Online supplementary material for
Recurrent speciation rates on islands decline with species number

Ryo Yamaguchi, Yoh Iwasa and Yuuya Tachiki

Appendix 1

Detailed assumptions and simulation scheme for recurrent species origination

In the Wright-Fisher process, we assume that the offspring of the next generation are reproduced more than the carrying capacity, and few individuals migrate before they reach sexual maturity. Then, individuals that develop sexual maturity in the source population are reduced to a number controlled by the carrying capacity. It means that immigrants mature sexually during migration, and the time required for their dispersal is more than that for population regulation. Thus, the number of immigrants may exceed the number of individuals in the source population and arrive at a different recipient population.

If no migration occurred between the two islands, then the two populations would diverge into different species rapidly, but the number of species would not increase further. However, in the presence of migration events, the number of species would increase to three or more. A combination of the accumulation of novel mutations on different islands and the rare migration between them provides a simple mechanism to produce novel species, without considering any ecological differentiation (Yamaguchi and Iwasa 2015).

Although migration events contribute deterministically to genetic differentiation in our model, we considered that colonisation events occurred stochastically with the migration rate m. The timing of migration events can be neglected in the mean dynamics of genetic differentiation because of its considerably longer time scale compared to colonisation dynamics. Successful migrations from one population to the other follow a Poisson process. We also assumed that colonising migrants could rapidly establish their population and the size could attain equilibrium simultaneously. Thus, population growth immediately after colonisation does not affect the formulation of genetic differentiation. Overall, when a speciation event occurs after calculating the deterministic dynamics of genetic distance, random numbers are drawn to determine the occurrence of a stochastic colonisation.

Appendix 2

The equilibrium of the number of species in a two-island system is reached without species extinction

Here, we considered two islands that were initially occupied by a single species and the accumulation of novel mutations in incompatibility-controlling loci. There exists a threshold fraction of the loci z_c for reproductive isolation. Let S_1 and S_2 be the number of species on each island, respectively. We assumed symmetric interspecific competition among species on each island, and the species on island i had an abundance of $N_i = K/\left(1 + a(S_i - 1)\right)$, based on the assumptions of Lotka–Volterra-type competition equations with a carrying capacity K and competition coefficient a.

Let us consider the increase in the number of species on Island 2, which occurred with the arrival of a species from Island 1 that did not exist on Island 2. This is possible when the asymptotic genetic distance is greater than the speciation threshold: $\hat{z}_2 > z_c$, where $\hat{z}_2$ has a suffix indicating the island, as it may differ between islands. In the absence of species extinction, the increase in the number of species S_2 will be realised as long as $\hat{z}_2 > z_c$, and will cease when $\hat{z}_2 \leq z_c$ holds. $\hat{z}_2$ is not a constant and should be dependent on the size of the resident population N_2, the number of individuals in a single migration event N', and the successful migration rate m. We can safely assume that N' is a constant determined by the social structure of the species. In contrast, N_2 depends on the number of species and the competition strength as $N_2 = K/\left(1 + a(S_2 - 1)\right)$. Here, we have:

$$\hat{z}_2 = \frac{u}{u+m\varepsilon} = \frac{u}{u+\frac{mN'}{N'+N_2}} \tag{A2.1}$$

The condition for S_2 to stop increasing is $\hat{z}_2 \leq z_c$. This can be rewritten as

$$\frac{u}{m}\frac{1-z_c}{z_c}\left(1 + \frac{K/N'}{1+a(S_2-1)}\right) \leq 1 \tag{A2.2}$$

In a similar manner, S_1 stops increasing when $\hat{z}_1 \leq z_c$ holds. This can be rewritten as

$$\frac{u}{m}\frac{1-z_c}{z_c}\left(1 + \frac{K/N'}{1+a(S_1-1)}\right) \leq 1. \tag{A2.3}$$

By setting $S_1 = S_2$, we have the stationary level when

$$S_1 = S_2 \geq \frac{1}{a}\left[\frac{K}{N'}\frac{u(1-z_c)}{mz_c - u(1-z_c)} - (1-a)\right]. \tag{A2.4}$$

This is achieved without extinction under the deterministic model. However, in the stochastic model, speciation occurs in a finite number of generations even when $\hat{z}_i \geq z_c$ (see Yamaguchi and Iwasa 2016).

