Topology and habitat assortativity drive neutral and adaptive diversification in spatial graphs

Victor Boussange[a,b,1] and Loïc Pellissier[a,b,2]

[a]Swiss Federal Research Institute WSL, CH-8903 Birmensdorf, Switzerland

[b]Landscape Ecology, Institute of Terrestrial Ecosystems, Department of Environmental System Science, ETH Zürich, CH-8092 Zürich, Switzerland

[1]Email: bvictor@ethz.ch

[2]Email: loic.pellissier@usys.ethz.ch

July 7, 2021

Abstract

Biodiversity results from differentiation mechanisms developing within biological populations. Such mechanisms are influenced by the properties of the landscape over which individuals interact, disperse and evolve. Notably, landscape connectivity and habitat heterogeneity constrain the movement and survival of individuals, thereby promoting differentiation through drift and local adaptation. Nevertheless, the complexity of landscape features can blur our understanding of how they drive differentiation. Here, we formulate a stochastic, eco-evolutionary model where individuals are structured over a graph that captures complex connectivity patterns and accounts for habitat heterogeneity. Individuals possess neutral and adaptive traits, whose divergence results in differentiation at the population level. The modelling framework enables an analytical underpinning of emerging macroscopic properties, which we complement with numerical simulations to investigate how the graph topology and the spatial habitat distribution affect differentiation. We show that in the absence of selection, graphs with high characteristic length and high heterogeneity in degree

promote neutral differentiation. Habitat assortativity, a metric that captures habitat spatial autocorrelation in graphs, additionally drives differentiation patterns under habitat-dependent selection. While assortativity systematically amplifies adaptive differentiation, it can foster or depress neutral differentiation depending on the migration regime. By formalising the eco-evolutionary and spatial dynamics of biological populations in complex landscapes, our study establishes the link between landscape features and the emergence of diversification, contributing to a fundamental understanding of the origin of biodiversity gradients.

biodiversity | coevolution | diversification | eco-evolutionary dynamics | habitat networks | heterogeneous environments | measure-valued individual-based model | mathematical modelling | neutral diversity | spatial graphs

Author contributions: V.B. and L.P. designed research; V.B. performed research; V.B. and L.P. wrote the paper. The authors declare no competing interest.

Contents

Significance statement

It is not clear how landscape connectivity and habitat heterogeneity influence differentiation in biological populations. To obtain a mechanistic understanding of underlying processes, we construct an individual-based model that accounts for eco-evolutionary and spatial dynamics over graphs. Individuals possess both neutral and adaptive traits, whose co-evolution results in differentiation at the population level. In agreement with empirical studies, we show that characteristic length, heterogeneity in degree and habitat assortativity drive differentiation. By using analytical tools that permit a macroscopic description of the dynamics, we further link differentiation patterns to the mechanisms that generate them. Our study provides support for a mechanistic understanding of how landscape features affect diversification.

1 Introduction

Biodiversity results from differentiation processes influenced by the features of the landscape over which populations are distributed [1, 2]. The documentation of high levels of species diversity in mountain region or riverine systems suggests that complex connectivity patterns and habitat heterogeneity foster diversity [3, 4, 5, 6, 7]. Correlative studies support this hypothesis, establishing concertedly that connectivity and habitat heterogeneity systematically emerge as core predictors of species richness across a wide range of regions and taxonomic groups [8, 9, 10, 11, 12, 13]. However, hypotheses formulated based on empirical evidence should be complemented by mechanistic models to crystallise a causal understanding between processes and patterns [14]. While the number of simulation studies is growing steadily [15], such studies lack a mathematical formalism to facilitate the interpretation of the model outcomes by providing an analytical underpinning to the simulation results [16].

Differentiation processes emerge as a result of mutation, selection and migration and can be classified as neutral or adaptive [17, 18, 19]. Neutral differentiation is initiated by the stochastic drift of local phenotypes when spatial isolation and limited dispersal create barriers to gene flow, allowing distinct phenotypes to emerge in spatially structured populations [18, 20, 21]. In contrast, adaptive differentiation results from selection pressure [22, 23], which promotes distinct, locally adapted phenotypes in populations occupying patches with different habitat conditions [24, 25]. The evolution of neutral phenotypes and of adaptive phenotypes are not independent, as selective forces can indirectly select for those neutral phenotypes that happen to be linked to the fittest adaptive phenotype, a mechanism called "genetic hitchhiking" [26]. Moreover, selection pressure can generate barriers to gene flow between populations in heterogeneous habitat landscapes [27, 28, 29], a phenomenon coined "isolation by environment", which can amplify neutral differentiation. Overall, landscape features constrain populations' dispersal, affecting neutral and adaptive processes and their interplay. This leads to complex feedback loops that are difficult to comprehend without a formalised mechanistic model [2, 30].

Models link patterns to processes [14], and the explicit representation of the landscape within an evolutionary model can lead to a causal understanding of how landscape features shape diversity [31]. Graphs provide a convenient mathematical representation of landscapes, where vertices represent demes inhabited by populations, and edges capture the connectivity between demes [32, 33, 34, 35]. Under ecological dynamics, metapopulation models have been used to study the role of graph topology on the persistence of population [36, 37, 38, 39, 40, 41] and community diversity [42, 43, 44]. Evolutionary mechanisms are nevertheless fundamental drivers of diversity, and should therefore be explicitly integrated within models [45]. Evolutionary game theory explores how graph topology impacts the fixation probability and the fixation time of a mutated phenotype [46, 47, 48, 49, 50, 51, 52].

However, the framework does not consider the continuous accumulation of mutations, and is therefore not suited to address the emergence of diversity. By combining a metapopulation model with a model of neutral evolution, [31, 53] investigated how graph topology affects neutral diversity. This phenomenological approach demonstrated the key role of topological properties in shaping diversity, and its predictions could be matched with empirical data from e.g. river basins [54]. Nevertheless, diversity results from the combination of neutral and adaptive processes developing at the population level [55, 56, 57]. A first principles modelling approach considering graphs but also building upon the elementary processes of ecological interactions, reproduction, mutation and migration may therefore be promising to investigate emergent patterns.

Stochastic models for structured populations, rooted in the microscopic description of individuals, offer a generic framework for modelling eco-evolutionary dynamics [58, 59]. In particular, such models can capture the interplay between population dynamics, spatial dynamics and phenotypic evolution, while providing a rigorous set-up for analytical investigation. By anchoring this modelling paradigm in a mathematical framework, the work of Champagnat et al. [58] generalises models of population genetics [60, 61, 62, 63] and quantitative genetics [64, 65, 66, 67], which have early on stimulated research into the link between spatial population structure and neutral differentiation. The framework embraces density-dependent selection, which could explain the emergence of sympatric speciation from competition processes [22], and how spatial segregation can emerge as a byproduct of these adaptive processes along environmental gradients [68]. Related models have addressed the effects of landscape dynamics and habitat heterogeneity [69, 70, 71] on diversification, with mathematical insights into the dynamics [72, 73, 74]. Because it accounts for finite population size, the baseline model can also capture neutral differentiation dynamics and therefore the coupling between neutral and adaptive processes [75, 76]. Nonetheless, the aforementioned studies considered geographical space as implicit [69, 70] or assumed regular spatial structures (regular lattices [72, 74] or continuous space [68, 73, 71]), therefore not addressing the topological complexity of landscapes. A stochastic individual-based model using graphs as a representation of the underlying landscape could help formalise the fundamental link between differentiation processes and landscape features.

A key challenge in comprehending biodiversity patterns is to understand how individual dynamics result in differentiation at the landscape level [55, 56, 57]. Here, we investigate how complex connectivity patterns and habitat heterogeneity affect both neutral and adaptive differentiation by constructing an individual-based model (IBM) that accounts for eco-evolutionary and spatial dynamics. In this modelling approach, individuals are spatially structured over a graph and possess co-evolving neutral and adaptive traits. The finite size of local populations generates neutral differentiation by inducing a stochastic drift in the neutral trait evolution,

while heterogeneous selection pressure gives rise to adaptive differentiation. Macroscopic properties of the model are analytically tractable, and we obtain a deterministic approximation of population and adaptive trait dynamics which connects the emerging patterns to the graph properties that generate them. However, neutral differentiation is stochastic by nature, which complicates its analytical underpinning. We therefore rely on numerical simulations of the IBM to measure the effect of graph topology on neutral differentiation. In the case where selection is absent (setting (1)), we investigate which graph topology metrics [77] affect neutral diversity. Second, in the case of heterogeneous selection pressure (setting (2)), we investigate the extent to which such topological properties, in combination with the spatial habitat distribution, affect levels of (i) adaptive and (ii) neutral differentiation. We expect to identify metrics that capture the effects of graph topology and habitat spatial distribution on differentiation, with the intention of relating them to the underlying processes. Overall, our study establishes the link between landscape features and population differentiation and contributes to a fundamental understanding of how landscape topology shapes biodiversity patterns.

2 Results

2.1 Graph eco-evolutionary model

We establish an individual-based model (IBM) where individuals are structured over a trait space and a graph representing a landscape. Individuals die, reproduce, mutate and migrate in a stochastic fashion, which together results in macroscopic properties. Our formalisation results in an analytical description of the dynamics at the population level, which links emergent properties to the elementary stochastic processes that generate them.

2.1.1 Microscopic description

The trait space $\mathcal{X} \subseteq \mathbb{R}^d$ is continuous and can be split into a neutral trait space $\mathcal{U}$ and an adaptive trait space $\mathcal{S}$. We refer to neutral traits $u \in \mathcal{U}$ as traits that are not under selection, in contrast to adaptive traits $s \in \mathcal{S}$, which experience selection. The graph denoted by G is composed of a set of vertices $\{V_1, V_2, \ldots, V_M\}$ that correspond to demes, and a set of edges that constrain the movement of individuals across the demes. By defining $\bar{u}^{(i)}$ as the mean trait on V_i, we quantify neutral differentiation as the variance of $\bar{u}^{(i)}$ across demes and adaptive differentiation as the variance of $\bar{s}^{(i)}$ across the demes, denoted by β_u and β_s respectively, so that $\beta_u = \sum_i \sum_j \frac{1}{2M} \left[\bar{u}^{(i)} - \bar{u}^{(j)}\right]^2$. This notation is motivated by the fact that one can regard the level of differentiation as the population β-diversity ([78] and Supplementary Methods).

Following the Gillespie update rule [79], an individual in a deme represented by V_i with trait $x_k \in \mathcal{X}$ is randomly selected to give birth at rate $b_i(x_k)$ and die at rate $d(N^{(i)}) = {}^{N^{(i)}}/_{K}$, where $N^{(i)}$ is the number of individuals on V_i and K is the local carrying capacity. The definition of d therefore assumes that competition is proportional to the number of individuals in a deme, and does not depend on the individuals' traits. Individuals reproduce asexually and the offspring resulting from a birth event inherits the parental traits, which can independently be affected by mutations with probability μ. A mutated trait differs from the parental trait by a random change that follows a normal distribution with variance σ_μ^2 (corresponding to the continuum of alleles model [62]). The offspring can further migrate to neighbouring demes by executing a simple random walk on G with probability m. Under no selection pressure (setting (1)), individuals are only characterised by neutral traits so that $\mathcal{X} = \mathcal{U}$. For an individual in a deme with trait $x_k \equiv u_k$ we define $b_i(x_k) \equiv b$, so that the birth rate is constant. This ensures that neutral traits do not provide any selective advantage. Under heterogeneous selection pressure (setting (2)), the vertices of the graph are further labelled by a set of habitats $\boldsymbol{\theta} = \{\theta_1, \theta_2, \ldots, \theta_M\}$ that specify the optimal adaptive trait value θ_i on V_i. It follows that, for an individual with traits $x_k = (u_k, s_k) \in \mathcal{U} \times \mathcal{S}$, we define

$$b_i(x_k) \equiv b_i(s_k) = b(1 - p(s_k - \theta_i)^2), \tag{2.1}$$

where p is the selection strength [74]. This ensures that on V_i the maximum birth rate on V_i is $s_k = \theta_i$, which results in a differential advantage that acts as an evolutionary stabilising force. For the sake of the study, we assume that habitats are binary and symmetric, so that $\theta_i \in \{\theta_{\bullet}, \theta_{\bullet}\}$, $\theta_{\bullet} = -\theta_{\bullet} = \theta$, where θ can be viewed as the habitat heterogeneity [74].

2.1.2 From stochastic dynamics to emergent deterministic properties

The model can be formulated as a measure-valued point process ([59] and Supplementary Methods). Under this formalism, we demonstrate in Supplementary Methods how the population size and trait dynamics show a deterministic behaviour when a stabilising force dominates. As a byproduct, the dynamics of the macroscopic properties can be readily expressed with deterministic differential equations, connecting emergent patterns to the processes that generate them. In particular, under the setting of no selection pressure competition stabilises the population size, and the dynamics can be considered as deterministic and expressed as

$$\partial_t N_t^{(i)} = N_t^{(i)} \left[b(1-m) - \frac{N_t^{(i)}}{K} \right] + mb \sum_{j \neq i} \frac{a_{i,j}}{d_j} N_t^{(j)}, \tag{2.2}$$

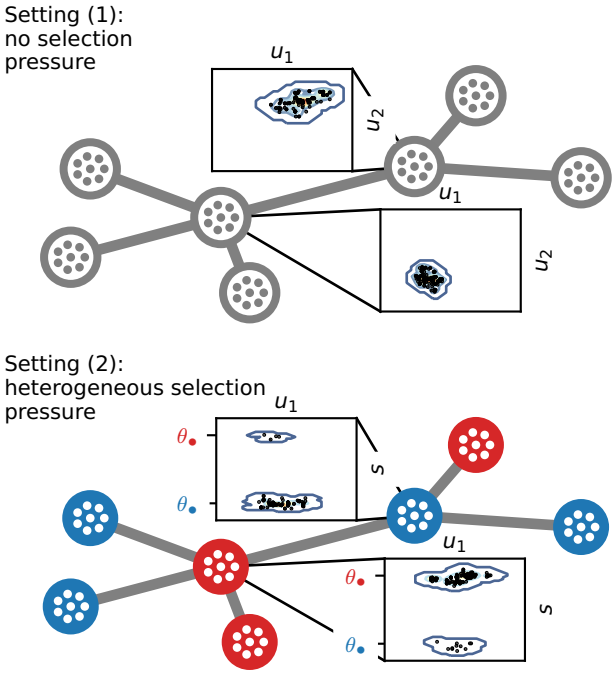

Figure 2.1: Graphical representation of the population structure. The upper panel corresponds to setting (1) with no selection pressure. In this case, individuals are characterised by a set of neutral traits $u \in \mathcal{U}$. The scatter plots represent a projection of the first two components of u for local populations at time $t = 500$, obtained from a simulation.
The lower panel corresponds to setting (2) with heterogeneous selection pressure. In this setting, individuals are additionally characterised by adaptive traits $s \in \mathcal{S}$. Blue vertices favour the adaptive trait optimum $\theta_\bullet$, while red vertices favour $\theta_\bullet$. The scatter plots represent a projection of the first component of u and s for local populations at time $t = 500$, obtained from a simulation. The majority of individuals are locally adapted and have an adaptive trait centred around the trait optimum, but some less adapted individuals originating from neighbouring vertices are also present. $m = 0.1$.

where $A = (a_{i,j})_{1 \leq i,j \leq M}$ is the adjacency matrix of the graph G and $D = (d_1, d_2, \ldots, d_M)$ is a vector containing the degree of each vertex. The first term on the right-hand side corresponds to a logistic growth, which accounts for birth and death events of non-migrating individuals. The second term captures the gains due to migrations, which depend on the graph topology. Assuming that all vertices with the same degree have an equivalent position on the graph, consisting in a "mean field" approach, one can obtain a closed-form solution from Eq. (2.2), which shows that local population size $N^{(i)}$ scales with $\langle\sqrt{k}\rangle^2/\langle k\rangle$, where $\langle k\rangle$ is the average vertex degree and $\langle\sqrt{k}\rangle$ is the average square-rooted vertex degree (see Methods). $\langle\sqrt{k}\rangle^2/\langle k\rangle$ relates to the heterogeneity in vertex degree and can therefore be viewed as a measure of heterogeneity in connectivity [80]. Complementary numerical simulations illustrate that this metric can explain differences in population size for complex graph topologies with varying migration regimes (Fig. 2.2A). This analytical result connects with theoretical work on reaction diffusion processes [80] and highlights that irregular graphs result in unbalanced migration fluxes, which affect the ecological balance between births and deaths. Highly connected demes present an oversaturated carrying capacity ($N^{(i)} > K$, see Methods), increasing local competition and lowering total population size compared with regular graphs (Fig. 2.2A). Because populations with small sizes experience more drift ([67], Fig. S6B), this result suggests that graph topology not only affects neutral differentiation through population isolation, but also has a non-trivial effect on neutral differentiation by affecting population size.

Nonetheless, the stochasticity of the processes at the individual level can propagate to the macroscopic level and significantly affect the evolutionary trajectory of the populations. In particular, the mean value of the populations' neutral trait evolves in a stochastic fashion because it is not influenced by a stabilising force, as neutral traits do not provide any selective advantage (see Supplementary Methods and Fig. S4). Those fluctuations complicate an analytical underpinning of the dynamics, and in this case numerical simulations of the IBM offer the most effective approach to evaluate the emergent properties.

2.2 Effect of graph topology on neutral differentiation under no selection pressure

We study a setting with no selection pressure (setting (1)), and investigate the effect of the graph topology on neutral differentiation. When migration is limited, individuals' traits within demes are coherent but stochastic drift at the population level generates neutral differentiation across the demes [81]. Migration attenuates neutral differentiation because it has a correlative effect on local trait distributions, and we expect that the intensity of the correlative effect depends on the topology of the graph. We consider various graphs with identical number of vertices and run simulations of the IBM to obtain the neutral differentiation level attained at a time long enough for transient population dynamics to vanish. We then examine how topology metrics summarising the graph properties explain the discrepancies in neutral differentiation β_u across the simulations.

We compute the 30 scalar graph metrics listed in [77] and assess their Pearson correlations with β_u for varying migration rate m (Fig. S1 and Table S1). We include mean population size and $\langle\sqrt{k}\rangle^2/\langle k\rangle$ (quantifying the graph heterogeneity in degree, see previous section) within the list of metrics, to disentangle whether heterogeneity in connectivity only affects neutral differentiation by reducing population size, or if it has more profound consequences for neutral differentiation. Although such correlations depend on the migration regime (Table S1 and Fig. S10), we find that betweenness centrality variance and characteristic length yield the strongest correlations overall ($\rho = 0.89$ and $\rho = 0.87$, respectively; Fig. 2.2B and Table S1), while population size shows a lower correlation ($\rho = 0.53$). Betweenness centrality variance measures heterogeneity in vertex centrality [77] and characterises the graph heterogeneity in connectivity. We further find that related metrics (eigenvector centrality variance, edge betweenness centrality variance $\langle\sqrt{k}\rangle^2/\langle k\rangle$; Fig. S1) also correlate with β_u ($|\rho| > 0.7$; Table S1). On the other hand, characteristic length measures the average vertex centrality and therefore quantifies the graph connectivity. We observe that related metrics (edge betweenness centrality mean, eigenvector centrality mean, closeness centrality mean, edge density and algebraic connectivity; Fig. S1) correlate with β_u ($|\rho| > 0.7$; Table S1).

We then perform a linear regression analysis with two predictor variables for different migration regimes in

order to evaluate the concurrent effect of the topology metrics on β_u. We identify metrics quantifying graph connectivity and heterogeneity in connectivity as the best combination of explanatory variables (Table S3 and Fig. S1). The best linear model involves the characteristic length of the graph and $\langle\sqrt{k}\rangle^2/\langle k\rangle$, with similar contributions to neutral differentiation ($R^2 = 0.88$; Fig. 2.2C). This result is consistent with insights from the graph metapopulation model of [31], which also proposes characteristic length as a good predictor of neutral β diversity. Nonetheless, in contrast to the authors conclusion that the star graph presents low differentiation due to its low characteristic length, we find that the star graph supports higher differentiation than most graphs because of competitive interactions (Fig. 2.2B). Our model assumes logistic growth, and unbalanced migration fluxes lead central vertices to host more individuals than allowed by their carrying capacity. This causes increased competition that results in a higher mortality rate, so that migrants have a lower probability of further spreading their trait. Highly connected vertices therefore behave as bottlenecks, increasing the isolation of peripheral vertices [28]. We conclude that vertices with a high degree in irregular graphs reduce effective migration and consequently amplify neutral differentiation. In the absence of selection pressure, graphs with high characteristic length and high heterogeneity in degree, or similarly graphs with low connectivity and high heterogeneity in connectivity, show high levels of neutral differentiation.

