

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

MS JEcol-2016-0422.R1

Pedro Jordano[†]

Sevilla, October 9, 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: dispersal, frugivory, plant-animal interactions, pollination, seed dispersal

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

Abstract

- 1 1. Dispersal is a key individual-based process influencing many life-history at-
2 tributes, scaling up to population-level properties (e.g., metapopulation connec-
3 tivity). A persistent challenge in dispersal ecology has been the robust characteri-
4 zation of dispersal functions (kernels), a fundamental tool to predict how dispersal
5 processes respond under global change scenarios. Especially the rightmost tail
6 of these functions, i.e. the long-distance dispersal (LDD) events, are difficult to
7 characterize empirically and to model in realistic ways.
- 8 2. But, when is it a LDD event? In the specific case of plants, dispersal has three
9 basic components: 1) a distinct (sessile) source, the maternal plant producing
10 the fruits or the paternal tree acting as a source of pollen; 2) a distance compo-
11 nent between source and target locations; and 3) a vector actually performing the
12 movement entailing the dispersal event. Here we discuss operative definitions of
13 LDD based on their intrinsic properties: 1) events crossing geographic boundaries
14 among stands; and 2) events contributing to effective gene flow and propagule
15 migration.
- 16 3. Strict-sense long distance dispersal involves movement both outside the stand
17 geographic limits and outside the genetic neighborhood area of individuals. Combi-
18 nations of propagule movements within/outside these two spatial reference frames
19 results in four distinct modes of LDD.
- 20 4. *Synthesis.* We expect truncation of seed dispersal kernels to have multiple
21 consequences on demography and genetics, following to the loss of key dispersal
22 services in natural populations. Irrespective of neighborhood sizes, loss of LDD

23 events may result in more structured and less cohesive genetic pools, with increased
24 isolation-by-distance extending over broader areas. Proper characterization of the
25 LDD events helps to assess, for example, how the ongoing defaunation of large-
26 bodied frugivores pervasively entails the loss of crucial LDD functions.

27

²⁸ Introduction

²⁹ Dispersal is a key individual-based process influencing many life-history attributes
³⁰ and scaling up to population-level properties (e.g., metapopulation connectivity,
³¹ Cousens *et al.* 2008). In the specific case of plants, largely sessile organisms,
³² dispersal has three basic components: 1) a distinct (sessile) source, the maternal
³³ plant producing the fruits or the paternal tree acting as a source of pollen; 2) a
³⁴ distance component between source and target locations; and 3) a vector actually
³⁵ performing the movement entailing the dispersal event. While realized dispersal
³⁶ also depends upon stages subsequent to dissemination (e.g., successful germination
³⁷ and seedling establishment) Schupp (1995), the three previous components fully
³⁸ characterize the dispersal process per se. Therefore, plant movement differs in
³⁹ important natural history details from animal dispersal, yet both can be assessed
⁴⁰ within a common conceptual framework (e.g., Nathan, 2006). Characteristically,
⁴¹ animal-assisted plant dispersal has three distinct, highly integrated, components
⁴² missing in the process of animal dispersal: the properties of the source (parental)
⁴³ plant, that mediate in the foraging of the animal vector (pollinator or frugivore),
⁴⁴ the intrinsic properties of the propagule, and the functional characteristics of the
⁴⁵ animal vector who performs the movement (Nathan *et al.*, 2008a).

⁴⁶ The movement of pollen and seeds by animals and its consequences have intrigued
⁴⁷ population geneticists and field ecologists since the infancy of both research dis-
⁴⁸ ciplines. Each has generated an impressive body of theoretical and empirical re-
⁴⁹ search through the past decades, yet advances have long been co-existing in parallel

