Use of $\alpha$, $\beta$, and $\gamma$ Diversity Measures to Characterize Seed Dispersal by Animals

Douglas G. Scofield,1,2 Peter E. Smouse,3 Jordan Karubian,4 and Victoria L. Sork1,5,*

1. Department of Ecology and Evolutionary Biology, University of California, Los Angeles, California 90095; 2. Umeå Plant Science Centre, Department of Plant Physiology, Umeå University, 901 87 Umeå, Sweden; 3. Department of Ecology, Evolution, and Natural Resources, School of Environmental and Biological Sciences, Rutgers University, New Brunswick, New Jersey 08901; 4. Department of Ecology and Evolutionary Biology, Tulane University, New Orleans, Louisiana 70118; 5. Institute of the Environment and Sustainability, University of California, Los Angeles, California 90095

Submitted March 28, 2011; Accepted July 27, 2012; Electronically published October 30, 2012


Abstract: Seed dispersal shapes ecological and evolutionary dynamics of plant populations. Here, we extend classical diversity measures to study the impact of disperser behavior on seed dispersal. We begin by extending our previous diversity structure approach, which partitioned seed source diversity within and among dispersal sites, into the more general framework of traditional diversity measures. This statistical approach allows an assessment of the extent to which foraging behavior shapes $\alpha$ and $\gamma$ diversity, as well as the divergence in seed sources among dispersal sites, which we call $\delta$. We also introduce tests to facilitate comparisons of diversity among dispersal sites and seed vectors and to compare overall diversity among sampled systems. We then apply these tools to investigate the diversity blend of parentage resulting from seed dispersal by two avian seed vectors with very different social and foraging behaviors: (1) acorn woodpeckers, transporting Quercus agrifolia acorns, and (2) long-wattled umbrellabirds, transporting Oenocarpus bataua palm nuts. Using these diversity and divergence measures, we test the hypothesis that different foraging behaviors generate distinctive diversity partitions for the two focal tree species. This approach provides a new tool for assessment of the impact of dispersal agents on the seed source structure of plant populations, which can be extended to include the impact of virtually any propagule vector for a range of systems.

Keywords: $\alpha$ diversity, $\beta$ diversity, frugivory, seed dispersal by animals, genetic diversity, foraging behavior.

Introduction

Seed dispersal is a critical process in plants (van de Pijl 1969; Howe and Smallwood 1982), influencing patterns of both species diversity within communities and genetic diversity within and among populations (e.g., Vellend 2005). The ecology of seed dispersal has received extensive attention since Janzen (1970) and Connell (1971) first posited that movement of seeds away from the parent plant shapes the distribution of seedlings through avoidance of density- and distance-dependent sources of mortality, such as localized pathogens and intense competition among sibs, and through increased probability of encountering a favorable establishment site (Howe and Smallwood 1982). The most frequent shape of the dispersal kernel is a leptokurtic curve emanating from the source tree (Janzen 1970; Levin 1981; Howe and Smallwood 1982; Levin et al. 2003)—a pattern often observed for abiotic and animal dispersal vectors alike. However, when the vector transports seeds to a specific destination (such as a caching site, perch site, or lek), the pattern of seed deposition created by that vector can be overdispersed relative to exponential decay (e.g., Janzen 1970; Wenny and Levey 1998; Gomez 2003; Karubian et al. 2010). Regardless of the shape of the kernel, one important consequence of dispersal is the resulting mixture of seed sources because some of the ecological benefits of dispersal, such as avoidance of localized pathogens or sib-sib competition, depend on the diversity of genotypes locally and across the site. This mixture has evolutionary implications as well, because the distribution and abundance of genetic variation is a direct consequence of how dispersal vectors move seeds.

The pattern of seed dispersal by animals is sensitive to the species-specific foraging and social behavior of the dispersal agents themselves (Fleming 1981; Wang and Smith 2002; Gomez 2003; Russo et al. 2006; Jordano et al. 2007; Muscarella and Fleming 2007). For example, seeds of a single plant species may be dispersed by several animal species, and each may disperse seeds different distances and directions and into different habitats (Jordano et al. 2007; Garcia et al. 2009). These studies demonstrate that ecological variables and vector species can be critical factors in determining where and how far seeds are dispersed.
Given that genetic and ecological consequences of dispersal are likely to be related (Vellend 2005), it would be useful to develop a generic diversity currency that would accommodate both within the same analytical framework, especially since genetic markers are commonly used to track dispersal. Ecological diversity measures should provide a useful analytic approach for seed dispersal and would be particularly useful for assessing the impact of vector-based dispersal on plant populations. Whitaker (1960, 1972) first introduced the concepts of $\alpha$, $\beta$, and $\gamma$ diversity to describe compositional diversity within sites, the turnover in species composition among sites, and cumulative diversity across sites. Many ecologists have contributed to the concepts and measurement of species diversity (e.g., MacArthur 1965; Pianka 1966; Pielou 1966; McIntosh 1967; Hurlbert 1971; Huston 1979; Hill and Gauch 1980; Hughes 1986; Lande 1996; Chao et al. 2005; Jost 2006, 2007, 2010; Colwell 2010; Tuomisto 2010, 2012). Jost (2006, 2007, 2008) and colleagues (e.g., Chao et al. 2008) have developed ecological diversity measures that share a common structural foundation with genetic diversity measures, thus constituting a familiar basis for numerical interpretation. For example, the index of species identity, calculated from the species’ frequencies (Simpson 1949), is mathematically identical to the standard estimate of allelic identity (Nei 1972), and both can be translated directly into “effective numbers” of elements (effective number of species for ecology, effective number of alleles for genetics). In previous work, we developed a set of metrics for characterizing patterns for both seeds (probability of maternal identity [PMI]; Grivet et al. 2005) and pollen (probability of paternal identity [PPI]; Smouse and Roldedo-Arnuncio 2005). Here, we use these measures to estimate the effective number of maternal seed sources, constructing a generic framework for diversity analysis of seed dispersal.