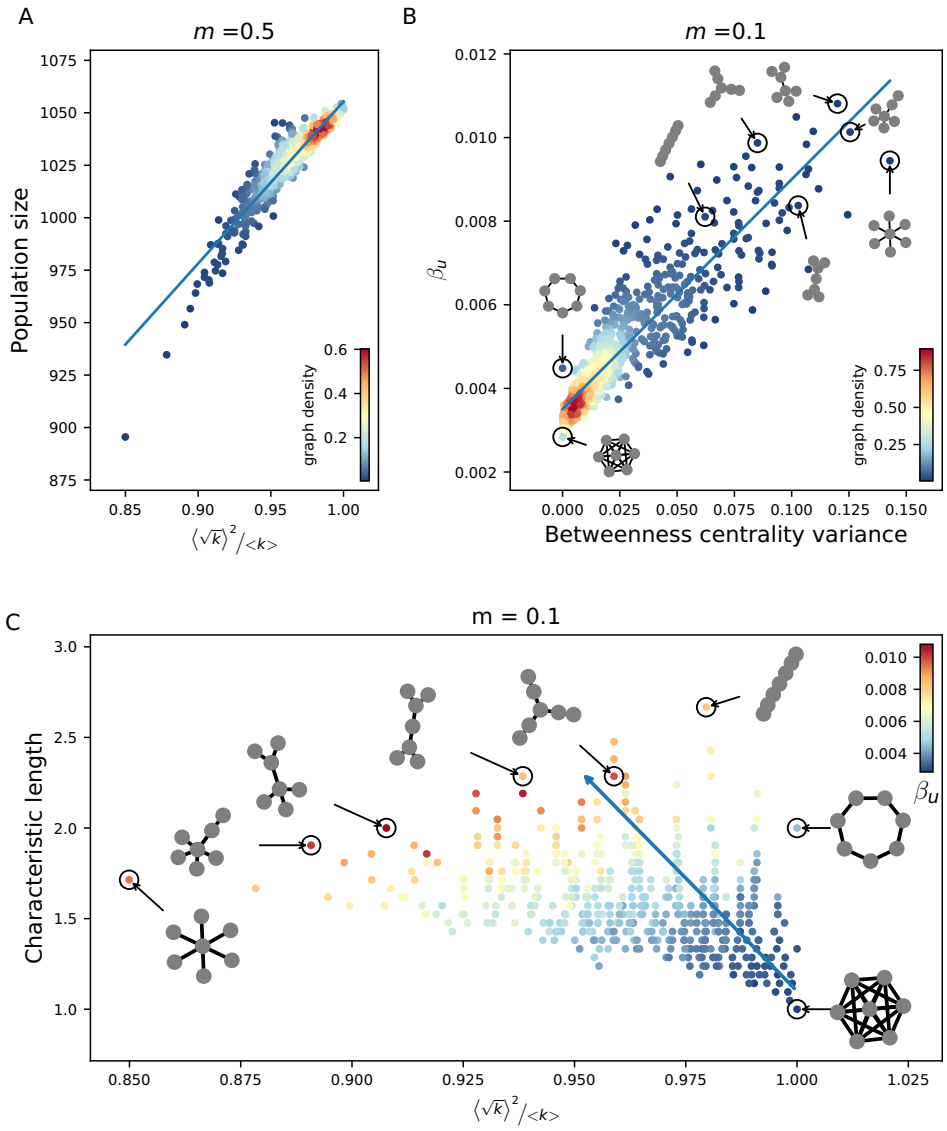

Figure 2.2: Effect of topology metrics on neutral differentiation β_u and population size under no selection pressure (setting (1)). Each plot represents results from simulations of the individual-based model (IBM) on all undirected connected graphs with seven vertices. (A) Response of neutral differentiation β_u to betweenness centrality variance for $m = 0.1$. This topology metric best correlates with β_u for varying migration regimes (Table S1 and Fig. S1). (B) Response of population size N to $\langle\sqrt{k}\rangle^2/\langle k\rangle$ (heterogeneity in degree), for $m = 0.5$. (B) shows that insights obtained from the mean field approximation (see Methods, Eq. (5.2)) hold for more complex graphs and varying migration regimes. In (A) and (B) the colour scale corresponds to the proportion of the graphs with similar x and y axis values (graph density), while the blue line corresponds to a linear fit. (C) Response of β_u to $\langle\sqrt{k}\rangle^2/\langle k\rangle$ and characteristic length. These two topology metrics appear in the best linear model with two predictor variables describing the effect of graph topology on β_u (Table S3). In (C), the colour scale corresponds to the average β_u observed for the graphs with similar $\langle\sqrt{k}\rangle^2/\langle k\rangle$ and characteristic length values, while the blue arrow represents the gradient ∇f where $f\,(^{<k>}/_{<\sqrt{k}>^2}, \text{characteristic length}) = \beta_u$ corresponds to the linear model.

2.3 Effect of complex spatial habitat distribution on adaptive differentiation

We next consider heterogeneous selection pressure and investigate the response of adaptive differentiation to complex spatial habitat distribution. Adaptive differentiation emerges from the adaptation to local habitat conditions, but migration destabilises adaptation because it brings maladapted migrants. We expect that differences in connectivity across habitats, captured by the spatial habitat distribution $\boldsymbol{\theta} = \{\theta_1, \theta_2, \ldots, \theta_M\}$ and the graph topology, influence the proportion of maladapted migrants and therefore affect the level of adaptive differentiation. Selection pressure, captured by the dependence of birth rate on θ_i, stabilises the adaptive trait distribution so that it can be approximated deterministically. We demonstrate under mean field assumption how a graph can be reduced to an equivalent two-deme model. From this simplification arises a quantity, coined the habitat assortativity r_θ, that further determines adaptive differentiation for complex topologies and spatial habitat distributions.

2.3.1 Insights under the mean field assumption

Due to the stabilising force of selection, the stochastic fluctuations of the adaptive trait distribution are small and one can approximate the number of individuals on V_i with traits $s \in \Omega \subset \mathcal{S}$ by the quantity $\int_\Omega n^{(i)}(s)ds$, where $n^{(i)}$ is a continuous function solution of the Partial Differential Equation (PDE)

$$\partial_t n_t^{(i)}(s) = n_t^{(i)}(s)\left[b_i(s)(1-m) - \frac{1}{K}\int_{\mathcal{S}} n_t^{(i)}(\mathbf{s})d\mathbf{s}\right] + m\sum_{j\neq i} b_j(s)\frac{a_{i,j}}{d_j} n_t^{(j)}(s) + \tfrac{1}{2}\mu\sigma_\mu^2\Delta_s\left[b_i(s)n_t^{(i)}(s)\right] \quad (2.3)$$

(see Supplementary Methods for a justification, and Figs. S3 and S5). Equation (2.3) is similar to Eq. (2.2), except that it accounts for selection through the dependence of birth rate on the adaptive trait and that it incorporates a last term corresponding to mutation processes. In the case where $m = 0$, [82, 83] show that Eq. (2.3) admits a stationary solution $n^{(i)}$ that is Gaussian with mean $\bar{s}^{(i)} = \theta_i$, so that $N^{(i)} = \int_{\mathcal{S}} n^{(i)}(s)d\,s = K(1 - \frac{1}{2}\sigma_\mu\sqrt{\mu})$. Hence, under low μ and σ, population size is similar to that in setting (1). It further indicates that, in contrast to β_u, β_s does not depend on population size, but is rather governed by $\mathrm{Var}(\boldsymbol{\theta})$, the variance of the habitat distribution.

We now show under mean field assumption how Eq. (2.3) can be reduced in the general case where $m \geq 0$ to an equivalent two-deme model. The mean field approach slightly differs from setting (1) because vertices are labelled with θ_i. In this case, we assume that vertices with similar habitats have an equivalent position on the graph (see Fig. S8 for a graphical representation), so that all vertices with habitat $\theta_\bullet$ are characterised by the

same adaptive trait distribution that we denote by $\bar{n}^{\bullet}$. Let $P(\bullet,\bullet)$ denote the proportion of edges connecting a vertex with habitat of type $\theta_{\bullet}$ to a vertex with habitat of type $\theta_{\bullet}$, and let $P(\bullet)$ denote the proportion of vertices with habitat $\theta_{\bullet}$. By additionally assuming that habitats are homogeneously distributed so that $P(\bullet) = P(\bullet) = \frac{1}{2}$, Eq. (2.3) transforms into

$$\begin{aligned}\partial_t \bar{n}_t^{\bullet}(s) = \bar{n}_t^{\bullet}(s)\left[b_{\bullet}(s)(1-m) - \frac{1}{K}\int_{\mathcal{S}} \bar{n}_t^{\bullet}(\mathbf{s})d\mathbf{s}\right] + \tfrac{1}{2}\mu\sigma_{\mu}^2(\Delta_s b_{\bullet}\bar{n}_t^{\bullet})(s) \\ + \frac{m}{2}\left[(1-r_\theta)b_{\bullet}(s)\bar{n}_t^{\bullet}(s) + (1+r_\theta)b_{\bullet}(s)\bar{n}_t^{\bullet}(t)\right]\end{aligned} \tag{2.4}$$

(see Methods), where we define:

$$r_\theta = 2\left(P(\bullet,\bullet) - P(\bullet,\bullet)\right) \tag{2.5}$$

as the habitat assortativity of the graph, which ranges from -1 to 1. The case where $r_\theta = -1$ corresponds to the case where all edges map dissimilar habitats (disassortative graph), while as $r_\theta \to 1$ the graph is composed of two clusters of vertices with homogeneous habitats (assortative graph). We numerically solve Eq. (2.4) and show in Fig. 2.3 how both m and r_θ influence population size and β_s. As expected, migration destabilises local adaptation and decreases population size. We find a critical threshold m^* that increases with r_θ; when $m > m^*$, local adaptation cannot be sustained anymore and β_s consequently vanishes (Fig. 2.3A, red curve). Figure 2.3A shows that increasing assortativity r_θ systematically increases β_s at an equivalent migration regime. Our analytical reduction thereby demonstrates that, under the mean field assumption, assortative graphs present high levels of adaptive differentiation. On the other hand, population size and β_s rapidly decline with increasing migration on disassortative graphs, until β_s vanishes completely when $m > m^*$.

2.3.2 Relaxing the mean field assumption

In order to verify the conclusions obtained with the mean field, deterministic approximation Eq. (2.4), we generate randomly different habitat distributions $\boldsymbol{\theta} = \{\theta_1, \theta_2, \ldots, \theta_M\}$ for varying graph topology, and we compare simulations of the IBM (see Methods) with results from the PDE model Eq. (2.4) under different migration regimes (Fig. 2.4). For each combination of $\boldsymbol{\theta}$ and graph, we compute the habitat assortativity r_θ, using the fact that r_θ can be generalised from Eq. (2.5) to any graph topology following the original definition of [84]

$$r_\theta = \frac{\mathrm{Cov}(\boldsymbol{\theta}_\times, \boldsymbol{\theta}_\wedge)}{\sigma_{\boldsymbol{\theta}_\times}\sigma_{\boldsymbol{\theta}_\wedge}}, \tag{2.6}$$

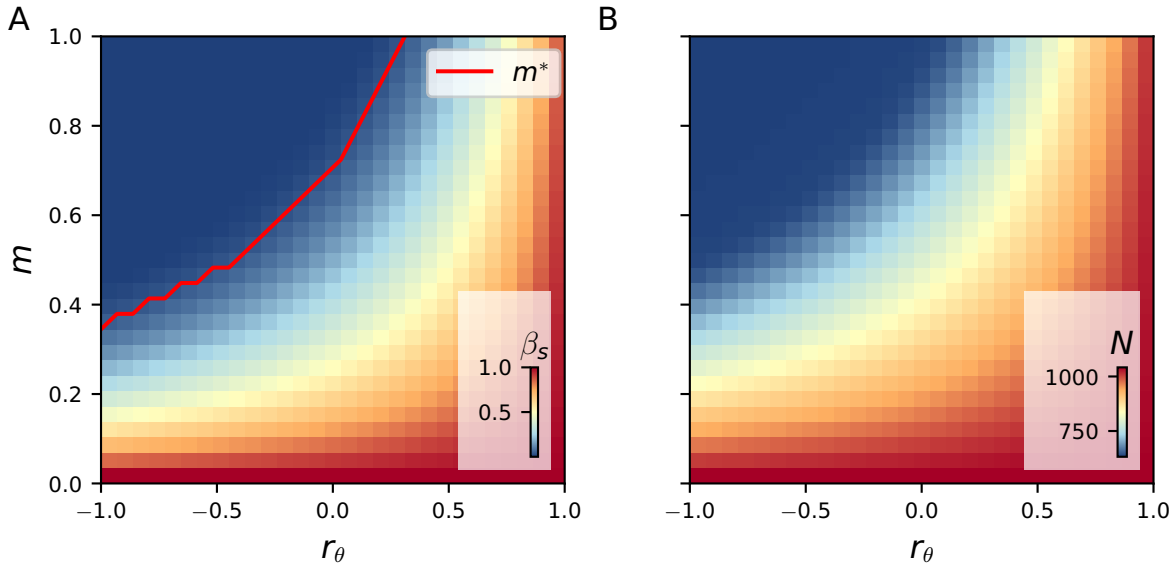

Figure 2.3: Effect of habitat assortativity r_θ and migration on adaptive differentiation β_s and population size under the mean field, deterministic approximation Eq. (2.4). (A) Effect of r_θ on β_s. The red line indicates the critical migration threshold m^* for which β_s vanishes when $m > m^*$. (B) Effect of r_θ on population size. (A) and (B) show that increasing r_θ systematically increases β_s and population size, irrespective of the migration regime.

where $\boldsymbol{\theta}_\times$ and $\boldsymbol{\theta}_\wedge$ denote the sets of habitats found at the toe and tip of each directed vertex of graph V, and $\langle\boldsymbol{\theta}_\times\rangle, \langle\boldsymbol{\theta}_\wedge\rangle$ and $\sigma_{\boldsymbol{\theta}_\times}, \sigma_{\boldsymbol{\theta}_\wedge}$ denote their respective means and standard deviations (see Supplementary Methods). Figure 2.4A,B confirm that the mean field, deterministic approximation Eq. (2.4) captures the response of β_s to r_θ for more general graph ensembles. On the other hand, we assess the correlation between β_s and topology metrics and perform a linear regression analysis to show that the topology metrics have a negligible effect on β_s (Tables S2 and S4).

Under strong selection, stochastic drift vanishes [85] and we find that the effect of complex connectivity patterns across habitat can be reduced to a simplified two-deme model where migration rates are altered by a factor that involves the habitat assortativity r_θ. Mirrahimi et al. [83, 74] investigated a deterministic two-deme model corresponding to the IBM approximation Eq. (2.4) in the special case where $r_\theta = -1$, with the difference that they assumed that individuals continuously migrate and mutate (no dependence of the migration and mutation terms on the birth rate b). The authors provided a closed form for m^*, yielding $m^* = p\theta^2$, that indicates that as selection or habitat heterogeneity θ increases, so does m^* and local adaptation is sustained under higher migration regimes. On the other hand, considering the stochastic dynamics of the IBM, as selection pressure decreases the birth rate b provides only a marginal advantage to adapted individuals, so that

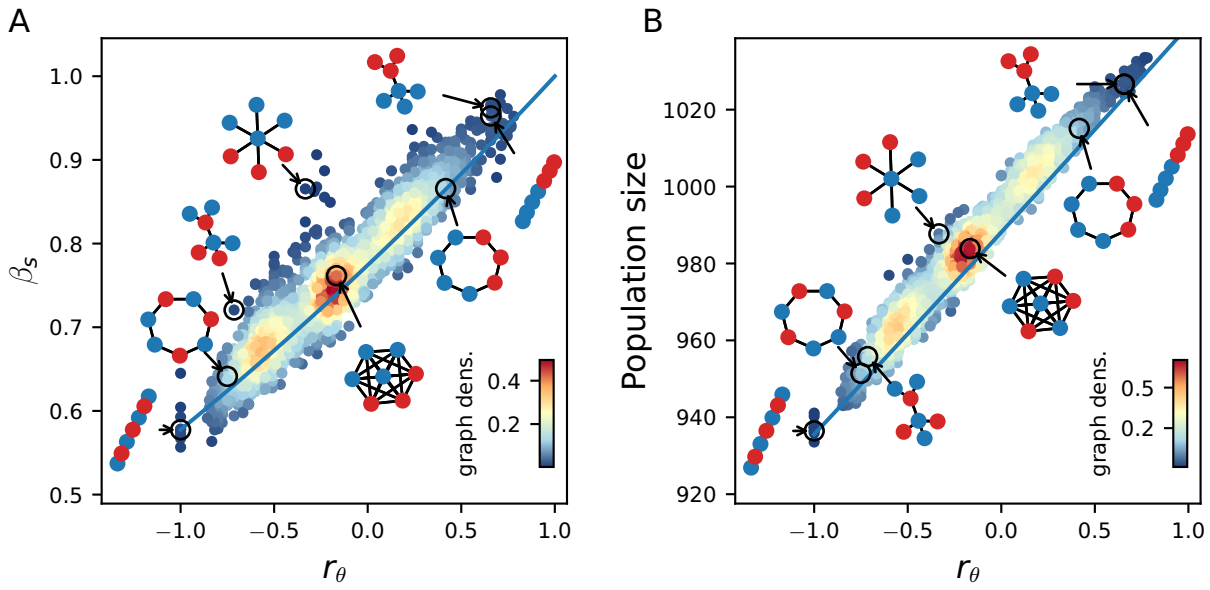

Figure 2.4: Effect of habitat heterogeneity r_θ on β_s and population size for general graph ensembles. Each plot represents results from simulations of the IBM on all undirected connected graphs with seven vertices and varying r_θ, for $m = 0.1$. The colour scale corresponds to the proportion of the graph with similar x and y axis values (graph density). The blue lines correspond to results obtained from the mean field approximation Eq. (2.4). (A) and (B) illustrate that insights from the mean field, deterministic approximation Eq. (2.4) remain valid in the stochastic setting for complex habitat connectivity patterns.

stochastic drift in the populations' mean adaptive trait rises up. In this case the role of habitat assortativity in determining the level of differentiation should be supplanted by the graph topology, so that characteristic length and heterogeneity in degree become the dominant drivers of differentiation patterns. Irrespective of the graph connectivity *per se*, we thus find that under strong enough selection the adaptive differentiation is mainly driven by the habitat assortativity r_θ and vanishes for $m > m^*$, where m^* depends on the habitat assortativity of the graph r_θ, the habitat heterogeneity θ and the selection pressure p.

2.4 Ambiguous effect of habitat assortativity on neutral differentiation

We eventually consider individuals carrying both neutral and adaptive traits (setting (2)). Under heterogeneous habitats, selection pressure promotes neutral differentiation by reducing the reproductive rate of maladapted migrants, reinforcing local population isolation [29]. We have shown that adaptive differentiation is driven by habitat assortativity, so we expect habitat heterogeneity, together with the topological metrics found in setting (1), to influence the level of neutral differentiation. We first investigate how the response of neutral differentiation to migration compares between setting (1) and setting (2) for graphs with an identical topology.

We then examine how the response compares between graphs with an identical topology but different habitat assortativity. We finally consider simulations on different graphs with varying habitat assortativity to assess the concurrent effect of assortativity and topology metrics on neutral differentiation.

