Diversity patterns and speciation processes in a two-island system with continuous migration

Débora Princepe,1,2 Simone Czarnobai,3 Thiago M. Pradella,1 Rodrigo A. Caetano,4 Flavia M. D. Marquitti,1,5 Marcus A. M. de Aguiar,1* and Sabrina B. L. Araujo4

1Instituto de Física ‘Gleb Wataghin’, Universidade Estadual de Campinas, Campinas, Brasil
2E-mail: princepe@ifi.unicamp.br
3Programa de Pós Graduação em Ecologia e Conservação, Universidade Federal do Paraná, Curitiba, Brasil
4Departamento de Física, Universidade Federal do Paraná, Curitiba, Brasil
5Instituto de Biologia, Universidade Estadual de Campinas, Campinas, Brasil

Received March 2, 2022
Accepted June 22, 2022

Geographic isolation is a central mechanism of speciation, but perfect isolation of populations is rare. Although speciation can be hindered if gene flow is large, intermediate levels of migration can enhance speciation by introducing genetic novelty in the semi-isolated populations or founding small communities of migrants. Here, we consider a two-island neutral model of speciation with continuous migration and study diversity patterns as a function of the migration probability, population size, and number of genes involved in reproductive isolation (dubbed as genome size). For small genomes, low levels of migration induce speciation on the islands that otherwise would not occur. Diversity, however, drops sharply to a single species inhabiting both islands as the migration probability increases. For large genomes, sympatric speciation occurs even when the islands are strictly isolated. Then species richness per island increases with the probability of migration, but the total number of species decreases as they become cosmopolitan. For each genome size, there is an optimal migration intensity for each population size that maximizes the number of species. We discuss the observed modes of speciation induced by migration and how they increase species richness in the insular system while promoting asymmetry between the islands and hindering endemism.

KEY WORDS: Continuous migration, island biogeography, neutral model, speciation with gene flow.
Migration between groups is a common behavior that prevents complete isolation of populations (Winker 2000; Nosil 2008; Chaîne & Clobert 2012; Turbek et al. 2018). For instance, estuarine-river environments and tide pools are systems under pulse-driven biotic events (Halas et al. 2005). These environments are cyclically isolated and expanded as a consequence of the water levels (Baggio et al. 2017). Thus, the system behaves like islands subject to periodic exchanges of migrants from neighboring sites. Similarly, species that have seasonal migration behavior are constantly mixing in a common area nesting or breeding place and irradiating after to other sites (Winker & Pruett 2006; Cooper & Uy 2017; Everson et al. 2019; Manthey et al. 2020). In this case, the migration itself (associated with dispersal abilities and spatial structure of the landscape) plays the role of an intermittent geographic barrier (Hey 2006; Claramunt et al. 2012; Agnarsson et al. 2014; Pinheiro et al. 2014; Linck & Battey 2019; Ashby et al. 2020). Such intermediate geographic situations can affect the process of speciation in complex ways since diversification in the presence of migration depends on a balance between colonization, local selection, and the homogenizing effects of gene flow (Garant et al. 2007). On the one hand, high levels of migration can constrain diversification via maintenance of gene flow (Mayr 1963; von Rönne et al. 2016) and hinder speciation (Fitpatrick et al. 2009; Claramunt et al. 2012). It has been suggested that even one or a few migrants should be sufficient to avoid genetic diversification altogether (Slatkin 1987). On the other hand, migration between partially isolated populations can foster speciation by increasing genetic variation (Cowie & Holland 2006; Smadja & Butlin 2011) and by promoting events (Spurgin et al. 2014), which is the establishment of a small number of migrants in a new area, favoring rapid genetic changes and eventual diversification into a new species (Barton & Charlesworth 1984; Gavrilets & Hastings 1996; Templeton 1980; 2008). Understanding how migration increases diversity in semi-isolated populations has become increasingly important in a world of fragmented environments (Mills & Allendorf 1996; Hagen et al. 2012).

Investigation of the role of migration on speciation has mainly focused on the time to speciation and the influence of population structure on the rate of species creation, with evolution essentially driven by mutation and genetic drift (Manzo & Peliti 1994; Yamaguchi & Iwasa 2013). Notably, migration increases the time for speciation (Gavrilets 2000), but species divergence can be favored by subdivision of the population, even in the absence of local adaptation (Gavrilets et al. 1998; 2000). When considering isolated populations under rare but recurrent migration, optimal intermediate migration intensities maximize the rate of species creation in two-islands systems (Yamaguchi & Iwasa 2013; 2016; 2017b). Still, some patterns are yet to be addressed and discussed in more detail, such as the distribution of species richness, endemism, and species persistence. In this aspect, models can help elucidate and differentiate processes from patterns. Furthermore, current advances in obtaining genome-wide datasets provide robust estimates of gene flow and demographic history (Hey 2010; Gyllenhaal et al. 2020), evincing that speciation with gene flow in insular systems is common and may rise even in the absence of ecological divergence (Gyllenhaal et al. 2020). Therefore, understanding how migration and diversification are related can provide a better interpretation of speciation events in these intermediary geographic configurations that are not fully sympatric or allopatric.

In the present study, we focus on how migration intensity affects speciation and species richness in insular communities. We consider a neutral model for two islands subject to frequent exchange of individuals, where mating occurs locally and is only restricted by genetic similarity. The model is based on the work of Manzo and Peliti (1994), who considered a similar dynamic process of local reproduction and migration between two islands in the limit of infinitely large genomes, and investigated the possibility of allopatric speciation in the presence of gene flow. Here, we show that population size and genetic features, particularly the number of loci involved in reproductive isolation, are key to determine not only the total number of species in the system but also the fraction of species that are endemic or cosmopolitan. Our simulations show that migration promotes the appearance of new species through two distinct processes: founding populations and induced sympatric speciation. In the first case, the small but continuous flow of migrants creates a sub-population that differentiates from the residents. In the second case, the migrants mix with the residents, creating a genetically diverse population that eventually splits into species.

