic plasticity for environmentally or dietary variation. Although COI and SNP data revealed no genetic differences between DA and RE types within the same populations, dietary availability of astaxanthin precursors, such as β-carotene and zeaxanthin[23,38], may generate a plastic response. Alternatively, an undetected genetic structure within the population may be another contributor. Regardless of whether the mechanism is the genetic or the plastic, the repeated emergence and coastal bias of the BL and RE types make the *G. dehaani* species complex a promising model for investigating how color traits respond to climate and dietary gradients.

Despite the incomplete correspondence between body color traits and populations, combining body color with collection locality data could help identify the population affinities of collected specimens. For example, the sKK population, exclusively comprising BL individuals, spans southern Kyushu and southern Kanto, indicating that body color and collection locality can together distinguish populations. Similarly, the HO and nKC populations, predominantly of DA or RE type, can often be identified this way, despite rare BL-type individuals (Table S1, No. 249 in the HO population area and No. 313 in the nKC area). These rare cases likely represent an individual-level color variant, considering the limited collection sites and specimen numbers, similar to findings for *Ryukyum yaeyamense*[39].

In contrast, the SHI and cK populations exhibited greater intraspecific variation in body coloration (Figs. 1 and 3). Although the SHI population predominantly comprises BL individuals, it also includes DA/RE types. Mismatched specimens sharing multiple common ancestors (Fig. 3) suggest that the genetic introgression through secondary contact contributes to the discordance in body color within the population. The cK population also displayed various coloration types, including DA, RE, OC, and even *G. amakusa* types, despite consistent genetic proportions (Fig. 3). Such variations in SHI and cK populations imply historical or ongoing gene flow and highlight the need for reevaluating the genetic and morphological boundaries in the *G. dehaani* species complex, including *G. amakusa*. Molecular methods, especially genome-wide analyses, are essential for resolving the genetic structure and identifying individuals in SHI and cK effectively in future research.

Discordance between mtDNA and nuDNA sequences

Our molecular analysis highlights the need for caution when interpreting previous studies relying on the combinations of partial mtDNA and nuDNA sequences. While our COI-based neighbor-net diagram aligns with earlier low-resolution studies using isozymes or partial sequences[14,15,17,27], the population structure revealed by our ADMIXTURE analysis using SNPs significantly differs. This discordance emphasizes the importance of genome-wide analyses alongside partial sequence analyses for accurately elucidating the genetic phylogeny and population structure of the *G. dehaani* species complex.

Discordances between mtDNA and nuDNA are common in various taxa and often attributed to incomplete lineage sorting or other biological processes for mtDNA (e.g., adaptive introgression, demographic disparities, or sex-related asymmetries)[20]. Although at present we cannot determine the mechanism and process of mito-nuclear discordance in the *G. dehaani* species complex, ILS is unlikely for this reason; there was a strong geographic discordance between patterns of nuDNA and mtDNA and geographically coherent distribution patterns (Figs. 1 and 3), whereas stochastic ILS should not be expected to produce such clear biogeographic pattern[20]. For the many taxa where mtDNA and nuDNA do not match, mtDNA is less structured than nuDNA[20]; a similar result was obtained for the *G. dehaani* species complex.

In contrast, population overlap areas shown by nuclear information, individuals exhibited genetic elements from both populations, indicate the past secondary contact between two populations (e.g., nKC and HO populations; Fig. 3). Breeding between isolated populations with divergent patterns of gene flow between mtDNA and nuDNA may promote the mito-nuclear discordance[20]. Our results imply that nuDNA may unify within populations with maintenance over time of mtDNA variation due to gene flow or ecological differences, such as sex-biased mobility. We, therefore, speculate that repeated contact and separation between populations created the complex structure observed in mtDNA COI, in contrast with the population structure derived from inter-simple sequence repeat (ISSR) SNPs. Our research indicates that gene flow may still be ongoing or might have historically occurred but has since ceased. This study did not analyze the gene flow dynamics in detail, as the samples were selected to investigate the overall genetic characteristics of *G. dehaani* populations. Future studies focusing on genetic boundary regions are needed to clarify these patterns.