Appendix 3

The cycle time is always a convex function of the migration rate

The sum of the mean time to speciation and the mean waiting time to the next colonisation event (1/*m*) fits a U-shaped curve, and the sum indicates the 'cycle time'. To obtain the cycle time analytically, we described the waiting to speciation. According to the model section, the ordinary differential equation for genetic distance is indicated by $dz/dt = 2u(1-z) - m\varepsilon z$, where t represents the time measured in terms of generations. Given the initial condition $z(0)$, the solution of the above-mentioned equation is

$$z(t) = \frac{1}{u+m\varepsilon}\left(e^{-2(u+m\varepsilon)t}\left(u\left(z(0) - 1 + e^{2(u+m\varepsilon)t}\right) + m\varepsilon z(0)\right)\right). \quad \text{(A3.1)}$$

The time until speciation τ is

$$\tau = \frac{\log[-u/(u(z_c-1)+m\varepsilon z_c)]}{2(u+m\varepsilon)}, \qquad \text{if } \hat{z} \geq z_c \quad \text{(A3.2)}$$

$$\tau = \infty, \qquad \text{if } \hat{z} < z_c \quad \text{(A3.3)}$$

under the assumption that $z(0) = 0$. Hereafter, we only considered the situation in which the waiting time was finite, as shown in Eq. (A3.2). In other words, when the dynamics is at a threshold for speciation, $dz/dt > 0$ holds; hence, this situation satisfies the following inequality: $0 < m\varepsilon z_c < u(1-z_c)$. As a function of the migration rate *m*, the cycle time *T*(*m*) is

$$T(m) = \frac{1}{m} + \frac{\log[-u/(u(z_c-1)+m\varepsilon z_c)]}{2(u+m\varepsilon)} \quad \text{(A3.4)}$$

If *T*(*m*) can be differentiated twice and the domain is the real line, then we can characterise it as follows: convex if and only if *T*″(*m*)>=0 for all *m*. It is noteworthy that all parameters and variables in this study are non-negative values. The first and second derivatives of *T*(*m*) are:

$$T'(m) = -\frac{1}{m^2} + \frac{\varepsilon z_c}{2(u+m\varepsilon)(u(z_c-1)+m\varepsilon z_c)} + \frac{\varepsilon \log[1-z_c(u+m\varepsilon)/u]}{(u+m\varepsilon)^2}, \quad \text{(A3.5)}$$

$$T''(m) = \frac{2}{m^3} + \frac{\varepsilon^2 z_c(u(3z_c-2)+3m\varepsilon z_c)}{2(u+m\varepsilon)^2(u(z_c-1)+m\varepsilon z_c)^2} + \frac{\varepsilon^2 \log[u/(u(1-z_c)-m\varepsilon z_c)]}{(u+m\varepsilon)^3} \geq 0, \quad \text{(A3.6)}$$

respectively. Thus, by Eq.(A3.6), the cycle time is always a U-shaped function of the migration rate, and the optimal migration rate can be obtained by numerically solving Eq.(A3.5), equalling to 0.

Appendix 4

Extended discussion: character displacement and niche specialisation in archipelagic avifauna

Despite the fact that these two taxa arrived at the Galapagos archipelago almost concurrently, the consequences of their diversification are completely different (15 species for finches vs. 4 allospecies for mockingbirds). A possible reason for the lack of diversification of mockingbirds is their generalist feeding behaviour. Grant and Grant (2008) argued that mockingbirds have long beaks, and their diets are broad and largely overlapped among species. Interestingly, Darwin's finches with proportionately long beaks do not coexist with close relatives. In contrast, tree and ground finch species with blunter beaks that differ in size are fully sympatric. For sustained coexistence without competitive exclusion on each island, generalist feeding behaviour may pose a disadvantage for producing strong interspecies competition. The discussion by Grant and Grant (2008) suggests that the manner in which birds exploit the environment is an important factor in determining the species diversity of the two groups of birds.

