Manuscript Draft

What is long-distance dispersal? And a taxonomy of dispersal events

Pedro Jordano[†]

Sevilla, June 26, 2016

[†] Integrative Ecology Group, Estación Biológica de Doñana, CSIC, Avda. Americo Vespucio, s/n, Isla de La Cartuja E-41092 Sevilla, Spain.

Corresponding author: Pedro Jordano. Integrative Ecology Group, Estación Biológica de Doñana, CSIC, Avda. americo Vespucio, s/n, E-41092 Sevilla, Spain. Email address: jordano@ebd.csic.es

Key words: ***

Manuscript information: ** Words; ** Chars; ** Pages, * Figures; * Tables.

Abstract

1 Dispersal is a key individual-based process influencing many life-history attributes,
2 scaling up to population-level properties (e.g., metapopulation connectivity). A
3 persistent challenge in dispersal ecology has been the robust characterization of
4 dispersal functions (kernels), a fundamental tool to predict how dispersal processes
5 respond under global change scenarios. Especially the rightmost tail of these func-
6 tions, i.e. the long-distance dispersal (LDD) events, are difficult to characterize
7 empirically and to model in realistic ways. But, when is it a LDD event? In the
8 specific case of plants, dispersal has three basic components: 1) a distinct (sessile)
9 source, the maternal plant producing the fruits or the paternal tree acting as a
10 source of pollen; 2) a distance component between source and target locations;
11 and 3) a vector actually performing the movement entailing the dispersal event.
12 Here we discuss operative definitions of LDD based on their intrinsic properties: 1)
13 events crossing geographic boundaries among stands; and 2) events contributing
14 to effective gene flow and propagule migration. Strict-sense long distance disper-
15 sal involves movement both outside the stand geographic limits and outside the
16 genetic neighborhood area of individuals. Combinations of propagule movements
17 within/outside these two spatial reference frames results in four distinct modes
18 of LDD. Beyond traditional statistical approaches to characterize distributions,
19 Extreme Value Analysis (EVA) can be used to properly and explicitly evaluate
20 the properties of frequency and extent of LDD events. We discuss conditions
21 where global change scenarios truncate dispersal processes, leading to the loss of
22 key dispersal services in natural populations. Proper characterization of the LDD
23 events helps to assess, for example, how the ongoing defaunation of large-bodied
24 frugivores pervasively entails the loss of crucial LDD functions.

Introduction

Dispersal is a key individual-based process influencing many life-history attributes and scaling up to population-level properties (e.g., metapopulation connectivity, Cousens *et al.* 2008). In the specific case of plants, largely sessile organisms, dispersal has three basic components: 1) a distinct (sessile) source, the maternal plant producing the fruits or the paternal tree acting as a source of pollen; 2) a distance component between source and target locations; and 3) a vector actually performing the movement entailing the dispersal event. While realized dispersal also depends upon stages subsequent to dissemination (e.g., successful germination and seedling establishment) Schupp (1995), the three previous components fully characterize the dispersal process per se. Therefore, plant movement differs in important natural history details from animal dispersal, yet both can be assessed within a common conceptual framework (e.g., Nathan, 2006). Characteristically, animal-assisted plant dispersal has three distinct, highly integrated, components missing in the process of animal dispersal: the properties of the source (parental) plant, that mediate in the foraging of the animal vector (pollinator or frugivore), the intrinsic properties of the propagule, and the functional characteristics of the animal vector who performs the movement (Nathan *et al.*, 2008a).

The movement of pollen and seeds by animals and its consequences have intrigued population geneticists and field ecologists since the infancy of both research disciplines. Each has generated an impressive body of theoretical and empirical research through the past decades, yet advances have long been co-existing in 'parallel worlds' and the great synergistic potential of population genetics and demography for the study of plant dispersal by animals remains little explored. Knowledge gaps still having the imprint of this conceptual disconnection include the idea of long distance dispersal, and the paradoxes of forest fragmentation effects on genetic diversity (Kramer *et al.*, 2008), survival and persistence of relict tree species (Hampe & Jump, 2011), rapid post-Pleistocene recolonization of vast continental areas in response to climate modification (Clark *et al.*, 1998; Clark, 1998), among other persisting issues. This conceptual isolation has been exacerbated by technical difficulties for the robust characterization of dispersal events,

especially those involving movement over long-distances (long-distance dispersal, LDD). LDD is a characteristically extreme event of propagule movement in any plant population, typically occurring with an extremely low probability but potentially reaching an extremely long distance. Some progress has recently been made through the fast-paced implementation of molecular tools in ecological research labs and the availability of cutting-edge technology for biotelemetry applications [REF]. But much of the population geneticist and ecologist communities remains unaware of the state of the art in each other and likely under-appreciates their potential to validate and enrich dispersal studies (Jones & Muller-Landau, 2008). In particular, LDD events remain difficult to assess, both technically- with serious methodological problems for its reliable estimation- and conceptually. Our aim here is to review the LDD concept with a specific emphasis on dispersal of plant propagules (seeds and pollen), providing an extended definition that might be helpful in the robust quantification of LDD events.

Two main conceptual approaches have been used to assess dispersal (Fig. 1). The "forward" approach attempts to track the dispersal events away from the known sources, e.g., by tracking the movement patterns of frugivores as they leave fruiting plants after feeding (Fig. 1A). This is the main approach used in the movement ecology framework (Nathan *et al.*, 2008a), with extensive application to animal movement based on the use of advanced biotelemetry. The "backward" approach attempts to reconstruct the most likely source of a dispersed propagule by inferring the sources given the propagule delivery pattern, the fecundity of potential sources, and the dispersal function (Fig. 1B), i.e., by using an inverse modeling approach. The main technical challenge in Fig. 1A is to sample enough dispersal events away from the source to be able to fully characterize the tail (LDD events) of the dispersal function. In Fig. 1B, the main challenge is to have a robust sampling scheme with propagule collectors (e.g., seed traps) and a good characterization of the potential sources to derive robust estimates of the actual sources. Both approaches are limited logistically by the difficulties to sample the vast areas required to assess LDD events from the focal source population.