Conclusion

Our study reveals the complex genetic population structure of the Japanese freshwater crab *G. dehaani* species complex, resulting from multiple geological events, including volcanic activity and sea level fluctuations, in the Japanese Islands. ADMIXTURE analysis identified five populations with distinct geographic boundaries. On Kyushu Island, three distinct populations (nKC, cK, and sKK) separated along a North–South gradient were likely influenced by large-scale disturbances, such as volcanic activity. The recent rapid expansion of the HO population, indicated by low genetic diversity, has contributed to the unique enclave distribution of the SHI population. Previously, a high genetic segmentation in *G. dehaani* sensu lato was expected because of its low dispersal ability and lack of planktonic phase; in fact, a recent study divided it into 10 populations or cryptic species across the Japanese Islands using a combination of partial mtDNA markers[17]. In contrast, our ADMIXTURE results indicate that *G. dehaani* species complex has a higher dispersal ability than expected. Thus, we hypothesized that adult crabs can move over wide distances, and the high adaptation of *G. dehaani* sensu lato to terrestrial life and terrestrial excursions with roaming behavior could support our hypothesis[16,40]. Further, regional body color variations partially correlate with SNP clades, suggesting the potential of combining collection locality and body color for population identification. However, further molecular and morphological approaches are needed for precise population-identification from specimens. Such integrating approaches will also contribute to resolve the taxonomic issues in *G. dehaani* species complex. Contrasting patterns between mtDNA and nuDNA highlight historical gene flow and introgression, underscoring the need for comprehensive genomic analysis. In addition, the discordance between nuclear SNPs and mtDNA classifications warrants caution when interpreting earlier phylogenetic studies based on combined mtDNA and nuDNA data. These findings provide new insights into the evolutionary history and genetic diversity of the *G. dehaani* species complex, providing a basis for future research on its taxonomy, phylogeny, and population dynamics.

Materials and methods

Sample collection and laboratory procedures

From 2019 to 2023, 504 *Geothelphusa* specimens, including *G. dehaani* species complex, *G. exiguia*, *G. koshikiensis*, *G. marmorata*, and *G. sakamotoana*, were collected from 217 sites, spanning from Hokkaido to the central Ryukyu Islands (Table S1). In this study, *G. dehaani*, *G. mutsu*, and *G. amakusa* were all considered as *G. dehaani* species complex without distinction due to ambiguous species boundaries and limited distribution knowledge. Almost all specimens were transported to the laboratory either alive or anesthetized, and photographs were captured to document live body coloration. To correct for color differences caused by photographic conditions, we performed color correction using Adobe Photoshop element 8.0 and a color scale bar (Casmatch, Bear Medic Co.) for quantitative comparisons. The specimens were categorized into four types based on Chokki[21,22] and Naruse and Ng[9]: BL (grayish blue or bluish green carapace, occasionally including dark reddish purple, with light grayish or light yellow thoracic legs), DA (dark purplish brown carapace and thoracic legs), RE (dark brown carapace, sometimes with orange or brown color on the posterolateral portion, and reddish orange thoracic legs), and "amakusa" (a sky blue carapace with whitish, brownish, or reddish thoracic legs). Specimens with unclassifiable color patterns were labeled as OC (other color). Specimens were fixed in 80% ethanol, and part of a leg was transferred to 99% ethanol for DNA analysis. Some specimens were directly fixed in 99.5% ethanol in the field. In addition, out-group specimens, including *G. exigua* (*n* = 2, collected at southern Kyushu), *G. marmorata* (*n* = 2, Yakushima Island), *G. koshikiensis* (*n* = 6, Koshiki Islands), and *G. sakamotoana* (*n* = 2, Okinawa Island and Tokunoshima Island) were treated similar to *G. dehaani* samples. The identification of out-group specimens was based on morphological characteristics, such as body color and male first pleopod shapes[10]. All fixed specimens were registered and are preserved at the Wakayama Prefectural Museum of Natural History (WMNH). Our research protocols complied with the current laws of Japan, including the animal welfare ones, and followed the guidelines for protecting and promoting decapod crustacean welfare in research, by the Insect Welfare Research Society[41].

