bioRxiv preprint doi: https://doi.org/10.1101/527028; this version posted January 22, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

Transmission bias's fundamental role in biodiversity change

Abstract

Understanding changes in biodiversity is a complex subject contingent on many interacting or poorly differentiated processes. As a result, it is desirable to organize processes in community ecology into a small number of high-level mechanisms that completely account for change in ecological communities. It has been suggested that all change in ecological communities can be partitioned into four mechanisms: 1) selection, 2) drift, 3) immigration and 4) speciation; however, the completeness of this framework requires testing. Here we use insights derived from one of the most fundamental equations in evolution, the Price equation, to quantify the strength of selection, drift, immigration and speciation in simulations. We show how the impacts of each of these mechanisms can be quantified using experimental and simulated data and find that these four mechanisms cannot account for large portions of the change in these simulated communities. This gap is a consequence of a fifth mechanism fundamental to evolutionary theory, transmission bias, which describes change in the measurements associated with organisms. Examples of transmission bias from evolution include phenotypic plasticity and selection within groups. Our results highlight the distinction between biodiversity change and the processes that change species' relative abundances. Selection, drift, immigration and speciation change species' relative abundances, but many diversity measures summarize an additional piece of information: a measurement of species' rarity. Species' rarity changes over time, and these changes are not included when the influence of selection, drift, immigration and speciation are

bioRxiv preprint doi: https://doi.org/10.1101/527028; this version posted January 22, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

quantified. Accounting for changes in species' rarity by adding transmission bias to our list of fundamental mechanisms leads to a complete accounting of biodiversity change.

Keywords: Price equation; theoretical ecology; evolutionary theory; biodiversity;

Introduction

One of the fundamental goals of ecology is to understand the origins and maintenance of biodiversity. Unfortunately, many mechanisms can shape biodiversity and it can be difficult to distinguish their effects. A promising avenue of research is to organize this complexity into a small number of high-level mechanisms that are easier to study (Vellend 2010, Scheiner and Willig 2011, Vellend 2016). A particularly influential example of this approach is the "four mechanisms" framework (Chase and Myers 2011, Weiher et al. 2011, Costello et al. 2012a, Hanson et al. 2012). Inspired by lists of high-level mechanisms in evolution, this framework attributes change in ecological communities to: 1) selection, the change in the frequency of species due to the tendency of some species to contribute more offspring to subsequent generations (Figure 1 A); 2) drift, the change in the frequency of species due to random variation in the number offspring contributed by each species (Figure 1 A); 3) dispersal, the change in frequency due to the arrival of individuals from other locations (Figure 1 B); and 4) speciation, the evolution of new species (Figure 1 C). This verbal model has great intuitive appeal, and has been adopted by researchers on topics as different as macroecology (Pontarp et al. 2018) and the human microbiome (Costello et al. 2012b). However it has not been clear how the effects of each

bioRxiv preprint doi: https://doi.org/10.1101/527028; this version posted January 22, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

mechanism can be quantified (Vellend 2016), and that they completely describe biodiversity change.

One of the most general ways to quantify mechanisms in evolutionary theory is in terms of change in average measurements of the organisms under study. In this interpretation an entire assemblage of organisms is broken down into separate types and we study the change in the average measurement across types. In evolution the types are often genotypes; in community ecology, types can be species (Collins and Gardner 2009, Genung et al. 2011, Norberg et al. 2012, Govaert et al. 2016). Each type is associated with a distinct measurement such as: body size, beak size, crop yield, or disease resistance. Nearly any quantitative attribute that is associated with organism types can count as a measurement (Price 1995, Frank 2012b). At any point in time, we can calculate the average measurement across all members of the community. Evolutionary mechanisms change this average (Price 1970, Price 1972, Rice 2004, Queller 2017, Lehtonen 2018). This interpretation of evolutionary theory has already been used to define three of the "four mechanisms". The effect of selection is defined as an association between fitness (the per-capita birth rate of a type divided by the average per capita birth rate for all types) and the measurement of each type at the start of the observation period (Robertson (1966). For example, Grant and Grant (2002) quantified natural selection on beak size of Darwin's finches by measuring the association between the beak sizes of different finch types (beak size being the relevant measurement) and the fitness of these types. Price (1970) generalized this approach, producing what is commonly known as the Price equation (Rice 2004, Frank 2012b, Queller 2017). This equation can be re-interpreted to analyse drift (Rice 2004 citation), and an extension derived by Kerr and Godfrey-Smith (2009) can be used to consider a community that can receive immigrants.

bioRxiv preprint doi: https://doi.org/10.1101/527028; this version posted January 22, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

When fitness is defined in terms of the number of offspring produced by each type, there is no current approach to incorporate speciation into the Price equation. Rankin et al. (2015) showed that some long standing questions in macroevolution can be answered by re-defining fitness as in terms of speciation rate rather than birth rate. This useful approach is not applicable to numerous problems in community ecology which concern change in species' frequency due to differences in the number of offspring produced. For example, consider a remote island observed from 2010 to 2015. Over this time, no speciation occurred, but an invasive species produced far more descendants than native species and so increased in frequency from 10 % of the community to 90 % of the community. The change in frequency is dramatic but Rankin's method will register no selection because no speciation has occurred.

To quantify the effect of each of the "four mechanisms" on biodiversity change we will emphasize one fundamental interpretation of measures of biodiversity: averages of measurements of species' rarity (Jost 2006). To take one example: Simpson's diversity index can be interpreted as probability of intraspecific encounter (a measure of rarity since members of common species are likely to encounter conspecifics while members of rare species are unlikely to encounter conspecifics), averaged across all individuals in the community (Hurlbert 1971). In this respect rarity can be understood as a measurement associated with each species in a community, subject to change through time. The interpretation of diversity as an average of rarity is fundamental to theories of biodiversity indices (Hill 1973, Jost 2006, 2007), just as the notion of change in averages is fundamental to evolutionary theory (Price 1970, Frank 2012c, b, Queller 2017).

A major gap remains to be filled before we can analyse biodiversity change using four mechanisms framework. The four mechanisms are thought to completely describe change in

bioRxiv preprint doi: https://doi.org/10.1101/527028; this version posted January 22, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

"any community property one might wish to examine" (Vellend 2016, section 4.4), but evolutionary theory recognizes one high level mechanism that has no analogue in the "four mechanisms" framework: transmission bias (Price 1970, Frank 2012c, b, Queller 2017). Transmission bias describes change in the measurement associated with each type of organism. One familiar example of transmission bias is when offspring tend to be larger than parents of the same type (Rice 2004). This effect can matter tremendously, say if drought stunts the yield of each type of plant in a meadow, reducing the productivity of the community, in addition to any effects of drift, selection, immigration or speciation.

To resolve this problem, we review an established interpretation of biodiversity indices as measurements of rarity averaged over all members of an ecological community. We show how evolutionary theory can be re-interpreted to partition change in this average. We use this approach to quantify the effects of selection, drift, immigration, and speciation on biodiversity change. When this approach is used, there is a discrepancy between total change and the combined effects of selection, drift, immigration and speciation, due to transmission bias (Price 1970, Kerr and Godfrey Smith). We show that transmission bias can strongly influence biodiversity and that it represents the effect of changes in species' rarity over time. To completely account for change in biodiversity, we propose that transmission bias be treated as a fundamental process alongside selection, drift, immigration and speciation.

The model

Selection and biodiversity change

Both evolutionary theory and diversity indices are concerned with species' frequencies: the abundance of species i (n_i) divided by the total number of individuals in the community

bioRxiv preprint doi: https://doi.org/10.1101/527028; this version posted January 22, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

$\left(p_i = \frac{n_i}{\sum_i n_i}\right)$. This quantity is also called relative abundance or proportional abundance. Many diversity measures are based on sums of functions of the species frequencies (Jost 2010). Such indices can be expressed as averages of species' rarity weighted by species' frequency:

(1) $D = \sum_i p_i z_i$,

where z_i is a measure of species' rarity (and hence a function of p_i). For simplicity we will refer to such measures as measures of biodiversity. We will use a trick to visualize the meaning of this (and subsequent) equations (Figure 2). D can be thought of as the area contained by *i* rectangles where each rectangle has a width (p_i) and a height z_i. The larger the area contained by all *i* rectangles the greater the total diversity of the community, the larger the area of the *i*th rectangle, the greater the contribution of the *i*th species to D.

