Species dispersal mediates opposing influences of a branching network on genetic variation in a metapopulation

Ming-Chih Chiu[1], Bin Li[1], Kei Nukazawa[2], Thaddeus Carvajal[1] and Kozo Watanabe[1,*]

[1]Department of Civil and Environmental Engineering, Ehime University, Ehime Prefecture 790-8577, Japan

[2]Department of Civil and Environmental Engineering, University of Miyazaki, Miyazaki Prefecture 889-2192, Japan

[*]Corresponding Author:

E-mail address: watanabe_kozo@cee.ehime-u.ac.jp

Running Head: Genetic effects of river branching

Abstract

In nature, ubiquitous fractal networks can have two but opposing influences, by increasing distal and confluent habitats, respectively, under raising branching complexity on metapopulations' genetic structure, although this remains poorly understood, particularly regarding the roles of species-specific traits. In this study, we evaluated the integrated influences of network complexity and species dispersal mode/ability on genetic divergence among populations at the catchment scale, using a theoretical framework with empirical genetic data from four sympatric stream macroinvertebrate species. Empirical patterns of spatial genetic structure were attributed to dispersal ability and the species' habitat specialisation levels. Our theoretical evidence showed that both greater landscape connectivity (via shorter watercourse distance) and greater isolation of distal habitats (e.g. headwater streams) occur in the more-branched networks. These two spatial features have negative and positve influences on genetic divergence, respectively, with their relative importance varying in different species. Watersheds harbouring a higher number of local populations have larger genetic divergence of metapopulations. Downstream- and upstream-biased asymmetric dispersals dictate increases and declines, respectively, in genetic divergence. In addition, distal populations (e.g. in headwaters) have higher genetic independence between themselves under higher levels of downstream-biased asymmetry. A strong association between species features and evolutionary processes (gene flow and genetic drift) mediates the pervasive influences of

branching complexity on metapopulation genetic divergence, which highlights the importance of considering species dispersal patterns when developing management strategies in rapid environmental change scenarios.

Keywords: distribution, fractal geometry, habitat fragmentation, isolation by distance, landscape complexity, macroinvertebrate

Introduction

There is growing interest in understanding how landscape architecture determines ecosystems' spatial biodiversity (Economo & Keitt, 2008; Albert *et al.*, 2013; Wilson *et al.*, 2016). Despite comprehensive findings about spatial biodiversity, revealed by substantial empirical and theoretical evidence (Chave, 2013), there is less information on spatial patterns of intraspecific genetic diversity (Paz-Vinas *et al.*, 2015). Eco-evolutionary evidence and theories derived from simplified landscapes are insufficient for understanding spatial genetic patterns in complex systems such as rivers (Campbell Grant *et al.*, 2007; Thomaz *et al.*, 2016; Terui *et al.*, 2018). Further explorations of the integrated genetic effects of species dispersal and landscape connectivity on metapopulations (here defined as groups of subpopulations with dispersal interactions) in complex habitats are needed.

In nature, ubiquitous fractal branching networks (e.g. with treelike patterns) have similar structural features (Green, 2006), and species dispersal can mediate landscape genetic structures (Paz-Vinas *et al.*, 2015; Thomaz *et al.*, 2016). Landscape connectivity shapes evolutionary processes, such as gene flow and genetic drift, driving spatial patterns of intraspecific genetic variation (McRae, 2006; Paz-Vinas *et al.*, 2015). Dendritic ecological networks (e.g. riverscape structures) constrain species dispersal (Grant *et al.* 2007). For example, ocean circulation patterns across seascapes shape their network connectivity and intraspecific spatial genetic patterns (Braunisch *et al.*, 2010; Ruiz-Gonzalez *et al.*, 2015; Chust

et al., 2016). The resulting spatial patterns within landscape networks are particularly pronounced in species with low dispersal ability; for example, the genetic structures of sea cucumbers (*Parastichopus californicus*) can be well explained by ocean circulation, which mediates larvae dispersal (Xuereb *et al.*, 2018). Branching networks can be characterised by distal and confluent habitats with fewer and more corridor linkages, respectively, and the two types of habitat have positive and negative influences on genetic divergence among local populations. Dendritic riverscape systems provide an excellent opportunity to reveal the roles of species dispersal in opposing influences of branching fractals and resulting consequences (i.e., either increasing or decreasing genetic divergence) based on their landscape spatial configuration (described below in detail).

Dispersal asymmetry (the situation in which dispersal tendency between two habitats is not necessarily equal to the tendency in the opposite direction) can dictate the isolation processes between pairs of populations, which provides mechanisms behind widely acknowledged patterns of spatial genetic diversity and differentiation (Kawecki & Holt, 2002). In river and stream systems, species dispersal ability and distribution pattern mediate their spatial genetic patterns (Pilger *et al.*, 2017). At all dispersal asymmetry levels, streamflow-connected populations have habitat connectivity based on gene flows predominantly in one direction or in both directions along a stream. Theoretically, more of these isolated tributaries within a network, under high river conditions, result in higher genetic

differentiation between local populations (Thomaz *et al.*, 2016). For example, populations in a river network's distal branches (e.g. different headwaters) are connected to a common source population in downstream confluences. Therefore, downstream-biased dispersal (a tendency for higher dispersal downstream than upstream) may lead to weak connections among headwaters and a large genetic divergence among riverine species such as fish and macroinvertebrates (Paz-Vinas *et al.*, 2013; Paz-Vinas & Blanchet, 2015).

In contrast, river branching can help enhance connectivity levels between demes by naturally increasing the number of confluences and shortening their watercourse distances (Labonne *et al.*, 2008). Stream-dwelling species with a strong tendency to migrate upstream, such as aquatic insects that disperse by flying during their terrestrial adult stages (Petersen *et al.*, 2004; Winterbourn *et al.*, 2007), can have low downstream-biased asymmetries or even upstream-biased gene flow. In this case, there is weaker isolation between distal populations in the river network when these sink populations receive higher gene flows from their shared source population at downstream confluences.

In this study, we evaluated the combined influences of landscape network and species dispersal on genetic divergences in ubiquitous fractal branching networks, which remain poorly understood. To the best of our knowledge, this is the first study to address how dispersal asymmetry mediates the countervailing influence of network branching with empirical genetic data, by which genetic divergence can potentially increase or decrease within natural

populations depending on species dispersal. First, we explored the spatial genetic variation of four macroinvertebrate species with flying adult stages in a shared river network, using a mechanistic model, based on evolutionary processes and asymmetric dispersal in northeastern Japan. These species all have substantially diverse habitat specificities and distributions within the network (Watanabe *et al*., 2014; Nukazawa *et al.*, 2015; Nukazawa *et al.*, 2017). Second, with the model empirically validated through Bayesian inference, we theoretically evaluated how branching complexity of random river networks, namely the network nodes' branching prevalence (Terui *et al*., 2018), differentially affects the global genetic differentiation throughout catchments in the context of different asymmetric gene flow modes. Here, we hypothesised that 1) widely distributed, generalist species associated with strong dispersal have smaller genetic divergence than specialist species with clumped, patchy or disjunctive distributions, and 2) increased river branching has positive effects on genetic divergence in species with downstream-biased dispersal but the opposite (negative) effect in those with symmetric dispersal, or upstream-biased asymmetric dispersal, which contributes to the dispersal-mediated consequences of the opposite effects of network branching (isolated habitats and landscape connectivity).