50 worlds and the great synergistic potential of population genetics and demography
51 for the study of plant dispersal by animals remains little explored. Knowledge
52 gaps still having the imprint of this conceptual disconnection include the idea of
53 long distance dispersal, and the paradoxes of forest fragmentation effects on ge-
54 netic diversity (Kramer *et al.*, 2008), survival and persistence of relict tree species
55 (Hampe & Jump, 2011), rapid post-Pleistocene recolonization of vast continental
56 areas in response to climate modification (Clark *et al.*, 1998; Clark, 1998), among
57 other persisting issues. This conceptual isolation has been exacerbated by techni-
58 cal difficulties for the robust characterization of dispersal events, especially those
59 involving movement over long-distances (long-distance dispersal, LDD). LDD is a
60 characteristically extreme event of propagule movement in any plant population,
61 typically occurring with an extremely low probability but potentially reaching
62 an extremely long distance. Some progress has recently been made through the
63 fast-paced implementation of molecular tools in ecological research labs and the
64 availability of cutting-edge technology for biotelemetry applications. But much of
65 the population geneticist and ecologist communities remains unaware of the state
66 of the art in each other and likely under-appreciates their potential to validate
67 and enrich dispersal studies (Jones & Muller-Landau, 2008). In particular, LDD
68 events remain difficult to assess, both technically- with serious methodological
69 problems for its reliable estimation- and conceptually. Our aim here is to review
70 the LDD concept with a specific emphasis on dispersal of plant propagules (seeds
71 and pollen), providing an extended definition that might be helpful in the robust
72 quantification of LDD events.

While my main focus is on population-level scenarios and the role of animal vectors, recent research has shown how relevant is habitat and landscape context in determining the occurrence and directionality of LDD (e.g. Schurr et al. 2008 J Ecol; Carlo et al. 2013). On the other hand, dealing with habitat effects seems obligated if aimed to discuss global change drivers (i.e. habitat loss and fragmentation). 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 logically 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

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) and of the specific organism considered. 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 their 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.

119 Long-distance dispersal is currently one of the most debated topics in dispersal
120 ecology; it defines the connectedness within the network of local populations and
121 the possibilities for range expansion and successful colonization events. We pro-
122 pose a first demogenetically-based, operational definition of what a long-distance
123 dispersal event actually is, and review existing empirical literature on distance
124 thresholds from population and genetic perspectives. We also show how molecular
125 tools have been used to identify the respective contributions of different animal
126 species to the LDD portion of dispersal kernels of pollen and seeds by setting
127 empirically-derived distance thresholds. Finally, we highlight potential applica-
128 tions of molecular markers beyond the quantification of just the dispersal distances
129 that prevails in current studies, e.g., experimental approaches to assess dispersal
130 limitation and Janzen-Connell effects.

131 **LDD within a demo-genetic perspective: a taxonomy of dis-**
132 **persal events**

133 Here we propose an explicit definition of LDD and what constitutes a LDD event.
134 Previous definitions of dispersal patterns emphasized only their distance compo-
135 nents and characterized LDD events basically in terms of geographic distance be-
136 tween a dispersed propagule (or an established early seedling) and its most likely
137 maternal or paternal (in case of pollen) source. Absolute and proportional defini-
138 tions for the LDD events have been proposed depending on arbitrary thresholds
139 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).

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

163 neighborhood (Hardesty *et al.*, 2006; Jones & Muller-Landau, 2008). If we classify
164 dispersal events according to these two spatial frameworks (Table 1) we end up
165 with four distinct types of events depending on whether or not dispersed propagules
166 are disseminated within these reference areas. Setting the limits of a population
167 can be problematic (Waples & Gaggiotti, 2006) yet we can identify with relative
168 ease the geographical limits of plant stands, patches, habitat spots or other types
169 of habitat or microhabitat discontinuities that determine landmark boundaries of
170 biological significance (see Kinlan *et al.*, 2005, for further discussion of boundaries
171 for dispersal). These "frontiers" set biological limits to what a LDD event is in
172 relation to the geographic limits of the source population. Most plants are dis-
173 tributed as clumped patches, discrete stands, or relatively isolated populations, so
174 we may distinguish between short-distance and long-distance dispersal events that
175 end up with dissemination within or beyond, respectively, the stand or population
176 geographic boundaries (Table 1, SDD_{loc} or LDD_{loc}) (Figure 2).