Many diversity indices can be expressed in the same form as Equation 1, or expressed as functions of Equation 1. For example, when studying Shannon entropy, the relevant measure of rarity is $z_i = -log(p_i)$. For this measure, high frequency (p_i) species have low rarity (z_i) values and low frequency species have high rarity. Recognizing this distinction between p_i and z_i helps to clarify why frequency change is described by one mechanism (selection) while rarity change is described by another (transmission). When we think about Shannon entropy's applications in fields other than ecology there is an intuitive reason for this distinction between frequency and rarity. Shannon entropy is a measure of information averaged over many events and z_i measures the information associated with event *i*, rare events encode a great deal of information (Frank 2012c). Here, we present worked examples involving hand calculations which we simplify by calculating logs in base 2, rather than the base e commonly employed by ecologists.

bioRxiv preprint doi: https://doi.org/10.1101/527028; this version posted January 22, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

When studying Simpson's diversity index the relevant measure of rarity is $z_i = p_i$. For ease of interpretation we will present Gini Simpson's index (i.e. 1- Simpsons's Diversity index), since Gini Simpson's index has the intuitive property that high values imply high diversity. Many other indices can be written as averages of species rarity including: species richness, the Berger-Parker index, and the Hurlbert-Smith-Grassle index for a sample size of 2. Other diversity measures are functions of such averages including Reyni entropies, Tsallis entropies, and the diversity of an equivalent number of uniformly distributed species i.e. Hill numbers (Hill 1973, Jost 2007).

Change in diversity can be thought of as the difference between *D* in one-time step and *D* in a subsequent time step:

(2) $\Delta(\sum_i p_i z_i) = \sum_i p_i' z_i' - \sum_i p_i z_i$

We will label the past frequency of species *i* as p_i and present frequency as p_i'. Similarly, the past measurement of rarity is z_i and the present measurement of rarity is z_i'. Difference is denoted Δ. We examine immigration, speciation and drift in the next section. But for now, we restrict ourselves to the case where the same species are present in the past and the present (eliminating the possibility of immigration and speciation) and we assume no sampling variation in the number of offspring produced by each species (eliminating drift). This definition of change works over any timescale (i.e. it is irrelevant whether the past observation was 1 day ago or one million years ago). By the same token the definitions of selection and other mechanisms that we define work over any timescale, though the importance of each mechanism can vary (Frank 2012a).

Under these assumptions, total change can be partitioned into two fundamental terms:

bioRxiv preprint doi: https://doi.org/10.1101/527028; this version posted January 22, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

$$(3)\ \overbrace{\Delta(\textstyle\sum_i p_i z_i)}^{Change} = \overbrace{\textstyle\sum_i \Delta p_i z_i}^{Selection} + \overbrace{\textstyle\sum_i p_i'(z_i' - z_i)}^{Transmission}$$

The selection term changes species' frequencies (Δp_i) based on the rarity of each species in the past time step (z_i). Selection does depend on the rarity of species in the present time step. Transmission bias can be thought of as the change in species' rarity ($z_i' - z_i$) weighted by the frequency of species in the present time step p_i'. Partitioning of change in Equation 3 into selection and transmission is mathematically exact, and well established in evolutionary theory (Frank 2012b, Queller 2017).

It is worth pausing here for a moment to think about why selection is not enough to completely describe change in diversity. Selection changes species' frequencies between the past and the present. Selection in the past acted on species' measurements in the past (i.e. it acted on species' rarity in the past). Selection in the past did not act on species' rarity in the present (selection is powerful, but not psychic). In contrast, diversity change does depend on the present rarity of each species. Thus, there is a gap between the information represented by diversity change and the information that was available to selection when it acted. Transmission bias fills this gap.

Transmission bias's role as a gap filler between the frequency of species in the past and the frequency of species in the present is particularly clear when we measure Shannon entropy. In this case the transmission bias reduces to a standard information theoretic measure for the difference between probability distributions, the Kullback Leibler divergence between the distribution of species frequencies in the past and the distribution of species frequencies in the present, multiplied by -1 (i.e. the transmission bias term reduces to $-\sum_i p_i' log\left(\frac{p_i'}{p_i}\right) = -D(p_i'||p_i)$).

bioRxiv preprint doi: https://doi.org/10.1101/527028; this version posted January 22, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

When studying Simpson's diversity, transmission represents change in the probability of a randomly chosen individual encountering another individual of the same species. When we assume that change is slow ($p_i' \approx p_i$) then transmission bias term $\sum_i p_i' \Delta p_i$ is nearly equal to the selection term $\sum_i \Delta p_i p_i$. With these observations of the importance of transmission bias in mind, the next step is to analyse the role of transmission bias when drift, immigration and speciation also operate.

Partition of the mechanisms changing biodiversity

Transmission bias also emerges when immigration, drift and speciation are included. Figure 3 provides an example of transmission and immigration jointly changing diversity. Using an extension of the Price equation we can account for change due to each mechanism. To facilitate indexing we first divide the community into three categories, 1) individuals who descended from members of the same species in the past community (the frequency of such individuals is ω), 2) individuals who immigrated to the community since the past observation period (the frequency of such individuals is μ), and 3) individuals who belong to a species which have emerged since the last observation period (the frequency of such individuals is σ). We assume that all individuals that arrive in the community are either immigrants or belong to a new species, such that $\omega + \mu + \sigma = 1$. This leads to the following formula for diversity in the present community:

(4) $D' = \omega \sum_i p_i' z_i' + \mu \sum_j a_j z_j^* + \sigma \sum_k \beta_k z_k^{\#}$

bioRxiv preprint doi: https://doi.org/10.1101/527028; this version posted January 22, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

$p_i{}'$ now represents the frequency of species *i* among individuals in the present community descended from individuals that were present in the past observation period a_j is the frequency of immigrants belonging to species *j* among all immigrants. β_k is the frequency of individuals belonging to new species *k* among all individuals belonging to new species. The z values in Equation 4 are measurements of the rarity of a given species in the present community. Thus $z_i{}'$ is the rarity of the ith descendant species, z_j^* is the rarity of the jth immigrant species and $z_k^\#$ is the rarity of the kth new species. Keep in mind that different indexes might refer to the same species (say if some European rabbits (*Oryctolagus cuniculus*) in the present community are immigrants while others are descendants of past community members). The rarity of a species then depends on the proportion of the present community belonging to that species. For example, when measuring Simpson's diversity the appropriate measure of rarity is $z_i' = z_j^* = \omega p'_{rabbits} + \mu a_{rabbits}$. When measuring Shannon entropy, the appropriate measure of rarity is $z_i' = z_j^* = -\log(\omega p'_{rabbits} + \mu a_{rabbits})$.

Change in diversity can now be partitioned using the "Extended Price equation":

$$(5)\ \Delta D = \underbrace{\omega \sum_i \Delta p_i z_i}_{\substack{lineage\ sorting \\ (selection \\ and\ drift)}} + \underbrace{\mu\left(\sum_j a_j z_j^* - \sum_i p_i z_i\right)}_{immigration} + \underbrace{\sigma\left(\sum_k \beta_k z_k^\# - \sum_i p_i z_i\right)}_{speciation} + \underbrace{\omega \sum_i p_i{}' \Delta z_i}_{\substack{transmission \\ bias}}$$

Derivations are found in Kerr and Godfrey-Smith (2009), and Frank (2012b). The lineage sorting term is analogous to the selection term in Equation 3, it still describes the difference between the frequency of individuals in species *i* in the past and the frequency of the descendants associated with species *i* in the present (Δp_i). However, we recognize that in the real-world species' frequencies change because of two interlinked causes: differences in the relative fitness

bioRxiv preprint doi: https://doi.org/10.1101/527028; this version posted January 22, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

of each type and sampling variation in the number of offspring produced by each type. Drift describes changes in species frequencies when the relative fitness of all species is identical, but random variation in the number of offspring produced in each time steps alters species' frequencies. Selection describes changes in species frequencies due to some species producing relatively more offspring than others (Rice 2004).

