

RESEARCH ARTICLE

Testing the effect of pollen exine rupture on metabarcoding with Illumina sequencing

Stephanie J. Swenson ^{*}, Birgit Gemeinholzer

Systematic Botany, Justus Liebig University Gießen, Giessen, Germany

^{*} stephanie.swenson-friedrich@bot1.bio.uni-giessen.de

Abstract

Pollen metabarcoding has received much attention recently for its potential to increase taxonomic resolution of the identifications of pollen grains necessary for various public health, ecological and environmental inquiry. However, methodologies implemented are widely varied across studies confounding comparisons and casting uncertainty on the reliability of results. In this study, we investigated part of the methodology, the effects of level of exine rupture and lysis incubation time, on the performance of DNA extraction and Illumina sequencing. We examined 15 species of plants from 12 families with pollen that varies in size, shape, and aperture number to evaluate effort necessary for exine rupture. Then created mock communities of 14 of the species from DNA extractions at 4 levels of exine rupture (0, 33, 67, and 100%) and two levels of increased lysis incubation time without exine rupture (2 or 24 hours). Quantities of these DNA extractions displayed a positive correlation between increased rupture and DNA yield, however increasing time of lysis incubation was associated with decreased DNA yield. Illumina sequencing was performed with these artificial community treatments with three common plant DNA barcode regions (*rbcl*, ITS1, ITS2) with two different primer pairings for ITS2 and *rbcl*. We found decreased performance in treatments with 0% or 100% exine rupture compared to 33% and 67% rupture, based on deviation from expected proportions and species retrieval, and increased lysis incubation was found to be detrimental to results.

 OPEN ACCESS

Citation: Swenson SJ, Gemeinholzer B (2021) Testing the effect of pollen exine rupture on metabarcoding with Illumina sequencing. PLoS ONE 16(2): e0245611. <https://doi.org/10.1371/journal.pone.0245611>

Editor: Berthold Heinze, Austrian Federal Research Centre for Forests BFW, AUSTRIA

Received: May 4, 2020

Accepted: January 4, 2021

Published: February 2, 2021

Copyright: © 2021 Swenson, Gemeinholzer. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All relevant data are within the manuscript and raw sequence data are accessible at NCBI Sequence Repository Archive (SRA) using the following accession numbers: SubmissionID: SUB8874015 BioProject ID: PRJNA690947 Title: Testing the effect of pollen exine rupture on metabarcoding with Illumina sequencing.

Funding: This research was funded by the Federal Ministry of Education and Research Germany, BMBF, FKZ: 01L11501H as part of The German Barcode of Life (GBOL) Phase II projects, received

Introduction

Accurate identification of the plant species present in mixed environmental pollen samples has important applications across several diverse disciplines including pollination ecology, forensics, paleobotany, and airborne allergen monitoring. Traditional pollen identification based on external morphology of the pollen grain is time consuming and requires a high level of training, and for many taxa, identification below the family level is not possible. Due to these limitations, metabarcoding has become an exciting alternative technique with the potential to provide faster identification of pollen, often to the species level. Despite this potential, metabarcoding in general and pollen metabarcoding in particular are not without complications. These complications must be accounted for, or overcome, before this technique can be reliably applied to large-scale biomonitoring or other programs that could affect conservation management, public health or government policy.

by BG. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: The authors have declared that no competing interests exist.

All metabarcoding programs consist of collection of environmental DNA samples, DNA extraction, amplification and sequencing. The methods chosen for each of these steps are most often dictated by user preference, cost, equipment available, and current level and/or availability of technology rather than the most rigorously tested and optimal protocol. Each one of these steps has the potential to skew results or introduce biases that affect downstream processes [1], it is therefore important for the development of metabarcoding programs, to perform comparisons of several methods in order to account for inconsistencies and provide quality controls of future investigations. The most critical aspect of improving metabarcoding performance is therefore the optimization of the initial steps. Each method of collection of environmental pollen samples inherently results in unintended capture of non-pollen plant material or non-plant taxa (e.g. fungal spores or insect fragments) and/or uses liquids, chemicals, or adhesives that can inhibit DNA extraction and amplification. There are very few ways in which pollen collection methods can be improved upon; therefore, optimizing the extraction method to overcome these potential inhibitors is the logical starting point for improvement and standardization of pollen metabarcoding.

The initial step of extraction of pollen DNA from non-germinated pollen is to break the strong outer wall of the pollen grain, called the exine, either by chemical or mechanical methods. This exine is composed of a biopolymer known as sporopollenin that is highly resistant to chemical, physical, and enzymatic degradation [2]. Breakage of the exine with the underlying cell walls is believed to be necessary to free the genetic material from these protective outer layers [3].

Biological samples that are heavily encased and not effectively disrupted by chemical lysis alone can undergo mechanical homogenization prior to DNA extraction in order to increase yield [4]. This is commonly performed by bead milling, which has maximum efficiency when the size and composition of the material being ground is matched to the container and grinding material implemented [4]. Undoubtedly, efficiency of exine rupture will be in part based on size of the individual pollen grains and the size and quantity of beads used; however, the strength and elasticity of the pollen grains' exine is highly variable. This variable strength is due to many factors beyond the sporopollenin composition, such as, the distribution of structural materials into smaller subunits, extra deposition of sporopollenin at points of greater stress, creation of striated patterns that distribute force, and the geometric arrangement of the apertures [5]. Pollen apertures are areas of reduced exine deposition, the patterns of which are species specific and can vary in size, shape, number and margination. They are the usual site of pollen tube development and likely have an important role in interaction between a pollen grain and the environment [2].

While disruption of the pollen exine, gametophytic cell walls, and nuclear membranes are necessary for efficient DNA extraction, the method implemented must also avoid degrading the DNA in the process. Simel et al. [3] evaluated several chemical and mechanical methods to achieve this and determined bead milling, to be the most efficient method to rupture the exine while maintaining abundant non-degraded DNA. This study examined each species individually however, and there is little information available concerning how to best perform this on the mixed species samples examined in metabarcoding. Due to the differential resiliency of the species' exines in mixed species samples, any method of exine maceration could result in some species' genetic material being retained in the pollen grain while other species' DNA would be over processed and sheared. This differential processing has the potential to greatly skew the diversity of species returned from sequencing.

Likely due to the results of Simel et al. [3], the majority of pollen metabarcoding studies have utilized microbead maceration as the method of exine rupture. However, there is little consistency among studies in the size, material or quantity of beads utilized, or in the time and

frequency of agitation employed (Table 1). In addition, there has been little examination into the effect this step has on quality and quantity of DNA produced, and the information that does exist is contradictory. Pornon et al. [6] saw no effect of bead milling on quantity or quality of DNA, while Prosser and Hebert [7] found unground samples or those ground with 3 mm tungsten beads produced only 1% of the DNA concentration of those ground manually with 100 mg of 0.3 mm sterile sand, and Leontidou et al. [8] showed a significant increase in yield ($p < 0.001$) with steel 5 mm beads over glass 0.5 mm beads. To date, no study has used mock communities to examine the effect of exine rupture methodology on species retrieval from sequencing.

In addition to mechanical disruption of the pollen exine, duration of incubation in lysis buffer in the DNA extraction protocol could also influence final DNA yield. It is possible this step in DNA extraction could greatly influence DNA yield from thin-walled pollen grain or those with multiple apertures that are potentially more susceptible to lysis buffer than thick walled or uniporate species.