177 A second consideration in terms of spatial boundaries, with effects on disper-
178 sal patterns, is the genetic neighborhood area N_e^b , i.e., the spatial extent includ-
179 ing a subset of panmictic individuals within a population (Wright, 1943, 1946).
180 Thus, the N_e^b area can be equal to the whole extent of the population whenever
181 the population is unstructured and there is evidence for random mating events
182 among all the individuals. However, most populations and stands of long-lived
183 trees show highly aggregated and clumped distributions (Seidler & Plotkin, 2006),
184 where relatively long distances may separate groups of individuals within the same
185 population. In these cases we might expect N_e^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 population, 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 SDD_{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.

206 Individual and Population Neighborhoods as Reference

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

$$N_e^b = 4\pi\sigma\delta \quad (1)$$

218 where δ is the density of adults per unit area and σ is the standard deviation of
 219 the distance between birth and breeding sites. This formulation is often called the
 220 neighborhood size and assumes a normal distribution of distances between parents
 221 and offspring (out in a perfect circular shape from the source). Thus, changes in
 222 the variance of dispersal distance can affect N_e^b (highly clumped populations will
 223 have reduced N_e^b). This is the basic model of "Isolation by Distance" proposed by
 224 Wright (1943, 1946). Under this type of model, migration (gene flow) is given by
 225 the variance in dispersal, and not by the proportion of the population that is com-

226 posed of migrants (denoted m), as is the case with island models (Slatkin, 1985).

227 With enough distance separating them, two plant individuals have a low probability
228 of mating and can be considered members of distinct genetic populations even
229 if they are not located in geographically distinct populations.

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

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

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

236 Neighborhood size in plants can be estimated by marking pollen and seeds with
237 fluorescent dyes, tags, or stable isotope enrichment (Carlo *et al.*, 2009). However,
238 these methods do not measure effective pollen or seed movement, but they may
239 be combined with genetic analysis to assess genetic identity and relatedness with
240 hypervariable DNA markers (Levin, 1988; Nason *et al.*, 1998; Godoy & Jordano,
241 2001) to achieve reliable estimates of both effective population size and neighbor-
242 hood area.

243 The extent of neighborhood area in plants can be extremely variable, depending
244 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 the size of neighborhood areas can in some cases exceed the geographic limits of local populations (Nason *et al.*, 1998). The size of neighborhood areas may encompass at least four orders of magnitude, $10^{-2} – 10^2$ km in radius, and include many individuals. Therefore, reference to this "genetic/evolutionary" paradigm and reference to the geographic boundaries (*sensu* Waples & Gaggiotti, 2006) may be instrumental to understand the actual role of LDD events in shaping the structuring of genetic pools and contributing to gene dispersal.

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

Empirical analysis of contributions to LDD

Empirical evaluation of differential contributions to the different forms of LDD events outlined in Table 1 requires identification of source trees as well as assignment of the dispersed propagules to specific vectors or functional groups of vectors (Jordano *et al.*, 2007). Recently, DNA-barcoding techniques have been developed and successfully applied to the identification of frugivore species contributing to specific seed dispersal events whose source can be identified with genetic, direct assignment techniques (González-Varo *et al.*, 2014). Otherwise, visual identifica-

tion can reliably assign the genotyped seeds to frugivore species groups based on specific characteristics of scats and regurgitations (Jordano *et al.*, 2007).

We inferred the frugivore groups contributing dispersal events by visually identifying scats and regurgitations in seed traps and line transects (see Jordano *et al.*, 2007, and Suppl. Mat. for additional details of methods). These frugivore functional groups include up to 38 bird and 4 mammal species feeding on *P. mahaleb* fruits (Jordano & Schupp, 2000). Here, we differentiate four major frugivore groups: large carnivorous mammals (such as foxes, badgers, and stone martens); two species of medium-sized frugivorous birds, mistle thrushes (*T. viscivorus*), 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 (Fig. S2). 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).

288 Empirically mapping of dispersal events for either pollen or seed disseminated by
289 animals may result in a complex pattern of different combinations of dispersal
290 events (Fig. S1), as animal movements are overlaid onto plant populations occu-
291 pying complex landscapes, resulting in different types of SDD and LDD events.

