

Title: Ecosystem size predicts the probability of speciation in migratory freshwater fish

Short title: Ecosystem size predicts speciation

Authors: Yo Y. YAMASAKI^{1,2*}, Hirohiko TAKESHIMA^{3,4}, Yuichi KANO⁵, Naoharu OSEKO⁶, Toshiyuki SUZUKI⁷, Mutsumi NISHIDA⁸, Katsutoshi WATANABE^{1*}

Author affiliations:

¹ Graduate School of Science, Kyoto University, Kitashirakawa-oiwake-cho, Sakyo, Kyoto 606-8502, Japan

² Ecological Genetics Laboratory, Department of Genomics and Evolutionary Biology, National Institute of Genetics, Yata 1111, Mishima, Shizuoka, 411-8540, Japan

³ Research Institute for Humanity and Nature, 457-4 Motoyama, Kamigamo, Kita, Kyoto 603-8047, Japan

⁴ Department of Marine Biology, Tokai University, Shimizu 3-20-2, Shizuoka 424-0902, Japan

⁵ Graduate education and research training program in Decision Science for a Sustainable Society, Kyushu University, 744, Motooka, Nishi, Fukuoka 819-0395, Japan

⁶ 3-18-5-203, Nagata Ginowan, Okinawa 901-2212, Japan

⁷ Kawanishi-midoridai Senior High School, 1-8 Kouyoudai, Kawanishi, Hyogo 666-0115, Japan

⁸ University of the Ryukyus, 1 Senbaru, Nishihara-cho, Nakagami-gun, Okinawa 903-0213, Japan

* Corresponding authors:

Yo Y. Yamasaki: yo.yamasaki@terra.zool.kyoto-u.ac.jp

Katsutoshi Watanabe: watanak@terra.zool.kyoto-u.ac.jp

ABSTRACT

Predicting speciation is a fundamental goal of research in evolutionary ecology. The probability of speciation is often positively correlated with ecosystem size. Although the mechanisms driving this correlation are generally difficult to identify, a shared geographic and ecological context provides a suitable condition to study the mechanisms that promote speciation in large ecosystems by reducing the number of factors to be considered. Here, we determined the correlation between speciation and ecosystem size, and discussed the underlying mechanisms of this relationship, using a probable parallel ecotype formation for freshwater fish. Our population genetic analysis revealed that speciation of the landlocked goby, *Rhinogobius* sp. YB, of the Ryukyu Archipelago, Japan, from its migratory ancestor, *R. brunneus*, occurred in parallel across five islands. Logistic regression analysis showed that speciation probability could be predicted using island size. The results suggest that ecosystem size predicts the occurrence of adaptation and reproductive isolation, likely through its association with three possible factors: divergent selection strength, population persistence, and occurrence probability of habitat separation.

Key words

ecological speciation, parallel evolution, island biogeography, approximate Bayesian computation, speciation-area relationship, *Rhinogobius*

1 | INTRODUCTION

Understanding the patterns and mechanisms of speciation is the fundamental goal of evolutionary ecology (Coyne & Orr, 2004; Reznick & Ricklefs, 2009). As macroevolutionary patterns of speciation, such as differences in diversification rate among lineages or regions, are caused by the accumulation of individual speciation events, the patterns are expected to be strongly influenced by speciation mechanisms (i.e., evolution of reproductive isolation). However, the relationships between macropatterns and the underlying mechanisms are not well understood, because they are usually investigated separately (Rosenblum et al., 2012; Cutter & Gray, 2016; Rabosky, 2016; but see Rabosky & Matute, 2013). Explicit consideration of the speciation mechanism would be helpful for obtaining a causal explanation of how macroevolutionary patterns of speciation have been generated (Harvey et al., 2017; Kisel et al., 2012).

Ecosystem size, often represented by habitat area, has long been considered as one of the best predictors for variation in species richness (MacArthur & Wilson, 1967). Recent studies on species–area relationships also recognize “in situ speciation” as an important factor, in addition to immigration and extinction rates (Losos & Parent, 2009; Losos & Schluter, 2000; Wagner, Harmon, & Seehausen, 2014; Weigelt, Steinbauer, Cabral, & Kreft, 2016). The positive correlation between ecosystem size and speciation rate is supported by several empirical (Kisel & Barraclough, 2010; Parent & Crespi, 2006) and theoretical (Gavrilets & Vose, 2005; Rosindell & Phillimore, 2011) studies.

Several hypotheses have been proposed that explain how speciation rate, or speciation probability, is affected by ecosystem size through speciation processes. For example, the opportunity for geographical isolation, which impedes gene flow, is likely to increase in larger areas (Kisel & Barraclough, 2010; Losos & Schluter, 2000). As larger areas contain heterogeneous environments, this phenomenon might also promote diversification to new environments following speciation (Parent & Crespi, 2006). Moreover, larger areas allow larger population sizes and longer persistence of newly emerged species (Kisel, McInnes, Toomey, & Orme, 2011). However, these hypotheses have been rarely tested empirically, because studies have usually assessed speciation by considering various mechanisms simultaneously. Thus, it is difficult to evaluate the influence of different factors on the speciation process. This problem could be addressed by targeting a group with parallel

diversification under the same mechanism (Lucek, Kristjánsson, Skúlason, & Seehausen, 2016; Vamosi, 2003). However, when using such an approach, it is important to first obtain evidence of repeated evolution under the same mechanism as the focal system.

Due to clear boundaries and replications, freshwater invasion by migratory fish in island habitats is a suitable case for testing the relationship between ecosystem size and occurrence of speciation. The East Asian goby, genus *Rhinogobius*, is one such fish group, in which the amphidromous (migratory) species *R. brunneus* might have repeatedly produced strictly freshwater populations (*Rhinogobius* sp. YB sensu Akihito, Sakamoto, Ikeda, & Aizawa, 2013) in the Ryukyu Archipelago of Japan (Kano, Nishida, & Nakajima, 2012; Nishida, 2001; Ohara, Takagi, Hashimoto, Miyazaki, & Hirashima, 2008). However, the number of times and geographical scales (i.e., islands, rivers, etc.) of the origin of the freshwater form of this goby have not yet been elucidated. Furthermore, the extent and factors of reproductive isolation between the amphidromous and freshwater forms have not been investigated.

We propose that the freshwater colonization of the goby in the Ryukyu Archipelago is a good model system for testing the correlation between ecosystem size and speciation probability for the following reasons. First, the ancestral amphidromous populations might have shared common evolutionary potential, because of possible gene flow through the sea during the larval period. Second, ecosystem age is similar among islands. According to geographical data, most islands in the Ryukyu Archipelago formed nearly simultaneously ~1.5 million years ago by the opening of straits between the islands (Osozawa et al., 2012). Third, ecological conditions are probably similar among the islands, because almost no freshwater fish species, except those of *Rhinogobius*, occur in the upper reaches of the rivers in the Ryukyu Archipelago (Kano et al., 2012; Yoshigou, 2014), resulting in few competitors and predators, which are important components of ecological opportunity (Wellborn & Langerhans, 2014; Yoder et al., 2010).

Using the simple system of *Rhinogobius* gobies, we primarily aimed to clarify whether the speciation probability of the freshwater form can be explained by ecosystem size. We first confirmed the existence of genetic isolation between the amphidromous and freshwater forms inhabiting the same islands or rivers. Next, we estimated where and how many times the freshwater forms originated, based on statistically explicit tests among demographic models that hypothesized parallel or single origin scenarios. We then examined whether speciation

probability could be predicted by ecosystem size. The correlation between independent origins of the freshwater form inferred from demographic modelling and ecosystem size (e.g., island area) was tested statistically (see Figure 1, concept diagram representing our approach). On the basis of these results, we discuss the mechanisms producing the relationships between ecosystem size and the speciation process in this system.

2 | MATERIALS AND METHODS

2.1 | Study system

The freshwater resident form derived from the amphidromous species *Rhinogobius brunneus* is recognized as an independent species but has not yet been described because of taxonomic confusion. Previous studies (e.g., Akihito et al., 2013) tentatively used the name “*Rhinogobius* sp. YB” for this freshwater form. However, until its evolutionary status is clarified, we will use the code “FF” (freshwater form) for *Rhinogobius* sp. YB and “AF” (amphidromous form) for *R. brunneus* based on their life history modes.

The FF completes its life in freshwater; conversely, AF larvae hatch in rivers and immediately drift down to the sea and return to rivers after 2–3 months (Mizuno, 2001). The eggs and larvae of FF are larger than those of AF (Supporting Information methods and results S1). These traits are assumed to be an adaptation to the river environment, where small food items (i.e., plankton) are scarce and higher level of swimming ability in river currents is needed (Hirashima & Tachihara, 2000; Nishida, 2001; Shinomiya, Sasabe, Sakurai, & Kishino, 2005; Yamasaki, Nishida, Suzuki, Mukai, & Watanabe, 2015).

Although the distribution of the two species sometimes overlapped in a single river, their distributions were generally different. AF mainly lived in the middle to lower reaches of rivers. In fact, AF populations are found only in rivers shorter than 18 km in the central Honshu region of Japan (Tamada, 2005). The larvae of AF must rapidly reach the sea to avoid starvation because of the scarcity of small food in rivers (Iguchi & Mizuno, 1999; Moriyama, Yanagisawa, Mizuno, & Omori, 1998). Thus, they might not be able to reproduce effectively in areas far from the sea (Iguchi & Mizuno, 1999). By contrast, FF mainly live in the upper reaches of rivers, and are often found in the upper areas of waterfalls where no other fish species are distributed (Kano et al., 2012).

2.2 | Sampling

Samples were collected on seven islands where populations of FF have been recorded previously (locality code 2. Amami-Oshima Is., 4. Tokunoshima Is., 5. Okinoerabujima Is., 6. Okinawajima Is., and 7. Kumejima Is. in Middle Ryukyu; 8. Ishigakijima Is. and 9. Iriomotejima Is. in South Ryukyu; Table S1; Figure 2; Ohara et al., 2008; Yoshigo, 2014). In addition, AF specimens were collected from three other islands (Figure 2). The specimens were identified by body colour patterns when fresh (Kano et al., 2012; Suzuki, Shibukawa, Yano, & Senou, 2004). Identification based on body colour was in agreement with the results of the genetic analysis, but some samples were difficult to identify or showed inconsistent genetic characteristics. When genetic data suggested that the focal specimens were not judged as hybrids based on the criteria described below and that incorrect identification based on morphology in the field was likely (see the next section and Supporting methods and results S1), these samples were removed from all subsequent analyses. If a sample was judged to be a hybrid, we tentatively named it based on morphology. As a result, 842 FF and 578 AF specimens from 52 points in 42 river systems were included in the analysis. Both forms were collected from nine river systems; they were sympatrically collected at seven points.

2.3 | Genetic data and population structure

Genetic data of *Rhinogobius* specimens were obtained by genotyping 20 nuclear microsatellite loci (Table S2; Ohara, Takahashi, & Takagi, 2004). Methods for molecular experiments and genotyping are described in the Supporting methods and results S1. Genotype data were deposited in Dryad (<https://doi.org/10.5061/dryad.mpg4f4qvh>). To assess the genetic features of the respective populations, we calculated the expected and observed heterozygosity (H_E and H_O , respectively) and allelic richness (R_s) for the respective populations. Pairwise F_{ST} and Jost's D were also calculated. We conducted analysis of molecular variance (AMOVA; Excoffier, Smouse, & Quattro, 1992) separately for AF and FF, to assess hierarchical genetic differentiation. To visualize relatedness among populations, we constructed a neighbour-joining tree among populations based on Nei's Da distance (Nei, Tajima, & Tateno, 1983) using poptree2 (Takezaki, Nei, & Tamura, 2010). The credibility of each clade was evaluated using 1,000 bootstrap resamplings.

To infer the existence and extent of gene flow and genetic isolation, we analysed the genetic structure of the population using structure version 2.3.4 (Lawson, van Dorp, & Falush, 2018; Pritchard, Stephen, & Donnelly, 2000). We prepared two data sets: one for both FF and AF populations from each island, and another for only AF populations from all islands. Evanno's ΔK (Earl & von Holdt, 2012; Evanno, Regnaut, & Goudet, 2005) was used as a reference for K value selection. Hybrid individuals were evaluated using q values; we arbitrarily judged specimens as hybrids (or individuals with hybrid origin) in cases where minor genetic components accounted for 10% or more of the aggregate (Vähä & Primmer, 2006). Detailed methods are provided in the Supporting methods and results S1.

2.4 | Testing the parallel formation of FF by ABC

Before testing the relationship between ecosystem size and speciation probability, we needed to infer the islands where speciation of FF occurred. To examine the single or parallel origins of FF statistically, we conducted a demographic model selection under the Approximate Bayesian Computation (ABC) framework (Beaumont, 2010). We briefly describe the procedures of model construction and model selection in this section (for details, see Supporting methods and results S1).

Our general model construction strategies for ABC analysis were as follows. First, we generated separate models using three geographical scales (i.e. within-islands, regional [Middle and South Ryukyu], and archipelago scales) and examined FF origins in each geographical scale to reduce the number of models and populations examined in a single analysis. Second, we examined both models with and without gene flow between AF and FF populations. This is because a simulation study suggests that gene flow could alter apparent phylogenetic relationships among populations when simple genetic distance was used (Bierne, Gagnaire, & David, 2013). Demographic model comparisons that consider gene flow can overcome this problem (Butlin et al., 2014).

In the within-islands analyses, we mainly examined two questions: whether FF had single or parallel origin, and whether gene flow between AF and FF existed on each island. We hypothesized several plausible divergence orders between populations based on the result of structure analysis, phylogeny and geographical locations for respective islands. Thereafter, we compared the models with single/parallel origins and with/without gene flow. Analyses were

conducted for the five islands where FF is distributed in multiple rivers (i.e., Amami-Oshima Is., Tokunoshima Is., Okinawajima Is., Kumejima Is. and Iriomotejima Is.). We selected three (or two) river populations for each island to reduce the model complexities. Criteria for population selection are described in the Supporting methods and results S1. The constructed models (a total of 34 models) are described in Figure S1a–e. Summary statistics are given in Table S3.

For the among-islands within-region analyses, we constructed several plausible models separately for the regions of Tokunoshima Is. and Okinoerabujima Is., Middle Ryukyu, and South Ryukyu. Five to 10 models were constructed on the basis of the phylogenetic tree and geographical distance between islands. Finally, in the between-region analyses, we compared 10 models generated from the results of the previous analyses. All constructed models are described in the Supporting methods and results S1 and Figure S1f–I. Scripts that describe the models have been deposited in Dryad (<https://doi.org/10.5061/dryad.mpg4f4qvh>).

For model selection, we chose the model that had the highest posterior probability in the respective analyses. To evaluate the accuracy of model selection, we calculated a misclassification rate (i.e., the rate at which the focal model was selected despite the other model being true [type 2 error rate]).

2.5 | Correlation between ecosystem size and speciation probability

We first identified the islands on which the speciation of FF occurred based on the results of structure (genetic isolation) and ABC model selection (independent origin). The relationship between ecosystem size and speciation was analysed for the Middle and South Ryukyu islands. This is because these islands are assumed to be similar in age, were never connected to the continent after separation, and were not catastrophically affected by volcanic activity (Osozawa et al., 2012; Supporting methods and results S1). We acquired distribution information of the two forms from published literature (Yoshigou, 2014) and our own field survey, and we only chose the islands where either or both forms were distributed. This is because we assumed that speciation was not likely to have occurred on the island where either of these species is not currently distributed. We coded the occurrence of speciation in each island as follows: 1 for the island where speciation occurred, and 0 for the island where speciation did not occur. As a proxy of ecosystem size, we used island area (km²; data from

the Geospatial Information Authority of Japan), maximum catchment area (km²), maximum river length (m) or the maximum number of waterfalls on each island (Supporting methods and results S1).

We performed one-parameter logistic regression analysis in the generalized linear model framework with binomial error structure for the response variable (speciation: 1, presence; 0, absence) using the glm package of r version 3.1.1 (R Core Team, 2014). Due to a strong positive correlation between the four variables (Pearson's correlation coefficient, 0.592–0.993, Supporting methods and results S1), we primarily used island area as the explanatory variable for the main result. Explanatory variables, except the number of waterfalls, were log₁₀-transformed. To test the significance of explanatory variables, we conducted a likelihood ratio test using 10,000 rounds of parametric bootstrap sampling with a constant model as a null model that assumes no relationship between ecosystem variables and speciation probability. We calculated McFadden's pseudo R^2 .

Furthermore, to examine the effect of the ecosystem size parameters other than island size (i.e., maximum catchment area, river length and number of waterfalls), we constructed another generalized linear model for speciation probability. We used the residuals from a regression between island size and each of the other ecosystem size parameters, together with the island size and their interaction, as explanatory variables. Stepwise model selections by Akaike information criterion (AIC) scores were conducted using the *step* function on r.

3 | RESULTS

3.1 | Basic summary statistics

Heterozygosity and allelic richness were generally higher in the AF populations than in the FF populations (Table S1, Figure S2a, b). The AF populations showed relatively little pairwise genetic differentiation within islands and among islands within regions (Tables S4–S6, Figure S2c, d). In comparison, the genetic differentiation of AF populations between regions was large (Tables S4–S6, Figure S2c, d). The FF populations often showed large genetic differentiation within islands and among islands within regions, as well as among regions (Tables S4–S6, Figure S2c, d).

3.2 | Genetic isolation and hybrids

In the AF data set, Evanno's ΔK showed the maximum value at $K = 2$; the populations in North plus Middle Ryukyu and South Ryukyu were clearly separated. In the data set for each island, including FF and AF populations, ΔK had a maximum value at $K = 2$ for all islands, except one (Amami-Oshima Is.). The two genetic clusters corresponded to the two forms, except on Okinawajima Is. (see below; Figure 3; Figure S3).

Suspected hybrid individuals were rare (2.0% of the total individuals). No suspected hybrid individuals were detected from five of the seven sympatric collection points. Several suspected hybrid individuals were detected in 11 FF and AF populations. Among these, all FF individuals that were identified based on morphology in the FFOK-4-1 and FFOK-4-2 populations (Okinawajima Is.; Figures 2 and 3; Figure S3) exhibited the genetic features of AF as the dominant component.

3.3 | Genetic relationships among populations

The neighbour-joining tree showed two divergent groups separated by the Kerama Gap in all of the FF and AF populations (bootstrap value, 97%; Figure 4). Each group included both forms, which we present as Group 1 (Honshu, North Ryukyu and Middle Ryukyu) and Group 2 (South Ryukyu), hereafter. In both groups, FF populations

on each island formed a monophyletic group (bootstrap value, 91%–100%), except those on two islands in Group 1 (Amami-Oshima Is., 67%; Okinawajima Is., 33%). In addition, FF populations on different islands (Tokunoshima Is. and Okinoerabujima Is.) formed a monophyletic group (89%). The AF populations within each group were closely related to each other, and the populations of each island in Group 1 tended to form a monophyletic group.

3.4 | Test for multiple origins of FF

In the within-island-scale analysis, the single origin model of FF was selected for all islands (Table 1 and Table S7; Figure 4 and Figure S4). Models including gene flow among all or partial populations of the two forms showed higher posterior probability than those not considering gene flow (Table 1 and Table S7; Figure 4 and Figure S4).

At the among-island scale, Group 1 in Middle Ryukyu and Group 2 in South Ryukyu were analysed separately. For Group 1, we first examined the single origin of FF in the adjoining islands, Tokunoshima Is. and Okinoerabujima Is., which was suggested by the phylogenetic analysis. ABC analysis supported their single origin (Table 1 and Table S7), and thus only FF on Tokunoshima Is. was used for subsequent analyses. The model selection for Group 1 strongly supported the four independent origins of FF (Table 1; Table S7). The model with gene flow between the two forms provided a higher Bayes factor (BF) than the model not considering gene flow (Table S7). In Group 2, the single origin of FF of Iriomotejima Is. and Ishigakijima Is. was supported (Table 1; Table S7). The model including gene flow between the AF and FF populations was better supported than that not considering gene flow (Table S7).

For the between-region scale, the neighbour-joining tree for populations suggested independent origins of FF in Groups 1 and 2. However, no conclusion could be drawn at this scale from ABC, because no single model was successfully selected, probably because of insufficient data (Table 1; Table S7).

3.5 | Correlation between speciation probability and ecosystem size

We confirmed the occurrence of AF or FF populations on 18 islands from the published literature and our field survey (Table S8). As we could not specify the islands where unique speciation occurred, we tested 20 alternative scenarios that hypothesized that speciation occurred in different combinations of the candidate islands (see Supporting methods and results S1, and Table S9 for details).

For all four data sets in the scenario of independent origins between the regions, speciation probability was significantly explained by island area (Table 2 and Table S9; Figure 5 and S5). The relationship was also significant for 12 of the 16 data sets in the nonindependent scenario (Table S9). Almost identical results were obtained for maximum catchment area and river length (Table 2 and Table S9; Figure S5). Furthermore, two of the four data sets for the independent origin scenario and five of the 16 data sets for the nonindependent scenario showed significant relationships between the number of waterfalls and speciation (Table 2 and Table S9). Model selection by AIC scores showed that the best model included no

variables, except for island area, as the predictors for speciation probability in most data sets when the variables were controlled by island area (Table S9).

4 | DISCUSSION

Our results revealed that FF repeatedly evolved from AF in the Ryukyu Archipelago, and that the occurrence of FF was highly correlated with island size. These results suggest that the speciation of FF was promoted in larger ecosystems. In the following sections, we discuss the mechanisms of speciation and the causes of correlation between ecosystem size and speciation of FF.

4.1 | Reproductive isolation between the two forms and its evolution

Genetic isolation between the two forms was suggested, even at sympatric sites (Figure 3 and Figure S3); the results support that FF and AF are reproductively isolated. This genetic isolation between AF and FF cannot be explained by isolation by distance (IBD), because geographically isolated FF populations in different river systems showed closer relationships than FF and AF in the same river. This relationship was consistent in the respective islands. Some behavioural mechanisms might be responsible for this reproductive isolation, as reported for other species pairs of *Rhinogobius* (Mizuno, 1987), which share common reproductive features with intersexual behavioural interactions (Takahashi & Kohda, 2004). Differences in the breeding season might also contribute to their current reproductive isolation (AF, January to May; FF, March to June; Hirashima, 2016; Katoh, 1996; Tachihara, 2009). Although the presence of pre- and post-zygotic isolations has not been tested for our studied pair of species and their position along a speciation continuum (Nosil, 2012) is not clear, genetic isolation, as well as distinct life history divergence, support that FF populations should be treated as different biological species from AF populations.

Although AF and FF currently exhibit genetic isolation, our ABC analysis suggested that reproductive isolation between these two ecologically divergent forms evolved under the existence of gene flow. This result suggests that the evolution of reproductive isolation was not caused simply by allopatry and mutation-order mechanisms but by divergent ecological adaptation between the AF and FF populations (Seehausen & Wagner, 2014).

What is the factor that led to the evolution of their reproductive isolation? Previous studies have focused on the differences in food size for larvae between seas and rivers (Nishida, 2001). AF larvae generally feed on small food items available in seas. If larvae cannot reach the sea rapidly, large larvae might be favoured by directional selection in rivers (Closs, Hicks, & Jellyman, 2013; Kishi, 1979; Kondo, Maeda, Hirashima, & Tachihara, 2013; Mizuno, 1963). A mathematical model has shown that speciation could occur due to differences in favourable egg and larval size between the upper and lower reaches (Omori et al., 2012). Such empirical and theoretical evidence suggests that the primary factor in the evolution of reproductive isolation is divergent selection between the optimal combinations of the within-rivers distribution and larval size.

The role of barriers in a river should also be emphasized, because they might promote adaptation to freshwater habitats through reduced gene flow and the invasion of predators and competitors. During repeated sea-level fluctuations during the Pleistocene, steep marine terraces of limestone or coral reef developed in the coastal areas of the Ryukyu Archipelago (Machida, Ota, Kawana, Moriwaki, & Nagaoka, 2001). Such terraces would form many waterfalls near river mouths. Nishida (2001) and Kano et al. (2012) hypothesized an important role of waterfall formation in the landlocking process of FF. The relationship between such barriers and divergent selection should be further investigated to understand the early processes of landlocking.

Despite the existence of current reproductive isolation, suspected hybrid individuals were occasionally detected. Some hybridization might be explained by artificial disturbances in the distribution and migration schemes of AF. Six of the 11 populations containing suspected hybrids were collected from the waters above artificial dams; they included FFOK-4-1 and FFOK-4-2 populations of FF, which primarily exhibited genetic features of AF. Artificial dams would promote hybridization in the following process.

Artificial dams have promoted the establishment of a landlocked AF population (Tachihara, 2009), because there are small food items in reservoirs that the larvae of AF and hybrids would be able to use. This would induce a sympatric distribution of the two forms and facilitate the survival of hybrids in the upper reaches of the dam. These cases might be examples of human-mediated reverse speciation (Grabenstein & Taylor, 2017; Taylor et al., 2006).

4.2 | Multiple origins of FF

The AF and FF populations showed large divergence between the North/Middle Ryukyu (Group 1) and South Ryukyu (Group 2). The AF populations were closely related among islands within each region (Figures 3 and 4), suggesting that gene flow by larval drifting is effective at the regional scale, but is prevented between regions by the Kerama Gap.

A series of ABC analyses supported the parallel derivation of FF populations on respective islands. The estimated number of times speciation occurred was four or five. The uncertainty in the number is attributed to the unresolved relationship among the FF populations across the Middle and South Ryukyus, which could not be resolved by ABC analysis, probably owing to the limitation of microsatellite markers, such as allele size homoplasy. However, the following reasons support that the FF populations originated independently in each region; that is, FF originated five times. First, the independent origins of FF populations among islands within regions were supported (Table 1; Figure 4). Second, the simple tree analysis supported the dichotomous relationship of the populations between the North/Middle and South Ryukyus. Third, the two regions are separated by a deep strait, the Kerama Gap, which is considered to have not emerged above water since the Ryukyu Archipelago separated from the continent (Osozawa et al., 2012). This geographical barrier has also been noted for many other terrestrial animals, such as amphibians and reptiles (Matsui, Ito, et al., 2005; Matsui, Shimada, et al., 2005; Ota, 1998; Tominaga, Matsui, Eto, & Ota, 2015).

Our results suggest that the multiple origins of FF were an example of parallel speciation, which is the phenomenon whereby reproductive isolation repeatedly evolves under the same mechanism (Schluter & Nagel, 1995). Nosil (2012) proposed four criteria to distinguish parallel speciation: nonmonophyletic pattern of an ecotype; nonmonophyly of ecologically similar forms representing multiple origins, rather than the effect of gene flow or incomplete lineage sorting; positive correlation between reproductive isolation and ecological divergence; and the genetic basis of reproductive isolation. The ABC analysis, as well as the egg and larval size difference between FF and AF, indicated that our study system satisfies the former two criteria. In addition, the latter two criteria were probably satisfied because of the clear differences in life history and reproductive traits between the two forms (Yamasaki et al., 2015), although they need to be confirmed by conducting mate choice and common garden experiments in the future.

4.3 | Correlation of ecosystem size and speciation probability

Our logistic regression analysis showed that speciation probability was significantly explained by island size, river length and catchment area in most of the data sets under the various analysis conditions. This result suggests that some mechanisms promoting the speciation process tend to be more effective in larger ecosystems. As the number of waterfalls also explained speciation probability in some data sets, it also partly supported that waterfalls promote speciation in the goby.

Why is the speciation probability of FF positively correlated with ecosystem size? To answer this question, understanding the relationships between ecosystem size and the following three components related to the mechanism of speciation is necessary: (a) establishment of reproductive isolation, (b) persistence of diverged populations and (c) enhancement of the former two components by population separation (Allmon, 1992; Mayer, 1963).

First, the establishment of reproductive isolation would be promoted in larger ecosystems, because the strength of divergent selection is expected to be positively correlated with ecosystem size. Divergent selection between flowing and staying larvae would become stronger in longer rivers (i.e., larger ecosystems), as the spawning sites of AF are restricted within a certain distance from the sea owing to the mortality of drifting larvae (see above; Moriyama et al., 1998; Omori et al., 2012). Furthermore, divergent selection might have been stronger during glacial periods, when sea level was lower, and hence rivers became longer (Kondo et al., 2013). Therefore, the landscape in the period of low sea level might have affected patterns of speciation. However, unfortunately, we could not test the correlation between speciation probability and island size at the low sea level, represented by the Last Glacial Maximum (i.e., when sea level was -122 m), because the 18 islands investigated were estimated to have been merged into eight, and the statistical power was too low to detect the correlation (data not shown). Although the correlation between the strength of divergent selection and river length should be tested in the future, larger ecosystems might enhance reproductive isolation via divergent selection, when the extent of the focal environmental gradient is positively related to ecosystem size. This relationship might generally lead to the development of a correlation between speciation probability and ecosystem size in other systems as well.

Second, the amount of available space might increase with ecosystem size, which could result in the greater persistence of freshwater-adapted populations through an increase in population size. In longer rivers, the freshwater population could use larger areas in the upper reaches of the rivers, because the optimal distance between the spawning area and sea for AF does not increase with river length. A larger population size might reduce the risk of demographic extinction. In addition, extinction risk due to fusion with the amphidromous population might be alleviated in the larger freshwater populations. Furthermore, the reinforcement of reproductive isolation might be more effectively developed when the population sizes of the two secondarily contacted species are similar (Liou & Price, 1994). A larger population size is also beneficial for retaining adaptive alleles (Gavrilets & Vose, 2005) and reducing maladaptive gene flow (Hanski, Mononen, & Ovaskainen, 2011).

Third, the probability of habitat separation between rivers and seas might be positively correlated with ecosystem size, promoting the previously described two processes during speciation. Correlation between the number of waterfalls and river length suggests the more frequent emergence of barriers to migration in longer rivers. Furthermore, greater river length might hinder migration to the upper reaches. Such habitat separation in longer rivers might be important for population divergence, because the separation would reduce gene flow between FF and AF populations, and would impede the entrance of other predatory and competitive fishes in the upper reaches.

5 | CONCLUSIONS

We have confirmed that speciation probability is explained by ecosystem size in our goby system. This result suggests positive correlations between ecosystem size and several components of the speciation process, including the strength of divergent selection, population persistence and population separation frequency. These correlations could promote the establishment of reproductive isolation and the persistence of new populations, although further explicit tests for these relationships are necessary. These correlations could also be important for understanding species diversification mechanisms at the phylogenetic level. Simple systems, such as our parallel species pairs, could contribute to the development of theories that link the mechanisms of each speciation to macroevolutionary patterns of speciation.

463

464 **ACKNOWLEDGMENTS**

465 We thank K. Nakao for collecting samples; S. Matsuzaki, N. Muto, S. Chiba, K. Nakayama
466 and S. Hirase for helping with the molecular experiment; and A. Iwata, M. Hosono, T. Fujisawa,
467 J. Kitano and members of Laboratory of Animal Ecology of Kyoto University for providing
468 valuable comments. We thank C. Riginos and three anonymous reviewers for providing
469 helpful comments. Computations were partially performed on the NIG supercomputer at
470 ROIS National Institute of Genetics. This study was partially supported by the Sasakawa
471 Scientific Grant (No. 27-507) and JSPS KAKENHI (Nos. 26250044, 16J08762, 17H03720).

472 **AUTHOR CONTRIBUTIONS**

473 Y.Y.Y., M.N. and K.W. planned the project. Y.Y.Y., Y.K., N.O. and T.S. performed field
474 surveys. Y.Y.Y. and H.T. conducted molecular experiments. Y.Y.Y. conducted population
475 genetic and statistical analysis. Y.Y.Y. and K.W. wrote the first draft manuscript, and all
476 authors contributed to the improvement of the manuscript.

477 **DATA AVAILABILITY STATEMENT**

478 The microsatellite genotype data and scripts of ABC analysis have been deposited in Dryad
479 (<https://doi.org/10.5061/dryad.mpg4f4qvh>). NGS data for developing microsatellite makers
480 have been deposited in DDBJ (accession no. DRA006388).

481

482 **ORCID**

483 *Yo Y. Yamasaki* <https://orcid.org/0000-0002-7495-2712>

484 *Katsutoshi Watanabe* <https://orcid.org/0000-0003-2244-2902>

485

486 **REFERENCES**

487 Akihito, Sakamoto, K., Ikeda, Y., & Aizawa, M. (2013). Gobioidae. In T. Nakabo (Ed.),
 488 *Fishes of Japan with pictorial keys to species*, 3rd ed. (pp. 1374–1608). Tokyo, Japan: Tokai
 489 University Press. (in Japanese).

490 Allmon, W. D. (1992). A causal analysis of stages in allopatric speciation. *Oxford Surveys in*
 491 *Evolutionary Biology*, 8, 219–257.

492 Beaumont, M. A. (2010). Approximate Bayesian computation in evolution and ecology.
 493 *Annual Review of Ecology, Evolution, and Systematics*, 41, 379–406.
 494 <https://doi.org/10.1146/annurev-ecolsys-102209-144621>

495 Bierne, N., Gagnaire, P. A., & David, P. (2013). The geography of introgression in a patchy
 496 environment and the thorn in the side of ecological speciation. *Current Zoology*, 59, 72–86.
 497 <https://doi.org/10.1093/czoolo/59.1.72>

498 Butlin, R. K., Saura, M., Charrier, G., Jackson, B., André, C., Caballero, A., ... Rolán-Alvarez,
 499 E. (2014). Adaptation and reproductive isolation in the face of gene flow. *Evolution*, 68, 935–
 500 949. <https://doi.org/10.1111/evo.12329>

501 Closs, G. P., Hicks, A. S., & Jellyman, P. G. (2013). Life histories of closely related
 502 amphidromous and non-migratory fish species: A trade-off between egg size and fecundity.
 503 *Freshwater Biology*, 58, 1162–1177. <https://doi.org/10.1111/fwb.12116>

504 Coyne, J. A., & Orr, H. A. (2004). *Speciation*. Sunderland, MA: Sinauer Associates.

505 Cutter, A. D., & Gray, J. C. (2016). Ephemeral ecological speciation and the latitudinal
 506 biodiversity gradient. *Evolution*, 70, 2171–2185. <https://doi.org/10.1111/evo.13030>

507 Earl, D. A., & vonHoldt, B. M. (2012). STRUCTURE HARVESTER: A website and program
 508 for visualizing STRUCTURE output and implementing the Evanno method. *Conservation*
 509 *Genetics Resources*, 4, 359–361. <https://doi.org/10.1007/s12686-011-9548-7>

510 Evanno, G., Regnaut, S., & Goudet, J. (2005). Detecting the number of clusters of individuals
 511 using the software STRUCTURE: A simulation study. *Molecular Ecology*, 14, 2611–2620.
 512 <https://doi.org/10.1111/j.1365-294X.2005.02553.x>

513 Excoffier, L., Smouse, P. E., & Quattro, J. M. (1992). Analysis of molecular variance inferred
514 from metric distances among DNA haplotypes: application to human mitochondrial DNA
515 restriction data. *Genetics*, 131, 479–491.

516 Gavrillets, S., & Vose, A. (2005). Dynamic patterns of adaptive radiation.
517 *Proceedings of the National Academy of Sciences of the United States of America*, 102,
518 18040–18045. <https://doi.org/10.1073/pnas.0506330102>

519 Grabenstein, K. C., & Taylor, S. A. (2017). Breaking barriers: Causes, consequences, and
520 experimental utility of human-mediated hybridization. *Trends in Ecology & Evolution*, 33,
521 198–212. <https://doi.org/10.1016/j.tree.2017.12.008>

522 Hanski, I., Mononen, T., & Ovaskainen, O. (2011). Eco-evolutionary metapopulation
523 dynamics and the spatial scale of adaptation. *American Naturalist*, 177, 29–43.
524 <https://doi.org/10.1086/657625>

525 Harvey, M. G., Seeholzer, G. F., Smith, B. T., Rabosky, D. L., Cuervo, A. M., & Brumfield,
526 R. T. (2017). Positive association between population genetic differentiation and speciation
527 rates in New World birds. *Proceedings of the National Academy of Sciences of the United*
528 *States of America*, 114, 6328–6333. <https://doi.org/10.1073/pnas.0506330102>

529 Hirashima, K. (2016). Variety of life history and strategy of *Rhinogobius* fishes. *Aquabiology*,
530 38, 363–369. (in Japanese with English abstract). Hirashima, K., & Tachihara, K. (2000).
531 Embryonic development and morphological changes in larvae and juveniles of two
532 land-locked gobies, *Rhinogobius* spp. (Gobiidae), on Okinawa Island. *Japanese Journal of*
533 *Ichthyology*, 47, 29–41. (in Japanese with English abstract). [https://](https://doi.org/10.11369/jji1950.47.29)
534 doi.org/10.11369/jji1950.47.29

535 Iguchi, K., & Mizuno, N. (1999). Early starvation limits survival in amphidromous fishes.
536 *Journal of Fish Biology*, 54, 705–712. <https://doi.org/10.1111/j.1095-8649.1999.tb02027.x>

537 Kano, Y., Nishida, S., & Nakajima, J. (2012). Waterfalls drive parallel evolution in a
538 freshwater goby. *Ecology and Evolution*, 2, 1805–1817. <https://doi.org/10.1002/ece3.295>

539 Katoh, M. (1996). Seasonal variation in gonadal activity of females among four species of
540 freshwater gobies in the *Rhinogobius brunneus* species complex in Okinawa, Japan.
541 *Ichthyological Research*, 43, 169–174.

542 Kisel, Y., & Barraclough, T. G. (2010). Speciation has a spatial scale that depends on levels
543 of gene flow. *American Naturalist*, 175, 316–334. <https://doi.org/10.1086/650369>

544 Kisel, Y., McInnes, L., Toomey, N. H., & Orme, C. D. L. (2011). How diversification rates
545 and diversity limits combine to create large-scale species–area relationships. *Philosophical*
546 *Transactions of the Royal Society B*, 366, 2514–2525. <https://doi.org/10.1098/rstb.2011.0022>

547 Kisel, Y., Moreno-Letelier, A. C., Bogarín, D., Powell, M. P., Chase, M. W., & Barraclough,
548 T. G. (2012). Testing the link between population genetic differentiation and clade
549 diversification in Costa Rican orchids. *Evolution*, 66, 3035–3052. [https://doi.](https://doi.org/10.1111/j.1558-5646.2012.01663.x)
550 [org/10.1111/j.1558-5646.2012.01663.x](https://doi.org/10.1111/j.1558-5646.2012.01663.x)

551 Kishi, Y. (1979). A geographical model of disruptive selection on offspring size and a
552 possible case of speciation in freshwater gobies characterized by egg-size difference.
553 *Researches on Population Ecology*, 20, 211–215.

554 Kondo, M., Maeda, K., Hirashima, K., & Tachihara, K. (2013). Comparative larval
555 development of three amphidromous *Rhinogobius* species, making reference to their habitat
556 preferences and migration biology. *Marine and Freshwater Research*, 64, 249–266.
557 <https://doi.org/10.1071/MF12234>

558 Lawson, D. J., van Dorp, L., & Falush, D. (2018). A tutorial on how not to over-interpret
559 STRUCTURE and ADMIXTURE bar plots. *Nature Communications*, 9, 3258.
560 <https://doi.org/10.1038/s41467-018-05257-7>

561 Liou, L., & Price, T. D. (1994). Speciation by reinforcement of premating isolation. *Evolution*,
562 48, 1451–1459.

563 Losos, J. B., & Parent, C. E. (2009). The speciation–area relationship. In J. B. Losos, & R. E.
564 Ricklefs (Eds.), *The theory of island biogeography revisited* (pp. 415–438). Princeton, NJ,
565 USA: Princeton University Press.

- 566 Losos, J. B., & Schluter, D. (2000). Analysis of an evolutionary species-area relationship.
567 *Nature*, 408, 847–850. <https://doi.org/10.1038/35048558>
- 568 Lucek, K., Kristjánsson, B. K., Skúlason, S., & Seehausen, O. (2016). Ecosystem size
569 matters: The dimensionality of intralacustrine diversification in Icelandic stickleback is
570 predicted by lake size. *Ecology and Evolution*, 6, 5256–5272. [https://doi.org/10.1002/](https://doi.org/10.1002/ece3.2239)
571 [ece3.2239](https://doi.org/10.1002/ece3.2239)
- 572 MacArthur, R. H., & Wilson, E. O. (1967). *The theory of island biogeography*. Princeton, NJ:
573 Princeton University Press.
- 574 Machida, H., Ota, Y., Kawana, T., Moriwaki, H., & Nagaoka, N. (2001). *Regional*
575 *geomorphology of the Japanese Islands, vol. 7 Kyushu and Southern Islands*. Tokyo, Japan:
576 University of Tokyo Press. (in Japanese).
- 577 Matsui, M., Ito, H., Shimada, T., Ota, H., Saidapur, S. K., Khonsue, W., ... Wu, G.-F. (2005).
578 Taxonomic relationships within the pan-oriental narrow-mouth toad *Microhyla ornata* as
579 revealed by mtDNA analysis (Amphibia, Anura, Microhylidae). *Zoological Science*, 22, 489–
580 495. <https://doi.org/10.2108/zsj.22.489>
- 581 Matsui, M., Shimada, T., Ota, H., & Tanaka-Ueno, T. (2005). Multiple invasions of the
582 Ryukyu Archipelago by Oriental frogs of the subgenus *Odorrana* with phylogenetic
583 reassessment of the related subgenera of the genus *Rana*. *Molecular Phylogenetics and*
584 *Evolution*, 37, 733–742. <https://doi.org/10.1016/j.ympev.2005.04.030>
- 585 Mayer, E. (1963). *Animal species and evolution*. Cambridge, MA, USA: Belknap Press.
- 586 Mizuno, N. (1963). Distribution of *Cottus japonicus* Okada (Cottidae) and *Tukugobius*
587 *flumineus* Mizuno (Gobiidae), with special references to their peculiarities in both the
588 land-locking and the speciation from amphidromous to ancestors. *Bulletin of Osaka Gakugei*
589 *University*, 11, 129–161. (in Japanese with English abstract).
- 590 Mizuno, N. (1987). Yoshinobori rui (genus *Rhinogobius*). In N. Mizuno, & A. Goto (Eds.),
591 *Nihon no tansuigyo – sono bunpu, hen'i, shubunka wo megutte (Freshwater fishes in Japan –*
592 *Distribution, variation, and speciation)* (pp. 179–188). Tokyo, Japan: Tokai University Press.
593 (in Japanese).

594 Mizuno, N. (2001). *Rhinogobius*. In H. Kawanabe, N. Mizuno, & K. Hosoya (Eds.),
595 *Freshwater fishes of Japan*, 3rd ed. (p. 584) Tokyo, Japan: Yama-Kei Publishers. (in
596 Japanese).

597 Moriyama, A., Yanagisawa, Y., Mizuno, N., & Omori, K. (1998). Starvation of drifting goby
598 larvae due to retention of free embryos in upstream reaches. *Environmental Biology of Fishes*,
599 52, 321–329. <https://doi.org/10.1023/A:1007333302864>

600 Nei, M., Tajima, F., & Tatenno, Y. (1983). Accuracy of estimated phylogenetic trees from
601 molecular data. *Journal of Molecular Evolution*, 19, 153–170.

602 Nishida, M. (2001). Dairanka eno purosusu: Yoshinobori rui (The process toward large egg
603 forms: *Rhinogobius*. In A. Goto, & K. Iguchi (Eds.), *Suisei doubutsu no ran-saizu:*
604 *Seikatsushi no hen'i, shubunka no seibutsugaku (Evolutionary biology of egg size in aquatic*
605 *animals)* (pp. 149–170). Tokyo, Japan: Kaiyusha. (in Japanese).

606 Nosil, P. (2012). *Ecological speciation*. Oxford, UK: Oxford University Press.

607 Ohara, K., Takagi, M., Hashimoto, M., Miyazaki, K., & Hirashima, K. (2008). DNA markers
608 indicate low genetic diversity and high genetic divergence in the landlocked freshwater goby,
609 *Rhinogobius* sp. YB, in the Ryukyu Archipelago, Japan. *Zoological Science*, 25, 391–400.
610 <https://doi.org/10.2108/zsj.25.391>

611 Ohara, K., Takahashi, D., & Takagi, M. (2004). Isolation of microsatellite loci in the
612 freshwater goby, *Rhinogobius* sp. (Gobiidae). *Molecular Ecology Notes*, 4, 449–451.
613 <https://doi.org/10.1111/j.1471-8286.2004.00677.x>

614 Omori, K., Ohnishi, H., Hamaoka, H., Kunihiro, T., Ito, S., Kuwae, M., ... Iguchi, K. (2012).
615 Speciation of fluvial forms from amphidromous forms of migratory populations. *Ecological*
616 *Modelling*, 243, 89–94. <https://doi.org/10.1016/j.ecolmodel.2012.06.006>

617 Osozawa, S., Shinjo, R., Armid, A., Watanabe, Y., Horiguchi, T., & Wakabayashi, J. (2012).
618 Palaeogeographic reconstruction of the 1.55 Ma synchronous isolation of the Ryukyu Islands,
619 Japan, and Taiwan and inflow of the Kuroshio warm current. *International Geology Review*,
620 54, 1369–1388. <https://doi.org/10.1080/00206814.2011.639954>

621 Ota, H. (1998). Geographic patterns of endemism and speciation in amphibians and reptiles of
622 the Ryukyu Archipelago, Japan, with special reference to their paleogeographical implications.
623 *Researches on Population Ecology*, 40, 189–204. <https://doi.org/10.1007/BF02763404>

624 Parent, C. E., & Crespi, B. J. (2006). Sequential colonization and diversification of Galapagos
625 endemic land snail genus *Bulimulus* (Gastropoda, Stylommatophora). *Evolution*, 60, 2311–
626 2328. <https://doi.org/10.1111/j.0014-3820.2006.tb01867.x>

627 Pritchard, J. K., Stephens, M., & Donnelly, P. (2000). Inference of population genetic structure
628 using multilocus genotype data. *Genetics*, 155, 945–959.

