Individual-based multiple-unit dissimilarity: novel indices and null model for assessing temporal variability in community composition

Running title: Individual-based multiple-unit dissimilarity

Ryosuke Nakadai*[123]

[1] Department of Environmental and Biological Sciences, Faculty of Science and Forestry, University of Eastern Finland, Yliopistokatu 7, 80101 Joensuu, Finland

[2] Department of Ecosystem Studies, Graduate School of Agricultural and Life Sciences, The University of Tokyo,1-1-1 Yayoi, Bunkyo-ku, Tokyo 113-8657, Japan

[3] Biodiversity Division, National Institute for Environmental Studies, Onogawa 16-2, Tsukuba, Ibaraki, 305-8506, Japan

* Correspondence: R. Nakadai

E-mail: r.nakadai66@gmai.com

Keywords: Individual, individual turnover, beta-diversity, multiple-temporal dissimilarity, randomization

Abstract

1. Beta-diversity was originally defined spatially, i.e., as variation in community composition among sites in a region. However, the concept of beta-diversity has since been expanded to temporal contexts. This is referred to as “temporal beta-diversity”, and most approaches are simply an extension of spatial beta-diversity.

2. The persistence and turnover of individuals over time is a unique feature of temporal beta-diversity. Nakadai (2020) introduced the “individual-based beta-diversity” concept, and provided novel indices to evaluate individual turnover and compositional shift by comparing individual turnover between two periods at a given site. However, the proposed individual-based indices are applicable only to pairwise dissimilarity, not to multiple-temporal (or more generally, multiple-unit) dissimilarity.

3. Here, individual-based beta-diversity indices are extended to multiple-unit cases. In addition, a novel type of random permutation criterion related to these multiple-unit indices for detecting patterns of individual persistence is introduced in the present study.

4. To demonstrate the usage the properties of these indices compared to average pairwise measures, I applied them to a dataset for a permanent 50-ha forest dynamics plot on Barro Colorado Island in Panama.

5. Information regarding “individuals” is generally missing from community ecology and biodiversity studies of temporal dynamics. In this context, the method proposed here is expected to be useful for addressing a wide range of research questions regarding temporal changes in biodiversity, especially studies using traditional individual-tracked forest monitoring data and bird banding data.

Introduction

The concept of beta-diversity was introduced by Whittaker (1960, 1972) to define the variation in community composition among sites in a region. Several other beta-diversity indices have subsequently been proposed (reviewed by Koleff et al., 2003, Anderson et al., 2011, Jost et al., 2011, Legendre & De Cáceres, 2013). Recent studies have focused on temporal changes in community composition at both single sites and multiple sites surveyed repeatedly over time (Magurran 2011, Legendre & Condit, 2019). Temporal changes in community composition are referred to as "temporal beta-diversity" (Hatosy et al., 2013, Legendre & Gauthier, 2014, Shimadzu et al., 2015), which is simply an extension of spatial beta-diversity. Only a few studies have focused on methodological developments of temporal beta-diversity (Shimadzu et al., 2015, Legendre, 2019, Nakadai, 2020). The persistence and turnover of individuals over time is a key feature of temporal beta-diversity (Magurran et al., 2019, Nakadai, 2020). The speed and frequency of compositional change over time is associated with the speed of individual turnover, and must be considered in community comparisons because even randomly high individual turnover can result in high temporal beta-diversity (Nakadai, 2020).

To address this issue, Nakadai (2020) proposed the "individual-based beta-diversity" concept, and provided novel indices to evaluate individual turnover and compositional shift by comparing individual turnover between two periods at a given site. Specifically, I developed an individual-based temporal beta-diversity approach for the Bray–Curtis dissimilarity index (i.e., an abundance-based temporal beta-diversity index) (Nakadai 2020). Over time, some individuals replace others, and therefore ecological communities are dynamic and vary to some degree both spatially and temporally (Mori et al., 2018). In addition, some individuals of long-lived species (e.g., trees and vertebrates) persist over the long term, whereas individuals of short-lived species are replaced with conspecifics or individuals of other species (Nakadai, 2020).

The average values of pairwise dissimilarity have been widely used to evaluate compositional heterogeneity among multiple sites but cannot accurately reflect the overall compositional heterogeneity within a pool of more than two sites (Baselga 2013a). Diserud and Ødegaard (2007) first proposed a multiple-site (or more generally, a multiple-unit) similarity measure, which has conceptual and methodological advantages over conventional

approaches based on average pairwise similarities (Koleff & Gaston, 2002; Gaston et al., 2007, Baselga et al., 2007). Many dissimilarity indices have been extended to multiple-unit indices (Baselga, 2013b, 2017). Multiple-unit dissimilarity indices can take into account components shared among more than three units (i.e., sites or temporal units) (Baselga et al., 2007). The pairwise individual-based indices proposed by Nakadai (2020) may be limited by the typical problems associated with the application of multiple-unit datasets; in particular, the presence of long-lived individuals resulted in underestimates of individual turnover when the average values of the pairwise indices were used. Extending individual-based indices to multiple-temporal datasets may allow accurate assessment of the temporal variability of community composition based on individual turnover.

A null model is a pattern-generating model based on randomization of ecological data and is designed to produce a pattern that would be expected in the absence of a particular ecological mechanism (Gotelli & Graves, 1996, Gotelli & McGill, 2006). Null models for temporal patterns in ecological communities have been evaluated previously (e.g., cyclic shift permutations; Hallett et al., 2014, 2016). Such models use a matrix of the abundance, presence or absence, or biomass of species over multiple temporal units (e.g., years) as an input. In each realization of the algorithm, the time series of every species is shifted forward by a random number of units, independently of other species (Kalyuzhny, 2020). However, no null model approach using individual randomization has been developed for community dynamics, and no research has been conducted on the patterns of individual persistence using either beta-diversity indices or null models.

Therefore, the purpose of this paper was to develop new multiple-unit indices that can quantify compositional variability across time according to the speed of individual turnover, as well as novel random permutation criteria to evaluate the deviation in individual persistence from the perspective of ecological drift. In the present study, I extended individual-based beta-diversity indices to multiple units after briefly reviewing the history of pairwise and multiple-unit indices and also introduced a newly developed null model. Then, I applied these indices to data from a 50-h forest dynamics plot on Barro Colorado Island (BCI) in Panama (Condit et al., 2019), as a case study to demonstrate the usage and properties of indices compared to average pairwise measures. Using both novel indices and a null model provides opportunities to answer the following questions pertaining to the BCI

plot. Are differences in the degree of compositional variability explained only by differences in the degree of individual turnover? Does the degree of deviation in individual persistence from each expected value of ecological drift lead to differences in compositional variability among forest plots?

Materials and methods

Conventional incidence-based and abundance-based indices and novel individual-based indices

Two types of dissimilarity indices have been proposed in previous studies: incidence-based and abundance-based dissimilarity indices (Baselga, 2013b, Nakadai, 2020). Specifically, the widely used Bray–Curtis dissimilarity index is an abundance-based extension of the Sørensen index (Legendre & Legendre, 2012, Baselga, 2013b). The Bray–Curtis dissimilarity index was originally proposed as a percentage difference index by Odum (1950), yet it is commonly misattributed to Bray & Curtis (1957). The incidence-based and abundance-based methods target the number of species and the number of individuals, respectively. Conventional beta-diversity indices are based on intersection (A, a) and relative complement (B, b and C, c) components. For example, the Sørensen dissimilarity index (d_{sor}; Sørensen, 1948) is formulated as follows:

$$d_{sor} = \frac{b+c}{2a+b+c} \qquad \text{equation 1}$$

where a is the number of species common to both units (i.e., sites or temporal units), b is the number of species that occur in the first but not the second unit, and c is the number of species that occur in the second but not the first unit. Similarity, the Bray–Curtis dissimilarity index (Bray & Curtis, 1957) is formulated as follows:

$$d_{BC} = \frac{B+C}{2A+B+C}. \qquad \text{equation 2}$$

The abundance-based components are as follows:

$$A = \sum_i \min\left(x_{ij}, x_{ik}\right) \qquad \text{equation 3a}$$

$$B = \sum_i \left[x_{ij} - \min\left(x_{ij}, x_{ik}\right)\right] \qquad \text{equation 3b}$$

$$C = \sum_i [x_{ik} - \min(x_{ij}, x_{ik})] \quad \text{equation 3c}$$

where x_{ij} is the abundance of species i at unit (i.e., site or time) j, and x_{ik} is the abundance of species i at unit (i.e., site or time) k. Therefore, A is the number of individuals of each species common to both units j and k, whereas B and C are the numbers of individuals unique to units j and k, respectively.

Therefore, it should be possible to distinguish the community dynamics component of persistence (P) from those that change over time, i.e., mortality and recruitment (M and R, respectively). M refers to individuals occurring at time 1 (abbreviated T1) that die before time 2 (abbreviated T2), thereby representing the death of individuals during the time between T1 and T2. R identifies individuals that were not alive at T1 (or were not yet in the size class counted, e.g., diameter breast height ≥ 1 cm) but were subsequently alive and counted at T2, thereby representing R between T1 and T2. Component P refers to the persistence of individuals from T1 to T2. Although P, M, and R relate to components A, B, and C, respectively, here I have used different characters to emphasize the processes represented by each component. The respective abundances were calculated as follows:

$$P = \sum_o \min(z_{o1}, z_{o2}) \quad \text{equation 4a}$$

$$M = \sum_o [z_{o1} - \min(z_{o1}, z_{o2})] \quad \text{equation 4b}$$

$$R = \sum_o [z_{o2} - \min(z_{o1}, z_{o2})] \quad \text{equation 4c}$$

where z_{o1} is individual o at time T1, and z_{o2} is individual o at time T2. Both z_{o1} and z_{o2} take a value only of 1 or 0 (presence and absence, respectively). Therefore, P is the total number of individuals present at both T1 and T2, whereas M and R are the numbers of individuals unique to T1 and T2, respectively (Nakadai 2020).