In this study, we examined the effects of level of exine rupture and lysis incubation time on the performance of DNA extraction and Illumina sequencing. We examined 15 species from 12 plant families with pollen that varies in size, shape, and aperture number and employed different bead mill bead sizes and durations of milling to find optimal rupture conditions for each species. DNA of 14 of the species from this rupture experiment were then extracted at four levels of exine rupture (0, 33, 67, and 100%) and two levels of longer duration of lysis incubation without exine rupture (2 or 24 hours). We measured quantity of these DNA extractions to test whether a positive correlation between exine rupture levels and quantity of DNA exists for these species. From these extractions, we performed Illumina MiSeq sequencing with three common plant DNA barcode regions (*rbcL*, ITS1, ITS2) as well as two different primer pairings for ITS2 and *rbcL*, in order to examine 1) the relationship between increased levels of exine rupture and lysis time on proportions of reads produced and 2) to evaluate the performance of common plant barcodes and primer pairings with a mock community containing equal quantities of pollen grains as starting material.

Materials methods

Exine rupture quantification

Fifteen species of pollen representing 12 families varying in size, shape, and aperture number were purchased (Bonapol a. s. Czech Republic) or collected by hand (Table 2). Species included

Table 1. Methods of exine homogenization implemented in previous pollen studies.

Method of homogenization	Bead size and duration of processing	Reference
Homogenization method not mentioned		[9–15]
Homogenization method other than bead milling		[16, 17]
Bead milling performed as homogenization method	Bead size and duration of processing not specified	[18–22]
	Beads 3–3.2 mm beads, 3–4 minutes processing in a bead mill	[23–28]
	Beads ≤ 1.0 mm, two minutes processing in a bead mill	[29–32]
	Beads ≤ 1.0 mm, 5–10 mins of manual processing	[33, 34]
	Beads a mix of sizes, two minutes processing in a bead mill	[8, 35, 36]
Bead milling performed as homogenization method, and bead size and/or duration tested		[6–8]

<https://doi.org/10.1371/journal.pone.0245611.t001>

Table 2. Morphology and collection information for species used in this study and optimal bead mill conditions for near complete exine rupture.

Species	Family	Species code	Shape	Size Category	Number of Apertures	Collection source	Optimal bead size (mm)	Time to 95% rupture (seconds)
<i>Abies concolor</i>	Pinaceae	AC	Saccate	Large (51–100 µm)	1	Bonapol	1.4	112
<i>Alnus glutinosa</i>	Betulaceae	AG	Oblate	Medium (26–50 µm)	5	Bonapol	2.8	605
<i>Brassica napus</i>	Brassicaceae	BN	Prolate	Medium (26–50 µm)	3	Bonapol	1.4	249
<i>Chenopodium album</i>	Amaranthaceae	CA	Spheroidal	Medium (26–50 µm)	25–70+	Bonapol	2.8	493
<i>Cupressus sempervirens</i>	Cupressaceae	CS	Spheroidal	Medium (26–50 µm)	0	Bonapol	1.4	231
<i>Dactylis glomerata</i>	Poaceae	DG	Spheroidal	Medium (26–50 µm)	1	Hand collected, Giessen, Germany	1.4	325
<i>Picea sp.</i>	Pinaceae	Pic	Saccate	Large (51–100 µm)	1	Hand collected, Giessen, Germany	2.8	211
<i>Plantago lanceolata</i>	Plantaginaceae	PL	Spheroidal	Small (10–25 µm)	9–15	Hand collected, Aschaffenburg, Germany	2.8	502
<i>Populus sp.</i>	Salicaceae	Pop	Spheroidal	Medium (26–50 µm)	0	Hand collected, Aschaffenburg, Germany	1.4	66
<i>Rumex acetosella</i>	Polygonaceae	RA	Spheroidal	Small (10–25 µm)	3	Bonapol	1.4	123
<i>Salix caprea</i>	Salicaceae	SC	Spheroidal	Small (10–25 µm)	3	Bonapol	1.4	482
<i>Taxus baccata</i>	Taxaceae	TB	Elliptic	Small (10–25 µm)	0	Hand collected, Rhodes, Greece	1.4	69
<i>Tilia platyphyllos</i>	Tiliaceae	TP	Oblate	Small (10–25 µm)	3	Bonapol	5.0	323
<i>Ulmus glabra</i>	Ulmaceae	UG	Spheroidal	Medium (26–50 µm)	5–6	Bonapol	2.8	660
<i>Zea mays</i>	Poaceae	ZM	Spheroidal	Large (51–100 µm)	1	Bonapol	5.0	114

<https://doi.org/10.1371/journal.pone.0245611.t002>

represent the highest taxonomic and morphological breadth that could be obtained in high quantities without contamination from other species or non-pollen plant material. Four replicates of each species were made by adding approximately 1 mg of pollen to 1 mL 50% glycerol solution [37] in a 2 mL SafeSeal microcentrifuge tube (Sarstedt AG & Co. KG).

Four bead size treatments were created by addition of 1) 1 x 5 mm steel bead, 2) 3 x 2.8 mm ceramic beads, 3) 1 g 1.4 mm ceramic beads, or 4) 0.5 g 0.5 mm glass beads to the 1 mg/mL sample. These sizes and quantities were chosen as they generally correspond to sizes used in previous publications (Table 1).

Intact pollen grains were counted in 20 microscope fields by transferring 24 µL of the pollen solution (2 x 12 µL) to a hemocytometer after 6 durations in a Retsch MM400 bead mill (0, 30, 60, 120, 300, and 600 seconds) or until no intact pollen grains could be found. From these time interval counts, time to 33%, 67%, 95%, and 100% rupture were calculated and best bead size was determined to be the size that resulted in the shortest duration to 95% rupture.

DNA extraction

Fourteen of the above species (*Cupressus sempervirens* L. was eliminated due to presence of non-pollen plant material in sample) were suspended as single species in sterile 50% glycerol and 10 replicate counts of 10 µL were performed in order to create 18 replicate aliquots of

Table 3. Experimental treatment conditions.

Estimated level of exine rupture	Duration of lysis incubation	Code for each replicate
0%	1 hour	0.1, 0.2, 0.3
0%	2 hours	2h.1, 2h.2, 2h.3
0%	24 hours	24h.1, 24h.2, 24h.3
33%	1 hour	33.1, 33.2, 33.2
67%	1 hour	67.1, 67.2, 67.3
100%	1 hour	100.1, 100.2, 100.3

<https://doi.org/10.1371/journal.pone.0245611.t003>

10,000 pollen grains for each species. From these 18 replicates, 6 treatments with 3 replicates (Table 3) were created by implementing the optimal bead mill conditions determined above. DNA extraction was performed with Macherey Nagel Nucleospin food kit following the instructions for honey with 400 μ L lysis buffer and a final elution volume of 50 μ L. Mixed species DNA extractions were created by combining equal volumes of each single species extractions for a final creation of 3 replicates of each of the 6 exine rupture treatments. The DNA quantity of these mixed species extractions were quantified with Qubit™ 4 fluorometer using the dsDNA HS Assay kit.

Barcode selection

In order to separate the effect of exine rupture from PCR bias, three common plant barcodes (ITS1, ITS2, and *rbcl*) were chosen. ITS2 and *rbcl* were performed with two different primer pairings to further examine PCR bias and performance with a mock community (Table 4). The primer pairs for *rbcl* create regions that accommodate the read length limitations of Illumina sequencing but they differ in the area of coverage. The ITS2 primer pairs differ in that the combination of ITS2-2SF/ITS4 (referred to hereafter as ITS2 universal primers) is a universal pairing that will also amplify fungus, the combination of ITS-3p62plF1/ITS-4unR1 (referred to hereafter as ITS2 plant specific primers) were created to target plant taxa more specifically in order to minimize the occurrence of false negatives.