No explicit definition of what constitutes an LDD event exists. Previous approaches (e.g., Nathan, 2006; Schurr *et al.*, 2009) include both absolute and

4

proportional definitions to characterize LDD events. This means providing information about the absolute distances moved by a given percentile of the events and/or providing data on the proportion of events exceeding a given distance threshold (Nathan *et al.*, 2008b). The exact proportional or absolute thresholds selected remain arbitrary, as no reference spatial frame is provided within the definition of LDD. This leaves the consideration of LDD as an extreme form of context-dependent phenomenon, strongly dependent upon the scale of the biological process studied (Kinlan *et al.*, 2005). For example, Kinlan *et al.* (2005) used a spatial reference frame to characterize LDD events of marine organisms, where sedentary adults and larvae differ enormously in the spatial scales of their dispersal (D'Aloia *et al.*, 2013). Therefore, any measure of extent and reach of LDD events requires reference to an explicit spatial frame or "local" scale (Kinlan *et al.*, 2005).

We aim at providing a general framework for the quantitative analysis of LDD events so that estimates of its frequency and extent could be comparable across different study systems. We argue that both demographic and genetic elements are needed for this framework, most likely requiring a combination of field-based movement data and genetic analyses. These elements can be overlaid on previous definitions based on absolute and proportional characterizations of LDD. We start with a definition of LDD events within a spatially-explicit mechanistic framework allowing an unambiguous meaning for setting long-distance thresholds. We then use a case study to assess differential contributions of animal frugivores performing LDD.

Long-distance dispersal is currently one of the most debated topics in dispersal ecology; it defines the connectedness within the network of local populations and the possibilities for range expansion and successful colonization events. We propose a first demogenetically-based, operational definition of what a long-distance dispersal event actually is, and review existing empirical literature on distance thresholds from population and genetic perspectives. We also show how molecular tools have been used to identify the respective contributions of different animal species to the LDD portion of dispersal kernels of pollen and seeds by setting empirically-derived distance thresholds. Finally, we highlight potential applications of molecular markers beyond the quantification of just the dispersal distances

5

that prevails in current studies, e.g., experimental approaches to assess dispersal limitation and Janzen-Connell effects.

LDD within a demo-genetic perspective: a taxonomy of dispersal events

Here we propose an explicit definition of LDD and what constitutes a LDD event. Previous definitions of dispersal patterns emphasized only their distance components and characterized LDD events basically in terms of geographic distance between a dispersed propagule (or an established early seedling) and its most likely maternal or paternal (in case of pollen) source. Absolute and proportional definitions for the LDD events have been proposed depending on arbitrary thresholds of either the distance beyond which a dispersal event is LDD or the proportion of events occurring beyond a specific distance (Nathan, 2005; Nathan *et al.*, 2008b). Thus, two key biological aspects of LDD events involve the transport of propagules outside a reference area: moving away from the source stand or population, and moving away from the area where relatives stand (Kinlan *et al.*, 2005). These two movements do not necessarily concur: a propagule may move over a very long distance yet still be disseminated within the reach of the neighborhood where parental individuals mate. Within a demo-genetic framework it is easy to envision a combination of situations concerning the spatial scale of the dispersal processes (Table 1) and unambiguously define different types of LDD events. The idea that dispersal occurs in reference to these two spatial reference frames, i.e., the population or stand and the genetic neighborhood area, is motivated by the fact that dispersal entails the movement of both an individual propagule (i.e., a pollen grain or a seed) and a distinct set of genes (i.e., the male genotype in case of pollen, or a seed genotype). Thus, dispersal entails simultaneous demographic and genetic effects through recruitment of new individuals in the population and through contributions to gene flow (Harper, 1977). When considered its combined influence on demography and population genetics, the concept of LDD nicely bridges these two paradigms embedded in the biological definition of population (Waples & Gaggiotti, 2006).

6

Two important components of plant dispersal ecology concern the movement of propagules away from the source population, a type of dispersal relevant to colonization ability and range expansion (Howe & Miriti, 2004), and the movement away from the location of close relatives, i.e., a movement away from the genetic neighborhood (Hardesty *et al.*, 2006; Jones & Muller-Landau, 2008). If we classify dispersal events according to these two spatial frameworks (Table 1) we end up with four distinct types of events depending on whether or not dispersed propagules are disseminated within these reference areas. Setting the limits of a population can be problematic (Waples & Gaggiotti, 2006) yet we can identify with relative ease the geographical limits of plant stands, patches, habitat spots or other types of habitat or microhabitat discontinuities that determine landmark boundaries of biological significance (see Kinlan *et al.*, 2005, for further discussion of boundaries for dispersal). These "frontiers" set biological limits to what a LDD event is in relation to the geographic limits of the source population. Most plants are distributed as clumped patches, discrete stands, or relatively isolated populations, so we may distinguish between short-distance and long-distance dispersal events that end up with dissemination within or beyond, respectively, the stand or population geographic boundaries (Table 1, SSD_{loc} or LDD_{loc}) (Figure 2).

A second consideration in terms of spatial boundaries, with effects on dispersal patterns, is the genetic neighborhood area N_b, i.e., the spatial extent including a subset of panmictic individuals within a population (Wright, 1943, 1946). Thus, the N_b area can be equal to the whole extent of the population whenever the population is unstructured and there is evidence for random mating events among all the individuals. However, most populations and stands of long-lived trees show highly aggregated and clumped distributions (Seidler & Plotkin, 2006), where relatively long distances may separate groups of individuals within the same population. In these cases we might expect N_b area to be substantially smaller than the total population area. Therefore, at least four possible scenarios exist with distinct implications in terms of consequences for dispersal (Table 1). In the case of dispersal events not extending beyond the geographic limits of the population or reference area, actual LDD events may involve dissemination beyond a reduced neighborhood area that is smaller than the geographic extent of the pop-

7

ulation, originating local long-distance (LDD_{loc}) dispersal events (Table 1, Fig. 2A). Actual short-distance dispersal would then involve those situations where the propagule is disseminated within *both* the population limits and the genetic neighborhood boundary (SDD_{loc}). Along a similar reasoning, dispersal events outside the population limits will not necessarily convey LDD (Table 1, Fig. 2B): this is expected in cases where the genetic neighborhoods are extensive, going beyond the geographic limits of local populations, as in fig trees (Nason *et al.*, 1998), generating LDD events within the genetic neighborhood (LDD_{neigh}). Finally, strict-sense LDD events would involve dissemination outside *both* the population limits and the genetic neighborhood boundary (LDD_{ss}) (Table 1, Fig. 2A).