629 R Core Team (2014). *R: A language and environment for statistical computing*. Vienna,
630 Austria: R Foundation for Statistical Computing. <http://www.R-project.org/>

631 Rabosky, D. L. (2016). Reproductive isolation and the causes of speciation rate variation in
632 nature. *Biological Journal of the Linnean Society*, 118, 13–25.
633 <https://doi.org/10.1111/bij.12703>

634 Rabosky, D. L., & Matute, D. R. (2013). Macroevoolutionary speciation rates are decoupled
635 from the evolution of intrinsic reproductive isolation in *Drosophila* and birds. *Proceedings of*
636 *the National Academy of Sciences of the United States of America*, 110, 15354–15359. <https://doi.org/10.1073/pnas.1305529110>

638 Reznick, D. N., & Ricklefs, R. E. (2009). Darwin's bridge between microevolution and
639 macroevolution. *Nature*, 457, 837–842. <https://doi.org/10.1038/nature07894>

640 Rosenblum, E. B., Sarver, B. A. J., Brown, J. W., Des Roches, S., Hardwick, K. M., Hether, T.
641 D., ... Harmon, L. J. (2012). Goldilocks meets Santa Rosalia: An ephemeral speciation model
642 explains patterns of diversification across time scales. *Evolutionary Biology*, 39, 255–261.
643 <https://doi.org/10.1007/s11692-012-9171-x>

644 Rosindell, J., & Phillimore, A. B. (2011). A unified model of island biogeography sheds light
645 on the zone of radiation. *Ecology Letters*, 14, 552–560.
646 <https://doi.org/10.1111/j.1461-0248.2011.01617.x>

- 647 Schluter, D., & Nagel, L. M. (1995). Parallel speciation by natural selection. *American*
648 *Naturalist*, 146, 292–301.
- 649 Seehausen, O., & Wagner, C. E. (2014). Speciation in freshwater fishes. *Annual Review of*
650 *Ecology, Evolution, and Systematics*, 45, 621–651.
651 <https://doi.org/10.1146/annurev-ecolsys-120213-091818>
- 652 Shinomiya, A., Sasabe, K., Sakurai, M., & Kishino, T. (2005). Morphological characters of
653 newly hatched larvae and their habitat selection of a fluviatile goby, *Rhinogobius* sp. YB in
654 the upper stream of Sumiyo River, Amami-oshima Island, Japan. *Japanese Journal of*
655 *Ichthyology*, 52, 1–8. (in Japanese with English abstract). [https://doi.org/10.11369](https://doi.org/10.11369/jji1950.52.1)
656 [/jji1950.52.1](https://doi.org/10.11369/jji1950.52.1)
- 657 Suzuki, T., Shibukawa, K., Yano, K., & Senou, H. (2004). *A photographic guide to the*
658 *gobioid fishes in Japan*. Tokyo, Japan: Heibonsha. (in Japanese).
- 659 Tachihara, K. (2009). Two landlocked *Rhinogobius* species in the Ryukyu Archipelago:
660 Conservation and the future of gobies endemic to isolated rivers. *Japanese Journal of*
661 *Ichthyology*, 56, 70–74. (in Japanese). <https://doi.org/10.11369/jji.56.70>
- 662 Takahashi, D., & Kohda, M. (2004). Courtship in fast water currents by a male stream goby
663 (*Rhinogobius brunneus*) communicates the parental quality honestly. *Behavioral Ecology and*
664 *Sociobiology*, 55, 431–438. <https://doi.org/10.1007/s00265-003-0693-1>
- 665 Takezaki, N., Nei, M., & Tamura, K. (2010). POPTREE2: Software for constructing
666 population trees from allele frequency data and computing other population statistics with
667 Windows interface. *Molecular Biology and Evolution*, 27, 747–752.
668 <https://doi.org/10.1093/molbev/msp312>
- 669 Tamada, K. (2005). Clutch and egg size in *Rhinogobius* sp. DA inhabiting small rivers.
670 *Japanese Journal of Ichthyology*, 52, 17–20. (in Japanese with English abstract).
671 <https://doi.org/10.11369/jji1950.52.17>
- 672 Taylor, E. B., Boughman, J. W., Groenenboom, M., Sniatynski, M., Schluter, D., & Gow, J. L.
673 (2006). Speciation in reverse: Morphological and genetic evidence of the collapse of a

674 three-spined stickleback (*Gasterosteus aculeatus*) species pair. *Molecular Ecology*, 15, 343–
675 355. <https://doi.org/10.1111/j.1365-294X.2005.02794.x>

676 Tominaga, A., Matsui, M., Eto, K., & Ota, H. (2015). Phylogeny and differentiation of
677 wide-ranging Ryukyu Kajika frog *Buergeria japonica* (Amphibia: Rhacophoridae):
678 Geographic genetic pattern not simply explained by vicariance through strait formation.
679 *Zoological Science*, 32, 240–247. <https://doi.org/10.2108/zs140227>

680 Vähä, J.-P., & Primmer, C. R. (2005). Efficiency of model-based Bayesian methods for
681 detecting hybrid individuals under different hybridization scenarios and with different
682 numbers of loci. *Molecular Ecology*, 15, 63–72.
683 <https://doi.org/10.1111/j.1365-294X.2005.02773.x>

684 Vamosi, S. M. (2003). The presence of other fish species affects speciation in threespine
685 sticklebacks. *Evolutionary Ecology Research*, 5, 717–730.

686 Wagner, C. E., Harmon, L. J., & Seehausen, O. (2014). Cichlid species– area relationships are
687 shaped by adaptive radiations that scale with area. *Ecology Letters*, 17, 583–592.
688 <https://doi.org/10.1111/ele.12260>

689 Weigelt, P., Steinbauer, M. J., Cabral, J. S., & Kreft, H. (2016). Late Quaternary climate
690 change shapes island biodiversity. *Nature*, 532, 99–102. <https://doi.org/10.1038/nature17443>

691 Wellborn, G. A., & Langerhans, R. B. (2014). Ecological opportunity and the adaptive
692 diversification of lineages. *Ecology and Evolution*, 5, 176– 195.
693 <https://doi.org/10.1002/ece3.1347>

694 Yamasaki, Y. Y., Nishida, M., Suzuki, T., Mukai, T., & Watanabe, K. (2015). Phylogeny,
695 hybridization, and life history evolution of *Rhinogobius* gobies in Japan, inferred from
696 multiple nuclear gene sequences. *Molecular Phylogenetics and Evolution*, 90, 20–33.
697 <https://doi.org/10.1016/j.ympev.2015.04.012>

698 Yoder, J. B., Clancey, E., Des roches, S., Eastman, J. M., Gentry, L., Godsoe, W., ... Harmon,
699 L. J. (2010). Ecological opportunity and the origin of adaptive radiations. *Journal of*
700 *Evolutionary Biology*, 23, 1581–1596. <https://doi.org/10.1111/j.1420-9101.2010.02029.x>

701 Yoshigou, H. (2014). Annotated checklist and bibliographic records of inland water fishes of
702 the Ryukyu Archipelago, Japan. *Fauna Ryukyana*, 9, 1–153. (in Japanese with English
703 abstract).

704

705 **TABLE 1** Results of model selection by ABC analysis

Tested geographic scale	Origins of freshwater form in the best model	Gene flow	Posterior probability of the best model	Bayes factors of the best model against all models	Bayes factor of the best model against another origin models	Type 2 error rate (all models)	Type 2 error rate (against another origin models)
Within-islands							
Amami-Oshima Is.	Single origin	Partial	0.5087	1.668-37.70	12.04-37.70	0.05-0.17	0.09-0.12
Tokunoshima Is.	Single origin	All	0.3128	1.355-2069	133.5 - 2069	0.02-0.11	0.05-0.07
Okinawajima Is.	Single origin	All	0.6357	1.754 - 4300	544.4 - 4300	0-0.04	0
Kumejima Is.	Single origin	Partial	0.3254	1.047 - 18735	4368- 18735	0.02-0.17	0.02-0.09
Iriomotejima Is.	Single origin	Partial	0.4544	1.321 - 309807	3247- 309807	0.01-0.15	0.01-0.15
Among-islands within-regions							
Middle Ryukyu	Originated in respective islands	All	0.4471	1.090-519.2	23.97-24.37	0-0.35	0
South Ryukyu	Single origin	All	0.8949	13.63 - 191.0	36.98 - 191.0	0-0.08	0-0.08
Tokunoshima Is. – Okinoerabujima Is.	Single origin	All	0.7461	4.680 - 122.6	9.419 -122.6	0.06-0.1	0.06-0.1
Among-regions							
Middle Ryukyu – South Ryukyu	Robust result could not obtain.	-	-	-	-	-	-

706

707

708

709 **TABLE 2** Results of the logistic regression analysis

Explanatory variable	Data Set	<i>P</i> value	Difference of deviance	R^2
Island Area	Larger island	0.0016	12.265	0.577
	Smaller island	0.0061	8.650	0.407
Catchment Area	Larger island	0.0026	12.055	0.567
	Smaller island	0.0187	6.923	0.325
River Length	Larger island	0.0038	11.699	0.550
	Smaller island	0.0172	7.028	0.330
No. Waterfalls	Larger island	0.0013	13.712	0.645
	Smaller island	0.3277	1.407	0.066

711 *Note:* The results presented in this table are based on the assumption that the freshwater form
712 (FF) has five independent origins. The data set column indicates the assumption of the islands
713 where speciation occurred. Our ABC analysis indicated that FF in Tokunoshima Is. and
714 Okinoerabujima Is., and Ishigakijima Is. and Iriomotejima Is. shared their origins. “Larger
715 island” in the data set column assumed that the speciation of FF occurred in the larger island
716 (i.e., Tokunoshima Is. and Iriomotejima Is.), whereas “smaller island” assumed that speciation
717 occurred in the smaller island (i.e., Okinoerabujima Is. and Ishigakijima Is.). The *p* values for
718 all combinations are presented in Table S9.

719

720

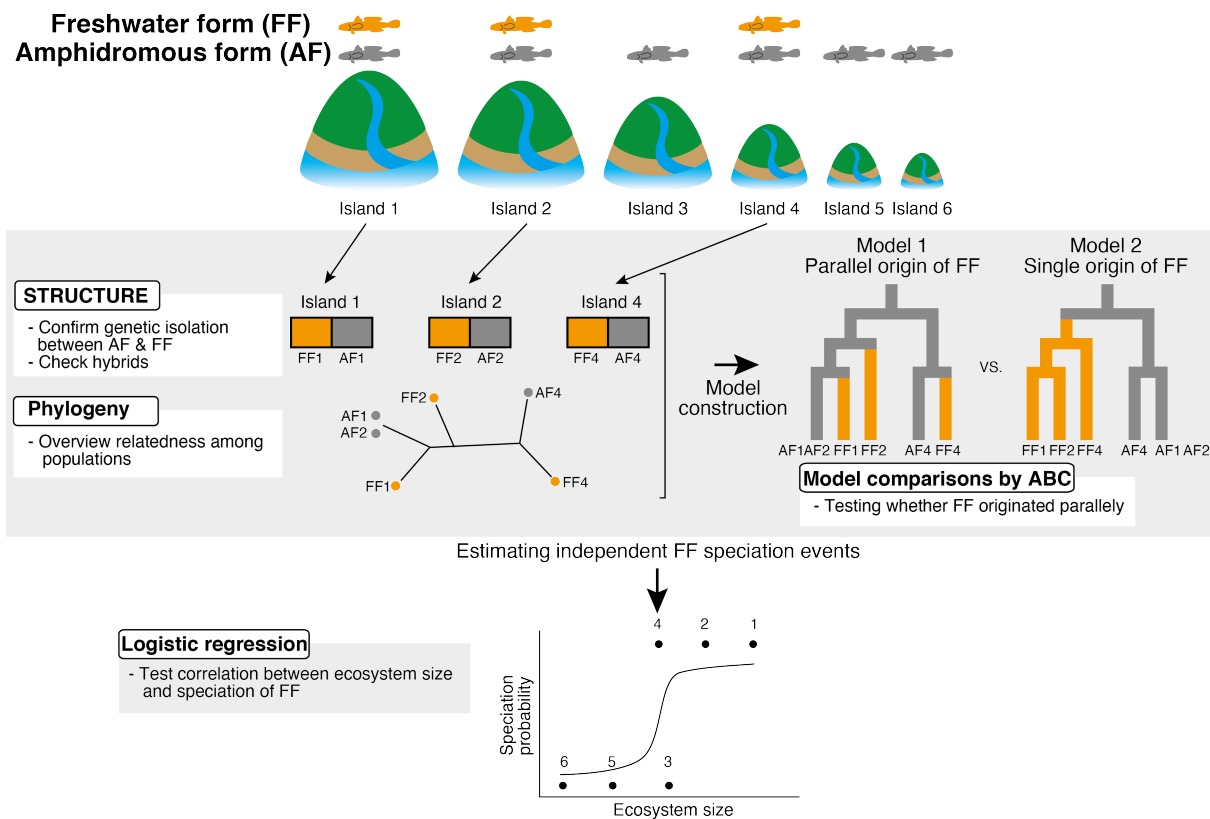


FIGURE 1 Concept diagram of this study based on a hypothetical example. Our final goal was to determine whether ecosystem size explains the occurrence of speciation of the freshwater form (FF). To explain the study workflow, we postulate six islands with the amphidromous form (AF). FF was distributed in three islands among them. Four main analyses were conducted in this study. First, we confirmed genetic isolation between AF and FF by structure analysis for each island. Second, we estimated phylogenetic relationships among populations. Third, we tested independent speciation of FF by demographic comparison based on ABC analysis. The phylogenetic tree and geographical locations were used to construct demographic models. The occurrence of speciation in focal islands was identified by these three analyses. Finally, we tested the correlation between ecosystem size and occurrence of speciation by logistic regression analysis.

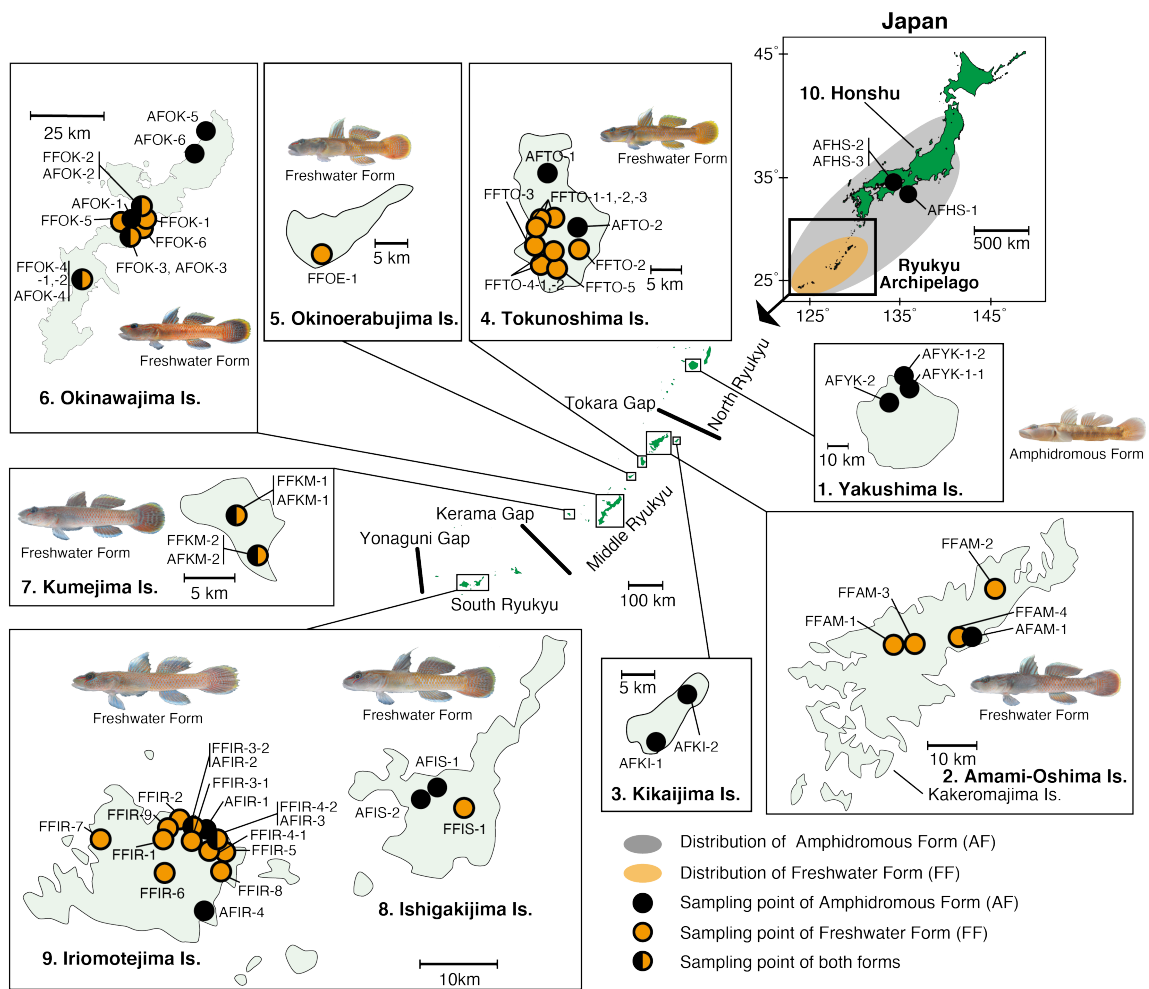


FIGURE 2

Sampling points in this study. The map in the upper right box indicates distribution ranges of the amphidromous form (AF; grey) and the freshwater form (FF; orange) in Japan as a whole. The central map is the whole area of the Ryukyu Archipelago. Black and orange circles indicate the points from where AF and FF were collected, respectively. Sympatrically collected points are indicated by both colours.

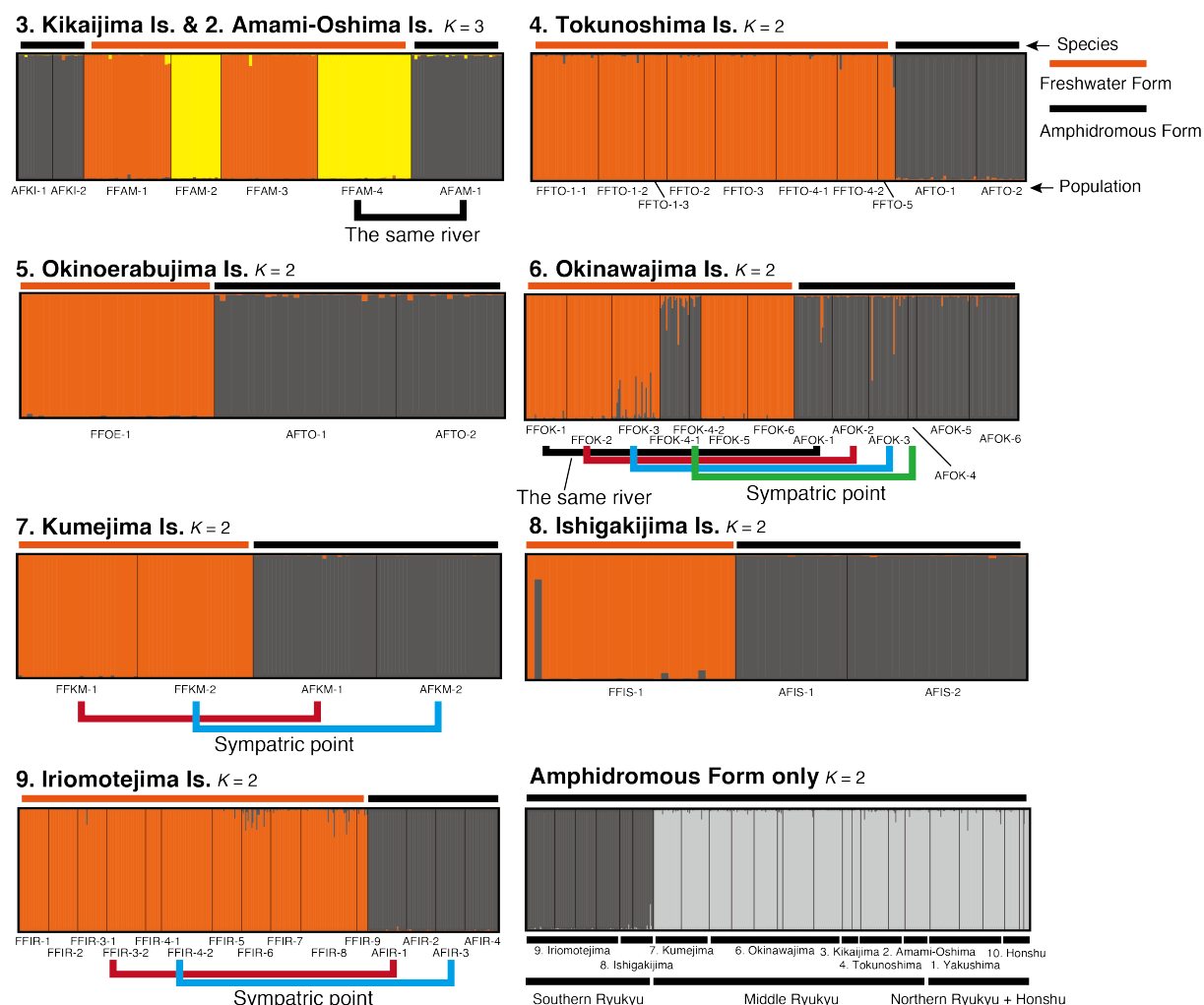


FIGURE 3 Results of genetic clustering analysis performed in structure (Amami-Oshima Is., $K = 3$; others, $K = 2$). Sampling points are indicated by black vertical bars in the respective bar plots. Horizontal bars at the top of the plot indicate morphological identification of respective specimens. Sympatric points are indicated by an additional annotation. The specimens morphologically identified as the freshwater form (FF) are presented in orange and yellow colours, whereas those identified as the amphidromous form (AF) are presented in grey.

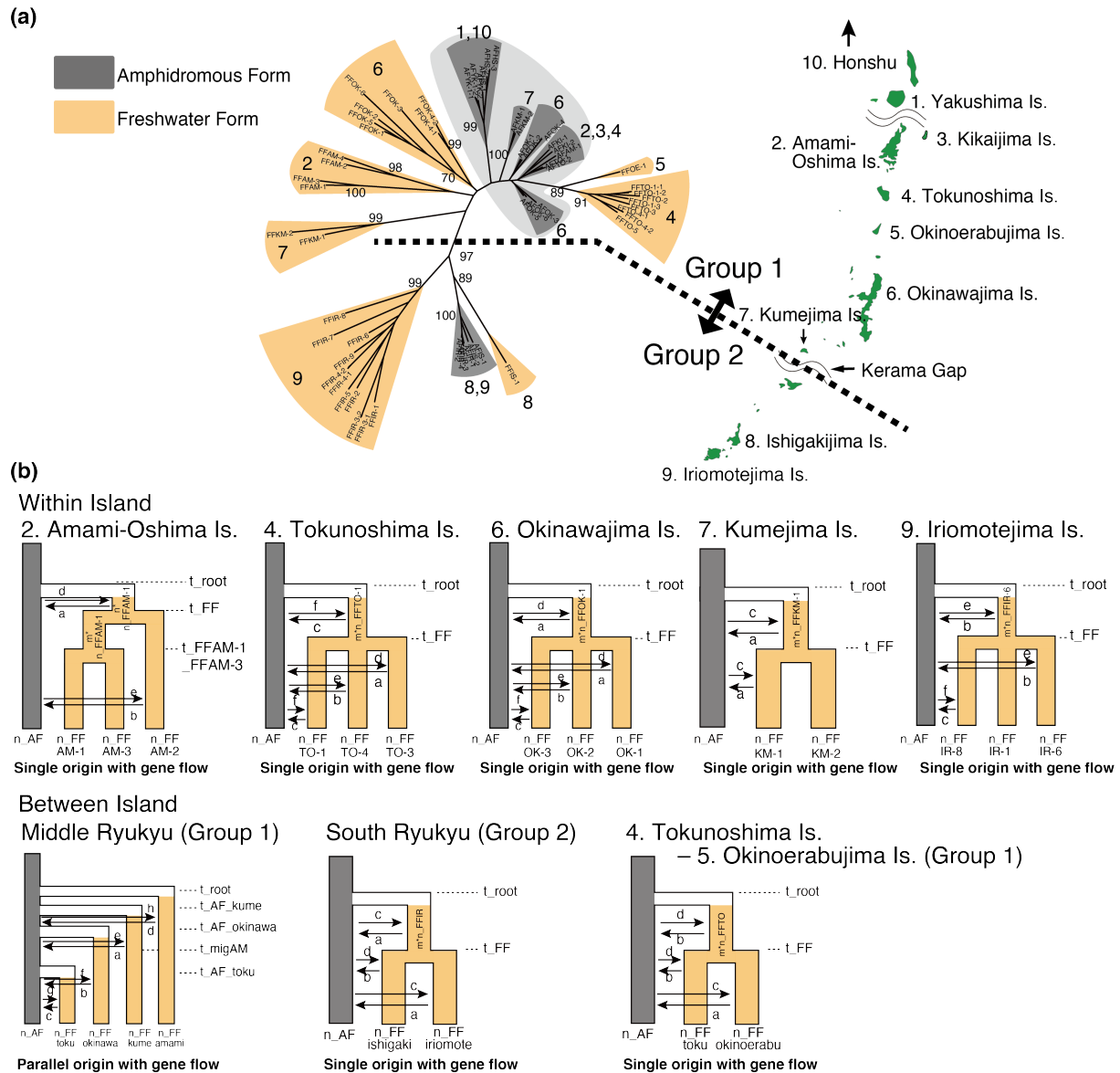


FIGURE 4 Population relationships and evolutionary scenarios of the amphidromous form (AF) and the freshwater form (FF). (a) Neighbour-joining tree based on Nei's *Da* distance of 20 microsatellites. Numbers on tree nodes indicate bootstrap values. Only bootstrap values >70% are indicated. Island numbers are indicated at tips. (b) The models selected by ABC analysis. AF and FF are represented by black and orange colours, respectively. Arrows between populations indicate gene flow. Only representative populations were used for ABC analysis to reduce the amount of calculation (see Materials and methods and Supporting methods and results S1).

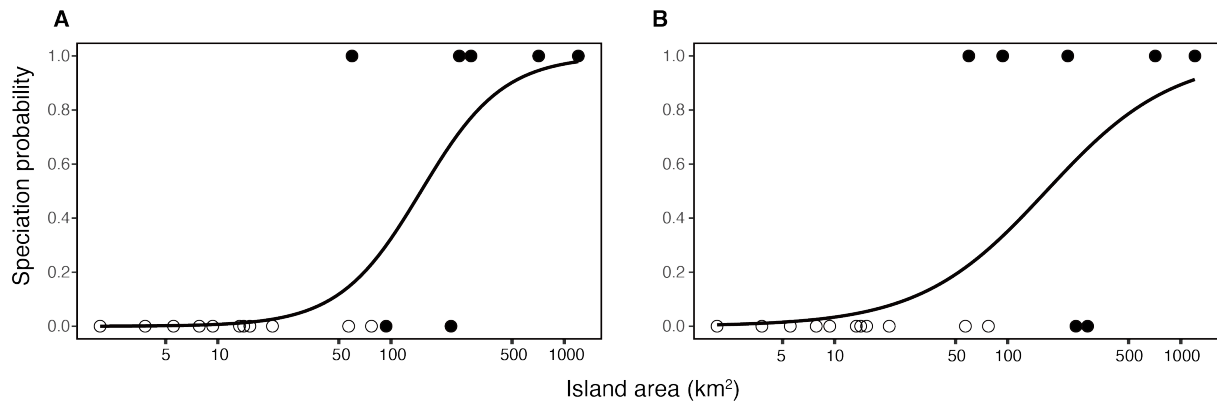


FIGURE 5 Results of logistic regression analysis between island size and speciation probability. Islands with the freshwater form (FF) are represented by filled circles, whereas those without FF are represented by open circles. Filled circles plotted on the x-axis (speciation probability = 0) are for the islands with FF that was assumed to be of nonindependent origin. Our ABC analysis indicated that FF in Tokunoshima Is. and Okinoerabujima Is., and Ishigakijima Is. and Iriomotejima Is. shared their origins. (a) Speciation of FF is assumed to have occurred in the larger islands (Tokunoshima Is. and Iriomotejima Is.). (b) Speciation is assumed to have occurred in the smaller islands (Okinoerabujima Is. and Ishigakijima Is.).

Supplemental methods and results S1: Complete version of Materials and Methods and Results

Materials and Methods

1. Study system

1.1 Study species

The genus *Rhinogobius* (Gobiidae) is the most speciose group of gobies occurring in freshwater. This group contains more than 85 species in East and Southeast Asia, and 18 species in Japan (Oijen et al., 2011; Yamasaki et al., 2015). The freshwater resident form derived from the amphidromous species *R. brunneus* is recognized as an independent species but has not been described to date because of taxonomic confusion. Previous studies (e.g., Akihito et al., 2013) tentatively used the name “*Rhinogobius* sp. YB” for this freshwater form. However, until its evolutionary status is clarified, we will use the code ‘FF’ (freshwater form) for *Rhinogobius* sp. YB and ‘AF’ (amphidromous form) for *R. brunneus* based on their life history modes.

The FF completes its life in freshwater; conversely, AF larvae hatch in rivers and immediately drift down to the sea and return to rivers after 2–3 months (Mizuno, 2001). The eggs and larvae of FF are larger than those of AF. The sizes of eggs and hatched larvae of respective forms and islands have already reported by some studies. As follows, the both egg and larval sizes are greater in FF than in AF. The mean and standard deviation of egg size of FF varies across islands; namely, long axis 3.07 and 3.36 mm, short axis 1.55 and 1.52 mm on Amami-Oshima Island (Kon et al., 2003); long axis 5.06 ± 0.39 mm, short axis 1.77 ± 0.10 mm on Amami-Oshima Is. (Shinomiya et al., 2005); long axis 3.4 and 4.4 mm, short axis 1.4 and 1.7 mm on Tokunoshima Island (measured from the Figure 2 of Oshiro, 1981), long axis 4.3 ± 0.2 mm, short axis 1.4 ± 0.1 mm on Okinawajima Island (Hirashima & Tachihara, 2000); long axis 5.4 mm, short axis 1.7 mm on Ishigakijima Island (measured from the Figure 2 of Oshiro, 1981). We also measured the egg size of FF spawned in a tank by one individual collected from Shirase River on Kumejima Island (long axis 3.52 ± 0.24 mm, short axis 1.31 ± 0.07 mm, $n=26$). The mean and standard deviation of egg size of AF was recorded on Okinawajima Island; namely, long axis 2.8 ± 0.1 mm, short axis 0.9 ± 0.04 mm (Kondo et al., 2013). The mean and standard deviation for notochord length of hatched larvae of FF varies across islands; namely, 5.0 mm on Amami-Oshima Is. (Kon et al., 2013);

5.96–6.60 mm on Amami-Oshima Is. (Shinomiya et al., 2004); 5.3 mm on Okinawajima Is. (Hirashima & Tachihara 2000). The mean and standard deviation in notochord length of hatched larvae of AF on Okinawajima Is. was 3.7 ± 0.2 mm (Kondo et al., 2013); 3.59 ± 0.087 mm on Shikoku Island in the Japanese archipelago (Moriyama et al., 1998). Fluvial species of this genus tend to produce larger eggs and larvae (Yamasaki et al., 2015). Such traits are considered to be an adaptation to river environments, where small food items (i.e., plankton) are scarce and higher level of swimming ability in river current is needed during the larval period (Nishida, 2001; Shinomiya et al., 2005; Hirashima & Tachihara, 2000; Yamasaki et al., 2015).

Although the distribution of the two species sometimes overlapped in a single river, the tendencies of their distributions were different. AF mainly lived in the middle to lower reaches of rivers. In fact, AF populations are actually only found in rivers shorter than 18 km in the central Honshu region of Japan (Tamada, 2005). The larvae of AF must rapidly reach the sea to avoid starvation because of the scarcity of small food in rivers (Moriyama, Yanagisawa, Mizuno, & Omori, 1998; Iguchi & Mizuno, 1999). Thus, they might not be able to reproduce effectively in areas far from the sea (Iguchi & Mizuno, 1999). On the other hand, FF mainly live in the upper reaches of rivers, and are often found in the upper areas of waterfalls where no other fishes are distributed (Kano et al., 2012).

1.2 Study area

The Ryukyu Archipelago is a series of islands located in the southwestern part of Japan (Figure 2). The islands separated simultaneously from the continent about 1.5 million years ago (Osozawa et al., 2012). Three deep straits (Tokara Gap, Kerama Gap, and Yonaguni Gap) divide this archipelago into three regions: North Ryukyu, Middle Ryukyu, and South Ryukyu. These gaps act as biogeographic barriers for terrestrial organisms (Ota 1998). On the islands of North Ryukyu, the biota was destroyed by pyroclastic flow 7300 years ago (Machida et al., 2001). AF is distributed in all three regions, as well as mainland Japan, whereas FF is only distributed in Middle and South Ryukyu (Yoshigou, 2014). In the Ryukyu Archipelago, primary freshwater fishes are almost absent, and most fishes are diadromous. Therefore, fish assemblages in freshwater areas are simple and similar among islands (Yoshigou, 2014).

2. Sampling

All animal experiments followed the rules, regulations, and guidelines relating to animal experimentation of Kyoto University. Sampling of FF and AF was conducted from June 2012 to March 2015 on seven islands where former populations have been recorded previously (Locality code 2. Amami-Oshima Is., 4. Tokunoshima Is., 5. Okinoerabujima Is., 6. Okinawajima Is., and 7. Kumejima Is. in Middle Ryukyu; 8. Ishigakijima Is. and 9. Iriomotejima Is. in South Ryukyu; Figure 2; Table S1; Ohara et al. 2008; Yoshigo, 2014). In addition, AF specimens were collected from Kikaijima Is. (Locality code 3) in Middle Ryukyu, Yakushima Is. (1) in North Ryukyu, and mainland Japan (10). The specimens were identified by body colour patterns under fresh conditions (Suzuki et al., 2004; Kano et al., 2012). Identification based on body colour was validated by the results of the genetic analysis; however, some samples were difficult to identify or showed inconsistent genetic characteristics (3.2% of specimens in total, including potential hybrids). When genetic data suggested the focal specimens were not judged as hybrids based on the criteria described below and that wrong identification based on morphology in the field was likely, these samples were removed from further analysis. In the case that a sample was judged to be a hybrid (see 4.2), we tentatively named it based on the morphology. It should be noted that we found that morphologically identified FFOK-4-1 and FFOK-4-2 populations of FF largely reflected the genetic features of AF (Figure 3 and Supporting Information Figure S3). We judged them as hybrid populations and did not remove any specimens from the analysis even if they were not judged as hybrid (see Discussion). Specimens were anesthetized with iced water or 2-phenoxy-ethanol, and then a fin was clipped and preserved in 99% ethanol. Most specimens were fixed in 10% formalin or 99% ethanol. As a result, 842 FF and 578 AF specimens from 52 points in 42 river systems were included in the population genetic analysis. The average numbers of localities and specimens per island were 5.2 and 142, respectively. Both forms were collected from nine river systems; they were sympatrically collected at seven points.

3. Molecular experiment

3.1 Microsatellite marker development

We developed new microsatellite markers following the methods of Takeshima et al.

(2017). Total genomic DNA was extracted from one AF specimen collected from DAIR-4 in Iriomotejima Island by using a Genomic DNA Purification Kit (Promega, Madison, Wisconsin, USA). We concentrated CA repeat regions by using the target capture method and sequenced them using the Roche 454 GS Junior system (Accession No. DRA006388). Primers and multiplex systems were designed using QDD ver. 2 (Meglec et al., 2010) and Multiplex Manager v. 1.2 (Holleley & Geerts 2009), respectively. One out of four universal tail sequences was added to each forward primer for fluorescent labelling during PCR (Blacket et al., 2012).

3.2 DNA extraction, PCR, and genotyping

Total genomic DNA was extracted from the fin clip by using the Genomic DNA Purification Kit. Before PCR amplification, a droplet of 1.25 μ L extracted DNA was placed on a 96 well plate and dried. Total 21 microsatellite markers (18 originally developed loci and three loci reported by Ohara et al. (2004)) were analyzed. We developed four multiplex sets and simultaneously amplified 4–10 loci (Table S2). For the markers that were developed, the final concentration of each forward and reverse primer was 1 and 2 μ M, respectively. The 1 μ M tailed fluorescent primer for each forward primer was also added to the multiplex reaction mixture. For the three loci reported by Ohara et al. (2004), 5'-fluorescent-labelled forward primers were used, and the final concentration of each forward and reverse primers was adjusted to 1 μ M. PCR amplification was performed in a 4 μ L volume containing 2.43 μ L of ultrapure water, 0.066 μ L of primer mix, and 1.5 μ L of Type-IT Microsatellite PCR Kit (Qiagen, Hilden, Germany) with the dried DNA. The PCR settings for the originally developed markers included a first step of denaturation at 95°C for 5 min, followed by 35 cycles of the second step (denaturation: 94°C, 30 s; annealing: 63°C, 90 s; and extension: 72°C, 75 s), and the final extension at 72°C for 30 min. For the PCR for the loci of Ohara et al. (2004), the second step was changed as follows: 28 cycles of denaturation at 94°C for 30 s; annealing at 53°C for 90 s; and extension at 72°C for 60 s. We directly added 10 μ L HiDi Formamide (Thermo Fisher Scientific, Waltham, MA, USA) and 0.2 μ L of GeneScan 500Liz dye Size Standard or GeneScan 400HD Rox dye Size Standard (Thermo Fisher Scientific) for the respective PCR products. Next, we subjected the products to heat shock treatment (95°C, 5 min; 0°C, 5 min). The peak size data were acquired using ABI 3130xl sequencer (Applied Biosystems, Foster City, CA, USA) and

were analyzed using GeneMapper v.4.0 (Applied Biosystems). The acquired allele size data were transformed for integral number by using tandem v.1.09 (Matschiner & Salzburger 2009).

We genotyped a total of 21 microsatellite loci for population genetic analyses (Table S2). Null alleles were checked using software Micro-checker ver.2.2.3 (van Oosterhout et al., 2004) under 100% interval condition for the 25 populations of *R. brunneus*, which were assumed to not be affected by the bottleneck effect. Null alleles were detected in several *R. brunneus* populations. We excluded locus br_b_04 from the analysis because null alleles were detected at this locus in both North Ryukyu and Middle Ryukyu populations. Null alleles in other loci occurred sporadically; hence, we used the other 20 loci for the following analysis to avoid reduction of information. Linkage disequilibrium among loci was tested using 10,000 times of permutation implemented in Arlequin ver.3.5 (Excoffier & Lischer 2010), and Bonferroni collection was applied to *P*-values to avoid type I errors. No significant linkage disequilibrium between loci was detected (adjusted $P \geq 0.05$). The final genotype data were deposited in dryad (doi: 10.5061/dryad.mpg4f4qvh).

4. Data analysis

4.1 Calculation of summary statistics and estimation of phylogenetic trees

We calculated the expected and observed heterozygosity (*He* and *Ho*, respectively), the effective number of alleles (*Ae*) for the respective populations, and population-pairwise *F_{ST}* and Jost's *D* by using GenoDive 2.0b27 (Meirmans & van Tienderen 2004). We also calculated allelic richness (*Rs*) of the respective populations by using FSTAT v2.9.3.2 (Goudet 1995). Phylogenetic relationships among populations were estimated based on *Nei's Da* distance by using the neighbor-joining method implemented in Poptree2 (Takezaki et al., 2010). The credibility of each clade was evaluated using 1,000 bootstrap resamplings. To assess hierarchical genetic differentiation, we conducted AMOVA (Excoffier, Smouse, & Quattro, 1992) implemented in GenoDive separately for AF and FF. The number of hierarchies to be tested was five. However, AMOVA could include only four hierarchies in one analysis. To deal with this problem, we conducted AMOVA separately in two geographic scales. In the among-islands within-regions scale, we set following four hierarchies: within-individuals, among-individuals, among-rivers, and among-islands. This analysis was separately

conducted for Middle Ryukyu and South Ryukyu. In the among-regions scale, we set the following four hierarchies: within-individuals, among-individuals, among-islands, and among-regions. F_{ST} -analog distance was applied. Significance was tested using 999 permutations.

4.2 STRUCTURE analysis

To infer the existence and extent of gene flow and genetic isolation, we analyzed the genetic structure of the population using STRUCTURE v.2.3.4 (Pritchard, Stephen, & Donnelly, 2000; Lawson, van Dorp, & Falush, 2018). We prepared two datasets: one for both FF and AF populations from each island, and another for only AF populations from all islands. Because AF was not found on Okinoerabujima Is., the FF on the island were analyzed along with the AF from the nearest island, Tokunoshima Is. Admixture model with correlated allele frequencies was assumed. Markov Chain Monte Carlo was performed with 100,000 generations for burn-in and final 1,000,000 generations. Ten independent runs for each K from 1 to 10 were evaluated. After a run finished, we calculated Evanno's ΔK by using STRUCTURE Harvester (Evanno, Regnaut, & Goudet, 2005; Earl & von Holdt, 2012) was used for a reference for K value selection. Hybrid individuals were evaluated using q -values; we arbitrarily judged specimens as hybrids (or individuals with hybrid origin) in cases where minor genetic components accounted for 10% or more of the aggregate (Vähä & Primmer 2006).

4.3 Testing parallel speciation of FF by using ABC

To statistically examine the single or parallel origins of FF, and to estimate which FF populations shared their origins we conducted a demographic model selection under the approximate Bayesian computation (ABC framework, Beaumont, 2010; Bertorelle et al., 2010) implemented in ABCtoolbox v.1 (Wegmann et al., 2010) and R package 'abc' v.2.1 (Csilléry et al., 2012). Coalescent simulation was conducted using fastsimcoal2 v.2.5.2.1.1 (Excoffier & Foll 2011; Excoffier et al., 2013). Summary statistics were calculated using arlsumstat v.3.5 for each simulation (Excoffier & Lischer 2010).

4.3.1 Prior distribution, summary statistics, and simulation

Prior distribution and summary statistics were determined after several preliminary analyses. When the range of the prior distribution of a parameter was across the digits,

the value was $\log_{10}$ transformed. A uniform distribution for each parameter was assumed. We set a larger prior range for population size of AF (10^2 – $10^{5.3}$) than for FF (10^2 – $10^{4.7}$). The prior distribution of coalescent time (generation) was set to 10^3 – $10^{6.2}$, considering that the Ryukyu Archipelago had separated from the continents 1.5 million years ago, and the generation time of the *Rhinogobius* species was one to two years (Yamasaki, personal observation). The prior distribution of the migration rate was set to 10^{-9} – 10^{-3} . In the case that the models of two or more populations were merged, the ancestral population size was set to the range of 0.5 to 2 times of either of the derived populations. We applied a generalized stepwise mutation model for microsatellite evolution. Our data did not match the simulated data produced using the models under a previously reported average mutation rate of microsatellites in fish (common carp: 5.56×10^{-4} /locus/generation, 95% interval, 1.52×10^{-4} – 1.63×10^{-3} , Yue et al., 2007) from principal component analysis (PCA) (see 4.3.3). Simulated data under several average mutation rates were tested, and then the value of 1.0×10^{-4} was selected, under which the generated data matched the acquired data. The prior distribution of the mutation rate for each locus was set to a Gamma distribution. The maximum difference of repeat number of alleles of each microsatellite in coalescent simulations was set to 45.

To capture information regarding population size and differentiation, we selected summary statistics listed in Table S3. Detailed formulae of respective summary statistics are described in the Arlequin 3.5 manual (Excoffier & Lischer 2010). A total of 46 to 132 summary statistics for each analysis were used according to the number of the analyzed populations. We performed 1.0×10^6 simulations for each demographic model.

4.3.2 General strategy for the construction of demographic models

Our general model construction strategies for ABC analysis were as follows. First, we generated separate models using three geographic scales, i.e. within-islands, regional (Middle and South Ryukyu), and archipelago scales, and examined FF origins in each geographic scale to reduce the number of models and populations examined in a single analysis. Second, we examined both models with and without gene flow between AF and FF populations. This is because a simulation study suggests that gene flow could alter apparent phylogenetic relationships among populations when simple genetic distance was used (Bierne, Gagnaire, & David, 2013). Demographic model comparisons that consider gene flow can overcome this problem (Butlin et al., 2014).

In the within-islands analyses, we mainly examined two questions: whether FF had single or parallel origin, and whether gene flow between AF and FF existed on each island. We hypothesized several plausible divergence orders between populations based on the result of STRUCTURE analysis, phylogeny, and geographic locations for respective islands. Thereafter, we compared the models with single/parallel origins and with/without gene flow. Analyses were conducted for the five islands where FF is distributed in multiple rivers (i.e., Amami-Oshima Is., Tokunoshima Is., Okinawajima Is., Kumejima Is., and Iriomotejima Is.). Ishigakijima Is. and Okinoerabujima Is. were not included in the analyses because the FF population was only found in one river system on each island. We selected three (or two) river populations for each island to reduce the model complexities and calculation loading. FF specimens from a single river system were merged to increase calculation accuracy. Populations were selected by the following criteria to extract information of population history as much as possible; geographically distant river systems, to avoid bias from shared population history among geographically nearby populations; and lower rates of missing data. The constructed models (a total of 34 models) are described in Supporting Information Figure S1A–E.