T1 and T2 can be generalized as unit j (Tj) and unit k (Tk), and generalized versions of P, M, and R for units j and k are simply formulated as follows:

$$P = \sum_o \min(z_{oj}, z_{ok}) \quad \text{equation 5a}$$

$$M = \sum_o [z_{oj} - \min(z_{oj}, z_{ok})] \quad \text{equation 5b}$$

$$R = \sum_o [z_{ok} - \min(z_{oj}, z_{ok})], \quad \text{equation 5c}$$

which are used for the main analysis of this study.

This formulation of the individual-based temporal beta-diversity index for Bray–Curtis (d_{MR}) dissimilarity can be expressed as follows:

$$d_{MR} = \frac{M+R}{2P+M+R} \qquad \text{equation 6}$$

In some cases, there is turnover of different individuals belonging to the same species even when the community composition is stable over time, which contributes to a dynamic compositional equilibrium across time periods. Therefore, the components that change over time (*M* and *R*) are reorganized into two further components: compositional equilibrium (*E*) and shift (*B* and *C*), respectively (Nakadai 2020). Component *M* can be partitioned into lost individuals that contribute to compositional change in a community (*B*) and lost individuals replaced by conspecifics, contributing substantially to equilibrium in community composition (E_{loss}). Similarly, component *R* can be partitioned into gained individuals that contribute to compositional change in a community (*C*) and gained individuals replacing conspecifics, thereby contributing substantially to equilibrium in community composition (E_{gain}). By definition, the numbers of E_{loss} and E_{gain} individuals are identical; hence, I replaced both E_{loss} and E_{gain} with the coefficient *E*, i.e., the component that contributes to dynamic compositional equilibrium (Nakadai 2020). Furthermore, components *M* and *R* can also be analyzed in the same manner as community data (Nakadai, 2020). Therefore, the Bray–Curtis dissimilarity between communities can be calculated based on *M* and *R* as follows:

$$v_s = \frac{B+C}{2E+B+C} = \frac{B+C}{M+R} \qquad \text{equation 7}$$

where v_s indicates the speed of compositional shifts in a community relative to the total speed of individual turnover associated with *M* and *R*. If v_s is 0, then individual turnover would not contribute to compositional shifts in a community; i.e., *M* and *R* are identical, and an

equilibrium state exists. If v_s is 1, all individual turnovers contribute to compositional shifts, and M and R are completely different. It is also possible to calculate v_s as d_{BC} divided by d_{MR}, as follows:

$$d_{BC} \div d_{MR} = \frac{B+C}{2A+B+C} \div \frac{M+R}{2P+M+R} = \frac{B+C}{2(P+E)+B+C} \div \frac{M+R}{2P+M+R} = \frac{B+C}{2P+M+R} \div \frac{M+R}{2P+M+R} = \frac{B+C}{2E+B+C} = v_s. \quad \text{equation 8}$$

Conventional multiple-unit beta-diversity indices

Multiple-site dissimilarity indices were first proposed by Diserud and Ødegaard (2007), and make it possible to consider species information shared among more than three units using the inclusion–exclusion principle (Erickson, 1996). Since then, various multiple-unit beta-diversity indices have been proposed (Baselga et al., 2007, Baselga, 2013b), including several new multiple-unit indices based on pairwise dissimilarity indices (Baselga et al., 2007, Baselga, 2010, 2012, 2017). Specifically, these multiple-unit dissimilarity indices share the A (a), B (b), and C (c) components of indices based on abundance (incidence). a_{mu}, b_{mu}, and c_{mu} correspond to a, b, and c, respectively, of incidence-based pairwise dissimilarity indices.

Although both b_{mu} and c_{mu} are simply sums of all components divided into two parts in the pairwise indices, a_{mu} is based on the inclusion–exclusion principle described above and is difficult to understand intuitively. For example, in the case of three units, the component a_{mu} can be calculated as follows (Baselga et al., 2007, Baselga, 2010):

$$a_{12} + a_{23} + a_{13} - \min(a_{12}, a_{23}, a_{13}) = \sum_{j<k} a_{jk} - a_{123} \quad \text{equation 9}$$

where a_{jk} is the number of species common to both units. Considering a general case with n units, the component a_{mu} can be formulated as follows:

$$\sum_{j<k} a_{jk} - \sum_{j<k<l} a_{jkl} + \sum_{j<k<l<m} a_{jklm} - \cdots . \quad \text{equation 10}$$

Equation 9 can be converted based on the inclusion–exclusion principle, and the resulting

components are as follows:

$$a_{mu} = \sum_j S_j - S_{\mathrm{T}} \qquad \text{equation 11a}$$

$$b_{mu} = \sum_{j<k} \min(p_{jk}, p_{kj}) \qquad \text{equation 11b}$$

$$c_{mu} = \sum_{j<k} \max(p_{jk}, p_{kj}) \qquad \text{equation 11c}$$

where S_j is the total number of species at site j, S_T is the total number of species at all sites considered together, and p_{jk} and p_{kj} are the numbers of species exclusive to units j and k, respectively, when compared pairwise (Baselga, 2010). Specifically, p_{jk} is the number of individuals included in unit j but not in unit k, and *vice versa*.

Similar to the case of a_{mu} described above, the component A_{mu} can be calculated for n units as follows (Baselga, 2010):

$$\sum_i^S \left[\sum_{j<k} \min\,(x_{ij}, x_{ik}) - \sum_{j<k<l} \min(x_{ij}, x_{ik}, x_{il}) + \sum_{j<k<l<m} \min(x_{ij}, x_{ik}, x_{il}, x_{im}) - \sum_{j<k<l<m<n} \min(x_{ij}, x_{ik}, x_{il}, x_{im}, x_{in}) \cdots,\right]. \qquad \text{equation 12}$$

Applying the inclusion–exclusion principle allows the A_{mu} component to be simplified. A_{mu}, B_{mu}, and C_{mu} correspond to A, B, and C, respectively, from abundance-based pairwise dissimilarity indices:

$$A_{mu} = T_{\mathrm{AB}} - \sum_i^S \max\,(x_{ij}, x_{ik}, x_{il}, x_{im}, x_{in} \ldots) \qquad \text{equation 13a}$$

$$B_{mu} = \sum_{j<k} \min\,\left(\sum_i^S \left[x_{ij} - \min\,(x_{ij}, x_{ik})\right], \sum_i^S \left[x_{ik} - \min\,(x_{ij}, x_{ik})\right]\right) \qquad \text{equation 13b}$$

$$C_{mu} = \sum_{j<k} \max\,\left(\sum_i^S \left[x_{ij} - \min\,(x_{ij}, x_{ik})\right], \sum_i^S \left[x_{ik} - \min\,(x_{ij}, x_{ik})\right]\right) \qquad \text{equation 13c.}$$

where T_{AB} is the total abundance in the data set (details see Baselga, 2017).

Multiple-unit indices for Sørensen and Bray–Curtis are formulated as follows (Baselga, 2010, 2013b),

$$d_{sor.mu} = \frac{b_{mu}+c_{mu}}{2a_{mu}+b_{mu}+c_{mu}} \qquad \text{equation 14}$$

$$d_{BC.mu} = \frac{B_{mu}+C_{mu}}{2A_{mu}+B_{mu}+C_{mu}}. \qquad \text{equation 15}$$

Novel individual-based multiple-unit beta-diversity indices

In this study, I developed individual-based components and novel indices for multiple-unit data sets. In accordance with existing methods of calculating multiple-unit dissimilarity, the persistence (P_{mu}) component can be calculated for three time steps as follows:

$$\min(z_{o1}, z_{o2}) + \min(z_{o2}, z_{o3}) + \min(z_{o1}, z_{o3}) - \min\,(z_{o1}, z_{o2}, z_{o3}). \quad \text{equation 16}$$

Considering a general case with n units, the component P_{mu} can be formulated as

$$\sum_i^I \Big[\sum_{j<k} \min\,(z_{oj}, z_{ok}) - \sum_{j<k<l} \min\big(z_{oj}, z_{ok}, z_{ol}\big) + \sum_{j<k<l<m} \min\big(z_{oj}, z_{ok}, z_{ol}, z_{om}\big) - \sum_{j<k<l<m<n} \min\big(z_{oj}, z_{ok}, z_{ol}, z_{om}, z_{on}\big) \cdots, \Big]. \quad \text{equation 17}$$

Using the inclusion–exclusion principle, P_{mu} can be simplified and the components P_{mu}, M_{mu} and R_{mu} reformulated as follows:

$$P_{mu} = \sum_j I_j - I_{\mathrm{T}} \quad \text{equation 18a}$$

$$M_{mu} = \sum_{\mathrm{j}<k} \min\big(q_{jk}, q_{kj}\big) \quad \text{equation 18b}$$

$$R_{mu} = \sum_{\mathrm{j}<k} \max\big(q_{jk}, q_{kj}\big). \quad \text{equation 18c}$$

where I_j is the total number of individuals in a unit j, I_T is the total number of individuals in all units considered together, and q_{jk} and q_{kj} are the numbers of individuals exclusive to units j and k, respectively, when compared pairwise.