**Methods**

Our model for insular populations is based on the theory proposed by Manzo and Peliti (1994), which is itself based on the sympatric speciation model by Higgs and Derrida (1991). In these models, each individual is characterized by a string of infinite biallelic genes that represents the loci involved in reproductive isolation, dubbed as the individual’s genome. Using infinite genes allows the derivation of several analytic results by mean-field calculations. However, important features of the infinite genes model disappear if the genome size is small, including the possibility of sympatric speciation (de Aguiar 2017). Here we review the model of speciation in islands as proposed by Manzo and Peliti and extend it to finite genomes (de Aguiar et al. 2009; Costa et al. 2018). We shall see that genome length plays a key role in the diversification process and that the limit of infinite genes is a very particular case.
Figure 1. Schematic of the model dynamics. We consider two insular populations of initially identical individuals, represented by dots. At each time step, the islands exchange individuals with a probability $\epsilon$. Generations are non-overlapping and sexual reproduction occurs between individuals in the same island having a minimal genetic similarity. Species are identified at the end of each cycle based on the possibility of gene flow, illustrated by links and enclosed in ellipses. The ellipses A, B, C, and D (full lines) correspond to species classification within each island. Species can be exclusive to an island (B and D), or can be common to both (A and C, dashed lines).

MODEL OVERVIEW

We consider two identical spatial regions (hereafter called islands) occupied by populations that can migrate and evolve through mutations and recombination. Each individual $\alpha$ is represented by a binary string of independent $B$ genes, $\{S^\alpha_1, S^\alpha_2, \ldots, S^\alpha_B\}$, where each locus $S^\alpha_i$ can assume the alleles $\pm 1$. Individuals are hermaphroditic and reproduction is sexual. Population size is kept constant at the carrying capacity $M$ in each island, except during the migration period when small fluctuations are allowed. The dynamics starts with $2M$ clonal individuals ($M$ on each island) and goes through cyclical events of migration, reproduction, and species identification (see Fig. 1), as detailed below:

Migration: at each generation, each individual has a probability $\epsilon$ of migrating to the other island. Thus the population size in each island can vary right after migration, but the total remains $2M$ individuals. We refer to this process as continuous migration, as opposed to rare migrations of larger groups (Yamaguchi & Iwasa 2013).

Reproduction: reproduction occurs after migration only between individuals on the same island. On each island, $M$ offspring are born and replace the previous population, compensating for fluctuations in populations size that may have occurred after migration. Generations do not overlap, and offspring are generated as follows: a first parent $P_1$ (focal parent) is chosen at random. A second parent $P_2$ is selected, also at random, among the remaining individuals. Reproduction between these two individuals will only occur if they are genetically compatible, i.e., if they have a minimal genetic similarity, as defined below. If the selected $P_2$ is not compatible with $P_1$, another second parent is randomly selected until this condition is met. If after $M$ trials no such individual is found, $P_1$ is discarded and a new first parent is selected. The offspring inherits, gene by gene, the allele of either parent with equal probability. The offspring’s genome is also subjected to a mutation rate $\mu$ per loci, which can be reversed during the dynamics.

Reproduction is restricted by the genetic compatibility of the mating pair. Two individuals $\alpha$ and $\beta$ can reproduce only if their genetic similarity, defined by

$$q^{\alpha\beta} = \frac{1}{B} \sum_{i=1}^{B} S^\alpha_i S^\beta_i,$$

satisfies the condition $q^{\alpha\beta} \geq q_{\text{min}}$, where the minimum similarity $q_{\text{min}}$ is a model parameter. If the genomes are identical then $q^{\alpha\beta} = 1$, while two genomes with random entries have $q^{\alpha\beta}$ close to zero. The model is neutral in the sense that there is no fitness assigned to the individuals, and the choice of mating partner is random, limited only by a minimum genetic similarity, i.e., a weak form of assortative mating.

Species identification: speciation occurs when the gene flow between groups of individuals is disrupted. The population maps to a network where the individuals (nodes) are connected if they are genetically compatible. Species correspond to the components of the network. Not all individuals of a species need to be compatible since the genetic flow can be established through intermediary individuals. In network terms, there is a path connecting them within the component, as illustrated in Figure 1. We identify species considering local and global classifications. In the first case, only individuals from the same island are
considered (components circled by continuous lines in Fig. 1), while in the second, the genetic compatibility network is built regardless of which island individuals belong (components circled by dashed lines in Fig. 1).

**ANALYTICAL RESULTS**

In the limit where genome size becomes infinite, the entire dynamics can be obtained by simply updating the similarity matrix. A number of simplifications occur and analytical results can be obtained (Higgs & Derrida, 1991), as we summarize below.

**One island model**

If there is no restriction on mating, $q_{\text{min}} = 0$, the overlaps $q^{\text{off}}$ are all very similar. The distribution of overlaps is a peaked function whose average value satisfies the evolution equation (Higgs & Derrida, 1991) (see section S1 A of the Supporting Material for a derivation (Princepe et al., 2022))

$$q_{t+1} = e^{-4\mu} \left( (1 - \frac{1}{M}) q_t + \frac{1}{M} \right) \equiv e^{-4\mu} Q(q_t, M)$$

(2)

where $t$ is measured in generations. Setting $q_{t+1} = q_t = \bar{q}$, we find the equilibrium $\bar{q} \equiv q_0 \approx (1 + 4\mu M)^{-1}$. If $q_{\text{min}} > q_0$, the distribution of similarities cannot reach the equilibrium and the population splits into species, formed by groups of individuals whose average similarity is larger than $q_{\text{min}}$. Inter-species similarities are smaller than $q_{\text{min}}$, tending to zero with time (Higgs & Derrida, 1991). Average species abundance $m$ can be estimated by setting $q_0 \rightarrow q_{\text{min}}$ and $M \rightarrow m$, as this gives the population size that would naturally equilibrate at $q_{\text{min}}$. We obtain $m = (q_{\text{min}}^{-1} - 1)/(4\mu)$ and the average number of species is $N_{\text{(one)}} = M/m = 4M\mu/(q_{\text{min}}^{-1} - 1)$. We find numerically that the distribution of abundances has an average slightly smaller than $m$, which implies average number of species slightly larger than $N_{\text{(one)}}$.