292 Long-Distance Dispersal: the ecology of extreme events

293 Long-distance dispersal (LDD) is a major component of the population dynamics,
294 genetic structure, and biogeographic history of plant species. It determines the
295 colonization ability of new habitats and the possibilities for fragmented populations
296 to sustain a cohesive metapopulation by immigration-emigration dynamics that
297 rely on LDD events (Nathan *et al.*, 2008b; Schurr *et al.*, 2009). Yet our current
298 understanding of the extent, frequency, and consequences of LDD is very limited.
299 On one hand, theoretical models fail to predict accurately the behavior of the tail
300 of the dispersal functions, and thus fail to predict very basic properties of LDD. On
301 the other hand, we still have very limited documentation of actual LDD events in
302 natural populations and we still see LDD as a sporadic, rarely far-reaching process
303 still marked with the stamp of natural history curiosity.

304 Combining spatially-explicit references to the geographic population limits and the
305 genetic neighborhood area extent (N_e^b) helps avoiding some imprecision in setting
306 distance thresholds to characterize LDD events (Jones & Muller-Landau, 2008).
307 In addition, the framework outlined in Table 1 bridges the combined demographic
308 and genetic effects of LDD events. When methods available to assign frugivore

309 taxa to the analyzed dispersal events, as in the study case with *P. mahaleb*, a
310 classification in the four categories of events is possible.

311 The frugivore assemblage of *P. mahaleb* is composed by a diversified set of animal
312 species spanning a wide size range, ca. 12-14000 g in body mass. We might expect
313 that this extreme variation translates in an ample pattern of foraging modes, move-
314 ment distances, and fruit/seed processing (Jordano & Schupp, 2000). If the results
315 for *P. mahaleb* are generalizable to other disperser assemblages, it seems that the
316 functional roles of frugivore species in terms of contributions to LDD events are
317 structured in two distinct groups: small-bodied frugivores, with substantial con-
318 tributions to SDD events, and large-bodied species with a disproportionate contri-
319 bution to LDD events. Both components of this sort of diplochorous (vander Wall
320 & Longland, 2004) dispersal system are very frequent in fleshy-fruited plants with
321 diversified frugivore assemblages (Galetti *et al.*, 2013). In such cases, small-bodied
322 frugivores largely contribute the short-distance dispersal key to support *in situ*
323 recruitment and population persistence. Yet the large-bodied frugivores distinctly
324 contribute LDD events that sustain the connectivity of metapopulation scenarios
325 (Urban & Keitt, 2001). As shown in Table 1, SDD and LDD events can be more
326 complex when we consider the contributions to gene flow via seed and the conse-
327 quences in terms of structure and spatial distributions of the genetic pools. For
328 example, local, within-population, dispersal events may vary enormously in terms
329 of genetic effects and local structuring of the genetic pools depending on whether
330 they specifically contribute SDD_{loc} or instead, LDD_{loc} . Note that only the latter
331 actually contribute erasing any form of local genetic structure by contributing to

332 increased genetic neighborhoods.

333 A number of classic studies have demonstrated that the activity of large frugivores
334 may also significantly contribute to SDD events and inefficient dispersal because
335 of, i.e., territorial defence, short gut retention times relative to on-tree foraging,
336 frequent revisit of same trees and perches, etc., resulting in substantial SDD
337 events (Pratt & Stiles, 1983; Pratt, 1984; Snow & Snow, 1984, 1988; Wheelwright,
338 1991). Yet these large-bodied frugivores are crucial for both LDD_{loc} and LDD_{ss} ,
339 given that extensive movement patterns and extremely large foraging ranges may
340 frequently contribute dissemination beyond distance thresholds defined with ei-
341 ther spatial landscape or genetic references. Recent analyses of the movement
342 ecology of large frugivores, coupled with results of their seed dispersal services
343 emphasize that LDD are by no means exceptional, either in terms of frequency
344 and extent (e.g., Westcott *et al.*, 2005; Bueno *et al.*, 2013; Morales *et al.*, 2013;
345 Carlo *et al.*, 2013). In addition, medium-sized birds such as thrushes (*Turdus* spp.)
346 can contribute substantial LDD_{loc} events, i.e., local LDD events contributing to
347 erase local population genetic structuring, effectively increasing the size of genetic
348 neighborhoods. In the case of *P. mahaleb* up to 55.49% of their dispersal events
349 are LDD_{loc} events. These birds are efficient seed dispersers of *P. mahaleb* and
350 other fleshy-fruited species (Snow & Snow, 1988; Jordano & Schupp, 2000; Carlo
351 *et al.*, 2013), also showing significant contributions of LDD_{ss} events.

