Who dispersed the seeds? The use of DNA barcoding in frugivory and seed dispersal studies

Juan P. González-Varo*, Juan M. Arroyo and Pedro Jordano

Integrative Ecology Group, Estación Biológica de Doñana, EBD-CSIC, Avda. America Vespucio S/N, Isla de La Cartuja, E-41092 Sevilla, Spain

Summary

1. Assessing dispersal events in plants faces important challenges and limitations. A methodological issue that limits advances in our understanding of seed dissemination by frugivorous animals is identifying ‘which species dispersed the seeds’. This is essential for assessing how multiple frugivore species contribute distinctly to critical dispersal events such as seed delivery to safe sites, long-distance dispersal and the colonization of non-occupied habitats.

2. Here, we describe DNA-barcoding protocols successfully applied to bird-dispersed seeds sampled in the field. Avian DNA was extracted from the surface of defecated or regurgitated seeds, allowing the identification of the frugivore species that contributed each dispersal event. Disperser species identification was based on a 464-bp mitochondrial DNA region (COI: cytochrome c oxidase subunit I).

3. We illustrate the possible applications of this method with bird-dispersed seeds sampled in the field. DNA-barcoding provides a non-invasive technique that allows quantifying frugivory and seed dispersal interaction networks, assessing the contribution of each frugivore species to seed rain in different microhabitats, and testing whether different frugivore species select different fruit/seed sizes.

4. DNA barcoding of animal-dispersed seeds can resolve the distribution of dispersal services provided by diverse frugivore assemblages, allowing a robust and precise estimation of the different components of seed dispersal effectiveness, previously unattainable to traditional field studies at individual seed level. Given that seeds are sampled at the end of the dispersal process, this technique enables us to link the identity of the disperser species responsible for each dispersal event to plant traits and environmental features, thereby building a bridge between frugivory and seed deposition patterns.

Key-words: Barcode of Life Data system, COI region, endozoochory, frugivores, interaction network, seed deposition patterns, fruit size selection, Mediterranean woodland, microhabitats

Introduction

Seed dispersal mediated by vertebrate frugivores is a central process in the dynamics and regeneration of many vegetation types (Fleming & Kress 2013). Frugivores ingest fleshy fruits and regurgitate or defecate seeds in conditions that may be suitable for germination and the early establishment of seedlings. This involves a mutualistic interaction with plants that can be identified in, for example, up to 98% of tropical rain forest or 60% of Mediterranean shrubland woody species (Jordano 2014). This mutualism is thus pivotal for supporting mega-diversified communities in which multiple species interact, thereby combining an extraordinary diversity of ecological services needed for forest regeneration (García, Zamora & Amico 2010), the colonization of vacant habitats after disturbance (Carlo & Yang 2011) and the inter-population connectivity mediated by long-distance seed dispersal (Nathan 2007).

Two methodological challenges that have hindered advances in our understanding of vertebrate-mediated dispersal are the identification of seed sources (‘from which fruit-bearing tree did the frugivore ingest the seeds?’) and the identification of the frugivore species disseminating seeds in a particular microsite (‘which frugivore species dispersed the seed?’) (Jordano et al. 2007; García & Griivet 2011; Cortés & Uriarte 2013; Jordano 2014). The identification of the source of dispersed seeds, which enables direct estimates of dispersal distances and the detection of habitat transitions, has been achieved by a diverse array of methods such as direct observations of disperser foraging activity (Gómez 2003), microsatellite DNA markers (Godoy & Jordano 2001; Ashley 2010), isotopic markers (Carlo et al. 2013), and the use of labelled seeds (Mack 1995) and colour-coded seed mimics (González-Varo, López-Bao & Guitián 2013). In contrast, the identification of the vector of animal-dispersed seeds has been much more challenging, usually undertaken by direct observation or by the visual identification of faecal remains in the field (e.g. Jordano et al. 2007; González-Varo, López-Bao & Guitián 2013). This poses obvious limitations, especially when attempting to identify closely related species, and obliges the grouping of frugivore species into functional groups (Jordano et al.)
use our data set of ‘individual seeds with DNA-identified frugivore–seed disperser’ in three application examples to answer the following questions: (i) Which frugivores disperse the seeds of which plant species? (ii) What is the contribution of different frugivore species to seed arrival in different microhabitats? and (iii) Do different frugivores select for different fruit/seed sizes?

Materials and methods

DNA region and primer design

Avian disperser identification was based on a 464-bp mitochondrial DNA region of the cytochrome c oxidase subunit I (COI) employing the ‘Barcode Of Life Data’ identification system (BOLD: http://www.boldsystems.org; Ratnasingham & Hebert 2007). BOLD accepts sequences from the 5’ region of the COI gene and returns species-level identification whenever possible and assigns a percentage of similarity to matched sequences.