We note that this dynamic behavior can change drastically for finite genomes. If the number of loci $B$ is small enough, the distribution of similarities becomes wider and, instead of splitting into separate peaks as it crosses $q_{\text{min}}$, it equilibrates as a single peak around $q_{\text{min}}$, preventing speciation. Only for $B$ above a critical value, which depends on population size, mutation rate and $q_{\text{min}}$, the above description (and therefore sympatric speciation) is recovered (de Aguiar, 2017).

**Two-islands model**

In the case of two islands, we need to distinguish the similarities between individuals inhabiting the same island, $q$, or different islands, $p$. If individuals $\alpha$ and $\beta$ were born on the same island, then their average $q$ similarity in the next generation, before migration, is given by Equation 2. If, on the other hand, they were born in different islands, then their $p$ similarity before migration is $p_{t+1} = e^{-4\mu} p_t$, as they cannot have a common parent (Manzo and Peliti, 1994) (see Equation 3 of Supporting Information (Princepe et al., 2022)). The dynamics of $q$ and $p$ after migration is, therefore, given by the following equations:

$$q_{t+1} = a(\epsilon) e^{-4\mu} Q(q_t, M) + b(\epsilon) e^{-4\mu} p_t$$

$$p_{t+1} = b(\epsilon) e^{-4\mu} Q(q_t, M) + a(\epsilon) e^{-4\mu} p_t$$

(3)

where $a(\epsilon) = (1 - \epsilon^2 + \epsilon^2)$ is the probability that both $\alpha$ and $\beta$ did not migrate or that they both migrated, therefore keeping their original geographic relation, and $b(\epsilon) = 2\epsilon(1 - \epsilon)$ is the probability that one of them exchanged places, altering the geographic relative position of the pair. The equilibrium solutions are obtained setting $q_{t+1} = q_t$ and $p_{t+1} = p_t$. For $\epsilon, \mu$ and $1/M$ all much smaller than 1, we obtain (Manzo and Peliti, 1994)

$$q_0(v, \sigma) = \frac{v + 2\sigma}{2\sigma(2v + 1) + v(v + 1)}$$

$$p_0(v, \sigma) = \frac{2\sigma}{2\sigma(2v + 1) + v(v + 1)}$$

(4)

where $v = 4\mu M$ and $\sigma = M\epsilon$ is the average number of exchanged migrants at each generation. Notice that for $\sigma \gg v$, we have $p_0(v, \sigma) = q_0(2v, 0)$, indicating that for a large migration intensity the two islands behave as a single island with twice the population.

Two new important results are derived here: first, the expected number of species in each island, $N$, can be estimated by replacing $q_0(v, \mu)$ by $q_{\text{min}}$ in Eq. 4, as similarly done before, and solving for $M \rightarrow m$. The number of species is $N = M/m$:

$$N = \frac{v(4\sigma + v)}{(2\sigma + v)(q_{\text{min}}^{-1} - 1)}.$$  

(5)

Second, the total number of species in the islands, $N_T$, can also be estimated as

$$N_T = N \left( 2 - \frac{p_0(v, \sigma)}{q_0(2v, 0)} \right).$$

(6)

When $\sigma = 0$ the islands are independent and $N_T = 2N$. When migration dominates over mutations, $\sigma \gg \mu$, $N_T = N$.

**NUMERICAL SIMULATIONS AND DATA ANALYSIS**

We performed simulations with both finite and infinite genomes to evaluate how migration probability ($\epsilon$), population size per island ($M$), and amount of loci involved in reproductive isolation ($B$) affects patterns of diversity in a two island system. The mutation rate per locus and the genetic threshold for reproduction were fixed for all simulations, $\mu = 0.001$ and $q_{\text{min}} = 0.9$, respectively. The parameters of interest were varied with the following values: $B = 1000, 2000, 3000, 10000$, and the infinite model; $M$ from 25 to 500 individuals; and $\epsilon$ from 0 to 0.5. Populations were evolved during 2000 generations, with observation of species richness at
When probing the species richness, we considered both local and global classifications (see Species identification in 2.1). In the local classification, the species richness per island were designated by \( N_1 \) and \( N_2 \). Under the global classification, the species richness in each island was \( N_1 \) and \( N_2 \), and the total number of species was \( N_T \). Local and global classifications differed when individuals in one island, say island 1, could establish gene flow with two species of island 2 that would otherwise be reproductively isolated, a configuration that resembles the ring species assemblage (Martins et al. 2013). In this case these two species that contributed to \( N_{1,2} \) would count as one for \( N_2 \). We quantified these events, named here “ring-like species”, by calculating \( N_{\text{ring}} = (N_{1,1} - N_1) + (N_{1,2} - N_2) \) (the local classification was exclusively employed for this analysis).

We also evaluated the asymmetry \( \Delta N \) between the islands as the proportion of unbalance in species richness (under the global classification) in respect to the average,

\[
\Delta N = \frac{\sqrt{(N_1 - N_2)^2}}{N},
\]

where \( \bar{N} = (N_1 + N_2)/2 \). The diversity between the islands was measured based on the Jaccard distance index, whereas we calculate the number of exclusive species given by \( (N_T - N_1) + (N_T - N_2) \). We call here the beta-diversity index \( \beta_I \) the normalization of this quantity, defined as

\[
\beta_I = \frac{2N_T - N_1 - N_2}{N_T}.
\]

If species are exclusive to each island, \( N_T = N_1 + N_2 \) and \( \beta_I = 1 \), indicating maximum endemism. In the other hand, if species are all common to both islands, we have \( N_T = N_1 = N_2 \) and \( \beta_I = 0 \), that is, all species are cosmopolitan.