Materials and methods

Empirical catchment and genetic data

In the Natori and Nanakita Rivers in northeastern Japan (integrated catchment area c. 1200 km^2; Fig. 1), the flow regime exhibits a seasonal pattern, with flooding due to snowmelt in spring. In the integrated catchment, the rivers flow from the western headwaters, at an elevation of 1350 m at Mount Kamuro, to the eastern river mouths at the Pacific Ocean, passing through Sendai City with a population of one million. Approximately 60% of this area is forested and mountainous. Two major reservoir dams (Kamafusa and Okura dams) are located there. The regional lowlands are farmlands (13%, primarily with rice paddy fields) and a mixture of residential and commercial areas (11%).

For both empirical and theoretical evidence, we used genetic data of neutral amplified fragment length polymorphism (AFLP) markers from four macroinvertebrate species in this catchment (Watanabe *et al.*, 2014). Three species were caddisflies (Trichoptera), namely, *Hydropsyche orientalis*, *Stenopsyche marmorata* and *Hydropsyche albicephala*, while the fourth was a mayfly, *Ephemera japonica* (Ephemeroptera). In this integrated catchment, the species distributions vary considerably, from the widespread *H. orientalis* to the narrowly distributed *E. japonica* (Fig. 1). These species have similar ecological functions in river ecosystems by feeding on fine organic matter (< 1 mm diameter). Approximately 18 to 20 individuals collected at each sampling site were genotyped (128 to 473 polymorphic AFLP loci

for each species). Based on the locus-specific genetic differentiation across this catchment, non-neutral loci identified by DFDIST (Beaumont & Nichols, 1996) and/or BayeScan (Foll & Gaggiotti, 2008) were removed, and 98 to 449 neutral AFLP loci for each species (Fig. S1) were retained and used for this study. Detailed protocols on the identification of non-neutral loci are described in our previous report (Watanabe *et al.*, 2014).

Metapopulation genetic modelling

We developed a metapopulation genetic model based on isolation by distance (lower gene flow with greater separation in terms of distance along the watercourse) and asymmetric dispersal (upstream- and downstream-biased movements) between local populations. This model was validated using the empirical data on neutral AFLP loci of the four macroinvertebrate species in the catchment. With its parameters estimated by Bayesian inference, this model was used to simulate the river branching influence on each species. We describe the model development and Bayesian estimation of parameters below.

Given a single locus with two allelic types, labelled ‘1’ and ‘2’ (e.g. an AFLP), $z_{k,l}$ denotes the number of type ‘1’ alleles (number of individuals with the allele type) at locus *l* from number of alleles $s_{k,l}$ (total number of types ‘1’ and ‘2’ together = total number of individuals) observed in local population *k*. Here, neutral AFLP loci of individuals from local

populations were used as the observed modelling output. This random sampling process can be characterised by a binomial distribution as follows (Guillot *et al.*, 2014):

$$z_{k,l} \sim \text{Binomial}\left(s_{k,l}, f_{k,l}\right) \qquad [1]$$

where, in this local population, the frequency of allele '1' is denoted by $f_{k,l}$. The allele frequencies $z_{k,l}$ are independent between loci.

For each locus, the frequency of allele '1' in local populations is determined by their genetic variations (related to genetic drift), the watercourse distance between local populations within the network (related to gene flow) and allele frequencies of the metapopulation. Without genetic drift and natural selection, gene flow leads to the genetic homogeneity among local populations, leading to allele frequencies at loci in local populations matching those of their metapopulation (Andrews, 2010). We denote by $\theta_{k,l}$ the random deviation of logit-transformed allele frequency in local populations from that of the metapopulation, and the allele frequency $f_{k,l}$ is obtained from the inverse logit transformation as follows:

$$f_{k,l} = \text{invLogit}\left(\theta_{k,l} + m_{k,l}\right) = \frac{1}{1+\exp\left(-\left(\theta_{k,l}+m_{k,l}\right)\right)} \qquad [2]$$

where m_{ij} denotes the metapopulation's transformed allele frequency. The deviation $\theta_{k,l}$ can be modelled by a multivariate normal distribution as follows (Bradburd *et al.*, 2013):

$$\theta_{k,l} \sim \text{MultiNormal}(\mu, \Omega) \qquad [3]$$

where μ denotes the mean of zero, and the covariance matrix Ω is a function of the watercourse distance between local populations and their spatial relationships (either

streamflow-connected or -disconnected). To model the covariance across local populations, we modelled this as a function of the shortest watercourse distance along the river network h_{ij} between populations *i* and *j* as follows (Ver Hoef & Peterson, 2010):

$$\Omega_{ij} = \begin{cases} \underbrace{\sigma_D^2}_{a} + \underbrace{\sigma_U^2}_{b} + \sigma_G^2 & \text{if flow-connected and } i = j \\ \underbrace{\sigma_D^2 \exp\left(-\left(c_D h_{ij}\right)\right)}_{a} + \underbrace{\sigma_U^2 \exp\left(-\left(c_U h_{ij}\right)\right)}_{b} & \text{if flow-connected and } i \neq j \\ \underbrace{\sigma_U^2 \exp\left(-\left(c_U h_{ij}\right)\right)}_{b} & \text{if flow-disconnected} \end{cases} \quad [4]$$

where part *a* or *b* describes the autocovariance, with the variance σ_D^2 or σ_U^2 and the scale parameter c_D or c_U related to the downstream (D) or upstream (U) movement, respectively. In part *a*, the autocovariance is set to zero for any two streamflow-disconnected populations (e.g. local populations in different headwaters). In other words, streamflow-disconnected populations are independent and have no gene flow between them via downstream movement. The nugget variance σ_G^2 describes the random error.

In the Bayesian framework 'Stan' (Stan Development Team, 2014b), the R interface 'RStan' (Stan Development Team, 2014a) was used to perform this metapopulation genetic modelling. For each species, four Markov Chain Monte Carlo chains (for numerical approximations of Bayesian inference) ran with 60,000 iterations each, and the first half of the iterations for each chain were discarded as burn-in. This was determined by modelling convergence when the R-hat statistic of each parameter approached a value of 1. To estimate the model parameters, 2,000 samples obtained, by collecting one sample every 60 iterations for

each chain, were used to build the each parameter's posterior distribution.

Simulation of river-branching influences

Before the simulations, we created artificial river networks with varying branching complexities (Terui *et al.*, 2018). The river networks were made up of nodes with scale length *e*, with each node representing a local population. These nodes were assigned to be either branching (or an upstream terminal) or non-branching with a probability of P or $1 - P$, respectively. As a series of non-branching nodes terminated at a branching (or terminal) node. The individual segments (watercourse stretches) were the geometric random variables with branching probability P. Before merging the segments to create a river network, the drawing process was repeated until the targeted number of notes (the number of local populations) and an odd number of segments were reached. To create the river network, these segments were put together as a pool merged hierarchically as follows (Fig. S2): Step 1): One segment was randomly selected as the root and its upstream end was merged with the downstream end of another two random segment selections. In this status, the semi-complete network had two unmerged upstream ends each for the next possible merger. Step 2) Two more segments were randomly selected and their downstream ends were merged together to the random draw one of two (or even more at subsequent steps) unmerged upstream ends of the semi-complete network. Step 3) Step 2 was repeated until there were no available segments in the pool.

We conducted stochastic simulations to illustrate the uncertainty of the global genetic differentiation among local populations, G_{ST} (Nei, 1973), under river branching. We created 1,000 river networks (with scale length *e* equal to 1 km) with the branching probability *P* and metapopulation size *N* (integer; the number of local populations in a river network) randomly drawn from 0 to 1 and from 100 to 500, respectively. Our Bayesian model of each species with median estimates was used to stimulate the global G_{ST} to be the metapopulation's genetic divergence in each of the 1,000 random river networks. We performed this simulation using the R packages 'stats' and 'base' (R Core Team, 2018).