Primers already designed for degraded avian DNA did not work successfully in our avian assemblage, as we found after checking sequences reported by Lijtmaer et al. (2012) with Sequencer 4.9 (Gene Codes Corp., Ann Arbor, MI, USA) and after testing primers designed by D. Pastor-Bévia et al. (unpublished data) in our samples. Indeed, Lohman, Prawiradilaga & Meier (2009) also described problems with amplifying the COI region using standard primers and recommended newly designed primers for Passeriformes. Consequently, we designed new primers (COI-fsdF: 5’– GCA TGA GCC CAA TAT TGG G–3’; COI-fsdR: 5’– TGT GAK AGG GCG TGG TT–3’) using one reference COI sequence of each of 16 avian species occurring in our study area, most of them reported as frugivorous seed dispersers (see Table S1). We acquired COI sequences from BOLD data bases by selecting individual samples that included the Iberian Peninsula as breeding and/or wintering grounds. Primers were designed by searching for low-variability regions spanning variable regions after the sequences were aligned using Sequencer 4.9. Degeneracy was included in one base position of each primer. This pair of primers was tested on non-degraded DNA isolated from feather, blood or muscle samples of 16 selected species (see Table S1). PCR amplifications of all species yielded a 464-bp product that was sequenced and verified for its matching with COI sequences from BOLD data bases.

Validation in bird faeces and defecated/regurgitated seeds

Validation material

In order to verify the correct identification of bird species in defecated or regurgitated material, we tested the primers on samples of known species provenance. We collected faeces without seeds and defecated or regurgitated seeds from birds captured in mist-netting sessions carried out within a routine bird-ringing programme in southern Spain (Cadiz province). Mist nets were operated in woodland habitats between November 2012 and January 2013. We collected bird droppings (i) from sterile filter paper placed within the cloth bags used to keep the birds during ringing sessions and (ii) from 1 m × 10 m plastic mesh (<0.5 mm pore) placed beneath the mist nets as we frequently observed fresh dropping beneath birds trapped in nets. We used sterile-disposable tweezers to pick up samples and place them in 1.5- or 2.0-mL sterile tubes (see Fig. S1). We obtained a total of 23 samples, 6 faeces and 17 defecated/regurgitated seeds (belonging to Pistacia...
lentiscus and Olea europaea), from a total of 23 individuals belonging to nine bird species (see Table 1), most of them known frugivores and legitimate seed dispersers (Herrera 1984). Additionally, we collected 17 defecated/regurgitated seeds of P. lentiscus and O. europaea beneath a roosting perch of spotless starlings (Sturnus unicolor). For validation, we checked the correspondence between the COI DNA sequence isolated from these samples and those from their known disperser species obtained in BOLD (see Table 1).

Sample processing

In the initial step of extraction, three methods for recovering avian DNA from the surface of the defecated/regurgitated seeds and faeces were evaluated and compared: ‘method 1’: the seed was wrapped with filter paper following Marrero et al. (2009); ‘method 2’: the seed surface was rubbed with a sterile cotton swab (Ramón-Laca & Gleeson 2014) soaked in phosphate buffer; and ‘method 3’: the seed was directly incubated in extraction buffer with rotation but no vortexing. Two DNA extraction protocols were tested on validation samples: ‘protocol 1’, a GuSCN/silica protocol for ancient DNA modified from Höss & Paabo (1993) and Rohland, Siedel & Hofreiter (2010) and ‘protocol 2’, the GuSCN/silica protocol of Marrero et al. (2009), originally used for regurgitated seeds and faeces from Canary Islands pigeons. Both protocols are based on GuSCN and silica but differ in extraction buffers, incubation times, amount of silica added and the silica pellet washing (see details in Table S2).

SAMPLING BIRD-DISPERSED SEEDS IN THE FIELD

We sampled bird-dispersed seeds in the field in order to apply the method and identify their disperser species. Sampling was carried out between 29 October 2013 and 30 January 2014 in Garrapilos, a Mediterranean lowland forest fragment (50 m a.s.l., c. 120 ha) located in Cádiz province, southern Spain (36°39’ N, 5°56’W). Vegetation consists of large holm- (Quercus ilex subsp. ballota) and cork- (Q. suber) oak, and an understorey dominated by treelets and shrubs, among which Quercus coccifera (Fagaceae), Pistacia lentiscus (Anacardiaceae), Olea europaea var. sylvestris (Oleaceae), Rhamnus alaternus (Rhamnaceae) and Crataegus monogyna (Rosaceae) are the dominant species.

We sampled bird-dispersed seeds in 40 seed traps with locations randomly stratified across three microhabitat types: 14 beneath oaks, 14 beneath treelets/shrubs bearing fleshy fruits (Pistacia female plants, Olea and Crataegus) and 12 beneath treelets/shrubs not bearing fleshy fruits (Pistacia male plants, Q. coccifera and R. alaternus, the latter a summer-ripening species). Seed traps consisted of plastic trays (40 cm × 55 cm, 8 cm height) with small holes (1 mm diameter) to allow the drainage of rainwater and covered with wire mesh (1 cm light) to prevent post-dispersal seed predation by vertebrates. Additionally, we set up six fixed transects (23- to 45-m long and 1-m wide) in open ground areas where bird-mediated seed rain is less likely (Jordano & Schupp 2000) and post-dispersal seed predation is typically low due to the lack of shelters for rodents (Fediriani & Manzaneda 2005). Finally, we also sampled bird-dispersed seeds through direct searches in order to increase sample sizes in under-sampled microhabitats. Seed traps and fixed transects were set up on 29 October 2013 and were monitored weekly or biweekly until 30 January 2014. The sampling period (November–January) spanned the peak abundances of different fleshy-fruited species in Mediterranean woodlands (Jordano 1988).

