Methods in Ecology and Evolution

British Ecological Society

Methods in Ecology and Evolution 2014, **5,** 806–814 doi: 10.1111/2041-210X.12212

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. Americo 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 & Grivet 2011; Côrtes & 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.*

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

2007). Yet specific identification is a crucial aspect when determining the particular role (e.g. redundancy vs. complementarity) of different frugivore species in the seed dispersal services provided by the whole disperser assemblage (Jordano *et al.* 2007; González-Varo, López-Bao & Guitián 2013). This information is essential for understanding the role of multiple mutualists in plant regeneration (Schupp, Jordano & Gómez 2010), ecosystem functioning (Lundberg & Moberg 2003) and, ultimately, biodiversity conservation (Trakhtenbrot *et al.* 2005; Montoya *et al.* 2008).

Developing non-invasive methods that allow for robust identification of the frugivore species contributing dispersed seeds collected in the field would open new research avenues for frugivory and seed dispersal studies, thereby allowing the link between the removal/departure (frugivory) and arrival (seed deposition) stages of animal-mediated dispersal to be made (Schupp, Jordano & Gómez 2010; Côrtes & Uriarte 2013). Different frugivore species may disperse seeds of different sizes (Rey *et al.* 1997), and seed size has strong effects on the early stages of plant recruitment (Alcántara & Rey 2003). Concomitantly, differences in the quality for recruitment of the microsites in which different frugivore species disperse seeds can compensate for between-species variation in the magnitude of frugivory interactions (Schupp 1995). This link between frugivory and seed deposition patterns is particularly necessary considering that almost all existing knowledge of seed dispersal networks is based on – and therefore biased towards – frugivory interactions detected by observations of animal visits to fruiting plants (e.g. Schleuning *et al.* 2014) or faecal analyses of mist-netted birds (Jordano 1988; Heleno *et al.* 2013).

Here, we develop a protocol for the application of DNA-barcoding techniques for identifying frugivore species from animal matter present on the surface of dispersed seeds sampled in the field. DNA barcoding has been widely used for biodiversity and animal diet studies (reviewed in Valentini, Pompanon & Taberlet 2009). However, its application when resolving ecological problems such as the functional features of different partner species within complex interaction networks is still emerging (see Jurado-Rivera *et al.* 2009 for a pioneer application for host–herbivore interactions). Marrero *et al.* (2009) pioneered the use of amplified DNA sequences to identify two pigeon species using faeces and dispersed seeds with the aim of characterizing habitat segregation (Nogales *et al.* 2009). Avian DNA can be successfully extracted from the minimal amount of material that is present on the surface of regurgitated or defecated seeds. Seeds can be sampled in seed traps or directly from the ground, as is habitual in seed rain studies (Jordano *et al.* 2007; García, Zamora & Amico 2010).

Avian species identification by DNA barcoding was based on a mitochondrial DNA region of the cytochrome c oxidase subunit I (COI; see Ratnasingham & Hebert 2007). In order to illustrate how DNA barcoding can be applied to unambiguously answer major questions in frugivory and seed dispersal, we conducted pilot surveys in Mediterranean woodland vegetation in SW Spain to sample bird-dispersed seeds. Then, we 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 Sequencher 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′-GCATGAGCCGGAATAGTRGG-3′; COI-fsdR: 5′-TGTGAKAGGGCAGGTGGTTT-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 Sequencher 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 (Cádiz 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 & Pääbo (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°397′N, 5°56′8W). Vegetation consists of large holm- (*Quercus ilex* subsp. *ballota*) and cork- (*Q. suber*) oaks, 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 the lack of shelters for rodents (Fedriani & 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.

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-ringing sessions, except those from *Sturnus unicolor*, which were obtained beneath resting sites of winter flocks. *Pl*: *Pistacia lentiscus* seeds; *Oe*: *Olea europaea* seeds

Bird species (*n* individuals)	*n* total samples	*n* faecal samples	*n* defecated or regurgitated seeds	Similarity (%)
Carduelis chloris	1	1	–	100
Fringilla coelebs	1	1	–	100
Erithacus rubecula	1	1	–	100
Parus major	1	1	–	100
Saxicola rubicola	1	1	–	99·4
*Sturnus unicolor**	17	–	17 (2 *Pl*, 15 *Oe*)	99·4–100
Sylvia atricapilla	14	1	13 (7 *Pl*, 6 *Oe*)	99·7–100
Sylvia melanocephala	3	–	3 (3 *Pl*)	100
Turdus philomelos	1	–	1 (1 *Oe*)	100
Total (9 species)	40	6	34 (12 *Pl*, 22 *Oe*)	99·4–100

*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.