It is easy to measure the joint effects of drift and selection, but it is difficult to tease apart their individual contributions. Rice (2004) clarifies this issue by pointing out that both drift and selection produce the same consequences—changes in the frequencies of descendants among species—but what separates these two mechanisms are their causes. Selection results from differences in the number of offspring produced by each type averaged over many replicates. In contrast drift represents sampling variation in the number of offspring produced by each type in a given community. In an individual community it will be difficult to determine when the increase in frequency of one species was due to higher fitness, random variation in the number of offspring produced or some combination of the two. Over numerous replicated experiments drift could in principle be identified as cases where all species have the same relative fitness, averaged across many replicates and regardless of their frequency. We know of no study which has distinguished drift and selection in using a Price equation framework similar to the one we have used, and it is quite likely that existing analyses of "selection" among species implicitly include drift (Collins and Gardner 2009, Rankin et al. 2015). Interested readers are referred to Engen and Sæther (2014) who distinguish drift and selection mathematically.

The effects of immigration depend on the proportion of individuals in the present community that are immigrants (μ), and the difference between the average rarity of immigrants in the present, and the average rarity of the past community. Similarly, the effects of speciation

bioRxiv preprint doi: https://doi.org/10.1101/527028; this version posted January 22, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

on changes in community properties depend on the proportion of fs that belong to new species (σ), and the difference between the average rarity of new species in the present, and the average rarity of the past community.

Our treatment of speciation in the Price equation is new, and so requires some justification. Essentially we treat species identity in the way described by Vellend (2010): "Species identity is a categorical phenotype assumed to have perfect heritability, except when speciation occurs, after which new species identities are assigned." When using the Price equation, a natural way to assign new identities is to treat members of new species as new arrivals to the community.

Transmission bias concerns the proportion of the community whose ancestors were present in the community in the past time step (ω). Species *i*'s contribution to transmission bias depends on the proportion of descendants belonging to species *i* ($p_i{}'$). It also depends on the change in rarity of species i (Δz_i). This change in rarity can reflect indirect consequences of all the other mechanisms. For example, in Figure 2 transmission bias emerges as an indirect consequence of selection changing the rarity of moose and squirrel, while in Figure 3 transmission bias emerges because the arrival of immigrants changes the rarity of other species.

Numbers equivalent of diversity indices

Thus far we have shown that transmission bias is needed for a general description of change in simple, familiar diversity indices. The units of these indices can be difficult to interpret. For example, there are cases where a large change in the number of species can result

bioRxiv preprint doi: https://doi.org/10.1101/527028; this version posted January 22, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

in a surprisingly small change in the corresponding diversity index (Jost 2006, 2007). To remedy this problem there are calls to move from studying diversity indices to the number of uniformly distributed species needed to produce the observed diversity index. We will refer to these metrics as the numbers equivalent of diversity indices.

Transmission bias's role in shaping the numbers equivalent of Shannon entropy is easiest to discern. The numbers equivalent of Shannon entropy is:

(6) $e^{\sum_i p_i z_i}$

Where $z_i = -log\ (p_i)$.We can derive an expression for change in the numbers equivalent of Shannon entropy by exponentiating both sides of Equation 5, and recalling that $e^{a-b} = \frac{e^a}{e^b}$, leaving us with:

(7) $\frac{e^{\sum_i p_i' z_i'}}{e^{\sum_i p_i z_i}} = e^{\omega \sum_i \Delta p_i z_i + \mu(\bar{z}_I - \bar{z}) + \sigma(\bar{z}_s - \bar{z}) + \omega \sum_i p_i' \Delta z_i}$

In words, the left-hand side is a measure of change in the numbers equivalent of diversity: the numbers equivalent of present Shannon entropy divided by the numbers equivalent of past Shannon entropy. The right-hand side is *e* raised to the exponent of the mechanisms changing the Shannon entropy. Transmission bias shapes change in numbers equivalent of Shannon entropy just as it does the Shannon entropy. We have simply adjusted the scale upon which we have measured these changes. Indeed, when change the numbers equivalent of Shannon entropy is not too large, Equation 7 can be approximated as:

(8) $\frac{e^{\sum_i p_i' z_i'}}{e^{\sum_i p_i z_i}} \approx 1 + \omega \sum_i \Delta p_i z_i + \mu(\bar{z}_I - \bar{z}) + \sigma(\bar{z}_s - \bar{z}) + \omega \sum_i p_i' \Delta z_i.$

bioRxiv preprint doi: https://doi.org/10.1101/527028; this version posted January 22, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

This indicates that change in the numbers equivalent of Shannon entropy (measured as the ratio of present measure to past measure) is approximately equal to change in Shannon entropy (measured as the difference between present entropy and past entropy) plus the constant 1. This approximation works because $e^x \approx 1 + x$ when $x \approx 0$.

More work will be needed to formally decompose change in numbers equivalent of other diversity indices, because these indices are not typically expressed as arithmetic means, as required by analyses using the Price equation. In any case we doubt that selection, drift, immigration, and speciation would be enough to account for change in numbers equivalent of other diversity indices since there will still be a need to account for the rarity of all species in the present, and as we have seen this information is often only accounted for by transmission.

Partitioning experimental change in diversity

Our approach translates readily to empirical measurements of change in diversity. We illustrate this using a simple two species experiment of interactions between two species of bacteria. As a part of a series of experiments (Remus-Emsermann et al. 2018) measured the frequency of two bacterial species *Escherichia coli* and *Pantoea eucalypti* on a leaf of mouse ear cress (*Arabidopsis thaliana*). Leaves were inoculated with sufficient *E. coli* to produce a population density of 2.5×10^6 colony forming units per 1 gram of tissue of *A. thaliana* (fresh weight) and with sufficient *P. eucalypti* to produce a population density of 4.95×10^6 colony forming units per 1 gram of tissue of *A. thaliana*. The population density was subsequently measured by assaying the number of colony forming units at 1, 3 and 7 days post infection. The experiment was established to assay the frequency with which *E. coli* obtain plasmids from *P.*

bioRxiv preprint doi: https://doi.org/10.1101/527028; this version posted January 22, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

eucalypti via conjugation, but conjugation events were not observed in the replicate and so are not considered here (Remus-Emsermann et al. 2018). No speciation was observed, and the experimental protocol was designed to exclude immigration. In this system lineage sorting is strong. Given the large ecological differences between the two species and the large population sizes we expect that selection outweighs drift, but to tease apart these factors we would need to measure the association between fitness and rarity across many replicates.

We use Equation 5 to partition change in diversity. To calculate Gini Simpson's index, we set the measurement of rarity to $z_i = p_i$. To partition change in the Shannon entropy set the measurement of rarity to $z_i = -\log(p_i)$. Here and elsewhere our measure of change in the numbers equivalent of Shannon entropy was not plotted since it is very close to 1 + change in Shannon entropy. To reflect the lack of immigration and speciation we set $\omega = 1, \mu = 0, \sigma = 0$.

Over the course of the experiment between *Escherichia coli* and *P. eucalypti*, diversity declined dramatically as lineage sorting favoured *P. ecalypti*, particularly between day 0 and day 1 (Figure 4 A). Gini Simpson's index was strongly influenced by transmission bias (Figure 4 B), notably between day 0 and day 1 where transmission bias reduced diversity more than selection did. On subsequent days lineage sorting and transmission had roughly equivalent effects on Gini Simpson's index. Transmission bias strongly influenced the Shannon entropy between days 0 and day 1, on subsequent time intervals its effect was less pronounced than lineage sorting (Figure 4 C).

Partitioning simulated change in diversity

To illustrate the importance of transmission bias when several mechanisms operate, we use modified versions of simulations Vellend used to illustrate his framework. Vellend (2016)

bioRxiv preprint doi: https://doi.org/10.1101/527028; this version posted January 22, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

presents R scripts to simulate change in communities due to due to drift, selection, immigration, and speciation (MacDonald and Vellend 2016). Most of these simulations consider two species communities and assume that the total number of individuals is fixed. This assumption has been demonstrated to lead to gaps in models of evolutionary change (Holt and Gomulkiewicz 1997, Day 2005, Mallet 2012). Indeed Vellend (2016, p 88) states that this assumption will be will be perceived by some as "unnecessarily crude". Instead, we explicitly modelled population growth using the familiar Ricker Model of competition (Ricker 1954, Otto and Day 2007, Luís et al. 2011). R scripts available upon request.