The overall goal of this article is to introduce and deploy the $\alpha$, $\beta$, and $\gamma$ diversity framework for ecological studies of seed dispersal and to use this diversity framework to assess the extent to which vertebrate social and foraging behavior affect the distribution of maternal progenies across the landscape. We first translate our previously developed measures of seed source diversity (Grivet et al. 2005, 2009; Scofield et al. 2010) into analogous indices of $\alpha$, $\beta$, and $\gamma$ diversity of maternal progenies within and among seed pools and introduce a new pairwise measure of divergence (diversity turnover), $\delta$. Taken together, these diversity and divergence indices can be deployed to represent any set of transported seeds found in a defined locality. For example, one could compare differences in the movement of seed, as summarized by $\alpha$, $\beta$, and $\gamma$ diversity, along with $\delta$ divergence, to assess “seed-vector effectiveness” of different vectors or vector guilds, which would complement the approach of Schupp et al. (2010), who advocate studying both quantity and quality of dispersal events.

In the second part (“Study Systems and Field Methods”), we contrast two seed dispersers and the impact of their divergent social and foraging behaviors on the diversity of maternal seed progenies within seed pools, overall diversity for the region, and the divergence (turnover) of seed source composition from seed pool to seed pool, using previously published data (Karubian et al. 2010; Scofield et al. 2010). The first data set is derived from a highly territorial seed forager, the acorn woodpecker (Melanerpes formicivorus), and their stored samples of acorns (Quercus agrifolia). The second data set is generated from the lek-breeding long-wattled umbrellabird (Cephalopterus penduliger) and dispersed seeds of the canopy palm tree (Oenocarpus bataua). The data include seed pools sampled inside the sexual display leks of male umbrellabirds, which incidentally regurgitate seeds at these sites, as well as seed pools sampled outside leks, representing the outcome of dispersal by multiple processes. We deploy statistical methods introduced in the first part (“Statistical Methods”) to evaluate the central hypothesis of this article: that different types of foraging behavior, attributable to distinctive social behaviors of the two seed vectors, yield profoundly different seed dispersion patterns for the two plant species. Such differences will have consequences for the genetic diversity of subsequent generations of those plant species. Specifically, we address three comparative questions. (1) How does $\alpha$ diversity within seed pools differ for different dispersal processes? (2) Do different dispersal agents sample and transport seed sources across a landscape in different ways, creating distinctive patterns in $\beta$ diversity and $\delta$ divergence, both measuring seed source turnover for different deposition/caching sites? (3) How does overall $\gamma$ diversity across the site change as a consequence of the divergent behaviors of different dispersal agents or situations?
Statistical Methods

Seed Source Diversity

After vertebrate dispersal agents move seeds across the landscape, we can determine the number and relative representations—the diversity—of seed sources found within each dispersal site (fig. 1A). Behavioral differences among dispersal agents may give rise to considerable intra- and interspecific differences in seed source diversity. To estimate the diversity of maternal seed sources in seed pools, we begin with a collection of \( N \) dispersed seeds collected from \( G \) seed pools, \( n_g \) from the \( g \)th pool. Each seed is unambiguously attributed to one of \( K \) identified (maternal) seed sources, on the basis of its genotype (Godoy and Jordano 2001) or direct behavioral observations (Koenig et al. 2008), with \( x_{gk} \) seeds sampled from the \( g \)th pool (Grivet et al. 2005). The contribution of the \( k \)th source to the entire collection of seeds collected from \( G \) pools is \( X_k \).

Grivet et al. (2005) provided an unbiased estimate of the probability of maternal identity (PMI), that is, the probability of drawing two seeds, both from the \( g \)th pool, from the same maternal source:

\[
 r_{gg} = \frac{\sum_{i=1}^{K} x_{g(i)}(x_{g(i)} - 1)}{n_g(n_g - 1)},
\]

also recognizable as Simpson’s (1949) unbiased estimator of species identity and as Nei’s (1972) unbiased estimator of the probability of allelic identity within the \( g \)th pool, the traditional measure of genetic homozygosity/identity. Jost’s (2007) translation of the identity index in equation (1) is a natural measure of \( \alpha \) diversity. For the \( g \)th seed pool, we define \( \alpha_g \) as the reciprocal of \( r_{gg} \),

\[
 \alpha_g = r_{gg}^{-1},
\]

the effective number of maternal sources (\( N_m \)) estimated by Grivet et al. (2005). The unweighted average within-pool \( \alpha \) diversity is estimated by

\[
 \alpha = \frac{1}{G} \sum_{g=1}^{G} \alpha_g,
\]

the probability that two randomly drawn seeds, both from any particular pool, have the same mother.

Grivet et al. (2005) also define an unbiased measure of PMI for the entire collection of \( G \) seed pools as the probability of drawing two seeds from the entire collection, without regard to its pool but from the same maternal source, as

\[
 R_0 = \frac{\sum_{g=1}^{G} X_k(X_k - 1)}{N(N - 1)},
\]

where \( X_k \) is the tally of seeds from the \( k \)th maternal source from among the total collection of \( G \) pools and \( N \) is the total sample size of seeds from those pools. Inversion of equation (3) yields our measure of \( \gamma \) (total site) diversity:

\[
 \gamma = R_0^{-1},
\]

Figure 1: Seed dispersal is the outcome of seed transport from seed sources to deposition sites, which may be survey plots, seed pools, burial sites, seed traps, or other seed pools, as shown in A. In B, the outcomes of those dispersal events may be summarized by the diversity of seed sources found within deposition sites, measured as \( \alpha \) diversity (\( \alpha_g \)) within those sites, \( \delta_{ij} \) divergence between pairs of sites, and \( \gamma \) diversity (\( \gamma \)) of all seed sources sampled. Also shown are relative magnitudes of all \( \alpha_i \) and \( \delta_{ij} \) estimates that may be derived from the figure.
the effective number of maternal seed sources contributing to the entire collection of $N$ seeds from $G$ pools.