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 & Pääbo 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-HCl 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

(50% Ethanol, 10 mM Tris-HCl pH 8·0, 1 mM EDTA pH 8·0 and 125 mM NaCl). Columns were placed in new tubes and DNA was eluted twice, first with 50 μL of ultrapure water and then with 50 μL of diluted TE buffer (5 mM Tris-HCl pH 8 and 0·1 mM EDTA pH 8·0).

PCR amplifications were performed by increasing the concentration of primers and Taq to overcome the expected low avian DNA amount in samples and by increasing the concentration of bovine serum albumin (BSA) to overcome the possible PCR inhibitors. The final 30-μL volume of the PCR cocktail contained 3·0 μL (1 ×) buffer (67 mM Tris-HCL pH 8·8, 16 mM $(NH_4)_2SO_4$, 10 mM KCl, 0·01% stabilizer), 1·2 μL (2·0 mM) $MgCl_2$, 0·75 μL (0·5 mg mL^{-1}) BSA (Roche Diagnostics, Barcelona, Spain), 0·3 μL (0·25 mM) dNTP, 1·8 μL (0·60 μM) × 2 primers (COI-fsdF and COI-fsdR; see above), 0·2 μL (1·0 U) Taq DNA polymerase (Bioline, London, UK), 12·95 μL ultrapure water and 8 μL of the DNA extract (mean ± SD: 8·4 ± 5·9 ng μL^{-1} of total DNA, $n = 39$ samples; quantified with NanoDrop® ND-1000 Spectrophotometer, NanoDrop Technologies Inc., Wilmington, DE, USA). Reactions were undertaken in a Bio-Rad DNA Engine® Peltier Thermal Cycler with an initial 4 min of denaturation at 94°C; 42 cycles at 94°C for 45 s, annealing at 54°C for 45 s and extension at 72°C for 45 s; and final extension of 6 min at 72°C. After verifying successful amplification by agarose gel electrophoresis, excess primers and dNTPs were removed using enzymatic reaction of Antarctic phosphatase buffer, Antarctic phosphatase and *Escherichia coli* exonuclease I (all New England Biolabs, Hitchin, UK). We only sequenced one strand (forward primer) of the amplified COI fragments because the electrophoretic patterns were clear (trimming initial 5′ region for low quality) and resulting sequences (390- to 420-bp length; average quality >90% in Sequencher) allowed successful discrimination between species (see Results). Sequencing reaction was carried out using the BigDye® Terminator v1.1 cycle sequencing kit (Applied Biosystems, Foster City, CA, USA) according to the manufacturer's instructions, and labelled fragments were cleaned on SephadexTM G-50 (GE Healthcare, Little Chalfont, UK) plates before electrophoresis in an ABI 3130xl Genetic Analyzer (Applied Biosystems). DNA fragments were aligned and edited using Sequencher 4.9, and the obtained sequences were identified using the BOLD identification system (http://www.boldsystems.org/index.php/IDS_OpenIdEngine).

DATA ANALYSES: APPLICATION EXAMPLES

We used our data set consisting of 'individual seeds whose disperser species was DNA-identified' to illustrate how DNA-barcoding techniques can be applied to unambiguously answer major questions in frugivory and animal-mediated seed dispersal.

Example 1. Which frugivores disperse the seeds of which plant species?

We used our data set to draw a weighted interaction network between frugivorous seed dispersers and fruiting plants which represents the identity and strength of interactions between partner species (i.e. the importance of fruit species for birds and the quantity of dispersal services provided by birds for plants). We used the R package 'bipartite' version 2.03 (Dormann, Fründ & Gruber 2014) to plot this network.

Example 2. What is the contribution of different frugivore species to seed arrival in different microhabitats?