(9a) $$N_1' = N_1 \exp(r_1 - a_{11}N_1 - a_{12}N_2 + \varepsilon)$$

(9b) $$N_2' = N_2 \exp(r_2 - a_{22}N_2 - a_{21}N_1 + \varepsilon)$$

In this discrete time model, N_i denotes the number of individuals of species i in the previous generation, while N_i' is the number of individuals in the current generation. N_i' depends on the number of individuals of that species at time t multiplied by a term describing population growth. The r_i parameters describe the intrinsic growth rates of species i. The a_{ij} terms describe how interacting with species j decreases the growth rate of species i. These terms represent interspecific interactions when i and j differ and intraspecific interactions when i and j are the same. We have included stochasticity in our model by adding a noise term ε. As a result of this term, some drift is included in each of our simulations, though we set parameter values such that it is frequently overwhelmed by selection. In our two species simulations, ε is normally distributed with a mean of 0 and a standard deviation of 0.02.

We will illustrate five distinct scenarios. To simulate drift's contribution in the absence of other mechanisms (scenario 1), we set the intrinsic growth rates of the two species to be equal

bioRxiv preprint doi: https://doi.org/10.1101/527028; this version posted January 22, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

and set all competition coefficients to be equal (Adler et al. 2007). This neutral model can be modified to include frequency-independent selection (scenario 2) by setting the intrinsic growth rate of species 1 to be higher than that of species 2 (Mallet 2012), or to include negative frequency-dependent selection (scenario 3), by setting interspecific competition coefficients to be less than intraspecific competition coefficients (Mallet 2012).

To consider the consequences of immigration into a patch (scenario 4), we modelled dispersal among a metacommunity of three patches (labelled L=1,2,3). To do this we divided the life cycle of each species into two stages. The first stage represents local population growth in a patch using equations Equation 9a and 9b (with an added index to indicate patch number). The density of species i after local population growth is denoted $N_{i,L}^*$. The second stage described dispersal among patches (Hedrick 2011):

(10a) $N_{1,L}' = m(\Sigma_k N_{1,L}^*)/3 + (1-m)N_{1,L}^*$

(10b) $N_{2,L}' = m(\Sigma_k N_{2,L}^*)/3 + (1-m)\, N_{1,L}^*$.

In this community, a proportion m of individuals joined a regional species pool. An equal proportion (1/3) of these individuals moved to each patch. A proportion (1-m) of individuals remain in the patch of their birth (i.e. they do not immigrate).

Lastly, we examined speciation (Scenario 5) using a variant of Hubbell (2001)'s neutral model where biodiversity is governed by drift and speciation. This model consists of a fixed metacommunity of j individuals. At each time step individuals die with some probability. They are replaced by the offspring of another individual selected at random from other individuals in the community (representing drift) or by a new species (representing speciation). In this model we retained the assumption that community size is fixed, as the alternative would be to

bioRxiv preprint doi: https://doi.org/10.1101/527028; this version posted January 22, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

complicate the model with assumptions about the population growth parameters of 1000s of new species, which would be difficult to justify empirically.

Our simulations demonstrated that transmission bias can strongly affect change in Gini Simpson's index (Figure 5). Transmission bias was roughly as strong as drift in the simulation of including only drift (scenario 1) and the simulation of speciation (scenario 2). It is also roughly as strong as selection in simulations of frequency-independent selection and frequency-dependent selection (scenario 2 and 3). Transmission bias strongly affects change in Shannon entropy (Figure 5, fifth column) in our simulations of frequency-dependent selection (scenario 3) and speciation (scenario 5). Transmission bias is weaker in other simulations of change in Shannon entropy, notably drift (scenario 1) and immigration (scenario 4).

Discussion

Our central result is that a complete accounting of change in biodiversity requires transmission bias along with selection, drift, immigration and speciation. We have illustrated how transmission bias's effect can be quantified in experiments (Figure 4) and used simulations to highlight cases where transmission bias is likely to emerge in ecological communities (Figure 5). In addition, we show how the mechanisms shaping biodiversity can be quantified. This formalization retains much of the heuristic utility of the "four mechanisms" approach, but increases its rigor, precision, and generality. Below we highlight the benefits and limitations of a "five mechanisms" framework.

Transmission bias tends to influence diversity when the rarity of one or more species changes between the past and the present. In our empirical example, *E. coli* was moderately common in day 0 and became more rare on day 1 (Figure 4), which resulted in strong transmission bias. This is echoed in our simulation of frequency-independent selection where

bioRxiv preprint doi: https://doi.org/10.1101/527028; this version posted January 22, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

selection in favour of species 1 leads to strong transmission bias in the first 10 generations (Figure 5 row 1). Immigration will to lead to transmission bias when it alter species' rarity (Figure 3). When speciation changes the rarity of other species (Harmon et al. 2009) it will have similar effects. Changing the diversity index changes the way that rarity is measured and this in turn will alter the importance of transmission bias. With Gini Simpson's, selection and transmission tend to be similar in magnitude because both terms depend strongly on the proportion of a given species in the community. We have less intuition about when transmission bias strongly affects change in Shannon entropy, but when immigration and speciation are excluded, transmission bias is simply a measure of how much species frequencies have changed between the past and the present. When these change little, transmission bias is weak, whereas when these change substantially, transmission bias is strong.

More work is needed to determine the effect of transmission bias on diversity in nature. As we have shown with our empirical example, selection and transmission can be distinguished using data on species' frequencies alone. This suggests that the next step is to empirically measure when transmission bias strongly influences biodiversity change. To identify the effect of speciation, the only additional data we need is the proportion of the descendant community belonging to new species. To partition the effect of immigration, we also need to know the proportion of the descendant community belonging to each species who are immigrants (or have descended from immigrants since the past observation period). This can be challenging but is possible, for instance when molecular methods identify the origin of individuals, or when experimental designs manipulate the proportion of individuals moving from one patch to another. Drift and selection remain more difficult to tease apart (Rice 2004). One can test for drift by determining if the fitness of each species (averaged over many replicates) is independent

bioRxiv preprint doi: https://doi.org/10.1101/527028; this version posted January 22, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

of its rarity. This requires repeated observations under comparable conditions but the difficulty in studying drift reflects the current state of community ecology rather than an idiosyncrasy of our framework. What our framework does simplify is the partitioning of the combined effects of drift and selection from immigration, speciation and transmission.

Our work is foreshadowed by other applications of the Price equation in community ecology. The Price equation can be used to study a huge number of variables that concern community ecology including phenology (Weis et al. 2015), tolerance to toxins (Collins and Gardner 2009), body size (Fox and Harpole 2008), ecosystem services (Winfree et al. 2015), temperature optima (Norberg et al. 2012), and productivity (Loreau and Hector 2001). Nevertheless, to our knowledge no one has layed out the general connection between the "four mechanisms" framework and the Price equation (Table 1). As a result, we suspect that the many readers of the four mechanisms framework will miss transmission bias's fundamental role alongside selection, drift immigration, and speciation. Treating transmission bias as an equal to the other four mechanisms expands the range of phenomena easily studied by community ecology. Evolution within species can be treated as a form of transmission bias. This approach has been used for simulated and empirical analyses of change in ecological communities (Collins and Gardner 2009, Genung et al. 2011, Norberg et al. 2012, Govaert et al. 2016). In all cases this form of transmission bias was a substantial portion of change in ecological communities. Environmentally induced changes in species' attributes resulting from phenotypic plasticity (Price et al. 2003) can also be treated as a form of transmission bias. (Agrawal 2001, Dzialowski et al. 2003, Engel and Tollrian 2009, Ozgul et al. 2009, Helanterä and Uller 2010, Bolnick et al. 2011, Turcotte and Levine 2016). By including transmission bias in the list of fundamental processes we can quantify the effects of evolution or phenotypic plasticity, which are two

bioRxiv preprint doi: https://doi.org/10.1101/527028; this version posted January 22, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

interesting mechanisms that do not fit neatly into the categories of species level selection, drift, immigration, or speciation.