We sampled each bird-dispersed seed (or individual faeces containing seeds) putting it with a minimum of handling into a 1.5- or 2.0-mL sterile tube (depending upon seed size; see Fig. S1) with the aid of the tube cap; alternatively, each seed can be collected with sterile, disposable tweezers. Tubes were labelled with information regarding seed identity, microhabitat identity and trap/transect code and then stored in a freezer at −20°C until DNA extraction.

DNA ISOLATION, AMPLIFICATION AND SEQUENCING

DNA extractions were performed in a laboratory dedicated to low DNA concentration procedures, and negative controls were included in each extraction to check for contamination. We followed ‘method 3’ (i.e. seed incubation in extraction buffer) combined with ‘protocol 1’ (i.e. a GuSCN/silica protocol for ancient DNA; modified from Höss & Paabo 1993 and Rohland, Siedel & Hofreiter 2010), which gave higher yields of PCR amplifications (see Results; Table S3). Nevertheless, a sterile swab soaked in phosphate buffer (‘method 2’) was used when the seed surface had an excess of faecal material and/or pulp residues.

A volume of 450 or 500 µL of extraction buffer (Longmire, Maltbie & Baker 1997) (0-1 M Tris-Cl pH 8.0, 0-1 M EDTA pH 8-0, 0-01 M NaCl, 0.5% SDS and 0.25 mg mL⁻¹ Proteinase K) was added to the 1.5- or 2.0-mL tubes, respectively, containing seeds or swabs, and incubated in rotation at 50°C for 2.5 h. Supernatant (c. 400-450 µL) was transferred to a new 2.0-mL tube; then, 1-4 mL of binding buffer (5 M GuSCN, 0.1 M Tris-HCl pH 6-4, 0.02 M EDTA pH 8.0 and 1-3%) Triton X-100) and 120 µL of silica suspension were added, and the mix was incubated in rotation at room temperature for 2.5 h in the dark. This step allows the binding of DNA to silica particles in the presence of high salt concentration. After centrifugation (2 min at 17000 g), the supernatant was discarded, and the silica pellet was resuspended in 400 µL of binding buffer and transferred to columns (MobiTec, Germany, product # M1002S) with a glass microfiber filter (Whatman Grade GF/B 1-0 µm) on the top of the 10-µm column filter. After centrifugation (1 min at 17 950 g), silica particles retained in the column were washed at least twice using 450 µL of washing buffer.

Table 1. Summary of DNA samples (faecal and defecated or regurgitated seeds) with known source species (passerine birds, most of them frugivores-seed dispersers) used for validation of the amplification of COI region and the subsequent species identification in BOLD data bases. All samples were obtained during bird-ripening sessions, except those from Sturnus unicolor, which were obtained beneath resting sites of winter flocks. Pl: Pistacia lentiscus seeds; Oc: Olea europaea seeds.

<table>
<thead>
<tr>
<th>Bird species (n individuals)</th>
<th>n total samples</th>
<th>n faecal samples</th>
<th>n defecated or regurgitated seeds</th>
<th>Similarity (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Carduelis chloris</td>
<td>1</td>
<td>1</td>
<td>–</td>
<td>100</td>
</tr>
<tr>
<td>Fringilla coelebs</td>
<td>1</td>
<td>1</td>
<td>–</td>
<td>100</td>
</tr>
<tr>
<td>Erithacus rubecula</td>
<td>1</td>
<td>1</td>
<td>–</td>
<td>100</td>
</tr>
<tr>
<td>Parus major</td>
<td>1</td>
<td>1</td>
<td>–</td>
<td>100</td>
</tr>
<tr>
<td>Saxicola rubicola</td>
<td>1</td>
<td>1</td>
<td>–</td>
<td>99.4</td>
</tr>
<tr>
<td>Sturnus unicolor*</td>
<td>17</td>
<td>–</td>
<td>17 (2 Pl, 15 Oc)</td>
<td>99.4–100</td>
</tr>
<tr>
<td>Sylvia atricapilla</td>
<td>14</td>
<td>1</td>
<td>13 (7 Pl, 6 Oc)</td>
<td>99.7–100</td>
</tr>
<tr>
<td>Sylvia</td>
<td>3</td>
<td>–</td>
<td>3 (3 Pl)</td>
<td>100</td>
</tr>
<tr>
<td>Turdus philomelos</td>
<td>1</td>
<td>–</td>
<td>1 (1 Oc)</td>
<td>100</td>
</tr>
<tr>
<td>Total (9 species)</td>
<td>40</td>
<td>6</td>
<td>34 (12 Pl, 22 Oc)</td>
<td>99.4–100</td>
</tr>
</tbody>
</table>

*Discerning between Sturnus vulgaris and S. unicolor was not possible owing to the low degree of genetic differentiation between these congeneric species; indeed, some authors treat them as subspecies (see Lovette et al. 2008 and references therein). We assigned our samples to S. unicolor based on field observations.
We used seeds sampled in seed traps and fixed transects to calculate the magnitude of seed rain (seeds m⁻² per microhabitat) but used all sampled seeds, including those collected during direct searches, to quantify the relative contribution (%) of each frugivore species to the seed rain in each microhabitat type (see Jordano & Schupp 2000).

Example 3. Do different frugivores select for different fruit/seeds sizes?