We used data from the two dominant fleshy-fruited species in the study site (*O. europaea* and *P. lentiscus*) to explore the patterns of seed deposition by different frugivore species over different microhabitats. 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/seed 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 ($r^2 = 0{\cdot}83$–$0{\cdot}90$; $P < 0{\cdot}001$, $n = 60$). 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: $Z = 3{\cdot}5$, $P < 0{\cdot}001$; 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 99·4–100% of sequence similarity (see results in Table 1). With the exception of *Sturnus unicolor* (see Table 1), the second species ranked by BOLD had a similarity of 89–94% with the scored sequences.

DISPERSER IDENTIFICATION IN BIRD-DISPERSED SEEDS

During the field-sampling operation, we collected a total of 221 seeds belonging to four fleshy-fruited species: 111 of *Pistacia lentiscus*, 105 of *Olea europaea*, four of *Crataegus monogyna* and one of *Myrtus communis* (Fig. S2). Five frugivorous bird species were successfully identified through DNA barcoding as 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. 99·4–100%). 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

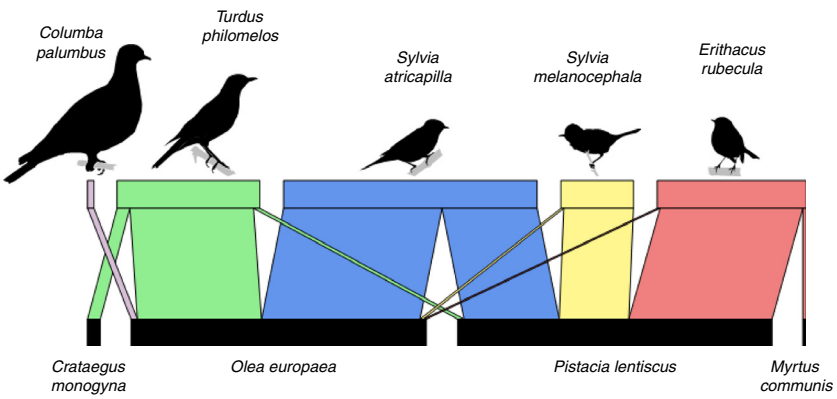

Fig. 1. Empirical interaction network of avian frugivores and fleshy-fruited plants determined by DNA-barcoding identification of frugivore species in dispersed seeds. Seeds were sampled from seed traps, transects and intensive searches, while avian DNA was extracted from remains present on the surface of dispersed seeds. Horizontal width of the nodes is proportional to the frequency of each species sampled in each trophic level. Horizontal width of the links between species is proportional to the frequency of seeds assigned to each avian-frugivore species.