Focusing attention on transmission bias helps us to see our summaries of ecological communities as more than just a summary changes in species abundances. Many properties of communities summarize additional measurements of the natural world, such as biomass, leaf area, reflectance, transpiration. Many measurements of the natural world change over time. We miss these changes when we only quantify selection, drift, immigration and speciation.

Nevertheless, transmission bias will be non-existent in some circumstances because some measurements remain constant over time. For example, the Price equation has been used to measure selection for one species (Nowak 2006) by defining z_i to be an indicator variable taking a value of 1 for individuals belonging to the species of interest and 0 otherwise. Descendants always have the same value for this as their ancestors except in rare instances of speciation. As a result, transmission bias does not occur for this measurement. Some readers have asked us how transmission bias can be fundamental when it only occurs for some variables and some assumptions? Keep in mind that other fundamental mechanisms will be absent in some communities. Speciation is rare over short time-scales. There are no immigrants from outside of planet earth (so far as we know). When the volcano Krakatoa exploded in 1883 it obliterated all vertebrate life. Immigrants could replenish this species pool, but until they did, drift and selection were absent. All five mechanisms are fundamental in the sense that omitting any one will produce an incomplete description of change in biodiversity.

bioRxiv preprint doi: https://doi.org/10.1101/527028; this version posted January 22, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

Acknowledgements

This paper emerged out of a Google Hangouts group discussing Vellend (2016). We thank Mark Westoby, Thuy Nguyen, and Buck Trible, the other participants in our discussion group whose thoughts shaped the current article. We are particularly indebted to Margaret Kosmala who had the brilliant idea to create and organize Internet-based discussions of Vellend (2016). Our paper would not have been possible without her effort. Helpful comments were provided by Mark Vellend, Steve Frank, Komei Kadowaki, Erol Ackay, Adrian Paterson, Bevin Brett and Robert Cruickshank. Images of rabbit and squirrel based on illustrations from by Anthony Caravaggi, https://creativecommons.org/licenses/by-nc-sa/3.0/. All illustrations from phylopic http://phylopic.org/.

Literature cited

Adler, P. B., J. HilleRisLambers, and J. M. Levine. 2007. A niche for neutrality. Ecology Letters **10**:95-104.

Agrawal, A. A. 2001. Phenotypic plasticity in the interactions and evolution of species. Science **294**:321-326.

Bolnick, D. I., P. Amarasekare, M. S. Araújo, R. Bürger, J. M. Levine, M. Novak, V. H. Rudolf, S. J. Schreiber, M. C. Urban, and D. A. Vasseur. 2011. Why intraspecific trait variation matters in community ecology. Trends in ecology & evolution **26**:183-192.

Chase, J. M., and J. A. Myers. 2011. Disentangling the importance of ecological niches from stochastic processes across scales. Philosophical Transactions of the Royal Society of London B: Biological Sciences **366**:2351-2363.

Collins, S., and A. Gardner. 2009. Integrating physiological, ecological and evolutionary change: a Price equation approach. Ecology Letters **12**:744-757.

Costello, E. K., K. Stagaman, L. Dethlefsen, B. J. Bohannan, and D. A. Relman. 2012a. The application of ecological theory toward an understanding of the human microbiome. Science **336**:1255-1262.

Costello, E. K., K. Stagman, L. Dethlefsen, B. J. M. Bohannan, and D. A. Relman. 2012b. The Application of Ecological Theory Toward an Understanding of the Human Microbiome. Science **336**:1255.

Day, t. 2005. Modelling the ecological context of evolutionary change: Deja vu or something new? Page 273 Ecological Paradigms Lost: Routes of Theory Change. Elsevier.

Dzialowski, A. R., J. T. Lennon, W. O'Brien, and V. H. Smith. 2003. Predator-induced phenotypic plasticity in the exotic cladoceran Daphnia lumholtzi. Freshwater Biology **48**:1593-1602.

bioRxiv preprint doi: https://doi.org/10.1101/527028; this version posted January 22, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

Engel, K., and R. Tollrian. 2009. Inducible defences as key adaptations for the successful invasion of Daphnia lumholtzi in North America? Proceedings of the Royal Society of London B: Biological Sciences **276**:1865-1873.

Engen, S., and B. E. Sæther. 2014. Evolution in fluctuating environments: decomposing selection into additive components of the Robertson–Price equation. Evolution **68**:854-865.

Fox, J. W., and W. S. Harpole. 2008. Revealing how species loss affects ecosystem function: the trait-based Price equation partition. Ecology **89**:269-279.

Frank, S. A. 2012a. Natural selection. III. Selection versus transmission and the levels of selection. Journal of evolutionary biology **25**:227-243.

Frank, S. A. 2012b. Natural selection. IV. The price equation. Journal of evolutionary biology **25**:1002-1019.

Frank, S. A. 2012c. Natural selection. V. How to read the fundamental equations of evolutionary change in terms of information theory. Journal of evolutionary biology **25**:2377-2396.

Genung, M. A., J. A. Schweitzer, F. Ubeda, B. M. Fitzpatrick, C. C. Pregitzer, E. Felker-Quinn, and J. K. Bailey. 2011. Genetic variation and community change–selection, evolution, and feedbacks. Functional Ecology **25**:408-419.

Govaert, L., J. H. Pantel, and L. De Meester. 2016. Eco-evolutionary partitioning metrics: assessing the importance of ecological and evolutionary contributions to population and community change. Ecology Letters **19**:839-853.

Grant, P. R., and B. R. Grant. 2002. Unpredictable evolution in a 30-year study of Darwin's finches. Science **296**:707-711.

Hanson, C. A., J. A. Fuhrman, M. C. Horner-Devine, and J. B. Martiny. 2012. Beyond biogeographic patterns: processes shaping the microbial landscape. Nature Reviews Microbiology **10**:497-506.

Harmon, L. J., B. Matthews, S. Des Roches, J. M. Chase, J. Shurin, and D. Schluter. 2009. Evolutionary diversification in stickleback affects ecosystem functioning. Nature **458**:1167-1170.

Hedrick, P. W. 2011. Genetics of populations. Jones & Bartlett Learning.

Helanterä, H., and T. Uller. 2010. The Price equation and extended inheritance. Philosophy & Theory in Biology **2**.

Hill, M. O. 1973. Diversity and evenness: a unifying notation and its consequences. Ecology **54**:427-432.

Holt, R. D., and R. Gomulkiewicz. 1997. How does immigration influence local adaptation? a reexamination of a familiar paradigm. The American Naturalist **149**:563-572.

Hubbell, S. P. 2001. A unified neutral theory of biodiversity and biogeography. Princeton University Press, Princeton, NJ.

Hurlbert, S. H. 1971. The nonconcept of species diversity: a critique and alternative parameters. Ecology **52**:577-586.

Jost, L. 2006. Entropy and diversity. Oikos **113**:363-375.

Jost, L. 2007. Partitioning diversity into independent alpha and beta components. Ecology **88**:2427-2439.

Kerr, B., and P. Godfrey-Smith. 2009. Generalization of the Price equation for evolutionary change. Evolution **63**:531-536.

bioRxiv preprint doi: https://doi.org/10.1101/527028; this version posted January 22, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

Lehtonen, J. 2018. The Price Equation, Gradient Dynamics, and Continuous Trait Game Theory. The American Naturalist **191**:146-153.

Loreau, M., and A. Hector. 2001. Partitioning selection and complementarity in biodiversity experiments. Nature **412**:72-76.

Luís, R., S. Elaydi, and H. Oliveira. 2011. Stability of a Ricker-type competition model and the competitive exclusion principle. Journal of Biological Dynamics **5**:636-660.

MacDonald, A., and M. Vellend. 2016. R code for Chapter 6: Simulating Dynamics in Ecological Communities.