A similar situation has been reported in the Hawaiian archipelago (Lovette et al. 2002). Like the mockingbirds on the Galapagos Islands, ecologically similar but unrelated *Myadestes* thrushes on the Hawaiian Islands have failed to show diversification beyond two sympatric species in 4.2 MY, whereas more than 50 species of honeycreeper finches evolved in as much as 6.4 MY (Fleischer and McIntosh 2001). None of the honeycreeper finches can be described as occupying a thrush niche and pre-empting further thrush evolution in the Hawaiian Islands, and this is similar to the observation in which none of Darwin's finches pre-empted mockingbird evolution in the Galapagos Islands. Moreover, continental relatives of mockingbirds and thrushes have also not diversified extensively (Lovette et al. 2002; Arbogast et al. 2006), whereas relatives of Darwin's finches (Burns et al. 2002) and honeycreepers (Lovette et al. 2002) have diversified.

These patterns can be explained by several different hypotheses (in the following paragraphs, the argument holds true for the Hawaiian Islands if we replace Darwin's finches with honeycreepers, and mockingbirds with thrushes). First, the pattern may be attributed to the extinction rate, depending on the number of competing species on the same islands. If each species tends to be specialised in resource use, the competition intensity is mitigated by character displacement, and the abundance of each species tends to be maintained higher. Hence, the extinction rate is lower than when species are not specialised. If Darwin's finches can rapidly engage in niche specialisation

than mockingbirds, the former groups of birds can exhibit an increase in their species number simply because of a lower extinction rate.

The analysis in the current study indicates that there exists an alternative process that also leads to a higher species number generated in the system. If Darwin's finches could be specialised in their resource use, they would be relatively less affected by interspecific competition and each species would be maintained at a higher abundance with the same number of species as mockingbirds. This would promote the rate of species origination because of the large population size of the recipient, which is less susceptible to gene flow, according to our theory. Thus, this mechanism may lead to a faster rate of species origination in finches than in mockingbirds.

Note that there is a third process that can explain the higher rate of species origination in Darwin's finches than in mockingbirds. Niche partitioning can be realised through behavioural changes, or adaptive phenotypic plasticity, rather than genetic evolution. If this is the case, then when the migrants arrive at a new island, migrant populations and resident populations can easily coexist through mutual niche segregation. If Darwin's finches performed this adaptation more rapidly than mockingbirds, the faster success rate of speciation in the former than in the latter is quite plausible.

These hypotheses are not mutually exclusive. It is possible that the hypothesized phenomena have occurred in reality to result in the differences observed between the species diversity of these groups of birds. Across the evolutionary time scale, the coefficient of interspecific competition can be reduced via character displacement and niche specialisation. The ease with which it occurs in evolutionary history affects the species diversity of a focal clade (Schluter 2000).

References

1. Arbogast BS, Drovetski SV, Curry RL, Boag PT, Seutin G, Grant PR, Grant BR, Anderson DJ. 2006 The origin and diversification of Galapagos mockingbirds. *Evolution* **60**, 370–382. (doi:10.1111/j.0014-3820.2006.tb01113.x)
2. Burns KJ, Hackett SJ, Klein NK. 2002 Phylogenetic relationships and morphological diversity in Darwin's finches and their relatives. *Evolution* **56**, 1240–1252. (doi:10.1111/j.0014-3820.2002.tb01435.x)
3. Fleischer RC, McIntosh CE. 2001 Molecular systematics and biogeography of the Hawaiian avifauna. *Stud. Avian Biol.* **22**, 51–60.

4. Lovette IJ, Bermingham E, Ricklefs RE. 2002 Clade-specific morphological diversification and adaptive radiation in Hawaiian songbirds. *Proc. R. Soc. B* **269**, 37–42. (doi:10.1098/rspb.2001.1789)
5. Schluter D. 2000 Ecological character displacement in adaptive radiation. *Am. Nat.* **156**, S4–S16.