PCR and sequencing

PCR was performed in three replicates per sample with an adaptation of the Canadian Centre for DNA Barcoding Platinum® Taq PCR protocol [44], with the addition of 0.25 μ L of BSA and 1.25 μ L of 50% DMSO in a total volume of 12.5 μ L per reaction. Primers utilized are listed

Table 4. Primers used in this study.

Barcode	Primer name	Primer sequence	Reference
<i>rbcl</i>	rbcl2	TGGCAGCATTYCGAGTAACTC	[38]
<i>rbcl</i>	rbclLaR	CTTCTGCTACAAATAAGAATCGATCTC	[39]
<i>rbcl</i>	rbclLaF	ATGTCACCACAAACAGAGACTAAAGC	[40]
<i>rbcl</i>	rbclr506	AGGGGACGACCATACTTGTTC	[41]
ITS1	ITS-1unF1	GGAAGKARAAGTCGTAACAAGG	Andreas Kolter (pers. comm.)
ITS1	ITS-4unR1	GCCDAGATATCCRTTGYCRRGAG	Andreas Kolter (pers. comm.)
ITS2	ITS2-2SF	ATGCGATACTTGGTGTGAAT	[42]
ITS2	ITS4	TCCTCCGCTTATTGATATGC	[43]
ITS2	ITS-3p62plF1	ACBTRGTGTGAATTGCAGRATC	Andreas Kolter (pers. comm.)
ITS2	ITS-4unR1	TCCTCCGCTTATTKATATGC	Andreas Kolter (pers. comm.)

<https://doi.org/10.1371/journal.pone.0245611.t004>

in Table 3. PCR cycling conditions were: 95°C for 3 minutes; followed by 35 cycles of 95°C for 30 seconds, 50°C for 30 seconds, 72°C for 45 seconds; and a final extension of 72°C for 10 minutes. A PCR negative control was included for each PCR. Following PCR cycling the three replicates were combined and purified with Thermo Scientific™ Exonuclease I. The pooled replicates of non-indexed PCR products were sent to LGC Genomics GmbH where indices were added and sequencing was performed with Illumina protocol in the 2 × 300 bp format.

Raw reads of ITS2 were filtered and paired with the pipeline of Sickles et al. [20] and *rbcL* reads were filtered and paired with the pipeline of Bell et al. [21], both of these pipelines perform pairing with QIIME v.1.8.0 [45] with default parameters, and low quality reads (<Q20, <150 bp, ambiguous base pairs) are removed with USEARCH v8 [46]. Both *rbcL* and ITS2 were classified with the UTX algorithm and the database created for the respective pipelines. ITS1 reads were filtered and paired with the same parameters of ITS2 and *rbcL*, however OTU clustering was performed with USEARCH v11 [46] and OTUs were identified with the SINTAX algorithm [47] and the curated reference database PlanITS1 [48].

Data analysis

A two-way ANOVA was performed using R software [49] (R Core Team 2016) to test for association between pollen grain size and number of apertures and time to 95% rupture.

Barcode and/or primer pair performance were evaluated by the number of species from the mock community returned from sequencing runs. Correct genus assignment was accepted as indication of species presence. A species was considered present when five or more reads per species were recovered in at least two of the three replicates in one or more exine rupture treatment.

In evaluation of the effect of exine rupture treatment on read proportions returned, any presence (i.e. singletons reads and above) of the known species was counted. The resulting read proportion of a species was subtracted from the expected read proportion and then totaled for each treatment and replicate. The optimal exine rupture level for each barcode was determined to be the treatment that produced the lowest additive value of deviation from expected proportion.

Results

Exine rupture

Time to 95% rupture varied widely among species (mean = 304.3 seconds, Standard deviation = 199.9, Range = 594) (Fig 1). No association was recovered between pollen grain size and time to 95% rupture, however the association of time to 95% rupture and aperture number was significant (ANOVA; $p < 0.001$).

For the majority of species rupture with 1.4 mm beads resulted in the fastest time to 95% (8 of 15) while rupture with 2.8 mm beads resulted in the fastest time to 95% for five species (Table 2). Rupture with a single 5.0 mm bead resulted in the fastest time to 95% for two species (*Tilia platyphyllos* Scop. and *Zea mays* L.) however in both of these species 1.4 mm and 2.8 mm beads resulted in nearly equivalent times. Rupture with 0.5 mm beads did not result in most efficient rupture in any of the species examined and in two instances projected time to 95% rupture exceeded 10 minutes.

DNA yield

Quantity scores, as the average of the 3 replicates, increased threefold when rupture level was increased from 0% to 100% (0%(1 hour lysis) = 0.531 ng/μl, 0%(2 hour lysis) = 0.588 ng/μl,

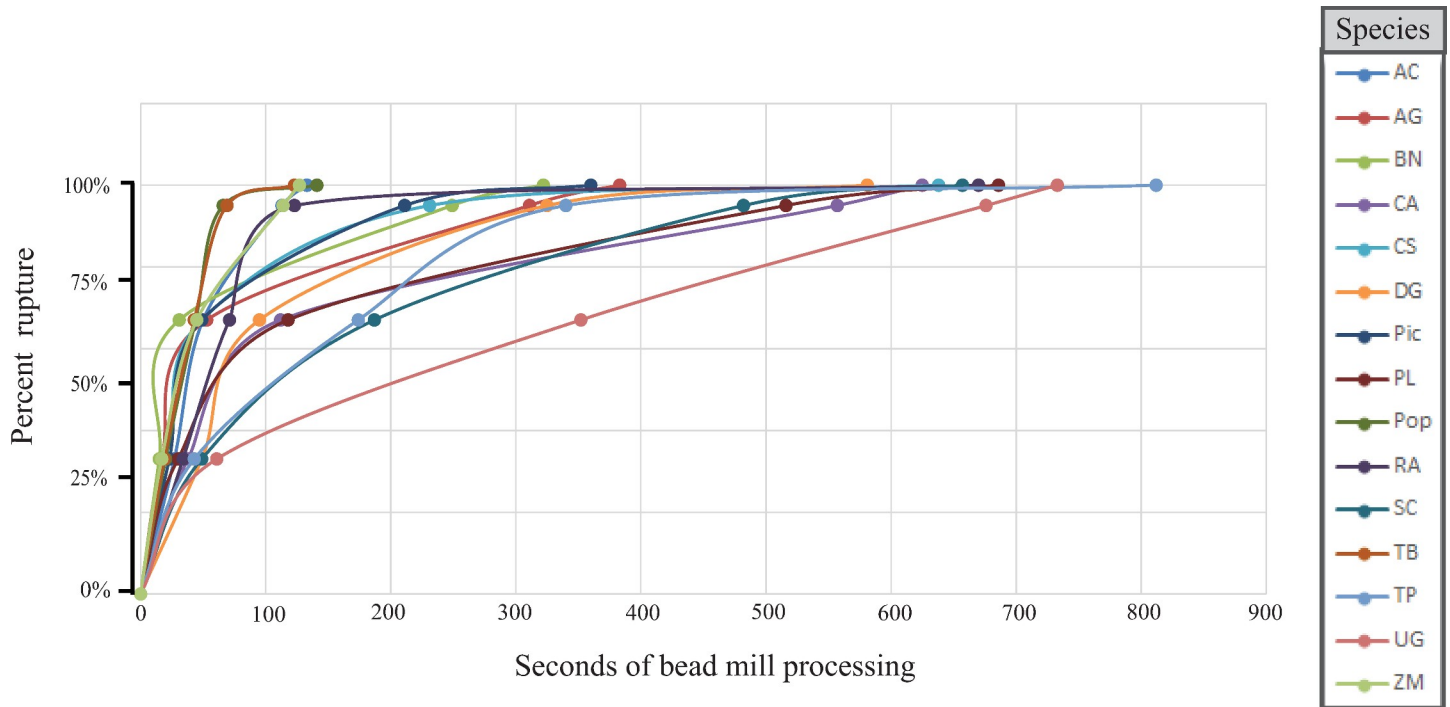


Fig 1. Curves produced from projected time to complete rupture for all species when optimal bead size is utilized. Species codes are listed in Table 2.