We used data from wild olive (O. europaea)-dispersed seeds to test whether different frugivore species select different fruit sizes. We chose the wild olive because its fruits are large enough to prevent fruit swallowing by small bird species (Rey et al. 1997). The gape width of local frugivorous birds ranges between 7.1 mm (Sylvia melanocephala) and 13.7 mm (Turdus philomelos) (Rey et al. 1997), whereas wild olive fruits in the study site ranged between 6.3 and 13.7 mm in diameter and 9.1 and 20.9 mm in length (spherical Pistacia fruits are typically <6 mm diameter). Seed dry weight is highly correlated with the weight, length and diameter of fresh wild olive fruits ($r^2 = 0.83-0.90; P < 0.001, n = 60$). Thus, once Olea (defleshed) seeds were processed for DNA-barcoding analysis and their avian-frugivore species was identified, we measured their dry weight to test for differences between frugivore species in fruit size consumption.

Results

SAMPLE PROCESSING AND VALIDATION

Validation tests were carried out on faeces and defecated/regurgitated seeds whose source frugivore species was known in advance. For recovering avian DNA from the surface of the dispersed seeds, ‘method 3’ (i.e. direct incubation of seeds in extraction buffer) resulted in satisfactory yields of PCR products for sequencing. Moreover, this was practical and feasible as all our seed species fitted into the 1.5- to 2.0-mL tubes (see Fig. S1). It also was very efficient as most seed samples had little avian-originated material (see Fig. S2). Wrapping seeds in filter paper (‘method 1’) was not suitable for Pistacia lentiscus seeds due to their small size (see Fig. S1), yielding insufficient PCR amplifications. The sterile swab (‘method 2’) was more efficient for recovering avian-originated material (DNA) from the seed surface but required more manipulation of the seeds that, in some cases, entailed the partial loss of material. However, it produced better results when the seed surface had an excess of faecal material and/or pulp remains, probably because avian-originated DNA is present on the external layers (those recovered with the swab) of the dispersed seeds.

‘Protocol 1’, which included different buffers for sample digestion and DNA binding steps, and columns for washing and elution steps, resulted in significantly higher (>2-fold) yields of PCR amplifications than ‘protocol 2’ (Mann–Whitney U-test: Z = 3.5, P < 0.001; Table S3). The resulting yields were also more consistent across individual samples (Table S3). The longer incubation times (in both digestion and DNA binding steps) probably improve the breakage of intact tissue structure and DNA binding to silica. Moreover, columns allow a more efficient washing for removing inhibitors and the elution steps.
The expected 464-bp PCR product was amplified, and all processed samples (faeces and defecated/regurgitated seeds whose source species were known in advance) were successfully validated (Table 1). After scoring, we obtained a 390- to 420-bp sequence length owing to the fact that templates were sequenced on one strand (forward primer). Species identification was correct in all cases based on a 99–100% of sequence similarity (see results in Table 1). With the exception of Sturnus unicolor, the dispersers of those seeds (Fig. 1): blackcap (Sylvia atricapilla, n = 81), European robin (Erithacus rubecula, n = 47), song thrush (Turdus philomelos, n = 45), Sardinian warbler (Sylvia melanocephala, n = 23) and woodpigeon (Columba palumbus, n = 2). The disperser species was successfully identified in 90% (197) of the samples (based on a similarity threshold > 99%, i.e. 99–100%). Most unsuccessful identifications were due to no DNA amplification (2.8% of samples) or to the presence of unspecific amplifications from exogenous DNA (i.e. non-avian DNA; 7.2% of samples).

Despite great differences in size, shape and coat texture (Fig. S1), the percentage of successful disperser identification was very similar for the two main seed species in the data set, P. lentiscus (89%; 99 seeds) and O. europaea (90%; 94 seeds). The disperser species was also identified for the four C. monogyna seeds and the single M. communis seed.

Identification success was higher for seeds sampled in November (97%) and December (95%), months dominated by non-rainy days, than for those sampled in January (81%), when rainy days prevailed and seeds in traps were frequently damp. Identification success was in general high and similar for seeds sampled in different microhabitats (Table S4), with all values between 81% and 100%, except for O. europaea seeds sampled beneath female Pistacia shrubs, with only 63% of seed successfully identified. This low value is likely to be related to the fact that most seeds collected in that microhabitat (9 out of 11) came from direct searches and the time period elapsed since deposition by birds was unknown.

**APPLICATION EXAMPLES**

Which frugivores disperse the seeds of which plant species?

Using seeds whose disperser was successfully DNA-identified (n = 198), we recorded a total of 11 (out of 20 possible) distinct ‘frugivory–seed dispersal’ interactions between species of fleshy-fruited plants and frugivorous birds (Fig. 1). The weighted interaction network shows how different bird species consume the fruits and disperse the seeds of plant species at different frequencies and how they vary in their role as seed dispersers (Fig. 1). It also illustrates how the fruit species varied in their contribution to the frugivorous diet of each bird species. For example, although both S. atricapilla and S. melanocephala ate the fruits and dispersed the seeds of the same species (O. europaea and P. lentiscus), S. atricapilla relied on these two plants more evenly than S. melanocephala. In turn, while most of the frugivorous diet of S. melanocephala consisted of P. lentiscus fruits (96%), the dispersal services of P. lentiscus only depended partially on this bird (22%). As well, despite being consumed by all five bird species, just two (T. philomelos and S. atricapilla) accounted for 96% of seed dispersal services (90 out of 94 seeds) in O. europaea, the contribution of the other three species being marginal (1–2% each; Fig. 1).