For the among-islands within-region analyses, we constructed several plausible models separately for the regions of Tokunoshima Is. and Okinoerabujima Is., Middle Ryukyu, and South Ryukyu. Five to ten models were constructed on the basis of the phylogenetic tree and geographic distance between islands. Gene flow settings followed the results of within-islands scale analysis. For each island, we selected only one FF population whose population size was the largest, as judged from *He* values. Finally, in the between-region analyses, we compared ten models generated from the results of the previous analyses. All constructed models are described in Figure S1F–I. Scripts that described models were deposited in Dryad (doi: 10.5061/dryad.mpg4f4qvh).

4.3.3 Prior check

We performed PCA by using *gfitpca* function implemented in *abc* package to check the validity of the prior distribution. We confirmed that the observed summary statistics included of the 90% envelope of each simulation result (data not shown).

4.3.4 Posterior probabilities of models

The posterior probability of each model based on the observed value was estimated using the *neuralnet* method, which is a non-linear regression method involving neural networks. We used *postpr* function implemented in the *abc* package. The regression step was performed using 30 neural networks and 10 hidden layer units. Five hidden layer units were set in the analysis of the archipelago scale, because an error occurred when 10 hidden layers were set. We used 1% of simulations with summary statistic values closest to the observed ones (tolerance rate = 1%). We conducted 10 calculations for each model selection by changing the tolerance rates from 0.1% to 1.0% with increments of 0.1% to evaluate the robustness of the results by using a tolerance rate of 1%.

4.3.5 Confidence in model selection

To evaluate the accuracy of the model selection, we calculated the misclassification rate by using function *cv4postpr* implemented in *abc* package. We randomly generated 100 pseudo-observed data sets (pods) of each model from the prior distribution and performed model selection for the respective pods with 10^5 simulations for each model. Model selection was performed using *neuralnet* method under 10 neural networks and 10 hidden layer units. Five hidden layer units were set in the analysis at the archipelago scale as in the calculation of posterior probabilities. Type 1 and type 2 errors were evaluated for each model: type 1 error rate is the probability that the other model was selected despite the focal model being true, and type 2 error rate is the probability that the focal model was selected despite the other model being true. For evaluating the accuracy of the model selection results, the type 2 error rate should be carefully considered (Cornuet et al., 2010; Lombert et al., 2014).

4.3.6 Estimating posterior distributions of parameters

We calculated the posterior distribution of each parameter under the selected model by using function *abc* implemented in the *abc* package. We retained 1% of simulations close to the observed summary statistics, and we performed regression analysis by using the *neuralnet* method under 30 neural networks and five hidden layer units. Parameters of population size and event time were log transformed before regression analysis.

4.3.7 Posterior predictive check

We checked whether the selected model could produce the observed value under the posterior distribution. We generated 1000 simulations under the posterior distribution of the selected model and calculated summary statistics. In addition to the originally used summary statistics, the following summary statistics were also calculated: the mean allelic range over loci for each population (R_{pop}) and its standard deviation (Rsd_{pop}), the mean allelic range over loci and populations ($MEAN_R$) and its standard deviation ($MEAN_Rsd$), and the mean total allelic range over loci and populations (TOT_R), and F_{IS} (FIS). Next, PCA using function *gfitpca* implemented in *abc* package was performed to check whether the observed summary statistics were included in the range of summary statistics generated from the posterior distribution.

5. Correlation between island size and speciation probability

We first identified the islands on which the speciation of FF occurred based on the results of STRUCTURE (genetic isolation) and ABC model selection (independent origin). The relationship between ecosystem size and speciation was analyzed for the Middle and South Ryukyu Islands between the Tokara Gap and Yonaguni Gap. This analysis was conducted because these islands are assumed to be similar in age, were never connected to any continent after separation, and were not catastrophically affected by volcanic activity (Osozawa et al., 2012). We acquired distribution information for the two forms from published literature (Yoshigou 2014) and our own field survey, and we only chose the islands where either or both forms were distributed. This is because we assumed that speciation was not likely to have occurred on the island where either of these species is not currently distributed. We coded the occurrence of speciation in each island as follows: 1 for the island where speciation occurred; 0 for the island where speciation did not occur. As a proxy of ecosystem size, we used island area (km^2 ; data obtained from Geospatial Information Authority of Japan), the maximum catchment area (km^2), the maximum river length (m), or the maximum number of waterfalls in each island. The data on rivers on the islands where FF is distributed were only obtained for rivers with FF. The latter three parameters were calculated using the 10 m mesh digital elevation model (Geographic Information Authority of Japan) by using ArcGIS10.3 (ESRI Japan, Tokyo). Because of strong positive correlations among the four variables (Pearson's correlation coefficient: island size and catchment area, 0.959; island size and river length, 0.951; catchment area and river length, 0.993; island size

and number of waterfalls, 0.592; catchment area and number of waterfalls, 0.865; river length and number of waterfalls, 0.865), we primarily used island area as the explanatory variable.

We performed one-parameter logistic regression analysis in the generalized linear model framework with binomial error structure for the response variable (speciation: 1, presence; 0, absence) using glm package of R v. 3.1.1 (R Core Team, 2014). Explanatory variables, except the number of waterfalls, were $\log_{10}$ transformed. To test the significance of explanatory variables, we conducted a likelihood ratio test using 10,000 rounds of parametric bootstrap sampling by using a constant model as the null model that assumes no relationship between ecosystem variables and speciation probability. We calculated McFadden's pseudo R^2 by using the following formula: $r^2 = 1 - (SSE/SST)$. In this formula, SSE and SST refer to the deviances of the focal and null models, respectively.

Furthermore, to examine the effect of the ecosystem size parameters other than island size (i.e., maximum catchment area, river length, and number of waterfalls), we constructed another generalized linear model for speciation probability. We used the residuals from a regression between island size and each of the other ecosystem-size parameters, together with the island size and their interaction, as explanatory variables. Stepwise model selections by AIC scores were conducted using *step* function on R.

Results

1. Basic summary statistics

Heterozygosity and allelic richness were generally higher in the AF populations than in the FF ones, possibly suggesting a larger current or historical population size in the former (AF: He , 0.296–0.554 and R_s , 1.94–3.65; FF: He , 0.070–0.390 and R_s , 1.21–3.08; Supporting Information Table S1, Figure S2A, B). The AF populations showed relatively little pairwise genetic differentiation within-islands ($F_{ST} = -0.008$ –0.145; Jost's $D = -0.009$ –0.102; Supporting Information Tables S4, 5, Figure S2C, D). In the case of the among-islands within-regions scale, pairwise genetic differentiation was moderate (F_{ST} , 0.022–0.168; Jost's D , 0.025–0.219). AMOVA showed significant differentiation in the among-rivers hierarchy in both Middle Ryukyu and South Ryukyu, and in the among-islands hierarchy in Middle Ryukyu (Supporting Information Table

S6). Most of the variation was explained by within-individuals hierarchy. Compared with this, the genetic differentiation of AF populations between regions was large (F_{ST} , 0.110–0.456; Jost's D , 0.124–0.613). AMOVA showed significant differentiation in both among-islands and among-regions hierarchies (Supporting Information Table S6). Variations were mainly explained by within-individuals, among-islands, and among-regions hierarchies. The FF populations often showed large genetic differentiation within-islands (F_{ST} , 0.003–0.804; Jost's D , 0.001–0.529; Supporting Information Tables S4, 5, Figure S2C, D) and among-islands within-regions (F_{ST} , 0.312–0.859; Jost's D , 0.146–0.778), as well as among-regions (F_{ST} , 0.507–0.914; Jost's D , 0.580–0.890). AMOVA for FF showed significant genetic differentiation in the hierarchies of among-individuals, among-rivers, among-islands, and among-regions, except among-islands in South Ryukyu (Supporting Information Table S6). In the among-islands within-regions scale analysis, genetic variation was mainly explained by within-individuals, among-rivers, and among-islands hierarchies.

2 Genetic isolation and hybrids

In the AF dataset, Evanno's ΔK showed the maximum value at $K = 2$; the populations in North plus Middle Ryukyu and South Ryukyu were clearly separated. In the dataset for each island, including FF and AF populations, ΔK had the maximum value at $K = 2$ for all islands, except one (Amami-Oshima Is.). The two genetic clusters corresponded to the two forms, except on Okinawajima Is. (see below; Figure 3 and Supporting Information Figure S3). ΔK on Amami-Oshima Is. showed the maximum value at $K = 3$, in which FF was further divided into two clusters. Suspected hybrid individuals (q -value, $\geq 10\%$) were rare (2.0% of total individuals). No suspected hybrid individuals were detected from five of the seven sympatric collection points. Several suspected hybrid individuals were detected in 11 FF and AF populations, including two FF populations on Tokunoshima Is. (FFTO-4-2, FFTO-5), three FF on Okinawajima Is. (FFOK-3, FFOK-4-1, and FFOK-4-2), one FF on Ishigakijima Is. (FFIS-1), three FF on Iriomotejima Is. (FFIR-3-1, FFIR-6, and FFIR-8), and two AF on Okinawajima Is. (AFOK-1, AFOK-3). Among these, all FF individuals that were identified based on morphology in the FFOK-4-1 and FFOK-4-2 populations (Okinawajima Is.; see Figures 2, 3 and Supporting Information Figure S3) exhibited the genetic features of AF as the dominant component. Except for these two populations, the proportion of suspected

hybrid individuals ranged from 3.6% to 21.2% of individuals in a population. Potential hybrids found in the FF populations contained 10.1%–80.5% of the AF component, whereas those of the AF populations contained 25.6%–69.0% of the FF component.

3 Genetic relationships among populations

The neighbor-joining tree showed two divergent groups separated by the Kerama Gap in all of the FF and AF populations (bootstrap value, 97%; Figure 4). Each group included both forms, which we present as Group 1 (Honshu, North Ryukyu, and Middle Ryukyu) and Group 2 (South Ryukyu), hereafter.

In both groups, FF populations on each island formed a monophyletic group (bootstrap value, 89%–100%), except those on two islands in Group 1 (Amami-Oshima Is., 67%; Okinawajima Is., 33%). In addition, FF populations on different islands (Tokunoshimajima Is. and Okinoerabujima Is.) formed a monophyletic group (89%). The AF populations within each group were closely related to each other, and the populations of each island in Group 1 tended to form a monophyletic group.

4 Test for multiple origins of FF

In the within-island scale analysis, the single origin model of FF was selected for all islands (Table 1 and Supporting Information Table S7; Figure 4 and Supporting Information Figure S4). Models including gene flow among all or partial populations of the two forms showed higher posterior probability than those not considering gene flow (Table 1 and Supporting Information Table S7; Figure 4 and Supporting Information Figure S4). Type 2 error rates for the selected single origin model against alternative parallel origin models were generally low (Table 1 and Supporting Information Table S7; Figure 4 and S4).

At the among-island scale, Group 1 in Middle Ryukyu and Group 2 in South Ryukyu were analyzed separately. For Group 1, we first examined the single origin of FF in the adjoining islands, Tokunoshima Is. and Okinoerabujima Is., which was suggested by the phylogenetic analysis. ABC analysis supported their single origin (posterior probability (PP), 0.746; Bayes factor (BF), 9.42–122.57; Type 2 error rate, 0.06–0.1; BFs and Type 2 error rates are of the selected model against parallel origin models; Table 1 and Supporting Information Table S7; Figure S4); thus, only FF on

Tokunoshima Is. was used for subsequent analyses. The model selection for Group 1 strongly supported the four independent origins of FF (PP, 0.447; BF, 23.97–519.23; Type 2 error rate, 0; BFs and Type 2 error rates are of the selected model against the other non-four origin models; Tables 1 and S7; Figure S4). The model with gene flow between the two forms obtained higher BF than the model not considering gene flow (BF, 4.72; Type 2 error rate, 0.125). In Group 2, the single origin of FF of Iriomotejima Is. and Ishigakijima Is. was supported (PP, 0.895; BF, 36.98–191.0; Type 2 error rate, 0–0.08; BFs and Type 2 error rates are of the selected model against parallel origin models Tables 1 and S7; Figure S4). The model including gene flow between the AF and FF populations was better supported than that not considering gene flow (BF, 13.63; Type 2 error rate, 0.06).

For the between-region scale, the neighbor-joining tree for populations suggested independent origins of FF in Groups 1 and 2. However, we could not obtain a conclusion at this scale from ABC, because no single model was successfully selected, probably because of insufficient data (PP, 0.0024–0.14; Table S7; Figure S4).

5 Correlation between speciation probability and ecosystem size

We confirmed the occurrence of AF or FF populations on 18 islands from the published literature and our field survey (Supporting Information Table S8). As we could not conclude from the ABC analysis whether the FF population originated independently between Groups 1 and 2, we analyzed independent and non-independent scenarios between the Middle and South Ryukyus. In addition, because the single origin of FF was supported for Tokunoshima Is. and Okinoerabujima Is. populations, as well as for Ishigakijima Is. and Iriomotejima Is. populations, we could not conclude which island FF originated in. We developed four datasets for the independent scenario between Groups 1 and 2 with all possible combinations of islands in which speciation would have occurred independently: dataset 1, speciation in Tokunoshima Is. and Iriomotejima Is.; dataset 2, in Tokunoshima Is. and Ishigakijima Is.; dataset 3, Okinoerabujima Is. and Iriomotejima Is.; and dataset 4, Okinoerabujima Is. and Ishigakijima Is. We also developed 16 datasets for the non-independent scenario (Supporting Information Table S9).

For all four datasets in the scenario of independent origins between the regions, speciation probability was significantly explained by island area ($P = 0.0016$ – 0.0061 ,

difference of deviance = 8.650–12.265, $r^2 = 0.407$ – 0.577 ; Table 2 and Supporting Information Table S9; Figure 5 and S5). The relationship was also significant for 12 of the 16 datasets in the non-independent scenario (Supporting Information Table S9). Almost identical results were obtained for the maximum catchment area and river length (Table 2 and Supporting Information Table S9). Furthermore, two of the four datasets for the independent origin scenario and five of the 16 datasets for the non-independent scenario, showed significant relationships between the number of waterfalls and speciation (Table 2 and Supporting Information Table S9). Model selection by AIC scores showed that the best model included no variables, except for island area, as the predictors for speciation probability in most datasets when the variables were controlled by the island area (Supporting Information Table S9).

References

- Akihito, Sakamoto, K., Ikeda, Y., & Aizawa, M. (2013). Gobioidae. In T. Nakabo (Ed.), *Fishes of Japan with Pictorial Keys to Species (Third ed.)*. (pp. 1374–1608). Tokyo, Japan: Tokai University Press. (in Japanese).
- Beaumont, M. A. (2010). Approximate Bayesian computation in evolution and ecology. *Annual Review of Ecology, Evolution, and Systematics*, 41, 379–406.
<https://doi.org/10.1146/annurev-ecolsys-102209-144621>
- Bertorelle, G., Benazzo, A., & Mona, S. (2010). ABC as a flexible framework to estimate demography over space and time: some cons, many pros. *Molecular Ecology*, 19, 2609–2625. <https://doi.org/10.1111/j.1365-294X.2010.04690.x>
- Blacket, M. J., Robin, C., Good, R. T., Lee, S. F., & Miller, A. D. (2012). Universal primers for fluorescent labelling of PCR fragments—an efficient and cost-effective approach to genotyping by fluorescence. *Molecular Ecology Resources*, 12, 456–463. <https://doi.org/10.1111/j.1755-0998.2011.03104.x>
- Cornuet, J.-M., Ravigné, V., & Estoup, A. (2010). Inference on population history and model checking using DNA sequence and microsatellite data with the software DIYABC (v1.0). *BMC Bioinformatics*, 11, 401.
<https://doi.org/10.1186/1471-2105-11-401>

- Csilléry, K., François, O., & Blum, M. G. B. (2012). abc: an R package for approximate Bayesian computation (ABC). *Methods in Ecology and Evolution*, 3, 475–479. <https://doi.org/10.1111/j.2041-210X.2011.00179.x>
- Earl, D. A., & vonHoldt, B. M. (2012). STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation Genetics Resources*, 4, 359–361. <https://doi.org/10.1007/s12686-011-9548-7>
- Evanno, G., Regnaut, S., & Goudet, J. (2005). Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Molecular Ecology*, 14, 2611–2620. <https://doi.org/10.1111/j.1365-294X.2005.02553.x>
- Excoffier, L., & Lischer, H. E. L. (2010). Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Molecular Ecology Resources*, 10, 564–567. <https://doi.org/10.1111/j.1755-0998.2010.02847.x>
- Excoffier, L., & Foll, M. (2011). fastsimcoal: a continuous-time coalescent simulator of genomic diversity under arbitrarily complex evolutionary scenarios. *Bioinformatics*, 27, 1332–1334. <https://doi.org/10.1093/bioinformatics/btr124>
- Excoffier, L., Dupanloup, I., Huerta-Sánchez, E., Sousa, V. C., & Foll, M. (2013). Robust Demographic Inference from Genomic and SNP Data. *PLoS Genetics*, 9, e1003905–17. <https://doi.org/10.1371/journal.pgen.1003905>
- Goudet, J. (1995). Fstat version 1.2: a computation program to calculate Fstatistics. *Journal of Heredity*, 86, 485–486. <https://doi.org/10.1093/oxfordjournals.jhered.a111627>
- Hirashima, K., & Tachihara, K. (2000). Embryonic development and morphological changes in larvae and juveniles of two land-locked gobies, *Rhinogobius* spp. (Gobiidae), on Okinawa Island. *Japanese Journal of Ichthyology*, 47, 29–41 (in Japanese with English abstract). <https://doi.org/10.11369/jji1950.47.29>
- Holleley, C., & Geerts, P. (2009). Multiplex Manager 1.0: a cross-platform computer program that plans and optimizes multiplex PCR. *Biotechniques*, 46, 511–517. <https://doi.org/10.2144/000113156>
- Iguchi, K., & Mizuno, N. (1999). Early starvation limits survival in amphidromous fishes. *Journal of Fish Biology*, 54, 705–712. <https://doi.org/10.1111/j.1095-8649.1999.tb02027.x>

- Kano, Y., Nishida, S., & Nakajima, J. (2012). Waterfalls drive parallel evolution in a freshwater goby. *Ecology and Evolution*, 2, 1805–1817.
<https://doi.org/10.1002/ece3.295>
- Kon, T., & Yoshino, T. (2003). Coloration and ontogenetic features of fluviatile species of *Rhinogobius* (Gobioidei: Gobiidae in Amami-oshima Island, Ryukyu Islands, Japan. *Ichthyological Research*, 50, 109–116.
<https://doi.org/10.1007/s10228-002-0145-3>
- Kondo, M., Maeda, K., Hirashima, K., & Tachihara, K. (2013). Comparative larval development of three amphidromous *Rhinogobius* species, making reference to their habitat preferences and migration biology. *Marine and Freshwater Research*, 64, 249–266. <https://doi.org/10.1071/MF12234>
- Lombert, E., Guillemaud, T., Lundgren, J., Koch, R., Facon, B., Grez, A., ...Estoup, A. (2014). Complementarity of statistical treatments to reconstruct worldwide routes of invasion: the case of the Asian ladybird *Harmonia axyridis*. *Molecular Ecology*, 23, 5979–5997. <https://doi.org/10.1111/mec.12989>
- Machida, H., Ota, Y., Kawana, T., Moriwaki, H., & Nagaoka, N. (2001). *Regional Geomorphology of the Japanese Island vol. 7 Kyushu and Southern Islands*. Tokyo, Japan: University of Tokyo Press. (in Japanese)
- Matschiner, M., & Salzburger, W. (2009). TANDEM: integrating automated allele binning into genetics and genomics workflows. *Bioinformatics* 25:1982–1983.
<https://doi.org/10.1093/bioinformatics/btp303>
- Meglecz, E., Costedoat, C., Dubut, V., Gilles, A., Malausa, T., Pech, N., & Martin, J. F. (2010). QDD: a user-friendly program to select microsatellite markers and design primers from large sequencing projects. *Bioinformatics*, 26, 403–404.
<https://doi.org/10.1093/bioinformatics/btp670>
- Meirmans, P. G., & van Tienderen, P. H. (2004). genotype and genodive: two programs for the analysis of genetic diversity of asexual organisms. *Molecular Ecology Notes*, 4, 792–794. <https://doi.org/10.1111/j.1471-8286.2004.00770.x>
- Mizuno, N. (2001). *Rhinogobius*. In H. Kawanabe, N. Mizuno, & K. Hosoya (Eds.), *Freshwater fishes of Japan* (Third ed.). (p. 584) Tokyo, Japan: Yama-Kei Publishers (in Japanese).

- Moriyama, A., Yanagisawa, Y., Mizuno, N., & Omori, K. (1998). Starvation of drifting goby larvae due to retention of free embryos in upstream reaches. *Environmental Biology of Fishes*, 52, 321–329. <https://doi.org/10.1023/A:1007333302864>
- Ohara, K., Takahashi, D., & Takagi, M. (2004). Isolation of microsatellite loci in the freshwater goby, *Rhinogobius* sp. (Gobiidae). *Molecular Ecology Notes*, 4, 449–451. <https://doi.org/10.1111/j.1471-8286.2004.00677.x>
- Oijen, M. J. P., Suzuki, T., & Chen, I. S. (2011). On the earliest published species of *Rhinogobius*. With a redescription of *Gobius brunneus* Temminck and Schlegel, 1845. *Journal of the National Taiwan Museum*, 64, 1–17.
- Omori, K., Ohnishi, H., Hamaoka, H., Kunihiro, T., Ito, S., Kuwae, M., ... Iguchi K. (2012). Speciation of fluvial forms from amphidromous forms of migratory populations. *Ecological Modelling*, 243, 89–94. <https://doi.org/10.1016/j.ecolmodel.2012.06.006>
- Oshiro, N. (1981). Ryukyu rettousan Yoshinobori no yomoyamabanashi (Some topics about *Rhinogobius* fishes in Ryukyu Island). *Tansuigyo*, 7, 14–18 (in Japanese).
- Osozawa, S., Shinjo, R., & Armid, A. (2012). Palaeogeographic reconstruction of the 1.55 Ma synchronous isolation of the Ryukyu Islands, Japan, and Taiwan and inflow of the Kuroshio warm current. *International Geology Review*, 54, 1369–1388. <https://doi.org/10.1080/00206814.2011.639954>
- Ota, H. (1998). Geographic patterns of endemism and speciation in amphibians and reptiles of the Ryukyu Archipelago, Japan, with special reference to their paleogeographical implications. *Researches on Population Ecology*, 40, 189–204. <https://doi.org/10.1007/BF02763404>
- Pritchard, J. K., Stephen, M., & Donnelly, P. (2000). Inference of population genetic structure using multilocus genotype data. *Genetics*, 155, 945–959.
- Shinomiya, A., Sasabe, K., Sakurai, M., & Kishino, T. (2005). Morphological characters of newly hatched larvae and their habitat selection of a fluvatile goby, *Rhinogobius* sp. YB in the upper stream of Sumiyo River, Amami-oshima Island, Japan. *Japanese Journal of Ichthyology*, 52, 1–8 (in Japanese with English abstract). <https://doi.org/10.11369/jji1950.52.1>
- Suzuki, T., Shibukawa, K., Yano, K., & Senou, H. (2004). A photographic giude to the gobioid fishes in Japan. Tokyo, Japan: Heibonsha (in Japanese).
- Takeshima, H., Muto, N., Sakai, Y., Ishiguro, N., Iguchi, K., Ishikawa, S., & Nishida, M.

- (2017). Rapid and effective isolation of candidate sequences for development of microsatellite markers in 30 fish species by using kit-based target capture and multiplexed parallel sequencing. *Conservation Genetics Resources*, 9, 479–490. <https://doi.org/10.1007/s12686-017-0730-4>
- Takezaki, N., Nei, M., & Tamura, K. (2010). POPTREE2: Software for constructing population trees from allele frequency data and computing other population statistics with Windows interface. *Molecular Biology and Evolution*, 27, 747–752. <https://doi.org/10.1093/molbev/msp312>
- Tamada, K. (2005). Clutch and egg size in *Rhinogobius* sp. DA inhabiting small rivers. *Japanese Journal of Ichthyology*, 52, 17–20 (in Japanese with English abstract). <https://doi.org/10.11369/jji1950.52.17>
- Vähä, J.-P., & Primmer, C. R. (2005). Efficiency of model-based Bayesian methods for detecting hybrid individuals under different hybridization scenarios and with different numbers of loci. *Molecular Ecology*, 15, 63–72. <https://doi.org/10.1111/j.1365-294X.2005.02773.x>
- van Oosterhout, C., Hutchinson, W. F., Willis, D. P. M., & Shipley, P. (2004). MICRO-CHECKER; software for identifying and correcting genotyping errors in microsatellite data. *Molecular Ecology Notes*, 4, 535–538. <https://doi.org/10.1111/j.1471-8286.2004.00684.x>
- Wegmann, D., Leuenberger, C., Neuenschwander, S., & Excoffier, L. (2010). ABCtoolbox: a versatile toolkit for approximate Bayesian computations. *BMC Bioinformatics*, 11, 116. <https://doi.org/10.1186/1471-2105-11-116>
- Yamasaki, Y. Y., Nishida, M., Suzuki, T., Mukai, T., & Watanabe, K. (2015). Phylogeny, hybridization, and life history evolution of *Rhinogobius* gobies in Japan, inferred from multiple nuclear gene sequences. *Molecular Phylogenetics and Evolution*, 90, 20–33. <https://doi.org/10.1016/j.ympev.2015.04.012>
- Yoshigou, H. (2014). Annotated checklist and bibliographic records of inland water fishes of the Ryukyu Archipelago, Japan. *Fauna Ryukyana*, 9, 1–153. http://w3.u-ryukyu.ac.jp/naruse/lab/Fauna_Ryukyuna.html
- Yue, G. H., David, L., & Orban, L. (2007). Mutation rate and pattern of microsatellites in common carp (*Cyprinus carpio* L.). *Genetica*, 129, 329–331 <https://doi.org/10.1007/s10709-006-0003-8>

A

2. Amami-Oshima Is.

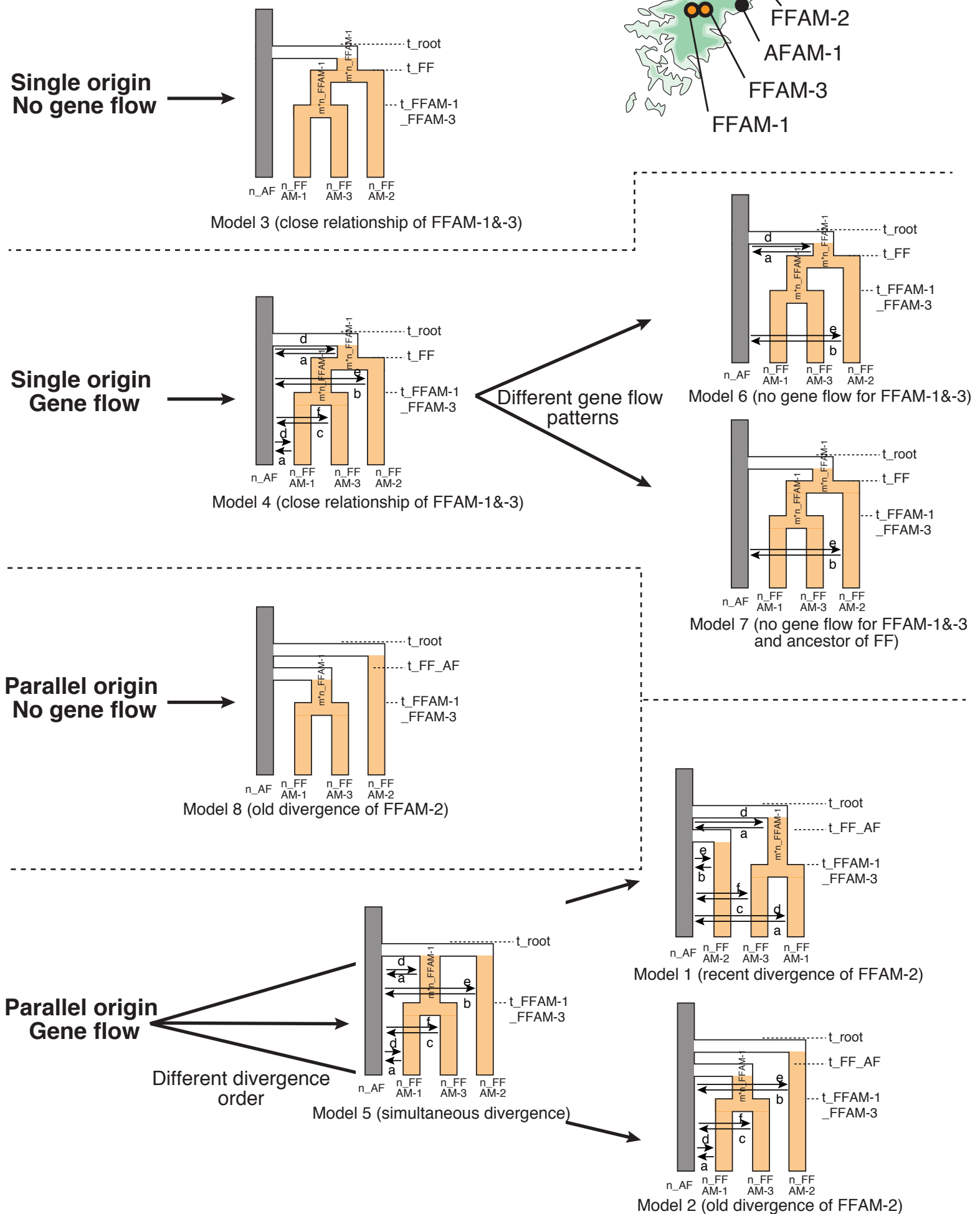


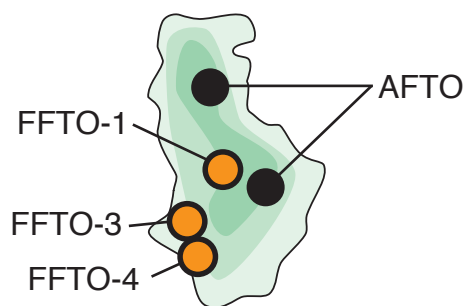
Figure S1 Tested demographic models and their relationships.

B

4. Tokunoshima Is.

**Single origin
No gene flow**

Model 5 (simultaneous divergence among FF)



Different divergence
order

Model 6 (close relationship of FFTO-3&-4)

**Single origin
Gene flow**

Different divergence order

Model 1 (simultaneous divergence among FF)

**Parallel origin
No gene flow**

Model 7 (old divergence of FFTO-1)

Model 3 (close relationship of FFTO-3&-4)

**Parallel origin
Gene flow**

Different divergence order

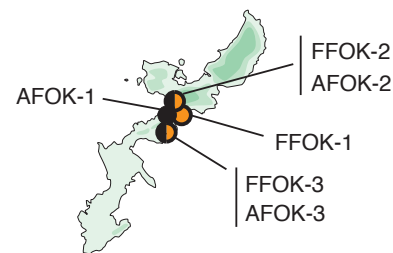
Model 2 (simultaneous divergence)

Model 4 (old divergence of FFTO-1)

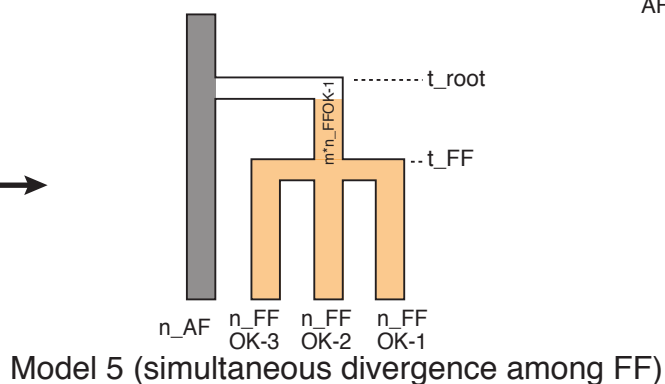
Figure S1 Continued.

C

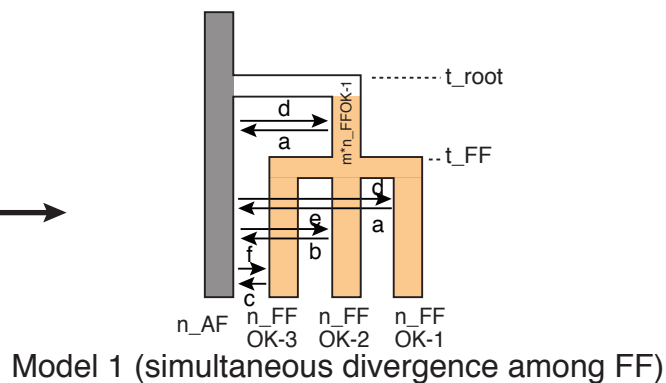
6. Okinawajima Is.



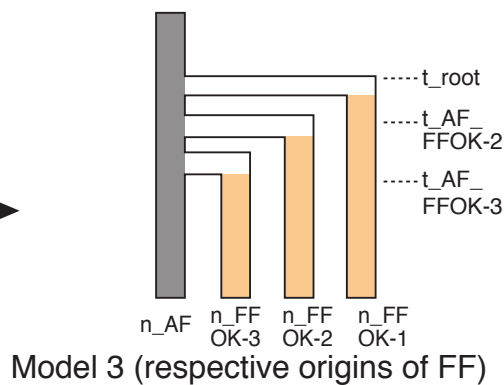
Single origin
No gene flow



Single origin
Gene flow



Parallel origin
No gene flow



Parallel origin
Gene flow

Different divergence order

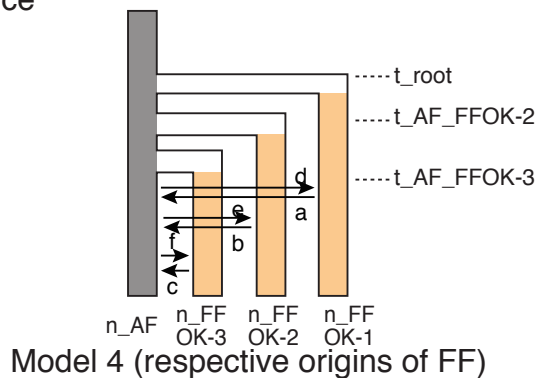
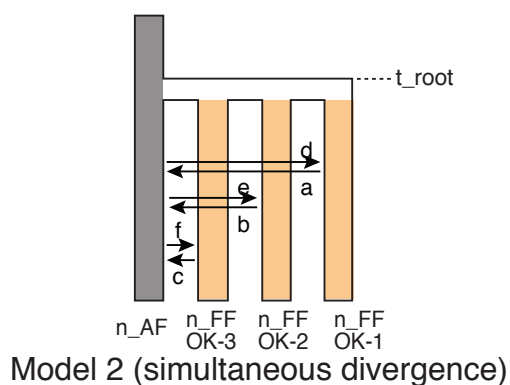


Figure S1 Continued.

D

7. Kumejima Is.

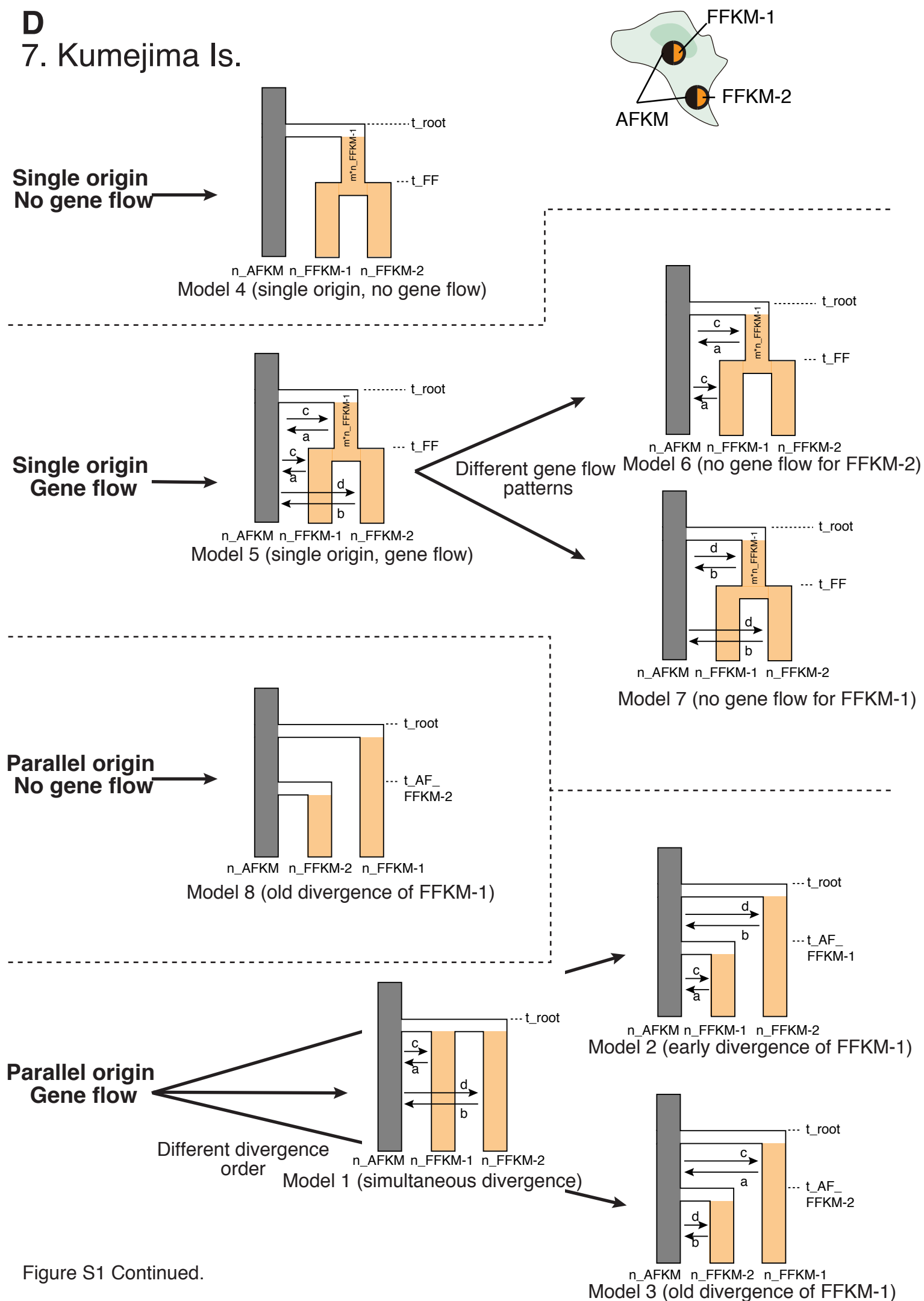


Figure S1 Continued.

E

9. Iriomotejima Is.

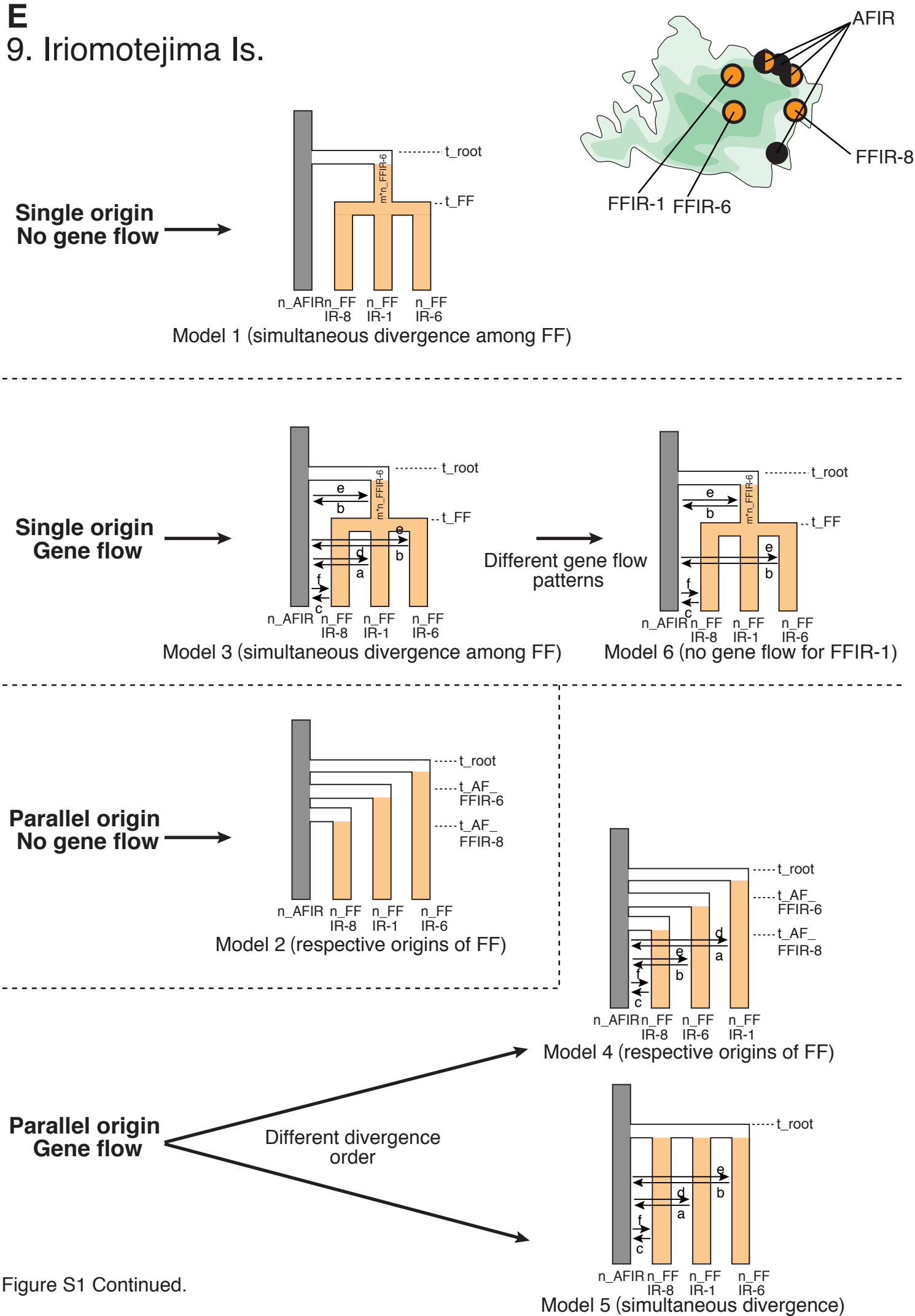
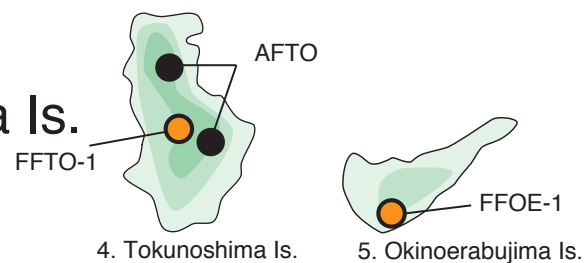


Figure S1 Continued.