DNA extraction, sequencing, and alignment

Mitochondrial DNA

A total of 485 specimens from 209 localities were selected for mtDNA analysis of the COI region, covering the distribution ranges of the *G. dehaani* species complex (Table S1). This selection criteria were set to include specimens as comprehensively as possible from Hokkaido to Tokara Islands, ensuring a variety of color patterns were included. Genomic DNA was extracted from the leg muscles of each specimen using a DNeasy Blood and Tissue Kit (Qiagen, Hilden, Germany). Partial sequences of the mtDNA COI-encoding region were amplified by PCR using the universal primers jgLCO1490 and jgHCO2198[42] with TaKaRa Ex Taq (TaKaRa, Shiga, Japan). The PCR protocol involved mixing 10 μL of Premix TaqTM (Takara Bio Inc., Tokyo, Japan), 1.2 μL of each primer (10 μM), 6.6 μL of Milli-Q water, and 1 μL of template DNA. PCR was performed using Applied Biosystems VeritiPro (Thermo Fisher Scientific K.K., Massachusetts, USA) under the following conditions: 94 °C for 1 min, followed by 35 cycles of 94 °C for 40 s, 51 °C for 40 s, and 72 °C for 1 min, and final elongation at 72 °C for 7 min. A 3-μL sample of the PCR product was mixed with 1 μL of Midori Green Direct DNA Stain (NIPPON Genetics, Tokyo, Japan) and electrophoresed on a 2% agarose gel at 100 V for 15 min. Then, the gel was visualized under UV light to confirm DNA amplification. Amplification products were purified by mixing 17 μL of PCR products with 1 μL of ExoSAP-IT solution (Thermo Fisher Scientific K.K., Massachusetts, USA) and incubating at 37 °C for 15 min and 80 °C for 15 min. The purified PCR products were then sent to Macro-gen Japan (Tokyo, Japan) for sequencing with the same primers used for PCR amplification. The resulting sequences were aligned using the MEGA 11 and ClustalW software[43]. The mtDNA sequences were registered in the DNA Data Bank of Japan (DDBJ) (accession numbers: LC864606–LC865102), and the haplotypes were determined (Table S2).

SNPs

Genome-wide SNPs were analyzed using multiplexed inter-simple sequence repeat genotyping by sequencing (MIG-seq)[44] for 154 individuals from 92 sites in the *G. dehaani* species complex utilizing the Illumina Hiseq system (Illumina) (Table S1). MIG-seq is a technology that uses ISSR as a universal multiplex PCR primer to simultaneously amplify and analyze multiple genomic regions to obtain genome-wide sequence information[44]. To ensure high-quality sequence data, reads 1 and 2, adapter regions, and low-quality reads were removed using fastp ver. 3[45]. The quality threshold was set to 30, with a required length of 134, trimming front 1 and front 2 by 14, cutting front/tail set to 20. The raw data obtained by MIG-seq were registered at DDBJ (accession numbers: DRR641568–DRR641721).

Population genetic analysis

Mitochondrial DNA analysis

In addition to the samples used in this study, Genbank samples from *G. dehaani* from Hachijyo Island (LC743147, LC743148), *G. dehaani* from Kuchinoshima Island (LC743174)[17], *G. dehaani* from Naju, South Korea (MG674171)[46], and *G. sakamotoana* from Takara Island (LC743301, LC743302)[17] were included in the analysis. To assess genetic diversity and neutrality, haplotype diversity (*h*), nucleotide diversity (*π*)[47], Tajima's *D* test[48], and Fu's *Fs* test[49] were calculated for each population using Arlequin version 3.5.2.2[50]. Genetic differentiation among populations was examined using the pairwise fixation index (F_{ST}) and F_{ST} *p*-value[51]. A network diagram was created using the neighbor-net method[52] with SplitsTree4[53].