Results

We first explored how migration probability, genome size, and population size affected diversity at the end of the simulations \((T = 2000)\) for genome lengths \( B = 1000, 2000, \) and \( 3000 \) with varying migration probability, \( 0 \leq \epsilon \leq 0.08 \). When the islands were completely isolated from each other, \( \epsilon = 0 \), speciation did not occur within the islands for the smallest genome size, \( B = 1000 \) (Fig. 2, upper left). In this case, the populations differentiated only through the accumulation of mutations in each island (allopatry), and each island had a unique exclusive species (indicated by \( \beta_I = 1 \) in Fig. 2, bottom left). This inter-island divergence was limited by the population size \((M \lesssim 350)\), as larger \( M \) required longer times for speciation (see Fig. S2).

For \( B = 2000 \), sympatric speciation in isolated islands was rare, occurring scarcely only for small values of \( M \) (Fig. 2, upper right). Increasing \( B \) also led to an increase in the maximum population size that supported \( \beta_I = 1 \), the maximum endemism (Fig. 2, middle). Strict allopatry \((\epsilon = 0)\) represented the peak of endemism for all scenarios (see Fig. 2, middle).

Migration even at low levels was sufficient to increase the number of species on each island. However, the beta-diversity decreased with migration, as expected, with a slower decay the larger the genome size. Therefore, although species in each island were more numerous, they were more likely to be shared. Regardless of the genome size, a curve depending on population size and migration probability maximized the per island and total species richness, with peak values increasing with \( B \) (Fig. 2, top; for the values of the total species richness, \( N_T \), see Fig. S2, S3, and S4). The increase of the average number of species per island, \( \bar{N} \), was followed by the formation of ring-like species (Fig. 2, middle), that is, groups of individuals with disrupted gene flow in the same island, but with the possibility of gene flow through individuals from the other island. The increase in \( N_{\text{ring}} \) was simultaneous to the loss of endemism, indicating the genetic homogenization due to more intense migration. Those patterns were consistent over time, with a direct dependence on the time for allopatry with \( M \) (see Fig. S2, S3, and S4).

We also simulated larger genomes \((B = 10000 \) and infinite loci\) to compare with the predictions from the analytical solutions. The results were analyzed in more detail by fixing the population size at \( M = 200 \) and varying the migration probability in the range \( 0 \leq \epsilon \leq 0.5 \). The theoretical prediction for the species richness in each island \( N \) (Equation 5), based on a model of infinite loci, was slightly lower than the average value from the simulations. It occurs because the prediction considers a uniform distribution of species abundances (constant \( m \)), and the simulated populations resulted, in fact, in heterogeneous species abundances. For a better comparison, we adjusted the result from Equation 5 for isolated islands \((N = 7.2 \) for \( \epsilon = \sigma = 0 \)) to the simulations by adding a correction \((\delta = 1.8 \) ), and shifted the whole curve by the same constant (black line in Fig. 3a). This theoretical expectation of the number of species in one island showed a monotonic increase with \( \epsilon \) and was also a reasonable estimate of the number of species for \( B = 10000 \). For large genomes \((B \geq 10000)\), the population split into several species even in strict isolation (sympatric speciation), and migration increased species richness further. However, intense migration, ranging from \( \epsilon = 0.1 \) to 0.5, led to the homogenization of the populations that evolved effectively as contiguous \( 2M \) individuals. For small genomes, \( B = 1000 \) and 2000, migration was essential to increment species richness, as seen before.
However, here we observe that this effect was limited to low levels of migration: above a critical point, the number of species within the island collapsed to a single one shared by the two islands ($\epsilon \geq 0.014$ for $B = 1000$ and $\epsilon \geq 0.03$ for $B = 2000$). $B = 3000$ had an intermediary behavior, presenting several species for $\epsilon = 0$, but also collapsed for high levels of migration ($\epsilon \geq 0.2$) with a continuous transition. The increment in total species richness ($N_T$) beyond the number of total species in strict allopatry ($\epsilon = 0$) showed that migration induced some form of speciation in the insular community, and the effect was more intense for small genome lengths (Fig. 3b). The approximation for $N_T$ from Equation 6 was a reasonable estimate for $B \geq 10000$, predicting the decay in $N_T$ with $\epsilon$ but in a steeper way. The qualitative patterns of species richness with respect to the migration probability.
Species richness and asymmetry between the islands at $T = 2000$ for varying migration probability ($\epsilon$) and genome sizes $B$ (colors) with fixed population size ($M = 200$). Results are shown for 50 independent simulations. The solid lines represent the average value of all realizations, and the shadowed areas show a confidence interval of 90%. The full black lines represent the expected values from Eqs. 5 and 6. (a) Average number of species per island, $\bar{N} = (N_1 + N_2)/2$. (b) Total number of species, $N_T$. (c) Species asymmetry $\Delta N$, calculated with Equation 7. The higher asymmetry occurs for values of $B$ for which migration is essential for speciation.

When increasing $B$, keeping all other parameters fixed, are reproduced in a similar way with decreasing $q_{\text{mix}}$ (Fig. S7a) and increasing $\mu$ (Fig. S7b) with all other parameters fixed – see Section S2 D of the Supplementary Material for a discussion of the trade-off between the model parameters not explored here (Princepe et al. 2022).

Although the migration probability was symmetric between islands, the number of species in each one was asymmetric, especially for low and intermediary migration intensities, $0 \leq \epsilon \leq 0.03$ (Fig. 3c). This asymmetry occurred most of the time but was ephemeral, as we observed an alternation of the island with more species and fluctuation of the species richness on both islands (videos included in the Supporting Information (Princepe et al. 2022)). Comparing to a null model of random distribution of the two species between the two islands with equal probability (see the Supporting Information, section S2E (Princepe et al. 2022)), we verified that some of the asymmetry results from inherent stochasticity of the system, but the highest asymmetry, beyond the expected by chance, occurred for values of $B$ and $\epsilon$ for which migration was the essential mechanism for the speciation. We inferred that asymmetry occurred when speciation induced by migration enhanced random imbalances between the number of species on each island. For instance, when one island was composed of a single species, migrants were likely to establish a new species on the arrival island as they were genetically compatible. On the other hand, as the other island then had more than one species, its migrants had a reduced chance of being conspecific and were more likely to go extinct on the island of arrival for lack of compatible mates, keeping the first island with a single species.