We use unbiased estimators here because more traditional diversity estimates from pools with small sample sizes may exhibit extreme biases, especially for pools containing seeds from several rare sources. We refer the reader to part A of the appendix, available online, for analogous treatments of the more traditional (biased) estimators.

$\beta$ Diversity and $\delta$ Divergence

For ecological studies, $\beta$ diversity, defined simply as $\beta = (\gamma / \alpha)$, provides a useful description of species turnover and thus compositional divergence among sites across the landscape (Whittaker 1960, 1972). Similarly, interpool $\beta$ diversity represents the turnover in maternal seed sources from pool to pool, the numerical equivalent of the "effective number" of sites (no compositional overlap) that would translate the average within-pool $\alpha$ diversity into the observed $\gamma$ diversity of the entire site. From a seed dispersal perspective, $\beta$ diversity reflects the foraging and social behavior of the dispersal vectors, which can depend on the scale of home ranges, spatial array of fruiting trees, landscape mosaic, number of fruit taken per visit, gut passage time, and many other factors that interact with the foraging behavior. Jost (2007) has suggested scaling $\beta$ relative to its maximum achievable value, $G$ (the number of pools being assessed). We can define a $[0, 1]$-scaled $\beta$ diversity measure as

$$0 \leq \Delta_\beta = \left[ 1 - \frac{1}{\beta} \right] \div \left[ 1 - \frac{1}{G} \right] = \frac{\beta - 1}{G - 1} \cdot \frac{G}{\beta} \leq 1. \quad (5)$$

We can also compute pairwise PMI, the probability that two randomly drawn seeds, one from each of the $g$th and $h$th pools, share the same maternal source:

$$r_{gh} = \sum_{i=1}^{S} \left( \frac{x_{i,gh}}{n_g} \cdot \frac{x_{i,gh}}{n_h} \right). \quad (6)$$

We next combine our within- and between-pool PMI estimators $r_{gg}$, $r_{gh}$, and $r_{hh}$ into a measure of source-compositional divergence for the $g$th and $h$th pools ($\delta_{gh}$), using a modification of the Horn (1966) index (see also Chao et al. 2008; Jost 2008):

$$\delta_{gh} = 1 - \frac{2 \cdot r_{gh}}{r_{gg} + r_{hh}}, \quad (7a)$$

where $\delta_{gh} = 1$ for total divergence (no shared seed sources) and $\delta_{gh} = 0$ for the case of total overlap (same seed source composition). In practice, the calculated values may be slightly less than 0, by virtue of the bias corrections for $r_{gg}$ and $r_{hh}$ (see part A of the appendix). We also compute an average divergence ($\bar{\delta}$) measure for the total collection of $G$ seed pools:

$$0 \leq \bar{\delta} = 1 - \frac{\sum_{g=1}^{G} r_{gh}}{(G - 1) \cdot \sum_{g=1}^{G} r_{gg}} \leq 1. \quad (7b)$$

These divergence ($\delta$) measures are affected by the same vector-specific behaviors as are the corresponding $\beta$ diversity (turnover) measures, but they provide greater flexibility and finer spatial resolution, both of which are useful in revealing additional features of the behaviors of interest. We refer the reader to part B of the appendix for a related estimator of pairwise compositional overlap ($\omega = 1 - \delta$), as well as its average value, $\bar{\omega} = (1 - \bar{\delta})$ (Chao et al. 2008; Jost 2008).

These estimators of diversity and divergence provide a hierarchical description of the number, relative representations, and divergence of maternal source composition among seed pools. We summarize the relationships between the physical process of dispersal and the resulting diversity of seed in figure 1.

Testing for Heterogeneity of $\alpha$ Diversity among Pools

Having sampled and characterized the compositions of a number of seed pools, the first question is whether all $G$ seed pools, regardless of whether they share maternal sources, have equivalent $\alpha$ diversities. Variation in the internal diversity of different seed pools might occur for a variety of biological reasons, such as inconsistent behavior of different disperser groups, different seed vector species, or landscape variation in conditions that affect either the seed disperser or the seeds being dispersed. The null hypothesis is that $(\alpha_1 = \alpha_2 = \cdots = \alpha_G = \alpha)$, and the alternative hypothesis is that the $\alpha_i$ diversities differ among seed pools. To test the null assertion of homogeneous within-pool $\alpha$ diversities, we use a slight modification of Bartlett’s test of within-pool variance homogeneity (see Snedecor and Cochran 1989):

$$T_a = (N - G) \cdot \ln (V) - \sum_{g=1}^{G} (n_g - 1) \cdot \ln (V_g), \quad (8)$$

where $V$ is an estimate of the variance in seed source composition within the $g$th seed pool and $V_g$ is the corresponding average within-pool estimate of that variance. Bartlett’s test, $2 \cdot T_a$, was designed for normally distributed variables and is typically compared with a $\chi^2$ distribution having $(G - 1)$ degrees of freedom. Our $x_a$ tallies are distributed as multinomial variables, however, so we compare the observed value of equation (8) against a null distribution constructed via permutation of individual deviations from their pool average compositions among the $G$ pools, holding observed sample sizes constant.
In part C of the appendix, we isolate each seed’s contribution to the point-source variance of its seed pool, deriving expressions for $V_g$ and $V_p$ in equation (8) and showing that

$$V_g = (1 - r_{gg})$$  \hspace{1cm} (9a)

and that

$$V_p = \sum_{g=1}^{G} \frac{(n_g - 1)}{(N - G)} \cdot V_g = \sum_{g=1}^{G} \frac{(n_g - 1)}{(N - G)} \cdot (1 - r_{gg}).$$  \hspace{1cm} (9b)