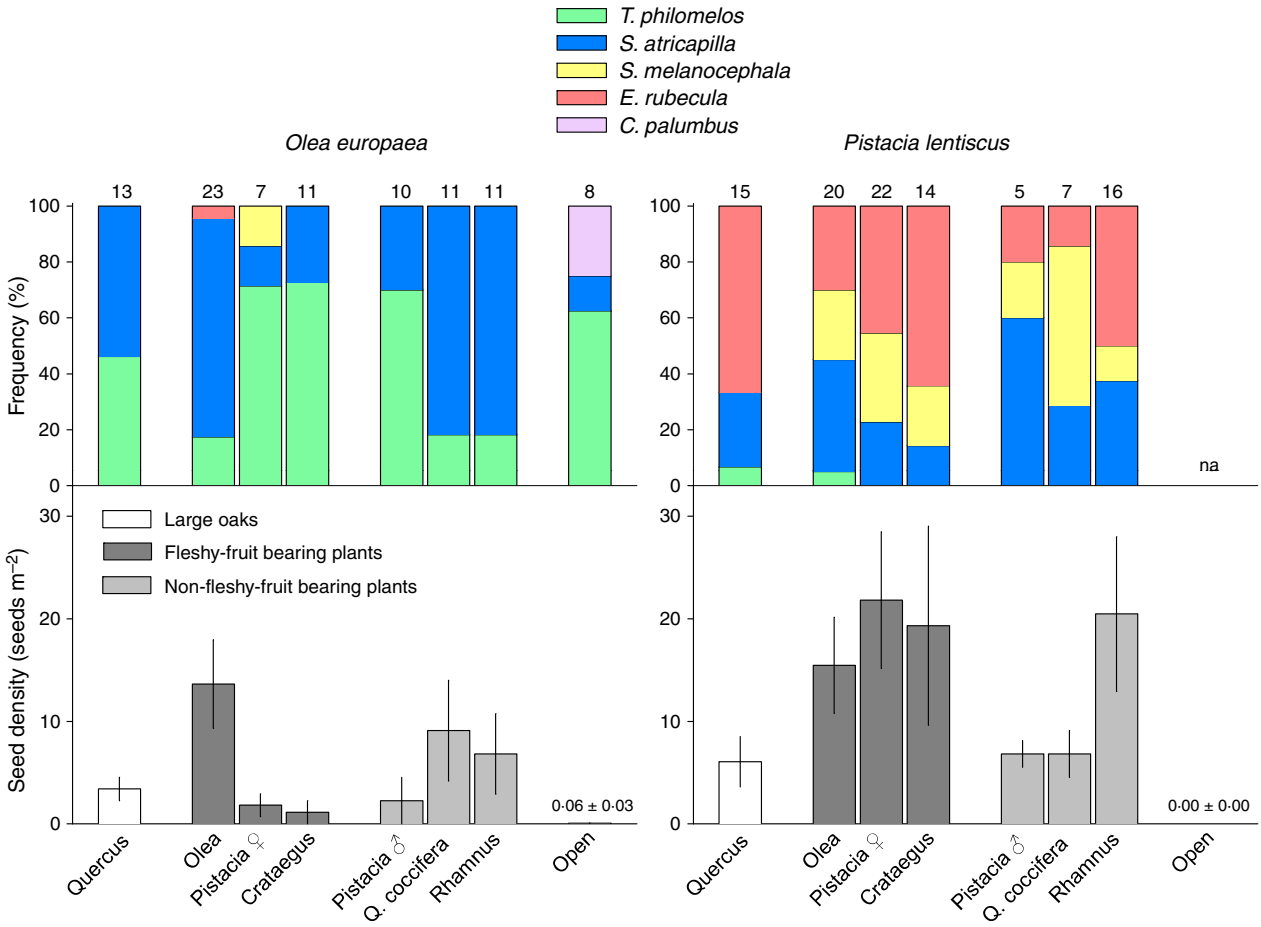

Fig. 2. Relative contribution (%) of different avian frugivores to seed rain of *O. europaea* and *P. lentiscus* in different microhabitats (top panels; numbers above bars denote sample sizes). Bird-dispersed seed density (bottom panels; mean ± SE) in the eight sampled microhabitats belonging to four microhabitat types: beneath *Quercus* trees (*Q. ilex*, *Q. suber*), beneath fleshy fruit-bearing plants (*O. europaea*, *P. lentiscus* female shrubs and *Crataegus monogyna*), beneath non-fleshy fruit-bearing plants (*P. lentiscus* male shrubs, *Q. coccifera* and *Rhamnus alaternus* – a summer fruiting species) and in open ground sites.

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 ± sd of 12·9 ± 5·5 (range 5–23) '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 ($n \leq 10$) did not allow the detection of significant differences (i.e. $CI_{95\%}$ 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

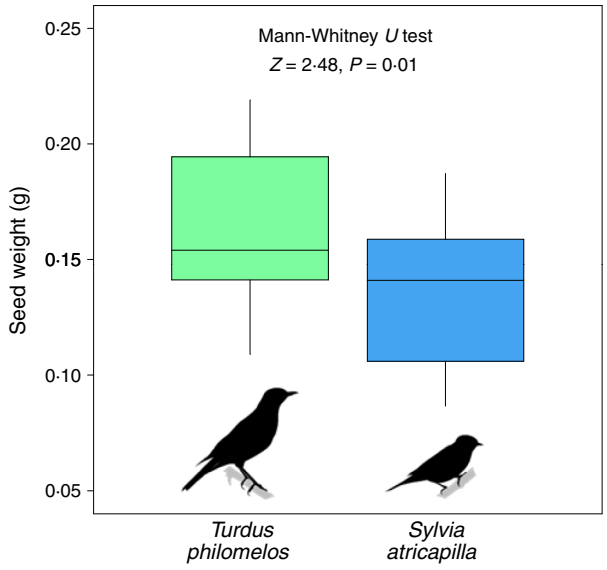

Fig. 3. Dry weight (line: median; box: quartiles; whisker: 90th percentile) of *Olea europaea* var. *sylvestris* seeds dispersed by blackcaps (*Sylvia atricapilla*) or song thrushes (*Turdus philomelos*).