While both SSD_{loc} and LDD_{loc} can be crucial for assuring the local persistence of populations, LDD_{neigh} and LDD_{ss} would be extremely important contributors to the structuring of genetic pools, realized gene flow, and maintaining connectivity in metapopulation scenarios. We argue that both the demographic and the genetic references are relevant for a proper definition of LDD.

Individual and Population Neighborhoods as Reference

Continuous populations can be modeled with the concepts of isolation by distance and neighborhood size(Wright, 1943, 1946). The former refers to the case that limited gene dispersal in continuous populations produces demes that are panmictic internally, but are isolated to some extent from adjacent demes. Each group of reproducing individuals is the neighborhood, defined as the population of a region in a continuum, from which the parents of individuals born near the center may be treated as if drawn at random (Wright, 1969). The importance and influence of the dispersal process in determining the size of the neighborhood is given by this equation, which shows how the spatial dispersion (pattern of spatial distribution) of the population influences the effective population size. This influence on the effective size is given by:

$$N_e^b = 4\pi\sigma\delta \tag{1}$$

8

where δ is the density of adults per unit area and σ is the standard deviation of the distance between birth and breeding sites. This formulation is often called the neighborhood size and assumes a normal distribution of distances between parents and offspring (out in a perfect circular shape from the source). Thus, changes in the variance of dispersal distance can affect N_e^b (highly clumped populations will have reduced N_e^b). This is the basic model of "Isolation by Distance" proposed by Wright (1943, 1946). Under this type of model, migration (gene flow) is given by the variance in dispersal, and not by the proportion of the population that is composed of migrants (denoted m), as is the case with island models (Slatkin, 1985). With enough distance separating them, two plant individuals have a low probability of mating and can be considered members of distinct genetic populations even if they are not located in geographically distinct populations.

For plants, gene flow may be accomplished by both seeds and pollen, so the variance may be decomposed to account for different patterns of seed and pollen dispersal, and to take into account the mating system (outcrossing rate, t). Thus, neighborhood size can be defined with the following equation (Crawford 1984) :

$$N_e^b = 4\pi(\sigma_s^2 + \frac{t\sigma_p^2}{2})\delta(1 + t) \tag{2}$$

where σ_s is the standard deviation of seed dispersal distance, σ_p is the standard deviation of pollen dispersal distance, and δ is the density of potential parents.

Neighborhood size in plants can be estimated by marking pollen and seeds with fluorescent dyes, tags, or stable isotope enrichment (Carlo *et al.*, 2009). However, these methods do not measure effective pollen or seed movement, but they may be combined with genetic analysis to assess genetic identity and relatedness with hypervariable DNA markers (Levin, 1988; Nason *et al.*, 1998; Godoy & Jordano, 2001) to achieve reliable estimates of both effective population size and neighborhood area.

The extent of neighborhood area in plants can be extremely variable, depending on life-history attributes such as life-span, spacing patterns, mating system, etc. Even a limited sample of available information (Table S1) highlights the fact that

237 the size of neighborhood areas can in some cases exceed the geographic limits of
238 local populations (Nason *et al.*, 1998). The size of neighborhood areas may en-
239 compass at least four orders of magnitude, $10^{-2} - 10^2$ km in radius, and include
240 many individuals. Therefore, reference to this "genetic/evolutionary" paradigm
241 and reference to the geographic boundaries (sensu Waples & Gaggiotti, 2006) may
242 be instrumental to understand the actual role of LDD events in shaping the struc-
243 turing of genetic pools and contributing to gene dispersal.

244 Whenever there is a large discrepancy between population area extent and N_e^b
245 we might expect the frequency of LDD_{loc} and LDD_{neigh} differ enormously. For
246 example, relatively small N_e^b may rise the importance of LDD_{loc} in preserving
247 scenarios of panmixia within a local population, as most distant dispersal events
248 will disseminate seeds outside the neighborhood of maternal plants.

Empirical analysis of contributions to LDD

250 Empirical evaluation of differential contributions to the different forms of LDD
251 events outlined in Table 1 requires identification of source trees as well as assign-
252 ment of the dispersed propagules to specific vectors or functional groups of vectors
253 (Jordano *et al.*, 2007). Recently, DNA-barcoding techniques have been developed
254 and successfully applied to the identification of frugivore species contributing to
255 specific seed dispersal events whose source can be identified with genetic, direct
256 assignment techniques (González-Varo *et al.*, 2014). Otherwise, visual identifica-
257 tion can reliably assign the genotyped seeds to frugivore species groups based on
258 specific characteristics of scats and regurgitations (Jordano *et al.*, 2007).

259 We inferred the frugivore groups contributing dispersal events by visually iden-
260 tifying scats and regurgitations in seed traps and line transects (see Jordano
261 *et al.*, 2007, and Suppl. Mat. for additional details of methods). These frugivore
262 functional groups include up to 38 bird and 4 mammal species feeding on *P. ma-*
263 *haleb* fruits (Jordano & Schupp, 2000). Here, we differentiate four major frugivore
264 groups: large carnivorous mammals (such as foxes, badgers, and stone martens);
265 two species of medium-sized frugivorous birds, mistle thrushes (*T. viscivorus*),

10

and carrion crows (*Corvus corone*); and a pool of small-sized frugivorous birds, including warblers, redstarts, and robins (Jordano *et al.*, 2007).