Since $r_{gg}$ is the reciprocal of $\alpha_g$, this test of the homogeneity of within-pool variance is also explicitly a test of homogeneity of $\alpha$ diversity values among the $G$ seed pools. Permuting the squared deviations of individual seeds from their within-pool average is tantamount to permuting the contributions of those seeds to their pool-specific ($\alpha_g)$ internal source diversities. The weighted average of those squared deviations over the full set of $G$ pools is $V_p$, irrespective of which squared deviation is allocated to which pool. Holding $V_p$ constant to its observed value, we compute the array of $V_g$ values for each of the $G$ pools under random permutation of single-seed deviations of seeds among pools and calculate the variation among $V_g$ estimates. From any single randomization of maternal seed sources into the $G$ pools, we then calculate the $T_g$ value from equation (8). We repeat the randomization 10,000 times to provide an empirical distribution of $T_g$ under the null hypothesis. An R language script that performs these calculations is available from https://github.com/douglasscofield/dispersal and from http://www.eeb.ucla.edu/Faculty/Sork/Sorklab.

**Intersystem Comparisons of $\alpha$ and $\gamma$ Diversity**

We extend our nonparametric $\alpha$ homogeneity test to a pair of additional contrasts that are useful here. For the first contrast, consider two study systems whose within-pool $\alpha$ diversities we wish to compare, such as the woodpecker-oak system and the umbrellabird-palm system or umbrellabirds dispersing palm seeds into random versus lek plots. Call the two seed collections under discussion A and B. The total sample of $N_a$ and $N_b$ seeds is subdivided into $G_a$ and $G_b$ seed pools, respectively. Compute the value of $V_g$ for each system, following equation (9b), and label them $V_{ag}$ and $V_{bg}$. The first represents $N_a$ single-seed deviations from their individual seed pool means and has $(N_a - G_a)$ degrees of freedom. The second has $N_b$ single-seed deviations from their seed pool means and $(N_b - G_b)$ degrees of freedom. The test criterion for a difference in the average $\alpha$ diversities of the two systems is

$$T_g(A \text{ vs. } B) = \frac{N_a + N_b - G_a - G_b}{(N_a - G_a)} \cdot \ln (V_{ag}) - \frac{N_a - G_a}{(N_a - G_a)} \cdot \ln (V_a)$$

$$- \frac{N_b - G_b}{(N_b - G_b)} \cdot \ln (V_b),$$  \hspace{1cm} (10)

where $V_{ag/b}$ is a “degree of freedom”–weighted average of $V_{ap}$ and $V_{bp}$, specifically

$$V_{ag/b} = \frac{(N_a - G_a) \cdot V_{ag} + (N_b - G_b) \cdot V_{bg}}{(N_a - G_a) + (N_b - G_b)},$$  \hspace{1cm} (11)

Nonparametric evaluation is achieved by repeatedly permuting the $(N_a + N_b)$ individual deviations from local seed pool means between the two strata, in numbers $N_a$ and $N_b$ respectively. The extension to three or more (pooled) strata is straightforward.

For the second contrast, we can use analogues of equations (9) to test whether the total ($\gamma$) diversity values are different for different systems. Again using A and B to label the two systems, we compute the total variation within either system as the sum of squared single-seed deviations from the overall mean composition of the collective system seed pool, which (using the same logic as in part C of the appendix) leads to

$$T_r(A \text{ vs. } B) = N_a + N_b - 2 \cdot \ln (V_{ag/b})$$

$$- (N_a - 1) \cdot \ln (V_a)$$

$$- (N_b - 1) \cdot \ln (V_b),$$  \hspace{1cm} (12)

where (for $\gamma$ diversities) we define

$$V_a = (1 - R_{ah}) \quad \text{and} \quad V_b = (1 - R_{bg}),$$  \hspace{1cm} (13a)

$$V_{ag/b} = \frac{(N_a - 1) \cdot (1 - R_{ah}) + (N_b - 1) \cdot (1 - R_{bg})}{(N_a + N_b - 2)},$$  \hspace{1cm} (13b)

where the $R$ values are the probabilities of global maternal identity for the two systems, defined as in equation (3). Note the differences between equations (10) and (11) and equations (12) and (13) in their degree of freedom weightings. We evaluate the null distribution for $\gamma$ diversity differences, evaluating equation (12), by computing single-seed deviations from their system means and then permuting them between systems while holding all sample sizes at their observed values. Again, extending this test to three or more systems is straightforward.

**Interpool Divergence of Seed Source Composition**

The comparison of two or more seed pools can reveal further outcomes of dispersal agent behavior. If a dispersal agent visits the same seed sources as other agents, then
their derivative seed pools will share maternal seed sources. Alternatively, different seed pools may contain one or more seed sources found in none of the other pools. The second relevant question that we may ask with respect to seed source diversity concerns the composition of seed sources in different dispersal sites—that is, whether seed vectors responsible for depositing seeds within different seed pools are sampling from different seed sources. In terms of diversity measures, the null hypothesis is that $\beta = 1$, equivalent to $\delta = 0$, either of which implies that $\alpha = \gamma$. If different proportions or subsets of seed sources are dispersed to different seed pools, however, then $1 < \beta \leq G$ and $\alpha < \gamma$, where $\beta$ is the effective number of nonoverlapping seed pools and $G$ is the total number of seed pools sampled. The issue here is the degree to which source composition diverges between seed pools. In addition to testing the average divergence, we can also conduct more sensitive pairwise tests among pairs of seed pools, which can reveal finer-scale differences in seed dispersal behavior.