Migration has a fitness cost because maladapted migrants present lower fitness [71]. Under an equivalent migration regime, migrants therefore have a lower probability of reproduction, increasing the populations' isolation compared with a scenario without selection pressure [29]. In Figure 2.5A we present simulation results obtained by varying m on the complete graph. This confirms that selection pressure in heterogeneous habitats reinforces neutral differentiation compared with a scenario without selection pressure. Nonetheless, previous results show that adaptive differentiation vanishes with a disassortative graph when $m > m^*$, implying that individuals become equally fit in all habitats. In this case, the isolation effect of heterogeneous selection pressure is lost, and Fig. 2.5A shows that β_u levels in setting (2) reach the same level as in setting (1) for $m > m^*$. This suggests that habitat assortativity reinforces β_u, as assortative graphs sustain higher levels of adaptive differentiation (Fig. 2.4). Simulations on the chain graph with varying spatial habitat distribution give support to this reasoning for high migration regimes, but show that neutral differentiation is higher for low habitat assortativity under low migration regimes (Fig. 2.5B). Assortative graphs are composed of large clusters of vertices with similar habitats, within which migrants can circulate without fitness losses. Local neutral trait distributions become more correlated within those clusters, resulting in a decline in neutral differentiation for assortative graphs compared with disassortative graphs. Figure 2.5B therefore highlights the ambiguous effect of habitat assortativity r_θ on neutral differentiation. r_θ reinforces neutral differentiation by favouring adaptive differentiation, but at the same time it decreases neutral differentiation by decreasing population isolation within similar habitat clusters. A criterion for determining when r_θ favours or depresses neutral differentiation should depend on habitat heterogeneity θ and selection pressure p, but is not trivial and remains to be determined.

We finally compare the effect of habitat assortativity on neutral differentiation to the effect of the connectivity metrics found in setting (1). We perform a linear regression analysis with three predictor variables on simulation results obtained for different graphs with varying habitat distribution, using as predictors r_θ, characteristic length and $\langle\sqrt{k}\rangle^2/\langle k\rangle$ (Fig. 2.5B,C and Table S5). The linear model explains the discrepancies in neutral differentiation across the simulations for varying migration regimes ($R^2 > 0.83$; Table S3), and we find that both r_θ and the two topology metrics contribute similarly to neutral differentiation. This analysis suggests that the effects of habitat assortativity and connectivity of the graph add up under heterogeneous selection pressure. A change in sign of the coefficient for r_θ appears when comparing models fitted for low and high migration regimes, verifying that the opposite effect of habitat assortativity on β_u found on the chain graph

holds for general graph ensembles. Characteristic length and heterogeneity in degree therefore drive neutral differentiation with and without heterogeneous selection pressure (setting (1) vs setting (2)). Habitat assortativity r_θ becomes an additional driver of neutral differentiation under heterogeneous selection pressure. In contrast to the non-ambiguous, positive effect of habitat assortativity on adaptive differentiation, r_θ can amplify or depress neutral differentiation depending on the migration regime considered.

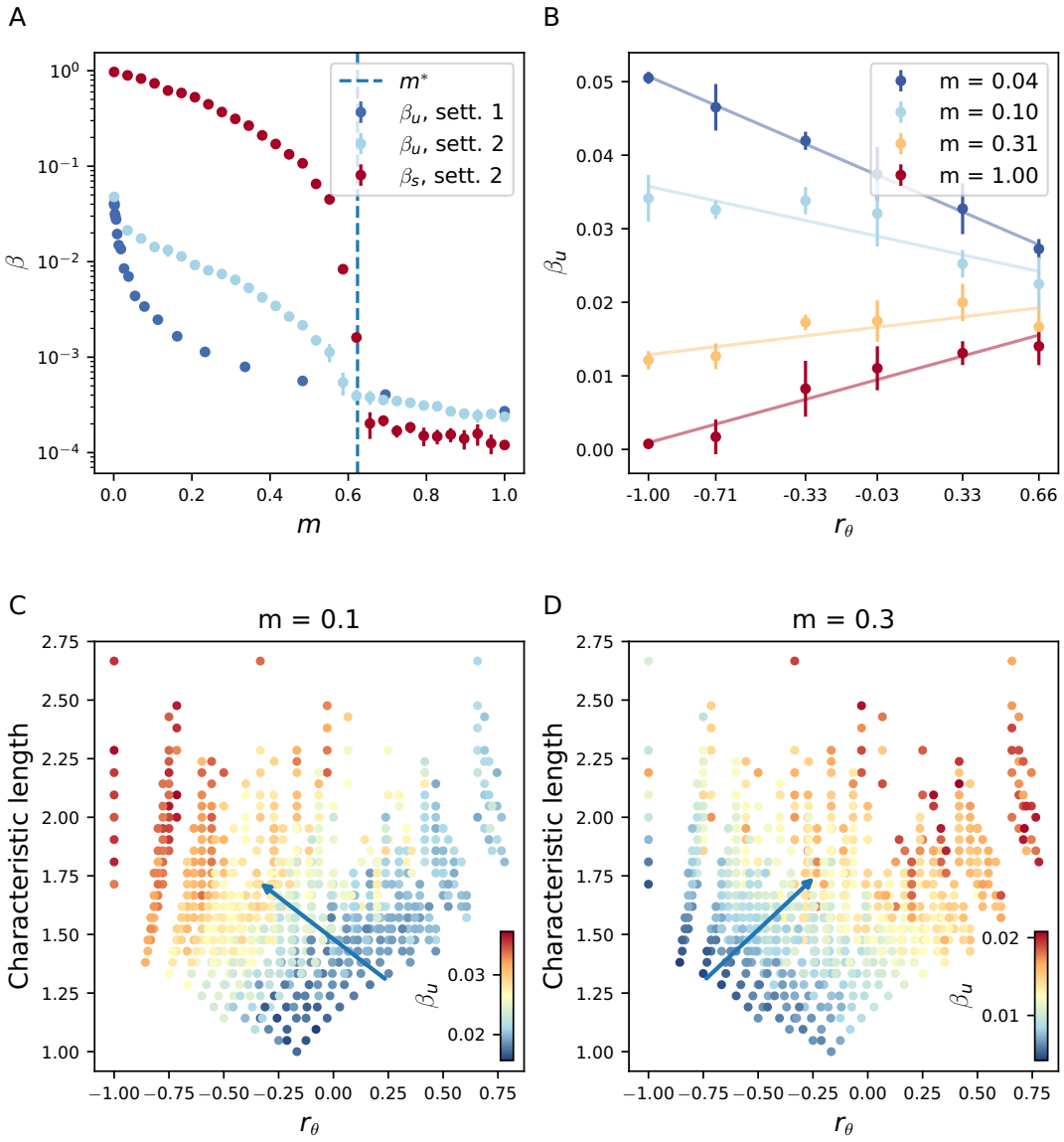

Figure 2.5: Effect of habitat assortativity r_θ and connectivity metrics on neutral differentiation β_u under heterogeneous selection pressure (setting (2)). (A) Comparison of the response of β_u to migration in setting (2) to the response of β_u in setting (1) (no selection) for the complete graph. (A) shows that β_u is higher in setting (2) than in setting (1) when $m < m^*$. However, adaptive differentiation β_s is lost for $m > m^*$, and in this case the β_u level in setting (2) reaches the level of β_u in setting (1). The dashed blue line corresponds to the critical migration regime m^* predicted by the mean field, deterministic approximation Eq. (2.4). (B) Response of β_u to r_θ and migration for the chain graph. (B) illustrates the ambiguous effect of r_θ on β_u and shows that under identical topology, r_θ correlates positively with β_u for high migration regimes, but correlates negatively for low migration regimes. (C) and (D) Concurrent effect of r_θ and characteristic length on β_u for low and high migration regimes. (C) and (D) show results from simulations of the IBM on all undirected connected graphs with seven vertices and varying r_θ (see Methods). (C) and (D) illustrate the additive effects of r_θ and connectivity metrics on β_u, and confirm that the ambiguous effect of r_θ found for the chain graph holds for general graph ensembles. In (C) and (D) the blue arrows represent the gradient ∇f, where $f(r_\theta, \text{characteristic length}, {}^{\langle\sqrt{k}\rangle^2}/_{\langle k\rangle}) = \beta_u$ corresponds to fitted linear model with three predictor variables.

3 Discussion

Using analytical tools and simulations, we have built upon an individual-based model to investigate how landscapes shape population differentiation on graphs, linking emergent patterns to the fundamental underlying processes [86]. Characteristic length, heterogeneity in degree and habitat assortativity are the main drivers of differentiation with respect to our modelling assumptions. Topology metrics can be obtained for any real landscape by projecting it on a graph (e.g. [35, 87]) and can improve our understanding of the origin of spatial biodiversity patterns. Empirical studies have suggested that high species richness stems from small-scale differentiation processes [88], and this pattern seems amplified within complex landscapes such as montane regions [89]. Correlative studies have already associated biodiversity with a variety of metrics that serve as surrogates for connectivity [7, 90, 5, 91], connectivity heterogeneity [10, 9, 11, 92] and habitat heterogeneity [93, 6, 94]. Our study provides a causal link between such metrics and the diverse mechanisms that underlie the differentiation processes [86, 95]. Overall, our results propose topology metrics that can support empiricists to better connect spatial biodiversity patterns to generating eco-evolutionary mechanisms and landscape features.

In the absence of selection pressure, neutral differentiation is more pronounced in graphs with high characteristic length, but is also associated with heterogeneity in degree (Fig. 2.2C). Characteristic length can be viewed as a measure of spatial dimensionality [35], and this result is consistent with empirical studies [7, 90, 5, 91], classical theoretical works based on the stepping stone model [62, 65, 66], as well as the graph metapopulation model of Economo et al. [31]. However, these works did not consider ecological interactions between individuals within populations, and our results show that such interactions can substantially influence the effect of graph topology on neutral differentiation. Our model assumes that population growth is limited by the local carrying capacity, which becomes saturated on highly connected vertices in irregular graphs. As a consequence, these vertices behave as bottlenecks and amplify the isolation of peripheral vertices [28], so that heterogeneity in connectivity appears as an equally strong driver of neutral differentiation in graphs (Fig. 2.2B,C). This behaviour should be prevalent in patchy landscapes where interspecific competition is high due to limiting resources [30]. While it has been acknowledged that reaction diffusion processes behave very differently in graphs with heterogeneous connectivity compared with homogeneous graphs [80, 96, 97], we further document how competitive interactions, as reaction processes, can influence the emergence of phenotypic differentiation in graphs. The role of heterogeneity in connectivity cannot be captured with classical metapopulation and quantitative genetics models [98] or with models of evolutionary dynamics in graphs [47], as they assume constant population size. Our study highlights that heterogeneity in connectivity can reinforce differentiation patterns through the

creation of unbalanced migration fluxes which affect ecological equilibrium.

Habitat assortativity r_θ is a useful indicator for assessing how complex spatial habitat distributions modulate local adaptation and adaptive differentiation [99, 100]. While adaptation has been extensively studied along environmental clines [81, 64, 101, 102, 71, 103, 68], landscapes can be patchy and assuming regularity might be unrealistic [35]. Graphs can capture irregularity in connectivity across habitats [104, 87], and our results indicate that, irrespective of the graph topology, adaptive differentiation is mainly driven by the habitat assortativity r_θ (Fig. 2.4), which can be regarded as the habitat spatial auto-correlation [84]. Habitat assortativity r_θ can be obtained for general habitat distributions that might not be neither binary nor symmetric [105, 84], and we expect our conclusions to hold in such cases. Low habitat assortativity reinforces the "swamping" effect of gene flow proposed by Haldane [102, 99] and limits local adaptation, so that r_θ can be viewed as the spatial scale of habitat variation that conditions local adaptation [106, 107]. Montane regions can support extremely varied habitats within a small spatial scale [8], and in this case habitat assortativity should be low. This implies that for taxa to be locally adapted, a low migration rate is required, which might explain why many species in such environments have small ranges [108, 109]. Our results predict that under an equivalent migration regime, populations structured over assortative habitats are larger, support higher adaptive differentiation, and can be locally adapted even in the case where migration rates are strong.

Spatial eco-evolutionary feedbacks in heterogeneous habitats critically affect diversification [110, 111, 69]. While most eco-evolutionary studies have investigated diversification by considering a unique adaptive trait [68, 101, 102, 71, 103, 112], distinguishing between neutral and adaptive differentiation is crucial [19] and our work underlines their distinct responses to landscape properties (Fig. 2.3A vs. Fig. 2.5C). Our study builds upon recent mathematical models that consider the co-evolution of neutral and adaptive traits [75, 113, 76] and extends those works to a spatial context. Our work provides an analytical framework to the concept of isolation by environment (IBE) [28], which has been suggested to be one of the most important mechanisms governing differentiation in nature [29]. Heterogeneous selection pressure leads to more isolation by modifying the fitness of migrants [71], and therefore affects the level of neutral differentiation (Fig. 2.5A) [30]. In empirical studies an ambiguous response of species richness to habitat heterogeneity has been observed [114, 93, 94]. Our work reconciles this apparent inconsistency and proposes a mechanism by which habitat assortativity, relative to the migration regime, controls the direction of the effect of habitat heterogeneity on diversification (Fig. 2.5C,D). Patchy, heterogeneous habitats can promote neutral differentiation as a result of selection pressure that reduces effective migration [94]. Nonetheless, adaptive differentiation decreases significantly when migration is high. In this case, neutral differentiation should be higher in landscapes with more aggregated habitats [106]. Our

study suggests that habitat assortativity should be considered a unifying metric for understanding the causal link between habitat heterogeneity and species richness in complex environments [94].

4 Conclusion

From a rigorous analytical description of micro-evolutionary processes explicitly accounting for spatial dynamics over graphs, we have established how diversification can emerge at the population level from eco-evolutionary feedbacks in complex landscapes. Our study formalises verbal propositions on how differentiation emerges on ecological time scales from the interplay between spatial dynamics, the co-evolution of neutral and adaptive traits, and landscape properties [115, 28, 29, 106]. In agreement with findings from empirical studies [7, 90, 5, 91, 93, 6, 94] and previous theoretical considerations [31, 62, 65, 66], characteristic length (which relates to landscape connectivity) and habitat assortativity (which relates to habitat spatial auto-correlation) emerge as core drivers of differentiation. We have further shown how the migration regime dictates whether neutral differentiation positively or negatively correlates with habitat heterogeneity. Additionally, our work highlights that heterogeneity in connectivity should equally be a strong driver of differentiation because highly connected demes behave as bottlenecks, increasing the isolation of peripheral demes. The framework laid out here is a promising approach for studying complex adaptive systems, as it can elucidate how macroscopic properties emerge from microscopic processes [14]. It could be used in other fields such as in linguistics [116] or economics [117], where agents interact and are also structured over complex spatio-evolutionary structures.

5 Methods

5.1 Mean field approximation

In setting (1), the mean field approach involves the assumption that all vertices having the same degree are equivalent. For this, let $P(k,k')$ denote the proportion of edges that map a vertex with degree k to a vertex with degree k', and consider the average population size $\bar{N}_t^{(k)}$ in each vertex with degree k. An individual has probability $P(k,k')/k'$ to migrate from a vertex with degree k' to a vertex with degree k. Viewing $a_{i,j}/d_j$ as the probability that an individual on V_i chosen for migration goes to V_j, Eq. (2.2) then transforms into

$$\partial_t \bar{N}_t^{(k)} = \bar{N}_t^{(k)} \left[b(1-m) - \frac{\bar{N}_t^{(k)}}{K} \right] + mbk \sum_{k' \in V} \frac{P(k,k')}{k'} \bar{N}_t^{(k')}. \tag{5.1}$$

Assuming uncorrelated networks for which ${}^{P(k,k')}/_{k'} = {}^{P(k')k'}/_{\langle k\rangle}$, where $\langle k\rangle$ denotes the average degree of the graph [118, 119], yields

$$\partial_t \bar{N}_t^{(k)} = \bar{N}_t^{(k)} \left[b(1-m) - \frac{\bar{N}_t^{(k)}}{K} \right] + mb \frac{k}{\langle k\rangle} \bar{N}_t, \tag{5.2}$$

where

$$\bar{N}_t = \sum_k P(k) \bar{N}_t^{(k)}. \tag{5.3}$$

When solving for the stationary state and setting $m = 1$, one obtains $\bar{N}^{(k)} = \sqrt{bK \frac{k}{\langle k\rangle} \bar{N}}$ from Eq. (5.2). Combining this with Eq. (5.3) yields $\bar{N} = bK \frac{\left[\sum_k P(k)\sqrt{k}\right]^2}{\langle k\rangle} = bK \langle\sqrt{k}\rangle^2 / \langle k\rangle$.

In setting (2), the mean field approach involves the assumption that all vertices with a similar habitat are equivalent. In this case, an individual from a vertex with habitat $\theta_\bullet$ has the probability ${}^{P(\bullet,\bullet)}/_{P(\bullet)}$ to migrate to a vertex with habitat $\theta_\bullet$, and therefore Eq. (2.3) transforms into

$$\begin{aligned} \partial_t \bar{n}_t^{\bullet}(s) = \bar{n}_t^{\bullet}(s) \left[b_\bullet(s)(1-m) - \frac{1}{K} \int_{\mathcal{S}} \bar{n}_t^{\bullet}(\mathbf{s}) d\mathbf{s} \right] + \tfrac{1}{2}\mu\sigma_\mu^2 \Delta_s \left[b_\bullet(s) \bar{n}_t^{\bullet}(s) \right] \\ + m \sum_{j \in \{\bullet,\bullet\}} b_j(s) \frac{P(\bullet, j)}{P(j)} \bar{n}_t^{(j)}(s). \end{aligned} \tag{5.4}$$

Considering that $P(\bullet) = P(\bullet) = \frac{1}{2}$ (habitats are equally distributed), $P(\bullet,\bullet) + P(\bullet,\bullet) = P(\bullet)$ (sum of conditional expectations) and $r_\theta = 2\left(P(\bullet,\bullet) - P(\bullet,\bullet)\right)$ (Eq. (2.5)), one obtains

$$P(\bullet,\bullet) = \tfrac{1}{4}(1 - r_\theta) \quad \text{and} \quad P(\bullet,\bullet) = \tfrac{1}{4}(1 + r_\theta). \tag{5.5}$$

Combining Eq. (5.5) with Eq. (5.4) yields Eq. (2.4). We show in Supplementary Methods how one can derive Eq. (2.5) from the general definition of assortativity given in Eq. (2.6) and initially introduced in [84].

5.2 Numerical simulations

We performed Monte Carlo simulations by running five replicate simulations for each result presented, with $b = 1$, total time span $t = 500$, local carrying capacity $K = 150$, mutation rate $\mu = 0.1$, and mutation range $\sigma_\mu = 5{\cdot}10^{-2}$. This parameter choice corresponds to 10^6 birth and death events, which made it possible to obtain results in a reasonable amount of time without including transient population dynamics. In settings (1) and (2), we ran simulations on all of the 853 undirected connected graphs with 7 vertices (listed at http://oeis.org/A001349).

For setting (2), we randomly generated different spatial habitat distributions and selected the ones with a unique r_θ value, corresponding to a total of 2537 different simulations for fixed m. We then computed diversity metrics that we further averaged over the last time steps and across the replicates. Since neutral diversity dynamics are characterised by large quadratic variations, we simulated individuals with $\mathfrak{u} = 300$ neutral traits, where each trait can independently be affected by mutations. Diversity metrics were then obtained from the average $\alpha_u, \beta_u, \gamma_u$ for each trait. This reduced the variance of the numerical simulations and is also biologically meaningful because populations are characterised by many traits, most of which are neutral [19]. As initial conditions, MK individuals were homogeneously distributed across all of the vertices, with traits centred at 0 and with standard deviation σ_μ. Graph metrics used for the meta-analysis were calculated using the **LightGraphs.jl** library [120]. We numerically solved the PDEs with a finite difference scheme using **DifferentialEquations.jl** [121], ensuring that the domain was large enough to avoid border effects.

EvoId.jl

We have implemented the modelling framework in a computationally efficient and user-friendly package, **EvoId.jl**, written in the Julia programming language and freely accessible at https://github.com/vboussange/EvoId.jl. The user can specify any combination of spaces over which the populations are structured, together with birth and death functions and specific update rules. The intention is to popularise the tool and encourage the community to investigate this framework for multi-disciplinary case studies.