Mallet, J. 2012. The struggle for existence: how the notion of carrying capacity, K, obscures the links between demography, Darwinian evolution, and speciation. Evolutionary Ecology Research **14**:627-665.

Norberg, J., M. C. Urban, M. Vellend, C. A. Klausmeier, and N. Loeuille. 2012. Eco-evolutionary responses of biodiversity to climate change. Nature Climate Change **2**:747-751.

Nowak, M. A. 2006. Evolutionary dynamics. Harvard University Press.

Otto, S. P., and T. Day. 2007. A biologist's guide to mathematical modeling in ecology and evolution. Princeton University Press.

Ozgul, A., S. Tuljapurkar, T. G. Benton, J. M. Pemberton, T. H. Clutton-Brock, and T. Coulson. 2009. The dynamics of phenotypic change and the shrinking sheep of St. Kilda. Science **325**:464-467.

Pontarp, M., L. Bunnefeld, J. S. Cabral, R. S. Etienne, S. A. Fritz, R. Gillespie, C. H. Graham, O. Hagen, F. Hartig, and S. Huang. 2018. The Latitudinal Diversity Gradient: Novel Understanding through Mechanistic Eco-evolutionary Models. Trends in ecology & evolution.

Price, G. R. 1970. Selection and covariance. Nature **227**:520-521.

Price, G. R. 1972. Extension of covariance selection mathematics. Annals of human genetics **35**:485-490.

Price, G. R. 1995. The nature of selection. Journal of Theoretical Biology **175**:389-396.

Price, T. D., A. Qvarnström, and D. E. Irwin. 2003. The role of phenotypic plasticity in driving genetic evolution. Proceedings of the Royal Society of London B: Biological Sciences **270**:1433-1440.

Queller, D. C. 2017. Fundamental Theorems of Evolution. The American Naturalist **189**:345-353.

Rankin, B. D., J. W. Fox, C. R. Barrón-Ortiz, A. E. Chew, P. A. Holroyd, J. A. Ludtke, X. Yang, and J. M. Theodor. 2015. The extended Price equation quantifies species selection on mammalian body size across the Palaeocene/Eocene Thermal Maximum. Proc. R. Soc. B **282**:20151097.

Remus-Emsermann, M. N., C. Pelludat, P. Gisler, and D. Drissner. 2018. Conjugation dynamics of self-transmissible and mobilisable plasmids into E. coli O157:H7 on Arabidopsis thaliana rosettes. bioRxiv.

Rice, S. H. 2004. Evolutionary theory: mathematical and conceptual foundations. Sinauer Associates Sunderland.

Ricker, W. E. 1954. Stock and recruitment. Journal of the Fisheries Board of Canada **11**:559-623.

Robertson, A. 1966. A mathematical model of the culling process in dairy cattle. Animal Science **8**:95-108.

bioRxiv preprint doi: https://doi.org/10.1101/527028; this version posted January 22, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

Scheiner, S. M., and M. R. Willig. 2011. The theory of ecology. University of Chicago Press.

Turcotte, M. M., and J. M. Levine. 2016. Phenotypic plasticity and species coexistence. Trends in ecology & evolution **31**:803-813.

Vellend, M. 2010. Conceptual synthesis in community ecology. The Quarterly review of biology **85**:183-206.

Vellend, M. 2016. The theory of ecological communities (MPB-57). Princeton University Press.

Weiher, E., D. Freund, T. Bunton, A. Stefanski, T. Lee, and S. Bentivenga. 2011. Advances, challenges and a developing synthesis of ecological community assembly theory. Philosophical Transactions of the Royal Society of London B: Biological Sciences **366**:2403-2413.

Weis, A. E., K. M. Turner, B. Petro, E. J. Austen, and S. M. Wadgymar. 2015. Hard and soft selection on phenology through seasonal shifts in the general and social environments: A study on plant emergence time. Evolution.

Winfree, R., J. W Fox, N. M. Williams, J. R. Reilly, and D. P. Cariveau. 2015. Abundance of common species, not species richness, drives delivery of a real-world ecosystem service. Ecology Letters **18**:626-635.

Table 1: Quantification of how selection, immigration, drift and speciation change ecological communities over time using variants of the Price equation. We note transmission bias's interpretation in each study as this can depend on the variables measured and the experimental design. None of these papers explicitly distinguish drift from selection, but theory suggests that their measurements of selection implicitly include drift (Rice 2004).

Study	Variable measured	Lineage sorting (of offspring)		Immigration	Speciation	Transmission bias
		Selection	Drift			
Collins and Gardner (2009)	Poison induced tolerance	Yes	?	No	No	Evolution within species + Physiology
Norberg et al. (2012)	Temperature optimum	Yes	?	Yes	No	Evolution within species
Rankin et al. (2015)	Body size	No[1]	?	Yes	Yes	Evolution within species
Govaert et al. (2016)	Age at first reproduction	Yes	?	Yes	No	Phenotypic plasticity-Evolution within species of plasticity

[1] Rankin et al. defines species level selection as the number of descendant species not the number of descendant individuals. With this definition, species level selection includes speciation but ignores changes in species' relative abundances.

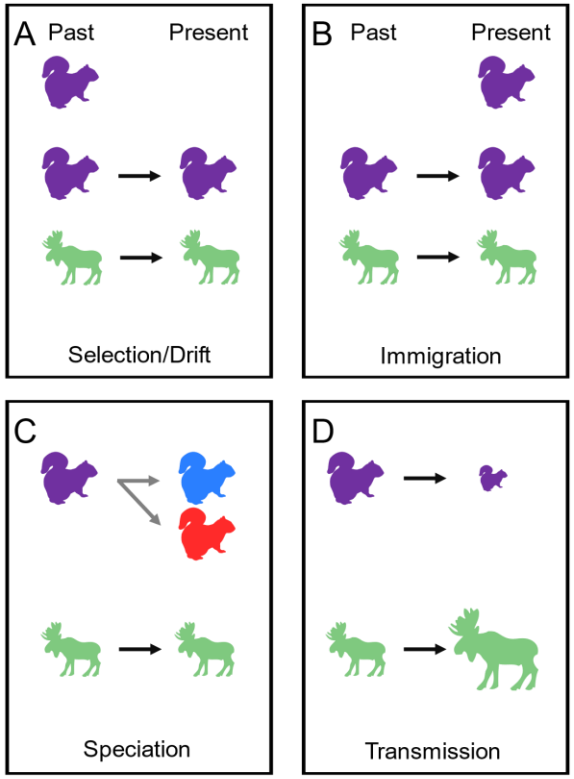
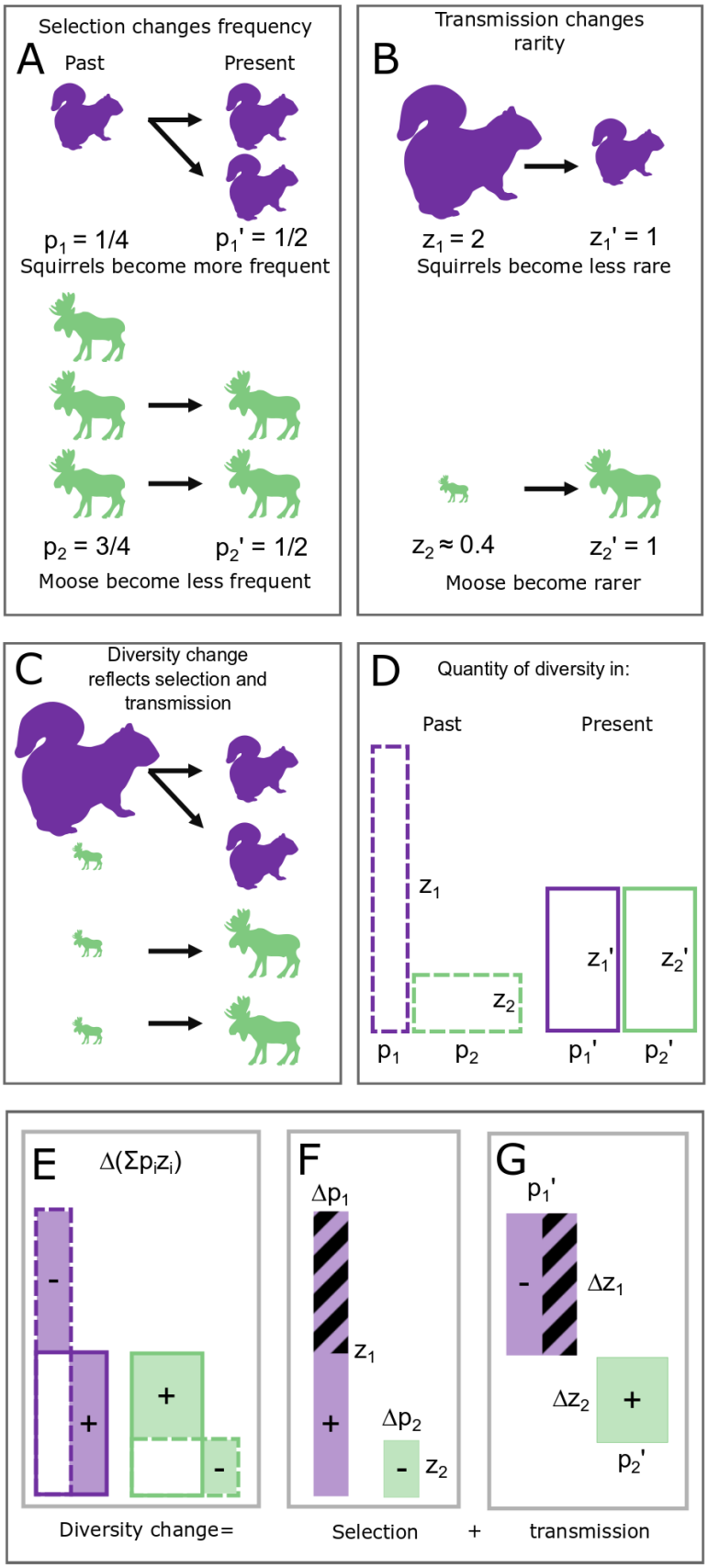