To a large extent, short-distance dispersal events (strict-sense, SDD_{loc} events) are contributed by small- and medium-sized (*Turdus*) frugivorous birds (Table 2). Given the relatively reduced N_e^b area of *P. mahaleb* (Suppl. Mat. Table S1), $< 1km^2$, well below the extent of the local study population (Garcia *et al.*, 2007, 2005), we cannot estimate LDD_{neigh} events (Table 2), as all LDD events outside the reference population occur, by definition, outside the N_e^b area. Larger frugivores such as corvids and the pigeon *Columba palumbus* contribute most LDD events, and most immigrant seeds potentially dispersed from other populations. Notably, strict-sense long-distance dispersal (LDD_{ss}) appears consistently associated with large-bodied frugivores (Table 2), most likely associated with a greater frequency of movements outside the local population (Fig. 4).

Long-Distance Dispersal: the ecology of extreme events

Long-distance dispersal (LDD) is a major component of the population dynamics, genetic structure, and biogeographic history of plant species. It determines the colonization ability of new habitats and the possibilities for fragmented populations to sustain a cohesive metapopulation by immigration-emigration dynamics that rely on LDD events. Yet our current understanding of the extent, frequency, and consequences of LDD is very limited. On one hand, theoretical models fail to predict accurately the behavior of the tail of the dispersal functions, and thus fail to predict very basic properties of LDD. On the other hand, we still have very limited documentation of actual LDD events in natural populations and we still see LDD as a sporadic, rarely far-reaching process still marked with the stamp of natural history curiosity.

Combining spatially-explicit references to the geographic population limits and the N_e^b area extent helps avoiding some imprecision in setting distance thresholds to characterize LDD events. In addition the framework outlined in Table 1 bridges the combined demographic and genetic effects of LDD events.

11

As defined in our framework (Table 1), LDD, and in particular LDD_{ss} events are a specific case of extreme events (García & Borda-de Água, 2017) consistently associated with large-sized frugivores, yet including also medium-sized and highly efficient frugivorous bird species.

- Large-bodied frugivores and LDD. Discuss lb frugivores not always contributing LDD events: Pratt, tapirs - Loss of large-bodied frugivores, consequences.

Challenges and future avenues for research

12

References

Carlo, T., Tewksbury, J., Martínez, C. & Mart (2009) A new method to track seed dispersal and recruitment using 15N isotope enrichment. *Ecology*, **90**, 3516–3525.

Clark, J.S., Fastie, C., Hurtt, G.C., Jackson, S.T., Johnson, C., King, G., Lewis, M., Lynch, J., Pacala, S., Prentice, C., Schupp, E., Webb III, T. & Wyckoff, P. (1998) Reid's paradox of rapid plant migration. *BioScience*, **48**, 13–24.

Clark, J. (1998) Why trees migrate so fast: Confronting theory with dispersal biology and the paleorecord. *The American Naturalist*, **152**, 204–224.

Cousens, R.D., Dytham, C. & Law, R. (2008) *Dispersal in plants. A population perspective.* Oxford University Press, Oxford, UK.

D'Aloia, C.C., Bogdanowicz, S.M., Majoris, J.E., Harrison, R.G. & Buston, P.M. (2013) Self-recruitment in a Caribbean reef fish: a method for approximating dispersal kernels accounting for seascape. *Molecular Ecology*, n/a–n/a, doi:10.1111/mec.12274.

Garcia, C., Arroyo, J., Godoy, J. & Jordano, P. (2005) Mating patterns, pollen dispersal, and the ecological maternal neighbourhood in a Prunus mahaleb L. population. *Molecular Ecology*, **14**, 1821–1830, doi:10.1111/j.1365-294X.2005.02542.x.

Garcia, C., Jordano, P. & Godoy, J.A. (2007) Contemporary pollen and seed dispersal in a Prunus mahaleb population: patterns in distance and direction. *Molecular Ecology*, **16**, 1947–1955, doi:10.1111/j.1365-294X.2006.03126.x.

García, C. & Borda-de Água, L. (2017) Extended dispersal kernels: insights from statistics of extremes. *Journal of Ecology*, **00**, 000–000.

Godoy, J.A. & Jordano, P. (2001) Seed dispersal by animals: exact identification of source trees with endocarp DNA microsatellites. *Molecular Ecology*, **10**(9), 2275–2283.

González-Varo, J.P., Arroyo, J.M. & Jordano, P. (2014) Who dispersed the seeds? The use of DNA barcoding in frugivory and seed dispersal studies. *Methods in Ecology and Evolution*, **5**(8), 806–814, doi:10.1111/2041-210X.12212.

Hampe, A. & Jump, A.S. (2011) Climate Relicts: Past, Present, Future. *Annual Review of Ecology Evolution and Systematics*, **42**(1), 313–333, doi:10.1146/annurev-ecolsys-102710-145015.

Hardesty, B.D., Hubbell, S.P. & Bermingham, E. (2006) Genetic evidence of frequent long-distance recruitment in a vertebrate-dispersed tree. *Ecology Letters*, **9**, 516–525.

Harper, J. (1977) *Population Biology of Plants*, 1st edn. Academic Press, London.

Howe, H.F. & Miriti, M.N. (2004) When seed dispersal matters. *BioScience*, **54**(7), 651–660.

Jones, F.A. & Muller-Landau, H.C. (2008) Measuring long-distance seed dispersal in complex natural environments: an evaluation and integration of classical and genetic methods. *Journal of Ecology*, **96**(4), 642–652, doi:10.1111/j.1365-2745.2008.01400.x.

Jordano, P., García, C., Godoy, J.A. & García-Castaño, J.L. (2007) Differential contribution of frugivores to complex seed dispersal patterns. *Proceedings Of The National Academy Of Sciences Of The United States Of America*, **104**(9), 3278–3282, doi:10.1073/pnas.0606793104.

Jordano, P. & Schupp, E.W. (2000) Seed disperser effectiveness: the quantity component and patterns of seed rain for Prunus mahaleb. *Ecological Monographs*, **70**(4), 591–615, doi:10.2307/2657187?ref=no-x-route:1d70d282eb4ed20cce6eac7274a631bb.