Acknowledgements

We thank Thomas Poulet, Sylvian Billiard, Sepideh Mirrahimi, Heike Lischke, Joshua Payne, Conor Waldock and Alexander Skeels for helpful discussions and comments on the manuscript. L.P. was supported by the Swiss National Science Foundation grant (N° 310030_188550).

Data availability

The code used in this article is available online at https://gitlab.ethz.ch/publications/neutral-and-adaptive-diversification-in-spatial-graphs.

References

[1] Stephen P Hubbell. *The unified neutral theory of biodiversity and biogeography*. Monographs in population biology 32. Princeton University press, Princeton [etc, 2001.

[2] Mark Vellend. Conceptual Synthesis in Community Ecology. *The Quarterly Review of Biology*, 85(2):183–206, jun 2010.

[3] Carsten Rahbek, Michael K. Borregaard, Alexandre Antonelli, Robert K. Colwell, Ben G. Holt, David Nogues-Bravo, Christian M.Ø. Rasmussen, Katherine Richardson, Minik T. Rosing, Robert J. Whittaker, and Jon Fjeldså. Building mountain biodiversity: Geological and evolutionary processes. *Science*, 365(6458):1114–1119, 2019.

[4] Wen-Na Ding, Richard H. Ree, Robert A. Spicer, and Yao-Wu Xing. Ancient orogenic and monsoon-driven assembly of the world's richest temperate alpine flora. *Science*, 369(6503):578–581, jul 2020.

[5] Murilo S. Dias, Thierry Oberdorff, Bernard Hugueny, Fabien Leprieur, Céline Jézéquel, Jean-François Cornu, Sébastien Brosse, Gael Grenouillet, and Pablo A. Tedesco. Global imprint of historical connectivity on freshwater fish biodiversity. *Ecology Letters*, 17(9):1130–1140, sep 2014.

[6] Jean-François Guégan, Sovan Lek, and Thierry Oberdorff. Energy availability and habitat heterogeneity predict global riverine fish diversity. *Nature*, 391(6665):382–384, jan 1998.

[7] M. Scheffer, G. J. van Geest, K. Zimmer, E. Jeppesen, M. Søndergaard, M. G. Butler, M. A. Hanson, S. Declerck, and L. De Meester. Small habitat size and isolation can promote species richness: second-order effects on biodiversity in shallow lakes and ponds. *Oikos*, 112(1):227–231, jan 2006.

[8] Alexandre Antonelli, W. Daniel Kissling, Suzette G.A. Flantua, Mauricio A. Bermúdez, Andreas Mulch, Alexandra N. Muellner-Riehl, Holger Kreft, H. Peter Linder, Catherine Badgley, Jon Fjeldså, Susanne A. Fritz, Carsten Rahbek, Frédéric Herman, Henry Hooghiemstra, and

Carina Hoorn. Geological and climatic influences on mountain biodiversity. *Nature Geoscience*, 11(10):718–725, 2018.

[9] Holger Kreft and Walter Jetz. Global patterns and determinants of vascular plant diversity. *Proceedings of the National Academy of Sciences*, 104(14):5925–5930, apr 2007.

[10] Carsten Rahbek and Gary R. Graves. Multiscale assessment of patterns of avian species richness. *Proceedings of the National Academy of Sciences*, 98(8):4534–4539, apr 2001.

[11] Richard G. Davies, C. David L. Orme, David Storch, Valerie A. Olson, Gavin H. Thomas, Simon G. Ross, Tzung-Su Ding, Pamela C. Rasmussen, Peter M. Bennett, Ian P.F. Owens, Tim M. Blackburn, and Kevin J. Gaston. Topography, energy and the global distribution of bird species richness. *Proceedings of the Royal Society B: Biological Sciences*, 274(1614):1189–1197, may 2007.

[12] Hanna Tuomisto. A diversity of beta diversities: straightening up a concept gone awry. Part 1. Defining beta diversity as a function of alpha and gamma diversity. *Ecography*, 33(1):2–22, feb 2010.

[13] Richard K. Grosberg, Geerat J. Vermeij, and Peter C. Wainwright. Biodiversity in water and on land. *Current Biology*, 22(21):R900–R903, nov 2012.

[14] Simon A Levin. Complex adaptive systems: Exploring the known, the unknown and the unknowable. *Bulletin of the American Mathematical Society*, 40(01):3–20, oct 2002.

[15] Juliano Sarmento Cabral, Luis Valente, and Florian Hartig. Mechanistic simulation models in macroecology and biogeography: state-of-art and prospects. *Ecography*, 40(2):267–280, feb 2017.

[16] Sébastien Lion. Moment equations in spatial evolutionary ecology. *Journal of Theoretical Biology*, 405:46–57, 2016.

[17] A R Templeton. Mechanisms of Speciation - A Population Genetic Approach. *Annual Review of Ecology and Systematics*, 12(1):23–48, nov 1981.

[18] Dolph Schluter. Evidence for ecological speciation and its alternative. *Science*, 323(5915):737–741, 2009.

[19] Rolf Holderegger, Urs Kamm, and Felix Gugerli. Adaptive vs. neutral genetic diversity: implications for landscape genetics. *Landscape Ecology*, 21(6):797–807, aug 2006.

[20] Montgomery Slatkin. ISOLATION BY DISTANCE IN EQUILIBRIUM AND NON-EQUILIBRIUM POPULATIONS. *Evolution*, 47(1):264–279, feb 1993.

[21] Theodosius Dobzhansky. *Genetics and the origin of species*. Columbia classics in evolution series. Columbia University Press, New York, repr. edition, 1982.

[22] Ulf Dieckmann and Michael Doebeli. On the origin of species by sympatric speciation. *Nature*, 400(6742):354–357, jul 1999.

[23] S. Gavrilets and J. B. Losos. Adaptive Radiation: Contrasting Theory with Data. *Science*, 323(5915):732–737, feb 2009.

[24] Dolph Schluter. *The ecology of adaptive radiation*. Oxford series in ecology and evolution. Oxford University Press, Oxford, 2000.

[25] Howard D. Rundle and Patrik Nosil. Ecological speciation. *Ecology Letters*, 8(3):336–352, jan 2005.

[26] Norman L Kaplan, R R Hudson, and C H Langley. The "hitchhiking effect" revisited. *Genetics*, 123(4):887–899, dec 1989.

[27] Patrik Nosil, Timothy H. Vines, and Daniel J. Funk. REPRODUCTIVE ISOLATION CAUSED BY NATURAL SELECTION AGAINST IMMIGRANTS FROM DIVERGENT HABITATS. *Evolution*, 59(4):705–719, apr 2005.

[28] Luisa Orsini, Joost Vanoverbeke, Ine Swillen, Joachim Mergeay, and Luc De Meester. Drivers of population genetic differentiation in the wild: Isolation by dispersal limitation, isolation by adaptation and isolation by colonization. *Molecular Ecology*, 22(24):5983–5999, 2013.

[29] Ian J. Wang and Gideon S. Bradburd. Isolation by environment. *Molecular Ecology*, 23(23):5649–5662, 2014.

[30] DANY GARANT, SAMANTHA E. FORDE, and ANDREW P. HENDRY. The multifarious effects of dispersal and gene flow on contemporary adaptation. *Functional Ecology*, 21(3):434–443, jun 2007.

[31] Evan P. Economo and Timothy H. Keitt. Species diversity in neutral metacommunities: a network approach. *Ecology Letters*, 11(1):071117033013001–???, nov 2007.

[32] Helmut Hillebrand and Birte Matthiessen. Biodiversity in a complex world: consolidation and progress in functional biodiversity research. *Ecology Letters*, 12(12):1405–1419, dec 2009.

[33] Jani Heino, Adriano S. Melo, Tadeu Siqueira, Janne Soininen, Sebastian Valanko, and Luis Mauricio Bini. Metacommunity organisation, spatial extent and dispersal in aquatic systems: patterns, processes and prospects. *Freshwater Biology*, 60(5):845–869, may 2015.

[34] Carlos J Melián, Blake Matthews, Cecilia S. de Andreazzi, Jorge P Rodríguez, Luke J Harmon, and Miguel A Fortuna. Deciphering the Interdependence between Ecological and Evolutionary Networks. *Trends in Ecology & Evolution*, 33(7):504–512, jul 2018.

[35] M. R.T. Dale and M. J. Fortin. From graphs to spatial graphs. *Annual Review of Ecology, Evolution, and Systematics*, 41:21–38, 2010.

[36] Michael Bode, Kevin Burrage, and Hugh P. Possingham. Using complex network metrics to predict the persistence of metapopulations with asymmetric connectivity patterns. *Ecological Modelling*, 214(2-4):201–209, jun 2008.

[37] Matthew D. Holland and Alan Hastings. Strong effect of dispersal network structure on ecological dynamics. *Nature*, 456(7223):792–794, dec 2008.

[38] Luis J. Gilarranz and Jordi Bascompte. Spatial network structure and metapopulation persistence. *Journal of Theoretical Biology*, 297:11–16, 2012.

[39] Lorenzo Mari, Renato Casagrandi, Enrico Bertuzzo, Andrea Rinaldo, and Marino Gatto. Metapopulation persistence and species spread in river networks. *Ecology Letters*, 17(4):426–434, apr 2014.

[40] Dominique Gravel, François Massol, and Mathew A. Leibold. Stability and complexity in model meta-ecosystems. *Nature Communications*, 7(1):12457, dec 2016.

[41] Andreas Brechtel, Philipp Gramlich, Daniel Ritterskamp, Barbara Drossel, and Thilo Gross. Master stability functions reveal diffusion-driven pattern formation in networks. *Physical Review E*, 97(3):032307, mar 2018.

[42] Francesco Carrara, Florian Altermatt, Ignacio Rodriguez-Iturbe, and Andrea Rinaldo. Dendritic connectivity controls biodiversity patterns in experimental metacommunities. *Proceedings of the National Academy of Sciences*, 109(15):5761–5766, apr 2012.

[43] Patrick L. Thompson, Bronwyn Rayfield, and Andrew Gonzalez. Loss of habitat and connectivity erodes species diversity, ecosystem functioning, and stability in metacommunity networks. *Ecography*, 40(1):98–108, jan 2017.

[44] Yuka Suzuki and Evan P. Economo. From species sorting to mass effects: spatial network structure mediates the shift between metacommunity archetypes. *Ecography*, page ecog.05453, feb 2021.

[45] F Pelletier, D Garant, and A P Hendry. Eco-evolutionary dynamics. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, 364(1523):1483–1489, jun 2009.

[46] Erez Lieberman, Christoph Hauert, and Martin A. Nowak. Evolutionary dynamics on graphs. *Nature*, 433(7023):312–316, jan 2005.

[47] Martin A. Nowak, Corina E. Tarnita, and Tibor Antal. Evolutionary dynamics in structured populations. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 365(1537):19–30, 2010.

[48] Marcus Frean, Paul B. Rainey, and Arne Traulsen. The effect of population structure on the rate of evolution. *Proceedings of the Royal Society B: Biological Sciences*, 280(1762):20130211, jul 2013.

[49] Josef Tkadlec, Andreas Pavlogiannis, Krishnendu Chatterjee, and Martin A. Nowak. Population structure determines the tradeoff between fixation probability and fixation time. *Communications Biology*, 2(1):138, dec 2019.

[50] Marius Möller, Laura Hindersin, and Arne Traulsen. Exploring and mapping the universe of evolutionary graphs identifies structural properties affecting fixation probability and time. *Communications Biology*, 2(1):137, dec 2019.

[51] Benjamin Allen, Gabor Lippner, Yu Ting Chen, Babak Fotouhi, Naghmeh Momeni, Shing Tung Yau, and Martin A. Nowak. Evolutionary dynamics on any population structure. *Nature*, 544(7649):227–230, 2017.

[52] Bahram Houchmandzadeh and Marcel Vallade. The fixation probability of a beneficial mutation in a geographically structured population. *New Journal of Physics*, 13, 2011.

[53] Evan P. Economo and Timothy H. Keitt. Network isolation and local diversity in neutral metacommunities. *Oikos*, 119(8):1355–1363, jul 2010.

[54] Rachata Muneepeerakul, Enrico Bertuzzo, Heather J. Lynch, William F. Fagan, Andrea Rinaldo, and Ignacio Rodriguez-Iturbe. Neutral metacommunity models predict fish diversity patterns in Mississippi–Missouri basin. *Nature*, 453(7192):220–222, may 2008.

[55] B. Neuffer and H. Hurka. Colonization history and introduction dynamics of Capsella bursa-pastoris (Brassicaceae) in North America: isozymes and quantitative traits. *Molecular Ecology*, 8(10):1667–1681, oct 1999.

[56] Lærke Stewart, Inger G. Alsos, Christian Bay, Amy L. Breen, Christian Brochmann, Noémie Boulanger-Lapointe, Olivier Broennimann, Helga Bültmann, Peder Klith Bøcher, Christian

Damgaard, Fred J.A. Daniëls, Dorothee Ehrich, Pernille Bronken Eidesen, Antoine Guisan, Ingibjörg Svala Jónsdóttir, Jonathan Lenoir, Peter C. le Roux, Esther Lévesque, Miska Luoto, Jacob Nabe-Nielsen, Peter Schönswetter, Andreas Tribsch, Liv Unn Tveraabak, Risto Virtanen, Donald A. Walker, Kristine B. Westergaard, Nigel G. Yoccoz, Jens Christian Svenning, Mary Wisz, Niels Martin Schmidt, and Loïc Pellissier. The regional species richness and genetic diversity of Arctic vegetation reflect both past glaciations and current climate. *Global Ecology and Biogeography*, 25(4):430–442, 2016.

[57] Stéphanie Manel, Pierre-Edouard Guerin, David Mouillot, Simon Blanchet, Laure Velez, Camille Albouy, and Loïc Pellissier. Global determinants of freshwater and marine fish genetic diversity. *Nature Communications*, 11(1):692, dec 2020.

[58] Nicolas Champagnat, Régis Ferrière, and Sylvie Méléard. Unifying evolutionary dynamics: From individual stochastic processes to macroscopic models. *Theoretical Population Biology*, 69(3):297–321, may 2006.

[59] Vincent Bansaye and Sylvie Méléard. Some stochastic models for structured populations : scaling limits and long time behavior. *Stochastic Models for Structured Populations: Scaling Limits and Long Time Behavior*, pages 1–107, jun 2015.

[60] R. A. FISHER. THE WAVE OF ADVANCE OF ADVANTAGEOUS GENES. *Annals of Eugenics*, 7(4):355–369, jun 1937.

[61] Sewal Wright. Isolation By Distance. *Genetics*, 1943.

[62] M Kimura and G H Weiss. The Stepping Stone Model of Population Structure and the Decrease of Genetic Correlation with Distance. *Genetics*, 49(4):561–76, apr 1964.

[63] B. Houchmandzadeh and M. Vallade. Fisher waves: An individual-based stochastic model. *Physical Review E*, 96(1):1–13, 2017.

[64] Montgomery Slatkin. Spatial patterns in the distributions of polygenic characters. *Journal of Theoretical Biology*, 70(2):213–228, jan 1978.

[65] R. Lande. Isolation by distance in a quantitative trait. *Genetics*, 128(2):443–452, 1991.

[66] T. Nagylaki. Geographical variation in a quantitative character. *Genetics*, 136(1):361–81, jan 1994.

[67] Reinhard Bürger. *The mathematical theory of selection, recombination, and mutation*. Wiley series in mathematical and computational biology. J. Wiley, Chichester [etc, 2000.

[68] Michael Doebeli and Ulf Dieckmann. Speciation along environmental gradients. *Nature*, 421(6920):259–264, jan 2003.

[69] Robin Aguilée, David Claessen, and Amaury Lambert. ADAPTIVE RADIATION DRIVEN BY THE INTERPLAY OF ECO-EVOLUTIONARY AND LANDSCAPE DYNAMICS. *Evolution*, 67(5):no–no, dec 2012.

[70] Fanny Gascuel, Régis Ferrière, Robin Aguilée, and Amaury Lambert. How Ecology and Landscape Dynamics Shape Phylogenetic Trees. *Systematic Biology*, 64(4):590–607, 2015.

[71] Jitka Polechová. Is the sky the limit? On the expansion threshold of a species' range. *PLoS Biology*, 16(6):1–18, 2018.

[72] F. Débarre, O. Ronce, and S. Gandon. Quantifying the effects of migration and mutation on adaptation and demography in spatially heterogeneous environments. *Journal of Evolutionary Biology*, 26(6):1185–1202, 2013.

[73] Jonas Wickman, Sebastian Diehl, Bernd Blasius, Christopher A. Klausmeier, Alexey B. Ryabov, and Åke Brännström. Determining Selection across Heterogeneous Landscapes: A Perturbation-Based Method and Its Application to Modeling Evolution in Space. *The American Naturalist*, 189(4):381–395, apr 2017.

[74] Sepideh Mirrahimi and Sylvain Gandon. Evolution of Specialization in Heterogeneous Environments: Equilibrium Between Selection, Mutation and Migration. *Genetics*, 214(2):479–491, feb 2020.

[75] Sylvain Billiard, Régis Ferrière, Sylvie Méléard, and Viet Chi Tran. Stochastic dynamics of adaptive trait and neutral marker driven by eco-evolutionary feedbacks. *Journal of Mathematical Biology*, 71(5):1211–1242, nov 2015.

[76] Niccolo Anceschi, Jorge Hidalgo, Carlos A. Plata, Tommaso Bellini, Amos Maritan, and Samir Suweis. Neutral and niche forces as drivers of species selection. *Journal of Theoretical Biology*, 483:109969, dec 2019.

[77] Gergana Bounova and Olivier de Weck. Overview of metrics and their correlation patterns for multiple-metric topology analysis on heterogeneous graph ensembles. *Physical Review E*, 85(1):016117, jan 2012.

[78] Michel Loreau. Are communities saturated? On the relationship between alpha, beta and gamma diversity. *Ecology Letters*, 3(2):73–76, mar 2000.

[79] Daniel T. Gillespie. A general method for numerically simulating the stochastic time evolution of coupled chemical reactions. *Journal of Computational Physics*, 22(4):403–434, dec 1976.

[80] Vittoria Colizza, Romualdo Pastor-Satorras, and Alessandro Vespignani. Reaction–diffusion processes and metapopulation models in heterogeneous networks. *Nature Physics*, 3(4):276–282, apr 2007.

[81] Montgomery Slatkin. GENE FLOW AND SELECTION IN A CLINE. *Genetics*, 75(4):733–756, dec 1973.

[82] M. Kimura. A stochastic model concerning the maintenance of genetic variability in quantitative characters. *Proceedings of the National Academy of Sciences*, 54(3):731–736, sep 1965.

[83] Sepideh Mirrahimi. A Hamilton-Jacobi approach to characterize the evolutionary equilibria in heterogeneous environments. *Mathematical Models and Methods in Applied Sciences*, 27(13):2425–2460, dec 2016.

[84] M. E. J. Newman. Mixing patterns in networks. *Physical Review E*, 67(2):026126, feb 2003.

[85] Thomas Lenormand, Denis Roze, and François Rousset. Stochasticity in evolution. *Trends in Ecology & Evolution*, 24(3):157–165, mar 2009.

[86] Volker Grimm. Pattern-Oriented Modeling of Agent-Based Complex Systems: Lessons from Ecology. *Science*, 310(5750):987–991, nov 2005.

[87] Vincent Miele and Catherine Matias. Spatially constrained clustering of ecological networks. pages 771–779, 2014.

[88] Mark Vellend and Monica A. Geber. Connections between species diversity and genetic diversity. *Ecology Letters*, 8(7):767–781, 2005.

[89] Carsten Rahbek, Michael K. Borregaard, Robert K. Colwell, Bo Dalsgaard, Ben G. Holt, Naia Morueta-Holme, David Nogues-Bravo, Robert J. Whittaker, and Jon Fjeldså. Humboldt's enigma: What causes global patterns of mountain biodiversity? *Science*, 365(6458):1108–1113, sep 2019.

[90] Cong Liu, Kenneth L. Dudley, Zheng-Hui Xu, and Evan P. Economo. Mountain metacommunities: climate and spatial connectivity shape ant diversity in a complex landscape. *Ecography*, 41(1):101–112, jan 2018.