Analysis of the beta-diversity index in the insular community shows that, as expected, increasing the migration probability led to a decline in $\beta_I$ (Fig. 4a), from the complete endemism when islands were isolated ($\epsilon = 0$ and $\beta_I = 1$) to the mixed populations, having most of the species in common ($\epsilon \approx 0.5$ and $\beta_I \approx 0$). Large genomes were more resilient, with slow decay, while the small genomes presented a sharp drop, following the collapse in the number of per island and total species to a single one. Although the proportion of exclusive species in the system always decreased, the absolute number of exclusive species increased with migration for small genomes (Fig. S6). The increase in the migration probability also caused the appearance of ring-like species (Fig. 4b). Figure 4c illustrates how it occurred during the simulations. In this example, using the local referential, the island at the left had one species and the other two species, $N_{I_1} = 1$ and $N_{I_2} = 2$. However, the two species in the second island were, in fact, a single one in the global classification, with gene flux between the two groups reconnected by individuals of island one, then $N_1 = N_2 = 1$. This assemblage is similar to ring species (Martins et al. 2013) on a much smaller scale and longing only through a few generations, although appearing throughout the dynamics and even at $T = 2000$, as shown in Fig. 4b. Such a configuration either preceded speciation (when the disrupted gene flow later led to the formation of species) or disappeared with the subsequently re-establishment of gene flow by the arrival of migrants or due to mutation. Therefore, we classified them as “ring-like species”. For small genomes, they occurred in a small number, restricted to the range of migration intensity that made $N_i > 1$. For large genomes, they were recurrent and appeared in significant numbers under intense migration.

Finally, we investigated the processes by which migration leads to speciation. Using genetic compatibility networks, we observed two speciation modes induced by migration, illustrated in Figure 5. Speciation by founding populations occurred when migrants arriving on an island could not reproduce with the resident population for being of a different species and, when they were able to accumulate through several generations and establish a
Figure 4. Beta-diversity index of the insular community and ring-like species at $T = 2000$ for varying migration probability ($\epsilon$) and genome sizes $B$ (colors), with fixed population size ($M = 200$). (a) The beta-diversity index $\beta$ decreases monotonically with migration from full allopatry ($\epsilon = 0$ and $\beta = 1$) to the mixed populations sharing most of the species ($\epsilon \approx 0.5$ and $\beta \approx 0$). (b) As migration increases, gene flow disrupted within one island can be re-established by individuals from the other island, in a configuration similar to ring species. (c) Example of ring-like species, for $B = 2000$ and $\epsilon = 0.02$ at $T = 430$. Locally there is one species at the left island and two species at the right one, but, at the global classification, individuals of the first island reconnect the two clusters from the second, and then the islands share a single species.

Figure 5. Evolution of the populations for $B = 1000$ and $\epsilon = 0.004$ (left) and $B = 2000$ and $\epsilon = 0.02$ (right) exemplifying the speciation modes induced by migration. Bars on the bottom indicate the distribution of species abundances in each island over time. Individuals of the same species have the same color and are connected by links or dashed lines when they belong to different islands. In the speciation through founder populations (left panel), a portion of migrants differentiates into a new species without mixing with the native population. In the second mode (right panel), migrants can incorporate into the resident population, then differentiation between islands is hindered, but migration promotes sympatric speciation.
population, ended up founding a new species (Fig. 5, left panel). We hypothesize that this was the dominant mode for low migration probabilities based on the examples observed with videos. The second mode, recognized as a “sympatric speciation induced by migration” (shown in the right panel), seems to occur more frequently under slightly larger migration probabilities. Here, the islands presented common species, then migrants could reproduce with the resident population, and their incorporated genetic novelties promoted speciation. The described effects can be observed in detail in the videos included with the Supporting Information (Princepe et al. 2022).

Discussion

The present model was based on the work of Manzo and Peliti (1994), who considered only genomes with infinitely many loci. They showed that when mating was not restricted by a minimum genetic similarity, that is, \( q_{\text{min}} = 0 \), the average values of intra and inter-island similarities (\( q \) and \( p \)) converged to the constant values \( q_0 \) and \( p_0 \) given by Equation 4. They proposed that setting \( p_0 < q_{\text{min}} < q_0 \) would drive the distribution of intra-island similarities to \( q_0 \), avoiding sympatric speciation and at the same time creating inter-island reproductive isolation. For the parameter range studied in Manzo and Peliti (1994), this scenario was only achieved by setting \( q_{\text{min}} \approx 0.2 \). For finite and sufficiently small genome sizes as explored here, on the other hand, sympatric speciation can be avoided even if \( q_{\text{min}} > q_0 \), as the distribution of intra-island similarities becomes stationary around \( q_{\text{min}} \) (de Aguiar 2017). The dynamics with finite genome sizes also brought up a richness of processes and outcomes that do not occur in the extreme case of infinite genomes. The time for allopatry, for example, does not depend on population size for infinite loci (Eq. 18 SM), but strongly depends on \( M \) for finite genomes (Fig. S2, S3, and S4). Furthermore, we extended the Manzo and Peliti model, obtaining expressions for the number of species and estimating the percentage of exclusive species on each island. Importantly, we studied the mechanisms of speciation induced by migration for infinite and finite genomes, showing how migration increases species richness while hindering endemism and promoting asymmetry between the islands and transient ring species.