What is the contribution of different frugivore species to seed arrival in different microhabitats?

The seed dispersal services provided by the different bird species to the two dominant fruiting plants (O. europaea and P. lentiscus) were not evenly distributed over the different microhabitats (Fig. 2). The magnitude of bird-mediated seed rain varied between the sampled microhabitats by nearly two orders of magnitude, ranging in O. europaea from a mean of 0.06 seeds m$^{-2}$ in open areas to 13.6 seeds m$^{-2}$ beneath Olea plants, and in P. lentiscus from zero seeds m$^{-2}$ in open areas to 21.8 seeds m$^{-2}$ beneath female Pistacia plants (Fig. 2). The
relative contribution of each disperser species to seed rain in each microhabitat varied enormously, ranging from 0% (i.e. no contribution) to 67% in *P. lentiscus* dispersed by *E. rubecula* under *Quercus* trees and to 82% in *O. europaea* dispersed by *S. atricapilla* beneath *Rhamnus* and *Q. coccifera* plants (see Fig. 2). Relative contributions were calculated from a mean CSD of 12/5 (range CSD 523) 'identified' seeds per microhabitat (see Fig. 2). Such sample sizes enabled us to detect that, for example, the relative contribution of *T. philomelos* to the seed rain of *O. europaea* beneath *Olea* plants (17%, CI 95% = 1–34%, n = 23) was significantly lower than beneath *Crataegus* plants (73%, CI 95% = 41–100%, n = 11). However, the reduced sample sizes in other microhabitats (n ≤ 10) did not allow the detection of significant differences (i.e. CI95% slightly overlapped). Combining (mean) seed rain per microhabitat and the relative contribution of each disperser species in that microhabitat allowed us to calculate the magnitude of seed rain in each microhabitat per disperser species. For example, due to important differences in the magnitude of seed rain, the mean density of *Olea* seeds dispersed by song thrushes beneath *Crataegus* plants (0.83 seeds m⁻²) was lower than beneath *Rhamnus* plants (1.24 seeds m⁻²), despite a much larger relative contribution under the former plant species (73% vs. 18%, respectively).

**Do different frugivores select for different fruit/seed sizes?**

As reported above, most *Olea* seeds whose avian disperser was successfully barcoded (96%; 90 out of 94) were dispersed by song thrushes (*T. philomelos*) and blackcaps (*S. atricapilla*), the sample size in the other three disperser species being too small to test fruit size selection (Fig. 1). Song thrushes and blackcaps significantly differed in the seed sizes they dispersed (see Fig. 3) and thus in the fruit sizes they consumed. Fruits consumed by song thrushes and blackcaps had, respectively, mean diameters of 8.14 and 7.80 mm, mean lengths of 11.88 and 11.27 mm and mean weights of 0.486 and 0.398 g, as estimated from linear regressions on dry seed weight \( r^2 = 0.83–0.90; P < 0.001, n = 60 \).

**Discussion**

Sufficient amount of amplifiable DNA can be recovered from the small amounts of animal material attached to the surface...
of vertebrate-dispersed seeds (sampled in seed traps or directly from the ground). DNA barcoding allowed us to successfully identify the frugivore species that contributed each single seed dispersal event. This technique proves especially useful in situations in which (i) direct observation is complicated owing to elusive character of the species and/or to habitat structure; (ii) removal rates of fruits are extremely low; and (iii) obtaining direct observations of the disperser’s activity in specific microhabitats or landscape settings is difficult. Accordingly, this technique can provide new data on plant–frugivore interactions in mega-diverse communities in which observations of several interactions are rare, improbable and, overall, time-consuming. In addition, this technique will help assess which particular species contribute to critical dispersal events such as long-distance dispersal and dispersal in non-occupied habitats.

METHODOLOGICAL CONSIDERATIONS

Three methodological issues are important for applying successfully the described protocols to other study systems. First, recovering avian DNA from the surface of the defecated/regurgitated seeds relied on our ability to directly incubate the seed in extraction buffer, because all seed species fitted into 1.5- to 2.0-mL tubes (see Fig. S1). For larger sizes, we recommend rubbing the seed surface with a sterile swab (Ramón-Laca & Gleeson 2014) soaked in phosphate buffer. We also recommend using the cotton swab when the seed surface has an excess of faecal material and/or pulp remains. In these cases, we obtained better results, probably because avian DNA is present on the external layers of the dispersed seeds (those recovered with the swab) and also because excess material leads to inhibition by bacterial DNA and/or plant secondary metabolites (see Marrero et al. 2008, 2009; Nogales et al. 2009).

Secondly, our data suggest that rainy conditions decrease the success of DNA-specific amplification (from 97% to 81% in our samples). Therefore, frequent monitoring (from every few days up to a week) and seed collection from traps is recommended in wet or rainy conditions. On the other hand, DNA amplification was less successful (63%) in a subset of Olea seeds that were mostly sampled by direct searches (see Table S4). The time since deposition for bird-dispersed seeds collected by direct searches is unknown. The longer the time a dispersed seed is exposed to environmental conditions, the greater the possibilities of degradation of the disperser’s DNA and contamination by other DNA sources (e.g. bacteria, fungi). Hence, if seed trap results are inefficient owing to the peculiarities of the study system, we recommend sampling seeds within fixed areas – for example transects or sampling quadrats – that can be monitored regularly and frequently throughout the fruiting period.