Appendix 5

Coexistence of all species and a relevant model in population genetics

In our present model, there was interspecific competition between species that was weaker than the intraspecific competition because $a < 1$. This is equivalent to the population genetics model that had a frequency-dependent selection that favoured rare alleles. Such a system has been explored in the context of the molecular evolution of major histocompatibility complex (MHC) alleles, the high diversity of which is maintained by selection that favours the rare alleles. They exhibit a much longer coalescence time with a slower turnover than neutral alleles (Takahata and Nei 1990), which was adopted as the basis of the coalescence theory for MHC (Klein et al. 1993).

References

6. Klein J, Satta Y, O'hUigin C, Takahata N. 1993 The molecular descent of the major histocompatibility complex. *Annu. Rev. Immunol.* **11**, 269–295.
7. Takahata N, Nei M. 1990 Allelic genealogy under overdominant and frequency-dependent selection and polymorphism of major histocompatibility complex loci. *Genetics* **124**, 967–978.

Appendix 6

When the number of migrants depends on the population size of the source population

In the main text, we have assumed and discussed the modelling of migration with abiotic processes such as storms and oceanic dispersal by rafting. Under this assumption, the number of immigrants is represented by N' constant. Alternatively, we can assume that migration occurs as a biotic process by the habits of organisms. In this case, m indicates the migration rate per individual and, thus, the number of immigrants can be rewritten as $N' = Nm$. Then, the impact of migration ε is represented by $\varepsilon = \frac{N'}{N+N'} = \frac{m}{1+m}$. Now, ε becomes an N-independent parameter through all the processes of recurrent speciation, and it will not contribute to the facilitation of gene flow as the number of species increases (see also Figure S5). Therefore, if migration is assumed to be a biotic process, speciation rates do not decrease due to interspecific competition, as observed in the results of the present study.