<https://doi.org/10.1371/journal.pone.0245611.g001>

33 = 0.808 ng/μl, 67 = 1.335 ng/μl, 100 = 1.512 ng/μl), however the 24 hour lysis with 0% exine rupture recovered very low quantities (0.222 ng/μl) and was not successful in amplification.

Metabarcoding results, barcode performance

ITS2 with either primer pairing returns all species (except *Zea mays*) above the threshold (Table 5). One primer pairing of *rbcL* (*rbcLaF*, *rbcLr506*) excludes only *Taxus baccata* L. and

Table 5. Species retrieved from Illumina MiSeq run above the threshold of a minimum of five reads in at least two replicates of at least one exine rupture treatment.

Species	Barcode and Primer Pairs				
	ITS1	ITS2 (ITS2-2SF/ITS4)	ITS2	<i>rbcL</i> (<i>rbcL2/rbcLaR</i>)	<i>rbcL</i> (<i>rbcLaF/rbcLr506</i>)
<i>Abies concolor</i>		✓	✓	✓	✓
<i>Alnus glutinosa</i>	✓	✓	✓		✓
<i>Brassica napus</i>	✓	✓	✓		✓
<i>Chenopodium album</i>	✓	✓	✓		✓
<i>Dactylis glomerata</i>	✓	✓	✓		✓
<i>Picea</i> sp.		✓	✓	✓	✓
<i>Plantago lanceolata</i>	✓	✓	✓		✓
<i>Populus</i> sp.	✓	✓	✓	✓	✓
<i>Rumex acetosella</i>	✓	✓	✓		✓
<i>Salix caprea</i>	✓	✓	✓	✓	✓
<i>Taxus baccata</i>		✓	✓	✓	
<i>Tilia platyphyllos</i>		✓	✓		✓
<i>Ulmus glabra</i>	✓	✓	✓		✓
<i>Zea mays</i>					

<https://doi.org/10.1371/journal.pone.0245611.t005>

Zea mays, while the other pairing (rbcL2, rbcLaR) returns only five of the 14 species present in this mock community. *Tilia platyphyllos* and *Zea mays* were not recovered above the threshold for ITS1 and due to the long length of ITS1 in most gymnosperms, *Abies concolor* (Gord. et Glend.) Lindl. ex Hildebr., *Picea sp.*, and *Taxus baccata* were not recovered from paired reads but could be evaluated if only single read directions are considered.

Metabarcoding results, evaluation of read proportions

All barcodes and primer pairings deviated from the expectation of equal proportion of each species at expected value of 0.0714. Due to the low number of species returned with the rbcL2/rbcLaR primer pairs, the data were eliminated from this evaluation. The treatments with 0% rupture plus increased duration of lysis incubation (2 and 24 hours) were also eliminated from further analyses due to suboptimal sequencing results or amplification failure that confounded comparable results.

No general trend that persisted across all primer pairs, treatments and species was recovered (Table 6, Fig 2), however there is little variation between treatment replicates. *Zea mays* is uniformly nearly eliminated across all primer pairs and is never present in more than two reads, *Taxus baccata* is also represented in very low read counts, often not above the five read threshold. No species is present at or near expected proportions across all treatments in all barcodes or primer pairs.

No species displays the trend of high deviation from expected values (either very low or very high) at 0% rupture and approaching the expected proportion at a higher level of rupture, across all barcodes or primer pairs. There are several instances of this within a single barcode or primer pair (Table 6).

Proportions evaluation

The mean of values for deviation from expected proportions was lowest in the 0% rupture treatment for ITS1, the 33% rupture treatment for ITS2 (universal and plant specific primers), and the 67% rupture for *rbcl*. Values for standard deviation of replicates indicate low variation in read proportions within individual treatments (Table 7, Fig 2).

Discussion

Exine rupture

Our results reflect the extreme variation and species' specific nature of the strength of the pollen grain exine. No single bead size was most efficient for exine rupture of all of the species examined and no relationship between size of pollen grain and bead size was recovered. This confirms that exine strength is comprised of more factors than surface to volume ratio alone. The two bead sizes that gave the shortest duration to 95% rupture (1.4 mm and 2.8 mm) reflect sizes used in several previous publications (Table 1). We observed a very wide range of time to achieve 95% rupture, from about one minute with *Populus sp.* and *Taxus baccata* to over 10 minutes in *Ulmus glabra* and *Alnus glutinosa*. This further illustrates the extreme variation in exine durability and architecture that occurs as a result of environmental pressures and strategies of dispersal, reproduction, and anti-desiccation [2]. We found an association between increased aperture number and increased time to 95% rupture corroborating the ideas of Payne [5] that arrangements and increased quantities of apertures contribute to the resistance of the exine to breakage.

Table 6. Trends of individual species' proportion of reads returned from Illumina sequencing of the different level of exine rupture treatments.

Trend	marker	species
Species is present above expected proportion across all treatments	<i>rbcL</i> (<i>rbcLaF</i> / <i>rbcLr506</i>)	<i>Picea sp.</i>
		<i>Populus sp.</i>
	ITS1	<i>Rumex acetosella</i>
	ITS2 (Universal)	<i>Plantago lanceolata</i>
	ITS2 (Plant specific)	-
Species is present below expected proportion across all treatments	<i>rbcL</i> (<i>rbcLaF</i> / <i>rbcLr506</i>)	<i>Taxus baccata</i>
		<i>Zea mays</i>
	ITS1	<i>Abies concolor</i>
		<i>Picea sp.</i>
		<i>Taxus baccata</i>
		<i>Tilia platyphyllos</i>
		<i>Ulmus glabra</i>
		<i>Zea mays</i>
	ITS2 (Universal)	<i>Abies concolor</i>
		<i>Taxus baccata</i>
		<i>Zea mays</i>
	ITS2 (Universal)	<i>Taxus baccata</i>
		<i>Zea mays</i>
Species is present at or near expected proportion across all treatments	<i>rbcL</i> (<i>rbcLaF</i> / <i>rbcLr506</i>)	<i>Plantago lanceolata</i>
	ITS1	-
	ITS2 (Universal)	-
	ITS2 (Plant specific)	-
Present with high deviation from expected proportion (either high or low) and approaches expected proportion at higher levels of rupture.	<i>rbcL</i> (<i>rbcLaF</i> / <i>rbcLr506</i>)	<i>Brassica napus</i>
		<i>Chenopodium album</i>
		<i>Dactylis glomerata</i>
		<i>Rumex acetosella</i>
		<i>Salix caprea</i>
		<i>Tilia platyphyllos</i>
	ITS1	<i>Alnus glutinosa</i>
		<i>Chenopodium album</i>
		<i>Dactylis glomerata</i>
	ITS2 (Universal)	<i>Alnus glutinosa</i>
		<i>Brassica napus</i>
		<i>Picea sp.</i>
		<i>Ulmus glabra</i>
ITS2 (Plant specific)	<i>Alnus glutinosa</i>	
	<i>Ulmus glabra</i>	