The third consideration is the species identification in the BOLD system (Ratnasingham & Hebert 2007). All bird species included in our study (European) had several records in BOLD. However, if there is a lack of BOLD records for any species in the disperser assemblage under study, the first step must be obtaining validation samples and the second to upload the COI sequences to BOLD. We successfully identified disperser species on the basis of a high similarity (>99%) between scored sequences and those stored in BOLD, with no discrepancy even between sympatric–congereneric species (similarity ≤94%). Indeed, the sequence divergence of the second-ranked species was >5% in all cases, and the BOLD initiative establishes a 2% threshold as a cut-off between species. The single exception was the species pair Sturnus vulgaris/unicolor (similarity of 99.4–99.7%) in validation samples (but see footnote in Table 1). In such cases, direct observations of disperser activities may help to discern between species, as we did in our validation samples (see Table 1). Finally, sequencing only one strand proved very cost-effective as the resulting sequences accounted for 84–90% of our COI fragment and allowed for successful between-species discrimination.

APPLICABILITY IN FRUGIVORY AND SEED DISPERSAL STUDIES

Our three application examples show how DNA barcoding provides a non-invasive technique for quantifying frugivory and seed dispersal interaction networks (Fig. 1), assessing the contribution of each frugivore species to the seed rain in different microhabitats (Fig. 2), and testing whether different frugivore species select different fruit sizes (Fig. 3). Given that seeds are sampled at the end of the dispersal process, this technique enables linking the identity of the disperser species responsible for each dispersal event to plant traits (e.g. fruit/seed size) and environmental features (e.g. habitat/microhabitat of destination), thus linking frugivory and seed deposition patterns (see Schupp, Jordano & Gómez 2010; Cortes & Uriarte 2013). It is precisely this bridge between phases at individual seed level that opens up new research avenues that were unavailable to traditional field studies. Combined with the analysis of post-dispersal seed fate, DNA barcoding allows us to calculate the quantitative (number of dispersed seeds; Fig. 1) and qualitative (microsite quality for recruitment; Fig. 2) components
of seed dispersal effectiveness for specific frugivore species (Schupp, Jordano & Gómez 2010). Moreover, combined with the genetic identification of the source tree of each seed (Godoy & Jordano 2001; Jordano et al. 2007), the protocols described here allow a full characterization of the dispersal process of any plant species, that is, the identity of the source trees of the dispersed seeds and the frugivore species that contributed each dispersal event. Given that dispersal distances and the contribution of specific frugivore species can be determined, both methods allow for the direct estimation of the Total Dispersal Kernel (Nathan 2007), that is, the relative contribution of different frugivores in distinct spatial sectors or distances. In conclusion, DNA barcoding can be used for characterizing the functional value of specific frugivore species within diverse mutualistic assemblages.

Acknowledgements

We thank J. Luis García-Madarro, Laura Soriano and Mireia Casas-Marec for useful suggestions on DNA isolation, amplification and sequencing, and Alejando Oraauba (bird ringer). Roger Jovani, David Serrano, Manuel Vázquez, José Luis Tellas and the staff of the scientific collection in Estación Biológica de Doñana (EBD-CSIC) for allowing access to the samples for non-degraded DNA isolation. Patricia Marrero and Manuel Nogales offered valuable insight on laboratory techniques. José A. Godoy provided useful comments and technical advice during all the stages of this project. Carolina Carvalho, Eugene W. Schupp, Néstor Pérez and Pablo González helped in the field. Logistical support was provided by the Molecular Ecology Laboratory, Estación Biológica de Doñana (LEM-EBD-CSIC). Michael Lockwood corrected the English style. This study was supported by funds from a Junta de Andalucía Excellence Grant (RNM-5731), as well as a Severo Ochoa Excellence Award from the Ministerio de Economía y Competitividad (SEV-2012-0262). Two anonymous reviewers contributed with constructive suggestions that greatly improved the paper.

Data accessibility

Data deposited in the Dryad repository (González-Varo et al. 2014).

References


DNA barcoding for seed disperser identification 813.


Received 8 April 2014; accepted 29 May 2014

Handling Editor: M. Gilbert

**Supporting Information**

Additional Supporting Information may be found in the online version of this article.

Fig. S1. Seeds of the four bird-dispersed species sampled in the field.

Fig. S2. Regurgitated and defecated bird-dispersed seeds.

Table S1. Non-degraded DNA samples used for primer design.

Table S2. Characteristics of the two DNA extraction protocols tested.

Table S3. Descriptive statistics for the quantity of PCR product obtained with DNA extraction methods ‘protocol 1’ and ‘protocol 2’.

Table S4. Sample sizes and success of disperser identification for seeds from the two main fleshy-fruited species collected in different microhabitats.
Who dispersed the seeds? The use of DNA barcoding in frugivory and seeds dispersal studies

Juan P. González-Varo¹*, Juan M. Arroyo¹ and Pedro Jordano¹

¹Integrative Ecology Group, Estación Biológica de Doñana, EBD-CSIC, Avda. Americo Vespucio S/N, Isla de La Cartuja, E-41092 Sevilla, Spain

* Correspondence author: E-mail: juanpe@ebd.csic.es

Short running title: DNA barcoding for seed disperser identification

Supporting Information

Table S1. Non-degraded DNA samples used for primer design.