The multiple-unit individual-based index is formulated as follows:

$$d_{MR.mu} = \frac{M_{mu}+R_{mu}}{2P_{mu}+M_{mu}+R_{mu}}. \quad \text{equation 19}$$

Furthermore, as an extension of equation 8, I also introduced $v_{\mathrm{s.mu}}$, which indicates the compositional variability in a community relative to the individual turnover, as follows:

$$v_{s.mu} = d_{BC.mu} \div d_{MR.mu}. \quad \text{equation 20}$$

297

298 ***Novel randomization criteria to construct null models of individual-tracked monitoring***
299 ***data***

300 Here, I introduce a novel method for constructing a null model to evaluate the degree of
301 individual persistence over time. Figure 2 is a conceptual diagram outlining the basis of the
302 new null model. Cases (i) and (ii) shown in Fig. 2 illustrate the patterns of individual
303 persistence of two individuals. The degree of individual persistence differed between the
304 two cases, even if the numbers of individuals involved in recruitment and mortality were
305 equal across the time steps. In both cases, one individual was recruited, and one individual
306 died. Specifically, in case (i), the individual z_1 persists over time, whereas individual z_2 was
307 recruited between T1 and T2 and died between T2 and T3. On the other hand, in case (ii)
308 described below, individual z_1 died between T2 and T3, and z_2 was recruited between T1 and
309 T2 and persisted, at least until the end of observation at T4. Such differences in the patterns
310 of individual persistence lead to differences in the value of the newly developed
311 individual-based temporal multiple-unit dissimilarity index ($d_{MR.mu}$), even when the numbers
312 of individuals undergoing mortality and recruitment are equal. The component P_{mu} can
313 account for differences in the pattern of individual persistence. Specifically, P_{mu} decreases,
314 and the index $d_{MR.mu}$ increases, when the number of long-lived individuals (i.e., more than
315 three time steps) increases, and *vice versa*. The null model approach introduced here uses
316 those properties of the component and the index.

317 In the novel null model, simulations using the same numbers of both recruited and
318 mortal individuals as the original data, but mortal individuals were randomly selected in
319 each time step. The simulations were repeated (e.g., 100 times) in each plot. The index
320 $d_{MR.mu}$ can be calculated from each repeated simulation. By comparing the observed and
321 simulated values of the index $d_{MR.mu}$, the standardized effect size of the degree of deviation in
322 individual persistence from the null model ($P_{ses.all}$) can be calculated as the observed value
323 minus the mean of the null distribution, divided by the standard deviation in the null
324 distribution. Calculating the deviation from the null model is commonly performed to
325 express biological differences regardless of the units of the indices (e.g., McCabe et al., 2012,
326 Nakadai & Kawakita, 2016). The value of $P_{ses.all}$ is close to 0, indicating random deaths of
327 individuals (i.e., ecological drift). In cases in which the value of $P_{ses.all}$ is greater than 0, the

proportion of long-lived individuals (more than three time steps) is larger than that in the null model.

Case study: the permanent 50-ha forest dynamics plot on BCI in Panama

To demonstrate the use of the new individual-based multiple-unit indices, I analyzed the data from Condit et al. (2019). The permanent 50-ha forest dynamics project plot on BCI in Panama was established in 1981 by Robin Foster and Stephen Hubbell (Hubbell & Foster, 1983). Data from this plot have been used in many scientific papers, and descriptions of the area and survey methods are widely available (e.g., Legendre & Condit, 2019). This 50-ha plot has been subjected to eight detailed surveys since its inception: in 1982, 1985, 1990, 1995, 2000, 2005, 2010, and 2015. The BCI data were divided into 1,250 quadrats (20 m × 20 m). Harms et al. (2001) classified quadrats into six habitat zones and a zone of mixed habitats (Fig. 3a, Table S1).

First, I calculated the values of the two individual-based beta-diversity indices for all possible pairs of time points and then determined their average values ($\overline{d_{MR}}$, $\overline{v_s}$) for each quadrat. In addition, I calculated the newly developed multiple-unit indices ($d_{MR.mu}$, $v_{s.mu}$) for each quadrat and tested their relationships using simple linear regression to characterize the novel indices. I predicted that the calculated average values would always be lower than the novel multiple-unit indices based on the results of Baselga (2013a), who compared the patterns of the average values of conventional beta-diversity indices and multiple-unit indices. Second, because high individual turnover can result in high compositional variability within a community, I tested the relationship between newly developed multiple-unit indices ($d_{MR.mu}$) and two previous ones ($d_{BC.mu}$, $d_{sor.mu}$). Third, I calculated the degree of deviation in individual persistence from the null model ($P_{ses.all}$) based on 100 randomizations via a novel random permutation method (details described above) for each quadrat. I analyzed differences among habitat types in compositional variability ($d_{BC.mu}$), the degree of individual turnover ($d_{MR.mu}$), compositional variability in individuals undergoing turnover ($v_{s.mu}$), and the degree of deviation in individual persistence from the null model ($P_{ses.all}$) using the Wilcoxon rank sum test, which allows for pairwise comparisons between groups with correction for multiple testing. The p-values were adjusted by the Holm method.

Finally, because the patterns of individual turnover and persistence may affect community variability, I tested the relationship between the degree of deviation in individual persistence from the null model ($P_{ses.all}$) and the conventional multiple-unit dissimilarity index ($d_{BC.mu}$) using a generalized additive model fitted by the restricted maximum likelihood (REML) method.

All analyses were conducted in R software (ver. 4.0.3; R Development Core Team, 2020). The *betapart* R package (Baselga & Orme, 2012) was used to calculate multiple-unit versions of the Bray–Curtis and Sørensen dissimilarity indices. The R package *mgcv* (Wood, 2017) was used to fit the generalized additive model to the relationship between the degree of deviation in individual persistence from the null model ($P_{ses.all}$) and the conventional multiple-unit dissimilarity index ($d_{BC.mu}$). The function *pairwise.wilcox.test* was used to compare the differences among habitats.

Results

Simple linear regression showed a significant correlation between $\overline{d_{MR}}$ and $d_{MR.mu}$ (*adjusted* $R^2 = 0.971$, $p < 0.0001$; Fig. S1a). A significant correlation between $\overline{v_s}$ and $v_{s.mu}$ (*adjusted* $R^2 = 0.899$, $p < 0.0001$; Fig. S1b) was also found. Both multiple-unit indices were higher than the average values of the pairwise indices at all times. $d_{BC.mu}$ and $d_{sor.mu}$ increased significantly with increasing $d_{MR.mu}$ (*adjusted* $R^2 = 0.756$, $p < 0.0001$; Fig. S2a and *adjusted* $R^2 = 0.444$, $p < 0.0001$, respectively; Fig. S2b).

Here, I summarize the results of the Wilcoxon rank sum test. The value of $d_{BC.mu}$ was significantly higher in group 4 (swamp) than in the other groups (Fig. 3b, f, Table S2a). In addition, no significant difference was found between groups 1 and 3, while the value of group 2 differed significantly from those of groups 1 and 3 (Fig. 3b, f, Table S2a). The $d_{MR.mu}$ values showed a similar pattern to that of $d_{BC.mu}$. $d_{MR.mu}$ was significantly higher in group 4 (swamp) than in the other groups (Fig. 3c, g, Table S2b). No clear difference was observed between groups 1 and 3, but the value of group 2 differed significantly from those of groups 1 and 3 (Fig. 3c, g, Table S2b). The $v_{s.mu}$ value was considerably and significantly higher in group 4 (swamp) than in all other habitats (Fig. 3d, h, Table S2c). In addition, no significant difference in $v_{s.mu}$ was found between groups 1 (low plateau) and 2 (high plateau), but the value of group 3 (slope) differed from those of groups 1 and 2 (Fig. 3d, h, Table S2c). There

were no significant differences in $P_{ses.all}$, among groups 4 (swamp), 5 (streamside), and 6 (young forest), but the values were significantly lower in groups 4 and 5 than in the other groups [i.e., group 1, 2, 3, and 7] (Fig. 3e, i, Table S2d). Moreover, no significant difference was found between groups 1 (low plateau) and 2 (high plateau), but group 3 (slope) differed significantly from groups 1 and 2 in terms of $P_{ses.all}$ (Fig. 3e, i, Table S2d).

In the generalized additive model, compositional variability ($d_{BC.mu}$) showed a significant response to the degree of deviation in individual persistence from the null model ($P_{ses.all}$: edf = 3.951, F = 31.59, *adjusted* R^2 = 0.112, p <0.0001), with the values of the compositional variability index constant from $P_{ses.all}$ of approximately 12 to 25 and tending to increase from $P_{ses.all}$ of approximately 12 to 0 (Fig. 4).

Discussion

Individual turnover is a major driver of compositional change and variability within a community, although it has rarely been discussed. Higher individual turnover clearly facilitates higher variability in community composition; thus, compositional variability over time can occur even when considering only ecological drift. Hubbell's unified neutral theory of biodiversity (Hubbell, 2001) postulates ecological neutrality among species within communities and emphasizes the role of ecological drift in shaping biodiversity patterns. To elucidate compositional variability, the impact of individual turnover (i.e., ecological drift) should be excluded. However, almost no methods exist to quantify the impact of individual turnover on compositional variability. The individual-based indices introduced herein can assess compositional variability considering the contribution of individual turnover. In addition, information about individual identity across multiple time steps provides opportunities to test individual persistence. Within and among species, some individuals live longer (i.e., individuals persist in the community for a longer period) than others. The patterns of individual persistence can be tested for their degree of deviation from expected patterns due to stochastic drift using the novel null model introduced here. This model can add a new assessment axis to empirical studies of community dynamics.