SNP analysis

Quality-filtered sequence data were used to identify SNPs using Stacks 2.41[54]. The Stacks pipeline denovo_map. pl was implemented with default values, except where indicated. The values for ustacks and population analysis

were set based on Onuki & Fuke[55]. In ustacks, the minimum coverage depth (m) was set to 5 and the maximum distance between stacks (M) to 2. For population analysis, the min-samples-overall (R) was set to 0.5, the minimum minor allele count (min-mac) to 2, and the maximum observed heterozygosity (max-obs-het) to 0.75; only the first SNP per locus was used, based on Onuki & Fuke[55]. We calculated the Ho, He[47], π, and F_{IS}[56] using the fstats module in the populations program in Stacks 2.41. Pairwise population differentiation indices F_{ST} and F'_{ST}[57] were calculated with the packages *hierfstat* 0.5.11 and *mmod* 1.3.3 in R software 4.5.0[58], respectively, with 1,000 bootstrap replications and Bonferroni correction.

Genetic structure analysis was performed with ADMIXTURE version 1.3.0[59], using C set to 0.0001 and S set to time. Output data were visualized in Excel. All individuals included in the SNP analysis (set 0) were analyzed by dividing the population into three mtDNA clades on popmap. Five populations were identified based on set 0 and ADMIXTURE analysis was performed on six adjacent sets to assess gene flow (Table S5; Fig. S3). This approach examined whether individuals with genetic elements from the multiple populations detected in set 0 are affected by gene flow or by the false genetic flow caused by analysis involving many groups. In each set, we selected individuals presumed to have been affected by gene flow and pure individuals from two of the five groups identified in set 0, from adjacent regions and surrounding areas (Fig. S4). Individuals included in the set were treated as a single group on popmap. Set 1 included populations in the Kanto region (17 individuals from Ibaraki, Tochigi, Gunma, Saitama, Tokyo, Kanagawa, and Shizuoka prefectures). Set 2 covered the Kinki region (16 individuals from Shiga, Mie, Nara, and Wakayama prefectures) and set 3 the Chugoku region (16 individuals from Kyoto, Hyogo, Tottori, Okayama, Shimane, Hiroshima, and Yamaguchi Prefectures), while set 4 included populations in the Shikoku region (21 individuals from Awaji Island-Hyogo Prefecture, Kagawa, Tokushima, Kochi, and Ehime Prefectures), set 5 represented northern Kyushu (31 individuals from Fukuoka, Saga, Oita, Nagasaki, and Kumamoto Prefectures), and set 6 included populations in southern Kyushu (39 individuals from Saga, Nagasaki, Kumamoto, Miyazaki, and Kagoshima Prefectures).

Estimation of the population demography history

To investigate the history of genetic diversification, we conducted a coalescent analysis using ABC with DIYABC Random Forest version 2.1[60]. Based on the results of our ADMIXTURE analysis (see results), we classified the populations into five groups: pop1 (Pacific coast from Boso Peninsula to Izu Peninsula + southern Kyushu and surrounding islands, hereafter referred to as the sKK population), pop2 (southern Shikoku + part of Kii Peninsula, SHI population), pop3 (central Kyushu, cK population), pop4 (northern Kyushu + Chugoku region of Honshu, nKC population), and pop5 (Hokkaido + Honshu from Tohoku to Chugoku + part of Shikoku, HO population). As ABC frameworks require populations without continuous gene flow, we focused the analysis on 118 specimens likely descending from a single ancestral population, based on the ADMIXTURE results (see results in detail). We used a Q-value threshold (< 0.95) often used to classify one genetic class (e.g[61]). In addition, we prepared a dataset with no missing values for DIYABC analysis, using 313 SNP loci. We used Hudson's algorithm[62] to calculate the minor allele frequency. Assuming a generation time of four years for *G. dehaani*[29] and an Early Pleistocene age (2 Mya) for the ancestral population's initial colony formation[63], we set the effective population sizes and generation parameters a priori at 100,000–500,000 (Table S7). All prior values were drawn from uniform distributions, under the condition t1 < t2 < t3 < t4 are generation time for merging of populations.

For scenario selection, the DIYABC assigns a vote for each scenario, reflecting how often it was chosen across a forest of n trees. The scenario with the highest vote count is considered the best fit for the dataset, alongside an estimate of the posterior probability for that scenario. A training set comprising 96,000 simulations and 500 trees was used to identify the most supported model scenario. Prior to that, we performed pre-scenario checks using PCA to detect potential model misspecification by comparing the prior and posterior parameter distributions.