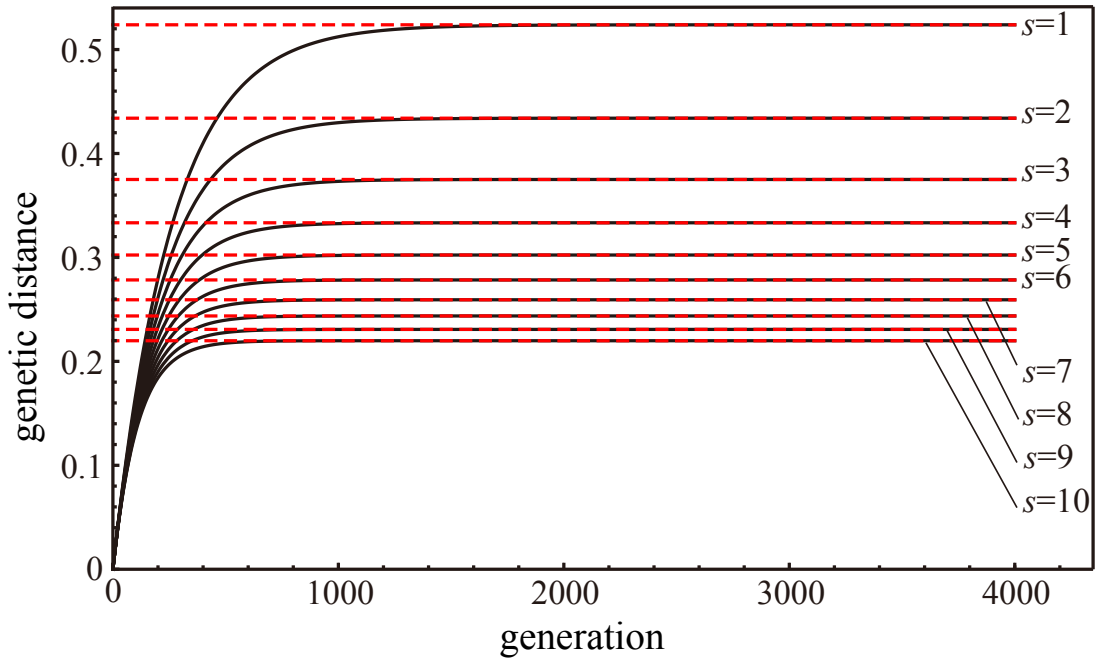

Figure S1. Dynamics in the incompatibility genetic distance between two populations. The horizontal axis indicates the time expressed in terms of the number of generations. The black solid lines represent the dynamics with different species numbers s. The dashed red lines correspond to the equilibrium values of each deterministic dynamics. Parameters include u=0.001, m=0.005, N'=10, and K=100.

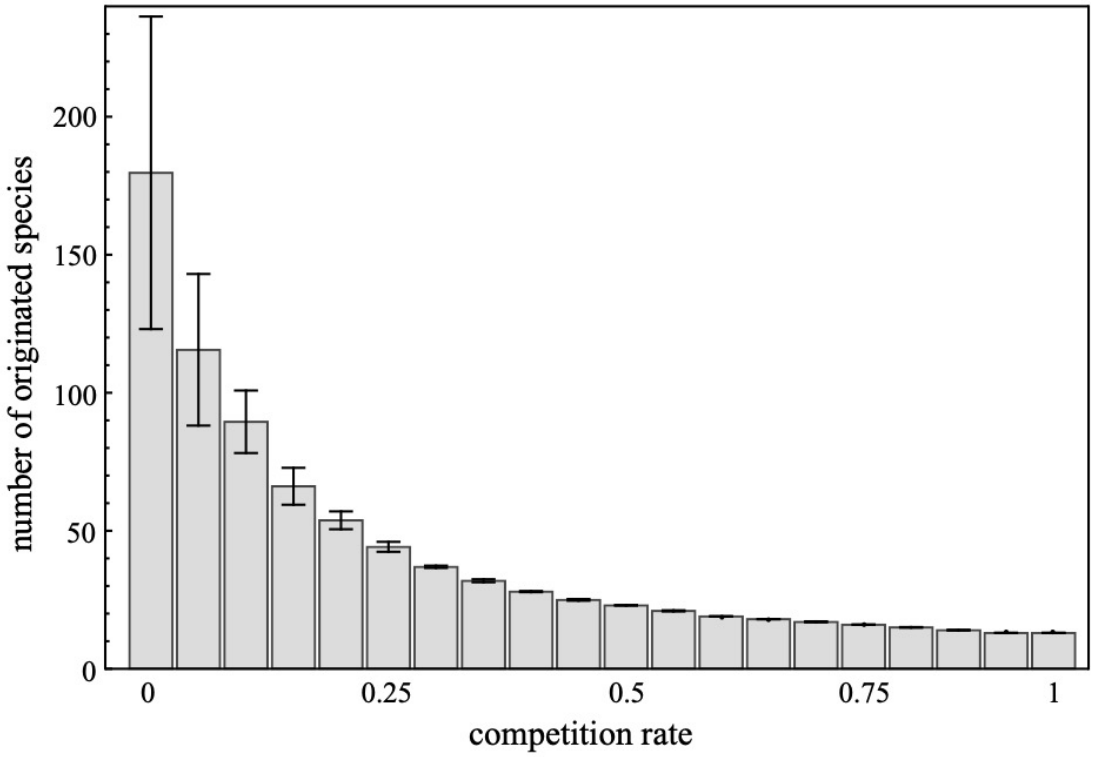

Figure S2. The number of originated species with different competition rates. The bar chart represents mean ± SE. An average of 100 simulations were run for each parameter set. Each simulation was run with 2500 generations. All other parameters are as those illustrated in Figure 2.