Kinlan, B., Gaines, S. & Lester, S. (2005) Propagule dispersal and the scales of marine community process. *Diversity and Distributions*, **11**, 139–148, doi:10.1111/j.1366-9516.2005.00158.x.

Kramer, A., Ison, J., Ashley, M. & Howe, H. (2008) The paradox of forest fragmentation genetics. *Conservation Biology*, 8, doi:10.1111/j.1523-1739.2008.00944.x.

Levin, D.A. (1988) The paternity pools of plants. *The American Naturalist*, **132**, 309–317, doi:10.2307/2461984.

Nason, J.D., Herre, E. & Hamrick, J.L. (1998) The breeding structure of a tropical keystone plant resource. *Nature*, **391**, 685–687.

Nathan, R. (2005) Long-distance dispersal research: building a network of yellow brick roads. *Diversity and Distributions*, **11**(2), 125–130, doi:10.1111/j.1366-9516.2005.00159.x.

Nathan, R. (2006) Long-distance dispersal of plants. *Science*, **313**, 786–788.

Nathan, R., Getz, W.M., Revilla, E., Holyoak, M., Kadmon, R., Saltz, D. & Smouse, P.E. (2008a) A movement ecology paradigm for unifying organismal movement research. *Proceedings of the National Academy of Sciences USA*, **105**(49), 19052–19059, doi:10.1073/pnas.0800375105.

Nathan, R., Schurr, F.M., Spiegel, O., Steinitz, O., Trakhtenbrot, A. & Tsoar, A. (2008b) Mechanisms of long-distance seed dispersal. *Trends In Ecology & Evolution*, **23**(11), 638–647, doi:10.1016/j.tree.2008.08.003.

Schupp, E. (1995) Seed–seedling conflicts, habitat choice, and patterns of plant recruitment. *American Journal of Botany*, **82**, 399–409, doi:10.2307%2F2445586.

Schurr, F.M., Spiegel, O., Steinitz, O., Trakhtenbrot, A., Tsoar, A. & Nathan, R. (2009) Long-distance seed dispersal. *Annual Plant Reviews*, **38**, 204–237.

Seidler, T.G. & Plotkin, J.B. (2006) Seed dispersal and spatial pattern in tropical trees. *PLoS Biology*, **4**, 2132–2137, doi:10.1371/journal.pbio.0040344.

Slatkin, M. (1985) Gene flow in natural populations. *Annual Review of Ecology Evolution and Systematics*, doi:10.2307/2097054.

Waples, R. & Gaggiotti, O. (2006) What is a population? An empirical evaluation of some genetic methods for identifying the number of gene pools and

15

their degree of connectivity. *Molecular Ecology*, **15**, 1419–1439, doi:10.1111/j. 1365-294X.2006.02890.x.

Wright, S. (1943) Isolation by distance. *Genetics*, **28**, 114–138.

Wright, S. (1946) Isolation by distance under diverse systems of mating. *Genetics*, **31**, 39–59.

Wright, S. (1969) *Evolution and the genetics of populations. Vol. 2: The theory of gene frequencies.* University of Chicago Press, Chicago, USA.

Table 1: Types of dispersal as a function of population area limits and genetic neighborhood limits. See Fig. 2 for a graphical representation of the four scenarios.

	Population geographic limit	
Genetic neighborhood limit	Within	Outside
Within	Local, short-distance dispersal, SDD_{loc}	Within neighborhood, long-distance dispersal, LDD_{neigh}
Outside	Local, long-distance dispersal, LDD_{loc}	Strict sense long-distance dispersal, LDD_{ss}

Table 2: Relative frequencies of *Prunus mahaleb* seed dispersal events for different frugivore groups according to population area limits and genetic neighborhood limits. See Fig. 2 for a graphical representation of the four scenarios. $N = 655$ seeds (see Table 1 in Jordano *et al.*, 2007). Given that the estimated neighborhood size is smaller than population area, LDD_{neigh} would be zero.

Frugivore group	Within-population, within-neighborhood SDD_{loc}	Within-population, long-distance LDD_{loc}	Outside-population, within-neighborhood LDD_{neigh}	Strict-sense long-distance LDD_{ss}	N seeds
Small-birds	0.7842	0.0171	0.00	0.1986	292
Turdus	0.2370	0.5549	0.00	0.2081	173
Large-birds	0.0435	0.3913	0.00	0.5652	23
Mammals	0.0120	0.2455	0.00	0.7425	167

Figures

Figure 1. The two approaches used in analyses of dispersal processes in plants. A, the "forward" approach attempts to track the dispersal events away from the known sources, e.g., by tracking the movement patterns of frugivores as they leave fruiting plants after feeding. B, the "backward" approach attempts to reconstruct the most likely source of a dispersed propagule by inferring the sources given the propagule delivery pattern, the fecundity of potential sources, and the dispersal function. The main technical challenge in A is to sample enough dispersal events away from the source to be able to fully characterize the tail (long-distance dispersal, LDD, events) of the dispersal function. In B, the main challenge is to have a robust sampling scheme with propagule collectors (e.g., seed traps) and a good characterization of the potential sources to derive robust estimates of the actual sources with inverse-modeling techniques.

Figure 2. Schematic representation of different types of long-distance dispersal events in relation to the geographical limits of local populations (dashed lines) and the genetic neighborhood area N_b (grey area) of specific individual plants (squares). Dispersal events (arrows) can be classified depending on their actual incidence on propagule movement outside these spatially-explicit reference areas (Table 1). Strict-sense long-distance dispersal events (LDD_{ss}) just include the LDD events that disseminate propagules out of *both* the population and genetic neighborhood boundaries. A, the neighborhood area is included within the geographic limits of the population; B, the neighborhood area is much larger than the geographic limits of the population.