[91] Enrico Bertuzzo, Francesco Carrara, Lorenzo Mari, Florian Altermatt, Ignacio Rodriguez-Iturbe, and Andrea Rinaldo. Geomorphic controls on elevational gradients of species richness. *Proceedings of the National Academy of Sciences*, 113(7):1737–1742, feb 2016.

[92] Jeremy T. Kerr and Laurence Packer. Habitat heterogeneity as a determinant of mammal species richness in high-energy regions. *Nature*, 385(6613):252–254, jan 1997.

[93] Joseph A. Veech and Thomas O. Crist. Habitat and climate heterogeneity maintain beta-diversity of birds among landscapes within ecoregions. *Global Ecology and Biogeography*, 16(5):650–656, sep 2007.

[94] Anke Stein, Katharina Gerstner, and Holger Kreft. Environmental heterogeneity as a universal driver of species richness across taxa, biomes and spatial scales. *Ecology Letters*, 17(7):866–880, 2014.

[95] Stéphanie Manel, Michael K. Schwartz, Gordon Luikart, and Pierre Taberlet. Landscape genetics: combining landscape ecology and population genetics. *Trends in Ecology & Evolution*, 18(4):189–197, apr 2003.

[96] Marc Barthélemy. Spatial networks, 2011.

[97] Romualdo Pastor-Satorras, Claudio Castellano, Piet Van Mieghem, and Alessandro Vespignani. Epidemic processes in complex networks. *Reviews of Modern Physics*, 87(3):925–979, aug 2014.

[98] François Rousset. *Genetic structure and selection in subdivided populations*. Monographs in population biology 40. Princeton University Press, Princeton [etc, 2004.

[99] Tadeusz J. Kawecki and Dieter Ebert. Conceptual issues in local adaptation. *Ecology Letters*, 7(12):1225–1241, dec 2004.

[100] Outi Savolainen, Tanja Pyhäjärvi, and Timo Knürr. Gene Flow and Local Adaptation in Trees. *Annual Review of Ecology, Evolution, and Systematics*, 38(1):595–619, dec 2007.

[101] Mark Kirkpatrick and N. H. Barton. Evolution of a Species' Range. *The American Naturalist*, 150(1):1–23, jul 1997.

[102] Jitka Polechová and Nicholas H. Barton. Limits to adaptation along environmental gradients. *Proceedings of the National Academy of Sciences of the United States of America*, 112(20):6401–6406, 2015.

[103] Martín Andrade-Restrepo, Nicolas Champagnat, and Régis Ferrière. Local adaptation, dispersal evolution, and the spatial eco-evolutionary dynamics of invasion. *Ecology Letters*, 22(5):767–777, may 2019.

[104] A.G Bunn, D.L Urban, and T.H Keitt. Landscape connectivity: A conservation application of graph theory. *Journal of Environmental Management*, 59(4):265–278, aug 2000.

[105] M. E. J. Newman. Assortative Mixing in Networks. *Physical Review Letters*, 89(20):208701, oct 2002.

[106] Jonathan L. Richardson, Mark C. Urban, Daniel I. Bolnick, and David K. Skelly. Microgeographic adaptation and the spatial scale of evolution. *Trends in Ecology & Evolution*, 29(3):165–176, mar 2014.

[107] Katalin Csilléry, Otso Ovaskainen, Christoph Sperisen, Nina Buchmann, Alex Widmer, and Felix Gugerli. Adaptation to local climate in multi-trait space: evidence from silver fir (Abies alba Mill.) populations across a heterogeneous environment. *Heredity*, 124(1):77–92, jan 2020.

[108] Jon Fjeldså, Rauri C.K. Bowie, and Carsten Rahbek. The Role of Mountain Ranges in the Diversification of Birds. *Annual Review of Ecology, Evolution, and Systematics*, 43(1):249–265, dec 2012.

[109] Carsten Rahbek, Michael K. Borregaard, Alexandre Antonelli, Robert K. Colwell, Ben G. Holt, David Nogues-Bravo, Christian M.Ø. Rasmussen, Katherine Richardson, Minik T. Rosing, Robert J. Whittaker, and Jon Fjeldså. Building mountain biodiversity: Geological and evolutionary processes. *Science*, 365(6458):1114–1119, 2019.

[110] Robin Aguilée, Fanny Gascuel, Amaury Lambert, and Regis Ferriere. Clade diversification dynamics and the biotic and abiotic controls of speciation and extinction rates. *Nature Communications*, 9(1):1–13, 2018.

[111] Katja Räsänen and Andrew P. Hendry. Disentangling interactions between adaptive divergence and gene flow when ecology drives diversification. *Ecology Letters*, 11(6):624–636, jun 2008.

[112] Michael Doebeli. *Adaptive diversification*. Monographs in population biology. Princeton University Press, Princeton, N.J, 2011.

[113] Clotilde Lepers, Sylvain Billiard, Matthieu Porte, Sylvie Méléard, and Viet Chi Tran. Inference with selection, varying population size, and evolving population structure: application of ABC to a forward–backward coalescent process with interactions. *Heredity*, 126(2):335–350, feb 2021.

[114] Jonathan M. Chase and Wade A. Ryberg. Connectivity, scale-dependence, and the productivity-diversity relationship. *Ecology Letters*, 7(8):676–683, aug 2004.

[115] DANY GARANT, SAMANTHA E. FORDE, and ANDREW P. HENDRY. The multifarious effects of dispersal and gene flow on contemporary adaptation. *Functional Ecology*, 21(3):434–443, jun 2007.

[116] Michael C. Gavin, Carlos A. Botero, Claire Bowern, Robert K. Colwell, Michael Dunn, Robert R. Dunn, Russell D. Gray, Kathryn R. Kirby, Joe McCarter, Adam Powell, Thiago F. Rangel, John R. Stepp, Michelle Trautwein, Jennifer L. Verdolin, and Gregor Yanega. Toward a Mechanistic Understanding of Linguistic Diversity. *BioScience*, 63(7):524–535, 2013.

[117] J Doyne Farmer and Duncan Foley. The economy needs agent-based modelling. *Nature*, 460(7256):685–686, aug 2009.

[118] Romualdo Pastor-Satorras, Alexei Vázquez, and Alessandro Vespignani. Dynamical and correlation properties of the internet. *Physical Review Letters*, 87(25):258701–1–258701–4, 2001.

[119] S.N. Dorogovtsev and J.F.F. Mendes. *Evolution of Networks*. Oxford University Press, Oxford, jan 2003.

[120] Seth Bromberger and other Contributors. JuliaGraphs/LightGraphs.jl. sep 2017.

[121] Christopher Rackauckas and Qing Nie. DifferentialEquations.jl – A Performant and Feature-Rich Ecosystem for Solving Differential Equations in Julia. *Journal of Open Research Software*, 5, 2017.

[122] Heiner Linke. Applications of Brownian Motion. 5114:199–213, aug 2015.

[123] M. Vidyasagar. Markov Processes. mar 1986.

Supplementary Methods

Formal description of the model

Mathematically, the model is a measure-valued point process [59], so that individuals are represented as dirac functions $\delta_{x_k^{(i)}}$, where $x_k^{(i)} \equiv (u_k, s_k) \in \mathcal{U} \times \mathcal{S}$ corresponds to the traits' value of an individual located on vertex V_i. Under this formalism, the population on V_i is represented as a sum of dirac functions $\nu^{(i)} = \sum_k^{N^{(i)}} \delta_{x_k^{(i)}}$, where $N^{(i)}$ is the local population size. It follows that $N^{(i)} = \int_{\mathcal{X}} \nu^{(i)}(dx)$ and one can equally express the neutral trait mean $\bar{u}^{(i)} = \frac{1}{N^{(i)}} \int_{\mathcal{X}} u\, \nu^{(i)}(dx) = \frac{1}{N^{(i)}} \sum_k^{N^{(i)}} u_k$ and the neutral trait variance $\mathrm{Var}_u(\nu^{(i)}) = \frac{1}{N^{(i)}} \int_{\mathcal{X}} u^2 \nu^{(i)}(dx) - \left[\bar{u}^{(i)}\right]^2$. The process is Markovian, and its dynamics after time t is described by the so called infinitesimal generator L that characterises the probabilistic dynamics of the process [122], given for all real valued functions ϕ by

$$
\begin{aligned}
L\phi(\nu^{(i)}) = & \sum_{k=1}^{N^{(i)}} b_i(1-\mu)(1-m)(\phi(\nu^{(i)} + \delta_{x_k^{(i)}}) - \phi(\nu^{(i)})) && \text{births w/o mutations, w/o migrations} \\
& + \sum_{k=1}^{N^{(i)}} b_i \mu (1-m) \int_{\mathcal{X}} (\phi(\nu^{(i)} + \delta_z) - \phi(\nu^{(i)})) \mathcal{M}(x_k^{(i)}, z) dz && \text{births w/ mutations, w/o migrations} \\
& + \sum_{k=1}^{N^{(i)}} \frac{N^{(i)}}{K} (\phi(\nu^{(i)} - \delta_{x_k^{(i)}})) - \phi(\nu^{(i)})) && \text{deaths} \\
& + \sum_{j \neq i} \frac{a_{i,j}}{d_j} \sum_{k=1}^{N^{(j)}} b_i \mu m \int_{\mathcal{X}} (\phi(\nu^{(j)} + \delta_{x_k^{(j)}}) - \phi(\nu^{(j)})) \mathcal{M}(x_k^{(j)}, z) dz && \text{migrations w/ mutations} \\
& + \sum_{j \neq i} \frac{a_{i,j}}{d_j} \sum_{k=1}^{N^{(j)}} b_i (1-\mu) m (\phi(\nu^{(j)} + \delta_{x_k^{(j)}}) - \phi(\nu^{(j)})). && \text{migrations w/o mutations}
\end{aligned}
\tag{S1}
$$

In the following sections we use the notation $\langle \nu^{(i)}, \phi \rangle = \int_{\mathcal{X}} \phi(x) \nu^{(i)}(dx)$.

Diversity partitioning

To characterise trait distributions, it is useful to introduce local α and metapopulation γ diversity, whose difference results in β diversity [78]. We define here α, β and γ within our modelling paradigm, which we use in the next section to show that the stochastic dynamics of the individual-based model (IBM) can be approximated deterministically under certain conditions.

Local α diversity observed on V_i is defined as the variance of the local trait distribution

$$\begin{aligned} A_t^{(i)} &= \frac{1}{N_t^{(i)}} \int_{\mathcal{X}} \left(x - \bar{x}_t^{(i)}\right)^2 \nu_t^{(i)}(dx) \\ &= \frac{1}{N_t^{(i)}} \sum_j^{N_t^{(i)}} \left[x - \bar{x}_t^{(i)}\right]^2 \\ &= \frac{1}{N_t^{(i)}} \sum_j^{N_t^{(i)}} \left[x_j^2 + \left(\bar{x}_t^{(i)}\right)^2 - 2x_j\bar{x}_t^{(i)}\right] \\ &= \frac{1}{N_t^{(i)}} \sum_j^{N_t^{(i)}} \left[x_j^2 + \left(\bar{x}_t^{(i)}\right)^2 - 2\left(\bar{x}_t^{(i)}\right)^2\right] \\ &= \frac{\langle \nu^{(i)}, \mathbf{1}^2\rangle}{N_t^{(i)}} - \left[\frac{\langle \nu^{(i)}, \mathbf{1}\rangle}{N_t^{(i)}}\right]^2. \end{aligned} \tag{S2}$$

Therefore, local α diversity corresponds in the model to the second moment of the local process minus the quadratic variations of the first moment, weighted by the local population size. We denote the mean α diversity as $A_t = \frac{1}{M}\sum_i^M A_t^{(i)}$ and we denote the expected mean α diversity as $\alpha_t = \mathbb{E}[A_t]$.

We define the observed β diversity as the variance of the local trait distribution

$$\begin{aligned} B_t &= \frac{1}{M}\sum_{i=1}^M \left[\bar{x}_t^{(i)} - \bar{x}_t\right]^2 \\ &= \frac{1}{M}\sum_{i=1}^M \left[\frac{\langle \nu_t^{(i)}, \mathbf{1}\rangle}{N_t^{(i)}}\right]^2 - \left[\frac{1}{M}\sum_{i=1}^M \frac{\langle \nu_t^{(i)}, \mathbf{1}\rangle}{N_t^{(i)}}\right]^2. \end{aligned} \tag{S3}$$

Therefore, β diversity is defined as the quadratic variations of the local processes minus the quadratic variation of the sum of the local processes, weighted by the local population size. It is also useful to

think of B_t as

$$\begin{aligned} B_t &= \frac{1}{2M^2}\sum_{i,j}^{M}\left[\bar{x}_t^{(i)} - \bar{x}_t^{(j)}\right]^2 \\ &= \frac{1}{2M^2}\sum_{i,j}^{M}\left[\left(\bar{x}_t^{(i)}\right)^2 + \left(\bar{x}_t^{(j)}\right)^2 - 2\bar{x}_t^{(i)}\bar{x}_t^{(j)}\right]. \end{aligned} \tag{S4}$$

This formulation reads that β diversity is the sum of the local variance less the co-variance of mean trait distribution between the demes. We further denote the expected β diversity as $\beta_t = \mathbb{E}[B_t]$.

Finally, we define γ diversity as the variance of the total trait distribution

$$\begin{aligned} \Gamma_t &= \frac{1}{M}\sum_{i}^{M}\frac{1}{N_t^{(i)}}\int_{\mathcal{X}}(x-\bar{x}_t)^2\,\nu_t^{(i)}(dx) \\ &= \frac{1}{M}\sum_{i}^{M}\frac{1}{N_t^{(i)}}\sum_{j}^{N_t^{(i)}}[x-\bar{x}_t]^2 \\ &= \frac{1}{M}\sum_{i}^{M}\frac{\langle\nu^{(i)},\mathbf{1}^2\rangle}{N_t^{(i)}} - \left[\frac{1}{M}\sum_{i=1}^{M}\frac{\langle\nu_t^{(i)},\mathbf{1}\rangle}{N_t^{(i)}}\right]^2, \end{aligned} \tag{S5}$$

where $\bar{x}_t = \frac{1}{M}\sum_i^M \bar{x}_t^{(i)}$and set $\gamma_t = \mathbb{E}[\Gamma_t]$.

Combining Eqs. (S2) and (S3), we obtain

$$\begin{aligned} A_t + B_t &= \frac{1}{M}\sum_{i=1}^{M}A_t^{(i)} + B_t \\ &= \Gamma_t. \end{aligned} \tag{S6}$$

Taking the expectation from both sides we obtain

$$\alpha_t + \beta_t = \gamma_t \tag{S7}$$

and therefore recover the classical additive diversity partitioning [78].

When does determinism emerge from stochasticity?

We demonstrate in a non-rigorous manner how deterministic dynamics can emerge at the macroscopic scale from the underlying stochastic dynamics when a stabilising force dampens the random fluctuations of the moments of the process $\nu^{(i)}$. Probability theory allows one to decompose $\nu^{(i)}$ as the solution of an integro-differential equation governed by the infinitesimal generator L and perturbed by a random variable describing the stochasticity of the process (a martingale, see [123, 58]). Given Eq. (S1), fixing $m = 0$ and further defining $\psi_t^{(i)} = \frac{1}{K}\mathbb{E}[\nu_t^{(i)}]$, Theorem 6.5 in [59] entails that for all real valued functions ϕ we have that

$$\begin{aligned}\langle \nu_t^{(i)}, \phi\rangle = K\Bigg[&\frac{1}{K}m_t^{(i)}(\phi) + \int_{\mathcal{X}} \phi(x)\psi_0^{(i)}(x)dx \\ &+ \int_0^t \int_{\mathcal{X}} \left\{\phi(x)\left(b_i(x)(1-\mu) - \langle \psi_s^{(i)}, 1\rangle\right) + b_i(x)\mu \int_{\mathcal{X}} \phi(x)\mathfrak{M}(x+z,z)dz\right\}\psi_s^{(i)}(x)dx\Bigg],\end{aligned} \tag{S8}$$

where the first term $m_t^{(i)}(\phi)$ is the random variable, the second term describes the initial conditions and the last term is similar to the infinitesimal generator in Eq. (S1), and corresponds to the deterministic part of the dynamics. For each t, $m_t^{(i)}(\phi)$ has mean zero and variance equal to

$$\left[m_t^{(i)}(\phi)\right]^2 = \int_0^t \int_{\mathcal{X}} \left\{\phi^2(x)\left(b_i(x)(1-\mu) + \langle \psi_s^{(i)}, 1\rangle\right) + b\mu \int_{\mathcal{X}} \phi^2(x)\mathfrak{M}(x+z,z)dz\right\}\psi_s^{(i)}(x)dx. \tag{S9}$$

Equation (S9) characterises the amplitude of the random fluctuations. From Eq. (S8), one can see that $m_t^{(i)}(\phi)$ is scaled by a factor $1/K$, and we discuss in the following two subsections the condition under which the random term becomes negligible in front of the deterministic term, depending on its amplitude $\left[m_t^{(i)}(\phi)\right]^2$.

Setting (1)

Recall that in setting (1) we define

$$b_i(x) \equiv b. \tag{S10}$$

One can obtain the stochastic fluctuations of the population size using Eq. (S9) with $\phi = 1$. Given that $\int_{\mathcal{X}} \mathfrak{M}(x,z)dz = 1$, we get

$$\begin{aligned}
[m_t(1)]^2 &= \int_0^t \int_{\mathcal{X}} \left\{ \left(b(1-\mu) + \langle \psi_s^{(i)}, 1 \rangle \right) + b\mu \int_{\mathcal{X}} \mathfrak{M}(x+z,z)dz \right\} \psi_s^{(i)}(x)dx \\
&= \int_0^t \langle \psi_s^{(i)}, 1 \rangle \left(b(1-\mu) + \langle \psi_s^{(i)}, 1 \rangle + b\mu \right) ds \\
&= \int_0^t \langle \psi_s^{(i)}, 1 \rangle \left(b + \langle \psi_s^{(i)}, 1 \rangle \right) ds.
\end{aligned} \tag{S11}$$

Taking the expectation from both sides of Eq. (S8) and deriving with respect to time, we obtain the deterministic trajectory of the process as

$$\begin{aligned}
\partial_t \langle \psi^{(i)}, \phi \rangle &= \partial_t \mathbb{E} \left[\langle \nu_t^{(i)}, \phi \rangle \right] \\
&= K \int_{\mathcal{X}} \left\{ \phi(x) \left(b_i(x)(1-\mu) - \langle \psi_t^{(i)}, 1 \rangle \right) + b\mu \int_{\mathcal{X}} \phi(x) \mathfrak{M}(x+z,z)dz \right\} \psi_t^{(i)}(x)dx.
\end{aligned} \tag{S12}$$

Using Eq. (S12) with $\phi = 1$ yields that at the quasi-equilibrium for which $\partial_t \langle \psi^{(i)}, 1 \rangle = 0$ one has $\langle \psi^{(i)}, 1 \rangle = b$, and we therefore get from Eq. (S11) that $[m_t(1)]^2 = \int_0^t 2b^2 ds$. This readily states that the quadratic variations of the population size are constant, and because $m_t^{(i)}(\phi)$ is scaled by $1/K$, it can be neglected in front of the deterministic term, provided that K is large enough. Figure S4 shows how the deterministic approximation given by Eq. (2.2) is valid for the population size in setting (1), and how this statement remains valid for $m > 0$.

One can repeat the same approach to quantify the amplitude of the stochastic fluctuations of the mean neutral trait, choosing $\phi = \mathbf{1}$ and using the fact that $\int_{\mathcal{X}} z^2 \mathfrak{M}(x,z)dz = \sigma_\mu^2$, for which we obtain

$$[m_t(\mathbf{1})]^2 = \int_0^t \left[\langle \psi_s^{(i)}, \mathbf{1}^2 \rangle \left(b(1-\mu) + \langle \psi_s^{(i)}, 1 \rangle \right) + b\mu\sigma^2 \langle \psi_s^{(i)}, 1 \rangle \right] ds. \tag{S13}$$

One can show from Eq. (S12) that at quasi-equilibrium the expected second moment of the density $\langle \psi_t^{(i)}, \mathbf{1}^2 \rangle$ grows in time at rate $\mu\sigma^2$, because no stabilising force maintains the evolution of the neutral trait. This means that the quadratic variations of the mean trait grow in time and therefore cannot be neglected when K is finite. Figure S2 shows how the IBM trajectory significantly differs from the deterministic trajectory, and Fig. S4 illustrates how the deterministic approximation fails at describing the diversity metrics in setting (1).