F

4. Tokunoshima Is.– 5. Okinoerabujima Is.



Single origin

Different gene flow patterns

Model 1 (single origin, no gene flow)

Model 2 (single origin, gene flow)

Parallel origin

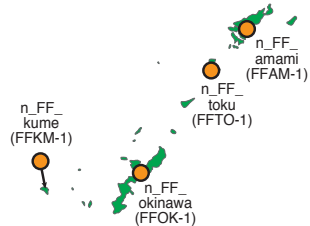
Different divergence order

Model 3 (old divergence of FFTO-1)

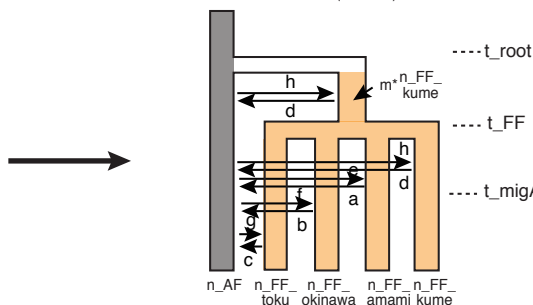
Model 4 (simultaneous divergence)

Model 5 (recent divergence of FFTO-1)

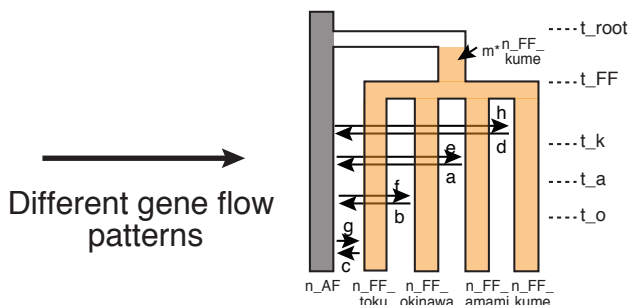
G Middle Ryukyu



Single origin

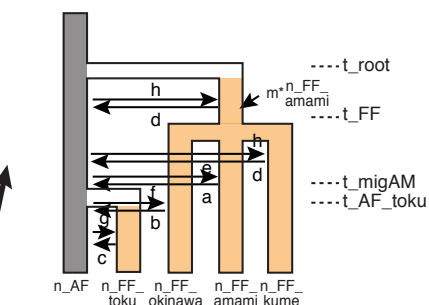


Model 3 (simultaneous divergence among FF)

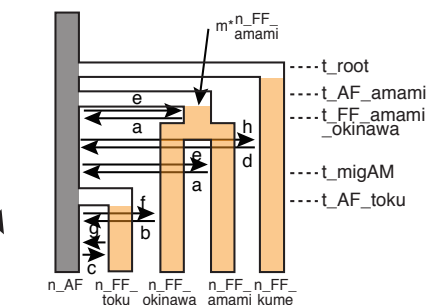


Model 7 (different gene flow timing among FF)

Two origins of FF



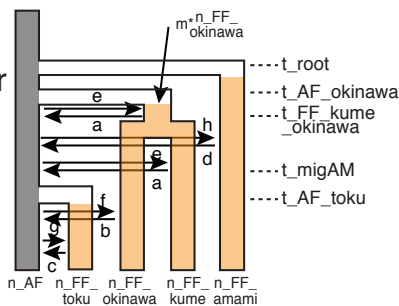
Model 1 (recent divergence of FF_toku and shared origin of others)



Model 5 (recent divergence of FF_toku and old divergence of FF_kume)

Three origins of FF

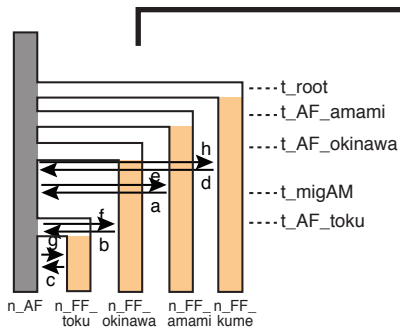
Different divergence order



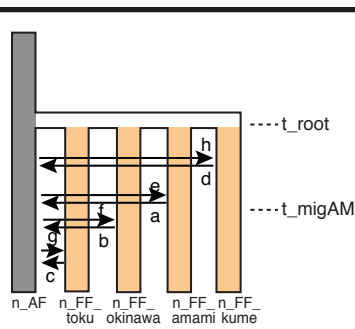
Model 10 (recent divergence of FF_toku and old divergence of FF_amami)

Parallel origin

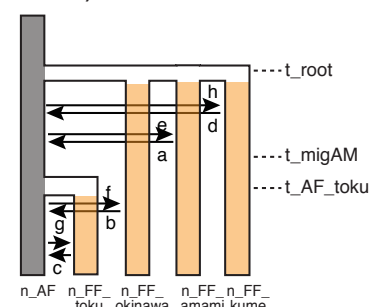
Four origins of FF



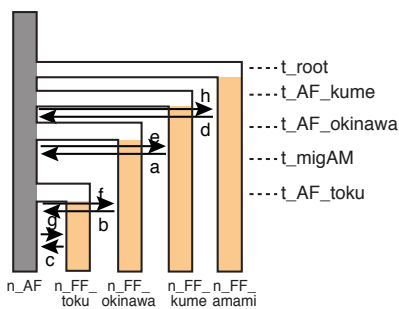
Model 2 (old divergence of FF_kume)



Model 4 (simultaneous divergence)

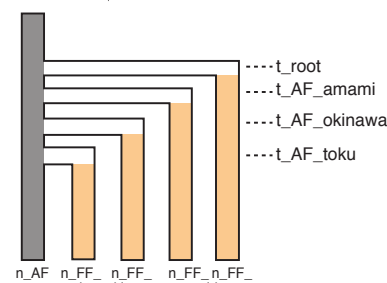


Model 6 (simultaneous divergence of the three pops)



Model 9 (old divergence of FF_amami)

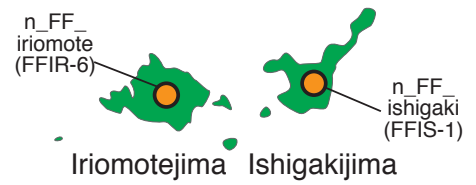
Different gene flow patterns



Model 8 (respective origins with no gene flow)

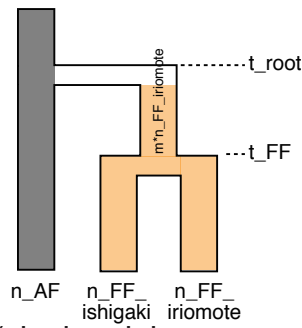
Figure S1 Continued.

H South Ryukyu

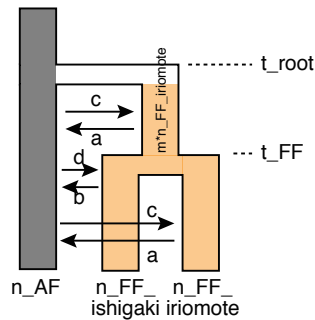


Single origin

Model 1 (single origin, no gene flow)

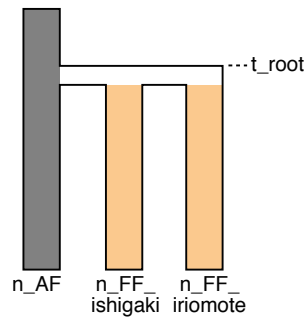


Model 5 (single origin, gene flow)

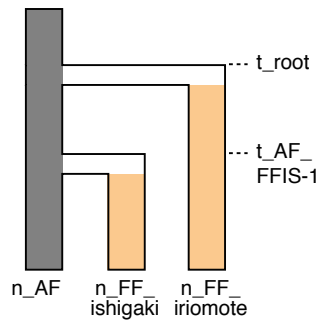


Parallel origin

Model 2 (simultaneous divergence among FF, no gene flow)



Model 3 (old divergence of FF_iriomote, no gene flow)



Model 4 (old divergence of FF_iriomote, gene flow)

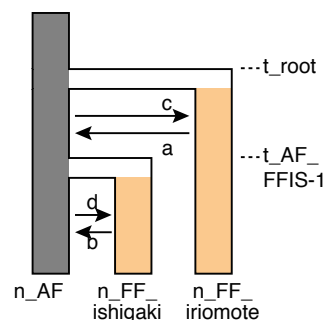


Figure S1 Continued.

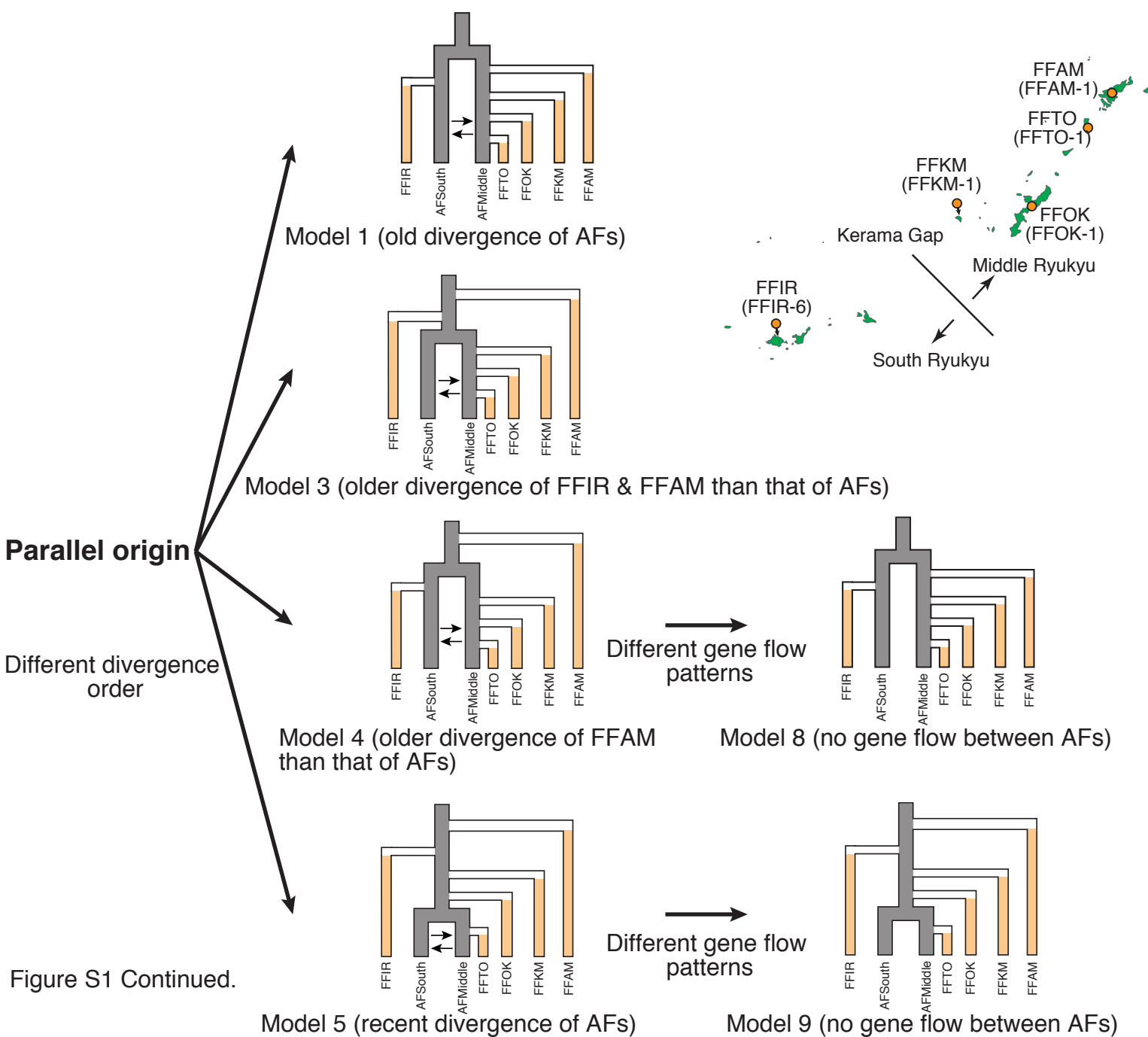
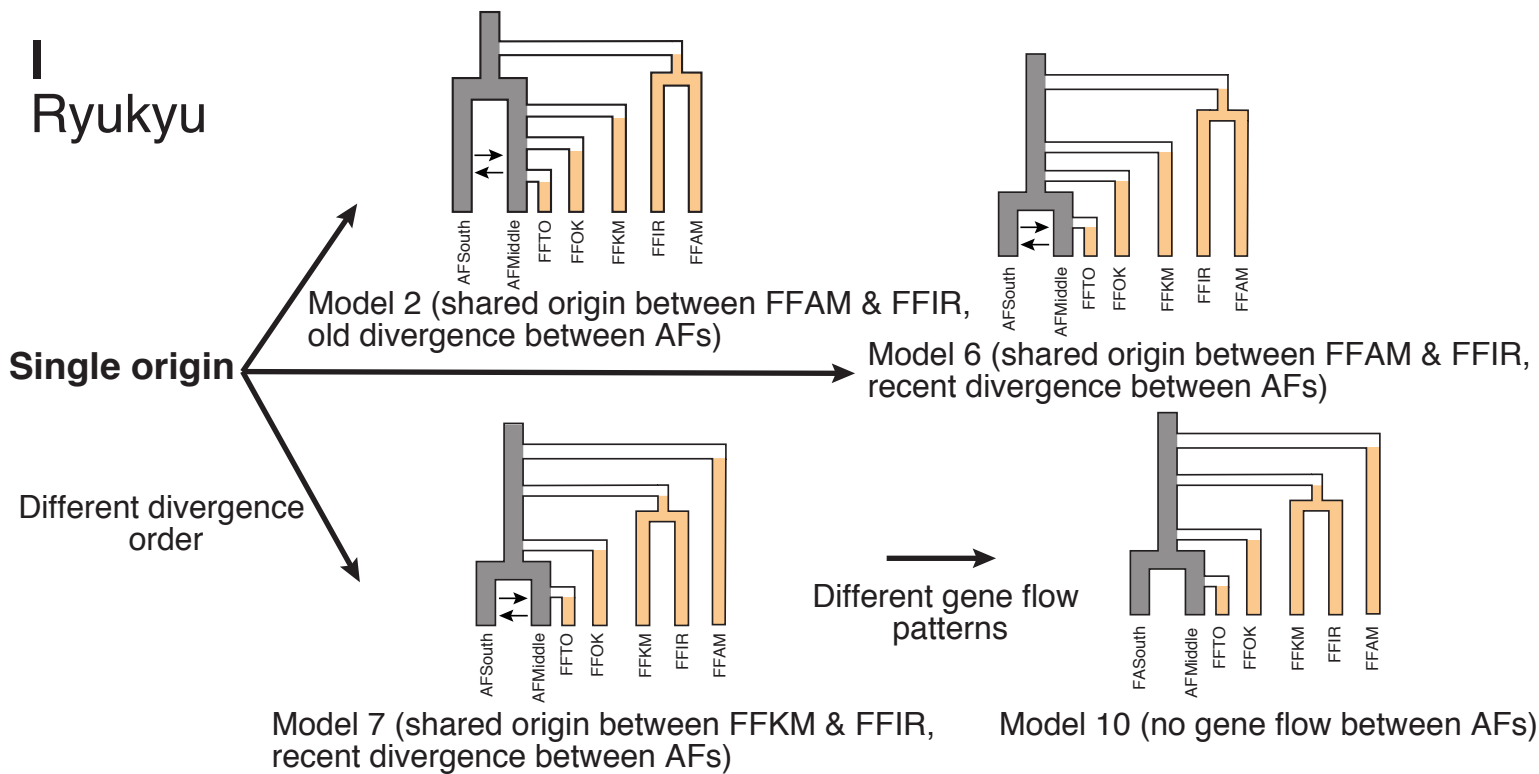


Figure S1 Continued.

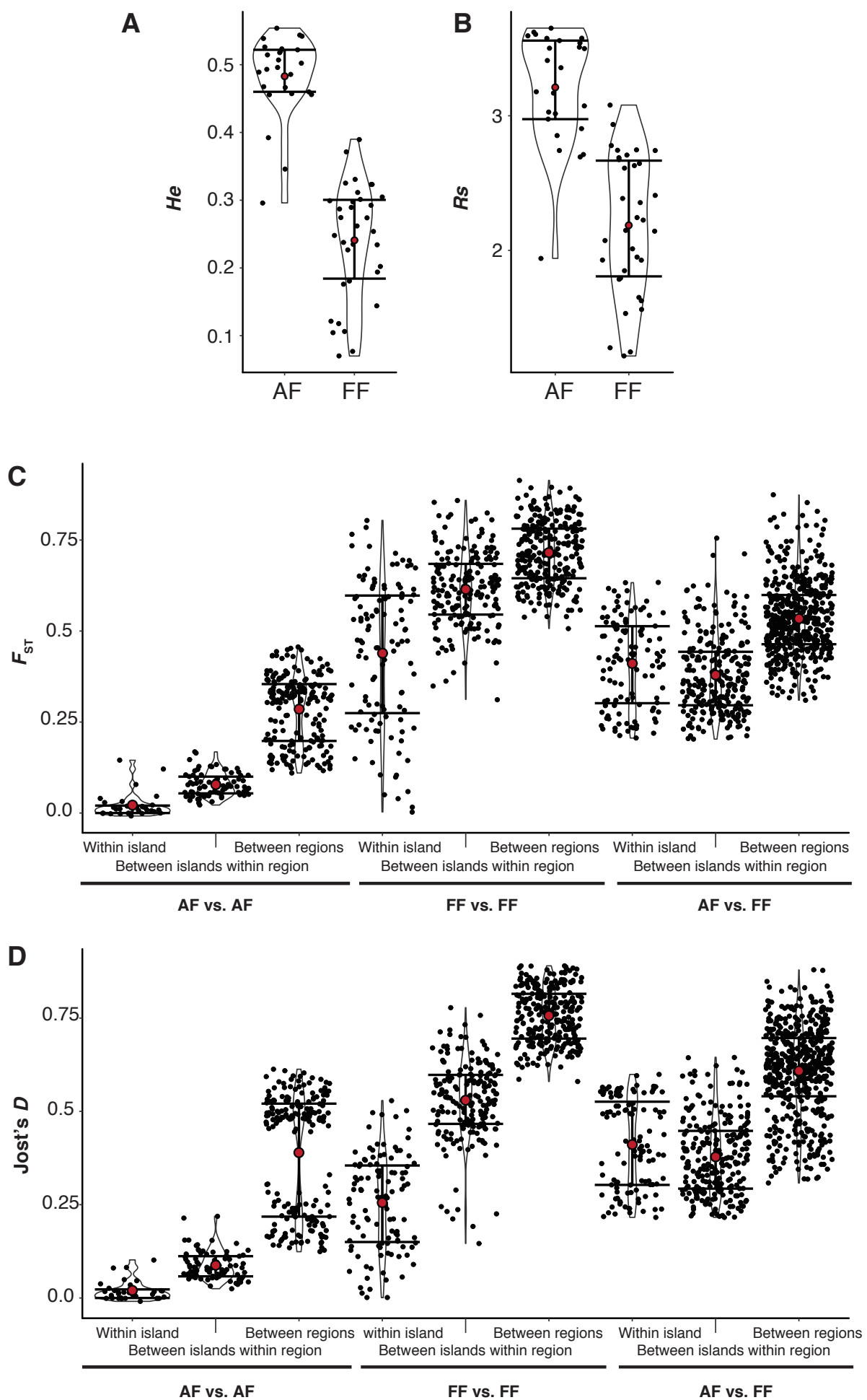


Figure S2 Distributions of summary statistics for genetic diversities within and between populations of the *Rhinogobius* species. Red dots and error bars indicate mean values and intermedian quantiles ranges, respectively. Genetic diversity values of populations were measured by expected heterozygosity (H_e , A) and allelic richness (R_s , B). Pairwise population differences were measured by F_{ST} (C) and Jost's D (D).

A
2. Amami-Oshima Is. & 3. Kikaijima Is.

$\Delta K = \text{mean}(|L''(K)|) / \text{sd}(L(K))$

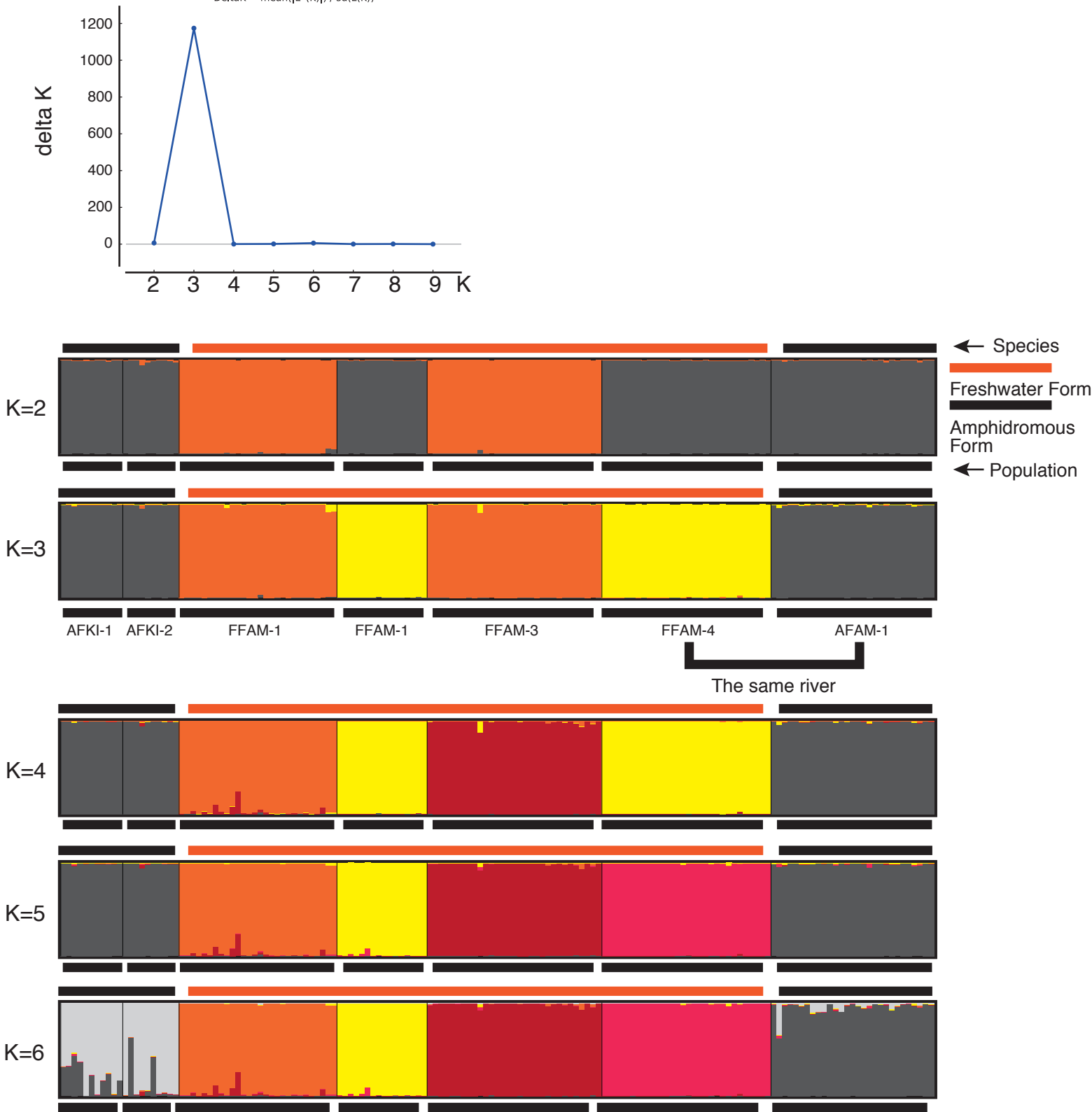


Figure S3 Results of STRUCTURE analysis for each island and AF. Sympatric points are indicated by additional annotation. Dominant genetic features in AF identified by morphology are shown in grey, whereas those in FF are shown in the other colours. Graphs of ΔK are also shown.

B
4. Tokunoshima Is.

$\Delta K = \text{mean}(|L'(K)|) / \text{sd}(L(K))$

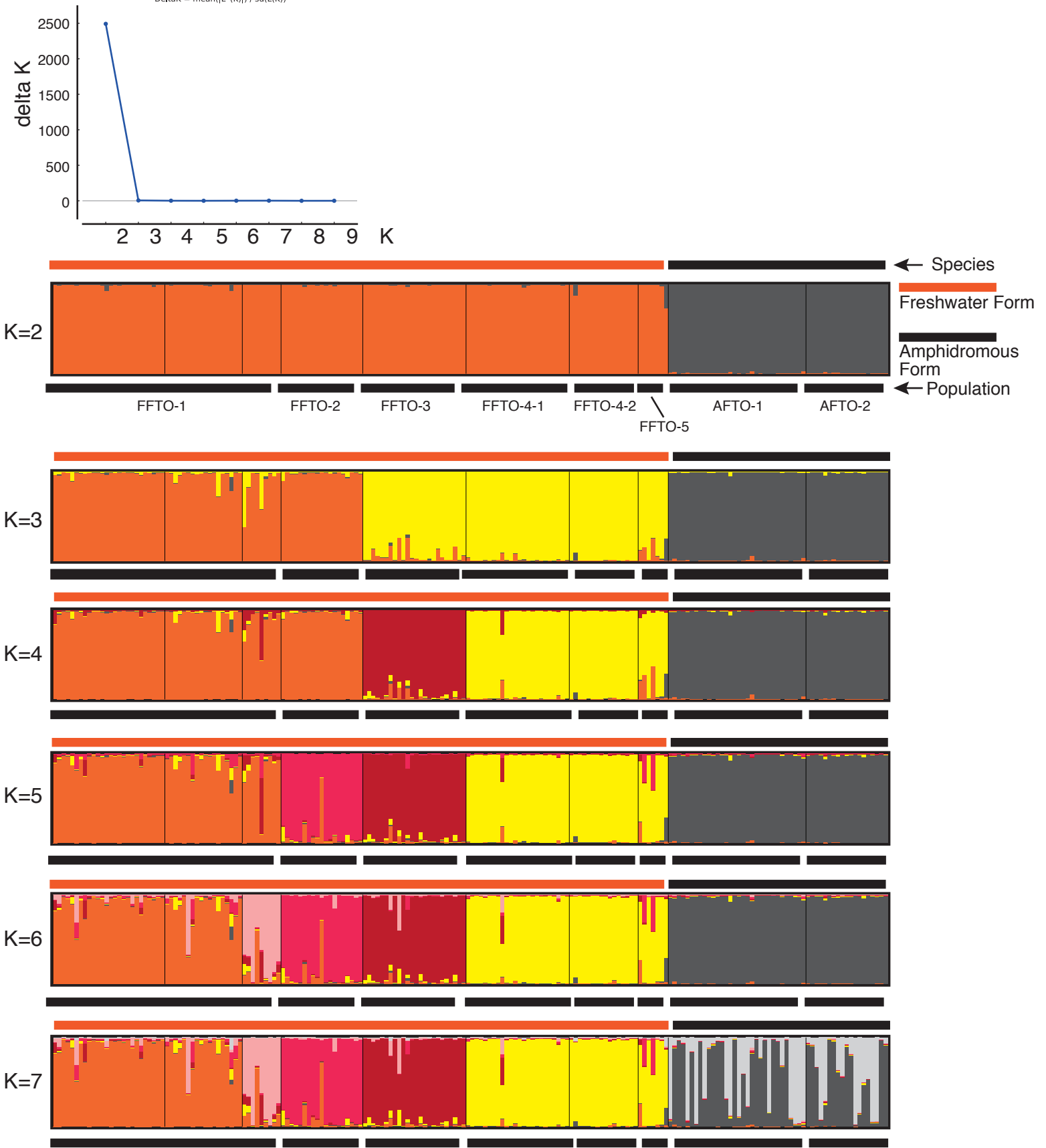


Figure S3 Continued.

C
5. Okinoerabujima Is. (but specimens of *R. brunneus* collected in Tokunoshima Is. were used)

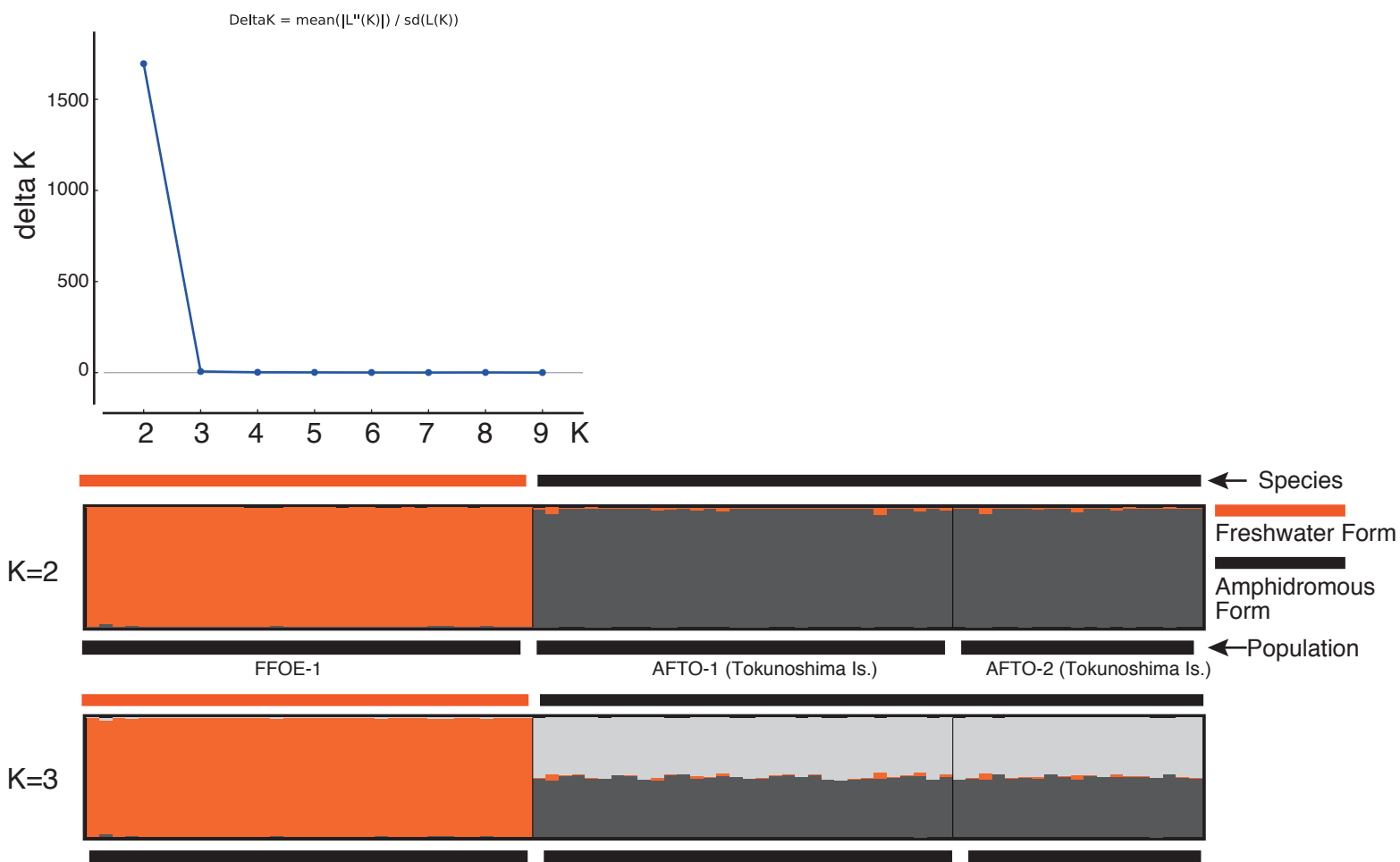


Figure S3 Continued.

D
6. Okinawajima Is.

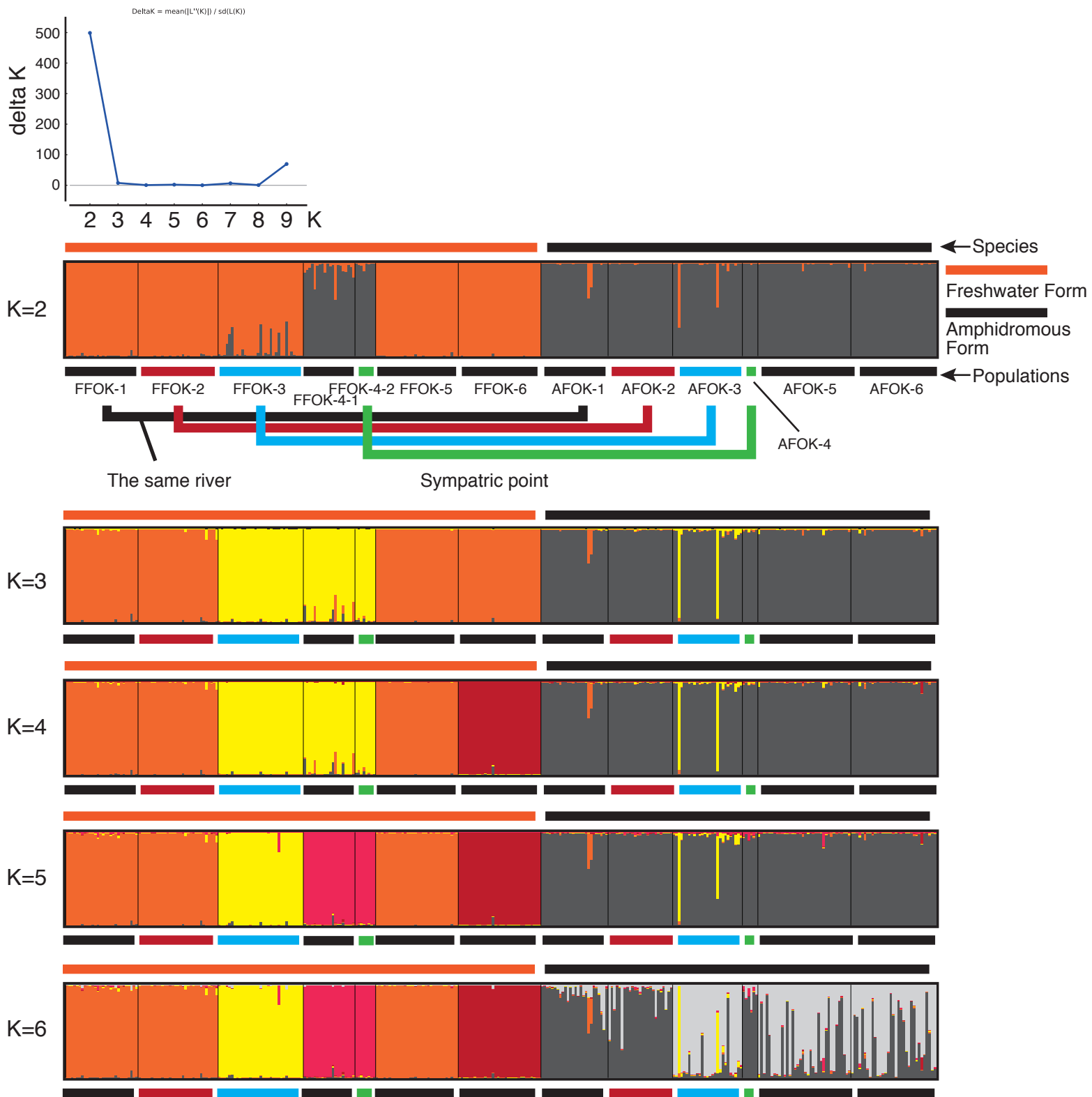


Figure S3 Continued.

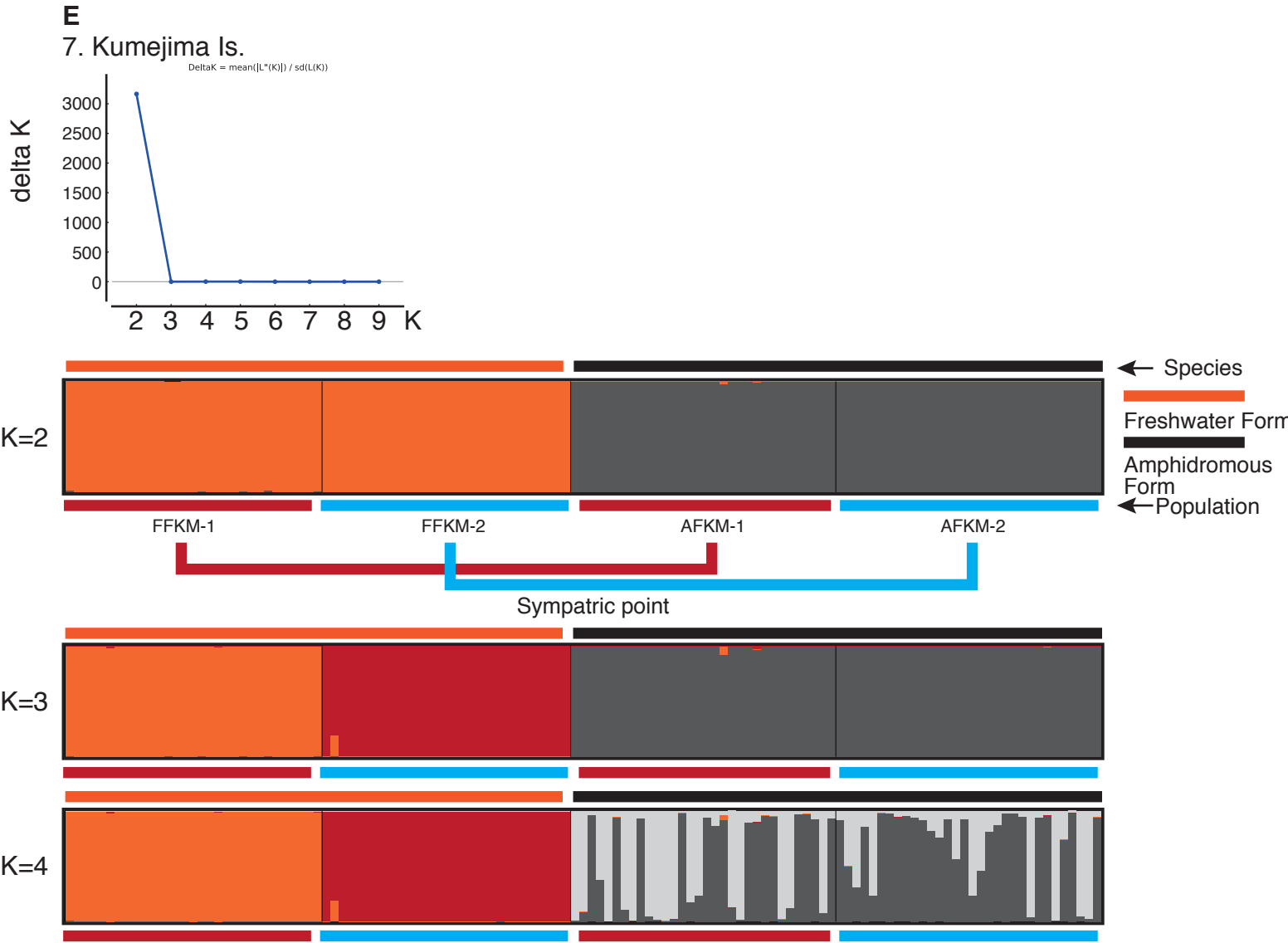


Figure S3 Continued.

F
8. Ishigakijima Is.

$\Delta K = \text{mean}(|L''(K)|) / \text{sd}(L(K))$

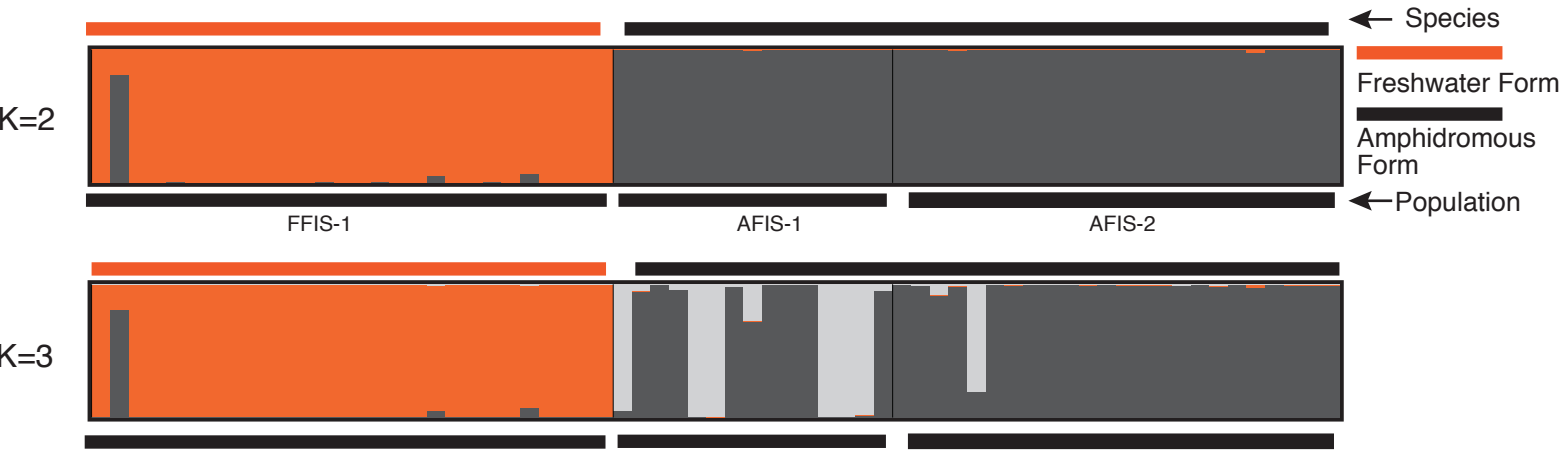
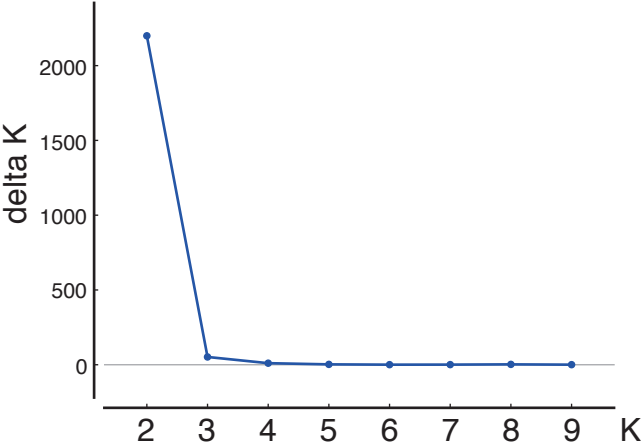


Figure S3 Continued.

G
9. Iriomotejima Is.

$\Delta K = \text{mean}(|L''(K)| / S(L(K)))$

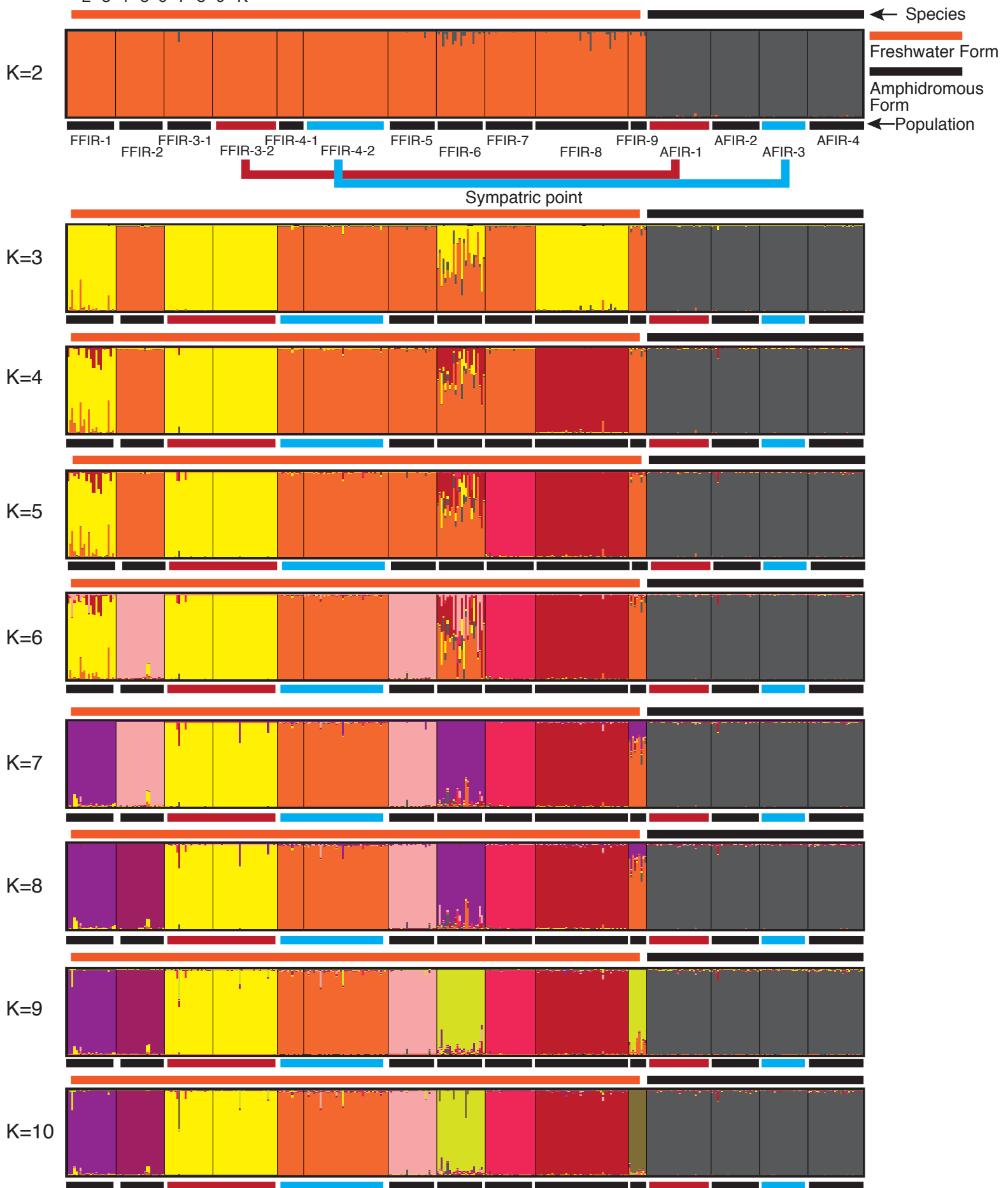
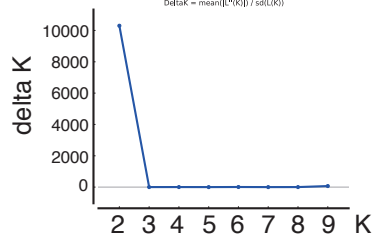
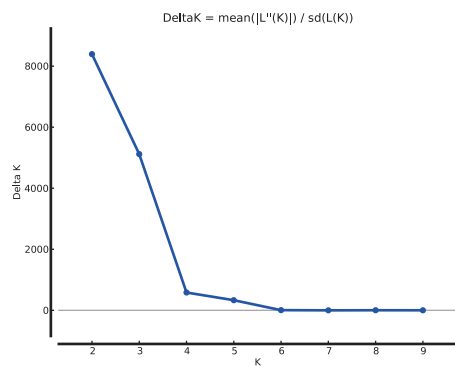


Figure S3 Continued.