Figure 1: Schematic illustrations of change in ecological communities. A) Selection and drift change the frequency of species due to differences in the numbers of offspring produced. B) The arrival of immigrants in the community (i.e. the top squirrel in the present column who has no ancestors in the past community). C) Speciation, here depicted as a change from purple squirrel to red and blue squirrels) alters species' frequency. D) Change in species attributes, in this case squirrels become smaller while moose become bigger. This form of change does not fit neatly into selection, drift, immigration or speciation, and its effect on biodiversity change is yet to be clarified.

bioRxiv preprint doi: https://doi.org/10.1101/527028; this version posted January 22, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.



bioRxiv preprint doi: https://doi.org/10.1101/527028; this version posted January 22, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

Figure 2: A visual guide to partitioning change in Shannon entropy. Here we consider an example community consisting of squirrels and moose. In the past, there was 1 squirrel for every 3 moose. At present there are equal numbers of moose and squirrels (arrows indicate descent; p_1 =1/4, p_2=3/4, p_1'=p_2'=1/2). A) selection changed the proportion of moose and squirrel. B) Over the same time interval, the rarity of each species changes. In this figure the height of moose and squirrel is proportional to its rarity (using $z_i = -\log_2(p_i)$: $z_1 = 2$, $z_2 \approx 0.4$, $z_1' = z_2' = 1$). C) Both selection and transmission alter diversity. The two processes differ in magnitudes and work in opposite directions (i.e. squirrels become more frequent and less rare). D) The role of transmission in diversity change can be verified graphically leading to a visual proof that total change equals selection plus transmission. Diversity in the past $\Sigma p_i z_i$ can be portrayed as the sum of areas contained by rectangles, one rectangle for each species, where each rectangle has a length p_i and height z_i. Diversity in the present is the sum of areas of rectangles which have lengths p_i' and heights z_i'. The change in diversity is the change in the area contained by the rectangles between present and past. E) We can visualize these changes by overlaying rectangles associated with each species. Diversity increases when present rectangles stretch beyond the past rectangle (shaded regions with a +). Diversity decreases when present rectangles cover less area than past rectangles (shaded regions with -). F) Selection's effect consists of the area contained by rectangles with base Δp_i and height z_i. G) Transmissions effects consists of the area contained by rectangles with base p_i' and height Δz_i. Total change (i.e. E) equals selection plus transmission (i.e. F+G). This can be checked by cutting out the shaded regions on each side and overlaying them. For moose one rectangle on the right-hand side corresponds to the effect of selection, while the other corresponds to the effect of transmission. For squirrels the gains due to

bioRxiv preprint doi: https://doi.org/10.1101/527028; this version posted January 22, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

selection partially cancel out the loss due to transmission (regions with black hatching), but the total area on the right-hand side equals the total area on the left-hand side. This graphical interpretation of diversity change can also be verified mathematically: total change in diversity $\Delta(\Sigma p_i z_i)=0.5\times1+0.5\times1-(0.25\times2+0.75\times0.4)=0.2$ is equal to change due to selection $\Sigma\Delta p_i z_i =(0.5-0.25)\times2+(0.5-0.75)\times0.4= 0.4$ plus change due to transmission $\Sigma p_i'\Delta z_i = 0.5\times(1-2)+ 0.5\times(1-0.4) = -0.2$.

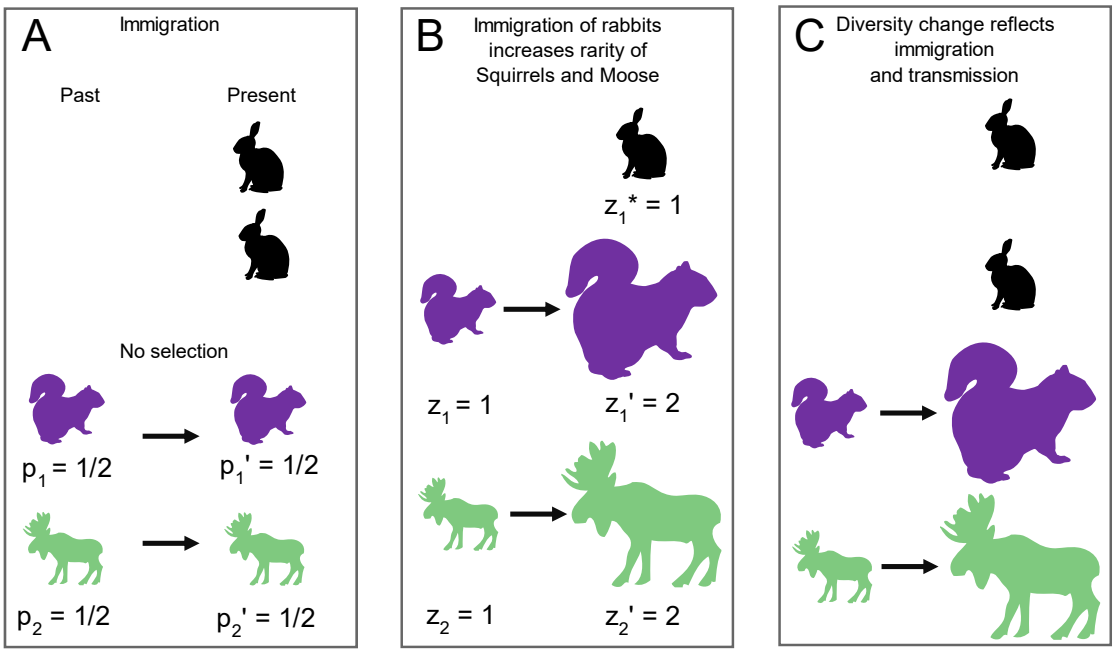
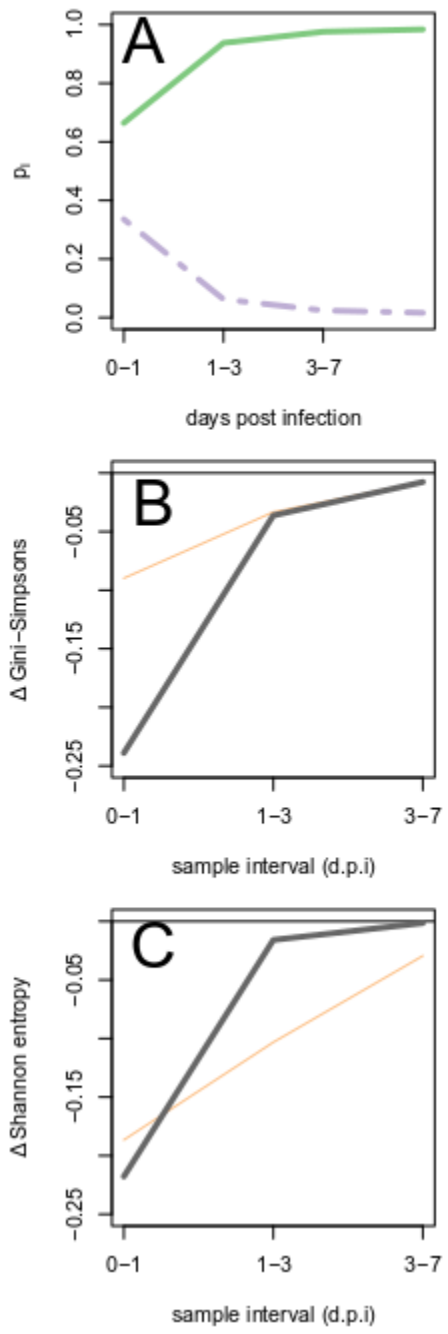