Table S2. Characteristics of the two DNA extraction protocols tested.

Table S3. Descriptive statistics for the quantity of PCR product obtained with DNA extraction methods ‘protocol 1’ and ‘protocol 2’.

Table S4. Sample sizes and success of disperser identification for seeds from the two main fleshy-fruited species collected in different microhabitats.

Fig. S1. Seeds of the four bird-dispersed species sampled in the field.

Fig. S2. Regurgitated and defecated bird-dispersed seeds.
Table S1. Non-degraded DNA samples of 16 bird species used for primer design. All species belong to order Passeriformes, except *Columba palumbus* (Columbiformes).

<table>
<thead>
<tr>
<th>Species name (common name)</th>
<th>Sample of non-degraded DNA</th>
<th>No. of samples</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Carduelis chloris</em> (Greenfinch)</td>
<td>Feather</td>
<td>1</td>
</tr>
<tr>
<td><em>Columba palumbus</em> (Woodpigeon)</td>
<td>Feather</td>
<td>2</td>
</tr>
<tr>
<td><em>Erithacus rubecula</em> (Robin)</td>
<td>Feather</td>
<td>5</td>
</tr>
<tr>
<td><em>Fringilla coelebs</em> (Chaffinch)</td>
<td>Feather</td>
<td>2</td>
</tr>
<tr>
<td><em>Parus major</em> (Great tit)</td>
<td>Feather</td>
<td>1</td>
</tr>
<tr>
<td><em>Passer domesticus</em> (House sparrow)</td>
<td>Blood</td>
<td>4</td>
</tr>
<tr>
<td><em>Passer hispaniolensis</em> (Spanish sparrow)</td>
<td>Blood</td>
<td>4</td>
</tr>
<tr>
<td><em>Phoenicurus ochruros</em> (Black redstart)</td>
<td>Feather</td>
<td>1</td>
</tr>
<tr>
<td><em>Saxicola rubicola</em> (European stonechat)</td>
<td>Blood</td>
<td>1</td>
</tr>
<tr>
<td><em>Sturnus unicolor</em> (Spotless starling)</td>
<td>Feather</td>
<td>1</td>
</tr>
<tr>
<td><em>Sturnus vulgaris</em> (Common starling)*</td>
<td>Muscle</td>
<td>2</td>
</tr>
<tr>
<td><em>Sylvia atricapilla</em> (Blackcap)</td>
<td>Feather</td>
<td>2</td>
</tr>
<tr>
<td><em>Sylvia melanocephala</em> (Sardinian warbler)*</td>
<td>Muscle</td>
<td>2</td>
</tr>
<tr>
<td><em>Turdus iliacus</em> (Redwing)*</td>
<td>Muscle</td>
<td>2</td>
</tr>
<tr>
<td><em>Turdus merula</em> (Blackbird)</td>
<td>Feather</td>
<td>2</td>
</tr>
<tr>
<td><em>Turdus philomelos</em> (Song thrush)</td>
<td>Feather</td>
<td>1</td>
</tr>
</tbody>
</table>

Table S2. Characteristics of the two DNA extraction protocols tested.

<table>
<thead>
<tr>
<th>Protocol components</th>
<th>Protocol 1</th>
<th>Protocol 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Extraction buffer</td>
<td>0.1 M Tris-HCl pH 8.0, 0.1 M EDTA pH 8.0, 0.01 M NaCl, 0.5% SDS and 0.25 mg/mL Proteinase K. (Longmire et al. 1997).</td>
<td>10 M GuSCN, 0.1 M Tris-HCl pH 6.4, 0.04 M EDTA pH 8.0 and 2.6% Triton X-100. (Marrero et al. 2009 cite Boom et al. 1990.)</td>
</tr>
<tr>
<td>Incubation for sample digestion</td>
<td>2.5 h</td>
<td>30 min</td>
</tr>
<tr>
<td>Binding buffer</td>
<td>5 M GuSCN, 0.1 M Tris-HCl pH 6.4, 0.02 M EDTA pH 8.0 and 1.3% Triton X-100. (Höss and Pääbo, 1993).</td>
<td>10 M GuSCN, 0.1 M Tris-HCl pH 6.4, 0.04 M EDTA pH 8.0 and 2.6% Triton X-100. (Marrero et al. 2009 cite Boom et al. 1990.)</td>
</tr>
<tr>
<td>Amount of silica</td>
<td>120 µL</td>
<td>15 µL</td>
</tr>
<tr>
<td>Incubation for DNA binding to silica</td>
<td>2.5 h</td>
<td>15 min</td>
</tr>
<tr>
<td>Washing buffer</td>
<td>50% Ethanol, 10 mM Tris-HCl pH 8.0, 1 mM EDTA pH 8.0 and 125 mM NaCl. (Rholand et al. 2010, Rholand and Hofreiter, 2007).</td>
<td>50% Ethanol, 20 mM Tris-HCl pH 8.0, 1 mM EDTA pH 8.0 and 200 mM NaCl.</td>
</tr>
<tr>
<td>Washing method</td>
<td>Column (MoBiTec, Germany, product # M1002S) with a glass microfibre filter (Whatman Grade GF/B 1.0 µm) on top of the 10 µm column filter. (Rholand et al. 2010).</td>
<td>Vortex in 1.5 mL tube.</td>
</tr>
<tr>
<td>Elution</td>
<td>First: 50 µL ultrapure water. Second: 50 µL TLE (5 mM Tris-HCl pH 8.0 and 0.1 mM EDTA pH 8.0).</td>
<td>50 µL TE (10 mM Tris-HCl pH 8.0 and 1 mM EDTA pH 8.0).</td>
</tr>
</tbody>
</table>

See section DNA ISOLATION, AMPLIFICATION AND SEQUENCING (Materials and Methods) for a detailed procedure of ‘Protocol 1’.