DIYABC-RF allows for simultaneous comparison of many alternative scenarios, because the Random-Forest classifier implemented in DIYABC-RF is designed for multi-class model choice and remains accurate even when analyzing dozens of competing scenarios at once[60]. In this study, we evaluated a total of 32 scenarios across eight groups, in a single run (Figure S5): 1, scenarios based on body color (i.e., BL vs. DA, scenarios 1–3); 2, scenarios based on geographic separation areas at first (i.e., Honshu vs. Kyushu, scenarios 4–7); 3–7, scenarios where each population diverged sequentially (i.e., from sKK to HO, scenarios 8–11, 12–15, 16–19, 20–23, and 24–28, respectively); and 8, scenarios involving nearly simultaneous multiple divergences (scenarios 29–32). Scenario groups 1 and 2 were derived from earlier hypotheses about geological (island) isolation and body color variation[13–15,17,27]. Furthermore, we added the scenario group 3–8 to obtain a detailed divergence order among all populations. Although we preliminary analyzed it by a two-step approach (i.e., DIYABC for each scenario group and a final DIYABC combining the best scenario for each group), the results were almost same as with the single approach. For conciseness, we present the single approach in this study.

Data availability

Raw data about collection localities of the G. dehaani species complex are available in supplementary information.

References

1. Avise, J. C. Molecular population structure and the biogeographic history of a regional fauna: A case history with lessons for conservation biology. *Oikos* **1992**, 62–76 (1992).
2. Haag, C. R., Riek, M., Hottinger, J. W., Pajunen, V. I. & Ebert, D. Founder events as determinants of within-island and among-island genetic structure of *Daphnia* metapopulations. *Heredity* **96**, 150–158 (2006).