Here, we built a regression model based on gradient boosting (GB) for each of the four macroinvertebrate species, identifying the importance of 1) the fraction of any two local populations being streamflow-disconnected in all combinations (any two being streamflow-connected or -disconnected), 2) the mean watercourse distance between local populations under different levels of river branching and 3) metapopulation size (number of local populations) for genetic divergence (G_{ST}). GBs are a type of machine-learning algorithm used for analysing unilinear relationships at the base of multiple decision trees and, in the boosting process, each next tree model generated is added to improve on the performance of the previous ensemble of models by minimising deviance (Friedman, 2001). Our GB modelling was performed using the R package 'gbm' (Greenwell *et al.*, 2018), in which the genetic divergence and other factors (the river features and metapopulation size) were

independent and dependent variables, respectively. We used the R package 'dismo' to assess the optimal number of boosting trees via a cross-validation procedure (Hijmans *et al.*, 2017).

We illustrated how the downstream and upstream dispersal-related parameters of the variances (σ_D^2 and σ_U^2, respectively) or scales (c_D and c_U, respectively) influence river branching on global genetic divergence. To illustrate this for each parameter type, we considered 3 × 3 (nine) combinations of two parameters each with the same upper, median and lower ends of ranges of their Bayesian median pooled estimates. For the same parameter type, we replicated the nine combinations (see Fig. 5 and 6) in each of the 1,000 random river networks, and parameters of the other type were fixed to the median of pooled estimates. In addition to the variances and scales, each of the other model parameters was set to its Bayesian median estimate.

Results

Metapopulation genetic modelling

Our Bayesian model was fitted to the empirical genetic data in the Natori and Nanakita catchment, and the R^2 values derived from the residual (differences between the observed and predicted numbers of type '1' alleles at a locus from the number of alleles observed in a local population; see Formula 1) are 0.97, 0.98, 0.97 and 0.93 for *H. orientalis*, *S. marmorata*, *H. albicephala* and *E. japonica*, respectively (Fig. 2). The metapopulation allele frequencies are species-specific, and the variation of allele frequency is greater in the widespread *H. orientalis* and *S. marmorata* than in the other two species with narrower habitat distributions (Fig. S1). The pairwise genetic difference between empirical local populations tended to increase with their watercourse distances throughout the four macroinvertebrate species (Fig. S3). Despite substantial variation in the scale parameter, amplifying the isolating effect of distance across study species (Fig. S4), there was a consistent decline in the genetic correlation between populations (the covariance Ω_{ij} divided by the variance $\sigma_D^2 + \sigma_U^2 + \sigma_G^2$; see Formula 4) with the increasing distance between local populations (Fig. S5). In addition, there was a greater decline in the genetic correlation with distance in the widely distributed *H. orientalis* than in the other species.

River-branching influence

We describe changes in the two landscape spatial configurations (fraction of any two local populations being streamflow-disconnected in all combinations and mean watercourse distance between local populations) with the increasing branching probability (*P*) in river networks (Fig. 3). Situations in which any two local populations are streamflow-disconnected (e.g. in different tributaries) across metapopulations occur at higher rates in heavily branched river networks. However, we found shorter watercourse distances between local populations in river networks with higher branching probability. The metapopulation size (the number of interacting subpopulations in a network) increases the values of both spatial configurations under the same level of river branching.

Notably, changes in the values of the two spatial configurations act synergistically on the genetic differentiation of metapopulations (G_{ST}) of the river network across four species (Fig. 4). Species-specific responses to the influence of river branching were identified. For example, increased branching probability decreased the genetic divergence of the metapopulation for three caddisflies (*H. orientalis*, *S. marmorata* and *H. albicephala*), but in the mayfly *E. japonica*, the opposite response (higher genetic divergence) occurred. In addition, both a low level and variation of genetic divergence are less likely to occur in the generalist *H. orientalis* than in the other three speices. The findings showed that the metapopulation size was positively correlated to genetic divergence in all species. According to the GB modelling results, the relative importance of streamflow-disconnected habitats, compared to the landscape

connectivity via a shorter watercourse distance, was higher in the mayfly *E. japonica* than in the other three caddisfly species (Table 1).

The genetic performances, varying across species, were illustrated by how these model parameters related to upstream and downstream dispersals take effect on the genetic divergences (Fig. 5 and 6). Branching complexity has various impacts on genetic divergence, which is determined by the relative values of the upstream and downstream parameters (scale and variance in genetic covariation function, see Formula 4). The positive and negative influences of branching complexity on the genetic divergence are conferred through the relatively high and low values of the upstream-dispersal scale parameter compared to the downstream one, respectively (Fig. 5). These, in turn, indicated higher and lower isolation effects of watercourse distance between local populations, respectively. Lower genetic divergence levels occurred in more-branched networks when there was higher variance related to upstream movement (σ_U^2) than downstream movement (σ_D^2) (Fig. 6). In other words, the populations in the distal branches (e.g. headwaters) have relatively strong genetic covariation between themselves, particularly in complex river networks. In addition, river branching has the opposite (positive) influence on the genetic divergences when σ_U^2 is lower than or equal to σ_D^2 (Fig. 6).

Discussion

In this study, we explored the integrated role of landscape architecture and species ecological strategy in shaping genetic divergence at neutral loci. We compared the landscape genetics of sympatric macroinvertebrate species in river networks, based on our Bayesian model, explicitly accounting for the effects of evolutionary processes among components of metapopulations on the spatial genetic structure. This model indicated that river-network connectivity predicted spatial genetic structures in four macroinvertebrate species. In addition, their empirical structuring patterns were determined by the species' intrinsic factors parameterised in this model. In this case, these factors can be associated with dispersal ability and mode, species distribution and effective population size (associated with the genetic variance in our model) in characterising relationships between genetic divergence and landscape connectivity, as shown in the discussion below (see the subsequent section '*Importance of species' intrinsic factors*').

In our simulations, these intrinsic factors could cause varying levels of overall genetic differentiation in river networks and induced increased river branching to have different or even opposite effects. Moreover, greater landscape connectivity (via shortened watercourse distance) and higher distal habitat isolation (e.g. headwater streams) simultaneously occur in more-branched river networks and have countervailing influences on genetic divergence; they also have different levels of relative importance across these sympatric species. This can

provide extensive insights into other complex networks (e.g. highly fragmented landscapes or those with corridors via ocean and atmospheric circulation). Our empirical and theoretical results highlight the fundamental importance of considering species' biological traits, which make different contributions to genetic connectivity, for the successful management of ecological corridors.

River branching and metapopulation genetic divergence

In dendritic river networks, our simulation results showed a species-dependent change in global genetic differentiation levels occurring with increases in network complexity and the number of local populations in a metapopulation. We theoretically showed that the differential downstream and upstream gene flows we considered in the model can act together to generate such relationships. Our finding that increased populations in the river network enhanced genetic differentiation is consistent with previous theoretical evidence (Thomaz *et al*., 2016). River branching's role has been documented, to some extent, in riverscape genetics, when higher genetic diversity is observed in downstream populations than in upstream ones (Paz-Vinas *et al.*, 2015) and greater river branching can increase the differences between such populations (Thomaz *et al.*, 2016). Little or no gene flow imposed by high river branching, because of strong isolation of headwater populations, can generally be observed for some riverine species with high or intermediate levels of downstream-biased vagility, such as fish

species (Osborne *et al.*, 2014; Pilger *et al.*, 2017). By adopting our mechanistic model validated by empirical data on macroinvertebrate species with flying adult stages, this theoretical evidence reveals their dispersal ability to overcome riverscape constraints, leading to low downstream-biased asymmetry and the opposite (negative) influence occurring under increased river branching. In addition, our comprehensive consideration of various branching river network topologies in simulations helped us to demonstrate the existence of opposing influences co-occurring under branching complexity. In one early theoretical study, not considering dispersal asymmetry (analogous to equal downstream and upstream dispersals in our study), the dendritic network structure was also documented to promote low genetic distances under high riverscape connectivity (Labonne *et al.*, 2008).