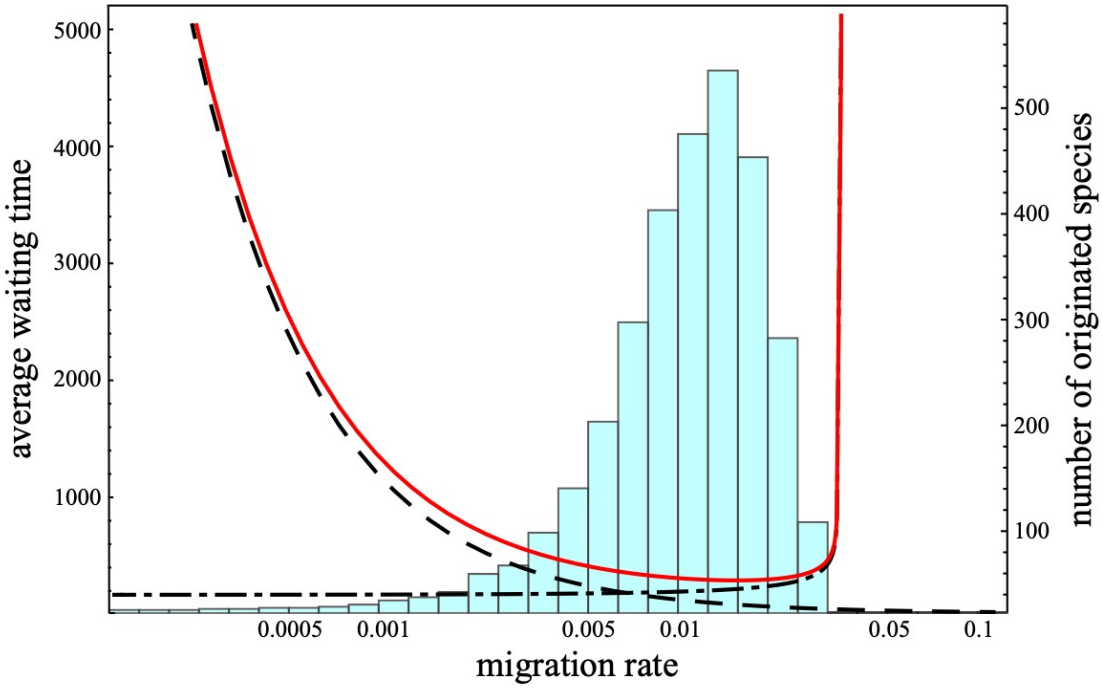

Figure S3. The optimal rate of migration that maximizes the rate of species origination in the case of $a = 0$. The horizontal axis indicates the migration rate in the logarithmic scale. The dashed line indicates the mean time interval between migration events, and the dashed-and-dotted line indicates the average waiting time until speciation. The solid red line represents their sum. The blue bar chart indicates the average number of species originated in the system composed of two islands after 10000 generations. One species existed at time 0. All other parameters are as those illustrated in Figure 2.