<https://doi.org/10.1371/journal.pone.0245611.t006>

Effect of exine rupture on DNA extraction quality

Much emphasis has been placed on the importance of the bead mill process in pollen DNA extraction with very little supporting evidence for the necessity of this process. While we did find that increased rupture did increase DNA quantity from 0.531 ng/μl to 1.512 ng/μl in our mixed samples, we expected the effect to be much more pronounced than observed. A complication in our evaluations is the low quantity of DNA that was produced from the extractions

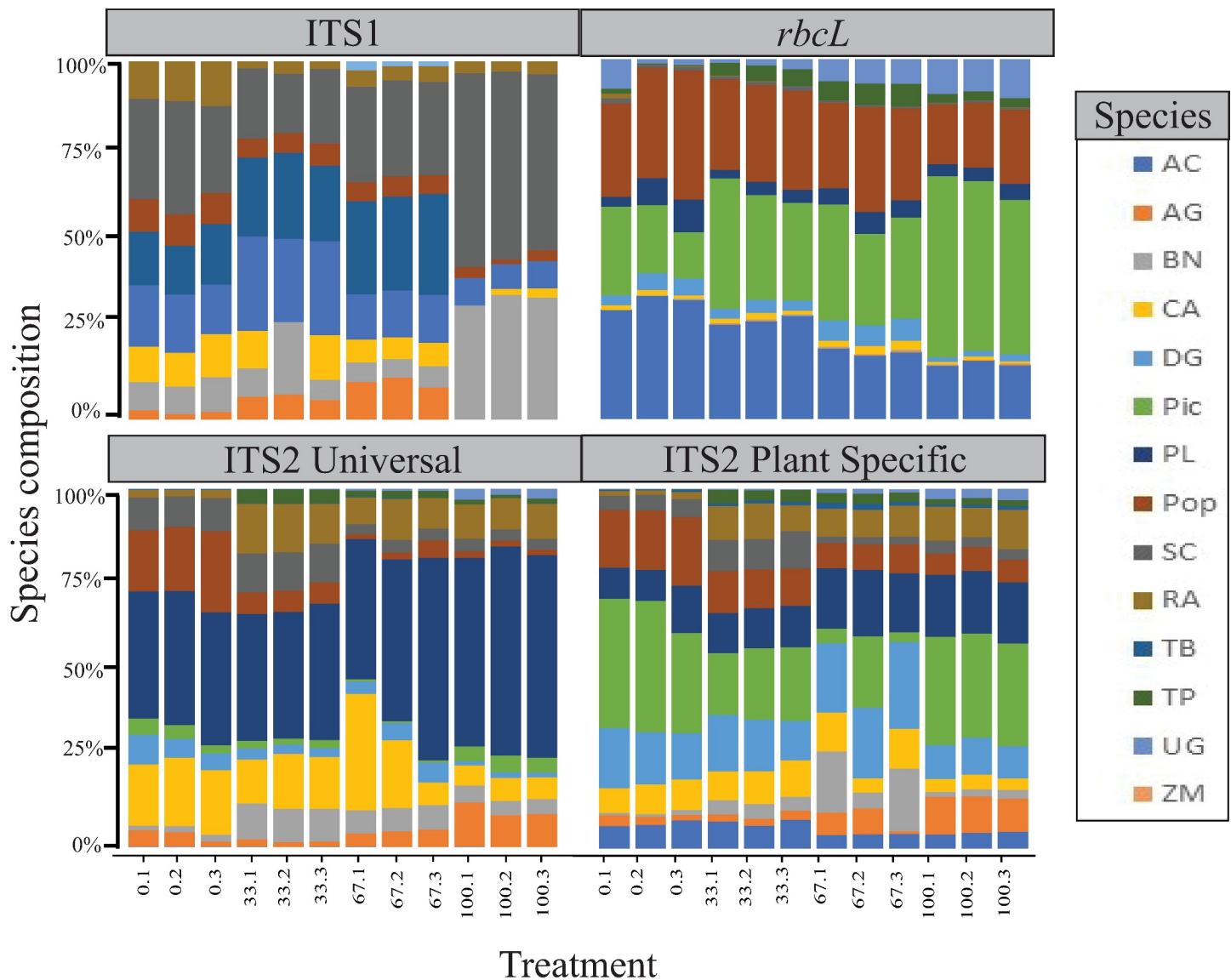


Fig 2. Species composition for all replicates of each treatment, codes for treatments are listed in Table 3 and species codes are listed in Table 2.

<https://doi.org/10.1371/journal.pone.0245611.g002>

of 10,000 pollen grains, higher quantities of starting material might have the potential to better illuminate the changes in quantity. An additional complication is that quantity was measured after species mixes were created and we are unable to evaluate the effect of increasing rupture on each species as an individual. It is possible that the positive effect of exine rupture on quantity scores will be reduced in multiporate species that are potentially more permeable to lysis buffer or thin walled species that are prone to breakage with liquid contact alone. We expect to see this positive effect more pronounced in species with typical Angiosperm morphology, where the 3 apertures lend to the structural integrity of the pollen wall but minimize the locations more vulnerable to chemical penetration. Unfortunately, due to the extreme complexity of collecting sufficient amounts of insect pollinated species, free from contamination of other species or non-pollen plant material, inclusion of representatives of these taxa was not realistic.

Regardless of bead size used or the time interval needed to reach complete rupture, the resulting curve produced for time to complete rupture (Fig 1) follows a similar pattern in all

Table 7. Cumulative deviation from expected proportions totaled for all species in each treatment and replicate.

Barcode	Treatment	Replicate 1	Replicate 2	Replicate 3	Mean	St. Dev.
<i>rbcL</i> (<i>rbcLaF/rbcLr506</i>)	0%	1.2094	1.2536	1.2436	1.2356	0.0232
	33%	1.3201	1.2346	1.2376	1.2641	0.0485
	67%	1.0824	1.0136	1.0129	1.0363	0.0399
	100%	1.2538	1.2362	1.2105	1.2335	0.0218
ITS2 (Universal)	0%	1.0274	1.0899	1.1689	1.0954	0.0709
	33%	0.9335	0.9686	0.9604	0.9542	0.0184
	67%	1.1533	1.0824	1.0132	1.0830	0.0700
	100%	1.0628	1.0914	1.0806	1.0783	0.0145
ITS2 (Plant specific)	0%	0.9779	0.9738	0.9291	0.9602	0.0271
	33%	0.6455	0.6748	0.6386	0.6530	0.0192
	67%	0.7216	0.7411	0.8588	0.7738	0.0742
	100%	0.8197	0.7846	0.7803	0.7949	0.0216
ITS1	0%	0.9477	0.9676	0.9564	0.9572	0.0100
	33%	1.0137	1.1108	1.0407	1.0547	0.0496
	67%	0.9418	0.9837	0.9612	0.9622	0.0210
	100%	1.444	1.4611	1.3872	1.4309	0.0387

<https://doi.org/10.1371/journal.pone.0245611.t007>

species examined here. The species reach 33% relatively quickly in relation to the time needed reach higher levels of rupture. This means that the additional time required to reach complete rupture of all pollen grains also greatly increases the time previously ruptured pollen grains are exposed to collisions that will degrade the DNA. This produces a trade-off where, at some point in the bead milling process, the benefit of increased genetic material released from the pollen grain wall will be outweighed by the detriment of shearing of exposed DNA. This could partially explain why we see increased DNA quantity yet diminished sequencing results in the 100% rupture treatments, as evidenced by increased divergence from expected proportions. For this reason, as well as the logistical impossibility of reaching complete exine rupture with a mixed pollen sample, we recommend conservative effort be used to break the pollen wall prior to DNA extraction.