352 Two-dimensional patterns in the *P. mahaleb* seed rain and the individual seed
353 shadows, accurately tracked with DNA-based genotyping methods, thus reflect
354 the complex effects of frugivore foraging, habitat preferences and heterogeneous

landscapes. This situation is probably generalizable to other plant–frugivore interactions (Morales and Carlo, 2006) where the combined spatial dynamics of habitat use and digestion processes determine complex seed shadows (Jordano et al., 2007; see Dennis and Westcott, [REF]; see Nathan, [REF]). Much of this complexity can be adequately handled by mechanistic models (Nathan et al., 2002; Morales and Carlo, 2006) incorporating very simple rules (Viswanathan et al., 2002). For example, earlier results (Jordano 2007) showed that the dispersal distances contributed by *P. mahaleb* frugivores closely map the spacing patterns of fruiting trees, but only up to a certain distance ($\leq 100m$; Fig. 10.3a in Jordano 2007). Beyond this, frugivores were probably responding to other major landscape elements (e.g. rock outcrops, forest edges, large patches of open grassland, etc) that cause the fat tail of the seed dispersal distribution, adding more frequent LDD events than expected from a Brownian random walk pattern generated by a tracking of the crops of the fruiting trees. For instance, the long flights performed by *T. viscivorus* (Jordano and Schupp, 2000) frequently faced the pine forest edge, at distances $\geq 100m$ of most *P. mahaleb* fruiting trees. If these medium-sized birds are selecting habitat with tall woody vegetation (e.g. pines $\geq 6m$ height), then they should be perceiving a much more patchy landscape, and thus requiring longer flights, than for example, small warblers seeking vegetation cover $< 0.5m$ (Fig. 10.3b in Jordano 2007).

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

378 efficient frugivorous bird species. Robustly characterizing the expected frequencies
379 and extent of those extreme events would be crucial to properly assess the func-
380 tional role of frugivores and the full range of influences (demographic, genetic) in
381 plant populations.

382 Challenges and future avenues for research

383 Pollen and seed dispersal in plants are essentially spatially-structured processes
384 for which the outcomes of interactions with dispersal vectors is intimately linked
385 to landscape features. Given this mechanistic link between the features of the
386 vector and the environments where its displacement occurs (Nathan *et al.*, 2008a),
387 consideration of landscape is key to understand the consequences of LDD events.
388 Yet these consequences hit two central aspects of plant life-histories: the demo-
389 graphic recruitment process (Harper, 1977), and the genetic signatures of pollen-
390 and seed-mediated gene flow in complex landscapes (Sork *et al.*, 1999). Recent
391 evidences point out that the selective extinction of large-bodied frugivores may
392 significantly impact plant populations dependent on frugivores both in terms of
393 recruitment (Traveset *et al.*, 2012; Pérez-Méndez *et al.*, 2015) and genetic con-
394 nectivity (Pérez-Méndez *et al.*, 2016). Frugivore downsizing represents a lasting
395 challenge for the collapse of seed dispersal processes where LDD_{ss} events are cru-
396 cial for population persistence and the cohesion of fragmented populations within
397 metapopulation scenarios.

398 We advocate (also see Jordano & Godoy, 2002; Nathan *et al.*, 2003; Jones &

399 Muller-Landau, 2008; Hardesty *et al.*, 2011) a combination of approaches includ-
400 ing large-scale biotelemetry to characterize animal movement, coupled with large-
401 scale genetic sampling of dispersed propagules, and demogenetic approaches that
402 combine both demographic and genetic research. A crucial aspect would be to
403 effectively associate the role of individual frugivore species to specific dispersal
404 outcomes, by identifying the actual disperser contributing a dissemination event
405 (González-Varo *et al.*, 2014) and simultaneously characterizing the source maternal
406 plant (Jordano & Godoy, 2002).