In the present study, I applied both conventional and novel indices, as well as the degree of deviation based on the null model to the BCI plot. The average values of both $\overline{d_{MR}}$ and $\overline{v_s}$ across all pairs are always lower than multiple-unit indices (i.e., $d_{MR.mu}$ and $v_{s.mu}$), in

accordance with the findings of Baselga (2013a). The results of the case study clearly demonstrated that the novel individual-based index was significantly correlated with conventional incidence- and abundance-based indices (Fig. S2); however, variations among communities still exist. The value of $v_{s.mu}$, which represents compositional variability after considering individual turnover, differed among habitat zones, especially for group 4 (swamp), which differed significantly from all other habitats (Fig. 3a, c, Table S1). Both the degree of compositional variability ($d_{BC.mu}$) and individual turnover ($d_{MR.mu}$) were significantly higher in group 4 than in all other habitats. There was significant compositional variability in the swamp even after controlling for individual turnover. Thus, individual turnover is a major driver, but a large component of variability remains unexplained by individual turnover. The environmental stability of each plot is a possible source of the unexplained effect. For example, Nakadai (2020) identified compositional changes after consideration of the individual turnover that was clearly associated with the degree of climate change (i.e., rate of temperature change) across the Japanese Archipelago. Compositional change over time in this swamp habitat has also been reported in terms of responses to a drought event at the BCI plot (Legendre and Condit 2019). In addition, although $d_{BC.mu}$ and $d_{MR.mu}$ were significantly higher in groups 1 (low plateau) and 3 (slope) than in group 2 (high plateau), $v_{s.mu}$ did not follow the same pattern. Instead, $v_{s.mu}$ was significantly lower in groups 1 (low plateau) and 2 (high plateau) than in group 3 (slope). These results suggest that the high plateau (group 2) has lower individual turnover, probably because it is isolated and stable and consequently has lower compositional variability. However, when considering individual turnover, the slope (group 3) differs from both the low and high plateau (groups 1 and 2).

The degree of deviation in individual persistence was lower in group 4 (swamp) than in groups 1, 2, 3, and 7, but not group 5 (streamside) or 6 (young forest), suggesting those three habitat types (i.e., swamp, streamside, and young forest) may be approaching expected mortal patterns due to ecological drift. Moreover, based on the generalized additive model, compositional variability was associated with the degree of deviation in individual persistence from the null model (Fig. 4). This suggests that an increase in the proportion of the long-lived individuals (i.e., higher values of $P_{ses.all}$) reduces compositional variability, even if the level of individual turnover remains the same. In previous studies, individual

persistence was recognized in community dynamics as a factor facilitating stable coexistence among species and buffering population growth (i.e., the storage effect) (Chesson & Warner, 1981, Chesson, 2000). The effect of buffered population growth is expected to be low compositional variability and turnover of individuals over time throughout the community. This prediction is consistent with the results related to $P_{ses.all}$. In other words, habitat types with higher $P_{ses.all}$ (except groups 4, 5, and 6) were better stabilized by the storage effect.

Although temporal beta diversity has attracted much attention, especially at large spatial scales (Brice et al., 2019, Magurran et al., 2019), it is simply an extension of spatial beta diversity. Both methodological and conceptual frameworks rooted in community dynamics should be used. Above all, information from "individuals" has been absent from empirical studies of temporal dynamics in community ecology and biodiversity. On the other hand, recent advances in modern coexistence theory (Chesson 2000) suggest that the impact of individuals across generations (e.g., individual persistence of adult trees) is a storage effect that facilitates species coexistence and stabilizes community dynamics. In this context, the method proposed here is expected to be useful for addressing a wide range of research questions related to temporal changes in biodiversity, especially studies using forest monitoring and bird banding data to track individuals, by bridging gaps between recent theoretical advances and accumulated empirical data.

Author contributions

R.N. conceived the study, designed the methods described herein, analyzed the empirical data, and wrote the manuscript.

Acknowledgments

Financial support was provided by the Japan Society for the Promotion of Science (grant 18J00093).

Data Availability

Data are available from the Dryad Digital Repository (https://doi.org/10.15146/5xcp-0d46;

accessed on 10th of April in 2020) (Condit et al., 2019).

References

Anderson, M.J., Crist, T.O., Chase, J.M., Vellend, M., Inouye, B.D., Freestone, A.L., Sanders, N.J., Cornell, H.V., Comita, L.S., Davies, K.F., Harrison, S. P., Kraft, N.J.B., Stegen, J. C. & Swenson, N. G. (2011). Navigating the multiple meanings of β diversity: a roadmap for the practicing ecologist. *Ecol. Lett.,* 14, 19–28.

Baselga, A. (2010). Partitioning the turnover and nestedness components of beta diversity. *Glob. Ecol. Biogeogr.*, 19, 134–143.

Baselga, A. (2012). The relationship between species replacement, dissimilarity derived from nestedness, and nestedness. *Glob. Ecol. Biogeogr.*, 21, 1223–1232.

Baselga, A. (2013a). Multiple site dissimilarity quantifies compositional heterogeneity among several sites, while average pairwise dissimilarity may be misleading. Ecography, 36, 124-128.

Baselga, A. (2013b). Separating the two components of abundance–based dissimilarity: balanced changes in abundance vs. abundance gradients. *Methods Ecol. Evol.*, 4, 552–557.

Baselga, A. (2017). Partitioning abundance-based multiple-site dissimilarity into components: Balanced variation in abundance and abundance gradients. *Methods Ecol. Evol*., 8, 799–808.

Baselga, A. & Orme, C. D. L. (2012). *betapart*: an R package for the study of beta diversity. *Methods Ecol. Evol.*, 3, 808–812.

Baselga, A., Jiménez-Valverde, A., & Niccolini, G. (2007). A multiple-site similarity measure independent of richness. *Biol. Lett*., 3, 642–645.

Bray, J.R., & Curtis, J.T. (1957). An ordination of the upland forest communities of southern Wisconsin. *Ecol. Monogr.*, 27, 325–349.

Brice, M.H., Cazelles, K., Legendre, P., & Fortin, M.J. (2019). Disturbances amplify tree community responses to climate change in the temperate–boreal ecotone. *Glob. Ecol. Biogeogr.*, 28, 1668–1681.

Condit, R., Perez, R., Aguilar, S., Lao, S., Foster, R., Hubbell, S.P. (2019). Complete data from the Barro Colorado 50-ha plot: 423617 trees, 35 years, 2019 version., Dryad,

Dataset, https://doi.org/10.15146/5xcp-0d46

Chesson, P.L., & Warner, R.R. (1981). Environmental variability promotes coexistence in lottery competitive systems. *Am Nat*, 117, 923-943.

Chesson P (2000) Mechanisms of maintenance of species diversity. *Annu. Rev. Ecol. Evol. Syst.*, 31, 343-366

Diserud, O.H. & Ødegaard, F. (2007). A multiple-site similarity measure. *Biol. Lett.*, 3, 20–22.

Erickson, M.J. (1996). Introduction to Combinatorics. Wiley, New York, NY, USA.

Gaston K.J, Davies, R.G., Orme, C.D.L., Olson, V.A., Thomas, G.H., Ding, T.S., Rasmussen, P.C., Lennon, J.J., Bennett, P.M., Owens, I.P.F. & Blackburn, T. M. (2007). Spatial turnover in the global avifauna. *Proc. R. Soc. B* 274, 1567–1574.

Gotelli, N. J., & Graves, G.R. (1996). Null models in ecology. Smithsonian Inst. Press.

Gotelli, N.J., & McGill, B.J. (2006). Null versus neutral models: what's the difference? *Ecography*, 29, 793-800.

Hallett, L. M., Hsu, J. S., Cleland, E. E., Collins, S. L., Dickson, T. L., Farrer, E. C., ... & Suding, K. N. (2014). Biotic mechanisms of community stability shift along a precipitation gradient. *Ecology*, 95, 1693-1700.

Hallett, L. M., Jones, S. K., MacDonald, A. A. M., Jones, M. B., Flynn, D. F., Ripplinger, J., ... & Collins, S. L. (2016). CODYN: An R package of community dynamics metrics. *Methods Ecol. Evol.*, 7, 1146-1151.

Harms, K.E., Condit, R., Hubbell, S. P., & Foster, R. B. (2001). Habitat associations of trees and shrubs in a 50-ha neotropical forest plot. *J. Ecol*., 89, 947–959.

Hatosy, S.M., Martiny, J.B., Sachdeva, R., Steele, J., Fuhrman, J.A., & Martiny, A.C. (2013). Beta diversity of marine bacteria depends on temporal scale. *Ecology*, 94, 1898–1904.

Hubbell, S.P. (2001). The unified neutral theory of biodiversity and biogeography. Princeton University Press.

Hubbell, S.P. & Foster, R.B. (1983). Diversity of canopy trees in a neotropical forest and implications for conservation. In: Whitmore, T., Chadwick, A., & Sutton, A. (eds). Tropical Rain Forest: Ecology and Management. The British Ecological Society, Oxford, pp. 25–41.

Jost, L., Chao, A. & Chazdon, R. (2011). Compositional similarity and beta diversity.

Biological Diversity: Frontiers in Measurement and Assessment (eds A. Magurran & B.J. McGill), Oxford University Press, Oxford, pp. 66–84.

Kalyuzhny, M. (2020). Null models for community dynamics: Beware of the cyclic shift algorithm. *Global Ecol. Biogeogra.*, 29, 1085-1093.

Koleff, P. & Gaston, K.J. (2002). The relationships between local and regional species richness and spatial turnover. *Global Ecol. Biogeogr.*, 11(5), 363–375.

Koleff, P., Gaston, K.J., & Lennon, J.J. (2003). Measuring beta diversity for presence–absence data. *J. Anim. Ecol.*, 72, 367–382.

Legendre, P. & Condit, R. (2019). Spatial and temporal analysis of beta diversity in the Barro Colorado Island forest dynamics plot, Panama. *For. Ecosyst.*, 6, 7.

Legendre, P. & De Cáceres, M. (2013). Beta diversity as the variance of community data: dissimilarity coefficients and partitioning. *Ecol. Lett*, 16, 951–963.

Legendre, P. & Gauthier, O. (2014). Statistical methods for temporal and space–time analysis of community composition data. *Proc. Roy. Soc. B.*, 281, 20132728.