Again, our observations $x_{gh}$ are multinomially distributed, so we use a sampling equivalent of Fisher’s (1958) $K \times G$ exact contingency test but with our $[0, 1]$-scaled divergence measure ($\delta$). We first calculate the entire $G \times G$ matrix of $r$ metrics, denoted $R = \{r_{gh}\}$ in equations (1) and (6), and then compute the array of pairwise divergence measures $\delta_{gh}$ for all pairs of seed pools, using equation (7a), as well as an estimate of their average pairwise divergence ($\bar{\delta}$), as in equation (7b).

To construct a null distribution for the test of no compositional divergence among seed pools ($\delta = 0$), we repeatedly permute the $N$ seed source identities among the $G$ pools, holding the sample sizes constant. To construct a null distribution for a test of no compositional divergence between the $g$th and $h$th seed pools ($\delta_{gh} = 0$), we repeatedly permute just the $(n_g + n_h)$ seed source identities for those two pools between them but holding the separate sample sizes constant. For both sorts of tests, we determine the probability that randomly permuted $\delta$ values exceed the observed values. Here, the essential patterns emerge from a consideration of the $\hat{\delta}$ values (see part D of the appendix).

**Study Systems and Field Methods**

**Acorn Woodpeckers and Quercus agrifolia**

The first focal vertebrate species is the acorn woodpecker (*Melanerpes formicivorus*), which gathers acorns from *Quercus* spp. and caches them in granaries constructed within the bark of *Quercus lobata* (valley oak) trees, other oak species, and even wooden structures, such as barns and telephone poles (Koenig and Mumme 1987). While acorn woodpeckers are poor dispersal agents (because most of the acorns they disperse will fail to germinate), they provide a good example of foraging by a territorial frugivorous species, characterized by family groups that tend to forage from distinct and nonoverlapping sets of seed sources (Koenig and Mumme 1987; Koenig et al. 2008; Scofield et al. 2010, 2011).

For this study, we will use acorn caches as seed pools of *Quercus agrifolia* Née (coast live oak) found in the Figueuroa Creek Valley of the Sedgwick Reserve. Located near Los Olivos, Santa Barbara County, California, this reserve is part of the University of California Natural Reserve System and is administered by the University of California, Santa Barbara. The study area includes $\sim180$ ha within the larger reserve area of 2,380 ha. The habitat is a matrix of oak woodland on the slopes and oak savanna on the valley floor, with *Quercus agrifolia, Quercus douglasii,* and *Q. lobata* as the primary oak species (Tyler et al. 2006). *Quercus agrifolia* occurred in large densities on the hillsides and hilltops, clustered individuals were found along the creek bed, and scattered adult trees occurred across the valley floor ($\sim2$/ha). We collected acorns from granaries during fall of 2006 and published our diversity findings, determined using PMI measures (Scofield et al. 2010). We genotyped the maternally derived pericarp tissue from each acorn, as well as leaf tissue from the putative maternal source trees for 10 microsatellite loci. Where possible, we assigned acorns to maternal parents on the basis of genotype matching between acorns and adults by means of an assignment method that takes into account common forms of genotyping error encountered with acorn pericarps (Scofield et al. 2010). Maternal assignments were accepted when there was a single maternal candidate or when the top maternal candidate was at least 100 times more likely than the second maternal candidate (LOD score of 2 or more). A total of 456 acorns were assigned maternal candidates. Because exact identification of maternal candidates was not necessary for quantifying the seed source diversities in this study, we examined a further 112 acorn genotypes not assigned to maternal candidates and assigned them to unknown seed sources, again taking into account common forms of genotyping error in pericarps. Further details of the study species, study site, and methods for identifying maternal acorn sources are available in Grivet et al. (2005) and Scofield et al. (2010).

**Long-Wattled Umbrellabirds and Oenocarpus bataua**

The second focal vertebrate species is the long-wattled umbrellabird (*Cephalopterus penduliger*), which is a large ($\sim550$ g), endangered, frugivorous bird endemic to Chocó rain forests of Ecuador and Colombia (Snow 1982, 2004). At the study site in northwest Ecuador, umbrellabirds are the most commonly observed species consuming the fruits
of Oenocarpus bataua, a large-seeded and widespread canopy palm species. Umbrellabirds swallow whole ripe Oenocarpus fruits and regurgitate viable seeds intact. Male umbrellabirds spend a large proportion of their time at the lek site but must leave the lek to forage. This results in a high frequency of palm nuts being transported to the lekking areas, where they are found beneath male display perches (Karubian et al. 2012). Oenocarpus seeds are also dispersed by other vectors, including other primary dispersal agents (toucans and female umbrellabirds), secondary dispersal agents (rodents), gravity, and water. We group these alternative dispersal modes into “background dispersal,” defined as the net effect of source-based dispersal processes by vectors other than lekking male umbrellabirds. Data used here were collected from a 30-ha study parcel (500 m × 600 m) at the Bilsa Biological Station (79°45′W, 0°22′N; elevation of 330–730 m); density of adult Oenocarpus trees is approximately two individuals per hectare in this study area. The study area is centered on a single umbrellabird lek attended by multiple males. This lek (~1 ha in area) represents a multimale seed dispersal area, although each seed plot (10 m in diameter) we sampled within the lek is likely to reflect dispersal by a single male because males hold stable display territories on the lek. The seed pools comprise seedling samples (with seeds still attached) from nine plots within the lek and from 12 “random” plots that reflect background dispersal, randomly located throughout the rest of the study parcel. To assign dispersed seeds to maternal parents, we matched genotypes from seed pericarp tissue and putative maternal source tree leaf tissue for seven microsatellite loci (Smouse et al. 2012). Details of sampling, genotyping, and other findings can be found in Karubian et al. (2010).