Figure 3. Empirical frequency distributions of seed dispersal events as a function of dispersal distance by animal frugivores consuming *Prunus mahaleb* fruits. In red, left (inset), frequencies of within-populations dispersal events inferred from direct assignment based on seed endocarp genotypes and maternal trees genotypes. Larger frame, left, contributions of four functional frugivore groups (small birds, medium- and large-sized birds, and mammals) to seed dissemination and propor-

³⁴² tional contributions (right bar) to dispersal of the inferred immigrant seeds (i.e.,
³⁴³ those not matching any maternal tree in the study population).
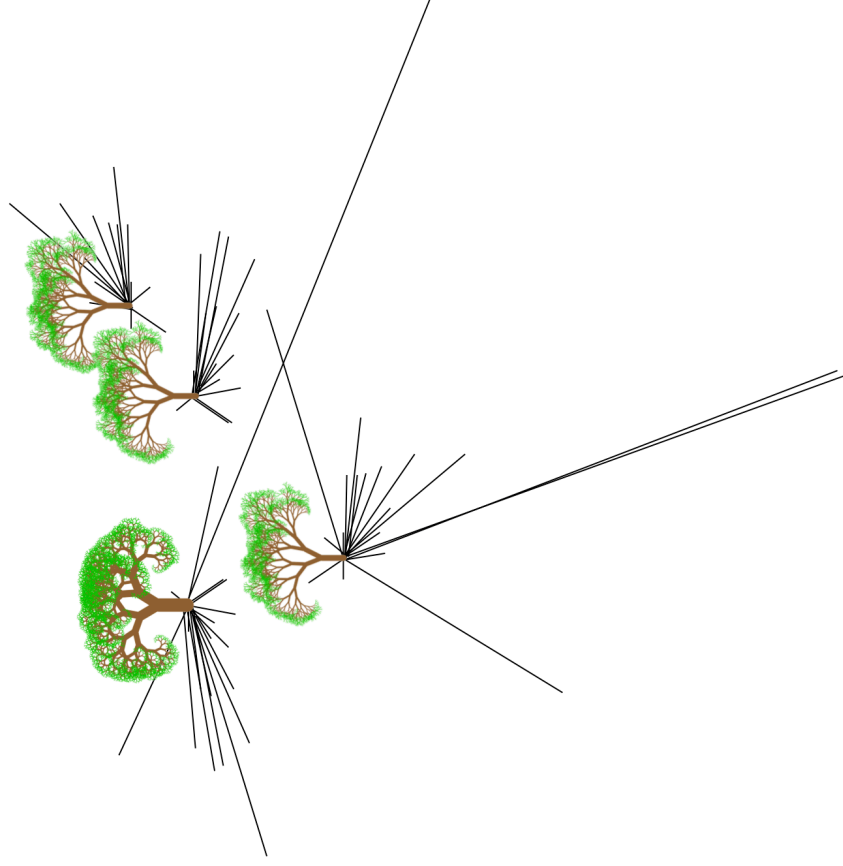
³⁴⁴

³⁴⁵ **Figure 4**. Differential contributions of functional groups of frugivores to the four
³⁴⁶ combinations of *Prunus mahaleb* seed dispersal events outlined in Table 1. These
³⁴⁷ result from dissemination within (yellow) or outside (blue) the population geo-
³⁴⁸ graphic limits (SDD_{loc}, LDD_{loc}, respectively) and within or outside the genetic
³⁴⁹ neigborhood area limits (SDD_{neigh}, LDD_{ss}, respectively).