407 The actual challenges to properly characterize the typologies of LDD events out-
408 lined in Table 1 will probably persist. We need more efficient quantitative ap-
409 proaches to assess these infrequent events, that occur over enormous spatial scales
410 and that need to be documented with sample sizes sufficient to facilitate modeling
411 efforts and robust statistical inferences. These are not trivial difficulties given the
412 urgency to assess how forest loss, defaunation, genetic purging due to logging, etc.,
413 alter plant populations.

414 *Acknowledgements.* I am indebted to Cristina Garca, Jos A. Godoy, Manolo Car-
415 rin, Juan Luis Garca-Castao, Jess Rodrguez and, especially, Juan Miguel Arroyo
416 for generous help with field and laboratory work and making possible this study.
417 I appreciate the help and advice of Cristina Garca and Etienne Klein during the
418 final stages of the manuscript. The study was supported by a Junta de Andaluca
419 Excellence Grant (RNM-5731), as well as a Severo Ochoa Excellence Award from
420 the Ministerio de Economia y Competitividad (SEV-2012-0262). The Agencia de
421 Medio Ambiente, Junta de Andaluca, provided generous facilities that made pos-
422 sible this study in the Andalusian natural parks (Sierra de Cazorla, Alcornocales)
423 and authorized my work there.

References

- Bueno, R.S., Guevara, R., Ribeiro, M.C., Culot, L., Bufalo, F.S. & Galetti, M. (2013) Functional Redundancy and Complementarities of Seed Dispersal by the Last Neotropical Megafrugivores. *PLoS ONE*, **8**(2), e56252, doi:10.1371/journal.pone.0056252.s003.
- Carlo, T.A., García, D., Martínez, D., Gleditsch, J.M. & Morales, J.M. (2013) Where do seeds go when they go far? Distance and directionality of avian seed dispersal in heterogeneous landscapes. *Ecology*, **94**(2), 301–307.
- Carlo, T., Tewksbury, J., Martínez, C. & Mart (2009) A new method to track seed dispersal and recruitment using ^{15}N isotope enrichment. *Ecology*, **90**, 3516–3525.
- Clark, J.S., Fastie, C., Hurttt, 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.

Galetti, M., Guevara, R., Côrtes, M.C., Fadini, R., Von Matter, S., Leite, A.B., Labecca, F., Ribeiro, T., Carvalho, C.S., Collevatti, R.G., Pires, M.M., Guimaraes, P.R., Brancalion, P.H., Ribeiro, M.C. & Jordano, P. (2013) Functional extinction of birds drives rapid evolutionary changes in seed size. *Science*, **340**(6136), 1086–1090, doi:10.1126/science.1233774.

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. *An-*

nual 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.

Hardesty, B.D., Metcalfe, S.S. & Westcott, D.A. (2011) Persistence and spread in a new landscape: Dispersal ecology and genetics of Miconia invasions in Australia. *Acta Oecologica*, **37**(6), 657–665, doi:10.1016/j.actao.2011.06.006.

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. & Godoy, J. (2002) Frugivore-generated seed shadows: a landscape view of demographic and genetic effects. *Seed dispersal and frugivory: ecology,*

evolution, and conservation (eds. D. Levey, W. Silva & M. Galetti). Book chapter, Wallingford, UK, pp. 305–321.

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.

Morales, J.M., García, D., Martínez, D., Rodríguez-Pérez, J. & Herrera, J.M. (2013) Frugivore Behavioural Details Matter for Seed Dispersal: A Multi-Species Model for Cantabrian Thrushes and Trees. *PLoS ONE*, **8**(6), e65216, doi:10.1371/journal.pone.0065216.s005.

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

Nathan, R., Perry, G., Cronin, J.T., Strand, A.E. & Cain, M.L. (2003) Methods for estimating long-distance dispersal. *Oikos*, **103**, 261–273.

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.

Pérez-Méndez, N., Jordano, P. & Valido, A. (2015) Downsized mutualisms: Consequences of seed dispersers' body-size reduction for early plant recruitment. *Perspectives In Plant Ecology Evolution And Systematics*, **17**(2), 151–159, doi: 10.1016/j.ppees.2014.12.001.