Results

Acorn Woodpecker Granaries

Each of the 17 granaries (seed pools) where acorn woodpeckers store acorns contained many genotypes, derived from different acorn source trees (fig. 2A). This low number of seed sources per granary yielded low mean α diversity ( \( \bar{\alpha}_w = 2.57 \pm 0.41 \), \( G_w = 17 \) granaries). The sample-size weighted average ( \( \bar{\alpha}_w = 2.15 \pm 0.27 \)) is a bit lower, and nine individual pools had low α diversity (1.0 ≤ \( \alpha_s \) ≤ 2.0). Our nonparametric test of α homogeneity, equation (8), was highly significant ( \( T_w = 766.4, P \leq 0.0001; \) fig. 3A), establishing that observed variation in \( \alpha_s \) among seed pools (fig. 2A) reflects substantial variation among acorn woodpecker territories in the degree and patterns of local seed source sampling.

We observed only four instances of seed source sharing (\( \delta_{\bar{\alpha}} > 0 \)) out of a total of 136 pairwise comparisons (3.7%; fig. 4A). Our mean δ divergence measure ( \( \bar{\delta}_w = 0.984, P \leq 0.0001 \)) was much greater than expected from our null expectation of no spatial structure in source usage (app. A, part D, fig. A2A); the null hypothesis ( \( \bar{\delta}_w = 0 \)) is clearly untenable. Typically, each granary contains mostly seed sources unique to that granary, so that the (collective) γ diversity, \( \gamma_w = 20.7 \), essentially reflects an arithmetic addition of α diversities across the G = 17 granaries. Adding granaries to the total collection amounts to adding α-sized increments to accumulated γ diversity across the site, basically a straight-line progression (fig. 4D).

Long-Wattled Umbrellabird Lek Plots and Random Forest Plots

Individual random plots, outside the leks, included many fewer seed sources per plot than did individual lek plots (fig. 2B and 2C, respectively). Mean α diversity of the 12 random plots ( \( \bar{\alpha}_r = 21.1 \pm 6.1, G_r = 12 \) ) was less than that for the nine lek plots ( \( \bar{\alpha}_l = 26.2 \pm 7.0, G_l = 9 \) ), while the sample-size-weighted mean α diversities for both sets of plots were similar ( \( \bar{\alpha}_w = 21.3 \pm 6.4 \) and \( \bar{\alpha}_l = 25.9 \pm 6.9 \), respectively). For the two lek plots containing just one seed per maternal source (1L06 and 1L07 in fig. 2C), bias correction resulted in an infinite estimate of α diversity, due to 0 values of \( r_w \) (eq. [1]); bias-corrected means exclude these plots. We observed a broad range of α diversities among random plots (1.2 < \( \alpha_s < 66.0 \), \( T_r = 9.98, P \leq 0.0001; \) fig. 3B), strongly suggesting that different random plots had experienced different deposition patterns of maternal seed sources. The range of α diversities among lek plots, by contrast, was a bit narrower and was nonsignificant (15.0 < \( \alpha_s < 68.0 \), \( T_l = 0.044, P \geq 0.8; \) fig. 3C), suggesting that male umbrellabirds within a lekking group are sampling from source tree arrays that are of approximately the same diversity, even if they are not the same arrays.

Among random plots of Oenocarpus seeds outside the lek, we observed only 12 instances of less than complete seed source divergence among 66 paired comparisons (18%; fig. 4B), and mean δ divergence of Oenocarpus seed sources among random plots ( \( \bar{\delta}_r = 0.922, P \leq 0.0001 \)) was considerably greater than our null expectation ( \( \bar{\delta} = 0; \) fig. A2B). By contrast, we observed less than complete seed source divergence for 31 of the 36 pairwise lek plot comparisons (86%; fig. 4C), with mean divergence noticeably lower ( \( \bar{\delta}_l = 0.606, P \leq 0.0001; \) fig. A2C). The (collective) γ diversity for random plots was \( \gamma_r = 37.3 \), but that for lek plots was \( \gamma_l = 59.7 \). The curves of accumulated γ diversity for both random and lek plots show a rapidly asymptotic trend as plots are added (fig. 4E, 4F).
**Figure 2:** Relative representations of seed sources among seeds deposited by two seed-dispersal agents. Filled or crosshatched patterns indicate proportions from seed sources shared across multiple discrete dispersal sites; white boxes indicate seed sources unique to a single dispersal site. 

A, *Quercus agrifolia* acorn sources found within 17 granaries (mean $\alpha = 2.57 \pm 0.41$) maintained by acorn woodpeckers (*Melanerpes formicivorus*) in Santa Barbara County, California (annotated version of figure originally presented in Scofield et al. 2010).

B, Seeds of the palm *Oenocarpus bataua* deposited in 12 random plots within the Chocó forest of Ecuador (mean $\alpha = 21.1 \pm 7.1$).

C, Seeds of *O. bataua* deposited by long-wattled umbrellabirds (*Cephalopterus penduliger*) in nine lek display sites in the Chocó forest of Ecuador (mean $\alpha = 26.2 \pm 6.0$ of noninfinite estimates).
Comparison of $\gamma$ diversities among the three situations yields the same order ($\gamma_L = 59.7 > \gamma_R = 37.3 > \gamma_W = 20.7$). The $\gamma$ diversity heterogeneity tests, using equation (11), yield mixed results, with one contrast being nonsignificant, $T_c(L$ vs. $R) = 0.0037$ ($P \leq .19$), but with the *Oenocarpus* versus *Quercus* comparisons being highly significant, $T_c(L$ vs. $W) = 0.055$ ($P \leq .001$) and $T_c(R$ vs. $W) = 0.026$ ($P \leq .008$). The combinations of $\alpha$ and $\gamma$ diversities in these three situations translate into corresponding $\beta$ diversity (turnover rates), which when [0, 1]-scaled by equation (12) yield the progression $\Delta_{R,W} = 0.946 > \Delta_{L,W} = 0.821 > \Delta_{L,R} = 0.738$, reflecting the decreasing deposition target specificity of the responsible seed vectors.