Setting (2)

Recall that in setting (2) we define for $x = (u, s) \in \mathbb{R}^2$

$$b_i(x) = b(1 - p(s - \theta_i)^2) \tag{S14}$$

By definition, neutral traits have no influence on the dynamics of the process, so that their dynamics can be left out to study the adaptive trait distribution and we use $\langle \psi^{(i)} \rangle$ in the following text to describe the expected adaptive trait distribution. In contrast to neutral differentiation, selection pressure acts as a stabilising force, which maintains the mean of the populations' adaptive trait to a fixed value. [74] shows that given Eq. (S14), Eq. (S12) accepts a steady state for all ϕ, and in particular one has $\langle \psi_s^{(i)}, \mathbf{1}^2 \rangle = \sqrt{\mu}\sigma_\mu^2/\sqrt{p}$. Therefore, the quadratic variations of the mean trait are constant and $m_t^{(i)}(\mathbf{1})$ can be neglected, provided that the selection strength p and the carrying capacity K are large enough. Figure S5 illustrates how the deterministic approximation succeeds at describing adaptive differentiation β_s in setting (2), even in the case where $m > 0$. Figure S3 shows how transient simulations support this claim and Fig. S5 illustrates that the deterministic approximation can also capture α_s and γ_s. One can therefore approximate the number of individuals with traits $s \in \Omega \subset \mathcal{S}$ with a continuous deterministic function $n_t^{(i)} = K\psi_t^{(i)}$ for Ω large enough, in the sense that $\int_\Omega n_t^{(i)}(x)dx = K \int_\Omega \psi_t^{(i)}(x)dx \approx \int_\Omega \nu_t^{(i)}(dx)$.

Simplification of the adaptive trait dynamics in the general case $m > 0$

Taking the expectation from both sides of Eq. (S1) one can show that $n_t^{(i)} = \mathbb{E}\left[\nu_t^{(i)}\right]$ satisfies

$$\begin{aligned}\partial_t n_t^{(i)}(s) = n_t^{(i)}(s)\left[b_i(s)(1-m)(1-\mu) - \frac{1}{K}\int_{\mathcal{S}} n_t^{(i)}(\mathbf{s})\, d\mathbf{s})\right] \\ + (1-m)\mu \int_{\mathcal{S}} b_i(\mathbf{s})\, n_t^{(i)}(\mathbf{s})\mathfrak{M}(s,\mathbf{s})\, d\mathbf{s} \\ + m\mu \sum_{j\neq i} \frac{a_{i,j}}{d_j}\int_{\mathbb{R}} b_j(\mathbf{s})\, n_t^{(j)}(s)\mathfrak{M}(s,\mathbf{s}) d\mathbf{s} \\ + m(1-\mu)\sum_{j\neq i}\frac{a_{i,j}}{d_j} b_j(s)\, n_t^{(j)}(s)\, d\mathbf{s}.\end{aligned} \tag{S15}$$

Assuming that the variance of the mutation kernel is small, one can use a diffusion approximation for the mutation term [82, 72, 74]

$$\int_{\mathcal{S}} b_i(\mathbf{s})\, n_t^{(i)}(\mathbf{s})\mathfrak{M}(s,\mathbf{s})\, d\mathbf{s} = b_i(s,t)\, n_t^{(i)}(s) + \tfrac{1}{2}\sigma_\mu^2 \Delta_s(b_i n_t^{(i)})(s). \tag{S16}$$

Neglecting the terms in $m\mu$, we obtain

$$\begin{aligned}\partial_t n_t^{(i)}(s) = n_t^{(i)}(s)\left[b_i(s,t)(1-m-\mu) - \frac{1}{K}\int_{\mathcal{S}} n_t^{(i)}(\mathbf{s})\, d\mathbf{s})\right] \\ + \mu\left[b_i(s,t)\, n_t^{(i)}(s) + \tfrac{1}{2}\sigma_\mu^2\Delta_s(b_i n_t^{(i)})(s)\right] \\ + m\sum_{j\neq i} b_j(s,t) n_t^{(j)}(s) a_{i,j},\end{aligned} \tag{S17}$$

which after rearranging terms yields the elegant deterministic description of the process

$$\partial_t n_t^{(i)}(s) = n_t^{(i)}(s)\left[b_i(s)(1-m) - \frac{1}{K}\int_{\mathcal{S}} n_t^{(i)}(\mathbf{s}) d\mathbf{s}\right] + m\sum_{j\neq i} b_j(s)\frac{a_{i,j}}{d_j} n_t^{(j)}(s) + \tfrac{1}{2}\mu\sigma_\mu^2\Delta_s\left[b_i(s) n_t^{(i)}(s)\right]. \tag{S18}$$

Derivation of the habitat assortativity metric r_θ in binary environments

We demonstrate here how the habitat assortativity r_θ relates to the conditional probability of habitats being connected, and show how r_θ simplifies under mean field assumption.

Following the original definition of [84], habitat assortativity r_θ is defined as the Pearson correlation of environmental conditions θ at either ends of the vertices V of graph G, that is

$$r_\theta = \frac{\mathrm{Cov}(\boldsymbol{\theta}_\times, \boldsymbol{\theta}_\wedge)}{\sqrt{\mathrm{Var}(\boldsymbol{\theta}_\times)\mathrm{Var}(\boldsymbol{\theta}_\wedge)}} = \frac{\langle \boldsymbol{\theta}_\times \boldsymbol{\theta}_\wedge \rangle - \langle \boldsymbol{\theta}_\times \rangle \langle \boldsymbol{\theta}_\wedge \rangle}{\sqrt{(\langle \boldsymbol{\theta}_\times^2 \rangle - \langle \boldsymbol{\theta}_\times \rangle^2)(\langle \boldsymbol{\theta}_\wedge^2 \rangle - \langle \boldsymbol{\theta}_\wedge \rangle^2}} \tag{S19}$$

where $\boldsymbol{\theta}_\times, \boldsymbol{\theta}_\wedge$ denote the sets of environmental conditions found at the toe and tip of each directed vertex of graph V, and $\langle \boldsymbol{\theta}_\times \rangle, \langle \boldsymbol{\theta}_\wedge \rangle$ denote their respective mean values.

Let $P(\bullet, \bullet)$ be the proportion of edges that connect a node with habitat $\theta_\bullet$ to a node with habitat $\theta_\bullet$. One can also view $P(\bullet, \bullet)$ as the conditional probability that a node of type $\theta_\bullet$ is connected to a node of type $\theta_\bullet$. Denote by $P(\bullet)$ the proportion of nodes that are of type $\theta_\bullet$. First observe that for undirected graphs, one has $\langle \boldsymbol{\theta}_\times \rangle = \langle \boldsymbol{\theta}_\wedge \rangle$, and $\langle \boldsymbol{\theta}_\times^2 \rangle = \langle \boldsymbol{\theta}_\wedge^2 \rangle$. Assuming that habitats are symmetric and binary, one has that $\theta_\bullet = -\theta_\bullet$. Then

$$\begin{aligned} \langle \boldsymbol{\theta}_\times \boldsymbol{\theta}_\wedge \rangle &= P(\bullet, \bullet)\theta_\bullet^2 + P(\bullet, \bullet)\theta_\bullet^2 + [P(\bullet, \bullet) + P(\bullet, \bullet)]\theta_\bullet \theta_\bullet \\ &= \theta_\bullet^2 \left(P(\bullet, \bullet) + P(\bullet, \bullet) - [P(\bullet, \bullet) + P(\bullet, \bullet)] \right), \end{aligned} \tag{S20}$$

$$\begin{aligned} \langle \boldsymbol{\theta}_\times \rangle &= P(\bullet)\theta_\bullet + P(\bullet)\theta_\bullet \\ &= \theta_\bullet \left[P(\bullet) - P(\bullet) \right], \end{aligned} \tag{S21}$$

$$\begin{aligned} \langle \boldsymbol{\theta}_\times^2 \rangle &= P(\bullet)\theta_\bullet^2 + P(\bullet)\theta_\bullet^2 \\ &= \theta_\bullet^2 \left[P(\bullet) + P(\bullet) \right] \\ &= \theta_\bullet^2. \end{aligned} \tag{S22}$$

Combining Eq. (S20), Eq. (S21), and Eq. (S22) with Eq. (S19) one gets

$$\begin{aligned} r_\theta &= \frac{\langle \boldsymbol{\theta}_\times \boldsymbol{\theta}_\wedge \rangle - \langle \boldsymbol{\theta}_\times \rangle \langle \boldsymbol{\theta}_\wedge \rangle}{\langle \boldsymbol{\theta}_\times^2 \rangle - \langle \boldsymbol{\theta}_\times \rangle^2} \\ &= \frac{P(\bullet,\bullet) + P(\bullet,\bullet) - [P(\bullet,\bullet) + P(\bullet,\bullet)] - (P(\bullet) - P(\bullet))^2}{P(\bullet) + P(\bullet) - (P(\bullet) - P(\bullet))^2} \\ &= \frac{P(\bullet,\bullet) + P(\bullet,\bullet) - [P(\bullet,\bullet) + P(\bullet,\bullet)] - (P(\bullet) - P(\bullet))^2}{1 - (P(\bullet) - P(\bullet))^2}. \end{aligned} \tag{S23}$$

Assuming that habitats are homogeneously distributed, we have $P(\bullet) = P(\bullet) = \frac{1}{2}$ and thus we obtain

$$r_\theta = P(\bullet,\bullet) + P(\bullet,\bullet) - [P(\bullet,\bullet) + P(\bullet,\bullet)]. \tag{S24}$$

The mean field approximation involves the assumption that all vertices with similar habitats are equivalent in terms of their connections with other habitats, so that $P(\bullet,\bullet) = P(\bullet,\bullet)$ and $P(\bullet,\bullet) = P(\bullet,\bullet)$, which yields $r_\theta = 2\,(P(\bullet,\bullet) - P(\bullet,\bullet))$.

Supplementary Tables

Table S1: Correlations between β_u and topology metrics for setting (1), ranked by absolute value.

metrics	$m = 0.01$	$m = 0.05$	$m = 0.10$	$m = 0.50$	cor. mean	cor. std.
var. betweenness centrality	0.83	0.90	0.90	0.88	0.88	0.03
mean edge betweenness centrality	0.75	0.82	0.84	0.87	0.82	0.05
characteristic length	0.71	0.80	0.83	0.89	0.81	0.07
var. eigenvector centrality	0.72	0.81	0.81	0.82	0.79	0.05
mean eigenvector centrality	−0.71	−0.81	−0.81	−0.82	−0.79	0.05
mean closeness centrality	−0.69	−0.76	−0.78	−0.83	−0.76	0.05
edges density	−0.75	−0.77	−0.76	−0.76	−0.76	0.01
$\langle\sqrt{k}\rangle^2/\langle k\rangle$	−0.82	−0.81	−0.77	−0.62	−0.75	0.09
algebraic connectivity	−0.70	−0.73	−0.73	−0.75	−0.73	0.02
var. edge betweenness centrality	0.57	0.71	0.76	0.87	0.73	0.12
graph energy	−0.69	−0.68	−0.67	−0.60	−0.66	0.04
diameter	0.53	0.63	0.67	0.75	0.65	0.09
mean deg. distrib.	−0.53	−0.61	−0.64	−0.69	−0.62	0.07
s max	−0.73	−0.67	−0.61	−0.45	−0.61	0.12
loop- 3	−0.53	−0.51	−0.50	−0.51	−0.51	0.01
mean neighb. deg.	−0.42	−0.48	−0.52	−0.61	−0.51	0.08
var. rich club	0.54	0.52	0.52	0.45	0.51	0.04
var. local clustering coef.	0.49	0.49	0.47	0.46	0.48	0.01
var. neighb deg	0.59	0.49	0.43	0.29	0.45	0.13
var. closeness centrality	0.46	0.47	0.43	0.32	0.42	0.07
var. deg. distrib.	−0.38	−0.37	−0.35	−0.37	−0.37	0.01
$\frac{1}{N}$	0.11	0.26	0.53	0.50	0.35	0.20
mean local clustering coefficient	−0.35	−0.34	−0.36	−0.34	−0.35	0.01
deg. var.	0.43	0.37	0.31	0.16	0.32	0.11
modules	0.29	0.32	0.32	0.31	0.31	0.01
clique4	−0.32	−0.30	−0.28	−0.29	−0.30	0.02
mean rich club	−0.24	−0.22	−0.22	−0.30	−0.25	0.04
deg. correlation	−0.31	−0.20	−0.14	0.01	−0.16	0.13
s elasticity	0.21	0.13	0.07	−0.06	0.09	0.11

Table S2: Correlations between β_s and topology metrics for setting (2), ranked by absolute value.

metrics	$m = 0.10$	$m = 0.31$	cor. mean	cor. std
r_θ	0.94	0.91	0.93	0.02
var. betweenness centrality	0.19	0.33	0.26	0.10
$\langle\sqrt{k}\rangle^2/\langle k\rangle$	0.17	0.29	0.23	0.09
s max	−0.19	−0.26	−0.22	0.05
algebraic connectivity	−0.14	−0.29	−0.22	0.11
var. eigenvector centrality	0.14	0.29	0.22	0.11
mean eigenvector centrality	−0.14	−0.29	−0.21	0.11
density	−0.13	−0.28	−0.20	0.11
mean edge betweenness centrality	0.12	0.28	0.20	0.11
characteristic length	0.12	0.28	0.20	0.11
var. edge betweenness centrality	0.13	0.26	0.20	0.09
mean closeness centrality	−0.11	−0.27	−0.19	0.11
var. neighb. deg	0.17	0.21	0.19	0.02
var. local clustering coef	0.13	0.22	0.17	0.07
var. rich club	0.13	0.20	0.17	0.05
graph energy	−0.09	−0.21	−0.15	0.08
var. closeness centrality	0.12	0.17	0.14	0.04
diameter	0.07	0.21	0.14	0.10
mean deg. distrib.	−0.07	−0.21	−0.14	0.10
var. degree	0.12	0.15	0.14	0.02
loop 3	−0.07	−0.18	−0.13	0.08
var. deg. distrib.	−0.08	−0.16	−0.12	0.05
mean neighb. deg.	−0.04	−0.17	−0.10	0.09
deg. correlation	−0.10	−0.09	−0.09	0.01
s elasticity	0.10	0.07	0.09	0.02
clique 4	−0.05	−0.11	−0.08	0.05
mean rich club	−0.03	−0.08	−0.05	0.03
modules	−0.00	0.07	0.03	0.05
mean local clustering coefficient	0.01	−0.07	−0.03	0.06

Table S3: Linear regression model coefficients for the effect of topology metrics on β_u, setting (1). *** $P < 0.001$

	$m = 0.01$			$m = 0.05$			$m = 0.1$			$m = 0.5$		
	(1)	(2)	(3)	(1)	(2)	(3)	(1)	(2)	(3)	(1)	(2)	(3)
(Intercept)	-0.000	-0.000	-0.000	-0.000	-0.000	-0.000	-0.000	-0.000	-0.000	-0.000	-0.000	-0.000
	(0.015)	(0.014)	(0.018)	(0.012)	(0.012)	(0.012)	(0.011)	(0.011)	(0.013)	(0.013)	(0.015)	(0.011)
characteristic length	0.432***		0.218***	0.609***		0.362***	0.549***		0.304***	0.773***		0.537***
	(0.017)		(0.027)	(0.013)		(0.018)	(0.013)		(0.020)	(0.015)		(0.016)
$\langle\sqrt{k}\rangle^2/\langle k\rangle$	0.619***	0.589***		0.488***	0.462***		0.561***	0.535***		0.263***	0.246***	
	(0.017)	(0.016)		(0.013)	(0.014)		(0.013)	(0.013)		(0.015)	(0.017)	
mean edge betw. centr.		0.464***			0.620***			0.564***			0.754***	
		(0.016)			(0.014)			(0.013)			(0.017)	
var. betw. centr.			0.670***			0.633***			0.675***			0.481***
			(0.027)			(0.018)			(0.020)			(0.016)
R^2	0.813	0.828	0.711	0.879	0.876	0.871	0.894	0.898	0.851	0.850	0.809	0.900

Table S4: Linear regression model coefficients for the effect of topology metrics on β_s, setting (2). *** $P < 0.001$

	$m = 0.1$			$m = 0.31$		
	(1)	(2)	(3)	(1)	(2)	(3)
(Intercept)	-0.000	-0.000	-0.000	0.000	0.000	0.000
	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.007)
$\langle\sqrt{k}\rangle^2/\langle k\rangle$	0.170***			0.293***		
	(0.006)			(0.006)		
r_θ	0.944***	0.935***	0.942***	0.910***	0.893***	0.907***
	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.007)
var. betw. centrality		0.137***			0.275***	
		(0.006)			(0.006)	
smax			-0.180***			-0.247***
			(0.006)			(0.007)
R^2	0.920	0.909	0.923	0.913	0.902	0.888

Table S5: Linear regression model coefficients for the effect of topology metrics on β_u, setting (2). *** $P < 0.001$

m	0.1	0.31
(Intercept)	-0.000	0.000
	(0.008)	(0.007)
characteristic length	0.471***	0.427***
	(0.009)	(0.008)
r_θ	-0.748***	0.655***
	(0.008)	(0.007)
$\langle\sqrt{k}\rangle^2/\langle k\rangle$	0.149***	0.322***
	(0.009)	(0.008)
R^2	0.832	0.867

Supplementary Figures

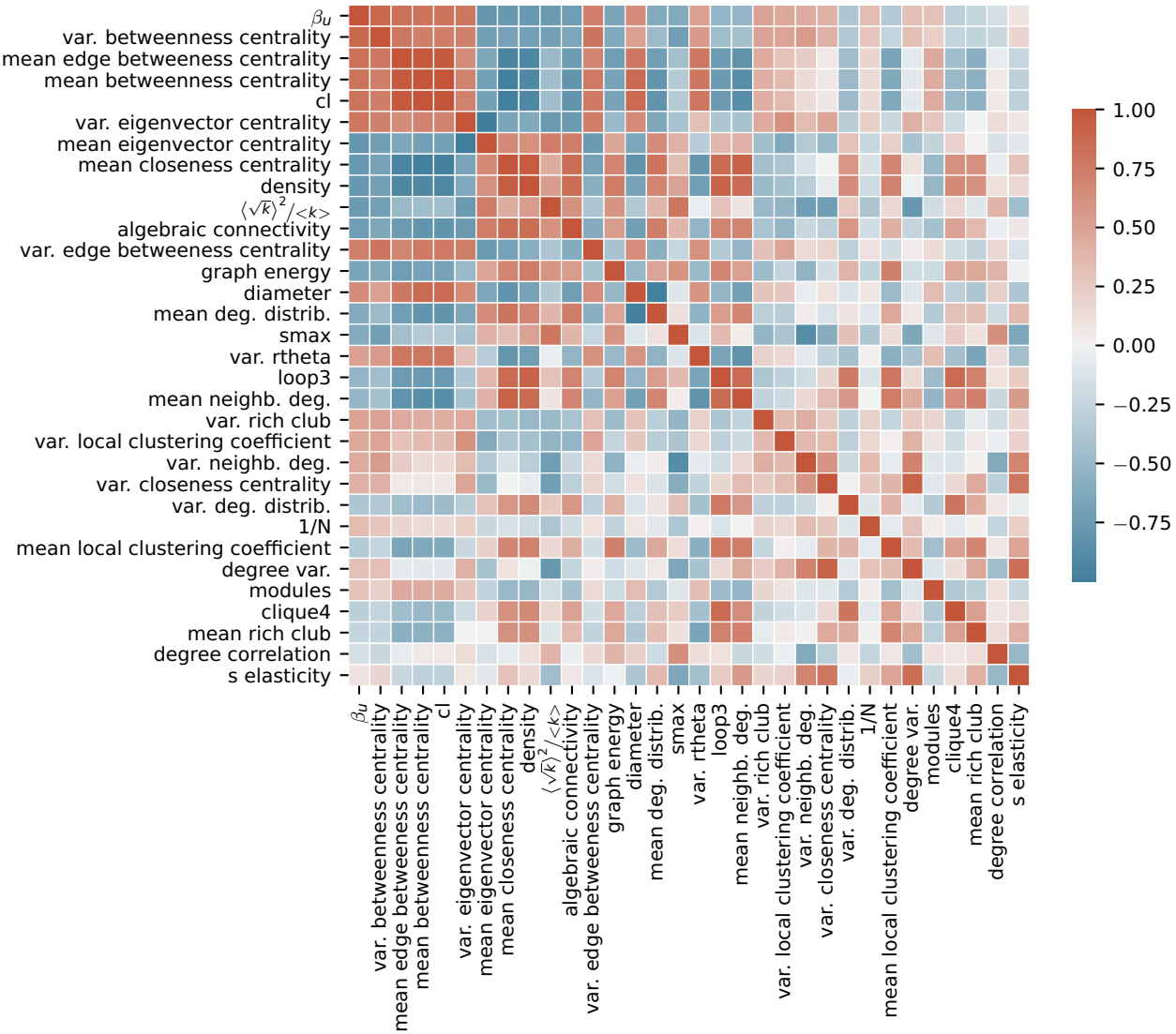

Figure S1: Pearson correlation matrix between β_u and topology metrics for setting (1). On top of a graphical illustration of the correlations provided in Table S1 (column "cor. mean"), the figure shows how topology metrics correlate between each other. Rows are ranked by the absolute value of the correlation between the topological metrics and β_u.