Additional support for being less aggressive in efforts to rupture pollen grain walls in a mixed sample is the comparatively successful results of our 0% rupture treatments with 1-hour lysis incubation to treatments with higher levels of rupture. Generally speaking, within each barcode or primer pairing, the same species were retrieved regardless of level of rupture. Higher rupture levels did increase the proportion of reads in species occurring in very low read numbers, but only increased species detection in a few cases. In addition to mechanical rupture, longer lysis incubation (2 hours) without exine rupture gave similar values of DNA quantity to the treatment with only one hour of lysis incubation, but in most cases numbers of paired reads were diminished to a point that necessitated elimination from analyses. The diminished results produced from a long lysis incubation are most evident in the 24-hour treatment, where DNA extractions produced quantities that were barely detectable and efforts at amplification failed. This indicates that in addition to taking precautions not to over-mill, caution should also be applied to not over expose samples to lysis buffer prior to DNA extraction and the lysis incubation step should not exceed one hour.

Performance of barcodes/primer pairs

Our results indicate that all three common plant barcodes used here have the potential to be accurate in species retrieval in Illumina sequencing, although proportions of the species' reads

retrieved are likely to never be duplicated across different barcodes or accurately reflect the proportions of the starting materials.

The only consistent result we observed across all barcodes and primer pairs was the nearly complete exclusion of *Zea mays*. This could be due to the high GC content of this species, DNA regions with a GC content in excess of 60% require special PCR protocols to overcome strong secondary structures that interfere with primer annealing [50]. We augmented our PCR protocol with the addition of 5% DMSO, and were successful with amplification of *Z. mays* when no other species were contained in the sample, but this addition does not seem to be sufficient for returning this species from a mixed species sample. This behavior of *Z. mays* has also been shown in previous studies [21, 51, 52], where it was completely excluded or only recovered when the number of species in a mixture did not exceed two. Regardless of the reasons for exclusion of *Z. mays*, this result indicates the great importance of tests with mock communities in metabarcoding, particularly when the absence or presence of one or more particular species are critical to the experimental question.

The results for *rbcL* indicate the critical aspect of primer pair choice, especially when creation of internal primers are necessary to accommodate for the read length limitation of Illumina sequencing and when the potential species contained in the sample represent very diverse taxonomies. With one primer pairing (*rbcLaF/rbclr506*) we retrieved 12 of 14 species (exclusion of *Taxus baccata* and *Zea mays*) adding confidence to the reliability of *rbcL* in pollen metabarcoding studies. With the other primer pairing (*rbcL2/rbclrR*) however, only five of the 14 species contained within the mock community were retrieved. This result is in contrast with results of previous studies that examined mock communities [21, 51] where species retrieval was adequate, and further investigation is necessary to elucidate the reduced success we achieved here.

Our results for ITS1 also excluded several species, and only 9 of the 14 species were retrieved. This result is expected for the three Gymnosperms (*Abies concolor*, *Picea sp.*, and *Taxus baccata*) in the sample, all of which have a read length too long for creation of paired reads, but could be added to analyses if only unidirectional reads are considered. Of more concern is the absence of *Tilia platyphyllos*, which cannot be explained by read length. This result indicates that ITS1 should be examined further with mock samples before it is widely utilized in pollen or plant metabarcoding, and it is not recommended for use in airborne samples that are likely to contain Gymnosperm species.

ITS2 resulted in the most consistent results with either primer pairing and returned the most species (all but *Z. mays*) indicating that this barcode gives reliable results across a high taxonomic breadth. These results demonstrate reliable species retrieval even with universal primers that contain instances of mismatches in several taxonomic groups. The plant specific ITS2 primers used here are recommended for use when a high taxonomic diversity is expected in a sample, as they lessen the potential for mismatches and may also result in an advantage for pollen samples with a high fungal contaminant component, such as airborne pollen samples.

Evaluation of proportions

The resulting trends produced across the barcodes and/or primer pairings were inconsistent and for this reason we cannot untangle the treatment effect from several other forces that have the potential to change the proportion of reads returned.

The first possible source of deviation from equal proportions is the possibility of unequal starting material. Our counting method was based on well-established protocol [53] and several replications, however the possibility of errors in counting or small changes in quantities aliquoted while creating replications cannot be eliminated. This is likely the case with *Taxus*

baccata, which was consistently under represented across all replicates, treatments, and barcodes. With the exception of *Zea mays*, whose absence can be explained by other factors, no other species displayed this trend (either consistently too low or too high), indicating that our methods for counting and distribution of near equal amounts of pollen grains was robust to error.

The intention of these experiments was to evaluate the effect of exine rupture on proportion of reads returned. Therefore, it was necessary to begin with a quantifiable ratio of broken to intact pollen grains rather than a predetermined amount of DNA from each species, based on individual quantities of DNA extraction and C-value of the species. We are aware that plant DNA barcodes are multi-copy markers with an unknown number of replicates in the pollen grains. Therefore, it is possible that the same quantity of starting material (ng/ μ l) of species in a mock community might not lead to equal sequence recovery. Furthermore, it is widely accepted that PCR bias strongly affects metabarcoding results in Eukarya.

Trends of individual species were not consistent across all barcodes and primer pairings and we are therefore unable to draw definitive conclusions of an ideal level of exine rupture, or how a species morphology and exine strength might affect its retrieval in Illumina sequencing. The deviation from expected proportions indicate however, with these barcodes and primer pairings, the 33% and 67% treatments produce the lowest deviation from expected proportions, with the exception of ITS1 which had the lowest deviation at 0%. These results strengthen the idea that exine rupture prior to DNA extraction does enhance results and aids in counteracting other sources of biases in downstream processes, however, this is not a linear relationship and extreme effort to lyse every pollen exine in a sample will also have detrimental effects on results.

Perhaps the greatest implication of our study is to further strengthen the idea that great care must be taken in attempting to extract conclusions about the quantity of starting material from the proportions of reads retrieved. The barcodes used here were adequate in retrieving all or a majority of species; however, results were completely inconsistent across treatments and barcodes in terms of proportions of these species retrieved. Our results indicate that Illumina sequencing can provide accurate information of presence or absence of a species, but should never be used to interpret quantities of those species within a sample.

Conclusions

- Exine rupture does appear to enhance results, but over-processing can be detrimental for results. We recommend bead milling for 2–4 minutes in a 30 hz bead mill with 1–3 mm beads or a mix that contains these sizes, and a lysis incubation should not exceed 1 hour.
- When creating artificial pollen samples, we recommend a minimum of 100,000 pollen grains be used for starting material in order to obtain the quantities necessary for experimentation.
- The lack of trends for any species across barcodes illustrates the confounding problems associated with pollen metabarcoding using plant multi copy markers, as well as the role every component of the collection and laboratory processes can have on the proportion of reads returned from sequencing.
- The lack of clear trends recovered here further display the inability for pollen metabarcoding with Illumina sequencing to be indicative of the proportions of species present in the initial samples, however, species identification, except of the problematic *Zea mays*, is very successful.

Acknowledgments

We would like to thank the the German Aerospace Center (DLR) for project support, especially Gerhard Michling, Uta von Witsch, Stefanie Vey, and Silvia Schneider; Wolfgang Wägele und Mathias Geiger for The German Barcode of Life (GBOL) Phase II project coordination; Katharina Diehl, Gunther Gerlach and Annalena Mehl for administrative support; Silke Winkler and Bertold Fartman for support with Illumina Sequencing at LGC Genomics GmbH Berlin; and all members of the Systematic Botany working group at Justus Liebig University for guidance and technical assistance, especially Andreas Kolter for the development of the ITS1 and ITS2 primers used in this study.