Legendre, P. & Legendre, L. (2012). Numerical Ecology, 3rd edn. Elsevier, Amsterdam.

Legendre, P. (2019). A temporal beta-diversity index to identify sites that have changed in exceptional ways in space–time surveys. *Ecol. Evol.*, 9, 3500–3514.

McCabe, D. J., Hayes-Pontius, E.M., Canepa, A., Berry, K.S., and Levine, B.C. (2012). Measuring standardized effect size improves interpretation of biomonitoring studies and facilitates meta-analysis. *Freshw. Sci.,* 31, 800–812.

Magurran, A.E. (2011). Measuring biological diversity in time (and space). *Biological Diversity: Frontiers in Measurement and Assessment* (eds A. Magurran & B.J. McGill), pp. 85–104. Oxford University Press, Oxford.

Magurran, A.E., Dornelas, M., Moyes, F., & Henderson, P. A. (2019). Temporal β diversity–A macroecological perspective. *Glob. Ecol. Biogeogr.*, 28, 1949–1960.

Mori, A.S., Isbell, F., & Seidl, R. (2018). β-diversity, community assembly, and ecosystem functioning. *Trends Ecol. Evol.*, 33, 549–564.

Nakadai, R. (2020). Degrees of compositional shift in tree communities vary along a gradient of temperature change rates over one decade: application of an individual-based temporal beta-diversity concept. *Ecol. Evol.*, 10, 13613–13623.

Nakadai, R., & Kawakita, A. (2016). Phylogenetic test of speciation by host shift in leaf

cone moths (*Caloptilia*) feeding on maples (*Acer*). *Ecol. Evol*., 6, 4958-4970.

Odum, E.P. (1950). Bird populations of the Highlands (North Carolina) Plateau in relation to plant succession and avian invasion. *Ecology*, 31, 587–605.

R Development Core Team (2020). R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. http://www.R-project.org/.

Shimadzu, H., Dornelas, M., & Magurran, A.E. (2015). Measuring temporal turnover in ecological communities. *Methods Ecol. Evol.*, 6, 1384–1394.

Sørensen, T. A. (1948). A method of establishing groups of equal amplitude in plant sociology based on similarity of species content and its application to analyses of the vegetation on Danish commons. *Biol. Skar.*, 5, 1–34.

Whittaker, R. H. (1960). Vegetation of the Siskiyou mountains, Oregon and California. *Ecol. Monogr.*, 30, 279–338.

Whittaker, R.H. (1972). Evolution and measurement of species diversity. *Taxon*, 21, 213–251.

Wood, S.N. (2017) Generalized additive models: An introduction with R (2nd edition). Chapman and Hall/CRC.

Legends

Figure 1 Comparison of species composition and individuals (20 individuals of four species total) among temporal units, showing the components (A, B, C, M, R, P, E_{loss}, *and* E_{gain}) used for assessing compositional variability over time. The squares, circles, pentagons, and hexagons indicate species 1, 2, 3, and 4, respectively. The numbers under each element (e.g., 12) represent the corresponding time step. For components A and P, the number 123 is also shown, indicating the shared species and individuals across the three time steps, respectively. Detailed explanations of each component are provided in Table 1.

Figure 2 Conceptual diagram showing the basis of the new null model. Both cases (i) and (ii) showed patterns of individual persistence for two individuals. The number of persistent individuals differed between the two cases, even if the numbers of individuals affected by recruitment and mortality were the same at each time step. Here, in both cases, one individual was recruited, and one individual died. Specifically, in case (i), individual Z_1 persisted over time, but individual Z_2 was recruited between times 1 and 2 and died between times 2 and 3. On the other hand, in case (ii), individual Z_1 died between times 2 and 3, and individual Z_2 was recruited between times 1 and 2 and continued to persist at least until time 4. The average value of the individual-based temporal beta-diversity index ($\overline{d_{MR}}$), which indicates the degree of individual turnover, may be the same in both cases, but the newly developed multiple-unit dissimilarity index will have different values ($d_{MR.mu}$) because the component P_{mu} can account for differences in the pattern of individual persistence. Specifically, P_{mu} decreases, and the index $d_{MR.mu}$ increases, when the number of long-lived individuals (i.e., those persisting for more than three time steps) increases. The sprout and the skull and crossbones symbols indicate recruitment and mortality, respectively.

Figure 3 Comparison of the conventional multiple-unit Bray–Curtis dissimilarity ($d_{BC.mu}$; (b) and (f)), individual-based temporal beta-diversity dissimilarity ($d_{MR.mu}$; (c) and (g)), composition variability ($v_{s.mu}$; (d) and (h)), and the degree of deviation in individual persistence from the null model ($P_{ses.all}$; (e) and (i)) among habitat zones (a) using the pairwise Wilcoxon rank sum test, which allows for pairwise comparisons between groups with correction for multiple testing. Habitat types labeled with the same letter did not

significantly differ in terms of that index based on the Wilcoxon rank sum test with Holm's method for multiple comparisons. Detailed information about the habitat zones and results of the Wilcoxon rank sum test are shown in Tables S1 and S3, respectively. In (a), (b), (c), (d), and (e), the horizontal and vertical axes indicate easting and northing, respectively.

Figure 4 The relationship between the degree of deviation in individual persistence from the null model ($P_{ses.all}$) and the conventional multiple-unit dissimilarity index ($d_{BC.mu}$). The solid black line was drawn using a generalized additive model fitted using the restricted maximum likelihood method (*adjusted* $R^2 = 0.112$, $p < 0.0001$, $edf = 3.951$, $F = 31.59$). The blue dashed lines represent the 95% confidence intervals.

Figure 1

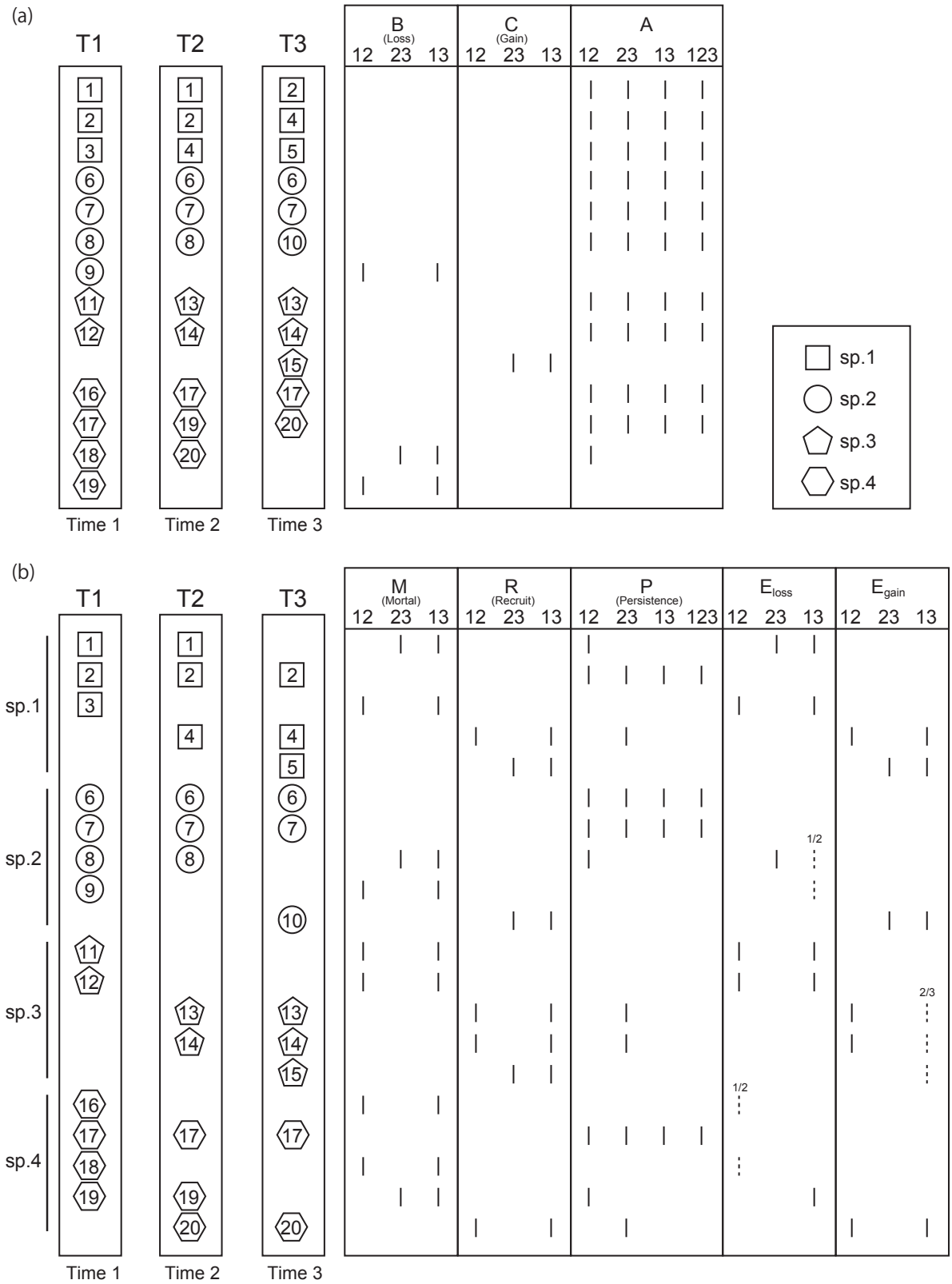

(a)
T1
T2
T3
B (Loss)
C (Gain)
A
12 23 13
12 23 13
12 23 13 123
Time 1
Time 2
Time 3
sp.1
sp.2
sp.3
sp.4
(b)
M (Mortal)
R (Recruit)
P (Persistence)
Eloss
Egain
1/2
2/3