Importance of species' intrinsic factors

The river networks' architecture can be one important extrinsic factor for explaining the observed and simulated genetic patterns, but there was strong variation among species with different intrinsic factors in our study. This finding was also previously observed; for example, two sympatric salmonid species were found to have remarkably different spawning locations, mating systems and population sizes, and these biological traits mediated the influences of riverscape features shaping their dispersal and genetic divergence in the Clark Fork River in the USA (Whiteley *et al.*, 2004). For each upstream and downstream dispersal tendency in our

model, there are two parameters (scale and variance in the genetic covariation function; see Formula 4) linked to species' intrinsic factors. In addition, these parameters together shape a mechanism behind the countervailing influences of river branching on the genetic divergence of metapopulations.

In our system, asymmetric (either downstream- or upstream-biased) dispersals could determine the direction of the resulting influence on riverscape complexity. Stronger dispersal can be associated with a lower value of the scale parameter since this parameter expands the isolation effects of distance within river networks (see Formula 4). Our modelling results showed that the widely distributed, generalist caddisfly (*H. orientalis*) has less intense genetic divergence than other species with clumped, patchy or disjunctive distributions. In addition, little change in genetic divergence, along with river branching, occurs in this caddisfly, which can be explained by the low isolation effect by watercourse distance.

Furthermore, our results suggested that, in the mayfly species (*E. japonica*) with high downstream-biased gene flow (based on a higher value of the scale parameter for upstream than for downstream), as typically shown in fish species (e.g., Pilger *et al.*, 2017), river branching has a positive influence on its genetic divergence, in which a higher number of isolated distal branches in river networks (e.g. headwaters) occur under this dispersal asymmetry. Mayfly species larvae are susceptible to drift during high river flow and have great potential to be strong downstream dispersers (Nukazawa *et al.*, 2017). In our system, the

opposite (negative) influence of the branching network on the other three caddisfly species with dispersal symmetry or even upstream-biased dispersal (based on the scale parameter for upstream being similar to or higher than that for downstream, respectively; Fig. S4 and Fig. 5) was identified. This might be attributable to their flying adults generally having a wide dispersal range, showing strong terrestrial movement at least in the upstream direction. This dampened the isolation between distal branches in river networks, compared to the case for mayflies (or even stoneflies), exhibiting restricted distributions to areas very close to their sources of emergence in the stream (Winterbourn *et al*., 2007).

Besides dispersal abilities, the effective population size was revealed to be a factor potentially influencing genetic drift and mediating the countervailing influences of river branching in our study. In our model, the uncertainty regarding allele frequency, determined by the variance parameter, can describe the levels of genetic drift, which can theoretically be associated with the effective population size (Nei & Tajima, 1981). Observational studies have documented that a smaller effective size of local populations can result in their higher genetic differentiation, induced by genetic drift (Weckworth *et al.*, 2013; Richmond *et al.*, 2018). In addition, there is an association between species features (dispersal ability and habitat requirements) and genetic divergence (Phillipsen *et al*., 2015). In our model, the variance parameter can be separated into two parts related to upstream and downstream dispersals in the genetic covariation function (Formula 4). Migration is one source of changes in population size,

and migrants themselves exhibit genetic variation derived from their source populations. Different migrant population sizes can have varying influences on the genetic drift of sink populations in the upstream or downstream direction. Therefore, the populations' locations (e.g. in tributaries or the main stem) determine how immigration routes in upstream and/or downstream directions act together in local populations, leading to their genetic drift. In our mayfly species (*E. japonica*), for example, a higher value of downstream-related variance, compared to the upstream one, theoretically determines the higher genetic variation among local populations in distal branches of river networks (Fig. S4 and 6). As documented in both theoretical and empirical studies, local populations of aquatic obligate species (e.g. fish constrained to river channels) in main stem confluences experience less genetic drift than those in isolated headwaters in river ecosystems (Thomaz *et al.*, 2016; Pilger *et al.*, 2017). As a result, in cases with an increased number of tributaries and augmented river branching levels, there is a higher likelihood that metapopulation genetic divergence will increase, for example, for this mayfly species in our study or other downstream-biased species elsewhere (Osborne *et al.*, 2014).

From a conservation and management perspective that takes spatial genetic structure into account (Luque *et al.*, 2012), it is crucial to understand the branching structure's role in driving metapopulation genetic divergence. Dispersal can dictate differences in landscape genetic diversification (Medina *et al.*, 2018), and predictive modelling, which can be validated by

empirical data based on asymmetric dispersals across networks, and shed light on the expected impacts of global climate change and the consequences of management practices. To manage native or even invasive species, our results shed light on the evolutionary importance of dispersal abilities and modes, suggesting that these intrinsic factors should be considered in decision-making processes when one managing strategy does not fit all species. For example, the same management and conversation practices can produce different, or even the opposite, results for species with varying levels of asymmetric gene flow and genetic drift (e.g. in dendritic river systems).

Acknowledgements

This study was partly funded by the Japan Society for the Promotion of Science (JSPS). We are grateful to Maribet Gamboa for her constructive comments on the manuscript.

References

Albert, E.M., Fortuna, M.A., Godoy, J.A., & Bascompte, J. (2013) Assessing the robustness of networks of spatial genetic variation. *Ecol Lett*, **16 Suppl 1**, 86-93.

Andrews, C.A. (2010) Natural selection, genetic drift, and gene flow do not act in isolation in natural populations. *Nature Education Knowledge*, **3**, 5.

Beaumont, M.A. & Nichols, R.A. (1996) Evaluating Loci for Use in the Genetic Analysis of Population Structure. *Proceedings: Biological Sciences*, **263**, 1619-1626.

Bradburd, G.S., Ralph, P.L., & Coop, G.M. (2013) Disentangling the effects of geographic and ecological isolation on genetic differentiation. **67**, 3258-3273.

Braunisch, V., Segelbacher, G., & Hirzel, A.H. (2010) Modelling functional landscape connectivity from genetic population structure: a new spatially explicit approach. *Molecular Ecology*, **19**, 3664-3678.

Campbell Grant, E.H., Lowe, W.H., & Fagan, W.F. (2007) Living in the branches: population dynamics and ecological processes in dendritic networks. *Ecology Letters*, **10**, 165-175.

Chave, J. (2013) The problem of pattern and scale in ecology: what have we learned in 20 years? *Ecology Letters*, **16**, 4-16.

Chust, G., Villarino, E., Chenuil, A., Irigoien, X., Bizsel, N., Bode, A., Broms, C., Claus, S., Fernández de Puelles, M.L., Fonda-Umani, S., Hoarau, G., Mazzocchi, M.G., Mozetič, P., Vandepitte, L., Veríssimo, H., Zervoudaki, S., & Borja, A. (2016) Dispersal similarly

shapes both population genetics and community patterns in the marine realm. *Scientific Reports*, **6**, 28730.