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–congeneric 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; Côrtes & 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-Mudarra, Laura Soriano and Mireia Casas-Marce for useful suggestions on DNA isolation, amplification and sequencing, and Alejandro Onrubia (bird ringer), Roger Jovani, David Serrano, Manuel Vázquez, José Luis Tella 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

Alcántara, J.M. & Rey, P.J. (2003) Conflicting selection pressures on seed size: evolutionary ecology of fruit size in a bird-dispersed tree, *Olea europaea*. *Journal of Evolutionary Biology*, **16**, 1168–1176.

Ashley, M.V. (2010) Plant parentage, pollination, and dispersal: how DNA microsatellites have altered the landscape. *Critical Reviews in Plant Sciences*, **29**, 148–161.

Carlo, T.A. & Yang, S. (2011) Network models of frugivory and seed dispersal: Challenges and opportunities. *Acta Oecologica*, **37**, 619–624.

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**, 301–307.

Côrtes, M.C. & Uriarte, M. (2013) Integrating frugivory and animal movement: a review of the evidence and implications for scaling seed dispersal. *Biological Reviews*, **88**, 255–272.

Dormann, C.F., Fründ, J. & Gruber, B. (2014) *Package 'bipartite': visualising bipartite networks and calculating some (ecological) indices*. R package, version 2.04. See http://cran.r-project.org/web/packages/bipartite/.

Fedriani, J.M. & Manzaneda, A.J. (2005) Pre- and post- dispersal seed predation by rodents: balance of food and safety. *Behavioural Ecology*, **16**, 1018–1024.

Fleming, T.H. & Kress, W.J. (2013) *The Ornaments of Life*. University of Chicago Press, Chicago.

García, C. & Grivet, D. (2011) Molecular insights into seed dispersal mutualisms driving plant population recruitment. *Acta Oecologica*, **37**, 632–640.

García, D., Zamora, R. & Amico, G.C. (2010) Birds as suppliers of seed dispersal in temperate ecosystems: conservation guidelines from real-world landscapes. *Conservation Biology*, **24**, 1070–1079.

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

Gómez, J.M. (2003) Spatial patterns in long-distance dispersal of *Quercus ilex* acorns by jays in a heterogeneous landscape. *Ecography*, **26**, 573–584.

González-Varo, J.P., Arroyo, J.M. & Jordano, P. (2014) Data from: who dispersed the seeds? The use of DNA barcoding in frugivory and seed dispersal studies. *Dryad Digital Repository*, doi:10.5061/dryad.qh3nq.

González-Varo, J.P., López-Bao, J.V. & Guitián, J. (2013) Functional diversity among seed dispersal kernels generated by carnivorous mammals. *Journal of Animal Ecology*, **82**, 562–571.

Heleno, R.H., Olesen, J.M., Nogales, M., Vargas, P. & Traveset, A. (2013) Seed dispersal networks in the Galápagos and the consequences of alien plant invasions. *Proceedings of the Royal Society B: Biological Sciences*, **280**, 20122112.

Herrera, C.M. (1984) A study of avian frugivores, bird-dispersed plants, and their interaction in Mediterranean scrublands. *Ecological Monographs*, **54**, 1–23.

Höss, M. & Pääbo, S. (1993) DNA extraction from Pleistocene bones by a silica--based purification method. *Nucleic Acids Research*, **21**, 3913–3914.

Jordano, P. (1988) Diet, fruit choice and variation in body condition of frugivorous warblers in Mediterranean scrubland. *Ardea*, **76**, 193–209.

Jordano, P. (2014) Fruits and frugivory. *Seeds: The Ecology of Regeneration of Plant Communities*, 3rd edn (ed R.S. Gallagher), pp. 18–61. CABI, Wallingford, UK.