Figure 2

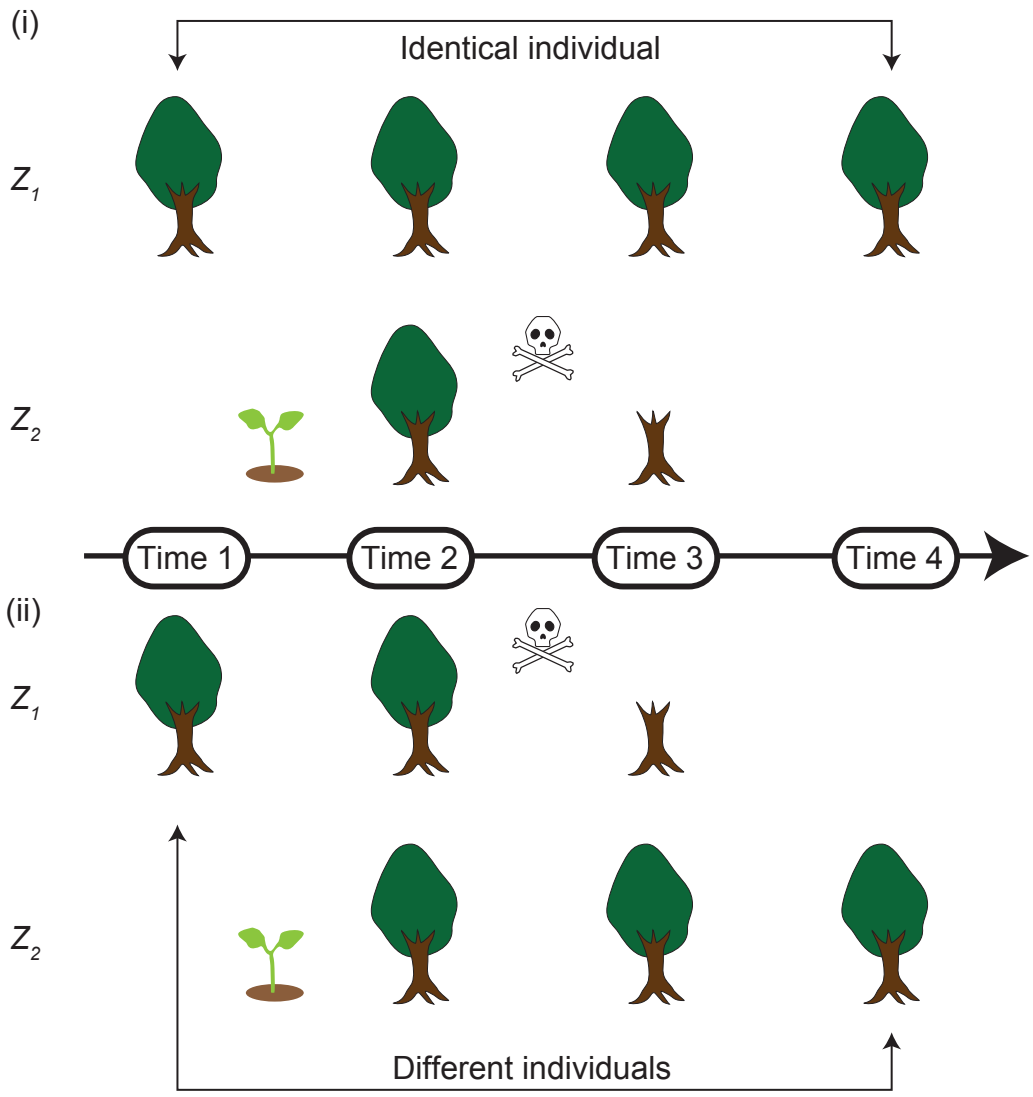
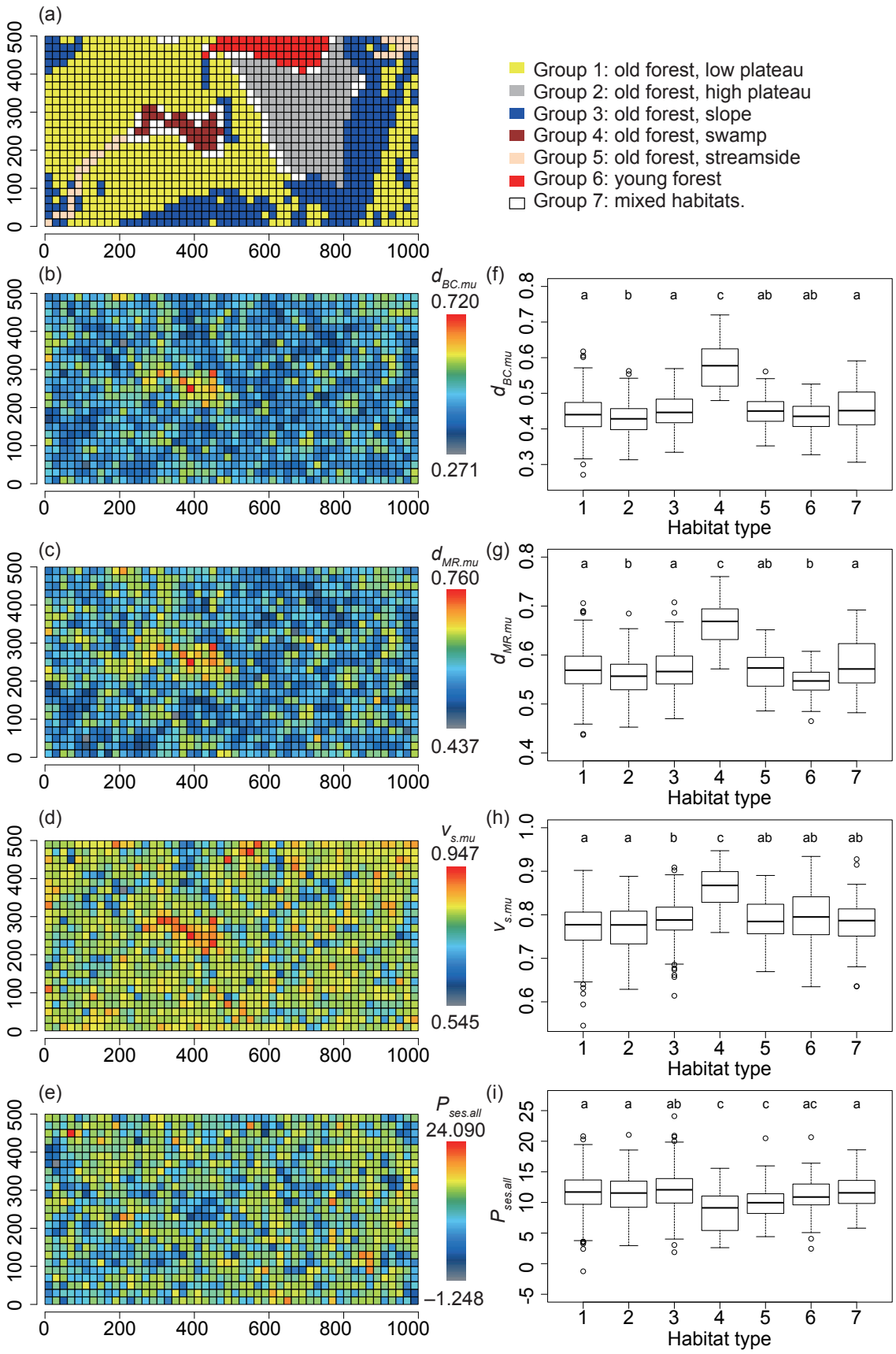

(i)
Identical individual
Z_1
Z_2
Time 1
Time 2
Time 3
Time 4
(ii)
Z_1
Z_2
Different individuals

Figure 3



645

646

Figure 4

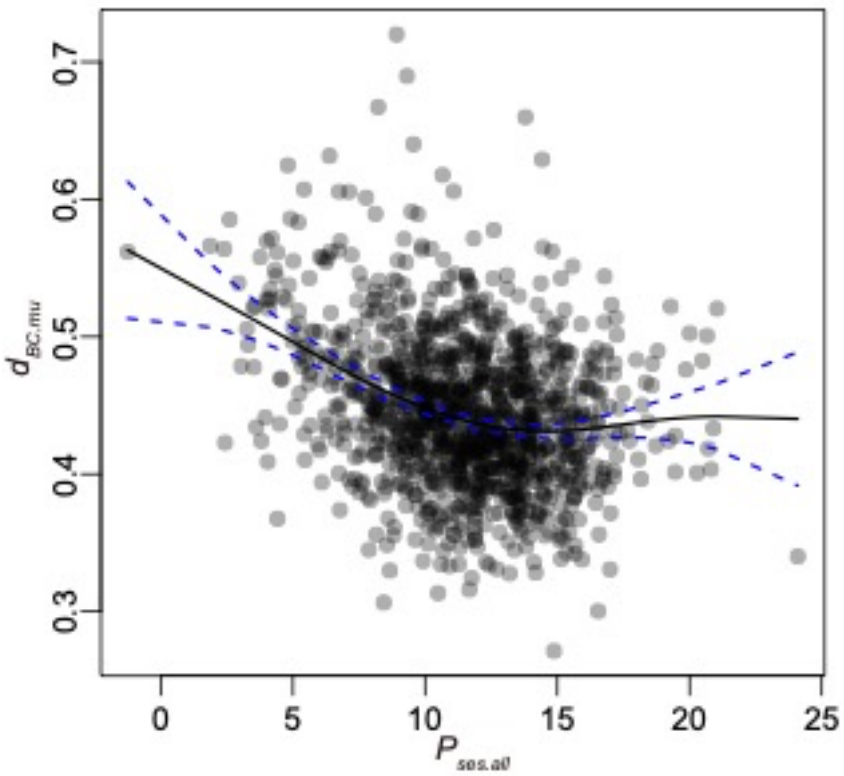


Table 1 Explanations of the components used in the conventional and novel indices of beta diversity. A visualization of those components is provided in Fig. 1.