**Discussion**

Seed dispersal is typically viewed through a “source-centric” lens that focuses on propagules moving away from the maternal source (Howe and Smallwood 1982). This approach has shaped our understanding of how seed dispersal impacts seed survival (Janzen 1970; Connell 1971), the size of the genetic neighborhood (Wright 1943; Levin and Kerster 1971; Epperson 2007), and connectivity across the landscape (Sork and Smouse 2006). It has less to say about the cumulative genetic effects of seed arrival from multiple source trees at specific dispersal sites. The central contribution here is the presentation of a hierarchical statistical framework for assessment and comparison of seed source diversity, viewed from the vantage point of those deposition sites. This framework allows better resolution of the composition of seed sources within seedling populations, facilitates comparisons of dispersal outcomes among populations and species of dispersal outcomes, and enhances the integration of seed-vector behavior into seed dispersal studies.

We illustrate the utility of this approach by addressing a long-standing question in seed dispersal: how does frugivore social behavior influence the spatial distribution of maternal progenies across the landscape (e.g., Russo et al. 2006; Jordano et al. 2007; Garcia et al. 2009)? Territorial and lek-breeding frugivores show predictable differences in where they forage, in how widely they forage, and in the resulting composition of the seed pools they ultimately deposit. Differences in social behavior influence the $\alpha$ diversity within a dispersal site, the $\delta$ divergence and the $\beta$ diversity among dispersal sites, and the overall $\gamma$ diversity of the landscape in predictable fashion.

Our findings suggest that a dispersal agent’s social and foraging behavior influence how widely it will forage, which then affects $\alpha$ diversity. Acorn woodpeckers live in highly territorial social groups, and they tend to forage in the trees proximal to their storage sites and with each territorial group gathering acorns from nonoverlapping...
Figure 4: Pairwise matrix of shared diversity ($\delta$ divergence) between pairs of seed pools (A–C; shading indicates amount of divergence) and the accumulation of $\gamma$ diversity as one adds successive seed pools to the collection (D–F). A, D, *Quercus agrifolia* seed sources within acorn woodpecker granaries. B, E, *Oenocarpus bataua* seed sources in random plots outside leks. C, F, *Oenocarpus bataua* seed sources within long-wattled umbrellabird lek plots.

territories, hence from different trees (Grivet et al. 2005; Scofield et al. 2010). This foraging pattern is reflected by $\alpha$ diversity measures at or close to unity for acorn woodpecker seed pools (in this case, granaries), indicating a disperser with a high fidelity to a single seed source when provisioning a pool. Other situations that may lead to qualitatively similar outcomes may occur when a seed source and dispersal site are in close proximity, as when cache-hoarding mammals forage close to their cache or fruit-processing sites, located beneath fruiting trees (Schupp et al. 2002).

In contrast, male umbrellabirds aggregate in groups of 5–15 individuals at a single lek site but forage individually or in small groups away from the lek. When they return to the lek, males regurgitate seeds, incidentally establishing separate seed pools, directly beneath their chosen display perches. These seed pools exhibit $\alpha$ diversity that is significantly greater than that of random plots located outside the lek and more than 10 times that in acorn woodpecker granaries, consistent with the idea that umbrellabird males forage over a much broader area and sample many more trees than do territorial acorn woodpeckers. When sampling random seed pools in the forest, we would expect considerable variation in $\alpha$ diversity among them as a function of variable source density, proximity of seed sources, or the foraging ecology of diverse dispersal vectors. The relative homogeneity of $\alpha$ values for lek plots, however, probably reflects qualitative similarity in the foraging ecology of different umbrellabird males, a point we expand upon below (see also Karubian et al. 2010).

The turnover in seed source composition from seed pool to seed pool can be described by $\beta = (\gamma/\alpha)$, a traditional diversity measure that yields a straightforward estimate of the effective number of (compositionally nonoverlapping)
seed pools. We can [0, 1]-scale that measure in the form of $\Delta_p$ via equation (12). Rather than focusing on $\beta$ or $\Delta_p$, however, we have followed a suggestion by Tuomisto (2010a, 2010b) and defined a construct that is more amenable to formal hypothesis testing—namely, the concept of $\delta$ divergence, which is particularly useful for our research systems and questions. We use $\delta$ to capture the degree of seed source divergence of separate deposition pools. This parameter, which also varies between 0 and 1, can be flexibly estimated in a pair- or groupwise manner and reveals more detail about the spatial consequences of resource partitioning than does $\beta$ diversity, although $\beta$ (or [0, 1]-scaled $\Delta_p$) provides an overview of the general level of partitioning across the site.

Our findings indicate a predictive linkage between the social system of the dispersal agent and the pattern of seed source distribution across the site. Large values of $\delta_{sh}$ between neighboring seed pools and $\delta \approx 1$, as observed for acorn woodpecker granaries, are consistent with territorial defense of seed sources. When more than two seed pools are provisioned by the same territorial forager(s), pairwise $\delta_{sh}$ values may have a bimodal distribution, with lower values observed between pools maintained by the same forager group. This occurs when a single acorn woodpecker territory contains paired granaries (e.g., granaries 107 and 108 are a pair, as are granaries 990 and 991; fig. 2A; also see Scofield et al. 2010). By contrast, umbrellabird lek sites exhibit substantially lower seed source divergence ($\delta_{h} = 0.606$) than do either random seedling plots ($\delta_{h} = 0.922$) or acorn woodpecker granaries ($\delta_{sh} = 0.984$). So, while acorn woodpeckers show divergence close to the maximal value of 1.0, different male umbrellabirds are foraging on some of the same source trees, consistent with field-based observations that they move in socially cohesive groups when foraging away from the lek (Karubian et al. 2010). Thus, we expect $\delta$ divergence (and scaled $\beta$ diversity) to vary as a function of the intensity of resource defense across social systems.