Strict allopatry with one species per island was only maintained for small genomes (\( B = 1000 \) and \( 2000 \)) in the absence of migration with \( M = 200 \). Even a few migrants were sufficient to promote speciation. For instance, with \( \epsilon = 0.004 \) (equivalent to an average number of \( \sigma = 0.8 \) migrants per generation), speciation by founding populations already occurred (see Fig. 5 left panel and videos in the Supplementary Material (Princepe et al. 2022)). This mode was mainly related to low migration intensity and benefited from the slow dynamics of those genomes sizes (Fig. S5). When the two islands still had the same species, the migrants were genetically similar and, in a small number, did not bring sufficient novelties to cause a break in gene flow. After the islands differentiated from each other (around \( T > 400 \) generations), recently arriving migrants in each island could not reproduce with the resident population and either accumulated through several generations or went extinct. When a population was established in the arrival island, they differentiated into a new species (Fig. 5, left panel). No sympatry was observed in those cases, in the sense that the island’s resident species did not branch. Therefore, the ancestry of the new species was always fully connected to migrants, that is, they did not share common ancestors with the former (native) species.

With a slightly higher level of migration, the islands had more than one species and shared some of them. Although there was a loss of endemism, sympatry induced by migration was favored, migrants could reproduce with residents, and their mixing led to speciation. Remarkably, despite the higher proportion of shared species, the absolute number of exclusive species increased with migration for small genomes (Fig. S6). This connection between high species richness with low endemism or low species richness and high endemism in insular systems has been observed in Darwin finches, for instance (Hamilton & Rubinoff 1963), and explained by the spatial structure of the islands (Gascuel et al. 2016). Low levels of migration can be understood as a great physical distance between the islands, while a moderate level indicates islands closer to each other, establishing an analogy between our results and previous findings. Also, an intermediary level of migration seemed to optimize the evolution of diversity, which has been suggested in different contexts (Garant et al. 2007; Yamaguchi & Iwasa 2013), for instance, in the intermediate dispersal model (Agnarsson et al. 2014; Ashby et al. 2020). Based on the theory of island biogeography and associating migration with dispersal abilities, it proposes that colonization promotes speciation at the same time that may increase the probability of extinction; therefore, moderate levels of migration optimize species richness in islands (Agnarsson et al. 2014; Ashby et al. 2020). Our model reproduces this effect while also providing a mechanistic view of the processes by which migration promotes speciation.

Small genome sizes were positively correlated to the asymmetry in species richness between the islands beyond the expected by chance (i.e., if the given total number of species was randomly distributed between the islands with equal probability). The effect was more significant for low levels of migration (Fig. 3c, also Fig. S9 and S10), and we hypothesized it was related to speciation by founding populations that, as mentioned before, enhanced random imbalances of species richness. When sympatric speciation took place, increasing the number of species, asymmetry was less observed. Asymmetry in the geographic range of recently branched sister species is expected.
under peripatric speciation and is more likely to occur in small populations (Barraclough & Vogler 2000). Here asymmetry in species richness resulted from gene flow in a symmetric setup (same number of individuals and migration probability in both islands) with a small number of species and not from landscape heterogeneity or different species ranges. Further investigation might address how such imbalances affect species abundances and the structure of the phylogenies.

Diversity under high migration could only be sustained with large genomes, which can be explained by the trade-off between population size and genome length (see also Fig. S8). In this model, when considering a single island, sympatric speciation does not occur if the genome is too small due to the low genetic variability and the slow accumulation of mutations (de Aguiar 2017). The required variability is readily provided by low and moderate migration. However, under higher exchange of migrants, the populations evolve as contiguous, and then the time to speciation increases for finite genomes (see (de Aguiar 2017) and Fig. S2, S3, and S4), hindering speciation. Large genomes, however, can sustain higher variability and present fast dynamics. Nevertheless, the robustness of large genomes made their response to migration less compelling.

Along the whole dynamics, we observed the formation of ring-like species, but they were more important for large genomes under high migration intensity, indicating how gene flow was actively connecting the populations. Previous models for islands with migration predicted the occurrence of ring species but in larger chains of patches (Gavrilets et al. 1998); here we find a similar effect with only two islands. However, they do not resemble the observed ring species (Martins et al. 2013), for having a much smaller extension and low durability, but can be later used as an indicative or step of the speciation process.

Finally, we note that the results presented here are conditioned to the regime of continuous migration and relatively large mutation probability (µM ≈ 0.1). Rare migrations of larger groups can lead to different outcomes if the time between migrations is larger than M and µ ≪ 1/M, allowing fixation of migrants alleles between migrations (Yamaguchi & Iwasa 2016; Yamaguchi et al. 2021). Hence, diversity in island systems depends on a large number of factors that include periodicity of migration, number of individuals, mutation probability and number of loci and alleles involved in reproductive isolation. Therefore, comparing models with data requires careful analysis of the situation at hand.

**AUTHOR CONTRIBUTIONS**

D.P., T.M.P., R.A.C., M.A.M.A., and S.B.L.A. designed the study; D.P. and S.B.L.A. ran the simulations; D.P., R.A.C., F.M.D.M., M.A.M.A., and S.B.L.A. analyzed the results; D.P., S.C., T.M.P., R.A.C., F.M.D.M., M.A.M.A., and S.B.L.A. wrote the original draft; D.P., F.M.D.M., M.A.M.A., and S.B.L.A. reviewed and edited the manuscript.

**ACKNOWLEDGMENTS**

This work was partly supported by the São Paulo Research Foundation (FAPESP), grants #2018/11187-8 (DP), #2019/24449-3 (DP), #2019/20271-5 (MAMA) and #2016/01343-7 (ICTP-SAIFR). FMDM was supported by Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brazil (Finance Code 001). MAMA and SBLA were supported by Conselho Nacional de Pesquisas Científicas (CNPq), grants #301082/2019-7 and #311284/2021-3 respectively. SBLA acknowledge the computational support from Professor Carlos M. de Carvalho at LFTC-DFis-UFPR.

**CONFLICT OF INTEREST**

The authors declare that they have no conflict of interests.

**DATA ARCHIVING**

Supplementary Information to this study is available in https://doi.org/10.5061/dryad.8931zcr8.

**REFERENCES**

Agnarsson, I., Cheng, R.C. & Kuntner, M. (2014) A multi-clade test supports the intermediate dispersal model of biogeography. *PLOS ONE*, 9(1), e86780.

Ashby, B., Shaw, A.K. & Kokko, H. (2020) An inordinate fondness for species with intermediate dispersal abilities. *Oikos*, 129(3), 311–319.