Pérez-Méndez, N., Jordano, P. & Valido, A. (2016) Persisting in defaunated landscapes: reduced plant population connectivity after seed dispersal collapse. *In 2nd. review.*

Pratt, T. (1984) Examples of tropical frugivores defending fruit-bearing plants. *Condor*, **86**, 123–129.

- Pratt, T. & Stiles, E. (1983) How long fruit-eating birds stay in the plants where they feed - implications for seed dispersal. *The American Naturalist*, **122**(6), 797–805.
- 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 and Systematics*, doi:10.2307/2097054.
- Snow, B. & Snow, D. (1984) Long-term defence of fruit by mistle thrushes *Turdus viscivorus*. *Ibis*, **126**, 39–49.
- Snow, B. & Snow, D. (1988) *Birds and berries*. Poyser, Calton, UK.
- Sork, V., Nason, J., Campbell, D. & Fernández, J. (1999) Landscape approaches to historical and contemporary gene flow in plants. *Trends In Ecology & Evolution*, **14**, 224–224.
- Traveset, A., González-Varo, J.P. & Valido, A. (2012) Long-term demographic consequences of a seed dispersal disruption. *Proceedings Of The Royal Society B-Biological Sciences*, doi:10.1098/rspb.2012.0535.

Urban, D. & Keitt, T. (2001) Landscape connectivity: a graph-theoretic perspective. *Ecology*, **82**, 1205–1218.

vander Wall, S. & Longland, W. (2004) Diplochory: are two seed dispersers better than one? *Trends In Ecology & Evolution*, **19**, 155–161, doi:10.1016/j.tree.2003.12.004.

Waples, R. & Gaggiotti, O. (2006) What is a population? An empirical evaluation of some genetic methods for identifying the number of gene pools and their degree of connectivity. *Molecular Ecology*, **15**, 1419–1439, doi:10.1111/j.1365-294X.2006.02890.x.

Westcott, D.A., Bentupperbaumer, J., Bradford, M.G. & McKeown, A. (2005) Incorporating patterns of disperser behaviour into models of seed dispersal and its effects on estimated dispersal curves. *Oecologia*, **146**, 57–67.

Wheelwright, N. (1991) How long do fruit-eating birds stay in the plants where they feed? *Biotropica*, **23**(1), 29–40.

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.

Genetic neighborhood limit	Population geographic limit
Within	Within
Outside	Outside
Within	Local, short-distance dispersal, SDD_{loc}
Outside	Local, long-distance dispersal, LDD_{loc}
Within	Within neighborhood, long-distance dispersal, LDD_{neigh}
Outside	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
<i>Turdus</i>	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

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

436

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

445 graphic limits of the population, with some dispersal events potentially contribut-
446 ing local LDD; B, the neighborhood area is much larger than the geographic limits
447 of the population.

448

449 **Figure 3.** Empirical frequency distributions of seed dispersal events as a function
450 of dispersal distance by animal frugivores consuming *Prunus mahaleb* fruits. In
451 red, left (inset), frequencies of within-population dispersal events inferred from di-
452 rect assignment based on seed endocarp genotypes and maternal trees genotypes.
453 Larger frame, left, contributions of four functional frugivore groups (small birds,
454 medium- and large-sized birds, and mammals) to seed dissemination and propor-
455 tional contributions (right bar) to dispersal of the inferred immigrant seeds (i.e.,
456 those not matching any maternal tree in the study population) (Jordano *et al.*,
457 2007).

458

459 **Figure 4.** Differential contributions of functional groups of frugivores to the
460 four combinations of *Prunus mahaleb* seed dispersal events outlined in Table 1.
461 These result from dissemination within (yellow) or outside (blue) the population
462 geographic limits; within-population dispersal events can either be short-distance
463 (SDD_{loc}) or local LDD (LDD_{loc}) depending on the size of the genetic neigborhood.
464 Dispersal outside the local population can entail short-distance dispersal, if within
465 the genetic neighborhood area limits (SDD_{neigh}) (yellow) or represent strict-sense
466 LDD (LDD_{ss}) (blue).

467

468 **Online Support Material and data accessibility**

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