3. Banks, S. C. et al. Genetic structure of a recent climate change-driven range extension. *Mol. Ecol.* **19**, 2011–2024 (2010).
4. Kobayashi, S. Distribution pattern and ecology of brachyuran crabs in the riverine environment: their significance in the ecosystem and present condition. *Ecol. Civil Eng.* **3**, 113–130 (2000). [in Japanese].
5. Shih, H. T., Hung, H. C., Schubart, C. D., Chen, C. A. & Chang, H. W. Intraspecific genetic diversity of the endemic freshwater crab *Candidiopotamon rathbunae* (Decapoda, Brachyura, Potamidae) reflects five million years of the geological history of Taiwan. *J. Biogeogr.* **33**, 980–989 (2006).
6. Shih, H. T. & Ng, P. K. Diversity and biogeography of freshwater crabs (Crustacea: Brachyura: Potamidae, Gecarcinucidae) from East Asia. *Syst. Biodivers.* **9**, 1–16 (2011).
7. Shih, H. T., Ng, P. K. & Chang, H. W. Systematics of the genus *Geothelphusa* (Crustacea, Decapoda, Brachyura, Potamidae) from Southern taiwan: a molecular appraisal. *Zool. Stud.* **43**, 561–570 (2004).
8. Shih, H. T., Ng, P. K., Schubart, C. D. & Chang, H. W. Phylogeny and phylogeography of the genus *Geothelphusa* (Crustacea: Decapoda, Brachyura, Potamidae) in Southwestern Taiwan based on two mitochondrial genes. *Zool. Sci.* **24**, 57–66 (2007).
9. Naruse, T. & Ng, P. K. The identity of the Japanese freshwater crab, *Geothelphusa dehaani* (White, 1847) (Decapoda: brachyura: Potamidae), with descriptions of two new species from Japan. *Crustac Res.* **53**, 53–71 (2024).
10. Toyota, K. & Seki, S. *Picture Book of Japanese Freshwater/brackish Water Shrimp and Crab* (Midorishobo, 2020).
11. Sugime, R., Okuzaki, Y., Furuse, K., Kishida, O. & Naruse, T. First specimen-based record of the freshwater crab *Geothelphusa dehaani* (White, 1847) (Decapoda: Brachyura: Potamidae) in Hokkaido, Northern Japan. *Crustacean Res.* **51**, 39–45 (2022).
12. Kobayashi, S., Yamada, K., Yoshino, K., Henmi, Y. & Archdale, M. V. Occurrence of the Japanese Potamid crab *Geothelphusa dehaani* in tidal rivers flowing into Ariake sea. *Crustac Res.* **53**, 159–164 (2024).
13. Nakajima, K. & Masuda, T. Identification of local populations of freshwater crab *Geothelphusa dehaani* (White). *Bull. Jpn Soc. Sci. Fish.* **51**, 175–181 (1985).
14. Aotsuka, T., Suzuki, T., Moriya, T. & Inaba, A. Genetic differentiation in Japanese freshwater crab, *Geothelphusa dehaani* (White): isozyme variation among natural populations in Kanagawa Prefecture and Tokyo. *Zool. Sci.* **12**, 427–434 (1995).
15. Ikeda, M., Suzuki, T. & Fujio, Y. Genetic differentiation among populations of Japanese freshwater crab, *Geotheiphusa dehaani* (White), with reference to the body color variation. *Benthos Res.* **53**, 47–52 (1998).
16. Okano, T., Suzuki, H., Hiwatashi, Y., Nagoshi, F. & Miura, T. Genetic divergence among local populations of the Japanese freshwater crab *Geothelphusa dehaani* (Decapoda, Brachyura, Potamidae) from Southern kyushu, Japan. *J. Crust Biol.* **20**, 759–768 (2000).
17. Takenaka, M., Yano, K. & Tojo, K. Phylogeography of the true freshwater crab, *Geothelphusa dehaani*: detected dual dispersal routes via land and sea. *Zool.* **160**, 126118 (2023).
18. Araki, Y. & Sota, T. Population genetic structure underlying the geographic variation in beetle structural colour with multiple transition zones. *Mol. Ecol.* **30**, 670–684 (2021).
19. Tominaga, A., Yoshikawa, N., Matsui, M., Nagata, N. & Sato, Y. The emergence of a cryptic lineage and cytonuclear discordance through past hybridization in the Japanese fire-bellied newt, *Cynops pyrrhogaster* (Amphibia: Urodela). *Biol. J. Linn. Soc.* **137**, 651–666 (2022).
20. Toews, D. P. & Brelsford, A. The biogeography of mitochondrial and nuclear discordance in animals. *Mol. Ecol.* **21**, 3907–3930 (2012).
21. Chokki, H. Preliminary report of the colouration of freshwater crab, *Geothelphusa dehaani* (White), with special reference to its distribution. *Res. Crustac.* **7**, 177–182 (1976). [**in Japanese with English abstract and captions**].
22. Chokki, H. Notes on the colouration of freshwater crab, *Geothelphusa dehaani* (White), in Northern districts of Japan. *Res. Crustac.* **10**, 57–60 (1980). [**in Japanese with English abstract and captions**].