Economo, E.P. & Keitt, T.H. (2008) Species diversity in neutral metacommunities: a network approach. *Ecol Lett*, **11**, 52-62.

Foll, M. & Gaggiotti, O. (2008) A Genome-Scan Method to Identify Selected Loci Appropriate for Both Dominant and Codominant Markers: A Bayesian Perspective. **180**, 977-993.

Friedman, J.H. (2001) Greedy function approximation: A gradient boosting machine. *Ann. Statist.*, **29**, 1189-1232.

Green, D. (2006) *Complexity in Landscape Ecology*.

Greenwell, B., Boehmke, B., Cunningham, J., & Developers, G. (2018) *gbm: Generalized Boosted Regression Models. R package version 2.1.4*.

Guillot, G., Vitalis, R., Rouzic, A.l., & Gautier, M. (2014) Detecting correlation between allele frequencies and environmental variables as a signature of selection. A fast computational approach for genome-wide studies. *Spatial Statistics*, **8**, 145-155.

Hijmans, R.J., Phillips, S., Leathwick, J., & Elith, J. (2017) *dismo: Species Distribution Modeling. R package version 1.1-4.* .

Kawecki, T.J. & Holt, R.D. (2002) Evolutionary consequences of asymmetric dispersal rates. *The American Naturalist*, **160**, 333-347.

Labonne, J., Ravigné, V., Parisi, B., & Gaucherel, C. (2008) Linking dendritic network

structures to population demogenetics: The downside of connectivity. **117**, 1479-1490.

Luque, S., Saura, S., & Fortin, M.-J.J.L.E. (2012) Landscape connectivity analysis for conservation: insights from combining new methods with ecological and genetic data. **27**, 153-157.

McRae, B.H. (2006) Isolation by resistance. *Evolution*, **60**, 1551-1561.

Medina, I., Cooke, G.M., & Ord, T.J. (2018) Walk, swim or fly? Locomotor mode predicts genetic differentiation in vertebrates. *Ecology Letters*, **21**, 638-645.

Nei, M. (1973) Analysis of gene diversity in subdivided populations. *Proceedings of the National Academy of Sciences of the United States of America*, **70**, 3321-3323.

Nei, M. & Tajima, F. (1981) Genetic drift and estimation of effective population size. *Genetics*, **98**, 625-40.

Nukazawa, K., Kazama, S., & Watanabe, K. (2015) A hydrothermal simulation approach to modelling spatial patterns of adaptive genetic variation in four stream insects. *Journal of Biogeography*, **42**, 103-113.

Nukazawa, K., Kazama, S., & Watanabe, K. (2017) Catchment-scale modeling of riverine species diversity using hydrological simulation: application to tests of species-genetic diversity correlation. *Ecohydrology*, **10**, e1778.

Osborne, M.J., Perkin, J.S., Gido, K.B., & Turner, T.F. (2014) Comparative riverscape genetics reveals reservoirs of genetic diversity for conservation and restoration of Great Plains

fishes. **23**, 5663-5679.

Paz-Vinas, I. & Blanchet, S. (2015) Dendritic connectivity shapes spatial patterns of genetic diversity: a simulation-based study. *Journal of Evolutionary Biology*, **28**, 986-994.

Paz-Vinas, I., Loot, G., Stevens, V.M., & Blanchet, S. (2015) Evolutionary processes driving spatial patterns of intraspecific genetic diversity in river ecosystems. *Mol Ecol*, **24**, 4586-604.

Paz-Vinas, I., Quéméré, E., Chikhi, L., Loot, G., & Blanchet, S. (2013) The demographic history of populations experiencing asymmetric gene flow: combining simulated and empirical data. *Molecular Ecology*, **22**, 3279-3291.

Petersen, I., Masters, Z., Hildrew, A.G., & Ormerod, S.J. (2004) Dispersal of adult aquatic insects in catchments of differing land use. **41**, 934-950.

Phillipsen, I.C., Kirk, E.H., Bogan, M.T., Mims, M.C., Olden, J.D., & Lytle, D.A. (2015) Dispersal ability and habitat requirements determine landscape-level genetic patterns in desert aquatic insects. *Molecular Ecology*, **24**, 54-69.

Pilger, T.J., Gido, K.B., Propst, D.L., Whitney, J.E., & Turner, T.F. (2017) River network architecture, genetic effective size and distributional patterns predict differences in genetic structure across species in a dryland stream fish community. *Mol Ecol*, **26**, 2687-2697.

R Core Team (2018) *R: A Language and Environment for Statistical Computing*. R Foundation

for Statistical Computing, Vienna, Austria. https://www.R-project.org/.

Richmond, J.Q., Backlin, A.R., Galst-Cavalcante, C., O'Brien, J.W., & Fisher, R.N. (2018) Loss of dendritic connectivity in southern California's urban riverscape facilitates decline of an endemic freshwater fish. **27**, 369-386.

Ruiz-Gonzalez, A., Cushman, S.A., Madeira, M.J., Randi, E., & Gómez-Moliner, B.J. (2015) Isolation by distance, resistance and/or clusters? Lessons learned from a forest-dwelling carnivore inhabiting a heterogeneous landscape. **24**, 5110-5129.

Stan Development Team (2014a) *RStan: the R interface to Stan, Version 2.5.* *http://mc-stan.org/rstan.html*.

Stan Development Team (2014b) *Stan Modeling Language Users Guide and Reference Manual, Version 2.5.0*.

Terui, A., Ishiyama, N., Urabe, H., Ono, S., Finlay, J.C., & Nakamura, F. (2018) Metapopulation stability in branching river networks. *Proc Natl Acad Sci U S A*, **115**, E5963-E5969.

Thomaz, A.T., Christie, M.R., & Knowles, L.L. (2016) The architecture of river networks can drive the evolutionary dynamics of aquatic populations. *Evolution*, **70**, 731-739.

Ver Hoef, J.M. & Peterson, E.E. (2010) A moving average approach for spatial statistical models of stream networks. *Journal of the American Statistical Association*, **105**, 6-18.

Watanabe, K., Kazama, S., Omura, T., & Monaghan, M.T. (2014) Adaptive genetic divergence

along narrow environmental gradients in four stream insects. *PLoS One*, **9**, e93055.

Weckworth, B.V., Musiani, M., Decesare, N.J., McDevitt, A.D., Hebblewhite, M., & Mariani, S. (2013) Preferred habitat and effective population size drive landscape genetic patterns in an endangered species. *Proc Biol Sci*, **280**, 20131756.

Whiteley, A.R., Spruell, P., & Allendorf, F.W. (2004) Ecological and life history characteristics predict population genetic divergence of two salmonids in the same landscape. **13**, 3675-3688.

Wilson, M.C., Chen, X.-Y., Corlett, R.T., Didham, R.K., Ding, P., Holt, R.D., Holyoak, M., Hu, G., Hughes, A.C., Jiang, L., Laurance, W.F., Liu, J., Pimm, S.L., Robinson, S.K., Russo, S.E., Si, X., Wilcove, D.S., Wu, J., & Yu, M. (2016) Habitat fragmentation and biodiversity conservation: key findings and future challenges. *Landscape Ecology*, **31**, 219-227.

Winterbourn, M.J., Chadderton, W.L., Entrekin, S.A., Tank, J.L., & Harding, J.S. (2007) Distribution and dispersal of adult stream insects in a heterogeneous montane environment. *Fundamental and Applied Limnology / Archiv f??r Hydrobiologie*, **168**, 127-135.