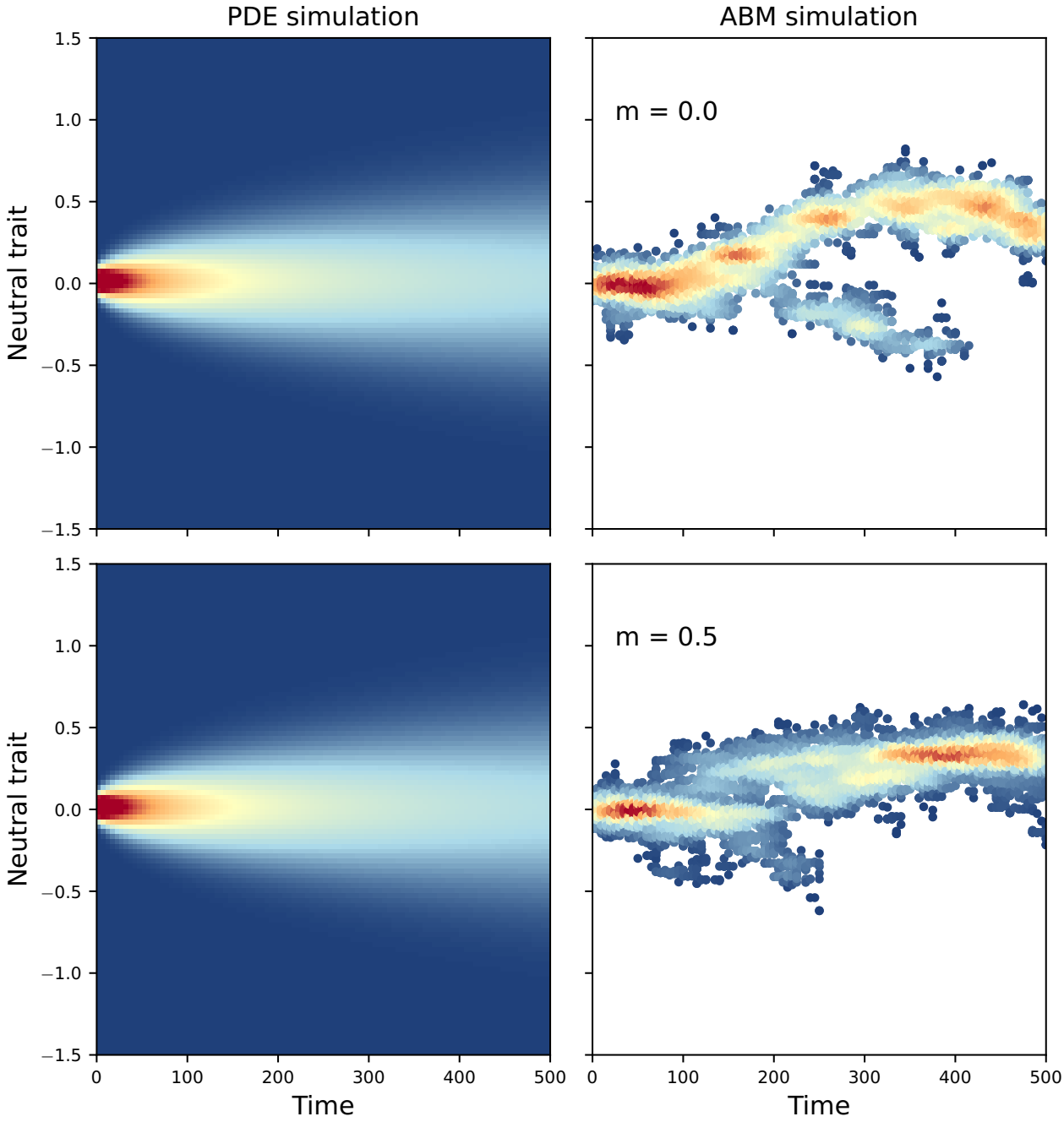

Figure S2: Comparison of the neutral trait density on one vertex obtained from Eq. (S18) with birth rate Eq. (S10) (left) and IBM (right) simulations in setting (1), for the chain graph. While the PDE model gives a density centred at 0 for $m = 0$, drift arises in the IBM simulation. As m grows (lower panels), correlation sets in the local trait pool and the density stabilises to a mean value centred at 0. $K_i = 150, \mu = 0.1, \sigma_\mu = 0.05$.

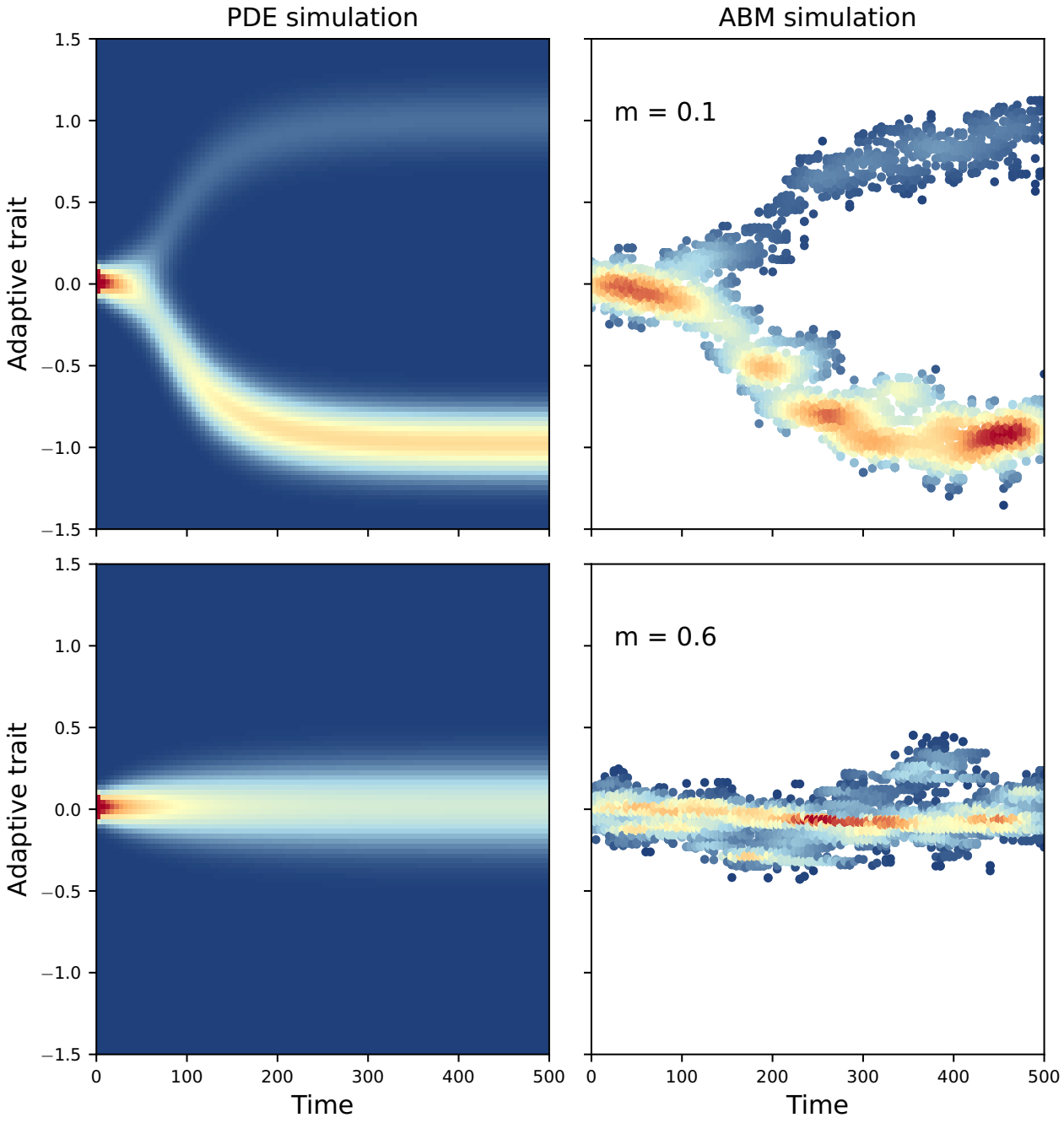

Figure S3: Comparison of the adaptive trait density on one vertex obtained from Eq. (S18) with birth rate Eq. (S14) (left) and IBM (right) simulations in setting (2), for the star graph. The densities obtained from the IBM and the PDE model look qualitatively similar. $K_i = 150, \mu = 0.1, \sigma_\mu = 0.05, p = 0.5$.

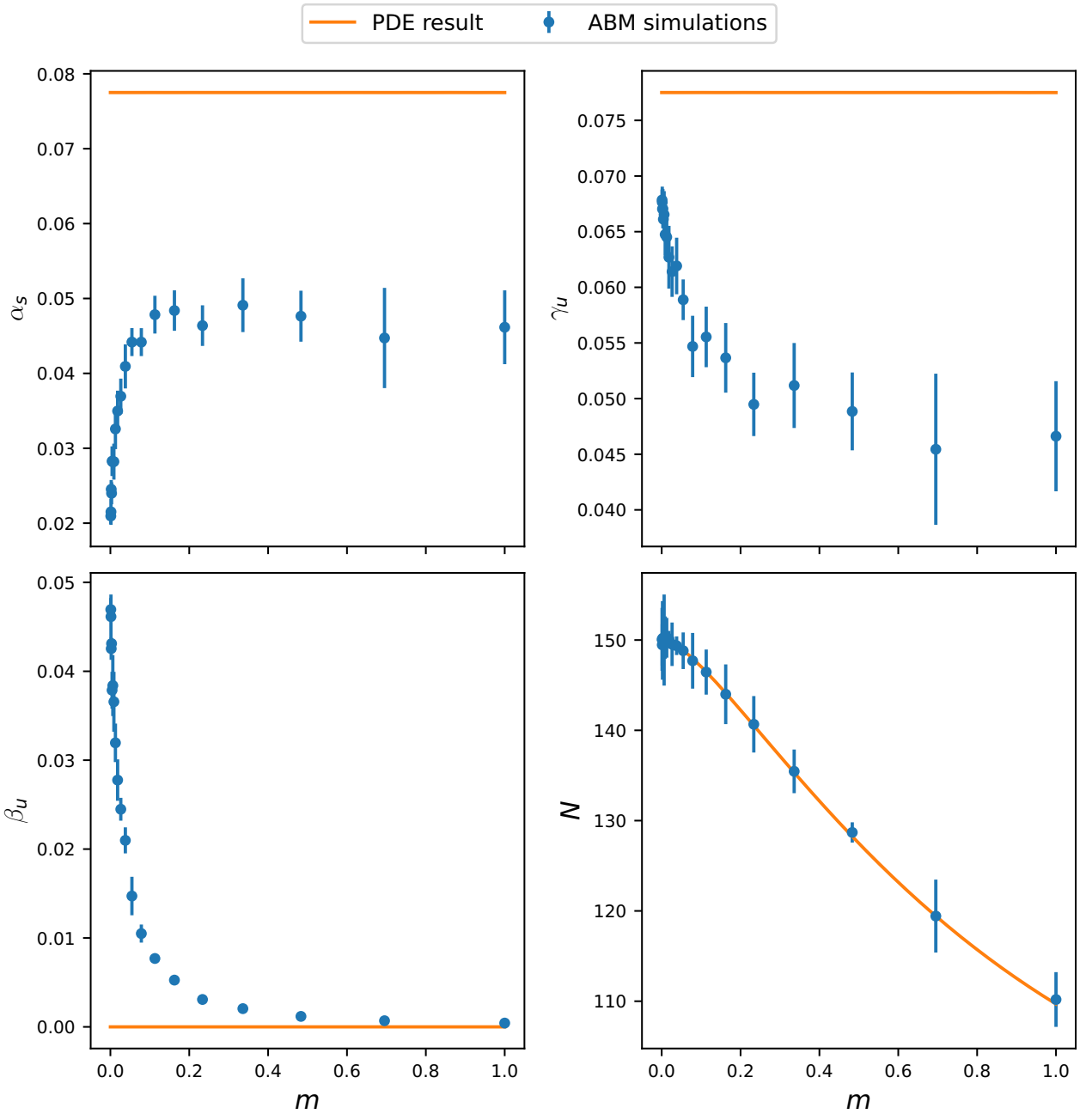

Figure S4: Comparison of neutral diversity metrics and population size obtained from Eq. (S18) with birth rate Eq. (S10) and IBM simulations in setting (1), on a star graph. The figure illustrates that while a deterministic approximation can capture population size, it is not able to capture α_u, β_u and γ_u diversity.

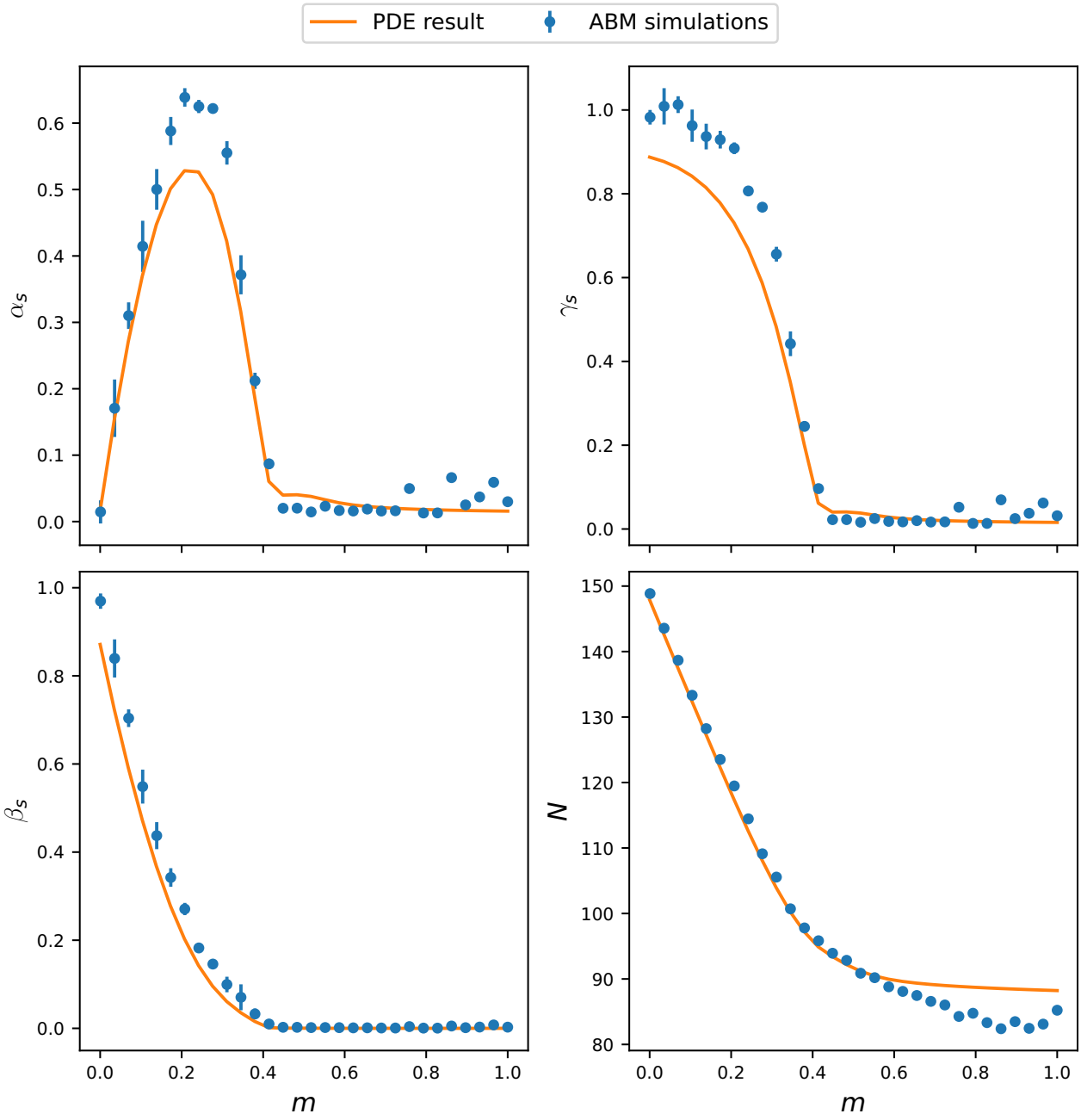

Figure S5: Comparison of adaptive diversity metrics and population size obtained from Eq. (S18) with birth rate Eq. (S14) and IBM simulations in setting (2), on a star graph. The figure illustrates that the metrics α_s, β_s, γ_s and population size obtained from the PDE closely match those from the IBM simulations.