23. Suzuki, H. & Tsuda, E. Study on the color variation and distribution of a freshwater crab, *Geothelphusa dehaani* (White), in Kagoshima Prefecture. *Benthos Res.* **1991**, 37-46 (1991).
24. Furuya, Y. & Yamaoka, J. *[Enigma of Blue Freshwater Crab]. 116 pp + 1 Map* (Minamino-kaze, 2017). [**in Japanese**].
25. Kunishima, T., Takata, K. & Matsuno, S. BL type *Geothelphusa dehaani* (Decapoda, Potamidae) from Southern Wakayama Prefecture representing the first specimens-based record from Kii Peninsula, Japan. *Bull. Biogeogr. Soc. Jpn.* **75**, 120–122 (2020). [**in Japanese abstract and figure caption**].
26. Endler, J. A. & Mappes, J. The current and future state of animal coloration research. *Philos. Trans. R Soc. Lond. B Biol. Sci.* **372**, 20160352 (2017).
27. De la Cruz Huervana, J. J., Kano, Y., Ando, D., Onikura, N. & Kurita, Y. Preliminary assessment of genetic variation in the Japanese endemic freshwater crab, *Geothelphusa dehaani*, based on mitochondrial DNA sequences. *Biodivers. Data J.* **11**, e97438 (2023).
28. Helyar, S. J. et al. Application of SNPs for population genetics of nonmodel organisms: new opportunities and challenges. *Mol. Ecol. Resour.* **11**, 123–136 (2011).
29. Araki, A. & Matsuura, S. Growth of a freshwater crab, *Geothelphusa dehaani* (White). *Sci. Bull. Fac. Agr. Kyushu Univ.* **49**, 125–132 (1995).
30. Hayashi, F., Dobata, S. & Futahashi, R. Disturbed population genetics: suspected introgressive hybridization between two *Mnais* damselfly species (Odonata). *Zool. Sci.* **22**, 869–881 (2005).
31. Matsui, M. et al. Systematics of the widely distributed Japanese clouded salamander, *Hynobius nebulosus* (Amphibia: Caudata: Hynobiidae), and its closest relatives. *Curr. Herpetol.* **38**, 32–90 (2019).
32. Kawai, K. et al. Mountain masses as a determinant to genetic structure of Char distributed in the Chugoku region, Japan. *Environ. Biol. Fishes.* **107**, 75–90 (2024).
33. Tominaga, K., Nagata, N., Kitamura, J. I., Watanabe, K. & Sota, T. Phylogeography of the Bitterling *Tanakia lanceolata* (Teleostei: Cyprinidae) in Japan inferred from mitochondrial cytochrome *b* gene sequences. *Ichthyol. Res.* **67**, 105–116 (2020).
34. Machida, H., Ota, Y., Kawana, T., Moriwaki, H. & Nagaoka, S. (eds) *Regional Geomorphology of the Japanese Islands* Vol. 7 (Univ. Tokyo, 2001). Geomorphology of Kyushu and the Ryukyus [**in Japanese**].
35. Tominaga, A. et al. Phylogeny and historical demography of *Cynops pyrrhogaster* (Amphibia: Urodela): Taxonomic relationships and distributional changes associated with climatic oscillations. *Mol. Phylogenet Evol.* **66**, 654–667 (2013).
36. Wielstra, B., Burke, T., Butlin, R. K. & Arntzen, J. W. A signature of dynamic biogeography: enclaves indicate past species replacement. *Proc. R. Soc. B: Biol. Sci* **284**, 20172014 (2017).
37. Wilkens, J. L. & Fingerman, M. Heat tolerance and temperature relationships of the fiddler crab, *Uca pugilator*, with reference to body coloration. *Biol. Bull.* **128**, 133–141 (1965).
38. Matsuno, T., Wakasa, Y. & Okubo, M. Carotenoids of a river crab *Potamon dehaani*. *Bull. Japan Soc. Sci. Fish.* **48**, 661–666 (1982). [**in Japanese with English abstract**].
39. Takata, K., Fukuyama, I. & Fukuyama, R. First record of blue individual of *Ryukyum yaeyamense* (Minei, 1973) from Iriomote island, Japan. *Fauna Ryukyuana.* **64**, 57–59 (2022). [**in Japanese with English abstract and captions**].
40. Shimotsukasa, H. & Wada, K. Distribution of the freshwater crab *Geothelphusa dehaani* in relation with season, sex and body size. *Biology Inland. Waters.* **10**, 18–25 (1995). [**In Japanese with English abstract**].
41. Crump, A. et al. Guidelines for Protecting and Promoting Decapod Crustacean Welfare Research. Insect insectwelfare.com/research-guidelines (2024).
42. Geller, J., Meyer, C., Parker, M. & Hawk, H. Redesign of PCR primers for mitochondrial cytochrome c oxidase subunit I for marine invertebrates and application in all-taxa biotic surveys. *Mol. Ecol. Resour.* **13**, 851–861 (2013).
43. Tamura, K., Stecher, G. & Kumar, S. MEGA11: molecular evolutionary genetics analysis version 11. *Mol. Biol. Evol.* **38**, 3022–3027 (2021).