Figure 3: an illustration of how transmission bias emerges from immigration. Here we consider a community that has recently experienced immigration of rabbits. A) In the past the community consisted of equal numbers of moose and squirrel. There is no selection because the proportion of ancestors and descendants of these species is unchanged ($p_1 = p_2= p_1'=p_2'=1/2$). B) However, moose and squirrel are rarer in the present than they were in the past (using $z_i = -\log_2(p_i)$: $z_1 = z_2 = 1=$, $z_1'= z_2' = 2$). This change in rarity is a form of transmission bias. C) To correctly account

bioRxiv preprint doi: https://doi.org/10.1101/527028; this version posted January 22, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

for diversity change we must include change in species' proportion due to immigration and change in species' rarity among descendants of the past community.



bioRxiv preprint doi: https://doi.org/10.1101/527028; this version posted January 22, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

Figure 4. A) Mechanisms changing diversity in an experiment concerning two species of bacteria (*P. ecalypti*, green; *E. coli*, purple). Leaves of *A. thaliana* were inoculated with both species and the frequency of each species was measured at days 0, 1, 3, and 7, resulting in three sampling intervals of which we can measure change in diversity (0-1, 1-3 and 3-7 days post inoculation). In the seven days following their inoculation the proportion of *P. eucalypti* increased while that of *E. coli* decreased. Change in the proportion of each species was particularly rapid between day 0 and day 1. B) Gini Simpsons diversity was strongly influenced by transmission bias (tan) and selection (grey). C) Shannon entropy was strongly influenced by transmission bias between day 0 and day 1 and weakly influenced in the subsequent two sampling intervals (tan). To facilitate comparisons with other analyses of Shannon entropy log base *e* was used.

bioRxiv preprint doi: https://doi.org/10.1101/527028; this version posted January 22, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

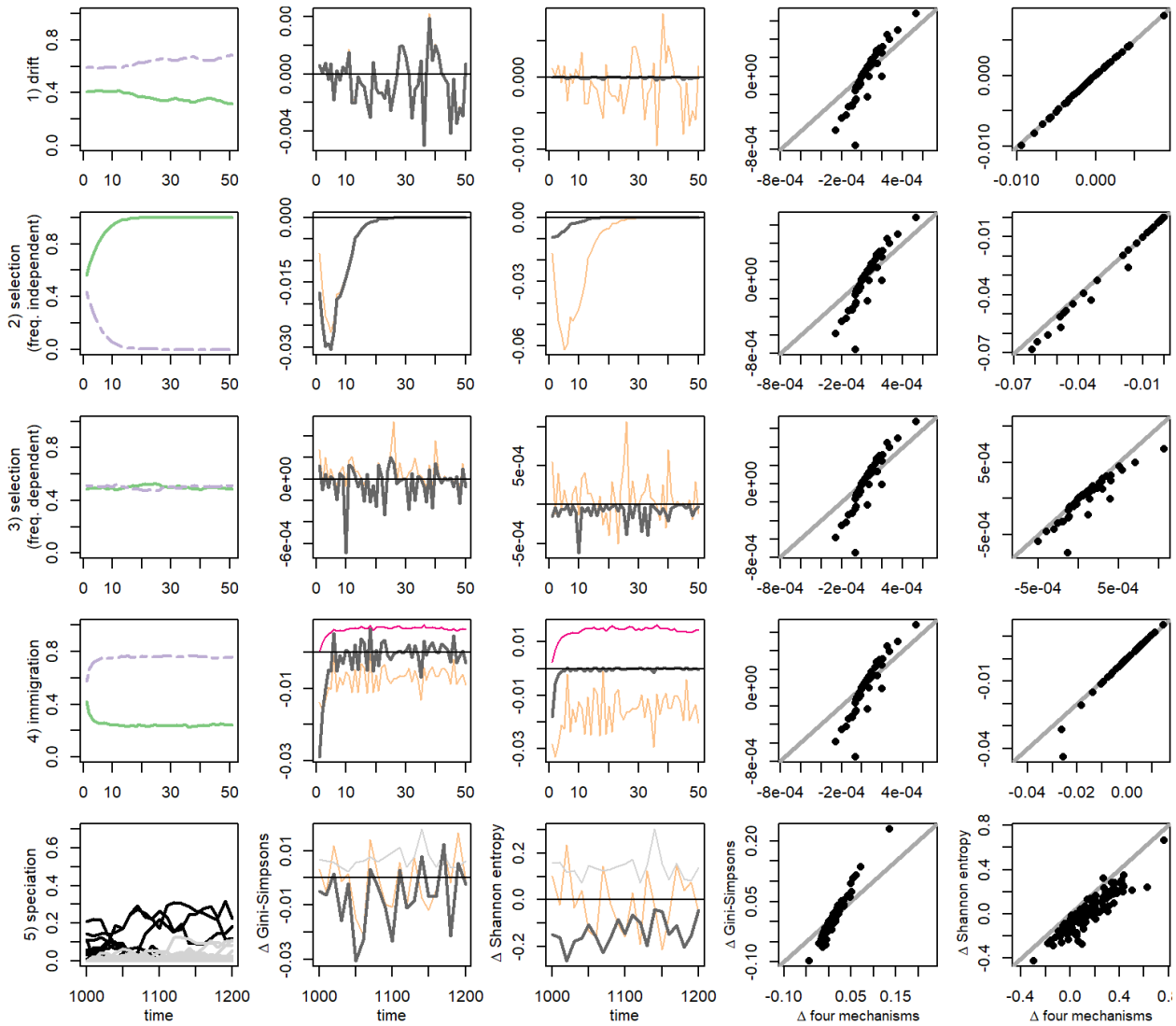


Figure 5. Transmission bias can substantially influence change in diversity over time. Each row depicts a different ecological scenario, with the first column depicting change in the proportion of species 1 (solid green line) and species 2 (dashed purple line), or in the case of the speciation simulation (final row) species present initially (black) and new species emerging over the course of the simulation (grey). The second column illustrates change in Gini Simpson's index over time. For this index transmission bias (thick dark grey) is often comparable in magnitude to immigration (fuschia), speciation (thin light grey), and the lineage sorting term encompassing

bioRxiv preprint doi: https://doi.org/10.1101/527028; this version posted January 22, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

selection and drift (tan). The third column illustrates change in Shannon entropy, where transmission bias is occasionally a substantial driver of change (frequency-dependent selection, and immigration). The fourth column illustrates that change in Gini Simpson's attributed to the four mechanisms is about ½ of the total change in Gini Simpson's across time steps. The fifth column illustrates change in Shannon entropy attributed to Vellend's four mechanisms (x-axis fifth column) is moderately smaller than total change (y-axis fifth column). To facilitate comparisons with other analyses of Shannon entropy log base *e* was used.