Jordano, P. & Schupp, E.W. (2000) Seed disperser effectiveness: the quantity component and patterns of seed rain for *Prunus mahaleb*. *Ecological Monographs*, **70**, 591–615.

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 USA*, **104**, 3278–3282.

Jurado-Rivera, J.A., Vogler, A.P., Reid, C.A.M., Petitpierre, E. & Gómez-Zurita, J. (2009) DNA barcoding insect–host plant associations. *Proceedings of the Royal Society B: Biological Sciences*, **276**, 639–648.

Lijtmaer, D.A., Kerr, K.C.R., Stoeckle, M.Y. & Tubaro, P.L. (2012) DNA barcoding birds: from field collection to data analysis. *DNA Barcodes: Methods and Protocols, Methods in Molecular Biology, Vol. 858* (eds W.J. Kress & D.L. Erickson), 158–171, Human Press, Berlin.

Lohman, D.J., Prawiradilaga, D.M. & Meier, R. (2009) Improved COI barcoding primers for Southeast Asian perching birds (Aves: Passeriformes). *Molecular Ecology Resources*, **9**, 37–40.

Longmire, J.L., Maltbie, M. & Baker, R.J. (1997) Use of "lysis buffer" in DNA isolation and its implications for museum collections. *Occasional Papers, The Museum of Texas Tech University*, **163**, 1–3.

Lovette, I.J., McCleery, B.V., Talaba, A.L. & Rubenstein, D.R. (2008) A complete species-level molecular phylogeny for the "Eurasian" starlings (Sturnidae: *Sturnus*, *Acridotheres*, and allies): recent diversification in a highly social and dispersive avian group. *Molecular Phylogenetics and Evolution*, **47**, 251–260.

Lundberg, J. & Moberg, F. (2003) Mobile link organisms and ecosystem functioning: implications for ecosystem resilience and management. *Ecosystems*, **6**, 87–98.

Mack, A.L. (1995) Distance and non-randomness of seed dispersal by the dwarf cassowary *Casuarius benneti*. *Ecography*, **18**, 286–295.

Marrero, P., Cabrera, V.M., Padilla, D.P. & Nogales, M. (2008) Molecular identification of two threatened pigeon species (Columbidae) using faecal samples. *Ibis*, **150**, 820–823.

Marrero, P., Fregel, R., Cabrera, V.M. & Nogales, M. (2009) Extraction of high-quality host DNA from feces and regurgitated seeds: a useful tool for vertebrate ecological studies. *Biological Research*, **42**, 147–151.

Montoya, D., Zavala, M.A., Rodríguez, M.A. & Purves, D.W. (2008) Animal versus wind dispersal and the robustness of tree species to deforestation. *Science*, **320**, 1502–1504.

Nathan, R. (2007) Total dispersal kernels and the evaluation of diversity and similarity in complex dispersal systems. *Seed Dispersal: Theory and its Application in A Changing World* (eds A.J. Dennis, E.W. Schupp, R.J. Green & D.A. Westcott), pp. 252–276. CAB International, Wallingford, UK.

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

Nogales, M., Marrero, P., Padilla, D.P. & Medina, F.M. (2009) The study of habitat use by censuses and molecular methods in birds: the case of two sympatric pigeons. *Bird Study*, **56**, 414–418.