See Marrero et al. (2009) for a detailed description of ‘Protocol 2’.
Table S3. Descriptive statistics for the quantity of PCR product amplified from DNA extracts obtained with ‘protocol 1’ and ‘protocol 2’. PCR product was measured with a QuantiFluor™-ST Fluorometer (Promega, Madison, WI, USA) using the Quant-iT™ PicoGreen® dsDNA Assay Kit (Invitrogen™, Life Technologies), which contains a fluorescent double-stranded-DNA-binding dye that enables sensitive and specific quantification of DNA amplification products in solution.

<table>
<thead>
<tr>
<th>DNA extraction protocols</th>
<th>n (validation samples)</th>
<th>PCR product (ng/µL)</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Mean</td>
<td>SD</td>
</tr>
<tr>
<td>Protocol 1</td>
<td>10</td>
<td>48.4</td>
<td>10.3</td>
</tr>
<tr>
<td>Protocol 2</td>
<td>11</td>
<td>20.3</td>
<td>14.9</td>
</tr>
</tbody>
</table>

Mann-Whitney $U$ test for the difference in mean values between protocols: $Z = 3.5$, $P < 0.001$
Table S4. Sample sizes ($n_{seeds}$) and success (%) of disperser identification through DNA-barcoding of seeds from the two main fleshy-fruited species in the study area (*Olea europaea* and *Pistacia lentiscus*) collected in different microhabitats.

<table>
<thead>
<tr>
<th>Microhabitat</th>
<th><em>Olea europaea</em></th>
<th></th>
<th><em>Pistacia lentiscus</em></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$n_{seeds}$</td>
<td>Seed disperser</td>
<td>$n_{seeds}$</td>
<td>Seed disperser</td>
</tr>
<tr>
<td></td>
<td>sampled (*)</td>
<td>identification (%)</td>
<td>sampled (*)</td>
<td>identification (%)</td>
</tr>
<tr>
<td><strong>Quercus trees</strong></td>
<td>14 (5)</td>
<td>93</td>
<td>18 (0)</td>
<td>83</td>
</tr>
<tr>
<td>Fruit-bearing shrubs</td>
<td>49 (21)</td>
<td>82</td>
<td>62 (0)</td>
<td>90</td>
</tr>
<tr>
<td><em>Olea</em></td>
<td>28 (2)</td>
<td>81</td>
<td>21 (0)</td>
<td>95</td>
</tr>
<tr>
<td><em>Pistacia ♀</em></td>
<td>11 (9)</td>
<td>63</td>
<td>24 (0)</td>
<td>92</td>
</tr>
<tr>
<td><em>Crataegus</em></td>
<td>11 (10)</td>
<td>100</td>
<td>17 (0)</td>
<td>82</td>
</tr>
<tr>
<td>Non-fruit-bearing shrubs</td>
<td>32 (16)</td>
<td>100</td>
<td>31 (1)</td>
<td>90</td>
</tr>
<tr>
<td><em>Pistacia ♂</em></td>
<td>10 (8)</td>
<td>100</td>
<td>6 (0)</td>
<td>83</td>
</tr>
<tr>
<td><em>Q. coccifera</em></td>
<td>11 (3)</td>
<td>100</td>
<td>8 (1)</td>
<td>88</td>
</tr>
<tr>
<td><em>Rhamnus</em></td>
<td>11 (5)</td>
<td>100</td>
<td>17 (0)</td>
<td>94</td>
</tr>
<tr>
<td>Open ground</td>
<td>9 (3)</td>
<td>89</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td><strong>TOTAL</strong></td>
<td>105 (45)</td>
<td>90</td>
<td>111 (1)</td>
<td>89</td>
</tr>
</tbody>
</table>

* Numbers in parentheses denote the number of seeds collected by means of directed sampling searches aimed to increase the sample size in some microhabitats.
Figure S1. The four species of bird-dispersed seeds sampled in the field. All species fitted within 1.5–2.0 mL tubes. For sampling, individual seeds of *O. europaea* and *C. monogyna* were introduced within 2.0 mL tubes, whereas individual seeds of *P. lentiscus* and *M. communis* were introduced within 1.5 mL tubes (shown in the photograph).
Figure S2. Appearance of the bird-dispersed seeds sampled in the field (see scale in Fig. S1). Some species (*O. europaea*, *C. monogyna* and *P. lentiscus*) were sampled as individual – regurgitated/defecated – seeds with tiny amounts of pulp and/or bird-originated material on their surface. Other species (*P. lentiscus* and *M. communis*) were sampled as bird faeces where the seeds were embedded.