As shown in figure 4D–4F, $\gamma$ diversity can either increase linearly as seed pools are added or plateau with additional seed pools if there is some sharing of seed sources between pools, with the degree of asymptotic behavior related to the overall level of seed source sharing. When testing whether pairwise values of $\delta_{sh}$ are different than predicted by a random foraging model, it is not surprising that we can reject an extremely conservative null hypothesis that there is no spatial structure to foraging (see app. A, part D). It is noteworthy that the observed $\delta$ value for acorn woodpecker granaries is far from the null distribution (app. A, part D, fig. A2A), while that for Oenocarpus seeds among random forest plots is somewhat closer to the null distribution (app. A, part D, fig. A2B) and that for lek plots is yet closer to the null distribution (app. A, part D, fig. A2C). Alternative formulations that consider spatial structure and/or more detailed examination of patterns of pairwise overlap between specific seed pools (Scofield et al. 2010) are likely to reveal further details regarding interactions between behavior of seed dispersal agents and seed source sharing. Thus, low $\delta$ divergence may be observed for a subset of plots, while for others each new plot will include new seed sources ($\delta \to 1$).

Prospects for Broader Application

By using diversity parameters to characterize the patterns of animal-vectored seed dispersal, we provide a useful
framework for assessing the impact of dispersal agents on the movement of propagules from maternal sources. To provide an example, consider the foragers dispersing *Prunus mahaleb* seeds, described by Jordano et al. (2007). Small frugivorous birds characterized by short distance dispersal may be classified as high-α/high-δ, medium-sized frugivorous birds that move seeds intermediate distances may be classified as low-α/high-δ, and carnivorous mammals characterized by a high incidence of long dispersal may be classified as high-α/low-δ. Also working with *P. mahaleb*, Garcia et al. (2009) showed that microsite effects can influence the diversity of seed sources that arrive through differences in attractiveness to a greater or lesser variety of frugivore species. The approach developed here can easily be adapted to seed dispersal within such multispecies systems, as demonstrated by our analyses of *Oenocarpus bataua* seeds dispersed into random plots throughout the forest via a variety of dispersal vectors.

This study emphasizes the diversity consequences of the movement of progeny, but we can extend this approach to allelic diversity consequences to explore not only how propagules are moved across a landscape but also the impact of that movement on allelic diversity within the subsequently germinating seedlings (V. L. Sork et al., unpublished manuscript). Moreover, we can extend the PPI approaches that estimate the number of pollen donors (Smouse and Robledo-Arnuncio 2005) to estimate the α, β, δ, and γ diversity consequences of pollen flow across that same landscape.

This framework also provides a means of assessing and predicting the biodiversity impacts of natural or anthropogenic phenomena expected to affect seed dispersal mutualisms, including frugivore population declines, spatiotemporal variation in fruit availability, and habitat alteration. We can mimic the effect of frugivore population decline or extirpation by removing (in silico) seed pools attributable to a given dispersal agent and examining the resulting effects on the diversity profile of the site. With low-α-diversity/high-δ-divergence seed vectors that behave like the woodpeckers illustrated here, a decline in the population would result in loss of seed sources from the pool of dispersed seeds. For high-α-diversity/low-δ-divergence seed vectors (here, umbrellabirds), a decline in the population would not necessarily result in a loss of seed sources from the pool of dispersed seeds, but extirpation would significantly change the overall movement of seeds. Finally, extensions of such diversity thinking and the deployment of α, β, δ, and γ diversity measures can help us refine our understanding of propagule movement in general for all sorts of vectored species, including plants and animals as well as viral, bacterial, and fungal organisms, and in particular for those using animate dispersal vectors.

Conclusions

The α, β, and γ diversity measures have had a long tradition in ecology as a means of summarizing species diversity and providing a framework for understanding the processes that lead to diversification. Here, we translate information from genetic markers to understand how propagules are distributed within and among regional sites with the same diversity metrics. The advantage of common currency is that future studies can now compare the structure of species diversity with the structure of individuals within a site or region—or even allelic diversity within populations, a site, or a region—and then address the question of whether the processes that promote diversity at one level apply to another. Here, we have introduced these translations of genetic measures to diversity measures to ask the question of whether social behavior shapes the consequences of dispersal. The answer is clearly yes. Future studies will be able to use this same framework both to gain a finer understanding of how this relationship may vary across seed dispersal systems and to answer related questions in a broad range of vectored systems.

Acknowledgments

Fieldwork for this study was performed at the University of California Natural Reserve System Sedgwick Reserve, administered by the University of California, Santa Barbara, and the Bilsa Biological Station, administered by the Jatun Sacha Foundation. We thank the directors and staff at both stations for their logistical support. The authors received awards from the National Science Foundation (NSF-DEB-0516529 to V.L.S., NSF-DEB-0514956 to P.E.S., and NSF-OISE-0402137 to J.K.). In addition, V.L.S. received funding from the University of California, Los Angeles; P.E.S. received funding from the U.S. Department of Agriculture and the New Jersey Agricultural Experiment Station (USDA/NJAES-17111); and J.K. received funding from the Disney Worldwide Conservation Fund and the National Geographic Society. We thank anonymous reviewers for superb and insightful feedback.

Literature Cited


Connell, J. H. 1971. On the roles of natural enemies in preventing...
competitive exclusion in some marine animals and rain forest trees


The long-wattled umbrellabird (Cephalopterus penduliger), an endangered frugivore endemic to Chocó rain forests of northwest South America, is an important seed dispersal agent for large-seeded tree species like palms. Photograph by Murray Cooper.