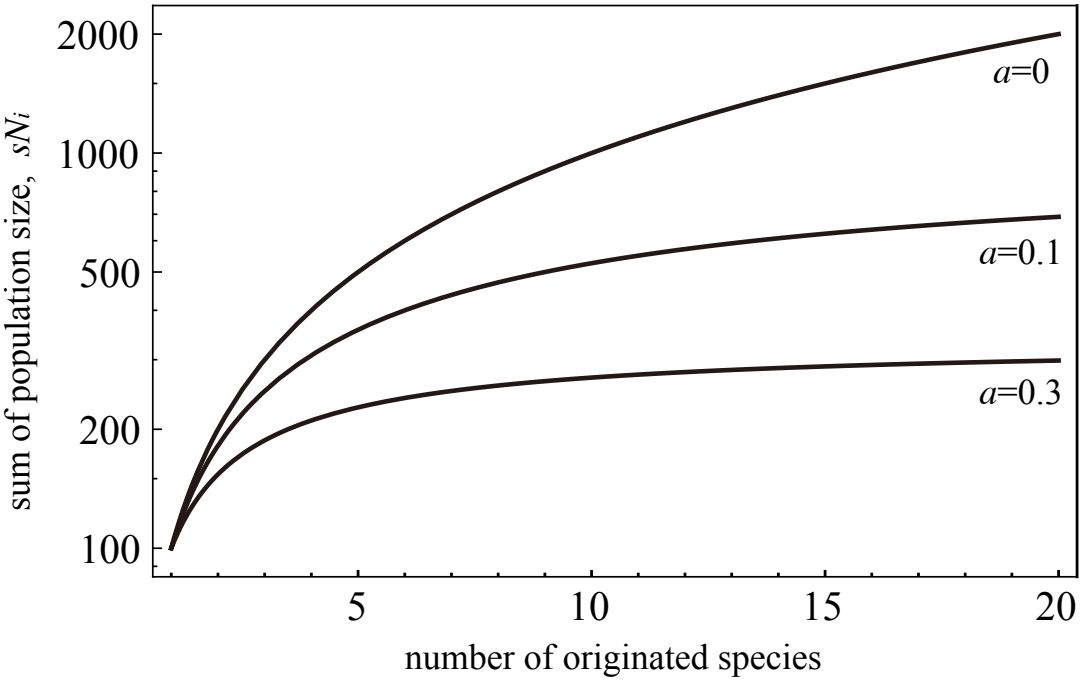

Figure S4. The total size of species s on an island, sN_i. The different curves represent the results with different interspecific competition coefficients a. We have set the initial population size as K=100 when only one species exists.

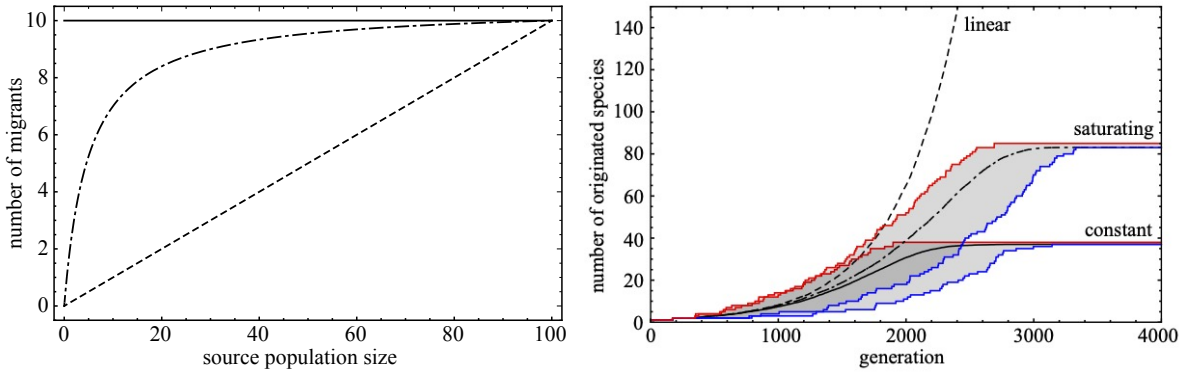

Figure S5.

Robustness in the results of the current model when the number of migrants *N'* is not assumed to be a constant, but is assumed as a saturating function of the population size *N*. For elucidating the saturating relationship, we define the number of migrants as $N' = M\frac{N}{N+c}$, where *M* is the maximum number of migrants and *c* is a half-velocity constant: the value of the source population size when N'/M=0.5. The left panel shows example relationships among the following three different source-migrant functions: constant (indicated by a solid line), linear (indicated by a dashed line), and saturating (indicated by a dashed-dotted line). We have assumed that the number of migrants is the same (10 individuals) among all the functions when there is a single species at the initial state with the source population size 100. Under this assumption, a waiting time to the occurrence of first speciation event is equivalent to each other. Figure on the right represents trajectories of the number of originated species like those illustrated in Fig.2 for each function shape. The trajectory of the case of 'constant' is exactly the same as that illustrated in Fig.2, and speciation rate does not decrease in the case of 'linear' as explained in Appendix 6. The 'saturating' case is mentioned in the middle of the other two. Parameters include *a*=0.3, *M*=10.5, *c*=5, and all other parameters are as those illustrated in Fig.2.