Component	Explanation
A	The number of individuals of a species common to both units j and k
B	The number of individuals of a species unique to unit j, absent from unit k ($j < k$)
C	The number of individuals of a species unique to unit k, absent from unit j ($j < k$)
M	The total number of individuals unique to unit j, absent from unit k ($j < k$)
R	The total number of individuals unique to unit k, absent from unit j ($j < k$)
P	The total number of individuals common to both units j and k
E_{loss}	The total number of lost individuals that were replaced by conspecifics, calculated as $M - B$
E_{gain}	The total number of gained individuals replacing conspecifics, calculated as $R - C$

Appendices

Figure S1 Correlations between novel multiple-unit indices and the average values of conventional pairwise indices ((a) d_{MR}-$d_{MR.mu}$; *adjusted* $R^2 = 0.971$, $p < 0.0001$, (b) v_s-$v_{s.mu}$; *adjusted* $R^2 = 0.899$, $p < 0.0001$)**.**

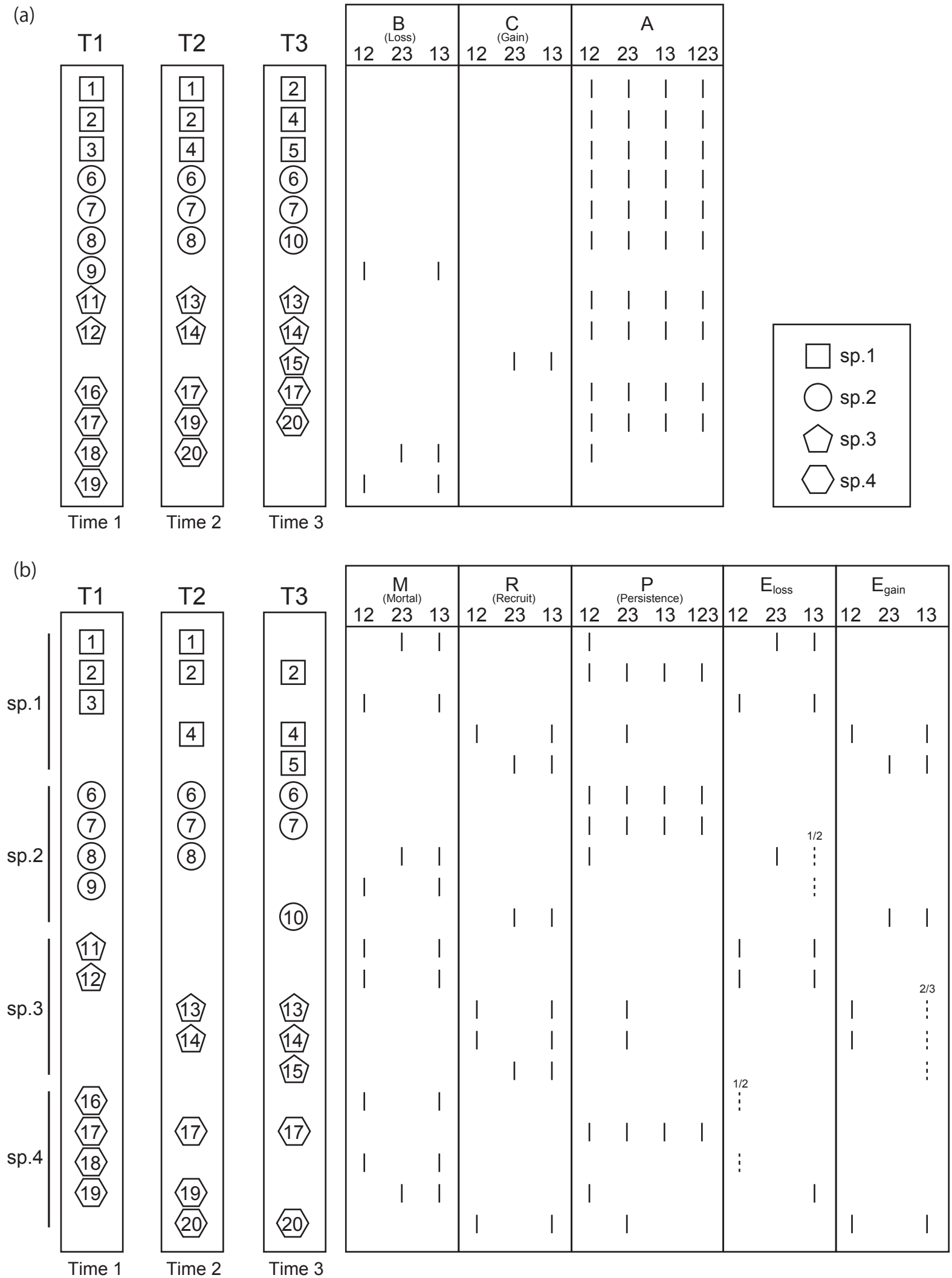
Figure S2 Correlations between a novel individual-based multiple-unit index ($d_{MR.mu}$) and conventional multiple-unit indices ($d_{BC.mu}$) ((a) $d_{BC.mu}$-$d_{MR.mu}$; *adjusted* $R^2 = 0.756$, $p < 0.0001$, (b) $d_{sor.mu}$-$d_{MR.mu}$; *adjusted* $R^2 = 0.444$, $p < 0.0001$).

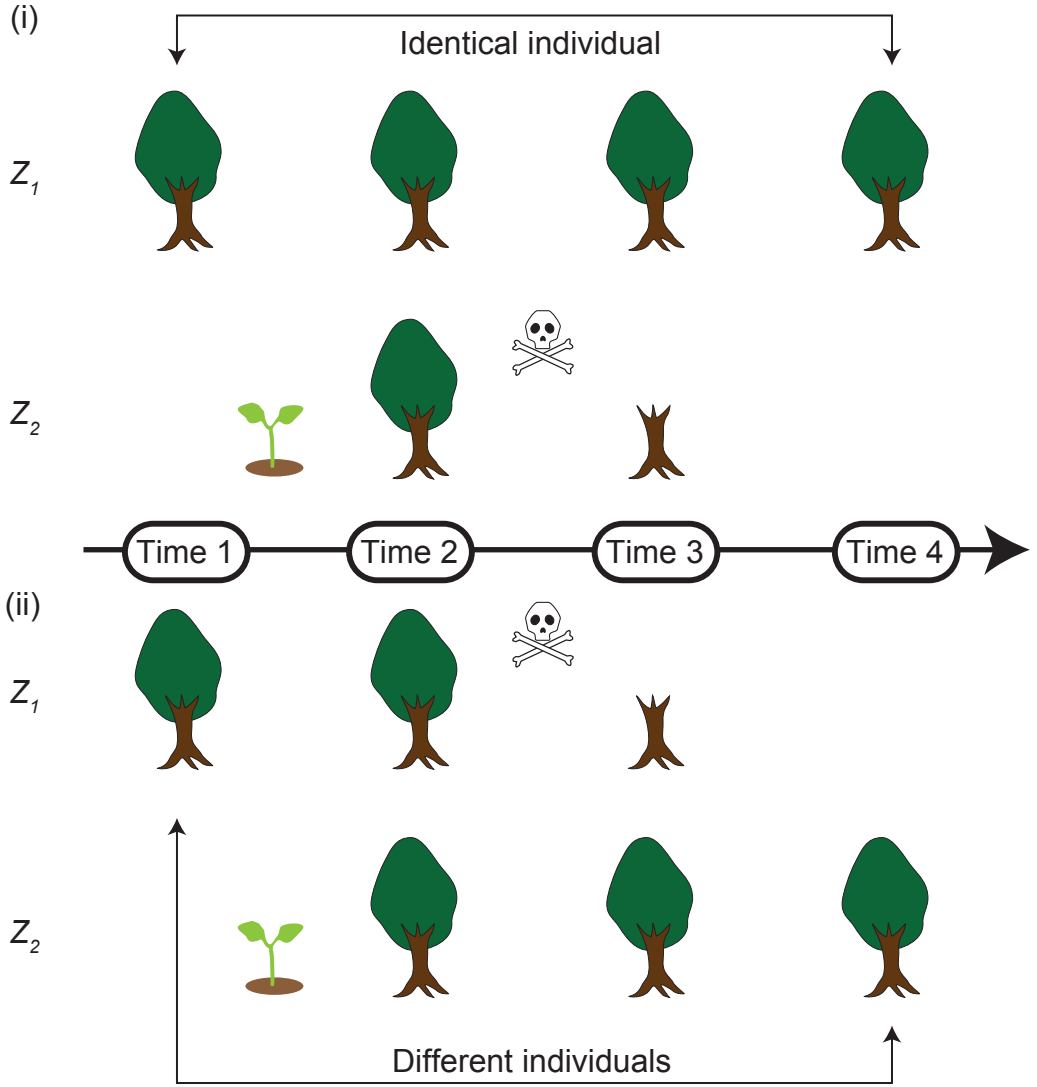
Figure S3 Examples of patterns of individual persistence with the corresponding values of $P_{ses.all}$. The x-axis indicates the number of time periods of individual persistence. The y-axis indicates the number of individuals persisting. (a) The case of Quadrat No. 1221 ($P_{ses.all}$ = 2.403), (b) the case of Quadrat No. 0322 ($P_{ses.all}$ = 24.090).