³⁵⁰

Online Support Material and data accessiblity

This review does not use new raw data, but includes some re-analyses of previously published material. All the original data supporting the paper, R code, supplementary figures, and summaries of analytical protocols is available at the author's GitHub repository (`https://github.com/pedroj/MS_LDD`), with DOI: `#/zenodo.#.`
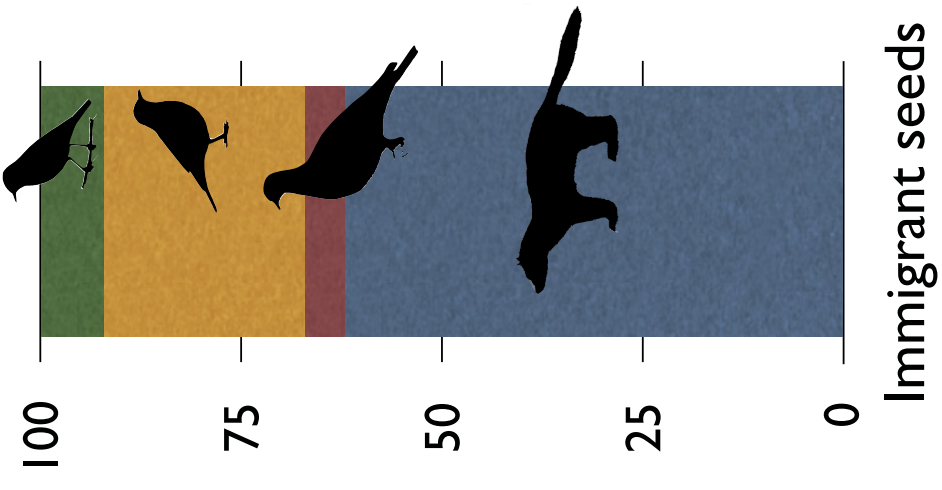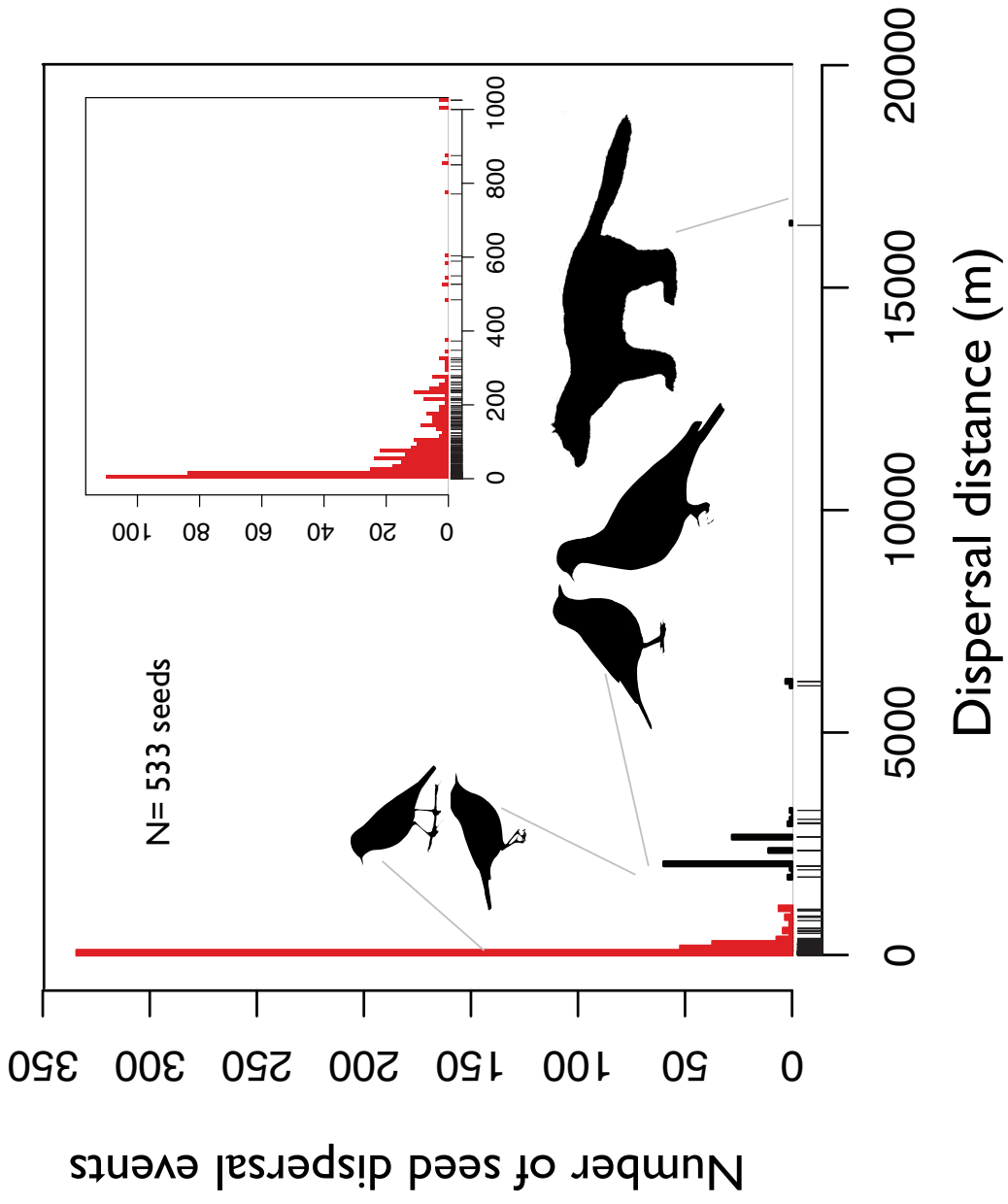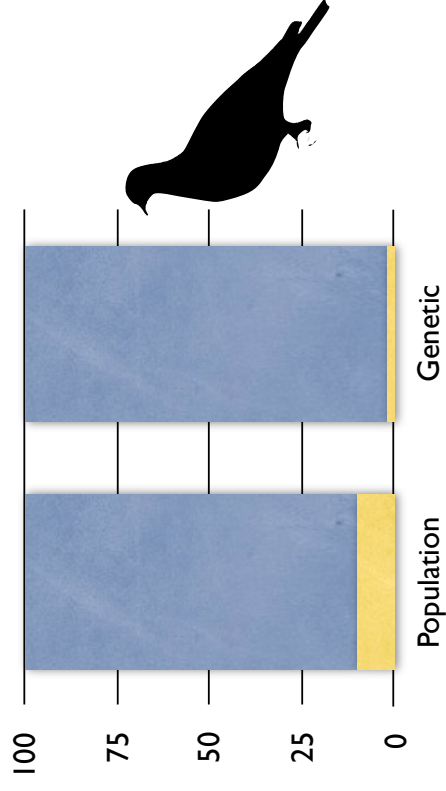
A. Emphasis on dispersal

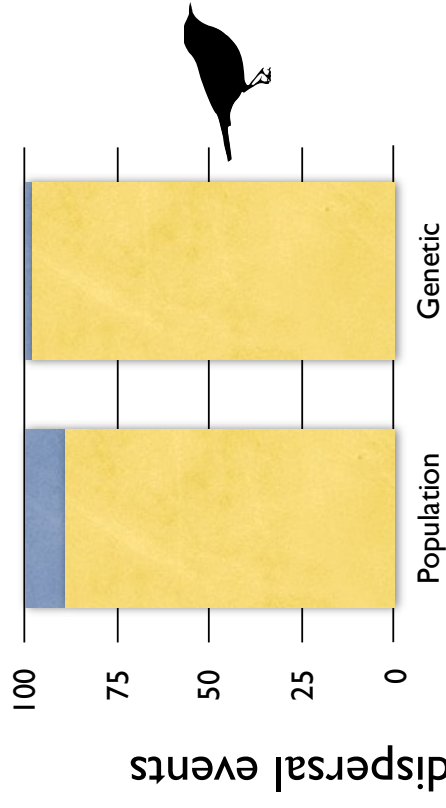B. Emphasis on arrival

A

LDD_{ss}

LDD_{loc}

SDD_{loc}

B

LDD_{neigh}

SDD_{loc}

LDD_{loc}

Number of seed dispersal events

Dispersal distance (m)