Author Contributions

Conceptualization: Stephanie J. Swenson, Birgit Gemeinholzer.

Data curation: Stephanie J. Swenson.

Formal analysis: Stephanie J. Swenson.

Funding acquisition: Birgit Gemeinholzer.

Investigation: Stephanie J. Swenson.

Methodology: Stephanie J. Swenson, Birgit Gemeinholzer.

Writing – original draft: Stephanie J. Swenson.

Writing – review & editing: Birgit Gemeinholzer.

References

1. Deiner K, Walser JC, Mächler E, Altermatt F. Choice of capture and extraction methods affect detection of freshwater biodiversity from environmental DNA. *Biol. Conserv.* 2015; 183: 53–63.
2. Wang R, Dobritsa AA. Exine and aperture patterns on the pollen surface: their formation and roles in plant reproduction. *Annu. Plant Reviews.* 2018; 1: 589–628.
3. Simel EJ, Saidak LR, Tuskan GA. Method of extracting genomic DNA from non-germinated Gymnosperm and Angiosperm pollen. *BioTechniques.* 1997; 22: 390–394. <https://doi.org/10.2144/97223bm02> PMID: 9067005
4. Gibbons LE, Brangs HCG, Burden DW. Bead beating: a primer. *Ran Primers.* 2014; 12: 1–20.
5. Payne WW. Structure and function in angiosperm pollen wall evolution. *Rev. Palaeobot. Palynol.* 1981; 35: 39–59.
6. Pornon A, Escaravage N, Burrus M, Holota H, Khimoun A, Mariette J, et al. Using metabarcoding to reveal and quantify plant–pollinator interactions. *Sci. Rep.* 2016; 6: 27282. <https://doi.org/10.1038/srep27282> PMID: 27255732
7. Prosser SWJ, Hebert PDN. Rapid identification of the botanical and entomological sources of honey using DNA metabarcoding. *Food Chem.* 2017; 214: 183–191. <https://doi.org/10.1016/j.foodchem.2016.07.077> PMID: 27507464
8. Leontidou K, Vernesi C, De Groeve J, Cristofolini F, Vokou D, Cristofori A. DNA metabarcoding of airborne pollen: new protocols for improved taxonomic identification of environmental samples. 2018; *Aerobiologia* 34: 63–74.
9. Longhi S, Cristofori A, Gatto P, Cristofolini F, Grando MS, Gottardini E. Biomolecular identification of allergenic pollen: a new perspective for aerobiological monitoring? *Ann. Allergy Asthma Immunol.* 2009; 103: 508–514. [https://doi.org/10.1016/S1081-1206\(10\)60268-2](https://doi.org/10.1016/S1081-1206(10)60268-2) PMID: 20084845
10. Valentini A, Miquel C, Taberlet P. DNA barcoding for honey biodiversity. *Diversity.* 2010; 2: (4)610–617.
11. Bruni I, Gallimberti A, Caridi L, Scaccabarozzi D, De Mattia F, Casiraghi M, et al. DNA barcoding approach to identify plant species in multiflower honey. *Food Chem.* 2015; 170: 308–315. <https://doi.org/10.1016/j.foodchem.2014.08.060> PMID: 25306350
12. Schnell IB, Fraser M, Willerslev E, Gilbert MTP. Characterisation of insect and plant origins using DNA extracted from small volumes of bee honey. *Arthropod Plant Interact.* 2010; 4: 107–116.

13. Danner N, Keller A, Härtel S, Steffan-Dewenter I. Honey bee foraging ecology: Season but not landscape diversity shapes the amount and diversity of collected pollen. *PLoS ONE*. 2017; 12(8): e0183716. <https://doi.org/10.1371/journal.pone.0183716> PMID: 28854210
14. Korpelainen H, Pietiläinen M. Biodiversity of pollen in indoor air samples as revealed by DNA metabarcoding. *Nord. J. Bot.* 2017; 35: 602–608.
15. Núñez A, de Paz GA, Ferencova Z, Rastrojo A, Guantes R, Garcia AM, et al. Validation of the Hirst-type spore trap for simultaneous monitoring of prokaryotic and eukaryotic biodiversity in urban air samples by NGS. *Appl. Environ. Microbiol.* 2017; 83: e00472–17. <https://doi.org/10.1128/AEM.00472-17> PMID: 28455334
16. Galimberti A, De Mattia F, Bruni I, Scaccabarozzi D, Sandionigi A, Barbuto M, et al. A DNA barcoding approach to characterize pollen collected by honeybees. *PLoS ONE*. 2014; 9(10): e109363. <https://doi.org/10.1371/journal.pone.0109363> PMID: 25296114
17. Galliot J-N, Brunel D, Bérard A, Chauveau A, Blanchetête A, Lanore L, et al. Investigating a flower-insect forager network in a mountain grassland community using pollen DNA barcoding. *J Insect Conserv.* 2017; 21: 827–837.
18. Guertler P, Eicheldinger A, Muschler P, Goerlich O, Busch U. Automated DNA extraction from pollen in honey. *Food Chem.* 2014; 149: 302–306. <https://doi.org/10.1016/j.foodchem.2013.10.129> PMID: 24295710
19. Keller A, Danner N, Grimmer G, Ankenbrand M, von der Ohe K, von der Ohe W, et al. Evaluating multiplexed next-generation sequencing as a method in palynology for mixed pollen samples. *Plant Biol.* 2015; 17: 558–566. <https://doi.org/10.1111/plb.12251> PMID: 25270225
20. Sickel W, Ankenbrand MJ, Grimmer G, Holzschuh A, Härtel S, Lanzen J, et al. Increased efficiency in identifying mixed pollen samples by meta-barcoding with a dual-indexing approach. *BMC Ecology.* 2015; 15: 20. <https://doi.org/10.1186/s12898-015-0051-y> PMID: 26194794
21. Bell KL, Loeffler VM, Brosi BJ. An *rbcL* reference library to aid in the identification of plant species mixtures by DNA metabarcoding. *Appl. Plant Sci.* 2017; 5(3): 1–7. <https://doi.org/10.3732/apps.1600110> PMID: 28337390
22. Gous A, Swanevelder DZH, Eardley CD, Willows-Munro S. Plant-pollinator interactions over time: Pollen metabarcoding from bees in a historic collection. *Evol. Appl.* 2019; 12: 187–197. <https://doi.org/10.1111/eva.12707> PMID: 30697333
23. Corman RS, Otto CRV, Iwanowicz D, Pettis JS. Taxonomic characterization of honey bee (*Apis mellifera*) pollen foraging based on non-overlapping paired-end sequencing of nuclear ribosomal loci. *PLoS ONE*. 2015; 10(12): e0145365. <https://doi.org/10.1371/journal.pone.0145365> PMID: 26700168
24. Kraaijeveld K, de Weger LA, Ventayol García M, Buermans H, Frank J, Hiemstra PS, et al. Efficient and sensitive identification and quantification of airborne pollen using next-generation DNA sequencing. *Mol. Ecol. Resources.* 2015; 15: 8–16. <https://doi.org/10.1111/1755-0998.12288> PMID: 24893805
25. De Vere N, Jones LE, Gilmore T, Moscrop J, Lowe A, Smith D, et al. Using DNA metabarcoding to investigate honey bee foraging reveals limited flower use despite high floral availability. *Sci. Rep.* 2017; 7: 42838. <https://doi.org/10.1038/srep42838> PMID: 28205632
26. Smart MD, Corman RS, Iwanowicz DD, McDermott-Kubeczko M, Pettis JS, Spivak MS, et al. A comparison of honey bee-collected pollen from working agricultural lands using light microscopy and ITS metabarcoding. *Environ. Entomol.* 2017; 1–12. <https://doi.org/10.1093/ee/nvw154> PMID: 28025220
27. Hawkins J, de Vere N, Griffith A, Ford CR, Allainguillaume J, Hegarty MJ, et al. Using DNA metabarcoding to identify the floral composition of honey: A new tool for investigating honey bee foraging preferences. *PLoS ONE*. 2015; 10: e0134735. <https://doi.org/10.1371/journal.pone.0134735> PMID: 26308362
28. Potter C, de Vere N, Jones LE, Ford CR, Hegarty MJ, Hodder KH, et al. Pollen metabarcoding reveals broad species-specific resource use by urban bees. *PeerJ.* 2019; 7: e5999. <https://doi.org/10.7717/peerj.5999> PMID: 30809427
29. Folloni S, Kagkli D-M, Rajcevic B, Guimaraes NCC, Van Droogenbroeck B, Valicente FH, et al. Detection of airborne genetically modified maize pollen by real-time PCR. *Mol. Ecol. Resour.* 2012; 12: 810–821. <https://doi.org/10.1111/j.1755-0998.2012.03168.x> PMID: 22805239
30. Richardson RT, Lin C-H, Quijia JO, Riusech NS, Goodell K, Johnson RM. Rank-based characterization of pollen assemblages collected by honey bees using a multi-locus metabarcoding approach. *Appl. Plant Sci.* 2015a; 3: 1500043.
31. Richardson RT, Lin C-H, Sponsler DB, Quijia JO, Goodell K, Johnson RM. Application of ITS2 metabarcoding to determine the provenance of pollen collected by honey bees in an agroecosystem. *Appl. Plant Sci.* 2015b; 3: (1)1400066.