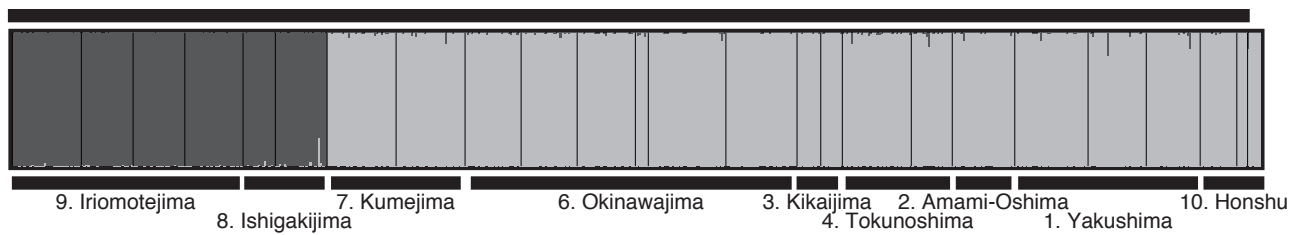
H

Rhinogobius brunneus



R. brunneus only K = 2

K=2



Southern Ryukyu

Middle Ryukyu

Northern Ryukyu + Honshu

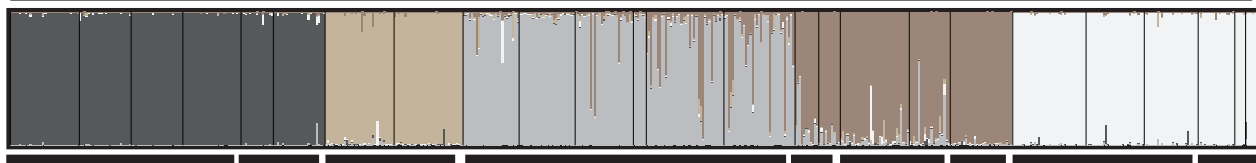
K=3



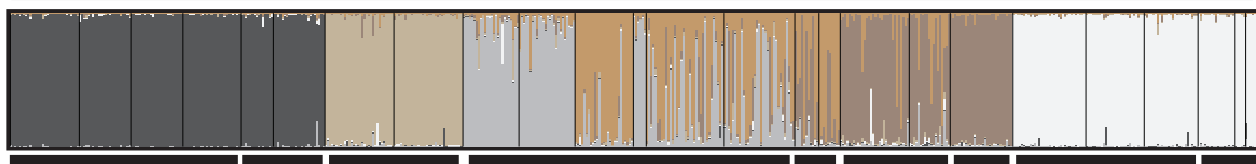
K=4



K=5



K=6



9. Iriomotejima

8. Ishigakijima

7. Kumejima

6. Okinawajima

3. Kikaijima

2. Amami-Oshima

4. Tokunoshima

1. Yakushima

10. Honshu

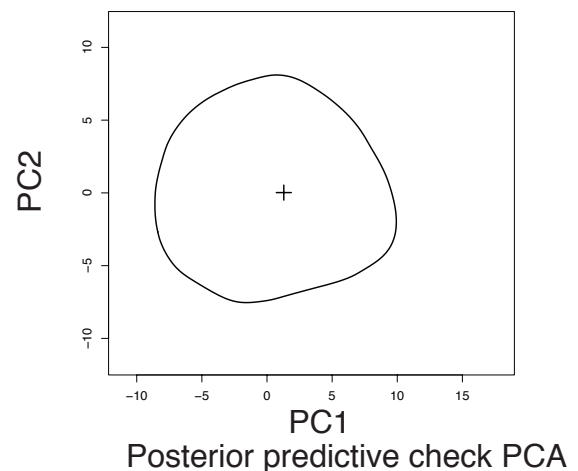
Southern Ryukyu

Middle Ryukyu

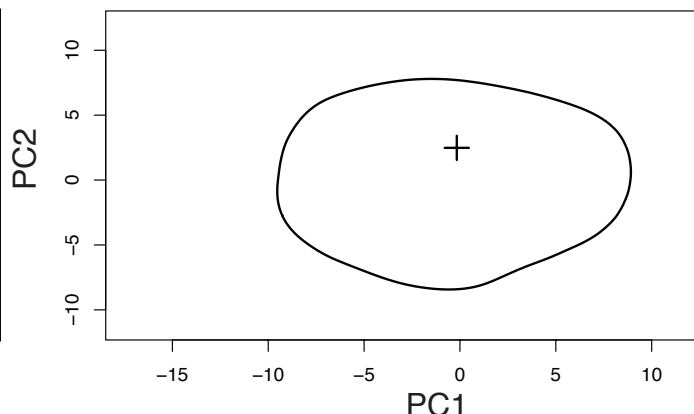
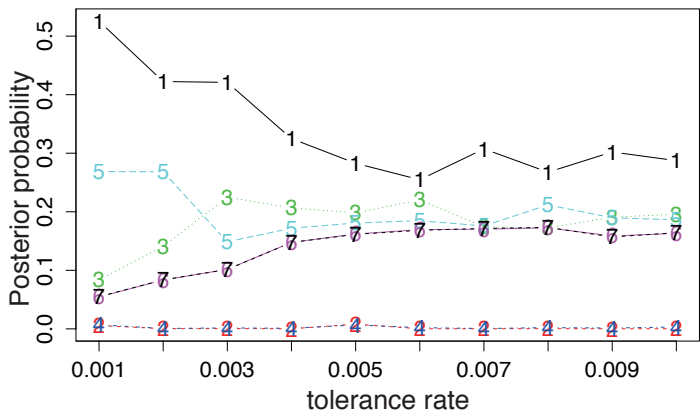
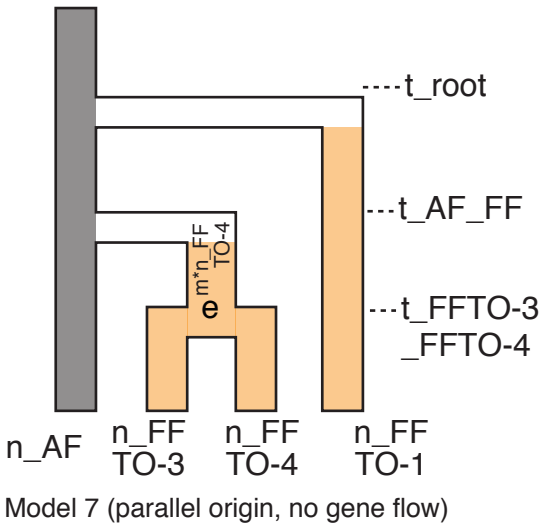
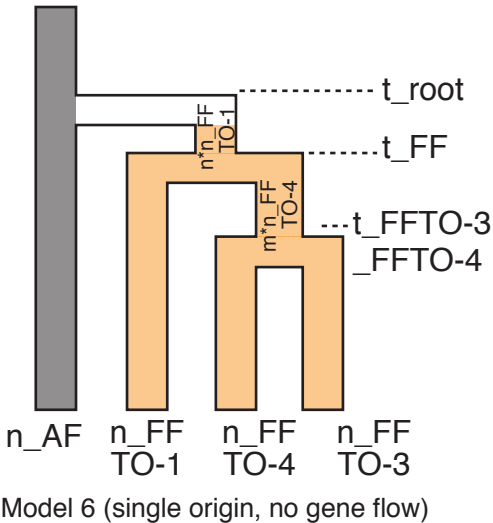
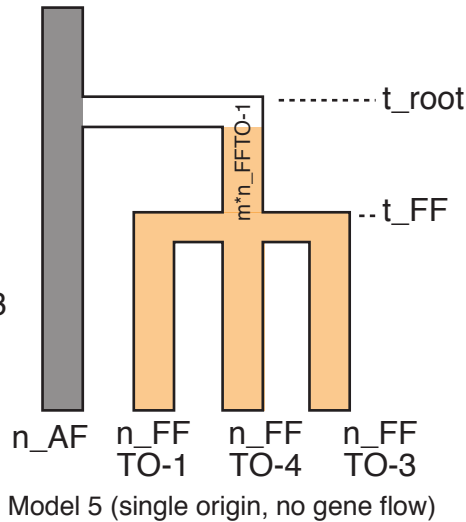
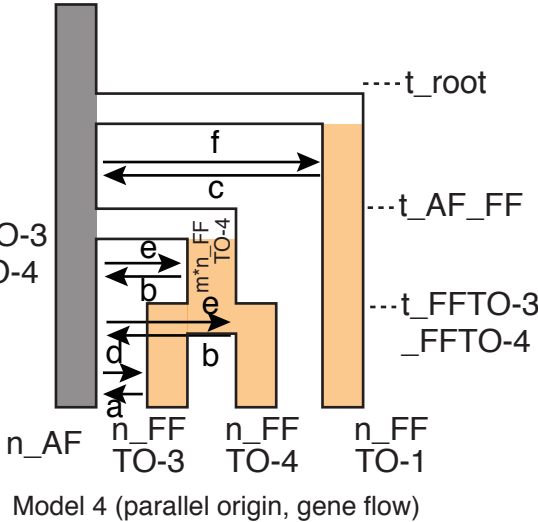
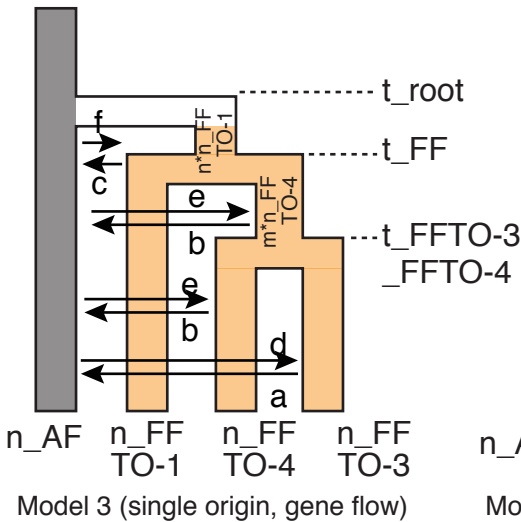
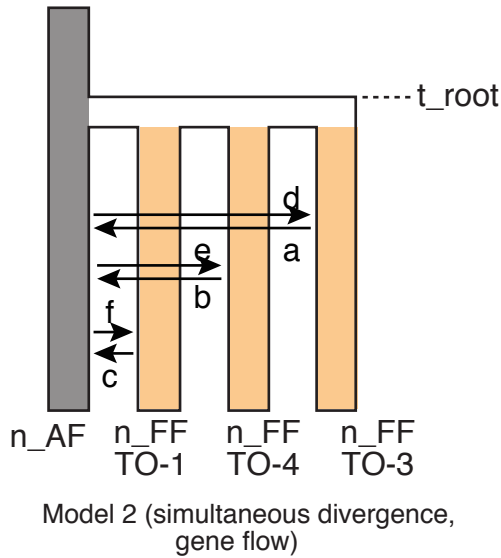
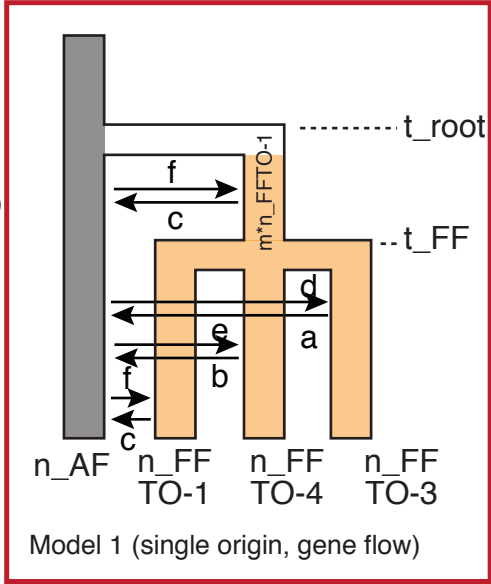
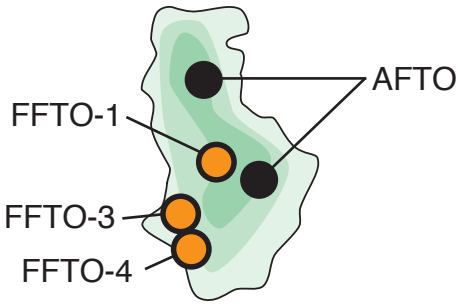
Northern Ryukyu + Honshu

Fig. S3 Continued.

A



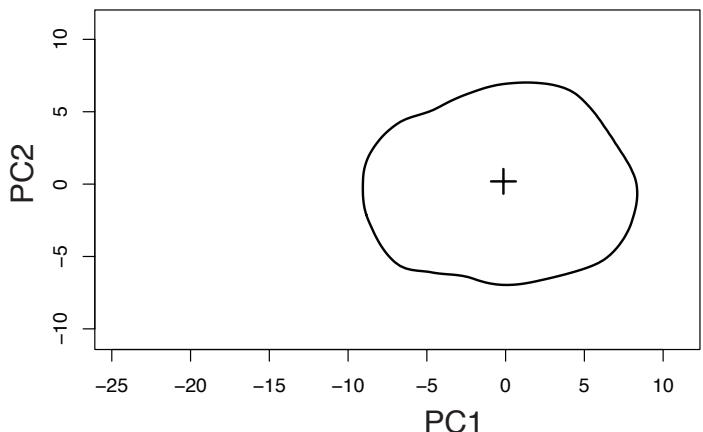
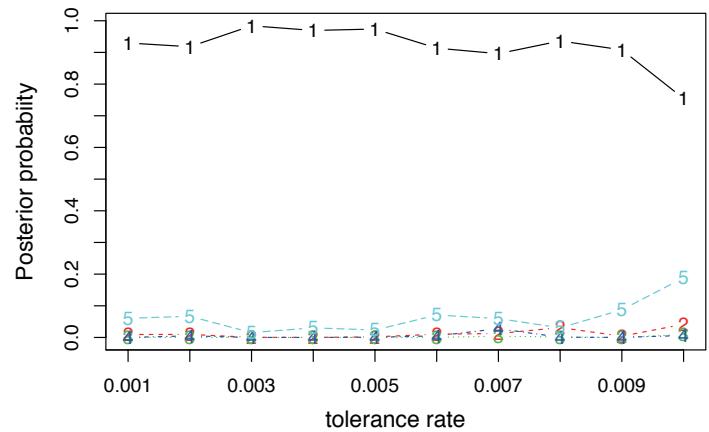
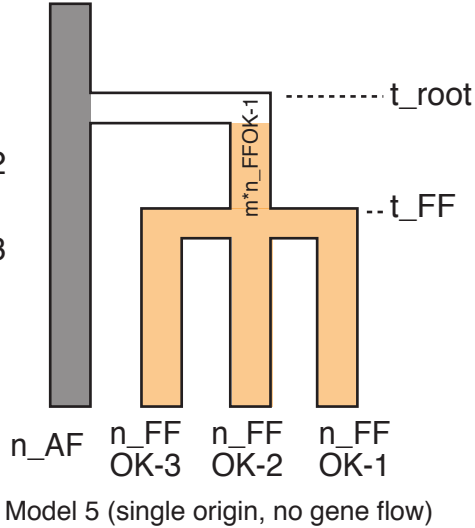
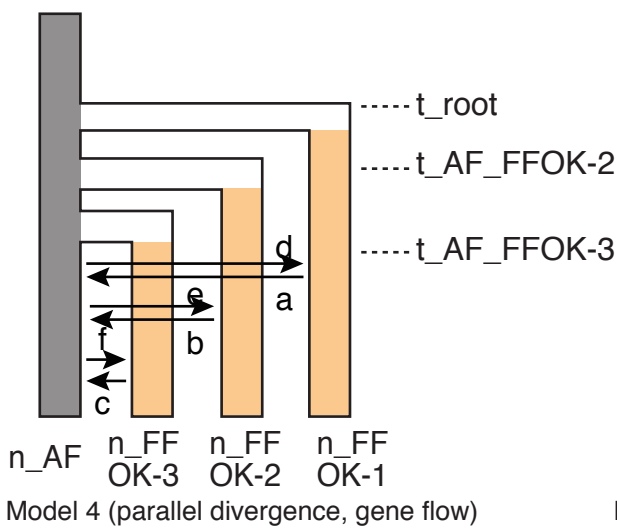
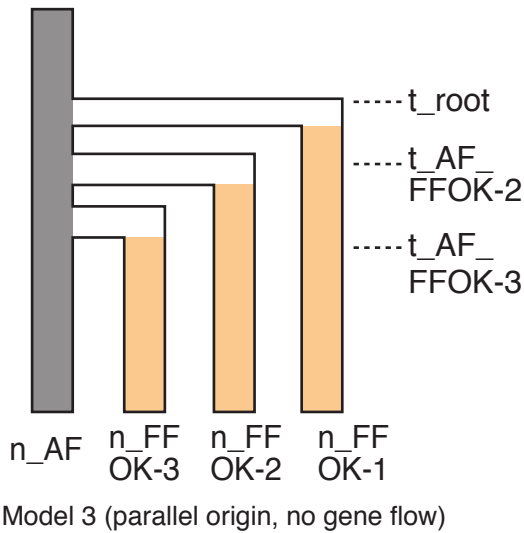
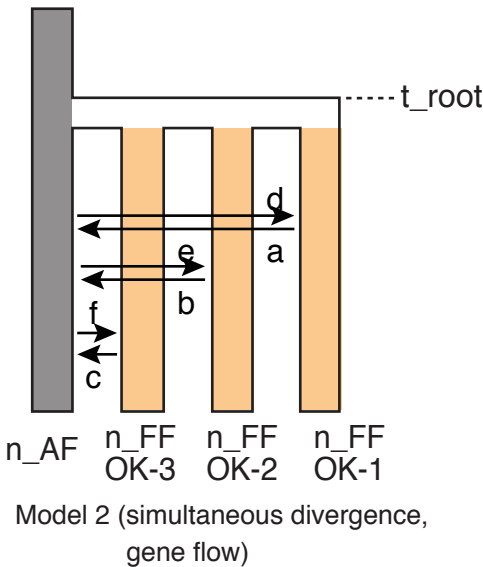
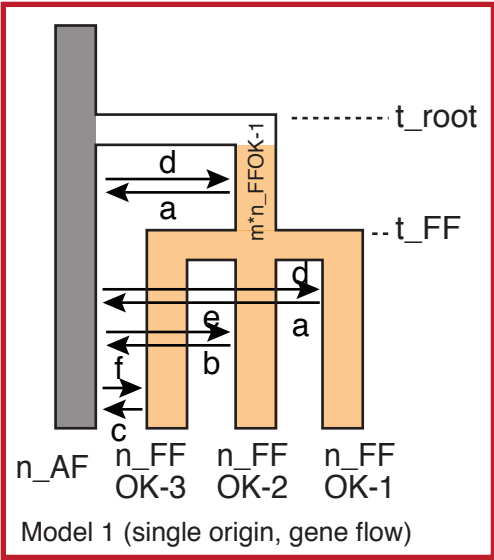
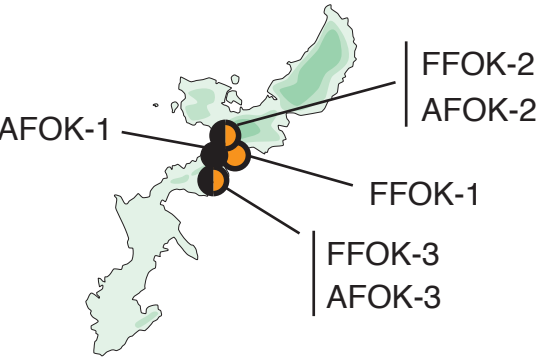
B
4. Tokunoshima Is.



Posterior probability of models

Posterior predictive check PCA

C
6. Okinawajima Is.



Posterior probability of models
Figure S4 Continued

Posterior predictive check PCA

D

7. Kumejima Is.

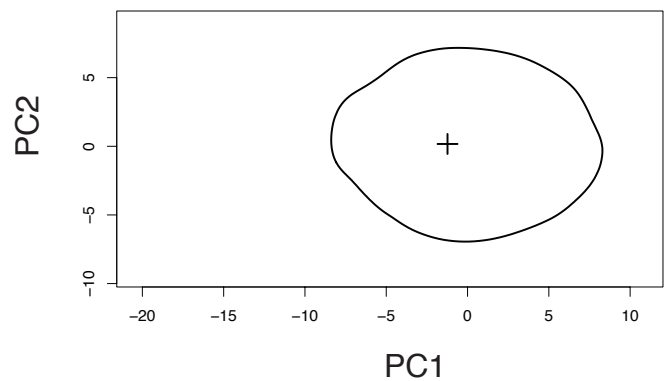
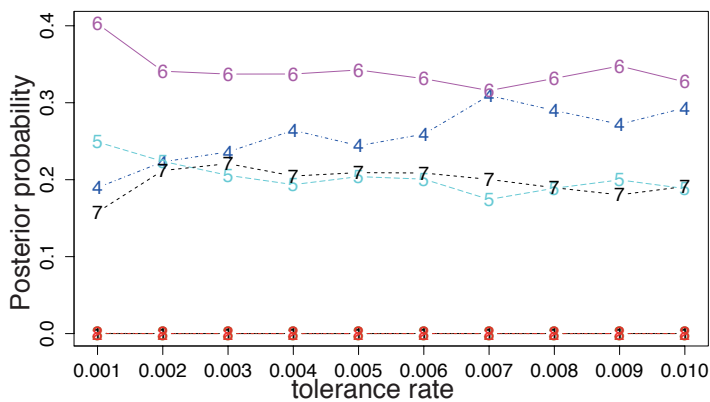
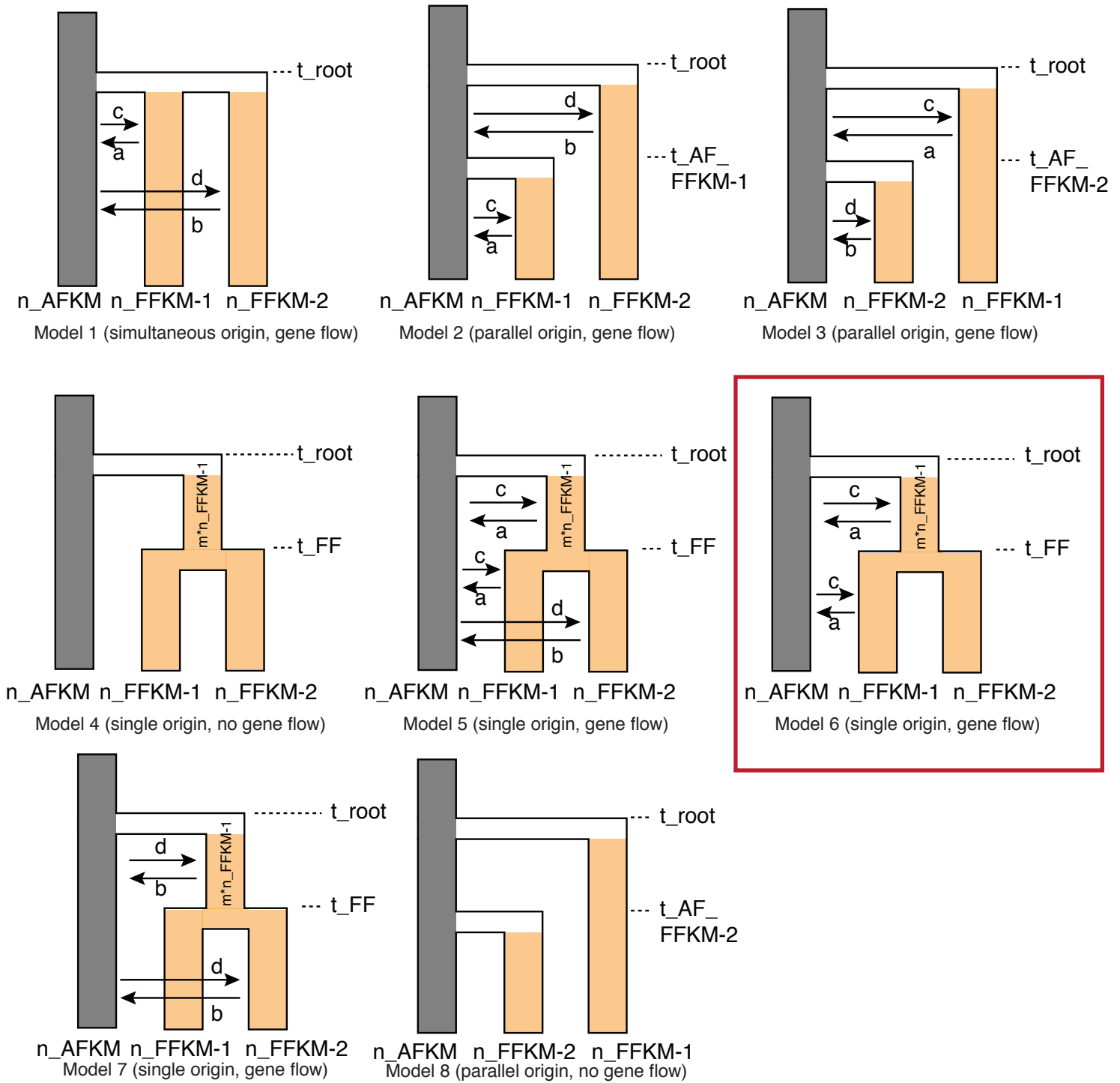
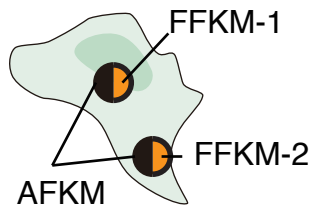
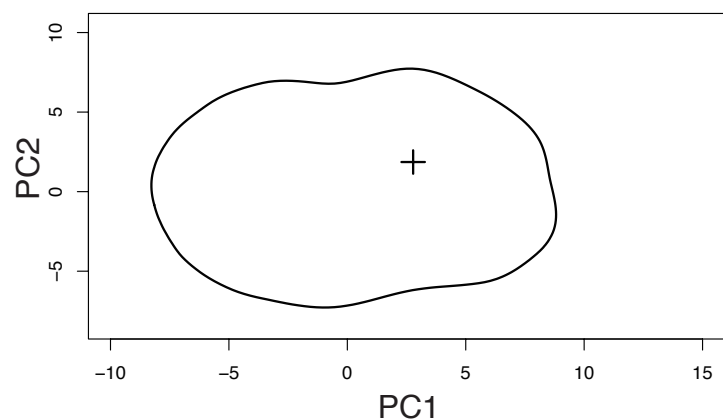
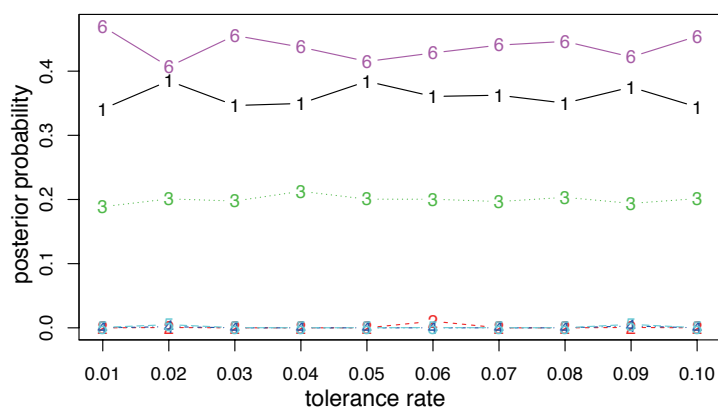
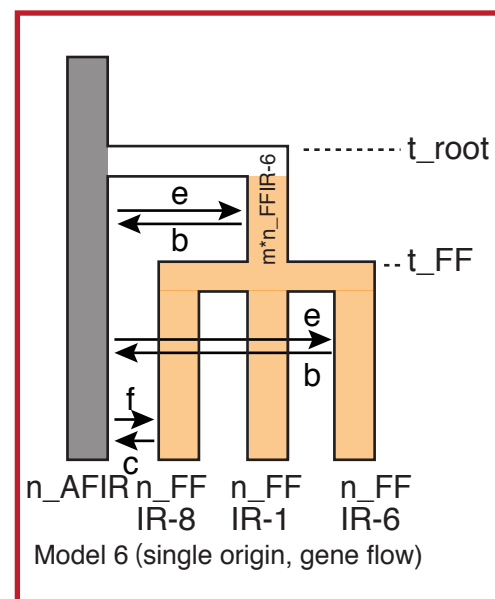
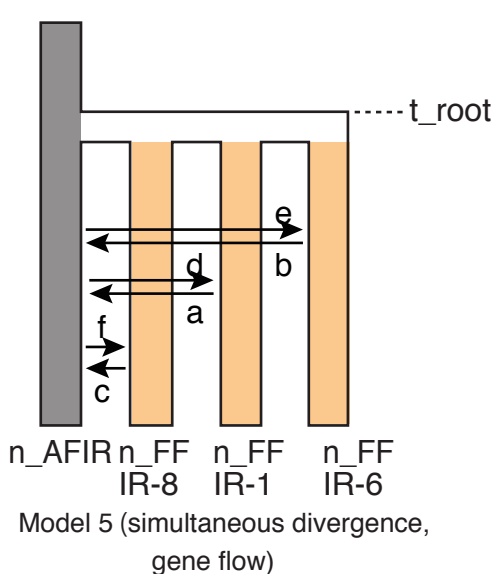
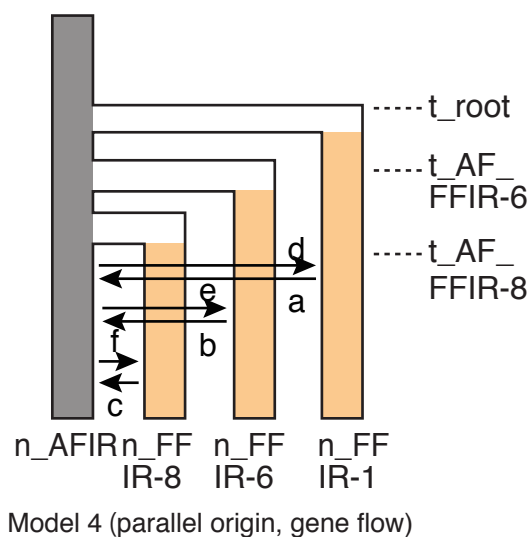
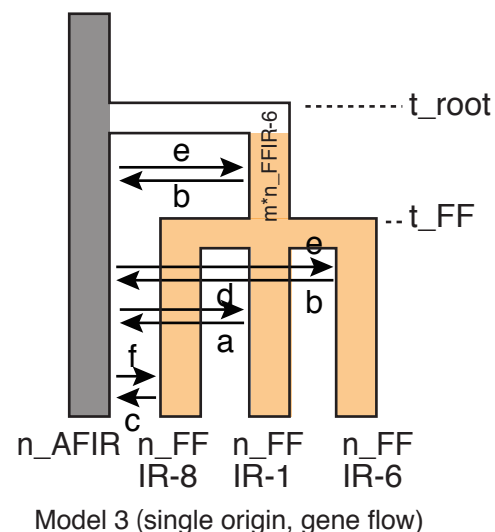
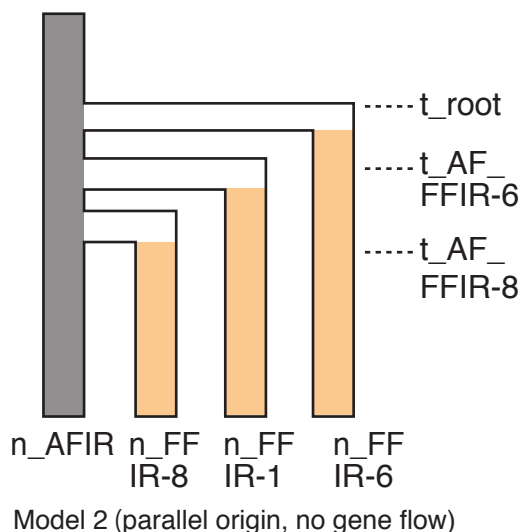
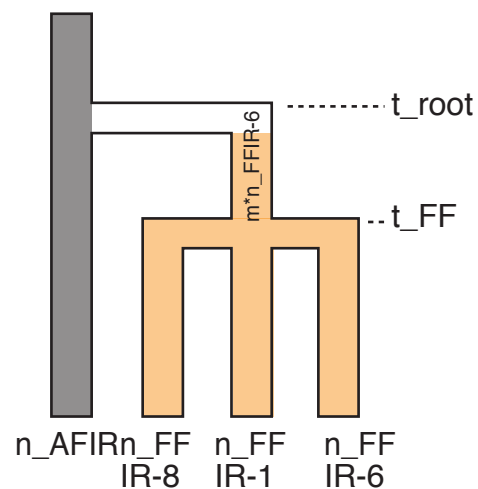
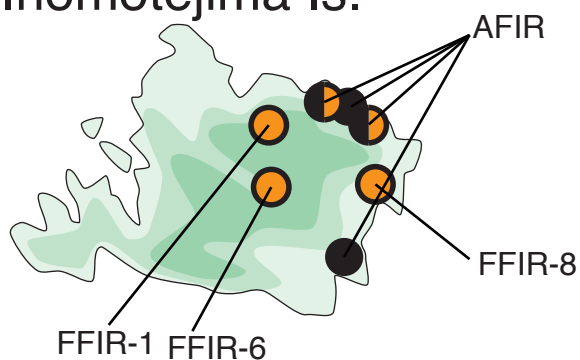


Figure S4 Continued

E

9. Iriomotejima Is.

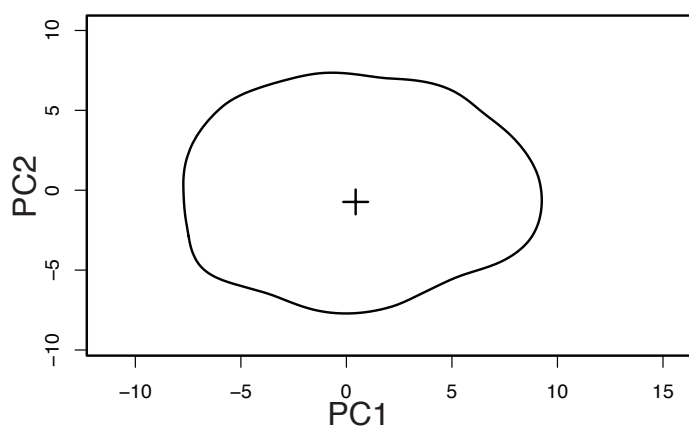
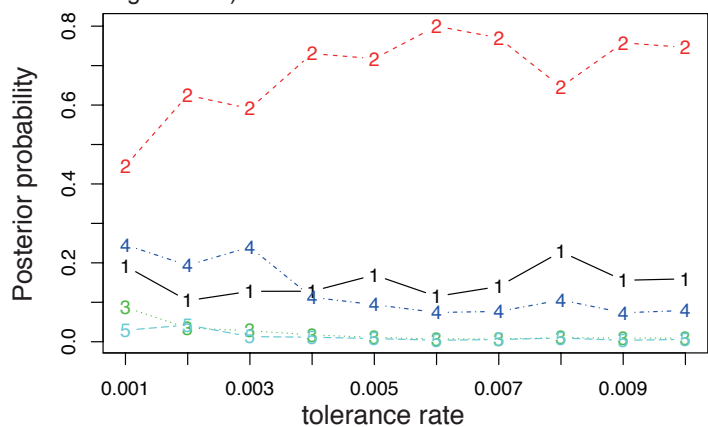
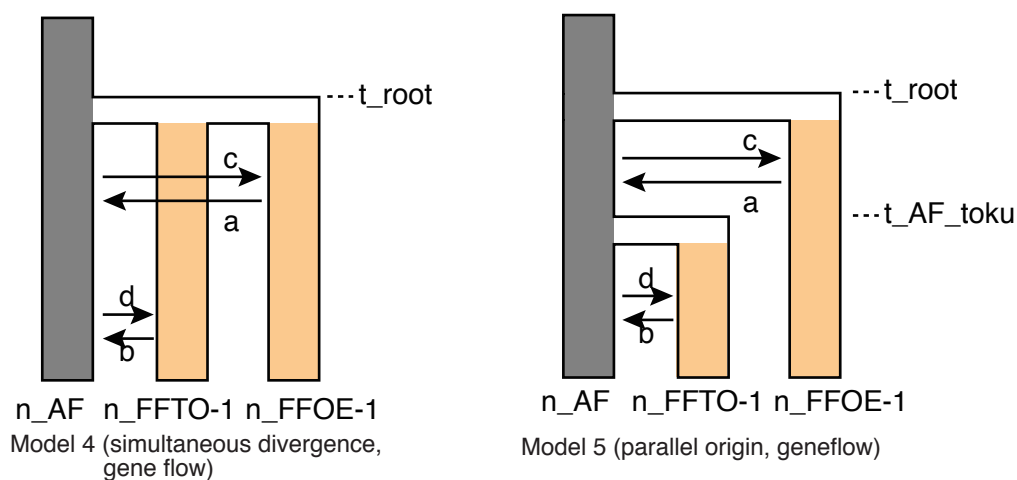
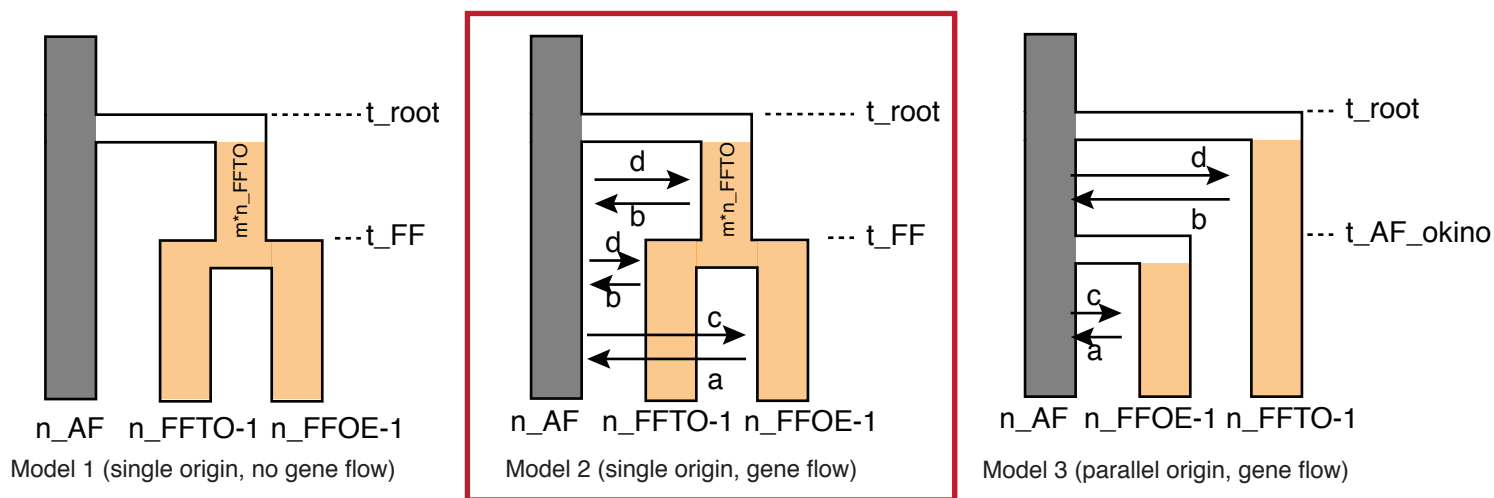
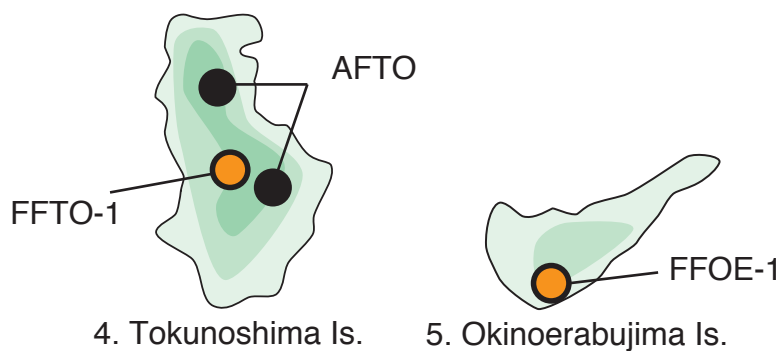


Posterior probability of models

Posterior predictive check PCA

F

4. Tokunoshima Is.– 5. Okinoerabujima Is.



Posterior probability of models

Posterior predictive check PCA

G Middle Ryukyu

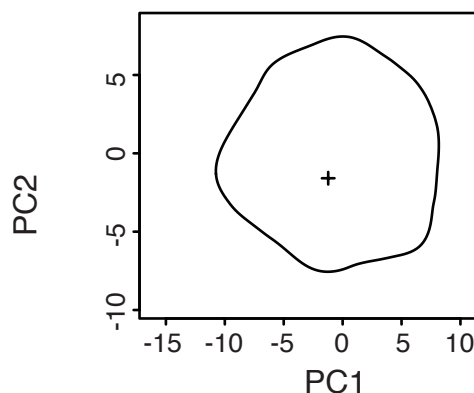
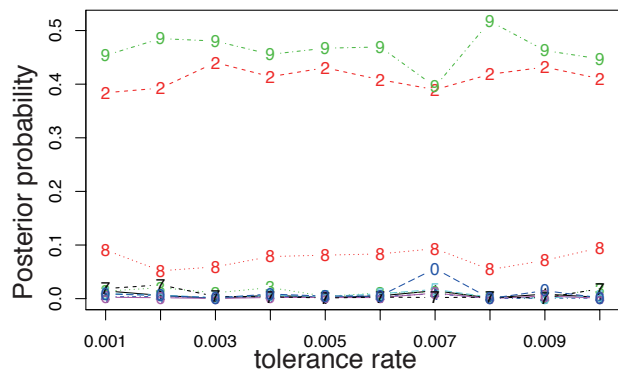
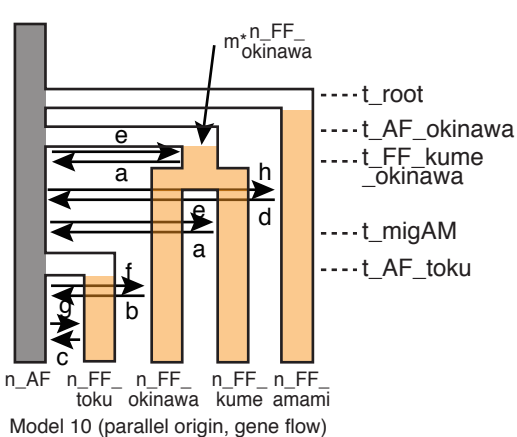
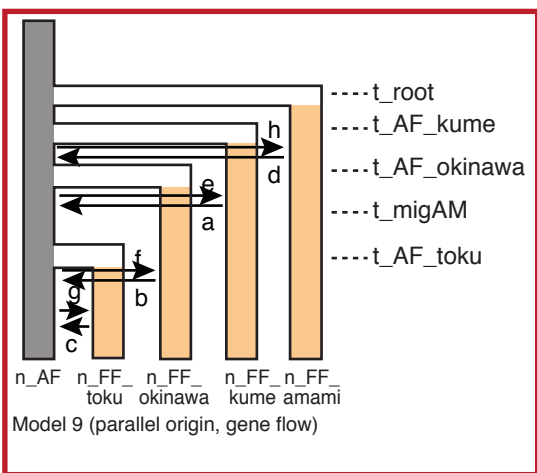
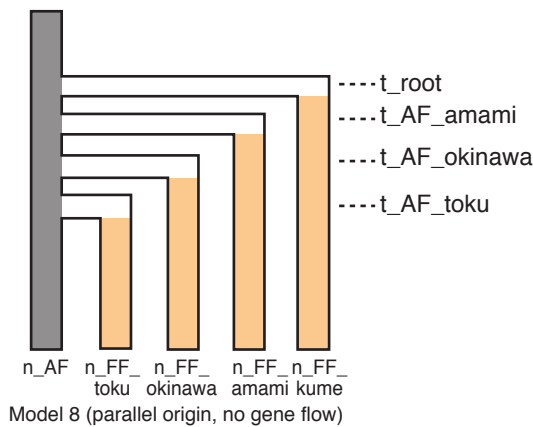
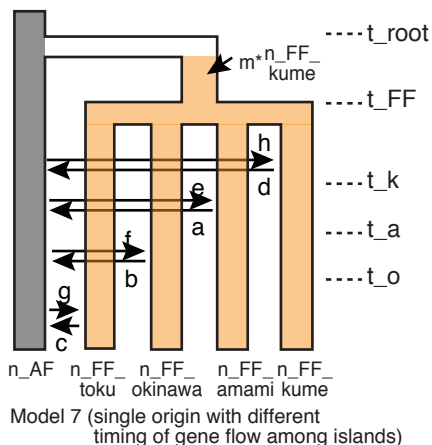
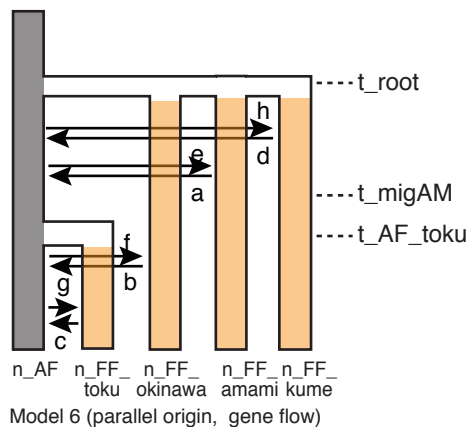
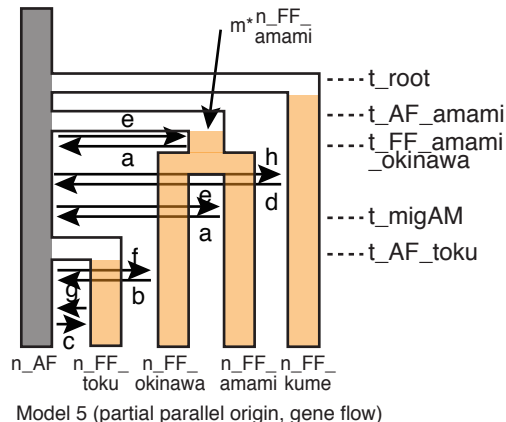
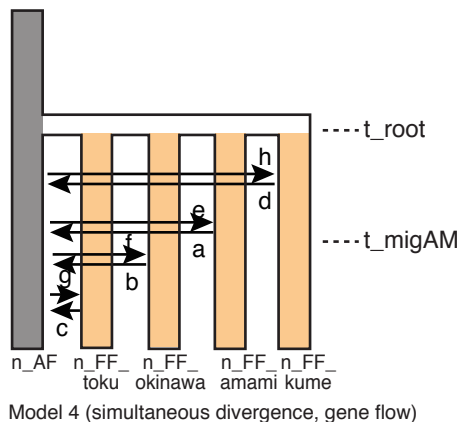
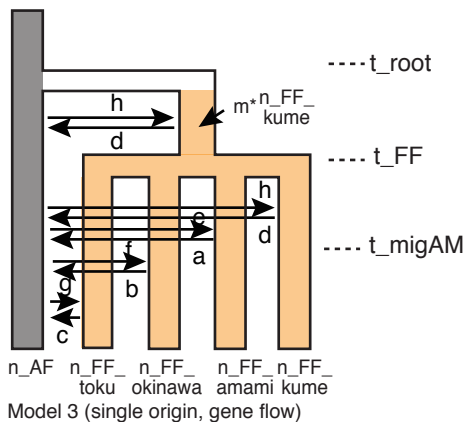
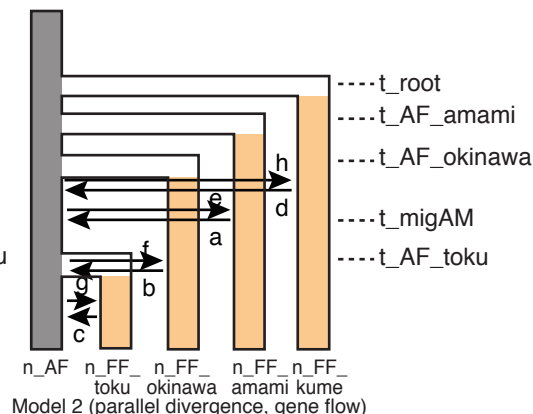
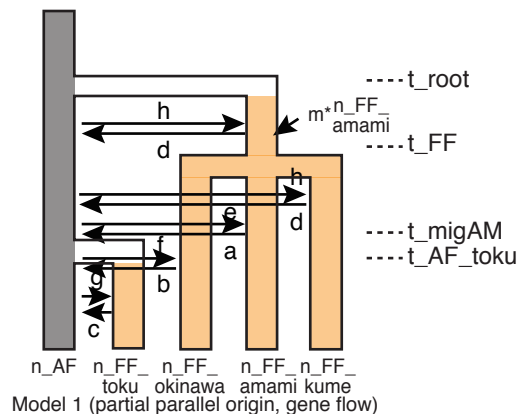
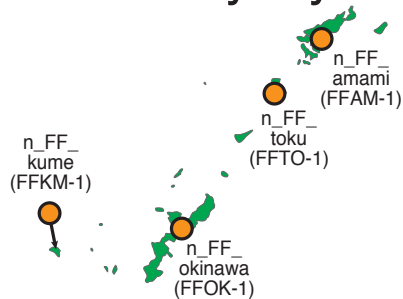
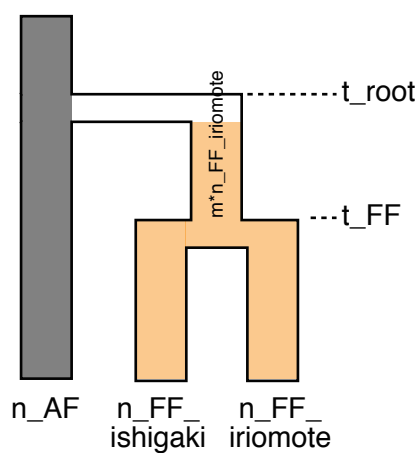
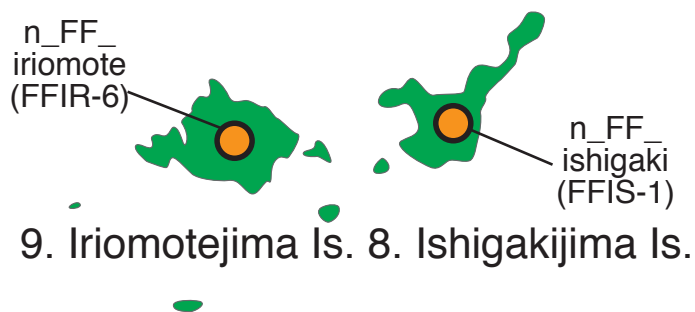