Xuereb, A., Benestan, L., Normandeau, É., Daigle, R.M., Curtis, J.M.R., Bernatchez, L., & Fortin, M.-J. (2018) Asymmetric oceanographic processes mediate connectivity and population genetic structure, as revealed by RADseq, in a highly dispersive marine

invertebrate (*Parastichopus californicus*). *Molecular Ecology*, **27**, 2347-2364.

Table 1. Performance measures and relative importance of predictors in gradient boosting models for *Ephemera japonica* (EJ), *Stenopsyche marmorata*, *Hydropsyche orientalis* (HO) and *Hydropsyche albicephala* (HA), in which the simulated genetic divergence of the metapopulation (global G_{ST}) is the response variable and fraction of any two local populations streamflow-disconnected, mean watercourse distance between populations and metapopulation size are predictor variables.

		Model performance		**Relative importance (%)**		
Species	**Number of trees**	**RMSE**	**R^2**	**Streamflow-disconnected fraction**	**Watercourse distance**	**Metapopulation size**
EJ	2250	0.03	0.94	90.4	4.5	5.1
HA	800	0.06	0.88	1.0	95.1	3.9
HO	600	0.05	0.64	4.3	85.3	10.4
SM	650	0.03	0.65	6.0	79.4	14.6

Figure legends

Fig. 1. The study catchment and distribution of (a) *Ephemera japonica* (EJ), (b) *Hydropsyche albicephala* (HA), (c) *Hydropsyche orientalis* (HO) and (d) *Stenopsyche marmorata* (SM) in northeastern Japan.

Fig. 2. Observed and predicted numbers of type '1' alleles at a locus from the number of alleles observed in a local population (see Formula 1) for (a) *Ephemera japonica* (EJ), (b) *Hydropsyche albicephala* (HA), (c) *Hydropsyche orientalis* (HO) and (d) *Stenopsyche marmorata* (SM) with posterior distribution in Bayesian modelling.

Fig. 3. Theoretical predictions for relationships of (a) mean watercourse distance between populations or (b) fraction of any two streamflow-disconnected populations (e.g. in headwaters) in all combinations with branching complexity under differential metapopulation sizes (range: 100 to 500, number of local populations).

Fig. 4. Theoretical predictions for relationships between metapopulation genetic divergence (global G_{ST}) and branching complexity under differential metapopulation sizes (range: 100 to 500, number of local populations) for (a) *Ephemera japonica* (EJ), (b) *Hydropsyche albicephala* (HA), (c) *Hydropsyche orientalis* (HO) and (d) *Stenopsyche marmorata* (SM).

Fig. 5. Theoretical predictions for relationships between metapopulation genetic divergence (global G_{ST}) and branching complexity under differential metapopulation sizes (range: 100 to 500, number of local populations) for combinations of dispersal-related scale parameters in

genetic covariation function (Formula 4), including (a, e and i) $c_D = c_U$, (b, c and f) $c_D < c_U$, and (d, g and h) $c_D > c_U$.

Fig. 6. Theoretical predictions for relationships between metapopulation genetic divergence (global G_{ST}) and branching complexity under differential metapopulation sizes (range: 100 to 500, number of local populations) for combinations of dispersal-related variances in genetic covariation function (Formula 4), including (a, e and i) $\sigma_D^2 = \sigma_U^2$, (b, c, and f) $\sigma_D^2 < \sigma_U^2$ and (d, g and h) $\sigma_D^2 > \sigma_U^2$.