Baggio, R.A., Stoive, S.B., Spach, H.L. & Boeger, W.A. (2017) Opportunity and taxon pulse: the central influence of coastal geomorphology on genetic diversification and endemism of strict estuarine species. *Journal of Biogeography*, 44(7), 1626–1639.

Barraclough, T.G. & Vogler, A.P. (2000) Detecting the geographical pattern of speciation from species-level phylogenies. *The American Naturalist*, 155(4), 419–434.

Barton, N.H. & Charlesworth, B. (1984) Genetic revolutions, founder effects, and speciation. *Annual Review of Ecology and Systematics*, 15(1), 133–164.

Bolnick, D.I. (2004) Waiting for sympatric speciation. *Evolution*, 58(4), 895–899.

Bolnick, D.I. & Fitzpatrick, B.M. (2007) Sympatric speciation: models and empirical evidence. *Annual Review of Ecology, Evolution, and Systematics*, 38, 459–487.

Caetano, R.A., Sánchez, S., Costa, C.L. & de Aguiar, M.A.M. (2020) Sympatric speciation based on pure assortative mating. *Journal of Physics A: Mathematical and Theoretical*, 53(15), 155601.

Chaine, A.S. & Clobert, J. (2012) Dispersal. In Behavioural responses to a changing world. Oxford University Press.

Claramunt, S., Derryberry, E.P., Remsen Jr., J.V., Brumfield, R.T., Remsen, J.V. & Brumfield, R.T. (2012) High dispersal ability inhibits speciation in a continental radiation of passerine birds. *Proceedings of the Royal Society B: Biological Sciences*, 279(1733), 1567–1574.

Cooper, E.A. & Uy, J.A.C. (2017) Genomic evidence for convergent evolution of a key trait underlying divergence in island birds. *Molecular Ecology*, 26(14), 3760–3774.

Costa, C.L.N., Lemos-Costa, P., Marquitti, F.M.D., Fernandes, L.D., Ramos, M.F., Schneider, D.M., Martins, A.B. & de Aguiar, M.A.M. (2018) Signatures of microevolutionary processes in phylogenetic patterns. *Systematic Biology*, 68(1), 131–144.

Cowie, R.H. & Holland, B.S. (2006) Dispersal is fundamental to biogeography and the evolution of biodiversity on oceanic islands. *Journal of Biogeography*, 33(2), 193–198.

Coyne, J.A. & Orr, H.A. (2004) Speciation (1 ed.). Sinauer Associates, Inc.
de Aguiar, M.A.M. (2017) Speciation in the derrida-higgs model with finite genomes and spatial populations. *Journal of Physics A: Mathematical and Theoretical*, 50(8), 085602.

de Aguiar, M.A.M., Baranger, M., Baptestini, E.M., Kaufman, L. & Bar-Yam, Y. (2009) Global patterns of speciation and diversity. *Nature*, 460(7253), 384–387.

Dieckmann, U. & Doebeli, M. (1999) On the origin of species by sympatric speciation. *Nature*, 400(6742), 354–357.

Everson, K.M., McLaughlin, J.F., Cato, I.A., Evans, M.M., Gastaldi, A.R., Mills, K.K., Shink, K.G., Wilbur, S.M. & Winker, K. (2019) Speciation, gene flow, and seasonal migration in Catharus thrushes (Aves:Turdidae). *Molecular Phylogenetics and Evolution*, 139, 106564.

Fitzpatrick, B., Fordyce, J. & Gavrilets, S. (2009) Pattern, process and geographic modes of speciation. *Journal of Evolutionary Biology*, 22(11), 2342–2347.

Garant, D., Forde, S.E. & Hendry, A.P. (2007) The multifarious effects of dispersal and gene flow on contemporary adaptation. *Functional Ecology*, 21(3), 434–443.

Gascuel, F., Laroche, F., Bonnet-Lebrun, A.S. & Rodrigues, A.S. (2016) Gene flow and rapid differentiation characterize a rapid insular radiation in the southwest Pacific (Aves: Zosterops). *Evolution*, 70(11), 2657–2666.

Gavrilets, S. (2000) Waiting time to parapatric speciation. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, 267(1461), 2483–2492.

Gavrilets, S. (2003) Perspective: models of speciation: what have we learned in 40 years? *Evolution*, 57(10), 2197–2215.

Gavrilets, S. (2006) The Maynard Smith model of sympatric speciation. *Journal of Theoretical Biology*, 239(2), 172–182.

Gavrilets, S. & Hastings, A. (1996) Founder effect speciation: a theoretical reassessment. *The American Naturalist*, 147(3), 466–491.

Gavrilets, S., Li, H. & Vose, M.D. (1998) Rapid parapatric speciation on highly adaptive landscapes. *Proceedings of the Royal Society B: Biological Sciences*, 265(1405), 1483–1489.

Gavrilets, S., Li, H. & Vose, M.D. (2000) Patterns of parapatric speciation. *Evolution*, 54(4), 1126–1134.

Gyllenhaal, E.F., Mapel, X.M., Naikatini, A., Moyle, R.G. & Andersen, M.I. (2020) A test of island biogeographic theory applied to estimates of gene flow in a Fijian bird is largely consistent with neutral expectations. *Molecular Ecology*, 29(21), 7058–7070.

Hagen, M., Kissling, W.D., Rasmussen, C., de Aguiar, M.A., Brown, L.E., Carstensen, D.W., Alves-Dos-Santos, I., Dupont, Y.L., Edwards, F.K., Genini, J., et al. (2012) Biodiversity, species interactions and ecological networks in a fragmented world. *Advances in Ecological Research*, 46, 89–210.

Halas, D., Zamparo, D. & Brooks, D.R. (2005) A historical biogeographic protocol for studying biotic diversification by taxon pulses. *Journal of Biogeography*, 32(2), 249–260.

Hamilton, T.H. & Rubinson, I. (1963) Isolation, endemism, and multiplication of species in the Darwin finches. *Evolution*, 17(4), 388.