Figure S4 Continued

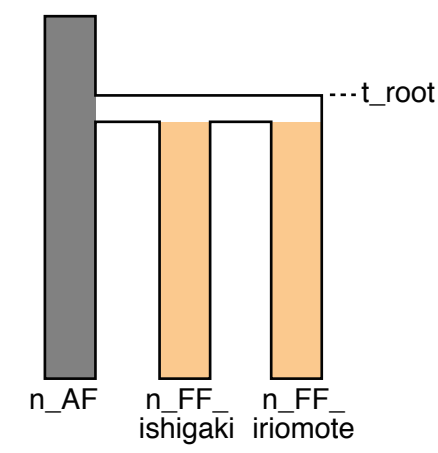
Posterior predictive check PCA

H

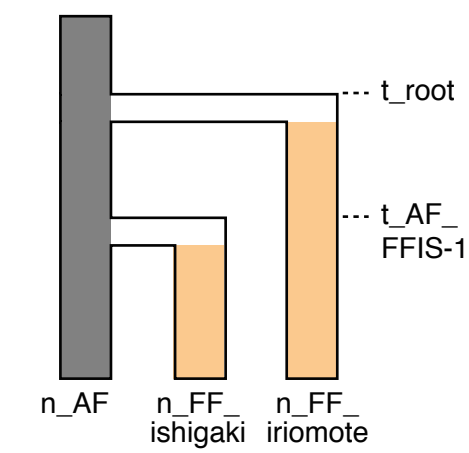
South Ryukyu



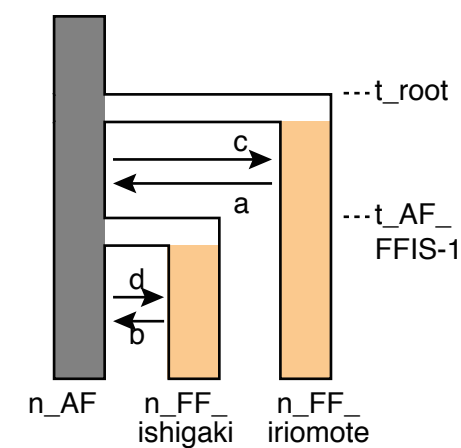
Model 1 (single origin, no gene flow)



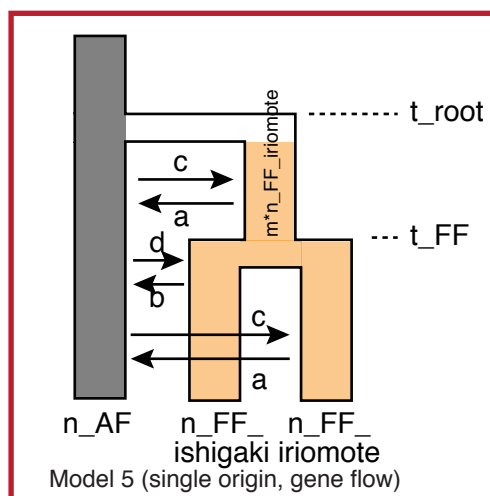
Model 2 (simultaneous divergence, no gene flow)



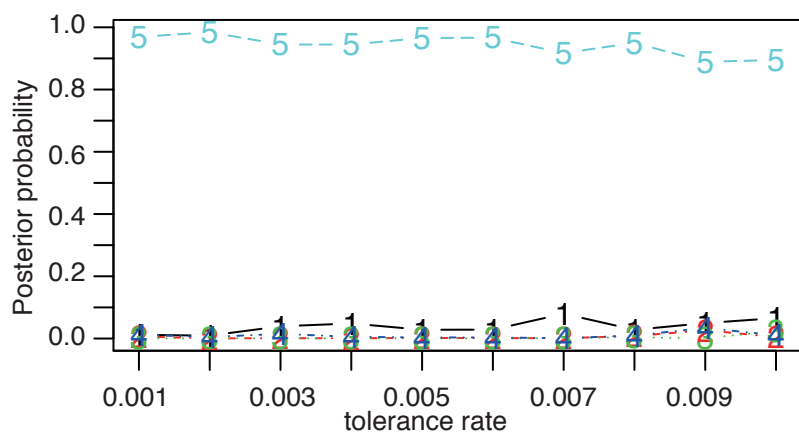
Model 3 (parallel origin, no gene flow)



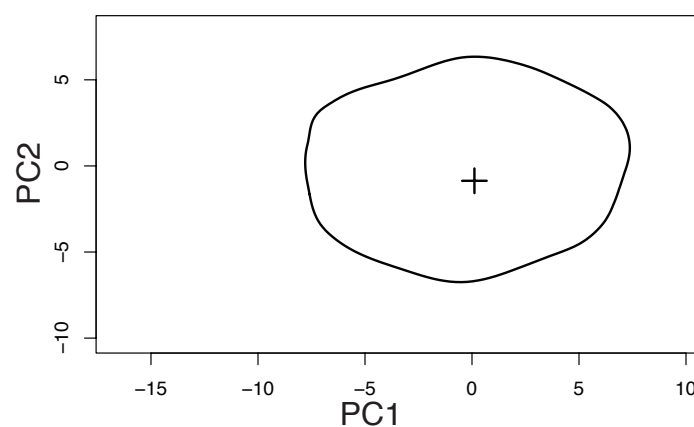
Model 4 (parallel origin, gene flow)



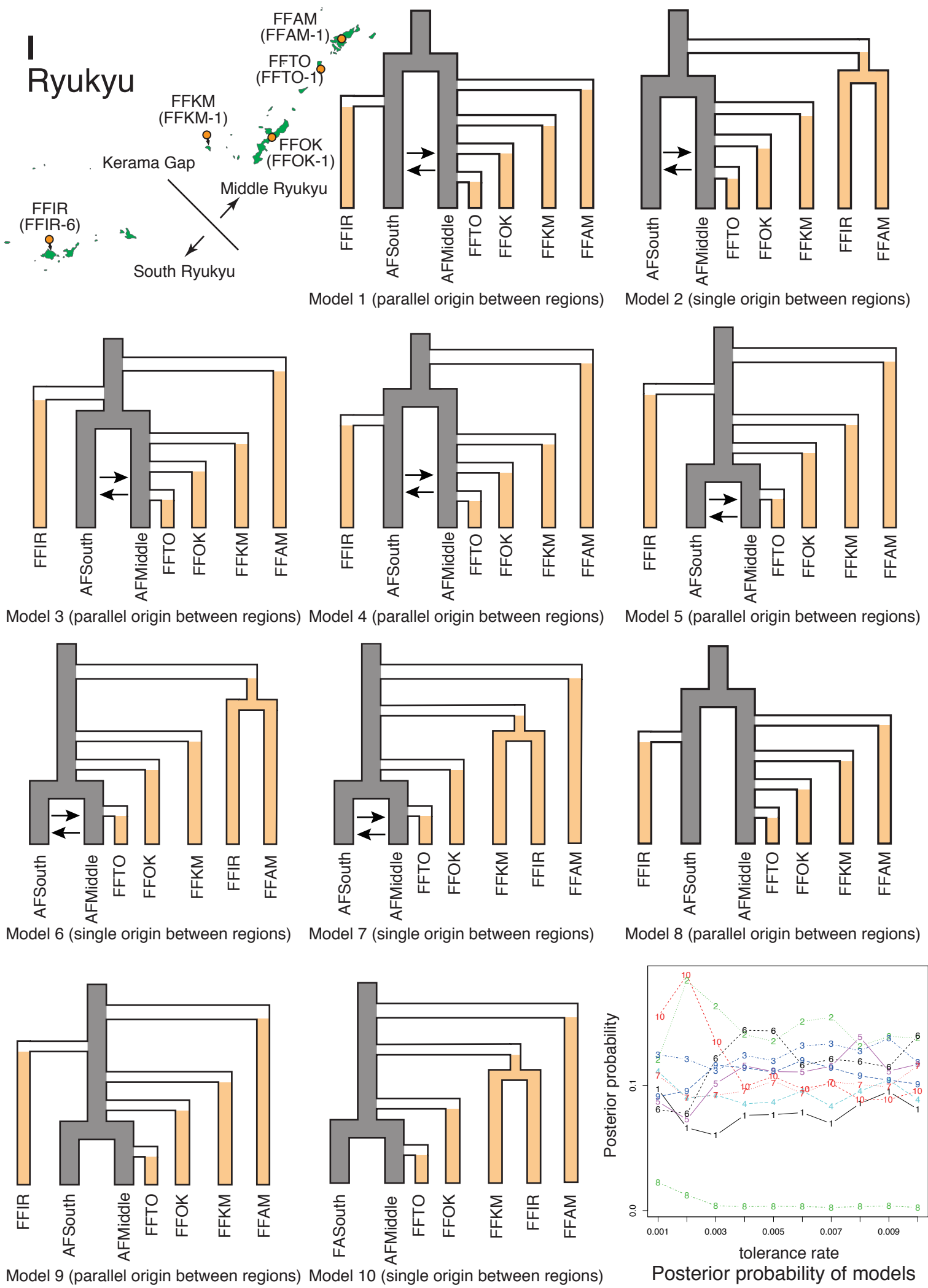
Model 5 (single origin, gene flow)



Posterior probability of models



Posterior predictive check PCA



Parameters except gene flow between amphidromous form in Middle and South Ryukyu were omitted in this figure.

Figure S4 Continued

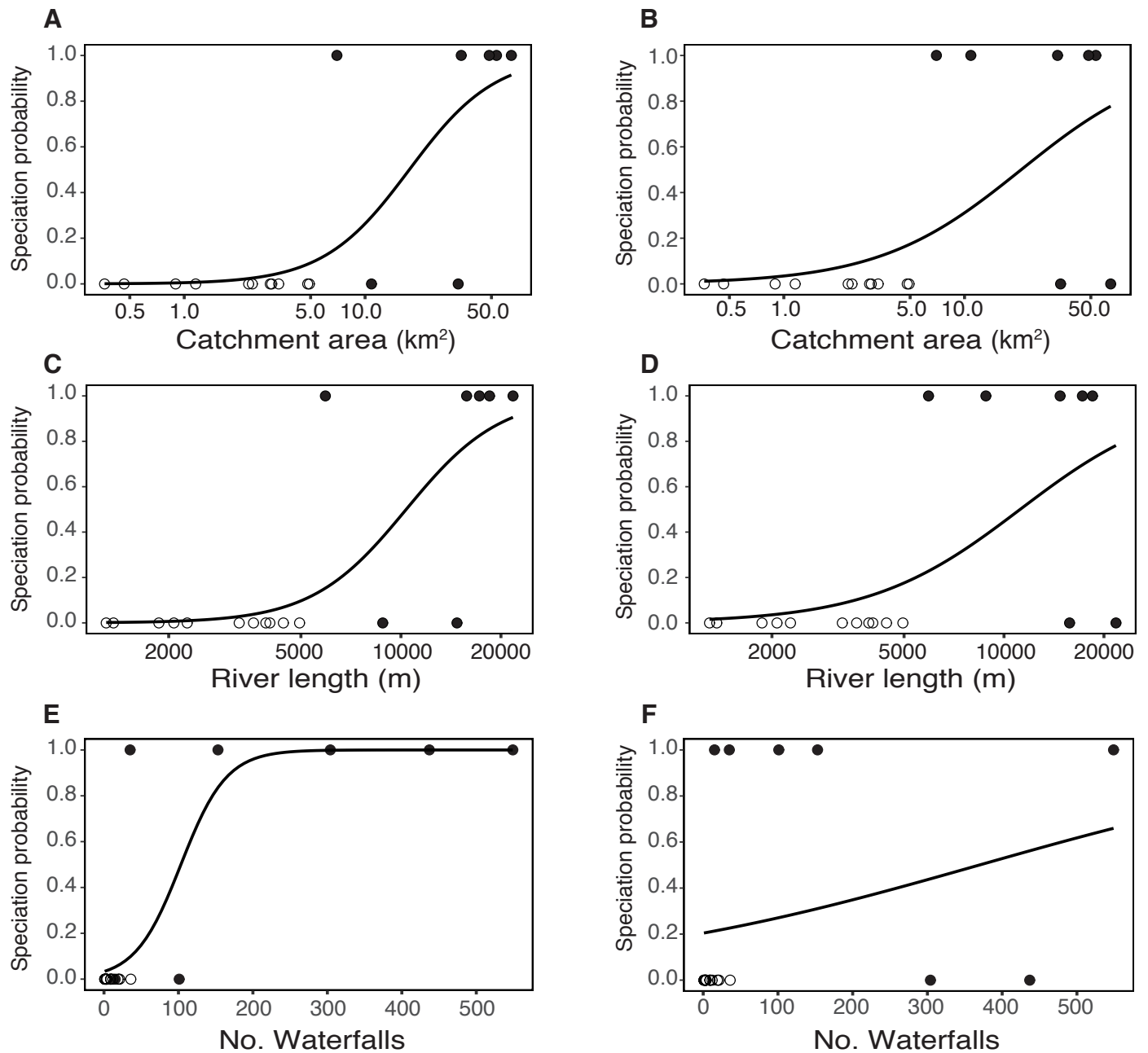


Figure S5 Results of logistic regression analysis between ecosystem size other than island size and speciation probability. Islands with the freshwater form (FF) are represented by filled circles, whereas those without FF are represented by open circles. Filled circles plotted on the x-axis (speciation probability = 0) are for the islands with FF that was assumed to be of non-independent origin. Our ABC analysis indicated that FF in Tokunoshima Is. and Okinoerabujima Is., and Ishigakijima Is. and Iriomotejima Is. shared their origins. Plots in the left column (A,C,E) assumed that the speciation of FF occurred in the larger island (Tokunoshima Is. and Iriomotejima Is.). Plots in the right column (B,D,F) assumed that speciation occurred in the smaller island (Okinoerabujima Is. and Ishigakijima Is.).

TableS1 Information about sampling point and summary statistics of each population.

Region	Island	Form	River system	Point	Population Name	River length m	Catchment area ha	N	Ho	He	Rs	ABC	
Middle Ryuky	Amami-Oshima	FF	Sumiyo	Sumiyo	FFAM-1	18478	4858	28	0.337	0.371	2.38575	Amami-Oshima/Middle Ryukyu/Ryukyu	
			Akina	Akina	FFAM-2	9969	1135	16	0.228	0.235	2.6299	Amami-Oshima	
			Kawauchi	Kawauchi	FFAM-3	12484	2817	31	0.239	0.234	1.65095	Amami-Oshima	
			Asado	Asado-2	FFAM-4	13155	3238	30	0.128	0.181	2.77865	-	
		AF	Asado	Asado-1	AFAM-1	13155	3238	29	0.507	0.507	3.5093	Amami-Oshima/Middle Ryukyu/Ryukyu	
		Kikaijima	AF	Kikai-1	Kikai-1	AFKI-1	1859	69	11	0.495	0.514	3.3566	-
			Kikai-2	Kikai-2	AFKI-2	1825	56	10	0.438	0.489	3.5747	-	
	Tokunoshima	FF	Akirigami	Akirigami-1	FFTO-1-1	15754	3400	26	0.312	0.325	2.70765	Tokunoshima/Tokunoshima-Okinoerabu/Middle Ryukyu/Ryukyu	
				Akirigami-2	FFTO-1-2			18	0.304	0.323	2.74675	Tokunoshima/Tokunoshima-Okinoerabu/Middle Ryukyu/Ryukyu	
				Akirigami-3	FFTO-1-3			9	0.3	0.323	2.61145	-	
			Oose	Oose	FFTO-2	8262	886	19	0.25	0.254	2.74265	-	
		AF	Uwanaru	Uwanaru	FFTO-3	6791	403	24	0.253	0.262	2.68755	Tokunoshima	
			Agon	Agon-1	FFTO-4-1	6958	492	24	0.3	0.311	2.0111	Tokunoshima	
				Agon-2	FFTO-4-2			16	0.256	0.287	1.78625	Tokunoshima	
			Shikaura	Shikaura	FFTO-5	9038	1184	7	0.279	0.298	1.7933	-	
		AF	Manda	Manda	AFTO-1	8817	1361	32	0.51	0.522	3.5405	Tokunoshima/Tokunoshima-Okinoerabujima/Middle Ryukyu/Ryukyu	
			Kametoku	Kametoku	AFTO-2	9034	969	19	0.526	0.539	3.49895	Tokunoshima/Tokunoshima-Okinoerabujima/Middle Ryukyu/Ryukyu	
	Okinoerabujima	FF	Amata	Amata	FFOE-1	8813	10.8	34	0.2	0.227	1.94955	Tokunoshima-Okinoerabujima	
	Okinawajima	FF	Sukuta	Sukuta-2	FFOK-1	5573	369	28	0.294	0.289	2.4092	Okinawajima/Middle Ryukyu/Ryukyu	
			Yofuke	Yofuke	FFOK-2	4442	403	31	0.098	0.121	2.2433	Okinawajima	
			Okukubi	Okukubi	FFOK-3	8967	1682	33	0.262	0.274	2.14875	Okinawajima	
			Hija	Hija-1	FFOK-4-1	17218	5330	20	0.307	0.331	2.14305	-	
		AF		Hija-2	FFOK-4-2			8	0.296	0.293	2.07295	-	
			Kyodafukuji	Kyodafukuji	FFOK-5	6445	479	32	0.169	0.176	2.35485	-	
			Kushiookawa	Kushi	FFOK-6	5299	662	32	0.133	0.144	1.92765	-	
			Sukuta	Sukuta-1	AFOK-1			26	0.506	0.554	3.62115	Okinawajima/Middle Ryukyu/Ryukyu	
			Yofuke	Yofuke	AFOK-2			25	0.533	0.542	3.55845	Okinawajima/Middle Ryukyu/Ryukyu	
			Okukubi	Okukubi	AFOK-3			27	0.477	0.523	3.6047	-	
			Hija	Hija-2	AFOK-4			6	0.475	0.493	3.16775	-	
			Sate	Sate	AFOK-5	7213	670	36	0.517	0.526	3.5023	-	
			Yona	Yona	AFOK-6	7700	1179	33	0.482	0.518	3.5941	-	
	Kumejima	FF	Shirase	Shirase	FFKM-1	5925	699	31	0.241	0.274	1.9285	Kumejima/Middle Ryukyu/Ryukyu	
			Suhara	Suhara	FFKM-2	3118	150	30	0.073	0.077	1.53105	Kumejima	
		AF	Shirase	Shirase	AFKM-1			32	0.491	0.544	3.65115	Kumejima/Middle Ryukyu/ryukyu	
			Suhara	Suhara	AFKM-2			32	0.494	0.52	3.5765	Kumejima	
South Ryukyu	Ishigakijima	FF	Miyara	Sokohara	FFIS-1	14750	3275	28	0.234	0.301	2.64575	South Ryukyu	
		AF	Arakawa	Arakawa	AFIS-1	2861	205	15	0.4	0.486	3.01515	South Ryukyu/Ryukyu	
		Nagura	Nagura	AFIS-2	8423	1604	24	0.505	0.496	3.02725	South Ryukyu/Ryukyu		
	Iriomotejima	FF	Pinai	Pinai	FFIR-1	4705	361	24	0.152	0.202	1.5609	Iriomotejima	
			Kura	Kura	FFIR-2	2937	160	24	0.094	0.104	1.21535	-	
			Geda	Geda-1	FFIR-3-1	2885	131	24	0.083	0.106	1.2767	-	
				Geda-2	FFIR-3-2			32	0.078	0.07	1.24555	-	
		AF	Yuchin	YuchinR-1	FFIR-4-1	4266	498	13	0.233	0.248	2.6738	-	
				YuchinR-2	FFIR-4-2			42	0.24	0.238	2.2249	-	
				YuchinL	FFIR-5			24	0.103	0.118	1.62565	-	
			Urauchi	Urauchi	FFIR-6	21736	6437	24	0.329	0.39	3.0796	Iriomotejima/South Ryukyu/Ryukyu	
			Arabara	Arabara	FFIR-7	3697	179	25	0.181	0.194	1.8484	-	
			Aira	Aira	FFIR-8	6130	557	46	0.289	0.305	2.9351	Iriomotejima	
			Nishida	Nishida	FFIR-9	4759	476	9	0.3	0.299	2.74405	-	
			Geda	Geda-2	AFIR-1	4266	498	32	0.46	0.456	2.8527	Iriomotejima/South Ryukyu/Ryukyu	
			Omija	Omija	AFIR-2	3395	227	24	0.508	0.502	3.1789	Iriomotejima/South Ryukyu/Ryukyu	
			Yuchin	YuchinR-2	AFIR-3	4266	498	24	0.437	0.46	2.7118	-	
			Nakama	Nakama	AFIR-4	15379	3355	27	0.437	0.466	2.9053	-	
North Ryukyu	Yakushima	AF	Takinokawa	Nunobiki-1	AFYK-1-1	2000	70	34	0.426	0.457	2.97495	-	
				Nunobiki-2	AFYK-1-2			27	0.417	0.456	3.07295	-	
			Isso	Isso	AFYK-2	10521	1471	25	0.424	0.468	3.4101	-	
Honshu	Honshu	AF	Esuno	Wakayama	AFHS-1			17	0.429	0.392	2.74215	-	
			Betto	Kagawa-1	AFHS-2			5	0.31	0.296	2.6938	-	
			Yoshida	Kagawa-2	AFHS-3			6	0.3	0.346	1.9406	-	
Total			42 river syster 50 points		59 populations			1420					

Table S2 Primer pairs used in this study.

Locus name	Motif	Forward primer sequence (5' →3') for microsatellite analysis.	Reverse primer sequence (5' →3')	Tm value	Fluorescent	Size range (bp)	No. of alleles
Ohara et al (2004)							
Rhi-3	(CA) ₅ (CA) ₄ (CA) ₁₀	GGATATTCTGTCTCTGTT	ATCTATTCCCTTTCTGTTGTCT	53	6-FAM	131-163	9
Rhi-8	(CA) ₁₈	ATACGCATAGTTTACCTTGA	CCTATGGTTTGAACCTGGGTGTG	53	HEX	155-195	8
Rhi-13	(TG) ₂ (TG) ₄ (GT) ₄	GACTCGCCATCAAATACAAAAA	AGTCTCCTCCTCACCCGCACACC	53	NED	99-132	7
New loci (This study)							
br_a_02	(AC) ₁₁	(GCCTTGCCAGCCCGC)ACTCCTAGCCTACAGCTCACTCG	GCGTCAATGCAGCACTATATTACC	63	VIC	82-126	22
br_a_05	(AC) ₁₀	(GCCTCCCTCGCGCCA)CCACTCAAGGCATTCTCCAGTTT	ATGTTTCCTCTCACAAACATCGC	63	6-FAM	195-243	14
br_a_06	(AC) ₉	(GCCTTGCCAGCCCGC)AACGTCATTATCAGATCCGCTCC	CTCCTAACTTGGCAATCACATGG	63	VIC	200-242	17
br_a_07	(AC) ₁₂	(CAGGACCAGGCTACCGTG)AGTTCATCGATCCATTACCAGA	CGTGGAGCTCTAAACAAGAGGTG	63	NED	174-226	14
br_a_08	(AC) ₁₃	(CGGAGAGCCGAGAGGTG)GTTCTGCTTCATCCATCACCAGT	TAAGATTTGTGCAGATGCGAAGG	63	PET	194-252	17
br_a_09	(AC) ₁₈	(GCCTCCCTCGCGCCA)CCAGCTGAACATGGTGTAGCTTT	TCTCAGCTGCCAGTGAACCTGAAC	63	6-FAM	278-334	29
br_a_10	(AC) ₁₁	(GCCTTGCCAGCCCGC)CGCTGTCAATCACGGTAAGAGTT	CCTATAAGCACGACCTCCATGTG	63	VIC	306-338	15
br_b_05	(AC) ₁₂	(GCCTCCCTCGCGCCA)GGAGCGGAGTTGTTGTGCTTAC	TGTCCAAGATAGTGACAAAATG	63	6-FAM	182-280	45
br_b_06	(AC) ₁₁	(GCCTTGCCAGCCCGC)ATCCTTAGGCCAACACAAAGCTC	TTAACCATATGCAAACCTCTCCC	63	VIC	190-243	22
br_c_01	(AC) ₁₃	(GCCTCCCTCGCGCCA)TGCAGTGGTTGTGTTGAAAGGTA	CCTTGACAGATGTGTCTTAGCAT	63	6-FAM	100-156	20
br_c_02	(AC) ₁₀	(GCCTTGCCAGCCCGC)GACAGCAGCACACTCCTAAGCTC	CAGATCCAGATCCTCTGTTGACAT	63	VIC	111-135	8
br_c_05	(AC) ₁₀	(GCCTCCCTCGCGCCA)TGGCTCTAGAACTCTTGATGATGG	GCACAGTATAGACGCTCTGCACAT	63	6-FAM	212-286	35
br_c_06	(AC) ₁₀	(GCCTTGCCAGCCCGC)CTGAGCAGGACAGGAAGGAAATC	TGAATTGTTTGGACCATGAGACAG	63	VIC	195-283	39
br_c_07	(AC) ₁₄	(CAGGACCAGGCTACCGTG)CCGTTTGACCGACTTCTTAAGGT	TTATCCTCACGCCTCCTTTCTTC	63	NED	209-251	17
br_c_08	(AC) ₁₀	(CGGAGAGCCGAGAGGTG)TCTTGCATGACCACAATGTCAAC	AAGCTTGGTTTGTTCCTCTTG	63	PET	222-258	15
br_c_09	(AC) ₁₀	(GCCTCCCTCGCGCCA)TCCCACCAAATAAGCCACAATA	GCTCCTCTGATGGTACTTGCTCA	63	6-FAM	334-386	22
br_c_10	(AC) ₁₁	(GCCTTGCCAGCCCGC)GACAGGAGCTCTGATCATCTCCA	CCACCATCCACCTCTATAACTCTTTC	63	VIC	340-380	17

The sequence in parentheses in forward primers indicate in tail sequence for universal fluorescent primers (Blacket et al., 2012).

Table S3 List of summary statistics used in ABC analysis.

Summary statistics	Abbreviation	Analysis the summary statistics used
Mean number of alleles over loci for each population	K_pop	All analysis
Standard deviation of Mean number of alleles over loci for each population	Ksd_pop	All analysis
Mean number of alleles over loci and population	MEAN_K	All analysis
Standard deviation of mean number of alleles over loci and population	SD_K	All analysis
Mean total number of alleles over loci	TOT_K	All analysis
Mean heterozygosity over loci for each population	H_pop	All analysis
Standard deviation of mean heterozygosity over loci for each population	Hsd_pop	All analysis
Mean heterozygosity over loci and population	MEAN_H	All analysis
Standard deviation of Mean heterozygosity over loci and population	SD_H	All analysis
Mean total heterozygosity	ALL_H	All analysis
Mean Graza–Williamson statistic over loci for each population	GW_pop	All analysis
Standard deviation of mean Graza–Williamson statistic over loci for each population	GWSD_pop	All analysis
Mean Graza–Williamson statistic over loci and population	MEAN_GW	All analysis
Standard deviation of mean Graza–Williamson statistic over loci and population	SD_GW	All analysis
Mean Graza–Williamson statistic over all populations	TOT_GW	All analysis
Mean modified Graza–Williamson statistic over loci for each population	NGW_pop	All analysis
Standard deviation of mean modified Graza–Williamson statistic over loci for each population	NGWSD_pop	All analysis
Mean modified Graza–Williamson statistic over loci and population	MEAN_NGW	All analysis
Standard deviation of mean modified Graza–Williamson statistic over loci and population	SD_NGW	All analysis
Global Fst	FST	All analysis
Global Fit	FIT	All analysis
Pairwise Fst	FST_pop1_pop2	All analysis
Mean number of pairwise differences over populations	PI_pop1_pop2	All analysis
Mean delta mu-square	DMUSQ_pop1_pop2	All analysis
Mean allele range over loci for each population	R_pop	Posterior predictive check only
Standard deviation of mean allele range over loci for each population	Rsd_pop	Posterior predictive check only
Mean allelic range over loci and populations	MEAN_R	Posterior predictive check only
Standard deviation of mean allelic range over loci and populations	MEAN_Rsd	Posterior predictive check only
Mean total allelic range over loci and populations	TOT_R	Posterior predictive check only
Global Fis	FIS	Posterior predictive check only

Table S4 Pairwise FST values.

Region		District	Sub-district	River system	Point	Population (No.)	Amani Omnia										Changirung/Changirung										Kunguma										Isiganga										Ishelemba										Yavuma										Huzulu																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																				
							FFM1	FFM2	FFM3	FFM4	FFM5	FFM6	FFM7	FFM8	FFM9	FFM10	FFM11	FFM12	FFM13	FFM14	FFM15	FFM16	FFM17	FFM18	FFM19	FFM20	FFM21	FFM22	FFM23	FFM24	FFM25	FFM26	FFM27	FFM28	FFM29	FFM30	FFM31	FFM32	FFM33	FFM34	FFM35	FFM36	FFM37	FFM38	FFM39	FFM40	FFM41	FFM42	FFM43	FFM44	FFM45	FFM46	FFM47	FFM48	FFM49	FFM50	FFM51	FFM52	FFM53	FFM54	FFM55	FFM56	FFM57	FFM58	FFM59	FFM60	FFM61	FFM62	FFM63	FFM64	FFM65	FFM66	FFM67	FFM68	FFM69	FFM70	FFM71	FFM72	FFM73	FFM74	FFM75	FFM76	FFM77	FFM78	FFM79	FFM80	FFM81	FFM82	FFM83	FFM84	FFM85	FFM86	FFM87	FFM88	FFM89	FFM90	FFM91	FFM92	FFM93	FFM94	FFM95	FFM96	FFM97	FFM98	FFM99	FFM100	FFM101	FFM102	FFM103	FFM104	FFM105	FFM106	FFM107	FFM108	FFM109	FFM110	FFM111	FFM112	FFM113	FFM114	FFM115	FFM116	FFM117	FFM118	FFM119	FFM120	FFM121	FFM122	FFM123	FFM124	FFM125	FFM126	FFM127	FFM128	FFM129	FFM130	FFM131	FFM132	FFM133	FFM134	FFM135	FFM136	FFM137	FFM138	FFM139	FFM140	FFM141	FFM142	FFM143	FFM144	FFM145	FFM146	FFM147	FFM148	FFM149	FFM150	FFM151	FFM152	FFM153	FFM154	FFM155	FFM156	FFM157	FFM158	FFM159	FFM160	FFM161	FFM162	FFM163	FFM164	FFM165	FFM166	FFM167	FFM168	FFM169	FFM170	FFM171	FFM172	FFM173	FFM174	FFM175	FFM176	FFM177	FFM178	FFM179	FFM180	FFM181	FFM182	FFM183	FFM184	FFM185	FFM186	FFM187	FFM188	FFM189	FFM190	FFM191	FFM192	FFM193	FFM194	FFM195	FFM196	FFM197	FFM198	FFM199	FFM200	FFM201	FFM202	FFM203	FFM204	FFM205	FFM206	FFM207	FFM208	FFM209	FFM210	FFM211	FFM212	FFM213	FFM214	FFM215	FFM216	FFM217	FFM218	FFM219	FFM220	FFM221	FFM222	FFM223	FFM224	FFM225	FFM226	FFM227	FFM228	FFM229	FFM230	FFM231	FFM232	FFM233	FFM234	FFM235	FFM236	FFM237	FFM238	FFM239	FFM240	FFM241	FFM242	FFM243	FFM244	FFM245	FFM246	FFM247	FFM248	FFM249	FFM250	FFM251	FFM252	FFM253	FFM254	FFM255	FFM256	FFM257	FFM258	FFM259	FFM260	FFM261	FFM262	FFM263	FFM264	FFM265	FFM266	FFM267	FFM268	FFM269	FFM270	FFM271	FFM272	FFM273	FFM274	FFM275	FFM276	FFM277	FFM278	FFM279	FFM280	FFM281	FFM282	FFM283	FFM284	FFM285	FFM286	FFM287	FFM288	FFM289	FFM290	FFM291	FFM292	FFM293	FFM294	FFM295	FFM296	FFM297	FFM298	FFM299	FFM300	FFM301	FFM302	FFM303	FFM304	FFM305	FFM306	FFM307	FFM308	FFM309	FFM310	FFM311	FFM312	FFM313	FFM314	FFM315	FFM316	FFM317	FFM318	FFM319	FFM320	FFM321	FFM322	FFM323	FFM324	FFM325	FFM326	FFM327	FFM328	FFM329	FFM330	FFM331	FFM332	FFM333	FFM334	FFM335	FFM336	FFM337	FFM338	FFM339	FFM340	FFM341	FFM342	FFM343	FFM344	FFM345	FFM346	FFM347	FFM348	FFM349	FFM350	FFM351	FFM352	FFM353	FFM354	FFM355	FFM356	FFM357	FFM358	FFM359	FFM360	FFM361	FFM362	FFM363	FFM364	FFM365	FFM366	FFM367	FFM368	FFM369	FFM370	FFM371	FFM372	FFM373	FFM374	FFM375	FFM376	FFM377	FFM378	FFM379	FFM380	FFM381	FFM382	FFM383	FFM384	FFM385	FFM386	FFM387	FFM388	FFM389	FFM390	FFM391	FFM392	FFM393	FFM394	FFM395	FFM396	FFM397	FFM398	FFM399	FFM400	FFM401	FFM402	FFM403	FFM404	FFM405	FFM406	FFM407	FFM408	FFM409	FFM410	FFM411	FFM412	FFM413	FFM414	FFM415	FFM416	FFM417	FFM418	FFM419	FFM420	FFM421	FFM422	FFM423	FFM424	FFM425	FFM426	FFM427	FFM428	FFM429	FFM430	FFM431	FFM432	FFM433	FFM434	FFM435	FFM436	FFM437	FFM438	FFM439	FFM440	FFM441	FFM442	FFM443	FFM444	FFM445	FFM446	FFM447	FFM448	FFM449	FFM450	FFM451	FFM452	FFM453	FFM454	FFM455	FFM456	FFM457	FFM458	FFM459	FFM460	FFM461	FFM462	FFM463	FFM464	FFM465	FFM466	FFM467	FFM468	FFM469	FFM470	FFM471	FFM472	FFM473	FFM474	FFM475	FFM476	FFM477	FFM478	FFM479	FFM480	FFM481	FFM482	FFM483	FFM484	FFM485	FFM486	FFM487	FFM488	FFM489	FFM490	FFM491	FFM492	FFM493	FFM494	FFM495	FFM496	FFM497	FFM498	FFM499	FFM500	FFM501	FFM502	FFM503	FFM504	FFM505	FFM506	FFM507	FFM508	FFM509	FFM510	FFM511	FFM512	FFM513	FFM514	FFM515	FFM516	FFM517	FFM518	FFM519	FFM520	FFM521	FFM522	FFM523	FFM524	FFM525	FFM526	FFM527	FFM528	FFM529	FFM530	FFM531	FFM532	FFM533	FFM534	FFM535	FFM536	FFM537	FFM538	FFM539	FFM540	FFM541	FFM542	FFM543	FFM544	FFM545	FFM546	FFM547	FFM548	FFM549	FFM550	FFM551	FFM552	FFM553	FFM554	FFM555	FFM556	FFM557	FFM558	FFM559	FFM560	FFM561	FFM562	FFM563	FFM564	FFM565	FFM566	FFM567	FFM568	FFM569	FFM570	FFM571	FFM572	FFM573	FFM574	FFM575	FFM576	FFM577	FFM578	FFM579	FFM580	FFM581	FFM582	FFM583	FFM584	FFM585	FFM586	FFM587	FFM588	FFM589	FFM590	FFM591	FFM592	FFM593	FFM594	FFM595	FFM596	FFM597	FFM598	FFM599	FFM600	FFM601	FFM602	FFM603	FFM604	FFM605	FFM606	FFM607	FFM608	FFM609	FFM610	FFM611	FFM612	FFM613	FFM614	FFM615	FFM616	FFM617	FFM618	FFM619	FFM620	FFM621	FFM622	FFM623	FFM624	FFM625	FFM626	FFM627	FFM628	FFM629	FFM630	FFM631	FFM632	FFM633	FFM634	FFM635	FFM636	FFM637	FFM638	FFM639	FFM640	FFM641	FFM642	FFM643	FFM644	FFM645	FFM646	FFM647	FFM648	FFM649	FFM650	FFM651	FFM652	FFM653	FFM654	FFM655	FFM656	FFM657	FFM658	FFM659	FFM660	FFM661	FFM662	FFM663	FFM664	FFM665	FFM666	FFM667	FFM668	FFM669	FFM670	FFM671	FFM672	FFM673	FFM674	FFM675	FFM676	FFM677	FFM678	FFM679	FFM680	FFM681	FFM682	FFM683	FFM684	FFM685	FFM686	FFM687	FFM688	FFM689	FFM690	FFM691	FFM692	FFM693	FFM694	FFM695	FFM696	FFM697	FFM698	FFM699	FFM700	FFM701	FFM702	FFM703	FFM704	FFM705	FFM706	FFM707	FFM708	FFM709	FFM710	FFM711	FFM712	FFM713	FFM714	FFM715	FFM716	FFM717	FFM718	FFM719	FFM720	FFM721	FFM722	FFM723	FFM724	FFM725	FFM726	FFM727	FFM728	FFM729	FFM730	FFM731	FFM732	FFM733	FFM734	FFM735	FFM736	FFM737	FFM738	FFM739	FFM740	FFM741	FFM742	FFM743	FFM744	FFM745	FFM746	FFM747	FFM748	FFM749	FFM750	FFM751	FFM752	FFM753	FFM754	FFM755	FFM756	FFM757	FFM758	FFM759	FFM760	FFM761	FFM762	FFM763	FFM764	FFM765	FFM766	FFM767	FFM768	FFM769	FFM770	FFM771	FFM772	FFM773	FFM774	FFM775	FFM776	FFM777	FFM778	FFM779	FFM780	FFM781	FFM782	FFM783	FFM784	FFM785	FFM786	FFM787	FFM788	FFM789	FFM790	FFM791	FFM792	FFM793	FFM794	FFM795	FFM796	FFM797	FFM798	FFM799	FFM800	FFM801	FFM802	FFM803	FFM804	FFM805	FFM806	FFM807	FFM808	FFM809	FFM810	FFM811	FFM812	FFM813	FFM814	FFM815	FFM816	FFM817	FFM818	FFM819	FFM820	FFM821	FFM822	FFM823	FFM824	FFM825	FFM826	FFM827	FFM828	FFM829	FFM830	FFM831	FFM832	FFM833	FFM834	FFM835	FFM836	FFM837	FFM838	FFM839	FFM840	FFM841	FFM842	FFM843	FFM844	FFM845	FFM846	FFM847	FFM848	FFM849	FFM850	FFM851	FFM852	FFM853	FFM854	FFM855	FFM856	FFM857	FFM858	FFM859	FFM860	FFM861	FFM862	FFM863	FFM864	FFM865	FFM866	FFM867	FFM868	FFM869	FFM870	FFM871	FFM872	FFM873	FFM874	FFM875	FFM876	FFM877	FFM878	FFM879	FFM880	FFM881	FFM882	FFM883	FFM884	FFM885	FFM886	FFM887	FFM888	FFM889	FFM890	FFM891	FFM892	FFM893	FFM894	FFM895	FFM896	FFM897	FFM898	FFM899	FFM900	FFM901	FFM902	FFM903	FFM904	FFM905	FFM906	FFM907	FFM908	FFM909	FFM910	FFM911	FFM912	FFM913	FFM914	FFM915	FFM916	FFM917	FFM918	FFM919	FFM920	FFM921	FFM922	FFM923	FFM924	FFM925	FFM926	FFM927	FFM928	FFM929	FFM930	FFM931	FFM932	FFM933	FFM934	FFM935	FFM936	FFM937	FFM938	FFM939	FFM940	FFM941	FFM942	FFM943	FFM944	FFM945	FFM946	FFM947	FFM948	FFM949	FFM950	FFM951	FFM952	FFM953	FFM954	FFM955	FFM956	FFM957	FFM958	FFM959	FFM960	FFM961	FFM962	FFM963	FFM964	FFM965	FFM966	FFM967	FFM968	FFM969	FFM970	FFM971	FFM972	FFM973	FFM974	FFM975	FFM976	FFM977	FFM978	FFM979	FFM980	FFM981	FFM982	FFM983	FFM984	FFM985	FFM986	FFM987	FFM988	FFM989	FFM990	FFM991	FFM992	FFM993

[illegible]

Table S6. Results of AMOVA analysis.

Geographic scales	Form	Source of variation	%var	Statistics	F values	P values
Among-islands within-regions	AF	Within-individuals	87.8	<i>FIT</i>	0.122	-
		Among-individuals	4.4	<i>FIS</i>	0.048	0.001
		Among-rivers in Middle Ryukyu	1.7	<i>FSC</i>	0.018	0.001
		Among-islands in Middle Ryukyu	6.2	<i>FCT</i>	0.062	0.001
		Within-individuals	91.8	<i>FIT</i>	0.082	-
		Among-individuals	2.8	<i>FIS</i>	0.03	0.016
		Among-rivers in South Ryukyu	1.7	<i>FSC</i>	0.018	0.001
		Among-islands in South Ryukyu	3.7	<i>FCT</i>	0.037	0.057
	FF	Within-individuals	34.9	<i>FIT</i>	0.651	-
		Among-individuals	3.2	<i>FIS</i>	0.083	0.001
		Among-rivers in Middle Ryukyu	30.9	<i>FSC</i>	0.448	0.001
		Among-islands in Middle Ryukyu	31	<i>FCT</i>	0.31	0.001
		Within-individuals	29.4	<i>FIT</i>	0.706	-
		Among-individuals	3.3	<i>FIS</i>	0.101	0.001
		Among-rivers in South Ryukyu	36.5	<i>FSC</i>	0.527	0.001
		Among-islands in South Ryukyu	30.8	<i>FCT</i>	0.308	0.1
Among-regions	AF	Within-individuals	68.1	<i>FIT</i>	0.319	-
		Among-individuals	4.1	<i>FIS</i>	0.057	0.001
		Among-islands	5.3	<i>FSC</i>	0.068	0.001
		Among-regions	22.5	<i>FCT</i>	0.225	0.002
	FF	Within-individuals	26.7	<i>FIT</i>	0.733	-
		Among-individuals	22.5	<i>FIS</i>	0.457	0.001
		Among-islands	34	<i>FSC</i>	0.409	0.001
		Among-regions	16.8	<i>FCT</i>	0.168	0.038

Table S7

Detail results of ABC analysis. For each ABC analysis, the following data are listed: prior distribution, posterior distribution, observed summary statistics, confusion matrix, mean model posterior probabilities and Bayes Factors.