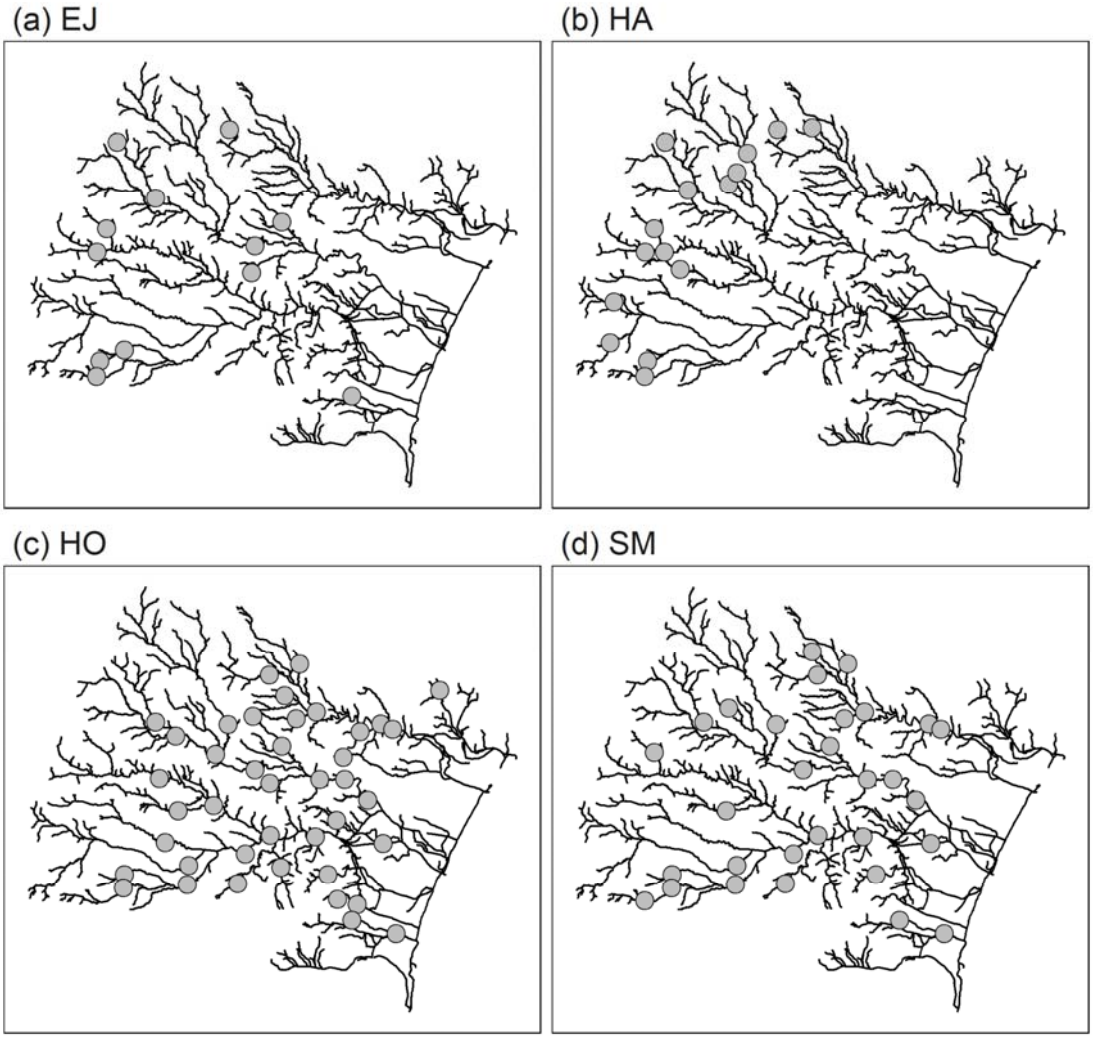


Fig. 1

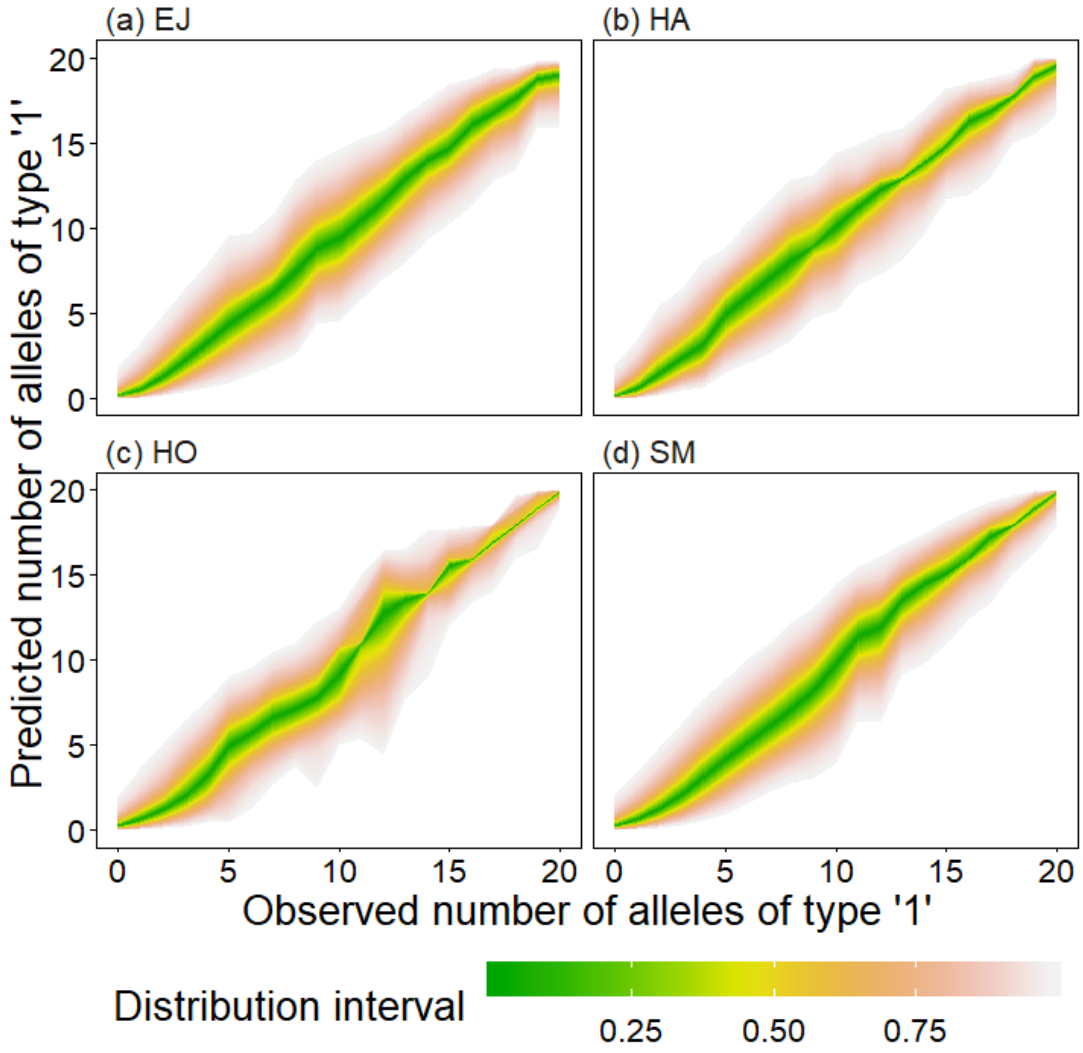
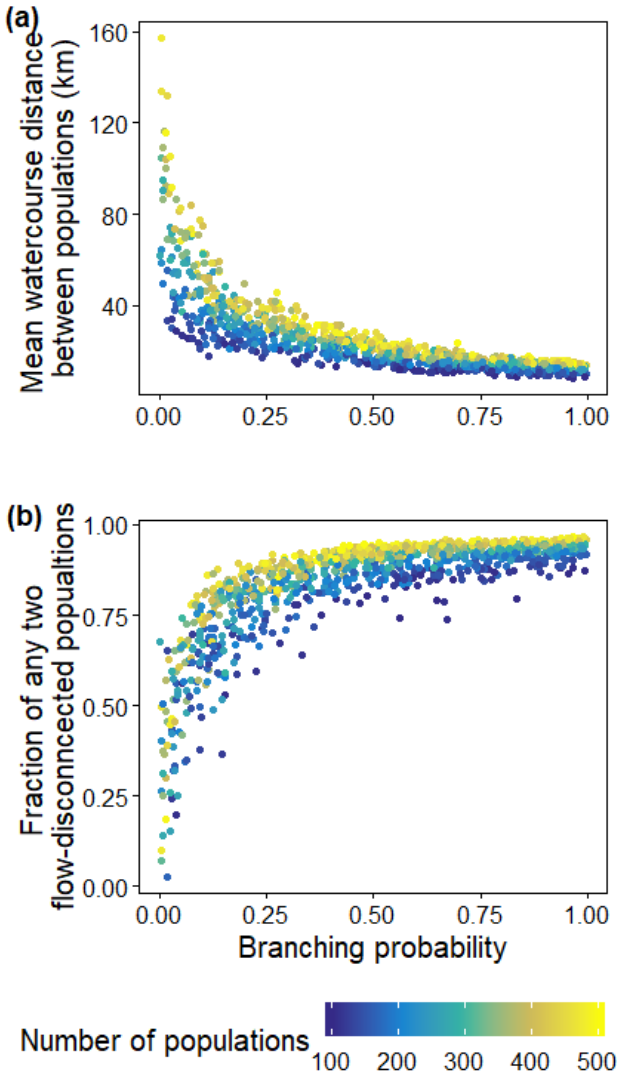

(a) EJ
(b) HA
(c) HO
(d) SM
Predicted number of alleles of type '1'
Observed number of alleles of type '1'
0
5
10
15
20
Distribution interval
0.25
0.50
0.75

Fig. 2


(a)
Mean watercourse distance between populations (km)
160
120
80
40
0.00
0.25
0.50
0.75
1.00
(b)
Fraction of any two flow-disconncected popualtions
1.00
0.75
0.50
0.25
0.00
0.00
0.25
0.50
0.75
1.00
Branching probability
Number of populations
100
200
300
400
500

570

571 **Fig. 3**

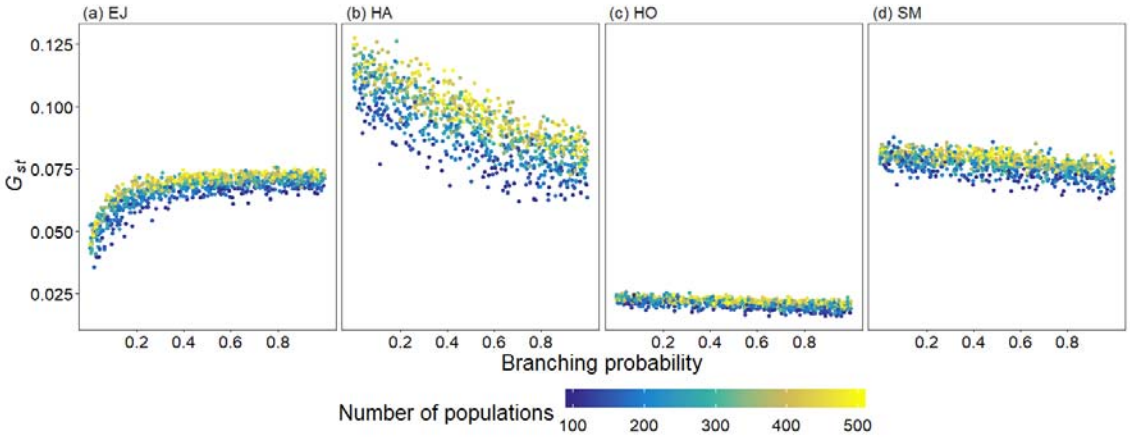

(a) EJ
(b) HA
(c) HO
(d) SM
0.125
0.100
0.075
0.050
0.025
G_{st}
0.2 0.4 0.6 0.8
Branching probability
Number of populations
100 200 300 400 500