Ramón-Laca, A. & Gleeson, D. (2014) Use of forensic genetics to detect a potential incursion of the brushtail possum onto Great Barrier Island. *New Zealand Journal of Ecology*, **38**, 152–155.
Ratnasingham, S. & Hebert, P.D.N. (2007) BOLD: The Barcode of Life Data System (http://www.barcodinglife.org). *Molecular Ecology Notes*, **7**, 355–364.
Rey, P., Gutiérrez, J., Alcántara, J. & Valera, F. (1997) Fruit size in wild olives: implications for avian seed dispersal. *Functional Ecology*, **11**, 611–618.
Rohland, N., Siedel, H. & Hofreiter, M. (2010) A rapid column-based ancient DNA extraction method for increased sample throughput. *Molecular Ecology Resources*, **10**, 677–683.
Schleuning, M., Böhning-Gaese, K., Dehling, D.M. & Burns, K.C. (2014) At a loss for birds: insularity increases asymmetry in seed-dispersal networks. *Global Ecology and Biogeography*, **23**, 385–394.
Schupp, E. (1995) Seed seedling conflicts, habitat choice, and patterns of plant recruitment. *American Journal of Botany*, **82**, 399–409.
Schupp, E.W., Jordano, P. & Gómez, J.M. (2010) Seed dispersal effectiveness revisited: a conceptual review. *New Phytologist*, **188**, 333–353.
Trakhtenbrot, A., Nathan, R., Perry, G. & Richardson, D.M. (2005) The importance of long-distance dispersal in biodiversity conservation. *Diversity and Distributions*, **11**, 173–181.
Valentini, A., Pompanon, F. & Taberlet, P. (2009) DNA barcoding for ecologists. *Trends in Ecology & Evolution*, **24**, 110–117.

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-Varo[1*], Juan M. Arroyo[1] and Pedro Jordano[1]

[1] *Integrative 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).

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

* Codes of specimens belonging to the scientific collection of the Estación Biológica de Doñana (EBD-CSIC): common starling (1900-200-008 and 2007-013-004), Sardinian warbler (2001-101-001 and 1000-001-325) and redwing (1900-150-103 and 1900-150-148).

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

Protocol components	Protocol 1	Protocol 2
Extraction buffer	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).	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.)
Incubation for sample digestion	2.5 h	30 min
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. (Höss and Pääbo, 1993).	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.)
Amount of silica	120 µL	15 µL
Incubation for DNA binding to silica	2.5 h	15 min
Washing buffer	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).	50% Ethanol, 20 mM Tris-HCl pH 8.0, 1 mM EDTA pH 8.0 and 200 mM NaCl.
Washing method	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).	Vortex in 1.5 mL tube.
Elution	First: 50 µL ultrapure water. Second: 50 µL TLE (5 mM Tris-HCl pH 8.0 and 0.1 mM EDTA pH 8.0).	50 µL TE (10 mM Tris-HCl pH 8.0 and 1 mM EDTA pH 8.0).

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.

DNA extraction protocols	*n* (validation samples)	PCR product (ng/µL)		
		Mean	SD	CV (%)
Protocol 1	10	48.4	10.3	21.3
Protocol 2	11	20.3	14.9	73.4

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.

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

* 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).

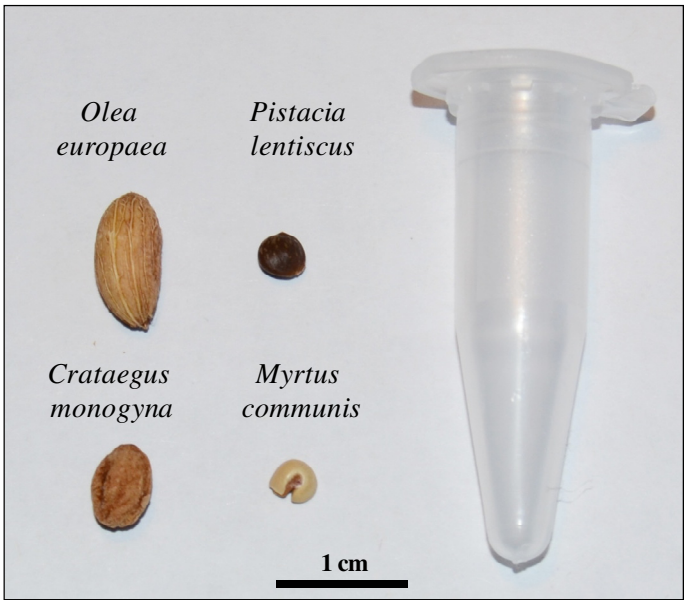

Figure S2. Appearance of the bird-dispersed seeds sampled in the field (see scale in Fig. S1). Some species (*O. europaea*, *C. monogyna* and *P. lentiscus*) were sampled as individual – regurgitated/defecated – seeds with tiny amounts of pulp and/or bird-originated material on their surface. Other species (*P. lentiscus* and *M. communis*) were sampled as bird faeces where the seeds were embedded.