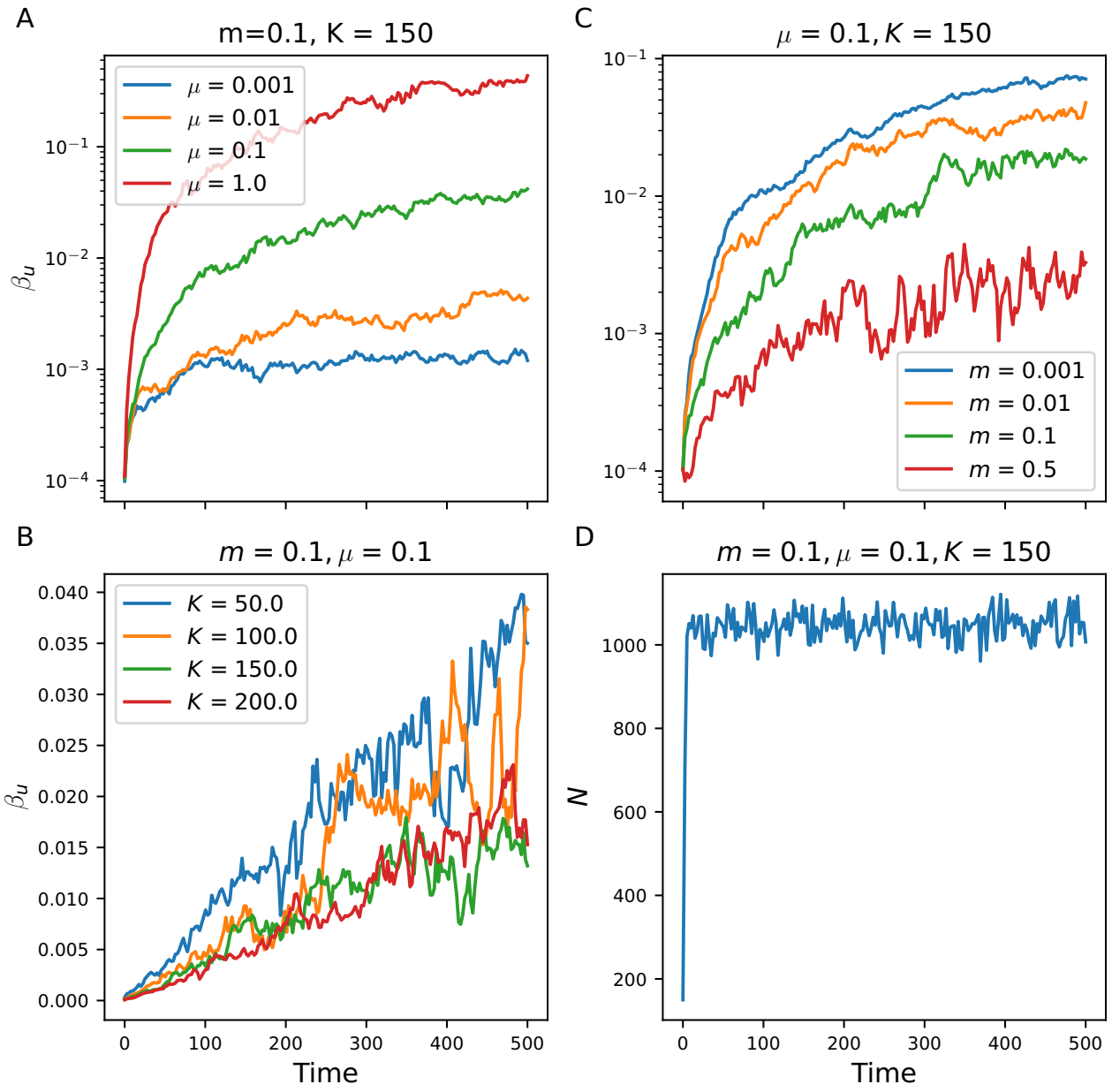

Figure S6: Transient dynamics of β_u and population size in setting (1), for different mutation rates μ, migration rates m, and carrying capacities K. (A) Transient evolution of β_u for varying μ. (A) shows that β_u increases with μ. (B) Transient evolution of β_u for varying K. (B) shows that β_u increases linearly with time at quasi-equilibrium, and illustrates that decreasing K increases β_u, as it favours drift. (C) Transient evolution of β_u for varying m. (C) shows that β_u decreases with increasing migration rate. (D) Transient evolution of population size. (D) shows that stochastic fluctuations in population size are negligible.

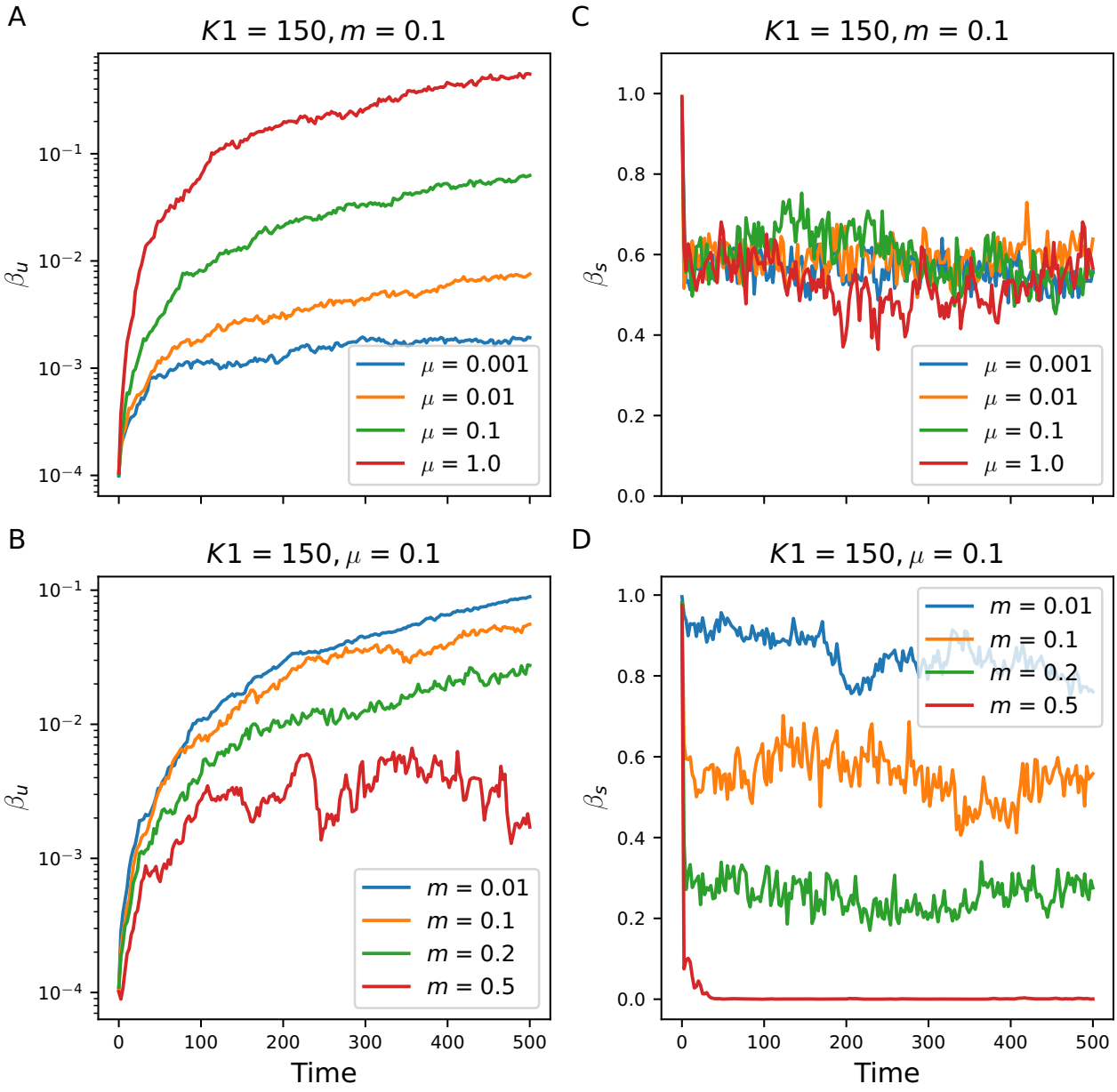

Figure S7: Transient dynamics for β_u and β_s in setting (2), for different mutation rates μ and migration rates m. (A) and (B) are similar to (A) and (C) in Fig. S6. (C) shows that β_s is not sensitive to μ. (D) shows that β_s decreases with increasing m. The drop in β_s for $r_\theta = -1 and m = 0.5$ corresponds to the case where the population becomes monomorphic.

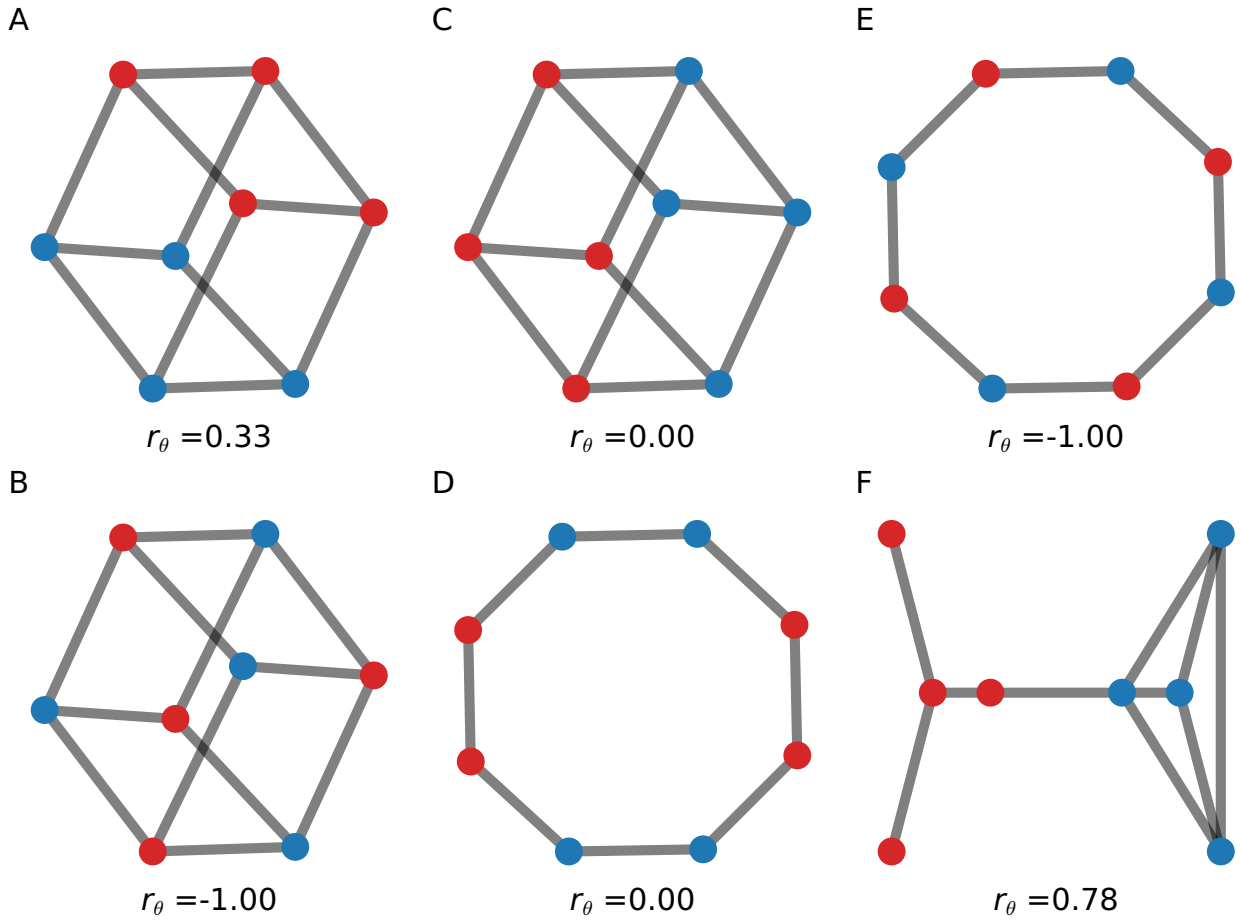

Figure S8: Graphs with different values for habitat assortativity r_θ. Graphs (A–D) can be described exactly with a mean field approach, as blue and red vertices have the equivalent position on the graph.

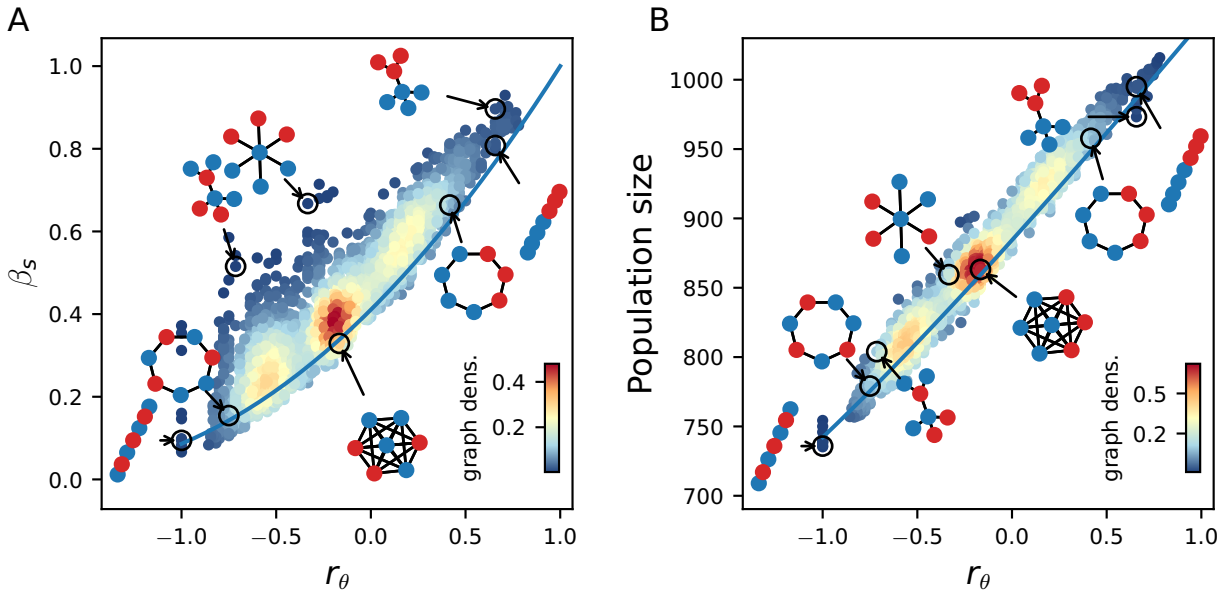

Figure S9: Effect of habitat heterogeneity r_θ on β_s and population size from simulations of the IBM. (A) and (B) show results for simulations on all undirected connected graphs with seven vertices and varying r_θ for $m = 0.31$.

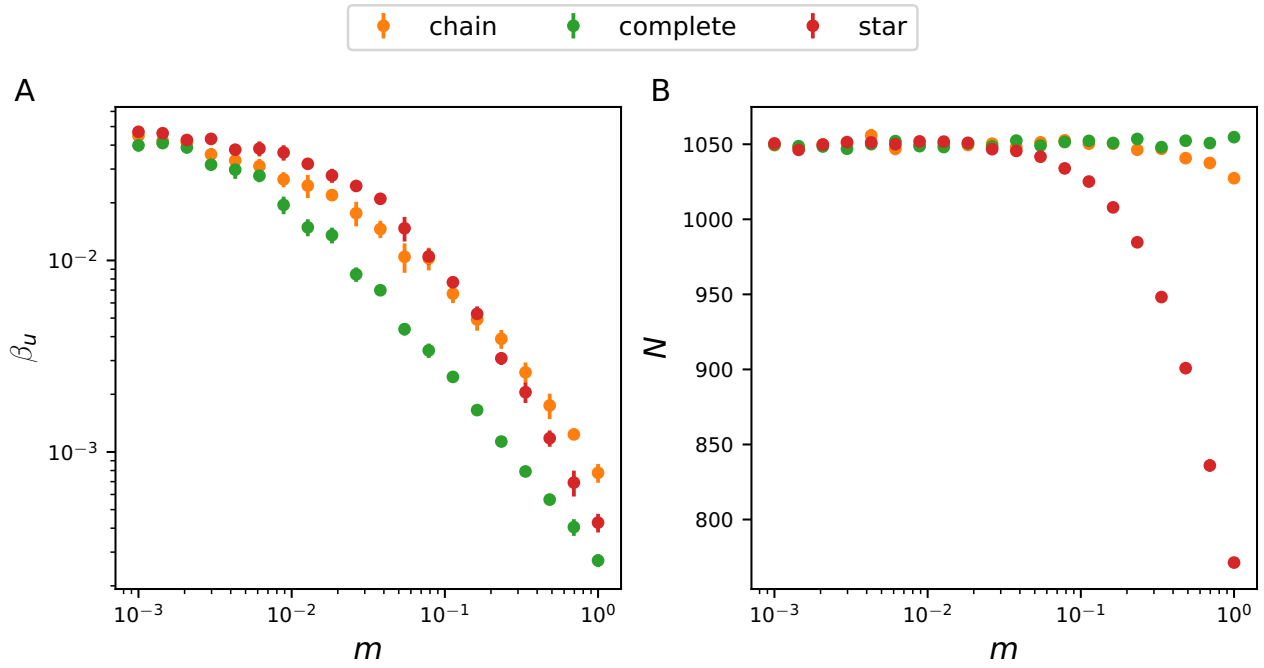

Figure S10: Response of β_u and population size to migration m in setting (1), for the chain graph, the complete graph and the star graph. (A) shows that β_u is higher in the star graph for low m, but becomes lower than in the chain graph under high migration regimes. (A) therefore suggests a dependency on m for the effect of topology metrics on β_u. This dependency appears in Table S1 and Table S3. (B) illustrates that the star graph experiences a large demographic loss under high migration regimes compared with the chain and complete graph.

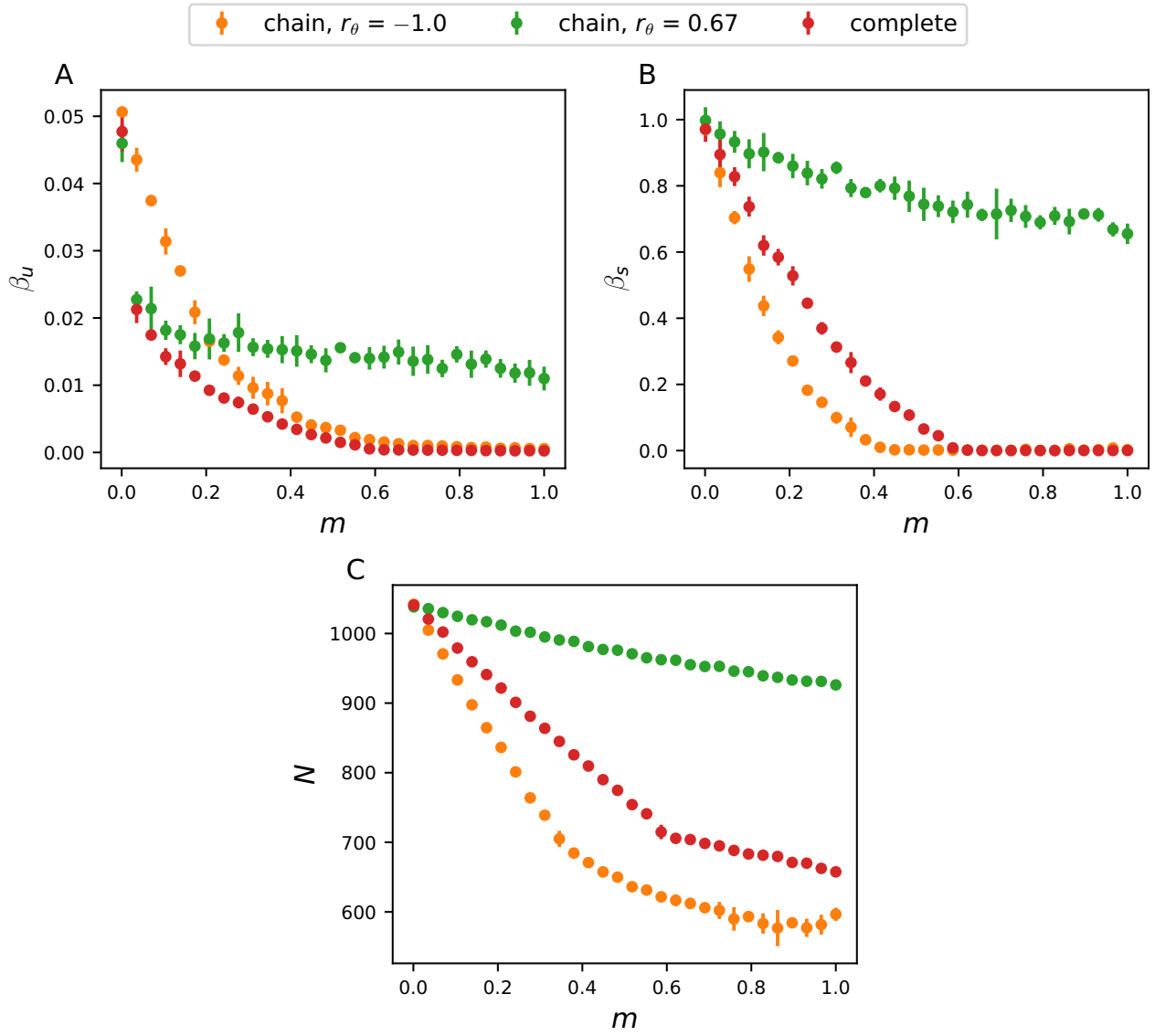

Figure S11: Response of β_u, β_s and population size to migration m in setting (2), for the chain graph with varying r_θ and for the complete graph ($r_\theta \approx 0$). (A) shows that β_u is higher in the disassortative chain graph for $m < 0.2$, but becomes lower than in the assortative chain graph for $m > 0.2$. (B–C) show that r_θ systematically amplifies β_s and population size irrespective of the migration regime, because populations are better adapted.