Fig. 4

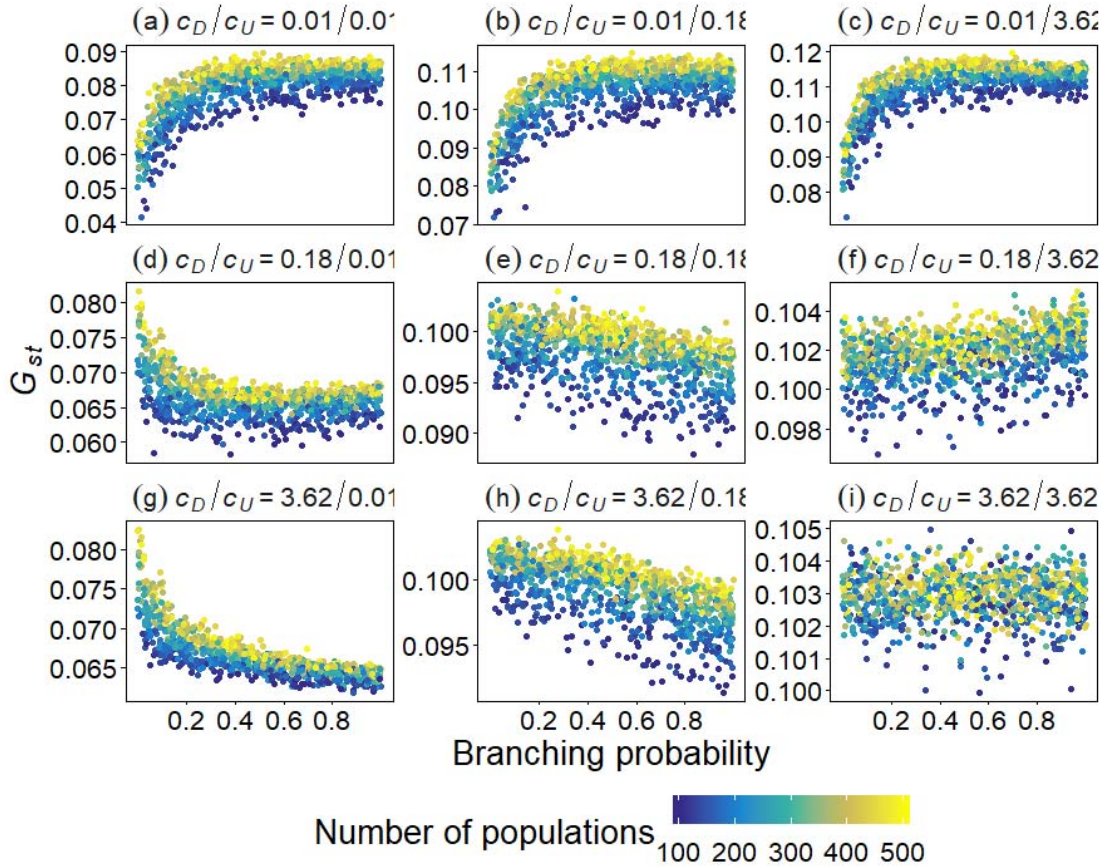
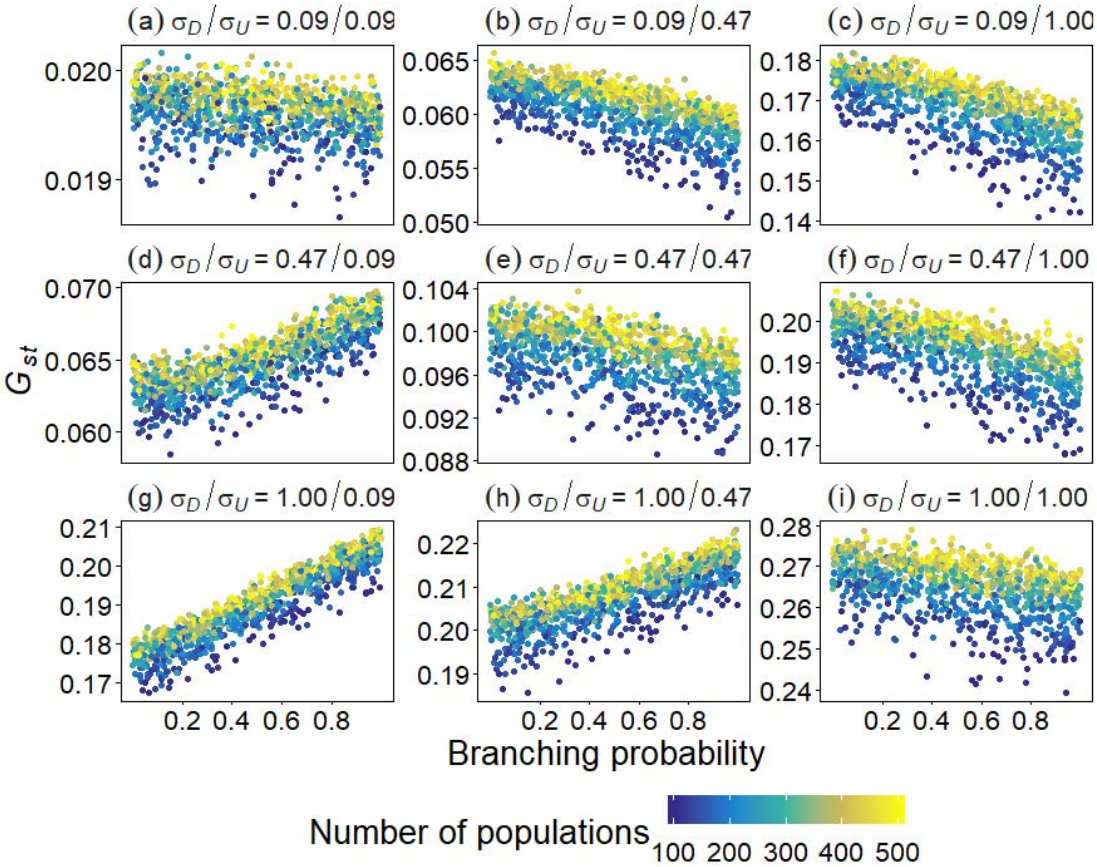

(a) c_D/c_U = 0.01/0.01
(b) c_D/c_U = 0.01/0.18
(c) c_D/c_U = 0.01/3.62
(d) c_D/c_U = 0.18/0.01
(e) c_D/c_U = 0.18/0.18
(f) c_D/c_U = 0.18/3.62
(g) c_D/c_U = 3.62/0.01
(h) c_D/c_U = 3.62/0.18
(i) c_D/c_U = 3.62/3.62
G_{st}
Branching probability
Number of populations
100 200 300 400 500

Fig. 5


(a) σD/σU = 0.09/0.09
(b) σD/σU = 0.09/0.47
(c) σD/σU = 0.09/1.00
(d) σD/σU = 0.47/0.09
(e) σD/σU = 0.47/0.47
(f) σD/σU = 0.47/1.00
(g) σD/σU = 1.00/0.09
(h) σD/σU = 1.00/0.47
(i) σD/σU = 1.00/1.00
Gst
Branching probability
Number of populations
100 200 300 400 500

576

577 **Fig. 6**