44. Suyama, Y. & Matsuki, Y. MIG-seq: an effective PCR based method for genome wide single nucleotide polymorphism genotyping using the next generation sequencing platform. *Sci. Rep.* **5**, 16963 (2015).
45. Chen, S., Zhou, Y., Chen, Y. & Gu, J. Fastp: an ultra fast all in one FASTQ preprocessor. *Bioinform.* **34**, i884–i890 (2018).
46. Lee, D. J., Lee, H. B. & Kim, I. C. The whole mitochondrial genome of the freshwater crab, *Geothelphusa* sp. (Stimpson, 1858) (Decapoda, Potamidae) and its phylogenetic relationship. *Mitochondr. DNA B* **4**, 1936–1937 (2019).
47. Nei, M. *Molecular Evolutionary Genetics* (Columbia University, 1987).
48. Tajima, F. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics* **123**, 585–595 (1989).
49. Fu, Y. X. Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. *Genetics* **147**, 915–925 (1997).
50. Excoffier, L. & Lischer, H. E. Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and windows. *Mol. Ecol. Resour.* **10**, 564–567 (2010).
51. Reynolds, J., Weir, B. S. & Cockerham, C. C. Estimation of the coancestry coefficient: basis for a short-term genetic distance. *Genetics* **105**, 767–779 (1983).
52. Bryant, D. & Moulton, V. Neighbor net: an agglomerative method for the construction of phylogenetic networks. *Mol. Biol. Evol.* **21**, 255–265 (2004).
53. Huson, D. H. & Bryant, D. Application of phylogenetic networks in evolutionary studies. *Mol. Biol. Evol.* **23**, 254–267 (2006).
54. Rochette, N. C., Rivera-Colón, A. G. & Catchen, J. M. Stacks 2: analytical methods for paired - end sequencing improve RADseq - based population genomics. *Mol. Ecol.* **28**, 4737–4754 (2019).
55. Onuki, K. & Fuke, Y. Rediscovery of a native freshwater shrimp, *Neocaridina denticulata*, and expansion of an invasive species in and around lake biwa, japan: genetic and morphological approach. *Conserv. Gen.* **23.5**, 967–980 (2022).
56. Wright, S. The interpretation of population structure by F-statistics with special regard to systems of mating. *Evolution* **19**, 395–420 (1965).
57. Meirmans, P. G. & Hedrick, P. W. Assessing population structure: FST and related measures. *Mol. Ecol. Resour.* **11**, 5–18 (2011).
58. R Core Team. *R: A Language and Environment for Statistical Computing* (R Foundation for Statistical Computing, 2025). https://www.R-project.org/.
59. Alexander, D. H., Novembre, J. & Lange, K. Fast model-based Estimation of ancestry in unrelated individuals. *Genome Res.* **19**, 1655–1664 (2009).
60. Collin, F. et al. Extending approximate bayesian computation with supervised machine learning to infer demographic history from genetic polymorphisms using DIYABC random forest. *Mol. Ecol. Resour.* **21**, 2598–2613 (2021).
61. Cordonnier, M., Gayet, T., Escarguel, G. & Kaufmann, B. From hybridization to introgression between two closely related sympatric ant species. *J. Zool. Syst. Evol. Res.* **57**, 778–788 (2019).
62. Hudson, R. Generating samples under a Wright- fisher neutral model of genetic variation. *Bioinform.* **18**, 337–338 (2002).
63. Shih, H. T., Ng, P. K., Naruse, T., Shokita, S. & Liu, M. Y. Pleistocene speciation of freshwater crabs (Crustacea: Potamidae: *Geothelphusa*) from northern Taiwan and southern Ryukyus, as revealed by phylogenetic relationships. *Zool. Anz.* **250**, 457–471 (2011).

Acknowledgements

Author contributions

T. K. and K.T. share to the first and corresponding authors, with equal to contribution. T.K. and K.T. designed the study, analyzed the data, visualized the results, and wrote the original manuscript. T.K., K.T., and K.N. collected the specimens. K.T., K.G., M.N. and A.I. conducted the molecular analyses. All authors reviewed the draft and approved the final manuscript.

Funding

Declarations

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-025-08085-0.

Correspondence and requests for materials should be addressed to T.K. or K.T.

Reprints and permissions information is available at www.nature.com/reprints.