Amami-Oshima

Prior distribution (Model0)																										
parameter name	log10_a	log10_b	log10_d	log10_e	log10_n_AFAM-1	log10_n_FFAM-2	log10_n_FFAM-3	log10_n_FFAM-1	log10_l_FF	log10_l_root	l_FFAM-1_FI	m	n	p	a	b	d	e	n_AFAM-1	n_FFAM-2	n_FFAM-3	n_FFAM-1	l_FF	l_root	l_FFAM-1	l_FFAM-3
distribution	uniform	uniform	uniform	uniform	uniform	uniform	uniform	uniform	uniform	uniform	uniform	uniform	uniform	uniform	uniform	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Min.	-3	-3	-3	-3	2	2	2	2	3	3	3	0.5	0.5	0	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	
Max.	-9	-9	-9	-9	5.3	4.7	4.7	4.7	6.2	6.2	6.2	2	2	1	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	
Complex parameter	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	10*(log10_a)	10*(log10_b)	10*(log10_d)	10*(log10_e)	n_AFAMlog10_n_FFAMlog10_n_FFAM	log10_l_FF	log10_l_root	log10_l_FFAM-1	log10_l_FFAM-3			

Posterior distribution (Model0)																			
parameter name	log10_a	log10_b	log10_d	log10_e	log10_n_AFAM-1	log10_n_FFAM-2	log10_n_FFAM-3	log10_n_FFAM-1	log10_l_FF	log10_l_root	l_FFAM-1_FI	m	n	p	a	b	d	e	n_AFAM-1
Min.	-9.688245188	-9.261748857	-9.447602948	-10.43091199	2.148707771	2.051890226	2.020025768	2.210023534	2.932256647	2.8493256	2.9351179	0.4689047	0.4833346	-0.0080614	2.01E-10	6.00E-10	3.56E-10	4.43E-11	143.36576
Weighted 2.5% Perc.	-8.880829517	-8.840722165	-8.688543283	-8.454426884	3.253432034	2.569599428	2.698025807	3.028594802	3.227351115	3.6598599	3.0116441	0.5043041	0.5307589	0.1804049	1.28E-09	1.46E-09	2.05E-09	4.14E-09	1781.4591
Weighted Median	-5.74920082	-6.03464377	-5.103433321	-5.131292527	4.162780391	3.141155135	3.397724724	3.71612343	3.715414707	4.7243364	3.2674309	1.1091171	1.1653536	0.5138616	1.82E-06	9.29E-07	8.01E-06	7.97E-06	14366.955
Weighted Mean	-5.851660078	-6.049782859	-5.457287905	-5.445428834	4.146442013	3.119337361	3.38485093	3.695395918	3.724955177	4.8218966	3.31287	1.1549767	1.194043	0.5004376	8.45E-05	5.89E-05	0.0001353	5.26E-05	20915.099
Weighted Mode	-3.961825684	-6.208226706	-3.646448158	-4.625548898	4.149616233	3.172005233	3.414502848	3.786201104	3.714934422	4.2238142	3.094914	0.6180184	0.6370036	0.5723542	4.05E-06	-6.10E-07	1.76E-06	1.32E-05	7608.3026
Weighted 97.5% Perc.	-3.139786642	-3.260633003	-3.024633339	-3.415798945	4.915542934	3.553001048	4.017396958	4.278223536	4.292053651	6.1541921	3.864486	1.9513652	1.952053	0.7499216	0.0007518	0.0005524	0.0009655	0.0003953	82097.119
Max.	-2.287622203	-2.393696873	-2.393696834	-1.834873376	5.348608663	4.305579685	6.613133261	4.67851295	5.288952643	6.6873637	4.6276388	2.0384319	2.021237	0.8216372	0.0051932	0.001373	0.0042044	0.0133041	214464.83

Observed Summary Statistics Value											
Summary statistics	K_FFAM-1	K_FFAM-2	K_FFAM-3	K_AF	Ked_FFAM-1	Ked_FFAM-2	Ked_FFAM-3	Ked_AF	mean_K	sd_K	tot_K
Observed Value	4.6	2.3	2.5	7.5	5.11345	1.83819	1.93309	5.64288	4.225	2.41851	10.75
Summary statistics	H_FFAM-1	H_FFAM-2	H_FFAM-3	H_AF	Hed_FFAM-1	Hed_FFAM-2	Hed_FFAM-3	Hed_AF	mean_H	sd_H	tot_H
Observed Value	0.370812	0.234677	0.233783	0.50702	0.319123	0.289914	0.28524	0.33781	0.336573	0.130605	0.53725
Summary statistics	GW_FFAM-1	GW_FFAM-2	GW_FFAM-3	GW_AF	GWad_FFAM-1	GWad_FFAM-2	GWad_FFAM-3	GWad_AF	mean_GW	sd_GW	tot_GW
Observed Value	0.809459	0.389957	0.685698	0.727478	0.192735	0.140586	0.256188	0.225266	0.653148	0.182836	0.718245
Summary statistics	NGW_FFAM-1	NGW_FFAM-2	NGW_FFAM-3	NGW_AF	NGWad_FFAM-1	NGWad_FFAM-2	NGWad_FFAM-3	NGWad_AF	mean_NGW	sd_NGW	
Observed Value	0.316684	0.186862	0.195131	0.513304	0.202424	0.117843	0.108306	0.201813	0.30302	0.152223	
Summary statistics	R_FFAM-1	R_FFAM-2	R_FFAM-3	R_AF	Rad_FFAM-1	Rad_FFAM-2	Rad_FFAM-3	Rad_AF	mean_R	sd_R	tot_R
Observed Value	6.6	8.8	5.81818	11.5263	7.37564	4.2374	6.64557	10.9769	8.18612	2.55973	14.5
Summary statistics	FIB	FST	FIT								
Observed Value	0.0284856	0.425806	0.440072								
Summary statistics	FST_FFAM-2_FFAM-1	FST_FFAM-3_FFAM-1	FST_FFAM-3_FFAM-2	FST_AF_FFAM-1	FST_AF_FFAM-2	FST_AF_FFAM-3					
Observed Value	0.518432	0.186551	0.615935	0.397943	0.354383	0.473872					
Summary statistics	PI_FFAM-2_FFAM-1	PI_FFAM-3_FFAM-1	PI_FFAM-3_FFAM-2	PI_AF_FFAM-1	PI_AF_FFAM-2	PI_AF_FFAM-3					
Observed Value	12.3831	7.01757	12.1764	14.8869	11.8646	15.853					
Summary statistics	MUSO_FFAM-2_FFAM-MUSQ	FFAM-3_FFAM-MUSQ	FFAM-3_FFAM-MUSQ	AF_FFAM-MUSQ	AF_FFAM-MUSQ	AF_FFAM-MUSQ					
Observed Value	15.2019	5.91293	6.84651	28.9158	5.34731	16.0727					

Confusion matrix based on 100 samples for each model (neuralnet, tolerance=0.01).								
model focal \ selected	1	2	3	4	5	6 (best model)	7	8
1 (Parallel origin, gene flow)	0.32	0.2	0	0.15	0.14	0.04	0.02	0.13
2 (Parallel origin, gene flow)	0.18	0.48	0.01	0.09	0.17	0.05	0	0.02
3 (Single origin, no gene flow)	0	0.01	0.81	0	0	0.03	0.04	0.11
4 (Single origin, gene flow)	0.17	0.16	0.07	0.28	0.18	0.05	0.07	0.04
5 (Parallel origin, gene flow)	0.18	0.27	0.03	0.14	0.18	0.12	0.02	0.06
6 (Single origin, restricted gene flow)	0.03	0.02	0.15	0.02	0.01	0.36	0.16	0.25
7 (Single origin, restricted gene flow)	0.03	0	0.27	0.01	0.01	0.2	0.29	0.19
8 (Parallel origin, no gene flow)	0.01	0	0.13	0	0	0.04	0.03	0.79

Mean model posterior probabilities (neuralnet, tolerance rate=0.01)								
model focal \ selected	1	2	3	4	5	6 (best model)	7	8
1 (Parallel origin, gene flow)	0.2295	0.1968	0.0092	0.1814	0.2052	0.0518	0.0389	0.0872
2 (Parallel origin, gene flow)	0.1891	0.3154	0.0158	0.1673	0.2271	0.0386	0.0249	0.0208
3 (Single origin, no gene flow)	0.0135	0.0147	0.4751	0.0446	0.0174	0.1342	0.2075	0.093
4 (Single origin, gene flow)	0.1786	0.1775	0.0463	0.2324	0.1934	0.0735	0.0681	0.0302
5 (Parallel origin, gene flow)	0.2015	0.2103	0.0181	0.1633	0.2146	0.0762	0.0524	0.0436
6 (Single origin, restricted gene flow)	0.0649	0.0503	0.1311	0.0849	0.0621	0.2387	0.2084	0.1596
7 (Single origin, restricted gene flow)	0.0571	0.0319	0.1672	0.0723	0.048	0.2252	0.2698	0.1284
8 (Parallel origin, no gene flow)	0.0632	0.0282	0.1034	0.0299	0.0379	0.1332	0.1141	0.49

Bayes Factor (neuralnet, tolerance rate=0.01)								
numerator \ denominator	1	2	3	4	5	6	7	8
1 (Parallel origin, gene flow)	1	0.5373	0.1342	0.3741	0.3947	0.0217	0.0384	0.0735
2 (Parallel origin, gene flow)	1.8611	1	0.2498	0.6962	0.7345	0.0404	0.0716	0.1368
3 (Single origin, no gene flow)	1.8611	1	0.2498	0.6962	0.7345	0.0404	0.0716	0.1368
4 (Single origin, gene flow)	7.4518	4.054	1	2.7875	2.9411	0.2965	0.5476	0.5476
5 (Parallel origin, gene flow)	2.5337	1.3614	0.34	0.9478	1	0.055	0.0974	0.1862
6 (Single origin, restricted gene flow)	46.0362	24.7355	6.1777	17.22	18.1691	1	1.77	3.3831
7 (Single origin, restricted gene flow)	26.0084	13.8748	3.4302	9.7288	10.265	0.265	1	1.9113
8 (Parallel origin, no gene flow)	13.9274	7.2315	1.608	5.09	5.3705	0.2666	0.3232	1

Table S7 (Continued)	
Tolkenwhale	
Prior distribution (Model1)	
parameter	
distribution	log10_n_AFTO log10_n_FFTO-4 log10_n_FFTO-1 log10_n_FFTO-3 log10_1_FF log10_1_root log_a log_b log_c log_d log_e log_f log_g log_h log_i m n a b c d e f n_AFTO n_FFTO-4 n_FFTO-1 n_FFTO-3 L_FF L_root
Min.	uniform uniform uniform uniform uniform uniform uniform uniform uniform uniform uniform uniform uniform uniform uniform uniform NA NA NA NA NA NA NA NA NA NA
Max.	2 2 2 2 3 3 -9 -9 -9 -9 -9 -9 -9 -9 0.5 0 NA NA NA NA NA NA NA NA NA NA
Med.	5.3 4.7 4.7 4.7 6.2 6.2 -3 -3 -3 -3 -3 -3 -3 -3 2 1 NA NA NA NA NA NA NA NA NA NA
Complex parameter	NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA NA 10*log10_a) 10*log10_b) 10*log10_c) 10*log10_d) 10*log10_e) 10*log10_n_AFTO) 10*log10_n_FFTO-4) 10*log10_n_FFTO-1) 10*log10_n_FFTO-3) 10*log10_L_FF) 10*log10_L_root)
Posterior distribution (Model1)	
parameter name	log10_n_AFTO log10_n_FFTO-4 log10_n_FFTO-1 log10_n_FFTO-3 log10_1_FF log10_1_root log_a log_b log_c log_d log_e log_f log_g log_h log_i m n a b c d e f n_AFTO n_FFTO-4 n_FFTO-1 n_FFTO-3 L_FF L_root
Min.	2.05921339 2.11618202 2.35155284 2.419102158 2.009713725 2.74163814 -9.45001829 -9.635812515 -9.478977609 -9.252902074 -9.629711499 -9.740295909 -0.48811003 -0.017780559 3.396-10 2.316-10 4.446-10 5.20E-10 4.18E-10 2.15E-10 675.2562373 118.8410321 157.1864122 263.8549956 584.8180336 642.5137943
Weighted 2.5% Perc.	3.426424279 2.911132051 2.854444308 3.088151093 2.991271137 3.260720287 -8.9174322 -8.899817743 -8.869917103 -8.809721254 -8.904184258 -8.939457709 0.032328496 0.21614744 1.20E-09 1.30E-09 1.89E-09 1.23E-09 1.31E-09 2.44E-09 2724.214665 768.2501065 1160.300136 962.9130787 2105.654271
Weighted Median	4.050183787 3.397146054 3.62401339 3.710313837 3.255686229 4.444697726 -5.871689247 -5.333141151 -5.648746588 -5.564528959 -4.8927291 0.031159975 0.038770074 1.07E-06 1.22E-06 4.51E-06 2.27E-06 2.83E-06 2.07E-06 11152.73624 3921.533151 4194.082041 5111.48018 1971.531927 28551.57589
Weighted Mean	4.047073388 3.584334357 3.674098775 3.701726334 3.374036875 4.558478855 -5.594912264 -5.354789488 -5.763838326 -5.769524192 -5.244238516 0.020794362 0.016996062 5.99E-05 5.71E-05 8.71E-05 9.72E-05 8.13E-05 0.00011989 14171.88937 5038.407337 6285.432626 6179.892249 163508.3018
Weighted Mode	4.105660297 3.637660558 3.634400247 3.715034348 3.16956544 3.904127149 -3.77003426 -3.63460428 -3.54499358 -3.60359397 -3.696817315 4.137987155 0.600678746 0.336684603 4.84E-07 7.25E-07 1.38E-06 1.00E-06 -3.50E-06 2571.630652 2141.203149 3924.487144 2462.91197 13984.66182
Weighted 97.5% Perc.	4.644603312 4.193663539 4.259667258 4.285119848 4.283031263 2.275147713 -3.254635203 -3.12963252 -3.240459918 -3.251728866 -3.148270183 1.932612241 0.7631989 0.000535864 0.000513347 0.00068436 0.00068436 0.00068436 0.00068436 42881.5475 16185.8368 18084.47646 18383.70345 21171.04172 114381.067
Max.	5.395078444 4.645828158 4.717598474 4.688302729 4.08931587 6.595095056 -2.975490895 -2.79832814 -2.89230284 -2.517720561 -2.234854856 2.541789927 0.931583562 0.00096041 0.001544548 0.00183594 0.001409813 0.001709343 0.000968659 238742.53939 48017.54928 51885.24818 48737.88337 1354889.487 3218707.589
Observed Summary Statistics Value	
Summary statistics	K_FFTO-3 K_FFTO-4 K_FFTO-1 K_AFTO Kwd_FFTO-3 Kwd_FFTO-4 Kwd_FFTO-1 Kwd_AFTO mean_K sd_K tot_K
Observed Value	3.15 3.41 3.45 3.45 2.9873 2.7923 2.7923 7.0957 4.7 2.5935 3.85
Summary statistics	N_FFTO-3 N_FFTO-4 N_FFTO-1 N_AFTO Nwd_FFTO-3 Nwd_FFTO-4 Nwd_FFTO-1 Nwd_AFTO mean_N sd_N tot_N
Observed Value	0.262238 0.305161 0.324758 0.526977 0.264581 0.262001 0.305237 0.338558 0.355554 0.118737 0.474844
Summary statistics	GW_FFTO-3 GW_FFTO-4 GW_FFTO-1 GW_AFTO GWwd_FFTO-3 GWwd_FFTO-4 GWwd_FFTO-1 GWwd_AFTO mean_GW sd_GW tot_GW
Observed Value	0.667812 0.685444 0.607993 0.758999 0.261519 0.238989 0.26461 0.163251 0.679404 0.0611817 0.739408
Summary statistics	NGW_FFTO-3 NGW_FFTO-4 NGW_FFTO-1 NGW_AFTO NGWwd_FFTO-3 NGWwd_FFTO-4 NGWwd_FFTO-1 NGWwd_AFTO mean_NGW sd_NGW
Observed Value	0.307558 0.272595 0.332555 0.545952 0.255957 0.265988 0.305916 0.235951 0.40545 0.187954
Summary statistics	R_FFTO-3 R_FFTO-4 R_FFTO-1 R_AFTO Rwd_FFTO-3 Rwd_FFTO-4 Rwd_FFTO-1 Rwd_AFTO mean_R sd_R tot_R
Observed Value	7.7 7.7 6.6531 8.7773 15.4444 0.0728 8.35347 8.35098 11.4509 8.63352 2.61254
Summary statistics	FIS FST FIT
Observed Value	0.0431308 0.262108 0.259381
Summary statistics	FBI_FFTO-1_FFTO-3 FBI_FFTO-1_FFTO-4 FBI_FFTO-1_FFTO-1 FST_AFTO_FFTO-3 FST_AFTO_FFTO-4 FST_AFTO_FFTO-1
Observed Value	0.233804 0.365177 0.234529 0.274302 0.267181 0.271897
Summary statistics	PI_FFTO-4_FFTO-3 PI_FFTO-1_FFTO-3 PI_FFTO-1_FFTO-4 PI_AFTO_FFTO-3 PI_AFTO_FFTO-4 PI_AFTO_FFTO-1
Observed Value	2.25054 2.26141 2.3052 11.335 11.335 11.335
Summary statistics	DMUSQ_FFTO-4_FFTO-3 DMUSQ_FFTO-1_FFTO-3 DMUSQ_FFTO-1_FFTO-4 DMUSQ_AFTO_FFTO-3 DMUSQ_AFTO_FFTO-4 DMUSQ_AFTO_FFTO-1
Observed Value	0.905784 1.3136 4.6883 14.3849 12.1673 6.54198
Confusion matrix based on 100 samples for each model (neutral, tolerance rate=0.01)	
model focal i selected	1best model) 2 3 4 5 6 7
1 (Single origin, gene flow)	0.4 0.37 0.06 0.06 0.09 0.01 0.01
2 (Parallel origin, gene flow)	0.16 0.06 0.02 0.08 0.07 0 0.01
3 (Single origin, gene flow)	0.16 0.08 0.23 0.26 0.01 0.04 0.02
4 (Parallel origin, gene flow)	0.07 0.12 0.17 0.48 0 0 0.16
5 (Single origin, no gene flow)	0.08 0.04 0 0 0.78 0.012 0.02
6 (Single origin, no gene flow)	0.01 0.01 0.01 0.01 0.35 0.61 0.011
7 (Parallel origin, gene flow)	0.01 0 0.07 0.07 0.04 0.02 0.79
Yellow mark indicates type2 error rate	
Mean model posterior probabilities (neutral, tolerance rate=0.01)	
model focal i selected	1best model) 2 3 4 5 6 7
1 (Single origin, gene flow)	0.3633 0.2337 0.1047 0.1066 0.0995 0.0306 0.0172
2 (Parallel origin, gene flow)	0.3654 0.4204 0.128 0.1232 0.0593 0.0171 0.0266
3 (Single origin, gene flow)	0.1908 0.1739 0.2687 0.2375 0.0302 0.0503 0.0457
4 (Parallel origin, gene flow)	0.0885 0.1214 0.2636 0.0061 0.0121 0.1532 0.0001
5 (Single origin, no gene flow)	0.1213 0.0087 0.0084 0.0078 0.5539 0.2074 0.0526
6 (Single origin, no gene flow)	0.0202 0.0159 0.0276 0.0276 0.4784 0.1162 0.1162
7 (Parallel origin, gene flow)	0.0176 0.0195 0.0545 0.1185 0.0379 0.1097 0.6442
Bayes Factor (neutral, tolerance rate=0.01)	
numerator i denominator	1 2 3 4 5 6 7
1 (Single origin, gene flow)	1.00 1.05309 1.27 86.30 1.24 1.76 1.7689
2 (Parallel origin, gene flow)	0.00 1.00 0.00 0.00 0.00 0.0016 0.0016
3 (Single origin, gene flow)	0.68 737.14 1.00 1.00 1.00 1.1003 1.1003
4 (Parallel origin, gene flow)	0.01 12.27 0.02 1.00 0.02 0.02 0.0198
5 (Single origin, no gene flow)	0.85 702.12 0.85 57.24 1.00 1.14 1.1537
6 (Single origin, no gene flow)	0.57 614.95 0.83 50.13 0.88 1.00 0.993
7 (Parallel origin, gene flow)	0.3718 619.302 0.8401 50.489 0.882 1.0071 1

Table S7 (Continued)
Oligonucleotides[illegible]

parameter name	log10 <i>n</i> AFQK	log10 <i>n</i> FFOK-3	log10 <i>n</i> FFOK-2	log10 <i>n</i> FFOK-1	log10 <i>n</i> FF	log10 <i>n</i> FOK	log10 <i>n</i> FOK-3	log10 <i>n</i> FOK-2	log10 <i>n</i> FOK-1	log10 <i>n</i> F	a	b	c	d	e	<i>n</i> AFQK	<i>n</i> FFOK-3	<i>n</i> FFOK-2	<i>n</i> FFOK-1	<i>n</i> FF	<i>n</i> FOK
Mean	2.103021171	2.103021171	2.103021171	2.103021171	2.103021171	2.103021171	2.103021171	2.103021171	2.103021171	2.103021171	0.42608363	0.00015961	0.00015961	0.00015961	0.00015961	8.88e-10	1.32e-10	2.69e-10	4.05e-14	1.87e-11	1.03e-10
Weighted 2.5% Percent	3.186550594	2.300697523	2.300697523	2.520250748	2.9115504281	3.23125048	4.88341782	8.90814178	4.89232538	3.42535668	-0.96686701	-0.88846263	0.51057032	0.29657171	1.49e-09	1.26e-10	1.42e-09	1.29e-10	5.19e-11	3.87e-08	446.53006242
Weighted 50% Percent	3.369896165	2.301212962	2.301212962	2.520250748	2.9115504281	3.23125048	4.88341782	8.90814178	4.89232538	3.42535668	-0.96686701	-0.88846263	0.51057032	0.29657171	1.49e-09	1.26e-10	1.42e-09	1.29e-10	5.19e-11	3.87e-08	446.53006242
Weighted 97.5% Percent	3.52081093	2.30106882	2.30106882	2.520250748	2.9115504281	3.23125048	4.88341782	8.90814178	4.89232538	3.42535668	-0.96686701	-0.88846263	0.51057032	0.29657171	1.49e-09	1.26e-10	1.42e-09	1.29e-10	5.19e-11	3.87e-08	446.53006242
Max	4.444144995	4.444144995	4.444144995	4.444144995	4.444144995	4.444144995	4.444144995	4.444144995	4.444144995	4.444144995	0.42608363	0.00015961	0.00015961	0.00015961	0.00015961	8.88e-10	1.32e-10	2.69e-10	4.05e-14	1.87e-11	1.03e-10

[illegible]

Observed Value	0.060408	0.373787	0.417008				
Summary statistics	FSL FPK-2 FPK-1	FSL FPK-3 FPK-1	FSL FPK-3 FPK-2	FST AFPK FPK-1	FST AFPK FPK-2	FST AFPK FPK-3	
Observed Value	0.369919	0.468347	0.056567	0.27425	0.390991	0.293323	
Summary statistics	PL FPK-2 FPK-1	PL FPK-3 FPK-1	PL FPK-3 FPK-2	PL AFPK FPK-1	PL AFPK FPK-2	PL AFPK FPK-3	
Observed Value	2.63785	10.2508	9.95937	13.7589	11.6712	11.7597	
Summary statistics	DMUSG FPK-2 FPK-1	DMUSG FPK-3 FPK-1	DMUSG FPK-3 FPK-2	DMUSG AFPK FPK-1	DMUSG AFPK FPK-2	DMUSG AFPK FPK-3	
Observed Value	2.1697	9.9437	9.95937	13.7589	11.6712	11.7597	

model	1 (selected)	2 (best model)	3	4	5
1 (Single origin, gene flow)	0.85	0	0.01	0	0.14
2 (Parallel origin, gene flow)	0	0.73	0.08	0.19	0
3 (Parallel origin, no gene flow)	0	0.03	0.86	0.11	0
4 (Parallel origin, gene flow)	0	0.28	0.17	0.55	0
5 (Single origin, no gene flow)	0.04	0	0	0	0.96

Mean model posterior probabilities (neuralnet, tolerance rate=0.01)					
model focal selected	1 (best model)	2	3	4	5
1 (Single origin, gene flow)		0.0073	0.0083	0.0094	0.1793
2 (Parallel origin, gene flow)	0.0071	0.5949	0.1312	0.2637	0.0031
3 (Parallel origin, no gene flow)	0.0003	0.0726	0.0748	0.1873	0.0022
4 (Parallel origin, gene flow)	0.0028	0.3331	0.1715	0.4481	0.0017
5 (Single origin, no gene flow)	0.1331	0.0032	0.0025	0.0021	0.8591

Bayes Factor (neuralnet, tolerance rate=0.01)					
numerator \ denominator	1	2	3	4	5
1 (Single origin, gene flow)	1.00	544.38	4300.22	1065.07	1.75
2 (Parallel origin, gene flow)	0.00	1.00	7.90	1.96	0.00
3 (Parallel origin, no gene flow)	0.00	0.13	1.00	0.25	0.00
4 (Parallel origin, gene flow)	0.00	0.51	4.04	1.00	0.00
5 (Single origin, no gene flow)	0.57	310.34	2451.54	607.19	1.00

Table S7 (Continued)

Kumejima																				
Prior distribution (Model 6)																				
parameter name	log10 n_AFKM	log10 n_FFKM-2	log10 n_FFKM-1	log10 t_FF	log10 t_root	log10 a	log10 b	log10 c	log10 d	m	p	a	b	c	d	n_AFKM	n_FFKM-2	n_FFKM-1	t_FF	t_root
distribution	uniform	uniform	uniform	uniform	uniform	uniform	uniform	uniform	uniform	uniform	uniform	NA	NA	NA	NA	NA	NA	NA	NA	NA
Min.	2	2	2	3	3	-9	-9	-9	-9	0.5	0	NA	NA	NA	NA	NA	NA	NA	NA	NA
Max.	5.3	4.7	4.7	6.2	6.2	-3	-3	-3	-3	2	1	NA	NA	NA	NA	NA	NA	NA	NA	NA
Complex parameter	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	10*(log10 a)	10*(log10 b)	10*(log10 c)	10*(log10 d)	10*(log10 n_AFKM)	10*(log10 n_FFKM-2)	10*(log10 n_FFKM-1)	10*(log10 n_t_FF)	10*(log10 t_root)

Posterior distribution (Model 6)																				
parameter name	log10 n_AFKM	log10 n_FFKM-2	log10 n_FFKM-1	log10 t_FF	log10 t_root	log a	log c	m	p	a	c	n_AFKM	n_FFKM-2	n_FFKM-1	t_FF	t_root				
Min.	2.416127171	2.004640794	2.195688172	2.571801336	2.506533198	-11.94135065	-11.94135622	0.350029579	0.119638824	9.19E-12	1.66E-12	256.9803544	101.0271525	155.7156975	356.0300794	320.6380399				
Weighted 2.5% Perc.	3.349392787	2.214602631	2.667326393	2.807337785	3.115649457	-8.863888873	-9.459927484	0.51226449	0.34271046	1.35E-09	3.76E-10	2210.427301	163.425521	466.8124526	616.3542291	1306.729313				
Weighted Median	3.90632869	2.812643055	3.191056472	3.222978094	4.226772144	-5.412463207	-6.082777466	1.207503815	0.573843902	3.82E-06	8.86E-07	8050.746861	646.0069959	1561.383519	1619.003405	16957.51468				
Weighted Mean	3.906568278	2.795631761	3.170139926	3.245222085	4.347789913	-5.74975406	-6.468813004	1.208351047	0.568498288	3.59E-05	4.18E-06	9966.727385	765.334502	1680.490703	2064.426182	92120.93428				
Weighted Mode	3.843857634	2.843821278	3.209075825	3.167078368	3.918224242	-4.657233417	-5.454441533	0.711134174	0.571936334	3.88E-06	7.67E-08	5824.187544	505.2645714	1472.223727	1031.97428	5801.871497				
Weighted 97.5% Perc.	4.442243706	3.314469105	3.565015759	3.798429659	5.809010545	-3.801695366	-4.586844626	1.927135382	0.784119195	0.000251621	2.74E-05	27658.72477	2064.018942	3687.98735	6016.463052	64994.7837				
Max.	5.217529632	4.101331001	3.972027084	4.490901997	6.377450717	-2.248523711	-3.633892552	2.06984598	0.866429859	0.005984198	0.000239331	195577.8217	12546.79442	9468.430179	29546.41896	2419445.977				

Observed Summary Statistics Value										
Summary statistics	K_FFKM-1	K_FFKM-2	K_AFKM	Ksd_FFKM-1	Ksd_FFKM-2	Ksd_AFKM	mean_K	sd_K	tot_K	
Observed Value	2.65	1.5	6.85	1.53125	0.688247	5.35355	3.69667	2.81617	8.25	
Summary statistics	H_FFKM-1	H_FFKM-2	H_AFKM	Hsd_FFKM-1	Hsd_FFKM-2	Hsd_AFKM	mean_H	sd_H	tot_H	
Observed Value	0.273834	0.0765779	0.537819	0.219506	0.148743	0.29866	0.296077	0.231424	0.569533	
Summary statistics	GW_FFKM-1	GW_FFKM-2	GW_AFKM	GWsd_FFKM-1	GWsd_FFKM-2	GWsd_AFKM	mean_GW	sd_GW	tot_GW	
Observed Value	0.954947	0.522009	0.671573	0.26537	0.328004	0.225123	0.586176	0.0770089	0.664262	
Summary statistics	NGW_FFKM-1	NGW_FFKM-2	NGW_AFKM	NGWsd_FFKM-1	NGWsd_FFKM-2	NGWsd_AFKM	mean_NGW	sd_NGW		
Observed Value	0.29725	0.175673	0.543082	0.247376	0.135949	0.280676	0.338668	0.187174		
Summary statistics	R_FFKM-1	R_FFKM-2	R_AFKM	Rsd_FFKM-1	Rsd_FFKM-2	Rsd_AFKM	mean_R	sd_R	tot_R	
Observed Value	5.77778	5.25	13.6471	5.07074	3.99106	12.879	8.22495	4.7031	13.95	
Summary statistics	FIS	FST	FIT							
Observed Value	0.0915018	0.478775	0.526468							
Summary statistics	FST_FFKM-2_FFKM-1	FST_AFKM_FFKM-1	FST_AFKM_FFKM-2							
Observed Value	0.541478	0.415315	0.514248							
Summary statistics	PI_FFKM-2_FFKM-1	PI_AFKM_FFKM-1	PI_AFKM_FFKM-2							
Observed Value	7.49005	14.3749	14.2251							
Summary statistics	DMUSQ_FFKM-2_FFKM-1	DMUSQ_AFKM_FFKM-1	DMUSQ_AFKM_FFKM-2							
Observed Value	17.6762	21.6765	31.2208							

Confusion matrix based on 100 samples for each model (neuralnet, tolerance rate=0.01).

model focal \ selected	1	2	3	4	5	6 (best model)	7	8
1 (Parallel origin, gene flow)	26	30	13	3	4	4	5	15
2 (Parallel origin, gene flow)	10	50	21	1	1	5	4	8
3 (Parallel origin, gene flow)	8	15	30	0	3	9	9	26
4 (Single origin, no gene flow)	2	2	0	75	0	9	9	3
5 (Parallel origin, gene flow)	1	24	12	13	19	12	13	6
6 (Single origin, restricted gene flow)	7	19	0	23	6	31	8	
7 (Single origin, restricted gene flow)	3	4	7	28	3	1	34	20
8 (Parallel origin, no gene flow)	1	1	7	4	0	0	3	84

Yellow mark indicates type2 error rate.

Mean model posterior probabilities (neuralnet, tolerance rate=0.01)								
model focal \ selected	1	2	3	4	5	6 (best model)	7	8
1 (Parallel origin, gene flow)	0.24	0.2216	0.1562	0.028	0.1068	0.0708	0.0679	0.1018
2 (Parallel origin, gene flow)	0.1801	0.2934	0.1855	0.0183	0.1169	0.0958	0.0539	0.0561
3 (Parallel origin, gene flow)	0.1604	0.153	0.2137	0.0121	0.1105	0.0707	0.0939	0.1857
4 (Single origin, no gene flow)	0.0308	0.0232	0.0153	0.4492	0.1015	0.1906	0.1533	0.0361
5 (Parallel origin, gene flow)	0.1108	0.1474	0.1272	0.0917	0.2154	0.1435	0.1129	0.0511
6 (Single origin, restricted gene flow)	0.0995	0.1232	0.0632	0.1674	0.1608	0.2594	0.0909	0.0357
7 (Single origin, restricted gene flow)	0.0754	0.0609	0.0921	0.1677	0.109	0.0935	0.2644	0.1369
8 (Parallel origin, no gene flow)	0.1039	0.0553	0.1594	0.0444	0.0429	0.0247	0.093	0.4785

Bayes Factor (neuralnet, tolerance rate=0.01)								
numerator \ denominator	1	2	3	4	5	6 (best model)	7	8
1 (Parallel origin, gene flow)	1	8.3992	6.1755	0.0002	0.0003	0.0002	0.0003	3.2952
2 (Parallel origin, gene flow)	0.1203	1	0.7432	0	0	0	0	0.3918
3 (Parallel origin, gene flow)	0.1619	1.3455	1	0	0.0001	0	0.0001	0.5271
4 (Single origin, no gene flow)	4463.2158	37085.7043	27562.3749	1	1.5558	0.8985	1.5339	14528.6928
5 (Parallel origin, gene flow)	2868.7132	23836.6803	17715.6004	0.6427	1	0.5763	0.8659	9338.2561
6 (Single origin, restricted gene flow)	4978.2329	41365.0782	30742.8378	1.1154	1.7354	1	1.7109	16205.18
7 (Single origin, restricted gene flow)	2909.7535	24177.6918	17969.0427	0.8519	1.0143	0.5845	1	9471.8508
8 (Parallel origin, no gene flow)	0.3072	2.5526	1.8971	0.0001	0.0001	0.0001	0.0001	1

Table S7 (Continued)

Prior distribution (Model 6)																			
parameter name	log10_n_AFIR	log10_n_FFIR-6	log10_n_FFIR-1	log10_n_FFIR-6	log10_l_FF	log10_l_root	log10_b	log10_c	log10_e	log10_f	m	p	b	c	e	f	n_AFIR	n_FFIR-6	n_FFIR-1
distribution	uniform	uniform	uniform	uniform	uniform	uniform	uniform	uniform	uniform	uniform	uniform	uniform	NA	NA	NA	NA	NA	NA	NA
Min.	2	2	2	2	3	3	-9	-9	-9	-9	0.5	0	NA	NA	NA	NA	NA	NA	NA
Max.	5.3	4.7	4.7	4.7	6.2	6.2	-3	-3	-3	-3	2	1	NA	NA	NA	NA	NA	NA	NA
Complex parameter	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	10*log10_b	10*log10_c	10*log10_e	10*log10_f	10*log10_n_AFIR	10*log10_n_FFIR-6	10*log10_n_FFIR-1

posterior distribution (Model 6)																			
parameter name	log10_n_AFIR	log10_n_FFIR-6	log10_n_FFIR-1	log10_n_FFIR-6	log10_l_FF	log10_l_root	log10_b	log10_c	log10_e	log10_f	m	p	b	c	e	f	n_AFIR	n_FFIR-6	n_FFIR-1
Min.	1.915487029	2.21968195	2.15476063	2.441619089	2.747147682	2.473732259	-11.21508222	-9.626500114	-9.801595982	-9.465448388	0.29780942	0.002710558	5.72E-12	2.07E-10	1.47E-10	3.15E-10	84.25277843	159.0992715	131.1922672
Weighted 2.5% Perc.	2.762904326	2.953914766	2.703967649	3.087023057	2.824921302	3.111579657	-9.584977969	-8.615824726	-9.115491716	-8.704787727	0.448573465	0.288991886	2.57E-10	2.54E-09	7.62E-10	1.91E-09	602.2293597	881.255959	518.6263397
Weighted Median	3.715029567	3.650292266	3.232449278	3.772329725	3.297647403	4.258229583	-6.102813347	-5.730094717	-5.695215583	-5.559136046	1.028367829	0.546874289	7.55E-07	1.98E-06	1.95E-06	2.70E-06	5346.47333	4408.27543	1688.73104
Weighted Mean	3.896881817	3.694443527	3.321828394	3.787036101	3.335202315	4.372951088	-6.454943357	-5.819289093	-5.969328487	-5.699428759	1.077953693	0.539573701	7.99E-06	6.28E-05	3.53E-05	8.11E-05	7832.198196	5961.144974	1883.967953
Weighted Mode	3.775550466	3.661780701	3.242585543	3.755713713	3.269077797	3.884817905	-5.046552658	-3.893446638	-4.25135968	-3.722118858	0.871091832	0.57290005	4.38E-07	6.69E-07	2.21E-06	-5.80E-08	2848.357543	2432.654168	1470.13855
Weighted 97.5% Perc.	4.45841304	4.322372895	3.632862996	4.404183987	3.940386625	5.850348299	-4.231739617	-3.323883092	-3.585472019	-3.197150551	1.858439401	0.752515902	5.61E-05	0.000496538	0.00025481	0.000609329	29675.48484	20143.94118	4299.770343
Max.	5.160344328	4.628358594	4.012198264	5.137285861	4.829166001	6.493866008	-3.291207014	-2.849310272	-2.591596757	-2.511623696	2.094541059	0.860199351	0.000477437	0.001502305	0.00244674	0.003012237	147579.5789	40978.71296	10703.12864

Observed Summary Statistics Value												
Summary statistics	K_FFIR-1	K_FFIR-6	K_FFIR-6	K_FFIR-6	Kd_FFIR-1	Kd_FFIR-6	Ksd_FFIR-6	Kd_AFIR	mean_K	sd_K	tot_K	
Observed Value	2.35	3.75	3.75	6.85	1.81442	2.4034	4.35135	4.51051	3.925	1.44309	6.85	
Summary statistics	H_FFIR-1	H_FFIR-6	FFIR-6	H_AFIR	Hd_FFIR-1	Hd_FFIR-6	Hd_FFIR-6	Hd_AFIR	mean_H	sd_H	tot_H	
Observed Value	0.201064	0.388779	0.30482	0.477427	0.223632	0.268172	0.302482	0.262433	0.343023	0.117997	0.542345	
Summary statistics	GW_FFIR-1	GW_FFIR-6	GW_FFIR-6	GW_AFIR	GWsd_FFIR-1	GWsd_FFIR-6	GWsd_FFIR-6	GWsd_AFIR	mean_GW	sd_GW	tot_GW	
Observed Value	0.713468	0.720055	0.542658	0.689957	0.273478	0.26706	0.263749	0.224724	0.666527	0.0835859	0.658672	
Summary statistics	NGW_FFIR-1	NGW_FFIR-6	NGW_FFIR-6	NGW_AFIR	NGWsd_FFIR-1	NGWsd_FFIR-6	NGWsd_FFIR-6	NGWsd_AFIR	mean_NGW	sd_NGW	tot_NGW	
Observed Value	0.216348	0.328606	0.265182	0.30789	0.132975	0.183321	0.148052	0.280571	0.329231	0.126981		
Summary statistics	R_FFIR-1	R_FFIR-6	R_FFIR-6	R_AFIR	Rad_FFIR-1	Rad_FFIR-6	Rad_FFIR-6	Rad_AFIR	mean_R	sd_R	tot_R	
Observed Value	5.92857	6.22222	16.7692	8.63158	10.4695	5.54718	27.7853	6.29118	9.3879	5.06769	17.45	
Summary statistics	FIS	FST	FIT									
Observed Value	0.0589495	0.399616	0.435008									
Summary statistics	FST_FFIR-6_FFIR-1	FST_FFIR-6_FFIR-1	FST_FFIR-6_FFIR-6	FST_FFIR-6_FFIR-1	FST_AFIRIR_FFIR-6	FST_AFIR_FFIR-6						
Observed Value	0.304243	0.455915	0.294781	0.48297	0.35409	0.39128						
Summary statistics	PI_FFIR-6_FFIR-1	PI_FFIR-6_FFIR-1	PI_FFIR-6_FFIR-6	PI_FFIR-6_FFIR-6	PI_AFIRIR_FFIR-6	PI_AFIR_FFIR-6						
Observed Value	8.36632	9.52423	9.57088	14.1875	13.5054	12.9078						
Summary statistics	DMUSQ_FFIR-6_FFIR-1	DMUSQ_FFIR-6_FFIR-1	DMUSQ_FFIR-6_FFIR-6	DMUSQ_FFIR-6_FFIR-1	DMUSQ_AFIRIR_FFIR-6	DMUSQ_AFIR_FFIR-6						
Observed Value	7.36113	11.5983	3.47136	25.9701	14.0454	20.1645						

Confusion matrix based on 100 samples for each model (neuralnet, tolerance rate=0.01).

model focal \ selected	1	2	3	4	5	6(best model)
1 (Single origin, no gene flow)	0.9	0.03	0.01	0	0	0.06
2 (Parallel origin, no gene flow)	0.02	0.88	0	0.07	0.02	0.01
3 (Single origin, gene flow)	0.16	0.03	0.29	0.1	0.22	0.2
4 (Parallel origin, gene flow)	0	0.25	0.03	0.48	0.19	0.05
5 (Parallel origin, gene flow)	0.04	0.09	0.09	0.16	0.47	0.15
6 (Single origin, gene flow)	0.14	0.11	0.1	0.07	0.1	0.48

Yellow mark indicates type2 error rate.