Immigrant seeds

N= 533 seeds

Percentage of seed dispersal events
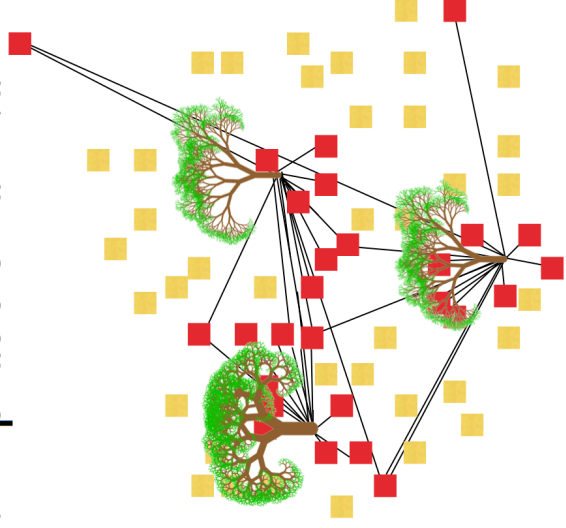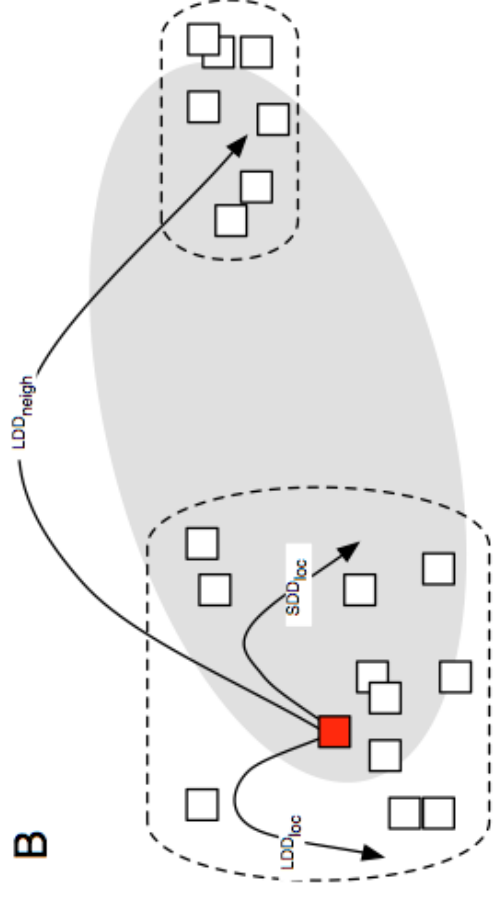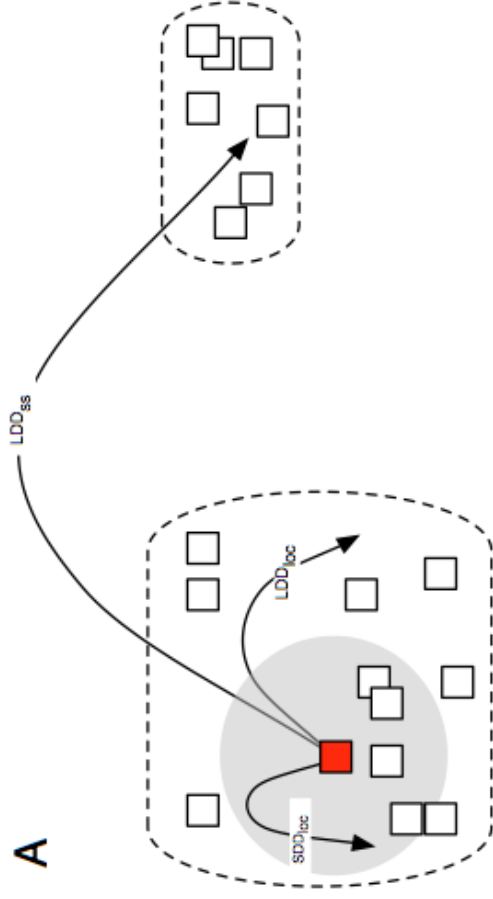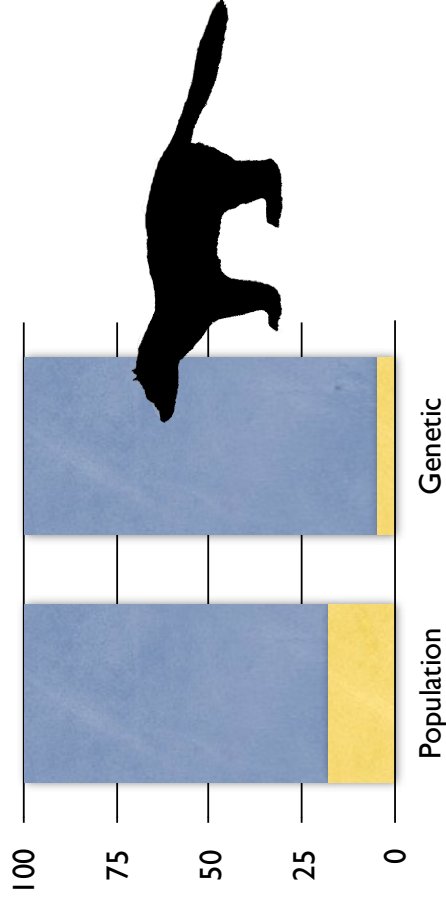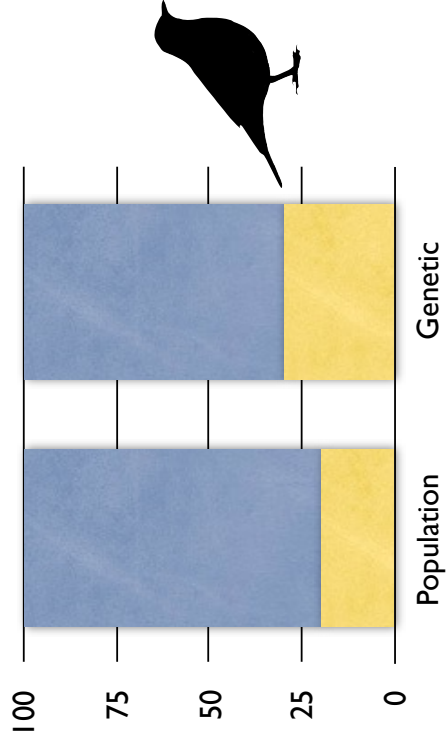
Within   Outside

Small birds   Large birds

Turdus   Mammals

Population   Genetic