Hey, J. (2006) Recent advances in assessing gene flow between diverging populations and species. *Current Opinion in Genetics & Development*, 16(6), 592–596.

Hey, J. (2010) Isolation with migration models for more than two populations. *Molecular Biology and Evolution*, 27(4), 905–920.

Higgs, P.G. & Derrida, B. (1991) Stochastic models for species formation in evolving populations. *Journal of Physics A: Mathematical and General*, 24(17), L985–L991.

Linck, E. & Battey, C.J. (2019) On the relative ease of speciation with periodic gene flow. *bioRxiv*, 758664.

Manthey, J.D., Oliveros, C.H., Andersen, M.J., Filardi, C.E. & Moyle, R.G. (2020) Gene flow and rapid differentiation characterize a rapid insular radiation in the southwest Pacific (Aves: Zosterops). *Evolution*, 74(8), 1788–1803.

Manzo, F. & Peliti, L. (1994) Geographic speciation in the Derrida-Higgs model of species formation. *Journal of Physics A: Mathematical and General*, 27(21), 7079.

Martins, A.B., de Aguiar, M.A.M. & Bar-Yam, Y. (2013) Evolution and stability of ring species. *Proceedings of the National Academy of Sciences*, 110(13), 5080–5084.

Mayr, E. (1963) Animal Species and Evolution. Belknap of Harvard University Press.

Mills, L.S. & Allendorf, F.W. (1996) The one-migrant-per-generation rule in conservation and management. *Conservation Biology*, 10(6), 1509–1518.

Nosil, P. (2008) Speciation with gene flow could be common. *Molecular Ecology*, 17(9), 2103–2106.

Pinheiro, F., Cozzolino, S., Draper, D., de Barros, F., Félix, L.P., Fay, M.F. & Palma-Silva, C. (2014) Rock outcrop orchids reveal the genetic connectivity and diversity of inselbergs of northeastern Brazil. *BMC Evolutionary Biology*, 14(1), 1–16.

Princke, D., Czarzarni, S., Pradella, T., Caetano, R.A., Marquitti, F.M.D., de Aguiar, M.A.M. & Araujo, S.B. (2022) Dryad Dataset – Diversity patterns and speciation processes in a two-island system with continuous migration, https://doi.org/10.5061/dryad.8931zor8.

Slatkin, M. (1987) Gene flow and the geographic structure of natural populations. *Science*, 236(4803), 787–792.

Smadja, C.M. & Butlin, R.K. (2011) A framework for comparing processes of speciation in the presence of gene flow. *Molecular Ecology*, 20(24), 5123–5140.

Smith, J.M. (1966) Sympatric speciation. *The American Naturalist*, 100(916), 637–650.

Spurgin, L.G., Illera, J.C., Jorgensen, T.H., Dawson, D.A. & Richardson, D.S. (2014) Genetic and phenotypic divergence in an island bird: Isolation by distance, by colonization or by adaptation? *Molecular Ecology*, 23(5), 1028–1039.

Templet, A.R. (1980) The theory of speciation via the founder principle. *Genetics*, 94(4), 1011–1038.

Templet, A.R. (2008) The reality and importance of founder speciation in evolution. *BioEssays*, 30(5), 470–479.

Turbek, S.P., Scordato, E.S. & Safran, R.J. (2018) The role of seasonal migration in population divergence and reproductive isolation. *Trends in Ecology & Evolution*, 33(3), 164–175.

von Rönn, J.A., Shafer, A.B. & Wolf, J.B. (2016) Disruptive selection without dispersal and gene flow on contemporary adaptation. *Journal of Evolutionary Biology*, 29(11), 2197–2215.

Winker, K. (2000) Migration and speciation. *Nature*, 404(6773), 36–36.

Winker, K. & Pruett, C.L. (2006) Seasonal migration, speciation, and morphological convergence in the genus catharus (Turdidae). *The Auk*, 123(4), 1052–1068.

Yamaguchi, R. & Iwasa, Y. (2013) First passage time to allopatric speciation. *Interface Focus*, 3(6), 20130026.

Yamaguchi, R. & Iwasa, Y. (2016) Smallness of the number of incompatibility loci can facilitate parapatric speciation. *Journal of Theoretical Biology*, 405, 36–45.

Yamaguchi, R. & Iwasa, Y. (2017a) Parapatric speciation in three islands: dynamics of geographical configuration of allele sharing. *Royal Society Open Science*, 4(2), 160819.

Yamaguchi, R. & Iwasa, Y. (2017b) A tipping point in parapatric speciation. *Journal of Theoretical Biology*, 421, 81–92.
Supplementing Information

Additional supporting information may be found online in the Supporting Information section at the end of the article.

Figure S1. Number of simulations (color scale) performed to compose Figure 2 of the main text.
Figure S2. Diversity patterns of the two-island system as a function of population size (M) and migration probability (ε) for B = 1000 at T = 500 (left column), 1000 (middle) and 2000 (right column) generations.
Figure S3. Diversity patterns of the two-island system as a function of population size (M) and migration probability (ε) for B = 2000 at T = 500 (left column), 1000 (middle) and 2000 (right column) generations.
Figure S4. Diversity patterns of the two-island system as a function of population size (M) and migration probability (ε) for B = 3000 at T = 500 (left column), 1000 (middle) and 2000 (right column) generations.
Figure S5. Distribution of species per island over time for different migration probabilities (indicated in each panel) and fixed population size M = 200, for B = 1000 (left column), 2000 (middle) and 3000 (left column).
Figure S6. Number of exclusive species in the system, K1 + K2.
Figure S7. Effect of the minimal genetic similarity qmin (a) and mutation rate μ (b) on the total species richness.
Figure S8. Selected sets of B and M from Figure 2 of the main text showing the behavior of NT with the number of migrants.
Figure S9. Asymmetry in species richness in the insular system (∆N) for finite genomes.
Figure S10. Histograms of the difference between the islands species richness, N1 - N2, obtained from the simulations (blue) and the null model (red) for 1000 replications when the migration probability is ε = 0.01.