Mean model posterior probabilities (neuralnet, tolerance rate=0.01)

model focal \ selected	1	2	3	4	5	6 (best model)
1 (Single origin, no gene flow)	0.6892	0.0294	0.0885	0.0044	0.0271	0.1743
2 (Parallel origin, no gene flow)	0.0028	0.8313	0.0099	0.1906	0.0067	0.0536
3 (Single origin, gene flow)	0.123	0.0442	0.2661	0.1296	0.2042	0.1939
4 (Parallel origin, gene flow)	0.0059	0.2157	0.1187	0.3654	0.2155	0.0808
5 (Parallel origin, gene flow)	0.0368	0.0825	0.205	0.197	0.3357	0.1429
6 (Single origin, gene flow)	0.1299	0.0732	0.2285	0.0924	0.1365	0.2595

Bayes Factor (neuralnet, tolerance rate=0.01)

numerator \ denominator	1	2	3	4	5	6
1 (Single origin, no gene flow)	1	9404.0848	1.7085	234566.4108	2458.9742	0.7571
2 (Parallel origin, no gene flow)	0	1	0	2.4947	0.0062	0
3 (Single origin, gene flow)	0.5853	55033.272	1	137294.1532	1439.2631	0.4432
4 (Parallel origin, gene flow)	0	0.4008	0	1	0.0105	0
5 (Parallel origin, gene flow)	0.0004	39.2371	0.0007	96.392	1	0.0003
6 (Single origin, gene flow)	1.3208	124183.7803	2.2565	309807.2552	3247.7286	1

Table S7 (Continued)
Tokunoshima, Okinoerabujima
Prior distribution (Model2)

parameter name	log10_a	log10_b	log10_c	log10_d	log10_n_AF	log10_n_okino	log10_n_toku	log10_t_FF	log10_t_root	m	p	a	b	c	d	n_AF	n_okino	n_toku	t_FF	t_root
distribution	uniform	uniform	uniform	uniform	uniform	uniform	uniform	uniform	uniform	uniform	uniform	NA	NA	NA	NA	NA	NA	NA	NA	NA
Min.	-9	-9	-9	-9	2	2	2	3	3	0.5	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Max.	-3	-3	-3	-3	5.3	4.7	4.7	6.2	6.2	2	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Complex parameter	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	10*(log10_a)	10*(log10_b)	10*(log10_c)	10*(log10_d)	10*(log10_n_AF)	10*(log10_n_okino)	10*(log10_n_toku)	10*(log10_t_FF)	10*(log10_t_root)

Posterior distribution (Model2)

parameter name	log10_a	log10_b	log10_c	log10_d	log10_n_AF	log10_n_okino	log10_n_toku	log10_t_FF	log10_t_root	m	p	a	b	c	d	n_AF	n_okino	n_toku	t_FF	t_root
Min.	-10.14958659	-11.11341166	-11.45255462	-11.16476095	2.776673139	2.061336903	2.360266768	2.074409034	2.401879147	0.404548429	0.003573449	7.15E-11	6.89E-12	3.48E-12	6.82E-12	591.8862858	114.8724309	227.2451764	118.0719184	250.5689687
Weighted 2.5% Perc.	-9.202913073	-9.023659958	-8.941294563	-8.899265843	3.43063722	2.558173848	3.069487412	2.585305711	3.074098215	0.514615452	0.242243218	6.31E-10	9.55E-10	1.13E-09	1.21E-09	2738.742395	360.0914213	1161.04389	382.5846214	1196.245616
Weighted Median	-5.961562108	-5.348512608	-5.909820868	-5.452163411	4.050707087	3.27214837	3.710535377	3.241885296	4.424051541	1.146281297	0.532173827	1.10E-06	4.49E-06	1.22E-06	3.45E-06	11172.29011	1869.353565	5143.961394	1746.746089	26421.55556
Weighted Mean	-6.116380114	-5.784742914	-6.09210827	-5.690588205	4.053608151	3.246556204	3.698837717	3.347693538	4.452163395	1.176857129	0.524192096	3.82E-05	4.56E-05	1.75E-05	3.99E-05	14562.28311	2244.978483	6264.689698	7467.92739	113792.1471
Weighted Mode	-4.3469019	-4.50111073	-5.950859573	-5.511585451	4.033503063	3.328487865	3.714023312	3.1567225	3.890017278	0.666153834	0.579110955	-2.36E-06	-2.27E-06	-8.21E-07	2.27E-06	7453.902105	1283.334464	3690.010605	661.3093564	8150.788215
Weighted 97.5% Perc.	-3.500079707	-3.449034485	-3.847886318	-3.449370538	4.67254187	3.822723853	4.251151613	4.79650326	5.847947266	1.924979731	0.757162432	0.000315004	0.000348579	0.000140908	0.000347709	46575.89176	6619.335656	17901.73924	62485.42911	701379.1444
Max.	-2.542248289	-2.359737163	-2.920104543	-2.606485662	5.102251631	4.405683442	4.688525192	5.832934866	6.368778078	2.070120686	0.920796071	0.002853665	0.004411571	0.001196782	0.002424363	124846.5407	24946.85325	47142.42174	675162.9522	2326504.085

Observed Summary Statistics Value

Summary statistics	K_okino	K_toku	K_AF	Ksd_okino	Ksd_toku	Ksd_AF	mean_K	sd_K	tot_K
Observed Value	2.35	4.4	8.45	1.53125	4.41767	7.08947	5.06667	3.10416	9.4
Summary statistics	H_okino	H_toku	H_AF	Had_okino	Had_toku	Had_AF	mean_H	sd_H	tot_H
Observed Value	0.226868	0.324758	0.526077	0.268593	0.305237	0.338558	0.359234	0.152555	0.4863
Summary statistics	GW_okino	GW_toku	GW_AF	GWsd_okino	GWsd_toku	GWsd_AF	mean_GW	sd_GW	tot_GW
Observed Value	0.586317	0.607993	0.796569	0.241392	0.26461	0.163251	0.650293	0.0926739	0.722554
Summary statistics	NGW_okino	NGW_toku	NGW_AF	NGWsd_okino	NGWsd_toku	NGWsd_AF	mean_NGW	sd_NGW	
Observed Value	0.246524	0.382195	0.649745	0.197962	0.202091	0.224729	0.426155	0.205173	
Summary statistics	R_okino	R_toku	R_AF	Rad_okino	Rad_toku	Rad_AF	mean_R	sd_R	tot_R
Observed Value	6.38462	8.27778	12.4444	7.98476	8.25908	11.4029	9.03561	3.10018	12.45
Summary statistics	FIS	FST	FIT						
Observed Value	0.0437566	0.299879	0.330513						
Summary statistics	FST_toku_okino	FST_AF_okino	FST_AF_toku						
Observed Value	0.343439	0.303634	0.271697						
Summary statistics	PI_toku_okino	PI_AF_okino	PI_AF_toku						
Observed Value	8.33941	10.9974	11.7025						
Summary statistics	DMUSQ_toku_okino	DMUSQ_AF_okino	DMUSQ_AF_toku						
Observed Value	1.69281	3.63525	6.54198						

Confusion matrix based on 100 samples for each model (neuralnet, tolerance rate=0.01).

model focal \ selected	1	2 (best model)	3	4	5
1 (Single origin, no gene flow)	0.85	0.08	0.01	0.05	0.01
2 (Single origin, gene flow)	0.23	0.29	0.16	0.17	0.15
3 (Parallel origin, gene flow)	0.02	0.07	0.37	0.24	0.3
4 (Parallel origin, gene flow)	0.08	0.1	0.22	0.37	0.23
5 (Parallel origin, gene flow)	0.04	0.06	0.17	0.15	0.58

Yellow mark indicates type2 error rate.

Mean model posterior probabilities (neuralnet, tolerance rate=0.01)

model focal \ selected	1	2 (best model)	3	4	5
1 (Single origin, no gene flow)	0.7214	0.1837	0.0197	0.0531	0.022
2 (Single origin, gene flow)	0.2064	0.3135	0.1605	0.184	0.1356
3 (Parallel origin, gene flow)	0.0195	0.1454	0.3236	0.2543	0.2572
4 (Parallel origin, gene flow)	0.08	0.1616	0.2381	0.2938	0.2265
5 (Parallel origin, gene flow)	0.0314	0.1381	0.2532	0.2236	0.3537

Bayes Factor (neuralnet, tolerance rate=0.01)

numerator \ denominator	1	2	3	4	5
1 (Single origin, no gene flow)	1.00	0.21	17.36	2.01	26.19
2 (Single origin, gene flow)	4.68	1.00	81.23	9.42	122.57
3 (Parallel origin, gene flow)	0.06	0.01	1.00	0.12	1.51
4 (Parallel origin, gene flow)	0.50	0.11	8.62	1.00	13.01
5 (Parallel origin, gene flow)	0.04	0.01	0.66	0.08	1.00

Table S7 (Continued)
South Ryukyu
Prior distribution (Model 5)

parameter name	log10_a	log10_b	log10_c	log10_d	log10_n_AF	log10_n_ishigaki	log10_n_iriomote	log10_t_FF	log10_t_root	m	p	a	b	c	d	n_AF	n_ishigaki	n_iriomote	t_FF	t_root
distribution	uniform	uniform	uniform	uniform	uniform	uniform	uniform	uniform	uniform	uniform	uniform	NA	NA	NA	NA	NA	NA	NA	NA	NA
Min.	-9	-9	-9	-9	2	2	2	3	3	0.5	0	NA	NA	NA	NA	NA	NA	NA	NA	NA
Max.	-3	-3	-3	-3	5.3	4.7	4.7	6.2	6.2	2	1	NA	NA	NA	NA	NA	NA	NA	NA	NA
Complex parameter	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	10*(log10_a)	10*(log10_b)	10*(log10_c)	10*(log10_b)	10*(log10_n_AF)	10*(log10_ishigaki)	10*(log10_iriomote)	10*(log10_t_FF)	10*(log10_t_root)

Posterior distribution (Model5)

parameter name	log10_a	log10_b	log10_c	log10_d	log10_n_AF	log10_n_ishigaki	log10_n_iriomote	log10_t_FF	log10_t_root	m	p	a	b	c	d	n_AF	n_ishigaki	n_iriomote	t_FF	t_root
Min.	-10.41018112	-10.10277534	-10.00713071	-10.50807508	2.611704817	1.993488455	2.135259187	2.789609709	2.818403745	0.43645659	0.0003027	3.36E-11	8.17E-11	9.95E-11	3.66E-11	431.1793174	99.24520694	138.3037164	630.2111403	618.7346758
Weighted 2.5 Perc.	-9.134330709	-8.683549288	-8.826107389	-8.853848892	3.247395058	2.612327062	2.906432549	3.250282209	3.60541131	0.541477846	0.29644687	7.21E-10	2.02E-09	1.49E-09	9.30E-09	1763.714134	413.9688356	819.0888356	1750.357478	4025.782063
Weighted Median	-5.349287518	-5.262047439	-5.017281743	-4.771362201	3.893367898	3.269444635	3.626680308	3.818297744	4.88055598	1.238407052	0.529120986	4.53E-06	5.50E-06	9.72E-06	1.69E-05	7846.83798	1857.881138	4234.648740	6589.028835	76234.0719
Weighted Mean	-5.846756557	-5.584088808	-5.528189056	-5.072548038	3.890743197	3.251866104	3.602818945	3.895147357	4.911209857	1.244904324	0.521514117	3.65E-05	7.64E-05	6.14E-05	0.000168655	10348.16041	2219.446264	5216.152089	20220.76185	316656.6385
Weighted Mode	-4.386709355	-4.239063231	-4.118923658	-4.025084745	3.887053211	3.292880034	3.694422564	3.780748221	4.37540734	0.745433858	0.532950304	2.87E-06	-7.07E-07	3.26E-06	7.90E-06	5052.568074	1390.244892	2791.688958	5319.192251	17784.51235
Weighted 97.5 Perc.	-3.622195581	-3.200261519	-3.022195581	-2.87117254	4.518010446	3.778079715	4.191778683	5.140157887	6.321592804	1.959613182	0.713528915	0.000248307	0.000639776	0.000406212	0.001303584	33213.58103	5983.949561	15377.31758	140239.5293	2094257.497
Max.	-2.730132231	-2.0709559	-2.358728293	-1.925656543	5.236961892	4.498616253	4.695800924	6.247419837	6.858398816	2.054198232	0.824551573	0.001899799	0.00860596	0.004533259	0.011842609	175350.0396	31366.23533	44667.67009	1823148.674	7157235.775

Observed Summary Statistics Value

Summary statistics	K_iriomote	K_ishigaki	K_AF	Ksd_iriomote	Ksd_ishigaki	Ksd_AF	mean_K	sd_K	tot_K
Observed Value	3.75	3.1	7.3	2.4034	1.33377	4.78044	4.71667	2.26072	8.8
Summary statistics	H_iriomote	H_ishigaki	H_AF	Hsd_iriomote	Hsd_ishigaki	Hsd_AF	mean_H	sd_H	tot_H
Observed Value	0.388779	0.299456	0.50118	0.288172	0.241195	0.260336	0.396472	0.101081	0.569886
Summary statistics	GW_iriomote	GW_ishigaki	GW_AF	GWsd_iriomote	GWsd_ishigaki	GWsd_AF	mean_GW	sd_GW	tot_GW
Observed Value	0.720055	0.660367	0.653008	0.26706	0.260751	0.245748	0.67771	0.0368415	0.648847
Summary statistics	NGW_iriomote	NGW_ishigaki	NGW_AF	NGWsd_iriomote	NGWsd_ishigaki	NGWsd_AF	mean_NGW	sd_NGW	
Observed Value	0.324162	0.267827	0.539543	0.216599	0.168378	0.223509	0.377044	0.143369	
Summary statistics	R_iriomote	R_ishigaki	R_AF	Rsd_iriomote	Rsd_ishigaki	Rsd_AF	mean_R	sd_R	tot_R
Observed Value	6.22222	5.05882	12.0526	5.54718	2.92555	7.13733	7.77789	3.74746	13.35
Summary statistics	FIS	FST	FIT						
Observed Value	0.0909702	0.343889	0.409029						
Summary statistics	FST_ishigaki_iriomote	FST_AF_iriomote	FST_AF_ishigaki						
Observed Value	0.501156	0.324191	0.331076						
Summary statistics	PI_ishigaki_iriomote	PI_AF_iriomote	PI_AF_ishigaki						
Observed Value	13.4829	13.3516	12.4218						
Summary statistics	DMUSQ_ishigaki_iriomote	DMUSQ_AF_iriomote	DMUSQ_AF_ishigaki						
Observed Value	13.1287	12.8723	7.13103						

Confusion matrix based on 100 samples for each model (neuralnet, tolerance rate=0.01).

model focal \ selected	1	2	3	4	5 (best model)
1 (Single origin, no gene flow)	0.79	0.14	0.01	0	0.06
2 (Parallel origin, no gene flow)	0.03	0.84	0.04	0.01	0.06
3 (Parallel origin, no gene flow)	0.02	0.33	0.62	0	0.03
4 (Parallel origin, gene flow)	0	0	0	100	0
5 (Single origin, Gene flow,best model)	0.3	0.11	0.04	0	0.55

Yellow mark indicates typic error rate

Mean model posterior probabilities (neuralnet, tolerance rate=0.01)

model focal \ selected	1	2	3	4	5 (best model)
1 (Single origin, no gene flow)	0.461226746	0.088258646	0.043990076	0.000303527	0.162031194
2 (Parallel origin, no gene flow)	0.088916008	0.552662745	0.234758924	0.000154347	0.068301858
3 (Parallel origin, no gene flow)	0.000454802	0.274154931	0.499342396	0.015675549	0.081498049
4 (Parallel origin, gene flow)	0.000361667	0.00125748	0.023494941	0.477026376	0.008149881
5 (Single origin, Gene flow,best model)	0.178577648	0.098446936	0.090713468	0.010823468	0.384182543

Bayes Factor (neuralnet, tolerance rate=0.01)

numerator \ denominator	1	2	3	4	5
1 (Single origin, no gene flow)	1	14.01	2.7125	6.1501	0.0733
2 (Parallel origin, no gene flow)	0.0714	1	0.1936	0.4418	0.0552
3 (Parallel origin, no gene flow)	0.3687	5.1649	1	2.282	0.027
4 (Parallel origin, gene flow)	0.1615	2.2633	0.4382	1	0.0118
5 (Single origin, Gene flow,best model)	13.6333	191.002	36.9807	84.3915	1

Table S7 (Continued)

Observed Summary Statistics Value																			
Summary statistics	K, Hromosome	K, koku	K, ohromosome	K, loku	K, amant	K, AF, north	K, AF, south	Kad, Hromosome	Kad, koku	Kad, ohromosome	Kad, loku	Kad, amant	Kad, AF, north	Kad, AF, south	mean, K	st, K	tot, K		
Observed Value	3.9	2.7	3.1	4.4	4.6	7.4	13.3	2.39273	1.39273	1.44732	4.43346	5.21728	4.17059	9.33717	5.52143	3.35057	18.73		
summary statistics	R, Hromosome	R, koku	R, ohromosome	R, loku	R, amant	R, AF, north	R, AF, south	Rad, Hromosome	Rad, koku	Rad, ohromosome	Rad, loku	Rad, amant	Rad, AF, north	Rad, AF, south	mean, R	st, R	tot, R		
Observed Value	0.420209	0.292053	0.326	0.338737	0.380259	0.330259	0.324718	0.28851	0.229413	0.231458	0.26033	0.232812	0.239803	0.281136	0.359027	0.102554	0.038989		
summary statistics	GW, Hromosome	GW, koku	GW, ohromosome	GW, loku	GW, amant	GW, AF, north	GW, AF, south	GWad, Hromosome	GWad, koku	GWad, ohromosome	GWad, loku	GWad, amant	GWad, AF, north	GWad, AF, south	mean, GW	st, GW	tot, GW		
Observed Value	0.58932	0.44932	0.47831	0.57851	0.581321	0.57852	0.57852	0.58131	0.57852	0.58131	0.57852	0.58131	0.57852	0.58131	0.58131	0.041321	0.00000		
summary statistics	NGW, Hromosome	NGW, koku	NGW, ohromosome	NGW, loku	NGW, amant	NGW, AF, north	NGW, AF, south	NGWad, Hromosome	NGWad, koku	NGWad, ohromosome	NGWad, loku	NGWad, amant	NGWad, AF, north	NGWad, AF, south	mean, NGW	st, NGW	tot, NGW		
Observed Value	0.10734	0.09431	0.09431	0.10734	0.10734	0.10734	0.10734	0.10734	0.10734	0.10734	0.10734	0.10734	0.10734	0.10734	0.10734	0.077322	0.00000		
summary statistics	R, Hromosome	R, koku	R, ohromosome	R, loku	R, amant	R, AF, north	R, AF, south	Rad, Hromosome	Rad, koku	Rad, ohromosome	Rad, loku	Rad, amant	Rad, AF, north	Rad, AF, south	mean, R	st, R	tot, R		
Observed Value	20.5557	14.4444	14.4444	14.4444	15.5555	24	43	33.3768	35.3333	35.3333	37.3333	35.3333	35.3333	34.3333	25.5555	15.4444	14.4444		
summary statistics	FSR	FSR	FSR	FSR	FSR	FSR	FSR	FSR	FSR	FSR	FSR	FSR	FSR	FSR	FSR	FSR	FSR		
Observed Value	0.13333	0.38889	0.44444	0.38889	0.44444	0.38889	0.44444	0.13333	0.38889	0.44444	0.38889	0.44444	0.38889	0.44444	0.38889	0.44444	0.38889		
summary statistics	FSR, Hromosome	FSR, koku	FSR, ohromosome	FSR, loku	FSR, amant	FSR, AF, north	FSR, AF, south	FSRad, Hromosome	FSRad, koku	FSRad, ohromosome	FSRad, loku	FSRad, amant	FSRad, AF, north	FSRad, AF, south	mean, FSR	st, FSR	tot, FSR		
Observed Value	0.13333	0.38889	0.44444	0.38889	0.44444	0.38889	0.44444	0.13333	0.38889	0.44444	0.38889	0.44444	0.38889	0.44444	0.38889	0.44444	0.38889		
summary statistics	PL, Hromosome	PL, koku	PL, ohromosome	PL, loku	PL, amant	PL, AF, north	PL, AF, south	PLad, Hromosome	PLad, koku	PLad, ohromosome	PLad, loku	PLad, amant	PLad, AF, north	PLad, AF, south	mean, PL	st, PL	tot, PL		
Observed Value	15.7151	15.7151	15.7151	15.7151	15.7151	15.7151	15.7151	15.7151	15.7151	15.7151	15.7151	15.7151	15.7151	15.7151	15.7151	0.00000	0.00000		
summary statistics	DMUS2, koku, Hromosome	DMUS2, koku, koku	DMUS2, koku, ohromosome	DMUS2, koku, loku	DMUS2, koku, amant	DMUS2, koku, AF, north	DMUS2, koku, AF, south	DMUS2, koku, Rad, Hromosome	DMUS2, koku, Rad, koku	DMUS2, koku, Rad, ohromosome	DMUS2, koku, Rad, loku	DMUS2, koku, Rad, amant	DMUS2, koku, Rad, AF, north	DMUS2, koku, Rad, AF, south	mean, DMUS2	st, DMUS2	tot, DMUS2		
Observed Value	15.7151	15.7151	15.7151	15.7151	15.7151	15.7151	15.7151	15.7151	15.7151	15.7151	15.7151	15.7151	15.7151	15.7151	15.7151	0.00000	0.00000		
summary statistics	DMUS2, AF, Hromosome	DMUS2, AF, koku	DMUS2, AF, ohromosome	DMUS2, AF, loku	DMUS2, AF, amant	DMUS2, AF, AF, north	DMUS2, AF, AF, south	DMUS2, AF, Rad, Hromosome	DMUS2, AF, Rad, koku	DMUS2, AF, Rad, ohromosome	DMUS2, AF, Rad, loku	DMUS2, AF, Rad, amant	DMUS2, AF, Rad, AF, north	DMUS2, AF, Rad, AF, south	mean, DMUS2	st, DMUS2	tot, DMUS2		
Observed Value	15.7151	15.7151	15.7151	15.7151	15.7151	15.7151	15.7151	15.7151	15.7151	15.7151	15.7151	15.7151	15.7151	15.7151	15.7151	0.00000	0.00000		
summary statistics	DMUS2, AF, AF, Hromosome	DMUS2, AF, AF, koku	DMUS2, AF, AF, ohromosome	DMUS2, AF, AF, loku	DMUS2, AF, AF, amant	DMUS2, AF, AF, AF, north	DMUS2, AF, AF, AF, south	DMUS2, AF, AF, Rad, Hromosome	DMUS2, AF, AF, Rad, koku	DMUS2, AF, AF, Rad, ohromosome	DMUS2, AF, AF, Rad, loku	DMUS2, AF, AF, Rad, amant	DMUS2, AF, AF, Rad, AF, north	DMUS2, AF, AF, Rad, AF, south	mean, DMUS2	st, DMUS2	tot, DMUS2		
Observed Value	15.7151	15.7151	15.7151	15.7151	15.7151	15.7151	15.7151	15.7151	15.7151	15.7151	15.7151	15.7151	15.7151	15.7151	15.7151	0.00000	0.00000		
summary statistics	DMUS2, AF, AF, AF, Hromosome	DMUS2, AF, AF, AF, koku	DMUS2, AF, AF, AF, ohromosome	DMUS2, AF, AF, AF, loku	DMUS2, AF, AF, AF, amant	DMUS2, AF, AF, AF, AF, north	DMUS2, AF, AF, AF, AF, south	DMUS2, AF, AF, AF, Rad, Hromosome	DMUS2, AF, AF, AF, Rad, koku	DMUS2, AF, AF, AF, Rad, ohromosome	DMUS2, AF, AF, AF, Rad, loku	DMUS2, AF, AF, AF, Rad, amant	DMUS2, AF, AF, AF, Rad, AF, north	DMUS2, AF, AF, AF, Rad, AF, south	mean, DMUS2	st, DMUS2	tot, DMUS2		
Observed Value	15.7151	15.7151	15.7151	15.7151	15.7151	15.7151	15.7151	15.7151	15.7151	15.7151	15.7151	15.7151	15.7151	15.7151	15.7151	0.00000	0.00000		
summary statistics	DMUS2, AF, AF, AF, AF, Hromosome	DMUS2, AF, AF, AF, AF, koku	DMUS2, AF, AF, AF, AF, ohromosome	DMUS2, AF, AF, AF, AF, loku	DMUS2, AF, AF, AF, AF, amant	DMUS2, AF, AF, AF, AF, AF, north	DMUS2, AF, AF, AF, AF, AF, south	DMUS2, AF, AF, AF, AF, Rad, Hromosome	DMUS2, AF, AF, AF, AF, Rad, koku	DMUS2, AF, AF, AF, AF, Rad, ohromosome	DMUS2, AF, AF, AF, AF, Rad, loku	DMUS2, AF, AF, AF, AF, Rad, amant	DMUS2, AF, AF, AF, AF, Rad, AF, north	DMUS2, AF, AF, AF, AF, Rad, AF, south	mean, DMUS2	st, DMUS2	tot, DMUS2		
Observed Value	15.7151	15.7151	15.7151	15.7151	15.7151	15.7151	15.7151	15.7151	15.7151	15.7151	15.7151	15.7151	15.7151	15.7151	15.7151	0.00000	0.00000		
summary statistics	DMUS2, AF, AF, AF, AF, AF, Hromosome	DMUS2, AF, AF, AF, AF, AF, koku	DMUS2, AF, AF, AF, AF, AF, ohromosome	DMUS2, AF, AF, AF, AF, AF, loku	DMUS2, AF, AF, AF, AF, AF, amant	DMUS2, AF, AF, AF, AF, AF, AF, north	DMUS2, AF, AF, AF, AF, AF, AF, south	DMUS2, AF, AF, AF, AF, AF, Rad, Hromosome	DMUS2, AF, AF, AF, AF, AF, Rad, koku	DMUS2, AF, AF, AF, AF, AF, Rad, ohromosome	DMUS2, AF, AF, AF, AF, AF, Rad, loku	DMUS2, AF, AF, AF, AF, AF, Rad, amant	DMUS2, AF, AF, AF, AF, AF, Rad, AF, north	DMUS2, AF, AF, AF, AF, AF, Rad, AF, south	mean, DMUS2	st, DMUS2	tot, DMUS2		
Observed Value	15.7151	15.7151	15.7151	15.7151	15.7151	15.7151	15.7151	15.7151	15.7151	15.7151	15.7151	15.7151	15.7151	15.7151	15.7151	0.00000	0.00000		
summary statistics	DMUS2, AF, AF, AF, AF, AF, AF, Hromosome	DMUS2, AF, AF, AF, AF, AF, AF, koku	DMUS2, AF, AF, AF, AF, AF, AF, ohromosome	DMUS2, AF, AF, AF, AF, AF, AF, loku	DMUS2, AF, AF, AF, AF, AF, AF, amant	DMUS2, AF, AF, AF, AF, AF, AF, AF, north	DMUS2, AF, AF, AF, AF, AF, AF, AF, south	DMUS2, AF, AF, AF, AF, AF, AF, Rad, Hromosome	DMUS2, AF, AF, AF, AF, AF, AF, Rad, koku	DMUS2, AF, AF, AF, AF, AF, AF, Rad, ohromosome	DMUS2, AF, AF, AF, AF, AF, AF, Rad, loku	DMUS2, AF, AF, AF, AF, AF, AF, Rad, amant	DMUS2, AF, AF, AF, AF, AF, AF, Rad, AF, north	DMUS2, AF, AF, AF, AF, AF, AF, Rad, AF, south	mean, DMUS2	st, DMUS2	tot, DMUS2		
Observed Value	15.7151	15.7151	15.7151	15.7151	15.7151	15.7151	15.7151	15.7151	15.7151	15.7151	15.7151	15.7151	15.7151	15.7151	15.7151	0.00000	0.00000		
summary statistics	DMUS2, AF, AF, AF, AF, AF, AF, AF, Hromosome	DMUS2, AF, AF, AF, AF, AF, AF, AF, koku	DMUS2, AF, AF, AF, AF, AF, AF, AF, ohromosome	DMUS2, AF, AF, AF, AF, AF, AF, AF, loku	DMUS2, AF, AF, AF, AF, AF, AF, AF, amant	DMUS2, AF, AF, AF, AF, AF, AF, AF, AF, north	DMUS2, AF, AF, AF, AF, AF, AF, AF, AF, south	DMUS2, AF, AF, AF, AF, AF, AF, AF, Rad, Hromosome	DMUS2, AF, AF, AF, AF, AF, AF, AF, Rad, koku	DMUS2, AF, AF, AF, AF, AF, AF, AF, Rad, ohromosome	DMUS2, AF, AF, AF, AF, AF, AF, AF, Rad, loku	DMUS2, AF, AF, AF, AF, AF, AF, AF, Rad, amant	DMUS2, AF, AF, AF, AF, AF, AF, AF, Rad, AF, north	DMUS2, AF, AF, AF, AF, AF, AF, AF, Rad, AF, south	mean, DMUS2	st, DMUS2	tot, DMUS2		
Observed Value	15.7151	15.7151	15.7151	15.7151	15.7151	15.7151	15.7151	15.7151	15.7151	15.7151	15.7151	15.7151	15.7151	15.7151	15.7151	0.00000	0.00000		
summary statistics	DMUS2, AF, AF, AF, AF, AF, AF, AF, AF, Hromosome	DMUS2, AF, AF, AF, AF, AF, AF, AF, AF, koku	DMUS2, AF, AF, AF, AF, AF, AF, AF, AF, ohromosome	DMUS2, AF, AF, AF, AF, AF, AF, AF, AF, loku	DMUS2, AF, AF, AF, AF, AF, AF, AF, AF, amant	DMUS2, AF, AF, AF, AF, AF, AF, AF, AF, AF, north	DMUS2, AF, AF, AF, AF, AF, AF, AF, AF, AF, south	DMUS2, AF, AF, AF, AF, AF, AF, AF, AF, Rad, Hromosome	DMUS2, AF, AF, AF, AF, AF, AF, AF, AF, Rad, koku	DMUS2, AF, AF, AF, AF, AF, AF, AF, AF, Rad, ohromosome	DMUS2, AF, AF, AF, AF, AF, AF, AF, AF, Rad, loku	DMUS2, AF, AF, AF, AF, AF, AF, AF, AF, Rad, amant	DMUS2, AF, AF, AF, AF, AF, AF, AF, AF, Rad, AF, north	DMUS2, AF, AF, AF, AF, AF, AF, AF, AF, Rad, AF, south	mean, DMUS2	st, DMUS2	tot, DMUS2		
Observed Value	15.7151	15.7151	15.7151	15.7151	15.7151	15.7151	15.7151	15.7151	15.7151	15.7151	15.7151	15.7151	15.7151	15.7151	15.7151	0.00000	0.00000		
summary statistics	DMUS2, AF, AF, AF, AF, AF, AF, AF, AF, AF, Hromosome	DMUS2, AF, AF, AF, AF, AF, AF, AF, AF, AF, koku	DMUS2, AF, AF, AF, AF, AF, AF, AF, AF, AF, ohromosome	DMUS2, AF, AF, AF, AF, AF, AF, AF, AF, AF, loku	DMUS2, AF, AF, AF, AF, AF, AF, AF, AF, AF, amant	DMUS2, AF, AF, AF, AF, AF, AF, AF, AF, AF, AF, north	DMUS2, AF, AF, AF, AF, AF, AF, AF, AF, AF, AF, south	DMUS2, AF, AF, AF, AF, AF, AF, AF, AF, AF, Rad, Hromosome	DMUS2, AF, AF, AF, AF, AF, AF, AF, AF, AF, Rad, koku	DMUS2, AF, AF, AF, AF, AF, AF, AF, AF, AF, Rad, ohromosome	DMUS2, AF, AF, AF, AF, AF, AF, AF, AF, AF, Rad, loku	DMUS2, AF, AF, AF, AF, AF, AF, AF, AF, AF, Rad, amant	DMUS2, AF, AF, AF, AF, AF, AF, AF, AF, AF, Rad, AF, north	DMUS2, AF, AF, AF, AF, AF, AF, AF, AF, AF, Rad, AF, south	mean, DMUS2	st, DMUS2	tot, DMUS2		
Observed Value	15.7151	15.7151	15.7151	15.7151	15.7151	15.7151	15.7151	15.7151	15.7151	15.7151	15.7151	15.7151	15.7151	15.7151	15.7151	0.00000	0.00000		
summary statistics	DMUS2, AF, AF, AF, AF, AF, AF, AF, AF, AF, AF, Hromosome	DMUS2, AF, AF, AF, AF, AF, AF, AF, AF, AF, AF, koku	DMUS2, AF, AF, AF, AF, AF, AF, AF, AF, AF, AF, ohromosome	DMUS2, AF, AF, AF, AF, AF, AF, AF, AF, AF, AF, loku	DMUS2, AF, AF, AF, AF, AF, AF, AF, AF, AF, AF, amant	DMUS2, AF, AF, AF, AF, AF, AF, AF, AF, AF, AF, AF, north	DMUS2, AF, AF, AF, AF, AF, AF, AF, AF, AF, AF, AF, south	DMUS2, AF, AF, AF, AF, AF, AF, AF, AF, AF, AF, Rad, Hromosome	DMUS2, AF, AF, AF, AF, AF, AF, AF, AF, AF, AF, Rad, koku	DMUS2, AF, AF, AF, AF, AF, AF, AF, AF, AF, AF, Rad, ohromosome	DMUS2, AF, AF, AF, AF, AF, AF, AF, AF, AF, AF, Rad, loku	DMUS2, AF, AF, AF, AF, AF, AF, AF, AF, AF, AF, Rad, amant	DMUS2, AF, AF, AF, AF, AF, AF, AF, AF, AF, AF, Rad, AF, north	DMUS2, AF, AF, AF, AF, AF, AF, AF, AF, AF, AF, Rad, AF, south	mean, DMUS2	st, DMUS2	tot, DMUS2		
Observed Value	15.7151	15.7151	15.7151	15.7151	15.7151	15.7151	15.7151	15.7151	15.7151	15.7151	15.7151	15.7151	15.7151	15.7151	15.7151	0.00000	0.00000		
summary statistics	DMUS2, AF, AF, AF, AF, AF, AF, AF, AF, AF, AF, AF, Hromosome	DMUS2, AF, AF, AF, AF, AF, AF, AF, AF, AF, AF, AF, koku	DMUS2, AF, AF, AF, AF, AF, AF, AF, AF, AF, AF, AF, ohromosome	DMUS2, AF, AF, AF, AF, AF, AF, AF, AF, AF, AF, AF, loku	DMUS2, AF, AF, AF, AF, AF, AF, AF, AF, AF, AF, AF, amant	DMUS2, AF, AF, AF, AF, AF, AF, AF, AF, AF, AF, AF, AF, north	DMUS2, AF, AF, AF, AF, AF, AF, AF, AF, AF, AF, AF, AF, south	DMUS2, AF, AF, AF, AF, AF, AF, AF, AF, AF, AF, AF, Rad, Hromosome	DMUS2, AF, AF, AF, AF, AF, AF, AF, AF, AF, AF, AF, Rad, koku	DMUS2, AF, AF, AF, AF, AF, AF, AF, AF, AF, AF, AF, Rad, ohromosome	DMUS2, AF, AF, AF, AF, AF, AF, AF, AF, AF, AF, AF, Rad, loku	DMUS2, AF, AF, AF, AF, AF, AF, AF, AF, AF, AF, AF, Rad, amant	DMUS2, AF, AF, AF, AF, AF, AF, AF, AF, AF, AF, AF, Rad, AF, north	DMUS2, AF, AF, AF, AF, AF, AF, AF, AF, AF, AF, AF, Rad, AF, south	mean, DMUS2	st, DMUS2	tot, DMUS2		
Observed Value	15.7151	15.7151	15.7151	15.7151	15.7151	15.7151	15.7151	15.7151	15.7151	15.7151	15.7151	15.7151	15.7151	15.7151	15.7151	0.00000	0.00000		
summary statistics	DMUS2, AF, AF, AF, AF, AF, AF, AF, AF, AF, AF, AF, AF, Hromosome	DMUS2, AF, AF, AF, AF, AF, AF, AF, AF, AF, AF, AF, AF, koku	DMUS2, AF, AF, AF, AF, AF, AF, AF, AF, AF, AF, AF, AF, ohromosome	DMUS2, AF, AF, AF, AF, AF, AF, AF, AF, AF, AF, AF, AF, loku	DMUS2, AF, AF, AF, AF, AF, AF, AF, AF, AF, AF, AF, AF, amant	DMUS2, AF, AF, AF, AF, AF, AF, AF, AF, AF, AF, AF, AF, AF, north	DMUS2, AF, AF, AF, AF, AF, AF, AF, AF, AF, AF, AF, AF, AF, south	DMUS2, AF, AF, AF, AF, AF, AF, AF, AF, AF, AF, AF, AF, AF, Rad, Hromosome	DMUS2, AF, AF, AF, AF, AF, AF, AF, AF, AF, AF, AF, AF, AF, Rad, koku	DMUS2, AF, AF, AF, AF, AF, AF, AF, AF, AF, AF, AF, AF, AF, Rad, ohromosome	DMUS2, AF, AF, AF, AF, AF, AF, AF, AF, AF, AF, AF, AF, AF, Rad, loku	DMUS2, AF, AF, AF, AF, AF, AF, AF, AF, AF, AF, AF, AF, AF, Rad, amant	DMUS2, AF, AF, AF, AF, AF, AF, AF, AF, AF, AF, AF, AF, AF, Rad, AF, north	DMUS2, AF, AF, AF, AF, AF, AF, AF, AF, AF, AF, AF, AF, AF, Rad, AF, south	mean, DMUS2	st, DMUS2	tot, DMUS2		
Observed Value	15.7151	15.7151	15.7151	15.7151	15.7151	15.7151	15.7151	15.7151	15.7151	15.7151	15.7151	15.7151	15.7151	15.7151	15.7151	0.00000	0.00000		
summary statistics	DMUS2, AF, AF, AF, AF, AF, AF, AF, AF, AF, AF, AF, AF, AF, Hromosome	DMUS2, AF, AF, AF, AF, AF, AF, AF, AF, AF, AF, AF, AF, AF, koku	DMUS2, AF, AF, AF, AF, AF, AF, AF, AF, AF, AF, AF, AF, AF, ohromosome	DMUS2, AF, AF, AF, AF, AF, AF, AF, AF, AF, AF, AF, AF, AF, loku	DMUS2, AF, AF, AF, AF, AF, AF, AF, AF, AF, AF, AF, AF, AF, amant	DMUS2, AF, AF, AF, AF, AF, AF, AF, AF, AF, AF, AF, AF, AF, AF, north	DMUS2, AF, AF, AF, AF, AF, AF, AF, AF, AF, AF, AF, AF, AF, AF, south	DMUS2, AF, AF, AF, AF, AF, AF, AF, AF, AF, AF, AF, AF, AF, AF, Rad, Hromosome	DMUS2, AF, AF, AF, AF, AF, AF, AF, AF, AF, AF, AF, AF, AF, AF, Rad, koku	DMUS2, AF, AF, AF, AF, AF, AF, AF, AF, AF, AF, AF, AF, AF, AF, Rad, ohromosome	DMUS2, AF, AF, AF, AF, AF, AF, AF, AF, AF, AF, AF, AF, AF, AF, Rad, loku	DMUS2, AF, AF, AF, AF, AF, AF, AF, AF, AF, AF, AF, AF, AF, AF, Rad, amant	DMUS2, AF, AF, AF, AF, AF, AF, AF, AF, AF, AF, AF, AF, AF, AF, Rad, AF, north	DMUS2, AF, AF, AF, AF, AF, AF, AF, AF, AF, AF, AF, AF, AF, AF, Rad, AF, south	mean, DMUS2	st, DMUS2	tot, DMUS2		
Observed Value	15.7151	15.7151	15.7151	15.7151	15.7151	15.7151	15.7151	15.7151	15.7151	15.7151	15.7151	15.7151	15.7151	15.7151	15.7151	0.00000	0.00000		
summary statistics	DMUS2, AF, AF, AF, AF, AF, AF, AF, AF, AF, AF, AF, AF, AF, AF, Hromosome	DMUS2, AF, AF, AF, AF, AF, AF, AF, AF, AF, AF, AF, AF, AF, AF, koku	DMUS2, AF, AF, AF, AF, AF, AF, AF, AF, AF, AF, AF, AF, AF, AF, ohromosome	DMUS2, AF, AF, AF, AF, AF, AF, AF, AF, AF, AF, AF, AF, AF, AF, loku	DMUS2, AF, AF, AF, AF, AF, AF, AF, AF, AF, AF, AF, AF, AF, AF, amant	DMUS2, AF, AF, AF, AF, AF, AF, AF, AF, AF, AF, AF, AF, AF, AF, AF, north	DMUS2, AF, AF, AF, AF, AF, AF, AF, AF, AF, AF, AF, AF, AF, AF, AF, south	DMUS2, AF, AF, AF, AF, AF, AF, AF, AF, AF, AF, AF, AF, AF, AF, AF, Rad, Hromosome	DMUS2, AF, AF, AF, AF, AF, AF, AF, AF, AF, AF, AF, AF, AF, AF, AF, Rad, koku	DMUS2, AF, AF, AF, AF, AF, AF, AF, AF, AF, AF, AF, AF, AF, AF, AF, Rad, oh									

Table S8 Geographic data of respective islands.

Island Name	Freshwater form distribution	Latitude	Longitude	Area (km ²)	Area log10	Catchment area (km ²)	Catchment area log10	Length (m)	Length log	No. Steep Slopes
Okinawajima Is.	1	26°35'	128°2'	1206.93	3.082	53.30	1.727	17218	4.236	152
Amami-Oshima Is.	1	28°16'	129°20'	712.5	2.853	48.58	1.686	18478	4.267	548
Iriomotejima Is.	1	24°20'	123°50'	289.6	2.462	64.37	1.809	21736	4.337	435
Tokunoshima Is.	1	27°45'	128°57'	247.9	2.394	34.00	1.531	15754	4.197	304
Ishigakijima Is.	1	24°24'	124°11'	222.2	2.347	32.75	1.515	14750	4.169	101
Okinoerabujima Is.	1	27°22'	128°35'	93.7	1.972	10.84	1.035	8813	3.945	15
Kakeromajima Is.	0	28°7'	129°14'	77.3	1.888	2.99	0.475	3601	3.556	26
Kumejima Is.	1	26°22'	126°46'	59.5	1.775	6.99	0.844	5925	3.773	35
Kikaijima Is.	0	28°18'	129°57'	57.0	1.755	4.81	0.683	4957	3.695	5
Iheyajima Is.	0	27°3'	127°58'	20.7	1.315	4.93	0.693	4428	3.646	2
Tokashikijima Is.	0	26°11'	127°21'	15.3	1.185	3.33	0.522	4029	3.605	19
Izenajima Is.	0	26°56'	127°56'	14.1	1.150	3.05	0.484	3918	3.593	3
Ukejima Is.	0	28°1'	129°15'	13.3	1.125	2.38	0.377	2275	3.357	21
Yorojima Is.	0	28°2'	129°9'	9.4	0.971	2.27	0.356	3257	3.513	12
Yagajijima Is.	0	26°40'	128°0'	7.8	0.893	1.16	0.063	2072	3.316	0
Miyagijima Is.	0	26°22'	128°0'	5.5	0.744	0.90	-0.048	1867	3.271	2
Akajima Is.	0	26°12'	127°17'	3.8	0.580	0.46	-0.333	1364	3.135	4
Hamahigajima Is.	0	26°19'	127°57'30"	2.1	0.320	0.36	-0.441	1297	3.113	0

Table S9 The results of statistical test and AIC model selection of logistic regression analysis for respective combinations of islands where speciation observed. 1 indicated speciation occurred, 0 indicated speciation did not occur.

Combinations of islands		Originate in respective regions				Share the origin in respective regions															
		Dataset 1	Dataset 2	Dataset 3	Dataset 4	Dataset 5	Dataset 6	Dataset 7	Dataset 8	Dataset 9	Dataset 10	Dataset 11	Dataset 12	Dataset 13	Dataset 14	Dataset 15	Dataset 16	Dataset 17	Dataset 18	Dataset 19	Dataset 20
Island name																					
Okinawajima Is.		1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	0	0
Amami-Oshima Is.		1	1	1	1	1	1	1	1	1	1	1	1	0	0	0	0	1	1	1	1
Tokunoshima Is.		1	1	0	0	1	1	0	0	0	0	1	0	1	1	0	0	1	1	0	0
Kumejima Is.		1	1	1	1	0	0	0	0	1	1	1	1	1	1	1	1	1	1	1	1
Okinoerabujima Is.		0	0	1	1	0	0	1	1	0	0	0	1	0	0	1	1	0	0	1	1
Iriomotejima Is.		1	0	1	0	1	0	1	0	1	0	0	0	1	0	1	0	1	0	1	0
Ishigakijima Is.		0	1	0	1	0	1	0	1	0	1	0	0	0	1	0	1	0	1	0	1
Kakeromajima Is.		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Kikajima Is.		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Tokashikijima Is.		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Iheyajima Is.		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Izenajima Is.		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Ukejima Is.		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Yorojima Is.		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Yagajijima Is.		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Miyajijima Is.		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Akajima Is.		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Hamahigajima Is.		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Maximum likelihood ratio																					
P values																					
Area		0.0016	0.0026	0.0054	0.0061	0.0001	0.002	0.005	0.0064	0.0068	0.0093	0.0073	0.0253	0.024	0.0317	0.0534	0.0606	0.0342	0.046	0.0733	0.0868
Catchment Area		0.0026	0.0069	0.008	0.0187	0.0003	0.0041	0.0046	0.0113	0.0105	0.025	0.0242	0.0637	0.0191	0.0322	0.0479	0.082	0.0214	0.0346	0.0518	0.0899
River Length		0.0038	0.0046	0.0077	0.0172	0.0006	0.0031	0.0057	0.0115	0.0101	0.0294	0.0239	0.0596	0.0212	0.0307	0.0502	0.0789	0.0191	0.0295	0.0464	0.0729
No. Waterfall		0.0013	0.0697	0.0449	0.3277	2.00E-04	0.0347	0.0203	0.2627	0.0177	0.2163	0.0628	0.3241	0.1269	0.587	0.5181	0.8131	0.0055	0.0879	0.0614	0.4257
Model Selection																					
AIC scores																					
Island Area only		19.478	13.005	13.922	15.965	8.953	4	9.145	12.471	13.366	13.939	14.678	14.382	16.682	16.532	17.063	18.315	18.729	17.553	18.016	19.113
Residuals of Catchment Area only		22.943	25.059	24.389	24.723	11.391	22.913	22.251	22.589	21.623	22.548	21.551	21.362	20.472	23.069	22.803	22.977	22.407	22.835	23.067	23.042
Island Area + Residuals of Catchment Area		21.405	14.835	15.868	17.96	6	6	6	14.454	15.092	15.926	16.417	16.015	17.809	17.646	18.991	20.131	20.701	17.371	19.46	20.316
Island Area + Residuals of Catchment Area		22.78	16.82	16.004	19.9	8	8	8	16.357	15.368	17.847	17.585	16.982	18.61	19.578	20.991	22.084	22.69	17.905	20.579	21.091
Island Area only		19.478	13.005	13.922	15.965	8.953	4	9.145	12.471	13.366	13.939	14.678	14.382	16.682	16.532	17.063	18.315	18.729	17.553	18.016	19.113
Residual of River Length only		23.067	25.018	24.386	24.88	11.646	22.888	22.274	22.76	22.026	22.34	21.326	21.368	21.006	23.058	22.734	23.012	22.567	22.726	23.062	22.873
Island Area + Residual of River Length		21.107	14.918	15.832	17.932	6	6	6	14.351	15.276	15.935	16.272	15.99	18.098	17.902	19.025	20.064	20.727	17.084	19.201	19.669
Island Area + Residual of River Length		22.611	16.908	15.996	19.764	8	8	8	15.937	14.921	17.852	17.32	17.138	18.757	19.893	21.019	22.051	22.615	17.652	20.199	20.856
Island Area only		19.478	13.005	13.922	15.965	8.953	4	9.145	12.471	13.366	13.939	14.678	14.382	16.682	16.532	17.063	18.315	18.729	17.553	18.016	19.113
Residuals of No. Waterfall only		20.894	24.196	20.881	22.692	9.13	22.551	19.758	21.339	16.818	21.989	18.335	20.969	18.718	20.977	15.981	18.731	9.806	23.03	22.305	22.913
Island Area + Residuals of No. of Waterfall		21.288	13.13	14.906	17.963	10.952	6	6	14.442	6	15.538	15.159	16.38	17.679	18.52	15.828	19.161	11.775	14.038	19.719	20.019
Island Area + Residuals of No. of Waterfall		15.168	15.081	16.756	19.49	8	8	8	16.087	8	17.538	16.986	18.377	18.832	20.438	16.985	21.052	13.672	8	18.523	14.982