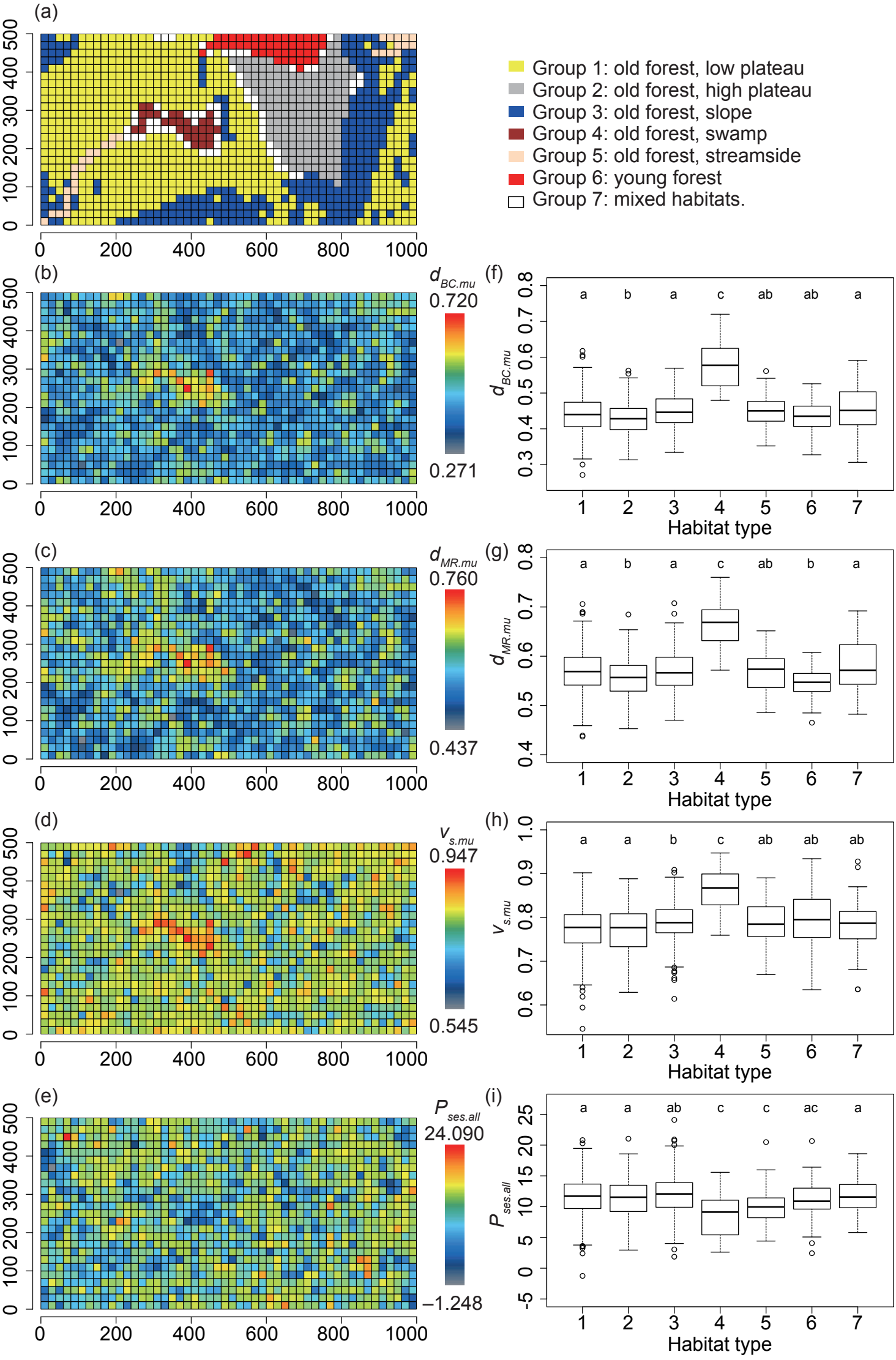
Table S1 Habitat clarification in Harms et al. (2001). Group 1: old forest, low plateau; group 2: old forest, high plateau; group 3: old forest, slope; group 4: old forest, swamp; group 5: old forest, streamside; group 6: young forest; and group 7: mixed habitats.

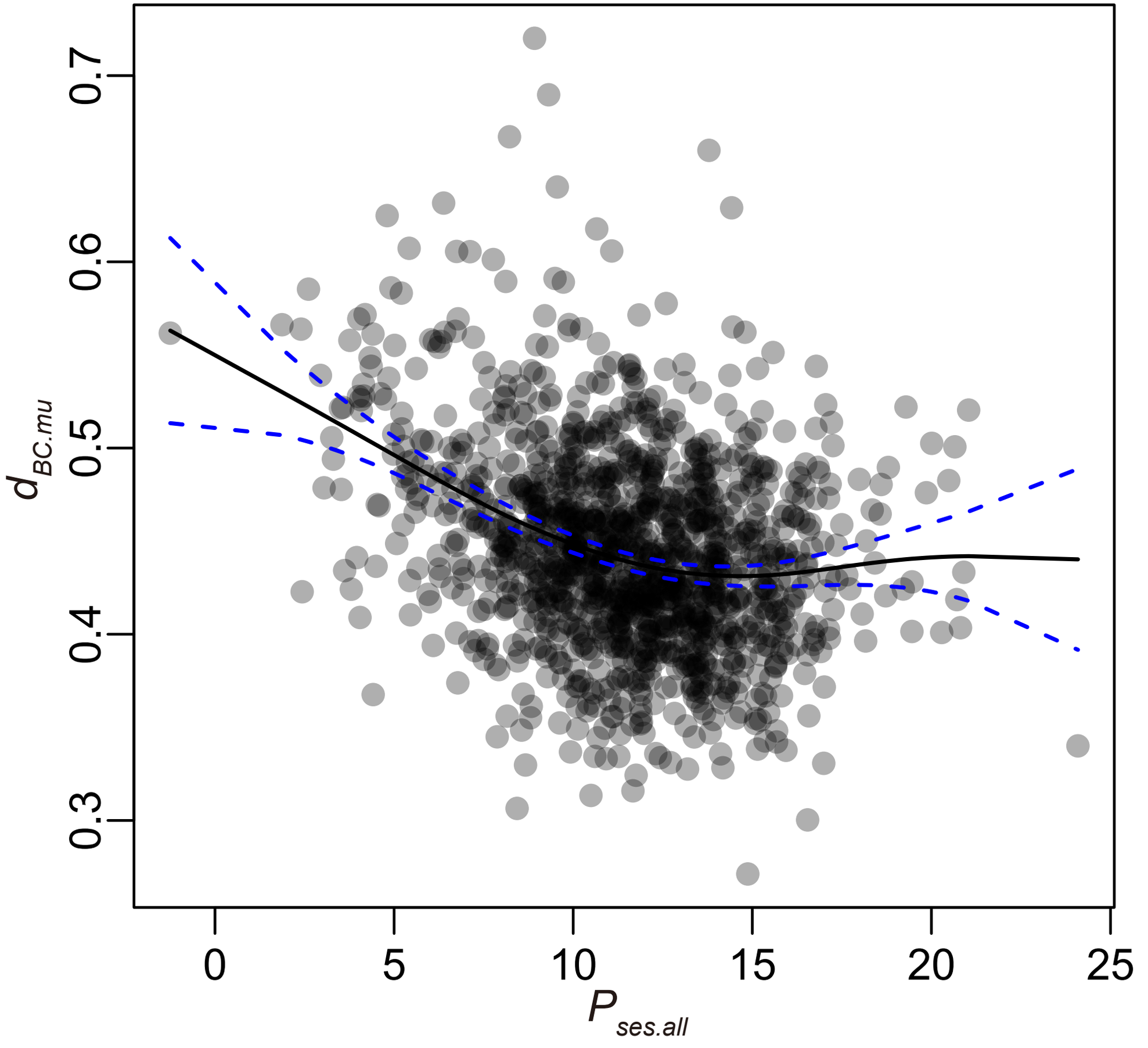
Table S2 Comparisons of the conventional multiple-unit Bray–Curtis dissimilarity ($d_{BC.mu}$: (a)), the individual-based temporal beta-diversity dissimilarity index ($d_{MR.mu}$: (b)), composition variability ($v_{s.mu}$: (c)), and the degree of deviation in individual persistence from the null model ($P_{ses.all}$: (d)) among habitat zones using the pairwise Wilcoxon rank sum test, which allows for pairwise comparisons between groups with correction for multiple testing. The tables correspond to (f), (g), (h), and (i) in Figure 3.

Supplementary File 1 R script for the individual-based multiple-unit beta-diversity calculations performed using sample data.


(a)
T1
T2
T3
Time 1
Time 2
Time 3
B (Loss)
C (Gain)
A
12 23 13
12 23 13
12 23 13 123
sp.1
sp.2
sp.3
sp.4
(b)
sp.1
sp.2
sp.3
sp.4
M (Mortal)
R (Recruit)
P (Persistence)
Eloss
Egain
1/2
2/3


(i)
Identical individual
Z_1
Z_2
Time 1
Time 2
Time 3
Time 4
(ii)
Z_1
Z_2
Different individuals


(a)
Group 1: old forest, low plateau
Group 2: old forest, high plateau
Group 3: old forest, slope
Group 4: old forest, swamp
Group 5: old forest, streamside
Group 6: young forest
Group 7: mixed habitats.
(b)
$d_{BC.mu}$
0.720
0.271
(f)
$d_{BC.mu}$
a b a c ab ab a
Habitat type
(c)
$d_{MR.mu}$
0.760
0.437
(g)
$d_{MR.mu}$
a b a c ab b a
Habitat type
(d)
$v_{s.mu}$
0.947
0.545
(h)
$v_{s.mu}$
a a b c ab ab ab
Habitat type
(e)
$P_{ses.all}$
24.090
−1.248
(i)
$P_{ses.all}$
a a ab c c ac a
Habitat type


0.7
0.6
0.5
0.4
0.3
$d_{BC.mu}$
0
5
10
15
20
25
$P_{ses.all}$