32. Peel N, Dicks LV, Clark MD, Heavens D, Percival-Alwyn L, Cooper C, et al. Semi-quantitative characterization of mixed pollen samples using MinION sequencing and Revers Metagenomics (RevMet). *Methods Ecol. Evol.* 2019; 10: 1690–1701.
33. Lalhmanghai R, Ghatak S, Laha R, Gurusubramanian G, Kumar NS. Protocol for optimal quality and quantity pollen DNA isolation from honey samples. *J Biomol Tech.* 2014; 25: 92–95. <https://doi.org/10.7171/jbt.14-2504-001> PMID: 25365793
34. Laha RC, De Mandal S, Lalhmanghai R, Raite L, Kumar NS, Gurusubramanian G, et al. Meta-barcoding in combination with palynological inference is a potent diagnostic marker for honey floral composition. *AMB Expr.* 2017; 7: 132.
35. McFrederick QS, Rehan SM. Characterization of pollen and bacterial community composition in brood provisions of a small carpenter bee. *Mol. Ecol.* 2016; 25: 2302–2311. <https://doi.org/10.1111/mec.13608> PMID: 26945527
36. Lang D, Tang M, Hu J, Zhou X. Genome-skimming provides accurate quantification for pollen mixtures. *Mol. Ecol. Resources.* 2019; 19: 1433–1446. <https://doi.org/10.1111/1755-0998.13061> PMID: 31325909
37. Wodehouse RP. Pollen grains: Their structure, identification and significance in science and medicine. McGraw-Hill, New York. 1935;574 pp.
38. Palmieri L, Bozza E, Giongo L. Soft fruit traceability in food matrices using real-time PCR. *Nutrients.* 2009; 1: 316–328. <https://doi.org/10.3390/nu1020316> PMID: 22253987
39. Kress WJ, Erickson DL. A two-locus global DNA barcode for land plants: The coding *rbcl* gene complements the non-coding *trnH-psbA* spacer region. *PLoS ONE.* 2007; 2: e508. <https://doi.org/10.1371/journal.pone.0000508> PMID: 17551588
40. Levin RA, Wagner WL, Hoch PC, Nepokroeff M, Pires JC, Zimmer EA, et al. Family-level relationships of Onagraceae based on chloroplast *rbcl* and *ndhF* data. *Am J Bot.* 2003; 90: 107–115. <https://doi.org/10.3732/ajb.90.1.107> PMID: 21659085
41. de Vere N, Rich TCG, Ford CR, Trinder SA, Long C, Moore CW, et al. DNA barcoding the native flowering plants and conifers of Wales. *PLoS ONE.* 2012; 7(6): e37945. <https://doi.org/10.1371/journal.pone.0037945> PMID: 22701588
42. Chen S, Yao H, Han J, Liu C, Song J, Shi L, et al. Validation of the ITS2 region as a novel DNA barcode for identifying medicinal plant species. *PLoS ONE.* 2010; 5(1): e8613. <https://doi.org/10.1371/journal.pone.0008613> PMID: 20062805
43. White TJ, Bruns T, Lees S, Taylor J. Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. In *PCR protocols a guide to methods and applications.* Academic Press, San Diego. 1990. pp. 315–322.
44. Ivanova NV, Grainger CM. CCDB Protocols, COI amplification. 2007. Available from http://ccdb.ca/site/wp-content/uploads/2016/09/CCDB_Amplification.pdf
45. Caporaso JG, Kuczynski J, Stombaugh J, Bittinger K, Bushman FD, Costello EK, et al. QIIME allows analysis of high throughput community sequencing data. *Nat. Methods.* 2010; 7: 335–336. <https://doi.org/10.1038/nmeth.f.303> PMID: 20383131
46. Edgar RC. Search and clustering orders of magnitude faster than BLAST. *Bioinformatics (Oxford, England).* 2010; 26: 2460–2461. <https://doi.org/10.1093/bioinformatics/btq461> PMID: 20709691
47. Edgar RC. SINTAX: a simple non-Bayesian taxonomy classifier for 16S and ITS sequences. *bioRxiv* 074161.
48. Banchi E, Ametrano CG, Greco S, Stanković D, Muggia L, Pallavicini A. PLANITS: a curated sequence reference dataset for plant ITS DNA metabarcoding. *Database* 2020, 2020. <https://doi.org/10.1093/database/baz155> PMID: 32016319
49. R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. 2016; URL <http://www.R-project.org/> [accessed 11 May 2017]
50. Flores-Juárez CR, González-Jasso E, Antaramain A, Pless RC. PCR amplification of GC-rich DNA regions using the nucleotide analog N4-methyl-2'-deoxycytidine 5'-triphosphate. *Biotechniques.* 2016; 61(4): 175–182. <https://doi.org/10.2144/000114457> PMID: 27712580
51. Bell KL, Burgess KS, Botsch JC, Dobbs EK, Read TD, Brosi BJ. Quantitative and qualitative assessment of pollen DNA metabarcoding using constructed species mixtures. *Mol. Ecol.* 2019; 28: 431–455. <https://doi.org/10.1111/mec.14840> PMID: 30118180
52. Nürnberger F, Keller A, Härtel S, Steffan-Dewenter I. Honey bee waggle dance communication increases diversity of pollen diets in intensively managed agricultural landscapes. *Mol Ecol.* 2019; 28: (15)3602–3611. <https://doi.org/10.1111/mec.15156> PMID: 31233637
53. Kääpylä M, Penttinen A. An evaluation of the microscopical counting methods of the tape in Hirst-Burkard pollen and spore trap. *Grana.* 1981; 20: (2)131–141.